**The Democracy of Dating: How Politics Shape Relationship Formation**

**[Not to be included in printed versions]**

**Online Appendix Table of Contents:**

1. Additional Background Literature/Theoretical Framework
2. APSA’s Recommended Reporting Standards for Experiments
3. More Details on the Survey and Treatments
4. Additional Preregistered Heterogeneities
5. Additional Results Not Included in the Paper

**1. Additional Background/Theoretical Framework**

*Affective party polarization* (or “affective polarization” for short) has been defined as animosity and mutual dislike between Republicans and Democrats. Scholars have argued that partisan affect may be driven by a complex web of ideological differences, social connections, social context, and/or simple disdain for members of an out-group (Iyengar et al. 2019; Robison and Moskowitz 2019; Mason 2018; Druckman et al. 2019; Rogowski and Sutherland 2016).[[1]](#footnote-1) While there is disagreement over why partisan affect is growing, several scholars have looked to the increasingly fractured and politicized media environment (Lau et al. 2017; Druckman et al. 2019; Henderson and Theodoridis 2018). Partisan affect may cause individuals to perceive that out-party individuals lack a variety of desirable individual characteristics, from intelligence to emotional aptitude (Chen and Kendrick 2002; see also Rothschild et al. 2019 and Goggin et al. 2019). This phenomenon may also shape how we understand basic facts about the world around us. Partisan animus is likely to touch all aspects of our lives. In theory, the role of partisan affect is not constrained to the political realm.

While affective polarization has taken center stage in recent years, this concept is not without its skeptics in the scientific community. Some have argued, for example, that measures of affective polarization are not fully informative of animosity towards out-parties, but rather a dislike for partisanship more generally (Klar and Krupnikov 2016, C.4; Klar, Krupnikov, and Ryan 2018; Levendusky and Malhotra 2016; Klofstad, McDermott, Hatemi 2012; Klofstad, McDermott, and Hatemi 2013). This prominent and growing hypothesis asserts that a core part of what appears to be partisan animosity is actually attributable to an aversion to engage with partisan extremists who convey their politics too loudly/strongly.[[2]](#footnote-2)

Given qualitative accounts of growing animus between political parties, a growing body of survey research has been conducted in an attempt to measure levels of affective polarization in the United States (for a thorough overview of this topic, see Iyengar et al. 2019). Doing so is no easy task as partisan animus is a somewhat slippery concept to fully capture. According to recent cross-national surveys, Republicans and Democrats increasingly loathe each other (Iyengar, Sood, and Lelkes 2012). Today, people are more likely to say that they would feel uncomfortable if a family member married someone in another political party (these effects are highest among those on the ideological poles). Individuals are also increasingly willing to ascribe negative adjectives to the out-party and positive adjectives to the in-party.

While directly asking people how much they dislike the out-party is one approach to measuring partisan affect, it is certainly not the only--nor the best--way to do so. Indeed, social desirability may mask individuals’ willingness to report their true feelings towards the out-party.[[3]](#footnote-3) Just as we might not trust individuals to truthfully report their levels of racial bias, we might also be skeptical of individuals’ willingness to say that they dislike Republicans/Democrats. Social desirability may help to explain why survey-based measures of partisan affect remain relatively low, despite a general sense that animus runs rampant. Indeed, while a growing number of Americans say they would be unhappy if an immediate family member were to marry someone in the other party, this still holds true for only a minority of Americans.[[4]](#footnote-4) While other measures--such as the ANES feeling thermometer--indicate higher (and growing) levels of affective polarization (especially among Republicans) it remains somewhat uncertain just how high levels of partisan animus are based on survey self-reports alone.[[5]](#footnote-5) In addition to social desirability, individuals may not be capable of evaluating their own levels of latent party animus. Doing so may require more self awareness than individuals possess. Finally, survey self-reports of partisan animus may actually pick up on other (parallel) attitudes like a general disdain for politics in general or other forms of bias. For example, Republicans may dislike Democrats because they perceive them as belonging to racial out-groups, being socially distant, and/or having a different set of core individual beliefs. When a Republican is told about a Democrat, they may imagine someone who is an atheist minority who lives in an urban environment in the northeast. These attributions may activate other forms of individual bias. If this were the case, then attempts to measure partisan affect could really be picking up on racial, geographic, or religious biases (for example).[[6]](#footnote-6)

For all these reasons, we may want to move beyond direct questioning to measure partisan affect. An alternate approach is to use randomized control trials. By making comparisons across groups exposed to various randomly-assigned partisan manipulations, we are able to 1) hold all else equal and 2) see how people behave in controlled social environments. While experiments designed to detect partisan animus are not easy or straightforward and still might be detecting separate forms of individual bias--depending on their design features--scholars have shown their promise in this area (e.g. Iyengar et al. 2019; Iyengar and Westwood 2015; Gift and Gift 2015; McConnell et al. 2018; Michelitch 2015; Shafranek Forthcoming; Mason 2016). Indeed, experimental approaches have (increasingly) become more common as a means of detecting a variety of forms of bias.

Our key area of interest in measuring any partisan biases is among (seemingly) non-political romantic relationships. We choose this domain given the vital importance of relationship building. Who individuals choose to associate with, date, and/or marry is a foundational part of modern and historical societies the world over. The early part of relationship formation is especially important given the potential for transformation and convergence once relationships have been well established (Iyengar et al. 2019). And, as we outline in greater detail in the next section, previous research has yielded vastly different estimates of just how much politics affects behavior in this domain.

Previous survey-based research has shown that politics can sometimes bleed not only into friendships, but also romantic relationships. Some of the earliest work on this topic in political science showed that there was substantial evidence of “assortative mating” in terms of married partners’ social attitudes (e.g. Alford et al. 2005; Funk et al. 2013; Hatemi et al. 2009; Hatemi et al. 2010; Hatemi et al. 2011; Martin et al. 1986; McDermott et al. 2014; Smith et al. 2011; Smith et al. 2012). This work connects to a literature in psychology and sociology that too shows that political preferences play a key role in family formation (e.g. Kandler et al. 2012; Luo and Klohnen 2005; Watson et al. 2004). More recently, work has found that “political attitudes display interspousal correlations that are among the strongest of all social and biometric traits” (Alford et al. 2011, 362).[[7]](#footnote-7) Scholars have also found that some people report a willingness to avoid dating people of the out-party (Petter 2018), which comports with validated levels of partisan intermarrying documented in voter files (Iyengar, Konitzer, and Tedin 2018; Hersh and Ghitza 2018). Further, some scholars have found that many people desire to avoid politics altogether when pursuing others romantically (Klofstad, McDermott, Hatemi 2012; Klofstad, McDermott, and Hatemi 2013). However, there is some evidence that this is changing; more recent evidence has reported an increase in individuals’ willingness to signal their own political leanings in dating environments (Frisbie 2016; Iyengar et al. 2019; Kiefer 2017). In total, however, most (but not all, as we note in the next section) of the research done on how politics does/does not change has been observational--leaving open the possibility that any documented patterns of sorting are due to other difficult to observe characteristics that actually are behind descriptive patterns in partisan homophily in relationship building.

Our results speak to the literature on affective polarization in important ways. Our combination of multiple experimental approaches gives us a more complete picture of the impact of partisan biases on (seemingly) non-political behaviors.Importantly, our results provide concrete evidence that partisan animus is *not* simply a dislike of politics that is independent from any partisan penalties themselves as recent scholars have argued (Klar and Krupnikov 2016, C.4; Klar, Krupnikov, and Ryan 2018; Klofstad, McDermott, Hatemi 2012; Klofstad, McDermott, and Hatemi 2013). We provide evidence that partisan biases in contemporary society are present, but these play less of a role than previously thought. Our results also provide depth to our understanding of how to experimentally detect levels of partisan bias, helping to rule out other potential counter-explanations like demand effects or other individual biases. Finally, they help to start to shine light an experimental literature that has, to this point, yielded vastly different results on the effect of politics on relationship building. That previous experiments have yielded effects as disparate from null to a whole (statistically) standard deviation necessitates the exploration that we provide here.

Our results find special meaning given the rapidly-growing role of online dating in contemporary society (Rosenfeld et al. 2019) and the increasingly willingness people have shown to signal their own political attachments in online dating formats (Frisbie 2016; Iyengar et al. 2019; Kiefer 2017). They speak to a core individual behavior, as who people choose to associate with, date, and/or marry is a foundational part of modern and historical societies the world over. The early part of relationship formation is especially important given the potential for transformation and convergence once relationships have been well established (Iyengar et al. 2019).

**2. APSA’s Experimental Research Sections’ Recommended Reporting Standards for Experiments**

Here we include (or reference the reader to) all of the information that the Experimental Standards Committee of APSA’s Organized Section on Experimental Research lists to include (at minimum) for all experimental research. Though some of the information listed below is provided at other points throughout the article and the Appendix, we have put all of the responses to the recommended items to report below in one place here, so as to facilitate the reader in finding access to this information. The response items are included in **blue**.

**A. Hypotheses**

* What question(s) was (were) the experiment designed to address? What are the specific hypotheses to be tested?

The article was designed to test this research question: to what extent do political identities (i.e. partisanship, ideology, and candidate support) affect relationship building via online dating? In our pre-analysis plan, we listed party, age, gender, and marital status as individual characteristics we wished to test for heterogeneities among.

**B. Subjects and Context**

* Eligibility and exclusion criteria for participants: Why was this subject pool selected? Who was eligible to participate in the study? What would result in the exclusion of a participant? Were any aspects of recruitment changed (such as the exclusion criteria) after recruitment began?

There are two subject pools included in the study--one from Amazon’s Mechanical Turk (MTurk) and one from the Cooperative Congressional Election Study (CCES, 2018 Brigham Young University individual module). These two subject pools were chosen to ensure that results from the MTurk sample generalized to a more representative population. In the MTurk pool, subjects who did not have an U.S. IP address or have a high approval rating were excluded. No aspects of recruitment changed after recruitment began. For more details about the sampling framework and design of the CCES, please see <https://cces.gov.harvard.edu/>.

* Procedures used to recruit and select participants. If there is a survey: Identify the survey firm used and describe how they recruit respondents.

As stated on their website, “The CCES is a 50,000+ person national stratified sample survey administered by YouGov. Half of the questionnaire consists of Common Content asked of all 50,000+ people, and half of the questionnaire consists of Team Content designed by each individual participating team and asked of a subset of 1,000 people...A large portion of the CCES respondents are YouGov panelists. These are people who have made an account on yougov.com to receive periodic notifications about new surveys. Others are recruited live from online advertisements or are recruited from another survey provider. Therefore, while panelists are prompted to participate in the CCES, they opt-in to being a YouGov panelist. In order to make the sample representative, not all respondents to the CCES questionnaire end up in the final dataset. To read more about the pruning process used to match to the target population, please refer to the guide.” For more information about the CCES, please see <https://cces.gov.harvard.edu/frequently-asked-questions>. Mechanical Turk subjects are drawn from the pool of individuals using the platform. All of the surveys were fielded via Qualtrics.

* Recruitment dates defining the periods of recruitment and when the experiments were conducted. Also list dates of any repeated measurements as part of a follow-up.

Study 1 took the survey on November 2, 2018. Study 2 (the CCES sample) took the survey between September 27, 2018 and November 05, 2018. Study 3 took the survey on March 6, 2019. Study 4 took the survey on April 22, 2019. Study 5 took the survey on April 6, 2020. All surveys were single shots, so there were no repeated measurements as part of a follow-up.

* Settings and locations where the data were collected. In the field, lab, classroom, or some other specialized setting? Other relevant specifics of the population: e.g., large public university vs. small private university; geographic location; etc.

All of the surveys were from online panels.

* If there is a survey: Provide response rate and how it was calculated.

The CCES’s response rate in 2018 ranged from 30.0-35.2%, depending on which AAPOR response rate is used (see [here](https://sda.berkeley.edu/sdaweb/docs/cces2018/DOC/CCES+Guide+2018.pdf)). To our knowledge, there is no way to calculate a response rate for MTurk workers given that the denominator for a specific time period is unknown.

**C. Subjects and Context**

* Details of the procedure used to generate the assignment sequence (e.g., randomization procedures).

In all experiments, individuals were block randomized at the individual level to a dating profile based on their political party (3 item) and gender that they most preferred to date (male or female). There was no clustering in the random assignment.

* If random assignment used, then details of procedure (e.g., any restrictions, blocking). Note the unit of randomization (individuals, groups, households, etc). Pay careful attention to report clustered random assignment if subjects were assigned at some level other than the individual subject.

In all experiments, individuals were block randomized at the individual level to a dating profile based on their political party (3 item) and gender that they most preferred to date (male or female). There was no clustering in the random assignment.

* If random assignment used, provide evidence of random assignment. If blocking was used, and group assignment proportions were not equal across blocks, provide table for each of the blocks.

Figure A8 provides evidence of random assignment by providing tests for covariate balance. With the blocking, group assignment proportions were equal across blocks.

* Blinding: Were participants, those administering the interventions, and those assessing the outcomes unaware of condition assignments? If blinding took place, include a statement regarding how it was accomplished and how the success of blinding was evaluated.

With the exception of the demand effects group in Study 3, all subjects were blind to the intervention assignment in the sense that they did not know what the research question was or the group to which they were being compared to. (In the demand effects condition, all subjects knew was our research question, which we told them.) The survey experiment was administered online and not by any proctor, hence there was no traditional blinding required.

**D. Treatments**

* Description of the interventions in each treatment condition, as well as a description of the control group. Descriptions should be sufficient to allow replication: Summary or paraphrasing of experimental instructions in the article text; verbatim instructions and/or other treatment materials provided in an appendix.

Figure A1 provides the images shown in the dating profiles (which also includes the control group). The question wording for the outcomes is included in Section 2 of this Appendix.

* How and when manipulations or interventions were administered. Method of delivery: Pen-and-paper vs. computer or internet vs. face-to-face communication vs. over the telephone. If computerized, the software should be described and cited. (If possible, programs should be included in appendix so as to be available for purposes of replication.)

The interventions were administered immediately after the baseline covariates (shown in the covariate balance tests) were taken. They were administered over the internet embedded in the online survey. The images were embedded on the same page as the outcome measures on the Qualtrics survey.

**E. Results**

* Outcome Measures and Covariates Provide precise definition of all primary and secondary measures and covariates. For indices, provide exact description of how they are formed. For survey items provide exact question wording in an appendix. Please provide a copy of the complete survey questionnaire (in an on-line appendix if it is long). Clearly state which of the outcomes and subgroup analyses were specified prior to the experiment and which were the result of exploratory analysis.

Additional information about the survey questions, wording, and flow is provided in section 2 of this Appendix. We do not use scales in the manuscript, but rather look at the outcome measures individually. In one of our analyses in the Appendix, we scale the outcome measures together to see if our findings are robust (see Figure A6). The main subgroup explored (respondent’s party) was prespecified in our pre-analysis plan. Heterogeneities tested in Section 4 of this Appendix are listed as having been prespecified or not.

* Report sample means and standard deviations for the outcome variables using intent-to treat (ITT) analysis (means for the entire collection of subjects assigned to a group, whether the treatment is successfully delivered or not).

This is the approach that we use in all of our figures. Our results hold with baseline controls preregistered.

* If the experiment uses block randomization with unequal assignment rates, present ITT analysis by block or present overall means using inverse probability weighting.

As noted earlier, our analysis uses block randomization, but with equal assignment rates.

* Note if level of analysis differs from level of randomization and estimate appropriate standard errors.

Our level of analysis is the individual level and so is our randomization.

* If there is attrition, discuss reasons for attrition and examine if attrition is related to pretreatment variables.

Our study is cross-sectional. Survey completion rates are equal across treatment groups (see the next bullet point).

* Report for other missing data (not outcome variables): Frequency or percentages of missing data by group. Methods for addressing missing data (e.g., listwise deletion, imputation methods). For each primary and secondary outcome and for each subgroup, provide a summary of the number of cases deleted from each analysis and rationale for dropping the cases.

Missingness is *very rare* for our outcomes and pre-treatment measures. Hence, we use

listwise deletion.

* For survey experiments: Describe in detail any weighting procedures that are used.

We use the standard CCES weights in our analysis using the CCES (see [here](https://cces.gov.harvard.edu/frequently-asked-questions) for more information on these). We use no weighting for MTurk samples.

**F. Other Information**

* Was the experiment reviewed and approved by an IRB?

The studies in this paper were approved by the Brigham Young University (E18118) and University of Virginia (3698) IRBs.

* If the experimental protocol was registered, where and how can the filing be accessed?

The pre-analysis plan is available on the OSF, see It's Not Me, It’s You: A Survey

Experiment on How Political Ideology Shapes Romantic Preferences (<https://osf.io/vufxe/>).

* What was the source of funding? What was the role of the funders in the analysis of the experiment? Were there any restrictions or arrangements regarding what findings could be published? Any funding sources where conflict of interest might reasonably be an issue?

The National Science Foundation (SES-1657821) provided funding for this project. The funders played no role in the analysis of the experiment. There were no restrictions or arrangements regarding what findings could be published. There were no funding sources where conflict of interest might reasonably be an issue.

**3. Other Details on the Survey**

We began each of our surveys with a series of demographic questions including age, gender, education, ethnicity, party, ideology, employment status, sexual orientation, whether someone is transgender, religion, political interest, and family income. To ensure our three surveys were parallel, we used the question wording from the CCES as a template for our MTurk questionnaires.

Our dating module started immediately thereafter with the question: “When dating, are you more inclined towards men or women?”. This was a forced-response question which allowed us to sort people into groups so that they would see someone which gender best describes them. We then prompted respondents with this statement, “On the following page, you will be presented with a hypothetical dating profile, similar to one you may encounter on a dating website or app. Please answer the following questions as though you are single and interested in viewing dating profiles.”[[8]](#footnote-8) After reading this short prompt, respondents were randomly shown one of our three hypothetical dating profiles (see below for these profile images). The dating profiles contained an image of the individual, just like standard online dating profiles (and previous experiments of this type). We intentionally held the image constant in our analyses and, in fact, used the image of the male individual to create an image of a female individual. Doing so allows us to hold constant subtle facial features that might vary across individuals/genders. Importantly, the dating profiles held constant complementary/covarying characteristics--like race--that have been shown to be simultaneously activated when one primes political parties (Westwood and Peterson 2019). This allows us to avoid potential compound treatments problems.

The profiles were designed to look like the cover image for a standard dating profile, providing about the same amount of information that someone would find on an online dating platform like Tinder. Like our profiles, Tinder pages usually only contain a few descriptive terms about individuals.[[9]](#footnote-9) Each of the profiles had identical images and biographies except for one key feature: their stated political party. One profile stated “Proud Republican”, one read “Proud Democrat” (though in experiment 4, we omit the “proud modifier”), and the third had no mention of political party.

Below the dating profile, we asked the following questions:

1. How attractive would you rate the person in this dating profile?
   * Continuous scale of 1-100, 1 = “not at all attractive”, 50 = “Neither attractive nor unattractive, and 100 = “Very attractive”.
2. If you were single, would you respond to a message sent from the person in this profile?
   * Seven point scale ranging from “Definitely” to “Definitely not”
3. If you were single, would you be interested in going on a date with the person in this profile?
   * Seven point scale ranging from “Definitely” to “Definitely not”
4. If you were single, could you see yourself in a relationship with the person in this profile?
   * Seven point scale ranging from “Definitely” to “Definitely not”
5. (Asked only on MTurk survey platforms due to CCES question constraints): If you weren’t single, would you set this person up with a close friend?
   * Seven point scale ranging from “Definitely” to “Definitely not”

We chose the first outcome measure (attractiveness) to parallel previous research from Nicholson et al. (2016).[[10]](#footnote-10) To see the scope of potential bias, we chose our other outcome measures to indicate various stages of a relationship beyond physical attractiveness—responding to a message, going on a date, and entering into a relationship. We added our fifth outcome on the MTurk surveys to better measure effects on individuals who may not find the profile personally attractive or desirable to date, but who are overall amicable towards the hypothetical person represented. This provides us with evidence of how any partisan animus may spread through social networks--something not previously explored in research on the topic at hand.

Figure A1: Survey Treatment Profiles



The profile picture was found via Google Images. To minimize bias between the male and female profiles, we put the male picture through an online generator that morphed it into a “female version”. We chose the name Sam because it is unisex and standardized all other information in the biography. In experiment 4, we used the control image (i.e. the last row of images) for all conditions (with the treatment being administered on a previous screen). The change for experiment 5 was to change the politics condition to be liberal, conservative, Trump, Sanders, democrat, republican, or control.

Experiment 3: Demand effects may occur in survey research because respondents identify research objectives and act in accordance with those objectives instead of providing a normal response (Mummolo and Peterson 2018). Ruling out demand effects could be especially important among highly-surveyed and highly-sampled groups like MTurk workers. And it could be important if our respondents intuited our research design was about politics and dating.

Our approach in Experiment #3 uses the exact same approach as Mummolo and Peterson (2018). This approach randomly assigns respondents to either receive the normal pretreatment primer before viewing the profile (see above), or to receive the following slight modification (emphasis added here only):

“On the following page, you will be presented with a hypothetical dating profile, similar to one you may encounter on a dating website or app. *We are researching how stated political party in dating profiles influences perceptions of dateability.* Please answer the following questions as though you are single and interested in viewing dating profiles.”

The idea behind this design was to see if MTurk workers moved in a compliant manner. That is, when we increased individuals’ knowledge in the research question, whether the treatment effects would be larger.

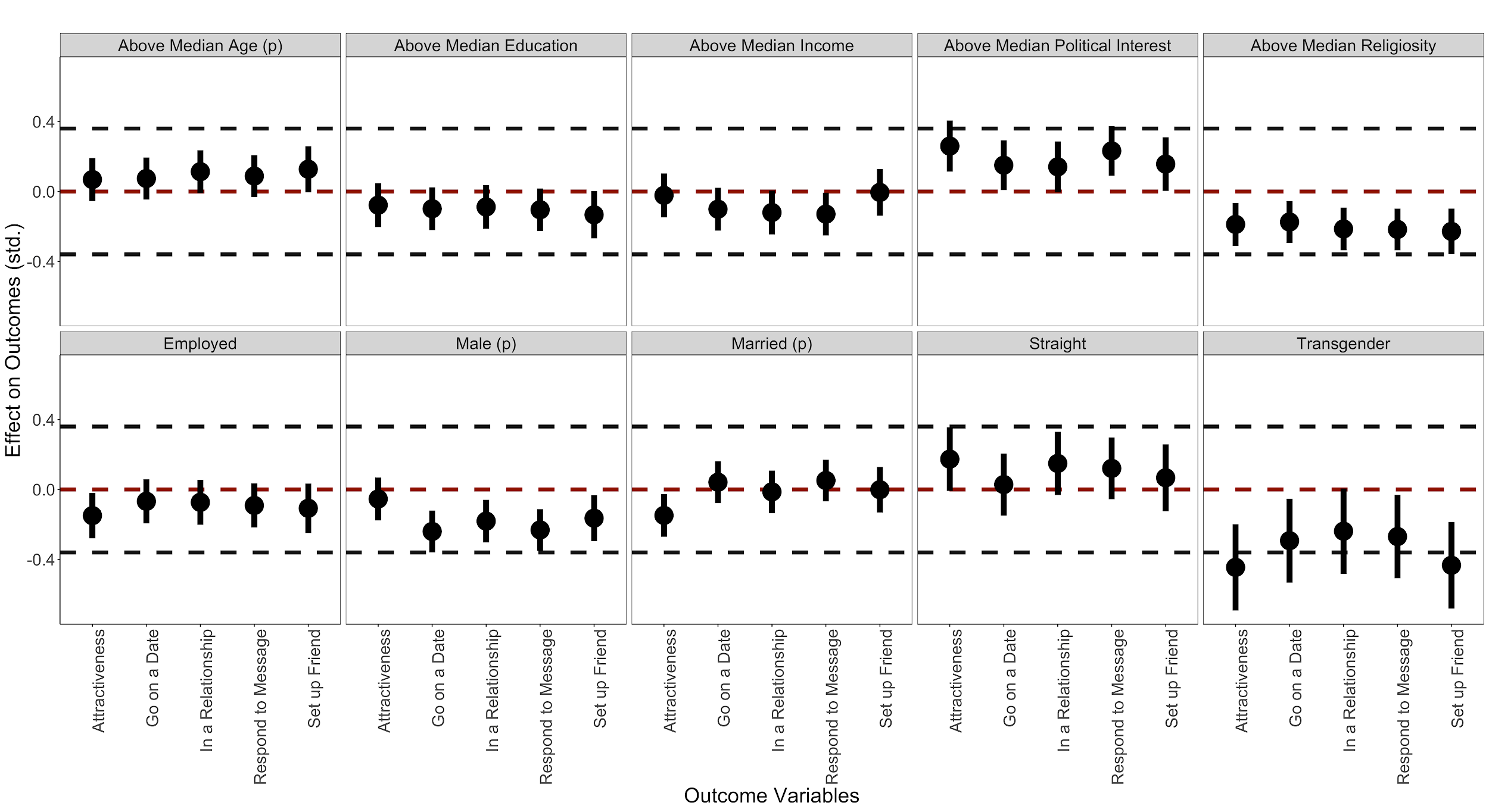
Experiment 4: the randomization structure in this experiment was exactly the same as our previous experiments. However, on an earlier screen--the one that told subjects that they would be viewing dating profiles--we told individuals the political party of the individual on the next screen. Again, this was meant to mimic a situation where someone learns of the politics of the prospective partner indirectly. This was meant to address the argument (explicitly made by those cited in the text) that partisan penalties are due, in large part, to individuals feeling uncomfortable with individuals who wear their politics on their sleeve--who are so engaged in politics that they would post in a (seemingly) non-political environment (i.e. a dating profile). Experiment 4’s design is meant to tone down the possibility of us picking up on this related but distinct channel.

Experiment 5: this final experiment had everything the previous studies had, with the exception of the nature of the treatment. In addition to testing for differences in party out-groups, this tested for ideological and candidate effects. To do so, we blocked on political and gender preferences and then assigned individuals to either view the profile of a conservative, liberal, republican, democrat, Trump supporter, or Sanders supporter. The reason we chose Sanders (instead of Biden--the presumptive Democratic nominee at the time of the study[[11]](#footnote-11)) was partially driven by the timing of our experiments. Recall from the paper that our experiments happened in the Fall of 2018 (Experiments 1 and 2), Spring of 2019 (Experiments 3 and 4), and the Spring of 2020 (Experiment 5). By the time that reviewers (at other journals) requested that we test party, ideology, and candidate effects at the same time, we had seen the results from studies 1-4, which showed a substantially smaller effect than what Nicholson et al. (2016) had shown. To give our experiment the best chance of replicating Nicholson et al.’s large effect, we chose the two candidates that had the largest ideological space between them. Even in spite of this, however, we observed an effect that is much smaller than that observed between Romney and Obama--candidates much closer on the ideological spectrum than Sanders and Trump.

**4. Testing for Treatment Effect Heterogeneity**

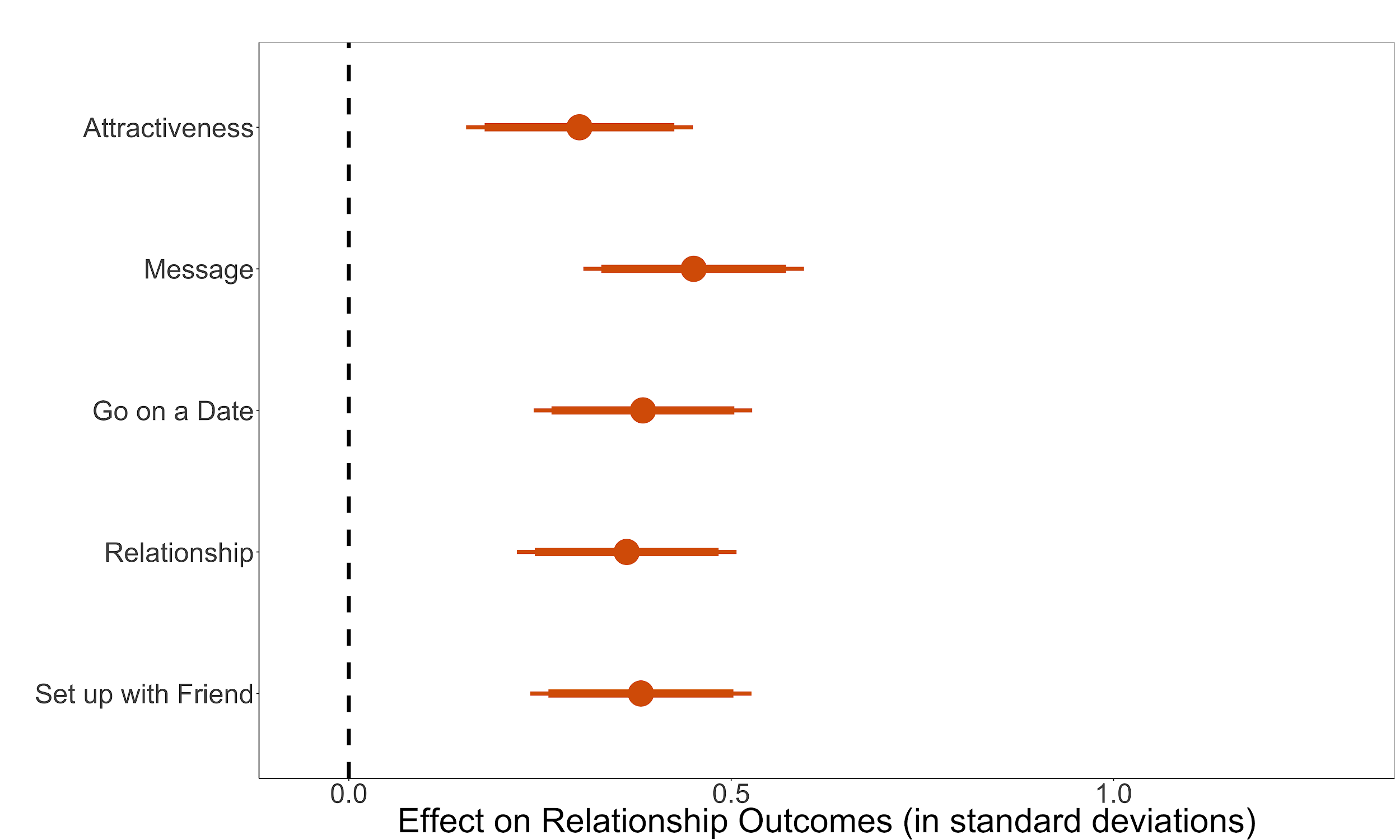
Figure A2 displays our results broken by various pretreatment characteristics. (We use the pooled sample to maximize statistical power given that many tests for heterogeneity are woefully underpowered; see Imai and Ratkovic 2013; Gelman 2015.) As can be seen, across our three preregistered tests for treatment heterogeneity (denoted with a (p)), there are generally minimal levels of effect heterogeneity. Older and young individuals and married and unmarried individuals see similar out party penalties (the one exception is the fact that married individuals’ attractiveness effect is smaller than for non-married individuals, but for the other outcomes the effect is the same). Interestingly, however, on one of our pretreatment variables--gender--we do see some significant heterogeneity. Males seem to punish out partisans less than females (with the exception of the attractiveness outcome). Across other (non-preregistered) pretreatment variables, there appears to be minimal heterogeneity by educational attainment, income, employment status, and straight/gay. Interestingly, there is some heterogeneity by political interest (higher political interest people penalize out partisans more), religiosity (higher religiosity people penalize out partisans less), and transgender individuals (who also punish outpartisans less). We should note that in some instances (e.g. Transgender), however, our tests for effect heterogeneity are underpowered--having very wide confidence intervals.[[12]](#footnote-12) And some of these tests are not preregistered, so they should be treated as such.

Figure A2: Treatment Effect Heterogeneities (Pooled Sample)

****

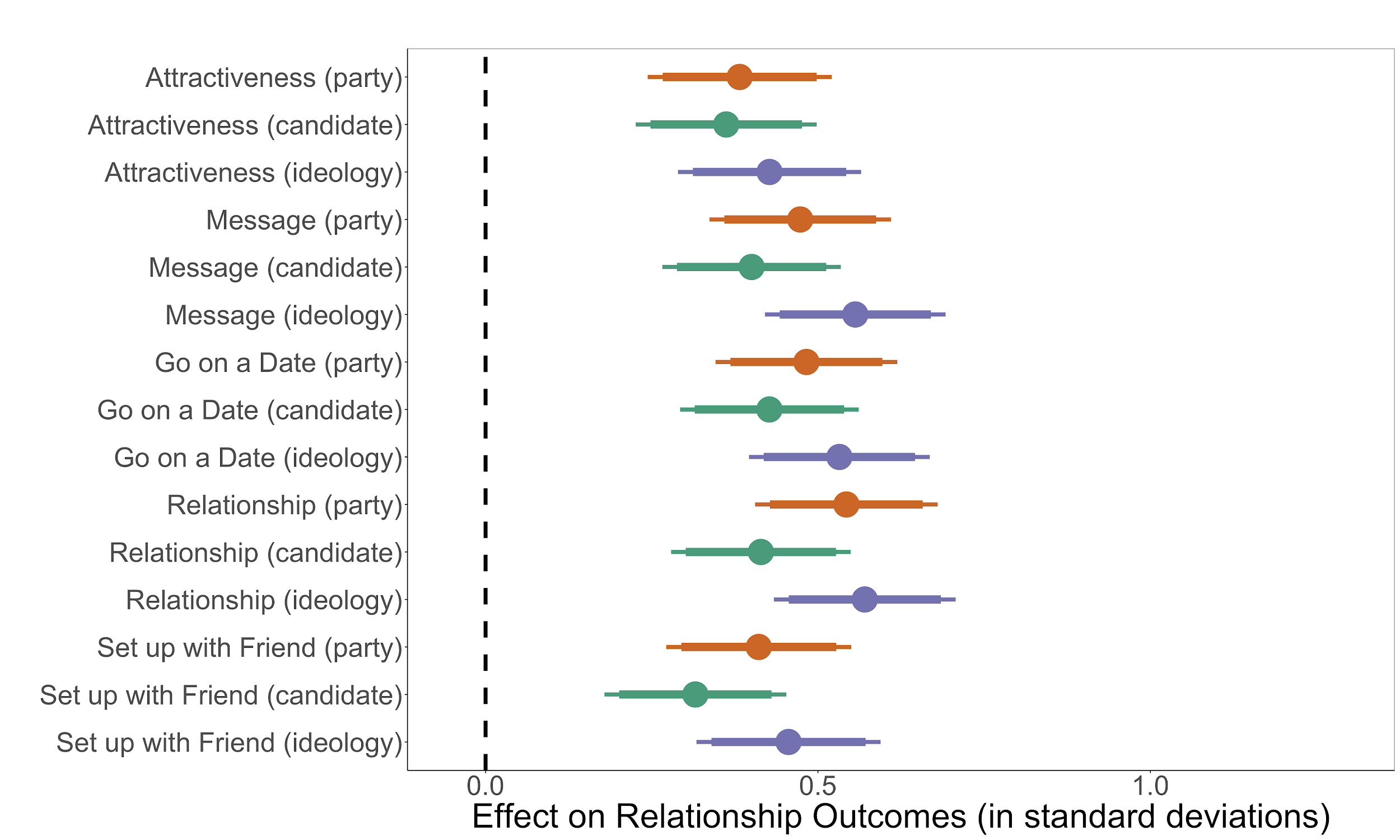
Note: figure displays the effect of being in a same party to the dating profile on the outcomes listed on the x-axis (outcomes are standardized in the pooled sample). Facets are pre-treatment measures. Preregistered heterogeneity tests are denoted with a (p). All effects are the interaction between the pre-treatment covariate and the same party treatment. Effects are plotted as points, with corresponding 90% (narrow line) and 95% (wider line) confidence interval. The middle reference line is the test against the null hypothesis of zero. The two outer reference lines are to the default equivalence testing values suggested by Hartman and Hidalgo (2018). To maximize statistical power, the Figure displays the results from the pooled sample. We include study fixed effects as controls. N’s range from 4052-4557.

**5. Additional Results Not Included in the Paper**

Figure A3: Full Set of Effect Estimates from Experiment 4 

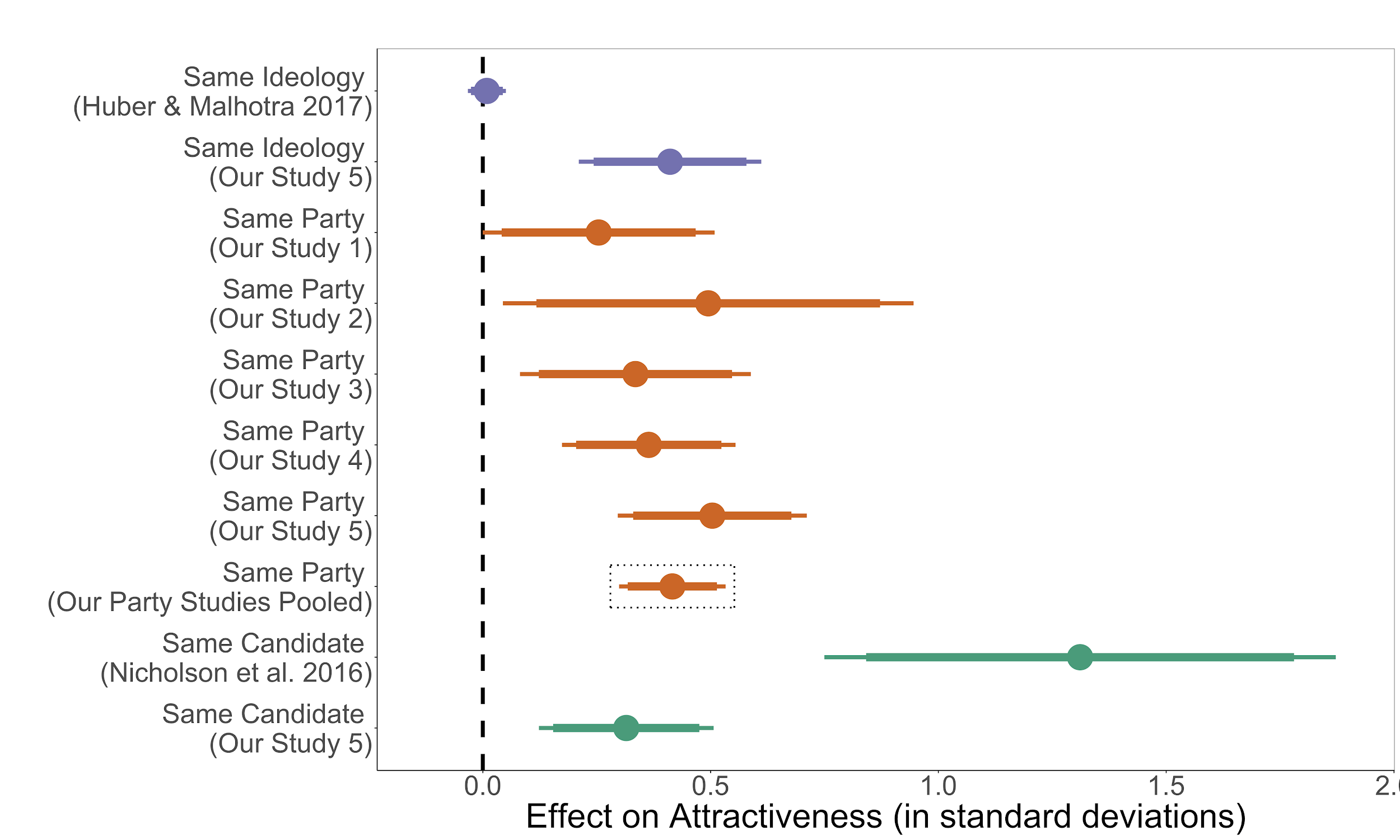
Note: figure displays the effect of being in a similar group on the outcomes displayed. Effects are plotted as points, with corresponding 90% (narrow line) and 95% (wider line) confidence interval. N=725

Figure A4: Full Set of Effect Estimates from Experiment 5



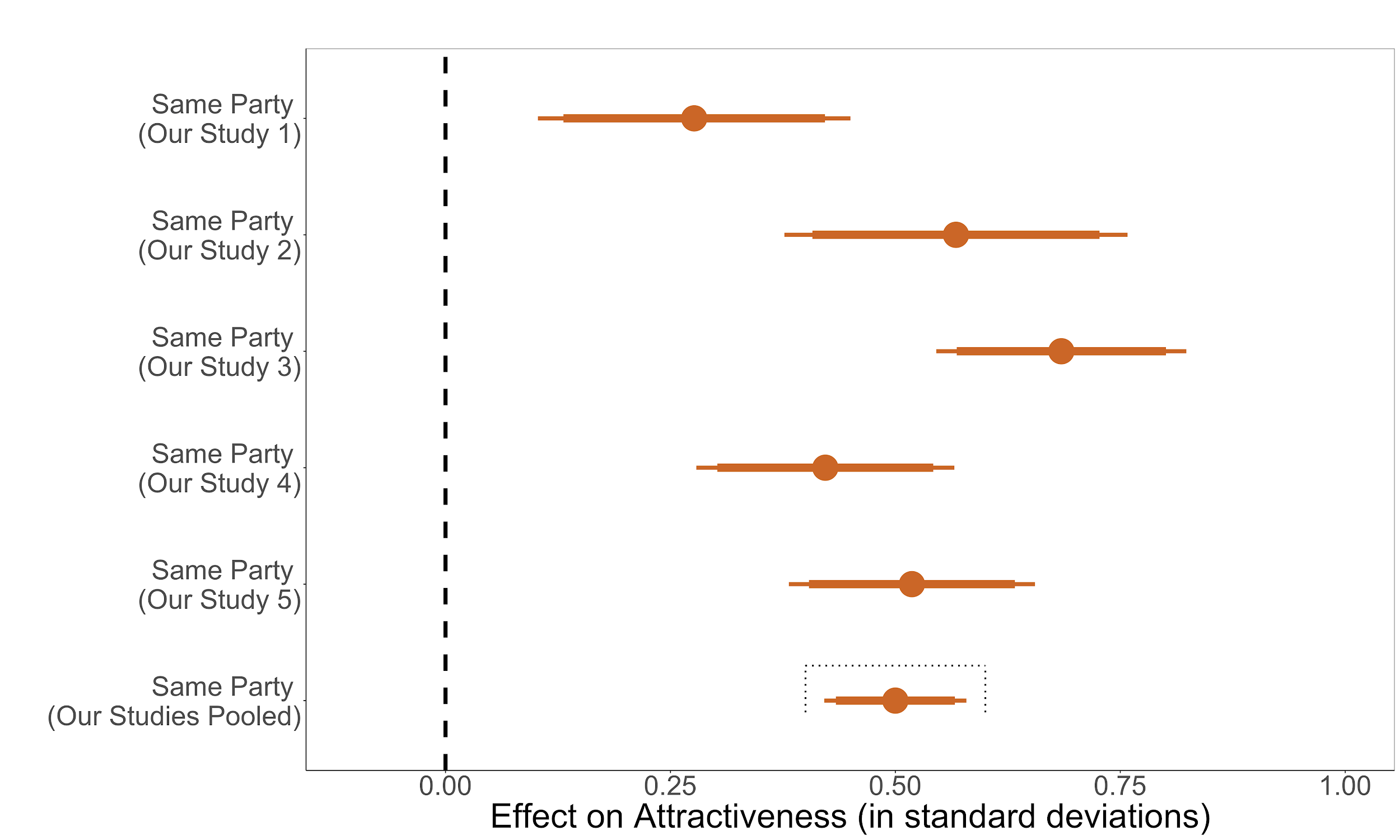
Note: figure displays the effect of being in a similar group on the outcomes displayed. Effects are plotted as points, with corresponding 90% (narrow line) and 95% (wider line) confidence interval. Figures shaded by the treatment. N=1786

Figure A5: Effect Estimates Among 18-35 Year Olds



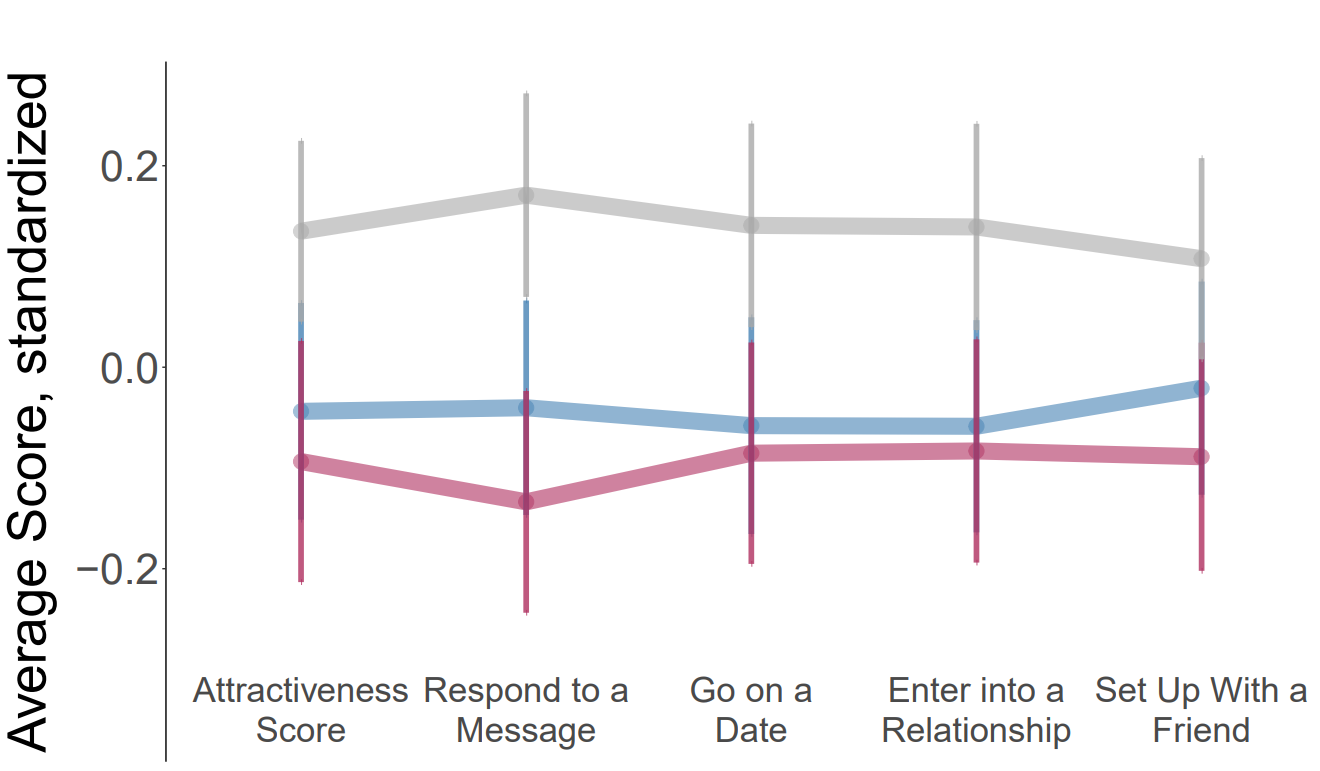
Note: the effect of being in a similar group to the dating profile on perceived levels of attractiveness. Effects are plotted as points, with corresponding 90% (narrow line) and 95% (wider line) confidence interval. Figure benchmarks our four estimates (middle) and a pooled estimate from a meta analysis of these (highlighted with a box) to an ideology treatment from Huber and Malhotra (2017) (on the top) and a candidate treatment from Nicholson et al. (2016) (on the bottom). In our study 3, we use only those in the no demand effects condition to parallel across the treatments. Here, our data and the Nicholson data are restricted to those aged 18-35 to most closely correspond to the Huber and Malhotra (2017) sample. N’s from top to bottom: 9790 (unit of analysis: individual-profile--the rest are individual level); 882, 293, 143, 238, 412, 882, 1968, 114, 882. We replicate the models used in Huber and Malhotra (2017) and Nicholson et al. (2016).

Figure A6: Effect Estimates Scaled Dependent Variable



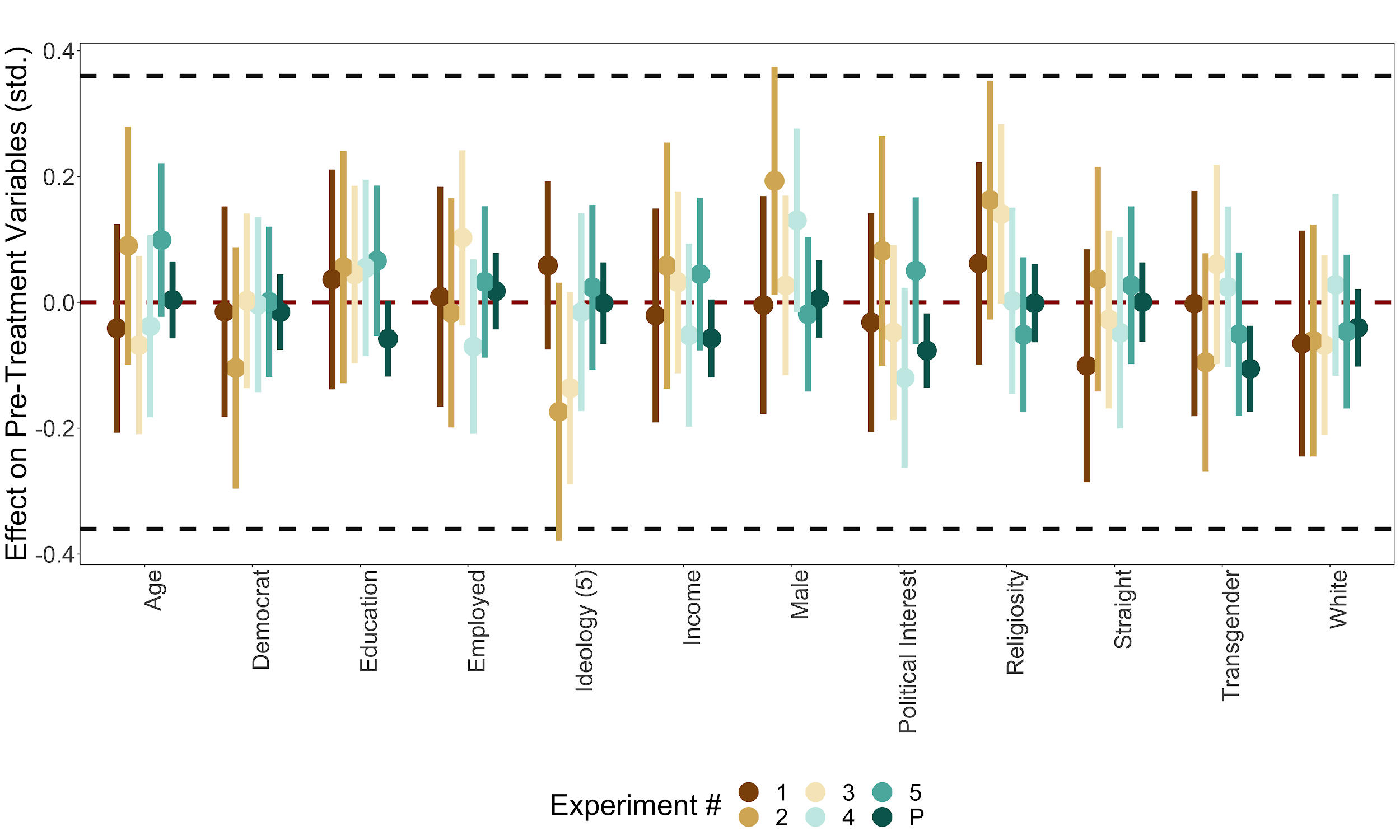
Note: the effect of being in a similar group to the dating profile on a scaled outcome of attractiveness, messaging, going on a date, and being in a relationship. Scale created using factor analysis. Effects are plotted as points, with corresponding 90% (narrow line) and 95% (wider line) confidence interval. In our study 3, we use only those in the no demand effects condition to parallel across the treatments. N’s from top to bottom: 513, 454, 758, 725, 2450.

Figure A7: Overall Findings by Treatment (Study 1)



Note: Figure displays results from Study 1 (MTurk) with all respondents pooled together (N=1,025). The y-axis shows mean levels. The x-axis shows our dependent variables. The randomly assigned treatment groups are broken by color (grey=control profile; blue=Democrat profile; red=Republican profile). When everyone is pooled together, both parties’ partisan profiles are penalized, on average, across our dependent variables. The penalty to Republicans and Democrats is statistically indistinguishable.

Figure A8: Covariate Balance Across Experiments



Note: the effect of being in the same party to the dating profile on pre-treatment variables. Effects are plotted as points, with corresponding 90% (narrow line) and 95% (wider line) confidence interval. Outcomes are standardized. Coefficients are organized by experiment number (p=pooled). Reference lines show tests against a 0 effect and 36% of a standard deviation--Hartman and Hidalgo’s (2018) default equivalence testing value. Only 3/72 coefficients (4.2%) are significant at the 5% level--about what we would expect by chance. Only 2/72 coefficients fail to reject Harmann and Hidalgo’s (2018) default equivalence testing value.

**5. Appendix References**

Abramowitz, Alan I., and Kyle L. Saunders. "Is polarization a myth?." The Journal of Politics 70, no. 2 (2008): 542-555.

Alford, J. R., Hatemi, P. K., Hibbing, J. R., Martin, N. G., & Eaves, L. J. (2011). The politics of mate choice. Journal of Politics, 73(02), 362–379.

Alford, John R., Carolyn L. Funk, and John R. Hibbing. "Are political orientations genetically

transmitted?." American political science review 99, no. 2 (2005): 153-167.

Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes. The american voter. University of Chicago Press, 1980.

Cassese, Erin C. "Partisan Dehumanization in American Politics." Political Behavior (2019): 1-22.

Chen, F. F., & Kenrick, D. T. (2002). Repulsion or attraction? Group membership and assumed attitude similarity. Journal of Personality and Social Psychology, 83(1), 111.

Coppock, Alexander, Thomas J. Leeper, and Kevin J. Mullinix. "Generalizability of heterogeneous treatment effect estimates across samples." Proceedings of the National Academy of Sciences 115, no. 49 (2018): 12441-12446.

Diermeier, Daniel, and Christopher Li. "Partisan Affect and Elite Polarization." American Political Science Review 113, no. 1 (2019): 277-281.

Druckman, James N., S. R. Gubitz, Ashley M. Lloyd, and Matthew S. Levendusky. "How Incivility on Partisan Media (De) Polarizes the Electorate." The Journal of Politics 81, no. 1 (2019): 291-295.

Engelhardt, Andrew M., and Stephen M. Utych. "Grand Old (Tailgate) Party? Partisan Discrimination in Apolitical Settings." Political Behavior (2018): 1-21.

Fiorina, Morris P., Samuel A. Abrams, and Jeremy C. Pope. "Polarization in the American public: Misconceptions and misreadings." The Journal of Politics 70, no. 2 (2008): 556-560.

Funk, Carolyn L., Kevin B. Smith, John R. Alford, Matthew V. Hibbing, Nicholas R. Eaton, Robert F. Krueger, Lindon J. Eaves, and John R. Hibbing. "Genetic and environmental transmission of political orientations." Political Psychology 34, no. 6 (2013): 805-819.

Garrett, R. Kelly, Shira Dvir Gvirsman, Benjamin K. Johnson, Yariv Tsfati, Rachel Neo, and Aysenur Dal. "Implications of pro-and counterattitudinal information exposure for affective polarization." Human Communication Research 40, no. 3 (2014): 309-332.

Gelman, Andrew. "The connection between varying treatment effects and the crisis of unreplicable research: A Bayesian perspective." (2015): 632-643.

Gift, Karen, and Thomas Gift. "Does politics influence hiring? Evidence from a randomized experiment." Political Behavior 37, no. 3 (2015): 653-675.

Gimpel, James G., and Iris S. Hui. "Seeking politically compatible neighbors? The role of neighborhood partisan composition in residential sorting." Political Geography 48 (2015): 130-142.

Goggin, Stephen N., John A. Henderson, and Alexander G. Theodoridis. "What goes with red and blue? Mapping partisan and ideological associations in the minds of voters." Political Behavior (2019): 1-29.

Hartman, Erin, and F. Daniel Hidalgo. "An equivalence approach to balance and placebo tests."

American Journal of Political Science 62, no. 4 (2018): 1000-1013.

Hatemi, Peter K., John R. Alford, John R. Hibbing, Nicholas G. Martin, and Lindon J. Eaves. "Is

there a “party” in your genes?." Political Research Quarterly 62, no. 3 (2009): 584-600.

Hatemi, Peter K., Nathan A. Gillespie, Lindon J. Eaves, Brion S. Maher, Bradley T. Webb, Andrew

C. Heath, Sarah E. Medland et al. "A genome-wide analysis of liberal and conservative political attitudes." The Journal of Politics 73, no. 1 (2011): 271-285.

Hatemi, Peter K., John R. Hibbing, Sarah E. Medland, Matthew C. Keller, John R. Alford, Kevin B.

Smith, Nicholas G. Martin, and Lindon J. Eaves. "Not by twins alone: Using the extended

family design to investigate genetic influence on political beliefs." American Journal of

Political Science 54, no. 3 (2010): 798-814.

Henderson, John A., and Alexander G. Theodoridis. "Seeing spots: Partisanship, negativity and the conditional receipt of campaign advertisements." Political Behavior 40, no. 4 (2018): 965-987.

Hersh, Eitan, and Yair Ghitza. "Mixed partisan households and electoral participation in the United States." PloS one 13, no. 10 (2018): e0203997.

Hersh, Eitan. 2016. “What The Yankees-Red Sox Rivalry Can Teach Us About Political Polarization.” FiveThirtyEight, available at: https://fivethirtyeight.com/features/ what-the-yankees-red-sox-rivalry-can-teach-us-about-political-polarization/.

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.

Imai, Kosuke, and Marc Ratkovic. "Estimating treatment effect heterogeneity in randomized program evaluation." The Annals of Applied Statistics 7, no. 1 (2013): 443-470.

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

Iyengar, Shanto, and Masha Krupenkin. "The strengthening of partisan affect." Political Psychology 39 (2018): 201-218.

Iyengar, Shanto, and Masha Krupenkin. 2018. “The Strengthening of Partisan Affect.” Advances in Political Psychology 39 (S1): 201–18.

Iyengar, Shanto, Yphtach Lelkes, Matthew Levendusky, Neil Malhotra, and Sean J. Westwood. "The Origins and Consequences of Affective Polarization in the United States". Annual Review of Political Science. Vol 22 (Publication date May 2019). Forthcoming.

Iyengar, Shanto, and Sean J. Westwood. "Fear and loathing across party lines: New evidence on group polarization." American Journal of Political Science 59, no. 3 (2015): 690-707.

Iyengar, Shanto, Tobias Konitzer, and Kent Tedin. "The home as a political fortress: family agreement in an era of polarization." The Journal of Politics 80, no. 4 (2018): 1326-1338.

Iyengar, Shanto, and Masha Krupenkin. "The strengthening of partisan affect." Political Psychology 39 (2018): 201-218.

Kandler, Christian, Wiebke Bleidorn, and Rainer Riemann. "Left or right? Sources of political

orientation: The roles of genetic factors, cultural transmission, assortative mating, and personality." Journal of personality and social psychology 102, no. 3 (2012): 633.

Klar, Samara, Yanna Krupnikov, John Barry Ryan; Affective Polarization or Partisan Disdain?: Untangling a Dislike for the Opposing Party from a Dislike of Partisanship, Public Opinion Quarterly, Volume 82, Issue 2, 26 June 2018, Pages 379–390.

Klar, Samara and Yanna Krupnikov (2016). Independent politics: how American disdain for parties leads to political inaction. New York: Cambridge University Press.

Klofstad, Casey, Rose McDermott, and Peter Hatemi. 2013. “The Dating Preferences of Liberals and Conservatives.” Political Behavior 35 (3): 519–538.

Klofstad, C. A., McDermott, R., & Hatemi, P. K. (2012). “Do bedroom eyes wear political glasses? The role of politics in human mate attraction.” Evolution and Human Behavior, 33(2), 100–108.

Lau, Richard R., David J. Andersen, Tessa M. Ditonto, Mona S. Kleinberg, and David P. Redlawsk. "Effect of media environment diversity and advertising tone on information search, selective exposure, and affective polarization." Political Behavior 39, no. 1 (2017): 231-255.

Lelkes, Yphtach, Gaurav Sood, and Shanto Iyengar. "The hostile audience: The effect of access to broadband internet on partisan affect." American Journal of Political Science 61, no. 1 (2017): 5-20.

Lelkes, Yphtach. "Policy over party: comparing the effects of candidate ideology and party on affective polarization." Political Science Research and Methods (2019): 1-8.

Levendusky, Matthew S. "Americans, not partisans: Can priming American national identity reduce affective polarization?." The Journal of Politics 80, no. 1 (2018): 59-70.

Levendusky, Matthew, and Neil Malhotra. "Does media coverage of partisan polarization affect political attitudes?." Political Communication 33, no. 2 (2016): 283-301.

Luo, Shanhong, and Eva C. Klohnen. "Assortative mating and marital quality in newlyweds: a couple-centered approach." Journal of personality and social psychology 88, no. 2 (2005): 304.

Martin, Nicholas G., Lindon J. Eaves, Andrew C. Heath, Rosemary Jardine, Lynn M. Feingold, and Hans J. Eysenck. "Transmission of social attitudes." Proceedings of the National Academy of Sciences 83, no. 12 (1986): 4364-4368.

Mason, Lilliana. "A cross-cutting calm: How social sorting drives affective polarization." Public Opinion Quarterly 80, no. S1 (2016): 351-377.

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

Mason, Lilliana. "“I disrespectfully agree”: The differential effects of partisan sorting on social and issue polarization." American Journal of Political Science 59, no. 1 (2015): 128-145.

McConnell, Christopher, Yotam Margalit, Neil Malhotra, and Matthew Levendusky. "The economic consequences of partisanship in a polarized era." American Journal of Political Science 62, no. 1 (2018): 5-18.

McDermott, Rose, Dustin Tingley, and Peter K. Hatemi. "Assortative mating on ideology could

operate through olfactory cues." American Journal of Political Science 58, no. 4 (2014): 997-1005.

Michelitch, Kristin. "Does electoral competition exacerbate interethnic or interpartisan economic discrimination? Evidence from a field experiment in market price bargaining." American Political Science Review 109, no. 1 (2015): 43-61.

Mummolo, Jonathan, and Clayton Nall. "Why partisans do not sort: The constraints on political segregation." The Journal of Politics 79, no. 1 (2017): 45-59.

Nicholson, Stephen P., Chelsea M. Coe, Jason Emory, and Anna V. Song. 2016. “The Politics of Beauty: The Effects of Partisan Bias on Physical Attractiveness.” Political Behavior 38 (4): 883–898.

Oswald, Frederick L., Gregory Mitchell, Hart Blanton, James Jaccard, and Philip E. Tetlock. "Predicting ethnic and racial discrimination: A meta-analysis of IAT criterion studies." Journal of personality and social psychology 105, no. 2 (2013): 171.

Pierpoint, George. “New York abortion law: Why are so many people talking about it?” *BBC News,* Washington. 28 Jan 2019.

Robison, Joshua, and Rachel L. Moskowitz. "The Group Basis of Partisan Affective Polarization." The Journal of Politics 81, no. 3 (2019): 000-000.

Rogowski, Jon C., and Joseph L. Sutherland. "How ideology fuels affective polarization." Political Behavior 38, no. 2 (2016): 485-508.

Rothschild, Jacob E., Adam J. Howat, Richard M. Shafranek, and Ethan C. Busby. "Pigeonholing partisans: Stereotypes of party supporters and partisan polarization." Political Behavior 41, no. 2 (2019): 423-443.

Shafranek, Richard M. "Political considerations in nonpolitical decisions: A conjoint analysis of roommate choice." Political Behavior (Forthcoming): 1-30.

Smith, Kevin B., Douglas R. Oxley, Matthew V. Hibbing, John R. Alford, and John R. Hibbing.

"Linking genetics and political attitudes: Reconceptualizing political ideology." Political Psychology 32, no. 3 (2011): 369-397.

Smith, Kevin, John R. Alford, Peter K. Hatemi, Lindon J. Eaves, Carolyn Funk, and John R.

Hibbing. "Biology, ideology, and epistemology: How do we know political attitudes are inherited and why should we care?." American journal of political science 56, no. 1 (2012): 17-33.

Soave, Robert. “The Media Wildly Mischaracterized That Video of Covington Catholic Students Confronting a Native American Veteran.” *Reason* Blog, 20 Jan 2019.

Watson, David, Eva C. Klohnen, Alex Casillas, Ericka Nus Simms, Jeffrey Haig, and Diane S. Berry.

"Match makers and deal breakers: Analyses of assortative mating in newlywed couples." Journal of personality 72, no. 5 (2004): 1029-1068.

Westwood, Sean, Peterson Erik, and Yphtach Lelkes. 2018. “Are There Still Limits on Partisan Prejudice?” Working Paper URL: <https://www.dartmouth.edu/~seanjwestwood/papers/stillLimits.pdf>

Westwood, Sean, and Erik Peterson. "Compound Political Identity: How Partisan and Racial Identities Overlap and Reinforce." Available at SSRN 3417476 (2019) URL: <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3417476>

1. Iyengar (2018) finds that it is not ideology but partisan identity that matters most. [↑](#footnote-ref-1)
2. In addition, from previous research it remains unclear the extent to which any bias towards political foes is really attributable to emotion (as the name “*affective* polarization” implies) or other individual processes (e.g. policy preferences) that are independent from these forces (Mason 2015). Relatedly, there are skeptics that the public has truly polarized over time (e.g. Fiorina, Abrams, and Pope 2008). However, this assertion has been contested (e.g. Abramowitz and Saunders 2008). Finally, while the literature on affective polarization has done much to answer positive questions, less work has been done on the normative considerations in play for a society increasingly sorting by politics. [↑](#footnote-ref-2)
3. For this reason, some scholars have tried to use IATs (Iyengar and Westwood 2015; Iyengar et al. 2018) to measure partisan animus. However, IATs have also faced strong criticism on whether they identify constructs of interest (e.g. Oswald et al. 2013). [↑](#footnote-ref-3)
4. See “Political Polarization in the American Public” *Pew Report* (2017) [↑](#footnote-ref-4)
5. See Iyengar et al. 2019; Iyengar, Sood, and Lelkes 2012; Garrett et al. 2014; Lelkes, Sood, and Iyengar 2017; Levendusky 2018; Levendusky and Malhotra 2016; Westwood and Lelkes 2018). See also Hersh 2016 for a substantive benchmark in the non-political arena. [↑](#footnote-ref-5)
6. However, recent research has shown that partisan affect is not totally a product of these factors. In an innovative conjoint design, Shafranek (Forthcoming) provides evidence that partisan animus is unlikely to be driven by these intuited factors. Among evaluations of potential roommates, individuals penalize individuals in the political outgroup more than individuals who go to bed at a late hour, who are unclean, and who are LGBTQ+ (to name a few) [↑](#footnote-ref-6)
7. The literature has identified evolutionary, child rearing, and bonding as reasons for seeking mates with similar political values (Alford et al. 2011; Klofstad, McDermott, Hatemi 2012; Klofstad, McDermott, and Hatemi 2013). [↑](#footnote-ref-7)
8. We included this mainly to prime the respondents with the hypothetical nature of our survey over concern that age or marital status would dissuade older individuals from considering the person in the profile eligible for dating. To further address this potential issue, we 1) preregistered to look at differences in treatment effects by relationship status and 2) added a question (#5 below) about setting the person up with one of their friends. [↑](#footnote-ref-8)
9. In reviewing Tinder profiles when designing our experiment, we found that it was *very* common for individuals to describe themselves with only a few words. These include their interests (e.g. “interested in hiking”) and/or adjectives that describe them (e.g. “adventurous” or “a good listener”). Only very rarely did individuals write more than a few words. [↑](#footnote-ref-9)
10. McConnell et al. (2018) show that self-reported outcomes benchmark well to real-world behaviors that also capture affective polarization. [↑](#footnote-ref-10)
11. At the time of experiment 5, Biden held a sizable lead in delegates over Sanders. [↑](#footnote-ref-11)
12. To give a sense of the power for these test of heterogeneity tests, all of the 95% confidence intervals range between 23.8%-49.6% of a standard deviation (i.e. this is the overall width of the 95% confidence intervals). Outside of the transgender effects, confidence intervals from 23.8 to 38.3% of a standard deviation. [↑](#footnote-ref-12)