July 2021 Update for: Can Video Advertisements Change Partisanship? Results From Panel Survey Experiments

Alexander Coppock, Donald P. Green, and Ethan Porter*

July 6, 2021

In the May 2, 2021 version of this paper (posted to SSRN on on May 7, 2021, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3838299), we described the results of the five experiments that indicated that exposing participants to pro-Democratic ads caused them to identify more strongly as Democrats, even 10 days post-treatment. We summarized these results for a general audience in an an op-ed for Politico (https://www.politico.com/news/magazine/2021/06/04/political-science-campaign-advertising-party-persuasion-491804).

In this update, we describe new reservations about differential attrition in our study that will require changes to the analytic strategy, the collection of new data, and revision of the paper. In advance of doing this reanalysis and any further experimentation, we wanted to warn readers of the problem and to communicate that we are working to address it.

By way of background, our experiments followed a placebo-controlled, panel survey experimental design over three waves. In wave 1, we measured pre-treatment covariate information. In wave 2, we administered treatments and collected wave 2 outcomes. In wave 3

*Alexander Coppock is Assistant Professor of Political Science, Yale University. Donald P. Green is Burgess Professor of Political Science, Columbia University. Ethan Porter is Assistant Professor of Media and Public Affairs, George Washington University.

approximately 10 days later, we re-interviewed subjects and collected wave 3 outcomes. We used placebo-controlled designs throughout. Subjects in the placebo groups were assigned to see an equivalent number of product adverstisements. We used a placebo-controlled designs because we anticipated that watching as many as six videos would cause some subjects to quit the survey and we want to maintain balance across treatment and control. That is, we wanted to prevent differential attrition as much as we could.

Study by study, we had checked response rates by condition at wave 3, finding no statistically significant differences (see Figure 1 for rates and Figure 2 for the effects of treatment on response rates). In a reorganization of the paper, we cut this analysis from the appendix and failed to mention it in the main text, which was an error.

We did not, however, check to see if response rates at wave 2 – the responses immediately post-treatment – suffered from differential attrition. Looking in to this problem, we discovered that indeed, subjects assigned to watch more videos were more likely to drop out of the survey. We only ever obtained "complete case" data for study 4, so we don't know the extent of the problem in that study and we are currently working to understand why Study 3 does not display this pattern. Additionally, in the process of investigating the wave 2 results, we uncovered a coding error that partially obscured the attrition problem, which we have now corrected. In light of the wave 2 results, we now view wave 3 with more suspicion as well.

We plan to take two approaches going forward. First, we will employ trimming bounds (also known as Lee bounds), stratified by pre-treatment party ID, to the estimates at wave 2 and wave 3. Preliminary analysis suggests that these bounds imply much greater statistical uncertainty about the effects of the ads. Figure 3, for example, shows the naive estimate and the trimming bounds on the effects among "always-reporters" at wave 3. Second, we will conduct fresh experiments that will work harder to address the attrition problem.

We wish to express our apologies to readers for not appreciating the extent to which our

Figure 1: Wave 2 and Wave 3 response rates by condition

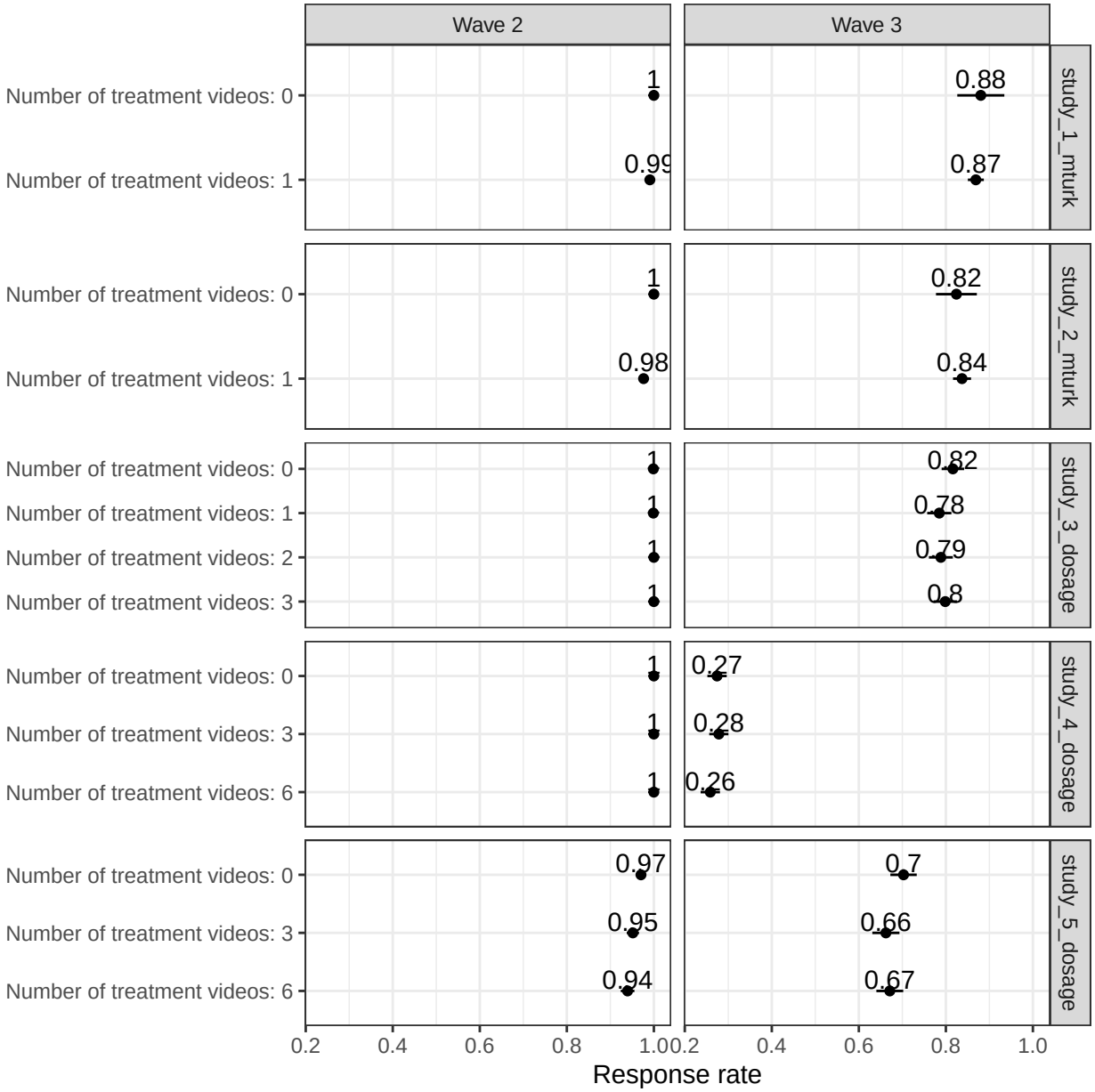


Figure 2: Effects of treatments on Wave 2 and Wave 3 response rates

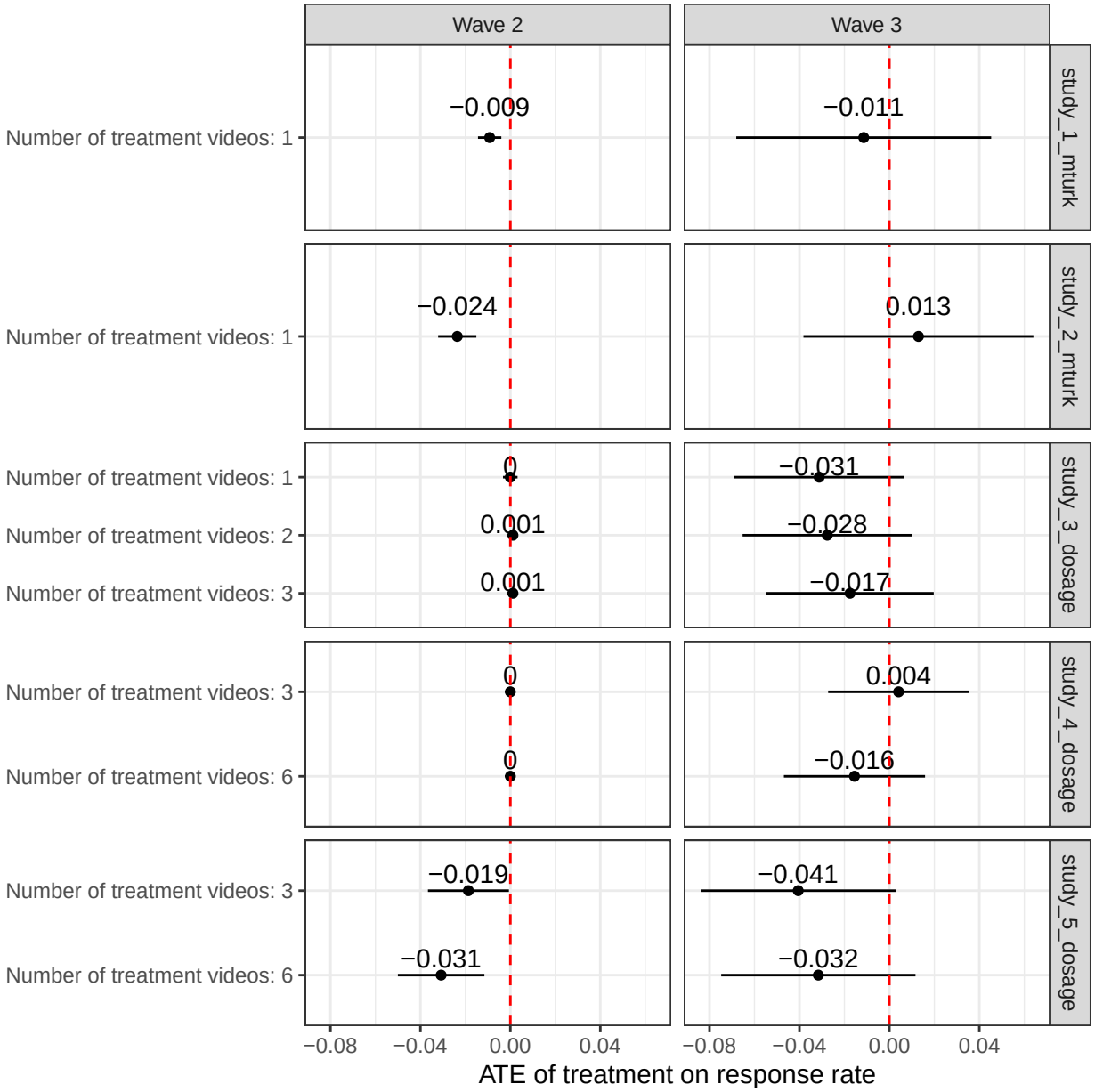
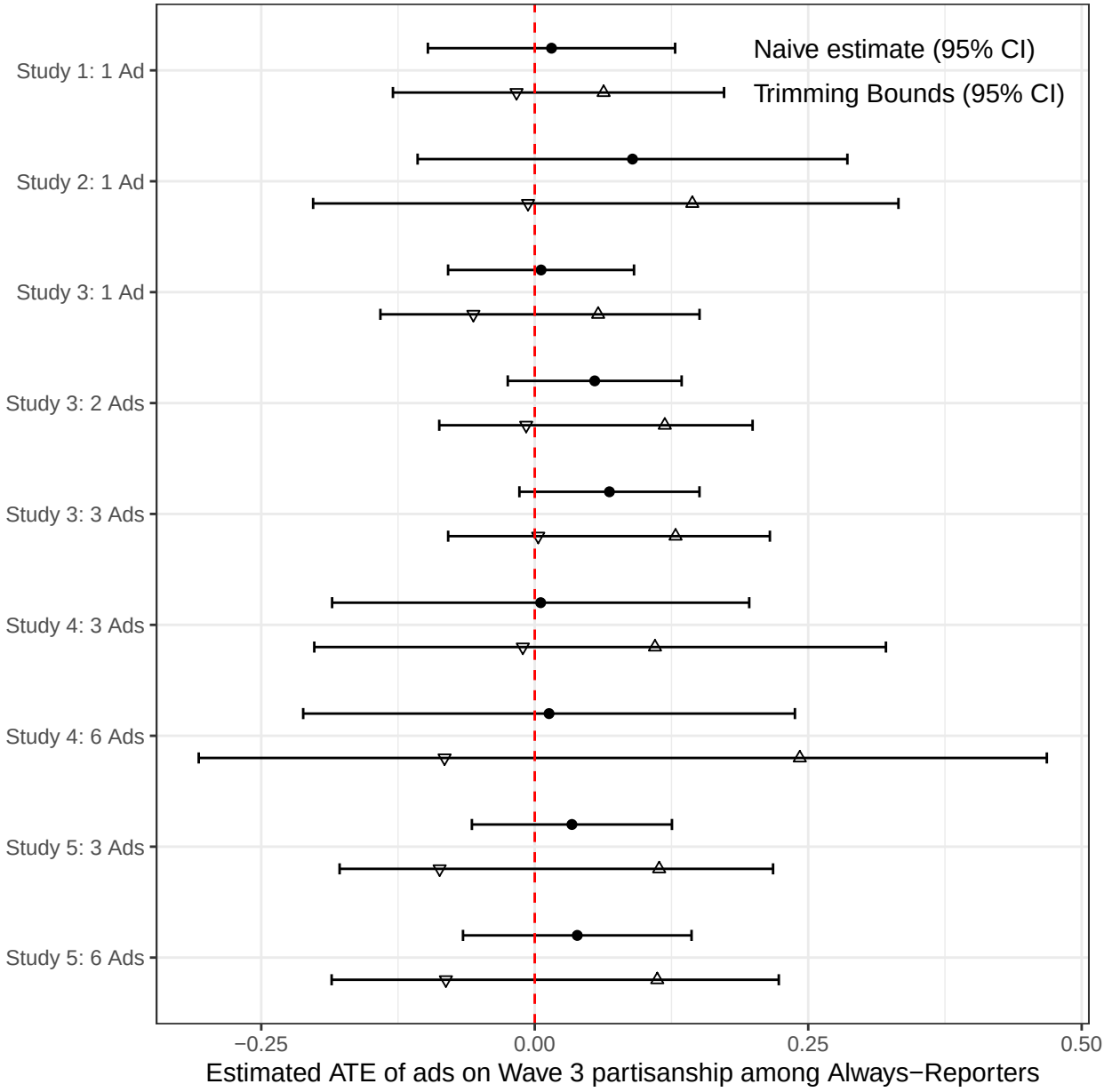


Figure 3: Estimated effects on 7-point party ID at Wave 3 using bounds



design was weakened by differential attrition, and we thank Peter Aronow for his help in diagnosing and detecting the problem.

Can Video Advertisements Change Partisanship? Results From Panel Survey Experiments

Alexander Coppock, Donald P. Green, and Ethan Porter*

May 2, 2021

Abstract

Although partisan orientations are sometimes characterized as the unmoved movers of political evaluations and vote choice, both panel surveys and aggregate time-series indicate that partisan change does occur. To date, however, researchers have seldom attempted to induce a change in partisanship experimentally. Guided by longstanding explanations of how people come to acquire and update their partisan attachments, we worked with media consultants to develop persuasive political advertising designed to win partisan converts. These video ads were deployed in a series of multi-wave survey experiments that tracked panel respondents over time in an effort to gauge the messages' immediate and persistent effects on party identification, as well as their downstream effects on political evaluations and voting preferences. We find that party-focused messaging operates as theoretically expected, especially at high dosages, producing small but demonstrable effects on partisanship with downstream consequences

*Alexander Coppock is Assistant Professor of Political Science, Yale University. Donald P. Green is Burgess Professor of Political Science, Columbia University. Ethan Porter is Assistant Professor of Media and Public Affairs, George Washington University. This research was reviewed and approved by the Institutional Review Board of Columbia University (IRB-AAAN3800). For studies 3-8, no university funds were used in the production of the treatments or the collection of survey responses. We thank Joshua Handelman, Jennifer Mueller and Emily Mellencamp-Smith for their assistance in securing funding, the Democratic Legislative Campaign Committee and the American Center for facilitating support, David Shor for fielding one of the studies, J.T. Harvat for assistance with video production and Frank Sesno for providing narration.

for political preferences. The fact that these effects dissipate over time helps illuminate individual- and aggregate-level patterns of partisan change.

Few literatures in political science are more extensive than the accumulated body of scholarship on partisan evaluations and attachments. Spanning several decades and drawing data from many countries, the study of voters' views about political parties encompasses a broad array of theoretical perspectives and research methods. The vitality of the literature on partisan attitudes – a capacious term that includes the affective sense of identification with a party as well as the cognitive assessments of parties' attributes and competencies – reflects their strong correlation with evaluations and preferences related to day-to-day politics. Partisans of different stripes often have sharply divergent opinions about how things are going, who is to blame, and what should be done. This pattern is especially pronounced in the United States, where partisan conflict has intensified in recent decades.

But where do partisan attachments come from? Although researchers since the 1950s have noted the correlation between one's partisan orientations and those of one's parents (e.g., Hyman, 1959), the literature on political socialization can be read as calling attention to the slippage that often occurs as political views are transmitted. The highly influential Youth-Parent Socialization Panel Study (Jennings et al., 2005) underscores how little correlation remains between generations' partisan coloration once the younger generation reaches middle age (Niemi and Jennings 1991; Jennings, Stoker and Bowers 2009. See also Hatemi and Ojeda 2020.) Although partisanship tends to change only gradually from age thirty on, a good deal of change occurs during one's twenties. Later in life, gradual changes can add up to important shifts over long stretches of time (Green and Palmquist, 1994).

Several leading theories offer insights into how and why partisan attitudes evolve. Fiorina (1981) and MacKuen, Erikson and Stimson (1989) argue that macroeconomic fluctuations and scandals cause citizens to reappraise the parties. Another leading theory focuses on policy positions rather than performance evaluations. Voters are said to select their preferred party by comparing their own location on leading issues to the positions advocated by party leaders (Jackson, 1975; Franklin and Jackson, 1983). A third leading theory, inspired by

evidence suggesting that many voters have limited understanding of the parties' ideological locations, stresses instead the personal appeal of prominent party figures (Harris, 1954). A fourth theory suggests that citizens are attentive to the social group imagery associated with party coalitions and political campaigns, and gravitate to the party whose partisans look most like them (Green, Palmquist and Schickler, 2002).

A residual perspective posits that partisan evaluations have theoretically nonspecific origins, and are instead simply byproducts of respondents' broad feelings about which party is vindicated by the "nature of the times." This school of thought can be traced back at least to Converse (1964). Such evaluations may be shaped by the tone of media coverage about the leading parties. While media coverage may encompass messages about performance, policy, and personae, the specific thematic content is less important than the overall impression about the nature of the times left by the balance of media coverage (McCombs, 2014; Kleinnijenhuis et al., 2006; Kepplinger et al., 1989).

Each of these arguments has generated a sizable empirical literature comprised largely of observational evidence. Drawing on a combination of cross-sectional and panel surveys, individual-level research suggests that partisanship is shaped by short-term forces, but critics charge that citizens' short-term perceptions are themselves a by-product of partisan attitudes (Green and Palmquist, 1990). The resulting literature is a thicket of methodological arguments about nonrecursive modeling and unobserved heterogeneity. Ultimately, although some evidence indicates that Americans have increasingly "sorted" themselves into distinct social and ideological groups (Levendusky, 2009; Brown and Enos, 2021), the causal ordering and isolation of the processes involved remains unsettled.

Uncertainty surrounds not only the causes of partisanship but also its consequences for vote choice. The vast majority of studies that assess the effects of party identification on vote choice or candidate evaluations do so by applying regression models to nonexperimental data. This approach inevitably provokes debates about whether vote-related preferences are

precursors to partisanship, in which case failing to control for them courts omitted variables bias. Alternatively, if candidate evaluations are shaped by partisanship, controlling for them risks post-treatment bias. In our view, these debates cannot be advanced by further observational data analysis, since the empirical conclusions drawn to date depend so deeply on contested assumptions.

Recent interest in experimentation among public opinion scholars offers an opportunity to break this methodological impasse. Experimental investigation in this domain involves two interrelated research programs. The first is to develop interventions that affect partisanship; the second is to assess whether this exogenous change to party identification sets in motion other downstream changes, such as shifts in how respondents evaluate political figures or current conditions. To date, however, lab studies have found party attachments to be unyielding. Cowden and McDermott (2000) attempted to change partisan attachments by exposing respondents to hypothetical election contests in which one party or the other took ideologically extreme stances; they also had lab subjects read and write briefs in support of or in opposition to the impeachment of Bill Clinton. Neither intervention moved party attachments. A subsequent field experiment (Gerber, Huber and Washington, 2010) had better success by sending letters to randomly selected Connecticut residents who declined to state a party affiliation when registering to vote; the letters urged them to register with a major party so that they could vote in the upcoming 2008 presidential primary. This encouragement did seem to increase participants' level of party identification when they were surveyed a few months later. Gerber, Huber and Washington (2010) also find some evidence that as new registrants attach themselves to a party, they become more likely to adopt its outlook. Beyond encouragements to register with a political party, other theoretically guided paths to partisanship have not been tested experimentally. This paper builds on these experimental foundations.

An experimental research agenda has the capacity to shed light not only on whether par-

tisanship changes in the wake of an intervention but also on which kinds of people move the most. Prior research leads us to expect that younger people have weaker partisan attachments than their older counterparts (Jennings et al., 2005). If those with weaker attachments are generally more open to change or if young people change because they are at an "impressionable" stage of life, they should display elevated treatment effects.

In contrast, effects may be attenuated among those who have strong partisan identities, regardless of age, prior to being exposed to our messages. Adherents of one party who encounter ads designed to increase identity with the other party may be unmoved or even deepen their existing party attachment. Such behavior would be in line with findings on partisan "motivated reasoning," whereby individuals adjust their beliefs to correspond with their partisan identities or actively resist arguments that challenge the superiority of their preferred party (e.g., Bolsen and Cook, 2014). With these possibilities in mind, we will consider below whether the persuasive effects of our treatments differ by respondents' age and pre-treatment partisanship.

We begin by laying out a set of perspectives that guided the development of short videos designed to change party identification. The videos were produced by professional ad-makers, all of whom have made ads for sitting federal officials. Importantly, the array of messages offered reasons for affiliating with the Democratic Party but did not mention specific candidates or weigh in on upcoming elections. A series of five multi-wave panel survey experiments involving 8,291 participants examine whether exposure to these messages in varying dosage affects party attachments in the short run (immediately after viewing the ad), medium run (a week or two later), or long run (more than a year later). These survey experiments also measure whether exposure to party-focused ads has subsidiary effects on vote preference and presidential approval, as would be predicted by theories that characterize party identification as a cause of evaluations. We find robust evidence of small but meaningful movements in party identification, especially at high dosages, accompanied by shifts in evaluations and vote

intentions. These effects are roughly halved a week or two later and disappear altogether a year later. We find some evidence that younger respondents and independents were most affected by our messages. We conclude by discussing the implications of these findings for our understanding of micro- and macro-level partisan change.

Dynamics of Partisan Change

Although party identification is sometimes caricatured as an "unmoved mover," even the most ardent defenders of partisan stability concede that change does occur at both the micro and macro levels. Analysis of panel data with three or more waves of interviews has, since the 1970s, become increasingly attentive to the distinction between wave-to-wave changes that occur due to sloppy survey responses and changes that occur due to true underlying changes in partisan attitudes (Asher, 1974; Achen, 1975).

The hallmark of the former type of change is that the correlations between attitudes measured at time t and time $t + k$ are well below 1 but diminish slowly if at all as the time elapsed between t and $t + k$ lengthens. For example, in the original ANES 1956-58-60 panel, the correlation between the seven-point party ID scale in 1956 and 1958 was 0.85, while the correlation between 1956 and 1960 was 0.83. Similarly, among the parents of the Youth-Parent Socialization panel, the correlation between the seven-point party ID scale in 1965 and 1973 was 0.78, while the correlation between 1965 and 1983 was 0.77 (Green and Palmquist, 1994). The hallmark of true change, by contrast, is substantial decline in these correlations as the time elapsed between t and $t + k$ lengthens. The outbreak of the First Gulf War seemed to cause a quickening of partisan change: in the ANES panel the correlation between the seven-point scale in 1990 and 1991 was 0.86 while the correlation between 1990 and 1992 fell to 0.80. The multi-wave panel studies from the 1950s to the early 1990s indicate that, net of measurement error, about one in eight panel respondents

changed their partisanship by more than 1 scale point over a two-year period (Green and Palmquist, 1994, p.455). More recent panel surveys seem to suggest a slightly elevated rate of change (Tucker, Montgomery and Smith, 2019), although these inter-temporal differences may reflect over-time changes in response rates and survey mode.

Although panel correlations help describe the relative positioning of a given set of respondents over time, they do not reflect changes that may occur over time in mean responses. (In principle, an over-time correlation of close to 1.0 could occur if everyone's partisanship increased by one scale point between waves.) The study of macropartisanship (MacKuen, Erikson and Stimson, 1989) is the investigation of how mean levels of partisanship evolve over time. The macropartisanship literature has produced two key findings: the first is that mean partisanship is responsive to short-term forces. The proportion of the electorate that identifies with a given party changes as that party's fortunes change. If that party holds the presidency, high levels of presidential approval attract more partisan identifiers. These short-term changes are seldom large, however, and even the cumulative effect of a sustained run of good (or bad) news changes the balance of party identification by only a few percentage points. Green and Schickler's 2009 (p.192) analysis of the Democratic share of party identifiers from 1976 to 2007 concluded that "a 20 point increase in presidential approval sustained over a four year period translates into 3.4 points of macropartisanship." The second is that when short-term forces dissipate (either because conditions change or because the public's focus changes), the partisan balance gradually returns to its long-term equilibrium. Evidently, memories fade slowly.

Or perhaps memories fade for most but not all of the electorate. Enduring macropartisan change in the wake of short-term forces may be concentrated among young voters just entering the electorate, as suggested by the "impressionable years" hypothesis (Sears, 1983). Two facts seem to support this hypothesis. The first is that panel studies that have tracked generational cohorts over time have found much more movement among younger cohorts

(again net of measurement error). For example, high school students interviewed in 1965 showed much more instability in partisanship by 1973 than did their parents. The second fact is that those who come of political age at moments when short-term forces seem to decisively favor one party (e.g., 1984 for Republicans or 2008 for Democrats) appear to bear the stamp of their early experiences for years to come. On average, the magnitude of these "generation" effects tends to be modest, however, amounting to only a few percentage points per presidency (Jacobson, 2016, Figure 1).[1]

While much of what we know about dynamics – the responsiveness to short-term forces and the speed with which a short-term disturbance's effect dissipates over time – comes from aggregate data, one important recent exception comes from Tucker, Montgomery and Smith (2019), who analyze a rare twenty-wave panel study that spans 2011 to 2016. One insight from their analysis is that shocks at the individual level dissipate quickly; a shock that moves party identification 0.21 scale points in one wave has an effect of just 0.04 scale points in the next wave (roughly four months later) and just 0.01 two waves later (Tucker, Montgomery and Smith, 2019, Figure 7). These findings represent an important departure from aggregate time-series analysis, which tends to show that shocks dissipate slowly. One motivation for our experimental approach is to corroborate these results using randomly assigned interventions, as opposed to a shock that occurs via an uncontrolled process that may be confounded by omitted variables.[2]

The Tucker, Montgomery and Smith (2019) findings also have an important design implication: if experimental researchers want to study the dynamics of treatment effect decay, they

[1]We find some echoes of these patterns in our experiments measuring the effects of pro-Democratic videos on partisanship: effects are mildly larger for people under thirty-five than for their older counterparts. However, contrary to the motivated reasoning literature, the effects we observe are not strongly conditioned by past party identification.

[2]Some prior work in this area has been stymied by a lack of control over exogenous shocks. For example, when Jennings and Markus (1984, p. 1001) do not observe young people becoming more stable in their views over time, as expected, they can only speculate that the lack of stability reflects the shock of the Vietnam War.

must invoke maximum-strength interventions with large samples so that they can reliably detect how the effect dissipates over time. We took this point to heart in our experimental design, which is two orders of magnitude larger than previous experiments in this literature. In the study of partisan change, Cowden and McDermott (2000) describe experiments that assigned participants to brief textual interventions, while Gerber, Huber and Washington (2010) mailed letters. More broadly, lab studies that examine how ads and vote preference often rely on one or two video treatments (Kahn and Geer, 1994; Brader, 2005; Coppock, Hill and Vavreck, 2020). Although the quantity of political ads aired during national elections dwarfs these totals (Fowler, Ridout and Franz, 2016), inattentiveness to commercials and the rise of multi-screening (Segijn et al., 2017) suggest that many viewers are not paying attention to many of the ads shown. By assigning an unusually high quantity of ads in a controlled survey experiment, we attempt to approximate exposure to large numbers of ads in real-world settings.

Dosage is also important for theories of persuasion. In general, message repetition makes subsequent retrieval from memory easier (Schwarz, 2004). Subjects exposed to a larger number of stimuli that convey the same message will have more opportunities to cognitively engage with the message and take up its recommendations, strengthening the longevity of effects (Petty and Cacioppo, 1986). If one message advocating for a party does not appeal to them or otherwise does not maintain their attention, another might. Indeed, some evidence shows that repeated exposure to different variations of the same message has more lasting effects than repeated exposure to identical versions of the same message (Unnava and Burnkrant, 1991). Whether messages are identical or distinct, increased dosage makes it more likely that messages are encoded in memory and available for later evaluations.

Methodological Challenges

In this section, we discuss three threats to validity that arise in the context of experiments that present subjects with persuasive messages concerning political parties. These threats concern (1) the measurement of outcomes, (2) control over the content of the message and the context in which it is received, and (3) pre-treatment measurement of individual differences that are thought to interact with the treatment. We also describe our responses to these challenges.

Outcome measurement

Our core hypothesis is that exposure to ads promoting one political party should increase respondents' self-identification with that party. But making a convincing case that attitude change has occurred means showing that the persuasive effects of communication persist over time. This requirement places important practical demands on any experiment, as the subjects must be exposed to the treatment at one point in time and outcomes measured at some point in the future. Although most lab and survey experiments present stimuli and measure outcomes immediately thereafter (perhaps after a few minutes of intervening questions), some notable studies have tracked participants over time. For example, Chong and Druckman (2010) traced the enduring effects of issue frames over the course of 10 days. Internet panels, in particular, facilitate this type of follow-up measurement (Clinton and Lapinski, 2004).

Separating in time the treatment from the measurement of outcomes serves other methodological purposes as well. Presenting subjects who know they are participating in a research study with a video about one or both political parties and shortly thereafter asking questions about the parties makes apparent the connection between the stimulus and the response. If participants sense that the intelligent or desired response is one that praises the party being

extolled in the video, its apparent effects may be spurious. To some extent, this concern may be addressed by providing subjects with a cover story (e.g., they are asked to watch the video as though they were evaluating its cinematography) or by burying the treatment video in a series of other irrelevant videos, but the fact that subjects are being paid to watch a party propaganda video in itself is sufficiently unusual and obtrusive to threaten the symmetry of outcome measurement in treatment and placebo groups. Conducting the treatment and outcome measurement at separate points in time does not fully eliminate this threat to validity, but it does strengthen the credibility of the results.

In lieu of a complex cover story, before vending our post-treatment measures we asked subjects to "reflect" on the videos they had seen and rate how much they had enjoyed them.

Control over messages and context

In principle, messaging experiments could be designed so as to manipulate each aspect of the script, images, and messengers, as well as overall length and the context in which messages are presented. In practice, the number of possible variations becomes unmanageable. As is the case for any experimental agenda, some manipulations must be given priority over others. Because the persuasive effects of party-focused video messages have seldom, if ever, been explored systematically, a basic question for early testing is whether any partisan appeal can generate change. If so, a followup research program can go on to vary subtle aspects of presentation.

Because our primary concern was to create ads that would influence partisan attachments, we turned creative control over to professional advertisers. Before production, the advertisers discussed with us at length the leading theories of partisan change, which we expected to motivate their ads. But the advertising experts used their own creative talents to produce ads that they thought would be influential. What we sacrifice in fidelity to each ad's theoretical inspiration, we more than gain back in realism. The resulting treatment stimuli have the

look and feel of professionally-produced advertisement, bolstering the external validity of our study. For a full description of how these theories guided the development of our treatments, please see the supplemental materials.

As this study unfolded, it became increasingly apparent that the treatment effects were so subtle that we would need many thousands more experimental subjects than MTurk could provide in order to have the power to reliably detect the *relative* effectiveness of different theory-inspired ads. Because we find no robust evidence of message-specific effects, the analysis that follows focuses on the immediate and downstream effects of dosage, but interested readers can find the message-specific effects in the appendix.

Minimizing Priming and Demand Effects

A final design consideration is how to facilitate the collection of respondents' background information in a manner that does not contaminate the experiment (Klar, Leeper and Robison, 2020). If prior partisan attitudes are measured immediately before treatment, respondents may be tipped off to the purpose of the study, or they may answer follow-up questions in ways that maintain consistency with their pre-treatment answers. For this reason, pre-treatment measurement is best conducted in a prior wave of a panel survey, so that time can wash out priming or anchoring effects. Such concerns led us to conduct the studies described here over multiple waves, with pre-treatment party identification gathered in a pre-treatment wave.

Survey Experimental Design

All five studies follow a multi-wave panel survey experimental design. Four of the five studies (Studies 1-3 and Study 5) were conducted on Mechanical Turk (MTurk). Some scholars take a negative view of studies conducted on MTurk because the population of "workers" is quite different from the national population of American adults in the sense that the distribution

of demographic characteristics on MTurk is different from the national population. We chose to use MTurk because conducting multi-wave experiments on the platform is – by an order of magnitude – less expensive than on putatively "nationally representative" samples (Gross, Porter and Wood, 2018). Further, recent investigations have demonstrated that both Average Treatment Effect (ATE) estimates (Mullinix et al., 2015; Coppock, 2019) and Conditional Average Treatment Effect (CATE) estimates (Coppock, Leeper and Mullinix, 2018) obtained on probability and MTurk samples correspond quite tightly. On this basis, our guess is that the results we present below will generalize to other populations of Americans.

The remaining study (Study 4) was conducted on Lucid, an alternative source of online respondents that quota samples to match U.S. Census margins. Coppock and McClellan (2019) show that treatment effect estimates obtained on Lucid closely follow those obtained on both MTurk and probability samples. Lucid has some under some criticism due to a recent increase in inattention (Aronow et al., 2020), but we conducted our study before this increase appears to have taken place.

While the first two studies evaluated the effects of individual ads, studies 3, 4, and 5 investigated how subjects respond to various *dosages* of pro-Democratic advertisements. In study 3, all subjects saw a total of three videos. We randomly assigned whether subjects saw 0, 1, 2, or 3 treatment videos; the remainder were placebo videos (advertisements for paper towels or ice cream). In studies 4 and 5, subjects could be assigned to see 0, 3, or 6 treatment videos.

Recruitment

For the studies conducted on Mechanical Turk, we conducted wave 1 baseline surveys that measured demographic characteristics (age, gender, education, race and ethnicity) political variables (ideology, interest in politics, political knowledge) and pretreatment versions of our outcome variables (feeling thermometers towards various leaders and groups, spending

attitudes, and 7-point party ID). In Wave 2, conducted approximately 10 days later, subjects were randomly assigned to treatment or placebo conditions before responding to an outcome survey. In Wave 3, conducted approximately 10-15 days later, we administered a second outcome survey. In these waves, most of the survey questions were identical across the two outcome surveys (and across the two studies), though we did add and subtract a few questions across implementations. Finally, more than a year later, we recontacted all subjects in Studies 1-3 and Study 5 (i.e., all the Mechanical Turk studies) and queried them about their partisan identity once again. See the supplementary materials for the wordings of all survey items. Mechanical Turk respondents were paid $1.00 for each wave of their participation.

The design of Study 4 conducted on Lucid differed in that we did not conduct a pre-treatment survey wave. Instead, we measured pre-treatment background characteristics in the same survey that treatments were allocated and immediate outcomes were solicited. Due to a miscommunication, we only have a 3-point pre-treatment measure of party identification for this study, which results in slightly smaller precision gains, but no bias. Subjects were recontacted 10-15 days later for a second round of post-treatment measurement. We paid Lucid $1.00 per complete. Lucid subjects themselves are compensated in a variety of ways, including website points, gift cards, and cash, so we do not know the precise amount individual Lucid subjects were paid for their participation in our studies.

Table 1 summarizes the dates and sample sizes of all studies at each wave.

Table 1: Dates of implementation

	Wave 1: Collect demographics and baseline attitudes	Wave 2: Allocate treatments and collect post-treatment attitudes	Wave 3: Collect post-treatment attitudes again
Study 1 (MTurk)	April 30, 2018 (N = 1,548)	May 7 - 10, 2018 (N = 1,1329)	May 23 - 29, 2018 (N = 1,154)
Study 2 (MTurk)	June 6, 2018 (1,599)	June 15 - 21, 2018 (N = 1,359)	July 2 - 9, 2018 (N = 1,138)
Study 3 (MTurk)	October 15-17, 2018 (N = 4,946)	October 24-26, 2018 (3,369)	November 4 - 12, 2018 (N = 2,947)
Study 4 (Lucid)		September 14 - October 2, 2019 (4,631)	September 25 - October 7, 2019
Study 5 (MTurk)	October 16-20, 2019 (N = 2,071)*	October 30 - November 7, 2019 (2,593)	November 12 - 19, 2019 (N = 1,933)

Respondents' evaluation of the treatments

Immediately after treatment, we assessed subjects' enjoyment of the ads, on the assumption that subjects paying attention to the manipulation would respond along predictable partisan lines. Specifically, we asked: "How would you rate your enjoyment of the ads you just watched, on a scale of 1 to 5, with 1 being no enjoyment and 5 being the most enjoyment?"' As expected, Democrats reported enjoying the ads more than Republicans and Independents (see Appendix E), suggesting that the thrust of the ads was clear to participants.

Treatments

We collaborated with two advertising agencies to develop advertisement-length videos that drew upon the four theoretical paths to partisanship that we outlined above. Both firms have previously produced ads for sitting Democratic officials, including prominent federal legislators. All ads were intended to increase identification with the Democratic Party. As mentioned above, political advertising firms almost always work for only one party or the other; enlisting firms to make ads for both parties would have been unworkable. To the batch of videos created by the advertising firms for this project, we added a video produced by the DNC that combines aspects of all four theories. In the supplementary materials, we provide scripts and links to all videos.

Figure 1 displays screen shots from the ads produced and tested. Our tests of issue proximity highlighted those issues about which Democrats are most closely identified, such as climate change, gay rights and gun control. The Democratic Party, contends one such ad, represents "the issues we believe in." To further test this theory, we fielded multiple versions of the same issue ad, swapping out one of the particular issues emphasized for another.

For messages inspired by the persona theory, the ad-makers presented viewers with images and audio of party leaders, from Kennedy through Obama. Ads motivated by the

social groups theory feature images of young and diverse people effectively endorsing the Democratic Party. "People like you are Democrats," explains the narrator in one such ad, while another ad claims that the Democratic Party represents "all colors, all creeds, and all genders."

The ads related to economic performance relied on empirical evidence concerning the economy. These ads credited the Democratic Party for overseeing greater job growth than Republicans. One of the tested performance ads includes a graph displaying GDP growth under both parties since the Second World War. "25 million more jobs, 44% more growth," the ad concludes, before turning to the Democratic Party logo.

Table 2 shows the number of subjects assigned to each treatment condition for studies 1 and 2. Study 2 omitted ineffective videos from Study 1 and added one new video. Firm B produced two versions of an "issues" ad that discussed Democratic positions on climate change and marriage equality. Version 1 also included material about gun control and version 2 included material about drug decriminalization. The advertising firm produced a third version that did not include any additional issue content over and above climate change and marriage equality, so that we could gauge what, if any, special impact the gun control or drug decriminalization versions may have exerted.

In all three dosage experiments, we used the Performance, Charisma, Social Identity, and Issues ads from Firm B and the generic ad from the DNC. The precise videos that were shown to subjects were chosen via sampling without replacement from all six possibilities, depending on the dosage to which the subject was assigned. Table 3 shows the number of subjects who saw each video, by dosage condition.

Table 2: Assignment of Treatments to Subjects, Studies 1 & 2

	Study 1	Study 2
No video	142	256
Issues (Firm A)	156	
Performance (Firm A)	108	240
Charisma (Firm A)	118	
Social Identity (Firm A)	140	
Generic (DNC)	126	250
Performance (Firm B)	111	
Charisma (Firm B)	154	
Social Identity (Firm B)	132	
Issues (Firm B, version 1)	132	259
Issues (Firm B, version 2)	127	230
Issues (Firm B, version 3)		249

Cell entries are numbers of subjects assigned to each condition.

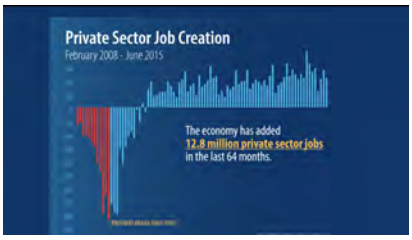
Table 3: Assignment of Treatments to Subjects, Studies 3, 4, & 5

Dosage	N	Charisma	Social Identity	Performance	Issues	Generic	Placebo
0	870	0	0	0	0	0	870
1	846	144	122	147	134	151	846
2	851	273	312	259	299	289	851
3	859	425	423	447	429	425	0
0	1550	0	0	0	0	0	1550
3	1581	786	789	800	800	779	0
6	1500	1500	1500	1500	1500	1500	0
0	878	0	0	0	0	0	878
3	891	471	427	431	450	465	0
6	894	894	894	894	894	894	0

Cell entries are numbers of subjects who saw each video.

Figure 1: Screen shots of the ads produced and tested.

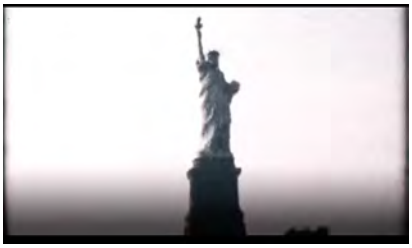
(a) Issues (Firm A)



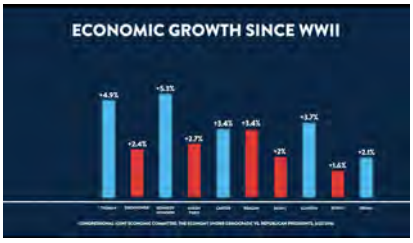
(b) Performance (Firm A)

(c) Charisma (Firm A)

(d) Social Identity (Firm A)

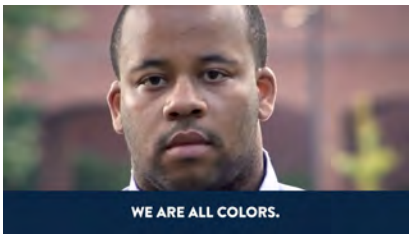


(e) Generic (DNC)



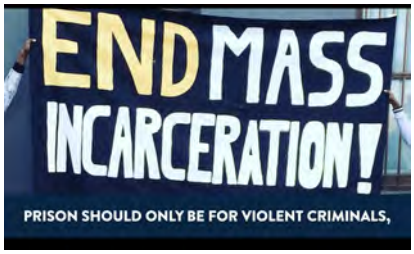
(f) Performance (Firm B)

(g) Charisma (Firm B)



(h) Identity (Firm B)

(i) Issues (Firm B, V1)



(j) Issues (Firm B, V2)

(k) Issues (Firm B, V3)

(l) Negative Partisanship (Firm B)

Outcome measurement

We measure five outcome variables. Party identification is measured using the standard branching 7-point American National Election Study format. We reverse code this variable so that higher values indicate stronger attachments to the Democrats. The placebo group's distribution for this measure (and all our outcome variables) may be found in Figure 2, which shows the overrepresentation of Democrats typical of Mechanical Turk samples (Huff and Tingley, 2015). In addition to the commonly used seven-point Party ID scale, we also deploy other measures of party identification that have been proposed and validated in previous research. Green and Schickler (1993) show that a self-labeling exercise produces a valid and reliable indicator of net identification with Democrats versus Republicans.[3] The resulting measure ranges from -9 to 9. Among placebo group subjects, the correlation between this measure and the seven-point party identification scale is 0.86.

Another measure of partisan social identity, drawing on the work of Iyengar, Sood and Lelkes (2012), assesses the social distance that partisans of one party feel toward partisans of the other party. This measure asks subjects if they would be "happy" if a member of their immediate family married either a Democrat or a Republican. The net response for the two questions creates a scale ranging from -2 to 2. Given that this is just one component of a social distance scale, it is not surprising that the correlation with the seven-point party identification scale is just 0.56. In part, the problem is that the scale is unable to detect gradations of partisanship, and the majority of respondents have a net score of zero. Still, it remains to be seen whether messages that stir up partisan sentiments have detectable effects on feeling of social distance.

Our final two measures gauge partisan sentiment rather than party identification per se. One measure is the "feeling thermometer" ratings of the Democratic party and Republican

[3]Respondents were also asked about other labels, in random order: an evangelical Christian, a feminist, a Southerner, and an environmentalist.

party commonly used in the American National Election Studies. Another measure that is similar in spirit but perhaps more straightforward in its wording asks subjects for their evaluation of each party on a seven-point Likert scale. As expected, these two measures are highly correlated with one another (0.91) and highly correlated with party identification (0.81 and 0.83, respectively), and in these samples the balance of evaluations favors the Democratic party.

For the dependent variables that are labelled "Net," we take the difference between the Democratic response and the Republican response. Figure 2 shows the distributions and bivariate correlations of these five outcome measures. Here we give exact wording of all five outcomes and indicate in which studies and waves they were asked. Because we added and subtracted some survey items across studies, the precise set of outcome measures is not identical in each instance.

- *Seven-point Party ID (reverse coded)* Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent, or what?

  Would you call yourself a strong Republican or a not very strong Republican?

  [if Republican or Democrat:] Would you call yourself a strong Democrat or a not very strong Democrat?

  [if Independent:] Do you think of yourself as closer to the Republican Party or to the Democratic party?

  [1: strong Republican, 7: strong Democrat; asked in all studies and waves]

- *Net Feeling Thermometer* Please use the sliders to indicate how warm or favorably you feel towards these political groups. Ratings between 50 degrees and 100 degrees mean that you feel favorable and warm toward the person. Ratings between 0 degrees and 50 degrees mean that you don't feel favorable toward the person and that you don't

care too much for that person. You would rate the person at the 50 degree mark if you don't feel particularly warm or cold toward the person. [-100 to 100; asked in all studies and waves]

- *Net Favorability* On a scale from 1 to 7, with 1 being the least favorable and 7 being the most favorable, how would you rate the [Republican/Democratic] party? [-14 to 14; asked in all studies and waves]

- *Net Self-Description* On a scale from 1 to 10, where '10' represents a description that is perfect for you, and '1' a description that is totally wrong for you, how well do each of the following describe you? [-10 to 10; not asked in study 1, wave 3]

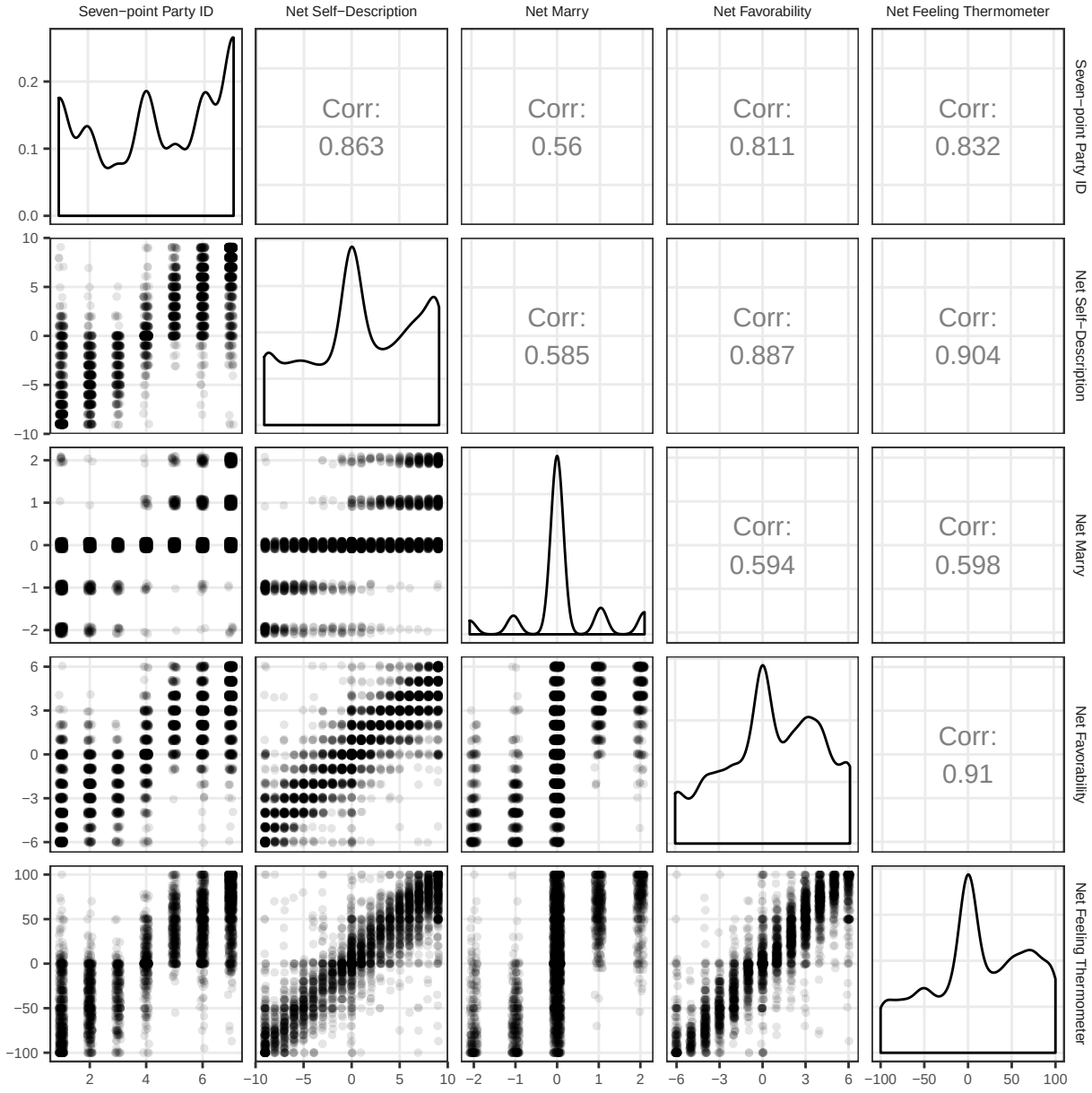
We aggregate these five measures into an index using principal components analysis; we report treatment effects on this outcome measure in the main text and effects on each separate outcome variable in the appendix.

In addition to these measures of partisanship, we also estimate the downstream effects our treatments may have had on vote preference for Democratic candidates. For studies conducted before the 2018 midterm elections, we asked "If the 2018 midterm elections were held today, would you want to see the Republican or Democratic Party win control of the House of Representatives?" After the midterms, the question was changed to "If the 2020 presidential election were held today, would you want to see the Republican or Democratic Party win control of the White House?" Finally, we also consider evaluations of Donald Trump assessed via feeling thermometer.

Analysis

We conduct all analyses among subjects who responded in all three waves in order to hold sample composition constant when comparing instantaneous outcomes and delayed outcomes. We estimate ATEs for each specific dosage by regressing the outcome on a vector of

Figure 2: Outcome Measures



Source: Pooled placebo group subjects excluding study 3 (Total N = 4,280). A very small amount of jitter has been added to each point to visually separate each observation.

treatment indicators. We also estimate regressions of the form

$$Y_i = \beta_0 + \beta_1 D_i + \beta_2 X_i + \epsilon$$

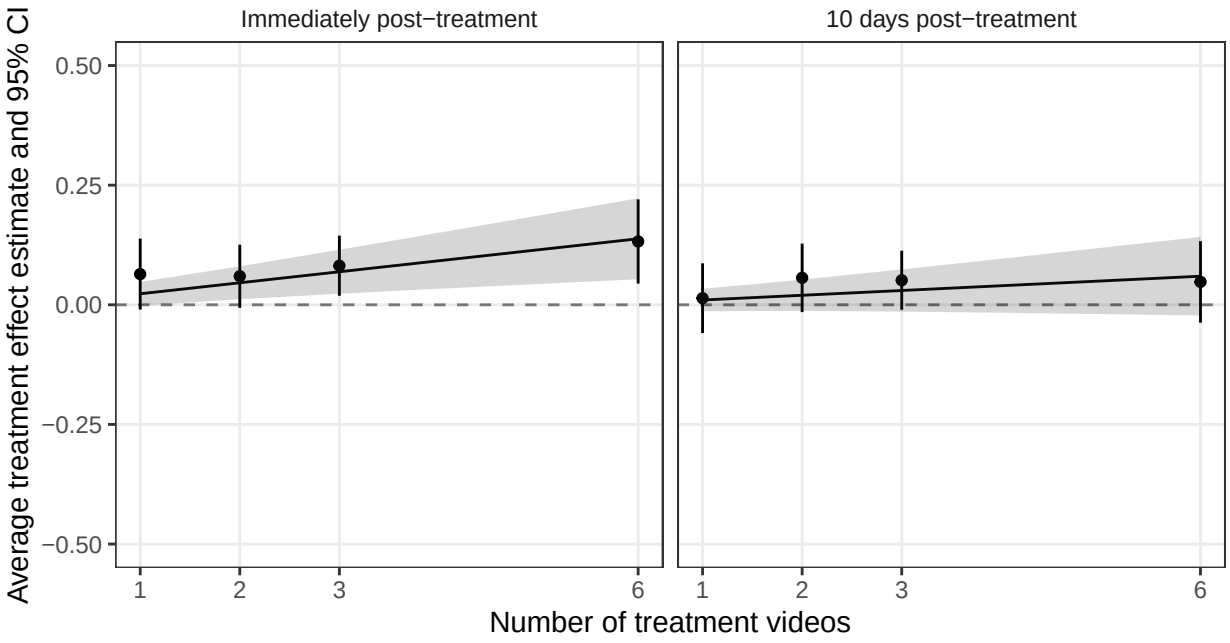
where D_i is the dosage of videos the subject was assigned to watch and X_i is party identification, measured pre-treatment. Because the coefficient β_1 is the best linear approximation to the dose-response curve, it serves as a useful summary of the average marginal effect of one additional video. Of course, the dose-response curve need not be exactly linear. In the supplementary materials, we also estimate effects separately by dosage level with substantively similar conclusions; within narrow range of 0 to 6 videos, we do not observe obvious nonlinearities.

Detecting Advertising Effects on Party Identification

Figure 3 shows how the dosage of pro-Democratic Party advertising affects the 7-point party identification measure. The point estimates (represented by the black circles) indicate that a single ad moves respondents by an average of 0.063 scale points, with a 95% confidence interval ranging from -0.011 to 0.137. The average effects of two or three ads are similar. It is only when the dosage reaches six ads that the upward trajectory becomes apparent, for here the estimated effect reaches 0.151 points with a 95% confidence interval ranging from 0.064 to 0.238. It seems clear that exposure to a series of ads extolling the Democratic Party has statistically significant effects on party identification measured a few minutes later. Regression estimates the average marginal increase per ad to be 0.025 points, which in this case yields a p-value of 0.0001.

Returning to these respondents two weeks later reveals a treatment effect that has decayed to roughly half its original size. The point estimate for a single ad is now 0.01, climbing to

Figure 3: Effects of advertising dosage on post-treatment 7-point party identification

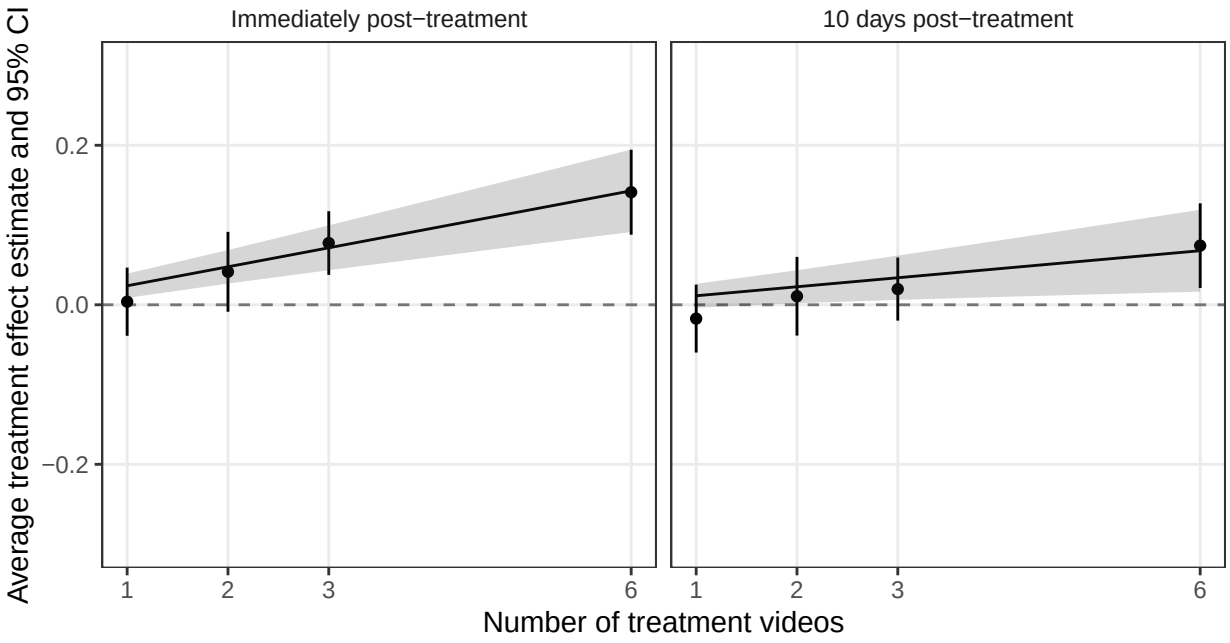


0.06 for two ads, 0.05 for three ads, and 0.07 for six ads. Regression estimates the average marginal increase per ad to be 0.010, which has a p-value of 0.15. Party identification is clearly movable, but it does not move much even after a salvo of six ads.

Other measures of partisanship show similar patterns. To conserve space, we relegate to the appendix our analysis of each specific partisanship measure. Here we report the results of an standardized index of all of the measures. The left panel of Figure 4 reaffirms the substantial and statistically significant effects of advertising dosage on party identification, measured shortly after treatment. The right panel of this figure again shows that several days later the regression slope declines by about half.

If the half-life of the advertising effect is one week, it comes as no surprise that a year later, the apparent effects have disappeared entirely. Regardless of how we measure party identification, the long-term effects appear to be zero, as respondents return to their pre-treatment baseline. For example, in studies 1 and 2, the pre-treatment measure of party ID

Figure 4: Effects of advertising dosage on post-treatment party ID index scale



shows a correlation of 0.95 with party ID measured immediately after treatment – which has perturbed partisan attachments in the treatment group – and 0.96 roughly two weeks later, after party ID has begun to return to its usual levels. A year later, the correlation with pre-treatment party ID is 0.87 for study 1 and 0.89 for study 2. This is the familiar pattern of gradual decay. The main difference between our panel results and those reported in the many panel surveys that have traced over-time correlations in partisanship is that the ours reveal the effects of a randomly induced perturbation at a particular point in time.

Detecting Advertising Effects on Vote Preferences and Presidential Ratings

With a cleanly-identified effect of an intervention on partisanship, we have a rare opportunity to see whether outcomes that are believed to be affected by party identification move as well. These so-called downstream outcomes are especially interesting theoretically *insofar as the*

treatments themselves do not refer to them directly. For example, the ads do not encourage audiences to vote for Democrats in the coming election, nor do they mention the incumbent president or give reasons for disapproving of his performance or policies. For partisanship, the key outcomes dating back to *The American Voter* (1960) are voting preferences and evaluation of the president, both of which present respondents with partisan stimuli.[4]

Figure 5 shows the relationship between ad dosage and vote preferences in the coming election. Although a single ad has no apparent effect, the effects grow larger with dosage; among respondents exposed to six ads, the probability of preferring Democrats is six percentage points higher than among the placebo group. In other words, the same high dosage intervention that shifts party identification produces a concomitant change in vote preference. As the right panel of Figure 5 suggests, when the intervention's effects on party identification fade, so too do their effects on vote preferences.

Figure 6 indicates that a similar pattern applies to ratings of President Trump. In keeping with other experiments showing that a single ad that directly attacks Trump's record in office has little effect on vote preference (Broockman and Kalla, 2020), a single party-centered ad seems to have negligible effect. However, increasing dosage does appear to diminish thermometer ratings of Trump. After two or three ads, the point estimate hovers around a one point reduction; after six ads, the point estimate is approximately -2.5 points. The downward slope is estimated with considerable sampling uncertainty but conforms to expectations: the same intervention that increases Democratic identification also lowers approval of the Republican incumbent. Although we lack the statistical precision to reliably detect declining effects over time, it appears that the effect on presidential ratings subsides, as expected.

[4]We assess effects on downstream outcome following the "intention-to-treat" principle rather than with an instrumental variables setup or formal meditation analysis, since those methods rely on assumptions (excludability for IV and sequential ignorability for mediation) that we do expect to hold in this setting.

Figure 5: Effects of advertising dosage on post-treatment vote preference

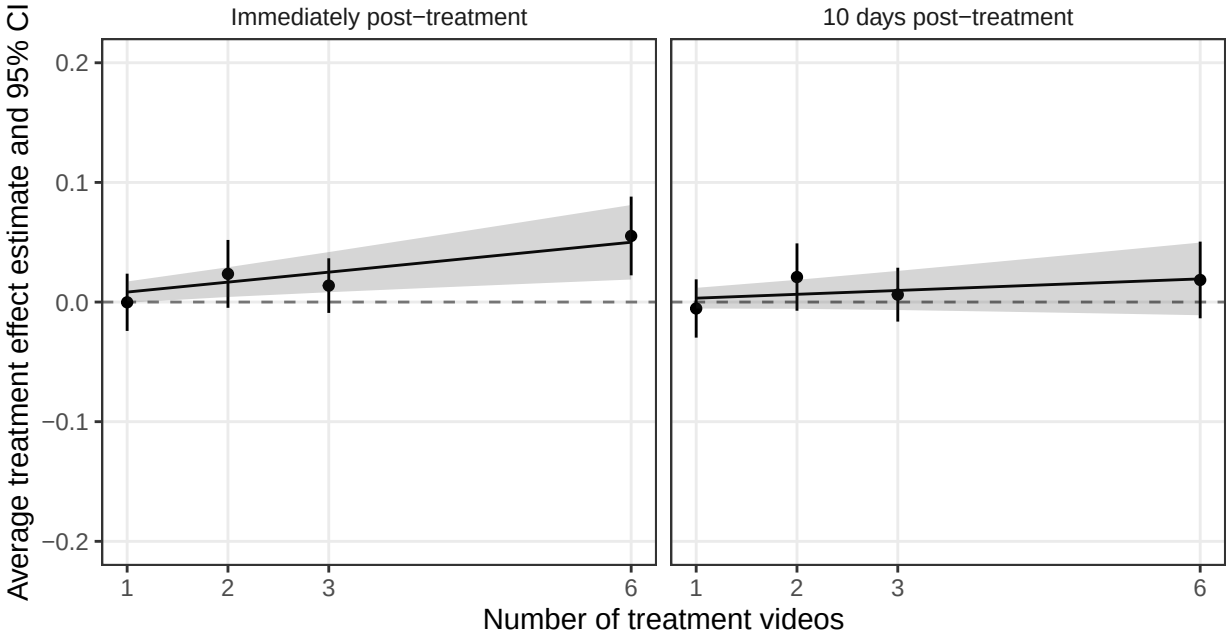
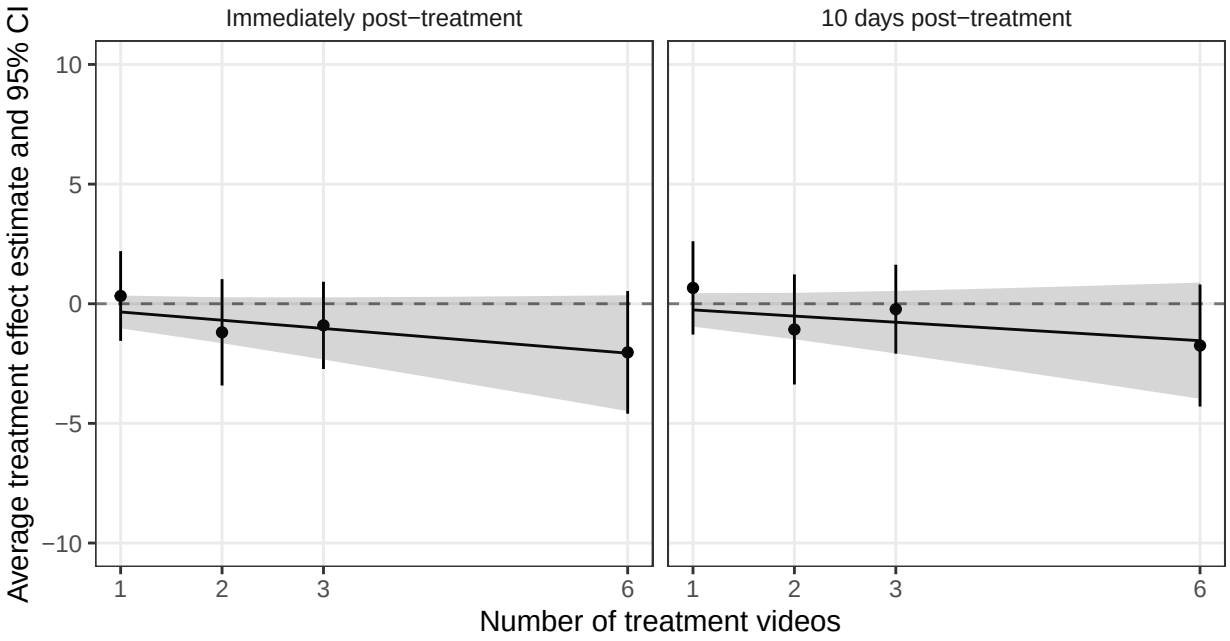


Figure 6: Effects of advertising dosage on post-treatment 101-point Trump feeling thermometer



Heterogeneous Effects?

A prominent theme in the literature on partisan change is the hypothesis that those with weaker partisan attachments are more susceptible to influence. As discussed above, this hypothesis points to two groups that are especially prone to partisan change: young people and independents.

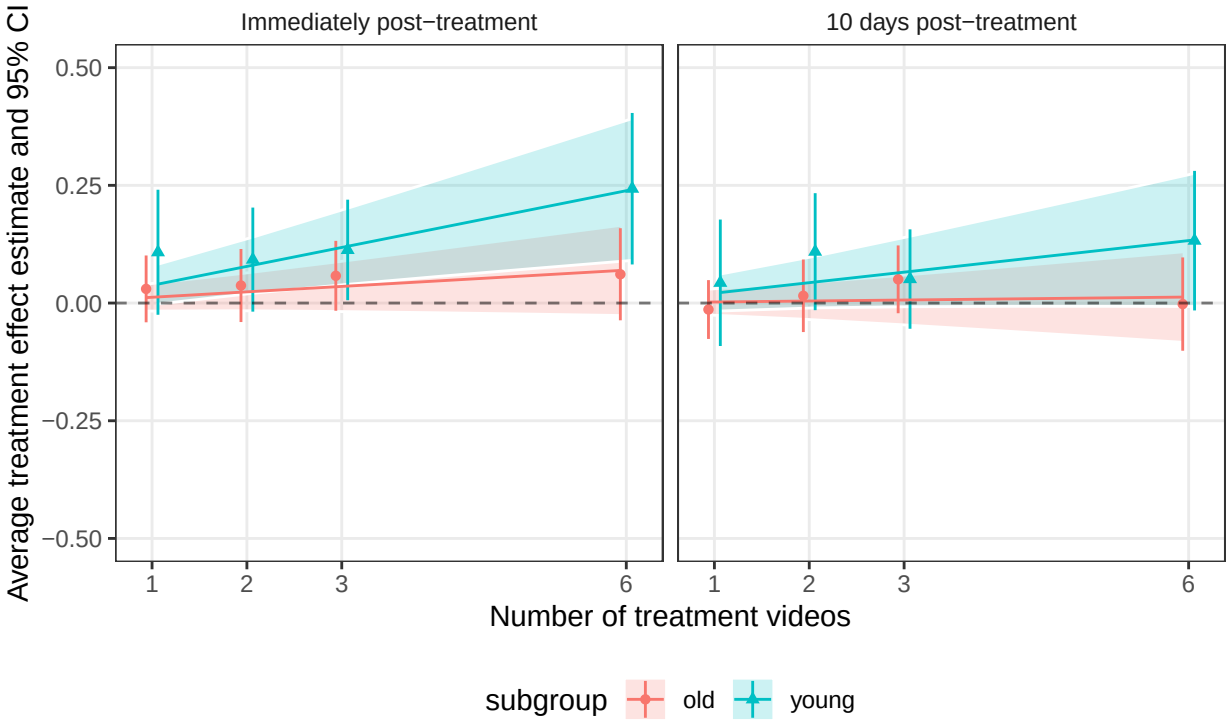
To test for treatment-by-covariate interactions, we partition our data into younger (18-35, N=4,533) versus older (36+, N=3,758) subgroups and by three categories of pre-treatment partisanship (Democrats, Republicans, and Independents).[5] The effects of dosage for each age and partisan subgroup are displayed graphically in Figure 7, and corresponding regression results are presented in the supplementary materials.

As expected, younger respondents were more responsive to messaging than their older counterparts. The average marginal effect of an additional ad is 0.040 points (SE: 0.013, $p = 0.001$) more effective for young subjects in the immediate wake of the treatment. Ten days later, young people continue to show more responsiveness (especially at high dosages), and the treatment-by-covariate interaction remains significant (estimate = 0.025, SE = 0.012, $p = 0.043$).

Examining responsiveness to messaging across partisan groups, we also see the expected pattern by which independents are most strongly influenced initially and after ten days. Relative to the effect for Democrats, the marginal effect for independents is 0.033 points larger (SE: 0.017, $p = 0.060$) immediately, but this difference shrinks to 0.007 points after 10 days. Somewhat unexpected are the results for Republicans and Democrats. Republicans move initially but show no signs of influence ten days later. Democrats, on the other hand, are unresponsive initially but (if anything) show hints of movement when reinterviewed. It

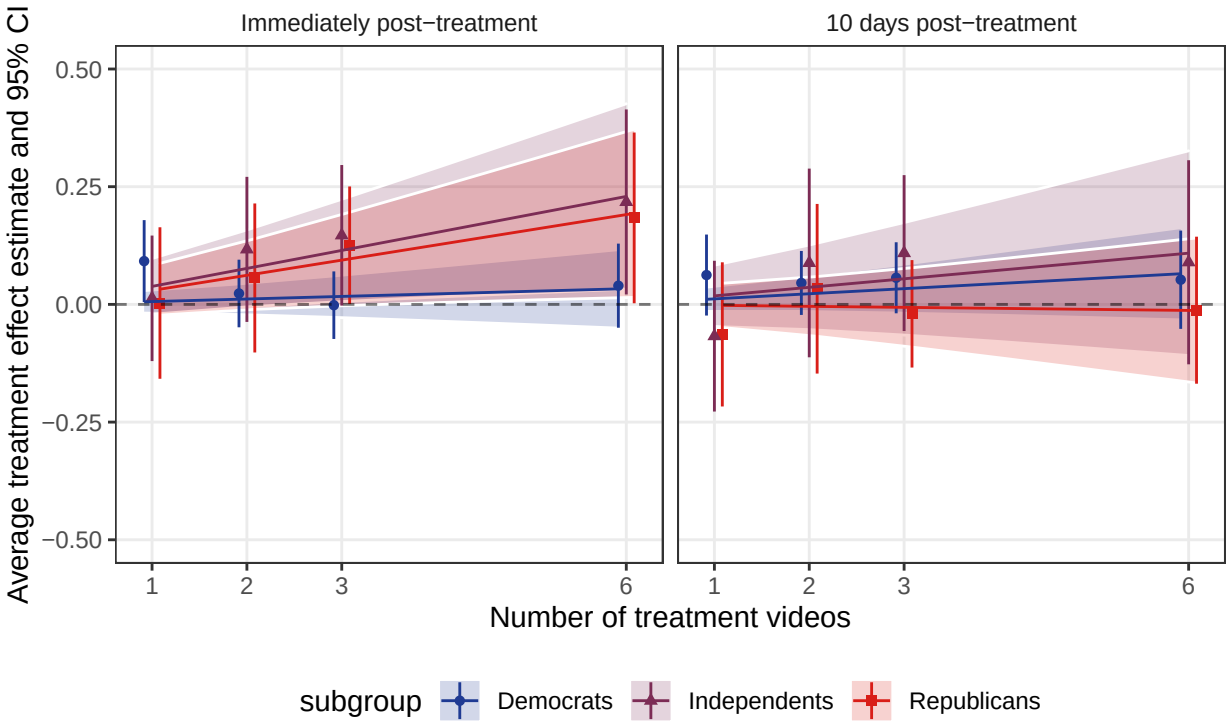
[5] Age was measured in ranges, with the lowest two ranges encompassing 18 to 24 and 25 to 34. We pooled these groups in order to dampen sampling variability, but the full results for each age range may be found in the supplementary materials.

Figure 7: Effects of advertising dosage on post-treatment 7-point party identification by age group



may be that out-partisans are initially won over by messages from the other side, to which they might not ordinarily be exposed, but snap back to their prior partisan orientation when they return to their social milieu.

Figure 8: Effects of advertising dosage on post-treatment 7-point party identification by pre-treatment partisanship



Discussion

The experimental study of party identification is a challenging and largely uncharted research endeavor. The discouraging results presented by Cowden and McDermott (2000) seem to have convinced a generation of scholars that inducing partisan change in a lab or survey context is hopeless. Our attempt to revive this line of research drew persuasive messages from an array of theoretical perspectives; our innovation is to convey these messages using recorded media so that the intervention would be both reproducible and realistic. Like Gerber, Huber and Washington (2010), we sought to assess the effects of the intervention both on party identification and evaluations that party identification is believed to influence. In addition, we use multiple waves of outcome measurement to assess the rate at which party attachments re-equilibrate in the wake of a short-term perturbation.

We find that a preponderance of pro-Democratic messages made the Democratic Party a more attractive object of identification. To be sure, the effects are not large, especially by the standards of the literature on partisan realignment. But these movements are large enough to detect reliably at high dosages. And the fact that high-dosage treatments also change vote preference and presidential approval lends support to the notion that identification with a political party is indeed a cause of such evaluations. Consistent with this interpretation is the fact that as effects on party identification fade ten days later, so too do the effects on vote preferences.

An important micro-level insight into aggregate partisan change is the mincing steps by which partisanship changes in the wake of messages favoring one party over the other. A key property of American macropartisanship is that the balance of partisanship adjusts very gradually, even in the wake of abrupt reversals in party fortunes and economic conditions. Partisan identification may thought of a persistent orientation that is susceptible to subtle but discernible change. While the effects of any one message may be small, the cumulative

effects of many such messages that disproportionately favor one party may be enough to affect the distribution of party attachments in the aggregate.

Finally, the experimental results shed light on longstanding debates about whether party identification reverts to its mean in the wake of some kind of exogenous shock. In the aggregate, the macropartisan balance reverts to its mean quite slowly (Box-Steffensmeier and Smith, 1996), whereas individual-level partisanship, as we have seen, sheds half the effect of a shock over the course of two weeks. The gradual undulation of the aggregate series may reflect the fact that identification-altering events and the discourse surrounding them unfold gradually over time, repeatedly "treating" individuals who would otherwise revert to their long-term means. This characterization in some ways splits the difference between scholars who emphasize partisanship's stability and those who emphasizes its potential for change.

These findings are hardly the last word on the determinants of partisan identification. Thanks to the role that professional ad makers played in their production, our treatments achieved a high level of realism, but they were delivered in the artificial and controlled environment of the survey experiment. Field experiments that deploy partisan messages on television and over the Internet, with outcome measurements collected periodically on seemingly unrelated surveys, would provide evidence of greater external validity (Broockman, Kalla and Sekhon, 2017). Similarly, as one implication of this work is that higher dosage yields sharper, longer-lasting effects, scholars should investigate the effects of dosages well beyond those studied here. Longer panel studies, with subjects exposed to more treatments, would allow researchers to better approximate how partisan messages are consumed in the real world. In addition, better understanding of how age and weak partisanship interact with treatment effects would require oversampling segments of the population that are expected to be most responsive to messages designed to instil partisanship. Finally, the research agenda described here may be fruitfully deployed across party systems that vary in terms of the number and durability of parties as well as the social and ideological bases of party support.

References

Achen, Christopher H. 1975. "Mass political attitudes and the survey response." *The American Political Science Review* 69(4):1218–1231.

Aronow, Peter Michael, Joshua Kalla, Lilla Orr and John Ternovski. 2020. "Evidence of Rising Rates of Inattentiveness on Lucid in 2020." *SocArXiv* .

Asher, Herbert B. 1974. "Some consequences of measurement error in survey data." *American Journal of Political Science* pp. 469–485.

Bolsen, Toby, James N. Druckman and Fay Lomax Cook. 2014. "The Influence of Partisan Motivated Reasoning on Public Opinion." *Political Behavior* 36:235–262.

Box-Steffensmeier, Janet M. and Renée M. Smith. 1996. "The Dynamics of Aggregate Partisanship." *American Political Science Review* 90(3):567–580.

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

Broockman, David E., Joshua L. Kalla and Jasjeet S. Sekhon. 2017. "The Design of Field Experiments With Survey Outcomes: A Framework for Selecting More Efficient, Robust, and Ethical Designs." *Political Analysis* 25(4):435–464.

Broockman, David and Joshua Kalla. 2020. "When and Why Are Campaigns' Persuasive Effects Small? Evidence from the 2020 US Presidential Election.".

Brown, Jacob R. and Ryan D. Enos. 2021. "The measurement of partisan sorting for 180 million voters." *Nature Human Behaviour* .
URL: *https://doi.org/10.1038/s41562-021-01066-z*

Chong, Dennis and James N. Druckman. 2010. "Dynamic Public Opinion: Communication Effects over Time." *American Political Science Review* 104(04):663–680.

Clinton, Joshua D. and John S. Lapinski. 2004. ""Targeted" Advertising and Voter Turnout: An Experimental Study of the 2000 Presidential Election." *Journal of Politics* 66(1):69–96.

Converse, Philip E. 1964. The Nature of Belief Systems in Mass Publics. In *Ideology and Discontent*, ed. David E. Apter. Vol. 18 Glencoe, Illinois: Free Press.

Coppock, Alexander. 2019. "Generalizing from Survey Experiments Conducted on Mechanical Turk: A Replication Approach." *Political Science Research and Methods* 7(3):613–628.

Coppock, Alexander and Oliver A. McClellan. 2019. "Validating the Demographic, Political, Psychological, and Experimental Results Obtained from a New Source of Online Survey Respondents." *Research & Politics* 6(1):1–14.

Coppock, Alexander, Seth J. Hill and Lynn Vavreck. 2020. "The small effects of political advertising are small regardless of context, message, sender, or receiver: Evidence from 59 real-time randomized experiments." *Science Advances* 6(36).
URL: *https://advances.sciencemag.org/content/6/36/eabc4046*

Coppock, Alexander, Thomas J. Leeper and Kevin J. Mullinix. 2018. "Generalizability of heterogeneous treatment effect estimates across samples." *Proceedings of the National Academy of Sciences* 115(49):12441–12446.

Cowden, Jonathan A. and Rose M. McDermott. 2000. "Short-term Forces and Partisanship." *Political Behavior* 22(3):197–222.

Fiorina, Morris. 1981. *Retrospective Voting in American National Elections*. New Haven: Yale University Press.

Fowler, Erika Franklin, Travis N. Ridout and Michael M. Franz. 2016. "Political Advertising in 2016: The Presidential Election as Outlier?" *The Forum* 14(4):445–469.
URL: *https://doi.org/10.1515/for-2016-0040*

Franklin, Charles H. and John E. Jackson. 1983. "The Dynamics of Party Identification." *American Political Science Review* 77(4):957–973.

Gerber, Alan S., Gregory A. Huber and Ebonya Washington. 2010. "Party Affiliation, Partisanship, and Political Beliefs: A Field Experiment." *American Political Science Review* 104(4):720–744.

Green, Donald P. and Bradley Palmquist. 1990. "Of Artifacts and Partisan Instability." *American Journal of Political Science* 34(3):872–902.

Green, Donald P., Bradley Palmquist and Eric Schickler. 2002. *Partisan Hearts and Minds*. New Haven: Yale University Press.

Green, Donald P. and Eric Schickler. 2009. A Spirited Defense of Party Identification Against its Critics. In *Political Parties and Partisanship*, ed. John Bartle and Paolo Bellucci. Routledge pp. 180–199.

Green, Donald Philip and Bradley Palmquist. 1994. "How stable is party identification?" *Political Behavior* 16(4):437–466.

Green, Donald Philip and Eric Schickler. 1993. "Multiple-measure Assessment of Party Identification." *Public Opinion Quarterly* 57(4):503–535.

Gross, Kimberly, Ethan Porter and Thomas J. Wood. 2018. "Identifying Media Effects Through Low-Cost, Multiwave Field Experiments." *Political Communication* 36(2):1–16.

Harris, Louis. 1954. *Is There a Republican Majority?: Political Trends, 1952-1956*. New York: Harper.

Hatemi, Peter K. and Christopher Ojeda. 2020. "The Role of Child Perception and Motivation in Political Socialization." *British Journal of Political Science* pp. 1–22. Forthcoming.

Huff, Connor and Dustin Tingley. 2015. ""Who are these people?" Evaluating the demographic characteristics and political preferences of MTurk survey respondents." *Research & Politics* 2(3):1–12.

Hyman, Herbert H. 1959. *Political Socialization: A Study in the Psychology of Political Behavior*. Glencoe, Illinois: Free Press.

Iyengar, Shanto, Gaurav Sood and Yphtach Lelkes. 2012. "Affect, Not Ideology: A Social Identity Perspective on Polarization." *Public Opinion Quarterly* 76(3):405–431.
URL: *https://doi.org/10.1093/poq/nfs038*

Jackson, John E. 1975. "Issues, Party Choices, and Presidential Votes." *American Journal of Political Science* 19(2):161–185.

Jacobson, Gary C. 2016. "The Obama legacy and the future of partisan conflict: Demographic change and generational imprinting." *The Annals of the American Academy of Political and Social Science* 667(1):72–91.

Jennings, M. Kent and Gregory B. Markus. 1984. "Partisan Orientations over the Long Haul: Results from the Three-Wave Political Socialization Panel Study." *American Political Science Review* 78(4):1000–1018.

Jennings, M. Kent, Gregory B. Markus, Richard G. Niemi and Laura Stoker. 2005. *Youth-Parent Socialization Panel Study, 1965-1997: Four Waves Combined*. Inter-university Consortium for Political and Social Research.

Jennings, M Kent, Laura Stoker and Jake Bowers. 2009. "Politics across generations: Family transmission reexamined." *The Journal of Politics* 71(3):782–799.

Kahn, Kim Fridkin and John G. Geer. 1994. "Creating Impressions: An Experimental Investigation of Political Advertising on Television." *Political Behavior* 16(1):93–116.
URL: *http://www.jstor.org/stable/586484*

Kepplinger, Hans Mathias, Wolfgang Donsbach, Hans-Bernd Brosius and Joachim Friedrich Staab. 1989. "Media Tone And Public Opinion: A Longitudinal Study Of Media Coverage And Public Opinion On Chancellor Kohl." *International Journal of Public Opinion Research* 1(4):326–342.

Klar, Samara, Thomas Leeper and Joshua Robison. 2020. "Studying Identities with Experiments: Weighing the Risk of Posttreatment Bias Against Priming Effects." *Journal of Experimental Political Science* 7(1):56–60.

Kleinnijenhuis, Jan, Anita M. J. van Hoof, Dirk Oegema and Jan A. de Ridder. 2006. "A Test of Rivaling Approaches to Explain News Effects: News on Issue Positions of Parties, Real-World Developments, Support and Criticism, and Success and Failure." *Journal of Communication* 57(2):366–384.

Levendusky, Matthew. 2009. *The Partisan Sort: How Liberals Became Democrats and Conservatives Became Republicans*. Chicago Studies in American Politics Chicago, Illinois: University of Chicago Press.

MacKuen, Michael B., Robert S. Erikson and James A. Stimson. 1989. "Macropartisanship." *American Political Science Review* 83(4):1125–1142.

McCombs, Maxwell. 2014. *Setting the agenda: Mass media and public opinion*. 2nd ed. Cambridge, UK: Polity Press.

Mullinix, Kevin J., Thomas J. Leeper, James N. Druckman and Jeremy Freese. 2015. "The Generalizability of Survey Experiments." *Journal of Experimental Political Science* 2:109–138.

Niemi, Richard G. and M. Kent Jennings. 1991. "Issues and inheritance in the formation of party identification." *American Journal of Political Science* 35(4):970–988.

Petty, Richard and John Cacioppo. 1986. The Elaboration Likelihood Model of Persuasion. In *Advances in Experimental Psychology*, ed. Leonard Berkowitz. Vol. 19 The name of the publisher pp. 123–205.

Schwarz, Norbert. 2004. "Metacognitive Experiences in Consumer Judgment and Decision Making." *Journal of Consumer Psychology* 14(4):332–348.
URL: *https://www.sciencedirect.com/science/article/pii/S1057740804701606*

Sears, David O. 1983. "The persistence of early political predispositions: The roles of attitude object and life stage." *Review of personality and social psychology* 4(1):79–116.

Segijn, Claire M., Hilde A. M. Voorveld, Lisa Vandeberg, Sjoerd F. Pennekamp and Edith G. Smit. 2017. "Insight into everyday media use with multiple screens." *International Journal of Advertising* 36(5):779–797.
URL: *https://doi.org/10.1080/02650487.2017.1348042*

Tucker, Patrick D, Jacob M Montgomery and Steven S Smith. 2019. "Party Identification in the Age of Obama: Evidence on the Sources of Stability and Systematic Change in Party Identification from a Long-Term Panel Survey." *Political Research Quarterly* 72(2):309–328.

Unnava, H. Rao and Robert E. Burnkrant. 1991. "Effects of Repeating Varied Ad Executions or Brand Name Memory." *Journal of Marketing Research* 28(4):406–416.
URL: *http://www.jstor.org/stable/3172781*