How Campaign Ads Stimulate Political Interest

Nathan Canen and Gregory J. Martin
University of Houston and Stanford Graduate School of Business

Preliminary Version

ABSTRACT. We empirically investigate two key dynamic features of advertising competition in elections, using a new dataset of very high-frequency, household-level television viewing matched to campaign advertising exposures. First, we demonstrate that exposure to campaign advertising stimulates increased consumption of news programming: the average effect is a 4 minute increase in news viewing time over the next 24 hours. Hence, in contrast to many existing models, advertising indirectly influences voters' information environment by changing voters' decisions to become informed about the campaign. Second, viewers' attention to political advertising declines over the campaign: the rate of viewer tune-out of political ads is about 20 percent higher the day before election day than two months out. These dynamic forces counteract the short shelf life of advertising's direct persuasive effects, and rationalize the observation that the large majority of candidate-sponsored television advertising occurs more than two weeks prior to election day.

KEY WORDS. Campaign advertising, electoral competition, dynamic effects.

JEL CLASSIFICATION: D72; D83; M37.

1. Introduction

The minimal effects hypothesis (Lazarsfeld et al., 1944; Berelson et al., 1954) suggests that campaigning has very small effects on voters' decisions. In this view, voters have persistent beliefs and do not change their voting or turnout decisions, even when exposed to new information from political campaigns. The early arguments for this position were based on observational evidence subject to the critique that one candidate's campaign

Date: October 2019.
We would like to thank Hugo Jales and Juan Felipe Riaño for valuable discussions.
Corresponding author: Martin <gjmartin@stanford.edu>.

effort is highly correlated with his/her opponent's, and thus the equilibrium marginal effect might be much lower than the true all-else-equal counterfactual. But modern field-experimental evidence (Kalla and Broockman, 2018; Gerber et al., 2011), circumventing this problem through random assignment, has found effects on vote choice that are either very small or very short-lived. The Gerber et al. (2011) study, which randomized a portion of television ads by one candidate in a gubernatorial race, found effects on vote intention that decayed to zero within a week. Other studies find similarly small effects on turnout and voting behavior (Ashworth and Clinton, 2007; Krasno and Green, 2008 and most recently, Spenkuch and Toniatti, 2018).

Politicians themselves do not appear to believe these findings. Campaigns devote enormous quantities of money and effort to advertising, debates, direct mail, door-to-door voter contact, and other forms of information transmission. Campaigns carefully target their advertising to programs whose audiences skew towards moderate swing voters, and voters likely to turn out (Lovett and Peress, 2015). And while effort and expenditures certainly ramp up towards the end of the campaign, candidates do not conserve their advertising budgets for an immediate pre-election-day blowout, as would be suggested by a straightforward reading of the experimental evidence: in our data, more than 80% of general election TV advertising impressions occur two or more weeks prior to election day.

Are these campaign efforts wasteful, from the candidate's perspective? We propose and empirically investigate a new mechanism that can rationalize early campaign advertising expenditures, even if the direct changes to turnout or vote intentions they induce are minimal or short-lived. We study whether exposure to campaigns spurs political interest, inducing voters to inform themselves by consuming more political news. Hence, while voters' preferences might be persistent, advertising can still change voters' information sets at election time, by changing if and when voters choose to become informed about the election. This opportunity declines as the election approaches, because there are comparatively fewer uninformed voters whose beliefs can be moved by new information.

Our explanation focuses on an under-studied aspect of campaigns: their dynamic nature. Both politicians and voters send and receive information on multiple occasions over the course of a campaign. A single TV ad or other voter contact in isolation may not have much direct impact on voter beliefs or actions, as the experimental evidence generally confirms. But voters' total accumulation of political information over an entire campaign – from campaign activities as well as from media coverage – dwarfs that of any single ad. Indeed, there is a large body of evidence that quasi-random variation in the macro-scale information environment has substantial impacts on voter behavior.[1] We connect these two seemingly inconsistent bodies of evidence by investigating whether an advertising intervention at time t can affect voting at time T by changing the voter's information acquisition decisions at $t+1, t+2, ..., T$.

To test this mechanism, we implement a differences-in-differences design exploiting variation in exposure to ads using very high-frequency set-top-box data. We are able to observe television viewership by approximately 200,000 households at 1 second intervals. We compare viewers who were watching a program at the instant a political ad began with a control group of viewers that tuned out from this same program slightly earlier, or tuned into it slightly later (and hence, did not see the ad). Our results show an average increase of approximately 4 minutes in news viewing over the next 24 hours following exposure to an ad, an effect that declines as election day approaches. The fact that political ads frequently run on news programs[2] generates a positive feedback, as ad viewers are more likely to encounter additional political ads in the future. This feedback loop keeps viewers exposed to political information in both the ads themselves and in surrounding news content. As a consequence, campaigns appear to be able to stimulate increased attention to politics through their campaigns, affecting both political interest and engagement in campaigns. This contrasts with results that interpret the lack of effects of advertisements on turnout to mean that campaigns do not spur engagement (Huber and Arceneaux, 2007).

Our effects are heterogeneous and related to both the type of race for which the ad was aired and the identity of the sponsor. The effects are strongest for ads in gubernatorial races, weakest for ads by outside groups, and are stronger for ads run by challenger candidates than those run by incumbents. The effect appears to be driven by increased viewership of local news (rather than national cable channels). These results together provide evidence for the mechanism of ads inducing voters to seek out information. Voters are more likely to recognize and already possess information about incumbents than challengers. And except for the highest-profile governors, local news is a more likely source to find coverage of state-level politicians.

We also test how viewer attention to ads themselves varies over the course of the campaign, and with viewer attributes. We measure, among the set of viewers tuned in at the instant a political ad begins, the fraction who change the channel before the end of the ad. As a baseline for comparison, we compute the viewer-specific tune-out

[1] Such informational variation could take the form of exposure to campaign advertising, as in Gordon and Hartmann (2013) or Spenkuch and Toniatti (2018); exposure to partisan media as in DellaVigna and Kaplan (2007) and Martin and Yurukoglu (2017); release of newspaper endorsements (Chiang and Knight, 2011); or release of evidence of corruption (Ferraz and Finan, 2008).

[2] About half of political ads in our sample occurred on a news program.

rate at another, randomly selected time during the same day. We find that early in the campaign, political ads induce about 25% *less* tune-out than average TV content. As the election approaches, however, the tune-out rate increases, and by election day the political ad tune-out rate is similar to that of other content. This result is consistent with the mechanism of viewer interest in the campaign being stimulated by political advertising, and also strengthens candidates' incentives to advertise early, before voters are saturated with advertising.

Our results connect and contribute to several strands of research. First, our results add subtlety to standard Bayesian voting models used in formal work on campaigning (e.g. Achen, 1992; Gerber and Green, 1998). If ads can be treated as (partially) informative signals and voters are Bayesian, willingness to search for information should decrease with every additional ad a viewer sees.[3] This is the opposite of the stimulating effect we observe in the data. Instead of viewing ads as containing information themselves, our results suggest a model of ads as increasing viewer interest in campaign information from other sources. Or, equivalently, reducing viewers' search costs for campaign-related information. This mechanism helps to explain why, for example, the randomized information campaign of Kendall et al. (2015), which provided information to voters exclusively about the incumbent mayor's record and positions, reduced the uncertainty of treated voters' beliefs about his *opponent*.

Second, we provide further results on the effects of television advertising in campaigns. Our very granular and high-dimensional data provides advantages to those mostly used in the literature[4], as we can explore very fine variation in the timing of viewing of political ads across citizens, while controlling for media markets, program characteristics and viewership preferences. Furthermore, this variation is very rich generating many

[3] As priors get tighter, the benefit of new information decreases. If costs of information acquisition are constant, then search should decline. This intuition also holds in a strategic setting where politicians can target ads at such citizens (Gratton et al., 2017). In such rational learning models, the strategic decision that politicians face when choosing ads has a straightforward solution: send good signals about themselves early on (that may persist through voter's beliefs during the campaign), and send unflattering signals about themselves as late as possible (so that there are less voters paying attention or acquiring such information), see Gratton et al. (2017). However, the implications of our results can be quite different, and more subtle. Politicians face different distributions of voters viewing political ads at any point in time, and these distributions change as a result of the ads they send out. As such, stimulating viewers to acquire information can be beneficial, as it convinces viewers who are not actively following the campaign, while also changing the distribution of viewers who will see later ads. Such complexities could be a reason for the delayed study of dynamic campaigning and even though we do not offer a full formal theory, our empirical results suggest the clear mechanisms by which these actions occur.

[4] As Kalla and Broockman (2018) describe, "First, the existing literature (and, by extension, our meta-analysis) provides only scarce evidence on the effects of television and digital advertising, which represent a great deal of campaign spending...more evidence about these mediums would clearly be welcome." (p.163)

quasi-natural experiments. This variation allows us to evaluate heterogeneity along several dimensions, including the effects of advertisements by different types of candidates, during different parts of the campaign, and on different subsets of viewers.

When compared to the existing literature on television advertisement on turnout and vote shares (e.g. Ashworth and Clinton, 2007; Krasno and Green, 2008; Spenkuch and Toniatti, 2018), we differ on multiple fronts. First, our focus is on how campaigns change viewers' consumption of political news. We investigate viewers' media consumption, a preliminary step before voting outcomes. While there is some literature on interactions between paid and "earned" media (Ridout and Smith, 2008; Lovett and Staelin, 2016), and on the influence of campaigning on the media environment more broadly (Vavreck, 2009), we provide direct evidence of this linkage, and a mechanism that works through induced changes in viewer demand for campaign information. Instead of focusing on the static effects of an ad over on viewers' contemporaneous preferences - which may well be minimal - we examine their dynamic influence on the complete informational picture that voters absorb. Second, we provide evidence that can inform models of information acquisition and accumulation. Campaign advertising does not appear to be well approximated as a partially informative signal about candidate attributes which voters process in Bayesian fashion. Instead, it appears to modulate the parameters of voters' dynamic information search problem (examples include Steiner et al., 2017; Matějka and Tabellini, 2017) Finally, our source of variation differs: we do not study differences across media markets (e.g. Ashworth and Clinton, 2007; Spenkuch and Toniatti, 2018), but instead explore variation within them through the *timing* of viewership for consumers who were watching the same channel, but happen to have watched/not watched an ad.

Every political ad within our sample is a quasi-experiment generating distinct treatment and control groups, generating over 100,000 such quasi-experiments. Traditional methods for event studies or difference-in-differences designs only work with a few of those. They also do not usually allow for heterogeneous treatment effects,[5] or heterogeneous timing of treatment exposure - both important features of our problem. We build appropriate ways to analyze the validity of our empirical strategy, and to aggregate such results to meaningful parameters. We are then able to present results on aggregate pretrends and balance tests across control and treatment groups. For our main results, we show the distribution of estimated effects across ads and across days.

The paper is organized as follows. In Section 2, we explore our data and discuss how we create our variables of interest. In Section 3, we outline our empirical strategy, including how we define treatment and control for each ad, and show evidence for the

[5]Which are very likely in our setting, given viewers are in different districts with different candidates, and receive ads at different points of the election and of the day.

validity of our design. Section 4 presents our results and analysis for our main specifications, together with heterogeneity analysis. We discuss the effects of advertisements on tune-out rates in Section 5 before concluding in Section 6. Additional details on the data, robustness checks and discussion are presented in Appendix.

2. Data

We use two primary data sets in our analysis: household-level television viewership from set-top boxes, and ad occurrences. The former comes from the vendor FourthWall Media; the latter comes from Nielsen. We briefly describe each dataset in turn.

2.1. Viewing Data

Our viewership data covers the subscribers to 10 cable providers (Multiple System Operators or MSO's) around the country. The dataset covers viewing by more than 200,000 anonymized households from June 2012 to January 2013,[6] spanning 60 Designated Market Areas (DMAs). Importantly, the data are not an opt-in sample, but cover the population of subscribers to each of the 10 MSO's.

The data are event-level, tracking tuning decisions (e.g. changing the channel or turning off the set-top box) to a resolution of 1 second. We observe the time of a tuning event and the channel switched to. A viewing interval is the time span between two tuning events. We use this data to determine the set of viewers exposed to a given political ad, using a method detailed below.

The outcome of interest is viewers' consumption of news programming, which can be measured in the same set-top-box data. We use a database of program classifications provided by FourthWall to determine which programs qualify as news and which do not. Specifically, we include any programs which FourthWall tags as in the "News" or "Politics" genres, except for any program which also includes any of "Entertainment", "Sports non-event", "Sports event", or "Religious" in the genre field. The latter is to exclude programs like ESPN SportsCenter or Access Hollywood; our interest is in measuring viewers' consumption of politically-relevant news. We use this classification to determine which viewing intervals corresponded to news programs, and measure consumption as the total minutes spent in viewing intervals corresponding to news programs.[7]

[6]The number of households covered varies by day, but the median day in our sample period has 234,834 active households.

[7]The Fourthwall program database is incomplete and has some missingness, particularly for local news programs. For example, in the Bend media market (where one of the MSOs covered by our data is located), most days in the sample period have no local news programs appearing in the program database. The cable news program schedule is much more consistent and complete but also has some idiosyncratic missingness.

The set-top box data also provides a rich set of demographic covariates. Further detail on the set of covariates in the data is in Table 6 in Appendix A. Covariates, like viewership, are measured at the household level. As the majority of households contain more than one individual, covariates that are measured at individual level (such as age, or race) are the values for the head of household.

2.2. Ads Data

Data on political advertising comes from Nielsen's Ad Intel database. We capture the universe of political advertisements aired in the 2012 cycle on spot TV.[8] Nielsen data includes the DMA in which the ad ran, the date and exact air time, the program and network on which the ad ran, the sponsor, the impressions and Gross Ratings Points (GRPs), and a descriptive title. Data are at the level of the individual ad occurrence. Using the sponsor information, we determined the type of sponsor (candidate or outside group) and, for candidates, the office sought. We limit to ads run in the 60 DMAs covered by the viewership data. The resulting dataset has 286,863 individual ad occurrences, the vast majority of which (96%) are standard 30-second length; the remainder are split approximately equally between 15 and 60 second lengths.

Because many DMAs in the set of 60 covered by the set-top-box data have only a small number of covered households, many of these ads have no households in our sample in either treatment or control groups. Dropping these ads with no viewing data reduces the number of ads in the final dataset to 116,475. Table 1 shows the distribution of ads with nonzero set-top-box households in both treatment and control groups, by sponsor.

2.3. Linking Viewership to Ad Occurrences

We consider a set top box in the viewing data to be exposed to a given political ad if its viewing history meets the following criteria. First, the box was tuned to the channel on which the ad ran in the second at which the ad *began*.[9] Second, that the box had tuned to the channel on which the ad ran no more than one hour prior to the ad start time. This second restriction is in place to remove viewers who leave their television on

We note here that this missingness affects the efficiency of our estimates but not consistency, as program schedules are shared across all viewers in a DMA and thus the measurement error it induces cannot be correlated with treatment group status. Appendix C presents results where we drop ads in markets with the most missing news program data, which are qualitatively similar to the main estimates.

[8]Ads in the sample ran on local network affiliates, each of which reach a single media market. The dataset does not cover national ad buys (which are rare in presidential campaigns and nonexistent in down-ballot races) or local cable buys.

[9]I.e., ours is an intent-to-treat estimate. It is possible that viewers in the treated group tune out before the end of the ad. Section 5 provides evidence on the rate of tune-out among viewers in the treated group.

TABLE 1. Distribution of Ads by Sponsor used in our Empirical Strategy

Sponsor Type	Ad Count
US House	25632
US Senate	16207
President	35304
Statewide	13389
Outside Group	25943

Notes: We present the ad count by sponsor in the Nielsen data, conditional on having at least one device in both treatment and control groups. The first four rows are candidate-sponsored ads, grouped by office sought. The last row are ads sponsored by third-party independent expenditure groups.

for long stretches but may not be actively watching. The set of set top boxes which meet these two criteria are the treatment group for a given ad.

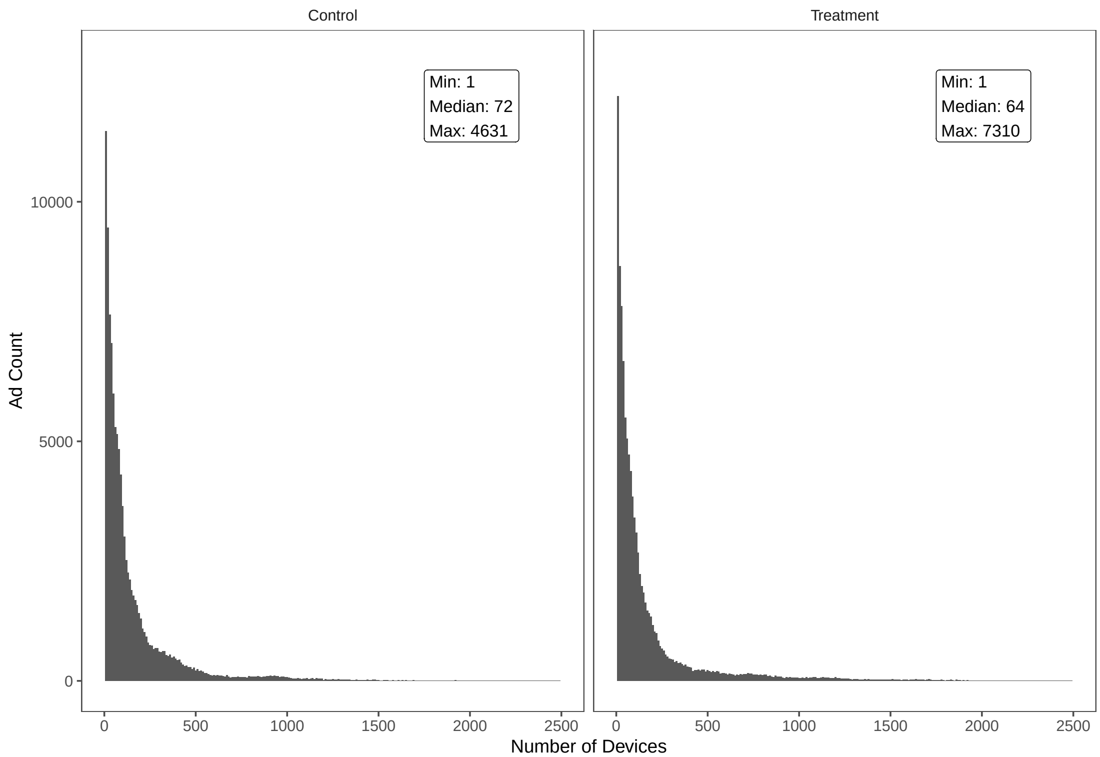
We define a control group of set top boxes at the ad level using the following criteria. First, the box was active (not off) at the second in which the ad began, but was not tuned to the channel on which the ad ran at that time. Second, the box must have been tuned to the same program and channel on which the ad ran at some point during the same half-hour time block in which the ad ran.[10] Third, it began viewing the channel on which the ad aired no more than one hour prior to the ad start time.

These conditions are necessary to generate a valid control group for the treated devices. The second condition is the most fundamental, and deals with the issue that campaigns target advertising purchases to specific programs on the basis of desired political characteristics of the program's audience (Lovett and Peress, 2015). Because targeting occurs at the program–time block–media market level, constructing a control group that also viewed the same program in the same time block and the same media market as the treated group ensures that treated and control groups do not differ on unobservables known to advertising buyers.

The first and third conditions are needed because we impose the same conditions on the treatment group. We require that a "treated" device is on at the time of ad airing, for obvious reasons, and had begun watching in the past hour, in order to increase the likelihood that the viewer was actively watching at the time of ad airing. If we did not also impose the same condition on control devices, we would put the parallel trends assumption at risk, a condition for identification of our model which we expand on in the next section.

[10] Our point estimates are very similar if we consider a five minute block for the control group instead of the half-hour one.

FIGURE 1. Distribution of treatment and control group size (number of devices)



Notes: We present the distribution of the group size in number of devices for treatment and control groups. Figure excludes 123 groups with size exceeding 2500 devices.

The typical size of treatment and control groups in the sample is around 70 devices. The median ad has 64 devices in treatment group and 72 in control group. Figure 1 shows the full distribution of group sizes in the sample across all 116,475 ads.

3. Empirical Strategy

Our aim is to study the effects of exposure to political ads on news consumption. Let us define the news viewing time[11] in time window $w \in \{0, 1\}$ relative to the air time of ad a for viewer i as $y_{a,i,w}$. We use a 48 hour window around ad time, so $w = 0$ indicates the 24 hours prior to air time and $w = 1$ indicates the 24 hours following air time. Viewer i can be in either the treatment or control groups, with treatment status indicated by the random variable $T_{i,a}$. $T_{i,a}$ equals 1 in case of treatment (device was viewing the program of ad airing at time of ad airing) and 0 in case of control (device viewed the

[11]The total time, in seconds or minutes, spent watching news programs.

same program but was viewing something else at the time of ad airing). The standard differences-in-differences regression model under homogeneous treatment effects is:

$$y_{a,i,w} = \alpha + \beta T_{i,a} + \gamma w + \delta T_{i,a} w + \varepsilon_{i,a,w}, \tag{3.1}$$

with δ being our parameter of interest. δ captures the average effect of being exposed unexpectedly to a political ad on news consumption in the following 24 hours (relative to the average change in news consumption among those not exposed).

For the moment, let us consider the case of one single experiment (ad). To run the within-ad analysis of equation (3.1), we must first identify an exogenous variation that introduces political ads to some viewers, but not others. As discussed above, we use the variation that some active viewers tuned into the political ad, while others tuned out just before the ad came on. Treated viewers had active viewership on their set-top boxes when the ad came on, while control groups had active TV viewing on another channel when the ad came on, but had viewed the same channel as treatment within the same half hour. As such, our variation is within program-time blocks and within media markets.

As is standard in such difference-in-differences studies, identification requires the parallel trends assumption: that both treatment and control would have had the same *change* in news viewership absent exposure to the ad. Parallel trends is fundamentally untestable, as we cannot observe counterfactual viewership for the treated group. However, we have available a rich set of viewer demographic characteristics and can check that treated and control groups do not exhibit substantive differences on observable dimensions. And, we can compare viewership trends in the pre-exposure period, to check that any divergence between groups appears only after the ad time. We show the results of these tests in the next subsection.

Before we do so, however, recall that we have multiple ads and treatments, allowing for a generalization of the specification in (3.1). We can allow for ad-specific effects, $\{\delta_a\}$ as in:

$$y_{a,i,w} = \alpha_a + \beta_a T_{i,a} + \gamma_a w + \delta_a T_{i,a} w + \varepsilon_{i,a,w}. \tag{3.2}$$

We now have a vector of parameters of interest, $\{\delta_a\}$, which can vary across subsamples - including when the ad is aired during a day (for example, effects of ads aired early in the morning could differ from those of ads aired in the evening), the timing of the ad in the campaign (whether in the beginning or the last few days), or the type of election (e.g. presidential vs. congressional). Our results in Section 4 illustrate that it is important to allow for such heterogeneity.

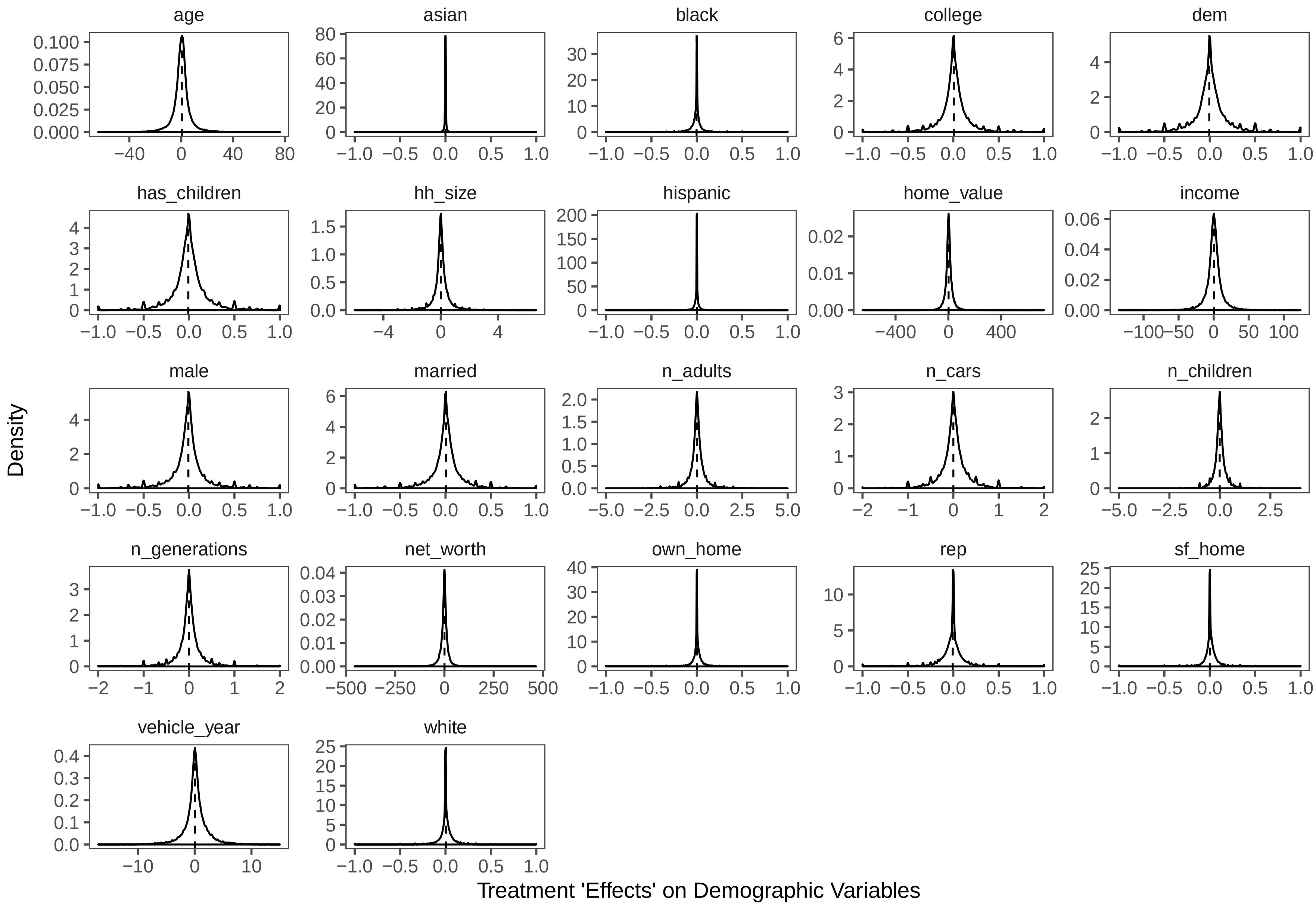
We pursue identification and estimation of equation (3.2). Identification assumptions here are relaxed relative to the simpler homogeneous model (3.1), in that we require parallel trends to hold only within-ad and not across treatment (T) and control (C) groups from different ads. This relaxation will prove to be important due to the cyclical nature of news viewing over the course of the day, and variation in audience composition across programs and time blocks. In the next subsection, we report *distributions* of test statistics from tests of balance and pre-trends across all ads in the sample.

3.1. Exogeneity of Treatment

3.1.1. **Balance Tests.** We now show that the covariates of treated and control groups are balanced, an important check for the validity of our experimental design. Due to the high-dimensionality of our data with multiple experiments, we must find ways to appropriately present such balance tests. We provide two alternatives: the first, shown in Figure 2, presents the distribution of the estimated "treatment" effects for each covariate across ads. Each observation underlying the reported kernel density estimates is the regression coefficient of some predetermined covariate (such as age or income) on treatment or control status, for a single ad. We report a distribution instead of a point estimate because we have many ads. Under the null that groups are balanced, this distribution should approximate a normal centered at zero.

An alternative is to plot the (distribution of) p-values across ads for the treatment effects on each covariate, rather than the coefficients themselves. Under the null hypothesis of no differences across treatment and control groups, we expect a uniform distribution of p-values. We report these distributions in Figure 6 in Appendix C.

FIGURE 2. Balance of Covariates



Notes: We present the distribution of estimated treatment effects for different covariates. Estimates are constructed by regressing each indicated covariate on an indicator for treatment status $T_{i,a}$. There is one such regression for every covariate and every ad in the sample. Our data is well balanced as the covariates are not significantly different across those groups. Additional graphs and details are presented in Appendix C.

Our covariates appear balanced across treatment and control groups. This can be clearly seen in both sets of graphs, where treatment effects have an estimated distribution precisely around 0 for all measured covariates. Given the large number of ads, it is natural that some would be drawn with statistically significant differences across characteristics, but the central tendency is very close to zero for all measured covariates. The distribution of p-values for treatment effects across ads (in Figure 6) closely follows the theoretical uniform distribution as well. The possible exceptions are for `age` and `income` in Figure 6, which have higher than expected mass below 0.05. Figure 2 suggests that these low p-values are driven by the very large sample size and corresponding precision of our estimates, as the differences on all covariates including age and income are substantively very small. We show in Appendix C that our results are qualitatively similar (and in fact increase in magnitude) when we drop the subset of ads with statistically significant $(p < 0.05)$ differences in age or income.

3.1.2. **Parallel Trends.** In the standard difference-in-differences (DiD) model, checking for the absence of pre-trends simply involves comparing trends in the outcome variable in the pre-treatment period across two groups. While the standard DiD model has two periods (pre and post-treatment) and two groups (treatment and control), we are faced with multiple periods - one for each ad, across multiple days, multiple dosages (i.e. the same subjects can be treated multiple times) and heterogeneous treatment effects (we do not expect the effect of an ad aired at 5am to be the same as that in prime time, or that the effects are the same across geographical areas).

A naïve way to visualize the validity of pre-trends would be to aggregate all ads and plot the average news consumption for treatment and control groups across windows before/after ad exposure. This would be the standard graphical representation in event-study designs: normalize all treatments to a period "0", and compare outcomes before and after that period across groups. However, as discussed in Abraham and Sun (2018), a spurious pre-trend may arise simply due to the aggregation of heterogeneous trends across treatments. The intuition is simple, within the context of heterogeneous treatment effects. For example, most people watch news in the morning and/or in the evening, so the time-pattern of viewing in a 24h window around 5am will look very different than that in a 24h window around 8pm. It follows that aggregating these effects might assign spurious pre-trends: the pre-exposure control group at 5am is not an appropriate comparison group for the pre-exposure treatment group at 8pm. Under different treatment effects, the weight attributed to those at 5am conflated with their differential can show up as an inappropriate pre-treatment trend break. In fact, aggregating heterogeneous effects across time and treatments is known to generate multiple problems, as the aggregation weights can be inappropriate (Goodman-Bacon, 2018; De Chaisemartin and D'HaultfŒuille, 2017, among others).

The solution in our context is to appropriately control for the "news viewing cycle" and the sources of heterogeneity in treatment effects. Most notably, these could be variations within a day and media markets, as well as the day/hour/timezone an ad is aired, for both treatment and control groups. Ideally, we would visualize such trends at the ad-level (the level upon which this assumption is made), or at least control flexibly for such confounding factors. Given we cannot present graphs or results for all experiments, we follow our approach from the last section and run a regression based analysis at the ad-level. We then aggregate the results for the presence of pre-trends as we did for the balance checks. More precisely, we run an analogous specification from our main estimating equation (3.2) with outcomes as the averages per group at hourly bins for periods *before* treatment occurs. We present the results in Figure 3 below. In Appendix B, we complement this analysis with graphs for the largest ad with nonmissing schedule data in our sample, as well as averages for the biggest DMA's in the hours with most data. There, we also present more details on the estimation procedure for Figure 3).

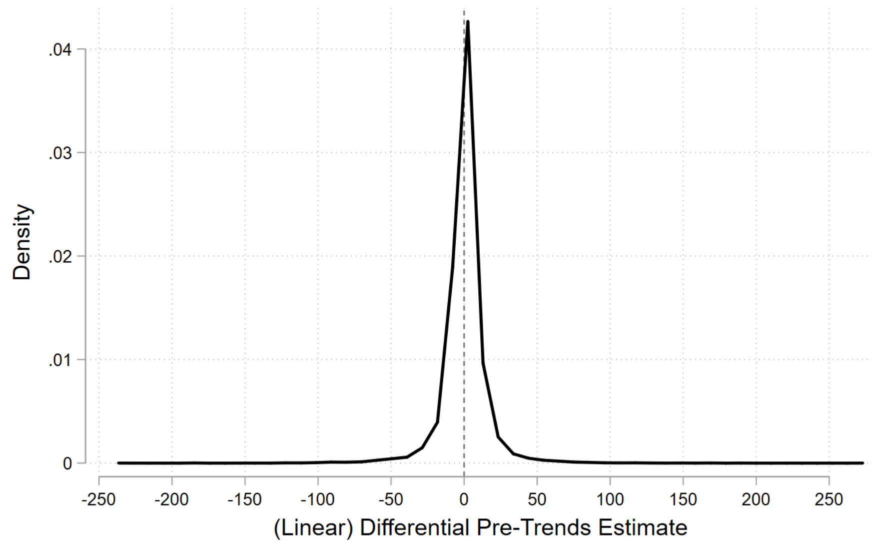
The results suggest the lack of systematic differential pre-trends across treatment and control groups. We also emphasize that anticipation of treatment due to unobservables that could explain the outcome (news viewing) is a priori implausible in our setting, given that "treatment" is defined as being tuned in at the instant an ad begins. Anticipation would require knowledge of the schedule of an ad block, which is unlikely.

Meanwhile, the additional Figures in Appendix illustrate that these results appear to hold in the most influential observations in our sample. We note that the increase in news viewership for both groups before treatment in those figures results from how we define treatment itself: devices must be active at the time of ad airing to be included in either group, implying viewership will be higher than average over the whole 48-hour period, during some or most of which the device may be off.

3.2. Estimation

We estimate equation (3.2) by computing the Ordinary Least Squares estimator for δ_a at the ad-level, exploring our quasi-random variation. For a certain ad k, we compute $\hat{\delta}_k = \frac{1}{N_{a,T}} \sum_{i\,:\,T_{i,a}=1} (y_{k,i,1} - y_{k,i,0}) - \frac{1}{N_{a,C}} \sum_{i\,:\,T_{a,i}=0} (y_{k,i,1} - y_{k,i,0})$. This represents the average change in news viewing time for the treated group minus the average change in news viewing time for the control group for this ad.

FIGURE 3. Pre-Trends



Notes: We present the distribution across ads of the estimated parameter that capture differential pre-trends. Estimates are constructed by regressing pre-treatment hourly average outcomes on a time trend, treatment dummy and an interaction of both. The coefficient we present is the latter, which captures whether treatment and control groups have differential trends in the sample before treatment occurs. Our results suggest that, on average, this is not the case. Additional graphs and details are presented in Appendix B.

We explore the fact that our research design is valid within each ad, and so equation (3.2) can be run at the ad-level. This also means we can estimate different $\{\delta_a\}_{a=1}^{A}$, across all ads $a = 1, ..., A$, identifying differential effects depending on ad characteristics. For example, we can classify ads within subgroups, such as by day of the campaign, time of the day in which it is aired, or the type of campaign it relates to. We can then aggregate all the $\{\delta_a\}$ within these particular subgroups to illustrate average weighted and unweighted effects of campaign ads on news viewership across samples. For standard errors, we proceed by an appropriate bootstrap. We stack all estimates of δ_a together. We then bootstrap from this stacked dataset of estimates and recompute our estimates each time. This bootstrap is valid under the assumption that all estimates we are using to aggregate are derived from the same (asymptotic) distribution.

The procedure above presents multiple advantages relative to standard DiD or event-studies models in our set-up, as we cannot run model (3.2) jointly across ads to simultaneously estimate all $\{\delta_a\}$. In our case, we have tens of millions of observations, making computation of the estimators computationally difficult. Second, doing so would include some households changing treatment status across experiments. This is because treatment and control groups may overlap across ads: someone who watched an ad today might have not watched the equivalent ad the next day.[12] Finally, since we can estimate our parameter of interest δ *for each ad*, we do not suffer the same aggregation problems as dynamic DiD/event-study models have when trying to estimate a single model with multiple treatment effects.[13]

4. Results

Our results across 3 main specifications are shown in Table 2. We present average results across all ads in the first row under different weighting schemes. The other rows present heterogeneous effects by ad-type (whether the ad was sponsored by a candidate for a House race, a presidential one and so forth, and whether the candidate was an incumbent or a challenger). In the first 2 columns, we weigh ads according to treatment and control size - the first, as a geometric mean $\sqrt{n_{t,C}n_{t,T}}/\sum_t \sqrt{n_{t,C}n_{t,T}}$, the second by the number of devices $(n_{t,C}+n_{t,T})/\sum_t(n_{t,C}+n_{t,T})$. These are our preferred specifications, as they account for the sample size within each ad. The third column presents the results by averaging ads without weighting. Standard errors are block bootstrapped.

Our results show that, on average, the treatment effect of a political ad is of approximately 4.5 minutes of news viewing the following day. This result is statistically significant and stable across specifications. As a benchmark, the average viewership of news is approximately 129m per 24h, so we are finding an increase of about 3.5%.[14]

[12] Unfortunately, we are not aware of works in DiD with multiple time periods that can accommodate such a set-up. For example, Callaway and Sant'Anna (2018) assumes that once a subject is treated, he remains treated throughout the sample, an assumption also maintained in Han (2019) and Athey and Imbens (2018). This is clearly not appropriate in our set-up, where viewers in the treated group are not necessarily viewing the next ad. Meanwhile, the procedure that would most closely relate to ours is that of De Chaisemartin and D'HaultfŒuille (2017). Unfortunately, an application of their estimator for fuzzy DiD (i.e. agents can change group status) would require us to track the treatment status of every individual at every ad, which is computationally infeasible.

[13] For example, the latter requires an appropriate definition of what is the parameter of interest (treatment effect relative to which control subgroup) and does not provide appropriate weighting of treatment groups under heterogeneous effects, see Goodman-Bacon (2018); De Chaisemartin and D'HaultfŒuille (2017) for instance. These are issues that do not come up in our estimation.

[14] The level of news viewing in our sample is substantially higher than average because, as noted previously, a large fraction of political ads run on news programs and hence our construction of treatment and control groups selects for types with above-average preference for news programs. Additionally, the period we study (less than 2 months prior to a presidential election) has elevated overall viewership of news programs: in our data, average ratings of news programs were approximately twice as high in October compared to July of 2012.

TABLE 2. Average Effects and Bootstrapped Confidence Intervals, by Subgroup.

	Subgroup	Effect (Minutes) Inv. Variance	Total Devices	Unweighted
		(1)	(2)	(3)
Full Sample		4.258 (4.055, 4.462)	4.339 (4.130, 4.542)	4.524 (4.173, 4.858)
Sponsor	House	3.760 (3.395, 4.165)	3.818 (3.447, 4.226)	1.599 (0.948, 2.217)
	President	4.421 (4.029, 4.814)	4.503 (4.116, 4.891)	6.264 (5.621, 6.919)
	Senate	4.639 (3.952, 5.300)	4.750 (4.062, 5.412)	5.145 (4.187, 6.090)
	Statewide	6.776 (6.216, 7.375)	6.951 (6.397, 7.529)	5.213 (4.378, 5.973)
	Outside Group	2.216 (1.758, 2.655)	2.249 (1.783, 2.694)	4.299 (3.532, 5.064)
Incumbency	Challenger	5.264 (4.911, 5.616)	5.337 (4.988, 5.690)	5.180 (4.596, 5.834)
	Incumbent	4.297 (4.011, 4.633)	4.402 (4.117, 4.749)	4.154 (3.703, 4.622)

Notes:Table reports average effects, weighted by 1) the geometric mean of treatment and control group size, 2) total devices in treatment and control groups and 3) a simple unweighted average. Confidence intervals are the central 95% interval of 1000 bootstrap replicates within each subgroup.

This result contrasts with many standard models of the effect of information on voter's beliefs. The Bayesian framework (e.g. Achen, 1992; Gerber and Green, 1998) would suggest that voters would want to decrease exposure to information with additional ads. In a set-up with Bayesian voters and ads that are (noisy) informative signals, the marginal return to new information has decreased with an additional ad, as viewers' priors have tightened, so more information might no longer be worth the cost. In contrast, we find a positive average effect - political ads are *stimulating* news viewership, and this effect is statistically significant.

Our result also speaks to work as Huber and Arceneaux (2007), which concludes that political advertisements do not engage voters into politics. Even though campaigns do not appear to affect turnout (as found in Ashworth and Clinton, 2007; Krasno and Green, 2008; Spenkuch and Toniatti, 2018), a first measure of engagement, we find that viewers do increasingly "engage" with campaigns after ads through increased news viewing. Our mechanism is both consistent with the literature's results on lack of effects on turnout, as well as their evidence that advertisements are able to persuade voters. In particular, beliefs can change because there is an increased viewership of political news through campaigns, and this effect through media consumption seemingly responds to the type and importance of those campaigns. These effects are significant, but possibly small enough that single ads might not generate observable effects on outcomes such as turnout. However, the composition of multiple ads over the campaign might. We explore the role of dynamics and dynamic effects in the next section.

One explanation for this effect could be voter inattention, since unexpected ads appear to stimulate interest in news. However, the patterns of response indicate a sophisticated rather than naïve inattention. From the heterogeneity effects in Table 2, the estimates are larger in presidential, Senate and especially statewide elections, as well as for ads sponsored by challengers. For example, while the average effect for an ad sponsored by a candidate for the House is just over 1 minute in our preferred specifications, it increases to 3 for presidential campaigns, and to 5 minutes for statewide campaigns. Hence, it appears that political ads have stronger effects for more salient campaigns with a higher marginal impact on policy (gubernatorial, senatorial, presidential), and those of which voters are less likely to be aware of ex-ante (challengers).

Matějka and Tabellini (2017) provide a straightforward theoretical framework that can accommodate such variation. Their model of electoral competition with rationally inattentive voters has variation in the quality of prior information that voters possess as well as variation in costs of acquiring new information across voters. We can think of advertisements as manipulating voters' acquisition-cost parameters. Sponsors with greater resources or greater intrinsic interest - like presidential or gubernatorial candidates - may be more effective in achieving such manipulation. And fixing the cost parameters, voter response will be larger for candidates about which priors are more dispersed, e.g. challengers.[15]

4.0.1. **Decomposing News Viewing: Effects by Channel.** Another dimension of heterogeneity we can explore is to compare effects across different news channels. We decompose our total news viewing outcome $y_{i,a,w}$ into channel specific news viewing outcomes $y^c_{i,a,w}$, where c indexes channels.[16] We recall that the source of ads in our dataset

[15]An alternative framework is that of cheap talk with multiple receivers (e.g. the theory in Farrell and Gibbons, 1989; Goltsman and Pavlov, 2011, and experimental work in Battaglini and Makarov, 2014). In this set-up, politicians are better informed than the public, and choose signals through ads to inform the latter. In practice, however, the real world complexity creates a wedge between this framework and what we can measure - empirically, our receivers do not necessarily receive the signals from the sender, and we face an election in which multiple politicians (senders) compete among themselves when sending signals.

[16]We break y into separate categories for each of the four main cable news channels CNN, Fox News Channel (FNC), Headline News (HLN), MSNBC; a combined category for the public channels CSPAN and

are local affiliates, as there are no ads on national cable channels like Fox News in this data.[17] The results are presented in Table 3. They show that increases in the viewing of political news are predominantly in local news through local affiliates.

One explanation consistent with these results is that treated viewers have preferences for local news to begin with. This is because treatment was defined exactly on those watching the local affiliates, and frequently local news programs, on which political ads ran. Hence, viewers seem to increase consumption of news on their initially-preferred channels. Table 3 also suggests that the increase in local news viewing is offset by a small amount of substitution away from national channels.

TABLE 3. Average Effects and Bootstrapped Confidence Intervals, by Channel.

	Effect (Minutes)						
	CNN	CSPAN	FNC	HLN	MSNBC	OTHER	NETWORK
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treated	−0.527	−0.063	−1.279	−0.372	−0.472	−0.208	7.466
	(−0.721, −0.333)	(−0.257, 0.131)	(−1.474, −1.085)	(−0.566, −0.178)	(−0.666, −0.278)	(−0.402, −0.014)	(7.272, 7.661)

Notes: Table reports average effects of news viewership on the indicated channel, weighted by the geometric mean of treatment and control group size. OTHER is all other national cable networks, and NETWORK is all local network affiliates. Confidence intervals are the central 95% interval of 1000 bootstrap replicates within each subgroup.

By increasing news viewing over the campaign, the treated group also changes their likelihood of future treatment (exposure to more ads), an effect we explore in the next section.

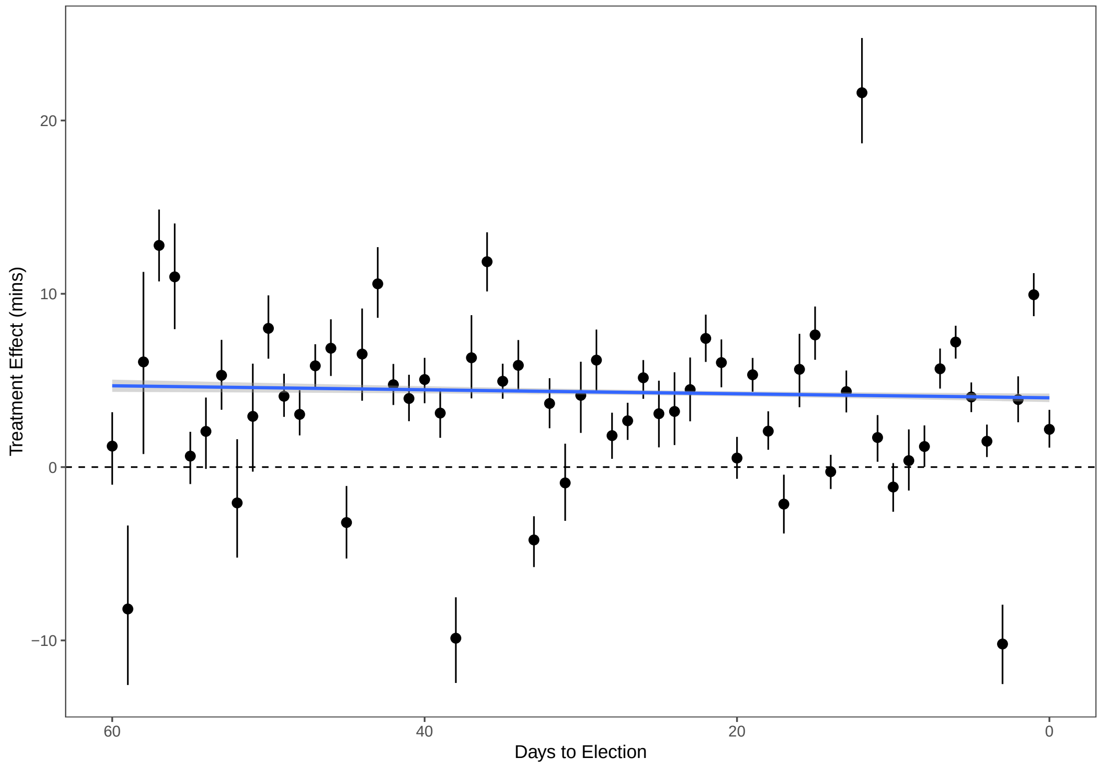
4.0.2. **Dynamics and Long-term Effects.** First, we plot the daily average effects $\{\bar{\delta}_1, \bar{\delta}_2, ..., \bar{\delta}_T\}$, disaggregated at the daily level from our results in Table 2. The daily averaged effects are shown in Figure 4, along with a fitted linear trend line which shows that the average short-term effects on news viewing decline (albeit slowly) over the course of the campaign.

While the short-term effects ($\bar{\delta}_t$, the effect of an ad on news viewing in the next 24 hours) are largely constant, that does not mean the overall effect of an ad remains so. This is because in the beginning of the campaign, an additional 4-5 minutes of political news in the next day can stimulate further news viewing in the following day, as well as the possibility of new ad exposures further on. Hence, being exposed early on allows many days of further news viewing, while exposure late in the campaign (for example, in the last day), would only allow the sum of few short-lived effects.

CSPAN2; a combined category for any local affiliate of the broadcast networks; and a combined category for all other cable channels.

[17]Our treatment and control groups were defined on viewers watching ads on local TV at the time, so they would not be watching national channels concurrently.

FIGURE 4. Treatment effects across the campaign



Notes: We present the average treatment effects at the daily level (average effects across all ads within a day, with each given an equal weight). 95% bootstrapped confidence intervals for ads by day. The blue line is the least squares fit of the estimated effect (at ad level) on days to election, weighted by the geometric mean of treatment and control group size. This regression has significantly positive slope (coefficient= 0.012, SE = 0.004), indicating that treatment effects fall over the course of the campaign (i.e. effects increase with the number of days remaining in the campaign).

To appropriately quantify the long-term effects of an ad over the campaign, we estimate the same model as in equation (3.2) using the number of political ads a viewer is exposed to in a given time window (again, ± 24 hours around the initial ad air time) as the outcome. Hence, the treatment effect is how many more ads does one political ad lead treated viewers to watch relative to the control over the course of the next 24 hours. Results are shown in Table 4.

We find, as expected, a positive and significant effect on future ad exposure. A viewer in the treatment group will expect to be exposed to about 0.6 additional political ads, on average, compared to a viewer in control group, over the next 24 hours. Hence, a straightforward estimate of the total (dynamic) effect of ads on news viewing can be obtained by computing the geometric sum $\sum_{\tau=0}^{T-t} \delta\rho^{\tau}$, where ρ is the estimated increase in future ad exposure over the control group. Plugging in estimates from Tables 2 and 4, we obtain a multiplier of 2.62 for ads aired sixty days prior to election time, compared to the day before the election.

TABLE 4. Average Effects and Bootstrapped Confidence Intervals, by Subgroup.

	Effect (Additional Ads)		
	Inv. Variance	Total Devices	Unweighted
	(1)	(2)	(3)
Treated	0.618	0.640	0.959
	(0.411, 0.824)	(0.434, 0.847)	(0.753, 1.166)

Notes: Table reports average effects, weighted by 1) the geometric mean of treatment and control group size, 2) total devices in treatment and control groups and 3) a simple unweighted average. Confidence intervals are the central 95% interval of 1000 bootstrap replicates within each subgroup.

5. Effects on Tune-Out

Our results to this point are all in intent-to-treat terms: we define "treatment" as having been tuned in to the program on which the ad aired at the instant the ad *began*. In this section, we investigate viewers' choices to select out of exposure to political ads, by focusing attention on the set of treated devices and examining (variation in) the likelihood of sitting through the full ad without changing the channel, conditional on having been viewing at the instant the ad began. This analysis illuminates another source of dynamic variation in the effects strategic incentives to advertise over the course of the campaign.

Interpreting tune-out frequencies requires an appropriately constructed baseline for comparison. Just as a post-treatment increase in an outcome cannot be interpreted as a causal effect without reference to the counterfactual change estimated by the corresponding increase in an untreated control group, the rate of tune-out for political ads only has meaning in relation to the comparable rate for other content.

We construct an ad-level baseline rate for comparison by the following procedure. For every device tuned in at the instant an ad aired (e.g., the treated group for some ad), we compute an indicator $o^A_{i,a} \in \{0, 1\}$ which is 1 if the device registered a tuning event prior to the end of the ad and 0 otherwise. We then use the device's viewing history to randomly select a device-specific time somewhere else in the day that satisfies our criteria for inclusion in an ad treatment group (device is on, and has had tuning event in the previous hour).[18] We denote this randomly selected time by $r_{i,a}$ and measure an indicator $o^R_{i,a} \in \{0,1\}$ which is 1 if the device registered a tuning event between $r_{i,a}$ and $r_{i,a} + 30$ and 0 otherwise. For each ad we construct average tune-out rates $O^A_a = \sum_{i\,:\,T_{i,a}=1} o^A_{i,a}/|T_{i,a} = 1|$ and $O^R_a = \sum_{i\,:\,T_{i,a}=1} o^R_{i,a}/|T_{i,a} = 1|$.

The difference between the two rates is a measure of the ad's (dis-)utility relative to the average utility of TV viewing.[19] We plot results aggregated to the daily level in Figure 5.

As we can see in the plot, tune-out rates for political ads are generally lower (by 20-50%), compared to the rate of tune-out among the same set of viewers at comparable times of the same day.[20] Viewers avoid political ads at rates that are low both in absolute terms and relative to the average content that the treated groups consume at other times.

Tune out rates at the randomly selected times are approximately flat over the campaign. Tune-out rates for political ads, on the other hand, have a statistically significant increasing trend as election day approaches. This differential trend suggests that (some) viewers indeed become saturated with political advertising over the course of the campaign, and actively avoid it more as the campaign goes on. Hence, the effective reach of political advertising is higher earlier in the campaign than later on.

We also examine the time trend in tune-out in a regression framework, in Table 5. This table shows results of regressions where the outcome is an indicator for a viewer tuning out prior to the end of a political ad, conditional on being active and tuned in to the ad's channel at air time (e.g., conditional on being in the "Treated" group for that ad). The columns split the sample by household partisanship: column (1) is households with unknown or independent partisanship; column (2) is households with a Republican-registering head of household, and column (3) is households with a Democratic-registering head of household. All specifications include household fixed effects, to account for differential baseline tune-out likelihood across different viewers.

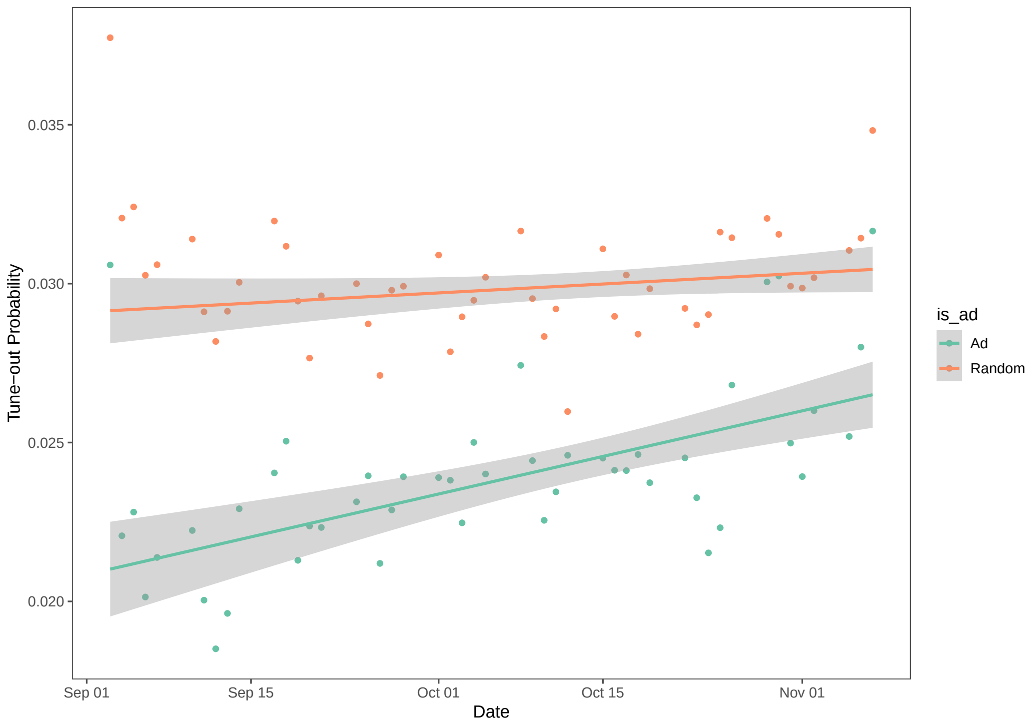
Across all three groups, the time trend is negative and statistically significant, indicating that tune-out is more likely the fewer days there are remaining until election. This is consistent with the visual evidence in Figure 5, and indicates that the trend is not solely driven by changes in composition of advertising-exposed viewers over time.

[18] We exclude the 30 minutes prior to the ad air time and the duration of the ad itself, and sample uniformly over all other seconds that satisfy the two stated criteria.

[19] This measure can be interpreted as an (up-to scale) approximation of consumer surplus under small shares of tuning-out and some additional structure (see Appendix E).

[20] Base rates for both categories are low, in the 2-3% range. "Compliance" with treatment is thus very close to 1, and our intent-to-treat estimates in the previous will be close to the average treatment effect on the treated.

FIGURE 5. Tune-Out Probability of Political Ads



Notes: We present estimates of the tune-out rate, the fraction of people who tuned out before the ad ends. We compare to the fraction who tuned out at a device-specific randomly selected time somewhere else in the day and that satisfies our criteria for treatment (device is on, and has had tuning event in the previous hour). This generates an average rate of tune out for other content for comparison to the rate among political ads. The difference between the two rates also approximates consumer surplus under more theoretical structure, see Appendix E.

Our individual-level data also allows us to assess whether viewers selectively expose themselves to ads from candidates of one or the other party. The first two rows in Table 5 show the effect of sponsor party on tune-out probability for each partisan group. There is some evidence of selective exposure on partisan congruence, as Republicans are less likely to tune out an ad by a Republican sponsor than a Democratic one. There is no similar effect for Democrats, an asymmetry which aligns with experimental findings on selective exposure in Henderson and Theodoridis (2018). Independents and viewers of unknown partisanship look similar to Republicans in this sample (likely reflecting the baseline rightward skew of the markets included in the dataset, which tend to encompass smaller regional cities and surrounding suburbs). Overall, though, the magnitude of partisan selective exposure is small: the party effects are comparable to moving an additional 10-20 days away from the election, and party explains only a small amount of variation in tune-out. Two ads on the same program, one sponsored by a Republican and one by a Democratic candidate, would have almost the same effective audience.

TABLE 5. Differential Tune-out by Party and Timing.

	Tuned Out		
	Indep / Unknown	Rep	Dem
	(1)	(2)	(3)
Dem Sponsor	0.0005**	−0.0001	−0.0004
	(0.0002)	(0.001)	(0.0003)
Rep Sponsor	−0.001***	−0.002***	−0.001***
	(0.0002)	(0.0005)	(0.0003)
Days to Election	−0.0001***	−0.0001***	−0.0001***
	(0.00000)	(0.00001)	(0.00001)
Household FE:	Y	Y	Y
Number of Households:	96979	22868	49267
N	9,625,597	1,456,304	4,239,205
R^2	0.050	0.056	0.051

*p < .1; **p < .05; ***p < .01
An observation is a household-ad. The sample is all households active and tuned in to the channel on which a political ad ran at the time the ad began. Column (1) restricts to households with independent or unknown party affiliation; column (2) restricts to Republican-identifying households, and column (3) restricts to Democratic-identifying households. The omitted category is ads sponsored by outside groups. Standard errors (clustered by household) in parentheses.

6. Conclusion

In this paper, we have provided evidence that exposure to political ads on television increases viewership of television news. We used quasi-random variation based on active viewers and those who tuned out of political ads just before they were screened. We find that an unexpected exposure to a political ad increases news viewing in the next day by approximately 4-5 minutes. Our effect appears robust across specifications, and does not appear to be driven by imbalance on predetermined characteristics or by differential pre-exposure viewing trends. With very fine data, we can look at the effects across different periods of the campaign and time of the day. We found that the effect was strongest from ads in presidential and statewide races, particularly from ads sponsored by challengers. Altogether, our results are consistent with viewer search costs for political information (or, equivalently, interest in the campaign) being altered by advertisements - a new mechanism.

The theoretical underpinning for this mechanism is different than a standard Bayesian model, where ads are noisy signals of candidate attributes or the state of the world. Here, a "randomized" stimulus from a campaign leads to more information acquisition activity, particularly so for candidates about which voters have relatively high value of information and relatively low initial precision. These findings are consistent with models of rational inattention where voters face nonzero costs of acquiring information, e.g. Matějka and Tabellini (2017). The heterogeneity in our results suggests that certain sponsors are more effective at using advertising to manipulate voter search costs than others.

Our mechanism suggests a role of advertisements as changing the distribution of viewers (voters) who are politically informed - not directly, but through inducing voters to inform themselves through other sources. Given the ebb and flow of political news throughout a campaign, the timing of advertising also has the potential to affect *what* voters learn. Our results suggest that candidates should advertise when news about them is good, and avoid it when news is bad. Even if campaigns cannot control the timing of information releases in the media (as in Gratton et al., 2017) they retain some control over which releases voters are likely to actually absorb. Hence, imbalances in advertising expenditures across opponents for the same office can alter the informational picture that voters have available at election time, allowing for effects of advertising advantages (Gordon and Hartmann, 2013; Spenkuch and Toniatti, 2018) on election outcomes even in the absence of any direct persuasive effect.

The dynamic variation in strategic incentives to advertise we uncover helps to explain the empirical fact that campaigns do not concentrate activity at the very end of a campaign, despite high-quality empirical evidence (Gerber et al., 2011; Kalla and Broockman, 2018) of rapid decay in ads' persuasive potential. There are strong trade-offs involved in waiting: campaigns miss the option to induce information acquisition at times when media signals are most favorable, and may find themselves at the end of the campaign facing an audience already saturated with campaign news and uninterested in hearing their message. The results that ads induce additional exposures to future ads, that the news consumption-stimulating effects decline over time, and tune-out rates increase over time, all push in the direction of advertising early.

On the methodological front, our work engages with the complexities of aggregating event studies with heterogeneous effects. With many quasi-experiments and a large panel, the implementation of improved event-studies estimators is far from trivial. As datasets increase, we expect that developing new methodologies might prove appropriate for such contexts.

While we do not provide a formal theoretical model, our results are informative for models of electoral competition and dynamic information acquisition. They also illustrate the complexity of fully embodying such campaign dynamics. A model of dynamic effects would have to incorporate multiple layers. First, politicians may anticipate that their ads will be watched by certain groups of viewers, and also anticipate that the distribution of such viewers will then change over time as a result. Second, there are multiple politicians who might be competing with one another when sending ads, not just over content or quantity but also on the dimension of congruence with media reports. Third, in contrast to existing models, politicians may target voters who are not actively acquiring information in a campaign, with the intention of engaging and them. Finally, viewers might have different degrees of prior information and baseline interest in political information. While outside of the scope of this work, this research agenda seems promising to us.

References

Abraham, S. and Sun, L. (2018). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Available at SSRN 3158747*.

Achen, C. H. (1992). Social psychology, demographic variables, and linear regression: Breaking the iron triangle in voting research. *Political behavior*, 14(3):195–211.

Ashworth, S. and Clinton, J. D. (2007). Does advertising exposure affect turnout? *Quarterly Journal of Political Science*, 2(1):27–41.

Athey, S. and Imbens, G. W. (2018). Design-based analysis in difference-in-differences settings with staggered adoption. Technical report, National Bureau of Economic Research.

Battaglini, M. and Makarov, U. (2014). Cheap talk with multiple audiences: An experimental analysis. *Games and Economic Behavior*, 83:147–164.

Berelson, B. R., Lazarsfeld, P. F., McPhee, W. N., and McPhee, W. N. (1954). *Voting: A study of opinion formation in a presidential campaign*. University of Chicago Press.

Callaway, B. and Sant'Anna, P. H. (2018). Difference-in-differences with multiple time periods and an application on the minimum wage and employment. *arXiv preprint arXiv:1803.09015*.

Chiang, C.-F. and Knight, B. (2011). Media bias and influence: Evidence from newspaper endorsements. *The Review of Economic Studies*, 78(3):795–820.

De Chaisemartin, C. and D'HaultfŒuille, X. (2017). Fuzzy differences-in-differences. *The Review of Economic Studies*, 85(2):999–1028.

DellaVigna, S. and Kaplan, E. (2007). The fox news effect: Media bias and voting. *The Quarterly Journal of Economics*, 122(3):1187–1234.

Farrell, J. and Gibbons, R. (1989). Cheap talk with two audiences. *The American Economic Review*, 79(5):1214–1223.

Ferraz, C. and Finan, F. (2008). Exposing corrupt politicians: the effects of brazil's publicly released audits on electoral outcomes. *The Quarterly Journal of Economics*, 123(2):703–745.

Gerber, A. and Green, D. P. (1998). Rational learning and partisan attitudes. *American Journal of Political Science*, 42:794–818.

Gerber, A. S., Gimpel, J. G., Green, D. P., and Shaw, D. R. (2011). How large and long-lasting are the persuasive effects of televised campaign ads? results from a randomized field experiment. *American Political Science Review*, 105(1):135–150.

Goltsman, M. and Pavlov, G. (2011). How to talk to multiple audiences. *Games and Economic Behavior*, 72(1):100–122.

Goodman-Bacon, A. (2018). Difference-in-differences with variation in treatment timing. Technical report, National Bureau of Economic Research.

Gordon, B. R. and Hartmann, W. R. (2013). Advertising effects in presidential elections. *Marketing Science*, 32(1):19–35.

Gratton, G., Holden, R., and Kolotilin, A. (2017). When to drop a bombshell. *The Review of Economic Studies*, 85(4):2139–2172.

Han, S. (2019). Identification in nonparametric models for dynamic treatment effects. *Journal of Econometrics*.

Henderson, J. A. and Theodoridis, A. G. (2018). Seeing spots: Partisanship, negativity and the conditional receipt of campaign advertisements. *Political Behavior*, 40(4):965–987.

Huber, G. A. and Arceneaux, K. (2007). Identifying the persuasive effects of presidential advertising. *American Journal of Political Science*, 51(4):957–977.

Kalla, J. L. and Broockman, D. E. (2018). The minimal persuasive effects of campaign contact in general elections: Evidence from 49 field experiments. *American Political Science Review*, 112(1):148–166.

Kendall, C., Nannicini, T., and Trebbi, F. (2015). How do voters respond to information? evidence from a randomized campaign. *American Economic Review*, 105(1):322–53.

Krasno, J. S. and Green, D. P. (2008). Do televised presidential ads increase voter turnout? evidence from a natural experiment. *The Journal of Politics*, 70(1):245–261.

Lazarsfeld, P. F., Berelson, B., and Gaudet, H. (1944). *The people's choice*. Duell, Sloan & Pearce.

Lovett, M. and Peress, M. (2015). Targeting Political Advertising on Television. *Quarterly Journal of Political Science*, 10(3):391–432.

Lovett, M. J. and Staelin, R. (2016). The role of paid, earned, and owned media in building entertainment brands: Reminding, informing, and enhancing enjoyment. *Marketing Science*, 35(1):142–157.

Martin, G. J. and Yurukoglu, A. (2017). Bias in cable news: Persuasion and polarization. *American Economic Review*, 107(9):2565–99.

Matějka, F. and Tabellini, G. (2017). Electoral competition with rationally inattentive voters. *Available at SSRN 3070204*.

Ridout, T. N. and Smith, G. R. (2008). Free advertising: How the media amplify campaign messages. *Political Research Quarterly*, 61(4):598–608.

Spenkuch, J. L. and Toniatti, D. (2018). Political Advertising and Election Results*. *The Quarterly Journal of Economics*, 133(4):1981–2036.

Steiner, J., Stewart, C., and Matějka, F. (2017). Rational inattention dynamics: Inertia and delay in decision-making. *Econometrica*, 85(2):521–553.

Vavreck, L. (2009). *The message matters: The economy and presidential campaigns*. Princeton University Press.

Appendix A. Additional data details

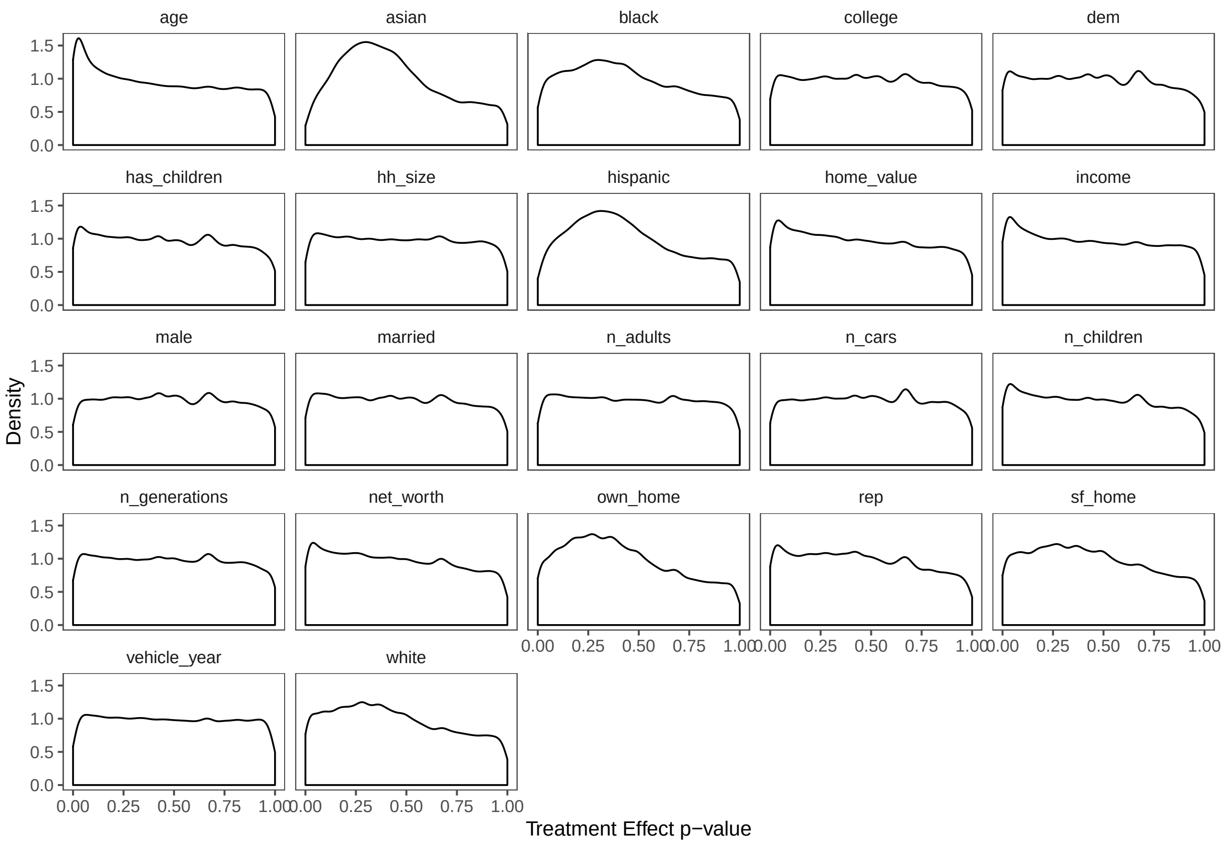
TABLE 6. Additional information on the covariate names, shown in Figure 2 and Figure 6

Name in the Graph	Covariate Details
age	Age of the head of household
white	Indicator variable if head of household is white
asian	Indicator variable if head of household is Asian
black	Indicator variable if head of household is African-American
college	Indicator variable if head of household is a college graduate
dem	Indicator variable if head of household is a Democrat (from party of registration)
rep	Indicator variable if head of household is a Democrat (from party of registration)
has_children	Indicator variable if household has children
hh_size	Number of individuals in the household
hispanic	Indicator variable if head of household is Hispanic
home_value	Home value of the household
income	Household income
male	Indicator variable if head of household is male
married	Indicator variable whether head of household is married
n_adults	Number of adults in the household
n_cars	Number of owned cars in the household
n_children	Number of children in the household
n_generations	Number of different family generations living in the household
net_worth	Estimated net worth at the household level
own_home	Indicator variable whether the home being lived in is owned
sf_home	Size of home in square feet
vehicle_year	Year of newest owned vehicle

Appendix B. Further Checks on Balance Tests and Pre-Trends

We begin with another representation of the results on the balance of covariates.

FIGURE 6. Additional Balance Tests - Estimated Treatment Effects on each Covariate



Notes: We present the distribution of p-values for different covariates under the null that the treatment effect of that covariate is equal to 0. The p-values should follow a uniform distribution if the null hypothesis that the treatment effect is 0 is true. We can see that, in general, our data is well balanced as the covariates are not significantly different across treatment and control. This graph complements Figure 6.

B.0.1. **Further Details about Figure 3.** Let us describe the specification which generates Figure 3. Denote $h \in \{-24, -23, -22, ..., -1, 0, 1, ..., 24\}$ as indexing hourly bins, with treatment occurring at period "0". In other words, bin $h = -1$ refers to the period between treatment and 1 hour before treatment. The average amount of news viewing in this interval for ad k for the treated group is given by $\bar{y}_{k,T,h} = \frac{1}{N_{k,T}} \sum_{i\,:\,T_{i,k}=1} y_{k,i,h}$. Then, the estimates from Figure 3 comes from running an analogous equation to (3.2) for all periods *before* treatment:

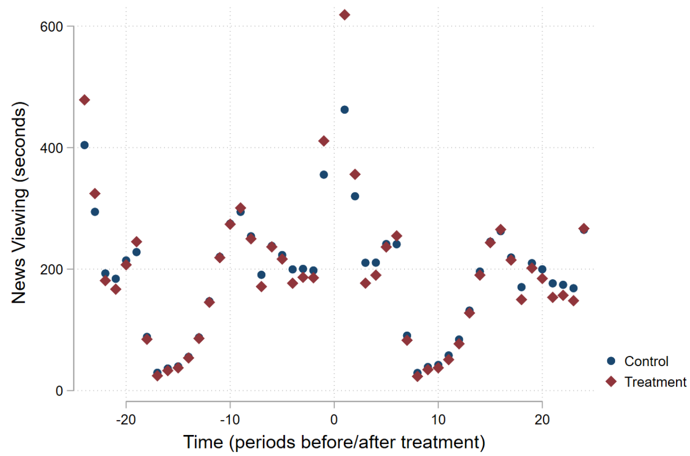
$$\bar{y}_{a,g,h} = \beta_{1,a} + \beta_{2,a} T_{g,a} + \beta_{3,a} h + \beta_{4,a} T_{g,a} h + \varepsilon_{g,a,h}, \tag{B.1}$$

for periods $h < 0$, with $g \in \{C, T\}$.

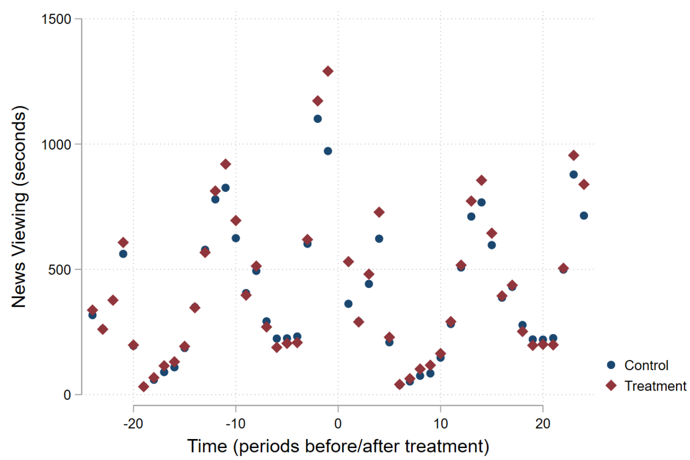
$\beta_{4,a}$ is our parameter of interest: it captures differential trends in outcomes between treated and control groups over time. Since the sample is restricted to pre-treatment observations, it is informative of differential pre-trends in outcomes. This is the parameter whose distribution across ads is plotted in Figure 3.

B.0.2. **Additional Figures for Pre-trends.** We now show graphs on the pre-trends for the biggest ad (with full data) in our sample, as well as for the hours with the most observations for the two the biggest DMA's in our sample.

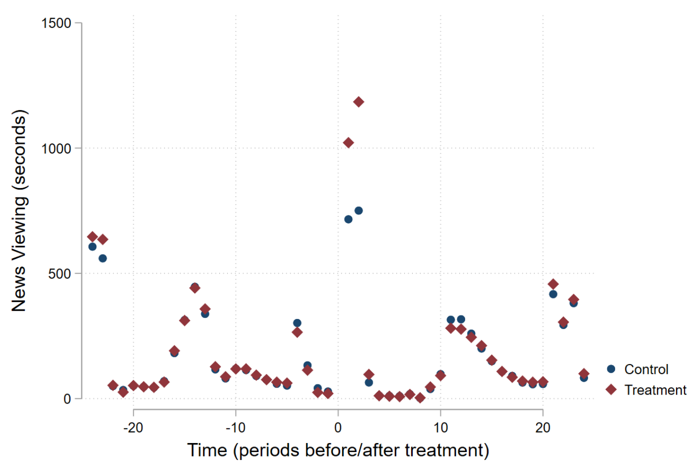
FIGURE 7. Additional Pre-Trend Graphs



(A) Largest Ad (with full data) - aired during College Football on an affiliate of FOX, DMA Charleston-Huntington



(B) Ads in DMA Tampa- St. Pete (Sarasota), 21pm (UTC time)



(C) Ads in DMA Roanoke-Lynchburg, 23pm (UTC time)

Appendix C. Robustness Checks

As shown in Figure 6, one concern with our empirical specification could be that some variables, most notably age and income, appear unbalanced across a larger share of treatment and control groups than what should be expected. To check the robustness of our results from Table 2 to this dimension, we run our main specifications again on the subset of ads (treatments) for which the p-value for treatment balance is larger than 0.05. The results are shown below in Tables 7-8. We can see that our main conclusions remain unchanged.

TABLE 7. Average Effects and Bootstrapped Confidence Intervals: Age-balanced ads only.

	Inv. Variance	Effect (Minutes) Total Devices	Unweighted
	(1)	(2)	(3)
Treated	6.817	6.976	7.126
	(6.671, 6.962)	(6.830, 7.122)	(6.980, 7.272)

Notes: Table reports average effects, weighted by 1) the geometric mean of treatment and control group size, 2) total devices in treatment and control groups and 3) a simple unweighted average. Confidence intervals are the central 95% interval of 1000 bootstrap replicates. The sample is restricted to the set of ads with p-value greater than 0.05 for the hypothesis that treatment and control groups have no difference in mean age.

TABLE 8. Average Effects and Bootstrapped Confidence Intervals: Age-balanced ads only.

	Inv. Variance	Effect (Minutes) Total Devices	Unweighted
	(1)	(2)	(3)
Treated	6.105	6.241	7.010
	(5.912, 6.299)	(6.047, 6.435)	(6.817, 7.204)

Notes: Table reports average effects, weighted by 1) the geometric mean of treatment and control group size, 2) total devices in treatment and control groups and 3) a simple unweighted average. Confidence intervals are the central 95% interval of 1000 bootstrap replicates. The sample is restricted to the set of ads with p-value greater than 0.05 for the hypothesis that treatment and control groups have no difference in mean income.

Another concern could be related to households receiving multiple treatments over the campaign. Current event studies models typically assume that treatment status does not change (e.g. once a household is treated, they remain so[21]), or the dosage of treatment is the same (e.g. the household would be treated once). In this regard, we note three features of our set-up and results. First, we have shown pre-trends in news consumption to be similar and groups to be balanced, which includes past viewership of ads. Second, our definition of treatment and control is local in nature (i.e. ads on local news), implying that overlap in treatment statuses across ads at similar times is small. Finally, our short-term outcomes treatment effect dissipates quickly, as can be seen by Figure 3. Hence, persistent effects of past treatments would only occur in changing the probability of treatment through increased media consumption. This exactly what we study in our section on dynamic effects of ads.

Appendix D. Missing Programming Data

As mentioned in the main text, one limitation of our data is that the program database (Fourthwall) is incomplete, an issue that is most pronounced for local network affiliates in certain DMAs, primarily in the West. National cable channels have nearly complete coverage. The national cable channels have an average of roughly 30 news programs per day in our data, and have some program entry in all 154 days of our coverage period. The local affiliates average about 9.6 news programs per day on days with coverage, but almost 70% of total station-days are missing. There are 71 stations in 15 DMAs (out of 663 and 60 respectively) that are largely unaffected by schedule missingness. These coincide with markets with the largest number of subscribers in our data. We restrict our analysis to this subset of DMAs in the robustness analysis below.

While solutions to this issue are limited by the data availability, we do not believe it poses a significant threat to our analysis, for several reasons.

First, this missingness cannot produce bias away from zero in the estimates in our set-up. This is because the variation we use is within media markets (i.e. within the same program schedules). Hence, the measurement error induced by such incompleteness cannot be correlated with treatment group status. It is possible that missing program data generates bias towards zero, as we are effectively plugging in zeros for the outcome variable on days with missing schedule data. In fact, our estimates rise when we include only markets with relatively complete schedule data; see Table 9 below.

Regarding the efficiency of our estimates, which would be affected by this incompleteness: our results in Table 2 are similar whether weighted by sample size or unweighted.

[21]Consider the leading example in Callaway and Sant'Anna (2018) of the effect of increases in the minimum wage across states in the U.S. Once a state has increased their minimum wage they remain with this higher value (i.e. treated), a status that is essentially irreversible.

This is suggestive that average estimates are similar across ads with smaller and larger number of viewers within media markets. Hence, data incompleteness does not appear to affect the efficiency of our estimates when this missingness is correlated with viewership numbers within markets.

However, data incompleteness could be correlated to the magnitude of effects in media markets. For example, data incompleteness could be higher in markets where ads are less effective. This does not appear to be the case: when we drop media markets with the most missing data in our sample and re-estimate the specifications in Table 2, we find even larger results than before. These are shown in Table 9 below.

TABLE 9. Average Effects and Bootstrapped Confidence Intervals: Dropping DMAs with missing local news program data.

		Effect (Minutes)	
	Inv. Variance	Total Devices	Unweighted
	(1)	(2)	(3)
Treated	8.078	8.313	14.513
	(7.883, 8.273)	(8.118, 8.508)	(14.318, 14.708)

Notes: Table reports average effects, weighted by 1) the geometric mean of treatment and control group size, 2) total devices in treatment and control groups and 3) a simple unweighted average. Confidence intervals are the central 95% interval of 1000 bootstrap replicates. Sample drops ads run in DMAs with extensive missingness of the local news program schedule.

Appendix E. An Empirical Measure of Distaste for Political Ads

Consider the set of consumers who are watching TV at period t, when the political ad comes on. They have the choice between continuing to watch TV and tuning out. Assume that the utility for a representative viewer in a district i, of tuning out at t, conditional on viewing at $t-1$ is given by:

$$V_{i,t} = \gamma_t + \varepsilon_{i,t},$$

where γ_t is a period fixed effect, and $\varepsilon_{i,t}$ is an i.i.d. standard Logit shock. For example, γ_t captures variation in the willingness to watch political ads during the day (during working hours, late night shows etc.). We normalize the utility from continuing to watch TV to 0.

In this case, viewer i tunes out if $V_{i,t} \geq 0$. The share of viewers that do so is given by:

$$\pi_t = \int \pi_{i,t} di = \int \frac{e^{\gamma_t}}{1+e^{\gamma_t}} di, \tag{E.1}$$

where $\pi_{i,t}$ is the probability that i tunes out. Equation (E.1) is derived from the distributional assumption on $\varepsilon_{i,t}$. To find the aggregate share, we integrate over the mass of consumers. In a large market, the estimate of π_t, given by $\hat{P}_t$ will also be close to $\pi_{i,t}$.

π_t can be estimated directly from the data, and plotted over time. A naïve way to estimate it would be to look at the share of viewers who switch out once the ad comes on at t. Nevertheless, viewers also tune out for reasons that are not just ad-specific. Hence, we estimate π_t by the fraction of people who tuned out before the ad ends minus the fraction who tuned out at a device-specific randomly selected time somewhere else in the day and that satisfies our criteria for treatment (device is on, and has had tuning event in the previous hour). This generates an average of viewers that tune out, accounting for tuning out behavior not driven by the political ad.

In this simple set-up, the expected consumer surplus is given by (up to a constant):

$$\begin{aligned} \mathbb{E}CS_{i,t} &\propto ln(1+e^{\gamma_t}) \\ &= ln\left(1+\frac{\pi_{i,t}}{1-\pi_{i,t}}\right), \end{aligned} \tag{E.2}$$

where the first line uses a standard property of the Logit distribution for discrete choice models, and the second line rewrites $\pi_{i,t}$ using equation (E.1). The scale comes from converting utility into dollars, such as the marginal utility from an extra dollar.

The consumer surplus (up-to scale) can then be estimated by just using $\hat{P}_t$ from before:

$$\hat{\bar{C}S_t} = ln\left(1 + \frac{\hat{P}_t}{1 - \hat{P}_t}\right), \tag{E.3}$$

and can also be plotted over time.

Hence, plotting the estimates of both (E.1) and (E.3) allows to compare not only how the share of consumers tuning out changes over time (a measure for distaste), but also the alternative measure of how the consumer surplus from that is changing, completing the view of the dynamic incentives for advertisement.

For low values of $\frac{\hat{P}_t}{1-\hat{P}_t}$, we find that our results in Figure 5 are approximately those in equation (E.3), since $ln(1+x) \approx x$ and $\frac{z}{1-z} \approx z$ for small x, z.