Do Online Advertisements Increase Political Candidates' Name Recognition or Favorability? Evidence from Randomized Field Experiments

David E. Broockman · Donald P. Green

Abstract Internet advertisements are an increasingly common form of mass communication and present fresh opportunities for understanding enduring questions about political persuasion. However, the effects of online ads on electoral choice have received little scholarly attention. We develop a new field experimental approach for assessing the effects of online advertisements and deploy it in two studies. In each study, candidates for legislative office targeted randomly selected segments of their constituencies for a high volume of Facebook advertising. Recall of the ads, candidate name recognition, and candidate evaluations were measured with ostensibly unrelated telephone surveys after weeklong advertising campaigns. Voters randomly exposed to the ads were in some cases more likely to recall them but no more likely to recognize or positively evaluate the candidates they depicted. From a theoretical standpoint, these findings suggest that even frequent exposure to advertising messages may be insufficient to impart new information or change attitudes.

Keywords Field experiments · Mass communication · Persuasion · Campaigns and elections

D. E. Broockman
Department of Political Science, University of California, Berkeley, 210 Barrows Hall, Berkeley, CA 94720, USA
e-mail: broockman@berkeley.edu

D. P. Green (✉)
Department of Political Science, Columbia University, 7th Floor International Affairs Bldg., New York, NY 10027, USA
e-mail: dpg2110@columbia.edu

Internet advertising is rapidly becoming a medium of choice for governments, political parties, corporations, activists, and others to win support, sell products, budge stubborn prejudices, and otherwise shape the public's perceptions, beliefs, and behavior. Nearly $40 billion was spent on online advertising in the United States in 2012, surpassing the amount spent on once-supreme print advertisements (eMarketer 2012a, b); meanwhile, the audience for online ads has become enormous—about 85 % of Americans use the internet (Pew Internet and American Life Project 2012), and currently one-third of the US adult population logs into Facebook alone at least once per day (Public Religion Research Institute 2012).

In tandem with this broader sea change in mass communication, political campaigns' advertising efforts have also increasingly focused on online media— indeed, the 2012 Obama and Romney campaigns appear to have spent roughly 25 % of their advertising dollars on Internet ads, or around 10–15 % of their overall budgets (Kaye 2012; see also Kaid 2012 for review). Yet in comparison to the extensive literatures investigating the political influence of traditional media such as newspaper coverage (e.g., Mondak 1995; Ladd and Lenz 2009) and television broadcasts (e.g., Iyengar 1991; Gerber et al. 2011a), the research literature on internet advertising remains sparse.

In addition to their growing political significance, online advertisements also present scholars fresh opportunities to investigate enduring questions about the effects of mass communication. Internet advertisements can be deployed with great frequency, nearly guaranteeing that individuals who visit specific sites are exposed; likewise, the ads often present information (such as a candidate's name) that one would not expect motivated reasoners (Taber and Lodge 2013) to reject. These conditions would seem to be ideal for persuasion under theories of attitude change that emphasize the impact of repeated exposure to even subtle messages (e.g., Zajnoc 1968; Atkin and Heald 1976; Bargh et al. 1992; Lodge et al. 1995; Grimmer et al. 2012; Kam and Zechmeister 2013), and the special potency of messages that are not at odds with prior beliefs or values (e.g., Zaller 1996; Taber and Lodge 2013; Ladd and Lenz 2009).[1]

On the other hand, for mass messages to bring about enduring attitude change exposed individuals may need to desire retaining the information the messages contain (e.g., Petty and Cacioppo 1986). Consistent with studies suggesting that individuals typically forget televised messages within a matter of days or hours (e.g., Gerber et al. 2011a; Hill et al. 2012; Patterson and McClure 1977; Sears and Kosterman 1994), we might expect online advertisements to leave at most a fleeting impression on most viewers. Online advertisements present a unique empirical opportunity to distinguish between these hypotheses as they nearly guarantee that individuals will be repeatedly exposed to acceptable messages, with only individuals' motivation to retain those messages remaining as a barrier.

Notwithstanding the significance of Internet ads for political practice and theories of mass persuasion, their effects have rarely been assessed systematically. Empirical

---

[1] In the context of political campaigns, candidates for office have long been hypothesized to rely on brief and relatively uninformative advertising to build name recognition and support (e.g., Stokes 1963; Stokes and Miller 1966).

research on online advertising in other disciplines has largely neglected the study of ads that seek to change minds or affect 'offline' behaviors; most existing research focuses on immediate online purchases, click behavior, installation of software applications, and online charitable giving (Goldfarb and Tucker 2011; Bakshy et al. 2012; Lewis et al. 2012; Ryan 2012; Aral and Walker 2010; Lacetera et al. 2012). To the extent experimental research has attempted to identify the effect of internet advertising on the public's perceptions, beliefs, evaluations, or 'offline' behavior, subjects are typically aware that they are being studied, having been previously enrolled in a research study or asked to browse the internet while sitting in a researcher's lab (e.g., Danaher and Mullarkey 2003; Buscher et al. 2009; Grimmer et al. 2012). One of the few unobtrusive studies finds that online ads boost in-store purchases among pre-existing customers (Lewis and Reiley 2012),[2] and the sole study to examine offline political outcomes finds that Facebook's reminders to vote nudge turnout upward, but only when accompanied by pictures of friends (Bond et al. 2012).

In sum, despite their substantive and theoretical import, the effects of online advertisements on the public's political perceptions, beliefs, and evaluations remain a largely open question. In this article, we develop an experimental research strategy that uses clustered random assignment to gauge the effects of online advertising in real-world settings. We demonstrate the practical advantages of this method in the context of two political campaigns leading up to the November 2012 election. In the first study, a little-known Republican candidate for state legislative office conducted a week-long Facebook advertising campaign one month before the election. In the second study, a viable Democratic candidate for Congress purchased a week's worth of Facebook ads one week before the election.

The article is organized as follows. We begin by describing the experimental protocol we developed for evaluating the effects of online ads on political perceptions, beliefs, and evaluations. Next, we discuss the political settings in which the experiments took place, the nature and frequency of the experimental ads, and our outcome measures. Statistical results from both studies suggest that the online ads had little effect on their viewers' recognition or evaluation of the advertising candidate. We conclude by discussing the implications of the findings and suggesting avenues for future research.

Experimental Design

One of this study's contributions is the development of a feasible method for studying the effects of online advertising on the public's beliefs, perceptions, evaluations, and offline behavior. Because the identities of Internet users are typically proprietary information, to date it has proven difficult for scholars and advertisers to rigorously investigate the impacts of online advertising.

---

[2] Bailard (2012) randomizes access to the Internet itself in Tanzania, which addresses a causal question that goes beyond the effects of specific Internet messages. See also Nickerson (2007) regarding the negligible effects of email reminders to vote and the growing research literature on how individuals use (and fail to use) the Internet for political purposes (e.g., Carlisle and Patton 2013).

Our method takes advantage of the fact that online advertising platforms typically permit advertisements to be targeted to individuals on the basis of pre-defined and mutually exclusive attributes. For example, Facebook, the platform we employ in the present studies, allows advertisements to be targeted on the basis of users' age, gender, and location; one could thus instruct Facebook to deliver a given ad only to 24-year-old males residing in San Francisco. As Ryan (2012) has shown, such demographic targeting permits researchers to conduct cluster randomized experiments on the platform, with each cluster referring to individuals who share the same age (or range of ages), gender, and location.

We build on Ryan (2012)'s approach by noting that these (and other) demographic characteristics targetable on websites like Facebook are also present in voter files, campaign finance reports, and a number of other publicly available registers of individuals. This demographic information forms a bridge between the targeting of online advertising and various public lists. When these data sources are merged, researchers can use the demographic information in these public records to ascertain which individuals were in the treatment and control groups of an online advertising campaign.

For example, if the cluster "24 year old males in San Francisco" were randomly assigned to be targeted for ads on Facebook, one would know that 24 year old males in the San Francisco voter file would be exposed to the ads if they used the website.[3] Subsequent telephone interviews with individuals residing in San Francisco could then be used to measure experimental outcomes. To estimate the effect of treatment assignment, the researcher need only compare individuals in clusters randomly assigned to be displayed the ads (e.g., 24 year old males in San Francisco) to individuals in clusters shown no ads (e.g., 24 year old males in Palo Alto, 25 year old females in San Francisco, etc.). In order to avoid priming subjects to draw the link between online ads and candidate evaluations, the telephone survey should ask questions regarding the key dependent variables (e.g., candidate name recall and vote choice) prior to any items about the use of online media or recall of the advertisements.

Previous attempts to assess the 'offline' effects of online advertisements have grown out of collaborations between scholars and online advertising firms that can target online advertisements at the individual level and match these users to individual-level outcomes such as voter turnout (Bond et al. 2012; Lewis and Reiley 2012). Unfortunately, few researchers have the opportunity to collaborate with firms like Facebook, and even online advertisers themselves often have difficulty identifying their users' offline identities. However, clustered assignment using demographic groupings does not require individually identifying internet users. As a result, any researcher with a modest budget may use the method we describe to conduct large-N randomized field experiments on the effects of online advertisements.[4] So long as researchers block on cluster size when randomly allocating

---

[3] There is reason to think that Facebook users provide accurate age information, as they enter their birth date when they first give their personal information to the site upon signing up and are offered the opportunity to hide this information from other users.

[4] Cluster random assignment has also rendered many other causal questions tractable: for example, many experiments on education randomize at the level of classrooms, not pupils; likewise, television and radio ads are typically randomized at the level of media markets, not individual devices (e.g., Panagopoulos and Green 2008).

subjects to treatment and control (Middleton and Aronow 2011), clustered assignment imposes no additional assumptions beyond those invoked by experiments randomized at the individual level; the main complication associated with clustered designs is proper estimation of sampling variability (Arceneaux 2005; Wooldridge 2003), an issue that we address below.

The Present Studies

We next describe two studies in which we deployed this method to gauge the effectiveness of political candidates' online advertisements on the popular website Facebook. In each of these studies, the campaigns we collaborated with initially supplied us with a list of voters in their districts. We then generated clusters of individuals with unique combinations of age,[5] gender, and location; e.g., "24 year old males in San Francisco." After selecting treatment clusters at random, we deployed each candidate's ads on Facebook (see below), targeting only these randomly assigned clusters of individuals. Voters in clusters that were assigned to the treatment group thus saw ads for the candidates all week if and when they logged on to Facebook, while the campaigns showed no advertisements to voters in the randomly selected control group.[6] Finally, after delivering these advertisements on Facebook over the span of a week, we conducted polls of registered voters in the candidates' constituencies that included questions about whether subjects knew the candidates' names, had favorable impressions of them, and recalled seeing material on the internet about the candidates. The survey also measured respondents' Facebook usage.

Comparing the responses given by individuals who were randomly assigned to be exposed to the ads to those who were randomly assigned to the control group allows us to identify the average causal effect of assignment to ad exposure on the candidate's name recognition, or the "intent-to-treat" effect. Since exposure to the experimental ads was nearly universal among those in the assigned treatment group who visited Facebook, examining differences among treatment and control users who report using Facebook provides an estimate of the average "treatment-on-treated" effect[7]—that is, the effect of exposure to the ads among those who were exposed (because they visited Facebook and because all Facebook visitors in the

---

[5] Age can be targeted on the basis of single-year categories (e.g, 30, 31, 32, 33) or multi-year ranges (e.g., 30–31, 32–33). In these studies age was calculated on the basis of the date of birth available in the voter file, with the age used for the study being the voter's age on the first day the ads were run.

[6] The only opportunity for spillover between the groups within the site itself is a notification that a friend in the treatment group has just "Liked" the candidate's Facebook page. Over the course of both studies fewer than a dozen new individuals "Liked" the candidate's pages during the study period, suggesting that the potential for such spillover is minimal.

[7] We assume that the ads had no effect on respondents' tendency to use Facebook itself or report using it.

treatment group were exposed)[8] or would have been exposed had they been randomized into the treatment group (because they visited Facebook and would have been exposed had their cluster been randomized into the treatment group).[9]

Study 1: Republican State Legislative Candidate in a Non-battleground State

In the first study, a candidate for state legislature deployed advertisements to randomly selected segments of his constituency. The collaborating candidate was a Republican running for state legislature. The candidate's opponent was a longstanding Democratic incumbent who was running for re-election in a newly drawn district with a partisan composition that leaned Republican, giving the challenger a reasonable chance to win the seat. Both candidates were white males. The district is predominately white and rural.

We expected several aspects of the experimental setting and treatments to be conducive to uncovering effects of online advertising. First, the ads were deployed through the website Facebook, the second most visited website in the United States (Fitzgerald 2012); according to Facebook's records, nearly 15,000 individuals were exposed to the advertisements in this study. Next, given the inexpensiveness of Facebook advertising and the frequency with which Facebook users visit the site (Hampton et al. 2011), the campaign could expose Facebook users to the ads at remarkable volume; treated voters were typically exposed to the ads many dozens of times over the course of the week (the maximum volume of advertising the platform could deliver). Last, the campaign context would seem to facilitate strong advertising effects. The candidate was running for office in a district in which a large segment of the electorate shared his party identification and would therefore be receptive to his advertising appeals. Moreover, because the candidate was relatively unknown at the time the ads ran, many politically interested voters could have learned about his candidacy for the first time from his ads.

Ad Treatments

The campaigns' Facebook ads appeared on the right side of computer users' screens on all pages on the site and were 125 pixels high by 255 pixels wide, as is standard

---

[8] About 1 in 5 individuals who use Facebook use it only on their mobile phones (Van Grove 2012). Because the mobile phone application does not show the kind of ads we purchased in the first study (although does include the 'news feed' ads in the second study), it is possible that some Facebook users in the treatment group were not exposed to the ads. If we incorrectly assumed exposure (or the potential for exposure) where none occurred (or would have occurred), our estimate of the treatment-on-treated effect would be biased toward zero, although this bias would be small.

[9] Note that this approach for estimating the treatment-on-treated effect differs from analyses of one-sided noncompliance in voter mobilization studies, which have no way to ascertain who would have been treated in the control group (e.g., which individuals in the control group would have answered their doors had a canvass contacting been attempted). By restricting both the treatment and the control group to the individuals who used Facebook in the prior week, we are able to compare those who were treated to those who would have been treated had they been randomly assigned to treatment. For a discussion of this 'placebo-controlled' design, see Nickerson (2005) and Gerber et al. (2010).

for such ads. When clicked, all ads brought individuals to the candidate's Facebook 'page' (though as is typical for online advertisements, click rates were well below 0.1 %).[10]

The first ad merely sought to build the candidate's name recognition and identify him as a proud resident of the area:

[Name of candidate] for [Name of office]
My family is one of few younger families to move to [region]. Find out why!
[Picture of candidate with his family]

A second ad included a more explicit character appeal stressing the candidate's business experience and military service:

[Name of candidate] for [Name of office]
I spent 12 years in [branch of the military] and grew a [region] small business. Connect with me today!
[Picture of candidate smiling and holding his campaign sign]

Finally, a third appeal sought to appeal to voters on a salient policy issue by stressing the candidate's desire to improve farming in the state:

[Name of candidate] for [Name of office]
Farming is crucial to [state]'s economy. [Candidate's first name]'s 4 WAYS to improve farming in [state] today!
[Picture of candidate giving a speech to a small crowd]

The candidate's constituency includes a large number of people connected to the farming industry, and thus the candidate expected this to be a particularly salient issue.

Random Assignment Procedure

We implemented the clustered random assignment procedure described in the previous section. First, we received a copy of the public list of voters from the campaign. At the campaign's request, we removed individuals under the age of 30 and over the age of 75 from the study, leaving 32,029 voters who were then assigned to 1,220 clusters across 18 age ranges,[11] the 34 towns in the candidates' district, and 2 genders.

We then blocked clusters into 244 groups of five based on cluster size[12] (which ranged from one person to five hundred), then age range, then town, and finally gender. Within these blocks of five clusters, we assigned two clusters to the control condition and one each to our three treatments: the name recognition appeal, the

---

[10] The ads garnered a total of about 150 clicks for a total click rate of about 0.02 % per impression.

[11] The ranges were 30–31, 32–33, 34–35,…62–63, 64, and 65 and above.

[12] Blocking on cluster size holds constant the ratio of treatment to control subjects. When this ratio is allowed to vary, clustered assignment may cause difference-in-means comparison to be biased. See Gerber and Green (2012, p. 84).

character appeal, and the policy appeal. (For further explanation of the mechanics of this blocked-clustered design, see Appendix 1.)

Treatment Delivery

We uploaded these ads to Facebook on Saturday, October 6, 2012 and the site approved them for delivery to individuals in the treatment clusters shortly thereafter. The ads were served on Facebook beginning on Monday, October 8, 2012 and ending Friday, October 12, 2012.

Online ads are purchased through 'mini-auctions' among potential advertisers. The campaign placed an extremely high bid for each ad, $1.51 per thousand impressions.[13] Although a bid of only 0.151 cents per impression may seem like a pittance, the market for targeted Facebook advertisements actually typically clears at prices well below $0.30 per thousand impressions, or 0.030 cents per impression ($0.0003 per impression). We selected this unusually high bid so as to be sure that we displayed as many ads as possible. According to Facebook's accounting records, although the campaign contracted with Facebook to spend up to $150 per day delivering impressions at up to this plum price, the platform was only able to deliver about $40 per day in advertising due to the finite supply of Facebook users from the targeted constituency. The campaign therefore ran as many Facebook advertisements to the treatment group as was possible.

During the course of the week-long advertising campaign, the Facebook ad interface reported that essentially every single person who could have seen the ads on Facebook did indeed see them (in Facebook parlance, the number of 'targeted' individuals was identical to the number of individuals 'reached')—5,012 users in the family treatment, 4,752 users in the character treatment, and 4,970 in the policy treatment, or 14,734 Facebook users in all. 19,377 voters on the voter file were assigned to these clusters. Facebook's records suggest that over the course of the week the typical targeted person saw the ads about three dozen times.

Outcome Measurement

To assess the impact of the ads, on Saturday, October 13 through Monday, October 15 the polling firm AMM Political Strategies completed live interviews with 2,984 individuals on the voter file (all of whom had associated phone numbers). The firm called the numbers in random order and did not have access to the treatment assignment status of the respondents. At our request, the campaign ceased all advertising on Facebook during this 3 day polling period.

The questionnaire, given in Appendix 2, asked respondents (1) whether they had a positive, negative, or no impression of the collaborating candidate, (2) whether they had a positive, negative, or no impression of the opposing candidate, (3) their vote intention in the upcoming election, (4) whether they recalled seeing any ads for the candidate on the internet, and (5) how often they used Facebook over the last

---

[13] Facebook sells its ads using second price auctions. As a result, the campaign rarely paid its stated bid, instead paying the price offered by the second bidder.

Variable	Statistic
Assignment to ad treatment	60 %
Heard of candidate	15 %
Have heard of candidate and have positive impression	11 %
Plan to vote for candidate	19 %
Self-reported Facebook use	46 %
Male	45 %
Republican	46 %
Democrat	29 %
Age (mean)	57
N (sample size)	2,984

Table 1 Descriptive statistics for survey respondents in study 1

week. As mentioned earlier, this question ordering was crucial to the credibility of the estimates: questions (4) and (5) were asked after the main dependent variables of interest so as to avoid priming respondents with the ads' content or tipping off treated users to the connection between the poll and the online advertisements.

Results: Descriptive Statistics

Descriptive statistics for the 2,984 voters who completed the poll are shown in Table 1. Recall that the original random assignment placed 60 % of the subjects in the treatment group. Most importantly, 60 % of the voters who completed the survey had been assigned to the treatment group and 40 % to the control group; there are no signs of differential attrition from the treatment group.[14] Although not necessary for unbiased estimation of the treatment effect within the sample, the fact that the overall partisan composition of the sample mirrors that of the district is encouraging for the generalizability of the results. Of the voters on the file furnished to us by the campaign, 29 % were registered Democrats and 44 % were registered Republicans; the sample is nearly identical, with 29 and 46 % of voters being registered with the Democratic or Republican parties, respectively. Other statistics describing the sample appear in the rows below. Most importantly for present theoretical purposes, the ads would be expected to be able to increase the

---

[14] Moreover, covariates remain balanced in the sample to the same degree as would be expected on the basis of chance. (If attrition in treatment and control groups were caused by the same factors, we would expect to see no deterioration in covariate balance between treatment and control groups among subjects who completed the survey.) In order to test the null hypothesis of covariate balance using randomization inference, we first generated 10,000 permutations of treatment assignment under the study's blocked and clustered randomization scheme. We then regressed (using OLS) each potential treatment assignment vector on gender, age, party, Facebook use, and turnout in the 2012 presidential primary. The F statistics from each of these regressions yields a distribution of covariate balance statistics under the null hypothesis of no systematic imbalance; this reference distribution allows us to compute the p-value for the F statistic from the experiment (see Gerber and Green 2012, pp. 298–299). The F statistic in the sample was larger than 45 % of the F statistics under the null, for an insignificant p-value of 0.45. The results are similar using logistic regression, although see Hansen and Bowers 2008 for a discussion of the pitfalls of using logistic regression to test for covariate balance in blocked and clustered experiments.

candidate's name recognition among voters in the sample as fully 85 % of the respondents reported that they had not heard of the candidate.

Experimental Results

Table 2 presents the experimental results estimating the causal effect of the online advertisements on various outcomes, each of which is a dichotomous indicator variable set to either 0 or 1.[15] To quantify the uncertainty associated with these estimates, the first two rows of each panel calculates (one-tailed) p values using randomization inference under the sharp null hypothesis of no effect. These calculations were conducted using the `ri` package for R (Aronow and Samii 2012). (Full replication code and data files are available from the authors.) This procedure takes account of the uncertainty generated by the blocked, clustered randomization process by replicating the original randomization process 20,000 times and then calculating the average treatment effect we would have estimated under each possible randomization were there no effect of the ads on the variable of interest. The p value captures the share of randomizations that, under the sharp null hypothesis, would yield an average treatment effect estimate at least as large as the estimated obtained from the actual experimental data. Confidence intervals for these specifications were generated by inverting the test of the null hypothesis using the method described in Rosenbaum (2002, pp. 45–46).

Table 2 presents three specifications in order to demonstrate the robustness of the results. The first specification compares means without covariate adjustment and uses randomization inference to calculate confidence intervals. The second specification compares means after regressing the outcome measures on covariates describing the subjects' party identification, age, and vote history, again using randomization inference to form confidence intervals and test the sharp null hypothesis of no effect. These covariates were selected prior to the launch of the survey, in accordance with the authors' ex ante analysis plan, registered at the EGAP website.[16] Finally, we present the results from regressions that control for blocks and form confidence intervals using clustered standard errors and an assumed normal sampling distribution.

Across four dependent variables, the results of the experiment are consistent: the ad treatments appeared to have no politically consequential effect on knowledge of the candidate, favorable evaluation of the candidate, or electoral support. Moreover, these results are precisely estimated, with confidence intervals that rule out

---

[15]   The *heard of candidate* variable is set to 1 if the respondent indicated hearing of the candidate and having a positive or negative impression and 0 if the respondent had not heard of the candidate (see Q1 in Appendix 2). The *positive impression of candidate* variable is set to 1 if the respondent reported having a positive impression of the candidate and 0 if the respondent had not heard of the candidate or had a negative impression (see Q1 in Appendix 2). The *vote for the candidate* variable is set to 1 if the respondent indicated intention to vote for the candidate (see Q3 in Appendix 2) and 0 otherwise. The *recall of the online advertisements* variable is set to 1 if the respondent recalled seeing online advertisements (see Q4 in Appendix 2) and 0 otherwise.

[16]   Best practices in randomized trials include the pre-registration of an analysis plan to limit researcher discretion (Casey et al. 2012).

Table 2 Treatment effect estimates and confidence intervals of Facebook advertising in study 1

Subgroup	Estimation procedure	Heard of candidate	Positive impression of candidate	Vote for candidate	Recall seeing online ad
All subjects (N = 2,948)	Difference-in-means and randomization inference, no covariates	−0.011 [−0.042, 0.020]	0.008 [−0.017, 0.034]	0.011 [−0.018, 0.040]	0.010 [−0.010, 0.032]
	Difference-in-means and randomization inference, covariate adjustment	−0.011 [−0.041, 0.018]	0.008 [−0.017, 0.034]	0.007 [−0.019, 0.034]	0.009 [−0.011, 0.031]
	OLS, clustered standard errors and block fixed effects	−0.010 [−0.039, 0.018]	0.010 [−0.013, 0.034]	0.016 [−0.012, 0.044]	0.006 [−0.013, 0.026]
Self-reported Facebook users only (N = 1,364)	Difference-in-means, no covariates	−0.010 [−0.055, 0.038]	0.000 [−0.039, 0.041]	0.012 [−0.032, 0.056]	0.012 [−0.026, 0.051]
	Difference-in-means, covariate adjustment	−0.012 [−0.056, 0.033]	−0.003 [−0.041, 0.037]	0.000 [−0.041, 0.039]	0.007 [−0.030, 0.046]
	OLS, clustered standard errors and block fixed effects	−0.020 [−0.067, 0.026]	−0.010 [−0.050, 0.030]	0.027 [−0.018, 0.072]	0.002 [−0.034, 0.038]

Each cell records the estimate of the effect of being treated with online advertising on the dependent variable at the top of the column. 95 % confidence intervals are shown in brackets below each estimate. No results are statistically significant at the 0.05 level (one-tailed). The first two rows in each panel employ randomization inference to estimate confidence intervals, with the first row applying the procedure to unadjusted difference in means and the second employing covariate adjusted values. Rosenbaum (2002) 95 % confidence intervals for these results are calculated taking into account the blocked, clustered randomization scheme. The final row shows estimates employing OLS with block fixed effects, with 95 % confidence intervals calculated based on clustered standard errors

politically meaningful impacts. Indeed, the results cast doubt on the proposition that dozens of online ads increased the candidate's name recognition in his district by more than approximately 1.8 % points, which marks the upper end of the 95 % confidence interval.

Because only Facebook users were exposed to the ads, we also estimate the treatment effect for the subset of the respondents that reported using Facebook over the previous week. These estimates are shown in the fourth, fifth, and six rows of Table 2. Although these effects are less precisely estimated due to the decreased sample size in this subgroup, the estimates are in accordance with the results calculated among the broader sample. The 95 % confidence intervals again rule out effects that would be politically consequential, a point we elaborate further in the "Discussion" section.[17]

Consistent with the finding that the ads had little or no effect on name recognition, the other columns of Table 2 show that those in the treatment group did not become significantly more favorable toward the candidate or more likely to vote for him. Indeed, Facebook users in the treatment group were not significantly more likely to recall seeing the online advertisements.

---

[17] Results are similar when we further restrict the sample to respondents who report using Facebook every day.

Study 2: Democratic Congressional Candidate in a Non-battleground State

Given the null findings from Study 1, we sought to replicate the experiment in a different context and alter the treatments in ways that might produce detectable effects. Study 2 differed from Study 1 in three ways.

First, rather than collaborating with a relatively unknown candidate running for state legislature, we collaborated with a viable candidate running for Congress. This candidate enjoyed much higher name recognition prior to the launch of our study, and the contest itself was of much higher salience.

Next, in addition to purchasing the sidebar ads deployed by the candidate in Study 1, the candidate in Study 2 also purchased so-called sponsored stories, Facebook ads that are displayed more prominently on users' screens and display information about which of a user's friends also "Like" the candidate's own Facebook Page. However, the platform only displays these ads to users with friends who have opted to "Like" the candidate; as a result, only about 10 % of Facebook users were eligible to view these premium ads. We did not purchase these ads in Study 1 given the small proportion of constituents who would be eligible to view them. However, feedback from social media advertising consultants who commented on the results of Study 1 led us to purchase these premium ads as well as the standard ads.

Finally, rather than randomizing at the town level, we instead randomized at the level of counties. The towns we randomized in Study 1 were widely dispersed; the Congressional district of the collaborating candidate for Study 2 included some areas that were more densely populated. To minimize misclassification of subjects' treatment status (e.g., a control user logging into Facebook from a treatment location), we decided to use entire counties as clusters. However, because Facebook does not facilitate county-level ad targeting, we instead assembled groups of zip codes that fell within county boundaries and targeted the ads on the basis of these zip code groups. The clusters in Study 2 thus comprised contiguous groups of zip codes.

Ad Treatments

The campaigns' sidebar ads appeared on the right side of Facebook users' screens on all pages on the site and were 125 pixels high by 255 pixels wide, as in Study 1. When clicked, all ads brought individuals to the candidates' Facebook page.

These ads sought to build the candidates' name recognition and identify him with an important issue in the campaign:

[Name of candidate] for [Name of office]
Democrat [name of candidate] will protect [state] from reckless fracking. Show your support!
[Picture of candidate smiling]

As mentioned above, the campaign also purchased sponsored story ads that were shown to users in the treatment group who had friends that "Liked" the candidate's Facebook page.[18] At least once per day, the candidate posted new updates to his Facebook page, and users eligible to receive these ads saw these updates in their Facebook "news feeds." These updates included stories about the candidate's visits to places in the district (e.g., a factory), rallies, television commercials, endorsements, and favorable news articles.

Random Assignment Procedure

We followed the clustered random assignment procedure described in the previous section and from Study 1. First, we received a copy of the public voter registration list from the campaign. We then removed voters age 65 and older and those without known phone numbers. Remaining voters were then assigned to 752 clusters across 47 values of age (each age 18–64), 8 counties, and 2 genders. We then blocked 752 clusters into groups of four based on cluster size (Middleton and Aronow 2011). Within each block of four clusters, we assigned three clusters to the control condition and one to receive the ad treatments. Our procedure thus randomly assigned 25 % of 261,150 identifiable individuals to an online treatment or a control condition.

Treatment Delivery

We uploaded these ads to Facebook on Monday, October 28, and the site approved them for delivery to the treatment clusters shortly thereafter. The ads were served on Facebook beginning early in the morning on Tuesday, October 29, 2012 and ending late in the evening on Sunday, November 4, 2012. As with the previous study, the campaign placed an unusually high bid for each ad, $0.80 per thousand impressions, more than triple the market price it actually paid. Users in the treatment group were thus exposed to the ads as many times as the platform would allow. Facebook's records indicate that 108,783 individuals were shown the ads an average of 36.6 times, corresponding to 3.98 million impressions overall.[19]

Outcome Measurement

To measure subjects' attitudes, the polling firm Winning Connections attempted automated interviews with 154,024 individuals on the voter file (all of whom had associated phone numbers) on the evening of Monday, November 5. The firm called the numbers in a random order and did not have access to data on the treatment

---

[18] These advertisements are shown to individuals who access Facebook exclusively on mobile devices (and meet the other criteria for seeing these ads).

[19] These figures include users who were also shown the 'sponsored story' ads. Only 8.5 % of targeted users could receive these premium ads because they can only be shown to individuals who have Facebook friends who 'Liked' the candidate's Facebook page. The click rate was similar to the ads in Study 1: the ads garnered a total of about 800 clicks, for a total click rate of about 0.02 % per impression.

assignment status of the individuals. At our request, the campaign ceased all online advertising on this day.

The poll's text is given in Appendix 2. It first asked respondents to enter their age, zip code, and gender. Because the poll was conducted automatically, we could not instruct the pollster to verify voters' names; the initial questions were therefore used to verify that we reached the intended voter on the phone. We then asked respondents (1) whether they had a positive, negative, or no impression of the collaborating candidate, (2) whether they had a positive, negative, or no impression of the opposing candidate, (3) whether they recalled the candidates' main campaign issue described in the sidebar ads (hydraulic fracking; subjects were also given the choices 'trade with China' and 'abortion'), (4) whether they recalled seeing any ads for the candidate on the internet, and (5) how often they used Facebook over the last week. Questions (4) and (5) were asked after the main dependent variables of interest so as to avoid priming respondents to recall the ads.

A total of 4,359 voters answered at least one question in the automated poll.[20] Of these responses, 3,557 were successfully matched back to the voter file based on an exact match on each subject's telephone number, age, gender, and zip code. In other words, 802 responses to the poll were from respondents who did not appear in the voter file under that phone number. When we impute treatment assignment to these individuals (on the basis of self-reported age, gender, and zip code), the results are nearly identical, but we exclude them from the analysis below in keeping with the cautious procedure laid out in our ex ante pre-analysis plan.

Results: Descriptive Statistics

Descriptive statistics for the 3,557 voters who both completed the poll and were matched back to the voter file are shown in Table 3. Reassuringly, 24 % of the sample had been assigned to be treated with ads, essentially the same share of voters who had been assigned to receive the treatment ex ante (25 %); we find no evidence of differential attrition from the treatment group.[21] Although not necessary for unbiased estimation of treatment effects, it is also reassuring that the sample's

---

[20] Although automated polls have lower response rates than live polls, low response rates do not threaten unbiased estimation of sample average treatment effects as long as poll non-response is independent of treatment assignment (as appears to be the case here; see next footnote). A secondary questions concerns the generalizability of experimental estimates from the kind of voters who answer automated polls to the broader public. Our sample does not appear particularly limited in this regard as the party registration figures on the voter file and in the sample are very similar.

[21] To assess whether we had the expected covariate balance across the treatment groups in the final sample we used the same procedure as described in Study 1, regressing each potential treatment assignment vector on Facebook use, gender, party, age, and turnout in the 2012 presidential primary in the sample to generate a distribution of F statistics under the null. The F statistic in the sample was smaller than 61 % of F statistics under the null, for a p value of 0.61. We cannot reject the null hypothesis that covariates remained balanced in the sample to the degree that would be expected by chance given the randomization scheme.

Table 3 Descriptive statistics for survey respondents in study 2

Variable	Statistic
Assigned to ad treatment	24 %
Heard of candidate	55 %
Have heard of candidate and have positive impression	34 %
Self-reported Facebook use	55 %
Male	37 %
Republican	42 %
Democrat	36 %
Age (mean)	51
N (sample size)	3,577

partisan composition looks largely similar to the district at large: 32 % of the voters in the district were registered with the Democratic party and 40 % with the Republican party; these statistics are 36 and 42 % in the sample, respectively. Other statistics of the sample appear in the rows below. As expected, the descriptive statistics show that the political context of Study 2 differed dramatically from Study 1—over half of the subjects reported having previously heard of the Congressional challenger.

Experimental Results

Table 4 presents the experimental results estimating the causal effect of the candidate's online advertisements. The statistical procedures used to generate estimates and 95 % confidence intervals were identical to those employed in Study 1: in the first two rows we compare means in the treatment and control groups to obtain our point estimates and quantify statistical uncertainty by simulating the sampling distribution of the blocked and clustered randomization procedure.[22] The final row uses OLS with block fixed effects and clustered standard errors. Table 4 follows the same format as Table 2 but introduces another dependent variable— campaign issue recall—that records whether subjects correctly recalled[23] the main issue featured in the candidate's campaign (hydraulic fracking, the subject of the online ad and the main focus of the candidate's other campaign communications).[24]

---

[22] As in Table 2, a total of 20,000 simulated random assignments were used in each simulation. Inverted tests were used to form 95 % confidence intervals (Rosenbaum 2002).

[23] As the Appendix shows, we asked subjects to recall which issue they thought the candidate's campaign mainly focused on from a list. An alternative measurement strategy would have asked for open-ended responses and coded them; unfortunately as this poll was administered robotically, we were unable to collect open-ended responses from subjects.

[24] Opposition to hydraulic fracking was the main issue in the candidate's campaign. Both the banner ads and many of the candidate's 'promoted' Facebook posts during the study period concerned the issue. On the other hand, the other issues presented as response options, trade with China and abortion, were not central to the campaign. Trade was not discussed at all on the candidate's Facebook site, and abortion was mentioned only once, when the candidate posted a news article that referenced his endorsement by a pro-choice group over three months before the study began.

Table 4 Treatment effect estimates and confidence intervals of Facebook advertising in study 2

Subgroup	Estimation procedure	DV = heard of candidate	DV = positive impression of candidate	DV = recall of campaign issue	DV = recall seeing online ad
All subjects	Difference-in-means and randomization inference, no covariates	0.011 [−0.036, 0.061]	0.015 [−0.033, 0.067]	0.022 [−0.024,0.069]	0.057*** [0.024, 0.088]
	Difference-in-means and randomization inference, covariate adjustment	0.004 [−0.036, 0.046]	0.011 [−0.028, 0.052]	0.020 [−0.024,0.064]	0.053*** [0.021, 0.083]
	OLS, clustered standard errors and block fixed effects	0.021 [−0.022, 0.065]	−0.000 [−0.047, 0.047]	0.020 [−0.021,0.062]	0.059*** [0.029, 0.089]
	N	3,085	3,085	2,320	2,459
Self-reported Facebook users only	Difference-in-means and randomization inference, no covariates	−0.003 [−0.069, 0.067]	0.011 [−0.045, 0.072]	0.012 [−0.051,0.077]	0.082*** [0.036, 0.126]
	Difference-in-means and randomization inference, covariate adjustment	−0.008 [−0.067, 0.053]	0.011 [−0.045, 0.072]	0.009 [−0.050, 0.070]	0.081*** [0.035, 0.125]
	OLS, clustered standard errors and block fixed effects	−0.013 [−0.082, 0.056]	−0.017 [−0.084, 0.049]	0.001 [−0.059, 0.061]	0.084*** [0.036, 0.134]
	N	1,337	1,337	1,221	1,328

Each cell records the estimate of the effect of being treated with online advertising on the dependent variable at the top of the column. 95 % confidence intervals are shown in brackets below each estimate. The first two rows in each panel employ randomization inference to estimate the uncertainty associated with the main point estimates, with the first row applying the procedure to unadjusted difference in means and the second employing covariate adjusted values. Rosenbaum (2002) 95 % confidence intervals for these results are calculated taking into account the blocked, clustered randomization scheme. The final row shows estimates employing OLS with block fixed effects, with 95 % confidence intervals calculated based on clustered standard errors

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (one-tailed)

The first panel in the Table reports estimated intent-to-treat effects for all subjects, while the second panel reports treatment-on-treated effects for those who report using Facebook.

The number of valid survey responses for each variable is also presented under the estimates and decrease slightly across columns. As typically occurs with automated polls, some subjects abandoned the calls after answering each question. (However, there is no evidence of differential attrition across the treatment groups during the course of the survey.) Low response rates reduce statistical power but do not threaten unbiased estimation of average treatment effects among those for whom outcome measures are available.

As in Study 1, we find no evidence that the ads had consequential effects on knowledge of the candidate or his favorability ratings. The new variable we included in this study, correct recall of the candidate's main campaign issue (fracking, which was featured in the ads), also generated substantively small treatment effects that are statistically indistinguishable from zero. Although the confidence intervals are larger than in the previous study, the point estimates and confidence intervals again cast doubt on the proposition that the ads have politically meaningful effects.

The results from Study 2 did depart in one important respect from the findings from Study 1. As the last column of Table 4 indicates, subjects randomly assigned to online ads were 5.3 percentage points more likely to recall seeing items about the candidate on the Internet, an estimate that is highly statistically significant ($p < 0.001$). Among Facebook users this effect is even more pronounced, an 8.1 percentage point increase.[25] Again, randomization inference places the p value at less than 0.001. As intuition would suggest, ad recall shows no treatment effect among non-Facebook users.

As a methodological matter, this finding gives us confidence that our clustered randomization methodology worked successfully. Study 1's null effects caused us to be concerned that Facebook had inadvertently delivered the ads to some members of the control group (although we meticulously checked and verified that the treatments had been delivered as intended in Study 1). Study 2 found that ad recall was significantly higher in the treatment group, putting this concern further to rest; nevertheless, the overall pattern of findings is consistent with the results obtained from Study 1.

Discussion: Pooled Estimates and Theoretical Implications

This article developed and implemented a relatively low-cost method for rigorously assessing the impacts of online advertising. The results bring the first field experimental data to the question of whether online advertisements shape the public's political beliefs, perceptions, and evaluations. In the two studies, candidates for legislative office presented randomly selected individuals with a heavy volume of online advertisements while those in control groups were shown no advertisements. Surprisingly, we found that voters randomly assigned to view the political candidates' online ads were no more likely to recall the candidates' names, did not significantly update their opinions of the candidates, and sometimes did not recall viewing the ads at all.

---

[25] A number of differences between the contexts for Studies 1 and 2 could account for the fact ads were recalled only in the congressional race. One possibility is the salience of the race for US House. Another is that individuals primarily recall seeing online ads for entities with which they are already familiar; thus most individuals in Study 1 may not have recalled the ads due to the first collaborating candidate's relatively low baseline name recognition. Further experimentation is necessary to explore these possibilities.

Table 5 Pooled estimates for candidates' name recognition and favorability in studies 1 and 2, Facebook users only

Estimation procedure	DV = heard of candidates	DV = positive impression of candidates
OLS, clustered standard errors, weights, and block fixed effects	−0.016	−0.013
	[−0.053, 0.022]	[−0.048, 0.022]
N	2,701	2,701

Each cell records the estimate of the effect of being treated with online advertising on the dependent variable at the top of the column. Results from pooled OLS regression with block fixed effects, inverse probability weights, and clustered standard errors. 95 % confidence intervals are shown below each estimate. Neither result is statistically significant at the 0.05 level (one-tailed)

Pooled Results and Comments on Assessing Cost Effectiveness

To quantify the uncertainty associated with the overall pattern of results, we pooled the two studies together in order to generate the estimates shown in Table 5.[26] Taking the top of the 95 % confidence interval as the maximum credible estimate, we interpret the results in Table 5 to mean that we can rule out effects of these candidates' online advertising greater than 2.2 percentage points on their name recognition and evaluations.

To put this finding in perspective, consider the scope of the intervention and the manner in which it was delivered. Recall that treated Facebook users were typically exposed to the ads about 38 times in these studies. Suppose that these 38 exposures did generate a cumulative effect of 2.2 percentage points, the top of the 95 % confidence interval. If 38 exposures were necessary for this 2.2 percentage point effect, it follows that each marginal exposure to the online political advertisements increased name recognition or candidate favorability by an average of (2.2/38=) 0.058 percentage points. In other words, even if the true average treatment effect were in fact at the top of the 95 % confidence interval, just (1/0.00058=) 1 in 1,700 people learned the name of the candidates or gained a favorable impression of them from each exposure to their Facebook advertisements. Based on this evidence, online advertisements appear unlikely to play a meaningful role in determining a candidate's success or failure.

Although these results generally run counter to the notion that online ads have substantively large effects on their viewers' attitudes and behavior, it remains possible that online advertising remains a *cost-effective* persuasion tactic given the ads' very low price (Lewis and Rao 2012). Suppose, for example, that the candidates' online ads attracted votes at a rate of $10 per vote, which would be efficient by comparison to most campaign tactics (Green and Gerber 2008, p. 139). Recall that the candidate in Study 1 was only able to purchase $200 worth of ads per week—Facebook was unable to provide any more advertising given the finite supply of Facebook users and the finite number of pages they each load on the site. If Facebook advertisements won votes at a respectable $10 per vote and the

---

[26] We derived these results by pooling the datasets and using OLS regression with block fixed effects, inverse probability weights, and clustered standard errors. Inverse weights are calculated as 1/the probability of assignment to treatment for the treatment group and 1/(1 − the probability of assignment to treatment) for the control group (Gerber and Green 2012).

candidate purchased $200 worth of ads, we would expect only 20 votes to change. To change merely 20 voters' minds out of the roughly 20,000 who were exposed to the ads would have rendered the $200 ad buy fairly cost effective (at $10/vote). However, to reliably detect the implied 0.1 percentage point effect (20/ 20,000 = 0.001) of such an ad would require an experiment of roughly 5 million voters. Our experiment (and indeed the legislative districts we studied) contained far fewer than 5 million individuals and could not have detected effects of this miniscule size. As with commercial online advertisements, understanding whether very cheap political online advertisements are cost effective will likely remain "nearly impossible" (Lewis and Rao 2012). At the same time, our experiments do cast doubt on the view that online advertising has substantively meaningful impacts on political attitudes or electoral outcomes.

Two experiments, of course, can hardly provide the last word on a phenomenon as complex as online political advertising, and further research is needed to assess whether online ads would prove more potent if delivered via different modes (e.g., video) or in different contexts. For example, although we anticipated that relatively low salience elections would be conducive to strong advertising effects, it is possible that online advertisements could prove more effective in higher salience contests or for candidates who are already relatively well known. Likewise, although we study political candidates' efforts at winning votes, perhaps other entities such as non-profit or issue advocacy organizations tend to be more persuasive; or, perhaps online ads are successful in altering attitudes beyond candidate evaluations or in affecting behavior. Online video advertisements or large and colorful banner ads could prove more effective than the static display ads we purchased. The Internet can also be deployed for a variety of political purposes other than persuasion; for example, the Obama campaign leveraged Facebook data to allow supporters to identify their friends in swing states and mobilize them to vote, a tactic that may prove more effective than impersonal display advertisements. The experimental methodology developed in this article will allow scholars and practitioners to shed light on important questions such as these in future research.

Theoretical Implications for Mass Communication

Our results also leveraged the unique potential of online advertisements to contribute to longstanding theoretical questions about the degree to which individuals' perceptions, beliefs, and evaluations can be influenced by impersonal mass appeals (e.g., Hartmann 1936; DellaVigna and Gentzkow 2009). Reigning theories of mass communication stress the potency of repeated exposure (e.g., Atkin and Heald 1976; Lodge et al. 1995; Grimmer et al. 2012)—even to subtle messages (Zajnoc 1968; Bargh et al. 1992; Kam and Zechmeister 2013)—and the power of communication that does not run counter to individuals' pre-existing values (e.g., Zaller 1992). From such perspectives, online advertisements would seem to represent a propitious way to generate sizeable shifts in the public's perceptions, beliefs, and evaluations, especially for outcomes such as candidate name recognition and candidate support (e.g., Zaller 1996; Ladd and Lenz 2009): it can be essentially guaranteed that individuals will be exposed to online ads dozens of

times, and the ads often contain messages that few individuals should be predisposed to resist (e.g., a graphic displaying the name of a candidate for office for the mere purpose of informing viewers of his candidacy). Our results present a challenge to the sufficiency of these theories: the treatments essentially guaranteed that subjects in our studies crossed these two theoretical hurdles to communication effects, but in neither study do we find evidence of increasing awareness of the candidates. It seems that a final hurdle for effective mass communication—individuals' interest in processing these messages and retaining their contents (e.g., Petty and Cacioppo 1986)—largely stymied these attempts at mass influence.

Campaigns to change the public's attitudes, beliefs, and behaviors increasingly rely upon online advertisements, yet evidence remains sparse that impersonal mass communications are able to effect large, enduring changes in individuals' attitudes and behaviors, online or not. To be sure, the evidence strongly suggests that mass communications can sometimes influence individuals in the short term, even in sufficient numbers to swing very close elections (e.g., Hill et al. 2012). Indeed, our findings may seem surprising in light of other field experiments that have found sizeable short-term effects of mass communication on candidate choice in both low-salience (Panagopoulos and Green 2008; Gerber et al. 2011b) and high-salience elections (Huber and Arceneaux 2007; Gerber et al. 2011a).[27] However, a rich research tradition also raises doubts about the ability of mass messages to leave more than a fleeting impression on the public (e.g., Campbell et al. 1960; Klapper 1960), even when sympathetic subjects are directly exposed to persuasive content (e.g., Hovland et al. 1949). Consistent with this "minimal effects" perspective, observational studies and field experiments suggest that individuals often forget televised messages within a matter of days or even hours (e.g., Gerber et al. 2011a; Hill et al. 2012; Sears and Kosterman 1994); lab studies of negative advertising (e.g., Mutz and Reeves 2005) and issue framing (e.g., Chong and Druckman 2010) often find that the effects of experimental stimuli decay rapidly; and impersonal behavioral interventions often exhibit rapid decay or fail altogether (e.g., Allcott and Rogers 2012; Galiani et al. 2012).

Whether it is because people typically do not attend closely to impersonal mass messages concerning subjects they have little interest in or because they quickly forget their content amid life's distractions, attempts at mass influence quite often have minimal effects. A growing conventional wisdom among internet advertisers and journalists suggests that online advertisements nonetheless have profound effects on the public—for example, Facebook has claimed that its ads have moved vote shares by around 20 percentage points in some cases (e.g., Facebook 2011), Google has suggested that Senator Scott Brown's online advertisements "seal[ed] his] upset victory in 2010" (Google 2013), and numerous journalistic retrospectives have credited online ads with political candidates' victories (e.g., Edwards 2012a, b). Our experimental design offers a way to rigorously evaluate such claims by assessing the efficacy of online messages. In light of our evidence, it appears that attempts to influence the mass public online warrant the same healthy skepticism as their offline counterparts.

---

[27] See also a growing experimental and quasi-experimental literature on the broader political and social effects of the mass media (e.g., Adena et al. 2013; Paluck and Green 2009).

Appendix 1

In order to illustrate how our cluster-randomization procedure works, this Appendix provides a brief stylized example.

Suppose we were to begin with the 10 individuals shown in Table 6, each of which has a corresponding age, gender, and location.

Table 6 Example individuals on voter file

Person	Age	Gender	Location
A	30	Male	New York
B	30	Male	San Francisco
C	30	Male	San Francisco
D	30	Female	San Francisco
E	31	Male	New York
F	31	Female	New York
G	31	Female	San Francisco
H	31	Female	New York

Table 7 Example clusters

Cluster	Person(s)	Age	Gender	Location
Cluster 1 (30-M-NY)	A	30	Male	New York
Cluster 2 (30-M-SF)	B, C	30	Male	San Francisco
Cluster 3 (30-F-SF)	D	30	Female	San Francisco
Cluster 4 (31-M-NY)	E	31	Male	New York
Cluster 5 (31-F-NY)	F, H	31	Female	New York
Cluster 6 (31-F-SF)	G	31	Female	San Francisco

Table 8 Example blocks

Block	Cluster
Block 1	2 (30-M-SF), 5 (31-F-NY)
Block 2	1 (30-M-NY), 4 (31-M-NY)
Block 3	3 (30-F-SF), 6 (31-F-SF)

Table 9 Example treatment assignment

Block	Cluster	Treatment assignment
Block 1	2 (30-M-SF)	Treatment (B, C)
Block 1	5 (31-F-NY)	Control (F, H)
Block 2	1 (30-M-NY)	Treatment (A)
Block 2	4 (31-M-NY)	Control (E)
Block 3	3 (30-F-SF)	Treatment (D)
Block 3	6 (31-F-SF)	Control (G)

Individuals B and C as well as individuals F and H cannot be randomized at the individual level because they share the same age, gender, and location. However, these individuals can be randomized at the cluster level with clusters shown in Table 7.

To conduct block randomization of these clusters, we combine these clusters in blocks of similarly sized clusters (see Gerber and Green 2012 for the rationale behind blocking on cluster size). In this case, Clusters 2 and 5 each contain two individuals, so they are put together in Block 1, shown in Table 8. The other four clusters are all of the same size, so we block them into pairs based on other attributes (in this case, prioritizing gender and location similarity).

Finally, we randomize treatment assignment to the clusters within these blocks. Table 9 shows an example of how treatment assignment might be realized within these blocks. For example, within Block 1, Cluster 2 was assigned to Treatment and Cluster 5 was assigned to control. This means that 30 year old males in San Francisco would be treated with advertisements but 31 year old females in New York would not be. Persons B and C would thus receive treatment, while persons F and H would not.

Appendix 2

Text of Telephone Survey in Study 1 (Live Interviewers)

Hi, I'm calling with a brief academic research study. Could I speak with [voter name] please?

Q1. [collaborating candidate] is a candidate for state legislature this year. Could you please tell me whether you've heard of [collaborating candidate] and, if so, whether you have a positive or negative impression of him? If you don't recognize [collaborating candidate]'s name, just let me know.

- Have *not* heard of him
- Have heard of him, have *positive* impression
- Have heard of him, have *negative* impression

Q2. [candidate's opponent] is also running for state legislature this year. Could you please tell me whether you've heard of [candidate's opponent] and, if so, whether you have a positive or negative impression of him? If you don't recognize [candidate's opponent]'s name, just let me know.

- Have *not* heard of him
- Have heard of him, have *positive* impression
- Have heard of him, have *negative* impression

Q3. If the election for state assembly were held today, would you vote for [party], [candidate's opponent], or the [party], [collaborating candidate]?

- [party] [candidate's opponent]
- [party] [collaborating candidate]
- Undecided, or other candidate

Q4. Thanks. Thinking about [collaborating candidate] for a moment – have you seen, heard, or read anything about [collaborating candidate] on the internet over the last week or so?

- Yes
- No

Q5. Thanks. Finally, on a different topic, how often have you logged onto the website Facebook in the last week? Just about every day, a few times, only once, or never?

- Just about every day
- A few times
- Only once
- Never/I don't have an account

Text of Telephone Survey in Study 2 (Automated Calls)

Hi, we're calling with a brief academic research study about tomorrow's election.

Q1. First, for verification purposes, could you please enter your current, two-digit age? [break off if the person is 17 or younger or 65 or older]

Q2. Thank you, and could you please enter 1 if you are male, or 2 if you are female?

Q3. And finally, for verification purposes, could you enter your zip code?

Q4. Thank you. On to our questions about the election, which should take just a moment. In tomorrow's Presidential election between Barack Obama and Mitt Romney, who do you plan to vote for? Press 1 if you plan to vote for Barack Obama, 2 if you plan to vote for Mitt Romney, 3 if you plan to vote for some other candidate, and 4 if you do not plan to vote.

Q5. Thank you. Now moving on to the race for Congress, [collaborating candidate]is a candidate for Congress this year. Could you please tell me whether you have a positive impression of [collaborating candidate], negative impression of [collaborating candidate], or have not heard of him before? Press 1 if you have a positive impression of [collaborating candidate], 2 if you have a negative impression of [collaborating candidate], and 3 if you are not familiar with the name [collaborating candidate].

Q6. Thank you. [candidate's opponent] is also candidate for Congress this year. Could you please tell me whether you have a positive impression of [candidate's opponent], negative impression of [candidate's opponent], or have not heard of him before? Press 1 if you have a positive impression of [candidate's opponent], 2 if you have a negative impression of [candidate's opponent], and 3 if you are not familiar with the name [candidate's opponent].

Q7. Thanks. Thinking about [collaborating candidate] for a moment, do you happen to recall if [collaborating candidate]'s campaign has focused mostly on the issue of abortion, hydraulic fracking, or free trade with China? Press 1 if you recall [collaborating candidate]'s campaign focusing on abortion, 2 if you recall [collaborating candidate]'s campaign focusing on fracking, and 3 if you recall [collaborating candidate]'s campaign focusing on free trade with China.

Q8. Thanks. Thinking about [collaborating candidate] for another moment – have you seen, heard, or read anything about [collaborating candidate] on the internet over the last week or so? Press 1 if you do recall items on the internet about [collaborating candidate], and press 2 if you do not recall seeing anything on the internet about [collaborating candidate].

Q9. Thanks. Finally, on a different topic, how often have you logged onto the website Facebook in the last week? Press 1 if you've logged on just about every day, press 2 if you've logged on at least once, press 3 if you haven't logged into Facebook at all over the past week, or press 4 if you don't have an account.

That's all, thank you for your time.

References

Adena, M., Enikolopov R., Petova M., Santarosa V., & Zhuravskaya E. (2013). Radio and the rise of the Nazis in Prewar Germany. SSRN Working Paper.

Allcott, H., & Rogers, T. (2012). The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. Working Paper, New York University.

Aral, S., & Walker, D. (2010). Creating social contagion through viral product design: A randomized trial of peer influence in networks. Working Paper, New York University.

Arceneaux, K. (2005). Using cluster randomized field experiments to study voting behavior. *Annals of the American Academy of Political and Social Science, 601*(1), 169–179.

Aronow, P.M., & Samii, C. (2012). ri: The R package for performing randomization-based inference for experiments. R package version 0.9. http://CRAN.R-project.org/package=ri.

Atkin, C., & Heald, G. (1976). Effects of political advertising. *Public Opinion Quarterly, 40*(2), 216–228.

Bailard, C. S. (2012). A field experiment on the internet's effect in an African election: Savvier citizens, disaffected voters, or both? *Journal of Communication, 62*(2), 330–344.

Bakshy, E., Eckles, D., Yan, R., & Rosenn, I. (2012). Social influence in social advertising: Evidence from field experiments. In *Proceedings of the 13th ACM Conference on Electronic Commerce (EC' 12)* pp. 146–161.

Bargh, J. A., Chaiken, S., Govender, R., & Pratto, F. (1992). The generality of the automatic attitude activation effect. *Journal of Personality and Social Psychology, 62*(6), 893–912.

Bond, R. M., Fariss, C. J., Jones, J. J., Adam, D. I., Kramer, C. M., Settle, J. E., et al. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature, 489*, 295–298.

Buscher, G., Cutrell, E., & Morris, M.R. (2009). What do you see when you're surfing? Using eye tracking to predict salient regions of web pages. Computer–human interaction.

Campbell, A., Converse, P. E., Miller, W. E., & Stokes, D. E. (1960). *The American voter*. New York: Wiley.

Carlisle, J.E. & Patton, R.C. (2013). Is social media changing how we understand political engagement? An analysis of Facebook and the 2008 presidential election. *Political Research Quarterly*. Forthcoming.

Casey, K., Glennerster, R., & Miguel, E. (2012). Reshaping institutions: Evidence on aid impacts using a pre-analysis plan. *Quarterly Journal of Economics, 127*(4), 1755–1812.

Chong, D., & Druckman, J. N. (2010). Dynamic public opinion: Communication effects over time. *American Political Science Review, 104*(4), 663–680.

Danaher, P. J., & Mullarkey, G. W. (2003). Factors affecting advertising recall: A study of students. *Journal of Advertising Research, 43*(3), 252–267.

DellaVigna, S., & Gentzkow, M. (2009). Persuasion: Empirical evidence. NBER Working Paper.

Edwards, J. (2012a). DATA: Google totally blows away Facebook on ad performance. Business insider. http://articles.businessinsider.com/2012-05-15/news/31704866_1_mark-zuckerberg-image-ads-advertising-platform. Accessed 24 May 2013.

Edwards, J. (2012b). Michigan Supreme Court judge says Facebook campaign won her election. http://www.businessinsider.com/michigan-supreme-court-judge-bridget-mccormacks-facebook-election-campaign-2012-12. Accessed 24 May 2013.

eMarketer. (2012a). Facebook usage beats TV during working hours. http://www.emarketer.com/Article.aspx?R=1008657. Accessed 16 Oct 2012.

eMarketer. (2012b). US online ad spending to close in on $40 billion. http://www.emarketer.com/Article.aspx?R=1008783. Accessed 16 Oct 2012.

Facebook (2011). Case study: Reaching voters with Facebook ads (vote no on 8). https://www.facebook.com/notes/us-politics-on-facebook/case-study-reaching-voters-with-facebook-ads-vote-no-on-8/10150257619200882. Accessed 24 May 2013.

Fitzgerald, B. (2012). Most popular sites 2012: Alexa ranks the 500 most-visited websites. http://www.huffingtonpost.com/2012/08/09/most-popular-sites-2012-alexa_n_1761365.html. Accessed 24 May 2013.

Galiani, S., Gertler, P., & Orsola-Vidal A. (2012). Promoting handwashing behavior in Peru: The effect of large-scale mass-media and community level interventions. Working Paper, The World Bank.

Gerber, A.S., & Green, D.P. (2012). *Field Experiments: Design, Analysis, and Interpretation.* New York: W. W. Norton.

Gerber, A. S., Green, D. P., Kaplan, E. H., & Kern, H. L. (2010). Baseline, placebo, and treatment: Efficient estimation for three-group experiments. *Political Analysis, 18*, 297–315.

Gerber, A. S., Gimpel, J. G., Green, D. P., & Shaw, D. R. (2011a). 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.

Gerber, A. S., Kessler, D. P., & Meredith, M. (2011b). The persuasive effects of direct mail: A regression discontinuity based approach. *Journal of Politics, 73*(1), 140–155.

Goldfarb, A., & Tucker, C. (2011). Search engine advertising: Channel substitution when pricing ads to context. *Management Science, 57*(3), 458–470.

Google (2013). Get out the vote—Google politics and elections ad toolkit. http://www.google.com/ads/elections/get-out-vote.html.

Green, D. P., & Gerber, A. S. (2008). *Get out the vote: How to increase voter turnout* (2nd ed.). Washington, D.C.: Brookings Institution Press.

Grimmer, J., Messing, S., & Westwood, S. (2012). How words and money cultivate a personal vote: The effect of legislator credit claiming on constituent credit allocation. *American Political Science Review, 106*(4), 703–719.

Hampton, K.N., Goulet, L.S., Rainie, L., & Purcell, K. (2011). Social networking sites and our lives. Pew Research Center for Internet and American Life Project. http://pewinternet.org/Reports/2011/Technology-and-social-networks.aspx.

Hansen, B. B., & Bowers, J. (2008). Covariate balance in simple, stratified and clustered comparative studies. *Statistical Science, 23*(2), 219–236.

Hartmann, G. W. (1936). Field experiment on the comparative effectiveness of 'emotional' and 'rational' political leaflets in determining election results. *Journal of Abnormal Psychology, 31*, 99–114.

Hill, S.J., Lo J., Vavreck, L., & Zaller, J. (2012). How quickly we forget: The duration of persuasion effects from mass communication. Working Paper, University of California, Los Angeles.

Hovland, C.I., Lumsdaine, A.A., & Sheffield, F.D. (1949). *Experiments on mass communication*. Studies in Social Psychology.

Huber, G.A. & Arceneaux, K. (2007). Identifying the persuasive effects of presidential advertising. *American Journal of Political Science 51*(4):961–981.

Iyengar, S. (1991). *Is anyone responsible? How television frames political issues*. Chicago: University of Chicago Press.

Kaid, L. L. (2012). Political advertising as political marketing: A retro-forward perspective. *Journal of Political Marketing, 11*, 29–53.

Kam, C.D., & Zechmeister, E.J. (2013). Name recognition and candidate support. *American Journal of Political Science*. Forthcoming.

Kaye, K. (2012). Digital ad spend gap widens between Obama and Romney. ClickZ. http://www.clickz.com/clickz/news/2192377/digital-ad-spend-gap-widens-between-obama-and-romney.

Klapper, J. T. (1960). *The effects of mass communication*. New York: Free Press.

Lacetera, N., Macis, M., & Mele, A. (2012). Viral altruism? A natural field experiment of social contagion in on-line networks. Working Paper. NET Institute.

Ladd, J. M., & Lenz, G. S. (2009). Exploiting a rare communication shift to document the persuasive power of the news media. *American Journal of Political Science, 53*(2), 394–410.

Lewis, R.A. & Reiley, D.H. (2012). "Does retail advertising work? Measuring the effect of advertising on sales via a controlled experiment on Yahoo! Working paper, Google.

Lewis, R.A., Reiley, D.H., & Schreiner, T.A. (2012). Ad attributes and attribution: Large scale field experiments measure online consumer acquisition. Working Paper, Google.

Lewis, R.A. & Rao, J.M. (2012). On the near impossibility of measuring advertising effectiveness. Working Paper, Google.

Lodge, M., Steenbergen, M. R., & Brau, S. (1995). The responsive voter: Campaign information and the dynamics of candidate evaluation. *American Political Science Review, 89*(2), 309–326.

Middleton, J.A., & Aronow, P.M. (2011). Unbiased estimation of the average treatment effect in cluster-randomized experiments. Yale University. Unpublished manuscript.

Mondak, J.J. (1995). *Nothing to read: Newspapers and elections in a social experiment*. Michigan.

Mutz, D. C., & Reeves, B. (2005). The new videomalaise: Effects of televised incivility on political trust. *American Journal of Political Science, 99*(1), 1–15.

Nickerson, D. W. (2005). Scalable protocols offer efficient design for field experiments. *Political Analysis, 13*(3), 233–252.

Nickerson, D. W. (2007). Does email boost turnout? *Quarterly Journal of Political Science, 2*, 369–379.

Paluck, E. L., & Green, D. P. (2009). Deference, dissent, and dispute resolution: An experimental intervention using mass media to change norms and behavior in Rwanda. *American Political Science Review, 103*(4), 622–644.

Panagopoulos, C., & Green, D. P. (2008). Field experiments testing the impact of radio advertisements on electoral competition. *American Journal of Political Science, 52*(1), 156–168.

Patterson, T. E., & McClure, R. D. (1977). *The unseeing eye*. New York: Berkley Publishing Group.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology, 19*, 124–181.

Pew Internet and American Life Project (2012). Changes to the way we identify internet users. http://www.pewinternet.org/∼/media//Files/Reports/2012/Commentary%20-%20counting%20internet%20users.pdf. Accessed 24 May 2013.

Public Religion Research Institute (2012). Few Americans use social media to connect with their faith communities. http://publicreligion.org/research/2012/08/july-2012-tracking-survey/. Accessed 24 May 2013.

Rosenbaum, P. (2002). *Observational studies* (2nd ed.). New York: Springer.

Ryan, T. J. (2012). What makes us click? Demonstrating incentives for angry discourse with digital-age field experiments. *Journal of Politics, 74*(4), 1138–1152.

Sears, D.O., & Kosterman, R. (1994). Mass media and political persuasion. In *persuasion: Psychological insights and perspectives*. Boston: Allyn and Bacon.

Stokes, D. E. (1963). Spatial models of party competition. *American Political Science Review, 57*(2), 368–377.

Stokes, D.E. & Miller, W.E. (1966). Party Government and the Saliency of Congress. In E.M. Warren, & E.S. Donald (Eds.), *Elections and the political order*, Angus Campbell, Phillip Converse, New York: Wiley.

Taber, C. S., & Lodge, M. (2013). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science, 50*(3), 755–769.

Van Grove, J. (2012). Nearly 20 % of Facebook's mobile users skip the website altogether. Venture Beat. http://venturebeat.com/2012/07/31/facebook-mobile-maus/.

Wooldridge, J. M. (2003). Cluster-sample methods in applied econometrics. *American Economic Review, 93*(2), 133–138.

Zajnoc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology Monograph Supplement, 9*(2), 1–27.

Zaller, J. R. (1992). *Nature and origins of mass opinion*. Cambridge: Cambridge University Press.

Zaller, J.R. (1996). The myth of massive media impact revived. In C.M. Diana, M.S. Paul, A.B. Richard (Eds.), *Political Persuasion and Attitude Change* (pp. 17–78). Ann Arbor: University of Michigan Press.