Supplementary Information for *Choosing to Avoid? A Conjoint Experimental Study to Understand Selective Exposure and Avoidance on Social Media*

Contents

A. Pilot Tests and Details of Stimuli Generation Process

A.1 Pilot tests overview

We conducted one pilot study to pretest the partisan cues embedded in the source and message attributes of our posts, and a second pilot study to pretest the partisan slants and the perceived emotionality of the images we used as thumbnails in the posts. We recruited 149 people for the first pilot test, and another 34 people for the second pilot test, both on Amazon's Mechanical Turk platform.

A.2 Selection of outlets

The choice of sources (or media outlets) was informed by the first pilot test which surveyed participants about their awareness and perceived political slant of various popular American media outlets. We used a simple questionnaire survey item to ask subjects their perceived partisan slants of each outlet and assess whether they were aware of the outlet or not. The exact wording of the question was, "Please select the option that, in your opinion correctly categorizes each of the following media outlets as Strongly Democratic, Democratic, Neutral, Republican, or Strongly Republican. If you haven't heard of an outlet or are not sure of their political slant, select the "Don't know/Can't say" option", (see Table A.1).

The five partisan categories were captured using a 5-point Likert scale ("Strongly Democrat" being 1, and "Strongly Republican" being 5). We discarded those outlets that were unknown to a substantial fraction of the subjects, and those whose perceived slants showed significant variance, depending on the partisan identity of the subject. In other words, we shortlisted only those outlets that most of the subjects in the pilot study were aware of, and those whose perceived partisan slants were universally agreed upon irrespective of the subjects' partisan identity. The final set (or levels) of outlets (sources) included Fox News, Associated Press (AP), USA Today, The New York Times, and CNN. The New York Times and CNN were universally perceived by our subjects to be left leaning; AP and USA Today were universally perceived to be of a centrist (neutral) disposition. Fox News was the only outlet on the right that enough subjects were acquainted with that was universally perceived to be right leaning. Our original pool of outlets included several other typically right leaning, centrist, and left leaning outlets (for example, The Wall Street Journal and The Washington Times on the right; Reuters in the center, and The Washington Post and MSNBC on the left), but these did not qualify for our final set owing to poor awareness levels and/or high variance in perceived political slant among our subjects). Because we had only one right leaning outlet, and two each of neutral and left-leaning outlets, we duplicated Fox News in our pool, in order to ensure a uniform probability of selection in each draw.

Table A.1 Perceived partisanship and awareness of outlets

Outlet	Unknown Percentage	Mean Ideology
HuffPost	17.12	2.41
MSNBC	13.51	2.51
TIME	18.02	2.59
CNN	**5.41**	**2.60**
The New York Times	**11.71**	**2.62**
The Washington Post	18.02	2.69
ABC	13.51	2.70
NBC	9.01	2.70
The New Yorker	23.42	2.74
Slate	37.84	2.75
Daily Kos	51.35	2.78
CNBC	16.22	2.81
CBS	17.12	2.88
New York Post	19.82	2.93
Politico	39.64	2.94
USA Today	**15.32**	**2.95**
Atlantic	39.64	2.99
AP	**21.62**	**2.99**
Hill	45.05	3.10
Bloomberg	27.03	3.12
Reuters	28.83	3.22

Forbes	18.02	3.29
The Washington Times	26.13	3.30
The Economist	27.03	3.41
The Wall Street Journal	14.41	3.42
National Review	51.35	3.46
The Drudge Report	27.93	3.83
The Federalist	35.14	3.97
Fox News	**4.50**	**4.16**
The American Conservative	18.02	4.19
Breitbart	23.42	4.28

A.3 Generation of headlines (messages)

We aimed to generate 6 messages for each partisan level per topic. Therefore, we first came up with 21 messages for each topic, in which seven of them were supposed to lean towards Republicans, seven of them were supposed to lean towards Democrats, and remaining seven were supposed to be neutral. We intentionally kept one extra message for each category as a backup, in case any of the first six did not pass our pilot pretests. Following pretests similar to those used in the past (Budak, Goel, & Rao, 2016; Knobloch-Westerwick & Meng, 2009), we asked each participant to evaluate the perceived partisan slant of 30 messages randomly drawn from the pool of messages, using the following question item: “In the next section of the survey you have to assess the political slant of 3 news headlines as “Democratic”, “Republican”, or “Neutral”. Please make your judgment based on the content of the headlines only, not potential sources where you think they could have appeared or any other cues that you might remember about them. If you are unable to assess the slant adequately, then please choose “Don't know/can't say””.

The partisan slants for the messages were operationalized using a 3-point scale (“Democratic” being 1, “Neutral” being 2, and “Republican” being 3). In our creation of “neutral” messages, we did not make the distinction between messages that are actually neutral (i.e. they take neither a Republican or a Democratic stance on the issue) and messages that are two-sided (i.e. they equally favor / disfavor the Democratic and Republicans stances on the issue. On average, each message was evaluated by 39 participants. To qualify for our main study, we had to test whether the average perceived partisan slant of each message was close to what we expected when we generated the messages. Because the analysis did not yield a sufficient number of satisfactorily slanted messages (as perceived by our subjects), we tested another three messages in the second pilot test. Messages that finally passed our pretest were chosen as stimuli in the main experiment. These are displayed in Table A.2 and their perceived partisan slants as evaluated by our subjects are shown in Figure A.1.

Table A.2 Perceived partisanship of messages

Message	Partisanship	Issue
Can the Left and the Right ever agree on immigration?	N	Immigration
Humanity versus security: making sense of the immigration debate	N	Immigration
Going beyond open borders and mass deportation: the problems with the current immigration debate	N	Immigration
The solution to the immigration debate is not a trade-off between a humanitarian ideal and national security	N	Immigration
Working towards an immigration policy that works for both sides	N	Immigration
Hearing the other side in the immigration debate	N	Immigration
Republicans are using hapless migrant children as a political tool	D	Immigration
Why peaceful demonstration along the border might deter Republican policymakers	D	Immigration
Trump wants to deport undocumented immigrants without any judicial process	D	Immigration
Trump administration defies court order to hold migrant kids "indefinitely"	D	Immigration
Child separation? What the ICE should be doing instead	D	Immigration
Why the negative effects of illegal immigration might be overstated	D	Immigration

House Republicans heap praise on Trump's no tolerance policy on illegal immigration	R	Immigration
President Trump's border wall must be built, and as quickly as possible	R	Immigration
Trump administration plans to expedite reunification of migrant children	R	Immigration
Why the number of undocumented immigrants in the US could be significantly more than previously thought	R	Immigration
Why Americans are justified in their anger towards illegal immigrants	R	Immigration
It's unthinkable to advocate for open borders when crimes committed by illegal immigrants are at an all-time high	R	Immigration
America's complex history of gun-control	N	Guns
The future of America's gun control debate	N	Guns
America's love-hate relationship with guns	N	Guns
Gun control: what if the other side is right?	N	Guns
The promise and perils of the Second Amendment	N	Guns
America can never free itself from its gun control debate	N	Guns
Gun deaths: How last month was the worst in US history	D	Guns
Here's how Democrats can fight the NRA lobby	D	Guns
The NRA's uncertain future is the one positive sign on the issue of gun control today	D	Guns
Gun deaths in 2018: The US has gun laws to blame	D	Guns
Why your Republican friend cannot be persuaded on gun control	D	Guns
No country has as many gun deaths as the US and we are not doing anything about it	D	Guns
People, not guns: the same old argument holds true today	R	Guns
Gun control: What the left fails to understand	R	Guns
Democrats and their love for gun control	R	Guns
The Left's attack on guns reveals a deeply rooted bias against American values	R	Guns
Jacksonville shooting: why gun control isn't the answer	R	Guns
Debunking the arguments for gun control	R	Guns

The abortion debate: how the issue has polarized American voters over the years	N	Abortion
For many, the pro-life-pro-choice debate isn't political; it's personal	N	Abortion
Why it's impossible to reconcile Republicans and Democrats on the issue of abortion	N	Abortion
Planned Parenthood in America: what the future holds	N	Abortion
The other side of Planned Parenthood	N	Abortion
Why the debate over abortion feels like a stalemate	N	Abortion
Opinion: Train more people to perform abortions	D	Abortion
The abortion debate shows that Republican are willing to politicize everything, including women's health	D	Abortion
Being pro-choice is about being a decent human and should not be a partisan issue in a civilized country	D	Abortion
Opinion: Religion should never be the reason for overturning Roe v Wade	D	Abortion
Why supporting abortion can be a pro-life position	D	Abortion
Republican senator willing to rethink his stance on abortion	D	Abortion
Opinion: Abortion is antithetical to morality	R	Abortion
Abortion is anti-Christian and undermines the spirit of America	R	Abortion
How Democrats have politicized the immoral issue of abortion	R	Abortion
Why Planned Parenthood ails the left	R	Abortion
Planned Parenthood isn't the answer to women's health	R	Abortion
Planned Parenthood: what the left doesn't understand	R	Abortion
What the future holds for US-Middle East ties	N	Middle East
Unpacking the long and troubling past of US-Middle East relations	N	Middle East
What Saudi Arabia's war in Yemen means for US foreign policy	N	Middle East
What the scrapping of the Iran deal could mean for the middle east	N	Middle East
It doesn't matter who gets elected, the American middle east policy won't change anytime	N	Middle East

US-Middle East relations can make or break the future of America's position in the world	N	Middle East
Trump's Iran policy will likely destabilize the middle east further	D	Middle East
Most of the middle-east problems today can be traced back to the Iraq War	D	Middle East
Why Trump's friendship with Saudi Arabia should scare us all	D	Middle East
How Trump is making the middle east situation worse	D	Middle East
Trump's recklessness in foreign policy could cost us in the long run	D	Middle East
It's short-sighted to blame the Obama administration for the middle east crisis	D	Middle East
Trump's middle east policies could be the best that America has seen in decades	R	Middle East
How the Obama administration set a bad precedent with middle east policy	R	Middle East
What if Trump is right about the Middle East?	R	Middle East
The Iran deal wasn't as successful as Democrats would like to think	R	Middle East
Opinion: How Trump is solving the Middle-East crisis	R	Middle East
The Iran deal was a disaster, and we should be thankful to Trump for scrapping it	R	Middle East
How the minimum wage became a partisan issue in the US	N	Min. Wage
By politicising the minimum wage, we look forget its economic implications	N	Min. Wage
Will America ever be able to settle the minimum age debate?	N	Min. Wage
What the future holds for America's minimum wage debate	N	Min. Wage
Should the minimum wage be raised?	N	Min. Wage
What a $1 increase in minimum wage means for the economy	N	Min. Wage
Raising the minimum wage isn't about politics, it's about valuing human dignity	D	Min. Wage
A compelling reason to raise minimum wage is economics 101	D	Min. Wage
Can we please reward human worth a bit more by raising the minimum wage?	D	Min. Wage

It's not late yet to raise the minimum wage	D	Min. Wage
Republicans doesn't care about workers, their attack on the minimum wage hike is proof	D	Min. Wage
The right's attack on raising the minimum wage shows how little they care about workers	D	Min. Wage
What the Democrats get wrong on minimum wage	R	Min. Wage
Raising the minimum wage won't help the economy, here's why	R	Min. Wage
Minimum Wage: Why the Left will never understand the economics	R	Min. Wage
The minimum wage issue isn't an economic issue for the Democrats, it's political	R	Min. Wage
The minimum wage hike will hurt workers more	R	Min. Wage
Raising minimum wages hasn't worked anywhere, it won't work in the US	R	Min. Wage

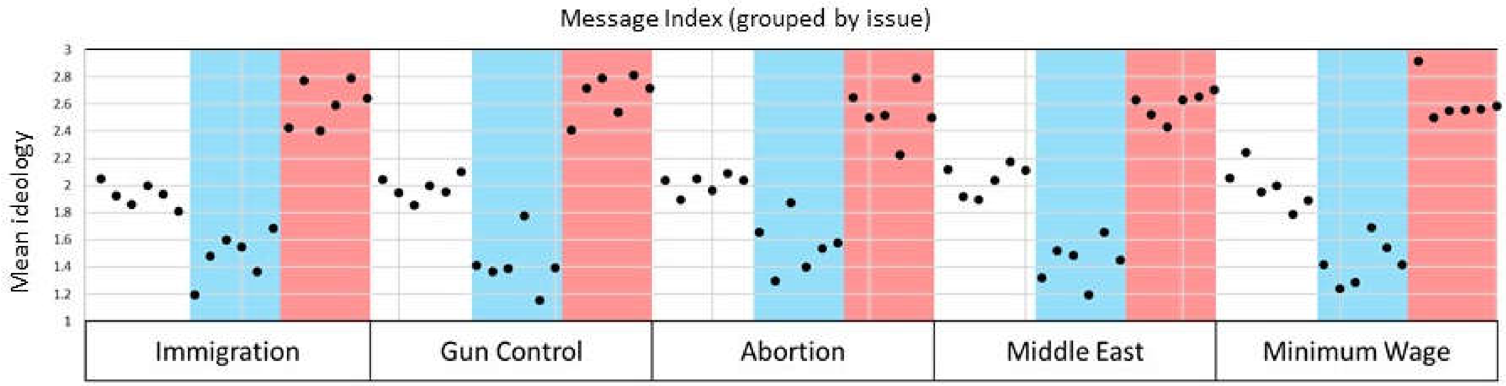


Figure A.1. Perceived partisanship of messages

Note: Points in blue areas indicate that the corresponding messages lean towards Democrats and points in red areas indicate that the corresponding messages lean towards Republicans. Other messages (points in white) were perceived by our pilot subjects to be neutral.

A.4 Image selection process

In the second pilot test, we also pretested the images that we used in our posts. This is because visual cues conveyed by means of images are strong determinants of what catches peoples' eyes which in turn drives their behavior. We first searched for images on the web using our topics as keywords. To decide which images to include in our pretest we considered two dimensions of emotion, arousal (the amount of emotion the image elicits) and valence (positive or negative). We used our judgement to manually select six images for each topic, which we deemed to be neutral, not just along these two emotional dimensions, but also along the dimension of partisanship. To verify our choices, we recruited subjects on Mechanical Turk and asked them to assess each of the images along these three dimensions. The exact question wording of the questions for measuring these are displayed in Table A.3 (the first two questions come from Kurdi, Lozano, & Banaji, 2017).

Table A.3 Question wordings for images

Dimension	Question wording
Arousal	We will ask you to rate a series of pictures in terms of the amount of emotion that they evoke. In other words, we would like to know how much **emotional intensity** the picture creates; whether the picture captures something good or bad doesn't matter. We are interested only in the degree of excitement, energy, or intensity of feeling it represents. Use the right side of the scale to mark you answer if the picture represents something that is strongly emotional. The words that we might use to describe the emotional state that the picture creates are **aroused, alert, activated, charged, or energized**. Use the left side of the scale to mark your answer if the picture represents something that is not strongly emotional. The words that we might use to describe the emotional state that the picture creates are **unaroused, slow, still, deenergized, calm, or peaceful**. Use the middle of the scale to indicate an image that is moderately arousing or halfway through the two extremes. 1 = Strongly emotional; 2 = Emotional; 3 = Moderately emotion; 4 = Unemotional; 5 = Strongly unemotional
Valence	We will ask you to rate a series of pictures in terms of how positive or negative they are.

	If the picture represents something good or positive, please use the right side of the scale to mark your answer. Positive images are those that represent things that make us **happy, satisfied, competent, proud, contented, delighted, and so on.** It doesn't matter what the specific picture is about, as long as it represents something positive or good. If the picture represents something bad or negative, please use the left side of the scale to mark your answer. Negative images those that make us **unhappy, upset, irritated, angry, sad, depressed, and so on**. It doesn't matter what the specific picture is about, as long as it is something negative or bad. Use the middle of the scale to indicate that the picture makes you feel **neutral**, that is, that you think that the picture is neither positive nor negative. 1 = Strongly positive; 2 = Positive; 3 = Neutral; 4 = Negative; 5 = Strongly Negative
Partisanship	In the next section of the survey you have to assess the political slant of 32 images as "Democratic", "Republican", or "Neutral". 1 = Strongly democratic; 2 = Democratic; 3 = Neutral; 4 = Republican; 5 = Strongly Republican; 99 = Don't know/Can't say

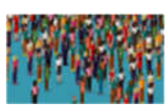
In order to minimize the effects that the image cues could have on news selection or ratings, we selected four images for each topic which had similar perceived partisan slants, valences, and arousal[1]. The images were further randomized (as part of the posts) during assignment to the subjects. While we tried to only keep images that were perceived to be neutral along all three dimensions, it was difficult to find realistic images for particularly polarizing issues like gun control that were not overtly emotion inducing or partisan. Therefore, while images used for gun control posts were not strictly neutral, they did have similar values for valence, arousal, and partisanship. A subject presented with a pair of posts about gun control thus saw two images that were roughly equal along the partisan dimension, thereby canceling out the effect that the partisanship might have had in driving their choice or ratings. Finally, we exercised caution in our design by having enough images for each issue so that any one subject

[1] We first prioritized selecting images with similar levels of partisanship. When there were more than the required number of images with similar partisanship levels, we considered valence and arousal values.

did not see the same image twice times and/or with other messages as part of other posts during the experiment. This precluded the possibility of induced artificiality. Table A.4 lists the images that we finally used in our main study.

Table A.4 Image partisanships, valences, and arousals

Issue	Image	Mean Partisanship	Mean Valence	Mean Arousal
Abortion		3.18	2.76	3.00
		***2.93***	***2.91***	***3.21***
		***2.77***	***2.74***	***3.29***
		***2.77***	***3.12***	***2.88***
		3.37	3.35	3.00
		***3.03***	***2.97***	***3.06***
Gun Control		***3.79***	***2.65***	***3.26***
		***3.97***	***2.65***	***3.21***

		2.31	3.38	3.35
		***4.03***	***2.38***	***2.97***
		***3.71***	***2.15***	***3.15***
		2.34	3.44	3.12
Immigration		2.59	3.56	3.09
		***2.74***	***3.71***	***3.56***
		***3.03***	***3.26***	***3.06***
		***2.76***	***3.35***	***2.85***
		2.66	3.53	2.94
		***3.27***	***3.44***	***2.71***
Middle East		***3.22***	***2.94***	***2.79***

		***3.32***	***3.09***	***2.71***
		3.53	3.18	2.82
		***3.24***	***3.03***	***2.82***
		3.12	3.65	2.62
		***3.00***	***2.85***	***2.65***
Minimum Wage		***3.16***	***3.74***	***2.74***
		2.28	3.26	2.94
		3.44	3.65	2.97
		***3.13***	***3.09***	***2.82***
		***3.10***	***2.71***	***3.35***
		***2.56***	***3.50***	***3.12***

Note: Those italic and bold numbers denoted that correspondent images were chosen.

A.5 Social endorsement generation process

The endorsement level was operationalized using the visible metrics of the number of reactions, the number of shares, and the number of comments that appear at the bottom of a Facebook post. We created several "high" and "low" endorsement cues by randomly generating these numbers (maintaining parity between the number of reactions, the number of comments, and the number of shares for each post). For each high endorsement cue, we generated a number between 2,000 to 3,000. This was used as the number of reactions. We then generated another two numbers between 2,000 to 3,000 and decided to use one-fourth of these numbers as the number of comments and shares respectively. For each low endorsement cue, the range of randomization was 20 to 100. These numeric ranges were informed by observing the number of reactions that posts by popular news pages like CNN and Fox News typically get on Facebook. A similar methodology was used by Messing and Westwood (2014) in their study on selective exposure as well. We generated 10 cues for each level, high and low accompanied with a similar set of emoji (see Table A.5).

Table A.5 Social endorsement cues

	Social endorsement levels and cues	
	2.6K	667 Comments 559 Shares
High	2.1K	651 Comments 561 Shares
	2.9K	661 Comments 559 Shares
	2.4K	735 Comments 726 Shares

	2.1K	525 Comments	625 Shares
	2.8K	668 Comments	705 Shares
	2.3K	650 Comments	748 Shares
	2.4K	545 Comments	638 Shares
	2.9K	504 Comments	556 Shares
	2.4K	607 Comments	644 Shares
Low	65	9 Comments	18 Shares
	88	19 Comments	19 Shares
	85	24 Comments	23 Shares
	72	25 Comments	17 Shares
	84	17 Comments	19 Shares
	71	20 Comments	21 Shares
	83	19 Comments	9 Shares
	70	11 Comments	9 Shares
	96	21 Comments	7 Shares
	92	18 Comments	6 Shares

We generated a sufficient number of these cues so that when we generated the Facebook posts, we could prevent the same participant from seeing identical numbers more than once, in order to preserve the realism of their experience.

A.6 Timestamp generation process

Another aspect of the treatments that we controlled for was the timestamp that appeared in the posts. We consciously kept them roughly similar yet randomized, ranging from "9 hours ago" to "15 hours ago", in order to prevent recency or immediacy of certain posts to confound our findings.

All other aspects of the artificial Facebook posts were identical, and customized to look exactly like real Facebook posts.

A.7 Process of generating stimuli

Since the platform we used to run the experiment, Qualtrics, could not dynamically generate images, we generated 1,500 pairs of Facebook posts for each topic and stored each pair as one image. For each pair, we first randomly generated their outlet partisan slants, source partisan slants, and social endorsement levels. Then, we randomly drew a pair of outlets, messages, social endorsement cues accordingly, from our pretested pools, as well as images and timestamps, and then programmatically stitched them together to form one image for each pair. Our randomization process made sure that each pair did not contain identical messages, social endorsement cues, or images. An example an image pair in displayed in Figure A.2.

Each subject who agreed to participate in our study (and who passed our screening questions) was shown one image pair of posts randomly drawn from the pool of generated image pairs for each topic. We noticed that one pair might be drawn twice if we have more than 1,500 participants, which might bias our estimates of standard errors. Therefore, we conducted robustness tests which preserved the first response and removed duplicated responses for each pair (duplicated responses constituted 15.5% of the total number responses). Our conclusions were identical (see the Robustness section for details).

Figure A.2. An example of pair image

B. Question wordings in the main experiment

B.1 Party identity

1. Generally speaking, do you usually think of yourself as a Democrat, a Republican, an independent, or what?
Democrat;

Republican;

Independent;

Some other party

Display This Question:

If Q1 = Independent

Or Q1 = Some other party

2. If you had to choose, do you think of yourself as closer to the Democratic Party or the Republican Party?
Democratic party;

Republican party;

Neither

Display This Question:

If Q1 = Democrat

3. Would you call yourself a strong Democrat or a not very strong Democrat?
Strong;

Not very strong

Display This Question:

If Q1 = Republican

4. Would you call yourself a strong Republican or a not very strong Republican?
Strong;

Not very strong

B.2 News selection

Forced response:
Which Facebook post are you more likely to click on when browsing Facebook?
Upper;

Lower

Rated response:
What are the chances that, while browsing Facebook, you would select the upper/lower Facebook post to read?
1.very unlikely;

2;

3;

4;

5;

6;

7.very likely

C. Robustness

We present five robustness checks in this section. For the first, we preserved the first response for each pair of posts, and discarded the subsequent responses, as we suspected that these duplicated responses might bias our estimates on standard errors. Thus, we ensured that each pair of posts had been evaluated by exactly one subject. Overall, 13,484, or 84.5% of all responses, were retained for the first robustness check. The results are shown in Figure A.3 (main effects) and Figure A.4 (interaction effects).

Second, we considered the quality of our subject pool. Since we recruited our subjects using an online platform (Amazon Mechanical Turk), it was possible that they provided random responses. Therefore, the second robustness check considered only those responses in which the subjects' forced response (choice) agreed with their corresponding rated responses. In other words, if a subject gave a higher rating to one post, but claimed to choose the post that she/he herself/himself gave a lower rating to, we would deem such a response as invalid. We then reran all analysis on the valid responses only (N = 14,840, 93.0% of the all responses). Results are demonstrated in Figure A.5 (main effects) and Figure A.6 (interaction effects).

Third, we split our responses into two pools, responses from self-reported Democrats (N = 9,464, 59.3% of all responses) and responses from self-reported from Republicans (N = 6,492, 40.7% of all responses). We again reran all analysis on these two pools separately. The results for Democrats are displayed in Figure A.7 (main effects) and Figure A.9 (interaction effects), and the results on Republicans are demonstrated in Figure A.8 (main effects) and Figure A.10 (interaction effects) (see Table A.6 for the summary of results of first three robustness checks).

The next robustness check focused on the first two research questions: we compared the two coefficients by running simulations in our main text. Here, we ran two regressions using linear combinations of our independent variables to isolate the differences between coefficients in our original regression model. In the first regression, to understand whether the absolute effect sizes of out-the party cues were significantly different from the absolute effect sizes of the in-party cues, we tested whether the *differences* between in-party and out-party outlet (δ_1) and between in-party and out-party message (δ_2) were significant or not (see equations 1 and 2)

$$\mathrm{Y} = \beta_0 + \beta_1 Endorsement + \beta_2 InOutlet + (-\beta_2 + \delta_1) OutOutlet + \beta_3 InMessage + (-\beta_3 + \delta_2) OutMessage \tag{1}$$

$$\mathrm{Y} = \beta_0 + \beta_1 Endorsement + \beta_2 (InOutlet - OutOutlet) + \beta_3 (InMessage - OutMessage) + \delta_1 OutOutlet + \delta_2 OutMessage \tag{2}$$

In the second regression, to understand whether the effect of out-party outlet was different from the effect of out-party message we tested whether their difference $\delta_1$was significant or not. Similarly, we tested for the significance of δ_2, or the difference in effect sizes of in-party outlet and message (see equations 3 and 4).

$$\mathrm{Y} = \beta_0 + \beta_1 Endorsement + \beta_2 OutOutlet + (\beta_2 + \delta_1) OutMessage + \beta_3 InOutlet + (\beta_3 + \delta_2) InMessage \tag{3}$$

$$\mathrm{Y} = \beta_0 + \beta_1 Endorsement + \beta_2 (OutOutlet + OutMessage) + \beta_3 (InOutlet + InMessage) + \delta_1 OutMessage + \delta_2 InMessage \tag{4}$$

These regressions were done to formally validate our results using the simulations, and they supported our conclusions in our main text (see Table A.7).

The last robustness check looked at the interaction effects of partisan strength and social media cues. Leeper, Hobolt, and Tilley (2019) suggested the use of an omnibus *F*-test and estimated difference in marginal means to measure the differences between subgroup preferences. We first conducted several *F*-tests, which revealed that preferences between strong partisans and weak partisan on both message cues (forced response: F (3,15950) = 6.341, $p < .001$; rated response: F (3, 15950) = 15.134, $p < .001$) and outlet cues (forced response: F (3,15950) = 5.578, $p < .001$; rated response: F (3, 15950) = 19.011, $p < .001$) were significantly different. With endorsement cues, this test yielded different conclusions for forced response (F (2,15952) = 0.005, $p < .995$) and rated response (F (2,15952) = 10.657, $p < .001$) respectively. Moreover, we calculated differences in conditional marginal means between strong partisans and weak partisans (using a method identical to one shown in Figure 5 in Leeper et al. (2019)). The results are shown in Figure A.11. The figure displays the differences in conditional means between two subgroups, which reveal the interaction effects between partisan strength and cues. The significant effect shown in this plot (which was significantly different from zero) suggests that the difference between two groups is discernible). We found a robust pattern that strong partisans had higher preferences for posts with in-party message cues (forced response: $b = .050$, $se = .014$, $p < .001$; rated response: $b = 0.322$, $se = 0.053$, $p < .001$) and in-party outlet cues (forced response: $b = .043$, $se = .014$, $p = .002$; rated response: $b = 0.350$, $se = 0.053$, $p < .001$), which aligned with our findings in the main text.

In summary, all the conclusions that we arrived at in our main text passed all of our robustness checks.

Table A.6 Results of first three robustness checks

	Deduplicated responses		Valid responses		Democratic responses		Republican responses	
	Forced response	Rated response	Forced response	Rated response	Forced response	Rated response	Forced response	Rated response
H1a: in-party	.010	0.009	.004	0.010	.004	−0.013	.014	0.013
outlet cues	(.011)	(0.041)	(.010)	(0.040)	(.013)	(0.047)	(.016)	(0.064)
H1b: in-party	$.030^{**}$	0.207^{***}	$.027^{*}$	0.221^{***}	$.032^{*}$	0.205^{***}	$.027^{\dagger}$	0.216^{***}
message cues	(.011)	(0.041)	(.011)	(0.040)	(.013)	(0.048)	(.016)	(0.060)
H2a: out-party	$-.102^{***}$	-0.465^{***}	$-.106^{***}$	-0.498^{***}	$-.149^{***}$	-0.661^{***}	$-.031^{*}$	-0.157^{**}
outlet cues	(.011)	(0.044)	(.011)	(0.043)	(.013)	(0.055)	(.015)	(0.057)
H2b: out-party	$-.125^{***}$	-0.425^{***}	$-.127^{***}$	-0.461^{***}	$-.126^{***}$	-0.482^{***}	$-.116^{***}$	-0.323^{***}
message cues	(.011)	(0.044)	(.011)	(0.042)	(.013)	(0.053)	(.015)	(0.060)
RQ1a: outlet cues	*** (out > in)	*** (out > in)	*** (out > in)	*** (out >in)	*** (out > in)	*** (out > in)	*n.s.*	*n.s.*
RQ1b: message cues	*** (out > in)	*** (out > in)	*** (out > in)	*** (out > in)	*** (out > in)	*** (out > in)	*** (out > in)	*n.s.*
RQ2: avoidance	*n.s.*	*n.s.*	*n.s.*	*n.s.*	*n.s.*	* (O > M)	*** (M > O)	* (M > O)
RQ2: selection	*n.s.*	*** (M > O)	*n.s.*	*** (M > O)	*n.s.*	*** (M > O)	*n.s.*	* (M > O)
RQ3: social	−.003	0.032	−.000	0.038	−.008	0.023	.005	0.030
en-dorsement	(.009)	(0.035)	(.008)	(0.033)	(.010)	(0.042)	(.012)	(0.048)
RQ4: social	.001	−0.004	−.003	−0.001	.015	0.096	−.021	-0.230^{*}
en-dorsement	(.017)	(0.070)	(.017)	(0.068)	(.021)	(0.084)	(.025)	(0.098)
RQ4: in-party	$.040^{\dagger}$	$0.147^{\dagger}$	$.049^{*}$	0.197^{*}	$.043^{\dagger}$	0.129	$.062^{\dagger}$	0.208
outlets	(.022)	(0.084)	(.021)	(0.081)	(.026)	(0.095)	(.033)	(0.129)
RQ4: in-party	$.062^{**}$	0.161^{*}	$.060^{**}$	0.184^{*}	$.053^{*}$	0.103	$.083^{**}$	0.269^{*}
messages	(.022)	(0.082)	(.021)	(0.080)	(.026)	(0.097)	(.032)	(0.118)
RQ4: out-	−.027	-0.206^{*}	−.031	-0.306^{***}	$-.058^{*}$	-0.333^{**}	.041	−0.123
party outlets	(.022)	(0.091)	(.021)	(0.089)	(.026)	(0.112)	(.032)	(0.120)
RQ4: out-	−.017	-0.215^{*}	−.025	-0.223^{*}	−.029	-0.287^{**}	.005	0.026
party	(.022)	(0.090)	(.022)	(0.087)	(.027)	(0.107)	(.032)	(0.124)

Note: Unstandardized coefficients with corrected standard errors in parentheses. RQ4 referred to interaction effects with strong partisanship. For RQ2, M denotes message cues and O refers to outlet cues.
† $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

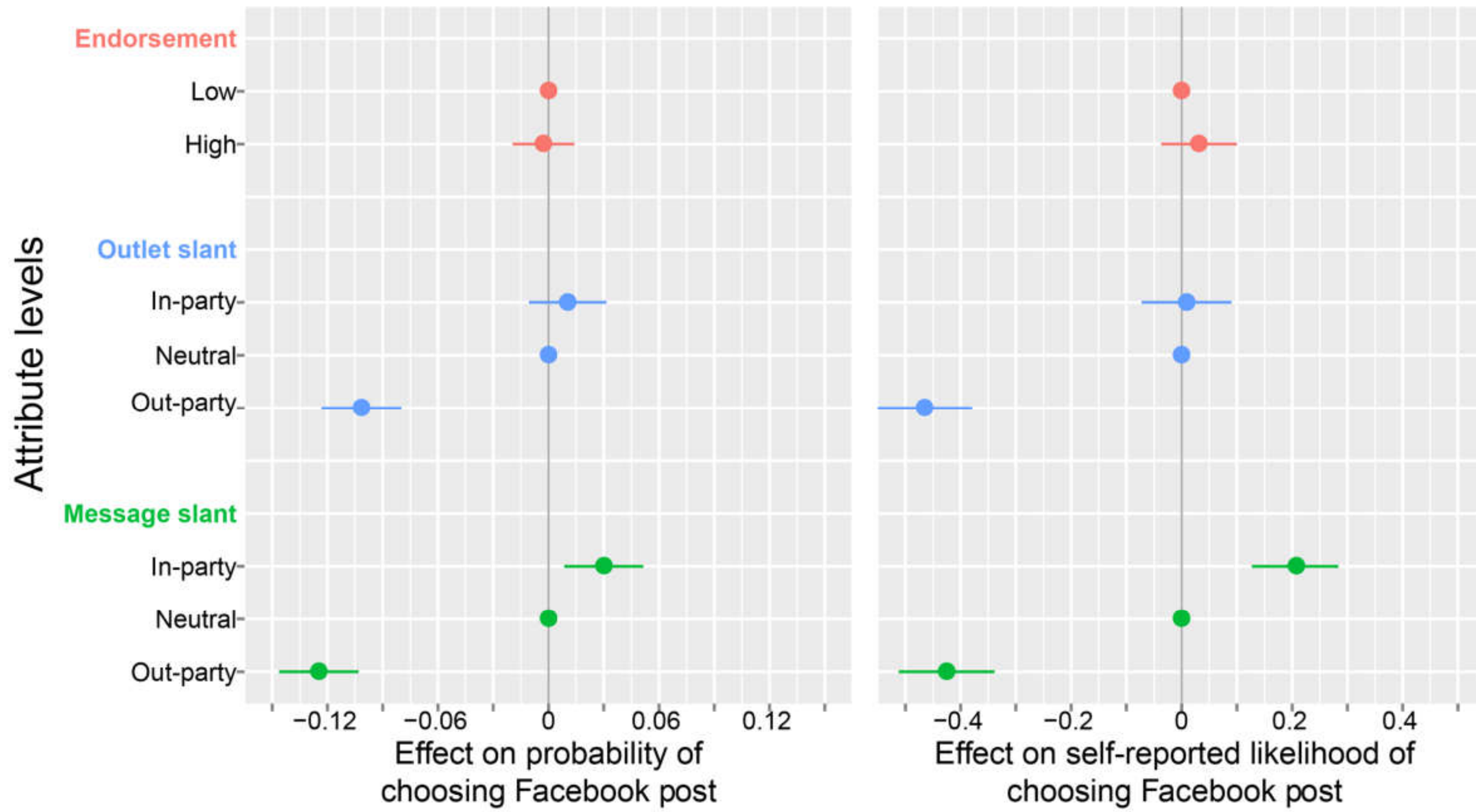


Figure A.3. Effects of Social Media Cues on News Selection (deduplicated responses only)
Note: The figure illustrates estimates of the effects of social media cues on news selection. Estimates are based on the OLS model with clustered standard errors. The bars denote 95% confidence intervals and the points without bars represent the attribute level that is the reference for each attribute.

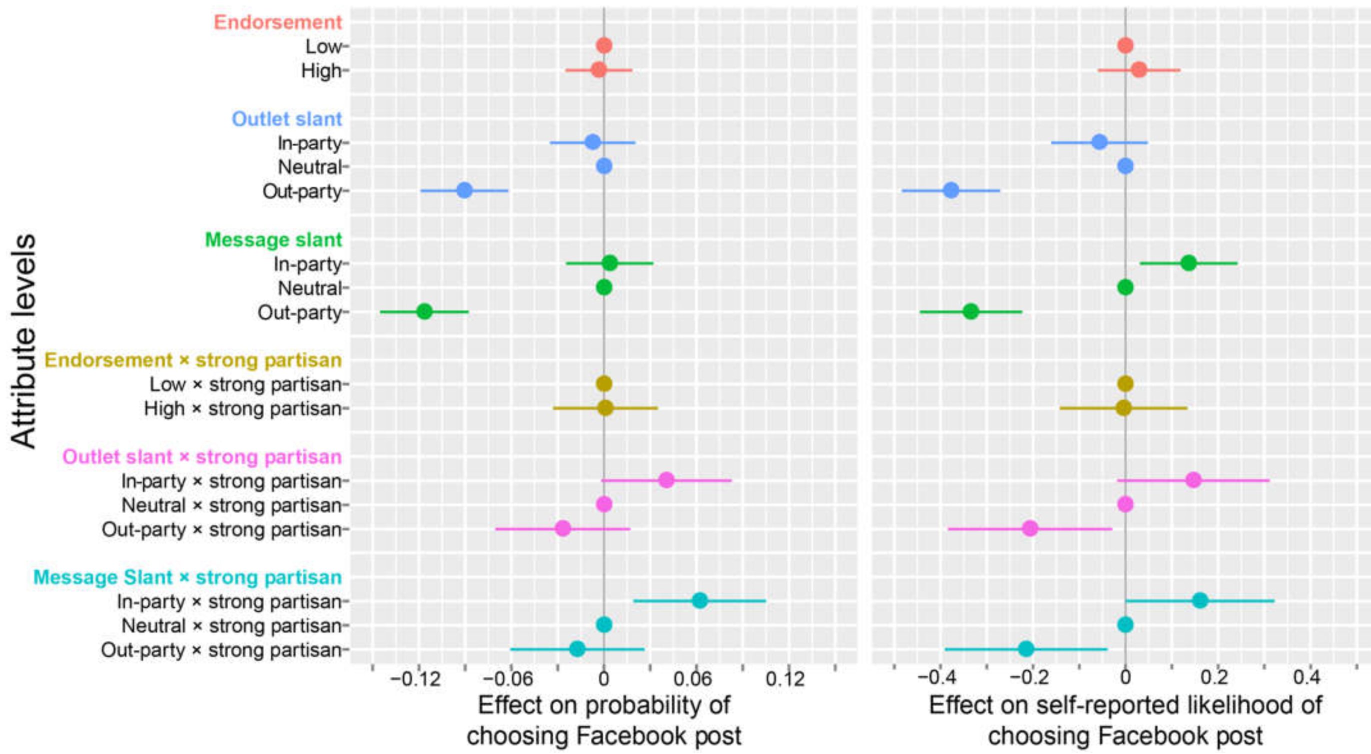


Figure A.4. Interaction Effects of Social Media Cues with Partisan Strength on News Selection (deduplicated responses only)
Note: The figure illustrates estimates of the effects of social media cues on news selection. Estimates are based on the OLS model with clustered standard errors. The bars denote 95% confidence intervals and the points without bars represent the attribute level that is the reference for each attribute.

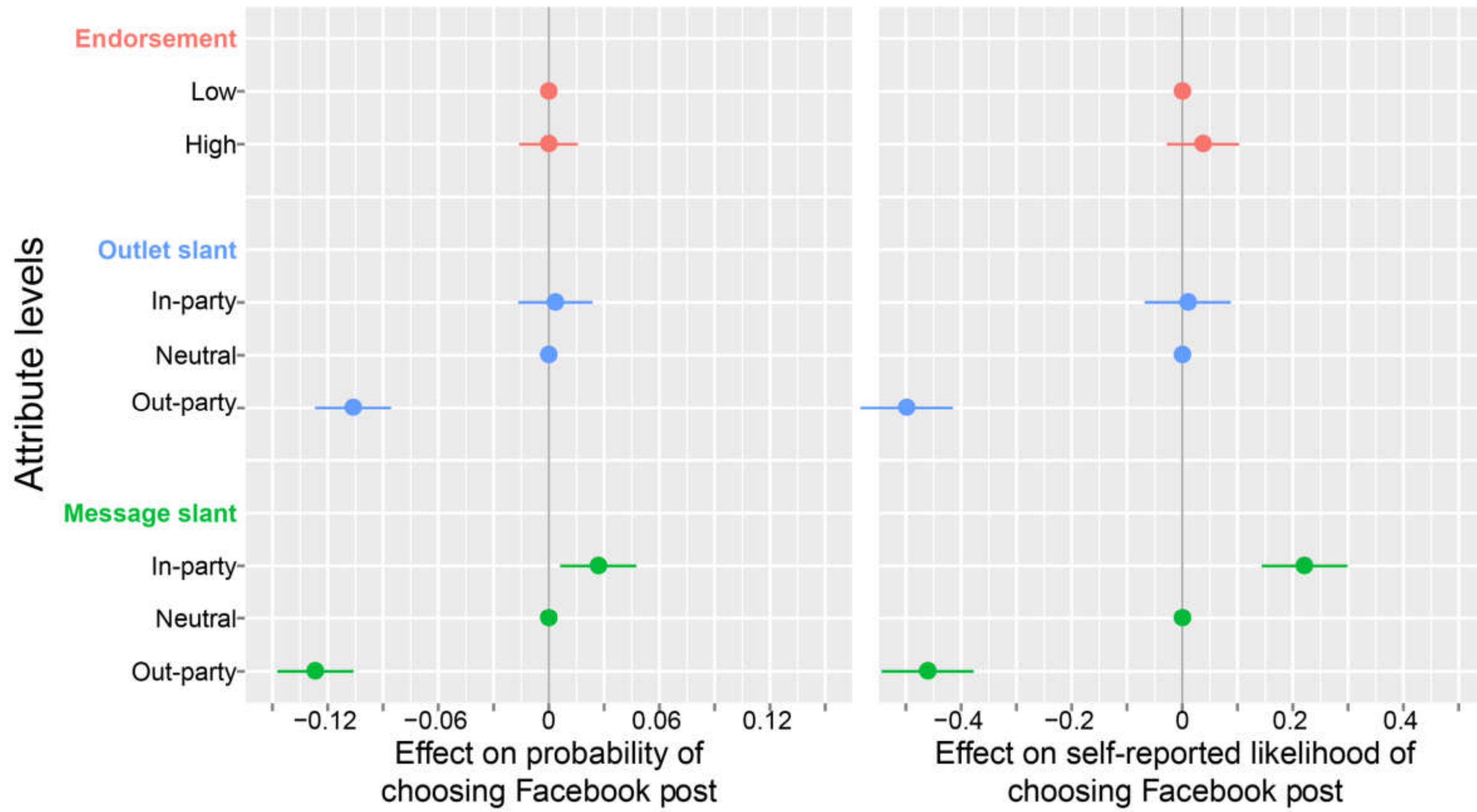


Figure A.5. Effects of Social Media Cues on News Selection (valid responses only)
Note: The figure illustrates estimates of the effects of social media cues on news selection. Estimates are based on the OLS model with clustered standard errors. The bars denotes 95% confidence intervals and the points without bars represent the attribute level that is the reference for each attribute.

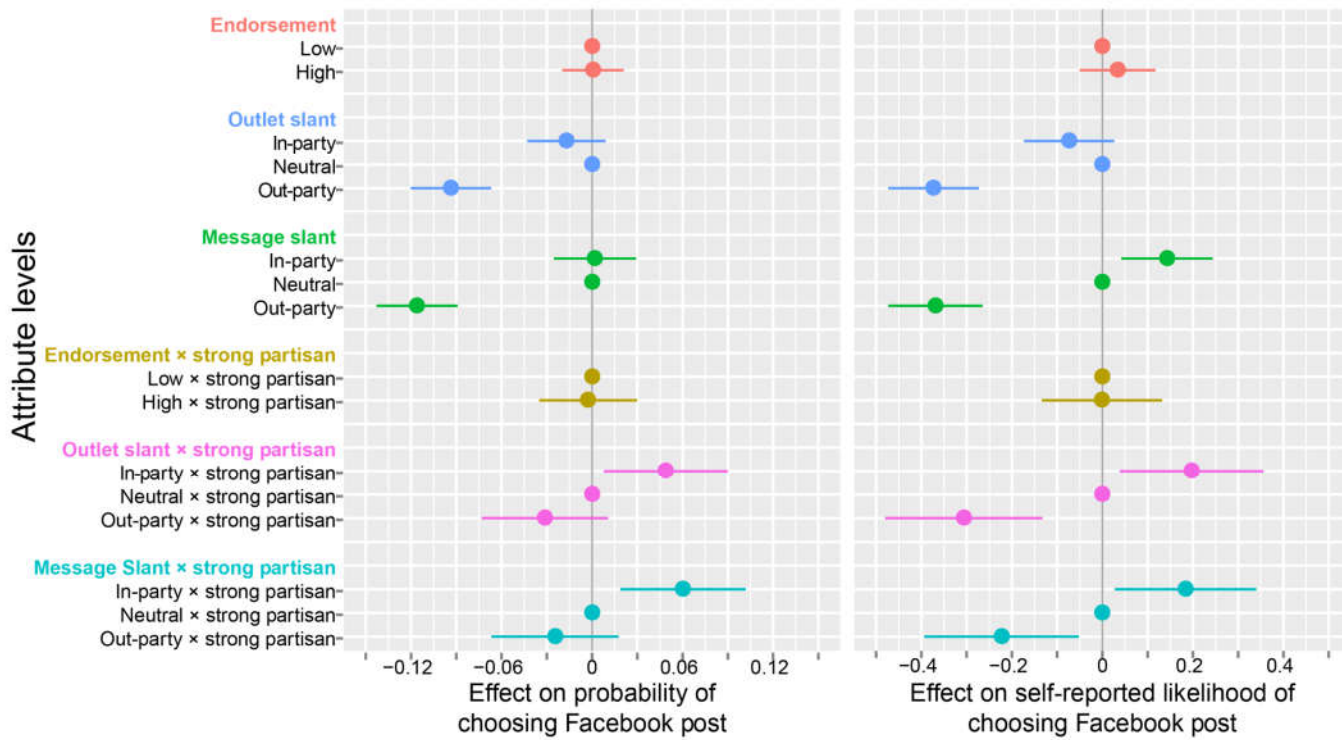


Figure A.6. Interaction Effects of Social Media Cues with Partisan Strength on News Selection (valid responses only)
Note: The figure illustrates estimates of the effects of social media cues on news selection. Estimates are based on the OLS model with clustered standard errors. The bars denote 95% confidence intervals and the points without bars represent the attribute level that is the reference for each attribute.

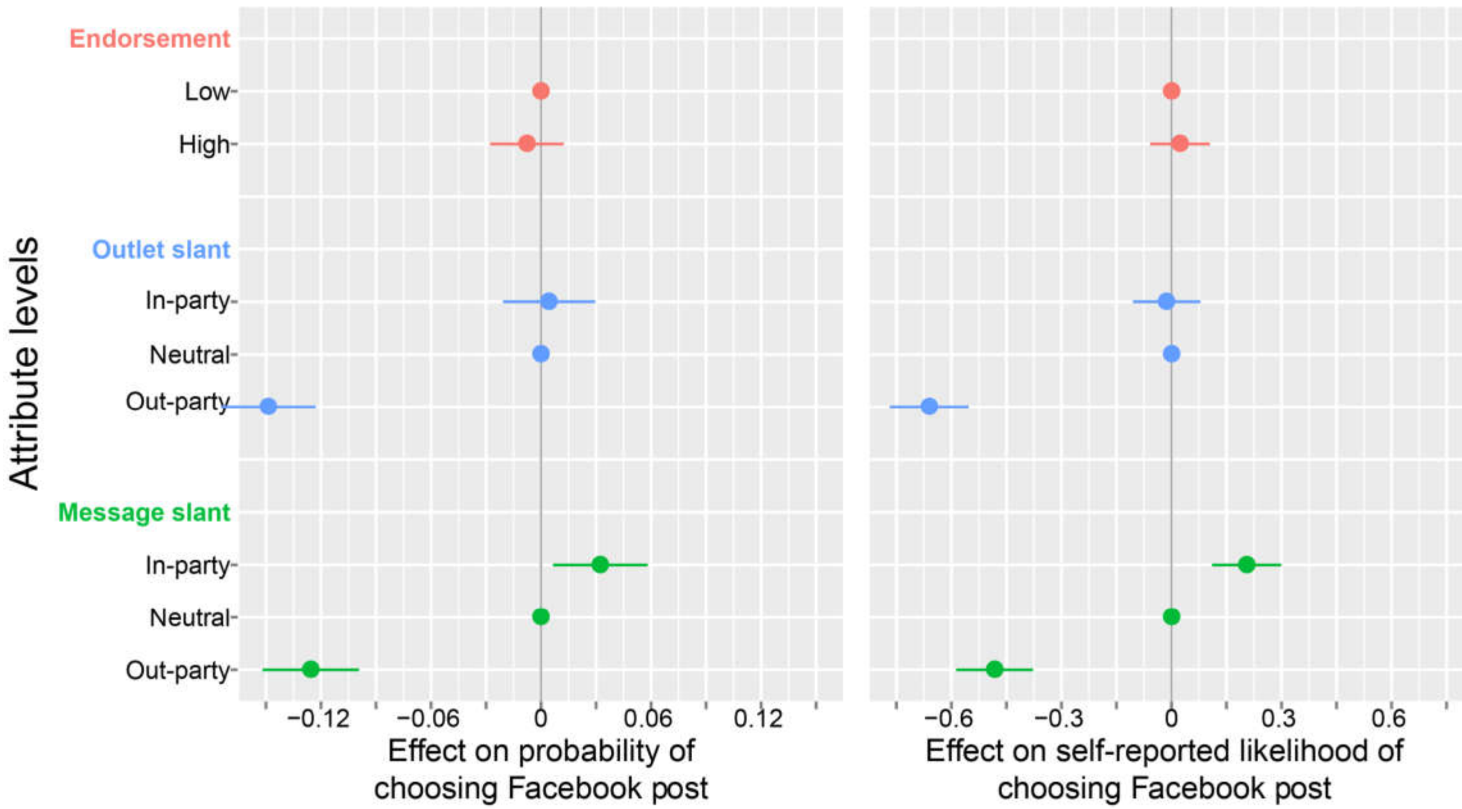


Figure A.7. Effects of Social Media Cues on News Selection (Democratic responses only)
Note: The figure illustrates estimates of the effects of social media cues on news selection. Estimates are based on the OLS model with clustered standard errors. The bars denotes 95% confidence intervals and the points without bars represent the attribute level that is the reference for each attribute.

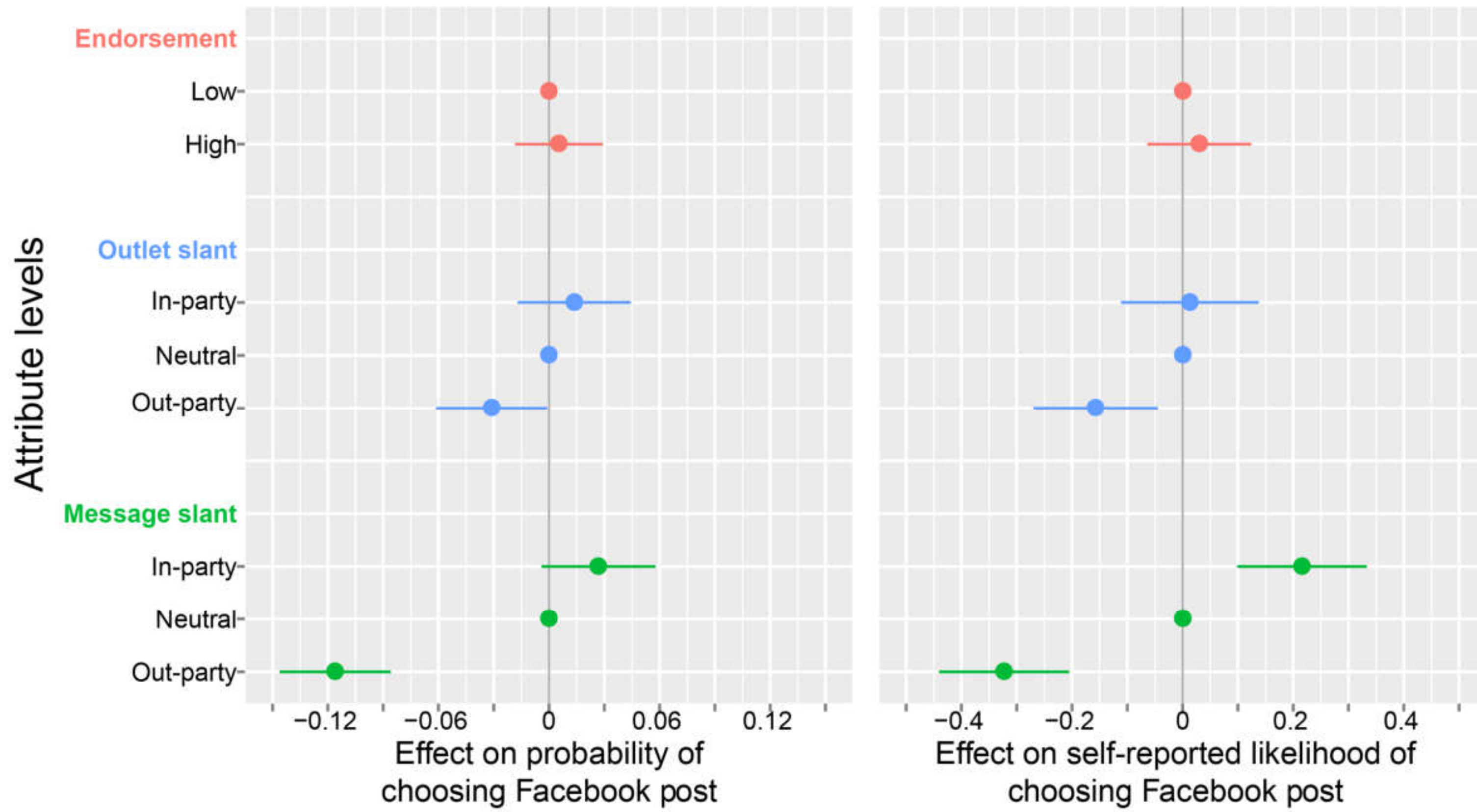


Figure A.8. Effects of Social Media Cues on News Selection among (Republican responses only)
Note: The figure illustrates estimates of the effects of social media cues on news selection. Estimates are based on the OLS model with clustered standard errors. The bars denote 95% confidence intervals and the points without bars represent the levels that are the references for each attribute.

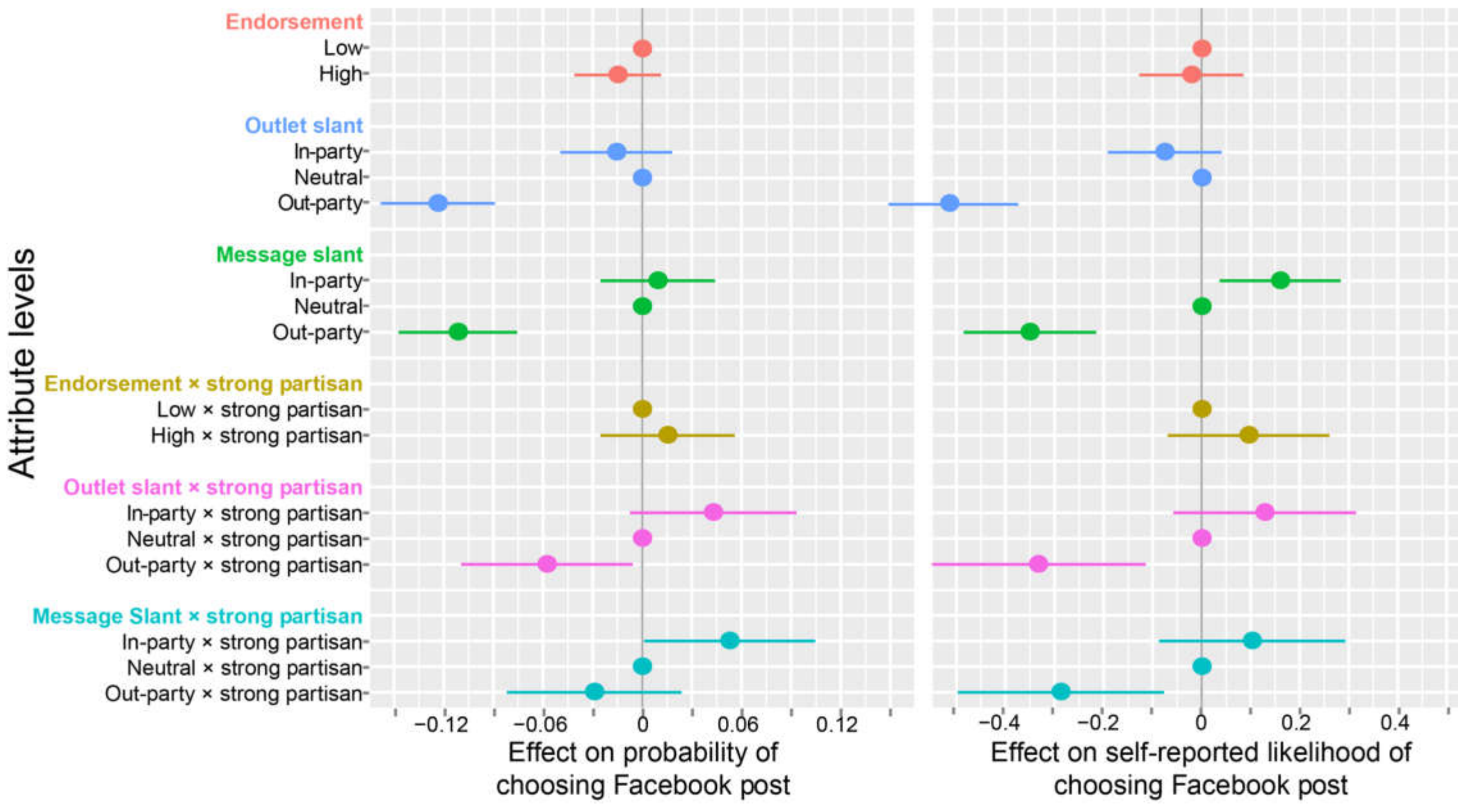


Figure A.9. Interaction Effects of Social Media Cues with Partisan Strength on News Selection (Democratic responses only)
Note: The figure illustrates estimates of the effects of social media cues on news selection. Estimates are based on the OLS model with clustered standard errors. The bars denote 95% confidence intervals and the points without bars represent the levels that are the references for each attribute.

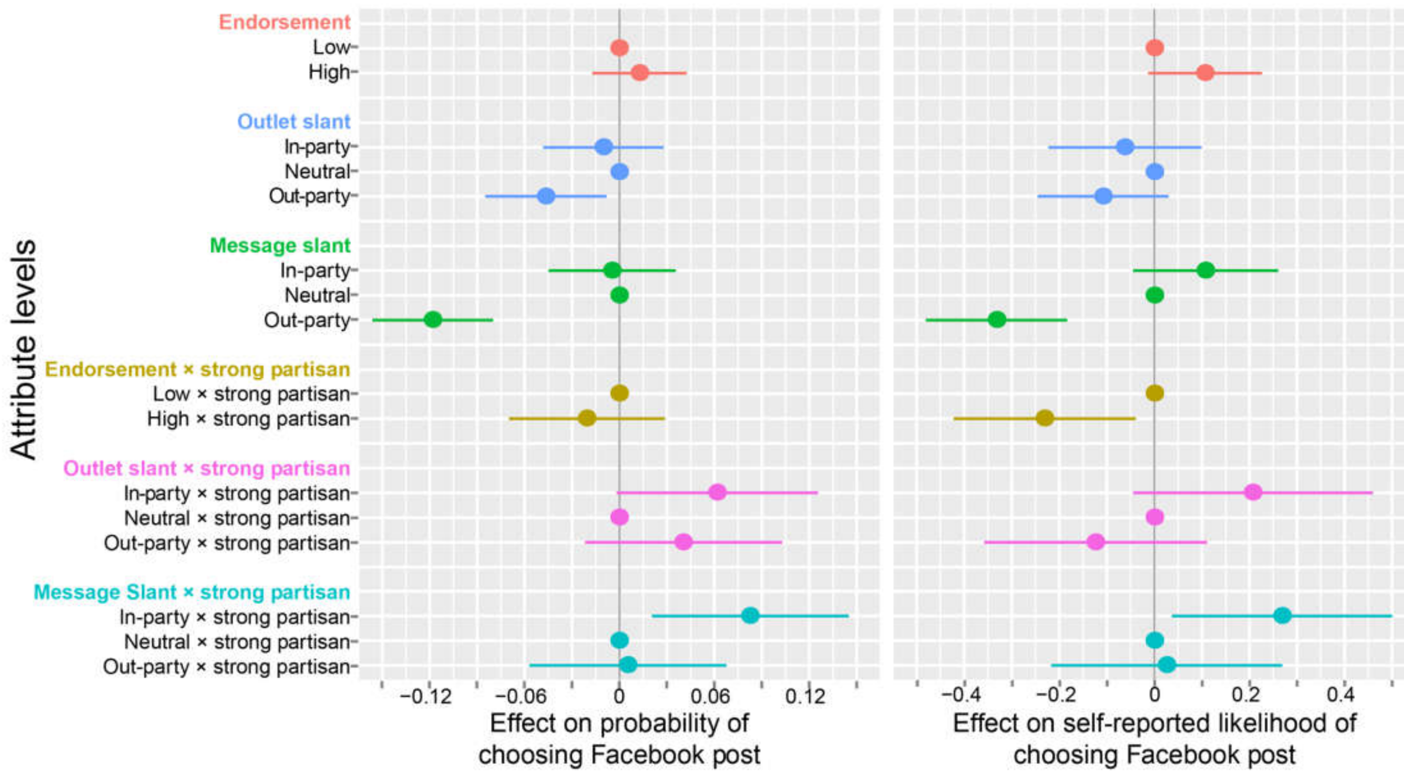


Figure A.10. Interaction Effects of Social Media Cues with Partisan Strength on News Selection (Republican responses only)
Note: The figure illustrates estimates of the effects of social media cues on news selection. Estimates are based on the OLS model with clustered standard errors. The bars denote 95% confidence intervals and the points without bars represent the levels that are the references for each attribute.

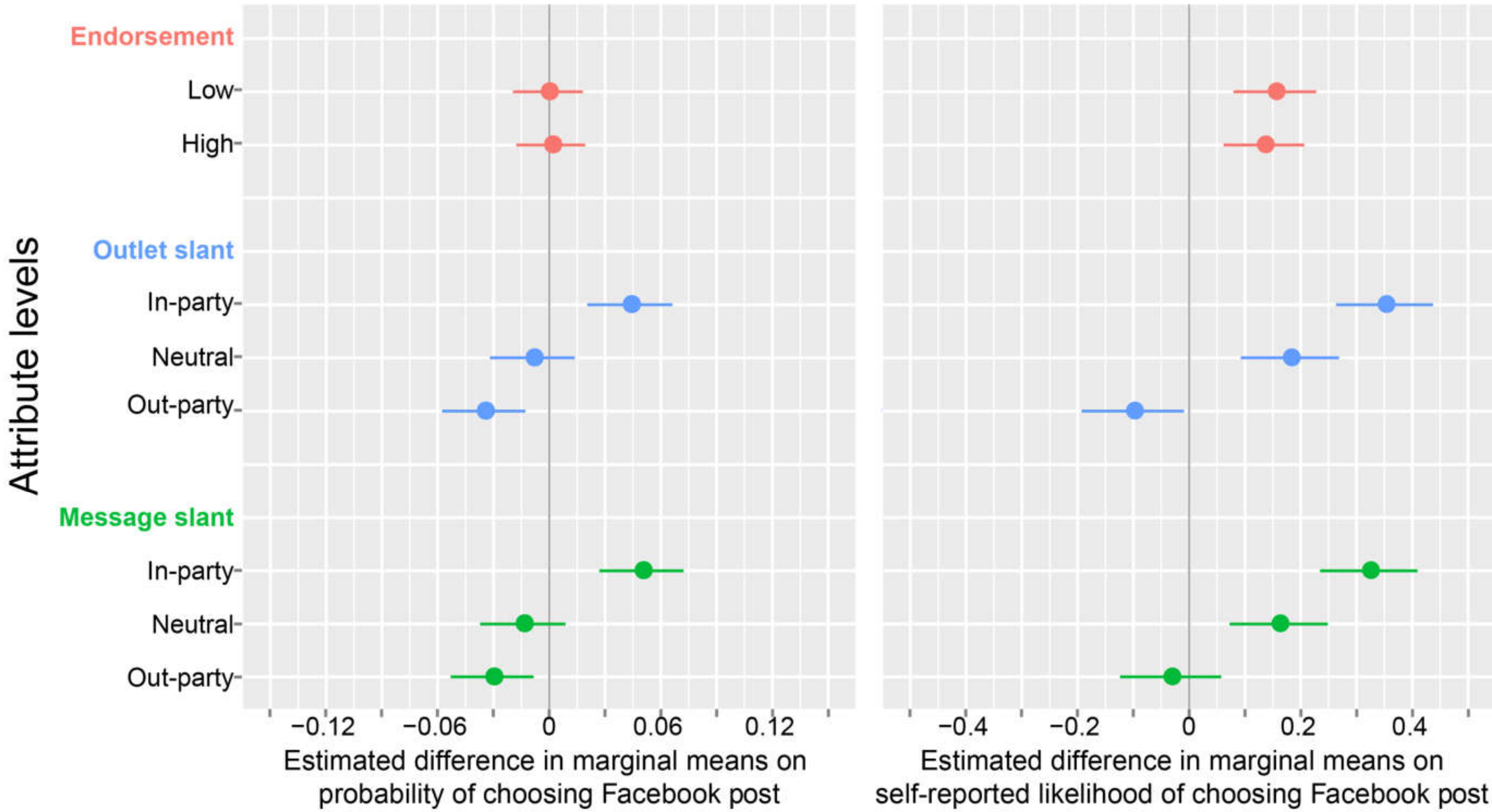


Figure A.11. Estimated Difference between Strong Partisans and Weak Partisans in Conditional Marginal Means on News Selection
Note: The figure was calculated by using methods suggested by Leeper and colleagues (2019). The bars denote 95% confidence intervals.

Table A.7 Results of robustness checks on coefficient differences

Model	Research question	Coefficient	Forced response	Rated response
$Y = \beta_0 + \beta_1 Endorsement + \beta_2(InOutlet - OutOutlet) + \beta_3(InMessage - OutMessage) + \delta_1 OutOutlet + \delta_2 OutMessage$	RQ1a: outlet cues	δ_1	$-.094^{***}$ (.017)	-0.458^{***} (0.067)
	RQ1b: message cues	δ_2	$-.091^{***}$ (.018)	-0.210^{**} (0.066)
$Y = \beta_0 + \beta_1 Endorsement + \beta_2(OutOutlet + OutMessage) + \beta_3(InOutlet + InMessage) + \delta_1 OutMessage + \delta_2 InMessage$	RQ2: avoidance	δ_1	$-.020$ (.014)	0.039 (0.055)
	RQ2: selection	δ_2	$.023^{\dagger}$ (.014)	0.209^{***} (0.053)

Note: Unstandardized coefficients with corrected standard errors in parentheses.
† $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

Table A.8 Means of Each Experimental Group

			Forced Response						Rated Response						*N*		
Endorse-ment	Outlet slant	Message slant	All		Dems		Reps		All		Dems		Reps		All	Dems	Reps
			M	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*			
High	Out-party	Out-party	0.341	0.016	0.299	0.019	0.403	0.025	3.208	0.074	2.930	0.094	3.621	0.114	912	545	367
High	Out-party	Neutral	0.433	0.016	0.394	0.021	0.484	0.025	3.555	0.070	3.235	0.088	3.985	0.108	950	545	405
High	Out-party	In-party	0.495	0.017	0.476	0.022	0.521	0.026	3.844	0.074	3.583	0.100	4.205	0.104	905	525	380
High	Neutral	Out-party	0.425	0.016	0.443	0.021	0.398	0.025	3.603	0.069	3.474	0.089	3.796	0.110	906	544	362
High	Neutral	Neutral	0.572	0.017	0.596	0.022	0.538	0.026	4.075	0.068	4.027	0.086	4.143	0.112	870	513	357
High	Neutral	In-party	0.615	0.016	0.631	0.020	0.592	0.025	4.346	0.068	4.403	0.092	4.267	0.100	901	526	375
High	In-party	Out-party	0.469	0.016	0.477	0.022	0.460	0.025	3.749	0.072	3.591	0.094	3.956	0.109	884	501	383
High	In-party	Neutral	0.564	0.017	0.568	0.023	0.559	0.025	4.108	0.065	4.033	0.086	4.211	0.099	885	511	374
High	In-party	In-party	0.588	0.017	0.599	0.022	0.570	0.028	4.243	0.070	4.198	0.089	4.310	0.114	839	504	335
Low	Out-party	Out-party	0.354	0.016	0.333	0.020	0.387	0.026	3.249	0.076	3.018	0.100	3.613	0.111	893	547	346
Low	Out-party	Neutral	0.459	0.017	0.433	0.022	0.502	0.027	3.616	0.075	3.389	0.097	3.997	0.113	870	545	325
Low	Out-party	In-party	0.506	0.016	0.473	0.021	0.555	0.026	3.749	0.072	3.467	0.094	4.173	0.104	884	531	353
Low	Neutral	Out-party	0.431	0.017	0.432	0.022	0.428	0.025	3.527	0.071	3.376	0.094	3.751	0.107	864	518	346
Low	Neutral	Neutral	0.568	0.017	0.598	0.022	0.525	0.025	4.092	0.069	4.108	0.086	4.070	0.113	866	508	358
Low	Neutral	In-party	0.583	0.016	0.600	0.020	0.554	0.028	4.324	0.068	4.204	0.086	4.529	0.108	890	563	327
Low	In-party	Out-party	0.438	0.016	0.460	0.021	0.406	0.024	3.616	0.074	3.508	0.095	3.778	0.115	899	539	360
Low	In-party	Neutral	0.595	0.016	0.611	0.021	0.573	0.025	4.031	0.072	4.012	0.093	4.057	0.114	861	491	370
Low	In-party	In-party	0.585	0.016	0.612	0.021	0.547	0.025	4.220	0.066	4.169	0.087	4.290	0.101	877	508	369

Note: *M* referred to mean value, and *SE* denoted standard error clustered by respondents.

References

Budak, C., Goel, S., & Rao, J. M. (2016). Fair and balanced? Quantifying media bias through crowdsourced content analysis. *Public Opinion Quarterly*, *80*(Special Issue I), 250–271. https://doi.org/10.1093/poq/nfw007

Knobloch-Westerwick, S., & Meng, J. (2009). Looking the other way. *Fire Rescue Magazine*, *27*(1957), 107–108. https://doi.org/10.1177/0093650209333030

Kurdi, B., Lozano, S., & Banaji, M. R. (2017). Introducing the Open Affective Standardized Image Set (OASIS). *Behavior Research Methods*, *49*(2), 457–470. https://doi.org/10.3758/s13428-016-0715-3

Leeper, T. J., Hobolt, S. B., & Tilley, J. (2019). Measuring Subgroup Preferences in Conjoint Experiments. *Political Analysis*, 1–15. https://doi.org/10.1017/pan.2019.30

Messing, S., & Westwood, S. J. (2014). Selective Exposure in the Age of Social Media: Endorsements Trump Partisan Source Affiliation When Selecting News Online. *Communication Research*, *41*(8). https://doi.org/10.1177/0093650212466406