Can Media Pluralism Be Harmful to News Quality?*

Federico Innocenti†

June 23, 2022

Abstract

I study a Bayesian persuasion model that connects two stylized facts characterizing the Internet: a great diversity of news sources and the proliferation of disinformation. I show that media pluralism reduces information quality when news consumers have limited attention because of the endogenous formation of echo chambers. According to the standard narrative, echo chambers arise because news consumers exhibit confirmation bias. I show that even unbiased and rational news consumers devote their limited attention to like-minded news sources in equilibrium. Confirmation bias thus arises endogenously because news sources have no incentive to provide valuable information.

Keywords: Bayesian Persuasion, Echo Chambers, Heterogeneous Beliefs, Limited Attention, Media Bias, Media Pluralism.

JEL Classification: D82, D83, L82

---

*I am grateful to Volker Nocke, Martin Peitz and Thomas Tröger for their encouragement and support. I thank Michele Bisceglia, Carl-Christian Groh, Melika Liporace, Niccolò Lomys, Nicolas Schutz, Konrad Stahl and André Stenzel for their helpful comments. I thank seminar participants in Mannheim and Toulouse as well as at the Louvain Economics of Digitization seminar, the 14th RGS Doctoral Conference, the 2021 MaCCI Annual Conference, the 17th CEPR/JIE School on Applied Industrial Organisation, the 2021 North American Summer Meeting of the Econometric Society, the Bavarian Young Economist Meeting 2021, the 8th CRC TR 224 Conference, the Annual Research Conference 2021 of the European Commission, the 2021 European Winter Meeting of the Econometric Society, the Asia-Pacific Industrial Organization Conference 2021, the Spring Meeting of Young Economists 2022 and the Oligo Workshop 2022. I gratefully acknowledge funding by the German Research Foundation (DFG) through CRC TR 224 (Project B05).

†University of Mannheim, Department of Economics and MaCCI, email: f.innocenti93@gmail.com.

1. Introduction

A critical problem for modern democracies is that those who control the information flow can influence political and economic outcomes. Ideally, the presence of competing sources of information is beneficial. The more information an individual can receive, the more she knows about the issue, and the smaller is the influence of a particular source. For a long time, the Internet has been considered a very effective way to guarantee pluralism in information (Keen, 2015). But is competition among news sources on the Internet undoubtedly beneficial? Empirical evidence suggests a deterioration of the quality of the information at one's disposal. For instance, it is hard to find reliable online information about health conditions (Swire-Thompson and Lazer, 2019). More generally, conspiracy theories and "fake news" proliferate online.[1] I suggest a novel explanation for the deterioration of information quality online: the endogenous formation of echo chambers even when news consumers are unbiased.

The Cambridge dictionary defines an echo chamber as "a situation in which people only hear opinions that are similar to their own". Echo chambers are a prominent feature of the Internet. Online networks show high homophily: an individual learns from those who share her worldview (Del Vicario et al., 2016; Halberstam and Knight, 2016). The existence of echo chambers is a policy concern, as it endangers meaningful debate in a democracy. Within echo chambers, each individual never questions her beliefs. As a consequence, society divides into opposing factions. Moreover, the presence of echo chambers affects the quality of news. As I show, the media have no incentive to provide informative news in echo chambers.

The standard explanation for the existence of echo chambers is preference-based, namely that individuals are subject to confirmation bias. Nickerson (1998) defines confirmation bias as "seeking or interpreting of evidence in ways that are partial to existing beliefs, expectations, or a hypothesis in hand". However, there are settings where confirmation bias does not seem a reasonable explanation for echo chambers: for instance, investment decisions (Cookson et al., 2021). I provide a complementary and novel channel: even if individuals seek the most informative news, echo chambers arise because of the interplay between limited attention of news consumers with heterogeneous beliefs and media bias of news sources.[2]

I study a Bayesian persuasion model with two states of the world and two actions (see Figure 1). There are two types of agents: experts and decision-makers. Each expert is biased: his preferred action is independent of the state of the world. In stage 1, each expert designs information about the state of the world to persuade decision-makers to take the expert's preferred action. Such information is public: all decision-makers that devote attention to the expert observe the same information. Each decision-maker is unbiased: she wants to match her action with the state. Decision-makers have partitioned into subgroups holding heterogeneous prior beliefs about the state of the world and have limited attention: each decision-maker can only devote attention to one expert. In stage 1, each decision-maker chooses which expert is worthy of attention and observes the information such an expert provides. Then, she updates her belief (stage 2) and takes the optimal action given such belief (stage 3). I show that the existence of experts with opposite biases is harmful to decision-makers when the latter strategically allocate their limited attention.

As a benchmark, I consider a single expert and two subgroups of decision-makers with

[1] Fake news are of public concern since the 2016 US presidential election (Allcott and Gentzkow, 2017). Using the taxonomy proposed by Molina et al. (2021), my model captures partisan news, misreporting and persuasive advertising. All these lie in the "grey area" between objectively real and false news.

[2] Lee et al. (2017) show that perceived information overload is positively associated with selective exposure in online news consumption. Internet users fail to discriminate news based on quality (Qiu et al., 2017). My results are in line with recent advances in psychology, showing that politically motivated reasoning does not drive selective exposure of online news consumers to confirmatory news (Pennycook and Rand, 2021).

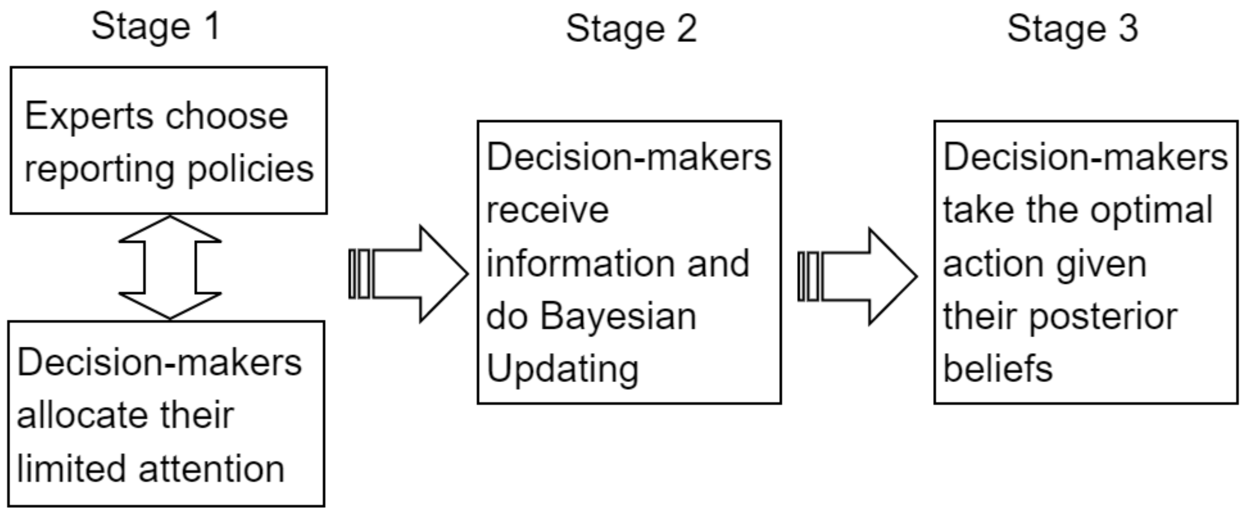


Figure 1: Stages of the game

different beliefs that I label "sceptics" and "believers". Without information, believers choose the expert's preferred action, whereas sceptics do not. Hence, the expert designs information to change sceptics' behaviour. Such information is public - i.e., all decision-makers receive the same information. Thus, any attempt to change a sceptic's belief affects a believer's belief as well. Being exposed to information could induce believers to take the expert's undesired action. Therefore, the expert faces a trade-off between persuading sceptics and retaining believers. I show that there are two candidates for the optimal information design (or reporting policy).

The *hard-news policy* focuses on persuading sceptics. For this purpose, a message must be sufficiently credible - i.e., it can be misleading only to a limited extent. Therefore, this policy entails the cost of revealing the unfavourable state to all decision-makers with positive probability. If this state is revealed, believers take the expert's undesired action.

The *soft-news policy* focuses on retaining believers. The expert sends two messages of different credibility. One is credible enough to persuade sceptics. The other one is not, but at the same time, it does not induce believers to take the expert's undesired action. With this second message, the expert leverages believers' credulity. This policy ensures that believers will continue to choose the expert's preferred action.

I show that the hard-news policy is more informative than the soft-news policy according to the order defined by Blackwell (1953). Nevertheless, the expert prefers the soft-news policy if decision-makers have sufficiently polarized beliefs. In a context of severe polarization, it is very costly to persuade sceptics. To be credible, the expert has to reveal the unfavourable state with high probability. At the same time, it is particularly tempting to retain believers because it is easy to leverage their credulity. Both these arguments imply that the soft-news policy is more favourable for the expert. A second key parameter is the expert's belief. The higher is the expert's belief of his unfavourable state, the more he values his ability to mislead (at least) believers, making the soft-news policy more appealing.

Next, I show how media pluralism (i.e., the presence of multiple experts with different biases) makes decision-makers worse off. Two experts with different preferred actions compete to persuade two subgroups of decision-makers with heterogeneous beliefs. Because of limited attention, each decision-maker can only devote attention to one expert. Therefore, each expert behaves like a monopolist given his audience. In other words, for any expert, the allocation of attention by decision-makers determines the distribution of beliefs such an expert has to confront, and his reporting policy must be optimal given such a distribution. Here, the novelty (compared to the benchmark) is the interaction between the optimal information design and the endogenous allocation of attention.

The allocation of attention depends on the policies of the experts. Each decision-maker allocates her attention to maximize her subjective probability of taking the correct action. This probability is at its minimum without information. An expert designs information to change decision-makers' behaviour. To be successful, the expert must provide sufficiently accurate information, and this makes decision-makers (weakly) better off. I define a decision-maker's information gain as the increase in her subjective probability of taking the correct action following information provision. Thus, each decision-maker allocates her attention to maximize her information gain.

It makes a difference for a decision-maker whether she is a *target* of an expert. An expert targets a subgroup of decision-makers if he tailors his policy to persuade them. For example, the sceptics are the targets when the expert uses his hard-news policy. An expert does not reveal more information than what is strictly necessary to change the behaviour of targets. Therefore, any target of a given expert receives zero information gain when devoting attention to him. Thus, each decision-maker aims to avoid being a target. At the same time, the optimal policy of each expert features (at least) one target, unless the expert faces only his believers. This tension determines which allocations of attention can support an equilibrium.

I label an equilibrium as "symmetric" if any two decision-makers of the same subgroup devote attention to the same expert. I show that the unique symmetric equilibrium of this game featuring two active experts is *echo chambers* with *babbling* (i.e., no information provision). In echo chambers, the audience of each expert is composed only of his believers. Therefore, the expert finds it optimal to leave their beliefs unchanged. Thus, babbling is the optimal policy for each expert. Given babbling by each expert, decision-makers have no incentive to deviate, as the information gain is zero in any case. In echo chambers, information quality is strictly lower than in monopoly for any decision-maker (whereas, in terms of information gains, targeted decision-makers are indifferent). This is because a monopolist uses either his hard-news policy or his soft-news policy. Both these policies produce some dispersion in posterior beliefs, hence have higher quality than babbling according to Blackwell (1953)'s criterion.

I extend the model to consider a general distribution of decision-makers' beliefs. I label an expert as "informative" if he uses either a hard-news policy or a soft-news policy. In any symmetric equilibrium, there is at most one informative expert. Indeed, if there are two informative experts, there is always (at least) one target who can get a positive information gain by changing her allocation of attention. Therefore, in any symmetric equilibrium, at least one expert is babbling. I label the audience of a babbling expert as an echo chamber. Limited attention makes media pluralism harmful to those decision-makers who cluster into an echo chamber by reducing the quality of the information they receive compared to a monopoly. In general, no decision-maker can benefit from media pluralism. For any equilibrium, there exists a monopoly outcome such that both information quality and information gain are (weakly) higher for any decision-maker.

My results show that the omnipresence of information - a characteristic of the Internet - could make all information useless. This negative result follows from the endogenous allocation of attention by decision-makers. As an extension, I study the problem of a platform that can allocate decision-makers' attention. The platform's goal is to maximize information quality. The platform can enable the coexistence of two informative experts. In particular, the platform can induce each expert to use his hard-news policy. In this way, such an altruistic platform can enhance information quality and make media pluralism beneficial.

1.1. Example

The widespread existence of misinformation about the COVID-19 vaccination provides a fitting example to illustrate my results. There are two possible states of the world: either a vaccine is safe or not (e.g., either it has long-run side effects or not). Each citizen wants to get vaccinated if and only if the vaccine is safe. Some citizens are sceptical about vaccinations being safe and are not willing to get vaccinated a priori (Paul et al., 2021). The government aims to reach herd immunity because the societal benefits of vaccination outweigh very rare private costs due to side effects. Therefore, a pro-government media wants to persuade citizens to get vaccinated.

In a monopoly, the supply of news by the pro-government media depends on its confidence about vaccinations' safety. If the pro-government media is very confident, it provides "hard evidence" (e.g., the evaluations by the European Medicines Agency based on clinical trials). The pro-government media attempts to persuade sceptics to get vaccinated because it expects persuasion to be very likely. If polarization is sufficiently high and the pro-government media is not confident enough, it also provides "soft evidence" (e.g., weaker statements such as "benefits are higher than risks"). In this way, the pro-government media is sure to retain those citizens who were already willing to get vaccinated.

In a plural environment, a no-vax media opposes vaccinations to make profits with alternative treatments (Ghoneim et al., 2020). An equilibrium could be as follows: the pro-government media produces "hard evidence", whereas the no-vax media is babbling within its echo chamber.[3] Citizens who are sceptical about vaccinations understand that the pro-government media designs information to change their attitudes. Therefore, these citizens do not benefit from the information provided by the pro-government media and thus rationally allocate their limited attention to confirmatory news. This type of news does not allow sceptics to learn about the nature of vaccinations and create a negative externality on our society. Indeed, the pro-government media cannot persuade these citizens to get vaccinated. The existence of a large no-vax echo chamber can help to explain why herd immunity is difficult to reach (Diamond et al., 2021).

The rest of the paper is organized as follows. In Section 2, I review the literature. In Section 3, I present the theoretical model. In Section 4, I study optimal information design in a monopoly. In Section 5, I describe the effects of media pluralism. In Section 6, I examine some extensions. In Section 7, I discuss the applicability of my model to the real world. In Section 8, I conclude.

2. Related Literature

I contribute to the literature by exploring how the endogenous supply of (potentially misleading) information to decision-makers with heterogeneous beliefs interacts with limited attention. Therefore, my paper connects with the following streams in the literature.

Limited attention

> "In an information-rich world, the wealth of information [...] creates a poverty of attention and a need to allocate that attention efficiently among the overabundance of information sources that might consume it." Simon (1971)

[3] Di Marco et al. (2021) find evidence of echo chambers about the COVID-19 pandemic. Jiang et al. (2021) show that segregation is stronger among far-right users.

The Internet has led to an information-rich economy as it allows news sources to reach more consumers at a lower per-consumer cost. The growth in consumers wealth and firms market power helped this process (Falkinger, 2008). Limited attention can explain many puzzling empirical patterns, for instance, asset-price dynamics (Peng and Xiong, 2006), the attraction effect (Masatlioglu et al., 2012), nominal rigidities (Matějka, 2016), persistently low inflation (Pfäuti, 2021), the superstar effect (Hefti and Lareida, 2021) and why minorities and extremists are very influential in the political process (Matějka and Tabellini, 2021).[4] In this paper, I offer new insights into the effects of limited attention. I show that limited attention can explain why rational and unbiased news consumers cluster into echo chambers and thus rationalizes the proliferation of low-quality information.

Limited attention influences price competition and advertising within and across industries (Anderson and de Palma, 2012; De Clippel et al., 2014; Hefti and Liu, 2020). My findings are complementary to Anderson and Peitz (2020), who show that increasing media diversity has the undesired effect of increasing advertising clutter and thus can make consumers worse off. Indeed, I show that media diversity can also harm news consumers by causing a reduction in information quality.

Bayesian persuasion A standard assumption in this literature - pioneered by Aumann and Maschler (1995) and Kamenica and Gentzkow (2011) - is the existence of a common prior belief. By contrast, I examine the problem of a sender (expert) who faces many receivers (decision-makers) endowed with heterogeneous beliefs.[5] In Guo and Shmaya (2019), a separating (soft-news) policy yields a higher payoff to the sender than a pooling (hard-news) policy if the receiver has sufficiently accurate private information. The distribution of private information is (strategically) equivalent to receivers holding heterogeneous beliefs. From this perspective, I show that more accurate private information can lead to less accurate public information. Indeed, if polarization is above a threshold, the sender provides information of lower quality.[6]

Gentzkow and Kamenica (2017a,b) argue that competition among senders increases information provision and benefits receivers. I show that this conclusion fails if receivers have limited attention. My model incorporates endogenous allocation of attention between senders and endogenous persuasion.[7] In Knoepfle (2020), senders compete to gather the attention of a receiver. By contrast, senders are concerned about receivers' actions in my model. This difference leads to opposite results: endogenous echo chambers in my model, whereas full revelation is the final outcome in Knoepfle (2020).

Echo chambers The existence of echo chambers is a distinctive feature of the Internet. Indeed, there is evidence of echo chambers even in non-partisan contexts such as climate

[4]Gabaix (2019) and Mackowiak et al. (2021) survey the literature on behavioural and rational inattention, respectively.

[5]Alonso and Camara (2016) study the consequences of beliefs heterogeneity between one sender and one receiver. Persuasion in my model is a special case of the general framework examined by Kolotilin et al. (2017). In my simpler model, I can explore the effect of beliefs on information quality. Beliefs are exogenous to the model, and it is beyond the purpose of this paper to study the origin of beliefs (Flynn et al., 2017). Bergemann and Morris (2019) and Kamenica (2019) survey the literature on information design.

[6]Away from this discontinuity point, higher polarization unambiguously increases the quality of media reporting. See also Gitmez and Molavi (2022).

[7]Che and Mierendorff (2019) and Leung (2020) study the problem of a receiver who has to allocate her limited attention between biased senders. In these papers, the information design is exogenous. Bloedel and Segal (2020), Lipnowski et al. (2020) and Wei (2021) study how limited attention by the receiver(s) affects optimal persuasion by a single sender.

change (Williams et al., 2015), vaccinations (Cossard et al., 2020) and the financial markets (Cookson et al., 2021). Echo chambers facilitate the proliferation of misinformation (Törnberg, 2018; Acemoglu et al., 2021). As a consequence, being part of an echo chamber affects individual behaviour. For instance, during the COVID-19 pandemic, Democrats and Republicans in the US show different attitudes towards social distancing (Allcott et al., 2020; Gollwitzer et al., 2020) and vaccinations (Fridman et al., 2021).

Jann and Schottmuller (2019) rationalize echo chambers in a many-to-many cheap talk model with biased decision-makers.[8] By contrast, even unbiased decision-makers may cluster into echo chambers in my model. Martinez and Tenev (2020) study a model where experts are unbiased. The experts are heterogeneous in terms of information precision. A decision-maker rationally infers that an expert has higher quality if he supplies information more in line with the decision-maker's belief. By contrast, experts are biased and precision is endogenous in my model. The strategic interaction between decision-makers and experts plays a crucial role in the formation of echo chambers.[9] Jann and Schottmuller (2019) and Martinez and Tenev (2020) argue that echo chambers can be helpful, either to enhance communication in a network or to separate high-quality and low-quality news. Instead, echo chambers have a negative effect in this paper. The reason is the endogenous supply of information by biased experts.

Detrimental competition in the market for news When biased media interact with rational and unbiased news consumers, the standard theoretical prediction is that media competition increases news consumers' welfare (Gentzkow et al., 2015). I show that media competition can harm news consumers when the latter have limited attention and heterogeneous beliefs. This result follows an endogenous market segmentation.[10] My contribution fits in an emerging literature about the downsides of competition in the market for news, which presents a number of complementary channels.[11] Information overload does not allow decision-makers to identify high-quality experts (Persson, 2018) and implies higher prices because consumers get lost in diversity (Hefti, 2018). Costly information acquisition or communication reduces each expert's effort in the presence of other experts: free-riding harms decision-makers (Kartik et al., 2017; Emons and Fluet, 2019). Competition increases informational specialization, which in turn increases social disagreement: in large enough societies, this reduces welfare (Perego and Yuksel, 2021). Competition also increases the pressure to publish without fact-checking because of the risk of being pre-empted (Andreottola and de Moragas Sánchez, 2020). Finally, because of the unbundling of journalism, online competition weakens the media's incentives to invest in news' quality, especially when news consumers' switching behaviour is particularly pronounced (Bisceglia, 2021).

[8]Similarly, in Giovanniello (2021) echo chambers arise because biased voters have incentives to communicate useful information only to like-minded peers.

[9]Alternatively, echo chambers may arise because the cost of processing information is increasing in its precision (Nimark and Sundaresan, 2019) or when decision-makers look for disapproving evidence eventually supplied by like-minded experts (Hu et al., 2021). Levy and Razin (2019) survey the economics literature on echo chambers.

[10]According to Mullainathan and Shleifer (2005), "reader heterogeneity is the crucial antidote to media bias". I show that reader heterogeneity exacerbates media bias when combined with limited attention.

[11]A broader literature shows that competition can backfire in many different settings. For instance, in a model where consumers have heterogeneous tastes for products, Chen and Riordan (2008) show that competition can increase prices. In particular, this can happen because consumers tend to sort themselves as customers of their preferred firms, making the demand less elastic in a duopoly than in a monopoly. This sorting effect resembles decision-makers' incentives for the allocation of attention in my paper. Similarly to Chen and Riordan (2008), sorting weakens experts' incentives to provide information. See also Spinnewijn (2013), Janssen and Roy (2014) and Heidhues et al. (2021).

3. Model

There are two states of the world and two actions. I denote with $\Omega := \{\omega_1, \omega_2\}$ the set of states and with $A := \{a_1, a_2\}$ the set of actions. Each agent l has a prior belief $\mu_l^0(\omega_1) \in (0,1)$ that the state is ω_1. Clearly, $\mu_l^0(\omega_2) = 1 - \mu_l^0(\omega_1)$ is the agent l's prior belief that the state is ω_2. There are two types of agents: experts and decision-makers. I denote with D the set of decision-makers and with J the set of experts. Decision-makers partition in homogenous subgroups: $D := \bigcup_{i \in I} D_i$ where I is the set of subgroups of decision-makers. Two decision-makers of the same subgroup share the same belief: $\mu_d^0(\omega_1) = \mu_{d'}^0(\omega_1) = \mu_i^0(\omega_1)$ for any $d, d' \in D_i$ and any $i \in I$.

Each decision-maker (she) takes an action $a \in A$, and her goal is to match the action with the state:

$$u(a, \omega_k) := \mathbb{1}\{a = a_k\} \tag{1}$$

Before taking an action, each decision-maker $d \in D$ pays attention to one expert $j_d \in J$ of her choice: she uses the information provided by the expert to update her belief. The allocation problem is analysed in greater detail in Section 5.

An expert (he) cannot implement an action on his own. Therefore, he designs information to manipulate decision-makers' behaviour. In particular, each expert $j \in J$ chooses a reporting policy $\pi_j : \Omega \to \Delta(S_j)$, that is, each expert commits to the probability $\pi_j(s \mid \omega)$ to send message s given state ω, for any message $s \in S_j$ and any state $\omega \in \Omega$.[12] Each expert j has a unique preferred action $a_j \in A$. For any state $\omega \in \Omega$, his payoff from a decision-maker who takes action $a \in A$ is:

$$u_j(a, \omega) = u_j(a) := \mathbb{1}\{a = a_j\} \tag{2}$$

In other words, each expert has state-independent preferences, and his payoff is 1 if and only if the action chosen by a decision-maker is the expert's preferred action.

The game has the following timing:

1. Each expert j chooses a policy π_j and, at the same time, each decision-maker d chooses which expert j_d to pay attention to.

2. Each decision-maker d observes the policy π_{j_d} of the expert she pays attention to, *and* the policy's realization $s \in S_{j_d}$ (that is, a message) chosen by Nature.

3. Given any posterior belief μ_d, each decision-maker d takes an optimal action. In case of indifference, I assume that decision-maker d chooses the preferred action of expert j_d.

The equilibrium notion is Perfect Bayesian Equilibrium. I consider the case that the preferred action of expert j_d is a_1.[13] By (1), the optimal action of decision-maker d with posterior belief μ_d is given by the following function:

$$\sigma(\mu_d) := \begin{cases} a_1 & \text{if } \mu_d(\omega_1) \geq \frac{1}{2} \\ a_2 & \text{otherwise} \end{cases}$$

Each decision-maker d forms the posterior belief μ_d using Bayesian updating:

$$\mu_d(\omega_1 \mid s) := \frac{\pi_{j_d}(s \mid \omega_1)\mu_d^0(\omega_1)}{\pi_{j_d}(s \mid \omega_1)\mu_d^0(\omega_1) + \pi_{j_d}(s \mid \omega_2)\mu_d^0(\omega_2)}$$

[12] I focus on straightforward policies without loss of generality (Kamenica and Gentzkow, 2011): the message set S_j contains two elements for any expert $j \in J$.

[13] The analysis is very similar when the preferred action of expert j_d is a_2.

Thus, for any decision-maker $d \in D_i$ to take action a_1, upon observing message s, the following condition must hold:

$$\mu_d(\omega_1 \mid s) \geq \frac{1}{2} \iff \pi_{j_d}(s \mid \omega_1)\mu_i^0(\omega_1) \geq \pi_{j_d}(s \mid \omega_2)\mu_i^0(\omega_2)$$

In words, the expert must ensure that state ω_1 is more likely than state ω_2 for a decision-maker of subgroup i after receiving the message s. I label this condition *persuasion constraint.*

Definition 1 (Persuasion constraints)**.** *The persuasion constraint for a decision-maker of subgroup $i \in I$, who devotes attention to expert $j \in J$ and observes message $s \in S_j$, in order for her to take action a_1 is:*

$$\pi_j(s \mid \omega_2) \leq \frac{\mu_i^0(\omega_1)}{\mu_i^0(\omega_2)}\pi_j(s \mid \omega_1) := \phi_i \pi_j(s \mid \omega_1) \tag{3}$$

I denote with $H_j := \{d \in D \mid j_d = j\}$ the set of decision-makers who pay attention to expert j. For any $i \in I$, I define g_{ij} as the fraction of decision-makers in H_j who are of subgroup i. Mathematically,

$$g_{ij} := \begin{cases} 0 & \text{if } H_j = \varnothing \\ \frac{|\{d \in H_j \mid d \in D_i\}|}{|H_j|} & \text{otherwise} \end{cases} \tag{4}$$

These decision-makers have the same posterior belief. Therefore, the payoff of expert j from these decision-makers, upon observing message s, is:

$$v_{ij}(\pi_j, s) := g_{ij} u_j\Big(\sigma\big(\mu_d(\omega_1 \mid s)\big)\Big)$$

The expert j maximizes the sum of expected utilities he derives from his audience H_j:

$$\max_{\pi_j} \sum_{i \in I} \sum_{s \in S_j} \sum_{\omega \in \Omega} \pi_j(s \mid \omega)\mu_j^0(\omega) v_{ij}(\pi_j, s) \tag{5}$$

The expert takes his audience H_j as given. Therefore, (5) is a best-response problem in a simultaneous-move game, where each decision-maker d chooses her allocation of attention j_d, and each expert j chooses his policy π_j.

This problem entails a trade-off for the expert. On the one hand, a message must be "credible" to induce a decision-maker to take the expert's preferred action. Formally, this message must satisfy the corresponding persuasion constraint. The former imposes an upper bound to the probability of observing such a message in the state associated with a different action. On the other hand, provided that a message is persuading, the expert would like to send this message as often as possible.

Lemma 1 (Persuasion constraint)**.** *Consider any expert j and assume without loss of generality that $a_j = a_1$. In any best response π_j, either 1.) there exist a subgroup $i \in I$ of decision-makers and a message $s \in S_j$ such that $\pi_j(s \mid \omega_2) = \phi_i \pi_j(s \mid \omega_1)$ or 2.) the expert is babbling, that is, $\pi_j(s \mid \omega_1) = \pi_j(s \mid \omega_2)$ for any $s \in S_j$.*

By Lemma 1, I can restrict the set of policies that can be best responses: when there is scope for persuasion, then at least one persuasion constraint must hold with equality. In the following section, I use this insight to find candidates for the optimal policy.

4. Media Monism

As a benchmark, I study the problem of one expert - that is given by (5) - abstracting from the attention allocation problem of decision-makers (that I study in Section 5). I assume without loss of generality that the expert's preferred action is a_1, and I omit the index j for simplicity. By (3), a message s persuades a decision-maker of subgroup i to take action a_1 if and only if $\pi(s|\omega_2) \leq \phi_i \pi(s|\omega_1)$. The ratio of prior beliefs ϕ_i for each subgroup $i \in I$ will play a crucial role in the following analysis. From the perspective of the expert, there are two categories of decision-makers: believers and sceptics.

Definition 2 (Believers and sceptics)**.** *Decision-makers of subgroup i are believers of state ω_1 relative to ω_2 if $\phi_i > 1$. Decision-makers of subgroup i are sceptics of state ω_1 relative to ω_2 if $\phi_i < 1$. I denote with $I_2 \subset I$ the set of subgroups of sceptics.*

Without information provision by the expert, believers choose the expert's preferred action, whereas sceptics do not. Therefore, sceptics require persuasion: the expert manipulates their beliefs through his policy π, to induce sceptics to take action a_1. However, the expert must account for the indirect effect that persuasion of sceptics has on the behaviour of believers, as all decision-makers receive the same information. Information provision could induce believers to take the expert's undesired action a_2. Therefore, the expert trades off between persuading sceptics and retaining believers.

In this section, I assume that there are two subgroups of decision-makers, that is, $I = \{1, 2\}$. I assume that subgroup 1 of decision-makers are believers i.e. $\phi_1 > 1$, whereas subgroup 2 are sceptics i.e. $\phi_2 < 1$.[14] Thus, the expert can use a message to persuade all decision-makers or only believers or nobody to take action a_1. In the optimal policy at least one persuasion constraint must hold with equality (Lemma 1). In particular, either only the persuasion constraint for sceptics holds with equality, or both persuasion constraints do so. Hence, I identify two candidates for the optimal policy: hard-news policy and soft-news policy.

Definition 3 (Hard-news policy)**.** *The hard-news policy π_h consists of a persuading message s and a residual message s' such that*

$$\pi_h(s|\omega_1) = 1, \quad \pi_h(s'|\omega_1) = 0,$$

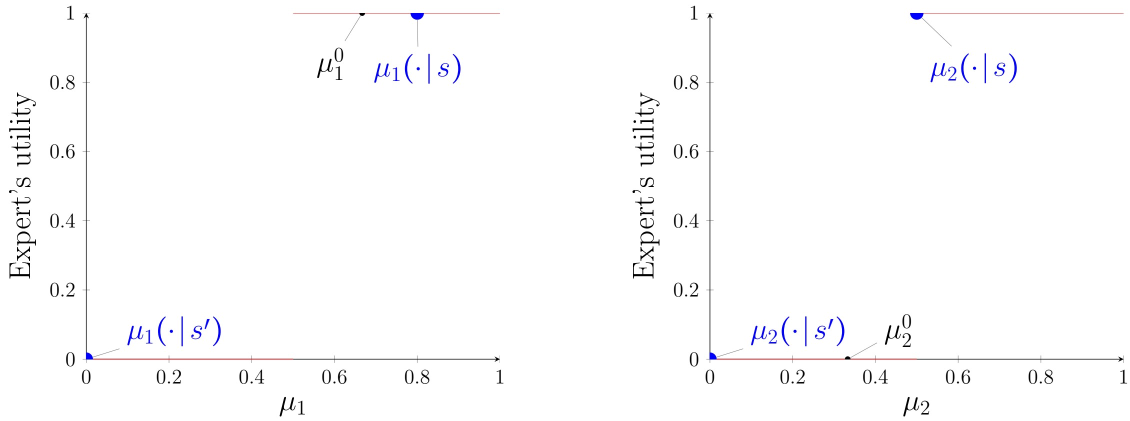
$$\pi_h(s|\omega_2) = \phi_2, \quad \pi_h(s'|\omega_2) = 1 - \phi_2$$

The hard-news policy implies the following posterior beliefs (Figure 2):

$$\begin{aligned} \mu_1(\omega_1|s) = \frac{\phi_1}{\phi_1 + \phi_2} > \mu_2(\omega_1|s) = \frac{1}{2} \\ \mu_1(\omega_1|s') = \mu_2(\omega_1|s') = 0 \end{aligned} \tag{6}$$

The hard-news policy persuades all decision-makers after seeing s and nobody after seeing s'. Thus, decision-makers choose the expert's preferred action in the state ω_1, and sometimes in the state ω_2. The expert provides sufficiently accurate information able to influence sceptics. However, this comes at a high cost to make the persuading message s credible. The credibility of s requires to send the residual message s' often enough when the state is ω_2. The message s' reveals the unfavourable state ω_2, inducing *all* decision-makers to choose the expert's undesired action.

[14] In Section 6.2 I consider the case of arbitrarily many subgroups of decision-makers.



(a) Believers (b) Sceptics

Figure 2: Posterior beliefs (in blue) with the hard-news policy.

Definition 4 (Soft-news policy)**.** *The soft-news policy π_s consists of two messages s, s' such that*

$$\pi_s(s\,|\,\omega_1) = k, \quad \pi_s(s'\,|\,\omega_1) = 1 - k$$
$$\pi_s(s\,|\,\omega_2) = \phi_2 k, \quad \pi_s(s'\,|\,\omega_2) = \phi_1(1 - k)$$

where $k := \frac{\phi_1 - 1}{\phi_1 - \phi_2}$ is strictly increasing in ϕ_1 and ϕ_2.

The soft-news policy implies the following posterior beliefs (Figure 3):

$$\begin{aligned} \mu_1(\omega_1\,|\,s) = \frac{\phi_1}{\phi_1 + \phi_2} &> \mu_2(\omega_1\,|\,s) = \frac{1}{2} \\ \mu_1(\omega_1\,|\,s') = \frac{1}{2} &> \mu_2(\omega_1\,|\,s') = \frac{\phi_2}{\phi_1 + \phi_2} \end{aligned} \tag{7}$$

The soft-news policy persuades all decision-makers after seeing s and believers after seeing s'. Thus, believers choose the expert's preferred action *with probability one*, whereas sceptics choose it with a positive probability (but smaller than one) in either state. The expert alternates information of different accuracy. The message s' is not credible enough to persuade sceptics but ensures that believers keep choosing the expert's preferred action. The expert leverages the believers' credulity without completely giving up on the persuasion of sceptics. The value of k is the maximal extent of persuasion of sceptics, which is possible without affecting believers' behaviour.

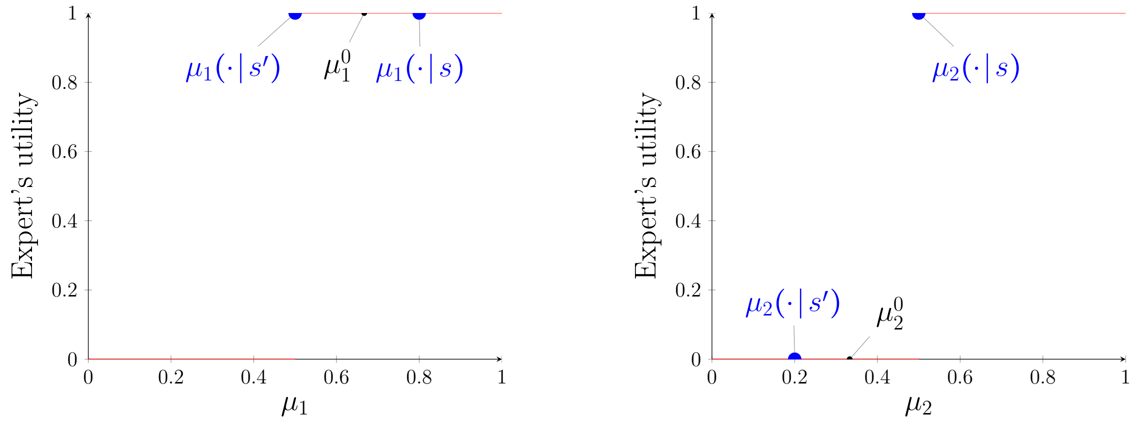
Proposition 1 (Optimal persuasion)**.** *Let $I = \{1, 2\}$, $\phi_1 > 1$ and $\phi_2 < 1$. The unique optimal policy is either the hard-news policy or the soft-news policy. The hard-news policy is optimal if and only if*

$$\mu^0(\omega_1) \geq \frac{\phi_1 g_1 - \phi_2}{1 - \phi_2 + (\phi_1 - 1)g_1} \tag{8}$$

In words, the hard-news policy is optimal if 1.) decision-makers have sufficiently similar beliefs or 2.) the fraction of believers is sufficiently small or 3.) the expert's favourable state is sufficiently likely from his perspective.

By Proposition 1, three parameters influence optimal persuasion:

1. *Decision-makers' polarization, that is, $\phi_1 - \phi_2$*: The larger ϕ_1 is, the higher is the incentive to use the soft-news policy. Indeed, it is easier to leverage believers' credulity



(a) Believers (b) Sceptics

Figure 3: Posterior beliefs (in blue) with the soft-news policy.

using the message s'. In other words, it is easier to prevent believers from taking the expert's undesired action. The smaller ϕ_2 is, the smaller is the incentive to use the hard-news policy. Indeed, it is more costly to persuade sceptics using the message s: the credibility of s requires revealing the unfavourable state with a higher probability. The difference $\phi_1 - \phi_2$ is a proxy for polarization, as the underlying beliefs become more extreme as such difference grows. Therefore, the higher polarization is, the higher the incentive to use the soft-news policy;

2. *Fraction of believers, that is,* g_1: The larger the subgroup of believers (the higher g_1), the higher is the incentive to retain believers (and the lower the incentive to persuade sceptics). This implies a higher incentive to use the soft-news policy;

3. *Expert's prior belief, that is,* $\mu^0(\cdot)$: The lower the expert's belief of his favourable state $\mu^0(\omega_1)$, the higher the cost of revealing the unfavourable state ω_2 to all decision-makers with the hard-news policy. In other words, the expert values his ability to mislead (at least) believers, especially when he is very unconfident about his favourable state being the true state of the world. It follows a higher incentive to use the soft-news policy.

Proposition 1 relates to Kamenica and Gentzkow (2011) in the following way. Kamenica and Gentzkow (2011) assume a common prior belief and, if the decision-maker is a sceptic, the hard-news policy is optimal. Heterogeneous beliefs give rise to a new type of optimal policy - the soft-news policy - pointing out the importance of decision-makers' polarization for optimal persuasion. Moreover, Kamenica and Gentzkow (2011) argue that if a decision-maker chooses the expert's undesired action, then it must be the case that the state is one where such action is optimal. However, this holds only if the expert uses the hard-news policy. With the soft-news policy, sceptics may choose the expert's undesired action even if it is not optimal for them. Finally, persuasion is always optimal when decision-makers have heterogeneous beliefs. The expert uses either the hard-news policy or the soft-news policy. Babbling is never optimal.

Lemma 2 (Blackwell's criterion)**.** *The hard-news policy is more informative than the soft-news policy, according to the order over distributions of posterior beliefs defined by Blackwell (1953).*

A policy π is more informative than π' according to Blackwell (1953) if the distribution of posterior beliefs induced by π constitutes a mean preserving spread of the distribution of posterior beliefs induced by π'. Following this definition, truth-telling is the most informative policy, as the posterior belief is either 0 or 1. Instead, babbling leaves beliefs unchanged, and thus it is the least informative policy. The hard-news policy is more informative than the soft-news policy, for all decision-makers. Indeed, it induces more dispersion in the posterior beliefs through the residual message, which reveals the unfavourable state for the expert.

As Figures 4a and 4b show, the effect of polarization on the informativeness of the monopolist's policy is non-monotonic. Polarization increases informativeness (i.e., the range of posterior beliefs). However, there is a discontinuity point, that is, when the expert shifts from the hard-news policy to the soft-news policy. Therefore, having some degree of heterogeneity in beliefs is beneficial, as it increases the quality of the information provided by the expert. However, if polarization becomes too high, the expert changes policy. Lemma 2 shows that the soft-news policy is less informative than the hard-news policy.

Example. I consider the example from the introduction. There are two states of the world: either a vaccine is safe or it has side effects. The pro-government media wants to persuade citizens that the vaccine is safe. There are two groups of citizens, 1 and 2, and $g_1 = g_2 = \frac{1}{2}$. Group 1 are believers whereas group 2 are sceptics, with prior beliefs $\mu_1^0(\text{Safe}) = 0.7$ and $\mu_2^0(\text{Safe}) = 0.2$ respectively. Therefore, $\phi_1 = \frac{7}{3}$ and $\phi_2 = \frac{1}{4}$. Each citizen decides whether to get vaccinated. Using Definition 3, the hard-news policy can be represented as follows:

ω	Safe	Side Effects	
s	*safe*		*side effects*
$\pi(s\mid\omega)$	1	0.25	0.75

The message *safe* persuades sceptics. To be credible, the pro-government media needs to commit to sending the message *side effects* often enough when the true state is "Side Effects".

Using Definition 4, the soft-news policy can be represented as follows:

ω	Safe		Side Effects	
s	*safe*			*anecdotal safe*
$\pi(s\mid\omega)$	0.64	0.36	0.16	0.84

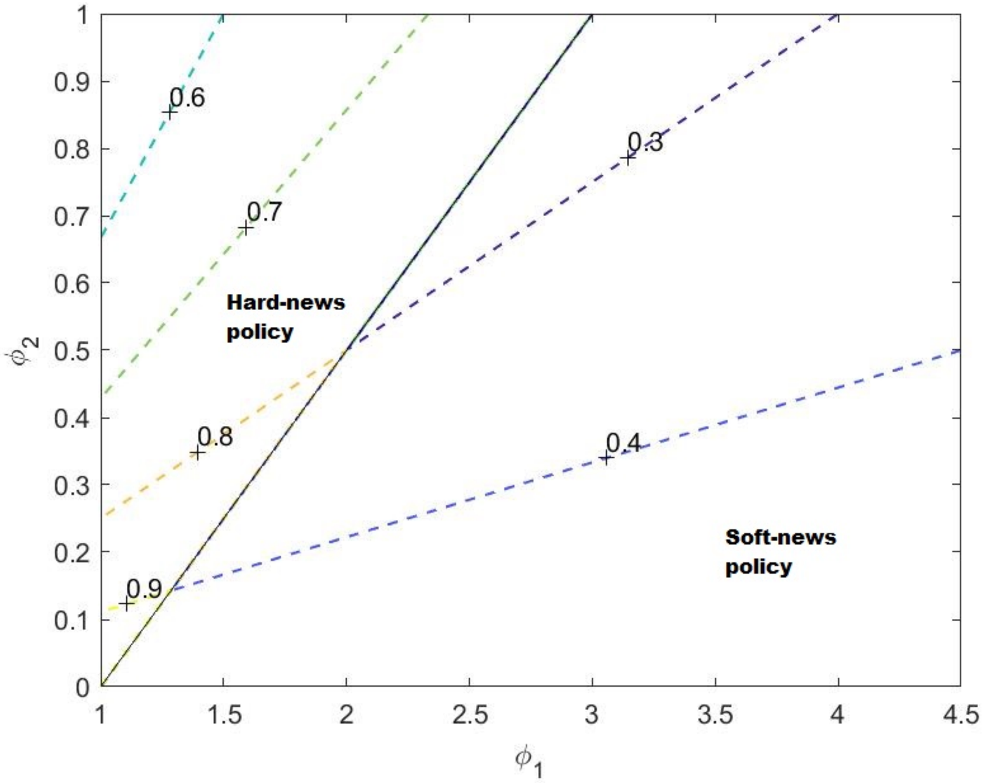
The soft-news policy consists of two messages. The message *safe* (e.g., clinical trials) persuades sceptics but has a low chance to be misleading (that is, to induce decision-makers to choose the wrong action). The message *anecdotal safe* (e.g., vague comparisons of benefits and risks) has a higher chance to be misleading but persuades only believers.

The advantage of the soft-news policy is that believers get vaccinated with probability one. With *anecdotal safe* the pro-government media leverages believers' credulity. Meanwhile, it does not give up entirely from the persuasion of sceptics (message *safe*).

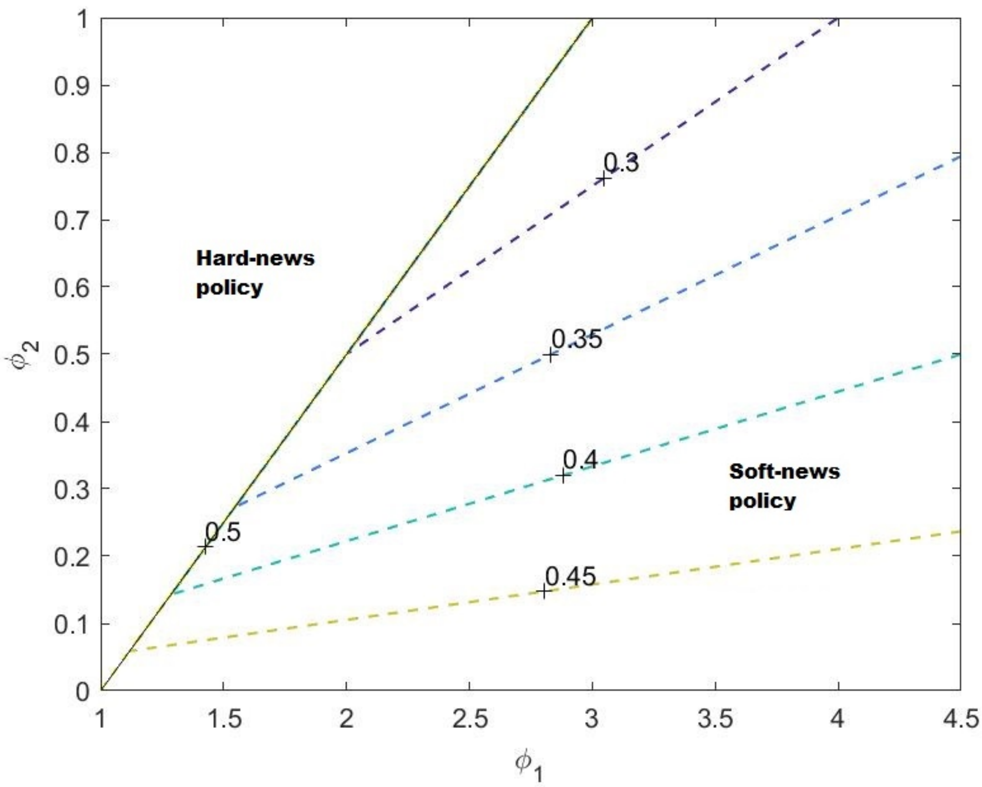
Given citizens' beliefs, whether the soft-news policy is better than the hard-news policy only depends on the pro-government media's belief. In particular, by (8) the pro-government media uses the hard-news policy only if its belief of the vaccine being safe is larger than $\frac{11}{17}$. When sufficiently uncertain about the existence of side effects and if citizens have sufficiently polarized beliefs, the pro-government media uses the soft-news policy.[15]

The natural question to ask is then: What happens if we allow a no-vax media to provide alternative information? The following section provides an answer.

[15] In Section 7, I discuss some possible caveats of this example.



(a) Believers ($i = 1$)



(b) Sceptics ($i = 2$)

Figure 4: Range of posterior beliefs $\mu_i(\omega_1|s) - \mu_i(\omega_1|s')$ by (6)-(7) when $\mu^0(\omega_1) = \frac{1}{2}$ and $g_1 = \frac{1}{2}$.

5. Media Pluralism

In this section, I study how the existence of multiple experts affects the quality of information and the welfare of decision-makers. I restrict attention to a setting with two experts with different preferred actions. Formally, $J = \{\alpha, \beta\}$ with $a_\alpha = a_1$ and $a_\beta = a_2$. Full revelation (i.e., truth-telling by both experts) is the equilibrium when decision-makers have unlimited attention (Gentzkow and Kamenica, 2017a,b; Ravindran and Cui, 2020). In the following, I introduce limited attention and show that full revelation is not an equilibrium. Media pluralism is actually harmful to decision-makers as it deteriorates the quality of information.

Limited attention implies that each decision-maker can only devote attention to one expert. In other words, either $j_d = \alpha$ or $j_d = \beta$ for any decision-maker $d \in D$. The problem for each expert j is identical to the one solved previously. However, the composition of his audience H_j is now endogenous. The distribution of prior beliefs each expert faces is the result of the optimal attention choices of decision-makers. The allocation of attention and the optimal policy are chosen simultaneously by each decision-maker and each expert, respectively.

The objective function of each decision-maker is her subjective probability of choosing the correct action (that is, her expected payoff). Suppose that a decision-maker $d \in D_i$ devotes attention to the expert $j \in J$. Mathematically, this probability can be expressed as follows:

$$\lambda_i(\pi_j) := \sum_{s \in S_j} \sum_{\omega_k \in \Omega} \pi_j(s \mid \omega_k) \mu_i^0(\omega_k) \mathbb{1}\left\{\sigma\big(\mu_d(\omega_1 \mid s)\big) = a_k\right\}$$

Lemma 3 (Decision-maker's payoff)**.** *The policy π_j is truth-telling if and only if $\lambda_i(\pi_j) = 1$. If π_j is babbling, then $\lambda_i(\pi_j) = \mu_i^0(\omega_m)$, where $m = \arg\max_{n \in \{1,2\}} \mu_i^0(\omega_n)$. It holds that $\lambda_i(\pi_j) \in [\mu_i^0(\omega_m), 1]$.*

Intuitively, the subjective probability of taking the correct action is maximal when an expert reveals the state of the world. Without information, a decision-maker of subgroup i chooses the action associated with her most plausible state given prior beliefs: $\mu_i^0(\omega_m)$ is the corresponding subjective probability of taking the correct action. Persuasion cannot decrease such a probability compared to the no information case. In particular, an expert can change a decision-maker's behaviour. However, this requires the expert to reveal some information and makes the decision-maker (weakly) better off. Therefore, $\Delta_{ij} := \lambda_i(\pi_j) - \mu_i^0(\omega_m) \geq 0$ is the subjective information gain from persuasion. I do not assume confirmation bias: babbling is the least desired policy by decision-makers. Even if the assessment of quality is subjective, each decision-maker prefers any informative policy (i.e., any policy with positive information gain) to babbling.

Definition 5 (Target)**.** *For any expert $j \in J$, a target is a subgroup $i \in I$ of decision-makers whose persuasion constraint holds with equality, given the policy of expert j. Let T_j be the set of targets for expert j.*

By Lemma 1, the set of targets is non-empty. A hard-news policy targets sceptics, whereas a soft-news policy targets sceptics and believers. A subgroup being a target means that the expert tailors his policy to persuade marginally decision-makers belonging to such subgroup and thus renders them exactly indifferent between the two actions.

Lemma 4 (Zero information gain for a target)**.** *For each expert $j \in J$ and each $i \in T_j$, it holds that $\Delta_{ij} = 0$.*

Lemma 4 states that when a subgroup is a target of an expert, such decision-makers receive zero information gain when devoting attention to this expert. Intuitively, an expert reveals only the information that is strictly necessary to persuade decision-makers of a targeted subgroup. Being a target is a sufficient condition for zero information gain from persuasion.[16]

Lemma 4 shapes decision-makers' incentives regarding the allocation of attention. The optimal allocation of attention for a decision-maker $d \in D_i$ is given by $j_d(\pi_\alpha, \pi_\beta)$, and $j_d(\cdot) = j$ requires that $j \in \arg\max_{j \in J} \Delta_{ij}$. In other words, each decision-maker devotes attention to the expert that grants her the highest information gain. Crucially, each decision-maker wants to avoid being a target, as in that case $\Delta_{ij} = 0$.

Any equilibrium is thus characterized by a vector $(\pi_\alpha, \pi_\beta, j_1, \ldots, j_{|D|})$. The set of decision-makers who pay attention to the expert j (his audience) is $H_j = \{d \in D \,|\, j_d(\cdot) = j\}$. Each policy must be a best response for the corresponding expert: for a given audience H_j, each expert j uses his optimal policy $\pi_j(H_j)$. At the same time, the allocation of attention must be consistent with decision-makers' incentives. In particular, for any expert $j \in J$ and any decision-maker $d \in H_j$, it must hold that $j_d(\pi_\alpha(H_\alpha), \pi_\beta(H_\beta)) = j$. I define two categories of equilibria:

Definition 6. *An equilibrium is "symmetric" if any two decision-makers of the same subgroup $i \in I$ pay attention to the same expert $j \in J$. Otherwise, the equilibrium is "asymmetric".*

Here, I continue to assume $I = \{1, 2\}$ with $\phi_1 > 1$ and $\phi_2 < 1$. Importantly, decision-makers of subgroup $i = 1$ ($i = 2$) are believers (sceptics) of ω_1 and sceptics (believers) of ω_2. There are three symmetric equilibrium candidates, namely:

1. *Monopoly.* All decision-makers devote attention to the same expert: $H_\alpha = D$ or $H_\beta = D$. The optimal policy follows Proposition 1. The non-active expert is indifferent between any policy;

2. *Echo chambers.* Each expert collects attention only by his believers: $H_\alpha = D_1$ and $H_\beta = D_2$. Therefore, for each expert the optimal policy is babbling;[17]

3. *Opposite-bias learning.* Each expert collects attention only by his sceptics: $H_\alpha = D_2$ and $H_\beta = D_1$. Therefore, for each expert the optimal policy is his hard-news policy.[18]

Proposition 2 (Equilibrium)**.** *Let $J = \{\alpha, \beta\}$ and $I = \{1, 2\}$, where decision-makers of subgroup* 1 *(*2*) are believers from the perspective of expert α (β). Echo chambers with babbling is the unique symmetric equilibrium such that both experts are active.*

In echo chambers, persuasion is unnecessary because each expert faces only his believers. At the same time, given babbling by both experts, decision-makers have no incentive to deviate because each expert provides zero information gain. Therefore, echo chambers are an equilibrium.

An equilibrium with a monopolist requires that the non-active expert provides zero information gain. Otherwise, the targets of the monopolist would find it beneficial to deviate. However, the non-active expert is indifferent between any policy, thus he could provide a

[16]However, it is not a necessary condition: decision-makers whose behaviour is not affected by beliefs updating have zero information gain as well.

[17]Each expert could use any policy that does not change believers' behaviour. I assume that each expert breaks indifference in favour of babbling. This assumption is without loss of generality because a policy is associated with positive information gain only if it changes decision-makers' behaviour. Babbling is the unique optimal policy when the expert pays a cost to generate information.

[18]The soft-news policy is useful to retain believers. Therefore, it cannot be optimal when only sceptics devote attention.

positive information gain. To support this equilibrium, the expert must break indifference in favour of babbling (or equivalent policies).

By Lemma 2, opposite-bias learning would be desirable as each expert would use his hard-news policy. However, opposite-bias learning cannot be an equilibrium because it is not coherent with each decision-maker's incentives. Each sceptic can get a strictly positive information gain by becoming a believer of her like-minded expert. Indeed, when a sceptic deviates and devotes attention to her like-minded expert, she is not a target given the like-minded expert's policy. In other words, the like-minded expert does not tailor information to manipulate his believers' behaviour. That is why sceptics benefits from the deviation.

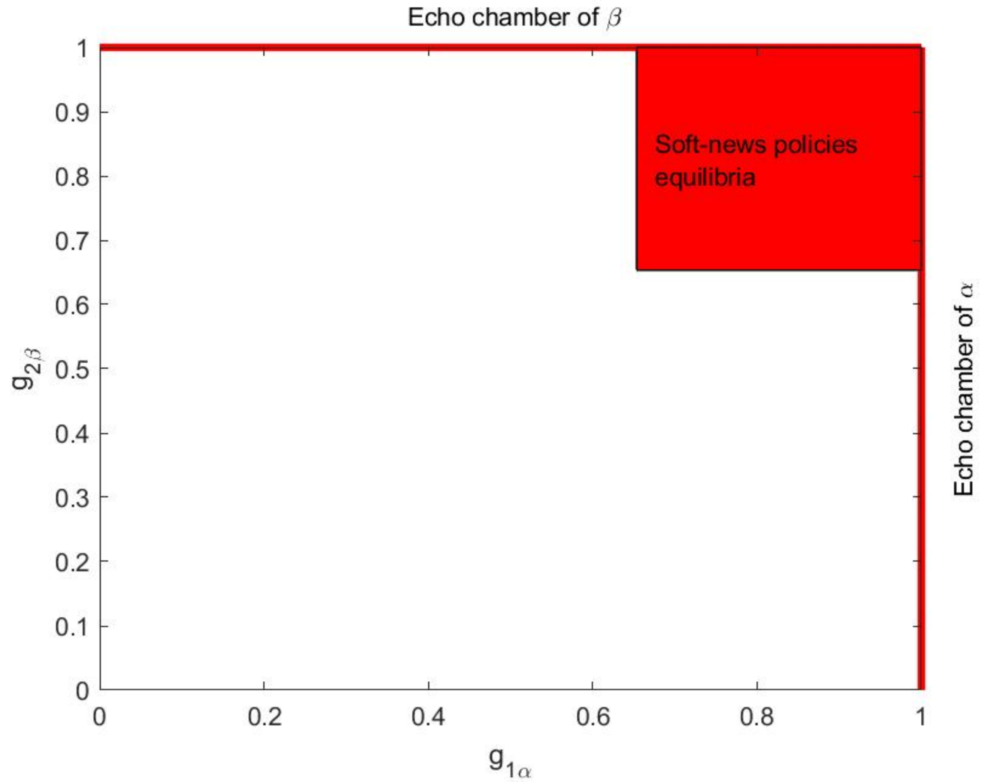


Figure 5: Allocations of attention that can support an equilibrium, when $\mu^0_\alpha(\omega_1) = \mu^0_\beta(\omega_2) = \frac{7}{10}$, $\phi_1 = 2$ and $\phi_2 = \frac{1}{2}$.

The game has also asymmetric equilibria (see Figure 5). A necessary condition is that decision-makers of the same subgroup are indifferent about the allocation of attention. There exist asymmetric equilibria where one expert uses either his hard-news policy or his soft-news policy (i.e., informative expert), whereas the other expert is babbling (i.e., babbling expert). To support these equilibria, the babbling expert must collect attention only from his believers. If this is not the case, babbling is not optimal (Proposition 1). Thus, the informative expert collects attention from all his believers and some of his sceptics. If the informative expert uses his hard-news policy, his sceptics are targets (i.e. zero information gain, from Lemma 4) and thus indifferent about the allocation of attention, whereas his believers are strictly better off by devoting attention to him. If the informative expert uses his soft-news policy, all decision-makers are targets and thus indifferent about the allocation of attention. There also exist asymmetric equilibria where each expert uses his soft-news policy. All decision-makers are targets of each expert. Thus, each decision-maker gets zero information gain independently of the allocation of attention. Any allocation of attention that makes it optimal for each expert to use his soft-news policy constitutes an equilibrium.

Proposition 3 (Harmful Media Pluralism)**.** *For any competitive equilibrium, there exists a monopoly outcome such that information gain and information quality are (weakly) higher for all decision-makers.*

Proposition 3 implies that decision-makers are worse informed when there exist multiple experts. When decision-makers have limited attention, media pluralism leads to an unravelling of information provision. Imagine two experts committing to informative reporting policies. Each decision-maker has an incentive to sort herself into the audience of her like-minded expert to achieve a higher information gain. However, this leads decision-makers to cluster into echo chambers which, in turn, decreases experts' incentives for information provision. In particular, an echo chamber is harmful because the expert faces only his believers, and the best response is babbling. Thus, those decision-makers who cluster in an echo chamber receive information of lower quality than in a monopoly. Indeed, a monopolist uses either his hard-news policy or his soft-news policy (Proposition 1): these policies produce some dispersion in posterior beliefs, whereas babbling leaves beliefs unchanged. Hence, babbling is less informative according to Blackwell (1953)'s order. Moreover, in terms of information gains, those decision-makers who are not targets of the monopolist are strictly worse off in echo chambers. The monopoly outcome also outperforms those asymmetric equilibria where each expert uses his soft-news policy. This result follows Lemma 2 and all decision-makers being targets in these asymmetric equilibria.

Example. An asymmetric equilibrium could fit the COVID-19 vaccination example. The pro-government media collects attention from believers and sceptics and, thus, uses his hard-news policy. The no-vax media exploits his echo chamber and provides information that amounts to babbling. Therefore, decision-makers in the no-vax echo chamber are less informed than in a monopoly. Citizens who are sceptical about vaccinations understand that the pro-government media tailors information to change their behaviour. Therefore, a sceptic has no advantage from devoting attention to the pro-government media and could decide to join the no-vax echo chamber. The number of citizens that the pro-government media can persuade to get vaccinated depends on the equilibrium allocation of attention. Sceptics may cluster into the no-vax echo chamber and get confirmatory news. Their worldview cannot change and, thus, they are not willing to get vaccinated. An implication of this result is that herd immunity is unachievable if the no-vax echo chamber is too large.

6. Extensions

6.1. Platform

The negative effect of media pluralism relates to the endogenous allocation of attention by decision-makers. In this section, I show that media pluralism can enhance information quality when the allocation of attention is exogenous for decision-makers. I assume that there exists a third agent (a platform) that chooses the allocation of attention to maximize the average quality of information that decision-makers receive, which I denote with Ψ and formally define with equation (13) in the Appendix. Let $J = \{\alpha, \beta\}$, $a_\alpha = a_1$, $a_\beta = a_2$ and $I = \{1, 2\}$. I assume that decision-makers of subgroup 1 (2) are believers of state ω_1 (ω_2), that is, $\phi_1 > 1$ and $\phi_2 < 1$. The timing is as follows. The platform chooses g_{ij} for any subgroup $i \in I$ and any expert $j \in J$. Then, each expert j solves (5). By Lemma 2, the most informative policy (among those that are compatible with each expert's incentives) is the hard-news policy. By Proposition 1 (in particular equation (12) in the Appendix), each expert uses his hard-news policy if there are not too many believers in his audience:

$$g_{1\alpha} \leq \hat{g}_\alpha := \frac{\mu^0_\alpha(\omega_1) + \phi_2\mu^0_\alpha(\omega_2)}{\mu^0_\alpha(\omega_1) + \phi_1\mu^0_\alpha(\omega_2)} \tag{9}$$

$$g_{2\beta} \leq \hat{g}_\beta := \frac{\mu^0_\beta(\omega_2) + \frac{1}{\phi_1}\mu^0_\beta(\omega_1)}{\mu^0_\beta(\omega_2) + \frac{1}{\phi_2}\mu^0_\beta(\omega_1)} \tag{10}$$

I label $\hat{g}_\alpha$ and $\hat{g}_\beta$ as the *degrees of tolerance* of experts α and β, respectively. The degree of tolerance is the maximum fraction of believers an expert can have in his audience without finding it optimal to use his soft-news policy.

The previous conditions represent a constraint for the platform that chooses the allocation of attention to induce each expert to use his hard-news policy. There is no equivalent constraint when the allocation of attention is chosen by decision-makers, and this explains echo chambers. Indeed, given that each expert uses his hard-news policy, decision-makers have incentives to become believers. However, this makes the hard-news policy suboptimal for each expert and traps decision-makers into echo chambers.

A hard-news policy is *more informative* for a believer than for a sceptic. Therefore, the platform would like to allocate believers to like-minded experts ($g_{1\alpha}, g_{2\beta} \uparrow$). However, this is effective only if each expert uses his hard-news policy, and this requires the presence of enough sceptics ($g_{1\alpha}, g_{2\beta} \downarrow$). Some believers can be allocated to each expert without affecting his incentives to use his hard-news policy: (9)-(10) must hold. The following proposition summarizes the cases where the platform can find an allocation of attention (where both experts receive attention) that outperforms a monopoly in terms of Ψ.

Proposition 4 (Platform). *A benevolent platform strictly prefers media pluralism to monopoly if any of the following conditions holds:*

1. *Each expert uses his soft-news policy as monopolist;*
2. *Each expert can tolerate more than one believer for each sceptic, that is $\hat{g}_\alpha, \hat{g}_\beta > \frac{1}{2}$.*
3. *The expert α (β) uses his hard-news policy as monopolist but $\hat{g}_\alpha < \frac{1}{2}$ ($\hat{g}_\beta < \frac{1}{2}$), whereas the expert β (α) has degree of tolerance $\hat{g}_\beta > \frac{1}{2}$ ($\hat{g}_\alpha > \frac{1}{2}$) but he uses his soft-news policy as monopolist.*

If condition 1 holds, then by Lemma 2 any allocation of attention that gives to each expert incentives to use his hard-news policy (for instance, opposite-bias learning) is better than any monopoly. If condition 2 holds, the platform can exploit the fact that each expert is willing to use his hard-news policy *even if* there are more believers than sceptics in his audience. Therefore, the platform can increase the mass of believers receiving a hard-news policy, compared to any monopoly. If condition 3 holds, the platform can induce the expert with the highest degree of tolerance to use his hard-news policy by allocating some of his believers to the other expert. This is beneficial because overall there are more believers than in monopoly.

As a final remark, opposite-bias learning is never optimal for the platform. Indeed, each expert uses his hard-news policy, but each decision-maker is a sceptic. The platform can increase Ψ by allocating some but not too many believers to like-minded experts. Alternatively, the platform can increase Ψ inducing a monopoly with hard-news policy. Therefore, even if opposite-bias learning is better than echo chambers (and any other equilibrium in Section 5), an heterogeneous audience is necessary to exploit fully media pluralism.

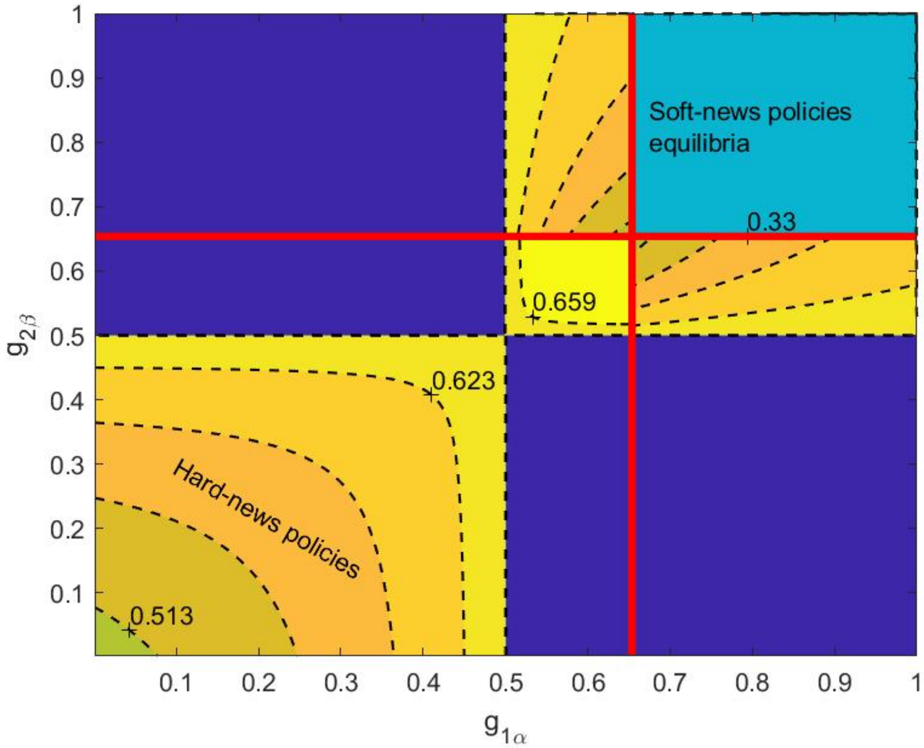


Figure 6: The value of Ψ when $\mu_\alpha^0(\omega_1) = \mu_\beta^0(\omega_2) = \frac{7}{10}$, $\phi_1 = 2$ and $\phi_2 = \frac{1}{2}$. In this example, assuming additionally that the two subgroups have equal size, a monopoly is outperformed by a setting with two experts where each expert uses his hard-news policy. In a monopoly Ψ amounts to $\frac{13}{20}$, whereas the platform can achieve Ψ equal to $\frac{87}{125}$. The platform allocates attention to expose as many believers as possible to hard-news policies, that is, $g_{1\alpha} = g_{2\beta} = 0.653$.

6.2. Many Decision-makers

In this section, I show that my results continue to hold with any arbitrary set I of subgroups of decision-makers. First of all, I consider finitely many subgroups, each one endowed with a different prior belief.

Proposition 5 (Optimal Persuasion)**.** *Let $I = \{1, \dots, R\}$ with $R > 2$, $\phi_1 < 1$ and $\phi_R > 1$. The unique optimal policy is either a hard-news policy or a soft-news policy. A hard-news (soft-news) policy is optimal if a subgroup of sceptics (believers) has the highest value of being persuaded marginally.*

Proposition 5 shows that optimal persuasion is robust to heterogeneity within believers and sceptics. The expert uses a hard-news policy if the subgroup with the highest value as a target is a subgroup of sceptics. Next, I use such insight to extend the analysis to a continuous distribution of decision-makers' beliefs.

Proposition 6 (Optimal persuasion)**.** *Let $F(x)$ be a distribution with support $[0, \infty)$ and density $f(x) > 0 \ \forall x$. Let $\phi_i := \frac{\mu_i^0(\omega_1)}{\mu_i^0(\omega_2)} \sim F$. Then, the expert j with ratio of prior beliefs ϕ_j uses a hard-news policy if a solution $\phi \in [0,1]$ to the following equation exists*

$$h(\phi) = \frac{1}{\phi_j + \phi} \tag{11}$$

and condition (17) *holds. Note that $h(x) := \frac{f(x)}{1-F(x)}$ is the hazard rate function.*

It is possible to evaluate the quality of the information in real-world settings using condition (11). A researcher needs to know the distribution of decision-makers' beliefs and the expert's belief.[19] Then, condition (11) predicts whether the expert uses a hard-news policy or a soft-news policy.[20]

As an example, I assume that F is the exponential distribution. In other words, $F(x;\eta) = 1 - e^{-\eta x}$ where η is a parameter. A special property of this distribution is a constant hazard rate, that is, $h(x) = \eta$. Therefore, equation 11 implies $\phi = \frac{1}{\eta} - \phi_j$ and, by Proposition 6, the expert uses a hard-news policy if $\eta \geq \frac{1}{1+\phi_j}$. Fixing $\phi_j = 1$, Figure 7 depicts two examples of density functions that imply different optimal policies.

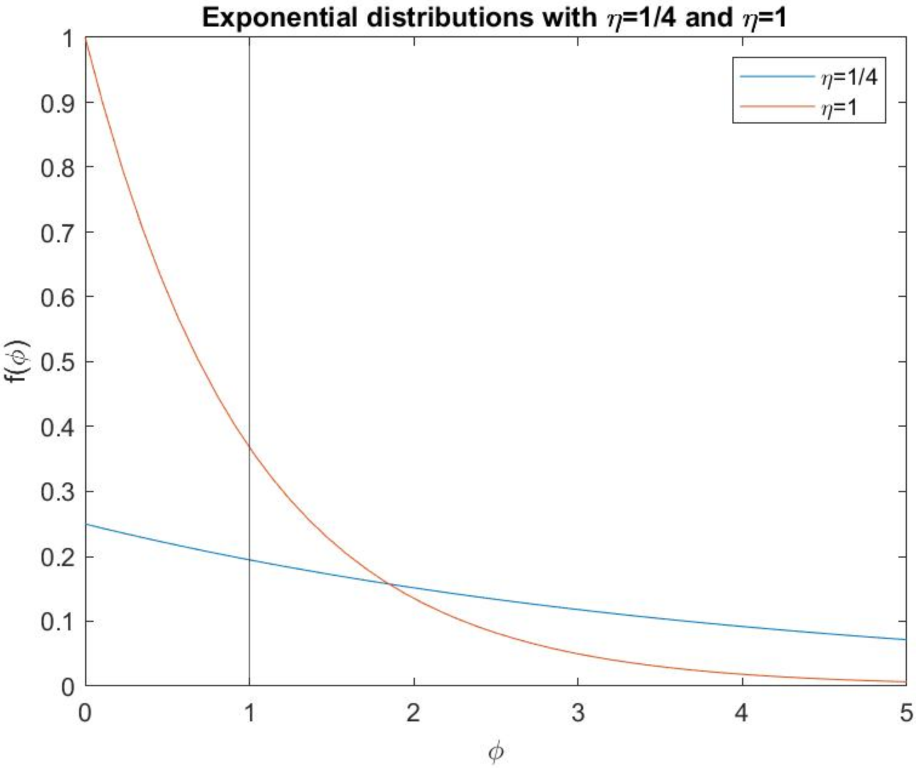


Figure 7: The black line at $\phi = 1$ separates sceptics (at the left) from believers. When $\eta = 1$, the majority of decision-makers are sceptics and, thus, a hard-news policy is optimal. By contrast, a soft-news policy is optimal when $\eta = \frac{1}{4}$, because many decision-makers are believers.

Lemma 5 (Blackwell's criterion)**.** *A hard-news (soft-news) policy is more informative the more extreme are the prior beliefs of its target(s). The ranking of the policies in terms of informativeness is subgroup specific.*

More extreme targets (i.e., targets with beliefs closer to either 0 or 1) induce a more disperse distribution of posterior beliefs: the policy moves closer to truth-telling. Lemma 5 extends Lemma 2: some decision-makers may find a soft-news policy more informative than a hard-news policy if the former targets more extreme sceptics. See condition (18) in the Appendix.

[19]Similar knowledge could derive, for instance, from surveys.

[20]Gitmez and Molavi (2022) show that a hard-news (soft-news) policy is optimal if the distribution of prior beliefs is single-peaked (single-dipped).

Proposition 7 (Equilibrium)**.** *In any symmetric equilibrium, at most one expert is informative.*

The key mechanism behind this result is the following: for any allocation of attention and respective optimal policies, when both experts are informative, there exists at least one target who can deviate and get a positive information gain. This implies that, in any equilibrium, at least one expert is babbling.

The existence of more than two subgroups of decision-makers generates additional symmetric equilibria, which I label *partial echo chambers*. In these equilibria, an ordered subset of believers (those with the most extreme prior beliefs) join the echo chamber of the babbling expert. The other expert gets attention from the remaining decision-makers, including some of his sceptics. Thus, he uses either a hard-news policy or a soft-news policy or, in other words, he is an informative expert. Given babbling, nobody outside the echo chamber wants to join it. At the same time, any believer within the echo chamber would become the most sceptical decision-maker of the informative expert in case of a deviation: given the informative expert's policy, her behaviour would not change. Therefore, this deviation would yield zero information gain, and this supports the equilibrium.

Proposition 8 (Harmful Media Pluralism)**.** *For any competitive equilibrium, there exists a monopoly outcome such that information gain and information quality are (weakly) higher for all decision-makers.*

The negative effect of media pluralism (Proposition 3) extends in a setting with any arbitrary distribution of decision-makers' beliefs. When comparing monopoly with partial echo chambers, a case distinction is necessary. If the informative expert as monopolist uses a hard-news policy, media pluralism is harmful because information gains are (weakly) lower, and those decision-makers who cluster into the echo chamber receive babbling. When the informative expert uses different soft-news policies in monopoly and partial echo chambers, some decision-makers might be better off in partial echo chambers. In this case, media pluralism is harmful to all decision-makers if the targets are strategic substitutes. In particular, decision-makers are worse off - in terms of both information gains and information quality - if both targets are less extreme in partial echo chambers than in monopoly. Intuitively, this sufficient condition should hold because the targeted sceptics are (by construction) less sceptical in partial echo chambers, and thus the expert might be tempted to retain less extreme believers.

6.3. Other Extensions

The results of this paper extend on many other dimensions, which I briefly describe in this section. The online Appendix includes a more detailed discussion. First, when attention is costly rather than limited, my results are robust for any positive cost. By contrast, full revelation is an equilibrium only when attention is costless (or, equivalently, unlimited). Second, I study what happens if decision-makers pay an entropy cost to process information. An entropy cost is a form of confirmation bias: any positive confirmation bias makes echo chambers the unique robust equilibrium. Third, if decision-makers can pay a cost to be second-movers, the results are robust if this cost is high enough (the higher polarization, the lower the threshold). By contrast, full revelation is the equilibrium only when the adjustment is costless. Fourth, if experts can only partial commit to their reporting policies, the results are robust, provided that experts have sufficient commitment power to persuade sceptics. Fifth, when decision-makers are subject to over-inference (i.e., they attribute more importance to the message rather than to their prior beliefs), the unique robust equilibrium is echo chambers. Sixth, the results are robust even if the experts are not exclusively biased, but they also care about gathering attention. Finally, the results are robust when considering a generic number of experts, a continuous state space and more than two actions.

7. Applications

Throughout the paper, I have considered the COVID-19 vaccination as an example to illustrate my results. Such an example could have some caveats. Perhaps it is controversial to assume that the pro-government media has state-independent preferences. There is a trade-off between economic outcomes and the time needed to eradicate COVID-19, which means that herd immunity is a goal. However, the pro-government media is also concerned about safety. My model applies to a vaccine that has been approved for administration. Thus, it is safe overall. However, the pro-government media could avoid disclosing possible side effects. Moreover, many citizens are irrational and cannot be persuaded. Hence, my model applies to the subset of the population that is rational. I show that endogenous echo chambers can explain why many rational citizens are still sceptical about vaccinations and can be a threat to reaching herd immunity.

In this section, I argue that the applicability of my results goes beyond the previous example. My findings require five assumptions: on the one hand, experts are biased and have commitment power; on the other hand, decision-makers have heterogeneous beliefs and limited attention. Finally, I assume that decision-makers and experts make their choices simultaneously. Here, I briefly discuss what is the outcome if I relax any of these assumptions:

1. Under unlimited attention, experts are in direct competition to persuade decision-makers. As a consequence, full revelation is the unique equilibrium as discussed at the beginning of Section 5.

2. When decision-makers share the same prior belief, experts do not face a trade-off between persuading sceptics and retaining believers. As a consequence, each decision-maker has zero information gain independently of the allocation of attention.

3. Trivially, an unbiased expert is truth-telling and collects all attention.

4. When experts have no commitment power, decision-makers anticipate that babbling is optimal for each expert. Thus, decision-makers are indifferent about the allocation of attention.

5. When the allocation of attention is more flexible than the reporting policies of experts, the latter are implicitly attention-seekers. As a consequence, full revelation is the unique equilibrium (Knoepfle, 2020).

Therefore, each assumption is necessary for my results to hold. These assumptions allow me to build a model able to offer insights into the real world. By contrast, the outcome when relaxing any assumption is either full revelation or not conclusive (that is, any outcome is an equilibrium).

My assumptions are realistic in many contexts. First of all, limited attention is a well-established fact. About two-thirds of Americans feel worn out by the excessive amount of news available to them (Pew Research Center, 2020). News consumers tend to interact with a very narrow set of news sources (Cinelli et al., 2020) and have an active role in determining this selective exposure (Bakshy et al., 2015). Second, the media may have commitment power, for instance, because of law or reputation concerns. Kamenica and Gentzkow (2011) discuss this assumption in detail. The media have incentives to build commitment power

(Min, 2021). Fréchette et al. (2019) provide experimental evidence that news consumers react to commitment power as predicted by the Bayesian persuasion theory. Third, there exist empirical evidence of the relative inflexibility of attention habits compared to the reporting policies of media. For instance, Eisensee and Strömberg (2007) show that politicians respond strategically to attention habits in the context of news coverage about natural disasters. Using data from Wikipedia, Ciampaglia et al. (2015) show that the demand for information (that is, the allocation of attention) precedes its supply. Attention habits make the commitment assumption weaker: an expert commits not to misreport but does not commit ex-ante to the quality of information.[21] In other words, news consumers may know how the media are biased but not the quality of news before devoting attention (e.g., news consumers see an article's headline). Media have incentives to exploit news consumers by providing tailored information and cannot commit otherwise. Fourth, heterogeneous beliefs are also very likely to exist in all situations where the objective probability for a claim to be true is ambiguous. For instance, politicians and bureaucrats may share the same goal but disagree about the best way to achieve it (Hirsch, 2016). Finally, whenever the true state of the world is disputed, there are likely competing interpretations of the current state of events. If this is true, the last requirement to apply my insights, namely competition between biased experts, is fulfilled. McCarthy and Dolfsma (2014) survey evidence that all media are biased, intentionally or unintentionally.

My model applies to the design of information about political issues. A politician (or a firm) wants to persuade voters to support a particular point of view. When the audience is heterogeneous, there is a trade-off between persuading sceptical voters and retaining loyalists. As a result, it is optimal to provide some information. However, when attention is limited, voters have incentives to sort themselves into more homogenous audiences - in the extreme case, into an echo chamber - and thus, endogenously, politicians provide less information. There is evidence of political information tailored to particular audiences.[22] However, this could be driven by preferences (i.e., confirmation bias) rather than limited attention. The two channels are complementary: my model shows that even rational voters could cluster into echo chambers. I claim that the confirmation bias channel becomes less credible (and, thus, limited attention is a more credible channel) when the information is about seemingly non-partisan issues such as, for instance, climate change.[23] My model explains the existence of echo chambers in all those contexts such that decision-makers should not have a confirmation bias. For instance, Cookson et al. (2021) provide evidence of the existence of echo chambers even among (professional) investors in the financial markets. In general, in settings where a decision-maker's utility depends primarily on her knowledge of the state of the world and a decision-maker cannot compensate her ignorance with the feeling of being right, limited attention is the most credible explanation for echo chambers.

Advertising represents an interesting application for my model. Chen and Riordan (2008) show that competition between differentiated products can lead to higher prices. My model adds another negative effect: the endogenous sorting of consumers into the audience of their

[21] Experts' reporting policies are not observable before decision-makers allocate attention. This idea resembles ambiguous persuasion (Beauchêne et al., 2019).

[22] A recent example is Trump's claim that the US Presidential election was fraudulent. My model offers an explanation for why Republicans believe Biden won because of a "rigged" election, even though Trump has failed to provide any evidence about that (Rutenberg et al., 2020).

[23] Scientists claim that climate change is real and warn that immediate intervention is necessary to avoid a sharp increase in mass disasters, whereas corporations (especially coal and oil producers) try to dispute such warnings. Limited attention and endogenous echo chambers can explain the existence of climate change deniers. Similarly, believers of a long list of debunked conspiracy theories can survive within echo chambers. The common root is widespread scepticism about Science (Achenbach, 2015).

respective preferred firm changes the informational content of advertising. Imagine that consumers need to search for the best product given their tastes but have limited time (or attention) to process advertising of two competing firms. Each consumer finds it optimal to devote attention to the advertising by the firm she believes supplies the best product. However, sorting reduces firms' incentives to provide informative advertising. As a result, both firms supply uninformative advertising (echo chambers equilibrium), or one firm invests in informative advertising, whereas the other enjoys its market niche (partial echo chambers equilibrium).

8. Conclusion

I show two main results about the quality of the information. First, it depends on agents' beliefs. When worldviews are sufficiently polarized, a monopolist provides lower quality information. Second, media pluralism backfires when attention is limited: increasing the diversity of information sources reduces information quality even further. Echo chambers arise endogenously, and as a consequence, the incentives for the media to provide valuable information vanish.

My findings suggest that increasing media pluralism is likely to have a non-monotonic effect on information quality. In particular, the effect is positive when there are few media (or their ownership is concentrated) but negative as soon as there is information overload. Limited attention introduces an additional choice for news consumers: the subset of information to process. Policymakers should account for news consumers' incentives to cluster into echo chambers. I show that supporting media pluralism is a good idea only if news consumers are sufficiently attentive to process information from diverse sources.

The standard explanation for the existence of echo chambers is demand-driven: news consumers are biased, selfish or have some cognitive limitation. I show that there exists a complementary and supply-driven explanation. Because of information overload, even rational and unbiased news consumers end up devoting their limited attention to like-minded media. The latter, then, find it optimal to confirm news consumers' beliefs. Therefore, I provide a rational foundation for confirmation bias. Goette et al. (2020) provide experimental evidence that limited attention reinforces confirmation bias.

Whether the formation of echo chambers is mainly demand-driven or supply-driven is a fundamental question to address with future research. Understanding which is the main channel is necessary to design policy remedies. When the formation of echo chambers is supply-driven, as I suggest in this paper, one solution is to enhance attention, but it is unclear how to do this. An alternative is to manipulate the allocation of attention to improve information quality. In Section 6.1, I have shown how a platform that wants to maximize the informativeness of news should allocate attention. Such a platform can design each expert's audience to give him incentives to use his hard-news policy. In this way, media pluralism can enhance the average quality of information that news consumers receive. Platforms such as news aggregators may have the ability to shape how their users allocate attention. However, there is no guarantee that such platforms behave as a social planner would do. Future research should assess the impact of platforms when their objective is to maximize profits.

References

Acemoglu, D., A. Ozdaglar, and J. Siderius (2021): "Misinformation: Strategic Sharing, Homophily, and Endogenous Echo Chambers," .

Achenbach, J. (2015): "Why Do Many Reasonable People Doubt Science?" *National Geographic Magazine*.

Allcott, H., L. Boxell, J. Conway, M. Gentzkow, M. Thaler, and D. Yang (2020): "Polarization and public health: Partisan differences in social distancing during the coronavirus pandemic," *Journal of Public Economics*, 191, 104254.

Allcott, H. and M. Gentzkow (2017): "Social media and fake news in the 2016 election," *Journal of Economic Perspectives*, 31, 211–36.

Alonso, R. and O. Camara (2016): "Bayesian persuasion with heterogeneous priors," *Journal of Economic Theory*, 165.

Anderson, S. P. and A. de Palma (2012): "Competition for attention in the Information (overload) Age," *The RAND Journal of Economics*, 43, 1–25.

Anderson, S. P. and M. Peitz (2020): "Ad clutter, time use, and media diversity," .

Andreottola, G. and A.-I. de Moragas Sánchez (2020): "Scandals, Media Competition and Political Accountability," .

Aumann, R. J. and M. Maschler (1995): *Repeated games with incomplete information*, MIT press.

Bakshy, E., S. Messing, and L. A. Adamic (2015): "Exposure to ideologically diverse news and opinion on Facebook," *Science*, 348, 1130–1132.

Beauchêne, D., J. Li, and M. Li (2019): "Ambiguous persuasion," *Journal of Economic Theory*, 179, 312–365.

Bergemann, D. and S. Morris (2019): "Information design: A unified perspective," *Journal of Economic Literature*, 57, 44–95.

Bisceglia, M. (2021): "The Unbundling of Journalism," *Available at SSRN 3885251.*

Blackwell, D. (1953): "Equivalent comparisons of experiments," *The Annals of Mathematical Statistics*, 265–272.

Bloedel, A. W. and I. Segal (2020): "Persuading a Rationally Inattentive Agent," .

Che, Y.-K. and K. Mierendorff (2019): "Optimal dynamic allocation of attention," *American Economic Review*, 109.

Chen, Y. and M. H. Riordan (2008): "Price-Increasing Competition," *The RAND Journal of Economics*, 39, 1042–1058.

Ciampaglia, G. L., A. Flammini, and F. Menczer (2015): "The production of information in the attention economy," *Scientific Reports*, 5, 1–6.

Cinelli, M., E. Brugnoli, A. L. Schmidt, F. Zollo, W. Quattrociocchi, and A. Scala (2020): "Selective exposure shapes the Facebook news diet," *PloS one*, 15, e0229129.

Cookson, J. A., J. Engelberg, and W. Mullins (2021): "Echo chambers," *Available at SSRN 3603107*.

Cossard, A., G. D. F. Morales, K. Kalimeri, Y. Mejova, D. Paolotti, and M. Starnini (2020): "Falling into the echo chamber: the Italian vaccination debate on Twitter," in *Proceedings of the International AAAI conference on web and social media*, vol. 14, 130–140.

Dasgupta, P. and E. Maskin (1986): "The Existence of Equilibrium in Discontinuous Economic Games, II: Applications," *The Review of Economic Studies*, 53, 27–41.

De Clippel, G., K. Eliaz, and K. Rozen (2014): "Competing for consumer inattention," *Journal of Political Economy*, 122, 1203–1234.

de Clippel, G. and X. Zhang (2020): "Non-bayesian persuasion," .

Del Vicario, M., A. Bessi, F. Zollo, F. Petroni, A. Scala, G. Caldarelli, H. E. Stanley, and W. Quattrociocchi (2016): "The spreading of misinformation online," *Proceedings of the National Academy of Sciences*, 113, 554–559.

Di Marco, N., M. Cinelli, and W. Quattrociocchi (2021): "Infodemics on Youtube: Reliability of Content and Echo Chambers on COVID-19," *arXiv preprint arXiv:2106.08684*.

Diamond, D., D. Keating, and C. Moody (2021): "Vaccination rates fall off, imperiling Biden's July Fourth goal," *The Washington Post*.

Eisensee, T. and D. Strömberg (2007): "News Droughts, News Floods, and U. S. Disaster Relief," *The Quarterly Journal of Economics*, 122.

Emons, W. and C. Fluet (2019): "Strategic communication with reporting costs," *Theory and Decision*, 87, 341–363.

Falkinger, J. (2008): "Limited Attention as a Scarce Resource in Information-Rich Economies," *The Economic Journal*, 118, 1596–1620.

Flynn, D., B. Nyhan, and J. Reifler (2017): "The nature and origins of misperceptions: Understanding false and unsupported beliefs about politics," *Political Psychology*, 38, 127–150.

Fréchette, G. R., A. Lizzeri, and J. Perego (2019): "Rules and commitment in communication: An experimental analysis," .

Fridman, A., R. Gershon, and A. Gneezy (2021): "COVID-19 and vaccine hesitancy: A longitudinal study," *PloS One*, 16.

Gabaix, X. (2019): "Behavioral inattention," in *Handbook of Behavioral Economics: Applications and Foundations 1*, Elsevier, vol. 2, 261–343.

Gentzkow, M. and E. Kamenica (2014): "Costly persuasion," *American Economic Review*, 104, 457–62.

——— (2017a): "Bayesian persuasion with multiple senders and rich signal spaces," *Games and Economic Behavior*, 104.

——— (2017b): "Competition in Persuasion," *The Review of Economic Studies*, 84.

Gentzkow, M., J. M. Shapiro, and D. F. Stone (2015): "Media bias in the marketplace: Theory," in *Handbook of Media Economics*, Elsevier, vol. 1, 623–645.

Ghoneim, A., S. Salihudin, I. Thange, A. Wen, J. Oledan, and J. N. Shapiro (2020): "Profiting from panic: the bizarre bogus cures and scams of the coronavirus era," *Bulletin of the Atomic Scientists*.

Giovanniello, M. A. (2021): "Echo chambers: voter-to-voter communication and political competition," *arXiv preprint arXiv:2104.04703*.

Gitmez, A. A. and P. Molavi (2022): "Polarization and Media Bias," *arXiv preprint arXiv:2203.12698*.

Goette, L., H.-J. Han, and B. T. K. Leung (2020): "Information overload and confirmation bias," *Cambridge Working Papers in Economics*.

Gollwitzer, A., C. Martel, W. J. Brady, P. Pärnamets, I. G. Freedman, E. D. Knowles, and J. J. Van Bavel (2020): "Partisan differences in physical distancing are linked to health outcomes during the COVID-19 pandemic," *Nature Human Behaviour*, 4, 1186–1197.

Guo, Y. and E. Shmaya (2019): "The interval structure of optimal disclosure," *Econometrica*, 87, 653–675.

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

Hefti, A. (2018): "Limited attention, competition and welfare," *Journal of Economic Theory*, 178, 318–359.

Hefti, A. and J. Lareida (2021): "Competitive attention, superstars and the long tail," *University of Zurich, Department of Economics, Working Paper*.

Hefti, A. and S. Liu (2020): "Targeted information and limited attention," *The RAND Journal of Economics*, 51, 402–420.

Heidhues, P., B. Koszegi, and T. Murooka (2021): "Procrastination Markets," .

Hirsch, A. V. (2016): "Experimentation and persuasion in political organizations," *American Political Science Review*, 110, 68–84.

Hu, L., A. Li, and X. Tan (2021): "A Rational Inattention Theory of Echo Chamber," *arXiv preprint arXiv:2104.10657*.

Jann, O. and C. Schottmuller (2019): *Why echo chambers are useful*, Verein für Socialpolitik.

Janssen, M. C. and S. Roy (2014): "Competition, Disclosure and Signalling," *The Economic Journal*, 125, 86–114.

Jiang, J., X. Ren, and E. Ferrara (2021): "Social media polarization and echo chambers: A case study of COVID-19," *arXiv preprint arXiv:2103.10979*.

Kamenica, E. (2019): "Bayesian persuasion and information design," *Annual Review of Economics*, 11, 249–272.

Kamenica, E. and M. Gentzkow (2011): "Bayesian Persuasion," *American Economic Review*, 101.

Kartik, N., F. X. Lee, and W. Suen (2017): "Investment in concealable information by biased experts," *The RAND Journal of Economics*, 48, 24–43.

Keen, A. (2015): *The Internet Is Not the Answer*, Grove Press.

Knoepfle, J. (2020): "Dynamic Competition for Attention," .

Kolotilin, A., T. Mylovanov, A. Zapechelnyuk, and M. Li (2017): "Persuasion of a privately informed receiver," *Econometrica*, 85, 1949–1964.

Lee, S. K., N. J. Lindsey, and K. S. Kim (2017): "The effects of news consumption via social media and news information overload on perceptions of journalistic norms and practices," *Computers in Human Behavior*, 75, 254–263.

Leung, B. T. K. (2020): "Limited cognitive ability and selective information processing," *Games and Economic Behavior*, 120, 345 – 369.

Levy, G. and R. Razin (2019): "Echo chambers and their effects on economic and political outcomes," *Annual Review of Economics*, 11, 303–328.

Lipnowski, E., L. Mathevet, and D. Wei (2020): "Attention management," *American Economic Review: Insights*, 2, 17–32.

Mackowiak, B., F. Matejka, and M. Wiederholt (2021): "Rational inattention: A review," .

Martinez, G. and N. H. Tenev (2020): "Optimal Echo Chambers," .

Masatlioglu, Y., D. Nakajima, and E. Y. Ozbay (2012): "Revealed Attention," *American Economic Review*, 102, 2183–2205.

Matějka, F. (2016): "Rationally inattentive seller: Sales and discrete pricing," *The Review of Economic Studies*, 83, 1125–1155.

Matějka, F. and G. Tabellini (2021): "Electoral competition with rationally inattentive voters," *Journal of the European Economic Association*, 19, 1899–1935.

Matysková, L. and A. Montes (2021): "Bayesian Persuasion With Costly Information Acquisition," .

McCarthy, K. J. and W. Dolfsma (2014): "Neutral media? Evidence of media bias and its economic impact," *Review of Social Economy*, 72, 42–54.

Min, D. (2021): "Bayesian persuasion under partial commitment," *Economic Theory*, 1–22.

Molina, M. D., S. S. Sundar, T. Le, and D. Lee (2021): ""Fake news" is not simply false information: a concept explication and taxonomy of online content," *American Behavioral Scientist*, 65, 180–212.

Mullainathan, S. and A. Shleifer (2005): "The market for news," *American economic review*, 95, 1031–1053.

Nickerson, R. S. (1998): "Confirmation bias: A ubiquitous phenomenon in many guises," *Review of General Psychology*, 2, 175–220.

Nimark, K. P. and S. Sundaresan (2019): "Inattention and belief polarization," *Journal of Economic Theory*, 180, 203–228.

Paul, E., A. Steptoe, and D. Fancourt (2021): "Attitudes towards vaccines and intention to vaccinate against COVID-19: Implications for public health communications," *The Lancet Regional Health-Europe*, 1, 100012.

Peng, L. and W. Xiong (2006): "Investor attention, overconfidence and category learning," *Journal of Financial Economics*, 80, 563–602.

Pennycook, G. and D. G. Rand (2021): "The psychology of fake news," *Trends in Cognitive Sciences*.

Perego, J. and S. Yuksel (2021): "Media competition and social disagreement," .

Persson, P. (2018): "Attention manipulation and information overload," *Behavioural Public Policy*, 2.

Pew Research Center (2020): "News Fatigue More Common Among Whites than Blacks and Hispanics," .

Pfäuti, O. (2021): "Inflation–who cares? Monetary Policy in Times of Low Attention," *arXiv preprint arXiv:2105.05297*.

Qiu, X., D. F. Oliveira, A. S. Shirazi, A. Flammini, and F. Menczer (2017): "Limited individual attention and online virality of low-quality information," *Nature Human Behaviour*, 1, 1–7.

Ravindran, D. and Z. Cui (2020): "Competing Persuaders in Zero-Sum Games," *arXiv preprint arXiv:2008.08517*.

Rutenberg, J., N. Corasaniti, and A. Feuer (2020): "Trump's Fraud Claims Died in Court, but the Myth of Stolen Elections Lives On," *The New York Times*.

Simon, H. A. (1971): in *Computers, Communication and the Public Interest*, ed. by M. Greenberger, chap. Designing organizations for an information-rich world.

Spinnewijn, J. (2013): "Insurance and Perceptions: How to Screen Optimists and Pessimists," *The Economic Journal*, 123, 606–633.

Swire-Thompson, B. and D. Lazer (2019): "Public health and online misinformation: Challenges and recommendations." *Annual Review of Public Health*, 41, 433–451.

Törnberg, P. (2018): "Echo chambers and viral misinformation: Modeling fake news as complex contagion," *PLoS one*, 13, e0203958.

Wei, D. (2021): "Persuasion under costly learning," *Journal of Mathematical Economics*, 94, 102451.

Williams, H. T., J. R. McMurray, T. Kurz, and F. H. Lambert (2015): "Network analysis reveals open forums and echo chambers in social media discussions of climate change," *Global Environmental Change*, 32, 126–138.

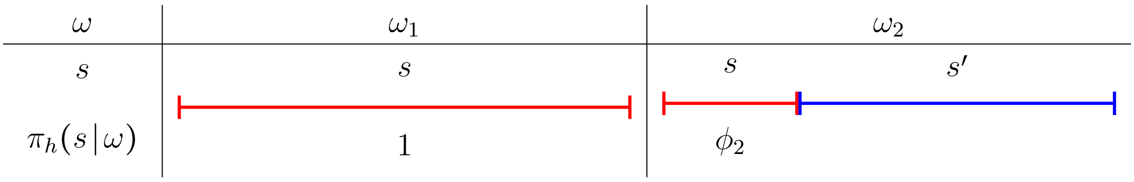
A. Appendix (Proofs)

Proof of Lemma 1

Proof. I assume there exists $i \in I$ such that $g_{ij} > 0$ and $\phi_i < 1$. Otherwise, persuasion is not necessary and babbling is the only optimal policy. I assume by contradiction that $\nexists\, s \in S_j$ such that $\pi_j(s|\omega_2) = \phi_i \pi_j(s|\omega_1)$ for some $i \in I$. Let $\{\phi_i\}$ be the ordered (in ascending order) set of constraints for each subgroup $i \in I$ such that $g_{ij} > 0$. If the n-th constraint holds for a message $s \in S_j$, then the m-th constraint holds too, for any $m > n$. Therefore, if n-th constraint holds there is more persuasion than if only the m-th constraint were holding, ceteris paribus. Thus, if the n-th constraint is slack, it is beneficial for the expert to increase the probability of the corresponding message, at the expense of the probability of a message which satisfy only the m-th constraint. There always exists a deviation for the expert unless at least one constraint holds with equality. $\square$

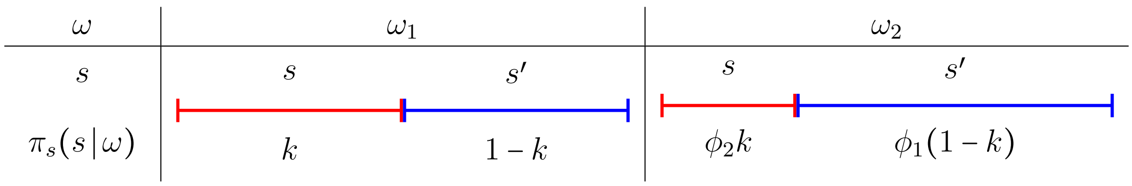
Proof of Proposition 1

Proof. The payoff for *Babbling* is $V_u := g_1$, whereas the payoff for the *Truth-telling policy* is $V_t := \mu^0(\omega_1)$. The *Hard-news policy* is as follows:



$$\implies V_h := \mu^0(\omega_1) + \mu^0(\omega_2)\phi_2$$

The *Soft-news policy* is as follows:



$$\implies V_s := kV_h + (1-k)\left[\mu^0(\omega_1) + \mu^0(\omega_2)\phi_1\right] g_1$$

where

$$1 - \phi_2 k = \phi_1(1-k) \iff k = \frac{\phi_1 - 1}{\phi_1 - \phi_2}$$

Any alternative policy with $\pi(s|\omega_1) < k$ is suboptimal, because the soft-news policy increases the probability of persuading sceptics without affecting the behaviour of believers.

Note that $V_h \geq V_t$. Hence, the expert does not use the truth-telling policy. Moreover, $V_s > V_u$ for any $g_1 \in (0,1)$. The hard-news policy is optimal if:

$$V_h \geq V_s \iff \mu^0(\omega_1) + \mu^0(\omega_2)\phi_2 \geq \left(\mu^0(\omega_1) + \mu^0(\omega_2)\phi_1\right) g_1$$

$$\iff \mu^0(\omega_1)(1-g_1) \geq \mu^0(\omega_2)\left(\phi_1 g_1 - \phi_2\right) \tag{12}$$

Note that the RHS of (12) is increasing in ϕ_1 and decreasing in ϕ_2. The difference of these two values is a proxy for decision-makers' polarization in terms of prior beliefs. The RHS (LHS) of (12) is increasing (decreasing) in g_1, the share of believers among decision-makers. Finally, the RHS (LHS) of (12) is decreasing (increasing) in $\mu^0(\omega_1)$, the expert's belief of his favourable state. $\square$

Proof of Lemma 2

Proof. First of all, the distributions of posterior beliefs induced by these two policies have the same mean, which coincides with $\mu_i^0(\omega_1)$ for any $i \in I$, following Bayesian plausibility. It follows by (6)-(7) that π_h is characterized by more dispersion then π_s. Indeed, with the hard-news policy:

$$\mu_1(\omega_1 \,|\, s) - \mu_1(\omega_1 \,|\, s') = \frac{\phi_1}{\phi_1 + \phi_2}$$

$$\mu_2(\omega_1 \,|\, s) - \mu_2(\omega_1 \,|\, s') = \frac{1}{2}$$

whereas with the soft-news policy:

$$\mu_1(\omega_1 \,|\, s) - \mu_1(\omega_1 \,|\, s') = \frac{\phi_1}{\phi_1 + \phi_2} - \frac{1}{2}$$

$$\mu_2(\omega_1 \,|\, s) - \mu_2(\omega_1 \,|\, s') = \frac{1}{2} - \frac{\phi_2}{\phi_1 + \phi_2}$$

Therefore, π_h is more informative than π_s following Blackwell (1953). □

Proof of Lemma 3

Proof. Assume that π_j is truth-telling. Hence, $\pi_j(s|\omega_1) = \pi_j(s'|\omega_2) = 1$ and $\pi_j(s|\omega_2) = \pi_j(s'\,|\,\omega_1) = 0$. This implies that $\lambda_i(\pi_j) = 1$. Assume that π_j is not truth-telling, and without loss of generality $\pi_j(s\,|\,\omega_2) > 0$. Note that either $\sigma\big(\mu_i(\omega_1\,|\,s)\big) = a_1$ or $\sigma\big(\mu_i(\omega_1\,|\,s)\big) = a_2$. It follows that $\lambda_i(\pi_j) < 1$.

If π_j is babbling then, for any $s \in S_j$, $\sigma\big(\mu_i(\omega_1\,|\,s)\big) = a_m$. It follows that $\lambda_i(\pi_j) = \mu_i^0(\omega_m)$. Assume that there exists $s \in S_j$ and $\omega_k \neq \omega_m$ such that $\pi_j(s\,|\,\omega_k) \neq \pi_j(s\,|\,\omega_m)$. By (3), $\sigma\big(\mu_i(\omega_1\,|\,s)\big) = a_k$ if $\pi_j(s\,|\,\omega_k) \geq \frac{\mu_i^0(\omega_m)}{\mu_i^0(\omega_k)}\pi_j(s\,|\,\omega_m)$, and this implies that $\lambda_i(\pi_j) \geq \mu_i^0(\omega_m)$. □

Proof of Lemma 4

Proof. Assume without loss of generality $a_j = a_1$. If π_j is a hard-news policy then $T_j = \{i\}$ and $\phi_i < 1$. This implies $\lambda_i(\pi_j) = \mu_i^0(\omega_1) + \mu_i^0(\omega_2)\,[1 - \phi_i] = \mu_i^0(\omega_2)$. If π_j is a soft-news policy then $T_j = \{i, i'\}$ and without loss of generality $\phi_{i'} > 1 > \phi_i$. Therefore, $\lambda_i(\pi_j) = \mu_i^0(\omega_1)k + \mu_i^0(\omega_2)\,[1 - \phi_i k] = \mu_i^0(\omega_2)$ and $\lambda_{i'}(\pi_j) = \mu_{i'}^0(\omega_1)$. □

Proof of Proposition 2

Proof. Echo chambers: Given $H_\alpha = D_1$ and $H_\beta = D_2$, babbling is optimal for each expert. Therefore, by Lemma 3, $\Delta_{ij} = 0$ for any $i \in I$ and $j \in J$. Therefore, $j_1 = \alpha$ and $j_2 = \beta$ is optimal for decision-makers.

Monopoly: I assume without loss of generality $H_\alpha = D$ and $H_\beta = \varnothing$. The subgroup $i = 2$ must be a target. By Lemma 4, sceptics get zero information gain, that is $\Delta_{2\alpha} = 0$. Therefore, $j_2 = \alpha$ is optimal only if $\Delta_{2\beta} = 0$. Note that β is indifferent between any policy. This equilibrium breaks down if π_β is such that $\Delta_{2\beta} > 0$.

Opposite-bias learning: Given $H_\alpha = D_2$ and $H_\beta = D_1$, the hard-news policy is optimal for each expert. By Lemma 4, $\Delta_{1\beta} = \Delta_{2\alpha} = 0$. However, $\Delta_{1\alpha}, \Delta_{2\beta} > 0$. Therefore, $j_1 = \beta$ and $j_2 = \alpha$ cannot be optimal for decision-makers. □

Proof of Proposition 3

Proof. An asymmetric equilibrium where for each subgroup $i \in I$ two decision-makers of the same subgroup devote attention to different experts requires each expert to use his soft-news policy. Indeed, in this case all decision-makers are targets and get zero information gain independently of the allocation of attention: $\Delta_{i\alpha} = \Delta_{i\beta} = 0$ for any $i \in I$. These equilibria are equivalent to echo chambers in terms of information gains. Decision-makers are (weakly) better off in a monopoly: if the expert uses his hard-news policy, believers are better off; whereas if he uses his soft-news policy all decision-makers are indifferent. There cannot exist an asymmetric equilibrium such that one expert (say α) uses his hard-news policy whereas the other expert (say β) uses his soft-news policy. With the hard-news policy, believers (say subgroup 1) get a positive information gain, that is, $\Delta_{1\alpha} > \Delta_{1\beta} = 0$. Therefore, they are not indifferent about the allocation of attention. The alternative asymmetric equilibria is such that one expert (say α) uses his hard-news policy whereas the other expert (say β) is babbling. This requires the second expert to collect attention only from his believers, that is, $g_{2\beta} = 1$. Such asymmetric equilibria are equivalent to a monopoly with the hard-news policy in terms of information gains. For these equilibria to exist, there must be at least one expert such that as a monopolist he would use his hard-news policy. In this case, a sufficiently small mass of sceptics can devote attention to the other expert without changing the monopolist's optimal policy. If each expert as monopolist would use his soft-news policy, the mass of believers must be reduced to switch in favour of his hard-news policy. However, this is not compatible with the second expert babbling.

In any equilibrium with (at least) a babbling expert, those who devote attention to the latter receive information of the lowest quality. Indeed, babbling is the least informative outcome following Blackwell (1953): posterior beliefs are equal to prior beliefs. Instead, the hard-news policy and the soft-news policy produce both some dispersion in posterior beliefs. In any asymmetric equilibrium where each expert uses his soft-news policy, each decision-maker is equally informed. By (6)-(7),

$$\mu_1(\omega_1 \,|\, s) - \mu_1(\omega_1 \,|\, s') = \mu_2(\omega_1 \,|\, s) - \mu_2(\omega_1 \,|\, s') = \frac{\phi_1 - \phi_2}{2\,[\phi_1 + \phi_2]} < \frac{1}{2}$$

Therefore, in a monopoly each decision-maker is better (equally) informed if the expert uses his hard-news (soft-news) policy. □

Proof of Proposition 4

Proof. I denote with g the fraction of decision-makers belonging to the subgroup $i = 1$, that is, $g := \frac{|\{d \in D_1\}|}{|D|}$. Note that $g = g_{1j}$ when j is the monopolist. When there are two experts, that is $J = \{\alpha, \beta\}$, $g = g_\alpha + g_\beta$ where $g_j := \frac{|\{d \in H_j | d \in D_1\}|}{|D|}$. Similarly, $1 - g$ is the fraction of decision-makers belonging to the subgroup $i = 2$ and $1 - g = g'_\alpha + g'_\beta$ where $g'_j := \frac{|\{d \in H_j | d \in D_2\}|}{|D|}$. Note that $g_{1\alpha} = \frac{g_\alpha}{g_\alpha + g'_\alpha}$ and $g_{2\beta} = \frac{g'_\beta}{g_\beta + g'_\beta}$. I define news informativeness ψ_{ij} as the range of posterior beliefs for any subgroup of decision-makers $i \in I$ and any expert $j \in J$:

$$\psi_{i\alpha} = \begin{cases} \frac{\phi_i}{\phi_i + \phi_2} & \text{if (9) holds} \\ \frac{\phi_1 - \phi_2}{2(\phi_1 + \phi_2)} & \text{otherwise} \end{cases} \qquad \psi_{i\beta} = \begin{cases} \frac{\phi_1}{\phi_1 + \phi_i} & \text{if (10) holds} \\ \frac{\phi_1 - \phi_2}{2(\phi_1 + \phi_2)} & \text{otherwise} \end{cases}$$

Then, I define Ψ as the weighted sum of decision-makers' ranges of posterior beliefs:

$$\Psi := g_\alpha \psi_{1\alpha} + g'_\alpha \psi_{2\alpha} + g_\beta \psi_{1\beta} + g'_\beta \psi_{2\beta} \tag{13}$$

If the expert j is the monopolist, then the average quality of information is:

$$\Psi_j^M := g\psi_{1j} + (1-g)\psi_{2j}$$

Here, I compare $\Psi_\alpha^M, \Psi_\beta^M$ with Ψ to determine whether a platform can make a setting with two experts more informative than a monopoly. There are two cases to consider:

1. If each expert as monopolist uses his soft-news policy - that is, (9)-(10) do not hold given g - then a setting with two experts is always better. By Lemma 2, opposite-bias learning is more informative than a monopoly with the soft-news policy. The platform can do even better than opposite-bias learning by allocating some believers to each expert, that is $g_\alpha, g'_\beta > 0$, making sure that (9)-(10) hold true.

2. When at least one expert as monopolist uses his hard-news policy, the result depends on the degrees of tolerance $\hat{g}_\alpha$ and $\hat{g}_\beta$. I assume without loss of generality that the expert α uses his hard-news policy as a monopolist. First of all, I show that a setting with two experts must be better if $\hat{g}_\alpha, \hat{g}_\beta > \frac{1}{2}$. Note that, by assumption, $g < \hat{g}_\alpha$ and $\Psi_\alpha^M = g\left(\frac{\phi_1}{\phi_1+\phi_2}\right) + \frac{1-g}{2}$. Consider a fraction $\epsilon \in (0, 1-g)$ of sceptics of α (believers of β) and set $g'_\beta = \epsilon$. In a setting with two experts, the expert β uses his hard-news policy if $g_{2\beta} = \frac{\epsilon}{\epsilon+g_\beta} \leq \hat{g}_\beta$. This is equivalent to $g_\beta \geq \left(\frac{1-\hat{g}_\beta}{\hat{g}_\beta}\right)\epsilon := \epsilon' < \epsilon$. Now, let $g_\alpha = g - \epsilon'$ and $g'_\alpha = 1 - g - \epsilon$ such that $g_{1\alpha} = \frac{g-\epsilon}{1-\epsilon-\epsilon'} \leq \hat{g}_\alpha$. It follows that:

$$\Psi = (g-\epsilon')\left(\frac{\phi_1}{\phi_1+\phi_2}\right) + \frac{1-g-\epsilon}{2} + \epsilon\left(\frac{\phi_1}{\phi_1+\phi_2}\right) + \frac{\epsilon'}{2}$$

and the change in average quality of information is positive:

$$\Delta\Psi := \Psi - \Psi_\alpha^M = \epsilon\left(\frac{\phi_1}{\phi_1+\phi_2}\right) + \frac{\epsilon'}{2} - \epsilon'\left(\frac{\phi_1}{\phi_1+\phi_2}\right) - \frac{\epsilon}{2} = (\epsilon-\epsilon')\left(\frac{\phi_1}{\phi_1+\phi_2} - \frac{1}{2}\right) > 0$$

If $\hat{g}_\alpha > \frac{1}{2}$ whereas $\hat{g}_\beta < \frac{1}{2}$, the steps are similar but the result is opposite. Indeed, $\epsilon' > \epsilon$ and therefore $\Delta\Psi < 0$. Hence, the monopoly (of expert α) is better. If $\hat{g}_\alpha < \frac{1}{2}$ whereas $\hat{g}_\beta > \frac{1}{2}$, there are two cases to consider. When each expert as monopolist uses his hard-news policy, it must be the case that $g < \frac{1}{2}$ and therefore $\Psi_\beta^M > \Psi_\alpha^M$. Then, the previous logic applies to the monopoly of expert β, which is the best outcome. Instead, when the expert β as monopolist uses the soft-news policy (that is, $1-g > \hat{g}_\beta$), the monopoly of β is not optimal. Here, I show that there exists a setting with two experts that outperforms the monopoly of expert α. The idea is to induce the expert β to use his hard-news policy. Let $g_\beta = g$. Then, it must hold $g_{2\beta} = \frac{g'_\beta}{g'_\beta+g} \leq \hat{g}_\beta$. This is equivalent to $g'_\beta \leq \left(\frac{\hat{g}_\beta}{1-\hat{g}_\beta}\right)g > g$. Let $g'_\beta = \left(\frac{\hat{g}_\beta}{1-\hat{g}_\beta}\right)g$ and, by definition, $g'_\alpha = 1-g-g'_\beta$. It follows that:

$$\Psi = g'_\beta\left(\frac{\phi_1}{\phi_1+\phi_2}\right) + \frac{1-g'_\beta}{2}$$

and the change in the average quality of information is positive:

$$\Delta\Psi = (g'_\beta - g)\left(\frac{\phi_1}{\phi_1+\phi_2}\right) - \left(\frac{g'_\beta - g}{2}\right) = (g'_\beta - g)\left(\frac{\phi_1}{\phi_1+\phi_2} - \frac{1}{2}\right) > 0$$

Finally, consider the case where $\hat{g}_\alpha, \hat{g}_\beta < \frac{1}{2}$. Assume by contradiction that each expert as monopolist uses his hard-news policy and $g < \hat{g}_\alpha < \frac{1}{2}$. Therefore, it must be the case that $1-g > \frac{1}{2} > \hat{g}_\beta$. But then the expert β uses his soft-news policy as monopolist, contradiction. Thus, the monopoly of expert α is better than the monopoly of expert β and of any setting with two experts.

□

Proof of Proposition 5

Proof. Let $|I_2| = R_2 < R$. I order the subgroups of decision-makers from the most sceptical to the least:

$$\phi_1 < \cdots < \phi_{R_2} < 1 < \cdots < \phi_R$$

For any subgroup $r \in I$, I define the value for the expert of persuading marginally subgroup r as

$$E_r := \left[\mu^0(\omega_1) + \mu^0(\omega_2)\phi_r\right] \sum_{i=r}^{R} g_i \tag{14}$$

For any $r, r' \in I$, it is possible to define the following policies:

Definition 7 (Hard-news policy). *A hard-news policy π_r, with target $T = \{r\}$ such that $r \leq R_2$, consists of a persuading message s and a residual message s' such that*

$$\pi_r(s \,|\, \omega_1) = 1 \quad \pi_r(s' \,|\, \omega_1) = 0$$

$$\pi_r(s \,|\, \omega_2) = \phi_r \quad \pi_r(s' \,|\, \omega_2) = 1 - \phi_r$$

The hard-news policy π_r implies the following posterior beliefs:

$$\mu_i(\omega_1 \,|\, s) = \frac{\phi_i}{\phi_i + \phi_r}, \quad \mu_i(\omega_1 \,|\, s') = 0 \quad \forall i \in I \tag{15}$$

Definition 8 (Soft-news policy). *A soft-news policy $\pi_{\{r,r'\}}$, with targets $T = \{r, r'\}$ such that $r \leq R_2$ and $r' > R_2$, consists of two messages s, s' such that*

$$\pi_{\{r,r'\}}(s \,|\, \omega_1) = k \quad \pi_{\{r,r'\}}(s' \,|\, \omega_1) = 1 - k$$

$$\pi_{\{r,r'\}}(s \,|\, \omega_2) = \phi_r k \quad \pi_{\{r,r'\}}(s' \,|\, \omega_2) = \phi_{r'}(1 - k)$$

where

$$k := \frac{\phi_{r'} - 1}{\phi_{r'} - \phi_r}$$

is strictly increasing in $\phi_r \in [0, 1]$ and $\phi_{r'} \in [1, \infty]$.

The soft-news policy $\pi_{\{r,r'\}}$ implies the following posterior beliefs:

$$\mu_i(\omega_1 \,|\, s) = \frac{\phi_i}{\phi_i + \phi_r}, \quad \mu_i(\omega_1 \,|\, s') = \frac{\phi_i}{\phi_i + \phi_{r'}} \quad \forall i \in I \tag{16}$$

The payoff of a hard-news policy is

$$V_r := E_r$$

whereas the payoff of a soft-news policy is

$$V_{\{r,r'\}} := kE_r + (1 - k)E_{r'}$$

The payoff from the truth-telling policy is $V_t = \mu^0(\omega_1)$ and $V_1 > V_t$. The payoff from babbling is $V_u = G_1 := \sum_{i=R_2+1}^{R} g_i$. Note that $V_{\{r,R_2+1\}} > V_u$. Therefore, babbling is not optimal. I assume that there exist a unique $r^* = \arg\max_r E_r$. It follows that a monopolist uses optimally either a hard-news policy or a soft-news policy. This assumption rules out, for instance, any linear combination of hard-news policies targeting different subgroups of sceptics. If $r^* \leq R_2$, a hard-news policy with $T = \{r^*\}$ is optimal. Clearly $V_{r^*} > V_r$ for any $r \leq R_2$ and $r \neq r^*$. Moreover $V_{r^*} > V_{\{r,r'\}}$ as $E_{r^*} \geq E_r$ and $E_{r^*} > E_{r'}$ for any $r \leq R_2$ and any $r' > R_2$. If $r^* > R_2$, clearly $V_{\{r,r^*\}} > V_r$ for any $r \leq R_2$. Therefore, a soft-news policy is optimal. However, r^* is not necessarily the target: for any $r \leq R_2$, $V_{\{r,r^*\}} < V_{\{r,r'\}}$ if there exists a subgroup of believers $r' < r^*$ such that the difference $E_{r^*} - E_{r'}$ is sufficiently small. □

Proof of Proposition 6

Proof. The value of being persuaded marginally - a generalization of expression (14) - is:

$$E_\phi := \left[\mu^0(\omega_1) + \mu^0(\omega_2)\phi\right][1 - F(\phi)]$$

As suggested by Proposition 5, the expert uses a hard-news policy or a soft-news policy depending on whether the solution to $\max_\phi E_\phi$ belongs to $[0,1]$ or to $[1,\infty)$, respectively. The F.O.C. is:

$$\mu^0(\omega_2)[1 - F(\phi)] - f(\phi)\left[\mu^0(\omega_1) + \mu^0(\omega_2)\phi\right] = 0$$

and implies condition (11), whereas the S.O.C. is:

$$-2\mu^0(\omega_2)f(\phi) - f'(\phi)\left[\mu^0(\omega_1) + \mu^0(\omega_2)\phi\right] < 0$$

which implies

$$\frac{f'(\phi)}{f(\phi)} > -\frac{2}{\phi_j + \phi} \tag{17}$$

Clearly, if the F.O.C. is always negative/positive (or the S.O.C. is violated) there exist a corner solution, namely the most valuable subgroup is $\phi = 0$ or $\phi = \infty$. Following Proposition 5, $\phi = 0$ implies the truth-telling policy, which is a special case of a hard-news policy in this setting. Instead, $\phi = \infty$ does not imply necessarily that such subgroup is a target. The actual targets of the soft-news policy depends on the shape of $F(\cdot)$. A sufficient condition for uniqueness is an increasing hazard rate function. □

Proof of Lemma 5

Proof. Let us consider two hard-news policies π_r and $\pi_{r'}$, with targets $T = \{r\}$ and $T = \{r'\}$ respectively, such that $r < r'$. Then, π_r is more informative than $\pi_{r'}$ for any $i \in I$, according to the order from Blackwell (1953). This follows by (15) and $\phi_r < \phi_{r'}$.

Now, let us consider two soft-news policies $\pi_{\{r,r'\}}$ and $\pi_{\{r,r''\}}$, with targets $T = \{r, r'\}$ and $T = \{r, r''\}$ respectively, such that $r' > r''$. Then, $\pi_{\{r,r'\}}$ is more informative than $\pi_{\{r,r''\}}$ for any $i \in I$, according to the order from Blackwell (1953). This follows by (16) and $\phi_{r'} > \phi_{r''}$.

Finally, let us consider a hard-news policy with target $T = \{r\}$ and a soft-news policy with targets $T = \{r', r''\}$. If $r < r'$, Lemma 2 extends. If $r > r'$, there are two opposite effects: on the one hand, moving from a hard-news policy targeting r to another targeting r' increases informativeness; on the other hand, moving from a hard-news policy to a soft-news policy reduces informativeness. For each subgroup $i \in I$, with the hard-news policy, by (15):

$$\mu_i(\omega_1 \,|\, s) - \mu_i(\omega_1 \,|\, s') = \frac{\phi_i}{\phi_i + \phi_r}$$

whereas with the soft-news policy, by (16):

$$\mu_i(\omega_1 \,|\, s) - \mu_i(\omega_1 \,|\, s') = \frac{\phi_i}{\phi_i + \phi_{r'}} - \frac{\phi_i}{\phi_i + \phi_{r''}}$$

The hard-news policy is more informative if the following holds:

$$\frac{\phi_i + \phi_{r'}}{\phi_i + \phi_r} > \frac{\phi_{r''} - \phi_{r'}}{\phi_{r''} + \phi_i} \tag{18}$$

This condition may fail, especially if subgroup i are sceptics. □

Proof of Proposition 7

Proof. If at least one expert gathers attention exclusively from believers, then his best response is babbling. This supports the existence of an equilibrium in some cases. More details in the main text. Here, I focus on showing that this is a necessary condition. I assume that both experts gather attention from some sceptics and some believers and show that this cannot be an equilibrium. By Proposition 5 each expert j uses either a hard-news policy with target r_j or a soft-news policy with targets $\{r_j, r_j'\}$. Consider a hard-news policy. It follows:

$$\lambda_i(\pi_j) = \begin{cases} \mu_i^0(\omega_2) & \text{if } i \leq r_j \\ \mu_i^0(\omega_1) + \frac{\mu_i^0(\omega_2)}{\mu_{r_j}^0(\omega_2)}\left[\mu_{r_j}^0(\omega_2) - \mu_{r_j}^0(\omega_1)\right] > \mu_i^0(\omega_2) & \text{if } i \in (r_j, R_2] \\ \mu_i^0(\omega_1) + \frac{\mu_i^0(\omega_2)}{\mu_{r_j}^0(\omega_2)}\left[\mu_{r_j}^0(\omega_2) - \mu_{r_j}^0(\omega_1)\right] > \mu_i^0(\omega_1) & \text{if } i > R_2 \end{cases}$$

Therefore, $\Delta_{ij} > 0 \iff i > r_j$.

Consider a soft-news policy. It follows:

$$\lambda_i(\pi_j) = \begin{cases} \mu_i^0(\omega_2) & \text{if } i \leq r_j \\ \mu_i^0(\omega_1)k + \frac{\mu_i^0(\omega_2)}{\mu_{r_j}^0(\omega_2)}\left[\mu_{r_j}^0(\omega_2) - \mu_{r_j}^0(\omega_1)k\right] > \mu_i^0(\omega_2) & \text{if } i \in (r_j, R_2] \\ \mu_i^0(\omega_1)k + \frac{\mu_i^0(\omega_2)}{\mu_{r_j'}^0(\omega_2)}\mu_{r_j'}^0(\omega_1)(1-k) > \mu_i^0(\omega_1) & \text{if } i \in (R_2, r_j') \\ \mu_i^0(\omega_1) & \text{if } i \geq r_j' \end{cases}$$

Therefore, $\Delta_{ij} > 0 \iff i \in (r_j, r_j')$.

There are three cases to analyse:

1. Each expert uses a hard-news policy. It follows that each expert targets a subgroup of sceptics, and they gets zero information gain. Such sceptics can deviate, become believers of the other expert, and get a positive information gain.

2. One expert uses a soft-news policy whereas the other uses a hard-news policy. The sceptics targeted by the soft-news policy can deviate, become believers of the other expert, and get a positive information gain.

3. Each experts uses a soft-news policy. Let $T_\alpha = \{r_\alpha, r_\alpha'\}$ and $T_\beta = \{r_\beta, r_\beta'\}$ be the set of targets for the experts α and β respectively. I assume without loss of generality that $r_\alpha < r_\beta' \leq R_2 < r_\beta < r_\alpha'$. By Lemma 4, each target experiences zero information gain. Those targets who have intermediate prior beliefs (in this case, r_β' and r_β) have incentives to deviate, to get a positive information gain.

□

Proof of Proposition 8

Proof. To prove the result, I distinguish between symmetric and asymmetric equilibria.

Symmetric equilibria In the following, I compare the optimal policies of an informative expert in two scenario: monopoly and partial echo chambers. The difference is that in partial echo chambers some sceptics devote attention to the other expert, who is babbling. I denote with $\hat{r}$ the most sceptical subgroup of decision-makers who in partial echo chambers devote attention to the informative expert. There are two cases to consider:

1. The expert uses a hard-news policy in monopoly. Let r be the target under monopoly. If $\hat{r} \le r$, by Proposition 5, the subgroup with the highest value of being marginal persuaded is still r. Therefore, the expert uses the corresponding hard-news policy. Decision-makers of any subgroup $i < \hat{r}$ are indifferent about the allocation of attention, that is, get zero information gain in any case. However, because they devote attention to the babbling expert, they get lower quality information. If $\hat{r} > r$, then the subgroup of sceptics that is targeted must change, and the new target is $r' > r$. The new policy could be either a hard-news policy or a soft-news policy. In both cases, all decision-makers have a (weakly) lower information gain and, by Lemma 5, receive information of lower quality.

2. The expert uses a soft-news policy in monopoly with targets $T = \{r, r'\}$. For any $\hat{r} \le R_2$, a subgroup of believers has the highest value of being marginal persuaded. Therefore, by Proposition 5, the expert uses a soft-news policy in partial echo chambers. If $\hat{r} \le r$, the expert's payoffs do not change, thus the expert uses the same soft-news policy. Decision-makers of any subgroup $i < \hat{r}$ are indifferent about the allocation of attention, but they get lower quality information. If $\hat{r} > r$, the new targets are $\hat{T} = \{i, i'\}$, where $i > r$. Now, if $i' \le r'$ all decision-makers have a (weakly) lower information gain and, by Lemma 5, receive information of lower quality.

In the following, I find a sufficient condition for $i' \le r'$. The optimal policy in monopoly is the soft-news policy with the highest payoff. Therefore, it is the solution of the following maximization problem:

$$\max_{\phi_r, \phi_{r'}} k\left[\mu_j^0(\omega_1) + \mu_j^0(\omega_2)\phi_r\right]\left[1 - F(\phi_r)\right] + (1-k)\left[\mu_j^0(\omega_1) + \mu_j^0(\omega_2)\phi_{r'}\right]\left[1 - F(\phi_{r'})\right]$$

subject to $k = \frac{\phi_{r'} - 1}{\phi_{r'} - \phi_r}$, $\phi_r \in [0,1]$ and $\phi_{r'} \in [1, \infty)$. The F.O.C. are:

$$\Psi^F_{\phi_r} := \frac{\partial k}{\partial \phi_r}\Big\{\left[\mu_j^0(\omega_1) + \mu_j^0(\omega_2)\phi_r\right]\left[1 - F(\phi_r)\right] - \left[\mu_j^0(\omega_1) + \mu_j^0(\omega_2)\phi_{r'}\right]\left[1 - F(\phi_{r'})\right]\Big\} +$$

$$+k\mu_j^0(\omega_2)\left[1 - F(\phi_r)\right] - kf(\phi_r)\left[\mu_j^0(\omega_1) + \mu_j^0(\omega_2)\phi_r\right] = 0$$

$$\Psi^F_{\phi_{r'}} := \frac{\partial k}{\partial \phi_{r'}}\Big\{\left[\mu_j^0(\omega_1) + \mu_j^0(\omega_2)\phi_r\right]\left[1 - F(\phi_r)\right] - \left[\mu_j^0(\omega_1) + \mu_j^0(\omega_2)\phi_{r'}\right]\left[1 - F(\phi_{r'})\right]\Big\} +$$

$$+(1-k)\mu_j^0(\omega_2)\left[1 - F(\phi_{r'})\right] - (1-k)f(\phi_{r'})\left[\mu_j^0(\omega_1) + \mu_j^0(\omega_2)\phi_{r'}\right] = 0$$

In partial echo chambers, the distribution of beliefs changes. In particular, I denote with $G(\cdot)$ the new distribution that the informative expert faces. By (4), it follows

$$g(\phi_i) = \begin{cases} 0 & \text{if } i < \hat{r} \\ \frac{f(\phi_i)}{1 - F(\phi_{\hat{r}})} & \text{if } i \ge \hat{r} \end{cases} \implies 1 - G(\phi_i) = \begin{cases} 1 & \text{if } i < \hat{r} \\ \frac{1 - F(\phi_i)}{1 - F(\phi_{\hat{r}})} & \text{if } i \ge \hat{r} \end{cases}$$

Therefore, $\Psi^F_{\phi_r} = \Psi^G_{\phi_r}$ and $\Psi^F_{\phi_{r'}} = \Psi^G_{\phi_{r'}}$ for any $i \ge \hat{r}$, which is the subset of possible targets of the informative expert. Because it must hold that the new targets as sceptics are a subgroup $i > r$, then $i' \le r'$ if the targets are strategic substitutes, that is if $\frac{\partial \Psi_{\phi_{r'}}}{\partial \phi_r} \le 0$.

There exist other symmetric equilibria where disjoint subsets of sceptics devote attention to the babbling expert. These equilibria do not differ significantly from partial echo chambers and, under the previous conditions, are worse for decision-makers than some monopoly outcome. In particular, there cannot exists an equilibrium where i devotes attention to the babbling expert and $i \ge r$, where r is the target of the informative expert.

Asymmetric equilibria Any symmetric equilibria described before is such that decision-makers devoting attention to the babbling expert are indifferent about the allocation of attention. Therefore, there exists asymmetric equilibria where decision-makers belonging to the corresponding subgroups behave differently in terms of allocation of attention. However, these equilibria do not differ significantly from the symmetric equilibria, and the result that media pluralism is harmful holds true under the previous conditions.

Finally, there could exists asymmetric equilibria where both experts use soft-news policies with the same targets. If targets were different, some targeted decision-makers would find optimal to deviate (for the same logic of the proof of Proposition 7). I denote with $F_\alpha(\cdot)$ and $F_\beta(\cdot)$ the distributions of beliefs that the two experts α and β face, respectively. If these distributions are atomless, then the two experts target the same subgroups only if they face the same distribution, that is $F_\alpha(\cdot) = F_\beta(\cdot) = F(\cdot)$, and have the same prior beliefs, almost surely. Therefore, $F(\cdot)$ must coincide with the distribution that a monopolist face. It follows that the monopolist must have the same targets. Hence, these equilibria are equivalent to a monopoly. □

B. Online Appendix

B.1. Costly Attention

The results in my paper are derived under the assumption that each decision-maker can devote attention to just one expert. Now, I endogenize this decision by allowing each decision-maker to devote attention to a second expert at a cost $c \geq 0$.

Proposition 9. *Full revelation is an equilibrium if and only if* $c = 0$.

Assume that π_α and π_β are truth-telling policies. It follows that $\lambda_i(\pi_\alpha) = \lambda_i(\pi_\beta) = \lambda_i(\pi_J) = 1$ for any $i \in I$. Therefore, it is sufficient to devote attention to one expert to maximize the subjective probability of taking the correct action. If $c = 0$, decision-makers can pay attention to both experts without any cost. This is equivalent to unlimited attention, and full revelation is indeed the equilibrium in such a setting. If $c > 0$, each decision-maker strictly prefers to devote attention to just one expert, as she gains no additional information from the second one. However, it is not optimal for the experts to reveal the true state when decision-makers pay attention to only one expert.

The equilibria of the game are robust for any $c \geq 0$. Given any equilibrium, it follows by Proposition 7 that there is no incentive to devote attention to a second expert. Multi-homing is not optimal because at least one expert is babbling. For instance, consider partial echo chambers with β babbling. For any $i \in H_\alpha$, it holds $\lambda_i(\pi_\alpha) = \lambda_i(\pi_J)$ because π_β does not affects posterior beliefs, hence optimal actions. For any $i \in H_\beta$ it must be the case that both experts are providing zero information gains, and $\lambda_i(\pi_\alpha) = \lambda_i(\pi_\beta) = \lambda_i(\pi_J) = \mu_i^0(\omega_m)$. Therefore, decision-makers are not willing to pay $c \geq 0$ to devote attention to a second expert.

B.2. Costly Information

In the paper, I assume that the information is costless to produce for experts and to process for decision-makers. Here, I study the effects of relaxing this assumption. In particular, I assume that either experts or decision-makers have to pay an entropy cost (Gentzkow and Kamenica, 2014; Matysková and Montes, 2021). For any policy π_j by expert j and any decision-maker d, its cost is proportional to the expected reduction in uncertainty:

$$c(\pi_j) := \chi[H(\mu_d^0) - \sum_{s \in S_j} \pi_j(s) H(\mu_d(\cdot|s))]$$

where $H(\mu) := -[\mu(\omega_1)\ln(\mu(\omega_1)) + (1 - \mu(\omega_1))\ln(1 - \mu(\omega_1))]$ is the entropy and $\chi > 0$ is a parameter. The cost of babbling is zero by definition. Following Bayesian plausibility and strict concavity of $H(\cdot)$,

$$H(\mu_d^0) = H\left(\sum_{s \in S_j} \pi_j(s)\mu_d(\cdot|s)\right) > \sum_{s \in S_j} \pi_j(s) H(\mu_d(\cdot|s))$$

Therefore, it holds that $c(\pi_j) > 0$ for any policy π_j different from babbling.

When decision-makers bear this cost, a decision-maker is not indifferent between being a target of an expert and receiving babbling: she prefers the second option. As a consequence, all the equilibria with one informative expert and one babbling expert, for instance partial echo chambers, are not robust to this extension. The unique symmetric equilibrium is echo chambers. Remarkably, an entropy cost by news consumers can be interpreted as a form of confirmation bias. In particular, news consumers bear a cognitive cost every time they change their beliefs. I show that even a very small confirmation bias makes echo chambers the unique robust equilibrium.

When experts bear this cost, the optimal policies change as shown by Gentzkow and Kamenica (2014). In particular, it is costly to induce extreme posterior beliefs. However, the objective of an expert is to persuade decision-makers i.e., to make them just indifferent. Extreme posterior beliefs are an indirect effect, see for instance the hard-news policy. It turns out that a posterior belief $\mu = \frac{1}{2}$ is the cheapest for an expert, that is $\arg\max_\mu H(\mu) = \frac{1}{2}$. Therefore, experts keep targeting decision-makers, unless information is so costly that babbling is the best option. Lemma 4 continue to hold and the incentives of decision-makers about the allocation of attention are not affected. The game has the same equilibria but costly information reduces the overall quality of information. Nevertheless, the negative effect of media pluralism on quality continues to exist.

B.3. Multi-Homing

In Section B.1, I show that decision-makers have no incentive to devote attention to a second expert in equilibrium. Here, I assume that some decision-makers are exogenously multi-homing and study the impact on the equilibria of the game.

The first result is that full revelation cannot be achieved unless all decision makers have unlimited attention. Given truth-telling by the rival, multi-homing decision-makers cannot be persuaded. Therefore, the expert can focus on single-homing decision-makers and persuade them. This incentive exists independently on the share of single-homing decision-makers. Indeed, there is no cost on the multi-homing side from using a policy different from truth-telling.

Differently, targeting multi-homing decision-makers is costly for an expert because it lowers the probability to persuade single-homing decision-makers by ϵ arbitrarily small but positive. Therefore, when the set of multi-homing decision-makers has zero measure, there is no incentive to deviate (by Lebesgue's theorem). In this case, the equilibria of the game are robust.

When there is a positive mass of multi-homing decision-makers and experts find it optimal to target them, this creates the same undercutting incentives that exist with competition under unlimited attention. Therefore, a policy involving persuasion of multi-homing decision-makers cannot be part of an equilibrium. Theorem 5b in Dasgupta and Maskin (1986) can be used to establish the existence of mixed-strategy equilibria. However, the characterization and the interpretation of mixed-strategies is a hard task. It is not possible to establish a priori whether the consequent equilibrium of the game is better or worse than a monopoly in terms of information quality.

B.4. Alternative Timing

In the main text, I assume that optimal persuasion and the allocation of attention are simultaneous. Now, I examine the possibility that the two are sequential.

If the allocation of attention is chosen *before* persuasion takes place, my results extend. Remarkably, a monopoly is a much more credible equilibrium in this case. The allocation of attention cannot react to optimal persuasion by a monopolist. Therefore, it does not matter what is the policy of the non-active expert in the second stage of the game.

If the allocation of attention is chosen *after* persuasion takes place, babbling by both experts (with any allocation of attention) is not an equilibrium. Suppose, by contradiction, the opposite. Believers take each expert's preferred action, but any expert can deviate and persuade also his sceptics with positive probability (for instance, with his soft-news policy).

To do so, it is sufficient to provide a strictly positive information gain, which requires to avoid targeting sceptics.

At the same time, truth-telling is the equilibrium policy. If any expert deviates, he does not collect attention. Therefore, he is not able to persuade, and indifference follows. This result is in line with Knoepfle (2020). Experts are implicitly attention-seekers: persuasion is effective only if an expert gets attention in the second stage. Optimal persuasion involves targeting of some decision-makers. However, by Lemma 4 a target gets zero information gain from persuasion. Therefore, she is unlikely to devote attention in the second stage of the game.

The latter setting is in line with the literature on media bias, where consumers buy news knowing the media's reputation or slant (Gentzkow et al., 2015). In turn, the latter is influenced by the incentive to steal consumers from the rival, and this is likely to generate beneficial competition. My approach is different because I assume that persuasion is rather flexible compared to the attention habits. Experts behave strategically taking as given the allocation of attention, and this is a source of persuasion power.

B.5. Second-movers

In this section, I maintain the timing as in the paper. However, decision-makers have the faculty of adjusting their allocation of attention at a cost $\zeta \geq 0$ after the reporting policies have been settled.

Proposition 10. *Full revelation is the equilibrium if and only if $\zeta = 0$.*

Full revelation requires all decision-makers to be second-movers. Assume by contradiction that there is one decision-maker who does not adjust her attention habit to experts' reporting policies. Then, the expert who receives her attention has an incentive to persuade her. Indeed, given truth-telling by the rival, the expert can deviate from truth-telling: he loses the attention of the second-movers, but this does not affect his payoff. At the same time, given full revelation, a decision-maker is not willing to pay a positive cost to be a second-mover. Indeed, she is already achieving the highest payoff, independently of whom she pays attention. Therefore, full revelation is the equilibrium only if $\zeta = 0$ and all decision-makers are second-movers.

The equilibria that I have identified in the paper are robust if ζ is large enough. I take the perspective of expert α without loss of generality. Expert α can attract second-movers of subgroup i only if $\zeta \leq \lambda_i(\pi_\alpha) - \lambda_i(\pi_\beta)$. As an illustration, I consider the echo chambers equilibrium. In this case, expert α can attract his sceptics $i = 2$ as second-movers if $\lambda_2(\pi_\alpha) \geq \mu_2^0(\omega_2) + \zeta$. Therefore, a sufficient condition for the robustness of echo chambers is $\zeta > \mu_2^0(\omega_1)$. Remarkably, the higher polarization (the more extreme believers' prior belief), the lower the threshold of ζ for echo chambers to be robust. If $\zeta \leq \mu_2^0(\omega_1)$ and $\mu_\alpha^0(\omega_1) \geq g_1$, then truth-telling is a beneficial deviation for expert α. Indeed, expert α prefers to attract his sceptics and reveal the truth to all decision-makers instead of exploiting his echo chamber. If this is not the case, expert α must persuade sceptics (to some extent) to find it optimal to deviate from echo chambers. However, this lowers sceptics' payoff and hence the threshold for ζ that makes echo chambers robust. Finally, when ζ is positive but small enough, there exist mixed-strategy equilibria, as in Section B.3.

B.6. Partial Commitment

In the paper, I assume that the experts can fully commit to their reporting policies. Trivially, in a cheap talk model the unique possible outcome is babbling by experts, and therefore the model has no predictive power. Here, I study the intermediate cases between cheap talk and Bayesian persuasion, in the spirit of Min (2021). In particular, I assume that with probability $\delta \in (0,1)$ the expert can deviate from his reporting policy. Therefore, each decision-maker expects that:

- with probability $1-\delta$, the message she receives is originated from the expert's optimal policy π;
- with probability δ, the message amounts to babbling.

This changes the way each decision-maker updates beliefs, and therefore changes the persuasion constraints. I assume that the expert wants to persuade to take action a_1. The persuasion constraint for subgroup i is:

$$\pi(s \mid \omega_2) \le \phi_i \pi(s \mid \omega_1) + \left(\frac{\delta}{1-\delta}\right)(\phi_i - 1) \tag{19}$$

Clearly, when $\delta = 0$ the persuasion constraint (19) becomes (3). One first result of partial commitment is that it may make impossible to persuade sceptics. An expert can design a hard-news policy targeting a subgroup i of sceptics only if $\phi_i \ge \delta$.

The second effect of partial commitment is that a targeted sceptic has a positive information gain. In order to see this, I generalize the definitions of hard-news and soft-news policies:

Definition 9 (Hard-news policy). *A hard-news policy π_r, with target $T = \{r\}$ such that $r \le R_2$ and $\phi_r \ge \delta$, consists of a persuading message s and a residual message s' such that*

$$\pi_r(s \mid \omega_1) = 1 \quad \pi_r(s' \mid \omega_1) = 0$$

$$\pi_r(s \mid \omega_2) = \frac{\phi_r - \delta}{1-\delta} \quad \pi_r(s' \mid \omega_2) = 1 - \frac{\phi_r - \delta}{1-\delta}$$

The corresponding subjective probability of taking the correct action is:

$$\lambda_i(\pi_r) = \begin{cases} \mu_i^0(\omega_2) & \text{if } i < r \\ (1-\delta)\mu_i^0(\omega_1) + \frac{\mu_i^0(\omega_2)}{\mu_r^0(\omega_2)}\left[\mu_r^0(\omega_2) - \mu_r^0(\omega_1)\right] + \delta\mu_i^0(\omega_2) & \text{if } i \in [r, R_2] \\ \mu_i^0(\omega_1) + \frac{\mu_i^0(\omega_2)}{\mu_r^0(\omega_2)}\left[\mu_r^0(\omega_2) - \mu_r^0(\omega_1)\right] & \text{if } i > R_2 \end{cases}$$

Therefore, $\Delta_i > 0 \iff i \ge r$.

Definition 10 (Soft-news policy). *A soft-news policy $\pi_{\{r,r'\}}$, with targets $T = \{r, r'\}$ such that $r \le R_2$, $\phi_r \ge \delta$ and $r' > R_2$, consists of two messages s, s' such that*

$$\pi_{\{r,r'\}}(s \mid \omega_1) = k \quad \pi_{\{r,r'\}}(s' \mid \omega_1) = 1 - k$$

$$\pi_{\{r,r'\}}(s \mid \omega_2) = \phi_r k + \left(\frac{\delta}{1-\delta}\right)(\phi_r - 1) \quad \pi_{\{r,r'\}}(s' \mid \omega_2) = \phi_{r'}(1-k) + \left(\frac{\delta}{1-\delta}\right)(\phi_{r'} - 1)$$

where

$$k := \frac{\phi_{r'} - c}{\phi_{r'} - \phi_r} \quad \textit{and} \quad c := \frac{1 - \delta(\phi_r + \phi_{r'} - 1)}{1-\delta}$$

The corresponding subjective probability of taking the correct action is:

$$\lambda_i(\pi_{\{r,r'\}}) = \begin{cases} \mu_i^0(\omega_2) & \text{if } i < r \\ (1-\delta)\mu_i^0(\omega_1)k + \frac{\mu_i^0(\omega_2)}{\mu_r^0(\omega_2)}\left[\mu_r^0(\omega_2) - \mu_r^0(\omega_1)(k(1-\delta)+\delta)\right] + \delta\mu_i^0(\omega_2) & \text{if } i \in [r, R_2] \\ (1-\delta)\mu_i^0(\omega_1)k + \mu_i^0(\omega_2)\left[\phi_{r'}((1-k)(1-\delta)+\delta) - \delta\right] + \delta\mu_i^0(\omega_1) & \text{if } i \in (R_2, r') \\ \mu_i^0(\omega_1) & \text{if } i \ge r' \end{cases}$$

Therefore, $\Delta_i > 0 \iff i \in [r, r')$.

As a consequence, all the equilibria which rely on targeted sceptics being indifferent about the allocation of attention (i.e., the asymmetric equilibria) do not exist with partial commitment. Instead, the symmetric equilibria (echo chambers and partial echo chambers) are robust to this extension.

Even if targeted sceptics have a positive information gain, the most moderate among targeted sceptics still have incentives to become believers of the other expert. For instance, when both experts α and β use hard-news policies with targets $T_\alpha = \{r_\alpha\}$ and $T_\beta = \{r_\beta\}$ such that $\phi_{r_\alpha} < 1 < \phi_{r_\beta}$, decision-makers of subgroup r_α (r_β) have incentives to deviate if $\phi_{r_\alpha} > \frac{1}{\phi_{r_\beta}}$ $\left(\phi_{r_\alpha} < \frac{1}{\phi_{r_\beta}}\right)$. A similar reasoning applies for any combination of experts' policies, and therefore Proposition 7 extends.

B.7. Non-Bayesian Persuasion

Decision-makers with limited attention are probably unwilling to use a complex updating rule such as Bayesian updating. Drawing from the insights in de Clippel and Zhang (2020), I study my model under the following generalized version of the persuasion constraint for subgroup i:

$$\pi(s \mid \omega_2) \leq \phi_i^\rho \pi(s \mid \omega_1) \tag{20}$$

where $\rho \geq 0$ is a parameter. Clearly, when $\rho = 1$ the persuasion constraint (20) becomes (3). When $\rho \in (0,1)$, decision-makers are subject to base-rate neglect or over-inference. Instead, when $\rho > 1$, decision-makers overweight priors or are subject to under-inference. Let $\hat{\phi}_i = \phi_i^\rho$. Given a distribution of beliefs' ratio ϕ_i, $\rho \in (0,1)$ makes the distribution of $\hat{\phi}_i$ more moderate, whereas $\rho > 1$ makes it more extreme. In particular, if $\rho \in (0,1)$ then $\hat{\phi}_i > \phi_i$ for any $i \leq R_2$ and $\hat{\phi}_i < \phi_i$ for any $i > R_2$, whereas if $\rho > 1$ then $\hat{\phi}_i < \phi_i$ for any $i \leq R_2$ and $\hat{\phi}_i > \phi_i$ for any $i > R_2$. This is important because $\hat{\phi}_i$ is relevant for the expert's information design, whereas decision-makers keep evaluating information based on their original priors. It follows that:

$$\lambda_i(\pi_r) = \begin{cases} \mu_i^0(\omega_2) & \text{if } i < r \\ \mu_i^0(\omega_2)(\phi_i + 1 - \hat{\phi}_r) & \text{if } i > r \end{cases}$$

$$\lambda_i(\pi_{\{r,r'\}}) = \begin{cases} \mu_i^0(\omega_2) & \text{if } i < r \\ \mu_i^0(\omega_2)(\phi_i k + 1 - \hat{\phi}_r k) & \text{if } i \in [r, r') \\ \mu_i^0(\omega_1) & \text{if } i \geq r' \end{cases}$$

Therefore, $\Delta_r < 0$ if $\phi_r < \hat{\phi}_r$, that is if $\rho \in (0,1)$. Targeted sceptics have a negative information gain when are subject to base-rate neglect or over-inference. In this scenario, the unique equilibrium of the game is echo chambers.

When $\rho > 1$, the targeted sceptics have a positive information gain. However, as in Section B.6, the most moderate among targeted sceptics still have incentives to become believers of the other expert. In particular, either $\lambda_{r_\alpha}(\pi_\beta) > \lambda_{r_\alpha}(\pi_\alpha) \iff (\phi_{r_\alpha}\phi_{r_\beta})^\rho > \phi_{r_\alpha}$ or $\lambda_{r_\beta}(\pi_\alpha) > \lambda_{r_\beta}(\pi_\beta) \iff (\phi_{r_\alpha}\phi_{r_\beta})^\rho < \phi_{r_\beta}$ hold. Therefore, Proposition 7 extends.

B.8. Profit-maximizing experts

In the paper, I assume that experts are biased. Each expert has a preferred action and achieves positive utility only if a decision-maker takes such an action. Here, I modify experts' preferences by introducing a second component that captures each expert's desire to gather attention. In particular, the payoff of expert j from a decision-maker who takes action $a \in A$ and devotes attention to expert $j_d \in J$ is:

$$u_j(a, j_d) := \mathbb{1}\{a = a_j\} + \gamma \mathbb{1}\{j_d = j\}$$

See (2) for a comparison. The models are equivalent when $\gamma = 0$. Each expert is better off the larger is his audience, but this does not affect the equilibria of the game. Indeed, when experts decide their reporting policies, they take as given their respective audiences. In other words, a change in π_j can influence the action a taken by a decision-maker d but does not affect her allocation of attention j_d. In particular, an expert does not expand his audience by making his reporting policy more informative. Therefore, his reporting policy is oriented uniquely by the persuasion motive, as in the baseline model.

B.9. Homogenous Experts

With unlimited attention, having two experts with the same preferences does not affect information provision compared to a monopoly.

Proposition 11 (Homogeneous experts)**.** *Consider $J = \{\alpha, \beta\}$ and assume $a_\alpha = a_\beta$ and $\mu^0_\alpha(\omega_1) = \mu^0_\beta(\omega_1)$. In the equilibrium one expert (say α) behaves as a monopolist whereas the other one (say β) is babbling.*

Given babbling by β, α uses the optimal policy as monopolist (Proposition 1). The two experts have the same preferences and the same belief. Therefore, the policy of α is optimal also for β. There is no incentive to change the posterior beliefs by providing further information. Hence, babbling is optimal for β.

The entry of (potentially many) experts with the same preferences and belief as the incumbent is not affecting information provision. The intuition is that the entrant cannot refine the optimal policy of the incumbent.[24]

With limited attention, two experts using the same policy can be active. Indeed, each decision-maker is indifferent about her allocation of attention, as each expert provides her the same information gain.[25] This allows to extend the prediction of my model beyond a duopoly. The existence of additional experts has the effect of splitting attention, but it does not affect the equilibria of the game qualitatively.

With costly attention, a decision-maker could rationally pay attention to multiple experts providing her a positive information gain. However, multi-homing triggers a strategic response by the experts (Proposition 11). In this setting, the unique equilibrium is a monopoly.

B.10. Micro-Targeting

In the paper, persuasion is public. By contrast here, I assume that decision-makers are micro-targeted: each expert uses a specific policy for each subgroup of decision-makers. Let π^i_j be the policy of expert $j \in J$ which targets subgroup $i \in I$. In a monopoly, π^i_j is babbling if subgroup i are believers, whereas it is the hard-news policy if subgroup i are sceptics. This follows from Kamenica and Gentzkow (2011). With multiple experts and single-homing decision-makers, $\lambda_i(\pi^i_j) = \mu^0_i(\omega_m)$ for any $i \in I$ and any $j \in J$. In words, there cannot be a positive information gain from persuasion, for any decision-maker. This follows from Lemma 3 and Lemma 4. Therefore, decision-makers are indifferent about the allocation of attention.

[24] Experts with heterogeneous beliefs can have different optimal policies (in monopoly). However, there is no incentive to undercut the rival because the preferred actions coincide.

[25] If the experts use different policies, then decision-makers have incentive to devote attention to the most informative one.

An expert benefits from the possibility to target many different decision-makers. By contrast, the effect of micro-targeting on decision-makers is ambiguous: believers are always worse off, but the sceptics might benefit. For instance, assume that public persuasion is given by a soft-news policy. With micro-targeting, each subgroup of sceptics is tailored with a specific hard-news policy, and she could be better informed by Lemma 5.

Here, the equivalence between public and private persuasion (Kolotilin et al., 2017) fails because the expert knows the prior beliefs of each decision-maker.

B.11. Many States

In this section, I examine how my model can be extended allowing for more than two states of the world. A first approach is to consider a continuous state space i.e. $\Omega := [0,1]$ while keeping the action binary i.e. $A := \{a_0, a_1\}$. Here, I adopt a setting similar to Guo and Shmaya (2019). Each agent $l \in I \cup J$ has distinct prior beliefs with full support: $\mu_l^0(\cdot) \in \Delta_+(\Omega)$, where $\mu_l^0(\omega)$ is agent l's belief that the state is ω. Following Bayesian updating, posterior beliefs are:

$$\mu_i(\omega \,|\, s) := \frac{\pi(s \,|\, \omega)\mu_i^0(\omega)}{\int_0^1 \pi(s \,|\, \omega')\mu_i^0(\omega')d\omega'}$$

I assume that each decision-maker follows a threshold rule: she wants to take action a_1 if and only if the state ω is above a threshold $\bar{\omega}$. It follows that the optimal action for each decision-maker of subgroup i becomes:

$$\sigma(\mu_i) = \begin{cases} a_1 & \text{if } \int_{\bar{\omega}}^1 \mu_i(\omega)d\omega \geq \frac{1}{2} \\ a_2 & \text{otherwise} \end{cases}$$

Upon receiving message s, the implied persuasion constraint is

$$\int_{\bar{\omega}}^1 \pi(s \,|\, \omega)\mu_i^0(\omega)d\omega \geq \int_0^{\bar{\omega}} \pi(s \,|\, \omega)\mu_i^0(\omega)d\omega$$

In such a setting, I keep the restriction of two subgroups of decision-makers, believers ($i = 1$) and sceptics ($i = 2$). A believer is such that $\int_{\bar{\omega}}^1 \mu_1^0(\omega)d\omega > \frac{1}{2}$, whereas a sceptic is such that $\int_{\bar{\omega}}^1 \mu_2^0(\omega)d\omega < \frac{1}{2}$. As in the baseline model, the optimal policy focuses either on persuading sceptics or on retaining believers. However, the structure of the optimal policy changes.

If the focus is to persuade sceptics (hard-news policy), then a candidate optimal policy must satisfy the following constraint:

$$\int_{\bar{\omega}}^1 \mu_2^0(\omega)d\omega = \int_0^{\bar{\omega}} \pi(s \,|\, \omega)\mu_2^0(\omega)d\omega \tag{21}$$

I denote with Π_H the subset of policies such that (21) holds. Note that in the baseline model Π_H is singleton, whereas here the expert has degrees of freedom on the distribution of probability for each state $\omega \in [0, \bar{\omega}]$. By (5), the incentive of the expert is to pool states with high $\mu_j^0(\omega)$, while fully revealing others.

If the focus is to retain believers (soft-news policy), then a candidate optimal policy must satisfy the following constraints:

$$\int_{\bar{\omega}}^1 \pi(s \,|\, \omega)\mu_2^0(\omega)d\omega = \int_0^{\bar{\omega}} \pi(s \,|\, \omega)\mu_2^0(\omega)d\omega \tag{22}$$

$$\int_{\bar{\omega}}^1 \pi(s' \,|\, \omega)\mu_1^0(\omega)d\omega = \int_0^{\bar{\omega}} \pi(s' \,|\, \omega)\mu_1^0(\omega)d\omega \tag{23}$$

I denote with Π_S the subset of policies such that (22)-(23) hold, and note that in the baseline model Π_S is singleton. In this case, the goal of the expert is to maximize the probability of persuading sceptics subject to the constraint that believers chooses the preferred action with probability one. The incentives of the expert are difficult to disentangle, as these depend on $\mu_j^0(\omega)$, $\mu_1^0(\omega)$ and $\mu_2^0(\omega)$.

However, even if the structure of the optimal policy changes, my results are not affected. In particular, Lemma 4 generalizes to this setting. Note that

$$\int_0^{\bar{\omega}} \mu_2^0(\omega)d\omega = \int_0^{\bar{\omega}} \pi(s|\omega)\mu_2^0(\omega)d\omega + \int_0^{\bar{\omega}} \pi(s'|\omega)\mu_2^0(\omega)d\omega$$

which implies

$$\int_0^{\bar{\omega}} \pi(s'|\omega)\mu_2^0(\omega)d\omega = \int_0^{\bar{\omega}} \mu_2^0(\omega)d\omega - \int_0^{\bar{\omega}} \pi(s|\omega)\mu_2^0(\omega)d\omega$$

It follows that sceptics get zero information gain. By (22),

$$\lambda_2(\pi) = \int_{\bar{\omega}}^1 \pi(s|\omega)\mu_2^0(\omega)d\omega + \int_0^{\bar{\omega}} \pi(s'|\omega)\mu_2^0(\omega)d\omega = \int_0^{\bar{\omega}} \mu_2^0(\omega)d\omega$$

Hence, $\Delta_2 = 0$. Lemma 4 characterizes the incentives of decision-makers about the allocation of attention. Therefore, the effect of media pluralism with limited attention is unchanged.

More than two actions The analysis of optimal persuasion becomes generally intractable when the cardinality of Ω is equal to the cardinality of A.[26] I define $\phi_i(\omega,\omega') := \frac{\mu_i^0(\omega)}{\mu_i^0(\omega')}$ for any $\omega,\omega' \in \Omega$. A message s persuades decision-makers of subgroup i that the state is ω if $\pi(s|\omega') \leq \phi_i(\omega,\omega')\pi(s|\omega)$ for any $\omega' \in \Omega$. Decision-makers of subgroup i are true believers (sceptics) of state ω if $\phi_i(\omega,\omega') \geq 1$ (< 1) for any $\omega' \in \Omega$. A hard-news policy can target true sceptics. A soft-news policy can solve the trade-off between persuading true sceptics and retaining true believers. Therefore, if an expert faces only true sceptics and true believers, the result of Proposition 5 extends. However, different policies could be optimal if there exist decision-makers who believe that some states are a priori more plausible than ω, whereas others are not. Even in this scenario, there exists a decision-maker i (who is sceptical of ω relative to some other states) who is the target of the expert. Therefore, her information gain is zero. For instance, consider the following policy π:

$$\pi(s|\omega) = 1, \quad \pi(s|\omega') = \phi_i(\omega,\omega'), \quad \pi(s'|\omega') = 1 - \phi_i(\omega,\omega') \quad \text{for any } \omega' \neq \omega$$

which implies:

$$\lambda_i(\pi) = \mu_i^0(\omega) + [1 - \phi_i(\omega,\omega'')]\,\mu_i^0(\omega'') = \mu_i^0(\omega'') \iff \Delta_i = 0$$

where $\omega'' = \arg\max_\omega \mu_i^0(\omega)$. Thus, targeted (true) sceptics have incentives to deviate, as in the baseline model. Therefore, my results should not be affected by the existence of many states of the world and corresponding actions. For instance, Proposition 2 extends to this setting. True believers clustering into echo chambers is an equilibrium. Indeed, no information is provided, and hence the decision-makers do not have incentives to deviate. Similarly, Proposition 7 holds true. Targeted sceptics have zero information gain also in this setting. Therefore, they want to deviate unless there is at most one informative expert.

[26] A full characterization of prior beliefs requires $|\Omega|!$ subgroups of decision-makers. Unlike Section 6.2, there is no useful ordering of the subgroups of decision-makers.

B.12. Biased Decision-makers

In the paper, decision-makers are unbiased in their utilities. All the results are driven exclusively by heterogeneous prior beliefs. Now, I show that the same results can be obtained in a setting where decision-makers share a common prior belief $\mu^0(\omega_1)$, but each subgroup of decision-makers i is endowed with a vector of biases $b_i := \{b_i^\omega\}_{\omega\in\Omega}$. The utility of a decision-maker of subgroup i is $u_i(a,\omega_k) := \mathbb{1}\{a = a_k\}b_i^\omega$. See (1) for a comparison. The corresponding optimal action is as follows:

$$\sigma(\mu, b_i) = \begin{cases} a_1 & \text{if } \mu(\omega_1) \geq \frac{b_i^{\omega_2}}{b_i^{\omega_1}+b_i^{\omega_2}} \\ a_2 & \text{otherwise} \end{cases}$$

Upon observing message s, action a_1 is chosen if and only if:

$$\mu(\omega_1 | s) \geq \frac{b_i^{\omega_2}}{b_i^{\omega_1} + b_i^{\omega_2}} \iff \pi_j(s|\omega_2) \leq \frac{\mu^0(\omega_1)}{\mu^0(\omega_2)} \frac{b_i^{\omega_1}}{b_i^{\omega_2}} \pi_j(s|\omega_1) \tag{24}$$

A model with unbiased decision-makers and heterogeneous beliefs is equivalent to a model with biased decision-makers and a common belief only if, for any $i \in I$ and any $\omega \in \Omega$, $b_i^\omega = \frac{\mu_i^0(\omega)}{\mu^0(\omega)}$. This follows immediately from the comparison of conditions (3) and (24). Note that $b_i^\omega > 1$ if and only if $\mu_i^0(\omega) > \mu^0(\omega)$. Hence, a larger bias is equivalent to a decision-maker having a higher prior belief that the state ω is the true state. Remarkably, this multiplicative bias is different from the common definition of bias. In the literature, the utility of biased decision-makers depends on the action, but not on the state. By contrast here, each decision-maker has a strict preference to take the correct action given the state. The bias is limited to each decision-maker valuing some states more than others ex ante.

Hu et al. (2021) consider a model where decision-makers have different default actions. Given a common belief, each decision-maker would take her default action. Decision-makers of subgroup i are characterized by a specific threshold $c_i \in [0,1]$ for the posterior belief which makes them indifferent:

$$\sigma(\mu, c_i) = \begin{cases} a_1 & \text{if } \mu(\omega_1) \geq c_i \\ a_2 & \text{otherwise} \end{cases}$$

Thus, the models are equivalent if $c_i = \frac{b_i^{\omega_2}}{b_i^{\omega_1}+b_i^{\omega_2}}$.