Better the Devil You Know:

Selective Exposure Alleviates Polarization in an Online Field Experiment

Donghee Jo*

March 3, 2020

Abstract

This paper empirically investigates the causal link between news consumers' self-selective exposure to like-minded partisan media and polarization. Through an online experiment, I find a lower likelihood of extreme policy views—an alleviation of polarization—among the subjects who were allowed to choose their like-minded partisan media than those who were not. I begin the paper by presenting a parsimonious model to formalize a traditionally neglected channel through which self-selective media choice leads to reduced polarization. The predictions of this model are supported by experimental evidence collected from a South Korean mobile news application that I created. Users of the app were given access to curated articles on key political issues and were regularly asked about their views on those issues. Some randomly selected users were allowed to choose the news sources from which to read articles; others were given randomly selected articles. The users who selected their news sources showed larger changes in their policy views and were less likely to have radical policy views—an alleviation of polarization—in comparison with those who read randomly provided articles. In support of the main mechanism in the model, an exogenous increase in familiarity enables news consumers to better adjust for the biases of news articles, and this in turn leads to more learning and less extreme policy views.

JEL Codes: D83, D91.
Keywords: Selective Exposure, Political Polarization, Online Experiment, Political Economy.

*Northeastern University. Contact: d.jo@northeastern.edu, 360 Huntington Avenue LA 301, Boston, MA 02115, (617) 373-8187. I truly appreciate the guidance of my advisors, Ben Olken, Daron Acemoglu, Esther Duflo, and Abhijit Banerjee, without whom this paper would not have been possible. Funding was provided by the George and Obie Shultz Fund and JPAL Incubator Funding. I am grateful to Manseok Jo, Dohyeon Kim, Young Jae Woo, and Jaehoon Yang, who helped build and operate a mobile application for the experiment. I also thank Enrico Cantoni, Bilge Erten, John Firth, Chishio Furukawa, Arda Gitmez, Nick Hagerty, Greg Howard, Peter Hull, Gabriel Kreindler, Matt Lowe, Rachael Meager, Yuhei Miyauchi, Arianna Ornaghi, Frank Schilbach, Cory Smith, Michael Wong, Muhamet Yildiz, and participants in the MIT Political Economy Seminar for helpful comments.

1 Introduction

> *"Increasingly we become so secure in our bubbles that we start accepting only information, whether it's true or not, that fits our opinions, instead of basing our opinions on the evidence that is out there." – President Barack Obama, in his farewell address, January 10, 2017*

Political polarization is on the rise, and many find this trend troubling.[1] Politicians and academics alike blame the media's adoption and dissemination of political biases, as well as the public's self-selected exposure to like-minded media, as the main drivers of political polarization. President Trump tweeted that some liberal media are "the enemy of the American people" for their coverage of "fake news." As quoted above, President Obama described people as being in their "bubbles" where we only accept information "that fits our opinions." Academics seem to agree. People are reportedly in their own "echo chambers" or "filter bubbles," exposing themselves to the information sources that are likely to confirm their preexisting views (Sunstein, 2001; Prior, 2007; Pariser, 2011). As such, many experts advocate enhancing chance encounters—random and balanced media exposure—to counteract the accelerating trend of polarization (e.g., Sunstein, 2017).[2]

In contrast to this view, this paper concerns a traditionally neglected channel through which selective exposure to partisan media can actually contribute to an alleviation—not exacerbation—of polarization. In a world where the media distort signals, familiarity with media biases is vitally important. By choosing like-minded partisan media, news consumers are exposed to familiar news sources through which they can arrive at better estimates of the underlying truth. In other words, selective exposure can contribute to an alleviation of polarization via facilitated learning.

In this paper, I (i) formalize this idea in a simple model, (ii) empirically demonstrate the causal link using a South Korean mobile news application (henceforth "app") that I created for an online field experiment, and (iii) investigate the evidence on the suggested mechanism. The results of the experiment show that selective exposure alleviates partisan polarization of policy views, likely via the suggested mechanism.

[1] For evidence on rising polarization, see Hetherington (2001); Jensen et al. (2012); Pew Research Center (2014); Iyengar and Westwood (2015); Gentzkow et al. (2016); Gentzkow (2016); McCarty et al. (2016); Boxell et al. (2017). Political polarization has increasingly infiltrated personal lives (YouGov, 2008; Iyengar et al., 2012; Iyengar and Westwood, 2015).

[2] Entrepreneurs and some news media organizations have responded to these requests. For example, there are applications such as *Read Across the Aisle* or *Flipfeed* that help news readers obtain balanced chance encounters. *The Wall Street Journal* provides a comparison between the Facebook news feeds of liberals and conservatives. *The New York Times* has a series, "Right and Left: Partisan Writing You Shouldn't Miss," that covers both left-wing and right-wing articles on a designated topic.

To begin, I present a parsimonious Bayesian model in which the agent's objective is to learn the true state of the world by selecting a news source for gaining information. The selected news source sends a contaminated signal: a mixture of the truth and the source's media bias. Since the media bias of a familiar source is better understood by the agent, the model predicts selection of a relatively familiar source. The most important and perhaps surprising prediction of the model is that receiving a signal from a relatively familiar source, compared with receiving it from a randomly chosen source, alleviates polarization. Specifically, agents who are allowed to select their news sources learn more about the truth as long as they have approximately correct assessments of the sources' media biases. This is more likely to be true for the agents who begin with especially erroneous policy views. Note that increased familiarity can be one of the pathways toward greater trust in a news source and, therefore, this model can be considered as a microfoundation of the potentially nebulous concept of "trust."

I then test the key implications of the model using experiments conducted in the aforementioned mobile app. The users of the app were given access to curated articles on key political issues in South Korea, and were regularly asked about their views on the issues.[3] I explore two types of random variation regarding the user experience. First, every user ($N_1 = 1420$) was given randomly chosen articles during at least a partial period of the experiment. This gives random variation in the characteristics of the articles and their sources. Second, the users who remained in the app past five rounds of reading articles ($N_2 = 367$) were divided into several treatment groups. Some users were allowed to choose the news source from which to read an article while others were given randomly selected articles. Furthermore, when allowed to choose their sources, readers were shown in the selection screen either (a) the names of the news sources, (b) the news sources' average positions on the issue at hand, or (c) both.

The belief updating patterns observed in the data are consistent with the predictions of general Bayesian models with heterogeneous priors. The reader's view reported after reading an article is affected both by the article's political position and the view reported prior to reading the article, and is less influenced by the article if she reported high confidence in her view prior to reading the article. For this analysis, I use the first type of random variation: random provision of news articles.

Next, I find evidence of selective exposure among the users who were allowed to choose their sources based on the sources' names. They chose news sources that were likely to represent the

[3] South Korea provides an ideal context to conduct low-cost online field experiments due to its mobile environment that is considered one of the best in the world. It also shares similar political and journalistic challenges with other developed democratic countries, such as political polarization and the prevalence of misinformation campaigns.

views of the parties that the users supported. This could be due to users' selecting based on their partisan preferences, as is assumed by most of the literature, or other factors like familiarity. The amount of selective exposure found in the data is comparable to that of the US in terms of the isolation index defined in Gentzkow and Shapiro (2011).

Most importantly, selective exposure to partisan media facilitates learning and reduces the likelihood of extreme policy views, consistent with my model's prediction. The distance between the users' positions reported before and after reading an article is 17% larger ($p < 0.05$) for users who could select based on source names compared with users who were given randomly selected articles. Users who could select based on source names also show a greater decline in their proportion of extreme policy views (a decline of 7 additional percentage points, $p < 0.05$).[4] The results suggest that selective exposure to partisan media causes an alleviation of attitude polarization, contrary to the traditional view.

I find evidence that the suggested channel—familiarity facilitates learning and reduces polarization—is one of the underlying mechanisms of how media selection results in moderated policy views. Every user was initially given randomly chosen articles from randomly chosen news sources for five days. This treatment randomly increased users' familiarity with the specific news sources they were allocated in the initial period. I find that the users better adjust for the news source's typical position on issues when updating their policy views—a sign of signal de-biasing—when they were given articles from news sources to which they had been pre-exposed. Furthermore, among users who were given randomly selected articles, those who were randomly re-assigned to pre-exposed news sources have larger changes in their policy views and greater declines in the proportion of extreme policy views compared with those who were randomly assigned to the news sources without pre-exposure. Although increased familiarity can be considered one of the theoretical pathways toward enhanced media trust, I provide evidence that self-reported trust alone does not explain the main results that I find in this paper.

I also find evidence against some alternative explanations, such as (i) differences in characteristics of the articles given, (ii) differences in engagement, (iii) general differential movement toward right-wing positions, and (iv) the role of selection itself on user attitudes.

I also explore alternative interpretations of the results. Although reduced extremism can theoretically be accompanied by strengthened partisan identification—an exacerbation of polarization—I do not find evidence supporting this. Finally, I conduct an analysis using a subsample as an alternative counterfactual and find similar main results.

[4] Policy views are considered to be extreme if the reader reported a view that is farthest to the left or right of the policy view spectrum bar in the app. See Section 3 to see how readers reported their policy views.

In a world where news curation services are ever more widespread (e.g., social network services such as Facebook and Twitter), this paper provides a easily applied insight relevant to both designers of curation algorithms and to regulators: When providing curated news articles, making the names and other characteristics of the news sources salient can help the news consumers make decisions about which articles to read, which in turn may facilitate learning.

Related Literature. The results in this paper speak to a rapidly growing literature that studies political polarization. There is abundant descriptive evidence about the increasing trend of polarization of political elites (Poole and Rosenthal, 1991; Jensen et al., 2012; Gentzkow et al., 2016; McCarty et al., 2016) and the public (Hetherington, 2001; YouGov, 2008; Iyengar et al., 2012; Pew Research Center, 2014; Iyengar and Westwood, 2015; Gentzkow, 2016; Boxell et al., 2017), and, naturally, there has been much conjecture about the causes of polarization. In particular, there has been speculation that selective exposure to like-minded media is one of the main drivers of political polarization (Sunstein, 2001), and partisan media have been shown to have the potential to change the behaviors of voters (DellaVigna and Kaplan, 2007; Martin and Yurukoglu, 2017; Adena et al., 2015; Durante et al., 2015).[5] There is also abundant descriptive evidence on the prevalence of selective exposure (Iyengar and Hahn, 2009; Gentzkow and Shapiro, 2011; Flaxman et al., 2016; Halberstam and Knight, 2016). However, rigorous causal evidence linking selective exposure and the public's partisan polarization is limited.

The closest studies to this paper are Boxell et al. (2017) and De Benedictis-Kessner et al. (2019). Boxell et al. (2017) reject the claim that the Internet, which makes selective exposure to like-minded media extremely easy, is a direct cause of rising polarization. The main evidence for their rejection is that the over-75 age group, unlikely subscribers to information technology, experienced the sharpest rise in attitude polarization. In this paper, I provide more rigorous causal evidence on the relationship between selective exposure and attitude polarization. In a survey experiment, De Benedictis-Kessner et al. (2019) study the causal effect of reading an article about Marijuana legalization in the US, focusing on identifying heterogeneous effects between populations of different media preference measured by asking people to choose between Fox News, MSNBC, and the Food Network at the baseline. In this paper, I instead focus on the causal effect of having different media choice environments and its mechanisms in a real world news consumption setting with variety of highly partisan issues and realistic media choice sets.

[5] While media markets have been studied in the context of political bias, the focus has mostly been on the supply side (Mullainathan and Shleifer, 2005; Gentzkow and Shapiro, 2006, 2010) or on political elites (Campante and Hojman, 2013), rather than on consumers of news articles and the evolution of their political views. A notable exception is Chopra et al. (2019).

The model in this paper derives from Sethi and Yildiz (2016a,b). Sethi and Yildiz (2016a,b) build on the insight of Acemoglu et al. (2016), who demonstrate the importance of understanding the receiver's belief about the messenger. Sethi and Yildiz (2016a,b) focus on studying the trade-off between expertise and familiarity in the selection of information sources, where familiarity is defined similarly as in this paper's model. Other related theoretical papers include Calvert (1985) and Suen (2004), who detail the theoretical foundations of rational selective exposure. In these papers, agents choose like-minded media to maximize the value of information in a limited number of signals, similar to the model in this paper. The difference is that the value of information in Calvert (1985) and Suen (2004) is driven by the discreteness of the actions in their models—the information is valuable only insofar as it can change the actions of the agents.

The mechanism emphasized in this paper is closely related to Chiang and Knight (2011), who provide convincing empirical evidence of bias deduction by news consumers. Relatedly, Durante and Knight (2012) also show that news consumers actively take into account media biases when selecting TV channels to watch.

The remainder of the paper proceeds as follows. Section 2 presents a parsimonious Bayesian model and its implications. Section 3 describes the setting, experimental design, and data. Section 4 provides the main results of this paper. Section 5 explores the mechanisms and alternative interpretations, and Section 6 concludes.

2 Theoretical Framework

In this section, I provide a theoretical framework to formally demonstrate a potential mechanism through which unregulated media selection leads to reduced polarization. The model not only produces testable predictions but also sheds light on the generalizability of the results of this paper by providing conditions under which this mechanism can operate.

2.1 Model

For each policy issue j, there is a common truth $\theta_j^* \in \mathbb{R}$, and the agents hold their own subjective beliefs about θ_j^* at the beginning of the period. This prior belief can be represented by a normal distribution with support $\mathbb{R}$: $\theta_j \underset{i}{\sim} N\left(\theta_{ij0}, \frac{1}{\tau_{ij0}}\right)$. θ_{ij0} is the mean, and τ_{ij0} is the precision of the prior belief. Higher θ_j means a more right-wing view on the issue. Higher τ_{ij0} means that the agent has greater confidence in her prior belief.

For the supply side, assume that there are two news sources, $p \in \{F, A\}$. F represents a familiar news source (e.g., *The New York Times* for liberals or Fox News for conservatives), and A represents an alien one.

A signal that news source p sends to agent i on issue j is given by:

$$s_{pij} = \theta_j^* + I_{pj}^* + \varepsilon_{ij} \tag{1}$$

The signal that p sends (s_{pij}) is a sum of θ_j^*—the common truth—and I_{pj}^*. I_{pj}^* is the "media bias" added to the truthful signal. It is the position p advocates. For example, a right-wing source can be thought of as having positive I_{pj}^*, adding right-wing biases to the articles it publishes. The addition of media bias (I_{pj}^*) can be achieved by omission of evidence or by careful control of tone or voice. I_{pj}^* is not directly observable to i, and each agent holds a subjective belief about it: $I_{pj} \underset{i}{\sim} N\left(I_{pji}, \frac{1}{\tau_{Ipj}}\right)$. I_{pji} is the mean, and τ_{Ipj} is the precision of the belief about I_{pj}^*. Higher τ_{Ipj} means that the agent has a better understanding of the bias p adds to the signal—the "style" of the news source is familiar to i. The level of understanding of the media bias (τ_{Ipj}) can be thought of as a microfoundation of the agent's "trust" in the news source. Having accurate expectations about the tone, intention, or bias of a news source enables the agent to learn more about the issue at hand (see below). This may in turn lead to higher trust.[6]

$\varepsilon_{ij} \sim N\left(0, \frac{1}{\tau_s}\right)$ is the idiosyncratic noise that is added to the signal. τ_s is the precision of the signal, and it represents the journalistic competence of the source in collecting and conveying information. It also includes the agent's ability to understand the signal. For simplicity, these idiosyncratic noise terms are assumed to have the same precision (τ_s).

In this section, I analyze a single-period belief updating with a timeline as follows.[7] An issue (j) is randomly selected by nature. The agent selects a news source, and she gets a contaminated signal—s_{pij} as in Equation (1)—from the chosen source. The agent updates her belief about θ_j^*

[6]There is no clear consensus about the definition of trust. One of the most relevant definitions of trust is that it "(...) entails a state of perceived vulnerability or risk that is derived from individuals' uncertainty regarding the motives, intentions, and prospective actions of others on whom they depend" (Kramer, 1999).

[7]A long-run extension of this model is discussed in Section 2.2.

using Bayes' rule, and chooses an action $a_{ij} \in \mathbb{R}$ to maximize utility:[8]

$$U(a_{ij} \mid \theta_j^*) = -(a_{ij} - \theta_j^*)^2 \tag{2}$$

This is equivalent to a setting where the agent genuinely tries to learn the truth as accurately as possible. In the language of political psychology, agents have *accuracy* motives, instead of *directional* motives (Leeper and Slothuus, 2014).

Understanding the selection of the news source requires backward induction—let us begin by assuming that news source p is selected by i to get a signal about issue j. The posterior distribution after getting a signal, s_{pij} is:[9]

$$\theta_j \,|\, s_{pij} \underset{i}{\sim} N\left(\theta_{ij1}\,,\, \sigma^2_{ij1}\right) \tag{3}$$

where

$$\theta_{ij1} = \frac{\tau_s \tau_{ij0} + \tau_{ij0}\tau_{Ipj}}{\tau_s \tau_{ij0} + \tau_{ij0}\tau_{Ipj} + \tau_{Ipj}\tau_s}\theta_{ij0} + \frac{\tau_{Ipj}\tau_s}{\tau_s \tau_{ij0} + \tau_{ij0}\tau_{Ipj} + \tau_{Ipj}\tau_s}\left(s_{pij} - I_{pji}\right) \tag{4}$$

$$\sigma^2_{ij1} = \frac{\tau_s + \tau_{Ipj}}{\tau_s \tau_{ij0} + \tau_{ij0}\tau_{Ipj} + \tau_{Ipj}\tau_s} \tag{5}$$

As can be seen in Equation (4), the mean of the posterior belief, θ_{ij1}, is a convex combination of the prior mean (θ_{ij0}) and the bias-deducted signal ($s_{pij} - I_{pji}$). $\frac{\tau_{Ipj}\tau_s}{\tau_s \tau_{ij0} + \tau_{ij0}\tau_{Ipj} + \tau_{Ipj}\tau_s}$ is the weight on the signal, and the remaining weight is on the prior. Propositions 1 and 2 describe this and provide comparative statics about the weight on the signal.

Proposition 1. *(Bayesian Updating) The posterior mean is a function of the prior mean (θ_{ij0}) and the bias-deducted signal ($s_{pij} - I_{pji}$). The weight on the signal is decreasing in τ_{ij0}.*

Proof. For the first sentence, see Equation (4). It is then straightforward to show that

[8] Although I assume common truth here, the citizens' utility-maximizing bliss points can vary in reality. We can extend the model to incorporate this by adding a term, b_{ij}, that represents individual characteristics to the utility function:

$$U(a_{ij} \mid \theta_j^*, b_{ij}) = -(a_{ij} - \theta_j^* - b_{ij})^2$$

Then the bliss point for i on issue j becomes $\theta_j^* + b_{ij}$. The model's predictions are not too different, as long as $|b_{ij}|$ is sufficiently small so that extreme positions (defined in Section 3) are still far from the optimal positions. I qualitatively argue that this is likely to be true in the context of the experiment conducted for this paper, given the definition of "extreme positions" (see Section 3.5).

[9] See Appendix A for the derivation. I assume that the prior distribution of I_{pj} and θ_j are independent. This implies that the prior information that i has on issue j is not from source p. Relaxing this assumption does not affect the main results of this paper—although it complicates the algebra.

$\frac{\partial}{\partial \tau_{ij0}} \left(\frac{\tau_{Ipj}\tau_s}{\tau_s\tau_{ij0}+\tau_{ij0}\tau_{Ipj}+\tau_{Ipj}\tau_s} \right) < 0.$ □

Proposition 2. *(Familiarity and Updating) The weight on the signal, used to form the posterior mean, is increasing in τ_{Ipj}.*

Proof. It is straightforward to show that $\frac{\partial}{\partial \tau_{Ipj}} \left(\frac{\tau_{Ipj}\tau_s}{\tau_s\tau_{ij0}+\tau_{ij0}\tau_{Ipj}+\tau_{Ipj}\tau_s} \right) > 0.$ □

The prediction that the posterior mean is a function of the prior and the signal is consistent with typical Bayesian models. The prediction that the weight on the signal decreases in prior confidence (τ_{ij0}) is also consistent with typical Bayesian models. The prediction on τ_{Ipj} (Proposition 2) is a unique feature of this model. The signal has greater influence on the posterior belief if it is from a source that the agent understands better. The predictions in Propositions 1 and 2 are empirically verified in later sections of this paper.

Now that we understand the final stage—belief updating—we can analyze the preceding stage: the agent's source selection. For this, I make the following assumption.

Assumption 1. *(Familiarity) $\forall j$, $\tau_{IFj} > \tau_{IAj}$*

This assumption indicates that the agent has a more accurate assessment of the bias of the familiar source than of the alien one. Theoretically, this assumption has been explored in Sethi and Yildiz (2016a,b). Empirically, it has been shown that people have more nuanced knowledge about the biases of in-group news sources compared to that of out-group sources (Stroud et al., 2014).[10] The assumption implies that a signal from the familiar source has better information value for i, which makes her choose the familiar source as described in Proposition 3.

Proposition 3. *(Selective Exposure) Suppose Assumption 1 is true. The source for which the agent has higher τ_{Ipj}—the familiar source—will be chosen.*

Proof. When τ_{Ipj} is higher, the signal is easier for the agent to interpret, making it more informative. For formal proof, see Appendix A. □

[10]Specifically, people appear to understand the media-specific bias of in-group sources, whereas they do not know the differences in biases between out-group sources. This is often called "out-group homogeneity." Formally, this can be directly linked to Assumption 1 if we take a Bayesian hierarchical model in which an individual media outlet has a separate bias I_{mj}, which is known to be a realization of $N(I_{Fj}, \tau)$ or $N(I_{Aj}, \tau)$—depending on whether the source is in-group (F) or out-group (A). If i has taken more signals from familiar sources ($m \in F$), she must have a better idea about the distribution of I_{mj}. On the other hand, if i has taken virtually no signals from alien sources, when asked about her best estimation of I_{mj} ($m \in A$), she will simply say I_{Aj} regardless of m, making her appear to have no understanding of the nuanced differences between out-group sources. This gives Assumption 1. Stroud et al. (2014) show that there is a positive correlation between media familiarity and the nuanced understanding of the media's bias.

In the mobile app experiment described in this paper, some subjects were allowed to choose their news sources and are predicted to choose the familiar source (Proposition 3). The remaining subjects who could not choose their news sources can be regarded as encountering the familiar source and the alien source with equal probability. The propositions below predict the difference in the evolution of beliefs between these two groups.

For Proposition 4, I also make the following assumption.

Assumption 2. *(Approximately Correct Bias Assessment)* $I_{pji} = I^*_{pj} + \nu_{pji}$, *where* $\left|\nu_{pji}\right| < M_j \frac{1}{\gamma_{pij}}$ *for some* $M_j > 0$, *where* $\gamma_{pij} \equiv \frac{\tau_{Ipj}\tau_s}{\tau_s\tau_{ij0} + \tau_{ij0}\tau_{Ipj} + \tau_{Ipj}\tau_s}$.

Assumption 2 implies that the bias assessment of the agents is not too erroneous, and the absolute value of the error ($\left|\nu_{pji}\right|$) is bounded by a constant that is proportional to the weight (γ_{pij}) that the agent puts on the signal when forming the posterior mean (Equation (4)). When the agent has a greater familiarity with the media bias (e.g., higher τ_{Ipj}), the error of her bias assessment is subject to a tighter bound. For example, a liberal is expected to have a smaller maximum error in her assessment on the media bias of *The New York Times* compared with her assessment on the bias of Fox News.

For notational convenience, let us define the error of the prior as $\eta_{ij0} \equiv \theta_{ij0} - \theta^*_j$ or, equivalently, $\theta_{ij0} = \theta^*_j + \eta_{ij0}$. I put no structure on the distribution of η_{ij0}, except that it is independent of ν_{pi} and ε_{ij}.

Proposition 4 shows that those who are allowed to select their news source learn more on average (i.e., the distance between the prior mean and the posterior mean is higher when source selection is allowed).

Proposition 4. *(Learning) (i) Suppose Assumptions 1 and 2 are true with some* M_j. *There exists* $M_j > 0$ *such that allowing source selection, as opposed to equal-chance exposure to either F or A, facilitates learning (*$|\theta_{ij0} - \theta_{ij1}|$ *is larger) on average. (ii) If we further assume that* $E\left[\left(\gamma_{Fij} - \gamma_{Aij}\right)\left|\eta_{ij0}\right|\right] > E\left[\left(\gamma_{Fij} + \gamma_{Aij}\right)\left|\varepsilon_{ij}\right|\right]$, *then agents who are allowed to choose their news sources get closer to the truth (*$\left|\theta_{ij1} - \theta^*_j\right|$ *is smaller) on average than agents who are given signals from randomly selected sources.*

Proof. See Appendix A. □

Proposition 4 implies that (i) an agent learns more and (ii) gets closer to the truth (θ^*_j) when she is allowed to choose her news source. Part (i) of Proposition 4 relies only on Assumption 2, whereas Part (ii) requires an additional assumption—$E\left[\left(\gamma_{Fij} - \gamma_{Aij}\right)\left|\eta_{ij0}\right|\right] > E\left[\left(\gamma_{Fij} + \gamma_{Aij}\right)\left|\varepsilon_{ij}\right|\right]$.

Note that this assumption is more likely to hold among the agents who are initially far away from the truth ($\left|\eta_{ij0}\right| = \left|\theta_{ij0} - \theta_j^*\right|$ is large). An example of such a subpopulation in the context of this paper's experiment is those who begin with extreme policy views that are unlikely to be close to the optimal position. Proposition 4 predicts that agents who are allowed to select their news sources are more likely to have moderated policy views after exposure to a signal.

For the derivation of Proposition 4, we needed to assume that the agents have approximately correct assessments of the media's biases, where the approximation is more accurate for the familiar news sources (Assumption 2). On the other hand, if the agents have a systematically incorrect perception about the biases of the news sources (e.g., imagine a liberal firmly believing that a left-wing source is telling the truth, the whole truth, and nothing but the truth), then selective exposure to the familiar news source may result in convergence to the position advocated by the party that the agent supports. Since this can be interpreted as an exacerbation of polarization,[11] I empirically test this in Section 5.3.

2.2 Discussion

The purpose of the model described in the previous subsection is to provide important insights that the empirical results of this paper can corroborate, while maintaining a standard structure of Bayesian models.

In a model where media can choose to be silent about certain issues, the predictions can be different (Banerjee and Somanathan, 2001). This is an unlikely scenario in the context of this paper because each issue has been covered in almost all of the media appearing in the mobile app.

Note that selection of familiar sources can be driven by other reasons than the model above demonstrates. In the model, the familiar news source is selected because the agent hopes to learn the truth. This objective is better achieved when the agent acquires information from news sources whose bias is more familiar to the agent, thereby making their signals more informative. It is important to acknowledge that there are other potential motives for selective exposure that are not included in the suggested model. For example, people may get direct utility from the act of reading articles from familiar media. Empirically distinguishing the motives of selective exposure is beyond the scope of this paper. Although it is theoretically concise to have the

[11] Convergence of policy views to the favored party's position is one of the most salient forms of attitude polarization in the US in the last two decades. (Gentzkow, 2016). Baldassarri and Gelman (2008) argue that this trend is not because the general public changed their views; it is because the public switched their supporting parties to have better aligned policy views, and parties adjusted their policy positions to better serve their supporters.

same factor—familiarity to media bias—driving both the selection of familiar news sources and facilitated learning, it is not strictly necessary for the main results. Specifically, the agent must be sophisticated to the extent that she takes the media biases into account when updating her views, but she does not need to choose the familiar news sources solely due to the desire for facilitated learning.

Learning can be a multi-stage process in reality. Even if the agent begins with the wrong perception of her familiar media's bias she may learn its true bias through receiving sufficient amounts of signals in the long run. The process of learning true biases will be facilitated if she attempts to learn deeply about a particular issue, which makes it optimal to eventually take signals from other available sources (Liang et al., 2017).[12] In the long run, therefore, the assumption of a correct bias assessment (Assumption 2) may be even more reasonable.

Due to my focus on consumers of news articles, this paper does not explore the strategic motives of the media. In particular, if the true biases of the sources are eventually revealed to the readers, why would the news media bother to add media biases to the signals? There are many possible explanations. For example, the news producers might get pleasure from adding media biases to the articles. It is also possible that a large enough part of the population is yet to reach the long-run steady state in which media biases are precisely known. In this case, adding biases can be beneficial for news producers either because it makes the readers biased or because it lets the media sustain its reputation of providing high-quality information.[13] Finally, the pleasure derived from reading like-minded news may be large enough for readers to choose like-minded news sources.

The "group," used in this paper to define polarization, can take many different forms depending on the specific context. The empirical sections of this paper emphasize a partisan divide, because that appears to be the most salient determinant of media selection in the context of this paper. In other contexts, groups can be divided by, for example, religious beliefs, ethnic identity, or regional boundaries, as long as such group identity is most relevant for media selection and political polarization in that context.

[12]This is intuitive because the signals from one source are inevitably correlated. See Liang et al. (2017) for theoretical exposition. The authors also prove that, under some regularity conditions, a myopic agent and a forward-looking agent select approximately (or exactly, with additional assumptions) the same news sources. This may justify the use of myopic utility maximization in the model above.

[13]When the target population is biased, it may be optimal for the media to provide biased information—see Gentzkow and Shapiro (2006).This paper does not take a stand on whether the bias of the media is supply- or demand-driven. For an empirical analysis on this, see Gentzkow and Shapiro (2010). The appendix in Chiang and Knight (2011) discusses the theoretical foundation of both channels.

2.3 Predictions

The model in Section 2.1 gives the following predictions, listed in order of their appearance in the empirical sections below. The Results Section—Section 4—provides evidence for predictions [1]-[3]. Prediction [4] is presented in Section 5 where I discuss mechanisms.

[1] *Belief updating patterns are consistent with standard Bayesian updating. The posterior mean is affected both by the prior mean and the signal, and it is affected less by the signal when the confidence on the prior is higher. (Proposition 1)*

[2] *Selective exposure exists (i.e., familiar news sources are selected when given a choice). (Proposition 3)*

[3] *When allowed to choose news sources (as opposed to getting randomly selected articles), readers change their views to a greater degree as long as they can properly deduct the biases of the articles (Proposition 4). Among those who start from positions that are highly unlikely to be close to the optimal position, readers who are allowed to choose their news source get closer to the truth (i.e., their initial extreme positions are more likely to become moderate).*

[4] *Readers are more affected by the article's position when the article is from familiar sources. (Proposition 2).*

3 Setting, Experimental Design, and Data

3.1 Background

South Korea has a representative democracy with a presidential system. Since the end of the military dictatorship in 1988, two parties have led the executive branch in turn. A right-wing party (called the "Saenuri Party" at the time of this experiment), which represents economic liberalism and social conservatism, brought four out of seven presidents to power. The remaining three presidents were from a center-left party (the Democratic Party), which advocates economic equality and social liberalism. In addition, there were two smaller parties, the People's Party (center) and the Justice Party (left) during the period of this experiment.[14]

[14] President Park of the Saenuri Party led the executive branch during most of the experimental period. The Saenuri Party also had a parliamentary majority for the first three months of the experiment, but they lost it in the April, 2016, General Election. Shortly before the end of the experiment, there was an outbreak of one of the

Political polarization is one of the most serious sources of societal conflict in South Korea (Korea Institute of Public Administration, 2015; p.165). There is survey evidence from 2004 and 2008 (Lupu, 2015) demonstrating that the partisan polarization of South Korea has been similar to or more severe than that of the US. Like many other countries, news sources in South Korea are believed to be politically biased, and are accused of being the main drivers of political polarization. In a survey (Korea Institute of Public Administration, 2015), more than half (61%) of the respondents said that "the press is unfair" (p.142).

Traditional media, such as TV channels and newspapers, remain the primary news sources in South Korea. According to a recent survey (Korea Press Foundation, 2016), South Koreans on average spent 50% of their daily news consumption time (83 minutes) on TV, and 40% on newspapers (including online newspapers). The rapid rise of news consumption via mobile phone is noteworthy and made this experiment possible. In 2016, South Koreans spent 17 minutes per day reading newspaper articles on their phones, a 148% increase from 2011.

Three top national conservative newspapers (*The Chosun Ilbo, The JoongAng Ilbo, The Dong-a Ilbo*) with a combined market share of 60% largely dominate the printed news subscriptions in South Korea. As for online news consumption (both mobile and PC), news aggregation services, online messaging services, and social network services are the main news sources, whereas visits to individual newspaper websites are rare. News aggregation services by Naver and Daum, two of South Korea's top IT companies, are the most common avenues to consume news online (Korea Press Foundation, 2016). This presumably contributes to a relatively less concentrated market share among original creators of news content.

With its exceptionally vibrant mobile environment, South Korea provides unique opportunities for researchers to study political persuasion through low-cost online field experiments in a country that shares similar political and journalistic problems with other developed countries such as the US.

3.2 Setting: Mobile Application

To study the impact of media consumption on policy views, I developed an iOS mobile application and distributed it to the general public of South Korea from February through November, 2016. The app, "Spoon," was primarily a news curation service. I collected articles daily on

largest political scandals in the history of South Korea, which resulted in the impeachment of then-president Park. The experiment was mostly unaffected by this scandal because only 1% of the article-reading instances in this app ocurred after the political scandal was exposed on October 24, 2016. None of these instances is included in the comparison between treatment groups.

eight politically disputed issues from a wide spectrum of news sources spanning left to right. At each point in time, there were 10-20 articles on each issue and the articles were consistently replenished on a weekly basis. Figure 1 presents the list of issues covered in this application.[15] All eight issues had been of great interest to the general public throughout the experimental period. Therefore, every issue had been covered by almost all the news sources that were featured in the app during the experimental period.

The users had additional incentives to use the app: I provided several types of aesthetically pleasant summary statistics on, for example, how their self-reported policy views evolved over time and which articles and news sources had been most influential to them. These features contributed not only to attracting and retaining users, but also to incentivizing truthful revelations of the users' attitudes. There was no monetary compensation to use the app, except for 37 users who responded to the intensive follow-up e-mails.[16] Most of the users connected voluntarily to the app, presumably to read well-written articles and to see summaries of their self-reported policy views.

The app was distributed to the general public for free; anyone who had access to an iPhone could participate in the study. Facebook advertisements were the primary source of user inflow. A total of 2,627 people installed the app; 1,420 of them finished reading the first article and were included in the sample.

3.3 Experimental Design

I exploit two sources of randomness in this experiment. First, an article was randomly selected for each user in a subset of the experimental period. All 1,420 users in the sample are included in the analyses that exploit this random provision of articles. Second, each user was assigned to one of seven treatment groups. Only those who completed the five-day pre-exposure period (367 users) were included in this randomization.

As shown in Figure 2, upon installation and baseline survey, each user had a five-day pre-exposure period. During the pre-exposure period, a randomly selected article from a randomly

[15] There were five additional issues ("temporary issues") covered in the mobile app for a relatively short period of time, and I provided fewer articles on these issues (3-5) at each point in time. These issues are excluded from the analysis in this paper, because we did not survey the baseline positions on these issues.

[16] This intensive follow-up is not exploited in this paper due to the limited response to it. For those who are curious about the exact procedure: there were low- and high-intensity groups in this follow-up attempt. The low-intensity group could choose between (i) a 4,000 Won ($\approx$ \$4) reward with 100% probability and (ii) a 50,000 won ($\approx$ \$50) reward with 10% probability. The high-intensity group had the same choice except the reward amounts were doubled.

selected news source about a randomly selected issue was provided each day.[17] Pre-exposure to randomly chosen news sources during this period serves as an important exogenous variation for the test of the mechanism advocated in this paper. This period also serves as a grace period, alleviating the attrition problem by screening out those who were going to drop out early. See Schilbach (2019) for the use of a grace period in a different context.

I assigned the 367 users who finished the pre-exposure period into one of three treatment groups as in Figure 2. The everyday experience of the treatment groups was slightly different. First—a common step for every treatment group—the issue of the day for each user was randomly selected by the app. Then, the Source-Name Group was allowed to select their news source based on source names. The Source-Position Group could not see the names (thus it was harder for them to identify the media that were likely to advocate the supported party's view) but could see source positions and select based on this information.[18] The No-Choice Group was not allowed to choose news sources and were given randomly selected articles.[19] Regardless of the treatment status, everyone could determine the news source at the time of reading the article with a minimal effort because the articles always indicated the source's name at the end of the text. Most articles also indicated the source's name at the beginning and in the middle of the text.

These three bigger groups had subgroups—G1-G7 in Figure 2. I do not explore these subgroups in this paper due to lack of statistical power, although I do provide analysis by finest subgroups in Appendix Tables whenever necessary.[20]

[17]The users were not required to use the app every day; they chose whether to use it on a particular day. The users could read multiple articles per day, and starting from her second article, a user could select the next issue of interest. The next article was also randomly selected for the chosen issue. Since I allowed people to read multiple articles per day, the median term between rounds on an issue was 5.8 days, which is shorter than 8 to 9 days—the expected term between rounds if everyone read only one article per day.

[18]The average position of a news source is the average of the positions of the articles written by the news source on the issue at hand. The average position of each article is calculated based on users' reports after reading the article—see Panel (e) of Figure 3.

[19]There was a cross-randomization where some of the users had access to the distribution of the positions of other users. This treatment arm is not the focus of this paper.

[20]Another reason not to explore these subgroups is the fact that the source positions that we provided in the app were inaccurate due to logistical problems. They almost always indicated source positions close to the center of the scale, and not surprisingly this indicator was largely ignored by our subjects according to anecdotal evidence. Furthermore, the names of the news sources were always salient in the reading screen, making the experience of G4-G7 highly indistinguishable. One can also argue that selection based on uninformative source positions (the Source-Position Group) is equivalent to having no choice at all (the No-Choice Group). Given the fundamental difference between these two groups (having the choice vs. not), I do not merge the groups; however, merging them generates similar results. The main comparison will be between the Source-Name Group and the No-Choice Group throughout this paper.

3.4 Data Collection

All the data that I use in this study were collected in the app. After the (random or deliberate) selection of a news source, the users read an article from that source. Given the news source, an article was randomly selected among the pool of articles that were written by the source. The users then reported their policy views on the issue and the level of confidence in their positions (Panel (d), Figure 3), and evaluated the quality and the position of the article (Panel (e), Figure 3). The positions are continuous measures between 0 and 1. For users' positions, the app asks the user to "move the scroll bar to denote your view on <*IssueName*>." The app provided a rough cardinal benchmark to the positions on the horizontal attitude bar. As the user moves the scroll bar, the app explained what each position meant in terms of specific policies or attitudes on the issue. To be more specific, the horizontal attitude bar was divided into five segments where each segment had a corresponding explanation. See Figure 4 for an example of such explanations. The confidence (Panel (d), Figure 3) and the article quality (Panel (e), Figure 3) are discrete "star" measures, where half stars are allowed. In the empirical results below, the confidence measures are also transformed into $[0, 1]$ for easier interpretation.

I also measured the total time that each user spent reading articles. In addition to the time spent on reading the article, the measure includes the time spent on the selection screen (Panels (a) and (b)), the media information screen (Panel (c)), and the evaluation screens (Panels (d) and (e)). Unfortunately, I do not have separate measures for the time spent on each screen.

3.5 Summary Statistics and Experimental Validity

In Appendix Table B1 and Table B2, the row for "Outcome Mean" provides summary statistics. As expected for typical app users, the sample is younger (31 years) than the South Korean population (41 years) on average. It is also more liberal, less likely to watch TV news (56-59% in my sample vs. 83% for the population), and slightly more likely to use the Internet for news consumption (88-89% vs. 74%). All population figures are from the Korea Press Foundation (2016). One should be careful in generalizing the results of this paper to a wider population or to other nations.

The median time each person spent on the app each day was 2.2 minutes. This is 3% of the self-reported news consumption time of South Koreans (Korea Press Foundation, 2016). However, considering the facts that (i) only 30% of the total news consumption was on political news and (ii) self-reported news consumption time is reportedly exaggerated by a factor of 7-8 for the app's target population (Prior, 2009), perhaps 60-70% of the total political news consumption can be accounted for by the news consumption in the app.

Many of the subjects reported extreme policy views at the baseline—37% of the reported policy views are either 0 or 1, the most extreme views in either direction that could be reported in the app. These extreme views are qualitatively highly unlikely to be the optimal positions for any subject, considering the strong language used to describe these positions. For example, for the issue of foreign policy toward North Korea (see Figure 5), the most conservative policy-view category was "North Korea is our enemy; we should be aggressive in all aspects." A policy view of 0 corresponds to the most extreme view even within this category, an unlikely optimal position. I use the likelihood of extreme views as one of the most important outcome variables in the Results Section of this paper.

Appendix Table B1 provides the randomization check for the random provision of articles, and Appendix Table B2 provides the check for random assignment to treatment groups. It appears to be balanced for both. Note that article pool and round fixed effects are included in the regressions of Appendix Table B1. The article position is exogenous only after including these fixed effects because the news articles are removed and replenished over time to keep the app up-to-date. The position of the article is crowd-sourced—it is the average of the user-evaluated article positions.

There is a concern about attrition for the analyses that explore the random assignment to treatment groups. When I compare the experimental groups to identify the causal effect of selective exposure on polarization, I use user$\times$issue combinations with at least one article read after the group assignment. As shown in Column 9 of Appendix Table B2, there is 40% attrition for this comparison; though encouragingly, the baseline characteristics or assignment to treatments as a whole are uncorrelated with the attrition (see Appendix Table B3). Note also that within a user, the order in which issues appear is randomly determined, so within-user (across-issue) attrition is less of a concern. The more severe type of attrition is across-user attrition. Reassuringly, 92% of the users who were assigned to a treatment group read at least one article after the pre-exposure period.

4 Results

Subsection 4.1 below establishes that the user's policy-view-updating patterns are consistent with typical Bayesian models, giving confidence in using a simple Bayesian model to analyze the mechanism. Subsection 4.2 investigates the news-source-selection patterns. Subsection 4.3 provides the main results of this paper: Unregulated media selection, as opposed to random exposure, facilitates learning and alleviates polarization.

4.1 Policy view updating patterns: consistent with typical Bayesian models (Prediction [1])

I begin the empirical analysis by establishing the consistency of the data with the predictions of general Bayesian models. In this subsection, I use the subsample that was given randomly selected articles. This subsample includes everyone during the pre-exposure period and the users in the No-Choice Group after the pre-exposure period (see Figure 2 for randomization design). Because the provision of articles was random, there is no spurious correlation between the article position and the reader's view reported after reading an article.[21]

I test some basic predictions of typical Bayesian models: (i) the reader's policy view reported after reading an article is affected by the article's position and the view reported prior to the reading, and (ii) the article has lower influence if the reader reported high confidence in her view prior to reading the article.

I first estimate:

$$Position_{ijr}^{post} = \alpha_r + \alpha_A + \gamma_1 \cdot Position_{ijr}^{prior} + \gamma_2 \cdot s_{ijr} + \mathbf{X_i}\beta + \varepsilon_{ijr} \tag{6}$$

where $Position_{ijr}^{post}$ is the posterior position on j reported by i after article-reading round r, $Position_{ijr}^{prior}$ is the position reported prior to reading this article, s_{ijr} is the position of the article (signal), and $\mathbf{X_i}$ is the vector of user controls. Prior positions are recorded either from a post-reading survey in the previous round, or a random survey that often appears after reading an article about other issues. The position of the article is crowd-sourced—it is the average of the user-evaluated article positions. The pool from which an article is randomly drawn changed over time—I add article pool fixed effects (α_A) to deal with this. Since article pools vary across issues, α_A contains the issue fixed effect. α_r is the round fixed effect ($r = 1$ for the first article read in the app, $r = 2$ for the second, etc.).

Columns 1-3 of Table 1 provide the results. Each column is derived from a separate OLS regression with a different combination of fixed effects and user controls as specified. Standard errors are clustered by user. Since the results are similar across the article-reading rounds, I pool them together.

User positions are affected by the article's position. The posterior position consists of ap-

[21] It was very rare for a user to drop out after getting an article without finishing it. Using only the observations with completed reading, Appendix Table B1 shows balance between sample with different article positions.

proximately 8-9% of the signal, 55-69% of the prior, and other unexplained factors such as pure statistical noise. Results are similar with and without fixed effects and user controls.[22]

Taking the theory more seriously, one must account for the bias deduction of the agents (see Equation (4)). To control for this, I add news source fixed effect interacted with issue fixed effect (Column 2 of Appendix Table B6) and its interaction with the dummy variable indicating the most trusted party at the baseline (Column 3 of Appendix Table B6). The results are similar—user positions are positively affected by the position of the article and the magnitude of the effect is similar. In the Mechanisms Section (Section 5), I also explicitly include the news source's position in the regression to report how readers adjust for the media's biases.

Typical Bayesian models also predict less influence of the signal when the agent has high confidence in the prior position. In Column 3 of Table 1, I interact both the prior and the signal with self-reported confidence on the prior. I control for the confidence, and the interaction of article pool fixed effect and confidence is also included. The result is consistent with the prediction, where the highest confidence (= 1) means almost zero coefficient on the signal $(0.26 - 0.24 \cdot 1 \approx 0)$.

The results in this subsection give confidence in using a Bayesian model to understand the mechanism of the main empirical results of this paper.

4.2 Selective exposure to partisan news sources (Prediction [2])

In this subsection, I report the observed media selection patterns. The first step is to establish the existence of selective exposure—selection of media that are likely to represent the favored political group's position. I show the result supporting this claim.

I first report the source selection of the group that could select based on source names (Source-Name Group). It was easy for them to identify and choose the media that are expected to represent the supported party's position if they wanted to do so, because the source names were shown on the selection screen (see Panel (a) of Figure 3). I estimate:

$$C_{ij} = \alpha_r + \alpha_A + \gamma_1 \cdot y_{ijr}^{prior} + \gamma_2 \cdot P_{ij} + \mathbf{X_i}\beta + \varepsilon_{ijr} \tag{7}$$

where C_{ij} is a proxy for the chosen source's expected position on issue j, and P_{ij} is the position

[22] One may wonder whether the effect is highly transient and if the knowledge or attitude will quickly dissipate as time goes by. This is not the case. As can be seen in Appendix Table B4, the effect of the first article persists even after reading the second article (median 5 days later) and after reading the third article (median 14 days later). The result above is also robust to dropping the observations where the users hold extreme prior views ($\in \{0,1\}$) before reading an article (Appendix Table B5).

on issue j of the i's most trusted political party at the baseline, which is proxied by taking the average of the baseline positions on issue j of other users who share the most trusted party with user i and have the baseline confidence ($\in [0,1]$) of 0.8 or larger. The user's level of trust for each party was surveyed at the baseline. y_{ijr}^{prior} is the user's position prior to making the news source selection. The pool of articles changed over time—I add article pool fixed effect (α_A) to deal with this—note that α_A includes issue fixed effect. Also included are $\mathbf{X_i}$, the vector of baseline user characteristics, and Round FE (α_r). They are included for statistical power. Standard errors are clustered by user. $\gamma_2 = 0$ indicates the scenario where there is no selective exposure to the partisan media.

In Columns 1-2 of Table 2, the position of the article that was read by the user is used as a proxy for the source's expected position. The chosen source's expected position is strongly associated with the supported party's position, indicating that the news sources that are expected to represent the supported party's position were selected. The association is strong even after controlling for the prior position (Column 2): Partisan news sources are selected rather than the news sources that are likely to confirm the preexisting views of the readers. In Columns 2 and 3 of Appendix Table B7, I show that even when the prior position is relatively far from the party position, the users choose their partisan media.

The result is similar in Column 3 of Table 2, where an alternative measure of expected source position is used. "Expected position of the chosen source" is the average position on an issue of the articles written by the news sources that share the same favorite party as the news source at hand. The favorite party of the news source is the party that has the smallest average position distance across issues with the news source. In Columns 4 and 5, I take only the first-round selection after the pre-exposure period for each user$\times$issue combination to minimize the concerns of sample selection due to dropouts. The results are similar.

As in many other contexts, it is expected that the readers have greater familiarity with their like-minded partisan news sources. Therefore, the selection patterns found in this subsection are consistent with the model's prediction that the familiar news sources are selected (Prediction [2]). However, there can be other reasons why like-minded partisan sources are selected, and I do not argue against those here. No matter what the motivations, the main implications of the model—comparison between the treatment groups—are valid as long as the selected news sources are *also* familiar news sources to the readers.

The magnitude of selective exposure found here is comparable to the level found in Gentzkow and Shapiro (2011). I calculate the isolation index defined in their paper, deliberately coarsening the information that I have on users to match what the authors report. As described in Gentzkow and Shapiro (2011), the isolation index equals "the average conservative exposure of conservatives minus the average conservative exposure of liberals."[23] I find an isolation index of 0.076 to 0.143, depending on how I match the parties with the binary variable indicating ideology. If I categorize the two most conservative parties as "conservative" and the two remaining as "liberal," the isolated index is 0.076. If I instead categorize the Saenuri Party as the sole conservative party, it is 0.143. I do not directly ask about the user's ideology in the survey. The standard error, calculated from a non-parametric bootstrapping over users, is 0.06. Although the measure is noisier than Gentzkow and Shapiro (2011) due to the smaller sample size, the magnitude is comparable to the isolation index they find for the Internet (0.075) or national newspapers (0.104). It is encouraging to observe a similar level of selective exposure in the context of this experiment—the effect of selective exposure found later in this paper may be applicable to other contexts, such as online news consumption in the US.

Table 3 compares the selection patterns of the three treatment groups. Columns 1 and 2 in Table 3 are benchmarks that correspond to the group that could select based on source names (Source-Name Group), a repetition of Table 2. Columns 3 and 4 in Table 3 show that selection of partisan media is effectively shut for the group that could select only based on source positions (Source-Position Group), presumably because it is hard to identify the media's party affiliations without seeing their names and the provided source positions were uninformative, as described in Section 3. Columns 5 and 6 present a placebo test. The observations that were given randomly selected articles should have no statistically significant correlation between the positions of the chosen sources and the predetermined variables, and this is verified in the Table.[24]

Provided that the Source-Name Group was selectively exposed to their like-minded partisan media in this experiment, a prediction that is consistent with the popular rhetoric would be stagnated learning and increased political polarization for the Source-Name Group, compared with No-Choice Group. In the next subsection, I test these hypotheses by comparing the treatment groups.

[23]See Appendix A of Gentzkow and Shapiro (2011) for the precise method used by the authors. I closely follow their method, including leave-one-out estimation to adjust for the small sample. The only notable adjustment that I make to their method is that I put higher weight on the conservative users, so that the measure is comparable to the nationally representative sample. Specifically, I normalize so that the conservative sample as a whole receives the same weight as the liberal sample as a whole.

[24]Appendix Table B8 compares the media selection patterns of the finest subgroups. The selection patterns of subgroups within the three bigger treatment groups are similar.

4.3 Allowing source selection facilitates learning and reduces extremism (Prediction [3])

In the last subsection, I show that the readers select news sources that are expected to represent their favored party's position (and not prior-confirming ones). I also observe that the three treatment groups showed quite different selection patterns. The group that could select based on source names (Source-Name Group) showed selective exposure, whereas the group that could select based only on source positions (Source-Position Group) revealed effective shutdown of selective exposure. The group that was given randomly selected articles (No-Choice Group) did not have any choice and thus mechanically showed no selective exposure.

In this subsection, I present the main findings of this paper—the impact of media selection on evolution of policy views. I compare the positions of three treatment groups that show different patterns of selective exposure as described above. Although I show the results of all three treatment groups for completeness, I focus on the comparison between the group that could select based on source names (Source-Name Group) and the group that was given randomly selected articles (No-Choice Group). The comparison between the two groups identifies the effect of unregulated media selection—a setting that resembles the real world—(Source-Name Group) as opposed to random and balanced media exposure (No-Choice Group). Comparison between the group that could select based only on source positions (Source-Position Group) and other groups is not highlighted because the results are generally noisy, as the group has the fewest subjects.

The model in Section 2 demonstrates that selective exposure enables the readers to select sources that they perceive to have high information value. The readers who were allowed to select familiar news sources will experience facilitated learning—the positions of the readers will change to a greater degree after reading articles (Prediction [3]).

In Table 4, I estimate:

$$\left|Position_{post} - Position_{pre}\right| = \gamma_0 + \gamma_1 \text{SourceName}_i + \gamma_2 \text{SourcePosition}_i + \mathbf{X_i}\beta + \alpha_j + \varepsilon_{ij} \quad (8)$$

where i is individual, and j is issue. $\mathbf{X_i}$ is the vector of user controls, and α_j is issue fixed effect. The level of observations is user$\times$issue, and the standard errors are clustered by user. The omitted group is the group that was given randomly selected articles (No-Choice Group).

I observe a 17% larger ($\frac{0.025}{0.15} \simeq 0.17$; $p < 0.05$) absolute distance between prior and posterior positions for the group that could select based on source names (Source-Name Group), compared with the group that was given randomly selected articles (No-Choice Group; the omitted category in the regression). In other words, unregulated source selection actually facilitates learning. The result is robust to the addition of user control variables (Column 2) and issue fixed effects

(Column 3). This result is consistent with Prediction [3]. The result is not driven by any single issue—the effect is robust to dropping any one of the issues (Appendix Table B9).

If the larger distance between prior and posterior positions is driven by facilitated learning about the underlying true states of the world—as demonstrated by the model in Section 2.1—and if we are willing to believe that the ideal positions of the readers are probably not at the extreme points, we should observe larger movements in positions for those who have extreme baseline views. Therefore, I next focus on the users who reported having extreme views at the baseline. A user is categorized as having an extreme policy view on an issue if she has a position of 0 or 1, which are the endpoints of the continuous horizontal scroll bar that the readers used to record their policy views in the app (See Panel (d) of Figure 3).

In Table 5, I investigate whether the proportion of people holding extreme views is affected by selective exposure. I run two different types of regressions: first difference (FD) and lagged dependent variable (LDV). Randomized controlled trials with good balance should, in principle, give similar results even though they make different kinds of assumptions on their data generation processes. To make sure that the baseline imbalance (not statistically significant) is not affecting the results, I ran both regressions, as recommended by Angrist and Pischke (2008) and Wooldridge (2010). The regression equations are as follows.

$$y_{ij,post} - y_{ij,pre} = \gamma_0^{FD} + \gamma_1^{FD}\text{SourceName}_i + \gamma_2^{FD}\text{SourcePosition}_i + \mathbf{X_i}\beta + \alpha_j + \varepsilon_{ij} \tag{9}$$

$$y_{ij,post} = \gamma_0^{LDV} + \gamma_1^{LDV}\text{SourceName}_i + \gamma_2^{LDV}\text{SourcePosition}_i + \gamma_3^{LDV} y_{ij,pre} + \mathbf{X_i}\beta + \alpha_j + \varepsilon_{ij} \tag{10}$$

where $\mathbf{X_i}$ is the vector of user controls, and α_j is issue fixed effect. "*pre*" indicates baseline, and "*post*" is after reading an article in the experimental period. The outcome variable (y_{ij}) in Table 5 is an indicator variable that equals 1 if the policy view of i on issue j is extreme (i.e., $Belief \in \{0,1\}$). The hypothesis that we want to test is whether free choice of media (SourceName Group) facilitates a reduction of extremism compared to providing randomly selected articles (No-Choice Group). In terms of regression coefficients, we want to test whether $\gamma_1^{FD} = 0$ and $\gamma_1^{LDV} = 0$.

Regardless of the treatment, extremism is alleviated on average after reading an article—there was a 9 percentage point decline in the proportion of extreme views (0.37 at the baseline, and 0.28 after the first round of reading in the experimental period). Columns 1-3 show a robust negative effect of unregulated media selection on extremism—the group that could select based on source names (Source-Name Group) has 6-8 percentage points fewer people with extreme attitudes compared with the group that was given randomly selected articles (No-Choice Group; the omitted category in the regressions). The result for the group that could select only based on source positions (Source-Position Group) is noisy and inconsistent.[25]

The results are not driven by a particular issue—they are robust to dropping any one issue (Appendix Table B10). According to Proposition 4, convergence to the optimal positions is most likely to be expected among people who have unreasonable positions to begin with. Consistent with this condition, when I instead use a continuous measure of extremism—$|Position - 0.5|$—I find weaker and often insignificant effects, as shown in Appendix Table B11. Note also that 0.5 is not necessarily an optimal position for everyone, so this is not a test for the prediction that readers are on average getting closer to the optimal positions.

The results can also be seen graphically in the histograms in Figure 6 in which the prior and posterior position distributions are contrasted for each group. The Kolmogorov-Smirnov equality-of-distributions test accepts the null hypothesis that the two groups have the same distribution *prior to* reading an article ($p = .21$). It rejects the null that the two groups have the same distribution *after* reading an article ($p = .04$).

The results in this subsection indicate that selective exposure to partisan media may play a role in making attitudes less polarized. The readers learn more when they are allowed to choose their news sources, and their policy views become more moderated.

5 Mechanisms and Interpretations

5.1 Mechanism: familiarity

In the Results Section above I find that unregulated media selection contributes to moderated policy views, an alleviation of polarization. The model in Section 2 explains a potential mechanism of this finding. When media selection is allowed, the selected media are likely to be the readers' familiar news sources with biases that are better understood by the readers. This enables better adjustment for the biases of the news articles, which facilitates learning.

In this section, I provide evidence that supports this mechanism. First, I show that readers' posterior policy views are affected more by the article if the article is written by a news source that is ideologically similar to the user, which I use as a proxy for familiarity. Acknowledging the

[25] Appendix Table B12 shows the comparison for the finest subgroups, as described in Figure 2. Although the results are imprecise due to the smaller size of each subgroup, the sign of the effects are generally consistent with the larger groups' direction.

limitation of this proxy, I then provide experimental evidence, using the design where subjects were randomly pre-exposed to some news sources, temporarily boosting the subjects' familiarity with the sources. I show that when a randomly provided article in the experimental period is from a source that the subject was pre-exposed to during the pre-exposure period, the subject (i) is influenced more by the article, (ii) is better at adjusting for the biases of the article, and (iii) shows signs of facilitated learning and reduced extremism.

Ideologically Closer News Sources: Articles' Influence is Greater. The model in Section 2.1 predicts that a more familiar news source has greater influence on the readers' beliefs. According to the model, readers perceive that the articles from more familiar news sources have higher information value. In Bayesian language, readers put more weight on articles produced by such media (Proposition 2).

To test this, I again use the sample that was given randomly selected articles as in Subsection 4.1. I use an empirical specification similar to Equation (6), except now the prior and the article positions are interacted with a proxy measure of familiarity—$Dist_{pi}$.

$$\begin{gathered} Position_{ijr}^{post} = \alpha_r + \alpha_A + \gamma_1 \cdot Position_{ijr}^{prior} + \gamma_2 \cdot Position_{ijr}^{prior} \cdot Dist_{pi} + \\ \gamma_3 \cdot s_{ijr} + \gamma_4 \cdot s_{ijr} \cdot Dist_{pi} + \gamma_5 \cdot Dist_{pi} + \mathbf{X_i}\beta + \varepsilon_{ijr} \end{gathered} \tag{11}$$

where the ideological distance $Dist_{pi}$ is the average attitude distance—across the app's eight issues—between the news source and the user at the baseline. This is a proxy measure of how unfamiliar the news source is to i. As before, $Position_{ijr}^{post}$ is the posterior position on j reported by i after article-reading round r, $Position_{ijr}^{prior}$ is the position reported prior to reading the article, and s_{ijr} is the position of the article (signal).

Proposition 2 predicts that i is affected less by the signal when the news source is ideologically farther from the reader's position on average (i.e., γ_4 is negative). The results are presented in Column 2 of Table 6 and are consistent with the model's prediction. When the source's and the user's ideologies are perfectly aligned, the article's coefficient is comparable to 40% of the coefficient put on the prior position. However, when the ideological distance is in its 95th percentile (≈ 0.54), the article has virtually no influence ($0.2 - 0.54 \cdot 0.35 \approx 0$).

Although this is suggestive evidence supporting Prediction [4] in Section 2.3, the connection is not airtight. Specifically, I cannot rule out the possibility that $Dist_{pi}$ is correlated with a variable other than familiarity with the news source, and that variable is what makes the article more influential. To alleviate this concern, I leverage an experimental design that provides exogenous boost in familiarity with the news source.

Pre-Exposure May Lead to Better Bias Adjustment. As described in Figure 2, there was a five-day pre-exposure period for all the subjects, where the subjects were exposed to randomly selected news sources. It is plausible that this exogenous exposure temporarily raised the subject's familiarity with the news sources.

Leveraging this experimentally-induced variation in familiarity with the news sources, I investigate whether increased familiarity affects readers' position-updating patterns. In Table 7, I report the results of regressions similar to Equation (11), except here I add a right-hand side variable—news source position. This is the typical position of the news source on the particular issue, measured by averaging all the article position evaluations on the news source for the issue. I then interact all three right-hand side variables (prior reader position, article position, and news source position) with the number of pre-exposures of the user on the particular news source during the pre-exposure period. Note that the variable, the number of pre-exposures, is well-defined only for the observations after the pre-exposure period. Accordingly, I include observations of the No-Choice Group during the experimental period (after the pre-exposure period). In Column 2, I exclude observations with news source $\times$ issue with less than 10 total article evaluations to mitigate the classic measurement error for news source position.

Note that a negative coefficient for the news source position indicates that the reader is correctly adjusting for the bias of the article—controlling for the article position, the news source position reflects the bias that the news source has on the particular issue. Watching Fox News endorsing Trump would have no impact on the audience if she takes into account (or "subtracts") the pro-Trump bias of Fox News.[26] In both Columns 1 and 2 of Table 7, there is suggestive evidence that bias adjustments are more likely to be done correctly by those who were pre-exposed to the news source—see the negative coefficients on the interaction of the news source positions and the number of pre-exposures. Those who were pre-exposed to the article are also influenced more by the article—see the positive coefficient on the interaction between the article position and the number of pre-exposure. These coefficients are marginally insignificant in Column 1, whereas in Column 2 they are statistically significant in a conventional level, presumably due to a decreased measurement error of the news source position. These results are consistent with the

[26] In the formal model of Section 2.1, the posterior mean is

$$\theta_{ij1} = (1-\gamma_{pij})\,\theta_{ij0} + \gamma_{pij}\,(s_{pij} - I_{pji}) = (1-\gamma_{pij})\,\theta_{ij0} + \gamma_{pij}s_{pij} - \gamma_{pij}I_{pji}$$

Note that the position of news source p on issue j, $E[s_{pij}] = \theta_j + I_{pji}$, or $I_{pji} = E[s_{pij}] - \theta_j$. Plugging this in, we get

$$\theta_{ij1} = (1-\gamma_{pij})\,\theta_{ij0} + \gamma_{pij}s_{pij} - \gamma_{pij}E[s_{pij}] + \gamma_{pij}\theta_j$$

which suggests negative coefficient for the news source position, as long as the "truth" is properly controlled for.

advocated mechanism in Section 2—increased familiarity with a news source means (i) higher influence of the articles from the news source and (ii) enhanced capability to de-bias the articles.

Pre-Exposure May Lead to Facilitated Learning and Reduced Extremism. Does temporarily increased familiarity with a news source ultimately affect the main outcome variables of this paper: learning and moderation?

Table 8 provides suggestive evidence that the answer is potentially yes. In this Table, I further divide the No-Choice Group into two groups—the *No-Choice Group with pre-exposure* consists of those who were randomly assigned to an article from a pre-exposed news source in the experimental period. The *No-Choice Group without pre-exposure* are those who were randomly assigned to an article from a source to which they had not been exposed prior to the experimental period (i.e., pre-exposure period).

The *No-Choice Group with pre-exposure* shows greater movement in positions (Column 1) and has less extreme policy views (Column 2) after reading an article than the *No-Choice Group without pre-exposure*. In other words, a temporary boost in familiarity resulted in facilitated learning and reduced extremism, consistent with the advocated mechanism. The magnitudes are 57-82% of the main effects found in Tables 4 and 5.

Although not statistically significant, it is also encouraging that the degree of movement (Column 1) and moderation of positions (Column 2) are both smaller for the *No-Choice Group with pre-exposure* than the Source-Name Group—even with pre-exposure, the level of familiarity will be greater for those who intentionally select a news source than for those who happen to be randomly assigned a news source to which they were pre-exposed.

As discussed in Section 2, increased familiarity with a news source can be one of the channels through which trust in the source is built. Therefore, the result in this subsection is not necessarily an argument against a mechanism through trust. However, it is important to understand that blind trust without careful consideration of media bias can lead to less learning and more partisanship and extremism. Perhaps partly due to these countervailing forces, I do not find evidence of more learning or reduced extremism among the users who were randomly exposed to more trusted news sources compared with those who were randomly exposed to less trusted news sources (Appendix Table B13).[27]

[27] I use the users' self-reported trust in news sources, which was occasionally measured in the app. Specifically, at the beginning of each day each user had a 50% chance of receiving this survey. Conditional on receiving this survey, three randomly selected news sources were chosen to be surveyed.

5.2 Other mechanisms

In this subsection, I briefly discuss other potential mechanisms that can lead to facilitated learning and reduced extremism for the Source-Name Group compared with the No-Choice Group. I do not find evidence supporting these mechanisms. The discussion focuses on the comparison between these two main groups, as the comparison with the Source-Position Group is not the primary interest of this paper.

Source-Name Group was exposed to higher-quality or less extreme articles? It may be possible that the Source-Name Group, compared with the No-Choice Group, was exposed to objectively higher-quality or less extreme articles, and that is the reason why they learn more or be persuaded to be moderate. To tackle this, I re-weight the observations so that each article within a group has the same proportional weight as the article's weight in the Source-Name Group. This exercise is equivalent to creating a counterfactual where reading frequency per article is kept constant while the articles are randomly shuffled. As shown in Table 9, the main results are robust to this re-weighting.[28]

Disengaged No-Choice Group? The No-Choice Group are more likely to be exposed to news sources that are not aligned with the policy views of the readers' supported party, as shown in Section 4.2. This could potentially lead to disengagement by readers who would stop reading the article relatively quickly, limiting their learning opportunity. I do not find evidence supporting this claim. As shown in Appendix Table B16, the No-Choice Group spends on average only about 15-20 seconds less time (not statistically significant) in selecting, reading, and evaluating an article than the the Source-Name Group. Furthermore, taking into account that the No-Choice Group did not have the source selection step before reading an article, the difference in time spent on reading the article is probably even smaller.[29]

Movement toward right-wing positions? Given that the sample is left-leaning, one may hypothesize that the reduced extremism is driven by the Source-Name Group's general directional movement toward right-wing positions. However, there is no evidence of differential directional movement toward right-wing positions, as shown in Appendix Table B17.

Selection itself? More learning by the Source-Name Group could stem from the fact that they

[28] I also conduct direct test of article characteristics comparisons in Appendix Tables B14 and B15. User-evaluated article qualities (see Figure 3, Panel (e)) and article extremism are balanced between the Source-Name Group and the No-Choice Group. I use residualized article quality in Appendix Table B14, controlling for the distance between the supporting party's position and the article position as well as the distance between the user's prior position and the article position.

[29] Unfortunately, I do not have separate time measures for article selection and article reading.

actively selected the source themselves, potentially making them have a more open-minded attitude toward the articles. It has been suggested to test this hypothesis by comparing learning by two users—one in the Source-Name Group and the other in the No-Choice Group—who read the same articles. Unfortunately, this comparison suffers from selection bias—those who selected a particular news source may be different from those who were exogenously selected to receive an article from the same news source (De Benedictis-Kessner et al., 2019). The closest empirically viable evidence that I have on this is the comparison between the No-Choice Group, who did not select the news source, and the Source-Position Group, who selected the news source based only on highly uninformative signals, as described in Section 3. Learning between these two groups is very similar, as can be seen in Table 4.

5.3 Other interpretations

Stronger partisan identification? It is possible that allowing media selection simultaneously leads to (i) less extreme policy views (the main finding of this paper) *and* (ii) stronger partisan identification by making individual attitudes become closer to the party's prevalent policy views.[30] The latter can be interpreted as an exacerbated polarization, and thus it warrants empirical examination. Inconsistent with this hypothesis, users' views are not differentially approaching the supported party's positions. In Table B18, I test this using the regressions in Equation (9) and Equation (10), except now the outcome variable is the distance between the user's policy view and the supported party's position. There is no statistically distinguishable difference between the group that was given randomly selected articles (No-Choice Group) and the group that could select based on source names (Source-Name Group). In Appendix Table B19, I try other definitions of parties' positions, including the average baseline position of (1) the users who share the same most-trusted party and have trust in that party of over 0.5 ($\approx$ median of the trust in the most-trusted party) and (2) all the users who share the same most trusted party at the baseline. There is no evidence of differential convergence regardless of the measure. I conclude that I find no evidence of increased polarization induced by selective exposure.[31]

Alternative counterfactual? In the previous sections, I compare the Source-Name Group and the No-Choice Group, taking the counterfactual world where people are chance-encountering

[30] See Duclos et al. (2004) for the typology of polarization.

[31] The distributions of policy views by treatment group $\times$ supporting party are provided in Appendix Figure B1. I do not find evidence of stronger party identification for the Source-Name Group, although it is difficult to interpret the figure due to limited statistical power.

political information from a given set of available articles. Another counterfactual that I consider above in Section 5.2 is where the article viewership is kept constant while the articles are randomly assigned to readers (Table 9). The main results are robust to this re-weighting exercise.

In the previous exercises, I am also keeping *the current level of media plurality*. Another counterfactual to consider is the world where the Internet and cable TV are not invented, and therefore the media plurality brought by the Internet and cable TV is not realized. Although this general equilibrium analysis is impossible, a partial analysis can be made: Even compared with those who were randomly exposed to historically famous news sources, do users who were allowed to select news sources have more learning and less extremism? For this analysis, I keep users of the No-Choice Group who were exposed to any one of the five historically famous newspapers,[32] or one of the three air-wave TV channels.[33] As shown in the Appendix Table in B20, the result on moderation is similar to before even with this restricted No-Choice Group (Columns 3 and 4), although the result on learning is smaller and now statistically insignificant in conventional levels.[34]

6 Conclusion

This paper theoretically and empirically demonstrates that selective exposure can contribute to an alleviation of partisan polarization. By selecting partisan media, news consumers are exposed to familiar news sources. This facilitates learning and reduces polarization. I show suggestive empirical evidence that experimentally boosted familiarity with news sources better enables readers to adjust for media biases, and this in turn leads to facilitated learning and moderated policy views.

Over the last 20 years, we have experienced increases in the scope of media from which we can select, mainly due to the introduction of the Internet and the success of many social network websites. At the same time, polarization has increased. The insight in this paper is not necessarily in contrast to these recent trends and might be a key to understanding them. The rapid increase in the sheer number of potential news sources may have eroded media loyalty and reduced media familiarity. The fame of news sources comes and goes quickly nowadays, to the extent that even

[32] These five are widely recognized as the most prominent newspapers in Korea (to the extent that there is a nationally recognized abbreviation indicating these five): *The Chosun Ilbo, The JoongAng Ilbo, The Dong-a Ilbo, The Hankyoreh*, and *The Kyunghyang Shinmun*.

[33] KBS, MBC, SBS.

[34] I cannot reject the null that $|position(post) - position(pre)|$ of the No-Choice Group with exposure to one of the traditionally famous sources and of the No-Choice Group with exposure to one of less famous sources is the same.

well-known news aggregators often fail to keep up with popular demands. According to the mechanism introduced in this paper, reduced media familiarity can result in less learning and exacerbated polarization.

The result in this paper provides timely policy implications. Large social network websites have immense influence on how news articles are curated to their users because their algorithms determine how news links are sorted and filtered on the users' news feeds. These companies are actively devising ways to regulate how the news articles are shared, in reaction to the recent emergence of fake news (e.g., see CyberScoop, 2017). At the same time, there are society-wide discussions on how to regulate these intermediaries (e.g., see Fortune, 2017). According to the mechanism described above, it may be beneficial to provide news articles from users' familiar news sources instead of trying to provide balanced articles from relatively unfamiliar media (see Bail et al., 2018 for a similar recommendation). When the names and characteristics of news sources are made more salient in news feeds, users are better able to adjust for media biases in the articles they choose to read.[35]

How do we reconcile the mechanism introduced in this paper with the traditional wisdom that selective exposure drives partisan polarization (Sunstein, 2001)? Although selecting news sources among the established media is a very important decision, news consumers can make other choices. Readers may choose between the traditional, established media and other forms of new media, including pundits on Twitter or Facebook, fake news websites, and online communities. The selection can be across policy issues—readers can choose to stop paying attention to certain issues (Golman et al., 2017) or to political issues altogether (Arceneaux and Johnson, 2013; Kim and Kim, 2017). These "extensive-margin" selections may be the main drivers of political polarization. It is beyond the scope of this paper to compare the consequences of these different types of selections, and is an area I hope to address in future research.

Until further studies are conducted in different contexts, it is an open question whether the results in this study will generalize. The model presented in Section 2 clarifies the conditions under which the channel can operate. Note that South Korea does not have a particularly harmo-

[35]It is important to note that partisan polarization may not be necessarily costly to society. In a model where the party and the reader have a set of shared underlying values that determine the "correct position" for each political issue, taking the arguments of partisan media at face value could be welfare enhancing, especially when information acquisition and verification of the argument is costly to the general public (as in Canes-Wrone et al. 2001, Maskin and Tirole 2004). On the other hand, if these traits of partisan readers are strategically exploited by elites, the readers could be manipulated. Coordinated responses by non-partisan voters are required to counteract such under-informed partisans (Feddersen and Pesendorfer, 1996), which is a difficult goal to achieve (Battaglini, 2017). It is plausible that the misalignment in policy positions on *most of the issues* may cause deepened conflict and misunderstanding across the aisle, potentially affecting even personal matters such as marriages or hiring decisions (Iyengar and Westwood, 2015).

nizing media environment. In fact, survey evidence shows that the level of partisan polarization is similar to or more severe than that of the US (Lupu, 2015). Also, as shown in Section 4.2, the level of selective exposure found in this study is similar to the level in the US, giving hope for wider applicability of the results. The theoretical insight can certainly be generalized to other contexts where a partisan divide is not the defining characteristic of polarization. For example, the idea can be applied to situations where there is a great divide in policy views between ethnic or religious groups.

References

Acemoglu, Daron, Victor Chernozhukov, and Muhamet Yildiz, "Fragility of asymptotic agreement under Bayesian learning," *Theoretical Economics*, 2016, *11* (1), 187–225.

Adena, Maja, Ruben Enikolopov, Maria Petrova, Veronica Santarosa, and Ekaterina Zhuravskaya, "Radio and the Rise of the Nazis in Prewar Germany," *The Quarterly Journal of Economics*, 2015, *130* (4), 1885–1939.

Angrist, Joshua D and Jörn-Steffen Pischke, *Mostly harmless econometrics: An empiricist's companion*, Princeton university press, 2008.

Arceneaux, Kevin and Martin Johnson, *Changing minds or changing channels?: Partisan news in an age of choice*, University of Chicago Press, 2013.

Bail, Christopher A, Lisa P Argyle, Taylor W Brown, John P Bumpus, Haohan Chen, MB Fallin Hunzaker, Jaemin Lee, Marcus Mann, Friedolin Merhout, and Alexander Volfovsky, "Exposure to opposing views on social media can increase political polarization," *Proceedings of the National Academy of Sciences*, 2018, *115* (37), 9216–9221.

Baldassarri, Delia and Andrew Gelman, "Partisans without Constraint: Political Polarization and Trends in American Public Opinion," *American Journal of Sociology*, 2008, *114* (2), 408–446.

Banerjee, Abhijit and Rohini Somanathan, "A simple model of voice," *Quarterly Journal of Economics*, 2001, pp. 189–227.

Battaglini, Marco, "Public protests and policy making," *The Quarterly Journal of Economics*, 2017, *132* (1), 485–549.

Benedictis-Kessner, Justin De, Matthew A Baum, Adam J Berinsky, and Teppei Yamamoto, "Persuading the enemy: Estimating the persuasive effects of partisan media with the preference-incorporating choice and assignment design," *American Political Science Review*, 2019, pp. 1–15.

Boxell, Levi, Matthew Gentzkow, and Jesse M Shapiro, "Greater Internet use is not associated with faster growth in political polarization among US demographic groups," *Proceedings of the National Academy of Sciences*, 2017, p. 201706588.

Calvert, Randall L, "The value of biased information: A rational choice model of political advice," *The Journal of Politics*, 1985, *47* (02), 530–555.

Campante, Filipe R and Daniel A Hojman, "Media and polarization: Evidence from the introduction of broadcast TV in the United States," *Journal of Public Economics*, 2013, *100*, 79–92.

Canes-Wrone, Brandice, Michael C Herron, and Kenneth W Shotts, "Leadership and pandering: A theory of executive policymaking," *American Journal of Political Science*, 2001, pp. 532–550.

Chiang, Chun-Fang and Brian Knight, "Media bias and influence: Evidence from newspaper endorsements," *The Review of Economic Studies*, 2011, p. rdq037.

Chopra, Felix, Ingar Haaland, and Christopher Roth, "Do People Value More Informative News?," 2019.

CyberScoop, "Facebook admits 'false news' is a problem, pledges to curb 'information operations'," April 2017. [Written by Chris Bing Online; posted 27-April-2017].

DellaVigna, Stefano and Ethan Kaplan, "The Fox News Effect: Media Bias and Voting," *The Quarterly Journal of Economics*, 2007, pp. 1187–1234.

Duclos, Jean-Yves, Joan Esteban, and Debraj Ray, "Polarization: concepts, measurement, estimation," *Econometrica*, 2004, *72* (6), 1737–1772.

Durante, Ruben and Brian Knight, "Partisan control, media bias, and viewer responses: Evidence from Berlusconi's Italy," *Journal of the European Economic Association*, 2012, *10* (3), 451–481.

— **, Paolo Pinotti, and Andrea Tesei**, "The Political Legacy of Entertainment TV," DP 10738, CEPR Discussion Papers 2015.

Feddersen, Timothy J and Wolfgang Pesendorfer, "The swing voter's curse," *The American economic review*, 1996, pp. 408–424.

Flaxman, Seth, Sharad Goel, and Justin M Rao, "Filter bubbles, echo chambers, and online news consumption," *Public opinion quarterly*, 2016, *80* (S1), 298–320.

Fortune, "Facebook and Google Need to Be Regulated, Says British News Industry," March 2017. [Written by Mathew Ingram Online; posted 09-March-2017].

Gentzkow, Matthew, "Polarization in 2016," *Mimeo*, 2016.

— **and Jesse M Shapiro**, "Media Bias and Reputation," *Journal of Political Economy*, 2006, *114* (2), 280–316.

— **and** — , "What drives media slant? Evidence from US daily newspapers," *Econometrica*, 2010, *78* (1), 35–71.

— **and** — , "Ideological segregation online and offline," *The Quarterly Journal of Economics*, 2011, *126*, 1799–1839.

— **,** — **, and Matt Taddy**, "Measuring polarization in high-dimensional data: Method and application to congressional speech," Working Paper 22423, National Bureau of Economic Research 2016.

Golman, Russell, David Hagmann, and George Loewenstein, "Information Avoidance," *Journal of Economic Literature*, 2017, *55* (1), 96–135.

Halberstam, Yosh and Brian Knight, "Homophily, group size, and the diffusion of political information in social networks: Evidence from Twitter," *Journal of Public Economics*, 2016, *143*, 73–88.

Hetherington, Marc J, "Resurgent mass partisanship: The role of elite polarization," *American Political Science Review*, 2001, *95* (3), 619–631.

Iyengar, Shanto and Kyu S Hahn, "Red media, blue media: Evidence of ideological selectivity in media use," *Journal of Communication*, 2009, *59* (1), 19–39.

— **and Sean J Westwood**, "Fear and loathing across party lines: New evidence on group polarization," *American Journal of Political Science*, 2015, *59* (3), 690–707.

— **, Gaurav Sood, and Yphtach Lelkes**, "Affect, Not Ideology: A Social Identity Perspective on Polarization," *Public opinion quarterly*, 2012, *76* (3), 405–431.

Jensen, Jacob, Suresh Naidu, Ethan Kaplan, Laurence Wilse-Samson, David Gergen, Michael Zuckerman, and Arthur Spirling, "Political Polarization and the Dynamics of Political Language: Evidence from 130 Years of Partisan Speech [with Comments and Discussion]," *Brookings Papers on Economic Activity*, 2012, pp. 1–81.

Kim, Eunji and Jin Woo Kim, "Switching On and Off: Rethinking Partisan Selective Exposure," *Working Paper*, 2017.

Korea Institute of Public Administration, "Korea Social Integration Survey," 2015.

Korea Press Foundation, "Media Audience Survey," 2016.

Kramer, Roderick M, "Trust and distrust in organizations: Emerging perspectives, enduring questions," *Annual review of psychology*, 1999, *50* (1), 569–598.

Leeper, Thomas J and Rune Slothuus, "Political parties, motivated reasoning, and public opinion formation," *Political Psychology*, 2014, *35* (S1), 129–156.

Liang, Annie, Xiaosheng Mu, and Vasilis Syrgkanis, "Optimal Learning from Multiple Information Sources," 2017.

Lupu, Noam, "Party polarization and mass partisanship: A comparative perspective," *Political Behavior*, 2015, *37* (2), 331–356.

Martin, Gregory J and Ali Yurukoglu, "Bias in cable news: Persuasion and polarization," *American Economic Review*, 2017, *107* (9), 2565–99.

Maskin, Eric and Jean Tirole, "The politician and the judge: Accountability in government," *The American Economic Review*, 2004, *94* (4), 1034–1054.

McCarty, Nolan, Keith T Poole, and Howard Rosenthal, *Polarized America: The dance of ideology and unequal riches*, mit Press, 2016.

Mullainathan, Sendhil and Andrei Shleifer, "The market for news," *American Economic Review*, 2005, pp. 1031–1053.

Pariser, Eli, *The filter bubble: What the Internet is hiding from you*, Penguin UK, 2011.

Pew Research Center, "Political Polarization in the American Public," 2014.

Poole, Keith T and Howard Rosenthal, "Patterns of congressional voting," *American journal of political science*, 1991, pp. 228–278.

Prior, Markus, *Post-broadcast democracy: How media choice increases inequality in political involvement and polarizes elections*, Cambridge University Press, 2007.

_ , "The immensely inflated news audience: Assessing bias in self-reported news exposure," *Public Opinion Quarterly*, 2009, *73* (1), 130–143.

Schilbach, Frank, "Alcohol and Self-Control: A Field Experiment in India," *American Economic Review*, 2019.

Sethi, Rajiv and Muhamet Yildiz, "Communication with unknown perspectives," *Econometrica*, 2016, *84* (6), 2029–2069.

_ **and** _ , "Culture and Communication," 2016.

Stroud, Natalie Jomini, Ashley Muddiman, and Jae Kook Lee, "Seeing media as group members: An evaluation of partisan bias perceptions," *Journal of Communication*, 2014, *64* (5), 874–894.

Suen, Wing, "The Self-Perpetuation of Biased Beliefs," *The Economic Journal*, 2004, *114* (495), 377–396.

Sunstein, Cass, *Republic. com*, Princeton, NJ: Princeton University Press, 2001.

Sunstein, Cass R, *# Republic: Divided Democracy in the Age of Social Media*, Princeton University Press, 2017.

Wooldridge, Jeffrey M, *Econometric analysis of cross section and panel data*, MIT press, 2010.

YouGov, "Anglo-Saxon Attitudes: A Survey of British and American Views of the World," 2008.

Table 1: Position updating patterns: consistent with standard Bayesian models

	Subsample: random provision		
	Position (post) (1)	Position (post) (2)	Position (post) (3)
Position (pre)	.69***	.55***	.27***
	(.012)	(.015)	(.04)
Position (pre) × Confidence in prior			.36***
			(.051)
Article position	.089***	.075***	.26***
	(.0097)	(.0096)	(.042)
Article position × Confidence in prior			-.24***
			(.051)
Constant and Round FE	Y	Y	Y
Article pool FE, user controls	N	Y	Y
Number of users	1420	1417	1417
Obs. (user × issue × round)	7792	7781	6382

Notes: Each column in this table originates from a separate OLS regression of the belief after reading an article on respective regressors. User-issue-round level observations are pooled together. Included are the observations where users were provided randomly selected articles; that is, all the observations of the group with random provision (No-Choice Group) and the observations collected during the grace period for the other groups. The article pool from which an article is randomly drawn changes over time—article pool FE is added to the regression. Article position is the average position of the article reported by all the users after reading it. Confidence in prior ($\in [0,1]$) is the reported confidence in the prior position. This variable is controlled for in the regressions. Also, the interaction of this variable and the article pool FE is added to the regression for the relevant columns. Controls (user controls) include all the baseline characteristics, described in the notes of Appendix Tables B1-B3. As more fixed effects are added, the number of users and observations decrease because singletons are dropped. Standard errors (in parentheses below coefficients) are clustered by user. *** p<0.01, ** p<0.05, * p<0.1.

Table 2: Readers select their partisan media (and not prior-confirming ones)

	Source-Name Group (source selection allowed with source names shown)			Source-Name Group, 1st round only	
	Article position (1)	Article position (2)	Expected position of chosen source (3)	Article position (4)	Expected position of chosen source (5)
Position (pre)		.011	.014	.016	.02
		(.025)	(.021)	(.045)	(.036)
User's party position	.27***	.26***	.16***	.27**	.18*
	(.058)	(.057)	(.041)	(.11)	(.093)
Pool FE, const.,	Y	Y	Y	Y	Y
Round FE	Y	Y	Y	N/A	N/A
Number of users	84	84	84	83	83
Obs. (user $\times$ issue $\times$ round)	1448	1448	1478	393	396

Notes: Each column in this table originates from a separate OLS regression of a proxy for the chosen news source's position on the latest position users held before making the choice (prior) and the party position of the user's most-trusted party on the issue. User-issue-round level observations are pooled together. Round FE are added when applicable. The article pool from which an article is randomly drawn changes over time—article pool FE is added to the regression. User's party position is the average baseline position of the users who share the same most-trusted party at the baseline and have baseline confidence ([0,1]) of 0.8 or more. Expected position of the chosen source is the average position on an issue of the articles written by the sources that share the same favorite party as the source at hand. The favorite party of the source is determined by the party that has the smallest squared position distance across issues with the source. Article position is the average position of the article reported by all the users after reading it. Here, this is used as a proxy for the chosen news source's position. Standard errors (in parentheses below coefficients) are clustered by user. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Source selection: comparison between treatment groups

	Source-Name Group: identifying partisan media is *easy* (selection allowed; with source names shown)		*Source-Position Group*: identifying partisan media is *hard* (selection allowed; with only source positions shown)		No-Choice Group (Placebo; articles are randomly selected)	
	Article position (1)	Expected position of chosen source (2)	Article position (3)	Expected position of chosen source (4)	Article position (5)	Expected position of chosen source (6)
Position (pre)	.011	.014	.00008	-.019	.00011	-.013
	(.025)	(.021)	(.043)	(.031)	(.015)	(.011)
User's party position	.26***	.16***	.068	.051	.054	.041
	(.057)	(.041)	(.13)	(.098)	(.042)	(.026)
Article pool FE, constant, round FE	Y	Y	Y	Y	Y	Y
Number of users	84	84	38	38	194	194
Obs. (user × issue × round)	1448	1478	631	634	3092	3150

Notes: Each column in this table originates from a separate OLS regression of position of the chosen article on the latest position users held before making the choice (prior) and the party position of the user's most-trusted party on the issue. User-issue-round level observations are pooled together. Round FE are added when applicable. The article pool from which an article is randomly drawn changes over time—article pool FE is added to the regression. User's party position is the average baseline position of the users who share the same most-trusted party at the baseline and have baseline confidence ([0,1]) of 0.8 or more. Expected position of the chosen source is the average position on an issue of the articles written by the sources that share the same favorite party as the source at hand. The favorite party of the source is determined by the party that has the smallest squared position distance across issues with the source. Article position is the average position of the article reported by all the users after reading it. Here, this is used as a proxy for the chosen news source's position. Standard errors (in parentheses below coefficients) are clustered by user. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: More learning caused by selective exposure

	User × issue combination w/ at least one article read after pre-exposure period		
	| Position (post)- position (pre) |	| Position (post)- position (pre) |	| Position (post)- position (pre) |
	(1)	(2)	(3)
Source-Name Grp.	.025**	.023**	.023**
	(.011)	(.01)	(.01)
Source-Position Grp.	-.01	-.0034	-.0025
	(.016)	(.016)	(.016)
Const. (Omit'd: No-Choice Grp.)	.13***	.2***	.15***
	(.0055)	(.022)	(.025)
p-value: Source-Name = Source-Position	.051	.14	.15
User controls	N	Y	Y
Issue FE	N	N	Y
Number of users		336	
Obs. (user × issue)		1774	
Sample mean		.14	
Sample s.d.		.17	

Notes: Each column in this table originates from a separate OLS regression of the absolute value of change in position on the treatment group dummies. The omitted group is the No-Choice Group (random provision of articles). The Source-Name Group is allowed to select the news source with the source names shown. The Source-Position Group is allowed to select the news source with only the source positions shown. Controls (user controls) include all the baseline characteristics, described in the notes of Appendix Tables B1-B3. User-issue level observations are pooled together. The first observation on each user-issue combination is included in the regressions. Issue FE is included in some columns as indicated. Position (pre) ($\in [0, 1]$) is the reported position on the issue at the baseline. Position (post) ($\in [0, 1]$) is the reported position on the issue after reading an article in the experimental period (after pre-exposure period). Standard errors (in parentheses below coefficients) are clustered by user. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Moderation of extreme positions

	Basic	First difference (FD)	Lagged dependent var. (LDV)	Baseline balance
	$\text{Extreme}_{\text{post}}$ (1)	$\text{Extreme}_{\text{post}}$ - $\text{Extreme}_{\text{pre}}$ (2)	$\text{Extreme}_{\text{post}}$ (3)	$\text{Extreme}_{\text{pre}}$ (4)
Source-Name Grp.	-.054*	-.075**	-.061**	.021
	(.032)	(.035)	(.028)	(.036)
Source-Position Grp.	.016	-.074	-.016	.089
	(.057)	(.052)	(.046)	(.063)
$\text{Extreme}_{\text{pre}}$			.35***	
			(.028)	
Const. (Omit'd: No-Choice Grp.)	.22***	.12*	.19***	.097
	(.074)	(.074)	(.063)	(.081)
p-value: Source-Name = Source-Position	.24	.98	.34	.3
User controls, issue FE	Y	Y	Y	Y
Number of users	336	336	336	336
Obs. (user × issue)	1774	1774	1774	1774
Sample mean	.28	-.087	.28	.37
Sample s.d.	.45	.48	.45	.48

Notes: Each column in this table originates from a separate OLS regression. The omitted group is the No-Choice Group (random provision of articles). The Source-Name Group is allowed to select the news source with the source names shown. The Source-Position Group is allowed to select the news source with only the source positions shown. Controls (user controls) include all the baseline characteristics, described in the notes of Appendix Tables B1-B3. User-issue level observations are pooled together. The first observation on each user-issue combination is included in the regressions. Issue FE is included. $\text{Extreme}_{\text{post}}$ is a dummy variable indicating whether the position ($\in [0,1]$) equals to 0 or 1 after reading an article in the experimental period (after pre-exposure period). $\text{Extreme}_{\text{pre}}$ is an analogous dummy variable for the baseline position. Standard errors (in parentheses below coefficients) are clustered by user. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Article has larger influence if it is written by ideologically closer media

	Subsample: random provision	
	Position (post) (1)	Position (post) (2)
Position (pre)	.55***	.54***
	(.015)	(.029)
Position (pre) $\times$ source's ideological distance to reader		.0074
		(.086)
Article position	.075***	.2***
	(.0096)	(.028)
Article position $\times$ source's ideological distance to reader		-.35***
		(.082)
Constant and Round FE	Y	Y
Article pool FE, user controls	Y	Y
Number of users	1417	1417
Obs. (user $\times$ issue $\times$ round)	7781	7781

Notes: Each column in this table originates from a separate OLS regression of the belief after reading an article on respective regressors. User-issue-round level observations are pooled together. Included are the observations where users were provided randomly selected articles; that is, all the observations of the group with random provision (No-Choice Group) and the observations collected during the grace period for the other groups. The article pool from which an article is randomly drawn changes over time—article pool FE is added to the regression. Article position is the average position of the article reported by all the users after reading it. Source's ideological distance to reader ($\in [0,1]$) is the mean absolute distance between the baseline user position and the source's average article position. This variable is controlled for in the regressions. Also, the interaction of this variable and the article pool FE is added to the regression for the relevant columns. Controls (user controls) include all the baseline characteristics, described in the notes of Appendix Tables B1-B3. Standard errors (in parentheses below coefficients) are clustered by user. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Pre-exposure may lead to better de-biasing for future articles from same source

	Subsample: No-Choice Group (after pre-exposure period)	———, # evaluations$\geq$10
	Position (post) (1)	Position (post) (2)
Position (pre)	.58***	.57***
	(.026)	(.027)
Position (pre) $\times$ # of pre-exposure	-.035	-.017
	(.03)	(.03)
Article position	.012	.017
	(.028)	(.028)
Article position $\times$ # of pre-exposure	.064	.082*
	(.046)	(.045)
News source position	.031	.016
	(.031)	(.031)
News source position $\times$ # of pre-exposure	-.065	-.087*
	(.049)	(.047)
Constant, round FE, # of pre-exposure	Y	Y
Article pool FE, user controls	Y	Y
Number of users	207	205
Obs. (user $\times$ issue $\times$ round)	3219	2857

Notes: Each column in this table originates from a separate OLS regression of the belief after reading an article on respective regressors. User-issue-round level observations are pooled together. The article pool from which an article is randomly drawn changes over time—article pool FE is added to the regression. Article position is the average position of the article reported by all the users after reading it. News source position is the average of the article positions written by the news source on the specific issue. Included are the observations of the No-Choice Group (random provision of articles), not including the pre-exposure period. For Column 2, observations where the user is exposed to a news source that has lower than 10 total position evaluations for the issue are excluded. Controls (user controls) include all the baseline characteristics, described in the notes of Appendix Tables B1-B3. Standard errors (in parentheses below coefficients) are clustered by user. *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Pre-exposure may lead to more learning and less extremism compared with No-Choice Grp. w/o pre-exposure

	User × issue combination w/ at least one article read after pre-exposure period Source-Name and No-Choice Grp. only	
	| Position (post)- position (pre) | (1)	$\text{Extreme}_{\text{post}}$ (2)
No-Choice Grp. with pre-exposure	.013	-.05*
	(.012)	(.029)
Source-Name Grp.	.022**	-.077***
	(.01)	(.029)
Const. (Omit'd: No-Choice Grp. w/o pre-exposure)	.14***	.19***
	(.026)	(.067)
p-value: Source-Name = No-Choice w/ pre-exposure	.51	.43
User controls, issue FE	Y	Y
$\text{Extreme}_{\text{pre}}$	N	Y
Number of users	297	297
Obs. (user × issue)	1572	1572
Sample mean	.14	.27
Sample s.d.	.17	.45

Notes: Each column in this table originates from a separate OLS regression. The omitted group is the No-Choice Group (random provision of articles) with an article from a news source without pre-exposure. The Source-Name Group is allowed to select the news source with the source names shown. Controls (user controls) include all the baseline characteristics, described in the notes of Appendix Tables B1-B3. User-issue level observations are pooled together. The first observation on each user-issue combination is included in the regressions. Issue FE is included. Position (pre) ($\in [0,1]$) is the reported position on the issue at the baseline. Position (post) ($\in [0,1]$) is the reported position on the issue after reading an article in the experimental period (after pre-exposure period). $\text{Extreme}_{\text{post}}$ is a dummy variable indicating whether the position ($\in [0,1]$) equals to 0 or 1 after reading an article in the experimental period (after pre-exposure period). $\text{Extreme}_{\text{pre}}$ is an analogous dummy variable for the baseline position. Standard errors (in parentheses below coefficients) are clustered by user. *** p<0.01, ** p<0.05, * p<0.1.

Table 9: Main results are robust to re-weighting article frequency

	User × issue combination w/ at least one article read after pre-exposure period (Re-weighted)	
	| Position (post)- position (pre) |	$\text{Extreme}_{\text{post}}$
	(1)	(2)
Source-Name Grp.	.033***	-.072**
	(.011)	(.031)
Source-Position Grp.	-.003	-.026
	(.017)	(.056)
Const. (Omit'd: No-Choice Grp.)	.12***	.14*
	(.025)	(.077)
p-value: Source-Name = Source-Position	.047	.41
User controls, issue FE	Y	Y
$\text{Extreme}_{\text{pre}}$	N	Y
Number of users	327	327
Obs. (user × issue)	1474	1474
Sample mean	.14	.28
Sample s.d.	.17	.45

Notes: Each column in this table originates from a separate OLS regression. The omitted group is the No-Choice Group (random provision of articles) with an article from a news source without pre-exposure. The Source-Name Group is allowed to select the news source with the source names shown. The Source-Position Group is allowed to select the news source with only the source positions shown. Observations in the No-Choice and Source-Position Group are reweighted so that frequency of an article within each group is equalized at the level of the Source-Name Group. Controls (user controls) include all the baseline characteristics, described in the notes of Appendix Tables B1-B3. User-issue level observations are pooled together. The first observation on each user-issue combination is included in the regressions. Issue FE is included. Position (pre) ($\in [0,1]$) is the reported position on the issue at the baseline. Position (post) ($\in [0,1]$) is the reported position on the issue after reading an article in the experimental period (after pre-exposure period). $\text{Extreme}_{\text{post}}$ is a dummy variable indicating whether the position ($\in [0,1]$) equals to 0 or 1 after reading an article in the experimental period (after pre-exposure period). $\text{Extreme}_{\text{pre}}$ is an analogous dummy variable for the baseline position. Standard errors (in parentheses below coefficients) are clustered by user. *** p<0.01, ** p<0.05, * p<0.1.

Figure 1: Issues covered in the mobile application

1. Foreign policy toward North Korea
2. Law governing legislative process (minority protection vs. faster process)
3. Cash support to unemployed youth
4. Single nationalized history textbook
5. Providing free lunch to every elementary school student
6. Who should pay (central vs. regional government) for free early education for ages 3-5
7. Minimum wage
8. Flexible labor market vs. job security

Figure 2: Randomization design

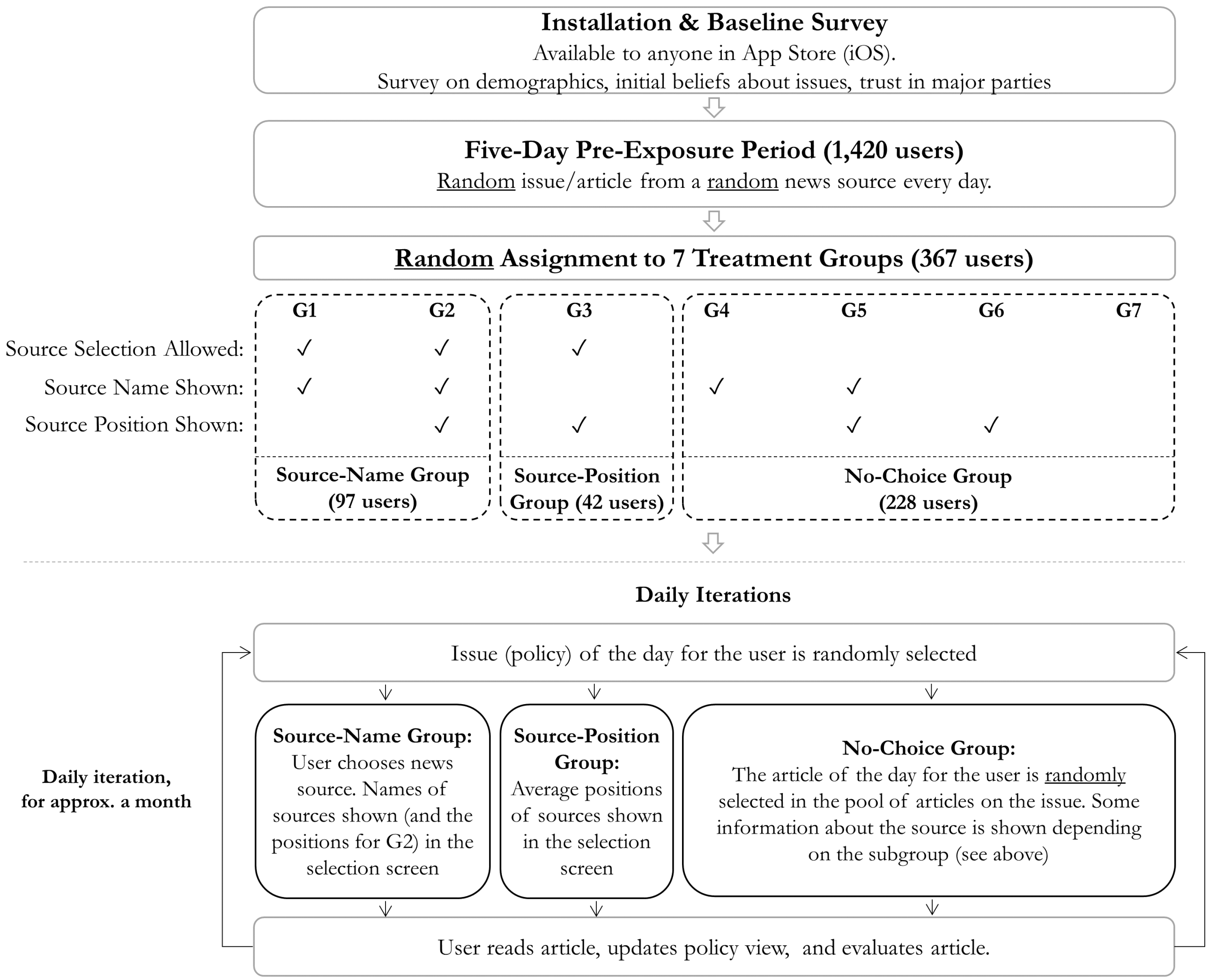


Figure 3: Screens of the app

(a) **Source-Name Group** chooses news source with names (circled) of the sources shown (and positions, if G2)

(b) **Source-Position Group** chooses source only based on source's positions (average position on the issue; circled); provider name is *not* shown

(c) **No-Choice Group** gets randomly selected article, but can also be informed about name and/or position of the source of the article

(d) After reading an article, user **updates** her **policy view** on the issue, and her **confidence** in the position.

(e) The user also reports the **subjective quality** and **position** of the article.

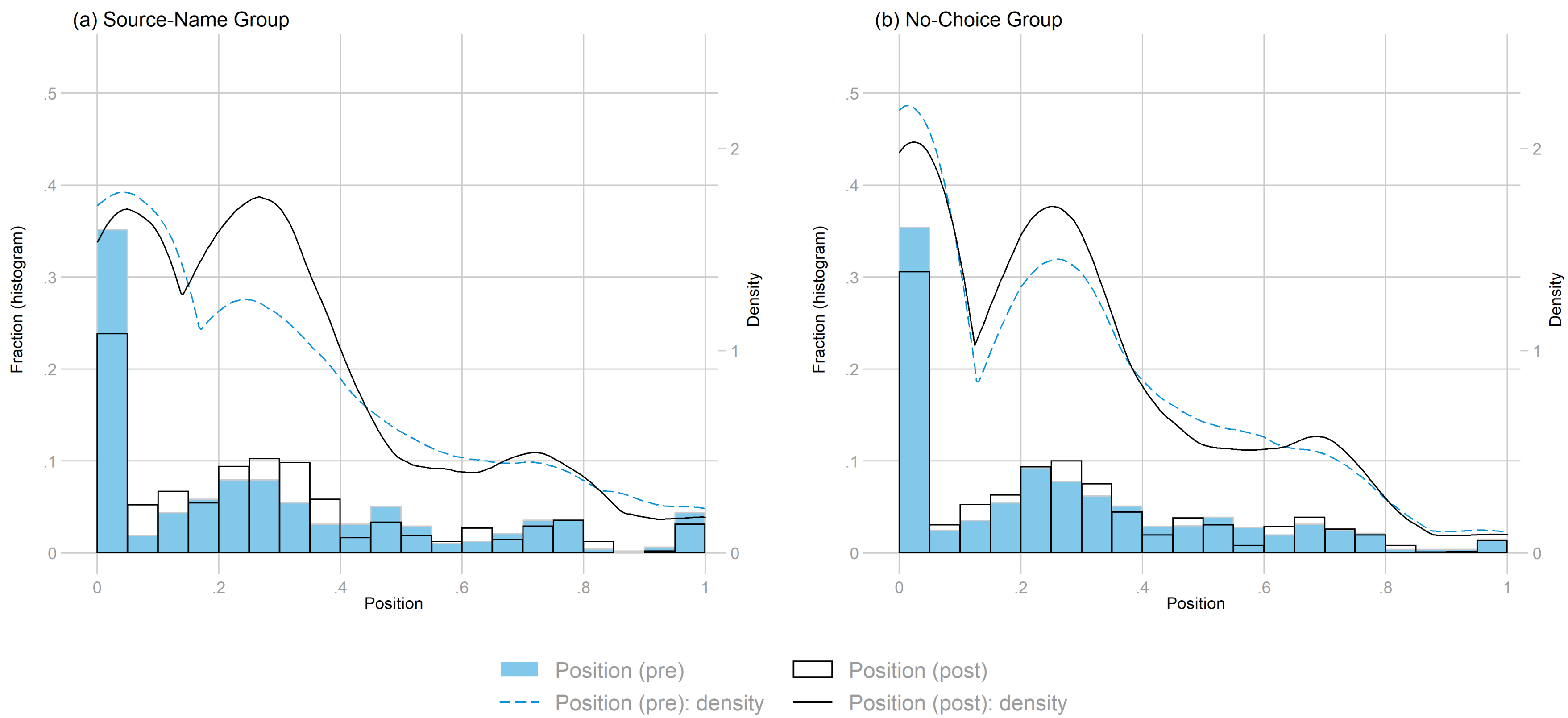
Figure 4: An example of cardinal description (foreign policy toward North Korea)

(0.0-0.2):	We should put *all* our effort into communication and collaboration
(0.2-0.4):	Communication and collaboration first, but aggression is also necessary
(0.4-0.6):	Both sides have their points; I am neutral
(0.6-0.8):	Supporting North Korea can be considered, but only after steps for denuclearization are taken
(0.8-1.0):	North Korea is our enemy; we should be aggressive in all aspects

Figure 5: Right- and left-wing extreme positions by issue

- Foreign policy toward North Korea
  - R: North Korea is our enemy; we should be aggressive in all aspects
  - L: We should put all our efforts into communication and collaboration
- Law governing legislative process (minority protection vs. faster process)
  - R: New laws with much stricter minority protections and checks on the legislative branch are "evil laws" that stymie the legislative process
  - L: Such laws should be fully and immediately enacted; intermediate steps are unnecessary
- Cash support to unemployed youth
  - R: It is a populist policy; it does not help solve the problem of youth unemployment at all
  - L: It is an absolutely crucial policy and should be prioritized over any other labor policies
- Single nationalized Korean history textbook
  - R: We should have it; many existing textbooks seriously distort Korean history
  - L: I am opposed to any type of nationalized history textbook
- Providing free lunch to every elementary school student
  - R: It is a populist policy to provide free lunch to every student, and it doesn't help society at all
  - L: It should be implemented right away; it is a right of students and a part of education
- Who should pay (central vs. regional government) for free early education for ages 3-5
  - R: Regional institutions are fully responsible for the cost of free early education
  - L: The central government is fully responsible for the cost of free early education
- Minimum wage
  - R: There should be no minimum wage at all; it should be fully left to the market economy
  - L: Minimum wage is a fundamental right of workers regardless of the economic consequences
- Flexible labor market vs. job security
  - R: Temporary and contract workers are inevitable; policies should focus on a flexible labor market for the greater good of the overall economy
  - L: We should have the fewest number of temporary and contract workers as possible

Figure 6: Histogram of positions, before and after article reading



Notes: Position (pre) ($\in [0, 1]$) is the reported position on the issue at the baseline. Position (post) ($\in [0, 1]$) is the reported position on the issue after reading an article in the experimental period (after pre-exposure period). User-issue level observations are pooled together. The No-Choice Group is given randomly selected articles. The Source-Name Group is allowed to select the news source with the source names shown.

Online Appendix

A Proofs

Derivation of Posterior Distribution

The density function of posterior belief after getting signal s_{pij} is proportional to the multiplication of the likelihood function, $l(s_{pij}|\theta_j)$, and the prior distribution, $f(\theta_j)$.

$$f(\theta_j|s_{pij}) \propto l(s_{pij}|\theta_j) \cdot f(\theta_j)$$

Since I_p is also unknown, we must add that component to get the relevant likelihood function.

$$\begin{aligned} f(\theta_j|s_{pij}) &\propto \int l(s_{pij}|\theta_j, I_{pj}) \cdot f(I_{pj}|\theta_j) dI_{pj} \times f(\theta_j) \\ &= \int l(s_{pij}|\theta_j, I_{pj}) \cdot f(I_{pj}) dI_{pj} \times f(\theta_j) \end{aligned}$$

where the final equality is due to the independence of the prior distribution of I_{pj} and θ_j. Plugging in the density functions and rearranging, we get:

$$\begin{aligned} f(\theta_j|s_{pij}) &\propto \int \exp\left(-\frac{\tau_s}{2}\left(s_{pij}-\theta_j-I_{pj}\right)^2-\frac{\tau_{Ipj}}{2}\left(I_{pji}-I_{pj}\right)^2\right) dI_{pj} \times f(\theta_j) \\ &\propto \int \exp\left(-\frac{1}{2}\left[\left(\tau_s+\tau_{Ipj}\right)I_{pj}^2-2\left\{\tau_s\left(s_{pij}-\theta_j\right)+\tau_{Ipj}I_{pji}\right\}I_{pj}+\tau_s\left(s_{pij}-\theta_j\right)^2\right]\right) dI_{pj} \times f(\theta_j) \\ &= \int \exp\left(-\frac{\tau_s+\tau_{Ipj}}{2}\left[I_{pj}^2-2\frac{\tau_s\left(s_{pij}-\theta_j\right)+\tau_{Ipj}I_{pji}}{\tau_s+\tau_{Ipj}}I_{pj}+\left(\frac{\tau_s\left(s_{pij}-\theta_j\right)+\tau_{Ipj}I_{pji}}{\tau_s+\tau_{Ipj}}\right)^2\right]-\frac{1}{2}\left[-\frac{\left(\tau_s\left(s_{pij}-\theta_j\right)+\tau_{Ipj}I_{pji}\right)^2}{\tau_s+\tau_{Ipj}}+\tau_s\left(s_{pij}-\theta_j\right)^2\right]\right) dI_{pj} \\ &\quad \times f(\theta_j) \\ &\propto \int \sqrt{\frac{\tau_s+\tau_{Ipj}}{2}} \exp\left(-\frac{1}{2}\left(I_{pj}-\frac{\tau_s\left(s_{pij}-\theta_j\right)+\tau_{Ipj}I_{pji}}{\tau_s+\tau_{Ipj}}\right)^2\right) dI_{pj} \times \exp\left(-\frac{1}{2}\left[-\frac{\left(\tau_s\left(s_{pij}-\theta_j\right)+\tau_{Ipj}I_{pji}\right)^2}{\tau_s+\tau_{Ipj}}+\tau_s\left(s_{pij}-\theta_j\right)^2\right]\right) \times f(\theta_j) \\ &= \exp\left(-\frac{1}{2}\left[-\frac{\left(\tau_s\left(s_{pij}-\theta_j\right)+\tau_{Ipj}I_{pji}\right)^2}{\tau_s+\tau_{Ipj}}+\tau_s\left(s_{pij}-\theta_j\right)^2\right]\right) \times f(\theta_j) \end{aligned}$$

where the last equality is because the integrand is a density of a normal distribution. Plugging in

θ_j, rearranging, and dropping a term unrelated to θ_j, we get:

$$
\begin{aligned}
f(\theta_j|s_{pij}) &\propto \exp\left(-\frac{1}{2}\left[-\frac{\tau_s^2}{\tau_s+\tau_{Ipj}}(\theta_j-s_{pij})^2+\frac{2\tau_s\tau_{Ipj}}{\tau_s+\tau_{Ipj}}I_{pji}(\theta_j-s_{pij})+\tau_s(\theta_j-s_{pij})^2+\tau_{ij0}\left(\theta_j-\theta_{ij0}\right)^2\right]\right) \\
&\propto \exp\left(-\frac{1}{2}\left[\frac{\tau_s\tau_{Ipj}+\tau_s\tau_{ij0}+\tau_{Ipj}\tau_{ij0}}{\tau_s+\tau_{Ipj}}\theta^2-2\left\{\frac{\tau_s\tau_{Ipj}}{\tau_s+\tau_{Ipj}}(s_{pij}-I_{pji})+\frac{\tau_s\tau_{ij0}+\tau_{Ipj}\tau_{ij0}}{\tau_s+\tau_{Ipj}}\theta_{ij0}\right\}\theta_j\right]\right) \\
&\propto \sqrt{\frac{\frac{\tau_s\tau_{Ipj}+\tau_s\tau_{ij0}+\tau_{Ipj}\tau_{ij0}}{\tau_s+\tau_{Ipj}}}{2}}\exp\left(-\frac{1}{2}\left\{\theta_j-\frac{\tau_s\tau_{Ipj}}{\tau_s\tau_{Ipj}+\tau_s\tau_{ij0}+\tau_{Ipj}\tau_{ij0}}(s_{pij}-I_{pji})-\frac{\tau_s\tau_{ij0}+\tau_{Ipj}\tau_{ij0}}{\tau_s\tau_{Ipj}+\tau_s\tau_{ij0}+\tau_{Ipj}\tau_{ij0}}\theta_{ij0}\right\}^2\right)
\end{aligned}
$$

which is the density of normal distribution with mean of $\frac{\tau_s\tau_{Ipj}}{\tau_s\tau_{Ipj}+\tau_s\tau_{ij0}+\tau_{Ipj}\tau_{ij0}}\left(s_{pij}-I_{pji}\right)-\frac{\tau_s\tau_{ij0}+\tau_{Ipj}\tau_{ij0}}{\tau_s\tau_{Ipj}+\tau_s\tau_{ij0}+\tau_{Ipj}\tau_{ij0}}\theta_{ij0}$ and precision of $\frac{\tau_s\tau_{Ipj}+\tau_s\tau_{ij0}+\tau_{Ipj}\tau_{ij0}}{\tau_s+\tau_{Ipj}}$.

Proof of Proposition 3

Proof. Minimizing the loss function is equivalent to minimizing the posterior variance.

$$
\begin{aligned}
\underset{p}{argmax} E_i\left[max_{a_{ij}}E_i\left[-(a_{ij}-\theta_j^*)^2 \mid s_{pij}\right]\right] &= \underset{p}{argmin}\left[Var_i\left(\theta_j^* \mid s_{pij}\right)\right] \\
&= \underset{p}{argmin}\frac{\tau_s+\tau_{Ipj}}{\tau_s\tau_{ij0}+\tau_{ij0}\tau_{Ipj}+\tau_{Ipj}\tau_s}
\end{aligned}
\tag{12}
$$

According to Assumption 1, $\tau_{IFj} > \tau_{IAj}$. The result immediately follows. □

Proof of Proposition 4

Proof. (i) Starting from Equation (4), it is trivial to derive:

$$\left|\theta_{ij1}-\theta_{ij0}\right| = \gamma_{pij}\left|s_{pij}-I_{pji}-\theta_{ij0}\right|$$

where $\gamma_{pij} \equiv \frac{\tau_{Ipj}\tau_s}{\tau_s\tau_{ij0}+\tau_{ij0}\tau_{Ipj}+\tau_{Ipj}\tau_s} \in [0,1]$. Plugging in Equation (1) and rewriting I_{pji} and θ_{ij0} with the error terms, we get

$$
\begin{aligned}
\left|\theta_{ij1}-\theta_{ij0}\right| &= \gamma_{pij}\left|\theta_j^*+I_{pj}^*+\varepsilon_{ij}-I_{pj}^*-\nu_{pi}-\theta_j^*-\eta_{ij0}\right| \\
&= \gamma_{pij}\left|\varepsilon_{ij}-\nu_{pji}-\eta_{ij0}\right|
\end{aligned}
$$

Taking expectations over the continuum of agents,

$$E\left[\left|\theta_{ij1}-\theta_{ij0}\right| \mid p=P\right] = \int\int\int\int \gamma_{Pij}\left|\varepsilon_{ij}-\nu_{Pji}-\eta_{ij0}\right| dF\left(\varepsilon_{ij}\right)dF\left(\nu_{Pji}\right)dF\left(\eta_{ij0}\right)dF\left(\gamma_{Fij},\gamma_{Aij}\right)$$

We need to prove that $\frac{E\left[\left|\theta_{ij1}-\theta_{ij0}\right| \mid p=F\right]}{E\left[\left|\theta_{ij1}-\theta_{ij0}\right| \mid p=A\right]} > 1$. Take M_j such that:

$$0 < M_j < \frac{\int\int\int \gamma_{Fij}\left|\varepsilon_{ij}-\eta_{ij0}\right| dF\left(\varepsilon_{ij}\right) dF\left(\eta_{ij0}\right) dF\left(\gamma_{Fij},\gamma_{Aij}\right) - \int\int\int \gamma_{Aij}\left|\varepsilon_{ij}-\eta_{ij0}\right| dF\left(\varepsilon_{ij}\right) dF\left(\eta_{ij0}\right) dF\left(\gamma_{Fij},\gamma_{Aij}\right)}{2} \tag{13}$$

which exists because $\gamma_{Fij} > \gamma_{Aij}$ according to Assumption 1. Then,

$$\begin{aligned}
\frac{E\left[\left|\theta_{ij1}-\theta_{ij0}\right| \mid p=F\right]}{E\left[\left|\theta_{ij1}-\theta_{ij0}\right| \mid p=A\right]} &= \frac{\int\int\int\int \gamma_{Fij}\left|\varepsilon_{ij}-\nu_{Fji}-\eta_{ij0}\right| dF\left(\varepsilon_{ij}\right) dF\left(\nu_{Fji}\right) dF\left(\eta_{ij0}\right) dF\left(\gamma_{Fij},\gamma_{Aij}\right)}{\int\int\int\int \gamma_{Aij}\left|\varepsilon_{ij}-\nu_{Aji}-\eta_{ij0}\right| dF\left(\varepsilon_{ij}\right) dF\left(\nu_{Aji}\right) dF\left(\eta_{ij0}\right) dF\left(\gamma_{Fij},\gamma_{Aij}\right)} \\
&> \frac{\int \gamma_{Fij}\left|\varepsilon_{ij}-\eta_{ij0}\right| dF(\cdot) - \int \gamma_{Fij}\left|\nu_{Fji}\right| dF(\cdot)}{\int \gamma_{Aij}\left|\varepsilon_{ij}-\eta_{ij0}\right| dF(\cdot) + \int \gamma_{Aij}\left|\nu_{Aji}\right| dF(\cdot)} \\
&> \frac{\int\int\int \gamma_{Fij}\left|\varepsilon_{ij}-\eta_{ij0}\right| dF\left(\varepsilon_{ij}\right) dF\left(\eta_{ij0}\right) dF\left(\gamma_{Fij},\gamma_{Aij}\right) - M_j}{\int\int\int \gamma_{Aij}\left|\varepsilon_{ij}-\eta_{ij0}\right| dF\left(\varepsilon_{ij}\right) dF\left(\eta_{ij0}\right) dF\left(\gamma_{Aij},\gamma_{Aij}\right) + M_j} \\
&> 1
\end{aligned}$$

I used the triangle inequality both on the denominator and the numerator for the first inequality. $dF(\cdot)$ is an abuse of notation for simplicity, indicating the distribution of all relevant variables. For the second inequality, I applied Assumption 2. The final inequality can be derived by applying Inequation (13).

$\frac{E\left[\left|\theta_{ij1}-\theta_{ij0}\right| \mid p=F\right]}{E\left[\left|\theta_{ij1}-\theta_{ij0}\right| \mid p=A\right]} > 1$ implies that there will be more learning when the reader receives the signal from the familiar source than from the alien source. It then immediately implies that selective exposure to the familiar source will give, on average, a bigger movement in beliefs than an equal-chance encounter.

(ii) Similarly, using Equation (4) and (1), and rewriting I_{pji} and θ_{ij0} with the error terms, it is trivial to derive:

$$\left|\theta_{ij1}-\theta_j^*\right| = \left|(1-\gamma_{pij})\eta_{ij0} + \gamma_{pij}(\varepsilon_{ij}-\nu_{pji})\right|$$

where $\gamma_{pij} \equiv \frac{\tau_{Ipj}\tau_s}{\tau_s\tau_{ij0}+\tau_{ij0}\tau_{Ipj}+\tau_{Ipj}\tau_s} \in [0,1]$. Taking expectations over the continuum of agents,

$$E\left[\left|\theta_{ij1}-\theta_j^*\right| \mid p=P\right] = \int\int\int\int \left|(1-\gamma_{Pij})\eta_{ij0} + \gamma_{Pij}(\varepsilon_{ij}-\nu_{Pji})\right| dF\left(\varepsilon_{ij}\right) dF\left(\nu_{Pji}\right) dF\left(\eta_{ij0}\right) dF\left(\gamma_{Fij},\gamma_{Aij}\right)$$

We need to prove that $\frac{E\left[\left|\theta_{ij1}-\theta_j^*\right| \mid p=F\right]}{E\left[\left|\theta_{ij1}-\theta_j^*\right| \mid p=A\right]} < 1$. First, take M_j such that:

$$0 < M_j < \frac{\int \left|(1-\gamma_{Aij})\eta_{ij0} + \gamma_{Aij}\varepsilon_{ij}\right| dF(\cdot) - \int \left|(1-\gamma_{Fij})\eta_{ij0} + \gamma_{Fij}\varepsilon_{ij}\right| dF(\cdot)}{2} \tag{14}$$

We can show that the upper bound of Inequation (14) is larger than 0 as follows.

$$
\begin{aligned}
&\int \left|(1-\gamma_{Aij})\eta_{ij0}+\gamma_{Aij}\varepsilon_{ij}\right| dF(\cdot) - \int \left|(1-\gamma_{Fij})\eta_{ij0}+\gamma_{Fij}\varepsilon_{ij}\right| dF(\cdot) \\
> &\int (1-\gamma_{Aij})\left|\eta_{ij0}\right| dF(\cdot) - \int \gamma_{Aij}\left|\varepsilon_{ij}\right| dF(\cdot) - \int (1-\gamma_{Fij})\left|\eta_{ij0}\right| dF(\cdot) - \int \gamma_{Fij}\left|\varepsilon_{ij}\right| dF(\cdot) \\
= &\int (\gamma_{Fij}-\gamma_{Aij})\left|\eta_{ij0}\right| dF(\cdot) - \int (\gamma_{Fij}+\gamma_{Aij})\left|\varepsilon_{ij}\right| dF(\cdot) \\
> &0
\end{aligned}
$$

where I used the triangle inequality for the first inequality, and the last inequality is from the assumption that $E\left[\left(\gamma_{Fij}-\gamma_{Aij}\right)\left|\eta_{ij0}\right|\right] > E\left[\left(\gamma_{Fij}+\gamma_{Aij}\right)\left|\varepsilon_{ij}\right|\right]$. Then,

$$
\begin{aligned}
\frac{E\left[\left|\theta_{ij1}-\theta_j^*\right| \mid p=F\right]}{E\left[\left|\theta_{ij1}-\theta_j^*\right| \mid p=A\right]} &= \frac{\int\int\int\int \left|(1-\gamma_{Fij})\eta_{ij0}+\gamma_{Fij}(\varepsilon_{ij}-\nu_{Fji})\right| dF(\varepsilon_{ij})\, dF(\nu_{Fji})\, dF(\eta_{ij0})\, dF(\gamma_{Fij},\gamma_{Aij})}{\int\int\int\int \left|(1-\gamma_{Aij})\eta_{ij0}+\gamma_{Aij}(\varepsilon_{ij}-\nu_{Aji})\right| dF(\varepsilon_{ij})\, dF(\nu_{Aji})\, dF(\eta_{ij0})\, dF(\gamma_{Fij},\gamma_{Aij})} \\
&< \frac{\int \left|(1-\gamma_{Fij})\eta_{ij0}+\gamma_{Fij}\varepsilon_{ij}\right| dF(\cdot) + \int \gamma_{Fij}\left|\nu_{Fji}\right| dF(\cdot)}{\int \left|(1-\gamma_{Aij})\eta_{ij0}+\gamma_{Aij}\varepsilon_{ij}\right| dF(\cdot) - \int \gamma_{Aij}\left|\nu_{Aji}\right| dF(\cdot)} \\
&< \frac{\int \left|(1-\gamma_{Fij})\eta_{ij0}+\gamma_{Fij}\varepsilon_{ij}\right| dF(\cdot) + M_j}{\int \left|(1-\gamma_{Aij})\eta_{ij0}+\gamma_{Aij}\varepsilon_{ij}\right| dF(\cdot) - M_j} \\
&< 1
\end{aligned}
$$

I used the triangle inequality both on the denominator and the numerator for the first inequality. $dF(\cdot)$ is an abuse of notation for simplicity, indicating the distribution of all relevant variables. For the second inequality, I applied Assumption 2. The final inequality can be derived by applying Inequation (14). $\square$

B Appendix Tables and Figures

See next page.

Table B1: Balance: random provision of articles

	Subsample: exposed to randomly chosen articles								
	Age (1)	Gender (2)	Polit. knowldg. Score (3)	Trust Dem (4)	Trust *Saenuri* (5)	TV as news source (6)	Internet as news source (7)	Prior position (8)	Prior conf. (9)
Article Position	-.73*	-.026	.022	-.027	-.0053	.025	-.0016	-.0029	.0011
	(.42)	(.025)	(.088)	(.028)	(.022)	(.027)	(.019)	(.013)	(.011)
Outcome Mean	31	1.6	2.9	.14	-.76	.56	.88	.25	.74
Outcome S.d.	7.7	.49	1.6	.53	.38	.5	.33	.26	.23
Pool FE, Round FE					Y				
Number of users					1420				
Obs. (user $\times$ issue $\times$ round)					7794				

Notes: Each column in this table originates from a separate OLS regression of baseline characteristics of the users on the position of the article to which the user was randomly exposed. Baseline characteristics are: (1) age, (2) gender ($1 = f$, $2 = m$), (3) political knowledge score ($[0,5]$ scale; based on a quiz conducted at the baseline), (4) trust in Democratic Party ($[-1,1]$ scale), (5) trust in Saenuri Party, (6) dummy variable indicating whether ever used TV as a news source over the last 7 days, (7) dummy variable indicating whether ever used Internet as a news source over the last 7 days, (8) the latest position reported before being exposed to the relevant article ($[0,1]$ scale), and (9) confidence in the lastest position ($[0,1]$ scale). User-issue-round level observations are pooled together. Article pool FE (includes issue FE) and Round FE are included. Article position is the average position of the article reported by all the users after reading it. Standard errors (in parentheses below coefficients) are clustered by user. *** p<0.01, ** p<0.05, * p<0.1.

Table B2: Balance and attrition: between treatment groups

	Subsample: assigned to a treatment group								
	Age (1)	Gender (2)	Polit. knowldg. Score (3)	Trust Dem (4)	Trust *Saenuri* (5)	TV as news source (6)	Internet as news source (7)	Initial extreme position (8)	Retention (9)
	Panel A: everyone who finished the pre-exposure period								
Source-Name Grp.	-.44	.056	-.12	.066	.039	.015	.051	.012	.024
	(.87)	(.058)	(.2)	(.064)	(.05)	(.06)	(.034)	(.037)	(.042)
Source-Position Grp.	1.2	.1	.29	.13	-.048	-.036	-.02	.12**	.054
	(1.3)	(.077)	(.29)	(.095)	(.062)	(.084)	(.058)	(.056)	(.058)
Const. (Omit'd: No-Choice Grp.)	31***	1.6***	2.9***	.085**	-.75***	.58***	.88***	.34***	.59***
	(.49)	(.032)	(.11)	(.036)	(.026)	(.033)	(.022)	(.021)	(.023)
Outcome Mean	31	1.6	2.9	.12	-.74	.58	.89	.36	.6
Outcome S.d.	7.4	.48	1.6	.54	.4	.49	.32	.48	.49
	Panel B: those who read an article after the pre-exposure period								
Source-Name Grp.	.087	.072	-.16	.016	.085	.024	.032	.0063	0
	(1.1)	(.066)	(.22)	(.076)	(.057)	(.068)	(.041)	(.044)	(.)
Source-Position Grp.	1.2	.093	.51*	.17	-.041	-.039	-.031	.14**	0
	(1.5)	(.09)	(.3)	(.1)	(.053)	(.095)	(.067)	(.065)	(.)
Const. (Omit'd: No-Choice Grp.)	31***	1.6***	3***	.11**	-.79***	.59***	.88***	.35***	1
	(.56)	(.038)	(.12)	(.042)	(.026)	(.038)	(.024)	(.024)	(.)
Outcome Mean	31	1.6	3	.13	-.77	.59	.89	.37	1
Outcome S.d.	7.5	.48	1.6	.55	.37	.49	.32	.48	0
Issue FE	Y								
Number of users	A: 367, B: 336								
Obs. (user × issue)	A: 2936, B: 1774								

Notes: Each column in this table originates from a separate OLS regression of baseline characteristics of the users on the treatment status. The omitted group is the No-Choice Group (random provision of articles). The Source-Name Group is allowed to select the news source with the source names shown. The Source-Position Group is allowed to select the news source with only the source positions shown. Baseline characteristics are the same as the previous table for Columns (1)-(7). The remaining are: (8) a dummy variable indicating whether the baseline position is either 0 or 1 and (9) a dummy varible indicating whether the user read at least 1 article on the issue after being assigned to a treatment group. User-issue level observations are pooled together. Issue FE is included. Standard errors (in parentheses below coefficients) are clustered by user. *** p<0.01, ** p<0.05, * p<0.1.

Table B3: Attrition is not explained by key variables

	Users assigned to treat. grp.		
	Attrition Dummy (1)	Attrition Dummy (2)	Attrition Dummy (3)
Source-Name Grp.	-.024		-.03
	(.041)		(.042)
Source-Position Grp.	-.054		-.044
	(.058)		(.057)
Init. extrm. posi.		-.034	.0032
		(.028)	(.028)
Init. confidence		.0025	.069
		(.06)	(.06)
Age			-.0051*
			(.0026)
Gender			.038
			(.041)
Poit. knowldg. score			-.019
			(.012)
Trust Dem.			.013
			(.043)
Trust Saenuri			.041
			(.05)
TV as news source			-.031
			(.08)
Internet as news source			.059
			(.093)
Other baseline char.	N	N	Y
p-value for F-stat	.6	.45	.16
Number of users	367	367	367
Obs. (user $\times$ issue)	2936	2936	2936

Notes: Each column in this table originates from a separate OLS regression of attrition dummy on key baseline characteristics and the treatment status. See the notes of the previous two tables for the explaination of each independent variable. Other baseline controls include (i) time spent on news media last week (minutes), (ii) trust in Justice Party (the most progressive party in Korea's National Assembly), (iii) dummy variables indicating whether spent time on, respectively, ground wave TV, cable TV, Internet new portal, newspaper websites, and on reading paper news for news consumption last week. Standard errors (in parentheses below coefficients) are clustered by user. *** p<0.01, ** p<0.05, * p<0.1.

Table B4: The effect of article: not transient

	Subsample: random provision		
	Position after 1st article	Position after 2nd article	Position after 3rd article
	(1)	(2)	(3)
Position (before reading 1st article)	.55***	.51***	.5***
	(.015)	(.023)	(.028)
Article position [Current]	.075***		
	(.0096)		
Article position [Lag 1]		.042***	
		(.014)	
Article position [Lag 2]			.033**
			(.015)
Constant, Round & Pool FE, controls	Y	Y	Y
Number of users	1417	599	278
Obs. (user $\times$ issue $\times$ round)	7781	3861	2424

Notes: Each column in this table originates from a separate OLS regression of the belief after reading an article on respective regressors. User-issue-round level observations are pooled together. Included are the observations where users were provided randomly selected articles; that is, all the observations of the group with random provision (No-Choice Group) and the observations collected during the grace period for the other groups. The article pool from which an article is randomly drawn changes over time—article pool FE is added to the regression. Note that lagged article pool FE is added for lagged regressions in Columns 2 and 3. Article position is the average position of the article reported by all the users after reading it. Controls (user controls) include all the baseline characteristics, described in the notes of Appendix Tables B1-B3. Standard errors (in parentheses below coefficients) are clustered by user. *** p<0.01, ** p<0.05, * p<0.1.

Table B5: Position updating patterns: dropping obs. w/ extreme priors

	Subsample: random provision excluding position (pre) $\in \{0,1\}$		
	Position (post) (1)	Position (post) (2)	Position (post) (3)
Position (pre)	.66***	.53***	.28***
	(.014)	(.018)	(.05)
Position (pre) $\times$ Confidence in prior			.34***
			(.066)
Article position	.12***	.11***	.28***
	(.014)	(.014)	(.05)
Article position $\times$ Confidence in prior			-.24***
			(.067)
Constant and Round FE	Y	Y	Y
Article pool FE, user controls	N	Y	Y
Number of users	1092	1089	1089
Obs. (user $\times$ issue $\times$ round)	4929	4913	4049

Notes: Each column in this table originates from a separate OLS regression of the belief after reading an article on respective regressors. User-issue-round level observations are pooled together. Included are the observations where users were provided randomly selected articles; that is, all the observations of the group with random provision (No-Choice Group) and the observations collected during the grace period for the other groups. The article pool from which an article is randomly drawn changes over time—article pool FE is added to the regression. Article position is the average position of the article reported by all the users after reading it. Confidence in prior ($\in [0,1]$) is the reported confidence in the prior position. This variable is controlled for in the regressions. Also, the interaction of this variable and the article pool FE is added to the regression for the relevant columns. Controls (user controls) include all the baseline characteristics, described in the notes of Appendix Tables B1-B3. As more fixed effects are added, the number of users and observations decrease because singletons are dropped. Standard errors (in parentheses below coefficients) are clustered by user. *** p<0.01, ** p<0.05, * p<0.1.

Table B6: Position updating patterns: news source fixed effects

	Subsample: random provision		
	Position (post) (1)	Position (post) (2)	Position (post) (3)
Position (pre)	.55***	.55***	.54***
	(.015)	(.015)	(.016)
Article position	.075***	.05**	.054**
	(.0096)	(.021)	(.022)
Pool & round FE, const., controls	Y	Y	Y
Issue × news source FE	N	Y	Y
Issue × news source × most trusted party FE	N	N	Y
Number of users	1417	1417	1297
Obs. (user × issue × round)	7781	7777	7254

Notes: Each column in this table originates from a separate OLS regression of the belief after reading an article on respective regressors. User-issue-round level observations are pooled together. Included are the observations where users were provided randomly selected articles; that is, all the observations of the group with random provision (No-Choice Group) and the observations collected during the grace period for the other groups. The article pool from which an article is randomly drawn changes over time—article pool FE is added to the regression. Article position is the average position of the article reported by all the users after reading it. Confidence in prior ($\in [0,1]$) is the reported confidence in the prior position. Controls (user controls) include all the baseline characteristics, described in the notes of Appendix Tables B1-B3. As more fixed effects are added, the number of users and observations decrease because singletons are dropped. Standard errors (in parentheses below coefficients) are clustered by user. *** p<0.01, ** p<0.05, * p<0.1.

Table B7: Readers select their partisan media: subsample analysis

	Source-Name Group	Source-Name Group, *Distance* > p(50)	Source-Name Group, *Distance* > p(75)
	Article position (1)	Article position (2)	Article position (3)
Position (pre)	.011	-.019	-.045
	(.025)	(.032)	(.052)
User's party position	.26***	.26***	.35***
	(.057)	(.077)	(.091)
Article pool FE, constant, round FE	Y	Y	Y
Number of users	84	79	65
Obs. (user $\times$ issue $\times$ round)	1448	691	311

Notes: Each column in this table originates from a separate OLS regression of position of the chosen article on the latest position users held before making the choice (prior) and the party position of the user's most-trusted party on the issue. The first column is a full-sample benchmark, including the entire Source-Name Group. The second column shows the subsample where the prior position and the party position are farther than the median distance (0.12). The third is similar, except the cutoff is at the 75th percentile (0.22). User-issue-round level observations are pooled together. Round FE are added to the regressions. The article pool from which an article is randomly drawn changes over time—article pool FE is added to the regression. User's party position is the average baseline position of the users who share the same most-trusted party at the baseline and have baseline confidence ([0,1]) of 0.8 or more. Article position is the average position of the article reported by all the users after reading it. Here, this is used as a proxy for the chosen news source's position. Standard errors (in parentheses below coefficients) are clustered by user. *** p<0.01, ** p<0.05, * p<0.1.

Table B8: Source selection: comparison between treatment groups (finest subgroups)

	All source selections after pre-exposure period	
	Article position (1)	Expected position of chosen source (2)
User's party position	.27***	.19***
	(.089)	(.057)
——— × Group 2 (subgroup of Source-Name Group)	.075	-.0089
	(.12)	(.088)
——— × Group 3 (Source-Position Group)	-.21	-.14
	(.15)	(.11)
——— × Group 4 (subgroup of No-Choice Group)	-.12	-.17***
	(.12)	(.064)
——— × Group 5 (subgroup of No-Choice Group)	-.16	-.14
	(.15)	(.11)
——— × Group 6 (subgroup of No-Choice Group)	-.33**	-.16*
	(.14)	(.091)
——— × Group 7 (subgroup of No-Choice Group)	-.4***	-.2*
	(.14)	(.11)
Constant, pool FE × subgroup FE, round FE × subgroup FE, position (pre) × subgroup FE	Y	Y
Number of users	315	315
Obs. (user × issue × round)	4922	5012

Notes: Each column in this table originates from a separate OLS regression of a proxy for the chosen news source's position on the latest position users held before making the choice (prior) and the party position of the user's most-trusted party on the issue, interacted with the indicator variable for the finest subgroups. User-issue-round level observations are pooled together. Round FE are added to the regressions. The article pool from which an article is randomly drawn changes over time—article pool FE is added to the regression. User's party position is the average baseline position of the users who share the same most-trusted party at the baseline and have baseline confidence ([0,1]) of 0.8 or more. Expected position of the chosen source is the average position on an issue of the articles written by the sources that share the same favorite party as the source at hand. The favorite party of the source is determined by the party that has the smallest squared position distance across issues with the source. Article position is the average position of the article reported by all the users after reading it. Here, this is used as a proxy for the chosen news source's position. Standard errors (in parentheses below coefficients) are clustered by user. *** p<0.01, ** p<0.05, * p<0.1.

Table B9: Learning: robust to drop any issue

	User × issue combination w/ at least one article read after pre-exposure period (each column excludes one issue) *Dependent variable*: | position (post) - position (pre) |							
	Exc.: issue 1 (1)	Exc.: issue 2 (2)	Exc.: issue 3 (3)	Exc.: issue 4 (4)	Exc.: issue 5 (5)	Exc.: issue 6 (6)	Exc.: issue 7 (7)	Exc.: issue 8 (8)
Source-Name Grp.	.021*	.029**	.025**	.018*	.022**	.022**	.025**	.019*
	(.011)	(.011)	(.011)	(.01)	(.011)	(.01)	(.011)	(.01)
Source-Position Grp.	-.0073	-.0072	.0016	-.0041	-.0069	.00094	.0028	-.00059
	(.017)	(.017)	(.018)	(.016)	(.016)	(.017)	(.018)	(.017)
Const. (Omit'd: No-Choice Grp.)	.18***	.14***	.16***	.16***	.14***	.16***	.15***	.15***
	(.024)	(.027)	(.027)	(.025)	(.026)	(.025)	(.025)	(.025)
p-value: Source-Name = Source-Position	.11	.055	.21	.21	.091	.25	.24	.28
User controls, issue FE	Y							
Number of users	330	336	330	328	332	330	333	333
Obs. (user × issue)	1546	1547	1548	1543	1561	1563	1547	1563
Sample mean	.15	.14	.14	.14	.14	.13	.14	.14
Sample s.d.	.17	.17	.17	.17	.16	.16	.17	.17

Notes: Each column in this table originates from a separate OLS regression of the absolute value of change in position on the treatment group dummies. The omitted group is the No-Choice Group (random provision of articles). The Source-Name Group is allowed to select the news source with the source names shown. The Source-Position Group is allowed to select the news source with only the source positions shown. Controls (user controls) include all the baseline characteristics, described in the notes of Appendix Tables B1-B3. User-issue level observations are pooled together. The first observation on each user-issue combination is included in the regressions. Issue FE is included in some columns as indicated. Position (pre) ($\in [0,1]$) is the reported position on the issue at the baseline. Position (post) ($\in [0,1]$) is the reported position on the issue after reading an article in the experimental period (after pre-exposure period). Standard errors (in parentheses below coefficients) are clustered by user. *** p<0.01, ** p<0.05, * p<0.1.

Table B10: Moderation of extreme positions: robust to drop any issue

	User × issue combination w/ at least one article read after pre-exposure period (each column excludes one issue) *Dependent variable*: $\text{Extreme}_{\text{post}}$							
	Exc.: issue 1	Exc.: issue 2	Exc.: issue 3	Exc.: issue 4	Exc.: issue 5	Exc.: issue 6	Exc.: issue 7	Exc.: issue 8
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Source-Name Grp.	-.053*	-.067**	-.061**	-.06**	-.071**	-.065**	-.054*	-.058**
	(.03)	(.032)	(.028)	(.029)	(.028)	(.03)	(.029)	(.029)
Source-Position Grp.	-.022	-.0012	-.022	-.01	-.029	-.0046	-.012	-.023
	(.045)	(.052)	(.044)	(.047)	(.046)	(.049)	(.05)	(.046)
$\text{Extreme}_{\text{pre}}$	.36***	.34***	.34***	.34***	.35***	.36***	.35***	.35***
	(.03)	(.029)	(.029)	(.029)	(.028)	(.029)	(.03)	(.03)
Const. (Omit'd: No-Choice Grp.)	-.047	.18***	.19***	.17***	.2***	.16**	.19***	.16**
	(.057)	(.068)	(.064)	(.065)	(.065)	(.066)	(.066)	(.064)
p-value: Source-Name = Source-Position	.52	.22	.39	.3	.37	.24	.41	.48
User controls, issue FE	Y							
Number of users	330	336	330	328	332	330	333	333
Obs. (user × issue)	1546	1547	1548	1543	1561	1563	1547	1563
Sample mean	.24	.3	.28	.29	.28	.3	.28	.28
Sample s.d.	.43	.46	.45	.45	.45	.46	.45	.45

Notes: Each column in this table originates from a separate OLS regression. The omitted group is the No-Choice Group (random provision of articles). The Source-Name Group is allowed to select the news source with the source names shown. The Source-Position Group is allowed to select the news source with only the source positions shown. Controls (user controls) include all the baseline characteristics, described in the notes of Appendix Tables B1-B3. User-issue level observations are pooled together. The first observation on each user-issue combination is included in the regressions. Issue FE is included. $\text{Extreme}_{\text{post}}$ is a dummy variable indicating whether the position ($\in [0,1]$) equals to 0 or 1 after reading an article in the experimental period (after pre-exposure period). $\text{Extreme}_{\text{pre}}$ is an analogous dummy variable for the baseline position. Standard errors (in parentheses below coefficients) are clustered by user. *** p<0.01, ** p<0.05, * p<0.1.

Table B11: Moderation of extreme positions (continuous measure of extremism)

	Basic	First difference (FD)	Lagged dependent var. (LDV)	Baseline balance
	$\text{Extreme}^{\text{CTS}}_{\text{post}}$	$\text{Extreme}^{\text{CTS}}_{\text{post}}$ - $\text{Extreme}^{\text{CTS}}_{\text{pre}}$	$\text{Extreme}^{\text{CTS}}_{\text{post}}$	$\text{Extreme}^{\text{CTS}}_{\text{pre}}$
	(1)	(2)	(3)	(4)
Source-Name Grp.	-.0013	-.019*	-.0086	.018
	(.011)	(.01)	(.0091)	(.012)
Source-Position Grp.	.0019	-.042***	-.016	.044**
	(.019)	(.015)	(.015)	(.018)
$\text{Extreme}^{\text{CTS}}_{\text{pre}}$			.41***	
			(.028)	
Const. (Omit'd: No-Choice Grp.)	.25***	.00097	.14***	.24***
	(.026)	(.023)	(.021)	(.026)
p-value: Source-Name = Source-Position	.87	.15	.64	.17
User controls, issue FE	Y	Y	Y	Y
Number of users	336	336	336	336
Obs. (user $\times$ issue)	1774	1774	1774	1774
Sample mean	.31	-.016	.31	.33
Sample s.d.	.16	.15	.16	.17

Notes: Each column in this table originates from a separate OLS regression. The omitted group is the No-Choice Group (random provision of articles). The Source-Name Group is allowed to select the news source with the source names shown. The Source-Position Group is allowed to select the news source with only the source positions shown. Controls (user controls) include all the baseline characteristics, described in the notes of Appendix Tables B1-B3. User-issue level observations are pooled together. The first observation on each user-issue combination is included in the regressions. Issue FE is included. $\text{Extreme}^{\text{CTS}}_{\text{post}}$ is the absolute distance between 0.5 and the position ($\in [0,1]$) reported after reading an article in the experimental period (after pre-exposure period) $\text{Extreme}^{\text{CTS}}_{\text{pre}}$ is an analogous variable for the baseline position. Standard errors (in parentheses below coefficients) are clustered by user. *** p<0.01, ** p<0.05, * p<0.1.

Table B12: Evolution of positions by finest subgroup

	User × issue combination w/ at least one article read after pre-exposure period	
	ǀ Position (post) - position (pre) ǀ	$\text{Extreme}_{\text{post}}$
	(1)	(2)
Group 2 (subgroup of Source-Name Group)	.012	-.08*
	(.018)	(.046)
Group 3 (Source-Position Group)	-.019	.0048
	(.021)	(.051)
Group 4 (subgroup of No-Choice Group)	-.021	.021
	(.017)	(.045)
Group 5 (subgroup of No-Choice Group)	.0053	-.012
	(.017)	(.045)
Group 6 (subgroup of No-Choice Group)	-.029*	.039
	(.016)	(.048)
Group 7 (subgroup of No-Choice Group)	-.016	.022
	(.018)	(.042)
$\text{Extreme}_{\text{pre}}$		.35***
		(.027)
Const. (Omit'd: Group 1 ∈ Source-Name Grp.)	.16***	.18**
	(.029)	(.072)
User controls, issue FE	Y	Y
Number of users	336	336
Obs. (user × issue)	1774	1774
Sample mean	.14	.28
Sample s.d.	.17	.45

Notes: Each column in this table originates from a separate OLS regression. The omitted group is the No-Choice Group (random provision of articles). The Source-Name Group is allowed to select the news source with the source names shown. The Source-Position Group is allowed to select the news source with only the source positions shown. Controls (user controls) include all the baseline characteristics, described in the notes of Appendix Tables B1-B3. User-issue level observations are pooled together. The first observation on each user-issue combination is included in the regressions. Issue FE is included. Position (pre) ($\in [0,1]$) is the reported position on the issue at the baseline. Position (post) ($\in [0,1]$) is the reported position on the issue after reading an article in the experimental period (after pre-exposure period). $\text{Extreme}_{\text{post}}$ is a dummy variable indicating whether the position ($\in [0,1]$) equals to 0 or 1 after reading an article in the experimental period (after pre-exposure period). $\text{Extreme}_{\text{pre}}$ is an analogous dummy variable for the baseline position. Standard errors (in parentheses below coefficients) are clustered by user. *** p<0.01, ** p<0.05, * p<0.1.

Table B13: Trust in exogneously selected news source does not explain learning or moderation

	User × issue combination w/ at least one article read after pre-exposure period (No-Choice Group only)			
	| Position (post)-position (pre) |	$\text{Extreme}_{\text{post}}$	| Position (post)-position (pre) |	$\text{Extreme}_{\text{post}}$
	(1)	(2)	(3)	(4)
Trust in news source (all)	-.0033	.002		
	(.015)	(.04)		
Trust in news source (exogenous)			.0041	.017
			(.024)	(.055)
Constant, user controls, issue FE	Y	Y	Y	Y
$\text{Extreme}_{\text{pre}}$	N	Y	N	Y
Number of users	157	157	133	133
Obs. (user × issue)	571	571	281	281
Sample mean	.14	.3	.14	.3
Sample s.d.	.16	.46	.16	.46

Notes: Each column in this table originates from a separate OLS regression. Included are the observations of the No-Choice Group (random provision of articles). Controls (user controls) include all the baseline characteristics, described in the notes of Appendix Tables B1-B3. User-issue level observations are pooled together. The first observation on each user-issue combination is included in the regressions. Issue FE is included. Position (pre) ($\in [0,1]$) is the reported position on the issue at the baseline. Position (post) ($\in [0,1]$) is the reported position on the issue after reading an article in the experimental period (after pre-exposure period). $\text{Extreme}_{\text{post}}$ is a dummy variable indicating whether the position ($\in [0,1]$) equals to 0 or 1 after reading an article in the experimental period (after pre-exposure period). $\text{Extreme}_{\text{pre}}$ is an analogous dummy variable for the baseline position. *Trust in new source* is self-reported trust ($\in [-1,1]$) in news sources. This was measured with pop-up surveys given to users at the beginning of the day with a random chance (each time, the user was asked about trust in three randomly selected news sources). *Exogenous* means trust is measured before the first exposure to the news source. *All* means the measure includes all trust measures regardless of the timing to maximize the statistical power. Standard errors (in parentheses below coefficients) are clustered by user. *** p<0.01, ** p<0.05, * p<0.1.

Table B14: Article quality is balanced between Source-Name Group and No-Choice Group

	User × issue combination w/ at least one article read after pre-exposure period (Column 2: # evaluations≥5)	
	Residualized article quality (1)	Residualized article quality (2)
Source-Name Grp.	-.037	.01
	(.031)	(.026)
Source-Position Grp.	-.1**	-.08*
	(.043)	(.043)
Const. (Omit'd: No-Choice Grp.)	-.00057	-.033
	(.07)	(.067)
p-value: Source-Name = Source-Position	.17	.049
User controls, issue FE	Y	Y
Number of users	332	302
Obs. (user × issue)	1456	872
Sample mean	.0051	-.051
Sample s.d.	.53	.37

Notes: Each column in this table originates from a separate OLS regression. The omitted group is the No-Choice Group (random provision of articles). The Source-Name Group is allowed to select the news source with the source names shown. The Source-Position Group is allowed to select the news source with only the source positions shown. Controls (user controls) include all the baseline characteristics, described in the notes of Appendix Tables B1-B3. User-issue level observations are pooled together. The first observation on each user-issue combination is included in the regressions. Issue FE is included. Residualized article quality is calculated from the regression of user-reported article quality on (i) the distance between user's supporting party's position and the article's position (ii) the distance between user's prior position and the article's position, and (iii) issue FEs. Included are user × issue level observations, taking only the first observation after the pre-exposure period. In column 2, articles with less than 5 user evaluations on quality are excluded. Standard errors (in parentheses below coefficients) are clustered by user. *** p<0.01, ** p<0.05, * p<0.1.

Table B15: Source-Name group wasn't more likely to read more moderate articles

	User × issue combination w/ at least one article read after pre-exposure period	
	Article's relative extremism $= sign(Position(pre) - 0.5) \times (ArticlePosition - position(pre))$ (1)	Article's absolute extremism $\|ArticlePosition - 0.5\|$ (2)
Source-Name Grp.	-.0085	-.0039
	(.015)	(.0058)
Source-Position Grp.	-.026	.012*
	(.022)	(.0065)
Const. (Omit'd: No-Choice Grp.)	-.18***	.26***
	(.036)	(.013)
p-value: Source-Name = Source-Position	.46	.038
User controls, issue FE	Y	Y
Number of users	336	336
Obs. (user × issue)	1757	1757
Sample mean	-.29	.18
Sample s.d.	.25	.11

Notes: Each column in this table originates from a separate OLS regression. The omitted group is the No-Choice Group (random provision of articles). The Source-Name Group is allowed to select the news source with the source names shown. The Source-Position Group is allowed to select the news source with only the source positions shown. Article position is the average position of the article reported by all the users after reading it. Position (pre) ($\in [0,1]$) is the reported position on the issue at the baseline. Controls (user controls) include all the baseline characteristics, described in the notes of Appendix Tables B1-B3. User-issue level observations are pooled together. The first observation on each user-issue combination is included in the regressions. Issue FE is included. Standard errors (in parentheses below coefficients) are clustered by user. *** p<0.01, ** p<0.05, * p<0.1.

Table B16: Source-Name and No-Choice Group spent similar amount of time reading articles

	Time spent < 10 minutes		Full sample		
	Linear regression	Quantile regression (median)	Quantile regression (median)	Linear Probability Model	Linear Probability Model
	(1)	(2)	(3)	(4)	(5)
	Time spent (minutes)	Time spent (minutes)	Time spent (minutes)	Time spent> 10 minutes	Time spent> 1 hour
Source-Name Grp.	.23	.29	.33	.046	.022
	(.18)	(.19)	(.23)	(.03)	(.023)
Source-Position Grp.	-.0031	.17	-.028	-.026	-.036
	(.22)	(.2)	(.27)	(.032)	(.027)
Const. (Omit'd: No-Choice Grp.)	3.7***	2.8***	3.6***	.3***	.25***
	(.4)	(.41)	(.55)	(.06)	(.052)
p-value: Source-Name = Source-Position	.35	.61	.27	.063	.058
User controls, issue FE	Y	Y	Y	Y	Y
Number of users	326	326	336	336	336
Obs. (user × issue)	1488	1488	1774	1774	1774
Sample mean	2.4	2.4	1218	.16	.12
Sample s.d.	1.8	1.8	9161	.37	.32

Notes: Each column in this table originates from a separate linear or quantile regression. The specific regression that I run is indicated above the column number. Time spent (minutes) is the total amount of time that the user spent (i) selecting the source, (ii) reading the article, and (iii) reporting the relevant policy view and evaluating the article. If the user spent more than 10 minutes on an article, it is likely that the app was inactive and the user digressed. I cannot detect and measure such inactivity due to technological limitations. The omitted group is the No-Choice Group (random provision of articles). The Source-Name Group is allowed to select the news source with the source names shown. The Source-Position Group is allowed to select the news source with only the source positions shown. Controls (user controls) include all the baseline characteristics, described in the notes of Appendix Tables B1-B3. User-issue level observations are pooled together. The first observation on each user-issue combination is included in the regressions. Issue FE is included. Standard errors (in parentheses below coefficients) are clustered by user. *** p<0.01, ** p<0.05, * p<0.1.

Table B17: There was no overall movement toward right-wing positions

	User × issue combination w/ at least one article read after pre-exposure period		
	Position (post) (1)	Position (post) (2)	Position (post) (3)
Source-Name Grp.	.027	.0067	.0062
	(.024)	(.012)	(.012)
Source-Position Grp.	-.043	.002	.0007
	(.026)	(.017)	(.017)
Const. (Omit'd: No-Choice Grp.)	.26***	.22***	.17***
	(.012)	(.03)	(.03)
p-value: Source-Name = Source-Position	.027	.79	.75
User controls	N	Y	Y
Issue FE	N	N	Y
Control for Position (pre)	Y	Y	Y
Number of users		336	
Obs. (user × issue)		1774	
Sample mean		.26	
Sample s.d.		.25	

Notes: Each column in this table originates from a separate OLS regression of the absolute value of change in position on the treatment group dummies. The omitted group is the No-Choice Group (random provision of articles). The Source-Name Group is allowed to select the news source with the source names shown. The Source-Position Group is allowed to select the news source with only the source positions shown. Controls (user controls) include all the baseline characteristics, described in the notes of Appendix Tables B1-B3. User-issue level observations are pooled together. The first observation on each user-issue combination is included in the regressions. Issue FE is included in some columns as indicated. Position (pre) ($\in [0,1]$) is the reported position on the issue at the baseline. Position (post) ($\in [0,1]$) is the reported position on the issue after reading an article in the experimental period (after pre-exposure period). Standard errors (in parentheses below coefficients) are clustered by user. *** p<0.01, ** p<0.05, * p<0.1.

Table B18: No evidence on convergence to party's position

	Basic	First difference (FD)	Lagged dependent var. (LDV)	Baseline balance
	ǀ Position (post)- party position ǀ	Δ ǀ User position- party position ǀ	ǀ Position (post)- party position ǀ	ǀ Position (pre)- party position ǀ
	(1)	(2)	(3)	(4)
Source-Name Grp.	-.000093	-.015	-.0055	.015
	(.011)	(.0099)	(.0096)	(.01)
Source-Position Grp.	-.0058	-.0049	-.0054	-.00088
	(.012)	(.013)	(.011)	(.012)
ǀ Position(pre) - party position ǀ			.36***	
			(.034)	
Const. (Omit'd: No-Choice Grp.)	.17***	.017	.12***	.15***
	(.026)	(.024)	(.022)	(.027)
p-value: Source-Name = Source-Position	.7	.46	1	.25
User controls, issue FE	Y	Y	Y	Y
Number of users	316	316	316	316
Obs. (user × issue)	1668	1668	1668	1668
Sample mean	.16	-.016	.16	.18
Sample s.d.	.14	.16	.14	.16

Notes: Each column in this table originates from a separate OLS regression. The omitted group is the No-Choice Group (random provision of articles). The Source-Name Group is allowed to select the news source with the source names shown. The Source-Position Group is allowed to select the news source with only the source positions shown. Controls (user controls) include all the baseline characteristics, described in the notes of Appendix Tables B1-B3. User-issue level observations are pooled together. The first observation on each user-issue combination is included in the regressions. Issue FE is included. Position (pre) ($\in [0, 1]$) is the reported position on the issue at the baseline. Position (post) ($\in [0, 1]$) is the reported position on the issue after reading an article in the experimental period (after pre-exposure period). User's party position is the average baseline position of the users who share the same most-trusted party at the baseline and have baseline confidence ([0,1]) of 0.8 or more. Standard errors (in parentheses below coefficients) are clustered by user. *** p<0.01, ** p<0.05, * p<0.1.

Table B19: No convergence to party's position (alternative definitions of party position)

	User × issue combination w/ at least one article read after pre-exposure period: Lagged dependent var. (LDV) model		
	| Position (post)- party position |	| Position (post)- party position (alt1) |	| Position (post)- party position (alt2) |
	(1)	(2)	(3)
Source-Name Grp.	-.0055	-.0075	-.0087
	(.0096)	(.0097)	(.009)
Source-Position Grp.	-.0054	-.0062	-.011
	(.011)	(.011)	(.01)
| Position(pre) - party position |	.36***		
	(.034)		
| Position(pre) - party position (alt1) |		.36***	
		(.034)	
| Position(pre) - party position (alt2) |			.33***
			(.033)
Const. (Omit'd: No-Choice Grp.)	.12***	.1***	.08***
	(.022)	(.022)	(.02)
p-value: Source-Name = Source-Position	1	.92	.85
User controls, issue FE	Y	Y	Y
Number of users	316	316	316
Obs. (user × issue)	1668	1668	1668
Sample mean	.16	.16	.17
Sample s.d.	.14	.14	.13

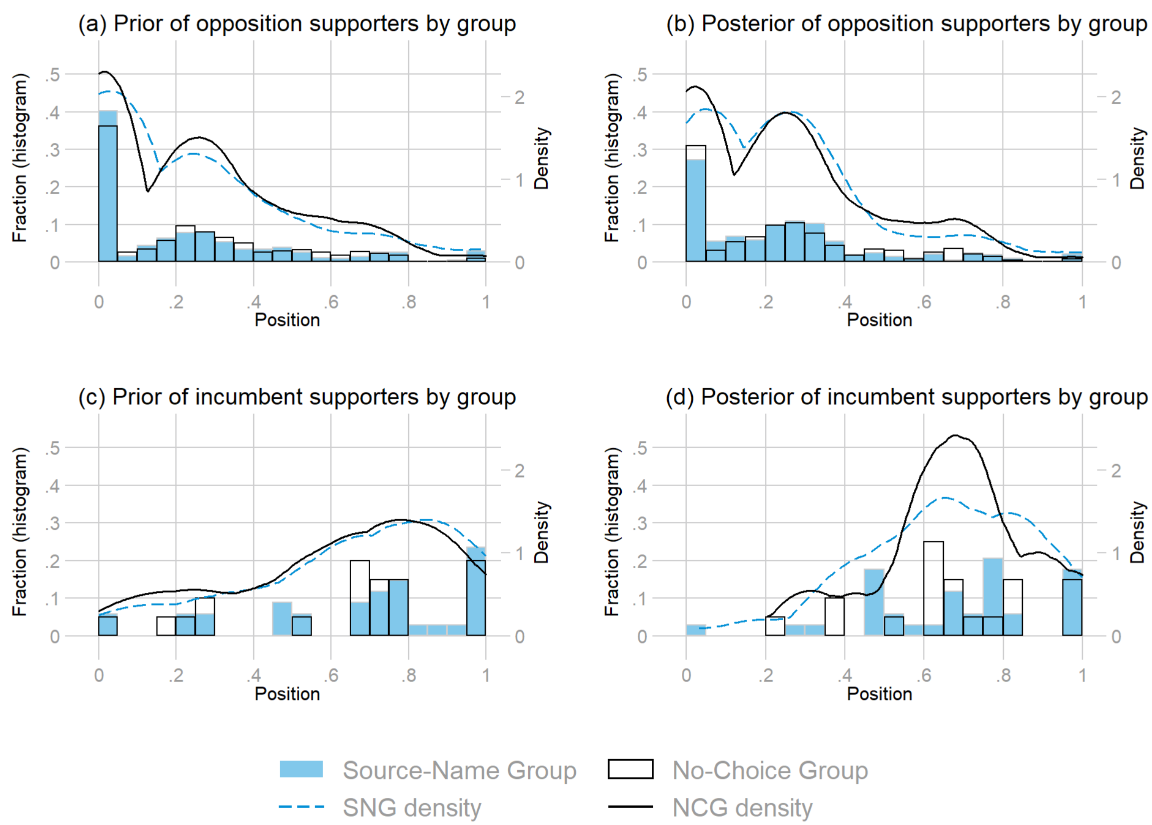
Notes: Each column in this table originates from a separate OLS regression. The omitted group is the No-Choice Group (random provision of articles). The Source-Name Group is allowed to select the news source with the source names shown. The Source-Position Group is allowed to select the news source with only the source positions shown. Controls (user controls) include all the baseline characteristics, described in the notes of Appendix Tables B1-B3. User-issue level observations are pooled together. The first observation on each user-issue combination is included in the regressions. Issue FE is included. Position (pre) ($\in [0,1]$) is the reported position on the issue at the baseline. Position (post) ($\in [0,1]$) is the reported position on the issue after reading an article in the experimental period (after pre-exposure period). User's party position is the average baseline position of the users who share the same most-trusted party at the baseline and have baseline confidence ([0,1]) of 0.8 or more. Party position (alt1) uses party trust of larger than 0.5 (median trust) as the criterion instead of confidence level. Party position (alt2) uses everyone with the same most-trusted party at the baseline. Standard errors (in parentheses below coefficients) are clustered by user. *** p<0.01, ** p<0.05, * p<0.1.

Table B20: 1990s counterfactual—similar results on moderation

	User × issue combination w/ at least one article read after pre-exposure period Source-Name and No-Choice only		Among No-Choice Group, including only those exposed to famous news sources	
	∣ Position (post)- position (pre) ∣ (1)	Extreme_{post} (2)	∣ Position (post)- position (pre) ∣ (3)	Extreme_{post} (4)
Source-Name Grp.	.021**	-.065**	.013	-.069**
	(.01)	(.028)	(.012)	(.032)
Extreme_{pre}		.34***		.3***
		(.03)		(.036)
Const. (Omit'd: No-Choice Grp.)	.14***	.17**	.13***	.13
	(.026)	(.066)	(.034)	(.082)
User controls, issue FE	Y	Y	Y	Y
Number of users	297	297	257	257
Obs. (user × issue)	1557	1557	896	896
Sample mean	.14	.27	.15	.26
Sample s.d.	.17	.45	.17	.44

Notes: Each column in this table originates from a separate OLS regression. The omitted group is the No-Choice Group (random provision of articles). In columns 3 and 4, users in the No-Choice Group who were randomly exposed to famous news sources (five historically famous newspapers and three air-wave TV channels) are included, and others are excluded in the regression. The Source-Name Group is allowed to select the news source with the source names shown. Controls (user controls) include all the baseline characteristics, described in the notes of Appendix Tables B1-B3. User-issue level observations are pooled together. The first observation on each user-issue combination is included in the regressions. Issue FE is included. Position (pre) ($\in [0, 1]$) is the reported position on the issue at the baseline. Position (post) ($\in [0, 1]$) is the reported position on the issue after reading an article in the experimental period (after pre-exposure period). Extreme_{post} is a dummy variable indicating whether the position ($\in [0, 1]$) equals to 0 or 1 after reading an article in the experimental period (after pre-exposure period). Extreme_{pre} is an analogous dummy variable for the baseline position. Standard errors (in parentheses below coefficients) are clustered by user. *** p<0.01, ** p<0.05, * p<0.1.

Figure B1: Histogram of positions by partisanship, before and after article reading



Notes: Position (pre) ($\in [0, 1]$) is the reported position on the issue at the baseline. Position (post) ($\in [0, 1]$) is the reported position on the issue after reading an article in the experimental period (after the pre-exposure period). User-issue level observations are pooled together. The No-Choice Group is given randomly selected articles. The Source-Name Group is allowed to select the news source with the source names shown. Incumbent supporters are users who indicated that they trust the *Saenuri* Party the most at the baseline. Opposition supporters are the remaining users.