

Thinking Fast and Furious: Emotional Intensity and Opinion Polarization in Online Media

David Asker* Elias Dinas^{†*}

Abstract

How do online media increase opinion polarization? The “echo chamber” thesis points to the role of selective exposure to homogeneous views and information. Critics of this view emphasize the potential of online media to expand the ideological spectrum news consumers encounter. Embedded in this discussion is the assumption that online media affects public opinion via the range of information that it offers to users. We show that online media can induce opinion polarization even among users exposed to ideologically heterogeneous views, by heightening the emotional intensity of the content. Higher affective intensity provokes motivated reasoning, which in turn leads to opinion polarization. We provide empirical evidence for this argument through an online experiment focusing on the comments’ section, a user-driven tool of communication whose effects on opinion formation remain poorly understood. We show that participants randomly assigned to read an online news article with a user comment section subsequently express more extreme views on the topic of the article than a control group reading the same article without any comments. Consistent with our expectations, this effect is driven by the emotional intensity of the comments, lending support to the idea that the mechanism behind this effect is motivated reasoning.

Keywords: polarization; comments section; motivated reasoning; echo chambers

*Department of Politics and International Relations, University of Oxford, Manor Road Building, OX1 3UQ, Oxford, david.asker@gmail.com.

[†]European University Institute, Department of Political and Social Sciences, Villa Sanfelice, San Domenico di Fiesole, I-50014, Italy, elias.dinas@eui.eu.

*Basically, Trump is what would happen
if the comments section became a human
and ran for president.*
--@momasnark, Twitter user

In 2006, *Time* magazine deviated from its current practice of naming a specific individual as its *Person of the Year*, instead bestowing that honor on “You.”¹ *Time* thus intended to recognize a new type of media user, one who was not just a passive recipient of information, but an active producer of it. At the time, such democratization of the media seemed, to *Time* and others, a cause for celebration. Yet in the ensuing decade, online media expanded, penetrated global markets and transformed long-term media consumption habits. Along this process, the initial enthusiasm gave way to increasing concern about the impact of online media on public civility and opinion polarization. Perhaps no example describes this turning point more succinctly than *Time*’s own cover ten years later, as the magazine announced in 2016 that “we’re losing the internet to the culture of hate.”²

Concern about the polarizing effects of the internet is not confined to news commentators and pundits. A rapidly growing literature in communication studies, political science and sociology has tried to shed light on the way in which computer-mediated communication leads to political polarization. For some, this comes as a direct consequence of the “echo-chamber” thesis (Bennett and Iyengar 2008; Quattrociocchi, Scala and Sunstein 2016), which portrays online media as virtual spaces that promote opinion similarity and generate “filter bubbles” against diverging views (Pariser 2011). Yet others have found that online media provide citizens with an increasingly diverse mix of pro- and counter-attitudinal information (Bakshy et al. 2012; Bakshy, Messing and Adamic 2015; Barberá et al. 2015). The underlying reasoning in both lines of research is that online media matter because they qualify (either shrink or expand) the information set available to users.

¹ *Time*, December 25, 2006.

² *Time*, August 18, 2016.

This study enters the debate by shifting attention away from the political homophily thesis and closer to the means through which information is disseminated. Drawing on theories of motivated reasoning, we posit that online media induce polarization because they encourage information to be expressed with heightened levels of emotional intensity. Due to a well-documented phenomenon of “toxic disinhibition” in computer-mediated communication (Lapidot-Leffer and Barak 2012), user-generated information on the internet is often emotionally charged (Kwon and Cho 2017). Compared to professional journalism, direct user interaction becomes more aggressive and potentially more polarizing, as motivated rejection of challenging views becomes more likely.

We test this claim by looking at a specific online format: the comments section. We see the comments section as one prevalent but far from unique manifestation of a new ecosystem that has developed around online interactivity (Ksiazek, Peer and Zivic 2015), in which professional journalism is juxtaposed with user-generated, opinionated communication.

We explore the effect of user comments on opinion extremity in an experiment using two real articles from *USA Today*, published both in hard copy and online. Both treated and control groups read the online version, but only the treated group has access to a random sample of (the actual) user-generated comments. We expect that, on average, user comments are more emotionally intense than journalistic content, and for this reason exposure to comments will increase polarization among readers.³ Building on this assumption, we compare treated and control group across both articles. To shed light on our proposed mechanism, namely emotional intensity, we examine the heterogeneity of treatment effects according to the emotional intensity of the comments. We expect effects to be higher for users confronted with more emotionally charged comments. Finally, building on the findings of a vast literature on motivated reasoning, we expect the emotional intensity of the comments to be more

³ Drawing on previous work (Anderson et al. 2014; Bischof and Wagner 2018), we understand polarization, as the aggregate outcome of individual-level change towards more extreme opinions. In particular, we expect the user-generated comments to drive respondents to more extreme stances towards either side of the issue dimension.

influential among politically sophisticated users. Consistent with our expectations, we find non-trivial average and heterogeneous treatment effects. Our results suggest that exposure to emotionally intense comments turns users towards more extreme views and their effect is particularly strong for politically sophisticated users.

Our study contributes to various streams of literature. First, although user-generated comments have received considerable attention in communication studies, we still know very little about their implications for readers’ political attitudes. Moreover, given that user-generated information extends to various facets of online media, including social media and blog posts, the findings speak more broadly to the role of the user as an influential mediator of online news. More generally, the evidence informs a vast literature in social psychology and public opinion about the conditions under which new information can lead to attitudinal change. The bulk of this work focuses on whether new information is congruent to or divergent from the recipient’s priors. We illustrate that what also matters is how this information is expressed. The more emotionally charged this information is the more likely it is to drive people towards more extreme views. We describe the design and present the results in more detail after we first explain the theoretical motivation of the study.

Online Media, Motivated Reasoning and Polarization

Some skeptical voices notwithstanding (Fiorina 2014), there is growing academic consensus that polarization in American politics has escalated. This phenomenon is often attributed to an increasingly fragmented media landscape (Levendusky 2013; Stroud 2010) allowing consumers to tailor exactly what information reaches them, resulting in ideologically homogeneous “echo chambers” (Sunstein 2009). Many scholars believe the Internet to exacerbate this tendency (Lelkes, Sood and Iyengar 2017), and to intensify the process of polarization through direct interaction between ideologically homogeneous users (Colleoni, Rozza and

Arvidsson 2014; Conover et al. 2011; Garcia et al. 2015; Halberstam and Knight 2016).⁴ Critics of this view point out that online networks tend to be more ideologically diverse than the echo chamber thesis suggests (Bakshy, Messing and Adamic 2015), and that even in networks where users are politically homogeneous, the news and information they exchange need not be, as users often share news they do not sympathize with (Messing and Westwood 2014). As a result, online media can have a moderating effect on opinions (Barberá 2014).

Even if they disagree about the level of cross-partisan penetration of online media, both strands in this literature seem to agree that what matters for political polarization is the level of similarity between the content of the cue and the cue recipient. Implicit in this line of argument is the idea that news messages can only have persuasive effects, i.e. strengthening users' priors when the content matches them and weakening them when it contrasts them. This is why political homophily is understood to increase polarization whereas diversity in exposure is expected to decrease it. That said, the empirical evidence does not seem to always support these stylized facts. Even when actual interchange of cross-cutting opinions takes place, it sometimes fails to moderate views. Yardi and Boyd (2010), for example, find that the interaction between different-minded individuals on Twitter reinforces accentuation effects, strengthening in-group/out-group distinctions (see also Levendusky 2013). Such results strongly suggest that there is more to the phenomenon of online opinion polarization than selective exposure alone.

We approach this question through the lens of motivated reasoning, which holds that individuals' responses to new information are conditioned by their priors (Kunda 1990). Broadly speaking, individuals are motivated to preserve their existing views, and will find reason to reject or rationalize new information which might undermine them (Taber and Lodge 2006). Numerous studies have shown how partisanship, ideology or issue positions

⁴ As Barberá et al. (2015) show, even if social networks initially form across the political space, they gradually sort into partisan or ideological clusters. Bright (2016) illustrates how ideological convergence and extremity shape the formation and dynamics of online networks: interactions take place more within than between ideological sides and centrist ideological pairs interact more frequently than extreme ideological pairs.

influence how voters weight and process information from the news (Campbell et al. 1960; Carsey and Layman 2006; Gaines et al. 2007; Levendusky 2013). Emotional intensity plays a central role in the motivated reasoning framework. Individuals, according to the theory, automatically reject messages they are predisposed to disagree with if they have negative emotional reactions to them. The “automaticity of affect” for such messages means that rejection is involuntary and irrational: the individual becomes emotionally motivated to reject the information, and then assembles a counter-argument to rationalize skepticism of it (Lodge and Taber 2005).

This affective dimension is particularly relevant to discussions of online media. As explained by Suler (2004), common features of online communication, like anonymity and lack of immediate interpersonal feedback, can induce “toxic disinhibition,” whereby users act impulsively, often resorting to hostile and uncivil communication. Networked environments facilitate high-arousal emotional expressions, as users are accustomed to blending and exchanging ostensibly discrete genres of information — facts, emotions and opinions (Papacharissi and de Fatima Oliveira 2012). The successful transmission of online messages is often contingent upon their level of emotional activation, with messages characterized by higher levels of arousal gaining more attention (Kwon and Cho 2017). Eventually, even in the absence of nonverbal cues such online textual interactions end up generating emotional contagion (Kramer, Guillory and Hancock 2014). We hypothesise that exposure to such high-emotion stimuli induces higher levels of opinion polarization.

To test this hypothesis, we focus on one feature of online news in particular, namely the comments section. Despite having received only scant interest from public opinion researchers, the comments section has received considerable attention by political commentators (Moosa 2014) and communication scholars (Lee 2012; Papacharissi 2004). Numerous studies have shown empirically that the disinhibition effects explained above are prevalent in online comments sections, resulting into “rampant incivility” (Ksiazek, Peer and Zivic 2015; Santana 2014). Although the extent of the phenomenon remains somewhat controversial

(Collins and Nerlich 2015; Papacharissi 2004; but see Coe, Kenski and Rains 2014), when it occurs, incivility appears to induce polarization in the risk assessment of new technologies (Anderson et al. 2014) and in attitudes towards political movements (Suhay et al. 2015). We go beyond incivility to look at the emotional intensity of user-driven comments. We examine the impact of emotional intensity on people’s opinions about political issues. Within this context, our main hypothesis can be expressed as follows:

H_1 : Readers exposed to more emotionally intense comments report more extreme opinions about political issues.

We try to delve more into the role of motivated reasoning in driving the effects of emotional intensity. A voluminous literature in motivated reasoning has provided ample empirical evidence suggesting that resistance to new information increases with political knowledge (see e.g. Gaines et al. 2007; Redlawsk 2002; Stoker 1993; Taber and Lodge 2006). As Miller, Saunders and Farhart (2016) point out, two mechanisms have been put forward in the literature. First, higher political sophistication implies higher ability to link abstract principles with more concrete attitudes. Second, people who care more about politics tend to hold stronger attitudes and are thus more keen on protecting these attitudes. In sum, political sophistication leads to quicker activation of affective responses to new stimuli, making motivated reasoning more likely (Lodge and Taber 2013). Accordingly, we expect that:

H_2 : The effect of emotionally intense comments will be greater for participants with greater political knowledge.

Having specified our theoretical expectations, we now turn to our empirical strategy. In the next section we describe both our key measures and our experimental setup.

Research Design

Sample

We conducted an experiment on the Amazon Mechanical Turk (MTurk) online crowdsourcing platform between December 2015 and February 2016.⁵ 660 MTurk workers, from 46 US states, aged between 20 and 69 (median 33), participated in the experiment. The experiment consisted of an online survey, hosted on the Qualtrics platform, and entailed three parts: a pre-treatment questionnaire, a treatment condition, and a post-treatment questionnaire.⁶

Covariates, Treatments and Outcomes

The pre-treatment questionnaire contained basic demographic questions as well as questions on party identification, ideology and political knowledge.⁷ The treatment conditions consisted of participants being asked to read a short online news article, either accompanied by reader comments sampled from the live comments section of the article (for the treatment group) or the same article without any reader comments (the control group). Participants were randomly assigned to one of the two conditions, and within the treatment group, comments were randomly sampled from a larger set of 18 reader-generated comments, such that each participant in the treatment group were shown three comments.⁸ On a subsequent page, participants were presented with an attention check asking them to summarize contents of

⁵ The suitability of MTurk for social science experiments has been well documented, and the platform has been shown to compare favorably to convenience samples in terms of data reliability (Buhrmester, Kwang and Gosling 2011). In Table 1 we present detailed descriptive statistics of our participant sample.

⁶ The experiment had received ethics approval from the Central University Research Ethics Committee (CUREC) of the University of Oxford.

⁷ Following Delli Carpini and Keeter (1993), we measure political sophistication directly through four knowledge questions on the US political system. See the exact wording and distribution of responses in Section A.1 of the SI.

⁸ Table 1 shows that the two groups are balanced across a series of demographics and other covariates.

the article, in order to verify they had read and understood it.

The articles used in the experiment were chosen from the official website of *USA Today*, the largest U.S. newspaper by circulation.⁹ *USA Today* is also the most ideologically diverse of the major US newspapers in terms of readership according to the Pew Research Center (2012). This diversity of readers is helpful to our design, because it provides a large ideological range of reader comments to draw from (see also Iyengar and Hahn 2009; Messing and Westwood 2014).

Participants were assigned to read one article on either new gun control measures proposed by the Obama White House, or negotiations surrounding the Trans-Pacific Partnership (TPP). The topics of the articles were chosen to vary in terms of partisanship and political saliency. Whereas the gun control issue is more politically salient and evokes clear partisan signals, the Trans-Pacific Partnership was, at the time of the experiment, less politically salient and less partisan. These differences are expected to generate variation in the degree of emotional intensity of the user-generated comments. By default, three comments were randomly sampled and shown to each respondent, using the “Advanced Randomization” tools in Qualtrics’ survey flow interface. In cases where comments formed short exchanges between readers, those exchanges were kept intact to ensure comments were not presented out of context.

Figures 1 and 2 present the full text of the articles as they appeared to participants. Table 1 presents descriptive statistics about the comments and participants’ priors. Approximately half of respondents agreed with the content of the article. This share is almost identical between treatment and control groups. Consistent with the random assignment of comments across the treated observations, the share of comments agreeing with the article (approximately one third), is orthogonal to participants’ priors (see Figure A2 of the SI).

[Figure 1 about here]

⁹ According to the Alliance for Audited Media figures for 2014, the last time AAM published head-to-head comparisons of newspaper circulations.

[Figure 2 about here]

[Table 1 about here]

The outcome is measured in the post-treatment survey, which asked participants for their attitude towards the issue discussed in the article: a) new gun safety regulations; or b) the Trans-Pacific Partnership. Participants were asked to give their attitude on a sliding scale from 0 (strongly opposed) to 100 (strongly supportive) with 50 as a neutral midpoint. For purposes of analysis, this scale was then “folded” at the midpoint such that each score was converted to its absolute-value distance from the midpoint, and then rescaled to the range $[0, 1]$. The resulting “polarization score” thus represented a maximally polarized attitude as one and a completely neutral attitude as zero.

Measuring Emotional Intensity

Intensity was coded through crowdsourcing on the MTurk platform, along the lines suggested by Benoit et al. (2016). MTurk workers were recruited to code each comment on a scale from 0 (not at all emotional) to 10 (very strongly emotional), and the mean score was assigned to each comment. In total, we employed 22 coders for the TPP article and 25 coders for the gun control article. We sampled from a population of eight comments in the case of the TPP article and a population of 10 comments in the case of gun control article. Within each article, every coder coded all comments. The Cronbach’s alpha is 0.90 and the average inter-coder correlation is 0.25.¹⁰ In Figure 3 we display the distribution of emotional intensity per comment. The first panel shows the comments for the gun control article, while the second panel shows the comments for the TPP article. The maximum standard deviation of the crowdsourced intensity estimates for a given comment was 2.81 (std. error 0.56). In the SI (Section A.3, Figures A3 and A4), we also provide the distribution of

¹⁰ Looking at each article separately, Cronbach’s alpha is 0.89 and 0.93 and the average inter-coder correlation is 0.25 and 0.39 in the gun control and the TPP article respectively.

emotional intensity scores for each rater. Although some coders used a wider range of scores than others, they all seemed to identify differences in terms of emotional intensity between the comments. The standard deviation of the scores given by each coder ranges from 0.7 to 3.7 with an average standard deviation of 1.86. Below, we provide indicative examples of comments scoring high and comments scoring low in the emotional intensity scale. The real names of the commenters have been replaced by pseudonyms.

Most intense comment (Average Score 8.0):

Allan Barclay

Obummer said HE TAUGHT constitutional law, where was that at?

Just another LIE, LIE ,LIE.....

John Herman

Think Chicago. This Racist POS is a traitor and a Muslim who

hates America and Americans continuing to destroy the very

fabric of our country

Least intense comment (Average Score 3.1):

Dave Cook

GW Bush has backed this deal.

Terry Sims

Dave, there was no TPP when GW was in office!

To ensure what is captured is emotional intensity rather than additional information provided by each comment, we repeated the same procedure, this time asking coders to code each comment in terms of its informativeness, again on a scale from zero (not informative) to ten (very informative). Below we reproduce as examples the most and least informative comments according to this coding.

Most informative comment (Average score 4.4):

Mike Phillips

Clinton gave us... ^ 26 million new jobs. Obama gave us.... 13 million new jobs.

Bush gave us the Great Recession and cost us 7 million jobs.

Least informative comment (Average score 0.9):

Tony Carraway

I thought the tears added a little more pandering to the emotional side which is how

the left wants to see things. Never facts just emotions.

[Figure 3 about here]

Results

Do comment sections polarize? Figure 4 presents the average polarization score across the treatment and control group, pooling both articles together. The first panel of the figure displays the distribution of the polarization score for the treatment and control groups. Although similar, there is a clear gap on both tails of the distribution, with the treatment group reporting on average more extreme positions.¹¹ The second panel presents the average polarization score for the two groups. Participants in the treatment group reported a 12.8% higher polarization score than those in the control group (0.51 in the treatment group and 0.45 in the control group).¹² In the Supplementary Information we test the sensitivity of the

¹¹ The Kolmogorov-Smirnov test which tests the equality of the two distributions yields a one-tailed p -value of 0.051.

¹² We used randomization inference to test the significance of this difference against the sharp null of no effect, and found a p -value of 0.035. We also checked whether this difference is due to outliers, driving the mean away from the median for any of the two groups. First, we implemented the Wilcoxon rank-sum test, which tests the probability that a member of the control group will score higher than a member of the treatment group (plus half the probability that the values are tied). The resulting estimate is 0.04 ($p = 0.038$, 95% CI 0.0-0.1). We also tested explicitly for the difference between medians with a quantile regression. The results appear in the SI (Section A.4, Figure A5) and show that there is a significant difference of medians. Moreover, we find that the effect of the comments is concentrated on higher values of the polarization scale.

treatment effects to alternative measures of the dependent variable. All analyses point to the same conclusion.¹³

[Figure 4 about here.]

As a next step, we try to examine whether these results hold when controlling for a series of covariates. First, we include a binary indicator to denote which of the two articles was shown to respondents, thus estimating within-article average treatment effects. The treatment effect remains substantively intact, i.e. 0.054 (std. error 0.026, two-tailed $p < 0.05$). To this specification we also add a series of individual-level pre-treatment characteristics: gender, age and age squared, partisanship and political involvement.¹⁴ The treatment effect becomes 0.050 (std. error 0.026, two-tailed $p < 0.1$). Insofar as we believe the comments to be more emotionally intense than the article alone, these results lend support to hypothesis H_1 . In what follows, we put this premise to the test.

Is exposure to the comments increasing polarization because of the emotional intensity emanating from user-driven communication? To address this question we implement three sets of tests. First, we break the analysis down by article. We find a significantly stronger effect for the gun control article than the TPP article. The first panel of Figure 5 illustrates this difference. This is consistent with our expectation that the effect should be stronger for more salient and more partisan issues, as we expect such issues would prompt readers to produce more emotionally intense commentary. Indeed, examining the average emotional intensity of comments on each article, we find that this expectation holds true: comments on the gun control article were about 21% more emotionally intense on average (two-tailed

¹³ We use nine alternative measures, including changing where the opinion scale is folded (at the median and the mean response in addition to the scale’s logical midpoint), how the distance from this folding point is measured (by squaring the distance and by creating a dichotomous distance measure), and by measuring the share of participants who picked an extreme endpoint of the opinion scale. We present these results in Tables A1 to A3 in Section A.5 of the SI.

¹⁴ Full results are shown in Section A.6 of the Online SI, Table A4.

$p < 0.01$). The second panel of Figure 5 illustrates the difference in a boxplot showing the distributions of intensities of comments shown to participants in each experiment. Taken together these results provide suggestive evidence for the role of emotional intensity underlying the treatment effects.

Second, to ensure that the exact same level of emotional intensity is assigned to the treatment group, we use the largest block of *MTurk* coders receiving exactly the same comments as a homogenous treatment group that is compared against the control group. The estimated level of intensity assigned to these comments is above the mean intensity score across all comments, i.e. 0.66 (mean is 0.61). We find a gap in the polarization score of 0.24 points in a 0 to 1 scale (two-tailed $p < 0.05$, see Table A6 of the SI).

[Figure 5 about here]

Third, we decompose the effects according to the level of emotional intensity to which every participant in the treatment group was exposed. To see the results of this analysis we turn to Figure 6. The first entry of the Figure denotes the average treatment effects, with and without individual-level controls (while always including article-fixed effects). Each of the remaining entries of Figure 6 decomposes these effects by comparing the control group against a specific subset of the treatment group, according to the level of emotional intensity of the comments they were exposed to. To facilitate the analysis we split the treatment group into three groups: low intensity (below the 33rd percentile of the emotional intensity distribution); medium intensity (between 33rd and 66th percentile of the intensity distribution); and high intensity (above the 66th percentile of the intensity distribution).¹⁵ We find a monotone increase in the magnitude of the comments effect as the level of emotional intensity increases. For non-emotionally intense comments, exposure to comments makes no difference in opinion

¹⁵ Figure A6 in Section A.7 of the SI shows that the results remain substantively identical when we use: a) a more fine-grained categorization of emotional intensity and; b) a more continuous estimation strategy to explore the heterogeneity of the effects across the level of emotional intensity.

polarization. Exposure to more intense comments however appears to increase polarization by approximately eight percentage points. These effects remain substantively intact when adding covariates and when we try to account for the level of uncertainty accompanying the intensity estimate of each comment (see Figure A7 in Section A.10 of the Online SI). Finally, the results do not seem to be driven by specific coders. Our jackknife analysis in the Online SI (Figures A8 and A9 of Section A.11) suggests a very similar pattern as the one shown in Figure 6.

[Figure 6 about here.]

The second panel of Figure 6 provides an alternative way of summarizing the affective intensity of comments. It may be that what matters is not the average degree of emotional intensity but how emotionally intense any one comment is. Put differently, a comment that is not emotionally intense does not necessarily cancel out an emotionally intense comment. Following this logic, we partition the analysis according to the level of emotional intensity of the most intense comment. The results appear in Figure 6 and largely confirm those presented thus far. The higher the level of intensity of the most emotionally intense comment, the greater is the effect of comments on opinion polarization.

We also replicate our analysis using weighted least squares, attaching weights inversely proportional to the average standard error of the intensity score estimates for each set of comments.¹⁶ Finally, to ensure that the results are not driven by outlier assessments on behalf of any single coder, we also repeat the analysis while iteratively removing one coder at a time from the intensity score calculation.¹⁷

Exploring alternative explanations and specifications

¹⁶ This procedure, which aims to account for variation in the confidence of the coding of each comment, is inspired by Benoit et al. (2016) and is explained in more detail in the Online SI, section A.10.

¹⁷ For details on this procedure, see Section A.11 of the Online SI.

Next, we try to examine alternative mechanisms through which user-driven comments could induce opinion polarization. First, it is plausible that comments provide additional information to readers and this information strengthens their views about the issue in question. We try to test this possibility in three ways. First, we code all comments in terms of the amount of information they convey.¹⁸ We then examine if the treatment effects vary according to the level of informativeness of the comments to which users were exposed. The results appear in the first panel of Figure 7. We find a curvilinear pattern which bears no resemblance to the pattern found when looking at emotional intensity. Emotional intensity does not seem to impact polarization simply because more emotionally intense comments are also comments with more information. Second, we use a different proxy for information, namely the length of the comment. The second panel of Figure 7 shows the heterogeneity of treatment effects according to the total number of words in the comments shown to each participant. We find no monotone increase in the magnitude of the effects as we move to lengthier comments.

A second alternative explanation could be that extreme positions stem from spending more time thinking about a topic and some comments may engage readers more than others. To address this possibility, we use the time each respondent spent on the task. This information is available for all respondents, not only those exposed to the comments section.¹⁹ We interact the treatment with the amount of time (measured in seconds) each respondent spent completing the survey. We find no substantive heterogeneity in the magnitude of the treatment effect, with the coefficient of the interaction term being 0.00003 (std.

¹⁸ We follow the same procedure as with emotional intensity: for the gun control and TPP articles, we ask 19 and 21 *MTurk* workers respectively to code each comment on a scale from zero to ten, with higher scores denoting more informative comments. We then use the average as the informativeness score of each comment. See the previous section for examples.

¹⁹ Table A6 in Section A.8 of the SI displays descriptive statistics of these additional comment characteristics.

error 0.0001).²⁰

[Figure 7 about here.]

Third, it is possible that all that is captured here is another manifestation of the echo chamber mechanism. Even if users cannot control the identity of the comment providers and thus have no way to control the content of the cues, readership can be heavily clustered on partisan or ideological lines. This means that comments may simply reflect additional information that matches well with the individual’s priors. This could be the case if, for instance, liberal comments were systematically more emotionally intense than conservative comments, and most of our participants were liberal, in which case emotional intensity may simply act as a proxy for agreement with the comments.

To test the plausibility of this alternative explanation, we constructed several measures. First, we tried to capture the degree of one-sidedness in the user comments. We coded each comment according to whether it supported or opposed the policy presented in the article.²¹ Next, we estimated the share of pro-policy comments to which every respondent in the treatment group was exposed. We then examined the heterogeneity of treatment effects according to the degree of exposure to pro-policy comments. The results appear in the first panel of Figure 8. We find no evidence that the comments matter more as the distribution of comments becomes more skewed towards supportive opinions.

[Figure 8 about here.]

Second, we try to measure the degree of agreement with the comments. We do this by interacting respondents’ prior opinion about the topic in question with the proportion of

²⁰ Full results available in Section A.9, Table A7, of the SI.

²¹ Although we intended to classify each comment into three categories—pro-policy, anti-policy, and neutral—, we found no comment that could be coded as neutral. Thus, all comments were coded as either pro- or anti-policy. This means that the share of pro-policy comments is the inverse of the share of anti-policy comments, and we only present the analysis using the former expression of the measure.

comments sharing this opinion. As discussed in the previous section, approximately 55% of respondents were in favour new gun control measures and the Trans-Pacific Partnership; 25% were undecided while the remaining were against (see Figure A1 in Section A.2 of the Online SI). We use the modal category, i.e. those who agree with the policy presented in the article and interact it with the share of pro-policy comments. This interaction term captures the degree of exposure to comments respondents agreed with. If agreement is driving the results, we should find an increase in the magnitude of comment effects for higher levels of agreement.²² We test whether this is the case in the second panel of Figure 8. The results provide little support for this hypothesis. The effect of comments does not seem to grow monotonically as the level of agreement with user-comments increases.

Is Motivated Reasoning Conditional on Political Knowledge?

Finally, we turn to H_2 . We start by looking at whether the comments induced polarization differentially, according to participants' levels of political knowledge. Figure 9 presents the polarization score by treatment status, separately for those who answered all four knowledge questions correctly (high level of political knowledge, 57.14% of respondents) and those who did not (low level of political knowledge). The first panel of the Figure includes only a binary indicator denoting the article shown to readers and the second panel adds individual-level covariates (the same as in previous analyses). In accordance with the motivated reasoning literature, the polarizing effect of the comments is considerably stronger among those with higher levels of political knowledge compared to those with low levels of political knowledge. The difference-in-differences estimate is 7.3 percentage points (std. error 5.5), implying that exposure to comments was more than four times as influential for those with high political knowledge (9.4 percentage points, std. error 4.1) than for those with low political knowledge

²² We use the modal category because it offers more variation in the agreement scale, hence allowing us to better grasp the heterogeneity of treatment effects across different levels of agreement.

(2.1 percentage points, std. error 3.7).

[Figure 9 about here]

Next, we look whether this difference is indeed driven by the emotional intensity of the comments. As shown in the remaining entries of both panels of Figure 9, political sophistication makes a difference but only when comments are emotionally intense. Comments that are not emotionally intense do not induce polarization and this is the case for both low and high levels of political knowledge. Put differently, emotional intensity increases polarization but only for respondents with highest levels of political knowledge. This provides support for hypothesis H_2 , and is consistent with an interpretation of the results based on motivated reasoning.

Discussion

We have argued that online opinion polarization can occur not only through “echo chamber” effects, whereby users filter out unwelcome information, but also through motivated reasoning effects, whereby more emotionally intense messages prompt individuals to reassert their prior views and thus express more extreme positions. We argued that online media has increased the prevalence of this effect by juxtaposing professional journalism with disinhibited, emotionally intense user commentary. This argument is supported empirically by the experimental results we present. Using an online experiment, we find that displaying reader comments below a newspaper article causes readers to report more extreme opinions on the topic of the article than when no such comments are displayed. We find that this effect is greater for more politically knowledgeable participants and when comments are more emotionally intense, as expected if the effects are based on motivated reasoning.

Our results should not be taken as a challenge to the echo chamber argument, but rather as a complement to it. Selective exposure to desirable information and motivated rejection of undesirable information constitute separate mechanisms whereby online news audiences may

develop more extreme views. Whereas there is already ample empirical evidence about the first mechanism, previous research on the second has been scant. Our contribution should thus be seen as an attempt to fill this gap.

The findings of this study also help to inform both the scholarly and media discussion about the threats arising from online political discussion. Incivility, flaming and trolling have emerged as major obstacles to public deliberation on the internet (Hmielowski, Hutchens and Cicchirillo 2014; Papacharissi 2004). Typically, such behaviour flourishes when there are no means to make users accountable for the content they produce. Traditionally this is the case with the online comments' sections.²³ Although incivility is often mediated via censoring by gatekeepers, disturbing comments are far from absent. Despite the widespread concern about the expansion of this behaviour in both online and social media, the attitudinal implications of exposure to such behaviour are not yet fully understood. Using two articles published in *USA Today* and the actual comments posted under these articles, we designed an experiment to gauge their effects. Our findings indicate that emotionally charged messages increase opinion extremity among subsequent readers.

A few objections can be raised regarding the external validity of the study. First, we find that the effects are mainly concentrated among more politically sophisticated news consumers. While this is consistent with our theoretical expectations, since motivated reasoning necessitates sufficiently strong priors, it raises the question of whether this phenomenon matters to politics more broadly. If most news consumers are casual observers of politics, the effects we describe may be quite limited in scope. However, regardless of what share of the public is susceptible to these effects, it is likely that this group of politically engaged individuals exerts an outsize influence on public discourse and political life. The rise in hard-line political activism in the U.S. in recent years, dominated by the Tea Party and Trump

²³ That said, efforts are made to link comments' sections with social media accounts as a way to make users more easily identifiable and hence civilize political discussion (Rowe 2015). Constraints to anonymity (Santana 2014) as well as the engagement of an identified news reporter in the discussion (Stroud et al. 2015) can help in reducing incivility and enhance the quality of citizens' discursive participation.

supporters on the right and the anti-Trump “resistance” on the left, may suggest as much.

Second, one may question whether the effects demonstrated in this experiment persist over time, or if they are merely fleeting expressions of indignity with little impact on long-term opinion formation. We do not test this directly and indeed exploring the long-term implications of these effects would be a fruitful avenue for future research. Nevertheless, we note that previous studies of motivated reasoning in politics have found that its effects can be detected several days later even after brief exposure (Levendusky 2013). Moreover, even if the specific treatment might not last long, exposure to similar user-generated content is likely to occur at relatively regular intervals, helping these effects persist.

Finally, one may object that if most news consumers seek out highly partisan news outlets with ideologically homogeneous readership, they may not actually be exposed to emotionally intense debate in the comment sections. If a strong version of the echo chamber argument holds, the audiences of many news outlets may be too homogeneous to produce aggressive, emotional commentary. Yet the fact that most news outlets are taking active measures to police the civility of their comments sections suggests these remain the scene of vitriolic debate. Similarly, research on online news suggests that unwelcome exposure to opposing views online is in fact commonplace. One of the main contributions of our paper is to demonstrate that, even if such exposure to ideologically heterogeneous information online is quite common, the effect of online news on public opinion can still be a polarizing one.

References

- Anderson, Ashley A, Dominique Brossard, Dietram A Scheufele, Michael A Xenos and Peter Ladwig. 2014. “The “Nasty Effect:” Online Incivility and Risk Perceptions of Emerging Technologies.” *Journal of Computer-Mediated Communication* 19:373–387.
- Bakshy, Eytan, Itamar Rosenn, Cameron Marlow and Lada Adamic. 2012. The Role of Social Networks in Information Diffusion. In *Proceedings of the 21st International Conference on World Wide Web*. ACM pp. 519–528.
- Bakshy, Eytan, Solomon Messing and Lada A Adamic. 2015. “Exposure to Ideologically Diverse News and Opinion on Facebook.” *Science* 348:1130–1132.
- Barberá, Pablo. 2014. “How Social Media Reduces Mass Political Polarization. Evidence from Germany, Spain, and the US.” *Job Market Paper, New York University* 46.
- Barberá, Pablo, John T Jost, Jonathan Nagler, Joshua A Tucker and Richard Bonneau. 2015. “Tweeting from Left to Right: Is Online Political Communication More than an Echo Chamber?” *Psychological Science* 26:1531–1542.
- Bennett, W Lance and Shanto Iyengar. 2008. “A New Era of Minimal Effects? The Changing Foundations of Political Communication.” *Journal of Communication* 58:707–731.
- Benoit, Kenneth, Drew Conway, Benjamin E Lauderdale, Michael Laver and Slava Mikhaylov. 2016. “Crowd-Sourced Text Analysis: Reproducible and Agile Production of Political Data.” *American Political Science Review* 110:278–295.
- Bischof, Daniel and Markus Wagner. 2018. “Do Voters Polarize When Extreme Parties Enter Parliament?” *Unpublished Manuscript*.
- Bright, Jonathan. 2016. “Explaining the Emergence of Echo Chambers on Social Media: the

Role of Ideology and Extremism.” *Available at SSRN* .

URL: <https://ssrn.com/abstract=2839728>

Buhrmester, Michael, Tracy Kwang and Samuel D Gosling. 2011. “Amazon’s Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data?” *Perspectives on Psychological Science* 6:3–5.

Campbell, Angus, Philip Converse, Warren Miller and Donald E Stokes. 1960. “The American Voter.” *Ann Arbor: University of Michigan Press* .

Carsey, Thomas M and Geoffrey C Layman. 2006. “Changing Sides or Changing Minds? Party Conversion, Issue Conversion, and Partisan Change on the Abortion Issue.” *American Journal of Political Science* 50:464–77.

Coe, Kevin, Kate Kenski and Stephen A Rains. 2014. “Online and Uncivil? Patterns and Determinants of Incivility in Newspaper Website Comments.” *Journal of Communication* 64:658–679.

Colleoni, Elanor, Alessandro Rozza and Adam Arvidsson. 2014. “Echo Chamber or Public Sphere? Predicting Political Orientation and Measuring Political Homophily in Twitter Using Big Data.” *Journal of Communication* 64:317–332.

Collins, Luke and Brigitte Nerlich. 2015. “Examining User Comments for Deliberative Democracy: A Corpus-Driven Analysis of the Climate Change Debate Online.” *Environmental Communication* 9:189–207.

Conover, Michael D, Jacob Ratkiewicz, Matthew Francisco, Bruno Gonçalves, Filippo Menczer and Alessandro Flammini. 2011. “Political polarization on twitter.”.

Delli Carpini, Michael X and Scott Keeter. 1993. “Measuring Political Knowledge: Putting First Things First.” *American Journal of Political Science* pp. 1179–1206.

- Fiorina, Morris. 2014. Americans Have Not Become More Politically Polarized. Technical report Washington Post.
- URL:** <https://www.washingtonpost.com/news/monkey-cage/wp/2014/06/23/americans-have-not-become-more-politically-polarized/>
- Gaines, Brian J, James H Kuklinski, Paul J Quirk, Buddy Peyton and Jay Verkuilen. 2007. “Same Facts, Different Interpretations: Partisan Motivation and Opinion on Iraq.” *Journal of Politics* 69:957–974.
- Garcia, David, Adiya Abisheva, Simon Schweighofer, Uwe Serdült and Frank Schweitzer. 2015. “Ideological and Temporal Components of Network Polarization in Online Political Participatory Media.” *Policy & Internet* 7:46–79.
- Halberstam, Yosh and Brian Knight. 2016. “Homophily, Group Size, and the Diffusion of Political Information in Social Networks: Evidence from Twitter.” *Journal of Public Economics* 143:73–88.
- Hmielowski, Jay D., Myiah J. Hutchens and Vincent J. Cicchirillo. 2014. “Living in an Age of Online Incivility: Examining the Conditional Indirect Effects of Online Discussion on Political Flaming.” *Information, Communication & Society* 17:1196–1211.
- URL:** <http://dx.doi.org/10.1080/1369118X.2014.899609>
- Iyengar, Shanto and Kyu S Hahn. 2009. “Red media, blue media: Evidence of ideological selectivity in media use.” *Journal of Communication* 59:19–39.
- Kramer, Adam DI, Jamie E Guillory and Jeffrey T Hancock. 2014. “Experimental Evidence of Massive-Scale Emotional Contagion Through Social Networks.” *Proceedings of the National Academy of Sciences* 111:8788–8790.
- Ksiazek, Thomas B, Limor Peer and Andrew Zivic. 2015. “Discussing the News: Civility and Hostility in User Comments.” *Digital Journalism* 3:850–870.

- Kunda, Ziva. 1990. "The Case for Motivated Reasoning." *Psychological Bulletin* 108:480.
- Kwon, K Hazel and Daegon Cho. 2017. "Swearing Effects on Citizen-to-Citizen Commenting Online: A Large-Scale Exploration of Political Versus Nonpolitical Online News Sites." *Social Science Computer Review* 35:84–102.
- Lapidot-Lefler, Noam and Azy Barak. 2012. "Effects of Anonymity, Invisibility, and Lack of Eye-Contact on Toxic Online Disinhibition." *Computers in Human Behavior* 28:434–443.
- Lee, Eun-Ju. 2012. "That's Not the Way It Is: How User-Generated Comments on the News Affect Perceived Media Bias." *Journal of Computer-Mediated Communication* 18:32–45.
- Lelkes, Yphtach, Gaurav Sood and Shanto Iyengar. 2017. "The Hostile Audience: The Effect of Access to Broadband Internet on Partisan Affect." *American Journal of Political Science* 61:5–20.
- Levendusky, Matthew S. 2013. "Why Do Partisan Media Polarize Viewers?" *American Journal of Political Science* 57:611–623.
- Lodge, Milton and Charles S Taber. 2005. "The Automaticity of Affect for Political Leaders, Groups, and Issues: An Experimental Test of the Hot Cognition Hypothesis." *Political Psychology* 26:455–482.
- Lodge, Milton and Charles S Taber. 2013. *The Rationalizing Voter*. Cambridge University Press.
- Messing, Solomon and Sean J Westwood. 2014. "Selective Exposure in the Age of Social Media: Endorsements Trump Partisan Source Affiliation when Selecting News Online." *Communication Research* 41:1042–1063.
- Miller, Joanne M, Kyle L Saunders and Christina E Farhart. 2016. "Conspiracy Endorsement as Motivated Reasoning: The Moderating Roles of Political Knowledge and Trust." *American Journal of Political Science* 60:824–844.

- Moosa, Tariq. 2014. "Comment Sections are Poison: Handle With Care or Remove Them." *The Guardian* .
- News Consumption Survey*. 2012. *Pew Research Centre* .
- URL:** <http://www.people-press.org/2012/06/03/2012-media-consumption-survey/>
- Papacharissi, Zizi. 2004. "Democracy Online: Civility, Politeness, and the Democratic Potential of Online Political Discussion Groups." *New media & society* 6:259–283.
- Papacharissi, Zizi and Maria de Fatima Oliveira. 2012. "Affective News and Networked Publics: The Rhythms of News Storytelling on # Egypt." *Journal of Communication* 62:266–282.
- Pariser, Eli. 2011. *The Filter Bubble: What the Internet is Hiding From You*. Penguin UK.
- Quattrociocchi, Walter, Antonio Scala and Cass R Sunstein. 2016. "Echo Chambers on Facebook." *Available at SSRN* .
- URL:** <https://ssrn.com/abstract=2795110>
- Redlawsk, David P. 2002. "Hot Cognition or Cool Consideration? Testing the Effects of Motivated Reasoning on Political Decision Making." *Journal of Politics* 64:1021–1044.
- Rowe, Ian. 2015. "Civility 2.0: A Comparative Analysis of Incivility in Online Political Discussion." *Information, Communication & Society* 18:121–138.
- Santana, Arthur D. 2014. "Virtuous or Vitriolic: The Effect of Anonymity on Civility in Online Newspaper Reader Comment Boards." *Journalism Practice* 8:18–33.
- Stoker, Laura. 1993. "Judging Presidential Character: The Demise of Gary Hart." *Political Behavior* 15:193–223.
- Stroud, Natalie Jomini. 2010. "Polarization and Partisan Selective Exposure." *Journal of Communication* 60:556–576.

- Suhay, Elizabeth, Allyson Blackwell, Cameron Roche and Lucien Bruggeman. 2015. "Forging Bonds and Burning Bridges: Polarization and Incivility in Blog Discussions About Occupy Wall Street." *American Politics Research* 43:643–679.
- Suler, John. 2004. "The Online Disinhibition Effect." *Cyberpsychology & Behavior* 7:321–326.
- Sunstein, Cass R. 2009. *Going to Extremes: How Like Minds Unite and Divide*. Oxford University Press.
- Taber, Charles S and Milton Lodge. 2006. "Motivated Skepticism in the Evaluation of Political Beliefs." *American Journal of Political Science* 50:755–769.
- Yardi, Sarita and Danah Boyd. 2010. "Dynamic Debates: An Analysis of Group Polarization Over Time on Twitter." *Bulletin of Science, Technology & Society* 30:316–327.

Tables - Figures

Obama announces gun actions in emotional plea for congressional action

WASHINGTON — President Obama announced a series of executive actions on guns Tuesday, focusing on the victims of gun violence in a White House event intended to prod Congress to take further action.

The executive actions — first previewed by the White House on Monday — would attempt to clamp down on unlicensed gun sellers who exploit an exception for hobbyists and collectors in order to avoid having to run criminal background checks on gun purchasers. Many of those sellers, Obama said, are running a business by selling guns at gun shows and online.

"The problem is that some gun sellers have been operating by a different set of rules," he said. "That doesn't make sense."

Obama said the administration is also beefing up enforcement, streamlining the background check system, investing \$500 million in mental health care and researching "smart gun" technology.

Obama said the actions he's taking are consistent with gun rights.

"I believe in the Second Amendment. It's there written on paper. It guarantees the right to bear arms. No matter how much people try to twist my words around, I taught constitutional law. I know a little bit about this. I get it," he said. "This is not a plot to take away everybody's guns."

Congressional Republicans had a mixed reaction to Obama's announcement. "The president's actions are out of bounds and vastly exceed his executive authority," said House Oversight Committee Chairman Jason Chaffetz, R-Utah, who promised vigorous oversight hearings. But House Majority Leader Kevin McCarthy, R-Calif., called the additional guidance on firearms licenses a "weak gesture" that falls short of what Obama really wanted to accomplish.

Figure 1: Text of Gun Control Article

U.S., allies strike Pacific Rim trade deal

The United States and 11 other Pacific Rim nations struck a tentative trade agreement Monday, a landmark deal that has the potential to transform the global economy, divide political parties in Congress and roil the U.S. presidential race.

As President Obama and aides began selling the agreement to Congress and the public, critics denounced it as yet another free-trade deal that will help ship American jobs overseas.

The Trans-Pacific Partnership will "promote economic growth" and "support higher paying jobs," said U.S. Trade Rep. Michael Froman, making the announcement along with other trade ministers in Atlanta after days of negotiations.

The massive proposed agreement — which faces months of debate in Congress — would tie together nearly 40% of the world's economy, from Canada to Chile to Japan to Australia; it would be the largest regional trade agreement in history.

In hailing the agreement, Obama said, "Congress and the American people will have months to read every word" before he signs the deal that he described as a win for all sides.

"If we can get this agreement to my desk, then we can help our businesses sell more Made in America goods and services around the world, and we can help more American workers compete and win," Obama said.

Critics said employers will use the agreement to move jobs to poorer countries that have lower wages and fewer regulations.

Arthur Stamoulis, executive director of the Citizens Trade Campaign, predicted Congress would reject the deal, especially on the eve of an election year.

"Heading into 2016 and beyond, Congress members know that American voters are not going to accept a massive trade agreement with undemocratic countries that offshores jobs and drives down wages," he said.

In pledging to work with Congress, Obama said the pact includes "the strongest commitments on labor and the environment of any trade agreement in history" and removes trade barriers that have blocked U.S. products.

Figure 2: Text of TPP Article

Table 1: Summary Statistics on Participants

	Mean						p
	All	Treatment	Control	Diff	CI low	CI high	
Age	36.00	35.91	36.10	0.20	-1.68	2.07	0.84
Female	0.56	0.53	0.59	0.06	-0.15	0.03	0.20
Political knowledge	0.49	0.51	0.46	-0.05	-0.04	0.14	0.30
Democrats	0.50	0.53	0.46	-0.07	-0.02	0.16	0.12
Republicans	0.17	0.15	0.18	0.03	-0.17	0.07	0.47
Independents	0.30	0.28	0.33	0.04	-0.15	0.04	0.30
Prior support for policy	0.54	0.55	0.53	-0.02	-0.06	0.11	0.64
N	548.00	281.00	267.00				

Note: Columns ‘All’, ‘Treatment’ and ‘Control’ report the mean value of each variable for the respective groups. Column ‘Diff’ reports the difference between the Treatment and Control group means. Columns ‘CI low’ and ‘CI high’ report the end points of the 95% confidence interval of the difference in means. Column ‘p’ reports the p-value of the difference in means when tested against a null hypothesis of zero difference.

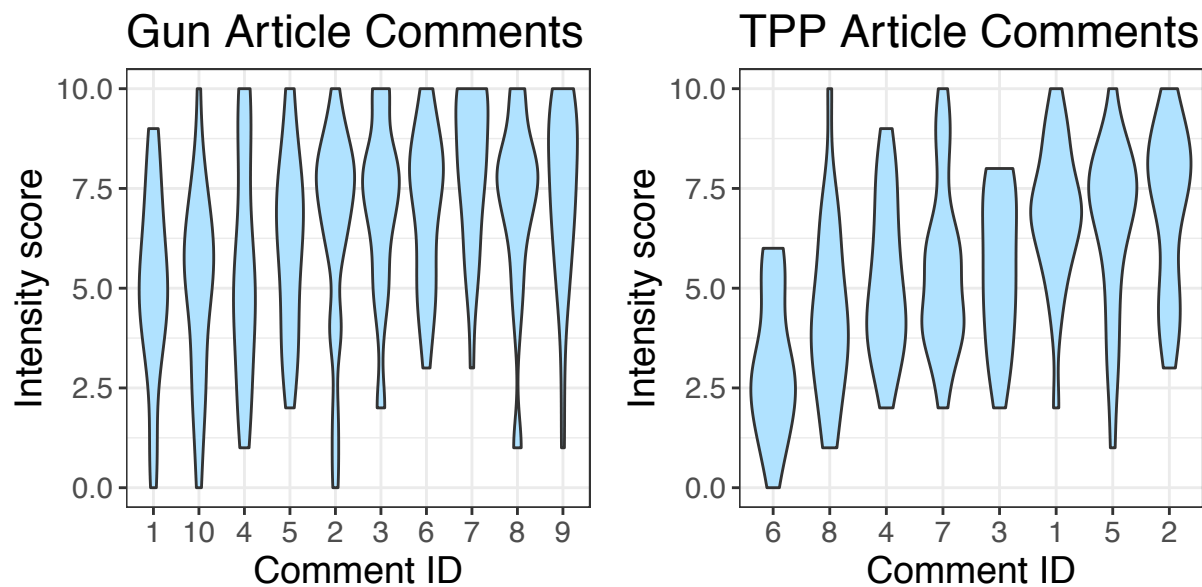


Figure 3: Distributions of intensity coding for comments

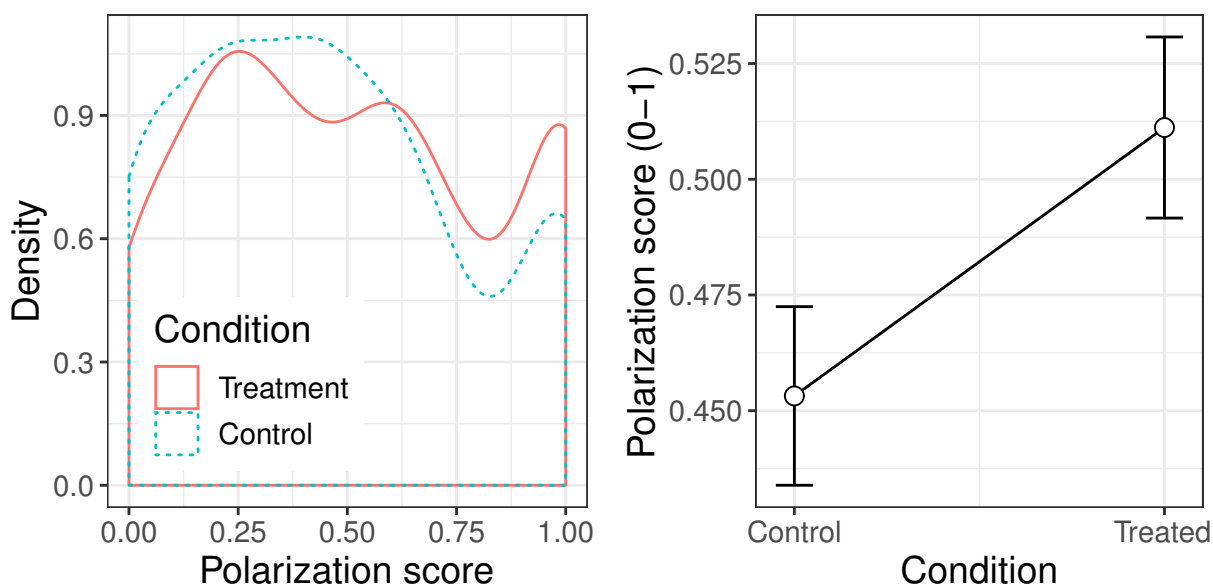


Figure 4: Difference between treatment and control group in their polarization score.

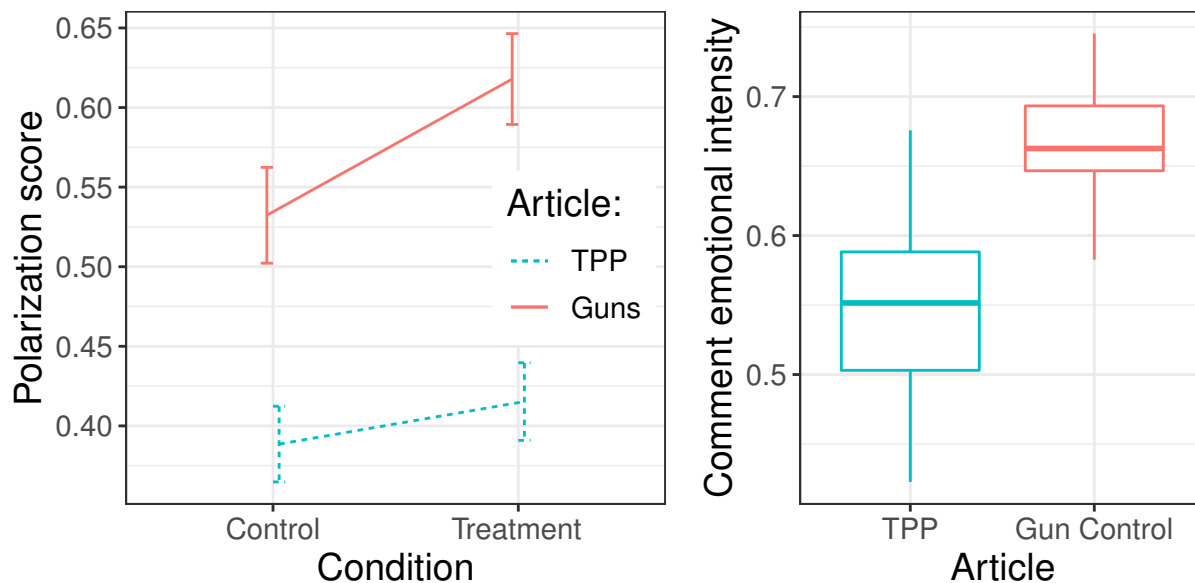


Figure 5: Effect of treatment and intensity of comments by article

Note: The first panel decomposes the average treatment effect, between the two articles. Vertical bars denote the 95% CIs. The second panel shows the two boxplots that summarize the distribution of emotional intensity among all comments for each article separately.

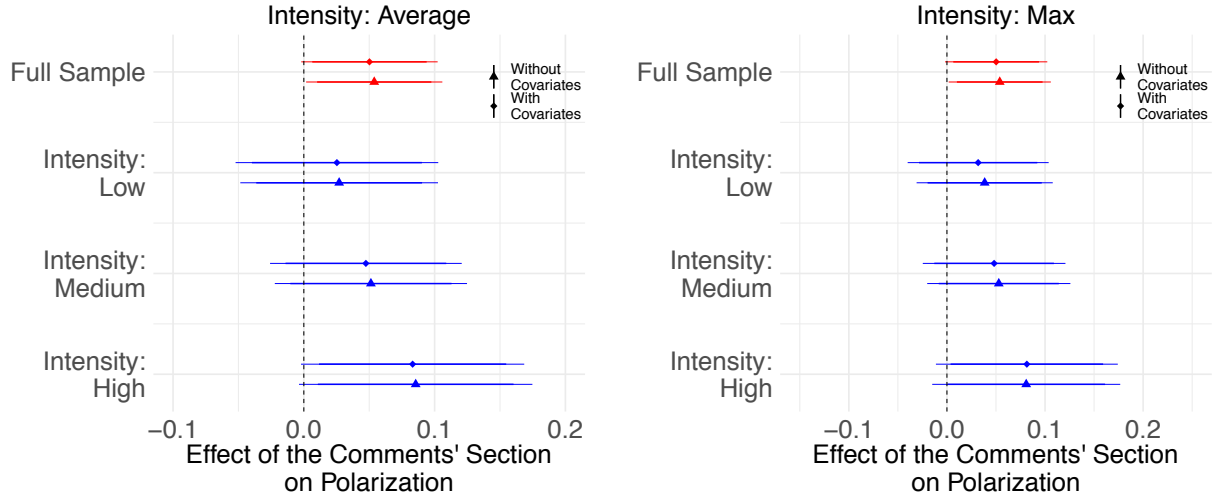


Figure 6: Difference between treatment and control group in their polarization score, according to: a) the overall level of emotional intensity (panel A) and b) the emotional intensity of the most-emotionally intense comment. Entries are average differences in opinion polarization between treatment and control group, with (dashed line) and without (solid line) covariates. Horizontal lines denote the 90% (thick line) and 95% (thin line) confidence intervals.

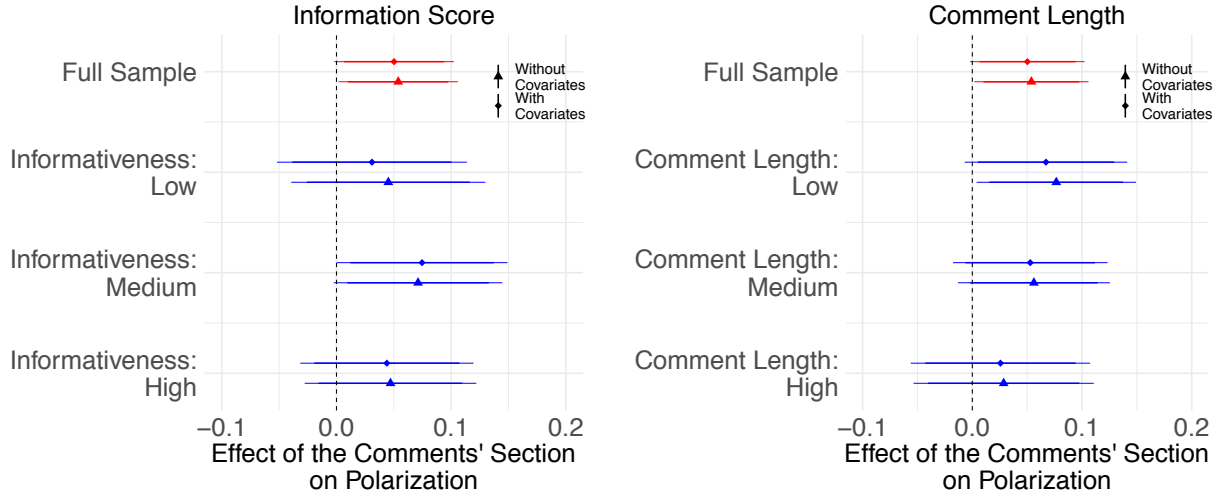


Figure 7: Difference between treatment and control group in their polarization score, according to: a) the amount of information displayed by each comment (panel a); and b) the length of each comment. Entries are average differences in opinion polarization between treatment and control group, with (dashed line) and without (solid line) covariates. Horizontal lines denote the 90% (thick line) and 95% (thin line) confidence intervals.

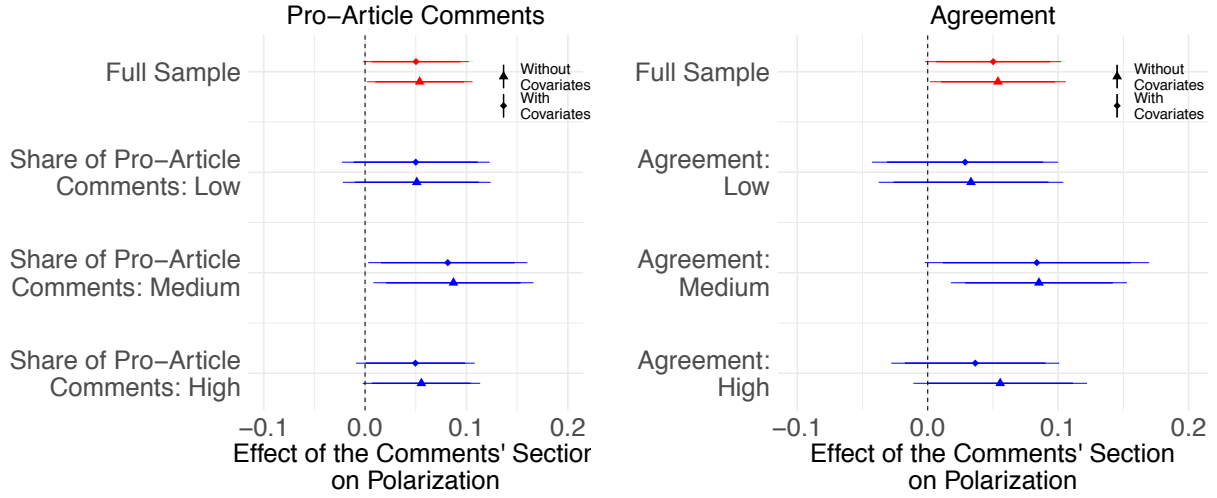


Figure 8: Difference between treatment and control group in their polarization score, according to: a) the share of pro-policy comments; and b) the degree of agreement between respondent and user-comments. Entries are average differences in opinion polarization between treatment and control group, with (dashed line) and without (solid line) covariates. Horizontal lines denote the 90% (thick line) and 95% (thin line) confidence intervals.

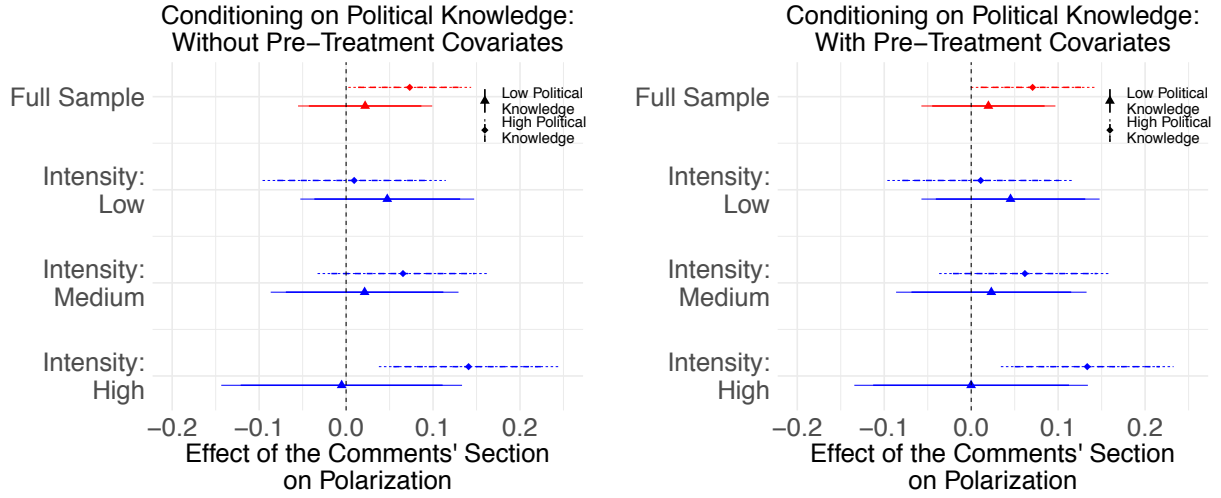


Figure 9: Difference between treatment and control group in their polarization score, according to individuals' level of political knowledge and the level of emotional intensity of the user comments. Entries are average differences in opinion polarization between treatment and control group, for individuals with low (dashed line) and high (solid line) political knowledge. Horizontal lines denote the 90% (thick line) and 95% (thin line) confidence intervals. The first panel includes only a dummy to disintinguish the two experiments, while the second panel includes also individual pre-treatment characteristics.