Using Big Data and Algorithms to Determine the Effect of Geographically Targeted Advertising on vote intention: Evidence from the 2012 U.S. Presidential Election

Tobias Konitzer*[1], David Rothschild†[2], Shawndra Hill‡[2], and Kenneth Wilbur§[3]

[1]Stanford University

[2]Microsoft Research

[3]University of California - San Diego

March 20, 2018

*tobiask@stanford.edu
†davidmr@microsoft.com
‡shawndra@microsoft.com
§kcwilbur@ucsd.edu

Abstract

We develop a new conceptualization of political advertising effects by looking at the effect of the marginal advertising dollar during the heat of presidential campaigns. We argue that in contrast to other studies investigating effects of political ads, our approach is more apt to capture the natural environment in which political ads are encountered during a presidential campaign. We focus on the intense inundation of political ads voters are confronted with in swing states in the weeks leading up to the presidential election, and argue that it is unclear a priori whether we should expect advertising to affect vote intention in that critical circumstance. We empirically validate this hypothesis using a trove of data from the 2012 campaign: daily polling in media markets around the country, detailed data on all registered voters in the country, all TV advertisements by market and exact air-time, and the entire Twitter corpus. We find that neither overall increases in advertising spending nor partisan imbalances in spending expanded the candidates' electorate. In fact, total DMA-level spending significantly moderates a negative relationship between spending advantages and advantages in vote intention, suggesting a boomerang effect of additional spending late in the campaign. In closing, we discuss the ramifications of our findings for future research, and stress the importance of research tracking advertising effects.

The Citizens United vs. Federal Election Commission ruling in 2010 marked a paradigm shift for the nature and scope of modern political campaigns. The next presidential election in 2012 was the first to see the rise of super PACs and with these the virtual elimination of spending limits. The Republican and Democratic campaigns combined spent over $2 billion during the 2012 presidential election (Vogel and Parti, 2013). The largest portion of this amount was poured into television advertising; combined, the candidates spent close to $900 million on media-market-level ads inserted into national broadcast and cable network programming between April 11, 2012 and election night on November 6, 2012.[1] National advertising accounted for almost 42% of overall campaign expenditures and was more than double the spending in 2008, making political advertising the single-most-expensive method of campaigning in 2012.

The focus on television advertising is understandable. While television ads are "about as precise as carpet-bombing" (Bradley, 2010), advertisers can reach a large audience in one shot. Further, campaigns can rely on audience demographics for stations and programs to reach desirable targets and thus benefit from the rise of niche television stations or programs (e.g., the Golf Channel), which create increasingly tightly partitioned sub-groups of the voting population. To be sure, technological advancements regarding the identifiability of online users mean that campaigns can more precisely microtarget online advertisements. Indeed, the 2012 presidential candidates spent an unprecedented $78 million on digital advertising, representing a 251% increase over 2008 figures (Ballard, Hillygus and Konitzer, 2016). By some

[1] These figures are estimated by Kantar Media and exclude costs for locally purchased cable ads.

reports, this figure was expected to have grown by 20% in 2016 (Sullivan-Jeniks, 2016). Nonetheless, the unique combination of broad reach and precise targeting, paired with the unabated popularity of TV and high levels of trust people place in the medium (Wilbur, 2015), make television the ideal place for political advertising. Consequently, expenditures for television advertising were expected to swell even further during the 2016 campaign (Kurtzleben, 2015).

However, in defiance of expectations, the 2016 campaigns saw a severe *decrease* in advertising spending. Total advertising spending in 2016 was slightly below $570 million – only slightly more than 60% of the total amount spent in 2012 (Kantar, 2016). Further, Republican candidate Donald Trump spent a mere $160 million – less than a third of what Republican candidate Mitt Romney spent in 2012, while Clinton spent about the same as Obama in 2012 – $410 Million vs. $404 Million (Kantar, 2016). In effect, 2016 represents the first campaign cycle that saw a huge imbalance in advertising spending: the Republican candidate spent less than 30% of overall two-party spending.[2] Despite this imbalance, Donald Trump handily won the Electoral College (while losing the popular vote).[3]

In this light, it is also unsurprising that conclusions about the effectiveness of political advertising with regard to vote intention and hence persuasion (both in terms of likeliness to vote and support conditional on voting) are inconsistent. Some studies have found persuasive effects (Johnston, Hagen and Jamieson, 2004; Shaw, 1999; Huber and Arceneaux, 2007; Gerber et al., 2011), while others have claimed

[2]These figures are based on Kantar expenditure estimates.

[3]Of course, this is by no means empirical evidence of a non-substantive ad effect or even null or boomerang effects, but the magnitude of the spending imbalance and a popular vote within historic parameters should spur renewed interest in rigorous empirical assessments of ad effects.

that campaign advertising does not change minds (Goldstein and Ridout, 2004, for an overview). We argue that such views are deficient for a multitude of reasons.

First, due to limited sample sizes in data, prior studies have been unable to track granular swings in vote intention at the media-market level, where most ad buys occur. Second, a number of studies have grappled with concerns of treatment endogeneity. Media markets that see heavy doses of advertising usually lie in swing states, making it difficult for scholars to disentangle effects of on-the-ground campaigning, such as canvassing, from effects of political ads (Johnston, Hagen and Jamieson, 2004). Third, most studies have looked at either learning effects or changes in vote intention but have rarely considered behavioral variables (but see Aldrich et al., 2015). Fourth, previous research has not looked at effects of ads against the backdrop of high ad spending been late in the campaign, but has been concerned with the hypothetical effect of advertising exposure (i.e., what would happen if individual i was exposed to a political ad) (Ansolabehere and Iyengar, 1995; Brader, 2005) or has tracked individual-level effects of advertising on vote intention after months of exposure to political ads in the midst of campaigns, documenting that consistent partisan asynchronicity in advertising can indeed lead to individual-level opinion change (Huber and Arceneaux, 2007). Of course, no study is able to investigate the effects of advertising on actual individual-level vote, given its proprietary nature. Fifth, previous research examines non-contemporary campaigns prior to the Citizens United ruling, such as the 2000 campaign analyzed by Huber and Arceneaux (2007).

Instead, we ask a different question. Given the intense saturation of the airwaves with political ads in modern campaigns, what are the effects of the marginal advertising dollar (i.e., what does the effect of an additional dollar spent per candidate look like?) on vote intention when the campaigns are in full swing come October and November, and relatedly, what does the effect of the marginal dollar look like for high- vs. low-spending DMAs? We argue that this reconceptualization of advertising effects captures the natural environment in which political ads are encountered much more realistically than previous studies.

In contrast to other studies cited above, relying on high-frequency polling allows us to leverage the sample size necessary to analyze granular effects of advertising the day after ads are aired. Two identification strategies – comparing those incidentally exposed to political ads with those residing in the same state who are not exposed to political ads, and analyzing trends among those incidentally exposed – yield effects of ad spending uncontaminated by on-the-ground efforts and make it possible for us to isolate the causal effect of advertising on vote intention. In looking at vote intention, we assume a tight relationship between vote intention and underlying hypothetical vote choice at that time. Indeed, there is plenty of evidence that reported vote intention close to election day is strongly related to actual vote choice (Kennedy et al., 2017). Finally, we leverage access to the entire Twitter corpus (dubbed Twitter Firehose) to track behavioral effects triggered by ad spending. Using behavioral outcomes can validate the nature of our spending variable, and more importantly would document that political advertising can generate free media exposure in the form of Tweets and re-Tweets. We further document that ad spending can generate *positive* tweets per candidate as well as negative tweets mentioning the opposing candidate. To our knowledge, this study is the first attempt to comprehensively map the effectiveness of political advertising of a modern presidential campaign.

Our data is uniquely comprehensive: daily polls in media markets around the country, detailed data on all registered voters in the country, data on all TV advertisements by markets and exact time, and the entire Twitter corpus. We track vote intention, both daily and by media markets, with polls administered on the Microsoft Xbox gaming system. We couple the responses with innovative Bayesian post-stratification techniques and access to detailed information on every voter in the United States (to our knowledge also a novelty) to track overall vote intention. We determine changes in vote intention as a function of daily changes in advertising volume of the two major presidential candidates using the Kantar Media data set of all TV advertising. For behavioral effects, access to the Twitter Firehose allows us to track whether political advertising prompts tweets mentioning the candidates, or tweets mentioning the candidate in a positive light or the opposing candidate in a negative light.

Regarding vote intention, we find that neither candidate benefited from TV advertising. Neither overall increases in advertising spending nor partisan imbalances in spending persuaded voters to switch camps or expanded the candidates' electorate when the campaign was in full swing. What's more, we find a boomerang effect in high-spending DMAs: If total spending goes up, the relationship between candidates' advantages in spending and candidates' advantages in vote intention becomes more negative, as indicated by a significant and negative interaction effect of total ad spending with spending imbalance on vote intention. In contrast, we find a positive effect in looking at tweeting behavior. We use the universe of all geotaggged tweets, drawn from the Twitter Firehose, to estimate the effect of ad spending on tweeting behavior. We find that both candidates *benefited* from additional advertising spending in terms of overall tweet volume, and in terms of share of tweets with a positive tone mentioning the candidate. In the final section, we discuss the implications of the findings for the future of presidential campaigns, and the importance of closely monitoring ad spending in the new era of unhinged spending.

Hypotheses

Research on the effects of political advertising has come a long way. Many contributions have raised theoretical expectations of null effects (e.g. Klapper, 1960; Holbrook, 1996), and others have validated them empirically (e.g. Lazarsfeld, Berelson and Gaudet, 1948). A different set of contributions has raised theoretical expectations of either persuasion or mobilization effects (e.g. Lasswell, 1927), and yet again others have validated them empirically (for emotional responses see Ansolabehere and Iyengar, 1995; Brader, 2005, for mobilization see Ansolabehere and Iyengar, 1995; Hillygus, 2005, for persuasion see Huber and Arceneaux, 2007). Our approach here is not to develop a novel theoretical account of political advertising effects, but to derive empirically testable hypotheses from structural changes in the legal framework that have in turn changed the game of political advertising.

In the wake of the Citizens United ruling issued by the Supreme Court on January 21, 2010, we argue that airwaves in swing-state DMAs are uniquely and unprecedentedly saturated with political ads during presidential campaigns and become even more saturated when a campaign nears its zenith in late October. For example, during one week in July 2012, political ads accounted for 10% of *overall* programming across major networks in the Columbus, Ohio DMA (Eggen and Farnam, 2012).[4] The already extreme saturation increased from this point. In late October 2012, a local CBS affiliate in the Cleveland, Ohio DMA ran 22 (consecutive) political ads in a 30-minute time frame, meaning that as much as 33% of the entire broadcast consisted of presidential campaign advertising (Lewison, 2012). Similarly, an NBC affiliate in the Columbus market broadcast 11 ads in just under 19 minutes during a prime-time broadcast, meaning that political ads comprised almost 30% of the programming. Figure 1 graphs the average airing time of political advertising in three high-spending DMAs in 2012, empirically reinforcing the anecdotal notion of over-saturation. In the Denver, Colorado DMA, for example, political ads easily made up over 12% of *overall* programming.

[4] This assumes a 30-second length of the average spot.

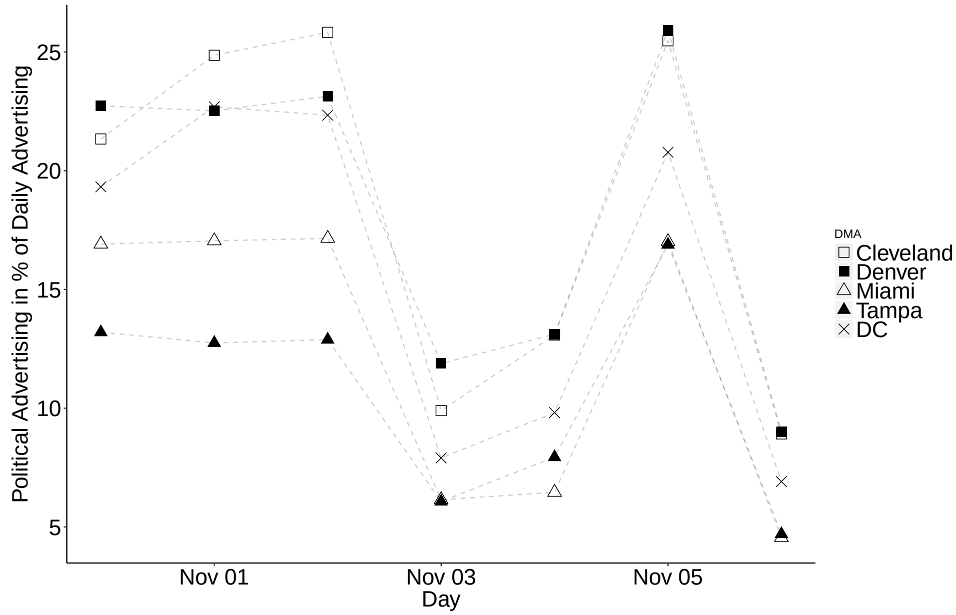

Figure 1: Airing Time of Political Ads in Three Swing State DMAs

While Huber and Arceneaux (2007) have convincingly shown that long-running asynchronicities in political advertising can lead to persuasive effects, we argue that evaluating the effectiveness of political ads against the backdrop of already intense saturation captures much better the natural environment in which political ads are usually encountered. For example, it is unclear if a marginal spot in late October will result in a positive net effect for a candidate. We argue that it is likely that viewers are overfed (read: annoyed) with political advertisements. Given the exceeding negativity of political ads and their power to demobilize (Ansolabehere and Iyengar, 1995), we hypothesize that null effects of political ads, with viewers penalizing candidates for over-saturation, are likely. While Brader (2005) has documented the potential of emotional appeals to affect political behaviors, our suggestion is that the emotional response to over-saturation is simply annoyance, mitigating the hypothetical power of the ad to mobilize or persuade under conditions of limited exposure. In the most extreme cases of over-saturation, we expect attitudinal boomerang effects to be possible, i.e. a defection from the candidate sponsoring the ad.

Regarding behavioral ramifications, our expectations are clear. Behavior as induced by political ads should be non-discriminative with regard to valence; additional ads by a particular candidate should prompt behavior such as tweeting that is either supportive of or critical of that candidate. Hence, we expect ad spending by either candidate to trigger increases in candidate-related tweet volume via a simple mechanism of priming saliency of the candidate in question. Likewise, candidate spending should not impact tweet-volume associated with the opponent candidate. Whether candidate-specific ad spending leads to increases in either *positive* tweets mentioning the sponsoring candidate or negative tweets mentioning the opposing candidate, for example due to a focus on negative ads (Ansolabehere and Iyengar, 1995), warrants investigations as well. We use this analysis first to demonstrate the conceptual validity of our advertising data and second to highlight the importance of these *behavioral* effects in the context of future political advertising.

Data and Methods

Data

The paper focuses on four data sets: daily polling in media markets around the country run on the Xbox gaming platform, detailed data on all registered voters in the country from the Target Smart voter file, all TV advertisements by markets and exact air-time from Kantar Media, and the entire Twitter corpus.

To assess aggregate change in vote intention, we use a unique data set collected on the Microsoft Xbox gaming system (Gelman et al., 2016). During the 2012 presidential campaign, we conducted 750,248 interviews with 345,858 respondents during the 45 days preceding election day and on election day itself.[5] Importantly, we also collected demographic and political variables for all respondents prior to their first response. We derive daily estimates of Obama, Romney, and "other" vote share at the DMA level, relying on methods of geo-identification via IP-addresses and Bayesian Multilevel Regression and Post-Stratification (MRP) as specified in detail in Appendix 1 (Gelman et al., 2016; Park, Gelman and Bafumi, 2004).[6] Briefly, we employ a Bayesian multi-level model to derive aggregate vote intention among 288 demographic subgroups per DMA (and later DMA-state combination) and day, and then project these estimates onto the known turnout population of 2012 derived from the TargetSmart voter files.[7] This brings us the unique advantage that we can assess advertising effects on aggregate vote intention at geographically granular levels.

We also collected Twitter data from the full Twitter Firehose. We collected all geotagged tweets that contained either *obama*, *romney*, or both for the 46-day study period. We used the given latitude and longitude to assign each geotagged tweet to a DMA. In total, there are over 494,133 tweets in DMAs for which we have

[5]See Appendix 4 for more details on the survey experience.

[6]We code those indicating to vote for Obama as Democratic vote intention, those indicating to vote for Romney as Republican vote intention, and those indicating that they are either unsure if they will vote at all or unsure as to what candidate to vote for and those indicating that they vote for a third-party candidate, as "other".

[7]See Appendix 1 for details, model specifications, and implementations in Stan.

advertising data. To gauge overall sentiment of tweets, we score sentiment by tweet using VADER (Valence Aware Dictionary and sEntiment Reasoner), a lexicon and rule-based sentiment analysis tool that is specifically attuned to sentiments expressed in social media (Hutto and Gilbert, 2014). VADER breaks down every tweet into composite scores of positive, negative and neutral sentiment. The scores sum to one and can be interpreted as "sentiment-share". We then construct the proportion positive over all tweets by day and DMA, as well as the proportion negative over all tweets by day and DMA, to get measures of overall positive and negative tweet sentiment.[8]

The advertising data is from Kantar Media and incorporates all campaign expenditures of PACs that spent more than $10,000 in total and of the campaigns themselves between September 21 and November 5, 2012. We collected advertising data for the 101 largest DMAs, which included ads worth $236,865,899 for the limited time frame.[9]

Identification strategy

Johnston, Hagen and Jamieson (2004) have introduced the idea of treating media-market (DMA) level ad buys as a natural experiment given that some key DMAs in battleground states are the target of most buys and tied cross-sectional and temporal volatility in intention to vote for Al Gore to ad spending (see Appendix 3 for a visualization of this set up for 2012). But identifying advertising effects is chronically

[8]The dictionary was able to score 404,242 tweets in total.

[9]Note that this number is much lower than the $900 Million reported in the introduction because a) we focus on a subset of DMAs, and b) because we focus on a limited time frame.

difficult due to various endogeneity concerns. Most pressingly, ad spending is highly correlated with on-the-ground campaigning efforts but also influences outcome measures such as vote intention (Huber and Arceneaux, 2007). In such a setting, the selection-on-observables identification assumption does not hold in any conventional linear modeling because treatment, in this case expenditure on political advertisement, is non-random. Huber and Arceneaux (2007) have focused on combinations of swing-state DMAs that fall into non-swing states, allowing them to disentangle general campaign effects from specific advertising effects.

In spirit close to Huber and Arceneaux (2007), we offer two models with differing rigidity regarding their identification assumptions – one for our Twitter data, for which a limited sample size prohibits us from relying on the less restrictive identification assumption, and a fully causal model for vote intention. For the twitter data, we look at within-DMA changes. Specifically, we look at the effect of increases in spending by either candidate on increases in tweeting behavior related to either candidate occurring the following day (see Appendix 3 for a visualization of this conceptualization, plotting the prior means for tweeting behavior (any tweet mentioning either candidate) for selected DMAs).[10] Our identifying assumption here rests on the belief that day-to-day spikes in spending are temporally removed and hence expected to be uncorrelated with day-to-day developments and strategies of the on-the-ground campaign. This is a) a reasonable assumption, given that advertising time is bought

[10]We have tested different lag models, specifically Auto Distributed Lag (ADL) models with different lag specifications and a Weibull function model to account for different assumptions of effect decay. Gerber et al. (2011) for example find that effects of campaign ads decay fast, while Broockman and Kalla (2016) find more durable effects, albeit in the field of political canvasing. Results of models with these different lag specifications are equivalent with the results presented here.

several weeks prior to airing, and b) a much less restrictive assumption than suggesting the *entire advertising campaign* is only exogenously related to ground-game efforts (Johnston, Hagen and Jamieson, 2004). Overall, our model is highly similar to a regular fixed-effects model, but we demean both variables and only consider the mean of the variable distribution up to the day in question to better capture the natural environment of inundation in which political ads are usually encountered.[11] To address whether spending resulted in increases of positive tweets for the candidate or negative tweets for the opposing candidate, we model the effect of increases in spending by either candidate on increases in the positive share of the overall tweet volume associated with the candidate, and on increases in the negative share of the overall tweet volume associated with the opposing candidate.

$$\begin{aligned} Tweets/Vote_{i,t} - \sum_{j=1}^{t-1} \frac{Tweets/Vote_{i,j}}{J_i} = \\ [Advertising_{i,t-1} - \sum_{j=1}^{t-2} \frac{Advertising_{i,j}}{J_i}]'\beta + \delta_t^{day} + \epsilon_{i,t,clustered} \end{aligned} \tag{1}$$

Here, i captures DMAs, t is a given day, J is the total of $t-1$ days , and δ_t^{day} are unit-common time shocks.

As Huber and Arceneaux (2007) have noted, DMA-level advertising can be correlated to ground-game efforts. While it is much more reasonable to assume that *day-to-day spikes* are uncorrelated to ground-game efforts as a function of planning

[11] We also display the regular two-way fixed-effect specification, including both time and market fixed effects, because readers may be more accustomed to this.

and coordination that applies to the former but not the latter, it is at least conceivable that the opposite holds true, i.e. ground game efforts are planned on a similar time scale as TV advertisements. Our second model for vote intention builds on a method that examines change over people who unintentionally receive the treatment (i.e., the ads), thus receiving the treatment but not the other endogenous efforts (i.e., the ground game). Specifically, we monitor *within* changes in 39 exogenous treatment regions –[12] the parts of swing-state DMAs that fall into non-swing states – we can identify and for which we have both, advertising and survey data (Huber and Arceneaux, 2007). Specifically, we model the effects of deviations from the partisan balance in ad spending on deviations from previous mean levels of vote intention the following day. Given that political advertising across campaigns is highly correlated, we cannot simply model the effect of candidate-specific ad spending because it is endogenously related to spending of the other candidate. Instead, we model the effect of deviations from the partisan balance in ad spending, the mean of the margin of candidate A over candidate B until the day in question, on shifts in Obama-over-Romney margin in vote intention occurring the next day. We hence substantively focus on the effect of the marginal dollar in ad spending. The model follows the notation given in Euqation 1, with the only difference that i now captures DMA-state combinations, instead of DMAs.[13] Here, our identifying assumption implies that day-to-day variations in advertising in exogenous treatment regions are idiosyncratic

[12]For a full list of exogenous treatment regions, see Appendix 4.

[13]We also present results for advertising effects on vote intention using the slightly more restrictive identification assumption underlying our analyses of ad effects on tweeting behavior. Results are shown in Appendix 5, Table 2, and are substantively identical to the results relying on the more causal identification assumptions presented in the main paper.

but allows correlation of day-to-day ad buys and day-to-day ground campaigning (see also Huber and Arceneaux, 2007).[14]

The alternative approach – comparing individuals in regions that are part of swing-state DMAs (i.e., exposed to the ads) and located in non-swing-states (i.e., not exposed to the ground game) with individuals directly across the DMA border in the same or contiguous non-swing-states (i.e., are not exposed to the ground game) who are not exposed to treatment (i.e., not exposed to the ads) (Shapiro, 2017) – is more difficult. We began monitoring vote intention only 45 days prior to election day, after a major part of the campaign had already been conducted, leading to endogenously different baseline conditions between swing-state DMAs located in non-swing-states and non-swing-state DMAs located in the same non-swing-state across the DMA border. But a difference-in-difference estimator allows us to test whether those exposed to ads have undergone differential trends predicted by ad spending compared to those in the same state, but residing in a non-swing-state DMA. For this analysis, we have paired every exogenous treatment region (i.e. swing-state DMA non-swing state combination) with comparable non-treatment regions (i.e. non-swing-state DMA non-swing state combination) that are closest to the original exogenous treatment region and for which we have advertising data (see Appendix 4). The difference-in-difference estimator is given in Equation 2.

[14]We tested the same lag models as specified in Footnote 10, with substantively equal results.

$$[Vote_{tr,t} - Vote_{co,t}] - [\sum_{j=1}^{t-1} \frac{Vote_{tr,j}}{J_{tr}} - \sum_{j=1}^{t-1} \frac{Vote_{co,j}}{J_{co}}] = \big([Advertising_{tr,t-1} - Advertising_{co,t-1}] - [\sum_{j=1}^{t-2} \frac{Advertising_{tr,j}}{J_{tr}} - \sum_{j=1}^{t-2} \frac{Advertising_{co,j}}{J_{co}}]\big)'\beta + \delta_t^{day} + \epsilon_{i,t,clustered} \tag{2}$$

Here, *tr* represents a treatment unit, and *co* represents a non-treatment, or control unit. The rest of notation follows this in Equation 1.

In all models, we introduce a unit-common time shock in each time period (time fixed effects) to deal with common confounds that should affect all geographical units equally. Substantively, the common time shocks capture national-level developments such as poor or strong debate performance correlated with spending efforts. In any model that does not account for time effects, a change in aggregate vote intention caused by debate performances would be artificially attributed to advertising spending, if spending on advertising and debate performances are correlated. In addition, we cluster standard errors at the unit of analysis, i.e. at the DMA level for our models on ad spending effects on tweeting behavior and at the DMA-state level for our models on ad spending effects on vote intention (see Equations 1, 2).

Results

Change in vote intention at the DMA level

As our post-stratification results reveal, candidate support at the DMA level remained largely stable over the 45-day period preceding the 2012 presidential election, with changes occurring predominantly in the North and Northeast in the Democratic direction (Figure 2) and many changes occurring in battleground areas. For example, in the Alpena, Michigan DMA, Romney's lead over Obama narrowed from 55% vs. 36% to 48% vs. 44% during the 45-day period prior to the election. Likewise, in the Flint, Michigan market, Obama's lead expanded from 47% vs. 45% to 55% vs. 37%. Outcomes for the two DMAs with the most contested airwaves are similar. In the Cleveland, Ohio market, Obama's lead expanded from 54% vs. 41% to 58% vs. 38%, and in the Miami, Florida market, Obama's lead expanded from 65% vs. 32% to 68% vs. 29%. In sum, our Bayesian post-stratification model produced realistic estimates of change in aggregate vote intention over the 46- day period.

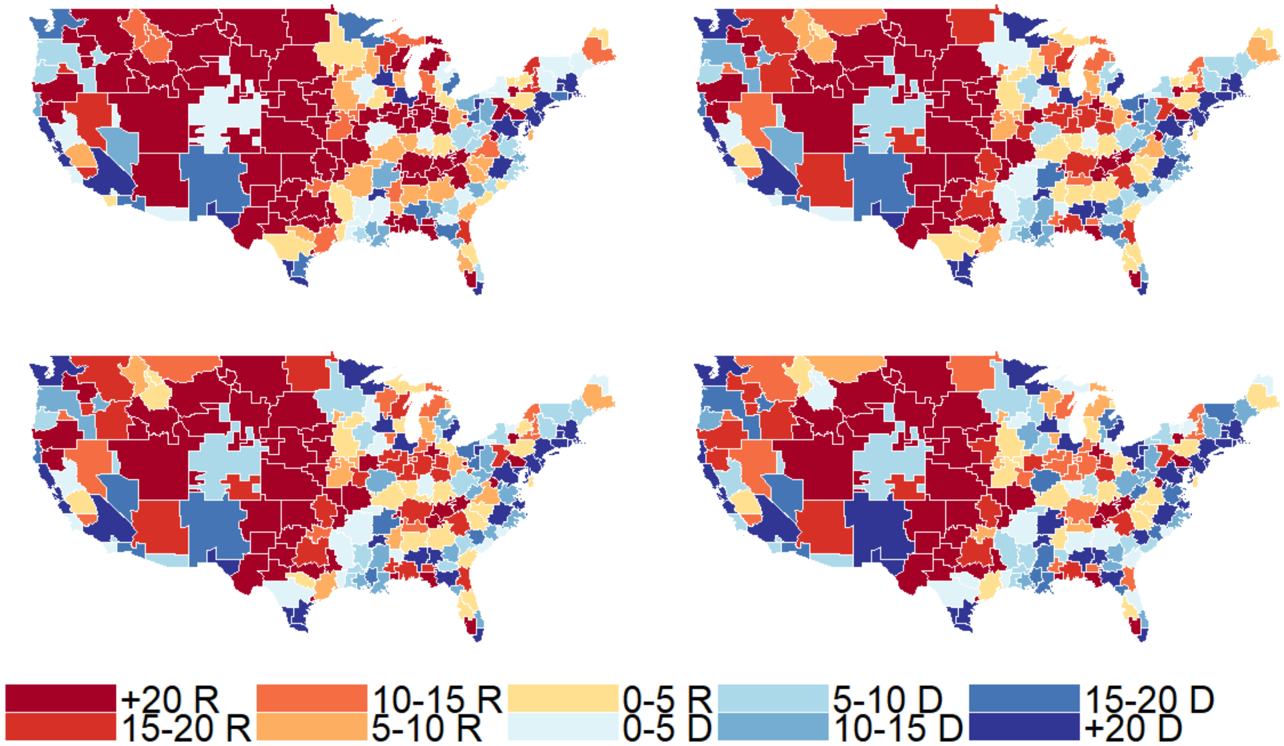

Figure 2: Dem-Rep margin in vote intention at the DMA level for September 22 (top-left), October 5, (top-right), October 22 (bottom-left), and November 6 (bottom-right) in Percentage Points

Effects of ad spending at the market level

Table 1 looks at effects of advertising on tweeting behavior. As predicted, marginal increases in advertising lead to more activity on Twitter, regardless of the candidate. Foremost, the positive association of tweeting volume and increases in ad spending validate the conceptualization of our treatment variable. Clearly, differences in advertising volume do have an effect that registers.[15]

[15]The results of our preferred model are presented in the top panel of Table 1, while the results of the more conventional fixed-effects specification are presented in the bottom panel of Table 1.

Besides validation, this is an interesting result in its own right. Our model predicts that an additional $1 million spent on ads supporting Obama will lead to an additional 707 tweets mentioning Obama and that an additional $1 million spent on ads supporting Romney ads will lead to an additional 223 tweets mentioning Romney.[16] Ad spending does also significantly register when it comes to increasing the share of positive sentiment among tweets mentioning the candidate in question, and when it comes to increasing the share of negative sentiment among tweets mentioning the opposing candidate. Our model predicts that an additional $1 million spent on ads supporting Obama will increase the share of positive sentiment among tweets mentioning the candidate by 6.6% and that an additional $1 million spent on ads supporting Romney will increase the share of positive sentiment among tweets mentioning the candidate by 7.0%. Further, our model predicts that increases in ad spending will increase the share of negative sentiment among tweets mentioning the opposing candidate. Our model predicts that an additional $1 million spent on ads supporting Obama will increase the share of negative sentiment among tweets mentioning Romney by 4.9% and that an additional $1 million spent on ads supporting Romney will increase the share of negative sentiment among tweets mentioning Obama by 4.7%.

[16] While we have estimates of the proportion of geotagged tweets in the population of all tweets, we are hesitant to extrapolate to the effect ad spending had on *all* tweets at the DMA level, given that the political nature of the tweet is endogenously related to the tweet's location in a swing state. As a result, we can confidently say that our estimates are *lower bounds* of advertising effects.

Model	Obama Tweets	Romney Tweets	Share Positive Obama Tweets	Share Positive Romney Tweets	Share Negative Tweets Oppos. Cand. (Obama)	Share Negative Tweets Oppos. Cand. (Romney)
Marginal In Candidate						
Spending in $1,000,000	707.301*	223.155*	0.066 *	0.070 *	0.047 *	0.049*
	(171.108)	(49.297)	(0.017)	(0.019)	(0.011)	(0.019)
Observations	4,545	4,545	4,545	4,545	4,545	4,545
Unit-Common Shock in Each Time Period						
Time-Common Shock in Each DMA	✓	✓	✓	✓	✓	✓
Standard Errors Clustered at the DMA Level	✓	✓	✓	✓	✓	✓
Demeaned In Candidate						
Spending in $1,000,000	722.580*	227.147*	0.066*	0.068 *	0.047 *	0.050 *
	(175.35)	(49.633)	(0.016)	(0.018)	(0.011)	(0.018)
Observations	4,646	4,646	4,646	4,646	4,646	4,646
Unit-Common Shock in Each Time Period	✓	✓	✓	✓	✓	✓
Time-Common Shock in Each DMA	✓	✓	✓	✓	✓	✓
Standard Errors Clustered at the DMA Level	✓	✓	✓	✓	✓	✓

Table 1: Effect of Advertising on Tweeting Behavior and Share Positive Tweets

Next, we look at vote intention in exogenous treatment regions that are part of swing-state DMAs that roll over into non-swing states but are not themselves in a swing state. Shifts in the advertising balance in favor of either candidate did provide a discernible boost to either candidate. Here, we are confident in causally attributing effects on our outcome variables to differences in advertising spending. To be more precise, shifts in the advertising spending balance in favor of Obama by $1 million would result in a 1.5% *decrease* in aggregated intention to vote for Obama within a given exogenous treatment region while for Romney, a shift in the advertising balance in his favor by an additional \$1 million would result in a 1% *decrease* in intention to vote for Romney within a given exogenous treatment region.[17]

Model	Obama Vote Intention	Romney Vote Intention
Marginal Candidate Advantage in Spending in \$1,000,000	-1.494	-1.055
	(1.017)	(1.000)
Observations	1,755	1,755
Unit-Common Shock in Each Time Period		
Time-Common Shock in Each DMA-State	✓	✓
Standard Errors Clustered at the DMA-State Level	✓	✓
Demeaned Candidate Advantage in Spending in \$1,000,000	-1.720	-1.302
	(1.028)	(0.995)
Observations	1,794	1,794
Unit-Common Shock in Each Time Period	✓	✓
Time-Common Shock in Each DMA-State	✓	✓
Standard Errors Clustered at the DMA-State Level	✓	✓

Table 2: Effect of Advertising on vote intention in Exogeneous Treatment Regions

These estimates are too imprecise to provide evidence of a negative relationship between marginal spending and vote intention, but serve as preliminary evidence

[17]Again, the results of our preferred model are presented in the top panel of Table 2, while the results of the more conventional fixed-effects specification are presented in the bottom panel of Table 2.

against persuasive effects of advertising. To shed further light on the differentiation of null effects and imprecise effects, we look at heterogeneous effects across high- and low-spending exogenous treatment regions. Further, if our hypothesis holds and over-saturation drives down the value of the marginal dollar, we should see a more pronounced boomerang effect in high-spending exogenous treatment regions. Indeed, if we interact the divergence in spending with the *overall* amount of advertising dollars spent in that DMA in which the exogenous treatment region in question falls, we find the expected negative interaction effect.[18] Specifically, for every $ 1 Million spent more in total, shifts in the advertising spending balance in favor of Obama by $1 million would result in a 1.4% *additional decrease* in aggregated intention to vote for Obama within a given exogenous treatment region. For Romney, for every $1 Million spent more in total, shifts in the advertising balance in his favor by an additional $1 million would result in a 1.3% *additional decrease* in intention to vote for Romney within a given exogenous treatment region.

[18]Given that total spending does not vary over time, we can only display our preferred specification, but not the two-way fixed-effect specification.

Model	Obama Vote Intention	Romney Vote Intention
Marginal Candidate Advantage		
in Spending in $1,000,000	3.305*	2.743
	(1.622)	(1.562)
Total Spending		
at DMA in $1,000,000	-0.000	0.001
	(0.003)	(0.004)
Total spending x		
Marginal Candidate Advantage		
in Spending in $1,000,000	-0.141*	-0.126*
	(0.006)	(0.057)
Observations	1,755	1,755
Unit-Common Shock in Each Time Period		
Time-Common Shock in Each DMA-State	✓	✓
Standard Errors Clustered at the DMA-State Level	✓	✓

Table 3: Effect of Marginal spending by overall Dollars spent

We replicate these findings relying on our alternative difference-in-difference specification comparing trends of those in exogenous treatment region and those residing across the DMA border in a non-swing-state DMA, and relying on DMA-level models slightly more restrictive in their identification assumptions. The full models are given in Appendix 5.

In addition, we offer a number of robustness checks. First, we show that general (i.e. two-party) spending on political advertising in the last 45 days before the election also does little to increase overall turnout by looking at the effect of overall spending on the aggregate percentage of those who are not sure what candidate to vote for, or indicate to not vote for either major-party candidate. We find no evidence of a negative relationship, indicating that general two-party ad spending in the last 45 days before the election does not increase two-party turnout.

Second, we look at demographic subgroups that are theorized to be the most persuadable: the median educated – people without party affiliation (i.e. independents) with moderate levels of education (Zaller, 1992). Finally, we look at the effects of changes in overall spending. Differences in the partisan spending balance may not capture differences in vote intention because of generalized advertising effects. In essence, it may be that increases in spending, *regardless* of the sponsoring candidate, hurt one candidate significantly while helping or having no effect on aggregated intention to vote for the other candidate. These three models confirm our main results presented here, and are presented in full in Appendix 5.

Conclusion and Discussion

In this study, we begin outlining an empirical framework by which advertising effects can be assessed, taking into account the reality in which political ads are encountered, using non-probability surveys, big data, and real-time data on tweeting and advertising behavior. In contrast to earlier studies, we evaluate the impact of ads against the natural environment in which ads are usually encountered on a day-by-day basis. In line with our expectations, we find that more advertising is not helpful for candidates' standing. Specifically, we suggest that increases in total candidate spending do not benefit either candidate. What's more, the overall spending-level of DMAs significantly moderates a negative relationship between candidates' spending advantages and candidates' vote intention, lending evidence to our hypothesis of boomerang effects induced by over-saturation. In addition, we find no discernible effect of a stark partisan imbalance in spending on aggregated vote intention. We confirm these analyses by looking at so-called median voters – voters who we believe are highly persuadable because of their median level of education combined with their independent affiliation.

Of course, it is entirely possible that political ad campaigns are more valuable in their early stages, for example when they help voters form opinions on candidates. For example, negative ads aired early in the campaign cycle might shrink candidates' electorates (Ansolabehere and Iyengar, 1995). However, this possibility does not change the validity of our finding that ads aired late in the campaign-cycle do little to expand candidates' electorates.

When validating our ad spending data against tweeting behavior, we find that both candidates do indeed get a boost in attention on Twitter with increased ad spending. First, we show that ad spending can trigger meaningful behavioral outcomes (see also Aldrich et al., 2015). We can also put a lower bound dollar value on the additional tweet. A marginal increase in ad spending related to the Democratic campaign by $1 million resulted in *at least* an additional 707 tweets mentioning Obama, and an additional $1 million spent on political ads by the Romney campaign resulted in *at least* an additional 223 tweets mentioning Romney. Second, candidate-specific spending increases the share of positive tweets for both candidates, and increases the share of negative tweets for the opposing candidate in both cases.

We highlight that due to the unique setup of our data – a rich survey of more than half a million respondents over the course of 46 days combined with high-end analytics to correct for non-response bias and access to proprietary data to construct an appropriate post-stratification space – our study is the first to assess the aggregate effect of political advertising at a granular geographical level. While important in its own right, the abundant sample size helps us to reconceptualize the treatment variable of ad spending. Against the backdrop of already intense saturation, marginal increases in ad spending do very little to sway voters one way or another. In high-spending DMAs, marginal spending increases by Romney lead to a boomerang effect, significantly decreasing aggregate Republican vote intention, and the same holds true for marginal spending increases by Obama.

With regard to the 2016 U.S. presidential election, we suspect that the results presented here indicate that there is still some value to political advertising. First, as we have shown, political advertising can generate free media exposure, for example, via tweets. Such free media exposure, especially if mediated by a social network (Aldrich et al., 2015), can yield powerful indirect persuasive effects, which we are unable to test here. Future research would do well in looking at mediation models on a two-step flow of communication – from advertising to social media to public opinion – especially in light of the 2016 elections. While many of the data sources used here are proprietary, the spirit of our analysis can be replicated. The kind of high-frequency non-probability polling we use here is much cheaper compared to more traditional polls, and aggregate information on the breakdown of the American public is easily available, for example via the Census 5-year estimates.[19] In addition, the power of ads likely depends on the make-up of individual clips. More micro-level studies are needed to elucidate what kinds of ads can produce what kinds of emotional responses. But when it comes to direct advertising effects, it is reasonable to conclude that the value of the marginal dollar spent on political advertising, in the heat of the campaign, is negligible at best.

References

Aldrich, John H., Rachel K. Gibson, Marta Cantijoch and Tobias B. Konitzer. 2015. "Getting Out the Vote in the Social Media Era: Are Digital Tools Changing the Extent, Nature and Impact of Party Contacting in Elections?" *Party Politics* 22(2):165–178.

Ansolabehere, Stephen and Shanto Iyengar. 1995. *Going Negative: How Political Ads Shrink and Polarize the Electorate*. New York: Free Press.

Ballard, Andrew O., Sunshine D. Hillygus and Tobias B. Konitzer. 2016. "Campaigning Online: Web Display Ads in the 2012 Presidential Campaign." *PS: Political Science and Politics* 49:627–631.

Brader, Ted. 2005. "Striking a Responsive Chord: How Political Ads Motivate and Persuade Voters by Appealing to Emotions." *American Journal of Political Science* 49(2):388–405.

[19] See `https://www.census.gov/programs-surveys/acs/`.

Bradley, Tom. 2010. "Google Ads Level the Playing Field.".
URL: *http://www.pcworld.com/article/188783/Google-Ads-Level-the-laying-Field.html/*

Broockman, David and Joshua Kalla. 2016. "Durably reducing transphobia: A field experiment on door-to-door canvassing." *Science* 352(6282):220–224.

Eggen, Dan and T.W. Farnam. 2012. "In 2012 campaign, Obama and Romney inundate swing states with ads.".
URL: *https://www.washingtonpost.com/politics/in-2012-campaign-obama-and-romney-inundate-swing-states-with-ads/2012/07/28/gJQACnHzGX-story.html*

Gelman, Andrew, Sharad Goel, Douglas Rivers and David Rothschild. 2016. "The Mythical Swing Voter." *Quarterly Journal of Political Science* 1(1):103–130.

Gerber, Alan S., James G. Gimpel, Donald P. Green and Daron R. Shaw. 2011. "How Large and Long-lasting Are the Persuasive Effects of Televised Campaign Ads? Results from a Randomized Field Experiment." *American Political Science Review* 105:135–150.

Goldstein, Kenneth and Travis N. Ridout. 2004. "Measuring the Effects of Televised Political Advertising in the United States." *Annual Review of Political Science* 7(1):205–226.

Hillygus, D. Sunshine. 2005. "Campaign Effects and the Dynamics of Turnout Intention in Election 2000." *Journal of Politics* 67(1):50–68.

Holbrook, Thomas M. 1996. *Do Campaigns Matter?* Thousand Oaks: Sage Publications.

Huber, Gregory A. and Kevin Arceneaux. 2007. "Identifying the Persuasive Effects of Presidential Advertising." *American Journal of Political Science* 51(4):957–977.

Hutto, Clayton J and Eric Gilbert. 2014. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Eighth international AAAI conference on weblogs and social media*.

Johnston, Richard, Michael J. Hagen and Kathleen H. Jamieson. 2004. *The 2000 Presidential Election and the Foundations of Party Politics*. New York: Cambridge University Press.

Kantar. 2016. "Election 2016: A Campaign Odyssey.".
URL: *http://www.kantarmedia.com/us/thinking-and-resources/blog/election-2016-a-campaign-odyssey*

Kennedy, Courtney, Mark Blumenthal, Scott Clement, Joshua D. Clinton, Claire Durand, Charles Franklin, Kyley McGeeney, Lee Miringoff, Kristen Olson, Doug Rivers, Lydia Saad, Evans Witt and Chris Wlezie. 2017. "An Evaluation of 2016 Election Polls in the U.S.".
URL: *http://www.aapor.org/Education-Resources/Reports/An-Evaluation-of-2016-Election-Polls-in-the-U-S.aspx*

Klapper, Joseph T. 1960. *The effects of mass communication*. Free Press of Glencoe.

Kurtzleben, Danielle. 2015. “2016 Campaigns Will Spend $4.4 Billion On TV Ads, But Why?”.
URL: *http://www.npr.org/sections/itsallpolitics/2015/08/19/432759311/2016-campaign-tv-ad-spending*

Lasswell, Harold D. 1927. “The Theory of Political Propaganda.” *American Political Science Review* 21:627–631.

Lazarsfeld, Paul Felix, Bernard Berelson and Hazel Gaudet. 1948. *The people’s choice: how the voter makes up his mind in a presidential campaign.* Columbia University Press.

Lewison, Jed. 2012. “22 consecutive political ads during half-hour Columbus, Ohio, newscast.”.
URL: *http://www.dailykos.com/story/2012/11/01/1153753/-22-consecutive-political-ads-during-half-hour-Columbus-Ohio-newscast*

Park, David K, Andrew Gelman and Joseph Bafumi. 2004. “Bayesian multilevel estimation with poststratification: state-level estimates from national polls.” *Political Analysis* 12(4):375–385.

Shapiro, Bradley T. 2017. “Positive Spillovers and Free Riding in Advertising of Prescription Pharmaceuticals: The Case of Antidepressants.” *Working Paper* .

Shaw, Daron R. 1999. “A Study of Presidential Campaign Event Effects from 1952 to 1992.” *The Journal of Politics* 61:387–422.

Sullivan-Jeniks, Megan. 2016. "What the Election Taught Us About Digital Advertising?". Online; accessed 2016-11-30.
URL: *http://blog.choozle.com/what-the-election-taught-us-about-digital-advertising/*

Vogel, Kenneth P. and Tarini Parti. 2013. "Obama, Romney both topped $1B.". Online; accessed 2015-09-01.
URL: *http://www.politico.com/story/2012/12/barack-obama-mitt-romney-both-topped-1-billion-in-2012-084737/*

Wilbur, Kenneth. 2015. Recent Developments in Mass Media: Digitization and Multitasking. In *Handbook of Media Economics, vol. 1.*, ed. S. Anderson, D. Stromberg and J. Waldfogel. North-Holland pp. 207–224.

Zaller, John R. 1992. *The Nature and Origins of Mass Opinion.* Cambridge University Press.