The Effects of Political Advertising on Facebook and Instagram before the 2020 US Election

Hunt Allcott,[1] Matthew Gentzkow,[1*] Ro'ee Levy,[2]
Adriana Crespo-Tenorio,[3] Natasha Dumas,[3] Winter Mason,[3] Devra Moehler,[3]
Pablo Barberá,[3] Taylor Brown,[3] Juan Carlos Cisneros,[1] Drew Dimmery,[3,4]
Deen Freelon,[5] Sandra González-Bailón,[5] Andrew M. Guess,[6]
Young Mie Kim,[7] David Lazer,[8] Neil Malhotra,[9]
Sameer Nair-Desai,[1] Brendan Nyhan,[10]
Ana Carolina Paixao de Queiroz,[1] Jennifer Pan,[11] Jaime Settle,[12]
Emily Thorson,[13] Rebekah Tromble,[14] Carlos Velasco Rivera,[3]
Benjamin Wittenbrink,[1] Magdalena Wojcieszak,[15,16] Shiqi Yang,[1] Saam Zahedian,[1]
Annie Franco,[3] Chad Kiewiet de Jonge,[3]
Natalie Jomini Stroud,[17] Joshua A. Tucker[18]

[1]Department of Economics, Stanford University, Stanford, CA, US
[2]School of Economics, Tel Aviv University, Tel Aviv, Israel
[3]Meta, Menlo Park, CA, US
[4]Data Science Lab, Hertie School, Berlin, Germany
[5]Annenberg School for Communication, University of Pennsylvania, Philadelphia, PA, US
[6]Department of Politics and School of Public and International Affairs,
Princeton University, Princeton, NJ, US
[7]School of Journalism and Mass Communication, University of Wisconsin–Madison,
Madison, WI, US
[8]NULab of Texts, Maps, and Networks, Northeastern University, Boston, MA, US
[9]Graduate School of Business, Stanford University, Stanford, CA, US
[10]Department of Government, Dartmouth College, Hanover, NH, US
[11]Department of Communication, Stanford University, Stanford, CA, US
[12]Government Department and Data Science, William & Mary, Williamsburg, VA, US
[13]Department of Political Science, Syracuse University, Syracuse, NY, US
[14]School of Media and Public Affairs and Institute for Data, Democracy,
and Politics, George Washington University, Washington, DC, US
[15]Department of Communication, University of California, Davis, CA, US
[16]University of Warsaw, Center for Excellence in Social Science, Warsaw, Poland
[17]Moody College of Communication, University of Texas at Austin, TX, US
[18]Wilf Family Department of Politics and Center for Social Media and Politics,
New York University, New York, NY, US

*To whom correspondence should be addressed; E-mail: gentzkow@stanford.edu.

We study the effects of social media political advertising by randomizing subsets of 36,906 Facebook users and 25,925 Instagram users to have political ads removed from their news feeds for six weeks before the 2020 US presidential election. We show that most presidential ads were targeted toward parties' own supporters and that fundraising ads were most common. On both Facebook and Instagram, we found no detectable effects of removing political ads on political knowledge, polarization, perceived legitimacy of the election, political participation (including campaign contributions), candidate favorability, and turnout. This was true overall and for both Democrats and Republicans separately.

Introduction

Digital advertising plays an increasingly important and controversial role in U.S. politics. Digital ads rose from 2-3 percent of political ad spending in 2016 to 18 percent in 2020, totaling $1.6 billion (*1*). Many commentators have suggested that online ads may have profound impacts on elections (*2–4*). For example, in his announcement that Twitter would not run political ads in 2020, then-CEO Jack Dorsey wrote that "internet advertising ... brings significant risks to politics, where it can be used to influence votes to affect the lives of millions" (*5*).

Moreover, there is significant worry about campaigns' ability to target political ads to specific individuals on social media (*6, 7*). In 2020, over three fourths of Americans said that it is not acceptable for social media companies to use data to target political ads, and over half said that social media companies should not allow political ads at all (*8*). Internal Meta documents described in (*9*) outline potential concerns: "targeted political content can potentially harm people by narrowly delivering divisive appeals to vulnerable audiences; inciting violence; intimidating, discouraging, or misleading voters; creating echo chambers; and decreasing ac-

countability for politicians."

We first present descriptive findings on the reach, targeting, and purpose of ads related to the U.S. 2020 presidential campaign using internal data from Meta. We then present results from a randomized experiment that removed all political ads from users' Facebook or Instagram feeds for the six weeks before Election Day and replaced them with non-political ads. To test concerns specific to targeted political ads, we also implemented a condition that removed and replaced only the subset of political ads on Facebook that were targeted with user data provided by the advertiser ("list-targeted ads"). We estimate effects on political knowledge, polarization, perceived legitimacy of the election, political participation (including campaign contributions), candidate favorability, turnout, and vote choice. Prior literature has proposed mechanisms by which political ads could substantially impact all of these outcomes (*10–12*). More details on our sample, experimental design, outcomes, and data can be found in the Methods Section.

We contribute to a distinguished literature on the effects of political advertising (*10, 11, 13–16*). Many papers report "minimal effects" of political advertising (*17–21*), but others find evidence of meaningful effects both online (*22–25*) and offline (*12, 26–28*). Our paper differs conceptually from most prior work because we remove all or most political ads from users' feeds instead of randomly varying exposure to specific ads across users. Thus, our paper complements previous literature by providing information on the effectiveness of an entire social media campaign in the weeks before the elections. Many hypothesized mechanisms by which ads could affect outcomes—including voter (de-)mobilization, polarization, and perceptions of election legitimacy—concern the aggregate impact of all ads to which voters are exposed (*11*), and our large-sample study can capture these aggregate effects for online ads. In addition, like (*28*), we are able to estimate the effects of a single party's ads by exploiting the fact that partisans saw almost exclusively ads from their own parties. We also go beyond prior work in surveying tens of thousands of participants multiple times, observing a large suite of out-

comes from both surveys and administrative data, and incorporating internal data from Meta that provides insights into the reach and targeting of political ads.

We emphasize several important limitations. First, our estimates are only directly informative about the set of people who agreed to participate in the study. Second, although we include direct measures of turnout and contributions, many of our outcome variables are self-reported. Third, our estimates are specific to removing the last six weeks of social media ads in the US 2020 general election. Existing evidence suggests that political advertising could have larger effects earlier in the election cycle and in elections that are less politically polarized and heavily covered in the media, either in the US (*29*) or in other countries (*24*).

This project is part of the U.S. 2020 Facebook and Instagram Election Study, a collaboration between Meta and academics. The lead academic authors had the final say on the pre-registered analysis plan, had control rights over analysis decisions and the text, and Meta could not block any results from being published. More details of this partnership are available in Suuplementary Notes 9 and 10.

Results

Descriptive Evidence

This section presents descriptive findings on the reach and targeting of political advertising on Facebook and Instagram in the six weeks before the 2020 election. Unless noted otherwise, these findings are based on ad exposure data for our Control group participants. We classify the ad affiliation and its type using two external sources: OpenSecrets (OS) and Wesleyan Media Project (WMP). OS provides a relatively narrow definition of presidential ads based on financial ties (unless stated otherwise, we use this definition). WMP provides a broader definition of presidential ads by predicting whether an ad can be associated with one of the campaigns according to its text. For more details on ad classification see the Methods Section.

Additional figures and tables are presented in Supplementary Note 6.

Ad loads were highly skewed. The top left panel of Figure 1 presents the distribution of political ad load during the study period (in impressions per week) on Facebook and Instagram. We define an "impression" to be a single instance in which an ad appeared in a user's feed; a given user may have multiple impressions of the same ad. The Facebook Control group mean, median, and 99th percentile of political ad impressions were 23, 9.1, and 175 per week; 79 percent averaged at least one impression per week. The Instagram Control group mean, median, and 99th percentile of political ad impressions were 6.2, 1.8, and 71 per week; 57 percent averaged at least one impression per week.

The variation in political ad load is partially predicted by observable characteristics; see Supplementary Figure 2. For example, on both Facebook and Instagram, participants in swing states and participants who voted in 2016 have higher ad loads.

Individual ads were often shown multiple times to the same user: over the full study period, the average Facebook Control group user was shown 132 political ad impressions corresponding to 72 unique ads.

Ads were seen primarily by their own partisans. The top right panel of Figure 1 presents the average political ad load in the Facebook Control group by party identification. The top line presents the overall political ad load, while the remaining four lines present presidential ads favoring each party. Using their broader and narrower definitions, respectively, WMP and OS coded 61 and 20 percent of Control group political ad impressions as presidential. Using the WMP definition, 67% of the presidential partisan ad impressions were Democratic (the corresponding figure among the full U.S. adult Facebook population is 65%).

While individuals saw on average more Democratic ads, there was substantial variation based on the audience: Democratic presidential ads were shown mainly to Democrats, while

Republican presidential ads were shown mainly to Republicans. Both types of ads were most likely to be shown to the strongest partisans. Independents were shown the fewest political ads overall (see Supplementary Figure 3) as well as the fewest presidential ads. Relatedly, Supplementary Figure 2 shows that participants who reported being undecided voters on the baseline survey see fewer political ads than those who reported that their minds were made up. Supplementary Figures 4 and 5 show that almost all ads on Instagram that we can associate with a party were pro-Democratic (based on the subsample of Instagram users for which we have ad load data). This may partly reflect Instagram's younger user base and the tendency of young voters to favor the Democratic party.

Most political ads were list-targeted. About 62 percent of all political ad impressions shown to the Facebook Control group were list-targeted. For Instagram, the share is 71 percent. When focusing specifically on Facebook presidential campaign ads, this share is especially high: 73 and 90 percent of presidential ads were list-targeted according to the presidential ad coding from WMP and OS, respectively. This heavy use of lists suggests that campaigns' targeting of their own partisans may be an intentional strategy.

Most list-targeted ads use multiple targeting strategies. About 54 percent of all political ad impressions shown to the Facebook Control group used customer list custom audiences, 25 percent used website custom audiences, and 20 percent used lookalike audiences; see Supplementary Table 7.

Fundraising ads were most common. The bottom panel of Figure 1 presents the distribution of ad goals for presidential ads coded by WMP separately based on the participant's party identification and whether the ad is pro-Democratic or pro-Republican. Overall, 46 percent of the presidential ad impressions seen by the Facebook Control group are categorized as seeking donations, 26 percent as persuasive, 17 percent as collecting user information, and 5.5 percent

as merchandise sales. Looking across the facets, own partisans primarily saw donation ads, while persuasion was the most common goal of ads shown to supporters of the other party.

The goals of the ads indicate which outcomes are more likely to be affected when ads are removed. Since persuasion and fundraising ads were the most common, we study the effects on attitudes and turnout, and the effects on engagement (including on-platform donations) and participation (including self-reported donations).

Beyond our sample, ads reached diverse ages and genders. Figure 2 shows characteristics of Facebook political ads shown to the full U.S. adult Facebook population (rather than just our study sample) in the 40 days before the election. The first panel shows the distribution across ads of the share of impressions that were seen by male users. Most ads reached fairly mixed audiences by gender, with 74 percent having between 40 and 60 percent impressions to males. For comparison, the vertical line on the figure shows that 44 percent of users are male. However, approximately 15 percent of political ads do appear to be narrowly targeted by gender, with male shares close to either 0 or 100 percent.

The second panel shows the distribution of ads by the standard deviation of age across their impressions. Most ads have a fairly large standard deviation, indicating that campaigns typically do not target a narrowly defined age group. Still, for most ads, the standard deviation across ages is slightly lower than the population standard deviation of 17, indicating that ads are not distributed at random and reach a somewhat narrower age range compared to the overall Facebook user base. This is consistent with Democrats and Republicans typically seeing different ads, as illustrated in Figure 1, and age being correlated with party identification.

Experimental Estimates

Estimating equation. We define Y_i as an outcome, T_i as a vector of treatment group indicators (with two elements for Facebook and one for Instagram), $\boldsymbol{X}_i$ as a vector of controls, and ν_s as a vector of randomization stratum indicators. As pre-specified, the controls $\boldsymbol{X}_i$ are the variables selected in a lasso regression of Y_i on the baseline value of Y_i (if available) and a vector of demographics and baseline survey variables. Controlling for these predictors both improves precision and reduces bias due to naturally occurring imbalances. The estimating equation is

$$Y_i = \tau T_i + \rho \boldsymbol{X}_i + \nu_s + \varepsilon_i, \tag{1}$$

where the coefficient τ captures the average treatment effect of ad removal. Supplementary Note 3 provides additional detail on control variables and variable construction.

Effects on ad load. The All Ad Removal groups mechanically saw zero political ads over the study period. As described above, 62 percent of political ads and 90 percent of presidential ads in the Control group would have been removed in the List-Targeted Ad Removal treatment. Some of these ads were replaced by other (non-list-targeted) political ads. The average Facebook user in the List-Targeted Ad Removal group saw 10 political ad impressions and 0.55 presidential ad impressions per week over the study period, compared to 23 political impressions and 4.7 presidential impressions in the Control group.

Effects on primary outcomes. We study the effects of ads on nine pre-specified primary outcomes: knowledge, affective polarization, issue polarization, perceived legitimacy, participation, engagement, Trump favorability, turnout, and Trump vote. *Knowledge* is the participant's score on factual questions related to the election, news, or general facts; *affective polarization* is the difference in attitudes toward the participant's own party versus the other party; *issue po-*

larization is the extent to which the participant holds opinions closely aligned with their party; *perceived legitimacy* is agreement with various statements implying that elections are free and fair; *participation* is self-reported political participation, including attending protests or contributing to candidates; *engagement* is engagement with political content on the focal platform, measured using platform data; *Trump favorability* is the sum of Trump's approval rating and the difference between Trump's and Biden's thermometer ratings; *turnout* is whether the participant reported voting; and *Trump vote* is defined as 1 for participants who voted for Trump, -1 for those who voted for Biden; and 0 otherwise. See the Methods Section and Supplementary Note 3 for the precise definitions of these variables.

Figure 3 presents the average treatment effects of ad removal on these outcome variables. The first seven variables are standardized into units of standard deviations within the Control groups. The bottom two variables (*turnout* and *Trump vote*) are reported in their original units.

We estimate statistically zero effects on all primary outcomes, with tight confidence intervals. None of the effects in Figure 3 are statistically significant at the 5-percent level after adjusting for multiple hypothesis testing; see Supplementary Tables 8, 9, and 10. The 95-percent confidence intervals rule out effects greater than 0.043 standard deviations for *knowledge*, *affective polarization* and *issue polarization*, *perceived legitimacy*, and *participation*. The confidence interval is even tighter for *Trump favorability*, ruling out effects greater than 0.024 standard deviations. The confidence intervals rule out turnout effects of greater than 0.0069 percentage points, and effects on *Trump vote* for the Facebook All Ad Removal and Instagram treatments of more than 0.027 percentage points.

Many of our null results are precise enough to rule out substantively meaningful effects. For comparison, college graduates in the Control groups have 0.67 standard deviations higher *knowledge* than non-college graduates, and (*30*) estimate that affective polarization has grown by an average of 0.021 standard deviations *per year* since 1978. We show below that our effects

on vote choice are substantially more precise than previous estimates in the literature.

The effects on two primary outcomes are marginally significant after adjusting for multiple hypothesis testing. First, the point estimate implies that All Ad Removal on Facebook may have slightly increased *engagement* by 0.02 standard deviations ($p = 0.01$, $q = 0.158$). This result could suggest that removing political ads from Facebook, while having the direct effect of removing some political content, may *increase* overall engagement with (non-advertising) political content on the platform. To unpack this finding, Supplementary Figure 6 shows effects on individual components of the *engagement* index; most estimates are positive, with particularly strong effects on engagement with candidates running for office. These effects on online engagement could arise because political ad removal increases time on platform (regardless of how people spent their time) or because political ad removal shifts online behavior (regardless of time on platform). Supplementary Figure 7 provides some support for both channels: removing political ads increases civic content views more than total content views and may increase time on platform. As our point estimate is only marginally significant after adjusting for multiple hypothesis testing, it is also possible that the removal of political ads on Facebook did not causally affect engagement.

Second, the point estimate implies that List-Targeted Ad Removal on Facebook may have increased *Trump vote* by 0.027 units ($p = 0.0069$, $q = 0.158$). However, several other results suggest that this may be idiosyncratic rather than a reflection of a true causal effect. First, we do not find that All Ad Removal on Facebook affected *Trump vote* (see Figure 3 and Supplementary Table 11), and most presidential ads were list-targeted. Second, any effects of Facebook List-Targeted Ad Removal seem to be driven by users with *below*-median predicted ad load (see Supplementary Figure 8), whereas we might expect users with lower ad load to have smaller treatment effects. Finally, we do not find any detectable effect of political ad removal on outcomes that one would expect would move in the same direction as vote choice: *Trump*

favorability and secondary outcomes Republican vote share in state-level races, pro-Republican affect, and pro-Republican issue positions (see Supplementary Table 12).

Secondary outcomes, including administrative data. In addition to primary outcomes, we also pre-specified a set of 40 secondary outcomes, including individual components of the primary outcome composite variables as well as administrative data on campaign contributions and voter turnout. While a few of the effects are individually statistically significant (as would be expected to occur by chance), none of the effects are significant after adjusting for multiple hypothesis testing; see Supplementary Note 7.2. For both outcomes where we have administrative data (campaign contributions and voter turnout), the estimated effects are not statistically different from the estimates using self-reported data. In Supplementary Note 7.3, we also analyze a set of pre-specified auxiliary outcomes.

Given the large share of ads focused on donations, effects on campaign contributions are of particular interest. We find no statistically significant effect of ad removal on contributions, with 95 percent confidence intervals (-7.7, 2.3) and (-2.5, 0.78) in units of dollars in the survey and administrative data, respectively. For example, we can rule out that seeing political ads over the study period increased contributions by more than $2.5 per person in the administrative data. In a non-pre-registered analysis in the Facebook sample, we also estimate effects on the *probability* of making campaign contributions. Compared to the continuous campaign contribution variable, this binary variable better captures possible treatment effects on small donations that might be induced by social media ads. We find no statistically significant effect of ad removal on the binary outcome, with 95 percent confidence intervals of (-0.016, 0.004) and (-0.0049, 0.0022) in the survey and administrative data, respectively. Thus, we can rule out that seeing political ads over the study period increased the probability of making a contribution by more than 0.49 percentage points in the administrative data. These insignificant effects may sug-

gest that few participants made contributions via donation ads, that those who did would have contributed through other channels in the absence of the ads, and/or that the effects on contributions were non-zero but smaller than we can detect even in our large sample. As we show below, however, our confidence intervals include contribution effects that would be consistent with a positive return to ad spending.

Subgroup analysis. As pre-specified, we estimate heterogeneous effects using four primary moderators: (i) political party, (ii) an indicator for undecided voters, (iii) an indicator for voters who identify as Black or Hispanic, and (iv) above-median predicted political ad load. The last of these is designed to focus on users who would have seen particularly high volumes of political ads absent our treatment, and who would therefore be most likely to exhibit large treatment effects. We define predicted political ad load based on data from our Control group, where we run a lasso regression of ad load in our study period on a set of predictors including baseline political ad load.

Supplementary Note 8 presents the subgroup analyses. Even among those above the median or 75th percentile of predicted political ad load, we do not find that the treatments had a significant effect on any primary outcome other than *engagement*. There is no statistically significant heterogeneity on any primary outcome by the other three primary moderators.

We also analyze heterogeneous effects by the predicted number of pro-Democratic ads (this moderator was not pre-specified). We test whether participants who would have seen more Democratic ads and were not exposed to these ads because of the treatment, were less likely to vote or support Trump, compared to participants who would have seen more Republican ads. We do not find evidence for such an effect.

Advertising effectiveness. To provide more context on the precision of the null effects on vote choice, turnout, and contributions, we compute the implied returns to advertising spend-

ing by Democratic and Republican presidential advertisers. Our experiment does not directly identify the returns for a single campaign or party because our treatments removed all ads simultaneously. However, Figure 1 showed that people mostly saw ads from their own party. This means that the effect of political ad removal on Democrats was similar to the effect of removing only Democratic ads, and the effect on Republicans was similar to the effect of removing only Republican ads.

To account for the fact that targeting to own-partisans was not perfect, we assume that ads from a person's own party and ads from the opposing party have equal and opposite effects. (If we instead assumed that the effect of own-party ads was greater than opposite-party ads, our estimates of returns to political ad spending would be closer to zero.) We also conservatively assume that non-presidential ads have no effect on our outcomes. We define τ to be the treatment effect of Facebook All Ad Removal among users from a particular party, Δ to be the incremental pro-party presidential ad impressions (i.e., pro-party impressions minus opposite party impressions) seen by users from that party in the Control group, and p to be the ad cost per impression. The treatment effect of ad removal per dollar of incremental pro-party ad spending is then $\tau/(p\Delta)$. We assume that p is equal to $\$25.10$ per thousand impressions, the average costs of the ads in our WMP sample, calculated using the cost data provided in the Meta Ad Library (to calculate costs we impute values at the midpoints of ranges provided by the Ad Library).

Table 1 presents the results. Panel (a) calculates Δ, the Control group incremental pro-party ad impressions, separately using the OS and WMP ad definitions. The first and second columns, respectively, show results for Democratic and Republican users (including independents who "lean" toward the respective party). For example, using the WMP presidential ad sample, the average Democratic user saw 74 more pro-Democrat presidential ads than pro-Republican presidential ads over the study period. The third column pools Democrats and Republicans to calculate incremental pro-party ad impressions across both parties.

Panel (b) computes the cost-effectiveness point estimates and standard errors for net presidential votes, turnout, and contributions. Turnout and contributions both use administrative data instead of self-reports. All outcomes are signed so positive effects benefit the party corresponding to the respective column. For example, Facebook All Ad Removal increased Democratic users' net votes for Biden (i.e., decreased *Trump vote*) by a point estimate of $\tau = 0.023$. Dividing that effect by the $\Delta p = \$1.86$ of additional pro-Democrat presidential ad spending (using the WMP coding) seen by the Control group implies that each Pro-Democratic incremental ad spending increased the Democratic vote by a point estimate of 0.012 votes per dollar.

Using the broader WMP presidential ad coding, and focusing on the pooled estimates, the 95-percent confidence intervals (based on the coefficients and standard errors reported in the table) reject that $1000 of incremental pro-party spending yielded more than 30 net votes, 8.7 additional voters, and $2,208 of additional contributions. As a point of comparison for the net vote effects, (*28*) estimate a cost per vote of $170 for television ads—i.e., a return of 6 votes per $1000. Thus, we cannot rule out reasonably positive returns to advertising. Supplementary Note 7.4 presents additional tables analyzing electoral effects, including focusing only on participants with above-median predicted ad load and analyzing self-reported outcomes instead of validated outcomes. The conclusions are similar: none of the results are statistically significant, but we cannot rule out positive returns for advertising.

Discussion

This paper provides experimental and descriptive evidence on the reach, targeting, goals, and impacts of political advertising on Facebook and Instagram in the six weeks before the U.S. 2020 presidential election.

Our descriptive results show that political advertising on Facebook and Instagram is not primarily used to target undecided voters with persuasive messaging. Instead, most social media

ads seek campaign contributions and voter information, not mobilization or persuasion, and the overwhelming majority of ads are seen by a party's own supporters. These patterns hold both for all presidential ads and for list-targeted presidential ads.

Our experimental results show that removing Facebook and Instagram political advertising had no detectable effects on outcomes including candidate favorability, turnout, participation (including contributions), political knowledge, polarization, and perceived legitimacy of the election. This is true for both all political ads and the subset of list-targeted ads. It is true even though political ad exposure was substantial—23 ads per week for the average Facebook user over the study period, and 175 ads per week at the 99th percentile. In comparison, (*28*) report that the average American was exposed to 8.4 political ads per week on television in the 60 days before the 2004, 2008, and 2012 elections. Our confidence intervals rule out even small to moderate effect sizes. Our null effects are consistent with the fact that relatively few ads were aimed at persuasion and mobilization. They stand in contrast to arguments in both the popular press (*2–7, 9*) and academic literature (*12, 29, 31–37*) that digital ads are likely to be key drivers of political attitudes and election outcomes.

These results echo some prior experimental literature that has failed to find significant effects of online political ads (*11, 29, 38*). However, the large scale of our experiment means that the precision of our estimates is greater than that in most prior work. We provide the large-scale evidence of the combined impact of social media advertising on outcomes like turnout, polarization, and perceived legitimacy of the election. For the effect on vote choice – arguably the parameter that has received the most emphasis in the literature – Table 2 shows that our estimates of the number of votes per 1,000 impressions are more precise than those in any prior study of which we are aware. Supplementary Table 19 presents a similar comparison focusing on the cost-effectiveness (votes per dollar) of ads. Supplementary Note 7.5 provides more details on the calculations behind these tables.

Our experiment advances the literature in several ways beyond sample size and precision. First, we combine both surveys and linked administrative data from a social media platform in an experimental study of political advertising. This allows us to provide a much broader and more precise picture of the mix of advertising to which participants were exposed than would be possible with surveys alone. Second, we are able to link our treatments to directly measured turnout and contribution outcomes, limiting concerns about bias in survey self-reports. Third, previous experiments have not studied political advertising on Instagram, one of the most popular social media platforms. Finally, our interventions were relatively light-touch. Participants were aware that their feeds may have been modified in some way, but they did not know that this would concern ads in particular and they were not made directly aware of their own treatment status. This may limit concerns about experimenter demand effects.

At least in the 2020 US presidential elections context, our results provide evidence against a number of specific hypotheses that have been prominent in the literature: for example, that political ads are de-mobilizing and would tend to reduce participation and turnout (*47–49*); that political ads can be polarizing and that this is likely to be especially true for digital ads (*31, 32*); that 2020 political ads and political micro-targeting were effective in undermining confidence in the electoral system (*33, 34*); and that "micro-targeted" (or in our case, list-targeted) ads are particularly effective in mobilizing supporters (*12, 29, 35–37*).

While we estimate precise null effects on many outcomes, our confidence intervals are nevertheless consistent with a positive return to advertising spending. At the upper end of our confidence intervals, even the effect on donations alone would exceed what advertisers paid for their ads. Taken together, our results thus suggest that the impact of Facebook and Instagram advertising on aggregate political outcomes was minimal at best, but that even very small effects could have been enough to make the advertising worthwhile for campaigns, given its low cost.

We end by reiterating that our results are subject to important limitations. The estimates

are only directly informative about the sample that selected into our study. They are specific to the last six weeks of a highly-contested US presidential election, a situation where advertising might have more limited effects compared to ads placed earlier or in elections where candidate evaluations are less solidified (*29*). Many of our outcomes can only be measured in surveys, and we cannot rule out that they are affected by measurement error and/or experimenter demand effects. Further improving on these dimensions remains an important goal for future research.

Methods

Ethics

The U.S. 2020 Facebook and Instagram Election Study, which includes this research, was evaluated and approved by the NORC Institutional Review Board (Protocol number 20.08.10, Project number 8870). Academic researchers coordinated with their specific university IRBs to ensure they followed regulations concerning human subject research when analyzing data collected by NORC and Meta, as well as when authoring papers based on the results. Additionally, the research group was provided ethical counsel by the independent company Ethical Resolve to inform the study designs.

All participants in the study provided informed consent and allowed access to data on their activity on the focal platform. Participants were informed that the experiment could involve changing aspects of their on-platform experiences, but they were not informed of the specific treatments for which they would be eligible or that the study was focused on political ads (several companion papers (*39–42*) study treatments unrelated to political advertising).

As a mitigation strategy to minimize unanticipated negative effects, we defined a pre-registered stopping rule, inspired by clinical trials, which would have ended a treatment if we detected that it was generating changes in specific variables relevant to individual welfare that were much larger than expected. The stopping rule conditions were not met and we did not have to stop the

treatment. For more information on ethical considerations see Supplementary Note 10.

Experimental Design

This section provides a high-level overview of the experimental design. Figures 4 and 5 present CONSORT diagrams summarizing the design. Additional details are in Supplementary Note 1.

We ran two parallel experiments, with Facebook and Instagram as the respective "focal platform." For each focal platform, Meta drew a stratified random sample of users who lived in the U.S., were at least 18 years old, and had logged into their account at least once in the past month. Meta placed survey invitations in these users' focal platform feeds from August 31 to September 12, 2020. Participants completed a series of surveys implemented by NORC at the University of Chicago, including baseline (responses collected September 8 - 21) and endline (November 4 - 18) surveys. The questionnaires are included in Supplementary Note 13. Participants received base payments of $5 for completing the baseline survey, $20 for completing the endline survey, and other payments described in Supplementary Note 1.

Just after the baseline survey, participants were randomized to a control group or one of several treatment groups that varied the content they saw on their focal platform. Our paper studies the two treatment conditions that removed political ads: "All Ad Removal" and "List-Targeted Ad Removal." We define "political ads" to be all ads that Meta's classifiers determine to be about social issues, elections, or politics (details here). In the All Ad Removal condition, Meta removed all political ads and replaced them with non-political ads. In the List-Targeted Ad Removal condition, Meta removed the subset of political ads that were targeted based on user data provided by the advertiser and replaced them with other political or non-political ads. More specifically, the List-Targeted Ad Removal condition removed and replaced ads targeted using any of the following three strategies: *customer list custom audiences* (described here), which are lists of people uploaded by the advertiser, *website custom audiences* (described here), which

are sets of people who have visited the advertiser's website, and *lookalike audiences* (described here), which are users that Facebook's internal algorithm estimates to be similar to users in a set of users defined by the advertiser.

In both the All Ad Removal and List-Targeted Ad Removal conditions users' total ad load was unchanged and the ads were replaced with the next non-removed ad assigned to the user based on the platform's algorithm, regardless of the ad's content. These two separate treatment conditions allow us to speak directly to concerns described in the introduction about political ads in general and targeted ads in particular.

In the Facebook experiment, we implemented both treatment conditions. In the Instagram experiment, to conserve on the total number of treatment arms, we implemented only All Ad Removal. In both experiments, the ads were removed starting on September 24th. We had originally planned to continue the experiment into December, but Meta made an unexpected decision (*43*) to stop serving political ads starting on Election Day (November 4th). We refer to September 24th through November 3rd as the "study period." Data collection and analysis were not performed blind to the conditions of the experiments.

Balance and attrition

As described in Supplementary Note 1.3, we stratified randomization on self-reported party identification, race/ethnicity, friend/following count, and an indicator for residence in a swing state. Chance imbalance on some other baseline covariates occurred in our Facebook treatment groups; see Supplementary Table 4. Other treatment conditions from the same randomization that we study in parallel papers were balanced on observables, making it unlikely that this reflects an error in randomization. (*44*) find that after controlling for observables that predict the outcome (which we do), correctly implemented randomizations with chance imbalance are no more likely to generate false hypothesis rejections than those without chance imbalance.

The Instagram treatment groups were balanced on all covariates at randomization. More than 70 percent of the primary analysis sample completed the endline survey. Supplementary Tables 5 and 6 show that our samples remain similarly balanced on covariates after attrition and that attrition rates did not differ significantly across control and treatment groups.

Participants and sample size

On Facebook, 14,643,120 users were invited to the study. Of these, 988,247 clicked the invitation, 193,880 consented to participate, and 75,276 completed the baseline survey. A total of 36,906 were randomized to one of the two ad removal conditions or to the Control group and did not delete their account or request data deletion. This final group is our "primary analysis sample." On Instagram, the analogous numbers are 4,618,628 invites, 531,164 clicks, 135,688 consents, 47,659 baseline completes, and 25,925 participants in the primary analysis sample. We excluded users who deleted their account or requested data deletion from the analysis for ethical reasons and since we could not access their data.

We determined the sample size to provide statistical power to detect small effects (with $beta = 0.8$ and $alpha = 0.05$). For example, our power calculations were designed to detect an effect of 1.5 percentage points on self-reported turnout in the Facebook sample and 1.4 percentage points in the Instagram sample.

Of the participants in our Facebook sample, 74% are white, 41% are male, 52% hold a college degree, 55% identify as Democrats or lean Democrat, and 33% identify as Republicans or lean Republican. In our Instagram sample, 61% are white, 38% are male, 58% hold a college degree, 72% identify as Democrats or lean Democrat, and 20% identify as Republicans or lean Republican.

For all analyses, including both descriptive findings and treatment effect estimation, we weight our samples to be representative of U.S. focal platform (Facebook or Instagram) users

on race/ethnicity, political party, education, baseline account activity, and (on Instagram only) whether the respondent was in a swing state. See Supplementary Note 1.5 for additional details on weighting.

Primary outcomes

We have nine pre-specified primary outcome variables. All primary outcomes other than *engagement* are based on survey responses. Supplementary Note 3 provides precise definitions of these primary outcomes as well as of our 40 secondary outcomes. Some of these outcomes are also defined in other papers in the U.S. 2020 Facebook and Instagram Election Study (*42*).

The variable *knowledge* is the average of standardized scores on three sets of factual questions: (i) *election knowledge* (knowledge of candidates' policy positions); (ii) *news knowledge* (distinguishing recent news events from plausible placebo events that had not happened); and (iii) *fact knowledge* (distinguishing true statements from misinformation that was circulating about topics such as COVID-19 and fraudulent ballots).

Affective polarization is the average of three underlying standardized variables: (i) *political supporters polarization* (the difference in participants' favorability toward people who support their own party versus the other party); (ii) *political candidates polarization* (the difference in participants' favorability toward their own party candidates running for office versus candidates in the other party); and (iii) *party smartness polarization* (the difference in participants' perceived smartness of people in their own party versus people in the other party).

Issue polarization is an index of participants' opinions on eight political issues (immigration, repeal of Obamacare, unemployment benefits, mask requirements, foreign policy, policing, racial justice, and gender relations). The signs of the variables are adjusted so that for each issue, a higher value is closer to the participant's own-party mean. A higher issue polarization value should be interpreted as having opinions more closely aligned with one's party.

Perceived legitimacy is an index of participants' agreement with the following six statements: (i) elections are free from foreign influence, (ii) all adult citizens have equal opportunity to vote, (iii) elections are conducted without fraud, (iv) government does not interfere with journalists, (v) government protects individuals' right to engage in unpopular speech, and (vi) voters are knowledgeable about candidates and issues.

Participation is the sum of indicators for whether a participant reported doing the following six activities: (i) attended a protest or rally, (ii) contributed money to a political candidate or organization, (iii) signed an online petition, (iv) tried to convince someone how to vote (online or in-person), (v) wrote and posted political messages online, and (vi) talked about politics with someone they know.

Engagement is the average of 15 standardized components measuring engagement with political content on the focal platform, in the following seven groups: (i) likes, comments, clicks, reshares, and reactions on content classified as "civic" (i.e., related to politics and social issues) by Meta's Civic classifier, described in Supplementary Note 2, (ii) views and clicks on Meta's Voter Hub (a tool that provided election-related information to voters), (iii) likes, comments, and reshares on content from politicians running for office, and on Facebook only, (iv) indication of interest in going to civic events, (v) clicking on a petition, (vi) donations to civic causes, and (vii) enabling a constituent badge. None of the *engagement* components include engagement with political ads, so political ad removal has no mechanical effect on these outcomes. We note that we had to modify the construction of this variable compared to the pre-analysis plan because some data was not available and because some engagement indicators were always zero in our sample (for details see Supplementary Note 11).

The *Trump favorability* variable is the sum of two standardized components: (i) Trump approval ratings and (ii) the difference between Trump and Biden feeling thermometer ratings.

Turnout is an indicator of whether the participant reported voting.

Trump vote is defined as 1 for people who reported voting for Trump, -1 for people who reported voting for Biden, and 0 for those who did not vote or voted for some other candidate.

Voting records and campaign contribution data

We matched participants to state voting records and public records of campaign donations. Supplementary Notes 1.6 and 1.7 provide details on this matching. These directly measured outcomes are an important supplement to our survey-based outcome variables, as they have different measurement errors (from matching errors instead of self-reporting errors) and are available for matched users regardless of whether they completed the endline survey.

Advertising data

We record all political ads seen by our participants during the study period and match them to metadata from the Meta Ad Library and to internal data on the type of targeting associated with each ad. We match a subset of these ads to additional metadata from two sources: OpenSecrets (OS) and the Wesleyan Media Project (WMP) (*45*).

OS provides data on all ads in the Meta Ad Library whose sponsoring page and/or disclaimer are known organizations that provide financial support to presidential campaigns, including Political Action Committees (PACs) and political parties. This provides a relatively narrow definition of presidential ads based on financial ties. Unless stated otherwise, we use the OS definition of "presidential ads."

WMP provides data on all ads in the Meta Ad Library whose sponsoring pages ran at least one ad containing keywords related to the 2020 presidential campaigns. They predict whether each ad favored Biden or Trump using the ad's text and a machine learning model trained on a sample of ads whose sponsor can be associated with one campaign or the other (as in the OS data). This provides a broader definition of presidential ads based on the ads' content. WMP

also trains a separate model to predict the primary "goal" of the ad from its text, based on a hand-labeled training sample with five categories: (i) donation, (ii) purchase/merchandise, (iii) prompting the user to sign up or provide information, (iv) persuasion, and (v) other.

We also use internal Meta data for the full set of political ads in the Meta Ad Library, including the distribution of exposure to ads by user characteristics and the targeting strategies associated with the ads. Supplementary Note 4 provides more details on the advertising data sources and how they were matched with Meta data.

Pre-analysis plan

The pre-analysis plan was registered on September 23, 2020 and updated on May 8, 2023, the day before endline data collection began. It specified our sample, how weights would be created, the handling of missing data, the outcomes studied, our main specification, and subgroup analyses. The plan also specified that we would use Benjamini-Hochberg sharpened False Discovery Rate (FDR) adjusted q-values (*46*) to control for multiple hypothesis testing, with $q < 0.05$. We did not substantively deviate from the pre-analysis plan. Supplementary Note 11 describes clarifications and minor modifications which are mainly driven by changes in data availability. These changes include focusing on ad exposure in Facebook due to missing Instagram data, a change in the construction of several variables (time use, income, urban-rural, engagement, knowledge) due to data availability and minor changes in the endline survey, not analyzing secondary and auxiliary variables for which data does not exist (additional polarization variables, civic events in the feed, distribution of ads by race/ethnicity), and fixing mistakes in the figure shells.

Acknowledgments

This paper was conducted as part of the *U.S. 2020 Facebook and Instagram Election Study*, a collaboration between a team of researchers at Meta and an independent set of external academic researchers. The team designed the project together (with lead academic authors having final say on the pre-registered analysis plan). Data collection was carried out by Meta and NORC. The pipeline code used to pre-process raw platform data was conducted by Meta and reviewed and approved by the external researchers; both the Meta team and external researchers analyzed the data. The paper was written by the external researchers with feedback from the Meta team, but the lead academic authors had final control rights over text and publication. More details on the collaboration appear in the Competing Interests Section and in Supplementary Note 9.

The Facebook Open Research and Transparency (FORT) team provided substantial support in executing the overall project. We are grateful for support on various aspects of project management from C. Nayak, S. Zahedi, I. Rosenn, L. Ahmad, A. Bhalla, C. Chan, A. Gruen, B. Hillenbrand, D. Li, P. McLeod, and D. Rice; engineering from Y. Chen, S. Chen, T. Lohman, R. Pyke, and Y. Wan; data engineering from S. Chintha, J. Cronin, D. Desai, Y. Kiraly, T. Li, X. Liu, S. Pellakuru, C. Xie, and B. Xiong; data science and research from H. Connolly-Sporing; academic partnerships from R. Mersey, M. Zoorob, L. Harrison, S. Aisiks, Y. Rubinstein and C. Qiao; privacy and legal assessment from K. Benzina, F. Fatigato, J. Hassett, S. Iyengar, P. Mohassel, A. Muzaffar, A. Raghunathan and A. Sun; and content design from C. Bernard, J. Breneman, D. Leto and S. Raj. NORC at the University of Chicago partnered with Meta on this project to conduct the fieldwork with the survey participants and pair the survey data with web tracking data for consented participants in predetermined aggregated forms. We are particularly grateful for the partnership of NORC principal investigator J. M. Dennis and NORC project

director M. Montgomery.

The costs associated with the research (such as participant fees, recruitment, and data collection) were paid by Meta, who collaborated with academics in this project as part of the U.S. 2020 Facebook and Instagram Election Study (for more details see Supplementary Note 9). Ancillary support (for example, research assistants and course buyouts), as applicable, was sourced by academics from (authors' initials in parenthesis): the Democracy Fund (NJS), the European Research Council Starting Grant (EXPO- 756301) (MW), the Guggenheim Foundation (BN), the Hewlett Foundation (NJS, JT), the Hopewell Fund (JT), the John S. and James L. Knight Foundation (DF, DL, NJS, JT, RT), the Charles Koch Foundation (JT), New York University (JT), the Alfred P. Sloan Foundation (MG), Stanford University (HA, MG, NM, JP), the Stanford Institute for Economic Policy Research (MG), the University of Texas at Austin (NJS), and the University of Wisconsin–Madison (YMK). These funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Table 1: Electoral Effects by Party (Facebook)

Panel (a): Control Group Ad Impressions

	(1) Democrats	(2) Republicans	(3) Pooled
Share of users over study period	0.60	0.40	1.00
Political ad impressions per user over study period	139	133	137
Using OpenSecrets definition of presidential ads			
Share coded presidential	0.182	0.243	0.206
Share of presidential coded pro-party	0.927	0.832	0.889
Incremental pro-party presidential impressions	22	21	22
Incremental pro-party presidential spending	$0.542	$0.538	$0.540
Using WMP definition of presidential ads			
Share coded presidential	0.630	0.602	0.619
Share of presidential coded pro-party	0.917	0.657	0.814
Incremental pro-party presidential impressions	74	27	55
Incremental pro-party presidential spending	$1.863	$0.674	$1.391

Panel (b): Cost Effectiveness of Ads

	(1) Democrats Point estimate	(2) Republicans Point estimate	(3) Pooled Point estimate
Net President votes per person seeing all ads vs. no ads	0.023	0.013	0.020
	(0.013)	(0.017)	(0.011)
per $ incremental ad spending (OpenSecrets)	0.041	0.024	0.037
	(0.024)	(0.033)	(0.020)
per $ incremental ad spending (WMP)	0.012	0.019	0.014
	(0.007)	(0.025)	(0.008)
Turnout (percentage points) per person seeing all ads vs. no ads	-0.016	0.002	-0.009
	(0.014)	(0.018)	(0.011)
per $ incremental ad spending (OpenSecrets)	-0.029	0.004	-0.017
	(0.026)	(0.034)	(0.021)
per $ incremental ad spending (WMP)	-0.009	0.003	-0.007
	(0.008)	(0.027)	(0.008)
Contributions ($) per person seeing all ads vs. no ads	3.000	-0.640	0.956
	(1.632)	(1.356)	(1.079)
per $ incremental ad spending (OpenSecrets)	5.432	-1.215	1.763
	(2.956)	(2.573)	(1.991)
per $ incremental ad spending (WMP)	1.610	-0.950	0.687
	(0.876)	(2.011)	(0.776)

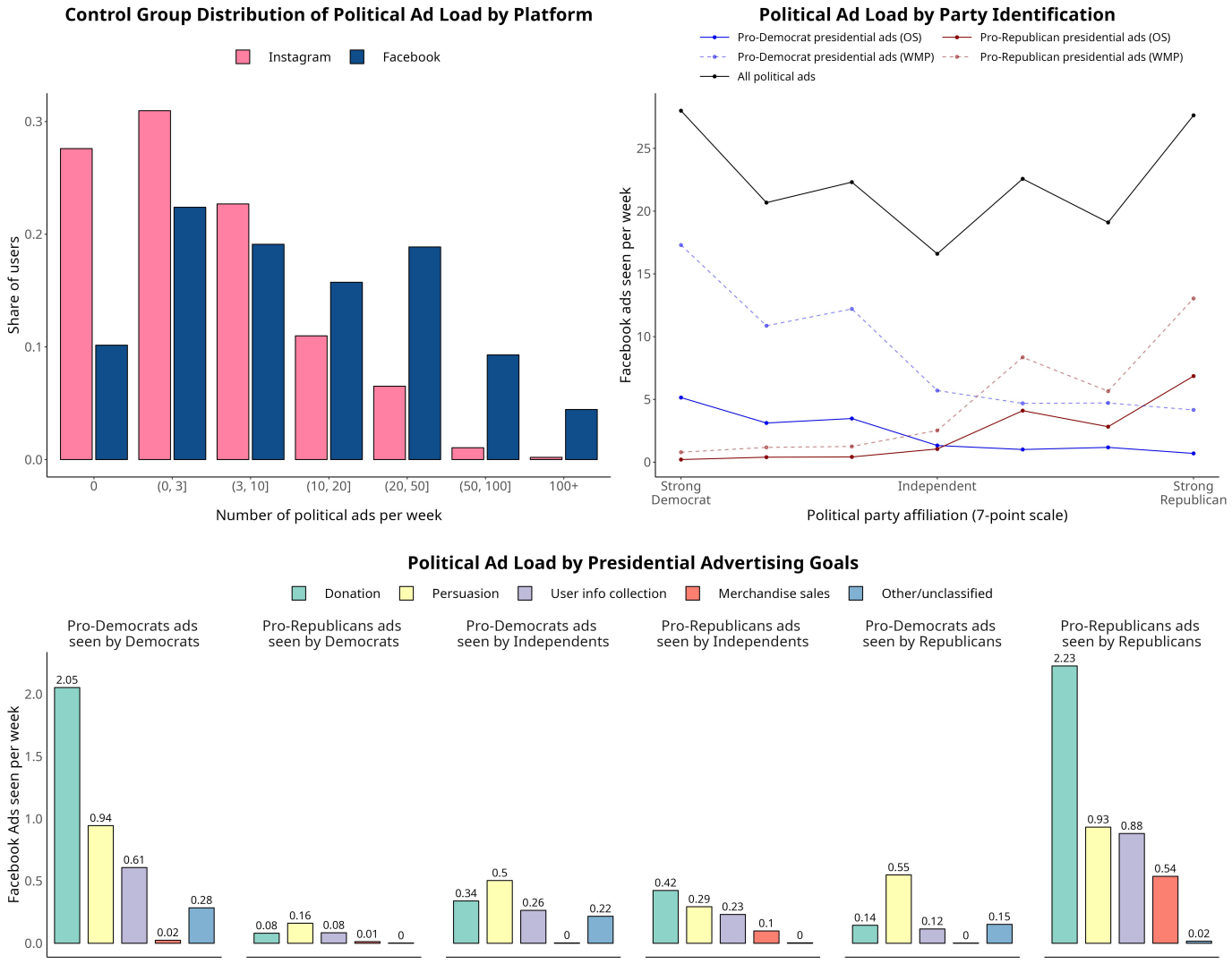
Note: This table computes estimates of returns to political ad spending on Facebook, for outcomes measured in administrative data. The first and second columns present separate results for Democratic and Republican users, including independents who "lean" toward the respective party. The third column presents estimates for the pooled sample. Panel (a) presents data on Control group political ad impressions. Panel (b) presents point estimates and standard errors for the effects of presidential ad spending on net presidential votes, validated turnout, and validated campaign contributions. "Net President votes" refers to the candidate of the respective party in each column (net Democratic votes in the first column and net Republican votes in the second column.) The treatment effect per dollar of incremental ad spending is $\tau/(p\Delta)$, where τ is the treatment effect of ad removal on an outcome, Δ is the treatment effect on net pro-party presidential ad impressions, and the ad cost p is equal to $\$25.10$ per thousand impressions.

Table 2: Comparison to the Literature, Votes per 1,000 Impressions

Paper	(1) Setting	(2) Treatment	(3) Measurement level	(4) N	(5) Votes per 1,000 impressions	(6) Std. error
This paper	2020 U.S. pres. elect., pooled (WMP)	Remove all political ads on FB	Individual	24,546	0.351	0.201
This paper	2020 U.S. pres. elect., pooled (OS)	Remove all political ads on FB	Individual	24,546	0.929	0.502
Broockman and Green (2014)	2012 non-battleground state leg. elect.	Pro-candidate FB ads	Individual	2,984	0.889	0.778
Coppock et al. (2022)	2018 Florida U.S. House elect.	Pro-Democrat ads on FB and IG	Precinct	853	-0.375	7.96
Enríquez et al. (2024)	2018 Mexican municipal elect.	Incumbent performance FB ads	Precinct	13,254	23.087	14.207
Hager (2019)	2016 Berlin state elect.	Pro-CDU FB + Google ads	Postal districts	189	0.918	0.656
Turitto et al. (2014)	2014 Texas Lt. Gov. Rep primary	Pro-candidate digital ad campaign	Individual	5,842	0.701	1.283

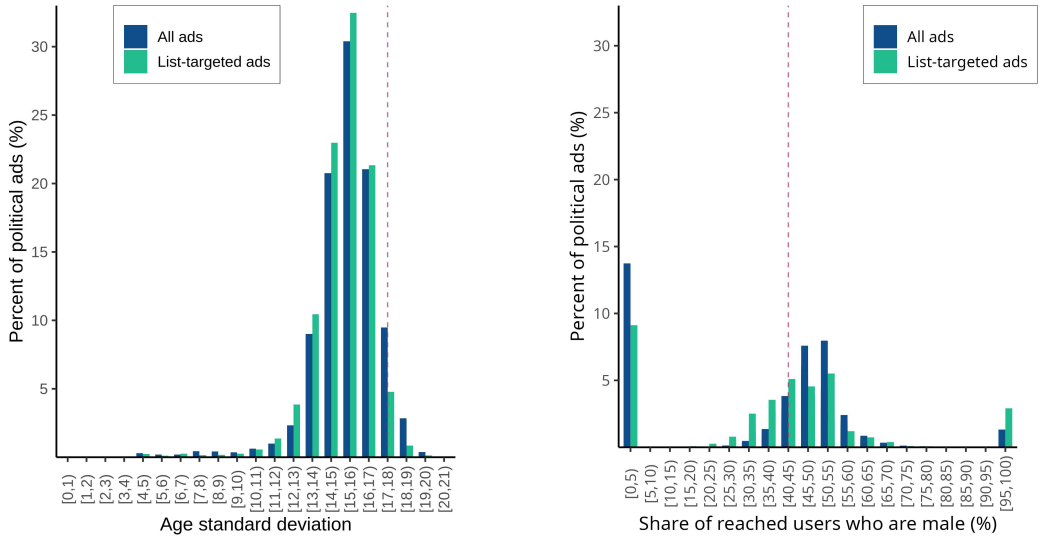
Note: This table compares point estimates and standard errors of the effect of impressions on net votes. The first two rows are based on this paper. They show estimates for pooled Democrats' and Republicans' incremental pro-party impressions, where the party leaning is coded using either WMP or OS. These estimates can be derived directly from Table 2 by dividing the effect of seeing all ads by the number of incremental pro-party impressions and rescaling by 1,000. For example, the number of additional votes per 1,000 impressions based on WMP data is (0.02/55) * 1000 = 0.4. The rest of the rows in the table provide benchmarks from other studies. Supplementary Note 7.5 details how these estimates were computed. Measurement level is the level at which the study observes outcomes. N is the number of units for which the outcome is observed.

Figure 1: Control Group Political Ad Characteristics



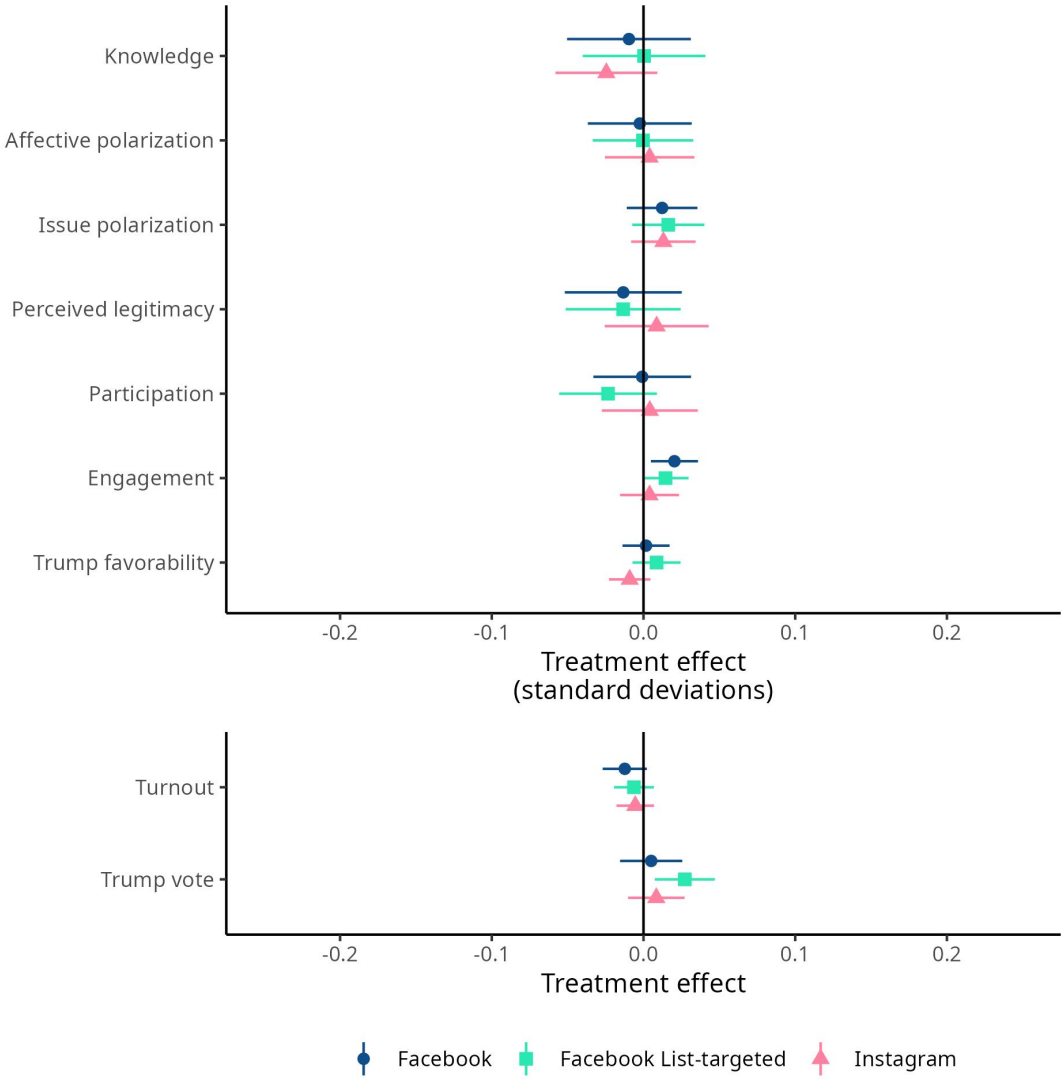
Note: The top left panel presents the distribution of political ad load in the Facebook and Instagram Control groups. The top right panel presents the average political ad load by political party identification in the Facebook Control group, as well as the average presidential ad load favoring each party, as coded by OpenSecrets and the Wesleyan Media Project. The bottom panel presents the distribution of presidential ad goals (as coded by Wesleyan Media Project) shown to the Facebook Control group. The panel shows the distribution separately based on the participant's party identification (with independents who "lean" toward either party included with that party) and whether the ad is pro-Democratic or pro-Republican.

Figure 2: Distribution of User Demographics Across Political Ad Impressions



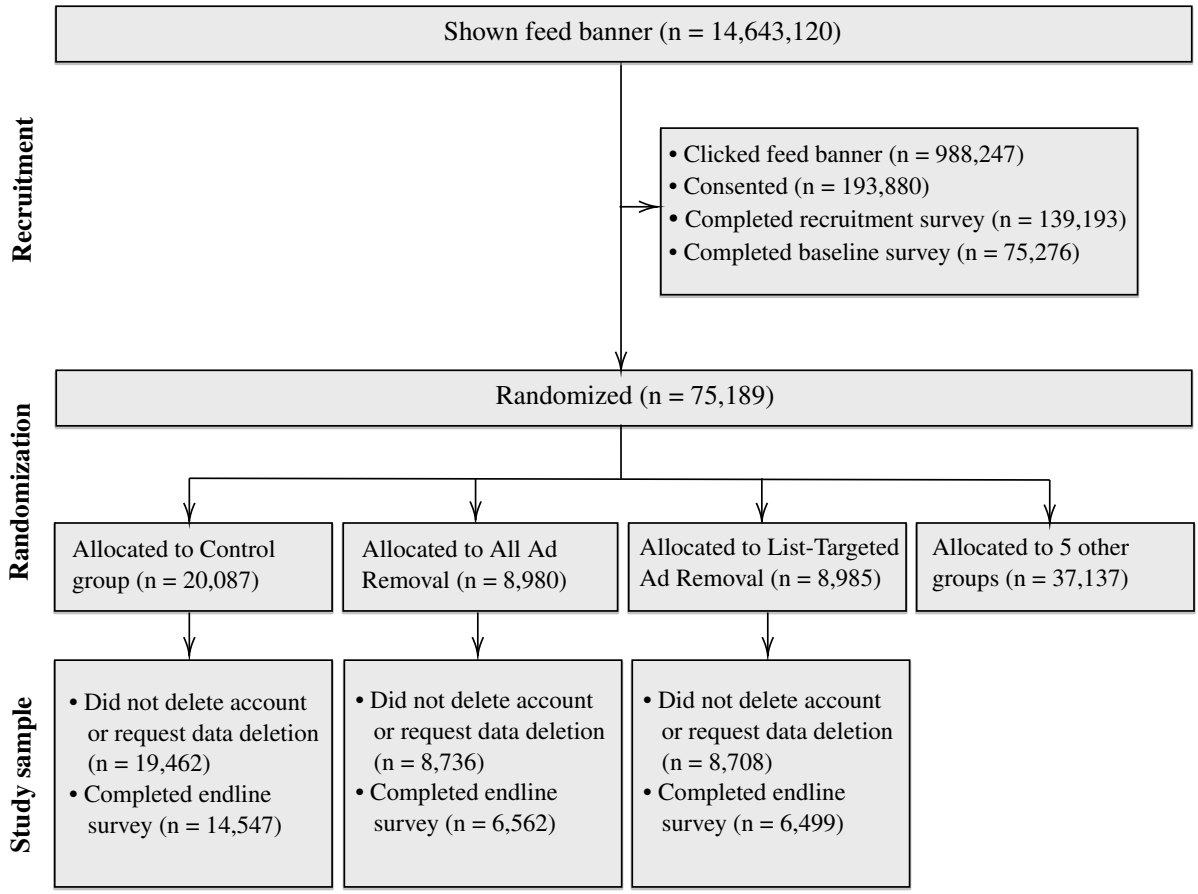
Note: This figure presents distributions of user demographics across political ads shown to U.S. adult Facebook users in the 40 days before the 2020 election. The unit of analysis is an ad, and the observations are weighted by the number of impressions. The dashed red lines display user-level averages across all users who saw at least one political ad, constructed using internal data from Meta.

Figure 3: Effects of Political Ad Removal on Primary Outcomes



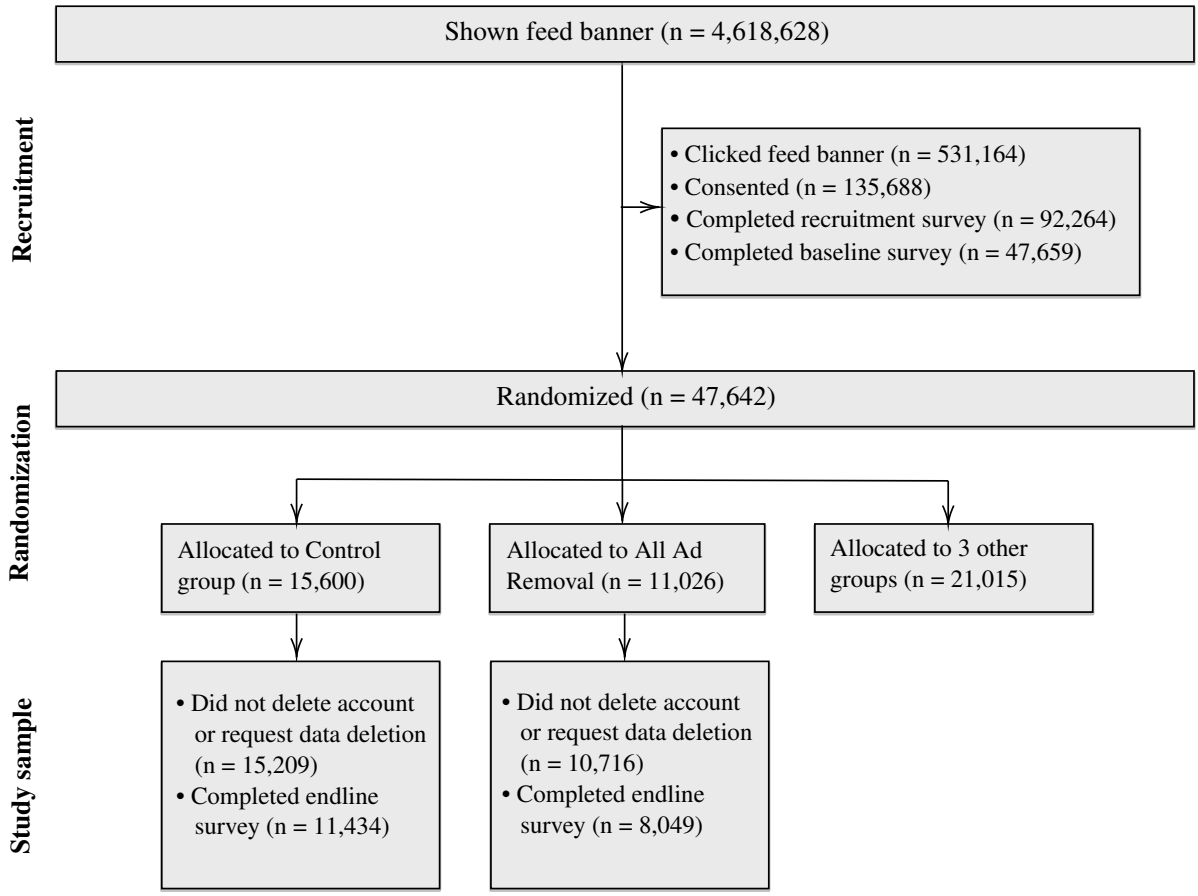
Note: This figure presents average treatment effects of political ad removal estimated using equation (1). Our Facebook sample consists of 36,906 participants, of whom 27,608 completed the endline survey. Our Instagram sample consists of 25,925 participants, of whom 19,483 completed the endline survey. The horizontal lines represent 95 percent confidence intervals. For detailed results see Supplementary Tables 8-10.

Figure 4: CONSORT Flow Diagram Facebook



Note: This figure presents a flow diagram of recruitment, randomization, and sample statistics for Facebook study participants.

Figure 5: CONSORT Flow Diagram Instagram



Note: This figure presents a flow diagram of recruitment, randomization, and sample statistics for Instagram study participants.

References and Notes

1. H. Homonoff, 2020 political ad spending exploded: Did it work?, *Forbes* (2020). Available at https://www.forbes.com/sites/howardhomonoff/2020/12/08/2020-political-ad-spending-exploded-did-it-work/?sh=584a80d73ce0.

2. R. Mac, C. Warzel, Congratulations, Mr. President: Zuckerberg secretly called Trump after the election, *BuzzFeed* (2018). Available at https://www.buzzfeednews.com/article/ryanmac/congratulations-zuckerberg-call-trump-election-2016?bfsplash.

3. M. Coppins, The billion-dollar disinformation campaign to reelect the president, *The Atlantic* (2020). Available at https://www.theatlantic.com/magazine/archive/2020/03/the-2020-disinformation-war/605530/.

4. K. Roose, S. Frenkel, M. Isaac, Don't tilt scales against Trump, Facebook executive warns, *The New York Times* (2020). Available at https://www.nytimes.com/2020/01/07/technology/facebook-trump-2020.html.

5. J. Dorsey. [@jack], We've made the decision to stop all political advertising on Twitter globally [Twitter post], Twitter (30 Oct, 2019). Available at https://twitter.com/jack/status/1189634360472829952?lang=en.

6. A. C. Madrigal, What Facebook did to American democracy, The Atlantic (2017). Available at https://www.theatlantic.com/technology/archive/2017/10/what-facebook-did/542502/.

7. J. Zacharia, Congress should pass Rep. Eshoo's bill restricting microtargeting of political ads, *San Francisco Chronicle* (2020). Available at https://www.sfchronicle.com/opinion/article/Congress-should-pass-Rep-Eshoo-s-bill-15371138.php.

8. B. Auxier, 54% of Americans say social media companies shouldn't allow any political ads, *Pew Research Center* (2020). Available at https://www.pewresearch.org/short-reads/2020/09/24/54-of-americans-say-social-media-companies-shouldnt-allow-any-political-ads/.

9. C. Lima, Facebook knew ads, microtargeting could be exploited by politicians. It accepted the risk., *The Washington Post* (2021). Available at https://www.washingtonpost.com/politics/2021/10/26/facebook-knew-ads-microtargeting-could-be-exploited-by-politicians-it-accepted-risk/.

10. E. F. Fowler, M. M. Franz, G. J. Martin, Z. Peskowitz, T. N. Ridout, Political advertising online and offline, *American Political Science Review* **115**, 130–149 (2021). Publisher: Cambridge University Press.

11. E. F. Fowler, M. M. Franz, T. N. Ridout, *Political advertising in the United States* (Routledge, 2021).

12. K. Haenschen, J. Jennings, Mobilizing millennial voters with targeted internet advertisements: A field experiment, *Political Communication* **36**, 357–375 (2019). Publisher: Taylor & Francis.

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

14. W. L. Bennett, S. Iyengar, A new era of minimal effects? The changing foundations of political communication, *Journal of communication* **58**, 707–731 (2008).

15. G. C. Jacobson, How do campaigns matter?, *Annual Review of Political Science* **18**, 31–47 (2015).

16. M. Aggarwal, *et al.*, A 2 million-person, campaign-wide field experiment shows how digital advertising affects voter turnout, *Nature Human Behaviour* **7**, 332–341 (2023).

17. D. E. Broockman, D. P. Green, Do online advertisements increase political candidates' name recognition or favorability? Evidence from randomized field experiments, *Political Behavior* **36**, 263–289 (2014). Publisher: Springer.

18. D. Broockman, J. Kalla, The manifold effects of partisan media on viewers' beliefs and attitudes: A field experiment with Fox News viewers, *OSF Preprints* (2023).

19. A. Coppock, D. P. Green, E. Porter, Does digital advertising affect vote choice? Evidence from a randomized field experiment, *Research & Politics* **9**, 20531680221076901 (2022). Publisher: SAGE Publications Sage UK: London, England.

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

21. C. Turitto, D. P. Green, B. Stobie, S. Tranter, Testing the persuasive effects of digital media: A cluster randomized field experiment, *Available at SSRN 3537287* (2014).

22. D. Shaw, C. Blunt, B. Seaborn, Testing overall and synergistic campaign effects in a partisan statewide election, *Political Research Quarterly* **71**, 361–379 (2018).

23. A. Hager, Do online ads influence vote choice?, *Political Communication* **36**, 376–393 (2019). Publisher: Taylor & Francis.

24. J. R. Enríquez, H. Larreguy, J. Marshall, A. Simpser, Mass political information on social media: Facebook ads, electorate saturation, and electoral accountability in Mexico, *Journal of the European Economic Association* p. jvae011 (2024).

25. M. M. Franz, T. N. Ridout, Does political advertising persuade?, *Political Behavior* **29**, 465–491 (2007). Publisher: Springer.

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

27. C. Kendall, T. Nannicini, F. Trebbi, How do voters respond to information? Evidence from a randomized campaign, *American Economic Review* **105**, 322–353 (2015).

28. J. L. Spenkuch, D. Toniatti, Political advertising and election results, *The Quarterly Journal of Economics* **133**, 1981–2036 (2018).

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

30. L. Boxell, M. Gentzkow, J. M. Shapiro, Cross-country trends in affective polarization, *The Review of Economics and Statistics* pp. 1–60 (2022).

31. M. Prior, Media and political polarization, *Annual Review of Political Science* **16**, 101–127 (2013).

32. P. Törnberg, How digital media drive affective polarization through partisan sorting, *Proceedings of the National Academy of Sciences* **119**, e2207159119 (2022).

33. N. Berlinski, *et al.*, The effects of unsubstantiated claims of voter fraud on confidence in elections, *Journal of Experimental Political Science* **10**, 34–49 (2023).

34. J. Matthes, *et al.*, Understanding the democratic role of perceived online political microtargeting: longitudinal effects on trust in democracy and political interest, *Journal of Information Technology & Politics* **19**, 435-448 (2022).

35. F. N. Ribeiro, *et al.*, *Proceedings of the conference on fairness, accountability, and transparency* (2019), pp. 140–149.

36. B. Zarouali, T. Dobber, G. De Pauw, C. de Vreese, Using a personality-profiling algorithm to investigate political microtargeting: assessing the persuasion effects of personality-tailored ads on social media, *Communication Research* **49**, 1066–1091 (2022). Publisher: SAGE Publications Sage CA: Los Angeles, CA.

37. F. Liberini, M. Redoano, A. Russo, A. Cuevas, R. Cuevas, Politics in the Facebook era-evidence from the 2016 US presidential elections, *CESifo Working Paper* (2020). Publisher: CESifo Working Paper.

38. A. Coppock, S. J. Hill, L. Vavreck, 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**, eabc4046 (2020). Publisher: American Association for the Advancement of Science.

39. A. M. Guess, *et al.*, How do social media feed algorithms affect attitudes and behavior in an election campaign?, *Science* **381**, 398–404 (2023).

40. B. Nyhan, *et al.*, Like-minded sources on facebook are prevalent but not polarizing, *Nature* **620**, 137–144 (2023).

41. A. M. Guess, *et al.*, Reshares on social media amplify political news but do not detectably affect beliefs or opinions, *Science* **381**, 404–408 (2023).

42. H. Allcott, *et al.*, The effects of facebook and instagram on the 2020 election: A deactivation experiment, *Proceedings of the National Academy of Sciences* **121**, e2321584121 (2024).

43. Facebook, A Look at Facebook and US 2020 Elections (2020). Available at https://about.fb.com/wp-content/uploads/2020/12/US-2020-Elections-Report.pdf.

44. M. Bruhn, D. McKenzie, In pursuit of balance: Randomization in practice in development field experiments, *American Economic Journal: Applied Economics* **1**, 200–232 (2009).

45. M. Neumann, *et al.*, Biden and Trump Mentions in Facebook Advertising from June 1 through Election Day 2020, DOI: 10.25438/wes02.23546064 [dataset].

46. Y. Benjamini, A. M. Krieger, D. Yekutieli, Adaptive linear step-up procedures that control the false discovery rate, *Biometrika* **93**, 491–507 (2006).

47. S. Ansolabehere, S. Iyengar, A. Simon, N. Valentino, Does attack advertising demobilize the electorate?, *American Political Science Review* **88**, 829–838 (1994).

48. S. Ansolabehere, S. Iyengar, *Going Negative: How Political Advertisements Shrink and Polarize the Electorate* (Free Press, 1995).

49. J. D. Clinton, J. S. Lapinski, "Targeted" advertising and voter turnout: An experimental study of the 2000 presidential election, *The Journal of Politics* **66**, 69–96 (2004).