Does Affective Polarization Undermine Democratic Norms or Accountability? Maybe Not*

David E. Broockman† Joshua L. Kalla‡ Sean J. Westwood§

First draft: December 21, 2020
This draft: May 18, 2021

Abstract

Scholars warn that affective polarization undermines democratic norms and accountability. If citizens increasingly detest the other party, are they more likely to endorse norm violations, overlook copartisan politicians' shortcomings, oppose compromise, adopt their party's views, or misperceive economic conditions? A large, influential literature speculates as such. However, such speculation remains difficult to test. We argue the contrary: affective polarization's consequences should be generally confined to interpersonal domains, with more circumscribed political implications. We support this argument with unique experiments which exogenously manipulate citizens' affective polarization and trace downstream consequences, such as their reaction to information about their actual representatives. In our experiments ($N = 9,837$) we produce the equivalent of three decades of change in affective polarization, but find no evidence that these changes influence a range of political behaviors—only interpersonal attitudes. Our results suggest caution about the widespread assumption that reducing affective polarization would meaningfully bolster democratic norms or accountability.

*This research was approved by the IRBs at the University of California, Berkeley, Stanford University, Yale University, and Dartmouth College. Broockman acknowledges the Civic Health Project for support.

†Associate Professor, Travers Department of Political Science, University of California, Berkeley. dbroockman@berkeley.edu.

‡Assistant Professor, Department of Political Science and Department of Statistics and Data Science, Yale University. josh.kalla@yale.edu.

§Assistant Professor, Department of Government, Darthmouth College. sean.j.westwood@dartmouth.edu.

"While we have focused here on the nonpolitical consequences of affective polarization and partisan animus, there is also the question of political consequences. Interestingly, little has been written on this..." – Iyengar et al. (2019, p. 139)

Affective polarization—citizens' growing negative sentiment towards supporters of opposing political parties and positive sentiment toward their own party(Iyengar, Sood and Lelkes 2012)—has been growing worldwide (Boxell, Gentzkow and Shapiro 2020; Gidron, Adams and Horne 2020). Research on this trend now constitutes one of the most influential literatures in contemporary social science scholarship, and has sown widespread alarm among scholars across disciplines (e.g., Finkel et al. 2020).

A bevy of research has documented the negative *interpersonal*, "apolitical" (Druckman et al. 2020*c*) consequences of affective polarization: as more citizens have come to detest outpartisans, they have discriminated against outpartisan job applicants (Gift and Gift 2015), prospective romantic partners (Huber and Malhotra 2017), workers (McConnell et al. 2018), and more (for review, see Iyengar et al. 2019).

But the gravest worries the literature raises are about democracy: that affective polarization has *political* consequences such as, changing the actions citizens incentivize politicians to take, or shifting the norms to which they expect their representatives to adhere. In Table 1, we give examples of over a dozen studies that express concern about the political consequences of affective polarization. Scholars speculate that affective polarization has corrosive effects across a number of political domains, including willingness to exercise electoral accountability, support for democratic norms, adopting one's party's policy positions, reductions in support for legislative bipartisanship, and inaccuracies in the perceptions of objective conditions. Seeing negative trends in many of these areas, scholars widely suspect that affective polarization is to blame.

Surprisingly, however, research investigating these potential political consequences of affective polarization is scarce. As Druckman et al. (2020*b*, p. 9) review, "we know little about its political effects" (see also Yair 2020). In other words, speculation is rife, but evidence is rare.

In this paper, we remedy this dearth of evidence in the context of a novel argument about

Table 1: Example Speculation On Political Implications of Affective Polarization in Prior Work

Outcome	Quote
Electoral Accountability	• "...growth in affective polarization may weaken the role of elections in moderating [elite] polarization. ...the willingness to punish one's own party's politicians for taking an extreme position will weaken..." (Pierson and Schickler 2020, p. 50) • "we suspect that affective polarization increases support for extremist politicians, or, at least, blinds partisans to the ideological extremity of candidates from their party" (Iyengar et al. 2019, p. 142) • "[P]olarization undermines the public's ability to serve as a democratic check." (Graham and Svolik 2020, p. 407) • "...heightened polarization has made it almost impossible for partisans to abandon their party's candidates, no matter their limitations" (Iyengar and Krupenkin 2018, p. 215)
Adopting Party's Policy Attitudes	• "For the mass public, we suspect that affective polarization increases partisans' willingness to conform to their party's policy positions" (Iyengar et al. 2019, p. 142) • "If affective polarization – and, most importantly, partisan animus – is associated with greater responsiveness to party cues, then elite behaviours could have tremendous capacity to change mass response to the pandemic" (Druckman et al. 2020*a*, p. 9).
Legislative Bipartisanship	• "Negative views of the opposing party among voters, in turn, encourage political elites to adopt a confrontational approach to governing." (Abramowitz and Webster 2016, p. 22) • "holding opposing partisans in contempt ...precludes innovative cross-party solutions and mutually beneficial compromises" (Finkel et al. 2020, p.533) • "...affective polarization makes governance more difficult." (Levendusky 2018, p. 59)
Democratic Norms	• "Understanding affective polarization is important because it erodes democratic norms and institutions." (Gidron, Adams and Horne 2020, p. 3) • "The cumulative effect of severe polarization as we have defined it here is a deterioration in the quality of democracy, leading to backsliding, illiberalism, and in some cases reversion to autocracy." (McCoy and Somer 2019, p. 258) • "As affective polarization increases, partisans may become more likely to ignore democratic norms..." (Kingzette et al. 2021, p. 2) • "affective polarization can have grave ramifications...Partisanship appears to now compromise the norms and standards we apply to our elected representatives, and even leads partisans to call into question the legitimacy of election results" (Iyengar et al. 2019, p. 143)
Condition Perceptions	• "These findings underscore the challenges that affective polarization poses for governance. In a polarized America, citizens may be willing to tolerate poor economic performance from their own party, or fail to reward the other side for apparently good economic stewardship, winnowing further already weak hopes that the public will be responsive to government action." (Freeder 2020, p. 28) • "[T]he increased level of affective polarization poses considerable challenges to the democratic process. Partisan bias in perceptions of economic conditions means that voters will fail to credit opposing-party incumbents when the economy grows under their stewardship and fail to penalize in-party incumbents whose economic performance is suspect." (Iyengar, Sood and Lelkes 2012, p. 428)

when and why affective polarization will have consequences in other domains. In particular, we argue that affective polarization's consequences should be generally confined to the interpersonal domain, and should rarely cross over into political domains. We articulate two reasons to be skeptical that affective polarization would "spill over" into strictly political behaviors such as voting decisions. First, citizens' *political* behavior is informed by many considerations beyond their affect towards outpartisans. Changing their political decisions on the basis of affective polarization would require them to undermine these other considerations. For example, would a citizen give up the chance to elect a politician who shares their values simply because this politician shares outpartisans' values, too? Second, previous research finds that citizens attempt to distinguish between outcomes that are attributable to particular politicians and those that are not (e.g., Brody and Sniderman 1977; Kayser and Peress 2012). This suggests that simply because people dislike outpartisans as a group does not mean they will attribute dislike to a particular outpartisan politician in any particular instance.

Previous research has had difficulty empirically investigating the political consequences of affective polarization because of the potential for omitted variable bias: researchers cannot estimate the causal effects of affective polarization simply by comparing citizens who are more or less affectively polarized, because they may differ from each other in other ways, too, confounding comparisons between them. Likewise, correlated time trends between changes in mass behavior and increases in affective polarization confound any potential effects of affective polarization with other changes, including whichever changes may have given rise to affective polarization itself, such as changes in the media environment (e.g., Hopkins 2018). Even as many hope that reducing affective polarization would have positive consequences for democracy (e.g., Levendusky 2018), it is therefore not straightforward to investigate its causal effects.

To probe the causal effects of affective polarization on political attitudes and behaviors, we conduct a series of unique experiments in the United States which manipulate citizens' levels of affective polarization and then probe potential downstream consequences, such as how citizens

react to information about their representatives. We manipulate affective polarization using a trust game in which respondents to a survey believe they are playing with an outpartisan. The putative outpartisan player either is highly cooperative, resulting in the respondent earning a bonus, or completely uncooperative, resulting in no bonus for the respondent.[1] This manipulation produces very large effects on affective polarization as measured in the same manner as the literature; the exogenous differences in affective polarization we create are similar in size to the three decades of increases in affective polarization in the United States documented by Iyengar, Sood and Lelkes (2012).

Deploying this paradigm across four surveys and $N = 9,837$ respondents, we investigate the causal effects of affective polarization on a variety of downstream outcomes. First, consistent with our argument, we show that increasing affective polarization has large effects on downstream outcomes in the *interpersonal* domain, such as being uncomfortable with having outpartisan friends. We then investigate effects in political domains. Although our intervention dramatically increases affective polarization towards other citizens and party elites, our results run contrary to nearly all predictions in the literature. They are, however, consistent with our argument: in political domains, our estimates of the causal effects of affective polarization are consistently null. We investigate potential downstream effects in five political domains most commonly named in the literature (as reviewed in Table 1): support for democratic norms, willingness to engage in electoral accountability, adopting one party's policy attitudes, legislative bipartisanship, and perceptions of objective conditions. Contrary to the literature's expectations, we observe similar null findings in all these domains.

Our approaches for measuring downstream political outcomes rely on a variety of measurement strategies. For example, to investigate implications for electoral accountability, we avoided hypotheticals, matching respondents to their Congressional districts and providing them

[1] A related paper about linkages between partisan affect and racial affect created this approach (Westwood and Peterson 2020).

with information on how their actual Members of Congress voted on recent legislation (McDonald 2020). We then examine whether manipulating citizens' degree of affective polarization changes how they use this information to hold their representatives accountable.

We also conduct a simple exercise that suggests that most of the correlational relationship between affective polarization and outcomes in the *interpersonal* domain appears to be causal, but that these correlational relationships in *political* domains appear driven by omitted variable bias. Indeed, essentially every single causal relationship we look for in the paper is null, even though almost every single correlational relationship goes in the direction the literature predicts. What researchers think are political consequences of affective polarization may be the consequence of omitted variable bias.

In concluding, we discuss potential alternative explanations, remaining limitations, and potential extensions of our work. As we discuss, we see two central implications.

First, our results suggest lessons for advocates of democratic reforms. Those seeking to bolster support for democratic norms or enhance elite accountability have suggested that reducing affective polarization could help advance these goals (e.g., Klein 2020; Levendusky 2018). However, to achieve these goals, it is valuable to understand both what does and what does not advance them. Our results suggest that efforts to reduce affective polarization, while potentially valuable for many reasons, may not be the most effective strategy to bolster these aspects of democracy. Indeed, a new working paper that conceptually replicates our findings (Voelkel et al. 2021) confirms just this: in line with our findings, interventions designed to decrease affective polarization do not appear to have the hoped-for salutary consequences for democratic norms.

Second, our work suggests that future research on the political implications of affective polarization may wish to concentrate on those areas where the interpersonal domain and the political domain are inherently linked, such as willingness to engage in deliberation or cooperative forms of collective action.

Theoretical Framework

Following the original work on affective polarization (Mason 2014; Iyengar, Sood and Lelkes 2012; Hetherington 2001), we consider the effects of the growing polarization in citizens' affect towards the out-party versus their own in-party.[2] Iyengar, Sood and Lelkes (2012) and Mason (2018) show that conceptualizing growing affective polarization in the context of social identity theory (Tajfel and Turner 1979) fits a number of patterns in the data. In particular, that when we identify with a political party, we negatively stereotype outparty members, seek to avoid interactions with them, and favor ingroup members when distributing resources. For example, Iyengar, Sood and Lelkes (2012, p. 407) find that affective polarization "permeate[s] judgments about interpersonal relations" such as views towards children marrying outpartisans and ratings of traits such as intelligence and selfishness. As described above, a number of highly influential studies have built on their findings to document the extent to which partisanship now "permeate[s] judgments about interpersonal relations" in a number of apolitical settings (e.g., Chen and Rohla 2018; Druckman and Shafranek 2020; Gift and Gift 2015; Hersh and Goldenberg 2016; Huber and Malhotra 2017; Klar, Krupnikov and Ryan 2018; Lelkes and Westwood 2017; Michelitch 2015; Panagopoulos 2016).

In light of the startlingly broad consequences of affective polarization for apolitical, interpersonal interactions, many scholars speculate (or simply assert) that affective polarization must permeate *political* judgments as well (see Table 1 for examples). However, on a theoretical level, this is not entirely obvious.

Most of all, it is not clear that attitudes towards outpartisans are an important consideration in many political decisions citizens make; and, to act on outpartisan affect, citizens necessarily must

[2]Affective polarization is often conceived as being specific to affect towards "rank-and-file" partisans, i.e., towards other citizens (Iyengar, Sood and Lelkes 2012). However, Druckman and Levendusky (2019) note that attitudes are even more negative towards political elites of the other party and that many survey items measuring affective polarization capture antipathy towards both citizens and elites in the outparty. Cognizant of the points Druckman and Levendusky (2019) raise, our manipulation checks distinguish between affect towards citizens and elites, and we find increases in affective polarization towards both (see Figure 2).

compromise these other considerations. In particular, interpersonal judgments may require weighing the anticipated costs and benefits of interactions with outgroup members, such as the discomfort of interacting with an outgroup member. However, political judgments are usually not closely connected to these sorts of interpersonal interactions. Citizens may indeed wish to avoid the discomfort of dating outpartisans. But it is less clear what discomfort affective polarization would lead citizens to want to avoid, or expressive benefits it would lead them to gain, by engaging in many of the political behaviors posited to be downstream of affective polarization. Moreover, political judgments should instead principally rely on a different set of considerations: citizens have a number of underlying predispositions, preferences, values, and interests they bring to bear on political judgments, such as whether a government action accords with their principles or whether a policy serves their interests. Crucially, in order for affective polarization to cause citizens to make different political decisions, it must lead citizens to compromise these other important considerations: if citizens make one decision on the basis of considerations beyond affective polarization and a different decision due to affective polarization, they must by definition be compromising whatever considerations would have led to the former decision had it not been for affective polarization. For example, would a citizen give up the chance to elect a politician who shares their values simply because this politician shares outpartisans' values, too?

The literature on citizens' attributions provides a second reason to be skeptical of the political impacts of affective polarization. When citizens are asked to judge a politician, this research finds that citizens attempt to separate outcomes and signals that are informative about a politician from those that are not (Brody and Sniderman 1977; Ashworth 2012; Kayser and Peress 2012). For instance, citizens do not immediately assume that negative impacts to their personal financial situation can be blamed on government (Brody and Sniderman 1977). Likewise, we expect citizens may not blame individual politicians or candidates for negative feelings they have about an outparty in general or citizens of that outparty in particular.[3] More generally, this mirrors social scientific

[3]This is distinct from arguments about a party brand: citizens may have *information* about policies that a particular

research that finds beliefs and attitudes in one domain rarely spill over into even closely adjacent domains (Coppock and Green 2021; Hopkins and Mummolo 2017).

On an empirical level, the evidence supporting the view that affective polarization has significant downstream consequences is also murky.

First, despite many scholars asserting that the causal effects of affective polarization on political outcomes are well-known, there is in fact surprisingly little empirical research exploring the role of affective polarization in explicitly political contexts.[4] Indeed, many reviews of the literature explicitly note this evidentiary lacuna (e.g., Iyengar et al. 2019; Yair 2020; Druckman et al. 2020*b*;*c*).

Second, observational relationships between affective polarization and other outcomes are difficult to interpret. For instance, the fact that affective polarization has grown in tandem with other trends in American politics, such as the decline in ticket-splitting (Hopkins 2018), may seem to imply that affective polarization plays a role in causing these trends. Likewise, other research has documented correlational differences between affective polarization and reaction to certain political stimuli (e.g., Lelkes and Westwood 2017; Hetherington and Rudolph 2015). However, there are other plausible, observationally equivalent explanations for these patterns in which affective polarization has no causal effect (Fowler 2020). In particular, the very same forces that have led to the increase in affective polarization could lead to these other behaviors. For example, among other factors, changes in the media landscape have been implicated in contributing to the rise of affective polarization (Lelkes, Sood and Iyengar 2017)[5] and the decline

party supports which they then infer politicians of that party support. This is different from the kind of inference citizens must make if they change their votes based on affective polarization; affect has no informational content for citizens to extrapolate in the way that information about a "party brand" does.

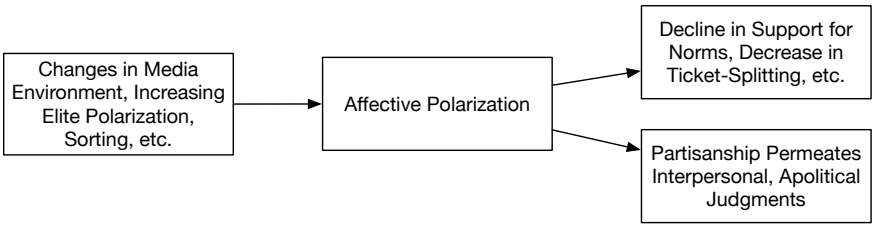
[4]Most research in Table 1 does not present any empirical evidence supporting its speculation, but there is some research which does: (see Abramowitz and Webster 2016; Hetherington 2001; Graham and Svolik 2020; Druckman et al. 2020*a*;*b*). All of this research is observational in nature, although some has examined panel data. See also Orr and Huber (2020) for alternative perspectives.

[5]We take no position in this paper on the causes of affective polarization, we only note that many of the posited causes would also likely influence other downstream political outcomes, creating the opportunity for omitted variable bias. There are a number of factors beyond the media, such as the increasingly hostile rhetoric of political elites (e.g., Hetherington and Rudolph 2015) or increased social sorting (Mason 2018), which scholars have argued contribute to

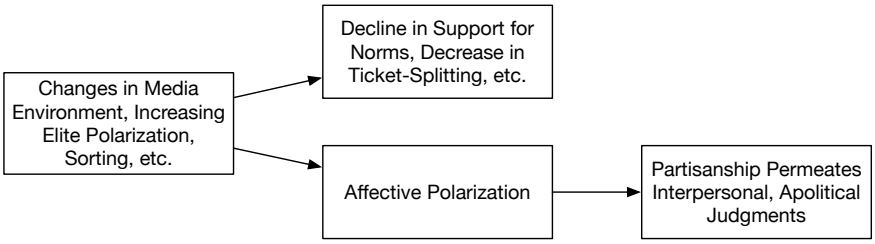
of ticket-splitting (Hopkins 2018), meaning the two may correlate across time or individuals even if affective polarization has no effect on ticket-splitting. Figure 1 visualizes this problem: even though the literature has speculated that changes such as the decline in ticket-splitting may be traced to affective polarization (Figure 1a), it is plausible that other factors cause both, and this would be observationally equivalent (Figure 1b).

Figure 1: Observational Equivalence Between Literature's Speculation and Alternative Hypotheses

(a) Literature's Speculation: Affective Polarization Has Political Consequences



(b) Observationally Equivalent Alternative: Other Factors Cause Both Affective Polarization, Other Political Changes



Understanding which of these two alternatives holds has important implications for efforts to reverse negative trends in democracies such as the decline in support for democratic norms: is affective polarization a cause of these trends, such that reducing it would stem them; or is it merely a symptom of other underlying causes? To break this observational equivalence and isolate the downstream effects of affective polarization, researchers must locate or produce an exogenous source of variation in affective polarization. We therefore deploy a paradigm to produce large changes in affective polarization and then trace its impact on a number of downstream outcomes of interest.

the rise of affective polarization (see also Boxell, Gentzkow and Shapiro 2017; 2020).

Data and Research Design

We trace the downstream impacts of affective polarization on a broad variety of political outcomes across four surveys with a total sample size of $N = 9,837$. Our surveys deploy a two-stage process in which we first randomly assign a task which changes respondent's level of affective polarization (either increasing or decreasing affective polarization) and then present a number of survey items posited to be downstream of affective polarization. This design facilitates a causal analysis of the effect of affective polarization on a number of constructs.

Surveys

Our data is drawn from four surveys conducted in 2019 and 2020 using the online survey vendor Dynata (formerly known as Survey Sampling International). In order to be eligible to participate, survey respondents needed to provide their informed consent, identify as a Democrat or a Republican (including as a leaner, as is typical in this literature), pass a pre-treatment attention check, and demonstrate understanding of the trust game (described below). All these criteria were assessed prior to random assignment. Our surveys resulted in a total of $N = 9,837$ completed responses. The four surveys were conducted in October 2019 (Survey 1 $N = 1,684$), December 2019 (Survey 2 $N = 2,499$), December 2019 (Survey 3 $N = 3,519$), and December 2020 (Survey 4 $N = 2,135$). The studies were all pre-registered with the Center for Open Science.[6]

Online Appendix Table B2 presents the demographics of the survey samples relative to the 2019 Cooperative Congressional Election Study. We find that our survey samples are broadly similar to the demographics of the 2019 CCES, with the largest difference being that our sample is slightly more educated, potentially because we require our respondents to successfully complete an attention check and understand the trust game. Survey item wordings appear in Appendix B. Although all surveys contained the trust game before the outcome measures, different surveys

[6]Pre-analysis plans are available at `https://osf.io/kde27/?view_only=9b450ae09e234f0ba3f7193432170530`, `https://osf.io/7ve49/?view_only=6d36403f548c43a6817e24c079d91d30`, and `https://osf.io/2ysp7/?view_only=1903af75960a4683a4238f835d5eff53`. They are also printed at the conclusion of the Appendix.

contained different outcome measures. Online Appendix Table B1 lists which items appeared on which surveys and the order in which they appeared.

With respect to ethics, all of our surveys were IRB approved and we obtained voluntary and informed consent from all subjects. All subjects were aware they were taking part in a research study. As described below, our research design required the use of deception, as is common in many political science survey- and laboratory-based experiments (Dickson 2011). Participants were aware that deception might be used and we debriefed all subjects immediately following the conclusion of the survey. The survey vendor offered a predetermined, fair level of compensation to participants based on the length of the survey. As shown in Online Appendix Table B2, the survey samples were also diverse. We did not intentionally oversample nor exclude members of any vulnerable or marginalized populations.

Inducing Variation in Affective Polarization with a Modified Trust Game

After collecting demographics, the surveys prompted participants to engage in a task from Westwood and Peterson (2020) which either increases or decreases affective polarization. There are few paradigms available for manipulating affective polarization as the existing literature has defined and measured it.[7] Most vignette designs are ineffective, with Levendusky and Stecula (2021, p. 18) noting that they attempted "close to a dozen techniques to heighten partisan animus" without success. The design we deploy is, to our knowledge, the only available approach for manipulating affective polarization without introducing other factors that could also affect our outcomes of interest. (In the discussion section, we discuss how this manipulation compares with other interventions or trends that create variation in affective polarization.)

This task uses a scripted set of allocations in a modified trust game (Berg, Dickhaut and McCabe 1995). In typical trust games, there are two players. Player 1 receives a cash allocation and is instructed to give "some, all, or none" of the money to Player 2. The player is also told that

[7] Some experiments have decreased affective polarization (Levendusky 2018; Druckman et al. 2019; Rossiter 2020), but would not be suitable for our purpose because they introduce other factors that might affect our outcomes of interest for reasons unrelated to affective polarization.

the researchers will *triple* any amount Player 2 gives to Player 2 and that Player 2 can return some, all, or none of the money back to Player 1. Therefore, the more Player 1 expects reciprocity from Player 2, the more money they should allocate to Player 2 in anticipation they will receive a larger sum in return, and the better off Player 2 will be. For example, if Player 1 gives all her money to Player 2, this sum would be tripled, and Player 2 could return half of the tripled amount to Player 1—leaving both players with 50% more than Player 1's initial allocation. But if Player 1 gives no money to Player 2, Player 1 leaves with only her initial allocation and Player 2 leaves with nothing.

Our studies deploy a version of this framework we alter in several ways. First, we always make participants take the role of Player 2. This means they always first observe an allocation another player makes to them. Second, across three consecutive rounds of game play, participants are told they are interacting with three other respondents of the opposite political party who have each been allocated \$10.[8] However, they are in fact are interacting with computerized opponents who offer allocations based on a pre-determined script. Participants randomized to the Positive Experience condition receive allocations from Player 1 of \$8, \$7 and \$8 (tripled to \$24, \$21 and \$24) respectively across the three rounds of the game. However, those in the Negative Experience condition receive \$0 allocations in all three rounds. Following each round, in both conditions, participants are told the other player said that they made their allocation due to the player's own partisanship. At the end of the game, participants saw a summary of the results for all rounds. At the very end of the survey, after all the dependent variables were measured, we debriefed respondents and informed them that they were playing against a computer. Consistent with our ex ante instructions to respondents, we also pay respondents a bonus equal to 0.05 times the amount they won.

In the survey, participants read instructions, saw three example rounds to ensure they

[8]The opponent's other profile attributes beyond partisanship were randomly assigned in each round of play. Gender was either 'Male' or 'Female', age was drawn between 25 and 54, and income was drawn from four brackets: '\$30,000 - \$39,999', '\$40,000 - \$49,999', '\$50,000 - \$59,999', and '\$60,000 - \$69,999.'

understood the game, and completed two comprehension questions.[9] They then participated in the trust game, with gameplay unfolding in a manner consistent with their random assignment. All game instructions are given in Appendix B.

In summary, participants assigned to the Positive Experience condition thought they were playing with three outpartisans who were generous to and trusting of them, resulting in the respondent having the opportunity to earn a reasonably sized bonus, and who named the participant's partisanship as the reason for their generosity. Participants in the Negative Experience condition experienced the opposite, with outpartisans denying them any opportunity to win a bonus three times in a row, and naming their partisanship as the reason.

Although we increase and decrease levels of affective polarization with a manipulation based on a trust game, a new working paper (Voelkel et al. 2021) that conceptually replicates our work found that other approaches for reducing affective polarization also do not appear to lead to downstream changes in support for democratic norms, indicating our findings are not an artifact of the manipulation we use. We return to this issue again in the discussion section.

Manipulation Checks

This manipulation produced large changes in affective polarization in all four of our surveys, as assessed by the main measure of affective polarization employed in the literature (the difference in feeling thermometers for citizens of the in- minus out-party). In particular, participants assigned to the Negative Experience condition rated outpartisan citizens and outpartisan politicians and elected officials much more negatively than those in the Positive Experience condition.[10] For example, in Survey 1, the average feeling thermometer difference for in- minus out-party citizens in the Positive Experience group was 23.8, but in the Negative Experience group it grew to 36.3. Across our four surveys, the average difference in affective

[9] If respondents missed a comprehension question, participants were given the answer and asked the questions again. Those failing the questions three times were removed from the survey prior to random assignment.

[10] To be clear, these feeling thermometer items ask about "People who are Democrats/Republicans" and "Democratic/Republican Politicians and Elected Officials" in general, and do *not* reference the individual players in the trust game.

polarization towards other citizens across the Negative and Positive Experience conditions is 14.3 degrees. The effect on affective polarization towards political elites (measured only in Survey 4) is 9.8 degrees. These differences are approximately equivalent to over three decades of increasing affective polarization in the United States as estimated by Iyengar, Sood and Lelkes (2012).[11] This means that the effect of the Negative Experience condition on our downstream variables can be interpreted as the effect of approximately three decades of increased affective polarization in the United States.

The top row of Figure 2 visualizes these effects on standardized versions of variables we asked as part of a manipulation check. Mirroring Iyengar, Sood and Lelkes's (2012) findings regarding changes over time in the electorate, we see no meaningful changes in sentiments towards inpartisans (if a small decrease), but large decreases in sentiments towards outpartisans. This produces a large increase in the difference between the outpartisan and inpartisan feeling thermometers. The intervention also meaningfully affected affective polarization towards "politicians and candidates." Importantly, this shows that the manipulation does not just affect measures of affective polarization that are social in nature. Finally, unsurprisingly, we show that respondents were much more likely to say they were treated unfairly in the trust game if in the Negative Experience condition.

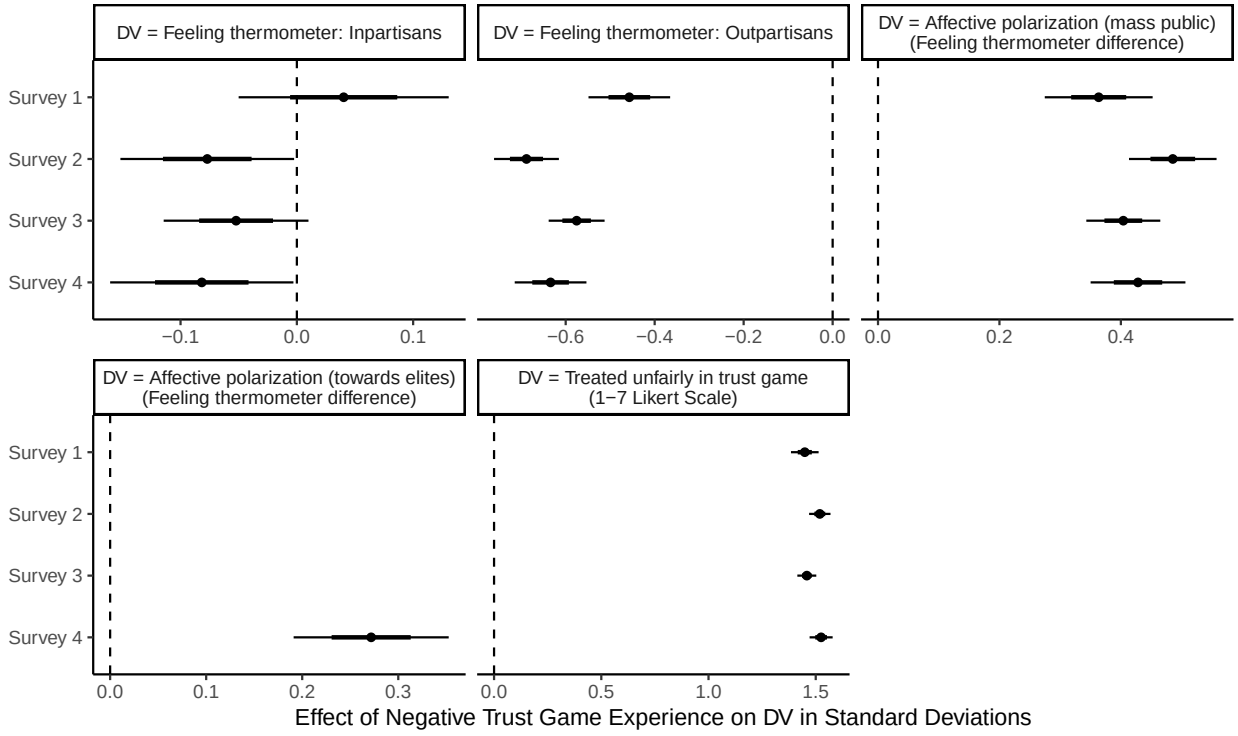
In summary, our experimental manipulation produces changes in affective polarization similar to the changes seen over the last three decades in the United States and about which scholars have expressed such concern.

Results: Downstream Consequences of Affective Polarization

We next exploit these experimentally-induced changes in affective polarization to analyze its downstream consequences across a variety of outcomes.

[11] The total increase in affective polarization Iyengar, Sood and Lelkes (2012, Table A1) observe from 1978 to 2008 is 15.35 degrees for Democrats and 10.16 degrees for Republicans, for an average of 12.76, similar to our observed effects of 14.3 and 9.8 degrees in affective polarization towards voters and elites, respectively.

Figure 2: Manipulation Checks: Effect of Negative Trust Game Experience on Manipulation Check DVs



Notes: Point estimates are surrounded by one standard error (thick tails) and 95% confidence intervals (thin tails). Point estimates are from multivariate regressions controlling for the pre-registered covariates identified in our pre-analysis plan.

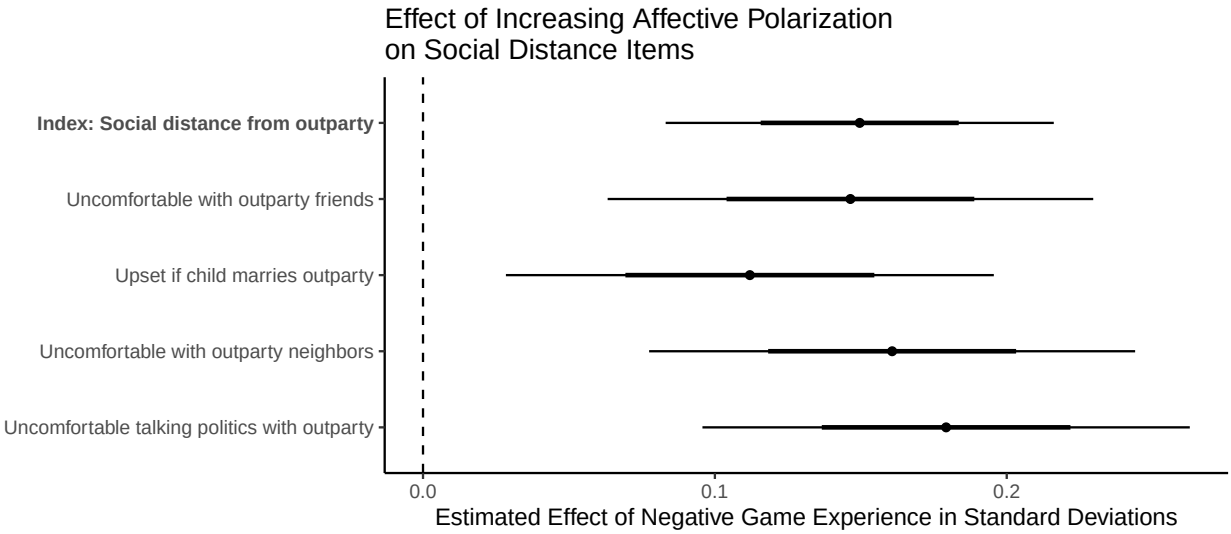
Apolitical, Interpersonal Items: Social Distance Measures

We first examine the causal effects of affective polarization in an apolitical domain where prior research has also documented its effects, social distance. Our argument also predicts we would observe effects of affective polarization since it is squarely within the interpersonal realm, and our research design allows us to replicate this prior research's findings within a causal framework.

Figure 3 reports the effects of the treatment on a set of social distance measures, most from Iyengar, Sood and Lelkes (2012) and Druckman and Levendusky (2019). These items probe discomfort with social interactions or social ties with outparty members (the items do *not* refer to the particular partisans in the trust game). These items appeared on Survey 4. We reverse code items as necessary such that the hypothesis that affective polarization has deleterious downstream

consequences corresponds always with positive estimates; in this case, positive estimates correspond with greater discomfort with social interactions with outparty members.

Figure 3: Effect of Increasing Affective Polarization on Social Distance Items



Notes: Point estimates are surrounded by one standard error (thick tails) and 95% confidence intervals (thin tails). Point estimates are from multivariate regressions controlling for the pre-registered covariates identified in our pre-analysis plan.

As Figure 3 shows, we observe large and statistically significant increases on all the social distance items. The first coefficient shows an index we pre-registered of all four of these items; the effect on this index is also significant. The treatment strongly affected all of the items we drew from previous literature, including expressing discomfort with having outpartisan friends and displeasure with one's child marrying an outpartisan.[12] We also asked an item not in previous literature: whether the respondent would be uncomfortable talking about politics with an outpartisan. Respondents grew more uncomfortable with this form of interaction as well, raising a theme we return to in the discussion: that there may be political implications of affective polarization in those areas where the interpersonal domain and the political domain are inherently linked, such as in willingness to engage in deliberation. However, it is also notable that the effects

[12]However, see Klar, Krupnikov and Ryan (2018) for an important critique of the marriage item.

on this item were no larger than the effects on the other, apolitical social distance items.

Electoral Accountability for Elected Representatives

Having demonstrated that our stimulus manipulates affective polarization and that this in turn creates downstream consequences in the interpersonal domain, we next turn to evaluating affective polarization's downstream consequences in political domains. Perhaps the concern scholars most commonly articulate about the political consequences of affective polarization is that it encourages voters to vote for their party's candidates, no matter their shortcomings. Many scholars express some version of this concern (see Table 1), such as concern that affective polarization "has made it almost impossible for partisans to abandon their party's candidates, no matter their limitations" (Iyengar and Krupenkin 2018, p. 215, see also Kalla and Broockman 2018).

Following from Little, Schnakenberg and Turner's (2020) framework, we consider two possible channels through which this concern might manifest: *divergence*, in which affective polarization simply increases party loyalty; and *desensitization*, in which affective polarization weakens voter's responses to information about incumbent's actions.

We sought to assess these consequences of affective polarization with an experiment with an unusual degree of realism. In particular, in Surveys 1 and 3, we first asked respondents for their full zip code so that we could determine their actual Congressional district and, in turn, their actual representative in Congress (McDonald 2020). We next asked them whether they personally supported or opposed a series of bills that Congress had previously voted on (see full text in Appendix B). Then, respondents played the trust game and were randomized to have a positive or negative experience.

Finally, we showed respondents how their actual Member of Congress actually voted on several bills in Congress, randomizing whether we showed respondents congruent votes (up to 3 votes where they *agreed* with how their Member of Congress voted) or incongruent votes (up to 3 votes where they *disagreed* with how their Member of Congress voted). Whether votes were congruent or incongruent was measured by comparing respondent's pre-treatment stated

preferences on each issue to our records of how their Member of Congress actually voted. That is, respondents were randomized to learn about votes their Member of Congress had cast that were in line with respondents' pre-treatment issue positions or out of step with respondents' pre-treatment positions. In all cases the votes were real.[13] (In Survey 3, we also included a pure control group where respondents received no information about how their Member of Congress voted.)

We concluded by asking a series of questions capturing vote intention in the next Congressional election and approval of their Member of Congress. As we pre-registered, we combine these into a Member of Congress (MC) Approval Index,[14] our primary outcome. This index is standardized to have mean zero and standard deviation one in each survey.

Respondents were told their MC's party when they learned about their votes and the MC's party affiliation was again given in the survey question asking for their vote intentions and approval (see Online Appendix B.2 for the full items).

This design produced the $2x2$ factorial design shown in Table 2. The first factor, shown at the top, is the affective polarization manipulation: whether respondents had a Positive or Negative Game Experience, which reduced or increased their levels of affective polarization. The second factor, shown at left, is the Member of Congress vote manipulation: whether respondents were shown votes their Member of Congress cast which were either congruent or incongruent with their own views. Respondents are therefore assigned to one of the four cells shown in the center of Table 2.

Divergence: Does Affective Polarization Increase Party Loyalty?

A first simple test we conduct of the political impacts of affective polarization is whether affective polarization increases party loyalty. To do so, for the moment we set aside and hold constant

[13] All the votes we showed respondents were accurate information on how their Member of Congress actually voted, there was no deception with respect to who represented respondents in Congress or how their representative voted. A small proportion of respondents either agreed or disagreed with their Member of Congress on every single vote we asked about and so were not eligible for this experiment.

[14] As we pre-registered, we standardize these three questions and compute a simple additive index of them. In Study 1, $\alpha = 0.93$; in Study 3, $\alpha = 0.83$.

Table 2: 2x2 Design: Member of Congress Accountability Experiment

		Affective Polarization Manipulation		
		Positive Game Experience (Reduced Affective Polarization)	Negative Game Experience (Increased Affective Polarization)	*Difference Between Columns*
Member of Congress Vote Manipulation	Show Congruent Votes (Votes Respondent Agrees With)			*Effect of Increasing Affective Polarization, Given Respondent Shown Congruent Votes*
	Show Incongruent Votes (Votes Respondent Disagrees With)			*Effect of Increasing Affective Polarization, Given Respondent Shown Incongruent Votes*
	Difference Between Rows	*Effect of Showing Incongruent Votes, Under Reduced Affective Polarization*	*Effect of Showing Incongruent Votes, Under Increased Affective Polarization*	

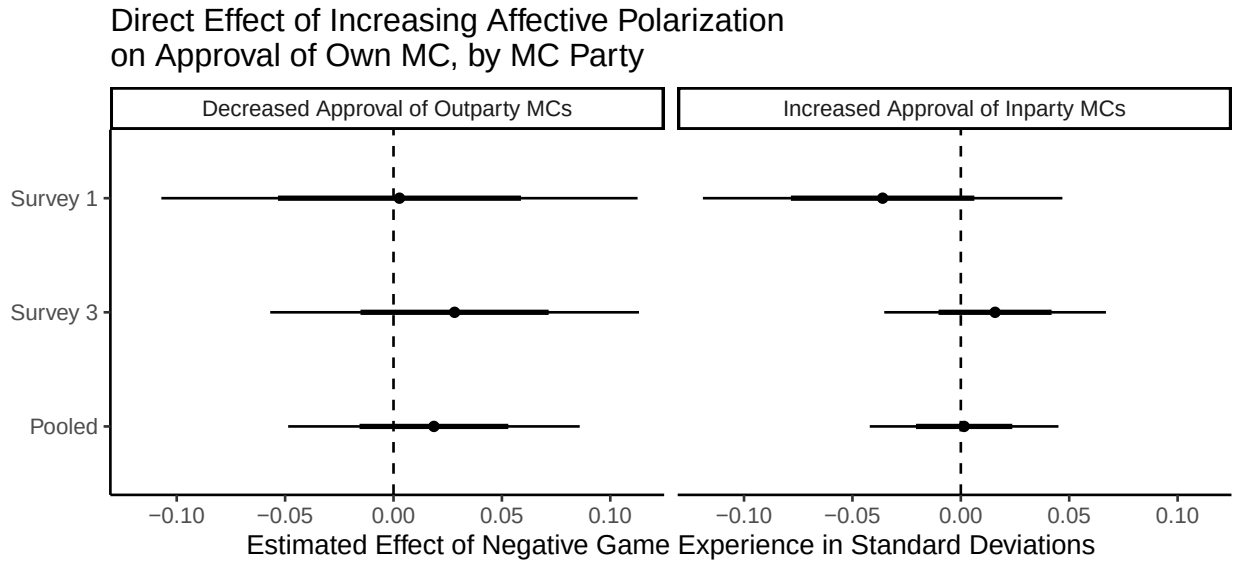
the Member of Congress vote manipulation and focus on the effects of the affective polarization manipulation.

The worry that affective polarization increases party loyalty would predict that, all else equal, individuals whose affective polarization is increased would be more approving of copartisan politicians and less approving of outpartisan politicians. We test whether this is the case by comparing the MC Approval Index in the Positive and Negative Game Experience conditions.

Figure 4 shows the results. As in our other Figures, we code the dependent variables such that positive coefficients correspond with prevailing predictions. We also show the results for both Survey 1 and Survey 3 separately, followed by the pooled results of both coefficients. The left panel in Figure 4 subsets to respondents with outpartisan MCs and shows that increasing participants' level of affective polarization does not make them disapprove of their outpartisan MCs any less. The point estimates in both studies are essentially zero, and the pooled point estimate is very close to zero. The right panel subsets to respondents with copartisan MCs and tests whether they are more approving of their MCs when their level of affective polarization is increased. We again find

that there is no effect: there is no evidence that respondents approve of copartisan MCs more when their level of affective polarizaton is increased, and the pooled estimate is almost exactly zero with a tight confidence interval.

Figure 4: Testing for Divergence: Does Increasing Affective Polarization Reduce Approval of Outpartisan MCs or Increase Approval of Copartisan MCs?



Notes: Point estimates are surrounded by one standard error (thick tails) and 95% confidence intervals (thin tails). Point estimates are from multivariate regressions controlling for the pre-registered covariates identified in our pre-analysis plan.

Desensitization: Does Affective Polarization Weaken Reactions to Information?

Another, more subtle possibility is that affective polarization weakens voter's reactions to politician's actions in office. Following Little, Schnakenberg and Turner (2020), we call this *desensitization*, referring to the possibility that affective polarization weakens the relationship between incumbent actions and voter behavior. To test for this, we first separately estimate the effects of showing respondents incongurent votes their MC cast by affective polarization condition. (These correspond with the first two "Differences Between Rows" shown in Table 2.) The desensitization hypothesis would predict that the negative effects of showing respondents incongurent votes would be smaller in the increased affective polarization (Negative Game

Experience) condition.

Consider first the results in Figure 5a, which shows results from Survey 1.

The left panel of Figure 5a plots the effects of showing incongruent (instead of congruent) votes on the MC Approval Index, separated by affective polarization condition. The first coefficient, above "Effect in Reduced Affpol Condition" shows that, given a respondent has been randomly assigned to the Reduced Affective Polarization condition, showing them incongruent instead of congruent votes their Member of Congress cast causes them to approve of their Member of Congress much less ($d = -0.57$). That is, among respondents in the Reduced Affective Polariaztion condition, those who were randomly assigned to see incongruent votes their Member of Congress cast approved of their MC significantly less than those who saw congruent votes their MC cast.

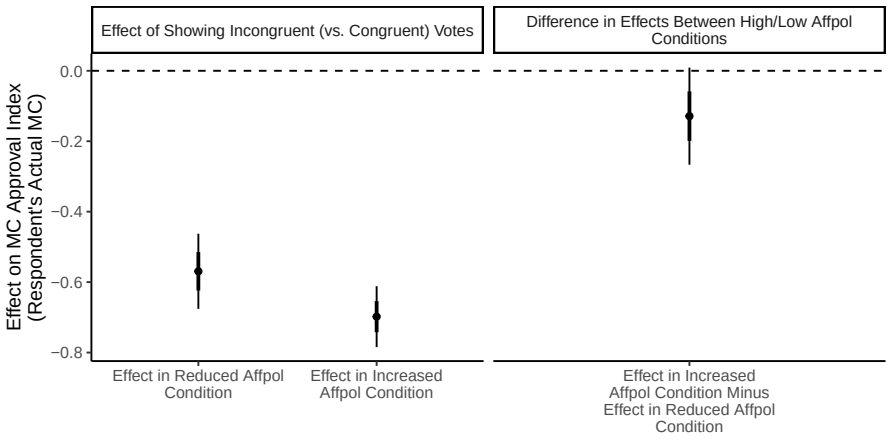
The key test for desensitization is how the effect of showing incongruent instead of congruent votes differs when we increase affective polarization. If the literature's worry is right, we should see respondents be less sensitive to this information when they are more affectively polarized, and therefore have a less negative reaction. However, the second coefficient shown in Figure 5a if anything indicates the opposite, as the effect of showing incongurent votes is actually slightly larger when respondent's level of affective polarization is increased ($d = -0.70$). The difference in these two effects between the high and low affective polarization condition is $d = -0.13$; respondents react if anything slightly more negatively to the incongruent votes when their level of affective polarization is increased.

In Survey 3 we replicated these results with a larger sample and included a pure control that was shown no information about how their Member of Congress voted.

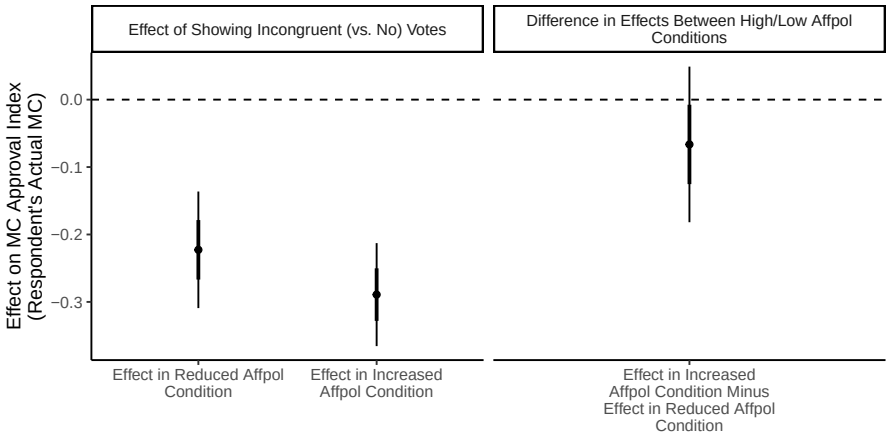
The results are generally similar. Figure 5b shows the effects of showing respondents incongruent votes their MC cast as compared to not showing them any votes their MC cast. As in Survey 1, we see similar effects regardless of whether respondent's level of affective polarization was reduced or increased; and, if anything, those in the increased affective polarization condition

Figure 5: Testing for Desensitization: Does Increasing Affective Polarization Reduce Voter's Sensitivity to Information about their Member of Congress' Votes?

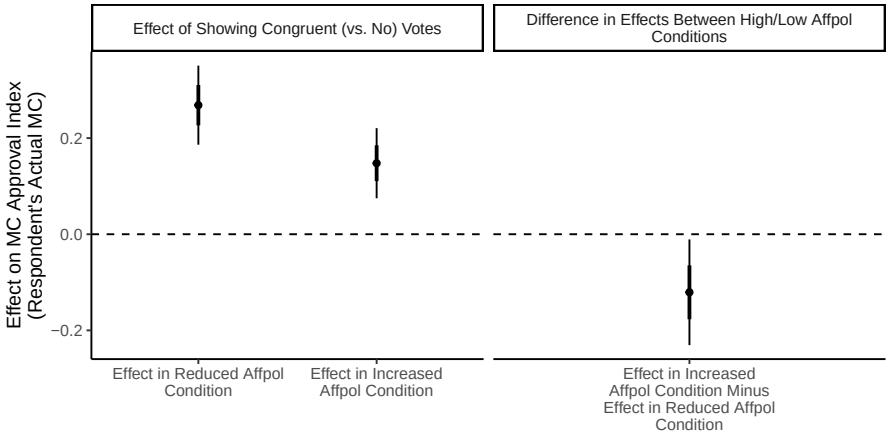
(a) Survey 1: Effect of Showing Incongruent (vs. Congruent) Votes



(b) Survey 3: Effect of Showing Incongruent (vs. Not Showing Votes) Votes



(c) Survey 3: Effect of Showing Congruent (vs. Not Showing) Votes



Notes: See Figure A3 for condition means and Figure A4 for robustness (e.g., effects split by whether MC is copartisan or outpartisan). Point estimates are surrounded by one standard error (thick tails) and 95% confidence intervals (thin tails). Point estimates are from multivariate regressions controlling for the pre-registered covariates identified in our pre-analysis plan.

punished their Member of Congress *more* for casting incongruent votes.

When examining the effects of casting congruent votes, we do see some sign that respondents rewarded their MCs less for casting congruent votes when in the increased affective polarization condition, a borderline statistically significant result ($p = 0.03$). However, there are several reasons to be skeptical of this result: 1) it goes in the opposite direction as the other two results, and pooling the three results yields a null result; 2) as we discuss at the end of the paper, adjusting this result for multiple comparisons renders it insignificant; and 3) while insignificant in both cases, the point estimate is similar in size for participants with copartisan and outpartisan MCs, inconsistent with what the standard view of affective polarization would hold (i.e., that the effect, if real, would be driven by affective polarization leading respondents to resist rewarding outpartisan politicians for congruent votes).

In summary, randomly inducing an increase in affective polarization does not appear to discourage voters from engaging in electoral accountability; if anything, on net, our point estimates on average indicate they engage in more of it. We also show in Figure A4 that these results hold if we look separately by whether the Member of Congress is a copartisan or outpartisan (indicating the results are not due to ceiling or floor effects), as well as if we focus on just a survey item capturing vote intentions and not our MC Approval Index.[15]

Adopting One's Party's Positions

Several scholars have also speculated that "affective polarization" might be "…associated with greater responsiveness to party cues" (Druckman et al. 2020*b*, p. 9). This could undermine accountability if legislators could avoid responsibility for their actions by more easily persuading affectively polarized copartisan citizens of the merits of their actions. We examine this possibility with a series of additional dependent variables we asked after the MC Accountability experiments

[15]Figure A5 also presents estimates that explore a series of potential mechanisms, finding no effects on other attitudes towards respondents' Members of Congress, such as whether they are loyal or compromise too much. We also found no evidence that a negative experience in the trust game affected the amount of time respondents spent looking at their Members' votes.

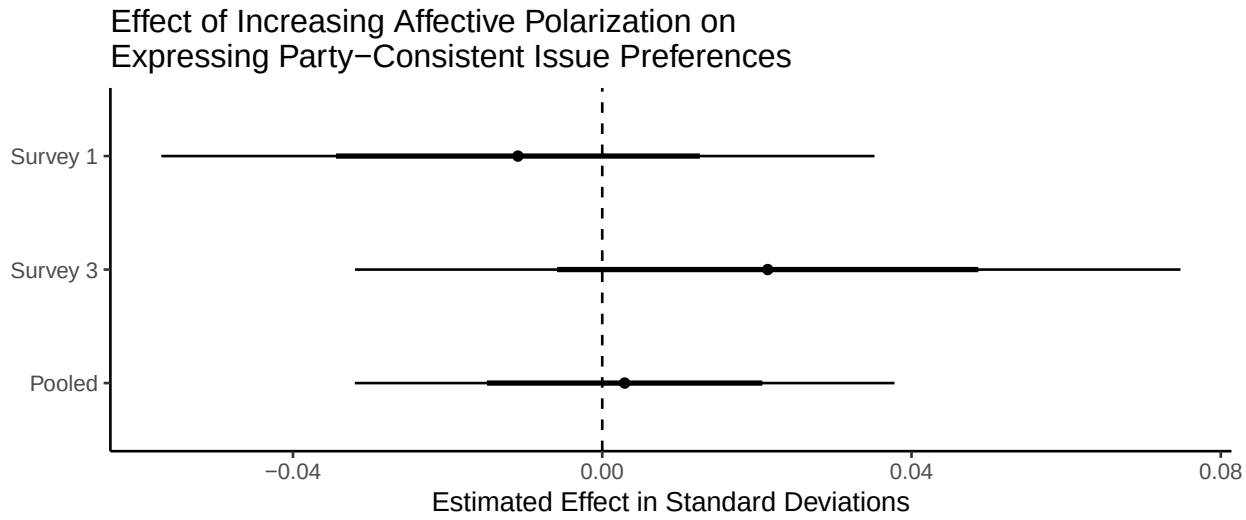
described above. In particular, after playing the trust game, learning their Member of Congress's legislative voting record, and answering questions about favorability towards their Member of Congress, we also asked respondents to again answer questions about their own issue preferences on the issues on which we showed them their Member of Congress' votes.

Are respondents more likely to adopt their party's positions on issues when their level of affective polarization is increased? To test this, we examine whether respondents who were assigned the negative trust game experience (high affective polarization) are more likely to express post-treatment party-consistent issue preferences than those assigned to the positive trust game experience (low affective polarization). This would be expected regardless of whether respondents saw information about how a copartisan or outpartisan MC voted, as partisans could also use an outpartisan MC's vote as a negative cue. To code the dependent variable, we compute the share of issues on which the respondent gave the same position as we had previously said their MC had taken; or, in the case when respondents had outpartisan MCs, the share of issues on which they gave the opposite position.

Figure 6 shows the results. We find no substantively or statistically significant effect of exogenously increasing affective polarization on whether respondents subsequently express party-consistent issue preferences in either Study 1 or Study 3.[16] The point estimates are essentially zero, and the confidence intervals are very small. We continue to find null effects when we separately examine respondents who were of the same party or different party as their Member of Congress.

[16]In this analysis, we omit respondents from the condition in Survey 3 where respondents were shown no votes from their Member of Congress because they received no party cues. However, we actually find that, among individuals shown no votes, if anything respondents in the increased affective polarization condition are slightly less likely to express party-consistent issue preferences. We have no condition where respondents were not asked for their own views pre-treatment, and so we cannot evaluate whether respondents could have felt some pressure to answer consistently with their prior answers, which would bias our estimates downward somewhat. However, the top of our 95% confidence interval is very close to zero, and we are well-powered to detect even very small effects. Nevertheless, future research on this hypothesis could include such a condition.

Figure 6: Testing for Greater Receptivity to Party Cues: Effect of Inducing Affective Polarization on Party-Consistent Issue Preferences



Notes: Point estimates are surrounded by one standard error (thick tails) and 95% confidence intervals (thin tails). Point estimates are from multivariate regressions controlling for the pre-registered covariates identified in our pre-analysis plan.

Support for Legislative Bipartisanship

Increases in affective polarization are also often blamed for partisan gridlock. For example, Levendusky (2018, p. 59) expresses concern that "affective polarization" may "make[] governance more difficult" by discouraging legislators from compromising.

To evaluate this prediction, Survey 2 deployed a manipulation developed by Harbridge and Malhotra (2011, Study 2). The authors found two Members of Congress, one of each party, who had each cast mostly party-line votes in one recent year but had cast a number of votes with the outparty in another year. The respondents are told about the Member of Congress of their party and it is randomly assigned whether they learn about the year of votes when this copartisan member cast party-line votes or often cast votes with the outparty. For example, Republican respondents saw the text below, with the last sentence varied depending on whether respondents were assigned to the bipartisan or partisan condition:

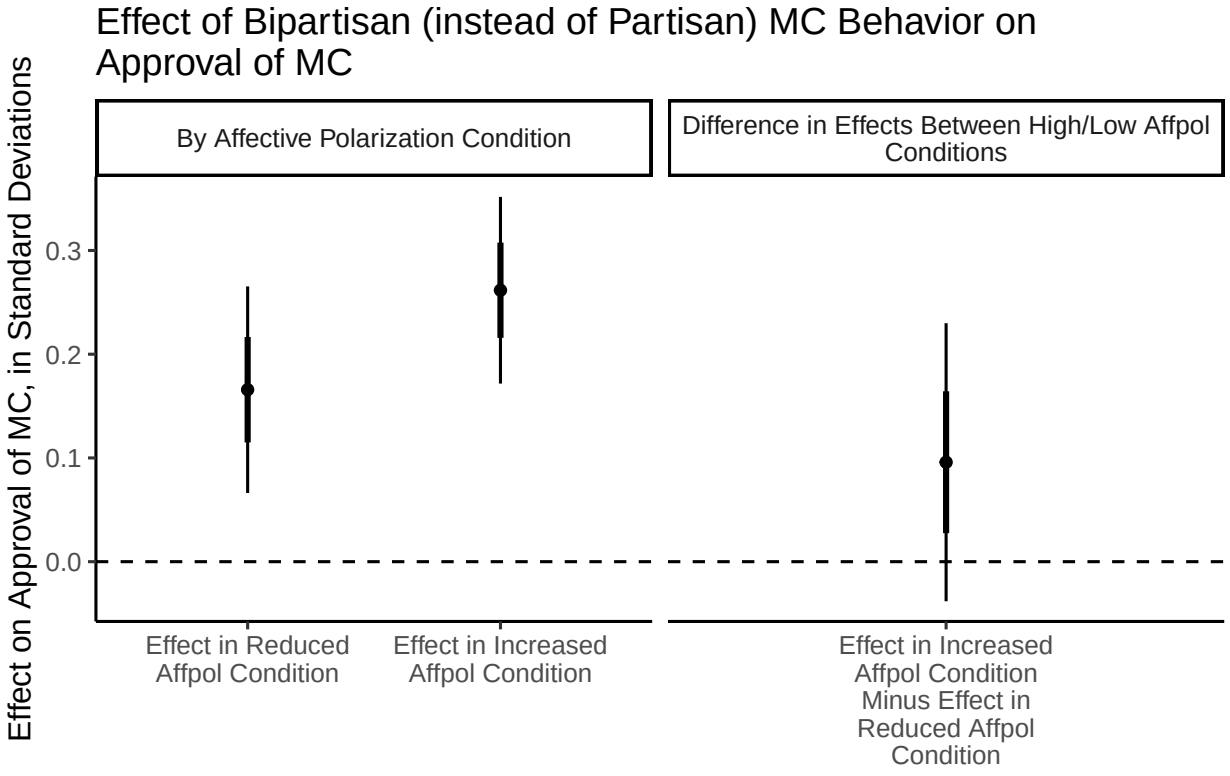
> Some members of Congress work with members of their own party almost all of the time. Other members work with members of both parties. In a recent year, on key issues identified by the American Conservative Union (ACU), Representative Steve LaTourette (a Republican from Ohio) [**bipartisan:** *took the Republican position on about half the votes and the Democratic position on about half the votes* / **partisan:** *almost always voted the Republican position*].

After the vignette, respondents are asked whether they approve or disapprove of the job the representative is doing in Congress. Harbridge and Malhotra (2011) find that respondents on average approve of the job the representative is doing more when they learn about the year of votes where the member voted in a bipartisan manner.

We build on Harbridge and Malhotra's (2011) work by also randomly assigning respondents to either the positive or negative trust game experience before this vignette. If increases in affective polarization decrease voters' preference for bipartisanship, or even lead them to prefer partisan behavior, we should expect to see a weaker or even reversed effect of the bipartisanship condition when voters are assigned to the increased affective polarization condition.

Figure 7 shows the results. The first coefficient shows the effect of the bipartisan condition for respondents in the reduced affective polarization condition. Similar to Harbridge and Malhotra (2011), we find that these respondents are $d = 0.19$ standard deviations more approving of the Member of Congress described as bipartisan instead of partisan. However, inconsistent with speculation in the literature on affective polarization, we do not find that respondents whose level of affective polarization has been increased desire more partisan behavior instead. If anything, as the second coefficient shows, respondents who are assigned to be more affectively polarized reward the bipartisan *even more*, showing an increase of $d = 0.31$ in approving of the job this MC is doing in Congress. The difference in these effects is not significant ($d = 0.11$, $p = 0.17$), but if anything indicates that respondents assigned to have increased affective polarization rewarded bipartisan behavior *more*.

Figure 7: Testing for Decreased Support for Bipartisanship in Congress



Notes: See Figure A6 for condition means. Point estimates are surrounded by one standard error (thick tails) and 95% confidence intervals (thin tails). Point estimates are from multivariate regressions controlling for the pre-registered covariates identified in our pre-analysis plan.

Support for Democratic Norms

We next consider whether increased affective polarization undermines support for a wide variety of democratic norms. A substantial amount of existing research has speculated that increases in affective polarization are causing a breakdown in support for democratic norms. However, we are unfamiliar with any prior work that has experimentally tested this causal claim by inducing exogenous variation in affective polarization.

In Surveys 1, 2, and 4, after respondents played the trust game, we asked a wide variety of questions using multiple approaches to probe support for democratic norms. In particular, we

asked respondents:

- An index probing opposition to democratic norms. This index included items such as "If a journalist accuses a [in party] politician of misconduct without naming their sources, the journalist should be criminally investigated." or "[in party]s should reduce the number of polling stations in areas that typically support [out party]s." (See Appendix B.2.8 for the items.)

- Several vignettes from Lelkes and Westwood (2017). In each of these vignettes, respondents would read a brief, fictional news article and asked for their reaction. In each vignette, we randomized whether the respondent read about the in- or out-party. (See Appendix B.2 for the question text.)

  - *Suppression Vignette.* Respondents read a brief fictional news story with the headline "Police Use Tear Gas on Peaceful Young [Party] Protest," where party was randomized to either the in- or out-party. Respondents were then asked "Do you agree or disagree with the decision to use tear gas on the [Party] protesters?"

  - *Corruption Tolerance Vignette.* The headline was "Donations from Millionaire Businessman to [Party] Super PACs in Question" and respondents were asked "Do you support the investigation of the businessman?"

  - *Antilocution Vignette.* The headline was "[Party]s Drive Congress to Do Less Than Last Years Record-Breaking Low" and respondents were asked "A large website that posts stories from many different news sources is considering sharing the article you just read. Do you think they should post this article?"

  - *Election Override Vignette.* The headline was "Local [Party] Candidate Calls for State Legislature to Decide Election" and respondents were asked "Do you support or oppose the [Party] state legislature determining the outcome of the election?"

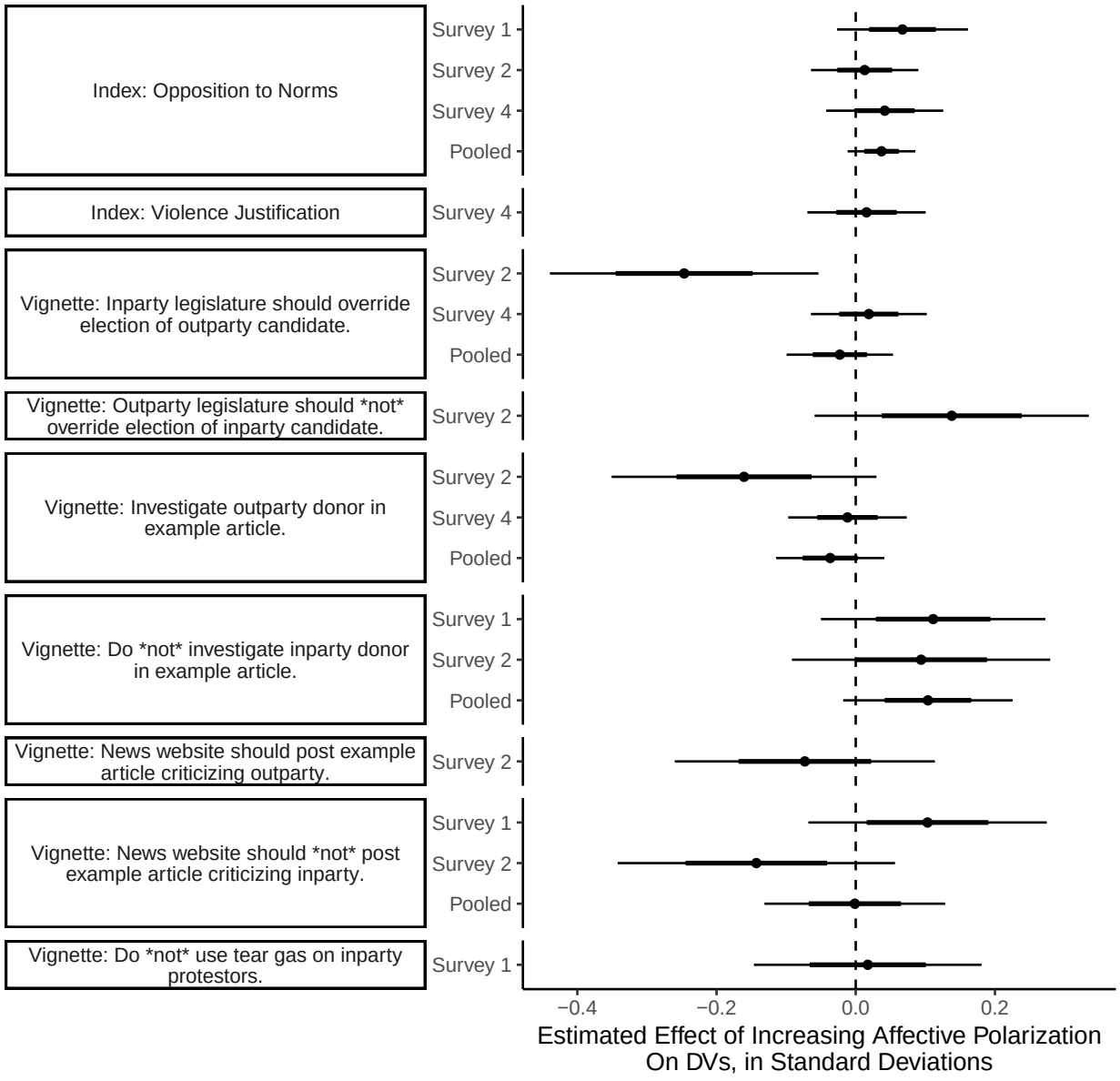
- A partisan violence justification index, four items from Kalmoe and Mason (2019). Statements include "How much would it be justified for [in party]s to use violence to advance their political goals these days?" and "When (if ever) do you think it is ok for an ordinary person who is a [in party] to send physical threats and intimidating messages to [out party] leaders?"

We estimate the effect of increasing affective polarization on these norms indices and vignettes by regressing each outcome on an indicator for the randomly assigned negative trust game experience and two pre-registered pre-treatment covariates (7-point partisanship scale and partisan strength). We code the direction of all the items such that higher values correspond with greater opposition to norms, harsher actions on outpartisans, and leniency towards copartisans.

In Figure 8, we report consistently small and statistically insignificant effects on each of the norms outcomes. We find no evidence that an exogenous increase in affective polarization causes a downstream increase in opposition to democratic norms. Even those estimates with positive point estimates have tight confidence intervals that overlap zero; for example, the top of the 95% confidence interval on the norms index is only a $d = 0.09$ standard deviation increase (and the bottom is below zero). Figure A1 shows that the same pattern holds on the individual items on the norms index.[17] We likewise see no evidence of a greater taste for norm violations in the vignettes; for example, participants in the increased affective polarization condition are no more likely, and if anything less likely, to say they support an inparty-controlled legislature overturning the election of an outparty candidate. Although not all the point estimates are exactly zero, the spread of the point estimates around zero is consistent with what we would expect from sampling variability.

[17] Kingzette et al. (2021) find the relationship between affective polarization and support for norms is only positive for Republicans, and we also examined the effects on the norms scale by party. We find no evidence of heterogeneity; the null effects hold for both parties, and our point estimate (while far from significant) is actually slightly larger for Democrats.

Figure 8: Testing for Undermined Support for Democratic Norms



Notes: Individual items in the norms index reported in Figure A1. Individual items in the violence index reported in Figure A2. Point estimates are surrounded by one standard error (thick tails) and 95% confidence intervals (thin tails). Point estimates are from multivariate regressions controlling for the pre-registered covariates identified in our pre-analysis plan.

Perceptions of Objective Conditions

The last set of outcomes we examine is whether an increase in affective polarization causes changes in perceptions of objective conditions. Prior work has suggested that increases in polarization should increase the "perceptual screen" (Campbell et al. 1960) through which partisans are thought to see the world, leading Democrats and Republicans to perceive objective conditions differently (although see Bullock et al. 2015). For example, Freeder (2020, p. 28) finds that "citizens may be willing to tolerate poor economic performance from their own party" and expresses concern that affective polarization could have produced this outcome, and Iyengar, Sood and Lelkes (2012, p. 428) explicitly worry about a connection between "the increased level of affective polarization" and "partisan bias in perceptions of economic conditions."

We test this prediction in Survey 4 by asking respondents questions about their perceptions of unemployment rates and deaths from COVID-19. After respondents completed the trust game, we asked them both

> "Donald Trump has been President of the United States for four years. What is your best guess about the current unemployment rate in the United States? This is the percent of people who want to work who dont have a job"

and

> "Donald Trump has been President of the United States for four years. What is your best guess about the number of Americans who have died from COVID-19?"

If increases in affective polarization cause differences in stated perceptions of objective conditions, we should observe differences in how Democrats and Republicans answer these questions depending on whether they were randomly assigned to the positive or the negative trust game experience.

As expected, there were large partisan differences in stated perceptions of the number of COVID deaths and of the unemployment rate, with Republicans stating perceptions far lower, and Democrats stating perceptions far higher, of both. However, increasing affective polarization did not appear to meaningfully affect this partisan bias. Figure 9 shows these results. First, for Republicans, we find that increasing affective polarization is associated with no change in stated perceptions of unemployment levels or the number of COVID deaths. Were affective polarization to exacerbate partisan bias, we would expect to see positive coefficients corresponding to stated perceptions of lower levels of both, but we do not. For Democrats, stated perceptions of the number of COVID deaths do not increase, but there is some evidence that Democrats in the high affective polarization condition say they perceive higher levels of unemployment, a result that just meets the threshold for statistical significance at conventional levels ($p = 0.03$). However, this result does not survive any approach to adjusting for multiple comparisons, an issue we return to in the next section.

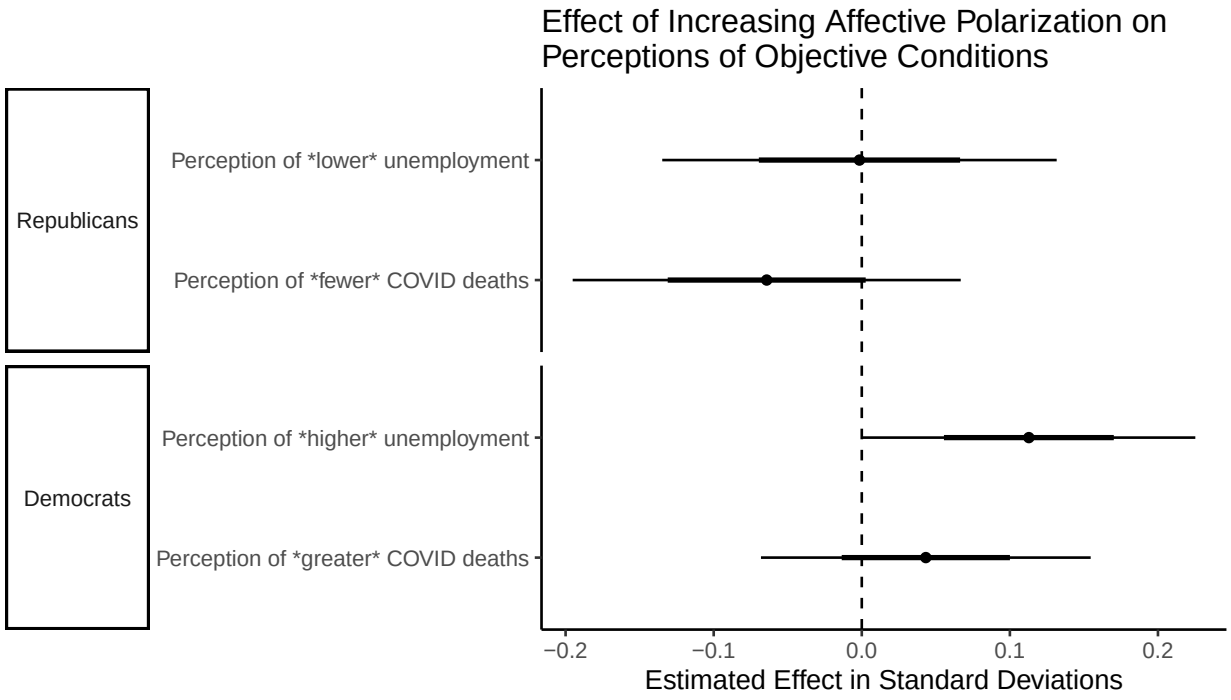
Summarizing the Results

We have presented a large number of estimates of the downstream effects of affective polarization across many outcome areas, most of which we have argued are null. To summarize our results and provide a sense of the consistency of our null results across the areas where our argument predicts, we computed the t-statistic associated with every result about the downstream effects of affective polarization shown or discussed in the paper, with the coefficients oriented in every case such that positive values aligned with the literature's expectations.

Figure 10 shows histograms of all these t-statistics, grouped by area. Each histogram also has a dashed red line at 1.96, the conventional statistical significance threshold.

The first two panels show that the manipulation check items and the social distance items all have highly statistically significant results. However, the distribution of t-statistics in the five political outcome areas are consistent with null results. Out of the 62 t-statistics across these five areas, only 3 are large enough to meet statistical significance at the conventional level, in line with

Figure 9: Testing for Increased Bias in Perceptions of Objective Conditions



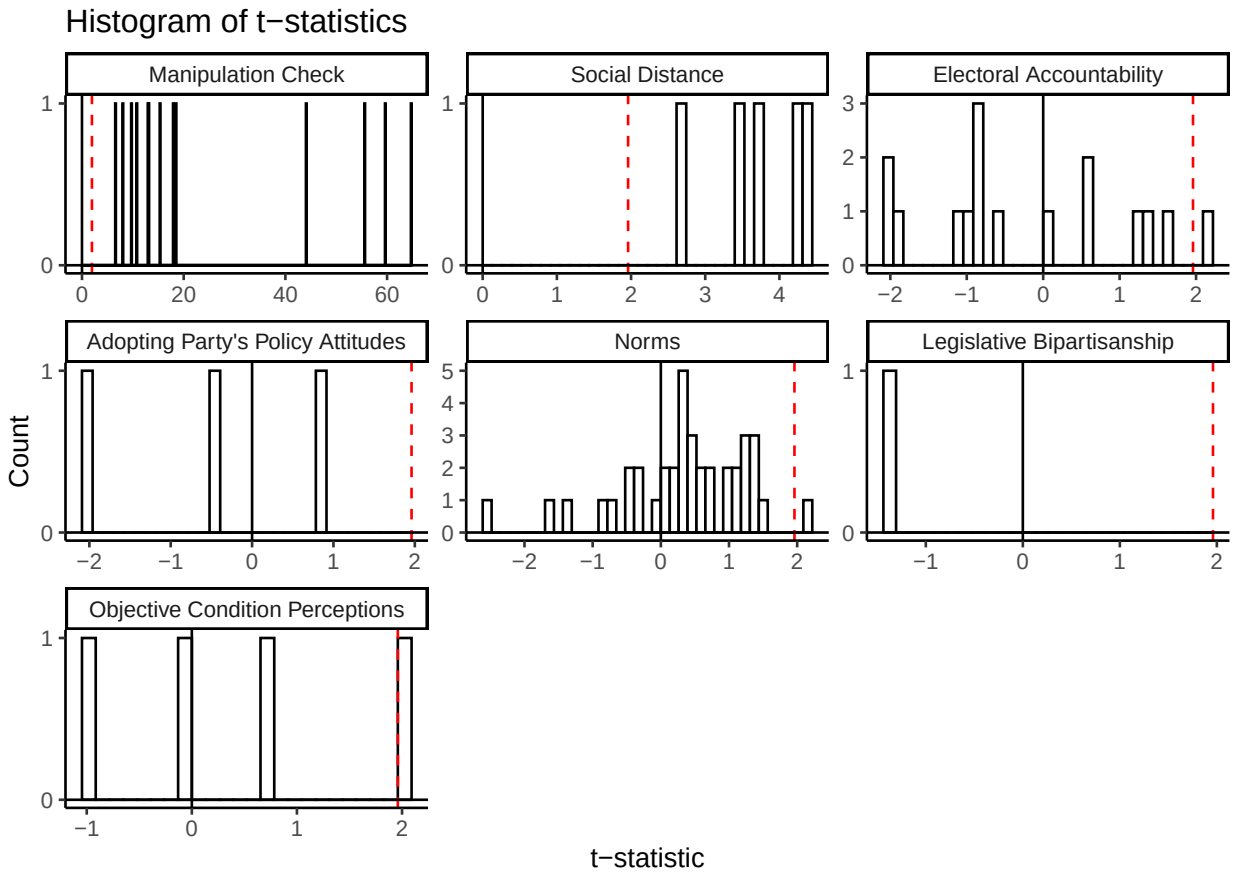
Notes: Point estimates are surrounded by one standard error (thick tails) and 95% confidence intervals (thin tails). Point estimates are from multivariate regressions controlling for the pre-registered covariates identified in our pre-analysis plan.

what would be expected by chance. Moreover each of these 3 statistically significant results are also accompanied by other clearly null results on either replications or closely related outcomes or tests, consistent with them being false positives.[18] Finally, when applying the Anderson's (2008) False Discovery Rate adjustment to all the p-values associated with these t-statistics, all the manipulation check and social distance items remain significant at the conventional threshold, but none of the formerly significant outcomes in the other areas do.

These consistently null results cannot be easily attributed to limited statistical power. Not only

[18]These are as follows. The first significant result is on the "investigate journalists" item on the norms index in Survey 1; however, this did not replicate in Survey 2 or 4 (see Figure A1). The second significant result is the decreased reward Members of Congress earned for casting congruent votes in Survey 3 (see Figure 5c), but the results for incongruent votes in both Survey 1 and Survey 3 (see Figures 5a and 5b) go in the opposite direction. Finally, the third significant result is the increased perception of the unemployment rate among Democrats, but we do not see these results for their perceptions of COVID deaths, nor for Republicans on either outcome (see Figure 9).

Figure 10: Histogram of t-statistics for results discussed and presented in paper, by outcome area



Notes: Histograms of t-statistics for results presented or discussed in the paper, organized by outcome area. The red vertical line is at 1.96, the conventional statistical significance threshold.

was our manipulation of affective polarization powerful (equal to approximately three decades of increased affective polarization) and our sample sizes large, we would expect to observe a larger proportion of statistically significant t-statistics.

Comparing Experimental and Endogenous Estimates

We next show that the null results our experimental approach surfaced stand in contrast to large and positive—but misleading—results one would reach without it. In particular, for every result we discussed or presented in the paper, we computed (1) the experimental estimate we show in the paper (on standardized versions of each outcome) to (2) a non-causal endogenous

estimate, computed by estimating the observational relationship between an "increase" in affective polarization equal to the size our experimental treatment created and each standardized outcome. This allows us to compare endogenous estimates of the form that the existing literature would compute to our causal estimates on the same scale.

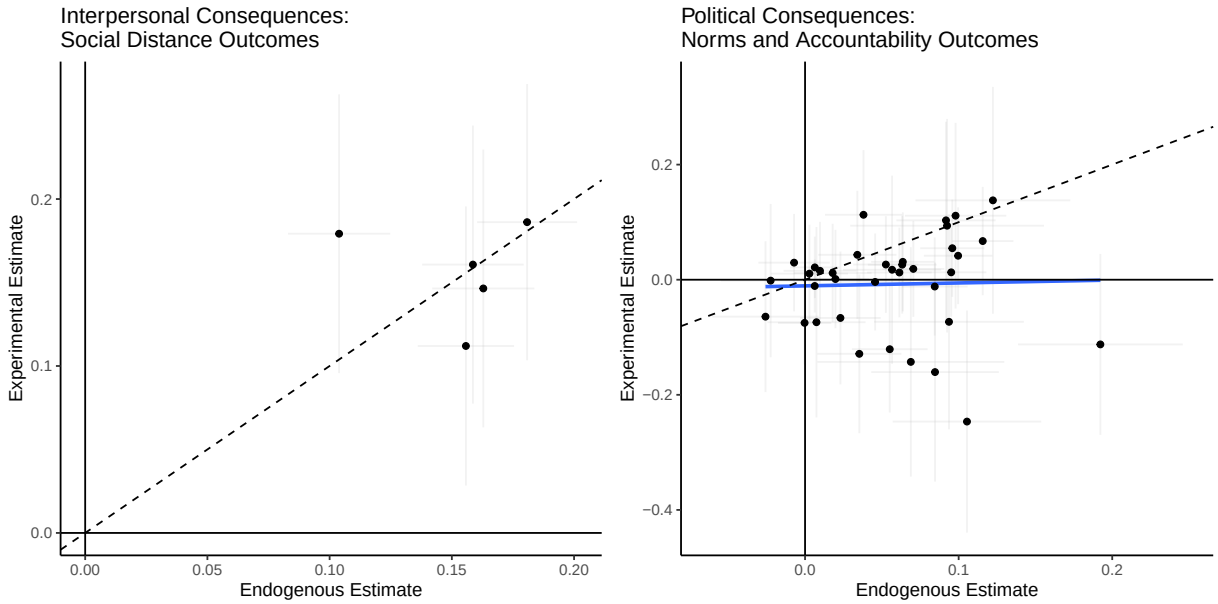
Figure 11 compares these endogenous estimates to our experimental estimates. The first panel, 11a, shows the social distance outcomes. In this domain, where our argument predicts and we find causal effects of affective polarization, the endogenous correlational estimates and the experimental estimates are nearly identical. This is good news for the literature on affective polarization, as it is consistent with affective polarization indeed having a causal effect on the interpersonal items in this literature.

Figure 11b shows the same relationship for the political outcomes. Approximately 89% of the endogenous estimates are positive, meaning that a version of this paper which followed existing literature in examining the observational relationship between affective polarization and these political outcomes would have reported nearly all positive results. However, only 48% of the experimental estimates are positive, essentially a coin flip. Moreover, 70% of the endogenous estimates are larger than the experimental estimates: the endogenous estimates consistently overestimate the causal effects of affective polarization. Finally, the blue line of best fit shows that there is no relationship between the size of the experimental and the endogenous estimates.

These results provide caution for analysts interested in affective polarization's impacts. Affective polarization correlates with a large number of other constructs, including the intensity of partisanship, policy preferences, and many more. But this makes it difficult, if not impossible, for analysts to estimate the causal effects of affective polarization with observational data, and very likely to overestimate them.

Figure 11: Comparing Endogenous Estimates (x-axes) with Experimental Estimates (y-axes)

(a) Social Distance Outcomes **(b)** Political Outcomes



Notes: The x-axis value corresponds with the experimental estimate presented in the paper on standardized versions of each outcome. The y-axis value is a non-causal endogenous estimate of the same effect, computed by estimating the observational relationship between an "increase" in affective polarization equal to the size our experimental treatment created and each standardized outcome. 95% confidence intervals in both regressions are shown surrounding the point estimates. The dotted line is a 45 degree line showing the slope $y = x$; if the correlational estimates were similar to the causal effects we should expect to see the estimates generally cluster around this 45 degree line. Panel 11b shows a line of best fit in blue.

Limitations and Potential Alternative Explanations

Is This Manipulation Informative About the Real World?

Some may wonder whether the increases in measured affective polarization the trust game produced are informative about the effects of increases in affective polarization occurring in the real world. For example, what does variation in affective polarization created by the trust game say about variation in it created by hostile partisan media? It is a common approach across many social science disciplines to experimentally vary a construct with a particular manipulation even though that construct might vary in the real world due to many factors. Nevertheless, this concern can be expressed more precisely in terms of three more specific concerns.

A first concern could be that affective polarization is a multidimensional construct and that the trust game manipulates different latent dimensions of affective polarization than research that has expressed concern about affective polarization has referred to, or that has led to the measured increase in affective polarization. However, a new working paper by Voelkel et al. (2021) exogenously reduces affective polarization using different interventions than we do and also finds null results on downstream outcomes such as support for democratic norms, suggesting that different interventions on affective polarization are unlikely to yield different results. In addition, we note that *none* of the research we cite in Table 1 described affective polarization as a multidimensional construct. Our goal is to assess the speculation offered by this research on its own terms. One would therefore need to "move the goal posts" by redefining the existing literature's claims in order to advance this critique. Although future research on whether affective polarization has multiple latent dimensions would be welcome, such a finding would not impeach our critique of existing research that assumes that it is not. Further, it is far from settled what causes affective polarization (for review, see Boxell, Gentzkow and Shapiro 2020), so it would be unclear what the "right" latent dimensions of affective polarization would be—that is, the literature has not yet pinpointed a source of variation in affective polarization that is the "right one" to study or that is known to produce the "right" variation. In addition, we document large effects on the main measures of affective polarization prior research has used, including towards both members of the mass public and towards political elites. Finally, we *did* find predicted downstream consequences of affective polarization for social distance items (see Figure 3), and Westwood and Peterson (2020) also found downstream predicted effects of this manipulation on racial affect.

A second concern considers the so-called "exclusion restriction": perhaps our trust game manipulation affected constructs other than affective polarization, leading to our results? For this to lead to our null findings even if affective polarization did undermine norms and accountability, the trust game would need to have *bolstered* norms and accountability to an extent that exactly

offset the negative effects of affective polarization. It is unclear why this would be the case. We see another alternative as more likely: as argued in Figure 1b, the real-world forces that have increased affective polarization themselves affect support for democratic norms and accountability, violating the "exclusion restriction" in research that asserts changes in norms and accountability result from changes in affective polarization.

Finally, our intervention and the real-world factors that are increasing affective polarization could increase affective polarization *among different populations*.[19] The most obvious version of this concern would be as follows: strong partisans have driven much of the increase in affective polarization, but what if "leaners" who only lean towards a party were most responsive to our manipulation, and these two groups have different reactions to affective polarization? Our evidence does not support this concern: Table A1 shows that our manipulation appears to have increased affective polarization for every single partisan group on the traditional 7-point scale of partisan identity strength. The manipulation affectively polarizes strong partisans, not strong partisans, and leaners alike; and the largest effects are among the populations (strong partisans) who have affectively polarized over the last three decades.

With these points noted, we would still welcome other research that replicated our findings with other approaches to manipulating affective polarization.[20] Any such research can use the approaches we have developed to measuring downstream political outcomes to consider the generalizability of our results.

[19]In other words, the local average treatment effect of our manipulation could be local to different individuals than the increase in affective polarization over the last several decades has occurred among. If these two groups of individuals had a different reaction to increased affective polarization, our results could differ.

[20]As of this writing, we are not aware of any other approaches for doing so that would be suitable for our research design. Most of the other approaches available in the literature did not have effects on affective polarization large enough to produce statistically precise estimates of the downstream impacts of affective polarization (e.g., Levendusky 2018). Some interventions show larger effects (Rossiter 2020), but could affect downstream political outcomes for other reasons.

Other Limitations and Alternative Explanations

One other important limitation of our research is that we only examine the immediate, short-run effects of affective polarization on outcomes that can be measured in a survey. It is possible that there are other long-run or indirect effects of affective polarization that our research design cannot measure, such as the information to which people are exposed in interpersonal conversations. For example, if affective polarization reduces interactions with outpartisans (Mosleh et al. 2021) and conversations with outpartisans expose people to rationales for the other party's policies (Druckman and Nelson 2003), affective polarization could have indirect effects on people's political attitudes (through social interaction) even if it does not have direct effects on them. This is worthy of future study.

In addition, while we tested for downstream effects on the outcomes most commonly named in the literature, future research can use the paradigm we present to test for effects on outcomes such as media choice, information processing, the decision to engage in identity-expressive actions, and more (e.g., Peterson, Goel and Iyengar 2019; Druckman et al. 2020*b*). Likewise, although our key accountability results had the added realism of providing individuals real information about how their actual Members of Congress voted, studies that manipulate affective polarization in the field or measure behavioral outcomes would also be welcome and speak to external validity.

A final potential concern is that the momentary increase in affective polarization caused by the intervention may have "worn off" by the time our measures were asked. We assess this concern by examining the outcomes asked immediately after the trust game portion of the survey. However, examining the first dependent variables asked after the manipulation in each survey reveals that they do not have consistently different or larger effects.[21]

[21]Table B1 shows the items on each survey and their order.

Discussion

Across many democratic societies, citizens express increasing dislike for those of opposing political parties (Boxell, Gentzkow and Shapiro 2020). Research on this so-called affective polarization is one of the most influential and active areas of research across multiple social science disciplines and has inspired two sets of worries among scholars.

A first worry, following from social identity theory and borne out in a great deal of empirical research, is that affective polarization may affect a variety of interpersonal judgments. For instance, citizens may discriminate against outpartisan citizens or avoid interactions with them, with implications for who people hire or date (e.g., McConnell et al. 2018; Huber and Malhotra 2017). Our results confirm the causal relationship between affective polarization and outcomes in the interpersonal domain.

A second worry goes beyond these strictly interpersonal implications: that affective polarization affects citizens' *political* behaviors. Seeing negative trends in many of these areas, a number of scholars have speculated, or sometimes simply asserted, that affective polarization is to blame (see examples in Table 1). However, little empirical research has assessed these concerns. Our results call for skepticism when it comes to this second worry.

To support our argument, we leveraged one of the only known approaches to introducing exogenous variation in affective polarization, a scripted trust game. Our manipulation produced large differences in affective polarization as the previous literature has measured and defined it. We used this exogenous variation in affective polarization to trace its downstream causal effects on a variety of outcomes. Consistent with our argument, we find large impacts on social distance measures, in line with previous literature. But we found no impacts in a variety of political domains: electoral accountability, propensity to adopt one's party's policy positions, support for legislative bipartisanship, support for democratic norms, or bias in perceptions of objective conditions.

We stress that our results are in no way inconsistent with the large body of empirical research to date on the interpersonal impacts of affective polarization—indeed, we found that the observational relationship between affective polarization and desire for social distance from outpartisans is essentially identical to our estimate of the causal effects, consistent with the relationships previous research has uncovered indeed being causal effects of affective polarization. But it is possible for a citizen to simultaneously not want their son or daughter to marry a member of the political outparty (Iyengar, Sood and Lelkes 2012) and still support their legislators voting in a bipartisan manner, or being willing to vote for an outpartisan politician if they do in fact represent their views better.

Our findings therefore suggest that affective polarization is best understood as a social phenomenon that regulates individual and group decision-making and not attitudes toward individual representatives and or broad democratic norms. While contrary to the literature's widespread suggestion and potentially counter-intuitive, this is consistent with partisanship transforming into a social identity that exists beyond ideology and issues of representation. Affective polarization remains a deeply troubling phenomenon, but our results should help allay concerns that it perverts our democratic system.

Implications for Future Research

Our results suggest several implications for future research and concrete implications for those seeking to improve American democracy.

First, our work suggests that future research on the political implications of affective polarization may wish to concentrate on those areas where the interpersonal domain and the political domain are inherently linked. This could be a fruitful and important area for study that holds important normative implications. For example, among other forms of interactions with outpartisans our respondents sought to avoid, we found that becoming more affectively polarized made them want to avoid discussing politics with outpartisan citizens which is consistent with other work (Druckman et al. 2020*c*; Klar, Krupnikov and Ryan 2018).

Second, our results suggest implications for efforts to improve democracy. There are a number of hypotheses for what is behind negative trends in democratic societies, such as the decline in ticket splitting in the United States (e.g., Hopkins 2018), the decay of certain democratic norms (e..g, Graham and Svolik 2020), or even other psychological constructs related to partisanship but conceptually distinct from affect or affective polarization, such as the strength of partisan identity. In order to reverse these trends, researchers must build an understanding of which factors are contributing to these trends and which are not. We see our research as a part of this effort. It suggests that attempting to reduce affective polarization—although potentially valuable for many reasons—may not be the most effective way to reverse these trends. Indeed, consistent with our findings, Voelkel et al. (2021) report in a new working paper that interventions that decrease affective polarization do not have any downstream consequences for democratic norms—a conceptual replication of our findings that underscore this important implication. Our findings in no way end the conversation about the potential political impacts of affective polarization on the functioning of American democracy, but rather represent a step towards more focused theorizing and rigorous empirical study of this important topic.

References

Abramowitz, Alan I and Steven Webster. 2016. "The rise of negative partisanship and the nationalization of US elections in the 21st century." *Electoral Studies* 41:12–22.

Anderson, Michael L. 2008. "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American statistical Association* 103(484):1481–1495.

Ashworth, Scott. 2012. "Electoral accountability: Recent theoretical and empirical work." *Annual Review of Political Science* 15:183–201.

Berg, Joyce, John Dickhaut and Kevin McCabe. 1995. "Trust, reciprocity, and social history."

Games and economic behavior 10(1):122–142.

Boxell, Levi, Matthew Gentzkow and Jesse M Shapiro. 2017. Is the internet causing political polarization? Evidence from demographics. Technical report National Bureau of Economic Research.

Boxell, Levi, Matthew Gentzkow and Jesse M. Shapiro. 2020. "Cross-Country Trends in Affective Polarization." Working paper, available at `https://web.stanford.edu/~gentzkow/research/cross-polar.pdf`.

Brody, Richard A and Paul M Sniderman. 1977. "From life space to polling place: The relevance of personal concerns for voting behavior." *British Journal of Political Science* 7(3):337–360.

Bullock, John G, Alan S Gerber, Seth J Hill, Gregory A Huber et al. 2015. "Partisan Bias in Factual Beliefs about Politics." *Quarterly Journal of Political Science* 10(4):519–578.

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

Chen, M. Keith and Ryne Rohla. 2018. "The effect of partisanship and political advertising on close family ties." *Science* 360(6392):1020–1024.

Coppock, Alexander and Donald P. Green. 2021. "Do Belief Systems Exhibit Dynamic Constraint?" *Journal of Politics* .

Dickson, Eric S. 2011. Economics versus Psychology Experiments. Stylization, Incentives, and Deception. In *Cambridge Handbook of Experimental Political Science*. Cambridge University Press.

Druckman, James N and Kjersten R Nelson. 2003. "Framing and deliberation: How citizens' conversations limit elite influence." *American Journal of Political Science* 47(4):729–745.

Druckman, James N and Matthew S Levendusky. 2019. "What do we measure when we measure affective polarization?" *Public Opinion Quarterly* 83(1):114–122.

Druckman, James N and Richard M Shafranek. 2020. "The Intersection of Racial and Partisan Discrimination: Evidence from a Correspondence Study of Four-Year Colleges." *The Journal of Politics* 82(4):1602–1606.

Druckman, James N, Samara Klar, Yanna Krupnikov, Matthew Levendusky and John Barry Ryan. 2020*a*. "Affective polarization, local contexts and public opinion in America." *Nature Human Behaviour* pp. 1–11.

Druckman, James N, Samara Klar, Yanna Krupnikov, Matthew Levendusky and John Barry Ryan. 2020*b*. "How Affective Polarization Shapes Americans' Political Beliefs: A Study of Response to the COVID-19 Pandemic." *Journal of Experimental Political Science* pp. 1–12.

Druckman, James N, Samara Klar, Yanna Krupnikov, Matthew Levendusky and John Barry Ryan. 2020*c*. "Mis-estimating affective polarization." *The Journal of Politics* .

Druckman, James N, SR Gubitz, Ashley M Lloyd and Matthew S Levendusky. 2019. "How incivility on partisan media (de) polarizes the electorate." *The Journal of Politics* 81(1):291–295.

Finkel, Eli J, Christopher A Bail, Mina Cikara, Peter H Ditto, Shanto Iyengar, Samara Klar, Lilliana Mason, Mary C McGrath, Brendan Nyhan, David G Rand et al. 2020. "Political sectarianism in America." *Science* 370(6516):533–536.

Fowler, Anthony. 2020. "Partisan Intoxication or Policy Voting?" *Quarterly Journal of Political Science* .

Freeder, Sean. 2020. "It's No Longer the Economy, Stupid: Selective Perception and Attribution

of Economic Outcomes." Available at https://seanfreeder.files.wordpress.com/2020/12/paper_economy.pdf.

Gidron, Noam, James Adams and Will Horne. 2020. *American Affective Polarization in Comparative Perspective*. Cambridge University Press.

Gift, Karen and Thomas Gift. 2015. "Does Politics Influence Hiring? Evidence from a Randomized Experiment." *Political Behavior* 37(3):653–75.

Graham, Matthew H and Milan W Svolik. 2020. "Democracy in America? Partisanship, Polarization, and the Robustness of Support for Democracy in the United States." *American Political Science Review* 114(2):392–409.

Harbridge, Laurel and Neil Malhotra. 2011. "Electoral incentives and partisan conflict in Congress: Evidence from survey experiments." *American Journal of Political Science* 55(3):494–510.

Hersh, Eitan D and Matthew N Goldenberg. 2016. "Democratic and Republican physicians provide different care on politicized health issues." *Proceedings of the National Academy of Sciences* 113(42):11811–11816.

Hetherington, Marc J. 2001. "Resurgent mass partisanship: The role of elite polarization." *American Political Science Review* 95(3):619–631.

Hetherington, Marc J and Thomas J Rudolph. 2015. *Why Washington won't work: Polarization, political trust, and the governing crisis*. University of Chicago Press.

Hopkins, Daniel J. 2018. *The increasingly United States: How and why American political behavior nationalized*. University of Chicago Press.

Hopkins, Daniel J and Jonathan Mummolo. 2017. "Assessing the Breadth of Framing Effects." *Quarterly Journal of Political Science* 12(1):37–57.

Huber, Gregory A and Neil Malhotra. 2017. "Political Homophily in Social Relationships: Evidence from Online Dating Behavior." *The Journal of Politics* 79(1):269–283.

Iyengar, Shanto, Gaurav Sood and Yphtach Lelkes. 2012. "Affect, Not Ideology: A Social Identity Perspective on Polarization." *Public Opinion Quarterly* .

Iyengar, Shanto and Masha Krupenkin. 2018. "Partisanship as social identity; Implications for the study of party polarization." *The Forum* 16(1):23–45.

Iyengar, Shanto, Yphtach Lelkes, Matthew Levendusky, Neil Malhotra and Sean J Westwood. 2019. "The origins and consequences of affective polarization in the United States." *Annual Review of Political Science* 22:129–146.

Kalla, Joshua L and David E Broockman. 2018. "The minimal persuasive effects of campaign contact in general elections: Evidence from 49 field experiments." *American Political Science Review* 112(1):148–166.

Kalmoe, Nathan P and Lilliana Mason. 2019. Lethal mass partisanship: Prevalence, correlates, and electoral contingencies. In *NCAPSA American Politics Meeting*.

Kayser, Mark Andreas and Michael Peress. 2012. "Benchmarking across borders: electoral accountability and the necessity of comparison." *American Political Science Review* pp. 661–684.

Kingzette, Jon, James N. Druckman, Samara Klar, Yanna Krupnikov, Matthew Levendusky and John B. Ryan. 2021. "How Affective Polarization Undermines Support for Democratic Norms." *Public Opinion Quarterly* .

Klar, Samara, Yanna Krupnikov and John Barry Ryan. 2018. "Affective polarization or partisan disdain? Untangling a dislike for the opposing party from a dislike of partisanship." *Public Opinion Quarterly* 82(2):379–390.

Klein, Ezra. 2020. *Why we're polarized*. Simon and Schuster.

Lelkes, Yphtach, Gaurav Sood and Shanto Iyengar. 2017. "The hostile audience: The effect of access to broadband internet on partisan affect." *American Journal of Political Science* 61(1):5–20.

Lelkes, Yphtach and Sean J Westwood. 2017. "The limits of partisan prejudice." *The Journal of Politics* 79(2):485–501.

Levendusky, Matthew and Dominik Stecula. 2021. *We Need to Talk: How Cross-Party Dialogue Reduces Affective Polarization*. Cambridge University Press.

Levendusky, Matthew S. 2018. "Americans, not partisans: Can priming American national identity reduce affective polarization?" *The Journal of Politics* 80(1):59–70.

Little, Andrew T, Keith Schnakenberg and Ian R Turner. 2020. "Motivated reasoning and democratic accountability." Working paper, available at `https://osf.io/preprints/socarxiv/esfy6/`.

Mason, Lilliana. 2014. ""I Disrespectfully Agree": The Differential Effects of Partisan Sorting on Social and Issue Polarization." *American Journal of Political Science* .

Mason, Lilliana. 2018. *Uncivil agreement: How politics became our identity*. University of Chicago Press.

McConnell, Christopher, Yotam Margalit, Neil Malhotra and Matthew Levendusky. 2018. "The economic consequences of partisanship in a polarized era." *American Journal of Political Science* 62(1):5–18.

McCoy, Jennifer and Murat Somer. 2019. "Toward a theory of pernicious polarization and how it harms democracies: Comparative evidence and possible remedies." *The Annals of the American Academy of Political and Social Science* 681(1):234–271.

McDonald, Jared. 2020. "Avoiding the Hypothetical: Why "Mirror Experiments" are an Essential Part of Survey Research.".

Michelitch, Kristin. 2015. "Does electoral competition exacerbate interethnic or interpartisan economic discrimination? Evidence from a field experiment in market price bargaining." *The American Political Science Review* 109(1):43.

Mosleh, Mohsen, Cameron Martel, Dean Eckles and David G. Rand. 2021. "Shared Partisanship Dramatically Increases Social Tie Formation in a Twitter Field Experiment." *Proceedings of the National Academy of Sciences* .

Orr, Lilla V and Gregory A Huber. 2020. "The policy basis of measured partisan animosity in the United States." *American Journal of Political Science* 64(3):569–586.

Panagopoulos, Costas. 2016. "All about that base: Changing campaign strategies in US presidential elections." *Party Politics* 22(2):179–190.

Peterson, Erik, Sharad Goel and Shanto Iyengar. 2019. "Partisan selective exposure in online news consumption: Evidence from the 2016 presidential campaign." *Political Science Research and Methods* pp. 1–17.

Pierson, Paul and Eric Schickler. 2020. "Madison's Constitution Under Stress: A Developmental Analysis of Political Polarization." *Annual Review of Political Science* 23:37–58.

Rossiter, Erin. 2020. "The Consequences of Interparty Conversation on Outparty Affect and Stereotypes." Working paper, available at `http://erossiter.com/files/conversations.pdf`.

Tajfel, Henri and John Turner. 1979. An Integrative Theory of Intergroup Conflict. In *The Social Psychology of Intergroup Relations*, ed. W.G. Austin and Stephen Worchel. Monterey, CA: Brooks Cole.

Voelkel, Jan G., James Chu, Michael N. Stagnaro, Joseph S. Mernyk, Sophia Pink, Chrystal Redekopp, James Druckman, David Rand and Robb Willer. 2021. "Interventions Reducing Affective Polarization Do Not Improve Anti-Democratic Attitudes." Working paper, Stanford University.

Westwood, Sean J and Erik Peterson. 2020. "The Inseparability of Race and Partisanship in the United States." *Political Behavior* pp. 1–23.

Yair, Omer. 2020. "A Note on the Affective Polarization Literature." Working paper, available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3771264.

Online Appendix

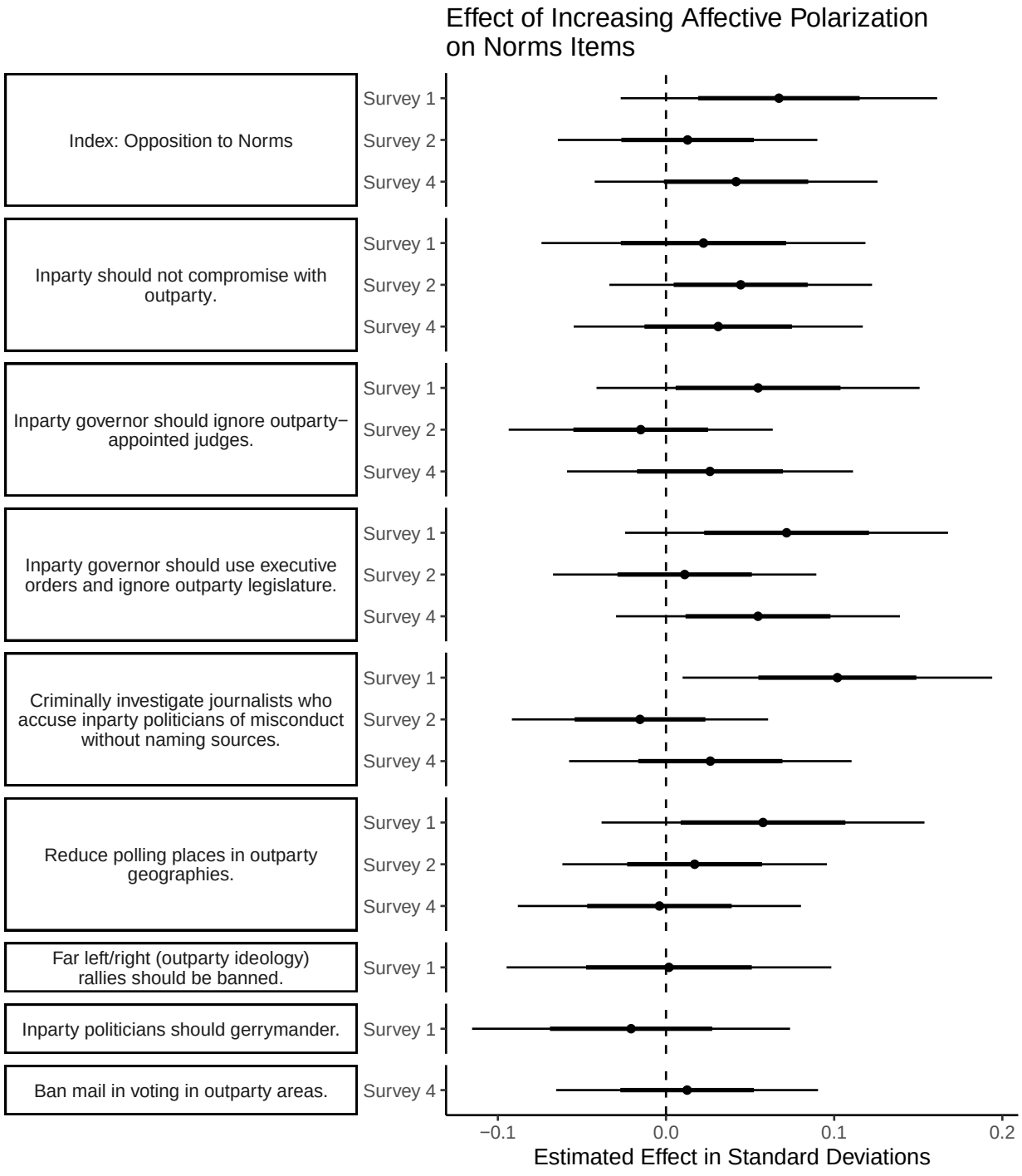
A Additional Tables and Figures

Table A1: Effect of Negative Trust Game Experience on Affective Polarization Towards Mass Public and Towards Elites, by Strength of Party ID

	DV = Outparty - Inparty Mass Public Feeling Therms	DV = Outparty - Inparty Elite Feeling Therms
PID = Not Strong Democrat	-20.3***	-22.3***
	(2.88)	(3.25)
PID = Lean Democrat	-22.3***	-19.0***
	(3.12)	(3.38)
PID = Lean Republican	-23.5***	-24.7***
	(3.76)	(4.12)
PID = Not Strong Republican	-29.8***	-32.2***
	(2.87)	(3.33)
PID = Strong Republican	-5.55	-4.15
	(3.29)	(3.55)
Strong Democrat X Negative Game Experience	**15.7*** (2.65)**	**9.86*** (2.88)**
Not Strong Democrat X Negative Game Experience	**13.9*** (3.07)**	**10.2** (3.24)**
Lean Democrat X Negative Game Experience	**10.2** (3.37)**	**4.46 (3.42)**
Lean Republican X Negative Game Experience	**10.8* (4.34)**	**5.21 (4.59)**
Not Strong Republican X Negative Game Experience	**16.5*** (3.44)**	**11.2** (3.71)**
Strong Republican X Negative Game Experience	**18.0*** (3.69)**	**12.6*** (3.71)**
Constant	37.0***	44.2***
	(1.98)	(2.19)
N	2135	2135

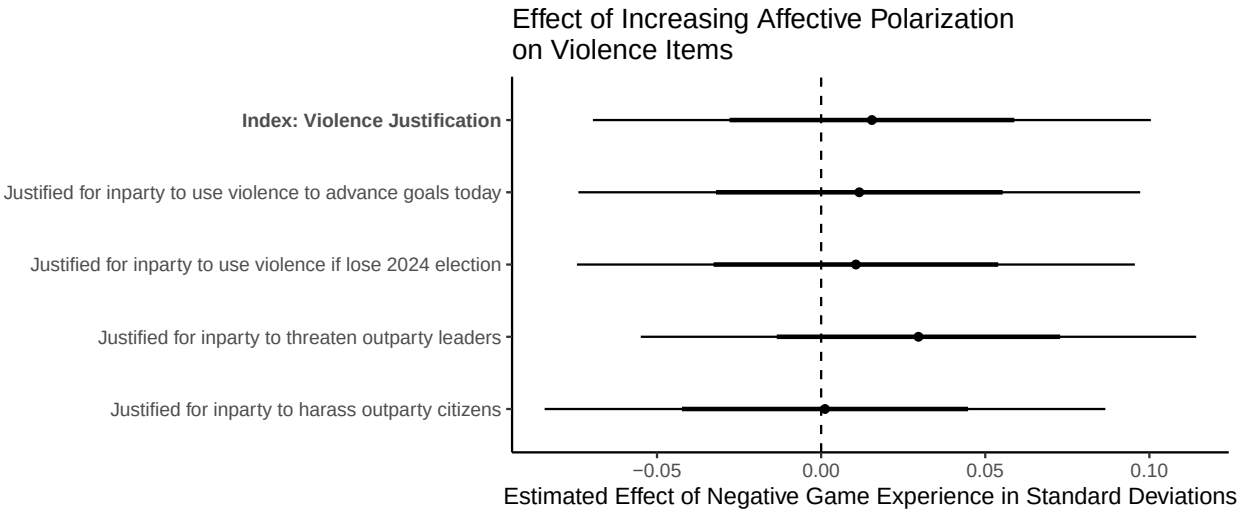
*Notes: The bolded rows show the effect of a negative game experience by party ID. The non-bolded rows show differences between levels of affective polarization between different partisan groups within the control (positive game experience) condition, with PID = Strong Democrat as the omitted baseline category. Data are from Survey 4, as Survey 4 was the only survey that asked about affective polarization towards elites. Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.*

Figure A1: Effect of Increasing Affective Polarization on Individual Items in Norms Index



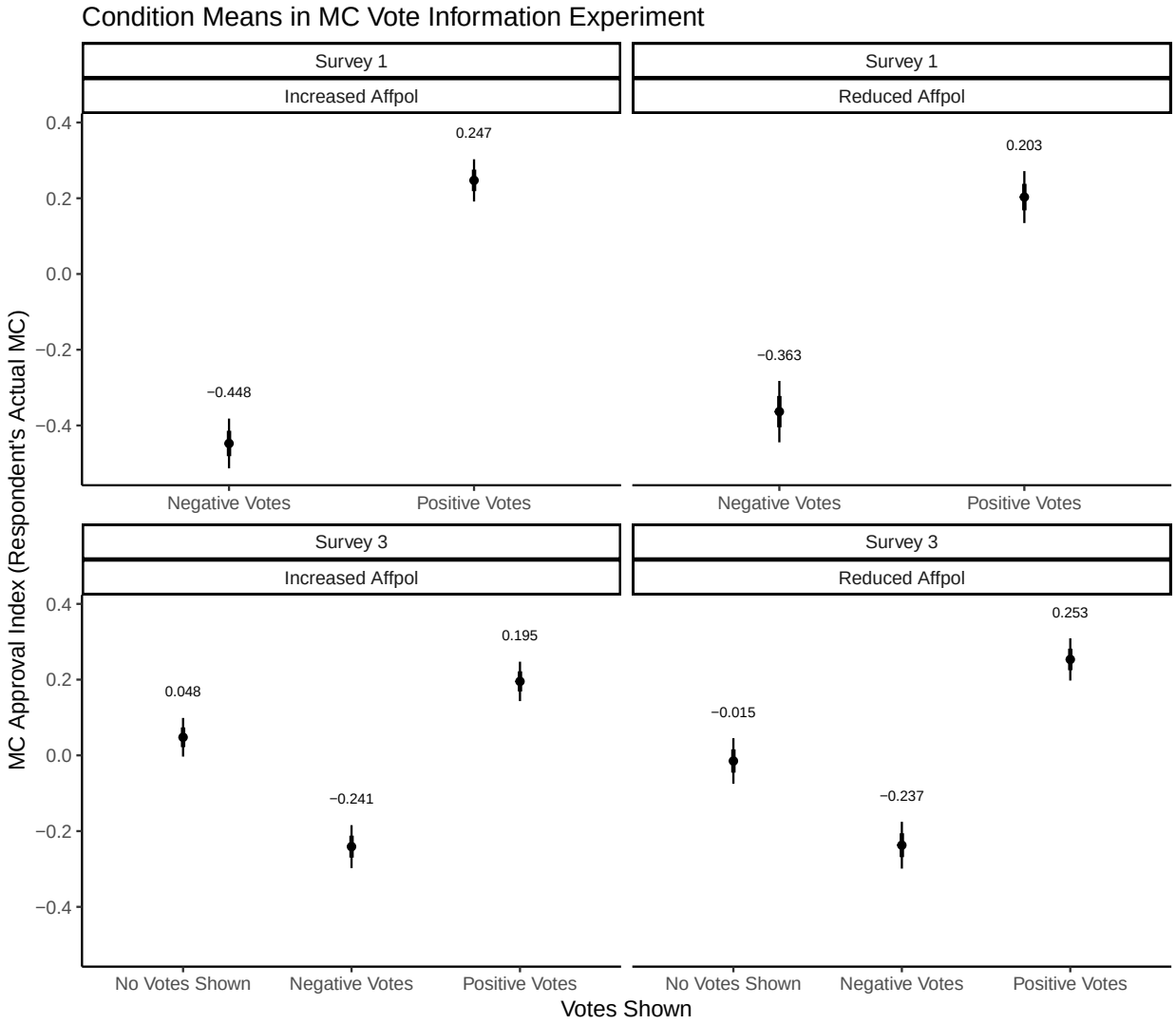
Notes: The first set of coefficients reproduces the estimates on the Norms index shown in Figure 8 in the main text. The remaining coefficients show estimates on the individual items in this index. For the full text of these items, see Appendix B.2.8. Point estimates are surrounded by one standard error (thick tails) and 95% confidence intervals (thin tails). Point estimates are from multivariate regressions controlling for the pre-registered covariates identified in our pre-analysis plan.

Figure A2: Effect of Increasing Affective Polarization on Individual Items in Violence Justification Index



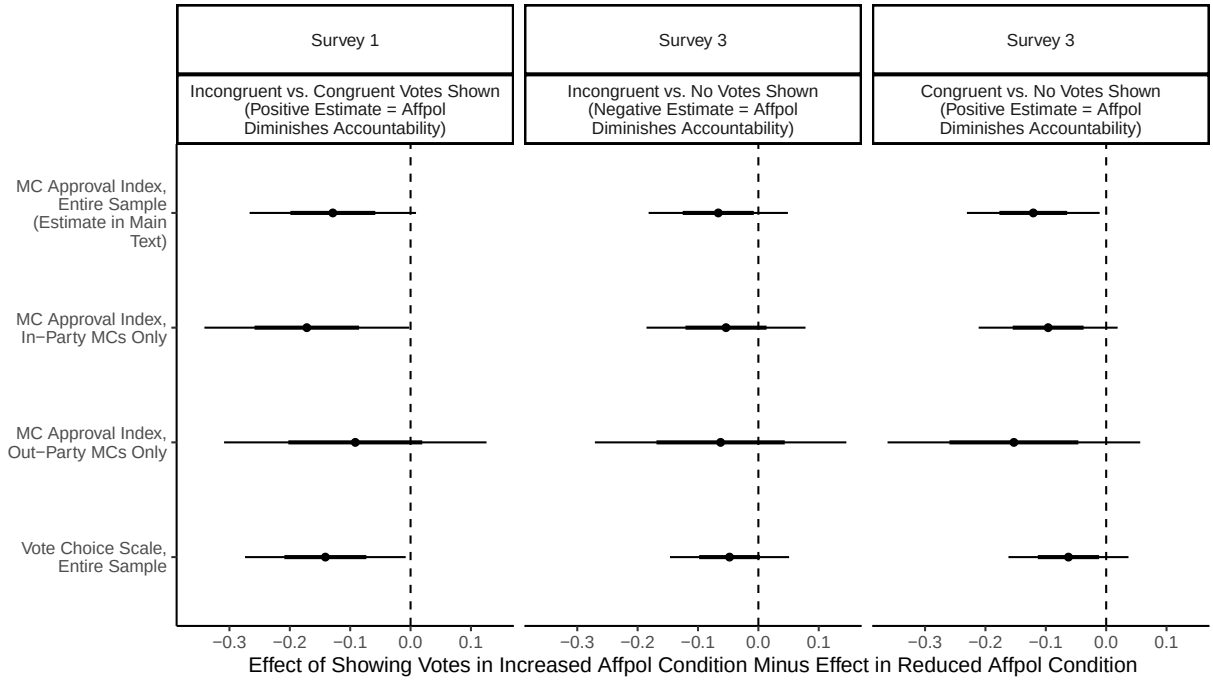
Notes: The first set of coefficients reproduces the estimates on the Violence Justification index shown in Figure 8 in the main text. The remaining coefficients show estimates on the individual items in this index. Point estimates are surrounded by one standard error (thick tails) and 95% confidence intervals (thin tails). Point estimates are from multivariate regressions controlling for the pre-registered covariates identified in our pre-analysis plan.

Figure A3: Condition Means in MC Vote Choice Experiment



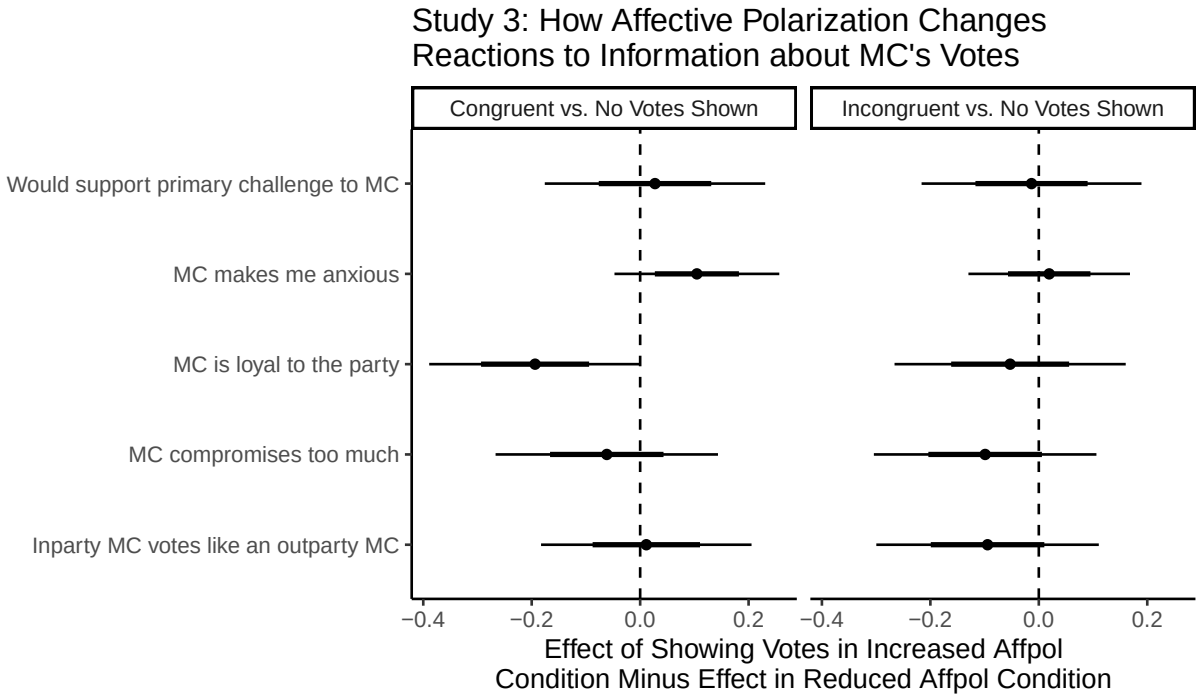
Notes: This Figure shows the condition means in the experiment reported in Figure 5 in the main text. Point estimates are surrounded by one standard error (thick tails) and 95% confidence intervals (thin tails). Point estimates are predicted probabilities from multivariate regressions controlling for the pre-registered covariates identified in our pre-analysis plan.

Figure A4: Testing for Desensitization: Robustness



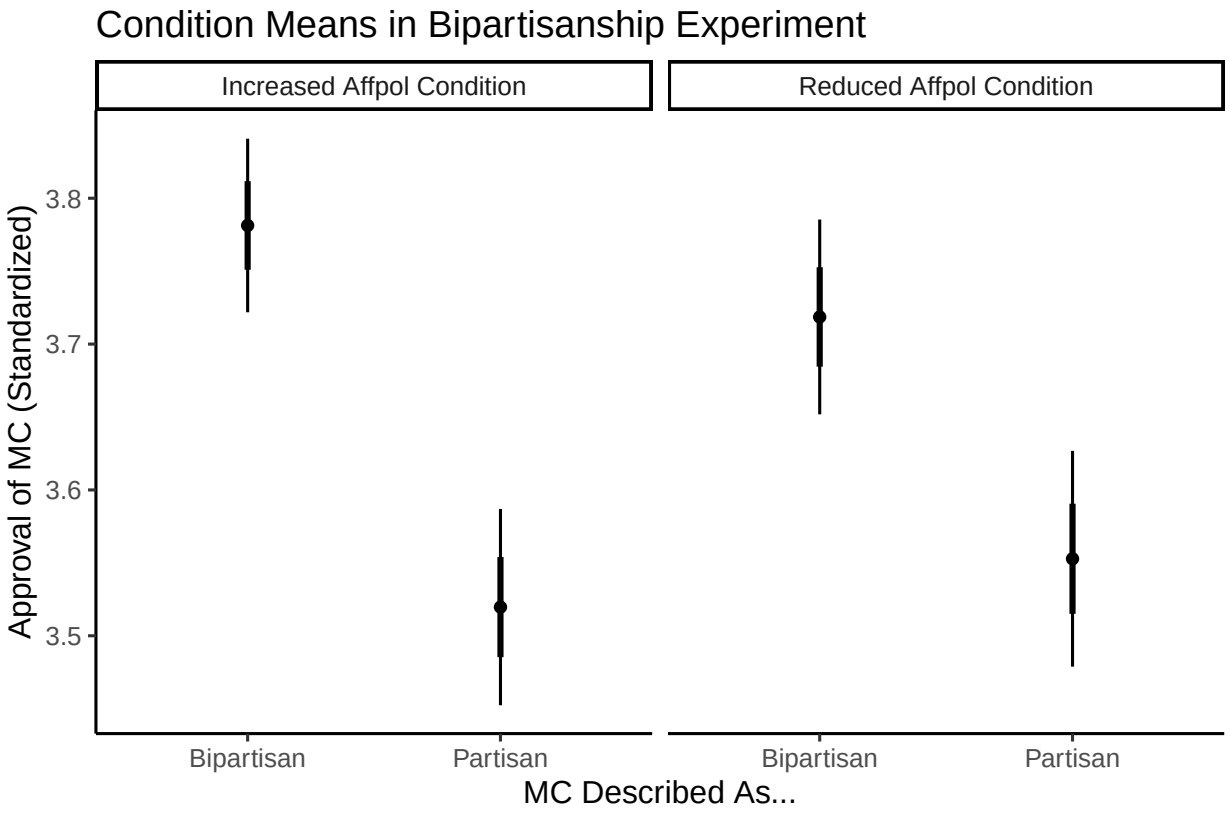
Notes: This Figure helps assess the robustness of the results reported in Figure 5 in the main text. The first row reports the same results as shown in the main text, the next row shows the estimates for respondents with in-party MCs only, the following row the estimates for respondents with out-party MCs only, and the final row the results in the entire sample when using just the vote choice scale as an outcome instead of the entire MC approval index. Point estimates are surrounded by one standard error (thick tails) and 95% confidence intervals (thin tails). Point estimates are from multivariate regressions controlling for the pre-registered covariates identified in our pre-analysis plan.

Figure A5: Testing for Desensitization: Potential Mechanisms



Notes: This Figure reports the difference-in-differences estimate from Figures 5b and 5c in the main text but on other "mechanism" items. Point estimates are surrounded by one standard error (thick tails) and 95% confidence intervals (thin tails). Point estimates are from multivariate regressions controlling for the pre-registered covariates identified in our pre-analysis plan.

Figure A6: Condition Means in Bipartisanship Experiment



Notes: This Figure shows the condition means in the bipartisanship experiment reported in Figure 7 in the main text. Point estimates are surrounded by one standard error (thick tails) and 95% confidence intervals (thin tails). Point estimates are from multivariate regressions controlling for the pre-registered covariates identified in our pre-analysis plan.

B Surveys

B.1 Survey Contents and Order

Table B1: Survey Contents and Order, by Survey

	Survey 1	**Survey 2**	**Survey 3**	**Survey 4**
Demographics Battery	1	1	1	1
Asked Own Views on Issues Pre-Treatment	2		2	
Trust Game	3	2	3	2
Manipulation Check: Treated Fairly	4	3	4	3
Manipulation Check: Mass Feeling Thermometers	5	4	5	4
Manipulation Check: Salience of Partisan Identity	6			
Manipulation Check: Elite Feeling Thermometers				5
Social Distance Items				6
Shown MC Votes & Asked MC DVs	7		6	
Norms Index	8	5		7
Asked Own Views on Issues Post-Treatment	9		7	
Suppression Vignette	10 (In-party only)			
Corruption Tolerance Vignette	11 (In-party only)	6 (Party randomized)		9 (Out-party only)
Antilocution Vignette	12 (In-party only)	8 (Party randomized)		
Election Override Vignette		7 (Party randomized)		8 (In-party only)
Bipartisanship Vignette		9		
Violence Justification Index				10
Perceptions of Objective Conditions				11
Final Demographics	13	10	8	12

B.2 Question Text

B.2.1 Demographics Battery

- What is your gender?
  - Male
  - Female
  - Other
- What is your race and ethnicity? Select all that apply.
  - White/Caucasian
  - Hispanic/Latino
  - African American
  - Asian
  - Native American
  - Pacific Islander
  - Other
- What is your combined annual household income?
  - Less than $30,000
  - $30,000 - 39,999
  - $40,000 - 49,999
  - $50,000 - 59,999
  - $60,000 - 69,999
  - $70,000 - 79,999

- $80,000 - 89,999
- $90,000 - 99,999
- $100,000 or more
- Prefer not to say

B.2.2 Own Views on Issues Pre-Treatment

Congress considered many important laws recently.

If you were in Congress would you vote FOR or AGAINST each of the following? (Options: FOR, AGAINST)

- **Working Families Flexibility Act of 2017.** Allows employers to give employees who worked overtime paid time off instead of only overtime pay.
- **Reducing Regulatory Burdens Act of 2017.** Allows pesticides to be sprayed near water sources without obtaining a permit.
- **Thin Blue Line Act.** Allows the death penalty in the case of a murder or attempted murder of police officers, correctional officers, firefighters, or other first responders.
- **Save Local Businesses Act.** If an employee working for a company through a 'temp' agency is injured, only the temp agency is responsible and not the company directing the worker day-to-day.
- **Kate's Law.** Increases criminal penalties for unauthorized immigrants who re-enter the United States after having been deported.
- **Promoting Cross Border Energy Infrastructure Act.** Allows oil and natural gas pipelines that cross into Canada or Mexico to be built without the President's permission.

- **Countering America's Adversaries Through Sanctions Act.** Places additional sanctions on Iran, Russia, and North Korea, as well as individuals who conduct business with these countries.

- **Department of Veterans Affairs Accountability and Whistleblower Protection Act of 2017.** Authorizes the Secretary of Veterans Affairs to demote, suspend, or fire senior Veterans Affairs employees for performance or misconduct, but forbids retaliation against whistleblowers.

- **Financial CHOICE Act of 2017.** Allows banks of sufficient size to take additional risk, and limits the power of the Consumer Financial Protection Bureau to investigate banks.

- **No Sanctuary for Criminals Act.** Prohibits giving federal grants to cities with "sanctuary" policies, policies cities enact to limit their cooperation with federal immigration law enforcement.

- **Ozone Standards Implementation Act of 2017.** Delays the implementation of a rule that would have reduced ozone pollution, allowing previous levels of pollution until 2026.

- **Tax Cuts and Jobs Act.** Reduces corporate taxes from 35% to 21% permanently. Temporarily reduces individual income taxes, with larger reductions for wealthier individuals. Increases the federal budget deficit by $1 trillion.

- **Sportsmen's Heritage and Recreational Enhancement (SHARE) Act of 2015.** Allows individuals to fish and hunt on federal lands without a license, unless the lands are closed for conservation, public safety, or national security.

- **Ozone Standards Implementation Act of 2017.** Delays the implementation of a rule that would have reduced ozone pollution, allowing previous levels of pollution until 2026.

- **No Taxpayer Funding for Abortion and Abortion Insurance Full Disclosure Act of 2017.** Prohibits the use of any federal funds for health insurance that provides abortion services.

- **Veterans 2nd Amendment Protection Act.** Allows any veteran deemed mentally incompetent to buy and own firearms and ammunition, unless a judge deems them dangerous.

- **Prohibits Use of Funds for Discrimination Based on Sexual Orientation or Gender Identity.** Prohibits the government from doing business with companies that discriminate against individuals based on sexual orientation or gender identity.

- **American Health Care Act of 2017.** Repeals "Obamacare": 1) Allows states to allow insurance companies to charge individuals more for insurance if they have a pre-existing condition. 2) Removes the requirement that Americans must carry health insurance. 3) Reduces amount given to low-income Americans to help them purchase health insurance.

B.2.3 Trust Game

Trust Game Instructions.

For our next study, we are going to ask you to play games with other survey respondents.

You will participate in several economic tasks called "games" over the next few minute.

You will be assigned to a different partner (someone else completing this survey) for each game.

You will receive some basic demographic information on each partner, but you will not find out who this person is, nor will he or she find out who you are (not now, nor after the survey is over).

You will work with money for each game. We will pay you an amount based on your final total.

(page break)

Instructions

This game is played by pairs of individuals. Each pair is made up of a Player 1 and a Player 2.

Game steps

We will give $10 to each Player 1. Player 1 then has the opportunity to give a portion of his or her $10 to Player 2. Player 1 could give some, all, or none of the $10. Whatever amount Player 1 decides to give to Player 2 will be tripled before it is passed on to Player 2. Player 2 then has the option of returning any portion of this tripled amount to Player 1. Each Player has 20 seconds to act.

Payment

Player 1 receives whatever he or she kept from their original $10, plus anything returned to him or her by Player 2. Player 2 receives whatever was given to him or her by Player 1, tripled, but then minus whatever they returned to Player 1. Note: We will multiply the final totals by 0.05 and give you a bonus for this survey of that amount. For example, if you win $20, we would pay you a bonus of $20 * 0.05 = $1. Please pay careful attention to these instructions. We will ask practice questions to ensure you understand.

(page break)

We will now run through 3 examples to show you how the game might be played.

Example 1 As always, Player 1 starts with $10. Imagine that Player 1 then gives $4 to Player 2. We triple this amount, so Player 2 gets $12 (3 times $4 equals $12). At this point, Player 1 has $6 and Player 2 has $12. Then Player 2 has to decide whether to give anything back to Player 1, and if so, how much. Suppose Player 2 decides to return $3 to Player 1. At the end of the game Player 1 will have $9 and Player 2 will have $9.

Example 2 Imagine that Player 1 gives all $10 to Player 2. We triple this amount, so Player 2 gets $30 (3 times $10 equals $30). At this point, Player 1 has $0 and Player 2 has $30. Then Player 2 has to decide whether to give anything back to Player 1, and if so, how much. Suppose Player 2 decides to return $15 to Player 1. At the end of the game Player 1 will have $15 and Player 2 will

have $15.

Example 3 Imagine that Player 1 gives $3 to Player 2. We triple this amount, so Player 2 gets $9 (3 times $3 equals $9). At this point, Player 1 has $7 and Player 2 has $9. Then Player 2 has to decide whether to give anything back to Player 1, and if so, how much. Suppose Player 2 decides to return $0 to Player 1. At the end of the game Player 1 will have $7 and Player 2 will have $9.

(page break)

Practice questions

Question 1: Player 1 starts with $10. Suppose that Player 1 gives $7 to Player 2. How much money will Player 2 get?

- $7
- $14
- $21

Question 2: After getting the money, what can Player 2 do with the money?

- Keep all the money
- Give some of the money to Player 1
- Give all of the money to Player 1
- All of the above

If correct: Correct!

If incorrect: Please try again. You must answer all questions correctly before you can continue. Player 1 starts with $10. Suppose that Player 1 gives $2 to Player 2. How much money will Player 2 get?

- $0

- $2
- $6

Question 2: After getting the money, what can Player 2 do with the money?

- Keep all the money
- Give some of the money to Player 1
- Give all of the money to Player 1
- All of the above

Trust Game.

You have been randomly assigned to play as Player 2. You will play as Player 2 for three rounds.

Each Player 1 will see the following information about you. They will use it to decide how trustworthy you are.

Age: [respondent's own age] Gender: [respondent's own gender] Income: [respondent's own income] Partisanship: [respondent's own party]

Round 1/2/3. (This is repeated for 3 rounds.)

You are Player 2.

Player 1 can give you some, all, or none of the $10. We will triple any amount Player 1 allocates to you. You are under no obligation to give anything back.

Information about who you are playing with (Player 1):

Age: [randomized age] Gender: [randomized gender] Income: [randomized income group] Partisanship: [respondent's out party]

Please wait while Player 1 decides your allocation.

(timer from 20 seconds counts downs, advances after about 5 seconds)

(page break)

If in Negative Experience group: *Results.* Player 1 allocated you $0. We are unable to triple this amount.

If in Positive Experience group:

Results. Player 1 allocated you $8. We have tripled this to $24. You can now return some, all or none of this money to Player 1. Put the number of dollars you wish to keep in the box labeled "Player 2." Put the dollars you wish to go to Player 1 in the box labeled "Player 1."

You gave $[amount given back] back to Player 1. Which factors, if any, were part of your decision making process?

- Age
- Gender
- Income
- Partisanship
- Something else

Player 1's reason for their allocation to you: your partisanship (all rounds), your income (Round 2)

Game Summary.

Round 1: Earnings: $[amount] Factors about you that Player 1 used when deciding how much money to give: Political Party.

Round 2: Earnings $[amount] Factors about you that Player 1 used when deciding how much money to give: Political Party, Income.

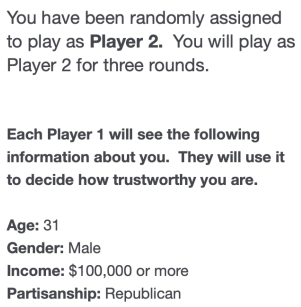
Round 3: Earnings $[amount] Factors about you that Player 1 used when deciding how much money to give: Political Party.

Your total earnings: $[amount]

Trust Game Screenshots

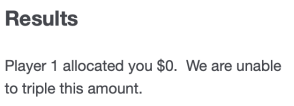
Figure B1: Steps in Trust Game: Screenshots

(a) Step 1



(b) Step 2

(c) Step 3



(d) Step 4

B.2.4 Manipulation Checks

Would you say that you were treated fairly or unfairly when playing the game?

- Very fairly
- Fairly
- Unfairly
- Very unfairly

We'd like you to rate how you feel towards some groups on a scale of 0 to 100. Zero means very unfavorable and 100 means very favorable. Fifty means you do not feel favorable or unfavorable. How would you rate your feeling toward each of the following?

- People who are [out party]s
- People who are [in party]s
- *Survey 4 only.* [out party] Politicians and Elected Officials
- *Survey 4 only.* [in party] Politicians and Elected Officials

- *Surveys 1-3 only.* People who are White
- *Surveys 1-3 only.* People who are Black
- *Surveys 1-3 only.* People who are Poor
- People who are Young
- People who are Old

How important is being a [in party] to how you see yourself as a person?

- Extremely important
- Very important
- Somewhat important
- Not very important
- Not at all important

B.2.5 Social Distance Items

How comfortable would you be to talk about politics with people who are [out party]s?

- Not at all comfortable
- Not too comfortable
- Somewhat comfortable
- Extremely comfortable

How comfortable are you having close personal friends who are [out party]s?

- Not at all comfortable

- Not too comfortable
- Somewhat comfortable
- Extremely comfortable

How comfortable are you having neighbors on your street who are [out party]s?

- Not at all comfortable
- Not too comfortable
- Somewhat comfortable
- Extremely comfortable

Suppose a son or daughter of yours was getting married. How would you feel if he or she married a supporter of the [out party]s?

- Not at all upset
- Not too upset
- Somewhat upset
- Extremely upset

B.2.6 Shown MC Votes and Asked MC DVs

Welcome to Study 3 of 3 in this survey. In this part, we will show you information about your Representative in Congress and ask for your reaction to it.

Your Representative in Congress is [Representative Name and Party].

From the bills in Congress we asked you about before, we have chosen three at random. We want to show you how [MC Name] actually voted on those bills.

- [*Bill title 1*]. Bill title description: [MC Name] Voted [Yes/No].
- [*Bill title 2*]. Bill title description: [MC Name] Voted [Yes/No].
- [*Bill title 3*]. Bill title description: [MC Name] Voted [Yes/No].

If the 2020 Congressional election were held today, who would you vote for?

- [MC Name and Party]
- The [Opposite party of MC] that runs against them
- For a third party
- Other
- I would not vote

If Third party, Other, or Would not vote is selected: If you had to choose in the 2020 election between [MC Name] and the [Opposite party of MC] who runs against them, who would you lean towards?

- [MC Name]
- The [Opposite party of MC] that runs against them
- Completely undecided

Do you approve or disapprove of the way [MC Name and Party] is handling their job as your representative in Congress?

- Strongly approve
- Approve
- Somewhat approve

- Neither approve nor disapprove
- Somewhat disapprove
- Disapprove
- Strongly disapprove

How favorable is your overall opinion of [MC Name and Party]?

- Extremely favorable
- Moderately favorable
- Slightly favorable
- Neither favorable nor unfavorable
- Slightly unfavorable
- Moderately unfavorable
- Extremely unfavorable

B.2.7 Social Distance Index

How comfortable would you be to talk about politics with people who are [outparty]s?

- Not at all comfortable
- Not too comfortable
- Somewhat comfortable
- Extremely comfortable

How comfortable are you having close personal friends who are [outparty]s?

- Not at all comfortable
- Not too comfortable
- Somewhat comfortable
- Extremely comfortable

How comfortable are you having neighbors on your street who are [outparty]s?

- Not at all comfortable
- Not too comfortable
- Somewhat comfortable
- Extremely comfortable

Suppose a son or daughter of yours was getting married. How would you feel if he or she married a supporter of the [outparty]s?

- Not at all upset
- Not too upset
- Somewhat upset
- Extremely upset

B.2.8 Norms Index

Do you agree or disagree with the statements below? (Options: Strongly agree, Agree, Somewhat agree, Neither agree nor disagree, Somewhat disagree, Disagree, Strongly disagree)

- When possible, [in party] politicians should try to compromise with [out party] politicians to get things done.

- *Study 1 only.* When determining the boundaries for Congressional districts, [in party] politicians should try to maximize the number of [in party]s elected.
- If a [in party] governor of a state can't get cooperation from [out party] legislators to pass new laws, the [in party] governor should issue executive orders on their own to accomplish their priorities.
- [in party] elected officials should sometimes consider ignoring court decisions when the judges are appointed by [out party] presidents.
- If a journalist accuses a [in party] politician of misconduct without naming their sources, the journalist should be criminally investigated.
- [in party]s should reduce the number of polling stations in areas that typically supports [out party]s.
- *Study 1 only.* Far-[left (if respondent is Republican) / right (if respondent is Democrat)] rallies should be banned.
- *Study 4 only, not included in norms scale.* Mail-in voting should be banned in areas that typically support [out party]s.

B.2.9 Own Views on Issues Post-Treatment

If you were in Congress would you vote FOR or AGAINST each of the following? (Options: For, Against)

- [*Bill title 1*]. Bill title description.
- [*Bill title 2*]. Bill title description.
- [*Bill title 3*]. Bill title description.

B.2.10 Vignette: Suppression

Please read the news story below.

Police Use Tear Gas on Peaceful Young [*Party*] *Protest*

Riot police broke up a group of Young [Party]s protesting a tuition increase at a college north of New York yesterday, using what appeared to be tear gas.

Students had been told earlier that ”all necessary force” would be used to ensure the protests did not interfere with the administration, in line with a recent court injunction requested and received by 53 of the schools students.

Witnesses reported that 30 police officers deployed to confront the 80 peaceful protesters. After some jostling and use of what appeared to be tear gas by police, the crowd blocking the doors broke and moved away.

Alan Krenshaw, a spokesperson for the Young [Party]s, complained that “The brutality was absolutely designed to chill the movement and literally try to beat and terrorize our right to criticize, to think critically and to act on that criticism.”

Do you agree or disagree with the decision to use tear gas on the [Party] protesters?

- Strongly agree
- Agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Disagree

- Strongly disagree

The cost of the police response is unknown at this point, but the city can fine the protests any amount up to $10,000. What amount, if any, do you think the city should fine the group of protesters? Enter a number between 0 and 10,000 below. (Do not include a $ sign.)

B.2.11 Corruption Tolerance

Please read the news story below.

> *Donations from Millionaire Businessman to* [*Party*] *Super PACs in Question*
>
> Federal election officials announced this morning that 6 million dollars in donations from Alan Gregory, a retired millionaire, to [Party] Super PACs are now under investigation.
>
> Officials allege that the donations were made through a questionable middle-man over a period of two months. The case is particularly important, as support from Mr. Gregory is credited in the victory of several [Party] state officials in the last election.
>
> Mr. Gregory and [Party] party officials deny the allegations and claim that although the donations were not transparent, they were not illegal.

Do you support the investigation of the businessman?

- Strongly support
- Support
- Somewhat support
- Neither support or oppose
- Somewhat oppose

- Oppose
- Strongly oppose

B.2.12 Antilocution Vignette

Please read the news article below.

> *OPINION: [Party]s Drive Congress to Do Less Than Last Year's Record-Breaking Low*
>
> Led by the [Party]'s refusal to compromise, the current U.S. Congress faces a backlog of unfinished business and sliding approval ratings.
>
> [Party]s must recognize that the country is not behind their obstructionist tactics. We need to move forward without letting crazy members of the opposition hold our nation hostage.
>
> Out of touch [Party]s are causing the gridlock by being unwilling and unable to engage in the normal process of negotiation and compromise. Major pieces of legislation, including a budget agreement and a farm and food-aid policy bill, are sitting on the sidelines.

A large website that posts stories from many different news sources is considering sharing the article you just read. Do you think they should post this article?

- Definitely
- Probably
- Maybe
- Probably not
- Definitely not

B.2.13 Election Override Vignette

Please read the news story below.

> *Local* [*Party*] *Candidate Calls for State Legislature to Decide Election*
>
> After the State Board of Elections announced that [Party] candidate Valerie Johnston lost by a few hundred votes, Johnson is petitioning the [Party]-controlled state legislature to take matters into its own hands and decide the election in her favor. According to the State Board of Elections, Johnston, the [Party] candidate, received 217 fewer votes than the [Other Party] candidate, Stephanie Cole. The result leaves the two candidates less than 0.1% apart after a bitter campaign.
>
> Talking to reporters, Johnston noted "that she had heard many claims that voting machines weren't working correctly in [Party] neighborhoods and that absentee ballots were not correctly handled." State law gives the legislature final say over certifying election outcomes, a provision meant to be used in cases of fraud or widespread errors. Johnson claims the law was intended to allow the state legislature to intervene in cases exactly like hers.
>
> "Whatever process that the state legislature chooses to go down, it's not going to change the overall number of votes," Cole said at a press conference. "In a democracy, we cannot allow a candidate that loses an election to ask her [Party] allies in the state legislature to just overturn election results. In a democracy, we must respect the will of the people."
>
> A spokesperson for the State Board of Elections told reporters that Johnston's complaint was "without basis." Meanwhile, the state legislature is debating whether to vote on the issue and determine the outcome of the election itself.

Do you support or oppose the [Party] state legislature determining the outcome of the election?

- Strongly support
- Support
- Somewhat support
- Neither support or oppose
- Somewhat oppose
- Oppose
- Strongly oppose

B.2.14 Bipartisanship Vignette

Note: This vignette is from Harbridge and Malhotra (2011, Study 2).

Republicans, Bipartisan Condition. Some members of Congress work with members of their own party almost all of the time. Other members work with members of both parties. In a recent year, on key issues identified by the American Conservative Union (ACU), Representative Steve LaTourette (a Republican from Ohio) took the Republican position on about half the votes and the Democratic position on about half the votes.

Republicans, Partisan Condition. Some members of Congress work with members of their own party almost all of the time. Other members work with members of both parties. In a recent year, on key issues identified by the American Conservative Union (ACU), Representative Steve LaTourette (a Republican from Ohio) almost always voted the Republican position.

Democrats, Bipartisan Condition. Some members of Congress work with members of their own party almost all of the time. Other members work with members of both parties. In a recent year, on key issues identified by the Americans for Democratic Action (ADA), Representative Bud Cramer (a Democrat from Alabama) almost always voted the Democratic position.

Democrats, Partisan Condition. Some members of Congress work with members of their own party almost all of the time. Other members work with members of both parties. In a recent year, on key issues identified by the Americans for Democratic Action (ADA), Representative Bud Cramer (a Democrat from Alabama) took the Democratic position on about half the votes and the Republican position on about half the votes.

Based on this information, do you approve or disapprove of the job [Representative LaTourette (R) / Representative Cramer (D)] is doing in Congress?

- Strongly approve
- Somewhat approve
- Neither approve nor disapprove
- Somewhat disagree
- Strongly disagree

B.2.15 Violence Justification Index

How much would it be justified for [in party]s. . . (Options: Not at all, A little, A moderate amount, A lot, A great deal)

- . . . to use violence to advance their political goals these days?
- . . . to use violence if they lose the 2024 election?

When (if ever) do you think it is ok for an ordinary person who is a [in party] . . . (Options: Never, Rarely, Sometimes, Often)

- . . . to harass an ordinary [out party] on the Internet in a way that makes the [out party] person feel unsafe?
- . . . to send physical threats and intimidating messages to [out party] leaders?

B.2.16 Perception of Objective Conditions

Donald Trump has been President of the United States for four years.

What is your best guess about the current unemployment rate in the United States? This is the percent of people who want to work who don't have a job.

- Over 10% unemployed
- 9-10% unemployed
- 8-9% unemployed
- 7-8% unemployed
- 6-7% unemployed
- 5-6% unemployed
- 4-5% unemployed
- Under 4% unemployed

Donald Trump has been President of the United States for four years.

What is your best guess about the number of Americans who have died from COVID-19?

- Fewer than 100,000
- 100,000 - 200,000
- 200,001 - 300,000
- 300,001 - 400,000
- 400,001 - 500,000
- More than 500,000

B.2.17 Final Demographics

What is the highest level of education you have completed?

- Less than High School
- High School / GED
- Some College
- 2-year College Degree
- 4-year College Degree
- Masters Degree
- Doctoral Degree
- Professional Degree (JD, MD)

Q21.2 When it comes to politics do you usually think of yourself as...

- Very conservative
- Conservative
- Moderate
- Liberal
- Very liberal
- Don’t know/ None of the above

B.3 Survey Demographics and Representativeness

Table B2: Comparing survey demographics with the 2019 CCES

	Survey 1	Survey 2	Survey 3	Survey 4	2019 CCES
Party ID - Strong Democrat	23.5%	22.6%	29.7%	30.0%	30.6%
Party ID - Not very strong Democrat	14.3%	17.2%	17.1%	17.3%	12.9%
Party ID - Lean Democrat	9.6%	12.6%	6.2%	11.5%	11.4%
Party ID - Independent*	0%*	0%*	0%*	0%*	0%*
Party ID - Lean Republican	11.2%	10.9%	6.6%	8.8%	10.6%
Party ID - Not very strong Republican	16.0%	16.2%	16.1%	13.5%	10.6%
Party ID - Strong Republican	25.6%	20.6%	24.3%	18.8%	23.9%
Ideology - Very liberal	8.2%	10.5%	10.8%	11.3%	15.5%
Ideology - Liberal	16.2%	17.8%	19.8%	20.4%	18.8%
Ideology - Moderate	34.9%	35.8%	31.6%	37.3%	28.7%
Ideology - Conservative	27.5%	23.5%	24.8%	20.0%	20.4%
Ideology - Very conservative	13.2%	12.5%	13.0%	10.9%	16.6%
Age	61.2	45.7	52.7	51.1	49.8
% Female	59.4%	52.1%	50.2%	54.8%	52.0%
% White	90.1%	82.0%	85.1%	71.6%	69.7%
% Hispanic or Latinx	1.7%	6.4%	4.0%	12.2%	11.6%
% Black	4.8%	7.0%	7.8%	14.6%	12.2%
% with income less than $50,000	33.9%	43.5%	31.4%	44.0%	41.9%
% with income greater than $100,000	23.4%	18.2%	28.2%	21.6%	16.9%
% with a 4-year college degree	52.1%	41.5%	57.4%	42.9%	32.7%

Note: Benchmark data comes from the 2019 Cooperative Congressional Election Study. The CCES's common weights were applied when calculating the CCES averages.
** Recall that we do not include pure independents in the surveys. To match our sampling frame, we also drop respondents to the CCES who stated they were Independents after being asked a lean question or were not sure of their party identification.*

C Pre-Analysis Plans

Pre-Analysis Plan: Affective Polarization and Democratic Accountability

October 28, 2019

Contents

1 Preliminary Notes

- This is the pre-analysis plan for a survey experiment on how changes in affective polarization affect citizens' willingness to engage in issue accountability in Congress.
- This is a revised version of our PAP. We had ideas for additional hypotheses before we gathered our data, and so are filing this revised PAP in advance of gathering our data.
- We have gathered approximately 50 pilot responses to test the survey. We will **not** include this pilot data in the dataset for the full study.
- All of the code excerpted below is included in our upload to OSF along with our PAP. We excerpt it into the PAP to facilitate peer review.

2 Data Cleaning

We will clean the data for the survey as follows:

```
// mc votes
insheet using "../code/pvs-heroku-affective/processed_MCs_for_api.csv", ///
    clear comma
keep candidateid vote*
rename vote* mcvote*
rename candidateid mccandidateid
tempfile mcvotes
save `mcvotes'


// main data
import delimited using "survey_data.csv", ///
clear rowrange(4:) varnames(2) bindquotes(strict) delim(",")

// merge in MC votes
destring mccandidateid, replace
merge m:1 mccandidateid using `mcvotes', nogen keep(1 3)

// PID
gen pid7 = 7 if wouldyousaythatyouarea == "Strong Republican"
replace pid7 = 6 if wouldyousaythatyouarea == "Not a strong Republican"
replace pid7 = 5 if strpos(pid_closerbranch, "Republican") != 0
replace pid7 = 4 if strpos(pid_closerbranch, "Neither") != 0
replace pid7 = 3 if strpos(pid_closerbranch, "Democrat") != 0
replace pid7 = 2 if strpos(v23, "Not a strong") != 0
replace pid7 = 1 if v23 == "Strong Democrat"
drop generallyspeakingdoyouusuallythi - pid_closerbranch
```

```
gen pid_strength = abs(pid7-4)


// What are the demographics of this sample?
tab whatisyourage
destring whatisyourage, replace force
rename whatisyourage age

tab whatisyourgender
gen female = whatisyourgender == "Female"
drop whatisyourgender

tab whatisyourcombinedannualhousehol
gen income = 1 if strpos(whatisyourcombinedannualhousehol, "Less than") != 0
replace income = 2 if strpos(whatisyourcombinedannualhousehol, "39,999") != 0
replace income = 3 if strpos(whatisyourcombinedannualhousehol, "49,999") != 0
replace income = 4 if strpos(whatisyourcombinedannualhousehol, "59,999") != 0
replace income = 5 if strpos(whatisyourcombinedannualhousehol, "69,999") != 0
replace income = 6 if strpos(whatisyourcombinedannualhousehol, "79,999") != 0
replace income = 7 if strpos(whatisyourcombinedannualhousehol, "89,999") != 0
replace income = 8 if strpos(whatisyourcombinedannualhousehol, "99,999") != 0
replace income = 9 if strpos(whatisyourcombinedannualhousehol, "100,000") != 0
drop whatisyourcombinedannualhousehol

tab whatisthehighestlevelofeducation
gen education = 1 if strpos(whatisthehighestlevelofeducation, "Less than") != 0
replace education = 2 if strpos(whatisthehighestlevelofeducation, "GED") != 0
replace education = 3 if strpos(whatisthehighestlevelofeducation, "Some Col") != 0
replace education = 4 if strpos(whatisthehighestlevelofeducation, "2-year College")
replace education = 5 if strpos(whatisthehighestlevelofeducation, "4-year College")
replace education = 6 if strpos(whatisthehighestlevelofeducation, "Masters") != 0 |
strpos(whatisthehighestlevelofeducation, "Doctoral") != 0 | ///
strpos(whatisthehighestlevelofeducation, "MD") != 0
drop whatisthehighestlevelofeducation

gen race_white = strpos(whatisyourraceandethnicityselect, "White") != 0
gen race_latinx = strpos(whatisyourraceandethnicityselect, "Latino") != 0
gen race_black = strpos(whatisyourraceandethnicityselect, "African") != 0
gen race_asian_or_other = strpos(whatisyourraceandethnicityselect, "Asian") != 0 |
strpos(whatisyourraceandethnicityselect, "Native") != 0 | ///
strpos(whatisyourraceandethnicityselect, "Pacific") != 0 | ///
strpos(whatisyourraceandethnicityselect, "Other") != 0
tabstat race_*
drop whatisyourraceandethnicityselect
```

```
// Own views on issues
drop votes1_dmc_rvoterworkingfamilies - votes3_rmc_dvoteramericanhealthc
destring ownview59180 - ownview59189, replace force

// Ideology
gen ideology = 5 if whenitcomestopoliticsdoyouusuall == "Very conservative"
replace ideology = 4 if whenitcomestopoliticsdoyouusuall == "Conservative"
replace ideology = 3 if whenitcomestopoliticsdoyouusuall == "Moderate"
replace ideology = 2 if whenitcomestopoliticsdoyouusuall == "Liberal"
replace ideology = 1 if whenitcomestopoliticsdoyouusuall == "Very liberal"
drop whenitcomestopoliticsdoyouusuall

// DVs
gen manip_check_fair = 4 if wouldyousaythatyouweretreatedfai == "Very fairly"
replace manip_check_fair = 3 if wouldyousaythatyouweretreatedfai == "Fairly"
replace manip_check_fair = 2 if wouldyousaythatyouweretreatedfai == "Unfairly"
replace manip_check_fair = 1 if wouldyousaythatyouweretreatedfai == "Very unfairly'
drop wouldyousaythatyouweretreatedfai

rename wedlikeyoutoratehowyoufeeltoward therm_outparty
rename v128 therm_inparty
drop v129 - v133 // therms for white, black, poor, young, old
destring therm_*, replace
gen affpol = therm_inparty - therm_outparty

gen pid_identity_importance = 5 if strpos(howimportantisbeingafieldinparty, "Extrem
replace pid_identity_importance = 4 if strpos(howimportantisbeingafieldinparty, "Ve
replace pid_identity_importance = 3 if strpos(howimportantisbeingafieldinparty, "So
replace pid_identity_importance = 2 if strpos(howimportantisbeingafieldinparty, "No
replace pid_identity_importance = 1 if strpos(howimportantisbeingafieldinparty, "No
tab pid_identity_importance
drop howimportantisbeingafieldinparty

gen mc_vote_choice = 5 if strpos(generic, "mclongname") != 0
replace mc_vote_choice = 4 if strpos(genericlean, "mcshortname") != 0
replace mc_vote_choice = 3 if strpos(genericlean, "Complet") != 0
replace mc_vote_choice = 2 if strpos(genericlean, "mcotherparty") != 0
replace mc_vote_choice = 1 if strpos(generic, "mcotherparty") != 0
tab mc_vote_choice, m
drop generic genericlean // this was messed up in pilot but is fixed

rename mcapprove mc_approve
replace mc_approve = "7" if mc_approve == "Strongly approve"
replace mc_approve = "6" if mc_approve == "Approve"
```

```
replace mc_approve = "5" if mc_approve == "Somewhat approve"
replace mc_approve = "4" if mc_approve == "Neither approve nor disapprove"
replace mc_approve = "3" if mc_approve == "Somewhat disapprove"
replace mc_approve = "2" if mc_approve == "Disapprove"
replace mc_approve = "1" if mc_approve == "Strongly disapprove"

rename mcfavorability mc_favorability
replace mc_favorability = "7" if mc_favorability == "Extremely favorable"
replace mc_favorability = "6" if mc_favorability == "Moderately favorable"
replace mc_favorability = "5" if mc_favorability == "Slightly favorable"
replace mc_favorability = "4" if mc_favorability == "Neither favorable nor unfavora
replace mc_favorability = "3" if mc_favorability == "Slightly unfavorable"
replace mc_favorability = "2" if mc_favorability == "Moderately unfavorable"
replace mc_favorability = "1" if mc_favorability == "Extremely unfavorable"
destring mc_approve mc_favorability, replace


// Covariate for analysis, MC PID match
gen voter_mc_same_party = substr(inparty, 1, 1) == mcpartyoneletter
drop inparty
gen mc_pid_match = pid7 if mcpartyoneletter == "R"
replace mc_pid_match = 8 - pid7 if mcpartyoneletter == "D"
bysort mcpartyoneletter: reg mc_approve_scale mc_pid_match, robust // sanity check
reg mc_approve_scale voter_mc_same_party, robust

// Norms battery
local normsvars norms1whenpossiblefieldinpartyic - norms2farfieldfaroutralliesshoul
foreach var of varlist `normsvars' action_suppression {
replace `var' = "7" if `var' == "Strongly agree"
replace `var' = "6" if `var' == "Agree"
replace `var' = "5" if `var' == "Somewhat agree"
replace `var' = "4" if `var' == "Neither agree nor disagree"
replace `var' = "3" if `var' == "Somewhat disagree"
replace `var' = "2" if `var' == "Disagree"
replace `var' = "1" if `var' == "Strongly disagree"
}
destring `normsvars', replace
alpha `normsvars', item std gen(norms_scale)

// Norms vignettes
rename (action_suppression action_suppression2 corruption_tolerance antilocution) n

destring normsv_action*, replace

replace normsv_corruption_tolerance = "7" if normsv_corruption_tolerance == "Strong
```

```
replace normsv_corruption_tolerance = "6" if normsv_corruption_tolerance == "Suppor
replace normsv_corruption_tolerance = "5" if normsv_corruption_tolerance == "Somewh
replace normsv_corruption_tolerance = "4" if normsv_corruption_tolerance == "Neithe
replace normsv_corruption_tolerance = "2" if normsv_corruption_tolerance == "Somewh
replace normsv_corruption_tolerance = "2" if normsv_corruption_tolerance == "Oppose
replace normsv_corruption_tolerance = "1" if normsv_corruption_tolerance == "Strong
destring normsv_corruption_tolerance, replace

// note: changed this to a 5-pt scale ranging from "definitely" to "definitely not'
replace normsv_antilocution = "1" if normsv_antilocution == "Yes"
replace normsv_antilocution = "0" if normsv_antilocution == "No"
destring normsv_antilocution, replace

// Post-treatment issue agreement
gen post_bill0 = v158 + v160
gen post_bill1 = ifyouwereincongresswouldyouvotef + v159
gen post_bill2 = v157
drop ifyouwereincongresswouldyouvotef - v161
tab post_bill0
tab post_bill1
tab post_bill2
foreach var of varlist ownview* {
local billid = substr("`var'", 8, 5)
disp "`billid'"
gen posttreat_ownview`billid' = .
forvalues b=0/2 {
// was respondent for or against?
replace posttreat_ownview`billid' = 1 if strpos(post_bill`b', "FOR") != 0 & randomv
replace posttreat_ownview`billid' = 0 if strpos(post_bill`b', "AGAINST") != 0 & ran
}

// pre-treat view on all issues
gen pretreat_agreewmc`billid' = ownview`billid' == mcvote`billid' if ///
!missing(ownview`billid') & !missing(mcvote`billid')

// pre-treat view on just issues also asked about post-treatment
gen pretrt_askedl8r_agreewmc`billid' = pretreat_agreewmc`billid' if ///
randomvote_id_0 == "`billid'" | randomvote_id_1 == "`billid'" | randomvote_id_2 ==

// post-treat view on issues (only up to 3 issues asked about)
gen posttreat_agreewmc`billid' = posttreat_ownview`billid' == mcvote`billid' if //
!missing(posttreat_ownview`billid') & !missing(mcvote`billid')
}
drop post_bill0 - post_bill1
```

```
egen pretreat_agreewmc_share = rowmean(pretreat_agreewmc*)
egen pretrt_justasked_agreewmc_share = rowmean(pretrt_askedl8r_agreewmc*)
egen posttreat_agreewmc_share = rowmean(posttreat_agreewmc*)


// Code treatments
replace vote_type_treatment = "1 Positive Votes" if vote_type_treatment == "agree"
replace vote_type_treatment = "2 Negative Votes" if vote_type_treatment == "disagre

replace trust_treatment = "1 Positive Game Experience" if trust_treatment == "posit
replace trust_treatment = "2 Negative Game Experience" if trust_treatment == "negat

encode trust_treatment, gen(treat_trust_game)
encode vote_type_treatment, gen(treat_vote_type)

drop end_comments apisuccess2

compress
save pilot_data.dta, replace
```

Note: We do not expect missing data because our Qualtics survey is set to "force response", but if there is missing data we will recode all missing data to its mean.

3 Dependent Variables

3.1 Primary DV: MC Approval Scale

Our primary dependent variable of interest will be an additive index of the standardized MC Approval, MC Favorability, and MC Vote Choice variables, as calculated with the code below. We will then standardize this to standard deviation one, as calculated with the code below.

```
alpha mc_approve mc_favorability mc_vote_choice, ///
        item std gen(mc_approve_scale)
sum mc_approve_scale
replace mc_approve_scale = mc_approve_scale / r(sd)
sum mc_approve_scale
```

3.2 Primary DV: Norms

We will use the norms_scale variable computed above to assess the effects on democratic norms.

3.3 Secondary DV: Vignettes

We will also ask three vignettes from Lelkes and Westwood (2017), with DVs captured in the variables `normsv_action_suppression`, `normsv_action_suppression2`, `normsv_corruption_tolerance`, and `normsv_antilocution`.

3.4 Secondary DV: Following the Leader

We also measured issue opinions on the issues where individuals were shown their MC's votes, and will test whether there is evidence voters "follow the leader" (Lenz 2012) more when they are exposed to the negative treatment that should increase affective polarization.

3.5 Exploratory DVs

We will also use the following variables as exploratory dependent variables:

- `mc_vote_choice`, vote choice in the next election (a component of the MC approval scale)
- `pid_identity_importance`, as this is another possible mechanism
- `mc_serious_threat`, whether the respondent says the MC is a serious threat to the United States.[1]

3.6 Manipulation Checks

We see the following as representing manipulation checks:

- `affpol`, to measure whether affective polarization increased as a result of the game treatment
- `manip_check_fair`, to measure whether people thought the game was fair

4 Analysis

4.1 Covariate

To increase precision, we will use mc_pid_match, pid_strength, and pretreat_agreewmc_share as covariates to increase precision in some analyses described below.

[1]This will be coded as a binary variable.

4.2 Treatments

The design is a 2x2 design:

1. There are two "MC Vote Treatments": showing positive votes, i.e., votes the MC cast that the respondent agreed with; and showing negative votes, i.e., votes the MC cast that the respondent disagreed with.

2. There are two "Trust Game Treatments": a positive game experience treatment, where players from the other party cooperated; and a negative game experience treatment, where players from the other party defected.

We will code the treatments as follows, using the "positive" treatments as the base category in both cases.

```
// Code treatments
replace vote_type_treatment = "1 Positive Votes" if
    vote_type_treatment == "agree"
replace vote_type_treatment = "2 Negative Votes" if
    vote_type_treatment == "disagree"

replace trust_treatment = "1 Positive Game Experience" if
    trust_treatment == "positive"
replace trust_treatment = "2 Negative Game Experience" if
    trust_treatment == "negative"

encode trust_treatment, gen(treat_trust_game)
encode vote_type_treatment, gen(treat_vote_type)
```

4.3 Hypothesis Tests: MC Approval Scale

We will test the effect on the MC approval DVs as follows, including the `mc_pid_match` covariate. We expect a positive interaction term between the two treatments, as we expect the negative effect of revealing an MC cast incongruent votes to become less negative under conditions of higher affective polarization. The main p-value of interest is the p-value on this interaction term, with the p-value on the primary dependent variable being of greatest interest.

```
foreach dv in mc_approve_scale mc_vote_choice mc_serious_threat {
disp "DV = `dv'"

// Expect positive interaction term because the negative effect of seeing
// negative votes will get less negative in high affective polarization
// condition. This is main hypothesis test.
reg `dv' i.treat_trust_game##i.treat_vote_type mc_pid_match ///
    pretreat_agreewmc_share, robust
```

```
// To make the effects easier for readers to understand, we may also report:
bysort treat_trust_game: reg `dv' i.treat_vote_type mc_pid_match ///
    pretreat_agreewmc_share, robust
}
```

We will also conduct this exploratory test:

```
// Exploratory: does showing negative votes increase affective
// polarization itself?
bysort voter_mc_same_party: reg affpol i.treat_vote_type mc_pid_match, robus
```

4.4 Hypothesis Tests: Norms and Vignettes

We will conduct the following analyses on variables capturing adherence to democratic norms and for the prejudice vignettes.

```
// Analysis 2: Norms
// Here, we are interested in the main effect of affpol. Do not expect
// any main effect of the MC votes or any interaction effects, since these
// questions are not about the MC.
foreach dv in norms_scale normsv_action_suppression normsv_action_suppressio
normsv_corruption_tolerance normsv_antilocution {
disp "DV = `dv'"

reg `dv' i.treat_trust_game pid_strength, robust
}

// We will also test the effects on the individual norms items in the norms
// scale this way, but don't expect these to be very well-powered.
foreach dv of varlist norms1* {
disp "DV = `dv'"

reg `dv' i.treat_trust_game pid_strength, robust
}
```

4.5 Hypothesis Tests: Following the Leader

As described above, we will test whether there is evidence voters "follow the leader" (Lenz 2012) more when they are exposed to the negative treatment that should increase affective polarization.

```
// Analysis 3: More Follow-ing The Leader with high affpol?
bysort voter_mc_same_party: reg posttreat_agreewmc_share i.treat_trust_game
pretrt_justasked_agreewmc_share
```

4.6 Heterogenous Treatment Effects

We have no clear predictions for heterogeneous treatment effects. However, we will explore whether the interaction term estimate differs depending on whether the MC is the same party as the voter or in the other party.

```
foreach dv in mc_approve_scale mc_vote_choice mc_serious_threat {
    disp "DV = `dv'"

    // Exploratory HTEs:
bysort voter_mc_same_party: reg `dv' i.treat_trust_game##i.treat_vote_type
    mc_pid_match pretreat_agreewmc_share, robust
bysort respondent_party: reg `dv' i.treat_trust_game##i.treat_vote_type ///
    mc_pid_match pretreat_agreewmc_share, robust
}
```

4.7 Manipulation Checks

We will conduct the following manipulation checks:

```
reg manip_check_fair i.treat_trust_game, robust
reg affpol i.treat_trust_game, robust

// Potential other mechanism
reg pid_identity_importance i.treat_trust_game pid_strength, robust
```

References

Lelkes, Yphtach and Sean J Westwood. 2017. "The limits of partisan prejudice." *The Journal of Politics* 79(2):485–501.

Lenz, Gabriel S. 2012. *Follow the Leader? How Voters Respond to Politicians' Performance and Policies*. Chicago: University of Chicago Press.

Pre-Analysis Plan: Affective Polarization and Democratic Accountability, Surveys 2 and 3

December 4, 2019

Contents

1 Preliminary Notes

- This is the pre-analysis plan for two survey experiments on how changes in affective polarization affect citizens' willingness to engage in issue accountability in Congress and norms.
- We already conducted one survey. We plan to conduct two more surveys with larger Ns for additional statistical power on several hypotheses. In addition, for "Survey 2", we are also repeating so that the norms questions occur closer to the trust game.
- All of the code excerpted below is included in our upload to OSF along with our PAP. We excerpt it into the PAP to facilitate peer review.

2 Survey 2 - Norms

2.1 Data Cleaning

We will clean the data as follows.

```
************** CLEANING ******************
// main data
import delimited using "test_data.csv", ///
clear rowrange(4:) varnames(2) bindquotes(strict) delim(",")

gen finished_survey = finished == "True"

// examine termination reasons
tab termreason, m
// apifail2 is if they don't agree on at least one thing and disagree on at least o
// political party is removal of independents
// practice2fail is people who fail both practice rounds of the game practice round
keep if termreason == ""

// drop unnecessary variables
drop startdate - userlanguage thissurveyincludesthreestudiesst
drop round1age - round3income
drop timingfirstclick - yougaveqid274choicenumericentryv // any reason we need thes
drop vigpartyic vigpartyother
drop psid pid end_comments termreason outparty inparty respondent_party inpartyic
drop farout outpartyic apisuccess2 pay bonus

// PID
gen pid7 = 7 if wouldyousaythatyouarea == "Strong Republican"
replace pid7 = 6 if wouldyousaythatyouarea == "Not a strong Republican"
replace pid7 = 5 if strpos(pid_closerbranch, "Republican") != 0
```

```
replace pid7 = 4 if strpos(pid_closerbranch, "Neither") != 0
replace pid7 = 3 if strpos(pid_closerbranch, "Democrat") != 0
replace pid7 = 2 if strpos(v21, "Not a strong") != 0
replace pid7 = 1 if v21 == "Strong Democrat"
drop generallyspeakingdoyouusuallythi - pid_closerbranch
gen pid_strength = abs(pid7-4)
recode pid7 (1/3=0) (5/7=1), gen(pid_republican)


// What are the demographics of this sample?
tab whatisyourage
destring whatisyourage, replace force
rename whatisyourage age

tab whatisyourgender
gen female = whatisyourgender == "Female"
drop whatisyourgender

tab whatisyourcombinedannualhousehol
gen income = 1 if strpos(whatisyourcombinedannualhousehol, "Less than") != 0
replace income = 2 if strpos(whatisyourcombinedannualhousehol, "39,999") != 0
replace income = 3 if strpos(whatisyourcombinedannualhousehol, "49,999") != 0
replace income = 4 if strpos(whatisyourcombinedannualhousehol, "59,999") != 0
replace income = 5 if strpos(whatisyourcombinedannualhousehol, "69,999") != 0
replace income = 6 if strpos(whatisyourcombinedannualhousehol, "79,999") != 0
replace income = 7 if strpos(whatisyourcombinedannualhousehol, "89,999") != 0
replace income = 8 if strpos(whatisyourcombinedannualhousehol, "99,999") != 0
replace income = 9 if strpos(whatisyourcombinedannualhousehol, "100,000") != 0
drop whatisyourcombinedannualhousehol

tab whatisthehighestlevelofeducation
gen education = 1 if strpos(whatisthehighestlevelofeducation, "Less than") != 0
replace education = 2 if strpos(whatisthehighestlevelofeducation, "GED") != 0
replace education = 3 if strpos(whatisthehighestlevelofeducation, "Some Col") != 0
replace education = 4 if strpos(whatisthehighestlevelofeducation, "2-year College")
replace education = 5 if strpos(whatisthehighestlevelofeducation, "4-year College")
replace education = 6 if strpos(whatisthehighestlevelofeducation, "Masters") != 0 |
strpos(whatisthehighestlevelofeducation, "Doctoral") != 0 | ///
strpos(whatisthehighestlevelofeducation, "MD") != 0
drop whatisthehighestlevelofeducation

gen race_white = strpos(whatisyourraceandethnicityselect, "White") != 0
gen race_latinx = strpos(whatisyourraceandethnicityselect, "Latino") != 0
gen race_black = strpos(whatisyourraceandethnicityselect, "African") != 0
gen race_asian_or_other = strpos(whatisyourraceandethnicityselect, "Asian") != 0 |
```

```
strpos(whatisyourraceandethnicityselect, "Native") != 0 | ///
strpos(whatisyourraceandethnicityselect, "Pacific") != 0 | ///
strpos(whatisyourraceandethnicityselect, "Other") != 0
tabstat race_*
drop whatisyourraceandethnicityselect


// Ideology
gen ideology = 5 if whenitcomestopoliticsdoyouusuall == "Very conservative"
replace ideology = 4 if whenitcomestopoliticsdoyouusuall == "Conservative"
replace ideology = 3 if whenitcomestopoliticsdoyouusuall == "Moderate"
replace ideology = 2 if whenitcomestopoliticsdoyouusuall == "Liberal"
replace ideology = 1 if whenitcomestopoliticsdoyouusuall == "Very liberal"
drop whenitcomestopoliticsdoyouusuall

// DVs
gen manip_check_fair = 4 if wouldyousaythatyouweretreatedfai == "Very fairly"
replace manip_check_fair = 3 if wouldyousaythatyouweretreatedfai == "Fairly"
replace manip_check_fair = 2 if wouldyousaythatyouweretreatedfai == "Unfairly"
replace manip_check_fair = 1 if wouldyousaythatyouweretreatedfai == "Very unfairly'
drop wouldyousaythatyouweretreatedfai

rename wedlikeyoutoratehowyoufeeltoward therm_outparty
rename v78 therm_inparty
rename v79 therm_young
rename v80 therm_old
destring therm_*, replace
gen affpol = therm_inparty - therm_outparty

// Norms battery
local normsvars normswhenpossiblefieldinpartyicp - normsfieldinpartysshouldreduceth
foreach var of varlist `normsvars' {
replace `var' = "7" if `var' == "Strongly agree"
replace `var' = "6" if `var' == "Agree"
replace `var' = "5" if `var' == "Somewhat agree"
replace `var' = "4" if `var' == "Neither agree nor disagree"
replace `var' = "3" if `var' == "Somewhat disagree"
replace `var' = "2" if `var' == "Disagree"
replace `var' = "1" if `var' == "Strongly disagree"
}
destring `normsvars', replace
alpha `normsvars', item gen(norms_scale) std
sum norms_scale
replace norms_scale = norms_scale / r(sd)
sum norms_scale
```

```
// Norms vignettes
rename v87 election_override
rename (corruption_tolerance election_override antilocution) normsv_=

foreach var of varlist normsv_corruption_tolerance normsv_election_override {
replace `var' = "7" if `var' == "Strongly support"
replace `var' = "6" if `var' == "Support"
replace `var' = "5" if `var' == "Somewhat support"
replace `var' = "4" if `var' == "Neither support or oppose"
replace `var' = "3" if `var' == "Somewhat oppose"
replace `var' = "2" if `var' == "Oppose"
replace `var' = "1" if `var' == "Strongly oppose"
destring `var', replace
replace `var' = 8 - `var' if vigpartytreat == "outparty" // prediction in opposite
}

tab normsv_antilocution
replace normsv_antilocution = "5" if normsv_antilocution == "Definitely"
replace normsv_antilocution = "4" if normsv_antilocution == "Probably"
replace normsv_antilocution = "3" if normsv_antilocution == "Maybe"
replace normsv_antilocution = "2" if normsv_antilocution == "Probably not"
replace normsv_antilocution = "1" if normsv_antilocution == "Definitely not"
destring normsv_antilocution, replace
replace normsv_antilocution = 6 - normsv_antilocution if vigpartytreat == "outparty

gen normsv_bipartisan_approve = .
foreach var of varlist malhar_* {
replace normsv_bipartisan = 5 if `var' == "Strongly approve"
replace normsv_bipartisan = 4 if `var' == "Somewhat approve"
replace normsv_bipartisan = 3 if `var' == "Neither approve nor disapprove"
replace normsv_bipartisan = 2 if `var' == "Somewhat disapprove"
replace normsv_bipartisan = 1 if `var' == "Strongly disapprove"
}
drop malhar_*

// Code treatments
replace trust_treatment = "1 Positive Game Experience" if trust_treatment == "posit
replace trust_treatment = "2 Negative Game Experience" if trust_treatment == "negat

encode trust_treatment, gen(treat_trust_game)
encode malhar, gen(treat_bipartisan)
encode vigpartytreat, gen(vigpartytreatenc)
```

Note: We do not expect missing data because our Qualtics survey is set to "force response", but if there is missing data we will recode all missing data to its mean.

2.2 Hypotheses

2.2.1 Norms Scale

```
// Our first main hypothesis is that the norms scale will be affected.
reg norms_scale i.treat_trust_game pid_strength pid7, robust

// We will also report individual items on norms scale. We don't expect thes
foreach dv of varlist normswhenpossiblefieldinpartyicp - normsfieldinpartys
disp "DV = `dv'"

reg `dv' i.treat_trust_game pid_strength pid7, robust

bysort pid_republican: reg `dv' i.treat_trust_game pid_strength pid7, robust
}
```

2.2.2 Vignettes

For the first three vignettes there are inparty and outparty versions. Our main hypotheses concern the effect of the trust game treatment pooled across the inparty and outparty versions, as well as the separate inparty/outparty effects.

```
// We will present results pooled and then separately by in/out party.
foreach dv in normsv_corruption_tolerance normsv_election_override normsv_ar
disp "DV = `dv'"

// Primary hypotheses:
reg `dv' i.treat_trust_game pid_strength pid7 i.vigpartytreatenc, robust
bysort vigpartytreat: reg `dv' i.treat_trust_game pid_strength pid7, robust

// Show conditional ATEs by respondent PID - no clear prediction but we know
bysort pid_republican: reg `dv' i.treat_trust_game pid_strength pid7 i.vigpa
bysort pid_republican vigpartytreat: reg `dv' i.treat_trust_game pid_strengt
}
```

There is also a fourth vignette about bipartisanship that we will analyze as follows.

```
// Preference for bipartisanship. Key hypothesis is interaction term.
reg normsv_bipartisan i.treat_trust_game##i.treat_bipartisan pid_strength pi
```

2.3 Manipulation Checks

We will conduct the following manipulation checks:

```
// Manip checks
reg manip_check_fair i.treat_trust_game, robust
reg affpol i.treat_trust_game pid_strength pid7, robust

reg pid_identity_importance i.treat_trust_game pid_strength, robust
```

3 Survey 3 - MC Accountability

We will clean the data as follows.

3.1 Data Cleaning

```
********** CLEANING ***********
// mc votes
insheet using "../../pvs-heroku-affective/processed_MCs_for_api.csv", clear comma
keep candidateid vote*
rename vote* mcvote*
rename candidateid mccandidateid
tempfile mcvotes
save `mcvotes'

// main data
import delimited using "test_data.csv", ///
clear rowrange(4:) varnames(2) bindquotes(strict) delim(",")

gen finished_survey = finished == "True"

// examine termination reasons
tab termreason, m
// apifail1 is if they are in a non-eligible zip code
// apifail2 is if they don't agree on at least one thing and disagree on at least o
// political party is removal of independents
// practice2fail is people who fail both practice rounds of the game practice round
keep if termreason == ""

// merge in MC votes
destring mccandidateid, replace
merge m:1 mccandidateid using `mcvotes', nogen keep(1 3)

// time looking at votes
destring v136, gen(mech_time_looking_at_votes)

// drop unnecessary variables
drop startdate - thissurveyincludesthreestudiesst
```

```
drop round1age - round3income
drop timingfirstclick - yougaveqid274choicenumericentryv
drop v134 - v136 v137 // more timing stuff
drop randomvote_vote_bill_title_0 - v207
drop psid pid mccandidateid mclongname mcshortname apisuccess1
drop outparty inpartyic farout outpartyic


// PID
gen pid7 = 7 if wouldyousaythatyouarea == "Strong Republican"
replace pid7 = 6 if wouldyousaythatyouarea == "Not a strong Republican"
replace pid7 = 5 if strpos(pid_closerbranch, "Republican") != 0
replace pid7 = 4 if strpos(pid_closerbranch, "Neither") != 0
replace pid7 = 3 if strpos(pid_closerbranch, "Democrat") != 0
replace pid7 = 2 if strpos(v23, "Not a strong") != 0
replace pid7 = 1 if v23 == "Strong Democrat"
drop generallyspeakingdoyouusuallythi - pid_closerbranch
gen pid_strength = abs(pid7-4)


// What are the demographics of this sample?
tab whatisyourage
destring whatisyourage, replace force
rename whatisyourage age

tab whatisyourgender
gen female = whatisyourgender == "Female"
drop whatisyourgender

tab whatisyourcombinedannualhousehol
gen income = 1 if strpos(whatisyourcombinedannualhousehol, "Less than") != 0
replace income = 2 if strpos(whatisyourcombinedannualhousehol, "39,999") != 0
replace income = 3 if strpos(whatisyourcombinedannualhousehol, "49,999") != 0
replace income = 4 if strpos(whatisyourcombinedannualhousehol, "59,999") != 0
replace income = 5 if strpos(whatisyourcombinedannualhousehol, "69,999") != 0
replace income = 6 if strpos(whatisyourcombinedannualhousehol, "79,999") != 0
replace income = 7 if strpos(whatisyourcombinedannualhousehol, "89,999") != 0
replace income = 8 if strpos(whatisyourcombinedannualhousehol, "99,999") != 0
replace income = 9 if strpos(whatisyourcombinedannualhousehol, "100,000") != 0
drop whatisyourcombinedannualhousehol

tab whatisthehighestlevelofeducation
gen education = 1 if strpos(whatisthehighestlevelofeducation, "Less than") != 0
replace education = 2 if strpos(whatisthehighestlevelofeducation, "GED") != 0
replace education = 3 if strpos(whatisthehighestlevelofeducation, "Some Col") != 0
```

```
replace education = 4 if strpos(whatisthehighestlevelofeducation, "2-year College")
replace education = 5 if strpos(whatisthehighestlevelofeducation, "4-year College")
replace education = 6 if strpos(whatisthehighestlevelofeducation, "Masters") != 0 |
strpos(whatisthehighestlevelofeducation, "Doctoral") != 0 | ///
strpos(whatisthehighestlevelofeducation, "MD") != 0
drop whatisthehighestlevelofeducation

gen race_white = strpos(whatisyourraceandethnicityselect, "White") != 0
gen race_latinx = strpos(whatisyourraceandethnicityselect, "Latino") != 0
gen race_black = strpos(whatisyourraceandethnicityselect, "African") != 0
gen race_asian_or_other = strpos(whatisyourraceandethnicityselect, "Asian") != 0 |
strpos(whatisyourraceandethnicityselect, "Native") != 0 | ///
strpos(whatisyourraceandethnicityselect, "Pacific") != 0 | ///
strpos(whatisyourraceandethnicityselect, "Other") != 0
tabstat race_*
drop whatisyourraceandethnicityselect


// Own views on issues
drop votes1_dmc_rvoterworkingfamilies - votes3_rmc_dvoteramericanhealthc
destring ownview59180 - ownview59189, replace force

// Ideology
gen ideology = 5 if whenitcomestopoliticsdoyouusuall == "Very conservative"
replace ideology = 4 if whenitcomestopoliticsdoyouusuall == "Conservative"
replace ideology = 3 if whenitcomestopoliticsdoyouusuall == "Moderate"
replace ideology = 2 if whenitcomestopoliticsdoyouusuall == "Liberal"
replace ideology = 1 if whenitcomestopoliticsdoyouusuall == "Very liberal"
drop whenitcomestopoliticsdoyouusuall

// DVs
gen manip_check_fair = 4 if wouldyousaythatyouweretreatedfai == "Very fairly"
replace manip_check_fair = 3 if wouldyousaythatyouweretreatedfai == "Fairly"
replace manip_check_fair = 2 if wouldyousaythatyouweretreatedfai == "Unfairly"
replace manip_check_fair = 1 if wouldyousaythatyouweretreatedfai == "Very unfairly'
drop wouldyousaythatyouweretreatedfai

rename wedlikeyoutoratehowyoufeeltoward therm_outparty
rename v128 therm_inparty
rename v129 therm_whites
rename v130 therm_blacks
rename v131 therm_poor
rename v132 therm_young
rename v133 therm_old
destring therm_*, replace
```

```
gen affpol = therm_inparty - therm_outparty

gen mc_vote_choice = 5 if strpos(generic, "mclongname") != 0
replace mc_vote_choice = 4 if strpos(genericlean, "mcshortname") != 0
replace mc_vote_choice = 3 if strpos(genericlean, "Complet") != 0
replace mc_vote_choice = 2 if strpos(genericlean, "mcotherparty") != 0
replace mc_vote_choice = 1 if strpos(generic, "mcotherparty") != 0
tab mc_vote_choice, m
drop generic genericlean

rename mcapprove mc_approve
replace mc_approve = "7" if mc_approve == "Strongly approve"
replace mc_approve = "6" if mc_approve == "Approve"
replace mc_approve = "5" if mc_approve == "Somewhat approve"
replace mc_approve = "4" if mc_approve == "Neither approve nor disapprove"
replace mc_approve = "3" if mc_approve == "Somewhat disapprove"
replace mc_approve = "2" if mc_approve == "Disapprove"
replace mc_approve = "1" if mc_approve == "Strongly disapprove"

rename mcfavorability mc_favorability
replace mc_favorability = "7" if mc_favorability == "Extremely favorable"
replace mc_favorability = "6" if mc_favorability == "Moderately favorable"
replace mc_favorability = "5" if mc_favorability == "Slightly favorable"
replace mc_favorability = "4" if mc_favorability == "Neither favorable nor unfavora
replace mc_favorability = "3" if mc_favorability == "Slightly unfavorable"
replace mc_favorability = "2" if mc_favorability == "Moderately unfavorable"
replace mc_favorability = "1" if mc_favorability == "Extremely unfavorable"
destring mc_approve mc_favorability, replace

gen mc_serious_threat = strpos(mcseriousthreat, "Yes, ") != 0
replace mc_serious_threat = . if mcseriousthreat == ""
drop mcseriousthreat

alpha mc_approve mc_favorability mc_vote_choice mc_serious_threat, item std gen(mc_
sum mc_approve_scale
replace mc_approve_scale = mc_approve_scale / r(sd)
sum mc_approve_scale


// Mechanisms
foreach var of varlist othermcdvsfieldmcshortnamevotest - anxious {
replace `var' = "7" if `var' == "Strongly agree"
replace `var' = "6" if `var' == "Agree"
replace `var' = "5" if `var' == "Somewhat agree"
replace `var' = "4" if `var' == "Neither agree nor disagree"
```

```
replace `var' = "3" if `var' == "Somewhat disagree"
replace `var' = "2" if `var' == "Disagree"
replace `var' = "1" if `var' == "Strongly disagree"
destring `var', replace
}
rename othermcdvs* mech_*
rename anxious mech_anxious
rename mech_fieldmcshortnamevotest mech_votes_like_outparty
rename mech_fieldmcshortnamecompro mech_compromises_too_much
rename mech_iwouldsupportanotherfi mech_support_prim_challenge
rename mech_fieldmcshortnameisaloy mech_is_loyal_to_party


// Covariate for analysis, MC PID match
gen voter_mc_same_party = substr(inparty, 1, 1) == mcpartyoneletter
drop inparty
gen mc_pid_match = pid7 if mcpartyoneletter == "R"
replace mc_pid_match = 8 - pid7 if mcpartyoneletter == "D"
bysort mcpartyoneletter: reg mc_approve_scale mc_pid_match, robust // sanity check
reg mc_approve_scale voter_mc_same_party, robust
gen voter_republican = respondent_party == "R"

// Own view post treatment
gen post_bill0 = v150 + v152 + ownviewpost3fieldrandomvote_vote
gen post_bill1 = ownviewpost1fieldrandomvote_vote + ownviewpost2fieldrandomvote_vot
gen post_bill2 = v149
drop ownviewpost1fieldrandomvote_vote - ownviewpost3fieldrandomvote_vote
tab post_bill0
tab post_bill1
tab post_bill2

// Party perc post treatment
gen pperc_post_bill0 = v156 + v158 + partyperc3fieldrandomvote_vote_b
gen pperc_post_bill1 = partyperc1fieldrandomvote_vote_b + partyperc2fieldrandomvote
gen pperc_post_bill2 = v155
drop partyperc1fieldrandomvote_vote_b - partyperc3fieldrandomvote_vote_b
tab pperc_post_bill0
tab pperc_post_bill1
tab pperc_post_bill2

// Calculate pre-treatment issue agreement ON ALL ISSUES
gen pretreat_agreewmc_count = 0
gen pretreat_disagreewmc_count = 0
foreach var of varlist ownview* {
local billid = substr("`var'", 8, 5)
```

```
disp "`billid'"

gen pretreat_agreewmc`billid' = ownview`billid' == mcvote`billid' if ///
!missing(ownview`billid') & !missing(mcvote`billid')

replace pretreat_agreewmc_count = pretreat_agreewmc_count + 1 if pretreat_agreewmc
replace pretreat_disagreewmc_count = pretreat_disagreewmc_count + 1 if pretreat_agr
}
egen pretreat_agreewmc_share = rowmean(pretreat_agreewmc*)

// Own view post treatment
foreach var of varlist ownview* {
local billid = substr("`var'", 8, 5)
disp "`billid'"
gen posttreat_ownview`billid' = .
gen posttrt_thinksagrpty`billid' = .
forvalues b=0/2 {
// was respondent for or against?
replace posttreat_ownview`billid' = 1 if strpos(post_bill`b', "FOR") != 0 & randomv
replace posttreat_ownview`billid' = 0 if strpos(post_bill`b', "AGAINST") != 0 & ran

// does the respondent think they are congruent with party?
replace posttrt_thinksagrpty`billid' = 1 if strpos(pperc_post_bill`b', "FOR") != 0
posttreat_ownview`billid' == 1 & randomvote_id_`b' == "`billid'"
replace posttrt_thinksagrpty`billid' = 0 if strpos(pperc_post_bill`b', "FOR") != 0
posttreat_ownview`billid' == 0 & randomvote_id_`b' == "`billid'"
replace posttrt_thinksagrpty`billid' = 1 if strpos(pperc_post_bill`b', "AGAINST") !
posttreat_ownview`billid' == 0 & randomvote_id_`b' == "`billid'"
replace posttrt_thinksagrpty`billid' = 0 if strpos(pperc_post_bill`b', "AGAINST") !
posttreat_ownview`billid' == 1 & randomvote_id_`b' == "`billid'"
}

// pre-treat view on just issues also asked about post-treatment
gen pretrt_askedl8r_agreewmc`billid' = pretreat_agreewmc`billid' if ///
randomvote_id_0 == "`billid'" | randomvote_id_1 == "`billid'" | randomvote_id_2 ==

// post-treat view on issues (only up to 3 issues asked about)
gen posttreat_agreewmc`billid' = posttreat_ownview`billid' == mcvote`billid' if ///
!missing(posttreat_ownview`billid') & !missing(mcvote`billid')
}

drop post_bill0 - post_bill2
egen pretrt_justasked_agreewmc_share = rowmean(pretrt_askedl8r_agreewmc*)
egen posttreat_agreewmc_share = rowmean(posttreat_agreewmc*)
egen posttreat_thinksinline_share = rowmean(posttrt_thinksagrpty*)
```

```
// Code treatments
replace vote_type_treatment = "1 No Votes" if vote_type_treatment == "control"
replace vote_type_treatment = "2 Positive Votes" if vote_type_treatment == "agree"
replace vote_type_treatment = "3 Negative Votes" if vote_type_treatment == "disagre

replace trust_treatment = "1 Positive Game Experience" if trust_treatment == "posit
replace trust_treatment = "2 Negative Game Experience" if trust_treatment == "negat

encode trust_treatment, gen(treat_trust_game)
encode vote_type_treatment, gen(treat_vote_type)
```

3.2 Hypotheses

3.2.1 Does higher affective polarization affect accountability?

We are testing for whether there is an interaction between the two treatments, as we expect the negative effect of revealing an MC cast incongruent votes could become more or less negative under conditions of higher affective polarization. The main p-value of interest is the p-value on this interaction term.

The `mc_approve_scale` is our primary dependent variable. We also have several mechanisms we will test. These are as follows:

- `mech_votes_like_outparty`
- `mech_compromises_too_much`
- `mech_support_prim_challenge`
- `mech_is_loyal_to_party`
- `mech_anxious`
- `mech_time_looking_at_votes`

The relevant code is below:

```
// Analysis 1: Interaction of trust game treatment and vote type treatment
// Effects on mc_approve_scale and mechanism questions primary.
foreach dv of varlist mc_approve_scale mech_* {
disp "DV = `dv'"

// Expect positive interaction term because the negative effect of seeing ne
// votes will get less negative in high affective polarization condition
// This is main hypothesis test.
```

```
disp "Main test"
reg ‘dv’ i.treat_trust_game##i.treat_vote_type mc_pid_match pretreat_agreewr
test (2.treat_trust_game#2.treat_vote_type == 0) (2.treat_trust_game#3.treat

// To make the effects easier for readers to understand, we may also report
bysort treat_trust_game: reg ‘dv’ i.treat_vote_type mc_pid_match pretreat_ag

// Based on Study 1, we will examine effect when voter and MC are of same pa
disp "Same party only"
reg ‘dv’ i.treat_trust_game##i.treat_vote_type mc_pid_match pretreat_agreewr

// Exploratory HTEs:
if substr("‘dv’", 1, 5) != "mech_" { // We don’t ask the mechanism questions
disp "Other party"
reg ‘dv’ i.treat_trust_game##i.treat_vote_type mc_pid_match pretreat_agreewr
}

// We have no prediction here but we know it will be of interest to readers
disp "By respondent party"
bysort respondent_party: reg ‘dv’ i.treat_trust_game##i.treat_vote_type mc_p
}
```

There are also two other patterns that would speak to mechanisms.

- If people use positive info about same-party MCs more and negative info about opposite party MCs more, this could represent a “confirming priors” mechanism.
- Another pattern we could see, within the the pure control: just strengthening the effect of the negative information relative to control (with no change in the effect of positive information). We would expect this if higher affective polarization just makes people more averse to those who disagree (potentially including in-party MCs that disagree).

To help us interpret the effects, we may also calculate the following:

```
// Exploratory quantity we will calculate to help guide interpretation.
// People might be punishing their MC for taking votes they view as non-part
// when they actually are. Here, we are just interested in the descriptive q
// Note, these are only asked when there are no MC votes shown.
bysort voter_mc_same_party: sum posttreat_thinksinline_share if votes_shown

// We can also see the truth of how often they are actually in line.
bysort voter_mc_same_party: sum party_congruent_share_pre if votes_shown ==
```

3.2.2 Changes in Issue Opinions

We will also test whether there are changes in issue opinions.

```
// Analysis 2: More Follow-ing The Leader with high affpol?
gen votes_shown = treat_vote_type != 1 // in the control condition votes are

// Take variable of share agreeing with MC and recode as 1 - share agreeing
// MC if they are of different party to get party congruent share.
gen party_congruent_share_post = posttreat_agreewmc_share * voter_mc_same_pa
(1 - posttreat_agreewmc_share) * (1 - voter_mc_same_party)
gen party_congruent_share_pre = pretrt_justasked_agreewmc_share * voter_mc_s
(1 - pretrt_justasked_agreewmc_share) * (1 - voter_mc_same_party)

// Main analyses:
// When votes_shown = 0, we measure direct effect of affpol on holding party
// with the treat_trust_game coefficient.
// When treat_trust_game = 0, we measure direct effect of showing votes on
// party-congruent positions with the votes_shown coefficient.
// Interaction term then calculates whether following the leader increases
reg party_congruent_share_post i.treat_trust_game##i.votes_shown ///
party_congruent_share_pre

// Secondary: Examine whether effects of showing votes is driven by in- or
// out-partisan MCs.
bysort voter_mc_same_party: reg posttreat_agreewmc_share i.treat_trust_game
pretrt_justasked_agreewmc_share if votes_shown == 1
```

3.3 Manipulation Check

```
// Manip checks
reg manip_check_fair i.treat_trust_game, robust
reg affpol i.treat_trust_game pid_strength pid7, robust
```

Pre-Analysis Plan: Affective Polarization and Democratic Accountability, Survey 4

December 3, 2020

1 Preliminary Notes

- This is the pre-analysis plan for a survey experiment on how changes in affective polarization affect citizens' endorsement of and enforcement of various democratic norms.
- We already conducted three surveys. We are conducting this last survey to replicate prior results and test four new hypotheses. We may pool the coefficients across surveys.
- Unless otherwise stated, all data will be cleaned the same was as described in prior PAPs.

2 Hypotheses

As in prior PAPs, all hypothesis tests will be conducted using linear regressions of the following form, where `'dv'` is replaced with the DV of interest.

```
reg 'dv' i.treat_trust_game pid_strength pid7, robust
```

2.1 Social Distance

We will analyze the four social distance items as separate outcomes, coding each like a Likert scale with each response option one unit apart. We will also analyze a simple additive index of these items formed using the `alpha, std` command in Stata.

2.2 Replication: Norms Scale

We will form and analyze the norms scale in the same manner as in the prior PAP.

We will analyze a new item about voting by mail separately and not include it in the main norms scale.

2.3 Replication: Vignettes

We are asking two vignettes:

- Overriding close elections - inparty seeking to override
- Corruption tolerance - outparty engaged in corruption

We will code the outcomes in the same manner as in the previous PAP.

2.4 Violence

We will analyze each of the four violence items separately, and also form an additive index of all four items using the `alpha, std` command in Stata.

2.5 Perceptions of Reality

We will analyze the two perceptions of reality questions as separate outcomes, analyzing each like a Likert scale with each response option being one unit apart.

3 Manipulation Checks

We will conduct the following manipulation checks, identical to prior PAPs:

```
// Manip checks
reg manip_check_fair i.treat_trust_game, robust
reg affpol i.treat_trust_game pid_strength pid7, robust
reg elite_affpol i.treat_trust_game pid_strength pid7, robust
```

Note that the $\texttt{elite}_affpolvariableisanewmanipulationcheck; wearecheckingtoseetowhatextentthetrustgameaffectsc$