Identity-Based Elections

Helios Herrera\(^*\) and Ravideep Sethi\(^{†‡}\)

April 19, 2022

Abstract

We study how media choice driven by political identity can influence electoral results. Citizens gather information from mainstream news but also from partisan media sources which filter news in particular predetermined ways. We assume that citizens process all information they receive correctly but choose their own media sources in a behavioral self-serving way to try to preserve their political faith/identity. That is, they attempt to rationally counteract mainstream news that they might view as unfavorable. In the baseline setup, we assume that citizens on either side of the political spectrum are exposed to different extents to non-partisan mainstream news, as in the U.S. case. This endogenous media choice generates an electoral advantage for the less exposed side, which can turn into a sure electoral victory even for the wrong candidate in a democracy. Results are robust to forms of media distrust and are stronger if citizens have biased priors. In illiberal democracies, where the government controls the media, official media propaganda works only if citizens are unaware of its bias or if the government can engage in censorship. Propaganda backfires in the presence of freely available chosen media.

JEL codes: D72, D83, D9

Keywords: Information design, Belief-based utility, Information Aggregation

\(^*\)University of Warwick and CEPR. Email: h.herrera@warwick.ac.uk

\(^†\)David Eccles School of Business, University of Utah. Email: ravi@eccles.utah.edu

\(^‡\)We thank seminar participants at UC Berkeley, Caltech, Princeton, Harvard, Georgetown, Rutgers, SMU, Vanderbilt, University of Utah, LSE, LBS, PSE, QMUL, University of Venice, CREST, Ashoka University, and the Delhi School of Economics for useful comments and suggestions. All mistakes are our own.
1 Introduction

“More often than not, citizens do not choose which party to support based on policy opinion; they alter their policy opinion according to which party they support.”

Mason (2018)

In the current environment, people can choose to consume news from a plethora of possible media sources. Although traditional mainstream news sources continue to exist, many new media sources have emerged in just the past two decades. The richness of new media has perhaps benefited people by allowing them to more accurately tailor their media choices to their wants. However, the incredible diversity of viewpoints on offer, combined with new technologies, has made it easier for citizens to form “echo chambers” or “filter bubbles” through which they become insulated from possibly contrary perspectives offered by traditional media outlets.

At the same time, trust in traditional media has declined markedly over the past two decades. Especially in the past five years, this media distrust has followed radically different paths on either side of the political spectrum in the United States. As we can see in figure 1 below, the gap in media trust between Republicans and Democrats is staggering.

Figure 1: Asymmetry in trust in mass media

Indeed, as noted by the Pew Foundation in Jurkowitz et al. (2020), “one of the clearest differences between Americans on opposing sides of the political aisle is that large portions of Democrats express trust in a far greater number of news sources.” For instance, during the 2016 US presidential campaign, according to Bond (2017) “Pro-Clinton audiences were highly attentive to traditional media outlets, which continued to be the most prominent outlets across the public sphere, alongside more left-oriented online sites. But
pro-Trump audiences paid the majority of their attention to polarized outlets that have
developed recently, many of them only since the 2008 election season. [...] Breitbart News
became the center of a distinct right-wing media ecosystem, surrounded by Fox News,
the Daily Caller, the Gateway Pundit, the Washington Examiner, Infowars, Conservative
Treehouse, and Truthfeed.1

These two phenomena — the emergence of a dense array of media outlets and partisan
distrust of media — have repercussions on how political beliefs are formed and updated
and, as a consequence, on people’s decisions on election day. But can this new informa-
tion environment generate aggregate beliefs biased enough to swing an election? We seek
to shed light on how this new asymmetric media landscape shapes electoral outcomes.

The influence of the above-mentioned phenomena on aggregate electoral outcomes
is compounded, especially in the U.S., by the presence of a very polarized landscape in
which traditional ideological, religious, and racial identities are being replaced by over-
lapping meta-identities captured almost entirely by the Democratic and Republican polit-
ical faiths. Citizens have become less responsive to new info or real national problems,
as if political affiliations determine what information people absorb, rather than the other
way around (see, for instance, Mason (2018)).

Taking heed from the above issues, we build a model where political identities drive
individual media choices. To model media choice, our basic premise is a dichotomy in
types of media. This premise, in addition, allows us to study the case of both liberal and
illiberal democracies under the same framework: we assume that each citizen chooses to
follow certain specific media, but that to some extent she is also exposed to the broader
media environment — to media she does not explicitly choose. Each citizen is aware of the
bias of the media sources she decides to follow and tailors that choice to her preferences:
we denote these as Inside media (or In-media, henceforth). However, she is aware she
is exposed somewhat to the ‘outside world’, i.e. to media not chosen or non voluntarily
consumed (Outside media or Out-media, henceforth).

Different assumptions on Outside media allow us to tackle different questions. In the
context of a liberal democracy, as the U.S., we can think of Outside media as mainstream
media,2 which citizens trust to possibly different extents, implying different exposure

---

1 Lee (2010) finds that trust in media is negatively correlated with conservatism and Republican-leaning.
Pennycook & Rand (2019) notes that Democrats trust mainstream news more than Republicans, with the
difference ranging between 11.5 and 14.7 percent. According to Jones (2004), “only 16.5 percent of Democrats
(including Democratic-leaning independents) can be classified as media skeptics compared with nearly 40 percent of
Republicans and Republican-leaning independents”.

2 To be sure, it is hard to define mainstream media in the current rapidly evolving changing media
environment. Broadly, however, when we use the term mainstream, we are referring to traditional corporate
legacy outlets; these have standard fact-checking processes and try to synthesize and diffuse responsibly
and, as a consequence, different impacts on their beliefs. In the context of an illiberal democracy, we will refer to Outside media as official government media that promotes a government propaganda message some, or most, of the time. We describe the U.S. case first and later focus on propaganda in an illiberal democracy.

The core setup of the model is decision-theoretic and very simple. Citizens receive two pieces of news: one from exogenous Outside media and one from Inside media, i.e., a media source endogenously chosen by the citizen. Different exposure to Outside information means that Outside information is, in our benchmark setup, noisier for certain citizens and thus affects their beliefs to a lesser extent. The In-media source is chosen in a self-serving behavioral way, as explained below. Citizens are fully rational in the way they process all information they receive and update their beliefs based on the two signals and vote according to their posterior for the better candidate. In sum, the election aggregates all votes, each based on two conditionally independent signals of which of two candidates is preferable.

The key behavioral assumption of our model regards not information processing but the preference that drives each citizen’s choice of In-media, i.e., the tailored media outlets, each of which we view as a particular known signal structure (under commitment). We assume that each citizen identifies with a party, on the left or the right, and aims to preserve their political identity/faith. They choose In-media to maximize the likelihood they will believe the party they identify with is the better match for the state of the world, after their beliefs are rationally updated. Such beliefs are based on just two signals: the exogenous Out-media signal and the signal of the In-media outlet of their choice. In other words, citizens make their media choice attempting to shield themselves from possibly unfavorable (from the point of view of their political affiliation) outside news to which they are somewhat exposed. Equivalently, one can think of agents as having two selves — a heart and a mind. The heart chooses the media to follow attempting to preserve its political identity/faith, whereas the mind processes all information it receives rationally and votes for the party she believes is better. Our setup allows for several interpretations. We may view In-media readership as a long-term choice that people make, and not one that is instrumental to a particular vote in a specific upcoming election (so the utility is not derived from the voting action but from sheer ex-post beliefs). Alternatively, the choice of Inside media may be thought of as one made instrumentally in the context of a particular vote, in which case we can think of agents deriving explicit utility from the

---

3Distrust may also mean that distrusting citizens believe Outside media is biased (whether or not it truly is biased). We tackle this question in the second part of the paper on (true or perceived) propaganda.
action of voting for the party aligned with their identity or not.

To highlight the electoral consequences of different exposure to mainstream media along partisan lines, we present our benchmark results for a setup in which the two sides are perfectly symmetric, except for how right- and left-affiliated citizens are exposed to information from mainstream media.

An asymmetry in the exposure of agents to mainstream information (Out-media) implies that the type of media they actively choose to consume (In-media) can be drastically different. This difference in media choice has surprising implications for electoral outcomes. In the example below, we study the size of the electoral advantage (winning margin) of the side less exposed to mainstream media assuming that each citizen votes for the party she rationally believes to be superior. Our main results are qualitatively unchanged if we assume that each citizen votes for her culturally affiliated party only if she believes it is better, and abstains otherwise, namely if we assume turnout/abstention margins determine electoral outcomes. In this case, all winning margins would simply be halved.¹

**Illustrative example:** Assume a symmetric benchmark in which a large number of partisans of either side, left $L$ and right $R$, are equal in number. There are two equally likely states of the world, $\omega = L, R$, denoting which of two candidates is the better one. The only asymmetry between the two sides is that the left is more exposed to mainstream news than the right. Assume, for instance, that the left-wing citizens receive i.i.d. symmetric binary signals from mainstream news with precision $t_L = 0.75$, while right-wing citizens receive noisier mainstream signals, namely with lower precision $t_R = 0.51$. As a baseline, we consider electoral outcomes in the absence of Inside media choice — that is, when the agent is only exposed to an Outside signal. The winning margin and winning probability for the R side are:

<table>
<thead>
<tr>
<th></th>
<th>$\omega = R$</th>
<th>$\omega = L$</th>
</tr>
</thead>
<tbody>
<tr>
<td>R Win Margin</td>
<td>0%</td>
<td>+26%</td>
</tr>
<tr>
<td>R Win Prob</td>
<td>50%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Thus, asymmetric exposure to mainstream media generates symmetric electoral outcomes. In this baseline case, the ideal candidate is always elected, i.e., information is perfectly aggregated. No personal media choice is made by citizens, and thus Political Faith, $R$ or $L$, plays no role.

¹We assume infinite population size so being pivotal is not an issue in our model.
Now, suppose that citizens can also curate Inside media sources optimally. Here, their voting decision is made after updating rationally on two signals, not one. If the media is chosen to maximize the chance of political faith preservation, then the outcome of the election is no longer symmetric. In fact, it may be drastically skewed. In this example, the winning margin and winning probability for the R side are:

<table>
<thead>
<tr>
<th></th>
<th>Ex-Ante</th>
<th>(\omega = R)</th>
<th>(\omega = L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R Win Margin</td>
<td>+28%</td>
<td>+54%</td>
<td>+2%</td>
</tr>
<tr>
<td>R Win Prob</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

In this case, the R side has an ex-ante winning margin advantage in the election, but, surprisingly, it also has an advantage ex-post. Namely, the R side wins the election in either state of the world in this example and information is not aggregated, despite agents voting based solely on (rational updating of) the information they received.

The key drivers of the vote aggregation result above are based on two key features of citizens’ media choice. First, because media choice is based on preserving, as often as possible, political identity, it is equivalent, in the aggregate, to maximizing the ex-ante winning margin of the party with which each citizen identifies. However, maximizing the ex-ante winning margin differs from maximizing the probability of their preferred party winning. The latter is more related to the ex-post winning margin conditional on the realized state.

Second, in a certain parameter range, the media choices made by citizens on either side are not only quantitatively but also, crucially, qualitatively different. The choice of the less-exposed side (the right in our example) is a one-sided signal structure, for which news favorable to one candidate are very frequent and thus not so informative, while unfavorable news are rare and hence damning. This is similar to a partisan outlet. Whereas the choice of the more exposed side (the left in our example) resembles more balanced news, which can be mixed, favorable or unfavorable, in either state.

Importantly, in a world without a rich set of signal structures (In-media) available to agents to select from, even if partisan biases still drove media choice, we would not see such a stark electoral aggregate bias as in the example above.

Failure of information aggregation occurs when the two sides are symmetric in every respect other than the fact that one side has average exposure to mainstream media and the other has little exposure to it. As we show, party \(R\) would also win regardless of the state even if there were slightly more agents of type \(L\) than \(R\), or when the prior, \(w\), leans towards the \(L\) side. Our symmetric common priors benchmark assumption stacks the
cards against us. If agents on the two sides have heterogeneous priors biased towards
their side, then the effect is much stronger, and the scope for the failure of information
aggregation expands. In our benchmark model the bias is in the preferences and is non-
cognitive, but the model is malleable to additional cognitive behavioral assumptions. Re-
results do not change in a misspecified model in which the R side distrusts the mainstream
signal, that is, believes that the mainstream signal is noisier than it actually is.

Our modeling strategy based on the dichotomy between Inside media and Outside
media can also be used to explore other key questions. We initially assume that the Out-
side media news is unbiased, but assuming it is biased allows us to think about a model
of government propaganda.

In the second part of the paper, we imagine a propaganda regime. Citizens know that
the information from the official media is biased and pro-government. More specifically,
the official media delivers an unbiased message only with some probability and a pro-
government message otherwise. That is, the government-influenced media has limited
commitment to the truthful reporting of unbiased news and this is common knowledge.
Further, citizens have limited exposure/attention to the government message.

In a media censorship regime, the official media is the only media available. In this
benchmark, information is aggregated correctly only if the bias is low. For a large enough
bias, the propaganda works, and the government wins an election regardless of the state.
Without censorship, however, citizens, who support or oppose the government, can also
obtain information from sources of their choice. Assuming that citizens make this choice
to try to preserve their political faith, then the electoral outcome is starkly different. First,
the propagandizing party does not win in all states in any region of the parameter space.
Second, if the government media bias is high enough, the opposing side can win regard-
less of the state. Thus, propaganda backfires without censorship of non-governmental
media. When there is a failure of information aggregation, it is in favor of the anti-
government side. This full reversion is surprising and occurs because of the qualitatively
different Inside media chosen by citizens on either political side. This suggests why crack-
downs of free media are key in authoritarian states for government propaganda to work
when citizens are aware of it.

The general setup can be adapted to explore salient misspecified versions of this model,
in which some of the citizens also have a cognitive bias. If some citizens behaviorally mis-
perceive the actual distribution of the Out-Media signal, then their In-Media choice may
be suboptimal individually. The case in which the official government Out-media signal
is biased, but pro-government citizens (mistakenly) believe it is unbiased is of interest. As
we show, without censorship of non-governmental media, propaganda can work in swing-
ing elections only if citizens are unaware of it and wrongly believe the government media is unbiased.

Lastly, the cognitive bias case in which the Out-media signal is unbiased, but citizens from one side (mistakenly) believe it is biased against their side, is of particular interest as it may arguably apply to some liberal democracies, such as the US. In this case we show the electoral outcomes are similar if we interpret partisan distrust of mainstream media as (wrongly) believing Outside media is biased adversely, or (wrongly) believing it is non-informative (noisy).  

2 Related Literature

This paper considers the electoral implications of behavioral biases in information acquisition following in the tradition of Levy & Razin (2015) and Ortoleva & Snowberg (2015). In both these papers, the bias considered is correlation neglect — agents underestimate the correlation between their information sources. The spirit of our exercise is similar: we use a simple decision-theoretic problem aggregated to derive expected electoral outcomes. However, our paper is fundamentally different because the behavioral bias in our benchmark case is not cognitive but resides in the preferences that drive information collection. Specifically, agents update rationally all the information that they receive in our model.

Though cast in a different context, our decision model has a similar flavor to the Köszegi (2006) model of overconfidence, in that agents are unbiased in their beliefs (as they start from a correct prior and update rationally) but end up with a systematic bias in their choice due to a bias in their information collection process. Intuitively, given that agents derive intrinsic utility from believing that “something” is the case (specifically, that they have superior ability in some task, as in Köszegi’s case, or that their party is the better one, as in our case) then they voluntarily tend to collect information that helps them believe that, as often as rationally possible. The tradition of agents deriving utility from their beliefs goes back to Akerlof & Dickens (1982), which incorporates beliefs explicitly in the decision maker’s utility function. In that paper, beliefs are a choice variable, whereas, in our paper, beliefs are a stochastic outcome of choosing a particular signal structure.

The partisan asymmetry in trust in mainstream media has, in recent years, been a well documented and robust phenomenon that has only been exacerbated by Donald Trump’s labeling of news from the mainstream as “Fake News” and the mainstream media itself as “Enemy of the People.”

For an excellent survey of the recent literature on the electoral outcomes of these and similar cognitive biases, see Levy & Razin (2019). This literature keeps growing, for instance, see Little et al. (2020) on motivated reasoning cognitive bias.

5The partisan asymmetry in trust in mainstream media has, in recent years, been a well documented and robust phenomenon that has only been exacerbated by Donald Trump’s labeling of news from the mainstream as “Fake News” and the mainstream media itself as “Enemy of the People.”

6For an excellent survey of the recent literature on the electoral outcomes of these and similar cognitive biases, see Levy & Razin (2019). This literature keeps growing, for instance, see Little et al. (2020) on motivated reasoning cognitive bias.
Bénabou & Tirole (2016) provide a recent survey on the main findings that emanate from economic models of motivated beliefs.

The objective to preserve political faith is also consistent with findings in an extensive psychology literature. Sherman & Cohen (2006) note that biases in the assimilation of information can come from the motivation to maintain and protect political identity. Identity preserving motivated reasoning in politics is a pervasive phenomenon that is also noted in Kahan (2017) and Kahan & Braman (2006). The ability to explicitly deliberate does not ameliorate this issue — Kahan (2012) finds that cognitive reflection exacerbates ideologically motivated reasoning.

We borrow the theoretical techniques of Bayesian persuasion and use them to study electoral outcomes: conditionally aggregating individual decisions is our focus. In order to do so, we adapt and distill results developed in Kolotilin (2018), who builds on Kamienica & Gentzkow (2011) by requiring the sender to choose an information structure while being uncertain about the receiver’s type, which is similar to our agent’s problem of choosing an information structure while bracing for an Outside signal. Lipnowski & Mathevet (2018) considers an information design problem with a benevolent sender choosing a signal structure for a receiver with psychological preferences; this approach relates to our model of information choice by agents who derive belief-based utility.

Our model assumes that citizens can choose their In-media from a dense distribution of sources, spanning all possible biases; thus, our media are passive and non-strategic, and media consumption is demand-driven only. Motivated by the fast-changing media landscape, there is a burgeoning literature inspecting media bias, albeit usually from the supply side, where media are strategic. Gitmez & Molavi (2022) consider the problem of a single strategic media source that aims to induce a heterogeneous population of receivers to support its preferred policy. They find that polarization among receivers can make the media source less biased. In Perego & Yuksel (2022), growing competition between information providers leads to news specialization, which thus amplifies social disagreement — a result with a similar flavor to ours. Gentzkow & Shapiro (2006) find that media outlets slant their reports to conform to prior beliefs held by agents and also find that media bias reduces in the presence of Outside information. In our paper, the presence of asymmetrically precise Outside information is one of the settings that can cause failures of information aggregation. Gentzkow et al. (2021) explain that belief divergence among agents seeking to learn a sequence of states results from small ideological differences in
agents’ trust in information sources and beliefs about the state of the world. A strategic media that biases its news to induce greater viewership by citizens may increase polarization, as in Bernhardt et al. (2008). Their results hold even though citizens are aware of the media bias and update rationally. In Mullainathan & Shleifer (2005), agents demand different news outlets depending on their desire to confirm their pre-existing biases and finds that media strategically offer slanted news — an effect exacerbated with more competition among media.

On the empirical front, our voting results are consistent with several studies that show how the introduction of new (and largely conservative) media outlets resulted in a persistent and significant increase in Republican vote share. DellaVigna & Kaplan (2007) and Martin & Yurukoglu (2017). Our model shows that wrong electoral outcomes may emerge without assuming fake news, which we do not model.\footnote{As Angelucci & Prat (2021) show, belief in fake news, while present, does not have a systematic partisan bias.} We describe the workhorse model in the next section. Then, we apply it to liberal democracies. We examine illiberal democracies last.

3 Model

3.1 Setup

There are countably infinite agents, and each is one of two types (R and L), which correspond to political identities. There are two states of the world (R and L) and two political sides/parties (R and L). All agents share a common and symmetric prior regarding the state of the world (P[ω = R] := w = 0.5).\footnote{As we show in subsection 4.3, the key results of the model are robust to biased and non-common priors.}

Each agent receives two signals about the state of the world: an Inside and Outside signal. The Inside signal is generated from a chosen signal structure, while the signal structure of the Outside signal is exogenously specified. We interpret the Inside signal as being generated from the media that the agent chooses to consume. The agent does not choose the Outside signal structure — that is simply a feature of the media environment that the agent lives in. We suppose that the kind of outside information that the agent is hit with and must brace for depends on the structure of the mainstream media which may or may not be biased. Further, the agent may be imperfectly exposed to the Outside signal, or may hold incorrect beliefs about it.

The timing of the game is as follows. First, each agent of either type chooses her
Inside signal structure. Second, the agent receives Inside and Outside signals. Third, agents form Bayesian posteriors and realize belief-based utility. An agent gains utility from holding a posterior belief that the state of the world more likely matches her type. Finally, agents vote sincerely. Because there is a countable infinity of agents, none of them is ever pivotal. Voting, is therefore, non-strategic. Moreover, the action of voting does not impact the agent’s utility, which is based on the agent’s beliefs.

We can specify an equivalent model in which actions are strategic. Suppose that each agent has two selves, a heart (sender) and a mind (receiver). The sender-self is of type $R$ or $L$ and chooses the In-media signal structure. The receiver-self votes sincerely based on her updated beliefs. The receiver-self forms her beliefs rationally after observing both the Inside and Outside signals. The objective of a type-$R$ ($L$) sender-self is to maximize the likelihood that the receiver-self votes for party $R$ ($L$). This alternative specification may be thought of as another interpretation of the agent’s problem. In the rest of this paper, we study the model described in the previous paragraph, in which each agent has a single self that has belief-based utility.

We show in appendix A5 all results in this model are robust to considering an additive function where the agent has a relatively small gain from voting for the correct party. This is because for low values of gain from being correct, the set of signal structures that solve the agent’s problem remain unchanged from the benchmark. If there is a region of information aggregation failure present in the benchmark, then, while that region might shrink, it does not disappear if we include a small gain for the agent from voting for the correct party.

Further, we show in appendix A4 that all results remain unchanged if the agent gains linearly from holding posteriors favorable to her party, in addition to the gain from political faith preservation. This is because the choice if Inside signal structure is invariant to the gain from favorable posteriors, and all electoral outcomes are determined by the choice of Inside signal structures. This robustness holds for all levels of that gain in the modified utility function.

Because the game is symmetric, we specify the problem for a type-$R$ agent; the speci-

---

10Both versions of the model described so far suppose that agents vote sincerely. Alternatively, we could allow agents to abstain by supposing that agents only vote if they can do so sincerely for their preferred party. If they believe that the state more likely matches the other party, they abstain. Under this specification, the results remain the same qualitatively — only the winning margins are halved.
fication for a type-\(L\) agent is analogous. The utility function of an agent of type \(R\) is:

\[
U_R = \begin{cases} 
1 & \text{if } \mathbb{P}[\omega = R|S, s] \geq 0.5 \\
0 & \text{otherwise}
\end{cases}
\]

Where \(S\) is the signal generated by the Inside signal structure, while \(s\) is the Outside signal. For the agent, maximizing her expected utility is equivalent to maximizing the likelihood that she holds a posterior in which she believes that her type more likely matches the state of the world. We call this form of a self-serving bias \textit{Political Faith Preservation}. The agent’s objective can, therefore, be more simply expressed as:

\[
\mathbb{E}[U_R] = \mathbb{P}
\begin{bmatrix}
\mathbb{P}[\omega = R|S, s] \geq 0.5
\end{bmatrix}
\]

The agent’s objective is to choose the signal structure that generates her Inside signal to maximize equation [1]. The action set available to the agent is the set of Inside signal structures specified in the paragraph below. We assume that in addition to choosing the Inside signal structure, the agent also takes the non-strategic action of sincere voting.\[11\] We can alternatively assume that the agent abstains if her beliefs do not allow her to sincerely vote for her preferred party. In each case, we get qualitatively the same results — the winning margins are simply halved in the latter case.

We are interested in electoral outcomes, particularly in situations where the \textit{incorrect} party wins. This happens if party \(R\) wins in state \(L\) or vice versa. We call these situations information aggregation failures. We do not link information misaggregation to welfare loss because, in our model, the agents gain no utility from voting or electoral outcomes.\[12\]

\textbf{Inside signal structure.} The Inside signal structure comprises a finite set of signals \((S = \{S_1, S_2, ..., S_n\})\) that are correlated with the state of the world. The choice of signal structure must abide by the Martingale property:

\[
\sum_{i=1}^{n} \mathbb{P}[S_i] \cdot \mathbb{P}[\omega = R|S_i] = \mathbb{P}[\omega = R] = 0.5
\]

Where \(S_i\) is a signal \(i\) that has the probability \(\mathbb{P}[S_i]\) of being generated given the signal

---

\[11\]In the case of a tie, she votes for her preferred party.

\[12\]The alternative interpretation of this model as a dual-self, intra-agent persuasion problem permits a welfare analysis. If we give at least as much weight to the utility of the receiver as we do to the utility of the sender-self, then the full revelation of the state followed by sincere voting will always be at least as good as the persuasion outcome.
structure, and \( P[\omega = R|S_i] \) is the interim posterior belief induced by \( S_i \) before the Outside signal is observed.

We interpret this choice of signal structure as an agent choosing which combination of news media to consume. The rich set of signal structures reflects the rich media landscape available to voters today. Further, although the agent can choose a news media (a signal structure), she cannot choose the programming (a realization of a signal).

**Outside signal structure** Each agent also receives an Outside signal \((s \in \{r, l\})\) with distribution:

\[
P[s = l|\omega = L] = k, \quad P[s = r|\omega = R] = m
\]

where \( k \in [0,1] \) and \( m \in [0,1] \). Depending on the structure imposed on \( k \) and \( m \), the Outside signal may be modified to study a variety of environments. In the baseline model illustrated in subsection 4.1, we consider an environment with unbiased Outside signal structures by setting \( k = m \); we capture asymmetric exposure to Outside media by assuming that agents of type \( R \) receive a less precise signal than agents of type \( L \). In subsection 4.2 we study media distrust by supposing that agents of type \( R \) believe incorrectly that they receive a less precise signal than agents of type \( L \). We model propaganda in favor of party \( L \) in section 5 by imposing \( k > m \). Here, we consider the implications of censorship, and of citizens being believing incorrectly that the mainstream media is unbiased. Finally, we consider perceived propaganda in favor of party \( L \), in subsection 6.2, by setting \( k = m \), and we assume that type-\( R \) agents believe incorrectly that \( k > m \).

We assume that both the Inside signal and the Outside signal are realized independently for all agents, conditional on the state. This is a reasonable assumption in a media-rich environment where different media choices might have the same signal structure, but different realizations of the signal. With the independent realization of signals, there is no aggregate uncertainty in outcomes. Introducing correlation is straightforward, and the results have a similar flavor, albeit with aggregate uncertainty.

### 3.2 Structure of solution

We solve the model with the general Outside signal in appendix A and then apply the solutions to specific structures of the Outside signal that have particular interpretations. We illustrate the solution for a type-\( R \) agent, and the solution for a type-\( L \) agent is analogous. The solution of the agent’s problem in this model applies the techniques developed in [Kamenica & Gentzkow (2011)] and [Kolotilin (2018)]. As such, we are merely distilling
existing results in the persuasion literature. The purpose of providing the solution structure below is to illustrate the intuition and drivers of our main results which concern the aggregation of beliefs and votes of a countable infinity of agents.

This is an optimization problem for an agent, where her objective is to maximize \( \text{subject to the Martingale constraint specified in equation} \) \( 2 \). Simply put, the agent chooses the Inside signal structure to maximize the likelihood of political faith preservation subject to the constraint on the Inside signal structure that the expectation of the posterior belief equals the prior belief. We assume that the agent also votes sincerely following her belief realization.

**Expected utility** To solve the agent’s problem of choosing an In-media signal structure to maximize her likelihood of preserving her political faith, we first calculate the agent’s expected utility as a function of her interim priors — after she has received the In-media signal, and before she has received the Out-media signal. We denote this interim posterior \( P[\omega = R | S_i] \).

![Figure 2: Expected utility as a function of interim posteriors](image)

In figure 2, we plot a type-\( R \) agent’s expected utility as a function of her interim posterior for an Outside signal structure such that \( k = m = 0.75 \). For values of the interim posterior between 0 and 0.25, regardless of the realization of the Outside signal, the agent is never able to preserve her political faith. If the agent has an interim posterior equal to 0.25, and if she receives a favorable outside signal, her posterior expectation that the state of the world is \( R \) then equals 0.5, and she is *just* able to preserve her political faith. For values of interim posteriors between 0.25 and 0.75, the agent can preserve her political
faith if she receives a favorable Outside signal, \( s = r \). The likelihood that she receives a favorable Outside signal increases in her expectation that the state of the world is \( \omega = R \). Finally, if the agent’s interim posterior is at least 0.75, she is able to preserve her political faith regardless of the Outside signal.

Three interim posteriors are key to solving the agent’s problem. The first is an interim posterior such that the agent is just able to preserve her political faith if she receives an unfavorable Outside signal (\( \Pr[\omega = R \mid S_i] = 0.75 \) in figure 2). We call an Inside signal that generates such an interim posterior a Good (G) signal. The second is an interim posterior that allows the agent to just preserve her political faith only if she receives a favorable Outside signal (\( \Pr[\omega = R \mid S_i] = 0.25 \) in figure 2). We call an Inside signal that generates such an interim posterior a Bad (B) signal. Finally, the third is an interim posterior such that the agent is certain that the state does not match with her preferred party (\( \Pr[\omega = R \mid S_i] = 0 \)). We call an Inside signal that generates such an interim posterior a Terrible (T) signal.

We show in Appendix A that the agent needs to consider only the three signals described above — Good, Bad, and Terrible — to solve her problem. Any other signals would be suboptimal. Further, those combined in two signal structures, GT or GB, are sufficient to solve the agent’s problem. Note that the GT signal structure is one-sided, in the sense that, in the favorable state of the world, only signal G is realized. We interpret GT as media that has a high political slant.\(^{13}\) Favorable news reported by these outlets is par for the course and, thus not very informative to Bayesian agents. When these outlets report news unfavorable to their side, it is very informative for Bayesian agents, and would convince such agents that the state of the world is not favorable to their side.

On the other hand, GB is two-sided in that either G or B may be realized in either state of the world, according to a predetermined probability distribution. This is akin to the agent consuming less one-sided media insofar as it provides her with a mix of positive and negative news in either state of the world.\(^{14}\) The mixing probability changes depending upon the state. The fact that a negative signal is shown by these media in either state of the world makes it less informative and allows it to be countered by a favorable Outside signal.

In figure 3, we plot the agent’s expected utility as a function of the interim posterior she has after observing the In-media signal and before observing the Out-media signal. The graph on the left considers a case where the Out-media has a less precise signal. Here,

\(^{13}\)For instance, media sources like Breitbart News for Republicans or Huffington Post for Democrats.

\(^{14}\)We interpret the GB signal structure as media sources like Wall Street Journal for Republicans or The New York Times for Democrats.
the interim posterior that allows an agent to just preserve her political faith if she receives a favorable Outside signal is close to her prior (at $P[\omega = R|S_i] = 0.4$). Similarly, the agent preserves her political faith regardless of the realization of the Outside signal if her interim posterior is at least $P[\omega = R|S_i] = 0.6$. The graph on the right side of figure 3 considers a more precise Outside signal structure, which implies that the two key interim posteriors described above are more spread out away from the prior of $\omega = 0.5$.

For lower values of precision of the Outside signal, such as the left side of figure 3, the concave closure of the expected utility function is such that it would be optimal for the agent to choose a signal structure that mixes between $G$ and $T$ — a $GT$ signal structure. On the other hand, if the Outside signal is more precise, as it is on the right-side graph in figure 3, we plot the concave closure of the expected utility and find that the optimal signal structure is $GB$.

In Appendix A.3, we show that the structure of this problem and the optimal signal structures are robust to a small amount of uncertainty or noise in the threshold of belief required for political faith preservation by the citizen. This kind of robustness is also noted in Kamenica & Gentzkow (2011), where the key results and the structure of the solution hold in the presence of noise, which makes the receiver’s belief threshold for choosing the sender-preferred action stochastic.

In the following applications, we consider specific structures of $k$ and $m$ that admit particular interpretations and aggregate to understand the electoral implications of motivated media choice.

\[15\text{We follow the definition of concave closure in Kamenica & Gentzkow (2011).}\]
4 Liberal democracy

Consider the case of a liberal democracy, where the media is not influenced by the state. The mainstream media is unbiased and is perceived to be unbiased by citizens. When the media is influenced by a political party, we term this as an illiberal democracy with propagandizing rulers. We study that case in subsections 5.1, 5.2, and 6.1. We also study a situation in which citizens incorrectly believe that the mainstream media is systematically biased in favor of one political party subsection 6.2.

4.1 Asymmetric exposure to mainstream media

This specification builds on the illustrative example in the introduction section. Suppose that exposure to mainstream media, or attention paid to it, differs systematically by party preference. Here, there is an asymmetry in the informational insularity of citizens of differing political persuasions. This may be due to a difference in the trust placed on mainstream news or due to the media ecosystem allowing information silos to be more prevalent on one side of the political spectrum. Citizens in this environment have access to a wide array of news and opinion outlets that allow them to consume a specific diet of chosen media. We are interested in their choice of In-media as a function of their exposure to Out-media and the electoral outcomes. In this section, both agents receive unbiased Outside signals that are not fully informative, and we suppose that an agent of type \( R \) receives a noisier signal than an agent of type \( L \) does. The reverse case yields analogous results.

The Out-media signal has the following structure

\[
\begin{align*}
P[s = l | \omega = L] &= \tau \cdot t + (1 - \tau) \cdot \frac{1}{2}, \\
P[s = r | \omega = R] &= \tau \cdot t + (1 - \tau) \cdot \frac{1}{2}
\end{align*}
\]

where \( t \in [0.5, 1] \) is the precision of the Outside signal and \( \tau \) is the extent of exposure to that signal. We model the asymmetry in exposure by supposing \( \tau = 1 \) for agents of type \( L \) and \( \tau \in [0, 1] \) for type-\( R \) agents. Another isomorphic setup could simply specify two different Outside signal precisions for the two types of agents such that the Outside signal precision for an agent of type \( R \) is lower than that of an agent of type \( L \).

One interpretation of the precision of and exposure to the Outside signal is attention. The attention, or lack thereof, could be a feature of the media landscape or social circle that determines the intensity, frequency, or clarity with which agents receive the signal from Outside their chosen media. An alternative interpretation is that the signal precision and exposure capture the openness of agents to receiving a signal from outside the chosen In-
media. The asymmetry in exposure can, therefore, be thought of as capturing the agent’s preference, rather than a feature of the media system.

In the illustrative example in the Introduction, we set $t_R = 0.51$ and $t_L = 0.75$. That is equivalent to setting $t = 0.75$ and $\tau = 0.04$. For these parameter values, we find that introducing the ability to choose media transforms a fully symmetric election into one with failure of information aggregation in that party $R$ wins the election in both states of the world. This result is obtained despite agents being rational in their information processing.

Table 1: Results without and with In-media

<table>
<thead>
<tr>
<th>Without In-media</th>
<th>With In-media</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{Ex-Ante}$</td>
<td>$\text{Ex-Ante}$</td>
</tr>
<tr>
<td>$\omega = R$</td>
<td>$\omega = R$</td>
</tr>
<tr>
<td>$\omega = L$</td>
<td>$\omega = L$</td>
</tr>
<tr>
<td>R Win Margin</td>
<td>R Win Margin</td>
</tr>
<tr>
<td>0%</td>
<td>+26%</td>
</tr>
<tr>
<td>-26%</td>
<td>+28%</td>
</tr>
<tr>
<td>+26%</td>
<td>+54%</td>
</tr>
<tr>
<td>+2%</td>
<td>+2%</td>
</tr>
<tr>
<td>R Win Prob</td>
<td>R Win Prob</td>
</tr>
<tr>
<td>50%</td>
<td>100%</td>
</tr>
<tr>
<td>100%</td>
<td>0%</td>
</tr>
<tr>
<td>0%</td>
<td>100%</td>
</tr>
<tr>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

As we can see in the table, without In-media, asymmetric exposure to mainstream media does not provide side $R$ with any winning margin or winning probability advantage. The correct side wins in each state of the world. When the state of the world is $R$, then 51% of type-$R$ agents and 75% of type $L$ agents receive an outside signal indicating that the state of the world is more likely $R$, and these agents vote for party $R$, which implies that it wins with a margin of 26%. Because the Outside signal is unbiased, party $L$ wins with the same margin in state $L$.

Figure 4: Mechanism behind results

---

16 This value of $\tau = 0.04$ is only relevant for type-$R$ agents because we have fixed $\tau = 1$ for type-$L$ agents.
The agents share symmetric priors about the state of the world and update rationally. There are an equal number of each type of agent, and both types of agents want to maximize the likelihood that they preserve their political faith. This is equivalent to the agent maximizing the likelihood that she votes for her preferred party if she votes sincerely. As such, one might not expect, ex-ante, that the introduction of the Inside signal in the lower half of table [1] would result in an information aggregation failure that allows party R to win regardless of the state of the world.

As we see in panel (a) of figure 4, a GT signal structure provides an agent with a higher utility if the perceived precision of the outside signal \( t' = \tau t + \frac{(1-\tau)}{2} \) is low enough. For higher values of perceived precision, a GB signal is optimal. Under the parameter values described above \((t = 0.75 \text{ and } \tau = 0.04)\), a type-R agent chooses a GT signal structure, whereas a type L agent chooses a GB signal structure.\(^{17}\)

Panel (b) of figure 4 illustrates the probability of political faith preservation conditional on the state of the world. In the illustrative example, we considered values parameter values such that the perceived precision for type-R agents is quite low, which corresponds to relatively closed informational silos. On the other had, type-L agents have a higher perceived precision of the Outside signal, which reflects more openness to outside information. In the favorable state, GT allows an agent to preserve her political faith with certainty because the agent receives the Good signal. Compared with a GT signal, a GB signal allows the agent to preserve her political faith with a higher likelihood in the unfavorable state, but with a probability less than 1 in the favorable state of the world. Given the parameter values in table [1] when the state of the world is R, all type-R agents vote for party R. Type-L agents receive a mix of Good and Bad signals and 43.75% of them vote for party L. Therefore, party R wins in state R. When the state of the world is L, type-R agents receive a mix between Good and Terrible signals, and 96.08% of them vote for party R. Type-L agents receive a more favorable mix of Good and Bad signals than they do in state R. 93.75% of them vote for party L. Therefore, party R wins even in state L.\(^{18}\)

Each agent individually maximizes her likelihood of political faith preservation, but, for some parameters, this implies that party L loses regardless of the state of the world. If, instead, type-L agents choose GT as their Inside signal structure, then party L could win

\(^{17}\)In Appendix B.1, we show that the signal structure GT is optimal for type-R agents if and only if \( \tau \cdot t + (1-\tau) \cdot \frac{1}{2} \leq \frac{1}{\sqrt{2}} \). Because type-L agents receive an Out-media signal with a perceived precision \( t \), they choose a GT signal structure for all values of \( t \leq \frac{1}{\sqrt{2}} \) and a GB signal otherwise.

\(^{18}\)Recall that upon receiving a Good signal, an agent is always able to preserve her political faith and vote for her preferred party. If the agent receives a Bad signal, she votes for her preferred party only if she also receives a favorable Outside signal – otherwise she votes for the other party. If she receives a Terrible signal, she is certain that the state does not match her preferred party and she always votes for her non-preferred party.
in the correct state — implying that an ex-ante suboptimal individual choice can allow ex-post optimal policy choice. In other words, party $L$ could benefit if it could convince its electorate to consume more politically slanted news.

It is crucial to note that information aggregation failure occurs for low values of exposure ($\tau$) and intermediate values of precision ($t$) of the Outside signal. If the outside signal were less precise, for instance, $t = 0.7$, then both types of agents would choose a $GT$ signal structure, and the correct party would win always. On the other hand, if the Outside signal were very precise, then too, the correct party would win always. If the state of the world is $L$, then compared to the outcomes in table \[\text{table1}\], type-$L$ agents are more likely to receive a $Good$ signal and vote for party $L$ in state $L$. Similarly, type-$R$ agents are more likely to receive a $Terrible$ signal and vote for party $L$. Party $L$, therefore, receives more votes than party $R$ in state $L$.

Figure 5: Signal choices and results with asymmetric exposure

![Figure 5: Signal choices and results with asymmetric exposure](image)

In panel (a) of figure 5, we consider all values of signal precision ($t$) and exposure ($\tau$). The area shaded in red corresponds to the region where both types of agents choose a $GT$ signal structure for their In-media. The blue-shaded area denotes the region where both types of agents choose a $GB$ signal structure. In the purple-shaded region, type-$R$ agents choose a $GT$ signal structure, while type-$L$ agents choose a $GB$ signal structure. There is no region where type-$R$ agents choose a $GB$ signal structure and type-$L$ agents choose a $GT$ signal structure. As we claim in proposition \[\text{proposition1}\] below, there is a region where information aggregation fails; it is contained within the purple-shaded region, where type-$R$ and type-$L$ agents choose a $GT$ and $GB$ signal structure, respectively.

20
Proposition 1 In the environment specified in subsection 4.1, the correct candidate wins except in a region with intermediate precision the of Outside signal and low exposure for type-R agents. In the region of information misaggregation, type-R agents choose a GT signal structure and type-L agents choose a GB signal structure.

As we show in appendix B.1, the result in the example above is not knife-edge, and the voting margins are continuous in \( \tau \) and \( t \) in the neighborhood of \( \tau = 0.04 \) and \( t = 0.75 \). In figure 5, we see that the light-purple-shaded area that denotes the region of information aggregation failure exists for low values of media exposure on the right (\( \tau \)) and moderate values of Out-media precision (\( t \)).

When both types of agents choose a GT signal structure for their respective In-media, if the state of the world is \( R \), all type-R agents and some type-L agents vote for party \( R \). If the state of the world is \( L \), then all type-L agents and some type-R agents vote for party \( L \). In this case, information aggregation failure is not possible. We show in appendix B.1 that the party that matches the state wins for all values of \( \tau \) and \( t \) such that a GB signal structure is optimal for both types. Further, the parameter space is such that there is no situation in which type-R agents choose a GB signal structure and type-L agents choose a GB signal structure.

Although the region of information aggregation failure in panel (a) of figure 5 seems small, we argue that those parameter values are particularly relevant to consider in the U.S. case. The intermediate values of Out-media precision where information misaggregation takes place are high enough that type-L agents choose a two-sided, GB In-media signal structure, but not so high that party \( L \) wins in state \( L \). Further, as we see in subsection 4.3, the parameter space where information misaggregation occurs expands greatly when agents have non-common priors in favor of their preferred parties. While party \( R \) obtaining an ex-ante winning margin advantage through the lower exposure type-R agents have to Out-media may be expected, this advantage can be large enough to swing an election in all states.

4.2 Distrust in mainstream media precision

We considered an environment in which type-R agents receive a less precise Outside signal than type-L agents, so the model is correctly specified. We now show that the results are identical if type-R agents face an Outside signal structure that is identical to the one type-L agents face, but type-R agents incorrectly believe that the precision of the Outside signal is lower than it actually is. This misspecified model is particularly applicable to the U.S. considering the asymmetry among voters in the two main political parties in trust in
mainstream media (as documented in the introduction). Crucially, note that Republicans have a lower level of trust and that key Republican leaders have sought to exacerbate that distrust. Here, we interpret asymmetric distrust as reflecting agents’ judgment on the quality of mainstream media. In subsection 6.2, we interpret the asymmetric distrust as a judgment on the bias of mainstream media.

The Outside signal has the following structure:

$$
P[s = l|\omega = L] = \tau \cdot t + (1 - \tau) \cdot \frac{1}{2}, \quad P[s = r|\omega = R] = \tau \cdot t + (1 - \tau) \cdot \frac{1}{2}\]
$$

where \(t \in [0.5, 1]\) is the true precision of the Outside signal and \(\tau\) is the extent of trust in the Outside signal agents have. We specify the asymmetry in distrust by supposing that type-L agents have \(\tau = 1\) and \(\tau \in [0, 1]\) for type-R agents.

The choice of signal structures by either type of agent is identical to the previous specification. Again, we find misaggregation of information within the region where type-R agents choose a GT signal structure and type-L agents choose a GB signal structure. Because type-L agents are correct about the process generating the Outside signal and because the Outside signal does not impact the likelihood of political faith preservation for type-R agents, the region of misaggregation of information where party R wins regardless of the state is identical to the region presented in panel (a) of figure 5.

We considered an interpretation of the lower precision of the Outside signal for agents of type R where that asymmetry comes not from less exposure, but through an incorrect belief regarding the Outside media. We find that party R is imbued with such a large electoral advantage by this asymmetry that, for some parameter values, it can win regardless of the state of the world. In sum, whether we consider an asymmetry in exposure to Out-media or distrust in the quality of Out-media (see figure 1), we get identical results.

### 4.3 Non-common priors

In addition to cultural-political affiliations that define the agent’s type, it may be reasonable to consider the possibility that agents who are culturally affiliated with a political party may hold more favorable priors towards it. We can very simply extend the baseline model to consider the implication of such non-common priors. Suppose that type-R agent holds a prior \(P_R[\omega = R] := w_R\) greater than that for a type-L agent \(w_L\). As we show here, by assuming common priors, we have stacked the cards against our results.

We see in panel (b) of figure 5 that the region of information aggregation failure, where party R wins regardless of the state of the world, is much larger with non-common priors.
than with common and symmetric priors ($P[\omega = R] := w = 0.5$). All calculations are detailed in Appendix B.3.

Here, the agent can optimally choose a non-informative In-media signal structure if the prior allows the agent to preserve her political faith regardless of the realization of the Out-media signal. We call that kind of In-media signal Neutral or Nonsense news. This is akin to a citizen consuming opinion news commentary that has a political bias does not claim to be providing journalistic facts.

We have, therefore, tied our hands by assuming that agents share common priors. The key result of information aggregation failure holds for a larger subset of parameters when agents have non-common priors that favor their preferred party.

In the next section, we consider the implications of a propagandized (biased) Outside signal on the agents’ choice of In-media as well as electoral outcomes.

5 Propaganda

State influence on mainstream media reporting is a telltale sign of a decaying democracy. Here, we consider propaganda, which we define as a bias in the Outside signal structure. In the model, this bias emanates from an asymmetry in the realization of favorable and unfavorable news for political parties as a function of the state of the world. We are interested in the impact of propaganda on the choice of Inside-media by agents as well as on electoral outcomes. In particular, this model allows us to study if an agent’s ability to choose Inside-media can counteract propaganda. Further, we learn whether that same ability can lead to information aggregation failure if the agent incorrectly believes that the mainstream media is biased.

Suppose that the Outside signal is biased in favor of party $L$ such that the signal $l$ is realized more often. Specifically

\[
\begin{align*}
    P[s = l|\omega = L] &= \tau \cdot t + (1 - \tau) \cdot 1, \\
    P[s = r|\omega = R] &= \tau \cdot t + (1 - \tau) \cdot 0
\end{align*}
\]

where $t \in [0.5, 1]$ is the precision of the Outside signal, and $\tau \in [0, 1]$ is the extent of unbiasedness of state-influenced media. A high level of $\tau$ corresponds to a less-biased signal, while a low level of $\tau$ corresponds to a very biased signal. This bias is commonly known by all agents and is the true process that generates the Outside signal.

Such a signal structure for the Outside or mainstream media is possible when the state exerts control over mainstream media outlets. Suppose, for instance, that a strongman leader of an Eastern European nation forces the state-run media to run positive stories
often, but he cannot prevent the occasional negative story from being run. In such an environment, it is also reasonable to consider the possibility that the strongman leader can censor media outlets so that agents cannot receive an Inside signal. In the next two subsections, we consider the implications of propaganda with and without censorship.

5.1 With censorship

Suppose that neither type of agent has access to any information other than her prior and the realized Outside signal. This scenario may be interpreted either as the party in power shutting down all media other than the propagandized state-controlled media. An alternative interpretation is that the media environment is not rich because we are considering a time before the environment became rich through the proliferation of news outlets and social media. We consider this benchmark of censorship, as it will help us highlight the role that the ability to choose Inside media plays.

Figure 6: Propaganda with and without censorship

It is straightforward to show that because the agents share common and symmetric priors, and because the Outside signal is informative, that the realization of the Outside signal determines whether the agent preserves her political faith. The condition for party $L$ to win in state $L$ is simply that the Outside signal $l$ is realized more often than the signal $r$. This always holds. The details of this and the following calculations are presented in appendix B.4.

If the state of the world is $R$, then the Outside signal $l$ is realized more often if $\tau t < 0.5$. If this condition holds, then we have information misaggregation, and party $L$ wins regardless of the realization of the state of the world. This region is illustrated in panel
(a) of figure 6 above. In the next subsection, we allow agents to design a chosen signal structure.

If a ruler of an illiberal democracy can influence the mainstream media and also prevent citizens from independently accessing information, then he or she can ensure victory regardless of the state of the world for a large subset of the parameter space. Censorship along with propaganda is, therefore, a powerful combination of tools in an illiberal democracy.

5.2 Without censorship

Here, the agents can curate their Inside signal structure. Because agents are rational, they find an Outside signal favorable to the propagandizing side (party L) to be less informative than an Outside signal unfavorable to that party. Recall that a Good signal from the Inside signal structure is designed to just counteract an unfavorable Outside signal. For type-R agents, the unfavorable Outside signal, $I$, is relatively easy to counteract because it is less effective. For type-L agents, the unfavorable Outside signal is more expensive to counteract. As we see in panel (b) of figure 6, there is a large subset of parameters within which type-R agents optimally choose a GT signal structure and type-L agents choose a GB signal structure. The region of information misaggregation lies within this subset of parameter values.

In the region of information misaggregation, party R wins regardless of the state of the world. The intuition behind this result is that, if the state of the world is R, then party R must win because all type-R agents and some type-L agents vote for party R. If the state of the world is L, then the informativeness of the Outside signal, $r$, implies that not enough type-L agents preserve their political faith and thereby vote for party L. For propaganda to backfire, it must be that the Out-media is biased enough, and that the precision of the Outside signal is strong enough to push type-L agents to choose a GB signal structure for their In-media. However, if the precision of the Out-media signal is high, then party L would win in state L, and no misaggregation of information is present.

<table>
<thead>
<tr>
<th>Parameters: $t = 0.6, \tau = 0.3$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Ex-Ante $\omega = R$ $\omega = L$</td>
</tr>
<tr>
<td>L Win Margin</td>
</tr>
<tr>
<td>L Win Prob</td>
</tr>
</tbody>
</table>
In table 2, we consider parameter values where $t = 0.6$ and $\tau = 0.3$. All calculations for these parameters are easily done by plugging in those values into the expressions derived in appendix B.4 and B.5. These parameter values correspond to a situation where the Outside signal is fairly precise, and the party-$L$–influenced media is known by citizens to be very biased. We see that, with censorship, party $L$ can win always, and with high margins reminiscent of electoral results seen in some Eastern European “democracies” controlled by strongmen leaders. On the other hand, without censorship, party $L$ loses in either state of the world. Therefore, in a substantial subset of parameters, we see a reversal in the electoral outcomes when censorship is disallowed. Suppose that party $L$ cannot perfectly target propaganda, and that there is a likelihood of mistakes that implies that there is a positive probability to all levels of bias and precision. Then, to benefit from propaganda, party $L$ must also institute censorship. Otherwise, with a positive chance, propaganda backfires.

**Proposition 2** In the environment specified in section 5.2, the correct candidate wins except in a region with low to intermediate precision and a highly biased Outside signal. In the region of information misaggregation, type-$R$ agents choose a GT signal structure, and type-$L$ agents choose a GB signal structure.

The calculations and the proof is detailed in appendix B.5. We find that, without censorship, propaganda is not simply weak. It backfires. Propaganda, by its very nature, implies that news favorable to the propagandizing party is discounted by Bayesian citizens and does not move their posteriors substantially. News unfavorable to the ruling party is very informative because it is rare. If citizens can independently access information, then type-$R$ citizens need very little to counteract the state-influenced media’s propaganda for party $L$. Type-$L$ citizens need a much stronger signal to counteract unfavorable news from the state-influenced media. For all parameters, party $R$ wins in the state it is supposed to (state $R$). For a substantial parameter space, party $R$ also wins in state $L$.

## 6 Misperceived Propaganda

Here, we consider two types of misperception of propaganda. In subsection 6.1, we consider an environment where propaganda is present and agents, or a subset of agents, believe that the Out-media provides an unbiased signal. This captures what happens when the ruler of an illiberal democracy manages to induce mainstream media into reporting propagandized news while also convincing citizens that it is unbiased. Although this
scenario may imbue the propagandizing side with some advantage, the fact that agents are not prevented from choosing In-media signals could counteract the propaganda. In subsection 6.2, we suppose that, while the Outside signal is unbiased, agents of type $R$ believe that it is biased in favor of party $L$. As noted earlier, this is an alternative interpretation of the asymmetric trust in mainstream media in the U.S. (see figure 1). The incorrect beliefs held by type $R$ agents influence their choice of In-media and their belief updating. This can have important electoral consequences.

6.1 Propaganda with oblivious citizens

We model the rather insidious environment in which the mainstream media is biased in favor of party $L$ but citizens are convinced that this mainstream media is unbiased. We first consider the case where both types of agents are oblivious to the bias in the Outside signal. Later in this subsection, we consider the cases in which either only type-$R$ or only type-$L$ agents are oblivious to the bias. In each case, at least some citizens choose their In-media signal structure under incorrect beliefs. The mainstream media, which is biased toward party $L$, generates a signal favorable to party $L$ more often than it does for party $R$.

The true process that generates the Out-media signal is biased in favor of party $L$ and takes the following form:

$$P[s = l|\omega = L] = \tau \cdot t + (1 - \tau) \cdot 1, \quad P[s = r|\omega = R] = \tau \cdot t + (1 - \tau) \cdot 0$$

where $t \in [0.5, 1]$ is the precision of the Outside signal and $\tau \in [0, 1]$ is the true unbiasedness of the Outside signal. The first case we consider is where both types of agents incorrectly believe that $\tau = 1$.

As we see in panel (a) of figure 7, both types of agents choose a GB signal structure for their In-media in the region of information aggregation failure. Detailed calculations are provided in appendix B.6.

The agents are Bayesian, but, because they hold incorrect beliefs about this process, they update incorrectly and choose their In-media signal structure sub-optimally. Type-$R$ agents update too much upon receiving an unfavorable Outside signal, and brace for the Outside signal by choosing an overly informative Inside signal structure. They thus preserve their political faith less often than they could have. Similarly, type-$L$ citizens update too little upon receiving unfavorable outside information — which works to their benefit. As we show in proposition 3, for low values of Out-media unbiasedness ($\tau$) and moderate to high values of Out-media precision ($t$), party $L$ enjoys such a substantial
Figure 7: Signal choices and results in propaganda with oblivious citizens

(a) Signal choices and results with all citizens oblivious

(b) Results with only type-R citizens oblivious

advantage that it can win regardless of the realized state of the world.

**Proposition 3** In the environment specified in subsection 6.1, the correct candidate wins, except in a region with intermediate precision and highly biased Outside signal. In the region of information misaggregation, both types of agents choose a GB signal structure.

The proof and detailed calculations are in appendix B.6. The existence of such a region of information aggregation failure shows us why propagandizing outlets go to great lengths to portray themselves as accurate and balanced.

**Censorship:** We consider the implications of censorship in an environment where the Out-media is biased but citizens wrongly believe that it is not. Party $L$ wins regardless of the state if the Out-media signal favorable to party $L$ is generated more often in either state of the world. As in subsection 5.1, this condition holds if $\tau < 0.5$. The electoral results are, therefore, also the same and are illustrated in panel (a) of figure 6. The only difference from subsection 5.1 is in the intensity of beliefs that citizens hold about the parties for which they vote.

$L$ agents oblivious: So far in this section, we have considered the case where both types of players are oblivious to the bias in the Outside signal. Suppose, now, that only $L$ type agents believe that the Outside signal is unbiased, while type-$R$ citizens know that the Outside signal is biased. Agents of type $L$ update less than they should when faced with an unfavorable Outside signal, and their Inside signal structure choice is the same as in
Agents of type R choose their Inside signal structure, as shown in figure 6. The correct side always wins in this specification, which, for party L, is an improvement over the backfiring of propaganda we saw in subsection 5.2.

**R agents oblivious:** Here, we suppose that only agents of type R are unaware of the bias in the Outside signal. Type-L agents know that the Outside signal structure is biased in favor of Party L. Type-L agents choose their Inside signal structure correctly, which is the same as in panel (b) of figure 6. Type-R agents hold incorrect beliefs and choose their Inside signal structure such that it is the same as in panel (a) of figure 7.

As we see in panel (b) of figure 7, the region where party L wins regardless of the state of the world expands if only type-R agents, rather than both types, are oblivious to the bias in the Outside signal. This illustrates that, for propaganda to benefit the ruler of an illiberal democracy, the opposing side must be unaware of it. Alternatively, the propagandizers could use censorship to win regardless of the state.

### 6.2 Perceived propaganda

What happens if no actual propaganda is present and yet some citizens believe that there is? This obviously reflects distrust in the unbiasedness of the Outside media. Here, we suppose that type-R agents believe the Outside media to be biased when it truly is not.

Suppose the Outside signal structure is unbiased and has a precision, $t \in [0.5, 1]$. Type-R agents hold an incorrect belief and perceive the media to be biased when it is not. They believe that the Outside signal structure follows

$$
P[s = l | \omega = L] = \tau \cdot t + (1 - \tau) \cdot 1,
$$

$$
P[s = r | \omega = R] = \tau \cdot t + (1 - \tau) \cdot 0
$$

where $\tau \in [0, 1]$ is the belief in the extent unbiasedness of the Outside signal held by type-R agents. Type-L agents correctly believe that the process that generates the Outside signal is unbiased, which is equivalent to them believing that $\tau = 1$ in the process described above.

In figure 8, we see the Inside signal structure choices for each type of agent as well as the region of information aggregation failure. This region occurs where type-R agents perceive the media to be highly biased, and where the Outside signal is at least moderately precise. The intuition for this result is similar to the one provided in subsection 5.2 — that the perception of bias in the Outside signal structure makes an unfavorable Outside signal less informative for type-R agents. Therefore, an asymmetry in the perception of bias by the mainstream imbues party R with an advantage that allows it to win regard-
less of the state for a subset of parameters. The calculations are provided in appendix B.7.

**Proposition 4** In the environment specified in subsection 6.2 the correct candidate wins except in a region with intermediate precision and low perceived unbiasedness of the Outside signal. In the region of information misaggregation, type-R agents choose a GT signal structure and type-L agents choose a GB signal structure.

We find that the perception of propaganda is a strong political implement that can allow the incorrect party to win. One implication of this specification of this model is that party R has an incentive to convince type-R voters that the mainstream media is biased in favor of party L. As noted in the introduction, in the U.S., President Trump railed against the mainstream media, accusing it of being biased. Our model suggests that if type-R voters believe that the mainstream media is biased in favor of party L, then that influences the In-media choices of type-R voters. This imbues party R with a big electoral advantage, and for a substantial subset of parameters, party R can win regardless of the state of the world.

### 7 Conclusion

We have tried to understand how instrumental media choice driven by political identity can drive aggregate electoral results. In the core setup, the bias we assume is only in the preferences that drive media choice and is not cognitive. Specifically, all citizens update
rationally all the information they receive and vote according to this information. Later, we added cognitive biases to see how results change.

For liberal democracies, such as the United States, our premise is a perfectly symmetric benchmark where the asymmetry between the two sides appears only in the exposure to Outside media or in the bias of Outside media. In these contexts, we have shown how skewed the electoral outcomes become and highlighted situations in which information aggregation fails. When information aggregation fails, one candidate wins in all states; thus, it also wins when it is the worse candidate, i.e., when *average/expected* rational beliefs after *any* media signal, chosen or not, are unfavorable to them. A key determinant of this failure of information aggregation is that one side has low exposure, or low trust, in mainstream media, while the other side has moderate exposure/trust in it. In this region, the presence of a rich new media environment allows for a winning margin advantage to the less exposed side so substantial that it can swing elections in all states. This misaggregation region (characterized by low and moderate exposure on either side of the political spectrum) seems particularly salient in democracies such as the U.S., and it suggests why some parties have incentives to sow distrust in mainstream media. If we consider in addition, cognitive biases such as biased priors, which may also be salient, this misaggregation region becomes more prominent.

Regarding illiberal democracies, our key insight is that for government propaganda to work, it is crucial that citizens are unaware of it. In other words, citizens also have a cognitive bias. If citizens are in fact aware of the bias, then the presence of free media undermines government propaganda spectacularly: propaganda may backfire entirely, making the non-propagandizing side win regardless of the state. This suggests why it is not enough for authoritarian governments to control the official media message and to promote government propaganda when citizens are aware of it. For such governments to maintain power, it is also fundamental to crack down on the free press and any dissenting news outlets. Regimes must wall off their populations from possibly inconvenient truths revealed by non-governmental media.

References


A Solution for general model

Suppose that the agent believes (correctly or incorrectly) that

\[ P[s = l|\omega = L] = k, \quad P[s = r|\omega = R] = m \]

The agent’s expected utility still follows equation 1 and the signal structure must satisfy the Martingale constraint specified in equation 2.

The agent’s interim posteriors on observing the outside signal and without observing the signal from the chosen signal structure are such that:

\[ P[\omega = L|s = l] = \frac{k}{1 + k - m}, \quad P[\omega = L|s = r] = \frac{1 - k}{1 + m - k} \]

The \( G \), \( B \), and \( T \) signals are described section 3. A type-R agent is able to preserve her political faith as long as her posterior upon observing both signals is such that \( P[\omega = R|s, S] \geq 0.5 \). Therefore, the \( G \), \( B \), and \( T \) signals must be such that:

\[ P[\omega = R|S = G] = \frac{k}{1 + k - m} \]
\[ P[\omega = R|S = B] = \frac{1 - k}{1 + m - k} \]
\[ P[\omega = R|S = T] = 0 \]

Claim 1 It is sufficient to consider three signals used in two possible signal structures when solving the agent’s problem described above.

Proof. The following proof follows directly from Kolotilin (2018) and Kamenica & Gentzkow (2011). We use the linear optimization technique of comparing marginal utility to price ratios of different signals. In figure 2, the marginal utility to price ratio of a signal is the slope of the line joining the origin to the point on the expected utility curve that corresponds to the interim posterior generated by that signal. In that sense, the next parts of
the proof are simply confirming the shape of the concave closure of the expected utility curve.

Lemma 2 Any signal $M_i$ which generates a posterior $\mathbb{P}[^{\omega = R}M_i] \in \left( \frac{k}{1+k-m}, 1 \right]$ is sub-optimal when compared to a signal $M_G$, where $M_G$ is such that $\mathbb{P}[^{\omega = R}M_G] = \frac{k}{1+k-m}$.

Proof. Regardless of whether the agent observes a signal $M_i$ or $M_G$, the agent will have the same expected utility. This is because the agent would be able to preserve her political faith regardless of the realization of the outside signal.

However, it is more costly (according to the Martingale constraint) to generate the signal $M_i$. Therefore, the Marginal utility to price ratio of generating $M_i$ is lower than for generating $M_G$. This implies that any signal structure where $\mathbb{P}[^{\omega = R}M_i] > 0$ will have a lower ex-ante expected utility than a signal structure which assigns $\mathbb{P}[^{\omega = R}M_i] = 0$ and adds $\mathbb{P}[^{\omega = R}M_G][^{\omega = R}M_i]$ to the probability that $M_G$ is generated. Therefore, no $M_i$ such that $\mathbb{P}[^{\omega = R}M_i] \in \left( \frac{k}{1+k-m}, 1 \right]$ will be chosen by the agent. Equivalently, $M_i$ is sub-optimal when compared to $M_G$.

Lemma 3 Any signal $M_i$ which generates a posterior $\mathbb{P}[^{\omega = R}M_i] \in \left( \frac{1-k}{1+m-k}, \frac{1-k}{1+m-k} \right]$ is sub-optimal compared to a signal $M_B$, where $M_B$ is such that $\mathbb{P}[^{\omega = R}M_B] = \frac{1-k}{1+m-k}$.

Proof. For $M_i$, the ex-ante expected utility of the agent is given by

$$\mathbb{P} \left[ \mathbb{P}[^{\omega = R}M_i, s] \geq 0.5 \right] = m\mathbb{P}[^{\omega = R}M_i] + (1-k)(1-\mathbb{P}[^{\omega = R}M_i])$$

$$= 1 - k + \mathbb{P}[^{\omega = R}M_i](m + k - 1)$$

This implies that the marginal utility to price ratio is

$$(m + k - 1) + \frac{(1-k)}{\mathbb{P}[^{\omega = R}M_i]}$$

For $M_B$, the ex-ante expected utility is

$$\mathbb{P} \left[ \mathbb{P}[^{\omega = R}M_B, s] \geq 0.5 \right] = m\left( \frac{1-k}{1+m-k} \right) + (1-k)(1-\left( \frac{1-k}{1+m-k} \right)) = 2m\left( \frac{1-k}{1+m-k} \right)$$

This means that the marginal utility to price ratio is...
\[2m = (m + k - 1) + \frac{(1-k)}{\frac{1-k}{1+m-k}}\]

Since \(P[\omega = R|M_i] > \frac{1-k}{1+m-k}\), the MU-Price ratio for generating a signal-structure posterior \(M_i\) is lower than for \(M_B\). Therefore, no \(M_i \in (\frac{1-k}{1+m-k}, \frac{k}{1+k-m})\) will be chosen by the agent. Equivalently, \(M_i\) would be sub-optimal when compared to \(M_B\).

\[\blacksquare\]

**Lemma 4** Any signal \(M_i\) which generates a posterior \(P[\omega = R|M_i] \in (0, \frac{1-k}{1+m-k})\) is sub-optimal compared to a signal \(M_T\), where \(M_T\) is such that \(P[\omega = R|M_T] = 0\).

**Proof.**

For \(M_i\), the ex-ante expected utility is zero. This is because regardless of the realization of the outside signal, the agent is never able to preserve her political faith. The same is true for \(M_T\). However \(M_i > 0\), which implies that the cost for generating a signal \(M_i\) is higher than for \(M_T\). Therefore, no \(M_i \in (0, \frac{1-k}{1+m-k})\) will be chosen by the sender-self. Equivalently, \(M_i\) would be sub-optimal when compared to \(M_T\).

\[\blacksquare\]

The agent requires only three signals to solve her problem. In fact, any signal that generates a posterior different from these three will be sub-optimal. \(M_G\), which generates a posterior \(P[\omega = R|M_G] = \frac{k}{1+k-m}\), is abbreviated to \(G\). \(M_B\), which generates a posterior \(P[\omega = R|M_B] = \frac{1-k}{1+m-k}\), is abbreviated to \(B\). Finally, \(M_T\), which generates a posterior \(P[\omega = R|M_T] = 0\), is abbreviated to \(T\).

A signal structure is a combination of signal realizations. The three possible signals are \(G\), \(B\), and \(T\). Therefore, the possible signal structures are \(GT\), \(GB\), and \(GBT\).\[19\]

We argue that while \(GBT\) is feasible according to the budget constraint, and can even be an optimal choice for some parameters, it is safe to ignore it. Whenever it is optimal, a simpler signal structure \((GB\) or \(GT\)) is also optimal and feasible. In other words, this signal structure never offers strictly greater expected utility to the agent (than the max of \(GB\), and \(GT\)), and is, therefore, not required to solve the agent’s problem. Either the MU-Price ratio of \(G\) is higher than of \(B\) in which case \(GT\) should be implemented, and not \(GBT\). Or the MU-Price ratio of \(B\) is higher than of \(G\) in which case \(GB\) should be implemented, and not \(GBT\). Or, finally, the MU-Price ratio of \(G\) and \(B\) are equal, in which

\[19\text{There are a number of signal structures that are ruled out because they violate the Martingale constraint. Specifically, } G, B, T, \text{ and } BT \text{ are ruled out for this reason. While we assume that the agents share a common symmetric prior belief that } P[\omega = R] = 0.5, \text{ this result is robust to values of } P[\omega = R] \text{ such that } \frac{1-k}{1+m-k} < P[\omega = R] < \frac{k}{1+k-m}.\]
case either GB or GT give the sender-self the same expected utility as GBT, and therefore, GBT can be safely ignored.

Therefore, we state that a set of signal structures that are sufficient to solve the agent’s problem are GT and GB.

\[ \text{A.1 Type-} R \text{ agent’s problem} \]

Recall that

\[ P[s = l|\omega = L] = k, \quad P[s = r|\omega = R] = m \]

This is a linear optimization problem. Therefore, the agent chooses to employ the signals with the highest marginal utility to price (MU-P) ratio.

For signal G, the MU is 1. This is because regardless of the outside signal, the agent is able to preserve her political faith. For signal B, the MU is equal to the likelihood that the outside signal is favorable (r, for a type-R agent) given that B is realized. This equals \(2m(\frac{1-k}{1+m-k})\). Finally, for T, the agent is never able to preserve her political faith, and therefore, the MU is 0.

The price of each of these signals is determined according to coefficient corresponding to it in the Martingale constraint \((\frac{k}{1+k-m}) \cdot P_G + (\frac{1-k}{1+m-k}) \cdot P_B + 0 \cdot P_T = 0.5\). This price is simply the intermediate posterior generated by the signal.

The MU-P ratio for the signal G is \(\frac{1}{1+k-m}\), for signal B, it is \(\frac{2m(\frac{1-k}{1+m-k})}{(1+m-k)} = 2m\), and for signal T, it is undefined.

The signal-structure GT is optimal when \(MU - P_G \geq MU - P_B\). This simplifies to \(1 + k - m - 2km \geq 0\). If \(1 + k - m - 2km \leq 0\), then signal-structure GB is optimal. This is equivalent to saying that the concave closure of the expected utility curve shown in the graphs in figure 3 has a kink if and only if \(1 + k - m - 2km < 0\). If a kink exists in that curve, then a GB signal gives the agent a higher expected utility than a GT signal.

We can now calculate the probability of realization of different signals, utility, and the likelihood of voting for the preferred party under the signal structures GT and GB. Our calculations are for the type-R agent, the calculations for a type-L agent are analogous. \(^{20}\)

The probabilities of realization of different signals will help us calculate expected utilities as well as the outcome of the election.

\[^{20}\text{All one would have to do is to replace} \ w \ \text{with} \ 1 - \ w \ \text{and replace} \ k \ \text{and} \ m.\]
Signal structure \( GT \): Unconditional on the state, the likelihood that the signal \( G \) is realized is \( P_G = \frac{1+k-m}{2k} \). This is also the ex-ante expected utility of the agent.

Conditional on the state being \( \omega = R \), the signal \( G \) is always realized and so the agent’s expected utility is \( E[U_R|G,T \cap \omega = R] = P[G|\omega = R] = 1 \).

Conditional on the state being \( \omega = L \), the likelihood that signal \( G \) is realized is \( \frac{1-m}{k} \). Whenever signal \( G \) is realized, the agent is able to preserve her political faith. Therefore, this also equals the agent’s expected utility \( (E[U_R|G{T} \cap |\omega = L]) \)

Signal structure \( GB \): This signal structure is a little more complicated. So, we use the following three equations.

\[
P[G|GB] + P[B|GB] = 1 \tag{3}
\]

\[
\frac{P[G|\omega = L]}{P[G|\omega = R]} = \frac{1 - m}{k} \tag{4}
\]

\[
\frac{P[B|\omega = L]}{P[B|\omega = R]} = \frac{m}{1 - k} \tag{5}
\]

Given that the signal structure is \( GB \)

\[
1 = P[G] + P[B]
\]

\[
= \left( P[G|\omega = R] \cdot \frac{1}{2} + P[G|\omega = L] \cdot \left( \frac{1}{2} \right) \right) + \left( P[B|\omega = R] \cdot \frac{1}{2} + P[B|\omega = L] \cdot \left( \frac{1}{2} \right) \right)
\]

\[
= \frac{1}{2} \left( \frac{P[G|\omega = R]}{P[G|\omega = R]} \right) \left( 1 + \frac{P[G|\omega = L]}{P[G|\omega = R]} \right) + \left( \frac{P[B|\omega = R]}{P[B|\omega = R]} \right) \left( 1 + \frac{P[B|\omega = L]}{P[B|\omega = R]} \right)
\]

This simplifies to

\[
1 = P[G|\omega = R] \left( \frac{1+k-m}{2k} \right) + \left( 1 - P[G|\omega = R] \right) \frac{(1+m-k)}{2(1-k)}
\]

Therefore, conditional on the state,

\[
P[G|GB \cap \omega = R] = k, \quad P[B|GB \cap \omega = R] = 1 - k
\]

\[
P[G|GB \cap \omega = L] = 1 - m, \quad P[B|GB \cap \omega = L] = m
\]

and unconditional on the state:

\[
P[G|GB] = \frac{1+k-m}{2}, \quad P[B|GB] = \frac{1+m-k}{2}
\]

To calculate the likelihood of political faith preservation, and therefore, expected utility,
it is helpful to recall that 

\[ \mathbb{P} [s = l | \omega = L] = k, \quad \mathbb{P} [s = r | \omega = R] = m \]

The agent’s expected utility conditional on the state being \( \omega = R \) is:

\[
E[U_{R|GB\cap \omega=R}] = \mathbb{P}[G|GB \cap \omega = R] + \mathbb{P}[r|B \cap \omega = R] \cdot \mathbb{P}[B|GB \cap \omega = R] = k + m(1-k)
\]

The agent’s expected utility conditional on the state being \( \omega = L \) is:

\[
E[U_{R|GB\cap \omega=L}] = \mathbb{P}[G|GB \cap \omega = L] + \mathbb{P}[r|B \cap \omega = L] \cdot \mathbb{P}[B|GB \cap \omega = L] = (1-m) + (1-k)m
\]

The unconditional expected utility is simply a weighted average of the conditional expected utilities. If the agent’s beliefs about the signal structure of the outside signal are correct, then:

\[
E[U_{R|GB}] = \frac{1}{2} \cdot E[U_{R|GB\cap \omega=R}] + \frac{1}{2} \cdot E[U_{R|GB\cap \omega=L}] = \frac{1 + k + m - 2km}{2}
\]

### A.2 Type-L agent’s problem

Recall that:

\[ \mathbb{P} [s = l | \omega = L] = k, \quad \mathbb{P} [s = r | \omega = R] = m \]

For signal \( G \), the MU is 1. This is because regardless of the outside signal, the agent is able to preserve her political faith. For signal \( B \), the MU is equal to the likelihood that the outside signal is favorable (\( l \), for a type-L agent) given that \( B \) is realized. This equals \( 2k(\frac{1-m}{1+k-m}) \). Finally, for \( T \), the agent is never able to preserve her political faith, and therefore, the MU is 0.

The price of each of these signals is determined according to coefficient corresponding to it in the Martingale constraint \( \frac{m}{1+m-k} \cdot P_G + \frac{1-m}{1+k-m} \cdot P_B + 0 \cdot P_T = 0.5 \). This price is simply the intermediate posterior generated by the signal.

The MU-P ratio for the signal \( G \) is \( \frac{1}{m} \), for signal \( B \), it is \( \frac{2k(\frac{1-m}{1+k-m})}{(\frac{1-m}{1+k-m})} = 2k \), and for signal \( T \), it is undefined.
The signal-structure $GT$ is optimal when $MU - P_G \geq MU - P_B$. This simplifies to $1 + m - k - 2km \geq 0$. If $1 + m - k - 2km \leq 0$, then signal-structure $GB$ is optimal.

We can now calculate the probability of realization of different signals, utility, and the likelihood of voting for the preferred party under the signal structures $GT$ and $GB$.

The probabilities of realization of different signals will help us calculate expected utilities as well as the outcome of the election.

**Signal structure $GT$:** Unconditional on the state, the likelihood that the signal $G$ is realized is $P[G] = \frac{1+m-k}{2m}$. This is also the ex-ante expected utility of the agent.

Conditional on the state being $\omega = L$, the signal $G$ is always realized and so the agent’s expected utility is $E[U_L|GT \cap \omega = L] = P[G|\omega = L] = 1$.

Conditional on the state being $\omega = R$, the likelihood that signal $G$ is realized is $\frac{1-k}{m}$. Whenever signal $G$ is realized, the agent is able to preserve her political faith. Therefore, this also equals the agent’s expected utility ($E[U_L|G \cap \omega = R]$)

**Signal structure $GB$:** This signal structure is a little more complicated. So, we use the following three equations.

\[
P[G|GB] + P[B|GB] = 1 \tag{6}
\]
\[
P[G|\omega = R] = \frac{1-k}{m} \tag{7}
\]
\[
P[G|\omega = L] = \frac{k}{1-m} \tag{8}
\]

Conditional on the state,
\[
P[G|GB \cap \omega = L] = m, \quad P[B|GB \cap \omega = L] = 1 - m
\]
\[
P[G|GB \cap \omega = R] = 1 - k, \quad P[B|GB \cap \omega = R] = k
\]

and unconditional on the state:
\[
P[G|GB] = \frac{1+m-k}{2}, \quad P[B|GB] = \frac{1+k-m}{2}
\]

To calculate the likelihood of political faith preservation, and therefore, expected utility, it is helpful to recall that
\[
P[s = l|\omega = L] = k, \quad P[s = r|\omega = R] = m
\]
The agent’s expected utility conditional on the state being $\omega = L$ is:

$$E[U_L|GB \cap \omega = L] = \mathbb{P}[G|GB \cap \omega = L] + \mathbb{P}[l|B \cap \omega = L] \cdot \mathbb{P}[B|GB \cap \omega = L]$$

$$= m + k(1 - m)$$

The agent’s expected utility conditional on the state being $\omega = R$ is:

$$E[U_L|GB \cap \omega = R] = \mathbb{P}[G|GB \cap \omega = R] + \mathbb{P}[l|B \cap \omega = R] \cdot \mathbb{P}[B|GB \cap \omega = R]$$

$$= (1 - k) + (1 - m)k$$

The unconditional expected utility is simply a weighted average of the conditional expected utilities. If the agent’s beliefs about the signal structure of the outside signal are correct, then:

$$E[U_L|GB] = \frac{1}{2} \cdot E[U_L|GB \cap \omega = R] + \frac{1}{2} \cdot E[U_L|GB \cap \omega = L]$$

$$= \frac{1 + k + m - 2km}{2}$$

### A.3 Robustness to noise

Figure 9: Expected utility as a function of interim posteriors

Here we show that the key results of the model remain unchanged if we incorporate a small amount of noise in the threshold that agents have for political faith preservation. This is very similar to, and follows directly from Extension A in Kamenica & Gentzkow.
wherein the authors find that the stochasticity in the receiver’s action makes the sender’s expected payoff function smooth, and that the key results are unchanged.

Suppose that the belief threshold that type-R agents have for the preservation of their political faith is stochastic with the mean at $P[\omega = R|S,s] = 0.5$. The noise in the threshold implies that the agent’s preservation of her political faith is now stochastic, and varies continuously with the agent’s interim posterior. The smoothness of the agent’s expected utility function in figure 9 is because of this stochasticity.

If we have a small amount of mean-zero, normally-distributed noise in this threshold, we see in figure 9 that for less precise Outside signal structure, a $G^+T$ signal is optimal. The $G^+$ signal generates an interim posterior that is more favorable to party $R$ than a $G$ signal and in the case of normally-distributed noise, it is optimal because it allows the agent to preserve her political faith for a large proportion of the possible realizations of the noise.

Similarly, for high precision Outside signals, a $G^+B^+$ signal structure is optimal for type-R agents. For minute levels of noise, the optimal signal structures remain almost identical and would result in almost the same signal structures being chosen.

Since the key results in propositions 1, 2, 3, and 4 were not knife-edge, they are robust to stochasticity in the belief threshold, as long as the variance of the noise is low.

### A.4 Robustness to gain from holding more favorable posteriors

Here we show that the if we include a utility gain for the agent from holding posteriors that are more favorable to the agent-preferred party, then for all levels of that gain in the utility function, the results of this model are identical.

The utility function of a type-R agent is now modified to:

$$U_R = \begin{cases} 
(1 - \lambda) + \lambda P[\omega = R|S,s] & \text{if } P[\omega = R|S,s] \geq 0.5 \\
\lambda P[\omega = R|S,s] & \text{if } P[\omega = R|S,s] < 0.5
\end{cases}$$

Here, $\lambda \in [0,1)$ captures the weight that the agent puts on holding more favorable posteriors. to which the agent gains utility from consuming favorable news as opposed to a signal that allows her to have a posterior which is favorable to her preferred party.$^{21}$

Based on the utility function above, there are three interim posteriors (generated after the agent observes her Inside signal, and before she observes the Outside signal) which

---

$^{21}$ $\lambda = 0$ is the benchmark model. $\lambda = 1$ corresponds to a case where the agent’s utility is linear in how favorable her posterior belief is towards her party.
are key to solving the agent’s problem. These are the same as in the benchmark model.

\[
P[\omega = R | S = G] = \frac{k}{1 + k - m}
\]
\[
P[\omega = R | S = B] = \frac{1 - k}{1 + m - k}
\]
\[
P[\omega = R | S = T] = 0
\]

As in the benchmark model, only two signal structures are required to solve the agent’s problem, GT and GB. Further, the trade-off between the two signals remains unchanged. Specifically, the agent chooses GT if the marginal utility to price ratio of the G signal is at least as much as that of the B signal. This simplifies to:

\[
\lambda \frac{k}{1 + k - m} + (1 - \lambda) \geq \lambda \frac{(1-k)}{1 + m - k} + (1 - \lambda) \frac{2m(1-k)}{1 + m - k}
\]
\[
\lambda + \frac{(1 - \lambda)(1 + k - m)}{k} \geq \lambda + (1 - \lambda)(2m)
\]
\[
1 + k - m - 2km \geq 0
\]

This condition is identical to the one in the benchmark model. Therefore, the agent’s choice of Inside signal structure remains unchanged in this modified model. This also implies that the results of this modified model are identical to the benchmark for all values of \( \lambda \in [0, 1) \).

### A.5 Robustness to gain from being correct

Here, we show for the general model that the key results are robust to including a small gain from being correct in the utility function for agents. In addition to a gain from political faith preservation, agents also gain utility from being correct about the state of the world. The utility function of a type-R agent is, therefore, modified to:
\[
U_R = \begin{cases} 
(1 - \gamma) + \gamma, & \text{if } \mathbb{P}[\omega = R|S, s] \geq 0.5 \text{ and } \omega = R \\
(1 - \gamma), & \text{if } \mathbb{P}[\omega = R|S, s] \geq 0.5 \text{ and } \omega = L \\
\gamma, & \text{if } \mathbb{P}[\omega = R|S, s] < 0.5 \text{ and } \omega = L \\
0, & \text{if } \mathbb{P}[\omega = R|S, s] < 0.5 \text{ and } \omega = R 
\end{cases}
\]

(10)

Here, \( \gamma \in (0, 1) \) captures the extent to which the agent gains utility from consuming favorable news as opposed to a signal that allows her to have a posterior which is favorable to her preferred party.\(^{22}\)

Here, we show two key results for an agent of type \( R \). First, that for a low value of \( \gamma \), \( GT \) and \( GB \) signal structures are sufficient to solve the agent’s problem. Second, that as \( \gamma \) increases, the threshold at which the agent switches from a \( GT \) to a \( GB \) signal changes continuously. These two results together imply that the key results in this model are robust to small values of \( \gamma \). For an agent of type \( L \), the same results hold analogously.

Based on the utility function above, there are four interim posteriors (generated after the agent observes her Inside signal, and before she observes the Outside signal) which are key to solving the agent’s problem. Three of those interim posteriors come from the \( G \), \( B \), and \( T \) signals, while the fourth is an \textit{Excellent} or \( E \) signal. The signals must be such that:

\[
\begin{align*}
\mathbb{P}[\omega = R|S = G] &= \frac{k}{1 + k - m} \\
\mathbb{P}[\omega = R|S = B] &= \frac{1 - k}{1 + m - k} \\
\mathbb{P}[\omega = R|S = T] &= 0 \\
\mathbb{P}[\omega = R|S = E] &= 1
\end{align*}
\]

The set of signal structure that satisfy the Martingale property is

\[ \{ET, GT, EB, GB, EBT, EGT, GBT, EGBT\} \]

We disregard signal structures \( EBT, EGT, GBT \), and \( EGBT \) because whenever one of them is optimal, a simpler signal structure would also be optimal.

In the next step, we compare the ex-ante expected utilities of each of these signal struc-

\(^{22}\) \( \gamma = 0 \) is the benchmark model. \( \gamma = 1 \) corresponds to a case where the agent only wants to know the correct state. In that case, the agent will choose a fully revealing echo chamber signal structure.
tures to show that for low values of $\gamma$, $GT$ and $GB$ are sufficient to solve the agent’s problem.

**Signal structure $ET$**

\[ E[U_R|ET\cap \omega=R] = 1, \quad E[U_R|ET\cap \omega=L] = \gamma \]

\[ E[U_R|ET] = \frac{1+\gamma}{2} \]

**Signal structure $EB$**

\[ E[U_R|EB\cap \omega=R] = 1 - \frac{(1-k)(1-m)}{m}, \quad E[U_R|EB\cap \omega=L] = 1 - \gamma - m + 2\gamma m \]

\[ E[U_R|EB] = 1 - \frac{1}{2} \left[ \frac{(1-k)(1-m)}{m} + \gamma + m - 2\gamma m \right] \]

**Signal structure $GT$**

\[ E[U_R|GT\cap \omega=R] = 1, \quad E[U_R|GT\cap \omega=L] = \gamma + \frac{1-m}{k} - 2\gamma \frac{1-m}{k} \]

\[ E[U_R|GT] = \frac{1+\gamma}{2} + \left( \frac{1-m}{k} \right) \left( \frac{1}{2} - \gamma \right) \]

**Signal structure $GB$**

\[ E[U_R|GB\cap \omega=R] = k + m - km, \quad E[U_R|GB\cap \omega=L] = \gamma + (1-2\gamma)(1-km) \]

\[ E[U_R|GB] = \frac{1-\gamma+k+m}{2} - (1-\gamma)km \]

**Claim 5** Suppose $\gamma \leq 0.5$. The signal structures $GT$ and $GB$ are sufficient to solve the agent’s problem.

**Proof.**

We show that the signal structure $GT$ provides the agent at least as much expected utility as $ET$ or $EB$ as long as $\gamma \leq 0.5$.

\[ E[U_R|GT] - E[U_R|ET] \geq 0 \] simplifies to \( \left( \frac{1-m}{k} \right) \left( \frac{1}{2} - \gamma \right) \geq 0 \). This holds if \( \gamma \leq 0.5 \).

Similarly, \( E[U_R|GT] - E[U_R|EG] \geq 0 \) simplifies to \( (1-m)(1-k) \left( \frac{1-2\gamma}{k} + \frac{1}{k} \right) \geq 0 \). This holds if \( \gamma \leq 0.5 \).
Because $E[U_R|_{GT}]$ and $E[U_R|_{GB}]$ are continuous functions of $\gamma$, the agent’s choice of Inside signal structure depends on a threshold that varies continuously with $\gamma$. Therefore, for low values of $\gamma$, the region of information aggregation failure doesn’t disappear completely, although that region might shrink.

B Applications

B.1 Asymmetric exposure to unbiased outside signal

An agent of type $L$ receives an outside signal such that

$$P[s = l|\omega = L] = t \in [0.5, 1], \quad P[s = r|\omega = R] = t \in [0.5, 1]$$

An agent of type $R$ receives a less precise outside signal.

$$P[s = l|\omega = L] = \tau \cdot t + (1 - \tau) \cdot \frac{1}{2} \in [0.5, 1], \quad P[s = r|\omega = R] = \tau \cdot t + (1 - \tau) \cdot \frac{1}{2} \in [0.5, 1]$$

where $\tau \in [0, 1]$

B.1.1 Type-$R$ agent

We use the results developed in appendix A.1 and simply plug in $k = m = \frac{1-\tau}{2} + \tau t$.

Signal structure $GT$: Chosen if $\tau \cdot t + (1 - \tau) \cdot \frac{1}{2} \leq \frac{1}{\sqrt{2}} \simeq 0.71$.

The agent’s expected utilities are the same as the likelihood of political faith preservation for the agent, and they equal:

$$E[U_R|_{GT \cap \omega = R}] = 1, \quad E[U_R|_{GT \cap \omega = L}] = \frac{1 + \tau - 2\tau t}{1 - \tau + 2\tau t}$$

Unconditioned on the realization of the state, the ex-ante expected utility is

$$E[U_R|_{GT}] = \frac{1}{1 - \tau + 2\tau t}$$

Signal structure $GB$: Chosen if $\tau \cdot t + (1 - \tau) \cdot \frac{1}{2} \geq \frac{1}{\sqrt{2}} \simeq 0.71$.

The agent’s expected utilities are the same as the likelihood of political faith preserva-
tion for the agent, and they equal:

\[
E[UR|GB \cap \omega = R] = \frac{3}{4} + \tau t - \frac{\tau}{2} - \tau^2 \left(\frac{1}{4} + t^2 - t\right), \quad E[UR|GB \cap \omega = L] = 1 - \left(\frac{1 - \tau + 2\tau t}{2}\right)^2
\]

Unconditioned on the realization of the state, the ex-ante expected utility is

\[
E[UR] = \frac{3}{4} - \tau^2 \left(t^2 + \frac{1}{4} - t\right)
\]

B.1.2 Type-\(L\) agent

Signal structure \(GT\): Chosen if \(t \leq \frac{1}{\sqrt{2}} \approx 0.71\).

\[
E[UL|GT \cap \omega = R] = \frac{1 - t}{t}, \quad E[UL|GT \cap \omega = L] = 1
\]

Signal structure \(GB\): Chosen if \(t \geq \frac{1}{\sqrt{2}} \approx 0.71\).

\[
E[UL|GB \cap \omega = R] = 1 - t^2, \quad E[UL|GB \cap \omega = L] = 2t - t^2
\]

B.1.3 Proof of proposition \([1]\)

Proof.

Suppose that \(\tau = 0.04\) and \(t = 0.75\). A type-\(R\) agent chooses a \(GT\) signal structure because \(\tau \cdot t + (1 - \tau) \cdot \frac{1}{2} = 0.51 < \frac{1}{\sqrt{2}}\). A type-\(L\) agent chooses a \(GB\) signal structure because \(t = 0.75 > \frac{1}{\sqrt{2}}\). These conditions hold in the neighborhood of the parameter values \(\tau = 0.04\) and \(t = 0.75\). Since the likelihood of political faith preservation is continuous in these parameter values, and since the margin of victory for party \(R\) in state \(L\) at these parameter values is bounded away from zero, the result of information aggregation failure holds for a non-trivial subset of values.

When both types of agents choose a \(GT\) type signal structure. If the state of the world is \(R\), then all type-\(R\) agents vote for party \(R\), and some type-\(L\) agents vote for party \(R\). If the state of the world is \(L\), then all type-\(L\) agents vote for party \(L\) and some type-\(R\) agents vote for party \(L\). Clearly, the correct party wins in either state.

It is not possible, given the parameters, to have any parameter space where type-\(R\) agents choose a \(GB\) signal structure while type-\(L\) agents choose a \(GT\) signals structure.

When both types of agents choose a \(GB\) signal structure, if the state of the world is \(R\), then party \(R\) wins if
\[
\frac{3}{4} + \tau t - \frac{\tau}{2} - \tau^2 \left( \frac{1}{4} + t^2 - t \right) > 1 - t^2
\]

This simplifies to

\[
\tau \left( t - \frac{1}{2} \right) \left( 1 - \tau \left( t - \frac{1}{2} \right) \right) + \left( t^2 - \frac{1}{4} \right) > 0
\]

which always holds

Similarly, if the state of the world is \( L \), then party \( L \) wins if

\[
1 - \left( \frac{1 - \tau + 2\tau t}{2} \right)^2 < 2t - t^2
\]

which simplifies to

\[
\left( \frac{1 - \tau + 2\tau t}{2} \right)^2 - (1 - t)^2 > 0
\]

This always holds because \( 1 - t \in [0, 0.5] \) while \( \left( \frac{1 - \tau + 2\tau t}{2} \right) \in [0.5, 1] \).

B.2 Distrust in mainstream media

Suppose that type-\( R \) agents believe (incorrectly) that the media is less precise than it actually is. In particular, an agent of type \( L \) correctly believes that the process generating the outside signal is such that

\[
P[s = l | \omega = L] = t \in [0.5, 1], \quad P[s = r | \omega = R] = t \in [0.5, 1]
\]

An agent of type \( R \) incorrectly believes that the process that generates the outside signal is more noisy

\[
P[s = l | \omega = L] = \tau \cdot t + (1 - \tau) \cdot \frac{1}{2} \in [0.5, 1], \quad P[s = r | \omega = R] = \tau \cdot t + (1 - \tau) \cdot \frac{1}{2} \in [0.5, 1]
\]

where \( \tau \in [0, 1] \)

This is very similar to an asymmetry in exposure to mainstream media. As such, the signal choices and expected\(^{23}\) utilities are identical to those calculated in subsection B.1.

\(^{23}\)For type-\( R \) agents, the expectations are based on incorrect beliefs. We will, therefore, separately calculate the probability of political faith preservation.
For type-$L$ agents, the probability of political faith preservation is identical to the expected utilities calculated in subsection B.1.

For type-$R$ agents, if the chosen signal is of type $GT$, then the probability of political faith preservation is identical to the expected utility calculated in subsection B.1. If the chosen signal is of type $GB$, then the type-$R$ agent’s probability of political faith preservation is calculated conditioned on the state.

If the state is $\omega = R$

$$\mathbb{P}[\text{PFP}_{R}|_{GB \cap \omega = R}] = \mathbb{P}[G|GB \cap \omega = R] + \mathbb{P}[r|B \cap \omega = R] \cdot \mathbb{P}[B|GB \cap \omega = R]$$

$$= \left(\frac{1-\tau}{2} + \tau t\right) + t \cdot \left(\frac{1+\tau}{2} - \tau t\right)$$

$$= \frac{1}{2} \left(1 - \tau + t + 3\tau t - 2\tau t^2\right)$$

If the state is $\omega = L$

$$\mathbb{P}[\text{PFP}_{R}|_{GB \cap \omega = L}] = \mathbb{P}[G|GB \cap \omega = L] + \mathbb{P}[r|B \cap \omega = L] \cdot \mathbb{P}[B|GB \cap \omega = L]$$

$$= \left(\frac{1+\tau}{2} - \tau t\right) + (1-t) \cdot \left(\frac{1-\tau}{2} + \tau t\right)$$

$$= \frac{1}{2} \left(2 - t + \tau t - 2\tau t^2\right)$$

### B.3 Non-common priors

Suppose that type-$R$ and $L$ have different priors. Specifically

$$\mathbb{P}_R[\omega = R] := w_R \quad \quad \mathbb{P}_L[\omega = R] := w_L$$

where $w_L < 0.5 < w_R$

As in subsection 4.1 an agent of type $L$ receives an outside signal such that

$$\mathbb{P}[s = l|\omega = L] = t \in [0.5, 1], \quad \mathbb{P}[s = r|\omega = L] = t \in [0.5, 1]$$

An agent of type $R$ receives a less precise outside signal.

$$\mathbb{P}[s = l|\omega = L] = \tau \cdot t + (1-\tau) \cdot \frac{1}{2} \in [0.5, 1], \quad \mathbb{P}[s = r|\omega = R] = \tau \cdot t + (1-\tau) \cdot \frac{1}{2} \in [0.5, 1]$$

where $\tau \in [0, 1]$
B.3.1 Type-\(R\) agent

We use the results developed in appendix A.1 and allow for \(w_R > 0.5\) while plugging in \(k = m = \frac{1 - \tau}{2} + \tau t\).

Signal structure \(N\): Chosen if \(\tau \cdot t + (1 - \tau) \cdot \frac{1}{2} \leq w_R\). The agent is able to preserve her political faith regardless of the realization of the outside signal.

\[
E[U_R | N \cap \omega = R] = 1, \quad E[U_R | N \cap \omega = L] = 1
\]

Signal structure \(GT\): Chosen if \(w_R < \tau \cdot t + (1 - \tau) \cdot \frac{1}{2} \leq \frac{1}{\sqrt{2}} \approx 0.71\).

The agent’s expected utilities are the same as the likelihood of political faith preservation for the agent, and they equal:

\[
E[U_R | GT \cap \omega = R] = 1, \quad E[U_R | GT \cap \omega = L] = \frac{w_R(1 + \tau - 2\tau t)}{(1 - w_R)(1 - \tau + 2\tau t)}
\]

Signal structure \(GB\): Chosen if \(\tau \cdot t + (1 - \tau) \cdot \frac{1}{2} \geq \max\{\frac{1}{\sqrt{2}}, w_R\}\).

The agent’s expected utilities are the same as the likelihood of political faith preservation for the agent, and they equal:

\[
E[U_R | GB \cap \omega = R] = \frac{(2\tau t + 1 - \tau)(2\tau t - 1 - \tau)}{4w_R\tau(2t - 1)} + \frac{(2\tau t + 1 - \tau)(2\tau t + 1 - \tau - 2w_R)(1 + \tau - 2\tau t)}{8w_R\tau(2t - 1)}
\]

and

\[
E[U_R | GB \cap \omega = L] = \frac{(1 + \tau - 2\tau t)(2\tau t + 1 - \tau)}{4(1 - w_R)\tau(2t - 1)} + \frac{(2\tau t + 1 - \tau)(2\tau t + 1 - \tau - 2w_R)(1 + \tau - 2\tau t)}{8(1 - w_R)\tau(2t - 1)}
\]

B.3.2 Type-\(L\) agent

Signal structure \(N\): Chosen if \(t \leq 1 - w_L\).

The agent is able to preserve her political faith regardless of the realization of the outside signal.

\[
E[U_L | N \cap \omega = R] = 1, \quad E[U_L | N \cap \omega = L] = 1
\]
Signal structure $GT$: Chosen if $1 - w_L < t \leq \frac{1}{\sqrt{2}} \simeq 0.71$.

\[
E[U_L|_{GT \cap \omega = R}] = \frac{(1 - w_L)(1 - t)}{w_L t}, \quad E[U_L|_{GT \cap \omega = L}] = 1
\]

Signal structure $GB$: Chosen if $t \geq \max\{\frac{1}{\sqrt{2}}, 1 - w_L\}$.

\[
E[U_L|_{GB \cap \omega = R}] = \frac{(1 - t)(1 - w)}{w(2t - 1)} + \frac{t(1 - t)(t + w - 1)}{w(2t - 1)}
\]

and

\[
E[U_L|_{GB \cap \omega = L}] = \frac{t(2t - 1 - t^2 + (1 - w)t)}{(1 - w)(2t - 1)}
\]

B.4 Propagandized outside signal with censorship

Suppose that the outside signal is biased towards party $L$. Suppose, also, that this bias is common knowledge. The true structure of the outside signal is

\[
\mathbb{P}[s = l|\omega = L] = \tau \cdot t + (1 - \tau) \cdot 1, \quad \mathbb{P}[s = r|\omega = R] = \tau \cdot t + (1 - \tau) \cdot 0
\]

where $t > 0.5$.

We can use the results from appendix A.1 and A.2 by simply plugging in $k = 1 + \tau t - \tau$ and $m = \tau t$.

Suppose, also, that neither type of agent has access to any information other than her prior and the realized outside signal. This scenario may be interpreted either as the party in power shutting down all media other than the propagandized state-controlled media. An alternative interpretation is that the media environment is not rich because we are considering a time before the environment became rich through the proliferation of news outlets and social media.

Because agents share common and symmetric priors, their posteriors on observing the outside signals are:

\[
Pr[\omega = L|s = l] = \frac{1 + \tau t - \tau}{2 - \tau}
\]

Because $t > 0.5$, $Pr[\omega = L|s = l] > 0.5$

and
Because \( t > 0.5 \), \( Pr[\omega = L|s = r] < 0.5 \). For party \( L \) to win in state \( R \), we require that \( Pr(s = l|\omega = R) > Pr(s = r|\omega = R) \) or \( 1 - \tau t > \tau t \). That is, \( \tau t < 0.5 \).

For party \( L \) to win in state \( L \), we require that \( Pr(s = l|\omega = L) > Pr(s = r|\omega = L) \) or \( 1 + \tau t - \tau > \tau - \tau t \). That is, \( \tau (1 - t) < 0.5 \). If \( \tau \in [0,1] \), and \( t > 0.5 \), or if \( \tau \in (0,1] \) and \( t \geq 0.5 \) then this always holds.

This benchmark assumes that there is no chosen signal structure - it demonstrates the strength of propaganda when it is accompanied by censorship that prevents agents from choosing a signal structure to inform them.

Next, we re-introduce the ability of agents to design a chosen signal structure. Our interest is primarily in the region where propaganda can backfire once agents are able to choose a signal structure.

### B.5 Propagandized outside signal without censorship

Suppose that the outside signal is biased towards party \( L \). Suppose, also, that this bias is common knowledge. The true structure of the outside signal is

\[
P[s = l|\omega = L] = \tau \cdot t + (1 - \tau) \cdot 1,
\]

\[
P[s = r|\omega = R] = \tau \cdot t + (1 - \tau) \cdot 0
\]

where \( t > 0.5 \).

We can use the results from appendix A.1 and A.2 by simply plugging in \( k = 1 + \tau t - \tau \) and \( m = \tau t \).

#### B.5.1 Type-\( R \) agent’s problem

**GT signal structure:** Chosen if \( 2 - \tau - 2\tau t - 2\tau^2t^2 + 2\tau^2t \geq 0 \).

The agent’s expected utilities are:

\[
E[U_R|GT \cap \omega = R] = 1,
\]

\[
E[U_R|GT \cap \omega = L] = \frac{1 - \tau t}{1 + \tau t - \tau}
\]

**GB signal structure:** Chosen if \( 2 - \tau - 2\tau t - 2\tau^2t^2 + 2\tau^2t \leq 0 \).

The agent’s expected utilities are:

\[
E[U_R|GB \cap \omega = R] = 1 - \tau + \tau t + \tau^2 t - \tau^2 t^2,
\]

\[
E[U_R|GB \cap \omega = L] = 1 - \tau t + \tau^2 t - \tau^2 t^2
\]
B.5.2 Type-\(L\) agent’s problem

**GT signal structure:** Chosen if \(\tau(1 - 2t - 2\tau t^2 + 2\tau t) \geq 0\) The agent’s expected utilities are:

\[
E[U_L|_{GT \cap \omega = R}] = \frac{1-t}{t}, \quad E[U_L|_{GT \cap \omega = L}] = 1
\]

**GB signal structure:** Chosen if \(\tau(1 - 2t - 2\tau t^2 + 2\tau t) \leq 0\) The agent’s expected utilities are:

\[
E[U_L|_{GB \cap \omega = R}] = 1 - \tau t + \tau^2 t - \tau^2 t^2, \quad E[U_L|_{GB \cap \omega = L}] = 1 - \tau + \tau t + \tau^2 t - \tau^2 t^2
\]

B.5.3 Proof of proposition \(^2\)

**Proof.**

Suppose that \(\tau = 0.3\) and \(t = 0.6\). A type-\(R\) agent chooses a GT signal structure because \(2 - \tau - 2\tau t - 2\tau^2 t^2 + 2\tau^2 t = 1.3832 > 0\). A type-\(L\) agent chooses a GB signal structure because \(\tau(1 - 2t - 2\tau t^2 + 2\tau t) = -0.0168 < 0\). These conditions hold in the neighborhood of the parameter values \(\tau = 0.3\) and \(t = 0.6\). Since the likelihood of political faith preservation is continuous in these parameter values, and since the margin of victory for party \(R\) in state \(L\) at these parameter values is bounded away from zero, the result of information aggregation failure holds for a non-trivial subset of values.

When both types of agents choose a GT type signal structure. If the state of the world is \(R\), then all type-\(R\) agents vote for party \(R\), and some type-\(L\) agents vote for party \(R\). If the state of the world is \(L\), then all type-\(L\) agents vote for party \(L\) and some type-\(R\) agents vote for party \(L\). Clearly, the correct party wins in either state.

It is not possible, given the parameters, to have any parameter space where type-\(R\) agents choose a GB signal structure while type-\(L\) agents choose a GT signals structure.

When both types of agents choose a GB signal structure, if the state of the world is \(R\), then party \(R\) wins if

\[
1 - \tau + \tau t + \tau^2 t - \tau^2 t^2 > 1 - \tau t + \tau^2 t - \tau^2 t^2
\]

This simplifies to

\[
\tau(2t - 1) > 0
\]

which always holds
Similarly, if the state of the world is $L$, then party $L$ wins if

$$1 - \tau + \tau t + \tau^2 t - \tau^2 t^2 > 1 - \tau t + \tau^2 t - \tau^2 t^2$$

which simplifies to

$$\tau(2t - 1) > 0$$

Which always holds.

\[\blacksquare\]

**B.6 Propaganda with oblivious citizens**

Suppose that the true process of the outside signal is

$$\mathbb{P}[s = l|\omega = L] = \tau \cdot t + (1 - \tau) \cdot 1, \quad \mathbb{P}[s = r|\omega = L] = \tau \cdot t + (1 - \tau) \cdot 0$$

where $\tau \in [0, 1]$ and $t \in [0.5, 1]$ Agents don’t know that the true process, and believe that the process that generates the outside signal is unbiased and is as follows

$$\mathbb{P}[s = l|\omega = L] = t, \quad \mathbb{P}[s = r|\omega = R] = t$$

**B.6.1 Type R agent’s problem**

**GT signal structure:** Chosen if $t \leq \frac{1}{\sqrt{2}}$

The agent’s expected utilities are:

$$E[U_R|GT \cap \omega = R] = 1, \quad E[U_R|GT \cap \omega = L] = \frac{1-t}{t}$$

When the agent chooses a GT type signal structure, the likelihood of political faith preservation doesn’t depend on the realization of the outside signal. Therefore, the likelihood of political faith preservation is the same as the expected utility.

**GB signal structure:** Chosen if $t \geq \frac{1}{\sqrt{2}}$ The agent’s expected utilities (under incorrect beliefs) are:

$$E[U_R|GB \cap \omega = R] = 2t - t^2, \quad E[U_R|GB \cap \omega = L] = 1 - t^2$$
Here, the likelihood of political faith preservation differs from the agent’s expected utility.

\[ \mathbb{P}[\text{PFP}_R|\text{GB}\cap\omega=R] = t(1+\tau-\tau t), \quad \mathbb{P}[\text{PFP}_R|\text{GB}\cap\omega=L] = (1+\tau t)(1-t) \]

**B.6.2 Type L agent’s problem**

*GT signal structure:* Chosen if \( t \leq \frac{1}{\sqrt{2}} \)

The agent’s expected utilities are:

\[
E[U_L|\text{GT}\cap\omega=R] = \frac{1-t}{t}, \quad E[U_L|\text{GT}\cap\omega=L] = 1
\]

When the agent chooses a GT type signal structure, the likelihood of political faith preservation doesn’t depend on the realization of the outside signal. Therefore, the likelihood of political faith preservation is the same as the expected utility.

*GB signal structure:* Chosen if \( t \geq \frac{1}{\sqrt{2}} \)

The agent’s expected utilities (under incorrect beliefs) are:

\[
E[U_L|\text{GB}\cap\omega=R] = 1-t^2, \quad E[U_L|\text{GB}\cap\omega=L] = 2t-t^2
\]

Here, the likelihood of political faith preservation differs from the agent’s expected utility.

\[ \mathbb{P}[\text{PFP}_L|\text{GB}\cap\omega=R] = 1-\tau t^2, \quad \mathbb{P}[\text{PFP}_L|\text{GB}\cap\omega=L] = 1-\tau(1-t)^2 \]

**B.6.3 Proof of proposition 3**

Proof.

Suppose that \( \tau = 0.1 \) and \( t = 0.75 \). Both types of agents choose a GB signal structure for their In-media because \( t > \frac{1}{\sqrt{2}} \). These conditions hold in the neighborhood of the parameter values \( \tau = 0.1 \) and \( t = 0.75 \). Since the likelihood of political faith preservation is continuous in these parameter values, and since the margin of victory for party L in state R at these parameter values is bounded away from zero, the result of information aggregation failure holds for a non-trivial subset of values.

When both types of agents choose a GT type signal structure. If the state of the world is R, then all type-R agents vote for party R, and some type-L agents vote for party R. If the state of the world is L, then all type-L agents vote for party L and some type-R agents vote for party L. Clearly, the correct party wins in either state.
It is not possible, given the parameters, to have any parameter space the two types of agents choose different signal structures for their respective In-media consumption.

When both types of agents choose a $GB$ signal structure, if the state of the world is $L$, then party $L$ wins if

$$1 - \tau(1 - t)^2 > (1 + \tau t)(1 - t)$$

This simplifies to

$$t + \tau t - \tau > 0$$

which always holds

Further, if the state of the world is $R$, then party $L$ wins if

$$1 - \tau t^2 > t(1 + \tau - \tau t)$$

which simplifies to

$$t < \frac{1}{1 + \tau}$$

Therefore, party $L$ wins regardless of the state if $t \in (\frac{1}{\sqrt{2}}, \frac{1}{1 + \tau})$. There are no parameter values where party $R$ can win in state $L$.

\[\blacksquare\]

### B.7 Perceived propaganda

Suppose that the true process of the outside signal is

$$\mathbb{P}[s = l|\omega = L] = t \in [0.5, 1], \quad \mathbb{P}[s = r|\omega = R] = t \in [0.5, 1]$$

$L$ type agents know the true process, while type-$R$ agents believe that the process of the outside signal is biased in the following way

$$\mathbb{P}[s = l|\omega = L] = \tau \cdot t + (1 - \tau) \cdot 1, \quad \mathbb{P}[s = r|\omega = R] = \tau \cdot t + (1 - \tau) \cdot 0$$

where $\tau \in [0, 1]$

#### B.7.1 Type $R$ agent’s problem

$GT$ signal structure: Chosen if $2 - \tau - 2\tau t - 2\tau^2 t^2 + 2\tau^2 t \geq 0$
The agent’s expected utilities are:

\[
E[U_R | GT \cap \omega = R] = 1, \quad E[U_R | GT \cap \omega = L] = \frac{1 - \tau t}{1 + \tau t - \tau}
\]

When the agent chooses a GT type signal structure, the likelihood of political faith preservation doesn’t depend on the realization of the outside signal. Therefore, the likelihood of political faith preservation is the same as the expected utility.

**GB signal structure:** Chosen if \(2 - \tau - 2\tau t - 2\tau^2 t^2 + 2\tau^2 t \leq 0\) The agent’s expected utilities (under incorrect beliefs) are:

\[
E[U_R | GB \cap \omega = R] = 1 - \tau + \tau t + \tau^2 t - \tau^2 t^2, \quad E[U_R | GB \cap \omega = L] = 1 - \tau + \tau^2 t - \tau^2 t^2
\]

Here, the likelihood of political faith preservation differs from the agent’s expected utility.

\[
P[PFP_R | GB \cap \omega = R] = 1 - \tau + 2\tau t - \tau^2 t^2, \quad P[PFP_R | GB \cap \omega = L] = 1 - \tau t^2
\]

**B.7.2 Type L agent’s problem**

**GT signal structure:** Chosen if \(t \leq \frac{1}{\sqrt{2}}\)

The agent’s expected utilities are:

\[
E[U_L | GT \cap \omega = R] = \frac{1 - t}{t}, \quad E[U_L | GT \cap \omega = L] = 1
\]

**GB signal structure:** Chosen if \(t \geq \frac{1}{\sqrt{2}}\)

The agent’s expected utilities are:

\[
E[U_L | GB \cap \omega = R] = 1 - t^2, \quad E[U_L | GB \cap \omega = L] = 2t - t^2
\]

**B.7.3 Proof of proposition**

**Proof.**

Suppose that \(\tau = 0.1\) and \(t = 0.75\). A type-R agent chooses a GT signal structure because \(2 - \tau - 2\tau t - 2\tau^2 t^2 + 2\tau^2 t \simeq 1.75 > 0\). A type-L agent chooses a GB signal structure because \(t = 0.75 > \frac{1}{\sqrt{2}}\). These conditions hold in the neighborhood of the parameter values \(\tau = 0.1\) and \(t = 0.75\). Since the likelihood of political faith preservation is continuous in these parameter values, and since the margin of victory for party R in
state $L$ at these parameter values is bounded away from zero, the result of information aggregation failure holds for a non-trivial subset of values.

When both types of agents choose a $GT$ type signal structure. If the state of the world is $R$, then all type-$R$ agents vote for party $R$, and some type-$L$ agents vote for party $R$. If the state of the world is $L$, then all type-$L$ agents vote for party $L$ and some type-$R$ agents vote for party $L$. Clearly, the correct party wins in either state.

It is not possible, given the parameters, to have any parameter space where type-$R$ agents choose a $GB$ signal structure while type-$L$ agents choose a $GT$ signals structure.

When both types of agents choose a $GB$ signal structure, if the state of the world is $R$, then party $R$ wins if

$$1 - \tau + 2\tau t - \tau t^2 > 1 - t^2$$

This simplifies to

$$\tau(2t - 1) + t^2(1 - \tau) > 0$$

which always holds

Similarly, if the state of the world is $L$, then party $L$ wins if

$$2t - t^2 > 1 - \tau t^2$$

which simplifies to

$$(2t - 1) - t^2(1 - \tau) > 0$$

Which holds for values of $t \in \left[\frac{1}{\sqrt{2}}, 1\right]$ and $\tau \in \left[\frac{2}{3}, 1\right]$. The region where both types of agents choose a $GB$ signal structure is a subset of the region where $t \in \left[\frac{1}{\sqrt{2}}, 1\right]$ and $\tau \in \left[\frac{2}{3}, 1\right]$. Therefore, if both types of agents choose a $GB$ signal structure, then the correct party wins.

\[\blacksquare\]