

Supplementary Information for “(Mis-)Estimating Affective Polarization”

Table of Contents

SI1 – Details of three wave survey	1
SI2 – Full question wordings for measures.....	3
SI3 – Perceptions and actual levels for Democrats and Republicans	6
SI4 – Manipulation check and benchmarking	7
SI5 – Pure Independents	8
SI6 – Differences in in-party and out-party ratings	9
SI7 – Results for each type of affective polarization measure.....	10
SI8 – Reanalyzing Table 2’s model with control variables.	11
SI9 – Condition-by-condition comparisons	12
SI10 – In-party ratings	14
SI11 – OLS model for figure 2	15
SI12 – Effect Sizes.....	16

Supplementary Information 1: Details of Three Wave Survey

The survey was conducted using Bovitz Inc. (<http://bovitzinc.com/index.php>). They provide an online panel of approximately one million respondents recruited through random digit dialing and empanelment of those with internet access. As with most internet survey samples, respondents participate in multiple surveys over time and receive compensation for their participation.

The survey took place over three waves. All participants who completed the first wave were invited to participate in the other two waves (i.e., they could participate in the third wave even if they skipped the second wave). In the first wave (N=5,191), we asked participants about their demographics, political positions, and political engagement. The second wave (N=4,076) contains our experiment as we asked participants the measures in our main analyses. The third wave (N=4,048) contains our questions about our perceptions.

All waves took place during the summer of 2019, with at least 1 week in between each wave (the break was to ensure that there were no spillover effects across waves). Wave 1 took place from July 9, 2019 to July 17, 2019, wave 2 took place from July 16, 2019 to July 25, 2019, and wave 3 took place from July 26, 2019 to August 2, 2019. Importantly, respondents were only invited to do a subsequent wave when at least six days had passed from their completion of the prior wave.

The tables below present demographics based on their wave 1 answers, and compares them to 2018 benchmarks from the U.S. Census Bureau, via the American Community Survey.

Age

Age Category	Our Sample (%)	Census Benchmark
18-24	9.72	12.08
25-34	19.79	17.87
35-50	33.74	24.54
51-65	25.02	24.88
Over 65	11.74	20.65

Gender Identity

Gender Identity	Our Sample (%)	Census Benchmark
Female	50.16	50.8
Male	48.88	49.2
Transgender/None	< 1	-- ¹

¹ The U.S. Census Bureau does not currently ask about transgender identity, so there is no government-provided benchmark for that quantity. Flores et al. (2016) estimate that less than 1 percent of Americans identify as transgender, consistent with our estimates here; see <http://bit.ly/2Nj5DZE> for more details.

Primary Racial Group

Primary Race	Our Sample (%)	Census Benchmark
Caucasian (White)	69.3	72.2
African-American	14.55	12.7
Hispanic or Latino	9.2	18.3
Asian-American	4.01	5.6
Native American	< 1	< 1
Other	2.06	5

Annual Family Income before Taxes

Income Category	Our Sample (%)	Census Benchmark (%) ²
\$30,000 or less	29.57	29.4
\$30,000 - \$69,999	37.39	30.3
\$70,000 - \$99,999	16.58	12.5
\$100,000 - \$200,000	14.36	20.9
Above \$200,000	2.10	6.9

Education Level

Educational Attainment	Our Sample (%)	Census Benchmark (%)
Did not complete high school	2.98	12
High school graduate	21.73	27.1
Associates Degree/Some College	42.07	28.9
Bachelor's Degree	24.35	19.7
Advanced Degree	8.87	12.3

Across categories, our sample closely matches the Census benchmarks. Our biggest discrepancies are that (1) we under-estimate senior citizens, (2) we we under-estimate the least well-educated (and over-estimate those with some college or a bachelor's degree), and (3) under-estimate the top quarter of the income distribute. These are well-known limitations of any survey sampling procedure, not just our own—problems #1 and #2 are linked in that those populations are not online, and those with high incomes are also typically under-represented across all survey modes.

The other significant gap is that we under-estimate the fraction of the population that is Hispanic or Latino, but this is in part a methodological difference. The Census asks about ethnicity (Hispanic/Latino) separately from race, whereas we combine them into one question. As a result, our estimates for Hispanic/Latino citizens are measuring a different construct from the Census benchmark.

Overall, however, our sample closely matches the Census benchmarks across these different categories.

² The Census categories for income are slightly different than the ones we use. They record income as: \$34,999 or below, \$35,00 - \$74,999, \$75,000 - \$99,999, \$100,000 - \$199,999, and \$200,0000 or greater.

Supplementary Information 2: Full question wordings for measures

Participants read the following introduction prior to answering the affective polarization questions. "We are next going to ask you a set of questions about ordinary people (e.g., voters) who are [Republicans and Democrats / Democrats and Republicans]. Please take your time, and do your best to answer the questions about these people."

The participants were then asked the following questions. Where the word "[CONDITION]" currently is placed, the participants saw one of the following options depending on which treatment group they were placed in.

1. [Republicans/Democrats]
2. [Republicans/Democrats] who rarely talk about politics.
3. [Republicans/Democrats] who occasionally talk about politics.
4. [Republicans/Democrats] who frequently talk about politics.
5. Moderate [Republicans/Democrats]
6. Moderate [Republicans/Democrats] who rarely talk about politics.
7. Moderate [Republicans/Democrats] who occasionally talk about politics.
8. Moderate [Republicans/Democrats] who occasionally talk about politics.
9. [Conservative Republicans/Liberal Democrats]
10. [Conservative Republicans/Liberal Democrats] who rarely talk about politics.
11. [Conservative Republicans/Liberal Democrats] who occasionally talk about politics.
12. [Conservative Republicans/Liberal Democrats] who frequently talk about politics.

Feeling Thermometer

We'd like you to rate how you feel towards [CONDITION] on a scale of 0 to 100, which we call a "feeling thermometer." On this feeling thermometer scale, ratings between 0 and 49 degrees mean that you feel unfavorable and cold (with 0 being the most unfavorable/coldest). Ratings between 51 and 100 degrees mean that you feel favorable and warm (with 100 being the most favorable/warmest). A rating of 50 means you have no feelings one way or the other. How would you rate your feeling toward these groups? Remember we are asking you to rate ordinary people (e.g., voters) and *not* elected officials or candidates.

Trait Questions

We'd like to know more about what you think about [CONDITION]. Below, we've given a list of words that some people might use to describe them. For each item, please indicate how well you think it applies to [CONDITION]: not at all well; not too well; somewhat well; very well; or extremely well.

Terms: Patriotic, Intelligent, Honest, Open-minded, Generous, Hypocritical, Selfish, Mean
Response Options: Not at all well, Not too well, Somewhat well, Very well, Extremely well

Trust

How much of the time do you think you can trust [CONDITION] to do what is right for the country?

Response Options: Almost never, Once in a while, About half the time, Most of the time, Almost always

Social Distance

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

Response Options: Not at all comfortable, not too comfortable, somewhat comfortable, extremely comfortable.

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

Response Options: Not at all comfortable, not too comfortable, somewhat comfortable, extremely comfortable.

Suppose a son or daughter of yours was getting married. How would you feel if he or she married someone who is a [CONDITION]?

Response Options: Not all all upset, Not too upset, Somewhat upset, Extremely upset

Perceptions of Out-Partisans (Condition 13)

To measure the perceptions of out-party members, the following questions were asked.

Which point on the scale below, best politically describes the typical [Republican/Democrat]

Response Options: Liberal, Moderate, Conservative

How often do you think [Republicans/Democrats] talk about politics?

Response Options: Rarely, Occasionally, Frequently

Pre-Test

We pre-tested the words in our treatments to ensure that the participants viewed the words as we hoped they would. The pre-test was conducted on Amazon's Mechanical Turk (N=660).

Pre-test participants were first asked about the frequency words—participants were randomly assigned to one word, as follows.

“Imagine that you were going to have dinner with someone who [rarely/occasionally/some times/frequently] talks about politics. In a 2-hour dinner, what percentage of the time do you think this person would spend talking about politics?”

The percent of time spent discussing politics looked like this:

Means: Rarely: 18%, Occasionally: 32%, Sometimes: 33%, Frequently 52%

Medians: Rarely: 5%, Occasionally: 20%, Sometimes: 22%, Frequently 52%

The less frequent discussion means are skewed by a few people stating they would talk about politics the entire time possibly because they would want to talk about politics.

They were then asked about the ideology measures—participants were randomly assigned to one type of person, as follows.

“Imagine now that you are having dinner with a different person, and this person describes him/herself as a [Democrat who is moderate/Republican who is moderate/ Democrat who is liberal/Republican who is conservative]. Where on the scale below would you think he/she falls in terms of overall ideology?”

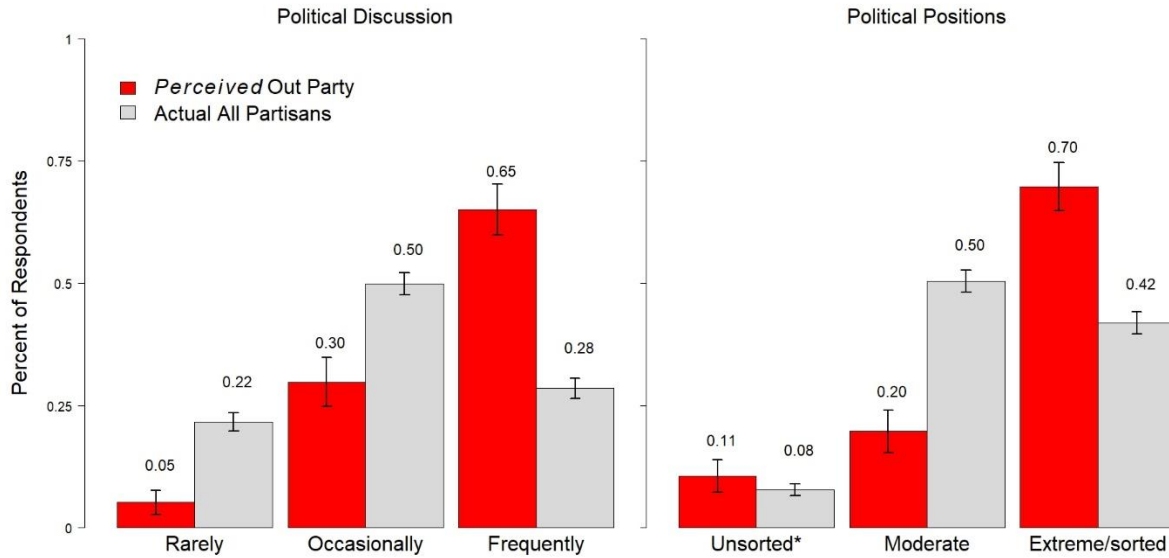
Response Options: (1) Very liberal, (2) Mostly liberal, (3) Somewhat liberal, (4) Moderate, (5) Somewhat conservative, (6) Mostly conservative, (7) Very conservative

The table presents the means for all respondents and then by the party of the respondent.

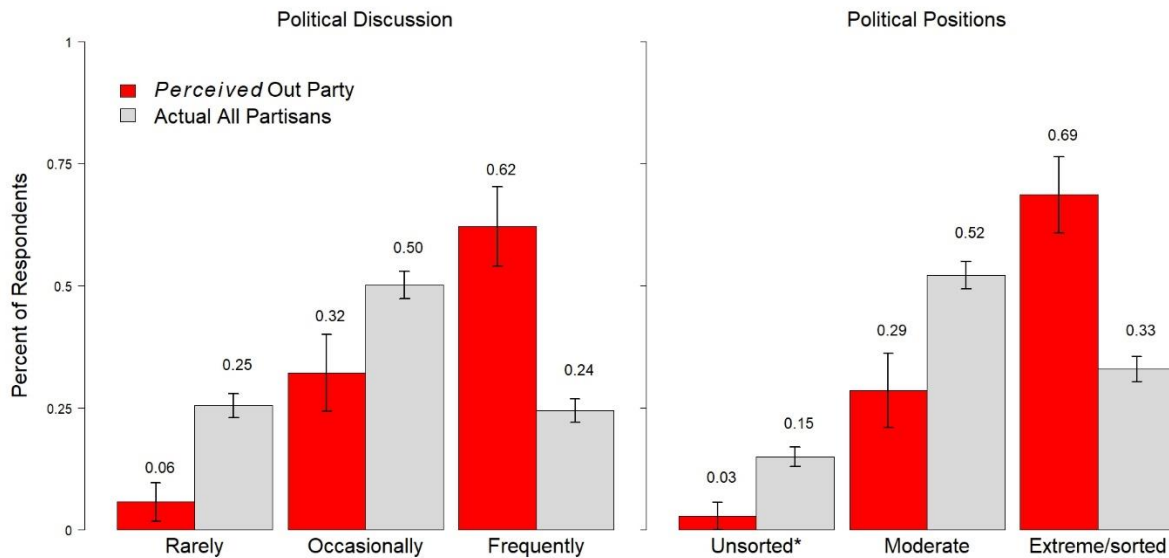
	Liberal Democrat	Moderate Democrat	Moderate Republican	Conservative Republican
All respondents	2.6	3.4	5.0	5.8
Democrats	2.6	3.3	5.0	5.9
Independents	2.2	3.3	5.1	6.0
Republicans	2.9	3.8	4.9	5.6

Supplementary Information 3: Perceptions and actual levels for Democrats and Republicans

A. Democrats (as perceived by Republicans and actual Democratic levels)



B. Republicans (as perceived by Democrats and actual Republican levels)



Supplementary Information 4: Manipulation Check and Benchmarking

We conducted two checks to ensure the validity of our data. The first is a manipulation check to ensure that our participants heeded our instructions to focus on ordinary voters—rather than elites—when rating partisans. Relying on a post-treatment measure that asks participants who they were thinking about when they rated partisans, we see strong evidence that participants were focusing on ordinary voters, suggesting that our results can speak directly to patterns of affective polarization in the electorate. Specifically, pooling conditions 1 and 12, we find that 88% of participants report that they thought of voters when rating the out-group and 89% report that they thought of voters when rating the in-group. The correlation between in-group and out-group categorizations is .82, suggesting most participants were keeping the same categories in mind as they rated them. Given that condition 13 is somewhat different from the other conditions, we consider it independently and again find that clearly respondents were thinking of voters rather than elites: 86% report that they were thinking of voters when rating the out-group and 85% report the same for the in-group; the correlation is again high, at .76. We also conducted a multinomial logit to consider whether categorizations differed significantly by condition; we find no evidence that individuals in a particular condition were more or less likely to categorize the targets differently.

Our second check focuses on benchmarking. Since our goal is to offer a re-interpretation of extant data collections, it is important that the patterns in condition 1—the condition that reflects traditional measurement practices—are similar to existing data. We compare the condition 1 ratings to Druckman and Levendusky (2019), which include similar measures. Our data reflect comparable levels of ratings (and also match other work on particular measures which cohered with Druckman and Levendusky's ratings). Specifically, Druckman and Levendusky (2019) report the following means for out-party *voter* conditions: thermometer: 28.79; traits: 2.33; trust: 1.89; and social distance: 3.22. The means in our condition 1 are thermometer: 30.29 (SD = 24.04; N=456); traits: 2.49 (SD = 0.81; N=452); trust: 1.94 (SD = 0.89; N=454); and social distance: 2.99 (SD = 0.76; N=454). They are thus similar albeit it a bit higher for the thermometer, traits, and trust.

Supplementary Information 5: Pure Independents

	Coef.	Std. Err.
<i>Talk Conditions</i>		
Rarely	0.051	0.018
Occasionally	-0.023	0.018
Frequently	-0.034	0.018
<i>Ideology Conditions</i>		
Moderate	0.016	0.017
Extreme	-0.005	0.015
Constant	0.502	0.013
N	951	
R ²	0.036	

OLS Model. Pure independents rated both parties and the level of analysis is the participant-party—that is, there are 2 cases for each participant. Standard errors are adjusted for 478 participants. Dependent variable is the mean rating for all affective polarization measure for each party. The variable ranges from 0 to 1 with larger values indicating greater positive views of the party's paritsans.

Supplementary Information 6: Differences in In-Party and Out-Party Ratings

	Coef.	Std. Err.
<i>Talk Conditions</i>		
Rarely	-0.128	0.012
Occasionally	-0.048	0.012
Frequently	-0.023	0.012
<i>Ideology Conditions</i>		
Moderate	-0.028	0.011
Extreme	0.011	0.011
Constant	0.285	0.009
N	2,871	
R ²	0.044	

OLS Model. Dependent variable is mean the difference between in-party and out-party ratings for all of the affective polarization questions. The variable can range from -1 to 1 with positive values indicating greater in-group preference.

Supplementary Information 7: Results for each type of affective polarization measure

	Feeling Thermometers		Traits		Trust		Social Distance	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
<i>Talk Conditions</i>								
Rarely	0.186	0.012	0.108	0.010	0.122	0.012	0.043	0.013
Occasionally	0.095	0.012	0.034	0.010	0.030	0.012	-0.047	0.013
Frequently	0.019	0.012	-0.004	0.010	-0.017	0.012	-0.098	0.013
<i>Ideology Conditions</i>								
Moderate	0.067	0.011	0.037	0.009	0.048	0.011	-0.004	0.011
Extreme	-0.019	0.011	-0.006	0.009	-0.006	0.011	-0.028	0.011
Constant	0.316	0.009	0.370	0.008	0.238	0.009	0.631	0.010
N	2,954		2907		2928		2925	
R ²	0.102		0.058		0.056		0.042	

All variables are coded 0-1. Traits and social distance are means of the answers to all questions of that type. In all cases, larger values indicate less out-group animus. All models are O.L.S. Models.

Supplementary Information 8: Reanalyzing Table 2's model with control variables.

	Coef.	Std. Err.
<i>Talk Conditions</i>		
Rarely	0.102	0.009
Occasionally	0.021	0.009
Frequently	-0.025	0.009
<i>Ideology Conditions</i>		
Moderate	0.030	0.008
Extreme	-0.010	0.008
<i>Control Variables</i>		
Age	-0.017	0.012
Woman	-0.003	0.007
White	0.025	0.014
Black	-0.008	0.016
Hispanic	0.017	0.017
Education	-0.034	0.023
Income	0.058	0.013
Partisan Strength	-0.063	0.008
Constant	0.445	0.019
N		2,852
R ²		0.109

OLS model. All variables are coded from 0 to 1.

Supplementary Information 9: Condition-by-condition comparisons

The following table provides the results of difference-of-means (t) tests for each pair of treatment conditions. This looks at the full out-party scale of affective polarization measures. In each cell, the top number reports the difference of means with positive values indicating that the column treatment's mean was greater than the row treatment's mean. The p-values are two-tailed p-values.

	1	2	3	4	5	6	7	8	9	10	11	12
Discussion	None	Rare	Occ.	Freq.	None	Rare	Occ.	Freq.	None	Rare	Occ.	Freq.
Ideology	None	None	None	None	Mod.	Mod.	Mod.	Mod.	Sorted	Sorted	Sorted	Sorted
1	XXXX	-0.08 p=0.00	0.00 p=1.00	0.04 p=0.00	0.00 p=0.87	-0.12 p=0.00	-0.04 p=0.01	0.00 p=.80	0.02 p=0.11	-0.07 p=0.00	0.01 p=0.70	0.05 p=0.00
2		XXXX	0.08 p=0.00	0.13 p=0.00	0.09 p=0.00	-0.04 p=0.01	0.05 p=0.00	0.09 p=0.00	0.11 p=0.00	0.02 p=0.28	0.09 p=0.00	0.13 p=0.00
3			XXXX	0.04 p=0.00	0.00 p=0.89	-0.12 p=0.00	-0.04 p=0.02	0.00 p=.83	0.02 p=0.16	-0.07 p=0.00	0.01 p=0.74	0.05 p=0.00
4				XXXX	-0.04 p=0.01	-0.17 p=0.00	-0.08 p=0.00	-0.04 p=0.01	-0.02 p=0.11	-0.11 p=0.00	-0.04 p=0.01	0.00 p=0.95
5					XXXX	-0.13 p=0.00	-0.04 p=0.02	0.00 p=0.95	0.02 p=0.24	-0.07 p=0.00	0.00 p=0.87	0.04 p=0.01
6						XXXX	0.09 p=0.00	0.13 p=0.00	0.14 p=0.00	0.05 p=0.00	0.13 p=0.00	0.17 p=0.00
7							XXXX	0.04 p=0.01	0.06 p=0.00	-0.03 p=0.04	0.04 p=0.00	0.08 p=0.00
8								XXXX	0.02 p=0.23	-0.07 p=0.00	0.00 p=0.91	0.04 p=0.01

9	XXXX	-0.09 p=0.00	-0.02 p=0.27	0.02 p=0.12
10		XXXX	0.07 p=0.00	0.11 p=0.00
11			XXXX	0.04 p=0.01
12				XXXX

Supplementary Information 10: In-Party Ratings

	Coef.	Std. Err.
<i>Talk Conditions</i>		
Rarely	-0.025	0.007
Occasionally	-0.028	0.007
Frequently	-0.047	0.007
<i>Ideology Conditions</i>		
Moderate	0.003	0.006
Extreme	-0.001	0.006
Constant	0.700	0.005
N	2,896	
R ²	0.015	

OLS Model. Dependent variable is the mean in-party rating for all affective polarization measure. The variable ranges from 0 to 1 with larger values indicating greater positive views of the in-party.

Supplementary Information 11: OLS model for figure 2

	Coef.	Std. Err.
Condition 6 (Rare/Moderate)	0.134	0.014
Condition 12 (Frequently/Sorted)	-0.041	0.015
Constant	0.423	0.008
N	883	
R ²	0.13	

OLS model of full out-party scale of affective polarization measures. Condition 1 (no discussion or ideology information given) is the reference category.

Supplementary Information 12: Effect Sizes

	Change due to treatment	Cohen's d	Cohen's d classification
<i>Thermometer Scores</i>			
No information about frequency of discussion vs. Rarely discuss	19 degrees (p<0.001)	0.77	Medium (Threshold 0.5)
No information about frequency/ideology vs. Rarely/moderate condition	25 degrees (p<0.001)	1.04	Large (Threshold 0.8)
Frequently/sorted vs. Rarely/moderate condition	24 degrees (p<0.001)	0.93	Large (Threshold 0.8)
No information about frequency/ideology vs. Frequently/sorted condition	1.5 degrees (p=0.45)	0.06	Very Small (Threshold 0.02)
<i>Full Index (0 to 1 scale)</i>			
No information about frequency of discussion vs. Rarely discuss	0.10 (p<0.001)	0.58	Medium (Threshold 0.5)
No information about frequency/ideology vs. Rarely/moderate condition	0.13 (p<0.001)	0.76	Medium (Threshold 0.5)
Frequently/sorted vs. Rarely/moderate condition	0.17 (p<0.001)	1.04	Large (Threshold 0.8)
No information about frequency/ideology vs. Frequently/sorted condition	0.06 (p=0.01)	0.22	Small (Threshold 0.2)