

To Discuss or Not to Discuss? How Selective Exposure to Political Discussion Conditions Experimental Findings on Polarization¹

Erin L. Rossiter² & Taylor N. Carlson³

September 21, 2023

Abstract

The question of how politically homogeneous and heterogeneous group discussion polarizes attitudes has received a great deal of scholarly attention, generally finding homogeneous group discussion can polarize and heterogeneous group discussion can depolarize attitudes. Some of the most rigorous evidence for these effects comes from experiments in which participants are asked to interact with other people in a randomly assigned group discussion setting (e.g., homogeneous or heterogeneous). However, this evidence has an important limitation. Americans have clear preferences for avoiding political discussions, and if they must discuss, most people have strong preferences for discussion with likeminded others. By assigning people to discussion settings without considering if they would ever choose that social setting themselves, prior experimental findings may have limited generalizability beyond the experimental setting. In this study, we incorporate self-selection into the standard political discussion experimental design. We use a Preference Incorporating Choice and Assignment (PICA) design to estimate a discussion setting's effect conditional on choice to receive (or avoid) it. We find that heterogeneous discussion has depolarizing effects for people who would and would *not* choose that social experience. Contrary to prior findings, we do not replicate that homogeneous discussion polarizes attitudes. Our findings have greater external validity than prior experimental work, better speaking to the effects of discussion we might see in the real world as people self-select into and out of different settings. As a result, the findings of this study help us understand how to best scale discussion-based interventions to reduce polarization.

¹We thank the Alexander and Diviya Magaro Peer Pre-Review Program at Harvard's Institute for Quantitative Social Science for valuable feedback during the design of this study. For generous research funding we thank the Social Science Research Council, the Francis and Kathleen Rooney Center for the Study of American Democracy at the University of Notre Dame, and the Institute for the Scholarship in the Liberal Arts, College of Arts and Letters at the University of Notre Dame. This research was approved by Washington University in St. Louis (ID 202304031) and University of Notre Dame (ID 23-04-7805) Institutional Review Boards and adheres to APSA's Principles and Guidance for Human Subjects Research. This experiment's design, hypotheses, and pre-analysis plan were pre-registered with OSF Registry (ID 052523) prior to data collection.

²Assistant Professor, University of Notre Dame, erinrossite@nd.edu, erossiter.com

³Associate Professor, Washington University in St. Louis, tncarlson@wustl.edu, sites.wustl.edu/tncarlson

1 Introduction

The question of how politically homogeneous and heterogeneous group discussion polarizes attitudes has received a great deal of scholarly attention (e.g., Klar 2014; Druckman, Levendusky and McLain 2017; Jost, Baldassarri and Druckman 2022), generally finding homogeneous group discussion can polarize attitudes and heterogeneous group discussion can depolarize. Many methodological approaches have been used to examine the relationship between political discussion composition (homogeneous vs. heterogeneous) and polarization (attitudinal, ideological, and affective), including: observational data of online social media networks (e.g., Barberá et al. 2015), survey data characterizing political discussion networks sometimes elicited via name generators and snowball sampling (Huckfeldt, Johnson and Sprague 2004; Mutz 2006; Minozzi et al. 2019), and deliberative polling (e.g., Fishkin et al. 2021). The most dominant approach, however, has been randomized experiments in which participants are asked to interact with other people in a randomly assigned political group context (Klar 2014; Druckman, Levendusky and McLain 2017; Levendusky and Stecula 2021; Ahn, Huckfeldt and Ryan 2014; Rossiter 2022; Broockman, Kalla and Westwood Forthcoming).

Experiments are a rigorous choice to estimate the effects of different discussion contexts because they allow researchers to estimate the causal effect of the features of the social experience of interest (e.g., the partisan composition of the group) on polarization. Other approaches, such as measuring the correlation between the partisan homogeneity within one’s political discussion network and their political attitudes, suffer from endogeneity issues. For example, people who *choose* to surround themselves with people from the other side of the political aisle might *already* hold less extreme attitudes or might be more open to changing their preferences than people who choose to surround themselves with only those who share their partisan identity. In this case, the political discussions themselves do not cause variation in polarization; rather the effects we observe are due to self-selection. Experiments overcome the problem self-selection causes when using observational data to understand the causal effects of political discussion contexts. By randomly assigning people to distinct

group contexts, often varying whether groups are politically homogeneous or heterogeneous, randomized experiments remove the threat of people selecting into treatment (i.e., preferable discussion settings). Removing the element of self-selection allows experiments to estimate a clean causal effect of a given political discussion on polarization outcomes.

This experimental design, sometimes called a “forced exposure” design because it assigns people to discussion contexts without considering if they would ever choose that social setting themselves, leads to an important limitation in interpreting the real-world implications of the results. If treatment effects of discussion settings are only observed amongst groups who would otherwise *not* join those settings in the real world, then findings from forced-exposure designs have limited generalizability beyond the experimental setting. Indeed, previous research has demonstrated that forced exposure designs sometimes overestimate treatment effects, relative to situations in which self-selection is allowed (e.g., Gaines and Kuklinski 2011; Arceneaux and Johnson 2015; Arceneaux, Johnson, and Murphy 2012; de Benedictis Kessner et al. 2019). For example, De Benedictis-Kessner et al. (2019) use a Preference Incorporating Choice and Assignment (PICA) design to examine the effects of exposure to partisan media on polarization, finding that partisan media can increase polarization both among people who would select into partisan media and those who would not, but the magnitude of this difference varied.

Political discussion is a particularly important context in which selection likely matters. Many people prefer to avoid political discussions in the first place. For example, Carlson and Settle (2022a) find that about one-third of Americans report that they prefer to avoid political discussion. Americans have clear preferences for avoiding political discussions in the first place and if they must discuss, most people report strong preferences for having those discussions with like-minded others. Moreover, these preferences are not uniformly distributed throughout the population. Demographic, political and psychological characteristics correlate strongly with these preferences. This further complicates the importance of selection in studies of political discussion.

In short, forced exposure designs could be overestimating the real-world effects of political discussion on polarization outcomes. In this study, we extend this line of work by seeking to understand differential effects of discussion based on individuals’ real-world tendencies to self-select into these settings, or even opt-out. When given the choice, people may choose like-minded discussants, people may choose to engage with a variety of viewpoints, or they may choose to avoid political discussion altogether. We expect whether and the extent to which discussion type (heterogeneous, homogeneous, or no discussion) polarizes or depolarizes attitudes will depend on people’s choice to engage in these social encounters.

In this study, we apply new experimental methods designed to incorporate self-selection to the political discussion experiment context. Because we expect that political discussion will have different effects for people who would and would not choose a given social experience, a central goal of this study is to estimate treatment effects of political discussion that incorporate this choice, and therefore likely have greater external validity than prior experimental work on the effects of political discussion. Specifically, we use an experimental design suited to understanding political discussion’s effects among different sub-groups of people who would choose to participate in the discussion or not. We apply the Preference Incorporating Choice and Assignment (PICA) experimental design (Knox et al. 2019; De Benedictis-Kessner et al. 2019), which allows a researcher to estimate an experimental intervention’s effect conditional on choice to receive it.

We find that heterogeneous political discussion had a durable, depolarizing effect for people who would self-select into this setting and people who would self-select out of it. Our results for heterogeneous discussion’s power to depolarize replicate prior experimental findings, and importantly, replicate these findings for people who are likely to receive this treatment in the real-world in addition to people who would not. This finding allows us to better understand the role self-selecting into depolarization interventions might play when interventions are used in the field. Interventions encouraging heterogeneous political discussion could still have an effect even if the only people to uptake are those that would seek out the opportunity.

While we replicate the finding that heterogeneous discussion depolarizes, we did not replicate the finding and conventional wisdom that homogeneous discussion polarizes. Even when disaggregating groups more likely to polarize from homogeneous settings, we find no evidence that homogeneous discussion polarizes.

2 Incorporating Political Discussion Preferences into its Polarizing Effects

This project reinvestigates the state of the literature on group discussion and polarization. Broadly, we argue that selection into different political discussion contexts is crucially important for interpreting the generalizability of treatment effects from forced-exposure group discussion experiments. In this section, we detail the dominant hypotheses from the literature, argue that self-selection could condition these effects, and describe the implications of understanding how self-selection does (or does not) condition the effects of group discussion contexts on polarization.

Because many group discussion contexts and types of polarization interest social scientists, we want to be clear about our language from the outset. First, we define "homogeneous" political discussion as conversations about political issues, candidates, or events with people who share similar views on the topics being discussed. In contrast, we define "heterogeneous" political discussion as conversations about political issues, candidates, or events with people who share different views on the topics being discussed. Note that we do not define homogeneity or heterogeneity in terms of partisan identity, as some past work has done (e.g., Klar 2014). Although partisan identity likely overlaps with views on the topics discussed in a given political discussion, we are reinvestigating claims about how discussion with others who are like-minded (not necessarily others who share an identity) affects policy preferences. When we say "polarization," we are referring to attitudinal polarization as the primary downstream outcome affected by different group discussion contexts.

Evidence from the political discussion and polarization literatures have generally points to two key findings: homogeneous political discussion increases polarization, while heterogeneous political discussion decreases polarization. However, previous research has not considered the potential differing effects of these distinct discussion contexts based on peoples' choice to join that kind of social setting. While forced exposure to heterogeneous discussion has positive effects for reducing polarization on average; and exposure to homogeneous discussion has negative effects for increasing polarization on average, the tendency to self-select into this exposure is an important moderator to for social scientists to understand because it structures the real-world implications of engaging in different political discussion settings.

2.1 Preferring Heterogeneous Discussion

First, consider people who actively choose to join heterogeneous discussions. These people may have higher political interest if they prefer talking politics to opting out of political discussion. Or, they may have political heterogeneity in their social networks that leads them to be comfortable with choosing mixed-view political discussion. Carlson and Settle (2022a) provide evidence for both of these possibilities. Despite the abundance of strong evidence that heterogeneous political discussion reduces polarization, it may be that those who opt-in to heterogeneous discussion enter the conversation having already reaped the benefits of being exposed to diverse views, and thus the effect of it will modest or even negligible.

When people who prefer heterogeneous discussion are forced into a homogeneous discussion, however, the results could be similarly interesting. Perhaps their inexperience with interactions with like-minded others could have a particularly strong polarizing effect. It may be that these people generally prefer heterogenous settings because they like the opportunity to persuade their peers, enjoy a good sparring match, or live for chaos. When the opportunity to argue is no longer there because the people in the discussion largely agree, this could push people farther to the extremes as they look for a way to create conflict. A different explanation could be that people who prefer heterogeneous discussion like it because they

are particularly open-minded, persuadable people who like hearing and learning from others. If so, a homogeneous discussion may depolarize these peoples' attitudes as they shift in light of seeing group unanimity on the issue.

2.2 Preferring Homogeneous Discussion

Next, consider people who would prefer homogeneous political discussion. Carlson and Settle (2022*b*) show that this is actually the most common political discussion preference, which is consistent with extensive previous research on political discussion networks indicating a strong pattern of politically homogeneous networks (e.g., Sinclair 2012; Huckfeldt, Johnson, and Sprague 2004; Ahn, Huckfeldt, and Ryan 2014; Carlson, Abrajano, and García Bedolla 2020; Minozzi et al. 2020). If people who prefer homogeneous discussion settings are less likely to be treated with heterogeneous discussion in the real world, this experience in a research setting could be particularly powerful, leading to large depolarizing effects. On the other hand, however, it is possible that this group, consisting largely of strong partisans, entrenches into their views and becomes resistant to the opposing views in the discussion. In this scenario, heterogeneous discussion could lead to no effects or even polarizing effects.

2.3 Preferring to Opt Out

Finally, consider people who would opt-out of political discussion. Carlson and Settle (2022*b*) show that about one-third of Americans fall into this category, reporting that they avoid political discussion at all costs, regardless of the views of the potential discussants. The desire to opt out of political discussion could be due to a distaste for politics (e.g., Klar and Krupnikov 2016; Krupnikov and Ryan 2022) or the anticipation of confrontation (Ulbig and Funk 1999; Carlson and Settle 2022; Mutz 2006; Connors and Howell n.d.), thus these people avoid political discussion when possible. However, in reality, political discussion may be more civil and informative than anticipated. This potential reality, coupled with the fact that people who prefer no discussion may not have strong political attitudes, could mean that

heterogeneous discussion serves its hypothesized role in reducing attitude polarization.

Similarly, undesired exposure to homogeneous discussion could result in strong polarizing effects. Some people prefer to avoid political discussion because they do not feel sufficiently knowledgeable about the topic. If someone would opt out of a homogeneous political discussion because they do not have enough information on the topic, the homogeneous conversation becomes a crucial information source for them. The information to which they are exposed in the discussion is likely to heavily favor one side, which could further polarize their attitudes.

In sum, there are strong theoretically-motivated reasons why selection into treatment could be particularly important in group discussion and polarization experiments. People who prefer heterogeneous, homogeneous, or no political discussion are fundamentally different from one another in ways that could moderate the effect each discussion context has on polarization.

3 Expectations

As discussed above, incorporating the role of self-selection into our understanding of the effects of discussion on polarization is important. The most rigorous evidence explaining the effects of political discussion on polarization comes from experiments, yet the "forced exposure" design of these experiments may limit how generalizable their findings are beyond the experimental setting. In this section, we outline expectations for how discussion may polarize beyond the experimental setting by incorporating that people have real-world tendencies for selecting *into* or *out of* different discussion contexts.

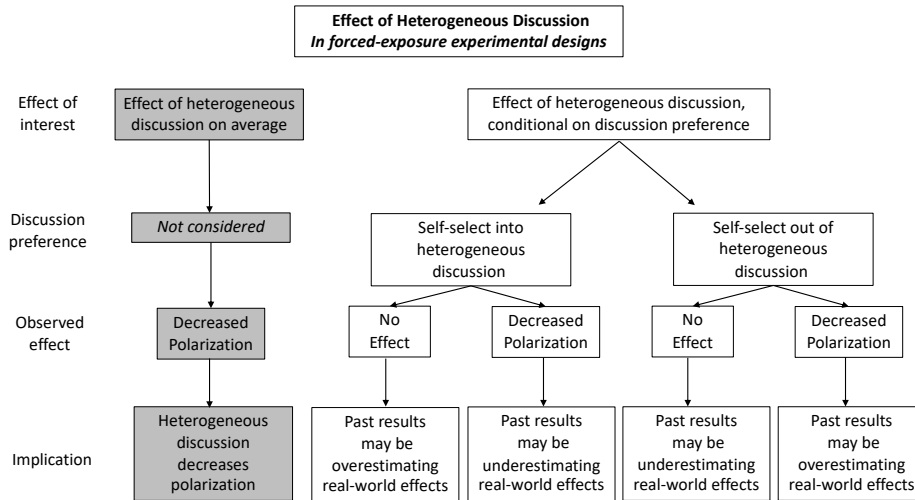
We visualize our main argument in Figures 1 (heterogeneous discussion) and 2 (homogeneous discussion). Both figures show that incorporating self-selection into our understanding of discussion's ability to (de)polarize is an important pursuit regardless of whether we find doing so replicates prior findings or shows prior findings are over- or under-estimating the real-world effects of discussion. Specifically, Figure 1 visualizes how the real-world effects

of heterogeneous discussion on reducing polarization may be over- or under-estimated in the literature depending on how the effects relate to self-selection. Figure 2 is a similar visualization for the real-world effects of homogeneous discussion on increasing polarization. We will use Figures 1 and 2 to outline our expectations and their implications.

3.1 Effect of Heterogeneous Discussion on Polarization

First, we will outline our expectations for the treatment effects of heterogeneous discussion amongst the different discussion preference groups discussed in Section 2—those who prefer heterogeneous discussion, homogeneous discussion, or opting out entirely. We illustrate our expectations in Figure 1. The left-most column in the flow chart, shaded in gray, represents the dominant trends from previous research. The remaining, unshaded portion of the chart reflects what we expect to observe if we were to consider individuals’ discussion preferences in the design.

Figure 1: Effect of Heterogeneous Discussion Incorporating Role of Self-Selection



Note: Gray boxes reflect the approach and dominant findings from previous research using forced exposure experimental designs. The remainder of the flowchart illustrates our expectations after accounting for discussion preferences.

As we highlight in gray in Figure 1, the dominant finding in the literature is that hetero-

geneous discussion is an important social setting for **depolarizing** attitudes. Importantly, this finding is an average treatment effect of heterogeneous discussion, thus is cannot speak to whether people with real-world tendencies to self-select into this setting might have a different treatment effect from heterogeneous discussion than those who would avoid this setting in real life. It may be that the findings from previous experiments overestimate the depolarizing effect for people who already prefer heterogeneous discussion, which is the very group most likely to engage in the behavior by choice.

Therefore, we first consider the effect of heterogeneous political discussion on polarization among those who prefer this type of discussion in the real world (the second path). In other words, this path of the flowchart examines possible outcomes of heterogeneous discussion for those who would self-select into this discussion context. The key implication is that if we find that heterogeneous discussion has **no effect** when we examine those who would opt-in to this setting in the real-world (the second path down the flowchart), past forced-exposure experimental findings are overestimating the real-world effects of heterogeneous discussion on reducing polarization.

However, if we instead observe that heterogeneous discussion **reduces polarization** for this group (the third path), then our results suggest that past results may be underestimating the real-world effects of heterogeneous discussion. If heterogeneous discussion reduces polarization among the very people most likely to be "treated" outside the experimental setting, past results may be downplaying the real-world implications of heterogeneous discussion for decreasing polarization.

Next, Figure 1 considers the effects of heterogeneous discussion on those who would not select into this setting—either they would prefer homogeneous discussion or would prefer to opt-out of political discussion entirely. We emphasize that this part of the flowchart replicates prior work by "forcing" exposure to people who would not encounter these settings in the real-world. If we find no effect of heterogeneous discussion among those who would not opt-in to heterogenous discussion (fourth path of the flowchart), then prior experimental

findings fail to replicate in this group. In this sense, observing no effect of heterogeneous discussion among this group would also imply prior work has underestimated the polarizing effects of heterogeneous discussion. This finding means the "treatment" does not work among people who would not receive it anyway in the real-world, so prior results, when being used to understand how heterogeneous discussion's effects generalize to the real-world, would be dampened by including this group. On the other hand, if we find heterogeneous discussion reduces polarization among this group (fifth path of the flowchart), then prior experimental findings were successfully replicated in this group. This suggests prior work may be overestimating real-world effects of heterogeneous discussion because the effects are being driven by people who otherwise would not engage in this kind of discussion in the real world.⁴

All together, we are interested in the treatment effects of different discussion settings conditional on the choice to join a heterogeneous discussion, homogeneous discussion, or no discussion at all. Knox et al. (2019) calls these causal quantities of interest the average choice-specific treatment effects (ACTEs). We are interested in three ACTEs of heterogeneous discussion. Specifically, we are interested in the effects of heterogeneous discussion, relative to no discussion, given discussion preference (i.e., the effect for those who would choose heterogeneous, homogeneous, and no discussion).

1. Effect of heterogeneous discussion, relative to no discussion, among those who prefer heterogeneous discussion (would self-select in, if given the choice)
2. Effect of heterogeneous discussion, relative to no discussion, among those who prefer homogeneous discussion (would self-select out, if given the choice)

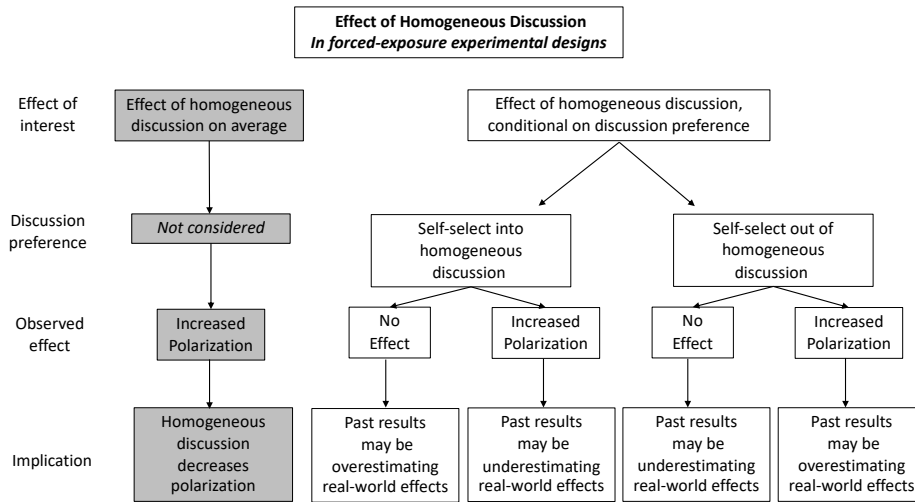
⁴We have not yet discussed the possibility that heterogeneous discussion actually polarizes attitudes. Our main expectations in Figure 1 consider the possibility that the depolarizing effect of heterogeneous discussion found on average in the literature either does or does not hold within different subgroups who have different discussion preferences. There is, of course, the possibility that these discussions backfire and polarize attitudes. We do not include this possibility explicitly in Figure 1, however, Figure 1 can be used to demonstrate the implications of this outcome as well. If heterogeneous discussion polarizes, we follow the logic of the "no effect" branches of the diagram because not only was there no effect in the expected direction, but it was in the opposite direction. Thus, the listed implications would only be even stronger if we found backfire effects in heterogeneous discussion.

3. Effect of heterogeneous discussion, relative to no discussion, among those who prefer no discussion (would self-select out, if given the choice)

3.2 Effect of Homogeneous Discussion on Polarization

Figure 2 similarly presents the implications of possible outcomes for the effect of homogeneous political discussions. The dominant finding in the literature, highlighted in gray on the left-hand side of Figure 2, is that engagement in homogeneous political discussion increases polarization. We investigate whether this pattern holds once we account for selection into homogeneous discussion.

Figure 2: Effect of Homogeneous Discussion Incorporating Role of Self-Selection



Note: Gray boxes reflect the approach and dominant findings from previous research using forced exposure experimental designs. The remainder of the flowchart illustrates our expectations after accounting for discussion preferences.

If we consider individuals' proclivity for such homogeneous discussion, we could come to different results. We present these possibilities in the unshaded portion of Figure 2. For example, consider the second path down the flowchart in Figure 2. Here, we consider people who prefer homogeneous discussion in the real world, so they would be likely to self-select into this context in reality and to be "treated" outside the experimental setting. If we observe that

homogeneous discussion does not increase polarization among those who would self-select into this context, then experimental findings from forced-exposure designs likely overestimate the real-world implications of homogeneous discussion for increasing polarization. That is, the popular narrative that echo-chambers increase polarization might be overstated because the very group most likely to be in an echo chamber in reality might be unaffected by any given homogeneous discussion. In contrast, if we observe that homogeneous discussion increases polarization among those who would self-select into this setting, then we have a successful replication of previous findings within this subgroup and provided evidence that forced-exposure experimental results may be underestimated.

Now, consider people who would not self-select into homogeneous discussion in reality, shown in the right-hand portion of Figure 2. If we observe that homogeneous discussion increases polarization among people who would not self-select into homogeneous discussion (i.e. people who prefer heterogeneous discussion or no discussion), then we have a successful replication of previous findings from forced-exposure experimental designs within this subgroup. This also has important real-world implications, suggesting that if people are inadvertently exposed to homogeneous group settings (i.e., they prefer no political discussion or heterogeneous political discussion, but are forced into a homogeneous setting, perhaps through social media algorithms or geographic sorting), polarization could indeed increase. However, if we observe no effect of homogeneous discussion on polarization among those who would self-select out of such a treatment in the real world, then we have failed to replicate past experimental findings within this subgroup, questioning how generalizable past results are beyond the experimental setting.

Like with heterogeneous discussion, we are interested in three ACTEs for homogeneous discussion. We are interested in the effects of homogeneous discussion, relative to no discussion, given discussion preference (i.e., the effect for those who would choose heterogeneous, homogeneous, and no discussion):

4. Effect of homogeneous discussion, relative to no discussion, among those who prefer

homogeneous discussion (would self-select in, if given the choice)

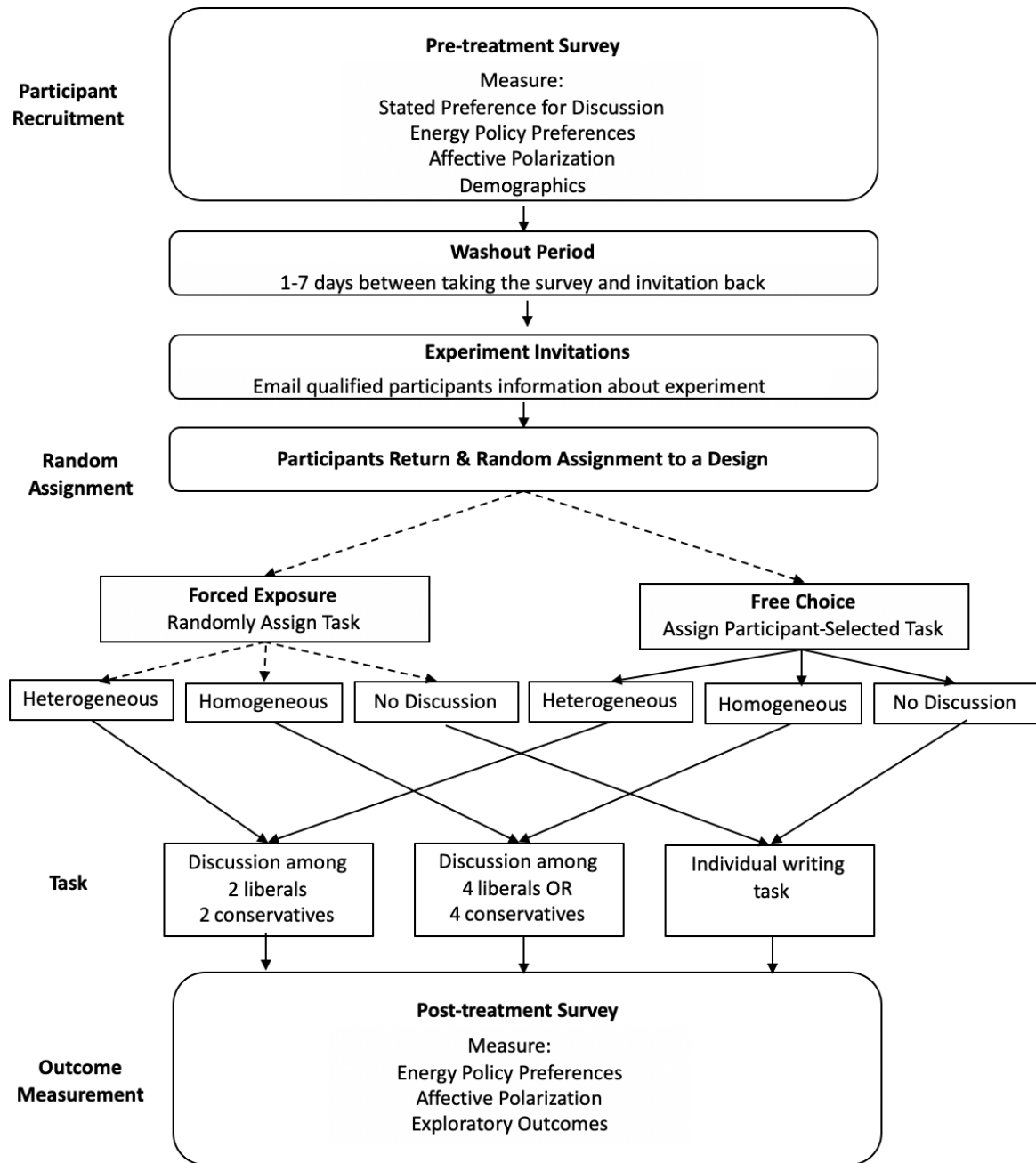
5. Effect of homogeneous discussion, relative to no discussion, among those who prefer heterogeneous discussion (would self-select out, if given the choice)
6. Effect of homogeneous discussion, relative to no discussion, among those who prefer no discussion (would self-select out, if given the choice)

4 Research Design

In order to understand how self-selection might condition the real-world effects of different discussion settings, we will use a Preference-Incorporating Choice and Assignment (PICA) experimental design, closely resembling the design in De Benedictis-Kessner et al. (2019) and introduced by Knox et al. (2019). This experimental design allows researchers to observe and compare the effects of their stimuli under free choice and forced exposure (i.e, randomization). This allows us to estimate the effects of discussion when people are "forced" to be there and when people select into the experience.

Figure 3 summarizes this design. Broadly, the design includes a recruitment survey where we ask about political discussion preferences, a washout period, and then an invitation to the experiment. When invited participants return, they are randomly assigned to a design: forced-exposure or free choice. Those in the **forced-exposure design** are randomized a second time. They are randomly assigned to heterogeneous, homogeneous, or no discussion, regardless of their pre-treatment stated preference for discussion. Those in the **free-choice design** are not randomized, but instead can choose their task (and will be assigned what they choose): heterogeneous, homogeneous, or no discussion. Participants then engage in either their randomly assigned or chosen task, and then take a post-treatment survey so we can learn about attitudinal polarization from the discussions.

Figure 3: Visualization of Group Political Discussion PICA Design



4.1 Recruitment, Washout Period, and Experiment Invitation

First, we begin by administering a recruitment survey in which we measure participants' stated preferences for three types of discussion (heterogeneous, homogeneous, or no discussion). We also collect demographics and pre-treatment measures of our outcomes of interest, which focus primarily on attitudes regarding the conversation topic in the design (energy policy) to measure attitudinal polarization. Specifically, we use an additive index based on four energy policy questions listed in Appendix B.1.

Key for the PICA design is collecting participants' stated preference for different discussion settings. This question allows us to understand what people who are randomized into a discussion setting would have freely chosen, if given the chance. Therefore, this question is our key operationalization of what discussion setting each participant prefers, and would choose if given the chance. The question wording is in Appendix B.2. It outlines that we will be posting a task in the future where participants read about energy policy, then engage with the topic in one of three ways: talking with people who agree with you, talking with a mix of people who agree and disagree with you, or writing about your views alone. After explaining the future task, we ask "If given the choice, in which of the following tasks would you choose to participate?" We also include additional measures of individuals' real-world discussion preferences, outside of the online study context, in order to demonstrate the preference for the study task correlates with these real-world behaviors.

The next step of the design is a washout period between the recruitment survey and the experimental portion of the study. Although (De Benedictis-Kessner et al. 2019) had a washout period of only a few minutes within the same survey, we will have a longer washout period of approximately 1-7 days. The washout period has a range because the recruitment survey will be live for several days, varying the length of time between when a participant takes the pre-treatment survey and returns for the experimental portion of the study.

During the washout period, we will determine which participants from our recruitment survey are qualified for the experimental portion of the study. The selection criteria are

as follows: (1) the participant answers "yes" to a question asking if we can contact them via the online panel notification to participate in future study, (2) the participant's energy policy preference not exactly equal to 4, on average, thus we know if they are conservative or liberal on the topic of discussion, and (3) the participant passes the open-ended text response quality check question. Moreover, we will invite an equal number of participants we classify as "liberal" vs. "conservative" on our energy policy index. After the washout period, qualified participants will be invited via a notification in the platform to return to the experimental task.

4.2 Experimental Randomization

The next section of Figure 3 demonstrates the randomized portion of the study. When invited participants return after the washout period, we will first randomize them to a design—either the forced exposure condition (assigned with probability .80) or the free choice condition (assigned with probability .20). If assigned to the forced exposure condition, participants will be randomly assigned to one of the three discussion types with equal probability—heterogeneous discussion, homogeneous discussion, or a solo task in lieu of discussion. If assigned to the free choice condition, participants will choose which type of task to join: a heterogeneous discussion, homogeneous discussion, or a solo task.⁵

After we have randomly assigned a participant to either the forced-choice or free-choice arm of the experiment, participants in the free-choice arm will be asked one survey question that participants in the forced-choice arm are not asked. Participants in the free-choice arm will be asked to choose their preferred task. This question exactly mirrors the preference question asked in the recruitment survey, except it makes clear the participant is indeed choosing what task to join in just a few moments (rather than asking which task they would

⁵Note that we make it clear to participants that all three tasks take the same amount of time to complete, have the same compensation rates, and are anonymous. The discussion tasks do not require extensive additional coordination effort or time spent waiting for discussion partners. We intend for these clarifications to remove extraneous motivations that might correlate with the kind of task, like thinking one task will be much faster than another, from the participants' decision calculus.

join hypothetically as in the recruitment survey).

As a final step prior to joining the assigned or chosen task, all participants in the experiment will be presented with a short description of U.S. energy policy tradeoffs shown in Appendix B.4. We modeled these policy descriptions after previous work on attitude polarization and group discussion by Klar (2014).

Next, participants will join their assigned or chosen task. Participants in the forced-choice design will be told at this point what their randomly assigned task is (i.e. that they will be having a heterogeneous or homogeneous discussion, or writing individually). Participants in the free-choice design will likewise be informed of the specifics of their chosen task. Participants will also be filtered into a specific chatroom according to the pre-treatment measure of respondents' attitudes on the topic (either liberal or conservative on the issue of energy policy). Specifically, liberal (conservative) participants assigned to or choosing homogeneous discussion will discuss with three other participants who we determined also have liberal (conservative) views on the issue. Heterogeneous discussions will have two participants with conservative views and two participants with liberal views, which follows closely from previous group discussion research (Klar 2014; Druckman, Levendusky and McLain 2017). Those who were assigned to or chose no discussion will write about the topic by themselves. The specific task prompts are shown in Appendix D.

Participants will discuss or individually write about the issue for 10 minutes. All discussions occur via text-chat using the Chatter app (Rossiter 2020). Chatter is an app that facilitates online chats. Its flexibility makes it uniquely compatible for studies like this that involve complex designs.

4.3 Post-treatment Surveys

Finally, after completing their discussion or essay, participants are asked to complete a short post-treatment survey. We again ask about energy policy preferences, and for exploratory purposes we ask about perceptions of and reactions to the task (was it a positive experience,

did you learn about the outparty, and did you learn about energy policy).

Three days after completing the experiment, participants will be recontacted via the platform’s notification feature to complete a short follow-up survey to determine the durability of treatment effects. We repeat the four core dependent variables about energy policy preferences.

5 Results

We conducted this experiment in June 2023 on CloudResearch’s online crowdsourcing platform, similar to Amazon Mechanical Turk, called Connect. Throughout one week, 2434 participants took our recruitment survey and qualified to be invited to the experiment. Of those participants, 940 were classified as having conservative attitudes on energy policy and 1494 were classified as having liberal attitudes. We invited all 940 conservative respondents and a random sample of 940 liberal participants. Of this sample, 997 participants returned and fully completed their assigned task, which also means their discussion group had all four members participate.

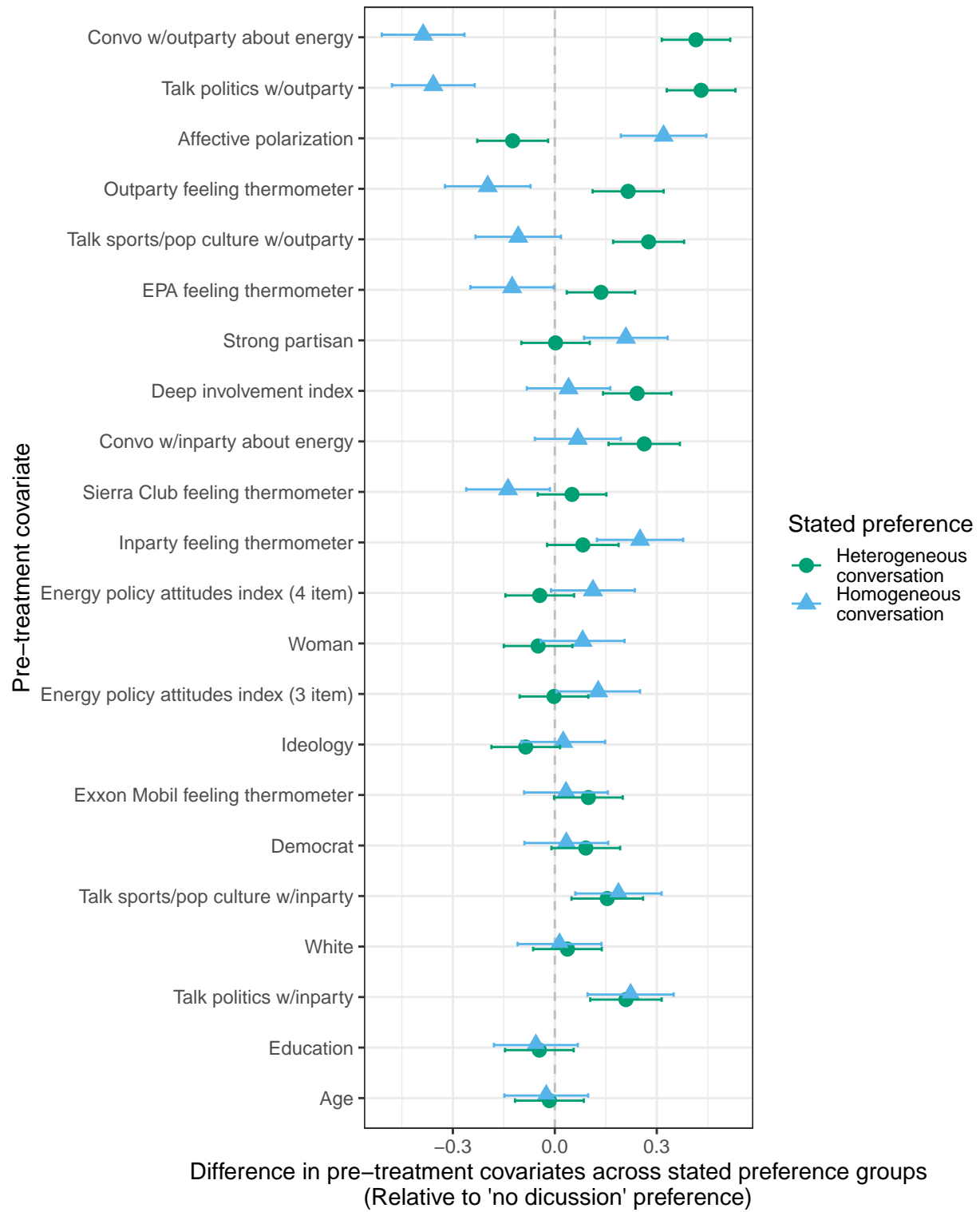
5.1 Describing Preference Groups

First, we investigate the free choice arm of the experiment. In our main sample of 997 participants, 39%, 20%, and 41% stated they would prefer heterogeneous, homogeneous, and no discussion respectively.⁶ First, we will assess how this preference correlates with other characteristics before we turn to examining the extent to which stated preference for a given discussion setting equates to what participants actually choose to do when given the choice.

Figure 4 reports linear regression coefficients describing the relationship between stated discussion preference and all pre-treatment covariates asked in the survey, standardized to have mean 0 and standard deviation of 1. We find that preferences for different discussion

⁶These numbers are consistent with the full sample invited to the study.

Figure 4: Pre-treatment Covariates by Preference Group



settings are most strongly correlated with willingness to talk to the outparty about both political and non-political topics. People who prefer heterogeneous discussion, relative to the group who prefer to opt-out of political discussion, are more likely to be willing to talk to the outparty about politics; whereas, those who prefer homogeneous discussion are less likely to be willing to talk politics with outparty members. Likewise, both the groups that prefer heterogeneous and homogeneous conversation are more likely to be willing to talk with inparty members about politics, relative to those who prefer to opt-out. These results validate our survey question measuring participants' stated preference for different discussion settings. (See Appendix C for additional validation of the stated preference question.)

We also find these discussion setting preferences to be highly correlated with partisan strength and affective polarization. People who prefer homogeneous discussion are stronger partisans and more affectively polarized. Finally, people who prefer homogeneous discussion also hold stronger views on energy policy. Notably, demographics such as age, education, race, gender, and partisanship are not strong predictors of political discussion preferences.

Figure 4 demonstrates the many ways selection into discussion settings correlates with covariates that also explain discussion's potential effects. Researchers who use randomized experiments to understanding the effects of different discussion on political attitudes and behaviors alleviate the threat self-selection imposes on these causal effects. However, these key differences across groups who prefer heterogeneous discussion, homogeneous discussion, and opting out entirely suggest that their reactions to different discussion settings might also vary in important ways.

However, people *saying* they prefer a certain discussion setting may not actually *choose* to engage in it when the opportunity presents itself. The free choice arm of the PICA design allows us to investigate whether stated preference aligns with actual behavior. In the free choice arm, we find that 65%, 44%, and 87% of those who stated preferring heterogeneous, homogeneous, and no conversation respectively chose that option in the experiment. When deviating, the people who stated they would prefer homogeneous or heterogeneous conversation

instead chose to opt-out (47% and 31% respectively). Very few people opted *for* a discussion setting that was against their stated preference. (Full results are in Appendix E.)

People who deviated from their stated preference of homogeneous or heterogeneous conversation, choosing to opt out when faced with a real choice, may have behaved inconsistent with their stated preference because the reality of having to share their views with others—likeminded or not—was intimidating or anxiety inducing, as political discussion can be in the real-world. However, an assumption of the PICA design is that a participants’ stated preference can substitute for observing their actual behavior. In a future draft, we will conduct sensitivity analyses introduced by Knox et al. (2019) to assess if our results hold under violations of this assumption.

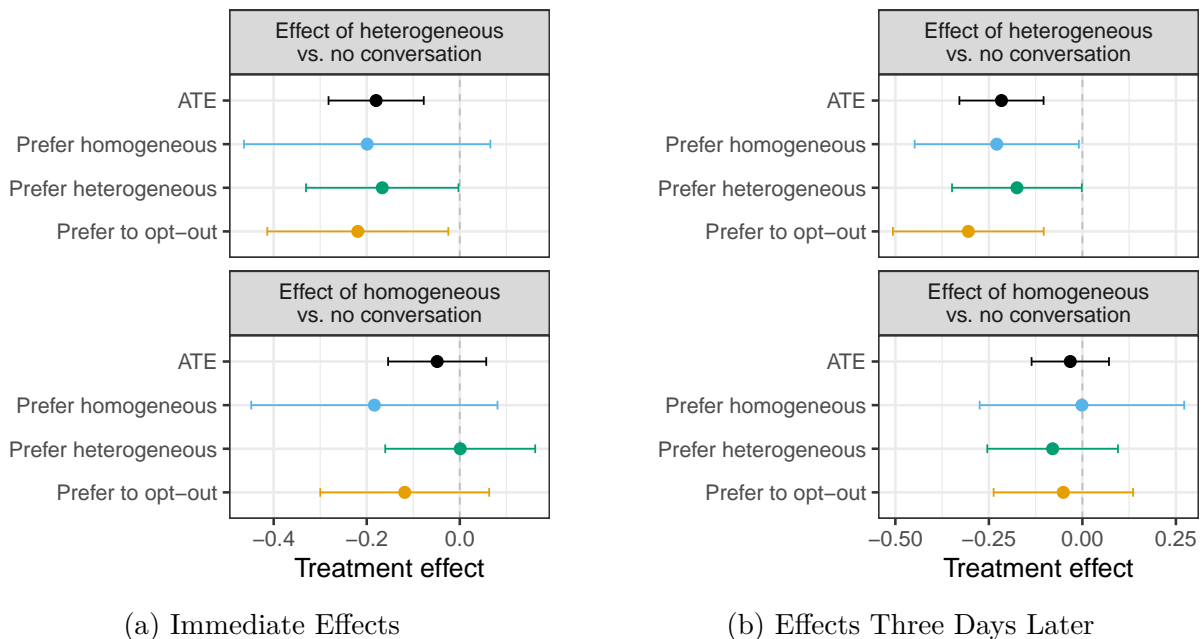
5.2 Forced Choice Results

Next we turn to the forced choice arm of the experiment. We estimate average treatment effects (ATEs) to replicate prior experimental work. We also estimate the ACTEs for the six quantities outlined in Section 3 by estimating subgroup treatment effects given stated preferences amongst participants in the forced-exposure arm of the experimental design (Knox et al. 2019).⁷ We use a linear model estimator for each test and cluster standard errors for participants in the same conversation. In all models, we control for pre-treatment covariates to increase precision. We control for participants’ pre-treatment measure of the outcome variable if applicable and the following demographics: age, gender, education, race and ethnicity, party identification, ideology.

Our main outcome of interest is attitude polarization. To measure polarization, we follow previous research and fold our energy policy index so that larger values mean the participant’s views are more polarized and smaller values mean less polarized. We fold the variable based on our pre-treatment determination of if the participant has liberal or conservative views on

⁷This is the naïve estimator of ACTEs because it assumes any deviations between stated preferences and actual choice are not associated with potential outcomes of the discussion treatments. We will probe the sensitivity of our results to deviations in this assumption in future drafts.

Figure 5: ACTE Estimates for Polarization of Energy Policy Attitudes for Each Discussion Preference Group



the issue.

Figure 5 shows results for our main pre-registered outcome of interest—attitudinal polarization regarding energy policy. The outcome is standardized to have a mean of 0 and a standard deviation of 1, so treatment effects can be interpreted in terms of standard deviations from the mean. First, we see that ATE of heterogeneous, relative to no discussion, depolarizes attitudes. This replicates prior experimental work. However, we see that the ATE for the effect of homogeneous discussion relative to no discussion is not significant, which fails to replicate prior findings on the polarizing effect of like-minded.⁸

The core of our inquiry is whether incorporating discussion preferences will change our understanding of discussion's effects. We find that heterogeneous conversation depolarizes regardless of whether one would self-select into this setting or not. Despite Figure 4 demonstrating that the preference subgroups being markedly distinct in ways that could affect their response to a discussion, the ACTE for each discussion preference group is negative. (The

⁸Heterogeneous discussion's depolarizing effect is stronger than homogeneous ($p=0.013$).

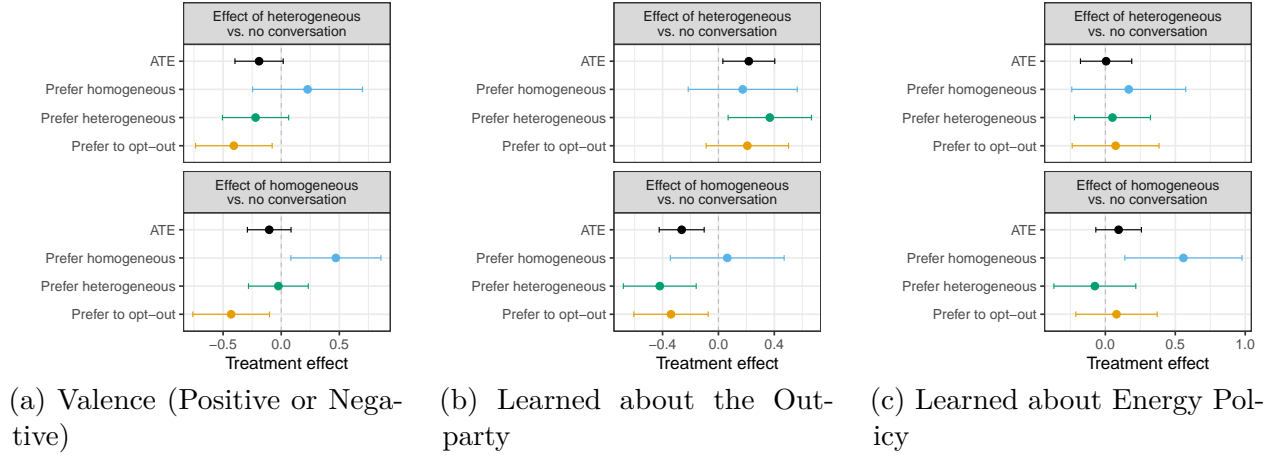
ACTE for those that prefer homogeneous discussion has high levels of uncertainty ($p=0.146$) because this is the smallest subgroup in our sample.)

Turning to the effect of homogeneous discussion among preference groups, we see suggestive evidence that those who want to be there depolarize the most. This runs contrary to the echo chamber literature that suggests homogeneous discussion can create a self-reinforcing cycle where people's opinions are amplified, rarely challenged, and features group think, which taken together can contribute to polarization. We note these results feature high levels of uncertainty, and future work should replicate if homogeneous discussion can indeed depolarize the people that choose those setting.

We resurveyed participants at least three days after their conversation to ask the same energy policy questions. We find that the depolarizing effect of heterogeneous conversation is durable across all preference subgroups. Any effects of homogeneous conversation, on the other hand, decayed to zero over time.

Despite finding no differences in treatment effect depending on your preference for the social setting, we find stark differences by preference subgroup regarding their experience in the discussion. Figure 6 reports exploratory outcomes we asked to probe potential mechanisms of discussion's effects. People who prefer to opt-out of political discussion felt either kind of discussion was a negative/bad experiences, relative to short essays. In stark contrast, there was a strong association for reporting homogeneous conversations as positive/good for those who wanted to be there. Likewise, the only group who had a positive effect on learning about energy policy were those who were exposed to homogeneous discussion and wanted to be there. Finally, people who wanted to be in a homogeneous setting had null effects on learning about the outparty, but the effect was negative for those who were forced into the likeminded setting. Despite these stark contrasts in how people experienced and perceived their discussion, we find no evidence of distinct effects of discussion on polarization based on their preference to be in that setting or not.

Figure 6: ACTE Estimates for Possible Mechanisms for Each Discussion Preference Group



6 Conclusion

This study makes two key contributions. First, from a theoretical perspective, this study seeks to better understand the extent to which preferences for different kinds of political discussions, or an aversion to political discussion altogether, can contribute to polarization. Previous research on political discussion highlights the importance of these decisions in explaining which political discussions are likely to occur (Carlson and Settle 2022a), but research on the relationship between political discussion and polarization has not yet adequately considered these preferences. We find that, despite preferences for discussion settings being correlated with important political characteristics, like affective polarization and partisan strength, heterogeneous discussion depolarizes attitudes regardless of tendency to self-select into this setting in the real world. Contrary to prior experimental findings, homogeneous discussion fails to polarize attitudes in our study, and we fail to find evidence for a polarizing effect among all preference subgroups, even those who would normally opt-out of political conversation and thus might have the most malleable views.

Second, from a more practical standpoint, understanding heterogeneity based on political discussion preferences will help us understand how interventions to reduce polarization may scale in the real world beyond the experimental setting. Many political discussion experiments

are justified as important because of their potential to help reduce polarization in the real world. However, if treatment effects are only observed among people who would not otherwise choose discussion in the real world, then interventions need to consider how to get people to come together for discussion in the first place. However, we do not find this to be the case.

Our results suggest heterogeneous discussion was a powerful depolarizing experience regardless of one's interest in joining this setting. This is important for depolarization practitioners. These results suggest that people willingly joining opportunities for open dialogue and diverse perspectives—the people walking through the door at depolarization events—may still see positive effects from depolarization interventions. It is not the case that their tendency to expose themselves to these kinds of settings in their everyday lives has put a ceiling on the positive effects of engaging with diverse perspectives and information. These findings are equally important for our understanding of compelling people to partake in heterogeneous settings and political discussion in general. Heterogeneous discussion had durable depolarizing effects among people who preferred a likeminded discussion or to avoid political talk altogether.

References

- Ahn, TK, Robert Huckfeldt and John Barry Ryan. 2014. *Experts, activists, and democratic politics: Are electorates self-educating?* Cambridge University Press.
- Barberá, Pablo, John T Jost, Jonathan Nagler, Joshua A Tucker and Richard Bonneau. 2015. “Tweeting from left to right: Is online political communication more than an echo chamber?” *Psychological science* 26(10):1531–1542.
- Broockman, David, Joshua Kalla and Sean Westwood. Forthcoming. “Does Affective Polarization Undermine Democratic Norms or Accountability? Maybe Not.” *American Journal of Political Science* .
- Carlson, Taylor N and Jaime E Settle. 2022a. *What goes without saying*. Cambridge University Press.
- Carlson, Taylor N. and Jaime E. Settle. 2022b. *What Goes Without Saying: Navigating Political Discussion in America*. Cambridge University Press.
- De Benedictis-Kessner, Justin, Matthew A Baum, Adam J Berinsky and Teppei Yamamoto. 2019. “Persuading the enemy: Estimating the persuasive effects of partisan media with the preference-incorporating choice and assignment design.” *American Political Science Review* 113(4):902–916.
- Druckman, James N, Matthew S Levendusky and Audrey McLain. 2017. “No Need to Watch: How the Effects of Partisan Media Can Spread via Interpersonal Discussions.” *American Journal of Political Science* .
- Fishkin, James, Alice Siu, Larry Diamond and Norman Bradburn. 2021. “Is deliberation an antidote to extreme partisan polarization? Reflections on “America in one room”” *American Political Science Review* 115(4):1464–1481.
- Huckfeldt, Robert, Paul E Johnson and John Sprague. 2004. *Political disagreement: The survival of diverse opinions within communication networks*. Cambridge University Press.
- Jost, John T, Delia S Baldassarri and James N Druckman. 2022. “Cognitive–motivational mechanisms of political polarization in social-communicative contexts.” *Nature Reviews Psychology* 1(10):560–576.
- Klar, Samara. 2014. “Partisanship in a social setting.” *American Journal of Political Science* 58(3):687–704.
- Knox, Dean, Teppei Yamamoto, Matthew A Baum and Adam J Berinsky. 2019. “Design, identification, and sensitivity analysis for patient preference trials.” *Journal of the American Statistical Association* 114(528):1532–1546.
- Levendusky, Matthew S and Dominik A Stecula. 2021. *We Need to Talk: How Cross-Party*

Dialogue Reduces Affective Polarization. Cambridge University Press.

Minozzi, William, Hyunjin Song, David MJ Lazer, Michael A Neblo and Katherine Ognyanova. 2019. "The Incidental Pundit: Who Talks Politics with Whom, and Why?" *American Journal of Political Science* .

Mutz, Diana C. 2006. *Hearing the other side: Deliberative versus participatory democracy*. Cambridge University Press.

Rossiter, Erin. 2022. "The Consequences of Interparty Conversation on Outparty Affect and Stereotypes." Working paper. <https://erossiter.com/files/conversations.pdf>.

Supplemental Information: To Discuss or Not to Discuss? How Selective Exposure to Political Discussion Conditions Experimental Findings on Polarization¹

Taylor N. Carlson² & Erin L. Rossiter³

August 21, 2023

¹We thank the Alexander and Diviya Magaro Peer Pre-Review Program at Harvard's Institute for Quantitative Social Science for valuable feedback during the design of this study. For generous research funding we thank the following the Social Science Research Council, the Francis and Kathleen Rooney Center for the Study of American Democracy at the University of Notre Dame, and the Institute for the Scholarship in the Liberal Arts, College of Arts and Letters at the University of Notre Dame. This research was approved by Washington University in St. Louis (ID 202304031) and University of Notre Dame (ID 23-04-7805) Institutional Review Boards and adheres to APSA's Principles and Guidance for Human Subjects Research. This experiment's design, hypotheses, and pre-analysis plan were pre-registered with OSF Registry (ID 052523) prior to data collection.

²Assistant Professor, Washington University in St. Louis, tncarlson@wustl.edu, sites.wustl.edu/tncarlson

³Assistant Professor, University of Notre Dame, erinrossite@nd.edu, erossiter.com

Contents

A Ethical Considerations	2
B Question Wording	4
B.1 Energy Policy Preferences	4
B.2 Stated Preference for Political Discussion (Recruitment Survey)	5
B.3 Free-Choice Arm Question Allowing Participants to Choose Task	5
B.4 Information about U.S. Energy Policy	6
C Validating the Stated Preference Question	7
D Experimental Conditions	10
D.1 Heterogeneous Discussion Instructions	10
D.2 Homogeneous Discussion Instructions	10
D.3 No Discussion - Individual Essay Instructions	10
E Free Choice Arm	12
F Inclusion/Exclusion Criteria	12
G Predicting Choice as an Alternative to Stated Preference	13

Appendix A Ethical Considerations

The human subjects research conducted for our study adheres to APSA's Principles and Guidance for Human Subjects Research. The study was approved prior to data collection by the authors' institutional review board under expedited review. The IRB deemed: "the criteria for approval are met per 45 CFR 46.111 and/or 21 CFR 56.111 as applicable. Project determined to be minimal risk per 45 CFR 46.102(i) and/or 21 CFR 56.102(i) as applicable." A modification to the study protocol was also approved in which the only change was to increase the number of participants enrolled in the study.

All participants were paid above federal minimum wage. All participants who completed the recruitment survey (3-5 minutes) were paid \$1. All participants who were invited to return to the conversation/short essay portion of the survey were paid \$5.00 if they returned. If they also completed their task the post-treatment survey that immediately followed the task, they were paid an additional \$1.00. This task took about 15 minutes. Participants who returned but could not complete the task because their conversation did not reach four participants or ran into a technical difficulty were still paid the full \$5.00 compensation. Finally, participants who completed the study were invited back to a follow-up survey asking only a few questions that took 1-2 minutes to answer. All participants who completed this question were paid \$0.75.

Prior to beginning the recruitment survey, all participants read an information sheet to obtain their informed and voluntary consent for the entire study. Participants had to check a box indicating "I consent" to proceed, which is how we documented consent in an online environment. Participants also had the option to select "I do not consent to participate," which would allow them to leave the survey immediately.

We will share the full consent information sheet upon request. Importantly, among other things, our information sheet shared how long we anticipated the tasks would take, the payment for completing the tasks, and how they would be paid. We also told participants that they could choose to stop participating at any time. We also told participants that we would keep the information they provided confidential, and that we would not be collecting any personally identifying information. Finally, we told participants that anonymized transcripts of their conversation or short essay would be made available for research purposes as described, involving removing any people's names, places,

religious or cultural backgrounds, occupations, family relationships, and any other potentially identifying information that they may have disclosed in their conversations (even though they were explicitly not prompted to do so).

Our participants came from CloudResearch’s panel called Connect, which has a subject pool broadly demographically diverse within the U.S. population and not comprised mainly of members of groups we should consider vulnerable or marginalized. This research did not differentially harm particular demographic groups.

Appendix B Question Wording

B.1 Energy Policy Preferences

We measure energy policy preferences following questions used in previous research on attitudinal polarization (Klar 2014). We ask four questions about energy policy, each with seven-point Likert scale response options. We use these responses to create an additive index. Based on pre-testing question wordings and topics on Mechanical Turk to evaluate preference distributions, we determined that we will consider participants to hold liberal attitudes on the issue if their average responses to the four questions are greater than 4. If their average response is less than 4, we will consider them to hold conservative attitudes. We will exclude respondents who have an average response of exactly 4.

1. Some people think it is important to protect the environment, even if it costs some jobs or otherwise reduces our standard of living. Other people think that protecting the environment is not as important as maintaining jobs and our standard of living. Using the scale below, which of these positions comes closest to your own point of view? [7 point scale ranging from 1 = Protect the environment, even if it costs some jobs; to 7 = Jobs are more important than the environment]
2. To what extent do you personally oppose or support efforts to increase drilling off the coastal waters of the United States? [7 point scale ranging from 1 = Strongly oppose; to 7 = Strongly support]
3. To what extent do you personally oppose or support efforts to increase drilling on federal lands? [7 point scale ranging from 1 = Strongly oppose; to 7 = Strongly support]
4. To what extent do you personally oppose or support the Biden Administration’s decision in 2022 to release oil from the country’s emergency reserve? [7 point scale ranging from 1 = Strongly support; to 7 = Strongly oppose]

B.2 Stated Preference for Political Discussion (Recruitment Survey)

In the future, we will post a task where you will first be asked to read a short summary of the energy policy positions preferred by Republicans and Democrats. After reading the summary, you may be able to choose which of three related tasks you would like to complete.

The tasks have many similarities. They all:

- take the same amount of time (10 minutes for the task, plus 1-2 minutes answering survey questions)
- have the same payment and bonuses available (up to \$6)
- award payments and bonuses based on your individual participation
- will be ready to complete right away – there would be no need to come back at another time or wait for other study participants to begin
- have a written (typed) component and do not require video or verbal communication
- allow you to remain anonymous

The main difference between the tasks is simply how you’d prefer to engage with the topic of energy policy.

If given the choice, in which of the following tasks would you choose to participate?

- A 10-minute chat with other Connect Workers where the views on energy policy are mixed (some people will have similar views as you and some people will have different views than you)
- A 10-minute chat with other Connect Workers where the views on energy policy are similar to yours (everyone will have similar views as you)
- A 10-minute individual essay-writing task about your views on energy policy

B.3 Free-Choice Arm Question Allowing Participants to Choose Task

Next, you will be asked to read a short summary of the energy policy positions preferred by Republicans and Democrats. Now, you will choose which of three related tasks you would like to complete after reading the summary.

The tasks have many similarities. They all:

- take the same amount of time (10 minutes for the task, plus 1-2 minutes answering survey questions)
- have the same payment and bonuses available (up to \$6)
- award payments and bonuses based on your individual participation
- will be ready to complete right away – there would be no need to come back at another time or wait for other study participants to begin
- have a written (typed) component and do not require video or verbal communication
- allow you to remain anonymous

The main difference between the tasks is simply how you'd prefer to engage with the topic of energy policy.

What would you like to do?

- A 10-minute chat with other Connect Workers where the views on energy policy are mixed (some people will have similar views as you and some people will have different views than you)
- A 10-minute chat with other Connect Workers where the views on energy policy are similar to yours (everyone will have similar views as you)
- A 10-minute individual essay-writing task about your views on energy policy

B.4 Information about U.S. Energy Policy

Many Americans are concerned with our country's dependence on foreign oil and our rising gas prices. American politicians have offered distinct approaches to improving energy policy. While these policies are not entirely at odds with each other, the government will likely prioritize one policy approach over the other.

One approach is to make it easier for companies to receive permits to drill for oil off the coast of the US (most significantly oil-rich Alaska). The argument is that easing the application process for drilling will increase oil production. It will remove the nearly endless regulations that oil producers must overcome, many of which have no apparent positive role – it is widely agreed that these excessive regulations do not protect the environment. Easing restrictions would expand the US energy supply, create American jobs, and help lower gas prices. Republicans often support this approach.

Another approach is to encourage the use of domestically produced alternative fuels (for example, wind and solar power) and increase incentives for designing more energy efficient vehicles. The argument is that many environmental and taxpayer savings would come from tightening fuel efficiency regulations and reducing reliance on gas-guzzling vehicles that harm the environment. The government would provide assistance to companies working to improve the efficiency and gas mileage of gas-powered cars and trucks. Emphasizing alternative energies would reduce our demand for oil, both at home and abroad, therefore reducing gas prices, while generating new jobs in renewable energy development. Democrats often support this approach.

Appendix C Validating the Stated Preference Question

Figure A.1: Usually Avoid or Enjoy Political Discussion and Stated Preference for Experiment Task

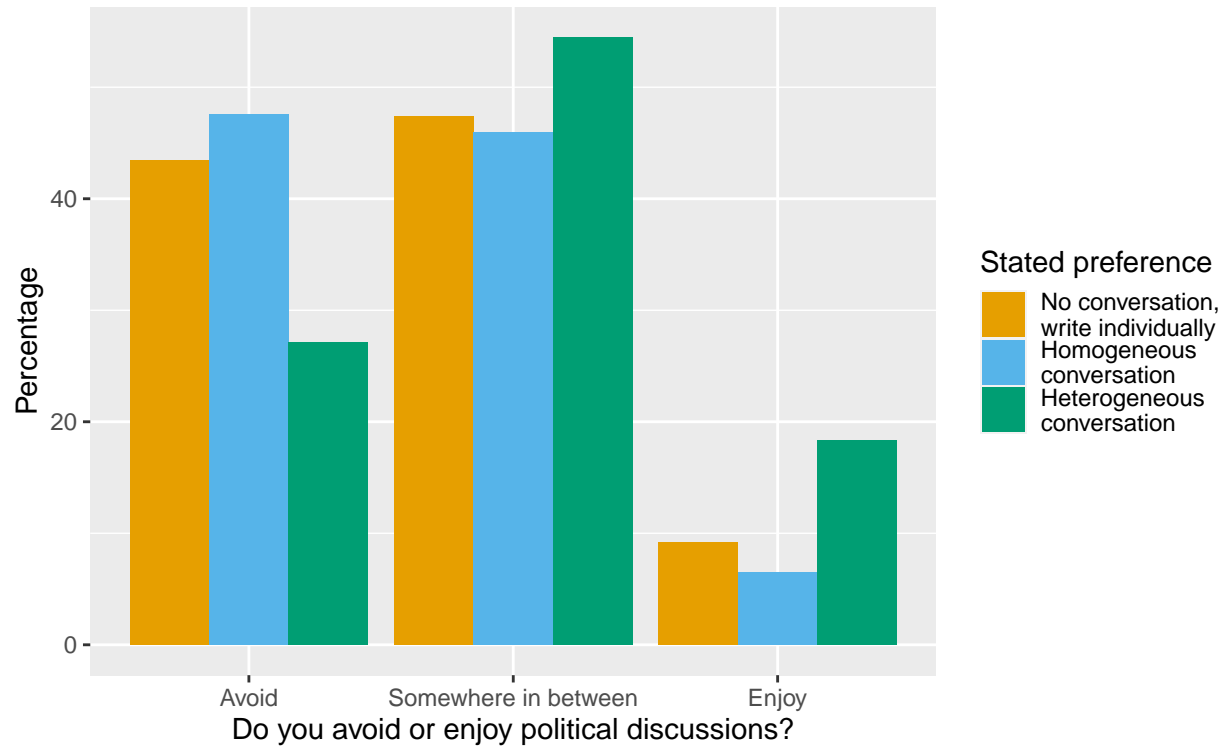


Figure A.2: Needing to Know Others' Views for Joining Political Discussions Stated Preference for Experiment Task

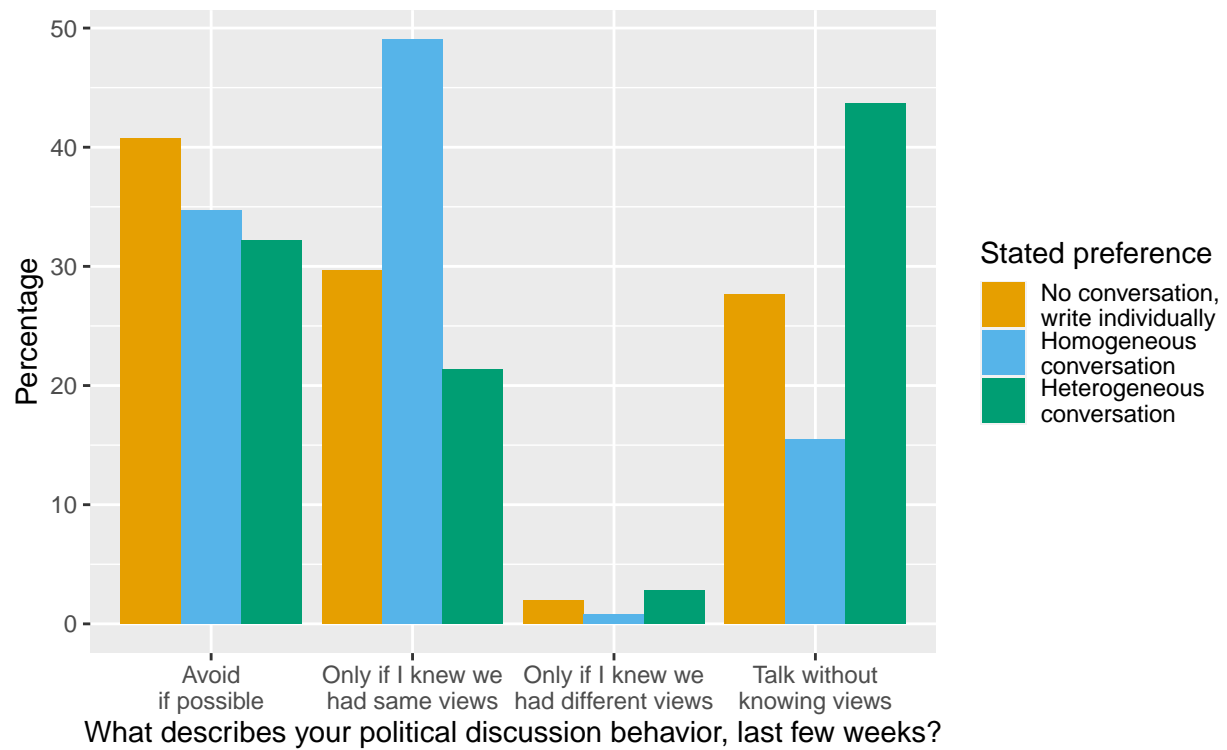
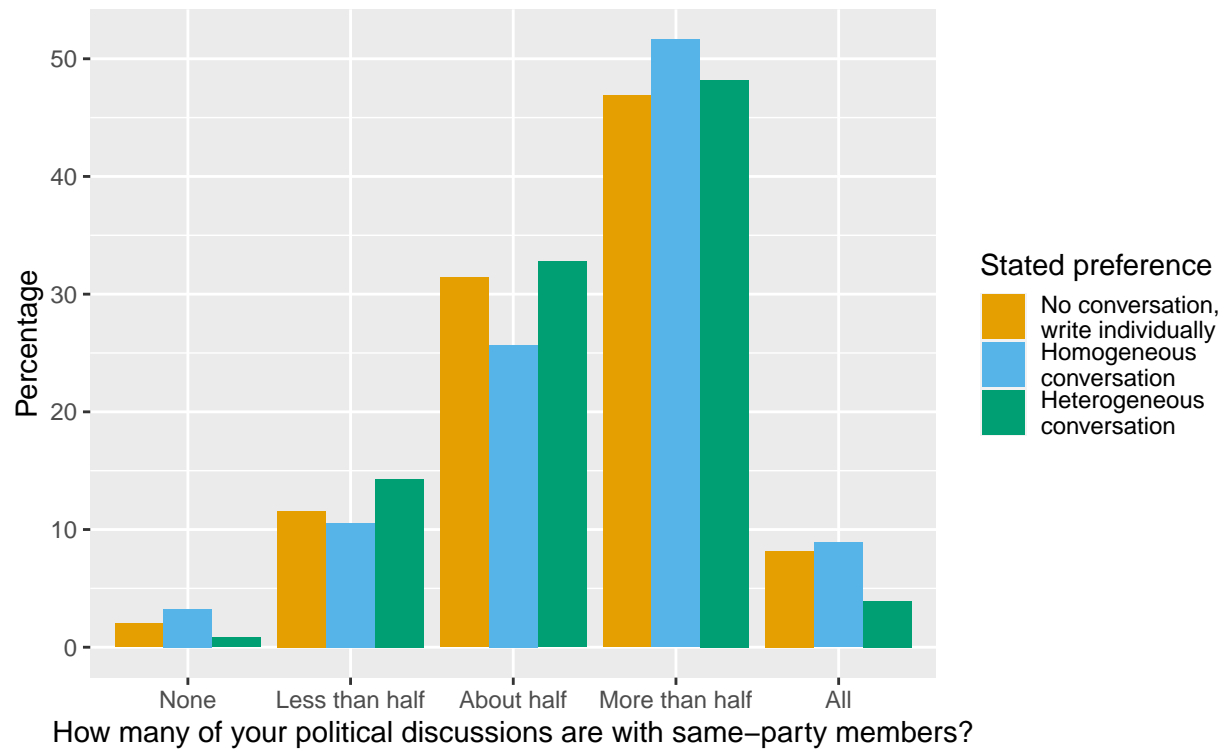


Figure A.3: How Often Political Discussion is with Same-Party and Stated Preference for Experiment Task



Appendix D Experimental Conditions

D.1 Heterogeneous Discussion Instructions

As you read a few moments ago, the U.S. government must balance environmental and economic concerns when creating energy policy. (You can reference that information by clicking [here](#) which will open a new tab in your browser.)

You are now in a chatroom with a group of people whose views on U.S. energy policy are mixed. Some people have similar views as you and others have different views than you.

Please have a conversation focusing on the following question: **Do you think the U.S. government should prioritize environmental or economic concerns when creating energy policy?**

There will be three participants joining you. We expect you to chat for 10 minutes. You can advance by clicking the Done button when the timer reaches 0.

Please start by sending a message so your group knows you've entered the chatroom.

D.2 Homogeneous Discussion Instructions

As you read a few moments ago, the U.S. government must balance environmental and economic concerns when creating energy policy. (You can reference that information by clicking [here](#) which will open a new tab in your browser.)

You are now in a chatroom with a group of people whose views on U.S. energy policy are similar to yours. Everyone in the group has similar views as you.

Please have a conversation focusing on the following question: **Do you think the U.S. government should prioritize environmental or economic concerns when creating energy policy?**

There will be three participants joining you. We expect you to chat for 10 minutes. You can advance by clicking the Done button when the timer reaches 0.

Please start by sending a message so your group knows you've entered the chatroom.

D.3 No Discussion - Individual Essay Instructions

As you read a few moments ago, the U.S. government must balance environmental and economic concerns when creating energy policy. (You can reference that information by clicking [here](#) which will open a new tab in your browser.)

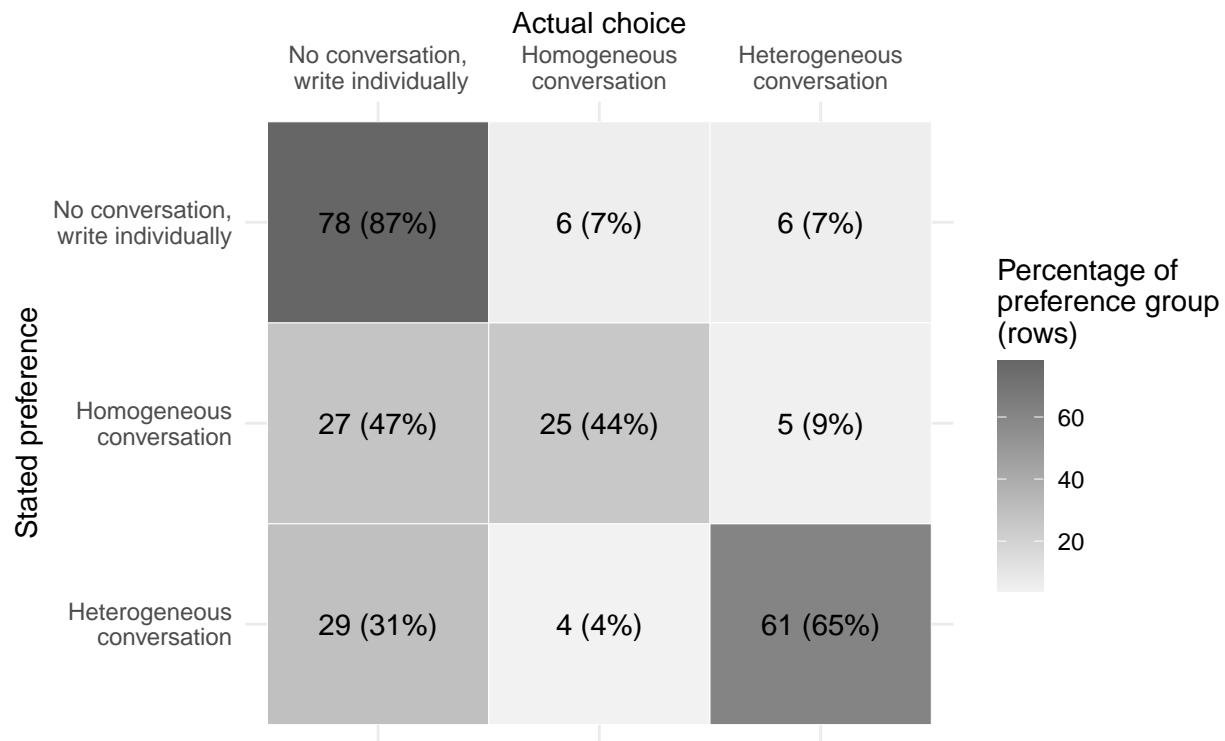
Now, please write individually about U.S. energy policy.

In your response, please focusing on the following question: **Do you think the U.S. government should prioritize environmental or economic concerns when creating energy policy?**

There will be three participants joining you. We expect you to chat for 10 minutes. You can advance by clicking the Done button when the timer reaches 0.

Appendix E Free Choice Arm

Figure A.4: Stated Preference and Actual Choice in Free Choice Arm



Appendix F Inclusion/Exclusion Criteria

If a person failed to take the follow-up survey (thus we do not have an observation for them in the analysis), but they joined their assigned treatment condition (i.e., entered the chatroom), this could be for two reasons. They could have had a technical error advancing to the final survey or they could have left the chat early.

Regardless, we do not exclude the other chat participants from our analyses. If a participant fails to complete the post-survey because of the second listed reason—they attrit—this sort of individual-level self-censorship is common in real-world political discussions. All four participants in a chatroom entered the homogeneous or heterogeneous setting, and if one (or more) people choose not to speak or choose to leave, that is akin to real-world discussions, thus it is a part of our bundled treatment and we include participants who experienced these kinds of discussions in our main analyses.

Appendix G Predicting Choice as an Alternative to Stated Preference

As we discuss in the main text, we used a pre-treatment survey question to measure what people would choose among our three experimental tasks if given the choice. In other words, we collected a stated preference. Using the free choice arm, we found that participants' stated preference did not predict actual behavior. In particular, people who stated they would prefer discussion opted out of discussion when given the choice, as discussed in Appendix E.

Collecting behavior using the free choice arm was critical for this experimental design. First, while we intended to use the stated preference question to infer what the forced choice arm would choose if given the choice, using the free choice arm, we were able to learn our stated preference question was in fact not a good measure of choice in this context. Second, to overcome this issue, in this Appendix we describe how we train a predictive model of choice using the free choice arm. Using this model, we then predict choice for those in the forced choice arm.

Recall we began with 264 observations in the free-choice arm, where 145 chose no conversation, 79 chose heterogeneous conversation, and 40 chose homogeneous conversation. These are our labelled data for the building a predictive model of experimental task choice to apply to participants in the forced choice arm for whom we did not observe their choice. The labelled data are unbalanced. Therefore, we begin by using the "SMOTE" method (Chawla et. al. 2002) used for unbalanced classification problems. We use this algorithm to artificially generate observations for the heterogeneous and homogeneous choice classes. The algorithm uses the five nearest neighbors to do so. We use all available pre-treatment covariates, including the stated preference question discussed above, demographics (age, gender, race, education, partisan identity, ideology), self-reported political discussion behaviors (frequency of political discussion with same party, tendency to avoid political discussion, and network composition), energy policy attitudes, deep involvement index, willingness to talk to outparty about energy policy, willingness to talk to inparty about energy policy, outparty feeling thermometer, inparty feeling thermometer, Exxon Mobil feeling thermometer, willingness to talk to outparty about politics, and willingness to talk to inparty about politics.

After generating the artificial observations, we have 140 and 144 observations in the heterogeneous and homogeneous classes respectively.

We use Breiman's random forest algorithm from the `randomForest` package in R.

.

References

Klar, Samara. 2014. "Partisanship in a social setting." *American Journal of Political Science* 58(3):687–704.