Discussion Paper Series – CRC TR 224

Discussion Paper No. 298
Project B 05

Can Media Pluralism Be Harmful to News Quality?

Federico Innocenti [1]

November 2021
(First version : May 2021)

[1] University of Mannheim, Department of Economics and CRC. email: finnocen@mail.uni-mannheim.de

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 is gratefully acknowledged.

Can Media Pluralism Be Harmful to News Quality?*

Federico Innocenti†

November 4, 2021

Abstract

Two stylized facts characterize the Internet: a great diversity of news sources and the proliferation of disinformation. I study a model of information design that connects these observations. I show that competition between news sources with opposite biases reduces information quality when news consumers have limited attention. The reason is the endogenous formation of echo chambers. The standard narrative is that 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. I show that the presence of many news sources and the widespread existence of misleading news are concurrent.

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 and Konrad Stahl 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 and the Bavarian Young Economist Meeting 2021. I gratefully acknowledge funding by the German Research Foundation (DFG) through CRC TR 224 (Project B05).

†University of Mannheim, Department of Economics and MaCCI.

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 echo chambers 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). 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 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". I provide an alternative explanation: 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]

[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).

I study a Bayesian persuasion model with two states of the world and two actions. 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. 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. The decision-maker chooses which expert is worthy of attention and observes the information such an expert provides. Then, she updates her belief and takes an action. I show that competition between experts 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 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) makes decision-makers worse off. Two experts with different preferred actions compete to persuade decision-makers. 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 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, 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 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 competitive 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 - can 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 solve the problem of harmful competition, and media pluralism can enhance information quality.

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 competitive setting, 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. The pro-government media cannot persuade these citizens to get vaccinated. The existence of a large no-vax echo chamber can help explaining 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

[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.

> “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)

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) and the superstar effect (Hefti and Lareida, 2021).[4] In this paper, I offer new insights into the effects of limited attention. I show that limited attention can explain why rational decision-makers 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. A similar effect arises in Gitmez and Molavi (2020). However, these

[4] Gabaix (2019) and Mackowiak et al. (2020) survey the literature on behavioural and rational inattention, respectively.

[5] Alonso and Camara (2016) consider a sender who persuades a receiver, and the two have heterogeneous beliefs. 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.

authors focus on the ability of a sender to gather attention from receivers with heterogeneous beliefs.

Gentzkow and Kamenica (2017a,b) argue that competition among senders weakly increases information provision and benefits receivers. I show that this conclusion fails if receivers have heterogeneous beliefs and limited attention. My model incorporates endogenous allocation of attention between competing senders and endogenous persuasion.[6] 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 change (Williams et al., 2015), vaccinations (Cossard et al., 2020) and the financial markets (Cookson et al., 2021). 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.[7] 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.[8] Jann and Schottmuller (2019) and Martinez and Tenev (2020) argue that echo chambers can be helpful, either to enhance communication

[6]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), Gitmez and Molavi (2020), Lipnowski et al. (2020) and Wei (2020) study how limited attention by the receiver(s) affects optimal persuasion by a single sender.

[7]See also Giovanniello (2021) where echo chambers arise because biased voters have incentives to communicate useful information only to like-minded peers.

[8]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.

in a network or to separate high-quality and low-quality news. Instead, echo chambers have a negative effect here. The reason is the endogenous supply of information by biased experts.

Detrimental competition. A broad literature shows that competition can backfire in many different settings. Chen and Riordan (2008) show that price-increasing competition occurs when products are sufficiently differentiated. Easier entry in a setting with procrastinating consumers and switching costs may lead to higher prices (Heidhues et al., 2021). In the insurance market, competition can increase distortions when agents have heterogeneous perceptions about risk (Spinnewijn, 2013). The "unravelling" effect of competition has been disputed: with vertically differentiated firms, only high-quality firms have incentives to disclose (Board, 2009), or there is no disclosure at all (Janssen and Roy, 2014). 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). I uncover a novel channel for detrimental competition: media bias when decision-makers have limited attention and heterogeneous beliefs.

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.[9] 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:

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

[9] In Section B.5, I discuss an extension with more than two states.

Before taking an action, each decision-maker $d \in D$ pays attention to one expert $j_d \in J$: 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 $\omega \in \Omega$.[10] 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:

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

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 pays attention to one expert j_d.[11]

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.

I solve the game by backward induction, and the equilibrium notion is Perfect Bayesian Equilibrium. I assume without loss of generality that the preferred action of expert j_d is a_1. 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:

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

[10] I assume that the message space S_j contains at least two elements for any expert $j \in J$.

[11] In Section B.2, I consider a sequential version of the game. The effect of competition is different when experts implicitly become attention-seekers, as in Knoepfle (2020).

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

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

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:*

$$\pi_j(s \,|\, \omega_2) \leq \frac{\mu_i^0(\omega_1)}{\mu_i^0(\omega_2)}\pi_j(s \,|\, \omega_1) := \phi_i \pi_j(s \,|\, \omega_1) \tag{2}$$

I denote with $H_j := \{d \in D \,|\, j_d = j\}$ 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,

$$g_{ij} := \begin{cases} 0 & \text{if } H_j = \varnothing \\ \frac{|\{d \in H_j \,|\, d \in D_i\}|}{|H_j|} & \text{otherwise} \end{cases} \tag{3}$$

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

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

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

$$\max_{\pi_j} \sum_{i \in I} \sum_{s \in S_j} \sum_{\omega \in \Omega} \pi_j(s \,|\, \omega)\mu_j^0(\omega) v_{ij}(\pi_j, s) \tag{4}$$

The expert takes his audience H_j as given. Therefore, (4) 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 $\pi_j(s\,|\,\omega_2) = \phi_i \pi_j(s\,|\,\omega_1)$ or 2.) $\pi_j(s\,|\,\omega_1) = \pi_j(s\,|\,\omega_2)$ for any $s \in S_j$.*

By Lemma 1, I can restrict the set of policies that can be best responses: if the expert's audience includes sceptics, 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 (4) - 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 (2), 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. $\phi_2 < 1$.[12] 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

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

constraint for sceptics holds with equality, or both persuasion constraints do so. Hence, I identify two candidates for the optimal 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*

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

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

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

$$\pi_s(s \mid \omega_1) = k, \quad \pi_s(s' \mid \omega_1) = 1 - k$$

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

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

The hard-news policy implies the following posterior beliefs:

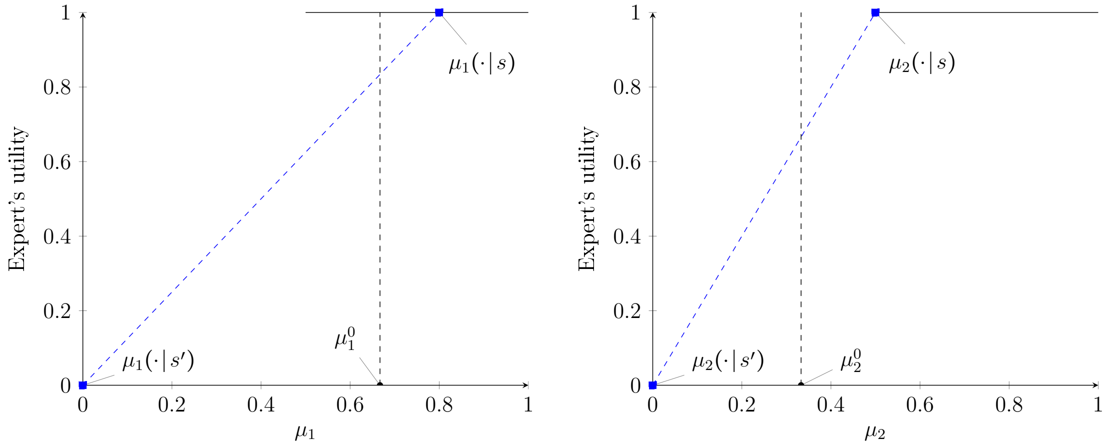
$$\mu_1(\omega_1 \mid s) = \frac{\phi_1}{\phi_1 + \phi_2} > \mu_2(\omega_1 \mid s) = \frac{1}{2}, \quad \mu_1(\omega_1 \mid s') = \mu_2(\omega_1 \mid s') = 0 \tag{5}$$

whereas the soft-news policy implies the following posterior beliefs:

$$\mu_1(\omega_1 \mid s) = \frac{\phi_1}{\phi_1 + \phi_2} > \mu_2(\omega_1 \mid s) = \frac{1}{2}, \quad \mu_1(\omega_1 \mid s') = \frac{1}{2} > \mu_2(\omega_1 \mid s') = \frac{\phi_2}{\phi_1 + \phi_2} \tag{6}$$

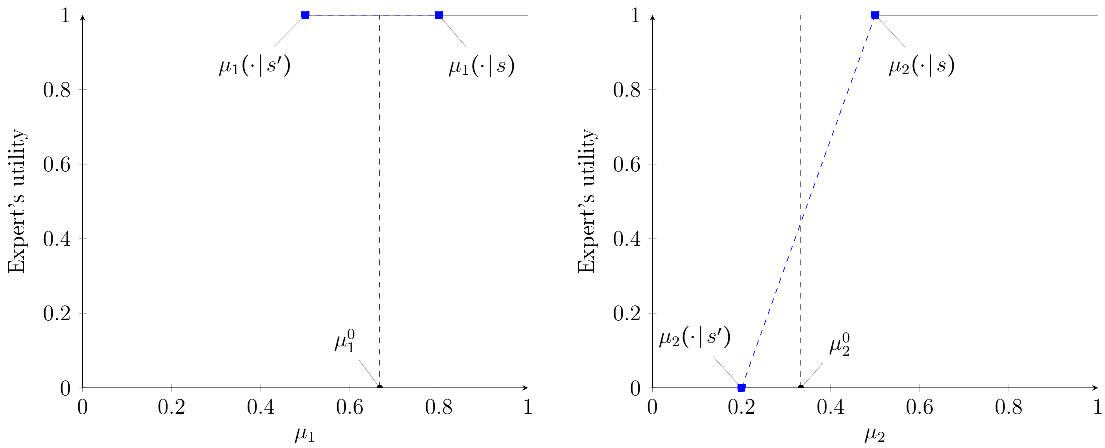
The hard-news policy persuades all decision-makers after seeing s and nobody after seeing s'. Thus, decision-makers choose the expert's preferred action in the state ω_1, and sometimes in the state ω_2. The expert provides sufficiently accurate information able to influence sceptics. However, this comes 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.

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.



(a) Believers (b) Sceptics

Figure 1: Posterior beliefs (blue squares) with the hard-news policy.



(a) Believers (b) Sceptics

Figure 2: Posterior beliefs (blue squares) with the soft-news policy.

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*

$$\mu^0(\omega_1) \geq \frac{\phi_1 g_1 - \phi_2}{1 - \phi_2 + (\phi_1 - 1) g_1} \tag{7}$$

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 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.[13]

[13] Formally, babbling is any policy π such that $\pi(s|\omega_1) = \pi(s|\omega_2)$ for any $s \in S$.

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.

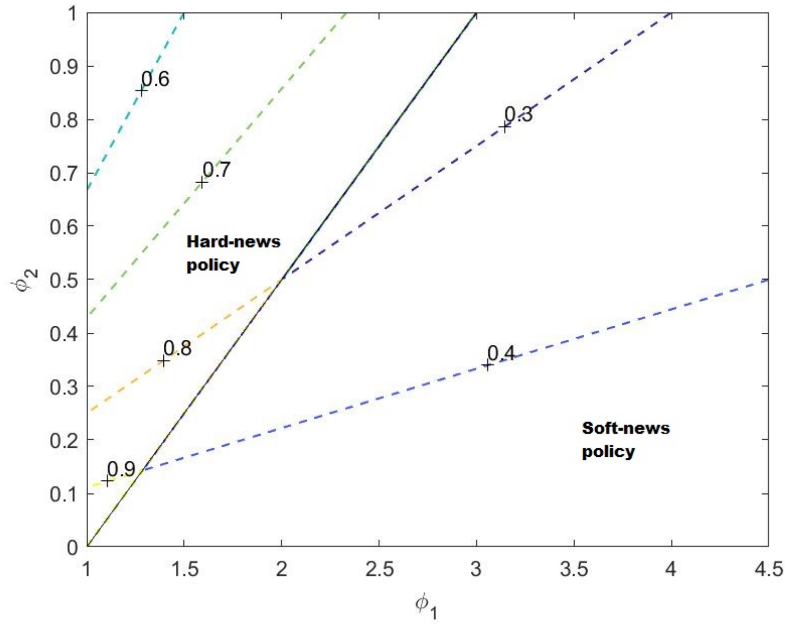
As Figures 3a and 3b show, the effect of polarization on the informativeness of the monopolist's policy is non-monotonous. Polarization increases informativeness (i.e., the range of posterior beliefs). However, there is a discontinuity point, that is, when the expert shifts from the hard-news policy 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. The hard-news policy is then defined as follows:

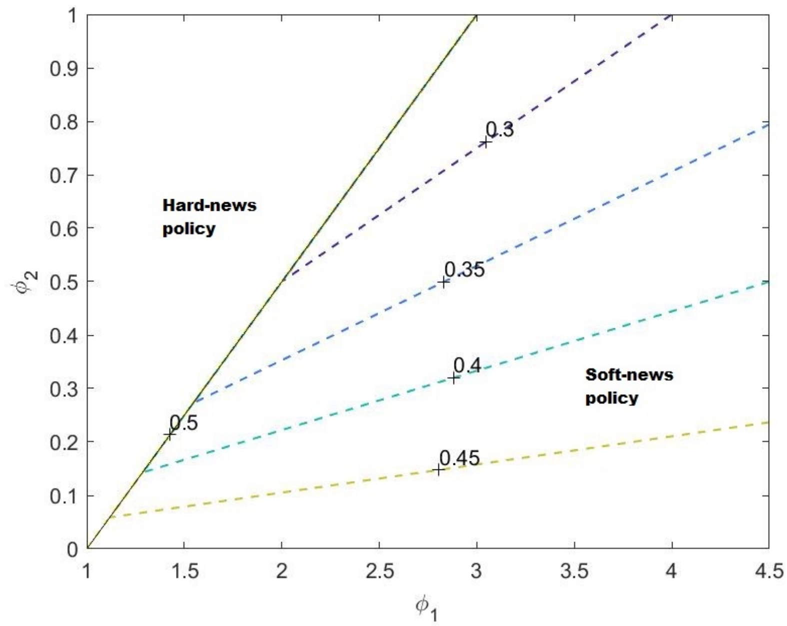
ω	Safe	Side Effects	
s	*safe*	*safe*	*side effects*
$\pi(s \mid \omega)$	1	0.25	0.75

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".

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



(a) Believers



(b) Sceptics

Figure 3: Range of posterior beliefs when $\mu^0(\omega_1) = \frac{1}{2}$ and $g_1 = \frac{1}{2}$.

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.

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

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 (7) 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.[14]

The natural question to ask is then: What happens if we allow competition by a no-vax media? The next section provides an answer.

5. Media Pluralism

In this section, I study how competition affects persuasion. I restrict attention to competition between two experts with different preferred actions.[15] The following lemma establishes the effect of competition with unlimited attention.

Lemma 3 (Competition). *Let $J = \{\alpha, \beta\}$ with $a_\alpha = a_1$ and $a_\beta = a_2$. For any $s_\alpha \in S_\alpha$ and any $s_\beta \in S_\beta$ such that $\pi_\alpha(s_\alpha \mid \omega_1), \pi_\beta(s_\beta \mid \omega_2) > 0$, it must hold $\pi_\alpha(s_\alpha \mid \omega_2) = \pi_\beta(s_\beta \mid \omega_1) = 0$.*

By Lemma 1, at least one persuasion constraint must hold with equality in any best response of any expert. The corresponding decision-makers are thus indifferent between either preferred action. Then, the rival has

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

[15] Information provision is not affected by the entry of experts with the same preferences as the incumbent. Indeed, the entrant cannot refine the optimal persuasion of the incumbent. See Section B.3 in the Appendix.

incentives to undercut the expert: it is sufficient to provide very little information to change such decision-makers' behaviour. Therefore, there cannot be an equilibrium unless any expert refrains from persuading under each other's favourable state. Full revelation (i.e., truth-telling by both experts) is the equilibrium when decision-makers have unlimited attention: given truth-telling by the rival, any attempt to persuade is futile.[16]

In the following, I introduce limited attention and show that full revelation is not an equilibrium. Competition 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.[17] 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:

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

Lemma 4 (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 $m = \arg\max_{m \in \{1,2\}} \mu_i^0(\omega_m)$. 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. 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. 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

[16]This result is coherent with the literature (Gentzkow and Kamenica, 2017a,b; Ravindran and Cui, 2020).

[17]In Section B.1, decision-makers can pay attention to the second expert at a cost. I show that full revelation is achievable if and only if this cost is zero. Such a case is equivalent to attention being unlimited. Instead, my results still hold under the weaker assumption that attention is costly rather than limited.

the corresponding subjective probability of taking the correct action. Therefore, $\Delta_{ij} := \lambda_i(\pi_j) - \mu_i^0(\omega_m) \geq 0$ is the subjective information gain from persuasion.

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.

Proposition 2 (Zero information gain for a target)**.** *For each expert $j \in J$ and each $i \in T_j$, it holds that $\Delta_{ij} = 0$.*

Proposition 2 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.[18]

Proposition 2 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".*

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

Here, I assume $I = \{1, 2\}$ with $\phi_1 > 1$ and $\phi_2 < 1$.[19] 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;

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.[20]

Proposition 3 (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, 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 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.

[19]In Section 6.2 I consider the case of more than two subgroups of decision-makers.

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

The game has also asymmetric equilibria (see Figure 4). 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 Proposition 2) 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.

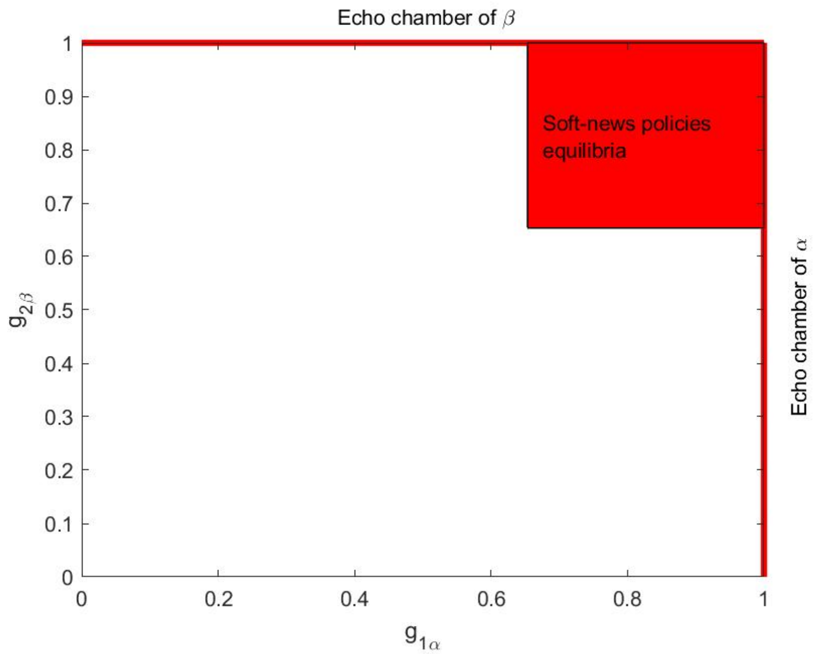


Figure 4: Allocations of attention that can support an equilibrium, when $\mu^0_\alpha(\omega_1) = \mu^0_\beta(\omega_2) = \frac{7}{10}$, $\phi_1 = 2$ and $\phi_2 = \frac{1}{2}$.

Proposition 4 (Harmful competition). *For any competitive equilibrium, there exists a monopoly outcome such that information gain and information quality are (weakly) higher for any decision-maker.*

Proposition 4 implies that decision-makers are worse informed with competition. Media pluralism harms decision-makers when the latter have limited attention and can freely allocate it between experts. Each decision-maker attempts to get positive information gain from persuasion by avoiding to devote attention to an expert who targets her. However, this leads decision-makers to cluster into echo chambers. Echo chambers are harmful because each expert faces only his believers, and the best response is babbling. Thus, decision-makers are better informed in a monopoly. Indeed, a monopolist uses either his hard-news policy or his soft-news policy: these policies produce some dispersion in posterior beliefs, whereas babbling leaves beliefs unchanged. Hence, babbling is less informative according to Blackwell (1953)'s order. The monopoly outcome also outperforms the asymmetric equilibria where each expert uses his soft-news policy. This 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 competition is related 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 aggregate informativeness (i.e., the average quality of news that decision-makers receive). In other words, the platform chooses g_{ij} for any subgroup $i \in I$ and any expert $j \in J$. Then, each expert j solves (4). 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 $\phi_2 < 1$. 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 (11) in the Appendix), each expert uses his hard-news policy if there are not too many believers in his audience:

$$g_{1\alpha} \leq \hat{g}_\alpha := \frac{\mu^0_\alpha(\omega_1) + \phi_2 \mu^0_\alpha(\omega_2)}{\mu^0_\alpha(\omega_1) + \phi_1 \mu^0_\alpha(\omega_2)} \tag{8}$$

$$g_{2\beta} \leq \hat{g}_\beta := \frac{\mu^0_\beta(\omega_2) + \frac{1}{\phi_1} \mu^0_\beta(\omega_1)}{\mu^0_\beta(\omega_2) + \frac{1}{\phi_2} \mu^0_\beta(\omega_1)} \tag{9}$$

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: (8)-(9) 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 aggregate informativeness.

Proposition 5 (Platform). *A platform with the objective to maximize aggregate informativeness prefers media pluralism 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 aggregate informativeness by allocating some but not too many believers to like-minded experts. Alternatively, the platform can increase each decision-maker's informativeness with a monopolist using his 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.

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 6 (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*

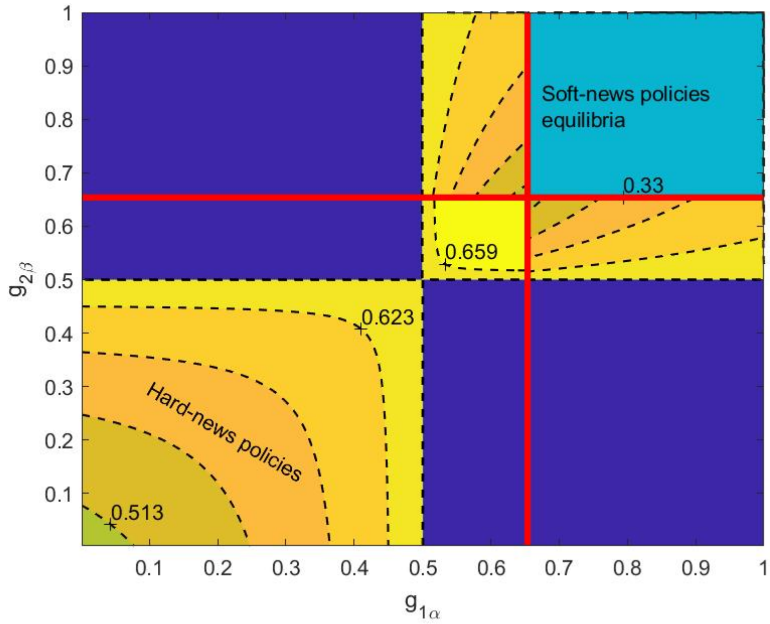


Figure 5: Aggregate informativeness (that is, the weighted sum of the ranges of posterior beliefs) when $\mu^0_\alpha(\omega_1) = \mu^0_\beta(\omega_2) = \frac{7}{10}$, $\phi_1 = 2$ and $\phi_2 = \frac{1}{2}$. In this example, assuming additionally that the two subgroups have equal size, the monopoly outcome is outperformed by a competitive setting where each expert uses his hard-news policy. The aggregate informativeness in monopoly amounts to $\frac{13}{20}$, whereas the platform can achieve aggregate informativeness equal to $\frac{87}{125}$. The platform allocates attention to expose as many believers as possible to hard-news policies, that is, $g_{1\alpha} = g_{2\beta} = 0.653$.

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 6 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 7 (Optimal persuasion)**.** *Let $F(x)$ be a distribution with support $[0,\infty)$ and density $f(x) > 0$ $\forall x$. Let $\phi_i := \frac{\mu^0_i(\omega_1)}{\mu^0_i(\omega_2)} \sim F$. Then, the expert j with ratio of prior beliefs ϕ_j uses a hard-news policy if a unique solution*

$\phi \in [0,1]$ *to the following equation exists*

$$h(\phi) = \frac{1}{\phi_j + \phi} \tag{10}$$

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 (10). A researcher needs to know the distribution of decision-makers' beliefs and the expert's belief.[21] Then, condition (10) predicts whether the expert uses a hard-news policy or a soft-news policy.

Gitmez and Molavi (2020) find a similar characterization of the optimal policy in a setting where the expert is trading-off between an extensive margin (how many decision-makers devote attention) and an intensive margin (how many decision-makers are persuaded). By contrast, in my setting devoting attention to one expert is costless, which means that all decision-makers devote attention.

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 10 implies $\phi = \frac{1}{\eta} - \phi_j$ and, by Proposition 7, the expert uses a hard-news policy if $\eta \geq \frac{1}{1+\phi_j}$. Fixing $\phi_j = 1$, Figure 6 depicts two examples of density functions that imply different optimal policies.

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.

Proposition 8 (Competition with limited attention)**.** *In any symmetric equilibrium, at least one expert is babbling.*

The key mechanism behind this result is the following: for any allocation of attention and corresponding optimal policies, there exists at least one target who can deviate and get a positive information gain, unless at least one expert is babbling.

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

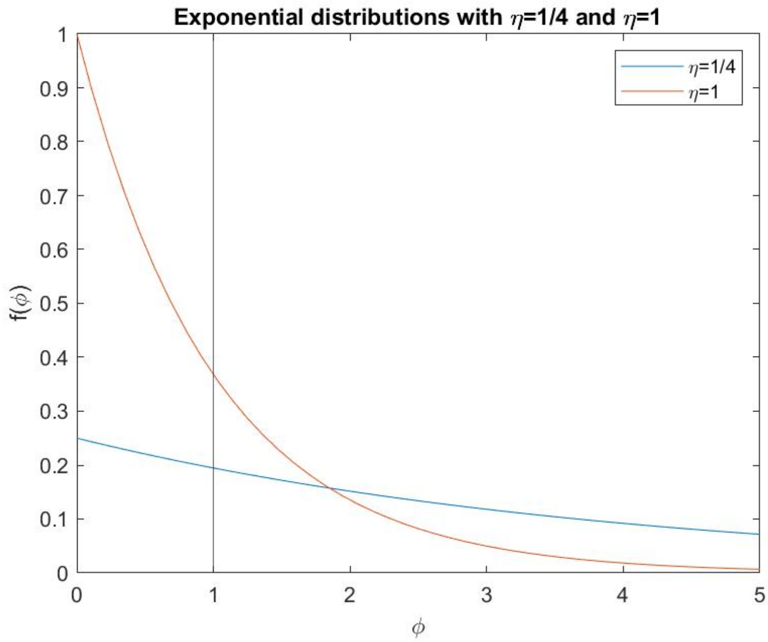


Figure 6: The black line at $\phi = 1$ separates sceptics (at the left) from believers. When $\eta = 1$, the majority of decision-makers are sceptics and, thus, a hard-news policy is optimal. By contrast, a soft-news policy is optimal when $\eta = \frac{1}{4}$, because many decision-makers are believers.

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 9 (Harmful competition)**.** *For any competitive equilibrium, there exists a monopoly outcome such that information gain and information quality are (weakly) higher for any decision-maker.*

The negative effect of competition (Proposition 4) 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 a monopolist uses a hard-news policy, competition is harmful because information gains are (weakly) lower, and those decision-makers who cluster into the echo chamber receive babbling. When an 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, competition 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.

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 four assumptions: on the one hand, experts are biased and have commitment power; on the other hand, decision-makers have heterogeneous beliefs and limited attention. Here, I briefly discuss what is the outcome if I relax any of these assumptions:

1. Under unlimited attention, by Lemma 3 experts are in direct competition to persuade decision-makers. As a consequence, full revelation is the unique equilibrium.

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.

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. The media may have commitment power, for instance, because of law or reputation concerns.[22] Limited attention is a well-established fact. Heterogeneous beliefs are also very likely to exist in all situations where the objective probability for a claim to be true is ambiguous. 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. In the following, I provide a non-exhaustive list of examples where my insights may be useful.

My model applies to the design of information about political issues. A politician wants to persuade voters to support a particular point of view. The optimal design of information trades off the desire of persuading sceptical voters and the goal of keeping loyalists. As a result, some information is provided. With competition and limited attention, some voters cluster into to echo chamber(s) and get no useful information.

A recent example is Trump's claim that the US Presidential election was fraudulent. The United States show increasing political polarization (Finkel et al., 2020). My model can explain 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).

Climate change is another relevant example. A vast majority of scientists claim that climate change is real. Many NGOs 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. Endogenous echo chambers can explain the existence of climate change deniers. Similarly, believers of a long list of debunked conspiracy theories can

[22] Nguyen (2017) and Fréchette et al. (2019) provide evidence in support of the Bayesian persuasion model.

survive within echo chambers. The common root is widespread scepticism about Science (Achenbach, 2015).

My model also applies to the advertising of differentiated products. A firm wants to persuade consumers to buy a product with uncertain value. Some consumers believe the product has a high value, whereas others believe it has low value. Each consumer buys if and only if she believes the product has high value. The firm designs the advertisement to maximize sales and then optimally provides some information about the product's value. With competition and limited attention, each consumer believes one product has a higher value than the other and may devote her attention only to the producer of this particular product. Echo chambers make it optimal for the firms to provide no information. My model can also rationalize asymmetric equilibria where one firm invest in informative advertising, whereas the other enjoys its market niche. If both firms design informative advertising, consumers rationally want to learn about their favoured products. But then providing informative advertising is not optimal for the firms. Cookson et al. (2021) provide evidence that investors behaviour in the financial markets is in line with this application.

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, competition 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 to provide valuable information vanish.

Whereas the literature has justified echo chambers with confirmation bias, I show that the opposite can be true. Even unbiased decision-makers end up devoting attention to like-minded experts. The latter, then, find it optimal to confirm decision-makers' beliefs. Therefore, I provide a rational foundation for confirmation bias.[23]

My findings provide a sobering insight into the effects of media pluralism: under media users' limited attention and heterogeneous beliefs, media pluralism leads to worse-informed media users. Information overload introduces an additional choice for decision-makers: the subset of information to process. Policymakers should account for decision-makers' incentives. Supporting media pluralism is a good idea only if decision-makers are sufficiently

[23]Goette et al. (2020) provide experimental evidence that limited attention reinforces confirmation bias.

attentive to process information from diverse sources.

My paper leaves an open question that requires further research. How can we mitigate the formation of echo chambers? One approach 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.

References

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," .

AUMANN, R. J. AND M. MASCHLER (1995): *Repeated games with incomplete information*, MIT press.

BERGEMANN, D. AND S. MORRIS (2019): "Information design: A unified perspective," *Journal of Economic Literature*, 57, 44–95.

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," Tech. rep., Working paper, Stanford University.

BOARD, O. (2009): "Competition And Disclosure," *The Journal of Industrial Economics*, 57, 197–213.

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.

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.

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

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.

Finkel, E. J., C. A. Bail, M. Cikara, P. H. Ditto, S. Iyengar, S. Klar, L. Mason, M. C. McGrath, B. Nyhan, D. G. Rand, L. J. Skitka, J. A. Tucker, J. J. Van Bavel, C. S. Wang, and J. N. Druckman (2020): "Political sectarianism in America," *Science*, 370, 533–536.

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 (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. and P. Molavi (2020): "Media Capture: A Bayesian Persuasion Approach," Tech. rep.

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," .

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," Tech. rep., CRC Discussion papers series.

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 (2020): "Rational Inattention: A Review," *CEPR Discussion Papers*.

MARTINEZ, G. AND N. H. TENEV (2020): "Optimal Echo Chambers," Tech. rep., arXiv. org.

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.

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.

NGUYEN, Q. (2017): "Bayesian persuasion: evidence from the laboratory," *Work. Pap., Utah State Univ., Logan.*

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*.

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

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.

Wei, D. (2020): "Persuasion under costly learning," *Journal of Mathematical Economics*.

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 A

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. □

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:

ω	ω_1	ω_2	
s	s	s	s'
$\pi_h(s\|\omega)$	1	ϕ_2	

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

The *Soft-news policy* is as follows:

ω	ω_1		ω_2	
s	s	s'	s	s'
$\pi_s(s\|\omega)$	k	$1-k$	$\phi_2 k$	$\phi_1(1-k)$

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

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:

$$\begin{aligned} V_h \geq V_s &\iff \mu^0(\omega_1) + \mu^0(\omega_2)\phi_2 \geq \left(\mu^0(\omega_1) + \mu^0(\omega_2)\phi_1\right) g_1 \\ &\iff \mu^0(\omega_1)(1-g_1) \geq \mu^0(\omega_2)\left(\phi_1 g_1 - \phi_2\right) \qquad (11) \end{aligned}$$

Note that the RHS of (11) 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 (11) is increasing (decreasing) in g_1, the share of believers among decision-makers. Finally, the RHS (LHS) of (11) is decreasing (increasing) in $\mu^0(\omega_1)$, the expert's belief of his favourable state. □

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 (5)-(6) 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. With unlimited attention, each decision-maker observes the policies of all experts and the corresponding messages. In particular, she observes $\pi(s\,|\,\omega) = \Pi_{j\in J}\pi_j(s_j\,|\,\omega)$ for any $s \in S_J := \times_{j\in J} S_j$ and any $\omega \in \Omega$, *and* a realization $s \in S_J$ chosen by Nature. I assume without loss of generality that decision-makers break the ties in favour of expert α. Thus, the *persuasion constraints* become:

$$\pi_\alpha(s_\alpha|\omega_2) \leq \frac{\mu_i^0(\omega_1)}{\mu_i^0(\omega_2)}\frac{\pi_\beta(s_\beta|\omega_1)}{\pi_\beta(s_\beta|\omega_2)}\pi_\alpha(s_\alpha|\omega_1) := \phi_i(s_\beta)\pi_\alpha(s_\alpha|\omega_1) \tag{12}$$

$$\pi_\beta(s_\beta|\omega_1) < \frac{\mu_i^0(\omega_2)}{\mu_i^0(\omega_1)}\frac{\pi_\alpha(s_\alpha|\omega_2)}{\pi_\alpha(s_\alpha|\omega_1)}\pi_\beta(s_\beta|\omega_2) := \phi_i(s_\alpha)\pi_\beta(s_\beta|\omega_2) \tag{13}$$

Let $I'_\alpha := I \times S_\beta$ and $I'_\beta := I \times S_\alpha$ be the hypothetical subgroups of decision-makers that experts α and β face, respectively. Indeed, in a competitive setting the ratio of priors each expert j faces depends on the prior of a subgroup $i \in I$ *and* on the particular message $s \in S_{-j}$ decision-makers receive. In other words, for any $i' \in I'_j$ such that $i' = (i, s)$, it holds $\phi_{i'} = \phi_i(s)$. Considering I'_j as the set of subgroups of decision-makers for the expert j, Lemma 1 extends to a competitive setting.

Now, I assume by contradiction that $\pi_\beta(s_\beta|\omega_2) > 0$ and $\pi_\beta(s_\beta|\omega_1) > 0$ for some $s_\beta \in S_\beta$. Thus, it holds $\phi_i(s_\beta) > 0$ for any $i \in I$, and by Lemma 1, $\pi_\alpha(s_\alpha|\omega_2) = \phi_{i'}(s_\beta)\pi_\alpha(s_\alpha|\omega_1)$ for some $i' \in I$ and some $s_\alpha \in S_\alpha$. It follows that $\phi_i(s_\alpha) > 0$ for any $i \in I$, and $\phi_{i'}(s_\alpha) = \frac{\pi_\beta(s_\beta|\omega_1)}{\pi_\beta(s_\beta|\omega_2)}$. To persuade i', s_β has to satisfy the following persuasion constraint:

$$\pi_\beta(s_\beta|\omega_1) < \phi_{i'}(s_\alpha)\pi_\beta(s_\beta|\omega_2)$$

which requires simply to decrease $\pi_\beta(s_\beta|\omega_1)$ by an amount $\epsilon > 0$ and small. This is a beneficial deviation because the expert β persuades an additional subgroup of decision-makers (i') with a negligible reduction in the probability of persuasion. By (12), it follows that the persuasion constraint for expert α becomes:

$$\pi_\alpha(s_\alpha|\omega_2) \leq \frac{\mu_{i'}^0(\omega_1)}{\mu_{i'}^0(\omega_2)}\frac{\pi_\beta(s_\beta|\omega_1)}{\pi_\beta(s_\beta|\omega_2)}\pi_\alpha(s_\alpha|\omega_1) < \phi_{i'}(s_\beta)\pi_\alpha(s_\alpha|\omega_1)$$

that is a contradiction, which follows from the fact that this is a zero-sum game for the experts. □

Proof of Lemma 4

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 (2), $\sigma\big(\mu_i(\omega_1 \,|\, s)\big) = a_k$ 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 Proposition 2

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 3

Proof. Echo chambers: Given $H_\alpha = D_1$ and $H_\beta = D_2$, babbling is optimal for each expert. Therefore, by Lemma 4, $\Delta_{ij} = 0$ 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 Proposition 2, 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 Proposition 2, $\Delta_{1\beta} = \Delta_{2\alpha} = 0$. 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 4

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 (5)-(6),

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

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

Proof of Proposition 5

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 $j \in J$:

$$\psi_{i\alpha} = \begin{cases} \frac{\phi_i}{\phi_i + \phi_2} & \text{if (8) holds} \\ \frac{\phi_1 - \phi_2}{2(\phi_1 + \phi_2)} & \text{otherwise} \end{cases} \qquad \psi_{i\beta} = \begin{cases} \frac{\phi_1}{\phi_1 + \phi_i} & \text{if (9) holds} \\ \frac{\phi_1 - \phi_2}{2(\phi_1 + \phi_2)} & \text{otherwise} \end{cases}$$

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

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

If the expert j is the monopolist, then aggregate informativeness 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 competitive setting more informative than a monopoly. There are two cases to consider:

1. If each expert as monopolist uses his soft-news policy - that is, (8)-(9) do not hold given g - then a competitive setting 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 (8)-(9) 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 competitive setting must be better if $\hat{g}_\alpha, \hat{g}_\beta > \frac{1}{2}$. Note that, by assumption, $g < \hat{g}_\alpha$ and aggregate informativeness in monopoly is $\Psi_\alpha^M = g\left(\frac{\phi_1}{\phi_1+\phi_2}\right) + \frac{1-g}{2}$. Consider a fraction $\epsilon \in (0, 1-g)$ of sceptics of α (believers of β) and set $g'_\beta = \epsilon$. In a competitive setting, 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 $g_{1\alpha} = \frac{g-\epsilon}{1-\epsilon-\epsilon'} \leq \hat{g}_\alpha$. Therefore, aggregate informativeness in a competitive setting is:

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

and the change in aggregate informativeness 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 a particular competitive setting 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, $g'_\alpha = 1 - g - g'_\beta$. The aggregate informativeness in this competitive setting is:

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

and the change in aggregate informativeness 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 competitive setting.

□

Proof of Proposition 6

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 7

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

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

As suggested by Proposition 6, 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:

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

and implies condition (10), 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 $x = 0$ or $x = 1$. Following Proposition 6, $x = 0$ implies the truth-telling policy, which is a special case of a hard-news policy in this setting. Instead, $x = 1$ 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 $f'(\phi) \geq 0$ for any $\phi \in [0,\infty)$. □

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 8

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 gathers attention from some sceptics and some believers. By Proposition 6 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:

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

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

Consider a soft-news policy. It follows:

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

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 Proposition 2, 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 9

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 $\hat{r} \leq r$, by Proposition 6, 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 6, 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 (3), 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 competition 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 8). 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. Appendix B

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 10. *Truth-telling 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. By Lemma 3, truth-telling 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 8 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. 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 an equilibrium. 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 Proposition 2 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.[24]

B.3. Competition with 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

[24]There exist empirical evidence that biased experts, for example politicians, respond strategically to attention habits. See for instance Eisensee and Strömberg (2007).

is that the entrant cannot refine the optimal policy of the incumbent.[25]

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.[26] 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.4. 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 sub-group of decision-makers. Let π_j^i be the policy of expert $j \in J$ which targets subgroup $i \in I$. In a monopoly, π_j^i 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 competition and single-homing, $\lambda_i(\pi_j^i) = \mu_i^0(\omega_m)$ 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 4 and Proposition 2. Therefore, decision-makers are indifferent about the allocation of attention.

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.

[25]Experts with heterogeneous beliefs can have different optimal policies (in monopoly). However, differently from Lemma 3, there is no incentive to undercut the rival because the preferred actions coincide.

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

B.5. 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:

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

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

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

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:

$$\int_{\bar{\omega}}^1 \mu_2^0(\omega)d\omega = \int_0^{\bar{\omega}} \pi(s \,|\, \omega)\mu_2^0(\omega)d\omega \tag{19}$$

I denote with Π_H the subset of policies such that (19) 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 (4), 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:

$$\int_{\bar{\omega}}^1 \pi(s \,|\, \omega)\mu_2^0(\omega)d\omega = \int_0^{\bar{\omega}} \pi(s \,|\, \omega)\mu_2^0(\omega)d\omega \tag{20}$$

$$\int_{\bar{\omega}}^1 \pi(s' \,|\, \omega)\mu_1^0(\omega)d\omega = \int_0^{\bar{\omega}} \pi(s' \,|\, \omega)\mu_1^0(\omega)d\omega \tag{21}$$

I denote with Π_S the subset of policies such that (20)-(21) 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, Proposition 2 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 (20),

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

Hence, $\Delta_2 = 0$. Proposition 2 characterizes the incentives of decision-makers about the allocation of attention. Therefore, the effect of competition with limited attention is unchanged.

The analysis of optimal persuasion becomes generally intractable when the cardinality of Ω is equal to the cardinality of A.[27] 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 6 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.

Example - I consider the COVID-19 vaccination example, and I assume that there exists a third state of the world: safe but with caution (simply caution now on). Therefore $\Omega = \{\omega_1, \omega_2, \omega_3\} = \{$caution, safe, not safe$\}$. I

[27] A full characterization of prior beliefs requires $|\Omega|!$ decision-makers. Unlike Section 6.2, there is no intuitive ordering of decision-makers. Optimal persuasion cannot be studied generically without restrictive assumptions on the distribution of beliefs.

assume that the monopolistic expert (say a politician) is biased towards caution. For instance, the politician might want to vaccinate only the elderly.

There are two subgroups of decision-makers as before: believers and sceptics, respectively, about the vaccine being safe. I assume $\phi_1(\omega_1, \omega_3) > 1 > \phi_1(\omega_1, \omega_2)$ and $\phi_2(\omega_1, \omega_2) > 1 > \phi_2(\omega_1, \omega_3)$. A soft-news policy is not useful because there are not true believers. Let π_h be a hard-news policy:

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

$$\pi_h(s|\omega_2) = \phi_1(\omega_1, \omega_2) \quad \pi_h(s'|\omega_2) = 1 - \phi_1(\omega_1, \omega_2)$$
$$\pi_h(s|\omega_3) = \phi_2(\omega_1, \omega_3) \quad \pi_h(s'|\omega_3) = 1 - \phi_2(\omega_1, \omega_3)$$

Let us consider as alternative π_s:

$$\pi_s(s|\omega_1) = k \quad \pi_s(s'|\omega_1) = 1 - k$$

$$\pi_s(s|\omega_2) = \phi_1(\omega_1, \omega_2)k \quad \pi_s(s'|\omega_2) \leq \phi_2(\omega_1, \omega_2)(1 - k)$$
$$\pi_s(s|\omega_3) = \phi_2(\omega_1, \omega_3)(1 - k) \quad \pi_s(s'|\omega_3) \leq \phi_1(\omega_1, \omega_3)k$$

The favourable state of the politician is caution, that is a compromise between opposite decision-makers' beliefs. If decision-makers have sufficiently polarized beliefs (and the politician is sufficiently uncertain about the true state), then it is optimal to use π_s. The intuition is similar to Proposition 1. With π_s, the politician randomizes between messages that either support one extreme state or the other. In other words, to persuade citizens that the best option is to take caution, a politician alternates positive and negative news about vaccinations. These news are not designed to move one group from one extreme to the other, but just from one extreme to a compromise. The alternative is to provide "hard evidence" that vaccinations are safe given precautions. This is extremely costly with high polarization, as both extreme views have to be contrasted at the same time. Note that π_s is not a soft-news policy, but it works similarly: the goal is to leverage believers' credulity.

The intractability of optimal persuasion does not allow to study the whole game. However, intuitively my results should not be affected by the existence of many states of the world and corresponding actions. For instance, let us consider Proposition 3. 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. Decision-makers are better informed with a monopoly, because the existence of heterogeneous beliefs makes optimal for the expert to use some informative policy, where informativeness is defined following Blackwell (1953).

B.6. 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:

$$\mu(\omega_1 \mid s) \geq \frac{b_i^{\omega_2}}{b_i^{\omega_1} + b_i^{\omega_2}} \iff \pi_j(s \mid \omega_2) \leq \frac{\mu^0(\omega_1)}{\mu^0(\omega_2)} \frac{b_i^{\omega_1}}{b_i^{\omega_2}} \pi_j(s \mid \omega_1) \tag{22}$$

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 (2) and (22). 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 $c_i = \frac{b_i^{\omega_2}}{b_i^{\omega_1}+b_i^{\omega_2}}$.