Do Online Ads Influence Vote Choice?

ANSELM HAGER

Do online ads influence vote choice? We partner with a German party to evaluate the effectiveness of online ads using a cluster-randomized experiment. During the 2016 Berlin state election, 189 postal districts were randomly assigned to (a) emotional ads; (b) factual ads; or (c) no ads. Analyzing electoral results at the postal district level, we find that the overall campaign weakly increased the party's vote share by 0.7 percentage points (p-value = 0.155). We also estimate a negative effect of the campaign on the vote share of the party's main competitors of 1.4 percentage points (p-value = 0.094). Turning to the mechanism of persuasion, we find that the factual ads, if anything, fared slightly better than the emotional ads. Our evidence thus provides tentative support that online ads positively affect vote choice.

Keywords   field experiment, online communication, campaigns, persuasion

Do online ads affect vote choice? Recent years have seen a dramatic surge in political advertisement on the Internet. In the United States alone it is estimated that Barack Obama's presidential campaign increased its spending on online ads from $22 million in 2008 to roughly $78 million in 2012 (Federal Election Commission [FEC], 2012). The use of Facebook ads during the 2016 U.S. presidential election, in particular, sparked a heated debate among academics, policymakers, and the public. Whether and how online ads influence voters, however, remains poorly understood. Studies assessing the causal effect of online ads on vote choice—using random assignment and administrative outcomes—are, to our knowledge, nonexistent.[1]

This article presents evidence from a cluster-randomized online field experiment. We were approached by the German Christian Democratic Union (henceforth, CDU) and asked to evaluate its online campaign during the 2016 Berlin state election. To estimate the campaign's effect on the party's vote share, we randomly assigned Berlin's 189 postal districts to three treatment conditions. First, in an *emotional* condition voters were exposed to ads designed to make voters proud of the party (Williams & DeSteno, 2008). Second, in a *factual* condition voters were exposed to ads designed to

Anselm Hager is Assistant Professor of Political Economy, Department of Politics and Public Administration, University of Konstanz, 78457 Konstanz, Germany.
Address correspondence to Anselm Hager, E-mail: afr2132@columbia.edu

communicate straightforward facts without activating specific emotions. Third, in a *control* condition no ads were shown. Ads were distributed via Facebook and Google, generating 17 million impressions.

Analyzing electoral outcomes aggregated at the postal district level, we find that the ads weakly raised the party's vote share by 0.7 percentage points (*p*-value of 0.155). We also find a negative effect of the campaign on the vote share of the party's main competitors of 1.4 percentage points (*p*-value = 0.094).[2] Our findings thus provide tentative evidence that the campaign positively affected the party's vote share. However, low statistical power coupled with the clustered nature of the treatment prohibit a firm conclusion.

Turning to the mechanisms of persuasion, we find no significant differences between the factual and the emotional campaigns. Specifically, postal districts in the factual condition see the party's vote share rise by 0.8 percentage points (*p*-value of 0.098). By contrast, the emotional campaign raised the vote share by 0.6 percentage points (*p*-value of 0.338). While the estimates are imprecise, the finding raises questions about a prevailing consensus among scholars of political behavior, namely, that emotions are central to how political advertising works (e.g., Bennett & Iyengar, 2008; Huddy & Gunnthorsdottir, 2000; Neuman, 2007; Panagopoulos, 2010).

Our article also adds to a literature on campaign finance and cost-effectiveness (Arceneaux & Nickerson, 2009; Green, Gerber, & Nickerson, 2003). Analyzing total votes won and taking our estimates at face value, we estimate that the cost per vote was 3 to 7 euros, which compares favorably to other media, including television and radio. This estimate also falls in line with one observational study from Europe. Specifically, Bekkouche and Cagé (2018) estimate that the price of a vote in French legislative elections is 6 euros.

The remainder of this article is organized as follows. In Section 1, we introduce the theoretical background of the study and develop predictions about the effects of the factual and emotional campaign ads. Section 2 introduces the data and design. In Section 3, we present our results and robustness checks. Section 4 presents additional analyses pertaining to turnout and persuasion. Section 5 concludes.

Background

Since the 1930s, political scientists have studied whether mass media ads can influence vote choice (Hovland, Lumsdaine, & Sheffield, 1949; Lazarsfeld, Berelson, & Gaudet, 1944). Early studies focused on newspaper ads and leaflets (e.g., Hartmann, 1936). Beginning in the 1970s, the advent of television marketing led to a surge in scholarly interest in this form of advertising. Empirical studies, however, have generated inconclusive results regarding the effectiveness of mass media (e.g., Kahn & Geer, 1994; Kaid & Boydston, 1987). A particular complication is that most empirical work relies on observational data, which makes the estimation of causal effects difficult (Arceneaux, 2010).

Within the small set of studies that randomly assign units to mass media ads and assess changes in administrative outcomes, the evidence is more favorable. In one pioneering study, Green and Vavreck (2007) randomly assign cable TV markets to advertisements intended to encourage young adults to turn out during the 2004 presidential election. Using public records of voter turnout, the authors find that the campaign produced a two-to-three-point increase in voter turnout (see also Panagopoulos & Green,

2008). Evidence regarding the effectiveness of online ads, however, is scarce. Indeed, we are only aware of four studies (cited earlier), which randomly distributed online ads. Studies using administrative electoral outcomes are, to our knowledge, absent.

Besides measuring the effectiveness of mass media ads, scholars of public opinion, early on, began to explore the mechanisms through which ads affect political behavior. One classic question expressed early on by Lazarsfeld and colleagues (1944) concerns the degree to which political advertising needs to activate emotions in order to move the electoral needle. It remains debated, however, *how* emotions affect voting behavior. Among the many theoretical accounts, four are particularly influential.

A first theory, put forth by Brader (2006, pp. 32–35), contends that the emotional content of ads—for example, pride—can rub off on individuals and thus affect voting behavior (known as the *affect-transfer hypothesis*). A second alternative conjecture— articulated in Roseman, Abelson, and Ewing (1986)—states that political ads work best when emotions coincide with voters' current state of mind. A third influential theory, known as *affective intelligence* (Marcus, Neuman, & MacKuen, 2000; Ryan, 2012), posits that emotions can serve as a trigger by activating a temporary state of arousal, which guides subsequent cognition.

A fourth and final theory, known as the elaboration likelihood model (ELM; Petty & Cacioppo, 1986), specifies two routes through which emotions affect voting behavior. In a central route, ad recipients are highly motivated and capable; attitude changes are the result of an in-depth rational assessment of arguments. In a peripheral route, ad recipients are less motivated and changes in attitudes are caused by affect—for example, whether a recipient likes the source of a message. Importantly, attitude change brought about via the central route is lasting and thus more likely to lead to sustained changes in behavior (Petty & Cacioppo, 1986).

While theoretical accounts about the role of emotions abound, empirical evidence— particulaly in the realm of online ads—is scarce. To fill this gap, we cooperated with the German CDU and implemented an online experiment during the 2016 Berlin state election. The party's online campaign relied on two strategies of political persuasion. In an "emotional" route, ads were designed to activate one particular emotion: pride.[3] By contrast, in a "factual" route, ads were designed to abstract from emotions. Rather, they communicated straightforward political information (more details later). The two treatments thus neatly map onto the ELM. The factual treatment presents straightforward political information, which allows voters to rationally assess political arguments before casting their ballot. By contrast, the emotional treatment was intended to make ad recipients feel proud. Pride, then, should lead to positive affect and thus, in theory, affect voting behavior.

While the precise mechanisms that link emotions to behavior are beyond the scope of this study, should one expect the emotional or the factual campaign to more effectively persuade voters? Early field experiments adopting similar designs provide evidence that emotional appeals perform substantially better. Hartmann (1936), for instance, finds that emotional leaflets—those with imagery intended to elicit positive emotional reactions—fare better than leaflets with straightforward political facts. More recently, Huddy and Gunnthorsdottir (2000) show that emotional visual imagery can increase positive views about specific policies. Finally, Brader (2005) finds that campaign ads can use images and music to successfully manipulate emotions, thereby affecting the behavior of voters. Given growing scholarly evidence and sentiments among campaign staff, we therefore formulated the following two hypotheses in the EGAP pre-registration document:[4]

H1: The overall campaign raises the vote share of the party.

H2: The emotional campaign generates a higher vote share.

Design

Population and Data

The population of the study is the city of Berlin, which has 3.6 million inhabitants and 2.5 million eligible voters. Berlin has 193 unique postal districts, which we obtained from the Berlin Statistical Office. We focus on the postal district level as it is the lowest geographic unit to which Facebook and Google permit the targeting of ads. We excluded four postal districts from the study as they contain fewer than 2,000 residents. The resulting 189 districts have an average of 19,327 residents, which ensured that a sufficient number of individuals could be treated. The postal districts are our primary unit of analysis.

The descriptive statistics of the overall sample, the two treatment groups and the control group, are given in Table A1 in the online supplemental Appendix. These variables, which are used in the analyses to control for pretreatment differences, were preregistered and provided to us by the Berlin Statistical Office. They include all standard items, which the offices regularly updates. The table also presents the electoral results from the 2011 election. The state parliament of Berlin, much like the German federal parliament (*Bundestag*), is elected on the basis of a mixed-member proportional electoral system. The first vote is for the local Member of Parliament (MP) who is elected using a plurality voting system. The second vote is for electoral party lists. The parliament is then filled with candidates, which won their electoral districts by first vote and candidates of the electoral lists in accordance with a given party's vote share in second votes.

Electoral results in Berlin are reported on the level of the voting district (*Stimmbezirk*), of which there are more than 2,200. Most of the voting districts lie within one postal district. In some cases, however, voting districts stretch two postal districts. This made it necessary to aggregate the electoral results at the postal district level. To do so, we followed a preregistered protocol, which we discuss in Section A.1 in the online supplemental Appendix.

Randomization

The 189 postal districts were randomly assigned to one of three treatment arms. First, 63 districts were randomly assigned to a pure control group, which did not receive any online ads. Second, 63 districts were randomly assigned to the emotional treatment. Third, 63 districts were randomly assigned to the factual treatment. As is laid out in the preregistration document, we used block randomization in order to minimize pretreatment imbalances of predictive variables.

Two blocking variables are particularly noteworthy: Google reachability and population size. The rationale for blocking on these variables was as follows. As opposed to Facebook, Google has imperfect information about individuals' home addresses.[5] Such data, however, are crucial in order to target ads to individuals that live and vote in a given postal district. We therefore reached out to Google before commencing the study to obtain a list of postal districts to which the company can target ads with certainty, which

were 114 out of 189 postal districts. Since we wanted to keep the exposure to Google ads constant across the factual and emotional treatment arms, we blocked on Google reachability. The decision to block on population size followed similar concerns: We wanted to ensure that small districts, where few individuals can be targeted, are similarly distributed across the experimental conditions since it otherwise threatens to bias difference-in-means estimation (Middleton, 2008).

Overall, online supplemental Appendix Table A1 demonstrates that block randomization produced excellent balance across the pre-registered control variables. A map of the postal districts and their respective treatment statuses is given in Figure A1 in the online supplemental Appendix. It similarly showcases geographic balance across treatment and control districts. Finally, in Table A2 in the online supplemental Appendix we estimate three models where we regress the treatment indicators (emotional, factual, and combined) on the pretreatment control variables. The results show that no demographic or electoral variables meaningfully predict treatment assignment. The F statistics of the three models, assessing joint significance, are all insignificant.

Treatment

Starting on September 9, 2016, all of the 126 treatment districts were exposed to online ads displayed via Facebook and Google. As explained, Google was only able to distribute ads in 76 of the 126 treatment districts due to technical limitations. This "failure to treat," however, was known ex ante and addressed, as discussed, by blocking on Google reachability. The funds for the campaign, which represent the treatment intensity, were evenly split across all districts proportional to their population size and taking into consideration whether Google is present. In addition, the funds were evenly split between video and banner ads.[6]

Importantly, the banners, but not the videos, were split into factual and emotional variants. This decision was made because we expected the bulk of online ads to be banners as they are typically—depending on ad demand and the provider—less expensive. In addition, designing emotional and factual variants of the campaign videos was not possible due to time constraints.[7]

The timing of the campaign was as follows. On days nine to seven before the election, the party distributed videos and banners relating to education. As can be seen in Figure A2 in the online supplemental Appendix, the factual campaign communicated key political facts on a plain orange background, namely, that the party will keep the *Gymnasium* school in tact—an advanced secondary school that is focused on preparing students to enter university. Compare this to the emotional banner, which shows the picture of a proud elementary school student, and adds the non-descriptive statement "safe learning." On days seven to five before the election, the party distributed videos and banners relating to the economy (Figure A3 in the online supplemental Appendix). Here, again, the factual banner stated matter-of-factly that the party "created 138,000 new jobs," while the emotional campaign showed the image of a large construction site in Berlin adding the non-descriptive statement "strong economy." On days five to three before the election, the party distributed videos and banners relating to public safety (Figure A4 in the online supplemental Appendix). The banners, again, were either factual —claiming "zero tolerance" against criminals—or emotional, showcasing the photo of joyful elderly citizens and the statement "safe traveling." On days three to the day of the election, the party distributed its main campaign video and banners communicating

a "call to action." Again, the banners were split into factual and emotional variants (see Figure A5 in the online supplemental Appendix).

Before moving to the empirical analysis, we want to briefly reflect on the process of designing the factual and emotional campaign. There is little doubt that broad terms like "emotional" and "factual" mean different things to different people. The two campaigns were designed by a professional media agency hired by the party. The two campaigns thus represent what online ad specialists would call "factual" (no images, focus on factual messages) and "emotional" (emotional images, focus on catchy messages intended to generate pride in the party).

However, the final ads differed in three important ways, which makes it difficult to pinpoint the precise mechanism underlying any observed differences between the two treatment routes. First, the factual and emotional banners (see Figures A2 through A5 in the online supplemental Appendix) differed in two salient ways: the use of images and the length of the slogan. The emotional ads all relied on emotional photos and used short, catchy phrases to convey key political messages. Compare this to the factual ads, which do not include any images and use longer, factual messages. As a result, we are not in a position to answer whether the images or the slogans (or both) led to different effect sizes.[8]

Second, one may object that the factual and emotional banners differed in the clarity of the message. For instance, the emotional education banner—which calls for "safe learning"—could have primed individuals to think about education as well as safety.[9] The factual education banner, on the other hand, is unlikely to have primed individuals to think about public safety. As such, the two treatment arms may have differed in the degree to which they activated political topics in voters' minds.

Third, a critic might argue that the valence between factual and emotional ads differed in meaningful ways. For instance, the factual ad on public safety casts a negative "zero tolerance" message, while its emotional counterpart focuses on a positive outcome—"safe travels"— and shows an image of joyful elders on a train. As such, the factual ad may have invoked a negative emotion, rather than abstracting from emotions altogether. This problem, however, is not apparent for the remaining two topics, economy and education. In the economy banner, for instance, the factual ad, if anything, transports a similarly positive message as the emotional banner.

Taken together, this lack of symmetry makes it impossible to delineate what factors of the ads—notably, color or content—give rise to any observed differences. As such, our study merely recovers the treatment effect of one campaign deemed *emotional* by party officials, and another campaign deemed *factual*.

Still, to assess whether the factual and emotional ads were received as intended, we conducted a manipulation check using a convenience sample of 149 participants in the summer of 2018. As we detail in the online supplemental Appendix Section A.5, the manipulation check broadly confirms that participants considered the emotional campaign more emotional than the factual one. Specifically, when asked whether a given emotional ad was emotional or fact-oriented, an overwhelming majority stated emotional. Moreover, in order to parse out the effect of the photos shown in the emotional ads, we constructed emotional ads with a neutral background and compared them to the emotional ads with a background image. Doing so allows us to hold constant the slogans, but vary whether photos are shown. We confirm that the emotional ads with images induce greater variance in emotional feelings than the emotional ads without images. The finding thus underlines that the images likely played a key role in activating emotions. Skepticism, however, remains whether and how the ads invoked emotions. While the manipulation check broadly

corroborates that the ads successfully captured factual and emotional domains, it remains unclear whether that very emotion was pride.

Results

To assess our first hyposthesis, we estimate the following linear model using OLS:

$$Vote\ Share_i = \beta_0 + \beta_1\ Combined_i + \epsilon_i \qquad (1)$$

where $Vote\ Share_i$ represents the party's result in postal district i. $Combined$ represents a dummy indicating whether a given district was exposed to any ads (factual or emotional) or no ads. In select robustness checks, we include 14 pre-registered covariates reported in Table A1 in the online supplemental Appendix. We estimate robust standard errors in all models and also report p-values based on randomization inference.

Vote Share

In Table 1, we present the results of our benchmark regressions. Comparing differences in means between treated and untreated districts (Column 1), we find that the online ad campaign raised the party's vote share by 1.7 percentage points (p-value of 0.156). The effect size is slightly lower for the second vote (1.3 percentage points; p-value of 0.186). When adding the preregistered control variables (Models 3 and 4 in Table 1), the effect sizes are slightly lower. Here, the campaign increased the party's first vote share by 0.9 percentage points (p-value of 0.196) and the second vote by 0.7 percentage points (p-value of 0.155).[10] Overall, the estimates thus yield tentative support for our main hypothesis, namely, that the "overall campaign raised the vote share of the party." In addition, the effect sizes are substantively meaningful, albeit not precisely estimated, due to the rather low number of clusters. Still, given that the party obtained only 19.8% for the first vote and 17.6% for the second vote—a historically low result—the campaign seemingly had an effect on the election.

A second question of interest is whether the campaign had an effect on the overall left/right political spectrum. Given that the party performed historically poorly during the

Table 1
Party's vote share in 2016

	First Vote	Second Vote	First Vote	Second Vote
	(1)	(2)	(3)	(4)
Combined estimate	0.017	0.013	0.009	0.007
S.E.	(0.012)	(0.010)	(0.006)	(0.005)
RI p-value	0.156	0.186	0.196	0.155
Controls	No	No	Yes	Yes
N (districts)	189	189	189	189

Notes. OLS regressions of the party's 2016 vote share (first and second vote) on the treatment dummy. The preregistered pretreatment controls (see Table A1 in the online supplemental Appendix) are included in Models 3 and 4, but not reported. p-values are based on randomization inference (randomly drawing 10,000 potential random assignments).

Table 2
Right-wing and left-wing vote shares in 2016

	First Vote		Second Vote	
	Conservative	Progressive	Conservative	Progressive
	(1)	(2)	(4)	(5)
Combined	0.014	–0.014	0.012	–0.014
S.E.	(0.011)	(0.011)	(0.010)	(0.010)
RI p-value	0.160	0.148	0.135	0.094
Controls	Yes	Yes	Yes	Yes
N (districts)	189	189	189	189

Notes. OLS regressions with 2016 right-wing and left-wing vote shares (first and second vote) on the treatment dummy. Conservative includes CDU, AfD, and FDP. Progressive includes SPD, Green, and Left. The preregistered pretreatment controls (see Table A1 in the online supplemental Appendix) are included, but not reported.

2016 election, it may have been the case that the ad campaign also raised the vote share of its immediate political competitors. Indeed, while administering the treatment, one campaign manager relayed to the authors that he expected the campaign to help other conservative parties given the sluggish performance of the party. In Table 2, we therefore combine the vote shares of parties traditionally considered right-leaning—the CDU, the *Alternative für Deutschland* (AfD), and the *Freie Demokratische Partei* (FDP)—into a conservative set, and parties considered left-leaning—the *Sozialdemokratische Partei Deutschlands* (SPD), *die Linke* (Left Party) and *Bündnis 90/Die Grünen* (Green Party)—into a progressive set.[11]

When analyzing the campaign's effect on the vote share of left- and right-leaning parties, we find that the factual campaign had a positive effect of 1.2 percentage points on the conservative set (p-value of 0.135) for the second vote (the more pivotal outcome). At the same time, the campaign reduced the vote share of the left-leaning parties by 1.4 percentage points (p-value of 0.094). The results hold across both the first and second vote. Moreover, they are robust to the inclusion of pretreatment covariates (see Models 2 and 4). These models thus provide additional evidence that the campaign affected the outcome of the election.

Robustness

A key robustness concern relates to the statistical significance of the findings. While estimates are weakly significant when analyzing right- and left-leaning parties in conjunction (Table 2), the party estimates reported in Table 1 are not significant and suffer from low statistical power. We address this uncertainty by estimating the following Bayesian model, which incorporates prior information on the effectiveness of mass media campaigns (Gelman & Hill, 2006):

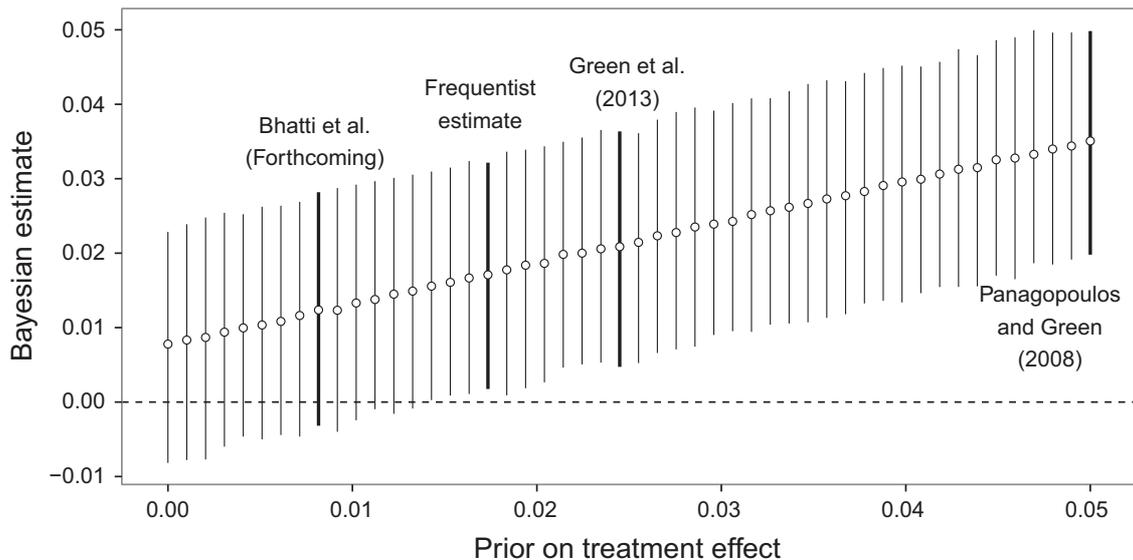$$y_i = N(\alpha + Treatment_{j[i]} + \epsilon_i, \sigma^2) \tag{2}$$

Figure 1. Bayesian models incorporating priors on treatment effectiveness.
Notes: The dots represent the estimate for the aggregate treatment effect, taken from a Bayesian regression of the party's vote share (first vote; y-axis) on the treatment indicator under different priors for the effect size (x-axis). The lines represent 95% highest density intervals. A total of 10,000 iterations were run per model.

where y is the vote share of the party (first vote) in postal district i, α is the party's mean result for the entire city, and j is a treatment dummy indicating whether the postal district received any advertisement. We let priors regarding the effectiveness of the campaign range from 0 percentage points (very skeptical) to 5 percentage points (very optimistic).[12] The maximum prior is inspired by one experimental study that randomly distributed radio ads, leading to an increase in turnout by 5.3 percentage points (Panagopoulos & Green, 2008). Most remaining studies—notably two meta-analyses assessing get-out-the-vote campaigns in the United States and Europe, cited earlier—motivate priors of 2.5 and 0.8 percentage points, respectively.

In Figure 1, we plot the estimates and 95% highest density intervals (HDI) from regressions that incorporate different priors regarding the treatment's effectiveness. The figure shows that the combined frequentist estimate of 1.7 percentage points is significantly bound away from 0 when postulating a prior of around 1.2 percentage points. Priors to achieve significance at the 10% level need to be around 0.8 percentage points for the aggregated treatment (see Figure A16 in the online supplemental Appendix).

A second robustness concern relates to one predictive covariate in particular: education. We did not include education as a potential control variable in our pre-analysis plan as the variable is not measured by Berlin's statistical agency. Still, one may be worried that the treatment, by chance, coincided with different levels of educational attainment. We therefore gained access to data on educational attainment at the postal district level from a commercial source. Table A1 in the online supplemental Appendix shows that educational levels are highly similar across the treatment groups. Moreover, when adding the variable as a control to the main regression, the key estimates are unaffected (see Table A8).

Table 3
Party's vote share in 2016

	First Vote	Second Vote	First Vote	Second Vote
	(1)	(2)	(3)	(4)
Factual estimate	0.022	0.016	0.011	0.008
S.E.	(0.014)	(0.012)	(0.007)	(0.005)
RI p-value	0.064	0.095	0.082	0.098
Emotional estimate	0.012	0.010	0.006	0.006
S.E.	(0.013)	(0.011)	(0.008)	(0.006)
RI p-value	0.314	0.319	0.329	0.338
Controls	No	No	Yes	Yes
N (districts)	189	189	189	189

Notes. OLS regressions of the party's 2016 vote share (first and second vote) on the two treatment indicators. Pretreatment controls (Table A1 in the online supplemental Appendix) are included in Models 3 and 4, but not reported. RI p-values are based on randomization inference (drawing 10,000 potential random assignments).

Mechanisms

Our second hypothesis concerns the difference between the factual and emotional campaigns. To assess this question, we next estimate the following linear model:

$$Vote\,Share_i = \beta_0 + \beta_1\,Factual_i + \beta_2\,Emotional_i + \epsilon_i \qquad (3)$$

We report results in Table 3. Interestingly, the factual campaign, if anything, fared slightly better across all models. However, differences between the two treatments are small and far from statistically significant. Specifically, the factual treatment raised the second vote by 1.6 percentage points (p-value of 0.095) in our benchmark model without control variables. When including control variables (Model 4), the estimate is 0.8 percentage points (p-value of 0.098). By contrast, the emotional banners raised the second vote by 1.0 percentage points (p-value of 0.319) in our benchmark model (Model 2) and by 0.6 percentage points (p-value of 0.338) when including control variables (Model 4). The effect sizes of the factual treatment are larger and statistically significant when analyzing the left-right political spectrum (see Table A10 in the online supplemental Appendix). Despite low statistical power, the evidence is thus contrary to our preregistered hypothesis, namely, that the "emotional campaign generated a higher vote share." We therefore interpret this evidence in favor of the null hypothesis, namely, that there is a negligible difference between the effectiveness of the two treatments. At a minimum, the evidence thus casts doubt on the truism that emotional campaigns fare better than factual campaigns (e.g., Huddy & Gunnthorsdottir, 2000).

Cost per Vote

Given precise information on ad spend, the present experiment can also provide information on the cost per vote.[13] Specifically, we can multiply the coefficients of the treatment dummies (based on Table 1) with the number of voters in the respective treatment areas.

Table 4
Cost per vote (euros)

	First Vote	Second Vote	First Vote	Second Vote
	(1)	(2)	(3)	(4)
Costs	2.94	3.91	5.74	7.50
95% CI	[1.23; ∞]	[1.55; ∞]	[2.39; ∞]	[3.17; ∞]
Controls	No	No	Yes	Yes

Notes. The table reports the estimated cost per vote for the different treatment groups. The 95% confidence interval is based on the estimates reported in Table 1. Given that the confidence intervals include zero, our upper bound is ∞.

Thereafter, we take the total budget of the online campaign and divide it by the number of additional votes. We report results in Table 4. The cost per vote ranges from 2.97 to 7.12 euros. The evidence thus showcases that online ads are a relatively cost-effective tool. The uncertainty around these estimates, however, is large.

Further Analyses

The present online experiment also allows us to bring new evidence to three additional questions at the heart of the study of political communication and behavior. The first is the degree to which campaigning has a "backlash" effect among rival parties. The second question is the degree to which the campaign affected turnout. A final question is whether the ads affected vote choice by persuading opponents or whether they merely turned ideologically proximate individuals to the polls.

Backlash

A common question among scholars of political communication is whether persuasive efforts produce backlash—either at the individual level (e.g., Nyhan & Reifler, 2010; Redlawsk, 2002; Taber & Lodge, 2006) or by directing ads to the wrong recipients. Given that the present experiment was conducted in a multiparty system, which is particularly well-balanced in Berlin, we are in a position to bring tentative evidence to this debate. In Table 5, we present models that estimate the effect of the two treatments on the vote share (first vote) of the six main parties that made it into parliament.[14] The results show that the ad campaign had a negative effect on all three left-leaning parties. Effect sizes are particularly sizable and statistically significant for the Green Party, which suffered a loss of up to 1.7 percentage points in the factual treatment. Since the campaign did not affect turnout (more on that later), it is thus unlikely that the campaign had a backlash effect among opponents. Rather, the campaign consistently reduced the vote share of all three left-leaning parties.

At the same time, the campaign slightly raised the vote share of other right-leaning parties. While the effect is negligible for the FDP—0.1 and 0.2 percentage points, respectively—there is a statistically significant effect for the right-wing populist AfD in the factual treatment of 0.8 percentage points. Taken together, these effects are in line with what marketing scholars label advertisement spillover—the inadvertent raise of sales

Table 5
Rival parties' vote share (first) and turnout

	CDU	Left	Green	SPD	FDP	AfD	Turnout
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Factual	0.011	–0.003	–0.017	–0.002	0.001	0.008	44.5
	(0.007)	(0.005)	(0.008)	(0.005)	(0.003)	(0.005)	(171.8)
Emotional	0.006	–0.005	–0.003	0.001	0.002	–0.000	155.5
	(0.008)	(0.005)	(0.008)	(0.005)	(0.003)	(0.005)	(185.7)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N (districts)	189	189	189	189	189	189	189

Notes. OLS regressions of the 2016 first vote on the factual and emotional treatment dummy, respectively. Robust SEs are given in parentheses. Preregistered pretreatment control variables are included (see Table A1 in the online supplemental Appendix), but not reported.

from close competitors as a result of advertising (Simonin & Ruth, 1998). The evidence, although estimates are often imprecise, is thus in favor of a spillover effect rather than a backlash effect. This finding matters insofar as it underlines that the targeting of ads to exclusively raise one party's vote share is a delicate task. Particularly in cases where coalition formation is at stake, campaign managers should carefully consider the spillover they are likely to produce.

Turnout

A related outcome of interest, motivated by a large literature on "get-out-the-vote" campaigns, is turnout. In the preregistration document we did not form any hypothesis regarding the campaign's effect on turnout. We did, however, commit to reporting this outcome. Analyzing results in absolute terms in column 7 of Table 5, we find that the ads had a negligible effect on turnout. The estimate showcases an increase by 155 voters in the factual treatment and 580 voters in the emotional treatment. The numbers fade in comparison to the overall increase in absolute votes of the party, which was 13,500.

Persuasion

One final related question of interest, not registered in our pre-analysis plan, was posed to us by party officials when designing the experiment. It is common sense that it is more cost-effective to get core voters to the polls as opposed to persuading swing voters or opponents (e.g., Chen & Hitt, 2002; Katz & Mair, 1995). This classic question is also posed in a literature on electoral fraud in developing countries where scholars weigh "turnout buying" against "vote buying" (Finan & Schechter, 2012; Nichter, 2008). While we cannot bring experimental evidence to this debate, we can provide exploratory observational evidence regarding the campaign's success in areas traditionally in favor of the CDU as compared with areas in favor of left-leaning parties. Berlin's history offers an excellent case study to approach this question.

From 1949 to 1990, East Berlin was the capital of the communist German Democratic Republic (GDR), while West Berlin was part of the Federal Republic of

Germany despite being located within the GDR. Since reunification in 1990, residents of East Berlin have continued to be staunch supporters of left-leaning parties, particularly the Left Party, which grew out of the main party of the GDR, the Socialist Unity Party. West Berlin, on the other hand, has historically been a stronghold of the CDU, given that the party was a driving force behind German reunification, which many West Berlin residents desired.

Figure A13 in the online supplemental Appendix plots the results across the three treatment conditions—control, emotional, and factual—across West (left graph) and East Berlin (right graph). The box plots show that electoral gains were minimal in East Berlin. Importantly, however, the campaign did not have a detrimental backlash effect. In West Berlin, on the other hand, the campaign had a pronounced positive effect. The analysis thus provides tentative evidence that the targeting of ads to areas where core supporters reside is more effective than targeting areas with a majority of opponents.[15]

Conclusion

The present article has evaluated a randomized online campaign experiment and found modest, insignificant effects on vote choice. A total of 189 postal districts in Berlin were randomly exposed to (a) emotional ads, (b) factual ads, (c) or no ads. We found that the overall campaign increased the party's vote share by 0.7 to 2.2 percentage points. Estimates, however, were not statistically significant—partly owing to a low number of clusters. We also reported a statistically significant negative effect of the campaign on the overall left-leaning political spectrum of 1.4 percentage points. Contrary to our expectations, the emotionally spirited campaign was, if anything, slightly *less* successful than the factual campaign. This result is surprising given a near-consensus in the literature that emotions are a pivotal persuasive mechanism.

Before reflecting on these findings and suggesting areas for improvement, we want to briefly revisit the precision of this article's estimates. While effect sizes are not miniscule and largely in line with theoretical expectations—down to the precise spillover within the political spectrum—few estimates reach conventional levels of statistical significance. While such uncertainty is less of a concern in an experimental study—above all, pretreatment balance was excellent—it does showcase the need to corroborate the present study by expanding the number of units. This comes at a cost. But advances in online technology, particularly the targeting of smaller geographical units, make online campaign experiments a fruitful area for future research.

Abstracting from these caveats, our findings present scholars of political behavior and communication with three insights and related pathways for future work. First, in contrast to a burgeoning literature highlighting the pivotal role of emotions in persuasion campaigns, we find that factual, non-emotional ads yielded slightly larger effects. Inasmuch as most mass media evaluations have focused on the United States, this may simply be an artifact of having studied a different population. Yet, German campaign managers similarly expected the emotional campaign to fare better. The results therefore demonstrate that voters may be equally receptive when approached with factual information as compared with suggestive, emotional ads. But, again, low statistical power does not allow one to draw firm conclusions.

A second insight relates to the spillover effects to other right-leaning parties generated by the campaign. Particularly in electoral systems that use proportional representation, campaign managers must think carefully about the results of rival parties. While we

found that spillover effects are largely in line with political expectations—right-leaning parties gain, left-leaning parties lose—the precise allocation of votes remains a challenge to model. In particular, we found tentative evidence that the ads spilled over primarily to parties on the far right (AfD), and to a lesser extent to parties closer to the center (FDP). The evidence thus implies that ad spillover seems to benefit more extreme parties rather than representing random "measurement error." Future evidence and theorizing, however, is necessary to model these losses with greater precision, and to explore whether ad spillover, indeed, follows such a polarizing pattern.

A third insight, observational in nature, is the finding that the campaign raised the vote share more by mobilizing supporters and de-mobilizing opponents as opposed to persuading opponents. In particular, we noted that the campaign performed significantly better in West Berlin—a CDU stronghold—without raising overall turnout. Future research, perhaps by conducting rigorous subgroup analyses, could help shed light on the precise process and population that gives rise to this finding. Specifically, researchers could couple administrative outcomes with survey-based data—tools that Facebook and Google have at their disposal—in order to parse out whether voters were persuaded or merely de-mobilized. Notwithstanding, the campaign did no electoral damage in areas that are fully opposed to the CDU, which highlights that the discussion between persuasion and differential mobilization is alive and well.

In concluding we would also like to situate our finding within the current public debate regarding online ads. Amid heated arguments about the use of online ads and bots during the U.S. 2016 presidential election, the present article has produced sobering results. Effect sizes were positive, but imprecise and not extraordinarily large. We are thus skeptical whether online ads have a decisive influence on elections. What is more, we showed that factual ads, if anything, fared better than emotional ads. Highly manipulative ads that activate voters' most spiteful emotions—a common characterization of ads in the public debate—may thus not be the most efficient way to move the needle. Perhaps voters, at least in Berlin, prefer a political process based on a rational exchange of ideas (Habermas, 1987).

Notes

1. Among the few empirical studies, Broockman and Green (2014) find weak evidence that voters are more likely to recognize candidates when randomly exposed to ads on Facebook. Relatedly, Collins, Kalla, and Keane (2014) randomly exposed voters to get-out-the-vote Facebook ads, and find null and negative effects, respectively. Only one study has linked online ads to administrative outcomes, namely, turnout. In particular, Bond and colleagues (2012) find that Facebook voting reminders do not increase turnout (see also Konitzer, Rothschild, Hill, & Wilbur, 2018).
2. This analysis was not preregistered.
3. Pride figures prominently among the emotions hypothesized to spur individuals' willingness to process political information. Panagopoulos (2010), for instance, uses a large-scale field experiment to show that the activation of pride motivates compliance with voting norms as compared with a different emotion, namely, shame (see also Gerber, Green, & Larimer, 2010). A related literature in social psychology has presented evidence that pride motivates prosocial behavior (e.g., Williams & DeSteno, 2008). Finally, scholars of advertising have argued that the activation of pride is a powerful emotion in facilitating persuasion (Ambler, 2000).
4. Given that the online experiment also enables us to analyze clicks rates across the two treatments, we added a final third hypothesis: "The emotional campaign generates a higher

click rate as compared with the factual campaign." To save space, this hypothesis is discussed in Section A.8 in the online supplemental Appendix.

5. Facebook can offer more reliable and fine-grained targeting as its users frequently share their whereabouts. Google users, on the other hand, do not typically communicate comparable information, thereby making local-level targeting less reliable. In general, administering ads to postal districts poses a significant technological challenge to ad providers. Targeting mobile machines (i.e., phones or tablets) is typically less difficult, given that these devices submit GPS signals (e.g., when using apps). This enables providers to locate individuals' likely home addresses. The case is more challenging for desktop computers. Here, Facebook and Google have to rely on IP addresses, which are more difficult to link to postal districts. Thus, unless ad providers can link phones or tablets to desktop computers—the most common and efficient technique—they have to rely on databases linking IP address ranges to postal districts. There can thus be no doubt that there was significant spillover across postal districts due to imperfect targeting by the ad providers. Unfortunately, neither Facebook nor Google offer measures that would allow one to assess the extent of the spillover. If anything, however, such fuzzy treatment assignment works to dampen potential treatment effects. Our effect sizes are thus likely a lower bound of the overall effectiveness. One potential, although imperfect, way to investigate possible spillovers is to assess whether the preregistered model (discussed later and shown in Table 1) systematically underestimates the party's vote share in control districts that share borders with treatment districts. We conduct such an analysis in Section A.7 in the online supplemental Appendix. Indeed, we confirm a positive correlation between model residuals and the percentage of control districts' shared borders with treatment districts. However, whether this spillover effect is due to fuzzy treatment assignment or due to social networks is unclear.

6. The exact allocation of funds is laid out in the preregistration document. Further details and key metrics of the online campaign are discussed in Section A.2 in the online supplemental Appendix.

7. The party requested that the online campaign also include videos. Importantly, as we detail in Section A.2 "Online campaign", the videos shared elements of *both* the factual and emotional condition. The videos, therefore, do not compromise the causal identification of the two main treatment conditions; they may, however, have washed out treatment effect differences across the two conditions.

8. A manipulation check, discussed in detail later, shows that the images are a likely source for the variation between the two ads.

9. We believe that such an interpretation is rather unlikely, given that the subheading of the banner explicitly states "We provide the most funding to schools, teachers and school renovations."

10. The reason for the reduction in effect sizes is the inclusion of the party's historic vote share. This can be seen when analyzing added variable plots, provided in Figures A14 and A15 in the online supplemental Appendix. In addition, in Table A12 in the online supplemental Appendix, we show that results are similar (and insignificant) when combining the first and second vote to a standardized index.

11. Table 2 was not included in the pre-analysis plan.

12. Specifically, the priors for the aggregate treatment are $Treatment_{j=1[i]} - Treatment_{j=0[i]} \sim N(\mu, \sigma^2)$, where $\mu \in [0, 0.05]$ and σ^2 is equal to the standard deviation of the frequentist estimate.

13. While we cannot communicate the precise number, in Section A.6 in the online supplemental Appendix we report the overall campaign spending by all major parties during the election so as to situate the study in the broader political context.

14. In Table A7 in the online supplemental Appendix we show that the results are corroborated when using the second vote as the outcome. This analysis was not preregistered.

15. We should note, however, that Berlin has an extraordinarily high number of swing voters (see Section A.6 in the online supplemental Appendix for details).

References

Ambler, T. (2000). Persuasion, pride and prejudice: How ads work. *International Journal of Advertising*, *19*(3), 299–315. doi:10.1080/02650487.2000.11104803

Arceneaux, K. (2010). The benefits of experimental methods for the study of campaign effects. *Political Communication*, *27*, 199–215. doi:10.1080/10584601003709407

Arceneaux, K., & Nickerson, D. W. (2009). Who is mobilized to vote? A re-analysis of 11 field experiments. *American Journal of Political Science*, *53*(1), 1–16. doi:10.1111/j.1540-5907.2008.00354.x

Bekkouche, Y., & Cagé, J. (2018). *The price of a vote: Evidence from France, 19932014*. Working paper. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3106818

Bennett, W. L., & Iyengar, S. (2008). A new era of minimal effects? The changing foundations of political communication. *Journal of Communication*, *58*(4), 707–731. doi:10.1111/jcom.2008.58.issue-4

Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D. I., Marlow, C., Settle, J. E., & Fowler, J. H. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature*, *489*(7415), 295–298. doi:10.1038/nature11468

Brader, T. (2005). Striking a responsive chord: How political ads motivate and persuade voters by appealing to emotions. *American Journal of Political Science*, *49*(2), 388–405. doi:10.1111/ajps.2005.49.issue-2

Brader, T. (2006). *Campaigning for hearts and minds: How emotional appeals in political ads work*. Chicago, IL: University of Chicago Press.

Broockman, D. E., & Green, D. P. (2014). Do online advertisements increase political candidates name recognition or favorability? Evidence from randomized field experiments. *Political Behavior*, *36*(2), 263–289. doi:10.1007/s11109-013-9239-z

Chen, P.-Y., & Hitt, L. M. (2002). Measuring switching costs and the determinants of customer retention in Internet-enabled businesses: A study of the online brokerage industry. *Information Systems Research*, *13*(3), 255–274. doi:10.1287/isre.13.3.255.78

Collins, K., Kalla, J. L., & Keane, L. (2014). *Youth voter mobilization through online advertising: Evidence from two GOTV field experiments*. Working Paper. Retrieved from https://igs.berkeley.edu/sites/default/files/biography-pdfs/kalla-cv.pdf

FEC. 2012. Campaign finance disclosure portal. Retrieved from http://fec.gov/pindex.shtml

Finan, F., & Schechter, L. (2012). Vote-buying and reciprocity. *Econometrica*, *80*(2), 863–881. doi:10.3982/ECTA9035

Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. New York, NY: Cambridge University Press.

Gerber, A. S., Green, D. P., & Larimer, C. W. (2010). An experiment testing the relative effectiveness of encouraging voter participation by inducing feelings of pride or shame. *Political Behavior*, *32*(3), 409–422. doi:10.1007/s11109-010-9110-4

Green, D. P., Gerber, A. S., & Nickerson, D. W. (2003). Getting out the vote in local elections: Results from six door-to-door canvassing experiments. *Journal of Politics*, *65*(4), 1083–1096. doi:10.1111/1468-2508.t01-1-00126

Green, D. P., McGrath, M. C., & Aronow, P. M. (2013). Field experiments and the study of voter turnout. *Journal of Elections, Public Opinion and Parties*, *23*(1), 27–48.

Green, D. P., & Vavreck, L. (2007). The targeting & mobilization of effects of television advertising. *Princeton Universitys Center for the Study of Democratic Politics*. Working paper. Retrieved from http://www.academia.edu/2789521/The_Targeting_and_Mobilization_Effects_OF_Television_Advertising

Habermas, J. (1987). *The theory of communicative action* (Vol. 2). Boston, MA: Beacon.

Hartmann, G. W. (1936). A field experiment on the comparative effectiveness of 'Emotional' vs. 'Rational' political leaflets in determining election results. *Journal of Abnormal and Social Psychology*, *31*, 99–114. doi:10.1037/h0056079

Hovland, C., Lumsdaine, A. A., & Sheffield, F. D. (1949). *Experiments on mass communications*. Princeton, NJ: Princeton University Press.

Huddy, L., & Gunnthorsdottir, A. H. (2000). The persuasive effects of emotive visual imagery. *Political Psychology*, *21*(4), 745–778. doi:10.1111/0162-895X.00215

Kahn, K. F., & Geer, J. G. (1994). Creating impressions: An experimental investigation of political advertising on television. *Political Behavior*, *16*(1), 93–116. doi:10.1007/BF01541644

Kaid, L. L., & Boydston, J. (1987). An experimental study of the effectiveness of negative political advertisements. *Communication Quarterly*, *35*(2), 193–201. doi:10.1080/01463378709369680

Katz, R. S., & Mair, P. (1995). Changing models of party organization and party democracy the emergence of the cartel party. *Party Politics*, *1*(1), 5–28. doi:10.1177/1354068895001001001

Konitzer, T., Rothschild, D., Hill, S., & Wilbur, K. C. (2018). Using big data and algorithms to determine the effect of geographically targeted advertising on vote intention: Evidence from the 2012 US presidential election. *Political Communication*, 1–16. doi:10.1080/10584609.2018.1467985

Lazarsfeld, P. F., Berelson, B., & Gaudet, H. (1944). *The people's choice*. New York, NY: Columbia University Press.

Marcus, G. E., Neuman, W. R., & MacKuen, M. (2000). *Affective intelligence and political judgment*. Chicago, IL: University of Chicago Press.

Middleton, J. A. (2008). Bias of the regression estimator for experiments using clustered random assignment. *Statistics and Probability Letters*, *78*, 2654–2659. doi:10.1016/j.spl.2008.03.008

Neuman, W. R. (2007). *The affect effect: Dynamics of emotion in political thinking and behavior*. Chicago, IL: University of Chicago Press.

Nichter, S. (2008). Vote buying or turnout buying? Machine politics and the secret ballot. *American Political Science Review*, *102*(01), 19–31. doi:10.1017/S0003055408080106

Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, *32*(2), 303–330. doi:10.1007/s11109-010-9112-2

Panagopoulos, C. (2010). Affect, social pressure and prosocial motivation: Field experimental evidence of the mobilizing effects of pride, shame and publicizing voting behavior. *Political Behavior*, *32*(3), 369–386. doi:10.1007/s11109-010-9114-0

Panagopoulos, C., & Green, D. P. (2008). Field experiments: Testing the impact of radio advertisements on electoral competition. *American Journal of Political Science*, *52*, 156–168. doi:10.1111/j.1540-5907.2007.00305.x

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology*, *19*, 123–205.

Redlawsk, D. P. (2002). Hot cognition or cool consideration? Testing the effects of motivated reasoning on political decision making. *The Journal of Politics*, *64*(04), 1021–1044. doi:10.1111/1468-2508.00161

Roseman, I., Abelson, R. P., & Ewing, M. F. (1986). Emotional and political cognition: Emotional appeals in political communication. In R. R. Lau & D. O. Sears (Eds.), *Political cognition* (pp. 279–294). Mahwah, NJ: Lawrence Erlbaum Associates.

Ryan, T. J. (2012). What makes us click? Demonstrating incentives for angry discourse with digital-age field experiments. *The Journal of Politics*, *74*(04), 1138–1152. doi:10.1017/S0022381612000540

Simonin, B. L., & Ruth, J. A. (1998). Is a company known by the company it keeps? Assessing the spillover effects of brand alliances on consumer brand attitudes. *Journal of Marketing Research*, *35*(1), 30–42. doi:10.2307/3151928

Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, *50*(3), 755–769. doi:10.1111/j.1540-5907.2006.00214.x

Williams, L. A., & DeSteno, D. (2008). Pride and perseverance: The motivational role of pride. *Journal of Personality and Social Psychology*, *94*(6), 1007. doi:10.1037/0022-3514.94.6.1007