Evolution of Beliefs in Networks under Bayesian Updating*

Antonio Jiménez-Martínez†

August 2012

Abstract

This paper studies the evolution of beliefs of a group of Bayesian updaters who are connected through a social network that enables them to listen to the opinions of others. Each agent observes a sequence of private signals about the value of an unknown parameter and receives private messages from others according to her connections in the network. A message conveys some information about the signal observed by the sender. 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. Messages cannot be transmitted through indirect connections in the network. We first characterize the long-run behavior of an agent's beliefs in terms of a measure that depends on the relative entropies of the conditional distributions of signals and messages available to the agent. Then, we show that the achievement of consensus in the society is closely related to the presence of prominent agents who are able to affect 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: C72, D82, D83.

*An earlier version was circulated under the title "Consensus in Communication Networks under Bayesian Updating." I am grateful to Luciana Moscoso-Boedo and an anonymous referee for extensive feedback which helped to substantially improve the paper.

†División de Economía, Centro de Investigación y Docencia Económicas.

1 Introduction

Coordinating decisions when individual payoffs depend on an unknown underlying parameter requires agents to reach similar beliefs about the parameter. To fix ideas, consider the following two-player investment game, based on an example from Cripps, Ely, Mailath, and Samuelson (2008). First, nature chooses a parameter $\theta \in \{\underline{\theta}, \overline{\theta}\}$ and each agent i assigns a prior probability p_i to $\underline{\theta}$ being the true parameter value. Then, in each period $t = 0, 1, 2, \ldots$, each agent can choose either action A, action B, or to wait until the next period. Waiting is costless. Simultaneous choices of A when the parameter is $\underline{\theta}$, or B when it is $\overline{\theta}$, give a payoff of 1 to each. Lone choices of A or B break joint investments opportunities and give a payoff of $-c$, for some $c > 0$. Joint choices that do not match the parameter give the agents a zero payoff when the parameter is $\underline{\theta}$ and a payoff of $-2c$ when the parameter if $\overline{\theta}$. Figure 1 depicts these payoffs for the choices of A and B. Suppose that, in each period t, each agent receives some additional (private) information about θ which she uses to update her beliefs. Under what circumstances do there exist equilibria of this investment game in which the agents coordinate their actions by choosing action A? Notice that action A will be optimal for an agent in some period *only if* the agent assigns probability at least $\frac{c}{1+2c} =: p$ to $\underline{\theta}$ being the true parameter value. Therefore, it is interesting to know whether the priors p_i will evolve over time so as to put eventually, both of them, at least probability p to $\underline{\theta}$ being the true parameter value. If this sufficient condition is not satisfied, then coordination in action A will fail.

	A	B
A	$1, 1$	$-c, -c$
B	$-c, -c$	$0, 0$

Parameter $\underline{\theta}$

	A	B
A	$-2c, -2c$	$-c, -c$
B	$-c, -c$	$1, 1$

Parameter $\overline{\theta}$

FIGURE 1.— Payoffs from a potential joint opportunity with actions A and B available to each agent.

In most environments, the evolution of beliefs over time depends on how agents are influenced by neighbors, friends, coworkers, local leaders, and political actors. Social networks are primary channels that carry news, information, and opinions about products, job vacancies, 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 beliefs about some common parameter of interest.

We consider 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 source and (b) there is some 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 is interpreted as the quality of the channel through which the agent receives information from some 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 weight 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 she has a directed 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. Thus, at each date, each agent receives some information about the signal that each of her neighbors is *currently* observing. Note, however, that the agent receives a message from her neighbor at each date from $t = 0$ to each given period t. Therefore, the agent obtains a stream of messages that conveys information about the stream of signals observed by her neighbor up to each given t.

The amount of information that is transmitted from the speaker to the listener depends only on the exogenously given weight of their directed connection. This weight can thus be interpreted as the quality of the channel that connects them. We further assume that the weight of each connection in the network is constant over time.[1]

At a more intuitive level, the network describes exogenously given conduits through which the agents listen to others speak about the signals that they observe. Given this framework, we ask under which conditions on the network structure will all agents eventually reach a consensus in beliefs about the parameter value. We also explore the conditions on the network under which the agents will aggregate correctly the decentralized information that they obtain from their external sources.

Building upon the model of network influence due to DeGroot (1974), a branch of the literature on social learning has recently studied how the network structure affects the transmission of opinions among connected agents. For instance, DeMarzo, Vayanos, and Zwiebel (2003) 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 group converge to the truth. They show that the attainment of limiting beliefs arbitrarily

[1]Therefore, in this paper we are not interested in the rich strategic interactions present in a sender-receiver game since we are assuming 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.

close to the true belief is characterized by the condition that the influence of the most influential agent vanishes as the size of the group grows.

In the DeGroot's model, agents update their beliefs by averaging their neighbors' beliefs according to the exogenous weights that describe the intensity of the connections between the agents. Hence, in this framework, agents are not Bayesian updaters and fail to adjust properly for repetitions and dependencies in information they hear several times. Instead, in the present paper we consider that the agents revise their beliefs, according to Bayes rule, using both the signals they receive from their external sources and the messages they hear from the agents to whom they have directed connections in the network. While our model considers that agents process the information they receive in a fully rational way, it is rigid in that the agents do not choose endogenously the weights (or intensities) of their connections and in that such weights are constant over time.

Our work is also related to the theoretical literature on common learning. The question that this line of research addresses is whether a group of agents commonly learn (at least approximately) the true parameter value as time evolves. 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' posteriors about the parameter when they start with possibly different priors and use Bayesian updating rules. In particular, we do not consider the ex-ante probabilistic assessments that the agents may make about the histories underlying their beliefs and we do not explore the evolution of the agents' higher order beliefs. In this respect, note that our notion of what constitutes similar beliefs departs from typical concepts of agreement used in the learning literature.[2]

To lay out the groundwork for our analysis, we need to address two 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 a measure appropriate to rank completely a set of available signals according to their informativeness. Following recent developments on the ranking of information value, we choose an entropy-based measure. More precisely, we use

[2]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.

the average of the relative entropy of the induced posterior (for a one-period 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 since 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 with the power of the corresponding signal and the weight of a directed connection with the power of the message associated with such a connection.[3]

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 her own private source and the information she hears from her connections in the network will 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. Notice that the aggregation of the decentralized information sources provides us with an estimate of the true parameter value which is arbitrarily accurate as the number of agents in the society tends to infinity.

Regarding our notion of correct beliefs, an important clarification is in order. We emphasize that the approach commonly pursued in models within the learning literature in which the agents observe sequences of signals (e.g., Parikh and Krasucki, 1990; Heifetz, 1996; Koessler, 2001; Steiner and Stewart, 2010; Cripps, Mailath, Ely, and Samuelson, 2008 and 2012) evaluates 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. Although this paper does make use of sequences of signals and messages, we are less interested in studying whether the agents either individually or commonly learn the true parameter value. In contrast, as in the approach pursued, following the DeGroot's model, by DeMarzo, Vayanos, and Zwiebel (2003) and by Golub and Jackson (2010), our main concern in this paper is to analyze whether the network structure allows for

[3]Entropy-based measures have been successfully applied to propose informativeness orderings. The power of a signal is used, for instance, by Sciubba (2005) to rank information in her work on survival of traders in financial markets under asymmetric information. In a recent paper, Cabrales, Gossner, and Serrano (2012) propose, for a class of no-arbitrage investment problems under ruin-averse preferences, an entropy-based measure that coincides with the power measure. Their measure is such 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. Thus, the power measure induces a complete order over signals for at least an interesting class of decision problems.

the aggregation of the decentralized sources of information possessed by the agents. We consider sequences of signals and messages not because we are interested in individual and/or common learning questions but simply because we wish to introduce an ingredient into the DeGroot's model that allows for Bayesian updating of beliefs.

Suppose that we defined instead correct beliefs by conditioning posteriors on a given (true) parameter value. Then, the results provided by Cripps, Mailath, Ely, and Samuelson (2008) would imply for our model that the agents commonly learn the true parameter value, regardless of the network structure. This is due to our assumptions that (a) the sets of signals and messages are finite and that (b) both signals and messages are independent across time. Nevertheless, as the society becomes large, the aggregation of the private signals constitutes an arbitrarily accurate estimate of the true parameter value. In addition, if we consider that the number of agents is sufficiently large, then the assumption that the sets of signals and messages are finite becomes less compelling. Thus, if we drop this assumption, it is no longer clear whether the agents commonly learn the true parameter value. In fact, the argument given by Rubinstein (1989) in his celebrated email game suggests that common learning of the true parameter is prevented with arbitrarily large sets of signals and messages. Thus, our approach to analyze correctness of beliefs seems an appropriate one when one focuses on large enough societies.

Our main results begin with a simple but complete characterization of an agent's limiting beliefs, in Proposition 1. In general, an agent's beliefs converge to some beliefs that favor (with probability one) one particular parameter value. This convergence is determined by both the informativeness of the private source available to the agent and the influence that she receives from her directed connections with other agents. The role of the information that the agent obtains from the private source and from her neighbors can be neatly described in terms of the the relative entropies of the (conditional) distributions of signals and messages available to her in the network. Then, using this result, we identify a necessary condition for the achievement of a consensus in the society, in Corollary 1. This condition, which we label as *Inter Group Connectedness (IGC)*, requires that each agent be connected with others in the network in a way such that the evolution of her beliefs can be altered by listening to her neighbors. This condition involves both the intensity of the agent's connections with her neighbors and the nature of the information that the neighbors receive from their sources. Intuitively, a consensus is achieved only if the are some prominent agents who can influence others so as to change their minds over time.

In Proposition 3, we provide a sufficient condition on the levels of informativeness of the directed connections in the network under which, provided that there is a consensus, the agents' limiting beliefs aggregate correctly the information initially available to all of them. 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 affect the evolution of beliefs that one obtains by aggregating the information initially held by all agents. Thus, by combining the necessary condition in Corollary 1 with the sufficient condition in Proposition 3, 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 by the prominent agents needs to be bounded and must not be able to manipulate the beliefs that obey to the aggregation of all the agents' private sources. In short, to attain consensus and correct limiting beliefs, a certain level of popularity is a bless, a disproportionate popularity is a curse.

The present paper relates to several lines of research on influence in networks other than the one that stems from the DeGroot's model. 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 consider consensus specifically, our model allows for insights with a similar flavor since some spread of beliefs among agents with different opinions is required for consensus. In our model, consensus can be prevented when an agent does not listen enough to agents with different opinions and, at the same time, is listened by others. Such an agent would play a similar role to a "forceful" agent in their model. The question of whether consensus is attained under a non-Bayesian updating rule is analyzed by Acemoglu, Como, Fagnani, and Ozdaglar (2010). 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 her to change her opinion over time, e.g., if the IGC condition is not satisfied.

Another branch 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. An important contribution within these models of observational learning is the work of Bala and Goyal (1998), in which the 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 her neighbors are earning payoffs different from her own. In addition, 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.

Finally, we note that existing models on communication and learning, based on Bayesian and non-Bayesian updating rules, typically lead to consensus when communication takes place over a strongly connected network (e.g., Acemoglu, Dahleh, Lobel and Ozdaglar, 2008; Bala and Goyal, 1998; DeMarzo, Vayanos and Zwiebel, 2003, Golub and Jackson, 2004; Acemoglu, Ozdaglar and ParandehGheibi, 2009). A strongly connected network need not satisfy condition IGC in our model, which could preclude consensus. This difference is explained by the fact that, in our set up, the information contained in the messages does not flow in *any* period through indirect connections in the network. This observation is important since it shows that restricting communication to flow over time only locally in the social network (i.e., within each pair of directly connected neighbors) interferes with the spread of information so as to prevent consensus. Nevertheless, the IGC condition seems relevant to identify influential agents in the presence of restrictions for the transmission of information through indirect connections in the network.

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

2 The Model

For any set X, $\Delta(X)$ denotes the set of all Borel probability measures on X.

2.1 Information Structure

Time is discrete and indexed by $t = 0, 1, 2, \ldots$. There is a finite set of agents $N := \{1, 2, \ldots, n\}$, with $n \geq 3$, who are connected through an exogenous directed social network. The agents care about a parameter $\theta \in \Theta := \left\{\underline{\theta}, \overline{\theta}\right\}$, which is selected by nature before period zero. Each agent i has a (subjective) prior distribution $p_i \in \Delta(\Theta)$ that describes her ex-ante beliefs about the parameter. We assume that $p_i(\underline{\theta}) \in (0, 1)$ for each $i \in N$. The realized parameter θ is not observed directly by any agent. Instead, in each period t, each agent i observes privately a *signal realization* $s \in S := \{\underline{s}, \overline{s}\}$ and receives privately a *message realization* $m \in M := \{\underline{m}, \overline{m}\}$ from each agent $j \in N$. When convenient, we will use further notation to provide more details about signals and messages. Specifically, we will sometimes find useful to use s_{it} to denote the signal received by agent i in period t and to use m_{ijt} to denote the message received by agent i from agent j in period t. Also, when no reference need be made to time periods, we will simply use s_i and m_{ij}, respectively.

As it will be explained in more detail below, a message received from some agent in some period t conveys information only about the private signal that such an agent observes at t. In principle, we allow each agent to receive messages from herself as well. However, since an agent already observes her own signals, the information conveyed by such messages is redundant. Also, even though an agent receives messages from all agents, the social network restricts the amount of information that she receives from such messages. The constraints that the network imposes on the information that the agents receive from others are described more precisely in the next subsection. A message vector received by agent i in period t is denoted by $\boldsymbol{m}_{it} := (m_{ijt})_{j\in N} \in M^n$, a *message profile* in period t is denoted by $\boldsymbol{m}_t := (\boldsymbol{m}_{it})_{i\in N} \in M^{n^2}$, and a *signal profile* in period t is denoted by $\boldsymbol{s}_t := (s_{it})_{i\in N} \in S^n$.

Both the signal profile $\boldsymbol{s}_t$ and the message profile $\boldsymbol{m}_t$ are independent and identically distributed across periods, conditional on the parameter. For each period t, the distribution over signals, conditional on the parameter value, observed by agent i is denoted by $\phi_i(\cdot \mid \theta)$. We use ϕ_i to denote the corresponding unconditional distribution over signals. Also, for each pair of agents i and j, and for each period t, the distribution over messages observed by agent i, conditional on the signal realization observed by agent j, is denoted by $\sigma_{ij}(\cdot \mid s)$. We use σ_{ij} to denote the corresponding unconditional distribution over messages.

An *informative signal* (associated with an external source) for agent i is a pair of conditional distributions over signals $\Phi_i := \{\phi_i(\cdot \mid \theta) \in \Delta(S) \,:\, \theta \in \Theta\}$. An *informative message*[4] from agent j to agent i (associated with the direct directed connection from agent i to agent j in the network) is a pair of conditional distributions over messages $\Sigma_{ij} := \{\sigma_{ij}(\cdot \mid s) \in \Delta(M) \,:\, s \in S\}$. Each Σ_{ij} is exogenously given and constant across periods.[5]

An informative message Σ_{ij} allows agent i to update her beliefs about θ by observing the sequence of messages that she receives from agent j. The distribution over the messages received by agent i from agent j, conditional on the parameter value, is denoted by $\psi_{ij}(\cdot \mid \theta)$. We use ψ_{ij} to denote the corresponding unconditional distribution. A (direct) *directed link from agent i to agent j* is a pair of conditional distributions over messages $\Psi_{ij} := \{\psi_{ij}(\cdot \mid \theta) \in \Delta(M) \,:\, \theta \in \Theta\}$.

Thus, each agent learns about the value of the parameter θ not only by observing her own sequence of signals but also by obtaining some information about the signals observed by the agents to whom she has a directed link. One way to interpret communication in this model is by considering that each agent i listens, with some exogenous and fixed degree of informativeness, to the opinions about θ that each other agent forms from her own private signals.

Of course, for any two agents $i, j \in N$, Ψ_{ij}, Σ_{ij}, and Φ_j are related through a consistency

[4]Note that we use the terms informative signal and informative message to refer to sets of conditional distributions while signals and messages are realizations of the corresponding random variables.

[5]Modeling communication between pair of agents connected in a network by means of a sender-receiver protocol has been recently considered, among others, by Hagenbach and Koessler (2010). Nevertheless, in contrast with a typical sender-receiver game, in our model an agent is not able to choose the amount of information that she reveals. Instead, information transmission from agent j to agent i is exogenously given by the degree of informativeness of Σ_{ij}.

requirement imposed by the total probability rule. Specifically,

$$\psi_{ij}(m \mid \theta) = \sum_{s \in S} \sigma_{ij}(m|s)\phi_j(s|\theta), \quad \forall \theta \in \Theta, \quad \forall m \in M. \tag{1}$$

We introduce now some information theory concepts that will be useful to analyze the transmission of information in our model. We begin with the definition of entropy of a distribution.

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

$$H(P) := -\sum_{x \in X} P(x) \log P(x).$$

The (Shannon) 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 our model, the entropy of the agents' priors will provide us with an upper bound on the degree of informativeness of the sets of informative signals and informative messages available to any agent.

To measure how informative are 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

$$D(P \,||\, Q) := \sum_{x \in X} P(x) \log \frac{P(x)}{Q(x)}.$$

The relative entropy is not a metric,[6] but, considering X as a sample space, it provides us with a useful measure of the gain of information in moving from Q to 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. We consider only a single-period Bayesian revision process.[7] Take an agent $i \in N$ and suppose that she receives information about the parameter only from her private source in the form of the informative signal Φ_i. For a one-period revision process, let $p_i(\cdot \mid s)$ denote agent i's posteriors about θ upon observing signal s (i.e., these are posteriors obtained solely from the information that the agent receives from her external source). Thus, using Bayes' rule, we have

$$p_i(\theta \mid s) = \frac{\phi_i(s|\theta)p_i(\theta)}{\sum_{\theta' \in \Theta} \phi_i(s|\theta')p_i(\theta')} \quad \forall \theta \in \Theta, \quad \forall s \in S.$$

We define the *power of the informative signal* Φ_i as the expectation of the relative entropy of the (one-period) posterior with respect to the prior.

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

[7] Recall that the distributions over signals and messages are constant over time.

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

$$\mathbb{P}(\Phi_i) := \sum_{s \in S} \phi_i(s) D\big(p_i(\cdot \,|\, s) \,||\, p_i\big). \tag{2}$$

The power measure allows us to rank completely any set of informative signals according to their degree of informativeness so that Φ_i is at least as informative as Φ_i' whenever $\mathbb{P}(\Phi_i) \geq \mathbb{P}(\Phi_i')$. Analogously, suppose that agent i receives information about the parameter only from the private messages sent by another agent j, in the form of the directed link Ψ_{ij}. For a one-period revision process, let $q_{ij}(\cdot \,|\, m)$ denote agent i's posteriors about θ upon receiving message m from agent j (i.e., these are posteriors obtained solely from the information that agent i receives from agent j in the network). Then, using Bayes' rule, we have

$$q_{ij}(\theta \,|\, m) = \frac{\psi_{ij}(m|\theta) p_i(\theta)}{\sum_{\theta' \in \Theta} \psi_{ij}(m|\theta') p_i(\theta')} \quad \forall \theta \in \Theta, \quad \forall m \in M.$$

We define the *power of the directed link* Ψ_{ij} as the expectation of the relative entropy of the (one-period) posterior with respect to the prior.

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

$$\mathbb{P}(\Psi_{ij}) := \sum_{m \in M} \psi_{ij}(m) D\big(q_{ij}(\cdot \,|\, m) \,||\, p_i\big). \tag{3}$$

We note that the informativeness of the directed link from agent i to agent j does not exceed the informativeness of agent j's own signals. In other words, suppose that agent i is allowed to use agent j's private signal to update her beliefs about the parameter. Then, the information about θ that agent i receives through her directed link to agent j is naturally less precise than the information that she obtains using directly agent j's signal. Thus, our model captures the presence of some (exogenous) decay in the transmission of information, which is associated to informative messages that do not fully transmit the private information available to the sender. When the informative message Σ_{ij} is such that agent i fully learns the signal that agent j observes, then, using such a directed link, agent i learns as much as agent j about the parameter. Only in this case there is no loss of information through the link. The following Lemma provides this intuitive result in terms of the power measure.

Lemma 1 (**Decay in the Transmission of Information).** *Suppose that agents $i, j \in N$ use the same informative signal Φ to obtain information about θ. Then, $\mathbb{P}(\Psi_{ij}) \leq \mathbb{P}(\Phi)$ for each directed link Ψ_{ij}. Moreover, $\mathbb{P}(\Psi_{ij}) = \mathbb{P}(\Phi)$ 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.*

It can be verified that, for each $i \in N$, $\mathbb{P}(\Phi_i) = H(p_i) - \sum_{s \in S} \phi_i(s) H(p_i(\cdot|s))$ so that $\mathbb{P}(\Phi_i) \leq H(p_i)$. In addition, we have that $\mathbb{P}(\Phi_i) = H(p_i)$ if and only if the average entropy of agent i's posteriors (obtained only from her private source) vanishes. In other words, $\mathbb{P}(\Phi_i) = H(p_i)$ whenever agent i obtains full information about the parameter from her private signal. Note that for agent i to obtain full information about the parameter from her 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 her own informative signal (i.e., $\mathbb{P}(\Phi_j) = H(p_j)$). Therefore, from the result in Lemma 1, we observe that $\max_{i \in N} H(p_i)$ imposes an upper bound on the degree of informativeness about θ that any agent in the society can obtain, regardless of the network structure.

2.2 Communication through the Network

The power measure allows us to rank completely any set of directed links according to their degree of informativeness so that Ψ_{ij} is at least as informative as Ψ'_{ij} whenever $\mathbb{P}(\Psi_{ij}) \geq \mathbb{P}(\Psi'_{ij})$. Using this, we identify the *weight of the directed link* from agent i to agent j with the power $\mathbb{P}(\Psi_{ij})$ of such a link. Notice that $\mathbb{P}(\Psi_{ij}) = 0$ if and only if agent i learns nothing about the parameter from the informative message received from agent j. We interpret this as agent i not listening to the opinions of agent j about parameter θ and model this situation as agent i not having a directed link to agent j (or, in other words, as having a directed link with zero weight). On the opposite extreme, it follows from Lemma 1 that $\mathbb{P}(\Psi_{ij}) \leq \mathbb{P}(\widehat{\Phi}) := \max_{i \in N} \mathbb{P}(\Phi_i)$. Then, using the power measure, we define a *directed weighted network* as an $n \times n$ matrix of weights $\mathbb{P}(\Psi) := \left[\, \mathbb{P}(\Psi_{ij}) \,\right]$, where $0 \leq \mathbb{P}(\Psi_{ij}) \leq \mathbb{P}(\widehat{\Phi})$ indicates the weight of the directed link from agent i to agent j. We set $\mathbb{P}(\Psi_{ii}) = \mathbb{P}(\Phi_i)$ for each $i \in N$ so that an agent transmits full information to herself. As mentioned earlier, this information is redundant since agent i already observes her own signals.

Agents can also be connected through indirect connections in a network. Nevertheless, we assume in this paper that information cannot be transmitted through indirect connections. Thus, each agent can hear only to the opinions of the agents to whom she has a direct (directed) link. A *directed path* from agent i_1 to agent i_K is a sequence of directed links $\gamma_{i_1 i_K} := (\Psi_{i_1 i_2}, \Psi_{i_2 i_3}, \ldots, \Psi_{i_{K-1} i_K})$ such that $\mathbb{P}(\Psi_{i_{k-1} i_k}) > 0$ for each $k \in \{2, \ldots, K\}$. A weighted network $\mathbb{P}(\Psi)$ is *strongly connected* if there is a directed path from any agent to any other agent.

2.3 Evolution of Beliefs, Consensus, and Correct Beliefs

We introduce a few additional concepts which are useful to analyze the evolution of the agents' beliefs. A *period-t history for agent i* is a sequence $h_{it} := ((s_{i0}, \boldsymbol{m}_{i0}), (s_{i1}, \boldsymbol{m}_{i1}), \ldots, (s_{it}, \boldsymbol{m}_{it})) \in (S \times M^n)^t$ of signals and message vectors. The posterior belief of agent i about parameter θ

in each period t is given by the random variable $\mu_i(\theta \,|\, h_{it}) : \Omega \to [0,1]$. For each agent i and each value of the parameter θ, the sequence of random variables $\{\mu_i(\theta \,|\, h_{it})\}_{t=0}^{\infty}$ is a bounded martingale,[8] which ensures that the agents' posterior beliefs converge almost surely (see, e.g., Billingsley, 1995, Theorem 35.5).

Definition 5. A *consensus is (asymptotically) reached 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 each $\theta \in \Theta$,

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

for some $p \in [0,1]$.

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 can access to the external sources available to all agents in the society. The observer's priors are given by a distribution $p \in \Delta(\Theta)$. A *period-t history for the external observer* is a sequence $h_t := (\boldsymbol{s}_0, \boldsymbol{s}_1, \dots, \boldsymbol{s}_t) \in (S^n)^t$ of signal profiles. The posterior belief of the external observer about parameter θ in each period t is given by the random variable $\mu_{\text{ob}}(\theta \,|\, h_t) : \Omega \to [0,1]$. Again, for each value of the parameter θ, the sequence of random variables $\{\mu_{\text{ob}}(\theta \,|\, h_t)\}_{t=0}^{\infty}$ is a bounded martingale so that the external observer's posteriors converge almost surely. 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. For large enough societies, the observer's limiting beliefs are arbitrarily accurate estimates of the true parameter value.

Definition 6. The directed network $\mathbb{P}(\Psi)$ attains *correct limiting beliefs* if a consensus is achieved in the society and, in addition, for each $i \in N$, and each $\theta \in \Theta$,

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

3 Results

3.1 Characterizing Limiting Beliefs

Given a value $\theta \in \Theta$ of the parameter, we will use from here onwards θ' to denote the other value of the parameter, i.e., $\{\theta'\} := \Theta \setminus \{\theta\}$. We use the relative entropies between the conditional distributions over both signals and messages to construct a set of measures that will be useful

[8] More formally, $\{\mu_i(\theta \,|\, h_{it})\}_{t=0}^{\infty}$ is a bounded martingale with respect to the measure on Ω, conditional on θ, induced by the priors $(p_i)_{i\in N}$, and the conditional distributions $\phi_i(\cdot \,|\, \theta)$, $\psi_{ij}(\cdot \,|\, \theta)$, $i, j \in N$.

to study the convergence of the agents' posteriors. For an agent $i \in N$, let

$$G_i(\theta||\theta') := p_i(\theta) D\big(\phi_i(\cdot|\theta) \,||\, \phi_i(\cdot|\theta')\big). \tag{4}$$

Also, for a pair of agents $i, j \in N$, let

$$F_{ij}(\theta||\theta') := p_i(\theta) D\big(\psi_{ij}(\cdot|\theta) \,||\, \psi_{ij}(\cdot|\theta')\big). \tag{5}$$

Intuitively, $G_i(\theta||\theta')$ measures the weight that agent i's private signal puts on i's posteriors (after a one-period revision) that θ is the true parameter instead of θ'. Analogously, $F_{ij}(\theta||\theta')$ measures the weight that the information transmitted from agent j to agent i puts on i's posteriors (for a one-period revision) that θ is the true parameter instead of θ'. Then, using the set of measures G_i and F_{ij} $(j \neq i)$, we propose a measure of the weight that both the private signal and the information received from her neighbors put on an agent's posteriors that the true parameter is a given value (against the alternative value).

Definition 7. For a directed network $\mathbb{P}(\Psi)$, the influence measure for agent i of parameter value θ with respect to parameter value θ' is

$$\xi_i(\theta||\theta') := G_i(\theta||\theta') + \sum_{j \neq i} F_{ij}(\theta||\theta').$$

Note that both influence measures $\xi_i(\underline{\theta}||\overline{\theta})$ and $\xi_i(\overline{\theta}||\underline{\theta})$ are nonnegative. The next proposition shows that the convergence of an agent's posteriors is characterized by the aggregation of the weights that her private signal and all her directed links place on a given parameter value being the true parameter. This aggregation of influences can be neatly expressed in terms of the influence measure defined above.

Proposition 1. *For a given weighted network $\mathbb{P}(\Psi)$ and for any sequence of histories $\{h_{it}\}_{t=0}^{\infty}$, agent i's limiting beliefs satisfy (i) $\lim_{t\to\infty} \mu_i(\theta|h_{it}) = 1$ if and only if $\xi_i(\theta||\theta') > \xi_i(\theta'||\theta)$ and (ii) $\lim_{t\to\infty} \mu_i(\theta|h_{it}) = p_i(\theta)$ if and only if $\xi_i(\theta||\theta') = \xi_i(\theta'||\theta)$.*

How do the levels of informativeness of signals and of directed links relate with the agents' limiting beliefs? What is the influence of the agents' priors on their limiting beliefs? The following lemma establishes useful relationships between the power of both informative signals and directed links in the network, on the one hand, and the measures G_i and F_{ij} that characterize the evolution of the agents' beliefs, on the other hand.

Lemma 2. *For each set of informative signals $\{\Phi_i\}_{i\in N}$ and each weighted network $\mathbb{P}(\Psi)$,*

(i) $\mathbb{P}(\Phi_i) - G_i(\theta'||\theta) \leq p_i(\theta')\big[G_i(\theta||\theta') - G_i(\theta'||\theta)\big]$;

(ii) $\mathbb{P}(\Psi_{ij}) - F_{ij}(\theta'||\theta) \leq p_i(\theta')\big[F_{ij}(\theta||\theta') - F_{ij}(\theta'||\theta)\big]$ for each $j \neq i$.

Take a given agent $i \in N$. By combining the results (i) and (ii) of Lemma 2 above with the definition of influence measure in Definition 1, we obtain

$$0 \leq \mathbb{P}(\Phi_i) + \sum_{j \neq i} \mathbb{P}(\Psi_{ij}) \leq p_i(\theta')\big[\xi_i(\theta||\theta') - \xi_i(\theta'||\theta)\big] + \xi_i(\theta'||\theta), \tag{6}$$

which, in turn, can be rewritten as

$$0 \leq \mathbb{P}(\Phi_i) + \sum_{j \neq i} \mathbb{P}(\Psi_{ij}) \leq p_i(\theta')\xi_i(\theta||\theta') + p_i(\theta)\xi_i(\theta'||\theta). \tag{7}$$

Note that the inequalities in (6) and (7) above specify, in different terms, an upper bound on the aggregate level of informativeness that agent i obtains from the external source and from her links in the network. Interestingly enough, such an upper bound is related with the influence measures that characterize the convergence of the agent's posteriors. Suppose for instance that, for the given agent i, $\lim_{t\to\infty} \mu_i(\underline{\theta}|h_{it}) = 1$ so that $\xi_i(\underline{\theta}||\overline{\theta}) > \xi_i(\overline{\theta}||\underline{\theta})$. Then, from the expression in (6), we observe that $p_i(\overline{\theta})\big[\xi_i(\underline{\theta}||\overline{\theta}) - \xi_i(\overline{\theta}||\underline{\theta})\big] + \xi_i(\overline{\theta}||\theta) > 0$ imposes an upper bound on agent i's aggregate level of informativeness. In this case, other things equal, lower values of $p_i(\overline{\theta})$ requires lower levels of aggregate informativeness for agent i to attain limiting posteriors that put probability one to $\underline{\theta}$ being the true parameter value. This conveys the intuitive message that, in order to end up believing that a certain parameter value is the true one, the agent requires less precise channels of information when her priors put low probability to the alternative value. In addition, the inequality in (7) reflects that, in order to restrict the size of the required aggregate level of informativeness for the agent, priors and influence measures interact in opposite directions. This is also intuitive. Observe that a situation with a low prior on a given parameter value θ and a high influence measure that favors the alternative parameter value θ' both increase the required upper bound. As a consequence, in this situation, high levels of aggregate informativeness for the agent are compatible with limiting beliefs that put probability one to the parameter value θ.

3.2 A Necessary Condition for Consensus

The achievement of consensus in the society requires that the evolution of some agents' posteriors can be substantially influenced from their communication with other agents. Roughly speaking, consensus requires that some agents change their minds over time by listening to other agents. From our results in Proposition 1, we observe that the measures $F_{ij}(\underline{\theta}||\overline{\theta})$ and $F_{ij}(\overline{\theta}||\underline{\theta})$ formalize in our model the extent to which agent i's posteriors are affected in the long run by the information that agent i receives from agent j.

It is useful for our analysis to consider first a benchmark case where each agent is isolated and receives no information from any of the other agents. Let $\mathbb{P}(\widehat{\Psi})$ denote the directed network specified as

$$\mathbb{P}(\widehat{\Psi}) := \begin{pmatrix} \mathbb{P}(\Phi_1) & 0 & \dots & 0 \\ 0 & \mathbb{P}(\Phi_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathbb{P}(\Phi_n) \end{pmatrix}.$$

The network $\mathbb{P}(\widehat{\Psi})$ describes a situation where each directed link has a zero weight so that there is no communication among the agents. In this case, the evolution of each agent's posteriors is determined solely by the information that she receives from her source. For the network $\mathbb{P}(\widehat{\Psi})$, application of Proposition 1 implies that $\lim_{t\to\infty} \mu_i(\theta|h_{it}) = 1$ if and only if $G_i(\theta||\theta') > G_i(\theta'||\theta)$. Then, for the case in which there is no communication among the agents, the set $N(\theta) := \{i \in N : G_i(\theta||\theta') > G_i(\theta'||\theta)\}$ gives us the set of agents whose limiting posteriors put probability one to θ being the true parameter value.

For simplicity, we assume that each agent's private signal is such that, for the case in which there is no communication among the agents, her posteriors converge so as to put probability one to either $\underline{\theta}$ or $\overline{\theta}$. Specifically, we assume that there is no agent i such that $G_i(\theta||\theta') = G_i(\theta'||\theta)$ so that $N = N(\underline{\theta}) \cup N(\overline{\theta})$. In other words, we ignore situations in which an agent's posteriors evolve in a way such that she finally ends up with her own initial priors. While this assumption is useful for presentation purposes, it plays no relevant role in the mechanisms that governs the model.[9] In addition, to make the problem interesting, we assume that $N(\underline{\theta}) \neq \emptyset$ and $N(\overline{\theta}) \neq \emptyset$. Thus, for the benchmark case without communication, we are splitting the society into two nonempty sets of agents, those whose long run posteriors favor $\underline{\theta}$, $N(\underline{\theta})$, and those whose long run posteriors favor of $\overline{\theta}$, $N(\overline{\theta})$.

We identify in the corollary below a key necessary condition for consensus in the society. Consensus requires that each of the agents who would favor a given parameter value, had she received no information from others, has a directed link to at least one agent whose information is able to change the evolution of her posteriors over time so as to eventually put probability one on the alternative value being the true one. We label this condition as *Inter Group Connectedness (IGC)*.

Definition 8. The directed network $\mathbb{P}(\Psi)$ satisfies *Inter Group Connectedness (IGC)* if either (i) each agent $i \in N(\underline{\theta})$ has a directed link to at least one agent j such that $F_{ij}(\overline{\theta}||\underline{\theta}) > F_{ij}(\underline{\theta}||\overline{\theta})$, or (ii) each agent from $k \in N(\overline{\theta})$ has a directed link to at least one agent from l such that $F_{kl}(\underline{\theta}||\overline{\theta}) > F_{kl}(\overline{\theta}||\underline{\theta})$, or both.

Corollary 1. *Suppose that a consensus is reached in a society described by a directed network $\mathbb{P}(\Psi)$. Then, the network $\mathbb{P}(\Psi)$ necessarily satisfies IGC.*

[9] Using standard genericity arguments, it can be shown that this assumption is generically satisfied for societies with a finite number of agents.

The corollary above follows directly from Proposition 1. Consider a directed network $\mathbb{P}(\Psi)$ which does not satisfy IGC. Then, there is some agent $i \in N(\underline{\theta})$ who does not have a directed link to any agent j such that $F_{ij}(\overline{\theta}||\underline{\theta}) > F_{ij}(\underline{\theta}||\overline{\theta})$. Also, there is some agent $k \in N(\overline{\theta})$ who does not have a directed link to any agent l such that $F_{kl}(\underline{\theta}||\overline{\theta}) > F_{kl}(\overline{\theta}||\underline{\theta})$ As a consequence, $\xi_i(\underline{\theta}||\overline{\theta}) > \xi_i(\overline{\theta}||\underline{\theta})$ and $\xi_k(\overline{\theta}||\underline{\theta}) > \xi_k(\underline{\theta}||\overline{\theta})$. From Proposition 1, it follows that $\lim_{t\to\infty} \mu_i(\underline{\theta} \,|\, h_{it}) = 1$ while $\lim_{t\to\infty} \mu_k(\overline{\theta} \,|\, h_{kt}) = 1$ so that a consensus is not attained.

In short, the achievement of consensus is closely related to the presence of prominent agents in the society. Through the links directed to them, these agents must have the ability to change the way in which some their listeners' beliefs evolve.

We end this subsection by noting that a strongly connected network need not satisfy the IGC condition. This is simply due to the fact that we assume that information cannot be transmitted through indirect connections in the network. Thus, we can have a strongly network in which some agent i is connected to another agent j only through a directed (indirect) path $\gamma_{ij} = (\Psi_{ii_1}, \Psi_{i_1 i_2}, \ldots, \Psi_{i_{K-1} j})$. The path γ_{ij} does not allow agent i's posteriors to be affected by any information available to agent j. Consequently, agent i may end up with posteriors different from agent j's limiting posteriors. In other words, a consensus may not be attained in a strongly connected network. This implication contrasts with the existing results in most of the related literature. Yet, we emphasize that this result is directly driven in the present paper by the assumption that information does not flow through indirect connections in any period.

3.3 Correct Limiting Beliefs and Influence of Prominent Agents

The next proposition characterizes the external observer's limiting beliefs. Recall that we assume that the external observer has access to the private sources available to all agents in the society. In this sense, the observer's posteriors constitute our aggregate of the information possessed by the agents. Our framework for studying the correctness of beliefs considers that the observer's beliefs do not depend on any particular network structure. Not surprisingly, the aggregation of the measures $G_i(\theta||\theta')$ across all agents play a key role in the observer's limiting beliefs.

Proposition 2. *For any sequence of histories $\{h_t\}_{t=0}^{\infty}$, the external observer's limiting beliefs satisfy (i)* $\lim_{t\to\infty} \mu_{\text{ob}}(\theta|h_t) = 1$ *if and only if* $\sum_{i\in N} G_i(\theta||\theta') > \sum_{i\in N} G_i(\theta'||\theta)$ *and (ii)* $\lim_{t\to\infty} \mu_{\text{ob}}(\theta|h_t) = p(\theta)$ *if and only if* $\sum_{i\in N} G_i(\theta||\theta') = \sum_{i\in N} G_i(\theta'||\theta)$.

The next proposition provides a sufficient condition on the levels of informativeness associated with 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 3. *Consider a directed network* $\mathbb{P}(\Psi)$ *and suppose that a consensus is reached in the society so that, for each* $i \in N$*, we have* $\lim_{t\to\infty} \mu_i(\theta|h_{it}) = 1$ *for some parameter value* θ*. If the following condition is satisfied*

$$\sum_{i\in N}\sum_{j\neq i}\left[F_{ij}(\theta'||\theta) - F_{ij}(\theta||\theta')\right] > 0,$$

with $\{\theta'\} := \Theta \setminus \{\theta\}$*, then the directed network* $\mathbb{P}(\Psi)$ *attains correct limiting beliefs in which* $\lim_{t\to\infty} \mu_{\text{ob}}(\theta|h_t) = 1$*.*

The sufficient condition identified in Proposition 3 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 the alternative parameter value θ'. The message conveyed by this result 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 from the one proposed here,[10] 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.

From our results, we observe that the presence of prominent agents is desirable to achieve consensus. Yet, the influence of the prominent agents must not be too high so as to change the agents' beliefs in the opposite direction with respect to what one learns by aggregating the agents' private sources of information.

4 Concluding Comments

We have assumed that there are only two possible values of the parameter. Nevertheless, the intuition underlying the results in Corollary 1 and in Proposition 3 is compelling and general. With finitely many parameter values, the IGC condition would generalize by requiring that each agent who would favor a given parameter value, had she received no information from others, has a connection with some given prominent agent who is able to change the evolution of her beliefs over time so as to end up favoring the same parameter value favored by the prominent agent. With more that two parameter values, the achievement of consensus is closely related to the presence of at least one prominent agent who favors in the long run a given parameter value.

[10] Their definition of correct beliefs requires that some external observer aggregates the pieces of information initial held by the agents, as we do. However, in contrast with our notion of correct beliefs, they also consider that the size of the society grows arbitrarily so that their definition is based on an asymptotic criteria as the number of agents tends to infinity.

In addition, this prominent agent must be able to spread her opinion to all agents in the society. Of course, this can be achieved also under the presence of several prominent agents, provided that they favor in the long run a common parameter value. If a consensus is achieved with more than two parameter values, then the agents' limiting beliefs will aggregate the decentralized sources of private information only if there is no prominent agent who favors in the long run a different parameter value and whose influence is too large. We would require that the influence of such a prominent agent be not so large so as to deviate the agents' limiting posteriors from the ones which obey to the aggregation of the information initially held by all the agents. In this sense, the message conveyed by Proposition 3 continues to hold with more that two parameter values.

As we discussed in the Introduction, our notion of correctness of beliefs is more compelling when one focuses on societies large enough. For small societies, using a definition of correct beliefs based on conditioning posteriors on a given parameter value would deliver the message that the agents always learn the truth. 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, 2010) 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 (2012) show that common learning is precluded when the messages that the agents receive are correlated across time. The present paper considers that messages are independent over time. Analyzing consensus and the evolution of correct beliefs for small societies when messages follow time dependence patterns remains and interesting an open question.

Another 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.

Finally, the assumption that messages are not transmitted through indirect connections seems more realistic in some environments than in others. For example, it is a natural assumption in networks within formal organizations where regulation restricts the transmission of information to indirectly connected members. This would be also the case in networks with physical restrictions to the flow of indirect information. This assumption, however, is less compelling in informal networks. It would be interesting to analyze the achievement of consensus and correct beliefs when this assumption is relaxed and some amount of information is allowed to flow through indirect connections.

Appendix

Proof of Lemma 1. Take two agents $i, j \in N$, and consider a directed link Ψ_{ij} from agent i to agent j. Suppose that agents i and j have the same informative signal Φ. Using the definition of power of a directed link in (3) and the relation between Ψ_{ij} and Σ_{ij} given by the consistency requirement in (1), we obtain

$$\begin{aligned}
\mathbb{P}(\Psi_{ij}) &= \sum_{m\in M} \psi_{ij}(m) D\big(q_{ij}(\cdot|m) \,||\, p_i\big) \\
&= \sum_{m\in M} \psi_{ij}(m) \sum_{\theta\in\Theta} q_{ij}(\theta|m) \log \frac{q_{ij}(\theta\,|\,m)}{p_i(\theta)} \\
&= \sum_{m\in M} \psi_{ij}(m) \sum_{\theta\in\Theta} \frac{\psi_{ij}(m|\theta) p_i(\theta)}{\psi_{ij}(m)} \log \frac{\psi_{ij}(m|\theta)}{\psi_{ij}(m)} \\
&= \sum_{\theta\in\Theta} \sum_{m\in M} p_i(\theta) \sum_{s\in S} \sigma_{ij}(m|s)\phi(s|\theta) \log \frac{\sum_{s'\in S} \sigma_{ij}(m|s')\phi(s'|\theta)}{\sum_{s'\in S} \sigma_{ij}(m|s')\phi(s')}.
\end{aligned}$$

Also, using the definition of power of an informative signal in (2), we have

$$\begin{aligned}
\mathbb{P}(\Phi) &= \sum_{s\in S} \phi(s) D\big(p_i(\cdot|s) \,||\, p_i\big) \\
&= \sum_{s\in S} \phi(s) \sum_{\theta\in\Theta} p_i(\theta|s) \log \frac{p_i(\theta|s)}{p_i(\theta)} \\
&= \sum_{s\in S} \phi(s) \sum_{\theta\in\Theta} \frac{\phi(s|\theta) p_i(\theta)}{\phi(s)} \log \frac{\phi(s|\theta)}{\phi(s)} \\
&= \sum_{\theta\in\Theta} \sum_{s\in S} \phi(s|\theta) p_i(\theta) \log \frac{\phi(s|\theta)}{\phi(s)} \Big[\sum_{m\in M} \sigma_{ij}(m|s) \Big].
\end{aligned}$$

By combining the two above expressions, it follows that

$$\mathbb{P}(\Psi_{ij}) = \mathbb{P}(\Phi) + \sum_{\theta\in\Theta} p_i(\theta) \sum_{m\in M} \sum_{s\in S} \sigma_{ij}(m|s)\phi(s|\theta) \log \frac{\phi(s) \sum_{s'\in S} \sigma_{ij}(m|s')\phi(s'|\theta)}{\phi(s|\theta) \sum_{s'\in S} \sigma_{ij}(m|s')\phi(s')} \tag{8}$$

Furthermore, since $\log x$ is a strictly concave function, from the expression above we have

$$\mathbb{P}(\Psi_{ij}) \le \mathbb{P}(\Phi) + \sum_{\theta\in\Theta} p_i(\theta) \log \sum_{m\in M} \sum_{s\in S} \sigma_{ij}(m|s)\phi(s|\theta) \frac{\phi(s) \sum_{s'\in S} \sigma_{ij}(m|s')\phi(s'|\theta)}{\phi(s|\theta) \sum_{s'\in S} \sigma_{ij}(m|s')\phi(s')}. \tag{9}$$

Now, notice that, for each given $\theta \in \Theta$,

$$\begin{aligned}
&\log \sum_{m\in M} \sum_{s\in S} \sigma_{ij}(m|s)\phi(s|\theta) \frac{\phi(s) \sum_{s'\in S} \sigma_{ij}(m|s')\phi(s'|\theta)}{\phi(s|\theta) \sum_{s'\in S} \sigma_{ij}(m|s')\phi(s')} \\
&\quad = \log \sum_{m\in M} \frac{\big[\sum_{s\in S} \sigma_{ij}(m|s)\phi(s) \big] \big[\sum_{s'\in S} \sigma_{ij}(m|s')\phi(s'|\theta) \big]}{\sum_{s'\in S} \sigma_{ij}(m|s')\phi(s')} \\
&\quad = \log \sum_{m\in M} \sum_{s'\in S} \sigma_{ij}(m|s')\phi(s'|\theta) = \log 1 = 0.
\end{aligned}$$

Therefore, from the inequality in (9), we obtain $\mathbb{P}(\Psi_{ij}) \leq \mathbb{P}(\Phi)$, as stated.

Moreover, 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 the two signal realizations $\underline{s}$ and $\overline{s}$ available to agent j. Without loss of generality, this is achieved if and only if $\sigma_{ij}(\underline{m}|\underline{s}) = \sigma_{ij}(\overline{m}|\overline{s}) = 1$. In this case, for each $\theta \in \Theta$, we have

$$\sum_{m \in M} \sum_{s \in S} \sigma_{ij}(m|s)\phi(s|\theta) \log \frac{\phi(s) \sum_{s' \in S} \sigma_{ij}(m|s')\phi(s'|\theta)}{\phi(s|\theta) \sum_{s' \in S} \sigma_{ij}(m|s')\phi(s')}$$
$$= \phi(\underline{s}|\theta) \log \frac{\phi(\underline{s})\phi(\underline{s}|\theta)}{\phi(\underline{s}|\theta)\phi(\underline{s})} + \phi(\overline{s}|\theta) \log \frac{\phi(\overline{s})\phi(\overline{s}|\theta)}{\phi(\overline{s}|\theta)\phi(\overline{s})} = 0.$$

Therefore, from the expression in (8), 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 $\mathbb{P}(\Psi_{ij}) = \mathbb{P}(\Phi)$. ∎

Proof of Proposition 1. 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 $\beta(m_{ij}; h_{it})$ be the number of periods in which agent i has received message m_{ij} from agent j before period t. Fix a sequence of histories $\{h_{it}\}_{t=0}^{\infty}$. Take a given $\theta \in \Theta$ and set $\{\theta'\} := \Theta \setminus \{\theta\}$. Application of Bayes rule gives

$$\mu_i(\theta|h_{it}) = \left(1 + \frac{p_i(\theta')}{p_i(\theta)} \prod_s \left[\frac{\phi_i(s|\theta')}{\phi_i(s|\theta)}\right]^{\alpha(s;h_{it})} \prod_{j \neq i} \prod_{m_{ij}} \left[\frac{\psi_{ij}(m_{ij}|\theta')}{\psi_{ij}(m_{ij}|\theta)}\right]^{\beta(m_{ij};h_{it})}\right)^{-1}.$$

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

$$\lim_{t\to\infty} \mu_i(\theta|h_{it}) = \lim_{t\to\infty} \left[1 + \frac{p_i(\theta')}{p_i(\theta)} \left(\prod_s \left[\frac{\phi_i(s|\theta')}{\phi_i(s|\theta)}\right]^{\phi_i(s)} \prod_{j \neq i} \prod_{m_{ij}} \left[\frac{\psi_{ij}(m_{ij}|\theta')}{\psi_{ij}(m_{ij}|\theta)}\right]^{\psi_{ij}(m_{ij})}\right)^t\right]^{-1}.$$

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

$$\prod_s \left[\frac{\phi_i(s|\theta')}{\phi_i(s|\theta)}\right]^{\phi_i(s)} \prod_{j \neq i} \prod_{m_{ij}} \left[\frac{\psi_{ij}(m_{ij}|\theta')}{\psi_{ij}(m_{ij}|\theta)}\right]^{\psi_{ij}(m_{ij})}$$

exceeds or not one. By taking logs, this is equivalent to studying whether

$$\sum_s \phi_i(s) \log \frac{\phi_i(s|\theta')}{\phi_i(s|\theta)} + \sum_{j \neq i} \sum_{m_{ij}} \psi_{ij}(m_{ij}) \log \frac{\psi_{ij}(m_{ij}|\theta')}{\psi_{ij}(m_{ij}|\theta)}$$

exceeds or not zero. Then, since $\phi_i(s) = \sum_\theta p_i(\theta)\phi_i(s|\theta)$ and $\psi_{ij}(m_{ij}) = \sum_\theta p_i(\theta)\psi_{ij}(m_{ij}|\theta)$, we can use the definitions of the measures G_i and F_{ij} in (4) and (5), respectively, to obtain that (i) $\lim_{t\to\infty} \mu_i(\theta|h_{it}) = 1$ if and only if $G_i(\theta||\theta') + \sum_{j\neq i} F_{ij}(\theta||\theta') > G_i(\theta'||\theta) + \sum_{j\neq i} F_{ij}(\theta'||\theta)$ and (ii) $\lim_{t\to\infty} \mu_i(\theta|h_{it}) = p_i(\theta)$ if and only if $G_i(\theta||\theta') + \sum_{j\neq i} F_{ij}(\theta||\theta') = G_i(\theta'||\theta) + \sum_{j\neq i} F_{ij}(\theta'||\theta)$, as stated. ∎

Proof of Lemma 2. (i) Consider a set of informative signals $\{\Phi_i\}_{i\in N}$ and a weighted network $\mathbb{P}(\Psi)$. Fix a given agent $i \in N$. Without loss of generality, take $\theta = \overline{\theta}$ so that $\theta' = \underline{\theta}$. From the definitions of power of an informative signal in (2) and of measure G_i in (4), we know that

$$\mathbb{P}(\Phi_i) - \big[G_i(\overline{\theta}||\underline{\theta}) - G_i(\underline{\theta}||\overline{\theta})\big] = \sum_\theta \sum_s p_i(\theta)\phi_i(s|\theta)\log\frac{\phi_i(s|\theta)}{\phi_i(s)}$$
$$- p_i(\overline{\theta})\sum_s \phi_i(s|\overline{\theta})\log\frac{\phi_i(s|\overline{\theta})}{\phi_i(s|\underline{\theta})} + p_i(\underline{\theta})\sum_s \phi_i(s|\underline{\theta})\log\frac{\phi_i(s|\underline{\theta})}{\phi_i(s|\overline{\theta})}$$
$$= \sum_\theta \sum_s p_i(\theta)\phi_i(s|\theta)\left[\log\frac{\phi_i(s|\theta)}{\phi_i(s|\overline{\theta})} - \log\frac{\sum_\theta \phi_i(s|\theta)p_i(\theta)}{\phi_i(s|\underline{\theta})}\right]$$
$$= \sum_\theta \sum_s p_i(\theta)\phi_i(s|\theta)\left[\log\frac{\phi_i(s|\theta)}{\phi_i(s|\overline{\theta})} - \log\left(p_i(\underline{\theta}) + p_i(\overline{\theta})\frac{\phi_i(s|\overline{\theta})}{\phi_i(s|\underline{\theta})}\right)\right].$$

Since $\log x$ is a strictly concave function, it follows from the expression above that

$$\mathbb{P}(\Phi_i) - \big[G_i(\overline{\theta}||\underline{\theta}) - G_i(\underline{\theta}||\overline{\theta})\big] \le \sum_\theta \sum_s p_i(\theta)\phi_i(s|\theta)\left[\log\frac{\phi_i(s|\theta)}{\phi_i(s|\overline{\theta})} - p_i(\overline{\theta})\log\frac{\phi_i(s|\overline{\theta})}{\phi_i(s|\underline{\theta})}\right]. \tag{10}$$

Furthermore, notice that the right-hand side of inequality (10) above can be rewritten as

$$\sum_\theta \sum_s p_i(\theta)\phi_i(s|\theta)\left[\log\frac{\phi_i(s|\theta)}{\phi_i(s|\overline{\theta})} - p_i(\overline{\theta})\log\frac{\phi_i(s|\overline{\theta})}{\phi_i(s|\underline{\theta})}\right] = p_i(\underline{\theta})\sum_s \phi_i(s|\underline{\theta})\log\frac{\phi_i(s|\underline{\theta})}{\phi_i(s|\overline{\theta})}$$
$$- p_i(\overline{\theta})\sum_s\left(\sum_\theta p_i(\theta)\phi_i(s|\theta)\right)\log\frac{\phi_i(s|\overline{\theta})}{\phi_i(s|\underline{\theta})}$$
$$= p_i(\underline{\theta})D(\phi_i(\cdot|\underline{\theta})\,||\,\phi_i(\cdot|\overline{\theta})) + p_i(\overline{\theta})\sum_s\big(p_i(\underline{\theta})\phi_i(s|\underline{\theta}) + p_i(\overline{\theta})\phi_i(s|\overline{\theta})\big)\log\frac{\phi_i(s|\underline{\theta})}{\phi_i(s|\overline{\theta})}$$
$$= p_i(\underline{\theta})D(\phi_i(\cdot|\underline{\theta})\,||\,\phi_i(\cdot|\overline{\theta})) + p_i(\overline{\theta})p_i(\underline{\theta})D(\phi_i(\cdot|\underline{\theta})\,||\,\phi_i(\cdot|\overline{\theta})) - \big[p_i(\overline{\theta})\big]^2 D(\phi_i(\cdot|\overline{\theta})\,||\,\phi_i(\cdot|\underline{\theta})).$$

Therefore, from (10), we obtain

$$\mathbb{P}(\Phi_i) - \big[G_i(\overline{\theta}||\underline{\theta}) - G_i(\underline{\theta}||\overline{\theta})\big] \le G_i(\underline{\theta}||\overline{\theta}) - p_i(\overline{\theta})\big[G_i(\overline{\theta}||\underline{\theta}) - G_i(\underline{\theta}||\overline{\theta})\big]$$
$$\Leftrightarrow \mathbb{P}(\Phi_i) - G_i(\underline{\theta}||\overline{\theta}) \le [1 - p_i(\overline{\theta})]\big[G_i(\underline{\theta}||\overline{\theta}) - p_i(\overline{\theta})G_i(\overline{\theta}||\underline{\theta})\big] = p_i(\underline{\theta})\big[G_i(\underline{\theta}||\overline{\theta}) - p_i(\overline{\theta})G_i(\overline{\theta}||\underline{\theta})\big],$$

as stated.

The proof of part (ii) can be done in a way totally analogous to the proof of part (i) above. One only needs to replicate the arguments given above for $\mathbb{P}(\Psi_{ij})$ instead of $\mathbb{P}(\Phi_i)$. The role played by $G_i(\theta||\theta')$ in the arguments above is now played, in a totally analogous way, by $F_{ij}(\theta||\theta')$. Therefore, we forego a formal statement. ∎

Proof of Proposition 2. The proof is similar to the proof of Proposition 1. For a history h_t, let $\alpha(s_i; h_t)$ be the number of periods in which the external observer has observed agent i's signal s_i before period t. Fix a sequence of histories $\{h_t\}_{t=0}^{\infty}$. Take a given $\theta \in \Theta$ and set $\{\theta'\} := \Theta \setminus \{\theta\}$.

Application of Bayes rule gives

$$\mu_{\rm ob}(\theta|h_t) = \left(1 + \frac{p(\theta')}{p(\theta)} \prod_i \prod_{s_i} \left[\frac{\phi_i(s_i|\theta')}{\phi_i(s_i|\theta)}\right]^{\alpha(s_i;h_t)}\right)^{-1}.$$

Since observed frequencies approximate distributions, i.e., $\lim_{t\to\infty} \alpha(s_i; h_t) = \lim_{t\to\infty}[t\,\phi_i(s_i)]$, we have

$$\lim_{t\to\infty} \mu_{\rm ob}(\theta|h_t) = \lim_{t\to\infty} \left[1 + \frac{p(\theta')}{p(\theta)} \left(\prod_i \prod_{s_i} \left[\frac{\phi_i(s_i|\theta')}{\phi_i(s_i|\theta)}\right]^{\phi_i(s_i)}\right)^t\right]^{-1}.$$

Therefore, upon taking logs, the converge of $\mu_{\rm ob}(\theta|h_t)$ is characterized by the fact that the term $\sum_i \sum_{s_i} \phi_i(s_i) \log[\phi_i(s_i|\theta')/\phi_i(s_i|\theta)]$ exceeds or not zero. Then, since $\phi_i(s_i) = \sum_\theta p_i(\theta)\phi_i(s_i|\theta)$, we can use the definitions of the measure G_i in (4), to obtain that (i) $\lim_{t\to\infty} \mu_{\rm ob}(\theta|h_t) = 1$ if and only if $\sum_{i\in N} G_i(\theta||\theta') > \sum_{i\in N} G_i(\theta'||\theta)$ and (ii) $\lim_{t\to\infty} \mu_{\rm ob}(\theta|h_t) = p(\theta)$ if and only if $\sum_{i\in N} G_i(\theta||\theta') = \sum_{i\in N} G_i(\theta'||\theta)$, as stated. ∎

Proof of Proposition 3. Consider a network $\mathbb{P}(\Psi)$ and suppose that, for each agent $i \in N$, $\lim_{t\to\infty} \mu_i(\theta|h_{it}) = 1$ for some given $\theta \in \Theta$. Then, Proposition 1 implies that

$$[G_i(\theta||\theta') - G_i(\theta'||\theta)] + \sum_{j\neq i} [F_{ij}(\theta||\theta') - F_{ij}(\theta'||\theta)] > 0,$$

where $\{\theta'\} := \Theta \setminus \{\theta\}$. Summing across all agents, we obtain that the condition

$$\sum_{i\in N} [G_i(\theta||\theta') - G_i(\theta'||\theta)] > \sum_{i\in N} \sum_{j\neq i} [F_{ij}(\theta'||\theta) - F_{ij}(\theta||\theta')]$$

is satisfied if a consensus, in which all agents believe in the long run that θ is the true parameter value, is achieved in the society. Therefore, if $\sum_{i\in N} \sum_{j\neq i} [F_{ij}(\theta'||\theta) - F_{ij}(\theta||\theta')] > 0$ holds, then we obtain that $\sum_{i\in N} G_i(\theta||\theta') > \sum_{i\in N} G_i(\theta'||\theta)$. In this case, by applying Proposition 2, it follows that $\lim_{t\to\infty} \mu_{\rm ob}(\theta|h_t) = 1$ so that correct limiting beliefs are attained in the society. ∎

References

[1] Acemoglu, D., Chernozhukov, V., and M. Yildiz (2009): "Fragility of Asymptotic Agreement under Bayesian Learning," mimeo.

[2] Acemoglu, D., Como, D., Fagnani, D., and A. Ozdaglar (2010): "Opinion Fluctuations and Disagreement in Social Networks," mimeo.

[3] Acemoglu, D., Dahleh, M. A., Lobel, I., and A. Ozdaglar (2011): "Bayesian Learning in Social Networks," mimeo, forthcoming in *Review of Economic Studies.*

[4] Acemoglu, D., Ozdaglar, O., and A. ParandehGheibi (2010): "Spread of (Mis)information in Social Networks," *Games and Economic Behavior*, 70, 194-227.

[5] Bala, V. and S. Goyal (1998): "Learning from Neighbors," *Review of Economic Studies*, 65, 595-621.

[6] Blackwell, D. and L. Dubins (1962): "Merging of Opinions with Increasing Information," *Annals of Mathematical Statistics*, 33, 882-887.

[7] Cabrales, A., Gossner, O., and R. Serrano (2012): "Entropy and the Value of Information for Investors," mimeo, forthcoming in *American Economic Review*.

[8] Cripps, M. W., Ely, J. C., Mailath, G. J., and L. Samuelson (2008): "Common Learning," *Econometrica*, 76, 4, 909-933.

[9] Cripps, M. W., Ely, J. C., Mailath, G. J., and L. Samuelson (2012): "Common Learning with Intertemporal Dependence," mimeo, forthcoming in *International Journal of Game Theory*.

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

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

[12] DeMarzo, P., Vayanos, D., and J. Zwiebel (2003): "Persuasion Bias, Social Influence, and Unidimensional Opinions," *Quarterly Journal of Economics*, 118:3, 909-968.

[13] 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.

[14] Hagenbach J. and F. Koessler (2010): "Strategic Communication Networks," *Review of Economic Studies*, 77, 3, 1072-1099.

[15] Heifetz A. (1996): "Comment on Consensus without Common Knowledge," *Journal of Economic Theory*, 70, 273-277.

[16] Koessler, F. (2001): "Common Knowledge and Consensus with Noisy Communication," *Mathematical Social Sciences*, 42, 139-159.

[17] Parikh, R., and P. Krasucki (1990): "Communication, Consensus, and Knowledge," *Journal of Economic Theory*, 52, 1, 178-189.

[18] Rubinstein, A. (1989): "The Electronic Mail Game: Strategic Behavior under 'Almost Common Knowledge'," *American Economic Review*, 79, 3, 385-391.

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

[20] Sciubba, E. (2005): "Asymmetric Information and Survival in Financial Markets," *Economic Theory*, 25, 353-379.

[21] Steiner, J., and C. Stewart (2010): "Communication, Timing, and Common Learning," mimeo.