Does Microtargeting Work? Evidence from an Experiment during the 2020 United States Presidential Election

Musashi Jacobs-Harukawa

Abstract

Political science research consistently shows that political advertisements have small and uniform effects. This contrasts with claims that microtargeted campaigning can sway electoral outcomes and poses a threat to society. This paper introduces a novel experimental design simulating a targeted campaign to empirically test whether targeting political advertisements can be effective. Participants in an online survey experiment view one of five anti-Biden advertisements. Respondents assigned to control are allocated advertisements at random. This data is used to train a model predicting Biden favorability from advertisement and pre-treatment traits. Respondents in the treatment group view the advertisement that this model predicts to be the most effective. The difference between targeted and random allocation is an 8.7 percentage point increase in Biden dislike and a 7.1 percentage point decrease in intent to vote Biden among unaligned voters who had not yet cast their vote ($N = 586$). The effect is negligible among partisan voters.

Word Counts: 3995 (150 in abstract)

Introduction

Following the Brexit referendum and the election of Donald Trump in 2016, observers raised questions surrounding the role of sophisticated digital campaign strategies employed by *Cambridge Analytica* and others (Simon 2019). Watchdogs, journalists, and scholars identified numerous issues stemming from these technologies: threats to privacy from the systematic harvesting of individual data by corporations (Zuboff 2015), threats to public accountability from personalization of political information (Fowler et al. 2021), and threats to autonomy from mass manipulation (Burkell and Regan 2019).

Although the normative discussion frequently proceeds with the assumption that digital campaigning strategies are effective (Krotoszynski 2020), empirical political science research indicates this conclusion is at best premature (Nickerson and Rogers 2020). A large panel study by Coppock, Hill, and Vavreck (2020) on US presidential election campaign advertisements finds not only small average persuasive effects, but a lack of heterogeneity that seems to leave little room for targeting to operate. In contrast, research in psychology finds that campaigns leveraging knowledge of receiver characteristics can improve message effectiveness (Zarouali et al. 2020). Efforts to reconcile these results are hampered by the challenges of studying campaigns and their activities online.[1]

I introduce a novel experimental design to directly estimate the effect of targeting political advertisements. The experiment runs in two immediately sequential stages. Respondents in the first stage are randomly assigned one of five real anti-Biden advertisements. I use this data to train an online targeting algorithm, which optimally allocates advertisements on the basis of respondent characteristics in the second stage. Provided the participants in the two stages are comparable, this design yields a consistent estimate of the effect of microtargeting on candidate preference and voting intention.

The results indicate that microtargeting can have an effect among specific groups. The treatment had no significant average effect on partisan respondents. Among unaligned respondents who had not already voted at the time of the survey, optimized allocation caused an 8.7 percentage point (pp) increase in Biden dislike and 7.1pp decrease in intent to vote for Biden relative to receiving a random advertisement.

The findings of this paper contribute to several strands of literature. In political science, these results suggest that prior studies finding no effect of political advertisements may have insufficiently considered heterogeneity on respondent characteristics. For the legal and ethical debate regarding this technology, the magnitude of the estimates found in this experiment implies that microtargeting may be sufficient to change the result of close elections with a sufficient number of unaligned voters.

[1] See, for example, difficulties faced by Edelson et al. (2019, 3).

Motivation

I begin by clarifying the distinction between *tailoring* and *targeting*. *Tailoring* a message is constructing it to appeal to a specific audience. *Targeting* is delivering a message only to the intended audience (Fowler et al. 2021). I use "microtargeting" to mean targeting on the basis of individual-level data such as gender or ethnicity, as opposed to aggregate data such as postal code. This article focuses on the effect of the advertisement allocation strategies of data-driven political campaigns, or *political microtargeting*.

A dramatic rise in online political campaigning brought with it a myriad of problems (Wood and Ravel 2017). Of particular concern is the ability of political campaigns, operating on platforms such as Facebook, to show narrowly tailored advertisements to the individuals they are designed to affect (Fowler et al. 2021). Legal and ethical scholars note systemic threats to democracy arising from this technology and the difficulty of limiting these harms (Burkell and Regan 2019; Hindman 2018). Although these arguments largely assume that microtargeting works, many political scientists remain skeptical of the claims of campaigners and their ability to alter electoral outcomes (Nickerson and Rogers 2020).

A decade of field experiments studies the effect of political advertising in various mediums (Gerber et al. 2011; Edelson et al. 2019), discovering largely small or null effects (Kalla and Broockman 2018). In a recent field experiment Coppock, Hill, and Vavreck (2020) conclude: "expensive efforts to target or tailor advertisements to specific audiences require careful consideration. The evidence… shows that the effectiveness of advertisements does not vary greatly from person to person or from advertisement to advertisement" (6). Correspondingly, targeting the delivery of advertisements should be unlikely to yield any gains in effectiveness. It is difficult to square these null findings with claims that microtargeting is a threat to democracy.

Research on *psychometric profiling*, the building of psychological voter "profiles" for tailoring advertisements, provides evidence that gains should be possible. Zarouali et al. (2020) use a personality profiling algorithm to label participants as introverts or extroverts, and then show a personality-congruent advertisement based on this label, finding that correctly matched advertisements have a stronger effect. The psychological models of persuasion underlying this technique (e.g. Petty and Cacioppo 1986; Cialdini 2007) emphasize that whether a message is processed *emotively* or *cognitively*[2] depends on an interaction between message content, receiver characteristics and receiver social context. Knowledge of receiver characteristics allows a campaign to tailor an advertisement to increase the likelihood that the message will be processed in a manner that results in opinion or attitude change.

The psychometric approach maximizes the persuasive effect of advertisements on the basis of a deductive

[2] *Elaboration Likelihood Model* (Petty and Cacioppo 1986): modes of processing information range from central (critical, rational) to peripheral (heuristic, emotive).

theoretical model associating psychological types and advertisement response. In contrast, in my approach I train a model to directly optimize the outcome (candidate preference), and thus avoid relying on the validity of a particular theoretical model to produce optimal respondent-advertisement matchings. By demonstrating whether a campaign can improve their performance by optimizing advertisement allocation on receiver traits, my theoretically agnostic approach shows whether there are possible gains from microtargeting regardless of the specific technique employed.

Design

Experiment

In order to estimate the effect of microtargeting, I compare the average outcomes between an untargeted control group and targeted treatment group. I operationalize targeting to mean that targeted respondents are shown the optimal advertisement conditional on their traits, while untargeted respondents are shown an advertisement independently of their traits (i.e. at random).

Respondents in both groups are shown one of five real anti-Biden advertisements (see table 1), with the allocation rule differing between treatment and control. These advertisements are chosen from the set of all Trump advertisements run in the final five months of the election, on the criteria that they are clearly tailored to different audiences, and no advertisement is likely to outperform all others.[3]

Table 1: Advertisements and Descriptions

Title	Description
They Mock Us	Clinton and Biden are mocking you (In-Group)
Why did Biden let him do it?	Hunter Biden's ostensible corruption
Biden will come for your guns	Second Amendment; Biden will steal guns
Insult	Biden: Black Trump supporters not Black
Real Leadership	Obama/Biden caused wars, neglected veterans

Respondents assigned to control are randomly allocated an advertisement using permuted block randomization with uniform probability. This data is used to train a model that predicts the outcome, Biden favorability, as a function of the respondent's characteristics[4] and the advertisement. For respondents in the treatment

[3]See SI-A: Case and Advertisement Selection for details.

[4]Characteristics measured are age, gender, race, income group, state, news interest, whether they think the country is on the right track, partisanship, and ideology.

group, this model generates predictions for each of the five advertisements and then allocates the one that minimizes Biden favorability.[5]

Algorithm

Five models are fitted to predict Biden favorability as a function of advertisement allocation and pre-treatment traits. The first three are decision tree-based algorithms chosen for their ability to learn highly conditional response surfaces: Random Forest (RF), AdaBoost and GradBoost. A Multi-Layer Perceptron and Support Vector Machine are included for comparison, with the latter hedging on the possibility of a smooth response surface. All models are standard, "off-the-shelf" algorithms implemented in the popular `scikit-learn` library (Pedregosa et al. 2011). This highlights that campaigns with low data science sophistication can easily implement similar approaches.

The models are compared on root mean squared error, maximum error and prediction time[6], in order to find the model best able to give quick and accurate predictions. RF and AdaBoost performed the best on all of these metrics. Between RF and AdaBoost, RF performed weakly better and was less likely to predict ties, and was thus chosen for predicting optimal advertisements[7].

Sample Balance

The first sample is necessarily assigned to the control group because randomized allocation of advertisements is necessary to train a model that produces unconfounded predictions of the outcome for targeting respondents in the treatment group. Ideally, respondents in the second sample would be randomly split between treatment and control (and the first sample used only for training the algorithm). Due to resource constraints, I instead used the first sample as the control group and assigned all respondents in the second sample to treatment. Thus treatment is assigned based on when the subject took the survey, raising the possibility that the two groups differ systematically in politically relevant ways.

The online format of the experiment allowed for the rapid deployment and simultaneous testing of many respondents, and the switch-over between stages where the targeting model was trained and uploaded was completed in under an hour. Given this and the results of a power analysis and simulation of the experiment based on the Coppock, Hill, and Vavreck (2020) data, I opted to assign the groups sequentially in order to have enough observations to train the targeting algorithm (300 per advertisement) and detect potentially

[5] In this case, I minimize the outcome because the advertisements are negative.

[6] Prediction time mattered because long queues during traffic bursts could create uncontrolled variation in the respondent's experience of the survey.

[7] In the case of ties, an optimum was randomly chosen.

small effects of targeting (900 in treatment).[8]

I conduct multiple checks for covariate imbalance; a chi-squared test, covariate balance check, and logistic regression predicting whether a respondent was collected in the first or second sample. None of these tests provide evidence of systematic difference between the samples.[9]

Hypotheses

I measure the effect of targeting on three outcomes that a campaign is likely to want to influence. For a given respondent, I hypothesize that targeting anti-Biden advertisements will *increase* Biden dislike, *decrease* intent to vote for Biden, and *decrease* turnout intention, relative to showing an advertisement at random. Note the latter two outcomes are only measurable for individuals who have not already voted at the time of the survey. Therefore, all hypotheses are tested for the subset of respondents who had not already voted at the time of the survey, but only Biden dislike is tested with the full sample.

In line with theories of motivated reasoning, whereby partisan voters resist information contrary to their identity orientations, and findings in related studies that the effects of advertisements depend on the partisan affiliation of the recipient (Coppock, Hill, and Vavreck 2020; Broockman and Kalla 2020), I test all hypotheses conditioning on partisanship.

Results

Effect of Targeting

The experiment took place on 28 October 2020. $1,500$ and 900 responses were collected for the control and treatment groups. After filtering for irregularities[10], the experiment yielded $2,261$ valid responses, with $1,416$ in control and 845 in treatment.

Figure 1 shows the estimated effect of targeting on Biden dislike, intent to vote for Biden, and turnout intention among voters who had not already voted at the time of the survey ($N = 1,160$), conditional on partisan self-identification. Each pair of bars shows the predicted level of the outcome untargeted and targeted respondents, with 95 percent confidence intervals, for Democrat, Unaligned and Republican respondents in the corresponding facet.

The numbers along the top show the estimated conditional average treatment effect (CATE) of targeting

[8] See Figure A.4 in SI-B.
[9] See SI-A: Randomization Check.
[10] Irregularities include incomplete surveys and failed attention checks. See SI-A.

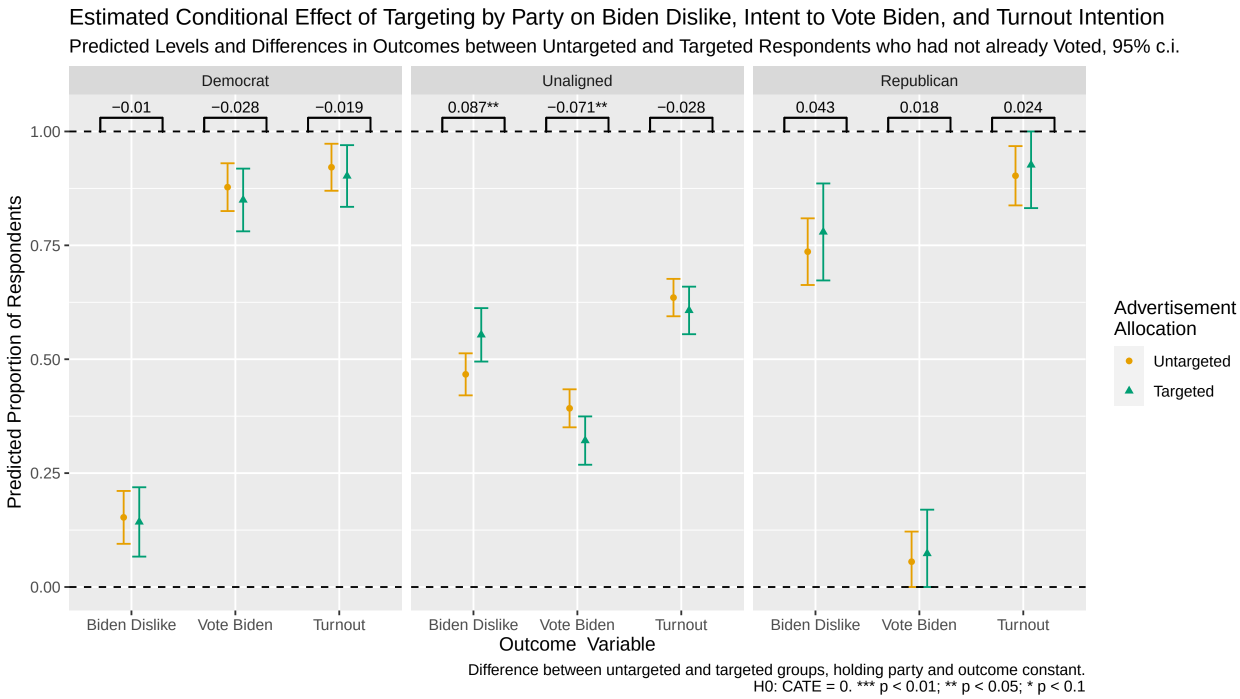

Figure 1: Effect of Targeting on Biden Dislike, Biden Vote, and Turnout; Predicted Levels and Differences.

for each outcome and partisanship group. The central facet shows that the estimated CATE of targeting among unaligned respondents who had not voted at the time of the survey on Biden dislike and intent to vote for Biden are 0.087 and -0.071 respectively. This translates to targeting resulting in an 8.7pp increase in the proportion of unaligned respondents who state that they dislike Biden, and a 7.1pp decrease in the proportion of unaligned respondents stating intent to vote for Biden, comparing to the untargeted scenario. The remaining differences are smaller and insignificant.

When I include respondents who had voted at the time of the survey, the estimated CATE of targeting on Biden dislike among unaligned voters is smaller ($\beta = -0.054$, $s.e. = 0.027$), but this result does not remain significant at $\alpha = 0.05$ if the outcome is operationalized in any other way. The estimated CATE of targeting is null for Democrat and Republicans, and the estimated ATE of targeting is null for all outcomes.

All results are checked for robustness to alternative operationalizations of the outcome as a five-point ordinal or continuous scale. Among unaligned voters who had not pre-voted, the effect of targeting on Biden favorability is robust to all checks, but the coefficient on treatment for intent to vote Biden does not remain significant at $\alpha = 0.05$ when operationalized as a logistic regression ($p = 0.08$). Robustness checks, regression tables and coefficient plots are in SI-C.

Decomposing the Effect

Microtargeting works by leveraging the fact that for a given individual, certain messages are more persuasive than others. By learning an individual's probable reaction to different advertisements, it becomes possible to match them with the most effective advertisement. This is the strategy of the design of this experiment.

Note that because each individual is matched with their optimal advertisement, this design allows the proportion of each advertisement shown to vary between stages. If one advertisement is the most effective for a greater proportion of people, then it will be shown more. This means we can decompose our estimator of the average effect of targeting into two parts.

The first part is the change due to improving individual allocations. At the individual level, the treatment "moves" the outcome from the average point on the advertisement response surface to the optimum. I refer to this as the improvement due to "better allocations". The second part is the change due to altering the proportions of the advertisements. If one advertisement is more effective and it is shown to a greater proportion of respondents, then the average outcome will correspondingly increase. I refer to this as "better advertisements" .[11]

This decomposition matters because our notion of microtargeting–showing the right advertisement to the right person–is closer to the effect of "better allocations". A campaign could achieve the effect of "better advertisements" without large amounts of individual data by using focus groups to find their best advertisement and showing it to everyone.

Figure 2 provides insight into this decomposition and its relation to within-respondent heterogeneity. For each respondent in the treatment group, the targeting model calculates their predicted Biden favorability given each of the five advertisements. Each of these individual predicted response surfaces is depicted in Panel A, Figure 2 by a thin line joining five points. The shape of this line determines which advertisement was shown to the respondent; the lowest point corresponds with the advertisement assigned. The lines are colored and grouped according to this assignment. The dot-and-whiskers connected by the thicker lines denote the group averages and 95 percent confidence intervals.[12]

The lines are not flat; this shows that each individual is predicted to respond differently to each of the five advertisements. We also see that no single advertisement is the most effective, nor is the relative effectiveness of each advertisement the same for all people. The dot-and-whiskers emphasize both of these facts, showing that the difference between the optimal advertisement and other advertisements is not marginal, nor are the individuals receiving the same advertisement homogeneous.

[11] As shorthand for "more of the better advertisements".

[12] A descriptive characterization of these groups is explored in SI-B: Characterization.

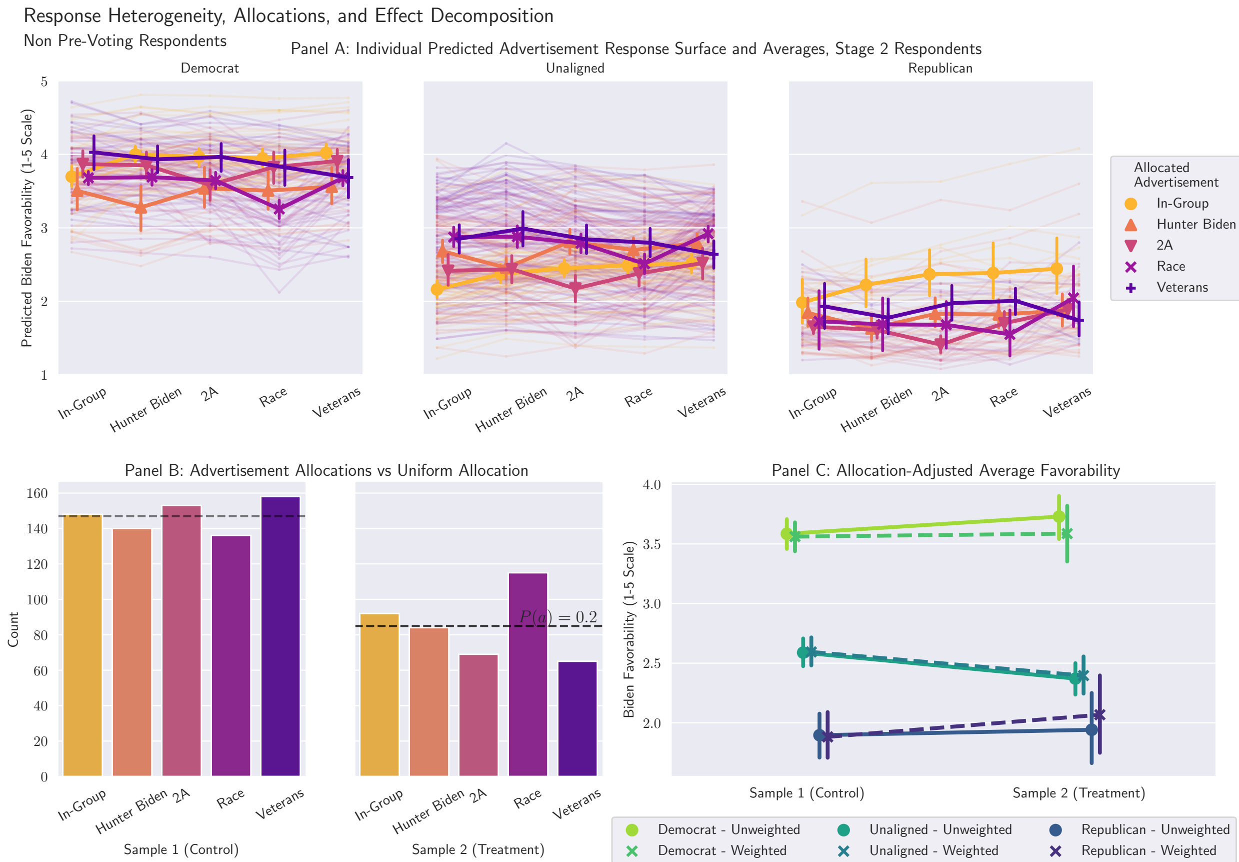

Figure 2: Response Heterogeneity, Advertisement Allocations and Effect Decomposition.

If the effect of targeting were entirely due to "better advertisements," then the model would have assigned everyone the same (best) advertisement. Panel B, Figure 2 shows this is clearly not the case; although the proportion of each advertisement shown varies between samples, it remains remarkably uniform in the targeted group. Therefore at least some of the effect is due to "better allocations".

To visualize the change in outcomes between samples holding the change due to "better advertisements" constant, we weight observations to make the advertisement assignment probability uniform in each stage. The circular markers connected by solid lines in Panel C, Figure 2 show the unweighted average Biden favorability conditional on partisan self identification and sample. The x-shaped markers connected by dotted lines show these averages after weighting each observation by the inverse probability of seeing the advertisement they were shown given the sample in which they were drawn.

There is little difference between the unweighted and weighted conditional means on the left because, as seen in panel B, the advertisement allocation in the first sample is nearly uniform. Notably, the conditional mean for unaligned respondents in the second sample hardly moves, indicating that there was little change in the proportions of each advertisement shown for unaligned respondents between stages. I find that 98 percent of the effect of targeting on unaligned voters who had not already voted is attributable to “better allocations” with the decomposition in SI-B. Thus the effect of microtargeting in this experiment is mostly attributable to optimized matching of dissimilar advertisements to heterogeneous respondents.

Discussion

In sum, this experiment shows that for these five advertisements, optimizing allocation on the basis of respondent traits causes a 7.1pp decrease among unaligned voters reporting that they intend to vote for Biden, and an 8.7pp increase in those reporting Biden dislike.

These magnitudes may have serious implications. Given that this experiment used a convenience sample[13] and does not account for decay (Gerber et al. 2011) a precise estimate of how these effects translate to electoral outcomes is difficult to give. Under moderately conservative assumptions, a 7.1pp decrease in unaligned voters voting for Biden is sufficient to flip the result of the 2020 presidential election in Arizona, where 35.1 percent of voters are registered as unaligned and Biden won by a 0.3 percent margin.[14]

Do these findings contradict prior research? I argue that this may not be the case. Coppock, Hill, and Vavreck (2020) demonstrate in their experiments that there is little variation in the effectiveness of advertisements, but mostly check for heterogeneous treatment effects for different advertisement types. Among respondent characteristics, the authors only test for heterogeneity on partisanship. In contrast, the design employed in this paper leverages response heterogeneity across nine respondent pre-treatment covariates, and searches the space of all possible covariate combinations for the optimal outcome. These results are likewise consistent with Broockman and Kalla (2020), who find that while advertisements about Trump demonstrate “identical, small effects,” “treatments for Biden show large, clear dispersion”.

Future research will look at the extent to which the results in this paper are specific to this setting and these advertisements. Policy researchers can utilize this experimental design to test the effectiveness of interventions and regulation designed to limit the harms of microtargeting, such as clearer disclaimers or civic awareness campaigns. So long as this technology continues to be a part of our elections, understanding its effects is crucial to safeguarding electoral integrity.

[13] Sample representativeness is addressed in SI-A.
[14] The assumptions behind this back-of-the-envelope calculation are explained in SI-B.

References

Broockman, David, and Joshua Kalla. 2020. "When and Why Are Campaigns' Persuasive Effects Small? Evidence from the 2020 US Presidential Election." OSF Preprints. https://doi.org/10.31219/osf.io/m7326.

Burkell, Jacquelyn, and Priscilla M. Regan. 2019. "Voter Preferences, Voter Manipulation, Voter Analytics: Policy Options for Less Surveillance and More Autonomy." *Internet Policy Review* 8 (4).

Cialdini, Robert B. 2007. *Influence: The Psychology of Persuasion*. Vol. 55. Collins New York.

Coppock, Alexander, Seth J Hill, and Lynn Vavreck. 2020. "The Small Effects of Political Advertising Are Small Regardless of Context, Message, Sender, or Receiver: Evidence from 59 Real-Time Randomized Experiments." *Science Advances* 6 (36).

Edelson, Laura, Shikhar Sakhuja, Ratan Dey, and Damon McCoy. 2019. "An Analysis of United States Online Political Advertising Transparency." *arXiv Preprint arXiv:1902.04385*.

Fowler, Erika Franklin, Michael M Franz, Gregory J Martin, Zachary Peskowitz, and Travis N Ridout. 2021. "Political Advertising Online and Offline." *American Political Science Review* 115 (1): 130–49.

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 (1): 135–50.

Hindman, Matthew. 2018. *The Internet Trap*. Princeton University Press. https://doi.org/10.1515/9780691184074.

Kalla, Joshua L, and David E Broockman. 2018. "The Minimal Persuasive Effects of Campaign Contact in General Elections: Evidence from 49 Field Experiments." *American Political Science Review* 112 (1): 148–66.

Krotoszynski, Ronald J., Jr. 2020. "Big Data and the Electoral Process in the United States: Constitutional Constraint and Limited Data Privacy Regulations." In *Big Data, Political Campaigning and the Law: Democracy and Privacty in the Age of Micro-Targeting*, 186–213. Routledge.

Nickerson, David W., and Todd Rogers. 2020. "Campaigns Influence Election Outcomes Less Than You Think." *Science* 369 (6508): 1181–82. https://doi.org/10.1126/science.abb2437.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, et al. 2011. "Scikit-Learn: Machine Learning in Python." *Journal of Machine Learning Research* 12: 2825–30.

Petty, Richard E, and John T Cacioppo. 1986. "The Elaboration Likelihood Model of Persuasion." In *Communication and Persuasion*, 1–24. Springer.

Simon, Felix M. 2019. "'We Power Democracy': Exploring the Promises of the Political Data Analytics Industry." *The Information Society* 35 (3): 158–69.

Wood, Abby K, and Ann M Ravel. 2017. "Fool Me Once: Regulating Fake News and Other Online Advertising." *S. Cal. L. Rev.* 91: 1223.

Zarouali, Brahim, Tom Dobber, Guy De Pauw, and Claes de Vreese. 2020. "Using a Personality-Profiling Algorithm to Investigate Political Microtargeting: Assessing the Persuasion Effects of Personality-Tailored Ads on Social Media." *Communication Research*.

Zuboff, Shoshana. 2015. "Big Other: Surveillance Capitalism and the Prospects of an Information Civilization." *Journal of Information Technology* 30 (1): 75–89.