Honors Thesis Collection

2016

Voter Discrimination in Democratic Elections

Suzanne Barth
sbarth@wellesley.edu

Follow this and additional works at: https://repository.wellesley.edu/thesiscollection

Voter Discrimination in Democratic Elections

Suzanne K. Barth
Advised by Professor Kyung H. Park, Economics

Submitted in partial fulfillment
of the
Prerequisite for Honors
in Economics

April 2016

Acknowledgments

There are many people that I would like to thank for their support and encouragement over the course of this year. I would first like to thank my advisor, Professor Kyung Park, for giving me the opportunity to work with him on this project for the 2015 Social Science Summer Research Program, and allowing me to continue this work as an honors thesis. Without his trust and mentorship, I could not have completed this thesis.

I am grateful for Casey Pattanayak of the Quantitative Analysis Institute, who helped me acquire the necessary data management skills to complete this project. I am equally indebted to the Library and Technology Support research librarians for helping me find many of the data my thesis required.

I would also like to thank Professor Robin McKnight and the students of the Economic Research Seminar. Their support throughout the year was instrumental to the success with this thesis, and my project benefited greatly from their thoughtful feedback.

Thank you to Corey Loeb, Abby Esrock, Mariya Patwa, Kady Shea, and the Severance Community. This thesis would have suffered greatly without your friendship. Kanika, I cannot begin to express my gratitude for your support throughout this process. Thank you for knowing what I needed, even when I did not.

Finally, Mom—thank you. This is for you.

Abstract

After the election of President Barack Obama in 2008, many have begun to doubt the continued relevance of candidate race in voter behavior. This study, however, hypothesizes that in low information elections, where candidate's policy preferences are not widely known, voters will discriminate against candidates they perceive as racial or ethnic minorities. This thesis finds evidence of racial and ethnic bias in voter behavior in statewide low profile elections in Texas between 1992 and 2010. Expected candidate race and ethnicity is assigned using U.S. Census Genealogical data and linear regression analysis will be used to measure the change in vote share for the Democratic Party when the candidate race and ethnicity varies.

Table of Contents

I.   Introduction

In 1982, Los Angeles Mayor and black Democrat, Tom Bradley ran for Governor of California against State Attorney General and white Republican, George Deukmejian. In the days leading up to the election, tracking polls reported that Bradley had a comfortable lead over Deukmejian. Exit polls on Election Day confirmed these results and projected a Bradley win. The vote count, however, registered a Deukmejian victory by a margin of less than one percent, or fewer than 50,000 votes (Lindsey, 1982). Bradley's surprising defeat begs the question: what drove the discrepancy between what individuals told pollsters, and their actions behind the curtain at their polling place? Researchers have speculated that voters' responses were impacted by social desirability bias. In other words, survey respondents wanted to appear more willing to vote for a black candidate than they actually were. The "Bradley Effect" has reared its ugly head many times since then, and suggests that voters do respond to the race or ethnicity of candidates for public office.

Just seven years later, the "Bradley Effect" struck another gubernatorial race. In 1989, the Virginia governor's race between Douglas Wilder, a black Democrat, and Marshall Coleman, a white Republican, was decided by just 6,700 votes (Hopkins, 2009). Although Wilder prevailed to become the first African American governor, his razor thin margin of victory is completely at odds with the 15 percentage point lead reported by opinion polls just two weeks before Election Day (Hopkins, 2009).

After the election of Barack Obama in 2008, however, many have begun to doubt the continued relevance of candidate race in voter behavior. In 2009, Harvard's Daniel Hopkins analyzed the "Bradley Effect" in 180 gubernatorial and senatorial elections between 1989 and 2006. Hopkins argues that while the "Bradley Effect" was very salient in the late 1980s and early

1990s, it had all but evaporated by the late 1990s. The dissipation of the gap between polling figures and election performance, however, does not necessarily indicate that candidate race is no longer a salient factor to voters. Voters may have become more comfortable expressing their true political preferences to pollsters, or the campaigns of the late 1990s and 2000s may have been significantly less racialized than those in the late 1980s. Neither of these potential scenarios mandates that voters have become more willing to vote for minority candidates.

While it does not have the last word on racial bias in voting behavior, Hopkins' study does offer an important clue: emphasis on high profile elections may make it difficult to empirically identify racial bias. While congressional and gubernatorial elections are important, they comprise only a small fraction of candidates for public office in the United States. Gubernatorial and congressional elections have distinct characteristics that are not necessarily representative of elections in general. The high-profile nature of the race for the governor's mansion or the U.S. Congress invites a higher level of scrutiny from the media. Given this heightened state of attention to high profile elections in the media, voters in these elections are more likely to know the candidate's race and ethnicity, but they are also more likely to know about the candidates' policy preferences. For example, most voters in the 2008 presidential election were aware that President Obama is black, but were also likely somewhat aware of his policy priorities, like health care reform. When a candidate's platform is so thoroughly documented in high profile elections, it becomes increasingly difficult to attribute differences in vote share to the race and ethnicity of a candidate and not to voter's policy preferences. Thus, high profile elections may conceal racial and ethnic bias in voter behavior. Candidates for public office in lower profile elections, however, generally do not benefit or suffer from large-scale media exposure or public interest. Typical voters likely do not know the platforms of candidates for these lower profile public offices, like

Comptroller of Public accounts, or Railroad Commissioner. In these lower profile elections then, it is more likely that heuristics like candidate race and ethnicity matter more explicitly to voters.

This thesis focuses on quantifying racial bias in voter behavior in statewide low profile elections. Candidate race and ethnicity is assigned using U.S. Census Genealogical data and linear regression analysis is used to measure the change in vote share for the Democratic Party when the candidate race and ethnicity varies.

The following section presents background information that is relevant to the content and structure of this study, which attempts to isolate and quantify racial bias in voter behavior using the universe of election results in Texas between 1992 and 2010. Section three includes a review of the literature on racial bias in voting behavior. The fourth section describes the makeup of the data. Section five offers descriptive statistics. Section six includes a discussion of the empirical strategy. The seventh section presents the main results of the analysis. Section eight includes discussion of secondary analysis that attempts to rule out policy based differences as a driving factor in differential vote share by race and ethnicity. Section nine concludes.

II.   Background

For the purpose of this study, "high information" elections are defined as presidential, gubernatorial, and U.S. House of Representatives, and U.S. Senate elections. These elections are high profile and command significant attention in mainstream media. Relevant elections for this study include: the presidential elections of 1992, 1996, 2000, 2004, and 2008; Texas gubernatorial elections in 1994, 1998, 2002, 2006, and 2010; elections for Texas' U.S. senators in 1994, 1996, 2000, 2002, 2006, and 2008; and House of Representatives elections every other year beginning in 1992.

The remaining public offices are defined as "low information" elections: attorney general, state treasurer, railroad commissioner, comptroller of public accounts, state senator, state representative, commissioner of the general land office, commissioner of agriculture, district attorney, criminal district attorney, district judge, family district judge, criminal district judge, court of criminal appeals judge or presiding judge, court of appeals judge or presiding judge, supreme court justice or chief justice, and member of the state board of education. These elections garner much less attention in the mainstream media and, therefore, are categorized separately from the high information elections.

Another important distinction among elections is between local and statewide elections. An election is considered local if the office's jurisdiction is a district that encompasses less than the entire state of Texas. The following elections are considered local: U.S. Representative, state representative, state senator, state board of education member, district attorney, criminal district attorney, district judge, family district judge, and criminal district judge. The following elections are considered statewide: U.S. senator, governor, attorney general, state treasurer, railroad commissioner, comptroller of public accounts, commissioner of the general land office, commissioner of agriculture, court of criminal appeals judge or presiding judge, and supreme court justice or chief justice.

III. Prior Literature

The literature surrounding racial bias and voting behavior provides evidence to suggest that candidate race does affect voter behavior. For example, Ebonya Washington, in her 2006 study, "How Black Candidates Affect Voter Turnout," examines congressional and gubernatorial elections between 1982 and 2000 for all states. She finds that the appearance of a black candidate on the ballot increased turnout among black voters by 2.3 percentage points and among white

voters by 2.2 percentage points. Washington also finds that both white Democrats and Republicans are less likely to vote for major party candidates when the candidates are black. Thus, the voter turnout effects are likely to disadvantage black candidates for public office because the average voting district is majority-white. These results directly contradict the notion that candidate race is no longer salient to voters, even in the high profile elections featured in Washington's study.

Seth Stephens-Davidowitz's 2013 study, "The Cost of Racial Animus on a Black Presidential Candidate," explores voters' racial bias in the 2008 and 2012 presidential elections. Stephens-Davidowitz uses Google search data to estimate the racial animus of a media market as a way to mitigate the social desirability bias of survey respondents. Using the frequency of searches that include racial epithets, Stephens-Davidowitz constructs a proxy for a local media market's racial animus. While controlling for a district's Democratic vote share in the 2004 (Kerry v. Bush) presidential election, he finds that this proxy for racial animus is a strong negative predictor of Barack Obama's vote share in both of his presidential elections. In fact, in both 2008 and 2012 Stephens-Davidowitz (2013) estimates that President Obama lost roughly 4 percentage points due to racial animus, giving his opponents the equivalent of a "home state advantage" nationwide. Clearly then, as recently as 2012, candidate race played a significant role in voter behavior, even in high profile elections.

While Washington's and Stephens-Davidowitz's studies offer evidence that candidate race affects voter behavior, neither considers low information elections where race and ethnicity may play a larger role in voter choices. These studies focus exclusively on high profile elections that are likely to receive large amounts of local and national media attention: presidential, gubernatorial, and congressional elections. In these elections it is extremely difficult to isolate racial bias in voting behavior, as there may be unobserved candidate-specific factors besides race

and ethnicity that are driving voter behavior. Because these elections are so well-publicized and information on a candidate's positions is more readily accessible, unless there is a well-established measure of racial bias like the one constructed by Stephens-Davidowitz, the assumption that voters' decisions are based solely on the candidate's race is weak. In low information elections, like railroad commissioner or district judge elections, however, the assumption that voters know little about candidates' ideologies is much stronger.

Monika McDermott examines the role of race and gender in low profile elections in her 1998 article, "Race and Gender Cues in Low-Information Elections." McDermott cites an extensive literature on the psychology of voter behavior that claims voters use political and social stereotypes to evaluate candidates in low information settings. In other words, in low information elections, voters may be more prone to rely on a candidate's partisan affiliation, and their gender or race and the associated stereotypes. Using quasi-experimental data from an LA Times survey, McDermott found that voters stereotype both black and female candidates as being more liberal than their white and male counterparts. These perceptions may provide rationalization for individuals who are unwilling to vote for candidates based on their race or gender. This study considerably improves upon McDermott's design, however, as it offers real-world evidence of racial bias in voting behavior, and is not be limited to hypothetical candidates for public office. This study also improves upon McDermott's design by investigating racial bias in voter behavior in non-urban, as well as urban areas. McDermott's study was limited to large metropolitan areas, which have lower levels of predicted racial prejudice.

In addition to its sociological and political nature, this thesis also serves as a specific case study of labor market discrimination. The labor market for public office is important, as elected officials control the distribution of governmental resources, enforce our laws, and ensure the equal

protection of all citizens. Adequate minority representation in public office is then important for the fair allocation of resources and protection for all citizens. Voter bias against minority candidates may then have much farther-reaching consequences than a lost election for an individual candidate. This thesis will contribute to the labor market discrimination literature in an important way, following in the footsteps of Bertrand and Mullainathan's study "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." In this study, Bertrand and Mullainathan randomly assigned white-sounding and African-American-sounding names to identical resumes, responded to help wanted ads in Boston and Chicago newspapers with the resumes, and measured callback rates for interviews. The authors found that white-sounding names receive 50 percent more callbacks for interviews. While Bertrand and Mullainathan's resume audit measured callback rates for interviews after an initial application, in the specialized labor market for public office in Texas, the measure of analysis is instead the differential vote shares for racial and ethnic minority candidates relative to white candidates.

This thesis will attempt to isolate and quantify racial bias in voter behavior using the universe of election results in Texas between 1992 and 2010. Racial bias, however, is difficult to quantify because the full set of information that voters consider before casting their ballots is unavailable. In high information elections, like those studied by Washington and Stephens-Davidowitz, racial bias is extremely difficult to isolate, as voters may have information about candidates' ideologies in addition to information about their race or ethnicity. Therefore, it is possible that voters spurn party because of politics and not prejudice. In low information elections, however, it will be assumed that voters have less information about the politics of candidates on the ballot. Thus voters rely more heavily on the political party, gender, and race or ethnicity of

the candidate rather than actual knowledge of the candidate's political ideology to make their selections. The results of McDermott's study suggest that this assumption is likely valid.

This thesis will follow in the footsteps of Bertrand and Mullainathan's audit study, and will capitalize on the critical difference in voter information between high and low information elections in an attempt to isolate racial bias in voting behavior.

IV.   Data

The data used for this analysis includes Texas election results from 1992 to 2010 (Office of the Texas Secretary of State, 2015), U.S. Census Genealogical data from 2000 (United States Census Bureau, 2014), voting district level demographic information from 1992 to 2010 (Minnesota Population Center, 2011), employment statistics for the oil and gas industries from the Census Bureau from 1998 to 2010 (United States Census Bureau, 2016), and Google Trends data on several candidates for public office from 2004 to the present (Google Trends, 2016).

The universe of election results for public office in Texas between 1992 and 2010 are publicly available from the Texas Secretary of State's website. The election results, which run from 1992 to 2010, are reported at the candidate-county-election-year level. In other words, vote counts are recorded for each candidate, in each election year, and in each county included in the voting district for their desired public office. For example, Robert Newsom ran for district judge of the 8th judicial district in 2008. Judicial district 8 includes Rains, Delta, and Franklin counties. For the 2008 primary election, Newsom has three observations in the dataset, one for each county in judicial district 8 (Rains, Delta, and Franklin). The breakdown of votes at this level will allow for the observation of heterogeneous voting patterns across different types of counties in statewide or large district elections, as well as enable aggregation to election wide totals. In

addition to vote totals, the election results dataset also includes each candidate's political party and incumbency status.

The genealogical data offers the racial and ethnic breakdown of surnames with an occurrence of more than 100 people in the 2000 Census. For example, in 2000, 92.7 percent of all the individuals with the surname "Rodriguez" self-identified as Hispanic. This dataset was merged on to the election results, matching each candidate to the genealogical record associated with their surname. Ninety-five percent of the candidates were matched to genealogical records, leaving only five percent unmatched. While this method will not reveal each candidate's true race or ethnicity, it does serve as a mechanism for assigning an expected race or ethnicity for each candidate, which should roughly reflect the level of information about a candidate's race or ethnicity held by voters in low information elections. Hereafter, when a candidate's race or ethnicity is referenced, it is understood that their true race and ethnicity is unknown; the study only assigns an expected race or ethnicity based on the method described below.

Individuals with surnames that have been identified as over 80 percent black, Hispanic, or Asian are assigned to that expected race or ethnicity. If, as it has been established above, 92.7 percent of all persons in the United States with the surname "Rodriguez" self-identify as Hispanic, it is plausible that voters would perceive a candidate with the surname to be Hispanic as well. Thus, candidates with the surname "Rodriguez" are assigned to the expected ethnicity of Hispanic. The default race is white, so all candidates whose surnames do not register above the 80 percent cutoff for black, Hispanic, or Asian are considered white. The relatively high cutoff of 80 percent for black, Hispanic, and Asian categories was chosen to ensure that the expected race and ethnicity of the candidates in the dataset aligned as closely as possible with the expectations of voters in low information elections.

There are 5,992 unique candidates for public office in Texas between 1992 and 2010. Based on their assignment using the Census genealogical records and the 80 percent cutoff, 5,257 (87.7 percent) are white, 6 (0.1 percent) are black, 706 (11.8 percent) are Hispanic, 20 (0.3 percent) are Asian, and 3 (0.05 percent) were not matched to any race or ethnicity. Thus, most of the variation in candidate race and ethnicity occurs between white and Hispanic candidates.

Demographic variables were obtained at the census tract level in 1990, 2000, and 2010 from the National Historical Geographic Information System (NHGIS) at the Minnesota Population Center. These variables include information on the population, urban/rural status, gender composition, age composition, racial composition, Hispanic or Latino origin, educational attainment, labor force and employment status, median household income, per capita income, and poverty rates. Once aggregated to the county level, these demographic variables were linearly interpolated for the intervening years between Census measurements.

There are six types of voting district boundaries in Texas: U.S congressional districts, state legislative districts (both upper and lower chamber), judicial districts, courts of appeals districts, and state board of education districts. Over the 18 year time period covered in this study there were numerous redistricting events in all but the judicial districts and courts of appeals districts. The boundaries and numbers of census tracts were updated in 2000 and 2010.

Given the fact that census tracts changed and were renamed with each census, it was impossible to interpolate values accurately between decades. Therefore, 1990 figures were used for the years 1992-1999, 2000 figures were used for the years 2000-2009, and 2010 figures were used for only 2010. Using census tract and district number crosswalk data from the Missouri Census Data Center, the demographic variables have been aggregated to the district level for each

redistricting event for the following district types: U.S. congressional districts, state legislative districts, judicial districts, and courts of appeals districts.[1]

Data on oil and gas employment by county comes from the U.S. Census' annual County and Business Patterns (CBP) series. The CBP series provides the number of employees in the oil and gas industries by county in the following years: 1998, 2000, 2002, 2004, 2006, 2008, and 2010. This measure is used as a rough approximation for how important the oil and gas industry is in an individual county in a given year. The greater the number of employees in the oil and gas industry in a county, the more deeply invested voters should be in the outcome of the Railroad Commissioner election.[2]

Google Trend data was obtained on the following individuals from 2004-the present: John Cornyn, Michael L. Williams, Victor G. Carrillo, Elizabeth Ames Jones, and Greg Abbott. Cornyn is a U.S. Senator for the state of Texas, while Williams, Carillo, and Jones all served as Railroad Commissioners. Greg Abbott is the current Governor of Texas and the former Attorney General. The rate of searches for these individuals signals the vast differences in demand for information about candidates between low and high information elections.

V.   Descriptive Statistics

Summary statistics for the racial and ethnic variation in candidate surname are shown in Figure 1 below. The four graphs associated with Figure 1 describe the overall expected racial and ethnic makeup of the universe of candidates for public office in Texas between 1992 and 2010 based on U.S. Census Genealogical records, as described above. On the X axis of each graph is

---

[1] Demographic variables were not aggregated to the district level for state board of education districts, because data was unavailable to match census tracts to state board of education districts.
[2] The Railroad Commission in Texas is the statewide regulatory agency for the oil and gas, natural gas and extraction industries. The three commissioners are elected to six year terms, with one commissioner standing for election every two years.

the proportion of individuals in the 2000 U.S. Census identifying with the given race or ethnicity given their last name, while the Y axis shows the density of occurrences in the data.

The graph on the top left demonstrates that the majority of candidates have last names that in the 2000 Census were associated with whites at a rate of 60 percent or higher. It also shows that there is a sizable number of candidates with last names that were associated with whites at much lower rates (less than 20 percent white). The graph on the top right reveals that almost no candidates had last names that were associated very strongly with blacks; only six candidates had last names that met the 80 percent cutoff. The graph on the bottom left indicates that there is a modest group of candidates whose last names exceed the 80 percent cutoff. Those candidates whose surnames exceed the 80 percent cutoff are "distinctively Hispanic" and will be assigned to the expected ethnicity of Hispanic. The graph on the bottom right tells much the same story as the top right. Only a very small number of candidates exceed the 80 percent cutoff for Asian. Thus, Figure 1 confirms that the primary variation in candidate race and ethnicity is between white candidates and Hispanic candidates.

Before proceeding with further analysis, however, it is important to verify that this study's assignment of expected candidate race and ethnicity appropriately mimics the expectations of voters in low information elections. Table 1 offers a look at specific candidate surnames as an indicator of race and ethnicity. Each column lists the last names with the highest proportion of that race or ethnicity, weighted by the proportion of people per 100,000 in the United States with that last name in the 2000 Census. This table confirms that the method of racial and ethnic assignment described above reasonably approximates the assumptions that average voters make about candidate race and ethnicity based on the low information setting of statewide low information elections. For example, it is reasonable to assume that the typical voter will associate

the surname "Garcia" (the first ranked Hispanic last name) with a Hispanic candidate, unless they have information to suggest otherwise. Voters are likely to assume candidates with the surname "Smith" (the first ranked white last name) are white, unless they have additional information about the candidate. In low information elections, however, this information is less readily available; this study's assignment of expected candidate race and ethnicity is reasonable for at least for majority white, Hispanic, and Asian surnames.[3]

Now that variation in candidate race has been established, it is prudent to ensure that this variation occurs across election types. If Hispanic candidates choose to run for only high information elections for example, there will be insufficient variation in candidate race and ethnicity in low information elections to analyze differential vote shares. Table 2 offers a look at the distribution of candidates across different types of elections. In high information elections, racial and ethnic variation is restricted almost exclusively to candidates for U.S. Representative. One hundred and twenty four Hispanic candidates ran for U.S. Representative, while only four Hispanic candidates ran for all other high information positions. This pattern suggests that Hispanic candidates may be choosing to run in districts that favor racial or ethnic minority candidates, as U.S. Representative is the only high information position with localized as opposed to statewide districts. Because smaller districts present the potential for selection effects for the candidates choosing to run, and information about candidate positions is assumed to be more readily available, it will be instructive to exclude these high information and local elections from analysis in the primary empirical strategy detailed below.

---

[3] Table 1 demonstrates that this study's methodology, which focuses on last names and not first names, does not convincingly assign race to black candidates. Bertrand and Mullainathan's audit study used variation in first names to signal for (fictional) black job-seekers. This study, however, did not have access to first name genealogy records, and so will focus on variation between candidates with white and Hispanic sounding last names.

Local low information elections offer fairly large numbers of Hispanic candidates, but these elections present the same challenges as candidates for U.S. Representative. Because elections are local, candidates may be aware of a racial or ethnic penalty from voters, and thus choose to run in districts where they believe they have a better chance of winning. Similarly to the case of the U.S. Representative races describe above, this selection strategy will make it difficult to observe racial and ethnic bias in voter behavior, despite the fact that in low information elections, information about a candidate is less readily available than in high information elections.

Statewide low information elections, however, offer some variation in candidate ethnicity and are voted for by all counties. Statewide low information elections solve the problem of selection effects in local elections like U.S. Representative and other low information local elections, and solve the problem of information availability. Because they are low information elections, voters will know less about these candidates' policy preferences and thus, will rely more heavily on heuristics like candidate race and ethnicity. Statewide low-information elections will be the primary focus of the analysis.

As statewide low information elections are the focus of this analysis, it is important to consider the possibility that voters only vote for candidates about whom they have knowledge. If this were true, we would expect to see much higher vote totals for presidential and congressional elections than for low information elections such as railroad commissioner or comptroller of public accounts. Figure 2A, however, challenges this expectation. Figure 2A shows vote totals across four election types in general elections. While there is some roll off in vote totals from high information to low information elections, these drops are usually relatively small, indicating that there are very likely uninformed voters voting for candidates in low information elections. This

further confirms that statewide low information elections should be the focus of this study's analysis.

Another important consideration is the distribution of variation in candidate race and ethnicity by political party. Table 3 shows the breakdown of race and ethnicity of candidates for the Democratic and Republican parties, as well as other parties. The overwhelming majority of Hispanic candidates for public office are Democrats (722, or 77 percent), while relatively few are Republican (167, or 18 percent).

Another assumption that must be addressed before detailing the empirical strategy is the assumption that average voters are less informed about candidates for low information elections than for high information elections. Figure 3A shows Google Trends data for four candidates from 2004 to 2014, John Cornyn, a U.S. Senator for the state of Texas, and three railroad commissioners: Victor G. Carrillo, Michael L. Williams, and Elizabeth Ames Jones.[4] Figure 3A clearly demonstrates that candidates for high information elections are searched more by the general public than candidates for low information elections.[5]

Figure 3B provides another clear example of the information gap between low and high information candidates. Figure 3B depicts the change in Greg Abbott's search rate from when he served as attorney general, a statewide low information office, and when he ran for governor of Texas, a high information office. Prior to the announcement of his campaign for governor in July of 2013, Abbott consistently received relatively few searches. After his announcement and successful gubernatorial campaign, his search rate spiked. This figure demonstrates that even with

---

[4] The Y-Axis for Figures 3A and 3B can be interpreted as the search interest relative to the highest point on the chart. The highest point on the chart is 100, indicating that at that time 10 percent of searches in Texas were for that particular candidate.

[5] The public officials selected for Figures 3A and 3B were selected to understate the differences in search rates between candidates for low information public offices, and candidates for high information public offices. Had more well-known public officials, like Governor George Bush or Senator Ted Cruz, been selected as examples, the differences in search rates would have been more extreme.

the same candidate, the average voter's knowledge varies widely depending on whether the election is low information or high information.

VI.   Empirical Strategy

The model for this study takes the following general specification:

$$Dem\_vs_{cet} = \beta_0 + \beta_1 Dem*Hisp_{et} + \beta_2 Rep*Hisp_{et} + \beta_3*Dem*Inc_{et} + \beta_4 Rep*Inc_{et} + \gamma X_{ct} + \gamma_e$$

$$+ \eta_t + \varepsilon_{cet}$$

where dem_vs$_{cet}$ is the percentage of the total votes awarded to the candidate for the Democratic Party in county c, in a given election type e, and election year t. Dem$_{et}$ and Rep$_{et}$ are dummy variables which signal whether a candidate is a Democrat or Republican, while Hisp$_{et}$ is a dummy variable for whether or not a candidate is Hispanic. β_0 indicates the expected vote share for the Democratic Party when both candidates are white. β_1 is the change in Democratic vote share when the Democratic candidate is Hispanic and the Republican candidate is white. β_2 is the change in Democratic vote share when the Republican candidate is Hispanic and the Democratic vote share is white. Inc$_{et}$ is a dummy variable that indicates the incumbency status of the candidate. X$_{ct}$ includes a set of county-level demographic controls. The model will be run with various combinations of the county-level demographic controls, and county fixed effects, to ensure the robustness of the regression results. The county-level demographic controls include: fraction Hispanic, fraction college-educated or more, unemployment rate, median household income (in 2010 dollars), fraction 65 years and older, and population (in 1000s). All regressions in the base model will include year fixed effects (η_t) and office fixed effects (γ_e). Standard errors are clustered at the office-year level.

The model above will be used to test the difference in Democratic Party vote share when Hispanic candidates run for office in statewide low information elections. The model will also be

used to assess differential effects on Democratic vote share by election type (mid-term versus presidential), by restricting the sample first to midterm (nonpresidential) election years, and then to presidential election years. If voters are swayed by racial or ethnic prejudice when information about a candidate's policies are unavailable, it is expected that Hispanic candidates are penalized more heavily during presidential election years because of the relatively higher numbers of uninformed voters in these years.

The model will also be used to assess differential effects on Democratic Party vote share by various measures of each county's predicted racial prejudice. This study provides two measures of predicted racial prejudice, and will use both to present the results of the model above. The first measure of racial prejudice is borrowed from Seth Stephens-Davidowitz. Stephens-Davidowitz, in his 2013 paper cited below, uses Google Trends data on areas' racially charged search rates to construct predictors of racial prejudice. He found that fraction 65 and older, fraction college-educated, fraction Hispanic, fraction black, and fraction black squared are predictors of the area's racial prejudice (Stephens-Davidowitz, 2013). The coefficients from Stephens-Davidowitz's model are multiplied by their respective county-level demographic variable to create a measure of racial prejudice for each county in each year. Counties will be separated into four quartiles of predicted racial prejudice. The model described above will be used to assess differential effects on Democratic Party vote share in statewide low information elections for each quartile of racial prejudice. Again, if voters are indeed swayed by racial or ethnic prejudice in low information elections, it is reasonable to expect that counties with higher levels of predicted racial prejudice will penalize Hispanic candidates more heavily than white candidates.

The second measure of racial bias is the change in Democratic vote share in 2008 relative to the average Democratic vote share prior to 2008. In 2008, Barack Obama became the first black

president of the United States. While President Obama received enough votes to be elected President, the results of Stephens-Davidowitz's study indicate he lost roughly four percentage points overall due to racial animus (Stephens-Davidowitz, 2013). Given the fact that racial bias was a motivating factor for many voters in 2008 because of Mr. Obama's presence on the ballot, a rough measure of racial bias is the change in Democratic vote share for a given county prior to 2008 and the Democratic vote share for the same county in 2008. Using this method, counties that experienced a decline in Democratic vote share in 2008 relative to the mean Democratic vote share prior to 2008 will have relatively high levels of predicted racial prejudice. Conversely, counties that did not experience a decline in Democratic vote share in 2008 relative to the mean Democratic vote share prior to 2008 will have relatively low levels of predicted racial prejudice.

Finally, the model will be adapted to assess the difference in vote share for Hispanic and white candidates for Railroad Commissioner. In this special case, the model will be run for each of the four quartiles of counties by the per-capita number of oil and gas industry employees. Standard errors in this model will be only clustered at the year level, as opposed to the office-year level, because only one office is under consideration. This variation on the primary model will attempt to rule out policy difference of Hispanic and white candidates as a driver of difference in Democratic vote share when a Hispanic candidate of either major party runs against a white candidate. In Texas, the Railroad Commission regulates the oil and gas industries. If Hispanic candidates for Railroad Commissioner systematically favored different policies than white candidates for Railroad Commissioner, one could expect to find a pattern of voting behavior by the county's quartile of oil and gas employees. If there was a systematic difference in white and Hispanic policy, one might expect counties with a high stake invested in the oil and gas

industry (using the measure of county-level industry employment as a proxy) to vote differently than counties with little at stake.

The empirical strategy of this study is flexible. It allows for the analysis of racial and ethnic bias in voting behavior in statewide low information elections through various lenses: midterm v. presidential election, county predicted racial prejudice, and potential policy relevance.

VII.   Main Results

Given the apparent and vast differentials in knowledge about candidates between low and high information elections, this study's analysis will center on statewide low information elections. Table 4 reports the regression results from the model described in the empirical strategy section above. The model produces a stubborn result across eight specifications of county-level demographic controls: when the Democratic Party's candidate is Hispanic, democratic vote share declines by 3.3 percentage points, indicating a relatively large disadvantage for Hispanic Democrats. This result is consistent across each of the various additional models. The results are just as stable across different specifications for Hispanic Republicans. Table 4 reveals that when the Republican candidate is Hispanic, the Democratic vote share increases by 1.2 percentage points, disadvantaging Hispanic Republicans, although not to the same degree as Hispanic Democrats.

Figure 4 presents these results graphically. The blue curve in the middle represents the distribution of residual democratic vote share in elections contested by a white Republican and white Democrat. In elections with Hispanic Democrats running against white Republicans, however, the distribution of residual democratic vote share shifts to the left, as shown by the red curve, indicating a loss in Democratic Party vote share in response to Hispanic Democrats. Similarly, in elections with Hispanic Republicans running against white Democrats, the

distribution of residual Democratic Party vote share shifts to the right (as shown by the green curve) indicating a gain in Democratic Party vote share in response to the Hispanic Republican candidate. Figure 4 presents a striking visualization of the gap in vote share for both Hispanic Democrats and Republicans.

Although these initial results are quite stark, one cannot yet conclude that voters discriminate based on candidate race and ethnicity. If the effects are more dramatic in presidential years, this could be a sign that the differences are due to racial bias, and not policy differences, because in presidential years there are likely to be greater numbers of uninformed voters, who are likely more influenced by candidate heuristics as identifiable by their name. It will also be necessary to use the measures of predicted racial prejudice to further investigate whether differential Democratic vote share is linked to racial or ethnic bias.

Table 5 reveals that the effects are indeed larger in presidential years for both Hispanic Democrats and Republicans. In midterm years, when it was hypothesized that the typical voter is better informed, neither the coefficient on Democratic*Hispanic or Republican*Hispanic is significant. Both coefficients are significant, however, during presidential election years. In presidential election years, when a Hispanic Democrat is running against a white Republican, the Democratic candidate loses 5.1 percentage points. Similarly, in presidential election years, when a Hispanic Republican is running against a white Democrat, the Democratic Party gains 5.8 percentage points. These respective losses and gains in response to Hispanic candidates are substantial.

VIII.   Does Policy Preference or Racial and Ethnic Bias the Gap?

The main results described in the section above indicate that Hispanic candidates of either political party lose vote share to their white opponent. This fact by itself does not necessarily

implicate voters in racial or ethnic discrimination. Voters could be making choices based solely on policy differences between white and Hispanic candidates, and not simply the race and ethnicity of candidates. The results below, however, do little to support this idea. This thesis uses the two measurements of county-level predicted racial prejudice described in the empirical strategy section above, and a case study of the Railroad Commissioner position to argue that the gap in vote share between Hispanic and white candidates is driven by racial prejudice and not distinct policy preferences between white and Hispanic candidates.[6]

To further investigate the cause of the differential vote shares for white and Hispanic candidates, the study turns to two measures of predicted racial prejudice. Table 6A reports the results of the base model for each quartile of predicted racial prejudice, based on the Stephens-Davidowitz predicted racial prejudice index described in the empirical strategy section above.

Table 6A reveals that differential vote shares are correlated with predicted racial prejudice. Counties in the top quartile of predicted racial prejudice respond much more dramatically and negatively to Hispanic candidates than counties in lower quartiles. In top quartile counties, when a Hispanic Democrat runs against a white Republican, the Democratic Party loses 6.7 percentage points. In those same counties, when a Hispanic Republican runs against a white Democrat, the Democratic Party gains 7.8 percentage points. Both of these results are statistically significant. Lower quartiles also appear to discriminate against Hispanic candidates, although magnitudes and significance levels are lower.

---

[6] Analysis was also conducted to measure own-race bias. Table 7 reports the difference in Democratic vote share by quartiles of counties' fraction of Hispanic residents. Counties in the top quartile (highest fraction Hispanic) have the lowest levels of discrimination for Hispanic Democrats. When Hispanic Democrats run against white Republicans in the top quartile, the Democratic vote share only declines by 2.5 percentage points, while in counties in the bottom three quartiles, the Democratic Party loses between 5.9 and 6 percentage points. Results for Hispanic Republicans are less stark. In the top quartile, the Democratic vote share increases by 4.2 percentage points when a Hispanic Republican runs against a white Democrat. In the bottom quartile, the Democratic vote share increases by 6.8 percentage points when a Hispanic Republican runs against a white Democrat.

Tables 6B tells much the same story: it uses the change in average Democratic vote share in 2008 relative to the Democratic vote share prior to 2008 as a measure of predicted racial prejudice. This strategy reveals that voters in counties with higher levels of predicted racial prejudice discriminate against Hispanic candidates when they run against white opponents. In other words, in the counties that the Democratic Party experienced large drop off in vote share in 2008, Hispanic candidates of both the Democratic and Republican parties lose greater vote shares.

Table 8 provides the results of analysis on candidates for Railroad Commissioner, as well as the number of employees employed in the oil and gas industry by county. No pattern can be clearly identified from these results, and indeed only two of the coefficients are even weakly statistically significant. This lack of a coherent story suggests that policy differences are not driving the differential vote shares for white and Hispanic candidates in elections for Railroad Commissioners. If policy differences between white and Hispanic candidates could explain the gap in vote share between white and Hispanic candidates for Railroad Commissioner, one could expect to see a pattern between voter behavior and the number of residents employed by the oil and gas industry. Counties with higher numbers of oil and gas employees have a greater policy interest in the Railroad Commission, as they have many jobs at stake. If either white candidates or Hispanic candidates routinely had more palatable policy preferences for oil and gas industry, then there should be a relationship between the quartile of oil and gas employees by county and the white-Hispanic vote share gap. This relationship does not appear to exist, however, suggesting that the gap between white and Hispanic candidates is not driven by policy differences. This supports this study's hypothesis that differences in vote share are indeed caused by racial and ethnic bias in voter behavior.

IX.   Conclusion

This study has found strong evidence of racial and ethnic bias in voter behavior in low information Texas elections. In counties of highest predicted racial prejudice, Hispanic Democrats lose 6.7 to 6.8 percentage points in vote share, while Hispanic Republicans lose 7.7 to 7.8 percentage points in vote share (Tables 6A and 6B). The consequences of racial and ethnic bias in voter behavior are difficult to fully assess, but are likely enormous. If minority candidates struggle to get elected to public office because of their race or ethnicity, then ethnic or racial minority communities will be inadequately represented in state and federal legislatures, and other state agencies. Poor representation could contribute to worse policy outcomes for these groups.

While this study has unearthed significant and substantial evidence of voter discrimination against Hispanic candidates in Texas, it is important to consider the generalizability of these results. The particular set of candidate surnames and genealogical records in this study limited analysis to white and Hispanic candidates. If discrimination exists against Hispanic candidates, it would be difficult to believe that voters do not discriminate against black, Asian, and other minority candidates. Further research should set about to investigate how other minority communities and women are affected by voter bias. For this research to be possible, first name genealogical data must be obtained, for it will help to identify women candidates, as well as distinctively black first names as done in Bertrand and Mullainathan's resume audit. Further research should also be conducted using low information elections across a broad swath of states. While it is unlikely that racial and ethnic bias in voter behavior is strictly limited to Texas, the state has a unique history and character that may raise questions about generalizability to the rest of the United States.

The results summarized above provide strong evidence that voters in low information elections discriminate against Hispanic candidates, and that this discrimination is tied to racial or

ethnic bias. These results represent a substantial contribution to the literature surrounding racial and ethnic bias in voter behavior. The focus on low information elections, rather than the high information elections that have been the focus of much of the previous literature, as well as the analysis undertaken using measures of predicted racial prejudice and policy interest, provide compelling evidence that voters exhibit racial and ethnic bias when casting their ballots on Election Day.

Table 1: Candidate Surnames as a Signal of Race and Ethnicity

Rank	White	Black	Hispanic	Asian	Unmatched
1	Smith	Washington	Garcia	Nguyen	Greytok
2	Johnson	Muldrow	Rodriguez	Tran	Hinckson
3	Miller	Grays	Martinez	Chen	Kohlhausen
4	Brown	Beckles	Hernandez	Wong	Magnis
5	Harper	Winkfield	Lopez	Le	Worldpeace
6	Jones	Amadi	Gonzalez	Liu	Yokie
7	Williams		Perez	Vu	Nuchia
8	Davis		Sanchez	Cheng	Malazzo
9	Anderson		Ramirez	Vo	Naishtat
10	Wilson		Torres	Hoang	Touzel
11	Martin		Flores	Chow	Cranberg
12	Taylor		Rivera	Yoo	Sarpalius
13	Moore		Gomez	Yao	Arashvand
14	Thompson		Diaz	Yau	Dorrycott
15	White		Reyes	Hsiao	Morovich
16	Clark		Morales	Sinha	Berriozabal
17	Thomas		Cruz	Chae	Alverez
18	Hall		Ortiz	Sakai	Fastuca
19	Holm		Gutierrez	Mahajan	Markantonis
20	Baker		Chavez	Shinoda	Deotte

Note: Census Genealogy records show both 1) the frequency of a given surname, and 2) the fraction of persons who belong to different racial or ethnic groups for a given surname. A candidate is categorized as black, Hispanic, or Asian if more than 80 percent of persons in the United States with her name belong to one of these racial or ethnic groups. All other surnames are categorized as white. Among the surnames that are not matched to the Census Genealogy records, a random subset is shown in the column labeled as Unmatched.

Table 2: Counts of Candidate Racial and Ethnic Groups by Election Type

High Information Elections	White	Black	Hispanic	Asian	Fraction Hispanic
President	47				0
U.S. Senator	43		3		0.065
U.S. Representative	834	1	124	3	0.129
Governor	22		1	1	0.042
Statewide "Low" Information Elections					
Attorney General	14		2		0.125
Lieutenant Governor	12		4		0.25
State Treasurer	2				0
Railroad Commissioner	27		7		0.206
Comptroller of Public Accounts	13		2		0.133
Commissioner of the General Land Office	12		3	1	0.188
Commissioner of Agriculture	17				0
Court of Criminal Appeals Presiding Judge	4		1		0.2
Court of Criminal Appeals Judge	56		4		0.067
Supreme Court Chief Justice	12				0
Supreme Court Justice	67		7		0.095
Local "Low" Information Elections					
State Senator	381		39		0.093
State Representative	2266	2	311	11	0.12
District Attorney	431		44		0.093
Criminal District Attorney	268		10		0.036
District Judge	2098	1	282	4	0.118
Family District Judge	199		22	1	0.099
Court of Appeals Chief Justice	66		7		0.096
Court of Appeals Justice	319		38	2	0.106
Member, State Board of Education	144		22		0.133
Criminal District Judge	70		5	1	0.066

Note: Race and ethnicity is assigned using Census Genealogy records. If more than 80 percent of individuals in the 2000 Census are black, Hispanic or Asian, then the candidate is assigned to that group; the remainder are assigned to be white. These are counts of racial and ethnic groups across all general election-by-year cells. Thus, some are counted multiple times.

Table 3: Political Party by Race and Ethnicity

Political Party	White	Black	Hispanic	Asian
Democrat	2,860	3	722	12
Republican	3,588	1	167	8
Other	976	0	49	4

Note: Race and ethnicity is assigned using Census Genealogy records. If more than 80 percent of individuals in the 2000 Census are black, Hispanic or Asian, then the candidate is assigned to that group; the remainder are assigned to be white. These are counts of candidate racial and ethnic groups across all general election-by-year cells. Thus, some are counted multiple times.

Dep Var: Democratic Candidate Vote Share (Mean Democratic Vote Share = .424)

Democratic*Hispanic	-0.033***	-0.033***	-0.033***	-0.033***	-0.033***	-0.033***	-0.033***	-0.033***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Republican*Hispanic	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Controls:								
Incumbent	Y	Y	Y	Y	Y	Y	Y	Y
Election Type Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Fraction Hispanic	N	Y	Y	Y	Y	Y	Y	N
Fraction College or More	N	N	Y	Y	Y	Y	Y	N
Unemployment Rate	N	N	N	Y	Y	Y	Y	N
Median Household Income (in 2010 $)	N	N	N	N	Y	Y	Y	N
Fraction 65+	N	N	N	N	N	Y	Y	N
Population (in 1000s)	N	N	N	N	N	N	Y	N
County Fixed Effects	N	N	N	N	N	N	N	Y
R-Squared	0.289	0.456	0.501	0.569	0.587	0.588	0.596	0.860

Note: N = 20,065. These regressions restrict the sample to statewide low information elections, which include elections for Attorney General, Lieutenant Governor, State Treasurer, Railroad Commissioner, Comptroller of Public Accounts, Commissioner of the General Land Office, and Commissioner of Agriculture. Both midterm and presidential elections are included. Standard errors in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 5: Differential Effects by Midterm Election

Dep Var: Democratic Candidate Vote Share

	Midterm Year	Presidential Year
Democratic*Hispanic	-0.020	-0.051***
	(0.014)	(0.011)
Republican*Hispanic	-0.004	0.058**
	(0.011)	(0.023)
Mean Democratic Vote Share	0.415	0.443
Observations	13,716	6,349
R-squared	0.607	0.597

Note: These regressions restrict the sample to statewide low information elections, which include elections for Attorney General, Lieutenant Governor, State Treasurer, Railroad Commissioner, Comptroller of Public Accounts, Commissioner of the General Land Office, and Commissioner of Agriculture. Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1

Table 6A: Effects by Predicted Racial Prejudice in Presidential Election Years

Dep Var: Democratic Candidate Vote Share

Panel A: Google Search Proxy

	Quartile of Predicted Racial Prejudice			
	Bottom	2nd	3rd	Top
Democratic*Hispanic	-0.019**	-0.055***	-0.063***	-0.067***
	(0.007)	(0.012)	(0.010)	(0.009)
Republican*Hispanic	0.016	0.073**	0.064**	0.078***
	(0.017)	(0.025)	(0.023)	(0.023)
Mean Democratic Vote Share	0.495	0.374	0.403	0.493
Observations	1,507	1,492	1,567	1,783
R-squared	0.796	0.552	0.597	0.608

Note: These regressions exclude counties in the top quartile of the fraction Hispanic distribution. The predicted racial prejudice is based on predictions from Stephens-Davidowitz. The sample includes statewide low information elections, e.g., Attorney General, Lieutenant Governor, State Treasurer, Railroad Commissioner, Comptroller of Public Accounts, Commissioner of General Land Office, and Commissioner of Agriculture. County level demographic data is constructed using NHGIS data. Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1

Table 6B: Effects by Predicted Racial Prejudice in Presidential Election Years

Dep Var: Democratic Candidate Vote Share

Panel B: Change in Democratic Vote Share in '08

	Quartile of Predicted Racial Prejudice			
	Bottom	2nd	3rd	Top
Democratic*Hispanic	-0.028***	-0.049***	-0.059***	-0.068***
	(0.010)	(0.013)	(0.011)	(0.011)
Republican*Hispanic	0.027	0.063**	0.066***	0.077***
	(0.019)	(0.024)	(0.022)	(0.026)
Mean Democratic Vote Share	0.47	0.407	0.42	0.475
Observations	1,575	1,600	1,575	1,599
R-squared	0.791	0.657	0.721	0.723

Note: These regressions exclude counties in the top quartile of the fraction Hispanic distribution. The predicted racial prejudice is based on predictions from Stephens-Davidowitz. The sample includes statewide low information elections, e.g., Attorney General, Lieutenant Governor, State Treasurer, Railroad Commissioner, Comptroller of Public Accounts, Commissioner of General Land Office, and Commissioner of Agriculture. County level demographic data is constructed using NHGIS data. Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1

Table 7: Own-Race Bias in Presidential Election Years

Dep Var: Democratic Candidate Vote Share
(Mean Democratic Vote Share = .424)

	Quartile of Fraction Hispanic			
	Bottom	2nd	3rd	Top
Democratic*Hispanic	-0.059***	-0.059***	-0.060***	-0.025**
	(0.008)	(0.012)	(0.013)	(0.010)
Republican*Hispanic	0.068***	0.061**	0.064**	0.042*
	(0.022)	(0.024)	(0.024)	(0.020)
Observations	1,730	1,531	1,564	1,524
R-squared	0.770	0.591	0.433	0.790

Note: These regressions restrict the sample to statewide low information elections, which include elections for Attorney General, Lieutenant Governor, State Treasurer, Railroad Commissioner, Comptroller of Public Accounts, Commissioner of General Land Office, and Commissioner of Agriculture. County level demographic data is constructed using NHGIS data. Standard errors in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 8: Elections for Railroad Commissioner

Dep Var: Democratic Candidate Vote Share

	Quartile of Per Capita Energy Employees			
	Bottom	2nd	3rd	Top
Democratic*Hispanic	-0.070*	-0.054*	-0.058	-0.069
	(0.031)	(0.026)	(0.038)	(0.039)
Republican*Hispanic	0.048	0.038	0.051	0.051
	(0.042)	(0.035)	(0.044)	(0.048)
Observations	576	567	576	567
R-squared	0.533	0.588	0.548	0.593

Note: These regressions focus only on elections for the Railroad Commissioner. Both midterm and presidential election years are included. Standard errors in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

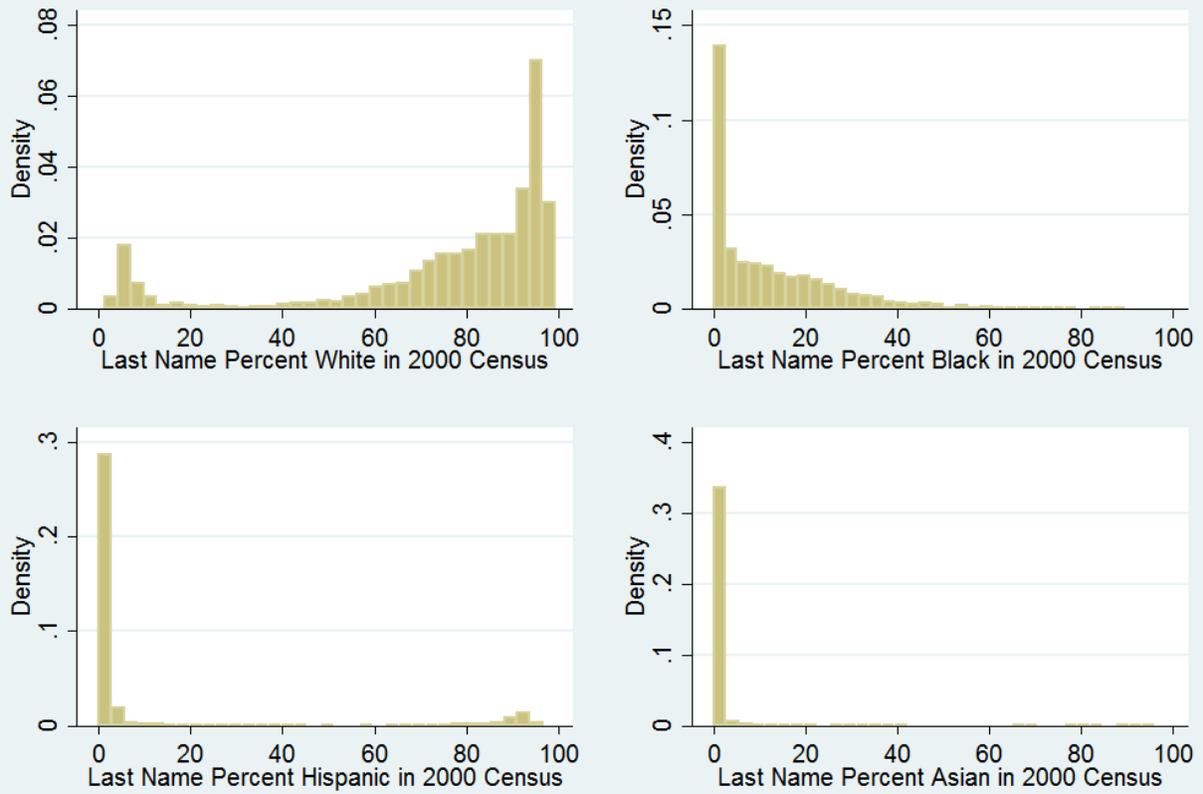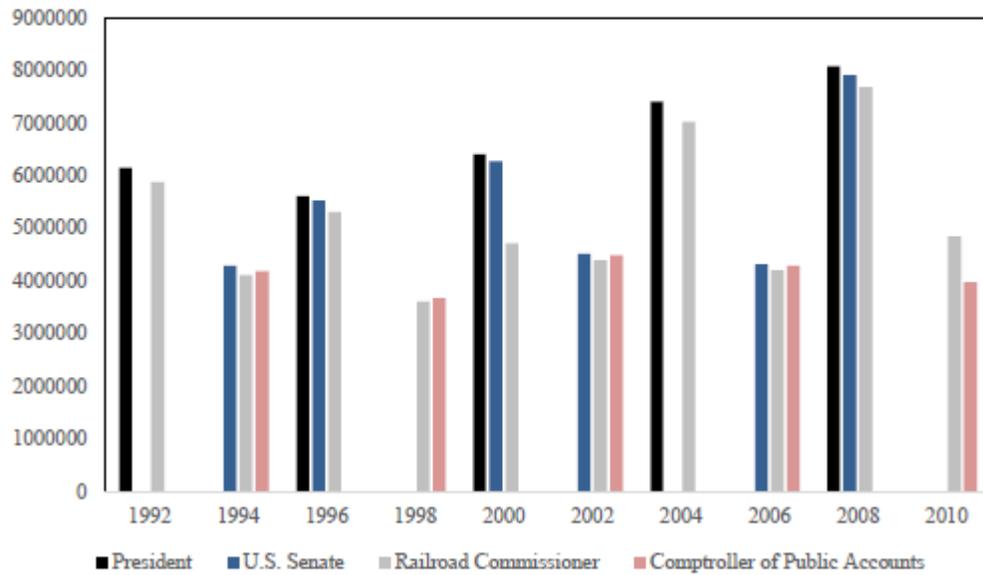Figure 1: Variation in Candidate Race and Ethnicity

Figure 2: Vote Totals Across Election Type

(a) Vote Totals in General Election

President ■ U.S. Senate ■ Railroad Commissioner ■ Comptroller of Public Accounts

(b) Vote Total for Railroad Commissioner: General vs. Primary Elections
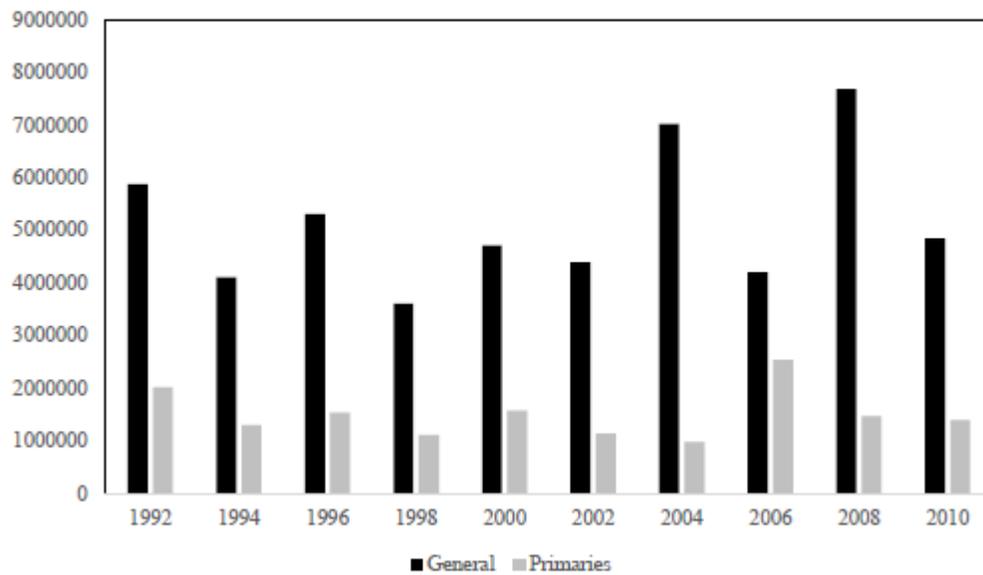

General ■ Primaries

Figure 3A: Google Trend Searches by Election Type 2004-Present
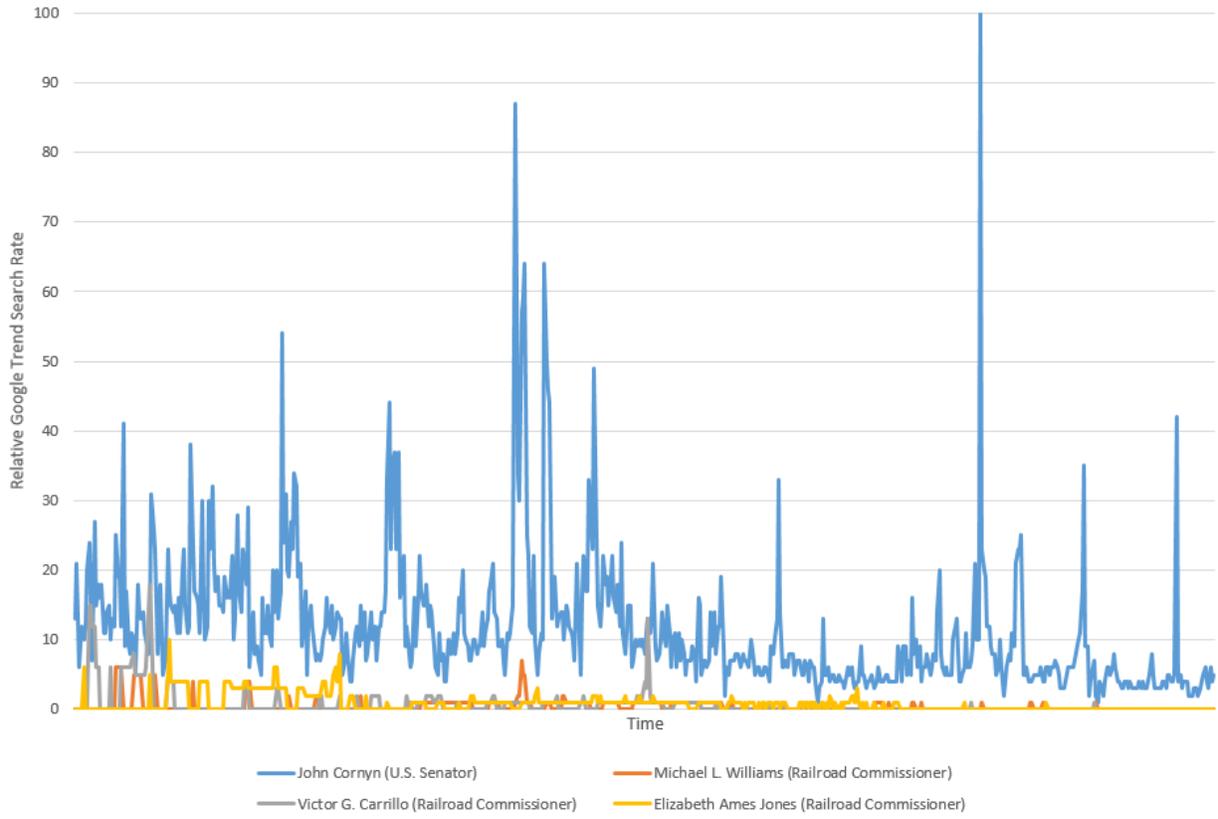U.S. Senate v. Railroad Commissioner

Figure 3B: Google Trend Searches by Election Type 2004-Present
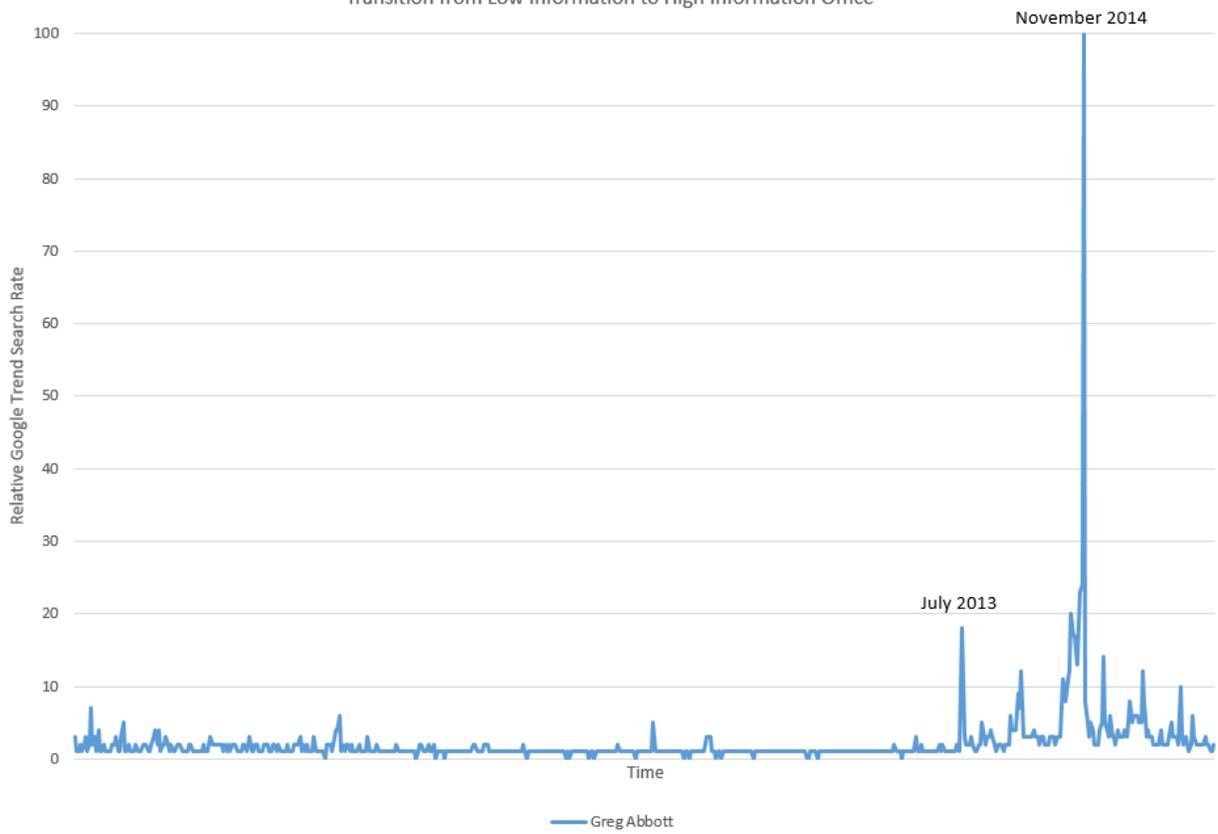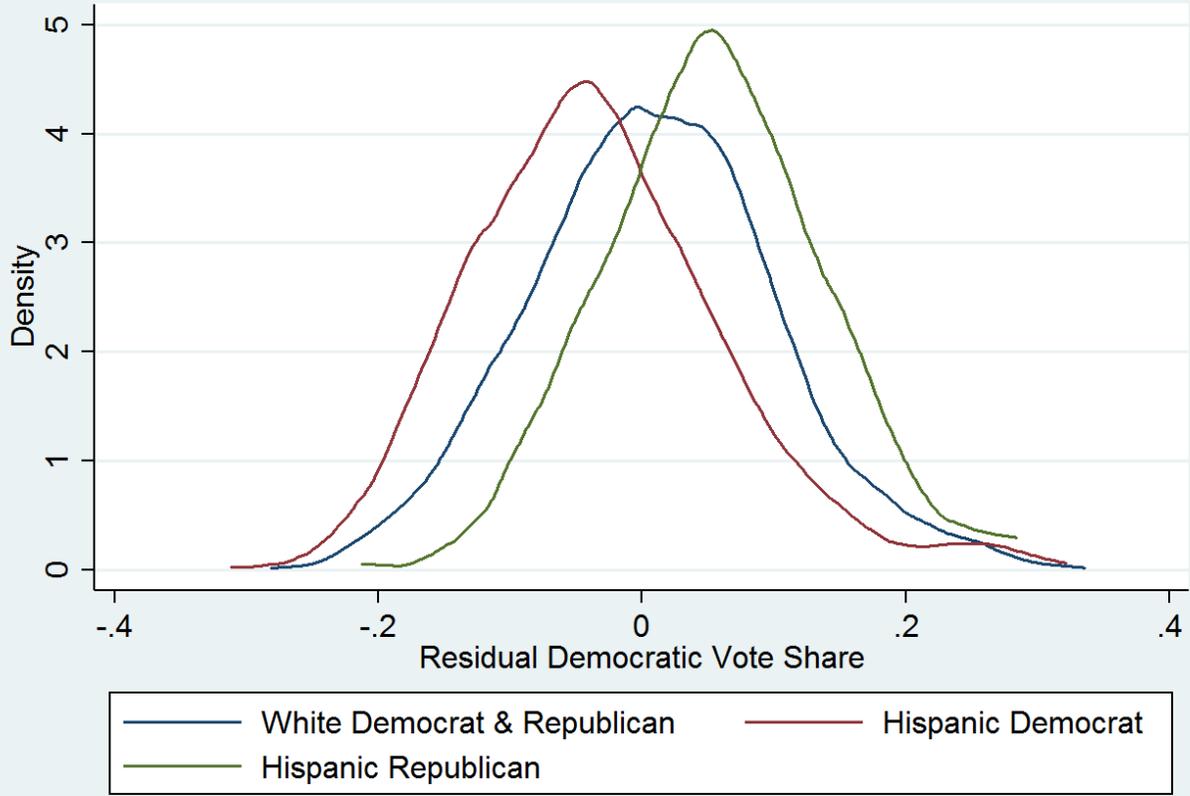Transition from Low Information to High Information Office

Figure 4: Residual Democratic Vote Share by Candidate Race and Ethnicity

<div align="center">References</div>

Bertrand, Marianne & Mullainathan, Sendhil. (2004). Are Emily and Greg More Employable
      than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *The*
      American Economic Review 94(4), 991-1013. Retrieved from
      http://www.jstor.org/stable/3592802

Google Trends. (2016). *Greg Abbott, Texas Governor; 2004-Present.* [Data file] Retrieved
      from https://www.google.com/trends/explore#q=%2Fm%2F055mw6&geo=US-
      TX&cmpt=q&tz=Etc%2FGMT%2B4

Google Trends. (2016). *John Cornyn, United States Senator; Michael L. Williams, Political*
      Leader; Victor G. Carrillo, Politician; and Elizabeth Ames Jones, Legislator; 2004-
      Present. [Data file] Retrieved from
      https://www.google.com/trends/explore#q=John%20Cornyn%2C%20%2Fm%2F0drxjr%
      2C%20%2Fm%2F080gxd2%2C%20%2Fm%2F05c2ghn&cmpt=q&tz=Etc%2FGMT%2
      B4

Hopkins, Daniel. (2009). No More Wilder Effect, Never a Whitman Effect: When and Why
      Polls Mislead about Black and Female Candidates. *The Journal of Politics*. Retrieved
      from
      http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1666187

Lindsey, Robert. (November 4, 1982). Bradley Loses Close Contest on Coast. *The New York*
      Times. Retrieved from http://www.nytimes.com/1982/11/04/us/bradley-loses-close-
      contest-on-coast.html

McDermott, Monika. (1998). Race and Gender Cues in Low-Information Elections. *Political*
      Research Quarterly 51, 895-918. Retrieved from
      http://prq.sagepub.com/content/51/4/895.short?rss=1&ssource=mfc

Minnesota Population Center, National Historical Geographic Information System: Version 2.0.
      Minneapolis, MN: University of Minnesota. (2011). *Census Tract Level Demographics,*
      1990-2010. [Data files and codebooks]. Retrieved from https://data2.nhgis.org/main

Missouri Census Data Center. (2014). *Listing of Geographic Correlations, 1990-2010.* [Data
      file]. Retrieved from http://mcdc.missouri.edu/websas/geocorr90.shtml

Office of the Texas Secretary of State. (2015). *1992 Current Election History, 1992-2010.* [Data
      file] Retrieved from http://elections.sos.state.tx.us/index.htm

Stephens-Davidowitz, Seth. (2013). The Cost of Racial Animus on a Black Presidential

> Candidate: Using google search data to find what surveys miss. Retrieved from
> http://static1.squarespace.com/static/51d894bee4b01caf88ccb4f3/t/51d89ab3e4b05a25fc
> 1f39d4/1373149875469/RacialAnimusAndVotingSethStephensDavidowitz.pdf

United States Census Bureau. (2014). *Frequently Occurring Surnames from the Census*

> *2000*. [Data file] Retrieved from
> https://www.census.gov/topics/population/genealogy/data/2000_surnames.html

United States Census Bureau. (2015). *2010 Census Gazetteer Files.* [Data file]. Retrieved from

> https://www.census.gov/geo/maps-data/data/gazetteer2010.html

United States Census Bureau. (2016). *County Business Patterns, 1998-2010.* [Data file and

> codebook]. Retrieved from http://www.census.gov/econ/cbp/index.html

Washington, Ebonya. (2006). How black candidates affect voter turnout. *The Quarterly Journal*

> *of Economics* 121(3), 973-998. Retrieved from
> http://www.jstor.org/stable/pdf/25098814.pdf?acceptTC=true