NÚMERO 572

ANTONIO JIMÉNEZ-MARTÍNEZ

A Model of Belief Influence in Large Social Networks

Importante

www.cide.edu
ABRIL 2014

Carretera México Toluca 3655, Col. Lomas de Santa Fe, 01210, Álvaro Obregón, México DF, México.
www.cide.edu

www.LibreriaCide.com

Dirección de Publicaciones
publicaciones@cide.edu
Tel. 5081 4003

A Model of Belief Influence in Large Social Networks*

Antonio Jiménez-Martínez†

April 2014

Abstract

This paper develops a model of evolution of beliefs through communication in an exogenous social network. We assume that the agents are Bayesian updaters and that the network enables them to listen to the opinion of others about some uncertain parameter of interest. We explore the effects of the network on the agents' long-run first-order beliefs about the parameter and investigate the aggregation of private information in large societies. Each agent observes private signals about the value of the unknown parameter and, according to his connections in the network, receives private messages from others as well. A message conveys some information about the signal observed by the sender and about the messages that the sender receives from other indirectly connected agents. The informativeness of a message is not strategically chosen but it is exogenously given by the intensity of the connection. Both signals and messages are independent and identically distributed across time but not necessarily across agents. We first characterize the long-run behavior of an agent's beliefs in terms of some entropy-based measures of the conditional distributions of signals and messages available to the agent. Then, we show that the achievement of a consensus in the society is closely related to the presence of prominent agents who are able to change the evolution of other agents' opinions over time. Finally, we show that the influence of the prominent agents must not be very high in order for the agents to aggregate correctly their private sources of information in the long-run.

Keywords: Communication networks, Bayesian updating, private signals, private messages, consensus, correct limiting beliefs.

JEL Classification: D82, D83, D85.

Resumen

Este artículo ofrece un modelo de evolución de creencias a través de comunicación en una red exógena. Suponemos que los agentes son Bayesianos, y que la red les permite escuchar las opiniones de otros sobre cierto parámetro incierto de interés. Estudiamos los efectos de la red sobre las creencias de primer orden de los individuos sobre el parámetro e investigamos la agregación de información privada en sociedades grandes. Cada agente observa señales privadas sobre el parámetro desconocido y, según sus conexiones en la red, recibe mensajes de otros también. Un mensaje transmite información sobre la señal observada por el emisor y sobre los mensajes que el emisor recibe de otros individuos conectados indirectamente. El grado de informatividad de un mensaje no es estratégico sino que está dado exógenamente por la intensidad de la conexión. Señales y mensajes son independientes e idénticamente distribuidos a lo largo del tiempo, pero no necesariamente entre agentes. Primero, caracterizamos las creencias de largo plazo de un agente en términos de algunas medidas de entropía de las distribuciones condicionadas de las señales y mensajes disponibles al agente. Después, demostramos que el logro de consenso en la sociedad está estrechamente relacionado con la presencia de agentes influyentes que son capaces de cambiar la evolución de las opiniones de otros en el tiempo. Por último, demostramos que la influencia de los agentes influyentes no debe de ser muy alta para que los individuos agreguen correctamente sus fuentes de información privada a largo plazo.

*A companion paper appeared in *Emile Borel and the Notion of Strategy: 90*th *Anniversary* under the title "Evolution of Beliefs in Networks." For very useful conversations, I am grateful to Rabah Amir, Luciana Moscoso-Boedo, and Larry Samuelson.

†División de Economía, Centro de Investigación y Docencia Económicas (*CIDE*).

1 Introduction

In most social environments, coordinating decisions when individual payoffs depend on some unknown parameter requires agents to reach similar beliefs about the parameter over time. Examples of such decisions include consumption, occupational, investment, and voting choices. The evolution of agents' beliefs about an uncertain variable usually depend on (a) how they are influenced by their own personal background, education, or personal learning about the variable (e.g., based on their own casual observations or private research), and (b) how they gather information from neighbors, friends, co-workers, local leaders, political actors, and prominent webpages. Social networks are primary channels that transmit opinions about products, job vacancies, investment opportunities, and political programs. The aim of this paper is to explore the relation between the network structure that connects a group of Bayesian updaters and the evolution of their first-order beliefs about some common parameter of interest.

We develop a stylized model of network-based dynamic belief formation where there are two types of information transmission: (a) each agent receives private information about the parameter from an external idiosyncratic source and (b) there is communication between connected agents about the information they are obtaining from their external sources.

More in detail, consider a group of agents who care about a payoff-relevant parameter. Each of them begins with some initial prior and observes over time a sequence of private signals about the parameter. The informativeness of such a stream of signals describes the quality of his learning about the parameter through his external source. In addition, suppose that the agents are connected through an exogenous (weighted and directed) social network that specifies a pattern of relations where each agent can listen to the opinions of others. Each directed connection is characterized by an exogenously given intensity that describes the quality of the information transmission from the speaker to the listener. Specifically, at each date, each agent receives a *non-strategic* message from each agent to whom he has a connection. Such a message is correlated with the sender's signal so that it conveys some information about the private signal that the speaker observes. In this way, each agent receives some information about the stream of signals observed by each of his neighbors over time. Given this framework, we investigate the conditions on the network structure under which the agents will eventually reach a consensus in their first-order beliefs about the parameter value. We also explore the conditions on the network under which the agents aggregate correctly the decentralized information that they obtain from their external sources.

At a more intuitive level, the network describes exogenously given conduits through which the agents listen to others speak about their personal learning. As a motivating example, consider a group of investors deciding their investment in a collective fund. Each investor has a prior about the potential profitability of the fund and also collects some further information by studying privately a number of characteristics of the fund. In addition, through communication, each investor can have some access to the private analyses of the fund features made by other investors.

The amount of information that is transmitted from the speaker to the listener depends only on the exogenous intensity of their directed connection and cannot be manipulated neither by the sender nor by the receiver.[1] We further assume that the intensity of each connection in the network is constant over time, which leads to stationary updating processes, and that signals are communicated along with their sources (i.e., the information sources are "tagged"). In the absence of communication between agents, the evolution of an agent's beliefs is governed only by the combination of his priors and the information that he receives from his private source. Since we allow the agents to begin with different priors, we must specify carefully how an agent i uses another agent j's source to update his beliefs over time once communication from j to i is taken into account. We assume that the messages that agent i receives from agent j allows i to have some access to j's information source but agent i does not incorporate in his belief-updating process any information about j's priors. Instead, in order to revise his beliefs through communication from j, agent i uses his own priors together with the information that he obtains when he places himself in j's position and uses j's private source. Intuitively, thorough communication, agent i maintains his own priors and has some access to the reports collected by agent j in his private analysis of the uncertain variable, but he makes no use whatsoever of j's priors to update his beliefs.

We allow for the transmission of indirect information through the network. In particular, each agent can pass the messages that he receives from his connections to any other agent who in turn has a connection with him.

To complete the groundwork for our analysis, we need to address two final modeling assumptions. First, we need to adopt a particular measure of the informativeness of signals and messages. In general, it is not obvious what constitutes an appropriate measure to rank a set of signals according to their informativeness. Following recent developments on the ranking of information value, we choose an entropy-based measure. More precisely, we use the average of the relative entropy of the induced posterior (for the stationary Bayesian revision process) with respect to the prior. This measure, which has some tradition in information theory, is known as the *power measure*. The power measure is an interesting measure because it induces a complete order over signals. At a more intuitive level, the power measure captures, from an ex-ante viewpoint, the gain of information in moving from the prior to the posterior. We then identify the informativeness of an external source

[1] In this paper, we are not interested in the rich strategic interactions present in a sender-receiver game. We assume that neither the sender nor the receiver choose the informativeness of the messages. Instead, such informativeness is exogenously given by the quality of the channel that connects speaker and listener in the network.

with the power of the corresponding signal and the intensity of a directed connection with the power of the message associated with such a connection.

Second, we need to adopt a notion of what constitutes correct beliefs in our framework. The beliefs of an outside Bayesian observer who begins with some priors and can use over time the external sources available to *all* agents could converge to some limiting beliefs. These limiting beliefs aggregate the decentralized information available to the agents in the sense that the evolution of the observer's beliefs over time obeys to the aggregation of the sources of information available to all the agents. Furthermore, the evolution of the observer's beliefs ignores the flows of information through the network. On the other hand, each agent using only his own private source and the information he obtains from his connections in the network could converge to some limiting beliefs. Suppose that all the agents' beliefs converge to some consensus limiting beliefs. Then, we ask for which networks will the agents' limiting beliefs coincide with the observer's limiting beliefs. A central observation that justifies our approach to correct limiting beliefs is that the aggregation of the decentralized information sources provides us with an estimate of the true parameter value which becomes arbitrarily accurate as the number of agents in the society grows (in the limit, tending to infinity).

Our main results begin by identifying the presence of decay in the flow of information along the connections in the network, in Lemma 1. Including some exogenous decay in the flow of information across connections has been common in the economic literature on networks since the seminal papers by Jackson and Wolinsky (1996), and by Bala and Goyal (2000). An interesting feature of our model is that the presence of decay can be described very precisely in terms of the power measures of signals and of messages in the network.

We then turn to provide a simple but complete characterization of an agent's limiting beliefs, in Theorem 1. In our model, an agent's beliefs converge generically to some beliefs that put probability one on a single parameter value. Our convergence results are determined by both the informativeness of the agent's source and by how he is influenced due to his connections with other agents. The role of the information that the agent obtains from his private source and from his connections in the evolution of his beliefs can be neatly described in terms of a set of measures that depend on the average likelihoods of the various parameter values. Armed with this result, we then provide a set of necessary and sufficient conditions, in Theorem 2, under which an agent j is able to influence another agent i in a way such that both of them end up with the same limiting beliefs that agent j would reach in the absence of communication, that is, due only to his learning from his private source. This can be naturally interpreted as agent j being able, through communication and as time evolves, to convince agent i to share his views about the uncertain parameter. Our characterization result is provided in terms of the power measure of the connections between the agents in the society and of other entropy-related measures. The advantages of using entropy-related measures are that they summarize very precisely a number of features of priors, signals, and messages (which, in fact, constitute the primitives of our model), and that they allow for complete rankings of these distributions in terms of their levels of informativeness.

Intuitively, the conditions identified in Theorem 2 require that the intensity of the connections (either direct or indirect) from i to j be sufficiently high so that agent i be influenced by agent j and, in addition, that the intensity of the connections from agent j to any agent k in the society be sufficiently low so that agent j be not in turn influenced by any other agent in the society. Furthermore, these conditions also relate the power measures of the various connections in the network to the likelihood measures that characterize the agents' limiting beliefs. This allows for further insights. In particular, in addition to the requirements on the intensities of the connections mentioned above, an agent j is a good candidate to influence another agent i when (a) i's priors have little ex-ante uncertainty (i.e., they depart from uniformity), (b) i's source places relatively low intensity on the parameter value that i would favor in the long-run in the absence of communication, (c) i's source places relatively high intensity on the parameter value that j would favor in the long-run in the absence of communication, and (d) the information that j receives from his connections in the network places high intensity on the parameter value that he would favor in the long-run in the absence of communication. Intuitively, these are conditions on how confident are the agents on the parameter values that they pick in the long-run as the most likely ones due only to their private learning, and conditions on the discrepancies between their levels of confidence on such parameter values. All these seem very natural requirements and Theorem 2 states them very precisely in terms of some entropy-based measures with a long tradition in information theory. Given a particular social network, our characterizations of limiting beliefs and of opinion influence (Theorems 1 and 2), combined with our results on the decay of the flow of information (Lemma 1), allow one to identify those prominent agents who might influence crucially others in the society and to assess whether a consensus could finally be achieved.

For applications, our approach seems very useful in those cases where observables can be used to estimate distributions over signals and messages. If we have information about such distributions, then the entropy-related measures used in our results can be easily computed.

Our last result, Proposition 1, provides a sufficient condition on the levels of informativeness of the connections in the network under which, provided that there is a consensus, the agents' limiting beliefs aggregate correctly the information available to all of them through their external sources. We show that a society with consensus attains correct limiting beliefs if the influence of the prominent agents is not so large so as to distortion the evolution of beliefs that results by aggregating the information that all agents collect from their sources. Thus, by combining our results about consensus with the sufficient condition in Proposition 1, we obtain the message that correct limiting beliefs are associated, on the one hand, with a certain degree of influence by some prominent agents. On the other hand, the influence of the prominent agents needs to be bounded so as to prevent any distortion of the belief evolution that obey to the aggregation of all the agents' sources. Our results bear a clear resemblance with the results of the model by Golub and Jackson (2010), whereby agents are not Bayesian and use instead a "rule of thumb" to update their beliefs. Although agents are Bayesian in our model, we also obtain their insight that, to attain consensus and correct limiting beliefs, a certain level of popularity is a bless whereas a disproportionate popularity is a curse.

1.1 Related Literature

The current paper is related to several strands of the economic literature on networks and information transmission, as well as to the statistics branch of information theory.

First, it is related to the literature that builds on the model of network influence due to DeGroot (1974) in order to study how the network structure affects the transmission of first-order beliefs among connected agents. This literature assumes that agents are non-Bayesian updaters and use some "rule of thumb" to incorporate the opinions of other agents into their belief updating. In the DeGroot's model, agents update their beliefs by averaging their neighbors' beliefs according to some exogenous weights that describe the intensity of the connections between the agents. Therefore, in these models, agents are "myopic" and fail to adjust properly for repetitions and dependencies in information they hear several times. A major advantage of these models lies in their tractability. A common feature of the present paper with these models is that the belief-revision processes have a stationary nature over time. A classical contribution in this literature is DeMarzo, Vayanos, and Zwiebel (2003) who propose a network-based explanation for the emergence of "unidimensional" opinions. Closer in spirit to the questions we propose, they also provide some insights on the correctness of learning. Within this literature, perhaps the paper closest to ours in terms of the questions asked is Golub and Jackson (2010). Using a version of the DeGroot's model, they examine whether all beliefs in a large society converge to the truth. They show that the attainment of limiting beliefs arbitrarily close to the true belief is characterized by the condition that the influence of the most influential agent vanishes as the size of the society tends to infinity.

Compared to these papers, our model assumes a perfectly rational protocol of information gathering by the agents. However, as it is also the case in this literature, our model is rigid in the fact that the agents do not choose endogenously the intensities of their connections and in that such intensities are constant over time.

Our work is also related to the theoretical literature on common learning. Considering higher-order beliefs, the question that this strand of the literature addresses is whether a group of agents commonly learn (at least approximately) the true parameter value as time evolves. In a seminal paper, for a setting where there is no communication among the agents, Cripps, Ely, Mailath, and Samuelson (2008) show that (approximate) common learning of the parameter is attained when signals are sufficiently informative and the sets of signals are finite. This result follows regardless of the pattern of correlations between the agents' signals. They assume that the agents start with common priors and ask whether each agent not only assigns sufficiently high probability to some given parameter value but also to the event that each other agent assigns high probability to such a value, and so on, ad infinitum. Our approach is different from theirs in that we focus on the agents' first-order posteriors about the parameter when they start with possibly different priors and use Bayesian updating rules. More specifically, we do not consider ex-ante probabilistic assessments that the agents could make over the histories underlying their beliefs, and we do not explore either the evolution of the agents' higher-order beliefs.[2]

Another difference between the current paper and the models within the learning literature where the agents observe sequences of signals (e.g., Parikh and Krasucki, 1990; Heifetz, 1996; Koessler, 2001; Steiner and Stewart, 2011; Cripps, Mailath, Ely, and Samuelson, 2008 and 2013) is in the fact that this literature usually evaluate the correctness of an agent's beliefs by conditioning the posteriors on a given value of the parameter, which is taken as the actual value. We do not consider this notion of belief correctness since our focus is not on analyzing whether the agents, either individually or commonly, learn the true parameter value by using their higher-order beliefs. In contrast, our model can be seen as an attempt to introduce Bayesian updating rules into the DeGroot's framework of influence and evolution of first-order beliefs. Accordingly, as in the approach pursued by DeMarzo, Vayanos, and Zwiebel (2003), and by Golub and Jackson (2010), our notion of belief correctness asks whether the network structure allows for the aggregation of the decentralized sources of private information of the agents.

Importantly, the result of common learning attainment by Cripps, Mailath, Ely, and Samuelson (2008) requires that the sets of signals and messages be finite. This is not surprising since the

[2] In this respect, our notion of what constitutes "similar beliefs" departs from typical concepts of agreement used in the learning literature. For instance, in their classical justification of the common prior assumption, Savage (1954, p. 48), and Blackwell and Dubins (1962) establish that Bayesian updaters who observe the same sequences of sufficiently informative signals will learn individually the true parameter value, and, as a consequence, they will reach an agreement. Individual learning in this context requires that, conditioned on a parameter value, the agent assigns probability one to the event that her limiting beliefs put probability one to that parameter value. Also, Acemoglu, Chernouzhukov, and Yildiz (2009) use a notion of agreement that requires that the agents assign probability one to the event that their posteriors converge to the same limiting beliefs.

approach followed by the literature on learning usually assumes that each agent is able to keep track of the higher-order beliefs of *all* agents about the signals that each of them is receiving at each period. Clearly, this approach is less appealing when one considers a society where the number of its members tends to infinity. In fact, the argument given by Rubinstein (1989) in his celebrated email game suggests that common learning of the true parameter is precluded with arbitrarily large societies. Our focus on first-order beliefs is justified by the fact that we consider large societies.

Another strand of the literature on learning in social networks considers that, in addition to observing signals, the agents are able to observe their neighbors' past payoffs or past actions. A classical contribution within these models of observational learning is Bala and Goyal (1998), whereby agents take repeated actions and can observe their neighbors' payoffs. They obtain consensus within connected components of the network since each agent can observe whether his neighbors are earning payoffs different from his own. Also, Acemoglu, Dahleh, Lobel, and Ozdaglar (2011) consider that agents can observe their neighbors' past actions and focus on studying asymptotic learning, defined as the convergence of the agents' actions to the right action as the social network becomes large. They provide conditions on the expansion of the network under which there is asymptotic learning when private beliefs are either bounded or unbounded.

The present paper relates also to several branches of the literature on influence in networks with non-Bayesian rules other than the one that stems from the DeGroot's model. For example, Acemoglu, Ozdaglar, and ParandehGheibi (2010) consider that the agents meet pairwise and adopt the average of their pre-meeting beliefs. They study how the presence of agents who influence the beliefs of others, but do not change their own beliefs, interferes with the spread of information along the network. Although they do not focus on consensus in particular, our model allows for insights with a similar flavor to theirs since some spread of beliefs among agents with different opinions is required for consensus in the current paper. In our model, consensus can be precluded when the intensities of the network connections do not allow an agent to listen enough to agents with different opinions. Such an agent plays a similar role to a "forceful" agent in Acemoglu, Ozdaglar, and ParandehGheibi (2010)'s model. Also, the question of whether consensus is attained under non-Bayesian updating rules is analyzed by Acemoglu, Como, Fagnani, and Ozdaglar (2013). They distinguish between "regular" agents, who update their beliefs according to the information they receive from their neighbors, and "stubborn" agents, who never update their beliefs. They show that consensus is never obtained when the society contains stubborn agents with different opinions. Again, this insight bears some resemblance with ours when the connections of some agent do not allow him to change his opinion over time, i.e., when he cannot be influenced by others.

The current paper is also related to the literature on *strategic communication* initiated by Crawford and Sobel (1982). The transmission of information in our model through signals and messages is modeled exact the same way in which a sender transmits information to a receiver in a *cheap talk* game. The crucial difference is that the amount of information transmitted in our model is not endogenously chosen but it is exogenously given by the description of the external sources and of the network connections, and it remains fixed over time. Some recent papers (e.g., Hagenbach and Koessler, 2010; Galeotti, Ghiglino and Squintani, 2013) have proposed single-period sender-receiver interactions to model information transmission in exogenously given networks. Acemoglu, Bimpikis, and Ozdalagar (2011) also consider sender-receiver interactions to model information transmission but they allow for multi-period communication and endogenize the network structure over which messages flow. In their model, agents choose the level of informativeness of the messages they send. Importantly, despite the modeling differences of the present paper with their model, we also obtain their insight that the presence of "influential" agents is important to aggregate information correctly. Nevertheless, in our model the influence of such agents must be bounded.

Finally, our paper is also related to the literature on information theory and to a growing diverse economic literature that uses entropy-based measures to describe levels of informativeness. The concept of power of a signal that we use in this paper was originally proposed by Shannon (1948) in his seminal paper on communication. Subsequently, entropy-based measures have been broadly used by applied mathematicians to model a number of aspects of communication, ranging from data compression and coding to channel capacity or distortion theory. Nevertheless, such measures have remained seldom used by economists for decades. Recently a number of papers are incorporating entropy-based measures to model communication and levels of informativeness in several economic phenomena. For example, Sciubba (2005) uses the power of a signal to rank information in her work on survival of traders in financial markets under asymmetric information. Cabrales, Gossner, and Serrano (2013) propose, for a class of no-arbitrage investment problems under ruin-averse preferences, an entropy-based measure which they call *entropy informativeness*. Using entropy informativeness, they obtain the interesting result that one information structure dominates another if and only if when the investment project associated with the first one is rejected at some price, then so is the project associated with the second. Nevertheless, entropy informativeness is not a novel concept in information theory since it coincides with the power measure of the signal associated to the corresponding information structure.

The rest of the paper is organized as follows. Section 2 presents the model, Section 3 analyses the attainment of consensus and of correct limiting beliefs in the society, and Section 4 concludes with a discussion of the results and of possible extensions. The proofs of all the results are grouped together in the Appendix.

2 The Model

We will use $\Delta(X)$ throughout the paper to denote the set of all Borel probability distributions on a given set X. For a probability distribution P, we will use $E_P[\cdot]$ to denote the expectation operator with respect to P.

There is a finite set of agents $N = \{1, 2, \ldots, n\}$ who care about the true value of an exogenous parameter $\theta \in \Theta = \{\theta_1, \theta_2, \ldots, \theta_L\}$.[3] Time is discrete and indexed by $t \in \{0, 1, 2, \ldots\}$. The true value of θ is selected by nature in period $t = 0$. Each agent i begins with a prior distribution $p_i \in \Delta(\Theta)$ that describes his (subjective) beliefs about the parameter in period $t = 0$.

2.1 Belief Revision from Signals

The realized parameter value θ is not observed directly by any agent. Instead, each agent i obtains some private (noisy) information about the parameter through an idiosyncratic external source of information. Each agent has access only to his own source of information. The interpretation of an agent's external source is that of his personal background, current education, or in general any source of information, aside from what he can hear from other agents, through which he can learn about the parameter. Agent i's source generates in each period $t \geq 1$ a *signal realization* $s_{it} \in S = \{s_1, s_2, \ldots, s_L\}$ which is privately observed by agent i.[4] When the true parameter value is θ, agent i's information source delivers signal $s \in S$ with probability $\phi_i^\theta(s)$. We use $\phi_i(s) = \sum_\Theta \phi_i^\theta(s) p_i(\theta)$ to denote the corresponding unconditional distribution. A *signal profile* in period t is denoted by $\boldsymbol{s}_t = (s_{it})_{i \in N} \in S^n$. Each sequence of signal profiles $\{\boldsymbol{s}_t\}_{t=1}^\infty$ is assumed to be independent and identically distributed across periods (conditional on the parameter). We then define an *informative signal* for agent i (associated with his external source) as a set of conditional distributions over signals $\Phi_i := \left\{\phi_i^\theta \in \Delta(S) : \theta \in \Theta\right\}$. Throughout the paper, we impose the following assumption on informative signals.

Assumption 1. For each agent i there exists at least a parameter value θ such that the conditional distribution $\phi_i^\theta \in \Delta(S)$ has full support.

Assumption 1 above guarantees that the agents' limiting beliefs are well defined. In addition, some results of the paper will require that we strengthen further our assumptions on signals by imposing the following assumption instead.

[3]Although the parameter space is assumed to be finite, the extension of our main results to a compact, but not necessarily finite, parameter space would only change sums to integrals in the appropriate formulae.

[4]We assume that $|S| = |\Theta|$ in order to allow an information source for full information disclosure.

Assumption 2. For each agent i and each parameter value θ, the conditional distribution $\phi_i^{\theta} \in \Delta(S)$ has full support.

Using Bayes rule, an informative signal Φ_i allows agent i to update, in each period $t \geq 1$, his information about the parameter θ. Because distributions over signals are constant over time, this Bayesian updating process is stationary. Let $q_i^s \in \Delta(\Theta)$ denote agent i's posteriors about θ, upon observing signal $s_i = s$, which the agent obtains using the informative signal Φ_i. We have:

$$q_i^s(\theta) = \frac{\phi_i^{\theta}(s)}{\phi_i(s)} p_i(\theta) = \frac{\phi_i^{\theta}(s) p_i(\theta)}{\sum_{\Theta} \phi_i^{\theta'}(s) p_i(\theta')}.$$

2.2 Directed Links in the Social Network

We consider that the agents receive information not only from their private information sources but they can also listen to the opinions of other agents. More precisely, the agents are connected through an exogenous social network which allows each of them to listen, in each period t, to other agents' opinions about the parameter θ. We focus on directed networks where links are one-sided.

A directed link from agent i to agent j is denoted by Ψ_{ij} and it allows i to receive messages from j. Specifically, we assume that, in each period $t \geq 1$, each agent i receives a (private) *message realization* $m_{ijt} \in M = \{m_1, m_2, \ldots, m_L\}$ from each agent j to whom he has a directed link.[5] The message m_{ijt} conveys some (noisy) information about the signal s_{jt} that the sender j is observing at that t. Given the information that agent i obtains about the signal s_{jt} that j observes, agent i uses j's signal Φ_j to update his beliefs about θ. Intuitively, through this type of communication, agent i has some (noisy) access to agent j's background or current education. The specification of a directed link Ψ_{ij}, which we will formally present in the next subsection, determines the level of informativeness of the message m_{ijt}.

Besides direct attention to the information sources of others, we consider that the network also allows for indirect attention. More precisely, we assume that messages can be transmitted indirectly through directed links. In other words, given two directed links Ψ_{ik} and Ψ_{kj}, agent k can receive a message from agent j and then pass it through to agent i. Therefore, through the links Ψ_{ik} and Ψ_{kj}, agent i receives, in each period, two different messages, one direct message from agent k (m_{ik}) and one indirect message from agent j (m_{ij}, which was previously received by agent k from agent j). The message m_{ik} conveys information about the signal s_k observed by agent k while the message m_{ij} conveys information about the signal s_j observed by agent j. Nonetheless, for the clarity of exposition, it is convenient to focus first on the description of the transmission only of direct

[5] We assume that $|M| = |S|$ so that each directed link allows for full information disclosure about the signal observed by the sender.

messages through links.

2.3 Information Transmission (only) with Direct Messages

Suppose for the moment that the agents receive only direct messages so that an agent cannot pass to another agent a message that he has received from a third agent. A *message vector*[6] received by agent i in period t is denoted by $\boldsymbol{m}_{it} = (m_{ijt})_{j \in N \setminus \{i\}}$ and a *message profile* in period t is denoted by $\boldsymbol{m}_t = (\boldsymbol{m}_{it})_{i \in N}$. We assume that each sequence of message profiles $\{\boldsymbol{m}_t\}_{t=1}^{\infty}$ is independent and identically distributed across periods.

2.3.1 Belief Revision from Direct Messages

For each period t, the distribution over messages observed by agent i, conditional on agent j's signal realization $s_{jt} = s$, is denoted by σ_{ij}^{s} and the corresponding unconditional distribution is denoted by σ_{ij}. We define an *informative message* from agent j to agent i as a set of conditional distributions over messages $\Sigma_{ij} := \left\{ \sigma_{ij}^{s} \in \Delta(M) : s \in S \right\}$.

Using Bayes rule, an informative message Σ_{ij} allows agent i to update, in each period t, his beliefs about j's private signal s_{jt} by observing the message m_{ijt}. As in the case of informative signals, this Bayesian updating process is stationary since distributions over messages are constant over time.

Given the informative signals and the informative messages, an agent i can use his information about the sequence $\left\{ s_{jt} \right\}_{t=1}^{\infty}$ to update over time his beliefs about the parameter θ. In other words, by combining the informative signal Φ_j with the informative message Σ_{ij}, agent i can use the information that he hears directly from agent j to update his beliefs about θ.

Since we consider that the priors of the agents may differ, we must specify very carefully how an agent i uses other agent j's informative signal Φ_j together with an informative message Σ_{ij} to update his priors p_i about θ. The difficulty comes from the fact that, since agents i and j may have different initial opinions about the occurrence of θ (i.e., $p_i \neq p_j$), then they may have different opinions about the unconditional distributions over signals and messages as well. With different priors, this disagreement appears even though the agents have the same (common knowledge) information about the corresponding conditional distributions, ϕ_i^{θ}, ϕ_j^{θ}, and σ_{ij}^{s}. Given this difficulty, we assume that, through a directed link Ψ_{ij}, agent i is able to use signal Φ_j, but with the restriction that his knowledge about the occurrence of s_j (conditional on the messages that he receives from

[6] In principle, our description of message vectors captures a situation in which each agent receives messages from each other agent in the society. Nevertheless, the specification of a link Ψ_{ij} will determine the degree of informativeness of the messages m_{ijt} that flow through it. In some cases, the corresponding degree of informativeness can be null, which is interpreted as if there is actually no directed link from agent i to agent j and, therefore, as if i receives no message whatsoever from j.

agent j) is given by the corresponding Σ_{ij} and that his prior knowledge of θ is given solely by p_i. In particular, agent i does not incorporate any information about p_j to update his beliefs about θ. For our setting, this seems to be the most natural approach with heterogeneous priors.

Specifically, for two agents $i, j \in N$, let

$$\phi_j[i](s) := \sum_{\Theta} \phi_j^{\theta}(s) p_i(\theta)$$

denote the probability that agent i observes signal s according to agent j's private signal Φ_j, but given that he uses his own priors p_i to compute the probability of occurrence of θ. Note that $\phi_i[i] = \phi_i$ for each agent i. Also, let $q_i^s[\Phi_j]$ denote agent i's posterior beliefs over Θ given that he knows that agent j observes signal s, and given that he uses agent j's signal Φ_j to update his own beliefs p_i about θ. Using Bayes rule, we have:

$$q_i^s[\Phi_j](\theta) = \frac{\phi_j^{\theta}(s)}{\phi_j[i](s)} p_i(\theta) = \frac{\phi_j^{\theta}(s) p_i(\theta)}{\sum_{\Theta} \phi_j^{\theta'}(s) p_i(\theta')}.$$

The expression above simply describes how agent i updates his priors when he has access to agent j's informative signal Φ_j and, in addition, he knows that the signal that j observes is s. However, the information that agent i receives about the signal that agent j observes is noisy, and the corresponding degree of informativeness is determined by the informative signal Σ_{ij}. Specifically, given the informative signal Σ_{ij}, let $\psi_{ij}^{\theta}(m) = \sum_S \sigma_{ij}^s(m) \phi_j^{\theta}(s)$ be the probability that agent i receives message m from agent j, conditional on θ being the true parameter value. Also, let $\psi_{ij}(m) = \sum_S \sigma_{ij}^s(m) \phi_j(s)$ be the corresponding unconditional probability. Then, if we denote by r_{ij}^m the posterior beliefs of agent i about the signal that agent j observes, given that i receives message m from j, we have:

$$r_{ij}^m(s) = \frac{\sigma_{ij}^s(m)}{\psi_{ij}(m)} \phi_j(s) = \frac{\sigma_{ij}^s(m) \phi_j(s)}{\sum_S \sigma_{ij}^{s'}(m) \phi_j(s')}.$$

Finally, we use $q_{ij}^m \in \Delta(\Theta)$ to denote agent i's posteriors about θ, upon receiving message $m_{ijt} = m$ from agent j. Using the conditional probabilities $q_i^s[\Phi_j]$ and r_{ij}^m obtained above, the posteriors q_{ij}^m are correctly specified as:

$$q_{ij}^m(\theta) := \sum_S r_{ij}^m(s) q_i^s[\Phi_j](\theta) = \frac{p_i(\theta)}{\psi_{ij}(m)} \sum_S \frac{\sigma_{ij}^s(m) \phi_j^{\theta}(s) \phi_j(s)}{\phi_j[i](s)}.$$

For convenience, we define

$$\widetilde{\psi}_{ij}^{\theta}(m) := \sum_S \sigma_{ij}^s(m) \phi_j^{\theta}(s) \left[\frac{\phi_j(s)}{\phi_j[i](s)} \right]. \tag{1}$$

The expression $\widetilde{\psi}_{ij}^{\theta}(m)$ above is the formal definition of the subjective probability (conditional on θ) that agent i assigns to receiving message m from agent j, under the additional condition that agent i

uses ϕ_j^θ and p_i (instead of p_j) to compute the occurrence of s_j. Thus, agent i uses his own priors p_i to compute the probability that agent j observes his signals or, in other words, he uses $\phi_j[i]$ instead of ϕ_j. Note that if agents i and j share the same priors, $p_i = p_j$, then $\phi_j[i] = \phi_j$ so that $\widetilde{\psi}_{ij}^\theta = \psi_{ij}^\theta$. Using the subjective conditional probability $\widetilde{\psi}_{ij}^\theta$, the posteriors above over the parameter space can be alternatively expressed as

$$q_{ij}^m(\theta) = \frac{\widetilde{\psi}_{ij}^\theta(m)}{\psi_{ij}(m)} p_i(\theta). \tag{2}$$

We can now define formally a directed link. The *directed link* from agent j to agent i associated with the signal Φ_j and the informative message Σ_{ij} is the set of conditional distributions over messages

$$\Psi_{ij} := \left\{ \psi_{ij}^\theta \in \Delta(M) : \theta \in \Theta, \text{ such that } \psi_{ij}^\theta(m) = \sum_S \sigma_{ij}^s(m)\phi_j^\theta(s) \right\}.$$

Note that the informational primitives of our model are (a) the agents' priors p_i, (b) the set of informative signals Φ_i, and (c) the set of informative messages Σ_{ij}. The directed links are obtained from these primitives. Nevertheless, it is very convenient to consider the set of directed links as our main analytical tool to formalize the transmission of information between agents. Intuitively, a directly link Ψ_{ij} allows agent i to have some access to agent j's private signal and, by doing so, to update his beliefs about the parameter θ. The amount of information that agent i receives in this way about θ is determined by the informativeness of the directed link Ψ_{ij}. In the next subsection we describe the approach that we propose to measure such informativeness.

Notice that we do not allow for the possibility that the agents manipulate strategically the messages they send, neither that they withhold any information they posses about their signals. The information they send to others is noisy but exogenously determined. As we mentioned earlier, our research interest in this paper is not on strategic decisions about information disclosure. One way to regard this benchmark is as one in which the agents previously made some investments in their links in a way such that only *hard information* can flow through them. Then, once the links are formed, their degrees of informativeness remain fixed and cannot be altered neither by senders nor by receivers.

2.3.2 Measuring the Informativeness of Signals and Links

To measure the degree of informativeness about θ attached to the agents' external sources (through informative signals) and to the communication between agents (through directed links), we make use of some entropy-based concepts commonly used in information theory. We begin by defining the entropy of a distribution.[7]

[7]In Definition 1, it follows the convention $0 \log 0 = 0$, which is justified by continuity.

Definition 1. Let X be a finite set. The *entropy* (or *Shannon entropy*) of a probability distribution $P \in \Delta(X)$ is

$$H(P) := -\sum_X P(x) \log P(x).$$

The entropy of a distribution is always nonnegative and measures the average information content one is missing from the fact that the true realization of the associated random variable is unknown. In other words, it measures the ex-ante uncertainty of the corresponding random variable. In our model, the entropy of an agent i's priors will serve as an upper bound on the degree of informativeness of his informative signal Φ_i.

To measure the information content of informative signals and messages, we rely on the concept of relative entropy between distributions.

Definition 2. Let X be a finite set and let $P, Q \in \Delta(X)$. The *relative entropy* (or *Kullback-Leiber distance*) of P with respect to Q is[8]

$$D(P \| Q) := \sum_X P(x) \log \frac{P(x)}{Q(x)}.$$

The relative entropy is not a metric,[9] but, considering X as a sample space, it constitutes a formal measure of the gain of information in moving from distribution Q to distribution P. The relative entropy is always nonnegative and equals zero if and only if $P = Q$ almost everywhere.

We apply the relative entropy to the agents' posteriors with respect to their priors. Then, we define the *power of the informative signal* Φ_i as the expectation of the relative entropy of the posterior q_i^s with respect to the prior p_i.

Definition 3 (**Power of the informative signal).**

$$\mathbb{P}(\Phi_i) := E_{\phi_i}\left[D(q_i^s \| p_i)\right] = \sum_S \phi_i(s) D\Big(q_i^s \| p_i\Big). \tag{3}$$

The power measure allows us to rank completely any set of informative signals according to their degree of informativeness. We say that Φ_i *is at least as informative as* Φ_i' if $\mathbb{P}(\Phi_i) \geq \mathbb{P}(\Phi_i')$. Analogously, we define the *power of the directed link* Ψ_{ij} as the expectation of the relative entropy of the posterior q_{ij}^m with respect to the prior p_i.

Definition 4 (**Power of the directed link).**

$$\mathbb{P}(\Psi_{ij}) := E_{\psi_{ij}}\Big[D(q_{ij}^m \| p_i)\Big] = \sum_M \psi_{ij}(m) D\Big(q_{ij}^m \| p_i\Big). \tag{4}$$

[8]The following conventions are used: $0 \log(0/0) = 0$ and, based on continuity arguments, $0 \log(0/a) = 0$ and $a \log(a/0) = \infty$.

[9]In particular, the relative entropy is not symmetric and it does not satisfy the triangle inequality either.

Using the power of a directed link, we say that Ψ_{ij} *is at least as informative as* Ψ'_{ij} if $\mathbb{P}(\Psi_{ij}) \geq \mathbb{P}(\Psi'_{ij})$. If $\mathbb{P}(\Psi'_{ij}) = 0$, then we interpret this as if there is actually no directed link from agent i to agent j. The power measure is an interesting measure to account for the amount of information that one gains by switching from his priors to his posteriors. Moreover, it induces a complete order over signals in a number of interesting decision problems.[10]

We are ready now to define a directed network. A *directed network* Ψ is a set of directed links which connects the agents in the society with a non-null degree of informativeness:

$$\Psi := \Big\{\Psi_{ij} : i, j \in N, i \neq j, \text{ such that } \mathbb{P}(\Psi_{ij}) > 0\Big\}.$$

As we mentioned earlier, a social network is determined by a set of informational primitives. For a social network Ψ, we denote the set of its primitives by $\Upsilon(\Psi) := \langle \{p_i\}_{i\in N}, \{\Phi_i\}_{i\in N}, \big\{\Sigma_{ij}\big\}_{i,j\in N} \rangle$.

We turn now describe how information is transmitted through indirect messages.

2.4 Information Transmitted (both) with Direct and Indirect Messages

A *directed path* from agent i to agent j is a sequence $\gamma_{ij} = (\Psi_{ii_1}, \Psi_{i_1 i_2}, \ldots, \Psi_{i_K j})$ of directed links such that $\mathbb{P}(\Psi_{ii_1}) > 0$, $\mathbb{P}(\Psi_{i_K j}) > 0$, and $\mathbb{P}(\Psi_{i_k i_{k+1}}) > 0$ for each $k \in \{1, \ldots, K-1\}$. We use $\Gamma_{ij}[\Psi]$ to denote the set of all directed paths from agent i to agent j under network Ψ and

$$N_i := \Big\{ j \in N : \text{there is some } \gamma_{ij} \in \Gamma_{ij}[\Psi] \Big\}$$

to denote set of agents to whom agent i has a directed path. We say that a directed network Ψ is *connected* if, for each agent $i \in N$, there is at least one directed path $\gamma_{ij} \in \Gamma_{ij}[\Psi]$ to each other agent $j \in N \setminus \{i\}$. Intuitively, a network is connected if it allows each agent to hear (either directly or indirectly) the opinions of each other agent in the society. Network connectedness can be regarded as a basic prerequisite to study the achievement of a consensus in a society.

For a given network Ψ, an agent i may receive messages from other agent j through (possibly) multiple paths $\gamma_{ij} \in \Gamma_{ij}[\Psi]$. We will restrict attention to those paths which convey the highest amount of information.

For the transmission of indirect messages, we make the natural assumption that an agent k uses the same conditional distribution $\sigma_{ik}^{s_k} = \sigma_{ik}^{m_{kj}}$ to transmit information to agent i both about the signal s_k that he observes and about the message m_{kj} that he receives from another agent j.

[10] As mentioned in the Introduction, entropy-based measures have a long tradition in the theory of informativeness orderings. In particular, the concept of power of a signal (under the label *entropy power of a signal*) dates back to Shannon (1948)'s seminal paper on communication. The power of a signal has been subsequently used in economics, for instance, by Sciubba (2005) to rank information in her work on survival of traders in financial markets under asymmetric information. Recently, Cabrales, Gossner, and Serrano (2013) propose, for a class of no-arbitrage investment problems under ruin-averse preferences, an entropy-based measure, which they call *entropy informativeness*, that coincides with the power measure of the signal associated to the corresponding information structure.

The interpretation of this assumption is that we consider the existence of a common technology for information transmission, which is equally used both for signals and for messages that pass from one agent to another.

Let $\psi^{\theta}_{ij}[\gamma_{ij}] \in \Delta(M)$ denote the distribution over messages received by i from j through the directed path γ_{ij}, conditional on the parameter value being θ. Let the corresponding unconditional distribution be denoted by $\psi_{ij}[\gamma_{ij}]$. For a directed path γ_{ij}, let $\pi^{s}_{\gamma_{ij}} \in \Delta(M)$ denote the conditional distribution over messages received by agent i from agent j, conditional on agent j observing signal $s_j = s$. Then, for a directed path specified as $\gamma_{ij} = (\Psi_{ii_1}, \Psi_{i_1 i_2}, \dots, \Psi_{i_K j})$, using the total probability rule, we obtain, for the message $m_{ij} = m$:

$$\pi^{s}_{\gamma_{ij}}(m) = \sum_{M} \dots \sum_{M} \sigma^{m_{i_1 i_2}}_{ii_1}(m) \prod_{k=1}^{K-2} \sigma^{m_{i_{k+1} i_{k+2}}}_{i_k i_{k+1}}(m_{i_k i_{k+1}}) \sigma^{m_{i_K j}}_{i_{K-1} i_K}(m_{i_{K-1} i_K}) \sigma^{s}_{i_K j}(m_{i_K j}). \tag{5}$$

Then, using the distribution $\pi^{s}_{\gamma_{ij}}$ specified above, we have $\psi^{\theta}_{ij}[\gamma_{ij}](m) = \sum_S \pi^{s}_{\gamma_{ij}}(m)\phi^{\theta}_j(s)$ and $\psi_{ij}[\gamma_{ij}](m) = \sum_S \pi^{s}_{\gamma_{ij}}(m)\phi_j(s)$.

For a directed path $\gamma_{ij} = (\Psi_{ii_1}, \Psi_{i_1 i_2}, \dots, \Psi_{i_K j})$, we can now extend straightforwardly the expression that gives us agent i's posteriors about parameter θ to the case where i receives messages indirectly from agent j through the path γ_{ij}. First, we extend the expression of agent i's subjective conditional probability given by (1) to the case where messages flow through paths:

$$\widetilde{\psi}^{\theta}_{ij}[\gamma_{ij}](m) := \sum_{S} \pi^{s}_{\gamma_{ij}}(m)\phi^{\theta}_j(s)\left[\frac{\phi_j(s)}{\phi_j[i](s)}\right].$$

Secondly, if we let $q^{m}_{ij}[\gamma_{ij}] \in \Delta(\Theta)$ denote agent i's posteriors about θ, upon receiving message $m_{ij} = m$ from agent j (through the path γ_{ij}), then we obtain the following expression,

$$q^{m}_{ij}[\gamma_{ij}](\theta) = \frac{\widetilde{\psi}^{\theta}_{ij}[\gamma_{ij}](m)}{\psi_{ij}[\gamma_{ij}](m)} p_i(\theta),$$

which is completely analogous to the expression in (2). Finally, we can also readily apply the concept of power of a link to a path. For the expression of the posterior beliefs given above, the *power of the path* γ_{ij} is defined as

$$\mathbb{P}(\gamma_{ij}) := E_{\psi_{ij}[\gamma_{ij}]}\Big[D\Big(q^{m}_{ij}[\gamma_{ij}] \,\|\, p_i\Big)\Big] = \sum_{M} \psi_{ij}[\gamma_{ij}](m) D\Big(q^{m}_{ij}[\gamma_{ij}] \,\|\, p_i\Big). \tag{6}$$

Of course, agent i can receive indirect messages from another agent j through several different paths in the network. We restrict attention to those paths to agent j which allows agent i to receive the highest amount of information about θ, that is, to those paths in the set

$$\Big\{\widehat{\gamma}_{ij} \in \Gamma_{ij}[\Psi] : \mathbb{P}(\widehat{\gamma}_{ij}) \geq \mathbb{P}(\gamma_{ij}) \ \ \forall \gamma_{ij} \in \Gamma_{ij}[\Psi]\Big\}.$$

If the set above is not singleton, then we randomly pick one of its elements as our path of interest and denote it by $\widehat{\gamma}_{ij}$. For future reference, we will denote $\psi^{\theta}_{ij}[\widehat{\gamma}_{ij}] =: \widehat{\psi}^{\theta}_{ij}$ and $\psi_{ij}[\widehat{\gamma}_{ij}] =: \widehat{\psi}_{ij}$.

2.5 Evolution of Beliefs, Consensus, and Correct Beliefs

We introduce now a few additional concepts to analyze the evolution of the agents' first-order posteriors. A *period-t history for agent i* is a sequence $h_{it} := ((s_{i0}, \boldsymbol{m}_{i0}), (s_{i1}, \boldsymbol{m}_{i1}), \dots, (s_{it}, \boldsymbol{m}_{it}))$ of signals and message vectors received by agent i. The posterior belief of agent i about parameter θ in each period t is given by the random variable $\mu_i(\theta \,|\, h_{it}) : \Theta \to [0,1]$. For each agent i and each value of the parameter θ, the sequence of random variables $\{\mu_i(\theta \,|\, h_{it})\}_{t=1}^{\infty}$ is a bounded martingale,[11] which ensures that the agents' posterior beliefs converge almost surely (see, e.g., Billingsley, 1995, Theorem 35.5).

Definition 5. A *consensus is (asymptotically) achieved in the society* if the posterior beliefs of all agents converge to the same value regardless of their priors, that is, if for each $i \in N$, each $p_i \in \Delta(\Theta)$, and for some (common) probability distribution $p \in \Delta(\Theta)$,

$$\lim_{t \to \infty} \mu_i(\cdot \,|\, h_{it}) = p.$$

Our notion of what constitutes correct beliefs requires that the network permits the aggregation of the pieces of information transmitted by the private signals available to the agents. Consider an external observer who has access to the external sources available to all agents in the society but cannot use any directed link in the social network. The observer's priors are given by a distribution $p_{\text{ob}} \in \Delta(\Theta)$. A *period-$t$ history for the external observer* is a sequence $h_t := (\boldsymbol{s}_0, \boldsymbol{s}_1, \dots, \boldsymbol{s}_t)$ of signal profiles. The posterior belief of the external observer about parameter θ in each period t is given by the random variable $\mu_{\text{ob}}(\theta \,|\, h_t) : \Theta \to [0,1]$.[12] With these preliminaries in hand, correct limiting beliefs require that the communication processes allowed by the network structure aggregate the diverse information obtained by the agents (from their external sources), exactly such as the external observer does. A key observation in our framework is that, for large enough societies, the observer's limiting beliefs are arbitrarily accurate estimates of the true parameter value.

Definition 6. The directed network Ψ attains *correct limiting beliefs* if a consensus is achieved in the society and, in addition, for each $i \in N$,

$$\lim_{t \to \infty} \mu_i(\cdot \,|\, h_{it}) = \lim_{t \to \infty} \mu_{\text{ob}}(\cdot \,|\, h_t).$$

[11] More formally, $\{\mu_i(\theta \,|\, h_{it})\}_{t=1}^{\infty}$ is a bounded martingale with respect to the (conditional) measure on Θ which is induced by the priors $(p_i)_{i \in N}$, and the conditional distributions ϕ_i^{θ}, ψ_{ij}^{θ}, for $i, j \in N$.

[12] Again, for each value of the parameter θ, the sequence of random variables $\{\mu_{\text{ob}}(\theta \,|\, h_t)\}_{t=1}^{\infty}$ is a bounded martingale so that the external observer's posteriors converge almost surely.

3 Results

3.1 Decay in the Flow of Information

An important result of our model is that the degree of informativeness of information decreases as it flows from one link to another. First, the informativeness of a directed link from agent i to agent j does not exceed the informativeness of agent j's own informative signal when both agents begin with the same priors, $p_i = p_j$. This is very intuitive. Our model captures the presence of some (exogenous) decay in the transmission of information, which is associated to informative messages that do not fully disclose the private information available to the sender. Only for the extreme case in which the informative message Σ_{ij} is such that agent i fully learns the signal that agent j observes, there is no loss of information through the directed link and agent i obtains the same amount of information about θ using the link Ψ_{ij} as agent j does using directly his informative signal Φ_j. Secondly, with the same logic, decay in the transmission of information also follows in the case where information is conveyed by indirectly transmitted messages, provided that the two agents connected through the corresponding path share the same priors. The following lemma provides these intuitive results in terms of the power measure.

Lemma 1 (**Decay in the Transmission of Information).** *(a) Consider a directed link in social network $\Psi_{ij} \in \Psi$, and suppose that agents i and j have the same priors p. Then, $\mathbb{P}(\Psi_{ij}) \leq \mathbb{P}(\Phi_j)$. Moreover, $\mathbb{P}(\Psi_{ij}) = \mathbb{P}(\Phi_j)$ if and only if the informative message Σ_{ij} associated with the directed link Ψ_{ij} is such that agent i fully learns the signals observed by agent j from the external source associated with Φ_j. (b) Consider a directed path in a social network $\gamma_{ij} = (\Psi_{ik}, \Psi_{kj}) \in \Gamma_{ij}(\Psi)$. Suppose that agents i and j have the same priors p. Then, $\mathbb{P}(\gamma_{ij}) \leq \mathbb{P}(\Psi_{kj})$. Moreover, $\mathbb{P}(\gamma_{ij}) = \mathbb{P}(\Psi_{kj})$ if and only if the informative message Σ_{ik} associated with the directed link Ψ_{ik} is such that agent i fully learns the messages received by agent k from agent j.*

For the case of different priors, the insights provided by Lemma 1 (a) continue to hold if the result is rephrased as follows. Suppose that, as an alternative to his private signal Φ_j, agent j places himself in agent i's position and uses the directed link Ψ_{ij} to update his beliefs about the parameter. Then, the information about θ that agent j receives through this directed link Ψ_{ij} is less precise than the information that he would obtain using directly his signal Φ_j. This reinterpretation of the result can be verified directly from the proof of the lemma. The insights provided by Lemma 1 (b) also continue to hold under the analogous restatement of the result. Nevertheless, if two agents i and j have different priors and we simply ask about the relation between $\mathbb{P}(\Psi_{ij})$ and $\mathbb{P}(\Phi_j)$, then it could be the case that $\mathbb{P}(\Psi_{ij}) > \mathbb{P}(\Phi_j)$. This is due to the role that the agents' priors have on the informativeness of signals and links.

It can be verified that, for any agent $i \in N$, $\mathbb{P}(\Phi_i) = H(p_i) - E_{\phi_i}[H(q_i^s)]$ so that $\mathbb{P}(\Phi_i) \leq H(p_i)$. Thus, $\mathbb{P}(\Phi_i) = H(p_i)$ if and only if the average entropy of agent i's posteriors (obtained only from his private source) vanishes. In other words, $\mathbb{P}(\Phi_i) = H(p_i)$ whenever agent i obtains full information about the parameter from his private signal. Suppose that all the agents in the society begin with some common priors p. Then, note that for an agent i to obtain full information about the parameter from a directed link to another agent j, it must be the case that (a) agent i obtains full information about agent j's informative signal (i.e., $\mathbb{P}(\Psi_{ij}) = \mathbb{P}(\Phi_j)$) and (b) agent j obtains full information about the parameter from his own informative signal (i.e., $\mathbb{P}(\Phi_j) = H(p)$). Therefore, from the result in Lemma 1, we observe that $\mathbb{P}(\Psi_{ij}) \leq H(p)$ for each $i, j \in N$. For the particular case of a finite parameter space, it is well known that the entropy of any distribution is bounded.[13] With common priors, this result leads to the implication that, for $L < \infty$, $H(p)$ constitutes an upper bound on the degree of informativeness about θ that any agent in the society can obtain, regardless of the network structure.[14]

3.2 Characterizing Limiting Beliefs and Consensus

For an agent i and a parameter value θ, we define the function $G_i : \Theta \to \mathbb{R}$ as

$$G_i(\theta) := E_{\phi_i}\left[\log \phi_i^{\theta}\right] = \sum_{S} \phi_i(s) \log \phi_i^{\theta}(s). \tag{7}$$

The value $G_i(\theta)$ is always negative and describes the (average) likelihood that the informative signal Φ_i (that is, through agent i's external source) assigns to θ being the true parameter value. Let $\Theta_i \subseteq \Theta$ be the set specified as $\Theta_i := \arg\max_{\theta\in\Theta} G_i(\theta)$. To account for the information that an agent i receives from another agent j, we define the function $F_{ij} : \Theta \to \mathbb{R}$ as

$$F_{ij}(\theta) := E_{\widehat{\psi}_{ij}}\left[\log \widehat{\psi}_{ij}^{\theta}\right] = \sum_{M} \widehat{\psi}_{ij}(m_{ij}) \log \widehat{\psi}_{ij}^{\theta}(m_{ij}). \tag{8}$$

The value $F_{ij}(\theta)$ is always negative and describes the (average) likelihood that the most informative directed path from agent i to agent j, $\widehat{\gamma}_{ij}$, assigns to θ being the true parameter value. For an agent i, we then specify the set $\Theta_i^* \subseteq \Theta$ as $\Theta_i^* := \arg\max_{\theta\in\Theta} \left\{G_i(\theta) + \sum_{j\in N_i} F_{ij}(\theta)\right\}$.

The next theorem shows that the convergence of an agent's posteriors is determined by the aggregation of likelihoods that his external source and all his directed paths in the network place on the various parameter values.

[13] A central result of information theory establishes for our model that $H(p) \leq \log L$ for each $p \in \Delta(\Theta)$.
[14] In general, the relative entropy between two distributions need not be bounded.

Theorem 1. *Consider a social network Ψ and suppose that Assumption 1 holds. Then, for each sequence of histories $\{h_{it}\}_{t=1}^{\infty}$, agent i's limiting beliefs satisfy:*

(i) $\lim_{t\to\infty}\mu_i(\theta|h_{it}) = 0$ for each $\theta \notin \Theta_i^*$;

(ii) if Θ_i^* is singleton so that $\Theta_i^* = \{\theta^*\}$ for some $\theta^* \in \Theta$, then $\lim_{t\to\infty}\mu_i(\theta^*|h_{it}) = 1$;

(iii) if Θ_i^* is not singleton, then $\lim_{t\to\infty}\mu_i(\theta|h_{it}) = p_i(\theta)/\sum_{\Theta_i^*} p_i(\theta')$ for each $\theta \in \Theta_i^*$.

We obtain the interesting feature that the various influences on a given agent's beliefs of his external information source and of the opinions that he receives from others are additively aggregated to determine his limiting beliefs. We then naturally interpret the parameter values in each Θ_i^* as the ones which are favored in the long-run by both agent i's private source and the information that he receives from others through his connections in the network.

Consider now the extreme case of our model which describes a society where the agents are completely isolated and no agent receives any information whatsoever from any other agent. In particular, this situation is obtained in our model if each informative message Σ_{ij} is such that messages do not depend on observed signals, i.e., $\sigma_{ij}^s(m) = \sigma_{ij}(m)$ for each $s \in S$. If this is the case, then it follows from the expression in (5) that $\pi_{\gamma_{ij}}^s(m) = \pi_{\gamma_{ij}}(m)$ for any directed path γ_{ij} so that $\psi_{ij}[\gamma_{ij}](m) = \psi_{ij}^{\theta}[\gamma_{ij}](m) = \pi_{\gamma_{ij}}(m)$. As a consequence, $F_{ij}(\theta) = -H(\widehat{\psi}_{ij})$ for each pair of different agents in the society. Since each F_{ij} does not depend on θ in this extreme case, we have $\Theta_i = \Theta_i^*$ for each agent $i \in N$. Then, using the results of Theorem 1 above, we obtain the intuitive insight that an agent i's limiting beliefs are governed only by his external information source, Φ_i.[15] We naturally interpret each Θ_i as the set of parameter values which agent i favors in the long-run due only to his own private learning, in the case of complete isolation.

It is convenient to consider this extreme case in the absence of communication as a reference situation in order to describe our results on consensus and correct limiting beliefs. We will refer to this case as the *complete isolation case*. Yet, in our model, we do actually allow for communication. Starting from the complete isolation case, we then study how some agents influence the evolution of others' beliefs when we allow for communication between them according to their connections in the network. From the results in Theorem 1, we observe that this question can be analyzed by studying the relation between the network structure Ψ and the differences existing between each Θ_i and Θ_i^*.

As a convenient step to explore the achievement of a consensus in the society, we would like to study the conditions on the network structure under which some agents are able to influence others' opinions in a way such that all of them end up with the same opinions about which parameter values are the most likely ones. Specifically, we wish to identify the features of the network which,

[15] Notice that this extreme case can be alternatively obtained if we simply exclude the possibility of network connections in our model.

starting from a reference situation $\Theta_i \neq \Theta_j$, induces $\Theta_i^* = \Theta_j^*$, with the additional requirement that $\Theta_j^* = \Theta_j$. The interpretation of this result is that the agents j are able to influence agents i's opinions so that all them put positive probability in the long-run to the same parameter values that agents j considered with positive probability based solely on their external sources, that is, the parameter values in Θ_j. From the results provided by Theorem 1, we see that $F_{ij}(\theta)$ gives us a convenient and intuitive measure of the extent to which agent j is able to influence agent i to believe that θ is the true parameter value.

Definition 7. Given a social network Ψ, and two distinct agents $i, j \in N$ such that $\Theta_i \neq \Theta_j$, we say that agent j influences agent i if $\Theta_i^* = \Theta_j^* = \Theta_j$.

In the definition above, for an agent to influence another, we require that he must not be in turn influenced by other agents. Also, from the results of Theorem 1, we note that if the sets Θ_i^* and Θ_j^* satisfying $\Theta_i^* = \Theta_j^*$ are not singleton and agents i and j begin with different priors, then their limiting beliefs differ. In this case, agents i and j do agree on the set of parameter values that have positive probability of occurrence. However, each Θ_i^* is generically singleton because non-singleton sets Θ_i^* are not robust to small perturbations of the network.[16]

The following example illustrates (a) how limiting beliefs are obtained for the complete isolation case and (b) how we measure the intensity with which an agent influences the evolution of others' beliefs in a way such that a consensus is finally achieved in the society.

Example 1. Consider a set of $n = 4$ agents who care about two possible parameter values, i.e., $\Theta = \{\theta_1, \theta_2\}$. The agents are connected through a social network $\Psi = \{\Psi_{13}, \Psi_{21}, \Psi_{24}, \Psi_{32}, \Psi_{43}\}$, which is depicted in Figure 1.

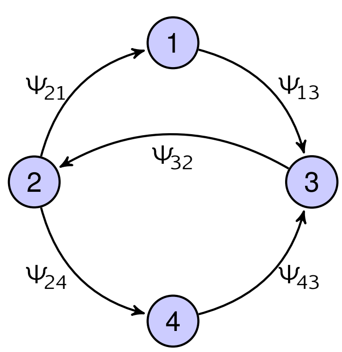

FIGURE 1

The agents begin with the (common) priors $p(\theta_1) = p(\theta_2) = 1/2$, so that $H(p) = -\log(1/2)$. The

[16]The set of networks for which some Θ_i^* is not singleton has Lebesgue measure zero in the set of all possible networks.

agents' private signals are specified as: (agent 1) $\phi_1^{\theta_1}(s_1) = 1/6$ and $\phi_1^{\theta_2}(s_1) = 1/2$; (agent 2) $\phi_2^{\theta_1}(s_1) = 1/3$ and $\phi_2^{\theta_2}(s_1) = 2/3$; (agent 3) $\phi_3^{\theta_1}(s_1) = 2/5$ and $\phi_3^{\theta_2}(s_1) = 9/10$; (agent 4) $\phi_4^{\theta_1}(s_1) = 2/3$ and $\phi_4^{\theta_2}(s_1) = 1/3$. With this information, we can compute: (agent 1) $G_1(\theta_1) = -0.7188$ and $G_1(\theta_2) = -0.6931$ so that $\Theta_1 = \{\theta_2\}$; (agent 2) $G_2(\theta_1) = -0.752$ and $G_2(\theta_2) = -0.752$ so that $\Theta_2 = \{\theta_1, \theta_2\}$; (agent 3) $G_3(\theta_1) = -0.7744$ and $G_3(\theta_2) = -0.8744$ so that $\Theta_3 = \{\theta_1\}$; (agent 4) $G_4(\theta_1) = -0.752$ and $G_4(\theta_2) = -0.752$ so that $\Theta_4 = \{\theta_1, \theta_2\}$. Therefore, when the agents use only their external information sources, there is some discrepancy in their limiting beliefs. In particular, agents 2 and 4 end up with their initial priors, agent 1 favors the parameter value θ_2, and agent 3 favors the parameter value θ_1.

To describe the links of the directed network, we specify the corresponding informative messages as: (link 13) $\sigma_{13}^{s_1}(m_1) = 1$ and $\sigma_{13}^{s_2}(m_1) = 0$; (link 21) $\sigma_{21}^{s_1}(m_1) = 4/5$ and $\sigma_{21}^{s_2}(m_1) = 1/5$; (link 24) $\sigma_{24}^{s_1}(m_1) = 1/4$ and $\sigma_{24}^{s_2}(m_1) = 0$; (link 32) $\sigma_{32}^{s_1}(m_1) = 1/3$ and $\sigma_{32}^{s_2}(m_1) = 2/3$; (link 43) $\sigma_{43}^{s_1}(m_1) = 9/10$ and $\sigma_{43}^{s_2}(m_1) = 1/10$. With this information, we can obtain the associated distributions $\widehat{\psi}_{ij}^{\theta} \in \Delta(\{m_1, m_2\})$, for $\theta \in \{\theta_1, \theta_2\}$. Observe that the network in Figure 1 is connected so that each agent can listen to the opinions of each other agent through some directed path. Also, some agents are connected through several paths. In particular, agent 2 can listen to agent 3 through the paths $\gamma_{23} = (\Psi_{21}, \Psi_{13})$ and $\gamma'_{23} = (\Psi_{24}, \Psi_{43})$. By computing the power of each path, we pick the paths which transmit the highest amount of information.

For the directed links, we obtain: (link 13) $F_{13}(\theta_1) = -0.7744$ and $F_{13}(\theta_2) = -0.8744$; (link 21) $F_{21}(\theta_1) = -0.6956$ and $F_{21}(\theta_2) = -0.6931$; (link 24) $F_{24}(\theta_1) = -0.3835$ and $F_{24}(\theta_2) = -0.3867$; (link 32) $F_{32}(\theta_1) = -0.6993$ and $F_{32}(\theta_2) = -0.6993$; (link 43) $F_{43}(\theta_1) = -0.7448$ and $F_{43}(\theta_2) = -0.7747$.

For the directed paths which transmit the highest amount of information, we obtain: (path 12) $F_{12}(\theta_1) = -0.6993$ and $F_{12}(\theta_2) = -0.6993$; (path 14) $F_{14}(\theta_1) = -0.662$ and $F_{14}(\theta_2) = -0.662$; (path 23) $F_{23}(\theta_1) = -0.7221$ and $F_{23}(\theta_2) = -0.7299$; (path 31) $F_{31}(\theta_1) = -0.6932$ and $F_{31}(\theta_2) = -0.6931$; (path 34) $F_{34}(\theta_1) = -0.662$ and $F_{34}(\theta_2) = -0.662$; (path 41) $F_{41}(\theta_1) = -0.6931$ and $F_{41}(\theta_2) = -0.6931$; (path 42) $F_{42}(\theta_1) = -0.6971$ and $F_{42}(\theta_2) = -0.6971$.

We observe that, for each agent $i \neq 3$, the value of $F_{i3}(\theta_1)$ is higher than the value of $F_{i3}(\theta_2)$. This indicates that, through communication, agents place a relatively high intensity on the parameter value that agent 3 favors in the complete isolation case. Then, we analyze whether agent 3 can be an influential agent in this society. This turns out to be the case. By computing the corresponding values of $G_i(\theta) + \sum_{j \neq i} F_{ij}(\theta)$, for each $i = 1, \dots, 4$ and each $\theta \in \{\theta_1, \theta_2\}$, using the values above, we obtain $\Theta_i^* = \{\theta_1\}$ for each $i = 1, \dots, 4$. Thus, the society achieves a consensus in which each agent believes in the long-run with probability one that θ_1 is the true parameter value.

Recall that the value $G_i(\theta)$ describes the intensity with which agent i's source induces him to believe in the long-run that θ is the true parameter value, and the value $F_{ij}(\theta)$ describes the intensity with which agent j induces agent i to believe that θ is the true parameter value. Then, it is natural that the achievement of a consensus in a society can be analyzed by studying the likelihood functions G_i and F_{ij} for $j \in N_i$. This can be done using directly the results of Theorem 1, which is the approach followed in Example 1 above.

A complementary way to study how influential are some agents in a network and the achievement of a consensus would involve to use the power of the paths in the network. We next provide two necessary and sufficient conditions, in terms of the power of the paths of the network, under which an agent is able to influence another agent. For an agent j to influence another agent i, it would be natural, on the one hand, to require that the informativeness of the (most informative) path from i to j be sufficiently high. On the other hand, it would be also natural to require that the informativeness of the (most informative) path from agent j to any other agent in the society be sufficiently low in order to prevent j from being influenced. This turns out to be the case, and such conditions are stated formally in Theorem 2 below. Other intuitive message of Theorem 2 is that some j is more likely to influence another agent i when agent j's private learning places high intensity on some parameter values. This can be interpreted as agent j being very convinced of his opinion about the true parameter values due to his background or personal analysis of the features of the parameter. Such an agent j can be viewed as a "self-confident" agent.

Theorem 2. *Consider a social network Ψ and two different agents $i, j \in N$ such that $\Theta_i \neq \Theta_j$. Suppose that Assumption 2 holds, then agent j influences agent i if and only if Ψ satisfies the following conditions:*

(i) for agents i and j:

$$\mathbb{P}(\widehat{\gamma}_{ij}) > \Big[G_i(\theta_i) - G_i(\theta_j)\Big] + \max_{\theta \notin \Theta_j} \sum_{h \in N_i} \Big[F_{ih}(\theta) - F_{ih}(\theta_j)\Big] + H(p_i) - E_{\widehat{\psi}_{ij}}\Big[H(q^m_{ij}[\widehat{\gamma}_{ij}])\Big], \tag{7a}$$

for any $\theta_i \in \Theta_i$ and any $\theta_j \in \Theta_j$.

(ii) for agent j:

$$\begin{aligned} &\max_{k \in N_j} \Big\{\mathbb{P}(\widehat{\gamma}_{jk}) + E_{\widehat{\psi}_{jk}}\Big[H(q^m_{jk}[\widehat{\gamma}_{jk}])\Big]\Big\} \\ &\qquad < G_j(\theta_j) + \sum_{h \in N_j} F_{jh}(\theta_j) - \max_{\theta \notin \Theta_j}\Big[G_j(\theta) + \sum_{h \in N_j} F_{jh}(\theta)\Big] + H(p_j), \end{aligned} \tag{7b}$$

for any $\theta_j \in \Theta_j$.

The conditions provided by Theorem 2 are intuitive. Condition (i) identifies a lower bound on the level of informativeness of the (most informative) directed path from agent i to agent j under which j is able to affect i's beliefs in a way such that i ends up favoring in the long-run the same parameter values that j favors in the complete isolation case, i.e., those parameter values in Θ_j. On the other hand, condition (ii) identifies an upper bound on the level of informativeness of the (most informative) directed path from agent j to any other agent in the society, which characterizes the situation where j continues to believe in the long-run that his most-favored parameter values in the complete isolation case, Θ_j, continue to be the most likely ones when he listens to others' opinions.

From inequality (7a) above, we observe that the required lower bound on $\mathbb{P}(\widehat{\gamma}_{ij})$ increases with the difference $G_i(\theta_i) - G_i(\theta_j)$. Intuitively, it is easer for us to influence some agent when the intensity that he puts on the parameter values that he considers the most likely ones does not differ much from the intensity that he puts on the values that we consider to be the most likely ones.

Inequality (7a) also states that the required lower bound on $\mathbb{P}(\widehat{\gamma}_{ij})$ decreases with each $F_{ih}(\theta_j)$, $h \in N_i$, which conveys the intuition that it is easer for us to influence some agent when the informative messages that he receives place a high intensity on the parameters value that we consider to be the most likely ones. Furthermore, such a lower bound increases with $\max_{\theta \notin \Theta_j} \sum_{h \in N_i} F_{ih}(\theta)$, which can be interpreted as a measure of the highest intensity that the informative messages received by agent i from the rest of the society places on a parameter value other than the ones favored by agent j. Then, we obtain that it is easer for agent j to influence agent i when the informative messages that i receives do not place a large intensity on parameter values different from those in Θ_j.

Finally, inequality (7a) also states that the required lower bound on $\mathbb{P}(\widehat{\gamma}_{ij})$ increases with the entropy $H(p_i)$ and decreases with the average entropy $E_{\widehat{\psi}_{ij}}[H(q_{ij}^m[\widehat{\gamma}_{ij}])]$. Therefore, for agent j to influence agent i, we need higher values of $\mathbb{P}(\widehat{\gamma}_{ij})$ when agent i's priors are very uncertain ex-ante[17] and when agent i's posteriors, based solely on the information that he receives from agent j, have in average little uncertainty.[18]

On the other hand, from inequality (7b), we observe that high values of the informativeness of agent j's path to another agent k in the society are compatible with j not being influenced by k when: (a) j's private source and/or the opinions that he receives from k put a high intensity on the parameter values that he considers the most likely ones in the complete isolation case (i.e., high values of $G_j(\theta_j)$ and/or of $F_{jh}(\theta_j)$), (b) j's private source and/or the opinions that he receives from other agents $h \in N_i$ do not place a high intensity on parameter values different from the ones that he favors in the complete isolation case (i.e., low values of $\max_{\theta \notin \Theta_j}\{G_j(\theta) + \sum_{h \in N_j} F_{jk}(\theta)\}$), (c) j's priors are very uncertain ex-ante (i.e., high values of $H(p_j)$), and (d) the ex-ante uncertainty in average of j's posteriors, conditioned on the messages that he receives from k, is low (i.e., low values of

[17] Higher values of $H(p_i)$ are associated with priors which are close to the uniform case $\overline{p}_i(\theta) = 1/L$ for each $\theta \in \Theta$.

[18] Higher values of $E_{\widehat{\psi}_{ij}}[H(q_{ij}^m[\widehat{\gamma}_{ij}])]$ are associated with posteriors which put large probabilities on a few parameter values.

$E_{\widehat{\psi}_{ij}}[H(q^m_{ij}[\widehat{\gamma}_{ij}])])$.

The conditions identified in Theorem 2 above have neat interpretations in terms of some measures of informativeness which have a long tradition in information theory. Moreover, if we know the set of primitives $\Upsilon(\Psi)$, then we can compute the entropy-based measures that appear in conditions (7a) and (7b). In this case, by using Theorem 2, we can determine with full precision whether an agent is able to influence another or not.

Example 2. Consider again the social network analyzed in Example 1. Figure 2 depicts the power of each directed link in the network. The power measures $\mathbb{P}(\Psi_{ij})$ below are computed using the set of primitives described in Example 1.

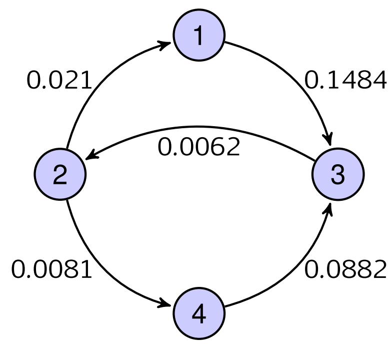

FIGURE 2

Recall that, in the complete isolation case agents 2, 3, and 4 favor in the long-run parameter value θ_1 but, unlike agent 3, agents 2 and 4 consider that parameter value θ_2 has also some positive probability of occurrence. Based only on their external sources, the discrepancy of opinions is relatively higher between agents 1 and 3. Agent 1 favors θ_2 with probability one while agent 3 favors θ_1 with probability one. The result obtained in Example 1 that agent 3 influences the rest of the society so as to achieve a consensus is not surprising now if we note the communication intensities described in Figure 2. We observe that the intensity with which agent 1 listens to agent 3's opinions is the highest in the society (0.1484). Also, agent 3 listens to the others' opinions exclusively through his link with agent 2, and the power of this link is the lowest in the society (0.0062). In addition, we obtain $\mathbb{P}(\widehat{\gamma}_{23}) = 0.0474$ while $\mathbb{P}(\Psi_{21}) = 0.021$. In other words, through agent 1, agent 2 pays more attention to the opinions of agent 3 than to the opinions of agent 1 himself. On the other hand, using the result in Lemma 1, we know that $\mathbb{P}(\widehat{\gamma}_{41}) < 0.021$ so that, given that $\mathbb{P}(\Psi_{43}) = 0.0882$, we observe that agent 4 also pays more attention to the opinions of agent 3 than to the opinions of agent 1. In short, using the power measure, we observe that agent 3 is a good candidate to influence the opinions of the rest of the society. Clearly, he is the agent in the best position, according to the directed links of the network and to their intensities, to do so.

Then, we examine whether condition (7a) of Theorem 2 holds for for the directed link Ψ_{13} or, in other words, whether agent 3 induces agent 1's limiting beliefs to put probability one on parameter value θ_1. Recall from Example 1 that $H(p) = -\log(1/2) = 0.6931$ Also, from the computations of Example 1, we observe that

$$G_1(\theta_2) - G_1(\theta_1) = 0.0257$$

and

$$F_{12}(\theta_2) - F_{12}(\theta_1) = 0, \;\; F_{13}(\theta_2) - F_{13}(\theta_1) = -0.1, \;\text{ and }\; F_{14}(\theta_2) - F_{14}(\theta_1) = 0.$$

Thus, to verify whether whether condition (7a) holds for the directed link Ψ_{13}, we only need to compute the expected entropy $E_{\psi_{13}}[H(q_{13}^m)]$. Using the informative messages specified in Example 1, we compute: $q_{13}^{m_1}(\theta_1) = 4/13$ and $q_{13}^{m_2}(\theta_1) = 6/7$. With these posteriors, we easily obtain $E_{\psi_{13}}[H(q_{13}^m)] = 0.5447$. Then, according to condition (7a), for agent 3 to influence agent 1, we need that the intensity of the directed link Ψ_{13} be above the bound

$$0.0257 - 0.1 + 0.6931 - 0.5447 = 0.0471.$$

This intensity is clearly exceeded in our example since we have $\mathbb{P}(\Psi_{13}) = 0.1484$.

Now, we turn to examine whether condition (7b) holds for agent 3 so that he is not influenced by any of the three other agents. First, from the computations of Example 1, we observe that

$$G_3(\theta_1) + \sum_{k\neq 3} F_{3k}(\theta_1) - \left[G_3(\theta_2) + \sum_{k\neq 3} F_{3k}(\theta_2)\right] + H(p) = 0.7932. \tag{9}$$

Second, we can easily compute

$$E_{\widehat{\psi}_{31}}[H(q_{31}^m[\widehat{\gamma}_{31}])] = 0.6909, \;\; E_{\psi_{32}}[H(q_{32}^m)] = 0.6869, \;\text{ and }\; E_{\widehat{\psi}_{34}}[H(q_{34}^m[\widehat{\gamma}_{34}])] = 0.6002.$$

We, therefore, obtain

$$\mathbb{P}(\Psi_{32}) + E_{\psi_{32}}[H(q_{32}^m)] = 0.0062 + 0.6869 = 0.6932.$$

Furthermore, by using the result in Lemma 1, we know that

$$\mathbb{P}(\widehat{\gamma}_{31}) + E_{\widehat{\psi}_{31}}[H(q_{31}^m[\widehat{\gamma}_{31}])] < 0.021 + 0.6909 = 0.7119, \;\text{ and}$$
$$\mathbb{P}(\widehat{\gamma}_{34}) + E_{\widehat{\psi}_{34}}[H(q_{34}^m[\widehat{\gamma}_{34}])] < 0.0081 + 0.6002 = 0.6083.$$

Since $\max_{k\neq 3}\left\{\mathbb{P}(\widehat{\gamma}_{3k}) + E_{\widehat{\psi}_{3k}}[H(q_{3k}^m[\widehat{\gamma}_{3k}])]\right\}$ is less than 0.7119, which exceeds not the required value 0.7932, identified in (9) above, we obtain that condition (7b) holds for agent 3.

In this example, one can analogously analyze the conditions in Theorem 2 for the most informative paths $\widehat{\gamma}_{23}$ and $\widehat{\gamma}_{43}$ to conclude that these conditions are satisfied in a way such that agent 3 influences agents 2 and 4 as well, and a consensus is achieved.

At this point we note the implication that, with heterogenous priors, the achievement of a consensus in the society is guaranteed by the existence of a set of agents $\overline{N} \subset N$ such that (a) each agent $j \in \overline{N}$ favors some given (common) parameter value $\Theta_j = \{\theta^*\}$ in the complete isolation case and (b) the agents in $\overline{N}$ influence the rest of agents in the society so that all the agents end up believing that the parameter value θ^* is the true one.

Corollary 1. *Consider a given social network Ψ. A consensus is attained in the society if there exists a set of agents $\overline{N} \subset N$ such that (i) for each $j \in \overline{N}$, we have $\Theta_j = \{\theta^*\}$ for some $\theta^* \in \Theta$, and (ii) for each $i \in N \setminus \overline{N}$ there is some agent $j \in \overline{N}$ such that agent j influences agent i.*

Following the related literature, if such set of agents $\overline{N}$ exists, then we refer to it as a set of *prominent agents*.

3.3 Correct Limiting Beliefs and Influence of Prominent Agents

Correct limiting beliefs refer in our model to the limiting beliefs of the external observer. Recall that we assume that the external observer has access to the private sources available to all agents in the society, Φ_i, $i \in N$, but has no access to the information that flows through the links of the social network. The observer's posteriors constitute an aggregate of the information possessed by the agents. More importantly, they become an arbitrarily accurate estimate of the true parameter value as the number of agents in the society tends to infinity.

The next proposition provides a sufficient condition on the levels of informativeness of the links of the network under which correct limiting beliefs are attained in the society. Recall that the achievement of a consensus in the society is a prerequisite to evaluate whether correct beliefs are attained.

Proposition 1. *Consider a directed network $\mathbb{P}(\Psi)$ and suppose that a consensus is achieved in the society in a way such that, for some $\theta^* \in \Theta$, we have $\Theta_i^* = \{\theta^*\}$ for each $i \in N$. If the following condition*

$$\sum_{i \in N} \sum_{j \in N_i} \left[F_{ij}(\theta^*) - F_{ij}(\theta) \right] < 0,$$

is satisfied for each $\theta \in \Theta \setminus \{\theta^*\}$, then the social network $\mathbb{P}(\Psi)$ attains correct limiting beliefs.

The sufficient condition identified in Proposition 1 is intuitive. Suppose that the aggregation of the pieces of information obtained from the private sources of all the agents leads one to believe in the long-run that a given parameter value θ^* is the true one. Then, the condition above imposes some restrictions on the influence of prominent agents. It requires that there is no agent whose influence on others be such that some agents' limiting posteriors favor alternative parameter values $\theta \in \Theta \setminus \{\theta^*\}$.

The message conveyed by the result of Proposition 1 above is reminiscent of the main results obtained by Golub and Jackson (2010) in their work without Bayesian updating (Propositions 2 and 3). Although they use a notion of correct beliefs that differs slightly from ours,[19] correctness of beliefs requires in their model that the influence of prominent agents vanish as the size of the society grows. In our setting, as well as in theirs, a disproportionate popularity by some agent(s) is the crucial obstacle to correct limiting beliefs.

We observe that the presence of prominent agents is desirable to achieve consensus. However, to attain correct beliefs, the influence of the prominent agents must not be too high so as to alter the agents' beliefs in a way such that they end up favoring parameter values different from the ones that one would favored by aggregating the agents' sources of information.

Yet, the fact that the condition stated in Proposition 1 is only sufficient is illustrated in the following example

Example 3. Consider again the social network described in Example 1. Recall that this society achieves a consensus in which all the agents' beliefs converge to a distribution that puts probability one on the parameter value θ_1. This consensus was propitiated by the fact that agent 3 is able to influence the rest of agents in the society. Using the computations of the functions F_{ij} provided in Example 1, it is easy to verify that

$$\begin{aligned}\sum_{i=1}^{4}\sum_{j\neq i}\Big[F_{ij}(\theta_1)-F_{ij}(\theta_2)\Big] &= [-0.7744+0.8744]+[-0.6956+0.6931]\\ &+[-0.7221+0.7299]+[-0.3835+0.3867]+[-0.6932+0.6931]\\ &= 0.1305>0,\end{aligned}$$

so that the sufficient condition of Proposition 1 is not satisfied. Nevertheless, we can check whether the consensus beliefs still coincide can with the limiting beliefs of the external observer who aggregates the information transmitted by all external sources. Observe that, from the result in Theorem 1, the parameter values that are favored in the long-run by the external observer are those in the set $\arg\max_{\Theta}\sum_{i\in N}G_i(\theta)$. Then, using the computations of the functions G_i provided in Example 1, we obtain $\sum_{i=1}^{4}G_i(\theta_1)=-2.9972$ and $\sum_{i=1}^{4}G_i(\theta_1)=-3.0719$ so that, for our social network, we have $\lim_{t\to\infty}\mu_{\mathrm{ob}}(\theta_1)=1$. Thus, although the sufficient condition in Proposition 1 is not satisfied, the influence of agent 3 does not interfere with the limiting beliefs that are obtained by aggregating the external sources, and correct limiting beliefs are attained in this social network.

[19]Their definition of belief correctness also requires that some external observer aggregates the pieces of information initially held by the agents.

4 Concluding Comments

Both our focus on first-order beliefs and our notion of belief correctness are appealing when one considers societies large enough. For small societies, the use of a definition of belief correctness based on conditioning posteriors on a given parameter value would deliver the message that, provided that the agents are allowed to use higher-order beliefs, they always learn the truth because both signals and messages are independent over time in our model. This implication would follow rather directly from the main result of Cripps, Ely, Mailath, and Samuelson (2008). Nevertheless, for the approach usually pursued in the learning literature, recent research (e.g., Parikh and Krasucki, 1990; Heifetz, 1996; Koessler, 2001; Steiner and Stewart, 2011) shows that the presence of communication among the agents may in some cases preclude common learning of the parameter. In particular, Cripps, Ely, Mailath, and Samuelson (2013) show that common learning is precluded when the messages that the agents receive are correlated across time. Analyzing consensus and the evolution of correct higher-order beliefs for small societies when messages follow time dependence patterns remains an interesting open question.

An interesting extension of the model would be that of endogenizing the listening behavior. To follow this approach, more structure should be added to the model so as to consider that the agents pursue the maximization of a payoff that depends on the unknown parameter. Then, by characterizing listening structures that are "stable," one could obtain some insights into the formation of communication networks in a dynamic framework of belief evolution.

Appendix

Proof of Lemma 1. (a) Consider a social network Ψ. Take two different agents $i, j \in N$ and a directed link $\Psi_{ij} \in \Psi$ from agent i to agent j. To allow for alternative interpretations of the result when priors are heterogenous, we first allow agents i and j to have different priors. Using the definition of power of a directed link in (4), we have

$$\begin{aligned}
\mathbb{P}(\Psi_{ij}) &= \sum_{M} \psi_{ij}(m) D\Big(q_{ij}^{m} \,\|\, p_i\Big) \\
&= \sum_{M} \psi_{ij}(m) \sum_{\Theta} q_{ij}^{m}(\theta) \log \frac{q_{ij}^{m}(\theta)}{p_i(\theta)} \\
&= \sum_{M} \psi_{ij}(m) \sum_{\Theta} \frac{\widetilde{\psi}_{ij}^{\theta}(m) p_i(\theta)}{\psi_{ij}(m)} \log \frac{\widetilde{\psi}_{ij}^{\theta}(m)}{\psi_{ij}(m)} \\
&= \sum_{\Theta} \sum_{M} p_i(\theta) \sum_{S} \sigma_{ij}^{s}(m) \phi_j^{\theta}(s) \left[\frac{\phi_j(s)}{\phi_j[i](s)} \right] \log \frac{\sum_S \sigma_{ij}^{s'}(m) \phi_j^{\theta}(s') \left[\frac{\phi_j(s')}{\phi_j[i](s')} \right]}{\sum_S \sigma_{ij}^{s'}(m) \phi_j(s')}.
\end{aligned} \tag{10}$$

Now, by applying, for each given $\theta \in \Theta$ and each given $m \in M$, the log-sum inequality to the expression in (10) above, we obtain

$$\mathbb{P}(\Psi_{ij}) \leq \sum_{\Theta} \sum_{M} p_i(\theta) \sum_{S} \sigma_{ij}^{s}(m) \phi_j^{\theta}(s) \left[\frac{\phi_j(s)}{\phi_j[i](s)} \right] \log \frac{\phi_j^{\theta}(s)}{\phi_j[i](s)}. \tag{11}$$

On the other hand, using the definition of power of an informative signal in (3), we have

$$\begin{aligned} \mathbb{P}(\Phi_j) &= \sum_{S} \phi_j(s) D\Big(q_j^s \,\|\, p_j\Big) \\ &= \sum_{S} \phi_j(s) \sum_{\Theta} q_j^s(\theta) \log \frac{q_j^s(\theta)}{p_j(\theta)} \\ &= \sum_{S} \phi_j(s) \sum_{\Theta} \frac{\phi_j^{\theta}(s) p_j(\theta)}{\phi_j(s)} \log \frac{\phi_j^{\theta}(s)}{\phi_j(s)} \\ &= \sum_{\Theta} \sum_{S} p_j(\theta) \phi_j^{\theta}(s) \log \frac{\phi_j^{\theta}(s)}{\phi_j(s)}. \end{aligned} \tag{12}$$

Now, suppose that agents i and j have common priors, $p_i = p_j = p$. Then, $\phi_j[i] = \phi_j$ so that, by combining the inequality in (11) with the expression in (12) above, we obtain

$$\begin{aligned} \mathbb{P}(\Psi_{ij}) &\leq \sum_{\Theta} \sum_{M} p(\theta) \sum_{S} \sigma_{ij}^{s}(m) \phi_j^{\theta}(s) \log \frac{\phi_j^{\theta}(s)}{\phi_j(s)} \\ &= \sum_{\Theta} \sum_{S} p(\theta) \phi_j^{\theta}(s) \log \frac{\phi_j^{\theta}(s)}{\phi_j(s)} \Big[\sum_{M} \sigma_{ij}^{s}(m) \Big] \\ &= \sum_{\Theta} \sum_{S} p(\theta) \phi_j^{\theta}(s) \log \frac{\phi_j^{\theta}(s)}{\phi_j(s)} = \mathbb{P}(\Phi_j), \end{aligned}$$

as stated.

Moreover, by combining the expressions in equations (10) and (12) for the case of common priors, $p_i = p_j = p$, we obtain

$$\mathbb{P}(\Psi_{ij}) = \mathbb{P}(\Phi_j) + \sum_{\Theta} p(\theta) \sum_{M} \sum_{S} \sigma_{ij}^{s}(m) \phi_j^{\theta}(s) \log \frac{\phi_j(s) \sum_{S} \sigma_{ij}^{s'}(m) \phi_j^{\theta}(s')}{\phi_j^{\theta}(s) \sum_{S} \sigma_{ij}^{s'}(m) \phi_j(s')}. \tag{13}$$

Now, note that the informative message Σ_{ij}, associated with the directed link Ψ_{ij}, allows agent i to learn fully the signal that agent j observes if and only if Σ_{ij} completely separates all the signal realizations $s \in S$ available to agent j. Without loss of generality, Σ_{ij} completely separates all the signal realizations in S if and only if $\sigma_{ij}(m_l|s_l) = 1$ for each $l \in \{1, \dots, L\}$. In this case, for each $\theta \in \Theta$, we obtain

$$\sum_{M} \sum_{S} \sigma_{ij}^{s}(m) \phi_j^{\theta}(s) \log \frac{\phi_j(s) \sum_{S} \sigma_{ij}^{s'}(m) \phi_j^{\theta}(s')}{\phi_j^{\theta}(s) \sum_{S} \sigma_{ij}^{s'}(m) \phi_j(s')} = \sum_{l=1}^{L} \phi_j^{\theta}(s_l) \log \frac{\phi_j(s_l) \phi_j^{\theta}(s_l)}{\phi_j^{\theta}(s_l) \phi_j(s_l)} = 0.$$

Therefore, from the expression in (13), we obtain that the informative message Σ_{ij} allows agent i to fully learn about the signal observed by agent j if and only if $\mathbb{P}(\Psi_{ij}) = \mathbb{P}(\Phi_j)$.

(b) The proof of part (b) uses exactly the same arguments given above for part (a). The only difference is that the role played in (a) by the informative signal Φ_j is now played by the directed link Ψ_{kj}. All the formal expressions required would replicate the previous ones used in (a) upon adaptation to the appropriate formulae. Therefore, we forego a formal statement. ∎

Proof of Theorem 1. Consider a given social network Ψ and take an agent $i \in N$. For a history h_{it}, let $\alpha(s; h_{it})$ be the number of periods in which agent i has observed signal s before period t and let $\beta(m; h_{it})$ be the number of periods in which agent i has received message m from agent j (through the directed path which transmits the highest amount of information from j to i) before period t. Fix a sequence of histories $\{h_{it}\}_{t=1}^{\infty}$. Take a given $\theta \in \Theta$. Application of Bayes rule gives

$$\mu_i(\theta|h_{it}) = \left(1 + \sum_{\theta' \neq \theta} \frac{p_i(\theta')}{p_i(\theta)} \prod_S \left[\frac{\phi_i^{\theta'}(s)}{\phi_i^{\theta}(s)}\right]^{\alpha(s;h_{it})} \prod_{j \neq i} \prod_M \left[\frac{\widehat{\psi}_{ij}^{\theta'}(m)}{\widehat{\psi}_{ij}^{\theta}(m)}\right]^{\beta(m;h_{it})}\right)^{-1}.$$

Since observed frequencies approximate distributions, i.e., $\lim_{t\to\infty} \alpha(s; h_{it}) = \lim_{t\to\infty}[t\,\phi_i(s)]$ and $\lim_{t\to\infty} \beta(m; h_{it}) = \lim_{t\to\infty}[t\,\widehat{\psi}_{ij}(m)]$, we have

$$\lim_{t\to\infty} \mu_i(\theta|h_{it}) = \left[1 + \sum_{\theta' \neq \theta} \frac{p_i(\theta')}{p_i(\theta)} \lim_{t\to\infty} \left(\prod_S \left[\frac{\phi_i^{\theta'}(s)}{\phi_i^{\theta}(s)}\right]^{\phi_i(s)} \prod_{j \neq i} \prod_M \left[\frac{\widehat{\psi}_{ij}^{\theta'}(m)}{\widehat{\psi}_{ij}^{\theta}(m)}\right]^{\widehat{\psi}_{ij}(m)}\right)^t\right]^{-1}.$$

Therefore, studying the converge of $\mu_i(\theta|h_{it})$ reduces to studying whether each term, for $\theta' \neq \theta$,

$$\prod_S \left[\frac{\phi_i^{\theta'}(s)}{\phi_i^{\theta}(s)}\right]^{\phi_i(s)} \prod_{j \neq i} \prod_M \left[\frac{\widehat{\psi}_{ij}^{\theta'}(m)}{\widehat{\psi}_{ij}^{\theta}(m)}\right]^{\widehat{\psi}_{ij}(m)}$$

exceeds or not one. By taking logs, this is equivalent to studying whether, for each $\theta' \neq \theta$, the expression

$$\sum_S \phi_i(s) \log \frac{\phi_i^{\theta'}(s)}{\phi_i^{\theta}(s)} + \sum_{j \neq i} \sum_M \widehat{\psi}_{ij}(m) \log \frac{\widehat{\psi}_{ij}^{\theta'}(m)}{\widehat{\psi}_{ij}^{\theta}(m)}$$

exceeds or not zero. Then, using the definitions of G_i and of F_{ij} in (7) and in (8), respectively, we obtain that:

(i) $\lim_{t\to\infty} \mu_i(\theta|h_{it}) = 0$ if

$$G_i(\theta) + \sum_{j \neq i} F_{ij}(\theta) < G_i(\theta') + \sum_{j \neq i} F_{ij}(\theta') \text{ for some } \theta' \neq \theta \quad \Leftrightarrow \quad \theta \notin \Theta_i^*;$$

(ii) $\lim_{t\to\infty} \mu_i(\theta|h_{it}) = 1$ if

$$G_i(\theta) + \sum_{j \neq i} F_{ij}(\theta) > G_i(\theta') + \sum_{j \neq i} F_{ij}(\theta') \text{ for each } \theta' \in \Theta \setminus \{\theta\} \quad \Leftrightarrow \quad \Theta_i^* = \{\theta\};$$

(iii)

$$\lim_{t\to\infty}\mu_i(\theta|h_{it})=\left[1+\sum_{\theta'\in\Theta_i^*\setminus\{\theta\}}\frac{p_i(\theta')}{p_i(\theta)}\right]^{-1}=\frac{p_i(\theta)}{\sum_{\theta'\in\Theta_i^*}p_i(\theta')}$$

if Θ_i^* is not singleton and $\theta\in\Theta_i^*$. ∎

Proof of Theorem 2. Consider a given social network Ψ, and take two different agents $i,j\in N$ and the directed path $\widehat{\gamma}_{ij}\in\Gamma_{ij}[\Psi]$ which conveys the highest amount of information from agent j to agent i. Using the definition of power of a directed path in (6), we have:

$$\begin{aligned}
\mathbb{P}(\widehat{\gamma}_{ij}) &= \sum_M \widehat{\psi}_{ij}(m)D\Big(q_{ij}^m[\widehat{\gamma}_{ij}]\,\|\,p_i\Big)=\sum_M \psi_{ij}(m)\sum_\Theta q_{ij}^m(\theta)\log\frac{q_{ij}^m[\widehat{\gamma}_{ij}](\theta)}{p_i(\theta)}\\
&=\sum_M \widehat{\psi}_{ij}(m)\sum_\Theta \frac{\widetilde{\psi}_{ij}^\theta[\widehat{\gamma}_{ij}](m)p_i(\theta)}{\widehat{\psi}_{ij}(m)}\log\frac{q_{ij}^m[\widehat{\gamma}_{ij}](\theta)}{p_i(\theta)}\\
&=\sum_\Theta\sum_M p_i(\theta)\widetilde{\psi}_{ij}^\theta[\widehat{\gamma}_{ij}](m)\log\frac{q_{ij}^m[\widehat{\gamma}_{ij}](\theta)}{p_i(\theta)}\\
&=-\sum_\Theta p_i(\theta)\log p_i(\theta)\Big[\sum_M \widetilde{\psi}_{ij}^\theta[\widehat{\gamma}_{ij}](m)\Big]+\sum_\Theta\sum_M p_i(\theta)\widetilde{\psi}_{ij}^\theta[\widehat{\gamma}_{ij}](m)\log q_{ij}^m[\widehat{\gamma}_{ij}](\theta)\\
&=H(p_i)+\sum_M\widehat{\psi}_{ij}(m)\sum_\Theta q_{ij}^m[\widehat{\gamma}_{ij}](\theta)\log q_{ij}^m[\widehat{\gamma}_{ij}](\theta)\\
&=H(p_i)-E_{\widehat{\psi}_{ij}}\Big[H(q_{ij}^m[\widehat{\gamma}_{ij}])\Big].
\end{aligned}\tag{14}$$

Using Definition 7, it follows that agent j influences agent i if and only if the two following conditions are satisfied:

(i) $\Theta_i^*=\Theta_j^*$. This condition is satisfied if and only if for any $\theta\in\Theta_j$,

$$G_i(\theta)+\sum_{h\in N_i}F_{ih}(\theta)\geq G_i(\theta')+\sum_{h\in N_i}F_{ih}(\theta')\quad\forall\theta'\in\Theta.$$

Since we know that, for each $\theta\in\Theta_i$, $G_i(\theta)\geq G_i(\theta')$ for each $\theta'\in\Theta$, the above condition is equivalent to require that for any $\theta_j\in\Theta_j$ and any $\theta_i\in\Theta_i$

$$\begin{aligned}
&G_i(\theta_j)+\sum_{h\in N_i}F_{ih}(\theta_j)\geq G_i(\theta_i)+\sum_{h\in N_i}F_{ih}(\theta)\quad\forall\theta\in\Theta.\\
\Leftrightarrow\ &G_i(\theta_j)+\sum_{h\in N_i}F_{ih}(\theta_j)>G_i(\theta_i)+\max_{\theta\notin\Theta_j}\sum_{h\in N_i}F_{ih}(\theta).
\end{aligned}$$

By adding the identity obtained in (14) to both sides of the inequality above, we obtain the following necessary and sufficient condition for $\Theta_i^*=\Theta_j^*$:

$$\mathbb{P}(\widehat{\gamma}_{ij})>G_i(\theta_i)-G_i(\theta_j)+\max_{\theta\notin\Theta_j}\sum_{h\in N_i}F_{ih}(\theta)-\sum_{h\in N_i}F_{ih}(\theta_j)+H(p_i)-E_{\widehat{\psi}_{ij}}\Big[H(q_{ij}^m[\widehat{\gamma}_{ij}])\Big],$$

which coincides with the condition stated in (7a).

(ii) $\Theta_j^* = \Theta_j$. This condition is satisfied if and only if, for any $\theta_j \in \Theta_j$,

$$G_j(\theta_j) + \sum_{h\in N_j} F_{jh}(\theta_j) \geq G_j(\theta) + \sum_{h\in N_j} F_{jh}(\theta) \quad \forall \theta \in \Theta.$$
$$\Leftrightarrow G_j(\theta_j) + \sum_{h\in N_j} F_{jh}(\theta_j) > \max_{\theta\notin\Theta_j}\Big[G_j(\theta) + \sum_{h\in N_j} F_{jh}(\theta)\Big].$$

By adding the identity obtained in (14) (upon changing the agents' subscripts to consider $\mathbb{P}(\widehat{\gamma}_{jk})$), where $k \in N_j$, to both sides of the inequality above, we obtain the condition

$$\mathbb{P}(\widehat{\gamma}_{jk}) < G_j(\theta_j) + \sum_{h\in N_j} F_{jh}(\theta_j) - \max_{\theta\notin\Theta_j}\Big[G_j(\theta) + \sum_{h\in N_j} F_{jh}(\theta)\Big] + H(p_j) - E_{\widehat{\psi}_{jk}}\Big[H(q_{jk}^m[\widehat{\gamma}_{ik}])\Big],$$

for each $k \in N_j$, which, by rearranging terms, coincides with the condition stated. ∎

Proof of Proposition 1. First, note that application of the result in Theorem 1 to the external observer leads directly to the result that, for each history h_t, $\lim_{t\to\infty}\mu_{\text{ob}}(\theta^*|h_{it}) = 1$ if and only if $\arg\max_{\theta\in\Theta}\sum_{i\in N} G_i(\theta)$ is singleton with $\arg\max_{\theta\in\Theta}\sum_{i\in N} G_i(\theta) = \{\theta^*\}$.

Second, suppose that at consensus is achieved in the society in a way such that, for some $\theta^* \in \Theta$, we have $\lim_{t\to\infty}\mu_i(\theta^*|h_{it}) = 1$ for each history h_{it}, for each agent $i \in N$. Then, by using the result in Theorem 1, it follows that, for each agent $i \in N$,

$$G_i(\theta^*) + \sum_{j\in N_i} F_{ij}(\theta^*) \geq G_i(\theta) + \sum_{j\in N_i} F_{ij}(\theta) \quad \forall \theta \in \Theta,$$

which, by summing over all agents, implies

$$\sum_{i\in N} G_i(\theta^*) - \sum_{i\in N} G_i(\theta) \geq -\sum_{i\in N}\sum_{j\in N_i}[F_{ij}(\theta^*) - F_{ij}(\theta)]. \tag{15}$$

Therefore, provided that the consensus described above is achieved in the society, if

$$\sum_{i\in N}\sum_{j\in N_i}[F_{ij}(\theta^*) - F_{ij}(\theta)] < 0 \quad \forall \theta \in \Theta \setminus \{\theta^*\}$$

holds, then the condition in (15) above implies that $\sum_{i\in N} G_i(\theta^*) \geq \sum_{i\in N} G_i(\theta)$ for each $\theta \in \Theta$, with strict inequality if $\theta \neq \theta^*$. As a consequence, correct limiting beliefs are attained in the society. ∎

References

[1] ACEMOGLU, D., BIMPIKIS, K., AND A. OZDAGLAR (2011): "Dynamics of Information Exchange in Endogenous Social Networks," mimeo, forthcoming in *Theoretical Economics.*

[2] ACEMOGLU, D., CHERNOZHUKOV, V., AND M. YILDIZ (2009): "Fragility of Asymptotic Agreement under Bayesian Learning," mimeo.

[3] ACEMOGLU, D., COMO, D., FAGNANI, D., AND A. OZDAGLAR (2013): "Opinion Fluctuations and Disagreement in Social Networks," *Mathematics of Operations Research*, 38(1), 1-27.

[4] ACEMOGLU, D., DAHLEH, M. A., LOBEL, I., AND A. OZDAGLAR (2011): "Bayesian Learning in Social Networks," *Review of Economic Studies*, 78(4), 1201-1236.

[5] ACEMOGLU, D., OZDAGLAR, O., AND A. PARANDEHGHEIBI (2010): "Spread of (Mis)information in Social Networks," *Games and Economic Behavior*, 70, 194-227.

[6] BALA, V. AND S. GOYAL (1998): "Learning from Neighbors," *Review of Economic Studies*, 65, 595-621.

[7] BALA, V. AND S. GOYAL (2000): "A Non-Cooperative Model of Network Formation," *Econometrica*, 68(5), 1181-1229.

[8] BLACKWELL, D. AND L. DUBINS (1962): "Merging of Opinions with Increasing Information," *Annals of Mathematical Statistics*, 33, 882-887.

[9] CABRALES, A., GOSSNER, O., AND R. SERRANO (2013): "Entropy and the Value of Information for Investors," *American Economic Review*, 103, 360-377.

[10] CRAWFORD, V. P. AND J. SOBEL (1982): "Strategic Information Transmission," *Econometrica*, 50(6) 1431-1451.

[11] CRIPPS, M. W., ELY, J. C., MAILATH, G. J., AND L. SAMUELSON (2008): "Common Learning," *Econometrica*, 76(4) 909-933.

[12] CRIPPS, M. W., ELY, J. C., MAILATH, G. J., AND L. SAMUELSON (2013): "Common Learning with Intertemporal Dependence," *International Journal of Game Theory*, 42(1), 55-98.

[13] BILLINGSLEY, P. (1995): *Probability and Measure* (Third Ed.), Wiley & Sons.

[14] DEGROOT, M. H. (1974): "Reaching a Consensus," *Journal of the American Statistical Association*, 69, 345, 118-121.

[15] DEMARZO, P., VAYANOS, D., AND J. ZWIEBEL (2003): "Persuasion Bias, Social Influence, and Unidimensional Opinions," *Quarterly Journal of Economics*, 118(3), 909-968.

[16] GALEOTTI, A., GHIGLINO, C., AND F. SQUINTANI (2013): "Strategic Information Transmission in Networks," *Journal of Economic Theory*, 148(5), 1751-1769.

[17] GOLUB, B. AND M. O. JACKSON (2010): “Naïve Learning in Social Networks and the Wisdom of Crowds,” *American Economic Journal: Microeconomics*, 2(1), 112-149.

[18] HAGENBACH J. AND F. KOESSLER (2010): “Strategic Communication Networks,” *Review of Economic Studies*, 77(3), 1072-1099.

[19] HEIFETZ A. (1996): “Comment on Consensus without Common Knowledge,” *Journal of Economic Theory*, 70, 273-277.

[20] JACKSON, M. O. AND A. WOLINSKY (1996): “A Strategic Model of Social and Economic Networks,” *Journal of Economic Theory*, 71, 44-74.

[21] KOESSLER, F. (2001): “Common Knowledge and Consensus with Noisy Communication,” *Mathematical Social Sciences*, 42, 139-159.

[22] PARIKH, R., AND P. KRASUCKI (1990): “Communication, Consensus, and Knowledge,” *Journal of Economic Theory*, 52(1), 178-189.

[23] RUBINSTEIN, A. (1989): “The Electronic Mail Game: Strategic Behavior under ‘Almost Common Knowledge’,” *American Economic Review*, 79(3), 385-391.

[24] SAVAGE, L. J. (1954): *The Foundations of Statistics*, Dover reprint, New York, 1972.

[25] SCIUBBA, E. (2005): “Asymmetric Information and Survival in Financial Markets,” *Economic Theory*, 25, 353-379.

[26] SHANNON, C. E. (1948): “A Mathematical Theory of Communication,” *The Bell System Technical Journal*, 27, 379-423.

[27] STEINER, J., AND C. STEWART (2011): “Communication, Timing, and Common Learning,” *Journal of Economic Theory*, 146(1), 230-247.