The Polarizing Effects of Online Partisan Criticism: Evidence from Two Experiments

Elizabeth Suhay, Emily Bello-Pardo, and Brianna Maurer

American University

July 19, 2017

Abstract

Affective and social political polarization—a dislike of political opponents and a desire to avoid their company—are increasingly salient and pervasive features of politics in many Western democracies, particularly the U.S. One contributor to these related phenomena may be increasing exposure to online political disagreements in which ordinary citizens criticize, and sometimes explicitly demean, opponents. This article presents two experimental studies that assessed whether U.S. partisans' attitudes became more prejudiced in favor of the in-party after exposure to online partisan criticism. In the first study, we draw on an online convenience sample to establish that partisan criticism that derogates political opponents increases affective polarization. In the second, we replicate these findings with a quasi-representative sample and extend the pattern of findings to social polarization. We conclude that online partisan criticism likely has contributed to rising affective and social polarization in recent years between Democrats and Republicans in the U.S. and perhaps between partisan and ideological group members in other developed democracies as well. We close by discussing the troubling implications of these findings in light of continuing attempts by autocratic regimes and other actors to influence democratic elections via false identities on social media.

Introduction

Academics, journalists, politicians, and citizens are preoccupied with growing partisan and ideological polarization in many democratic nations. While the U.S. is not particularly remarkable in its level of political polarization among elected officials (Hopkins & Sides 2015), recent scholarship has documented a sharp increase over the last several decades in polarization in the American *public* (Hetherington 2009; Levendusky 2009), particularly with respect to interparty prejudice (Iyengar, Sood, & Lelkes 2012; Iyengar & Westwood 2015; Mason 2015). In the contemporary U.S., most Democrats and Republicans express unfavorable views of the other party, with mutual levels of contempt intensifying in recent years (Pew Research Center 2014, 2016a). Scholars argue that these feelings of animosity are rooted to a significant degree in the fact that party identities are *social* identities. The dynamics of social identity contribute to partisans' intense dislike of their opposition (*affective* polarization) (Iyengar, Sood, & Lelkes 2012; Iyengar & Westwood 2015; Mason 2016), which tends to coincide with a motivation and tendency to avoid their company (what we refer to as *social* polarization).[1]

Paralleling the rise of affective and social partisan polarization has been a precipitous growth in online communication tools that facilitate political discourse. Whether through Facebook, Twitter, online discussion forums, or comment sections of more traditional news websites,

---

[1] There is some debate over the best terminology to use in describing these related processes, with more agreement over the term *affective polarization* than *social polarization* (see, for example, Mason 2015, Footnote 1). In our read of the literature, using the two terms in this way best describes and distinguishes the relevant processes, which together we refer to as *affective/social polarization*. While our study does not address the causal relationship between these two phenomena, we assume that affective polarization tends to precede social—given that attitudes tend to precede relevant behavioral inclinations in general.

individuals increasingly use online tools to voice their opinions about political and other topics (e.g., Margetts, John, Hale, & Yasseri 2015; Pew 2016b; Rainie & Wellman 2012). Several aspects of this new media landscape have the potential to contribute to partisan polarization, including high levels of opinionation, criticism of political adversaries, and uncivil discourse (Anderson, Brossard, Scheufele, Xenos, & Ladwig 2014; Levendusky 2013). Yet, few scholars have examined whether these characteristics of online discussions contribute to increasing affective and/or social polarization between partisan—and potentially other—political groups.

Drawing on data from two original experimental studies with U.S. citizens, we seek to understand whether partisan criticism online exacerbates affective and social polarization. In the first study, we establish that partisan criticism that derogates political opponents increases affective polarization in an online convenience sample. In the second study, using an improved sample and alternate design, we replicate and extend these results.

We focus on the U.S. case for two main reasons. First, theories of affective and social polarization were developed primarily while studying this case, and these phenomena have been extensively documented with U.S. data. In this article, our main advance is to extend this work by carefully testing whether affective and social polarization are affected by exposure to partisan criticism online. Second, there are several characteristics of U.S. politics that make online affective/social polarization likely: a political sphere marked by intense conflict, a large proportion of the electorate communicating online, and a citizenry with strong partisan identities. Of course, these characteristics are not unique to the U.S. We say more about the possibilities for future research on this topic in other nations in the Discussion and Conclusion section.

Theoretical Framework

Affective and Social Polarization

The topic of political polarization is well-traveled among scholars of democratic politics, particularly those who study U.S. politics (see Hopkins & Sides 2015; Thurber & Yoshinaka 2016). There is little question that the Democratic and Republican parties have drifted apart ideologically on average (e.g., Fiorina & Abrams 2008; Hetherington 2009).[2] However, scholars of U.S. politics have recently taken notice of a different type of party polarization: sharp differences in how partisans feel about the in-party vs. out-party, accompanied by a behavioral phenomenon—interparty social segregation (Iyengar, Sood, & Lelkes 2012; Mason 2015). We use the terms "affective" and "social" polarization to describe these phenomena; however, one could also use the labels "prejudice" and "discrimination" (see Gift & Gift 2015; Iyengar & Westwood 2015).

In the U.S., partisans routinely express both explicit and implicit biases against the out-party, feeling more negatively toward opponents, worrying that they represent a threat to the nation, hoping family members won't marry them, and even refusing to hire them. Scholars have begun to document this phenomenon in other nations as well (Garrett, Gvirsman, Johnson, Tsfati, Neo, & Dal 2014). Such biases have increased in the U.S. recently (Iyengar, Sood, & Lelkes 2012; Iyengar & Westwood 2015). Approaching the 2016 elections, Americans were aware of these divisions, and a majority expected them to persist or even worsen post-election (Pew 2016c).

Generally speaking, there are two potential contributors to affective/social polarization— diverging views over the substance of policy (i.e., ideology) and diverging political identities. In

---

[2] These growing differences have been driven both by "party sorting" (Levendusky 2009) and "conflict extension" (Layman, Carsey, & Horowitz 2006).

the U.S., there is little question as to the sharply different issue agendas of the two political parties (Noel 2013). Democrats and Republicans in Congress today are more polarized than they have been in over 100 years (Mansbridge & Martin 2013). While partisan issue polarization among ordinary citizens is less pronounced, lay Democrats and Republicans increasingly differ on the issues on average (Levendusky 2009; Pew 2014). These ideological differences likely contribute to affective/social polarization due to the simple fact that people tend to feel quite negatively toward those with whom they disagree about important political issues (Ryan 2014).

While ideological differences between Democrats and Republicans may be the most intuitive explanation for affective/social polarization in the U.S., the intensity of interparty prejudice there greatly exceeds the relatively modest levels of issue disagreement between ordinary Democrats and Republicans (see Iyengar, Sood, & Lelkes 2012). Recent scholarship has emphasized social identity as a contributor to affective/social polarization (Iyengar, Sood, & Lelkes 2012; Mason 2015). It is well-established that partisans in democratic systems tend to share a sense of group identity (Bankert, Huddy, & Rosema 2017; Campbell, Converse, Miller, & Stokes 1960; Greene 1999)—a feeling of affinity with co-partisans and an incorporation of the group identity into one's personal identity (Huddy, Mason, & Aarøe 2015). Social identity theory argues that the simple existence of distinct social groups will lead to affective/social polarization between group identifiers. Individuals tend to hold their own group in higher regard than others, a tendency rooted in a desire for self-esteem (Haslam, Ellemers, Reicher, et al. 2010; Tajfel & Turner 1979).

But why might citizens' tendency to socially identify with their political party lead to *increasing* interparty dislike and social avoidance over time? One explanation is that citizens in democratic nations today are much more likely than in previous eras to encounter media messages that criticize the political opposition. There are multiple reasons as to why this is the case. In the U.S., Congressional majorities are more insecure than in the past, leading to much more partisan messaging (Lee 2016). In addition, the weakening of campaign finance laws has

increased the amount of money available to fund campaign ads (see Hansen, Rocca, & Ortiz 2015). Less particular to the U.S. are high rates of participation in online discussions via social and other new media; it may be the case that, around the world, more citizens are being exposed to casual political talk. We say more about this particular phenomenon in the next section.

Social identity theory (Tajfel & Turner 1979) and related frameworks such as intergroup emotions theory (Mackie, Maitner, & Smith 2009) strongly suggest that encountering partisan criticism—toward the opposition, or by the opposition toward one's own group—will spur further affective and social polarization. Because party identities tend to be internalized, the group's failures and victories are interpreted as personal ones (Huddy, Mason, & Aarøe 2015). Criticisms of the in-group are perceived as attacks on the self (Mackie, Maitner, & Smith 2009), motivating efforts to regain self-esteem by denigrating the opposition and bolstering one's own partisan group.[3] Likewise, criticisms of the opposition reinforce notions of an inferior out-group that poses a threat to the self. The result is that what we might call an "affective party differential" (evaluating the in-group more highly than the out-group) grows even larger, as does an associated "social party differential" (preferring to interact with the in-group over the out-group).

There is some scholarly evidence of this dynamic at work. Iyengar, Sood, and Lelkes (2012) demonstrate that increasing exposure to political campaign messages, especially their negative and critical aspects, has contributed to affective and social polarization between partisans in the U.S. Mason (2016) finds that simply presenting partisans with fictional user comments that

---

[3] Generally speaking, it would not be unreasonable to expect some measure of persuasion or conformity among participants exposed to criticism from opponents. However, the type of communication we examine—ad hominem criticism—is unlikely to persuade in the usual sense. And, as conformity is contingent on shared identity (*citation omitted*), such an effect is unlikely as well.

threaten electoral loss or defeat on issues they care about is enough to arouse negative emotions. These contributions are important, particularly their coverage of both traditional and new media messages. However, more experimental work is needed to persuasively link encountering partisan criticism to affective and social polarization specifically. Following Mason, we argue that it is particularly important to investigate this phenomenon in the context of online communication, given the prevalence of both.

Partisan Criticism Online

The media landscape has rapidly changed in the last several decades. One of the most important changes is the fact that the media is more *accessible* today than ever before due to the Internet. Citizens' use of online platforms to comment on and discuss political topics began in earnest with the advent of political blogs and chat rooms in the 1990s and has exploded in recent years via social media platforms, such as Facebook and Twitter (Margetts, John, Hale, & Yasseri 2015; Rainie & Wellman 2012). Commentary not only occurs in parallel to traditional media organizations. Citizens also discuss politics and current events in comment sections hosted by news websites. For example, in the U.S., over 90% of newspaper and TV news sites hosted comment sections until recently (Stroud, Scacco, & Curry 2016).

Such interactive technologies allow citizens to engage the journalists creating news content and to discuss political topics and current events with a wide circle of individuals (Ruiz, Domingo, Micó, et al. 2011; Stromer-Galley 2014; Wojcieszak & Mutz 2009). Although in some cases such discussions occur among a homogeneous group of people (as in the case of many blogs), discussions on newer social media and particularly in online newspaper comment sections include substantial ideological diversity (Conover, Ratkiewicz, Francisco, et al. 2011; Lee, Choi, Kim, & Kim 2014; Suhay, Blackwell, Roche, & Bruggeman 2015).

All else equal, participation in political discussion with diverse others represents a positive development in a democratic polity. However, academics, journalists, and other observers have expressed concern over two particular features of many online discussions. First, much online discourse is low quality, with posts displaying group think, emphasizing opinion over fact, and including illogical argument and inaccurate factual claims (Groshek & Bronda 2016; Ruiz, Domingo, Micó, et al. 2011; Singer 2011). Second, online political discussions are overwhelmingly negative, dominated by criticism of political opponents (Berry & Sobieraj 2014; Park 2015; Suhay, Blackwell, Roche, & Bruggeman 2015). With the prior point in mind, much of this criticism is not substantive but, rather, *ad hominem* (Berry & Sobieraj 2014). Some of this online negativity is overtly disrespectful and demeaning (Coe, Kenski, & Rains 2014; Gervais 2014; Mutz 2005). Such uncivil political discourse has been shown to delegitimize the political opposition (Mutz 2007), decrease open-mindedness (Borah 2012), and encourage copycat incivility (Gervais 2014).

The conflictual nature of online discussion does not go unnoticed by people. A recent study found that, among Americans who engage in political discussion online, large numbers are turned off by the negative tone and find it "stressful and frustrating" to talk to political opponents. A majority say that, after conversations with opponents, they feel they have less in common with them (Pew 2016b).

Empirics

We designed two experiments to ascertain the impact of partisan criticism in online comment sections on affective and social partisan polarization. In each experiment, all participants read a news article that discussed an issue debate between Democrats and Republicans followed by reader comments. Participants in the control groups read two reader comments which were negative in tone but nonpartisan. Treated participants were exposed to these comments as well

as *additional comments which contained partisan criticism*. Note that this particular contrast is essential to properly test our hypotheses, which zero in on the effects of *partisan* negativity. For example, if the control group did not receive any comments and we observed experimental effects, we would not know whether those effects were due to the presence of comments, the presence of negatively valenced comments, or—our real interest—*partisan negative comments*. In addition, given our interest in testing whether partisan identity-based criticism—as opposed to issue-based criticism—is a source of affective/social polarization, the "treatment" comments criticized partisan opponents but contained no substantive issue content. After reading the articles and comments, participants' affective evaluations of Democrats and Republicans were measured to assess affective polarization. In Study 1, we also assessed evaluations of President Obama, which we discuss further below. In Study 2, we again measured affective evaluations of members of the two parties and then extended our investigation to participants' desire for social distance from political opponents, i.e., desire for inter-party social segregation.

Study 1: The Sequester Experiment

Sample & Study Design

We conducted the first experiment using Amazon Mechanical Turk (AMT) (N=424) in the spring of 2013.[4] The content of the experiment focused on the federal sequester, i.e. the automatic cuts in federal spending that occurred on March 1, 2013, as a consequence of Congress failing to reach a deficit-reduction deal. While numerous political issues could have provided the necessary content for the study, the sequester had certain advantages: first, it clearly divided Democrats and

---

[4] All participants were required to be residents of the United States and, according to AMT regulations, be at least 18 years old. We further restricted the sample by only sampling from participants who had at least a 90% approval rating on AMT (see Berinsky, Huber, & Lenz 2012).

Republicans over government spending and taxation; second, at the time, the sequester represented a new, current, and salient debate (Pew Research Center 2013), increasing the external validity of the study and, presumably, the engagement of participants.

Study 1 includes two treatment groups and a control group, all randomly assigned. Table 1 gives an overview of the experimental design. The control group was exposed to a news article on the sequester battle between President Obama and Congressional Republicans as well as two negative but politically neutral reader comments. Treatment groups also read the control condition material but were randomly assigned to read additional comments. One treatment group read three additional comments that were critical and uncivil toward Republicans; the other read three additional comments that were critical and uncivil toward Democrats. The comments represent edited versions of actual comments posted on partisan political blogs shortly preceding the study, lending the study good external validity. Note that none of the uncivil comments included substantive political information, such as providing evidence that opponents were governing poorly or supporting problematic policies, thus ensuring a test of affective polarization in response to uncivil criticism, not ideological or policy-based criticism. Appendix A shows the full treatments. (The uncivil comments—added to the baseline stimuli and received by participants in the treatment groups—are highlighted in gray.) In the analyses below, each treatment group is represented by a dummy variable, and the control group is the excluded, comparison group.[5]

---

[5] Note that, after completing the data collection, but *prior to analyzing the results*, we excluded participants who advanced past the article/comments stimulus page before 30 seconds, a minimum determined by multiple timed readings by the researchers. Quality and attention checks are particularly important when drawing participants from online, paid panels such as AMT, where participants will

[Table 1 here]

In the pre-test, we measured age, gender, race, and party as key pre-treatment covariates across the treatment groups.[6] Similar to typical AMT samples, our sample is predominantly male (approximately 62%), White (about 76%), and Democratic (about 46% Democrat (n=194), 35% Independent (n=147), 13% Republican (n=54), and 7% Other (n=29)). The average age in our sample was 33 years old, with a standard deviation of 11.6.

In the post-test, participants were asked to rate Democrats and Republicans using feeling thermometers. Employing feeling thermometers to measure affective polarization is now well-established (see Iyengar, Sood, and Lelkes 2012; Iyengar and Westwood 2015; Mason 2015). The feeling thermometer values were re-coded to range from 0 to 1. The Democratic feeling thermometer score was then subtracted from the Republican score, yielding a final comparative measure ranging from -1 (Democrats receive highest score possible and Republicans lowest score possible) to 0 (neutral) to +1 (Republicans receive highest score possible and Democrats lowest score possible). Because the sequester news article focused much attention on then-President

---

sometimes act unscrupulously (e.g., ignoring instructions and questions) to complete the study quickly and be paid. "Timing" requirements are one way to ensure that the analytical sample is exposed to the stimulus (Berinsky, Huber, & Lenz 2012). Of the original 458 observations, 34 (or about 7%) were dropped. While the relevant coefficients are slightly smaller on average when all observations are retained, the differences are not statistically significant and our substantive conclusions are unchanged.

[6] Chi-Square and ANOVA tests show that these variables are statistically equivalent across the experimental groups ($p>.2$). Therefore, even when we ran the regressions below using control variables, observed empirical patterns were nearly equivalent (and no statistically significant differences observed), allowing us to use the simplified models in our data presentations.

Obama—even including him in the article title—participants were also asked the extent to which they approved or disapproved of the way President Obama was handling the sequester as well as the job Obama was doing in office generally. The two Obama ($r=.80$) approval questions were additively combined into an approval scale, ranging from 0 to 1. Higher values indicate *dis*approval of Obama.[7] Appendix A includes question wording for the dependent variables. Finally, partisan identification of participants is measured on a seven-point scale coded to range from -1 (strong Democrats) to 0 (independents) to 1 (strong Republicans).

Results

Throughout the analyses, we use Ordinary Least Squares regression. We begin by regressing the two dependent variables—the comparative feeling thermometer and Obama Disapproval—on the two treatment variables, party identification, and the interactions between them. Note that, because party identification is scored such that Independent=0, the coefficients on the two treatment variables indicate the estimated treatment effects among Independents. Because we expect Republicans and Democrats to react to the stimuli differently—with Republicans moving upward on the DVs and Democrats downward—we are most interested in the coefficients on the interaction terms. Should the treatments contribute to polarization (i.e., drive Republicans and Democrats further apart on the DVs), *the coefficients on the interaction terms will be positive*.

---

[7] To make sure our effects were unique to *affective* polarization (e.g., not issue attitude polarization as well), participants were also asked three policy questions, which were substituted for the affective dependent variables in separate analyses. The first two were statements indicating that the sequester was harming the economy and ordinary Americans. The third was a statement indicating that the sequester cuts should remain in place. There is no sign that the treatments led to issue attitude polarization. Only affective polarization occurred. Analyses from authors available upon request.

[Table 2 here]

See Table 2.[8] In the first column (Rep/Dem feeling thermometer comparison), we observe that both of the coefficients on the interaction terms are positive. The coefficient on the anti-Republican term is significant at $p<.05$ ($b_{\text{AntiRepXPID}}=.102$); the coefficient on the anti-Democratic term does not reach standard thresholds for statistical significance. In the second column (Obama disapproval), both of the coefficients on the interaction terms are positive and reach standard levels of significance ($b_{\text{AntiDemXPID}}=.15$; $b_{\text{AntiRepXPID}}=.125$, both $p<.01$).[9]

To better understand the effects across the range of partisan identification, we create effects plots and display the associated 90% confidence intervals. See Figures 1 (feeling thermometer) and 2 (Obama disapproval). The left-most panel in each displays effects of the anti-Democrat treatment and the right-most displays effects of the anti-Republican treatment; both sets of effects are relative to the control group. The treatment effects increase as we move from left (Democrats)

---

[8] The models assume a linear relationship between independent and dependent variables. We probed for nonlinear patterns with respect to the seven-point partisan identification scale by employing a series of dummy variables representing each point on the scale (strong Democrats excluded) interacted with the experimental treatments. The subgroups are small, meaning some inconsistency is to be expected; however, the results were roughly linear. See Appendix C.

[9] It is worth noting that the coefficients on the treatment variables are positive and significant ($p<.05$). This suggests that Independents reacted to the treatments by increasing their positivity toward Republicans (and disapproval of Obama). One possible reason is that Study 1 Independents were more ideologically conservative; however, we found that Study 2 Independents had a slightly higher conservatism score (although the differences does not reach standard levels of significance, p=.11). Given this, and that this same pattern does not emerge in Study 2, we do not speculate further as to the cause of these results.

12

to right (Republicans) on the X axis, as expected—those further to the right are more likely to respond to the treatments by evaluating Republicans more positively than Democrats. Interestingly, at the left side of the figures, effects are close to zero; this suggests that Democrats, on average, did not react to the stimuli. Independents and Republicans are driving the results.

Study 2: The Middle Class Experiment

We conduct a second study to extend the reach of Study 1 and address its limitations. First, Study 2 extends beyond Study 1 by adding two measures of social polarization—a measure related to inter-party marriage and another related to partisan social segregation. Second, the comments that made up the experimental treatments in Study 1 were drawn from the "real world," giving the study good external validity but lower internal validity, as the anti-Democratic and anti-Republican comments differed in their specific details. In Study 2, the comments directed at the two sides are perfectly parallel, with only the name of the relevant party changing across treatment conditions. Third, and finally, the Study 1 data are a convenience sample. Although recent studies attest to the general quality of MTurk samples (Berinsky, Huber, & Lenz 2012; Clifford, Jewell, & Waggoner 2015), a more representative sample is desirable, particularly given the relatively small proportion of Republicans in Study 1.[10]

Sample & Study Design

We conducted the second experiment in the spring of 2015, using data from Qualtrics Panels (N=542). The dataset is representative for age, gender, and census region. We also obtained a sample with equal numbers of Republicans, Democrats, and Independents. In this study, again,

---

[10] Clifford, Jewell, and Waggoner's study (2015) finds that liberals and conservatives who use MTurk are similar in character to liberals and conservatives in the general population. This suggests Study 1's estimates of treatment effects among Democrats and Republicans are likely unbiased.

all participants read a neutral news article, this time on the decline of middle class incomes and proposed policy solutions, followed by two negative, but nonpartisan, comments. Treated participants read either: two additional comments which were critical of Democrats or two additional comments critical of Republicans. Comments were more concise than in the previous study and the content less offensive overall, making this a more conservative test of our hypotheses.[11] The *anti-Democratic* version of the treatment is available in Appendix B.[12] Again,

---

[11] There were four distinct treatment groups in the original study—two anti-Democratic (one uncivil and one civil) and two anti-Republican (one uncivil and one civil). The intent was to test whether civil criticism contributed to affective/social polarization as much as uncivil criticism; however, the effects across the uncivil-civil treatments were nearly identical, perhaps owing to the fact that the uncivil treatments only included two additional "uncivil" words (see Appendix B). Formal Wald tests comparing the uncivil and civil coefficients verify their statistical equivalence ($p>.2$). (There is one exception; however, in this case, results work against theoretical expectations, with incivility producing *less* polarization among participants ($p<.01$).) Thus, participants in these paired treatment groups were analyzed together, categorized as having read "anti-Democratic criticism" or "anti-Republican criticism."

[12] To ensure quality responses, we followed standard Qualtrics procedure by employing a simple attention check in the post-test: "We want to be sure you are paying attention. Please select 'disagree'." We also included an open-ended manipulation check following the treatment that asked participants to describe the main points of the article. This check was coded independent of any analyses and checked by a second coder; given the difficulty of open-ended recall items, participants only needed to provide the correct topic to be retained in the study. We dropped 42 respondents due to the attention and manipulation check (or about 7% of the original sample, similar to Study 1). Again, effect sizes weaken only slightly on average and are not statistically different from those we report.

the additional stimuli received by treated participants (relative to the control group) are highlighted in gray. Party labels were switched for those in the anti-Republican treatment group.

By design, the sample included approximately equal numbers of women and men, as well as of Republicans, Democrats, and Independents. The average age in our sample was 46 years old, with a standard deviation of 16.7 years. A majority of the sample self-identified as White (82%).[13]

The evaluative opinions assessed in the post-test are similar to Study 1: feeling thermometers assessing evaluations of Democrats and Republicans. Again, the Democratic feeling thermometer score was subtracted from the Republican score, yielding a final comparative measure ranging from -1 (Democrats receive highest score possible and Republicans lowest score possible) to 0 (neutral) to +1 (Republicans receive highest score possible and Democrats lowest score possible). Further, we added two additional dependent variables to Study 2 that gauged *social* aspects of polarization. First, similar to Iyengar, Sood, and Lelkes (2012), we asked participants how happy they would be if a member of their family married (A) a Republican, and (B) a Democrat. Happiness about a Democratic marriage was subtracted from happiness about a Republican one, with a final measure ranging from -1 (very happy with marriage to a Democrat and very unhappy with marriage to a Republican) to 0 (similarly happy/unhappy with both) to 1 (very happy with marriage to a Republican and very unhappy with marriage to a Democrat). Second, we asked how important it was to the participant to live in a community "where most people held political views similar to your own" (ranging from 0=not important to 1=very important). [14] Question wording for these dependent variables is in Appendix B. Partisan

---

[13] These pre-treatment covariates are statistically equivalent across the five experimental groups ($p>.2$).

[14] This item is based on similar items long used to measure racial prejudice.

identification is again measured on a seven-point scale coded to range from -1 (strong Democrats) to 0 (independents) to 1 (strong Republicans).[15]

Results

As in Study 1, the anti-Democratic and anti-Republican treatment groups are represented by two dummy variables, and the control group is the excluded/comparison group. We regress the first two dependent variables—the comparative feeling thermometer and marriage preference—on the anti-Republican and anti-Democratic treatment variables, party identification, and the interactions between them. See columns 1 and 2 of Table 3.

[Table 3 here]

The coefficients on the Anti-Republican interaction terms are positive and significant in both models ($b_{AntiRepXPID}$=.105, $b_{AntiRepXPID}$=.08; p<.05); however, the Anti-Democratic interaction terms are close to zero (and not significant). These findings mirror Study 1, where the Anti-Republican stimulus also yielded stronger results than the Anti-Democrat stimulus (although the difference is more apparent here). Again, graphing the results gives a clearer picture of the direction and size of effects across the range of partisan identification. See Figures 3 and 4. In these figures, the null effects of the Anti-Democratic treatment on the dependent variables are clear. With respect to the Anti-Republican treatment: While effects are significantly different from zero only among Democrats, what is more important is that the effects among those on the

---

[15] Again, to ensure that we were capturing affective/social polarization uniquely, we examined results for two policy dependent variables linked to the news story in the stimuli: a scale assessing whether taxes on wealthy people and corporations should be increased or decreased and a second scale assessing whether spending ought to be increased or decreased on a range of government programs associated with the Democratic Party. Results of these analyses (available upon request) were null.

left (Strong Democrats/Democrats) are significantly different from those on the right (Strong Republicans/Republicans) ($p<.05$). It is also worth noting that the effects are more balanced between Democratic and Republican participants here than in Study 1.[16]

[Figures 3 & 4 here]

Returning to Table 3, the third column contains the estimates from the model regressing the "likeminded community" dependent variable on the treatment indicator variables and partisan identification. Because this DV does not have a left-right ideological direction—again, it asks all participants whether they would prefer to live among politically likeminded others—*no interaction terms are needed to assess polarization*. Both coefficients on the treatment variables are positive and significant ($b_{AntiDem}=.12$, $p<.01$; $b_{AntiRep}=.08$, $p<.05$). These coefficients are straight-forward to interpret: across the sample, those exposed to online negativity aimed at Democrats or Republicans were more likely to say they wanted to live near people who shared their political views.[17]

---

[16] Again, we probed for nonlinear patterns with respect to the partisan identification scale by employing a series of dummy variables (strong Democrats excluded) interacted with the experimental treatments. Even where we observed statistically significant linear effects (i.e., in response to the anti-Republican treatment), there is some inconsistency across party categories; however, the only consistent nonlinear pattern appears to be unexpectedly weak effects among strong Republicans. See Appendix C.

[17] As the dependent variable has only four answer options, we also carried out ordered probit analysis. See Table A1 in Appendix C. This table also includes a model with interaction terms to test for heterogeneous effects by party identification. (There are none.)

Discussion & Conclusion

To revisit our key hypotheses: we find that partisan criticism encountered online led to affective and social polarization among U.S. partisan identifiers. Participants who read two or three reader comments critical of their own or the opposing party were, on average, more likely to: (1) rate the in-party higher than the out-party on a feeling thermometer, (2) express greater happiness at the prospect of a family member marrying a co-partisan (as opposed to someone in the out-party), and (3) state they prefer living near politically likeminded people. In Study 1, Independents and Republicans were also (4) more likely to disapprove of President Obama after exposure to partisan criticism. While evaluations of political figures are not normally considered a part of the affective polarization phenomenon, this pattern suggests such an extension may be warranted, perhaps with respect to politicians so well-known that they symbolically represent their partisan group (other current U.S. examples would be Hillary Clinton and, of course, President Trump).

These reactions were to online comments that criticized partisans' identities, *not* their issue stands, suggesting partisan social identification as the source of the polarization. Note that all participants (including the control group) were exposed to news articles that mentioned partisan conflict as well as two negative (nonpartisan) reader comments, ruling out partisan priming as well as online negativity in general as alternative explanations for these findings.

This said, there were several unexpected results that deserve discussion. First, in Study 1, on average, there were no treatment effects among Democratic participants. This was not true of Study 2, however. Second, the anti-Republican stimulus consistently "outperformed" the anti-Democrat stimulus. Given this, we decided to explore the data further, breaking down the feeling thermometer dependent measure into its component parts in each study and re-estimating the key models. See Tables A2 and A3 in Appendix C. Interestingly, we see an identical pattern across the two studies: effects are strongest among those who received the anti-Republican stimulus and

with respect to the *Democratic* feeling thermometer. In short, there is a pattern of heightened reaction to Democrats complaining about the opposition party that registered mainly as party polarization with respect to the Democratic feeling thermometer. Perhaps this is a novelty effect; in 2013 and 2015, when our studies were conducted, much partisan criticism was by Republicans toward Democrats, given that they held the Presidency. We can only speculate as to the reasons for these interesting asymmetries but encourage scholars to investigate further. There is no reason to expect that all partisans will polarize in equal measure (see, e.g., Grossmann & Hopkins 2016); various particulars—differences in identity strength between the parties as well as cultural or historical factors—could lead to a lack of parallel results.

This study is limited in two main ways. However, in each case, these limitations offer clear opportunities for future research. First and foremost, we have focused on U.S. politics and have not tested our hypotheses in any other national context. However, we would predict similar results wherever there exist rancorous politics, large numbers of individuals communicating online, and partisans with strong social identities. These characteristics are true of many nations (with respect to partisan identities in particular, see Bankert, Huddy, & Rosema 2017; Lupu 2015). Questions for comparative research on affective/social polarization include whether there may be differences in multi-party, as opposed to two-party systems, or whether this phenomenon is more or less common in economically developed nations. Testing how the framework we have described travels would help to complicate it and understand its boundaries.

A second limitation is our specific focus on comment sections of news websites, particularly at a moment when many news organizations are eliminating comment sections (Gross 2014). Yet, we think it unlikely that our findings are relevant only to comment sections. There are many online spaces where people encounter political criticism in casual conversation (and many others still to be invented). For example, Reddit hosted vigorous political discussions across many types of forums leading up to the 2016 U.S. elections (Barthel 2016). The

interactivity of the online sphere—and the many occasions when users will be exposed to political criticism—represents an opportunity for researchers who wish to replicate and extend our findings. For example, we were able to marginally increase affective and social partisan polarization with our "one-shot studies." It is likely such effects build over time, but further research is needed to test whether this is the case. Scholars might also explore whether the converse of our findings hold: do citizens exposed to expressions of interparty respect and camaraderie online become *less* affectively and socially polarized?

In itself, this natural clash between partisans online represents a problem for democratic polities. Making matters worse is the fact that many individuals and groups—ranging from subcultural groups to unethical politicians to foreign governments—take advantage of the relative anonymity of the online sphere to manipulate online conversation via "trolls" and "bots" (Chen 2016; Howard & Kollanyi 2016; Marwick & Lewis 2017). At the time of this writing, there are various ongoing investigations into Russian influence in the 2016 U.S. election. There is little question, however, that Putin's administration and its allies have used online trolls working for hire and bots to both destabilize American politics and help Donald Trump win the Presidency (Chen 2016; National Intelligence Council 2017). Fanning the flames of existing inter-party animosity has been one method of accomplishing these ends. This phenomenon is not limited to Russia as perpetrator or the U.S. as victim, of course, with governments around the world employing such tactics against citizens in other nations as well as their own (see Benedictus 2016 for an overview), for example, in Venezuela (Melendez & Aguilar 2016).

There is no obvious or easy solution to this situation. In our view, trying to restrict partisan discussion online—for example, by eliminating comment sections, as some have done—is a "solution" that currently raises many more problems than it solves. In democratic nations, decreased anonymity and increased attention by the hosts of online political discussions, including social media apps, to verifying users' identities represents one step forward. Reduced

anonymity not only helps to decrease the number of "fake" accounts online but also has been shown to increase civility in online conversation (Rowe 2014; Santana 2013). In addition, there is not much cost associated with focused efforts by those hosting online conversations to inculcate community norms that emphasize respect, open-mindedness to opposition views, and substantive political discussion (see, e.g., Manosevitch, Steinfeld, & Lev-On 2014). Online political discussion is still a relatively new phenomenon. More efforts to understand its effects on participants across the globe, with an eye toward improving civic culture as well as democratic deliberation and outcomes, would be welcome.

References

Anderson, Ashley, Dominique Brossard, Dietram A. Scheufele, Michael A. Xenos, and Peter Ladwig. 2014. The 'Nasty Effect:' Online Incivility and Risk Perceptions of Emerging Technologies. *Journal of Computer-Mediated Communication 19*(3): 373-387.

Bankert, Alexa, Leonie Huddy, and Martin Rosema. 2017. Measuring Partisanship as a Social Identity in Multi-Party Systems. *Political Behavior 39*(1): 103-132.

Barthel, Michael. 2016. How the 2016 Presidential Campaign Is Being Discussed on Reddit. *Pew Research Center*. May 26.

Benedictus, Leo. 2016. Invasion of the Troll Armies: From Russian Trump Supporters to Turkish State Stooges. *The Guardian*. November 6.

Berinsky, Adam J., Gregory A. Huber, and Gabriel S. Lenz. 2012. Evaluating Online Labor Markets for Experimental Research: Amazon.com's Mechanical Turk. *Political Analysis 20*: 351-368.

Berry, Jeffrey M., and Sarah Sobieraj. 2014. *The Outrage Industry: Political Opinion Media and the New Incivility*. New York: Oxford.

Borah, Porismita. 2012. Does It Matter Where You Read the News Story? Interaction of Incivility and News Frames in the Political Blogosphere. *Communication Research 41*(6): 809-827.

Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes. 1960. *The American Voter*. New York: Wiley.

Chen, Adrian. 2016. The Agency. *New York Times Magazine*. June 2.

Clifford, Scott, Ryan M. Jewell, and Philip D. Waggoner. 2015. Are Samples Drawn from Mechanical Turk Valid for Research on Political Ideology? *Research and Politics* (October-December): 1-9.

Coe, Kevin, Kate Kenski, and Stephen A. Rains. 2014. Online and Uncivil? Patterns and Determinants of Incivility in Newspaper Website Comments. *Journal of Communication 64*(4): 658-679.

Conover, Michael, Jacob Ratkiewicz, Matthew R. Francisco, Bruno Goncalves, Filippo Menczer, and Alessandro Flammini. 2011. Political Polarization on Twitter. *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media 133*: 89-96.

Fiorina, Morris P., and Samuel J. Abrams. 2008. Political Polarization in the American Public. *Annual Review of Political Science 11*: 563-588.

Garrett, R. Kelly, Shira Dvir Gvirsman, Benjamin K. Johnson, Yariv Tsfati, Rachel Neo, & Aysenur Dal. 2014. Implications of Pro- and Counterattitudinal Information Exposure for Affective Polarization. *Human Communication Research 40*: 309-332.

Gervais, Bryan T. 2014. Following the News? Reception of Uncivil Partisan Media and the Use of Incivility in Political Expression. *Political Communication 31*(4): 564-583.

Gift, Karen, and Thomas Gift. 2015. Does Politics Influence Hiring? Evidence from a Randomized Experiment. *Political Behavior 37*(3): 653-675.

Greene, Steven. 1999. Understanding Party Identification: A Social Identity Approach. *Political Psychology 20*(2): 393-403.

Groshek, Jacob, and Serena Bronda. 2016. How Social Media Can Distort and Misinform When Communicating Science. *The Conversation*. June 30.

Gross, Doug. 2014. Online Comments Are Being Phased Out. *CNN.com*. November 21.

Grossmann, Matt, and David A. Hopkins. 2016. *Asymmetric Politics: Ideological Republicans and Group Interest Democrats*. New York: Oxford University Press.

Hansen, Wendy L., Michael S. Rocca, and Brittany Leigh Ortiz. 2015. The Effects of Citizens United on Corporate Spending in the 2012 Presidential Election. *The Journal of Politics 77*(2): 535-45.

Haslam, S. Alexander, Naomi Ellemers, Stephen D. Reicher, Katherine J. Reynolds, and Michael T. Schmitt. 2010. The Social Identity Perspective Today: An Overview of its Defining Ideas. In *Rediscovering Social Identity*, Tom Postmes and Nyla R. Branscombe (eds.). New York: Psychology Press (pp. 341-356).

Hetherington, Marc J. 2009. Putting Polarization in Perspective. *British Journal of Political Science 39*(2): 413-448.

Hopkins, Daniel J., and John Sides, eds. 2015. *Political Polarization in American Politics*. New York: Bloomsbury Academic.

Howard, Philip N., and Bence Kollanyi. 2016. Bots, #Strongerin, and #Brexit: Computational Propaganda During the UK-EU Referendum. *Social Science Research Network*.

Huddy, Leonie, Lilliana Mason, and Lene Aarøe. 2015. Expressive Partisanship: Campaign Involvement, Political Emotion, and Partisan Identity. *American Political Science Review 109*(1): 1-17.

Iyengar, Shanto, Gaurav Sood, and Yphtach Lelkes. 2012. Affect, not Ideology: A Social Identity Perspective on Polarization. *Public Opinion Quarterly 76*(3): 405-431.

Iyengar, Shanto, and Sean J. Westwood. 2015. Fear and Loathing Across Party Lines: New Evidence on Group Polarization. *American Journal of Political Science 59*(3): 690-707.

Layman, Geoffrey C., Thomas M. Carsey, and Juliana Menasce Horowitz. 2006. Party Polarization in American Politics: Characteristics, Causes, and Consequences. *Annual Review of Political Science 9*: 83-110.

Lee, Frances E. 2016. *Insecure Majorities: Congress and the Perpetual Campaign*. Chicago, IL: University of Chicago Press.

Lee, Jae Kook, Jihyang Choi, Cheonsoo Kim, and Yonghwan Kim. 2014. Social Media, Network Heterogeneity, and Opinion Polarization. *Journal of Communication 64*(4): 702-722.

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

Levendusky, Matthew. 2013. *How Partisan Media Polarize America*. Chicago: University of Chicago Press.

Lupu, Noam. 2015. Party Polarization and Mass Partisanship: A Comparative Perspective. *Political Behavior* 37(2): 331-356.

Mackie, Diane M., Angela T. Maitner, and Eliot R. Smith. 2009. Intergroup Emotions Theory. In *Handbook of Prejudice, Stereotyping, and Discrimination*, Todd D. Nelson (ed.). New York: Psychology Press.

Manosevitch, Edith, Nili Steinfeld, and Azi Lev-On. 2014. Promoting Online Deliberation Quality: Cognitive Cues Matter. *Information, Communication & Society 17*(10): 1177-1195.

Mansbridge, Jane, and Cathie Jo Martin, Eds. 2013. Negotiating Agreement in Politics: Report of the Task Force on Negotiating Agreement in Politics. *American Political Science Association*. Washington, D.C.

Margetts, Helen, Peter John, Scott Hale, and Taha Yasseri. 2015. *Political Turbulence: How Social Media Shape Collective Action*. Princeton, NJ: Princeton University Press.

Marwick, Alice, and Rebecca Lewis. 2017. Media Manipulation and Disinformation Online. *Data & Society*. Data & Society Research Institute.

Mason, Lilliana. 2015. 'I Disrespectfully Agree': The Differential Effects of Partisan Sorting on Social and Issue Polarization. *American Journal of Political Science 59*(1): 128-145.

Mason, Lilliana. 2016. A Cross-Cutting Calm: How Social Sorting Drives Affective Polarization. *Public Opinion Quarterly 80*(Special Issue): 351-377.

Melendez, Rafael, and Flor Aguilar. 2016. Manipulation Of Public Opinion in Venezuela Using Political Bots. *Emergency Journalism*. May 25.

Mutz, Diana C. 2005. The New Videomalaise: Effects of Televised Incivility on Political Trust. *The American Political Science Review 99*(1): 1-15.

Mutz, Diana C. 2007. Effects of In-Your-Face Television Discourse on Perceptions of a Legitimate Opposition. *American Political Science Review 101*(4): 621-635.

National Intelligence Council. 2017. Assessing Russian Activities and Intentions in Recent U.S. Elections. Intelligence Community Assessment.

Noel, Hans. 2013. *Political Ideologies and Political Parties in America*. New York: Cambridge University Press.

Park, Chang Sup. 2015. Applying 'Negativity Bias' to Twitter: Negative News on Twitter, Emotions, and Political Learning. *Journal of Information Technology & Politics 12*(4): 342-359.

Pew Research Center. 2013. If No Deal is Struck, Four-in-Ten Say Let the Sequester Happen. February 21.

Pew Research Center. 2014. Political Polarization in the American Public: How Increasing Ideological Uniformity and Partisan Antipathy Affect Politics, Compromise, and Everyday Life.

Pew Research Center. 2016a. Partisanship and Political Animosity in 2016: Highly Negative Views of the Opposing Party and Its Members. June 22.

Pew Research Center. 2016b. The Political Environment on Social Media. October 25.

Pew Research Center. 2016c. As Election Nears, Voters Divided Over Democracy and 'Respect.' October 27.

Rainie, Lee, and Barry Wellman. 2012. *Networked: The New Social Operating System*. Cambridge, MA: The MIT Press.

Rowe, Ian. 2014. Civility 2.0: A Comparative Analysis of Incivility in Online Political Discussion. *Information, Communication & Society 18*(2): 121-138.

Ruiz, Carlos, David Domingo, Josep Lluís Micó, Javier Díaz-Noci, Koldo Meso, and Pere Masip. 2011. Public Sphere 2.0? The Democratic Qualities of Citizen Debates in Online Newspapers. *The International Journal of Press/Politics 16*(4): 463-487.

Ryan, Timothy J. 2014. Reconsidering Moral Issues in Politics. *The Journal of Politics 76*(2): 380-397.

Santana, Arthur D. 2014. Virtuous or Vitriolic: The Effect of Anonymity on Civility in Online Newspaper Reader Comment Boards. *Journalism Practice 8*(1): 18-33.

Singer, Jane B. 2011. Taking Responsibility: Legal and Ethical Issues in Participatory Journalism. In *Participatory Journalism: Guarding Open Gates at Online Newspapers*, Jane P. Singer, Alfred Hermida, David Domingo, Ari Heinonen, Steve Paulussen, Thorsten Quandt, Zvi Reich, and Marina Vujnovic (eds.). Wiley-Blackwell.

Stromer-Galley, Jennifer. 2014. Political Discussion and Deliberation Online. In *The Oxford Handbook of Political Communication*, Kate Kenski and Kathleen Hall Jamieson (eds.). New York: Oxford University Press. DOI: 10.1093/oxfordhb/9780199793471.013.015

Stroud, Natalie Jomini, Joshua M. Scacco, and Alexander L. Curry. 2016. The Presence and Use of Interactive Features on News Websites. *Digital Journalism 4*(3): 339-358.

Suhay, Elizabeth, Allyson Blackwell, Cameron Roche, and Lucien Bruggeman. 2015. Forging Bonds and Burning Bridges: Polarization and Incivility in Blog Discussions About Occupy Wall Street. *American Politics Research 43*(4): 643-679.

Tajfel, Henri, and John C. Turner. 1979. An Integrative Theory of Intergroup Conflict. *The Social Psychology of Intergroup Relations 33*(47): 74.

Thurber, James A. and Antoine Yoshinaka, eds. 2016. *American Gridlock: The Sources, Character, and Impact of Political Polarization*. New York: Cambridge University Press.

Wojcieszak, Magdalena E., and Diana C. Mutz. 2009. Online Groups and Political Discourse: Do Online Discussion Spaces Facilitate Exposure to Political Disagreement. *Journal of Communication 59*: 40-59.

Tables

Table 1: Design of Study 1

Experimental group	Participants read...		
	Baseline		*Treatment*
Control group	Sequester news article	Two negative, nonpartisan comments	
Anti-Democrat Treatment	Sequester news article	Two negative, nonpartisan comments	3 comments with anti-Democrat incivility
Anti-Republican Treatment	Sequester news article	Two negative, nonpartisan comments	3 comments with anti-Republican incivility
N	424		

Table 2: Treatment Effects by Partisan Identification (Study 1)

	Rep/Dem FT Comparison	Obama Disapproval
Anti-Democrat Incivility	0.063**	0.081***
	(0.030)	(0.030)
Anti-Republican Incivility	0.070**	0.089***
	(0.030)	(0.030)
Party ID	0.550***	0.263***
	(0.030)	(0.030)
Anti-Dem Incivility * Party ID	0.050	0.153***
	(0.045)	(0.044)
Anti-Rep Incivility * Party ID	0.102**	0.125***
	(0.045)	(0.044)
Constant	−0.095***	0.508***
	(0.020)	(0.020)
Observations	415	418
R^2	0.720	0.484
Adjusted R^2	0.717	0.478
Residual Std. Error	0.223 (df=409)	0.221 (df=412)
F Statistic	210.275*** (df=5;409)	77.359*** (df=5;412)

Note: Ordinary Least Squares regression. Numbers in parentheses are standard errors. Party ID coded

-1 to 1 with Republicans receiving higher scores. *p<0.1; **p<0.05; ***p<0.01

Table 3: Treatment Effects by Partisan Identification (Study 2)

	Rep/Dem FT Comparison	Marriage Preference	Likeminded Community
Anti-Democrat Criticism	−0.013	−0.021	0.124***
	(0.035)	(0.029)	(0.039)
Anti-Republican Criticism	−0.033	−0.020	0.077**
	(0.035)	(0.029)	(0.039)
Party ID	0.470***	0.166***	0.023
	(0.040)	(0.032)	(0.019)
Anti-Dem Criticism * PID	−0.011	0.014	
	(0.048)	(0.039)	
Anti-Rep Criticism * PID	0.105**	0.084**	
	(0.048)	(0.039)	
Constant	0.004	0.019	0.282***
	(0.029)	(0.023)	(0.032)
Observations	515	529	531
R^2	0.638	0.305	0.021
Adjusted R^2	0.634	0.298	0.016
Residual Std. Error	0.290 (df=509)	0.239 (df=523)	0.328 (df=527)
F Statistic	179.123*** (df=5;509)	45.892*** (df=5;523)	3.783** (df=3;527)

Note: Ordinary Least Squares regression. Numbers in parentheses are standard errors. Party ID coded -1
to 1 with Republicans receiving higher scores. *p<0.1; **p<0.05; ***p<0.01
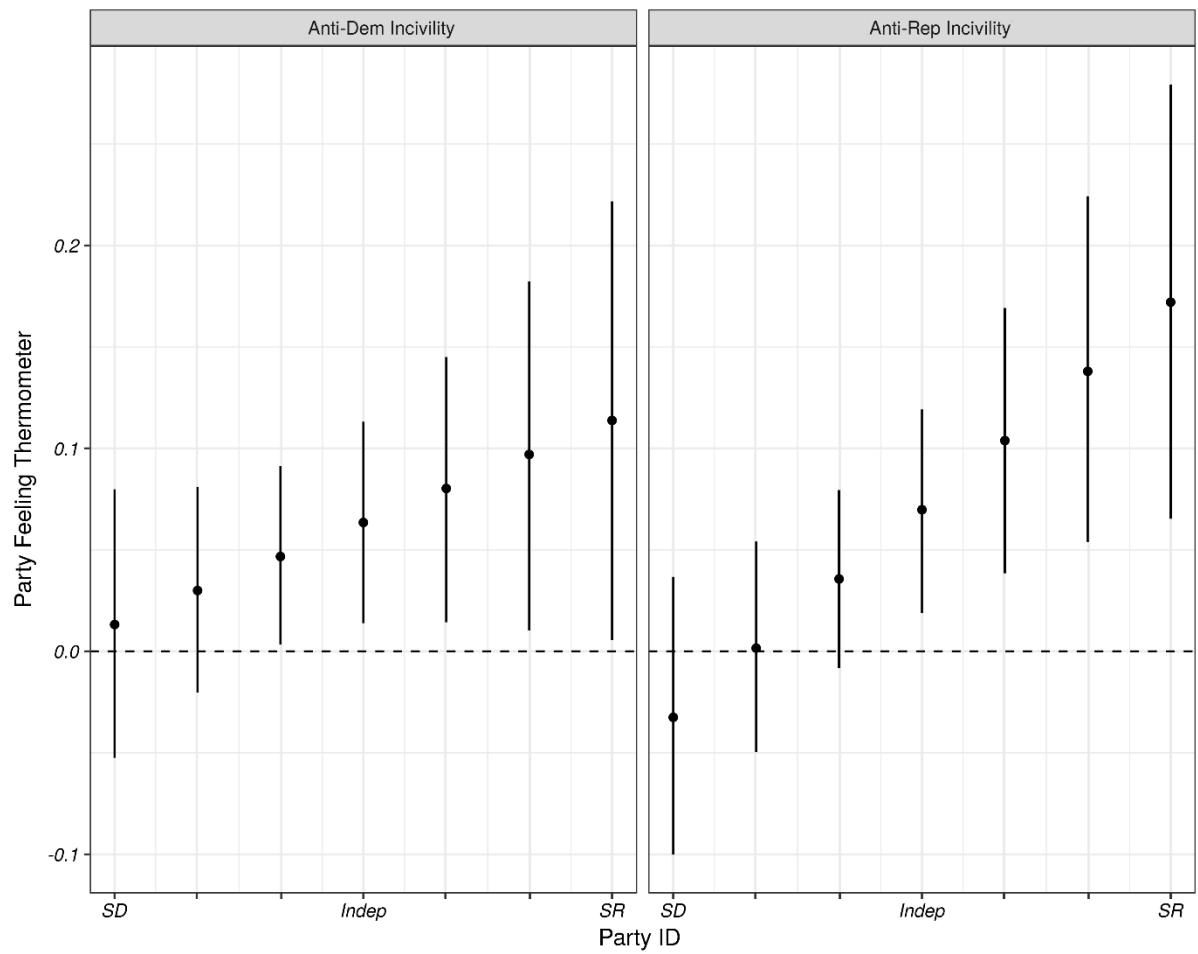
Figures

Figure 1: Treatment effects with respect to Rep/Dem FT Comparison dependent variable, by level of partisan identification. 90% CIs represented. (Study 1)
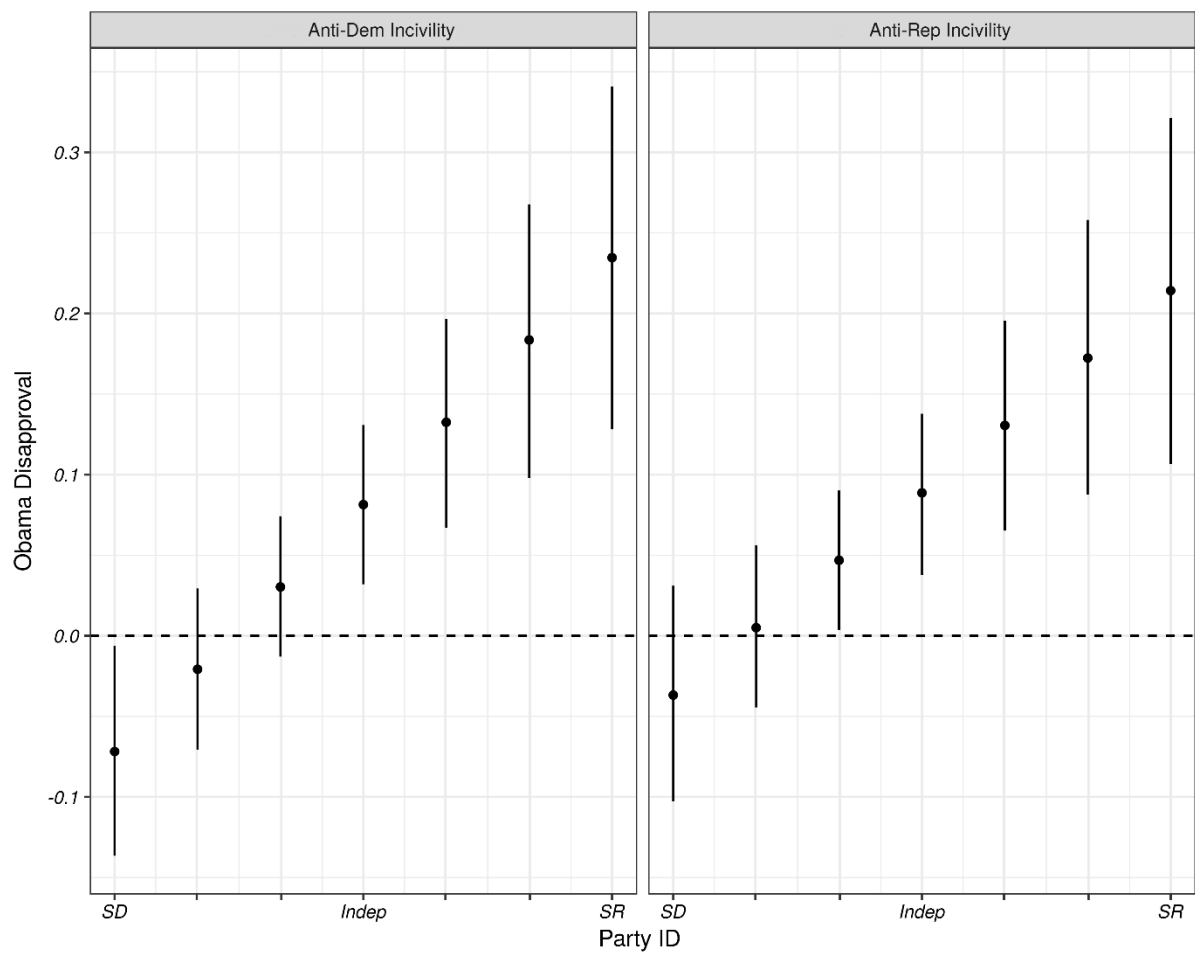
Figure 2: Treatment effects with respect to Obama Disapproval dependent variable, by level of partisan identification. 90% CIs represented. (Study 1)
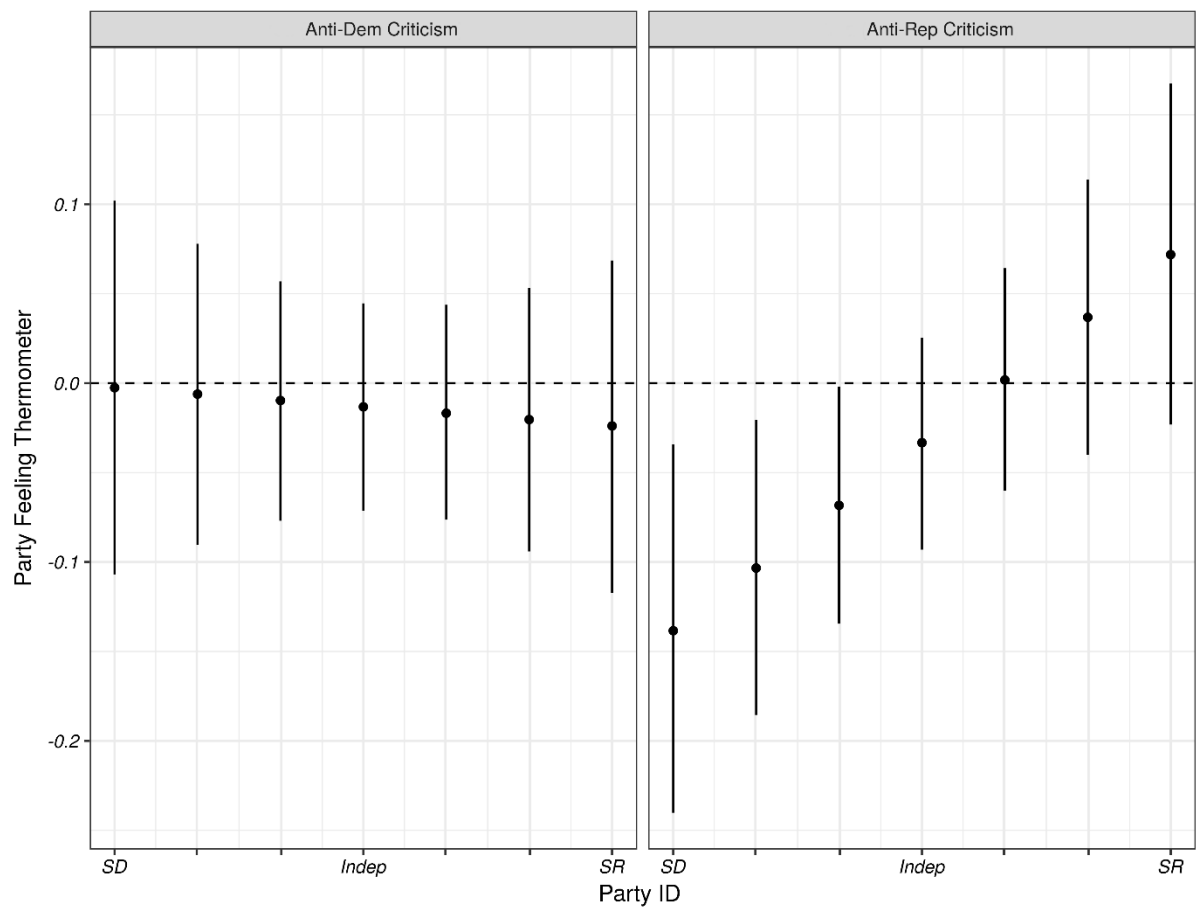
Figure 3: Treatment effects with respect to Rep/Dem FT Comparison dependent variable, by level of partisan identification. 90% CIs represented. (Study 2)
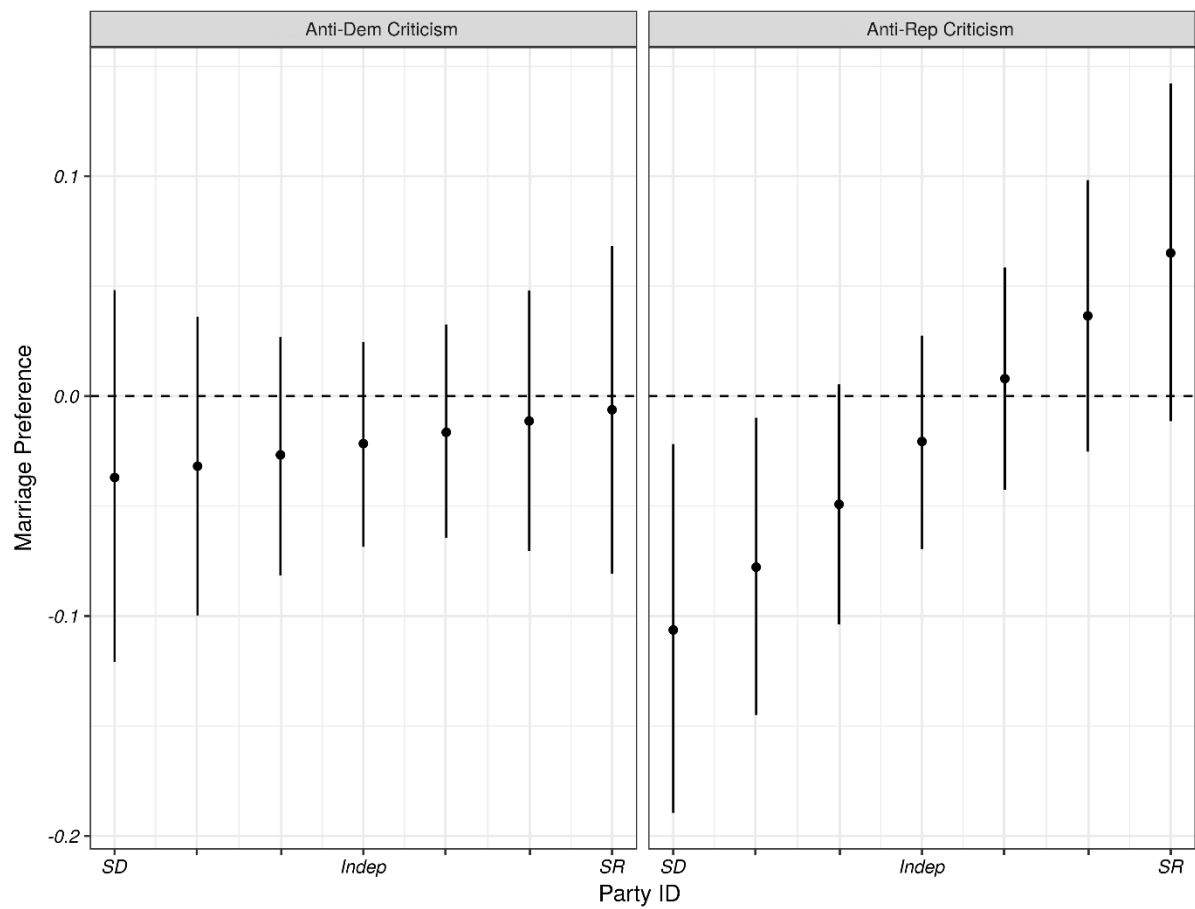
Figure 4: Treatment effects with respect to Marriage Preference dependent variable, by level of partisan identification. 90% CIs represented. (Study 2)