A Rational Inattention Theory of Echo Chamber

Lin Hu* Anqi Li† Xu Tan‡

July 2023

Abstract

We develop a rational inattention theory of echo chamber, whereby players gather information about an uncertain state by allocating limited attention capacities across biased primary sources and the other players. The resulting Poisson attention network transmits information from the primary source to a player either directly or indirectly through the other players. Rational inattention generates heterogeneous demands for information among players who are initially biased towards different decisions. In an echo chamber equilibrium, each player restricts attention to his own-biased source and like-minded friends, as the latter attend to the same primary source as his, and so could serve as secondary sources in case the information transmission from the primary source to him is disrupted. We provide sufficient conditions that give rise to echo chamber equilibria, characterize the attention networks within echo chambers, and use our results to inform the design and regulation of information platforms.

Keywords: rational inattention, echo chamber, information platform design and regulation

JEL codes: D83, D85

*Research School of Finance, Actuarial Studies and Statistics, Australian National University, lin.hu@anu.edu.au.

†Department of Economics, University of Waterloo, angellianqi@gmail.com.

‡Department of Economics, University of Washington, tauxu@uw.edu.

1 Introduction

The Cambridge English Dictionary defines echo chambers as "environments in which people encounter only beliefs or opinions that coincide with their own, so that their existing views are reinforced and alternative ideas are not considered." Examples that fit this description have recently flourished on the Internet and social media, and their economic and social consequences have been the subject of heated debate in the academia and popular press (Bakshy, Messing, and Adamic, 2015; Del Vicario, Bessi, Zollo, Petroni, Scala, Caldarelli, Stanley, and Quattrociocchi, 2016; Barberá, 2020; Cossard, Morales, Kalimeri, Mejova, Paolotti, and Starnini, 2020). Most ongoing discussions of echo chambers focus on their behavioral roots (Levy and Razin, 2019). This paper develops a rational theory of echo chamber with clear testable predictions and relevant normative implications.

Our premise is rational inattention (RI), i.e., the rational and flexible allocation of limited attention capacities across information sources. Such a premise has become increasingly relevant in today's digital age, as people are inundated with information on the one hand, but can selectively choose which information sources to visit using personalization technologies on the other hand. Since Sunstein (2007) and Pariser (2011), it has been long suspected that RI may engender a selective exposure to content and a formation of homogeneous opinion clusters. The current paper formalizes this idea by prescribing conditions that are conducive to echo chamber formation among rationally inattentive decision makers. We also characterize the attention networks within echo chambers, and use our results to inform the design and regulation of information platforms.

To create a role for RI, we embed the analysis in a simple model of decision making under uncertainty. To illustrate our framework, consider the problem faced by new parents who are about to feed their babies with solid food. Each parent must choose between a traditional approach denoted by A (e.g., spoon feeding), and a new approach denoted by B (e.g, baby-led weaning). Which of the two approaches is better for baby development is modeled as a random state that is either A or B with equal probability. A parent earns the highest level of utility if he chooses the best approach for baby development. Otherwise he incurs a loss, whose magnitude depends on whether the adopted approach matches his own preference or not (e.g., baby-led weaning may be more preferred because it is easy to prepare). Given the prior belief about the state, the *default decision* is to adopt the parent's preferred approach.

To gather information about the state, a parent must pay attention to information sources. We distinguish *primary sources* from *secondary sources*. The former generate original data about the state, and they take the form of scientific experiments published in pediatric journals in our leading example. As in Che and Mierendorff (2019), we model two biased primary sources called A-revealing and B-revealing. The ω-revealing source, $\omega \in \{A, B\}$, is an experiment designed to reject the null hypothesis that the state is $\omega' \neq \omega$. It works by revealing that the state is ω when it is and keeping silent otherwise. Secondary sources are our innovation. In the leading example, they constitute parents who consume and pass along primary source content to the other parents via online support groups. A parent is endowed with a limited attention capacity, or *bandwidth*, that can be allocated across information sources in a flexible manner. A feasible attention strategy specifies a nonnegative amount of attention that he pays to each source, subject to the constraint that the total amount of paid attention must not exceed his bandwidth.

The subject of our study is the Poisson attention network generated by parents' attention strategies. After these strategies are specified, the state ω is realized, and messages thereof are circulated in the society for two rounds. In the first round, the ω-revealing primary source disseminates a message "ω" to parents. The message reaches each parent independently with a Poisson probability that increases with the amount of attention that the latter pays to the primary source. In the second round, those parents who received a message in the previous round pass it along to the other parents using Poisson technologies. The probability of a successful information transmission between a parent pair increases with the sender's *visibility* as a secondary source (i.e., the rate of his Poisson technology), as well as the amount of attention that the recipient pays to the sender. After that, parents update beliefs and make final decisions. We study the attention networks that can arise in equilibrium.

Our notion of echo chamber has two defining features. The first feature is a selective exposure to content and a formation of homogeneous clusters. Specifically, we define a parent's *own-biased source* as the primary source that favors his default decision as its null hypothesis, and call two parents *like-minded friends* if they share the same default decision. We say that echo chambers arise in an equilibrium if all parents restrict attention to their own-biased sources and like-minded friends on the equilibrium path. The second feature of echo chambers is a belief polarization coupled with an occasional and yet drastic belief reversal. It is easy to see that after playing an echo chamber equilibrium, each parent receives no message from any source most of the time and updates the belief in favor of his default decision in that event. With a small complementary probability, the opposite happens, and the parent feels strongly about departing from his default approach. As discussed in Section 5, both features of echo chambers have solid empirical supports.

Our main results exploit the trade-off between primary and secondary sources. The first result prescribes sufficient conditions for the rise of echo chamber equilibria. Since a parent can always make his default decision without paying attention, paying attention is only useful if it sometimes convinces him to act differently. When attention is limited, the parent should, intuitively, focus on his own-biased source, as the latter generates the exact kind of the information that disapproves of his default decision. Likewise, he should attend only to his like-minded friends but no one else, as the former share the same primary source with him, and so could serve as secondary sources in case the information transmission from the primary source to him is disrupted. A caveat to this argument is that in a strategic environment like ours, a parent may gain strategically from switching sides, e.g., when many other parents are gathering the opposite kind of information, he may want to follow suit so as to gain access to many secondary sources in case he misses the primary source message. Nevertheless, such a gain is shown to be limited when parents are sufficiently biased towards their default decisions, when they can attend to many people, and when attention is scarce. These conditions accurately describe the world we are living in, whereby technology advances have turned the entire globe into a village and inundated people with more information than what they can process in a lifetime. These trends are conducive to echo chamber formation, especially when people are sufficiently biased to begin with.

Our second result characterizes the attention networks within echo chambers. We define a parent's *informedness* as a secondary source as the amount of attention he pays to the primary source, while capturing his *influence* on public opinion by the amount of attention he attracts from his friends. We find that the game among like-minded friends exhibits strategic substitutability, as raising one's informedness attracts more attention of his friends away from the primary source to him. Based on this finding, together with other basic model properties, we develop a method for investigating the comparative statics of the equilibrium attention network. Among other things, we find that increasing a parent's bandwidth promotes his informedness and influence while diminishing that of any other parent. This equilibrium mechanism is shown to magnify even a small difference between parents' bandwidths into a very uneven distribution of opinions, whereby some parents act as opinion leaders, while others act as opinion followers. An important feature of today's news landscape is that while most Americans are interested in many topics such as science, economy, and politics, only a minority of them are serious news consumers, due to the distractions that stem from the overabundant entertainment opportunities (Funk, Gottfried, and Mitchell, 2017). This attentional gap between the majority and minority may generate patterns such as the law of the few and fat-tailed distributions of opinions,[1] whose presences in the social media sphere have recently been detected by Lu, Zhang, Cao, Hu, and Guo (2014) and Del Vicario, Bessi, Zollo, Petroni, Scala, Caldarelli, Stanley, and Quattrociocchi (2016).

We use our results to inform the design and regulation of information platforms. First and foremost, we find that interventions that target misinformation and fake news through modulating source visibility may backfire if they are not well calibrated according to the underlying environment.[2] We also evaluate the usefulness of anti-polarization platforms such as Allsides.com, which operate through exposing users to balanced viewpoints from both sides. We model such platforms as a mega source that obtains from merging the A-revealing source and B-revealing source of the baseline model together. On the one hand, the use of a mega source does fulfill the purpose of dissolving echo chambers, as it forces different types of parents to attend to each other as secondary sources. Yet making more secondary sources available to parents discourages information acquisition from primary sources. The resulting free-riding problem can render the overall welfare impact ambiguous.

Even in arguably more controversial environments, our model still captures some facets of reality and sheds cautioning light on several normative issues. In Section 5, we study a political economy application whereby partisan voters must choose

[1]The law of the few refers to the phenomenon that information is disseminated by a few key players to the rest of the society. It was originally discovered by Katz and Lazarsfeld (1955) in their classical study of how personal contacts facilitate the dissemination of political news, and has since then been rediscovered in numerous areas such as the organization of online communities.

[2]For example, Facebook caps the number of daily posts by an individual account to 25 articles, beyond which the visibility of the account will be negatively affected.

between a Democratic candidate and Republican candidate based on their qualities. Information about candidates' qualities is generated by biased media outlets, and is passed along from one voter to another through social media platforms. We use our results to match additional empirical patterns, and to speak to media market regulations such as the FCC's viewpoint diversity objectives.

In the online appendix, we examine the (in)efficiency of echo chambers, demonstrate the robustness of our model to more complex decisions problems and more general information technologies, and address several technical points that are omitted from the main text.

1.1 Related literature

The current paper contributes to three strands of the economic literature: rational inattention, social network, and rational theories of echo chamber.

Rational inattention. A central question studied by the RI literature concerns the flexible acquisition of information about a payoff relevant state by a single decision maker (see Maćkowiak, Matějka, and Wiederholt 2023 for a literature survey). For that purpose, it is useful to model information acquisition strategies as signal structures that map each fundamental state to a lottery over final decisions. This compact representation, however, abstracts away from how decision makers can learn from multiple information sources, especially when the latter are themselves the players of a strategic game. In games with strategic complementarities, Hellwig and Veldkamp (2009), Denti (forthcoming, 2017), and Hébert and La'O (2021), demonstrate the usefulness of acquiring information about the endogenous signals gleaned by the other players, as the latter affect one's payoff through the other players' final actions. This is not the case in our model, where a player's utility depends on his own action and a payoff-relevant state, but nothing else. In equilibrium, a player is attended by the other players because he has a better information dissemination technology than the primary sources. Evidence for this assumption is discussed in Sections 2.2 and 5.

The idea of *filtering bias*, namely even a rational decision maker can exhibit a preference for biased information when constrained by information processing capacities, dates back to Calvert (1985) and is later expanded on by Suen (2004), Che and Mierendorff (2019), and Hu, Li, and Segal (forthcoming) among others, in single-agent decision making problems. We examine the validity of this idea in a strategic setting, where one's choice of attention strategy depends not only on his own preference, but also on the attention strategies of the other players.

The closest work to ours: Che and Mierendorff (2019), studies a dynamic information acquisition problem where a decision maker repeatedly allocates a limited attention capacity between biased primary sources that disseminate Poisson signals.[3] We instead focus on the trade-off between allocating attention to primary sources and to the other players in a static game. Our model becomes a special case of the stage decision problem studied by Che and Mierendorff (2019) if players are forbidden from attending to each other.

We are not the first to study Poisson attention networks. Dessein, Galeotti, and Santos (2016) analyze the efficient attention network between nonstrategic members of an organization with adaptation and coordination motives. We focus on the equilibrium attention network between strategic players, although we also characterize the efficient attention network as an extension. Our players' objective functions also differ from that of Dessein, Galeotti, and Santos (2016).

Social network. Inside an echo chamber, our game combines (i) the strategic formation of an information sharing network, with (ii) a network game in which players' investments (in their informedness) exhibit negative externalities. Studies of noncooperative network formation games without endogenous investments were pioneered by Jackson and Wolinsky (1996) and Bala and Goyal (2000), and have recently been advanced by Calvó-Armengol, de Martí, and Prat (2015) and Herskovic and Ramos (2020) among others. The last two papers bestow players with exogenous signals and focus on the formation of information sharing networks, whereas signals are endogenous in our model. Existing network formation models work mainly with discrete links. Links are divisible in our model, as well as in Bloch and Dutta (2009), Baumann (2021), and Elliott, Golub, and Leduc (2022), among others.

There is also a large literature studying how to play games with negative externalities on a fixed network (see Jackson and Zenou 2015 for a survey). Most methodological contributions to this literature concern the uniqueness and stability of equilibrium, with many early contributions assuming linear best response functions or a symmetric influence matrix between players (see, e.g., Parise and Ozdaglar 2019 and Bramoullé, Kranton, and D'Amours 2014, respectively). We develop a toolkit

[3]Zhong (2022) demonstrates the optimality of Poisson attention strategy in dynamic information acquisition problems with continuous time and discrete terminal actions.

for investigating equilibrium comparative statics under different assumptions from the aforementioned ones.

A few recent papers study hybrid games that are akin to ours (see Sadler and Golub 2021 for a survey). The closest work to ours: Galeotti and Goyal (2010), also combines endogenous information acquisition with strategic information sharing. Yet these authors work with homogeneous players and quantitative information acquisition (i.e., how much information is acquired rather than what kind of information is acquired), so their model cannot be immediately applied to the study of echo chamber formation among heterogeneous players. Their main result: the law of the few, is obtained under a different setting and using different arguments from ours.[4]

Rational theories of echo chamber. A small but growing literature investigates the rational origins of echo chambers.[5] The closest work to ours: Baccara and Yariv (2013) (BY), studies a model of group formation, followed by the production and sharing of information among group members. The main differences between BY and the current work are threefold. First, BY models information as a local public good that is automatically shared among group members. Here, the decision to acquire secondhand information from other players is private and strategic. Second, BY codes players' preferences for information in their utility functions. Here, such a preference arises endogenously from limited attention capacities. Third, BY's reasoning exploits the sorting of people with similar preferences into groups of limited sizes. We impose no restriction on the sizes of echo chambers and do not invoke the sorting logic.

Other rational theories of echo chamber fall broadly into two categories: (i) strategic communication or persuasion between players with conflicting objectives (Galeotti, Ghiglino, and Squintani, 2013; Jann and Schottmüller, 2021; Innocenti, 2021; Meng, 2021); (ii) learning from sources with unknown biases (Sethi and Yildiz, 2016; Williams, 2021). Neither consideration is present in our model, where sources are nonstrategic, and their biases are commonly known.

[4]The argument of Galeotti and Goyal (2010) exploits two properties of their environment: (i) the total amount of acquired information in the society is independent of players' population size; and (ii) links for information sharing are discrete. Together, these properties imply that only a small number of players can acquire information from the original source and disseminate it to the other players in equilibrium. Our setup and reasoning are very different.

[5]Other authors have used the term "echo chamber" to refer to: the excessive sharing of misinformation in a homophilous, exogenous, network (Acemoglu, Ozdaglar, and Siderius, 2022), correlation neglect (Levy and Razin, 2019), as well as one's exogenous neighbors in network learning games (Bowen, Dmitriev, and Galperti, 2023).

2 Baseline model

In this section, we first describe the model setup and then present an illustrative example. Discussions of model assumptions can be found in Footnotes 6-9.

2.1 Setup

A finite set $\mathcal{I}$ of players faces two equally likely states A and B. Each player $i \in \mathcal{I}$ has a *type* $t_i \in \{A, B\}$ (also called his *default decision*), and makes a final decision $d_i \in \{A, B\}$. His utility equals zero if his decision matches the true state. If the two objects differ, then the player incurs a loss of magnitude $\beta_i \in (0, 1)$ by making the default decision. Otherwise the loss has magnitude 1. Formally,[6] [7]

$$u_i(d_i, \omega) = \begin{cases} 0 & \text{if } d_i = \omega, \\ -\beta_i & \text{if } d_i \neq \omega \text{ and } d_i = t_i, \\ -1 & \text{if } d_i \neq \omega \text{ and } d_i \neq t_i. \end{cases}$$

The assumption $\beta_i \in (0, 1)$ implies that the player most prefers his default decision given the prior belief about the state distribution. The preference becomes stronger as we decrease β_i, which is hereinafter referred to as the player's *horizontal preference parameter*. Let $\mathcal{A}$ and $\mathcal{B}$ denote the sets of type A players and type B players, respectively. Assume throughout that $|\mathcal{A}|, |\mathcal{B}| \in \mathbb{N} - \{1\}$.

There are two *primary sources* called A-revealing and B-revealing. In state $\omega \in \{A, B\}$, the ω-revealing primary source announces a message "ω," whereas the other primary source is silent. The message "ω" fully reveals that the state is ω, since any player would assign probability one to the state being ω given this message. To gather information about the state, a player can pay attention to the primary sources, i.e., spend valuable time on them. In addition, he can pay attention to the other players as potential *secondary sources*. For each player $i \in \mathcal{I}$, $\mathcal{C}_i = \{\text{A-revealing}, \text{B-revealing}\} \cup \mathcal{I} - \{i\}$ denotes the set of the sources he can pay attention to. His attention strategy

[6] The following reparameterization of the model captures situations in which players differ in prior beliefs about the state, while leaving the analysis unaffected: let $u_i(d_i, \omega) = 0$ if $d_i = \omega$ and -1 otherwise $\forall i \in \mathcal{I}$; let player i assign probability $\frac{1}{1+\beta_i}$ to $\omega = t_i$ and probability $\frac{\beta_i}{1+\beta_i}$ to $\omega \neq t_i$.

[7] We specify the utility function in a way that eases parameterization and interpretation. For general utility functions, define the default decision as one's most preferred decision ex ante, i.e., $t_i \coloneqq \arg\max_{d \in \{A,B\}} \mathbb{E}[u_i(d, \omega)]$, and all arguments below will go through.

$x_i = (x_i^c)_{c \in \mathcal{C}_i}$ specifies a nonnegative amount $x_i^c \geq 0$ of attention that he pays to each source $c \in \mathcal{C}_i$. It is feasible if it satisfies the *bandwidth constraint* $\sum_{i \in \mathcal{C}_i} x_i^c \leq \tau_i$, which stipulates that the total amount of paid attention must not exceed the player's *bandwidth* $\tau_i > 0$. Let $\mathcal{X}_i$ denote the set of feasible attention strategies for player i. To capture the flexibility of attention allocation, we allow for the adoption of any strategy in $\mathcal{X}_i$.

After players specify their attention strategies, the state $\omega \in \{A, B\}$ is realized, and information about ω is circulated in the society for two rounds.

- In the first round, the ω-revealing primary source disseminates "ω" to players using a Poisson technology with rate 1. The message reaches each player $i \in \mathcal{I}$ independently with probability $1 - \exp(-x_i^{\omega\text{-revealing}})$, which increases with the amount $x_i^{\omega\text{-revealing}}$ of attention that player i pays to the primary source and is strictly bounded above by one. The last property captures the scarcity of attention relative to the available information in the world.

- In the second round, each player i who received a message in the first round passes it along to the other players using a Poisson technology with rate $\lambda_i > 0$. The message reaches each player $j \in \mathcal{I} - \{i\}$ independently with probability $1 - \exp(-\lambda_i x_j^i)$, which increases with the amount x_j^i of attention that player j pays to player i. The parameter λ_i captures player i's visibility as a secondary source, and is hereinafter referred to as his *visibility parameter*.

After two rounds of information transmission, players update beliefs about the state and make final decisions. The game sequence is summarized as follows.

1. Players choose attention strategies.

2. The state ω is realized.

3. (a) The ω-revealing source disseminates message "ω" to players.

   (b) Those players who received messages in Stage 3(a) pass them along to the other players.

4. Players update beliefs and make final decisions.

The solution concept is *pure strategy perfect Bayesian equilibrium* (PSPBE), or *equilibrium* for short.

2.2 Illustrative example

In this section, we illustrate our framework using an example of science news consumption among new parents.

Example 1. A group of new parents is about to feed their babies with solid foods. There are two approaches: A = traditional spoon-feeding, and B = new baby-led weaning. Under the traditional approach, parents spoon-feed babies first with purée food and then with different stages of baby food until babies are strong enough to eat on their own. The new approach skips traditional baby food and puts babies in charge of their mealtime. Babies are given chucks of suitable food such as banana or bread, and they hold the food in their hands and feed themselves. Which of the two approaches is better for baby development is an open question. For example, a possible downside of baby-led weaning is that babies tend to eat less and choke more in the first few months. Whether this issue has any long-term health consequence and how it should be weighed against the upside of the new approach (e.g., practice motor skills earlier and build healthy eating habits), are topics of active research. The uncertainty surrounding the truth is captured by the random state ω.

A parent's preference has two dimensions. The vertical dimension concerns which of the two approaches is better for baby development. The horizontal dimension—which is parameterized by β_i—captures the parent's own preference: some parents prefer the traditional approach because spoon-feeding is less messy, while others prefer the new approach because it skips purée foods and so is easier to prepare. A parent earns the highest level of utility when the best approach for baby development is being used, and incurs a loss otherwise. The loss is smaller in case the adopted approach is more preferred by the parent.

Information about the uncertain state is provided by two kinds of scientific experiments: those designed to reject the traditional approach A, and those designed to reject the new approach B. These experiments constitute the primary sources in our model. Parents can directly read about the experiments published in scientific journals. Alternatively, they can learn from the secondhand opinions shared by other parents on online support groups—Academic Moms, BabyCenter, to name a few.

Parents may differ in the amount of available time τ_is for information gathering, which depends on the nature of their work, the length of parental leave, etc. They may also differ in their visibility λ_is as secondary sources. For example, parents who are well-educated and good at explaining science to layman attract many followers on online platforms; the decision on whether to post a video on Youtube depends on how enthusiastic a parent is about helping others and how tech savvy he or she is. ♢

3 Analysis

This section conducts model analysis. We first formalize the problems faced by players in Section 3.1, and introduce key concepts to the analysis in Section 3.2. These are followed by three main theorems regarding the formation of echo chambers, the attention networks within echo chambers, and the comparative statics thereof, in Sections 3.3, 3.4, and 3.5, respectively.

3.1 Player's problem

Consider the problem faced by a typical player $i \in \mathcal{I}$, taking the strategy profile $x_{-i} \in \mathcal{X}_{-i} := \times_{j \in \mathcal{I}-\{i\}} \mathcal{X}_j$ among the other players as given. In case player i uses an attention strategy $x_i \in \mathcal{X}_i$, the information transmission from the ω-revealing source to him is disrupted with probability

$$\delta_i^{\omega\text{-}revealing} := \exp(-x_i^{\omega\text{-}revealing}),$$

and the information transmission from any player $j \in \mathcal{I} - \{i\}$ to him is disrupted with probability

$$\delta_i^j := \exp(-\lambda_j x_i^j).$$

Let $\mathcal{U}_i$ denote the event in which player i receives no message at Stage 4 of the game (hereinafter, the *decision making stage*), which happens if the information transmission from the ω-revealing source to player i—either directly or indirectly through another player—is completely disrupted. Its probability in state ω is given by

$$\mathbb{P}_x\left(\mathcal{U}_i \mid \omega\right) := \delta_i^{\omega\text{-}revealing} \prod_{j \in \mathcal{I}-\{i\}} (\delta_j^{\omega\text{-}revealing} + (1 - \delta_j^{\omega\text{-}revealing})\delta_i^j),$$

where $x := (x_i, x_{-i})$ denotes the joint attention strategy profile across all players.

At the decision making stage, player i earns zero utility if he learns the state and acts accordingly. Otherwise event $\mathcal{U}_i$ happens, and he must choose between the two available options. The optimal choice yields the following ex ante expected utility:

$$\max \left\{ -\frac{\beta_i}{2} \mathbb{P}_x \left(\mathcal{U}_i \mid \omega \neq t_i \right), -\frac{1}{2} \mathbb{P}_x \left(\mathcal{U}_i \mid \omega = t_i \right) \right\},$$

where the first and second terms in the big bracket constitute the expected utilities in case one makes his default decision and the other decision in event $\mathcal{U}_i$, respectively. At Stage 1 of the game (hereinafter the *attention paying stage*), the player chooses a feasible attention strategy to maximize the above expression, taking the other players' attention strategies x_{-i} as given.[8]

3.2 Key concepts

We first define the biases of primary sources, as in Che and Mierendorff (2019).

Definition 1. *A primary source is* biased toward decision $d \in \{A, B\}$ *(or simply* d-biased*) if attending to that source but receiving no message from it increases one's belief that the state favors decision d. A player's* own-biased source *is the primary source that is biased toward his default decision.*

In the leading example, the A-revealing source is an experiment designed to reject the null hypothesis that the state is B. It does so through generating A-revealing evidence, absent of which state B becomes more likely, hence the name "B-biased." A player's own-biased source favors his default decision as its null hypothesis. In case the player attends to that source but receives no message from it, he reinforces the belief that the state favors his default decision. For a more natural interpretation, one can think of no message as a "babbling message" that favors the player's default decision. In what follows, we will denote the A-revealing (resp. B-revealing) source by b (resp. a). Type A (resp. B) players' own-biased source is a (resp. b).

Next is the notion of like-minded friends.

Definition 2. *Two players are* like-minded friends *if they share the same default decision (and hence the same own-biased source).*

Next is our key notion of echo chamber equilibrium.

[8]The above formulation remains valid even if the player can first decide how much attention to pay to the primary sources, and, in case he receives from no message from them, divides the remaining attention capacity across the other players.

Definition 3. *An equilibrium is an* echo chamber equilibrium *if each player attends only to his own-biased source and like-minded friends on the equilibrium path.*

An echo chamber equilibrium has two noteworthy features. The first feature is a selective exposure to content and a formation of homogeneous clusters. Indeed, casual inspections suggest that parents who prefer the traditional baby-feeding approach focus on learning the upside of the new approach and share information among each other. The second feature of echo chamber equilibrium is a belief polarization coupled with an occasional yet drastic belief reversal. It is easy to see that after playing an echo chamber equilibrium, each parent receives no message from any source most of the time and updates the belief in favor of his default decision in that event. With a small complementary probability, the opposite happens, and the parent feels strongly about trying a different approach. As discussed in Section 5, both features of echo chambers have solid empirical supports.

While our analysis is mainly conducted in a general environment, we sometimes restrict attention to symmetric environments and symmetric equilibria.

Definition 4. *A society is* symmetric *if the two types of players have the same population size N, and all players have the same characteristic profile (β, λ, τ). An equilibrium is* symmetric *if the equilibrium strategy depends only on the amounts of attention that a typical player pays to his own-biased source, the other primary source, each like-minded friend of his, and any other player, respectively.*

We conclude this section by defining two useful functions and stating their main properties (a full list of properties can be found in Lemma 3 of Appendix A).

Definition 5. *For each $\lambda \geq 0$, define*

$$\phi\left(\lambda\right) := \begin{cases} \log(\frac{\lambda}{\lambda-1}) & \text{if } \lambda > 1, \\ +\infty & \text{if } \lambda \in [0,1]. \end{cases}$$

For each $\lambda > 1$ and $x \in [\phi\left(\lambda\right), +\infty)$, define

$$h\left(x;\lambda\right) := \frac{1}{\lambda}\log\left[\left(\lambda-1\right)\left(\exp\left(x\right)-1\right)\right].$$

Observation 1. *$\phi'(\lambda) < 0$ on $(1, +\infty)$ and $\lim_{\lambda \downarrow 1} \phi(\lambda) = +\infty$. For each $\lambda > 1$, $h(\phi(\lambda), \lambda) = 0$ and $h_x\left(x;\lambda\right) \in (0,1)$ on $\left(\phi\left(\lambda\right), +\infty\right)$.*

3.3 Echo chamber formation

Our first theorem provides sufficient conditions that give rise of echo chamber equilibria. It shows that every equilibrium of the game must be an echo chamber equilibrium when players have sufficiently strong horizontal preferences, holding everything else constant. Alternatively, in a symmetric society, the unique symmetric equilibrium of the game must be an echo chamber equilibrium when the population size is large, or when attention is scarce.

Theorem 1. *(i) Fix any population sizes $|\mathcal{A}|, |\mathcal{B}|$ and characteristic profiles $(\lambda_i, \tau_i)_{i \in \mathcal{A} \cup \mathcal{B}}$ as in Section 2. There exists $\underline{\beta} \in (0,1)$ such that when $\beta_i \in (0, \underline{\beta})$ $\forall i \in \mathcal{I}$, every equilibrium of the game must be an echo chamber equilibrium. (ii) Consider a symmetric society parameterized by $(N, \beta, \lambda, \tau)$.*

(a) For any $\beta \in (0,1)$, $\lambda > 0$, and $\tau > 0$, there exists $\underline{N} \in \mathbb{N} - \{1\}$ such that for any $N > \underline{N}$, the unique symmetric equilibrium of the game is an echo chamber equilibrium.

(b) For any $\beta \in (0,1)$, $\lambda > 0$, and $N \in \mathbb{N} - \{1\}$, there exists $\underline{\tau} > 0$ such that for any $\tau < \underline{\tau}$, the unique symmetric equilibrium of the game is an echo chamber equilibrium.

The conditions prescribed Theorem 1 accurately describe the world we are living in. As more and more people turn to the Internet and social media where the amount of available information is vastly greater than what an individual can process in a lifetime, it is reasonable to model bandwidth τ_is as finite, if not small, numbers. Meanwhile, the use of automated systems has destroyed the physical boundaries between people, enabling virtual connections between like-minded friends who have never met before in reality. In terms of modeling, this means that we can look at the case of a large N, and assume that the allocation of attention across information sources is flexible. According to Theorem 1, these conditions are conducive to echo chamber formation, especially when people are sufficiently biased to begin with.

The remainder of this section explains the ideas behind Theorem 1 in two steps.

Basic intuitions. Consider first a benchmark case in which players can only attend to the primary sources but not to each other. The next lemma, also proven by Che and Mierendorff (2019) as part of their preliminary analysis, solves the optimal decision problems that players face.

Lemma 1. *Let everything be as in Section 2 except that Stage 3(b) is removed from the game. Then each player attends only to his own-biased source but not the other source at the attention paying stage.*

The idea behind Lemma 1 is as follows: since a player can always make his default decision without paying attention, paying attention is useful only if it sometimes convince him to act differently. Achieving this goal requires that the player disapproves of his default decision, using the very kind of the information generated by his own-biased source. Moreover, since one can never be certain that the state is ω even if he focuses exclusively on the ω-revealing source, it is optimal to attend only to his own-biased source but not the other source. The takeaway from this exercise is that rational inattention generates heterogeneous demands for information among people who are initially biased towards different decisions.

We next allow players to attend to each other. Assuming that every player makes his default decision in event $\mathcal{U}_i$ (where he doesn't hear from any source), the argument articulated above remains valid, namely it is optimal to attend only to one's own-biased source but not the other source. Likewise, it is optimal to attend only to one's like-minded friends but no one else, as the former share the same own-biased source with him, and so could serve as secondary sources in case the information transmission from the primary source to him is disrupted. This is the basic intuition for why echo chambers could arise in equilibrium.

Limit strategic gains from switching sides. Yet in a strategic environment like ours, players may benefit strategically from switching sides, as illustrated by the next example.

Example 2. There is one player of each type called Alice (A) and Bob (B). If both players attend only to their own-biased sources, then Alice's expected utility is $-\beta_A \exp(-\tau_A)/2$. Now suppose that Alice decides to switch sides, i.e., make decision B in event $\mathcal{U}_A$. To facilitate decision making, she pays $x \in [0, \tau_A]$ units of attention to the other primary source and $\tau_A - x$ units of attention to Bob. Her problem becomes

$$\max_{x \in [0,\tau_A]} -\frac{1}{2} \exp(-x) \left[\exp(-\tau_B) + (1 - \exp(-\tau_B)) \exp(-\lambda_B(\tau_A - x))\right].$$

Solving the above problem shows that $x = \max\{\tau_A - h(\tau_B, \lambda_B), 0\}$, and that Alice benefits from switching sides if $\lambda_B > 1$, $\tau_B > \phi(\lambda_B)$, and $h(\tau_B, \lambda_B) - \tau_B + \phi(\lambda_B) < \log \beta_A$. The first two conditions guarantee that $h(\tau_B, \lambda_B) > 0$ (recall Observation 1), and they are easy to understand: if $\lambda_B \leq 1$, then Bob is less visible than the primary source and so should be ignored by Alice in any equilibrium. If $\tau_B \leq \phi(\lambda_B)$, then Bob lacks the capacity to absorb enough information and pass it along to Alice. In both cases, Alice will attend only to source A if she decides to switch sides, i.e., $x_A = \tau_A$. Given her horizontal preference for decision A, switching sides is unprofitable.

The last condition is most likely to hold when β_A is large, and so Alice has only a mild preference for her default decision; and when τ_B is large (recall from Observation 1 that $h_x \in (0, 1)$ on the relevant domain), and so Bob can absorb a lot of information and pass it along to Alice. These are the exact confounding forces we wish to rule out in order to sustain echo chambers in equilibrium. $\diamondsuit$

The construction presented above is relatively straightforward because there is only one player of each type. As a consequence, Bob's strategy in an echo chamber equilibrium is to attend to source B, and we used this fact to calculate Alice's expected utility gain from switching sides. The proof presented in the appendix handles the multi-agent case, in which players' equilibrium strategies are much harder, if not impossible, to compute explicitly.

Part (i) of Theorem 1 is the most intuitive: when players' horizontal preferences are sufficiently strong, making the default decision in event $\mathcal{U}_i$ becomes a dominant strategy, regardless of the belief he holds. As a result, any equilibrium must be an echo chamber equilibrium, for the same reason as given in the benchmark case. The existence of an equilibrium will become clear in the next section.[9]

When players' horizontal preferences are mild, we make progress by studying symmetric societies, as in Part (ii) of Theorem 1. To gain insights into Part (ii-a) of the theorem, notice that when all players except $i \in \mathcal{A}$ adopt equilibrium strategies, player i faces two choices: attend to his own-biased source and $N-1$ like-minded

[9] Part (i) of Theorem 1 is robust to several model variations. First and foremost, it is easy to see that the result remains valid even if players can communicate for more than one round. While the complete segregation between different types of players is an artifact of binary states and decisions, as well as the assumption that messages fully reveal the true state, the idea that rational inattention leads like-minded people to focus on similar information sources and on each other should and indeed has a life of its own. In Online Appendix O.3, we extend the baseline model to arbitrarily finite decisions and states. We establish a pattern called *semi echo chamber*, whereby players pay most, but not all their attention to own-biased sources and like-minded friends. The same pattern can also be established when primary source messages entail small false positive and false negative rates.

friends and make the default decision in event $\mathcal{U}_i$; or attend to the B-biased source and N type B players and make decision B in event $\mathcal{U}_i$. When N is small, switching sides significantly increases the amount of information that player i can glean from secondary sources. Yet the resulting gain vanishes and eventually becomes a loss as N grows to infinity.

Part (ii-b) of Theorem 1 can be understood as follows: when attention is abundant, as captured by a large τ, switching sides exposes player i to one more highly informed secondary source, and the resulting gain is shown to increase with τ. Interestingly and importantly, this is true despite that raising τ also makes player i's like-minded friends more informed as secondary sources. To sustain echo chambers in equilibrium, τ mustn't be too large.

3.4 Inside an echo chamber

Our second theorem gives a complete characterization of the equilibrium attention network within an echo chamber. Without loss of generality (w.l.o.g.), consider the echo chamber among type A players.

Theorem 2. *The following are true for any $i \in \mathcal{A}$ in any echo chamber equilibrium.*

(i) If all type A players attend to their own-biased source, i.e., $x_j^a > 0\ \forall j \in \mathcal{A}$, then the following are equivalent: (a) $x_j^i > 0$ for some $j \in \mathcal{A} - \{i\}$; (b) $x_i^a > \phi(\lambda_i)$; (c) $x_j^i \equiv h(x_i^a; \lambda_i)\ \forall j \in \mathcal{A} - \{i\}$.

(ii) $x_i^a = [\tau_i - \sum_{j \in \mathcal{A}-\{i\}} \underbrace{\frac{1}{\lambda_j} \log \max\left\{(\lambda_j - 1)(\exp(x_j^a) - 1), 1\right\}}_{=x_i^j \text{ if } x_i^a > 0}]^+$

(iii) If all type A players attend to each other, i.e., $x_j^k > 0\ \forall j \in \mathcal{A}$ and $k \in \mathcal{A} - \{j\}$, then player i's ex ante expected utility equals

$$-\frac{\beta_i}{2} \exp\Big(-\sum_{j \in \mathcal{A}} x_j^a + \sum_{j \in \mathcal{A}-\{i\}} \phi(\lambda_j)\Big).$$

Part (i) of Theorem 2 shows that if player i wishes to be attended by a like-minded friend, then he must first cross his *threshold of being visible* $\phi(\lambda_i)$, i.e., pay at least $\phi(\lambda_i)$ units of attention to the primary source. After that, he receives the same amount $h(x_i^a; \lambda_i)$ of attention from all his friends, which increases with the amount of attention x_i^a that he pays to the primary source. For this reason, we shall hereinafter call x_i^a player i's level of *informedness* as a secondary source. The following observations are important.

Core-periphery architecture. Fix any equilibrium as in Theorem 2. Define $\mathcal{COR} = \{i \in \mathcal{A} : x_i^a > \phi(\lambda_i)\}$ as the set of the players who are attended by their like-minded friends, and $\mathcal{PER} = \{i \in \mathcal{A} : x_i^a \leq \phi(\lambda_i)\}$ as the set of the players who are ignored by their like-minded friends. When both sets are nonempty, a *core-periphery architecture* emerges, whereby $\mathcal{COR}$ players acquire information from the primary source and share results among each other, whereas $\mathcal{PER}$ players tap into $\mathcal{COR}$ for secondhand information but are themselves ignored by any player. In order for player i to belong to $\mathcal{COR}$, he must be, first of all, more visible than the primary source, i.e., $\lambda_i > 1$. In addition, he must have a large enough bandwidth to satisfy $\tau_i > \phi(\lambda_i)$. This suggests that a core-periphery architecture is most likely to emerge when players are heterogeneous, so that those players with large bandwidths and high visibility parameters form $\mathcal{COR}$, and the remaining players form $\mathcal{PER}$.[10] Finally, notice that the horizontal preference parameter β_is do not affect the division between $\mathcal{COR}$ and $\mathcal{PER}$. Indeed, they have no impact on the equilibrium attention network within an echo chamber.

Informedness as strategic substitutes. For any $\mathcal{COR}$ player, we define his *influence* on public opinion as the amount $h(x_i^a; \lambda_i)$ of attention he receives from any other player. From $h_x > 0$ (recall Observation 1), it follows that different players' informedness levels are strategic substitutes: as a player becomes more informed, his like-minded friends pay more attention to him and less attention to the primary

[10] A sizable economic literature pioneered by Bala and Goyal (2000) examines when equilibrium social networks exhibit core-peripheral structures. In our case, numerical analysis suggests that a small difference between players is often enough to sustain a core-periphery architecture in equilibrium; see Figure 1 of Appendix B. The intuition behind this finding—which exploits the strategic substitutability between players' informedness levels—will be explained together with that of Theorem 3, in the next section.

Recently, Perego and Yuksel (2016) study a related trade-off between producing information and searching for information in a dynamic learning model. In their setting, a player either produces information or searches undirectedly for information, but not both, and he cannot learn directly from other information producers in the first case. Here, the interactions between core players are central to the comparative statics analysis in Section 3.5. Peripheral players may learn directly from the primary source, and they actively decide how to allocate attention capacities among core players.

source. In the next section, we examine the consequences of this observation for equilibrium.

Parts (i) and (ii) of Theorem 2 prescribe a two-step algorithm for computing all echo chamber equilibria. The first step is to pin down players' informedness levels, by solving the system of equations in as Part (ii) of the theorem. The second step is to back out the attention network between players using Part (i) of the theorem. Specifically, if player i pays a positive amount of attention to his own-biased source, i.e., $x_i^a > 0$, then the amount of attention he pays to a different player j equals

$$\frac{1}{\lambda_j}\log\max\left\{(\lambda_j-1)(\exp(x_j^a)-1),1\right\} = \begin{cases} h(x_j^a;\lambda_j) & \text{if } j \in \mathcal{COR}, \\ 0 & \text{if } j \in \mathcal{PER}. \end{cases}$$

If, instead, $x_i^a = 0$, then the above expression must be scaled by the Lagrange multiplier associated with the nonnegative constraint $x_i^a \geq 0$. Regardless of which case we end up with, the equilibrium attention network between players is always fully determined by their informedness levels. Since the system of equations that governs the latter has a solution by the Brouwer fixed point theorem, an echo chamber equilibrium must exist when β_is are small, which completes the proof of Theorem 1(i).

Part (iii) of Theorem 2 shows that when all players belong to $\mathcal{COR}$, their equilibrium expected utilities depend positively on the total amount of attention that the entire echo chamber pays to the primary source, and negatively on the visibility thresholds of their like-minded friends. Intuitively, members of an echo chamber become better off as they collectively acquire more information from the primary source, and as they become more capable of disseminating information to each other.

3.5 Comparative statics

This section investigates the comparative statics of the equilibrium attention network within an echo chamber, focusing again on the case among type A players. For ease of notation, we write $\{1, \cdots, N\}$ for $\mathcal{A}$, θ_i for (λ_i, τ_i), and $\boldsymbol{\theta}$ for $[\theta_1, \cdots, \theta_N]^\top$. The next regularity condition is maintained throughout this section.

Assumption 1. *The game among type A players has a unique equilibrium, and all type A players attend to each other in that equilibrium.*

Assumption 1 has two parts. The first part will be elaborated upon in Online Appendix O.4, where we provide sufficient conditions for the uniqueness of equilibrium. The second part requires that all players belong to $\mathcal{COR}$, and its sole purpose is to simplify the exposition: as demonstrated in Online Appendix O.5, introducing $\mathcal{PER}$ players into the analysis does not affect any qualitative prediction of ours.

We conduct two exercises. The first exercise fixes players' population size and varies their individual characteristics. The second exercise assumes that players are homogeneous and varies their population size.

3.5.1 Individual characteristics

The next theorem examines the effects of perturbing a single player's characteristics on the equilibrium attention network.

Theorem 3. *Fix any $N \in \mathbb{N} - \{1\}$, and let Θ be any neighborhood in $\mathbb{R}^{2N}_{++}$ such that for any $\boldsymbol{\theta} \in \Theta$, the game among a set $\mathcal{A}$ of type A players with population size N and characteristic profile $\boldsymbol{\theta}$ satisfies Assumption 1. Then the following must hold for any $i \in \mathcal{A}$, $j \in \mathcal{A} - \{i\}$, and $k \in \mathcal{A} - \{j\}$ ($i = k$ is allowed), at any $\boldsymbol{\theta}^\circ \in int(\Theta)$.*

(i) $\partial x_i^a/\partial \tau_i|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ} > 0$, $\partial x_j^i/\partial \tau_i\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ} > 0$, $\partial x_j^a/\partial \tau_i\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ} < 0$, *and* $\partial x_k^j/\partial \tau_i\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ} < 0$.

(ii) One of the following situations happens:

- *(a)* $\partial x_i^a/\partial \lambda_i|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ} > 0$, $\partial x_j^i/\partial \lambda_i\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ} > 0$, $\partial x_j^a/\partial \lambda_i\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ} < 0$, *and* $\partial x_k^j/\partial \lambda_i\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ} < 0$;
- *(b) the inequalities in Part (a) are all reversed;*
- *(c) the inequalities in Part (a) are all replaced with equalities.*

(iii) If $\theta_n \equiv \theta \; \forall n \in \mathcal{A}$, *then* $\partial \sum_{n \in \mathcal{A}} x_n^a/\partial \tau_i\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ} > 0$ *and* $\operatorname{sgn}\left(\partial \sum_{n \in \mathcal{A}} x_n^a/\partial \lambda_i\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ}\right) = \operatorname{sgn}\left(- \partial x_i^a/\partial \lambda_i|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ}\right)$.

Part (i) of Theorem 3 shows that increasing a player's bandwidth raises his informedness and influence as a secondary source, and so promotes him to become an opinion leader. More surprisingly, the change diminishes the informedness and influence of any other player, who thus becomes an opinion follower. As depicted in Figure 1 of Appendix B, this equilibrium mechanism can magnify even a small difference between people's bandwidths into a very uneven distribution of opinions, whereby some people occupy the center of attention, while others are barely visible. An important feature of today's news landscape is that while most Americans express curiosity in many topics such as science and politics, only a minority of them are willing to spend time on hard news consumption (Prior, 2007; Funk, Gottfried, and Mitchell, 2017). According to Theorem 3(i), this attentional gap between the majority and minority may lead the former to consume most firsthand news, and the latter to rely mainly on the secondhand opinions that are passed along to them from the former. Patterns consistent with this prediction, such as the law of the few and fat-tailed distributions of opinions, have recently been detected in the social media sphere (Lu, Zhang, Cao, Hu, and Guo, 2014; Del Vicario, Bessi, Zollo, Petroni, Scala, Caldarelli, Stanley, and Quattrociocchi, 2016; Néda, Varga, and Biró, 2017).

Part (ii) of Theorem 3 shows that increasing a player's visibility parameter by a small amount may promote his informedness and influence while diminishing that of the other players. But the opposite can also happen. Alternatively, the effect can be neutral. Two countervailing effects are at work here. On the one hand, raising player i's visibility parameter λ_i reduces the threshold $\phi(\lambda_i)$ at which he starts to exert influences on the other players (recall from Observation 1 that $\phi' < 0$). We refer to this effect as the *intercept effect*. On the other hand, now that player i has become a better information disseminator, the amount of attention that his friends pay to him no longer varies as sensitively with his informedness as it used to. We refer to this effect as the *slope effect*, and note that it goes in the opposite direction of the intercept effect. In general, either effect can dominate the other (as depicted in Figure 2 of Appendix B), which renders the comparative statics ambiguous. To counter the rising threat from misinformation and fake news, many social media companies have taken measures that operate through modulating account visibility. Among others, Facebook imposes a daily posting limit of 25 articles, beyond which the reach of the posts will be negatively affected. Theorem 3(ii) warns that such measures may backfire and, as demonstrated in the next paragraph, undermine consumer welfare, if they are not well calibrated according to the underlying environment.

Part (iii) of Theorem 3 concerns the total amount of attention that the entire echo chamber pays to the primary source, which is a crucial determinant of players' equilibrium expected utilities. In general, nothing clear-cut can be said, which is unsurprising given how few assumptions we have made about the magnitudes of the strategic substitution effects across people. This finding suggests that the aforementioned policy interventions could entail ambiguous welfare consequences, and so should be implemented with caution and care. When players are homogeneous, our predictions are clear, as increasing a player's bandwidth makes everyone in the echo chamber better off.[11] As for the consequences of increasing one's visibility parameter, our result depends on whether that player ends up being an opinion leader or an opinion follower: the entire echo chamber pays less attention to the primary source in the first case, and more attention to the primary source in the second case.

Proof sketch. Proving Theorem 3 requires a new method that we now develop. Due to space limitations, we only sketch the proof for $\partial x_1^a/\partial\tau_1|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ} > 0$ and $\partial x_j^a/\partial\tau_1\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ} < 0$ $\forall j \neq 1$, starting off from the case of two players. In that case, differentiating the system of equations concerning players' equilibrium informedness against τ_1 yields

$$\begin{bmatrix} 1 & \frac{\partial x_1^2}{\partial x_2^a} \\ \frac{\partial x_2^1}{\partial x_1^a} & 1 \end{bmatrix} \begin{bmatrix} \frac{\partial x_1^a}{\partial \tau_1} \\ \frac{\partial x_2^a}{\partial \tau_1} \end{bmatrix} \Bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ} = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

where the term $\partial x_i^j/\partial x_j^a\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ}$ captures how perturbing player j's informedness affects his influence on player i. Write g_j for $h_x(x_j^a;\lambda_j)\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ}$, and recall that

$$\frac{\partial x_i^j}{\partial x_j^a}\Bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ} \underbrace{=}_{\text{Theorem 2}} g_j \underbrace{\in}_{\text{Observation 1}} (0,1),$$

i.e., increasing player j's informedness by one unit raises his influence on player i by less than one unit. From $g_j > 0$, i.e., informedness levels are strategic substitutes, it follows that in order for the perturbation to have zero effect on player 2's bandwidth, it must cause one and only one player to pay more attention to the primary source. Now, since $g_j < 1$, i.e., the strategic substitution effects are sufficiently mild, that player must be player 1, as the direct effect stemming from increasing his bandwidth dominates the indirect effects that he and player 2 could exert on each other.

Extending the above argument to more than two players is a nontrivial task, as it requires that we trace out how the strategic substitution effects reverberate across

[11] To be precise, all we need is that core players be almost homogeneous. Online Appendix O.5 incorporates (heterogeneous) peripheral players into the analysis.

a large and endogenous attention network. Mathematically, we must solve

$$[\mathbf{I}_N + \mathbf{G}_N] \nabla_{\tau_1} [x_1^a \ \cdots \ x_N^a]^\top \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ} = [1, 0, \cdots, 0]^\top,$$

where $\mathbf{I}_N$ is the $N \times N$ diagonal matrix, and $\mathbf{G}_N$ is the *marginal influence matrix* defined as

$$[\mathbf{G}_N]_{i,j} = \begin{cases} 0 & \text{if } i = j, \\ \frac{\partial x_i^j}{\partial x_j^a}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ} & \text{else.} \end{cases}$$

A seemingly innocuous fact proves its usefulness here, namely an individual attracts the same amount of attention from all his friends, i.e., $x_i^j \equiv h(x_j^a; \lambda_j)$ $\forall i \neq j$. Taking derivatives yields $\partial x_i^j / \partial x_j^a \big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ} \equiv g_j$ $\forall i \neq j$, i.e., the off-diagonal entries of $\mathbf{G}_N$ are constant column by column:

$$\mathbf{G}_N = \begin{bmatrix} 0 & g_2 & \cdots & g_N \\ g_1 & 0 & \cdots & g_N \\ \vdots & \vdots & \ddots & \vdots \\ g_1 & g_2 & \cdots & 0 \end{bmatrix}.$$

Based on this fact, as well as $g_j \in (0,1)$ $\forall j \in \{1, \cdots, N\}$, we develop a method for solving $[\mathbf{I}_N + \mathbf{G}_N]^{-1}$ and determining the signs of its entries. Our findings are reported in the next lemma, from which Theorem 3 follows.

Lemma 2. *Fix any $N \in \mathbb{N} - \{1\}$ and $g_1, \cdots, g_N \in (0,1)$, and let $[\mathbf{G}_N]_{i,j} = g_j$ $\forall i \neq j$ in the marginal influence matrix. Then $\mathbf{A}_N := \mathbf{I}_N + \mathbf{G}_N$ is invertible and satisfies $\forall i \in \{1, \cdots, N\}$: (i) $\left[\mathbf{A}_N^{-1}\right]_{i,i} > 0$; (ii) $\left[\mathbf{A}_N^{-1}\right]_{i,j} < 0$ $\forall j \neq i$;(iii) $\sum_{j=1}^N \left[\mathbf{A}_N^{-1}\right]_{i,j} > 0$.*

To the best of our knowledge, Lemma 2 is to new to the literature on network games with negative externalities, as it generates sharp comparative statics predictions without invoking the usual assumptions made in the literature, such as linear best response functions or a symmetric influence matrix. The lemma is useful for other purposes, such as evaluating the consequences of a common shock to players' characteristics; see online Appendix O.2 for such an exercise. While the assumption that an individual must be equally visible to all his friends is certainly crucial, it can be relaxed as long as the environment is sufficiently close to the one considered above (as shown in Online Appendix O.6).

3.5.2 Population size

This section examines the comparative statics regarding the population size N. To obtain the sharpest insights, we assume, unlike in the previous section, that players are homogeneous. Under this assumption, it is easy to see that if the game has a unique equilibrium (as required by Assumption 1), it must be symmetric. Let $x(N)$ denote the amount of attention that a typical player pays to the primary source in that equilibrium. The next proposition investigates the comparative statics of $x(N)$.

Proposition 1. *Take any $\lambda, \tau > 0$, and $N'' > N' \geq 2$ such that the game among $N \in \{N', N''\}$ type A players with visibility parameter λ and bandwidth τ satisfies Assumption 1. As N increases from N' to N'', $x(N)$ decreases, whereas $Nx(N)$ may either increase or decrease.*

As an echo chamber grows in size, each member of it has access to more secondary sources and so pays less attention to the primary source. Depending on the severity of this free-riding problem compared to the population size effect, the overall effect on the amount of attention that the entire echo chamber pays to the primary source and, hence, players' equilibrium expected utilities, is in general ambiguous (as confirmed by numerical analysis).

Recently, several information platforms, including Allsides.com, have been built to combat the rising polarization through exposing users to diverse viewpoints from both sides. In Online Appendix O.2, we model such a platform as a *mega source* that results from merging the A-revealing source and B-revealing source of the baseline model together. We find that the use of a mega source does achieve the purpose of dissolving echo chambers, as it forces different types of players to attend to each other as secondary sources. Yet making more secondary sources available to players discourages information acquisition from primary sources and so exacerbate the free-riding problem. In a symmetric society, merging biased sources into a mega source is mathematically equivalent to doubling the population of type A players in Proposition 1. Its effect on players' equilibrium expected utilities is in general ambiguous.

4 Extensions

This section reports main extensions of the baseline model, in addition to the ones we have already discussed. Due to space constraints, we postpone the analysis to the online appendix, and focus here on results and takeaways.

(In)efficiency of echo chambers. In Online Appendix O.1, we examine the efficient attention network that maximizes the utilitarian welfare of a symmetric society. We focus on the case where players are sufficiently biased, and so making one's default decision is efficient in event $\mathcal{U}_i$. The main qualitative difference between the efficient attention network and an echo chamber equilibrium is, then, that the former can mandate that all players attend to both primary sources. In this way, a lot more players are qualified as secondary sources, despite that each individual cares only about the information generated by one primary source. Such an allocation is efficient when players have large bandwidths and high visibility parameters, and so are good at absorbing information and disseminating it to the others as secondary sources. Yet it cannot be sustained in any equilibrium, because a self-interested individual would only gather information about state A or state B, but not both.

General primary sources. In Online Appendix O.2, we extend the baseline model to multiple primary sources with general visibility parameters. The framework proposed there nests many interesting situations as special cases. For example, if a source is visible to all players in both states, then it is the mega source discussed at the end of Section 3.5.2. If it is only visible to a single player in a single state, then it is a private experiment conducted by that player.

Two findings are noteworthy. First, introducing multiple identical and independent primary sources into the model does not affect the total amount of attention one pays to each kind of source. All it does is to dilute players' attention across the same kind of sources. In our leading example, this finding suggests that increasing the number of independent experiments without improving their qualities may have limited impacts on public opinion and consumer welfare.

Second, when the visibility parameter of primary sources differs from one, all we need to do is to rescale things properly. As for comparative statics, we find that increasing the visibility of primary sources effectively diminishes the visibility of all secondary sources. Then using the toolkit developed in Lemma 2, we find the same ambiguous effect as in Theorem 3(ii). Thus in practice, factors that affect primary sources' visibility—such as the increasing reliance of scientific journals on digital technologies and AI to boost distribution and reach—could have ambiguous effects on public opinion and consumer welfare.

5 Further Application

This section studies a political economy application of our model. We use our results to match additional empirical patterns, and to shed light on several policy issues.

Example 3. Each player $i \in \mathcal{I}$ is now a voter who is affiliated with either the Democratic Party or the Republican Party. His decision is to vote for either the Democratic candidate or the Republican candidate, $d_i \in \{L, R\}$, and his utility is the highest if the candidate he supports has the best quality. In case the voter supports the lowest quality candidate, he incurs a loss of magnitude $\beta_i \in (0, 1)$ if that candidate comes from his own party. If the candidate comes from the opposing party, then the loss has magnitude 1. Expressive voting is an important motive for consuming political news according to Prat and Strömberg (2013).

Candidate quality is uncertain and is modeled as a random state $\omega \in \{L, R\}$. News about ω is generated by an L-revealing source and an R-revealing source. The game begins with voters specifying how much attention they wish to pay to each primary source and to each other. After that, news transmits from the primary source to voters, and then among voters themselves in a nonstrategic manner. While stylized, the last assumption helps us focus on the optimal attention allocation problem while still capturing important facets of reality: according to a recent study by Dewey (2016), 6 out of 10 people share links after glancing quickly at the headlines.

Voters can differ in their bandwidth τ_is, depending on how much time they are willing to spend on hard news consumption versus doing other things, such as consuming entertainment. They can also differ in their visibility parameter λ_is, as public figures like Oprah Winfrey are more capable of deciphering and disseminating the complex messages of elite newspapers to the less educated audience than ordinary people (Baum and Jamison, 2006). $\diamondsuit$

We interpret primary sources as journalists or media outlets that produce original reporting about the state. Following Che and Mierendorff (2019), we call the L-revealing source R-biased, and the R-revealing source L-biased. This practice is inspired by Chiang and Knight (2011), who find that newspaper endorsements for the presidential candidates in the United States are most effective in shaping voters' beliefs and decisions if the endorsement goes against of the bias of the newspaper. Additional evidence for this view is surveyed by DellaVigna and Gentzkow (2010).

Our model predicts after playing an echo chamber equilibrium, a majority of the voters will have more faith in their own-party candidates than before, while a minority of them will become supportive of the opposing party candidate. The coexistence of a belief polarization and an occasional yet drastic belief reversal is a hallmark of Bayesian rationality. Its presence among voters after social media consumption has recently been detected by Flaxman, Goel, and Rao (2016), Balsamo, Gelardi, Han, Rama, Samantray, Zucca, and Starnini (2019), and Allcott, Braghieri, Eichmeyer, and Gentzkow (2020), among others.

As for comparative statics, we note that today's high-choice media environment and the resulting separation of news junkies from the majority of people may generate the fat-tailed distribution of opinions that Farrell and Drezner (2008) find among political blogs. In the aftermath of the 2021 U.S. Capitol attack, there have been calls to modify Section 230 of the Communications Decency Act of 1996 so as to allow Internet companies to exercise more account control (Romm, 2021). We caution that account control via modulating source visibility may backfire if it is not well calibrated according to underlying environment.

Our results speak to the effectiveness of the FCC's viewpoint diversity objectives and, more specifically, the eight voice rule. The latter mandates that at least eight independent media outlets must be operating in the same digital media area (Ho and Quinn, 2008). As shown in Online Appendix O.2, increasing the number of independent primary sources without improving their qualities has no real impact in the setting we consider.

6 Concluding remarks

This paper develops a rational inattention theory of echo chamber. To single out the economic forces of our interest, we abstract away from alternative considerations such as strategic information sources, behavioral players, etc. To the extent that the latter are believed to foster echo chambers, our model prescribes a lower bound for the degree of opinion segregation that can arise and persist in reality. Its normative implications, especially those that question the effectiveness of interventions that target the rising misinformation and polarization, are relevant in more general, complex, environments.

Our analysis generates several new predictions, including conducive conditions to echo chamber formation and the comparative statics of equilibrium attention networks, in addition to patterns that have been documented by other authors in different contexts, such as the belief distribution after social media consumption. Testing these predictions together in a rigorous manner is an interesting avenue for future research. One way to make progress is to run a controlled experiment on social media, as has done by scholars working on related topics (see, e.g., Allcott, Braghieri, Eichmeyer, and Gentzkow 2020).

A Proofs

A.1 Useful lemmas and their proofs

Proof of Lemma 1. When players can only attend to the primary sources but not to each other, the ex ante problem faced by a type A player who makes decision A in event $\mathcal{U}_i$ is

$$\max_{x^a,x^b} -\frac{\beta_i}{2}\exp(-x^a) \text{ s.t. } x^a, x^b \geq 0 \text{ and } \tau_i \geq x^a + x^b.$$

If, instead, the player makes decision B in event $\mathcal{U}_i$, then his ex ante problem becomes

$$\max_{x^a,x^b} -\frac{1}{2}\exp(-x^b) \text{ s.t. } x^a, x^b \geq 0 \text{ and } \tau_i \geq x^a + x^b.$$

The solutions to these problems are $(x^a, x^b) = (\tau_i, 0)$ and $(x^a, x^b) = (0, \tau_i)$. The first solution generates a higher expected utility and so is optimal. □

Proof of Lemma 2. We proceed in three steps.

Step 1. Solve for $\mathbf{A}_N^{-1}$. We conjecture that

$$\det(\mathbf{A}_N) = 1 + \sum_{s=1}^{N} (-1)^{s-1}(s-1) \sum_{\substack{(k_l)_{l=1}^s : k_l \in \{1,\cdots,N\},\\ k_1<\cdots<k_s}} \prod_{l=1}^{s} g_{k_l}, \tag{1}$$

and that the following hold $\forall i \in \{1, \cdots, N\}$ and $j \in \{1, \cdots, N\} - \{i\}$:

$$\left[\mathbf{A}_N^{-1}\right]_{i,i} = \frac{1}{\det(\mathbf{A}_N)}[1 + \sum_{s=1}^{N-1} (-1)^{s-1}(s-1) \sum_{\substack{(k_l)_{l=1}^s : k_l \in \{1,\cdots,N\}-\{i\},\\ k_1<\cdots<k_s}} \prod_{l=1}^{s} g_{k_l}] \tag{2}$$

and

$$\left[\mathbf{A}_N^{-1}\right]_{i,j} = \frac{-g_j}{\det\left(\mathbf{A}_N\right)} \prod_{k \in \{1,\cdots,N\}-\{i,j\}} \left(1 - g_k\right). \tag{3}$$

We omit most algebra, but note that $\mathbf{A}_N$ can be rewritten as $\mathbf{B}_N + \mathbf{u}_N \mathbf{v}_N^\top$, where $\mathbf{B}_N \coloneqq \text{diag}(1 - g_1, \cdots, 1 - g_N)$, $\mathbf{u}_N$ is the N-vector of ones, and $\mathbf{v}_N \coloneqq [g_1, \cdots, g_N]^\top$. Applying the Sherman-Morrison formula (Sherman and Morrison, 1950) shows that

$$\mathbf{A}_N^{-1} = \mathbf{B}_N^{-1} - \frac{\mathbf{B}_N^{-1} \mathbf{u}_N \mathbf{v}_N^\top \mathbf{B}_N^{-1}}{1 + \mathbf{v}_N^\top \mathbf{B}_N^{-1} \mathbf{u}_N},$$

and simplifying the last expression by doing lengthy algebra gives the desired result.

Step 2. Show that $\det\left(\mathbf{A}_N\right) > 0$, i.e,

$$\sum_{s=1}^{N} (-1)^{s-1} (s-1) \sum_{\substack{(k_l)_{l=1}^s : k_l \in \{1,\cdots,N\}, \\ k_1 < \cdots < k_s}} \prod_{l=1}^{s} g_{k_l} > -1.$$

Denote the left-hand side of the above inequality by $\text{LHS}\left(g_1, \cdots, g_N\right)$. Since the function $\text{LHS} : [0,1]^N \to \mathbb{R}$ is linear in each g_i, holding $(g_j)_{j \neq i}$ constant, its minimum is attained at an extremal point of $[0,1]^N$. Indeed, since the function is symmetric across g_is, the following must hold for any $(g_1, \cdots, g_N) \in \{0,1\}^{N-1}$ such that $\sum_{i=1}^N g_i = n$:

$$\text{LHS}\left(g_1, \cdots, g_N\right) = f(n) \coloneqq \sum_{k=1}^{n} (-1)^{k-1} (k-1) \binom{n}{k}.$$

It remains to show that $f(n) \geq -1$ $\forall n = 0, 1, \cdots, N$. When $n = 0$ and 1, $f(n) = 0$. For each $n \geq 2$, define

$$p(n) \coloneqq \sum_{k=1}^{n} (-1)^k k \binom{n}{k}.$$

Below we prove by induction that $f(n) = -1$ and $p(n) = 0$ $\forall n = 2, \cdots, N$.

Our conjecture is clearly true when $n = 2$, as

$$f(2) = -\binom{2}{2} = -1 \text{ and } p(2) = -\binom{2}{1} + 2\binom{2}{2} = 0.$$

Now suppose that it is true for some $n \geq 2$. Then

$$
\begin{aligned}
& f(n+1) \\
& = \sum_{k=1}^{n+1} (-1)^{k-1} (k-1) \binom{n+1}{k} \\
& = \sum_{k=1}^{n} (-1)^{k-1} (k-1) \binom{n+1}{k} + (-1)^n n \binom{n+1}{n+1} \\
& = \sum_{k=1}^{n} (-1)^{k-1} (k-1) \left(\binom{n}{k} + \binom{n}{k-1} \right) + (-1)^n n \qquad \left(\because \tbinom{n+1}{k} = \tbinom{n}{k} + \tbinom{n}{k-1}\right) \\
& = f(n) + 0 \binom{n}{0} + \sum_{k=1}^{n-1} (-1)^k k \binom{n}{k} + (-1)^n n \binom{n}{n} \\
& = f(n) + p(n) \\
& = -1,
\end{aligned}
$$

and

$$
\begin{aligned}
& p(n+1) \\
& = \sum_{k=1}^{n+1} (-1)^k k \binom{n+1}{k} \\
& = \sum_{k=1}^{n} (-1)^k k \left(\binom{n}{k} + \binom{n}{k-1} \right) + (-1)^{n+1} (n+1) \qquad \left(\because \tbinom{n+1}{k} = \tbinom{n}{k} + \tbinom{n}{k-1}\right) \\
& = p(n) + \sum_{k=0}^{n-1} (-1)^{k+1} (k+1) \binom{n}{k} + (-1)^{n+1} (n+1) \\
& = 0 + \sum_{k=1}^{n-1} (-1)^{k+1} k \binom{n}{k} + \sum_{k=0}^{n-1} (-1)^{k+1} \binom{n}{k} + (-1)^{n+1} (n+1).
\end{aligned}
$$

Then from

$$
\begin{aligned}
\sum_{k=1}^{n-1} (-1)^{k+1} k \binom{n}{k} & = \sum_{k=1}^{n} (-1)^{k+1} k \binom{n}{k} - (-1)^{n+1} n \\
& = -p(n) - (-1)^{n+1} n \\
& = -(-1)^{n+1} n,
\end{aligned}
$$

it follows that

$$p\left(n+1\right)=\sum_{k=0}^{n-1}\left(-1\right)^{k+1}\binom{n}{k}+\left(-1\right)^{n+1},$$

and so $p(n+1)=0$ if and only if

$$q\left(n\right):=\sum_{k=0}^{n-1}\left(-1\right)^{k+1}\binom{n}{k}=\left(-1\right)^{n}.$$

When n is odd, simplifying $q(n)$ using the fact that

$$(-1)^{k+1}\binom{n}{k}+(-1)^{n-k+1}\binom{n}{n-k}=0\ \forall k\in\{1,\cdots,n-1\},$$

yields

$$q(n)=(-1)\binom{n}{0}+0=(-1)^{n},\ \text{as desired.}$$

When n is even, expanding $q\left(n\right)$ yields

$$\begin{aligned}
q\left(n\right)&=\sum_{k=1}^{n-1}\left(-1\right)^{k-1}\left(\binom{n-1}{k-1}+\binom{n-1}{k}\right)-\binom{n}{0} && \left(\because \tbinom{n}{k}=\tbinom{n-1}{k-1}+\tbinom{n-1}{k}\right)\\
&=\sum_{k=1}^{n-1}\left(-1\right)^{k-1}\binom{n-1}{k-1}+\sum_{k=0}^{n-1}\left(-1\right)^{k-1}\binom{n-1}{k}. && \left(\because -\tbinom{n}{0}=(-1)^{-1}\tbinom{n-1}{0}\right)
\end{aligned}$$

Then from

$$\sum_{k=1}^{n-1}\left(-1\right)^{k-1}\binom{n-1}{k-1}=\sum_{k=0}^{n-2}\left(-1\right)^{k}\binom{n-1}{k}=-q\left(n-1\right)=1\quad(\because q(n-1)=-1)$$

and

$$\begin{aligned}
\sum_{k=0}^{n-1}\left(-1\right)^{k-1}\binom{n-1}{k}&=\sum_{k=0}^{n-2}\left(-1\right)^{k-1}\binom{n-1}{k}+(-1)^{n-2}\binom{n-1}{n-1}\\
&=q\left(n-1\right)+1\\
&=0,
\end{aligned}$$

it follows that $q\left(n\right)=1=(-1)^{n}$, as desired.

Step 3. Verify that $\left[\mathbf{A}_N^{-1}\right]_{i,i} > 0$, $\left[\mathbf{A}_N^{-1}\right]_{i,j} < 0$, and $\sum_{j=1}^{N} \left[\mathbf{A}_N^{-1}\right]_{i,j} > 0$ $\forall i \in \{1, \cdots, N\}$ and $j \in \{1, \cdots, N\} - \{i\}$.

$\left[\mathbf{A}_N^{-1}\right]_{i,j}$ is clearly negative. To show that $[\mathbf{A}_N^{-1}]_{i,i}$ is positive, define $\mathbf{D}_{N,i}$ as the principal minor of $\mathbf{A}_N$ that obtains from deleting the i^{th} row and column of $\mathbf{A}_N$. Since $\mathbf{D}_{N,i}$ satisfies all the properties stated in Lemma 2, $\det(\mathbf{D}_{N,i})$ must be positive by Step 2. Then from the fact that $\left[\mathbf{A}_N^{-1}\right]_{i,i} = \det(\mathbf{D}_{N,i}) / \det(\mathbf{A}_N)$, it follows that $[\mathbf{A}_N^{-1}]_{i,i} > 0$, as desired. Finally, summing $\left[\mathbf{A}_N^{-1}\right]_{i,j}$ over $j \in \{1, \cdots, N\}$ and doing lengthy algebra yields

$$\sum_{j=1}^{N} \left[\mathbf{A}_N^{-1}\right]_{i,j} = \frac{1}{\det(\mathbf{A}_N)} \prod_{k \in \{1,\cdots,N\}-\{i\}} (1 - g_k) > 0, \text{ as desired.} \quad \square$$

Lemma 3. *$\phi'(\lambda) < 0$ on $(1, +\infty)$ and $\lim_{\lambda \downarrow 1} \phi(\lambda) = +\infty$. For each $\lambda > 1$, $h(\cdot; \lambda)$ satisfies (i) $h(\phi(\lambda); \lambda) = 0$; (ii) $h_x(x; \lambda) \in (0, 1)$ on $(\phi(\lambda), +\infty)$ and $\lim_{x \downarrow \phi(\lambda)} h_x(x; \lambda) = 1$; and (iii) $h_{xx}(x, \lambda) < 0$ on $(\phi(\lambda), +\infty)$.*

Proof. The result follows from straightforward algebra. $\square$

Lemma 4. *Fix any $\lambda > 1$ and $\tau > \phi(\lambda)$. For each $N \in \mathbb{N} - \{1\}$ and $x \in [\phi(\lambda), +\infty)$, define*

$$\varphi^N(x) \coloneqq \tau - (N-1)h(x; \lambda).$$

Then φ^N has a unique fixed point $x(N)$ that satisfies $x(N) \in (\phi(\lambda), \tau)$, $\lim_{N \to +\infty} x(N) = \phi(\lambda)$, $\lim_{\tau \to +\infty} x(N) = +\infty$, and $dx(N)/d\tau \in (1/N, 1)$.

Proof. Since $h(\phi(\lambda); \lambda) = 0$ and $h_x(x; \lambda) > 0$ on $(\phi(\lambda), +\infty)$ by Lemma 3, $\varphi^N(x) = \tau - (N-1)h(x; \lambda)$ has a unique fixed point $x(N)$ that belongs to $(\phi(\lambda), \tau)$ (drawing a picture will make this point clear). In order to satisfy $x(N) = \varphi^N(x(N)) > \phi(\lambda)$, $h(x(N); \lambda)$ must converge to zero as $N \to +\infty$, which, together with Lemma 3, implies that $\lim_{N \to +\infty} x(N) = \phi(\lambda)$.

Turning to the relationship between $x(N)$ and τ, notice that $x(N)$ must grow to infinity as $\tau \to +\infty$ in order to satisfy $x(N) + (N-1)h(x(N); \lambda) = \tau$. Taking the total derivative of $\varphi^N(x(N)) = x(N)$ with respect to τ yields $dx(N)/d\tau = (1 + (N-1)h_x(x(N); \lambda))^{-1}$, where the last term lies in $(1/N, 1)$ because $h_x \in (0, 1)$. $\square$

A.2 Proofs of theorems and propositions

Proof of Theorems 1(i) and 2. We proceed in four steps.

Step 1. Show that for any type A player i, making the default decision in event $\mathcal{U}_i$ is a dominant strategy when β_i is sufficiently small.

If the player attends only to source a and makes decision A in event $\mathcal{U}_i$, then his ex ante expected utility equals $-\beta_i \exp(-\tau_i)/2$. If, instead, he makes decision B in event $\mathcal{U}_i$, then his ex ante expected utility is at most $-\exp(-\overline{\lambda}\tau_i)/2$, where $\overline{\lambda} := \max\{1, \lambda_j, j \in \mathcal{I} - \{i\}\}$. The reason is that, in a hypothetical situation where all the other players knows for sure when state A occurs, player i essentially faces the original primary source b, together with $|\mathcal{I}| - 1$ primary sources with visibility parameter λ_js, $j \neq i$. His optimal strategy is to focus on the source with the highest visibility parameter $\overline{\lambda}$, and the resulting ex ante expected utility equals $-\exp(-\overline{\lambda}\tau_i)/2$. The last term is smaller than $-\beta_i \exp(-\tau_i)/2$ when $\beta_i < \exp((1 - \overline{\lambda})\tau_i)$. The remainder of the proof assumes that this is the case.

Step 2. Show that any equilibrium must be an echo chamber equilibrium. It suffices to show that $x_i^b = 0$ and $x_i^j = 0$ $\forall i \in \mathcal{A}$ and $j \in \mathcal{B}$. Fix any $x_{-i} \in \mathcal{X}_{-i}$. Rewrite player i's problem: $\max_{x_i \in \mathcal{X}_i} -\beta_i \mathbb{P}_x(\mathcal{U}_i \mid \omega = B)/2$, as follows:

$$\max_{(x_i^c)_{c \in \mathcal{C}_i}} - x_i^a - \sum_{j \in \mathcal{I} - \{i\}} \log(\delta_j^a + (1 - \delta_j^a)\delta_i^j) \tag{4}$$
$$\text{s.t. } x_i^c \geq 0 \ \forall c \in \mathcal{C}_i \text{ and } \tau_i \geq \sum_{i \in \mathcal{C}_i} x_i^c,$$

where $\delta_i^a := \exp(-x_i^a)$ and $\delta_i^j := \exp(-\lambda_j x_i^j)$. Since (4) is a concave problem, it can be solved using the Lagrangian method. Let $\eta_{x_i^c} \geq 0$ and $\gamma_i \geq 0$ denote the Lagrange multipliers associated with the constraints $x_i^c \geq 0$ and $\tau_i \geq \sum_{c \in \mathcal{C}_i} x_i^c$, respectively. The first-order conditions regarding x_i^a, x_i^b, and x_i^j, $j \in \mathcal{I} - \{i\}$, are

$$1 - \gamma_i + \eta_{x_i^a} = 0, \tag{$\text{FOC}_{x_i^a}$}$$
$$-\gamma_i + \eta_{x_i^b} = 0, \tag{$\text{FOC}_{x_i^b}$}$$
$$\text{and } \frac{\lambda_j (1 - \delta_j^a)\delta_i^j}{\delta_j^a + (1 - \delta_j^a)\delta_i^j} - \gamma_i + \eta_{x_i^j} = 0, \tag{$\text{FOC}_{x_i^j}$}$$

respectively. $\text{FOC}_{x_i^a}$ and $\text{FOC}_{x_i^b}$ together imply that $\gamma_i = \eta_{x_i^b} \geq 1 > 0$, and so $\sum_{c \in \mathcal{C}_i} x_i^c = \tau_i$ and $x_i^b = 0$. In words, player i must exhaust his bandwidth but ignore source b. The opposite is true for any type B player j, who must ignore source a, i.e., $x_j^a = 0$. Letting $\delta_j^a := \exp(-x_j^a) = 1$ in $\text{FOC}_{x_i^j}$ yields $\eta_{x_i^j} = \gamma_i > 0$, and so $x_i^j = 0$.

Step 3. Characterize the equilibrium attention network among type A players. Simplifying $\mathrm{FOC}_{x_i^j}$ shows that

$$x_i^j = \frac{1}{\lambda_j} \log \max\{\left(\frac{\lambda_j}{\gamma_i} - 1\right)(\exp(x_j^a) - 1), 1\} \; \forall i \in \mathcal{A} \text{ and } j \in \mathcal{A} - \{i\}, \tag{5}$$

where the term γ_i in the above expression satisfies $\gamma_i \geq 1$, and the inequality is strict if and only if $x_i^a = 0$ (as shown in Step 2). Simplifying (5) accordingly yields

$$x_i^a = [\tau_i - \sum_{j \in \mathcal{L} - \{i\}} \frac{1}{\lambda_j} \log \max\{(\lambda_j - 1)(\exp(x_j^a) - 1), 1\}]^+ \; \forall i \in \mathcal{A}. \tag{6}$$

Equations (5) and (6) together pin down all equilibria of the game among type A players. They can be further simplified when $x_i^a > 0 \; \forall i \in \mathcal{A}$, and when $x_i^j > 0 \; \forall i \in \mathcal{A}$ and $j \in \mathcal{A} - \{i\}$.

- In the first case, $\gamma_i \equiv 1 \; \forall i \in \mathcal{A}$, and so (5) becomes

$$x_i^j = \frac{1}{\lambda_j} \log \max\{(\lambda_j - 1)(\exp(x_j^a) - 1), 1\} \; \forall i \in \mathcal{A} \text{ and } j \in \mathcal{A} - \{i\}. \tag{7}$$

A close inspection of (7) reveals the equivalence between (a) $x_j^a > \phi(\lambda_j)$, (b) $x_i^j > 0$, and (c) $x_k^j \equiv h(x_j^a; \lambda_j) \; \forall k \in \mathcal{A} - \{j\}$, as in Part (i) of Theorem 2.

- In the second case, (5) and (6) become

$$x_i^j = h(x_j^a; \lambda_j) \; \forall i \in \mathcal{A} \text{ and } j \in \mathcal{A} - \{i\} \tag{8}$$

$$\text{and } x_i^a = \tau_i - \sum_{j \in \mathcal{A} - \{i\}} h(x_j^a; \lambda_j) \; \forall i \in \mathcal{A}, \text{ respectively.} \tag{9}$$

Simplifying $\mathbb{P}_x(\mathcal{U}_i \mid \omega = B)$ accordingly yields

$$\begin{aligned}
&\mathbb{P}_x(\mathcal{U}_i \mid \omega = B) \\
&= \delta_i^a \prod_{j \in \mathcal{A} - \{i\}} (\delta_j^a + (1 - \delta_j^a)\delta_i^j) \\
&= \exp(-x_i^a) \prod_{j \in \mathcal{A} - \{i\}} \exp(-x_j^a) + (1 - \exp(-x_j^a)) \exp(-\lambda_j \cdot \frac{1}{\lambda_j} \log(\lambda_j - 1)(\exp(x_j^a) - 1)) \\
&= \exp(-\sum_{j \in \mathcal{A}} x_j^a + \sum_{j \in \mathcal{A} - \{i\}} \phi(\lambda_j)),
\end{aligned}$$

as in Part (iii) of Theorem 2.

Step 4. Show that the game among type A players has an equilibrium. Recall from Step 3 that all equilibria of this game can be obtained by first solving (6) and then substituting the result(s) into (5). To show that (6) admits a solution, write $\{1, \cdots, N\}$ for $\mathcal{A}$. For each $\mathbf{x}^a := [x_1^a \cdots x_N^a]^\top \in \times_{i=1}^N [0, \tau_i]$, define $T(\mathbf{x}^a)$ as the N-vector whose i^{th} entry is given by the right-hand side of (6), and rewrite (6) as $T(\mathbf{x}^a) = \mathbf{x}^a$. Since $T : \times_{i=1}^N [0, \tau_i] \to \times_{i=1}^N [0, \tau_i]$ is a continuous mapping from a compact convex set to itself, it has a fixed point according to the Brouwer fixed point theorem. □

Proof of Theorem 1(ii). If $\tau \leq \phi(\lambda)$, then the game has a unique equilibrium in which players attend to their own-biased sources but nothing else. The remainder of the proof focuses on the more interesting case where $\tau > \phi(\lambda)$.

Our starting observation is that each player must attend only to a single primary source in any equilibrium. Thus in a symmetric equilibrium, either (i) all type A players attend to source a and their like-minded friends, and make decision A in event $\mathcal{U}_i$s; or (ii) all type A players attend to source b and their like-minded friends, and make decision B in event $\mathcal{U}_i$s.

Part (a): We proceed in two steps.

Step 1. Show that the game has a unique symmetric equilibrium of the first kind when N is large.

Let x_i^a and x_i^j denote the amounts of attention that a typical type A player i pays to source a and each like-minded friend of his, respectively. If $x_i^j = 0$, then $x_i^a = \tau$. But then $x_j^i = h(\tau, \lambda) \neq x_i^j$, which violates symmetry. As a result, $x_i^j > 0$ must hold, and x_i^a must solve $x = \tau - (N-1)h(x; \lambda) := \varphi^N(x)$ and so equals $x(N)$ by Lemma 4. Substituting this result into Parts (ii) and (iii) of Theorem 2(ii) yields $x_i^j = h(x(N); \lambda) > 0$, and $-\beta \exp(-Nx(N) + (N-1)\phi(\lambda))/2$ as the player's ex ante expected utility.

To sustain the above outcome on the equilibrium path, player i mustn't benefit from attending to source b and type B players and making decision B in event $\mathcal{U}_i$. In case player i commits such a deviation, solving his best response to type B players' equilibrium strategies yields $y_i^j = h(x(N); \lambda)$ as the amount of attention that he pays to each type B player, and $y_i^b = \tau - Nh(x(N); \lambda)$ as the amount of attention that he pays to source b. The last term is positive when N is large, because $y_i^b = \varphi^N(x(N)) - h(x(N);\lambda) = x(N) - h(x(N);\lambda) \to \phi(\lambda) > 0$ as $N \to \infty$ by Lemma 4. The ex ante expected utility generated by this best response function equals

$$-\frac{1}{2}\exp(-\tau + \frac{N(\tau - x(N))}{N-1} - Nx(N) + N\phi(\lambda)). \tag{10}$$

Comparing (10) with the on-path expected utility, we find that the former is smaller than the latter (and so the deviation is unprofitable) if and only if

$$\frac{\tau}{N-1} + \phi(\lambda) - \frac{N}{N-1}x(N) > \log\beta. \tag{11}$$

Since the left-hand side of (11) converges to zero as N grows to infinity by Lemma 4, it must exceed the right-hand side when N is sufficiently large.

Step 2. Show that no symmetric equilibrium of the second kind exists. In any equilibrium as such, any type A player can strictly benefit from attending to source a and $N-1$ type B players (who pay $x(N)$ unit of attention to source a) rather than adopting the equilibrium strategy.

Part (b): By Lemma 4, the left-hand side of (11), hereinafter denoted by LHS(τ), satisfies $\lim_{\tau\downarrow\phi(\lambda)}\text{LHS}(\tau) = 0$ and $\frac{d\text{LHS}(\tau)}{d\tau} = \frac{1}{N-1} - \frac{N}{N-1}\frac{dx(N)}{d\tau} \in (-1,0)$. Thus there exists $\underline{\tau} \in (\phi(\lambda), +\infty)$ such that (11) holds if and only if $\tau \in (\phi(\lambda), \underline{\tau})$. The remainder of the proof is the same as that of Part (i) and is therefore omitted for brevity. □

Proof of Theorem 3. Write $\{1, \cdots, N\}$ for $\mathcal{A}$. Since $x_i^a > \phi(\lambda_i)\ \forall i \in \mathcal{A}$ in equilibrium (as required by Assumption 1), $g_i := h_x(x_i^a;\lambda_i) \in (0,1)$ must hold by Lemma 3. Let $[\mathbf{G}_N]_{i,j} = g_j\ \forall j \in \{1,\cdots,N\}$ and $i \in \{1,\cdots,N\} - \{j\}$ in the marginal influence matrix. The corresponding matrix $\mathbf{A}_N := \mathbf{I}_N + \mathbf{G}_N$ satisfies the properties stated in Lemma 2.

Part (i): We only prove the result for τ_1. Under Assumption 1, $(x_i^a)_{i=1}^N$ and $(x_i^j)_{i,j}$ must solve (9) and (8) among type A players, respectively. Differentiating (9) with respect to τ_1 yields

$$\nabla_{\tau_1}\mathbf{x}^a = \mathbf{A}_N^{-1}\left[1\ 0\ \cdots\ 0\right]^\top,$$

where $\mathbf{x}^a \coloneqq [x_1^a \ \cdots \ x_N^a]^\top$. From Lemma 2, it follows that

$$\frac{\partial x_1^a}{\partial \tau_1} = \left[\mathbf{A}_N^{-1}\right]_{1,1} > 0 \text{ and } \frac{\partial x_j^a}{\partial \tau_1} = \left[\mathbf{A}_N^{-1}\right]_{j,1} < 0 \ \forall j \neq 1.$$

Substituting this result into (8) yields

$$\frac{\partial x_j^1}{\partial \tau_1} = g_1 \frac{\partial x_1^a}{\partial \tau_1} > 0 \text{ and } \frac{\partial x_k^j}{\partial \tau_1} = g_j \frac{\partial x_j^a}{\partial \tau_1} < 0 \ \forall j \neq 1 \text{ and } k \neq j.$$

Part (ii): We only prove the result for λ_1. Differentiating (9) with respect to λ_1 yields

$$\nabla_{\lambda_1} \mathbf{x}^a = \kappa \mathbf{A}_N^{-1} \left[0 \ 1 \ \cdots \ 1\right]^\top,$$

where $\kappa \coloneqq -h_\lambda(x_1^a; \lambda_1)$ has an ambiguous sign in general. From Lemma 2, it follows that

$$\operatorname{sgn}(\frac{\partial x_1^a}{\partial \lambda_1}) = \operatorname{sgn}(\kappa \underbrace{\sum_{i \neq 1} \left[\mathbf{A}_N^{-1}\right]_{1,i}}_{<0}) = \operatorname{sgn}\left(-\kappa\right)$$

and that

$$\operatorname{sgn}(\frac{\partial x_j^a}{\partial \lambda_1}) = \operatorname{sgn}(\kappa(\underbrace{\sum_{i=1}^{N} \left[\mathbf{A}_N^{-1}\right]_{j,i}}_{>0} - \underbrace{\left[\mathbf{A}_N^{-1}\right]_{j,1}}_{<0})) = \operatorname{sgn}\left(\kappa\right) \ \forall j \neq 1.$$

Substituting the second result into (8) yields

$$\operatorname{sgn}(\frac{\partial x_k^j}{\partial \lambda_1}) = \operatorname{sgn}(g_j \frac{\partial x_j^a}{\partial \lambda_1}) = \operatorname{sgn}\left(\kappa\right) \ \forall j \neq 1 \text{ and } k \neq j.$$

Finally, differentiating x_j^1 with respect to λ_1 yields

$$\operatorname{sgn}(\frac{\partial x_j^1}{\partial \lambda_1}) = \operatorname{sgn}(\kappa[g_1 \underbrace{\sum_{i \neq 1} [\mathbf{A}_N^{-1}]_{i,1}}_{<0} - 1]) = \operatorname{sgn}\left(-\kappa\right).$$

Thus in total, only three situations can happen, depending on whether κ is negative, positive, or zero. Specifically,

(a) if $\kappa < 0$, then $\partial x_1^a/\partial\lambda_1 > 0$, $\partial x_j^a/\partial\lambda_1 > 0$, $\partial x_j^a/\partial\lambda_1 < 0$, and $\partial x_k^j/\partial\lambda_1 < 0$ $\forall j \neq 1$ and $k \neq j$;

(b) if $\kappa > 0$, then reverse all the inequalities in case (a);

(c) if $\kappa = 0$, then replace the inequalities in case (a) with equalities.

Part (iii): Write $\overline{x}$ for $\sum_{i=1}^N x_i^a$. From

$$\frac{\partial \overline{x}}{\partial \tau_1} = [1\ 1\ \cdots\ 1]\,\nabla_{\tau_1}\mathbf{x}^a = \underbrace{\sum_{i=1}^N \left[\mathbf{A}_N^{-1}\right]_{i,1}}_{X}$$

$$\text{and } \frac{\partial \overline{x}}{\partial \lambda_1} = [1\ 1\ \cdots\ 1]\,\nabla_{\lambda_1}\mathbf{x}^a = \kappa\underbrace{(\sum_{i,j=1}^N \left[\mathbf{A}_N^{-1}\right]_{i,j} - \sum_{i=1}^N \left[\mathbf{A}_N^{-1}\right]_{i,1})}_{Y},$$

it follows that $\text{sgn}\left(\partial\overline{x}/\partial\tau_1\right) = \text{sgn}(X)$, and that $\text{sgn}\left(\partial\overline{x}/\partial\lambda_1\right) = \text{sgn}(\kappa) = \text{sgn}\left(-\partial x_1^a/\partial\lambda_1\right)$ if and only if $Y > 0$. It remains to show that $X, Y > 0$. For starters, notice that if the environment is symmetric and the game has a unique equilibrium (as required by Assumption 1), then the equilibrium and the corresponding matrix $\mathbf{A}_N$ must also be symmetric. Based on this fact, we can simplify X to $\sum_{i=1}^N \left[\mathbf{A}_N^{-1}\right]_{1,i}$ and Y to $(N-1)\sum_{i=1}^N \left[\mathbf{A}_N^{-1}\right]_{1,i}$. The last terms are positive by Lemma 2(iii). □

Proof of Proposition 1. We made three assumptions in the statement of Proposition 1: the environment is symmetric; the game has a unique equilibrium; and all players attend to each other in equilibrium. The first two assumptions imply that the equilibrium is symmetric. The last assumption implies that each player pays $x(N)$ units of attention to his own-biased source and $h(x(N);\lambda)$ units of attention to each like-minded friend of his (as shown in the proof of Theorem 1(ii)). Differentiating both sides of $x(N) = \varphi^N(x(N))$ with respect to N yields

$$\frac{dx(N)}{dN} = -\frac{1}{\lambda}\left[1 + \frac{N-1}{\lambda}\frac{\exp(x(N))}{\exp(x(N)) - 1}\right]^{-1} \log\left[(\lambda - 1)\left(\exp(x(N)) - 1\right)\right] < 0,$$

where the last inequality uses the fact that $x(N) > \phi(\lambda) := \log(\frac{\lambda}{\lambda-1})$. The term $Nx(N)$ is in general nonmonotonic in N, as is confirmed by numerical methods. □

B Figures

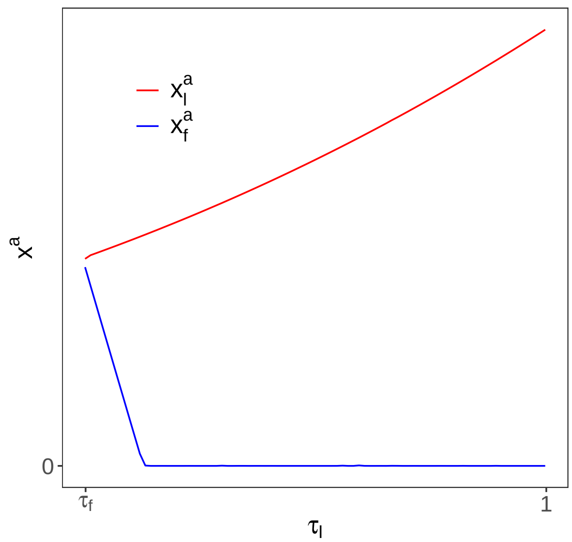
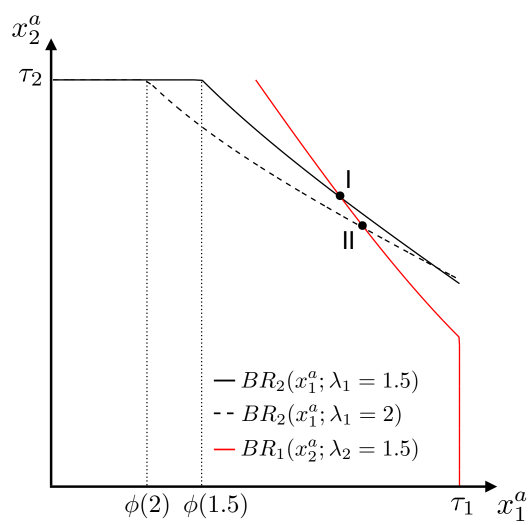
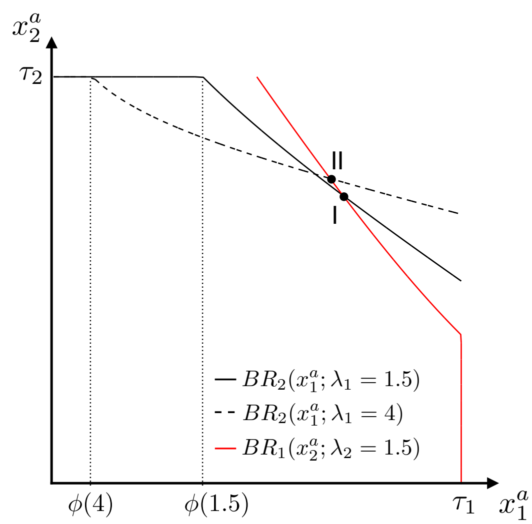


Figure 1: τ_l and τ_f denote the bandwidths of opinion leaders and followers; x_l^a and x_f^a denote their levels of informedness in equilibrium: $\tau_f = 0.16$, $\lambda_l = \lambda_f = 10$, the numbers of opinion leaders and followers are 10 and 90, respectively.



(a) Increase λ_1 from 1.5 to 2.



(b) Increase λ_1 from 1.5 to 4.

Figure 2: The red and black curves represent player 1 and player 2's best response functions, respectively: $\lambda_2 = 1.5$, $\tau_1 = \tau_2 = 3$.

References

Acemoglu, D., A. Ozdaglar, and J. Siderius (2022): "A model of online misinformation," *Working paper.*

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

Baccara, M., and L. Yariv (2013): "Homophily in peer groups," *American Economic Journal: Microeconomics*, 5(3), 69–96.

Bakshy, E., S. Messing, and L. A. Adamic (2015): "Exposure to ideologically diverse news and opinion on Facebook," *Science*, 348(6239), 1130–1132.

Bala, V., and S. Goyal (2000): "A noncooperative model of network formation," *Econometrica*, 68(5), 1181–1229.

Balsamo, D., V. Gelardi, C. Han, D. Rama, A. Samantray, C. Zucca, and M. Starnini (2019): "Inside the echo chamber: Disentangling network dynamics from polarization," *arXiv preprint arXiv:1906.09076.*

Barberá, P. (2020): "Social media, echo chambers, and political polarization," in *Social Media and Democracy: The State of the Field, Prospects for Reform*, ed. by N. Persily, and J. A. Tucker. Cambrige University Press.

Baum, M. A., and A. S. Jamison (2006): "The Oprah effect: How soft news helps inattentive citizens vote consistently," *Journal of Politics*, 68(4), 946–959.

Baumann, L. (2021): "A model of weighted network formation," *Theoretical Economics*, 16(1), 1–23.

Bloch, F., and B. Dutta (2009): "Communication networks with endogenous link strength," *Games and Economic Behavior*, 66(1), 39–56.

Bowen, T. R., D. Dmitriev, and S. Galperti (2023): "Learning from shared news: When abundant information leads to belief polarization," *Quarterly Journal of Economics*, 138(2), 955–1000.

Bramoullé, Y., R. Kranton, and M. D'Amours (2014): "Strategic interaction and networks," *American Economic Review*, 104(3), 898–930.

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

Calvó-Armengol, A., J. de Martí, and A. Prat (2015): "Communication and influence," *Theoretical Economics*, 10(2), 649–690.

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

Chiang, C.-F., and B. Knight (2011): "Media bias and influence: Evidence from newspaper endorsements," *Review of Economic Studies*, 78(3), 795–820.

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, pp. 130–140.

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(3), 554–559.

DellaVigna, S., and M. Gentzkow (2010): "Persuasion: empirical evidence," *Annual Review of Economics*, 2(1), 643–669.

Denti, T. (2017): "Network effects in information acquisition," *Working paper.*

——— (forthcoming): "Unrestricted information acquisition," *Theoretical Economics.*

Dessein, W., A. Galeotti, and T. Santos (2016): "Rational inattention and organizational focus," *American Economic Review*, 106(6), 1522–1536.

Dewey, C. (2016): "6 in 10 of you will share this link without reading it, a new, depressing study says," *The Washington Post, June 16.*

Elliott, M., B. Golub, and M. V. Leduc (2022): "Supply network formation and fragility," *American Economic Review*, 112(8), 2701–2747.

Farrell, H., and D. W. Drezner (2008): "The power and politics of blogs," *Public Choice*, 134(1), 15–30.

Flaxman, S., S. Goel, and J. M. Rao (2016): "Filter bubbles, echo chambers, and online news consumption," *Public Opinion Quarterly*, 80(S1), 298–320.

Funk, C., J. Gottfried, and A. Mitchell (2017): "Science news and information today," *Pew Research Center, September 20.*

Galeotti, A., C. Ghiglino, and F. Squintani (2013): "Strategic information transmission networks," *Journal of Economic Theory*, 148(5), 1751–1769.

Galeotti, A., and S. Goyal (2010): "The law of the few," *American Economic Review*, 100(4), 1468–1492.

Hébert, B. M., and J. La'O (2021): "Information acquisition, efficiency, and non-fundamental volatility," *Working paper*.

Hellwig, C., and L. Veldkamp (2009): "Knowing what others know: Coordination motives in information acquisition," *Review of Economic Studies*, 76(1), 223–251.

Herskovic, B., and J. Ramos (2020): "Acquiring information through peers," *American Economic Review*, 110(7), 2128–2152.

Ho, D. E., and K. M. Quinn (2008): "Viewpoint diversity and media consolidation: An empirical study," *Stanford Law Review*, 61, 781.

Hu, L., A. Li, and I. Segal (forthcoming): "The politics of personalized news aggregation," *Journal of Political Economy Microeconomics*.

Innocenti, F. (2021): "Can media pluralism be harmful to news quality?," *Working paper*.

Jackson, M. O., and A. Wolinsky (1996): "A strategic model of social and economic networks," *Journal of Economic Theory*, 71(1), 44–74.

Jackson, M. O., and Y. Zenou (2015): "Games on networks," in *Handbook of Game Theory with Economic Applications*, vol. 4, pp. 95–163. Elsevier.

Jann, O., and C. Schottmüller (2021): "Why are echo chambers useful," *Working paper*.

Katz, E., and P. F. Lazarsfeld (1955): *Personal Influence: The Part Played by People in the Flow of Mass Communications*. New York, NY: The Free Press.

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

Lu, Y., P. Zhang, Y. Cao, Y. Hu, and L. Guo (2014): "On the frequency distribution of retweets," *Procedia Computer Science*, 31, 747–753.

Maćkowiak, B., F. Matějka, and M. Wiederholt (2023): "Rational inattention: A review," *Journal of Economic Literature*, 61(1), 226–273.

Meng, D. (2021): "Learning from like-minded people," *Games and Economic Behavior*, 126, 231–250.

Néda, Z., L. Varga, and T. S. Biró (2017): "Science and Facebook: The same popularity law!," *PLoS One*, 12(7), e0179656.

Parise, F., and A. Ozdaglar (2019): "A variational inequality framework for network games: Existence, uniqueness, convergence and sensitivity analysis," *Games and Economic Behavior*, 114, 47–82.

Pariser, E. (2011): *The Filter Bubble: How the New Personalized Web is Changing What We Read and How We Think*. New York, NY: Penguin Press.

Perego, J., and S. Yuksel (2016): "Searching for information and the diffusion of knowledge," *Working paper.*

Prat, A., and D. Strömberg (2013): "The political economy of mass media," in *Advances in Economics and Econometrics: Theory and Applications, Tenth World Congress*, ed. by D. Acemoglu, M. Arellano, and E. Dekel. Cambrige University Press.

Prior, M. (2007): *Post-Broadcast Democracy: How Media Choice Increases Inequality in Political Involvement and Polarizes Elections.* New York, NY: Cambridge University Press.

Romm, T. (2021): "Facebook, Twitter could face punishing regulation for their role in U.S. Capitol riot, Democrats say," *The Washington Post, January 8.*

Sadler, E., and B. Golub (2021): "Games on endogenous networks," *arXiv preprint arXiv:2102.01587.*

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

Sherman, J., and W. J. Morrison (1950): "Adjustment of an inverse matrix corresponding to a change in one element of a given matrix," *Annals of Mathematical Statistics*, 21(1), 124–127.

Suen, W. (2004): "The self-perpetuation of biased beliefs," *Economic Journal*, 114(495), 377–396.

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

Williams, C. (2021): "Echo chambers: Social learning under unobserved heterogeneity," *Working paper.*

Zhong, W. (2022): "Optimal dynamic information acquisition," *Econometrica*, 90(4), 1537–1582.

Online Appendix for "A Rational Inattention Theory of Echo Chamber"

by Lin Hu, Anqi Li, and Xu Tan

O.1 Efficient attention network

In this appendix, we examine the efficient attention network that maximizes the utilitarian welfare of a symmetric society, and highlight its main qualitative difference from the equilibrium attention network.

We focus on the case where players have strong horizontal preferences, and so making one's default decision is efficient in event $\mathcal{U}_i$. The efficient attention network solves

$$\max_{x \in \times_{i \in \mathcal{I}} \mathcal{X}_i} -\frac{\beta}{2} \sum_{i \in \mathcal{I}} \mathbb{P}_x(\mathcal{U}_i \mid \omega \neq d_i).$$

The next theorem shows that when players are good at absorbing and disseminating information (as captured by a large bandwidth τ and a high visibility parameter λ), they are mandated to attend to both primary sources and to each other under the efficient attention network. Doing so qualifies all players as secondary sources and so facilitates information transmission between them, although the outcome cannot be sustained in any equilibrium. The reason is that in our model, a self-interested individual would only gather information about state A or B, but not both. Mandating that players attend to both primary sources is socially beneficial but individually wasteful.

Theorem O1. *Consider a symmetric society parameterized by* $(N, \beta, \lambda, \tau)$. *For each* $N \in \mathbb{N} - \{1\}$, *there exist* $\underline{\beta}$, $\underline{\lambda}$, *and* $\underline{\tau} > 0$ *such that when* $\beta < \underline{\beta}$, $\lambda > \underline{\lambda}$, *and* $\tau > \underline{\tau}$, *the efficient attention network features* $x_i^a, x_i^b, x_i^j > 0$ *for all* i, j, *and so cannot be sustained in any equilibrium.*

Proof. Omitted proofs from this appendix are gathered in subsection O.7. □

O.2 General primary sources

This appendix extends the baseline model to a finite set $\mathcal{S}$ of primary sources. In state $\omega \in \Omega := \{A, B\}$, source $s \in \mathcal{S}$ disseminates message "ω" to player i at Poisson rate $\lambda_i^s(\omega) \geq 0$. The baseline model is a special case of this general framework, where

$$\mathcal{S} = \{\omega\text{-}revealing : \omega \in \Omega\}, \text{ and } \lambda_i^{\omega'\text{-}revealing}(\omega) = \begin{cases} 1 \text{ if } \omega = \omega' \\ 0 \text{ else} \end{cases} \quad \forall i \in \mathcal{I} \text{ and } \omega' \in \Omega.$$

Other special cases include, but are not limited to the following.

Example O1. If $\forall i \in \mathcal{I}$, $\lambda_i^s(\omega) = \begin{cases} \lambda^s > 0 & \text{if } \omega = \omega' \in \Omega, \\ 0 & \text{else,} \end{cases}$ then s is a *public source* that specializes in revealing state ω' to all players at rate λ^s. ♢

Example O2. If $\exists i' \in \mathcal{I}$ and $\omega' \in \Omega$ such that $\lambda_i^s(\omega) = \begin{cases} \lambda^s > 0 & \text{if } i = i' \text{ and } \omega = \omega', \\ 0 & \text{else,} \end{cases}$ then s is a *private source* that specializes in revealing state ω' to player i'. ♢

Example O3. If $\lambda_i^s(\omega) = 1$ $\forall i \in \mathcal{I}$ and $\omega \in \Omega$, then s is a *mega source* that reveals both states to all players at rate 1, and it can be obtained from merging the biased primary sources in the baseline model together. ♢

For each $i \in \mathcal{I}$ and $\omega \in \Omega$, define $\overline{\lambda}_i(\omega) \coloneqq \max\{\lambda_i^s(\omega) : s \in \mathcal{S}\}$ as the highest rate at which state ω is revealed to player i, and assume that $\overline{\lambda}_i(\omega) > 0$ to make the analysis interesting. Define $\mathcal{S}_i(\omega) \coloneqq \{s : \lambda_i^s(\omega) = \overline{\lambda}_i(\omega)\}$ and $\mathcal{S}_i \coloneqq \cup_{\omega \in \Omega} \mathcal{S}_i(\omega)$. Our starting observation is that in equilibrium, each player i attends only to the sources in $\mathcal{S}_i$, and only the total amount of attention that he pays to these sources matters.

Lemma O1. *$\forall i \in \mathcal{I}$ and $s \in \mathcal{S}$, $x_i^s = 0$ if $s \in \mathcal{S} \backslash \mathcal{S}_i$ in any equilibrium. Moreover, if $((x_i^s)_{s \in \mathcal{S}_i}, (x_i^j)_{j \in \mathcal{I}-\{i\}})_{i \in \mathcal{I}}$ is an equilibrium strategy profile, then so is every strategy profile $((y_i^s)_{s \in \mathcal{S}_i}, (y_i^j)_{j \in \mathcal{I}-\{i\}})_{i \in \mathcal{I}}$ that satisfies $\sum_{s \in \mathcal{S}_i(\omega)} y_i^s = \sum_{i \in \mathcal{S}_i(\omega)} x_i^s$ $\forall i \in \mathcal{I}$ and $\omega \in \Omega$, as well as $y_i^j = x_i^j$ $\forall i \in \mathcal{I}$ and $j \in \mathcal{I} - \{i\}$.*

The remainder of this appendix examines two special cases: specialized sources and mega source. As in Examples O1 and O2, we say that primary sources are *specialized* if each of them reveals at most one state at a positive rate to each player, i.e., $\forall i \in \mathcal{I}$ and $s \in \mathcal{S}$, $|\{\omega : \lambda_i^s(\omega) > 0\}| \leq 1$. The next proposition extends the baseline model to encompass specialized sources.

Proposition O1. *When sources are specialized, (i) it is w.l.o.g. to assume that each player faces two primary sources, each revealing a distinct state $\omega \in \Omega$ to him at rate $\overline{\lambda}_i(\omega)$. (ii) In the case where $\overline{\lambda}_i(\omega) \equiv \nu > 0$ $\forall i \in \mathcal{I}$ and $\omega \in \Omega$, Theorems 1 and 2 remain valid after we replace x_i^c, λ_i, and τ_i with νx_i^c, λ_i/ν, and $\nu\tau_i$, respectively, $\forall i \in \mathcal{I}$ and $c \in \mathcal{C}_i$.*

Part (i) of Proposition O1 shows that introducing multiple (public or private) specialized sources of the same quality into the baseline model does not impact the equilibrium attention network in any meaningful way. All it does is to dilute players' attention across the same kind of sources.

When $\overline{\lambda}_i(\omega) \neq 1$, we must rescale player's bandwidths and visibility parameters properly to make the equilibrium characterization work. Part (ii) of Proposition O1 examines the consequences of applying a common shock ν to the visibility of the primary sources. Among other things, we find that increasing the visibility of primary sources effectively diminishes that of secondary sources. The equilibrium and welfare consequences of this change are in general ambiguous by Theorem 3.

Consider next the case of a mega source as in Example O3. We denote this source by m as in the main text, and normalize its visibility parameter to 1 for simplicity. The next proposition establishes the isomorphism between two interesting games.

Proposition O2. *Let everything be as in the baseline model except that sources a and b are merged into m. If $(x_i^m, (x_i^j)_{j \in \mathcal{I}-\{i\}})_{i \in \mathcal{I}}$ is an equilibrium of this augmented game, then the strategy profile $(y_i^a, (y_i^j)_{j \in \mathcal{I}-\{i\}})_{i \in \mathcal{I}}$ with $y_i^a = x_i^m$ and $y_i^j = x_i^j$ $\forall i \in \mathcal{I}$ and $j \in \mathcal{I} - \{i\}$ is an equilibrium of the game among a set $\mathcal{I}$ of type A players with characteristics $(\beta_i, \lambda_i, \tau_i)$s and access to source a. Moreover, the converse is also true, and players' expected utilities are the same under the two equilibria.*

We demonstrate the usefulness of Proposition O2 in a symmetric society. There, merging sources a and b into m is mathematically equivalent to doubling the size N of the echo chamber among type A players in Proposition 1. From the previous discussion, we know that the welfare consequence of this change is in general ambiguous.

O.3 Finite decision problems

This appendix extends the baseline model to decision problems with more than two states and actions. Suppose that the state ω is distributed uniformly on a finite set $\{1, \cdots, M\}$ with $M \in \mathbb{N} - \{1\}$. There are M types of players, each has a population $N \in \mathbb{N} - \{1\}$ and can make one of the decisions in $\{1, \cdots, M\}$. In case a type m player makes decision d, his utility in state ω equals zero if $d = \omega$, -1 if $\omega = m$ and $d \neq m$, and $-\beta$ if $\omega \neq m$ and $d = m$. Assume that $\beta \in (0, 1)$, and so m is the default decision of type m players. Also assume that all players share the same visibility parameter $\lambda > 0$ and bandwidth $\tau > 0$.

There are M primary sources called 1-*revealing*, $\cdots$, M-*revealing*. In state $\omega \in \{1, \cdots, M\}$, the ω-revealing source announces a message "ω," whereas the other sources are silent. To make informed decisions, players attend to the primary sources and to each other as potential secondary sources.

We analyze the symmetric PSPBE of the game. An equilibrium as such can be parameterized by four quantities: Δ^*, x^*, y^*, and z^*. For a type m player:

(i) Δ^* denotes the amount of attention that he pays to the m-revealing source;

(ii) x^* denotes the amount of attention that he pays to each other primary source;

(iii) y^* denotes the amount of attention that he pays to each like-minded friend of his;

(iv) z^* denotes the amount of attention that he pays to any other player.

Call an equilibrium a *semi echo chamber equilibrium* if $\Delta^* = 0$ and $y^* > z^*$. That is, no player wastes time on learning the state that favors his default decision, and all players prioritize like-minded friends over the other players when deciding whom to pay attention to.

The next theorem proves an analog of Theorem 1: when players are sufficiently biased towards their default decisions, the unique symmetric PSPBE of the game must be a semi echo chamber equilibrium.

Theorem O2. *For any $M, N \in \mathbb{N} - \{1\}$, $\lambda > 1/(M-1)$ and $\tau > (M-1)\phi(\lambda(M-1))$, there exist $\underline{\beta} \in (0,1)$ such that for any $\beta < \underline{\beta}$, the unique PSPBE of the game must be a semi echo chamber equilibrium.*

O.4 Uniqueness of equilibrium

This appendix provides sufficient conditions for the game among type A players to admit a unique equilibrium. Let the set $\mathcal{PV} := \{i \in \mathcal{A} : \tau_i > \phi(\lambda_i)\}$ gather those players who are potentially visible to their like-minded friends in equilibrium. The next observation is immediate.

Observation O1. *The game among type A players has a unique equilibrium if and only if the system (6) of equations among $\mathcal{PV}$ players has a unique solution.*

Proof. All equilibria of the game among type A players can be obtained as follows.

Step 1. Solve (6) among $\mathcal{PV}$ players. For each solution $(x_i^a)_{i\in\mathcal{PV}}$, define $\mathcal{COR} = \{i \in \mathcal{PV} : x_i^a > \phi(\lambda_i)\}$ and $\mathcal{PER} = \mathcal{A} - \mathcal{COR}$.

Step 2. For each pair $i, j \in \mathcal{COR}$, let $x_i^j = h(x_j^a; \lambda_j)$. For each pair $i \in \mathcal{A}$ and $j \in \mathcal{PER}$, let $x_i^j = 0$. For each pair $i \in \mathcal{PER}$ and $j \in \mathcal{COR}$, let $x_i^j = \frac{1}{\lambda_j} \log \max\{\left(\frac{\lambda_j}{\gamma_i} - 1\right)\left(\exp(x_j^a) - 1\right), 1\}$ and $x_i^a = \tau_i - \sum_{j\in\mathcal{COR}} x_i^j$, where $\gamma_i \geq 1$ is the Lagrange multiplier associated with the constraint $x_i^a \geq 0$, and $\gamma_i > 1$ if and only if $x_i^a = 0$. □

Observation O1 implies that the game among type A players has a unique equilibrium if $|\mathcal{PV}| = 1$. The remainder of this appendix assumes that $|\mathcal{PV}| \geq 2$. The analysis differs, depending on whether $\mathcal{PV}$ players are homogeneous or not. In the first case, equilibrium is unique when $\mathcal{PV}$ players' bandwidth is large relative to their population size and, roughly speaking, when they have a high visibility parameter.

Theorem O3. *In the case where $(\lambda_i, \tau_i) \equiv (\lambda, \tau)$ $\forall i \in \mathcal{PV}$, the game among type A players has a unique equilibrium if $\tau - (|\mathcal{PV}| - 2)h(\tau; \lambda) > \phi(\lambda)$.*

When $\mathcal{PV}$ players are heterogeneous, we cannot establish the uniqueness of equilibrium in the setting of the baseline model. The reason is that, when $x_i^a \approx \phi(\lambda_i)$, the marginal influence $h_x(x_i^a; \lambda_i)$ of player i on the other players equals approximately 1 (recall Lemma 3), which is too big for the contraction mapping theorem to work. To bound players' marginal influences on each other, we enrich the baseline model by assuming that each player i has $\overline{\tau}_i > 0$ units of attention to spare and yet must pay at least $\underline{\tau}_i \in [0, \overline{\tau}_i)$ units of attention to his own-biased source. If $\underline{\tau}_i \equiv 0$ $\forall i \in \mathcal{I}$, then we are back to the baseline model. The next proposition establishes the analog of Theorem 2 in this new setting. The fact that Theorems 1 is unaffected by the change is easy to see.

Proposition O3. *Let everything be as above. Then the following are true for any $i \in \mathcal{A}$ in any echo chamber equilibrium.*

(i) If $x_j^a > \underline{\tau}_j$ $\forall j \in \mathcal{A}$, then Part (i) of Theorem 2 remains valid.

(ii) $$x_i^a = \max\{\overline{\tau}_i - \sum_{j\in\mathcal{A}-\{i\}} \underbrace{\frac{1}{\lambda_j} \log\max\left\{(\lambda_j - 1)(\exp(x_j^a) - 1), 1\right\}}_{=x_i^j \text{ if } x_i^a > \underline{\tau}_i}, \underline{\tau}_i\}.$$

(iii) If $x_j^k > 0$ $\forall j \in \mathcal{A}$ and $k \in \mathcal{A} - \{j\}$, then Part (iii) of Theorem 2 remains valid.

Proof. The proof is analogous to that of Theorem 2 and is omitted for brevity. $\square$

We provide sufficient conditions for the augmented game among type A players to admit a unique equilibrium. Redefine $\mathcal{PV} \coloneqq \{i \in \mathcal{A} : \overline{\tau}_i > \phi(\lambda_i)\}$. Suppose that $\underline{\tau}_i > \phi(\lambda_i)\ \forall i \in \mathcal{PV}$, and define $\overline{g} \coloneqq \max_{i \in \mathcal{PV}} h_x(\underline{\tau}_i; \lambda_i)$. By Lemma 3, $\overline{g} < 1$ is a uniform upper bound for the marginal influences that $\mathcal{PV}$ players can exert on each other. The next theorem shows that when $\overline{g}$ is small relative to the population size of $\mathcal{PV}$ players, the augmented game among type A players has a unique equilibrium.

Theorem O4. *Let everything be as above. Then the game among type A players has a unique equilibrium if* $\overline{g} < 1/(|\mathcal{PV}| - 1)$.

O.5 Comparative statics with peripheral players

When proving Theorem 3, we assumed, for convenience, that all type A players must attend to each other, or, equivalently, $\mathcal{A} = \mathcal{COR}$. We now relax this assumption as follows.

Assumption O1. *The game among type A players has a unique equilibrium, whereby all players attend to source a,* $|\mathcal{COR}| \geq 2$ *(to make the analysis interesting), and no* $\mathcal{PER}$ *player is a borderline player, i.e.,* $x_i^a < \phi(\lambda_i)\ \forall i \in \mathcal{PER}$.

It is clear that perturbing the characteristics of a $\mathcal{PER}$ player has no impact on any other player under Assumption O1. The next proposition examines the effect on $\mathcal{PER}$ players as we perturb the characteristics of a $\mathcal{COR}$ player.

Proposition O4. *Let everything be as in Theorem 3 except that Assumption 1 is replaced with Assumption O1. At any* $\boldsymbol{\theta}^\circ \in int(\Theta)$*, the following are true for any* $i, j \in \mathcal{COR}$ *(*$i = j$ *is allowed) and any* $k \in \mathcal{PER}$.

(i) $\mathrm{sgn}\left(\frac{\partial x_k^a}{\partial \tau_i}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ}\right) = \mathrm{sgn}\left(-\frac{\partial x_i^a}{\partial \tau_i}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ}\right)$ *and* $\mathrm{sgn}\left(\frac{\partial x_k^j}{\partial \tau_i}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ}\right) = \mathrm{sgn}\left(\frac{\partial x_j^a}{\partial \tau_i}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ}\right)$.

(ii) $\mathrm{sgn}\left(\frac{\partial x_k^a}{\partial \lambda_i}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ}\right) = \mathrm{sgn}\left(-\frac{\partial x_i^a}{\partial \lambda_i}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ}\right)$ *and* $\mathrm{sgn}\left(\frac{\partial x_k^j}{\partial \lambda_i}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ}\right) = \mathrm{sgn}\left(\frac{\partial x_j^a}{\partial \lambda_i}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ}\right)$.

As we increase the bandwidth of a $\mathcal{COR}$ player i, a $\mathcal{PER}$ player k pays less attention to the primary source, more attention to player i, and less attention to any other $\mathcal{COR}$ player than i. As we increase the visibility parameter of player i, the effect on player k depends on whether the perturbation makes player i an opinion leader or an opinion follower. In the first case, player k pays less attention to the primary source, more attention to player i, and less attention to any other $\mathcal{COR}$ player than i. The opposite happens in the second case.

O.6 Pairwise visibility parameter

This appendix extends the baseline model to encompass pairwise visibility parameters. Let $\lambda_i^j \geq 0$ parameterize the visibility of player j to player i, and write $\boldsymbol{\lambda}_i$ for $(\lambda_i^j)_{j \in \mathcal{I}-\{i\}}$. The next proposition establishes the analog of Theorem 2 in this new setting. The fact that Theorem 1 is unaffected by the change is easy to see.

Proposition O5. *The following are true for any $i \in \mathcal{A}$ in any echo chamber equilibrium.*

(i) If all type A players attend to source a, then the following are equivalent: (a) $x_j^a > \phi(\lambda_i^j)$; (b) $x_i^j > 0$; (c) $x_i^j = h(x_j^a; \lambda_i^j)$.

(ii) $x_i^a = [\tau_i - \sum_{j \in \mathcal{A}-\{i\}} \underbrace{\frac{1}{\lambda_i^j} \log \max\left\{(\lambda_i^j - 1)(\exp(x_j^a) - 1), 1\right\}}_{= x_i^j \text{ if } x_i^a > 0}]^+.$

(iii) If all type A players attend to each other, then the ex ante expected utility of player i equals

$$-\frac{\beta_i}{2} \exp(-\sum_{j \in \mathcal{A}} x_j^a + \sum_{j \in \mathcal{A}-\{i\}} \phi(\lambda_i^j)).$$

Proof. The proof is analogous to that of Theorem 2 and is therefore omitted. □

With pairwise visibility parameters, player j must cross a personalized threshold $\phi(\lambda_i^j)$ in order to be attended by player i. After that, the amount of influence $h(x_j^a; \lambda_i^j)$ that he exerts on player i depends on his informedness x_j^a as a secondary source, as well as his visibility λ_i^j to player i. Player i's equilibrium expected utility depends positively on the total amount of attention that the entire echo chamber pays to the primary source, and negatively on the visibility threshold $\phi(\lambda_i^j)$s that prevent his like-minded friends from disseminating information to him.

Turning to equilibrium comparative statics, we write $\mathcal{A} = \{1, \cdots, N\}$, $\theta_i = (\boldsymbol{\lambda}_i, \tau_i)$ $\forall i \in \mathcal{A}$, and $\boldsymbol{\theta} = [\theta_1 \ \cdots \ \theta_N]^\top$. Note that with pairwise visibility parameters, the off-diagonal entries of the marginal influence matrix are no longer constant column by column. Nevertheless, if that matrix still satisfies the properties stated in Lemma 2, then we can establish an analog of Theorem 3 as follows.

Proposition O6. *Fix any $N \in \mathbb{N} - \{1\}$. Let Θ be any neighborhood in $\mathbb{R}^{N^2}_{++}$ such that for any $\boldsymbol{\theta} \in \Theta$, the game among a set $\mathcal{A}$ of type A players with population size N and characteristic profile $\boldsymbol{\theta}$ satisfies Assumption 1, and the matrix $\mathbf{A}_N := \mathbf{I}_N + \mathbf{G}_N$ satisfies the properties stated in Lemma 2. Then at any $\boldsymbol{\theta}^\circ \in int(\Theta)$, the following must hold for any $i \in \{1, \cdots, N\}$, $j, k \in \{1, \cdots, N\} - \{i\}$ ($j = k$ is allowed), and $m \in \{1, \cdots, N\} - \{k\}$.*

(i) $\partial x_i^a / \partial \tau_i|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ} > 0$, $\partial x_k^i / \partial \tau_i|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ} > 0$, $\partial x_k^a / \partial \tau_i|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ} < 0$, *and* $\partial x_m^k / \partial \tau_i\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ} < 0$.

(ii) One of the following situations happens:

(a) $\partial x_i^a / \partial \lambda_i^j\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ} > 0$, $\partial x_k^i / \partial \lambda_i^j\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ} > 0$, $\partial x_k^a / \partial \lambda_i^j\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ} < 0$, *and* $\partial x_m^k / \partial \lambda_i^j\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^\circ} < 0$*;*

(b) the inequalities in Part (a) are all reversed;

(c) the inequalities in Part (a) are replaced with equalities.

O.7 Proofs

Proof of Theorem O1. When β is sufficiently small, it is efficient to make one's default decision in event $\mathcal{U}_i$. The social planner's problem then becomes

$$\max_{x \in \times_{i \in \mathcal{I}} \mathcal{X}_i} - \sum_{i \in \mathcal{A}} \delta_i^a \prod_{j \in \mathcal{I} - \{i\}} \left(\delta_j^a + \left(1 - \delta_j^a\right) \delta_i^j \right) - \sum_{i \in \mathcal{B}} \delta_i^b \prod_{j \in \mathcal{I} - \{i\}} \left(\delta_j^b + \left(1 - \delta_j^b\right) \delta_i^j \right).$$

Since the above problem has a strictly concave maximand and a compact convex choice set, it has a unique solution. In case the solution is interior, it is fully determined by the first-order conditions. If, in addition, it is symmetric, then it is parameterized by

(i) $x^* > 0$: the amount of attention that a typical player pays to his own-biased source;

(ii) $y^* > 0$: the amount of attention he pays to the other source;

(iii) $z^* > 0$: the amount of attention he pays to each like-minded friend of his;

(iv) $\Delta^* > 0$: the amount of attention he pays to any other player.

The corresponding attention network cannot arise in any equilibrium because $y^* > 0$.

We provide conditions under which the solution to the planner's problem is interior and symmetric. For ease of notation, write X for $\exp(-x^*)+(1-\exp(-x^*))\exp(-\lambda z^*)$, Y for $\exp(-y^*)+(1-\exp(-y^*))\exp(-\lambda\Delta^*)$, $\tilde{a}$ for $\exp(x^*)-1$, $\tilde{b}$ for $\exp(y^*)-1$, $\tilde{c}$ for $\exp(\lambda z^*)-1$, and $\tilde{d}$ for $\exp(\lambda\Delta^*)-1$. Fix any type A player (name him i), and let $\gamma > 0$ denote the Lagrange multiplier associated with his bandwidth constraint, which must be binding under the efficient allocation. The first-order conditions regarding x_i^a, x_i^b, x_i^j with $j \in \mathcal{A}$, and x_i^k with $k \in \mathcal{B}$, are

$$\delta_i^a X^{N-1}Y^N + \sum_{j\in\mathcal{A}-\{i\}} \delta_i^a\delta_j^a(1-\delta_j^i)X^{N-2}Y^N = \gamma \tag{FOC$_{x_i^a}$}$$

$$\sum_{j\in\mathcal{B}} \delta_i^b\delta_j^b(1-\delta_j^i)X^{N-1}Y^{N-1} = \gamma \tag{FOC$_{x_i^b}$}$$

$$\lambda\delta_i^a(1-\delta_j^a)\delta_i^j X^{N-2}Y^N = \gamma \tag{FOC$_{x_i^j}$}$$

$$\text{and } \lambda\delta_i^a(1-\delta_k^a)\delta_i^k X^{N-1}Y^{N-1} = \gamma, \tag{FOC$_{x_i^k}$}$$

respectively. Letting $x_i^a = x^*$, $x_i^b = y^*$, $x_i^j = z^*$, and $x_i^k = \Delta^*$ in the FOCs yields

$$\begin{cases} (\lambda-1)\tilde{a} = N\tilde{c}+1 \\ N\tilde{d} = \lambda\tilde{b} \\ \lambda\tilde{a}(\tilde{b}+\tilde{d}+1) = N\tilde{d}(\tilde{a}+\tilde{c}+1) \\ \log(\tilde{a}+1)+\log(\tilde{b}+1)+\frac{N-1}{\lambda}\log(\tilde{c}+1)+\frac{N}{\lambda}\log(\tilde{d}+1) = \tau. \end{cases}$$

The solution to the first three linear equations is

$$\tilde{b} = \frac{N\tilde{a}}{N-1-\tilde{a}},\ \tilde{c} = \frac{(\lambda-1)\tilde{a}-1}{N},\ \text{and } \tilde{d} = \frac{\lambda\tilde{a}}{N-1-\tilde{a}}.$$

Simplifying the last equation accordingly yields

$$\begin{aligned} \log(\tilde{a}+1)+\log\Big(\frac{N\tilde{a}}{N-1-\tilde{a}}+1\Big)+\frac{N-1}{\lambda}\log\Big(\frac{(\lambda-1)\tilde{a}-1}{N}+1\Big) \\ +\frac{N}{\lambda}\log\Big(\frac{\lambda\tilde{a}}{N-1-\tilde{a}}+1\Big) = \tau. \end{aligned} \tag{12}$$

It remains to find conditions on (λ,τ,N) such that (12) admits a solution $\tilde{a}(\lambda,\tau,N)$

satisfying

$$\tilde{a}(\cdot) > 0, \ \frac{N\tilde{a}(\cdot)}{N-1-\tilde{a}(\cdot)} > 0, \ \frac{(\lambda-1)\tilde{a}(\cdot)-1}{N} > 0, \text{ and } \frac{\lambda\tilde{a}(\cdot)}{N-1-\tilde{a}(\cdot)} > 0,$$

or, equivalently,

$$\lambda > \frac{N}{N-1} \text{ and } \tilde{a}(\cdot) \in (\frac{1}{\lambda-1}, N-1).$$

To make progress, notice that the left-hand side of (12) as a function of $\tilde{a}$ (i) is well-defined on $(0, N-1)$, (ii) is negative when $\tilde{a} \approx 0$, (iii) $\to +\infty$ as $\tilde{a} \to N-1$, (iv) is strictly increasing in $\tilde{a}$, and (v) is independent of τ. Thus for any $N \geq 2$ and $\lambda > N/(N-1)$, there exists a threshold $\tau(\lambda, N)$ such that the solution to (12) belongs to $(1/(\lambda-1), N-1)$ for any $\tau > \tau(\lambda, N)$, as desired. □

Proof of Lemma O1. For each $i \in \mathcal{I}$, redefine $\mathcal{C}_i$ as $\mathcal{S} \cup \mathcal{I} - \{i\}$ and $\mathcal{X}_i$ as $\{(x_i^c)_{c\in\mathcal{C}_i} : \sum_{c\in\mathcal{C}_i} x_i^c \leq \tau_i\}$. Replacing the term $x_i^{\omega\text{-revealing}}$ with $\sum_{s\in\mathcal{S}} \lambda_i^s(\omega) x_i^s := y_i^{\omega\text{-revealing}}$ in the original expression for $\mathbb{P}_x(\mathcal{U}_i \mid \omega)$ gives its new expression. If a type A player i makes the default decision A in event $\mathcal{U}_i$, then his ex ante problem can be obtained from replacing the term $x_i^{B\text{-revealing}}$ in (4) with $y_i^{B\text{-revealing}}$. Since $y_i^{B\text{-revealing}}$, the nonnegative constraint $x_i^c \geq 0 \ \forall c \in \mathcal{C}_i$, and the bandwidth constraint $\sum_{c\in\mathcal{C}_i} x_i^c \leq \tau_i$, are all linear in x_i^ss, $x_i^s > 0$ only if $s \in \mathcal{S}_i(B)$, and only $\sum_{s\in\mathcal{S}_i(B)} x_i^s$ matters for the equilibrium analysis. The proof for the opposite case where decision B is made in event $\mathcal{U}_i$ is analogous and is omitted for brevity. □

Proof of Proposition O1. Part (i) of the proposition follows immediately from Lemma O1. Part (ii) of the proposition can be obtained from replacing the terms x_i^c, λ_i, and τ_i, in the proofs of Theorems 1 and 2 with νx_i^c, λ_i/ν, and $\nu\tau_i$, respectively, $\forall i \in \mathcal{I}$ and $c \in \mathcal{C}_i$. □

Proof of Proposition O2. In the augmented game with source m, redefine $\mathcal{C}_i$ as $\{m\} \cup \mathcal{I} - \{i\}$ and $\mathcal{X}_i$ as $\{(x_i^c)_{c\in\mathcal{C}_i} : \sum_{c\in\mathcal{C}_i} x_i^c \leq \tau_i\}$. Replacing the term $x_i^{\omega\text{-revealing}}$ with x_i^m in the original expression for $\mathbb{P}_x(\mathcal{U}_i \mid \omega)$ gives its new expression. Since player i's posterior belief equals the prior in event $\mathcal{U}_i$, he will make the default decision in that event. His ex ante problem is thus $\max_{x_i\in\mathcal{X}_i} -\beta_i\mathbb{P}_x(\mathcal{U}_i \mid \omega \neq t_i)/2$, or, equivalently,

$$\max_{(x_i^c)_{c\in\mathcal{C}_i}} - x_i^m - \sum_{j\in\mathcal{I}-\{i\}} \log(\delta_j^m + (1-\delta_j^m)\delta_i^j)$$
$$\text{s.t. } x_i^c \geq 0 \ \forall c \in \mathcal{C}_i \text{ and } \tau_i \geq \sum_{i\in\mathcal{C}_i} x_i^c.$$

Relabeling x_i^m as x_i^a in the above problem turns it into an analog of (4), with the only difference being that the set of type A players is $\mathcal{I}$ rather than $\mathcal{A}$. □

Proof of Theorem O2. In the setting described in Online Appendix O.3, the set of feasible sources for player i is $\mathcal{C}_i = \{1\text{-}revealing, \cdots, m\text{-}revealing\} \cup \mathcal{I} - \{i\}$, and the set $\mathcal{X}_i$ of feasible attention strategies for him is $\{(x_i^c)_{c\in\mathcal{C}_i} \in \mathbb{R}_+^{|\mathcal{C}_i|} : \sum_{c\in\mathcal{C}_i} x_i^c \leq \tau_i\}$. We focus on the case where β is small, and so all players must make default decisions in event $\mathcal{U}_i$s. The ex ante problem faced by a type m player is thus

$$\max_{x_i\in\mathcal{X}_i} -\frac{\beta}{M}\sum_{\omega\neq m}\delta_i^{\omega\text{-}revealing}\prod_{j\in\mathcal{I}-\{i\}}(\delta_j^{\omega\text{-}revealing} + (1-\delta_j^{\omega\text{-}revealing})\delta_i^j).$$

Using the parameterization proposed in Online Appendix O.3 and solving, we obtain that $\Delta^* = 0$, $(M-1)x^* + (N-1)y^* + (M-1)Nz^* = \tau$, $y^* = g_1(x^*)$, and $z^* = g_2(x^*)$, where

$$g_1(x) \coloneqq \frac{1}{\lambda}\log\max\{(\lambda(M-1)-1)(\exp(x)-1), 1\}$$
$$\text{and } g_2(x) \coloneqq \frac{1}{\lambda}\log\max\{(\lambda(M-2)-1)(\exp(x)-1), 1\}.$$

Thus x^* is the unique fixed point of $\frac{1}{M-1}[\tau - (N-1)g_1(x) - (M-1)Ng_2(x)]$, and the following are equivalent: (i) $y^* > 0$; (ii) $y^* > z^*$; (iii) $\lambda > 1/(M-1)$ and $\tau > (M-1)\phi(\lambda(M-1))$. Drawing a picture will make the last point clear. □

Proof of Theorem O3. Write $\{1, \cdots, N\}$ for $\mathcal{PV}$, and note that $\lambda_i > 1 \ \forall i \in \mathcal{PV}$. Simplifying the system (6) of equations among $\mathcal{PV}$ players accordingly yields

$$x_i^a = \max\{\tau_i - \sum_{j\in\mathcal{PV}-\{i\}}\max\left\{h(x_j^a;\lambda_j), 0\right\}, 0\} \ \forall i \in \mathcal{PV}. \tag{13}$$

Below we demonstrate that if $(\lambda_i, \tau_i) \equiv (\lambda, \tau)\ \forall i \in \mathcal{PV}$, and if $\tau - (|\mathcal{PV}|-2)h(\tau;\lambda) > \phi(\lambda)$, then the unique solution to (13) is $x_i^a \equiv x(N)\ \forall i \in \mathcal{PV}$, where $x(N)$ is the unique fixed point of $\varphi^N(x) = \tau - (N-1)h(x;\lambda)$, and it belongs to $(\phi(\lambda), \tau)$ as shown in Lemma 3.

Consider first the case $N = 2$. In that case, (13) is simply

$$\begin{aligned} x_1^a &= \max\left\{\tau - \max\left\{h\left(x_2^a;\lambda\right), 0\right\}, 0\right\} \\ \text{and } x_2^a &= \max\left\{\tau - \max\left\{h\left(x_1^a;\lambda\right), 0\right\}, 0\right\}, \end{aligned} \tag{14}$$

and drawing a picture makes it clear that (14) has a unique solution $(x(2), x(2))$. For each $N \geq 3$, we fix any pair $i \neq j$. Define $\hat{\tau} \coloneqq \tau - \sum_{k\neq i,j} \max\left\{h\left(x_k^a;\lambda\right), 0\right\}$, and note that $\hat{\tau} \in (\phi(\lambda), \tau)$ by assumption. Drawing a picture makes it clear that the solution to

$$\begin{aligned} x_i^a &= \max\left\{\hat{\tau} - \max\left\{h\left(x_j^a;\lambda\right), 0\right\}, 0\right\} \\ \text{and } x_j^a &= \max\left\{\hat{\tau} - \max\left\{h\left(x_i^a;\lambda\right), 0\right\}, 0\right\}, \end{aligned}$$

must satisfy $x_i^a = x_j^a \in (\phi(\lambda), \hat{\tau})$, and repeating this argument for all (i,j) pairs shows that $x_i^a = x_j^a \in (\phi(\lambda), \tau)\ \forall i,j \in \mathcal{PV}$. Simplifying (13) accordingly yields $x_i^a = \varphi^N(x_i^a)\ \forall i \in \mathcal{PV}$, or, equivalently, $x_i^a \equiv x(N)$. □

Proof of Theorem O4. Write $\{1, \cdots, N\}$ for $\mathcal{PV}$. For each $i \in \mathcal{PV}$, define $y_i \coloneqq x_i^a - \underline{\tau}_i$ and $\Delta\tau_i \coloneqq \overline{\tau}_i - \underline{\tau}_i$. Since $\overline{\tau}_i > \underline{\tau}_i > \phi(\lambda_i)\ \forall i \in \mathcal{PV}$, we can simplify the best response function of any $i \in \mathcal{PV}$:

$$x_i^a = \max\{\overline{\tau}_i - \sum_{j\in\mathcal{PV}-\{i\}} \frac{1}{\lambda_i} \log\max\{(\lambda_j - 1)(\exp(x_j^a) - 1), 1\}, \underline{\tau}_i\},$$

to the following:

$$y_i = \max\{\Delta\tau_i - \sum_{j\in\mathcal{PV}-\{i\}} h(y_j + \underline{\tau}_j; \lambda_j), 0\}.$$

For each $\mathbf{y} = [y_1 \ \cdots \ y_N]^\top \in Y \coloneqq \times_{i=1}^N [0, \Delta\tau_i]$, define $F(\mathbf{y})$ as the N-vector whose i^{th} entry equals $y_i + \sum_{j\in\mathcal{PV}-\{i\}} h(y_j + \underline{\tau}_j; \lambda_j) - \Delta\tau_i$. The function $F : Y \to \mathbb{R}^N$ is strongly monotone,[1] because for any $\mathbf{y}, \mathbf{y}' \in Y$,

$$\begin{aligned}
&(\mathbf{y}-\mathbf{y}')^\top (F(\mathbf{y})-F(\mathbf{y}')) \\
&= \sum_{i=1}^N (y_i - y_i')^2 + \sum_{i=1}^N \sum_{j\neq i} (y_i - y_i')[h(y_j + \underline{\tau}_j; \lambda_j) - h(y_j' + \underline{\tau}_j; \lambda_j)] \\
&\geq \|\mathbf{y}-\mathbf{y}'\|^2 - \overline{g} \sum_{i=1}^N \sum_{j\neq i} |y_i - y_i'||y_j - y_j'| \qquad (\because h_x \in (0, \overline{g})) \\
&\geq \underbrace{[1-(N-1)\overline{g}]}_{>0 \text{ by assumption}} \|\mathbf{y}-\mathbf{y}'\|^2.
\end{aligned}$$

Then by Proposition 1 of Naghizadeh and Liu (2017), the game among $\mathcal{PV}$ players has a unique equilibrium. □

Proof of Proposition O4. We only prove that $\text{sgn}\,(\partial x_k^a/\partial \tau_i) = \text{sgn}\,(-\partial x_i^a/\partial \tau_i)$ and $\text{sgn}\,(\partial x_k^a/\partial \lambda_i) = \text{sgn}\,(-\partial x_i^a/\partial \lambda_i)$ for an arbitrary pair of $k \in \mathcal{PER}$ and $i \in \mathcal{COR}$. The remaining results follow immediately from what we already know and so won't be proven again.

Write $\{1, \cdots, N\}$ for $\mathcal{COR}$. Let $\mathbf{G}_N$ denote the marginal influence matrix among $\mathcal{COR}$ players, and redefine $\mathbf{A}_N \coloneqq \mathbf{I}_N + \mathbf{G}_N$. W.l.o.g. let $i = 1$. Under the assumption that player k pays a positive amount of attention to source a, the following must hold:

$$x_k^a = \tau_k - \sum_{j=1}^N h(x_j^a; \lambda_j). \tag{15}$$

Differentiating both sides of (15) with respect to τ_1 yields

$$\frac{\partial x_k^a}{\partial \tau_1} = \sum_{j=1}^N -g_j \frac{\partial x_j^a}{\partial \tau_1} = (1-g_1)\frac{\partial x_1^a}{\partial \tau_1} - 1,$$

where the last equality follows from doing lengthy algebra, based on the fact that $\nabla_{\tau_1}[x_1^a \cdots x_N^a]^\top = \mathbf{A}_N^{-1}[1\ 0 \cdots 0]^\top$ (as shown in the proof of Theorem 3). Since $\partial x_1^a/\partial \tau_1 > 0$ by Theorem 3(i), $\text{sgn}(\partial x_k^a/\partial \tau_1) = \text{sgn}(-\partial x_1^a/\partial \tau_1)$ holds if and only if $\partial x_1^a/\partial \tau_1 < 1/(1-g_1)$. To establish the last inequality, recall from the proof of

[1] A function $f : K \to \mathbb{R}^n$ defined on a closed convex set $K \subset \mathbb{R}^n$ is *strongly monotone* if there exists $c > 0$ such that $(\mathbf{x}-\mathbf{y})^\top (f(\mathbf{x}) - f(\mathbf{y})) > c\|\mathbf{x}-\mathbf{y}\|^2\ \forall \mathbf{x}, \mathbf{y} \in K$.

Theorem 3 that $\partial x_1^a/\partial\tau_1 = \left[\mathbf{A}_N^{-1}\right]_{1,1}$ =(2). Tedious but straightforward algebra shows that

$$\left[\mathbf{A}_N^{-1}\right]_{1,1} - \frac{1}{1-g_1} = \frac{-g_1}{\det\left(\mathbf{A}_N\right)(1-g_1)} \prod_{j=2}^{N} (1-g_j) < 0,$$

as desired.

Meanwhile, differentiating both sides of (15) with respect to λ_1 yields

$$\frac{\partial x_k^a}{\partial \lambda_1} = -\sum_{j=1}^{N} g_j \frac{\partial x_j^a}{\partial \lambda_1} + \kappa,$$

where $\kappa := -h_\lambda(x_1^a;\lambda_1)$. Lengthy algebra reduces the right-hand side of the above expression to $\frac{\kappa}{\det(\mathbf{A}_N)} \prod_{j=2}^{N}(1-g_j)$, based the fact that $\nabla_{\lambda_1}[x_1^a \ \cdots \ x_N^a]^\top = \kappa \mathbf{A}_N^{-1}[0 \ 1 \ \cdots \ 1]^\top$ (as shown in the proof of Theorem 3). Thus $\text{sgn}(\partial x_k^a/\partial\lambda_1) = \text{sgn}(\kappa) = \text{sgn}(-\partial x_1^a/\partial\lambda_1)$, as desired, where the last equality was established in the proof of Theorem 3. □

Proof of Proposition O6. Write $\{1,\cdots,N\}$ for $\mathcal{A}$. Under the assumptions stated in Proposition O6, the following must hold $\forall i \in \{1,\cdots,N\}$ and $j \in \{1,\cdots,N\}-\{i\}$:

$$x_i^a = \tau_i - \sum_{j\in\mathcal{A}-\{i\}} h(x_j^a;\lambda_i^j) \tag{16}$$

$$\text{and } x_i^j = h(x_j^a;\lambda_i^j); \tag{17}$$

and $[\mathbf{G}_N]_{i,j} := h_x(x_j^a;\lambda_i^j) \in (0,1)$. Moreover, $\mathbf{A}_N := \mathbf{I}_N + \mathbf{G}_N$ must satisfy the properties stated in Lemma 2.

Part (i) of the proposition closely resembles Theorem 3(i), so its proof is omitted for brevity. For Part (ii) of the proposition, it suffices to prove the result for $i=1$ and $j=2$. Differentiating (16) with respect to λ_1^2 yields

$$\nabla_{\lambda_1^2}\mathbf{x}^a = \kappa_1^2 \mathbf{A}_N^{-1}\left[1 \ 0 \ \cdots \ 0\right]^\top,$$

where $\mathbf{x}^a := [x_1^a \ \cdots \ x_N^a]^\top$, and $\kappa_1^2 := -h_\lambda(x_2^a;\lambda_1^2)$ has an ambiguous sign in general. Since $\left[\mathbf{A}_N^{-1}\right]_{1,1} > 0$ and $\left[\mathbf{A}_N^{-1}\right]_{k,1} < 0 \ \forall k \neq 1$, by assumption, the following must hold:

$$\text{sgn}(\frac{\partial x_1^a}{\partial \lambda_1^2}) = \text{sgn}(\kappa_1^2\left[\mathbf{A}_N^{-1}\right]_{1,1}) = \text{sgn}(\kappa_1^2)$$

and

$$\text{sgn}(\frac{\partial x_k^a}{\partial \lambda_1^2}) = \text{sgn}(\kappa_1^2 \left[\mathbf{A}_N^{-1}\right]_{k,1}) = \text{sgn}(-\kappa_1^2)\ \forall k \neq 1.$$

Substituting these results into (17) yields

$$\text{sgn}(\frac{\partial x_k^1}{\partial \lambda_1^2}) = \text{sgn}(h_x(x_1^a; \lambda_k^1)\frac{\partial x_1^a}{\partial \lambda_1^2}) = \text{sgn}(\kappa_1^2)\ \forall k \neq 1$$

and

$$\text{sgn}(\frac{\partial x_m^k}{\partial \lambda_1^2}) = \text{sgn}(h_x(x_k^a; \lambda_m^k)\frac{\partial x_k^a}{\partial \lambda_1^2}) = \text{sgn}(-\kappa_1^2)\ \forall k \neq 1 \text{ and } (m,k) \neq (1,2)\,.$$

Finally, differentiating $x_1^2 = h(x_2^a; \lambda_1^2)$ with respect to λ_1^2 yields

$$\text{sgn}(\frac{\partial x_1^2}{\partial \lambda_1^2}) = \text{sgn}(\kappa_1^2 \left[h_x(x_2^a; \lambda_1^2)[\mathbf{A}_N^{-1}]_{2,1} - 1\right]) = \text{sgn}(-\kappa_1^2),$$

where the second equality follows from the assumption that $\left[\mathbf{A}_N^{-1}\right]_{2,1} < 0$. Thus in total, only the three situations listed in the proposition can happen, depending on whether κ_1^2 is positive, negative, or zero.

References

Naghizadeh, P., and M. Liu. (2017): "On the uniqueness and stability of equilibria of network games," in *2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 280–286. IEEE.