A Natural Experiment on Discrimination in Elections*

David E. Broockman
UC Berkeley

Evan J. Soltas
MIT

January 15, 2021

Abstract

We exploit a natural experiment to study discrimination in elections. In Illinois Republican presidential primaries, voters vote for delegates bound to presidential candidates, but delegates' names convey information about their race and gender. We identify discrimination from variation in vote totals among delegates bound to the same presidential candidate and who face the same voters. Examining delegate vote totals from 2000 to 2016, we estimate nonwhite delegates receive 9 percent fewer votes. We find essentially no gender discrimination. Negligible incentives for statistical discrimination, costs to preferred presidential candidates, and heterogeneity are consistent with an interpretation of this behavior as taste-based.

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

*Broockman: UC Berkeley (dbroockman@berkeley.edu). Soltas: MIT Department of Economics (esoltas@mit.edu). We are grateful to Dave Wasserman for the news article that served as inspiration for this study. We thank the editor Thomas Fujiwara and the anonymous referees for their comments and suggestions. For helpful feedback, we thank Avi Acharya, Hunt Allcott, David Autor, Tim Besley, Kate Casey, Stefano DellaVigna, Dana Foarta, Matthew Gentzkow, Don Green, Dan Ho, Saumitra Jha, Ilyana Kuziemko, Horacio Larreguy, Michael Luca, Jon Levin, Pablo Montagnes, Sergio Montero, Nicola Persico, Ken Shotts, Erik Snowberg, Jorg Spenkuch, Seth Stephens-Davidowitz, Francesco Trebbi, Ebonya Washington, Ismail White, and seminar participants at LSE, the NBER, Oxford, POLMETH, Queen's University, SPSA, Stanford University, UC Berkeley, UC Davis, UC Riverside, the University of Toronto, the University of Zurich, the University of Southern California, the Wallis Institute, and Yale. For their able research assistance, we thank Hadil Almowafak, Julia Espero, Corey Garcia, Chris McConnell, Leela Srinivasan, KS Tey, and Yao Yang. Soltas gratefully acknowledges support from the National Science Foundation Graduate Research Fellowship under Grant No. 1122374.

1 Introduction

Racial and ethnic minorities and women are underrepresented among elected officials in many countries.[1] A large body of research indicates that such underrepresentation contributes to disparities by race, ethnicity, and gender in a variety of economic and political outcomes. Electing minority or female officials 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 (Beaman et al., 2012; Clots-Figueras, 2012; Iyer et al., 2012; Fujiwara, 2015). One potential explanation for the underrepresentation of minorities and women in elected office is voter discrimination, wherein voters are less likely to vote for a minority or female candidate than an otherwise-identical non-minority or male candidate.

Understanding whether voters discriminate by race and gender and the mechanisms for any such discrimination are central questions in political economy and would inform significant policy debates. Research finds that individuals engage in racial discrimination in product and labor markets (for review, see Bertrand and Duflo, 2017), but credible evidence on whether they do so when voting is limited. On average, minorities run in different electoral districts, at different times, on different platforms, with different party affiliations, for different offices, with different pre-election experience and campaign resources, and so on, making credible identification of voter discrimination difficult in most electoral settings (Stephens-Davidowitz, 2014). In settings beyond politics, researchers have identified discrimination with a variety of credible research designs.[2] However, it is challenging to adapt many of these strategies to elections.

In this paper we analyze a natural experiment to study voter discrimination against nonwhite and female political candidates. This natural experiment occurred in four recent Illinois Republican presidential primary elections: 2000, 2008, 2012, and 2016.[3] Unique features of the institutional environment means discrimination in these elections can plausibly be causally identified. For voters to fully express their preferences in Illinois Republican presidential primary elections, 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 $N \in \{2, 3, 4\}$ of delegate candidates

[1] For example, in the United States as of 2015, 38 (51) percent of the population was nonwhite (female), compared to 17 (20) percent of the U.S. Congress (Manning, 2016). For brevity, in 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] 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), and lab and field experiments (for review, see Guryan and Charles, 2013; Bertrand and Duflo, 2017).

[3] 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.

who appear on the ballot for each presidential candidate, and voters cast votes for up to N delegates. The top N vote-getting delegates in each district win and are bound to vote for their presidential candidate at the convention. However, the delegates' names also appear on ballots, and the delegates' names convey information about their race and gender.

To identify racial and gender discrimination in this setting, we exploit the fact that we observe the vote totals of multiple delegates 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 voters should select to fully support their preferred presidential candidate. Our identification strategy is to examine variation in delegate vote totals by delegate race and gender within such groups of delegates. For example, suppose the delegates for Mitt Romney in Illinois' first congressional district in 2012 were named Tom, Dick, and José. To maximize the value of their ballot, a Romney supporter should cast their three votes for Tom, Dick, and José. However, to the extent Romney supporters engage in racial or ethnic discrimination, some may vote for Tom and Dick but not for José, leaving José's vote totals lower than Tom and Dick's vote totals. We observe 816 unique natural experiments of this form.

This election design has important advantages for studying discrimination. In typical elections, voters may value myriad dimensions of candidates—such as ideology, competence, or past performance—many of which may correlate with candidate demographics and few of which are perfectly observable. Here the voter's problem is dramatically more straightforward: For voters seeking to fully support their preferred presidential candidate, a delegate candidate's sole relevant dimension is the presidential candidate to whom they are bound,[4] which is clearly printed on ballots, such that both voters and researchers can perfectly observe it. White and male delegate candidates running alongside nonwhite and female delegate candidates on the same platform, for the same office, on the same ballots, in front of the same voters therefore provide a naturally-occurring control group that allow us to rule out factors that would confound other research designs.

Analyzing variation in delegate vote totals among 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 9 percent fewer votes when their names indicate they are not white. We also find, however, essentially no discrimination against women: Delegates whose names indicate they are female receive on average about the same number of votes as delegates whose names indicate they are male, if not slightly more in some specifications.

[4] In Section 4.3, we consider other dimensions voters may value besides the presidential candidate to whom delegates are bound, such as if a delegate is an existing elected official, is a "local notable," or is listed higher on the ballot. We find these dimensions are uncorrelated with delegate race. Serving as a delegate most resembles hobbyist consumption, rather than a career investment. See Appendix H for discussion, including quotes from convention delegates about why they volunteered to attend.

Examining the results by ethnicity, we find clear evidence of discrimination against delegates whose names indicate they are Hispanic, East Asian, Middle Eastern, or Indian. Our estimates for discrimination against black delegates are similar, although there are few clearly black delegates, making our estimates of anti-black discrimination less precise.

We next consider what theoretical mechanisms could account for this discrimination. Discrimination may be taste-based—wherein voters act as if they have preferences over candidate race and gender (Becker, 1957). It may also be statistical—wherein voters accurately use candidates' race and gender to infer non-racial or non-gender dimensions such as ideology and can advance these non-demographic preferences by discriminating (Phelps, 1972; Akerlof, 1976). Taste-based discrimination is of particular interest because it implies that voters act as if they pay a "psychic cost" (Becker, 1957) of voting for candidates from disfavored demographic groups and accept trade-offs on candidates' non-demographic dimensions to avoid paying these "psychic costs." As compared to markets (List, 2004), taste-based discrimination may be especially likely in elections because individual voting decisions are usually inconsequential for outcomes. Nevertheless, taste-based discrimination may be consequential in aggregate if many voters engage in it, reducing both minority representation and the appeal of election winners on other dimensions.

Institutional features of the setting we study suggest tastes—and, in particular, the "psychic costs" of voting for delegates of a disfavored race or gender—as the likely mechanism for the discrimination we detect. In most electoral settings it would be difficult to distinguish between taste-based and statistical discrimination.[5] Taste-based discrimination occurs when individuals behave as if they prefer candidates inferior on non-racial or non-gender dimensions in order to avoid incurring "psychic costs" from selecting otherwise-preferred candidates with disfavored demographics. Voters must behave in precisely this manner to discriminate against delegates in this environment. If voters do not vote for all their preferred presidential candidates' nonwhite delegates, this advantages delegates for presidential candidates they prefer less, undermining the nomination prospects of voters' preferred presidential candidates. In addition, although we cannot rule out all alternative interpretations, incentives for statistical discrimination should be naturally absent. Under convention rules, delegates have essentially no discretion. Even if a rational voter were unaware that delegates had no discretion, to engage in statistical discrimination, she would need to maintain beliefs we view as implausible: that nonwhite delegates bound to her presidential candidate of choice would be less likely to vote for the voter's presidential candidate of choice at the

[5] Existing approaches in the gender literature include comparisons of vote totals controlling for observables, survey-based experiments (e.g., Teele et al., 2018), and testing implications of voter bias on politician quality (Anzia and Berry, 2011; Ferreira and Gyourko, 2014; Vogl, 2014).

convention than white delegates bound to an opposing presidential candidate.[6]

Heterogeneity in the magnitude of discrimination across candidates, elections, and geography is also consistent with predictions of a taste-based interpretation. Most significantly, consistent with taste-based discrimination also affecting voters' choices of presidential candidates, and not only delegate candidates, we estimate that voters for nonwhite presidential candidates harbor significantly weaker racially-discriminatory tastes than voters for white presidential candidates. We also find that voters for a female presidential candidate actually discriminate in favor of female delegate candidates. These results are consistent with voters' racial and gender tastes having stakes for their choices of presidential candidates, as voters appear to endogenously select out of voting for nonwhite and female presidential candidates in a manner strongly correlated with our estimates of their collective tastes. In addition, consistent with Becker (1957), discrimination also appears to decrease when it is more costly to voters' preferred presidential candidates: Voter discrimination against nonwhites is less when voters are more likely to be decisive, although we find it 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.

We present a variety of robustness checks on our results. We show the results are similar when we use each of three different strategies to measure the racial 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; our results are robust to excluding delegates about whom voters might have had other information and to controlling for this information. We also consider alternative interpretations of the results, such as residual incentives for statistical discrimination, voter misunderstandings about the primary, voter signaling to presidential candidates or party elites, voter indifference across presidential candidates, and voter inferences about presidential candidates on the basis of nonwhite delegates. Although we cannot rule out alternative mechanisms beyond taste-based discrimination related to voter misunderstanding and signaling, Section 4.4 and

[6] It is therefore unlikely that voters with mistaken beliefs about the primary and delegates would perceive incentives for statistical discrimination. We consider the plausibility of statistical discrimination in more detail in Sections 2 and 4.4, as well as in Appendix J. 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. Moreover, Becker (1957) defined taste-based discrimination to encompass inaccurate beliefs about minorities: "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).

Appendix J provide further evidence on and discussion of these alternatives.[7]

We find voter racial discrimination likely had aggregate consequences for delegate election outcomes. It is likely that discrimination against several presidential candidates' nonwhite delegates reduced their vote totals sufficiently that white delegates for less-preferred presidential candidates won and served instead.[8] This illustrates the two consequences of taste-based racial discrimination in elections appear to have been realized in this environment: In reducing nonwhite representation, voters who discriminated also elected candidates less appealing to them on other dimensions. To contextualize the magnitude of our point estimate, we also offer back-of-the-envelope calculations that apply our estimate to Republican primary elections for the U.S. House of Representatives. Discrimination of this magnitude would decrease the share of nonwhites in the U.S. House by about 3 percentage points. Voter discrimination against racial-minority candidates therefore plausibly contributes to their underrepresentation in government, which other research has found lies at the root of important social disparities.

Like much previous research on discrimination (e.g., Goldin and Rouse, 2000; Price and Wolfers, 2010; Doleac and Stein, 2013; Rubinstein and Brenner, 2014; Glover et al., 2017), we exploit an institution whose unique properties facilitate otherwise-elusive causal inference. Therefore, we carefully consider the generalizability of our findings as to how detecting discrimination in this setting should inform views of the plausible extent of discrimination in other elections. The underrepresentation of minorities and women in the U.S. is most extensive among Republican elected officials, making Republican primaries—the de-facto elections in about half of U.S. electoral districts—of particular substantive importance. Our finding that voters appear to endogenously select into voting for presidential candidates in a manner strongly correlated with our estimates of their collective tastes is consistent with voters' racial and gender tastes having stakes for their choices of presidential candidates. However, discrimination could be greater in other primaries where voters may have less information or weaker preferences. The presidential primaries we study are relatively high-stakes elections, determining the Republican presidential nominee, and where evidence indicates voters have strong candidate preferences. In addition, our estimates can capture only the "psychic cost" of voting for nonwhites and women, not any "psychic costs" of being represented by them nor statistical discrimination against them. Accounting for such distinctions, the total disadvantage for nonwhite candidates in other elections due to voter discrimination may well

[7] There we discuss in more detail why voter misunderstanding should not produce incentives for statistical discrimination. We also fielded a survey of Illinois Republican primary voters that found limited voter misunderstanding or perceived incentives for statistical discrimination. However, 10 percent said they felt "uncomfortable" voting for nonwhites, consistent with taste-based discrimination.

[8] In Section 3.6 we discuss why presidential campaigns may nominate nonwhite delegates despite this cost.

be larger. On the other hand, discrimination could be smaller in general elections where partisan preferences may be more important, or in elections where voters have on average weaker racial tastes.

The difference in our results for race versus for gender merits an additional comment. In particular, our finding that voters do not appear to discriminate against women is consonant with evidence from gender quotas (Baltrunaite et al., 2014, 2019; Esteve-Volart and Bagues, 2012; Casas-Arce and Saiz, 2015; Besley et al., 2017) and survey experiments (Schwarz and Coppock, 2019) that suggest political institutions, rather than voters, as the principal obstacle to increasing female political representation in Western democracies. On the other hand, the discrimination we detect against nonwhites fits with Washington (2006), who shows that a fraction of white voters turns out specifically to vote against nonwhite U.S. House candidates. Our results therefore lend support to claims in the literature on gender that underrepresentation likely results from different combinations of contributing factors for nonwhites and women.

We explain the context and natural experiment in greater detail in Section 2. Section 3 introduces our main data sources. Section 4 presents our empirical strategy, results, and robustness checks. Section 5 examines the pattern of heterogeneity in discrimination and argues it is consistent with a taste-based interpretation. Section 6 discusses implications for other elections, U.S. election law, and strategies for increasing nonwhite and female representation.

2 The Illinois Republican Presidential Primary

2.1 Why Study Republican Primaries?

Why are nonwhites and women underrepresented among U.S. officeholders? Stylized facts about U.S. politics suggest discrimination among Republican voters in primary elections may play a significant role. First, U.S. voters have strong partisan preferences, with relatively weaker preferences among candidates of the same party (Green et al., 2002), implying that any taste for a candidate's race, gender, or ethnicity may be especially determinative in primary rather than general elections. Second, white Republican voters have been more racially conservative than white Democratic voters since the 1960s civil rights realignment (Kuziemko and Washington, 2018). Nevertheless, Republican primaries constitute the de-facto election in about half of U.S. electoral districts, where Republican nominees reliably win in general elections. To the extent voter taste-based discrimination affects the demographic composition of U.S. officeholders, one thus might expect it to do so especially through Republican

Table 1: Racial and Gender Composition of Officeholders and Voters by Party

	% Nonwhite		% Female	
	Republicans	Democrats	Republicans	Democrats
U.S. House Members	4%	31%	9%	31%
Primary Election Voters	11%	35%	51%	60%
Party Identifiers	15%	43%	53%	59%
General Election Voters	23%		54%	

Notes: This table reports the national shares nonwhite and female of U.S. House members, primary voters, party identifiers, and general election voters who are Republicans or Democrats. Data on the racial and ethnic composition of U.S. House Members from 2006-2014 was collected by Fraga (2013). Data on the gender composition of U.S. House Members was collected by the Center for American Women and Politics (2016). Data on the racial, ethnic, and gender composition of primary-election voters, general-election voters, and party identifiers is from the 2016 Cooperative Congressional Election Study (CCES), a large survey of American voters (N = 64,600) (Ansolabehere and Schaffner, 2017). Whether CCES participants voted is measured from administrative records. Party identification is measured by the survey question "Generally speaking, do you think of yourself as a...?" with answers that include "Democrat" and "Republican."

primary elections.

Consistent with this possibility, Table 1 shows that both nonwhites and women are underrepresented among Republican U.S. House members relative to the populations of Republican primary voters and adults who identify as Republicans. These disparities are present but notably smaller among Democrats. Discrimination in Republican primaries could contribute to these descriptive patterns because, in a considerable share of U.S. electoral districts, whoever the Republican Party nominates is likely to win the general election. Any racial, ethnic, or gender discrimination in Republican primaries in such Republican-leaning districts may therefore alter the demographic composition of U.S. elected officials. While of course other plausible explanations exist for these descriptive patterns, the natural experiment we analyze provides new evidence consistent with voter racial discrimination in Republican primaries as a contributor.

2.2 Design of the Primary

We study Illinois Republican primary elections, taking advantage of their unique design. The "delegate loophole primary" design of the Illinois Republican presidential primary is unique within the United States.[9] Voters vote separately for some number of delegates—

[9]Delegate loophole primaries were once common in the U.S. but were largely replaced by candidate-based primaries in the 1970s as part of reforms intended to empower voters in primary elections (Shafer, 1988). Difficulty in locating election returns 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

Figure 1: Section of Sample Ballot

FOR DELEGATE TO THE
NATIONAL NOMINATING CONVENTION
EIGHTEENTH 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)

- **ROBERT BROWNING** (CHRISTIE)
- **MARY K. BROOKHART** (CHRISTIE)
- **DONNA K. THOMPSON** (CHRISTIE)
- **JIM EDGAR** (BUSH)
- **BILL BRADY** (BUSH)
- **RAYMOND POE** (BUSH)
- **KENT GRAY** (TRUMP)
- **SANDRA YEH** (TRUMP)
- **WILLIAM GRAFF** (TRUMP)

Notes: This figure shows a relevant delegate-selection section of the 2016 Republican primary ballot from McLean County in Illinois' 18th Congressional District. See Appendix Figure A1 for a copy of the full ballot.

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 for the candidate to which the delegate was bound. Additional features of these elections, reviewed in Section 4.4, make it unlikely that even voters unaware of this rule would perceive incentives for statistical discrimination. In all the elections we study, about 80 percent of the Illinois delegation is allocated in this manner.[10]

Direct votes for these delegates occur at the congressional-district level as follows. In 2016, Illinois had 18 congressional districts, and 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 institutional environment permit the same inferences as 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. We are not aware of similar natural experiments in other countries.

[10] 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.

the same presidential candidate. For example, Figure 1 provides part of the relevant delegate-candidate section of the ballot from McLean County in the 18th congressional district of Illinois in the 2016 election. While the Jeb Bush and Chris Christie campaigns nominated three delegates with names voters likely perceived as white, the Donald Trump campaign nominated two likely-white delegates and one, Sandra Yeh, who voters likely perceived as nonwhite—0 out of 30 Amazon Mechanical Turk workers we showed this name expected Yeh to be white. A Trump voter who pays "psychic costs" in voting for nonwhites could vote for only Trump's two likely-white delegate candidates and not for Yeh. The performance of Yeh versus her white counterparts on the same ballot is thus one "natural experiment" of the 816 we observe, within which we analyze variation in delegate vote totals.

As Figure 1 shows, names of delegates are printed in large, bold font, followed by the last name of the presidential candidate to whom they are bound. Ballots provide no other information to voters. Delegates are grouped by presidential candidate, such that voters can easily identify delegates bound to their preferred presidential candidate.[11]

We exploit three unique features of this setting to study discrimination. First, to identify discrimination, 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 naturally-occurring control group for estimating discrimination. Second, indicating tastes as the likely mechanism for any observed discrimination, 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 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, such as ideology or competence; here the only relevant dimension is the presidential candidate to which the delegate is bound, which is printed on the ballot. Third, any such discrimination entails trade-offs—the essence of taste-based discrimination—as it undermines voters' preferred presidential candidates' nomination prospects.[12] The election design therefore creates opportunities to credibly identify voter discrimination in this context—and, given the costs of discrimination and the natural absence of incentives for statistical discrimination, suggest taste-based discrimination as the

[11]The same electoral process occurs for alternate delegates, for whom voters vote simultaneously. Our sample pools normal and alternate delegates.

[12]To the extent voters receive expressive utility for voting for their chosen presidential candidate (Pons and Tricaud, 2018; Spenkuch, 2018), engaging in racial discrimination also denies them this utility. See Appendix B for further discussion.

likely mechanism. To fix ideas, we present a formalization of the voter's decision problem in these elections in Appendix B. Section 4.4 and Appendix J discuss further the plausibility of alternative interpretations of the discrimination we observe beyond tastes.

2.3 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. Except in 2000, these primaries occurred relatively early in the primary season, before a "presumptive nominee" was established but nevertheless with a clear front-runner. The median district-level contest was decided by only 2,541 voters. For further context on the elections we study, see Appendix C.

2.4 Candidates for Presidential Convention Delegates

There were 2,386 unique delegate candidacies in Illinois across the four presidential primaries included in this study. We drop from the sample six who ran as uncommitted to a presidential candidate and 62 whose names cannot be coded by gender, as we discuss below.

If voters are more likely to vote for delegates whose names they recognize, and if white or male delegate candidates are especially likely to hold other political positions that would generate name recognition, we may uncover spurious relationships between delegate candidates' race, ethnicity, or gender 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 meticulously detailed "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 non-political posts. 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 2 reports the number of unique delegates in each category. 40.7 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.[13] White delegates are not significantly more likely to be officials than nonwhite delegates, nor is their distribution across levels of office notably different from nonwhite delegates. However, male delegates are somewhat more likely to be officials than

[13]Due to data limitations, these positions reflect the highest office we detected at any time for a delegate. See Appendix Table A1 for descriptive statistics on the number of unique delegates holding each office broken down by specific office.

female delegates, and they are distributed differently across office levels, though the statistical significance of this difference is marginal. Considering the distribution over specific offices in Appendix Table A1, however, we reject the null of independence with respect to delegate race and gender. These results suggest a need to check our estimates of discrimination for bias due to differential officeholding, as we do in Section 4.3.

Table 2: Officeholding by Race and Gender of Delegate

	Count and Percentage of Column				
	By Race		By Gender		
Level of Office	White	Nonwhite	Male	Female	Total
Major Office	18	4	19	3	22
	0.84	2.22	1.14	0.47	0.95
State Legislature	166	11	131	46	177
	7.76	6.11	7.83	7.14	7.64
Minor Office	687	57	553	191	744
	32.13	31.67	33.03	29.66	32.10
No Office	1,267	108	971	404	1,375
	59.26	60.00	58.00	62.73	59.32
Total	2,138	180	1,674	644	2,318

Notes: This table reports the distribution of the delegate population across four office categories by race and gender. 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 MTurk race measure and SSA gender measure at 0.5. The $\chi^2(4)$ statistics by race and gender are respectively 3.95 ($p = 0.27$) and 5.88 ($p = 0.12$). See Appendix Table A1 for a detailed tabulation of delegates by type of official and race: Results differ in the test of independence with respect to gender but not with respect to race. See Appendix D for a discussion of the data collection process.

2.5 The Illinois Republican Primary Electorate

Mirroring national patterns, Illinois Republican primary voters appear to be almost entirely non-Hispanic whites but are approximately balanced with respect to gender. Two separate data sources suggest this same conclusion. First, Appendix Table A2 presents demographic summary statistics on Republican primary voters in Illinois from complete-count administrative voter records from the Illinois Secretary of State with demographics estimated by the firm Catalist. Using their names and neighborhood racial composition, Catalist estimates that, of the people that official records indicate voted in the 2008, 2012, or 2016 Republican primaries, over 95 percent are non-Hispanic white and 51 percent are men. Data from the

2000 election are unavailable. Second, of Illinois Republicans who participated in a large-sample national survey, the Cooperative Congressional Election Study (Ansolabehere, 2010), and who administrative records show voted in the 2008 primary, 97 percent indicated they were white, and 55 percent indicated they were men ($N = 189$ Illinois Republicans who voted in the 2008 primary).

Given the paucity of nonwhite voters, these elections 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 coethnic biases. Conversely, because of statewide and county-level approximate gender balance, we emphasize our results for gender discrimination are *net* results and that we are mute as to whether men and women discriminate in favor or against co-gender delegates.

Catalist data also indicate mean voter age was 60. Voters lived in Census block groups where, on average, one third of residents were college graduates and median annual per capita income was about $70,000. These averages fit with national data which finds that primary voters tend to be whiter, older, more educated, and higher-income than nonvoters.

Propitiously for the external validity of our findings, available data suggest Illinois seems 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 A3 reports data on rates of racially-charged Google searches, 2.4 million results of self-administered Race Implicit Association Tests, the per-capita number of active hate organizations identified by the Southern Poverty Law Center, and the per-capita rate of race-related hate crimes as reported by the Federal Bureau of Investigation. None of these measures 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. We show this in two ways. First, comparing presidential "beauty contest" and delegate vote totals, Appendix Figure A2 shows that delegates receive about 84 percent of the votes cast for their candidate in the "beauty contest" totals, suggesting the vast majority of "beauty contest" voters also participate in delegate contests. Second, Appendix Table A4 shows that primary election turnout was 4–7 p.p. (18–27 percent) higher on average as a share of the voting-eligible population in presidential than in non-presidential election years from 1980 to 2016, controlling for Senate election years, even though the non-presidential years during this period featured contested Illinois governor primaries and the presidential years did not.

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. 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 for each delegate.[14] These county–district intersections are mutually exclusive and exhaustive of Illinois voters and geography. Our unit of observation is each delegate and county–district. Importantly, we do not observe voting at the ballot level, and so we cannot study individuals' joint voting decisions, nor decisions to "undervote" (i.e., not exhaust all N votes for delegates) versus spreading votes among the delegates of multiple presidential candidates.[15]

Our sample spans 2,318 unique delegate candidates and 19,711 vote-count observations, as we observe how a delegate candidate performed in multiple county-congressional district intersections, representing a total of 22.3 million votes. 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 of their preferred 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 specifications we include fixed effects for each cell.

3.2 Inferring Delegate Race and Ethnicity

We measure how voters should perceive delegate candidates' race and ethnicity from their names in three ways: using 2000 U.S. Census data, using a proprietary anthropological

[14] The same geographic unit is used, for example, in Autor et al. (2016). Total votes by district, not county, determine election outcomes. A delegate candidate can only run in one congressional district.

[15] Appendix Figure A2 shows that delegates receive about 84 percent of the votes cast for their candidate in the "beauty contest."

database of full-name frequencies (Onolytics), and using guesses of workers on Amazon Mechanical Turk (MTurk). These three sources yield two objective measures of likely race and ethnicity (Census and Onolytics) and one measure of subjective racial and ethnic perception (MTurk). For all three measures, we distinguish between delegates who are white, black, Hispanic/Latino, and Asian; for Onolytics and MTurk, we further make further distinctions among Asians. Throughout the paper we present estimates using all our measures of the racial and ethnic information in voters' names to demonstrate the robustness of our results.

3.2.1 Census Data on Last Names

Public-use tabulations from the 2000 U.S. Census report, for each last name occurring 100 or more times in Census returns, the count and racial and ethnic composition of individuals with the last name. 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.[16]

Similar to Fryer and Levitt (2004), we match delegates in our sample with the Census racial-composition data. Our measure is the racial composition of the U.S. population with the same last name and is thus continuous. About 87 percent (2,073 of 2,380) of delegates' last names match an entry in the Census data exactly. For the remaining names, we identify the nearest match in the Census data for each delegate last name by minimizing the Jaro–Winkler distance, a common measure of string similarity in record matching. Appendix Table A5 presents estimates including these inexact matches. The results remain similar, confirming that our results are unaffected by omitting delegates with rare names.[17]

3.2.2 Onolytics Classifier of Full Names

We also use a commercial software package to estimate the races of the delegate population. Onolytics is developed in Mateos (2014) and 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,

[16]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.

[17]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 in our context, as Asian and Hispanic names send opposite class signals, yet we find discrimination against both groups.

Indian, Middle Eastern, non-Hispanic white ethnic, and non-ethnic white.[18] 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 MTurk Perceptions of Full Names

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. We paid for 30 guesses for each delegate name to yield reasonably precise estimates of perceived race. Our measure of race is the racial composition of these guesses and is thus continuous. See Appendix F for MTurk survey details and Appendix I for an analysis of attenuation bias due to measurement error.

An advantage of the MTurk measure is we could ask MTurk workers to provide their perceptions in finer ethnic categories than available from the U.S. Census. MTurk workers coded the ethnicity 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 of East Asian, Indian, Middle Eastern, and Pacific Islander heritage.

3.3 Inferring Delegate Gender and Age

We also capture information about delegate gender and age contained in first names.

American first names robustly predict gender. To map delegates to likely genders, we use the baby-name file of the Social Security Administration (SSA) from 1930 to 2012, which covers all individuals born after 1930 and issued a Social Security card. 95 percent of delegates have first names that are either more than 95-percent male or more than 95-percent female in the SSA data. 62 delegate first names cannot be gender-coded and are dropped from our main analysis: These first names are either entirely missing from the ballot and thus our data or cannot be exact-matched in the SSA file. We use a continuous estimate of likely gender, the probability that a delegate is female is the proportion of U.S. citizens born with the same first name who are female at birth.

[18] See Appendix E 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.

First names are also informative about age due to changes over time in the relative name popularity. To map delegates to their likely ages, we use the SSA baby-name file to find the year in which a name attained its maximal share of births. About half of all births occur within 10 years of this "modal year." Among delegates, the median modal year is 1955, with a standard deviation of 20 years.[19]

3.4 Measuring Other Delegate Attributes

We also gathered information on delegates that voters can less readily infer from names and, as such, that they may be less likely to know. We determined delegates' home counties and Census block groups from the residential addresses reported in official candidacy filings. We use these data in two ways below. First, they allow us to restrict the sample to voters who, due to physical distance, are least likely to have information about a delegate beyond what is reported on the ballot. Second, we use block-group characteristics as a proxy for delegate socioeconomic status: in particular, the share with at least a bachelor's degree and log median household income.

3.5 Descriptive Statistics on Delegate Race and Ethnicity

To illustrate the racial and ethnic information in delegate names, Table 3 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 A5, 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.[20]

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 throughout the results presented in Section 4. We report the results from the principal component analysis, including for the detailed race categories, in Appendix Table A7. Overall, we find that same-race, different-measure correlations—for instance, the MTurk black measure's correlation with the Census black measure—are robustly positive. 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

[19]See Appendix Figures A3 and A4 for histograms of SSA percent female and modal year of birth.
[20]Appendix Table A6 gives examples of highly suggestive names for all racial and ethnic categories.

Table 3: Whitest and Least White Names of Delegate Candidates

	Census		MTurk	
	Name	Whiteness	Name	Whiteness
Whitest	Carol Hornickle	0.9946	Jill Bess	1
	Brian Milleville	0.9956	Helen Manson	1
	Sherry Hellmuth	0.9958	Mike Marron	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	Angel Garcia	0
	Neil V. Patel	0.0155	Gustavo Gonzalez	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. Ties are resolved by random draws. The categorical definition of Onolytics race variable means there is no equivalent ranking of names by informativeness. For further detail by race category, see Appendix Table A6.

our estimates toward zero, motivating the use of principal component analysis to extract the common signal.

Using the modal guesses of delegate race from the MTurk data, the delegate 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 delegates. Appendix Table A9 shows that likely-nonwhite delegates are nominated by nearly all the presidential candidates in our sample. Our results are thus not driven by a single presidential candidate's voters.

3.6 Where and Why Minority Delegates Are Nominated

Campaigns recruit and nominate delegates. This introduces the potential for two varieties of selection. First, the nonwhite or female delegates that candidates select may differ from the white and male delegates. We take up this concern in Section 4.3. Our results weigh strongly against the presence of important confounding delegate-level unobservables.

Another form of selection is that campaigns may be more likely to nominate nonwhite or female delegates in areas of Illinois with specific characteristics. This presents a threat to external validity but, importantly, would not bias our results due to the presence of fixed effects which restrict our comparisons to within-cell variation. We evaluate this external-validity concern in Appendix G. We regress the shares of nonwhite and female delegates on several county–district observable characteristics, including the white and college-educated shares of population, white per-capita income, and the Republican two-party vote share. We find that campaigns frequently nominate nonwhites and women throughout the state, although relatively more frequently in less-Republican areas. We detect no other differences. The degree of selection of cells into the identifying set is sufficiently mild that the set of cells that contribute to identification closely matches of Illinois on average. In addition, we provide direct evidence that selection on county–district observables is not a threat to external validity with respect to Illinois statewide. Using coarsened exact matching, we reweight the sample so that cells with likely-nonwhite delegates match Illinois statewide on the four observables above. We estimate a similar penalty against nonwhites, implying that nonwhites are not more likely to run in areas where voter discrimination is lower.

If voters discriminate against nonwhite candidates, why would campaigns nominate them? In discussions with several officials responsible for recruiting delegate candidates, we found that recruitment costs were a common explanation, consistent with our finding in Appendix G that nonwhites were more likely to be nominated in less-Republican areas, where the supply of Republicans who could serve as delegate candidates may be more limited.[21]

4 Racial and Gender Discrimination by Voters

4.1 Empirical Strategy

To estimate the effects of discrimination on voting behavior, we compare the vote totals of delegates who differ in race, ethnicity, or gender but are in the same "cell": delegates running 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—all of whom voters should select to maximize the value of their ballot. Our baseline Poisson regression specification is:

$$\mathbb{E}[\text{Votes}_{ipct}] = \exp\left(\beta \cdot \text{Nonwhite}_i + \gamma \cdot \text{Female}_i + \mathbf{X}'_{ipct}\delta + \alpha_{pct}\right), \tag{1}$$

where Nonwhite_i and Female_i are our proxy variables for whether voters believe delegate i is nonwhite and female, respectively. α_{pct} is a vector of cell fixed effects, where p denotes the

[21] Campaigns must secure three supporters to run in every congressional district months before the primary takes place, and these delegates must agree to pay their own travel and lodging to attend the convention if elected. Campaigns therefore face search costs in locating delegates willing to serve. In addition, Illinois campaign managers as agents may not fully internalize the presidential candidates' incentives. For example, the Trump campaign appears to have recruited from a campaign email list (Brueggeman, Brian. 5 March 2016. "Meet your delegates: the people who will vote for presidential candidates at the conventions." *Belleville* [IL] *News-Democrat*). Appendix H further argues delegate service is a form of hobbyist consumption.

presidential candidate, c the county–district, and t the election year. For our estimates of β and γ to be unbiased, race and gender must be uncorrelated with unobservable determinants of votes. To show robustness, we add a vector of covariates $\mathbf{X}_{ipct}$ in some specifications.

Estimates of β and γ can be interpreted as the average percentage of votes lost or gained due to discrimination by race and gender. 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. In alternative specifications, we replace Nonwhite_i with variables for specific nonwhite race and ethnicity groups. The dependent variable is the vote count for a delegate, and the unit of observation is the county–district-delegate-year.[22] Standard errors are clustered at the delegate level, as this is the level at which the "treatment," a delegate's race, is assigned.[23] The presence of cell fixed effects in all regressions ensures that coefficients reported in the tables only exploit variation in the performance of delegate candidates within the same cell.

4.2 Main Results

Table 4 presents our main results. Column 1 reports estimates of Equation 1 using the share of MTurkers who perceived each delegate as nonwhite. The coefficient 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 8.7 percent fewer votes in these elections. Column 2 breaks down these estimates by delegate race and ethnicity. Due to the small number of likely-black delegate candidates, our estimate of discrimination against black candidates is relatively imprecise. Delegates with names MTurkers perceive as Asian and especially as Indian and Middle Eastern appear to perform much worse than their counterparts running to represent the same candidates, receiving about 15 percent fewer votes than white delegates in the same cell. Using our SSA data on the female share of first names to code delegates who are objectively likely to be female, we find little evidence for discrimination against women in both Columns 1 and 2. Conditional on being in the

[22]Other specifications, such as unweighted OLS on the number of votes or on log votes, would not estimate a meaningful quantity of interest due to heterogeneity in the number of votes by presidential candidate. See Appendix Tables A10 and A11 for OLS regressions with two alternative dependent variables, respectively vote shares and $\ln(1 + \text{votes})$, weighted by votes. Both find similar results.

[23]We cluster standard errors at this level because the implied randomized experiment is that the same individual delegate candidates were randomly assigned to switch race or gender with other delegate candidates running in their same cell. We present a permutation test later in the paper that implements this implied experiment under the sharp null hypothesis. Appendix Table A14 reports our main results clustering at the level of congressional district, presidential candidate, alternate or regular, and year, which replicates the "slate" of delegates chosen by the same presidential candidate and competing for the same position. We find this increases standard errors by only about 20 percent on average and thus leaves the significance of our results almost entirely unchanged.

Table 4: Effect of Delegate Candidate Race and Gender on Votes, by Race and Ethnicity Measure

	MTurk		Census		Onolytics		Rescaled PC1	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	0.006	0.006	0.004	0.001	0.005	0.004	0.005	0.001
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Nonwhite	-0.087***		-0.045***		-0.035***		-0.092***	
	(0.016)		(0.014)		(0.007)		(0.016)	
Black		-0.033		0.018		-0.040***		-0.094***
		(0.050)		(0.025)		(0.015)		(0.032)
Hispanic/Latino		-0.057***		-0.061***		-0.045***		-0.079***
		(0.018)		(0.018)		(0.012)		(0.018)
Any Asian				-0.069***				-0.110***
				(0.014)				(0.022)
East Asian		-0.093***				-0.055***		
		(0.025)				(0.008)		
Indian		-0.174***				-0.076***		
		(0.040)				(0.013)		
Middle Eastern		-0.160***				-0.088***		
		(0.031)				(0.033)		
White Ethnic						-0.016*		
						(0.009)		
N	18,958	18,958	16,945	11,166	18,639	18,639	16,668	11,049
Pseudo-R^2	0.991	0.991	0.991	0.992	0.991	0.991	0.991	0.992

Notes: This table presents the results of estimating Equation 1, yielding estimates of the percentage vote penalties by nonwhite race or ethnicity and by gender. "Any Asian" uses the Census definition of Asian race, which spans our subsequent categories of East Asian, Indian, and Middle Eastern. Column 7 provides our preferred estimates throughout this paper. In all regressions the dependent variable is the vote count for the delegate. The unit of observation is the county–district-delegate-year. All regressions include cell-level FEs. Delegates are defined as in the same cell if they 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. Sample sizes change because not all delegate names can be classified using the Census data or Onolytics algorithm. Appendix Table A13 shows versions of Columns 7 and 8 that estimate the female and race/ethnicity coefficients in separate rather than combined regressions. Standard errors are clustered at the delegate level. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

same cell, likely-female delegates receive approximately the same average number of votes as likely-male delegates. Appendix Table A15 shows that interactions between delegate race and gender are insignificant.

In Columns 3 and 4, we present our estimates using Census data to code the racial information in delegates' last names. The sample is limited to delegates whose last names match the Census data exactly. Similar to the MTurk measure, we find that nonwhite delegate candidates receive fewer votes. On average, a delegate who was objectively likely to be nonwhite would receive approximately 4.5 percent fewer votes than a delegate objectively likely to be white. We find significant discrimination against delegates likely to be Hispanic or Asian, but Census data struggle to identify delegates likely to be black by last name.[24] We again find a tight zero for discrimination against women.

In Columns 5 and 6, we present estimates using the dichotomous Onolytics race categories. As we found using the MTurk and Census race variables, we estimate that delegate candidates identified by the Onolytics algorithm as nonwhite receive fewer votes. Broken down by ethnicity, we find significant shares of voters do not vote for black, Hispanic/Latino, Asian, Indian, and Middle Eastern delegate candidates. We also break out a "white ethnic" category but find only a weakly significant difference in votes between "ethnic" and "non-ethnic" whites. The null result for women is unchanged.

Columns 7 and 8 present estimates using the rescaled PC1 measure. In Column 7, which provides the preferred estimates of this paper, the coefficient on nonwhite implies that delegate candidates who are generally identified as nonwhite across the three measures receive 9.2 percent fewer votes.[25] We also obtain a precise point estimate of zero discrimination against female delegates. In Column 8, we break this result down by the three racial categories common across our three measures. We find significant discrimination against delegates of all three racial categories. We also treat these estimates as our preferred results for the detailed categories throughout the rest of the paper.

The large change in the black coefficient when using the PC1 index is due to the substantial measurement error in our measures for this category and by a small number of outlier observations that influence our rescaling of the PC1 index. Given these challenges, our estimates for discrimination against black delegates should be interpreted with caution. More generally, differences in sample coverage and differences in how variables measure race and

[24] Although the mean percentage black by last name in Census data is 9.8 percent, the 95th-percentile black last name is only 35 percent black.

[25] To guard against bias in 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. Appendix Figure A6 plots 10,000 draws from a Monte Carlo simulation of the main regression specification with treatment status permuted in this way. Our standard errors appear unbiased relative to the bootstrap. The estimated effect remains significant.

ethnicity explain differences in estimated coefficients.[26]

The PC1 index is missing when any of its constituent parts are missing, as the Onolytics or Census exact measures sometimes are. The gender measure is also missing for some delegates, as previously noted. Appendix Table A16 shows that imputing missing values for the PC1 index and for gender does not change the results. In Appendix I, 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.[27] Appendix Table A17 shows that our main results are essentially unchanged if the race and gender variables are dichotomized.

4.3 Robustness Checks

Our research design absorbs into fixed effects all attributes that may affect voting behavior, such as attributes of presidential candidates, but do not vary among delegate candidates who appear on the same ballots, in front of the same voters, in the same geographic areas, and bound to the same presidential candidates. A potential confound must therefore cause some delegate candidates to receive more or fewer votes than same-race or same-gender delegate candidates in the same cell and be correlated with delegate-candidate race or gender.

4.3.1 Differences in Local Political Networks

One such possible confound is a racial or gender differential in local political networks and serves to illustrate the main empirical challenge to our results. If white delegate candidates are better known to voters than nonwhite delegate candidates, for instance, 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 local political networks

[26]For example, the Census race measure is constructed exclusively from delegate last names, although for some groups (e.g., blacks), first names are a stronger signal of race. In addition, the MTurk measure captures some differences between how individuals perceive race and ethnicity from names that diverge from objective data. For instance, a majority of MTurk respondents perceived a delegate with the last name "Mercadante" as Hispanic, but the name, per Onolytics, is Italian in origin and is coded by the Census as white.

[27]We find that attenuation due to sampling is trivial. Attenuation due to noisy proxies for perceived race, however, may be substantial. A correction using Cronbach's (1951) α implies that being perceived as nonwhite may reduce the number of votes a delegate receives by 11 percent. Being perceived as black, most notably, may reduce the number of votes by 41 percent. Estimates for Hispanics/Latinos and Asians, whose names more clearly indicate race, rise by comparatively less. These estimates require the strong assumption that disagreement among the race measures is entirely classical measurement error. If the non-common components of these measures affect voting with the same sign as the common component, the reliability correction will overstate the true magnitude of discrimination. The estimates we report in Appendix I are therefore most reasonably viewed as upper bounds.

as well as the direct effect of delegate-candidate race. A parallel logic could apply for gender. 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 and gender that they use to determine their votes. Columns 1 and 2 of Table 5, which use the PC1 race measure, show that our results for both nonwhites and women are essentially unchanged when we control for repeat candidates. While repeat delegate candidates receive significantly more votes, repeat delegate status is not strongly correlated with either delegate race or gender.[28]

Second, Columns 3–6 of Table 5 control for voters' delegate-level prior information using our delegate "background checks," as described in Section 2.4, which exhaustively collected information on other offices or political roles held by delegates. In Columns 3 and 4 we include indicators for each of the 17 types of offices we recorded. We find significant returns to officeholding, in line with the literature on candidate name recognition (e.g., Panagopoulos and Green, 2008). Accounting for differential officeholding by race and gender modestly increases our estimate of discrimination against nonwhites but, in contrast, suggests that women receive 2 percent more votes than comparable men, a significant difference.[29] In Columns 7 and 8, we drop from the sample every delegate for whom we were able to find had held or previously ran for any office, no matter how minor. 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 10 percent fewer votes than white delegates. Such female delegates receive about 2–3 percent more votes than comparable male delegates.

It remains possible that white or female delegates are more likely to be highly-connected individuals in ways our "background checks" could not capture but that would increase their vote totals. We provide further evidence against this possibility by exploiting the facts that some congressional districts in our sample span large areas, often hundreds of miles from end to end, and that we can observe outcomes by county within each congressional district. Highly-connected individuals should benefit from connections principally in their home counties, and indeed we find delegates receive 8 percent more votes in their home counties as recorded in official candidacy filings. In Columns 7 and 8 of Table 5, we show

[28] Appendix Table A18 reports the coefficients on the covariate terms.

[29] This point estimate matches the small advantage female candidates enjoy on average in survey-based experiments, per a recent meta-analysis (Schwarz and Coppock, 2019).

Table 5: Robustness Checks

	With Controls For:									
	Repeat Candidates		Officeholders				Non-Home Counties Only		Ballot Order	
			Control		Non-Officeholders Only					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Female	0.006	0.002	0.020***	0.020***	0.031***	0.019***	-0.002	-0.002	0.010*	0.009*
	(0.006)	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)	(0.005)	(0.005)
Nonwhite	-0.088***		-0.099***		-0.099***		-0.096***		-0.101***	
	(0.017)		(0.015)		(0.019)		(0.022)		(0.014)	
Black		-0.095***		-0.070***		-0.056		-0.081**		-0.073***
		(0.033)		(0.026)		(0.040)		(0.035)		(0.027)
Hispanic/Latino		-0.074***		-0.080***		-0.059***		-0.078***		-0.087***
		(0.017)		(0.015)		(0.017)		(0.024)		(0.008)
Asian		-0.100***		-0.095***		-0.173**		-0.224***		-0.113***
		(0.023)		(0.020)		(0.083)		(0.057)		(0.022)
N	16,668	11,049	16,668	11,049	8,091	5,068	14,422	9,536	16,668	11,049
Pseudo-R^2	0.991	0.992	0.992	0.992	0.995	0.994	0.990	0.990	0.991	0.992

Notes: This table reports the results of estimating Equation 1 with varying sets of controls or sample restrictions. In all regressions above, the dependent variable is the vote count for the delegate and the race measure used is the rescaled PC1 measure. The unit of observation is the county–district-delegate-year. For coefficients on controls, see Appendix Table A18. All regressions include cell-level FEs. Delegates are defined as in the same cell if they 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. Standard errors are clustered at the delegate level. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

our results are robust to dropping delegates' home counties, suggesting our results are likely robust to any racial or gender differences in unobservable local political networks.

In summary, these tests of observable and unobservable differences among candidates suggest that our estimates of discrimination are unlikely to be confounded by differences in officeholding or voter information. We find that our estimates of racial discrimination remain unchanged with controls for possible sources of this variation and that these possible sources appear uncorrelated with delegate race and ethnicity.[30] On the other hand, controlling for officeholding leads us to find significant, if small, discrimination towards women, rather than the null result before controlling for officeholding, as female delegates are less likely than male delegates to hold offices with substantial electoral returns in the context of this primary. We interpret these results as ambiguous with respect to whether there is no discrimination against women or discrimination towards them.

4.3.2 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, Appendix Table A12 shows that, conditional on the number of ballot slots available (2, 3, or 4), delegate race and ballot order are uncorrelated, as are delegate gender and ballot order. We augment our specification in Equation 1 with controls for ballot order using dummy variables for the rank (1–4) of a delegate among those in the same cell and interact these with the maximum number of delegates (2–4) for whom a voter may vote in a given congressional district and year. Columns 9 and 10 of Table 5 report these results. Our findings change little when controlling for ballot-order effects.

4.3.3 Other Types of Discrimination

Voters may have preferences over other attributes of candidates, such as age, education, or income. In Section 3, we describe our data on these attributes: We use the age that voters might infer from a delegate's first name and the block-group characteristics of the delegate's official residential address as proxies of their education and income. To the extent that voters know the socioeconomic status of delegates, these measures allow us to detect whether there

[30] This implies nothing about whether nonwhites are unconditionally 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.

is discrimination on these dimensions. Appendix Table A19 reports our results: We find precise nulls on discrimination by estimated age, education, and income.

These results suggest our estimates of racial discrimination are unlikely to be contaminated by discrimination on characteristics correlated with race, such as age or socioeconomic status. Fryer and Levitt (2004) demonstrate this threat to inference by showing that black Americans whose parents give them racially-distinctive first names are lower on average in socioeconomic status. In our context, this concern is relatively unlikely a priori. 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 status from distinctively-black first names do not 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.

4.4 Residual Incentives for Statistical Discrimination

While the design of the primary suggests taste-based discrimination as the likely mechanism for our findings, we discuss here the possibility that voters have residual incentives for statistical discrimination. In Appendix J, we also discuss other alternative interpretations: (1) that voters are making inferences about presidential candidates from delegates, (2) unobserved confounds in general, (3) the implications of voter indifference between Presidential candidates, (4) whether voters use delegate voting to send signals to presidential candidates or party elites, and (5) whether voters understand whether delegate voting has stakes. There we also report an original survey of self-identified Illinois Republican primary voters.

An interpretation of the discrimination we observed as taste-based relies upon the assumption that rational voters have minimal incentives to engage in statistical discrimination. Following Becker's (1957) definitions of taste-based and statistical discrimination as with respect to how one would interpret observed behavior as if it were undertaken by rational agents, it is difficult to see why a rational agent in this setting would perceive incentives for statistical discrimination. Even a rational agent who misunderstood the institution, not knowing that delegates were bound, would need to maintain very unlikely beliefs: A substantial fraction of rational agents would need to believe that white delegates selected by the opposing candidates would be more likely to vote for their candidate of choice at the convention than nonwhite delegates selected by their candidate of choice. Moreover, as the elections we study had largely narrowed to two contenders, voters who believed delegates were likely to abstain or defect to their preferred presidential candidate's rival have only the alternative of helping that rival. Under such conditions, it would not be sufficient for a voter to believe nonwhite delegates were more likely to defect than white delegates for the same presidential candidate. For example, to explain our results in 2012, 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.[31] That convention rules do not allow delegates to engage in this behavior in the first place makes it all the more implausible.[32]

Several pieces of empirical evidence also are inconsistent with the presence of statistical discrimination. First, that our 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—is inconsistent with voters inferring nonwhites are more likely to abandon conservative presidential candidates to vote for liberals; voters voting for more liberal Republican presidential candidates discriminate no less. This result also suggests our findings are unlikely to come from voters concerned that nonwhite delegates would move party ideology left or right through their presence at the convention or in potential future political careers. Second, in Appendix J we present results from a survey we conducted of a convenience sample of Illinois Republican primary voters that finds that perceived differences in the loyalty of white versus nonwhite delegates are much too small to plausibly explain our finding of discrimination against nonwhite delegates. In contrast, about 10 percent of voters instead said they would avoid voting for nonwhite delegates because they were "uncomfortable" doing so, despite potential social desirability bias, consistent with "psychic costs." Third, although the risk of further rounds of convention balloting (sometimes called a "contested convention") was heightened in 2016, Table A25 finds that voter discrimination does not vary meaningfully across years, including in 2000, 2008, and 2012 when the risk of further rounds of balloting at the convention was not considered plausible by the time

[31] A corollary of this observation is that even if one were to adopt a broader definition of statistical discrimination that includes choices stemming from mistaken beliefs (Bohren et al., 2019), it seems unlikely that behavioral voters with such beliefs would perceive incentives to discriminate.

[32] The sole case in which delegates do have discretion is if the convention is contested and goes to a second round of voting. 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. However, as reviewed in Appendix C, the primary race had progressed sufficiently to a two-candidate race by the time Illinois voted in 2000, 2008, and 2012 that multiple rounds of balloting were essentially impossible, and our estimates remain largely unchanged when examining the two frontrunners in 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.

Illinois voted. Finally, our discussion of the nature of convention in Appendix H shows that voters who discriminate are unlikely to be trying to prevent political power from accruing to nonwhites.

5 Heterogeneity in Discrimination

We next estimate how racial and gender discrimination varies along several dimensions. Although some of these tests have low statistical power and we caution that they are fundamentally observational in nature, their results are consistent with a taste-based interpretation.

5.1 By Presidential Candidate Race and Gender

Would the discrimination we observe in these delegate elections also manifest in other voting decisions, such as in choices between presidential candidates themselves? If our estimates reflect discriminatory tastes, and if these discriminatory tastes also influence voters' choices of presidential candidates, we would expect presidential primary candidates who are nonwhite or female to attract voter populations with weaker tastes against nonwhites and stronger tastes towards women on average than the voter populations of white or male presidential candidates. We can test this prediction because the design of the primary allows us to separately estimate the magnitudes of race and gender discrimination among voters for each presidential candidate.[33]

In Column 1 of Table 6, we show that voters for nonwhite presidential candidates indeed do not appear to have any racially discriminatory tastes on average, whereas voters for white presidential candidates do. Column 1 estimates Equation 1 with an interaction term for the race and ethnicity of the presidential candidate with the race of the delegate candidate as well as controls for delegate officeholding and ballot order. As with delegate candidates, for concision we define presidential candidates as white if they are non-Hispanic whites alone. We estimate that nonwhite delegates of white presidential candidates lost 10 percent of votes due to discrimination. A χ^2-test rejects equality between delegates of white and nonwhite presidential candidates. This suggests that voters for nonwhite presidential candidates have, on average, weaker racial tastes than voters for white presidential candidates. Although we cannot rule out other explanations for this pattern, it is consistent with our estimates of discrimination as reflecting tastes that also affect voters' presidential candidate choices.

A similar result applies to women, consistent with the prediction that voters for female

[33]Appendix Figure A7 presents the estimated level of discrimination among voters for individual presidential candidates.

Table 6: Heterogeneous Treatment Effects by Presidential Candidate Race and Gender

	(1)	(2)	(3)
Nonwhite Delegate			
× White Pres. Cand.	-0.101***		-0.104***
	(0.017)		(0.017)
× Nonwhite Pres. Cand.	0.052		0.053
	(0.064)		(0.069)
Female Delegate			
× Female Pres. Cand.		0.335***	0.335***
		(0.096)	(0.109)
× Male Pres. Cand.		0.003	0.003
		(0.006)	(0.007)
χ^2 test of equality	7.07***	12.49***	15.73***
N	17,126	18,958	16,668
Pseudo-R^2	0.990	0.991	0.991

Notes: This table reports the results of estimating Equation 1, interacting with the race and gender of presidential candidates. We code Ben Carson, Ted Cruz, Alan Keyes, and Marco Rubio as nonwhite (including Hispanic), Carly Fiorina as female, and the remaining 19 candidates as white men. The χ^2 test results also include interactions of the nonwhite and gender variables with the year, thus exploiting only within-year comparisons of presidential candidates by attribute. In all regressions the dependent variable is the vote count for the delegate. The unit of observation is the county–district-delegate-year. All regressions include controls for ballot order and detailed officeholding and cell-level FEs. Delegates are defined as in the same cell if they 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. Standard errors are clustered at the delegate level. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

presidential candidates may similarly be selected towards relatively pro-female tastes. However, there was only one female Republican presidential candidate whose delegates won any votes in Illinois from 2000 to 2016, Carly Fiorina in 2016, and so our results come with the immediate caveat that they rely upon vote totals for one female candidate. Comparing the vote totals of female and male Fiorina delegates in the same cell, we find substantial discrimination towards female delegates: We show in Column 2 of 6 that female Fiorina delegates win about 35 percent more votes than otherwise-similar male Fiorina delegates, a statistically significant but imprecise estimate. The difference with non-Fiorina delegates is also significant in a χ^2-test.[34]

We also investigate whether discrimination varies by presidential candidate ideology, categorizing moderate and conservative presidential candidates on either side of the median

[34]See Appendix Table A15 for evidence that the presence of Fiorina voters with the highest estimated bias against nonwhites and for women is not explained by missing interaction terms.

of ideology estimates inferred from the identities of campaign donors from Bonica (2013). Reported in Appendix Table A21, the relationship is insignificant, suggesting little sorting of voters who discriminate more according to left–right ideology. Importantly, this null result is also inconsistent with statistical discrimination wherein voters believe white delegates are more conservative than nonwhite delegates. Voters for moderate presidential candidates would have a weaker incentive to discriminate than voters for conservative presidential candidates, but we do not find this.[35]

5.2 By Competitiveness and the Costs of Discrimination

Taste-based theories of discrimination predict individuals discriminate less when it is more costly for them to do so (Becker, 1957; Hedegaard and Tyran, 2018). In this setting, voters face a trade-off between any 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' minority delegates would prove decisive, causing the delegates to lose and impairing their preferred presidential candidates' nomination prospects. Other components of the cost of discrimination, such as the intrinsic utility voters gain from the act of voting for their chosen candidate (Pons and Tricaud, 2018; Spenkuch, 2018), do not vary with the probability voters will be decisive here.

To evaluate this prediction, we split presidential candidates by whether they received above- or below-median shares of statewide votes in each respective election year. In our context, below-median candidates have essentially no chance of winning delegates, and thus their voters face lower instrumental costs of discrimination than voters for above-median candidates. Appendix Table A21 presents results consistent with a downward-sloping demand curve for taste-based discrimination: We estimate a 9-percent penalty for the nonwhite delegates of above-median presidential candidates, compared to a penalty of 57 percent for below-median candidates. A χ^2-test confirms this difference is significant. When rational voters with tastes against nonwhites vote in an election they expect to be close, they appear to be more likely to prioritize the victory of their preferred presidential candidate over avoiding the "psychic cost" of voting for nonwhites than when no delegates are plausibly at stake. We interpret these patterns as suggestive evidence in favor of viewing the discrimination we detect as taste-based and as politically consequential even in competitive elections.[36]

[35] A lack of other proxies for ideology which cover Republican presidential candidates constrains further analysis of the relationship between ideology and discrimination.

[36] Discrimination persists even among the top two presidential candidates' voters in each year ($\beta = -0.087$, $p < 0.001$) as well as for the top presidential candidates' voters ($\beta = -0.083$, $p < 0.001$). We repeat this and the ideology analyses for gender in Appendix Table A21. Preventing further analysis of voter responses to instrumental incentives, nearly all variation in pivot probability is at the candidate level, and

5.3 By Geography

We also examine whether our area-level estimates of racial discrimination vary with 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 they correlate with discrimination, we modify Equation 1 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 counties within delegates, adding delegate fixed effects allows electorate demography to vary while holding delegate identities constant.

Appendix Table A23 reports these results. There is no significant association between our county-level estimates of racial discrimination and the white share of population. Discrimination is lower in areas with higher college-educated white shares of population, with higher per-capita income of whites, and with lower two-party Republican presidential vote shares, although significance is often sensitive to specification. These findings are consistent with historical patterns wherein American political parties that explicitly appeal to racial prejudice tend to perform better in lower-income areas (e.g., Mulkern, 1990). We find little evidence that the effect varies 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.[37]

A parallel analysis of gender discrimination reaches similar conclusions. Appendix Table A24 finds that discrimination towards female delegates is stronger in areas where the college share of adults is higher and per-capita income is higher. These results, however, are not robust in within-delegate analyses. In lieu of local measures of gender bias, we consider three variables drawn from U.S. Census data: the adult sex ratio, the log male-female difference in annual labor earnings, and the male-female difference in the percentage of individuals over age 25 with at least a bachelor's degree. We also consider the average Republican vote share.

within-candidate geographic variation is trivial. Another source of variation in the instrumental cost of discrimination is whether or not the delegate position is for a regular or an alternate delegate. See Appendix Table A22. We find only weak evidence for the hypothesis that racial and gender discrimination is greater against alternate delegates than against regular delegates. In tension with our results for competitiveness, the similarity of our estimates for alternates and delegates suggests an important role for disutility from the mere act of voting for a nonwhite delegate, irrespective of their probability of serving. Given these null results and the possibility that voters for presidential candidates who receive relatively few votes may have relatively stronger tastes against minorities, we regard this analysis of instrumental costs as suggestive.

[37] The lack of correlation across these measures may be attributed variously to multidimensionality in the 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 potentially other factors. The correlation between the Stephens-Davidowitz (2014) and Xu et al. (2014) measures is only 0.1.

We find no association with gender discrimination for any of these variables in either the pooled or within-delegate analyses.

5.4 Over Time

We estimate Equation 1 but interacting the nonwhite and female variables with year dummies. Appendix Table A25 presents the estimates by year. An χ^2-test of the equality restriction across years shows that we cannot reject either null hypothesis that discrimination is constant over time. We also impose linear time trends on racial and gender discrimination, starting from Equation 1, by interacting the variables with the continuous year variable. The point estimates on these interaction terms are essentially zero for both race and gender.

6 Discussion

Our estimate of racial discrimination in these elections is large, comparable in magnitude to the benefits in these elections of being a state legislator or running in one's home county. This section reviews the political and policy implications of our findings.[38]

We first evaluate the consequences of racial discrimination in Illinois Republican delegate selection by simulating counterfactual outcomes absent discrimination. To simplify, we assume delegates more likely than not to be nonwhite according to the MTurk race measure all lost the same fraction of votes due to discrimination and then vary the estimated penalty from 0 to -0.3. We augment their vote totals by these fractions, which assumes that discriminating voters either undervote or proportionally vote for other delegates when nonwhites are nominated. We then calculate which delegates would have won under these augmented vote totals. This exercise suggests that discrimination reduced the number of nonwhites who served: Appendix Figure A8 shows that, at our preferred estimate of discrimination against nonwhite delegates of 9.2 percent, 5 additional nonwhite delegates would have won, relative to the baseline of 34. In addition, discrimination reduces the appeal of election winners on non-racial dimensions to voters who discriminate: We estimate that, due to discrimination, several nonwhite delegates lost to white delegates bound to other presidential candidates. Appendix Table A27 provides an example of one such likely change in outcomes. Voters in

[38] We caution these are partial-equilibrium analyses. For instance, incentives of presidential candidates to nominate a nonwhite as a delegate, and those of nonwhite candidates to run for other offices, may also change in a counterfactual without voter racial discrimination. Plausible general-equilibrium effects in other elections, such as on the margins of nonwhite candidate entry or institutional design (Trebbi et al., 2008), could mean our estimate of the increase in nonwhite representation from the elimination of racial discrimination could be over- or understated.

this setting appear willing to discriminate despite the potential for discrimination to change outcomes and advantage disfavored presidential candidates.

Beyond the narrow context of Illinois Republican primaries, discrimination of the magnitude we observed would also be large enough to represent a substantial barrier to the election of racial minorities in other elections. Appendix Table A26 presents back-of-the-envelope calculations that use U.S. House primaries from 1990 to 2010 to illustrate the potential magnitude of that barrier. As these calculations require the strong assumption that our estimate of discrimination in Illinois delegate-primary elections is valid for U.S. House primaries, we offer them to contextualize the substantive importance of our estimated magnitude of discrimination, not to reach precise conclusions about nonwhite representation in the U.S. House in a counterfactual world without racial discrimination. We estimate nonwhites would have won 19 additional Republican U.S. House primaries from 1990 to 2010 absent racial discrimination of the size we observed in this setting. This would result in an increase of about 9 percent in the number of nonwhites winning Republican primaries ($19/218 = 0.087$). These comparisons suggest racial discrimination is plausibly a critical barrier to minority political officeholding, which other research shows lies at the root of important political and economic racial disparities.

Would such discrimination manifest in other elections? Encouragingly for the external validity of our findings, we found that voters appear to select out of voting for nonwhite and female presidential candidates in a manner correlated with our estimates of their collective racial and gender tastes, consistent with discriminatory tastes having stakes for presidential candidate choices. Of course, many considerations influence whether discrimination is likely to be greater or smaller in magnitude in other elections. Both costs and benefits of discrimination to voters may vary. On the one hand, discrimination could be greater in other elections where voters may have less information or weaker preferences. For example, voters may know less about candidates in state legislative primaries than in presidential primaries and therefore have weaker candidate preferences; all else equal, we would expect greater discrimination there. Likewise, our estimates would not capture behavior arising from the anticipated psychic costs of having nonwhite officeholders, nor from any statistical discrimination against them. On the other hand, discrimination could be smaller in general elections where partisan preferences may be more important, or in elections where voters have on average weaker racial tastes. For example, in a general election between a nonwhite Republican and a white Democrat, Republican voters with strong racial tastes may still prefer paying the psychic costs of voting for the nonwhite candidate to voting for a Democrat.

In the U.S., credible evidence about discrimination in elections also has significant policy implications. The U.S. Congress and the Supreme Court regularly consult academic assessments of voter discrimination in crafting and reviewing American election laws. Our results are most relevant to the debate over the Voting Rights Act of 1965, passed in part to facilitate the election of nonwhites to political office.[39] In the subsequent decades, Congress and the Court have evaluated the continuing necessity of the Act in part by attempting to answer whether nonwhite candidates still encounter discrimination in elections. In decisions in 2009 and 2013, the Supreme Court struck down portions of the Act, finding only "decades-old data" in support of federal claims of continued discrimination. Our findings are particularly applicable to vote-dilution cases under the Act in which plaintiffs seek the construction of majority–minority districts. First, our finding of racial and ethnic discrimination in voting behavior may meet evidentiary standards for injury often difficult to meet with existing correlational evidence. Second, by estimating significant discrimination against several nonwhite minority groups, we provide evidence of shared injury, often the relevant legal burden in "coalitional" cases brought jointly by minority groups under the Act.

Looking beyond the Act, our findings suggest policy responses that operate only on informational margins to reduce statistical discrimination, however effective (Casey, 2015), are likely to leave intact substantial barriers to the election of racial and ethnic minorities arising from taste-based discrimination. However, our work lends support to other work that finds political parties may be a more relevant constraint on female representation than voters (Baltrunaite et al., 2014, 2019; Esteve-Volart and Bagues, 2012; Casas-Arce and Saiz, 2015; Besley et al., 2017; Bhalotra et al., 2017; Schwarz and Coppock, 2019). The contrasting evidence that voters discriminate by race and ethnicity suggests that different interventions may be necessary to address nonwhite and female underrepresentation in political office.

References

Akerlof, George, "The Economics of Caste and of the Rat Race and Other Woeful Tales," *Quarterly Journal of Economics*, 1976, pp. 599–617.

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

_ **and Brian F. Schaffner**, "CCES Common Content, 2016," 2017.

Anzia, Sarah F and Christopher R Berry, "The Jackie (and Jill) Robinson Effect: Why Do Congresswomen Outperform Congressmen?," *American Journal of Political Science*, 2011, *55* (3), 478–493.

[39] We thank Christopher Elmendorf and David Schleicher for helpful discussions about implications of our results with respect to the Act.

Autor, David, David Dorn, Gordon Hanson, and Kaveh Majlesi, "Importing Political Polarization? The Electoral Consequences of Rising Trade Exposure," Working Paper, NBER 2016.

Baltrunaite, Audinga, Alessandra Casarico, Paola Profeta, and Giulia Savio, "Let the Voters Choose Women," *Journal of Public Economics*, 2019, *180*, 104085.

_ , Piera Bello, Alessandra Casarico, and Paola Profeta, "Gender Quotas and the Quality of Politicians," *Journal of Public Economics*, 2014, *118*, 62–74.

Beaman, Lori, Esther Duflo, Rohini Pande, and Petia Topalova, "Female Leadership Raises Aspirations and Educational Attainment for Girls: A Policy Experiment in India," *Science*, 2012, *335* (6068), 582–586.

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

Bertrand, Marianne and Esther Duflo, "Field Experiments on Discrimination," in "Handbook of Economic Field Experiments" 2017, pp. 309–393.

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

Besley, Timothy, Olle Folke, Torsten Persson, and Johanna Rickne, "Gender Quotas and the Crisis of the Mediocre Man: Theory and Evidence from Sweden," *American Economic Review*, 2017, *107* (8), 2204–42.

Bhalotra, Sonia, Irma Clots-Figueras, and Lakshmi Iyer, "Pathbreakers? Women's Electoral Success and Future Political Participation," *Economic Journal*, 2017.

Bohren, J Aislinn, Kareem Haggag, Alex Imas, and Devin G Pope, "Inaccurate Statistical Discrimination," Working Paper, NBER 2019.

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

Casas-Arce, Pablo and Albert Saiz, "Women and Power: Unpopular, Unwilling, or Held Back?," *Journal of Political Economy*, 2015, *123* (3), 641–669.

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

Center for American Women and Politics, "Women in the US Congress 2016," 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, "Are Female Leaders Good for Education? Evidence from India," *American Economic Journal: Applied Economics*, 2012, *4* (1), 212–244.

Cronbach, Lee J, "Coefficient Alpha and the Internal Structure of Tests," *Psychometrika*, 1951, *16* (3), 297–334.

Doleac, Jennifer L and Luke CD Stein, "The Visible Hand: Race and Online Market Outcomes," *Economic Journal*, 2013, *123* (572), F469–F492.

Esteve-Volart, Berta and Manuel Bagues, "Are Women Pawns in the Political Game? Evidence from Elections to the Spanish Senate," *Journal of Public Economics*, 2012, *96* (3), 387–399.

Ferreira, Fernando and Joseph Gyourko, "Does Gender Matter for Political Leadership? The Case of US Mayors," *Journal of Public Economics*, 2014, *112*, 24–39.

Fraga, Bernard L, "Race, Party, and Candidate Prospects Across the Multiple Stages of Congressional Elections," 2013.

Fryer, Roland G and Steven D Levitt, "The Causes and Consequences of Distinctively Black Names," *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.

Glover, Dylan, Amanda Pallais, and William Pariente, "Discrimination as a Self-Fulfilling Prophecy: Evidence from French Grocery Stores," *Quarterly Journal of Economics*, 2017, *132* (3), 1219–1260.

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

Green, Donald, Bradley Palmquist, and Eric Schickler, *Partisan Hearts & Minds: Political Parties and the Social Identities of Voters*, Yale University Press, 2002.

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

Hedegaard, Morten and Jean-Robert Tyran, "The Price of Prejudice," *American Economic Journal: Applied Economics*, 2018, *10* (1).

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.

Kuziemko, Ilyana and Ebonya Washington, "Why Did the Democrats Lose the South? Bringing New Data to an Old Debate," *American Economic Review*, 2018, *108* (10), 2830–67.

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

List, John A, "The Nature and Extent of Discrimination in the Marketplace: Evidence from the Field," *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.

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," *American Economic Review*, 2003, *93* (4), 1132–1151.

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

Pons, Vincent and Clémence Tricaud, "Expressive Voting and Its Cost: Evidence from Runoffs with Two or Three Candidates," *Econometrica*, 2018, *86* (5).

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

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

Rubinstein, Yona and Dror Brenner, "Pride and Prejudice: Using Ethnic-Sounding Names and Inter-Ethnic Marriages to Identify Labour Market Discrimination," *Review of Economic Studies*, 2014, *81* (1), 389–425.

Schwarz, Susanne and Alexander Coppock, "What Have We Learned About Gender From Candidate Choice Experiments? A Meta-Analysis of 42 Factorial Survey Experiments," 2019.

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

Spenkuch, Jörg L, "Expressive vs. Strategic Voters: An Empirical Assessment," *Journal of Public Economics*, 2018, *165*, 73–81.

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.

Teele, Dawn Langan, Joshua Kalla, and Frances Rosenbluth, "The Ties That Double Bind: Social Roles and Women's Underrepresentation in Politics," *American Political Science Review*, 2018, *112* (3), 525–541.

Trebbi, Francesco, Philippe Aghion, and Alberto Alesina, "Electoral Rules and Minority Representation in US Cities," *Quarterly Journal of Economics*, 2008, *123* (1), 325–357.

Vogl, Tom S, "Race and the Politics of Close Elections," *Journal of Public Economics*, 2014, *109*, 101–113.

Washington, Ebonya, "How Black Candidates Affect Voter Turnout," *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 Additional Tables and Figures

Table A1: Detailed Office by Race of Delegate Candidate

	Count of Delegate Candidates				
	By Race		By Gender		
	White	Nonwhite	Male	Female	Total
Board of Education	14	0	11	3	14
City Council	64	5	59	10	69
County Office	110	3	87	26	113
Court Clerk	4	0	3	1	4
Governor	3	0	3	0	3
Judge	8	1	9	0	9
Local Official	81	4	55	30	85
Mayor	45	2	45	2	47
Other Notable	10	0	8	2	10
Party Office	251	19	187	83	270
Past Candidate	67	17	55	29	84
Sheriff	13	2	15	0	15
State House	109	4	84	29	113
State Senate	57	7	47	17	64
State's Attorney	13	1	12	2	14
Statewide Office	17	3	15	5	20
US House	15	4	16	3	19
No Office	1,257	108	963	402	1,365
Total	2,184	180	1,674	644	2,318

Notes: This table reports the distribution of delegate candidates, split by race, among detailed public-office categories. As in Table 2, which provides a higher-level summary of the officeholding data, we dichotomize both the MTurk measure of race and the SSA measure of gender at 0.5. 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. The $\chi^2(4)$ statistics by race and gender are respectively 37.8 ($p = 0.003$) and 41.7 ($p = 0.001$).

Table A2: 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 elections. 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 A3: 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 A4: Turnout in Presidential and Non-Presidential Primary Elections in Illinois

	Primary Turnout as a % of Voting-Eligible Pop.			
	(1)	(2)	(3)	(4)
Presidential Year	0.066**	0.044	0.068**	0.061**
	(0.025)	(0.026)	(0.025)	(0.022)
Senate Year	0.031	0.020	0.033	0.025
	(0.026)	(0.027)	(0.025)	(0.022)
Lagged Turnout		-0.140		
		(0.242)		
Year			-0.001	
			(0.001)	
HP Filter				X
N	18	17	18	18
R^2	0.350	0.312	0.366	0.368

Notes: This table analyzes the relationship between primary voter turnout as a share of the voting-eligible population and the offices up for election in Illinois from 1980 to 2016. Mean turnout was 24.9 percent. Data on turnout are from the Official Canvass of the Illinois State Board of Elections and voting-eligible population from the U.S. Elections Project. Illinois runs gubernatorial and presidential elections in opposite years. In Column 4, we apply a Hodrick–Prescott filter to the data following Ravn and Uhlig (2002). All specifications suggest that primary turnout in presidential years is about 4–7 p.p. higher than in non-presidential (i.e., gubernatorial) years.

Table A5: Inclusion or Exclusion of Inexact Census Matches Does Not Affect Results

	Exact Matches Only		All Matches	
	(1)	(2)	(3)	(4)
Female	0.004	0.001	0.004	0.001
	(0.006)	(0.006)	(0.006)	(0.006)
Nonwhite	-0.045***		-0.043***	
	(0.014)		(0.013)	
Black		0.018		-0.005
		(0.025)		(0.024)
Hispanic/Latino		-0.061***		-0.058***
		(0.018)		(0.016)
Asian		-0.069***		-0.076***
		(0.014)		(0.014)
N	16,945	11,166	18,932	11,649
R^2	0.991	0.992	0.991	0.992

Notes: This table estimates Equation 1 using the Census race measure. In Columns 1 and 2, we restrict the sample to delegates whose last names can be matched exactly in the 2000 Census tabulation, reproducing our results in Table 4 for ease of comparison. In Columns 3 and 4, we allowing for matches to similar last names according to the Jaro–Winkler distance. In all regressions, the dependent variable is the vote count, and the unit of observation is the county–district-delegate-year. All regressions include cell-level FEs. Delegates are defined as in the same cell if they 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. Standard errors are clustered at the delegate level. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

Table A6: Most Racially Informative 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 are most informative 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 A7: 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/Latino	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 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
	Continuous	Age (years)	53.9	18.9	18	86	19,223

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). Dichotomized measures of race, ethnicity, and gender are in this table at the 50-percent probability threshold. See Section 3 for discussion of these measures.

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: Effect of Delegate Candidate Race and Gender on Vote Share, Weighted

	MTurk		Census		Onolytics		Rescaled PC1	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	0.005	0.006	0.004	0.004	0.005	0.004	0.005	0.001
	(0.006)	(0.006)	(0.007)	(0.007)	(0.006)	(0.006)	(0.007)	(0.007)
Nonwhite	-0.079***		-0.041**		-0.033***		-0.085***	
	(0.017)		(0.016)		(0.008)		(0.019)	
Black		-0.029		0.017		-0.038**		-0.085**
		(0.054)		(0.031)		(0.017)		(0.041)
Hispanic/Latino		-0.052**		-0.057**		-0.042***		-0.073***
		(0.020)		(0.022)		(0.013)		(0.023)
Any Asian				-0.065***				-0.086***
				(0.018)				(0.041)
East Asian		-0.085***				-0.052***		
		(0.027)				(0.009)		
Indian		-0.157***				-0.071***		
		(0.042)				(0.014)		
Middle Eastern		-0.144***				-0.079**		
		(0.033)				(0.035)		
White Ethnic						-0.015		
						(0.010)		
N	19,220	19,220	17,410	12,732	18,928	18,928	17,178	12,647
R^2	0.545	0.548	0.561	0.702	0.540	0.542	0.564	0.702

Notes: This table presents the results of estimating an OLS variant of Equation 1 in which the dependent variable is the delegate's share of votes of the maximum vote-getter within his or her cell. "Any Asian" uses the Census definition of Asian race, which spans our subsequent categories of East Asian, Indian, and Middle Eastern. The unit of observation is the county–district-delegate-year. All regressions include cell-level FEs and are weighted by the maximum number of votes won by a delegate candidate in the cell. Failing to weight the sample would dramatically overweight the delegates of presidential candidates who receive very few votes. With this weighting, the vote share specification estimates the same quantity of interest as the Poisson, the average share of votes nonwhite or female delegates lose because voters discriminate. Delegates are defined as in the same cell if they 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. Standard errors are clustered at the delegate level. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

Table A11: Effect of Delegate Candidate Race and Gender on Log Votes, Weighted

	MTurk		Census		Onolytics		Rescaled PC1	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	0.002	0.003	-0.001	0.001	0.001	0.001	0.000	0.000
	(0.010)	(0.010)	(0.012)	(0.009)	(0.010)	(0.010)	(0.012)	(0.010)
Nonwhite	-0.091***		-0.059**		-0.038***		-0.108***	
	(0.023)		(0.025)		(0.011)		(0.026)	
Black		-0.059		0.016		-0.040**		-0.133**
		(0.076)		(0.041)		(0.018)		(0.054)
Hispanic/Latino		-0.056**		-0.065***		-0.047***		-0.090***
		(0.027)		(0.028)		(0.015)		(0.026)
Any Asian				-0.074***				-0.119***
				(0.021)				(0.032)
East Asian		-0.092***				-0.055***		
		(0.032)				(0.010)		
Indian		-0.181***				-0.077***		
		(0.055)				(0.016)		
Middle Eastern		-0.160***				-0.085*		
		(0.042)				(0.049)		
White Ethnic						-0.021		
						(0.013)		
N	19,220	19,220	17,410	12,732	18,928	18,928	17,178	12,647
R^2	0.987	0.987	0.987	0.990	0.987	0.987	0.987	0.990

Notes: This table presents the results of estimating an OLS variant of Equation 1 in which the dependent variable is the logarithm of the delegate's cell vote count plus one. It is a reproduction of Table 4 in previous drafts of the paper, which used this dependent variable. "Any Asian" uses the Census definition of Asian race, which spans our subsequent categories of East Asian, Indian, and Middle Eastern. The unit of observation is the county–district-delegate-year. We increment all vote totals by one before taking the natural logarithm because 42 of the 19,711 observations (0.2%) have zero votes. All regressions include cell-level FEs and are weighted by the maximum number of votes won by a delegate candidate in the cell. Due to the logarithmic transformation of the dependent variable, failing to weight the sample would dramatically overweight the delegates of presidential candidates who receive very few votes. With this weighting, the log-count specification estimates the same quantity of interest as the Poisson, the average share of votes nonwhite or female delegates lose because voters discriminate. For example, suppose that, in a given cell, a white delegate received a total of 1,000 votes, and a nonwhite delegate in the same cell received a total of 850. From these two data points, we would estimate a penalty of 16.2 percent of votes due to racial discrimination, as $\ln(850+1) - \ln(1,000+1) = -0.162$. To preempt any concerns that weights and incrementation affect our results, we have switched our main specification to the Poisson regression in this version of the paper. To maximize transparency and demonstrate the robustness of the results, this table shows the previous results using this alternative specification are similar. Delegates are defined as in the same cell if they 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. Standard errors are clustered at the delegate level. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

Table A12: Delegate Race and Gender Is Nearly Uncorrelated with Ballot Order

	Correlation Coefficient	
	With Race	With Gender
Ballot Order	-0.0266	0.0258

Notes: This table reports correlation coefficients of the PC1 race measure with ballot order rank. The ballot order correlation we report is a partial correlation conditional on the number of total ballot slots available in the district. The maximum number of votes that voters may cast varies from 2 to 4 across districts, always equaling the number of delegates elected in that district, and is determined ex ante by party rules. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

Table A13: Effect of Delegate Candidate Race and Gender on Votes, Separate Regressions for Gender and Race

	(1)	(2)	(3)
Female	0.004		
	(0.006)		
Nonwhite		-0.088***	
		(0.016)	
Black			-0.101***
			(0.031)
Hispanic/Latino			-0.079***
			(0.018)
Asian			-0.108***
			(0.021)
N	18,958	17,126	11,528
Pseudo-R^2	0.991	0.990	0.991

Notes: This table presents the results of estimating Equation 1, yielding estimates of the percentage vote penalties by nonwhite race or ethnicity and by gender. By comparison to Table 4, we estimate race and gender effects in separate regressions. Our coefficients differ slightly because the regressions reported in Table 4 omit observations for which either race or gender are missing, whereas the regressions reported here only omit observations for which the relevant independent variable is missing. In all regressions the dependent variable is the vote count for the delegate. We use the PC1 measures of race, as in Columns 7 and 8 of Table 4. The unit of observation is the county–district-delegate-year. All regressions include cell-level FEs. Delegates are defined as in the same cell if they 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. Sample sizes change because not all delegate names can be classified using the Census data or Onolytics algorithm, leaving the PC1 index missing. Standard errors are clustered at the delegate level. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

Table A14: Effect of Delegate Candidate Race and Gender on Votes, Alternative Clusters

	MTurk		Census		Onolytics		Rescaled PC1	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	0.006	0.006	0.004	0.001	0.005	0.004	0.005	0.001
	(0.007)	(0.007)	(0.008)	(0.007)	(0.007)	(0.007)	(0.008)	(0.007)
Nonwhite	-0.087***		-0.045**		-0.035***		-0.092***	
	(0.019)		(0.017)		(0.008)		(0.021)	
Black		-0.033		0.018		-0.040***		-0.094**
		(0.063)		(0.032)		(0.014)		(0.042)
Hispanic/Latino		-0.057***		-0.061**		-0.045***		-0.079***
		(0.022)		(0.024)		(0.015)		(0.024)
Any Asian				-0.069***				-0.110***
				(0.015)				(0.025)
East Asian		-0.093***				-0.055***		
		(0.019)				(0.006)		
Indian		-0.174***				-0.076***		
		(0.048)				(0.017)		
Middle Eastern		-0.160***				-0.088**		
		(0.036)				(0.036)		
White Ethnic						-0.016*		
						(0.009)		
N	18,958	18,958	16,945	11,166	18,639	18,639	16,668	11,049
Pseudo-R^2	0.991	0.991	0.991	0.992	0.991	0.991	0.991	0.992

Notes: This table presents the results of estimating Equation 1, but with the standard errors clustered at the level of the "slate" of delegate or alternate candidates—which is constant at the year, congressional district, presidential candidate level—instead of the delegate level. We also find similar standard errors if we define clusters by congressional district, year, and alternate status. "Any Asian" uses the Census definition of Asian race, which spans our subsequent categories of East Asian, Indian, and Middle Eastern. In all regressions, the dependent variable is the vote count, and the unit of observation is the county–district-delegate-year. All regressions include cell-level FEs. Delegates are defined as in the same cell if they 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. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

Table A15: Interaction Effects Between Delegate Race, Delegate Gender, and Presidential-Candidate Race

	(1)	(2)	(3)	(4)
Female Deleg.	-0.000	0.015**	-0.000	0.018***
	(0.008)	(0.006)	(0.009)	(0.007)
Nonwhite Deleg.	-0.105***	-0.116***	-0.114***	-0.124***
	(0.019)	(0.017)	(0.018)	(0.017)
Nonwhite Deleg. × Female Deleg.	0.040	0.024	0.034	0.019
	(0.031)	(0.027)	(0.034)	(0.028)
Female Deleg. × Female Pres. Cand.			0.318***	0.290***
			(0.110)	(0.107)
Female Deleg. × Nonwhite Pres. Cand.			-0.011	-0.029***
			(0.015)	(0.011)
Nonwhite Deleg. × Female Pres. Cand.			-0.360*	-0.357*
			(0.206)	(0.203)
Nonwhite Deleg. × Nonwhite Pres. Cand.			0.159**	0.144**
			(0.074)	(0.065)
Controls		✓		✓
N	16,668	16,668	16,668	16,668
Pseudo-R^2	0.991	0.992	0.991	0.992

Notes: This tables presents the results of estimating Equation 1, modified by interacting the PC1 race measure with the SSA gender measure. Controls, when included, are for detailed delegate officeholding and ballot order. In all regressions, the dependent variable is the vote count, and the unit of observation is the county–district-delegate-year. All regressions include cell-level FEs. Delegates are defined as in the same cell if they 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. Standard errors are clustered at the delegate level. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

Table A16: Results Robust To Imputation of Missing Values

	(1)	(2)
Female	0.005	0.005
	(0.006)	(0.006)
Nonwhite	-0.095***	
	(0.015)	
Black		-0.057
		(0.038)
Hispanic/Latino		-0.051***
		(0.017)
Asian		-0.162***
		(0.021)
N	19,460	19,460
Pseudo-R^2	0.990	0.990

Notes: This table reports the results of estimating Equation 1, imputing missing values of the PC1 race measure and the SSA gender measure. For race, approximately 10.7 percent of values were missing and were predicted from the bivariate relationship between the PC1 and MTurk race measures. For gender, about 2.5 percent of values were missing and were replaced using the sample mean and an indicator for missing. In all regressions, the dependent variable is the vote count, and the unit of observation is the county–district-delegate-year. All regressions include cell-level FEs. Delegates are defined as in the same cell if they 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. Standard errors are clustered at the delegate level. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

Table A17: Effect of Delegate Candidate Race and Gender, Dichotomized, on Votes

	MTurk				Census			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Female	0.007	0.008	0.007	0.008	0.007	0.003	0.006	0.005
	(0.005)	(0.006)	(0.005)	(0.006)	(0.006)	(0.007)	(0.006)	(0.007)
Nonwhite	-0.086***	-0.072***			-0.055***	-0.046***		
	(0.015)	(0.011)			(0.009)	(0.008)		
Black			0.128***	-0.123***			-0.105***	-0.103***
			(0.043)	(0.020)			(0.031)	(0.021)
Hispanic/Latino			-0.041***	-0.037***			-0.048***	-0.044***
			(0.011)	(0.014)			(0.011)	(0.013)
Any Asian							-0.074***	-0.053***
							(0.018)	(0.011)
East Asian			-0.064***	-0.053***				
			(0.018)	(0.016)				
Indian			-0.129***	-0.095**				
			(0.026)	(0.048)				
			(0.026)	(0.043)				
Subsample		✓		✓		✓		✓
N	19,460	15,685	19,460	16,688	17,443	12,955	17,443	14,538
Pseudo-R^2	0.990	0.991	0.990	0.991	0.990	0.991	0.990	0.991

Notes: This table presents the results of estimating Equation 1, but with two dichotomizations of the race and gender variables. In odd columns, each indicator equals 1 if the probability of group membership exceeds 0.5 and equals 0 otherwise. In even columns, each indicator equals 0 if the probability is between 0 and 0.25, 1 if the probability is between 0.75 and 1, and is missing otherwise. Even columns thus use the subsample of delegates for which voters can be reasonably certain of their race or gender. The "black" coefficient is positive in Column 3 because members of the prominent LaHood family (former U.S. Secretary of Transportation Ray LaHood and Congressman Darin LaHood) were perceived as black by about half of MTurkers and both ran for delegate in nearly every election. Removing members of the LaHood family from the regression in Table 3 restores the significant negative coefficient, as does controlling for officeholding (see Table 5). "Any Asian" uses the Census definition of Asian race, which spans our subsequent categories of East Asian, Indian, and Middle Eastern. In all regressions, the dependent variable is the vote count, and the unit of observation is the county–district-delegate-year. All regressions include cell-level FEs. Delegates are defined as in the same cell if they 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. Standard errors are clustered at the delegate level. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

Table A18: Results of Officeholding and Ballot Order Controls

	Office Level (1)	Detailed Office (2)	Ballot Order (3)	Detailed Ballot Order (4)	Office Level & Det. Ballot Order (5)	All Controls (6)
Major Office	0.360***				0.323***	
	(0.036)				(0.038)	
State Legislature	0.096***				0.078***	
	(0.012)				(0.011)	
Minor Office	0.018***				0.016***	
	(0.005)				(0.005)	
Board of Education		0.041**				0.047***
		(0.018)				(0.017)
City Council		0.006				0.012
		(0.010)				(0.010)
County Office		0.025***				0.022***
		(0.009)				(0.008)
Court Clerk		0.106***				0.145***
		(0.032)				(0.027)
Governor		0.294***				0.256***
		(0.090)				(0.085)
Judge		0.164**				0.131*
		(0.068)				(0.068)
Local Official		0.025**				0.007
		(0.010)				(0.009)
Mayor		0.023				0.027**
		(0.014)				(0.012)
Other Notable		0.032**				0.016
		(0.016)				(0.015)
Party Office		0.002				0.001
		(0.006)				(0.006)
Past Candidate for Other Office		0.017				0.014
		(0.011)				(0.010)
Sheriff		0.136***				0.120***
		(0.023)				(0.024)
State House		0.091***				0.079***
		(0.011)				(0.011)
State Senate		0.118***				0.093***
		(0.018)				(0.016)
State's Attorney		0.090**				0.108***
		(0.039)				(0.032)
Statewide Office		0.088***				0.087***
		(0.029)				(0.029)
US House		0.387***				0.355***
		(0.036)				(0.038)
Ballot Order = 2			-0.040***	-0.079***		
			(0.005)	(0.012)		
Ballot Order = 3			-0.074***	-0.116***		
			(0.006)	(0.014)		
Ballot Order = 4			-0.109***	-0.146***		
			(0.009)	(0.015)		
Ballot Order = 1 × No. of Votes = 3				-0.049***	0.059***	0.059***
				(0.016)	(0.005)	(0.005)
Ballot Order = 1 × No. of Votes = 4				-0.030*	0.056***	0.052***
				(0.016)	(0.010)	(0.010)
Ballot Order = 2 × No. of Votes = 3				-0.003	0.036***	0.036***
				(0.010)	(0.005)	(0.005)
N	16,668	16,668	16,668	16,668	16,668	16,668
Pseudo-R^2	0.992	0.992	0.991	0.991	0.992	0.992

Notes: This table reports the results of estimating Equation 1 with controls for delegate-candidate officeholding and ballot order, displaying coefficients associated with these controls. Table 5 presents the coefficients from the same regressions for race, ethnicity, and gender. In all columns the dependent variable is the vote count, and the unit of observation is the county–district-delegate-year. We do not report coefficients on nonwhite and female variables in the above table but include them 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. All regressions include FEs at the cell level. Delegates are defined as in the same cell if they 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. Standard errors are clustered at the delegate level. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

Table A19: Do Voters Discriminate on Other Delegate Attributes?

	(1)	(2)	(3)	(4)
Female	0.019***	0.018***	0.017***	0.018***
	(0.005)	(0.005)	(0.005)	(0.005)
Nonwhite	-0.107***	-0.108***	-0.109***	-0.108***
	(0.014)	(0.014)	(0.014)	(0.014)
Age	0.000			0.000
	(0.000)			(0.000)
% College		0.008		-0.010
		(0.016)		(0.023)
Household Income			0.000	0.002
			(0.006)	(0.008)
N	16,668	16,668	15,316	15,316
Pseudo-R^2	0.992	0.992	0.992	0.992

Notes: This table investigates discrimination on three other delegate attributes: age, education, and income. Delegate age is measured in years using the year in which the delegates' first name attained its maximal share of births. The college share and log median household income refer to characteristics of the block group of the delegate's official residential address as reported in election filings. The unit of observation is the county–district-delegate-year. All regressions include cell-level FEs and controls for delegate office level and ballot order. Delegates are defined as in the same cell if they 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. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

Table A20: Heterogeneous Treatment Effects by Delegate Ballot Order

	(1)
Nonwhite Delegate	
× Ballot Order = 1	-0.097***
	(0.025)
× Ballot Order = 2	-0.114***
	(0.015)
× Ballot Order = 3	-0.119***
	(0.024)
× Ballot Order = 4	-0.096
	(0.078)
Female Delegate	
× Ballot Order = 1	0.021**
	(0.009)
× Ballot Order = 2	0.020**
	(0.008)
× Ballot Order = 3	0.020**
	(0.009)
× Ballot Order = 4	0.031**
	(0.014)
Ballot Order × No. of Votes Fixed Effects	✓
N	16,668
Pseudo-R^2	0.992

Notes: This table reports the results of estimating Equation 1, interacting with the delegate's order on the ballot. The regression includes fixed effects, unreported, for the interactions of ballot order and the number of delegate votes voters can cast. There are few contests in which voters could cast four delegate votes, making the estimate for the Ballot Order = 4 coefficients less precise. In all regressions, the dependent variable is the vote count, and the unit of observation is the county–district-delegate-year. All regressions include cell-level FEs and controls for ballot order and detailed officeholding. Delegates are defined as in the same cell if they 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. Standard errors are clustered at the delegate level. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

Table A21: Heterogeneous Treatment Effects by Other Presidential Candidate Attributes

	Statewide Votes (1)	Ideology (2)
Female × Below-Median Votes	-0.021	
	(0.057)	
Female × Above-Median Votes	0.020***	
	(0.004)	
Nonwhite × Below-Median Votes	-0.572***	
	(0.145)	
Nonwhite × Above-Median Votes	-0.093***	
	(0.013)	
Female × Left of Median CFScore		0.024***
		(0.006)
Female × Right of Median CFScore		0.009
		(0.006)
Nonwhite × Left of Median CFScore		-0.108***
		(0.019)
Nonwhite × Right of Median CFScore		-0.105***
		(0.019)
N	16,668	16,668
Pseudo-R^2	0.992	0.992

Notes: This table reports the results of estimating Equation 1, interacting with attributes of presidential candidates. The Nonwhite × Below-Median Votes and Nonwhite × Above-Median Votes interaction coefficients do not average to the coefficient in Table 4 because, naturally, delegates of presidential candidates with an above-median number of votes receive many more votes than delegates presidential candidates with a below-median number of votes. The reason that the coefficient for Nonwhite × above-median Votes is similar to the overall effect in Table 4 is because, given the Poisson specification, the estimates in Table 4 mostly reflect the behavior of voters for candidates who receive an above-median number of votes. Discrimination against nonwhites also persists even among the top two presidential candidates' voters in each year ($\beta = -0.087$, $p < 0.001$) as well as for the top presidential candidates' voters ($\beta = -0.083$, $p < 0.001$). In all regressions, the dependent variable is the vote count, and the unit of observation is the county–district-delegate-year. All regressions include cell-level FEs and controls for ballot order and detailed officeholding. Delegates are defined as in the same cell if they 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. Standard errors are clustered at the delegate level. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

Table A22: Heterogeneous Treatment Effects by Regular or Alternate Delegate Status

	(1)	(2) χ^2 test of equality
Nonwhite × Regular	-0.088***	0.07
	(0.028)	
Nonwhite × Alternate	-0.096***	
	(0.014)	
Female × Regular	-0.011	8.52***
	(0.011)	
Female × Alternate	0.024***	
	(0.005)	
N	16,668	
Pseudo-R^2	0.991	

Notes: This tables presents the results of estimating Equation 1, modified by interacting the PC1 race measure and gender measure with an indicator for whether the delegate is a regular or an alternate. The entire table reports results from one regression. Column 2 reports statistics for χ^2 tests of equality of coefficients between regular and alternate delegates, by race and gender respectively. The results show that taste-based racial discrimination is statistically indistinguishable between regular and alternate delegates, but that taste-based discrimination in favor of women appears weakly concentrated in alternate delegates. The statistic on the joint test is 1.29, not significant at conventional thresholds. In all regressions, the dependent variable is the vote count, and the unit of observation is the county–district-delegate-year. 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. Delegates are defined as in the same cell if they 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. Standard errors are clustered at the delegate level. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

Table A23: Heterogeneous Treatment Effects for Nonwhite Delegates
by Geographic Attributes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: No Delegate FEs							
Nonwhite	-0.098***	-0.099***	-0.105***	-0.103***	-0.099***	-0.098***	-0.101***
	(0.019)	(0.017)	(0.016)	(0.017)	(0.018)	(0.019)	(0.016)
× % White	0.003			0.006			
	(0.020)			(0.020)			
× % College		0.017*		0.005			
		(0.010)		(0.011)			
× Income			0.029**	0.028**			
			(0.014)	(0.013)			
× Animus					0.003		
					(0.010)		
× IAT Score						0.003	
						(0.011)	
× % Republican							-0.026**
							(0.010)
SD of Covariate	0.399	0.125	0.281		13.852	0.029	0.136
N	17,126	17,126	17,126	17,126	17,126	13,657	17,126
Pseudo-R^2	0.991	0.991	0.991	0.991	0.991	0.991	0.991
Panel B: Delegate FEs							
Nonwhite × % White	0.028			0.080			
	(0.042)			(0.050)			
Nonwhite × % College		0.031**		0.032*			
		(0.015)		(0.018)			
Nonwhite × Income			0.031*	0.010			
			(0.017)	(0.021)			
Nonwhite × Animus					-0.015*		
					(0.008)		
Nonwhite × IAT Score						0.005	
						(0.007)	
Nonwhite × % Republican							-0.024**
							(0.010)
SD of Covariate	0.404	0.125	0.273		14.266	0.030	0.126
N	16,812	16,812	16,812	16,812	16,812	13,343	16,812
Pseudo-R^2	0.994	0.994	0.994	0.994	0.994	0.994	0.994

Notes: In all regressions above, the dependent variable is the vote count, the race measure used is the rescaled PC1 measure, and the unit of observation is the county–district-delegate-year. Income is defined as the mean per-capita income of non-Hispanic whites. All covariates are standardized to zero mean and unit standard deviation. In Panel A, all regressions include controls for ballot order, delegate officeholding, and cell-level FEs. In Panel B, all regressions also include delegate-level FEs. As we observe vote counts across counties within delegates, adding delegate FEs allows electorate demography to vary while holding delegate identities constant. Delegates are defined as in the same cell if they 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. Standard errors are two-way clustered at the delegate and county due to the county-level interaction terms. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

Table A24: Heterogeneous Treatment Effects for Female Delegates
by Geographic Attributes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: No Delegate FEs							
Female	0.016***	0.016***	0.016***	0.016***	0.016***	0.016***	0.016***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
× Sex Ratio	0.002			0.004			
	(0.003)			(0.003)			
× % College		0.010***		0.007*			
		(0.004)		(0.004)			
× Income			0.009**	0.006			
			(0.004)	(0.004)			
× Income Gap					0.000		
					(0.003)		
× College Gap						0.007**	
						(0.003)	
× % Republican							-0.004
							(0.004)
SD of Covariate	0.067	0.124	0.282		0.141	0.036	0.138
N	18,958	18,958	18,958	18,958	18,958	18,958	18,958
Pseudo-R^2	0.992	0.992	0.992	0.992	0.992	0.992	0.992
Panel B: Delegate FEs							
Female × Sex Ratio	0.002			0.002			
	(0.003)			(0.003)			
Female × % College		0.005		0.007			
		(0.005)		(0.004)			
Female × Income			0.002	-0.002			
			(0.008)	(0.008)			
Female × Income Gap					0.003		
					(0.004)		
Female × College Gap						0.001	
						(0.004)	
Female × % Republican							-0.004
							(0.004)
SD of Covariate	0.067	0.124	0.272		0.135	0.037	0.138
N	18,563	18,563	18,563	18,563	18,563	18,563	18,563
Pseudo-R^2	0.994	0.994	0.994	0.994	0.994	0.994	0.994

Notes: In all regressions above, the dependent variable is the vote count, and the unit of observation is the county–district-delegate-year. The sex ratio is defined as the number of adult men divided by the number of adult women. Income is defined as the mean per-capita income of non-Hispanic whites. The income gap is defined as the log male-female difference in annual labor earnings. The college gap is defined as the difference in share of men less of women with a bachelor's degree or more. All covariates are standardized to zero mean and unit standard deviation. In Panel A, all regressions include controls for ballot order, detailed officeholding, and cell-level FEs. In Panel B, all regressions also include delegate-level FEs. Delegates are defined as in the same cell if they 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. Standard errors are two-way clustered at the delegate and county due to the county-level interaction terms. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

Table A25: Voter Discrimination Over Time

	(1) Nonwhite	(2) Female
2000	-0.075	0.011
	(0.060)	(0.015)
2008	-0.103**	0.000
	(0.048)	(0.013)
2012	-0.091***	0.021**
	(0.021)	(0.009)
2016	-0.095***	-0.003
	(0.026)	(0.012)
χ^2 statistic (equality across years)	3.51	0.14
N	16,668	
R^2	0.991	
Panel B: Linear Time Trend		
Level in 2000	-0.080	0.011
	(0.049)	(0.013)
Annual Trend	-0.001	-0.001
	(0.004)	(0.001)
N	16,668	
R^2	0.991	

Notes: This tables presents the results of estimating Equation 1, modified by interacting the PC1 race measure with year dummies. In all regressions, the dependent variable is the vote count, and the unit of observation is the county–district-delegate-year. All regressions include cell-level FEs. Delegates are defined as in the same cell if they 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. Standard errors are clustered at the delegate level. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

Table A26: Simulated Effect of Eliminating Racial Discrimination of 9 p.p. in U.S. House Primary Elections, 1990–2010

	All Primaries		Open Primaries in Favored Seats	
	Rep.	Dem.	Rep.	Dem.
# with Nonwhite <9 p.p. from White Winner	19	14	6	6
# Won by Nonwhites	218	372	9	21
# with Nonwhite Candidates	378	540	35	50
# of Close Primaries (<9 p.p. Margin)	160	130	32	24
# of Primaries (Total)	4,300	4,302	239	188
Increase in Nonwhite Winners (%)	8.7%	3.8%	66.7%	28.6%
Increase in Nonwhite Win Rate (p.p.)	5.6%	2.8%	20.0%	14.0%
% of Close Primaries Changed	11.9%	12.0%	18.75%	25.0%
% of All Primaries Changed	0.5%	0.3%	2.9%	3.7%

Notes: Rows 1–5 show numbers of U.S. House primaries from 1990 to 2010 with specified characteristics. In particular, Row 1 is the number of primaries for which, given our point estimate, taste-based racial discrimination changes the race of the winner. To contextualize this change, Rows 6–9 divide Row 1 by Rows 2–5 respectively. Rows 8 and 9 refer to primaries in which the race of the winner is changed. Data are from Pettigrew et al. (2014). Candidate race is coded using Census data as in Section 3.2.

Table A27: Example Delegate Selection Outcome:
Top Six Delegates by Vote, Illinois 6th Congressional District 2016

Delegate Name	Presidential Candidate	# Votes	Won Delegate Election?
Paul Minch	Trump	37,150	Won
Barbara Kois	Trump	36,838	Won
Patrick Brady	Kasich	34,072	Won
Ronald Sandack	Kasich	33,538	Lost
Aaron Del Mar	Kasich	32,228	Lost
Nabi Fakroddin	Trump	32,136	Lost

Notes: This Table provides an example of a delegate election outcome. It shows the top six vote-getting delegates in Illinois' 6th Congressional district in the 2016 Republican Presidential primary. In this contest, the top three vote-getting delegates won and served. Two of the top vote-getting delegates were bound to Donald Trump, but the third winning delegate was bound to John Kasich. The third Trump delegate, Nabi Fakroddin, received approximately 5,000 fewer votes than the other Trump delegates, or about 13 percent fewer. This lower vote total led Fakroddin's vote total to be below the vote totals of all three Kasich delegates. As a result, this district sent two Trump delegates and one Kasich delegate to the Republican National Convention. 93% of MTurkers we showed Fakroddin's name guessed Fakroddin was not white. Most MTurkers perceived his name as Middle Eastern; per Onolytics, the name is in fact Middle Eastern in origin. Note that, per Table 4, our findings are not limited to Middle Eastern names.

Table A28: Selection into the Identifying Set

	(1)	(2)	(3)	(4)	(5)
Panel A: Nonwhite Delegate					
White Share	0.015				0.030*
	(0.021)				(0.017)
College Share		0.033*			0.063**
		(0.018)			(0.031)
Log Per-Capita Income			0.027***		-0.022
			(0.009)		(0.024)
Rep. 2-Party Vote Share				-0.141***	-0.157***
				(0.039)	(0.055)
N	6,175	6,175	6,175	6,175	6,175
R^2	0.013	0.014	0.017	0.043	0.045
Panel B: Female Delegate					
White Share	0.043				0.010
	(0.038)				(0.046)
College Share		-0.150***			-0.056
		(0.042)			(0.088)
Log Per-Capita Income			-0.047*		-0.071*
			(0.027)		(0.041)
Rep. 2-Party Vote Share				-0.108**	-0.180***
				(0.051)	(0.044)
N	6,208	6,208	6,208	6,208	6,208
R^2	0.026	0.030	0.028	0.029	0.037

Notes: This table characterizes how the share of delegates who are nonwhite or female varies with the demographic and political characteristics of the county–districts in which they run. In Panel A, the dependent variable is the cell mean of the PC1 race measure; in Panel B, it is the cell mean of the SSA female measure. Delegates are defined as in the same cell if they 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. The independent variables are the white share of population, the college-educated share of whites, white per-capita income, and the Republican presidential candidates' share of the two-party vote in each county in the presidential election held that year. The first three variables are observed at the county–district level, and the fourth at the county–year level. We measure presidential vote share at the county level because to our knowledge it is only available at the county and district levels, but not the county–district levels. The unit of observation is the cell. All regressions include year fixed effects. Observations are weighted by the maximum number of votes for a delegate candidate in the cell. Standard errors are clustered at the county–district level. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

Figure A1: Full Ballot in McLean County, Illinois

SAMPLE REPUBLICAN PRIMARY BALLOT

Kathy Michael

Kathy Michael, County Clerk

MARCH 15, 2016
050-TOWANDA 01
MCLEAN COUNTY, ILLINOIS

Judge's Initials ______

To vote, darken the oval to the LEFT of your choice, like this ⬬. To cast a write-in vote, darken the oval to the LEFT of the blank space provided and write the candidate's name in that space. For specific information, refer to the card of instruction posted in the voting booth. If you tear, spoil, deface or erroneously mark this ballot, return it to the election judge and obtain another.

FEDERAL

FOR PRESIDENT OF THE UNITED STATES
(Vote for one)

- JEB BUSH
- CHRIS CHRISTIE
- DONALD J. TRUMP
- TED CRUZ
- MARCO RUBIO
- RAND PAUL
- CARLY FIORINA
- MIKE HUCKABEE
- RICK SANTORUM
- JOHN R. KASICH
- BEN CARSON
- ______ Write-in

FOR UNITED STATES SENATOR
(Vote for one)

- JAMES T. MARTER
- MARK STEVEN KIRK

STATE

FOR COMPTROLLER
(For an unexpired two year term)
(Vote for one)

- LESLIE GEISSLER MUNGER

CONGRESSIONAL

FOR REPRESENTATIVE IN CONGRESS
EIGHTEENTH CONGRESSIONAL DISTRICT
(Vote for one)

- DARIN LaHOOD

CONGRESSIONAL

FOR DELEGATE TO THE NATIONAL NOMINATING CONVENTION EIGHTEENTH 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)

- ROBERT BROWNING (CHRISTIE)
- MARY K. BROOKHART (CHRISTIE)
- DONNA K. THOMPSON (CHRISTIE)
- JIM EDGAR (BUSH)
- BILL BRADY (BUSH)
- RAYMOND POE (BUSH)
- KENT GRAY (TRUMP)
- SANDRA YEH (TRUMP)
- WILLIAM GRAFF (TRUMP)
- H. LEE NEWCOM (CRUZ)
- MICHAEL FLYNN (CRUZ)
- CHRISTIAN H. GRAMM (CRUZ)
- KRISTINA RASMUSSEN (FIORINA)
- PHIL CHILES (FIORINA)
- CHUCK WEAVER (FIORINA)
- JIL TRACY (KASICH)
- RANDY E. FRESE (KASICH)
- ERIK M. WOEHRMANN (KASICH)
- JUDITH A. HANKS (CARSON)
- MATTHEW HOPPOCK (CARSON)
- STEVEN A. WAILAND (CARSON)
- ELIZABETH BLANKENSHIP (PAUL)
- MIKE BROOKS (PAUL)
- TIMOTHY TARVIN (PAUL)
- DARIN LaHOOD (RUBIO)
- JASON BARICKMAN (RUBIO)
- MICHAEL D. UNES (RUBIO)

TURN PAGE TO CONTINUE VOTING ☞

FRONT Card 174 RptPct 670-860 "050-TOWANDA 01"

CONGRESSIONAL

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

- DAVID L. BENDER (BUSH)
- WILLIAM L. KEMPINERS (BUSH)
- CONNIE NORD (BUSH)
- EDWARD HENDRICKS (CHRISTIE)
- DAVID R. HEPLER (CHRISTIE)
- NATHAN CHARLES BYRNE (CHRISTIE)
- DAVID L. HULLINGER (CRUZ)
- JAMES S. FISHER (CRUZ)
- BRIAN THIELEN (CRUZ)
- DIANE VESPA (TRUMP)
- MATTHEW MAU (TRUMP)
- BRIAN DENNEY (TRUMP)
- CHRISTINE ARRA (FIORINA)
- MICHAEL McGAUGHAN (FIORINA)
- JAMES KENNY JR. (FIORINA)
- AMANDA TARVIN (PAUL)
- ALEXANDER DRUMMOND (PAUL)
- SHELLY I. HRANKA (CARSON)
- LAURICE LYNETTE FELD (CARSON)
- WAYNE R. MILLER (CARSON)
- CHRISTOPHER "CD" DAVIDSMEYER (RUBIO)
- KYLE A. MOORE (RUBIO)
- CHUCK ERICKSON (RUBIO)
- RONALD J. KELLER (KASICH)
- MARY F. SHEPHERD (KASICH)
- GEORGE O. WENDT (KASICH)

LEGISLATIVE

FOR STATE SENATOR
FIFTY-THIRD LEGISLATIVE DISTRICT
(Vote for one)

- JASON BARICKMAN

REPRESENTATIVE

FOR REPRESENTATIVE IN THE GENERAL ASSEMBLY
ONE HUNDRED AND FIFTH REPRESENTATIVE DISTRICT
(Vote for one)

- DAN BRADY

COUNTY

FOR CIRCUIT CLERK
(Vote for one)

- DON EVERHART

FOR STATE'S ATTORNEY
(Vote for one)

- JASON CHAMBERS

COUNTY

FOR AUDITOR
(Vote for one)

- MICHELLE L. ANDERSON

FOR CORONER
(Vote for one)

- GARY L. MOREFIELD
- RYAN D. GIBSON
- KATHY DAVIS

FOR MEMBERS OF THE COUNTY BOARD
DISTRICT 2
(Vote for one)

- MATT SORENSEN
- ______ Write-in

JUDICIAL

FOR JUDGE OF THE CIRCUIT COURT
ELEVENTH JUDICIAL CIRCUIT
(To fill the vacancy of the Hon. Charles G. Reynard)
(Vote for one)

- MARK A. FELLHEIMER

FOR JUDGE OF THE CIRCUIT COURT
ELEVENTH JUDICIAL CIRCUIT
(To fill the vacancy of the Hon. Elizabeth A. Robb)
(Vote for one)

- CASEY COSTIGAN

PRECINCT COMMITTEEMAN

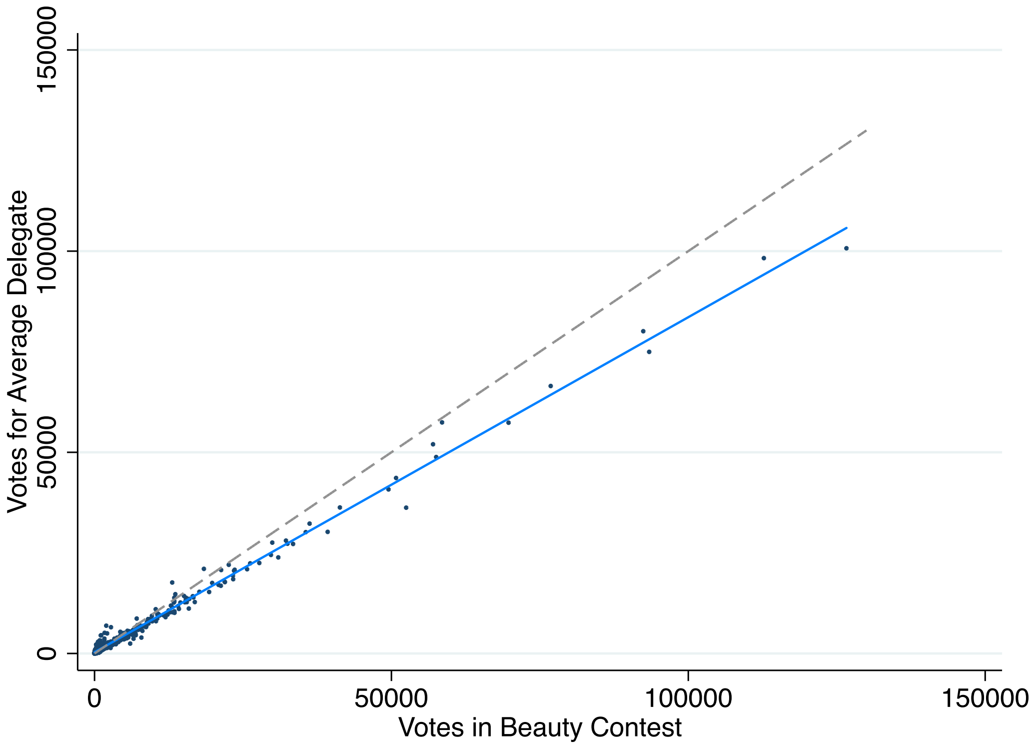
FOR PRECINCT COMMITTEEMAN
TOWANDA 01
(Vote for one)

No Candidate

BACK Card 174 RptPct 670-860 "050-TOWANDA 01"

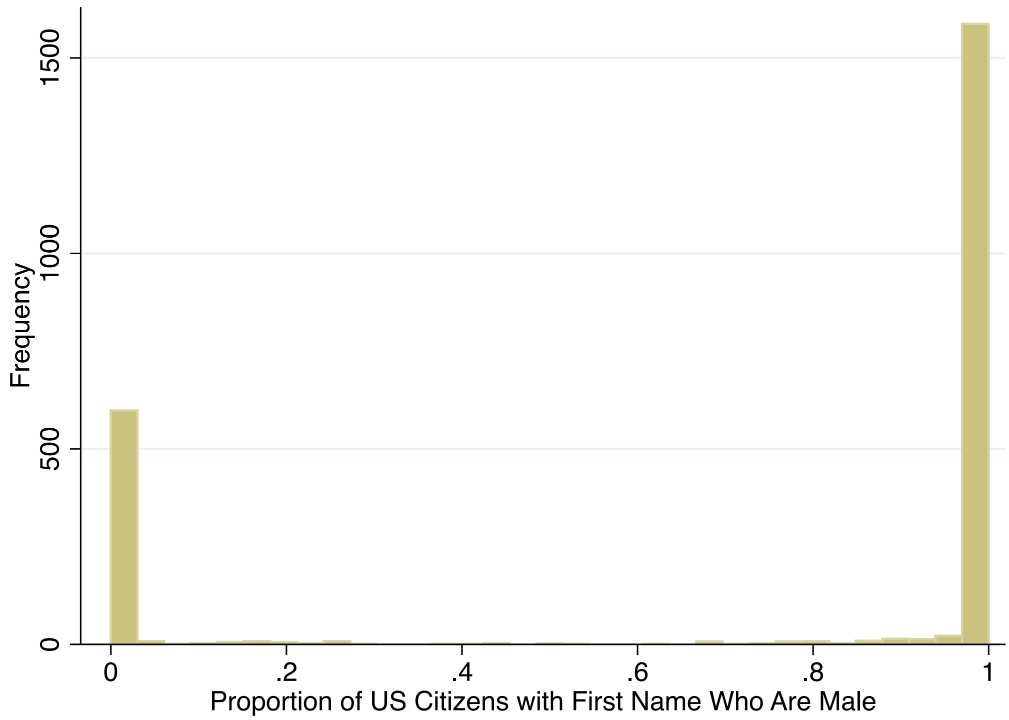
Notes: This figure shows in full the 2016 Republican primary ballot from McLean County in Illinois' 18th Congressional District, including the delegate-selection section. As described in the text, voters vote in both a "beauty contest" election for president which allocates few delegates and in the delegate elections that primarily determine outcomes.

Figure A2: Participation in Delegate Selection versus Presidential "Beauty Contest"



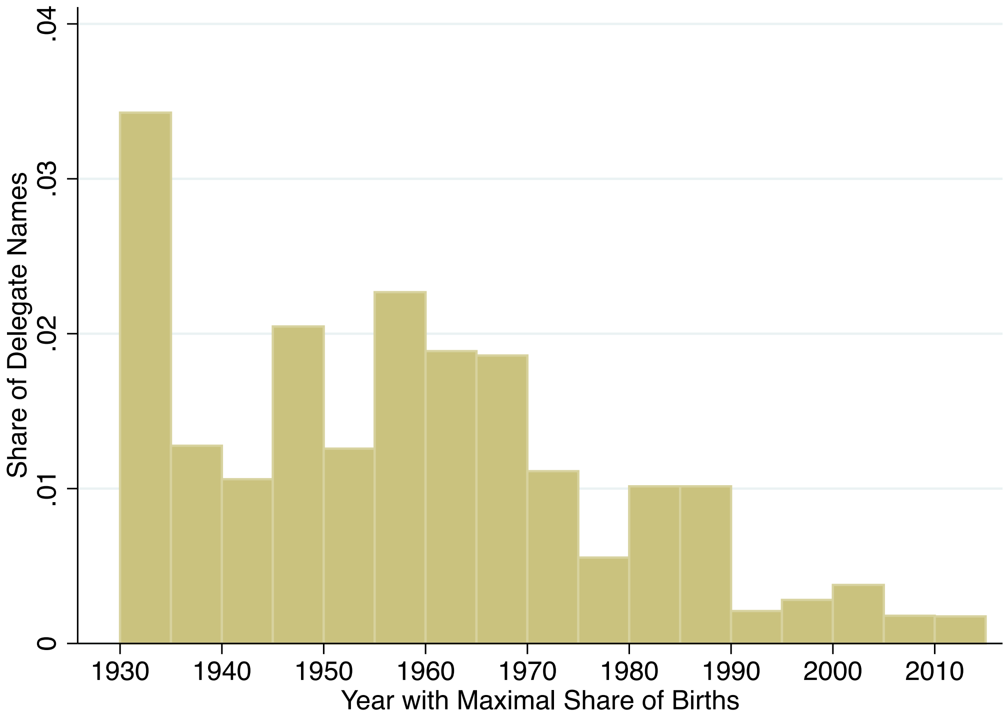
Notes: This figure plots the relationship between the number of votes cast for each Presidential candidate in each county in the "beauty contest" at the top of the ballot and the average number of votes that the same candidate's delegates received in that county. Data are at the county–presidential candidate level. The slope of the blue line is 0.84, with the gray dashed line of unit slope provided for comparison. This indicates that, on average, for every one additional vote cast for a presidential candidate in a county, there are 0.84 additional votes cast for that candidates' delegate in that county.

Figure A3: Histogram of Delegate Candidate First Names by Gender



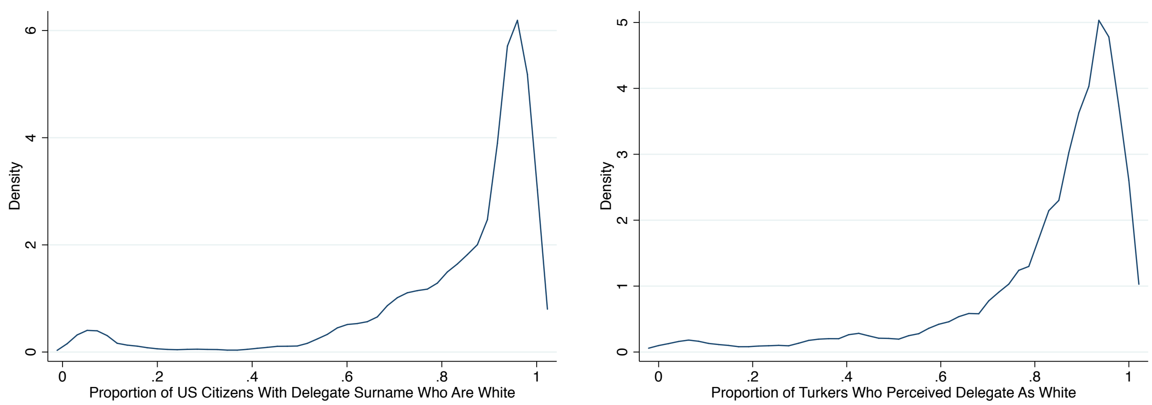
Notes: This figure plots a histogram of our measure of the perceived gender of delegates, which is the proportion of U.S. citizens with the same first name who are male at birth, according to the baby-name file of the U.S. Social Security Administration. Almost all delegate candidates' first names clearly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

Figure A4: Histogram of Delegate Candidate First Names by Modal Year of Birth



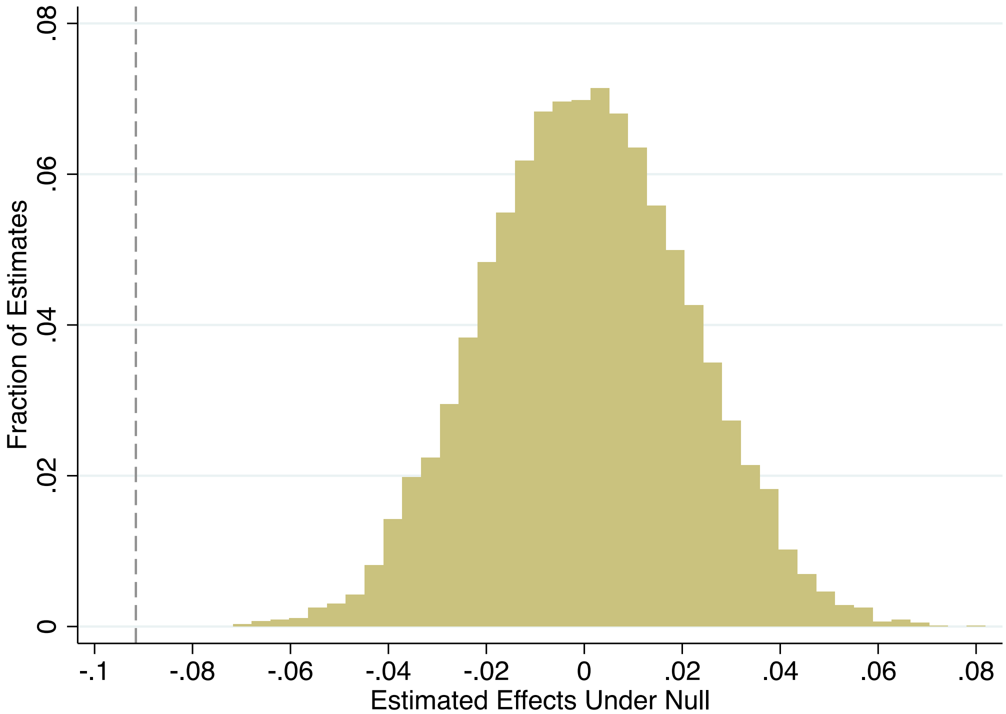
Notes: This figure plots a histogram of our measure of the perceived age of delegates, which is the year in which their first name attains its maximal share of births between 1930 and 2012, according to the baby-name file of the U.S. Social Security Administration.

Figure A5: Kernel Density Plots of Delegate Candidate Names by Whiteness



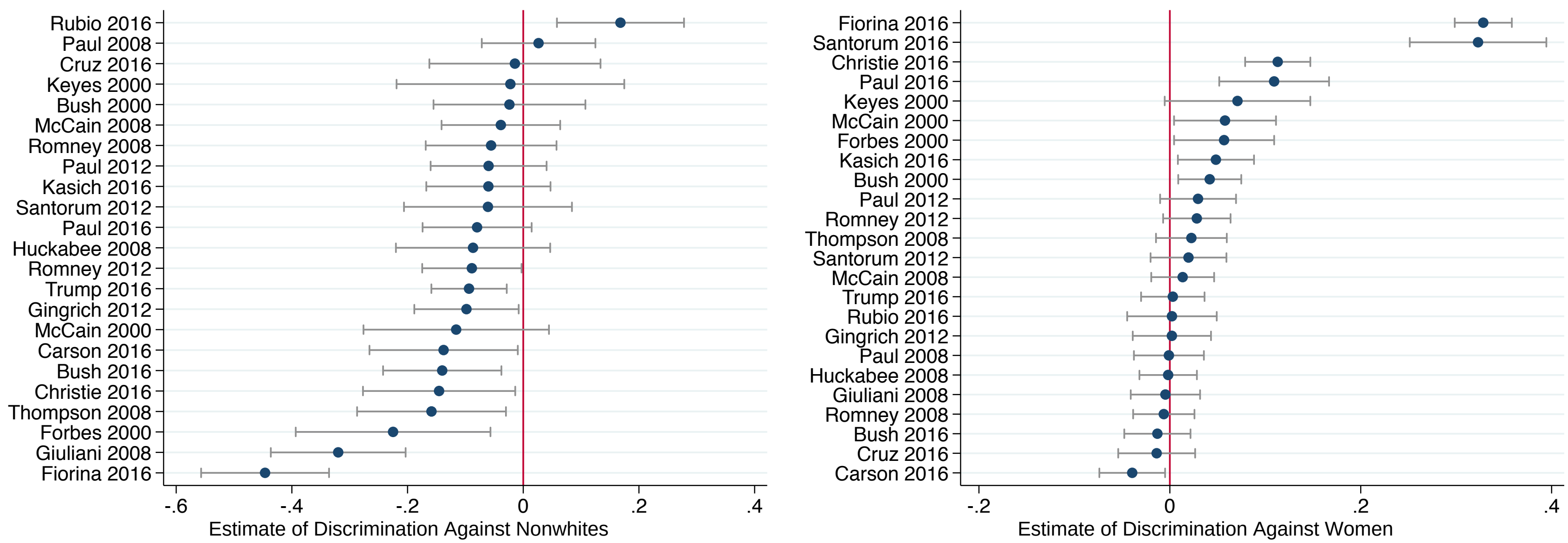
Notes: This figure plots the kernel density estimates of delegate-candidate name whiteness according to two distinct measures. The left panel matches delegate candidate last names to 2000 U.S. Census data on the percent of U.S. citizens with that last name who are non-Hispanic white. The right panel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

Figure A6: Permutation Test for Estimated Effect of a Nonwhite Name



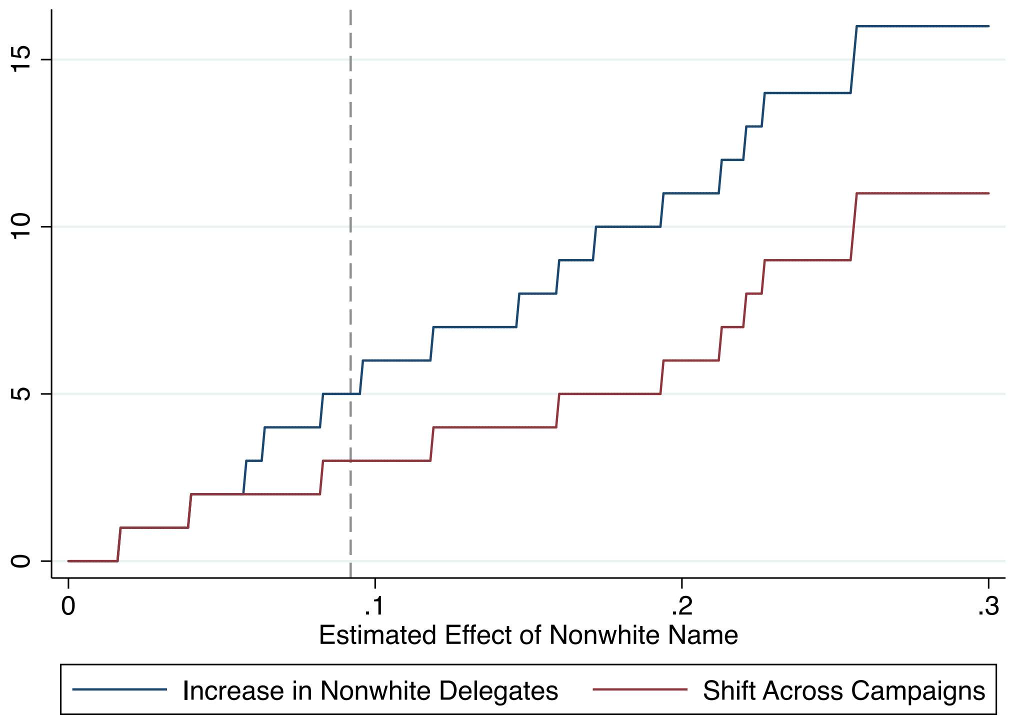
Notes: This figure plots 10,000 estimates of the main regression specification with the treatment re-randomized at the delegate level within presidential candidate–district strata. The coefficient estimated on the true treatment data using the specification in Equation 1 is shown as the dashed line at the left. The coefficient on the PC1 race measure is far from the distribution of the test statistic under the sharp null hypothesis.

Figure A7: Random-Effect Estimates of Discrimination by Candidate



Notes: This figure depicts random-effects (RE) estimates of the level of taste-based discrimination against nonwhites and women by candidate. We use a random-effects model as the estimates would otherwise be too imprecise. To estimate the model, we first partial out cell FEs and controls for ballot order and detailed delegate officeholding and then estimate the RE model with only random slopes by candidate interacting by delegate race and gender. To simplify the estimation of this model, we use the weighted OLS specification with $ln(1 + votes)$ as the dependent variable, as in Table A11. In plotting, we drop respective estimates for presidential campaigns that did not nominate nonwhite or female delegates.

Figure A8: Counterfactual Number of Nonwhite Delegates Without Voter Racial 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. To simplify, we assume delegates more likely than not to be nonwhite according to the MTurk race measure all lost the same fraction of votes due to discrimination and then vary the estimated penalty from 0 to -0.3. We augment their vote totals by these fractions, which assumes that discriminating voters either undervote or proportionally vote for other delegates when nonwhites are nominated. We then calculate which delegates would have won under these augmented vote totals. We use the MTurk measure of race here, as it covers the full population of delegates. 448 delegates and alternates won across the elections we study, only 34 of whom were likely nonwhite. We conduct this analysis to estimate the number of counterfactual nonwhite winners, as regressions where the dependent variable is a binary indicator for whether a delegate candidate wins would be badly underpowered.

B Conceptual Framework

To fix ideas, we present a simple conceptual framework of voter behavior when voters have racial tastes. The framework also motivates the use of vote totals in our empirical analysis as a strategy to identify taste-based discrimination.

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

Voters, each residing in a single county–district, vote for k unique delegate candidates and cannot vote for any given delegate more than once.[40] We express a voter's choice as $c = \{c_1, \ldots, c_k\}$, noting that the ordering of delegates is immaterial. Voter preferences for presidential candidate p are assumed to be additively separable in an expressive utility α_p of voting for p's delegate, an instrumental utility β_p which the voter receives if p wins, and discriminatory-taste parameter δ, reflecting the "psychic cost" of voting for a delegate $i : i \in N$.[41] We restrict these parameters by assuming that $\alpha_i \neq \alpha_j$ and $\alpha_i + \delta \neq \alpha_j$ for any distinct i, j and that $\delta \geq 0$.[42]

The voter's expressive utility of voting for delegate i is $\alpha_{\phi(i)} - \delta \cdot \mathbf{1}(i \in N)$. The voter's problem is

$$\max_{\{i,j,k\}: i \neq j \neq k} \mathbb{E} \left[\beta_p + \sum_{c \in \{i,j,k\}} [\alpha_{\phi(c)} - \delta \cdot \mathbf{1}(c \in N)] \right], \tag{2}$$

where p denotes the winning presidential candidate. We denote the solution $c^* = \{c_1^*, \ldots, c_k^*\}$.

There are two cases to consider: The voter's ith vote is pivotal or not pivotal. If the

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

[41] Such tastes could arise from a variety of underlying behavioral or psychological mechanisms, including "aversive racism" (Dovidio and Gaertner, 2000), "implicit bias" (Greenwald et al., 1998), or others. Another mechanism for such behavior might be termed "negative altruism": that voters do not vote for nonwhite delegates because they derive utility from harming nonwhites (e.g., by preventing nonwhites from gaining any non-political, private benefits associated with attending the convention). These mechanisms are consistent with Becker's (1957) definition of taste as long as voters act "as if" they have preferences over candidate race or gender.

[42] These assumptions imply respectively that the solution to the voter's problem is unique and that no voters prefer nonwhites to whites. As we discuss in Appendix J, 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.

voter's ith vote is not pivotal, then the solution satisfies

$$\alpha_{\phi(i)} - \delta \cdot \mathbf{1}(i \in N) \geq \max_{i'} \left\{ \alpha_{\phi(i')} - \delta \cdot \mathbf{1}(i' \in N) \right\}. \tag{3}$$

If the voter's ith vote is pivotal, then the solution satisfies

$$\alpha_{\phi(i)} + \beta_{\phi(i)} - \delta \cdot \mathbf{1}(i \in N) \geq \max_{i'} \left\{ \alpha_{\phi(i')} + \beta_{\phi(i)} - \delta \cdot \mathbf{1}(i' \in N) \right\}. \tag{4}$$

Then, as the voter maximizes their expected utility, the solution satisfies

$$\alpha_{\phi(i)} + q_{\phi(i)}\beta_{\phi(i)} - \delta \cdot \mathbf{1}(i \in N) \geq \max_{i'} \left\{ \alpha_{\phi(i')} + q_{\phi(i')}\beta_{\phi(i')} - \delta \cdot \mathbf{1}(i' \in N) \right\}, \tag{5}$$

where $q_{\phi(i)}$ is the probability of the voter's ith vote is pivotal for any presidential candidate $\phi(i)$.

A solution to (2) always exists by our assumption that the delegate set is finite. For any $i, j : c_i, c_j \in c^*$,

$$\phi(c_i) \neq \phi(c_j) \implies \delta > \left| \left(\alpha_{\phi(i)} - \alpha_{\phi(j)}\right) + \left(q_{\phi(i)}\beta_{\phi(i)} - q_{\phi(j)}\beta_{\phi(j)}\right) \right|, \tag{6}$$

which means that voters only vote for the delegate candidates of more than one presidential candidate if their racial tastes are sufficiently strong—in particular, if the taste parameter δ is larger than the difference in the expressive and instrumental expected utilities of the two presidential candidates i, j.

As $q_{\phi(i)}$ is very small in most elections, it follows that the strength of tastes δ relative to dispersion in expressive utilities over presidential candidates determines the extent to which taste-based discrimination will occur. Notably, for sufficiently small pivot probability, there may exist votes such that $\alpha_{\phi(j)} + \beta_{\phi(j)} > \alpha_{\phi(j)} + \beta_{\phi(j)} - \delta > \alpha_{\phi(i)} + \beta_{\phi(i)} > \alpha_{\phi(j)} - \delta$, where $c_i \in c^*$ but $c_j \notin c^*$. Even if, when a voter is pivotal, she would prefer voting for a nonwhite delegate of candidate j to a white delegate of candidate i, the probability of pivotality can be so low that she is unwilling to pay the psychic costs δ in expectation to achieve this outcome. A pivotal voter must exist, however, and so if all voters had such preferences, their Nash equilibrium strategy would be not to vote for nonwhites even if all doing so would achieve a Pareto improvement.

C 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 A29 reports the statewide vote share of each presidential candidate in the Illinois primary as well as their polling average in pre-election polls. Table A30 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 A29: 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

Table A30: 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.

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 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 campaign the day after the Illinois primary.

D Coding Delegate Biographies

As argued in Section 4.3, 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.[43]

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 2 and Appendix Table A1.

[43]The complete search instructions provided to RAs are available in the replication materials.

E 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 information in each name. 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.

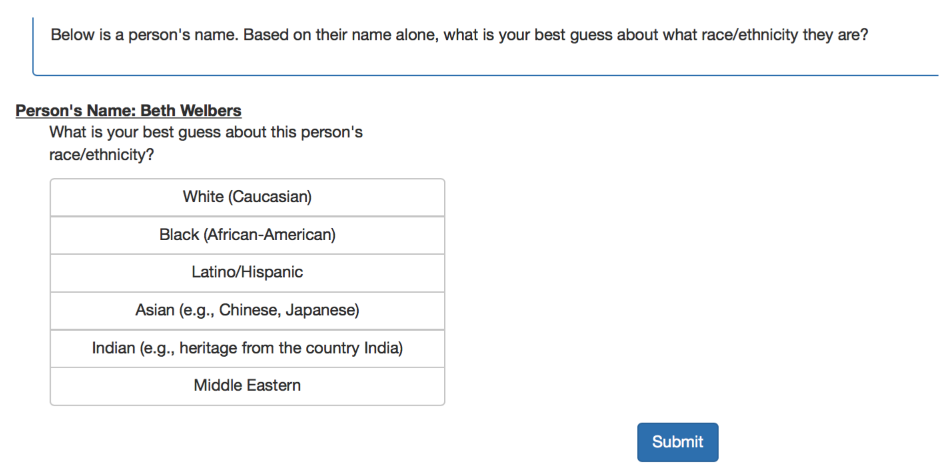
F MTurk Survey for Perceived Delegate Race

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 A9 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 I estimates how attenuated our estimates our by the fact that a finite number of MTurk workers coded their perception of each name.

Figure A9: Coding Interface for Mechanical Turk Workers



Notes: MTurk workers recorded their perceptions using this form.

G Where Nonwhite Candidates Run

As 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 relatively more likely to have nonwhite delegates differ on average from cells more likely to have white delegates, or such differences exist between cells with male and female delegates, then estimates from the identifying set may not generalize well to the full population.

We address this external-validity concern in two ways. First, we examine patterns of selection on observables into the identifying set. Second, we rebalance the sample so that the presence of likely-nonwhite delegates is uncorrelated with county–district observables and that this reweighted sample matches Illinois statewide on these observables. We emphasize here that the existence of selection—in the sense of geographically-nonrandom nonwhite entry, where selection is either observable or unobservable—is not a threat to bias due to the presence of fixed effects which restrict us to within-cell comparisons. Nevertheless, selection of this kind presents a potential risk to external validity, and we show empirically that it does not seem a significant concern in our context.

Detecting Selection in Nonwhite Delegate Entry

We regress the cell means of the PC1 race measure on various county–district observables. Appendix Table A28 reports results. The areas of Illinois with relatively more nonwhite delegate candidates are less Republican, as measured by the share of two-party vote in the general election the Republican presidential candidate received, but are closely similar in terms of the white share of population, the college-educated share of whites, and average white per-capita income. Areas that nominate relatively more women also appear to be less Republican, but vary little on other observables. More generally, the results are inconsistent with the possibility of strong selection on observables in the identifying set, as the racial and gender mix of the delegate-candidate population varies relatively little with demographic variables.[44]

Conversations with individuals involved in Illinois Republican Party politics suggested that search costs and supply constraints play important roles in explaining why campaigns nominate nonwhites in the first place if these delegates receive fewer votes. In particular, the alternative to a nonwhite or female delegate may be leaving the delegate slot unfilled, rather

[44]For further evidence, we split the sample in halves by the cell-average nonwhite or female name probabilities and compare the Republican two-party vote share between halves, the variable for which we found the most notable differences. The Republican vote share was 3.8 p.p. lower in the half with more nonwhite delegates, and 2.4 p.p. lower in the half with more female delegates.

than a white male delegate who campaigns believe would face little risk of discrimination by a Republican primary electorate. The data also suggest some further explanations. 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 2 and Appendix Table A1, 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. The minimal political value of serving as a delegate, as indicated by the apparent supply constraints campaigns face (see Footnote 21 for evidence that campaigns recruit delegates from their email lists), also suggests that service as a delegate is unlikely to meaningfully further one's political career. It is unclear why service at a national party convention in a different state would advance one's career in local politics (e.g., if running for town or county council). None of the campaign staff we spoke with indicated this would be a motivation for serving as a delegate. State parties select delegates in most other states and usually select donors; this indicates that serving as a delegate is likely a form of consumption for political activists who enjoy participating in politics. Finally, an additional potential form of statistical discrimination is that voters want to undermine the future political careers of nonwhites because they expect them to be more liberal. However, our finding in Table A21 that voters for more liberal presidential candidates discriminate no less than voters for more conservative presidential candidates is inconsistent with this channel.

Rebalancing Procedure to Address Selection in Entry

We employ an extension of coarsened exact matching (CEM), originally developed in Iacus et al. (2012), to rebalance our sample so that, in the presence of heterogeneous treatment effects, our design estimates the average treated effect of a likely-nonwhite name across Illinois statewide, not weighted by actual patterns of nonwhite delegate entry.

We coarsen the same four county–district demographic observables—the white share of population, the college-educated share of whites, the per-capita income of whites, and the Republican two-party vote share in the most recent presidential election—by splitting the sample into equal-frequency bins for each variable. We use two equal-frequency bins for the first three variables and, given the significant partial correlations of the Republican vote share in the preceding analysis, five for this variable. These coarsened covariates form 35 bins, as 40 ($= 5 \times 2^3 - 35$) combinations contain no observations. We also coarsen a cell-level version treatment variable, the cell-level maximum MTurk nonwhite variable, into three equal-frequency bins.[45] The CEM weights are defined as

$$\omega_c = \frac{N_j}{N_{ij}} \cdot \frac{N_i}{N},$$

where $N_j = \sum_{c \in j} N_c$ is the sum of the vote counts of cells c that are assigned to covariate bin j. N_{ij} is for cells assigned to covariate bin j and with a coarsened nonwhite variable in bin i. N_i is for all cells whose coarsened nonwhite variable in bin i. N is the total sum of the weights.

Appendix Table A31 reproduces our baseline results in Column 1 and, in Column 2, uses the CEM weights which rebalance the sample so that cells in which nonwhites run match Illinois statewide on county–district observables. The estimates are almost identical. Selection on these observables does not appear to be a critical issue for external validity.

Table A31: Comparison of Results with Standard and CEM Weights

	(1) Unweighted	(2) CEM Weights
Nonwhite	-0.099***	-0.094***
	(0.013)	(0.013)
N	17,126	17,126
Pseudo-R^2	0.991	0.991

Notes: In all regressions above, the dependent variable is the vote count, the unit of observation is the county–district-delegate-year, and the race measure used is the rescaled PC1 measure. All regressions include cell-level FEs. In Column 1, weights by the maximum number of votes won by a delegate candidate in the cell. In Column 2, we adjust these weights by the coarsened exact matching procedure explained above. Delegates are defined as in the same cell if they 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. Standard errors are clustered at the delegate level. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

[45] The MTurk variable is not missing for any delegates, simplifying the coarsened exact matching. We continue to use the PC1 race measure in the regressions.

H Why Does Anyone Run for Delegate?

It may be worthwhile to explain why anyone, whatever their race, runs for a delegate position, especially because delegates themselves pay for the privilege of attending the convention. Based on firsthand experience (one of the authors of this paper has served as a Democratic delegate) and a detailed review of interviews of delegate candidates in local newspapers, we view this as primarily a form of consumption: Political participation often resembles an enjoyable hobby (Hersh, 2020). Conventions are a festive event, with many activities designed to entertain delegates, and delegates can be seen dancing or laughing on the convention floor. In this way, conventions are similar to other political rallies for candidates that people attend to enjoy, or other forms of political expression such as campaign contributions. Even for officeholders, every interview we could find suggested delegate service was little if any value to their political careers.

In what follows, we provide an assortment of news clippings which speak to the motives of delegates:

- Russell Berman, "What Actually Happens at the U.S. Presidential Conventions?", *The Atlantic*, 10 July 2016:

  > "It's the ability to participate in a historic political moment," [said Julian Zelizer said]. "Barring some really unexpected turn at the Republican convention," he added, "they don't really have much power other than to raise their hand and vote. It's not as if when they go back to their states they're these power brokers because they've been at the convention. It really is largely symbolic."

- Stephen Ohlemacher, "So, you want to be a delegate to the Republican National Convention?", *PBS Newshour*, 9 April 2016:

  > "Suppose that your passion in life is helping out on the local level with political campaigns or with party work," he [Ben Ginsberg] said. "This is the reward at the end of a four-year rainbow."

- Marni Pyke, "'Nothing is more important to me': Delegate candidates push petitions, pump up prospective voters in the suburbs", *Arlington Heights* [IL] *Daily Herald*, 20 January 2020:

  > It cost Democratic state Sen. Cristina Castro of Elgin about $2,200 to attend the 2012 Democratic National Convention, she told attendees. "But to be

> finally part of the casting of the vote . . . it's an amazing experience," Castro said."

- Scott Fitzgerald, "'Terri Bryant says she has the experience for 115th House seat", *The Southern Illinoisian*, 14 March 2014:

> During her career as a state employee, [Terri] Bryant worked in her limited free time to build a political career for herself...."This has all been on my own dime and time. I love it," Bryant said.

- Maureen O'Reilly, "Illinois delegates enjoy anti-Clinton speeches, but also want pro-Trump talk at convention," *Belleville* [IL] *News–Democrat*, 16 July 2020

> For Hough, Logan and Kozanecki, the picture-perfect moments didn't come from the appearances from Republicans of celebrity-like stature, a former model and her reality-TV husband. Instead, it came from former service-members of the Benghazi security team. The trio of Illinois delegates bumped into Mark Geist and John Tiegan in the Quicken Loans arena after the Marines' joint speech ended. Geist and Tiegan posed for selfies with the three women.

I 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. Throughout, we use the weighted OLS specification where the dependent variable is the logarithm of the vote count plus one, not our main Poisson specification, to simplify the implementation of these corrections. 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 A7 reports both reliability estimates, and Table A32 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 Appendix Table A11. 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 discrimination against nonwhites is only modestly attenuated. 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 A7 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 A32 presents estimates for the detailed nonwhite race categories. Consistent with differential reliabilities, we find that the estimate of 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. These results require the assumption that our three measures capture the same latent construct. The instability of the black coefficient across race measures provides a further caution against emphasizing this large point estimate.

Table A32: 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 $\ln(1 + \text{votes})$ for the delegate candidates. The unit of observation is the county–district-delegate-year. 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. * = $p < 0.10$, ** = $p < 0.05$, *** = $p < 0.01$.

J Other Alternative Interpretations

In this Appendix, we consider several additional alternative interpretations of our results. In the first subsection, we discuss these alternatives and the extent to which our main results address them. In the second subsection, we report results of an original survey of self-reported Illinois Republican primary voters, providing additional evidence on the plausibility of some of these alternatives.

Discussion

Inferences about presidential candidates. It is possible voters make inferences about the presidential candidates from the race, ethnicity, and gender of the delegates they nominate. However, our fixed effects mean that such inferences must affect the delegates of a given presidential candidate heterogeneously 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]

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 (2019), we therefore evaluate the general plausibility of the claim that unobserved variables could explain our finding of 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 racial discrimination to be zero, demeaning our data at the cell level and assuming that controlling for unobservables would raise the within-cell R^2 by 30 percent, the ratio of selection on unobservables to selection on observables would need to be larger than 24.5, well above the threshold of unity Oster (2019) recommends and more robust than nearly all studies in Oster's sample.

Implications of indifference for the cost of discrimination. Voters indifferent between

[46] A variant of this concern involves voters voting for white delegates who appear high on a ballot before they realize a presidential candidate nominated a nonwhite delegate lower on the ballot, at which point voters infer something about the presidential candidate which causes them to select out. Were this behavior responsible for our results, our estimates would 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, Appendix Table A20 reports similar, significant discrimination regardless of ballot order. We conclude our results are unlikely to be affected by voter inferences about presidential candidates from delegate race.

presidential candidates may split their votes across delegates for multiple presidential candidates. Those who do so randomly will not bias our estimates away from zero. Indifferent voters could also lexicographically choose white delegates over nonwhite delegates due to arbitrarily-weak racial tastes against nonwhites. While this behavior would allow us to retain an interpretation of the results as revealing that many voters have racial tastes stronger than their presidential candidate preference, these racial tastes could then be weak in absolute terms, raising questions about whether this behavior would manifest in other elections. Relative to other primaries, few voters in these primaries, however, are likely to be indifferent between the presidential candidates: As discussed in Section 2.5, 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.[47] Moreover, voter preferences may well be stronger in presidential primary elections than in other primary elections. We show in Appendix Table A4 that Illinois primary turnout is higher in presidential years than in non-presidential years. In addition, a general feature of presidential elections is that many voters vote for presidential candidates and leave the rest of the ballot blank, and less than 1 percent of presidential election voters indicate they are indifferent, an order of magnitude less than in 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 primary elections where their preferences over candidates are likely weaker.

Are voters signaling to presidential candidates and party elites? Another alternative interpretation of the results is that voters do not vote for certain delegate candidates as a way to send a signal about their policy preferences to presidential candidates or party elites. For example, if a voter opposes policies that benefit Asians, refusing to vote for a presidential candidate's Asian delegates could signal those views to that presidential candidate or other party elites. We conducted a survey of a convenience sample of Illinois Republican primary voters that sought to evaluate this possibility. In the next subsection we show that very few voters in this survey indicated that would engage in this behavior, and that far more indicated that voting for nonwhite delegates would "make them uncomfortable." However, although our survey cannot rule out signaling, it seems unlikely that voters believe presidential candidates and/or elites both inspect delegate vote totals in this detail and would interpret their behavior in this way. We are also unaware of any empirical evidence that voters use their votes to signal to politicians.

Do voters understand that delegate voting has stakes? Another possibility is that voters do not understand that they should vote for all delegates of their preferred presidential

[47]By the primary dates in 2000, 2008, and 2012, the contest had narrowed to only two plausible candidates. Voters were therefore unlikely to be indifferent between two candidates who they both preferred to a third.

candidate to maximize the value of their ballot. First, because the ballots clearly indicate that voters get N votes, it seems unlikely that voters misunderstand how many delegate votes they can cast. The independence of our estimates of discrimination from ballot order, reported in Appendix Table A20, suggests that the discrimination we observe is not driven by an appreciable share of voters who believe they only have one vote selecting the first white delegate that appears and no other delegates. Consistent with this, in the same survey as reported below, essentially all voters shown a ballot where $N = 3$ perceive themselves to have three votes. Second, do voters understand their delegate votes matter for their preferred presidential candidate's election prospects? We suspect that most voters who do not perceive an incentive to vote for delegates would simply not cast any votes for delegates, removing themselves from the sample entirely, rather than only voting for white delegates. In addition, in the survey reported below we show that most voters in our survey understood that not voting for all their preferred presidential candidates' delegates is costly and that those who did understand this still appear to engage in discrimination in a survey version of the ballot. Voters might also misunderstand that delegates are bound to presidential candidates, and we discuss these above as residual incentives for statistical discrimination. Finally, our results in Section 5 that discrimination varies across candidates and elections as taste-based theories predict are also consistent with our estimates reflecting taste-based discrimination.

Survey of Illinois Republican Primary Voters

We surveyed a convenience sample of approximately 700 Illinois Republican primary voters recruited by targeted advertisements on Facebook, as in Sances (2017) and Zhang et al. (2017). The purpose of the survey was to further assess the plausibility of several alternative explanations for our paper's interpretation of the observed racial and ethnic discrimination in Illinois Republican primary elections as taste-based. In particular, we designed survey questions to assess whether voters understood the election environment or whether voter behavior can be reconciled with statistical discrimination or signaling to presidential candidates. In this Appendix, we provide detail on our survey procedures and its results. The full survey and data are available in the replication materials.

We note here for completeness that in a separate survey on MTurk discussed in Appendix F, we collected subjective opinions on the likely race and ethnicity of delegates.

Survey Recruitment Procedures

Identifying Republican primary voters in a general population survey would have been prohibitively costly, requiring a survey of tens of thousands of Illinois residents to yield a sufficient sample of Illinois Republican primary voters. We therefore relied on Facebook's ad-targeting tools to attempt to specifically recruit respondents who had a high probability of being Illinois Republican primary voters and then asked respondents to self-report whether they were in fact in this population.

To do so, we first bought Facebook advertisements that advertised the survey that were targeted to individuals who lived in Illinois and were "interested in" any of the 2016 Republican Presidential candidates, or matched the "interest" categories "Republican Party (United States)," "Illinois Republican Party", "US politics (conservative)," "US politics (very conservative)," or "Likely to engage with political content (conservative)." The construction of these "interest" categories is proprietary to Facebook but is based on user's posts on the platform, the content they engage with by reading and "liking" it, and administrative data that Facebook imports into the platform based on user's names. Facebook determines a user's state of residence from the city that they report on their Facebook profile and from device and connection data.

A sample of Facebook users from this population saw one of the two ads in Figure A10, both of which ask voters who participated in the 2016 Republican primary to complete "a 3 minute survey about your experience" and indicate the survey is "from Stanford University." Individuals who clicked the Facebook advertisements were directed to a Qualtrics survey. The first set of questions in the survey screened individuals by their self-reported current state of residence, whether they said they have voted at least once in an Illinois Republican presidential primary,[48] and a basic attention check. There were 698 individuals who met all these criteria.

The survey contained 19 questions, with one additional question for some respondents. About 75 percent of respondents who began the survey completed it. The median respondent took about 5.5 minutes to complete the survey. Those who did not complete the survey answered on average 48 percent of the questions.

Table A33 compares the demographics of our sample to the Illinois Republican electorate in the 2016 voter file provided by the Illinois Secretary of State, a processed version of which we accessed through the firm Catalist. After finding an overrepresentation of men in a pilot sample, we stratified our sample collection by gender to approximate the gender composition in voter file data. We find that the resultant Facebook sample approximately matches the Illinois Republican primary voter population by age, sex, and race.

[48] We asked "Have you ever voted in a Republican Presidential primary election in Illinois, the election in Illinois between candidates to be the Republican nominee for President?" and only accepted respondents who answered "Yes, I have voted in an Illinois Republican Presidential primary election at least once before" and not "I have only voted in Republican Presidential primary elections outside of Illinois" or "No, I have never voted in a Republican Presidential primary election."

Figure A10: Facebook Ads to Recruit Survey respondents

Notes: This figure shows the Facebook ads that appeared for users in Illinois to recruit survey respondents.

Results

Below we report the survey's results relevant to four alternative explanations for the paper's results about racial and ethnic discrimination.

We encourage caution when interpreting the survey results. Responses were unincentivized, and we use a convenience sample. In addition, measurement error in survey responses biases sample quantities of binary outcomes away from the extremes of 0 and 100 percent Ansolabehere et al. (2015). A final caveat is that our survey took place in November 2017, 20 months after the 2016 Illinois Republican presidential primary was held. Campaigns and the media in Illinois likely attempt to educate voters about the ballot and their incentives prior to the election, but our survey respondents may have forgotten this information in the nearly two years since the election. Our results may therefore understate the share of voters who correctly understand the election environment.

Do Voters Understand They Have Three Delegate Votes?

We first asked respondents to select the presidential candidate they preferred in the 2016 presidential primary from a list.

We then showed survey respondents the sample ballot in Figure 1 and asked them to read the instructions, advising them they would receive questions about the voting procedure:

Table A33: Facebook Sample versus Illinois Republican Primary Voter Population

	By Source	
Variable	Facebook	Voter File
Age (Mean)	51.8	56.0
Female (%)	48.0	49.0
Non-Hispanic White (%)	90.7	95.2

Notes: This table reports demographic summary statistics on the Facebook sample in comparison to administrative voter file data on Illinois voters in 2016, also reported in Table A2.

> Below is an example of part of the ballot for the 2016 Republican Presidential primary in Illinois. Please take a moment to look at the ballot, read the instructions, and imagine you tried to use it to vote. Then we will ask you several questions about it.

On the ballot, voters are instructed they can vote for up to three delegates. One alternative explanation for the paper's results is that voters are unaware they have three delegate votes and so cast only one vote, which would allow arbitrarily weak taste-based discrimination to generate the results—that is, a voter indifferent between three delegate candidates who believes they only have one vote chooses the white candidate.

To test whether respondents understood they had three votes, we asked them for the maximum number of delegate votes they are allowed to cast:

> How many of [*respondent's preferred presidential candidate*]'s three delegates do you think you are allowed to vote for on the ballot above?

We find that essentially all respondents understand that they have three votes: 96 percent of respondents (95% CI: [94.3%, 97.5%]) responded correctly that they can cast up to 3 delegate votes. Among those who said that they remembered previously voting for delegates in the primary,[49] 97 percent said they had three votes (95% CI: [95.5%, 98.9%]).

Do Voters Understand That Not Voting for Nonwhite Delegates Is Costly?

We next asked a question designed to elicit whether respondents understood that not voting for all of their preferred presidential candidate's delegates could be costly to their preferred

[49] We asked respondents "When you lasted voted in a Republican Presidential primary, do you recall having voted for delegates using a part of the ballot like the above?" Those who answered "Yes, I remember that I did vote for delegates last time I voted in the Republican Presidential primary" we refer to as having remembered previously voting for delegates.

presidential candidate. In particular, we asked respondents to imagine a scenario in which the primary was very close between two presidential candidates:

> Suppose the Illinois Presidential primary election was very close between [*preferred presidential candidate*] and another candidate.
>
> If many people did not vote for one of [*preferred presidential candidate*]'s delegates, might [*preferred presidential candidate*] end up with fewer delegates from Illinois at the Republican National Convention?
>
> - Yes, if one of [*preferred presidential candidate*]'s delegates received fewer votes, this might reduce the number of [*preferred presidential candidate*] delegates that win in Illinois.
> - No, [*preferred presidential candidate*] will receive the same number of delegates from Illinois regardless of how many votes [*his/her*] delegates receive on the ballot.

65 percent of respondents (95% CI: [61.6%, 69.1%]) answered this question correctly. Among those who said they remembered previously voting for delegates, 72 percent answered the question correctly (95% CI: [67.1%, 76.5%]). This difference may reflect voters who do not realize they have incentives to vote for delegates simply not filling out the relevant part of the ballot.

Together with our result that almost all respondents understand that they get three delegate votes, this indicates that most voters likely to enter our election sample understand the electoral environment.

We next provided respondents a list of three delegates and asked them to suppose they are the delegates for their preferred presidential candidate. One of these names was selected at random from names coded as likely white by the PC1 race variable, and one was selected at random from names coded as likely nonwhite. The third name was "Bill Hadley," a likely-white name we left constant across respondents:

> Suppose the delegates on the ballot for [*preferred presidential candidate*] were as follows:
>
> - [*Randomly assigned white name*] ([*preferred presidential candidate*])
> - [*Randomly assigned nonwhite name*] ([*preferred presidential candidate*])
> - Bill Hadley ([*preferred presidential candidate*])
>
> Which would you vote for? (Vote for not more than three.)

We find that about 6.3 percent (95% CI: [4.3%, 8.3%]) of respondents who voted for both white delegates do not vote for the nonwhite delegate. Restricting the sample to respondents who understood that the election environment created costs to discrimination, we find that they still discriminated: 4.9 percent (95% CI: [2.7%, 7.0%]) of such respondents who selected both white delegates did not select the nonwhite delegate. This is consistent with our interpretation of the results as representing conscious taste-based discrimination.[50]

Do Voters Perceive Incentives for Statistical Discrimination?

We next evaluate whether voters perceive incentives for statistical discrimination because they believe nonwhite delegates are more likely to not vote for their Presidential candidate at the Republican National Convention than white delegates. Voters did not have had any actual incentives for statistical discrimination in 2000, 2008, and 2012: Defection is impossible on the first ballot under Convention rules, and in these years, the Convention was essentially guaranteed to be resolved on the first ballot because it was down to a two-candidate race. Moreover, all the primaries we study had essentially narrowed to two contenders. As a consequence, voters who believe nonwhite delegates are likely to defect to their preferred presidential candidate's rival have only the alternative of themselves advantaging that rival. Finally, voters attempting to statistically discriminate on the basis of incorrect beliefs are considered to be engaged in taste-based discrimination under the Becker (1957) model (see our discussion in Footnote 6). However, as a matter of external validity, if the taste-based discrimination we observe is due to voters' attempts to engage in statistical discrimination on the basis of incorrect beliefs, this behavior might not generalize well to primary elections for candidates.

To assess whether voters believe that nonwhite delegates are more likely than white delegates to not vote for their specified presidential candidate at the convention, we asked respondents:

> Again suppose [*randomly assigned white name*] and [*randomly assigned nonwhite name*] were two of the delegates listed on the ballot as for [*preferred presidential candidate*].
>
> Also suppose both won and served as delegates at the Republican National Convention.
>
> When the Republican National Convention held a vote on who to nominate for

[50]Unsurprisingly, respondents who understood discrimination was costly were 4.9 percentage points (95% CI: [-0.0 p.p., 9.9 p.p.]) less likely to discriminate than those who did not understand, even though they still discriminated.

President, how do you these two delegates would be most likely to vote? Do you think each would be most likely to vote for [*preferred presidential candidate*], for some other candidate, or just not vote?

This was followed by a grid with the randomly assigned white and nonwhite names followed by the options "Vote for [*preferred presidential candidate*] at the Republican Convention," "Vote for some other candidate at the Republican Convention," and "Just not vote at the Republican Convention."

We find only very small differences in the perceived probability of disloyalty between white and nonwhite delegates: About 3.6 percent of respondents expect white delegates not to vote for their preferred presidential candidate, as compared to about 5.9 percent for nonwhite delegates. The difference of 2.7 p.p. (95% CI: [1.1 p.p., 4.2 p.p.]) is too small for perceived differences in loyalty to plausibly explain our finding of discrimination against nonwhite delegates. In addition, 31 percent (95% CI: [8%, 54%]) of respondents who say they expect nonwhites, but not whites, to be disloyal nevertheless vote for the nonwhite delegate, meaning the share who plausibly attempted to engage in statistical discrimination is even smaller. Moreover, we suspect that this represents an overestimate, as the question was placed after we asked respondents how they would have voted, and some respondents may have been attempting to rationalize a discriminatory response.

Furthermore, as a reminder, note that even if voters thought nonwhite delegates for their Presidential candidate were more likely to defect than white delegates for their Presidential candidates, this alone would not be enough for voters to perceive incentives to engage in statistical discrimination. Voters would also need to believe that the delegate bound to the other Presidential candidate who would win instead would be *more* likely to vote for their preferred candidate. For example, a voter preferring Mitt Romney in 2012 would need to believe that a white delegate bound to Rick Santorum was more likely to vote for Mitt Romney than a nonwhite delegate bound to Mitt Romney. We suspect this would be a small fraction of the 2.7 percent of respondents who perceived nonwhite delegates for their candidate to be more likely to defect than the nonwhite delegates for their candidate.[51]

Are Voters Discriminating As a Signal To Their Preferred Candidate?

Another alternative explanation for our results is that voters did not vote for a nonwhite delegate as a way to send their preferred Presidential candidates a signal about what policies to support if elected. To assess this possibility, after the other questions, we also asked all respondents:

[51]We did not explicitly ask about respondent's beliefs about the likely behavior of delegates bound to other Presidential candidates because we thought it would be too difficult to clearly explain to respondents.

> Sometimes, [*Asian/Hispanic/Indian/Middle Eastern/black*] people appear on the Illinois Republican primary ballot as possible delegates for Presidential candidates such as [*preferred presidential candidate*]. Do you agree or disagree with the statements below?
>
> - I would feel uncomfortable voting for [*Asian/Hispanic/Indian/Middle Eastern/black*] delegates.
> - If there were a [*Asian/Hispanic/Indian/Middle Eastern/black*] delegate listed for [*preferred presidential candidate*], I would not vote for that delegate as a way to send [*preferred presidential candidate*] a message about what policies [*he/she*] should support if elected.

The answer choices associated with each bullet were "Strongly agree," "Somewhat agree," "Neither agree nor disagree," "Somewhat disagree," and "Strongly disagree." For the block labeled as "[*Asian/Hispanic/Indian/Middle Eastern/black*]," we randomly assigned which race we asked about.

Given social desirability bias, we did not expect respondents to truthfully report that they "felt uncomfortable" voting for nonwhites, our phrase to elicit discriminatory tastes. Nevertheless, respondents were significantly more likely to answer that voting for nonwhites would make them "feel uncomfortable" than answer that they would "to send a message about what policies" their preferred candidate "should support if elected." 10.4 percent of respondents (95% CI: [8.2%,12.7%]) agreed with the "uncomfortable" phrase, as compared to 6.4 percent (95% CI: [4.6%,8.3%]) with the "send a message" phrase, a statistically significant difference ($p < 0.05$). It is noteworthy that such a large share of voters selected the "uncomfortable" phrase, as if anything this should be the answer most subject to desirability bias, meaning the true share who agree with this statement may be larger. Those who opt into completing an academic survey may also be more educated on average, and therefore less likely to harbor discriminatory tastes. The arguably more desirable response of "send[ing] a message" is itself not likely to be gravely understated by desirability bias. This survey evidence therefore suggests that signaling voters are likely few by comparison to voters who act on racial tastes.

Summary: Survey Results

A limitation to these results is that they come from an unincentivized survey environment rather than an actual election, raising the potential concern of experimenter demand effects. To the extent our survey participants apply more effort in understanding the ballot design than actual voters in elections, our results will understate voter misunderstanding of the election design. Recent research, however, suggests these experimenter demand effects (De Quidt et al., 2018; Mummolo and Peterson, 2017) are generally small in experiments similar to ours where participant effort is not directly encouraged or discouraged. In addition, it is possible that the sample of voters who selected into our survey was unusually politically informed, although primary electorates are also more informed than average registered voters in general as well.

Repeating caveats about the reliability of survey data for studying discrimination, our convenience-sample survey of Illinois Republican primary voters is consistent with several conclusions: (1) most Illinois Republican primary voters understand the relevant features of the election environment, and those that do still discriminate; (2) beliefs of differential defection of nonwhite delegates at the convention are not widespread enough to plausibly rationalize our results; and (3) voter attempts to send a signal to presidential candidates through delegate votes also cannot plausibly rationalize our results. By implication, we view the survey data as favorable to our preferred explanation of our main results as voters reflecting engaging in taste-based discrimination against nonwhite delegates.

References for Appendices

Ansolabehere, Stephen, Samantha Luks, and Brian F Schaffner, "The perils of cherry picking low frequency events in large sample surveys," *Electoral Studies*, 2015, *40*, 409–410.

De Quidt, Jonathan, Johannes Haushofer, and Christopher Roth, "Measuring and Bounding Experimenter Demand," *American Economic Review*, 2018, *108* (11), 3266–3302.

Dovidio, John F and Samuel L Gaertner, "Aversive racism and selection decisions: 1989 and 1999," *Psychological science*, 2000, *11* (4), 315–319.

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

Greenwald, Anthony G, Debbie E McGhee, and Jordan LK Schwartz, "Measuring individual differences in implicit cognition: the implicit association test.," *Journal of personality and social psychology*, 1998, *74* (6), 1464.

Hersh, Eitan, *Politics Is for Power: How to Move Beyond Political Hobbyism, Take Action, and Make Real Change*, Scribner, 2020.

Iacus, Stefano M, Gary King, and Giuseppe Porro, "Causal inference without balance checking: Coarsened exact matching," *Political analysis*, 2012, *20* (1), 1–24.

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

Mummolo, Jonathan and Erik Peterson, "Demand Effects in Survey Experiments: An Empirical Assessment," 2017.

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 Evidence," *Journal of Business & Economic Statistics*, 2019, *37* (2), 187–204.

Pettigrew, Stephen, Karen Owen, and Emily Wanless, "US House Primary Election Results (1956-2010)," 2014.

Ravn, Morten O and Harald Uhlig, "On adjusting the Hodrick-Prescott filter for the frequency of observations," *Review of Economics and Statistics*, 2002, *84* (2), 371–376.

Sances, Michael W, "Ideology and Vote Choice in US Mayoral Elections: Evidence from Facebook Surveys," *Political Behavior*, 2017, pp. 1–26.

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.

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).

Zhang, Baobao, Matto Mildenberger, Peter D. Howe, Jennifer Marlon, Seth Rosenthal, and Anthony Leiserowitz, "Quota Sampling Using Facebook Advertisements," 2017.