Learning from Shared News: When Abundant Information Leads to Belief Polarization*

Renee Bowen[†], Danil Dmitriev[‡], Simone Galperti[§]

June 17, 2020

Abstract

We study social learning in networks with agents who may share information selectively. This creates echo chambers with heterogeneous news diets across agents. We show that this can cause agents' beliefs to diverge—contrary to standard learning results—if (and only if) agents hold even minor misperceptions about selective sharing and information quality is sufficiently low. This belief polarization is exacerbated by abundant information. Specifically, when quantity of external information grows indefinitely different agents can hold extreme opposite beliefs. Polarization can also worsen when the echo chambers become larger, despite providing the agents with more information. We show how information aggregation can mitigate polarization—even if it involves losing some information.

*We thank our discussant, Myles Ellis, Joel Sobel and participants in the UCSD Theory and Behavioral Workshop and the PhDEI workshop for helpful comments and suggestions

[†]UC San Diego and NBER, trbowen@ucsd.edu

[‡]UC San Diego, ddmitrie@ucsd.edu

[§]UC San Diego, sgalperti@ucsd.edu

1 Introduction

Polarization has been linked to negative economic outcomes, including, political gridlock, low legislative quality, economic inequality, and a reduction in security of property rights, trust, investment, and growth (Bishop, 2009; Barber and McCarty, 2015; McCarty et al., 2009; Bartels, 2008; Gilens, 2012; Zak and Knack, 2001; Keefer and Knack, 2002). In recent decades, authors have noted increasing polarization in policymaking, media, and in public opinions (Pew Research Center (2014, 2020); Tucker et al. (2019); Desmet and Wacziarg (2018)). According to Alesina et al. (2020) "Evidence is growing that Americans are polarized not only in their views on policy issues and attitudes towards government and society, but also about their perceptions of the same, factual reality." Economists have, thus, been concerned with understanding the determinants of belief polarization.

Several authors have made a connection between increasing polarization of the American public and growing use of the Internet as a source of information (Periser, 2011; Sunstein, 2017; Azzimonti and Fernandes, 2018; Tucker et al., 2019). Some of these studies have highlighted the role of fake news, bots, and bad actors in driving polarization, leading to discussions about regulating social media to minimize these elements.[1] This raises the question of whether such regulation will completely alleviate the problem of polarization, even if successful. More generally, even in the absence of *misinformation*, can the way in which people consume and share information on networks lead to polarization? In addition, can the preponderance of information introduced by the expansion of such networks lead to more polarization? This paper provides a theoretical framework to answer these questions.

Evidence shows that information flows on social networks in specific ways. First, people tend to share information *selectively* in networks (Shin and Thorson, 2017; Weeks et al., 2017; Shin et al., 2018; Pogorelskiy and Shum, 2019). For example, people share only the information that favors their preferred political candidate, or their preferred views on the importance of protecting the environment. Second, agents are likely to have friends who select different kinds of information to share. Nonetheless, there is a wealth of evidence on the existence of echo chambers or media bubbles (see Levy and Razin, 2019a), which feature some *imbalance* in the kind of information provided by friends.

It seems intuitive that unbalanced selective sharing of information may lead to polarization. After all, if a person takes at face value what her friends say and they selectively support one view on an issue, her opinion can be swayed accordingly. However, if a person is fully aware of how her friends select what to share, she will adjust for this selection (or ignore the information received altogether) and her belief will not be distorted. Thus, what we assume about how aware people are about the selectivity of shared news is crucial. Pogorelskiy and Shum (2019) study experimentally how people exchange and perceive information in networks and find that people do share information selectively. They also find that agents often fail to fully account for the selective nature of received information. It is unrealistic

[1] See, for example, "Should the Government Regulate Social Media?", Wall Street Journal, June 25, 2019 and "Facebook Throws More Money at Wiping Out Hate Speech and Bad Actors", Wall Street Journal, May 15, 2018.

to expect agents to be completely naive about selective sharing, but they also seem to not fully take it into account either. Thus, our model allows for some *misperception* of selective sharing consistent with this evidence.

We consider a simple model of learning from shared information on a network. Agents learn about some underlying state of the world, which can take one of two values. The state realizes once and for all in the first period. In every subsequent period, each agent receives an independent signal about the state. She can then either share it with her neighbors in the network (hereafter, called friends) or remain silent. This rules out fake news in the form of fabricated signals. The choice of when to remain silent gives rise to selective sharing as follows. While some agents—called normal—share every signal, others share only signals that favor a specific state. We call these types of agents dogmatic. To fix ideas, consider the issue of whether or not to vaccinate children. Some individuals (possibly a tiny minority) may dogmatically support a particular view and share only favorable information about that view, while other agents simply share any information they have. When receiving signals from friends, the agents are aware of selective sharing, but *misperceive* its degree. We model such misperceptions in a way that renders agents partially unresponsive to selective sharing (as found in Pogorelskiy and Shum (2019)), prevent it from unraveling, and is inspired by the psychology literature on "illusory superiority" and the "better-than-average" effect (Cross, 1977; Svenson, 1981; Odean, 1998; Zuckerman and Jost, 2001). In a nutshell, while agents correctly interpret the signals they received, they misinterpret their friends' silence by allowing for the possibility that they simply have not received information. This is akin to an agent assuming that her friends do not read the newspaper as often as they do. While agents in our model update priors according to Bayes' rule, they do so with a (common) misspecified model of selective sharing.

We are interested in how agents learn in this environment and the overall effects on their belief polarization. Consider any agent and suppose her network has an imbalance—however small—in the composition of her dogmatic friends. That is, besides possibly many normal friends, she has more dogmatic friends who favor one state over the other. We refer to this composition as the agent's echo chamber. To examine short-run learning (i.e., with little information), we consider the agent's *expected* posterior belief after one period of signals. We find that this expected posterior differs from the agent's prior in the same direction of the imbalance in her echo chamber, at least when the quality of the original signals is sufficiently low (i.e., they are sufficiently uninformative about the state). Note that this departs from correctly specified Bayesian updating, under which the expected posterior always equals the prior. Intuitively, rather than correctly interpreting her friends' silence as suppressing information against their view, the agent incorrectly assumes that her friends may be uninformed. Thus, perhaps counterintuitively, it is the agent's misperception that her friends are less informed that gives them the power to distort her belief. Note that the imbalance in echo chambers can be arbitrarily small, yet counteract the impact of many unbiased signals depending on the quality of information. This demonstrates the role of information quality in generating belief polarization: If external information is very precise, it would be impossible for friends to bias an agent's beliefs.

We also examine long-run learning (i.e., with abundant information). We find that if the quality of external information is sufficiently low, then with certainty an agent's asymptotic belief assigns probability one to the state corresponding to the direction of imbalance in her echo chamber, *irrespective* of the true state. That is, the agent is essentially guaranteed to be pulled towards the preferred state of her echo chamber. This conclusion may be counter-intuitive, as more information does not alleviate, but rather exacerbates incorrect learning. We also show that greater misperception and more unbalanced networks can exacerbate incorrect learning. To the extent that the Internet has served to increase the frequency of information arrival where (as a result) each bit of information may be of low quality, our analysis suggests that this may be a driver of incorrect learning, even without misinformation (i.e., false information). In turn, to the extent that agents live in echo chambers with imbalances favoring different states, the previous results suggest mechanisms that can lead people's beliefs to diverge, thus giving rise to polarization.

We build on the link between unbalanced selective sharing and misperception thereof to further understand how networks interact with information to drive polarization. A natural question is whether both aspects are necessary for polarization to appear. We show that if *either* there is no imbalance in the network *or* there is no misperception, beliefs converge to the true state and the expected posterior is equal to the prior after every round of updating.

Next we show that the expansion of connections as networks grow can be another driver of polarization. We identify sharp conditions of the growth rates of the different types of friends under which network expansion curbs or exacerbates polarization. For instance, we find that by simply scaling any agent's network while keeping the proportions of her normal and dogmatic friends raises the threshold of information quality below which polarization occurs. That is, for a fixed level of information quality, it is possible that a person whose belief is not polarized in a smaller network become polarized when her network expands.

Finally, we explore interventions that may reduce polarization in our networks. We find that allowing signals to be aggregated before being released to the agents can alleviate polarization. Despite losing information relative to the set of original signals, the aggregated signal has higher quality than each original signal. Thus, for each level of misperception, there exists a level of aggregation which ensures that all agents learn correctly, thereby eliminating polarization. This result suggests a rationale for public information aggregators or for committing to releasing information only rarely to ensure that it is of high quality.

Related Literature

The economics literature discusses at least three possible causes for belief polarization. One strand of papers that is most closely related (e.g., Levy and Razin, 2019b; Hoffmann et al., 2019; Enke et al., 2019) has studied polarization arising from behavioral biases. Our main contribution to this literature is to highlight the importance of the network structure coupled with selective sharing and misperception in generating polarization. Another strand studies polarization arising as a result of heterogeneity in preferences (see Dixit and Weibull (2007) and Pogorelskiy and Shum (2019)). Such heterogeneity of preferences would exacerbate the polarization we find, but is not required for generating our preliminary results about polarization. Finally, another reason for polarization that has been studied is biased or multidimensional information sources (e.g. Mullainathan and Shleifer, 2005; **?**; Levendusky, 2013; Conroy-Krutz and Moehler, 2015; Reeves et al., 2016; Perego and Yuksel, 2018), where bias usually comes from media competition over viewers. In our model, the external information sources are assumed to be unbiased when reporting information and bias is generated in the network. **?** show that one-dimensional opinions can diverge with two-dimensional information. We provide conditions under which one-dimensional opinions diverge with one-dimensional information, based on selective sharing and misperception.

This project also fits into the growing literature on the effects of model misspecification on social learning. Bohren (2016), Bohren and Hauser (2018), and Frick et al. (2019) analyze how model misspecification impacts long-term learning in environments where agents learn from private signals and the actions of other agents. In particular, Bohren and Hauser (2018) study when agents with different, yet reasonable, models have no limit beliefs (i.e. beliefs cycle) or have different limit beliefs (disagreement). The result about disagreement is close in spirit to what we find, but the driving mechanism is fundamentally different: In our model, all agents have the same misspecified model of the world and polarization results from exposure to selectively shared information through the social network. We also emphasize the important role played by the quality of information, which may suggest a simple way to alleviate polarization by aggregating signals. Molavi et al. (2018) study long-run learning on social networks when non-Bayesian agents exhibit imperfect recall. They show that such agents may overweigh evidence encountered early on relative to later information, which can lead to mislearning in the limit. Again, while this conclusion is related to ours, unlike these authors, we assume that agents update beliefs via Bayes' rule, albeit with a misspecified model of the world.

A key element of our model is the network structure. There is a large literature on both Bayesian (Acemoglu et al., 2010; Pogorelskiy and Shum, 2019; Spiegler, 2019) and non-Bayesian learning in networks (Golub and Jackson, 2010; DeMarzo et al., 2003; Azzimonti and Fernandes, 2018; Eyster and Rabin, 2010). One closely related paper in this literature is Levy and Razin (2019a), which considers the effect of what they call a "Bayesian Peer Influence" updating heuristic on the limit beliefs in the network. One of the main results shows how beliefs in society can become polarized as a result of agents using this updating rule. However, the meaning of polarization in that paper is different from ours: Instead of groups in society becoming polarized, it is the entire society's consensus that shifts towards extreme beliefs, whereas we provide conditions under which beliefs of agents in society diverge upon arrival of new information.

Recent empirical studies show that social media is an important source of news for people and can lead to divergence of beliefs and attitudes (e.g. Allcott and Gentzkow, 2017; Bursztyn et al., 2019; Mosquera et al., 2019; Levy, 2020). Still other empirical evidence by Boxell et al. (2018) suggests that the Internet does not drive polarization. Our model can predict in which environments we expect to see the Internet drive polarization. We, thus, contribute to this literature by providing a theoretical framework to better understand how social media can contribute to polarization and to guide future empirical investigations of this phenomenon.

The remainder of the paper is organized as follows. Section 2 presents our stylized model of information sharing in a network. Section 3 presents the main results on learning in the short and long run. Section 4 considers what happens when the network expands. Section 5 considers the polarization in the entire network. Section 6 demonstrates how aggregating signals can mitigate polarization. We conclude with a discussion of the results in Section 7.

2 Model

We provide a stylized model of information sharing in a network. Time is discrete and denoted by $t = 0, \ldots, T$, where T may be infinite. There are two possible states of the world, $\omega \in \{L, R\}$, and the state is determined by nature in period 0. For example, the state could represent whether the preservation of the environment requires lower or higher national spending than the current level, or whether vaccinations can be harmful. There is an exogenously fixed network of agents who seek to match the state. Each period $t \geq 1$, each agent privately receives a signal $s_{it} \in \{l, r\}$ about the state, which is i.i.d. across agents and time.[2] The information quality of s_{it} is denoted by q, namely, the probability that the signal matches the true state. Formally,

$$\begin{aligned} \mathbb{P}(s_{it} = l | \omega = L) &= \mathbb{P}(s_{it} = r | \omega = R) = q, \\ \mathbb{P}(s_{it} = r | \omega = L) &= \mathbb{P}(s_{it} = l | \omega = R) = 1 - q. \end{aligned}$$

We assume that $q > \frac{1}{2}$, so that the signals are at least partially informative. As q increases the signal becomes more precise, so we refer to q as the information *quality*.

Once an agent receives a signal, she can either stay silent or share the signal with her neighbors in the network, which we will call her friends. She cannot lie when sharing a signal, which rules out fake news. For example, an agent can choose to share a news article or not, but cannot edit the article's contents. In addition, she cannot choose with whom to share her information, i.e. no targeting: She either shares it with all her friends, or with none. By ruling out fake news and targeted sharing, we highlight the role of selective sharing in a baseline model to which these other effects can be added.

Our model aims to capture the evidence of selective sharing that Pogorelskiy and Shum (2019) and others find. To this end, we introduce three types of agents. We will call an agent *dogmatic left* if she shares only signals $s_{it} = l$ and *dogmatic right* if she only shares signals $s_{it} = r$. One interpretation is that such agents choose to share only information that supports the state they dogmatically believe in.[3] We will call an agent *normal* if she shares

[2] In reality, people receive news that are correlated. However, strong evidence suggests that people often neglect correlation, especially in network environments (Enke and Zimmermann (2017); Eyster et al. (2018); Pogorelskiy and Shum (2019)). Under correlation neglect, we can allow for arbitrary correlation between the agents' signals within each period without qualitatively changing the results.

[3] Dogmatic agents can be thought of as having degenerate beliefs that do not change with new information.They can be thought of as stubborn, narrow minded, individuals who blindly follow and promote their

any signal she receives. Each agent's type is fixed and exogenous. Formally, each period the left and right dogmatic types will transmit signals as follows:

$$a_L(s_{it}) = \begin{cases} \text{transmit}, & \text{if } s_{it} = l \\ \text{stay silent}, & \text{if } s_{it} = r \end{cases}$$

$$a_R(s_{it}) = \begin{cases} \text{stay silent}, & \text{if } s_{it} = l \\ \text{transmit}, & \text{if } s_{it} = r \end{cases} .$$

The timing of period $t \geq 1$ is as follows: (1) signals s_{it} are independently realized; (2) agent i observes s_{it} privately; (3) agent i shares s_{it} with her friends according to her type and transmission strategy; (4) agents update beliefs based on all signals they have received, including their own signal.

We are interested in studying the beliefs of normal agents. We will assume that all normal agents have a common prior that assigns probability $\pi \in (0,1)$ to state L. Although our results can be adjusted to allow for different priors, this assumption helps emphasize the drivers of polarization that we discover. Let i be any normal agent. The composition of her friends will play a key role. Suppose i has $d_{Li} \geq 0$ dogmatic-left friends, $d_{Ri} \geq 0$ dogmatic-right friends, and $n_i \geq 0$ normal friends. We will refer to i's composition of friends (d_{Li}, d_{Ri}, n_i) as her *echo chamber*, since the behavior of dogmatic friends to reinforce a particular view is similar to the common understanding of how echo chambers work on social networks. Our results will show that the composition of i's echo chamber has a crucial effect on her posterior beliefs.

We argued that, besides selective sharing, a key part of our analysis is to allow for misperceptions about it. Selective sharing involves suppressing some information that one receives and replacing it with silence. Thus, one plausible way to misperceive selective sharing in our setting is to misinterpret the silence of a friend as the possibility that she genuinely received no information (i.e., no signal). This form of misperception has another interpretation that is consistent with our analysis. A manifestation of the so-called "illusory superiority" or "better-than-average" heuristic is thinking that other people may be less informed than they actually are (see, for example, Cross, 1977; Svenson, 1981; Odean, 1998; Zuckerman and Jost, 2001). People often have unjustifiably favorable views of themselves relative to the population average or even in person-to-person comparisons on various characteristics, which may include how well informed they are or how good they are at getting and understanding information.

We model such misperceptions as follows. Each period, agent i observes her signal s_{it} and the signals shared by her friends. When updating, i knows the types of her friends and is aware of their transmission strategies. However, she thinks that they are less informed than they are: In her view, in every period each friend observes a signal only with probability $\gamma < 1$ and no signal with probability $1 - \gamma$. Thus, from i's perspective, each of her friends faces the following timing in period $t \geq 1$: (1) signals s_{it} are independently realized; (2)

convictions.

friend i privately observes s_{it} with probability γ; (3) if s_{it} is observed, it is either shared or not according to i's type and strategy; (4) the agents update beliefs. Note that γ is the same for all normal agents. This way of modeling misperceptions of selective news sharing may be viewed as a minimal departure from a standard Bayesian model, where the agents continue to be Bayesian but relative to a slightly misspecified model of the world. This way of modeling mispercecеption also ensures that selective sharing does not unravel, which would prevent it from having any effect on beliefs.

3 Learning in the short and long run

To examine the effects of selective sharing and misperceptions in the short run, our first result looks at the expected posterior belief of agent i that results from one round of updating (i.e., $T = 1$). Since the focus of this section is normal agent i, we drop all i subscripts. Let $\mu(\mathbf{s}^1)$ be normal agent i's Bayesian posterior probability assigned to state L given all the information she obtains after one period, which we denote by $\mathbf{s}^1$ (i.e., her signal, her friends' signals, and their silence). We show that whenever there is an imbalance in i's echo chamber, i's expected posterior conditional on *any* state will be distorted towards the state preferred by the majority of her dogmatic friends when the signal quality q is sufficiently low.

Proposition 1. *Fix an agent such that $d_L > d_R$. There exists $\bar{q}_1(d_L, d_R, n, \gamma) > \frac{1}{2}$ such that, if $q < \bar{q}_1(d_L, d_R, n, \gamma)$, then for $\omega \in \{L, R\}$*

$$\mathbb{E}\left[\mu(\mathbf{s}^1)|\omega\right] > \pi.$$

The proof of Proposition 1, and henceforth all omitted proofs, can be found in the Appendix.

An immediate implication of this result is that a defining property of Bayesian updating no longer holds in our setting. Since the expected posterior is distorted conditional on any true state of the world, the *unconditional* expected posterior will also be distorted in the same direction. This is in contrast to Bayesian updating with a correctly specified model of the world, where the expected posterior would be *equal* to the prior.

We provide some intuition for the result by first deriving agent i's posterior after one period. Consider that agent i has echo chamber (d_L, d_R, n). Let ℓ_L denote the number of left signals that dogmatic-left friends of i have received, and ℓ_R denote the number of left signals that dogmatic-right friends of i have received. Since there are d_L left-dogmatic and d_R right-dogmatic friends (and they all receive independent signals), ℓ_L is distributed as a Binomial random variable with probability $1-q$ and sample size d_L, whereas ℓ_R is distributed as a Binomial random variable with probability $1-q$ and sample size d_R. Note that i also receives a private signal and n independent signals from her normal friends. This is equivalent to receiving $n+1$ independent private signals. Letting ℓ_N denote the number of signals $s = l$ that are contained within them, ℓ_N is a Binomial random variable with probability $1-q$ and sample size $N = n+1$. Note that agent i's information $\mathbf{s}$ is summarized by (ℓ_L, ℓ_R, ℓ_N).

Given a fixed realization (ℓ_L, ℓ_R, ℓ_N), by Baye's rule agent i's posterior belief is given by

$$\mu^1(\omega = L|\ell_L, \ell_R, \ell_N) \;=\; \frac{\pi}{\pi + (1-\pi)Q^M\Gamma^S}, \tag{1}$$

where

$$\begin{aligned} Q &\equiv \frac{1-q}{q} \\ M &\equiv \ell_L - (d_R - \ell_R) + 2\ell_N - N \\ \Gamma &\equiv \frac{\gamma(1-q) + (1-\gamma)}{\gamma q + (1-\gamma)} \\ S &\equiv \ell_R - (d_L - \ell_L). \end{aligned}$$

Note that Q^M captures the correct interpretation of signals received. Agent i receives $\ell_L + \ell_N$ left signals from dogmatic left and normal friends. She receives $(d_R - \ell_R) + (N - \ell_N)$ right signals from dogmatic right and normal friends which counteract the left signals (and hence enter negatively in the exponent).

In addition to signals received, agent i observes silence from dogmatic left and right friends, which is incorrectly interpreted. This is captured by Γ^S. Agent i observes silence from ℓ_R dogmatic right friends, each of which agent i (incorrectly) attributes either to a left signal with probability γ or to no signal with probability $1-\gamma$. Similarly, agent i observes silence from $d_L - \ell_L$ dogmatic left friends and (incorrectly) attributes these to either a right signal or no signal. Counterintuitively, the incorrectly interpreted silence captured by Γ^S provides a "boost" to the correctly interpreted signals Q^M. Note also that if $\gamma = 1$ then $\Gamma = (1-q)/q$ so the boost disappears and silence is correctly interpreted.

Although the term Γ^S distorts the posterior, it is still unclear that the *average* distortion goes in any specific direction. For instance, they could inflate or deflate updating but still be correct on average. Intuitively, if i's echo chamber has more dogmatic-left than dogmatic-right friends ($d_L > d_R$), she tends to receive more signals that support L than R as the true state due to their selective sharing. However, i is fully aware of this and would be able to fully unravel the effect of selective sharing, except that her misperception complicates this. When hearing silence from a dogmatic friend, i could either (correctly) assume her friend is censoring, or (incorrectly) assume that friend simply received no signal. It is this later incorrect assumption that allows i's belief to be biased. Clearly, the quality of the signals cannot be perfect (i.e., $q = 1$) for i's expected posterior to deviate from her prior. Yet, the result shows that the imbalance between dogmatic friends—however small—can always overcome the unbiased information coming from normal friends and i's own signal when the agents get signals of sufficiently low quality. Put differently, the misperceived selective sharing dominates even though dogmatic friends' signals become almost uninformative.

Proposition 1 has immediate consequences for belief polarization in societies. If i has more dogmatic-left than dogmatic-right friends and i' has more dogmatic-right than dogmatic-left friends—where again the imbalances can be small—then in expectation their posteriors will move apart, at least if the quality of signals is small. Thus, our result suggests that the composition of one's echo chamber can be a driver of belief polarization. This is consistent with the narrative that social media can give rise to "echo chambers" where people are exposed to an unbalanced diet of opinions and as a result develop polarized beliefs. However, our result qualifies this narrative in two ways: First, it does not require the presence of fake news; second, it stresses that echo chambers also require low quality information in order to distort beliefs from the truth.

Our first result showed that one round of information arrival is enough to give rise to polarization, at least in expectation. However, it leaves open the possibility that when information becomes abundant (i.e., in the long run after many signals) polarization disappears. In fact, the opposite can occur. As long as information quality is sufficiently low, abundant information can exacerbate the effect of misperceived selective sharing and cause i's beliefs to be incorrect *with probability 1.* To show this, we consider the limit in probability of the posterior belief $\mu(\mathbf{s}^T)$ as $T \to \infty$, which we denote by $\mathbf{plim}_{T\to\infty}\,\mu(\mathbf{s}^T)$. If q is sufficiently small, i's posterior belief is pushed towards the degenerate belief on the state that is favored by the majority of her dogmatic friends, with probability 1 and irrespective of the true state. Denote, $I_{\{\omega=L\}}$ as the indicator function that equals 1 if $\omega = L$ and 0 otherwise.

Proposition 2. *Fix an agent such that $d_L > d_R$.*

1. *There exists a unique $1 > \bar{q}_2(d_L, d_R, n, \gamma) > \frac{1}{2}$ such that*

$$\mathbf{plim}_{T\to\infty}\,\mu(\mathbf{s}^T) = \begin{cases} 1 & \text{if } q < \bar{q}_2(d_L, d_R, n, \gamma) \\ I_{\{\omega=L\}} & \text{if } q > \bar{q}_2(d_L, d_R, n, \gamma) \end{cases}$$

2. *$\bar{q}_2(d_L, d_R, n, \gamma)$ is increasing in $d_L - d_R$ and decreasing in n and γ.*

Consider the first part of Proposition 2. Intuitively, with many rounds of information arrival agent i's own signals provide a more accurate estimate of the state, which would result in perfect learning in a standard setting. However, i combines her signals with the signals from her friends, which provide more information, but this information is biased in ways she does not correctly take into account. Once again, the outcome of this race between the two kinds of information is a priori not clear. Yet, with low quality signals, the distortion introduced in each step of updating unveiled in Proposition 2.1 accumulates over time leading the posterior astray with certainty.

To see this more precisely, consider updating at each period t. Dogmatic friends tend to provide more left than right signals. Normal friends provide unbiased signals. Who wins the race as $T \to \infty$? Recall the correct updating term Q^M and incorrect updating term Γ^S. When q is close to $\frac{1}{2}$ (low informativeness), the correct updating term Q^M is close to 1 and thus the misperception boost to informativeness Γ^S dominates. By contrast, when q is close to 1 (maximal informativeness), the correct updating term Q^M is close to zero and hence the misperception boost to informativeness Γ^S is curtailed.

The second part of Proposition 2 provides important comparative statics on the lower bound of informativeness that avoids polarization. This lower bound increases with the absolute difference between the number of dogmatic left and right friends, or the magnitude of the echo chamber imbalance, and increases with the degree of misperception. The lower bound decreases with number of normal friends. Therefore, given q, it is possible that if an agent belongs to a social group with a large number of normal friends, she learns correctly as information becomes abundant. By contrast, if the agent belongs to a smaller social group with fewer normal friends, the effect of dogmatic friends may prevail for the same q, causing her belief to polarize relative to other agents.

With regard to polarization, Propositions 1 and 2 can shed further light on the role of "echo chambers." In sum, they highlight the importance of the composition of an echo chamber, rather than its size. To the extent that interacting online promotes the formation of more unbalanced social groups, it may contribute to belief polarization even when the quality of information is unchanged.

Propositions 1 and 2 show that imbalance in selective sharing and misperception are sufficient for polarization to arise. But are they necessary? The answer is yes.

Proposition 3. *If either $d_L = d_R$ or $\gamma = 1$, then*

$$\mathbb{E}[\mu(\mathbf{s}^1)] = \pi \qquad \textit{and} \qquad \operatorname*{plim}_{T\to\infty} \mu(\mathbf{s}^T) = I_{\{\omega=L\}}.$$

Intuitively, if i has an equal number of dogmatic-left and dogmatic-right friends, they offset each other when sharing information selectively, even in the presence of misperception. Alternatively, if i fully understands their selective sharing—always interpreting their silence as the arrival of unfavorable signals—she fully unravels its effects even if there are more dogmatic friends on one side. Note that even a minor degree of misperceptions (γ close to but strictly less than 1) will make polarization possible.

4 Network expansion

Consider agent i once again and suppose that due to advances in technology (e.g. development of social media), the agent is now connected with more agents of all types, while still receiving the same quantity of information on her own. Specifically, assume that the number of normal, dogmatic-left and dogmatic-right friends expands according to proportions $(\lambda_L, \lambda_R, \lambda_N)$, respectively. Proposition 4 characterizes the impact on the informativeness required to avoid polarization.

Proposition 4. *Fix an agent such that $d_L > d_R$, then $\bar{q}_2(\lambda_L d_L, \lambda_R d_R, \lambda_N n, \gamma) \geq \bar{q}_2(d_L, d_R, n, \gamma)$ if and only if*

$$(\lambda_N - 1) \leq \left(\frac{\lambda_L d_L - \lambda_R d_R}{d_L - d_R} - 1\right) \cdot \left(1 + \frac{1}{n}\right)$$

Proposition 4 states that the range of information quality q leading to incorrect learning shrinks if and only if normal friends grow sufficiently fast. Intuitively, there is a trade-off between access to information and the scope for echo-chambers to bias beliefs. Note that if scaling is proportional (i.e. $\lambda_L = \lambda_R = \lambda_N$) then $1 < 1 + \frac{1}{n}$ and hence the range of information quality leading to incorrect beliefs *expands*. Thus a proportional scaling of the network leads to more opportunities for polarization.

This result suggests that when there are changes in the network environment that increase connectivity of people (keeping proportions of different types of friends the same), but do not increase the amount of information people get on their own, it can contribute to belief polarization. For a fixed information quality q, it is possible that a person whose beliefs are not biased in a smaller network becomes polarized when the network expands.

The previous proposition deals with how large λ_N needs to be in order to decrease an individual's $\bar{q}_2$. This reduces the range of q in which the individual gets biased, but may still not be enough if the true q is far lower. The next proposition explores how large λ_N needs to be in order to drop $\bar{q}_2$ below any given value of q, so as to un-bias a given agent. Put another way: Given some $\hat{q}$, what λ_N suffices to ensure correct learning?

Proposition 5. *Fix $d_L > d_R$ and $\hat{q} \in \left(\frac{1}{2}, \bar{q}_2(d_L, d_R, n, \gamma)\right)$. Then $\bar{q}_2(\lambda d_L, \lambda d_R, \lambda_N n, \gamma) < \hat{q}$ if*

$$\lambda_N > \frac{d_L - \hat{q}(d_L + d_R)}{(2\hat{q} - 1)n}\lambda - \frac{1}{n}.$$

Given echo chamber (d_L, d_R, n) and growth rates of friends (which may be determined by data or algorithms), this result predicts if and when $\bar{q}_2$ crosses $\hat{q}$. This points to how link-formation algorithms may possibly be designed to minimize polarization. i.e. the result suggests the growth rate of normal friends required. Note that the right side of the inequality is increasing in d_L and λ, decreasing in d_R and $\hat{q}$. Moreover, the right side of the inequality is decreasing in n when it is positive (when $\hat{q} < \bar{q}_2$).

5 Network polarization

In this section, we move from the analysis of individual agents to the analysis of the network as a whole. Our focus is the set of normal agents in the network (i.e., excluding dogmatic ones). Interpreting dogmatic agents as those whose beliefs are degenerate implies that only the beliefs of normal agents respond to arriving information. Let the set of all normal agents be denoted by $\mathcal{N}$.

Before proceeding, we define a measure of polarization for our network. Without loss of generality, we will define it for beliefs in state L. Define polarization in period t as the average sum of differences of beliefs across all agents in $\mathcal{N}$:

$$\Pi^t = \frac{2}{|\mathcal{N}|^2} \sum_{i,j \in \mathcal{N}} \left| \mu_i(\mathbf{s}_i^t) - \mu_j(\mathbf{s}_j^t) \right|.$$

This expression captures the extent to which disagreement is present in the network. The scaling factor of 2 is introduced in order to ensure $\Pi^t \in [0, 1]$.

We focus on asymptotic polarization of beliefs (as $T \to \infty$), when all normal agents' beliefs have already converged. Let $\mathcal{N}_R$ and $\mathcal{N}_L$ denote the sets of normal agents whose asymptotic beliefs put probability 1 on states R and L, respectively. We will also refer to these sets as "eventually incorrect" and "eventually correct" populations, depending on which state is true.[4] Define the limit polarization Π as

$$\Pi \equiv \operatorname*{plim}_{t\to\infty} \Pi^t = \frac{2}{|\mathcal{N}|^2} \cdot 2|\mathcal{N}_R||\mathcal{N}_L| \cdot |1-0| = \frac{4|\mathcal{N}_R||\mathcal{N}_L|}{|\mathcal{N}|^2}$$

The value of Π varies from 0 to 1, attaining its maximum when $\mathcal{N}_R$ and $\mathcal{N}_L$ have the same cardinality.

The first result in this section concerns the dynamics of $\mathcal{N}_R$ and $\mathcal{N}_L$ as q increases from $1/2$ to 1. Assume the true state is R and fix some $q \in (1/2, 1)$. Then, $\mathcal{N}_R$ consists of two sets of agents: (1) all agents for whom $\bar{q}_{2i}$ is below q (these agents learn correctly), *and* (2) all agents for whom dogmatic-right friends dominate dogmatic-left friends and $\bar{q}_{2i}$ is above q (these agents place probability 0 on $\omega = L$ in the long run irrespective of true state). Meanwhile, $\mathcal{N}_L$ consists of all agents for whom dogmatic-left friends dominate dogmatic-right friends and whose $\bar{q}_{2i}$ is above q.

As q increases, all agents that are already in $\mathcal{N}_R$ will remain there. Each agent for whom $\bar{q}_{2i}$ is below q will continue to be below q, and other agents for whom dogmatic-right friends dominate dogmatic-left friends may remain in that category or pass into the first category and still remain in $\mathcal{N}_R$. However, agents in $\mathcal{N}_L$ may switch to the other set. When q becomes larger than $\bar{q}_{2i}$ for an agent in $\mathcal{N}_L$, he or she is no longer biased by dogmatic-left friends, and thus will place probability 0 on $\omega = L$ in the limit instead of 1, passing into $\mathcal{N}_R$. This argument is formalized in the following lemma.

Lemma 1. *As q increases, the set of "eventually correct" agents is weakly expanding and the set of "eventually incorrect" agents is weakly contracting.*

What implications does this have for network polarization? A common intuition would suggest that as people receive better information, disagreements should decline. Lemma 1 places some limits on this intuition: if for low q the set of "eventually incorrect" agents is sufficiently large, polarization in society will be low. However, as q increases, it will cause a gradual shift into the set of "eventually correct" agents, and polarization in society will actually *increase* initially. Once the set of "eventually correct" agents outnumbers the set of "eventually incorrect" ones, polarization will start to decline as q increases. This implies that polarization in a social network may temporarily increase as quality of information is slowly improving.

The following proposition summarizes the intuition above and provides a necessary and sufficient condition for such non-monotonicity of network polarization to take place. Let $\mathcal{N}_R(q)$ and $\mathcal{N}_L(q)$ denote the sets of agents whose beliefs converge to state R and L, respectively, for a given value of q. Additionally, let $\mathcal{D}_R$ and $\mathcal{D}_L$ denote sets of agents who have a

[4] If R is the true state, then $\mathcal{N}_R$ is the "eventually correct" population, whereas $\mathcal{N}_L$ is the "eventually incorrect" population—vice versa if the true state is L.

right- and left-slanted echo chambers, respectively. The remaining agents do not have any imbalance in their echo chamber.

Proposition 6. *Fix network with $\bar{q}_{2i} \neq \bar{q}_{2j}$ for all i, j and ω. Π is decreasing in q over $\left(\frac{1}{2}, 1\right)$ iff $|\mathcal{D}_{-\omega}| \leq \frac{1}{2}(|\mathcal{N}|+1)$. Otherwise, Π is single peaked.*

Intuitively, if $|\mathcal{D}_{-\omega}| > \frac{1}{2}(|\mathcal{N}|+1)$, more than half of normal agents are initially biased towards the wrong state. As q increases, this set of "eventually incorrect" agents will lose agents in favor of the "eventually correct" set. This will increase polarization in the network at first. It will achieve its maximum when the sets of "eventually incorrect" and "eventually correct" agents are equal in size, and will start decreasing afterwards.

6 Mitigating polarization

We have shown that selective sharing and misperceptions are enough to generate polarization of beliefs in a network of agents. How could a social planner combat this polarization, if he or she could affect only the information the agents receive (taking selective sharing and misperception as given)? That is, we want to see whether it is possible to eliminate polarization by changing only the external information structure of the agents. One obvious way to do this is to *directly* increase the quality of information q at the source. However, this may not be feasible due to technological or economic reasons. For example, this may involve forcing newspapers to spend more on reporters, data gathering, and fact checking. However, it is possible to increase the quality of information that the agents receive without changing the quality of the primitive signals s_{it}. This involves *signal aggregation*.

Signal aggregation consists in summarizing a set of signals into a single message. This summary can, of course, be done in many ways, emphasizing some aspects of the original signals and omitting others. It is important to note, however, that signal aggregation involves some loss of information relative to the totality of the aggregated signals. Nonetheless, the resulting message can have higher quality than each aggregated signal individually. It is this distinction that renders signal aggregation useful for our goal of reducing polarization, despite the loss of information. As such, signal aggregation involves a trade-off between slowing short-run learning and debiasing long-run learning, which is a novel aspect of our learning environment.

Consider the following form of signal aggregation. Let M be an odd number and divide time into blocks of M periods. Define $\hat{s}^i_{Mt}$ as the aggregated signal that is released to agent i at the end of each time block and reports whether more left or more right signals realized over the previous M periods:

$$\hat{s}^i_{Mt} = \begin{cases} 0, \text{ if } \sum_{k=(t-1)M+1}^{tM} 1_{\{s_{ik}=L\}} < \frac{M}{2} \\ 1, \text{ if } \sum_{k=(t-1)M+1}^{tM} 1_{\{s_{ik}=L\}} > \frac{M}{2}. \end{cases}$$

A natural question is why a social planner would prefer to coarsen each agent's information structure in such a way. Clearly, $\hat{s}^i_{Mt}$ conveys less information than do the aggregated

M signals together. However, on the one hand, if all agents observe only $\hat{s}^i_{Mt}$ every other M periods, they have *fewer* opportunities to selectively share information with their friends. But this is irrelevant in the long run. On the other hand—and more importantly—$\hat{s}^i_{Mt}$ has higher quality than each individual s_{it}. As a result, by substituting s_{it} information structure with $\hat{s}^i_{Mt}$ for each agent, we are worsening information quality for each agent individually, but we are also reducing the influence of selective sharing on agents' beliefs. Hence, this might restore the standard result of convergence to the truth in asymptotic learning. To see this, suppose $M = 3$, consider the first block of three periods, and assume $\omega = L$. We have

$$\begin{aligned}
\mathbb{P}(\hat{s}^i_3 = 0|\omega = L) &= \mathbb{P}\left(\sum_{t=1}^{3} s_{it} \leq 1 \middle| \omega = L\right) \\
&= q^3 + 3q^2(1-q) \\
&= q \cdot \left(q^2 + 2q(1-q) + q(1-q)\right) \\
&> q \cdot \left(q^2 + 2q(1-q) + (1-q)^2\right) \\
&= q = \mathbb{P}(s_{it} = l|\omega = L).
\end{aligned}$$

Thus, $\hat{s}^i_3$ is more informative than each s_{it} of quality q. In contrast to standard models where the quality of information does not matter in the long run, we saw that in our model a sufficiently high quality of information can allow all agents to learn correctly, thereby removing polarization. The remaining question is then whether aggregating signals according to $\hat{s}^i_{Mt}$ can achieve that quality level. The next proposition shows that this is always possible.

Proposition 7. *Fix any network and q such that $\Pi > 0$. There exists M such that, if each agent i observes signals $\hat{s}^i_{Mt}$, then the resulting limit polarization $\hat{\Pi}$ equals zero.*

7 Discussion

We explored conditions under which learning from shared news may lead to belief polarization. We develop a model reflecting two important facts about news sharing in networks: unbalanced selective transmission of information and misperception. We include misperception to account for the evidence that people are neither completely naive nor fully aware of selective sharing. We found that with these features polarization occurs if (and only if) information quality is sufficiently low. Thus, our results emphasize the importance of distinguishing between quality and quantity of information, a distinction that is irrelevant in standard models.

With this understanding, we also investigated ways to mitigate polarization. On the one hand, our analysis of the effects of network expansion suggests ways to influence network dynamics so as to limit the scope for polarization to occur. On the other hand, we show that aggregating the signals that agents receive may mitigate polarization. A benevolent authority may use this insight to stop the effects of misperceived selective sharing and debias beliefs. In practice, one way to achieve this is to create institutional intermediaries or platforms that aggregate news, namely, that provide people with information in larger "digested" batches rather than many raw bits. Even if these less frequent news reports *lose* some of the information in the summarized news, they can mitigate polarization by reducing the scope for echo chambers to cause people's beliefs to go awry.

Our analysis goes to the heart of why new technologies and formats of communication enabled by the Internet may increase polarization. Tweets and social-media posts tend to be short and hence of low quality. Moreover, overwhelmed by the abundance of information, people may spread their limited attention across more sources, thereby absorbing less accurate information from each of them. Finally, we also shed new light on two ways in which malevolent actors who thrive on social polarization benefit from how news is shared in social networks. One obvious way is to use fake news to directly lower the information quality or bias echo chambers in a particular direction so as to exacerbate their effects. A more subtle way is to release bits of news with high frequency so as to leverage the biasing power of misperceived selective sharing that we uncover and to draw attention away from high-quality information sources. Even though removing fake news is important for addressing polarization, our analysis suggests that it is not sufficient.

Our analysis opens several directions for future research. So far, we have exogenously assumed selective sharing by having dogmatic agents who do not update their beliefs and so do not change how they selectively transmit their signals. However, in reality people often choose what information to share so as to promote their views or beliefs, which are not set in stone. One can interpret the current model as describing situations where dogmatic agents can change their beliefs, yet they do so much more slowly than normal agents. As a result, how they selectively share information is very persistent, which should leave the insights of our results unchanged.

Our results about the beliefs of a single agent can be used to further examine the distribution of beliefs in the network as a whole. We conjecture that there is a relationship between how the composition of echo chambers is distributed in the network and how beliefs are distributed across agents. For instance, if half the population has a majority of dogmatic-right friends, half a majority of dogmatic-left friends, and information quality is sufficiently low, then the beliefs of each half should move apart as information arrives. This is consistent with evidence showing that people on the left and on the right of the political spectrum tend to have more like-minded friends than not, and with the view that this may be a cause of the ongoing polarization (e.g., see Pew Research Center (2014)). Note that, according to our analysis, such polarization does not require that people look at the world in fundamentally incompatible ways, but only that they have different news diets—which may be easier to address.

Finally, the role of the network structure in our analysis begs the question of what would change if we allowed for endogenous network formation. On the one hand, people may tend to form more links with like-minded friends, which may enlarge the imbalance in their echo chambers. This would reinforce and magnify the effects of selective sharing and misperceptions we highlight. On the other hand, people may be more likely to link with normal friends than dogmatic friends, which would have opposite implications. Either way, this is ultimately an empirical question: Once we measure the rates of link formation with different types of friends, we can use our model to predict their consequences. Nonetheless, understanding the incentives to form links so as to obtain information is also important to answer these questions and guide further empirical investigations of polarization.

References

Acemoglu, D., A. Ozdaglar, and A. ParadejGheibi (2010). Spread of (Mis)information in Social Networks. Games and Economic Behavior 70, 194–227.

Alesina, A., A. Miano, and S. Stantcheva (2020). The Polarization of Reality. In AEA Papers and Proceedings, Volume 110, pp. 324–28.

Allcott, H. and M. Gentzkow (2017). Social Media and Fake News in the 2016 Election. Journal of Economic Perspectives 31(2), 211–36.

Azzimonti, M. and M. Fernandes (2018). Social Media Networks, Fake News, and Polarization. *Working paper.*

Barber, M. and N. McCarty (2015). Causes and Consequences of Polarization. Political Negotiation: A Handbook 37, 39–43.

Bartels, L. M. (2008). Unequal Democracy: The Political Economy of the New Gilded Age. Princeton University Press.

Bishop, B. (2009). The Big Sort: Why the Clustering of Like-Minded America is Tearing us Apart. Houghton Mifflin Harcourt.

Bohren, A. and D. Hauser (2018). Social Learning with Model Misspecification: A Framework and a Robustness Result. PIER Working Paper Archive 18-017, Penn Institute for Economic Research, Department of Economics, University of Pennsylvania.

Bohren, J. A. (2016). Informational Herding with Model Misspecification. Journal of Economic Theory 163(C), 222–247.

Boxell, L., M. Gentzkow, and J. Shapiro (2018). Greater Internet Use is Not Associated with Faster Growth of Political Polarization Among US Demographic Groups. Proceedings of the National Academy of Sciences 115(3).

Bursztyn, L., G. Egorov, R. Enikolopov, and M. Petrova (2019). Social Media and Xenophobia: Evidence from Russia. NBER Working Paper No. 26567.

Conroy-Krutz, J. and D. C. Moehler (2015). Moderation from Bias: A Field Experiment on Partisan Media in a New Democracy. The Journal of Politics 77(2), 575–587.

Cross, P. (1977). Not Can But Will College Teachers Be Improved? New Directiosn for Higher Education 17, 1–15.

DeMarzo, P., D. Vayanos, and J. Zweibel (2003). Persuasion Bias, Social Influence, and Unidimensional Opinions. Quarterly Journal of Economics 118(3), 909–968.

Desmet, K. and R. Wacziarg (2018). The Cultural Divide. CEPR Discussion Papers 12947, C.E.P.R. Discussion Papers.

Dixit, A. and J. Weibull (2007). Political Polarization. Proceedings of the National Academy of Sciences 104(18), 7351–7256.

Enke, B. and F. Zimmermann (2017). Correlation Neglect in Belief Formation. The Review of Economic Studies 86(1), 313–332.

Enke, B., F. Zimmermann, and F. Schwerter (2019). Associative Memory and Belief Formation. *Working paper.*

Eyster, E. and M. Rabin (2010). Na'ive Herding in Rich-Information Settings. American Economic Journal: Microeconomics 2, 221–243.

Eyster, E., M. Rabin, and G. Weizsäcker (2018). An Experiment On Social Mislearning. Rationality and Competition Discussion Paper Series 73, CRC TRR 190 Rationality and Competition.

Frick, M., R. Iijima, and Y. Ishii (2019). Misinterpreting Others and the Fragility of Social Learning. Cowles Foundation Discussion Papers 2160, Cowles Foundation for Research in Economics, Yale University.

Gilens, M. (2012). Affluence and Influence: Economic Inequality and Political Power in America. Princeton University Press.

Golub, B. and M. O. Jackson (2010). Naïve Learning in Social Networks and the Wisdom of Crowds. American Economic Journal: Microeconomics 2(1), 112–49.

Hoffmann, F., K. Khalmetski, and M. Le Quement (2019). Disliking to Disagree. *Working paper.*

Keefer, P. and S. Knack (2002). Polarization, Politics and Property Rights: Links between Inequality and Growth. Public choice 111(1-2), 127–154.

Levendusky, M. S. (2013). Why Do Partisan Media Polarize Viewers? American Journal of Political Science 57(3), 611–623.

Levy, G. and R. Razin (2019a). Echo Chambers and Their Effects on Economic and Political Outcomes. Annual Review of Economics forthcoming.

Levy, G. and R. Razin (2019b). Information Diffusion in Networks with the Bayesian Peer Influence Heuristic. *Working paper*.

Levy, R. (2020). Social Media, News Consumption and Polarization: Evidence from a Field Experiment. *Working paper*.

McCarty, N., K. T. Poole, and H. Rosenthal (2009). Does Gerrymandering Cause Polarization? American Journal of Political Science 53(3), 666–680.

Molavi, P., A. Tahbaz-Salehi, and A. Jadbabaie (2018). A Theory of Non-Bayesian Social Learning. Econometrica 86(2), 445–490.

Mosquera, R., M. Odunowo, T. McNamara, X. Guo, and R. Petrie (2019). The Economic Effects of Facebook. Available at SSRN: https://ssrn.com/abstract=3312462 or http://dx.doi.org/10.2139/ssrn.3312462.

Mullainathan, S. and A. Shleifer (2005). The Market for News. American Economic Review 95(4), 1031–1053.

Odean, T. (1998). Volume, Volatility, Price and Profit When All traders Are Above Average. Journal of Finance 53(6), 1887–1934.

Perego, J. and S. Yuksel (2018). Media Competition and Social Disagreement. *Working paper*.

Periser, E. (2011). The Filter Bubble: What the Internet is Hiding from You. Penguin, London.

Pew Research Center (2014). Political Polarization and Media Habits. pp. October, 2014.

Pew Research Center (2020). U.S. Media Polarization and the 2020 Election: A Nation Divided. pp. January, 2020.

Pogorelskiy, K. and M. Shum (2019). News We Like to Share: How News Sharing on Social Networks Influences Voting Outcomes. Available at SSRN: https://ssrn.com/abstract=2972231 or http://dx.doi.org/10.2139/ssrn.2972231.

Reeves, A., M. McKee, and D. Stuckler (2016). 'It's The Sun Wot Won It': Evidence of Media Influence on Political Attitudes and Voting from a UK Quasi-Natural Experiment. Social science research 56, 44–57.

Shin, J., L. Jian, K. Driscoll, and F. Bar (2018). The Diffusion of Misinformation on Social Media: Temporal Pattern, Message, and Source. Computers in Human Behavior 83, 278–287.

Shin, J. and K. Thorson (2017). Partisan Selective Sharing: The Biased Diffusion of Fact-Checking Messages on Social Media. Journal of Communication 67(2), 233–255.

Spiegler, R. (2019). Behavioral Implications of Causal Misperceptions. *Working paper.*

Sunstein, C. (2017). Divided Democracy in the Age of Social Media. Princeton University Press.

Svenson, O. (1981). Are We all Less Risky and More Skillful Than Our Fellow Drivers? Acta Psychologica 47(2), 143–148.

Tucker, J., A. Guess, P. Barbera, C. Vaccari, A. Siegel, S. Sanovich, D. Stukal, and B. Nyhan (2019). Social Media, Political Polarization, and Political Disinformation: A Review of Scientific Literature. Available at SSRN: https://ssrn.com/abstract=3144139 or http://dx.doi.org/10.2139/ssrn.3144139.

Weeks, B. E., D. S. Lane, D. H. Kim, S. S. Lee, and N. Kwak (2017). Incidental Exposure, Selective Exposure, and Political Information Sharing: Integrating Online Exposure Patterns and Expression on Social Media. J. Computer-Mediated Communication 22, 363–379.

Zak, P. J. and S. Knack (2001). Trust and Growth. The economic journal 111(470), 295–321.

Zuckerman, E. and J. Jost (2001). What Makes you Think you are so Popular? Social Psychology Quarterly 64(3), 207–223.

Appendix

A Proof of Proposition 1

We consider agent i who has d_L dogmatic-left friends, d_R dogmatic-right friends and n normal friends. We drop all subscripts since we are focusing on a single agent.

Proof of Proposition 1. First, assume that the true state is $\omega = R$. As we noted in Equation 2, given fixed realizations of ℓ_L, ℓ_R and ℓ_N, agent i's posterior belief is given by the following expression:

$$\mu_i(\omega = L|\ell_L, \ell_R, \ell_N) \quad = \quad \frac{\pi}{\pi + (1-\pi)Q^M\Gamma^S}, \tag{2}$$

where

$$\begin{aligned} Q &\equiv \frac{1-q}{q} \\ M &\equiv \ell_L - (d_R - \ell_R) + 2\ell_N - N = \ell_L + \ell_R - d_R + 2\ell_N - N \\ \Gamma &\equiv \frac{\gamma(1-q) + (1-\gamma)}{\gamma q + (1-\gamma)} \\ S &\equiv \ell_R - (d_L - \ell_L) = \ell_L + \ell_R - d_L. \end{aligned}$$

In terms of random variables, $\mu_i(\omega = L|\ell_L, \ell_R, \ell_N)$ depends only on the sum of ℓ_L and ℓ_R and on ℓ_N. Let $\ell_D = \ell_L + \ell_R$ and note that due to independence of individual signals, ℓ_D is a Binomial r.v. with probability $(1-q)$ and sample size $(d_L + d_R)$.

Agent i's posterior belief, given fixed realizations of ℓ_D and ℓ_N, is then given by:

$$\mu_i(\omega = L|\ell_D, \ell_N) = \frac{\pi}{\pi + (1-\pi)Q^{\ell_D - d_R + 2\ell_N - N}\Gamma^{\ell_D - d_L}}$$

Recall that ℓ_D has Binomial distribution with parameters $(1-q)$ and $(d_L + d_R)$. Therefore, i's expected posterior belief, conditional on ℓ_N and $\omega = R$, is given by the following expression:

$$\begin{aligned} \mathbb{E}\left[\mu_i(\omega = L) \mid \ell_N, \omega = R\right] = \sum_{k=0}^{d_L+d_R} \quad & \left[\frac{(d_L+d_R)!}{k!(d_L+d_R-k)!} \cdot (1-q)^k q^{d_L+d_R-k}\right. \\ & \left. \cdot \frac{\pi}{\pi + (1-\pi)Q^{k-d_R+2\ell_N-N}\Gamma^{k-d_L}}\right]. \end{aligned}$$

Recall that ℓ_N is a Binomial r.v. with parameters $(1-q)$ and N. Hence, the expected posterior belief conditional on $\omega = R$ is given by

$$\mathbb{E}\left[\mu_i(\omega = L) \mid \omega = R\right] = \sum_{k=0}^{N} \frac{N!}{k!(N-k)!}(1-q)^k q^{N-k} \mathbb{E}\left(\mu_i(\omega = L)|k, \omega = R\right).$$

We will prove that there exists $\bar{q}_1 > \frac{1}{2}$ such that if $q \in \left(\frac{1}{2}, \bar{q}_1\right)$, $\mathbb{E}\left[\mu_i(\omega = L)|\omega = R\right] > \pi$. For that, we will first find the derivative of $\mathbb{E}\left[\mu_i(\omega = L)|\ell_N, \omega = R\right]$ with respect to q at $q = \frac{1}{2}$ for any $\ell_N \geq 0$, and then show that the derivative of $\mathbb{E}\left[\mu_i(\omega = L)|\omega = R\right]$ w.r.t. q at $q = \frac{1}{2}$ is positive. Then, using continuity of the expected posterior in q and the fact that at $q = \frac{1}{2}$ we have $\mathbb{E}\left[\mu_i(\omega = L)|\omega = R\right] = \pi$, this will imply the desired statement.

If $d_L + d_R$ is odd[5], we can rewrite one of the sums above as follows:

$$\mathbb{E}\left[\mu_i(\omega = L)|\ell_N, \omega = R\right] = \sum_{k=0}^{\frac{d_L+d_R-1}{2}} \frac{(d_L+d_R)!}{k!(d_L+d_R-k)!}\Big(f(k,q,\ell_N) + f(d_L+d_R-k,q,\ell_N)\Big),$$

where

$$f(k,q,\ell_N) = (1-q)^k q^{d_L+d_R-k} \frac{\pi}{\pi + (1-\pi)\left(\frac{1-q}{q}\right)^{k-d_R+2\ell_N-N}\left(\frac{\gamma(1-q)+(1-\gamma)}{\gamma q+(1-\gamma)}\right)^{k-d_L}}$$

For simplifying subsequent algebra, define a function $z(q,\gamma)$ as

$$z(q,\gamma) = \ln(\Gamma)\left[\ln(Q)\right]^{-1} = \ln\left(\frac{\gamma(1-q)+(1-\gamma)}{\gamma q+(1-\gamma)}\right)\left[\ln\left(\frac{1-q}{q}\right)\right]^{-1}.$$

Consider the derivative of $f(k,q,\ell_N)$ w.r.t. q at $q = \frac{1}{2}$:

$$\begin{aligned}
&\frac{\partial}{\partial q} f(k,q,\ell_N) = \Big((d_L+d_R-k)(1-q)^k q^{d_L+d_R-k-1} - k(1-q)^{k-1} q^{d_L+d_R-k}\Big)\cdot \\
&\frac{\pi}{\pi + (1-\pi)\left(\frac{1-q}{q}\right)^{k-d_R+2\ell_N-N+(k-d_L)z(q,\gamma)}} \\
&\quad + (1-q)^k q^{d_L+d_R-k}\cdot\pi(1-\pi)\cdot\Bigg[\frac{(k-d_R+2\ell_N-N)Q^{k-d_R+2\ell_N-N-1+(k-d_L)z(q,\gamma)}\frac{1}{q^2}}{\left(\pi+(1-\pi)Q^{k-d_R+2\ell_N-N}\Gamma^{k-d_L}\right)^2} \\
&\qquad + \frac{(k-d_L)Q^{k-d_R+2\ell_N-N+(k-d_L-1)z(q,\gamma)}\frac{(2-\gamma)\gamma}{(\gamma q+(1-\gamma))^2}}{\left(\pi+(1-\pi)Q^{k-d_R+2\ell_N-N}\Gamma^{k-d_L}\right)^2}\Bigg] \\
&\overset{(q=\frac{1}{2})}{=} \left(\frac{1}{2}\right)^{d_L+d_R-1}(d_L+d_R-2k)\frac{\pi}{\pi+(1-\pi)} + \left(\frac{1}{2}\right)^{d_L+d_R} 4\pi(1-\pi)\cdot \\
&\qquad \cdot\frac{(k-d_R+2\ell_N-N)+(k-d_L)\frac{\gamma}{2-\gamma}}{(\pi+(1-\pi))^2} \\
&= \left(\frac{1}{2}\right)^{d_L+d_R-1}\cdot\Big((d_L+d_R-2k)\pi + 2\pi(1-\pi)\Big((k-d_R+2\ell_N-N)+(k-d_L)\frac{\gamma}{2-\gamma}\Big)\Big).
\end{aligned}$$

[5] We consider the case of $d_{Li} + d_{Ri}$ being even at the end of the proof.

Therefore, at $q = \frac{1}{2}$ we have

$$\begin{aligned}
\frac{\partial}{\partial q}\mathbb{E}\left[\mu_i(\omega = L)|\ell_N, \omega = R\right] &= \sum_{k=0}^{\frac{d_L+d_R-1}{2}} \frac{(d_L+d_R)!}{k!(d_L+d_R-k)!}\left(\frac{\partial}{\partial q}f(k,q,\ell_N) + \frac{\partial}{\partial q}f(d_L+d_R-k,q,\ell_N)\right) \\
&= \sum_{k=0}^{\frac{d_L+d_R-1}{2}} \frac{(d_L+d_R)!}{k!(d_L+d_R-k)!}\left(\frac{1}{2}\right)^{d_L+d_R-1} 2\pi(1-\pi)\cdot \\
&\quad \cdot\left(2(2\ell_N - N) + (d_L - d_R) - (k - d_{Li})\frac{\gamma}{2-\gamma}\right) \\
&= 4\pi(1-\pi)(2\ell_N - N) + \\
&\quad + \sum_{k=0}^{\frac{d_L+d_R-1}{2}} \frac{(d_L+d_R)!}{k!(d_L+d_R-k)!}\left(\frac{1}{2}\right)^{d_L+d_R-1} 2\pi(1-\pi)\left(d_L - d_R\right)\tfrac{2-2\gamma}{2-\gamma}.
\end{aligned}$$

Note that the second term (a sum) in the expression above is strictly positive for any $\ell_N \geq 0$, since $d_L > d_R$ and $\gamma \in (0,1)$ hold. Thus, when we compute the derivative of $\mathbb{E}\left[\mu_i(\omega = L)|\omega = R\right]$, this part of the sum will remain strictly positive. The only term to be concerned about is the one with $(2\ell_N - N)$ in it.

Let

$$H = \sum_{k=0}^{\frac{d_L+d_R-1}{2}} \frac{(d_L+d_R)!}{k!(d_L+d_R-k)!}\left(\frac{1}{2}\right)^{d_L+d_R-1} 2\pi(1-\pi)\left(d_L - d_R\right)\frac{2-2\gamma}{2-\gamma}.$$

Then the derivative of the conditional expected posterior at $q = \frac{1}{2}$ is equal to

$$\begin{aligned}
\frac{\partial}{\partial q}\mathbb{E}\left[\mu_i(\omega = L)|\omega = R\right] = \sum_{k=0}^{N} \frac{N!}{k!(N-k)!}\left(\frac{1}{2}\right)^{N-1}\left(-k + (N-k)\right)\mathbb{E}\left[\mu_i(\omega = L)|k, \omega = R, q = \frac{1}{2}\right] + \\
+ \sum_{k=0}^{N} \frac{N!}{k!(N-k)!}\left(\frac{1}{2}\right)^{N}\Big[H + 4\pi(1-\pi)(2k - N)\Big].
\end{aligned}$$

We know $\mathbb{E}\left[\mu_i(\omega = L)|k, \omega = R, q = \frac{1}{2}\right] = \pi > 0$ and $H > 0$. Thus,

$$\begin{aligned}
\frac{\partial}{\partial q}\mathbb{E}\left[\mu_i^1(\omega = L)|\omega = R\right] &= H + \sum_{k=0}^{N} \frac{N!}{k!(N-k)!}\left(\frac{1}{2}\right)^{N-1}\left(\pi - 2\pi(1-\pi)\right)(N - 2k) \\
&= H + \left(\pi - 2\pi(1-\pi)\right)\left(\frac{1}{2}\right)^{N-1}\sum_{k=0}^{N} \frac{N!}{k!(N-k)!}(N - 2k)
\end{aligned}$$

Note that the sum $\sum_{k=0}^{N} \frac{N!}{k!(N-k)!}(N-2k)$ is symmetric around $k = \frac{N}{2}$. Specifically, for $k \leq \frac{N}{2}$ there is exactly identical term with an opposite sign for $j = N - k$. This implies that the sum will annihilate itself into zero. This will mean that the derivative is simply equal to H, which is positive, i.e. the conditional expected posterior is *increasing* at $q = \frac{1}{2}$.

Now consider the case $\omega = L$. The updating rule of the agent will be the same as in the previous case. What changes is the distribution of ℓ_L, ℓ_R and ℓ_N, as they now follow Binomial distributions with probability parameter q (rather than $1-q$). Hence,

$$\mathbb{E}\left[\mu_i(\omega = L) \mid \ell_N, \omega = L\right] = \sum_{k=0}^{d_L+d_R} \left[\frac{(d_L+d_R)!}{k!(d_L+d_R-k)!} \cdot q^k (1-q)^{d_L+d_R-k} \cdot \frac{\pi}{\pi + (1-\pi)Q^{k-d_R+2\ell_N-N}\Gamma^{k-d_L}}\right].$$

Recall that ℓ_N is a Binomial r.v. with parameters q and N. Hence, the expected posterior belief conditional on $\omega = R$ is given by

$$\mathbb{E}\left[\mu_i(\omega = L) \mid \omega = L\right] = \sum_{k=0}^{N} \frac{N!}{k!(N-k)!} q^k (1-q)^{N-k} \mathbb{E}\left(\mu_i(\omega = L) | k, \omega = L\right).$$

Following a similar proof method, suppose $d_L + d_R$ is odd and rewrite one of the conditional expected posteriors above as follows:

$$\mathbb{E}\left[\mu_i(\omega = L) | \ell_N, \omega = L\right] = \sum_{k=0}^{\frac{d_L+d_R-1}{2}} \frac{(d_L+d_R)!}{k!(d_L+d_R-k)!} \Big(\hat{f}(k, q, \ell_N) + \hat{f}(d_L + d_R - k, q, \ell_N)\Big),$$

where

$$\hat{f}(k, q, \ell_N) = q^k (1-q)^{d_L+d_R-k} \frac{\pi}{\pi + (1-\pi)Q^{k-d_R+2\ell_N-n_i}\Gamma^{k-d_L}}$$

Consider the derivative of $\hat{f}(k, q, \ell_N)$ w.r.t. q at $q = \frac{1}{2}$:

$$\begin{aligned}
\frac{\partial}{\partial q}\hat{f}(k, q, \ell_N) &= \Big(-(d_L+d_R-k)q^k(1-q)^{d_{Li}+d_R-k-1} + kq^{k-1}(1-q)^{d_L+d_R-k}\Big) \cdot \\
&\quad \frac{\pi}{\pi + (1-\pi)Q^{k-d_R+2\ell_N-N+(k-d_L)z(q,\gamma)}} \\
&\quad + q^k(1-q)^{d_L+d_R-k} \cdot \pi(1-\pi) \cdot \Bigg[\frac{(k-d_R+2\ell_N-N)Q^{k-d_R+2\ell_N-N-1+(k-d_L)z(q,\gamma)}\frac{1}{q^2}}{\left(\pi + (1-\pi)Q^{k-d_R+2\ell_N-N}\Gamma^{k-d_L}\right)^2} \\
&\quad + \frac{(k-d_L)Q^{k-d_R+2\ell_N-N+(k-d_L-1)z(q,\gamma)}\frac{(2-\gamma)\gamma}{(\gamma q+(1-\gamma))^2}}{\left(\pi + (1-\pi)Q^{k-d_R+2\ell_N-N}\Gamma^{k-d_L}\right)^2}\Bigg] \\
&\overset{(q=\frac{1}{2})}{=} \left(\frac{1}{2}\right)^{d_L+d_R-1} (2k-d_L-d_R)\frac{\pi}{\pi+(1-\pi)} + \left(\frac{1}{2}\right)^{d_L+d_R} 4\pi(1-\pi) \cdot \\
&\quad \cdot \frac{(k-d_R+2\ell_N-N) + (k-d_L)\frac{\gamma}{2-\gamma}}{(\pi+(1-\pi))^2} \\
&= \left(\frac{1}{2}\right)^{d_L+d_R-1} \cdot \Big((2k-d_L-d_R)\pi + 2\pi(1-\pi)\Big((k-d_R+2\ell_N-N) + (k-d_L)\frac{\gamma}{2-\gamma}\Big)\Big).
\end{aligned}$$

Therefore, at $q = \frac{1}{2}$ we have

$$\begin{aligned}
\frac{\partial}{\partial q}\mathbb{E}\left[\mu_i(\omega = L)|\ell_N, \omega = L\right] &= \sum_{k=0}^{\frac{d_L+d_R-1}{2}} \frac{(d_L+d_R)!}{k!(d_L+d_R-k)!}\left(\frac{\partial}{\partial q}\hat{f}(k,q,\ell_N) + \frac{\partial}{\partial q}\hat{f}(d_L+d_R-k,q,\ell_N)\right) \\
&= \sum_{k=0}^{\frac{d_L+d_R-1}{2}} \frac{(d_L+d_R)!}{k!(d_L+d_R-k)!}\left(\frac{1}{2}\right)^{d_L+d_R-1} 2\pi(1-\pi)\cdot \\
&\quad \cdot\left(2(2\ell_N - N) + (d_L - d_R) - (k - d_L)\frac{\gamma}{2-\gamma}\right) \\
&= 4\pi(1-\pi)(2\ell_N - N) + \\
&\quad + \sum_{k=0}^{\frac{d_L+d_R-1}{2}} \frac{(d_L+d_R)!}{k!(d_L+d_R-k)!}\left(\frac{1}{2}\right)^{d_L+d_R-1} 2\pi(1-\pi)\left(d_L - d_R\right)\tfrac{2-2\gamma}{2-\gamma}.
\end{aligned}$$

Note that the second term (a sum) in the expression above is strictly positive for any $\ell_N \geq 0$, since $d_L > d_R$ and $\gamma \in (0,1)$ hold. Thus, when we compute the derivative of $\mathbb{E}\left[\mu_i(\omega = L)|\omega = L\right]$, this part of the sum will remain strictly positive. The only term to be concerned about is the one with $(2\ell_N - N)$ in it.

Let

$$H = \sum_{k=0}^{\frac{d_L+d_R-1}{2}} \frac{(d_L+d_R)!}{k!(d_L+d_R-k)!}\left(\frac{1}{2}\right)^{d_L+d_R-1} 2\pi(1-\pi)\left(d_L - d_R\right)\frac{2-2\gamma}{2-\gamma}.$$

Then the derivative of the conditional expected posterior at $q = \frac{1}{2}$ is equal to

$$\begin{aligned}
\frac{\partial}{\partial q}\mathbb{E}\left[\mu_i(\omega = L)|\omega = L\right] &= \sum_{k=0}^{N} \frac{N!}{k!(N-k)!}\left(\frac{1}{2}\right)^{N-1}(k-(N-k))\,\mathbb{E}\left[\mu_i(\omega = L)|k, \omega = L, q = \frac{1}{2}\right] + \\
&\quad + \sum_{k=0}^{N} \frac{N!}{k!(N-k)!}\left(\frac{1}{2}\right)^{N}\left[H + 4\pi(1-\pi)(2k-N)\right] \\
&= H + \sum_{k=0}^{N} \frac{N!}{k!(N-k)!}\left(\frac{1}{2}\right)^{N-1}\left[2\pi + 4\pi(1-\pi)\right](2k-N)
\end{aligned}$$

We know that $H > 0$, whereas the second sum is symmetric around $k = \frac{N}{2}$ and will therefore annihilate to 0. Thus, the derivative of the expected posterior is positive at $q = \frac{1}{2}$, which then implies that the expected posterior is bigger than prior for some range above $q = \frac{1}{2}$. In fact, given that $\omega = L$ is the true state, it is reasonable to expect that it is bigger than prior for all $q > \frac{1}{2}$, as both the received information and the dogmatic-friends bias work in the same direction.

We have proved that both conditional expected posteriors will be larger than the prior π in some range $\left(\frac{1}{2}, \bar{q}_1\right)$. This also implies that the unconditional expected posterior belief of i in state $\omega = L$ must be strictly higher than π for any $q \in \left(\frac{1}{2}, \bar{q}_1\right)$.

The final part of the proof will touch on the case where $d_L + d_R$ is even. We will do it for the case $\omega = R$, with $\omega = L$ being completely analogous. When $d_L + d_R$ is even, we can rewrite the expression for $\mathbb{E}\left[\mu_i(\omega = L)|\omega = R\right]$ as follows:

$$\begin{aligned}
\mathbb{E}\left[\mu_i(\omega = L)|\ell_N, \omega = R\right] &= \sum_{k=0}^{d_L+d_R} \frac{(d_L+d_R)!}{k!(d_L+d_R-k)!} \cdot (1-q)^k q^{d_L+d_R-k} \cdot \\
&\quad \cdot \frac{\pi}{\pi + (1-\pi)Q^{k-d_R+2\ell_N-N}\Gamma^{k-d_L}} \\
&= \sum_{k=0}^{\frac{d_L+d_R}{2}-1} \frac{(d_L+d_R)!}{k!(d_L+d_R-k)!} \Big(f(k,q,\ell_N) + f(d_L+d_R-k,q,\ell_N) \Big) + \\
&\quad + \frac{(d_L+d_R)!}{\left(\frac{d_L+d_R}{2}\right)! \left(\frac{d_L+d_R}{2}\right)!} \left(f\left(\tfrac{d_L+d_R}{2}, q, \ell_N\right) + f\left(\tfrac{d_L+d_R}{2}, q, \ell_N\right) \right)
\end{aligned}$$

The derivative of the first sum at $q = \frac{1}{2}$ will be positive, as we showed above. As for the second term, note that at $q = \frac{1}{2}$ we have

$$\begin{aligned}
\frac{\partial}{\partial q} f\left(\frac{d_L+d_R}{2}, q, \ell_N\right)\bigg|_{q=\frac{1}{2}} &= \left(\frac{1}{2}\right)^{d_L+d_R-1} \\
&\quad \cdot \left(0 + 2\pi(1-\pi)\left(\left(\tfrac{d_L+d_R}{2} - d_R + 2\ell_N - N\right) + \left(\tfrac{d_L+d_R}{2} - d_L\right)\tfrac{\gamma}{2-\gamma}\right)\right) \\
&= 2\pi(1-\pi)\left(\frac{1}{2}\right)^{d_L+d_R-1} \cdot \left(2\ell_N - N + \tfrac{d_L-d_R}{2}\left(1 - \tfrac{\gamma}{2-\gamma}\right)\right).
\end{aligned}$$

Note that the term $(2\ell_N - N)$ is symmetric around $\ell_N = \frac{N}{2}$, and so will annihilate to zero when we compute expectation over ℓ_N. The second term is positive for all $\ell_N \geq 0$. Hence, $\frac{\partial}{\partial q}\mathbb{E}\left[\mu_i(\omega = L)\right] > 0$ still holds at $q = \frac{1}{2}$, and the argument follows exactly as in the case of $d_L + d_R$ being odd. We will omit it here for brevity.

This completes the proof. ■

B Proof of Proposition 2

Proof of Proposition 2. Suppose that the true state is $\omega = R$. Recall that i has d_{Li} dogmatic-left friends, d_{Ri} dogmatic-right friends, and n_i "normal" friends. Suppose that T periods have passed; this means that S has sent T time-independent and individual-independent signals to all agents in the network. Denote the number of signals $s = l$ received by:

- agent i as l_i
- dogmatic-left friends of agent j as l_j^L, $j \in \{1, 2, \ldots, d_L\}$
- dogmatic-right friends of agent j as l_j^R, $j \in \{1, 2, \ldots, d_R\}$

- normal friends of agent i as l_j^n, $j \in \{1, 2, \ldots, n\}$

Then the number of signals $s = r$ received by the same agents will be given by:

- agent i – it is $(T - l_i)$
- dogmatic-left friends of agent i – it is $(T - l_j^L)$, $j \in \{1, 2, \ldots, n_i\}$
- dogmatic-right friends of agent i – it is $(T - l_j^R)$, $j \in \{1, 2, \ldots, m\}$
- normal friends of agent i – it is $(T - l_j^n)$, $j \in \{1, 2, \ldots, k\}$

Additionally, a dogmatic-left friend j has stayed silent exactly $(T - l_j^L)$ times, and a right-leaning friend k has stayed silent exactly l_k^R times. Thus, agent i's posterior beliefs at the end of period T is given by the following:

$$\begin{aligned}\mu_i(\omega = L|\mathbf{s}_i^T) &= \pi \cdot \Bigg[\pi + (1-\pi) \cdot Q^{(2l_i - T) + \sum_{j=1}^{n}(2l_j^n - T) + \sum_{j=1}^{d_L} l_j^L - \sum_{j=1}^{d_R}(T - l_j^R)} \\ &\qquad \cdot \Gamma^{\sum_{j=1}^{d_R} l_j^R - \sum_{j=1}^{d_L}(T - l_j^L)}\Bigg]^{-1},\end{aligned}$$

where $Q = \frac{1-q}{q}$ and $\Gamma = \frac{\gamma(1-q)+(1-\gamma)}{\gamma q + (1-\gamma)}$.

This posterior belief converges to 1 with probability 1 as $T \to \infty$ if and only if

$$\left(\frac{1-q}{q}\right)^{(2l_i - T) + \sum_{j=1}^{n}(2l_j^n - T) + \sum_{j=1}^{d_L} l_j^L - \sum_{j=1}^{d_R}(T - l_j^R)} \cdot \left(\frac{\gamma(1-q)+(1-\gamma)}{\gamma q + (1-\gamma)}\right)^{\sum_{j=1}^{d_R} l_j^R - \sum_{j=1}^{d_L}(T - l_j^L)}$$

converges to zero with probability 1 as $T \to \infty$. This is equivalent to requiring that the natural log of this expression converges to $-\infty$ with probability 1 as $T \to \infty$, where the natural log equals $\ln\left(\frac{1-q}{q}\right) K(\mathbf{x}, T; q, \gamma)$, where

$$\begin{aligned}K(\mathbf{x}, T; q, \gamma) &= (2l_i - T) + \sum_{j=1}^{n}(2l_j^n - T) + \sum_{i=1}^{d_L} l_j^L - \sum_{j=1}^{d_R}(T - l_j^R) \\ &\quad + \left(\sum_{j=1}^{d_R} l_j^R - \sum_{j=1}^{d_L}(T - l_j^L)\right) z(q, \gamma),\end{aligned}$$

after we define

$$\mathbf{x} = (l_i, (l_j^{n_i})_{j=1}^{n}, (l_j^L)_{j=1}^{d_L}, (l_j^R)_{j=1}^{d_R})$$

and

$$z(q, \gamma) = \ln\left(\frac{\gamma(1-q)+(1-\gamma)}{\gamma q + (1-\gamma)}\right)\left[\ln\left(\frac{1-q}{q}\right)\right]^{-1}.$$

Given $\ln\left(\frac{1-q}{q}\right) < 0$, we require that $K(\mathbf{x}, T; q, \gamma)$ converges to $+\infty$ with probability 1 as $T \to \infty$. Now note that

$$\lim_{T\to\infty} K(\mathbf{x}, T; q, \gamma) = \lim_{T\to\infty} T\left(\frac{K(\mathbf{x}, T; q, \gamma)}{T}\right).$$

By the Law of Large Numbers, we have that

$$\begin{aligned}
\operatorname*{plim}_{T\to\infty} \frac{K(\mathbf{x}, T; q, \gamma)}{T} &= (2(1-q)-1) + \sum_{j=1}^{n}(2(1-q)-1) \\
&\quad + \sum_{j=1}^{d_L}(1-q) - \sum_{j=1}^{d_R}(1-(1-q)) + \\
&\quad + \left(\sum_{j=1}^{d_R}(1-q) - \sum_{j=1}^{d_L}(1-(1-q))\right) z(q,\gamma) \\
&= (d_L + 2n + 2)(1-q) - (n+1) - d_R q \\
&\quad + (d_R(1-q) - d_L q)\, z(q,\gamma).
\end{aligned}$$

Given this, $\operatorname{plim}_{T\to\infty} K(\mathbf{x}, T; q, \gamma) = +\infty$ if and only if

$$(d_L + 2n + 2)(1-q) - (n+1) - d_R q + (d_R(1-q) - d_{Li} q)\, z(q,\gamma) > 0,$$

which is equivalent to

$$q < \bar{q}_2(q) \equiv \frac{1 + d_L + n + d_R z(q,\gamma)}{2 + 2n + (d_L + d_R)(1 + z(q,\gamma))}.$$

To check whether there are values of $q \in (\frac{1}{2}; 1)$ for which the last inequality holds, let us check whether it holds strictly for $q = \frac{1}{2}$. If it does, then due to continuity of $\bar{q}_2(q)$, there will be a range of values of q for which the inequality must hold.

Before we proceed, note the following:

$$\begin{aligned}
\lim_{q\to\frac{1}{2}^+} z(q,\gamma) &= \lim_{q\to\frac{1}{2}^+} \frac{\ln(\gamma(1-q) + (1-\gamma)) - \ln(\gamma q + (1-\gamma))}{\ln(1-q) - \ln(q)} \\
&= \lim_{q\to\frac{1}{2}^+} \frac{-\frac{\gamma}{\gamma(1-q)+(1-\gamma)} - \frac{\gamma}{\gamma q+(1-\gamma)}}{-\frac{1}{1-q} - \frac{1}{q}} \\
&= \frac{-\frac{\gamma}{\frac{1}{2}\gamma+(1-\gamma)} - \frac{\gamma}{\frac{1}{2}\gamma+(1-\gamma)}}{-\frac{1}{1-\frac{1}{2}} - \frac{1}{\frac{1}{2}}} \\
&= \frac{2\gamma}{4\cdot\left(\frac{1}{2}\gamma + (1-\gamma)\right)} = \frac{\gamma}{2-\gamma}
\end{aligned}$$

Thus,

$$\begin{aligned}
\lim_{q\to\frac{1}{2}} \bar{q}_2\,(q) &= \frac{1+d_L+n+d_R\cdot\frac{\gamma}{2-\gamma}}{2+2n+(d_L+d_R)(1+\frac{\gamma}{2-\gamma})} \\
&= \frac{(1+d_L+n)(2-\gamma)+d_R\gamma}{2(2-\gamma)(1+n)+2(d_L+d_R)} \\
&= \frac{(2-\gamma)(1+n)+(2-\gamma)d_L+d_{Ri}\gamma}{2(2-\gamma)(1+n)+2(d_L+d_R)} \\
&= \frac{(2-\gamma)(1+n)+(d_L+d_R)+(d_L-d_R)(1-\gamma)}{2(2-\gamma)(1+n_i)+2(d_L+d_R)} \\
&= \frac{1}{2}+\frac{(d_L-d_R)(1-\gamma)}{2(2-\gamma)(1+n)+2(d_L+d_R)}
\end{aligned}$$

This is strictly bigger than $\frac{1}{2}$ in any situation where $d_L > d_R$ and $\gamma < 1$. Thus, $q < \bar{q}_2(q)$ holds at $q = \frac{1}{2}$, and, due to continuity of $\bar{q}_2(q)$, it holds for an interval to the right of $q = \frac{1}{2}$. Defining $\bar{q}_2$ through the implicit equation $q = \bar{q}_2(q)$, we can conclude that the posterior belief of agent i converges to 1 with probability 1 as $T \to \infty$ for any q in the range $\left(\frac{1}{2}, \bar{q}_2\right)$.

Note that

$$\begin{aligned}
\lim_{q\to 1^-} z(q,\gamma) &= \lim_{q\to 1^-} \frac{\ln\left(\gamma(1-q)+(1-\gamma)\right)-\ln\left(\gamma q+(1-\gamma)\right)}{\ln\left(1-q\right)-\ln\left(q\right)} \\
&= \lim_{q\to 1^-} \frac{-\frac{\gamma}{\gamma(1-q)+(1-\gamma)}-\frac{\gamma}{\gamma q+(1-\gamma)}}{-\frac{1}{1-q}-\frac{1}{q}} = 0
\end{aligned}$$

Thus,

$$\begin{aligned}
\lim_{q\to 1^-} \bar{q}_2\,(q) &= \frac{1+d_L+n}{2+2n+d_L+d_R} \\
&= \frac{1+d_L+n}{1+d_L+n+1+d_R+n} \in \left(\frac{1}{2},1\right),
\end{aligned}$$

because $d_L > d_R$ by assumption. By continuity of $\bar{q}_2(q)$, this implies that there is an interval to the left of $q = 1$ such that $q > \bar{q}_2(q)$. Note that if $q > \bar{q}_2$, we have $\operatorname{plim}_{T\to\infty} K(\mathbf{x}, T; q, \gamma) = \infty$, which means that $\operatorname{plim}_{T\to\infty} \mu_i(\omega = L|\mathbf{s}_i^T) = 0$ and therefore agent i learns the state $\omega = R$ correctly.

When $q = \bar{q}_2$, the limit $\operatorname{plim}_{T\to\infty} K(\mathbf{x}, T; q, \gamma)$ is indeterminate, but we can ignore this knife-edge case.

In the last part of the proof, we will touch on the case $\omega = L$. This only affects distribution of ℓ_L, ℓ_R and ℓ_N in their probability parameter, which changes from $1-q$ to q. Therefore, the proof above still applies up to the application of the Law of Large numbers.

The posterior belief of the agent will converge to 1 with probability 1 as $T \to \infty$ if and only if

$$K(\mathbf{x}, T; q, \gamma) = (2l_i - T) + \sum_{j=1}^{n}(2l_j^n - T) + \sum_{j=1}^{d_L} l_j^L - \sum_{j=1}^{d_R}(T - l_j^R) + \left(\sum_{j=1}^{d_R} l_j^R - \sum_{j=1}^{d_L}(T - l_j^L)\right) z(q, \gamma)$$

converges to $+\infty$ with probability 1 as $T \to \infty$. Similarly to before, note

$$K(\mathbf{x}, T; q, \gamma) = T \cdot \frac{K(\mathbf{x}, T; q, \gamma)}{T}$$

and apply the Law of Large Numbers to the fraction:

$$\begin{aligned}\operatorname*{plim}_{T\to\infty} \frac{K(\mathbf{x}, T; q, \gamma)}{T} &= (2q-1) + \sum_{j=1}^{n}(2q-1) + \sum_{j=1}^{d_L} q - \sum_{j=1}^{d_R}(1-q) + \left(\sum_{j=1}^{d_R} q - \sum_{j=1}^{d_L}(1-q)\right) z(q, \gamma) \\ &= (d_L + 2n + 2)q - (n+1) - d_R(1-q) + (d_R q - d_L(1-q))\, z(q, \gamma).\end{aligned}$$

Given this, $\operatorname{plim}_{T\to\infty} K(\mathbf{x}, T; q, \gamma) = +\infty$ if and only if

$$(d_L + 2n + 2)q - (n+1) - d_R(1-q) + (d_R q - d_{Li}(1-q))\, z(q, \gamma) > 0,$$

which is equivalent to

$$\frac{1 + n + d_R + d_L z(q, \gamma)}{2 + 2n + (d_L + d_R)(1 + z(q, \gamma))} < q.$$

Note

$$\frac{1 + n + d_R + d_L z(q, \gamma)}{2 + 2n + (d_L + d_R)(1 + z(q, \gamma))} = \frac{1}{2} - \frac{\left(1 - \frac{1 + z(q,\gamma)}{2}\right)(d_L - d_R)}{2 + 2n + (d_L + d_R)(1 + z(q, \gamma))} < \frac{1}{2},$$

which implies that the desired inequality on q holds for any $q > \frac{1}{2}$. Thus, given $\omega = L$, $d_L > d_R$ and $\gamma \in (0, 1)$, the agent's posterior belief converges to probability 1 on state $\omega = L$ as $T \to \infty$ with probability 1 for any $q \in \left(\frac{1}{2}, 1\right)$.

Therefore, provided that $q \in \left(\frac{1}{2}, \bar{q}_2\right)$, the posterior belief will converge to $\mu_i^\infty(\omega = L) = 1$ with probability 1 as $T \to \infty$ irrespective of what the true state is.

■

Moreover, as Lemmas 2 and 3 show below, the function $\bar{q}_2(q)$ is increasing in q, which implies that we can put a lower bound on the value of $\bar{q}_2$. As it is defined by the equation $q = \bar{q}_2(q)$, we can conclude that $\bar{q}_2 \geq \bar{q}_2\left(\frac{1}{2}\right)$, i.e.

$$\bar{q}_2 \geq \frac{1}{2} + \frac{(d_L - d_R)(1-\gamma)}{2(2-\gamma)(1+n) + 2(d_L + d_R)}.$$

Lemma 2. $\bar{q}_2(q)$ *is increasing in* q.

Proof. To prove Lemma 2 we first establish the following intermediate result.

Lemma 3. $z(q,\gamma)$ *is weakly decreasing for any* $q \in \left(\frac{1}{2};1\right)$.

Proof. Note:

$$
\begin{aligned}
\frac{\partial z}{\partial q} &= \frac{\partial}{\partial q}\frac{\ln\left(\frac{\gamma(1-q)+(1-\gamma)}{\gamma q+(1-\gamma)}\right)}{\ln\left(\frac{1-q}{q}\right)} \\
&= \frac{\frac{\gamma q+(1-\gamma)}{\gamma(1-q)+(1-\gamma)}\cdot\frac{-\gamma(\gamma q+(1-\gamma))-\gamma(\gamma(1-q)+(1-\gamma))}{(\gamma q+(1-\gamma))^2}\cdot\ln\left(\frac{1-q}{q}\right)-\ln\left(\frac{\gamma(1-q)+(1-\gamma)}{\gamma q+(1-\gamma)}\right)\cdot\frac{q}{1-q}\cdot\left(-\frac{1}{q^2}\right)}{\ln^2\left(\frac{1-q}{q}\right)} \\
&= \frac{-\frac{\gamma^2 q+\gamma(1-\gamma)+\gamma^2(1-q)+\gamma(1-\gamma)}{(\gamma(1-q)+(1-\gamma))(\gamma q+(1-\gamma))}\cdot\ln\left(\frac{1-q}{q}\right)+\ln\left(\frac{\gamma(1-q)+(1-\gamma)}{\gamma q+(1-\gamma)}\right)\cdot\frac{1}{q(1-q)}}{\ln^2\left(\frac{1-q}{q}\right)} \\
&= \frac{-\frac{\gamma(2-\gamma)}{(\gamma(1-q)+(1-\gamma))(\gamma q+(1-\gamma))}\cdot\ln\left(\frac{1-q}{q}\right)+\ln\left(\frac{\gamma(1-q)+(1-\gamma)}{\gamma q+(1-\gamma)}\right)\cdot\frac{1}{q(1-q)}}{\ln^2\left(\frac{1-q}{q}\right)} \\
&= \frac{-\frac{\gamma(2-\gamma)}{\gamma^2 q(1-q)+(1-\gamma)}\cdot\ln\left(\frac{1-q}{q}\right)+\ln\left(\frac{\gamma(1-q)+(1-\gamma)}{\gamma q+(1-\gamma)}\right)\cdot\frac{1}{q(1-q)}}{\ln^2\left(\frac{1-q}{q}\right)} \\
&= \frac{-\frac{\gamma(2-\gamma)}{\gamma^2 q(1-q)+(1-\gamma)}\cdot\ln\left(\frac{1-q}{q}\right)+\ln\left(\frac{1-q}{q}\right)\cdot\frac{z(q,\gamma)}{q(1-q)}}{\ln^2\left(\frac{1-q}{q}\right)} \\
&= \frac{-\gamma(2-\gamma)q(1-q)+\gamma^2 q(1-q)z(q,\gamma)+(1-\gamma)z(q,\gamma)}{\ln\left(\frac{1-q}{q}\right)\cdot\left(\gamma^2 q(1-q)+(1-\gamma)\right)q(1-q)} \\
&= \frac{(1-\gamma+\gamma^2 q(1-q))z(q,\gamma)-(2-\gamma)\gamma q(1-q)}{\ln\left(\frac{1-q}{q}\right)\cdot\left(\gamma^2 q(1-q)+(1-\gamma)\right)q(1-q)} \\
&= \frac{\left(\frac{1-\gamma}{q(1-q)}+\gamma^2\right)z(q,\gamma)-(2-\gamma)\gamma}{\ln\left(\frac{1-q}{q}\right)\cdot\left(\gamma^2 q(1-q)+(1-\gamma)\right)}
\end{aligned}
$$

Note that $z\left(\frac{1}{2},\gamma\right) = \frac{\gamma}{2-\gamma} > 0 = z(1,\gamma)$. As $z(q,\gamma)$ is continuously differentiable for $q \in \left(\frac{1}{2},1\right)$, it is enough to prove that there are no local maxima on that interval in order to show that $\frac{\partial z}{\partial q} \leq 0$ holds on that interval. At an intermediate local maximum, $\frac{\partial z}{\partial q} = 0$ must hold. Consider:

$$
\begin{aligned}
\frac{\partial z}{\partial q} = 0 \;\Rightarrow\; \left(\frac{1-\gamma}{q(1-q)} + \gamma^2\right) z(q,\gamma) - (2-\gamma)\gamma = 0 \\
z(q,\gamma) &= \frac{\gamma(2-\gamma)}{\gamma^2 + \frac{1-\gamma}{q(1-q)}} \\
&\leq \frac{\gamma(2-\gamma)}{\gamma^2 + \frac{1-\gamma}{\frac{1}{4}}} \\
&= \frac{\gamma(2-\gamma)}{(2-\gamma)^2} \\
&= \frac{\gamma}{2-\gamma}
\end{aligned}
$$

Hence, $z(q,\gamma) \leq \frac{\gamma}{2-\gamma}$ must hold at any intermediate local maximum in $\left(\frac{1}{2}, 1\right)$. This immediately rules our the possibility that $z(q,\gamma)$ is increasing at $q = \frac{1}{2}$, since otherwise it would need to achieve a local maximum with value above $\frac{\gamma}{2-\gamma}$.

Consider again the implication of $\frac{\partial z}{\partial q} = 0$:

$$
z(q,\gamma) = \frac{\gamma(2-\gamma)}{\gamma^2 + \frac{1-\gamma}{q(1-q)}}
$$

Note that the right-hand side is strictly decreasing in q for $q \in \left(\frac{1}{2}; 1\right)$. If $z(q,\gamma)$ was to decrease at first (as q rises from $\frac{1}{2}$) and then increase before going down to 0, the value of $z(q,\gamma)$ at the corresponding local maximum would be necessarily above the right-hand side. This is a contradiction, and thus $z(q,\gamma)$ cannot be strictly increasing for any $q \in \left(\frac{1}{2}; 1\right)$.

One final case to rule out is that $z(q,\gamma)$ is decreasing at first, passing through a local minimum, and then is increasing all the way until $q = 1$. That, however, would mean that the value at the local minimum is less than $z(1,\gamma)$, which is equal to 0. Since $z(q,\gamma) > 0$ for any $q \in (\frac{1}{2}; 1)$ and $\gamma \in (0; 1)$, this case is also impossible.

These considerations prove that $z(q,\gamma)$ is weakly decreasing for any $q \in \left(\frac{1}{2}; 1\right)$. ■

To complete the proof of Lemma 2, note that

$$
\begin{aligned}
\frac{\partial \bar{q}_2(q)}{\partial q} &= \frac{d_R \frac{\partial z}{\partial q}(2+2n+(d_L+d_R)(1+z(q,\gamma))) - (1+d_L+n+d_R z(q,\gamma))(d_L+d_R)\frac{\partial z}{\partial q}}{(2+2n+(d_L+d_R)(1+z(q,\gamma)))^2} \\
&= \frac{\frac{\partial z}{\partial q}\left(2d_R+2d_R k+(d_L d_R+d_R^2)(1+z(q,\gamma))-((1+d_L+n)(d_L+d_R)+(d_L+d_R)d_R z(q,\gamma))\right)}{(2+2n+(d_L+d_R)(1+z(q,\gamma)))^2} \\
&= \frac{\frac{\partial z}{\partial q}\left(2d_R(1+n)+d_R(d_L+d_R)-(1+d_L+n)(d_L+d_R)\right)}{(2+2n+(d_L+d_R)(1+z(q,\gamma)))^2} \\
&= \frac{\frac{\partial z}{\partial q}(d_R-d_L)(1+d_R+n_i)}{(2+2n+(d_L+d_R)(1+z(q,\gamma)))^2} \geq 0,
\end{aligned}
$$

since $(d_R - d_L) < 0$ and, $\frac{\partial z}{\partial q} \leq 0$ by Lemma 3 we have the result.

$\blacksquare$

Proof that $\bar{q}_2$ is unique.

We would like to prove that $\bar{q}_2(q)$ is convex in q. This, along with $\bar{q}_2(\frac{1}{2}) > \frac{1}{2}$ and $\bar{q}_2(1) < 1$, will imply that $\bar{q}_2(q) = q$ has a unique solution. For the purposes of this argument, we will still assume $d_L > d_R$ and a corresponding functional form of $\bar{q}_2(q)$.

Note that

$$
\bar{q}_2(q) = \frac{1+d_L+n+d_R z(q,\gamma)}{2+2n+(d_L+d_R)(1+z(q,\gamma))} = \frac{A+B\cdot z(q,\gamma)}{C+D\cdot z(q,\gamma)},
$$

where

$$
A = 1+d_L+n,\ B = d_R,\ C = 2+2n+d_L+d_R,\ D = d_L+d_R.
$$

Therefore, we can represent

$$
\bar{q}_2(q) = \frac{B}{D} + \frac{A-\frac{BC}{D}}{C+D\cdot z(q,\gamma)} = \frac{B}{D} + \frac{AD-BC}{D(C+Dz(q,\gamma))}.
$$

Note that

$$
\begin{aligned}
AD - BC &= (1+d_L+n)(d_L+d_R) - d_R(2+2n+d_L+d_R) \\
&= (d_L-d_R)(d_L+d_R) + (1+n)(d_L+d_R) - 2d_R(1+n) \\
&= (d_L-d_R)(d_L+d_R) + (1+n)(d_L-d_R) \\
&= (d_L-d_R)(1+n+d_L+d_R) > 0.
\end{aligned}
$$

$\bar{q}_2(q)$ is convex in q if and only if $\frac{AD-BC}{D(C+Dz(q,\gamma))}$ is convex in q. Since $AD-BC > 0$ and $D > 0$, it follows that $\bar{q}_2(q)$ is convex in q if and only if

$$
g(q) = \frac{1}{C+Dz(q,\gamma)} \quad \text{is convex in } q.
$$

Note:

$$g'(q) = -\frac{D}{(C + Dz(q,\gamma))^2} \cdot z_q(q,\gamma)$$

$$\begin{aligned} g''(q) &= \frac{2D^2}{(C + Dz(q,\gamma))^3} \cdot \left(z_q(q,\gamma)\right)^2 - \frac{D}{(C + Dz(q,\gamma))^2} \cdot z_{qq}(q,\gamma) \\ &= \frac{2D^2 \left(z_q(q,\gamma)\right)^2 - D(C + Dz(q,\gamma))z_{qq}(q,\gamma)}{(C + Dz(q,\gamma))^3} \end{aligned}$$

If we are able to prove that $z_{qq}(q,\gamma) < 0$ for all $q \in \left(\frac{1}{2}, 1\right)$, then we will have proven that $g''(q) > 0$, implying that $\bar{q}_2(q)$ is convex for all $q \in \left(\frac{1}{2}, 1\right)$.

Recall from Lemma 3:

$$z_q(q,\gamma) = \frac{\left(\frac{1-\gamma}{q(1-q)} + \gamma^2\right) z(q,\gamma) - \gamma(2-\gamma)}{\ln\left(\frac{1-q}{q}\right)\left(\gamma^2 q(1-q) + (1-\gamma)\right)}$$

Let $K(q) = \frac{1}{\ln^2\left(\frac{1-q}{q}\right)\left(\gamma^2 q(1-q) + (1-\gamma)\right)^2}$. Then:

$$\begin{aligned} z_{qq}(q,\gamma) = K(q)\Bigg[&\left(-\frac{(1-\gamma)(1-2q)}{q^2(1-q)^2} + \left(\frac{1-\gamma}{q(1-q)} + \gamma^2\right) z_q(q,\gamma)\right) \ln\left(\frac{1-q}{q}\right)\left(\gamma^2 q(1-q) + (1-\gamma)\right) \\ &- \left(\left(\frac{1-\gamma}{q(1-q)} + \gamma^2\right) z(q,\gamma) - \gamma(2-\gamma)\right)\left(\frac{-1}{q(1-q)}\left(\gamma^2 q(1-q) + (1-\gamma)\right) + \ln\left(\frac{1-q}{q}\right)\gamma^2(1-2q)\right) \\ = K(q)\Bigg[&\left(\frac{(1-\gamma)(2q-1)}{q^2(1-q)^2} + \left(\frac{1-\gamma}{q(1-q)} + \gamma^2\right) z_q(q,\gamma)\right) \ln\left(\frac{1-q}{q}\right)\left(\gamma^2 q(1-q) + (1-\gamma)\right) + \\ &+ \left(\left(\frac{1-\gamma}{q(1-q)} + \gamma^2\right) z(q,\gamma) - \gamma(2-\gamma)\right)\left(\frac{1}{q(1-q)}\left(\gamma^2 q(1-q) + (1-\gamma)\right) + \ln\left(\frac{1-q}{q}\right)\gamma^2(2q-1)\right) \end{aligned}$$

Let

$$\begin{aligned} C_1(q) &= \frac{(1-\gamma)(2q-1)}{q^2(1-q)^2} + \left(\frac{1-\gamma}{q(1-q)} + \gamma^2\right) z_q(q,\gamma) \\ C_2(q) &= \left(\frac{1-\gamma}{q(1-q)} + \gamma^2\right) z(q,\gamma) - \gamma(2-\gamma) \\ C_3(q) &= \frac{1}{q(1-q)}\left(\gamma^2 q(1-q) + (1-\gamma)\right) + \ln\left(\frac{1-q}{q}\right)\gamma^2(2q-1) \end{aligned}$$

Then we can write

$$z_{qq}(q,\gamma) = K(q)\left[C_1(q) \ln\left(\frac{1-q}{q}\right)\left(\gamma^2 q(1-q) + (1-\gamma)\right) + C_2(q)C_3(q)\right]$$

Consider $C_1(q)$:

$$\begin{aligned}
C_1(q) &= \frac{(1-\gamma)(2q-1)}{q^2(1-q)^2} + \left(\frac{1-\gamma}{q(1-q)} + \gamma^2\right) z_q(q,\gamma) \\
&= \frac{(1-\gamma)(2q-1)}{q^2(1-q)^2} + \frac{\gamma^2 q(1-q) + (1-\gamma)}{q(1-q)} \cdot \frac{\left(\frac{1-\gamma}{q(1-q)} + \gamma^2\right) z(q,\gamma) - \gamma(2-\gamma)}{\ln\left(\frac{1-q}{q}\right)(\gamma^2 q(1-q) + (1-\gamma))} \\
&= \frac{(1-\gamma)(2q-1)\ln\left(\frac{1-q}{q}\right) + \left(\gamma^2 q(1-q) + (1-\gamma)\right) z(q,\gamma) - \gamma(2-\gamma)q(1-q)}{q^2(1-q)^2 \ln\left(\frac{1-q}{q}\right)} \\
&= \frac{(1-\gamma)(2q-1)\ln\left(\frac{q}{1-q}\right) + \gamma(2-\gamma)q(1-q) - \left(\gamma^2 q(1-q) + (1-\gamma)\right) z(q,\gamma)}{q^2(1-q)^2 \ln\left(\frac{q}{1-q}\right)}
\end{aligned}$$

Therefore,

$$\begin{aligned}
C_1(q)\ln\left(\frac{1-q}{q}\right)\Big(\gamma^2 q(1-q) + (1-\gamma)\Big) &= -\Big(\gamma^2 q(1-q) + (1-\gamma)\Big) \cdot \\
&\cdot \frac{(1-\gamma)(2q-1)\ln\left(\frac{q}{1-q}\right) + \gamma(2-\gamma)q(1-q) - \left(\gamma^2 q(1-q) + (1-\gamma)\right) z(q,\gamma)}{q^2(1-q)^2}
\end{aligned}$$

Similarly, consider:

$$\begin{aligned}
C_2(q)C_3(q) = &\frac{\left(\gamma^2 q(1-q) + (1-\gamma)\right) z(q,\gamma) \cdot \left[\left(\gamma^2 q(1-q) + (1-\gamma)\right) + \ln\left(\frac{1-q}{q}\right)\gamma^2(2q-1)q(1-q)\right]}{q^2(1-q)^2} \\
&- \gamma(2-\gamma)\frac{\left(\gamma^2 q(1-q) + (1-\gamma)\right) q(1-q) + \ln\left(\frac{1-q}{q}\right)\gamma^2(2q-1)q^2(1-q)^2}{q^2(1-q)^2}
\end{aligned}$$

Therefore, we can write:

$$\begin{aligned}
\frac{z_{qq}(q,\gamma)q^2(1-q)^2}{K(q)} &= \left(2\left(\gamma^2 q(1-q)+(1-\gamma)\right)+\ln\left(\frac{1-q}{q}\right)\gamma^2(2q-1)(1-q)\right)\cdot \\
&\qquad\qquad\qquad\qquad \cdot\left(\gamma^2 q(1-q)+(1-\gamma)\right) z(q,\gamma) \\
&\quad+\left(\gamma^2 q(1-q)+(1-\gamma)\right)\left[(1-\gamma)(2q-1)\ln\left(\frac{1-q}{q}\right)-2\gamma(2-\gamma)q(1-q)\right] \\
&\quad+\ln\left(\frac{q}{1-q}\right)\gamma^3(2-\gamma)(2q-1)q^2(1-q)^2 \\
&=2\left(\gamma^2 q(1-q)+(1-\gamma)\right)^2 z(q,\gamma)+\ln\left(\frac{q}{1-q}\right)\gamma^3(2-\gamma)(2q-1)q^2(1-q)^2 \\
&\quad-\left(\gamma^2 q(1-q)+(1-\gamma)\right)\ln\left(\frac{q}{1-q}\right)(2q-1)\left[\gamma^2(1-q)z(q,\gamma)+(1-\gamma)\right] \\
&\quad-2\left(\gamma^2 q(1-q)+(1-\gamma)\right)\gamma(2-\gamma)q(1-q)
\end{aligned}$$

Let

$$D_1(q)=2\left(\gamma^2 q(1-q)+(1-\gamma)\right)z(q,\gamma)-\ln\left(\frac{q}{1-q}\right)(2q-1)(1-\gamma)-2\gamma(2-\gamma)q(1-q)$$

and

$$D_2(q)=\gamma^3(2-\gamma)q^2(1-q)^2-\left(\gamma^2 q(1-q)+(1-\gamma)\right)\gamma^2(1-q)z(q,\gamma)$$

Then we have

$$\frac{z_{qq}(q,\gamma)q^2(1-q)^2}{K(q)}=\left(\gamma^2 q(1-q)+(1-\gamma)\right)D_1(q)+\ln\left(\frac{q}{1-q}\right)(2q-1)D_2(q).$$

Note:

$$\begin{aligned}
D_1(q) &\le 2\left(\gamma^2 q(1-q)+(1-\gamma)\right)\frac{\gamma}{2-\gamma}-\ln\left(\frac{q}{1-q}\right)(2q-1)(1-\gamma)-2\gamma(2-\gamma)q(1-q) \\
&=\frac{1}{2-\gamma}\left[2\gamma^3 q(1-q)+2\gamma(1-\gamma)-\ln\left(\frac{q}{1-q}\right)(2q-1)(1-\gamma)(2-\gamma)-2\gamma(2-\gamma)^2 q(1-q)\right] \\
&=\frac{1}{2-\gamma}\left[2\gamma q(1-q)\left(\gamma^2-(2-\gamma)^2\right)+2\gamma(1-\gamma)-\ln\left(\frac{q}{1-q}\right)(2q-1)(1-\gamma)(2-\gamma)\right] \\
&=\frac{1}{2-\gamma}\left[-8q(1-q)\gamma(1-\gamma)+2\gamma(1-\gamma)-\ln\left(\frac{q}{1-q}\right)(2q-1)(1-\gamma)(2-\gamma)\right] \\
&=\frac{1-\gamma}{2-\gamma}\left[2\gamma(1-4q(1-q))-\ln\left(\frac{q}{1-q}\right)(2q-1)(2-\gamma)\right] \\
&=\frac{1-\gamma}{2-\gamma}E(q)
\end{aligned}$$

for $E(q) = 2\gamma(1-4q(1-q)) - \ln\left(\frac{q}{1-q}\right)(2q-1)(2-\gamma)$. Differentiating this expression with respect to q, observe:

$$\begin{aligned}
E'(q) &= 2\gamma \cdot 4(2q-1) - \frac{1}{q(1-q)}(2q-1)(2-\gamma) - 2\ln\left(\frac{q}{1-q}\right)(2-\gamma) \\
&= (2q-1)\left(4\gamma - \frac{2-\gamma}{q(1-q)}\right) - 2\ln\left(\frac{q}{1-q}\right)(2-\gamma) \\
&< (2q-1)\left(4\gamma - 4(2-\gamma)\right) - 2\ln\left(\frac{q}{1-q}\right)(2-\gamma) \\
&< 0 \text{ for any } q \in \left(\frac{1}{2}, 1\right)
\end{aligned}$$

Therefore, $E(q) < E\left(\frac{1}{2}\right)$ for any $q \in \left(\frac{1}{2}, 1\right)$. Note that

$$E\left(\frac{1}{2}\right) = 2\gamma\left(1 - 4\cdot\frac{1}{4}\right) - \ln(1)\left(2\cdot\frac{1}{2} - 1\right)(2-\gamma) = 0.$$

Therefore, we can conclude

$$D_1(q) \leq \frac{1-\gamma}{2-\gamma}E(q) < 0 \text{ for any } q \in \left(\frac{1}{2}, 1\right).$$

Returning to $D_2(q)$, note:

$$\begin{aligned}
D_2(q) &= \gamma^2(1-q)\left[\gamma(2-\gamma)q^2(1-q) - \left(\gamma^2 q(1-q) + (1-\gamma)\right)z(q,\gamma)\right] \\
&< \gamma^2(1-q)\left[\gamma(2-\gamma)q(1-q) - \left(\gamma^2 q(1-q) + (1-\gamma)\right)z(q,\gamma)\right]
\end{aligned}$$

The expression in the brackets is the negative of the numerator in $z_q(q,\gamma)$. Given that $z_q(q,\gamma)$ is negative and that it includes $\ln\left(\frac{1-q}{q}\right)$, it follows that the numerator has to be positive. This implies that the expression above is negative, and therefore, $D_2(q)$ must be negative as well.

Combining $D_1(q) < 0$ and $D_2(q) < 0$ (for any $q \in \left(\frac{1}{2}, 1\right)$ and recalling

$$\frac{z_{qq}(q,\gamma)q^2(1-q)^2}{K(q)} = \left(\gamma^2 q(1-q) + (1-\gamma)\right)D_1(q) + \ln\left(\frac{q}{1-q}\right)(2q-1)D_2(q),$$

we can conclude that $z_{qq}(q,\gamma) < 0$ for any $q \in \left(\frac{1}{2}, 1\right)$.[6]

Returning to the function $\bar{q}_2(q)$, recall that we showed it is convex in q if $z(q,\gamma)$ is concave. Given that $z_{qq}(q,\gamma) < 0$, it follows that $\bar{q}_2(q)$ is convex. Combining this result with the inequalities

$$\bar{q}_2\left(\frac{1}{2}\right) > \frac{1}{2} \text{ and } \bar{q}_2(1) < 1,$$

[6] This is due to the fact that $K(q) > 0$.

we can finally conclude that the equality $\bar{q}_2(q) = q$ is satisfied only once, which implies that $\bar{q}_2$ is unique.

Comparative statics of $\bar{q}_2$.

Here we will prove that $\bar{q}_2$ is increasing in d_L and decreasingi in d_R, n and γ. The value of $\bar{q}_2$ is determined by the equation

$$q = \frac{1 + d_L + n + d_R z(q,\gamma)}{2 + 2n + (d_L + d_R)(1 + z(q,\gamma))}.$$

We can rewrite it as

$$q - \frac{1}{2} = \frac{(d_L - d_R)\frac{1-z(q,\gamma)}{2}}{2 + 2n + (d_L + d_R)(1 + z(q,\gamma))}$$

The right-hand side is strictly decreasing in d_R and n (while the left-hand side is unaffected), which implies that the fixed point of this equation, $\bar{q}_2$, is decreasing in these two variables.

For γ, recall that we can rewrite the equation as

$$q - \frac{B}{D} = \frac{AD - BC}{D(C + Dz(q,\gamma))},$$

where

$$A = 1 + d_L + n,\ B = d_R,\ C = 2 + 2n + d_L + d_R,\ D = d_L + d_R.$$

We will show that $z(q,\gamma)$ is increasing in γ, which will imply[7] that the right-hand side is decreasing in γ. This, in turn, will imply that $\bar{q}_2$ is decreasing in γ. Consider:

$$\begin{aligned}
\frac{\partial}{\partial\gamma} z(q,\gamma) &= \left[\ln\left(\frac{1-q}{q}\right)\right]^{-1} \cdot \frac{-(1+q)(\gamma q + (1-\gamma)) + (1-q)(\gamma(1-q) + (1-\gamma))}{(\gamma(1-q) + (1-\gamma))(\gamma q + (1-\gamma))} \\
&= \left[\ln\left(\frac{1-q}{q}\right)\right]^{-1} \cdot \frac{-2q(1-\gamma) - (1+q)\gamma q + \gamma(1-q)^2}{(\gamma(1-q) + (1-\gamma))(\gamma q + (1-\gamma))} \\
&= \left[\ln\left(\frac{1-q}{q}\right)\right]^{-1} \cdot \frac{-2q + \gamma(1-q)}{(\gamma(1-q) + (1-\gamma))(\gamma q + (1-\gamma))} > 0,
\end{aligned}$$

where the last inequality holds because $(1-q) < q$, $\gamma < 2$ and $\ln\left(\frac{1-q}{q}\right) < 0$.

Finally, consider d_L. Return to the equation

$$q = \frac{1 + d_L + n + d_R z(q,\gamma)}{2 + 2n + (d_L + d_R)(1 + z(q,\gamma))}.$$

[7] Recall that $AD - BC$ is positive.

Denote the right-hand side by h and take the derivative w.r.t. d_L:

$$\begin{aligned}\frac{\partial h}{\partial d_L} &= \frac{2+2n+(d_L+d_R)(1+z(q,\gamma))-(1+z(q,\gamma))(1+d_L+n+d_R z(q,\gamma))}{(2+2n+(d_L+d_R)(1+z(q,\gamma)))^2} \\ &= \frac{(1-z(q,\gamma))\,(1+n+d_R(1+z(q,\gamma)))}{(2+2n+(d_L+d_R)(1+z(q,\gamma)))^2} > 0,\end{aligned}$$

since $z(q,\gamma) \leq \frac{\gamma}{2-\gamma} < 1$ for any $\gamma \in (0,1)$ and $q \in (\frac{1}{2},1)$.

C Proof of Proposition 3

Proof of Proposition 3. First, consider Proposition 2, since it is simpler. To prove that there is no polarization, it is sufficient to prove that $\bar{q}_2$ becomes $\frac{1}{2}$, because that will mean that the range $(\frac{1}{2},\bar{q}_2)$ collapses to nothing. The value of $\bar{q}_2$ is implicitly determined by the equation $q = \bar{q}_2(q)$, where

$$\bar{q}_2(q) = \frac{1+d_L+n+d_R z(q,\gamma)}{2+2n+(d_L+d_R)(1+z(q,\gamma))}$$

Note that

$$\bar{q}_2\left(\frac{1}{2}\right) = \frac{1}{2} + \frac{(d_L-d_R)(1-\gamma)}{2(2-\gamma)(1+n)+2(d_L+d_R)},$$

which collapses to $\bar{q}_2\left(\frac{1}{2}\right) = \frac{1}{2}$ iff either $d_L = d_R$ or $\gamma = 1$. Thus, if either of these conditions holds, the fixed point of the equation $q = \bar{q}_2(q)$ is achieved at $q = \frac{1}{2}$, which means that $\bar{q}_2 = \frac{1}{2}$.

Now consider Proposition 1. Assume that i initially holds a correct prior: $\mathbb{P}(\omega = L) = \mu_i^0(\omega = L) = \pi$.

If $\gamma = 1$, then the situation is equivalent to agent i having $n+d_L+d_R$ normal neighbors, since she is able to completely unravel selective sharing. In that case, i's expected posterior belief $\mathbb{E}[\mu_i(\omega = L)]$ must be equal to her prior π due to a standard argument about a fully Bayesian agent who is correctly specified with respect to her information structure.

Now suppose $\gamma < 1$ and $d_L = d_R$. Using the conditional expected posterior beliefs from the proof of Proposition 1, we can show that i's unconditional expected posterior belief from ex-ante perspective (before ω is determined) is given by

$$\begin{aligned}\mathbb{E}\left[\mu_i(\omega=L)\right] &= (1-\pi)\mathbb{E}\left[\mu_i^1(\omega=L)\mid\omega=R\right]+\pi\mathbb{E}[\mu_i^1(\omega=R)\mid\omega=L]\\ &= (1-\pi)\Bigg[\sum_{k=0}^{\frac{d_L+d_R}{2}-1}\frac{(d_L+d_R)!}{k!(d_L+d_R-k)!}\Big(f(k,q)+f(d_L+d_R-k,q)\Big)+\\ &\qquad+\frac{(d_L+d_R)!}{\left(\frac{d_L+d_R}{2}\right)!\left(\frac{d_L+d_R}{2}\right)!}f\left(\frac{d_L+d_R}{2},q\right)\Bigg]+\\ &\quad+\pi\Bigg[\sum_{k=0}^{\frac{d_L+d_R}{2}-1}\frac{(d_L+d_R)!}{k!(d_L+d_R-k)!}\Big(\hat{f}(k,q)+\hat{f}(d_L+d_R-k,q)\Big)+\\ &\qquad+\frac{(d_L+d_R)!}{\left(\frac{d_L+d_R}{2}\right)!\left(\frac{d_L+d_R}{2}\right)!}\hat{f}\left(\frac{d_L+d_R}{2},q\right)\Bigg],\end{aligned}$$

where

$$f(k,q)=\frac{\pi(1-q)^k q^{d_L+d_R-k}}{\pi+(1-\pi)\left(\frac{1-q}{q}\right)^{k-d_R}\left(\frac{\gamma(1-q)+(1-\gamma)}{\gamma q+(1-\gamma)}\right)^{k-d_L}}$$

and

$$\hat{f}(k,q)=\frac{\pi q^k(1-q)^{d_L+d_R-k}}{\pi+(1-\pi)\left(\frac{1-q}{q}\right)^{k-d_R}\left(\frac{\gamma(1-q)+(1-\gamma)}{\gamma q+(1-\gamma)}\right)^{k-d_L}}.$$

Note that the expected belief in $\omega=L$ is completely symmetric to the expected belief in $\omega=R$ with respect to numbers of friends and priors. That is, if we were to let $\hat{\pi}=1-\pi$, $\hat{d}_L=d_{Ri}$ and $\hat{d}_R=d_{Li}$, we would get the same expression for $\mathbb{E}[\mu_i^1(\omega=R)]$ as for

$\mathbb{E}[\mu_i^1(\omega = L)]$ above:

$$\mathbb{E}\left[\mu_i(\omega = R)\right] = (1-\hat{\pi})\mathbb{E}\left[\mu_i(\omega = R) \mid \omega = L\right] + \hat{\pi}\mathbb{E}[\mu_i(\omega = L) \mid \omega = R]$$

$$= (1-\hat{\pi})\left[\sum_{k=0}^{\frac{\hat{d}_R+\hat{d}_L}{2}-1} \frac{(\hat{d}_R+\hat{d}_L)!}{k!(\hat{d}_R+\hat{d}_L-k)!}\Big(f(k,q)+f(\hat{d}_R+\hat{d}_L-k,q)\Big)+\right.$$

$$\left. + \frac{(\hat{d}_R+\hat{d}_L)!}{\left(\frac{\hat{d}_R+\hat{d}_L}{2}\right)!\left(\frac{\hat{d}_R+\hat{d}_L}{2}\right)!} f\left(\frac{\hat{d}_R+\hat{d}_L}{2}, q\right)\right]+$$

$$+\hat{\pi}\left[\sum_{k=0}^{\frac{\hat{d}_R+\hat{d}_L}{2}-1} \frac{(\hat{d}_R+\hat{d}_L)!}{k!(\hat{d}_R+\hat{d}_L-k)!}\Big(\hat{f}(k,q)+\hat{f}(\hat{d}_R+\hat{d}_L-k,q)\Big)+\right.$$

$$\left. + \frac{(\hat{d}_R+\hat{d}_L)!}{\left(\frac{\hat{d}_R+\hat{d}_L}{2}\right)!\left(\frac{\hat{d}_R+\hat{d}_L}{2}\right)!} \hat{f}\left(\frac{\hat{d}_R+\hat{d}_L}{2}, q\right)\right],$$

where

$$f(k,q) = \frac{\hat{\pi}(1-q)^k q^{\hat{d}_R+\hat{d}_L-k}}{\hat{\pi}+(1-\hat{\pi})\left(\frac{1-q}{q}\right)^{k-\hat{d}_L}\left(\frac{\gamma(1-q)+(1-\gamma)}{\gamma q+(1-\gamma)}\right)^{k-\hat{d}_R}}$$

and

$$\hat{f}(k,q) = \frac{\hat{\pi}q^k(1-q)^{\hat{d}_R+\hat{d}_L-k}}{\hat{\pi}+(1-\hat{\pi})\left(\frac{1-q}{q}\right)^{k-\hat{d}_L}\left(\frac{\gamma(1-q)+(1-\gamma)}{\gamma q+(1-\gamma)}\right)^{k-\hat{d}_R}}.$$

This symmetry implies the following conclusion. Suppose that for some d_L and d_R the expected belief $\mathbb{E}[\mu_i(\omega = L)]$ is biased in some direction; for the sake of argument, assume $\mathbb{E}[\mu_i(\omega = L)] > \pi$. Because of the symmetry above, it then follows that for $\hat{d}_L = d_R$ and $\hat{d}_R = d_L$ (swapping the number of friends), we will have $\mathbb{E}[\mu_i(\omega = R)] > \hat{\pi} = 1-\pi$. In other words, if dogmatic-left friends are able to bias i's expected belief towards state $\omega = L$, dogmatic-right friends would be able to do the same with state $\omega = R$ under symmetric circumstances.

This implies that, when $d_L = d_R$ holds, if $\mathbb{E}[\mu_i(\omega = L)] > \pi$, then $\mathbb{E}[\mu_i(\omega = R)] > 1-\pi$ should also hold. However, this means that the expected posterior probabilities do not sum up to 1, which is a contradiction. Therefore, when $d_L = d_R$ holds, it must be the case that $\mathbb{E}[\mu_i(\omega = L)] = \pi$ and $\mathbb{E}[\mu_i(\omega = R)] = 1-\pi$. This finishes the proof of the proposition. ■

D Proof of Proposition 4

Proof of Proposition 4. To prove the result, we need to consider the implicit equation that defines $\bar{q}_2(d_L, d_R, n, \gamma)$:

$$q = \bar{q}_2(q) \Leftrightarrow q = \frac{1 + d_L + n + d_R z(q,\gamma)}{2 + 2n + (d_L + d_R)(1 + z(q,\gamma))} = \frac{1}{2} + \frac{d_L\left(1 - \frac{1+z(q,\gamma)}{2}\right) + d_R\left(z(q,\gamma) - \frac{1+z(q,\gamma)}{2}\right)}{2 + 2n + (d_L + d_R)(1 + z(q,\gamma))}$$

Consider the right-hand side of this equation when d_L, d_R and n_i increase by a factor of $\lambda > 1$:

$$\begin{aligned}
\frac{1}{2} + \frac{\lambda d_L\left(1 - \frac{1+z(q,\gamma)}{2}\right) + \lambda d_R\left(z(q,\gamma) - \frac{1+z(q,\gamma)}{2}\right)}{2 + \lambda\left(2n + (d_L + d_R)(1 + z(q,\gamma))\right)} &> \frac{1}{2} + \frac{\lambda d_L\left(1 - \frac{1+z(q,\gamma)}{2}\right) + \lambda d_R\left(z(q,\gamma) - \frac{1+z(q,\gamma)}{2}\right)}{2\lambda + \lambda\left(2n + (d_L + d_R)(1 + z(q,\gamma))\right)} \\
&= \frac{1}{2} + \frac{d_L\left(1 - \frac{1+z(q,\gamma)}{2}\right) + d_R\left(z(q,\gamma) - \frac{1+z(q,\gamma)}{2}\right)}{2 + 2n + (d_L + d_R)(1 + z(q,\gamma))} \\
&= \frac{1 + d_L + n + d_R z(q,\gamma)}{2 + 2n + (d_L + d_R)(1 + z(q,\gamma))} \\
&= \bar{q}_2(q)
\end{aligned}$$

Hence, the right-hand side of the equation $q = \bar{q}_2(q)$ increases when d_L, d_R and n all increase by a factor of $\lambda > 1$, which means that we need a higher q on the left-hand side to satisfy the equation. Therefore, if $\bar{q}_2(d_L, d_R, n, \gamma)$ solves $q = \bar{q}_2(q)$, then

$$\bar{q}_2(\lambda d_L, \lambda d_R, \lambda n, \gamma) > \bar{q}_2(d_L, d_R, n, \gamma).$$

Consider arbitrary $\lambda_L > 1$, $\lambda_R > 1$ and $\lambda_N > 1$ and look at the expression for $\bar{q}_2(\lambda d_L, \lambda d_R, \lambda_N n, \gamma)$:

$$\bar{q}_2(\lambda_L d_L, \lambda_R d_R, \lambda_N n, \gamma) = \frac{1}{2} + \frac{\lambda_L d_L\left(1 - \frac{1+z(q,\gamma)}{2}\right) + \lambda_R d_R\left(z(q,\gamma) - \frac{1+z(q,\gamma)}{2}\right)}{2 + 2\lambda_N n + (\lambda_L d_L + \lambda_R d_R)(1 + z(q,\gamma))}$$

The inequality $\bar{q}_2(d_L, d_R, n, \gamma) \leq \bar{q}_2(\lambda_L d_L, \lambda_R d_R, \lambda_N n, \gamma)$ will hold if and only if the following holds:

$$\begin{aligned}
\frac{(\lambda_L d_L - \lambda_R d_R)\left(1 - \frac{1+z(q,\gamma)}{2}\right)}{2+2\lambda_N n + (\lambda_L d_L + \lambda_R d_R)(1+z(q,\gamma))} &\geq \frac{(d_L - d_R)\left(1 - \frac{1+z(q,\gamma)}{2}\right)}{2+2n+(d_L+d_R)(1+z(q,\gamma))} \\
\frac{(\lambda_L d_L - \lambda_R d_R)}{2+2\lambda_N n + (\lambda_L d_L + \lambda_R d_R)(1+z(q,\gamma))} &\geq \frac{(d_L - d_R)}{2+2n+(d_L+d_R)(1+z(q,\gamma))} \\
(\lambda_L d_L - \lambda_R d_R)(2+2n+(d_L+d_R)(1+z(q,\gamma))) &\geq (d_L - d_R) \times \\
&\quad\; (2+2\lambda_N n + (\lambda_L d_L + \lambda_R d_R)(1+z(q,\gamma))) \\
(\lambda_L d_L - \lambda_R d_R)(2+2n) &\geq (d_L - d_R)(2+2\lambda_N n) \\
(\lambda_L d_L - \lambda_R d_R)(1+n_i) &\geq (d_L - d_R)(1+\lambda_N n) \\
((\lambda_L - 1)d_L - (\lambda_R - 1)d_R) + n(\lambda_L d_L - \lambda_R d_R) &\geq (d_L - d_R)\lambda_N n \\
\frac{((\lambda_L - 1)d_L - (\lambda_R - 1)d_R) + n(\lambda_L d_{Li} - \lambda_R d_R)}{(d_L - d_R)n} &\geq \lambda_N \\
\frac{((\lambda_L - 1)d_L - (\lambda_R - 1)d_R) + n((\lambda_L - 1)d_L - (\lambda_R - 1)d_R)}{(d_L - d_R)n} &\geq \lambda_N - 1 \\
\frac{((\lambda_L - 1)d_L - (\lambda_R - 1)d_R)(1+n)}{(d_L - d_R)n} &\geq \lambda_N - 1 \\
\frac{(\lambda_L - 1)d_L - (\lambda_R - 1)d_R}{(d_L - d_R)} \cdot \left(1 + \frac{1}{n}\right) &\geq \lambda_N - 1
\end{aligned}$$

■

E Proof of Proposition 5

Proof of Proposition 5. Recall that $q = \bar{q}_2(d_{Li}, d_{Ri}, n_i, \gamma)$ is defined by

$$q = \frac{1 + d_{Li} + n_i + d_{Ri} z(q,\gamma)}{2 + 2n + (d_{Li} + d_{Ri})(1 + z(q,\gamma))}$$

Fix $d_{Li}, d_{Ri}, n_i, \lambda$ and $\hat{q}$. We need to find λ_N such that

$$\hat{q} \geq \frac{1 + \lambda d_{Li} + \lambda_N n_i + \lambda d_{Ri} z(q,\gamma)}{2 + 2\lambda_N n_i + \lambda(d_{Li} + d_{Ri})(1 + z(q,\gamma))} = \frac{1}{2} + \frac{\lambda(d_{Li} - d_{Ri})\frac{1-z(q,\gamma)}{2}}{2 + 2\lambda_N n_i + \lambda(d_{Li} + d_{Ri})(1 + z(q,\gamma))}$$

Note that the right-hand side is decreasing in $z(q,\gamma)$. So a sufficient condition would be to impose the inequality for the lowest value of $z(q,\gamma)$, which is 0. This gives us the following inequality:

$$
\begin{aligned}
\hat{q} &\geq \frac{1}{2} + \frac{\frac{1}{2}\lambda(d_{Li} - d_{Ri})}{2 + 2\lambda_N n_i + \lambda(d_{Li} + d_{Ri})} \\
2\hat{q} - 1 &\geq \frac{\lambda(d_{Li} - d_{Ri})}{2 + 2\lambda_N n_i + \lambda(d_{Li} + d_{Ri})} \\
\lambda(d_{Li} - d_{Ri}) &\leq (2\hat{q} - 1)(2 + 2\lambda_N n_i + \lambda(d_{Li} + d_{Ri})) \\
\lambda(d_{Li} - d_{Ri}) &\leq (2\hat{q} - 1)(2 + \lambda(d_{Li} + d_{Ri})) + 2(2\hat{q} - 1)\lambda_N n_i \\
\lambda_N &\geq \frac{2(1 - \hat{q})\lambda d_{Li} - 2(2\hat{q} - 1) - 2\hat{q}\lambda d_{Ri}}{2(2\hat{q} - 1)n_i} \\
\lambda_N &\geq \frac{(1 - \hat{q})\lambda d_{Li} - \hat{q}\lambda d_{Ri} - (2\hat{q} - 1)}{(2\hat{q} - 1)n_i} \\
\lambda_N &\geq \frac{(1 - \hat{q})d_{Li} - \hat{q}d_{Ri}}{(2\hat{q} - 1)n_i}\lambda - \frac{1}{n_i} \\
\lambda_N &\geq \frac{d_{Li} - \hat{q}(d_{Li} + d_{Ri})}{(2\hat{q} - 1)n_i}\lambda - \frac{1}{n_i}
\end{aligned}
$$

■

F Proof of Proposition 6

Proof. First, we will quickly prove Lemma 1. Consider agent i who is "eventually correct" at a given $q = \hat{q}$. There are two possibilities: either $\bar{q}_{2i} > \hat{q}$ or $\bar{q}_{2i} < \hat{q}$ (we omit the knife-edge case). If $\bar{q}_{2i} > \hat{q}$, then the agent must have dogmatic imbalance towards the correct state. Increasing q beyond $\bar{q}_{2i}$ will lead to the agent learning correctly and having beliefs converge to the same state as before. Hence, the agent will not leave the set of "eventually correct" agents as q increases. If $\bar{q}_{2i} < \hat{q}$, then the agent is already learning correctly, and increasing q further will not alter her asymptotic beliefs. Thus, she will remain "eventually correct" as q goes up. All of this implies that the set of "eventually correct" agents is not contracting as q increases.

Now consider agent j who is "eventually incorrect" at a given $q = \hat{q}$. This can only occur if the agent is learning incorrectly, meaning she has a dogmatic imbalance towards the wrong state that is sufficiently large, i.e. $\bar{q}_{2i} > \hat{q}$. Increasing q beyond $\bar{q}_{2i}$ will make the agent begin to learn correctly, which means she will leave the "eventually incorrect" set of agents and will now be in the "eventually correct" set. This finishes the proof of Lemma 1.

Fix some $q = \hat{q}$ and considers sets $\mathcal{N}_L(q)$ and $\mathcal{N}_R(q)$. For definiteness, let $\omega = L$ be the true state. Let $\bar{q}_{min}(\hat{q}) = \min_{i \in \mathcal{N}}\{\bar{q}_{2i} \mid \bar{q}_{2i} > \hat{q}\}$ be the lowest $\bar{q}_{2i}$ among agents who are "eventually incorrect" at $q = \hat{q}$. As q increases and reaches $\bar{q}_{min}(\hat{q})$, that agent will flip from being "eventually incorrect" to being "eventually correct". Since $\omega = L$ is the true state, this implies

$$|\mathcal{N}_L(\bar{q}_{min}(\hat{q}))| = |\mathcal{N}_L(\hat{q})| + 1 \quad \text{and} \quad |\mathcal{N}_R(\bar{q}_{min}(\hat{q}))| = |\mathcal{N}_R(\hat{q})| - 1.$$

Consider the network polarization $\Pi(q)$ at $q = \hat{q}$ and $q = \bar{q}_{min}(\hat{q})$:

$$\Pi(\hat{q}) = \frac{4}{|\mathcal{N}|} \cdot |\mathcal{N}_L(\hat{q})||\mathcal{N}_R(\hat{q})|$$

$$\Pi(\bar{q}_{min}(\hat{q})) = \frac{4}{|\mathcal{N}|} \cdot (|\mathcal{N}_L(\hat{q})| + 1)(\mathcal{N}_R(\hat{q}) - 1)$$

Note that $\Pi(\hat{q}) \geq \Pi(\bar{q}_{min}(\hat{q}))$ if and only if

$$|\mathcal{N}_L(\hat{q})||\mathcal{N}_R(\hat{q})| \geq (|\mathcal{N}_L(\hat{q})| + 1)(\mathcal{N}_R(\hat{q}) - 1),$$

which is equivalent to

$$|\mathcal{N}_R(\hat{q})| \leq |\mathcal{N}_L(\hat{q})| + 1.$$

Hence, the network polarization weakly decreases with q if and only if the set of "eventually incorrect" agents initially (at $q = \hat{q}$) is smaller than the set of "eventually correct" agents plus one. Since $\mathcal{N}_R(q)$ is weakly contracting in q, a necessary and sufficient condition for $\Pi(q)$ to be always weakly decreasing in q is that $|\mathcal{N}_R(\frac{1}{2})| = |\mathcal{D}_R|$ is weakly smaller than $|\mathcal{N}| - |\mathcal{D}_R| + 1$, which is equivalent to $|\mathcal{D}_R| \leq \frac{1}{2}(|\mathcal{N}| + 1)$.

Therefore, the network polarization $\Pi(q)$ is weakly decreasing in q over $q \in \left(\frac{1}{2}, 1\right)$ if and only if $|\mathcal{D}_{-\omega}| \leq \frac{1}{2}(|\mathcal{N}| + 1)$. ■

G Proof of Proposition 7

Proof. The proof of this proposition is short, and requires us to prove that by aggregating enough signals, we can push quality of information above any given threshold. In other words, if $\hat{s}^i_{Mt}$ denotes an M-aggregated signal, we need to prove

$$\lim_{M\to\infty} \mathbb{P}(\hat{s}^i_{Mt} = l|\omega = L) = \lim_{M\to\infty} \mathbb{P}(\hat{s}^i_{Mt} = r|\omega = R) = 1.$$

Suppose that we aggregate M (odd number) signals together, and offer agents the following information structure:

$$\hat{s}^i_{Mt} = \begin{cases} 0, & \text{if } \sum_{t=1}^{M} s_{it} < \frac{M}{2} \\ 1, & \text{if } \sum_{t=1}^{M} s_{it} > \frac{M}{2} \end{cases}$$

For simplicity, assume that the true state is $\omega = L$, and consider the quality of signal $\hat{s}^i_{Mt}$:

$$\begin{aligned}
\mathbb{P}\left(\hat{s}^i_{Mt} = l \middle| \omega = L\right) &= \mathbb{P}\left(\sum_{t=1}^{M} s_{it} < \frac{M}{2} \middle| \omega = L\right) \\
&= \mathbb{P}\left(\frac{1}{M}\sum_{t=1}^{M} s_{it} < \frac{1}{2} \middle| \omega = L\right) \\
&= \mathbb{P}\left(\frac{1}{M}\sum_{t=1}^{M} s_{it} < \frac{1}{2} \middle| \omega = L\right) \\
&= 1 - \mathbb{P}\left(\frac{1}{M}\sum_{t=1}^{M} s_{it} \geq \frac{1}{2} \middle| \omega = L\right) \\
&= 1 - \mathbb{P}\left(\frac{1}{M}\sum_{t=1}^{M} s_{it} - (1-q) \geq \frac{1}{2} - (1-q) \middle| \omega = L\right) \\
&= 1 - \mathbb{P}\left(\ell_M - (1-q) \geq \frac{1}{2} - (1-q) \middle| \omega = L\right) \\
&\text{as } \overset{M\to\infty}{\to} 1 - 0 = 1,
\end{aligned}$$

where the last line is due to the weak Law of Large Numbers on a Binomial random variable ℓ_M with parameters M and $(1-q)$. Hence, the limit of signal's quality as $M \to \infty$ is equal to 1, meaning that there must exist a finite number of signals such that the quality surpasses the bound $\bar{q}_1$. For such aggregated signal, information quality is too high for Proposition 2 to apply, which implies that there is no divergence of beliefs in the network. ■