Beliefs in echo chambers: Identification and characterization

Pranjal Bhushan[*] Abhinash Borah[†] Siddarth Venkatesh[†]

December 8, 2023

Abstract

We introduce an interactive decision model capturing the functioning of echo chambers in society. Our starting point is an environment of uncertainty where individuals have preferences over Savage acts. All individuals have core beliefs, but these do not fully determine their behavior. We distinguish a decision maker's core beliefs from her behavioral beliefs, with the latter incorporating the former but, at the same time, being influenced by the behavioral beliefs of others within the echo chamber. Therefore, echo chambers feature interactive behavioral beliefs and, accordingly, interactive decision-making. We provide conditions on the model set-up that characterize the exact identification of the model parameters from observed behavior. Specifically, under these conditions, we can uniquely identify the composition of the echo chambers in society, the core beliefs of individuals, and the degree to which they are immune to influence. Further, we behaviorally characterize the version of the model that permits exact identification. We relate our model to leading empirical themes on echo chambers like elite influence, the stickiness of beliefs within echo chambers, and the polarization of beliefs across such segments of society.

JEL codes: D01, D81, D91

Keywords: echo chambers, subjective expected utility, core beliefs, behavioral beliefs, influence, behavioral identification, behavioral characterization

[*]HEC Paris (email: pranjal.bhushan@hec.edu)

[†]Ashoka University (emails: abhinash.borah@ashoka.edu.in, siddarthv2011@gmail.com)

Correspondence address: Department of Economics, Ashoka University. Rajiv Gandhi Education City, Sonepat. Haryana - 131029, India.

1 Introduction

Research across several disciplines has shown that *echo chambers* that form in society impact a range of important social and economic outcomes like polarization, populism, inequality, and asset prices in financial markets (Barberá 2020; Cookson et al. 2023; McCarty et al. 2006). A key reason echo chambers drive these outcomes is the beliefs that form within them and how they get transmitted. This paper looks at beliefs formed within echo chambers from a decision-theoretic perspective. As our starting point, we consider decision makers' (DMs') preferences over uncertain acts in the spirit of Savage (1954). The decision model we develop incorporates the idea that a DM's preferences over these acts may draw on the structure of the echo chamber she resides in through its influence on her beliefs.

The key idea underlying the model is the following. We consider a society that is partitioned into a set of echo chambers. Each individual in society has some core beliefs over the states of the world. However, her behavior is not determined exclusively by her core beliefs. Instead, her behavioral beliefs also draw on the behavioral beliefs of others within her echo chamber. Specifically, an individual's behavioral beliefs is a weighted average of her core beliefs and the average behavioral beliefs in her echo chamber. The weight put on the former measures the degree to which she is immune from her echo chamber's influence and is an important behavioral parameter. Therefore, echo chambers feature interactive beliefs and, accordingly, interactive decision-making. In particular, every DM is assumed to be a subjective expected utility maximizer with respect to her interactive behavioral beliefs. We refer to this representation of the preferences of the individuals in society as an *echo chamber representation.*

We present two key results in the paper. The first pertains to the behavioral identification of the model. Suppose the behavior of individuals in society is consistent with an echo chamber representation. Such a representation is based on several parameters. First, there is the way in which society is partitioned into echo chambers. Second, there are the core beliefs of all individuals in society. Third, is the degree to which individuals are immune to influence. Is it possible to identify all these parameters uniquely from behavior? Our first key result lays down conditions under which a positive answer can be provided to this question. What is noteworthy about these conditions is that they relate to an often-reported characteristic of how echo chambers operate in society. It has been reported that echo chambers form between people with similar beliefs over some events they are certain about. Moreover, individuals outside the echo chamber do not share this certainty. Formalizing this observation is our key identification strategy for identifying the echo chambers that form in society. At the same time, even within an echo chamber, arguably, there are differences in views regarding certainty. These differences are what produce influence amongst individuals, and is what allows us to identify details about an individual DM like the degree of her immunity from influence and her core beliefs.

Our second key result provides a behavioral foundation for a class of echo chamber representations, which we refer to as sharp echo chamber representations. These representations impose the additional restrictions discussed above that guarantee unique identification of the model parameters. Besides the subjective expected utility hypothesis, two axioms characterize this model. The first is an axiom we call validation, which essentially introduces the restriction that for a DM to assign probability 1 to an event according to her behavioral beliefs, she must have the validation of someone else who does likewise. In turn, any individual who provides such a validation is revealed to be her friend. The second axiom is based on a notion of incredulous events that we develop. Essentially, incredulous events are those that a DM is revealed to consider as unlikely by her core beliefs. We can determine such events from her behavior once we have factored out the role of her friends' influence on her beliefs. The axiom introduces the idea that there can be some disagreements amongst friends, so that an event may be incredulous for a DM and not so for a friend.

Finally, we relate our model to empirical observations regarding echo chambers that have been reported in the literature. Our model distinguishes between individuals more susceptible to influence and those less susceptible. We show that individuals who are influenced less, in turn, are the ones who have the most influence on beliefs within an echo chamber, and vice versa. Therefore, a key insight that the model develops is the disproportionate effect a few can have on the many within an echo chamber—a form of elite influence highlighted in the context of societal polarization (Wojcieszak et al. 2021). Moreover, whether a group can act on new information crucially depends on who receives this information. For instance, if individuals more susceptible to influence receive new information, then not only does the new information not get incorporated into their behavior to any great extent, but even to the extent that it does, it fails to impact the group's beliefs and behavior significantly. A consequence is that groups may have sticky beliefs that do not respond to information. Across echo chambers, we find that prior beliefs and the structure of new information can result in divergent belief updating and polarization. Arceneaux and Johnson (2015) argue that both partisan and mainstream media can polarize the beliefs of ideologically opposed individuals, the former by conveying differing information and the latter due to the way people with different beliefs respond to new information. Our model captures this insight by allowing for polarization between echo chambers upon receiving information through both channels.

The literature on echo chambers is primarily divided between examining the process of segregation into homogeneous groups and the transmission of beliefs and biases through these groups thereafter. Segregation into various groups is attributed to a variety of reasons, including economic, social, and cultural (Levy and Razin 2019). Baccara and Yariv (2016) examine the conditions under which segregation results either in homogeneous groups or polarization. Our model assumes the segregation of society into echo chambers to be exogenous, and we shift our focus to the transmission of beliefs within existing clusters, and provide behavioral identification and characterization of this process.

Golub and Jackson (2010) and Acemoglu et al. (2014) characterize convergence to the truth in a network where individuals communicate noisy signals with each other. In addition, Acemoglu et al. (2014) incorporate endogenous network formation into their model. Levy and Razin (2015a, 2015b) examine the effects of correlation neglect between signals on polarization and political outcomes. In particular, they argue that it is possible to achieve better outcomes compared to a rational electorate. Martinez and Tenev (2022) similarly argue that echo chambers could improve the process of learning if the quality of various sources of information is uncertain. A common characteristic of this literature is that information is transmitted within echo chambers by assuming the possibility of directly sharing signals amongst peers, whereas we restrict the channel of influence to observed behavior. On these lines, Eyster and Rabin (2010) model herding behavior based on sequential observation of actions. They find that in settings where agents assume that the observed actions are based solely on private information, they may converge to incorrect actions with confidence.

There is a substantial body of recent literature that studies the process of social influence from a choice theoretic perspective (Fershtman and Segal 2018; Kashaev and Lazzati 2019; Lazzati 2020; Borah and Kops 2018; Chambers et al. 2019; Chambers et al. 2021; Cuhadaroglu 2017). In particular, the structure of our model draws inspiration from Fershtman and Segal (2018). Whereas we model influence through belief transmission, they look at influence in tastes. They consider two sets of preferences for each individual, represented by her core and behavioral utilities, and an influence function such that her behavioral utilities can be represented as a function of her core utility and others' behavioral utilities. Analogous to their model, core beliefs are private and behavioral beliefs are observable in our model.

The rest of the paper is organized as follows. The next section introduces our setup and formally defines an echo chamber representation. Section 3 introduces the conditions that characterize the exact identification of the model parameters. Section 4 discusses the behavioral foundation of the model. Finally, in Section 5, we elaborate on some properties and empirical content of the model. Proofs of all results are available in the Appendix.

2 Setup

Our stylized society consists of a set $I = \{1, \ldots, n\}$ of individuals. These individuals are assumed to be partitioned into the sets $\langle E_1, \ldots, E_k \rangle$, with each element of the partition denoting an echo chamber (or chamber, for short) in society. To keep the setup meaningful, we assume that none of the echo chambers is a singleton. For any $i \in I$, we let $E(i)$ denote the echo chamber to which individual i belongs.

Let S be a finite set of states. Any subset of S is referred to as an event. Z denotes a set of prizes. An act is a mapping, $f : S \to Z$, from the set of states to the set of prizes. We denote the set of acts by H. Each individual i has preferences over the set of acts, denoted by $\succcurlyeq_i \subseteq H \times H$. We denote the symmetric and asymmetric component of $\succcurlyeq_i$ by $\sim_i$ and $\succ_i$, respectively.

Given the underlying uncertainty in the environment, we imagine that any individual decision maker (DM) forms some beliefs over the set of states. The way these beliefs are influenced by the echo chamber this DM is a part of is the focus of our model. We assume that any such DM, $i \in I$, is endowed with some *core belief* μ_i on S. However, her behavior may not be driven exclusively by her core beliefs. We will call the beliefs that drive this DM's behavior her *behavioral belief*, denoted by π_i. Behavioral beliefs, of course, depend on core beliefs. But, additionally, the working of influence within her echo chamber implies that her behavioral beliefs may be influenced by the behavioral beliefs of others in her echo chamber. We use the average behavioral belief prevailing within her echo chamber, $\frac{1}{|E(i)|} \sum_{j \in E(i)} \pi_j$, as a summary statistic capturing this aspect of influence. We assume that the dependence on the two takes a linear weighted average form, with the weights determined by a parameter $\alpha_i \in (0, 1)$ that captures the degree to which this DM is immune from influence, i.e., higher is α_i, the less susceptible is this DM to influence. Specifically, for any state $s \in S$, we assume that her behavioral belief is given by,

$$\pi_i(s) = \alpha_i \mu_i(s) + (1 - \alpha_i) \frac{1}{|E(i)|} \sum_{j \in E(i)} \pi_j(s)$$

This, therefore, makes beliefs and, accordingly, behavior interactive within an echo chamber.

Our representation of the collection of preferences $(\succcurlyeq_i)_{i \in I}$ captures these interactions. It highlights two key ideas. First, it captures the steady state of this process of interactions by requiring mutually consistent behavioral beliefs within an echo chamber. Second, it clarifies that individuals' knowledge of the behavioral beliefs of others in their cluster is drawn from their behavior, just like the analyst's.

Definition 1 *The collection of preferences* $(\succsim_i)_{i\in I}$ *has an echo chamber representation if there exists a partition* $\langle E_1, E_2, ..., E_k\rangle$ *of* I*, and for each* $i \in I$*:*

- *a (non-constant) utility function* $u_i : Z \to \mathbb{R}$*,*
- *a core probability measure* μ_i *on* S*, and*
- *an immunity from influence parameter* $\alpha_i \in (0, 1)$

such that defining the collection of behavioral probability measures $(\pi_i)_{i\in I}$ *on* S*, by*

$$\pi_i(s) = \alpha_i\mu_i(s) + (1-\alpha_i)\frac{1}{|E(i)|}\sum_{j\in E(i)}\pi_j(s), \ \ i \in I,$$

we have that for each $i \in I$*, the function* $U_i : H \to \mathbb{R}$ *given by*

$$U_i(f) = \sum_{s\in S}\pi_i(s)u_i(f(s))$$

represents $\succsim_i$*.*

In other words, DMs are subjective expected utility maximizers, with their beliefs formed through the process of interaction within echo chambers outlined above. Like in an equilibrium notion, we close the interactions by assuming that individuals hold correct beliefs about the preferences of others in their echo chamber and, through it, their behavioral beliefs. This allows them to correctly forecast the average belief about any event in their echo chamber.

Remark 1 It is straightforward to establish that the steady state notion captured by the echo chamber representation doesn't suffer from concerns about non-existence. That is, given a collection $(\mu_i)_{i\in I}$ of core beliefs, it is immediate to establish that there exists a collection of behavioral beliefs $(\pi_i)_{i\in I}$ that simultaneously satisfy the system of equations:

$$\pi_i(s) = \alpha_i\mu_i(s) + (1-\alpha_i)\frac{1}{|E(i)|}\sum_{j\in E(i)}\pi_j(s), \ \ i \in I$$

To see this, note that the equation determining the behavioral beliefs of individual i, depends only on the beliefs of the individuals belonging to $E(i)$. Thus, it is sufficient to prove existence for a single echo chamber. Given a chamber E_k, subtract both sides by the average behavioral belief, and sum over all $i \in E_k$ to yield:

$$\frac{1}{|E_k|}\sum_{j\in E_k}\pi_j(s) = \sum_{i\in E_k}\frac{\alpha_i\mu_i(s)}{\sum_{j\in E_k}\alpha_j}$$

Substituting this expression in the earlier system of equations, we get:

$$\pi_i(s) = \alpha_i \mu_i(s) + (1 - \alpha_i) \sum_{j \in E_k} \frac{\alpha_j \mu_j}{\sum_{l \in E_k} \alpha_l}$$

This process can be used for each echo chamber, and the resulting collection $(\pi_i)_{i \in I}$ satisfies the system of equations. Since this is also true of all $(\pi_i)_{i \in I}$ that solve the equations, given a collection of $(\mu_i)_{i \in I}$, $(\alpha_i)_{i \in I}$ and partition $\langle E_1, ..., E_k \rangle$, the resultant $(\pi_i)_{i \in I}$ must be unique.

3 Identification

We now address the question about the behavioral identification of the model parameters. That is, suppose $(\succcurlyeq_i)_{i \in I}$ has an echo chamber representation. Is this representation unique, or are multiple such representations possible? In this section, we introduce two conditions on the underlying structure of the model that characterize uniqueness of the representation. To do so, we introduce some terminology. We say that i is fundamentally certain about an event A if $\mu_i(A) = 1$. We say that a chamber E is fundamentally certain about an event A if $\mu_i(A) = 1$, for all $i \in E$. The first condition provides a statement about differing views on certainty across echo chambers.

Condition 1 *For all chambers E and E', there exists an event that one of the chambers is fundamentally certain about, but the other chamber is not.*

We consider this a defining property of echo chambers—the fact that, at a fundamental level, echo chambers are formed by people who share a certain view of "reality" that is not necessarily shared by those outside the echo chamber. The condition above expresses this idea in a fairly weak form. At the same time, we don't take the position that there is complete agreement on matters of certainty within an echo chamber. Our next condition captures this viewpoint, leaving open the possibility that there is room for disagreement and influence within an echo chamber.

Condition 2 *For all $i \in I$, there exists an event she is fundamentally certain about, but someone in her chamber is not*

Theorem 1 *Suppose $(\succcurlyeq_i)_{i \in I}$ has an echo chamber representation. The following statements are equivalent.*

1. *Conditions 1 and 2 hold*

2. *If* $(E = \langle E_1, \ldots, E_k \rangle, (u_i, \mu_i, \alpha_i)_{i\in I})$ *and* $(\tilde{E} = \langle \tilde{E}_1, \ldots, \tilde{E}_l \rangle, (\tilde{u}_i, \tilde{\mu}_i, \tilde{\alpha}_i)_{i\in I})$ *are both echo chamber representations of* $(\succcurlyeq_i)_{i\in I}$*, then* $E = \tilde{E}$*, and for each* $i \in I$*,* $u_i = a_i\tilde{u}_i + b_i$ *with* $a_i > 0$*,* $\mu_i = \tilde{\mu}_i$*, and* $\alpha_i = \tilde{\alpha}_i$*.*

Proof: Please refer to Appendix Section A.1

One of our main concerns in this paper is identification. In general, identifying echo chambers from behavior, along with beliefs that underlie these chambers, is a challenging task. Our two conditions identify fairly intuitive and empirically relevant conditions under which such identification is possible. We, therefore, are particularly interested in considering the echo chamber model with these two additional conditions. Henceforth, we refer to an echo chamber representation that satisfies conditions 1 and 2 as a *sharp echo chamber representation.*

4 Behavioral characterization

Given the discussion above, we now provide a behavioral foundation for the sharp echo chamber model. The first axiom is obvious. An echo chamber representation implies that any DM assesses acts according to a subjective expected utility (SEU) criterion w.r.t. a utility function on the set of prizes and a probability measure on the state space. Since behavioral foundation of an SEU representation is a well-studied problem, we do not get into those details here. Rather, we directly assume that individual preferences have an SEU representation

Axiom A1 (SEU) $\succcurlyeq_i$ *is non-degenerate and has an SEU representation, for each* $i \in I$*.*

$\succcurlyeq_i$ non-degenerate means that $\succ_i \neq \emptyset$. Given that $\succcurlyeq_i$ is non-degenerate and has an SEU representation, we know that each i has a unique behavioral probability measure π_i on the state space. Our next axioms makes use of these elicited beliefs to make two important observation.

The first observation pertains to the calculus of certainty within any echo chamber and the mutual influence that the process involves. It says that any individual considers an event to be certain by her behavioral probabilities only if there exists another individual who does likewise, with this influence being mutual. In other words, even if an individuals fundamentally thinks of an event as certain, it is not guaranteed to translate into her behavior unless it receives validation from someone else, with this process of validation being mutual.

Axiom A2 (Validation) *$\forall i \in I$, $\exists j \in I$ such that, for any event A, $\pi_i(A) = 1$ if and only if $\pi_j(A) = 1$.*

To introduce our final axiom, we first introduce some definitions. First, we define the revealed friends of any $i \in I$ by

$$R(i) = \{j \in I : \pi_i(A) = 1 \iff \pi_j = 1, \text{for any event } A\}$$

The idea behind this revealed elicitation is quite straightforward. An individual's friends, presumably, are those she seeks validation from. Observe that, by definition, $i \in R(i)$, and the validation axiom guarantees that there exists at least one other $j \in R(i)$.

Now, given $R(i)$, we can define the average belief about an event amongst i's friends. Specifically, for any event A, let

$$\overline{\pi}_i(A) = \frac{1}{|R(i)|} \sum_{j \in R(i)} \pi_j(A)$$

With this in place, we now want to develop a notion of events that a DM may consider unlikely by her core beliefs. To elicit this from her behavioral beliefs, we essentially have to wash away the role of her friends' influence on the latter. To do so we introduce the following key notion. We say that $i \in I$ finds an event A *incredulous* if,

$$\frac{\pi_i(A)}{\overline{\pi}_i(A)} \leq \frac{\pi_i(B)}{\overline{\pi}_i(B)}, \text{ for all events } B \text{ with } \pi_i(B) > 0$$

The idea behind this notion is straightforward. $\pi_i(A)$ measures how much i is willing to pay for a bet that pays one util on this event. On the other hand $\overline{\pi}_i(B)$ captures how much her friendship network, on average, will pay for such a bet. Therefore, the lower this ratio is the lower must be i's assessment according to her core beliefs about its likelihood.

Our final axiom reiterates the point that even amongst friends there are some disagreements when it comes to core beliefs.

Axiom A3 (Some disagreement amongst friends) *For any $i \in I$, there exists an event that i is incredulous about but at least one of her friends is not*

We can now present our behavioral characterization result.

Theorem 2 $(\succcurlyeq)_{i\in I}$ *has a sharp echo chamber representation if and only if axioms SEU, validation and some disagreement amongst friends hold.*

Proof: Please refer to Appendix Section A.2

5 Properties of the model

In this section, we highlight some important properties of the model. These pertain to the scope of elite influence, belief updating with the possibility of sticky beliefs, and the polarization of beliefs.

5.1 Elite Influence

A theme that has featured prominently in recent times is that of elite influence. For instance, it has been pointed out in the context of partisan politics that each side of the partisan divide has elites who have a disproportionate influence on their respective sides. In other words, when there is influence at play, it is typically marked by a great degree of heterogeneity in terms of the ability to influence. Such an effect shows up in our model. The key feature in this regard that our model demonstrates is that within each echo chamber, the individuals who are the least influenced (high α-s) happen to be the ones who end up having the greatest influence in terms of shaping beliefs within their echo chamber.

The nature of linear influence in the model gives the average belief in the echo chamber a unique structure. Recall the following equality we derived in Remark 1.

$$\overline{\pi}_i(s) = \sum_{j\in E(i)} \frac{\alpha_j}{\sum_{k\in E(i)} \alpha_k} \mu_j(s)$$

That is, the average belief in an echo chamber can be represented as the weighted average of core beliefs, with the weights capturing relative influence. In particular, the weight attached to i's core belief is given by $\frac{\alpha_i}{\sum_{k\in E(i)} \alpha_k}$. It is relative because it depends on the ratio of the DM's own α_i to the sum of all α_j. It also measures the degree of influence as the more a DM is immune to influence, the greater the weight placed on her core belief in the determination of the average behavioral belief. However, the more others are immune from influence, the more they influence the average belief, thus reducing the relative influence exhibited by the DM. A way of capturing the influence exhibited by a

DM is the difference between her core beliefs and the average behavioral beliefs, which is expressed as follows.

$$|\overline{\pi}_i(s) - \mu_i(s)| = \left| \frac{1}{\sum_{k \in E(i)} \alpha_k} \left(\sum_{j \in E(i) \setminus \{i\}} \alpha_j(\mu_j(s) - \mu_i(s)) \right) \right|$$

This is a measure of her influence because it captures how close average behavioral beliefs in her echo chamber are pulled towards her core beliefs. Note that α_i appears only in the denominator, which means the difference is decreasing in α_i, and for any collection $(\mu_i)_{i \in I}$, the average behavioral belief is influenced more by i's core belief if she is more immune to influence. Note also, that this difference depends on the relative influence exhibited by others. As α_j increases for $j \neq i$, the average belief moves closer to j's core belief, and $|\overline{\pi}_i(s) - \mu_j(s)|$ decreases. Since individual behavioral beliefs are a weighted average of core beliefs and average behavioral beliefs, each individual's behavioral belief moves closer to j's core belief. Thus, $|\pi_i(s) - \mu_j(s)|$ decreases too. In this manner, individuals who are more immune to influence and exhibit greater influence on the echo chamber act as a group of elites, whose core beliefs play a large role in determining average behavioral beliefs. This feature of elite influence also manifests itself in the way private information is incorporated into behavior and transmitted to others in the echo chamber.

5.2 Belief Updating

5.2.1 Mechanism

We now examine the ways in which beliefs are updated in our model upon the arrival of new information. The key assumption we maintain in the model is that this information is conveyed to any individual through private signals that cannot be shared with anyone else in the echo chamber. This means that the core probabilities of only the information recipient get updated, and the response within the echo chamber is purely driven by the behavioral changes she exhibits. Define a signal function $\sigma_i : S \to \mathbb{R}$ for each individual i such that for any signal $c \in \mathbb{R}$ that is received, the realized state s must be in $\sigma_i^{-1}(c)$. Denoting by c^i a signal received by i, we have $\cap_{i \in I} \sigma_i^{-1}(c^i) \neq \emptyset$. We assume that for any $c^i \in \mathbb{R}$, $\mu_i(\sigma_i^{-1}(c^i)) > 0$. Since σ_i-s are different for every individual, the content and precision of information they receive can differ.

Now given any signal structure $(\sigma_i)_{i \in I}$, an individual $i \in I$ will update her core beliefs upon the realization of any signal according to Bayes' rule. Suppose c is the signal received by i. Then, her posterior core belief is given by:[1]

$$\mu_i(s|c) = \frac{\mathbb{1}\left[s \in \sigma_i^{-1}(c)\right]\mu_i(s)}{\sum_{s' \in \sigma_i^{-1}(c)} \mu_i(s')}$$

As remarked earlier, while signals cannot be transmitted to others in the echo chamber, the way behavioral beliefs are updated throughout the network is through changes in observed behavior. Suppose each individual $i \in I$ receives a signal c^i. Denote their updated core belief $\mu_i(.|c^i)$ by μ_i' and their behavioral belief $\pi_i(.|c^i)$ by π_i'. Then, the change observed in each individual's behavioral belief is given by:

$$\pi_i'(s) - \pi_i(s) = \alpha_i(\mu_i'(s) - \mu_i(s)) + (1-\alpha_i)\sum_{j \in E(i)} \frac{\alpha_j(\mu_j'(s) - \mu_j(s))}{\sum_{k \in E(i)} \alpha_k}$$

5.2.2 Incorporation of private information

First, consider the case that individual i is the only one who receives private information in the echo chamber, which will happen as long as σ_j is a constant function for all $j \neq i \in E(i)$. In this case, the ratio of the change in her behavioral beliefs to the change in core beliefs can be written as:

$$\frac{\pi_i'(s) - \pi_i(s)}{\mu_i'(s) - \mu_i(s)} = \alpha_i\left(1 + \frac{1-\alpha_i}{\sum_{k \in E(i)} \alpha_k}\right)$$

Note that this ratio is increasing in α_i and is equal to 0 for $\alpha_i = 0$ and 1 for $\alpha_i = 1$. Note also, that it is bounded below by α_i. Taking the second derivative will reveal that it is concave in α_i. To interpret this, we can decompose the effect of α_i on this ratio into the direct effect $D(\alpha_i) = \alpha_i$, and the indirect effect defined as:

$$I(\alpha_i) = \frac{\alpha_i(1-\alpha_i)}{\sum_{j \in E(i)} \alpha_j}$$

The indirect effect captures the impact of the change in her beliefs on others' behavioral beliefs, which she then incorporates into her own behavioral beliefs. While $D'(\alpha_i) = 1$ for all $\alpha_i \in (0, 1)$, $I''(\alpha_i) < 0$ and $I'(\alpha_i) < 0$ for some $\alpha_i \in (0, 1)$. While $D(\alpha_i)$ captures the effect of μ_i on π_i through the DM's immunity from influence, $I(\alpha_i)$ captures the effect of the DM's core beliefs that influence her through the behavioral beliefs of others in her echo chamber. Since $I(0) = I(1) = 0$, there exists some $a \in (0, 1)$ such that $I'(a) = 0$.

[1] Note that each $\mathbb{P}(c = c_k \mid s = s_i)$ evaluates to either 1 or 0, which is equivalent to the indicator function $\mathbb{1}\left[s_i \in \sigma_i^{-1}(c_k)\right]$. Using the law of total probability, each $\mathbb{P}(c = c_k)$ can be written as $\sum_{s_i \in S} \mathbb{P}(c = c_k \mid s = s_i)\mu(s_i)$, which is then simplified to the sum of the core probabilities over all possible states (i.e. $s_i \in \sigma_i^{-1}(c_k)$).

Denote by $\overline{a} = \sum_{j \in E(i) \setminus \{i\}} \alpha_j$. Then we can uniquely define a:[2]

$$a = \sqrt{\overline{a}(\overline{a}+1)} - \overline{a}$$

We know that $I'(\alpha_i) > 0$ for $\alpha_i < a$ and $I'(\alpha_i) < 0$ for $\alpha_i > a$. What this means is that for $\alpha_i < a$, the rise in the DM's influence on her echo chamber's behavioral beliefs outweighs the fall in the influence of the echo chamber's behavioral beliefs on her own. For high α_i, the direct effect is high, but the indirect effect starts decreasing as $1 - \alpha_i$ goes to 0. We can also note that for $\alpha_i \in (0,\, 1)$, $D(\alpha_i) + I(\alpha_i)$ is never 1.

5.2.3 Belief updation and transmission

We now focus attention on the transmission of updated beliefs. Since DMs in our model require those in their chamber to agree on sure and null events, they also require agreement on new information they receive for it to translate into behavior fully. Without such agreement, beliefs may be sticky and fail to incorporate the full extent of the available information.

Proposition 1 (Sticky Beliefs) *For any $i \in I$, $\pi_i(\sigma_i^{-1}(c^i) \,|\, c^i) = 1$ if and only if $\mu_j(\sigma_i^{-1}(c^i) \,|\, c^j) = 1$, for all $j \in E(i)$.*

Proof: Please refer to Appendix Section A.3

For Bayesian DMs whose beliefs are solely based on private information, the way they update their beliefs upon receiving new information would imply that if $\mathbb{P}(c \,|\, s) = 0$, then $\mathbb{P}(s \,|\, c) = 0$, while for s such that $\mathbb{P}(s) = 0$, $\mathbb{P}(s \,|\, c) = 0$. However, in our model, $\pi_i(A) = 0$ if and only if $\overline{\pi}_i(A) = 0$, which holds if and only if $\mu_j(A) = 0$ for all $j \in E(i)$. Since DMs seek the validation of all members of their echo chamber, they are unable to assign beliefs according to their private information, which would demand that they rule out all states $s \notin \sigma_i^{-1}(c^i)$. Instead, they assign 0 probability to only those states that are ruled out by everyone else and null states according to their prior. In other words:

$$\pi_i\left(S \setminus \cup_{j \in E(i)} \sigma_j^{-1}(c^j)\right) = 0$$

A special case of this is when the signal received by the DM suggests that she should rule out a state, but not only does she assign it a positive probability, her posterior belief on the state can actually be greater than that under her prior. Since DMs in our model seek

[2]To do so, we must find the roots of the $I'(\alpha_i)$, the solution to which is given by $\sum_{j \in E(i) \setminus \{i\}} \alpha_j = -\frac{\alpha_i^2}{2\alpha_i - 1}$. This function is invertible in the domain $(0,\, 0.5)$.

their echo chamber's validation when assigning 0 probability to any event, they are unable to rule states out solely on private information. However, if high-influence individuals believe that such a state is more likely given the signals they receive, the DM may end up considering the state to be more likely even though her private information implies otherwise. Denote by π_i' and μ_i' the posterior behavioral and core beliefs of individual i. Then, $\pi_i'(s) > \pi_i(s)$ for some $s \notin \sigma_i^{-1}(c^i)$ if the following inequality is satisfied.

$$\frac{\alpha_i(1 + \sum_{k\in E(i)\setminus\{i\}} \alpha_k)}{1-\alpha_i}\mu_i(s) < \sum_{k\in E(i)\setminus\{i\}} \alpha_k(\mu_k'(s) - \mu_k(s))$$

This inequality suggests that this effect is most likely to be prevalent for individuals with low α_i and for states where their prior core belief was low. While low immunity from influence makes it more difficult for DMs to incorporate their private information, the seeking of validation from all members in the cluster, i.e. $\pi_i(A) = 0 \iff \pi_j(A) = 0$ for $j \in E(i)$, comes in the way of the DM assigning $\pi_i(A)$ equal to 0 if $\mu_j(A|c^j) - \mu_j(A)$ is strongly positive for some $j \in E(i)$ with a high α_j. In that case, the DM ends up acting against her private information.

We can now extend the statement of Proposition 1 to how the behavioral beliefs of an entire echo chamber are updated when new information is received. Let E_k be some echo chamber and denote by S_0^i the set of states such that $\mu_i(s_0) = 0$ for all $s_0 \in S_0^i$.

Corollary 1 $\pi_i(\sigma_i^{-1}(c^i)\,|\,c^i) = 1$ *for all* $i \in E_k$ *if and only if for any* $i, j \in E_k$

$$\sigma_i^{-1}(c^i) \setminus S_0^i \subseteq \sigma_j^{-1}(c^j)$$

Proof: Please refer to Appendix Section A.4

Corollary 1 asserts that agents can assign non-zero probability to only those states that are not ruled out by the signals received by everyone else in the echo chamber. The result follows directly from Proposition 1 in that it guarantees that $\mu_i(\sigma_j^{-1}(c^j)\,|\,c^i) = 1$ for all $j \in E(i)$ for every choice of $i \in E_k$. This is also equivalent then to all non-null states in the posterior being in the intersection of all $\sigma_j^{-1}(c^j)$ for $j \in E(i)$. That means that every agent in the echo chamber receives the same information about non-null states. Contrarily, in the presence of heterogenous information, DMs are unable to fully incorporate their private information. Instead, they require that everyone's observed beliefs agree with their private information, which is only possible if everyone receives the same private information. The only potential disagreements in signals are then restricted to the domain of null states, which do not affect posterior beliefs.

5.3 Polarization across echo chambers

Whereas the issue of stickiness is induced by the nature of communication of beliefs within an echo chamber, a notable phenomenon that is seen across different echo chambers is polarization. The prevalence and applications of polarization are well documented. We define a measure of polarization in the context of our model and examine the conditions under which it is observed. There is no precise quantification of polarization that is available in the context of our setup, but to capture the notion of the divergence of beliefs, we consider the absolute difference in average behavioral probabilities between two chambers:[3]

$$\rho_{(k,l)} = \sum_{s \in S} \frac{|\overline{\pi}_{E_k}(s) - \overline{\pi}_{E_l}(s)|}{2}$$

We can also define the extent to which polarization is observed over beliefs regarding some subset of S, and this may be greater or lower than the polarization over all of S.

$$\rho_{(A)(k,l)} = \sum_{s \in A} \frac{|\overline{\pi}_{E_k}(s) - \overline{\pi}_{E_l}(s)|}{\sum_{s' \in A}(\overline{\pi}_{E_k}(s') + \overline{\pi}_{E_l}(s'))}$$

This idea of polarization, then, can be revised upon the realization of some signal. After the realization of any particular set of signals, the extent of polarization between two clusters is then given by defining $A = \cup_{i \in E_k \cup E_l} \sigma_i^{-1}(c^i)$, and considering the updated $\rho_{(A)(k,l)}$. The goal, then, is to see the cases in which there may be greater extent of polarization than others. Consider the following motivating examples.

Example: Consider the state space $S = \{1, 2, ..., 100\}$ and set of individuals $I = \{1, ..., n\}$ with two echo chambers E_1 and E_2 that partition I. Assume that individuals have a uniform distribution over the state space as their core beliefs, and thus identical priors. In this scenario, the extent of their polarization $\rho_{(1,2)} = 0$. As such, we are starting with the least polarized society possible. Now consider that everyone in E_1 receives a signal that the realized state is a prime number, while everyone in E_2 receives a signal that the state is an even number. The only point of intersection between the signals is the state 2. Upon updating their beliefs, $\overline{\pi}_{E_1}(2) = \frac{1}{25}$ and $\overline{\pi}_{E_2}(2) = \frac{1}{50}$. However, $\overline{\pi}_{E_1}(\text{Odd primes}) = \frac{24}{25}$, while $\overline{\pi}_{E_2}(\text{Even}) = 1$. Defining A to be the union of the prime numbers and even numbers, we get the new polarization to be $\rho_{(A)(1,2)} = 0.99$.

To formalize this intuition, we introduce vector notation for signals, such that $\boldsymbol{\sigma}_k : S \to \mathbb{R}^{|E_k|}$ is the vector of all signal functions and $\boldsymbol{c}_k$ is a vector of signals received by individuals in echo chamber E_k. Then $\boldsymbol{\sigma}_k^{-1}(\boldsymbol{c}_k)$ gives a vector of sets of states deemed possible by each individual, with its i^{th} element denoted by $\boldsymbol{\sigma}_k^{-1}(\boldsymbol{c}_k)(i)$. Then define the set of states

[3] We interchangeably use notation denoting average echo chamber beliefs as per convenience, $\overline{\pi}_i = \overline{\pi}_{E(i)}$

deemed possible by at least one person in the chamber as follows.

$$\hat{S}_k(\boldsymbol{c}_k) = \bigcup_{i \in E_k} \boldsymbol{\sigma}_k^{-1}(\boldsymbol{c}_k)(i)$$

Particularly, if $\overline{\pi}_{E_k}(s) > 0$ then $s \in \hat{S}_k(\boldsymbol{c}_k)$. Let $\rho'_{(k,l)}$ denote the polarization between E_k and E_l upon receiving the signals $(\boldsymbol{c}_k, \boldsymbol{c}_l)$. Then $\rho'_{(k,l)} = \rho'_{(A)(k,l)}$ where $A = \hat{S}_k(\boldsymbol{c}_k) \cup \hat{S}_l(\boldsymbol{c}_l)$. The example above also suggests that a potential cause for polarization is the difference in information received by one echo chamber as opposed to another. Thus, it is meaningful to model settings where individuals in different echo chambers receive different signals, particularly ones such that $\hat{S}_k(\boldsymbol{c}_k) \Delta \hat{S}_l(\boldsymbol{c}_l) \neq \emptyset$.[4] We find that polarization increases in such settings as there is greater disagreement between signals.

We next introduce a notion of misinformation and highlight its connection to polarization, which has often been referenced in both academic work and popular media. Then any state $s \notin \bigcap_{i \in I} \sigma_i^{-1}(c^i)$, can be ruled out from being the true realized state. Assuming that the initial preferences of all the individuals had a sharp echo chamber representation, we can identify states that had a non-null prior and a null posterior, thus identifying the states each individual was able to rule out through their signal. The union of all such states can be ruled out from being the true realized state by an observer. Putting this together, we can deem any subset of $\sigma_i^{-1}(c^i)$ contained in this union to be untrue, and capturing misinformation in the signal.

Since any state s that does occur must be in $\cap_{i \in I} \sigma_i^{-1}(c^i)$, disagreement can only occur over misinformation. We can then limit our observations to any two clusters at once. Say a state s can be deduced untrue by E_k and E_l if $s \in \hat{S}_k(\boldsymbol{c}_k) \Delta \hat{S}_l(\boldsymbol{c}_l)$. Then we have the following result. To state it, define a function $\hat{\sigma}_i$ for some $i \in E_l$ such that:[5]

$$\hat{\sigma}_i(s) = \begin{cases} \sigma_i(s) & s \neq s^* \\ c^* \notin \mathcal{R}(\sigma_i) & s = s^* \end{cases}$$

given some state $s^* \in S$. We then define the vector $\hat{\boldsymbol{\sigma}}_l$ such that $\hat{\boldsymbol{\sigma}}_l[j] = \sigma_j$ for $j \in E_l$, $j \neq i$ and $\hat{\boldsymbol{\sigma}}_l[i] = \hat{\sigma}_i$. Denote by $\rho_{(k,l)}(s)$ the polarization between E_k and E_l when the signals faced by individuals in the echo chambers are a function of the realized state being s.

Proposition 2 *Define $\rho^1_{(k,l)}(s)$ and $\rho^2_{(k,l)}(s)$ to be the levels of polarization between E_k and E_l when individuals in E_l face the signals $\boldsymbol{\sigma}_l(s)$ and $\hat{\boldsymbol{\sigma}}_l(s)$ respectively. Then $\rho^2_{(k,l)}(s) \leq$*

[4] $A \Delta B$ denotes the symmetric difference of the two sets, which is defined as $(A \setminus B) \cup (B \setminus A)$. That is, elements that can be found in either set but not both.

[5] $\mathcal{R}(f) = \{y \in Y : \exists x \in X \text{ s.t. } f(x) = y\} \subseteq Y$ denotes the range set of a function $f : X \to Y$

$\rho^1_{(k,l)}(s)$ *if* s^* *can be deduced untrue by* E_k *and* E_l*.*

Proof: Please refer to Appendix Section A.5

What Proposition 2 tells us, from another perspective, is that making signals less precise, or signalling to even one individual that another state not contained in either $\hat{S}_k(\boldsymbol{c}_k)$ or $\hat{S}_l(\boldsymbol{c}_l)$ could occur can only increase polarization. Simply put, differential misinformation across echo chambers leads to greater polarization due to a lack of communication of beliefs between them. Furthermore, as Section 5.1 makes evident, if the individual who is more or less misinformed is relatively immune from influence, then the effect on polarization is greater than through others. Therefore, if one were to convey information that could otherwise be deduced untrue between any two echo chambers to a high-influence individual in either cluster, they could increase polarization by much more.

A Appendix

A.1 Proof of Theorem 1

We first prove **sufficiency** of the two conditions for exact identification. We begin by showing the uniqueness of behavioral probabilities and utility functions and then establish the uniqueness of the echo chambers. Thereafter, we show that the fundamentally certain events are invariant across representations. Finally, we show that the influence parameters and core probabilities are also unique. To that end, suppose $\big(E = \langle E_1, ..., E_k\rangle, (u_i)_{i\in I}, (\mu_i)_{i\in I}, (\alpha_i)_{i\in I}\big)$ and $(\tilde{E} = \langle \tilde{E}_1, ..., \tilde{E}_l\rangle, (\tilde{u}_i)_{i\in I}, (\tilde{\mu}_i)_{i\in I}, (\tilde{\alpha}_i)_{i\in I})$ are two echo chamber representations of $(\succsim_i)_{i\in I}$ satisfying Conditions 1 and 2.

<u>$(\pi_i)_{i\in I}$ **are unique and so are** $(u_i)_{i\in I}$ **upto positive affine transformations**</u>

First, note that (π_i, u_i) and $(\tilde{\pi}_i, \tilde{u}_i)$ are two SEU representations of $\succsim_i$. Therefore, from the uniqueness properties of an SEU representation, we know that $\tilde{\pi}_i = \pi_i$ and $\tilde{u}_i = a_i u_i + b_i$ with $a_i > 0$.

<u>**The echo chambers are unique,** $E = \tilde{E}$</u>

First, note from the observation made in Remark 1, it follows that for any i, $\pi_i(A) = 1$ if and only if $\mu_j(A) = 1$ for all $j \in E(i)$. That is, $\pi_i(A) = 1$ if and only if $\pi_j(A) = 1$, for all $j \in E(i)$. Now, consider some chamber $E(i)$ under the first representation, and suppose there exists j such that $j \in E(i)$ but $j \notin \tilde{E}(i)$. Since $\tilde{E}(i) \neq \tilde{E}(j)$, by Condition 1, there exists an event A that is fundamentally certain for one of the chambers but not the other.

Suppose A is fundamentally certain for $\tilde{E}(i)$ but not $\tilde{E}(j)$. Then, $\tilde{\pi}_i(A) = 1$ and $\tilde{\pi}_j(A) \neq 1$. Given the uniqueness of behavioral probabilities, this means that $\pi_i(A) = 1$ and $\pi_j(A) \neq 1$. But, $\pi_i(A) = 1 \neq \pi_j(A)$ is impossible as $j \in E(i)$! By a similar argument, A being fundamentally certain for $\tilde{E}(j)$ but not $\tilde{E}(i)$ is also not possible, violating Condition 1. Hence, if $j \in E(i)$, then $j \in \tilde{E}(i)$, and $E(i) \subseteq \tilde{E}(i)$. By a symmetric argument, there cannot exist any $j \in \tilde{E}(i)$ such that $j \notin E(i)$. Thus $\tilde{E}(i) \subseteq E(i)$, which together imply that $E(i) = \tilde{E}(i)$. Since this is true for all $i \in I$, it follows that $E = \tilde{E}$.

Individuals' fundamentally certain events are unique, $\mu_i(A) = 1$ iff $\tilde{\mu}_i(A) = 1$

Suppose $\mu_i(A) = 1$, i.e., i is fundamentally certain about A under the first representation. There are two cases to consider. First, suppose $\mu_j(A) = 1$, for all $j \in E(i)$. Then, $\pi_j(A) = 1$, for all $j \in E(i)$, from which it follows that $\tilde{\pi}_j(A) = 1$, for all $j \in E(i) = \tilde{E}(i)$. Accordingly, $\tilde{\mu}_j(A) = 1$, for all $j \in \tilde{E}(i)$, i.e., $\tilde{\mu}_i(A) = 1$. Second, consider any A for which $\mu_i(A) = 1$, but $\mu_j(A) < 1$, for some $j \in E(i)$. Call such an event Class 2 μ_i-FC. For such an event $\pi_i(A) < 1$ and $\overline{\pi}_i(A) < 1$. To see that class 2 FC events—which exist given Condition 2—are invariant across representations, note that rearranging the equation determining i's behavioral probabilities, we get that for any event B with $\pi_i(B) > 0$,

$$\frac{\pi_i(B)}{\overline{\pi}_i(B)} = \alpha_i \frac{\mu_i(B)}{\overline{\pi}_i(B)} + (1 - \alpha_i) \tag{A.1.1}$$

From this equation, we can see that $\frac{\pi_i(B)}{\overline{\pi}_i(B)} = 1 - \alpha_i$, for any event B that is the complement of a Class 2 μ_i-FC event. On the other hand for any other event B with $\pi_i(B) > 0$, $\frac{\pi_i(B)}{\overline{\pi}_i(B)} > 1 - \alpha_i$. Accordingly, the set of Class 2 μ_i-FC events are given by

$$\left\{ A : A^C \in \operatorname*{argmin}_{B:\pi_i(B)>0} \frac{\pi_i(B)}{\overline{\pi}_i(B)} \right\}$$

Likewise, under the second representation, the set of class 2 $\tilde{\mu}_i$-FC events are given by

$$\left\{ A : A^C \in \operatorname*{argmin}_{B:\tilde{\pi}_i(B)>0} \frac{\tilde{\pi}_i(B)}{\overline{\tilde{\pi}}_i(B)} \right\}$$

Since, $\pi_j = \tilde{\pi}_j$, for any $j \in I$, it follows that the set of Class 2 FC events are the same under both representations. Hence, $\mu_i(A) = 1$ iff $\tilde{\mu}_i(A) = 1$, for all i.

α_i is determined uniquely by fundamentally certain events

By condition 2, for each i, there exists a Class 2 fundamentally certain event, i.e., there exists A with $\mu_i(A) = 1$ and $\pi_i(A) < 1$, or equivalently, $\mu_i(A^c) = 0$ and $\pi_i(A^c) > 0$.

Then:

$$\begin{aligned}
\frac{\pi_i(A^c)}{\overline{\pi}_i(A^c)} &= \alpha_i \frac{\mu_i(A^c)}{\overline{\pi}_i(A^c)} + (1-\alpha_i) \\
&= 1-\alpha_i \\
\implies \alpha_i &= 1 - \frac{\pi_i(A^c)}{\overline{\pi}_i(A^c)} = 1 - \frac{\tilde{\pi}_i(A^c)}{\tilde{\overline{\pi}}_i(A^c)} = \tilde{\alpha}_i,
\end{aligned}$$

where the last equality follows from our conclusion above that the set of Class 2 FC events are the same across the two representations.

μ_i is unique

Behavioral probabilities are given by:

$$\pi_i(A) = \alpha_i \mu_i(A) + (1-\alpha_i)\overline{\pi}_i(A)$$

We can rearrange this to write:

$$\mu_i(A) = \frac{1}{\alpha_i}\left(\pi_i(A) - (1-\alpha_i)\overline{\pi}_i(A)\right)$$

From the uniqueness of α_i, π_i, and $\overline{\pi}_i$ established above, it follows that μ_i must be unique. This completes the proof of sufficiency.

We next show the **necessity** of Conditions 1 and 2 for unique identification. To that end, suppose $\left(E = \langle E_1, ..., E_k \rangle, (u_i)_{i\in I}, (\mu_i)_{i\in I}, (\alpha_i)_{i\in I}\right)$ is the unique echo chamber representation of $(\succcurlyeq_i)_{i\in I}$.[6] We first prove that Condition 2 holds, followed by Condition 1.

Condition 2

Suppose condition 2 is violated. That is, there exists some individual i such that $\mu_i(A) = 0$ if and only if $\mu_j(A) = 0$, for all $j \in E(i)$. This, in turn, implies that $\mu_i(A) = 0$ if and only if $\pi_i(A) = 0$. We will show that this leads to a contradiction of unique identification. In particular, we will show that we can find $\tilde{\alpha}_i$ and $\tilde{\mu}_i$ such that $\big(E = \langle E_1, ..., E_k\rangle, (u_i)_{i\in I}$, $(\tilde{\mu}_i, \mu_{-i}), (\tilde{\alpha}_i, \alpha_{-i})\big)$ also represents $(\succcurlyeq_i)_{i\in I}$.

First, we identify the desired $\tilde{\alpha}_i$. For this, pick $\tilde{\alpha}_i$ in the interval $\left(1 - \min \frac{\pi_i(A)}{\overline{\pi}_i(A)}, 1\right)$, $\tilde{\alpha}_i \neq \alpha_i$. This is well-defined as $\min \frac{\pi_i(A)}{\overline{\pi}_i(A)} > 0$, since $\pi_i(A) = 0$ if and only if $\overline{\pi}_i(A) = 0$ (Further, the number of events is finite and hence the minimum exists). First, for events such that $\pi_i(A) = \mu_i(A) = 0$ or $\pi_i(A) = \mu_i(A) = 1$, define $\tilde{\mu}_i(A) = \mu_i(A)$. Clearly, for these events,

$$\pi_i(A) = \tilde{\alpha}_i \tilde{\mu}_i(A) + (1-\tilde{\alpha}_i)\overline{\pi}_i(A)$$

[6] Of course, utilities are unique up to positive affine transformation.

Next, for any event A for which $\pi_i(A) > 0$, define $\tilde{\mu}_i$ as:

$$\tilde{\mu}_i(A) = \frac{\pi_i(A) - (1-\tilde{\alpha}_i)\overline{\pi}_i(A)}{\tilde{\alpha}_i}$$

Note that $\tilde{\alpha}_i > 1 - \min \frac{\pi_i(A)}{\overline{\pi}_i(A)}$, or, $1 - \tilde{\alpha}_i < \min \frac{\pi_i(A)}{\overline{\pi}_i(A)}$. Accordingly, $1 - \tilde{\alpha}_i < \frac{\pi_i(A)}{\overline{\pi}_i(A)}$, for all A such that $\pi_i(A) > 0$. Hence, for all such A, we get $(1-\tilde{\alpha}_i)\overline{\pi}_i(A) < \pi_i(A)$. Along with the fact that $\tilde{\alpha}_i \in (0,\, 1)$, this implies that $\tilde{\mu}_i(A) \geq 0$. Now considering this formulation for singleton events, we get $\sum_{s\in S} \tilde{\mu}_i(s) = \frac{1}{\tilde{\alpha}_i}\big(\sum_{s\in S}\pi_i(s) - (1-\tilde{\alpha}_i)\sum_{s\in S}\overline{\pi}_i(s)\big) = \frac{1-(1-\tilde{\alpha}_i)}{\tilde{\alpha}_i} = 1$. Thus, noting that $\tilde{\mu}_i(A) = \sum_{s\in A}\tilde{\mu}_i(s)$, allows us to conclude that $\tilde{\mu}_i$ is a valid probability distribution. Finally, rearranging the above equation gives us that

$$\pi_i(A) = \tilde{\alpha}_i\tilde{\mu}_i(A) + (1-\tilde{\alpha}_i)\overline{\pi}_i(A)$$

This, therefore, establishes that $\big(E = \langle E_1,\, ...,\, E_k\rangle,\, (u_i)_{i\in I},\, (\tilde{\mu}_i, \mu_{-i}),\, (\tilde{\alpha}_i, \alpha_{-i})\big)$ also represents $(\succsim_i)_{i\in I}$, contradicting the statement of unique identification.

Condition 1

Assume condition 1 is violated. Then there exist two echo chambers, say E_1 and E_2 w.l.o.g., such that for every event for which $\mu_i(A) = 1$ for all $i \in E_1$, $\mu_j(A) = 1$ for all $j \in E_2$. This implies that $\pi_i(A) = 1$ if and only if $\pi_j(A) = 1$ for all $i \in E_1$ and $j \in E_2$. Now consider $\tilde{E}_1 = E_1 \cup E_2$. Let all other echo chambers remain unchanged, so $\tilde{E}_k = E_k$. Then this new echo chamber satisfies $\pi_i(A) = 1$ if and only if $\pi_j(A) = 1$ for all $i,\, j \in \tilde{E}_1$. Given this new chamber, we have a newly defined $\tilde{\overline{\pi}}_i$ for all $i \in \tilde{E}_1$. Then like done previously, we can define $\tilde{\alpha}_i \in (\underline{\alpha}_i,\, 1)$ where $\underline{\alpha}_i = 1 - \min \frac{\pi_i(A)}{\tilde{\overline{\pi}}_i(A)}$, and accordingly a new $\tilde{\mu}_i$ for each $i \in \tilde{E}_1$, such that these new parameters represent $(\succsim_i)_{i\in I}$. This contradicts exact identification, thus completing the proof.

A.2 Proof of Theorem 2

We first show that the axioms are sufficient for a sharp echo chamber representation.

Step 1. Defining behavioral beliefs: To begin with, A1 implies that, for each $i \in I$, $\succsim_i$ has a subjective expected utility representation. That is, there exists a function $u_i : Z \to \mathbb{R}$ and a probability measure π_i on S, such that the function $U_i : H \to \mathbb{R}$ given by

$$U_i(f) = \sum_{s\in S}\pi_i(s)u_i(f(s))$$

represents $\succsim_i$ for each i. Let π_i denote the behavioral beliefs of an individual i.

Step 2. Defining chambers: Next, we define chambers. Partition the set of individuals into two sets: $I_1 = \{i \in I : \text{there exists } A \subsetneq S \text{ such that } \pi_i(A) = 1\}$, and $I_2 = I \setminus I_1$. Denote by $E(i)$ a chamber containing i. For all $i \in I$ define $E(i) := \{j \in I : \pi_i(A) = 1 \iff \pi_j(A) = 1, \text{for any event } A\}$. Notice that by this definition, all $j \in I_2$ form a chamber within themselves. That is $E(i) = I_2$ for all $i \in I_2$.

A2 implies that $\forall i, \exists j$ such that $\pi_i(A) = 1 \iff \pi_j(A) = 1$. Given how the chambers are defined, by A2, we get that for any i and $E(i)$, $\exists j \in E(i)$. Note that this is also true for I_2. To see this suppose $i' \in I_2$. Since $\pi_i(S) = 1 \; \forall i \in I$, by A2, for i' there must exist $j \in I$, such that $\pi_j(A) < 1$, $\forall A \subsetneq S$. Thus implying $j \in I_2$. Therefore, all chambers are non-singleton.

Step 3. Defining core beliefs and immunity from influence parameter: Now, for every individual $i \in I$, we define their core probability measure, μ_i, and their immunity from influence parameter, α_i. For all $i \in I$ and for all $A \subseteq S$ such that $\pi_i(A) = 1$, define $\mu_i(A) = 1$. And, for all $i \in I$ and for all $A \subseteq S$ such that $\pi_i(A) = 0$, define $\mu_i(A) = 0$. Using A3, for every i, $\exists A^0 \subsetneq S$ that i finds incredulous but $\exists j \in R(i)$ for whom A^0 is not incredulous. That is $\frac{\pi_i(A^0)}{\overline{\pi}_i(A^0)} = \min_{B \subsetneq S} \frac{\pi_i(B)}{\overline{\pi}_i(B)}$. For each such i, define $\alpha_i = 1 - \frac{\pi_i(A^0)}{\overline{\pi}_i(A^0)}$ and $\mu_i(A^0) = 0$.

Now for events with $\pi_i(A) \in (0, 1)$ that i does not find incredulous, we have that

$$\frac{\pi_i(A)}{\overline{\pi}_i(A)} > \frac{\pi_i(A^0)}{\overline{\pi}_i(A^0)} \iff \alpha_i > 1 - \frac{\pi_i(A)}{\overline{\pi}_i(A)}$$

That is, $\pi_i(A) - (1 - \alpha_i)\overline{\pi}_i(A) > 0$. Now, for all such events A, define

$$\mu_i(A) = \frac{\pi_i(A) - (1 - \alpha_i)\overline{\pi}_i(A)}{\alpha_i}$$

Note that, for such events, $\mu_i(A) > 0$ since $\pi_i(A) - (1 - \alpha_i)\overline{\pi}_i(A) > 0$ and $\alpha_i > 0$. Further, since $\sum_s \pi_i(s) = 1$ and $\sum_s \overline{\pi}_i(s) = 1$, we have that $\sum_s \mu_i(s) = 1$. Since $\mu_i(A) = \sum_{s \in A} \mu_i(s)$, we have $\mu_i(A) \in (0, 1)$.

Step 4. Establishing the echo chamber representation: We now establish the echo chamber representation using the objects defined in the previous steps.

For all non-incredulous events with $\pi_i(A) \in (0, 1)$, we have

$$\mu_i(A) = \frac{\pi_i(A) - (1 - \alpha_i)\overline{\pi}_i(A)}{\alpha_i}$$

$$\iff \pi_i(A) = \alpha_i \mu_i(A) + (1 - \alpha_i)\overline{\pi}_i(A)$$

For all incredulous events A^0, we have $\mu_i(A^0) = 0$ and $\alpha_i = 1 - \frac{\pi_i(A^0)}{\overline{\pi}_1(A^0)}$. This implies that

$$\pi_i(A^0) = \alpha_i \mu_i(A^0) + (1 - \alpha_i)\overline{\pi}_i(A^0)$$

For events with $\pi_i(A) = 1$, we have that $\mu_i(A) = 1$ and $\pi_j(A) = 1$ for all $j \in E(i) \iff \overline{\pi}_i(A) = 1$. Therefore, for such events,

$$\pi_i(A) = \alpha_i \mu_i(A) + (1 - \alpha_i)\overline{\pi}_i(A) = 1$$

Finally, for events with $\pi_i(A) = 0$, we have that $\mu_i(A) = 0$ and $\pi_i(A^c) = 1 \iff \pi_j(A^c) = 1$ for all $j \in E(i) \iff \overline{\pi}_i(A^c) = 1 \iff \overline{\pi}_i(A) = 0$. Hence, we have

$$\pi_i(A) = \alpha_i \mu_i(A) + (1 - \alpha_i)\overline{\pi}_i(A) = 0$$

This completes the proof of the sufficiency of axioms for an echo chamber representation.

Step 5. Establishing this is a sharp echo chamber representation: We now show that the echo chamber representation established is such that conditions 1 and 2 hold.

First, consider condition 1. Note that all agents find S to be fundamentally certain because $\pi_i(S) = 1 \implies \mu_i(S) = 1$ for all $i \in I$. Suppose towards a contradiction of condition 1, there exist chambers E and E' such that they have the same events they are fundamentally certain about. This implies that for every $i \in E$, $\pi_i(A) = 1 \iff \pi_j(A) = 1$ for all $j \in E'$. This would lead to a contradiction that all $i \in E$ and all $j \in E'$ are in fact in the same chamber! Therefore, condition 1 holds.

Next, consider condition 2. Since the echo chamber representation established above satisfies A3, we have that for all E and for all $i \in E$, there exists an event A that i finds to be incredulous, but some $j \in E$ does not. In our representation, for such events, $\mu_i(A) = 0$ but $\mu_j(A) \neq 0$. Which is equivalent to saying that $\exists A^c$ such that $\mu_i(A^c) = 1$ but $\mu_j(A^c) \neq 1$ for some $j \in E$. Therefore, condition 2 holds.

This completes the proof of the sufficiency of axioms for a sharp echo chamber representation.

We now show that the axioms A1-A3 are necessary for the representation.

Axiom A1: We start by recognizing that the preferences of each DM, i, are represented by a subjective expected utility function with respect to the probability distribution π_i. As such, $\succcurlyeq_i$ must satisfy A1.

<u>Axiom A2</u>: Given the representation, for any i we have that $\pi_i(A) = 1 \iff \pi_j(A) = 1$ for all $j \in E(i)$. Since the chambers are non-singleton, we get that for all $i \in I$, $\exists j \in I$ such that $\pi_i(A) = 1 \iff \pi_j(A) = 1$.

<u>Axiom A3</u>: By condition 2, for every E and $i \in E$ there exists an event A such that $\mu_i(A) = 1$ and $\mu_j(A) < 1$ for some $j \in E$. That is, $\mu_i(A^c) = 0$ and $\mu_j(A^c) > 0$. This implies, $\frac{\pi_i(A^c)}{\bar{\pi}_i(A^c)} \leq \frac{\pi_i(B)}{\bar{\pi}_i(B)}$ for all events B. Hence, i finds A^c incredulous. Again, by condition 2, there exists an event A' that j finds fundamentally certain (i.e. $\mu_j(A'^c) = 0$) but someone in E does not. This implies that $\frac{\pi_j(A'^c)}{\bar{\pi}_j(A'^c)} < \frac{\pi_j(A^c)}{\bar{\pi}_j(A^c)}$. That is, j does not find A^c incredulous. Thus, axiom A3 is satisfied.

A.3 Proof of Proposition 1

Proposition 1 can be proved as follows.

$$\begin{aligned} \pi_i(\sigma_i^{-1}(c^i) \,|\, c^i) = 1 &\iff \bar{\pi}_i(\sigma_i^{-1}(c^i) \,|\, (c_1, ..., c_k)) = 1 \\ &\iff \mu_j(\sigma_i^{-1}(c^i) \,|\, c^j) = 1 \,\forall\, j \in E(i) \end{aligned}$$

A.4 Proof of Corollary 1

Take some state s such that $\mu_i(s) > 0$ and $s \in \sigma_i^{-1}(c^i)$. Then $\mu_i(s \,|\, c^i) > 0$. This means, however, that $\pi_j(s \,|\, c^j) > 0$ for all $j \in E(i)$. Then, if $\pi_j(\sigma_j^{-1}(c^j) \,|\, c^j) = 1$ for all $j \in E(i)$, it must be that $s \in \sigma_j^{-1}(c^j)$. This implies that $\sigma_i^{-1}(c^i) \setminus S_0^i \subseteq \sigma_j^{-1}(c^j)$ for all $j \in E(i)$. By a symmetric argument, we can extend this to any i, $j \in E_k$ some echo chamber.

Let us now assume that $\sigma_i^{-1} \setminus S_0^i \subseteq \sigma_j^{-1}(c^j)$ for all i, $j \in E_k$. Note that $\mu_i(\sigma_i^{-1}(c^i) \setminus S_0^i \,|\, c^i) = 1$ for all i, as $\mu_i(s_0 \,|\, c^i) = 0$ if $\mu_i(s_0) = 0$. By the antecedent, $\mu_i(\sigma_j^{-1}(c^j) \,|\, c^i) = 1$ for all i, j as this is a superset of $\sigma_i^{-1}(c^i) \setminus S_0^i$. Then it must mean that $\bar{\pi}_i(\sigma_i^{-1}(c^i)) = 1$ for all $i \in E_k$. This completes the proof.

A.5 Proof of Proposition 2

Note first that we can write the polarization between E_k and E_l as follows:

$$\begin{aligned}2\rho_{(k,l)} &= \sum_{s\in S} |\overline{\pi}_{E_k}(s) - \overline{\pi}_{E_l}(s)| \\ &= \sum_{s\in S} \max\{\overline{\pi}_{E_k}(s) - \overline{\pi}_{E_l}(s),\, \overline{\pi}_{E_l}(s) - \overline{\pi}_{E_k}(s)\} \\ &= \sum_{s\in S} (\max\{\overline{\pi}_{E_k}(s) - \overline{\pi}_{E_l}(s),\, 0\} + \max\{0,\, \overline{\pi}_{E_l}(s) - \overline{\pi}_{E_k}(s)\})\end{aligned}$$

Since $\sum_{s\in S} \overline{\pi}_{E_k}(s) = \sum_{s\in S} \overline{\pi}_{E_l}(s) = 1$, we can rewrite ρ as:

$$\begin{aligned}\rho_{(k,l)} &= \sum_{s\in S} \max\{\overline{\pi}_{E_k}(s) - \overline{\pi}_{E_l}(s),\, 0\} \\ &= \sum_{s\in S} \max\{-\overline{\pi}_{E_l}(s),\, -\overline{\pi}_{E_k}(s)\} + \overline{\pi}_{E_k}(s) \\ &= 1 - \sum_{s\in S} \min\{\overline{\pi}_{E_k}(s),\, \overline{\pi}_{E_l}(s)\}\end{aligned}$$

If $\overline{\pi}_{E_l}(s^*) = 0$ then the given $\hat{\sigma}_i$ cannot be defined, so assume it is strictly greater than 0. Since $s^* \notin \hat{S}_k(\boldsymbol{c}_k)$, it must be that $\overline{\pi}'_{E_k}(s^*) = 0$. Then $\rho^1_{(k,l)}(s^*) = 1$, whereas $\rho^2_{(k,l)}(s^*) = 0$. Let $\overline{\pi}'_{E_l}$ and $\overline{\pi}''_{E_l}$ denote the posterior cluster probability distribution upon facing the signals $\boldsymbol{\sigma}_k$ and $\hat{\boldsymbol{\sigma}}_k$. Note that $\rho^1_{(k,l)}$ can be written as follows:

$$\rho^1_{(k,l)} = 1 - \sum_{s\in S} \min\left\{\overline{\pi}'_{E_k}(s), \overline{\pi}'_{E_l}(s)\right\}$$

Note that for $\rho^2_{(k,l)}$, on the other hand:

$$\rho^2_{(k,l)} = 1 - \sum_{s\in S} \min\left\{\overline{\pi}'_{E_k}(s),\, \overline{\pi}''_{E_l}(s)\right\}$$

Since $\overline{\pi}'_{E_k}(s^*) = 0$, we can remove s^* from the sum. Since $\overline{\pi}''_{E_l}(s) \geq \overline{\pi}'_{E_l}(s)$ for all $s \neq s^*$, it must be that $\min\left\{\overline{\pi}'_{E_k}(s),\, \overline{\pi}''_{E_l}(s)\right\} \geq \min\left\{\overline{\pi}'_{E_k}(s),\, \overline{\pi}'_{E_l}(s)\right\}$. Then, $\rho^1_{(k,l)} \geq \rho^2_{(k,l)}$.

References

Acemoglu, Daron, Kostas Bimpikis, and Asuman Ozdaglar. 2014. "Dynamics of information exchange in endogenous social networks." *Theoretical Economics* 9 (1): 41–97.

Arceneaux, Kevin, and Martin Johnson. 2015. "More a Symptom Than a Cause: Polarization and Partisan News Media in America." In *American Gridlock: The Sources, Character, and Impact of Political Polarization,* edited by Antoine Yoshinaka and James A Thurber, 309–336. Cambridge University Press.

Baccara, Mariagiovanna, and Leeat Yariv. 2016. "Choosing peers: Homophily and polarization in groups." *Journal of Economic Theory* 165 (September): 152–178.

Barberá, Pablo. 2020. "Social Media, Echo Chambers, and Political Polarization." In *Social Media and Democracy: The State of the Field and Prospects for Reform,* edited by Joshua A Tucker and Nathaniel Persily, 34–55. Cambridge University Press.

Borah, Abhinash, and Christopher Kops. 2018. "Choice via social influence." *Unpublished manuscript.*

Chambers, Christopher P, Tugce Cuhadaroglu, and Yusufcan Masatlioglu. 2019. "Behavioral influence." *Unpublished manuscript.*

Chambers, Christopher P, Yusufcan Masatlioglu, and Christopher Turansick. 2021. "Influence and Correlated Choice." *arXiv preprint arXiv:2103.05084.*

Cookson, J Anthony, Joseph E Engelberg, and William Mullins. 2023. "Echo Chambers." *The Review of Financial Studies* 36 (2): 450–500.

Cuhadaroglu, Tugce. 2017. "Choosing on influence." *Theoretical Economics* 12 (2): 477–492.

Eyster, Erik, and Matthew Rabin. 2010. "Naïve Herding in Rich-Information Settings." *American Economic Journal: Microeconomics* 2 (4): 221–243.

Fershtman, Chaim, and Uzi Segal. 2018. "Preferences and Social Influence." *American Economic Journal: Microeconomics* 10 (3): 124–142. ISSN: 1945-7669. https://doi.org/10.1257/mic.20160190.

Golub, Benjamin, and Matthew O Jackson. 2010. "Naïve Learning in Social Networks and the Wisdom of Crowds." *American Economic Journal: Microeconomics* 2 (1): 112–149.

Kashaev, Nail, and Natalia Lazzati. 2019. "Peer effects in random consideration sets." *arXiv preprint arXiv:1904.06742.*

Lazzati, Natalia. 2020. "Codiffusion of Technologies in Social Networks." *American Economic Journal: Microeconomics* 12 (4): 193–228.

Levy, Gilat, and Ronny Razin. 2015a. "Correlation Neglect, Voting Behavior, and Information Aggregation." *American Economic Review* 105 (4): 1634–1645.

Levy, Gilat, and Ronny Razin. 2015b. "Does Polarisation of Opinions Lead to Polarisation of Platforms? The Case of Correlation Neglect." *Quarterly Journal of Political Science* 10 (3): 321–355.

———. 2019. "Echo Chambers and Their Effects on Economic and Political Outcomes." *Annual Review of Economics* 11 (1): 303–328.

Martinez, Gabriel, and Nicholas H. Tenev. 2022. *Optimal Echo Chambers,* October.

McCarty, N, K T Poole, and H Rosenthal. 2006. *Polarized America: The Dance of Ideology and Unequal Riches.* MIT Press.

Savage, Leonard J. 1954. *The Foundations of Statistics.* John Wiley & Sons.

Wojcieszak, Magdalena, Andreu Casas, Xudong Yu, Jonathan Nagler, and Joshua A. Tucker. 2021. "Echo chambers revisited: The (overwhelming) sharing of in-group politicians, pundits and media on Twitter." *Unpublished manuscript.*