working paper

16 04

Differences of Opinions

Dionissi Aliprantis

FEDERAL RESERVE BANK OF CLEVELAND

Working papers of the Federal Reserve Bank of Cleveland are preliminary materials circulated to stimulate discussion and critical comment on research in progress. They may not have been subject to the formal editorial review accorded official Federal Reserve Bank of Cleveland publications. The views stated herein are those of the authors and are not necessarily those of the Federal Reserve Bank of Cleveland or the Board of Governors of the Federal Reserve System.

Working papers are available on the Cleveland Fed's website:
https://clevelandfed.org/wp

Working Paper 16-04 January 2016

Differences of Opinions

Dionissi Aliprantis

This paper presents a generalization of the DeGroot learning rule in which social learning can lead to polarization, even for connected networks. I first develop a model of biased assimilation in which the utility an agent receives from past decisions depends on current beliefs when uncertainty is slow to resolve. I use this model to motivate key features of an agent's optimization problem subject to scarce private information, which forces the agent to extrapolate using social information. Even when the agent extrapolates under "scientific" assumptions and all individuals in the network process and report their private signals in an unbiased way, the possibility of biased processing or reporting leads agents to process social signals differently depending on the sender. The resulting solution to the agent's problem is a heterogeneous confidence learning rule that is distinct from bounded confidence learning rules in that the agent may actually move her beliefs away from, and not only discard, signals from untrustworthy senders.

Keywords: Biased Assimilation, Social Learning, Network, DeGroot Learning Rule, Bounded Confidence, Heterogeneous Confidence, Self-Affirmation.

JEL Classification: D83.

Suggested citation: Aliprantis, Dionissi, 2016. "Differences of Opinions," Federal Reserve Bank of Cleveland Working Paper, no. 16-04.

Dionissi Aliprantis is at the Federal Reserve Bank of Cleveland. Many of the ideas in this paper originated from discussions with Alon Bergman and Gregorio Caetano. Several colleagues at the Cleveland Fed have also made helpful comments.

1 Introduction

Beliefs are a critical feature of economic models of choice. Consider a proposition, or a statement that is true or false. How does a population's beliefs regarding the proposition evolve over time? How individuals learn from the experiences of others, or social learning, is likely to be a key mechanism determining the answer to this question.

Models of social learning typically include a learning rule and a network structure. Learning rules determine how individuals interpret the information contained in the choices of others, and network structures determine which choices an individual observes. A wide literature studies opinion dynamics for rule of thumb learning rules, as fully Bayesian learning can require the decision-maker to make unrealistically complex computations (Molavi et al. (2015)). A key result is that beliefs updated by the canonical DeGroot (1974) learning rule, which updates opinions based on weighted averages of neighbors' opinions, converge for connected network structures (Jackson (2008), Chapter 8).[1]

The empirical evidence, however, suggests that there can be persistent disagreement even on connected networks. Individuals are exposed to many sources of information contradicting their beliefs (Gentzkow and Shapiro (2011)), and yet we still observe persistent disagreement regarding propositions like Iraq had an active WMD program, President Obama was born in the US, or global warming is occurring. This evidence suggests that differences of opinions arise not only from access to different information, but also from differential processing of the same information.

This paper studies how a non-degenerate distribution of beliefs can be sustained over time, even on connected networks, when agents process the same information differently depending on the sender. As one motivation for the differential assimilation of social information, I first present a model that can explain the biased assimilation of private information. In the model, uncertainty is slow to resolve, and the expected utility an agent receives from previous decisions changes with her current beliefs about the state of the world due to pride. This backward-looking utility is consistent with the large literature in social psychology on self-affirmation, a model in which individuals have an objective of maintaining a positive view of themselves and their past decisions (Steele (1988), Sherman and Cohen (2006)). If the utility gained from evaluating past decisions under false beliefs that justify those decisions outweighs the utility gained from making future decisions under unbiased beliefs, the agent will process private signals in a biased manner.

I then consider an agent's optimization problem in which she must determine her beliefs about a set of propositions subject to scarce private information. Taken as exogenous are the agent's weight on her prior, the private signals she observes (including the weights she attaches to them), and the social signals to which she is exposed (ie, the network structure). The agent's endogenous decision is how to interpret and then weight the social signals of others in her network, knowing that some agents may process or report their information in a biased way.[2] The agent in the model extrapolates "scientifically," interpreting signals based on the sender's past performance on their

[1]Time to convergence, however, is not invariant to the network structure (Golub and Jackson (2012)).

[2]The agent is not allowed to coarsen the sets of propositions over which she determines her beliefs.

common observation set, and placing more weight on her interpretation of signals from sources for which she has more evidence.

The possibility of biased assimilation or strategic revelation is sufficient for agents to process social signals differently depending on the sender. The solution to the agent's problem is then a heterogeneous confidence learning rule that is distinct from bounded confidence learning rules in that the agent may actually move her beliefs away from, and not only discard, signals from untrustworthy senders. I present simulations showing that the learning rule solving the agent's problem can generate persistent disagreement and polarization for a connected network even when all individuals process their private signals in an unbiased way and truthfully reveal their beliefs. The level of distrust individuals have in the signals they receive from others in the network is shown to be a key parameter in determining whether a distribution of beliefs will ultimately reach a consensus.

Generalizing the DeGroot rule in a way that delivers persistent or even increasing disagreement over time has been the goal of a large literature (Acemoglu and Ozdaglar (2011)). Lorenz (2005) demonstrates that models of bounded confidence, in which individuals only weight signals close to their own, can generate a distribution of opinions for certain network structures.[3] Dandekar et al. (2013) show, however, that convergence still obtains on connected networks for many existing generalizations of DeGroot in the absence of biased assimilation. Biased assimilation can also be motivated by its widespread empirical documentation (See Sherman and Cohen (2006) or Kandel and Pearson (1995).), although it should be stressed that the results in this paper can be generated in the absence of biased assimilation through strategic revelation of information.

The dimensionality of the learning problem is central to the model in this paper, if slightly different in nature from the related issues studied in Al-Najjar (2009) and DeMarzo et al. (2003). Solving the analogue to the agent's problem might be thought of as an alternative when case-based or coarse decision making is not sufficient for the agent's purposes (Gilboa and Schmeidler (2001), Mohlin (2014), Al-Najjar and Pai (2014)). The models in this paper are also related to Fryer et al. (2013)'s theory of biased assimilation when an agent must (coarsely) categorize ambiguous signals, as well as Jadbabaie et al. (2012)'s generalization of DeGroot to allow for the arrival of new private information over time.

The paper is organized as follows: Section 2 presents an economic model of biased assimilation. This model is then used to motivate key features of an agent's maximization problem subject to scarce information developed in Section 3. Section 4 presents simulation results, and Section 5 concludes. The Appendices work out some simple special cases of the model and discuss their relation to previous literature.

[3] Hegselmann and Krause (2002) developed the original bounded confidence model; a recent version closer to the model in this paper is in Sotiropoulos et al. (2015).

2 A Model of Biased Assimilation

2.1 Backward-Looking Utility with Pride and Slow-to-Resolve Uncertainty

Suppose there are two states of the world, $s \in \{0, 1\}$, and a decision-maker has beliefs at time t, $\lambda_t \in \{0, 1\}$. Given a signal $\sigma \in \{0, 1\}$ that is informative about the true state of the world (in the sense that $Pr(s = 1|\sigma_t = 1) > 0.5$ and $Pr(s = 0|\sigma_t = 0) > 0.5$), a decision-maker forms beliefs as $\lambda_{ti} = f_i(\sigma_t)$. Why would two decision-makers i and j process the same signal differently and reach different beliefs $\lambda_{ti} \neq \lambda_{tj}$? In other words, what could justify decision-makers i and j choosing $f_i \neq f_j$?

The following model can generate such biased assimilation of new information. Assume that there is a decision between one of two choices, $d \in \{0, 1\}$, and suppose there are two time periods, $t \in \{1, 2\}$. The decision-maker forms beliefs in the first sub-period, and makes choices based on her beliefs using expected utility where utility is a function of decisions and states of the world, $u(d, s)$. Expected utility given beliefs is defined as

$$\begin{aligned} U(d, \lambda_1) &= \mathbb{E}_{\lambda_1}[u(d, s)] = \sum_{s \in \{0,1\}} \lambda_s u(d, s) \\ &= \lambda_1 u(d, 1) + (1 - \lambda_1) u(d, 0) \end{aligned}$$

In the first sub-period, the agent chooses beliefs λ_1, knowing that if uncertainty is resolved, the realized value of those beliefs in the second sub-period will be:

$$V(\lambda_1) = su(d_1^\star, 1) + (1 - s)u(d_1^\star, 0)$$

where

$$d_1^\star = \max_{d_1} U(d_1, \lambda_1) = \max_{d_1} \left[\lambda_1 u(d_1, 1) + (1 - \lambda_1) u(d_1, 0)\right].$$

The realized value of correct beliefs will be:

$$V(\lambda_1|\lambda_1 = s) = su(d_1^*, 1) + (1 - s)u(d_1^*, 0)$$

where

$$d_1^* = \max_{d_1} U(d_1, \lambda_1 = s) = \max_{d_1} \left[su(d_1, 1) + (1 - s)u(d_1, 0)\right].$$

Note that $V(\lambda_1|\lambda_1 = s) \geq V(\lambda_1)$, and so there is an incentive for beliefs to be correct. At the very least, λ_1 is constrained so that $d_1^\star = d_1^*$. Formally, $\lambda_1^* = s$ always solves the problem

$$\lambda_1^* = \max_{\lambda_1} \quad V(\lambda_1) \tag{1}$$

$$s.t. \;\; \lambda_1 \in \{0, 1\}. \tag{2}$$

Given an informative signal σ_1 about the true state s, the decision-maker will process the information in an unbiased way. Formally, if $f^+(\sigma_t) = \sigma_t$, $f^-(\sigma_t) = 1 - \sigma_t$, $f^T(\sigma_t) = 1$, and $f^F(\sigma_t) = 0$, then f^+ will solve

$$f_1^* = \max_{f_1 \in \{f^+, f^-, f^T, f^F\}} V(\lambda_1)$$
$$s.t.\ \lambda_1 = f_1(\sigma_1).$$

Moving on to period 2, suppose that the uncertainty about s has unexpectedly not resolved, and so the decision maker must make the same decision again. We assume that the decision maker will now get utility in period 2 from current beliefs based on her past decisions, as well as expected future utility from decisions based on those beliefs:

$$V(\lambda_2, d_1) = \alpha\Big([\lambda_2 u(d_1, 1) + (1 - \lambda_2)u(d_1, 0)] \Big) + \beta\, [s u(d_2^\star, 1) + (1 - s)u(d_2^\star, 0)]\,. \tag{3}$$

This backward-looking utility is consistent with the large literature in social psychology on self-affirmation, a model in which individuals have an objective of maintaining a positive view of themselves and their past actions (Steele (1988), Sherman and Cohen (2006)).

Note that in the first sub-period of period 2, it is now possible for $V(\lambda_2, d_1|\lambda_2 = s) < V(\lambda_2, d_1)$ or $V(\lambda_2, d_1|\lambda_2 = s) \geq V(\lambda_2, d_1)$, whereas in period 1 it was always the case that $V(\lambda_1|\lambda_1 = s) \geq V(\lambda_1)$. The key is that this inequality now depends on the decision-maker's past action d_1, her level of "pride" or "ideology" as expressed by α, and the potential loss in period 2 utility from choosing d_2 incorrectly. Thus, depending on the previous signals they received, and the resulting differences in their choice of d_1, two decision makers could have an incentive to process the information in σ_2 differently, or to choose different f_2^* to solve:

$$f_2^* = \max_{f_2 \in \{f^+, f^-, f^T, f^F\}} V(\lambda_2, d_1)$$
$$s.t.\ \lambda_2 = f_2(\sigma_2).$$

2.1.1 Example: Weapons of Mass Destruction and the Invasion of Iraq

Consider an example in which the states of the world are $s = 1$ if Saddam Hussein has/had an active program to produce Weapons of Mass Destruction (WMD) as of March 2003 - or $s = 0$ if he does/did not. The related decision would be $d = 1$ to invade/occupy Iraq - or $d = 0$ not to. Suppose that a decision was already made to invade Iraq in March of 2003, and so that considering the problem of signal processing at some later date can be interpreted as period 2 in the model.

Given the decision to invade, getting it right would have had a very high payoff $u(d_1 = 1, s = 1) = 1$, but getting it wrong would have had a very low payoff $u(d_1 = 1, s = 0) = -1$. Similarly, given the decision *not* to invade, getting it right would have had a very high payoff $u(d_1 = 0, s = 0) = 1$, but getting it wrong would also have had a very low payoff $u(d_1 = 0, s = 1) = -1$.

Suppose that the utility for period 2's decision to continue to occupy Iraq does not depend on the state s, or that $u(d_2, 1) \approx u(d_2, 0)$. Then recalling Equation 3, the processing of period 2's signal would depend heavily on the past decision to invade or not invade:

$$V(\lambda_2, d_1) = \alpha\Big([\lambda_2 u(d_1, 1) + (1 - \lambda_2)u(d_1, 0)] \Big) + \beta\, [su(d_2^\star, 1) + (1 - s)u(d_2^\star, 0)]$$
$$\approx \alpha\Big([\lambda_2 u(d_1, 1) + (1 - \lambda_2)u(d_1, 0)] \Big).$$

Those who had advocated for invasion would now choose $f_2^* = f^T$, and those who had advocated against the invasion would now choose $f_2^* = f^F$. What could align incentives for unbiased processing of informative signals ($f_2^* = f^+$), even given decision-makers with different past actions, would be if the decision-makers either were not too "proud" or "ideological" (ie, $\alpha \approx 0$), if the stakes from biased assimilation in terms of foregone utility were too high (ie, $u(d_2^*, s) >> u(d_2^\star, s)$), or some combination of the two.

This theory matches empirical observations well. There was contentious disagreement about whether Saddam Hussein had an active WMD program prior to the US-led invasion of Iraq in March 2003. Over a decade later, 42 percent of Americans believed that US forces found an active WMD program in Iraq after the invasion. Furthermore, such sustained "uncertainty" about the state of the world matters not only for current and future decisions, but also for our evaluation of past decisions. Most Americans' opinion about the decision to invade Iraq depends on whether it had an active WMD program.

One can imagine many additional scenarios in which the psychic costs of evaluating past actions in light of the true state are too high relative to the payoffs from biased assimilation.

3 A Model of Social Learning

3.1 The Agent's Problem and Its Analogue

Now suppose there is a network of $J + 1$ individuals. We focus on the way agent i processes information received from the J other individuals in the network. For a set of $\mathcal{K}$ statements, the agent considers the truth value of the simple proposition $p^k \in \{0, 1\}$ for $k = 1, \ldots, K = card(\mathcal{K})$.[4] At time t agent i has beliefs $\lambda_{it}^k = \Pr(p^k = 1)$, receives her own private signal $\sigma_{it}^k \in [0, 1]$ of relative strength $\theta_{it}^k \in [0, 1]$, and interprets the signal as

$$\widehat{\sigma}_{it}^k = f_i^k(\sigma_{it}^k) \in [0, 1].$$

Ideally, the agent would have experience with respect to every proposition $p^k \in \mathcal{K}$, so that all beliefs would be based on information processed according to her rule f_i^k. In this case, she would receive an ideal stream of information $\{\sigma_{it}^{k*}\}_{t=1}^T$ with informative signals $\theta_{it}^{k*} = 1$ for all t and for all $p^k \in \mathcal{K}$.

[4]Simple means that p^k cannot be generated as a compound proposition from p^m and p^n in $\mathcal{K}$.

If she gave δ_t weight to her prior beliefs, she would arrive at λ^{k*}_{iT+1} after updating according to

$$\begin{aligned}\lambda^{k*}_{it+1} &= (1-\delta_t)\lambda^k_{it} + \delta_t \theta^{k*}_{it}\widehat{\sigma}^{k*}_{it} \\ &= (1-\delta_t)\lambda^k_{it} + \delta_t \widehat{\sigma}^{k*}_{it}. \end{aligned} \tag{4}$$

While general δ_t might be of interest in non-stationary environments, by setting $\delta_t = 1/t$, λ^{k*}_{it} is just the average of agent i's signals. Assuming the σ^{k*}_{it} are iid draws, and f^k_i is continuous, the uniform law of large numbers ensures that λ^{k*}_{it} converges in probability. We can think of λ^{k*}_{iT+1} as the agent's truth, and note that a distribution of f^k_i could be generated by a model of biased assimilation like the one presented in Section 2, strategic revelation of information, or simply due to adopting different axioms.[5]

The agent's objective is to approximate λ^{k*}_{iT+1} subject to the constraint that she does not have access to an ideal stream of information. That is, the sequence of private signals $\{\sigma^k_{it}\}^T_{t=1}$ observed by the agent has exogenous quality $\{\theta^k_{it}\}^T_{t=1}$, with $\theta^k_{it} < 1$ for at least one $t \in \{1, \ldots, T\}$. It is possible that no private signals will be observed for some propositions, making $\theta^k_{it} = 0$ for some k.

Relevant for her objective, the agent also observes the set of signals reported by others in some set $\mathcal{J}^k$ with cardinality J^k, where

$$\overline{\sigma}^k_{jt} = f^k_j(\sigma^k_{jt}) \in [0,1] \quad \text{for} \quad j \in \mathcal{J}^k \quad \text{and} \quad j \neq i.$$

Note that whereas f^k_i is just about processing, f^k_j is about both processing and truthful revelation. Throughout the analysis $\widehat{\sigma}$ will denote a signal interpreted by agent i, $\overline{\sigma}$ will denote a signal interpreted and reported (possibly strategically or mistakenly) by individual j, and σ_i or σ_j will denote the true signals. We restrict belief processing $f^k : [0,1] \to [0,1]$ to be in

$$\left\{ \; f^k(\sigma^k) = m^k\sigma^k + b^k \;\middle|\; (m^k, b^k) \in [0,1]^2 \;\text{ and }\; m^k = (1-b^k) \;\text{ if }\; b^k \in (0,1] \; \right\}$$

The agent engages in social learning by choosing how to interpret and weight others' reported signals. The interpretation agent i gives to j's reported signal is denoted by

$$\widehat{\sigma}^k_{jt} = g_{jt}(\overline{\sigma}^k_{jt}, \;\cdot\;)$$

and the weight she places on this information is denoted by $\theta^k_{jt} \in \mathbb{R}$. Note here that the domain of g_{jt} is also considered to be a choice when specifying the function. The agent updates her prior as

$$\lambda^k_{it+1} = (1-\delta_t)\lambda^k_{it} + \delta_t \left[\theta^k_{it}\widehat{\sigma}^k_{it} + \sum_{j\in\mathcal{J}^k} \theta^k_{jt}\widehat{\sigma}^k_{jt} \right]. \tag{5}$$

[5]For example, f^k_i might capture whether the agent accepts or rejects Euclid's Fifth Postulate, and therefore whether she interprets information in terms of Euclidean or non-Euclidean geometry, and the associated representations of physical space.

Suppose that the weight the agent places on her prior is exogenous, as is the stream of private and social signals. However, only the weight she attaches to the private signals is exogenous, and her problem is to determine how to interpret and weight the social signals she receives. Given the loss function $\mathcal{L}$, the agent chooses the functions g_{jt} and the endogenous weights θ_{jt}^k to solve the problem

$$
\begin{aligned}
\min_{\{g_{jt}\}_{t=1,j=1}^{T,J}\ ,\ \{\theta_{jt}^1,\ldots,\theta_{jt}^K\}_{t=1,j\in\mathcal{J}^k}^T} \quad & \sum_{k=1}^{K} \mathcal{L}\Big(\lambda_{iT+1}^{k*} - \lambda_{iT+1}^{k}\Big) \\
s.t.\ \ & \{f_i^k\}_{k=1}^K \\
& \{\theta_{it}^k, \sigma_{it}^k\}_{t=1}^T \qquad \text{for}\ \ k=1,\ldots,K \\
& \{\lambda_{jt}^k, \overline{\sigma}_{jt}^k\}_{t=1}^T \qquad \text{for}\ \ k=1,\ldots,K\ \ \text{and}\ \ j\in\mathcal{J}^k \\
& \{\delta_t\}_{t=1}^T = \tfrac{1}{t} \qquad \text{for}\ \ t=1,\ldots,T \\
& \theta_{it}^k + \sum_{j\in\mathcal{J}^k} \theta_{jt}^k = 1 \qquad \text{for}\ \ k=1,\ldots,K\ \ \text{and}\ \ t=1,\ldots,T.
\end{aligned}
$$

Alternatively, Equations 4 and 5 can be combined to state the agent's problem recursively:

$$
\begin{aligned}
\min_{\{g_j\}_{j=1}^J\ ,\ \{\theta_j^1,\ldots,\theta_j^K\}_{j\in\mathcal{J}^k}} \quad & \sum_{k=1}^{K} \mathcal{L}\left(\widehat{\sigma}_i^{k*} - \left[\theta_i^k\widehat{\sigma}_i^k + \sum_{j\in\mathcal{J}^k}\theta_j^k g_j(\cdot)\right]\right) \\
s.t.\ \ & \{f_i^k\}_{k=1}^K \\
& X \\
& \theta_i^k + \sum_{j\in\mathcal{J}^k}\theta_j^k = 1 \text{ for }\ k=1,\ldots,K
\end{aligned}
\tag{6}
$$

where the private and social information sets available to the agent at a given point in time are defined as

$$
\begin{aligned}
X_i &\equiv \Big\{ \left(\lambda_i^1,\ldots,\lambda_i^K\right)\ ,\ \left(\theta_i^1,\ldots,\theta_i^K\right)\ ,\ \left(\sigma_i^1,\ldots,\sigma_i^K\right) \Big\}, \\
X_j &\equiv \Big\{\ \{\lambda_j^1,\overline{\sigma}_j^1\}_{j\in\mathcal{J}^1}\ ,\ \ldots\ ,\ \{\lambda_j^K,\overline{\sigma}_j^K\}_{j\in\mathcal{J}^K}\ \Big\},
\end{aligned}
$$

and $X = \{X_i, X_j\}$. Because the ideal information stream $\{\theta_{it}^{k*}, \sigma_{it}^{k*}\}_{t=1}^T$ is not observed by the agent, she does not know $\widehat{\sigma}_{it}^{k*}$, and so (6) does not represent a well-posed problem. For a given model m generating predictions s_{it}^{k*} of $\widehat{\sigma}_{it}^{k*}$ based on X, the following analogue to the agent's problem is well-posed:

$$\begin{aligned} \min_{\{g_j\}_{j=1}^{J}, \{\theta_j^1,\ldots,\theta_j^K\}_{j\in\mathcal{J}^k}} \quad & \sum_{k=1}^{K} \mathcal{L}\left(s_i^{k*} - \left[\theta_i^k \widehat{\sigma}_i^k + \sum_{j\in\mathcal{J}^k} \theta_j^k g_j(\cdot)\right]\right) \\ s.t. \quad & \{f_i^k\}_{k=1}^{K} \\ & X \\ & \theta_i^k + \sum_{j\in\mathcal{J}^k} \theta_j^k = 1 \text{ for } \; k = 1,\ldots,K \\ & s_i^{k*} = m(X). \end{aligned} \tag{7}$$

If we restrict the model to the functional form

$$s_i^{k*} = m(X) = \theta_{it}^k \, \widehat{\sigma}_i^k + \sum_{j\in\mathcal{J}^k} w_j^k(X) s_j^k(X),$$

then the function $s_j^k(X)$ can be interpreted as the agent's prediction of $\widehat{\sigma}_i^{k*}$ based on source j's reported signal and X, and the weighting function $w_j^k(X)$ can be considered the relative confidence in that prediction.

Assuming that the agent must make a decision based upon a choice of λ_{T+1}^{k*}, her options are to set λ_{T+1}^{k*} either arbitrarily, using only her private information, using information other than the social signals she has received, or by solving her problem's analogue for a given model (7).[6] λ_{T+1}^{k*} might be chosen to minimize an expected or maximum loss function over a set of models (Manski (2011)), but here we focus on the first step of solving the analogue problem for one model. It is worth noting that combining a model with observations to construct unobserved quantities is a method used in causal inference to overcome the fundamental problem of evaluation (Holland (1986)), and is at the heart of the problem of induction (Aliprantis (2015)).

3.2 Using the Agent's Model to Interpret Signals

Since she does not observe f_j^k, the agent must interpret signals by extrapolating from the signals that are both privately and socially observed. For this reason, the interpretation functions g_j have a domain that is not simply the domain of the social signals $\overline{\sigma}_j$ (ie, [0,1]), but rather also include the information inferred to be in the sender's signal, as well as the agent's current signal and prior beliefs. Let $\mathcal{K}_i \subseteq \mathcal{K}$ be the set of propositions with private signals at the given time period (ie, $\theta_i^k \in (0,1]$), so that the agent does not receive private signals about propositions $p^m \in \mathcal{K} \setminus \mathcal{K}_i$. Define $\mathcal{K}_j \subseteq \mathcal{K}$ to be the set of propositions for which the agent has observed j's processed signal

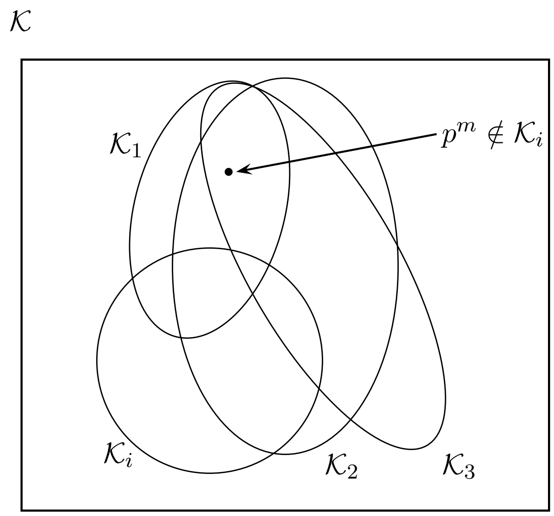
[6] If the agent could make her decision based upon her choice of λ_{T+1}^{m*} for some $p^m \neq p^k$ for which she had better information, she could restrict her learning to a set of propositions not including p^k and therefore coarser than $\mathcal{K}$ (See Al-Najjar (2009), Mohlin (2014), or Al-Najjar and Pai (2014) for related discussions.). The assumption here is that the agent must make a decision based on λ_{T+1}^{k*}, and she must do so with the information she has, and not the information she might want or wish she had.

and prior (assuming priors and signals are always observed together), and $\mathcal{K}_{ij} = \mathcal{K}_i \bigcap \mathcal{K}_j$.

Define the average disagreement over $\mathcal{K}_{ij}$, the shared set of propositions between the agent and sender j, as

$$\overline{\triangle}_{ij} = \frac{\sum_{k \in \mathcal{K}_{ij}} \ \overline{\theta}_i^k \ \left|\lambda_i^k - \lambda_j^k\right|}{|\mathcal{K}_{ij}|}, \tag{8}$$

where $\overline{\theta}_i^k$ is the average signal quality the agent has received about proposition p^k up to period t. Note that the information about $p^m \in \mathcal{K}_j$ is defined entirely over the space of propositions in $\mathcal{K}_{ij}$ (See Figure 1 below.).



Agent i's Privately Observed Propositions and Social Signals from Senders $j \in \{1, 2, 3\}$ where $p^m \notin \mathcal{K}_i$ and $|\mathcal{K}_{i2}| > |\mathcal{K}_{i1}| > |\mathcal{K}_{i3}|$

Figure 1: The Set of Propositions $\mathcal{K}$

The inductive assumption made by the agent is that $\overline{\triangle}_{ij}$ is informative for $\left|\widehat{\sigma}_i^{k*} - \overline{\sigma}_j^k\right|$. Specifically, the agent extrapolates by assessing the credibility of sender j's signal, using the observed disagreement over their common information under the following assumptions:

Interpretation 1: $\left|\widehat{\sigma}_i^{k*} - \overline{\sigma}_j^k\right|$ is increasing in $\overline{\triangle}_{ij}$

Interpretation 2: $\overline{\triangle}_{ij} = 0 \quad \Longrightarrow \quad \left|\widehat{\sigma}_i^{k*} - \overline{\sigma}_j^k\right| = 0$

Interpretation 3: $\overline{\triangle}_{ij} = 1 \quad \Longrightarrow \quad \begin{cases} \widehat{\sigma}_i^{k*} < \lambda_i^k & \text{if } \ \overline{\sigma}_j^k \geq \lambda_i^k; \\ \widehat{\sigma}_i^{k*} > \lambda_i^k & \text{if } \ \overline{\sigma}_j^k < \lambda_i^k. \end{cases}$

The agent's model predicts the value of σ_i^{k*} based on source j's signal and the other information available to the agent using the function:

$$m_j : [0,1]^3 \longrightarrow [0,1]$$

defined by

$$s_j^k = s_j(\overline{\sigma}_j^k, \overline{\triangle}_{ij}, \lambda_i^k).$$

To summarize **Interpretation 2**, if the agent is in perfect agreement with the sender over their common information set, then the agent will interpret their signal as being reliable irregardless of her beliefs:

$$s_j^k = m_j(\overline{\sigma}_j^k, 0, \lambda_i^k) = \overline{\sigma}_j^k.$$

Interpretation 3 pertains to the other extreme in which the agent is in perfect disagreement with the sender. In this case, she will actually move away from the sender's signal:

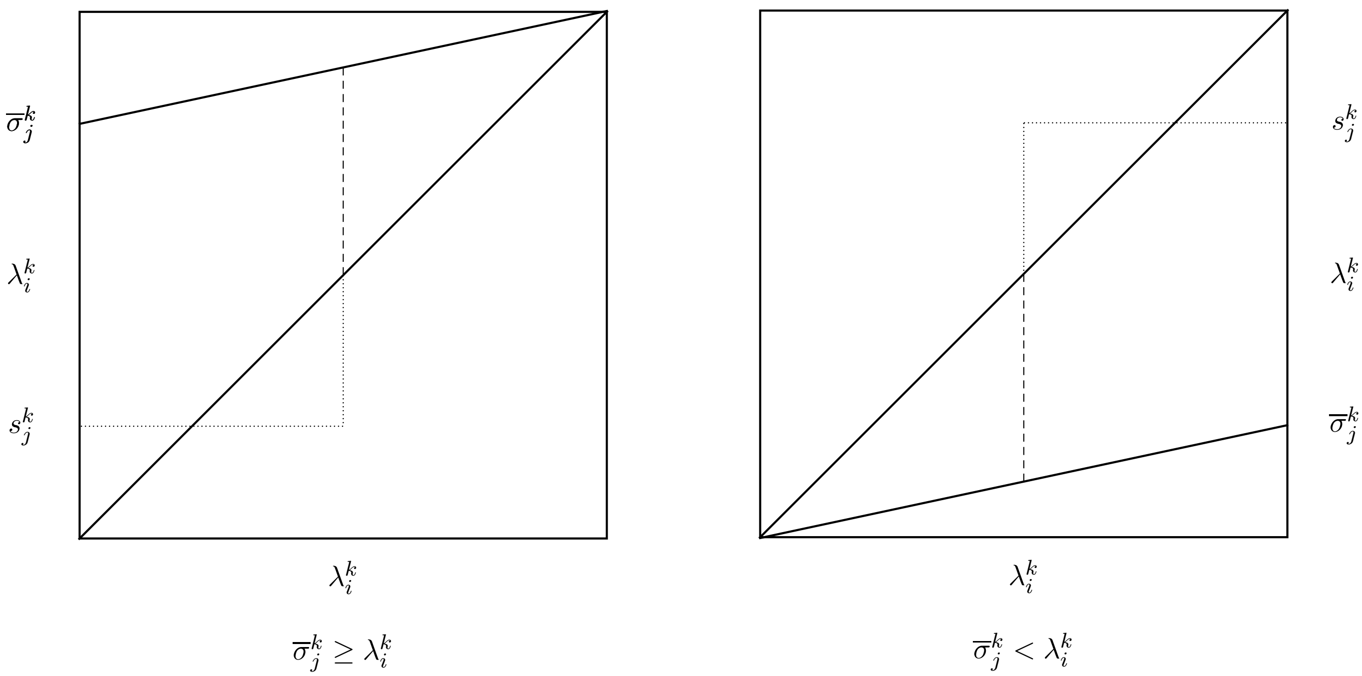


Figure 2: Interpretation under Perfect Disagreement

3.3 Using the Quantity of Empirical Evidence to Weight Interpreted Signals

Given a signal from sender j that has been interpreted, s_{jt}^k, how much weight does the agent attach to it? In other words, how confident is the agent that she has accurately interpreted the information in sender j's signal? In this context, an inductive assumption could be stated that as the evidence on a sender increases, so does the relative confidence in their signal:

Confidence 1:

$$w_j^k \text{ is increasing in } \frac{|\mathcal{K}_{ij}|}{\sum_{q\in\mathcal{J}^k}|\mathcal{K}_{iq}|}$$

3.4 Solving the Analogue to the Agent's Problem

Given the agent's model defined by $w_j^k(X)$ and $s_j^k(X)$, the solution to the analogue to the agent's problem:

$$\begin{aligned}
\min_{\{g_j\}_{j=1}^J, \{\theta_j^1,\ldots,\theta_j^K\}_{j\in\mathcal{J}^k}} \quad & \sum_{k=1}^K \mathcal{L}\left(s_i^{k*} - \left[\theta_i^k \hat{\sigma}_i^k + \sum_{j\in\mathcal{J}^k} \theta_j^k g_j(\cdot)\right]\right) \qquad (9)\\
s.t. \quad & \{f_i^k\}_{k=1}^K \\
& X \\
& \theta_i^k + \sum_{j\in\mathcal{J}^k} \theta_j^k = 1 \quad \text{for} \quad k = 1,\ldots,K \\
& s_i^{k*} = \theta_{it}^k \hat{\sigma}_i^k + \sum_{j\in\mathcal{J}^k} w_j^k(X) s_j^k(X)
\end{aligned}$$

is just

$$\begin{aligned}
\theta_j^k(X) &= w_j^k(X) \\
g_j^k(X) &= s_j^k(X).
\end{aligned}$$

4 Simulations

For the sake of a concrete example, I now show some simulation results for particular specifications of the agent's problem. I start by assuming that there is no strategic behavior on the part of agents; they all truthfully their beliefs. Furthermore, all agents process their own signal in an unbiased manner.

I use the continuous interpretation function satisfying **Interpretation 1-3** defined by

$$s_j^k = s_j(\overline{\sigma}_j^k, \overline{\triangle}_{ij}, \lambda_i^k) = \max\left\{ \min\left\{ \overline{\sigma}_j^k - 2\overline{\triangle}_{ij}^{\gamma}(\overline{\sigma}_j^k - \lambda_i^k) \ , \ 1 \right\} , \ 0 \right\}. \qquad (10)$$

Recall that average disagreement $\overline{\triangle}_{ij} \in [0,1]$. If disagreement is low ($\overline{\triangle}_{ij} \approx 0$), s_j^k pulls λ_i^k towards $\overline{\sigma}_j^k$. On the other hand, if disagreement is high ($\overline{\triangle}_{ij} \approx 1$), s_j^k pushes λ_i^k away from $\overline{\sigma}_j^k$. The strength of the attraction and repulsion depends on the distrust agent i has towards others, as set in the parameter $\gamma \in (0,\infty)$.

Finally, I use the weighting function satisfying **Confidence 1** defined by:

$$w_j^k(X) = \frac{|\mathcal{K}_{ij}|}{\sum_{q\in\mathcal{J}^k} |\mathcal{K}_{iq}|}(1-\theta_i^k). \qquad (11)$$

4.1 2-Agent Network with 1 Proposition

I begin with a very simple scenario: There are $N = 2$ agents in the network, and they must each determine the truth value of 1 proposition.

In the first column of Figure 3, I consider a scenario in which Agent 1 initially has beliefs tending toward false ($\lambda_{11} = 0.4$) and Agent 2 has initial beliefs tending toward true ($\lambda_{21} = 0.6$). Furthermore, in this first column, I assume that the agents have low levels of trust in the other agent's processing or truthful revelation of information ($\gamma = 0.1$), and that the sequence of private signals $\{\sigma_{ti}\}_{t=1}^{100}$ is constant for all time periods. Finally, I assume that the quality of private signals is also constant throughout all time periods, and that it is low ($\theta_{t1} = \theta_{t2} = 0.1$ for all t).

Figure 3a shows that in such a scenario, a constant stream of private signals equal to $\sigma_{ti} = 0.5$ results in polarization, or increasing disagreement over time. Whereas the agents begin at 0.4 and 0.6, within 10 periods they have moved apart, and by 100 periods they are approximately 0.3 and 0.7. Now if the constant stream of private signals is $\sigma_{ti} = 0.6$, the agents will move to similarly polarized beliefs (Figure 3c). The social learning mechanism still has an influence on beliefs, even when the private signals are constant at $\sigma_{ti} = 0.8$ and $\sigma_{ti} = 1.0$ (Figures 3e and 3g).

In the second column of Figure 3, I consider a scenario in which Agent 1 initially has strong beliefs that the proposition is false ($\lambda_{11} = 0.0$) and Agent 2 has strong initial beliefs that the proposition is true ($\lambda_{21} = 1.0$). I fix the sequence of private signals to be constant at 0.5, and vary the trust parameter γ. At low levels of trust ($\gamma = 0.2$), the agents barely move toward their private signals of 0.5. The remainder of the second column shows that as the level of trust increases, agents move closer and closer to their sequence of private signals 0.5.

4.2 2,000-Agent Network with 2 Propositions

Now suppose that we complicate the scenario, allowing for a network of $N = 2,000$ in which each individual must determine beliefs for $K = 2$ propositions. I construct the initial distribution of beliefs as

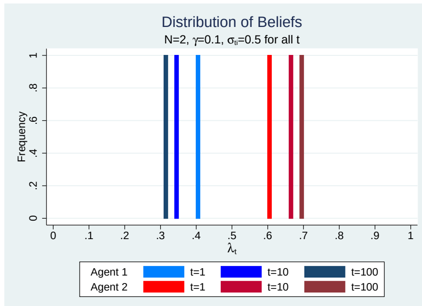
$$\boldsymbol{x} \sim \mathcal{N}\left(\begin{bmatrix} -0.50 \\ 0.50 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right), \tag{12}$$

where agents' initial beliefs are constructed using standard normal CDF applied to each vector:

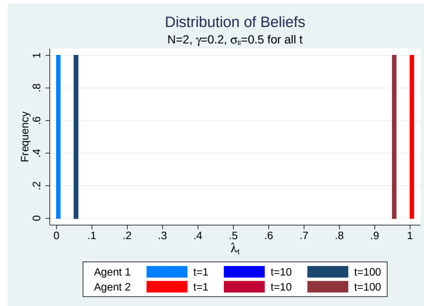
$$\lambda_{it=1}^{k} = \Phi(x_i^k).$$

Figures 4-6 show belief dynamics for this initial distribution of beliefs when the quality of private signals decreases, trust is low $\gamma = 0.1$, and the sequence of private signals is always constant at $\sigma_{ti} = 0.5$,. Assuming agents always have the same quality signal throughout time so that $\theta_{ti} = \theta_i$ for all t, these Figures show the dynamics of beliefs when $\theta_i \sim U[0, 1]$, $\theta_i \sim U[0, 0.5]$, and $\theta_i \sim U[0, 0.1]$. The lower the quality of private signals, the quicker and the stronger is the polarization of beliefs.

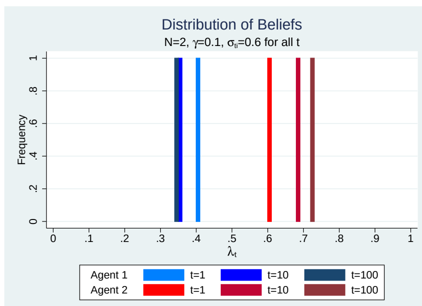
Figures 7 and 8 illustrate the social learning mechanism by comparing what would happen under the intermediate level of private signal quality ($\theta_i \sim U[0, 0.5]$) when beliefs about proposition 2 are fixed. We see that a fixed distribution of beliefs regarding p^2 can generate polarization with respect to p^1.



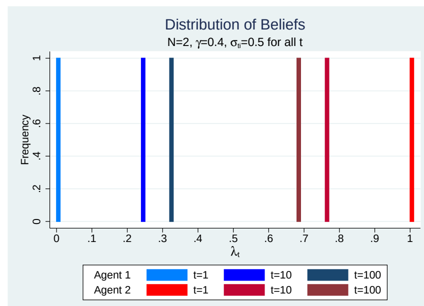
(a) Lowest Trust: $\gamma = 0.1$, $\sigma_{ti} = 0.5$



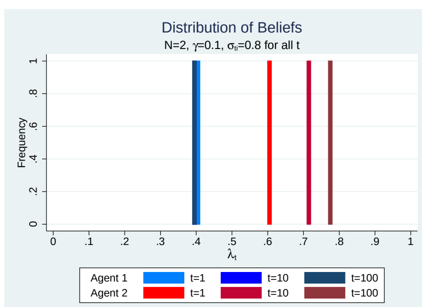
(b) Lower Trust: $\gamma = 0.2$, $\sigma_{ti} = 0.5$



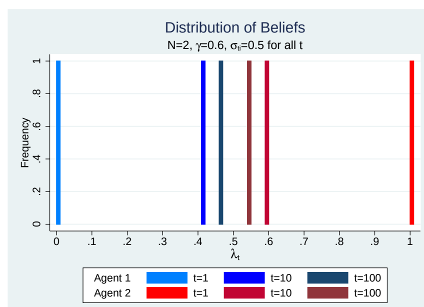
(c) Lowest Trust: $\gamma = 0.1$, $\sigma_{ti} = 0.6$



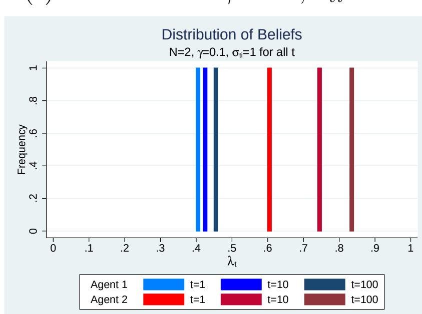
(d) Middle Trust: $\gamma = 0.4$, $\sigma_{ti} = 0.5$



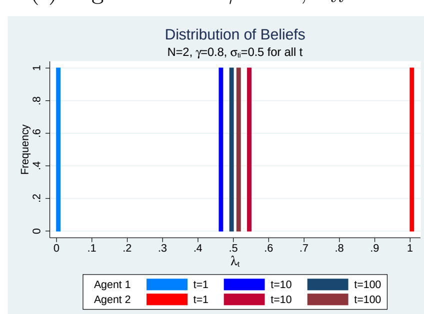
(e) Lowest Trust: $\gamma = 0.1$, $\sigma_{ti} = 0.8$



(f) Higher Trust: $\gamma = 0.6$, $\sigma_{ti} = 0.5$

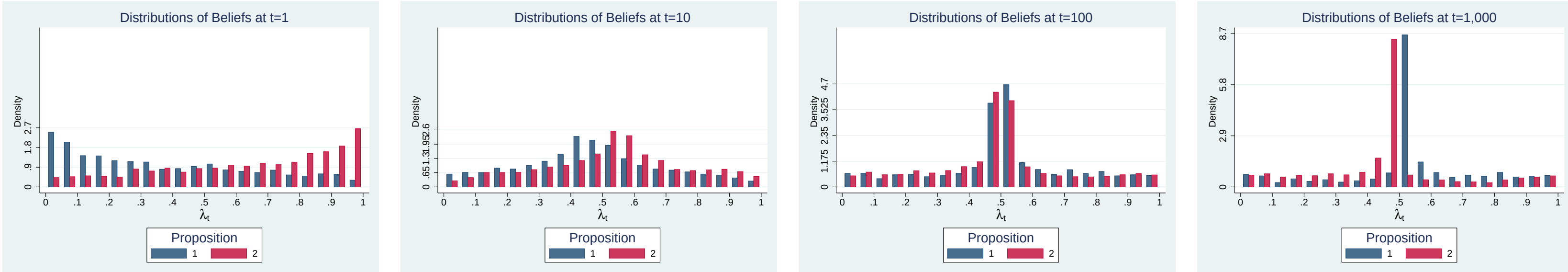


(g) Lowest Trust: $\gamma = 0.1$, $\sigma_{ti} = 1.0$



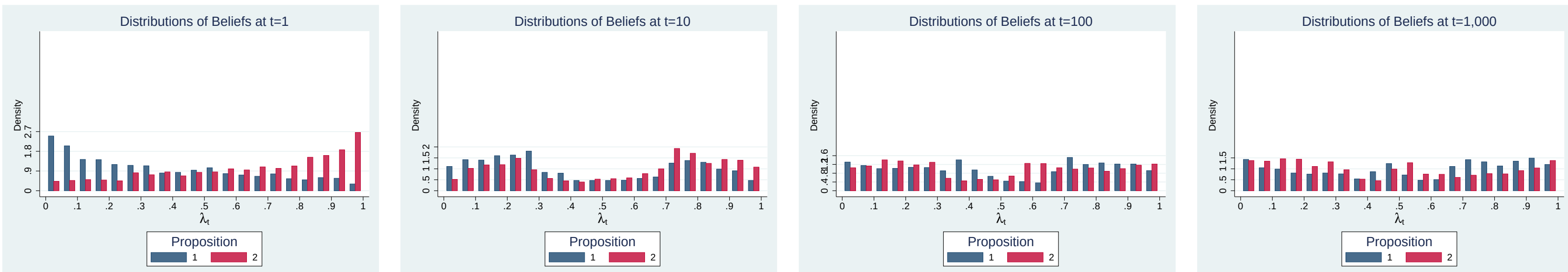
(h) Highest Trust: $\gamma = 0.8$, $\sigma_{ti} = 0.5$

Figure 3: Belief Dynamics when $N = 2$ and $K = 1$
for Various Levels of Trust γ and Private Signals σ_{ti}



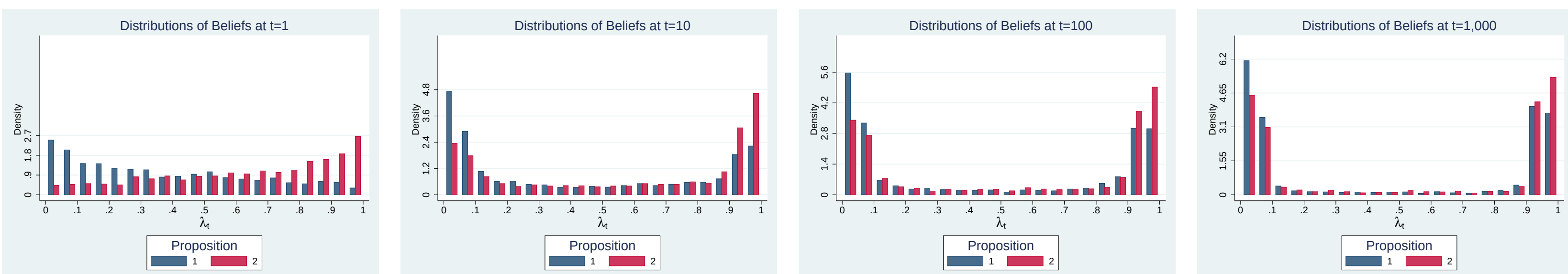
(a) $t = 1$ (b) $t = 10$ (c) $t = 100$ (d) $t = 1,000$

Figure 4: High-Quality Private Signals: $\theta_i \sim U[0, 1]$



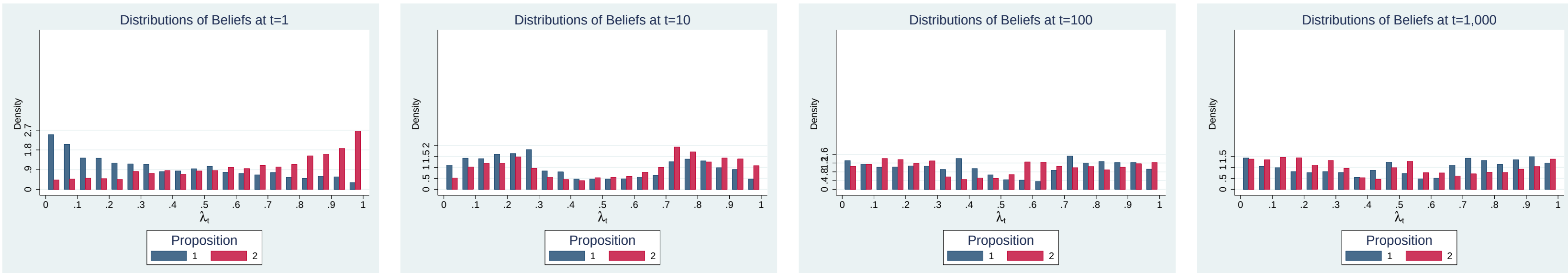
(a) $t = 1$ (b) $t = 10$ (c) $t = 100$ (d) $t = 1,000$

Figure 5: Medium-Quality Private Signals: $\theta_i \sim U[0, 0.5]$



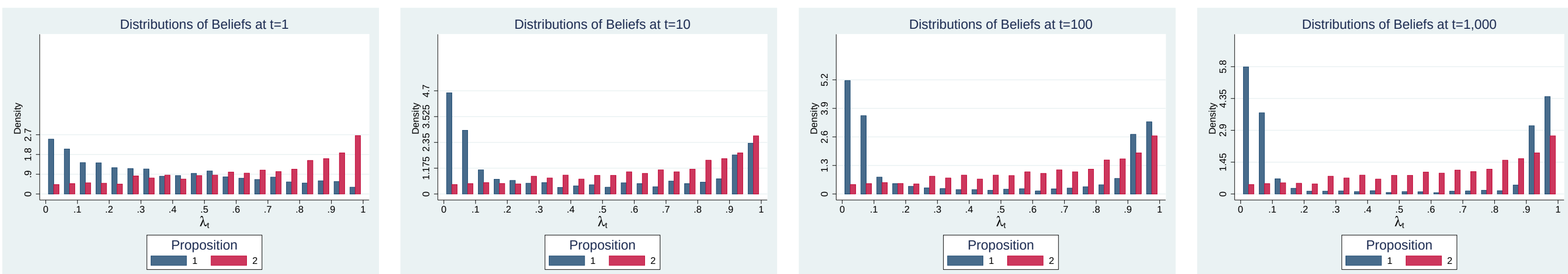
(a) $t = 1$ (b) $t = 10$ (c) $t = 100$ (d) $t = 1,000$

Figure 6: Poor-Quality Private Signals: $\theta_i \sim U[0, 0.1]$



(a) $t = 1$ (b) $t = 10$ (c) $t = 100$ (d) $t = 1,000$

Figure 7: Medium-Quality Private Signals: $\theta_i \sim U[0, 0.5]$ with Updating of λ^2



(a) $t = 1$ (b) $t = 10$ (c) $t = 100$ (d) $t = 1,000$

Figure 8: Medium-Quality Private Signals: $\theta_i \sim U[0, 0.5]$ without Updating of λ^2

5 Conclusion

This paper has made two contributions to the literature on social learning with rule of thumb updating rules. First, I have presented a model of biased assimilation. I have shown that when uncertainty is slow to resolve, backward-looking utility consistent with self-affirmation theory can lead to biases in an agent's processing of information about the state of the world.

This model of biased assimilation motivates key features of the second contribution of the paper, a rule of thumb updating rule generalizing the DeGroot updating rule. The updating rule is the solution to an agent's optimization problem in which she must determine her beliefs about a set of propositions subject to scarce private information. The possibility of either biased assimilation or strategic revelation is sufficient for agents to process social signals differently depending on the sender. This results in a heterogeneous confidence learning rule that is distinct from bounded confidence learning rules in that the agent may actually move her beliefs away from, and not only discard, signals from untrustworthy senders. This learning rule can generate both persistent disagreement and polarization for a connected network in which all individuals process and report their private signals in an unbiased way and extrapolate using social signals "scientifically," interpreting signals based on the sender's past performance on their common observation set, and placing more weight on her interpretation of signals from sources for which she has more evidence.

References

Acemoglu, D. and A. Ozdaglar (2011). Opinion dynamics and learning in social networks. *Dynamic Games and Applications 1*(1), 3–49.

Al-Najjar, N. I. (2009). Decision makers as statisticians: Diversity, ambiguity, and learning. *Econometrica 77*(5), 1371–1401.

Al-Najjar, N. I. and M. M. Pai (2014). Coarse decision making and overfitting. *Journal of Economic Theory 150*, 467–486.

Aliprantis, D. (2015). Covariates and causal effects: The problem of context. *Federal Reserve Bank of Cleveland Working Paper 13-10R.*

Andreoni, J. and T. Mylovanov (2012). Diverging opinions. *American Economic Journal: Microeconomics 4*(1), 209–232.

Dandekar, P., A. Goel, and D. T. Lee (2013). Biased assimilation, homophily, and the dynamics of polarization. *Proceedings of the National Academy of Sciences 110*(15), 5791–5796.

DeGroot, M. H. (1974). Reaching a consensus. *Journal of the American Statistical Association 69*(345), 118–121.

DeMarzo, P. M., D. Vayanos, and J. Zwiebel (2003). Persuasion bias, social influence, and unidimensional opinions. *The Quarterly Journal of Economics 118*(3), 909–968.

Fryer, Jr., R. G., P. Harms, and M. O. Jackson (2013). Updating beliefs with ambiguous evidence: Implications for polarization. *NBER Working Paper 19114*.

Gentzkow, M. and J. M. Shapiro (2011). Ideological segregation online and offline. *The Quarterly Journal of Economics 126*(4), 1799–1839.

Gilboa, I. and D. Schmeidler (2001). *A Theory of Case-Based Decisions*. Cambridge University Press.

Golub, B. and M. O. Jackson (2012). How homophily affects the speed of learning and best-response dynamics. *The Quarterly Journal of Economics 127*(3), 1287–1338.

Hegselmann, R. and U. Krause (2002). Opinion dynamics and bounded confidence: Models, analysis, and simulation. *Journal of Artificial Societies and Social Simulation 5*(3).

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association 81*(396), 945–960.

Jackson, M. O. (2008). *Social and Economic Networks*. Princeton: Princeton University Press.

Jadbabaie, A., P. Molavi, A. Sandroni, and A. Tahbaz-Salehi (2012). Non-Bayesian social learning. *Games and Economic Behavior 76*(1), 210 – 225.

Kandel, E. and N. D. Pearson (1995). Differential interpretation of public signals and trade in speculative markets. *Journal of Political Economy 103*(4), 831–872.

Lorenz, J. (2005). A stabilization theorem for dynamics of continuous opinions. *Physica A 355*, 217–223.

Manski, C. F. (2011). Choosing treatment policies under ambiguity. *Annual Review of Economics 3*, 25–49.

Mohlin, E. (2014). Optimal categorization. *Journal of Economic Theory 152*, 356–381.

Molavi, P., A. Tahbaz-Salehi, and A. Jadbabaie (2015). Foundations of non-Bayesian social learning. *Mimeo., Columbia University*.

Sherman, D. K. and G. L. Cohen (2006). The psychology of self-defense: Self-affirmation theory. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology*, Volume 38, pp. 183–242. San Diego, CA: Academic Press.

Sotiropoulos, D. N., C. Bilanakos, and G. M. Giaglis (2015). A time-variant and non-linear model of opinion formation in social networks. *Mathematical Social Sciences*. Forthcoming.

Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology*, Volume 21, pp. 261–302. San Diego, CA: Academic Press.

Vapnik, V. N. and A. Y. Chervonenkis (1971). On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probability and Its Applications 16*, 264–280.

A Appendix: Analytic Results for Special Cases

We now consider the agent's behavior in tractable, special cases. We begin with the baseline case of one agent and one sender (for a total network size of $N = J + 1 = 2$) and one proposition. We then consider how the agent's optimizing behavior changes when there is still a network size of $N = 2$, but more propositions ($K > 1$). And we finish by considering how the agent's behavior changes when there is still only one proposition ($K = 1$), but the network has more than just two individuals ($N > 2$).

This exercise illustrates that changes in agent i's interpretation of sender j's signal about any given proposition can arise from growing or shrinking the entire space propositions. Similarly, changes in the relative weights agent i gives to sender j's signals once interpreted can be attributed to changes in the size of the network. The basic idea is that information about more propositions allows the agent to better interpret the sender's signal relative to their ideal (ie, to determine the degree of disagreement the two individuals tend to have), and that more individuals allow the agent to better assess the relative reliability of their interpretation(s). Algebraically, comparing Equations 13 and 16, we can see that adding more propositions increases the information available to determine the average disagreement with a given sender (holding fixed the network structure). And comparing Equations 15 and 21, we can see that adding more individuals to the network increases the information available to determine the relative weight given to a given sender (holding fixed the space of propositions).

A.1 Baseline Case: $N = 2$ and $K = 1$

Recall that $K = card(\mathcal{K})$ is the size of the space of propositions, and J is the number of individuals sending signals to the agent. Thus when $J = 1$ there are $N = J + 1 = 2$ individuals in the network. From individual 1's perspective as the agent, and assuming there is always some public information, at a given t we have that

$$\overline{\triangle}_{12} = \overline{\theta}_1^1 \left|\lambda_1^1 - \lambda_2^1\right|. \tag{13}$$

Furthermore, agent 1 interprets sender 2's signal about proposition 1 as

$$s_2^1 = \max\left\{ \ \min\left\{ \ \overline{\sigma}_2^1 - 2\overline{\triangle}_{12}^{\gamma}(\overline{\sigma}_2^1 - \lambda_1^1) \ , \quad 1 \ \right\} \ , \quad 0 \ \right\} \tag{14}$$

and weights that interpretation by

$$w_2^1 = 1 - \theta_1^1. \tag{15}$$

A.2 More Propositions: $N = 2$ and $K \in \mathbb{N}$

Consider the case in which there are still $N = J + 1 = 2$ individuals in the network, but $K \in \mathbb{N}$ propositions over which they determine their disagreement. From individual 1's perspective as the agent, and assuming there is always some public information, at a given t we have that

$$\overline{\triangle}_{12} = \frac{\sum_{k \in \mathcal{K}_{12}} \overline{\theta}_1^k \left|\lambda_1^k - \lambda_2^k\right|}{|\mathcal{K}_{12}|}, \tag{16}$$

As before, agent 1 interprets sender 2's signal about proposition k as

$$s_2^k = \max\left\{ \min\left\{ \overline{\sigma}_2^k - 2\overline{\triangle}_{12}^{\gamma}(\overline{\sigma}_2^k - \lambda_1^k) \;,\; 1 \right\} \;,\; 0 \right\} \tag{17}$$

and weights that interpretation by

$$w_2^k = 1 - \theta_1^k. \tag{18}$$

What has changed here is the calculation of disagreement in Equation 16 relative to the baseline calculation in Equation 13.

A.3 Larger Network: $N \in \mathbb{N}$ and $K = 1$

Consider the case in which there are $N = J + 1$ individuals in the network for $J \in \mathbb{N}$, but again only one proposition with which the agent can calculate disagreement with each sender. From individual 1's perspective as the agent, and assuming there is always some public information, at a given t we have that average disagreement for each sender j is

$$\overline{\triangle}_{1j} = \overline{\theta}_1^1 \left|\lambda_1^1 - \lambda_j^1\right|. \tag{19}$$

Analogously to the $K = 1$ and $J = 1$ case, agent 1 interprets sender j's signal about the only proposition p^1 as

$$s_j^1 = \max\left\{ \min\left\{ \overline{\sigma}_j^1 - 2\overline{\triangle}_{1j}^{\gamma}(\overline{\sigma}_j^1 - \lambda_1^1) \;,\; 1 \right\} \;,\; 0 \right\}. \tag{20}$$

What has changed relative to the baseline case in Equation 15 is that now for each sender j, the agent ($i = 1$) weights their interpretation of j's signal about the sole proposition p^1 by

$$w_j^1 = \frac{|\mathcal{K}_{1j}|}{\sum_{q \in \mathcal{J}^1} |\mathcal{K}_{1q}|}(1 - \theta_1^1). \tag{21}$$

B Appendix: Relationship to the Literature

B.1 Al-Najjar (2009)

Al-Najjar (2009) uses Vapnik and Chervonenkis (1971) theory to describe how information scarcity helps to determine the set of uniformly learnable propositions. Note that information scarcity is the constraint faced by the agent in the model in this paper, but she cannot avoid determining beliefs about p^k, and so she is forced to break the impasse by solving her problem's analogue.

Let A_k be the event that $p^k = 1$ and A_k^C be the event that $p^k = 0$. This paper is focused on how an agent determines her beliefs for the elements of $\{A_1, \ldots, A_K\}$, rather than the algebra $\mathcal{A}$ generated by $\{A_1, \ldots, A_K\}$. Since many different probability measures assigning the same probability to each A_k could be used to assign probability to elements of $\mathcal{A}$, in this sense the problem studied in Al-Najjar (2009) is more difficult than the one studied here.

B.2 Andreoni and Mylovanov (2012)

The special case where $K \in \mathbb{N}$ and $J = 1$ is related to the model in Andreoni and Mylovanov (2012) when we consider compound propositions comprising simpler propositions in $\mathcal{K}$ (ie, assigning a measure to elements of $\mathcal{A}$). For example, suppose that there are two simple propositions in $\mathcal{K}$, p^1 and p^2, and let A_1 be the event that $p^1 = 1$, and A_2 be the event that $p^2 = 1$. Both individuals receive an idiosyncratic private signal about the truth value of p^1, and while both agents receive a private signal about p^2, it is the same (ie, $\sigma^2_{i=1} = \sigma^2_{i=2}$). Determining the realization of the state of nature $\theta = (\alpha, \beta)$ in Andreoni and Mylovanov (2012)'s model is equivalent to determining beliefs about the truth value of the compound proposition

$$(p^1 \wedge p^2) \;\; \vee \;\; (\sim p^1 \wedge \sim p^2),$$

or the probability measure of the event

$$B = (A_1 \cap A_2) \;\; \cup \;\; (A_1^C \cap A_2^C).$$

In other words,

$$\begin{aligned} \lambda\Big((p^1 \wedge p^2) \;\; \vee \;\; (\sim p^1 \wedge \sim p^2)\Big) &= \mu\Big((A_1 \cap A_2) \;\; \cup \;\; (A_1^C \cap A_2^C)\Big) \\ &= 1 - \mu(A_1) - \mu(A_2) - \mu(A_1 \cap A_2). \end{aligned}$$

The insight from Andreoni and Mylovanov (2012) is that private information about one of the simple propositions can lead to disagreement about a compound proposition even when the agents receive the same information about the other simple proposition. The focus in this paper has been on learning about the simple propositions $\{A_1, \ldots, A_K\}$.