# Breaking echo chambers with personalized news

Jiemai Wu[*,†]

January 18, 2022

**Abstract**

When a news platform such as Google News or Apple News selects personalized news for its user, will it select news that conforms to its user's bias, thus creating an "echo chamber"? To answer this question, this paper studies a game between a click-maximizing platform and a rational user who tries to learn the true state of the world. In equilibrium, driven by user demand, the platform recommends news that contradicts the user's bias. This result stands in contrast with theories of media bias in the literature and is consistent with recent empirical findings.

Keywords: media bias, personalized news, echo chambers, news aggregator, platform

JEL classifications: C72, D83, L82

## 1 Introduction

Advancements in data analytics have enabled news platforms such as Google News or Apple News to learn about their users' individual biases and deliver personalized news based on this information. To maximize clicks, should these platforms provide personalized news that challenges our existing view of the world, or news that panders to it? If the latter were true, personalized news feeds could increase the segregation and polarization of people's views, which may eventually lead to misunderstanding and conflict between different groups. This problem is often referred to as a media "echo chamber" (Sunstein, 2001a,b, 2007) or a digital "filter bubble" (Pariser, 2011), both of which refer to the possibility that platforms feed its users only news that echoes their existing views and filter out opposing stories. Echo

[*]School of Economics, University of Sydney. Email: jiemai.wu@sydney.edu.au

[†]I deeply thank Junmin Liao, John Nachbar, Murali Agastya, Tilman Börgers, Andrew McLennan, Claudio Mezzetti, one anonymous editor and three anonymous referees for their valuable comments.

chambers and filter bubbles have been a matter of public concern and have been widely discussed in major news outlets globally. [1]

However, the fact that digital platforms can deliver news that panders to a user's bias does not necessarily mean that it is in their best interest to do so. This paper shows that, in order for a platform to maximize clicks, the optimal strategy is the exact opposite - recommending news that *contradicts* a user's bias. This paper novelly models a user's two-step click-and-read news consumption process and interprets a news title as indicating a truncation of the potential news articles. Under this model, a user prefers to click on a title that contradicts her prior bias. This result reversed the predictions of previous media-bias theories and is consistent with recent empirical findings.

For intuition, consider the following motivating example. I also compare popular alternative theories of media bias within the context of the same example later on.

Suppose that numerous news sources on the internet are divided into two types with equal probability: left-biased and right-biased. They publish binary articles that either praise the left-leaning party or the right-leaning party. When the left-leaning party is more competent, a left-biased news source praises it with probability 1; a right-biased source praises it with probability 1/4 (and praises the less-competent right-leaning party with probability 3/4). When the right-leaning party is more competent, a left-biased news source praises it with probability 1/4, and a right-biased source praises it with probability 1. This implies that *articles vary in their informativeness*, e.g., a left-praising article from a left-biased source is less informative than a left-praising article from a right-biased source.

A platform chooses whether to present a left-praising or right-praising title to a user who is trying to learn the relative competency of the parties, so that she can cast the correct vote. The user must pay a (time or monetary) cost to click on the title and read the full article in order to find out its exact source and informativeness. The user prefers to vote for whichever party that is more likely to have a higher level of competence. She gets a payoff of 1 if she votes correctly and 0 otherwise. Suppose that the user has a left-biased prior belief: she thinks the left-leaning party is more competent with probability 0.8. Does such a left-biased user prefer to click on a left-praising or a right-praising title on the platform?

If she clicks on *a left-praising title that echos her bias*, she expects to either

(a) see a left-praising article from the left-biased source with probability 0.83 (rounded to two decimal places), in which case she believes the left-leaning party is more competent with probability 0.76; she votes for the left-leaning party and her expected utility is 0.76;

or,

[1] They include, but are not limited to, BBC, the Guardian, the New York Times, Forbes, Washington Post, and Financial Times.

(b) see a left-praising article from the right-biased source with probability 0.17, in which case she believes the left-leaning party is more competent with probability 1; she votes for the left-leaning party and her expected utility is 1.

Note that the user votes for the left-leaning party in both cases (a) and (b). Moreover, she would have voted for the this party under her prior belief even if she did not click on the left-praising title. Therefore, a click on a left-praising title has no impact on her action, and her ex-ante expected utility from this click is 0.8, the same as her expected utility if she does not click.

If she clicks on *a right-praising title that contradicts her bias*, she expects to either

(c) see a right-praising article from the left-biased source with probability 0.06, in which case she believes the left-leaning party is more competent with probability 0; she votes for the right-leaning party and her expected utility is 1;

or,

(d) see a right-praising article from the right-biased source with probability 0.94, in which case she believes the left-leaning party is more competent with probability 0.83; she votes for the left-leaning party and her expected utility is 0.83.

The user's ex-ante expected utility from a click on the right-praising title is 0.84, which is higher than her expected utility if she does not click. Because she strictly prefers to switch from her default left vote to a right vote in case (c), the information from a click on the right-praising title is strictly beneficial to her in expectation.

Therefore, in this example, the left-biased user is strictly more willing to click on a right-praising title. Consequently, the platform chooses to recommend her a right-praising, prior-contradicting title in order to maximize the chance of a click.

While the motivating example is extremely stylized, its prediction is robust. In mathematical terms, the title represents a binary (left or right, upward or downward) truncation of the news article distribution, and the user must pay a cost if she wishes to obtain a realization from this truncated distribution. This paper studies a general model in which the news articles follow an arbitrary distribution with monotonic likelihood ratio. The variance of the articles' informativeness may come from any cause, including the biases of news sources (as described in the example above) or varying volumes of hard evidence. The user can have a biased prior belief (as described in the example), a bias towards an action (e.g., she prefers to vote for the left-leaning party even when it is probably less competent), or a combination of the two. If she has a biased prior belief, that belief can be arbitrarily moderate so that she may vote in either direction with zero or positive probability regardless of the title that she clicks on.

This paper identifies the conditions for the equilibrium title to contradict the user's prior bias: (1) The most informative left-praising articles are better at inducing a left vote than the least informative right-praising articles. (2) The user is not indifferent between the two types of titles except when she is indifferent between the left and right votes at her prior. When these two relatively weak assumptions are satisfied, the user exhibits a preference for the title that contradicts her prior bias. Driven by user demand, the platform presents a prior-contradicting title in equilibrium.

This result is non-trivial because it makes the opposite prediction compared to theories based on the traditional models of media where a consumer chooses a news source and consumes all of its current and future news (e.g., subscriptions to newspapers, TV channels, and repeated direct visits to websites of news sources such as foxnews.com and nytimes.com). To make this point clear, consider two modifications of the motivating example that yield the opposite result:

*Alternative Model A.* Suppose that, instead of maximizing expected utility, the user seeks to subscribe to a platform *with the highest probability of revealing the true state*. By comparing cases (b) and (c) above, one can see that the user expects to learn the true state with a higher probability when the platform publishes a left-praising title that *echos* her bias. Therefore, she is more willing to subscribe to a platform that echos her bias if accuracy is what she is after, and the platform publishes prior-conforming titles in equilibrium. The highly-cited paper of Gentzkow and Shapiro (2006) built a model of media bias and reputation with this intuition. It showed that news outlets slant their reports toward the prior beliefs of their customers in order to build a reputation for reporting the true state and attract future subscriptions. Gentzkow et al. (2014) and Gentzkow and Shapiro (2010) documented this media slant in US newspaper markets in 1924 and early 2000s. Beyond Gentzkow and Shapiro (2006), the fact that a Bayesian user expects a prior-conforming outlet to generate accurate signals with a higher probability is observed in many settings, including this paper.[2] Nevertheless, the user in this paper still prefers the prior-contradicting title because she expects its associated article to be more influential on her action, even though it likely contains more noise.[3]

*Alternative Model B.* Suppose that the user seeks to *subscribe to a left-biased or right-biased news source* based on only her prior belief (i.e., before she sees any title or article). She compares her expected utility from the news sources rather than the anonymous titles. Once she pays, she learns whether the article from the chosen source is left-praising or right-praising. Let all the other assumptions (probabilities of articles conditional on news

[2]This is formally addressed by Remark 1.

[3]For example, a health-conscious user is unlikely to click on a title that says "exercise is good for your health" even though she believes it to be true.

sources, user's prior and decision rule) remain the same. In this setting, suppose that the left-biased user subscribes to a left-biased source. With probability 0.95, she receives a left-praising article, in which case she believes the left-leaning party is more competent with probability 0.842 and votes left. With probability 0.05, she receives a right-praising article, in which case she believes the right-leaning party is more competent with probability 1 and votes right. Her ex-ante expected utility from a subscription to the left-biased source is 0.85. Suppose, instead, that the left-biased user subscribes to a right-biased source. With probability 0.2, she receives a left-praising article, in which case she believes the left-leaning party is more competent with probability 1 and votes left. With probability 0.8, she receives a right-praising article, in which case she believes the left-leaning party is more competent with probability 0.75 and votes left. Her ex-ante expected utility from a subscription to the right-biased source is only 0.80. Therefore, under this model, the user prefers to subscribe to the left-biased source, the one that *echos* her prior bias. Versions of this model have been studied in the literature. The given example is an application of Suen (2004). Calvert (1985) discussed a similar problem with the same intuition. The lab experiments by Charness et al. (2021) contained the same setup; they found that a substantial percentage of subjects correctly chose the source with the same bias as theirs, and this percentage is higher for subjects with high scores in cognitive tests. Burke (2008) expanded Suen (2004) by introducing an endogenous market of news sources as well as a dynamic analysis of information acquisition and provision. Che and Mierendorff (2019) focused on a stopping problem in a dynamic version of alternative model $B$.

Both alternative models $A$ and $B$ shed light on the biases of traditional news outlets (e.g., newspapers, TV stations, and their corresponding websites). In comparison, the model in this paper better fits the free-to-view, click-to-read business model of a click-maximizing platform (e.g., online news aggregators) whose main income is from advertisements. Advertisement income increases with page visits, click-through rates, and average visit duration of users, all of which are simplified into a binary click decision in this paper. Users of such platforms are exposed to a range of news sources with varying informativeness. Because of the different business models and information environment, this paper predicts a media bias against the consumer's prior, while alternative models $A$ and $B$ predict a media bias towards the consumer's prior.

This difference is consistent with empirical data. Flaxman et al. (2016) studied online news consumption through four different channels: direct and independent visits to news domains such as nytimes.com ("Direct"), referrals from Google News ("Aggregator"), social media, and search engines. Among these four channels, Aggregator best fits this paper, while Direct fits alternative models A and B. They found that news consumption through

Aggregator exhibited the lowest segregation even though news referrals are personalized. Moreover, the percentage of opposing political news that users read was the highest for Aggregator and lowest for Direct (approximately 18% vs 2.5%). Fletcher et al. (2021) analyzed tracking data from the United Kingdom and found that the more people use distributed news access through aggregators, search engines, and social media, the more diverse their news repertoires. Furthermore, they are also more politically diverse, with news use spread across left- and right-leaning outlets. Dubois and Blank (2018) analyzed survey data and found that consumers are less likely to be in an echo chamber if they are exposed to news from more sources. It is worth noting that, while this paper does not study social media in particular, there is a large and growing empirical literature that investigates echo chambers on social media such as Facebook and Twitter. The majority sentiment is that social media encourage the formation of echo chambers more often than not (Allcott et al. (2020), Bakshy et al. (2015), Bessi et al. (2016), Levy (2021), Quattrociocchi et al. (2016)), although there is also documentation of the opposite (Guilbeaulta et al. (2018), Beam et al. (2018)). This paper does not claim to predict broad news consumption patterns on social media for two reasons: (1) subscription (to pages or accounts) is a significant feature on most social media, for which the alternative model A or B is a better fit; (2) the motives of users who post and share news on social media can be entirely different from a news platform's.

Theoretical papers that study the click-to-read model of news platforms are limited. The most related is a paper by Allon et al. (2021), who also studied a two-step process of news consumption on platforms. A user first chooses a title ("post") and then reads the article to digest its content. Among many other differences such as the state space, signal generating process, consumer choices and dynamics, the most crucial difference between this paper and theirs lies in the interpretation of titles and articles. In their paper, a title is the realization of a random variable that is correlated with the true state, but the variance of this random variable is unknown. By reading the article, the user learns the variance of the title. In my paper, fixing the user's prior belief, a title is a truncation of the news distribution with a known variance. By reading the article, the user obtains a realization from the truncated distribution. Because of this difference, the two models yield the opposite results. In their paper, the user optimally seeks a title with the smallest variance according to her own belief. At a high level, this accuracy-seeking motive is similar to that in Gentzkow and Shapiro (2006), so the prediction is also similar: the user chooses the title whose value is the closest to her current belief. In this paper, the user seeks a title whose realizations are the most influential on her action according to her own belief, and it is the one with the prior-contradicting title. Kranton and McAdams (2021) studied news producers whose revenues come from clicks on social media and showed that news veracity is endogenously determined by user networks. However, in their paper, news is modeled as either completely false or completely true, and biases of news are not discussed.

In sum, this paper is novel in both its modeling design and its result, and it provides a rational justification for the higher consumption level of opposing news and lower segregation level on news platforms compared to other types of media. The contrast between this paper's result and those under alternative models A and B suggests that one should not adopt a uniform approach when estimating (anti-)echo chamber effects across all channels of news consumption. The anti-echo chamber equilibrium outcome in this paper is also relevant for policy debates about regulation of news platforms.

In the remainder of this paper, Section 2 sets up the model, Section 3 presents the results, and Section 4 addresses several alternative scenarios, including when a platform can condition its recommendation on the true state, or does not have the personalization technology; when the user gets entertainment value from news or has a confirmation bias.

# 2 Model setup

### News articles and their titles

There are two possible states of the world, $\omega \in \{\omega_L, \omega_H\}$. In each state, there is an exogenous distribution of news articles. Each article contains an i.i.d. signal $s \in S \subseteq \mathbb{R}$ about the true state of the world. The support $S$ can be finite or infinite, but $\{s \in S \mid s < 0\}$ and $\{s \in S \mid s > 0\}$ must not be singletons so that the sign of $s$ does not fully reveal its value. Let $f_{\omega_L}(s)$ and $f_{\omega_H}(s)$ denote the conditional probability density function of $s$ in states $\omega_L$ and $\omega_H$, respectively. Let $F_{\omega_L}(s)$ and $F_{\omega_H}(s)$ denote the cumulative distribution functions. This paper makes three assumptions about these distributions:

(1) $f_{\omega_L}(s) = f_{\omega_H}(-s)$ for all $s \in S$. This assumption of symmetry ensures that the underlying information environment is the same in both states; any asymmetry in the user's posterior belief distribution is a result of only her prior belief and the platform's equilibrium strategy;

(2) $f_{\omega_H}(s) / f_{\omega_L}(s)$ is strictly increasing in $s$, so that a higher $s$ is more indicative of state $\omega_H$ and a lower $s$ is more indicative of state $\omega_L$;

(3) $\Pr(s = 0 \mid \omega_L) = \Pr(s = 0 \mid \omega_H) = 0$ (i.e., neither $f_{\omega_L}$ nor $f_{\omega_H}$ has an atom at 0). This assumption is created for technical convenience (see footnote 4), but it also implies that all articles are at least somewhat informative with probability 1.

Each article has a binary *title* that reveals the sign of $s$. An article has title $t_L$ if $s \leq 0$

and $t_H$ if $s \geq 0$. [4]

**The user**

Let $p_0 \in [0,1]$ denote the user's prior belief of $\Pr(\omega_H)$. She makes two decisions sequentially.

Firstly, the user observes a title recommended by the platform and must decide whether to click on it at the cost $c > 0$ to learn the more refined signal $s$ behind the title. $c$ can be a number or a random variable with an exogenous distribution on $(0, \infty)$. One can interpret $c$ as the time spent on reading, a monetary cost to access an article, or a combination of the two. If the user clicks on the recommended title, she Bayesian updates her belief based on the title and the revealed $s$. If she decides not to click, then her posterior belief is simply equal to her prior belief.[5] Without loss of generality, assume that she clicks when she is indifferent.

Secondly, the user chooses a binary action, either $a_L$ or $a_H$, that best matches her posterior belief about the true state. Without loss of generality, assume that she chooses $a_H$ when she is indifferent.

The user's payoff depends on her click decision, action, and the true state. Her payoff function, stated below, is separable in the action-related utility and the click cost:

$$U - c \cdot \mathcal{I}_c$$

where $\mathcal{I}_c = 1$ if she clicks and 0 if she does not. The *action-related* utility $U$ is described by the following table:

| | $\omega_L$ | $\omega_H$ |
|---|---|---|
| choose $a_L$ | $u$ | 0 |
| choose $a_H$ | 0 | 1 |

where the exogenous parameter $u \in (0, \infty)$ represents the user's biased taste towards action $a_L$ relative to $a_H$. This payoff table implies that the user chooses action $a_H$ if and only if her posterior belief satisfies $\Pr(\omega_H) \geq \frac{u}{u+1}$ or, equivalently, $\frac{\Pr(\omega_H)}{\Pr(\omega_L)} \geq u$.

[4] When $s = 0$, the article has both $t_L$ and $t_H$ titles by assumption. The results in this paper are unaffected by this assumption because $\Pr(s = 0) = 0$. If one had allowed $\Pr(s = 0)$ to be strictly positive, in order to maintain symmetry, one would need to assume that an article with $s = 0$ has a random title $t = t_L$ or $t_H$ with equal probability. This ad hoc assumption would complicate calculations without adding new insights.

[5] In the main model, the user does not update her belief based on the revealed title $t$ alone. Because the platform does not know the true state, its title recommendation does not reveal information about the state. This assumption is relaxed in Section 4.1.

### The platform

The platform's strategy is to design an algorithm that recommends either a positive or negative title to the user conditional on the user's prior belief $p_0$ and her payoff parameters $u$, $c$.

Specifically, the platform chooses the probability that it recommends $t_H$ or $t_L$ given each realization of $(p_0, u, c)$. If it decides to recommend a title, it takes a random draw from the distribution of articles with that title. The user initially sees only the title $t$; she must click on it in order to learn the refined signal $s$.

The platform's only objective is to maximize the probability of a click. Its payoff is equal to 1 if the user chooses to click and 0 otherwise.

### Timeline

The following list summarizes the timeline of the game. $p_0$ and $p_2$ denote the user's prior and posterior beliefs at different points of the timeline.[6]

1. The platform chooses a recommendation algorithm, and this algorithm is observed by the user.

2. Nature reveals the true state $\omega$, the user's prior belief $p_0$, and the user's payoff parameters $(u, c)$.

3. Based on the algorithm, the platform recommends a title $t$ to the user.

4. The user decides whether to click on it to learn $s$.

5. The user updates her posterior belief $p_2$ conditional on $s$ (or the absence of this information) and chooses an action $a$.

6. The platform and the user receive their payoffs.

# 3 Equilibrium

## 3.1 User's preference

The first step to finding the equilibrium is to derive the user's preference over titles. In this subsection, I study when the user prefers the prior-contradicting title, and how this preference depends on the article distributions. The answer is given by Theorem 1. This

[6]Notation "$p_1$" is reserved to denote the user's interim belief in the extension in Section 4.1.

subsection is the centerpiece of the paper. As later shown in Section 3.2 (platform's optimal algorithm), the equilibrium outcome is demand-driven and crucially shaped by the user's preference.

**Definition 1.** Given the user's prior belief $p_0$,

for $i = L, H$, let $U_i$ denote the user's action-related expected utility if the recommended title is $t_i$ and she intends to click on it;

let $U_0$ denote the user's action-related expected utility if she does not click on any title.

The rest of this section explores the properties and ranking of $U_L$, $U_H$, $U_0$, as well as the comparison between $U_i - c$ (total payoff from a click) and $U_0$ (total payoff without a click). One can express the $U$'s more explicitly with a bit of work, as shown below.

Suppose that $t = t_L$. The user learns a signal $s \leq 0$ if she clicks on it. Each $s \in (-\infty, 0] \cup S$ is drawn with probability or density

$$q_{t_L}(s) = \frac{f_{\omega_H}(s)}{F_{\omega_H}(0)} p_0 + \frac{f_{\omega_L}(s)}{F_{\omega_L}(0)} (1 - p_0)$$

After learning $s \leq 0$, the user's posterior belief for state $\omega_H$ becomes

$$\begin{aligned} p_2(s \mid t_L) &= \frac{f_{\omega_H}(s \mid s \leq 0) p_0}{f_{\omega_H}(s \mid s \leq 0) p_0 + f_{\omega_L}(s \mid s \leq 0)(1 - p_0)} \\ &= \frac{\frac{f_{\omega_H}(s)}{F_{\omega_H}(0)} p_0}{\frac{f_{\omega_H}(s)}{F_{\omega_H}(0)} p_0 + \frac{f_{\omega_L}(s)}{F_{\omega_L}(0)} (1 - p_0)} \end{aligned}$$

Given $p_0$, the set of feasible negative $s$ can be partitioned into two subsets.

**Definition 2.** For $i = L, H$, when the title is $t_i$, let the feasible $s$ be partitioned into two subsets $S_1^i$ and $S_2^i$ such that the user chooses $a_L$ if $s \in S_1^i$ and $a_H$ if $s \in S_2^i$, i.e.,

$$S_1^L \equiv S \cap \left\{ s \leq 0 \mid p_2(s \mid t_L) < \frac{u}{u+1} \right\} \text{ and } S_2^L \equiv S \cap \left\{ s \leq 0 \mid p_2(s \mid t_L) \geq \frac{u}{u+1} \right\}.$$

$$S_1^H \equiv S \cap \left\{ s \geq 0 \mid p_2(s \mid t_H) < \frac{u}{u+1} \right\} \text{ and } S_2^H \equiv S \cap \left\{ s \geq 0 \mid p_2(s \mid t_H) \geq \frac{u}{u+1} \right\}.$$

This partition changes with $p_0$. Because $p_2$ is increasing in $s$, all of the elements in $S_1^i$ are smaller than all of the elements in $S_2^i$.

The user's expected utility after clicking on $t_L$ is

$$
\begin{aligned}
U_L &= \sum_{s \in S_1^L} q_{t_L}(s)\, u\left[1 - p_2(s \mid t_L)\right] + \sum_{s \in S_2^L} q_{t_L}(s)\, p_2(s \mid t_L) \\
&= \sum_{s \in S_1^L} u\left[\frac{f_{\omega_L}(s)}{F_{\omega_L}(0)}(1 - p_0)\right] + \sum_{s \in S_2^L} \frac{f_{\omega_H}(s)}{F_{\omega_H}(0)} p_0
\end{aligned}
$$

If $f_{\omega_L}$ and $f_{\omega_H}$ are continuous then we can replace the sums with integrals. If $0 \notin S$, then replace $F_{\omega_L}(0)$ with $\Pr(s < 0 \mid \omega_L)$ and $F_{\omega_H}(0)$ with $\Pr(s < 0 \mid \omega_H)$; the same applies whenever $F_{\omega_L}(0)$ and $F_{\omega_H}(0)$ show up in the rest of the paper.

Similarly, if $t = t_H$, then each $s \in [0, \infty) \cup S$ is drawn with probability or density

$$
q_{t_H}(s) = \frac{f_{\omega_H}(s)}{1 - F_{\omega_H}(0)} p_0 + \frac{f_{\omega_L}(s)}{1 - F_{\omega_L}(0)}(1 - p_0)
$$

After learning $s \geq 0$, the user's posterior belief for state $\omega_H$ becomes

$$
p_2(s \mid t_H) = \frac{\frac{f_{\omega_H}(s)}{1 - F_{\omega_H}(0)} p_0}{\frac{f_{\omega_H}(s)}{1 - F_{\omega_H}(0)} p_0 + \frac{f_{\omega_L}(s)}{1 - F_{\omega_L}(0)}(1 - p_0)}
$$

The user's expected utility after clicking on $t_H$ is

$$
\begin{aligned}
U_H &= \sum_{s \in S_1^H} q_{t_H}(s)\, u\left[1 - p_2(s \mid t_H)\right] + \sum_{s \in S_2^H} q_{t_H}(s)\, p_2(s \mid t_H) \\
&= \sum_{s \in S_1^H} u\left[\frac{f_{\omega_L}(s)}{1 - F_{\omega_L}(0)}(1 - p_0)\right] + \sum_{s \in S_2^H} \left[\frac{f_{\omega_H}(s)}{1 - F_{\omega_H}(0)} p_0\right]
\end{aligned}
$$

For any discrete or continuous distribution $f$, $U_L$ and $U_H$ are: (a) continuous, and (b) equal if the user is indifferent between $a_L$ and $a_H$ at her prior belief (i.e., if $p_0 = \frac{u}{u+1}$). Lemmas 3 and 4 formally state these results in the Appendix. Proofs are also deferred to the Appendix.

The goal of the remainder of this section is to identify conditions under which the user prefers a title that contradicts her prior bias. To gain intuition, Example 1 formally restates the motivating example from the Introduction using the framework of the model. It plots the example's underlying probability distributions as well as the utility functions. Observations from these graphs motivate Assumption 1 and 2 (to be introduced after the example), which generate preferences for titles that contradict the user's prior bias.

**Example 1.** Motivating example revisited

The motivating example in the Introduction can be modeled with a discrete distribution that has four possible realizations:

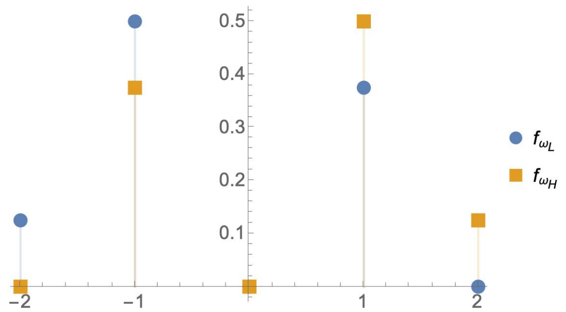
$s = -2$ represents a left-praising article from a right-leaning source;

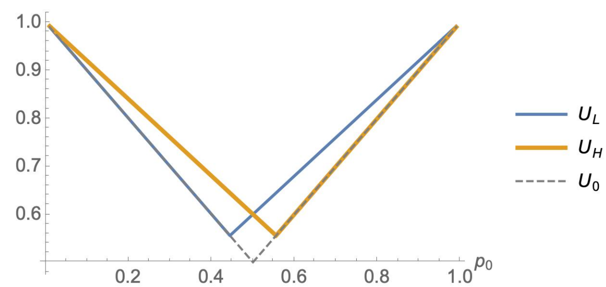
$s = -1$ represents a left-praising article from a left-leaning source;

$s = 1$ represents a right-praising article from a right-leaning source;

$s = 2$ represents a right-praising article from a left-leaning source.

Let $\omega_L$ be the state in which the left-leaning party is more competent and $\omega_H$ be the state in which the right-leaning party is more competent. Then, given that the two types of sources are equally likely in both states, $f_{\omega_L}(-2) = \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8}$, $f_{\omega_L}(-1) = \frac{1}{2}$, $f_{\omega_L}(1) = \frac{3}{4} \cdot \frac{1}{2} = \frac{3}{8}$, $f_{\omega_L}(2) = 0$, and $f_{\omega_H}(s) = f_{\omega_L}(-s)$. The graph on the left plots this distribution.




The graph on the right plots $U_L$, $U_H$, and $U_0$ when $u = 1$. Notice that if the user has a left-leaning bias $\left(p_0 < \frac{1}{2}\right)$, $U_H$ is strictly higher. If she has a right-leaning bias $\left(p_0 > \frac{1}{2}\right)$, $U_L$ is strictly higher. This shows that the user always prefers the title that contradicts her prior. The same result stands if $u$, $p_0$, and the news sources' probability of misreporting take different values.

Also notice two other details in this plot:

(a) For the smallest values of $p_0$, $U_H > U_L = U_0$. For the largest values of $p_0$, $U_L > U_H = U_0$. To understand why, recall that the user's default action is $a_L$ if $p_0 < \frac{1}{2}$ and $a_H$ if $p_0 \geq \frac{1}{2}$. In this example, for the smallest values of $p_0$, $U_H > U_L = U_0$ because no $s$ behind $t_L$ can make the user switch her default action, but some $s$ behind $t_H$ can make her strictly prefer to switch to $a_H$. For the largest values of $p_0$, $U_L > U_H = U_0$ for similar reasons. Loosely speaking, the negative articles are better at inducing $a_L$ than the positive articles, and the positive articles are better at inducing $a_H$ than the negative articles.

(b) $U_L \neq U_H$ except when $p_0 = \frac{u}{u+1}$.

These two details are important to generate a preference towards prior-contradicting titles. Below, Assumption 1 (or, equivalently, Assumptions 1a and 1b) is a general statement of (a), and Assumption 2 is a weaker version of the single-crossing property in (b).

**Assumption 1.** *For all* $p_0 \in (0,1)$,

$S_1^L = \emptyset \Rightarrow S_1^H = \emptyset$,

$S_2^H = \emptyset \Rightarrow S_2^L = \emptyset$,

*and the converses are not true for some* $p_0$.

The partitions $\{S_1^L, S_2^L\}$ and $\{S_1^H, S_2^H\}$ were defined in Definition 2. Assumption 1 requires that if no negative article behind $t_L$ can ever induce $a_L$ then the same is true for the positive articles behind $t_H$; if no positive article behind $t_H$ can ever induce $a_H$ then the same is true for the negative articles behind $t_L$. In addition, for some $p_0$, the user chooses $a_L$ with positive probability after clicking on only $t_L$ but not $t_H$, and chooses $a_H$ with positive probability after clicking on only $t_H$ but not $t_L$. Note that Assumption 1 excludes the following cases: (1) only some positive article can make the user choose $a_L$, and (2) only some negative article can make the user choose $a_H$. One could argue that if Assumption 1 fails, then the "negative" or "positive" labels of $s$ do not match the articles' real impact on the user's action and should be swapped.

While Assumption 1 is a statement about the partition of $s$, it has two equivalent alternative statements about the partition of $p_0$ (Assumption 1a) and the extreme values of $s$ (Assumption 1b). These alternative versions offer different perspectives on the restriction imposed by Assumption 1, and all three versions are used interchangeably in different proofs.

Lemma 1 states that when the user's prior is too low (below threshold $q_1$) or too high (above threshold $q_2$), no value of $s$ behind title $t_i$ can make her switch her default action. These thresholds change with the title. Assumption 1a requires that the low threshold $q_1$ is lower when the title is $t_H$, meaning that positive articles are better at making a low-prior user switch to $a_H$. Assumption 1a also requires that the high threshold $q_2$ is higher when the title is $t_L$, meaning that negative articles are better at making a high-prior user switch to $a_L$.

**Lemma 1.** *For* $i = L, H$ *and* $p_0 \in (0,1)$, *there exists thresholds* $q_1^i, q_2^i \in [0,1]$ *such that* $q_1^i \leq q_2^i$ *and*

*the user with prior belief* $p_0$ *chooses* $a_H$ *with positive probability after clicking on* $t_i$ *if and only if* $p_0 \geq q_1^i$, *and she chooses* $a_H$ *with probability 1 after clicking on* $t_i$ *if and only if* $p_0 \geq q_2^i$.

**Assumption 1a.** *The thresholds defined in Lemma 1 satisfy* $q_1^H < q_1^L$ *and* $q_2^H < q_2^L$.

Because the distribution of $s$ is assumed to have a monotonic likelihood ratio, to satisfy Assumption 1a, it is both sufficient and necessary to focus on the extreme values of $s$.

Assumption 1b requires that the most negative $s$ has a stronger negative impact on the user's belief than the least positive $s$, and the most positive $s$ has a stronger positive impact on the user's belief than the least negative $s$.

**Assumption 1b.** *Let $s_0 \geq 0$ denote the smallest non-negative element in $S$ and $-s_0 \leq 0$ be the largest non-positive element in $S$. When $p_0 = \frac{1}{2}$,*

*there exists $s \leq 0$ such that $p_2(s \mid t_L) < p_2(s_0 \mid t_H)$.*

*there also exists $s \geq 0$ such that $p_2(s \mid t_H) > p_2(-s_0 \mid t_L)$.*

Lemma 2 states the equivalence of the three versions of Assumption 1.

**Lemma 2.** *Assumptions 1, 1a, and 1b are equivalent.*

The next and last assumption is a weaker version of observation (b) in Example 1. It requires a single-crossing property of $U_L$ and $U_H$ for moderate prior belief $p_0$. The condition $\max\left\{q_1^L, q_1^H\right\} < p_0 < \min\left\{q_2^L, q_2^H\right\}$ implies that $S_1^L$, $S_1^H$, $S_2^L$, $S_2^H$ are all non-empty, i.e., the user may choose either action with positive probability after clicking on either type of title.

**Assumption 2.** *For all $p_0$ such that $\max\left\{q_1^L, q_1^H\right\} < p_0 < \min\left\{q_2^L, q_2^H\right\}$, $U_L \neq U_H$ except when $p_0 = \frac{u}{u+1}$.*

Assumption 2 is a technical sufficient condition for the main result. In contrast, Assumption 1 plays a bigger role in the interpretation of the main result, and is both necessary and sufficient to generate a preference for prior-contradicting titles in almost all cases. Theorem 1 and its discussion elaborate this.

All the building blocks for the comparison between $U_L$, $U_H$, and $U_0$ are now in place. Below, Proposition 1 compares the user's action-related expected utility without a click ($U_0$) with her action-related expected utility if she clicks on $t_i$ ($U_i$). Unsurprisingly, $U_i > U_0$ when and only when the user expects the revealed $s$ to be influential on her action. The more clueless she is (a more moderate $p_0$), the bigger the gap between $U_i$ and $U_0$ because information is more valuable for a user with a weaker prior. Proposition 1 is a general result that holds regardless of whether Assumption 1 or 2 is satisfied.

**Proposition 1.** *Let $q_1^H, q_1^L, q_2^H, q_2^L$ be the thresholds defined in Lemma 1. Then, for $i = L, H$,*

*(a) $U_i = U_0$ if $p_0 \in [0, q_1^i] \cup [q_2^i, 1]$;*

*(b) $U_i > U_0$ if $p_0 \in (q_1^i, q_2^i)$;*

*(c) $U_i - U_0$ is strictly increasing on $\left(q_1^i, \frac{u}{u+1}\right)$ and strictly decreasing on $\left(\frac{u}{u+1}, q_2^i\right)$.*

Finally, Theorem 1 states that Assumptions 1 and 2 are the sufficient conditions for the user to prefer a prior-contradicting title. Assumption 1 is also a necessary condition except at the knife-edge cases when $q_1^L = q_1^H$ or $q_2^L = q_2^H$ (or, equivalently, when the least negative $s$ is exactly as informative as the most positive $s$, or the most negative $s$ is exactly as informative as the least positive $s$). When Assumptions 1 and 2 hold, if the user's default action under her prior is $a_L$ $\left(p_0 < \frac{u}{u+1}\right)$, she weakly prefers $U_H$ and strictly so if $p_0 \in \left(q_1^H, \frac{u}{u+1}\right)$. If her default action under her prior is $a_H$ $\left(p_0 \geq \frac{u}{u+1}\right)$, she weakly prefers $U_L$ and strictly so if $p_0 \in \left(\frac{u}{u+1}, q_2^L\right)$.

**Theorem 1.** *Let $q_1^H, q_1^L, q_2^H, q_2^L$ be the thresholds defined in Lemma 1.*

*When Assumptions 1 and 2 hold,* $0 \leq q_1^H < q_1^L < \frac{u}{u+1} < q_2^H < q_2^L \leq 1$ *and*

$$\begin{cases} U_H > U_L & \text{if } p_0 \in \left(q_1^H, \frac{u}{u+1}\right) \\ U_L > U_H & \text{if } p_0 \in \left(\frac{u}{u+1}, q_2^L\right) \end{cases} \qquad (\triangle)$$

*If $q_k^L \neq q_k^H$ for $k = 1, 2$, Assumption 1 is necessary for $(\triangle)$ to hold.*

By comparing $t_L$ with $t_H$, the user is essentially choosing between two truncated distributions of $s$, one with only positive values and the other with only negative values. If one takes the perspective of Assumption 1 or 1a, Theorem 1 states that the user prefers the truncated distribution that is better at inducing a switch from her default action. If one takes the perspective of Assumption 1b, Theorem 1 states that the user prefers the truncated distribution that contains realizations capable of swinging her belief to the opposite direction by the largest magnitude.

It is worth noting that, when Assumptions 1 and 2 hold, the truncated distribution preferred by the user is not necessarily associated with more accurate articles. On the contrary, from the perspective of the user, the accuracy of articles (ranked by $|s|$) behind a prior-conforming title first-order stochastically dominates articles behind a prior-contradicting title, as shown in Remark 1. This adds another layer of insight to Theorem 1: the user prefers the prior-contradicting title *even if its associated articles are less accurate in expectation.*

*Remark* 1. Given $p_0$, let $G_L$ be the expected c.d.f. of $|s|$ for $s < 0$ and $G_H$ be the expected c.d.f. of $|s|$ for $s > 0$. Then,

if $p_0 < \frac{1}{2}$, $G_H(|s|) > G_L(|s|)$ for all $|s| > 0$;
if $p_0 > \frac{1}{2}$, $G_H(|s|) < G_L(|s|)$ for all $|s| > 0$.

Theorem 1 is the most important result of this paper. All the equilibrium results are built upon it. To gain a deeper understanding of Theorem 1, Examples 2-3 give two applications of it, and Examples 4-5 give two counter examples that illustrate the consequence when Assumption 1 or 2 fails.

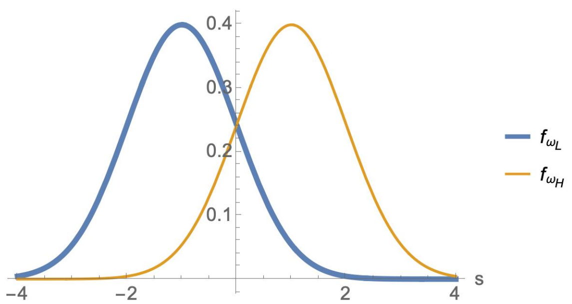
**Example 2.** Normally distributed articles

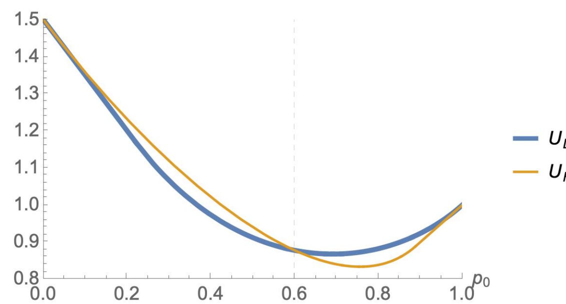
Suppose that the values of news articles follow a normal distribution:

$$s \sim \begin{cases} N(\mu, \sigma) & \text{in } \omega_H \\ N(-\mu, \sigma) & \text{in } \omega_L \end{cases}$$

for some $\mu > 0$ and $\sigma > 0$. One can interpret $|s|$ as an article's informativeness or quantity of hard evidence.

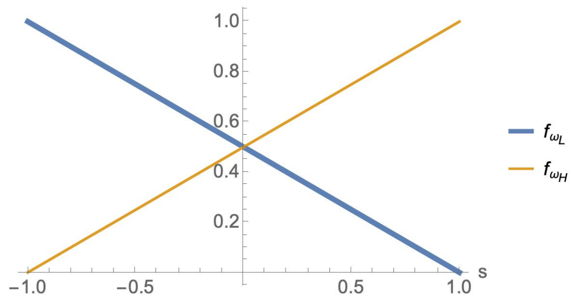
Because negative $s$ is bounded from above but not below, and positive $s$ is bounded from below but not above, Assumption 1 is satisfied because $S_1^L$ and $S_2^H$ are never empty but $S_1^H$ and $S_2^L$ are sometimes empty (in other words, $0 = q_1^H < q_1^L$ and $q_2^H < q_2^L = 1$). Assumption 2 is also satisfied: $U_L - U_H$ is strictly convex in $p_0$ when $p_0 \in \left(q_1^L, \frac{u}{u+1}\right)$ and strictly concave when $p_0 \in \left(\frac{u}{u+1}, q_2^H\right)$. Given that $U_L - U_H < 0$ on $\left(0, q_1^L\right]$ and $U_L - U_H > 0$ on $\left[q_2^H, 1\right)$, this implies that for $p_0 \in (0,1)$, $U_L = U_H$ at only $p_0 = \frac{u}{u+1}$.

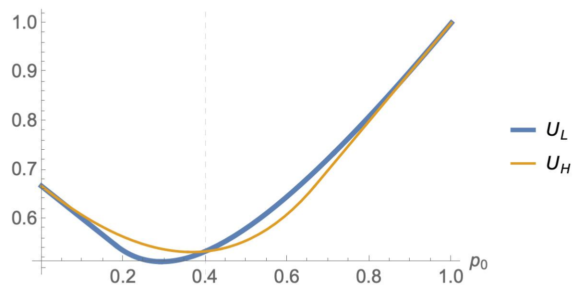



The graphs above plot the case of $u = 1.5$ (the user has a biased taste towards $a_L$), $\mu = 1$, and $\sigma = 1$. As the second graph shows, $U_H$ is higher whenever the user prefers $a_L$ under her prior belief ($p_0 < 0.6$), and $U_L$ is higher whenever the user prefers $a_H$ under her prior belief ($p_0 > 0.6$).

**Example 3.** Suppose that $f_{\omega_L}$ decreases linearly and $f_{\omega_H}$ increases linearly. Let the support of $s$ be normalized to $[-1, 1]$ and let the maximum of $f_{\omega_L}(s)$ and $f_{\omega_H}(s)$ be normalized to 1. The graphs below plot $f_{\omega_L}$, $f_{\omega_H}$, as well as the expected utility functions when $u = 2/3$ (the user has a biased taste towards $a_H$).

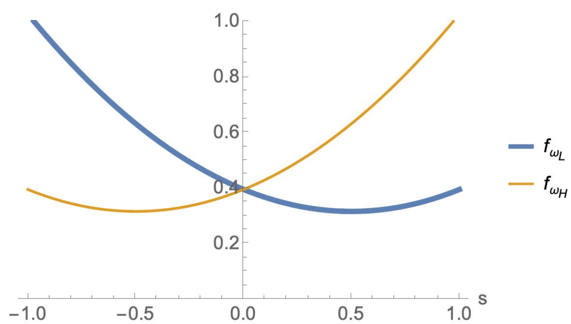


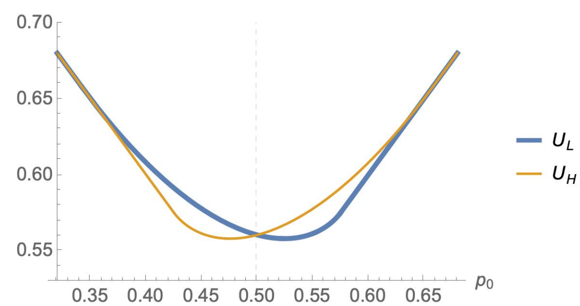



Assumption 1 is satisfied because $0 = q_1^H < q_1^L$ and $q_2^H < q_2^L = 1$. Moreover, $U_L = U_H$ has only one root $(0, 1)$ at $p_0 = \frac{u}{u+1}$ for all $u$, which satisfies Assumption 2. As shown in the graph on the right, $U_H$ is higher whenever the user prefers $a_L$ under her prior belief ($p_0 < 0.4$), and $U_L$ is higher whenever the user prefers $a_H$ under her prior belief ($p_0 > 0.4$).

**Example 4.** When Assumption 1 fails

Suppose that $S = [-1, 1]$. $f_{\omega_L}(s) = \frac{6}{19}\left[\left(s - \frac{1}{2}\right)^2 + 1\right]$ and $f_{\omega_H}(s) = \frac{6}{19}\left[\left(s + \frac{1}{2}\right)^2 + 1\right]$.

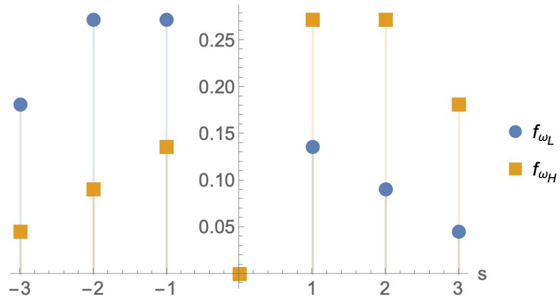


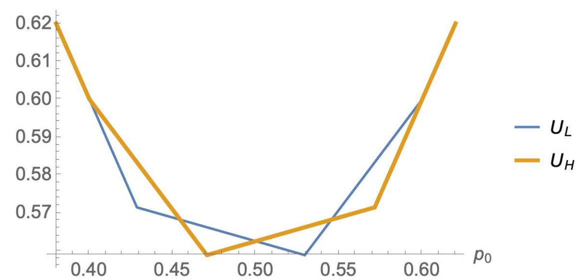



Let $u = 1$, which represents an unbiased taste for action. In this case, Assumption 1 fails and Assumption 2 holds. To see why Assumption 1 fails, observe that for a user with an unbiased prior belief $p_0 = \frac{1}{2}$, the lowest posterior belief that she can have after clicking on $t_L$ is approximately 0.425 (if $s = -1$), but the lowest posterior belief that she can have after clicking on $t_H$ is even lower, at approximately 0.342 (if $s = 0$). This violates Assumption 1b. Alternatively, one can look at the values of thresholds $q$ to find a violation of Assumption 1a: $q_1^L \approx 0.342$, $q_1^H \approx 0.425$, $q_2^L \approx 0.575$, $q_2^H \approx 0.658$. Because $q_1^L < q_1^H$ and $q_2^L < q_2^H$, a click on $t_L$ can induce $a_H$ under a wider range of prior belief and a click on $t_H$ can induce $a_L$ under a wider range of prior belief, which is a contradiction to Assumptions 1a and 1. Consequently, as shown in the graph on the right, $U_L \geq U_H$ when $p_0 < \frac{1}{2}$ and $U_L \leq U_H$ when $p_0 > \frac{1}{2}$. The user always weakly prefers a prior-conforming title, and strictly so when her prior belief is moderate.

**Example 5.** When Assumption 2 fails

Suppose that $S = \{-3,-2,-1,1,2,3\}$. $f_{\omega_L}(-3) = \frac{4}{22}$, $f_{\omega_L}(-1) = f_{\omega_L}(-2) = \frac{6}{22}$, $f_{\omega_L}(1) = \frac{3}{22}$, $f_{\omega_L}(2) = \frac{2}{22}$, $f_{\omega_L}(1) = \frac{1}{22}$, and $f_{\omega_H}(s) = f_{\omega_L}(-s)$.




Let $u = 1$, which represents an unbiased taste for action. In this case, Assumption 1 holds but Assumption 2 fails. As one can see in the graph on the right, because $U_L$ and $U_H$ cross multiple times, there are intervals of $p_0$ where the user prefers a prior-contradicting title, but there are also intervals where she prefers a prior-conforming title.

## 3.2 Platform's optimal algorithm

Recall that the platform's goal is to maximize the chance that the user clicks on the recommended title. When Assumptions 1 and 2 hold, because the user prefers a prior-contradicting title, she is more willing to pay the cost $c$ and click on it. It follows immediately that the platform should always recommend a prior-contradicting title in equilibrium. This section formalizes this intuition.

For $i = H, L$, let $\Delta U_i \equiv U_i - U_0$ denote the user's increase in expected utility if she clicks on title $t_i$, without taking cost into consideration. When Assumption 1 and 2 hold, Figure 1 describes the relative positions of $\Delta U_L$ and $\Delta U_H$ as results of Proposition 1 and Theorem 1. Note that the linearity in Figure 1 is not necessarily accurate, but the relative ranking of $\Delta U_L$ and $\Delta U_H$, as well as the increasing-decreasing pattern, are accurate.

Recall that the user clicks on title $t_i$ if and only if $\Delta U_i \geq c > 0$. Proposition 2, as a direct result of Figure 1, states that it is always optimal for the platform to recommend a prior-contradicting title regardless of the clicking cost $c$. Note that when $c$ is above $\max\{\Delta U_L(p_0), \Delta U_H(p_0)\}$, the user never clicks regardless of the recommendation. When $c$ is below $\min\{\Delta U_L(p_0), \Delta U_H(p_0)\}$, the user is willing to click on both types of titles. In these two cases, the platform's optimal strategy is not unique because it is indifferent between all strategies. However, if $c$ is a number or a random variable[7] whose value is in between $\min\{\Delta U_L(p_0), \Delta U_H(p_0)\}$ and $\max\{\Delta U_L(p_0), \Delta U_H(p_0)\}$ with positive probability, the prior-contradicting recommendation strategy is uniquely optimal for the platform.

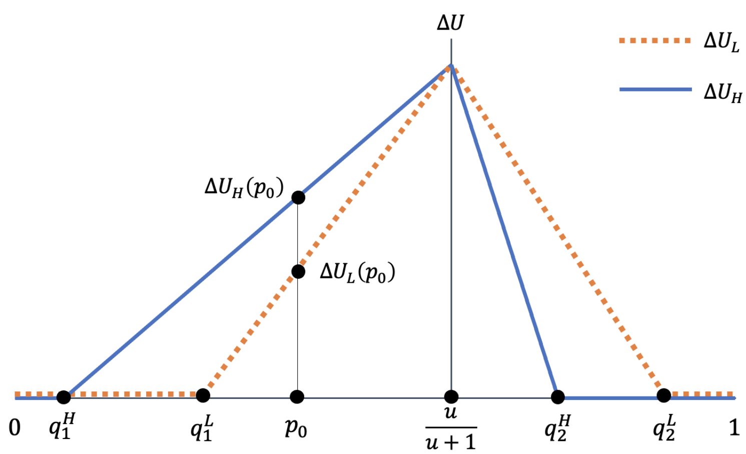



Figure 1: $\Delta U_L = U_L - U_0$ and $\Delta U_H = U_H - U_0$

**Proposition 2.** *When Assumptions 1 and 2 hold, the following algorithm maximizes the platform's expected payoff in equilibrium:*

*Always recommend $t_H$ if $p_0 < \frac{u}{u+1}$ and $t_L$ if $p_0 \geq \frac{u}{u+1}$.*

*Moreover, if $p_0 < \frac{u}{u+1}$ and $c \in (\Delta U_L(p_0), \Delta U_H(p_0)]$ with positive probability, the platform's unique equilibrium strategy is to recommend $t_H$; if $p_0 > \frac{u}{u+1}$ and $c \in (\Delta U_H(p_0), \Delta U_L(p_0)]$ with positive probability, the platform's unique equilibrium strategy is to recommend $t_L$.*

Proposition 2 is the formal response to the motivating question: "does a click-maximizing digital platform provide personalized news that challenges our existing view of the world, or news that panders to it?" It gives a positive answer: "It provides news that challenges our existing view of the world, because we demand it."

# 4 Extensions and Discussions

In this section, I explore three alternative modeling assumptions and discuss the robustness of the result with respective to these alternatives.

[7]One can argue that it is more realistic to treat $c$ as a random variable if it represents the user's subjective dis-utility for the time spent on reading at the particular moment.

## 4.1 If the platform knows the true state

In the main model, the platform cannot condition its recommendation on the true state $\omega$. Consequently, its title recommendation does not convey any information on $\omega$, and the user's belief stays at $p_0$ unless she clicks.

Consider, instead, that the platform can condition its recommendation on $\omega$.[8] The user can update her belief twice in the game: when she observes the title $t_i$, she updates her belief from $p_0$ to an interim belief $p_1$ based on the platform's algorithm and the realized $t_i$; then, if she clicks on $t_i$, she updates $p_1$ to her posterior belief $p_2$ based on the realization of $s$.

The platform's recommendation has a two-fold impact on the user: (1) by choosing a state-dependent algorithm, it can manipulate the user's interim belief; (2) by choosing which title to recommend, it chooses which truncated distribution of articles is available to the user.

The optimization problem in (2) is the same as that in the baseline model. The optimization problem in (1) is new, and is an application of the Bayesian persuasion problem as in Kamenica and Gentzkow (2011).

Specifically, every result in Section 3.1 (user's preference) continues to hold except that one should change $p_0$ to $p_1$ in every statement. The patterns of $\Delta U_L$ and $\Delta U_H$ as depicted in Figure 1 are also the same. When Assumptions 1 and 2 hold and $c \leq \max\left\{\Delta U_L\left(p_0\right), \Delta U_H\left(p_0\right)\right\}$, it is optimal for the platform to adopt the same algorithm as in Proposition 2: always recommend $t_H$ if $p_0 < \frac{u}{u+1}$ and $t_L$ if $p_0 \geq \frac{u}{u+1}$. Because this algorithm is state-independent, the realization of the title conveys no information about the true state, and $p_1 = p_0$. The platform's expected payoff is already maximized, so no belief manipulation (or concavification, in the words of Kamenica and Gentzkow 2011) is needed.

Let $\Delta U_{max} = \Delta U_L\left(\frac{u}{u+1}\right) = \Delta U_H\left(\frac{u}{u+1}\right)$ be the largest feasible utility gain from a click, which is achievable only when $p_0 = \frac{u}{u+1}$. The main departure from the baseline model's equilibrium outcome occurs when $c \in \left(\max\left\{\Delta U_L\left(p_0\right), \Delta U_H\left(p_0\right)\right\}, \Delta U_{max}\right]$. In this case, the user will not click on any title under her prior belief. However, if the platform can manipulate her belief so that $p_1$ is sometimes sufficiently close to $\frac{u}{u+1}$ and $\Delta U$ is sufficiently close to $\Delta U_{max}$, she will click with positive probability. Below, I apply the solution in Kamenica and Gentzkow (2011) to identify the platform's optimal algorithm.

Without loss of generality, suppose that $p_0 < \frac{u}{u+1}$. Let $c$ be a number in $\left(\Delta U_H\left(p_0\right), \Delta U_{max}\right]$. Define $p_H$ such that $\Delta U_H\left(p_H\right) = c$. By construction, $p_0 < p_H \leq \frac{u}{u+1}$, and $p_H$ is the lowest belief at which the user will click on any title (which is $t_H$ at $p_H$). The website's maximizes

[8] For example, suppose that the platform observes the ratio $\alpha$ between the numbers of positive and negative articles in a large sample with size $n$. Then, by the law of large numbers, as $n \to \infty$, whether $\alpha$ is bigger or smaller than $\frac{1}{2}$ reveals the true state. By conditioning its recommendation on whether $\alpha > \frac{1}{2}$ or $\alpha < \frac{1}{2}$, the platform can effectively condition its recommendation on $\omega$.

the probability that $p_1 \geq p_H$ by adopting the following algorithm: always recommend $t_H$ in state $\omega_H$; recommend $t_H$ with probability $\frac{p_0(1-p_H)}{(1-p_0)p_H}$ in state $\omega_L$. When the user sees $t_H$, her interim belief is $p_1 = p_H$ and she clicks. When she sees $t_L$, her interim belief is $p_1 = 0$ and she does not click. Without changing the user's ex-ante expected utility, this algorithm increases the user's ex-ante probability of a click from 0 to $\frac{p_0}{p_H}$.

The solution when $p_0 > \frac{u}{u+1}$ is similar. The recommended title always contradicts the user's prior when her default action under the prior is wrong. The recommended title contradicts the user's prior with positive probability even if her default action under the prior is correct. For these cases when $c \in (\max\{\Delta U_L(p_0), \Delta U_H(p_0)\}, \Delta U_{max}]$, driven by both user's demand and optimal belief manipulation, the platform recommends prior-contradicting titles more often than not.

## 4.2 If the platform cannot personalize news

Suppose that the platform does not know $p_0$ or $u$ and cannot make any personalized recommendation. Then, the optimal recommendation can, at most, be conditioned on the population distribution of users' types: if it is more likely that a user satisfies $p_0 < \frac{u}{u+1}$, the platform recommends $t_H$; otherwise, the platform recommends $t_L$. This means that for a user with the minority bias, the platform does not recommend the welfare-maximizing, prior-contradicting title to her. This can be problematic if the population weights are similar for the two directions of prior biases, as observed in the 2016 Brexit vote and the 2016 US presidential election. This paper suggests that the adoption of a personalization technology leads to a Pareto improvement for the user body, and provides the minority users with more prior-contradicting news as well as higher expected utility.

## 4.3 If the user also enjoys entertainment value from news

The user in this paper consumes news only for the instrumental value of its information. However, one can easily modify the model to accommodate a user who also gets direct psychological utility from reading an article. I call this the "entertainment value" of news articles, to separate it from the instrumental value of their information. To model the entertainment value, one can simply change the clicking cost $c$: make it lower for an article with greater entertainment value, and higher for an article with less entertainment value. If the platform does not have much information about a user's entertainment value, it will treat $c$ as a random variable, and apply the optimal algorithm in Proposition 2.

If the user has different entertainment values for different types of articles and the platform knows this, the equilibrium algorithm may change. For example, suppose that the user has a *confirmation bias*: she gets pleasure from clicking on a title that echos her bias (e.g., it feels good to read positive coverage of the political candidate she supports). Then, $c$ is higher for the prior-contradicting title and lower for the prior-conforming title. Proposition 2 holds if this confirmation bias is sufficiently moderate, but will fail if the bias is large. Nevertheless, this paper is useful even in the latter case. The model in this paper provides a framework for econometricians and experimentalists to separate the instrumental value of news from the entertainment value. When Assumptions 1 and 2 hold, the results in this paper set the prior-contradicting equilibrium outcome as a benchmark when a user is rational and does not enjoy any entertainment value from news consumption; any deviation from this outcome can be identified as a result of the entertainment value or other behavioral biases. In contrast, this identification would have been difficult under alternative models $A$ or $B$ in the Introduction. In those models, rational users who care only about information and users who have a confirmation bias both exhibit a preference for prior-conforming news.

# 5 Appendix

## 5.1 Continuity of $U_L$ and $U_H$

**Lemma 3.** *$U_L$ and $U_H$ are continuous in $p_0$ for all $p_0 \in [0,1]$.*

*Proof.* Take any arbitrary $p_0 \in [0,1]$. If $S_1^L = \emptyset$ then $U_L = \sum_{s \le 0} q_{t_L}(s)\, p_2(s \mid t_L)$. If $S_2^L = \emptyset$ then $U_L = \sum_{s \le 0} q_{t_L}(s)\, u\left[1 - p_2(s \mid t_L)\right]$. In both cases, $U_L$ is continuous at $p_0$ because $q_{t_L}(s)$ and $p_2(s \mid t_L)$ are both continuous in $p_0$ for all $s$.

Next, suppose that neither $S_1^L$ or $S_2^L$ is empty. $U_L = \sum_{s \in S_1^L} q_{t_L}(s)\, u\left[1 - p_2(s \mid t_L)\right] + \sum_{s \in S_2^L} q_{t_L}(s)\, p_2(s \mid t_L)$. Define $s_L^*(p_0)$ such that $p_2(s_L^*(p_0) \mid t_L) = \frac{u}{u+1}$. If $\nexists\, s_L^*(p_0) \in S$ that satisfies this equation, then an infinitesimal drift from $p_0$ does not change the elements in $S_1^L$ or $S_2^L$ because the new $p_0$ yields the same partition. Hence, in this case, $\lim_{p \to p_0} U_L(p) = U_L(p_0)$. If $\exists s_L^*(p_0) \in S$ that satisfy $p_2(s_L^*(p_0) \mid t_L) = \frac{u}{u+1}$ then there are two possible cases. In the first case, if neither $f_{\omega_L}$ nor $f_{\omega_H}$ has an atom at $s_L^*(p_0)$ then an infinitesimal drift from $p_0$ yields an infinitesimal shift in the probability weights on $S_1^L$ and $S_2^L$ and, hence, the value of $U_L$. This implies that $U_L$ is continuous at $p_0$. In the second case, suppose that $f_{\omega_L}$ or $f_{\omega_H}$ has an atom at $s_L^*(p_0)$. Note that the user is indifferent between $a_L$ and $a_H$ when the signal realization is exactly $s_L^*$. This means that $u\left[1 - p_2(s_L^*(p_0) \mid t_L)\right] = p_2(s_L^*(p_0) \mid t_L)$. Let $\hat{p_0} = p_0 + \Delta$ for some $\Delta > 0$ and suppose that $\hat{S}_1^L$ and $\hat{S}_2^L$ are the corresponding sets after the new partition. Because $\hat{p_0}$ is higher than $p_0$, there exists $s_L^*(\hat{p_0}) \le s_L^*(p_0)$

such that the signals in the interval $I = [s_L^*(\hat{p_0}), s_L^*(p_0))$ now belong in $\hat{S}_2^L$ (they belong to $S_1^L$ when $\Delta = 0$). As $\Delta \to 0$, for any $s \in I$, $s \to s_L^*(p_0)$ and $u[1 - p_2(s \mid t_L)] \to u[1 - p_2(s_L^*(p_0) \mid t_L)] = p_2(s_L^*(p_0) \mid t_L)$. Therefore,

$$\begin{aligned}
\lim_{\Delta \to 0} U_L(p_0 + \Delta) - U_L(p_0) &= \lim_{\Delta \to 0} \sum_{s \in I} - q_{t_L}(s)\, u[1 - p_2(s \mid t_L)] \\
&\quad + \lim_{\Delta \to 0} \sum_{s \in I} q_{t_L}(s)\, p_2(s \mid t_L) \\
&= \sum_{s \in I} - q_{t_L}(s)\, p_2(s_L^*(p_0) \mid t_L) \\
&\quad + \sum_{s \in I} q_{t_L}(s)\, p_2(s_L^*(p_0) \mid t_L) \\
&= 0
\end{aligned}$$

Using essentially the same argument, one can show that the same result holds for $\Delta < 0$. This proves that $U_L$ is continuous at any $p_0 \in [0, 1]$.

Finally, one can apply the same method of proof to the case of title $t_H$ to show that the user's expected utility after clicking, $U_H$, is also continuous at any $p_0 \in [0, 1]$. □

## 5.2 $U_L = U_H$ when the user is indifferent between the two actions under her prior belief

**Lemma 4.** *$U_L = U_H$ when $p_0 = \frac{u}{u+1}$.*

*Proof.* Suppose that $p_0 = \frac{u}{u+1}$. Because the user is currently indifferent between $a_L$ and $a_H$ at $p_0$, the realized $s$ behind $t_L$ or $t_R$ can swing her to either side of indifference. In other words, $S_1^L$, $S_2^L$, $S_1^H$, $S_2^H$ are all non-empty. Note that $p_2(s \mid t_L) \leq \frac{u}{u+1}$ if and only if $\frac{f_{\omega_H}(s)}{F_{\omega_H}(0)} \leq \frac{f_{\omega_L}(s)}{F_{\omega_L}(0)}$ and $p_2(s \mid t_R) \geq \frac{u}{u+1}$ if and only if $\frac{f_{\omega_H}(s)}{1-F_{\omega_H}(0)} \geq \frac{f_{\omega_L}(s)}{1-F_{\omega_L}(0)}$. Define $s_L^*$ to be the largest element in $S$ that satisfies $\frac{f_{\omega_H}(s)}{F_{\omega_H}(0)} \leq \frac{f_{\omega_L}(s)}{F_{\omega_L}(0)}$, and $s_H^*$ to be the smallest element in $S$ that satisfies $\frac{f_{\omega_H}(s)}{1-F_{\omega_H}(0)} \geq \frac{f_{\omega_L}(s)}{1-F_{\omega_L}(0)}$. Recall that by the symmetry assumption, $f_{\omega_L}(s) = f_{\omega_H}(-s)$ and $F_{\omega_L}(s) = 1 - F_{\omega_H}(-s)$ for all $s \in S$. Therefore, $s_L^* = -s_H^*$.

When $p_0 = \frac{u}{u+1}$ or, equivalently, $u = \frac{p_0}{1-p_0}$, the user's expected utilities after clicking on

$t_L$ or $t_H$ are

$$\begin{aligned} U_L &= \sum_{s \in S_1^L} u \left[ \frac{f_{\omega_L}(s)}{F_{\omega_L}(0)} (1-p_0) \right] + \sum_{s \in S_2^L} \frac{f_{\omega_H}(s)}{F_{\omega_H}(0)} p_0 \\ &= \left[ \frac{F_{\omega_L}(s_L^*)}{F_{\omega_L}(0)} u (1-p_0) \right] + \frac{F_{\omega_H}(0) - F_{\omega_H}(s_L^*)}{F_{\omega_H}(0)} p_0 \\ &= \left[ \frac{F_{\omega_L}(s_L^*)}{F_{\omega_L}(0)} + \frac{F_{\omega_H}(0) - F_{\omega_H}(s_L^*)}{F_{\omega_H}(0)} \right] p_0 \end{aligned}$$

$$\begin{aligned} U_H &= \sum_{s \in S_1^H} u \left[ \frac{f_{\omega_L}(s)}{1 - F_{\omega_L}(0)} (1-p_0) \right] + \sum_{s \in S_2^H} \left[ \frac{f_{\omega_H}(s)}{1 - F_{\omega_H}(0)} p_0 \right] \\ &= \left[ \frac{F_{\omega_L}(s_H^*) - F_{\omega_L}(0)}{1 - F_{\omega_L}(0)} u (1-p_0) \right] + \left[ \frac{1 - F_{\omega_H}(s_H^*)}{1 - F_{\omega_H}(0)} p_0 \right] \\ &= \left[ \frac{F_{\omega_H}(0) - F_{\omega_H}(s_L^*)}{F_{\omega_H}(0)} + \frac{F_{\omega_L}(s_L^*)}{F_{\omega_L}(0)} \right] p_0 \text{ by symmetry} \\ &= U_L \end{aligned}$$

This proof has one caveat. The calculations of $U_L$ and $U_H$ above assumed that $\sum_{s \in S_1^L} f_{\omega_L}(s) = F_{\omega_L}(s_L^*)$, $\sum_{s \in S_2^L} f_{\omega_H}(s) = F_{\omega_H}(0) - F_{\omega_H}(s_L^*)$, $\sum_{s \in S_1^H} f_{\omega_L}(s) = F_{\omega_L}(s_H^*) - F_{\omega_L}(0)$, and $\sum_{s \in S_2^H} f_{\omega_H}(s) = 1 - F_{\omega_H}(s_H^*)$. These equations are satisfied except in one case defined by two conditions: (1) $\frac{\Pr(s_L^* | \omega_H)}{F_{\omega_H}(0)} = \frac{\Pr(s_L^* | \omega_L)}{F_{\omega_L}(0)}$ and $\frac{\Pr(s_H^* | \omega_H)}{1 - F_{\omega_H}(0)} = \frac{\Pr(s_H^* | \omega_L)}{1 - F_{\omega_L}(0)}$. Either one of these equations implies the other by symmetry, and they jointly imply indifference between actions at $s_L^*$ and $s_H^*$, as well as the relation $s_L^* = -s_H^*$. (2) Additionally, assume that $\Pr(s_L^* \mid \omega_L)$ and $\Pr(s_H^* \mid \omega_L)$ are both strictly positive.

Below, I slightly alter the calculation of $U_L$ and $U_H$ to show that $U_L = U_H$ when conditions (1) and (2) are met. For this special case, change the user's tie-breaking rule so that she chooses $a_H$ when she is indifferent and the title is $t_L$ but chooses $a_L$ when she is indifferent and the title is $t_H$. This change of the tie-breaking rule does not change the value of the user's expected utilities $U_L$ and $U_R$. The new tie-breaking rule implies that $s_L^* \in S_2^L$ but $s_H^* \in S_1^H$, which leads to the following formulae:

$$\begin{aligned} U_L &= \left[ \frac{F_{\omega_L}(s_L^*) - \Pr(s_L^* \mid \omega_L)}{F_{\omega_L}(0)} u (1-p_0) \right] + \frac{F_{\omega_H}(0) - F_{\omega_H}(s_L^*)}{F_{\omega_H}(0)} p_0 \\ &= \left[ \frac{F_{\omega_L}(s_L^*)}{F_{\omega_L}(0)} + \frac{F_{\omega_H}(0) - F_{\omega_H}(s_L^*)}{F_{\omega_H}(0)} \right] p_0 - \frac{\Pr(s_L^* \mid \omega_L)}{F_{\omega_L}(0)} p_0 \end{aligned}$$

$$
\begin{aligned}
U_H &= \sum_{s\in S_1^H} u\left[\frac{f_{\omega_L}(s)}{1-F_{\omega_L}(0)}(1-p_0)\right] + \sum_{s\in S_2^H}\left[\frac{f_{\omega_H}(s)}{1-F_{\omega_H}(0)}p_0\right] \\
&= \left[\frac{F_{\omega_L}(s_H^*)-F_{\omega_L}(0)}{1-F_{\omega_L}(0)}u(1-p_0)\right] + \left[\frac{1-F_{\omega_H}(s_H^*)-\Pr(s_H^*\mid\omega_H)}{1-F_{\omega_H}(0)}p_0\right] \\
&= \left[\frac{F_{\omega_H}(0)-F_{\omega_H}(s_L^*)}{F_{\omega_H}(0)} + \frac{F_{\omega_L}(s_L^*)}{F_{\omega_L}(0)}\right]p_0 - \frac{\Pr(s_H^*\mid\omega_H)}{1-F_{\omega_H}(0)}p_0 \text{ by previous proof} \\
&= \left[\frac{F_{\omega_H}(0)-F_{\omega_H}(s_L^*)}{F_{\omega_H}(0)} + \frac{F_{\omega_L}(s_L^*)}{F_{\omega_L}(0)}\right]p_0 - \frac{\Pr(s_L^*\mid\omega_L)}{F_{\omega_L}(0)}p_0 \text{ by symmetry} \\
&= U_L
\end{aligned}
$$

This concludes the proof that $U_L = U_H$ when $p_0 = \frac{u}{u+1}$. □

## 5.3 Proof of Lemma 1 (thresholds $q_1^i, q_2^i$)

After clicking on $t_L$ or $t_H$ and learning the negative or positive $s$, the user's posterior belief for state $\omega_H$ becomes $p_2(s \mid t_L)$ and $p_2(s \mid t_H)$ with

$$
\begin{aligned}
\frac{p_2(s\mid t_L)}{1-p_2(s\mid t_L)} &= \frac{\frac{f_{\omega_H}(s)}{F_{\omega_H}(0)}}{\frac{f_{\omega_L}(s)}{F_{\omega_L}(0)}} \cdot \frac{p_0}{1-p_0} \\
\frac{p_2(s\mid t_H)}{1-p_2(s\mid t_H)} &= \frac{\frac{f_{\omega_H}(s)}{1-F_{\omega_H}(0)}}{\frac{f_{\omega_L}(s)}{1-F_{\omega_L}(0)}} \frac{p_0}{1-p_0}
\end{aligned}
$$

Both of these likelihood ratios are strictly increasing in $s$ and $p_0$. They are strictly increasing in $s$ because $f_{\omega_H}(s)/f_{\omega_L}(s)$ is strictly increasing in $s$ by assumption.

If $\inf\{f_{\omega_H}(s)/f_{\omega_L}(s) \mid s \leq 0\} = 0$ then the user will always choose $a_L$ with positive probability unless $p_0 = 1$, which implies that $q_2^L = 1$. If $f_{\omega_H}(s)/f_{\omega_L}(s) \to \infty$ as $s \geq 0$ increases then the user will always choose $a_H$ with positive probability unless $p_0 = 0$, which implies that $q_1^H = 0$. In all other cases, $\sup\{f_{\omega_H}(s)/f_{\omega_L}(s) \mid s \leq 0\}$, $\sup\{f_{\omega_H}(s)/f_{\omega_L}(s) \mid s \geq 0\}$, $\inf\{f_{\omega_H}(s)/f_{\omega_L}(s) \mid s \leq 0\}$, and $\inf\{f_{\omega_H}(s)/f_{\omega_L}(s) \mid s \geq 0\}$ are all strictly positive and finite. For these cases, define $q_1^i, q_2^i$ for $i = L, H$ such that

$$\begin{aligned}
\sup\left\{f_{\omega_H}(s)/f_{\omega_L}(s) \mid s \leq 0\right\} \cdot \frac{F_{\omega_L}(0)}{F_{\omega_H}(0)} \cdot \frac{q_1^L}{1-q_1^L} &= u \\
\sup\left\{f_{\omega_H}(s)/f_{\omega_L}(s) \mid s \geq 0\right\} \cdot \frac{1-F_{\omega_L}(0)}{1-F_{\omega_H}(0)} \cdot \frac{q_1^H}{1-q_1^H} &= u \\
\inf\left\{f_{\omega_H}(s)/f_{\omega_L}(s) \mid s \leq 0\right\} \cdot \frac{F_{\omega_L}(0)}{F_{\omega_H}(0)} \cdot \frac{q_2^L}{1-q_2^L} &= u \\
\inf\left\{f_{\omega_H}(s)/f_{\omega_L}(s) \mid s \geq 0\right\} \cdot \frac{1-F_{\omega_L}(0)}{1-F_{\omega_H}(0)} \cdot \frac{q_2^H}{1-q_2^H} &= u
\end{aligned}$$

$q_1^i, q_2^i \in (0,1)$ and $q_1^i \leq q_2^i$ for $i = L, H$. Moreover, $\frac{p_2(s|t_L)}{1-p_2(s|t_L)} \geq u$ for some feasible $s \leq 0$ if and only if $p_0 \geq q_1^L$; $\frac{p_2(s|t_L)}{1-p_2(s|t_L)} \geq u$ for some feasible $s \geq 0$ if and only if $p_0 \geq q_1^H$; $\frac{p_2(s|t_L)}{1-p_2(s|t_L)} \geq u$ for all $s \leq 0$ if and only if $p_0 \geq q_2^L$; $\frac{p_2(s|t_L)}{1-p_2(s|t_L)} \geq u$ for all $s \geq 0$ if and only if $p_0 \geq q_2^H$. This implies that for $i = L, H$, the user with prior belief $p_0$ chooses $a_H$ with positive probability after clicking on $t_i$ if and only if $p_0 \geq q_1^i$, and she always chooses $a_H$ after clicking on $t_i$ if and only if $p_0 \geq q_2^i$, which is the statement of the Lemma.

## 5.4 Proof of Lemma 2 (equivalence of Assumptions 1, 1a, and 1b)

*Equivalence of 1 and 1a:*

The definition of $q_1^H, q_1^L, q_2^H, q_2^L$ implies the following: for $i = L, H$

$$\begin{cases} S_1^i = S,\ S_2^i = \emptyset & \text{when } p_0 < q_1^i \\ S_1^i \neq \emptyset,\ S_2^i \neq \emptyset & \text{when } q_1^i \leq p_0 < q_2^i \\ S_1^i = \emptyset,\ S_2^i = S & \text{when } p_1 \geq q_2^i \end{cases}$$

Then, given $p_0$, $S_1^i = \emptyset \Leftrightarrow p_0 \geq q_2^i$ and $S_2^i = \emptyset \Leftrightarrow p_0 < q_1^i$. Assumption 1 is equivalent to the following: for any $p_0 \in [0,1]$, if $p_0 \geq q_2^L$ then $p_0 \geq q_2^H$ but the converse is not always true; if $p_0 < q_1^H$ then $p_0 < q_1^L$ but the converse is not always true. This statement is true if and only if $q_1^H < q_1^L$ and $q_2^H < q_2^L$, which is Assumption 1a.

*Equivalence of 1a and 1b:*

Let $s_0$ denote the smallest $s \in S$ such that $s \geq 0$. Let $s_\infty$ denote the largest element in $S$. If $S$ does not have an upper bound, let $s_\infty = \infty$.

By symmetry, $-s_0$ is the largest $s \in S$ such that $s \leq 0$ and $-s_\infty$ is the smallest element in $S$ or $-\infty$ if $S$ does not have a lower bound.

Then, by the monotonic likelihood ratio assumption,

$$
\begin{aligned}
\sup\left\{f_{\omega_H}(s)/f_{\omega_L}(s) \mid s \le 0\right\} &= f_{\omega_H}(-s_0)/f_{\omega_L}(-s_0) \\
\sup\left\{f_{\omega_H}(s)/f_{\omega_L}(s) \mid s \ge 0\right\} &= \lim_{s\to s_\infty} f_{\omega_H}(s_\infty)/f_{\omega_L}(s_\infty) \\
\inf\left\{f_{\omega_H}(s)/f_{\omega_L}(s) \mid s \le 0\right\} &= \lim_{s\to -s_\infty} f_{\omega_H}(-s_\infty)/f_{\omega_L}(-s_\infty) \\
\inf\left\{f_{\omega_H}(s)/f_{\omega_L}(s) \mid s \ge 0\right\} &= f_{\omega_H}(s_0)/f_{\omega_L}(s_0)
\end{aligned}
$$

Based on the definitions of $q_1^H, q_1^L$ and $q_2^H, q_2^L$ in the proof of Lemma 1, one can see that Assumption 1a holds if and only if

$$
\frac{f_{\omega_H}(-s_\infty)/F_{\omega_H}(0)}{f_{\omega_L}(-s_\infty)/F_{\omega_L}(0)} < \frac{f_{\omega_H}(s_0)/\left[1-F_{\omega_H}(0)\right]}{f_{\omega_L}(s_0)/\left[1-F_{\omega_L}(0)\right]} \text{ and } \frac{f_{\omega_H}(s_\infty)/\left[1-F_{\omega_H}(0)\right]}{f_{\omega_L}(s_\infty)/\left[1-F_{\omega_L}(0)\right]} > \frac{f_{\omega_H}(-s_0)/F_{\omega_H}(0)}{f_{\omega_L}(-s_0)/F_{\omega_L}(0)}.
$$

Given the formula of $p_2$, when $p_0 = \frac{1}{2}$, the inequalities above are equivalent to

$$
p_2(-s_\infty \mid t_L) < p_2(s_0 \mid t_H) \text{ and } p_2(s_\infty \mid t_H) > p_2(-s_0 \mid t_L). \qquad (*)
$$

This implies that Assumption 1b holds, thus proving that Assumption 1a $\Rightarrow$ Assumption 1b.

Suppose that Assumption 1b holds. Because of the monotonic likelihood assumption, if there exists some $s \le 0$ such that $p_2(s \mid t_L) < p_2(s_0 \mid t_H)$, then $p_2(-s_\infty \mid t_L) < p_2(s_0 \mid t_H)$; if there exists some $s \ge 0$ such that $p_2(s \mid t_H) > p_2(-s_0 \mid t_L)$ then $p_2(s_\infty \mid t_H) > p_2(-s_0 \mid t_L)$. When $p_0 = \frac{1}{2}$, Assumption 1b $\Rightarrow$ Condition $(*)$ $\Leftrightarrow$ Assumption 1a.

This concludes the proof for the equivalence of Assumptions 1, 1a and 1b.

## 5.5 Proof of Proposition 1 (compare $U_i$ with $U_0$)

For $i = L, H$:

(a) When $p_0 \in [0, q_1^i]$, $U_0$ is the user's expected utility if she always chooses $a_L$. If she clicks on title $t_i$, she also always chooses $a_L$ if $p_0 \in [0, q_1^i)$, so her expected utility $U_i$ is the same as $U_0$. If $p_0 = q_1^i$, she chooses $a_L$ except when she learns the largest $s$ behind $t_i$, in which case she is indifferent between $a_L$ and $a_H$ (she chooses $a_H$ because of the tie-breaking rule, but her utility would be the same if she chooses $a_L$ in this event). This implies that $U_i = U_0$.

When $p_0 \in [q_2^i, 1]$, $U_0$ is the user's expected utility if she always chooses $a_H$. If she clicks on title $t_i$, she also always chooses $a_H$, so $U_i = U_0$.

(b) When $p_0 \in (q_1^i, q_2^i)$, let $a_0$ denote the user's choice of action under $p_0$ if she does not click on any title. After she clicks on title $t_i$, there exists non-empty sets $\{\mu_i\}$ and $\{\mu_j\}$ of posterior belief realizations such that each $\mu_i$ makes the user strictly prefer $a \neq a_0$ and each $\mu_j$ makes her weakly prefer $a_0$. Let $q_i$ be the associated probability of posterior belief realization $\mu_i$ and let $q_j$ be the associated probability of posterior belief realization $\mu_j$. For $\mu$ and $q$ to be well-defined, they must satisfy two conditions: (1) $\sum\limits_i q_i + \sum\limits_j q_j = 1$ and (2) $\sum\limits_i q_i \mu_i + \sum\limits_j q_j \mu_j = p_0$ (the expectation of posterior belief is equal to the prior belief). Then,

$$\begin{aligned}
U_i &= \sum_i q_i \cdot EU\left(\text{action } a \neq a_0 \text{ under belief } \mu_i\right) + \sum_j q_j \cdot EU\left(\text{action } a_0 \text{ under belief } \mu_j\right) \\
&> \sum_i q_i \cdot EU\left(\text{action } a_0 \text{ under belief } \mu_i\right) + \sum_j q_j \cdot EU\left(\text{action } a_0 \text{ under belief } \mu_j\right) \\
&= EU\left(\text{always choose } a_0\right) \\
&= U_0
\end{aligned}$$

Therefore, $U_i > U_0$ when $p_0 \in (q_1^i, q_2^i)$.

(c) When $p_0 \in (q_1^i, q_2^i)$, neither $S_1^i$ nor $S_2^i$ is empty. Define $s_i^*$ as the supremum of $S_1^i$. Then,

$$\begin{aligned}
U_L &= \sum_{s \in S_1^L} u\left[\frac{f_{\omega_L}(s)}{F_{\omega_L}(0)}(1-p_0)\right] + \sum_{s \in S_2^L} \frac{f_{\omega_H}(s)}{F_{\omega_H}(0)} p_0 \\
&= \frac{F_{\omega_L}(s_L^*)}{F_{\omega_L}(0)} u(1-p_0) + \frac{F_{\omega_H}(0) - F_{\omega_H}(s_L^*)}{F_{\omega_H}(0)} p_0
\end{aligned}$$

$$\begin{aligned}
U_H &= \sum_{s \in S_1^H} u\left[\frac{f_{\omega_L}(s)}{1-F_{\omega_L}(0)}(1-p_0)\right] + \sum_{s \in S_2^H} \left[\frac{f_{\omega_H}(s)}{1-F_{\omega_H}(0)} p_0\right] \\
&= \frac{F_{\omega_L}(s_H^*) - F_{\omega_L}(0)}{1-F_{\omega_L}(0)} u(1-p_0) + \frac{1-F_{\omega_H}(s_H^*)}{1-F_{\omega_H}(0)} p_0
\end{aligned}$$

When $p_0 < \frac{u}{u+1}$, $U_0 = u(1-p_0)$.

$$U_L - U_0 = \underbrace{\left[\frac{F_{\omega_L}(s_L^*)}{F_{\omega_L}(0)} - 1\right]}_{<0} u(1-p_0) + \underbrace{\frac{F_{\omega_H}(0) - F_{\omega_H}(s_L^*)}{F_{\omega_H}(0)}}_{>0} p_0$$

is strictly increasing in $p_0$, and

$$U_H - U_0 = \underbrace{\left[\frac{F_{\omega_L}(s_H^*) - F_{\omega_L}(0)}{1 - F_{\omega_L}(0)} - 1\right]}_{<0} u(1-p_0) + \underbrace{\frac{1 - F_{\omega_H}(s_H^*)}{1 - F_{\omega_H}(0)}}_{>0} p_0$$

is also strictly increasing in $p_0$.

When $p_0 > \frac{u}{u+1}$, $U_0 = p_0$.

$$U_L - U_0 = \underbrace{\frac{F_{\omega_L}(s_L^*)}{F_{\omega_L}(0)}}_{>0} u(1-p_0) + \underbrace{\left[\frac{F_{\omega_H}(0) - F_{\omega_H}(s_L^*)}{F_{\omega_H}(0)} - 1\right]}_{<0} p_0$$

is strictly decreasing in $p_0$, and

$$U_H - U_0 = \underbrace{\frac{F_{\omega_L}(s_H^*) - F_{\omega_L}(0)}{1 - F_{\omega_L}(0)}}_{>0} u(1-p_0) + \underbrace{\left[\frac{1 - F_{\omega_H}(s_H^*)}{1 - F_{\omega_H}(0)} - 1\right]}_{<0} p_0$$

is also strictly decreasing in $p_0$.

This proves statement (c) and concludes the proof of the Proposition.

## 5.6 Proof of Theorem 1 (user prefers a prior-contradicting title)

### 5.6.1 Proof that the statement in Theorem 1 holds if Assumptions 1 and 2 hold

This proof is done in two steps:

1. Ranking of thresholds $q$

   Lemma 1 proves that $q_1^L < q_2^L$ and $q_1^H < q_2^H$. Assumption 1 implies that $q_1^L > q_1^H$ and $q_2^L > q_2^H$ (Assumption 1a). Moreover, $q_1^i < \frac{u}{u+1} < q_2^i$ for $i = L, H$. To show the last point, note that when $p_0 = \frac{u}{u+1}$, the user is indifferent between $a_L$ and $a_H$ if she does not click on any title. If she clicks on title $t_i$, her posterior belief will be a mean-preserving spread of $p_0$. This implies that her posterior belief can be strictly higher or lower than $\frac{u}{u+1}$ with positive probability. In other words, the probability that she chooses $a_i$ is strictly positive for both $i = L, H$. By the definition of $q$, this implies that $q_1^i < \frac{u}{u+1} < q_2^i$.

These inequalities imply the following ranking of thresholds $q$:

$$0 \leq q_1^H < q_1^L < \frac{u}{u+1} < q_2^H < q_2^L \leq 1.$$

2. Compare $U_L$ with $U_H$

   By Proposition 1, $U_H > U_L$ when $p_0 \in \left(q_1^H, q_1^L\right]$ because $U_H > U_0$ but $U_L = U_0$ when $p_0 \in \left(q_1^H, q_1^L\right]$. Similarly, $U_L > U_H$ when $p_0 \in \left[q_2^H, q_2^L\right)$.

   Because $U_H > U_L$ at $p_0 = q_1^L$, $U_H = U_L$ at $p_0 = \frac{u}{u+1}$ (Lemma 4), $U_H \neq U_L$ for any $p_1 \in \left(q_1^L, \frac{u}{u+1}\right)$ (Assumption 2), by continuity of $U_H$ and $U_L$ (Lemma 3), $U_H > U_L$ for any $p_1 \in \left(q_1^L, \frac{u}{u+1}\right)$. Similarly, because $U_H < U_L$ at $p_0 = q_2^H$, $U_H = U_L$ at $p_0 = \frac{u}{u+1}$, $U_H \neq U_L$ for any $p_1 \in \left(\frac{u}{u+1}, q_2^H\right)$ (Assumption 2), by continuity of $U_H$ and $U_L$, $U_H < U_L$ for any $p_1 \in \left(\frac{u}{u+1}, q_2^H\right)$.

   The two paragraphs above imply that $U_H > U_L$ when $p_1 \in \left(q_1^H, \frac{u}{u+1}\right)$ and $U_H < U_L$ when $p_1 \in \left(\frac{u}{u+1}, q_2^L\right)$.

### 5.6.2 When Assumption 1 or 2 fails

Suppose that Assumption 1 fails. This means that Assumption 1a also fails. If $q_k^L \neq q_k^H$ for $k = 1, 2$, and the ranking of thresholds $q$ satisfies $q_1^L < q_1^H < \frac{u}{u+1}$ or $\frac{u}{u+1} < q_2^L < q_2^H$. By Proposition 1, this ranking implies that $U_L > U_H$ at $p_0 = q_1^H$ or $U_H > U_L$ at $p_0 = q_2^L$, which is a violation of condition ($\triangle$) in Theorem 1. When $q_k^L = q_k^H$ for $k = 1$ or 2, $U_L = U_H$ at $p_0 = q_1^H$ or $U_H = U_L$ at $p_0 = q_2^L$. Condition ($\triangle$) may or may not hold.

Suppose that Assumption 2 fails and there exist one or multiple $p_0^{'} \in \left(\max\left\{q_1^L, q_1^H\right\}, \min\left\{q_2^L, q_2^H\right\}\right)$ such that $p_0^{'} \neq \frac{u}{u+1}$ but $U_L = U_H$. Then, it is possible (but not inevitable) that $\left(q_1^H, \frac{u}{u+1}\right)$ and $\left(\frac{u}{u+1}, q_2^L\right)$ are both divided into sub-intervals where $U_L - U_H$ is positive on some sub-intervals but negative on others, as illustrated in Example 5.

## 5.7 Proof of Remark 1 (prior-conforming articles first-order stochastically dominate prior-contradicting articles in accuracy)

Given $p_0$, let $g_L$ and $G_L$ be the expected p.d.f. and c.d.f. of $|s|$ for $s < 0$. Let $g_H$ and $G_H$ be the expected p.d.f. and c.d.f. of $|s|$ for $s > 0$. Then,

$$\begin{aligned} g_L(|s|) &= p_0 f_{\omega_H}(-|s| \mid s<0) + (1-p_0) f_{\omega_L}(-|s| \mid s<0) \\ &= \frac{p_0}{F_H(0)} \cdot f_{\omega_H}(-|s|) + \frac{1-p_0}{F_L(0)} \cdot f_{\omega_L}(-|s|) \\ &= \frac{p_0}{F_H(0)} \cdot f_{\omega_L}(|s|) + \frac{1-p_0}{F_L(0)} \cdot f_{\omega_H}(|s|) \end{aligned}$$

$$G_L(|s|) = \frac{p_0 \left[F_{\omega_L}(|s|) - F_{\omega_L}(0)\right]}{F_{\omega_H}(0)} + \frac{(1-p_0)\left[F_{\omega_H}(|s|) - F_{\omega_H}(0)\right]}{F_{\omega_L}(0)}.$$

$$\begin{aligned} g_H(|s|) &= p_0 f_{\omega_H}(|s| \mid s>0) + (1-p_0) f_{\omega_L}(|s| \mid s>0) \\ &= \frac{p_0}{1-F_{\omega_H}(0)} \cdot f_{\omega_H}(|s|) + \frac{1-p_0}{1-F_{\omega_L}(0)} \cdot f_{\omega_L}(|s|) \end{aligned}$$

$$G_H(|s|) = \frac{p_0 \left[F_{\omega_H}(|s|) - F_{\omega_H}(0)\right]}{1-F_{\omega_H}(0)} + \frac{(1-p_0)\left[F_{\omega_L}(|s|) - F_{\omega_L}(0)\right]}{1-F_{\omega_L}(0)}.$$

Because $F_{\omega_H}(0) = 1 - F_{\omega_L}(0)$ and $F_{\omega_L}(0) = 1 - F_{\omega_H}(0)$,

$$G_L(|s|) = \frac{p_0 \left[F_{\omega_L}(|s|) - F_{\omega_L}(0)\right]}{1-F_{\omega_L}(0)} + \frac{(1-p_0)\left[F_{\omega_H}(|s|) - F_{\omega_H}(0)\right]}{1-F_{\omega_H}(0)}.$$

$$G_H(|s|) - G_L(|s|) = \frac{(1-2p_0)\left[F_{\omega_L}(|s|) - F_{\omega_L}(0)\right]}{1-F_{\omega_L}(0)} + \frac{(2p_0-1)\left[F_{\omega_H}(|s|) - F_{\omega_H}(0)\right]}{1-F_{\omega_H}(0)}.$$

When $p_0 < \frac{1}{2}$, $1 - 2p_0 > 0$. When $p_0 > \frac{1}{2}$, $1 - 2p_0 < 0$. Therefore, if

$$\frac{F_{\omega_L}(|s|) - F_{\omega_L}(0)}{1 - F_{\omega_L}(0)} > \frac{F_{\omega_H}(|s|) - F_{\omega_H}(0)}{1 - F_{\omega_H}(0)} \qquad (\ast)$$

then the following statement in the Remark is true:
if $p_0 < \frac{1}{2}$, $G_H(|s|) > G_L(|s|)$ for all $|s| > 0$;
if $p_0 > \frac{1}{2}$, $G_H(|s|) < G_L(|s|)$ for all $|s| > 0$.
The remaining steps prove $(\ast)$.[9]
Because $f$ has the monotone likelihood ratio property, for all $|s_1| > |s_2| > 0$,

$$\frac{f_{\omega_H}(|s_1|)}{f_{\omega_L}(|s_1|)} > \frac{f_{\omega_H}(|s_2|)}{f_{\omega_L}(|s_2|)}$$

or

$$f_{\omega_H}(|s_1|)\, f_{\omega_L}(|s_2|) > f_{\omega_H}(|s_2|)\, f_{\omega_L}(|s_1|) \quad \cdots\cdots \quad \text{(A)}$$

Integrate (or sum up, if $f$ is discrete) inequality (A) in two different ways. Firstly, integrate both sides with respect to $|s_2|$ from 0 to $|s_1|$ to get

$$\begin{aligned} f_{\omega_H}(|s_1|)\left[F_{\omega_L}(|s_1|) - F_{\omega_L}(0)\right] &> f_{\omega_L}(|s_1|)\left[F_{\omega_H}(|s_1|) - F_{\omega_H}(0)\right] \\ \frac{f_{\omega_H}(|s_1|)}{f_{\omega_L}(|s_1|)} &> \frac{F_{\omega_H}(|s_1|) - F_{\omega_H}(0)}{F_{\omega_L}(|s_1|) - F_{\omega_L}(0)} \end{aligned}$$

Because $|s_1|$ can be any positive number, for any $|s| > 0$,

$$\frac{f_{\omega_H}(|s|)}{f_{\omega_L}(|s|)} > \frac{F_{\omega_H}(|s|) - F_{\omega_H}(0)}{F_{\omega_L}(|s|) - F_{\omega_L}(0)} \quad \cdots\cdots \quad \text{(B)}$$

Secondly, integrate both sides of (A) with respect to $|s_1|$ from $|s_2|$ to $\infty$ to get

$$\begin{aligned} \left[1 - F_{\omega_H}(|s_2|)\right] f_{\omega_L}(|s_2|) &> f_{\omega_H}(|s_2|)\left[1 - F_{\omega_L}(|s_2|)\right] \\ \frac{f_{\omega_H}(|s_2|)}{f_{\omega_L}(|s_2|)} &< \frac{1 - F_{\omega_H}(|s_2|)}{1 - F_{\omega_L}(|s_2|)} \end{aligned}$$

[9] Source of proof: https://en.wikipedia.org/wiki/Monotone_likelihood_ratio

Because $|s_2|$ can be any positive number, for any $|s| > 0$,

$$\frac{f_{\omega_H}(|s|)}{f_{\omega_L}(|s|)} < \frac{1 - F_{\omega_H}(|s|)}{1 - F_{\omega_L}(|s|)} \quad \cdots\cdots \quad \text{(C)}$$

(B) and (C) imply that for any $|s| > 0$,

$$\frac{1 - F_{\omega_H}(|s|)}{1 - F_{\omega_L}(|s|)} > \frac{F_{\omega_H}(|s|) - F_{\omega_H}(0)}{F_{\omega_L}(|s|) - F_{\omega_L}(0)}$$

which is equivalent to $(\ast)$. This concludes the proof.

## 5.8 Proof of Proposition 2 (platform's optimal strategy)

Suppose that Assumptions 1 and 2 hold. By Lemma 4, Proposition 1, and Theorem 1, $0 \le q_1^H < q_1^L < \frac{u}{u+1} < q_2^H < q_2^L \le 1$ and

$$\begin{cases} \Delta U_L = \Delta U_H = 0 & \text{if } p_1 \in \left[0, q_1^H\right] \\ \Delta U_H > \Delta U_L = 0 & \text{if } p_1 \in \left(q_1^H, q_1^L\right] \\ \Delta U_H > \Delta U_L > 0 & \text{if } p_1 \in \left(q_1^L, \frac{u}{u+1}\right) \\ \Delta U_L = \Delta U_H > 0 & \text{if } p_1 = \frac{u}{u+1} \\ \Delta U_L > \Delta U_H > 0 & \text{if } p_1 \in \left(\frac{u}{u+1}, q_2^H\right) \\ \Delta U_L > \Delta U_H = 0 & \text{if } p_1 \in \left[q_2^H, q_2^L\right) \\ \Delta U_L = \Delta U_H = 0 & \text{if } p_1 \in \left[q_2^L, 1\right] \end{cases}$$

Moreover, for $i = L, H$, $\Delta U_i$ is strictly increasing on $\left(q_1^i, \frac{u}{u+1}\right)$ and strictly decreasing on $\left(\frac{u}{u+1}, q_2^i\right)$. These characteristics of $\Delta U_L$ and $\Delta U_H$ are visualized by Figure 1.

The user clicks on title $t_i$ if and only if $\Delta U_i \ge 0$. When $p_0 \le \frac{u}{u+1}$, she is willing to click on both $t_H$ and $t_L$ if $c \le \Delta U_L(p_0)$, neither $t_H$ nor $t_L$ if $c > \Delta U_H(p_0)$, and only $t_H$ if $c \in (\Delta U_L(p_0), \Delta U_H(p_0)]$. When $p_0 \ge \frac{u}{u+1}$, she is willing to click on both $t_H$ and $t_L$ if $c \le \Delta U_H(p_0)$, neither $t_H$ nor $t_L$ if $c > \Delta U_L(p_0)$, and only $t_H$ if $c \in (\Delta U_H(p_0), \Delta U_L(p_0)]$.

Therefore, for a click-maximizing platform, it is weakly dominant to recommend $t_H$ whenever $p_0 < \frac{u}{u+1}$ and $t_L$ whenever $p_0 \ge \frac{u}{u+1}$. If $p_0 < \frac{u}{u+1}$ and $c \in (\Delta U_L(p_0), \Delta U_H(p_0)]$ with positive probability, it strictly prefers to recommend $t_H$. If $p_0 > \frac{u}{u+1}$ and $c \in (\Delta U_H(p_0), \Delta U_L(p_0)]$ with positive probability, it strictly prefers to recommend $t_L$. This proves the Proposition.

# References


Allcott, H., L. Braghieri, S. Eichmeyer, and M. Gentzkow (2020). The welfare effects of social media. *American Economic Review 110*(3), 629–676.

Allon, G., K. Drakopoulos, and V. Manshadi (2021, November-December). Information inundation on platforms and implications. *Operations Research 69*(6), ii–iv, 1651–1959, C2–C3.

Bakshy, E., S. Messing, and L. A. Adamic (2015). Exposure to ideologically diverse news and opinion on facebook. *Science 348*(6239).

Beam, M. A., M. J. Hutchens, and J. D. Hmielowski (2018). Facebook news and (de)polarization: reinforcing spirals in the 2016 us election. *Information, Communication and Society 21*(7), 940–958.

Bessi, A., F. Zollo, M. D. Vicario, M. Puliga, A. Scala, G. Caldarelli, B. Uzzi, and W. Quattrociocchi (2016). Users polarization on facebook and youtube. *PLoS ONE 11*(8).

Burke, J. (2008, Fall). Primetime spin: Media bias and belief confirming information. *Journal of Economics and Management Strategy 17*(3), 633–665.

Calvert, R. L. (1985, June). The value of biased information: A rational choice model of political advice. *The Journal of Politics 47*(2), 530–555.

Charness, G., R. Oprea, and S. Yuksel (2021). How do people choose between biased information sources? evidence from a laboratory experiment. *Journal of the European Economic Association 19*(3), 1656–1691.

Che, Y.-K. and K. Mierendorff (2019). Optimal dynamic allocation of attention. *American Economic Review 109*(8), 2993–3029.

Dubois, E. and G. Blank (2018). The echo chamber is overstated: the moderating effect of political interest and diverse media. *Information, Communication and Society 21*(5), 729–745.

Flaxman, S., S. Goel, and J. M. Rao (2016). Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly 80*(Special Issue), 298–320.

Fletcher, R., A. Kalogeropoulos, and R. K. Nielsen (2021). More diverse, more politically varied: How social media, search engines and aggregators shape news repertoires in the united kingdom. *New Media and Society 1*(22).

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

Gentzkow, M. and J. M. Shapiro (2010). What drives media slant? evidence from u.s. daily newspapers. *Econometrica 78*(1), 35–71.

Gentzkow, M., J. M. Shapiro, and M. Sinkinson (2014). Competition and ideological diversity: Historical evidence from us newspapers. *American Economic Review 104*(10), 3073–3114.

Guilbeaulta, D., J. Beckera, and D. Centola (2018, September). Social learning and partisan bias in the interpretation of climate trends. *Proceedings of the National Academy of Sciences 115*(39).

Kamenica, E. and M. Gentzkow (2011). Bayesian persuasion. *American Economic Review 101*(October), 2590–2615.

Kranton, R. and D. McAdams (2021, October). Social connectedness and the market for news. *Working paper*.

Levy, R. (2021). Social media, news consumption, and polarization: Evidence from a field experiment. *American Economic Review 111*(3), 831–870.

Pariser, E. (2011). *The Filter Bubble: What The Internet Is Hiding From You.* The Penguin Group.

Quattrociocchi, W., A. Scala, and C. R. Sunstein (2016, June). Echo chambers on facebook. *Working paper available at SSRN: https://ssrn.com/abstract=2795110*.

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

Sunstein, C. R. (2001a). *Echo Chambers: Bush v. Gore, Impeachment, and Beyond.* Princeton University Press.

Sunstein, C. R. (2001b). *Republic.com.* Princeton University Press.

Sunstein, C. R. (2007). *Republic.com 2.0.* Princeton University Press.