Names: Ryan Thomas, Abrar Bagalb, Robert Schultz

HW2

Read the paper:

Stephens-Davidowitz (2014). The Cost of Racial Animus on a Black Candidate Evidence using Google Search Data. Journal of Public Economics, 118, 26-40

Send back this file to: econometrics.methods@gmail.com

Background/Significance:

1) What is the motivation for this paper? What is the focus of inquiry (i.e. in a general way, what is the broad question of interest)? Does this question have policy relevance?

This paper explores the question of how much racial animus affects the ability of a black presidential candidate to be elected. The question does not have direct policy relevance, but it suggests a new proxy to quantify racial animus in future studies. The study helps understand the extent of contemporary prejudice as well as factors that determine a vote.

Overview:

2) What is the main causal question being asked in this paper?

a. This paper estimates the impact of racial animus (X) on the vote share a black presidential candidate receives in an election(Y).

b. How are X and Y measured?

(X) The researchers measure the amount of racially-charged Google searches and compare the percentage of people who voted for Obama with the percentage who voted for John Kerry. The impact of racial animus is measured by using a fixed effects framework to compare the outcomes of two democratic candidates - one white and one black - who are otherwise similar.

(Y) Vote share percentage is publically-available data from the presidential election.

3) What is the basic empirical challenge that the paper faces in tackling this causal question?

a. Let's say you know the association between X and Y from a large observational retrospective dataset. What are the sources of bias here (OVB/confounding, selection, etc)?

One source of bias is the assumption that individuals searching for [Word 1] on Google is a good proxy for racial animus. The author acknowledges the fact that Google searchers are more likely to be affluent. Only 70% of Americans had access to internet service. If African Americans make up a larger percentage of the households without internet, then racial animus might be overestimated. Additionally, only half of all searches were done in Google, so the Trends data might not be representative of the whole US depending on what demographics are more likely to search in Google. An additional source of bias could be the different circumstances surrounding 2004 vs 2008 elections.

Identification Strategy:

4) What is the identification strategy? (Overview)

a. What is the general class of identification strategy (RCT, diff-in-diff, IV, etc)?

The paper uses fixed effects to find the approximate percentage of votes lost by Obama due to racial animus.

b. How does the paper propose to obtain an unbiased (or relatively unbiased) estimate of impacts? Describe either the intuition or "thought experiment"

The main idea of this paper is that Google Trends data on searches for [Word 1] can be used as a proxy for racial animus. By combining the racial animus data with voter outcome data for elections for both a white presidential candidate (John Kerry) and a black presidential candidate (Barack Obama), the paper hopes to obtain an unbiased estimate of the effect of racial animus on vote percentage when controlling for a variety of additional factors.

c. If the paper is an RCT, give an overview of the intervention. If it evaluates a policy, provide an overview of the policy.

This paper is not an RCT, and does not consider, or measure a policy or intervention policy.

5) What is the identification Strategy? (Technical)

a. What are the main regressions for this strategy?

$$(\%Obama2008 - \%Kerry2004)_j = B_o + B_1 . Racially\ Charged\ Search\ Rate_j + \chi_j\, \phi^1 + \mu_j$$

b. What do each of the key regressors represent, and how are their coefficients interpreted?

The coefficient B_o represents a countrywide shock to Democratic popularity from 2004 to 2008, B_1 measures the impact of increasing one unit of standard deviation in Racially charged search rate j on Obama's vote share, χ_j represents the controls of the area-level that should influence change in supporting the democratic presidential candidate from 2004-2008. Racially Charged Search Rate j is as described in question 2b, using z-score to the normalization of data, and μ_j is noise.

c. Who are the treatment and control group?

Media markets with a greater-than-average search rate for racially charged words can be considered the treatment group, while media markets with the minimum search rate for racially charged words can be considered the control.

d. Who are the compliers?

Since this is a natural experiment and not an RCT, there isn't a clear-cut "complier" group. The media markets with greater-than-average search rates for racially charged words who voted mostly Republican can be considered the compliers.

6) What is the exclusion restriction?

a. What would cause the exclusion restriction to fail?

This paper does not utilize an instrumental variable, so this question is not applicable.

Data:

7) How is the dataset constructed?

a. What kind of dataset is this (observational, experimental, etc)?

Data is aggregated from Google Trends. This is collected observational data. Data is also collected from the General Society Survey (GSS) which consists of samples of the general population.

b. At what level(s) are the data measured (individual, household, village, etc)?

Data is collected and sorted by all 50 states and 196 of 210 media markets. Google Trends data is collected by IP address, so it could be at the individual or household level.

c. At what level(s) are the data grouped (household, village, school, etc.)?

Data is grouped by various demographics such as Education, state and political leaning. Some of the data is aggregated by media market.

d. Are there issues with measurement error? Attrition?

Measurement error occurs when searching words that aren't related to animus differ at the area-level. The author adds Google controls in Column (4) -(6) of Table 6 to reduce the measurement error in the area level proxy. The result shows that adding control for "African Americans", "niggas", and "profane language" increase the coefficient with a negative sign in each specification comparing to Column (1)-(3) of Table 6.

Google Trends data can also be skewed by individuals searching [Word 1] repeatedly. As discussed earlier, the data is collected by IP address. Searchers could be too young to vote as well, so a parent's child might search [Word 1] but their parent might not have any racial animus.

e. Do all variables actually capture the intended concept? E.g. does data on income adequately capture household consumption?

As with any study using google data, a search is merely an observation. A search could be for various reasons. In this study, some of the terms used also have other terms such as "joke" associated with it, so they have to use certain keywords. However, the Google Trend data is likely the best proxy for racial animus available because it is generally socially unacceptable to hold these beliefs and thus is unlikely to be self-reported on a survey. Searching on Google is perceived to be more anonymous.

Findings:

8) What are the main findings?

a. Interpret the magnitude of the coefficients.

Column 6 of Table 6 displays the results of the regression with all controls. For each standard deviation increase in racially charged search rate, Obama's vote share decreases by 1.776%. Since media markets' mean racially charged search rate is about 2.34 standard deviations higher than the minimum racially charged search rate, the percentage vote loss that can be attributed to racial animus is about 4.2%. As a conservative estimated range, the author estimates a 3-5% effect overall.

b. Does the magnitude seem reasonable? Does it seem like a meaningful effect size? (e.g. is it so small that it would never make a difference?)

The magnitude is reasonable and significant considering all of the controls employed in the regression. An election could easily be decided by 3-5 percentage points' difference in vote share.

c. What do you think of how the main findings are presented? (e.g. are they only in tables and should be presented graphically? Are the tables hard to interpret?)

The main findings are presented with a mix of tables and in-depth discussion of relevant research to prove that the estimate is as unbiased as possible. The tables and graphs are interesting and relevant while the discussion is accessible and thorough.

d. If relevant (e.g. for IV or D-D): do you think the identification strategy is presented convincingly?

Identification is not discussed in detail. However, we can gather from the author's discussion that a fixed effects framework is used. This makes sense after the author controlled for the changes in unemployment rates over time. There are inherent characteristics of the two candidates that can't be measured that could bias the results otherwise.

e. How do the findings compare to previous research (if relevant)? Are the differences between these findings and previous results what you would have expected?

This study builds on some other research testing racial attitudes on black candidates running for office. Also, it is believed to show bigger impact on race and voting than other various survey studies. Typically, racial attitudes are self-reported and are therefore inaccurate because of the

social pressure to hide negative racial attitudes. This study uses data from Google which we can assume is more accurate based on anonymity.

Theory/Mechanisms:

9) Does the paper discuss its theory of change or proposed mechanism? That is, what is the proposed pathway from intervention to outcome?

a. Is the proposed mechanism plausible?

The paper proposes that Google Trend data be used more often in difficult social science research where experiments are not feasible. The author suggests a few examples this method could be used with - voting; the use of illicit drugs; sexual practices; income; and embarrassing health conditions. It is plausible to use Google Trend data to analyze these topics because searchers are more open to type about these topics since it is virtually anonymous and there is no perceived risk to searching these terms. However, it is uncertain whether Google Trend data provides good representative samples. Further research on this topic is needed.

10) Is the paper able to assess the theory or mechanism? For example, does it collect data on intermediate outcomes or behaviors? Or is it a black box where T goes in, Y goes out, and we don't really know why?

The paper is only able to make assumptions that searches for [Word 1] are correlated with racial animus. The assumption becomes credible when the author controls for words 2 and 3 to account for African Americans searching the term and to control for areas which are more prone to using profane language.

11) If the intervention is found to have a significant impact, what evidence is provided for why the program was successful? If the program was not successful, what evidence is provided for why the program was not successful?

This study did not look at an intervention or policy change. The study considered voter patterns and searches.

Improvements to the Study:

12) What would you have done differently?

a. If you had the same or similar dataset, is there an alternative identification strategy? A better/more realistic intervention to analyze?

Understanding voter preferences is somewhat of a black box. Other than asking people why they voted a certain way, you can't get much feedback. Obviously some in a survey would give another answer as to why they perhaps didn't vote for a candidate of a certain race than "they were X race." Voer and poll data is also hard to use because it's always a sample and can be extremely cherry picked of the respondents they use to get a result.

b. Could you answer the same question with a different kind of data? Example: use better data on final or intermediate outcome. See whether results generalize in a different population.

This paper discusses how other researchers have attempted to explore this question. One issue is that in surveys, respondents tend to try to appear in a better light, so using a

survey is not always the best. As discussed earlier, Google data is merely searches and a lot of assumptions need to be made in order to convincingly prove causality. The paper does a good job of arguing how they were able to introduce controls and and split by demographics to make the best use of the data. A larger sample size is probably not possible, given the fact that millions of people use Google to search every day and its market share as a search engine is about 50%.

c. Would you have discussed limitations or advantages of the design differently?

The paper overall discusses how they built on previous research a bit as well as how their methods were believed to be more accurate. Specifically, they discuss the limits of using survey data in previous research that looks at things like race. To counter this, they add more data that is less subjective than self reported. They do still combine some survey data into their analysis. They clearly discuss how a google search is only a search and how they use this as an instrument. For example, they do control and look into whether some states are more prone to certain words by testing certain word search usage outside of racially charged language.