NBER WORKING PAPER SERIES

DYNAMICS OF INFORMATION EXCHANGE IN ENDOGENOUS SOCIAL NETWORKS

Daron Acemoglu
Kostas Bimpikis
Asuman Ozdaglar

Working Paper 16410
http://www.nber.org/papers/w16410

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2010

We thank seminar participants at Columbia, Cornell, Microsoft Research, MIT, Stanford, University of Chicago, University of Michigan, the SED Meeting at Montreal and the Quebec Workshop on the Economics of Social Networks for useful comments and suggestions. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Dynamics of Information Exchange in Endogenous Social Networks
Daron Acemoglu, Kostas Bimpikis, and Asuman Ozdaglar
NBER Working Paper No. 16410
September 2010
JEL No. D83,D85

ABSTRACT

We develop a model of information exchange through communication and investigate its implications for information aggregation in large societies. An underlying state determines payoffs from different actions. Agents decide which others to form a costly communication link with incurring the associated cost. After receiving a private signal correlated with the underlying state, they exchange information over the induced communication network until taking an (irreversible) action. We define asymptotic learning as the fraction of agents taking the correct action converging to one in probability as a society grows large. Under truthful communication, we show that asymptotic learning occurs if (and under some additional conditions, also only if) in the induced communication network most agents are a short distance away from "information hubs", which receive and distribute a large amount of information. Asymptotic learning therefore requires information to be aggregated in the hands of a few agents. We also show that while truthful communication may not always be a best response, it is an equilibrium when the communication network induces asymptotic learning. Moreover, we contrast equilibrium behavior with a socially optimal strategy profile, i.e., a profile that maximizes aggregate welfare. We show that when the network induces asymptotic learning, equilibrium behavior leads to maximum aggregate welfare, but this may not be the case when asymptotic learning does not occur. We then provide a systematic investigation of what types of cost structures and associated social cliques (consisting of groups of individuals linked to each other at zero cost, such as friendship networks) ensure the emergence of communication networks that lead to asymptotic learning. Our result shows that societies with too many and sufficiently large social cliques do not induce asymptotic learning, because each social clique would have sufficient information by itself, making communication with others relatively unattractive. Asymptotic learning results if social cliques are neither too numerous nor too large, in which case communication across cliques is encouraged.

Daron Acemoglu
Department of Economics
MIT, E52-380B
50 Memorial Drive
Cambridge, MA 02142-1347
and CIFAR
and also NBER
daron@mit.edu

Asuman Ozdaglar
Department of Electrical Engineering
and Computer Science
Massachusetts Institute of Technology
77 Massachusetts Ave, E40-130
Cambridge, MA 02139
asuman@mit.edu

Kostas Bimpikis
MIT - Sloan School of Management
50 Memorial Drive
Cambridge, MA 02142
kostasb@MIT.EDU

1 Introduction

Most social decisions, ranging from product and occupational choices to voting and political behavior, rely on information agents gather through communication with friends, neighbors and co-workers as well as information obtained from news sources and prominent webpages. A central question in social science concerns the dynamics of communication and information exchange and whether such dynamics lead to the effective aggregation of dispersed information. Our objective in this paper is to develop a tractable benchmark model to study the dynamics of belief formation and information aggregation through communication and the choices that individuals make concerning whom to communicate with. A framework for the study of these questions requires communication to be strategic, time-consuming and/or costly, since otherwise all information could be aggregated immediately by simultaneous communication among the agents. Our approach focuses on dynamic and costly communication (and we also allow strategic communication, though this turns out to be less important in the present context).

An *underlying state* of the world determines which action has higher payoff (which is assumed to be the same for all agents). Because of discounting, earlier actions are preferred to later ones. Each agent receives a *private signal* correlated with the underlying state. In addition, she can communicate with others, but such communication first requires the formation of a *communication link*, which may be costly. Therefore, our framework combines elements from models of social learning and network formation. The network formation decisions of agents induce a *communication graph* for the society. Thereafter, agents communicate with those whom they are connected to until they take an irreversible action. Crucially, information acquisition takes time because the "neighbors" of an agent with whom she communicates acquire more information from their own neighbors over time. Information exchange will thus be endogenously limited by two features: the communication network formed at the beginning of the game, which allows communication only between connected pairs, and discounting, which encourages agents to take actions before they accumulate sufficient information.

We characterize the equilibria of this network formation and communication game and then investigate the structure of these equilibria as the society becomes large (i.e., for a sequence of games). Our main focus is on how well information is aggregated, which we capture with the notion of asymptotic learning. We say that there is *asymptotic learning* if the fraction of agents taking the correct action converges to one (in probability) as the society becomes large.

Our analysis proceeds in several stages. *First*, we take the communication graph as given and assume that agents are non-strategic in their communication, i.e., they disclose truthfully all the information they possess when communicating. Under these assumptions, we provide a condition that is sufficient and (under an additional mild assumption) necessary for asymptotic learning. Intuitively, this condition requires that most agents are a short distance away from *information hubs*, which are agents that have a very large (in the limit, infinite) number of connections.[1] Two different types of information hubs can be conduits of asymptotic learning in our benchmark model. The first are *information mavens* who receive communication from many other agents, enabling them to aggregate information. If most agents are close to an information maven, asymptotic learning is guaranteed. The second type of hubs are *social connectors* who communicate to many agents, enabling them to spread their information widely.[2] Social connectors are only useful for asymptotic learning, if they are close to mavens so that they can distribute their information. Thus, asymptotic learning is also obtained if most agents are close to a social connector, who is in turn a short distance away from a maven. The intuition for why such information hubs and almost all agents being close to information hubs are necessary for asymptotic learning is instructive: were it not so, a large fraction of agents would prefer to act before waiting for sufficient information to arrive. But then a nontrivial fraction of those would take the incorrect action, and moreover, they would also disrupt the information flow for the agents to whom they are connected. The advantage of the first part of our analysis is that it enables a relatively simple characterization of equilibria and the derivation of intuitive conditions for asymptotic learning.

Second, we show that even if individuals misreport their information (which they may want to do in order to delay the action of their neighbors and obtain more information from them in future communication), it is an equilibrium of the strategic communication game to report truthfully whenever truthful communication leads to asymptotic learning. Interestingly, the converse is not necessarily true: strategic communication may lead to asymptotic learning in some special cases in which truthful communication precludes learning. From a welfare perspective, we show a direct connection between asymptotic learning and the maximum aggregate welfare that can be achieved by any strategy profile: when asymptotic learning occurs, all equilibria are (asymptotically) socially efficient, i.e., they achieve the maximum welfare. However, when asymptotic learning does not occur, equilibrium behavior can lead to inefficiencies that arise from the fact that agents do not internalize the positive effect of delaying their action and continuing information exchange. Thus, our analysis identifies a novel information externality that is a direct product of the agents being embedded in a network: the value of an agent to her peers does not only originate from her initial information but also from the paths she creates between different parts of the network through her social connections. It is precisely the destruction of these paths when the agent takes an action that may lead to a welfare loss in equilibrium.

Our characterization results on asymptotic learning can be seen both as "positive" and "negative". On the one hand, to the extent that most individuals obtain key information from either individuals or news sources (websites) approximating such hubs, efficient aggregation of information may be possible

[1] We also derive conditions under which ϵ, δ-asymptotic learning occurs at an equilibrium strategy profile. We say that ϵ, δ-asymptotic learning occurs when at least $1 - \epsilon$ fraction of the population takes an ϵ-optimal action with probability at least $1 - \delta$.

[2] Both of these terms are inspired by Gladwell (2000).

in some settings. We show in particular that hierarchical graph structures where agents in the higher layers of the hierarchy can communicate information to many agents at lower layers lead to asymptotic learning.[3] On the other hand, communication structures that do not feature such hubs appear more realistic in most contexts, including communication between friends, neighbors and co-workers.[4] Our model thus emphasizes how each agent's incentive to act sooner rather than later makes information aggregation significantly more difficult.

Third, armed with the analysis of information exchange over a given communication network, we turn to the study of the endogenous formation of this network. We assume that the formation of communication links is costly, though there also exist *social cliques*, groups of individuals that are linked to each other at zero cost. These can be thought of as "friendship networks" that are linked for reasons unrelated to information exchange and thus act as conduits of such exchange at low cost. Agents have to pay a cost at the beginning in order to communicate (receive information) from those who are not in their social clique. Even though network formation games have several equilibria, the structure of our network formation and information exchange game enables us to obtain relatively sharp results on what types of societies lead to endogenous communication networks that ensure asymptotic learning. In particular, we show that societies with too many (disjoint) and sufficiently large social cliques induce behavior inconsistent with asymptotic learning. The reason why relatively large social cliques may discourage efficient aggregation of information is that because they have enough information, communication with others (from other social cliques) becomes unattractive, and as a consequence, the society gets segregated into a large number of disjoint social cliques that do not share information. In contrast, asymptotic learning obtains in equilibrium if social cliques are neither too numerous nor too large so that it is worthwhile for at least some members of these cliques to communicate with members of other cliques, forming a structure in which information is shared across (almost) all members of the society.

These results also illustrate an interesting feature of the information exchange process: an agent's willingness to perform costly search (which here corresponds to forming a link with another social clique) is decreasing with the precision of the information that is readily accessible to her. This gives a natural explanation for informational segregation: agents do not internalize the benefits for the group of forming an additional link, leading to a socially inefficient information exchange structure. It further suggests a form of *informational Braess' paradox*,[5] whereby the introduction of additional information may have adverse effects for the welfare of a society by discouraging the formation of additional links for information sharing (see also Morris and Shin (2002) and Duffie, Malamud, and Manso (2009) for

[3] An additional challenge when significant information is concentrated in the hands of a few hubs may arise because of misalignment of interests, which our approach ignores.

[4] In particular, the popular (though not always empirically plausible) random graph models such as preferential attachment and Poisson (Erdős-Renyi) graphs do not lead to asymptotic learning.

[5] In the original Braess' paradox, the addition of a new road may increase the delays faced by all motorists in a Nash equilibrium.

a related result). Consider, for example, the website of a film critic that can be viewed as a good but still imprecise information source (similar to a reasonable-sized social clique in our model). Other agents can access the critic's information and form an opinion about a movie quickly. However, this precludes information sharing among the agents and may lead to a decrease in the aggregate welfare.

Our paper is related to several strands of the literature on social and economic networks. First, it is related to the large and growing literature on *social learning.* Much of this literature focuses on Bayesian models of observational learning, where each individual learns from the actions of others taken in the past. A key impediment to information aggregation in these models is the fact that actions do not reflect all of the information that an individual has and this can induce a pattern reminiscent to a "herd," where individuals ignore their own information and copy the behavior of others (see, for example, Bikhchandani, Hirshleifer, and Welch (1992), Banerjee (1992), and Smith and Sørensen (2000), as well as Bala and Goyal (1998), for early contributions, and Smith and Sørensen (2010), Banerjee and Fudenberg (2004) and Acemoglu, Dahleh, Lobel, and Ozdaglar (2010) for models of Bayesian learning with richer observational structures). While observational learning is important in many situations, a large part of information exchange in practice is through communication.

Several papers in the literature study communication, though typically using non-Bayesian or "myopic" rules (for example, Ellison and Fudenberg (1995), DeMarzo, Vayanos, and Zwiebel (2003) and Golub and Jackson (2010)). A major difficulty faced by these approaches, often precluding Bayesian and dynamic game theoretic analysis of learning in communication networks, is the complexity of updating when individuals share their ex-post beliefs (because of the difficulty of filtering out common sources of information). We overcome this difficulty by adopting a different approach, whereby individuals can directly communicate their signals and there is no restriction on the total "bits" of communication. This leads to a tractable structure for updating of beliefs and enables us to study perfect Bayesian equilibria of a dynamic game of network formation, communication and decision-making. It also reverses one of the main insights of these papers, also shared by the pioneering social learning work by Bala and Goyal (1998), that the presence of "highly connected" or "influential" agents, or what Bala and Goyal (1998) call a "royal family," acts as a significant impediment to the efficient aggregation of information. On the contrary, in our model the existence of such highly connected agents (information hubs, mavens or connectors) is crucial for the efficient aggregation of information. Moreover, the existence of such "highly connected" also reduces incentives for non-truthful communication, and is the key input into our result that truthful communication can be an equilibrium. The recent paper by Duffie, Malamud, and Manso (2009) is also noteworthy: in their model agents are randomly matched according to endogenously determined search intensities, and because they focus on an environment with a continuum of agents, communication of beliefs in their setup is equivalent to exchanging signals, and thus enables them to avoid the issues arising in the previous literature. Their main focus is on characterizing equilibrium search intensities as a function of the information that an agent already has access to. In contrast to our work, there is no explicit network structure. Finally, Mobius, Phan, and Szeidl (2010) empirically compare a non-Baysian model of communication (similar to the one adopted by Golub and Jackson (2010)) with a model in which, similar to ours, signals are communicated and agents are Bayesian. Although their study is not entirely conclusive on whether agents behave according to one or the other model, their evidence broadly supports the Bayesian alternative.

Our work is also related to the growing literature on *network formation*, since communication takes place over endogenously formed networks. Bala and Goyal (2000) model strategic network formation as a non-cooperative game and study its equilibria under various assumptions on the benefits of forming a link. In particular, they distinguish between *one-way* and *two-way* flow of benefits, depending on whether a link benefits only the agent that decides to form it or both participating agents. They identify a number of simple structures that arise at equilibrium: the empty network, the wheel, the star and the complete network. More recently, Galeotti, Goyal, and Kamphorst (2006) and Galeotti (2006) study the role of heterogeneity among agents in the network structures that arise at equilibrium. Closer to our work is Hojman and Szeidl (2008) who study a network formation model where the benefits from connecting to other agents have decreasing returns to scale (which is also the case in our model of information exchange because of endogenous reasons). The main focus of the network formation literature is on characterizing equilibrium structures and comparing them with patterns observed in real world networks (e.g., small distances between agents, high centrality etc.). However, in most of the literature the benefits and costs associated with forming a link are exogenous. A novelty in our work is that the benefits of forming links are endogenously determined through the subsequent information exchange. Our focus is also different: although we also obtain characterization results on the shape of the network structures that arise in equilibrium (which are similar to those in the literature), our focus is on whether these structure lead to asymptotic learning. Interestingly, while network formation games have a large number of equilibria, the simple structure of our model enables us to derive relatively sharp results about environments in which the equilibrium networks lead to asymptotic learning.

Finally, our paper is related to the literature on *strategic communication*, pioneered by the cheap talk framework of Crawford and Sobel (1982). While cheap talk models have been used for the study of information aggregation with one receiver and multiple senders (e.g. Morgan and Stocken (2008)) and multiple receivers and single sender (e.g. Farrell and Gibbons (1989)), most relevant to our paper are two recent papers that consider strategic communication over general networks, Galeotti, Ghiglino, and Squintani (2010) and Hagenbach and Koessler (2010). A major difference between these works and ours is that we consider a model where communication is allowed for more than one time period, thus enabling agents to receive information outside their immediate neighborhood (at the cost of a delayed decision) and we also endogenize the network over which communication takes place. On the other hand, our framework assumes that an agent's action does not directly influence others' payoffs, while such payoff interactions are the central focus of Galeotti, Ghiglino, and Squintani (2010) and Hagenbach and Koessler (2010). Our paper is also related to the existing work by Ambrus, Azevedo, and Kamada (2010), where the sender and the receiver communicate strategically through a chain of intermediators. Their primary focus is information intermediation, thus communication takes place over multiple rounds but it is restricted on a ordered line from the sender to the receiver, where each agent only sends information once.

The rest of the paper is organized as follows. Section 2 develops a general model of information exchange among rational agents, that are embedded in a communication network. Also, it introduces the two main environments we study. Section 3 contains our main results on social learning given a communication graph. It also includes a welfare discussion that draws the connection between learning and efficient communication. Finally, it illustrates how our results can be applied to a number of random graph models. Section 4 incorporates endogenous network formation to the information exchange model. Our main result in this section shows the connection between incentives to form communication links and asymptotic learning. Section 5 concludes. All proofs are presented in the Appendix.

2 A Model of Information Exchange in Social Networks

In the first part of the paper, we focus on modelling information exchange among agents over a given communication network. In the second part (Section 4), we investigate the question of endogenous formation of this network. We start by presenting the information exchange model for a finite set $\mathcal{N}^n = \{1, 2, \cdots, n\}$ of agents. We also describe the limit economy as $n \to \infty$.

2.1 Actions, Payoffs and Information

Each agent $i \in \mathcal{N}^n$ chooses an irreversible action $x_i \in \mathbb{R}$. Her payoff depends on her action and an underlying *state of the world* $\theta \in \mathbb{R}$, which is an exogenous random variable. In particular, agent i's payoff when she takes action x_i and the state of the world is θ is given by $f(x_i, \theta) = \pi - (x - \theta)^2$, where π is a constant.

The state of the world θ is unknown and agents observe noisy signals about it. In particular, we assume that θ is drawn from a Normal distribution with known mean μ and precision ρ. Each agent receives a *private signal* $s_i = \theta + z_i$, where the z_i's are idiosyncratic and independent from one another and θ, with common mean $\bar{\mu}$ (normalized to 0) and precision $\bar{\rho}$.

2.2 Communication

Our focus is on information aggregation, when agents are embedded in a network that imposes communication constraints. In particular, agent i forms beliefs about the state of the world from her private signal s_i, as well as information she obtains from other agents through a given *communication* *network* G^n, which, as will be described shortly, represents the set of communication constraints imposed on them. We assume that time $t \in [0, \infty)$ is continuous and there is a common discount rate $r > 0$. Communication times are stochastic. In particular, communication times are exponentially distributed with parameter $\lambda > 0$.[6] At a given time instant t, agent i decides whether to take action x_i (and receive payoff $f(x_i, \theta)$ discounted by e^{-rt}) or "wait" to obtain more information in subsequent communication rounds from her peers. Throughout the rest of the paper, we say that the agent "exits" at time t, if she chooses to take the irreversible action at time t. Discounting implies that an earlier exit is preferred to a later one. We define U_i^n as the discounted payoff of agent i (from the viewpoint of time $t = 0$) when the size of the society is n. For example, when the underlying state is θ and the agent takes action x_i at time t, we would have that

$$U_i^n = e^{-rt}\left(\pi - (x_i - \theta)^2\right).$$

As mentioned above, each agent obtains information from other agents through a communication network represented by a directed graph $G^n = (\mathcal{N}^n, \mathcal{E}^n)$, where $\mathcal{E}^n$ is the set of directed edges with which agents are linked. We say that agent j can *obtain information* from i or that agent i can *send information* to j if there is an edge from i to j in graph G^n, i.e., $(i, j) \in \mathcal{E}^n$. Let $I_{i,t}^n$ denote the *information set* of agent i at time t and $\mathcal{I}_{i,t}^n$ denote the set of all possible information sets. Then, for every pair of agents i, j, such that $(i, j) \in \mathcal{E}^n$, we say that agent j *communicates* with agent i or that agent i sends a *message* to agent j, and define the following mapping

$$m_{ij,t}^n : \mathcal{I}_{i,t}^n \to \mathcal{M}_{ij,t}^n \text{ for } (i, j) \in \mathcal{E}^n,$$

where $\mathcal{M}_{ij,t}^n \subseteq \mathbb{R}^n$ denotes the set of messages that agent i can send to agent j at time t. Note that without loss of generality the k-th component of $m_{ij,t}^n$ represents the information that agent i sends to agent j at time t regarding the signal of agent k[7]. Moreover, the definition of $m_{ij,t}^n$ captures the fact that communication is directed and is only allowed between agents that are linked in the communication network, i.e., j communicates with i if and only if $(i, j) \in \mathcal{E}^n$. The direction of communication should be clear: when agent j communicates with agent i, then agent i sends a message to agent j, that could in principle depend on the information set of agent i as well as the identity of agent j.

Importantly, we assume that the cardinality ("dimensionality") of $\mathcal{M}_{ij,t}^n$ is such that communication can take the form of agent i sharing all her information with agent j. This has two key implications. First, an agent can communicate (indirectly) with a much larger set of agents than just her immediate neighbors, albeit with a time delay. For example, the second time agent j communicates

[6]Equivalently, agents "wake" up and communicate simultaneously with their neighbors, when a Poisson clock with rate λ ticks.

[7]As will become evident in subsequent discussion, we assume that communication involves exchange of signals and not posterior beliefs. Moreover, information is tagged, i.e., the receiver of the message understands that its k-th component is associated with agent k.

with agent i, then j can send information not just about her direct neighbors, but also their neighbors (since presumably she obtained such information during the first communication). Second, mechanical duplication of information can be avoided. In particular, the second time agent j communicates with agent i, she can repeat her original signal, but this is not recorded as an additional piece of information by agent j, since given the size of the message space $\mathcal{M}_{ij,t}^n$, each piece of information is "tagged". This ensures that there need be no confounding of new information and previously communicated information.

Let T_t denote the set of times that agents communicated with their neighbors before time t. That defines the information set of agent i at time $t > 0$ as:

$$I_{i,t}^n = \{s_i, m_{ji,\tau}^n, \text{ for all } \tau \in T_t \text{ and } j \text{ such that } (j,i) \in \mathcal{E}^n\}$$

and $I_{i,0}^n = \{s_i\}$. In particular, the information set of agent i at time $t > 0$ consists of her private signal and all the messages her neighbors sent to i in previous communication times. Agent i's action at time t is a mapping from her information set to the set of actions, i.e.,

$$\sigma_{i,t}^n : \mathcal{I}_{i,t}^n \to \{\text{"wait"}\} \cup \mathcal{R}.$$

The tradeoff between taking an irreversible action and waiting, should be clear at this point. An agent would wait, in order to communicate indirectly with a larger set of agents and choose a better action. On the other hand, future is discounted, therefore, delaying is costly.

We close the section with a number of definitions. We define a *path* between agents i and j in network G^n as a sequence $i_1, \cdots, i_K$ of distinct nodes such that $i_1 = i$, $i_K = j$ and $(i_k, i_{k+1}) \in \mathcal{E}^n$ for $k \in \{1, \cdots, K-1\}$. The *length* of the path is defined as $K - 1$. Moreover, we define the distance of agent i to agent j as the length of the shortest path from i to j in network G^n, i.e.,

$$dist^n(i,j) = \min\{\text{length of } \mathcal{P} \mid \mathcal{P} \text{ is a path from } i \text{ to } j \text{ in } G^n\}.$$

Finally, the k-step *neighborhood* of agent i is defined as

$$B_{i,k}^n = \{j \mid dist^n(j,i) \leq k\},$$

where $B_{i,0}^n = \{i\}$, i.e., $B_{i,k}^n$ consists of all agents that are at most k *links* away from agent i in graph G^n. Intuitively, if agent i waits for k communication steps and all of the intermediate agents receive and communicate information truthfully, i will have access to all of the signals of the agents in $B_{i,k}^n$.

2.3 Equilibria of the Information Exchange Game

We refer to the game defined above as the Information Exchange Game. We next define the equilibria of the information exchange game $\Gamma_{info}(G^n)$ for a given communication network G^n. We use the standard notation σ_{-i} to denote the strategies of agents other than i and we let $\sigma_{i,-t}$ denote the vector of actions of agent i at all times except t. Also, let $\mathbb{P}_\sigma$ and $\mathbb{E}_\sigma$ denote the conditional probability, conditional expectation respectively when agents behave according to profile σ.

Definition 1. *An action strategy profile $\sigma^{n,*}$ is a pure-strategy perfect Bayesian Equilibrium of the information exchange game $\Gamma_{info}(G^n)$ if for every $i \in \mathcal{N}^n$ and time t, $\sigma_{i,t}^{n,*}$ maximizes the expected payoff of agent i given the strategies of other agents $\sigma_{-i}^{n,*}$, i.e.,*

$$\sigma_{i,t}^{n,*} \in \arg\max_{y \in \{\text{“}wait\text{”}\} \cup \mathcal{R}} \mathbb{E}_{((y,\sigma_{i,-t}^{n,*}),\sigma_{-i}^{n,*})}(U_i^n \big| I_{i,t}^n).$$

We denote the set of equilibria of this game by $INFO(G^n)$.

For the remainder, we refer to a pure-strategy perfect Bayesian Equilibrium simply as an equilibrium (we do not study mixed strategy equilibria). It is important to note here that although equilibria depend on the discount rate r, we do not explicitly condition on r (through the use of a subscript) for convenience.

If agent i decides to exit and take an action at time t, then the optimal action would be:

$$x_{i,t}^{n,*} = \arg\max_x \mathbb{E}[f(x,\theta)\big|I_{i,t}^n] = \mathbb{E}[\theta\big|I_{i,t}^n],$$

where the second equality holds as $f(x,\theta) = \pi - (x-\theta)^2$. Since actions are irreversible, the agent's decision problem reduces to determining the timing of her action. It is straightforward to see that at equilibrium an agent takes the irreversible action immediately after some communication step concludes. Thus, an equilibrium strategy profile σ induces an equilibrium timing profile $\tau^{n,\sigma}$, where $\tau_i^{n,\sigma}$ designates the communication step after which agent i exits by taking an irreversible action. The τ notation is convenient to use for the statement of some of our results below. Finally, similar to $B_{i,k}^n$, we define the k-step neighborhood of agent i under equilibrium σ as follows: a path $\mathcal{P}^\sigma$ between agents i and j in G^n under σ is a sequence $i_1, \cdots, i_K$ of distinct nodes such that $i_1 = i$, $i_K = j$, $(i_k, i_{k+1} \in \mathcal{E}^n)$ and $\tau_{i_k}^{n,\sigma} \geq k-1$, which ensures that the information from j will reach agent i before any of the agents in the path take an irreversible action. Then, we can define

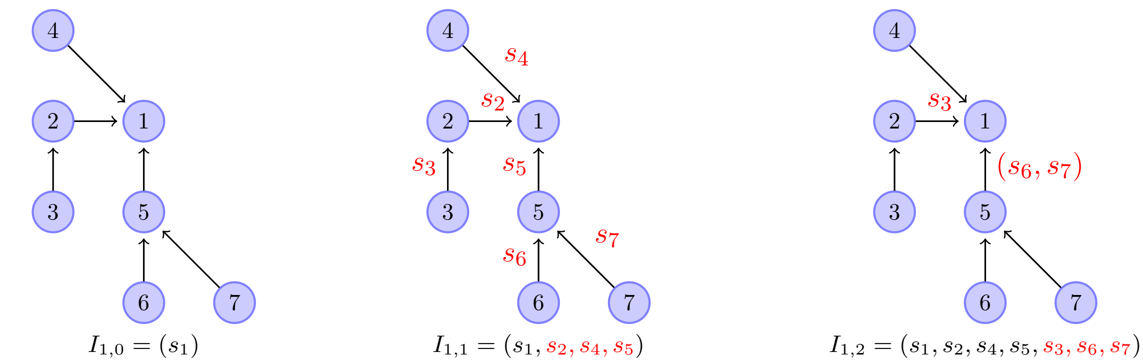
$$dist^{n,\sigma}(i,j) = \min\{\text{length of } \mathcal{P}^\sigma \;\big|\; \mathcal{P}^\sigma \text{ is a path from } i \text{ to } j \text{ in } G^n \text{ under equilibrium } \sigma\}$$

and

$$B_{i,k}^{n,\sigma} = \{j \;\big|\; dist^{n,\sigma}(j,i) \leq k\}.$$

2.4 Assumptions on the Information Exchange Process

The communication model described in Section 2.2 is fairly general. In particular, the model does not restrict the set of messages that an agent can send. Throughout, we maintain the assumption that the communication network G^n is *common knowledge*. Also, we focus on the following two environments



(a) Time $t = 0$. (b) First communication step. (c) Second communication step.

Figure 1: The information set of agent 1 under truthful communication.

defined by Assumptions 1 and 2 respectively.

Assumption 1 (Truthful Communication)**.** *Communication between agents is truthful, i.e.,*

$$m^n_{ij,t} = \begin{cases} \hat{m^n}_{ij,t} & \text{if } |T_t| \leq \tau_i^{n,\sigma} \\ \hat{m^n}_{ij,\tau_i^{n,\sigma}} & \text{otherwise.} \end{cases}$$

and

$$(\hat{m^n}_{ij,t})_\ell = \begin{cases} s_\ell & \text{if } dist^{n,\sigma}_{i,\ell} \leq |T_t| \\ \in \mathbb{R} & \text{otherwise} \end{cases}$$

Intuitively, this assumption compactly imposes three crucial features: (1) Communication takes place by sharing signals, so that when agent j communicates with agent i at time t, then agent i sends to j all the information agent i has obtained thus far (refer to Figure 1 for an illustration of the communication process centered at a particular agent); (2) Agents cannot strategically manipulate the messages they sent, i.e., an agent's private signal is *hard information*. Moreover, they cannot refuse to disclose the information they possess; (3) When an agent takes an irreversible action, then she no longer obtains new information and, thus, can only communicate the information she has obtained until the time of her decision. The latter feature captures the fact that an agent, who engages in information exchange to make a decision, would have weaker incentives to collect new information after reaching that decision. Nevertheless, she can still communicate the information she had previously obtained to other agents. An interesting consequence of this feature is that it imposes dynamically new constraints to communication: agent i can communicate with agent j only if there is a directed path between them in the original communication network G^n *and* the agents in the path do not exit early. We call this type of communication *truthful* to stress the fact that the agents cannot strategically manipulate the information they communicate.[8] We discuss the implications of relaxing Assumption 1 by allowing

[8]Yet another variant of this assumption would be that agents exit the social network after taking an action and stop communicating entirely. In this case, the results are essentially identical when their action is observed by their neighbors.

strategic communication in Subsection 3.4.

2.5 Learning in Large Societies

We are interested in whether equilibrium behavior leads to information aggregation. This is captured by the notion of "asymptotic learning", which characterizes the behavior of agents over communication networks with growing size.

We consider a sequence of communication networks $\{G^n\}_{n=1}^{\infty}$, where $G^n = \{\mathcal{N}^n, \mathcal{E}^n\}$, and refer to this sequence of communication networks as a *society*. A sequence of communication networks induces a sequence of information exchange games, and with a slight abuse of notation we use the term *equilibrium* to denote a sequence of equilibria of the information exchange games, or of the society $\{G^n\}_{n=1}^{\infty}$. We denote such an equilibrium by $\sigma = \{\sigma^n\}_{n=1}^{\infty}$, which designates that $\sigma^n \in INFO(G^n)$ for all n. For any fixed $n \geq 1$ and any equilibrium of the information exchange game $\sigma^n \in INFO(G^n)$, we introduce the indicator variable:

$$M_i^{n,\epsilon} = \begin{cases} 1 & \text{if agent } i \text{ takes an action that is } \epsilon\text{-close to the optimal,} \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

In other words, $M_i^{n,\epsilon} = 1$ (for some ϵ) if and only if agent i chooses irreversible action x_i, such that $|x_i - \theta| \leq \epsilon$.

Next definition introduces ϵ, δ-asymptotic learning for a given society.[9]

Definition 2. *We say that ϵ, δ-asymptotic learning occurs in society $\{G^n\}_{n=1}^{\infty}$ along equilibrium σ if we have:*

$$\lim_{n\to\infty} \mathbb{P}_\sigma \left(\left[\frac{1}{n} \sum_{i=1}^{n} (1 - M_i^{n,\epsilon}) \right] > \epsilon \right) < \delta.$$

This definition states that ϵ, δ-asymptotic learning occurs when the probability that at least $(1 - \epsilon)$-fraction of the agents take an action that is ϵ-close to the optimal action (as the society grows infinitely large) is at least $1 - \delta$.

Definition 3. *We say that perfect asymptotic learning occurs in society $\{G^n\}_{n=1}^{\infty}$ along equilibrium σ if we have:*

$$\lim_{n\to\infty} \mathbb{P}_\sigma \left(\left[\frac{1}{n} \sum_{i=1}^{n} (1 - M_i^{n,\epsilon}) \right] > \epsilon \right) = 0.$$

for any $\epsilon > 0$.

Perfect asymptotic learning is naturally a stronger definition (corresponding to ϵ and δ being arbitrarily small in the definition of ϵ, δ-asymptotic learning) and requires all but a negligible fraction of the agents taking the optimal action in the limit as $n \to \infty$.

However, if their action is not observable, then the analysis needs to be modified in particular, there exist other equilibria where several agents might exit together expecting others to exit. We do not analyze these variants in the current version to save space.

[9] Note that we could generalize Definition 2 by introducing yet another parameter and study ϵ, δ, ζ-asymptotic learning, in which case we would require that $\lim_{n\to\infty} \mathbb{P}_\sigma \left(\left[\frac{1}{n} \sum_{i=1}^{n} (1 - M_i^{n,\epsilon}) \right] > \zeta \right) < \delta$.

3 Learning and Efficient Communication

In this section, we present our main results on learning and discuss their implications for the aggregate welfare. Before doing so, we discuss the decision problem of a single agent, i.e., determining the best time to take an irreversible action given that the rest of the agents behave according to strategy profile σ. Later, we contrast the single agent problem with that of a social planner, whose objective is to maximize the expected aggregate welfare. The analysis in the next three subsections assumes that communication is truthful (cf. Assumption 1).

3.1 Agent i's problem

The (non-discounted) expected payoff of agent i taking an action after observing k truthful private signals (including her own) is given by:

$$\pi - \frac{1}{\rho + \bar{\rho}k},$$

where recall that ρ, $\bar{\rho}$ are the precisions of the state θ and the idiosyncratic noise respectively. To see this, note that if agent i takes her irreversible action, then the optimal such action would be $\hat{\theta} = \mathbb{E}[\theta \big| I_{i,t}^n]$ and the associated non-discounted payoff would be equal to:

$$\mathbb{E}[\pi - (\hat{\theta} - \theta)^2 \big| I_{i,t}^n] = \pi - var(\hat{\theta} - \theta \big| I_{i,t}^n) = \pi - var\left(\sum_{i=1}^{k} \frac{s_{(i)}}{k} - \theta \big| I_{i,t}^n\right) = \pi - \frac{1}{\rho + \bar{\rho}k},$$

where $s_{(i)}$ denotes the i-th signal observed by the agent and $\hat{\theta}$ is equal to the sum of k private signals normalized by k.

By the principle of optimality, the value function for agent i at information set $I_{i,t}^n$ and assuming that the rest of the agents behave according to profile σ is given by:

$$\mathbb{E}_{\sigma}(U_i^n \big| I_{i,t}^n) = \max \begin{cases} \pi - \frac{1}{\rho + \bar{\rho}k_{i,t}^{n,\sigma}} & \text{(when she takes the optimal irreversible action),} \\ e^{-rdt}\mathbb{E}[\mathbb{E}_{\sigma}(U_i^n \big| I_{i,t+dt}^n) \big| I_{i,t}^n] & \text{(when she decides to wait, i.e., } x = \text{“}wait\text{”),} \end{cases}$$

where $k_{i,t}^{n,\sigma}$ denotes the number of distinct private signals agent i has observed up to time t. The first line is equal to the expected payoff for the agent when she chooses the optimal irreversible action under information set $I_{i,t}^n$, i.e., $\mathbb{E}[\theta | I_{i,t}^n]$, and she has observed $k_{i,t}^{n,\sigma}$ private signals, while the second line is equal to the discounted expected continuation payoff.

The following lemma states that an agent's optimal action takes the form of a threshold rule: there exists a threshold $(k_{i,T_{|t|}}^{n,\sigma})^*$, such that an agent decides to take an irreversible action at time t as long as she has observed more that $(k_{i,T_{|t|}}^{n,\sigma})^*$ private signals. Like all other results in the paper, the proof of this lemma is provided in the Appendix.

Lemma 1. *Suppose Assumption 1 holds. Given communication network G^n and equilibrium $\sigma \in INFO(G^n)$, there exists a sequence of signal thresholds for each agent i, $\{(k_{i,\tau}^{n,\sigma})^*\}_{\tau=0}^{\infty}$, that depend on the current communication round, the identity of the agent i, the communication network G^n and σ*

such that agent i maximizes her expected utility at information set $I^n_{i,t}$ by taking action $x^n_{i,t}(I^n_{i,t})$ defined as

$$x^n_{i,t}(I^n_{i,t}) = \begin{cases} \mathbb{E}[\theta \big| I^n_{i,t}], & \textit{if } k^{n,\sigma}_{i,t} \geq (k^{n,\sigma}_{i,|T_t|})^*, \\ \textit{“wait”}, & \textit{otherwise,} \end{cases}$$

A consequence of Lemma 1 is that an equilibrium strategy profile σ defines both a time in which agent i acts (immediately after communication step $\tau^{n,\sigma}_i$), but also the number of signals that agent i has access to when she acts.

3.2 Asymptotic Learning

We begin the discussion by introducing the concepts that are instrumental for asymptotic learning: *the observation radius* and *k-radius sets*. Recall that an equilibrium of the information exchange game on communication network G^n, $\sigma^n \in INFO(G^n)$, induces a timing profile $\tau^{n,\sigma}$, such that agent i takes an irreversible action after $\tau^{n,\sigma}_i$ communication steps. We call $\tau^{n,\sigma}_i$ the *observation radius* of agent i under equilibrium profile σ^n. We also define agent i's *perfect observation radius*, τ^n_i, as the communication round that agent i would exit assuming that all other agents never exit. Note that an agent's perfect observation radius is equilibrium independent and depends only on the network structure. On the other hand, $\tau^{n,\sigma}_i$ is an endogenous object and depends on both the network as well as the specific equilibrium profile σ. Given the notion of an observation radius, we define *k-radius sets* (and similarly *perfect k-radius sets*) as follows.

Definition 4. *Let $V^{n,\sigma}_k$ be defined as*

$$V^{n,\sigma}_k = \{i \in \mathcal{N} \mid \big|B^{n,\sigma}_{i,\tau^{n,\sigma}_i}\big| \leq k\}.$$

We refer to $V^{n,\sigma}_k$ as the k-radius set (along equilibrium σ). Similarly, we refer to

$$V^n_k = \{i \in \mathcal{N} \mid \big|B^n_{i,\tau^n_i}\big| \leq k\}$$

as the perfect k-radius set.

Intuitively, $V^{n,\sigma}_k$ includes all agents that take an action before they receive signals from more than k other individuals at equilibrium σ. Equivalently, the size of their (indirect) neighborhood by the time they take an irreversible action is no greater than k. From Definition 4 it follows immediately that

$$i \in V^{n,\sigma}_k \Rightarrow i \in V^{n,\sigma}_{k'} \text{ for all } k' > k. \tag{2}$$

The following proposition provides a necessary and a sufficient condition for ϵ, δ-asymptotic learning to occur in a society under equilibrium profile σ. Recall that $erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$ denotes the *error function* of the normal distribution.

Proposition 1. *Suppose Assumption 1 holds. Then,* *(a) ϵ, δ-asymptotic learning does not occur in society $\{G^n\}_{n=1}^{\infty}$ under equilibrium profile σ if there exists $k > 0$ such that*

$$\eta = \limsup_{n\to\infty} \frac{1}{n} \cdot \left|V_k^{n,\sigma}\right| > \epsilon \text{ and } erf\left(\epsilon\sqrt{\frac{k\bar{\rho}}{2}}\right) < (1-\delta)(1-\epsilon/\eta). \tag{3}$$

(b) ϵ, δ-asymptotic learning occurs in society $\{G^n\}_{n=1}^{\infty}$ under equilibrium profile σ if there exists $k > 0$ such that

$$\zeta = \liminf_{n\to\infty} \frac{1}{n} \cdot \left|V_k^{n,\sigma}\right| < \epsilon \text{ and } erf\left(\epsilon\sqrt{\frac{k\bar{\rho}}{2}}\right) > 1 - \frac{\delta(\epsilon-\zeta)}{1-\zeta}. \tag{4}$$

This proposition provides conditions such that ϵ, δ-asymptotic learning takes place (or does not take place). Intuitively, asymptotic learning is precluded if there exists a significant fraction of the society that takes an action before seeing a large set of signals, since in this case there is a large enough probability that these agents will take an action far away from the optimal one. The proposition quantifies the relationship between the fraction of agents taking actions before seeing a large set of signals and the quantities ϵ and δ. Because agents are estimating a normal random variable from noisy observations (where the noise is also normally distributed), their probability of error is captured by the error function $erf(x)$, which is naturally decreasing in the number of observations. In particular, the probability that an agent with k signals takes an action at least ϵ away from the optimal action is no less than $erf\left(\epsilon\sqrt{\frac{k\bar{\rho}}{2}}\right)$ (see Lemma 2 in the Appendix), and this enables us to characterize the fraction of agents that will take an action at least ϵ away from the optimal one in terms of the set $V_k^{n,\sigma}$ as well as ϵ and δ. We thus obtain sufficient conditions for both ϵ, δ-learning to take place and for it to be incomplete. Finally, recall that equilibria and subsequently k-radius sets depend on the discount rate (thus, different discount rates result in different answers for ϵ, δ-learning).

Proposition 1 is stated in terms of the sets $V_k^{n,\sigma}$, which depend on the equilibrium (as the conditioning on σ makes clear). Our next proposition provides a necessary and sufficient condition for perfect asymptotic learning to occur in any equilibrium profile as a function of only exogenous objects, i.e., the perfect k-radius sets, that depend exclusively on the original network structure. Before stating the proposition, we define the notion of *leading agents*. Intuitively, a society contains a set of leading agents if there is a negligible fraction of the agents (the leading agents) whose actions affect the equilibrium behavior of a much larger set of agents (the followers). Let $indeg_i^n = |B_{i,1}^n|$, $outdeg_i^n = \left|\{j \middle| i \in B_{j,1}^n\}\right|$ denote the in-degree, out-degree of agent i in communication network G^n respectively.

Definition 5. *A collection $\{S^n\}_{n=1}^{\infty}$ of sets of agents is called a set of leading agents if*

(i) There exists $k > 0$, such that $S^{n_j} \subseteq V_k^{n_j}$ for all $j \in J$, where J is an infinite index set.

(ii) $\lim_{n\to\infty} \frac{1}{n} \cdot \left|S^n\right| = 0$, i.e., the collection $\{S\}_{n=1}^{\infty}$ contains a negligible fraction of the agents as the

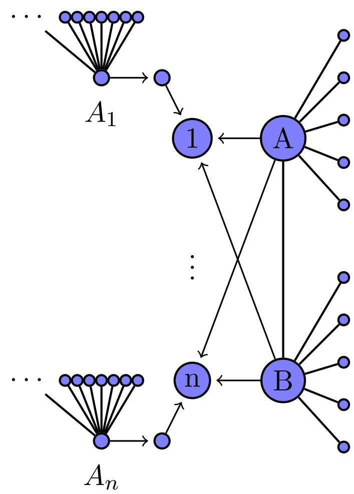


Figure 2: Leading agents and asymptotic learning.

society grows.

(iii) $\lim_{n\to\infty} \frac{1}{n} \cdot \left|S^n_{follow}\right| > \epsilon$*, for some* $\epsilon > 0$*, where* S^n_{follow} *denotes the set of followers of* S^n*. In particular,*

$$S^n_{follow} = \{i \big| \text{ there exists } j \in S^n \text{ such that } j \in B^n_{i,1}\}.$$

Proposition 2. *Suppose Assumption 1 holds. Then,*

(i) Perfect asymptotic learning occurs in society $\{G^n\}_{n=1}^{\infty}$ *in any equilibrium* σ *if*

$$\lim_{k\to\infty}\lim_{n\to\infty}\frac{1}{n}\cdot\left|V^n_k\right| = 0. \tag{5}$$

(ii) Conversely, if condition (5) does not hold for society $\{G^n\}_{n=1}^{\infty}$ *and the society does not contain a set of leading agents, then perfect asymptotic learning does not occur in any equilibrium* σ*.*

Proposition 2 is not stated as an if and only if result because the fact that condition (5) does not hold in a society does not necessarily preclude perfect asymptotic learning in the presence of leading agents. In particular, depending on their actions, a large set of agents may exit early before obtaining enough information to learn, or delay their actions and learn. Figure 2 clarifies this point: if the leading agents (agents A and B) delay their irreversible decision for one communication round, then a large fraction of the rest of the agents (agents 1 to n) may take (depending on the discount rate) an irreversible action as soon as they communicate with the leading agents and their neighbors (i.e., after the second communication round concludes), thus, perfect asymptotic learning fails. However, if the leading agents do not "coordinate," then they exit early and this may lead the rest of the agents to take a delayed (after the third communication round), but more informed action. Generally, in the presence of leading agents, asymptotic learning may occur in all or some of the induced equilibria, even when condition (5) does not hold.

In the rest of this section, we present two corollaries that help clarify the intuition of the asymptotic learning result and identify the role of certain types of agents on information spread in a given society. We focus on perfect asymptotic learning, since we can obtain sharper results, though we can state similar corollaries for ϵ, δ-asymptotic learning for any ϵ and δ. All corollaries are again expressed in terms of the original network topology.[10]

In particular, Corollary 1 identifies a group of agents, that is crucial for a society to permit asymptotic learning: *information mavens*, who have high in-degrees and can thus act as effective aggregators of information (a term inspired by Gladwell (2000)). Information mavens are one type of hubs the importance of which is clearly illustrated by our learning results. Our next definition formalizes this notion.

Definition 6. *Agent i is called an information maven of society $\{G^n\}_{n=1}^{\infty}$ if i has an infinite in-degree, i.e., if*

$$\lim_{n\to\infty} indeg_i^n = \infty.$$

Let $\mathcal{MAVEN}(\{G^n\}_{n=1}^{\infty})$ denote the set of mavens of society $\{G^n\}_{n=1}^{\infty}$.

For any agent j, let $d_j^{\mathcal{MAVEN},n}$ denote the shortest distance defined in communication network G^n between j and a maven $k \in \mathcal{MAVEN}(\{G^n\}_{n=1}^{\infty})$. Finally, let W^n denote the set of agents that are at distance at most equal to their perfect observation radius from a maven in communication network G^n, i.e., $W^n = \{j \mid \; d_j^{\mathcal{MAVEN},n} \leq \tau_j^n\}$.

The following corollary highlights the importance of information mavens for asymptotic learning. Informally, it states that if almost all agents have a short path to a maven, then asymptotic learning occurs.

Corollary 1. *Suppose Assumption 1 holds. Then, asymptotic learning occurs in society $\{G^n\}_{n=1}^{\infty}$ if*

$$\lim_{n\to\infty} \frac{1}{n} \cdot \left|W^n\right| = 1.$$

Corollary 1 thus clarifies that asymptotic learning is obtained when there are information mavens and almost all agents are at a "short distance" away from one (less than their observation radius).

As mentioned in the Introduction, a second type of information hub also plays an important role in asymptotic learning. While mavens have high in-degree and are thus able to effectively aggregate dispersed information, they may not be in the right position to distribute this aggregated information. If so, even in a society that has several information mavens, a large fraction of the agents may not benefit from their information. *Social connectors*, on the other hand, are defined as agents with a high

[10] The corollaries are stated under the additional assumption, that the in-degree of an agent is non-decreasing with n. This is simply a technicality that allows us to simplify the statement of the corollaries.

out-degree, and thus play the role of spreading the information aggregated by the mavens. Before stating the proposition, we define *social connectors.*

Definition 7. *Agent i is called a social connector of society $\{G^n\}_{n=1}^{\infty}$ if i has an infinite out-degree, i.e., if*

$$\lim_{n\to\infty} outdeg_i^n = \infty.$$

The following corollary illustrates the role of social connectors for asymptotic learning.

Corollary 2. *Suppose Assumption 1 holds. Consider a society $\{G^n\}_{n=1}^{\infty}$, such that the set of information mavens does not grow at the same rate as the society itself, i.e.,*

$$\lim_{n\to\infty} \frac{\left|\mathcal{MAVEN}(\{G^n\}_{n=1}^{\infty})\right|}{n} = 0.$$

Then, for asymptotic learning to occur, the society should contain a social connector within a short distance to a maven, i.e.,

$$d_i^{\mathcal{MAVEN},n} \leq \tau_i^n, \text{ for some social connector } i.$$

Corollary 2 thus states that unless a large fraction of the agents belongs to the set of mavens and, subsequently, the rest can obtain information directly from a maven, then, information aggregated at the mavens is spread through the out-links of a connector (note that an agent can be both a maven and a connector). These two corollaries highlight two ways in which society can achieve perfect asymptotic learning. First, it may contain several information mavens who not only collect and aggregate information but also distribute it to almost all the agents in the society. Second, it may contain a sufficient number of information mavens, who pass their information to social connectors, and almost all the agents in the society are a short distance away from social connectors and thus obtain accurate information from them. This latter pattern has a greater plausibility in practice than one in which the same agents collect and distribute dispersed information. For example, if a website or a news source can rely on information mavens (journalists, researchers or analysts) to collect sufficient information and then reach a large number of individuals, then information can be aggregated.

The results summarized in Propositions 1 and 2 as well as in Corollaries 1 and 2 can be seen both as positive and negative, as already noted in the Introduction. On the one hand, communication structures that do not feature information mavens (or connectors) do not lead to perfect asymptotic learning, and information mavens may be viewed as unrealistic or extreme. On the other hand, as already noted above, much communication in modern societies happens through agents that play the role of mavens and connectors (see again Gladwell (2000)). These are highly connected agents that are able to collect and distribute crucial information. Perhaps more importantly, most individuals obtain

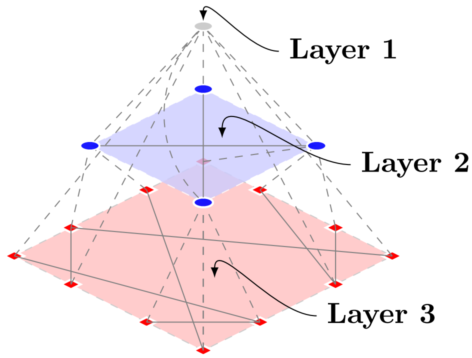


Figure 3: Hierarchical Society.

some of their information from news sources, media, and websites, which exist partly or primarily for the purpose of acting as information mavens and connectors.[11]

3.3 Asymptotic Learning in Random Graphs

As an illustration of the results we outlined in Subsection 3.2, we apply them to *hierarchical graphs*, a class of random graphs defined below. Note that in the present section we assume that communication networks are bidirectional, or equivalently that if agent $i \in B^n_{j,1}$ then $j \in B^n_{i,1}$.

Definition 8 (Hierarchical graphs)**.** *A sequence of communication networks* $\{G^n\}_{n=1}^{\infty}$*, where* $G^n = \{\mathcal{N}^n, \mathcal{E}^n\}$*, is called* ζ*-hierarchical (or simply hierarchical) if it was generated by the following process:*

(i) Agents are born and placed into layers. In particular, at each step $n = 1, \cdots,$ *a new agent is born and placed in layer* ℓ*.*

(ii) Layer index ℓ *is initialized to 1 (i.e., the first node belongs to layer 1). A new layer is created (and subsequently the layer index increases by one) at time period* $n \geq 2$ *with probability* $\frac{1}{n^{1+\zeta}}$*, where* $\zeta > 0$*.*

(iii) Finally, for every n *we have*

$$\mathbb{P}\left((i,j) \in \mathcal{E}^n\right) = \frac{p}{|\mathcal{N}^n_\ell|}, \text{ independently for all } i,j \in \mathcal{N}^n \text{ that belong to the same layer } \ell,$$

where $\mathcal{N}^n_\ell$ *denotes the set of agents that belong to layer* ℓ *at step* n *and* p *scalar, such that* $0 < p < 1$*. Moreover,*

$$\mathbb{P}\left((i,k) \in \mathcal{E}^n\right) = \frac{1}{|\mathcal{N}_{<\ell}|} \text{ and } \sum_{k \in \mathcal{N}_{<\ell}} \mathbb{P}\left((i,k) \in \mathcal{E}^n\right) = 1 \text{ for all } i \in N^n_\ell, k \in N^n_{<\ell}, \ell > 1,$$

where $\mathcal{N}^n_{<\ell}$ *denotes the set of agents that belong to a layer with index lower than* ℓ *at step* n*.*

[11] For example, a news website such as cnn.com acts as a connector that spreads the information aggregated by the journalists-mavens to interested readers. Similarly, a movie review website, e.g., imdb.com, spreads the aggregate knowledge of movie reviewers to interested movie aficionados.

Intuitively, a hierarchical sequence of communication networks resembles a pyramid, where the top contains only a few agents and as we move towards the base, the number of agents grows. The following argument provides an interpretation of the model. Agents on top layers can be thought of as "special" nodes, that the rest of the nodes have a high incentive to connect to. Moreover, agents tend to connect to other agents in the same layer, as they share common features with them (*homophily*). As a concrete example, academia can be thought of as such a pyramid, where the top layer includes the few institutions, then next layer includes academic departments, research labs and finally at the lower levels reside the home pages of professors and students.

Proposition 3. *Suppose Assumption 1 holds and consider society $\{G^n\}_{n=1}^{\infty}$. There exist $\bar{r} > 0$ and a function $\zeta(\eta)$ such that perfect asymptotic learning occurs in society $\{G^n\}_{n=1}^{\infty}$ with probability at least $1 - \eta$, if the sequence of communication networks $\{G^n\}_{n=1}^{\infty}$ is $\zeta(\eta)-$hierarchical and the discount rate $r < \bar{r}$.*

The probability η that perfect asymptotic learning fails is related here to the stochastic process that generated the graph. The results presented provide additional insights on the conditions under which asymptotic learning takes place. It can also be proved that the popular preferential attachment and Erdős-Renyi graphs do not lead to asymptotic learning (we omit these results to save space). This can be interpreted as implying that asymptotic learning is unlikely in several important networks. Nevertheless, these network structures, though often used in practice, do not provide a good description of the structure of many real life networks. In contrast, our results show that asymptotic learning takes place in hierarchical graphs, where "special" agents are likely to receive and distribute information to lower layers of the hierarchy. Although this result is useful in pointing out certain structures where information can be aggregated efficiently, our analysis on the whole suggests that the conditions for both perfect asymptotic learning and for ϵ, δ-learning are somewhat stringent.

3.4 Strategic Communication

Next we explore the implications of relaxing the assumption that agents cannot manipulate the messages they send. In particular, we replace Assumption 1 with the following:

Assumption 2 (Strategic Communication)**.** *Communication between agents is strategic if*

$$m_{ij,t}^{n} \in \mathbb{R}^n,$$

for all agents i, j and time t.

This assumption makes it clear that in this case the messages need not be truthful. Allowing strategic communication adds an extra dimension in an agent's strategy, since the agent can choose to "lie" about (part) of her information set in the hope that this increases her expected payoff. Note that, in contrast with "cheap talk" models, externalities in our framework are purely informational

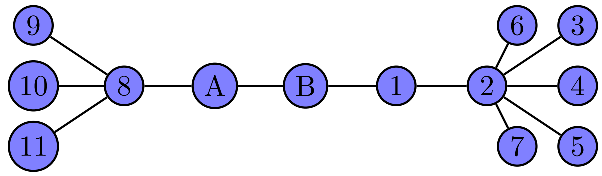


Figure 4: Agents may have an incentive to misreport/not disclose their information.

as opposed to payoff relevant. Thus, an agent may have an incentive to "lie" as a means to obtain more information from the information exchange process (by inducing a later exit decision from her neighbors).

Figure 4 illustrates how incentives for non-truthful communication may arise. Here, agent B may have an incentive not to disclose her information to agent A. In particular, for a set of parameter values we have that if agent B is truthful to A, then A takes an action after the first communication round. On the other hand, if B does not disclose her information to A, then A waits for an additional time period and B obtains access to the information of agents 9, 10 and 11.

Let (σ^n, m^n) denote an action-message strategy profile, where $m^n = \{m_1^n, \cdots, m_n^n\}$ and $m_i^n = [m_{ij,\tau}^n]_{t=0,1,\ldots}$, for j such that $i \in B_{j,1}^n$. Also let $\mathbb{P}_{\sigma^n, m^n}$ refer to the conditional probability when agents behave according to the action-message strategy profile (σ^n, m^n).

Definition 9. *An action-message strategy profile $(\sigma^{n,*}, m^{n,*})$ is a pure-strategy perfect Bayesian Equilibrium of the information exchange game $\Gamma_{info}(G^n)$ if for every $i \in \mathcal{N}^n$ and communication round τ, we have*

$$\mathbb{E}_{(\sigma^{n,*}, m^{n,*})}(U_i^n | I_{i,t}^n) \geq \mathbb{E}_{((\sigma_{i,\tau}^n, \sigma_{i,-\tau}^n, \sigma_{-i}^{n,*}), (m_{i,\tau}^n, m_{i,-\tau}^n m_{-i}^{n,*}))}(U_i^n | I_{i,t}^n),$$

for all $m_{i,\tau}^n, m_{i,-\tau}^n$, and $\sigma_{i,\tau}^n, \sigma_{i,-\tau}^n$. We denote the set of equilibria of this game by $INFO(G^n)$.

Similarly we extend the definitions of asymptotic learning [cf. Definitions 2 and 3]. We show that strategic communication does not harm perfect asymptotic learning. The main intuition behind this result is that it is weakly dominant for an agent to report her private signal truthfully to a neighbor with a high in-degree (maven), as long as others are truthful to the maven.

Proposition 4. *If perfect asymptotic learning occurs in society $\{G^n\}_{n=1}^{\infty}$ under Assumption 1, then there exists an equilibrium (σ, m), such that perfect asymptotic learning occurs in society $\{G^n\}_{n=1}^{\infty}$ along equilibrium (σ, m) when we allow strategic communication (cf. under Assumption 2).*

This proposition therefore implies that the focus on truthful reporting was without much loss of generality as far as perfect asymptotic learning is concerned. In any communication network in which there is perfect asymptotic learning, even if agents can strategically manipulate information, there is arbitrarily little benefit in doing so. Thus, the main lessons about asymptotic learning derived above apply regardless of whether communication is strategic or not.

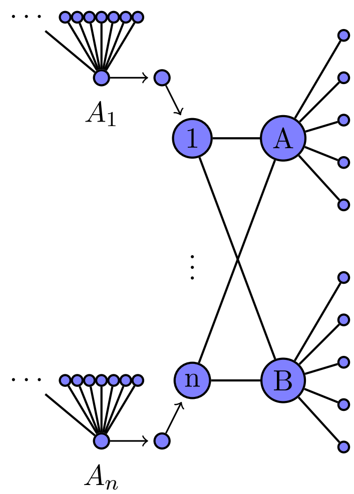


Figure 5: Strategic communication may lead to better actions.

However, this proposition does not imply that all learning outcomes are identical under truthful and strategic communication. In particular, interestingly, as illustrated in Figure 5, strategic communication may lead agents to take a better action with higher probability than under non-strategic communication (cf. Assumption 1). The main reason for this (counterintuitive) fact is that under strategic communication an agent may delay taking an action compared to the non-strategic environment. Therefore, the agent obtains more information from the communication network and, consequently, chooses an action, that is closer to optimal. In particular, in the example illustrated in Figure 5, if agents A, B decide not to disclose their information, then agents $1, \cdots, n$ may delay their action so as to communicate with the neighbors of $A_1, \cdots, A_n$ and thus take an action based on more information.

3.5 Welfare

In this subsection, we turn to the question of *efficient* communication and compare equilibrium allocations (communication and action profiles in equilibrium) with the timing of agents' actions and communications that would be dictated by the welfare-maximizing social planner. We identify conditions under which a social planner can improve over an equilibrium strategy profile. In doing so, we illustrate that communication over social networks might be inefficient because agents do not internalize the positive externality that delaying their action generates for their peers.

A social planner whose objective is to maximize the aggregate expected welfare of the population of n agents can implement the timing profile that is a solution to the optimization:

$$\max_{sp^n} \sum_{i=1}^{n} \mathbb{E}_{sp^n}[U_i^n] \tag{6}$$

We call the resulting timing profile as the *optimal allocation* and we denote it by $sp^n = (\tau_1^{n,sp}, \cdots, \tau_n^{n,sp})$.

Similarly with the asymptotic analysis for equilibria, we define a sequence of optimal allocations for societies of growing size, $sp = \{sp^n\}_{n=1}^{\infty}$. We are interested in identifying conditions under which the social planner can / cannot achieve an *asymptotically* better allocation than an equilibrium (sequence of equilibria) σ, i.e., we are looking at the expression:

$$\lim_{n\to\infty} \frac{\sum_{i\in N^n} \mathbb{E}_{sp^n}[U_i^n] - \sum_{i\in N^n} \mathbb{E}_{\sigma}[U_i^n]}{n}.$$

The next proposition shows a direct connection between learning and efficient communication.

Proposition 5. *Consider society $\{G^n\}_{n=1}^{\infty}$. If perfect asymptotic learning occurs at the optimal allocation $sp = \{sp^n\}_{n=1}^{\infty}$, then all equilibria are asymptotically efficient, i.e.,*

$$\lim_{n\to\infty} \frac{\sum_{i\in N^n} \mathbb{E}_{sp^n}[U_i^n] - \sum_{i\in N^n} \mathbb{E}_{\sigma}[U_i^n]}{n} = 0,$$

for all equilibria σ.

Therefore if perfect learning occurs at the optimal allocation, then perfect learning occurs in all equilibria σ.

We next provide a partial converse to Proposition 5. Before stating this result, we contrast the decision problem an individual agent i with that of the social planner. With a slight abuse of notation, $U_i^n(k,\sigma)$ denotes the expected payoff of agent i when agents behave according to profile σ and the agent has observed k signals. Agent i decides to take an irreversible action at time t and not to wait for an additional dt, when other agents behave according to σ, if (cf. Appendix)

$$\frac{r+\lambda}{\lambda}\left(\pi - \frac{1}{\rho + \bar{\rho} k_{i,t}^{n,\sigma}}\right) \geq U_i^n(k_{i,t}^{n,\sigma} + |B_{i,|T_t|+1}^{n,\sigma}| - |B_{i,|T_t|}^{n,\sigma}|, \sigma) \tag{7}$$

Similarly, in the corresponding optimal allocation agent i exits at time t and does not wait if:

$$\frac{r+\lambda}{\lambda}\left(\pi - \frac{1}{\rho + \bar{\rho} k_{i,t}^{n,sp}}\right) \\ \geq U_i^n(k_{i,t}^{n,sp} + |B_{i,|T_t|+1}^{n,sp}| - |B_{i,|T_t|}^{n,sp}|, sp) + \sum_{j\neq i} \mathbb{E}_{sp}[U_j^n \big| i \text{ “waits” at } t] - \mathbb{E}_{sp}[U_j^n \big| i \text{ “exits” at } t], \tag{8}$$

The comparison of (7) to (8) shows the reason for why equilibria may be inefficient in this setting: when determining when to act, agent i does not take into account the positive externality that a later action exerts on others. This externality is expressed by the summation on the right hand side of (8). We next derive sufficient conditions under which a social planner outperforms an equilibrium allocation σ. Consider agents i and j such that $i \in B_{j,1}^n$ and $\tau_j^{n,\sigma} > \tau_i^{n,\sigma} + 1$, which implies that $B_{j,\tau_j^{n,\sigma}}^n \supset B_{i,\tau_i^{n,\sigma}}^n$ (i.e., agent j communicates with a superset of the agents that i communicates with before taking an action). Also, let $k_{ij,\tau_i^{n,\sigma}}^{n,\sigma}$ denote the additional agents that j would observe if i delayed her irreversible action by dt and communication took place. Then, the aggregate welfare of the two agents increases if the following condition holds:

$$U_j^n(k_{j,\tau_j^{n,\sigma}}^{n,\sigma} + k_{ij,\tau_i^{n,\sigma}}^{n,\sigma}) + U_i^n(k_{i,\tau_i^{n,\sigma}}^{n,\sigma} + k_{ij,\tau_i^{n,\sigma}}^{n,\sigma}) > U_j^n(k_{j,\tau_j^{n,\sigma}}^{n,\sigma}) + \frac{r+\lambda}{\lambda} U_i^n(k_{i,\tau_i^{n,\sigma}}^{n,\sigma}), \tag{9}$$

Let set $D_{k,\ell}^{n,\sigma}$ denote the following set of agents: $j \in D_{k,\ell}^{n,\sigma}$, if

(i) $k_{j,\tau_j^{n,\sigma}}^{n,\sigma} \leq k$.

(ii) There exists an agent $i \in B_{j,1}^{n,\sigma}$ such that

  (i) $\tau_j^{n,\sigma} > \tau_i^{n,\sigma} + 1$.

  (ii) If i exits at $\tau_i^{n,\sigma} + 1$, then j gains access to at least an additional ℓ signals.

Intuitively, set $D_{k,\ell}^{n,\sigma}$ contains agents that would obtain higher payoff in expectation if one of their neighbors delayed taking her irreversible action. In particular, under equilibrium profile σ, agent $j \in D_{k,\ell}^{n,\sigma}$ takes an action after observing at most k signals. If her neighbor i delayed her action by one communication round, then she would have access to at least $k + \ell$ signals by the time of her action.

The following proposition provides a sufficient condition for an equilibrium to be inefficient.

Proposition 6. *Consider society $\{G^n\}_{n=1}^{\infty}$ and equilibrium $\sigma = \{\sigma^n\}_{n=1}^{\infty}$. Assume that $\lim_{n\to\infty} \frac{|D_{k,\ell}^{n,\sigma}|}{n} > \xi > 0$, for k, ℓ that satisfy the following:*

$$\frac{r}{\lambda}\pi + \frac{2}{\rho + \bar{\rho}(k+\ell)} < \left(2 + \frac{r}{\lambda}\right)\frac{1}{\rho + \bar{\rho}k}.$$

Then, there exists an $\zeta > 0$, such that

$$\lim_{n\to\infty} \frac{\sum_{i\in N^n} \mathbb{E}_{sp^n}[U_i^n] - \sum_{i\in N^n} \mathbb{E}_{\sigma}[U_i^n]}{n} > \zeta,$$

i.e., equilibrium σ is asymptotically inefficient. Moreover, there exist ϵ, δ such that ϵ, δ-asymptotic learning fails at equilibrium σ.

We close this section with a discussion on the implications of increasing the information that agents have access to at the beginning of the information exchange process. Consider the following setting: agents at time $t = 0$ have access to k public signals in addition to their private signal. This results in the following tradeoff: on the one hand, agents are better informed about the underlying state, but then, on the other hand, they will have less incentive to delay taking an action and thus obtain and share information with others. In particular, one can show that when all agents have access to the same k public signals, then information sharing will be reduced compared to a setting without public signals, in the sense that agents take an irreversible action earlier. Moreover, in some cases the presence of public signals leads to a strictly smaller aggregate welfare. Thus, more information is not necessarily better for the aggregate welfare of the agents. This result is similar to those in Duffie, Malamud, and Manso (2009) and in Morris and Shin (2002), both of which show how greater availability of public information may reduce welfare.

4 Network Formation

We have so far studied information exchange among agents over a given communication network $G^n = (\mathcal{N}^n, \mathcal{E}^n)$. We now analyse how this communication network emerges. We assume that link formation is costly. In particular, *communication costs* are captured by an $n \times n$ nonnegative matrix C^n, where C^n_{ij} denotes the cost that agent i has to incur in order to form the directed link (j,i) with agent j. As noted previously, a link's direction coincides with the direction of the flow of messages. In particular, agent i incurs a cost to form in-links. We refer to C^n as the *communication cost matrix*. We assume that $C^n_{ii} = 0$ for all $i \in \mathcal{N}^n$. Our goal in this section is to provide conditions under which the network structures that emerge as equilibria of the network formation game defined below guarantee asymptotic learning. Our results indicate that easy access to information (i.e., low cost to form links with some information sources) may preclude asymptotic learning, as it reduces the incentives for further information sharing. Moreover, asymptotic learning may depend on how well agents coordinate at equilibrium: we show that there may be multiple equilibria that induce sparser/denser network structures and lead to different answers for asymptotic learning.

We define agent i's *link formation strategy*, g^n_i, as an n-tuple such that $g^n_i \in \{0,1\}^n$ and $g^n_{ij} = 1$ implies that agent i forms a link with agent j. The cost agent i has to incur if she implements strategy g^n_i is given by

$$\text{Cost}(g^n_i) = \sum_{j \in \mathcal{N}} C^n_{ij} \cdot g^n_{ij}.$$

The link formation strategy profile $g^n = (g^n_1, \cdots, g^n_n)$ induces the communication network $G^n = (\mathcal{N}^n, \mathcal{E}^n)$, where $(j,i) \in \mathcal{E}^n$ if and only if $g^n_{ij} = 1$.

We extend our environment to the two-stage *Network Learning Game* $\Gamma(C^n)$, where C^n denotes the communication cost matrix. The two stages of the network learning game can be described as follows:

Stage 1 [*Network Formation Game*]: Agents choose their link formation strategies. The link formation strategy profile g^n induces the communication network $G^n = (\mathcal{N}^n, \mathcal{E}^n)$.

We refer to stage 1 of the network learning game, when the communication cost matrix is C^n as the *network formation game* and we denote it by $\Gamma_{net}(C^n)$.

Stage 2 [*Information Exchange Game*]: Agents communicate over the induced network G^n as studied in previous sections.

We next define the equilibria of the network learning game $\Gamma(C^n)$. Note that we use the standard notation g_{-i} and σ_{-i} to denote the strategies of agents other than i. Also, we let $\sigma_{i,-t}$ denote the vector of actions of agent i at all times except t.

Definition 10. *A pair $(g^{n,*}, \sigma^{n,*})$ is a pure-strategy perfect Bayesian Equilibrium of the network learning game $\Gamma(C^n)$ if*

(a) $\sigma^{n,*} \in INFO(G^n)$, where G^n is induced by the link formation strategy $g^{n,*}$.

(b) For all $i \in \mathcal{N}^n$, $g_i^{n,*}$ maximizes the expected payoff of agent i given the strategies of other agents $g_{-i}^{n,*}$, i.e.,

$$g_i^{n,*} \in \arg\max_{g_i^n \in \{0,1\}^n} \mathbb{E}_\sigma[\Pi_i(g_i^n, g_{-i}^{n,*})] \equiv \mathbb{E}_\sigma(U_i^n \big| I_{i,0}^n) - Cost(g_i^n).$$

for all $\sigma \in INFO(\tilde{G}^n)$, where $\tilde{G}^n$ is induced by link formation strategy $(g_i^n, g_{-i}^{n,*})$.

We denote the set of equilibria of this game by $NET(C^n)$.

Similar to the analysis of the information exchange game, we consider a sequence of communication cost matrices $\{C^n\}_{n=1}^\infty$, where for fixed n,

$$C^n : \mathcal{N}^n \times \mathcal{N}^n \to \mathcal{R}^+ \text{ and } C_{ij}^n = C_{ij}^{n+1} \text{ for all } i, j \in \mathcal{N}^n. \tag{10}$$

For the remainder of the section, we focus our attention to the *social cliques communication cost structure*. The properties of this communication structure are stated in the next assumption.

Assumption 3. *Let $c_{ij}^n \in \{0, c\}$ for all pairs $(i,j) \in \mathcal{N}^n \times \mathcal{N}^n$, where $c < \frac{1}{\rho+\bar{\rho}}$. Moreover, let $c_{ij} = c_{ji}$ for all $i, j \in \mathcal{N}^n$ (symmetry), and $c_{ij} + c_{jk} \geq c_{ik}$ for all $i, j, k \in \mathcal{N}^n$ (triangular inequality).*

The assumption that $c < \frac{1}{\rho+\bar{\rho}}$ rules out the degenerate case where no agent forms a costly link. The symmetry and triangular inequality assumptions are imposed to simplify the definition of a social clique, which is introduced next. Suppose Assumption 3 holds. We define a *social clique* (cf. Figure 6) $H^n \subset \mathcal{N}^n$ as a set of agents such that

$$i, j \in H^n \quad \text{if and only if} \quad c_{ij} = c_{ji} = 0.$$

Note that this set is well-defined since, by the triangular inequality and symmetry assumptions, if an agent i does not belong to social clique H^n, then $c_{ij} = c$ for all $j \in H^n$. Hence, we can uniquely partition the set of nodes $\mathcal{N}^n$ into a set of K^n pairwise disjoint social cliques $\mathcal{H}^n = \{H_1^n, \cdots, H_{K^n}^n\}$. We use the notation $\mathcal{H}_k^n$ to denote the set of pairwise disjoint social cliques that have cardinality greater than or equal to k, i.e., $\mathcal{H}_k^n = \{H_i^n, i = 1, \dots, K^n \mid |H_i^n| \geq k\}$. We also use $SC^n(i)$ to denote the social clique that agent i belongs to.

We consider a sequence of communication cost matrices $\{C^n\}_{n=1}^\infty$ satisfying condition (10) and Assumption 3, and we refer to this sequence as a *communication cost structure*. As shown above, the communication cost structure $\{C^n\}_{n=1}^\infty$ uniquely defines the following sequences, $\{\mathcal{H}^n\}_{n=1}^\infty$ and

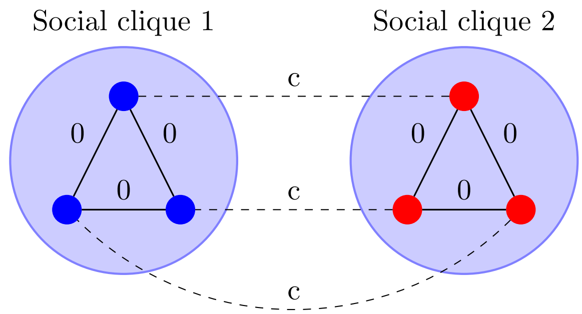


Figure 6: Social cliques.

$\{\mathcal{H}_k^n\}_{n=1}^\infty$ for $k > 0$, of sets of pairwise disjoint social cliques. Moreover, it induces network equilibria $(g, \sigma) = (g^n, \sigma^n)_{n=1}^\infty$ such that $(g^n, \sigma^n) \in NET(C^n)$ for all n. Our goal is to identify conditions on the communication cost structure that lead to the emergence of networks which guarantee asymptotic learning. We focus entirely on perfect asymptotic learning, as this enables us to obtain sharp results. Similar results can be obtained for ϵ, δ-asymptotic learning.

Proposition 7. *Let $\{C^n\}_{n=1}^\infty$ be a communication cost structure and let Assumptions 1 and 3 hold. Then, there exists a constant $\bar{k} = \bar{k}(c)$ such that the following hold:*

(a) Suppose that

$$\limsup_{n\to\infty} \frac{\left|\mathcal{H}_{\bar{k}}^n\right|}{n} \geq \epsilon \text{ for some } \epsilon > 0. \tag{11}$$

Then, perfect asymptotic learning does not occur in any network equilibrium (g, σ).

(b) Suppose that

$$\lim_{n\to\infty} \frac{\left|\mathcal{H}_{\bar{k}}^n\right|}{n} = 0 \text{ and } \lim_{n\to\infty} \left|H_\ell^n\right| = \infty \text{ for some } \ell. \tag{12}$$

Then, perfect asymptotic learning occurs in all network equilibria (g, σ) when the discount rate r satisfies $0 < r < \bar{r}$, where $\bar{r} > 0$ is a constant.

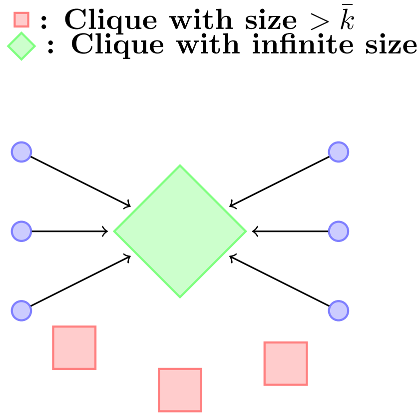
(c) Suppose that there exists $M > 0$ such that

$$\lim_{n\to\infty} \frac{\left|\mathcal{H}_{\bar{k}}^n\right|}{n} = 0 \text{ and } \limsup_{n\to\infty} \left|H_\ell^n\right| < M \text{ for all } \ell, \tag{13}$$

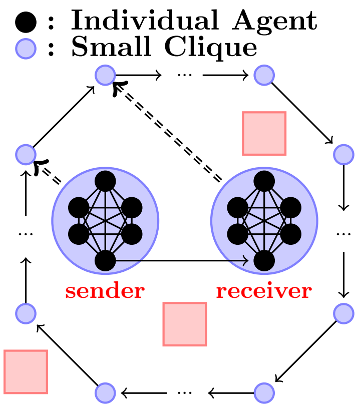
and let agents be patient, i.e., consider the case, when the discount rate $r \to 0$. Then, there exists a $\bar{c} > 0$ such that

(i) If $c \leq \bar{c}$, perfect asymptotic learning occurs in all network equilibria (g, σ).

(ii) If $c > \bar{c}$, there exists at least one network equilibrium (g, σ), where there is no perfect



(a) Equilibrium network, when (12) holds.



(b) Equilibrium network, when (13) holds.

Figure 7: Network formation among social cliques.

asymptotic learning and there exists at least one network equilibrium (g, σ) *where perfect asymptotic learning occurs.*

The results in this proposition provide a fairly complete characterization of what types of environments lead to the formation of networks that subsequently induce perfect asymptotic learning. The key concept is that of a *social clique*, which represents groups of individuals that are linked to each other at zero cost. These can be thought of as "friendship networks," which are linked for reasons unrelated to information exchange and thus can act as conduits of such exchange at low cost. Agents can exchange information without incurring any costs (beyond the delay necessary for obtaining information) within their social cliques. However, if they wish to obtain further information, from outside their social cliques, they have to pay a cost at the beginning in order to form a link. Even though network formation games have several equilibria, the structure of our network formation and information exchange game enables us to obtain relatively sharp results on what types of societies lead to endogenously formed communication networks that ensure perfect asymptotic learning. In particular, the first part of Proposition 7 shows that perfect asymptotic learning cannot occur in any equilibrium if the number of sufficiently large social cliques increases at the same rate as the size of the society. This is intuitive; when this is the case, there are many social cliques of sufficiently large size that none of their members wish to engage in further costly communication with members of other social cliques. But since several of these do not contain an information hub social learning is precluded.

In contrast, the second part of the proposition shows that if the number of disjoint and sufficiently large social cliques is limited (grows less rapidly than the size of the society) and some of them are large enough to contain information hubs, then perfect asymptotic learning takes place (provided that future is not heavily discounted). In this case, as shown by Figure 7(a), sufficiently many social cliques connect to the larger social cliques acting as information hubs, ensuring effective aggregation of information for the great majority of the agents in the society. It is important that the discount factor is not too small, otherwise smaller cliques do not find it beneficial to form links with the larger cliques.

Finally, the third part of the proposition outlines a more interesting configuration, potentially leading to perfect asymptotic learning. In this case, many small social cliques form an "informational ring"(Figure 7(b)). Each is small enough that it finds it beneficial to connect to another social clique, provided that this other clique also connects to others and obtain further information. This intuition also clarifies why such information aggregation takes place only in some equilibria. The expectation that others do not form the requisite links leads to a coordination failure. Interestingly, however, if agents are sufficiently patient and the cost of link formation is not too large, the coordination failure equilibrium disappears, because it becomes beneficial for each clique to form links with another one, even if further links are not forthcoming. Finally, the ring structure is a direct consequence of the fact that agents are patient (and has been shown to emerge as an equilibrium configuration in other models of network formation, e.g., Bala and Goyal (2000))

5 Conclusion

We have developed a framework for the analysis of information exchange through communication and investigated its implications for information aggregation in large societies. An underlying state determines the payoffs from different actions. Agents decide which agents to form a communication link with incurring the associated cost. After receiving a private signal correlated with the underlying state, they exchange information over the induced communication network until taking an (irreversible) action.

Our focus has been on asymptotic learning, defined as the fraction of agents taking the correct action converging to one in probability as a society grows large. We showed that asymptotic learning occurs if and, under some additional mild assumptions, only if the induced communication network includes information hubs and most agents are at a short distance from a hub. Thus asymptotic learning requires information to be aggregated in the hands of a few agents. This kind of aggregation also requires truthful communication, which we show is an equilibrium of the strategic communication in large societies (partly as a consequence of the fact there is no conflict among the agents concerning which action is best).

Our analysis also provides a systematic investigation of what types of cost structures, and associated social cliques which consist of groups of individuals linked to each other at zero cost (such as friendship networks), ensure the emergence of communication networks that lead to asymptotic learning. Our main result on network formation shows that societies with too many (disjoint) and sufficiently large social cliques do not form communication networks that lead to asymptotic learning, because each social clique would have sufficient information to make communication with others not sufficiently attractive. Asymptotic learning results if social cliques are neither too numerous nor too large so as to encourage communication across cliques. Our analysis was conducted under a simplifying assumption that all agents have the same preferences. Interesting avenues for research include investigation of similar dynamic models of information exchange and network formation in the presence of ex ante or ex post heterogeneity of preferences as well as differences in the quality of information available to different agents, which may naturally lead to the emergence of hubs.

Appendix

Proofs from Section 3

Proof of Lemma 1.

Recall that, by the principle of optimality, agent i's optimal continuation payoff at information set $I^n_{i,t}$, when the rest of the agents behave according to strategy profile σ, is given by:

$$\mathbb{E}_\sigma(U^n_i \big| I^n_{i,t}) = \max \begin{cases} \pi - \frac{1}{\rho + \bar{\rho} k^{n,\sigma}_{i,t}} & \text{(when she takes the optimal irreversible action)}, \\ e^{-rdt}\ \mathbb{E}[\mathbb{E}_\sigma(U^n_i | I^n_{i,t+dt}) \big| I^n_{i,t}] & \text{(when she decides to wait, i.e., } x = \text{“}wait\text{”)}, \end{cases}$$

where $k^{n,\sigma}_{i,t}$ denotes the number of distinct private signals agent i has observed up to time t. The first line is equal to the expected payoff for the agent when she chooses the optimal irreversible action under information set $I^n_{i,t}$, i.e., $\mathbb{E}[\theta | I^n_{i,t}]$, and she has observed $k^{n,\sigma}_{i,t}$ private signals, while the second line is equal to the discounted expected continuation payoff.

For the latter, we have that with probability λdt, communication takes place in time interval $[t, t+dt]$, thus the information set of agent i expands; with probability $(1 - \lambda dt)$, there is no communication and the value function for agent i remains unchanged. If communication takes place in interval $[t, t+dt]$, then agent i observes $|B^{n,\sigma}_{i,|T_t|+1}| - |B^{n,\sigma}_{i,|T_t|}|$ additional signals.

Note that since we assume that signals are identically distributed and independent, the value function can simply be expressed as a function of the number of distinct signals in $I^n_{i,t}$, $k^{n,\sigma}_{i,t}$ and profile σ. The agent chooses to take an irreversible action and not to wait if

$$\begin{aligned} \pi - \frac{1}{\rho + \bar{\rho} k^{n,\sigma}_{i,t}} &\geq e^{-rdt}\ \mathbb{E}[\mathbb{E}_\sigma(U^n_i | I^n_{i,t+dt}) \big| I^n_{i,t}] \\ &\geq e^{-rdt} \left(\lambda dt U^n_i (k^{n,\sigma}_{i,t} + |B^{n,\sigma}_{i,|T_t|+1}| - |B^{n,\sigma}_{i,|T_t|}|, \sigma) + (1 - \lambda dt) U^n_i (k^{n,\sigma}_{i,t}, \sigma) \right). \end{aligned}$$

Thus, we obtain that the agent chooses not to wait if:

$$U^n_i (k^{n,\sigma}_{i,t} + |B^{n,\sigma}_{i,|T_t|+1}| - |B^{n,\sigma}_{i,|T_t|}|, \sigma) \leq \frac{r + \lambda}{\lambda} \left(\pi - \frac{1}{\rho + \bar{\rho} k^{n,\sigma}_{i,t}} \right). \tag{14}$$

The left hand side of (14) is upper bounded by π, whereas the right hand side is increasing in the number of private signals $k^{n,\sigma}_{i,t}$ and in the limit is equal to $\frac{r+\lambda}{\lambda}\pi > \pi$. This establishes the lemma. ∎

The next lemma will be used in the rest of the Appendix. It shows that the probability of choosing an action that is more than ϵ away from the optimal for agent $i \in V^{n,\sigma}_k$, i.e., $\mathbb{P}_\sigma\ (M^{n,\epsilon}_i = 0)$, is uniformly bounded away from 0 in terms of the error function.

Lemma 2. *Let $k > 0$ be a constant, such that the k-radius set $V^{n,\sigma}_k$ is non-empty. Then,*

$$\mathbb{P}(M^{n,\epsilon}_i = 0) \geq erf\left(\epsilon \sqrt{\frac{k\bar{\rho}}{2}} \right) \ \textit{for all } i \in V^n_{k,\sigma},$$

where $erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$ *is the error function.*

Proof. Note that because of our normality assumption the empirical mean $\hat{\theta}$ after observing ℓ private signals is normally distributed around θ with precision $\rho_{\hat{\theta}} = \ell\bar{\rho}$. Then, the probability that $M_i^{n,\epsilon} = 0$ is simply equal to the probability that the error does not belong to the interval $[-\epsilon, \epsilon]$, i.e.,

$$\mathbb{P}(M_i^{n,\epsilon} = 0) = erf\left(\epsilon\sqrt{\frac{\ell\bar{\rho}}{2}}\right).$$

The lemma follows since agent $i \in V_k^{n,\sigma}$, thus she takes an irreversible action after observing at most k private signals. ■

Proof of Proposition 1. First, we show that learning fails if condition (3) holds, i.e., there exists a $k > 0$, such that

$$\eta = \limsup_{n\to\infty} \frac{1}{n} \cdot \left|V_k^{n,\sigma}\right| > \epsilon \text{ and } erf\left(\epsilon\sqrt{\frac{k\bar{\rho}}{2}}\right) < (1-\delta)(1-\epsilon/\eta). \tag{15}$$

From condition (15) we obtain that there exists an infinite index set J such that

$$\left|V_k^{n_j}\right| \geq \eta \cdot n_j \text{ for } j \in J.$$

Now restrict attention to index set J, i.e., consider $n = n_j$ for some $j \in J$. Then,

$$\begin{aligned} \mathbb{P}_\sigma\left(\tfrac{1}{n}\textstyle\sum_{i=1}^n M_i^{n,\epsilon} > 1-\epsilon\right) &= \mathbb{P}_\sigma\left(\tfrac{1}{n}\left[\textstyle\sum_{i\in V_k^{n,\sigma}} M_i^{n,\epsilon} + \sum_{i\notin V_k^{n,\sigma}} M_i^{n,\epsilon}\right] > 1-\epsilon\right) \\ &\leq \mathbb{P}_\sigma\left(\tfrac{1}{n}\left[\textstyle\sum_{i\in V_k^{n,\sigma}} M_i^{n,\epsilon} + n - \left|V_k^{n,\sigma}\right|\right] > 1-\epsilon\right) \\ &= \mathbb{P}_\sigma\left(\tfrac{1}{n}\textstyle\sum_{i\in V_k^{n,\sigma}} M_i^{n,\epsilon} > \frac{\left|V_k^{n,\sigma}\right|}{n} - \epsilon\right) \end{aligned},$$

where the inequality follows since we let $M_i^{n,\epsilon} = 1$ for all $i \notin V_k^{n,\sigma}$. Next we use Markov's inequality

$$\mathbb{P}_\sigma\left(\frac{1}{n}\sum_{i\in V_k^{n,\sigma}} M_i^{n,\epsilon} > \frac{\left|V_k^{n,\sigma}\right|}{n} - \epsilon\right) \leq \frac{\mathbb{E}_\sigma\left[\sum_{i\in V_k^{n,\sigma}} M_i^{n,\epsilon}\right]}{n \cdot \left(\left|V_k^{n,\sigma}\right|/n - \epsilon\right)}.$$

We can view each summand above as an independent Bernoulli variable with success probability bounded above by $erf\left(\epsilon\sqrt{\frac{k\bar{\rho}}{2}}\right)$ from Lemma 2. Thus,

$$\begin{aligned} \frac{\mathbb{E}_\sigma\left[\sum_{i\in V_k^{n,\sigma}} M_i^{n,\epsilon}\right]}{n\cdot\left(\left|V_k^{n,\sigma}\right|/n-\epsilon\right)} &\leq \frac{\left|V_k^{n,\sigma}\right| erf\left(\epsilon\sqrt{\frac{k\bar{\rho}}{2}}\right)}{n\cdot\left(\left|V_k^{n,\sigma}\right|/n-\epsilon\right)} \\ &\leq \frac{\eta}{\eta-\epsilon} erf\left(\epsilon\sqrt{\frac{k\bar{\rho}}{2}}\right) < 1-\delta, \end{aligned}$$

where the second inequality follows from the fact that n was chosen such that $\left|V_k^{n,\sigma}\right| \geq \eta \cdot n$. Finally, the last expression follows from the choice of k (cf. Condition (3)). We obtain that for all $j \in J$ it holds that

$$\mathbb{P}_\sigma\left(\left[\frac{1}{n_j}\sum_{i=1}^{n_j}\left(1-M_i^{n_j,\epsilon}\right)\right]>\epsilon\right)\geq\delta.$$

Since J is an infinite index set we conclude that

$$\liminf_{n\to\infty}\mathbb{P}_\sigma\left(\left[\frac{1}{n}\sum_{i=1}^{n}(1-M_i^{n,\epsilon})\right]>\epsilon\right)\geq\delta,$$

thus ϵ,δ-asymptotic learning is incomplete when (3) holds.

Next, we prove that Condition (4) is sufficient for ϵ,δ-asymptotic learning. As mentioned above, if agent i takes an irreversible action after observing ℓ signals, then the probability that $M_i^{n,\epsilon}=1$ is equal to

$$\mathbb{P}_\sigma(M_i^{n,\epsilon}=1)=erf\left(\epsilon\sqrt{\frac{\ell\bar{\rho}}{2}}\right). \tag{16}$$

Similarly with above, we have

$$\begin{aligned}\mathbb{P}_\sigma\left(\left[\frac{1}{n}\sum_{i=1}^{n}(1-M_i^{n,\epsilon})\right]>\epsilon\right) &\leq \mathbb{P}_\sigma\left(\left[\frac{1}{n}\sum_{i\notin V}(1-M_i^{n,\epsilon})\right]>\epsilon-\frac{|V|}{n}\right)\\ &\leq \frac{\mathbb{E}_\sigma\left[\sum_{i\notin V}(1-M_i^{n,\epsilon})\right]}{n\left(\epsilon-|V|/n\right)},\end{aligned} \tag{17}$$

where $V=\left\{i\mid\ \left|B^n_{i,\tau_i^{n,\sigma}}\right|\leq k\right\}$ and the second inequality follows from Markov's inequality. By combining Eqs. (16) and (17) and letting $k_i^{n,\sigma}$ denote the number of private signals that agent i observed before taking an action,

$$\frac{\mathbb{E}_\sigma\left[\sum_{i\notin V}(1-M_i^{n,\epsilon})\right]}{n\left(\epsilon-|V|/n\right)}\leq\frac{\sum_{i\notin V}1-erf\left(\epsilon\sqrt{\frac{k_i^{n,\sigma}\bar{\rho}}{2}}\right)}{n\left(\epsilon-|V|/n\right)}. \tag{18}$$

We have

$$erf\left(\epsilon\sqrt{\frac{k_i^{n,\sigma}\bar{\rho}}{2}}\right)>1-\frac{\delta(\epsilon-\zeta)}{1-\zeta}, \tag{19}$$

for all $i\notin V$ from the definition of k (cf. Condition (4)). Thus, combining Eqs. (17),(18) and (19), we obtain

$$\mathbb{P}_\sigma\left(\left[\frac{1}{n}\sum_{i=1}^{n}(1-M_i^{n,\epsilon})\right]>\epsilon\right)<\delta \text{ for all } n>N,$$

where N is a sufficiently large constant, which implies that condition (4) is sufficient for asymptotic learning. ■

Similar to perfect k-radius sets, we define sets X_k^n for scalar k as

$$X_k^n=\{i\in\mathcal{N}^n\,\big|\text{ there exists }\ell\in B^n_{i,\tau_i^n}\text{ with }indeg_\ell^n>k\},$$

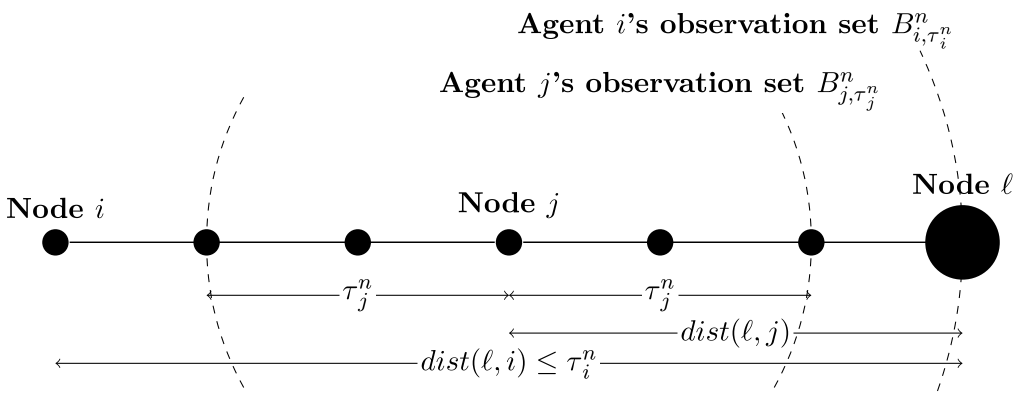


Figure 8: Proof of Proposition 8.

i.e., the set X_k^n consists of all agents, which have an agent with in-degree at least k within their perfect observation radius.

Proposition 8. *Suppose Assumption 1 holds. Then, perfect asymptotic learning occurs in society* $\{G^n\}_{n=1}^{\infty}$ *in any equilibrium* σ *if*

$$\lim_{k\to\infty}\lim_{n\to\infty}\frac{1}{n}|X_k^n| = 1.$$

Proof. Consider equilibrium profile σ and society $\{G^n\}_{n=1}^{\infty}$ such that $\lim_{k\to\infty}\lim_{n\to\infty}\frac{1}{n}|X_k^n| = 1$.

Define $Z_k^{n,\sigma}$ as the following set of agents

$$Z_k^{n,\sigma} = \{i \in \mathcal{N}^n \big| \text{ there exists } \ell \in B^{n,\sigma}_{i,\tau_i^{n,\sigma}} \text{ with } indeg_\ell^n > k\},$$

i.e., the agents that at equilibrium σ, communicate with an agent with in-degree at least k. Next, we show that for k large enough (and consequently n large enough), $X_k^n = Z_k^{n,\sigma}$.

Consider $i \in X_k^n$ and let $\mathcal{P}^n = \{\ell, i_1, \cdots, i_K, i\}$ denote the shortest path in communication network G^n between i and any agent ℓ, with $indeg_\ell^n \geq k$. First we show the following (refer to Figure 8)

$$i \in X_k^n \Rightarrow j \in X_k^n \text{ for all } j \in \mathcal{P}^n. \tag{20}$$

Assume for the sake of contradiction that condition (20) does not hold. Then, let

$$j = \arg\min_{j'}\{dist^n(\ell, j')\big| j' \in \mathcal{P}^n \text{ and } dist^n(\ell, j') > \tau_{j'}^n\},$$

where recall that τ_i^n denotes the perfect observation radius of agent i. For agents i, j we have $\tau_i^n > \tau_j^n$ and $dist(j,i) + d_j^n < dist(\ell, i) \leq \tau_i^n$, since otherwise $j \in X_k^n$. This implies that $B^n_{j,\tau_j^n} \subset B^n_{i,\tau_i^n}$. Furthermore,

$$\pi - \frac{1}{\rho + \bar{\rho}(|B^n_{j,\tau_j^n}|)} > \left(\frac{\lambda}{\lambda + r}\right)^{dist(\ell,j)-\tau_j^n}\left(\pi - \frac{1}{\rho + \bar{\rho}(|B^n_{j,\tau_j^n}| + k)}\right), \tag{21}$$

In particular, the left hand side is equal to the expected payoff of agent j if she takes an irreversible action at time τ_j^n after receiving $|B^n_{j,\tau_j^n}|$ observations, whereas the right hand side is a lower bound on the expected payoff if agent j delays taking an action until after she communicates with agent ℓ. The inequality follows, from the definition of the observation radius for agent j. On the other hand, since for agent i, $\ell \in B^n_{i,\tau_i^n}$, we have

$$\pi - \frac{1}{\rho + \bar{\rho}(|B^n_{j,\tau_j^n}|)} < \left(\frac{\lambda}{\lambda + r}\right)^{dist(\ell,i)-dist(j,i)-\tau_j^n} \left(\pi - \frac{1}{\rho + \bar{\rho}(|B^n_{j,\tau_j^n}| + k + k')}\right), \text{ for some } k' > 0. \quad (22)$$

For k large enough we conclude that $dist(\ell, j) < dist(\ell, i) - dist(j, i)$, which is obviously a contradiction. This implies that (20) holds.

Next we show, by induction on the distance from agent ℓ with in-degree $\geq k$ that $X_k^n = Z_k^{n,\sigma}$ for equilibrium σ. The claim is obviously true for all agents with distance equal to 0 (agent ℓ) and 1 (her neighbors). Assume that the claim holds for all agents with distance at most t from agent ℓ, i.e., if $i \in X_k^n$ and $dist(\ell, i) \leq t$ then $i \in Z_k^{n,\sigma}$. Finally, we show the claim for an agent i such that $i \in X_k^n$ and $dist(\ell, i) = t + 1$. Consider a shortest path $\mathcal{P}^n$ from i to ℓ. Condition (20) implies that all agents j in the shortest path are such that $j \in X_k^n$, thus from the induction hypothesis we obtain $j \in Z_k^{n,\sigma}$. Thus, for k sufficiently large we obtain that $i \in Z_k^{n,\sigma}$, for any equilibrium σ.

Finally, by the hypothesis of the proposition, i.e., $\lim_{k\to\infty} \lim_{n\to\infty} \frac{1}{n}|X_k^n| = 1$, we conclude that $\lim_{k\to\infty} \lim_{n\to\infty} \frac{1}{n}|Z_k^{n,\sigma}| = 1$, for any equilibrium σ. The latter implies that $\lim_{k\to\infty} \lim_{n\to\infty} \frac{1}{n}|V_k^{n,\sigma}| = 0$, thus asymptotic learning occurs along equilibrium σ from Proposition 1. ■

Proof of Proposition 2.

The first part of Proposition 2 follows directly from Proposition 8, since

$$\lim_{k\to\infty} \lim_{n\to\infty} \frac{1}{n}|V_k^n| = 0 \Rightarrow \lim_{k\to\infty} \lim_{n\to\infty} \frac{1}{n}|X_k^n| = 1.$$

To conclude the proof we need to show that if asymptotic learning occurs along some equilibrium σ when condition (4) does not hold, then the society contains a set of leading agents. In particular, consider a society $\{G^n\}_{n=1}^\infty$ in which condition (4) does not hold and equilibrium $\sigma = \{\sigma^n\}_{n=1}^\infty$ along which asymptotic learning occurs in the society. This implies that there should exist a subset $\{R^{n,\sigma}\}_{n=1}^\infty$ of agents such that $\lim_{n\to\infty} \frac{1}{n}|R^{n,\sigma}| > \epsilon$ and there is an infinite index set J for which

$$i \in R^{n_j,\sigma} \text{ and } \tau_i^{n_j} < \tau_i^{n_j,\sigma}, \text{ for } j \in J. \quad (23)$$

This further implies that

$$|B^{n_j}_{i,\tau_i^{n_j}}| > |B^{n_j,\sigma}_{i,\tau_i^{n_j}}|. \quad (24)$$

From equations (23) and (24) we obtain that there should exist a collection of agents $\{S^n\}_{n=1}^\infty$ such that (we restrict attention to index set J):

(i) $R^{n,\sigma} \subseteq S^n_{follow}$.

(ii) There exists a $k > 0$ such that $S^n \subseteq V_k^{n,\sigma}$.

(ii) $\lim_{n\to\infty} \frac{1}{n}|S^n| = 0$, since otherwise asymptotic learning would not occur under equilibrium σ.

Note that collection $\{S^n\}_{n=1}^{\infty}$ satisfies the definition of a set of leading agents [cf. Definition 5] and Proposition 2 (ii) follows. ■

Proof of Proposition 3. Consider the following two events A and B.

Event A: Layer 1 (the top layer) has more than k agents, where $k > 0$ is a scalar.

Event B: The total number of layers is more than k.

From the definition of a hierarchical sequence of communication networks, we have

$$\mathbb{P}(A) = \prod_{i=2}^{k}\left(1 - \frac{1}{i^{1+\zeta}}\right) < \exp\left(-\sum_{i=2}^{k}\frac{1}{i^{1+\zeta}}\right). \tag{25}$$

Also,

$$\mathbb{P}(B) \leq \frac{\mathbb{E}(\mathcal{L})}{k} = \frac{1}{k}\sum_{i=2}^{\infty}\frac{1}{i^{1+\zeta}}, \tag{26}$$

from Markov's inequality, where $\mathcal{L}$ is a random variable that denotes the number of layers in the hierarchical society. Let $\zeta(\eta)$ be small enough and k (and consequently n) large enough such that $\sum_{i=2}^{k}\frac{1}{i^{1+\zeta}} > \log\frac{4}{\eta}$ and $\sum_{i=2}^{\infty}\frac{1}{i^{1+\zeta}} < \frac{k\cdot\eta}{4}$. For those values of ζ and k we obtain $\mathbb{P}(A) < \eta/4$ and $\mathbb{P}(B) < \eta/4$. Next, consider the event $C = A^c \cap B^c$, which from Eqs. (25) and (26) has probability $\mathbb{P}(C) > 1 - \eta/2$ for the values of ζ and k chosen above. Moreover, we consider

Event D: The agents on the top layer are information hubs, i.e.,

$$\lim_{n\to\infty}|B_{i,1}^n| = \infty, \text{ for all } i \in \mathcal{N}_1^n.$$

We claim that event D occurs with high probability if C occurs, i.e., $\mathbb{P}(D\big| C) > 1-\eta/2$, which implies

$$\mathbb{P}(C \cap D) = \mathbb{P}(D\big| C)\mathbb{P}(C) > (1-\eta/2)^2 > 1-\eta. \tag{27}$$

In particular, note that conditional on event C occurring, the total number of layers and the total number of agents in the top layer is at most k. From the definition of a hierarchical society, agents in layers with index $\ell > 1$ have an edge to a uniform agent that belongs to a layer with lower index, with probability one. Therefore, if we denote the degree of an agent in a top layer by $\mathcal{D}_1^n$ we have

$$\mathcal{D}_1^n = \sum_{i=1}^{\mathcal{T}_2^n}\mathcal{I}_{i,1}^{level2} + \cdots + \sum_{i=1}^{\mathcal{T}_{\mathcal{L}}^n}\mathcal{I}_{i,1}^{level\mathcal{L}}, \tag{28}$$

where $\mathcal{T}_i^n$ denotes the random number of agents in layer i and $\mathcal{I}_{i,1}^{levelj}$ is an indicator variable that takes value one if there is an edge from agent i to agent 1 (here $levelj$ simply denotes that agent i

belongs to level j). Again from the definition, we have $\mathbb{P}(I_{i1}^{levelj} = 1) = \frac{1}{\sum_{\ell=1}^{j-1} \mathcal{T}_\ell^n}$, where the sum in the denominator is simply the total number of agents that lie in layers with lower index, and finally, $\mathcal{T}_1^n + \cdots \mathcal{T}_{\mathcal{L}}^n = n$.

We can obtain a lower bound on the expected degree of an agent in the top layer conditional on event C by viewing (28) as the following optimization problem:

$$\begin{array}{ll} \min & \dfrac{x_2}{x_1} + \cdots + \dfrac{x_k}{x_1 + \cdots + x_{k-1}} \\ s.t. & \sum_{j=1}^k x_j = n, \\ & 0 \le x_1 \le k, \\ & 0 \le x_2, \cdots, x_{k-1}, \end{array}$$

where we make use of the fact that the total number of layers is bounded by k, since we condition on event C. By solving the problem we obtain that the objective function is lower bounded by $\phi(n)$, where $\phi(n) = O(n^{1/k})$ for every n. Then,

$$\begin{aligned} &\mathbb{E}[\mathcal{D}_1^n \big| C] = \\ &= \sum_{\ell=2}^{k} \sum_{\substack{k_1 \le k, \cdots, k_\ell \\ k_1 + \cdots + k_\ell = n}} \mathbb{P}(\mathcal{L} = \ell, \mathcal{T}_1^n = k_1, \cdots, \mathcal{T}_\ell^n = k_\ell | C) \cdot \mathbb{E}[\mathcal{D}_1^n \big| C, \mathcal{L} = \ell, \mathcal{T}_1^n = k_1, \cdots, \mathcal{T}_\ell^n = k_\ell] \\ &\ge \sum_{\ell=2}^{k} \sum_{\substack{k_1 \le k, \cdots, k_\ell \\ k_1 + \cdots + k_\ell = n}} \mathbb{P}(\mathcal{L} = \ell, \mathcal{T}_1^n = k_1, \cdots, \mathcal{T}_\ell^n = k_\ell | C) \cdot \phi(n) = \phi(n), \end{aligned} \tag{29}$$

where Eq. (29) follows since $\mathbb{E}[\mathcal{D}_1^n \big| C, \mathcal{L} = \ell, \mathcal{T}_1^n = k_1, \cdots, \mathcal{T}_\ell^n = k_\ell] \ge \phi(n)$ for all values of ℓ ($2 \le \ell \le k$) and $k_1, \cdots, k_\ell$ ($k_1 \le k, k_1 + \cdots + k_\ell = n$) from the optimal solution of the optimization problem. The same lower bound applies for all agents in the top layer. Similarly we have for the variance of the degree of an agent in the top layer (we use $\ell, k_1, \cdots, k_\ell$ as a shorthand for $\mathcal{L} = \ell, \mathcal{T}_1^n = k_1, \cdots, \mathcal{T}_\ell^n = k_\ell$)

$$\begin{aligned} Var[\mathcal{D}_1^n \big| C] &= \sum_{\ell=2}^{k} \sum_{\substack{k_1 \le k, \cdots, k_\ell \\ k_1 + \cdots + k_\ell = n}} \mathbb{P}(\ell, k_1, \cdots, k_\ell | C) \cdot Var[\mathcal{D}_1^n \big| C, \ell, k_1, \cdots, k_\ell] \\ &= \sum_{\ell=1}^{k} \sum_{\substack{k_1 \le k, \cdots, k_\ell \\ k_1 + \cdots + k_\ell = n}} \mathbb{P}(\ell, k_1, \cdots, k_\ell | C) \cdot \Big(k_2 Var(I_{i,1}^{level2}) + \cdots + k_\ell Var(I_{i,1}^{level\ell}) \Big) \quad (30) \\ &\le \sum_{\ell=1}^{k} \sum_{\substack{k_1 \le k, \cdots, k_\ell \\ k_1 + \cdots + k_\ell = n}} \mathbb{P}(\ell, k_1, \cdots, k_\ell | C) \cdot \Big(k_2 \mathbb{E}(I_{i,1}^{level2}) + \cdots + k_\ell \mathbb{E}(I_{i,1}^{level\ell}) \Big) = \mathbb{E}[\mathcal{D}_1^n \big| C], \quad (31) \end{aligned}$$

where Eq. (30) follows by noting that conditional on event C and the number of layers and the agents in each layer being fixed, the indicator variables (defined above) are independent and Eq. (31) follows since the variance of an indicator variable is smaller that its expectation. We conclude that the variance of the degree is smaller than the expected value and from Chebyschev's inequality we conclude that

$$\mathbb{P}(D) \geq \mathbb{P}(\bigcap_{i \in \mathcal{N}_1^n} \frac{\mathcal{D}_i^n}{\phi(n)} > \zeta) > 1 - \eta/2,$$

where $\zeta > 0$, i.e., with high probability all agents in the top layer are information hubs (recall that $\lim_{n\to\infty} \phi(n) = \infty$).

We have shown that when event $C \cap D$ occurs, there is a path of length at most k (the total number of layers) from each agent to an agent at the top layer, i.e., an information hub with high probability. Therefore, if the discount rate r is smaller than some bound ($r < \bar{r}$), then perfect asymptotic learning occurs. Finally, we complete the proof by noting that $\mathbb{P}(C \cap D) > (1 - \eta/2)^2 > 1 - \eta$. ■

Proof of Proposition 4.

Proposition 4 is a direct consequence of the next lemma, which intuitively states that there is no incentive to lie to an agent with a large number of neighbors, assuming that everybody else is truthful.

Lemma 3 (Truthful Communication to a High Degree Agent)**.** *There exists a scalar $k > 0$, such that truth-telling to agent i, with $indeg_i^n \geq k$, in the first time period is an equilibrium of $INFO(G^n)$. Formally,*

$$(\sigma^{n,truth}, m^{n,truth}) \in INFO(G^n),$$

where $m_{ji,0}^{n,truth} = s_j$ for $j \in B_{i,1}^n$.

Proof. The proof is based on the following argument. Suppose that all agents in $B_{i,1}^n$ except j report their signals truthfully to i. Moreover, let $|B_{i,1}^n| \geq k$, where k is a large constant. Then, it is an weakly dominant strategy for j to report her signal truthfully to i, since j's message is not *pivotal* for agent i, i.e., i will take an irreversible action after the first communication step, no matter what j reports. ■

Proof (sketch) of Proposition 5. Assume that asymptotic learning occurs at the optimal allocation $sp = \{sp^n\}_{n=1}^{\infty}$. Then,

$$\lim_{k\to\infty} \lim_{n\to\infty} \frac{1}{n} |V_k^n| = 0. \tag{32}$$

This follows since if Equation (32) were not true, then a social planner could replicate the allocation induced by the perfect observation radius and achieve a higher aggregate welfare. This is possible since $B_{i,\tau}^n \supseteq B_{i,\tau}^{n,sp}$ for every τ, where *sp* denotes the socially optimal strategy profile. From Equation (32) and Proposition 1 we obtain that asymptotic learning occurs in all equilibria σ. Finally, the proposition follows using similar arguments as those used in the proof of Proposition 8. ■

Proof of Proposition 6. The claim follows by noting that the social planner could choose the following strategy profile: for each $j \in D_{k,\ell}^{n,\sigma}$ delay i's irreversible action by at least one time period, where i is an agent such that if i delays then j gains access to a least ℓ additional signals. Moreover, it is straightforward to see that there exist ϵ, δ for which ϵ, δ-learning fails. ■

Proofs from Section 4

Proof of Proposition 7

First we make an observation which will be used frequently in the subsequent analysis. Consider an agent i such that $H^n_{SC(i)} \in \mathcal{H}^n_{\bar{k}}$, where $\bar{k}$ is an integer appropriately chosen (see below), i.e., the size of the social clique of agent i is greater than or equal to $\bar{k}$, $|H^n_{SC(i)}| \geq \bar{k}$. Suppose agent i does not form a link with cost c with any agents outside her social clique. If she makes a decision at time $t = 0$ based on her signal only, her expected payoff is $\pi - \frac{1}{\rho+\bar{\rho}}$. If she waits for one period, she has access to the signals of all the agents in her social clique (i.e., she has access to at least $\bar{k}$ signals), implying that her expected payoff would be bounded from below by $\frac{\lambda}{r+\lambda}\left(\pi - \frac{1}{\rho+\bar{\rho}\bar{k}}\right)$. Hence, her expected payoff $\mathbb{E}[\Pi_i(g^n)]$ satisfies

$$\mathbb{E}[\Pi_i(g^n)] \geq \max\left\{\pi - \frac{1}{\rho+\bar{\rho}}, \frac{\lambda}{r+\lambda}\left(\pi - \frac{1}{\rho+\bar{\rho}\bar{k}}\right)\right\},$$

for any link formation strategy g^n and along any $\sigma \in INFO(G^n)$ (where G^n is the communication network induced by g^n). Suppose now that agent i forms a link with cost c with an agent outside her social clique. Then, her expected payoff is bounded from above by

$$\mathbb{E}[\Pi_i(g^n)] < \max\left\{\pi - \frac{1}{\rho+\bar{\rho}}, \left(\frac{\lambda}{\lambda+r}\right)^2 \pi - c\right\},$$

where the second term in the maximum is an upper bound on the payoff she could get by having access to the signals of all agents she is connected to in two time steps (i.e., signals of the agents in her social clique and in the social clique that she is connected to). Combining the preceding two relations, we see that an agent i with $H^n_{SC(i)} \in \mathcal{H}^n_{\bar{k}}$ *will not* form any costly links in any network equilibrium, i.e.,

$$g^n_{ij} = 1 \text{ if and only if } SC(j) = SC(i) \text{ for all } i \text{ such that } |H^n_{SC(i)}| \geq \bar{k}. \tag{33}$$

for $\bar{k}$ such that

$$\frac{\lambda}{r+\lambda}\left(\pi - \frac{1}{\rho+\bar{\rho}\bar{k}}\right) \geq \left(\frac{\lambda}{\lambda+r}\right)^2 \pi - c.$$

(a) Condition (11) implies that for all sufficiently large n, we have

$$\left|\mathcal{H}^n_{\bar{k}}\right| \geq \xi n, \tag{34}$$

where $\xi > 0$ is a constant. For any ϵ with $0 < \epsilon < \xi$, we have

$$\begin{aligned}\mathbb{P}\left(\sum_{i=1}^{n} \frac{1-M_i^{n,\epsilon}}{n} > \epsilon\right) &= \mathbb{P}\left(\left[\sum_{i|\ |H^n_{SC(i)}|<\bar{k}} \frac{1-M_i^{n,\epsilon}}{n} + \sum_{i|\ |H^n_{SC(i)}|\geq\bar{k}} \frac{1-M_i^{n,\epsilon}}{n}\right] > \epsilon\right) \\ &\geq \mathbb{P}\left(\sum_{i|\ |H^n_{SC(i)}|\geq\bar{k}} \frac{1-M_i^{n,\epsilon}}{n} > \epsilon\right). \end{aligned} \tag{35}$$

The right-hand side of the preceding inequality can be re-written as

$$\begin{aligned}\mathbb{P}\left(\sum_{i|\ |H^n_{SC(i)}|\geq\bar{k}}\frac{1-M_i^{n,\epsilon}}{n}>\epsilon\right) &= 1-\ \mathbb{P}\left(\sum_{i|\ |H^n_{SC(i)}|\geq\bar{k}}\frac{1-M_i^{n,\epsilon}}{n}\leq\epsilon\right)\\ &= 1-\ \mathbb{P}\left(\sum_{i|\ |H^n_{SC(i)}|\geq\bar{k}}\frac{M_i^{n,\epsilon}}{n}\geq r-\epsilon\right),\end{aligned}$$

where $r=\sum_{i|\ |H^n_{SC(i)}|\geq\bar{k}}\frac{1}{n}$. By Eq. (34), it follows that for n sufficiently large, we have $r\geq\xi$. Using Markov's inequality, the preceding relation implies

$$\mathbb{P}\left(\sum_{i|\ |H^n_{SC(i)}|\geq\bar{k}}\frac{1-M_i^{n,\epsilon}}{n}>\epsilon\right)\geq 1-\frac{\sum_{i|\ |H^n_{SC(i)}|\geq\bar{k}}\mathbb{E}[M_i^{n,\epsilon}]}{n}\cdot\frac{1}{r-\epsilon}. \tag{36}$$

By Lemma 2 and observation (33), $\mathbb{E}[M_i^{n,\epsilon}]$ for an agent i with $|H^n_{SC(i)}|\geq\bar{k}$ is upper bounded by

$$\mathbb{P}(M_i^{n,\epsilon}=0)\geq erf\left(\epsilon\sqrt{\frac{|H^n_{SC(i)}|\bar{\rho}}{2}}\right),$$

and therefore

$$\mathbb{E}[M_i^{n,\epsilon}]\leq 1-erf\left(\epsilon\sqrt{\frac{|H^n_{SC(i)}|\bar{\rho}}{2}}\right).$$

Now assuming that social cliques are ordered by size (H^n_1 is the biggest), we can re-write Eq. (36) as

$$\begin{aligned}&\mathbb{P}\left(\sum_{i|\ |H^n_{SC(i)}|\geq\bar{k}}\frac{1-M_i^{n,\epsilon}}{n}>\epsilon\right)\geq\\ &\geq 1-\frac{\sum_{j=1}^{|\mathcal{H}^n_{\bar{k}}|}|H^n_j|\left(1-erf\left(\epsilon\sqrt{\frac{|H^n_j|\bar{\rho}}{2}}\right)\right)}{(r-\epsilon)\cdot n}\\ &\geq 1-\frac{r\cdot(1-\zeta)}{r-\epsilon}\geq 1-\frac{\xi\cdot(1-\zeta)}{\xi-\epsilon}>\delta\end{aligned} \tag{37}$$

Here, the second inequality is obtained since the largest value for the sum is achieved when all summands are equal and $\zeta=erf\left(\epsilon\sqrt{\frac{\bar{k}\bar{\rho}}{2}}\right)$. The third inequality holds using the relation $r\geq\xi$ and choosing appropriate values for ϵ,δ.

This establishes that for all sufficiently large n, we have

$$\mathbb{P}\left(\sum_{i=1}^{n}\frac{1-M_i^{n,\epsilon}}{n}>\epsilon\right)>\delta>0,$$

which implies

$$\limsup_{n\to\infty} \mathbb{P}\left(\sum_{i=1}^{n} \frac{1-M_i^{n,\epsilon}}{n} > \epsilon\right) > \delta,$$

and shows that perfect asymptotic learning does not occur in any network equilibrium.

(b) We show that if the communication cost structure satisfies condition (12), then asymptotic learning occurs in all network equilibria $(g,\sigma) = (\{g^n,\sigma^n\})_{n=1}^{\infty}$. For an illustration of the resulting communication networks, when condition (13) holds, refer to Figure 7(a). Let $B_i^n(G^n)$ be the *neighborhood* of agent i in communication network G^n (induced by the link formation strategy g^n),

$$B_i^n(G^n) = \{j \mid \text{ there exists a path } \mathcal{P} \text{ in } G^n \text{ from } j \text{ to } i\},$$

i.e., $B_i^n(G^n)$ is the set of agents in G^n whose information agent i can acquire over a sufficiently large (but finite) period of time.

We first show that for any agent i such that $\limsup_{n\to\infty} \left|H_{SC(i)}^n\right| < \bar{k}$, her neighborhood in any network equilibrium satisfies $\lim_{n\to\infty} \left|B_i^n\right| = \infty$. We use the notion of an isolated social clique to show this. For a given n, we say that a social clique H_ℓ^n is *isolated* (at a network equilibrium (g,σ)) if no agent in H_ℓ^n forms a costly link with an agent outside H_ℓ^n in (g,σ). Equivalently, a social clique H_ℓ^n is not isolated if there exists at least one agent $j \in H_\ell^n$, such that j incurs cost c and forms a link with an agent outside H_ℓ^n.

We show that for an agent i with $\limsup_{n\to\infty} \left|H_{SC(i)}^n\right| < \bar{k}$, the social clique $H_{SC(i)}^n$ is not isolated in any network equilibrium for all sufficiently large n. Using condition (12), we can assume without loss of generality that social cliques are ordered by size from largest to smallest and that $\lim_{n\to\infty} |H_1^n| = \infty$. Suppose that $H_{SC(i)}^n$ is isolated in a network equilibrium (g,σ). Then the expected payoff of agent i is upper bounded (similarly with above)

$$\mathbb{E}[\Pi_i(g^n)] \leq \max\left\{\pi - \frac{1}{\rho+\bar{\rho}}, \frac{\lambda}{r+\lambda}\left(\pi - \frac{1}{\rho+\bar{\rho}(\bar{k}-1)}\right)\right\}$$

Using the definition of $\bar{k}$, it follows that for some $\epsilon > 0$,

$$\mathbb{E}[\Pi_i(g^n)] \leq \max\left\{\pi - \frac{1}{\rho+\bar{\rho}}, \left(\frac{\lambda}{r+\lambda}\right)^2 \pi - c - \epsilon\right\} \tag{38}$$

Suppose next that agent i forms a link with an agent $j \in H_1^n$. Her expected payoff $\mathbb{E}[\Pi_i(g^n)]$ satisfies

$$\mathbb{E}[\Pi_i(g^n)] \geq \left(\frac{\lambda}{r+\lambda}\right)^2 \cdot \left(\pi - \frac{1}{\rho+\bar{\rho}\left|H_1^n\right|}\right) - c,$$

since in two time steps, she has access to the signals of all agents in the social clique H_1^n. Since

$\lim_{n\to\infty} |H_1^n| = \infty$, there exists some N_1 such that

$$\mathbb{E}[\Pi_i(g^n)] > \left(\frac{\lambda}{\lambda + r}\right)^2 \pi - c - \epsilon \qquad \text{for all } n > N_1.$$

Comparing this relation with Eq. (38), we conclude that under the assumption that $r < \bar{r}$ (for appropriate $\bar{r}$), the social clique $H_{SC(i)}^n$ is not isolated in any network equilibrium for all $n > N_1$.

Next, we show that $\lim_{n\to\infty} |B_i^n| = \infty$ in any network equilibrium. Assume to arrive at a contradiction that $\limsup_{n\to\infty} |B_i^n| < \infty$ in some network equilibrium. This implies that $\limsup_{n\to\infty} |B_i^n| < |H_1^n|$ for all $n > N_2 > N_1$. Consider some $n > N_2$. Since $H_{SC(i)}^n$ is not isolated, there exists some $j \in H_{SC(i)}^n$ such that j forms a link with an agent h outside $H_{SC(i)}^n$. Since $\limsup_{n\to\infty} |B_i^n| < |H_1^n|$, agent j can improve her payoff by changing her strategy to $g_{jh}^n = 0$ and $g_{jh'}^n = 1$ for $h' \in H_1^n$, i.e., j is better off deleting her existing costly link and forming one with an agent in social clique H_1^n. Hence, for any network equilibrium, we have

$$\lim_{n\to\infty} |B_i^n| = \infty \quad \text{for all } i \text{ with } \limsup_{n\to\infty} |H_{SC(i)}^n| < \bar{k} \tag{39}$$

We next consider the probability that a non-negligible fraction (ϵ-fraction) of agents takes an action that is at least ϵ-away from optimal with probability at least δ along a network equilibrium (g, σ). For any n, we have from Markov's inequality

$$\mathbb{P}\left(\sum_{i=1}^{n} \frac{1 - M_i^{n,\epsilon}}{n} > \epsilon\right) \leq \frac{1}{\epsilon} \cdot \sum_{i=1}^{n} \frac{\mathbb{E}[1 - M_i^{n,\epsilon}]}{n} \tag{40}$$

We next provide upper bounds on the individual terms in the sum on the right-hand side. We have

$$\mathbb{E}[1 - M_i^{n,\epsilon}] \leq erf\left(\epsilon\sqrt{\frac{\bar{\rho}|B_i^n|}{2}}\right). \tag{41}$$

Consider an agent i with $\limsup_{n\to\infty} |H_{SC(i)}^n| < \bar{k}$ (i.e., $|H_{SC(i)}^n| < \bar{k}$ for all n large). By Eq. (39), we have $\lim_{n\to\infty} |B_i^n| = \infty$. Together with Eq. (41), this implies that for some $\zeta > 0$, there exists some N such that for all $n > N$, we have

$$\mathbb{E}[1 - M_i^{n,\epsilon}] < \frac{\epsilon\,\zeta}{2} \qquad \text{for all } i \text{ with } \limsup_{n\to\infty} |H_{SC(i)}^n| < \bar{k}. \tag{42}$$

Consider next an agent i with $\limsup_{n\to\infty} |H_{SC(i)}^n| \geq \bar{k}$, and for simplicity, let us assume that the limit exists, i.e., $\lim_{n\to\infty} |H_{SC(i)}^n| \geq \bar{k}$.[12]

[12] The case when the limit does not exist can be proven by focusing on different subsequences. In particular, along any subsequence $\mathcal{N}_i$ such that $\lim_{n\to\infty, n\in\mathcal{N}_i} |H_{SC(i)}^n| \geq \bar{k}$, the same argument holds. Along any subsequence $\mathcal{N}_i$ with $\lim_{n\to\infty, n\in\mathcal{N}_i} |H_{SC(i)}^n| < \bar{k}$, we can use an argument similar to the previous case to show that $\lim_{n\to\infty, n\in\mathcal{N}_i} |B_i^n| = \infty$, and therefore $\mathbb{E}[1 - M_i^{n,\epsilon}] < \frac{\epsilon\,\zeta}{2}$ for n large and $n \in \mathcal{N}_i$.

This implies that $|H^n_{SC(i)}| \geq \bar{k}$ for all large n, and therefore,

$$\sum_{i|\ \limsup_{n\to\infty} |H^n_{SC(i)}|\geq \bar{k}} \frac{\mathbb{E}[1-M_i^{n,\epsilon}]}{n} \leq \sum_{j=1}^{|\mathcal{H}^n_{\bar{k}}|} |H^n_j| \cdot erf\left(\epsilon\sqrt{\frac{\bar{\rho}|H^n_j|}{2}}\right) \leq \frac{|\mathcal{H}^n_{\bar{k}}|}{n}\cdot \bar{k},$$

where the first inequality follows from Eq. (41). Using condition (12), i.e., $\lim_{n\to\infty} \frac{\left|\mathcal{H}^n_{\bar{k}}\right|}{n} = 0$, this relation implies that there exists some $\tilde{N}$ such that for all $n > \tilde{N}$, we have

$$\sum_{i|\ \limsup_{n\to\infty} |H^n_{SC(i)}|\geq \bar{k}} \frac{\mathbb{E}[1-M_i^{n,\epsilon}]}{n} < \frac{\epsilon\,\zeta}{2}. \tag{43}$$

Combining Eqs. (42) and (43) with Eq. (40), we obtain for all $n > \max\{N, \tilde{N}\}$,

$$\mathbb{P}\left(\sum_{i=1}^{n} \frac{1-M_i^{n,\epsilon}}{n} > \epsilon\right) < \zeta,$$

where $\zeta > 0$ is an arbitrary scalar. This implies that

$$\lim_{n\to\infty} \mathbb{P}\left(\sum_{i=1}^{n} \frac{1-M_i^{n,\epsilon}}{n} > \epsilon\right) = 0,$$

for all ϵ, showing that perfect asymptotic learning occurs along every network equilibrium.

(c) The proof proceeds in two parts. First, we show that if condition (13) is satisfied, learning occurs in at least one network equilibrium (g, σ). Then, we show that there exists a $\bar{c} > 0$, such that if $c < \bar{c}$, then learning occurs in all network equilibria. We complete the proof by showing that if $c > \bar{c}$, then there exist network equilibria, in which asymptotic learning fails, even when condition (13) holds. We consider the case when agents are patient, i.e., the discount rate $r \to 0$. We consider $\bar{k}$, such that $c > \frac{1}{\rho+\bar{\rho}\bar{k}}$ and $c < \frac{1}{\rho+\bar{\rho}(\bar{k}-1)} - \epsilon'$, for some $\epsilon' > 0$ (such a $\bar{k}$ exists). Finally, we assume that $c < \frac{1}{\rho+\bar{\rho}}$, since otherwise no agent would have an incentive to form a costly link.

Part 1: We assume, without loss of generality, that social cliques are ordered by size (H^n_1 is the smallest). Let $\mathcal{H}^n_{<\bar{k}}$ denote the set of social cliques of size less than $\bar{k}$, i.e., $\mathcal{H}^n_{<\bar{k}} = \{H^n_i, i = 1, \ldots, K^n \mid |H^n_i| < \bar{k}\}$. Finally, let $rec(j)$ and $send(j)$ denote two special nodes for social clique H^n_j, the *receiver* and the *sender* (they might be the same node). We claim that (g^n, σ^n) described below and depicted in Figure 7(b) is an equilibrium of the network learning game $\Gamma(C^n)$ for n large enough and δ sufficiently close to one.

$$g^n_{ij} = \begin{cases} 1 & \text{if } SC(i) = SC(j), \text{ i.e., } i, j \text{ belong to the same social clique,} \\ 1 & \text{if } i = rec(\ell-1) \text{ and } j = send(\ell) \text{ for } 1 < \ell \leq |\mathcal{H}^n_{<\bar{k}}|, \\ 1 & \text{if } i = rec(|\mathcal{H}^n_{<\bar{k}}|) \text{ and } j = send(1), \\ 0 & \text{otherwise} \end{cases}$$

and $\sigma^n \in INFO(G^n)$, where G^n is the communication network induced by g^n. In this communication network, social cliques with size less than $\bar{k}$ are organized in a directed *ring*, and all agents i, such that $|H^n_{SC(i)}| < \bar{k}$ have the same neighborhood, i.e., $B^n_i = B^n$ for all such agents.

Next, we show that the strategy profile (g^n, σ^n) described above is indeed an equilibrium of the network learning game $\Gamma(C^n)$. We restrict attention to large enough n's. In particular, let N be such that $\sum_{i=1}^{|\mathcal{H}^N_{<\bar{k}}|} |H^N_i| > \bar{k}$ and consider any $n > N$ (such N exists from condition (13)). Moreover, we assume that the discount rate is sufficiently close to zero. We consider the following two cases.

Case 1: Agent i is not a connector. Then, $g^n_{ij} = 1$ if and only if $SC(j) = SC(i)$. Agent i's neighborhood as noted above is set B^n, which is such that $\pi - \frac{1}{\rho+\bar{\rho}|B^n|} > \pi - c$ from the assumption on n, i.e., $n > N$, where N such that $\sum_{i=1}^{|\mathcal{H}^N_{<\bar{k}}|} |H^N_i| > \bar{k}$. Agent i can communicate with all agents in B^n in at most $|\mathcal{H}_{<\bar{k}}|$ communication steps. Therefore, her expected payoff is lower-bounded by

$$\mathbb{E}[\Pi_i(g^n)] \geq \left(\frac{\lambda}{\lambda + r}\right)^{\left|\mathcal{H}^n_{<\bar{k}}\right|} \cdot \left(\pi - \frac{1}{\rho + \bar{\rho}\bar{k}}\right) > \pi - c,$$

under any equilibrium σ^n for r sufficiently close to zero. Agent i can deviate by forming a costly link with agent m, such that $SC(m) \neq SC(i)$. However, this is not profitable since from above her expected payoff under (g^n, σ^n) is at least $\pi - c$ (which is the maximum possible payoff if an agent chooses to form a costly link).

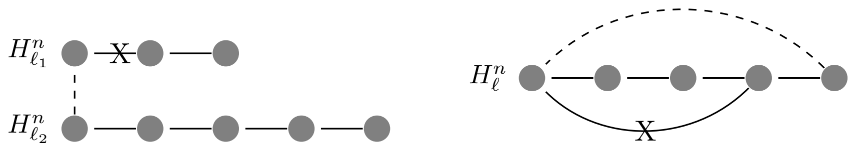
Case 2: Agent i is a connector, i.e., there exists exactly one j, such that $SC(j) \neq SC(i)$ and $g^n_{ij} = 1$. Using a similar argument as above we can show that it is not profitable for agent i to form an additional costly link with an agent m, such that $SC(m) \neq SC(i)$. On the other hand, agent i could deviate by setting $g^n_{ij} = 0$. However, then her expected payoff would be

$$\begin{aligned}\mathbb{E}[\Pi_i(g^n)] &= \max\left\{\pi - \frac{1}{\rho + \bar{\rho}}, \frac{\lambda}{r + \lambda}\left(\pi - \frac{1}{\rho + \bar{\rho}|H^n_i|}\right)\right\} \\ &\leq \max\left\{\pi - \frac{1}{\rho + \bar{\rho}}, \frac{\lambda}{r + \lambda}\left(\pi - \frac{1}{\rho + \bar{\rho}(\bar{k} - 1)}\right)\right\} < \pi - c - \epsilon' \\ &< \left(\frac{\lambda}{r + \lambda}\right)^{\left|\mathcal{H}^n_{<\bar{k}}\right|}\left(\pi - \frac{1}{\rho + \bar{\rho}|B^n|}\right) - c - \epsilon,\end{aligned} \tag{44}$$

for discount rate sufficiently close to zero. Therefore deleting the costly link is not a profitable deviation. Similarly we can show that it a (weakly) dominant strategy for the connector not to replace her costly link with another costly link.

We showed that (g^n, σ^n) is an equilibrium of the network learning game. Note that we described a link formation strategy, in which social cliques connect to each other in a specific order (in increasing size). There is nothing special about this ordering and any permutation of the first $|\mathcal{H}^n_{<\bar{k}}|$ cliques is an equilibrium as long as they form a directed ring. Finally, any node in a social clique can be a receiver or a sender.

Next, we argue that asymptotic learning occurs in network equilibria $(g, \sigma) = \{(g^n, \sigma^n)\}_{n=1}^{\infty}$, where for all $n > N$, N is a large constant, g^n has the form described above. As shown above,



(a) Deviation for $i \in H^n_{\ell_1}$ - property (i). (b) Deviation for $i \in H^n_\ell$ - property (ii).

Figure 9: Communication networks under condition (13).

all agents i for which $H^n_{SC(i)} < \bar{k}$ have the same neighborhood, which we denoted by B^n. Moreover, $\lim_{n\to\infty} |B^n| = \infty$, since social cliques with size less than $\bar{k}$ are connected to the ring and, by condition (13), $\lim_{n\to\infty} \sum_{i|\ |H^n_i|<\bar{k}} |H^n_i| = \infty$. For discount rate r sufficiently close to zero and from arguments similar to those in the proof of part (b), we conclude that asymptotic learning occurs in network equilibria (g, σ).

Part 2: We have shown a particular form of network equilibria, in which asymptotic learning occurs. The following proposition states that for discount rate sufficiently close to zero network equilibria fall in one of two forms.

Proposition 9. *Suppose Assumptions 1, 3 and condition (13) hold. Then, an equilibrium (g^n, σ^n) of the network learning game $\Gamma(C^n)$ can be in one of the following two forms.*

(i) ***(Incomplete) Ring Equilibrium:*** *Social cliques with indices $\{1, \cdots, j\}$, where $j \leq |\mathcal{H}^n_{<\bar{k}}|$, form a directed ring as described in Part 1 and the rest of the social cliques are isolated. We call those equilibria ring equilibria and, in particular, a ring equilibrium is called complete if $j = |\mathcal{H}^n_{<\bar{k}}|$, i.e., if all social cliques with size less than $\bar{k}$ are not isolated.*

(ii) ***Directed Line Equilibrium:*** *Social cliques with indices $\{1, \cdots, j\}$, where $j \leq |\mathcal{H}^n_{<\bar{k}}|$, and clique with index $|H^n_{K^n}|$ (the largest clique) form a directed line with the latter being the endpoint. The rest of the social cliques are isolated.*

Proof. Let (g^n, σ^n) be an equilibrium of the network learning game $\Gamma(C^n)$. Monotonicity of the expected payoff as a function of the number of signals observed implies that if clique H^n_ℓ is not isolated, then no clique with index less than ℓ is isolated in the communication network induced by g^n. In particular, let $conn(\ell)$ be the connector of social clique H^n_ℓ and $\mathbb{E}[\Pi_{conn(\ell)}(g^n)]$ be her expected payoff. Consider an agent i such that $SC(i) = \ell' < \ell$ and, for the sake of contradiction, $H^n_{\ell'}$ is isolated in the communication network induced by g^n. Social cliques are ordered by size, therefore, $|H^n_{\ell'}| \leq |H^n_\ell|$. Now, we use the monotonicity mentioned above. Consider the expected payoff of i:

$$\begin{aligned}
\mathbb{E}[\Pi_i(g^n)] &= \max\left\{\pi - \frac{1}{\rho + \bar{\rho}}, \frac{\lambda}{\lambda + r}\left(\pi - \frac{1}{\rho + \bar{\rho}|H^n_{\ell'}|}\right)\right\} \\
&\leq \max\left\{\pi - \frac{1}{\rho + \bar{\rho}}, \frac{\lambda}{\lambda + r}\left(\pi - \frac{1}{\rho + \bar{\rho}|H^n_{\ell}|}\right)\right\} < \ \mathbb{E}[\Pi_{conn(\ell)}(g^n)], \qquad (45)
\end{aligned}$$

where the last inequality follows from the fact that agent $conn(\ell)$ formed a costly link. Consider a deviation, $g_i^{n,deviation}$ for agent i, in which $g_{i,conn(\ell)}^{n,deviation} = 1$ and $g_{ij}^{n,deviation} = g_{ij}^n$, i.e., agent i forms a costly link with agent $conn(\ell)$. Then,

$$\mathbb{E}[\Pi_i(g^{n,deviation})] \geq \frac{\lambda}{\lambda + r}\, \mathbb{E}[\Pi_{conn(\ell)}(g^n)] > \mathbb{E}[\Pi_i(g^n)],$$

from (45) and for discount rate sufficiently close to zero. Therefore, social clique $H_{\ell'}^n$ will not be isolated in any network equilibrium (g^n, σ^n).

Next, we show two structural properties that all network equilibria (g^n, σ^n) should satisfy, when the discount rate r is sufficiently close to one. We say that there exists a path $\mathcal{P}$ between social cliques $H_{\ell_1}^n$ and $H_{\ell_2}^n$, if there exists a path between some $i \in H_{\ell_1}^n$ and $j \in H_{\ell_2}^n$. Also, we say that the in-degree (out-degree) of social clique $H_{\ell_1}^n$ is k, if the sum of in-links (out-links) of the nodes in $H_{\ell_1}^n$ is k, i.e., $H_{\ell_1}^n$ has in-degree k if $\sum_{i \in H_{\ell_1}^n} \sum_{j \notin H_{\ell_1}^n} g_{ij}^n = k$.

(i) Let $H_{\ell_1}^n, H_{\ell_2}^n$ be two social cliques that are not isolated. Then, there should exist a directed path $\mathcal{P}$ in G^n induced by g^n between the two social cliques.

(ii) The in-degree and out-degree of each social clique is at most one.

Figure 9 provides an illustration of why the properties hold for patient agents. In particular, for property (i), let $i = conn(H_{\ell_1}^n)$ and $j = conn(H_{\ell_2}^n)$ and assume, without loss of generality, that $|B_i^n| \leq |B_j^n|$. Then, for discount rate sufficiently close to zero and from monotonicity of the expected payoff, we conclude that i has an incentive to deviate, delete her costly and form a costly link with agent j. Property (ii) follows due to similar arguments. From the above, we conclude that the only two potential equilibrium topologies are the (incomplete) ring and the directed line with the largest clique being the endpoint under the assumptions of the proposition. ■

So far we have shown a particular form of network equilibria that arise under condition (13), in which asymptotic learning occurs. We also argued that under condition (13) only (incomplete) ring or directed line equilibria can arise for network learning game $\Gamma(C^n)$. In the remainder we show that there exists a bound $\bar{c} > 0$ on the common cost c for forming a link between two social cliques, such that if $c < \bar{c}$ all network equilibria (g, σ) that arise satisfy that g^n is a complete ring equilibrium for all $n > N$, where N is a constant. In those network equilibria asymptotic learning occurs as argued in Part 1. On the other hand, if $c > \bar{c}$ coordination among the social cliques may fail and additional equilibria arise in which asymptotic learning does not occur. Let

$$\bar{c}^n = \min_k \left\{ -\frac{1}{\rho + \bar{\rho}(\sum_{j=1}^k |H_j^n| + |H_{k+1}^n|)} + \frac{1}{\rho + \bar{\rho}|H_{k+1}^n|} \right\} \tag{46}$$

where $k_1 \leq k < |H^n_{<\bar{k}}|$ and $\sum_{j=1}^{k_1} |H^n_j| \geq |H^n_{K^n}|$ (size of the largest social clique). Moreover, let

$$\bar{c} = \lim \inf_{n\to\infty} \bar{c}^n.$$

The following proposition concludes the proof.

Proposition 10. *Suppose Assumptions 1, 3 and condition (13) hold. If $c < \bar{c}$ asymptotic learning occurs in all network equilibria (g, σ). Otherwise, there exist equilibria in which asymptotic learning does not occur.*

Proof. Let the common cost c be such that $c < \bar{c}$, where $\bar{c}$ is defined as above, and consider a network equilibrium (g, σ). Let N be a large enough constant and consider the corresponding g^n for $n > N$. We claim that g^n is a complete ring equilibrium for all such n. Assume for the sake of contradiction that the claim is not true. Then, from Proposition 9, g^n is either an incomplete ring equilibrium or a directed line equilibrium. We consider the former case (the latter case can be shown with similar arguments). There exists an isolated social clique H^n_ℓ, such that $|H^n_\ell| < \bar{k}$ and all cliques with index less than ℓ are not isolated and belong to the incomplete ring. However, from the definition of $\bar{c}$ we obtain that an agent $i \in H^n_\ell$ would have an incentive to connect to the incomplete ring, thus we reach a contradiction. In particular, consider the following link formation strategy for agent i: $g^{n,deviation}_{im} = 1$ for agent $m \in H^n_{\ell-1}$ and $g^{n,deviation}_{ij} = g^n_{ij}$ for $j \neq m$. Then,

$$\begin{aligned}\mathbb{E}[\Pi^n_i(g^{n,deviation})] &\geq \left(\frac{\lambda}{\lambda + r}\right)^{|H^n_{<\bar{k}}|}\left(\pi - \frac{1}{\rho + \bar{\rho}(\sum_{j=1}^{\ell-1}|H^n_j| + |H^n_\ell|)}\right) - c \\ &> \max\left\{\pi - \frac{1}{\rho+\bar{\rho}}, \frac{\lambda}{\lambda+r}\left(\pi - \frac{1}{\rho + \bar{\rho}|H^n_\ell|}\right)\right\} = \mathbb{E}[\Pi^n_i(g^n)],\end{aligned}$$

where the strict inequality follows from the definition of $\bar{c}$ for r sufficiently close to zero. Thus, we conclude that if $c < \bar{c}$, g^n is a complete ring for all $n > N$, where N is a large constant, and from Part 1 asymptotic learning occurs in all network equilibria (g, σ). On the contrary, if $c > \bar{c}$, then there exists an infinite index set W, such that for all n in the (infinite) subsequence, $\{n_w\}_{w\in W}$, there exists a k, such that

$$\frac{1}{\rho + \bar{\rho}(\sum_{j=1}^{k}|H^n_j| + |H^n_{k+1}|)} - c < \frac{1}{\rho + \bar{\rho}|H^n_{k+1}|}. \tag{47}$$

Moreover, $|H^n_{k+1}| < \bar{k}$ and $\sum_{j=1}^{k} |H^n_j| \geq |H^n_{K^n}|$. We conclude that for (47) to hold it has to be that $\sum_{j=1}^{k} |H^n_j| < R$, where R is a uniform constant for all n in the subsequence. Consider $(g, \sigma)^\infty_{n=1}$, such that for every n in the subsequence, g^n is such that social cliques with index greater than k (as described above) are isolated and the rest form an incomplete ring or a directed line and $\sigma^n = INFO(G^n)$, where G^n is the communication network induced by g^n. From above, we obtain that for $c > \bar{c}$, (g^n, σ^n) is an equilibrium of the network learning game $\Gamma(C^n)$. perfect assymptotic learning, however, fails in such an equilibrium, since for every $i \in N^n$, $|B^n_i| \leq R$, where B^n_i denotes the neighborhood of agent i. ■

References

Acemoglu, D., M. Dahleh, I. Lobel, and A. Ozdaglar (2010): "Bayesian learning in social networks," *Working paper*.

Ambrus, A., E. Azevedo, and Y. Kamada (2010): "Hierarchical Cheap Talk," *Working paper*.

Angeletos, G., and A. Pavan (2007): "Efficient use of information and social value of information," *Econometrica*, 75(4), 1103–1142.

Austen-Smith, D. (1990): "Information transmission in debate," *American Journal of Political Science*, 34(1), 124–152.

Austen-Smith, D., and J. Banks (1996): "Information aggregation, rationality, and the Condorcet jury theorem," *American Political Science Review*, 90(1), 34–45.

Bala, V., and S. Goyal (1998): "Learning from neighbours," *Review of Economic Studies*, 65(3), 595–621.

——— (2000): "A noncooperative model of network formation," *Econometrica*, 68(5), 1181–1229.

Banerjee, A. (1992): "A simple model of herd behavior," *Quarterly Journal of Economics*, 107(3), 797–817.

Banerjee, A., and D. Fudenberg (2004): "Word-of-mouth learning," *Games and Economic Behavior*, 46, 1–22.

Barabási, A., and R. Albert (1999): "Emergence of scaling in random networks," *Science*, 286(5439), 509–512.

Bikhchandani, S., D. Hirshleifer, and I. Welch (1992): "A theory of fads, fashion, custom, and cultural change as informational cascades," *Journal of Political Economy*, 100(5), 992–1026.

Crawford, V., and J. Sobel (1982): "Strategic information transmission," *Econometrica*, 50(6), 1431–1451.

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

Duffie, D., S. Malamud, and G. Manso (2009): "Information Percolation with Equilibrium Search Dynamics," *Econometrica*, 77(5), 1513–1574.

Ellison, G., and D. Fudenberg (1995): "Word-of-mouth learning," *The Quarterly Journal of Economics*, 110, 93–126.

Farrell, J., and R. Gibbons (1989): "Cheap talk with two audiences," *American Economic Review*, 79(5), 1214–1223.

Gale, D., and S. Kariv (2003): "Bayesian learning in social networks," *Games and Economic Behavior*, 45(2), 329–346.

Galeotti, A. (2006): "One-way flow networks: the role of heterogeneity," *Economic Theory*, 29(1), 163–179.

Galeotti, A., C. Ghiglino, and F. Squintani (2010): "Strategic information transmission in networks," *Working paper.*

Galeotti, A., and S. Goyal (2010): "The law of the few," *forthcoming in the American Economic Review.*

Galeotti, A., S. Goyal, and J. Kamphorst (2006): "Network formation with heterogeneous players," *Games and Economic Behavior*, 54(2), 353–372.

Gladwell, M. (2000): *The Tipping Point: How little things can make a big difference.* Little Brown.

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

Goyal, S. (2007): *Connections: an introduction to the economics of networks.* Princeton University Press, Princeton, New Jersey.

Hagenbach, J., and F. Koessler (2010): "Strategic communication networks," *Review of Economic Studies*, 77(3), 1072–1099.

Hojman, D., and A. Szeidl (2008): "Core and periphery in networks," *Journal of Economic Theory*, 139(1), 295–309.

Jackson, M. (2004): *A survey of models of network formation: stability and efficiency. In group formation in Economics; Networks, clubs and coalitions, edited by Gabrielle Demange and Myrna Wooders.* Cambridge University Press.

——— (2008): *Social and economic networks.* Princeton University Press, Princeton, New Jersey.

Jackson, M., and B. Rogers (2006): "Search in the formation of large networks: How random are social networks?," *American Economic Review*, 97(3), 890–915.

Mobius, M., T. Phan, and A. Szeidl (2010): "Treasure Hunt: Social Learning in Real-World Social Networks," *Working paper.*

Morgan, J., and P. Stocken (2008): "Information aggregation in polls," *American Economic Review*, 98(3), 864–896.

Morris, S., and H. S. Shin (2002): "Social value of public information," *American Economic Review*, 92(5), 1521–1534.

Smith, L., and P. Sørensen (2000): "Pathological outcomes of observational learning," *Econometrica*, 68(2), 371–398.

——— (2010): "Rational social learning with random sampling," *Working paper.*