Partisanship and Survey Refusal

Mark Borgschulte Heepyung Cho Darren Lubotsky

October 2019

Abstract

Survey refusal in the Current Population Survey (CPS) has tripled over the last decade. This rise coincides with the emergence of rhetoric, largely from the political right, questioning the accuracy and integrity of government statistics. We examine how support for the Tea Party and the Republican party have affected CPS refusal rates and whether households are more likely to participate in the survey when their preferred political party holds the White House. Using state and metro vote shares or an individual-level model based on the longitudinal structure of the CPS, we find no evidence that Republican or Tea Party supporters drive the long-term upward trend in refusals. We do find evidence of a political cycle in response rates. Refusal rates since 2015 exhibit polarization, with the fastest growth in refusals among those *least likely* to support Trump and the Tea Party. Evidence from an analysis which generates exogenous variation in Tea Party support using rain on the day of the first Tea Party rally indicates that exposure to anti-survey rhetoric decreases refusal rates, consistent with the findings from our other analyses.

Keywords: Current Population Survey, survey refusal, unit non-response, unemployment rate
JEL Classifications: C81, J64, P48

We thank Jacob Bastian, David Card, Kyra Linse, Steven Rivkin, Jim Ziliak and seminar audiences at the University of Illinois at Chicago, University of Illinois at Urbana-Champaign, and the University of Wisconsin for helpful discussion and comments. Affiliations: Borgschulte, Univeristy of Illinois at Urbana-Champaign and IZA; Cho, University of Illinios at Urbana-Champaign; Lubotsky, University of Illinois at Chicago and NBER.

1 Introduction

Over the last decade, prominent Republican politicians have expressed an increasing degree of skepticism regarding the accuracy and integrity of government statistics. These concerns have largely focused on the unemployment rate, a key metric of labor market performance often used to assess the effectiveness of economic policy.[1] A related set of concerns arose during the lead-up to the 2010 Decennial Census, during which several Tea Party politicians and pundits suggested that people should only partially respond, or refuse altogether. Rising skepticism comes at a time in which the quality of survey data is at risk due to increasing difficulties in getting households to respond to surveys. The share of households that refuse to fill out surveys has been rising across most surveys over the last 25 years, imperiling the quality of statistics used to administer government programs and conduct social science.[2] Although attitudes towards government have been put forward as a possible explanation (e.g., in National Research Council (2013)), there are few empirical tests of this or other theories that could explain rising refusal rates.

In this paper we examine whether individuals' partisan affiliation and ideology can explain the upwards trend in refusal rates in the Current Population Survey (CPS). We also study whether participation in the survey is related to the political affiliation of the President. There are several reasons to focus on the CPS. First, the CPS is administered by the Census Bureau and is the source for the official unemployment rate statistics, meaning it is closely connected to several key elements of the recent anti-survey rhetoric. Second, the CPS is considered among the best-run surveys measured by size and response rates, however, CPS refusal rates have increased sharply in recent years. Indeed, writing several years ago, Krueger, Mas and Niu (2017) assert that "irresponsible politicians" were to blame for a surge in non-response in the CPS beginning around 2010. The CPS is among the most commonly used surveys to study labor market behavior, poverty, wage inequality, and other social and economic phenomena, and non-response has been shown to gener-

[1] For example, during their presidential campaigns both Mitt Romney and Donald Trump stated that the reported unemployment rate significantly understated the true state of the labor market. We describe in detail the concerns that have been raised, including the emergence of the "real unemployment rate" rhetoric , in Section 2.

[2] These trends and the related literature are surveyed in Meyer, Mok and Sullivan (2015) and discussed in Section 3.

ate important biases.[3] Finally and most importantly for our analysis, the CPS is conducted with a short longitudinal structure. Households are interviewed once per month for four months, ignored for eight months, and then interviewed again for another four months. We use the responses in other rounds to examine the characteristics of households that refuse to fill out the survey. Our analysis indicates that the vast majority of households who refuse to respond to one round of the CPS do respond to at least one other round of the survey.

We begin with a cross-state comparisons of refusal rates based on voting patterns in the 2008, 2012 and 2016 presidential elections. We find that refusal rates in so-called "red states" rose earlier and faster than in "blue states" in the first years of the Obama presidency, though blue states do show a marked increase in refusal rates over the Obama presidency, as well. These patterns then reverse themselves late in the Obama presidency, with blue state refusal rates growing relatively faster than red states since 2014. Taking a longer view, the results suggest a modest political cycle in refusal rates over the last 25 years, with the blue-red state gap in refusal rates closing by one-half to one percent during Clinton and Obama's presidencies.[4] These patterns are consistent with an effect of political affiliation, though they do not support the hypothesis that the recent rise in refusals can be attributed to rhetoric questioning the accuracy and integrity of government statistics, largely associated with the political right. It is possible, however, that we require information on the household- and individual-level to detect the hypothesized effects, for example, because survey refusal is a relatively rare outcome or geographic correlations can exhibit the ecological fallacy.

For these reasons, we next turn to micro-level evidence using the longitudinal structure of the CPS. As the CPS lacks data on partisanship, we employ a proxy-variable strategy based on questions in the American National Election Survey (ANES) asking specifically about support for Republican Party presidential candidates (in 2004, 2008, 2012, and 2016) and the Tea Party (in 2010 and 2012). We map predicted responses to these questions to the CPS using the rich set of

[3] Many studies of CPS non-response bias focus on item non-response and misreporting/measurement error, e.g. Hokayem, Bollinger and Ziliak (2015), Bollinger et al. (2019), and Meyer and Mittag (2019). Bee, Gathright and Meyer (2015) and Heffetz and Reeves (2019) study biases from unit non-response, which is our focus.

[4] Similar patterns are found in a comparison across large metropolitan areas. Finer geographic analysis is not possible, as CPS geography is non-representative and masked at most sub-state levels.

covariates that overlap between the surveys. This strategy helps alleviate concerns about potential ecological inference issues in examining geographic correlates and is effective for addressing attenuation due to errors in the responses to the ANES (Hyslop and Imbens, 2001). Intuitively, this specification relies on a type of panel identification, in which we examine changes in a group over time, controlling for group and time effects; a similar strategy is used by Boxell, Gentzkow and Shapiro (2017).

We find that refusals grew across the political spectrum, consistent with what we found in our prior analysis. However, we also find evidence of an increase in polarization, particularly over the last five years. Relative refusal rates among GOP and Tea Party supporters increased starting in 2009, particularly in percentage terms (i.e. compared to the low rates of refusal in this group in the preceding decade). This reversed more recently, when Republican and Tea Party support predicts a *slower* growth rate of refusals beginning around 2014. A difference-in-difference analysis isolating within-household/across-round changes in refusals around the 2016 election reveals that the Democratic-leaning households became relatively more likely to refuse following Trump's election. We conclude this section with a comparison of these partisanship-based patterns to the household- and individual (i.e. household-head) predictors of refusals in a Blinder-Kitagawa-Oaxaca decomposition analysis of the rise in refusal rates. The decomposition finds the most important predictor of the recent rise in refusals to be the household employment rate, which the ANES model does not assign an important role in predicting partisanship. This analysis also indicates that over half of the rise in non-response is not explained by a rich set of observable characteristics, which indicates that it would be difficult to develop population weights to properly adjust for non-response.

Finally, we conduct an analysis to examine the effect of local support for the Tea Party movement, leveraging rainfall on April 15, 2009, the day of the first national Tea Party rally. Madestam et al. (2013) shows that rain on this day had important effects on Tea Party support, including on vote outcomes in the 2010 midterm elections. Given the patterns established in this previous work, this analysis isolates a causal effect of expanding Tea Party support around the time of the trend break in CPS refusals. Tea Party supporters are likely to be among the most receptive supporters of the Republican party to anti-government rhetoric, and a number of leading figures in the Tea Party endorsed anti-survey and anti-Census rhetoric, including conspiracy theories about the unemployment rate. We find, however, refusal rates in 2009 and 2010 briefly *increased* in areas where Tea Party support was *depressed* by rain on April 15, 2009; stated differently, the effect of the Tea Party on marginal supporters was to increase the likelihood they respond to the CPS. Moving later in the decade, we find that these areas with depressed Tea Party support continue to show higher refusal rates in 2014 and later. We speculate that these effects may reflect an increase in the salience of the CPS, and discuss alternative explanations in Section 6.

To sum up, we find evidence of a political cycle and emerging polarization in refusal rates. While the political cycle in refusals may have strengthened somewhat in the previous decade, it appears to be a feature of the data going back to the mid-1990s. Political cycles have previously been documented in trust in government (Pew Research Center, 2019), stock market returns (Santa-Clara and Valkanov, 2003; Pastor and Veronesi, 2017), the receipt of foreign aid (Faye and Niehaus, 2012) and government deficits (Shi and Svensson, 2006). The findings also reflect a differential non-response theory of survey response, in which non-response reflecting political sentiment induces selection bias (Gelman et al., 2016). We discuss alternative explanations for the overall rise in refusals since 2010 in the conclusion.

The paper proceeds as follows. In Section 2, we discuss the rise in anti-survey rhetoric and the origins of the term the "real unemployment rate." In Section 3, we discuss the CPS data and recent trends in CPS refusal, before turning to the state and metro analysis in Section 4. In Section 5.2 we present our main estimates of the household model based on the ANES prediction of partisanship, and contrast it to a Blinder-Kitagawa-Oaxaca decomposition. In Section 6, we analyze local Tea Party support. Section 7 concludes with a summary of the findings and a discussion of alternative explanations for the recent rise.

2 Skepticism of Government Data Collection and Survey-based Statistics

The rhetoric surrounding skepticism of government surveys over the last decade had its genesis in legitimate debates among economists and policy makers about the measurement of the state of the economy during the Great Recession. There are well-known conceptual issues with the unemployment rate (Card, 2014). The Great Recession and recovery were notable for a large increases in long-term unemployment, discouraged workers, and labor force nonparticipation, none of which are perfectly captured by the official U-3 unemployment rate. With this in mind, a number of commentators suggested that attention be refocused on the U-6 rate, which counts discouraged and under-employed (i.e. part-time employees who would like to find a full-time job) as "unemployed." The U-6 rate was introduced with the CPS redesign in 1994, but its use as a secondary measure of the health of the national labor market had mostly been limited to economic policy makers and central bankers. The term "real unemployment rate" seems to have been coined in early 2009 to refer to the U-6 rate; by mid-2009 this term can be found in many economics blogs.[5]

Following Obama's inauguration, rhetoric surrounding the "real unemployment rate" escalated quickly. By mid-2009, the monthly announcement of the official unemployment rate was met with regular and loud claims by administration critics that the *real* unemployment rate was twice as high. Although the origin of the term suggested a focus on the U-6 measure ahead of the U-3 measure, it did not take long before the connotation that the real rate was much higher than what people were being told morphed into attacks on the veracity of the numbers themselves. The mixture of truth and conspiracy found a receptive audience in the newly-formed Tea Party, which coalesced around a number of conservative anti-tax positions opposing the Obama administration's health care and macroeconomic policies. Tea Party supporters, including the initiator of the movement

[5] Florida, Richard. 2009. "Uneven Unemployment." *The Atlantic*. May 8. https://www.theatlantic.com/national/archive/2009/05/uneven-unemployment/17333/. Florida writes "But the real unemployment rate is as high as 15.8 percent according to the BLS U-6 measure which includes marginally attached and discouraged workers".

Rick Santelli, embraced conspiracy-theory versions of the unemployment rate debate.[6]

By early 2010, concerns about the integrity of the unemployment rate was echoed in skepticism of government data collection and opposition to questions in the 2010 Census that went beyond the narrow goals of enumerating the population.[7] In the run-up to the 2010 Census, for example, U.S. Representative and founder of the House Tea Party Caucus Michelle Bachmann argued that the Decennial Census had become too intrusive and received significant media coverage for her boast that she would only fill out information on the number of individuals who lived in her home.[8] These themes were picked up by a number of right-wing and Tea Party-affiliated media outlets, such as Glenn Beck, Sean Hannity and other conservative media outlets.[9] Controversy surrounding the 2010 Census was one of the concerns that caused Republican Judd Gregg to withdraw as the nominee for Commerce Secretary in the Obama administration.[10] Republican politicians also issued statements claiming that the Census subcontracted enumeration services to ACORN, a liberal activist group. While previous Censuses had dealt with controversy, especially over sample and re-weighting to address low response rates in high-poverty and high-immigrant areas, the 2010 Census marked a turn, encouraging refusal based on political beliefs.[11]

Since the 2010 Census, skepticism of how the government collects and reports the unemploy-

[6]New Yorker talks about Rick Santelli saying unemployment rate numbers are cooked: Cassidy, John. 2012. "Obama, the Job Figures, and the Conspiracy Theorists." *The New Yorker*. Oct 5. https://www.newyorker.com/news/john-cassidy/obama-the-job-figures-and-the-conspiracy-theorists. Santelli in 2013 saying the BLS surveyors are manipulating the numbers: Hains, Tim. 2013. "Santelli On "Fake" Census Data & Unemployment Numbers: "If We Knew Now What We Knew Then"." *RealClear Politics*. Nov 19. https://www.realclearpolitics.com/video/2013/11/19/santelli_on_fake_census_data__unemployment_numbers_if_we_knew_now_what_we_knew_then.html.

[7]Nasaw, Daniel. 2010. "Barack Obama opponents urge census boycott." *The Guardian*. Jan 29. https://www.theguardian.com/world/2010/jan/29/barack-obama-opponents-census-boycott

[8]Bachmann cited data privacy concerns and the use 1940 Census data in the roundup of Japanese and Japanese-American internees as an example of Census data misuse (Anderson, 2015).

[9]Nasaw, Daniel. 2010. "Barack Obama opponents urge census boycott." *The Guardian*. Jan 29. https://www.theguardian.com/world/2010/jan/29/barack-obama-opponents-census-boycott. Schwen, Christine. 2009. "Media conservatives target the 2010 census, encourage audience not to complete forms." *Media Matters*. Jun 26. https://www.mediamatters.org/fox-nation/media-conservatives-target-2010-census-encourage-audience-not-complete-forms.

[10]Baker, Peter. 2009. "A Nominee's Exit and the Nation's Nose Count." *The New York Times*. Sep 20. https://www.nytimes.com/2009/02/20/us/politics/20memo.html

[11]According to a 2010 New York Times/CBS poll, only one percent of self-identified supporters of the Tea Party said they would not fill out the 2010 Census, compared to three percent in the population as a whole. Zernike, Kate and Thee-Brenan Megan. 2010. "Poll Finds Tea Party Backers Wealthier and More Educated." *The New York Times*. Apr 14. https://www.nytimes.com/2010/04/15/us/politics/15poll.html.

ment rate has become a standard element of Republican rhetoric, with increasing levels of acceptance of the more conspiracy-minded view by the Republican mainstream. During the 2012 campaign, Mitt Romney claimed the real unemployment rate was “likely 12 to 14 percent,” and affiliated Political Action Committees claimed the real unemployment rate was 19 percent.[12] Fox News ran a number of news stories casting doubt on the integrity of the unemployment rate in anticipation of the October 2012 jobs report.[13] Former General Electric CEO and Republican donor Jack Welch said, regarding positive news in the October job report, that President Obama and his campaign aides "will do anything ... can't debate so change numbers." Many of these claims had come unmoored from the U-3 versus U-6 distinction, and expressed a tone of doubt about the reported numbers. These attacks returned at greater volume in the 2016 election when then-candidate Trump claimed the true unemployment rate was “as high as 42 percent.”[14] Trump also made regular attacks on polling data and the integrity of pollsters.

To put things in empirical terms, in Figure 2 we plot Google searches for the terms “real unemployment rate,” “Tea Party,” and “Trump” using data from Google Trends.[15] Searches for “real unemployment rate” became common in 2009 before surging in 2010. Following a brief lull in 2013 and the first half of 2014, the search term regained popularity, before collapsing with the election of Trump in late 2016. The emergence of searches is timed with the instigating events we highlight in this section. The rather dramatic fall after Trump's election is consistent with Republican voters' skepticism of government statistics produced during the Obama Administration.

[12]Jacobson, Louis. 2012. "Mitt Romney-aligned group says 'real' unemployment rate is 19 percent." *Politi Fact*. Sep 21. https://www.politifact.com/truth-o-meter/statements/2012/sep/21/restore-our-future/mitt-romney-aligned-group-says-real-unemployment-r.

[13]Groch-Begley, Hannah and Dimonda, Alessandra. 2012. "Fox Pushes Jobs Trutherism In Anticipation Of October Jobs Report." *Media Matters*. Nov 1. https://www.mediamatters.org/fox-news/fox-pushes-jobs-trutherism-anticipation-october-jobs-report.

[14]Kessler, Glenn. 2009. "Trump's claim that the unemployment rate is 23 percent." *The Washington Post*. Jan 20. https://www.washingtonpost.com/news/fact-checker/wp/2016/01/20/trumps-claim-that-the-unemployment-rate-is-23-percent. Zumbrun, Josh. 2015. "Donald Trump Is Right: About 42% of Americans Are Unemployed (If You Include My 88-Year-Old Grandma)." *The Wall Street Journal*. Aug 20. https://blogs.wsj.com/economics/2015/08/20/donald-trump-is-right-about-42-of-americans-are-unemployed-if-you-include-my-88-year-old-grandma.

[15]We take averages within quarter to smooth the series.

3 CPS Data and Trends in Refusal

In this section, we describe the CPS data, focusing on refusals and their measurement, and also provide comparisons to other surveys to place the rise in CPS refusals in context. Survey refusal occurs when a resident of the sampled household is contacted but refuses to fill out the survey. This is distinct from item non-response, which is when a survey is completed but the respondent chose not to answer a particular question or questions. Rising non-response and refusal is a problem because it raises the costs of administering the survey and makes it more difficult to ensure that the data is representative of the population. Refusal at the unit level can be particularly problematic as the data for these households then lacks core variables needed to construct weights to address item non-response in key outcomes.[16]

We use the Current Population Survey (CPS) Basic Monthly Data from the Integrated Public Use Microdata Series (Flood et al., 2018) in our analysis. The CPS is administered at the address level. A multistage stratified sample of approximately 72,000 housing units are randomly chosen from 824 sample areas to participate in the survey. One person in a household usually responds for all eligible members in the household, who is called the "reference person." The reference person is generally the person who owns or rents the housing units. If the reference person is not knowledgeable about other household members, the Census attempts to interview other knowledgeable adult member(s) in a household. If the household moves out of the sampled address, they leave the sample. If a new household moves into the housing unit, they are eligible for inclusion in the survey.[17]

Each household is interviewed for four consecutive months, not interviewed for the following eight months, and then interviewed again for four consecutive months, which correspond to the same four calendar months of the initial four interview. Thus in any given month, there are eight "cohorts" in the CPS, corresponding those participating in their first through eighth monthly in-

[16]This is less of a problem for the CPS than for other surveys, as we show below. A fast-growing literature documents the consequences of these trends for biases in survey measures of the unemployment and poverty rates, inequality, obesity, etc.

[17]The complete technical documentation for the Current Population Survey is available at https://www.census.gov/programs-surveys/cps/technical-documentation/complete.html.

terview. The Bureau of Labor Statistics puts enormous effort into maximizing survey completion. The first and fifth interviews with a household are generally conducted face-to-face with a Census Bureau enumerator. Other interviews generally take place over the phone. Enumerators follow-up multiple times during a month when a survey has not been completed to attempt to obtain a completed survey.

When a household fails to take a survey in a certain month during the interview period, it is recorded as "non-response" of Type A, B, or C. Type A indicates a household that was eligible to be interviewed but did not because they refused, were not at home, or were temporarily absent from the home. The subset of Type A non-response that is coded as a refusal is the focus of our analysis. The IPUMS-CPS, however, does not have detailed information on the specific reasons of Type-A non-responses. Thus, we additionally use the raw CPS data from the Census Bureau along with the program by Center for Economic and Policy Research (Center for Economic and Policy Research, 2019) and match with the IPUMS-CPS using the unique household identifiers. We drop from our analysis non-response Types B and C, which were ineligible for inclusion in the survey.[18]

Figure 1 shows trends in Type-A non-response and its subcomponents from January, 1994 to January, 2019. Non-response rates rose from 6.5 percent in January 1994 to 16.9 percent in January 2019, with a distinct trend break around 2010, the year in which Tea Party supporters began raising questions about the Decennial Census. Survey refusal accounts for virtually all of the near-doubling in unit non-response in the Current Population Survey over the last decade. Refusal rates ranged from 3 to 5 percent until 2010 and rose to 15 percent at the end of the sample period. The figure also depicts several auxiliary categories of non-response: no one at home, temporarily absent, and other.[19] These categories show almost no trend, with a recent but modest decline in noncontacts ("no one at home" and "temporarily absent"). None of the movements in these auxiliary non-response codes can explain the rapid growth in refusal.

[18]Type B nonresponse indicates a housing unit that is intended to be occupied, but there are currently no persons living there who are eligible for interview (e.g., the unit is vacant or is entirely occupied by people whose usual residence is elsewhere). Type C nonresponse refers to housing units that are ineligible for an interview, such as units there were demolished, converted to storage or business use.

[19]CPS added "language barrier" and "unable to locate" categories starting from 2010, and we include them in "other" category in Figure 1.

Other US and international cross-sectional surveys have exhibited increasing non-response rates dating back several decades; the CPS has, until recently, stood apart from this trend. In Figure A2, we depict non-response rates in three other large-scale U.S.-based cross-sectional surveys: the American Community Survey (ACS), National Health Interview Survey (NHIS), and General Social Survey (GSS). Beginning in the mid-1990s, non-response rates have risen in both the GSS, from 25 percent to close to 40 percent, and from below 10 percent to above 30 percent in the NHIS. The ACS has much lower non-response rates, likely because response is mandatory; however, non-response rates have grown from 3 to above 6 percent since 2010. By comparison, the CPS exhibits later and slower growth in non-response than the GSS and NHIS.[20] Biases associated with nonresponse have also been found in the University of Michigan's Survey of Consumers Curtin, Presser and Singer (2005); **?**. International surveys have also exhibited rising non-response, though again, this trend pre-dates recent patterns in the CPS by several decades (De Heer and De Leeuw, 2002; de Leeuw, Hox and Luiten, 2018). A unique advantage of studying the Current Population Survey, relative to most other surveys, is that the large sample size, sub-state geography, and short longitudinal structure allow us to measure how geographic and household characteristics correlate with changes in patterns of survey response.

It is possible that the increase in non-response reflects technological or methodological changes, particularly due to the increasing replacement of landlines with cell phones. However, we show in Appendix Figure A3 that refusal rates in the CPS increased by almost the same rate in both in-person and telephone surveys.[21] As Brick and Williams (2013) note, this similarity suggests that technological changes, including increased cell phone usage or the decline in social capital, are not likely the driving cause of the recent rise in the non-response rates. Moreover, there was no major change in CPS survey methodology or questionnaire which may significantly affect refusal rates, after the redesign in 1994 (U.S. Census Bureau, 2006).[22] Since technological or methodological

[20] In contrast to cross-sectional surveys, panel survey in the United States and other countries have not seen a rise in non-response (Schoeni et al., 2013).

[21] Also, the proportion of in-person and phone interviews have remained mostly unchanged after 1994.

[22] Technical Paper 66 (published in 2006) is the latest version of the technical paper from the BLS that discusses the CPS design and methodology in detail. See also `https://www.kansascityfed.org/research/kcdc/cps/techdoc/~/link.aspx?_id=A8FC2592EF0D4B67A5620D351644569C&_z=z` for a full list of changes from 1994. Ex-

changes cannot fully explain the rise, we next turn our focus to the political explanation of the rise in the subsequent sections.

4 State and Metropolitan Area Analysis

Our goal is to examine the relationship between political party affiliation and CPS response. However, the CPS lacks data on political preferences and party support. We thus pursue two complementary strategies. In this section we examine the relationship between CPS non-response and state and metropolitan area voting outcomes. In the next section we use a proxy variable strategy to predict political preferences at the individual level in the CPS. Finally, in Section 6 we assess whether arguably exogenous variation in Tea Party support in an area is related to CPS non-response.

We define the area-level refusal rate as the number of households who refused divided by the number of households eligible to be interviewed in a state or metropolitan area (excluding Type-B and Type-C non-interview households). Notably, geographic information is available for all households, even if they never respond to a survey. We include all eight potential interviews in this calculation. When we aggregate by metropolitan area, we also separately include the aggregate of all non-metropolitan areas within a state as a separate geographic unit.[23]

We first present refusal rates by state in maps in Figure 3. Panels A and B show the average state-level refusal rates in 2010 and 2018. Panel C shows the difference in state-level refusal rates between the two periods. Refusal rates increased in all states during this 8-year time period, but the size of the increase is quite heterogeneous across states. For example, the refusal rate in Washington, DC increased by 21 percentage points, whereas it increased only by 2 percentage points in Iowa.

Next in Figure 4, we show trends in monthly refusal rates separately by "Blue States," "Red

cept for the expansion of the monthly samples that took place in 2001 (from 50,000 households to 60,000 eligible households), we find that there was no major change in sampling, design or methodology in CPS.

[23] We drop households whose metropolitan area status is unknown (about one percent of the sample after 2000).

States" and "Purple States," which are classified by states' party affiliation in the recent three presidential elections (2008, 2012 and 2016). The 14 Blue States are those where the Democratic candidates had a vote margin greater than 10 percentage points over the Republican candidate. Similarly for the 14 Red States. The remaining 23 states are classified as "Purple States." The black solid line represents the difference in refusal rates between Blue and Red States.[24]

Figure 4 indicates that refusal rates have been lower historically in Red States. Starting in 2010 both the Blue and Red States experienced the rapid rise in refusal rates, which indicates that the partisanship itself cannot entirely explain the rise or acceleration in the recent years. However, there is suggestive evidence of political cycle in refusal rates, where the difference in refusal rates between Blue and Red States changes according to the party in control. During the most of George W. Bush and Donald Trump presidencies, for example, refusal rates generally grew faster in Blue States. On the contrary, refusal rates increased faster in Red States during most of the Obama years, closing some of the gap between the Blue and Red States.

Finally in Figure 5, we examine refusal rates across metropolitan areas (including non-metro areas of each state). We show the trends in metropolitan area refusal rates by quartile of McCain vote share (Panel A) and Trump vote share (Panel B) in the 2012 and 2016 presidential elections. These patterns generally mirror those in our state level analysis. Refusal rates for each group increased rapidly around 2010. We again find suggestive evidence of a political cycle in refusal rates. Especially after the 2016 election, the refusal rates in "Blue Metros" grew faster than those "Red Metros", resulting in fanning out of refusal rates across quartiles. Though perhaps less conspicuously in the figure, refusal rates in "Red Metros" grew faster during the years of the Obama presidency.

We also fit linear splines on refusal rates of Blue and Red metros, allowing for trend breaks at the times of election (Nov 2008, Nov 2012, Nov 2016). During the first Obama term, refusal rates in Blue Metros increased on average by 0.02 percentage points per month, compared to 0.03 points in Red Metros. After Trump was elected, refusal rates of Blue Metros increased by 0.21 percentage

[24] The difference in refusal rates is averaged at the quarterly level to smooth the series.

points per month, compared to only 0.15 percentage points in Red Metros. These differences in the growth rate of refusal rates between Blue and Red metros are statistically significant.

5 Does Household Partisanship Explain the Rise in Refusals?

In this section we examine the effects of political affiliation at the household level. To do so, we rely on on the short panel design of the CPS to identify households which refuse to answer some but not all rounds of the survey. Following Abraham, Helms and Presser (2009) and others we use the panel structure, in which households are surveyed eight times during the 16-month period, to construct a household-level refusal rate, defined as the number of refusals divided by the number of surveys eligible for interview (the number of completed interviews plus the number of Type A non-responses).[25] We drop always-refuse households (i.e. a refusal rate of 100%) since information on their characteristics is not available. Thus, the sample is composed of never-refuse and sometimes-refuse households, and the possible range of refusal rate is from 0% (no refusal) to 87.5% (refusing 7 out of 8 surveys).

It is, of course, possible that households which respond to at least one round of the survey are fundamentally different from those which never respond. We first address this concern in Figure 6, which compares the overall refusal rate in the CPS, the refusal rate after excluding the never-responders, and the share of households that never respond to the CPS. The horizontal axis is the interview cohort; i.e. the first month that a sampled household was supposed to be surveyed. The prevalence of never-responders was about three percent prior to 2010 and then began to rise to just over five percent by 2018. Importantly, the trend in refusal rate of households excluding never-responders (dashed line) is similar to the overall refusal rates including never-responders (solid line). Thus, sometimes-refuse households are an important part of the rise in refusals and understanding the correlates of their response rates is important for understanding trends in overall response rates. Furthermore, in the Appendix (Figure A4) we show that average metro-level Re-

[25]To be clear, a household may be eligible for fewer than eight interviews if they move into or out of an eligible address during the sixteen months of the survey period.

publican vote shares are quite similar between households that refuse all rounds of the survey and those that refuse to six or seven out of eight rounds (i.e. respond to two or one round). These tests further strengthen our confidence that we lose little by dropping always-refuse households.

A major challenge is that the CPS does not contain any direct information about individuals' political beliefs or attitudes. We therefore use survey data from the American National Election Studies (ANES) in 2004 2008, 2010, 2012 and 2016 to map political attitudes and party affiliations to the individual characteristics that are measured in both the ANES and CPS (University of Michigan, Stanford University and National Science Foundation, 2018). The ANES is a nationally representative survey that gathers data through in-person or online interviews conducted twice for each respondent before and after presidential elections (pre-election and post-election surveys).[26] The data contain a rich set of variables on an individual's voting behavior and political beliefs, including support for the Republican candidate in the most recent presidential elections, and for the Tea Party. Additionally, the data include standard socio-demographic characteristics of individuals.

The household-level analysis begins with the estimation of OLS regression models in the ANES samples of the form

$$z_i = \alpha + X_i \gamma^{ANES} + e_i \tag{1}$$

where z are binary variables capturing support for the Republican presidential candidate in 2004, 2008, 2012, and 2016; and support for the Tea Party in 2010 and 2012. X_i is the vector of individual characteristics which overlap in the ANES and CPS: age, sex, race, marital status, veteran status, educational attainment, labor force status and family income.[27] We then use the estimated coefficients, $\gamma^{\hat{ANES}}$, from the regression along with individual characteristics of the reference person in each household from the CPS (the person whose name the housing unit is owned or rented

[26]The 2010 supplemental survey, referred to as the Evaluation of Government and Society Study, was administered in October 2010.

[27]Family income is in 2018 dollars. For 2010, household income is instead used in the analysis.

who usually responds to the survey), X_i, to predict the probability of z_i, $\widehat{Pr(z)}$ in year t.[28] The construction of the prediction using an OLS regression also serves to address measurement error in the ANES response (Hyslop and Imbens, 2001). In particular, although our predicted partisanship is a mismeasured proxy variable, the measurement error is uncorrelated with the proxy and thus does not bias our estimates.

In Table 1 we report the coefficient estimates in the predictive model, Equation 1, for GOP support in the presidential years 2004 through 2016, and Tea Party support in 2010 and 2012. The coefficients are broadly stable over time with several exceptions. Blacks and Hispanics are less likely than the omitted group, whites, to support either the GOP or Tea Party. Marriage is positively associated with GOP and Tea Party support. Women are less likely to support the Tea Party and GOP in 2016. Other variables make smaller and less consistent contributions to the predictions. Much of our subsequent analysis focuses on GOP support in 2008 (McCain) and 2016 (Trump), as well as Tea Party support in 2010. However, given the similarities of the coefficients across the predictions, we should not expect large differences between the models.

5.1 Main Household-level Results

We begin with a graphical analysis in Figure 7, which depicts the response patterns in the Current Population survey based on the quartiles of predicted support, $\widehat{Pr(z)}$, for McCain (Panel A), the Tea Party (Panel B), and Trump (Panel C) estimated with Equation 1. The predictive model has a binary outcome, meaning the quartiles least likely to support the GOP candidate are the most likely to support the Democratic candidate. As with the state and metro results, we find that refusal rates have risen across all quartiles, though there is noteworthy heterogeneity.

Starting in 2010, refusal rates begin moving distinctly upwards across all quartiles and all three political measures. There is modest convergence in levels between 2009 and 2012. However, the baseline refusal rate was quite low for the most-likely GOP supporters and thus there was

[28] According to U.S. Census Bureau (2006), every effort is made to interview the same respondent (usually the reference person) every month.

considerably stronger convergence in percentage terms, with refusal rates doubling among the most likely GOP and Tea Party supporters during the first four years of Obama's first term. These results broadly correspond to the surge in refusals noted by Krueger, Mas and Niu (2017).

Beginning around 2012, there is a fanning out of refusal rates across quartiles. The divergence is strongest for McCain and Trump support, though it appears in the Tea Party classification as well. The fanning out becomes distinctly stronger after 2014, especially for the quartile least likely to support the GOP or Tea Party. Again, these results are consistent with the state and metro-level results, and suggest a slow-moving political cycle in refusal rates. Crucially, we have not incorporated on any geographic element in the predictive model, so these estimates reflect a distinct source of variation from the results presented in Section 4.

We next quantify the relationship between CPS refusal in the 2000 to 2017 cross-sections and the predicted ANES variables. We estimate regression models of the form

$$RefuseRate_{iym} = \delta_0 \widehat{Pr(z_i)} + \sum_{y=2006}^{2017} \delta_y (\widehat{Pr(z_i)} \times Cohort_y) + \phi_{ym} + \varepsilon_{iym} \tag{2}$$

where $Cohort_y$ is the dummy for starting the first CPS interview in year y. ϕ_{ym} is the survey cohort fixed effects (year-month of first interview) that capture the overall time trend of refusals. We fully interact the predicted ANES variables, $\widehat{Pr(z_i)}$ with $Cohort_y$ to allow their effects to vary across different years. Thus, δ_y is the effect of the predicted ANES variables for cohort y, relative to the average baseline effect in cohorts 2000 to 2005.[29] Since $\widehat{Pr(z_i)}$ is a particular linear combination of the covariates, X_{iym}, we also estimate a less-restrictive model in which we replace the main effect of $\widehat{Pr(z_i)}$ in the regression with the vector of covariates themselves.

Table 2 presents the estimates for the household-level model, Equation 2. The predicted probability of voting for John McCain is the ANES variable used in columns 1 and 2; the predicted

[29] When predicting the probability of z_i for each household, we use the first non-missing characteristics obtained from the completed interviews and drop households with missing or imputed values for entire rounds of surveys. Family income in the CPS has a somewhat higher item non-response rate than other variables, so we also estimated a specification in which we included households in which family income is missing or imputed in all periods and included an indicator variable, rather then their family income, in the model. The results are largely unchanged by this alternative specification.

probability of supporting the Tea Party is used in columns 3 and 4; and the predicted probability of voting for Donald Trump is used in columns 5 and 6. The odd-numbered columns control for the main effect of the predicted ANES variable. The even-numbered columns replace the ANES variable with the the vector of covariates. Our conclusions below are largely insentive to this specification choice.

In the first row of the odd columns we can see that GOP and Tea Party support are associated with lower refusal rates in the omitted 2000 to 2005 pre-period. For example, voting for John McCain is associated with a reduction of 1.58 percentage points in the propensity to refuse to respond to the CPS in the 2000 to 2005 period. These results correspond to the level differences between red and blue areas in Section 4. Households which we predict to support the GOP and Tea Party have a 1 to 2 percent lower refusal rate in the pre- period, with the largest difference associated with Tea Party support.

The interaction coefficients indicate changes across cohorts in the association between survey refusal and political support for the GOP or Tea Party. All three political measures indicate that GOP/Tea Party supporters who began the CPS survey in 2006 through 2008 were less likely to refuse to complete the survey, compared to those in the 2000 to 2005 cohorts. That is, during the George W. Bush presidency likely-GOP voters were more likely to participate in the CPS, an effect that became stronger throughout his tenure. The 2009 CPS cohort – the first of the Obama presidency – begins a reversal, with the relative GOP and Tea Party refusal rates rising. By the 2010 cohort the point estimates are positive and sometimes statistically significant.

Beginning perhaps as early as 2011 the estimates show a remarkable departure from the previous patterns as the gap in refusals between Republican and Democrat supporters widends significantly. For example, GOP/Tea Party supporters in the 2015 cohort have refusal rates of 2.48 to 4.46 percentage points lower than similar people in the 2000 to 2005 cohorts, with the largest estimates based on Tea Party support. The size of the gap more-or-less doubles in every cohort between 2014 and 2017, reaching six to ten percentage points in the 2017 cohort, depending on the measure. (It is important to keep in mind that the average refusal rate approximately triples over this time period, so these point estimates would be less striking if reported in percentage terms.)

In Table 3 we take a longer view of the political cycle in refusal. Using CPS data from 1994 to 2017, this table reports the results of an analysis in which we regress household refusal rates on the ANES variable for GOP support and an interaction with an indicator for a Republican in power; year effects are included in all specifications. The base term for GOP support shows that Republicans have, on average, been over two percent less likely to refuse over the sample period. The interaction term reveals that the gap between Republican and Democratic refusal rates increases by around one percentage point when a Republican is president. Thus, the political cycle we documented in Section 4 and Figure 4 also appears when using demographic characteristics, rather than geography, to predict partisan affiliation.

Does the recent surge in refusals since Donald Trump's election reflect partisanship or an increasing trend that happen to coincide the 2016 election? To answer this we "zoom in" on the Trump election. There was considerable uncertainty about the outcome of this election and Trump's victory conveyed unexpected information about partisan control of the government after 2016. To leverage this variation, we conduct a difference-in-difference analysis focusing on the change in the refusal rate between the first four interviews and last four interview for the cohorts that straddle the 2016 election. We also report the analogous change in refusal rates for the same month-cohorts from the preceding two years. Standard errors are bootstrapped to address the presence of predicted support, a generated regressor.[30]

The results in Table 4 and Figure 8 show that the Trump election was associated with a growing divergence between GOP and Democratic supporters. For cohorts that began the CPS in December 2015 through July 2016, refusal rates among those we predict to have voted for Trump were about 1.2 percentage points lower in interviews conducted after the 2016 selection. By comparison, predicted Trump voters who began the CPS between December 2013 and July 2014, or between December 2014 and July 2015, had no change in response rates after November of 2014 or 2015.

[30] To increase precision, we do not drop households with imputed family income (all households with missing family income in the CPS are assigned imputed income after 2010). Dropping those households gives similar results with larger standard errors.

In column 4 we pool together data from all three cohorts and run a differences-in-differences model to get a standard error on the differences in responses across cohorts. The estimates indicate that the nearly 1.4 percentage point decline among the cohort that began between December 2015 and July 2016 is indeed statistically different from the two preceding cohorts.

5.2 Decomposition: Who are the "new refusers"?

In this subsection we perform a Blinder-Kitagawa-Oaxaca decomposition to better understand the growth in the refusal rate using the same sample of sometimes- and never-refuse households. (Blinder, 1973; Kitagawa, 1955; Oaxaca, 1973) This model is equivalent to an unrestricted version of Equation 2, where instead of imposing that $Pr(z_i) = X_i\gamma^{ANES}$, we allow the regression to choose the γ vector that bets fits the data. We also drop the requirement that variables appear in both the ANES and CPS, and expand the set of characteristics to include CPS-only variables, such as household type or metropolitan status. To simplify the presentation of results, we pool 1994-2009 as a pre-period and consider the total change between 1994-2009 and 2010-2017.

The decomposition is informative about three elements. First, the decomposition tells us which variables are most associated with growth in refusals. This type of analysis is only possible because of the short panel design of the CPS. Second, by comparing the coefficients estimated for γ^{ANES} to those from an unrestricted model, we can describe the forces that are captured, or not, by the restricted political model above. Finally, the results also tell us about the share of variation that could possibly be explained by the X's, as the unrestricted model is an upper bound on the share of the variance that can be explained by the restricted model.

The decomposition results using the household characteristics (X) and the above regression coefficients (β) are shown in Table 5. We first document the mean household characteristics (X) of cohorts in 1994 to 2009 (column 1) and 2010 to 2017 (column 2), and show the difference between the two cohorts, ΔX, in column 3. Households became older, more Hispanic, more educated, less employed and more metropolitan. Are these changes in the household characteristics large enough to explain the increase in refusal rates? To answer this question, in column 7 we multiply the change in household characteristics (column 3) by the coefficients (β) that are estimated from the baseline cohorts in 1994 to 2009 (column 4). The penultimate row shows that the sum of $\Delta X\dot{\beta}$ equals -0.046. That is, the change in characteristics would predict a slight decline in refusals between the two time periods.[31] The results are not surprising given that there was no dramatic change in demography over the time periods.

We next focus on how much the changes in the coefficients of household characteristics (β) can explain the rise in refusal rates. This part of the decomposition can be interpreted as the unrestricted version of the previous ANES proxy analysis, where we allow the regression of refusal rates on characteristics to choose the coefficients that best predict the data, rather than first creating a linear combination of the characteristics that best predicts political preferences. The estimated regression coefficients for cohorts 1994 to 2009 and 2010 to 2017 are presented in columns 4 and 5 (without standard errors[32]). In particular, households with higher refusal rates tend to be non-white, native-born, and working-age individuals. Households with higher education are less likely to refuse than those with only a high school degree or some college education. People out of the labor force are less likely to refuse than those who are either unemployed or employed. Households with higher family income are less likely to refuse.

The difference in the regression coefficients between the two periods is shown in column 6. We find that those 55 and older (relative to those 30 to 54), Asians (relative to whites), and the foreign born are less likely to refuse, whereas blacks (relative to whites), high school graduates (relative to those with a BA or more), those who are employed (or enrolled in school) or unemployed (relative to not in labor force) and households with lower family income are more likely to refuse. Column 8 shows the results from multiplying the change in the coefficients (column 6), $\Delta\beta$, by the mean characteristics of the baseline cohorts, 1994-2009 (column 7). Notably, the coefficients of household size, the share age 55 or older and family income predicts *decreases* in the refusal rates. On the contrary, share high school graduates, share employed/enrolled and in metropolitan areas

[31] When using β in 2010 to 2017 as the baseline, the total change explained by the change in X is -0.150 percentage point.

[32] We find that most of the characteristics are statistically significantly associated with refusals. For a full list of estimated coefficients and their standard errors, refer to Appendix Table A1.

predict *increases* in the refusal rates. The share employed/enrolled itself explains 0.9 percentage point of the 2.1 percentage point increase. In total, these changes in β can explain 0.7 percentage point of the increase in the refusal rates, which is about 35% of the actual change in refusal rates.[33]

The estimates of the characteristics of households that drive the rising trend may be of interest as descriptive measures, independent of our main hypothesis. A number of previous studies examine the characteristics of individuals and households that refuse surveys, often relying on soft refusals, as we do here.[34]. Fewer studies examine the change in nonresponse or refusal over time. One such study, Brick and Williams (2013) finds that growing nonresponse rates between 1997 and 2007 in four surveys (the National Heatlh Interview Survey, the General Social Survey, and National Household Education Survey, and the National Immunization Survey) are negatively associated with households with children under 6, and local violent crime rate (surprisingly, lower violent crime rates are associated with faster growth in nonresponse), and positively associated with single-person households and the travel time to work. By comparison (using the results in Table 5, we find no evidence that the presence of children or household structure contributes to the recent growth in refusals. We do not include local violent crime rates in our regressions, but do not find that growth in refusals correlates in the proposed direction with other measures likely to be associated with violent crime, such as being in a metro area or demographic variables related to household structure and race/ethnicity. Finally, the result on transit time is quite interesting, since it relates to the growing positive association between employment and refusal. Thus, it may be the effect of employment on refusals has been growing for some time and across many surveys.

The results from this unrestricted decomposition can also highlight what variables were underweighted in our ANES analysis above. A comparison of the unrestricted βs in the decomposition to the ANES models in Table 1 indicates that the most important difference is the role of the employment measure. The share of the household that are employed has the largest explanatory power

[33] When using X in 2010-2017 as the baseline, the total change explained by the change in β is 0.636 percentage points.

[34] In the CPS context, see Abraham, Maitland and Bianchi 2006; Abraham, Helms and Presser 2009; Fricker and Tourangeau 2010; Brault 2014; Bee, Gathright and Meyer 2015. In other contexts, examples include Heffetz and Rabin 2013; Heffetz and Reeves 2019; Cheng, Zamarro and Orriens 2016 and Meyer, Mok and Sullivan 2015

in our decomposition. By contrast, employment is a weak predictor of partisanship in the ANES prediction.

More broadly, the Blinder-Kitagawa-Oaxaca decomposition provides important information on the ability of reweighting procedures to correct for the potential biases that arise from the rise in refusals (Kline, 2011). Our results in Table 5 indicate that less than half of the rise in the refusal rate is explained by changes in observable variables and changes in the correlation between these variables and refusals. A majority of the rise is due to unobserved individual and household characteristics that are poorly proxied by the observed characteristics. Thus, reweighting respondents based on observed variables is unlikely to produce a sample that fully characterizes the non-responders.

6 Local Tea Party Activism and Refusal Rates

So far we have examined the relationship between refusal rates and geographic and household characteristic-based predictors of political affiliation. We have shown that rising refusal rates are not monotonically related to Republican or Tea Party support. Instead, refusal rates have risen in both Republican and Democratic-leaning geographic areas and households, with relatively faster growth in Democratic areas after 2014, growth which accelerated after Trump's election. Since these estimates rely on panel-based identification, a natural concern is that they reflect coincident changes that are correlated with, but not caused by, political preferences. As discussed in Section 2, both Census refusal and skepticism of the integrity of government statistics were strongly associated with the leadership of the Tea Party. With this in mind, in this section we estimate instrumental variables models which more narrowly examine the causal effect of Tea Party engagement on CPS response rates.

We follow Madestam et al. (2013) and leverage rainfall on the day of the first Tea Party rally as an predictor of Tea Party support in a geographic area. They show that good weather on on the day of this first rally increased the size of the rally, which eventually led to increased public support for the Tea Party. While this analysis resembles an instrumental variables setting, it is somewhat unclear how to define the first stage because there could be a number of causal channels link rain to refusals, including new information, increased political engagement, and altered political views. Given this uncertainty, we believe the long-run evidence is most informative about a reduced-form "push" of Tea Party messaging into a metropolitan area. We thus focus on reduced form estimates before discussing the scaling and interpretation of magnitudes.

We use the CPS survey cohorts from 2005 to 2017 to estimate the effect of rain on the day of the Tea Party rally on refusal rates in the following regression:

$$RefuseRate_{iyma} = \delta_0 RainyRally_a + \sum_{y=2006}^{2017} \delta_y (RainyRally_a \times Cohort_y) + ProbRain_a + X_{iyma}\beta + \phi_{ym} + \varepsilon_{iyma} \quad (3)$$

where *RainyRally* is a dummy equal to 1 if there was more than 0.1 inch of rain in the metropolitan area a on the day of the Tea Party rally (Apr 15, 2009). As in the preceding sections, we allow the effects of a rainy rally to vary across survey cohort by interacting *RainyRally* with survey year cohort fixed effects, $Cohort_y$. We additionally control for the probability of rain in April, $ProbRain_a$, household characteristics, X, and survey cohort fixed effects, ϕ_{ym}.[35] The coefficient of interest, δ_y, identifies the effects of the rainy Tea Party Rally in survey year y, relative to the effect in the baseline survey year of 2005. The household characteristics are the same as those that appear in Table 5. Identification in this model hinges on the question of whether rain is as good as randomly assigned conditional on the probability of rain, cohort effects (which capture a national time trend), and the household characteristics. We test this assumption below.

The results appear in Table 6. The coefficients in the first row indicate the effect of a rainly Tea Party rally on refusals among the 2005 cohort. Each subsequent coefficient represents the interaction of the year with the rainy rally indicator, δ_y in Equation 3. In column 1, we report results for a larger sample in which we include always-refuse households (and thus must omit the

[35]The probability of rain is at the metropolitan area level and this data is obtained from XYZ.

controls for household characteristics). In columns 2 and 3 we drop the always-refuse households and (in column 3) include household-level controls in the model. We drop non-metro areas in all columns due to the limitations of the CPS geographic identifiers. The top four rows, corresponding to the 2005 through 2008 cohorts, represent the pre-period before rainfall on April 15, 2009. Future rainfall is not predictive of refusal in these years, as would be expected if the rainy rally indicator is as good as randomly assigned.

The main results begin with the 2009 cohort. This cohort, which responds to the survey between Jan 2009 and March 2011, shows an increase in refusal rates in response to a rainy rally. The result is of similar magnitude across the columns and shows that refusal rates were about 0.36 to 0.45 percentage points higher among households in this cohort who lived in areas where there was rain on the day of the Tea Party rally. That is, the causal effect of the Tea Party was to reduce refusal rates in areas where rallies were particularly successful. These estimates have time fixed effects, so they should be interpreted relative to a growing trend: as we showed above, areas with strong Tea Party support experienced slower growth in refusal rates.

Looking at the later years, the effect of the Tea Party returns in 2012, then shows a steady rise after 2014. By the 2017 cohort, a rainy rally in 2009 predicts about a two percentage point higher refusal rates. The estimates in the final years of the panel are about one-fifth as large as those in Table 2, suggesting that a rainy rally offsets a meaningful portion of the Tea Party-related decrease in refusal rates. Note that these results are larger than the 2009 estimates in level terms, but of similar magnitudes in percentages. Assuming that the 2009 rainfall variation is truly exogenous, it appears that exposure to the Tea Party has had a lasting causal effect on refusal rates.

The pattern of results is consistent with those in our panel analysis above, and we found no indication of pre-trends, we may still be concerned that the relationship between refusals and rain on the day of the Tea Party rally reflects a spurious correlation between rain and some characteristics of these areas that predict refusal. One indirect test of instrument exogeneity is reported in column 3, in which we add household characteristics to the set of controls. In Section 5.2 we document that these variables explain a significant portion of rising refusals. Their inclusion in column 3 has no effect on the estimated coefficients, consistent with the observables being uncorrelated with omitted variables.

How should the magnitude of these results be interpreted? Given an estimate of a "first stage," these results can be interpreted as local average treatment effects on the group of compliers, who are in this case those whose exposure to anti-survey rhetoric is manipulated by the rain-influenced size of Tea Party rallies on April 15, 2009. Ideally, the first stage would tell us the size of the group of people who are exposed to the anti-survey messaging, and how much the Tea Party rally contributed to it.[36] The most immediate group of affected individuals is likely those people who did not attend the rally because of the rain; Madestam et al. (2013) reports that rallies were about 10,000 people on average, which is 0.082 percent of the county population. It appears, however, that these effects grew quite dramatically over the 18 months between the rally and the 2010 election, in which rainy rally areas had 1.04 percent (as a share of the voting age population) smaller Republican vote shares, and 5.7 percent lower Tea Party support (estimated by Madestam et al. (2013) in the ANES data). If we consider these as various measures of the first stage, we could then divide the estimates in Table 6 by the share of the population affected to estimate the treatment on the treated. Such estimates would be at best suggestive, as the treatment on the treated estimates vary significantly and is well above one for the most narrow definition of treated individuals. An additional complication is that Madestam et al also shows an effect on the leadership of the Tea Party, suggesting that intensity of support was also manipulated.

The interpretation of results for the 2012 through 2017 cohorts depends on how the Tea Party and its political descendants have influenced attitudes towards survey responses. Given that the Tea Party itself has largely dissolved, it is somewhat unclear how to define the treated group in later years. The group of affected individuals could expand over time if the survey-rhetoric messaging reaches a larger audience, beyond the initial Tea Party members. The affected group could also contract over time if the least-affected people (possibly even the marginal rally-goers) becomes less

[36] One consequence of the local average treatment effect interpretation is that we are not necessarily estimating the same quantity as in the previous ANES-proxy analysis, in which we examined the behavior based on the likelihood of supporting the Tea Party, including among inframarginal regarding their support for the Tea Party.

affected by the Tea Party and its political descendants. We show in the Appendix that the effect of a rainy rally on GOP vote shares is not statistically different from zero in later years. However, the confidence intervals on these estimates do not exclude relevant effect sizes. Thus, it seems likely that longer-term effects of rain on the day of the rally do not reflect longer-term effects on political preferences.

It is surprising that the Tea Party messaging, which included elements of rhetoric encouraging adherents to distrust government data collection, led to lower refusal rates. One possibility is that the main effect of the Tea Party was to increase patriotism and hence likelihood to comply with the request for data, while the rhetoric regarding survey refusal was less prominent. Similarly, studies in psychology find that Republicans demonstrate greater adherence to authority (Jost, West and Gosling, 2009; Womick et al., 2019). As the rainy rally analysis isolates effects on marginal individuals, this story would still leave room for inframarginal Tea Party supporters to increase refusal rates. One problem with this story is, however, the time pattern of results. If the Tea Party increased patriotism or adherence to authority, it is unclear why these patterns disappear before re-emerging in the later years.

An alternative possibility is that the Tea Party messaging increased the salience of government surveys. The leverage-salience theory of survey participation posits that respondents are less likely to refuse when they perceive a reason to participate (i.e. they have leverage), and this is made clear to them before the administration of the survey (i.e. the importance of the survey is salient) (Groves, Singer and Corning, 2000). In the theory, salience is often conveyed by the enumerator, though in the Tea Party case, salience may have been communicated by political rhetoric. This theory is closely related to benefit-cost theories of survey participation (e.g. Singer (2011)). Political engagement has been show to predict survey response (Keeter et al., 2006). By tying survey response to political outcomes, such as assessment of the Obama administration's economic policies, the Tea Party may have also increased the perceived value of time spent responding. Indeed, many Republicans responded to the "irresponsible" call to boycott the 2010 Census by emphasizing the government funding at stake.[37] Recent work shows that relatively few people are aware of the government funds tied to Census response, and response rates rise when this information is conveyed. That survey respondents consider the salience of the survey has been raised in previous work, but has not previously, to the best of our knowledge, been quasi-experimentally tested in a large survey.

A final possibility, of course, is that the instrument is invalid. Daily rainfall conditional on the long-run average is likely random, however, it is possible that rainfall on the specific date of April 15, 2009 was correlated with area characteristics by chance. Our indirect tests of exogeneity provide some assurance that this is not the case. Some additional confirmation comes from the correspondence of the rainy rally results with the panel-based approach after 2014. These analyses use very different sources of identification, and the agreement between them greatly reduces the probability that the findings are spurious. While the evidence is consistent with an effect of political cycles on response rates, the strongest conclusions are likely warranted in 2009 and 2010, in the immediate aftermath of the rainy rally treatment. The results in these years support the conclusion that the Tea Party reduced refusal rates in areas with particularly strong rallies.

7 Conclusion

Survey refusal rates have been on the rise across most surveys for several decades. The Current Population Survey, previously an outlier in its maintenance of low refusal and non-reponse rates, has experienced rapid growth in refusal rates since 2010. We find that the growth in CPS refusal is not explained by the coincident political rhetoric questioning the value and integrity of government data collection. Instead, patterns in survey refusal support a modest political cycle dating back several decades. The political cycle appears to have grown modestly larger over time, to the point where it is detectable in within-household patterns of survey response before and after Trump's election.

[37] Amy Sullivan. 2009. "Why the 2010 Census Stirs Up Partisan Politics." New York Times. Feb 15. http://content.time.com/time/nation/article/0,8599,1879667,00.html.

We also examine local Tea Party activism, leveraging variation in rainfall on the day of the initial Tea Party rally. Surprisingly, we find that Tea Party activism decreased refusal rates both around the time of the initial rally, and also in the lead-up and following the 2016 election. These results are consistent with an increase in salience or perceived value in responding to the CPS.

Despite the broad awareness of survey response trends and their implications among survey researchers, there are few causal studies of the theories put forward for why households refuse to fill out surveys. There are several alternative explanations for the rise in CPS refusals that, we believe, merit greater consideration. Other less-explored reasons for the recent rise are concerns about data security, changes in trust in institutions, and survey fatigue. Finally, we find that one of the most important predictors of refusal is the households' employment share, suggesting that some households are simply too busy to respond to the CPS.

In the last few years issues related to non-response, refusal, misreporting, and the integrity of statistical surveys have become an important part of public discourse. Events peaked around and after the 2016 election, when pollsters are thought to have underestimated Trump's chances of winning due to a combination of voters reluctance to respond and respond truthfully to public opinion polls, and biased methodologies or herding behavior on the part of the pollsters. Following the election, the Trump Administration announced plans to add a citizenship question, raising concerns about non-response in immigrant and Hispanic communities. Our work indicates that the politicization of government surveys is an important, and potentially growing, problem.

References

Abraham, Katharine G., Aaron Maitland, and Suzanne M. Bianchi. 2006. “Nonresponse in the American time use survey: Who is missing from the data and how much does it matter?” *Public Opinion Quarterly*, 70(5): 676–703.

Abraham, Katharine G., Sara Helms, and Stanley Presser. 2009. “How Social Processes Distort Measurement: The Impact of Survey Nonresponse on Estimates of Volunteer Work in the United States.” *American Journal of Sociology*, 114(4): 1129–1165.

Anderson, Margo. 2015. “Public Management of Big Data: Historical Lessons from the 1940s.” *Fed. Hist.*, 7: 17.

Bee, C. Adam, Graton M. R. Gathright, and Bruce D. Meyer. 2015. “Bias from Unit Non-Response in the Measurement of Income in Household Surveys.”

Blinder, Alan S. 1973. “Board of Regents of the University of Wisconsin System Wage Discrimination : Reduced Form and Structural Estimates.” *The Journal of Human Resources*, 8(4): 436–455.

Bollinger, Christopher R., Barry T. Hirsch, Charles M. Hokayem, and James P. Ziliak. 2019. “Trouble in the Tails? What We Know about Earnings Nonresponse 30 Years after Lillard, Smith, and Welch.” *Journal of Political Economy*, 127(5): 2143–2185.

Boxell, Levi, Matthew Gentzkow, and Jesse M Shapiro. 2017. “Greater Internet use is not associated with faster growth in political polarization among US demographic groups.” *Proceedings of the National Academy of Sciences*, 114(40): 10612–10617.

Brault, Matthew W. 2014. “Non-response Bias in the 2013 CPS ASEC Content Test.”

Brick, J. Michael, and Douglas Williams. 2013. “Explaining Rising Nonresponse Rates in Cross-Sectional Surveys.” *Annals of the American Academy of Political and Social Science*, 645(1): 36–59.

Card, David. 2014. “Origins of the Unemployment Rate : The Lasting Legacy of Measurement without Theory.” *American Economic Review: Papers and Proceedings*, 101(3): 552–557.

Center for Economic and Policy Research. 2019. “CPS ORG Uniform Extracts, Version 2.4.”

Cheng, Albert, Gema Zamarro, and Bart Orriens. 2016. “Personality as a Predictor of Unit Nonresponse in Panel Data: An Analysis of an Internet-Based Survey.”

Curtin, Richard, Stanley Presser, and Eleanor Singer. 2005. “Changes in telephone survey nonresponse over the past quarter century.” *Public opinion quarterly*, 69(1): 87–98.

De Heer, W, and E De Leeuw. 2002. “Trends in household survey nonresponse: A longitudinal and international comparison.” *Survey nonresponse*, 41: 41–54.

de Leeuw, Edith, Joop Hox, and Annemieke Luiten. 2018. "International nonresponse trends across countries and years: An analysis of 36 years of labour force survey data." *Survey Methods: Insights from the Field*, 1–11.

Faye, Michael, and Paul Niehaus. 2012. "Political aid cycles." *American Economic Review*, 102(7): 3516–30.

Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, and J. Robert Warren. 2018. "Integrated Public Use Microdata Series, Current Population Survey: Version 6.0." Minneapolis, MN:IPUMS, 2018.

Fricker, Scott, and Roger Tourangeau. 2010. "Examining the relationship between nonresponse propensity and data quality in two national household surveys." *Public Opinion Quarterly*, 74(5): 934–955.

Gelman, Andrew, Sharad Goel, Douglas Rivers, David Rothschild, et al. 2016. "The mythical swing voter." *Quarterly Journal of Political Science*, 11(1): 103–130.

Groves, Robert M., Eleanor Singer, and Amy Corning. 2000. "Leverage-Saliency Theory of Survey Participation: Description and an Illustration." *The Public Opinion Quarterly*, 64(3): 299–308.

Heffetz, Ori, and Daniel B Reeves. 2019. "Difficulty of Reaching Respondents and Nonresponse Bias: Evidence from Large Government Surveys." *Review of Economics and Statistics*, 101(1): 176–191.

Heffetz, Ori, and Matthew Rabin. 2013. "Conclusions regarding cross-group differences in happiness depend on difficulty of reaching respondents." *American Economic Review*, 103(7): 3001–21.

Hokayem, Charles, Christopher Bollinger, and James P. Ziliak. 2015. "The Role of CPS Nonresponse in the Measurement of Poverty." *Journal of the American Statistical Association*, 110(511): 935–945.

Hyslop, Dean R, and Guido W Imbens. 2001. "Bias from classical and other forms of measurement error." *Journal of Business and Economic Statistics*, 19(4): 475–481.

Jost, John T, Tessa V West, and Samuel D Gosling. 2009. "Personality and ideology as determinants of candidate preferences and "Obama conversion" in the 2008 US presidential election." *Du Bois Review: Social Science Research on Race*, 6(1): 103–124.

Keeter, Scott, Courtney Kennedy, Michael Dimock, Jonathan Best, and Peyton Craighill. 2006. "Gauging the impact of growing nonresponse on estimates from a national RDD telephone survey." *International Journal of Public Opinion Quarterly*, 70(5): 759–779.

Kitagawa, Evelyn M. 1955. "Components of a Difference Between Two Rates." *Journal of the American Statistical Association*, 5: 1168–1194.

Kline, Patrick. 2011. "Oaxaca-Blinder as a reweighting estimator." *American Economic Review: Papers and Proceedings*, 101(3): 532–37.

Krueger, Alan B., Alexandre Mas, and Xiaotong Niu. 2017. "The evolution of rotation group bias: Will the real unemployment rate please stand up?" *Review of Economics and Statistics*, 99(2): 258–264.

Madestam, Andreas, Daniel Shoag, Stan Veuger, and David Yanagizawa-Drott. 2013. "Do Political Protests Matter? Evidence from the Tea Party Movement." *The Quarterly Journal of Economics*, 128(4): 1633–1685.

Meyer, Bruce D, and Nikolas Mittag. 2019. "Using Linked Survey and Administrative Data To Better Measure Income :." *American Economic Journal: Applied Economics*, 11(2): 176–204.

Meyer, Bruce D., Wallace K. C. Mok, and James X. Sullivan. 2015. "Household Surveys in Crisis." *Journal of Economic Perspectives*, 29(4): 199–226.

National Research Council. 2013. *Nonresponse in social science surveys: A research agenda.* National Academies Press.

Oaxaca, Ronald. 1973. "Male-Female Wage Differentials in Urban Labor Markets." *International Economic Review*, 14(3): 693–709.

Pastor, Lubos, and Pietro Veronesi. 2017. "Political cycles and stock returns." National Bureau of Economic Research.

Pew Research Center. 2019. "Public Trust in Government: 1958-2019."

Santa-Clara, Pedro, and Rossen Valkanov. 2003. "The presidential puzzle: Political cycles and the stock market." *The Journal of Finance*, 58(5): 1841–1872.

Schoeni, Robert F, Frank Stafford, Katherine A McGonagle, and Patricia Andreski. 2013. "Response rates in national panel surveys." *The Annals of the American Academy of Political and Social Science*, 645(1): 60–87.

Shi, Min, and Jakob Svensson. 2006. "Political budget cycles: Do they differ across countries and why?" *Journal of public economics*, 90(8-9): 1367–1389.

Singer, Eleanor. 2011. "Toward a benefit-cost theory of survey participation: Evidence, further tests, and implications." *Journal of Official Statistics*, 27(2): 379–392.

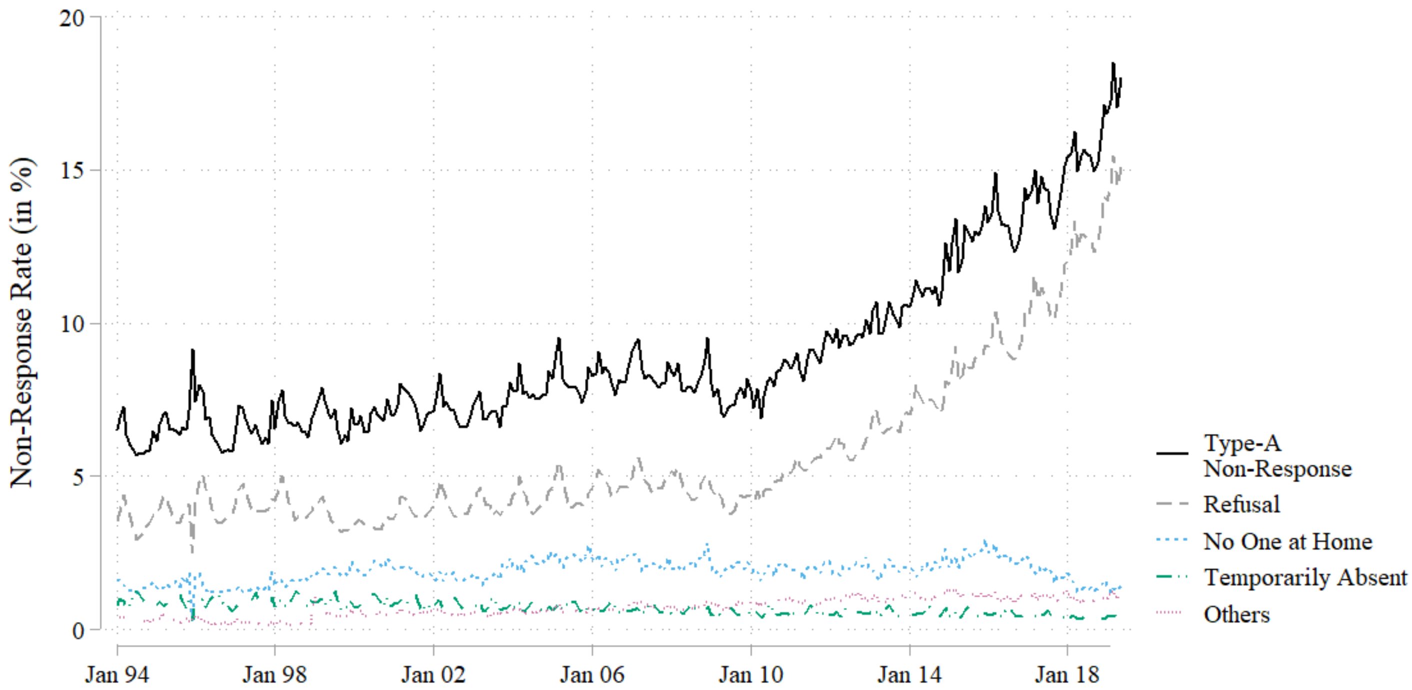
University of Michigan, Stanford University, and National Science Foundation. 2018. "The American National Election Studies."

U.S. Census Bureau. 2006. "Technical Paper 66: Current Population Survey: Design and Methodology." October.

Womick, Jake, Tobias Rothmund, Flavio Azevedo, Laura A King, and John T Jost. 2019. "Group-based dominance and authoritarian aggression predict support for Donald Trump in the 2016 US presidential election." *Social Psychological and Personality Science*, 1948550618778290.

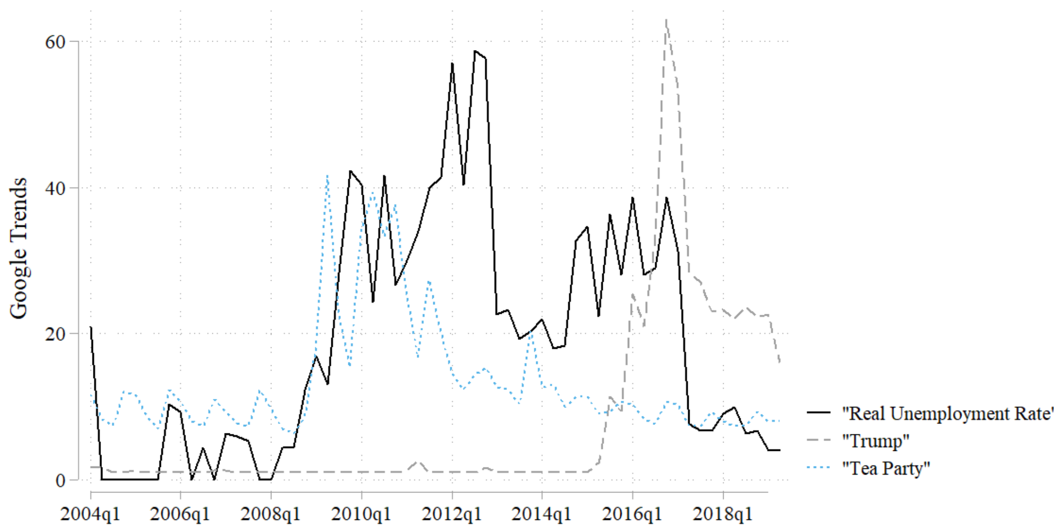
Figures and Tables

Figure 1: CPS Type-A Non Interview Trends



Notes: Type-A Non-interview rate is defined as (Number of Type-A Non Interviews)/(Number of Interviews+Number of Type A Non Interviews). Type-A nonresponse households represent housing units suitable for inclusion in the survey whose residents were not interviewed for reasons such as refusal to participate and temporary absence. Dec, 1995 and Jan, 1999 data have a lot of missing data on specific types of Type-A non-response, which results in a large drop in refusal rates. CPS added "language barrier" and "unable to locate" categories starting from 2010, and we include them in "others" category.

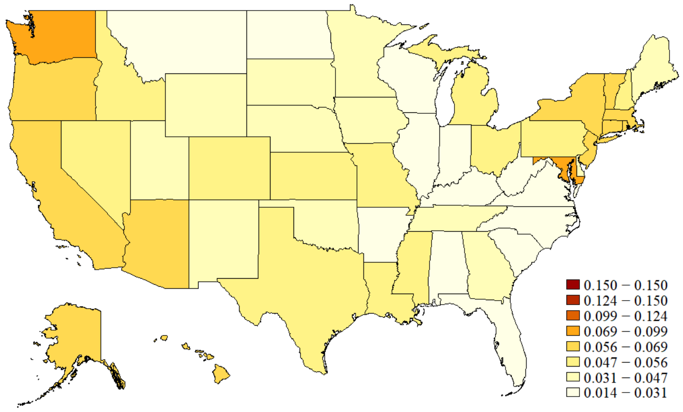
Figure 2: Google Trends



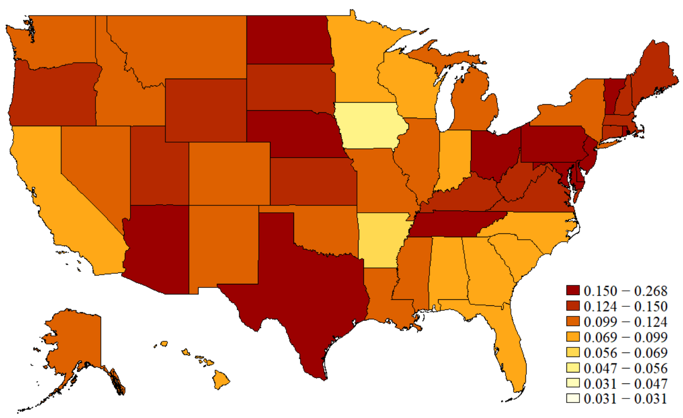
Notes: Monthly Google Trends data are collapsed by quarter. Taken from Google: "Numbers represent search interest relative to the highest point on the chart for the given region and time. A value of 100 is the peak popularity for the term."

Figure 3: State-level Refusal Rate

(a) 2010



(b) 2018



(c) Difference

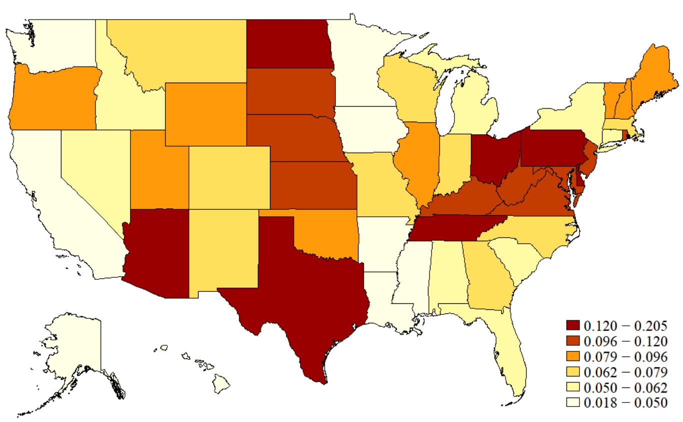
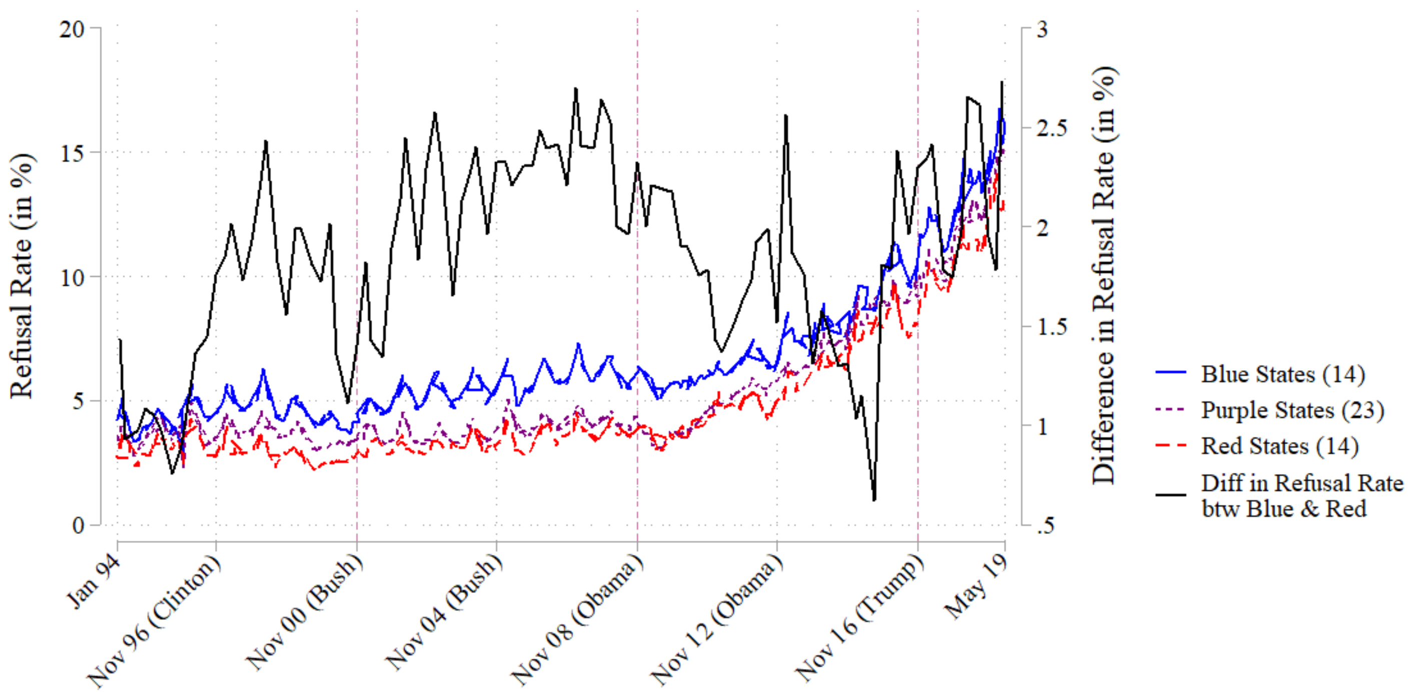


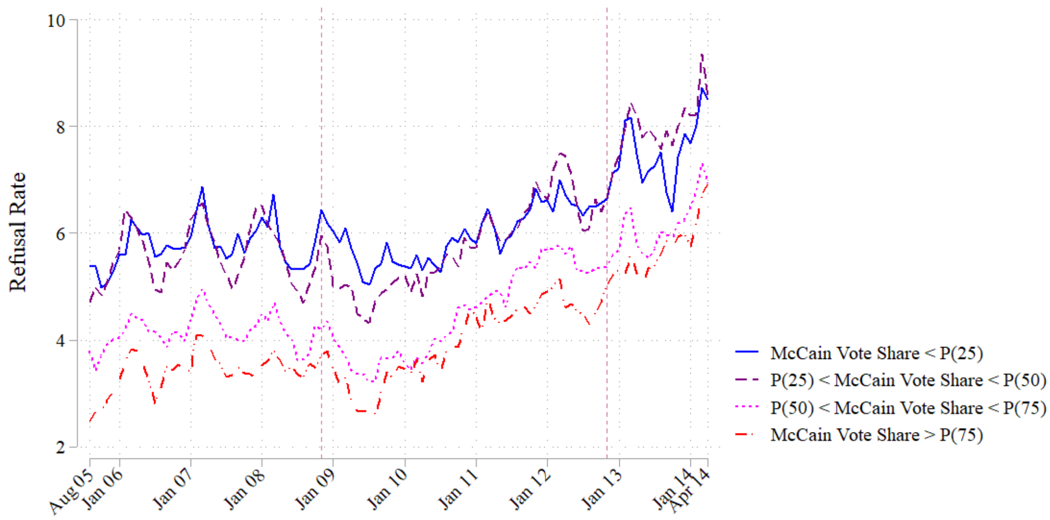
Figure 4: Non-Response Rates, Blue vs Red States



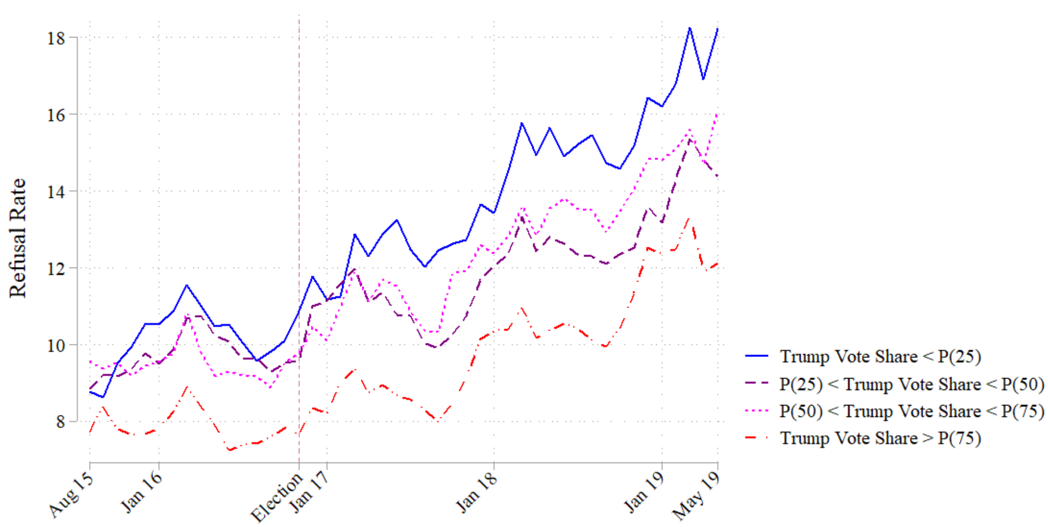
Notes: Blue States are 14 states (California, Connecticut, Delaware, District of Columbia, Hawaii, Illinois, Maryland, Massachusetts, New Jersey, New York, Oregon, Rhode Island, Vermont and Washington) where Democratic candidates have always won over Republican candidates by more than 10 percentage point margin (2008, 2012, 2016 elections). Red States are 14 states (Alabama, Alaska, Arkansas, Idaho, Kansas, Kentucky, Louisiana, Mississippi, Nebraska, Oklahoma, Tennessee, Utah, West Virginia and Wyoming) where Republican candidates have always over democratic candidates by more than 10 percentage point margin. The remaining states are purple states (23 states). The black line represents the difference in non-response rate between the blue and red states, collapsing by quarter.

Figure 5: Refusal Rate by Quartile of GOP Share (Metropolitan Areas)

(a) 2005-2014

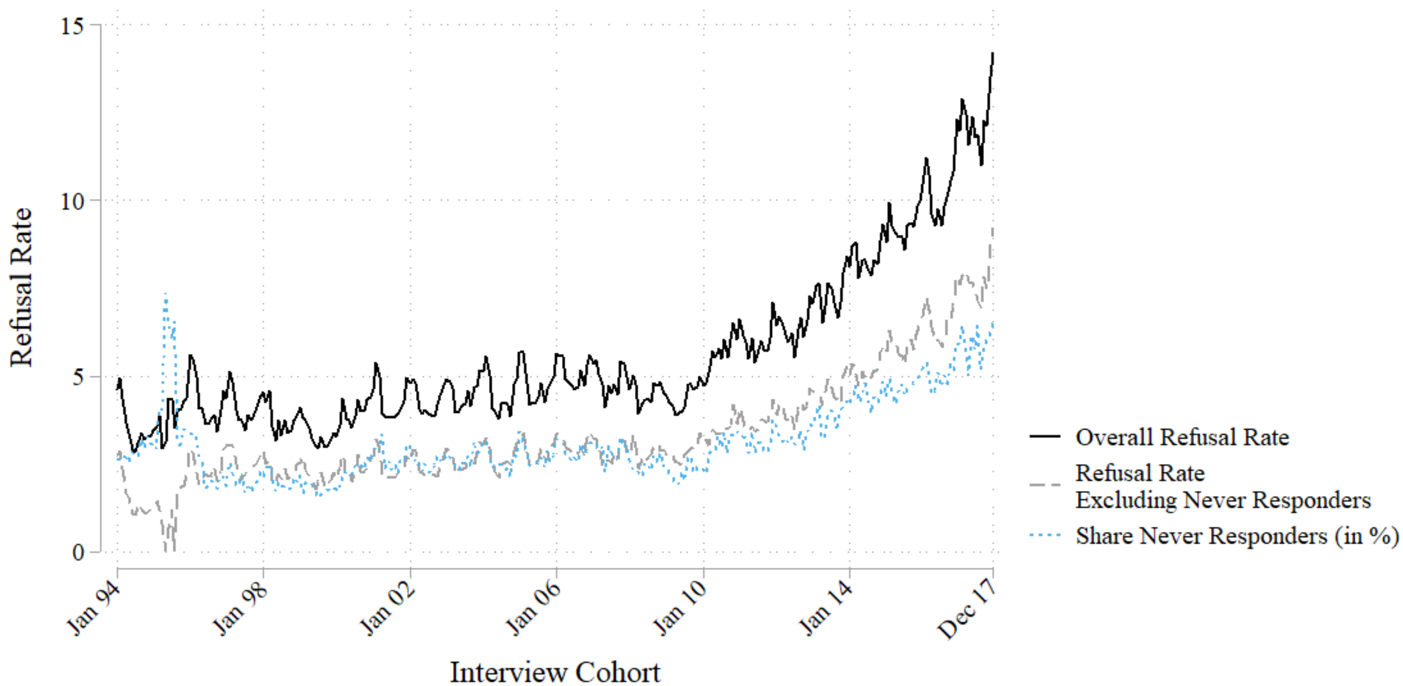


(b) 2015-2019



Notes: We stratify metropolitan areas by quartile of McCain vote share (2008) in Panel A and Trump vote share (2016) in Panel B. In Panel A, P(25)=0.35; P(50)=0.44; P(75)=0.54 In Panel B, P(25)=0.35; P(50)=0.49; P(75)=0.59. Non-metropolitan areas, identified by their states, are also included in the figures. We exclude non-metropolitan areas in Alaska. The vertical lines represent the time of elections (Nov 2008, Nov 2012 and Nov 2016)

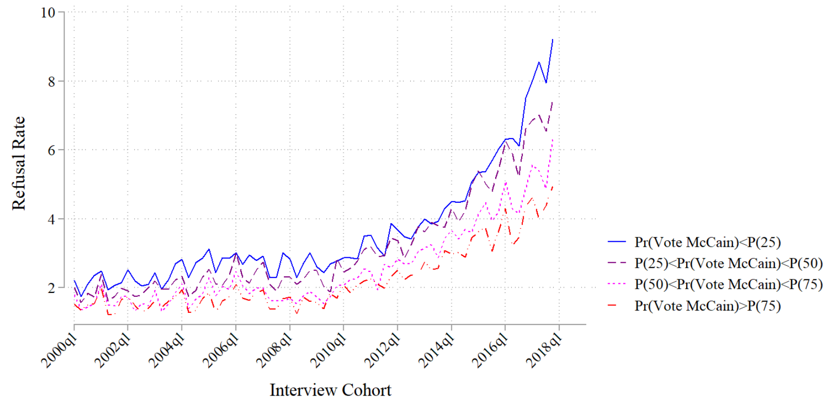
Figure 6: CPS Refusal Rates by Responder Type



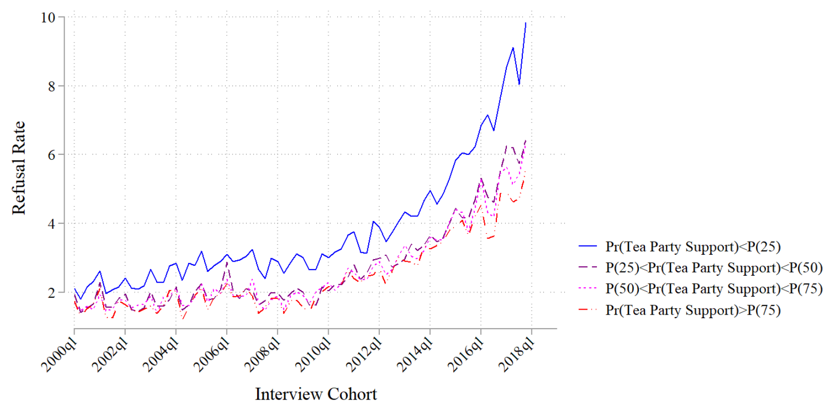
Notes: CPS interview cohorts are defined by the starting month-year that households start their first interviews. For example, Jan, 10 Cohorts are households who started their fist interviews in Jan, 10. Never responders are households which had Type-A non interviews throughout whole month in samples. Households who had their first interview in May 1996 were only interviewed once, which resulted in a big drop in the average number of times that households appear in the samples and surge in the share never responders.

Figure 7: Refusal Rate by Quartile of Predicted ANES Variables

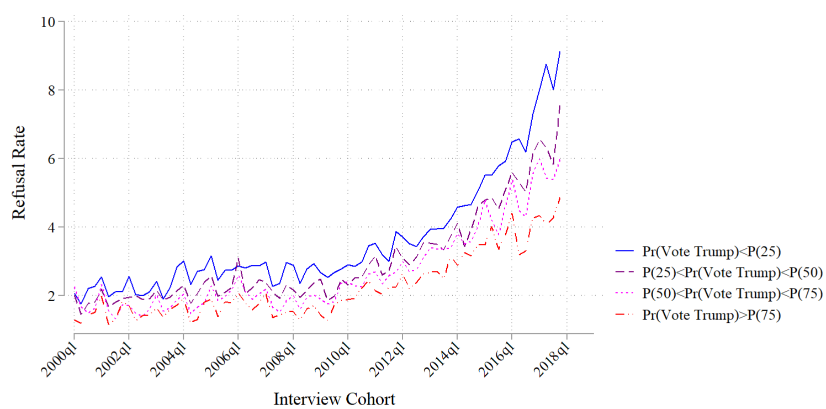
(a) Vote McCain (2008)



(b) Support Tea Party (2010)

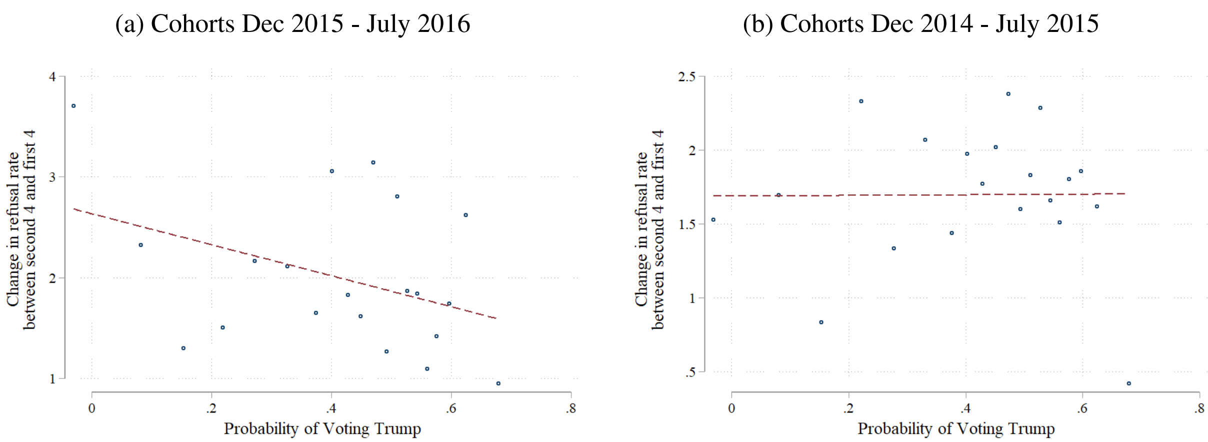


(c) Vote Trump (2016)



Notes: We predict the probability of voting McCain (Panel A), supporting Tea Party (Panel B) and voting Trump (Panel C) in the CPS samples using the coefficients estimated in Table 1. For each ANES variable, we stratify CPS households by quartile of the predicted probability and collapse them by interview cohort (year-quarter). When calculating the predicted probability, we drop households whose family income is missing or impute for all rounds of interviews.

Figure 8: Difference in Refusal Rate between Second and First 4 by Probability of Voting Trump



Notes: The unit of observations is interview cohort (year-month) by metropolitan area (ex: interview cohort Dec 2015 in Boston). The observations are bin-scattered. Y-axis is the difference in the average refusal rate between the second 4 interviews (MIS 5,6,7 and 8) and the first 4 interviews (MIS 1,2,3 and 4). The probability of voting Trump for each reference person in a household is obtained using the coefficients in column 4 of Table 1.

Table 1: Predicting Political Variables from ANES

	Vote/Prefer GOP Presidential Candidate				Support Tea Party	
	2004	2008	2012	2016	2010	2012
	(1)	(2)	(3)	(4)	(5)	(6)
Age	0.01**	-0.00	-0.00	0.00	0.01**	-0.00
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Age Squared	-0.00**	0.00	0.00	0.00	-0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Female	-0.05	-0.01	-0.00	-0.06***	-0.11***	-0.03***
	(0.03)	(0.02)	(0.01)	(0.02)	(0.03)	(0.01)
Black	-0.40***	-0.39***	-0.38***	-0.42***	-0.19***	-0.12***
	(0.04)	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)
Asian	-0.14	-0.20***	-0.12**	-0.20***	-0.07	0.06
	(0.10)	(0.07)	(0.05)	(0.04)	(0.13)	(0.05)
Other Race	-0.17**	-0.22***	-0.17***	-0.12***	0.05	-0.04*
	(0.09)	(0.08)	(0.03)	(0.04)	(0.06)	(0.02)
Hispanic	-0.09	-0.25***	-0.22***	-0.30***	-0.04	-0.08***
	(0.06)	(0.03)	(0.02)	(0.02)	(0.04)	(0.01)
Married, spouse present	0.10**	0.07**	0.11***	0.12***	-0.03	0.05***
	(0.04)	(0.03)	(0.02)	(0.02)	(0.03)	(0.01)
Married, spouse absent	0.01	0.01	0.04**	0.08***	-0.04	-0.01
	(0.05)	(0.03)	(0.02)	(0.03)	(0.04)	(0.01)
Have Kids	0.01	0.01	0.02	0.01	0.08***	-0.00
	(0.04)	(0.02)	(0.01)	(0.02)	(0.03)	(0.01)
Veteran	0.00	0.04	0.04**	0.05*	-0.02	0.04**
	(0.05)	(0.03)	(0.02)	(0.03)	(0.04)	(0.02)
Some College	-0.01	0.06*	0.05***	0.02	0.01	0.03***
	(0.05)	(0.03)	(0.01)	(0.02)	(0.03)	(0.01)
BA or More	-0.07**	-0.00	-0.00	-0.11***	-0.03	0.02
	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.01)
Unemployed	-0.12*	-0.04	-0.01	-0.04	-0.00	-0.00
	(0.07)	(0.03)	(0.02)	(0.03)	(0.04)	(0.02)
Not in Labor force	-0.01	-0.01	-0.03*	-0.02	0.01	-0.02**
	(0.04)	(0.02)	(0.01)	(0.02)	(0.03)	(0.01)
Log(Family Income)	0.02	0.06***	0.03***	-0.01	0.04***	-0.01**
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Constant	0.12	-0.19	-0.01	0.50***	-0.42**	0.32***
	(0.22)	(0.13)	(0.08)	(0.10)	(0.18)	(0.07)
R^2	0.11	0.22	0.16	0.14	0.06	0.04
N	1,099	1,877	5,204	3,413	1,172	5,595
Mean	0.49	0.28	0.32	0.40	0.23	0.16

Notes: Robust standard errors are shown in parentheses. A respondent is considered to vote/prefer for Republican presidential candidate if she voted for or preferred (for those who did not vote) a Republican candidate in presidential elections. Years in column titles represent ANES survey years that are used in the regressions. In 2010, we use Log(Household Income) instead of Log(Family Income) because of the data constraint.

Table 2: Predicted ANES Variables and Refusal Rate (2000-2017)

Dependent Variable	Vote McCain		Support Tea Party		Vote Trump	
: Refusal Rate (in %)	(1)	(2)	(3)	(4)	(5)	(6)
Predicted ANES Variable	-1.58***		-2.19***		-1.79***	
	(0.07)		(0.11)		(0.07)	
× Cohort2006	-0.34*	-0.32	-0.65**	-0.64**	-0.47**	-0.44**
	(0.20)	(0.20)	(0.32)	(0.32)	(0.20)	(0.20)
× Cohort2007	-0.66***	-0.59***	-1.31***	-1.28***	-0.64***	-0.58***
	(0.20)	(0.20)	(0.31)	(0.31)	(0.20)	(0.20)
× Cohort2008	-0.88***	-0.81***	-1.29***	-1.26***	-0.82***	-0.77***
	(0.20)	(0.20)	(0.32)	(0.32)	(0.19)	(0.19)
× Cohort2009	-0.65***	-0.55***	-1.06***	-1.00***	-0.63***	-0.56***
	(0.19)	(0.19)	(0.31)	(0.31)	(0.19)	(0.19)
× Cohort2010	-0.34	-0.20	-0.80**	-0.64*	-0.29	-0.20
	(0.21)	(0.21)	(0.34)	(0.34)	(0.21)	(0.21)
× Cohort2011	-0.91***	-0.69***	-1.20***	-0.92***	-0.86***	-0.68***
	(0.22)	(0.22)	(0.35)	(0.35)	(0.22)	(0.21)
× Cohort2012	-0.73***	-0.53**	-1.01***	-0.74**	-0.66***	-0.51**
	(0.22)	(0.22)	(0.36)	(0.36)	(0.22)	(0.22)
× Cohort2013	-1.11***	-0.90***	-2.05***	-1.74***	-1.03***	-0.87***
	(0.24)	(0.23)	(0.39)	(0.38)	(0.23)	(0.23)
× Cohort2014	-1.47***	-1.22***	-1.89***	-1.54***	-1.53***	-1.33***
	(0.24)	(0.24)	(0.39)	(0.39)	(0.24)	(0.24)
× Cohort2015	-2.75***	-2.48***	-4.46***	-4.06***	-2.84***	-2.61***
	(0.26)	(0.26)	(0.43)	(0.43)	(0.26)	(0.26)
× Cohort2016	-4.07***	-3.79***	-6.62***	-6.20***	-4.15***	-3.92***
	(0.29)	(0.29)	(0.46)	(0.46)	(0.28)	(0.28)
× Cohort2017	-6.76***	-6.48***	-10.76***	-10.32***	-6.97***	-6.73***
	(0.33)	(0.33)	(0.54)	(0.54)	(0.33)	(0.33)
R2	0.02	0.03	0.02	0.03	0.02	0.03
N	1,492,083	1,492,083	1,492,083	1,492,083	1,492,083	1,492,083
Cohort Fixed Effects	Y	Y	Y	Y	Y	Y
Controls ANES Characteristics		Y		Y		Y

Notes: Robust standard errors are shown in parentheses. We use households which started their first interviews between 2000 and 2017. Estimates are weighted by the number of eligible surveys (interview or Type-A noninterview) for each household. Dependent variable is household-level refusal rate (in %). We predict the probability of ANES variables (shown in column titles) for each household's reference person (household head) using Table 1. The ANES variables are vote McCain (ANES 2008), support Tea Party (ANES 2010) and vote Trump (ANES 2016). In odd columns, we include the ANES variables and their interaction with cohort-year dummies (2006-2017). In even columns we flexibly control for ANES personal characteristics, instead of the predicted ANES variables that are collinear with the characteristics. We include cohort (year-month) fixed effects in all columns. We drop households whose family income is missing or impute for all rounds of interviews.

Table 3: Refusal Rate, Republican and Republican President (1994-2017)

ANES Year Used in Prediction	ANES 2008			ANES 2016		
Dependent Variable: Refusal Rate	(1)	(2)	(3)	(4)	(5)	(6)
Pr(Vote GOP)	-2.86***	-2.59***		-3.00***	-2.72***	
	(0.04)	(0.06)		(0.04)	(0.06)	
Pr(Vote GOP) $\times$ GOP in Power		-0.63***	-0.72***		-0.67***	-0.76***
		(0.09)	(0.09)		(0.09)	(0.09)
R2	0.02	0.02	0.03	0.02	0.02	0.03
N	2,030,486	2,030,486	2,030,486	2,030,486	2,030,486	2,030,486
Cohort Fixed Effects	Y	Y	Y	Y	Y	Y
Controls ANES Characteristics			Y			Y

Notes: Robust standard errors are shown in parentheses. Estimates are weighted by the number of eligible surveys (interview + Type-A non-interview) for each household. Probability of voting Republican candidate is predicted from ANES. The ANES years used in the prediction of Pr(Vote GOP) are shown in column titles. "GOP in Power" is an indicator variable for a household to be surveyed in time with Republican president (The cutoff point is November elections). If there is an overlap of Democratic and Republican presidents, we use the number of eligible survey months with Republican candidates divided by the number of total eligible surveys for each household.

Table 4: Difference in Refusal Rate and Probability of Voting Trump

Dependent Variable : Change in refusal rate between second 4 and first 4	Dec 2013 -July 2014 (1)	Dec 2014 -July 2015 (2)	Dec 2015 -July 2016 (3)	Pooled (4)
Vote Trump	-0.27	-0.00	-1.55**	-0.13
	(0.67)	(0.61)	(0.68)	(0.46)
Vote Trump X Dec 15-Jul 16 Cohorts				-1.42*
				(0.76)
N	40,716	48,754	49,862	139,332
Cohort Fixed Effects	Y	Y	Y	Y

Notes: Bootstrapped standard errors shown in parentheses. Estimates are weighted by the number of eligible surveys (interview + Type-A noninterview) for each household. Dependent variable is the difference in refusal rate between the second 4 interviews and first 4 interviews. The samples are households who were first interviewed in Dec 2013-July 2014 (column 1), Dec 2014-July 2015 (column 2), Dec 2015-July 2016 (column 3) and Dec 2013-July 2016 (column 4), respectively.

Table 5: Decomposition of Household Characteristics and Refusal Rate

	1994-2009 (1) X	2010-2017 (2) X	(3) ΔX	1994-2009 (4) β	2010-2017 (5) β	(6) $\Delta\beta$	(7) $\Delta X \cdot \beta$	(8) $X \cdot \Delta\beta$
Unmarried civilian male	0.04	0.05	0.01	0.03	0.12	0.09	0.000	0.004
Unmarried civilian female	0.12	0.12	0.00	0.11	0.36	0.25	0.000	0.029
Civilian male primary individual	0.15	0.16	0.02	-0.00	0.09	0.10	-0.000	0.014
Civilian female primary individual	0.17	0.18	0.01	0.15	0.57	0.42	0.001	0.072
Household Size	2.53	2.46	-0.07	-0.11	-0.33	-0.22	0.007	-0.554
Number of Children (Age 0-14)	0.55	0.47	-0.07	-0.09	0.12	0.21	0.007	0.117
Share of Age 15-29	0.21	0.19	-0.02	-0.01	0.44	0.45	0.000	0.095
Share of Age 55+	0.32	0.40	0.08	-1.03	-1.74	-0.71	-0.077	-0.226
All Black	0.10	0.10	0.00	0.72	1.32	0.60	0.003	0.058
All Hispanics	0.07	0.09	0.02	0.19	0.17	-0.02	0.004	-0.002
All Asians	0.03	0.03	0.01	0.51	0.13	-0.39	0.004	-0.010
Mixed Race/Others	0.06	0.09	0.02	0.42	0.61	0.19	0.010	0.012
All Foreign Born	0.05	0.06	0.01	-0.14	-0.41	-0.27	-0.001	-0.014
Mixed Nativity	0.09	0.10	0.02	-0.13	-0.39	-0.26	-0.002	-0.023
Share of High School Dropouts	0.15	0.11	-0.04	0.06	0.37	0.31	-0.003	0.045
Share of High School Graduates	0.33	0.30	-0.03	0.50	1.09	0.59	-0.014	0.193
Share of Some College	0.27	0.28	0.02	0.33	0.73	0.40	0.005	0.106
Share of Employed or Enrolled	0.66	0.63	-0.03	0.60	1.97	1.37	-0.016	0.900
Share of Unemployed	0.03	0.04	0.00	1.52	2.94	1.42	0.005	0.047
Log(Family Income)	10.76	10.75	-0.01	-0.04	-0.07	-0.03	0.000	-0.307
In Metropolitan Areas	0.76	0.79	0.03	0.64	0.87	0.24	0.020	0.182
Total Change Explained							-0.046	0.740
Actual Change in Refusal Rate							2.144	

Notes: We use households that started interview between Jan 1994 and December 2017. Estimates are weighted by the number of eligible surveys (interview or Type-A noninterview) for each household. Survey cohorts from May 1995 to Aug 1995 are dropped, since they were mostly interviewed only once. Households which do not have consistent sex, age, race information of household heads across interview periods are dropped. For the share of employed or enrolled, we use the mean of employment and enrollment rate across the months that they were interviewed. The omitted category of educational attainment is the share of BA or more. The omitted category of employment status is the share of not in labor force, excluding those who are enrolled in schools. Columns 1 and 2 are the average household characteristics of CPS survey cohorts in 1994-2009 and 2010-2017, respectively. Column 3 is the difference between columns 2 and 1. Columns 4 and 5 show the regression coefficients estimated in columns 4 and 5 of Table A1, separately for 1994-2009 and 2010-2017 cohorts, where dependent variable is refusal rate (in percent). Column 6 is the difference between columns 5 and 4. Column 7 is the multiplication of columns 3 and 4 (baseline β: 1994-2009). Column 8 is the multiplication of columns 1 and 6 (baseline X: 1994-2009). The sum of each row in columns 7 and 8 are shown in "Total Change Explained." If we change the baselines to 2010-2017, the total change explained become -0.150 and 0.636, respectively.

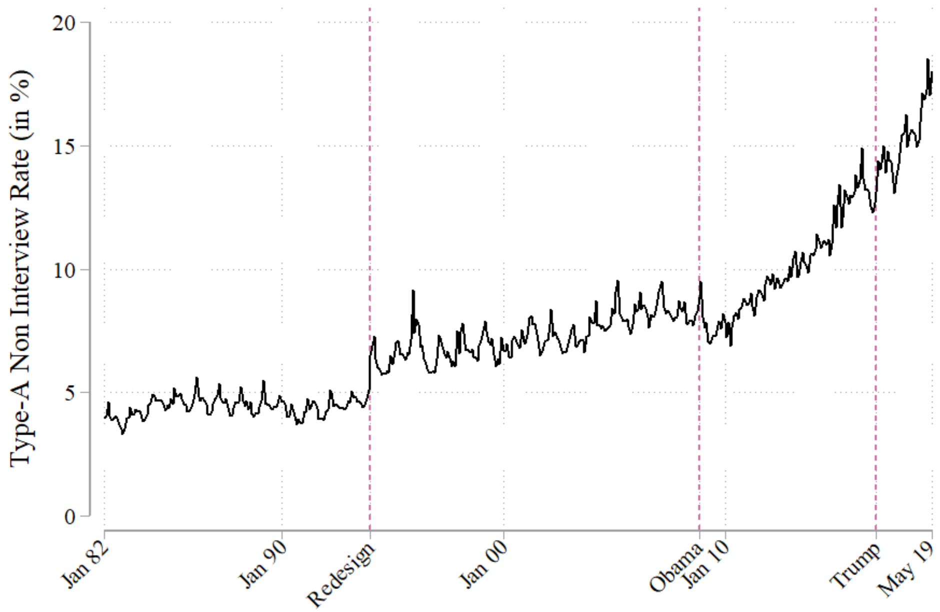
Table 6: Reluctant Responders and Tea Party Rally (2005-2017)

Dependent Variable: Refusal Rate (in %)	(1)	(2)	(3)
Rainy Tea Party Rally	0.09	-0.06	-0.10
	(0.41)	(0.27)	(0.24)
× Cohort2006	-0.29*	-0.17	-0.03
	(0.15)	(0.12)	(0.11)
× Cohort2007	-0.16	-0.06	-0.12
	(0.22)	(0.14)	(0.14)
× Cohort2008	0.10	0.22	0.21
	(0.20)	(0.20)	(0.17)
× Cohort2009	0.36	0.45**	0.40**
	(0.23)	(0.18)	(0.16)
× Cohort2010	-0.10	0.08	0.13
	(0.29)	(0.20)	(0.18)
× Cohort2011	-0.01	0.03	0.06
	(0.32)	(0.26)	(0.25)
× Cohort2012	0.64*	0.52**	0.44**
	(0.37)	(0.24)	(0.21)
× Cohort2013	0.22	0.29	0.29
	(0.34)	(0.28)	(0.24)
× Cohort2014	0.95**	0.83**	0.68**
	(0.38)	(0.32)	(0.29)
× Cohort2015	1.03**	0.90**	0.78**
	(0.48)	(0.40)	(0.38)
× Cohort2016	1.50**	1.35**	1.38***
	(0.59)	(0.53)	(0.49)
× Cohort2017	1.94**	2.06**	1.85**
	(0.94)	(0.85)	(0.78)
R2	0.02	0.02	0.03
N	881,108	854,165	740,479
Cohort Fixed Effects	Y	Y	Y
Controls Probability of Rain	Y	Y	Y
Excludes Non-Metros	Y	Y	Y
Excludes Never Responders		Y	Y
Controls HHD Characteristics			Y

Notes: Standard errors are clustered by metropolitan area and shown in parentheses. Estimates are weighted by the number of eligible surveys (interview or Type-A noninterview) for each household. We use households which were first interviewed between 2005 and 2017. Rainy tea party rally is based on the precipitation amount in the county on the rally day (April 15, 2009). The indicator variable is equal to 1 if there was significant rain in the metro (at least 0.1 inch) and 0 otherwise. We control for the probability of rain in April. Column 2 and 3 exclude never responders. Column 3 additionally controls for household characteristics shown in Table 5.

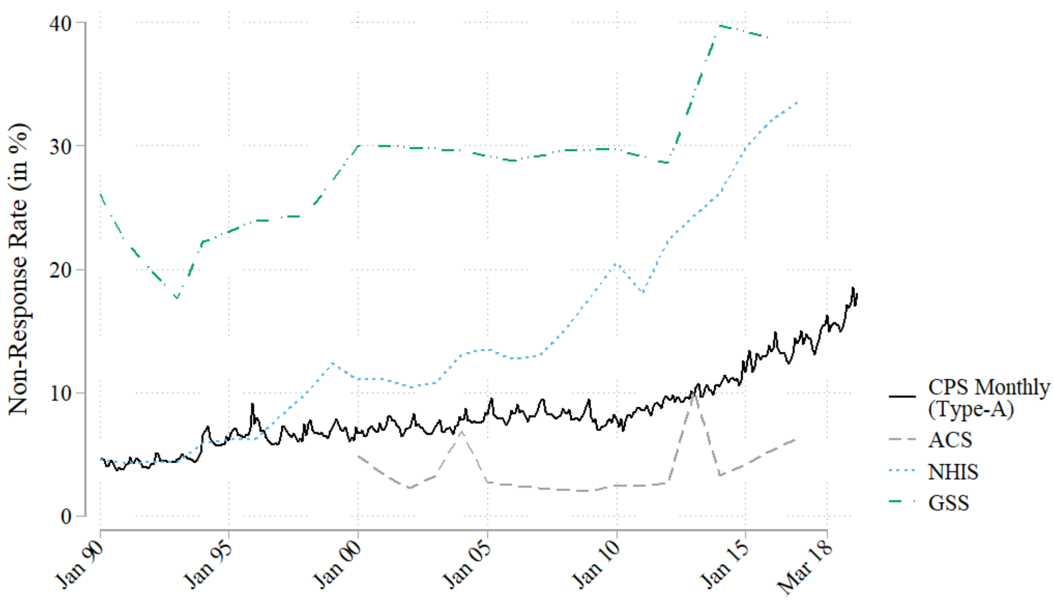
Appendix Figures and Tables

Figure A1: CPS Type-A Non Interview Trends



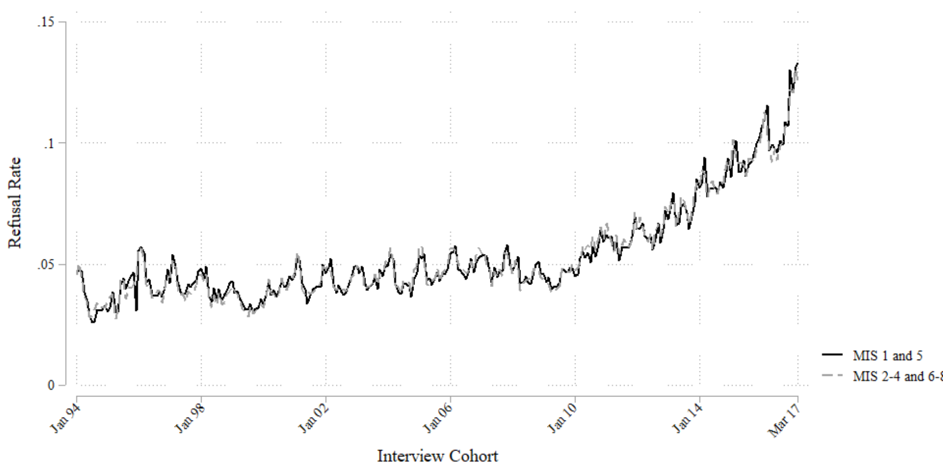
Notes: Type-A Non-interview rate is defined as (Number of Type A Non Interviews)/(Number of Interviews+Number of Type A Non Interviews). Type A nonresponse households represent housing units suitable for inclusion in the survey whose residents were not interviewed for reasons such as refusal to participate and temporary absence.

Figure A2: Non Response Rates in Other Surveys



Notes: CPS Monthly: Current Population Survey, Basic Monthly. ACS: American Community Survey. NHIS: National Health Interview Survey. GSS: General Social Survey. The two peaks in non-response rates in the ACS are due to the government shutdowns.

Figure A3: CPS Refusal Rate Trends by Month in Sample



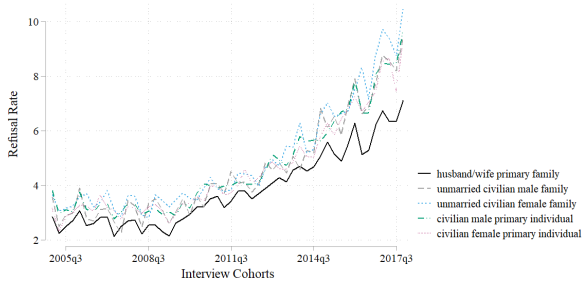
Notes: Interview cohort is defined using the first year-month that the household first enters the CPS. Refusal rate is defined as (Number of Refusals)/(Number of Interviews+Number of Type A Non Interviews). MIS 1 and 5 are default to in-person interview, whereas MIS 2-4 and 6-8 are default to telephone interview. If requested by a respondent, however, the means of contact can be changed.

Figure A4: Average Trump Share by Number of Refusals

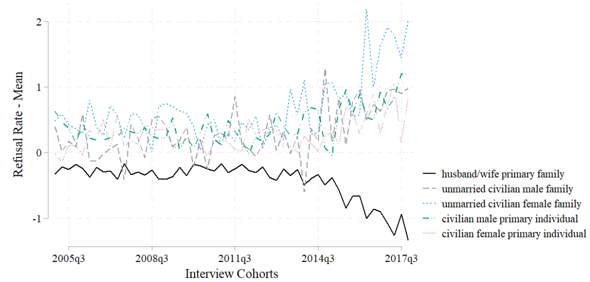
Notes: Figure plots the average metro-level Trump vote share in the 2016 presidential elections by the number of refusals, using households surveyed in 2016. We restrict to households who had full 8 rounds of surveys.

Figure A5: Trends of Refusal Rates by Household Characteristics

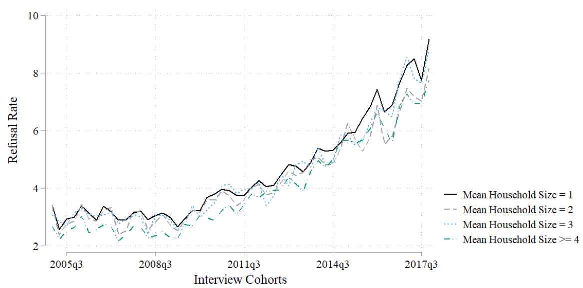
(a) Household Type



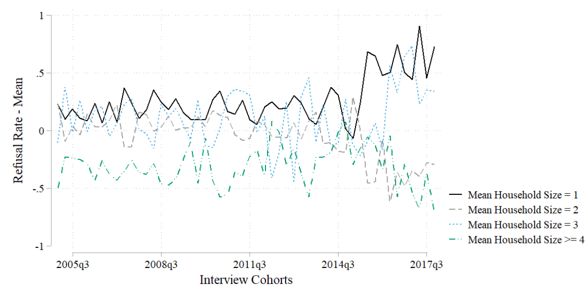
(b) Household Type (Differencing by Mean)



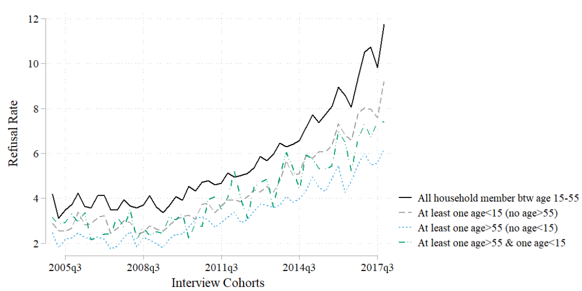
(c) Household Size



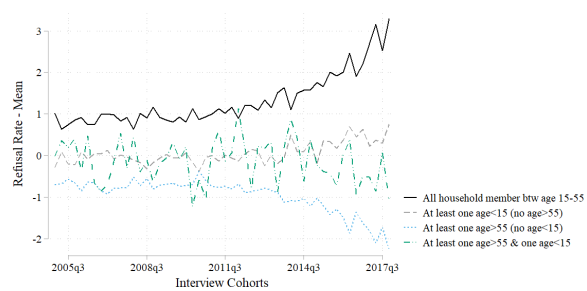
(d) Household Size (Differencing by Mean)



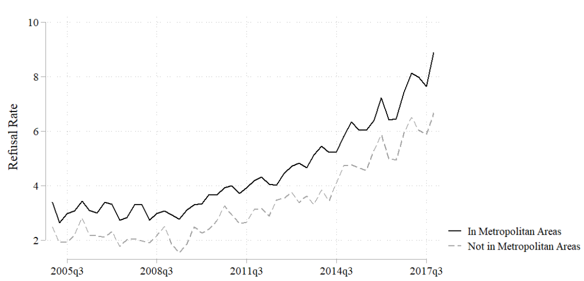
(e) Household Age Structure



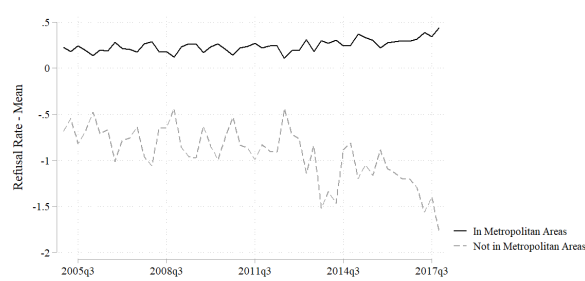
(f) Household Age Structure (Differencing by Mean)



(g) In Metro Area



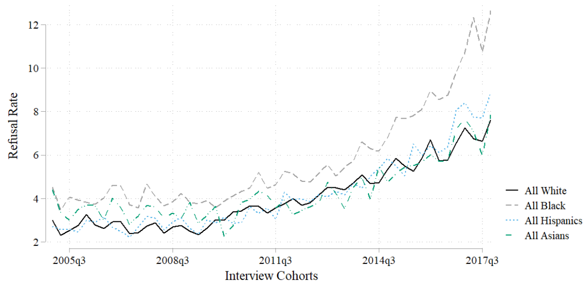
(h) In Metro Area (Differencing by Mean)



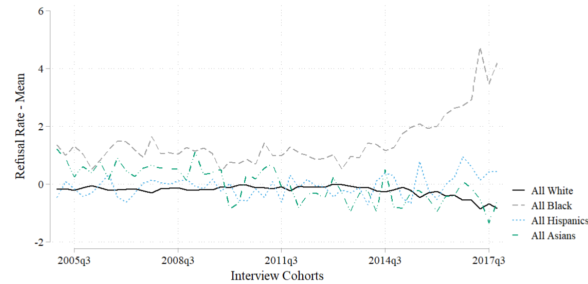
Notes: CPS interview cohorts are defined by the starting month-year that households start their first interviews. For example, Jan, 10 Cohorts are households who started their interviews in Jan, 10. We drop never responders.

Figure A6: Trends of Refusal Rates by Household Characteristics

(a) Race and Ethnicity

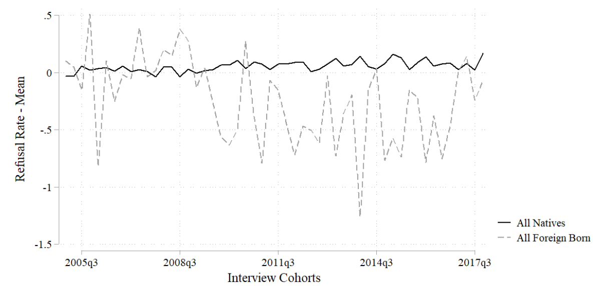


(b) Race and Ethnicity (Differencing by Mean)

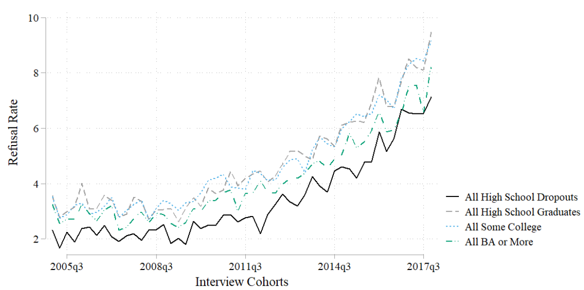


(c) Nativity

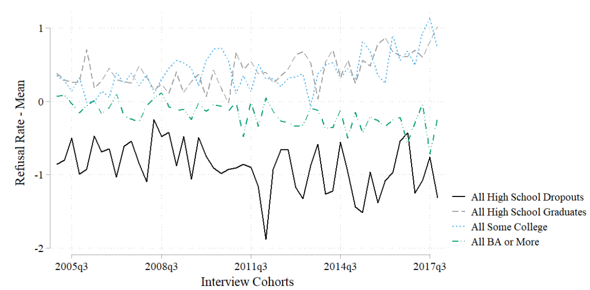
(d) Nativity (Differencing by Mean)



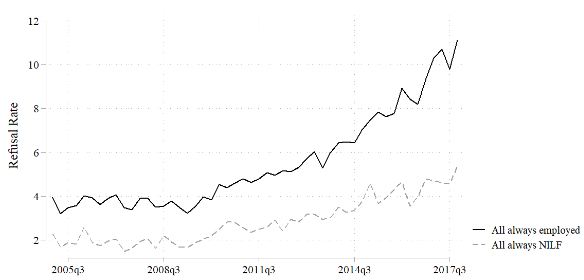
(e) Education



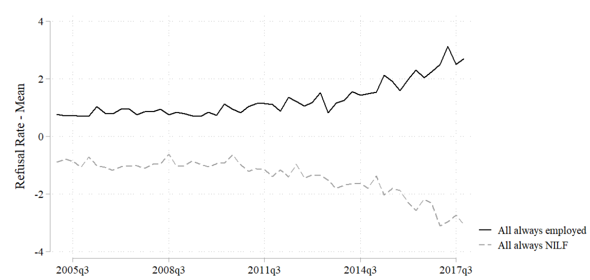
(f) Education



(g) Employment



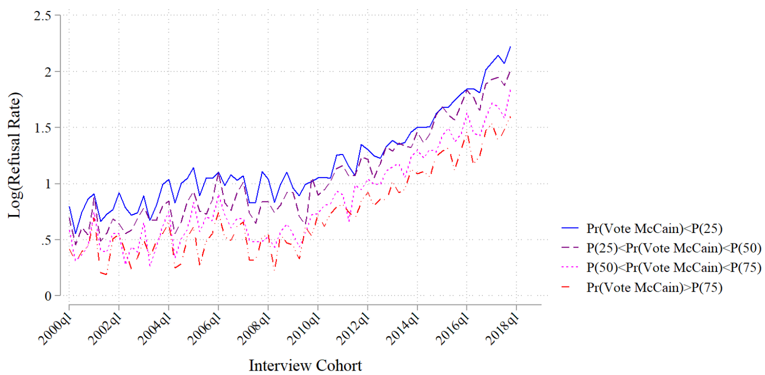
(h) Employment (Differencing by Mean)



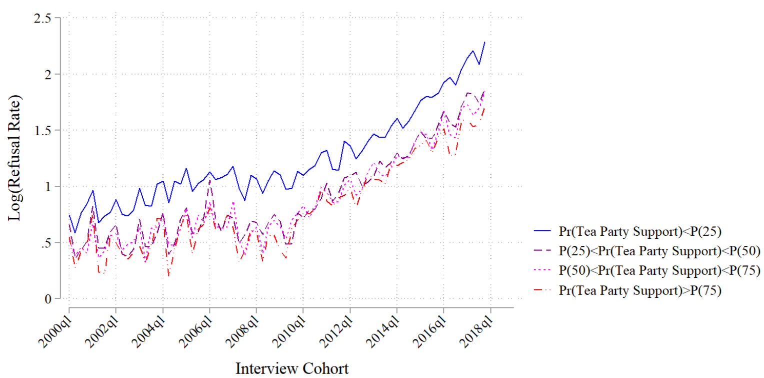
Notes: CPS interview cohorts are defined by the starting month-year that households start their first interviews. For example, Jan, 10 Cohorts are households who started their interviews in Jan, 10. We drop never responders.

Figure A7: Refusal Rate by Quartile of Predicted ANES Variables, Log Scale

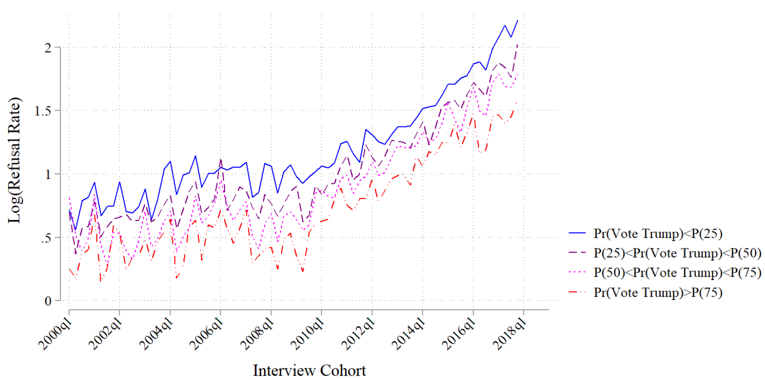
(a) Vote McCain (2008)



(b) Support Tea Party (2010)



(c) Vote Trump (2016)



Notes: We predict the probability of voting McCain (Panel A), supporting Tea Party (Panel B) and voting Trump (Panel C) in the CPS samples using the coefficients estimated in Table 1. For each ANES variable, we stratify CPS households by quartile of the predicted probability and collapse them by interview cohort (year-quarter).

Table A1: Household Characteristics and Refusal Rate

Dependent Variable:	1994-2017				1994-2009	2010-2017
Refusal Rate (in %)	(1)	(2)	(3)		(4)	(5)
	β	β	β	Post 2010	β	β
Constant	3.24***	2.40***	2.39***		2.51***	2.06***
	(0.03)	(0.03)	(0.03)		(0.03)	(0.06)
Post 2010	-0.04	-0.08***	-0.12***			
	(0.03)	(0.03)	(0.03)			
Year	0.07***	0.05***	0.04***		0.04***	0.56***
	(0.00)	(0.00)	(0.00)		(0.00)	(0.01)
Year X Post 2010	0.54***	0.51***	0.54***			
	(0.01)	(0.01)	(0.01)			
Unmarried civilian male		0.08*	0.00	0.05*	0.03	0.12
(family)		(0.04)	(0.04)	(0.02)	(0.04)	(0.08)
Unmarried civilian female		0.19***	0.09***	0.08***	0.10***	0.36***
(family)		(0.03)	(0.03)	(0.02)	(0.03)	(0.06)
Civilian male primary individual		0.03	-0.03	0.05***	-0.01	0.09*
		(0.03)	(0.03)	(0.02)	(0.03)	(0.06)
Civilian female primary individual		0.32***	0.15***	0.11***	0.15***	0.57***
		(0.03)	(0.03)	(0.02)	(0.03)	(0.06)
Household Size		-0.18***	-0.11***	-0.06***	-0.11***	-0.33***
		(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
Number of Children (Age 0-14)		-0.03**	-0.08***	0.05***	-0.09***	0.12***
		(0.01)	(0.01)	(0.01)	(0.01)	(0.03)
Share of Age 15-29		0.13***	-0.03	0.15***	-0.01	0.44***
		(0.03)	(0.03)	(0.02)	(0.03)	(0.07)
Share of Age 55+		-1.29***	-1.02***	-0.20***	-1.03***	-1.74***
		(0.03)	(0.03)	(0.02)	(0.03)	(0.05)
All Black		0.92***	0.67***	0.20***	0.72***	1.32***
		(0.03)	(0.03)	(0.02)	(0.03)	(0.06)
All Hispanics		0.18***	0.16***	0.01	0.19***	0.17**
		(0.04)	(0.04)	(0.02)	(0.04)	(0.07)
All Asians		0.35***	0.46***	-0.07***	0.51***	0.13
		(0.05)	(0.05)	(0.03)	(0.06)	(0.09)
Mixed Race/Others		0.51***	0.40***	0.07***	0.42***	0.61***
		(0.03)	(0.03)	(0.02)	(0.04)	(0.06)
All Foreign Born		-0.24***	-0.15***	-0.07***	-0.14***	-0.41***
		(0.04)	(0.04)	(0.02)	(0.04)	(0.07)
Mixed Nativity		-0.22***	-0.14***	-0.06***	-0.13***	-0.39***
		(0.03)	(0.03)	(0.02)	(0.03)	(0.06)
Share of High School Dropouts		0.22***	0.08**	0.07***	0.06*	0.37***
		(0.03)	(0.03)	(0.02)	(0.03)	(0.07)
Share of High School Graduates		0.68***	0.52***	0.14***	0.49***	1.08***
		(0.02)	(0.03)	(0.02)	(0.03)	(0.05)
Share of Some College		0.46***	0.36***	0.09***	0.33***	0.73***
		(0.03)	(0.03)	(0.02)	(0.03)	(0.05)
Share of Employed or Enrolled		1.09***	0.62***	0.36***	0.60***	1.97***
		(0.03)	(0.03)	(0.02)	(0.03)	(0.05)
Share of Unemployed		2.02***	1.53***	0.46***	1.51***	2.93***
		(0.10)	(0.11)	(0.07)	(0.11)	(0.19)
Log(Family Income)		-0.06***	-0.03***	-0.01*	-0.04***	-0.07***
		(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
In Metropolitan Areas		0.72***	0.63***	0.07***	0.64***	0.87***
		(0.02)	(0.02)	(0.01)	(0.02)	(0.04)
R2	0.02	0.03	0.03		0.01	0.03
N	2,287,594	2,021,721	2,021,721		1,366,290	655,431
Month Fixed Effects	Y	Y	Y		Y	Y
Year X Post 2010 X Household Charcteristics			Y			

Notes: Robust standard errors are shown in parentheses. I use households that started interview between Jan 1994 and March 2017. I drop survey cohorts from May 1995 to Aug 1995, since they were mostly interviewed only once. Dependent variable is refusal rate (in percent), which is the defined as (number of refusals)/(total number of interviews+Type A non-interviews) for each household. I drop households which do not have consistent sex, age, race information of household heads across interview periods. For the share of employed or enrolled, I use the mean of employment and enrollment rate across the months that they were interviewed. Variable year is normalized on 2010. The omitted category of educational attainment is the share of BA or more. The omitted category of employment status is the share of not in labor force, excluding those who are enrolled in schools.

Table A2: Tea Party Rally and GOP Vote Share in House Elections: Metropolitan-level

Dependent Variable : Republican Votes (% of Pop)	2010 (1)	2012 (2)	2014 (3)	2016 (4)	2018 (5)
Rainy Tea Party Rally	-1.52***	-1.25**	-0.56	0.39	-0.66
	(0.44)	(0.59)	(0.52)	(0.68)	(0.58)
R2	0.68	0.72	0.59	0.66	0.60
N	868	868	867	867	868
Region Fixed Effects	Y	Y	Y	Y	Y
Controls Probability of Rain	Y	Y	Y	Y	Y
Election Controls in 2008	Y	Y	Y	Y	Y

Notes: The election data come from Dave Leip's Atlas of U.S. Presidential Elections (https://uselectionatlas.org/). Robust standard errors shown in parentheses. Estimates are weighted by metro population. We use the 2010 version of metro definition, excluding non-metro areas. We include region fixed effects. For few cases where metros overlap two regions, we use region with more population. The 2008 election controls include Republican votes in 2008 per capita, Democratic votes in 2008 per capita and Republican vote share in 2008.