Bayesian Social Learning in a Dynamic Environment*

Krishna Dasaratha Benjamin Golub Nir Hak†

January 7, 2018

Abstract

Bayesian agents learn about a moving target, such as a commodity price, using private signals and the past estimates of their neighbors in an arbitrary network. The weights they place on these sources of information are endogenously determined by the precisions and correlations of individuals' estimates; these weights, in turn, determine future correlations. We study stationary equilibria—ones in which all of these quantities are constant over time. Equilibria in linear updating rules always exist. This yields a fully Bayesian learning model as tractable as the commonly-used DeGroot heuristic. Equilibria and the comparative statics of learning outcomes can be readily computed even in large networks. Substantively, we identify pervasive inefficiencies in Bayesian learning. In any stationary equilibrium where agents put positive weights on neighbors' actions, learning is Pareto inefficient in a generic network: agents rationally overuse social information and underuse their private signals.

*We are grateful to Nageeb Ali, Drew Fudenberg, Matthew O. Jackson, Eric Maskin, Matthew Rabin, Evan Sadler and Omer Tamuz for valuable discussions.

†Department of Economics, Harvard University, Cambridge, U.S.A., dasarath@g.harvard.edu, bgolub@fas.harvard.edu, nirhak@g.harvard.edu.

1 Introduction

Consider a group of agents engaged, over a period of time, in learning about an evolving fundamental, such as a commodity price. Some have noisy information about the fundamental. Agents also have access to the estimates of some of the others—their contacts or neighbors in a network—and they learn from one another over time. A concrete example comes from learning about market prices in a setting with information frictions. Consider a crop whose price evolves over time due to shocks (weather, demand, etc.). Farmers who grow this crop would like to learn about its price, but most do not have frictionless access to up-to-date information about it. Many have *some* relevant information—for example, they have local observations of weather, or have seen a transaction. Some people have better information than others—e.g., some visit a marketplace often and observe the prevailing price. Others can learn from them. There is a network describing who can learn from whom: a directed link from Ann to Bob means that Ann can observe Bob's current estimate of the state, though *not* his private signal. This may be through observation of Bob's behavior (e.g., what price he asks for the crop) or through communication. Settings of this type have been studied in the development literature, where information about prices is highly imperfect, yet important for households making production and consumption decisions (see, e.g., Jensen, 2007; Srinivasan and Burrell, 2013).

Basic questions in such a setting include: When will the community learn well? What kinds of errors is it prone to? Will some benefit from social information more than others? The network's role in the answers has been an important focus of recent research (see, e.g., Mossel, Sly, and Tamuz (2015); Alatas, Banerjee, Chandrasekhar, Hanna, and Olken (2016)), especially because data on networks is increasingly available. A model for the analysis of these issues should, ideally, be capable answering questions in four categories:

1. *Positive predictions about influence and learning:* How will network position and signal precision translate into influence?

2. *Identification:* How can the quality of underlying information be identified from observed outcomes?

3. *Welfare analysis:* What are the inefficiencies inherent in learning processes?

4. *Policy analysis and counterfactuals:* Is learning improved by natural interventions such as injecting information or adding communication opportunities? What are the unintended consequences? Who should be targeted with interventions?

Key features of the model

We propose a new model of social learning in a dynamic environment suited to the challenge of addressing questions in all four categories. The environment corresponds to the example described above with the crop price: network neighbors repeatedly learn from one another, in an arbitrary directed or undirected network structure. The agents' updating rules have standard foundations: learning is Bayesian and minimizes the error of individuals' estimates of the state of interest (e.g., the price) given their information. Under a convenient parametric structure that we discuss below, equilibrium learning rules (the weights agents place on each other) take a simple, time-invariant form—the same one found in the widely-used heuristic DeGroot model—and these quantities can all be computed efficiently.

As a result, the model fulfills the requirements above. First, it yields predictions about how the environment determines influence and learning, as well as simple equations characterizing (constant) equilibrium weights in terms of structural parameters. This makes it feasible to identify structural parameters from observed outcomes. Since agents' preferences for minimizing error are explicitly modeled and reflected in their behavior, the model is suited to standard welfare analysis based on revealed preferences. Perhaps more importantly, the fact that agents are maximizing their utilities allows us to model how their learning rules will react to changes in the environment, permitting standard counterfactual analysis. In Section 7, we compare our findings to classical social learning models such as sequential learning and discuss how these features, taken together, go beyond existing models.

The key to the tractability is the following parametric assumption: the state (commodity price in our example) drifts around according to a stationary process—namely, an AR(1) process—and agents receive Gaussian signals of its current value. Each agent makes an estimate of the state using her own signal and the estimates of her neighbors; this estimate is then used by her neighbors in the next round, etc. Using the resulting stationary structure, our first result gives conditions for the existence of a stationary equilibrium with linear learning rules. Indeed, the laws of motion of individuals' estimates implied by Bayesian updating happen to be closely related to those of the DeGroot (1974) model, widely used for its tractability.[1] We discuss this connection further in Section 7.4.2. The use of a changing state evolving according to an AR(1) law has technical advantages for us, permitting the tractable solutions just discussed. But more importantly, it fits certain applications (e.g.,

[1]Those rules are usually studied in a different environment from ours—information comes in only once—where they are *not* part of an equilibrium (DeMarzo et al., 2003).

learning about a price) substantially better than some standard modeling choices such as a fixed, binary state.

Results

Equilibrium existence and uniqueness First, we prove existence of linear equilibria and analyze them by studying a fixed-point problem that characterizes any such equilibrium. When an agent uses her information to learn about the state, she must think about how precise and how correlated his signals about it are. Agents' updating rules define a variance-covariance structure of each agent's observations and the state of interest, and thus determine each agent's weights on neighbors and signals about the fundamental. That, in turn, determines the variance-covariance structure of next period's estimates. Thus we obtain a self-mapping on the space of variance-covariance structures, which we prove has a fixed point; this yields the stationary learning rules we study. We also show uniqueness of equilibria for a class of networks that includes complete bipartite networks such as the star network.

Characterization of equilibrium in terms of structural parameters With or without uniqueness, equilibrium is characterized by simple equations relating each agent's weights, and therefore her precision, to the precisions and correlations of her neighbors' estimates. Thus, assumption of equilibrium weights can be tested against alternatives, such as naive rules, which we also discuss.

Within the model, structural parameters can be estimated readily from observed behavior. More precisely, an econometrician who observes a panel consisting only of the estimates of the agents and the realized states can use these data and a VAR model to recover the network structure and the precisions of the underlying signals available to agents in the network. We will discuss below the role these estimates can play in analyzing counterfactuals, such as injecting information in a particular place. The precisions of individuals' private information may also be an object of independent interest.

General inefficiency Our main substantive focus is on externalities that occur when Bayesian agents learn from each other over time, and the associated inefficiencies. Our most general result is that in any stationary equilibrium on virtually any connected network, assuming agents put positive weights on all neighbors, the learning strategies agents use are necessarily Pareto inefficient. That is, there is some (non-equilibrium) profile of constant learning weights that agents could use under which everyone would obtain higher utility. The inefficiency involves *overuse of social information*: if agents were to place more weight on their own signals, it would be easier to extract the private signals of each individual, making learning better for everyone.

The reason for this inefficiency and others that we document comes from a kind of confounding that is central to our analyses. We illustrate it with an example. Consider an agent, Ann, who observes several friends, all of whom observe a common source, Zeke (who, for simplicity, does not observe anybody). Ann would like to take Zeke's signal into account, but because Zeke's private signal is no better than any other individual's, she would not like to give it any extra weight. If she could observe the estimates of Zeke's that play a role in all friends' estimates, she would efficiently subtract them out from the friends' estimates. However, since everyone Ann sees observes Zeke, Ann cannot avoid overweighting Zeke's signals relative to how much she weights the equally good information of her friends, and as a result she gets less value from her friends' information. She would be better off if her friends paid more attention to their own information and less to Zeke's. The role that Zeke's signals play in this example can be played by any information that Ann does not observe (and so cannot optimally "subtract out"), and it turns out, in our model, that this force is powerful and generic enough to cause some inefficiency in virtually any network.

Perverse consequences of information improvements Another sense in which the system does not make the best possible use of information is that there are situations in which improving one agent's private signals by making them more precise can make everyone (including that agent) worse off. The idea is closely related to the example just sketched: the improvement drives some to increase their weights on that agent, which confounds the learning of many others. The key is an externality in learning: While weighting a new source of information is in the interest of every individual, their increased attention to it worsens the signal extraction problems of the others (directly and indirectly), and the latter effect can dominate, even for the agent whose information was improved.

A final kind of inefficiency is that adding links may harm everyone. On the one hand, links give each individual access to more information, and they help information to flow faster throughout the system. However, in some cases, adding a single link or completing the graph hurts all agents: it makes all agents' estimates less precise in equilibrium, again by exacerbating signal extraction problems.

Learning in large networks We next ask how well agents learn in large networks. Our asymptotic benchmark is the accuracy achieved by an agent with access to his own private signals and *perfect* knowledge of the previous period's state. This could be achieved by an agent with a large number of independent social observations. The obstacle to achieving this benchmark is that agents actions' indirectly weight guesses of past states, and these past states deviate from today's state. As long as there is some reciprocal observation, the confounding force we have identified will lead to correlated error throughout the network and prevents agents from achieving the benchmark.

Outline

Section 2 sets up the basic model and discusses interpretation of the model. Section 3 defines our equilibrium concept, shows that equilibria exist, and gives our results about uniqueness and convergence. In Section 4, we describe how an analyst observing actions generated according to our model can estimate informational parameters. Section 5 gives our general result on Pareto inefficiency of equilibrium and exhibits two perverse comparative statics. In Section 6, we document obstructions to information aggregation on large networks. Section 7 discusses other real-world examples, a version of our model with naive agents, and how information shocks propagate before concluding with a discussion of related literature.

2 Model

2.1 Description

State of the world At each discrete instant (also called period) of time, $t \in \{\ldots, -2, -1, 0, 1, 2, \ldots\}$, there is a state of the world, a random variable θ_t taking values in $\mathbb{R}$. This state evolves as an AR(1) stochastic process. That is,

$$\theta_{t+1} = \rho\theta_t + \epsilon_{t+1},$$

where $\rho \in (0, 1]$ is a constant and $\epsilon_{t+1} \sim \mathcal{N}(0, \sigma_\epsilon^2)$ are independent innovations. We can write explicitly

$$\theta_t = \sum_{\ell=0}^{\infty} \rho^\ell \epsilon_{t-\ell},$$

and thus $\theta_t \sim \mathcal{N}\left(0, \frac{\sigma_\epsilon^2}{1-\rho^2}\right)$. We make the normalization $\sigma_\epsilon = 1$.

Information and observations The set of *nodes* is $N = \{1, 2, \ldots, n\}$. A *network* G represents observation opportunities. It is given by the set of nodes N and a set of directed edges $E \subseteq N \times N$. We denote the neighbors of node i by N_i, which can be taken to include or exclude i itself.

At each time t, the *agent* (i, t) at node i is called to take a single action, $a_{i,t}$. (We will sometimes refer simply to agent i, when the time period at which she is acting is understood.) Before taking her action, however, the agent receives some information. First, there is a private signal,

$$s_{i,t} = \theta_t + \eta_{i,t},$$

where $\eta_{i,t} \sim \mathcal{N}(0, \sigma_i^2)$ has a variance $\sigma_i^2 > 0$ that depends on the agent but not on the time period. All the $\eta_{i,t}$ and ϵ_t are independent of each other.

Agent (i, t), in addition to observing her private signal, also observes the time-$(t - \ell)$ actions $a_{j,t-\ell}$ at all neighboring nodes $j \in N_i$ for $\ell \in \{1, 2, \ldots, m\}$, where m is the *memory*.[2] One simple interpretation of this is that an agent is born at period $t - m$, and experiences m periods of passive childhood, observing the actions in her neighborhood during these periods. Then as an adult, at period t, she makes her own decision, $a_{i,t}$, and then dies. For any time t, we denote by $\boldsymbol{a}_{N_i,t}$ the vector $(a_{j,t})_{j \in N_i}$. This observation structure is common knowledge.

A vector of all of agent (i, t)'s observations—$s_{i,t}$ and the neighbors' past actions—defines an agent's information. An important special case will be $m = 1$, where there is one period of memory, so that the agent's information is $(s_{i,t}, (a_{j,t-1})_{j \in N_i})$.

Actions, preferences, and choices As stated above, in each period t, agent (i, t) at each node i must choose an action $a_{i,t} \in \mathbb{R}$. Utility is given by

$$u_{i,t}(a_{i,t}) = -(a_{i,t} - \theta_t)^2.$$

The agent dies after making her decision, and so she makes the optimal choice for the current period given her information. By a standard fact about squared-error loss functions, given the distribution of $(\boldsymbol{a}_{N_i,t-\ell})_{\ell=1}^m$, she sets:

[2]We assume that the memory m is the same for all observed neighbors and that agents do not remember past private signals, but these assumptions do not substantially change results.

$$a_{i,t} = \mathbb{E}[\theta_t \mid s_{i,t}, (\boldsymbol{a}_{N_i,t-\ell})_{\ell=1}^m].$$

We will discuss the solution concept, and in particular the agent's beliefs about the distribution of $(\boldsymbol{a}_{N_i,t-\ell})_{\ell=1}^m$, below.

2.2 Interpretation

The agents are fully Bayesian given the information they have access to. Much of our analysis is done for arbitrary finite m; we view the restriction to finite memory as an assumption that avoids technical complications but has little substantive content. The model generalizes "Bayesian without Recall" agents from the engineering and computer science literatures (e.g., Rahimian and Jadbabaie, 2017), which, within our notation, is the case of $m = 1$.

Though an agent can only directly observe actions from the past m periods, these actions will indirectly incorporate information from further in the past, and so in many cases actions can be very accurate. We view this as a stylized representation of a realistic feature: it is prohibitively costly to keep durable records of all one's past observations and impressions that are relevant to current decisions, but current practices reflect accumulated information.

Note that an agent does not have access to the past private signals observed either at her own node or at neighboring ones. This is not a critical choice—similar models could be solved taking some signals to be observable—but it is worth explaining. Whereas $a_{i,t}$ is an observable choice, such as a published evaluation of an asset or a mix of inputs actually used by an agent, the private signals are not shareable. Though we model the signals for convenience as real numbers, a more realistic interpretation of these is an aggregation of all of an agent's experiences, impressions, etc., and these may be difficult to summarize or convey.

Finally, our agents act once and do not consider future payoffs of their dynasties. These assumptions are made so that an agent's equilibrium action reflects her best current guess about the state, and turn off the possibility that she may try to strategically manipulate the future path of social learning. Substantively, like Gale and Kariv (2003) and Harel, Mossel, Strack, and Tamuz (2017)[3], we view these types of assumptions as a clean way of capturing that in our applications, such strategic considerations—if present at all—are

[3]See also Manea (2011) and Talamàs (2017) in the bargaining literature.

likely to be secondary to matching the state. Equivalently, we could simply assume that agents sincerely announce their subjective expectations of the state, as in Geanakoplos and Polemarchakis (1982).

3 Equilibrium

In this section we define equilibrium, establish its existence, and discuss uniqueness and convergence to equilibrium.

3.1 Equilibria in Linear Strategies

A strategy of an agent is *linear* if the action taken is a linear combination of the random variables in her information set (and a constant). We will study only *stationary equilibria in linear strategies*—ones in which all agents' strategies are linear with time-invariant coefficients—though, of course, we will allow agents to consider deviating to arbitrary strategies.

We will now argue that in studying agents' best responses to stationary linear strategies, we may restrict attention to linear strategies. If stationary linear strategies have been played up to time t, we can express each action up until time t as a summation of past signals.[4] Because all innovations ϵ_t and signal errors $\eta_{i,t}$ are independent and Gaussian, it follows the joint distribution of any finite random vector of the past errors $(a_{i,t-\ell'} - \theta_t)_{i \in N, \ell' \geq 1}$ is multivariate Gaussian. Thus, $\mathbb{E}[\theta_t \mid s_{i,t}, (\boldsymbol{a}_{N_i,t-\ell})_{\ell=1}^m]$ is a linear function of $s_{i,t}$ and $(\boldsymbol{a}_{N_i,t-\ell})_{\ell=1}^m$ (Kay, 1993, Example 4.4). It follows that solving for equilibrium can be reduced to searching for the weights agents place on the variables in their information set.

3.2 Covariance Matrices

Having reduced the problem to one of choosing weights, our next observation is that the optimal weights for an agent to place on her sources of information depend on the accuracy and correlation of these sources, as we explain in this subsection and the next.

Let $\boldsymbol{V}_t$ be the $nm \times nm$ covariance matrix of the vector $(\rho^\ell a_{i,t-\ell} - \theta_t)_{\substack{i \in N \\ 0 \leq \ell \leq m-1}}$. The entries of this vector are the distances between the actions in the past m periods (or, more precisely, the best predictors of θ_t given those actions) and the current state of the world.

[4] To ensure this series is almost surely convergent, note in any best response of agent (i,t), the random variable $a_{i,t} - \theta_t$, has a finite variance: each player seeks to minimize the variance of this error and always has the option of relying on her own private signal, in which case her error has finite variance.

Denote the space of such covariance matrices by $\mathcal{V}$. In the case $m = 1$, the typical element is simply the covariance matrix $\boldsymbol{V}_t = \mathrm{Cov}(a_{i,t} - \theta_t)$. The matrix $\boldsymbol{V}_t$ records covariances of action errors: diagonal entries measure the accuracy of each action, while off-diagonal entries indicate how correlated the two agent's action errors are. We will sometimes use the shorthand as $\nu_{i,t}^2$ for the diagonal entries $V_{ii,t} = \mathrm{Var}(a_{i,t} - \theta_t)$.

3.3 Best-Response Weights

A strategy profile is an equilibrium if weights agents place on the variables in their information set minimize their posterior precision. We now characterize these in terms of the covariance matrices we have defined. Let $\boldsymbol{V}_{N_i,t-1}$ be a sub-matrix of $\boldsymbol{V}_{t-1}$ that contains only the rows and columns corresponding to neighbors of i[5] and let

$$\boldsymbol{C}_{i,t-1} = \begin{pmatrix} & & & 0 \\ & \boldsymbol{V}_{N_i,t-1} & & 0 \\ & & & \vdots \\ 0 & 0 & \ldots & \sigma_i^2 \end{pmatrix}.$$

Conditional on observations $(\boldsymbol{a}_{N_i,t-\ell})_{\ell=1}^m$ and $s_{i,t}$, the state θ_t is normally distributed with mean

$$\frac{\mathbf{1}^T \boldsymbol{C}_{i,t-1}^{-1}}{\mathbf{1}^T \boldsymbol{C}_{i,t-1}^{-1} \mathbf{1}} \cdot \begin{pmatrix} \rho \boldsymbol{a}_{N_i,t-1} \\ \vdots \\ \rho^m \boldsymbol{a}_{N_i,t-m} \\ s_{i,t+1} \end{pmatrix} \tag{1}$$

(see Example 4.4 of Kay (1993)).[6] This gives $\mathbb{E}[\theta_t \mid s_{i,t}, (\boldsymbol{a}_{N_i,t-\ell})_{\ell=1}^m]$ (recall that this is the $a_{i,t}$ the agent will play). Expression (1) is a linear combination of the agent's signal and the observed actions; the coefficients in this linear combination depend on the matrix $\boldsymbol{V}_{t-1}$ (but not on realizations of any random variables).

We denote by (W_t, w_t^s) the weights agents use in period t, with $w_t^s \in \mathbb{R}^n$ being the weights agents place on their private signals and W_t recording the weights they place on their other information. We do not describe the indexing of coefficients in W_t explicitly

[5] Explicitly, $\boldsymbol{V}_{N_i,t-1}$ are the covariances of $(\rho^\ell a_{j,t-\ell} - \theta_t)$ for all $j \in N_i$ and $\ell \in \{1, ..., m\}$.

[6] We implicitly assume here that agents' prior beliefs about the state are an improper distribution giving all states equal weight. With $\rho < 1$, agents' priors could instead be equal to the stationary distribution of the state and the analysis would not change substantially except in the discussion of equilibrium uniqueness.

in general, but when $m = 1$, we refer to the weight agent i places on $a_{j,t-1}$, j's action yesterday, as $W_{ij,t}$ and the weight on $s_{i,t}$, her private signal, as $w^s_{i,t}$.

In view of the formula (1) for the optimal weights, we can compute the resulting next-period covariance matrix $\boldsymbol{V}_t$ from the previous covariance matrix. This defines a map $\Phi : \mathcal{V} \to \mathcal{V}$, given by

$$\Phi : \boldsymbol{V}_{t-1} \mapsto \boldsymbol{V}_t \tag{2}$$

which we will study, for instance, in characterizing equilibria.

In the case of $m = 1$, we will write out the map explicitly using our notation for weights above:

$$\boldsymbol{V}_{ii,t} = (w^s_{i,t})^2\sigma_i^2 + \sum W_{ik,t}W_{ik',t}(\rho^2\boldsymbol{V}_{kk',t-1} + 1) \text{ and } \boldsymbol{V}_{ij,t} = \sum W_{ik,t}W_{i'k',t}(\rho^2\boldsymbol{V}_{kk',t-1} + 1). \tag{3}$$

There is an analogous (but more cumbersome) expression for $m > 1$, which we omit.

A natural question is how agents would, in practice, form the beliefs over others' play that are needed to carry out the inferences discussed above. The simplest foundation is that there is common knowledge (perhaps passed down from generation to generation) of the equilibrium covariance matrix $\boldsymbol{V}_{t-1}$, which is sufficient for playing optimally. Indeed, even if the covariance matrix is changing over time, the covariance matrix $\boldsymbol{V}_{t-1}$ of the previous period's errors is sufficient to compute the next period's $\boldsymbol{V}_t$ when agents play optimally. Thus there is a manageable information that must be carried from period to period. Alternately, at equilibrium the empirical covariances over many periods converge to the entries of the analytic covariance matrix. Intuitively, after many periods of play agents likely develop good estimates of how accurate each player is and how correlated two neighbors are. These estimates are sufficient to determine best responses.

3.4 Equilibrium Existence

Consider the map Φ defined in (2). Stationary equilibria in linear strategies correspond to fixed points of the map Φ. More generally, one can use this map to study how covariances of actions evolve given any initial distribution of play. Note that $\boldsymbol{V}_t$ and the map Φ are deterministic, and we can study the resulting dynamical system without considering the particular realizations of signals.

Our first result concerns the existence of equilibrium:

Proposition 1. *There is a matrix $\widehat{\boldsymbol{V}}$ of covariances such that $\Phi(\widehat{\boldsymbol{V}}) = \widehat{\boldsymbol{V}}$. Equivalently, a stationary equilibrium in linear strategies exists.*

At the stationary equilibrium, the covariance matrix and all agent strategies are time-invariant. As in the DeGroot model, agent actions are linear combinations of observations with stationary weights (which we refer to as $\widehat{W}_{ij,t}$ and $\widehat{w}_i^s$). We discuss the relationship between our model and DeGroot learning further in Section 7.

The idea of the argument is as follows. First, all variances are bounded when agents best-respond to any beliefs about prior play, because all agents' actions must be at least as precise in estimating θ_t as their private signals, and cannot be more precise than estimates given perfect knowledge of yesterday's state combined with the private signal. Because the Cauchy-Schwartz inequality bounds covariances in terms of the corresponding variances, all entries of the image of Φ are bounded. Once these bounds are established, the result follows from the Brouwer fixed point theorem.

When $m = 1$, the proof gives bounds $\widehat{V}_{ii} \in [\frac{1}{1+\sigma_i^{-2}}, \sigma_i^2]$ on equilibrium variances and $\widehat{V}_{ij} \in [-\sigma_i\sigma_j, \sigma_i\sigma_j]$ on equilibrium covariances.

The equilibrium variances and covariances $\widehat{\boldsymbol{V}}$ satisfy

$$\widehat{V}_{ii} = (\widehat{w}_i^s)^2\sigma_i^2 + \sum \widehat{W}_{ik}\widehat{W}_{ik'}(\rho^2\widehat{V}_{kk'} + 1) \text{ and } \widehat{V}_{ij} = \sum \widehat{W}_{ik}\widehat{W}_{jk'}(\rho^2\widehat{V}_{kk'} + 1)$$

(or corresponding equations with $m > 1$). Substituting equation (1) and rearranging terms, we find the entries of $\widehat{\boldsymbol{V}}$ are the solutions to a system of polynomial equations. These equations have large degree and cannot be solved analytically except in very simple cases, but can be used to numerically solve for equilibria.

3.5 Convergence and Equilibrium Uniqueness: Special Cases

Now that we have shown that an equilibrium exists, we ask whether the equilibrium is unique and whether play converges to the equilibrium. While these questions are difficult in general, we give affirmative answers for two special classes of networks.

To specify our notion of convergence to the equilibrium, suppose that we start with some distribution of actions $\boldsymbol{V}_0$ in period 0. For example, we could assume that all players choose actions $a_{i,t}$ equal to their private signal in period 0. We can then ask whether, under the updating process yielding $\boldsymbol{V}_t = \Phi(\boldsymbol{V}_{t-1})$ (recall Section 3.3) play converges to the equilibrium covariance matrix $\widehat{\boldsymbol{V}}$.

Our next result shows that for a particular class of networks, both uniqueness and convergence hold:

Proposition 2. *Suppose that $m = 1$ and whenever $i, i' \in N_j$, we have $N_i = N_{i'}$. Then $\widehat{\boldsymbol{V}}$ is unique and $\boldsymbol{V}_t$ converges to $\widehat{\boldsymbol{V}}$ from any initial values.*

The condition on the network says that any two agents observed by j have the same neighborhoods. The key implication is that if agent j observes i and i', then i and i' have the same information about previous period actions. For undirected graphs, the graphs that satisfy this condition are precisely the complete bipartite graphs (without self-links) and complete graphs (with self-links).

The idea of the proof is that we can think of an agent putting a weight on her private signal and a *social signal*, which is the linear combination of $\boldsymbol{a}_{N_i,t-1}$ (her social observations) yielding the best estimate of θ_t. Then, instead of looking at all of $\boldsymbol{V}_t$, we can consider the space of variances of these social signals. The condition on the network implies that these estimates determine the updating process, so instead of studying $\mathcal{V}$ we can consider the map $\widetilde{\Phi}$ induced by Φ on this new space. The proof shows $\widetilde{\Phi}$ is a contraction. The advantage to this new space is that all coordinates are variances of social signals, which are the minima of optimization problems (for estimating θ_t). This means we can apply the envelope theorem to understand $\widetilde{\Phi}$: when variances change slightly, we can calculate the change in the updated variances as if agents do not change their weights. The result follows from a (somewhat involved) calculation of the derivatives of $\widetilde{\Phi}$.

A directed network is a *directed acyclic graph* if there is no finite sequence of nodes $i_1, ..., i_m$ with $i_1 = i_m$ such that there is an edge from i_k to i_{k+1} for each k. The same uniqueness and convergence conclusions apply to directed acyclic graphs:

Proposition 3. *Suppose G is a directed acyclic graph. Then $\widehat{\boldsymbol{V}}$ is unique and $\boldsymbol{V}_t$ converges to $\widehat{\boldsymbol{V}}$ from any initial values.*

The proof is simpler than for the previous proposition; it is by induction on the network size. Agents with no observations follow their private signals in every period, so their variances converge immediately. After these agents' variances and covariances with each other become constant, the variances in the next level of agents (those whose neighbors have no observations themselves) will converge as well. We can iterate this argument until the whole network has converged.

On more general networks, we might hope to show that Φ induces a contraction on a suitable normed vector space. Usually that space must record not only variances, which can be understood using the envelope theorem as in Proposition 2, but also covariance terms that are not as well-behaved on complex networks.

Example 1. Consider four agents in an undirected line with $\sigma_i^2 = 1$ for each i and let $\alpha = \frac{19}{20}$. Proposition 2 does not apply, but we can check numerically that the system of polynomial equations defining the equilibrium has a unique solution. However, we find numerically that Φ does not induce a contraction on V in any of several standard matrix norms—$\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty$, the sup norm or the Froebenius norm—even when we restrict attention to the image $\Phi(\mathcal{V})$.

4 Identification and Testable Implications

This section discusses how our model can be tested and how its parameters can be identified using empirical data from a social learning process. The key feature making the model econometrically well-behaved is that, in the solutions we focus on, agents' actions are linear functions of the random variables they observe; moreover, the evolution of the state and arrival of information creates exogenous variation. We leverage these features both for testing and identification purposes.

Assume the following. The analyst obtains noisy measurements $\overline{a}_{i,t} = a_{i,t} + \xi_{i,t}$ of agent's actions (where $\xi_{i,t}$ are i.i.d., mean-zero error terms) in some set of periods. He knows the parameter ρ governing the stochastic process, but may not know the network structure or the qualities of private signals $(\sigma_i)_{i=1}^n$. Suppose also that, ex post, the analyst can observe the state θ_t.[7] In many applications, the state may be difficult to observe at the time and place agents are making their estimates, but can be measured separately. For example, in the situation discussed in the introduction, where farmers are estimating a market price, we can elicit their current expectations and later compare that with data on prices separately collected at marketplaces.

Now, consider *any* steady state in which agents put constant weights W_{ij} on their neighbors and w_i^s on their private signals over time. We will discuss the case of $m = 1$ to save on notation, though all the statements here generalize readily to arbitrary m. As we will now discuss, linear regression can be used to identify the weights agents are using, and to back out the structural parameters our model when it applies. The strategy does not rely on uniqueness of equilibrium.

[7] In some cases the analyst would instead observe (a proxy for) the private signal $s_{i,t}$ of agent i. Such a regressor can be used below in place of θ_t; our analysis will go through essentially unchanged.

More precisely, we can identify the weights agents are using through a regression. In steady state,

$$\overline{a}_{i,t} = \sum_j W_{ij} \rho \overline{a}_{j,t-1} + w_i^s \theta_t + \zeta_{i,t}, \tag{4}$$

where $\zeta_{i,t} = w_i^s \eta_{i,t} - \sum_j W_{ij} \rho \xi_{j,t-1} + \xi_{i,t}$ are error terms i.i.d. across time. The first term of the expression for $\zeta_{i,t}$ is the error of the signal that agent i receives at time t, and the summation combines the measurement errors from the observations $\widetilde{a}_{j,t}$ the previous period. This system defines a VAR(1) process (or generally VAR(m) for memory length m). Thus, we can obtain consistent estimators $\widetilde{W_{ij}}$ and $\widetilde{w}_i^s$ for W_{ij} and w_i^s, respectively. [8]

We now turn to the case in which agents are using equilibrium weights. First, and most simply, our estimates of agents' equilibrium weights allow us to recover the network structure. If the weight $\widehat{W_{ij}}$ is non-zero for any i and j, then agent i observes agent j. Generically the converse is true: if i observes j then the weight $\widehat{W_{ij}}$ is non-zero. Thus, tests of whether the recovered social weights are nonzero generically identify network links. For such tests, and generally, the standard errors in the estimators can be obtained by standard techniques.[9]

Now we examine how to identify the structural parameters assuming an equilibrium is played, and also how to test the assumption of equilibrium.

The first step is to compute the empirical covariances $\widetilde{\widehat{V}}_{ij}$ of action errors from observed data. Under the assumption of equilibrium, we now show how to determine the signal variances using the fact that equilibrium is characterized by $\Phi(\widehat{\boldsymbol{V}}) = \widehat{\boldsymbol{V}}$ and recalling the explicit formula (3) for Φ. In view of this formula, the signal variances σ_i^2 are uniquely determined by the other variables:

$$\widehat{V}_{ii} = \sum_j \sum_k \widehat{W}_{ij} \widehat{W}_{ik} (\rho^2 \widehat{V}_{jk} + 1) + (\widehat{w}_i^s)^2 \sigma_i^2.$$

Replacing the model parameters other than σ_i^2 by their empirical analogues, we obtain a consistent estimate $\widetilde{\sigma}_i^2$ of σ_i. This estimate could be directly useful—for example, to

[8]For our purpose here of recovering the weights, we may also replace θ_t in the regression with any unbiased estimate of θ_t such that all errors in the model are independent of the error in this estimate. The ensuing analysis in this section can be adapted to this case; assuming away this analyst error streamlines the exposition.

[9]Methods involving regularization may be practically useful in identifying links in the network. Manresa (2013) proposes a regularization (LASSO) technique for identifying such links (peer effects). In a dynamic setting such as ours, with serial correlation, such techniques will generally be more complicated.

an analyst who wants to choose an "expert" from the network and ask about her private signals directly.

Recall that our basic regression for recovering the weights agents use relies only on constant linear strategies and does not assume that agents best respond. This fact can be used to econometrically examine the behavior each agent is using. For instance we can test if an agent behaves rationally or according to the naive model of Section 7.2. Consider the hypothesis that agent i is behaving according to the rational model in this paper. Our proposed test of this hypothesis is as follows. First, extract the graph as discussed above. Then perform the regression in (4) to obtain regression coefficients $\widetilde{W}$, empirical covariances $\widetilde{\widehat{\boldsymbol{V}}}$, and estimated individuals' signal variances $\widetilde{\boldsymbol{\sigma}}$. Assuming that agent i behaves according to the model in a steady state, the weights i puts on neighbors are a best response to $\widehat{\boldsymbol{V}}$ and $\boldsymbol{\sigma}$. In view of this, we can calculate the weights that are best responses to the empirical covariances $\widetilde{\widehat{\boldsymbol{V}}}$ and estimated individuals' signal variances $\widetilde{\boldsymbol{\sigma}}$. Denote these best-response weights by W_{ij}^{model}. Under the null hypothesis that the agents behave according to the model, these extrapolated updating weights W_{ij}^{model} and the updating weights $\widetilde{W}_{ij}$ estimated from the regression should be the same. We can thus perform an F-test to check whether the difference $W_{ij}^{model} - \widetilde{W}_{ij}$ is significantly different than zero. The standard errors may be bootstrapped in the standard way.

5 Welfare

5.1 Pareto-Inefficiency of Equilibrium Weights

At equilibrium, each agent chooses her weights by a Bayesian calculation. We compare welfare under these weights to welfare when all agents use another set of weights in each period. Theorem 1 gives a general condition under which there exist stationary weights that are better for all agents than the equilibrium weights.

Definition 1. The *steady state* associated with weights W and w^s is the covariance matrix $\boldsymbol{V}_t$ such that if next-period actions are set using weights (W, w^s), then $\boldsymbol{V}_{t+1} = \boldsymbol{V}_t$.

In this definition of steady state, instead of optimizing (as at equilibrium) agents use fixed weights in all periods. By a straightforward application of the contraction mapping theorem, any non-negative weights under which covariances remain bounded at all times determine a unique steady state.

Theorem 1. *Suppose the network G is strongly connected and some agent has more than one neighbor. If all weights are positive at an equilibrium, then the variances at that equilibrium are Pareto-dominated by variances at another steady state.*

The result says that there exist stationary weights which are better for all agents (in terms of the associated steady-state welfare) than the weights at any equilibrium. The basic idea is that if agents place marginally more weight on their private signals, this introduces more independent information that eventually benefits everyone.

Next, we briefly discuss the sufficient conditions. It is clear that some condition on neighborhoods is needed: If every agent has exactly one neighbor, there are no externalities and the equilibrium weights are Pareto optimal. In fact, the same result (with the same proof) applies to a larger class of networks: it is sufficient that, starting at each agent, there are two paths of some length k to two distinct agents. Finally, the condition on equilibrium weights says that no agent anti-imitates any of her neighbors. This assumption makes the analysis tractable, but is unrelated to the basic intuition guiding the proof.

The idea of the proof is to begin at equilibrium and then marginally shift an appropriate agent's weights toward their private signal. By the envelope theorem, this means agents' actions are less correlated but not significantly worse in the next period. We show that if all agents continue using these new weights, the decreased correlation eventually benefits everyone.

In the last step, we use the absence of anti-imitation, which implies that the updating function associated with agents using fixed weights (instead of an optimization procedure) is monotonic. To first order, some covariances decrease while others do not change after one period under the new weights. Monotonicity of the updating function and strong connectedness imply that eventually all agents' variances decrease.

Example 2. Consider $n = 100$ agents in an undirected circle—i.e., each agent observes the agent to her left and the agent to her right. Let $\sigma_i^2 = \sigma^2$ be equal for all agents and $\rho = .9$. The equilibrium strategies place weight $\widehat{w}^s$ on private signals and weight $\frac{1}{2}(1 - \widehat{w}^s)$ on each observed action.

When $\sigma^2 = 10$, the equilibrium weight is $\widehat{w}^s = 0.192$ while the welfare-maximizing symmetric weights have $w^s = 0.234$. That is, weighting private signals substantially more is Pareto improving. When $\sigma^2 = 1$, the equilibrium weight is $\widehat{w}^s = 0.570$ while the welfare maximizing symmetric weights have $w^s = 0.586$. The inefficiency persists, but the equilibrium strategy is now closer to the optimal strategy.

In a review of sequential learning experiments, Weizsäcker (2010) find that subjects weight their private signals more heavily than is empirically optimal. Theorem 1 implies that in our environment with optimizing agents, deviations in the same direction are welfare-improving.

5.2 Value of Private Information

In this subsection and the next, we give examples where apparent informational improvements are harmful to all agents at equilibrium. This subsection shows that decreasing the precision of an agent's private signal can give a Pareto improvement.

More precisely:

Proposition 4. *There exists a network G and an agent $i \in G$ such that increasing σ_i^2 gives a Pareto improvement in equilibrium variances.*

To sketch the intuition, suppose a central agent observes many neighbors with private signals confounded by some common information sources. If one of these neighbors gave up her private signal, then her action would help the central agent parse other observations by isolating the common confound. This central person then aggregates information better and distributes a more accurate guess of the state to everyone (including the neighbor whose signal was worsened).

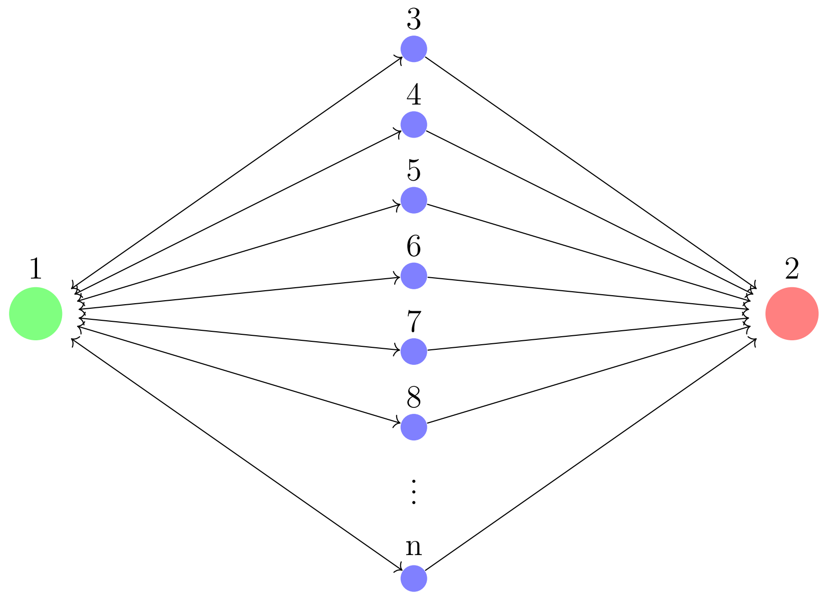
Our proof uses the network in Figure 1. The network has n agents, where n is large. Agent 1 observes agents $3, ..., n$, agents $3, ..., n$ observe agents 1 and 2, and there are no other connections. Agent 1 has no signal, i.e. $\sigma_1 = \infty$, while agents $4, ..., n$ have identical signals with variance σ_4^2.[10] We show in the appendix that for appropriate parameter values, making agent 3's private signal worse can be Pareto improving at equilibrium. In the language of the previous paragraph, agent 1 is the central agent, agent 3 is the neighbor and the first two agents' past actions are the confounding information.

Note that the example is not minimal, and the same comparative statics can be obtained on a star network if the center does not remember her previous actions. The example is chosen to emphasize that the inefficiencies driving this effect can come from network structure rather than finite memory.

A sufficiently large improvement in the precision of an agent's private information, however, must help that agent. For example, increasing an agent i's private signal precision

[10] A modification of the argument from Proposition 2 shows that this network has a unique equilibrium.

Figure 1: Value of Private Information



Network from the example in Section 5.2.

by 1 must help that agent because any estimate of θ_t based on observed past actions has precision at most 1 (recall the normalization that σ_ϵ^2, the variance of updates to the state, is 1).

5.3 Value of Connections

We next show numerically that adding links to the network can harm all agents. Like better private information, more communication can worsen confounds.

Adding links between agents can improve information sharing and therefore improve welfare, but can also increase correlation between signals and cause harm. In this part we provide an example for a network in which either adding a single link or completing the graph is Pareto-damaging in the sense of increasing all agents' equilibrium action variances.

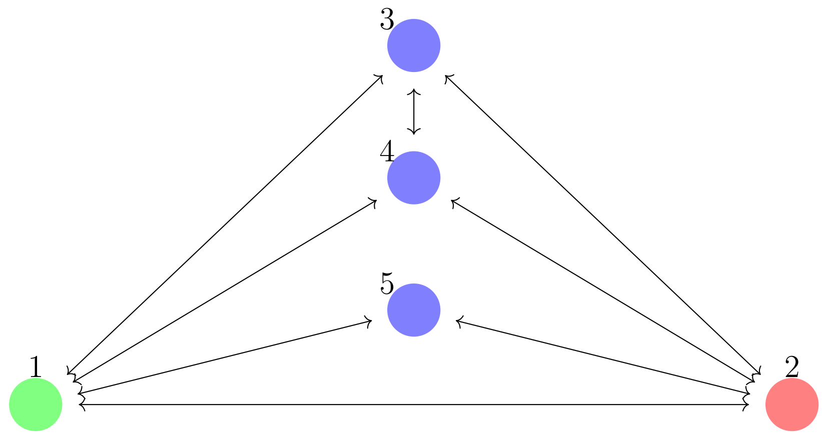
Denote by G_a the undirected graph in figure 2, in which agents 1 and 2 are connected to all other agents, agent 3 is connected to agent 4, and there are no other connections. Let $\rho = \frac{19}{20}$ and consider the vector of signal variances $\boldsymbol{\sigma}^2 = (20, 20, 1000, 1000, 1000)$. We will compare G_a to the network which is the same as G_a with the addition of an undirected link between 4 and 5, which we call G_b. We will also compare G_a to the complete network (without self-links), which we call G_c.

Table 1 shows that the all equilibrium variances are higher under G_a than under either

Table 1: Value of Connections

	$\widehat{\nu}_1^2$	$\widehat{\nu}_2^2$	$\widehat{\nu}_3^2$	$\widehat{\nu}_4^2$	$\widehat{\nu}_5^2$
G_a	3.1544	3.1544	3.4690	3.4690	3.4914
G_b	3.2140	3.2140	3.5267	3.5280	3.5280
G_c	3.2139	3.2139	3.5273	3.5273	3.5273

Figure 2: Value of Connections



Network G_a from the example in Section 5.3.

G_b or G_c. The first comparison says that adding a link hurts all agents, and the second says that the complete network is worse for all agents than another incomplete network.

The idea is that agents 1 and 2 see agent 5's previous action $a_{5,t-1}$ in each period, and for that action agent 5 puts most of their weight on 1 and 2's actions from two periods ago. Because these period $t-2$ actions are observed by several players, agents 1 and 2 benefit from anti-imitating $a_{5,t-1}$. But when we add a link between agents 4 and 5 these period $t-2$ actions are contaminated by 4's past actions, which makes the action $a_{5,t-1}$ less useful for agent 1.

6 Asymptotic Learning

In this section, we consider how well agents learn the state on large networks. Our benchmark is the action variance an agent would obtain given her private signal and perfect knowledge of the state in the previous period, reflecting the fact that other then then private signal all signal arrive with delay (hence perfect learning about today is impossible).

Our results in this section are as follows: On undirected graphs, an agent cannot achieve this benchmark asymptotically unless her signal about θ_{t-1} becomes perfectly accurate. On directed graphs, the same result holds if the network has enough cycles or agents are close to a peer with an accurate private signal.

6.1 Undirected Graphs

We first set up the asymptotic notation needed to examine our benchmarks. Fix ρ and let $\{G_n\}_{n=1}^{\infty}$ be a sequence of networks with n agents in each G_n. Let $\sigma_i^2(n)$ refer to the signal variance of agent i in G_n and $\widehat{\nu}_i^2(n)$ be the equilibrium variance of agent i in G_n.[11] For each i, we can discuss the asymptotic behavior of $\widehat{\nu}_i^2(n)$.

Because the state of the world is changing in each period, the benchmark cannot be asymptotically perfect learning: we cannot expect all error to disappear except in the trivial case where the variance of agents' private signals converges to zero. The following benchmark is the best we can hope for, and we show in an example following the definition that it can, indeed, be attained.

Definition 2. Agent i *learns asymptotically* if the limit of her equilibrium variance exists and satisfies

$$\lim_{n\to\infty} \widehat{\nu}_i^2(n) = \lim_{n\to\infty} (\sigma_i^{-2}(n) + 1)^{-1}.$$

This says that asymptotically, agent i does as well at equilibrium as if she knew her private signal and yesterday's state. Note she can never infer yesterday's state perfectly from observed actions, and so we must have $\widehat{\nu}_i^2 > (\sigma_i^{-2} + 1)^{-1}$ on any fixed network.

The following example shows that asymptotic learning is possible. In the next subsection we give a more interesting example in the case of directed networks.

Example 3. Suppose each G_n for $n \geq 2$ has a connected component with two agents, 1 and 2, with $\sigma_1^2(n) = 1$ and $\sigma_2^2(n) = 1/n$. Then agent 2's weight on her own signal converges to 1 as $n \to \infty$. So $\widehat{\nu}_i^2$ converges to $(\sigma_1^{-2}(n) + 1)^{-1} = \frac{1}{2}$. Agent 1 learns asymptotically.

On the other hand, if signal precisions are bounded, then no agent learns asymptotically:

Proposition 5. *Suppose there exists a constant $\underline{\sigma}^2 > 0$ such that $\sigma_i(n)^2 > \underline{\sigma}^2$ for all i and n. Then on any sequence of undirected graphs, no agent learns asymptotically.*

[11] If there are multiple equilibria for some values of n, a selection of an equilibrium for each n is part of the description of the asymptotic setting. The results below hold for any selection of equilibria.

The basic idea is that yesterday's actions all put weight on actions from period $t-2$, and so yesterday's actions are correlated through the state θ_{t-2} from two periods ago. Even with a very large number of observations, this confound—which agents cannot remove, because they do not know θ_{t-2}—prevents agents from learning yesterday's state precisely.

To make the argument more precise, assume toward a contradiction that agent i learns asymptotically. Because of the confounding discussed in the last paragraph, she would have to observe many neighbors who place almost all of their weight on their private signals. Because the network is undirected, though, these neighbors themselves see i. Since i's actions in this hypothetical reflect the state very accurately, the neighbors would do better by placing substantial weight on agent i and not just their private signals. So we cannot have such an agent i.

In summary, bidirectional observation presents a fundamental obstruction to attaining the best possible benchmark of aggregation. This is related to a basic observation about learning from multivariate Gaussian signals about a parameter: if the signals (here, social observations), conditional on the state of interest (θ_t) are all correlated and the correlation is bounded below, away from zero, (here this occurs because all involve some indirect weight on θ_{t-2}) then the amount one can learn from these signals is bounded, even if there are infinitely many of them. A related observation plays a role in Harel et al. (2017)—we discuss this more in Section 7.

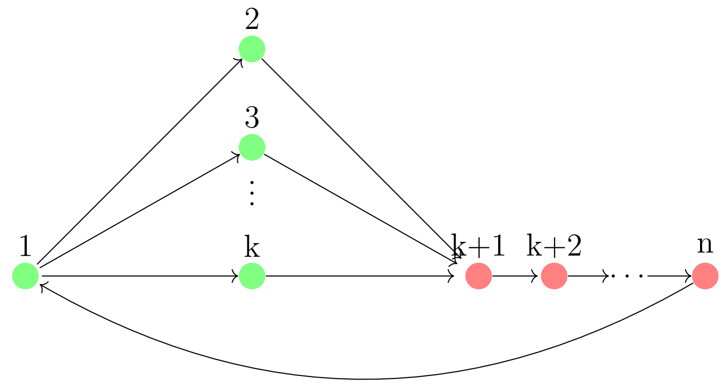
6.2 Directed Graphs

With directed networks, an agent can learn asymptotically even when all signal precisions are bounded. In our first example, a central agent observes a large number of peripheral agents while each peripheral agent's only information is her private signal. We say a directed edge (i, j) is *unidirectional* in a network G if $(i, j) \in E$ but $(j, i) \notin E$, meaning that i observes j.

Example 4. Consider a sequence of star networks G_n with n agents and unidirectional links from the center to the peripheral agents. Let $\sigma_i^2(n) = 1$ for all i and all n. Then agent 1 observes her own private signal and $n-1$ neighbors' actions that are independent (because they are based only on those agents' own private signals) with mean θ_{t-1}, so agent 1 learns asymptotically.

The next example shows that the assumption of strong connectedness does not disrupt the possibility of asymptotic learning.

Figure 3: Example of Asymptotic Learning



Network G_n from Example 5.

Example 5. Consider a sequence of networks G_n constructed as follows (see Figure 5). Let $k = \lfloor n/2 \rfloor$. All directed edges are unidirectional. In each network G_n there are $2n$ agents, $\sigma_i^2(n) = 1$ for $i \in \{1, \ldots, k\}$ while $\sigma_i^2(n) = n$ for $i \in \{k+1, \ldots, n\}$. Agent 1 observes agents $2, ..., k$. Each agent $k+1, ..., n$ observes agent $k+1$. Each agent $k+1, \ldots, n-1$ observes the next agent, and agent n observes agent 1. As n grows large, the weight agents $2, \ldots, k$ place on their private signal approaches 1. So agent 1 learns asymptotically.

Agent n can learn yesterday's state precisely even on a strongly connected network if later agents are further and further from any useful information.

Proposition 6. *Suppose there exists a constant $\underline{\sigma}^2 > 0$ such that $\sigma_i(n)^2 > \underline{\sigma}^2$ for all i and n. Suppose also that there exists an integer $M > 0$ and a constant $\overline{\sigma}^2 > 0$ such that for every edge from agent i to agent j, there exists a non-trivial path of length at most M from agent j to (i) agent i or (ii) an agent with signal variance $\sigma_k^2 < \overline{\sigma}^2$. Then no agent learns asymptotically.*

Each agent mustobserve signals from a reasonably close peer with accurate information, and the two conditions correspond to that peer's information coming from observed actions and their private signal. The second condition is implied by the (d, L)-egalitarian condition of Mossel, Sly, and Tamuz (2015) for any value of d (possibly infinite). The second condition is also implied if the network is strongly connected and all signal variances are bounded above by $\overline{\sigma}^2$.

In the undirected case, we used the fact that if agent i learned asymptotically then each of i's neighbors would see at least one useful past action and so would weight past actions substantially. Condition (i) or (ii) now replaces this argument, and the proof is similar otherwise.

7 Discussion and Related Literature

7.1 Other Motivating Cases

Our model is not limited to the example in the introduction of learning about crop prices. We now discuss several other practical situations where the same sort of learning model applies. First, consider financial analysts trying to predict the prices of stocks or commodities such as gold or oil at a future date in order to trade on them. The analysts have private signals in the form of their own research and also observe the predictions of other analysts, either by direct communication or via available publications. Maggio, Franzoni, Kermani, and Sommavilla (2017) document that substantial social learning occurs in networks of analysts, probably mediated by dealers, and that the relevant network has many links but is very far from complete.

Another example comes from the case of intelligence agencies around the world studying some dynamic quantity of mutual interest—say, another country's military capability in a particular area. Each agency has its own private signal (e.g., through its intelligence services' sources) while diplomatic and military links between allied countries facilitate partial information-sharing.

We believe our model fits these settings better than a standard social learning approach. Because the relevant state variables are continuous and evolve over time, assuming a fixed or binary state misses much of what there is to learn about. Indeed, prices and other macroeconomic variables are often modeled as AR(1) processes. For other applications the AR(1) model is not a perfect fit, but still a useful alternative to a once-and-for-all realization of uncertainty at a commonly known time, which is the standard approach in social learning models. Moreover, unlike models of sequential learning, our model allows for mutual and simultaneous learning by all agents in every period. These features are very relevant for estimation and prediction in empirical applications.

7.2 Naive Agents

While our main model assumes that agents are Bayesian, we can also accomodate misspecified beliefs. A case of particular interest is when all agents treat all observed actions as independent conditional on today's state. That is, when the covariance matrix is in fact $\boldsymbol{V}_t$ agents best respond to a covariance matrix with diagonal entries equal to $\boldsymbol{V}_{ii,t}$ and off-diagonal entries equal to zero. This model is the analogue of best-response trailing naive inference (Eyster and Rabin (2010)) in our dynamic environment, and is used by Alatas, Banerjee, Chandrasekhar, Hanna, and Olken (2016) to estimate learning about changing state variables in Indonesian villages.

Most of our results, including equilibrium existence (Proposition 1), continue to hold with naive agents. It can be shown that the equilibrium weights are Pareto-dominated as in Theorem 1, though the proof is different: it takes advantage of the fact that agents fail to optimize properly instead of using the envelope theorem.[12]

7.3 Shock Propagation

We now discuss how signal error propagates through the network over time. The effect on the population of error in an agent's private signal be characterized by a special case of Bonacich centrality.

Fix a period, without loss of generality $t = 0$; we will study how shocks $e_i := \eta_{i,0} = s_{i,0} - \theta_0$ received by agents at this time propagate. Let $\mathrm{diag}(\widehat{w}^s)$ denote the diagonal matrix with entries $\widehat{w}_i^s$, agents' weights on their private signals. The influence of the errors e_i will be to change agents' actions at $t = 0$ by $\mathrm{diag}(\widehat{w}^s)\boldsymbol{e}$, at $t = 1$ by $\rho\widehat{W}\mathrm{diag}(\widehat{w}^s)\boldsymbol{e}$, etc. In general the errors $\boldsymbol{e}$ change the vector of period $k-1$ actions by

$$(\rho\widehat{W})^k\mathrm{diag}(\widehat{w}^s)\boldsymbol{e}.$$

Note that, because some of the values in $\widehat{W}$ can be negative, errors can bias agents' predictions up or down.

We consider one very simple measure of the total error introduced by a shock: Imagine all signals outside period t are exactly correct (i.e. $s_{i,t} = \theta_t$), and sum the total error over all periods for each agent. For arbitrary $\widehat{W}$, the vector of total errors is

$$\sum_{k=0}^{\infty} \rho^k |\widehat{W}^k| \mathrm{diag}(\widehat{w}^s)|\boldsymbol{e}|,$$

where $|\cdot|$ is the entry-by-entry absolute value.

[12]Because naive agents will never anti-imitate, we no longer need to impose the requirement of positive weights as an assumption.

If there is no anti-imitation at equilibrium, then all entries of $\widehat{W}$ are positive the formula can be written as

$$(I - \rho\widehat{W})^{-1}\text{diag}(\widehat{w}^s)\boldsymbol{e}.$$

This can be interpreted as the Bonacich centrality with respect to the matrix $\rho\widehat{W}$ and vector $\text{diag}(\widehat{w}^s)\boldsymbol{e}$. The influence of an agent's private signal is equal to that agent's Bonacich centrality in the network that has edge weights equal to the equilibrium updating weights.

7.4 Related Literature

We now put our contribution in the context of the extensive literature on social learning and learning in networks.

7.4.1 Classic Social Learning Models

A classical approach to social learning involves infinitely many agents guessing a fixed, binary state, acting in sequence with access to (some) predecessors' actions (Banerjee (1992); Bikhchandani, Hirshleifer, and Welch (1992)). The original models were worked out with observation of *all* predecessors, but recent papers have developed analyses where a directed network structure describes the *subset* of predecessors seen by each agent Acemoglu et al. (2011); Eyster and Rabin (2014); Lobel and Sadler (2015a,b).[13]

The sequential structure is essential to tractability in these models, which rely on the fact that future agents' behavior does not affect current agents' learning choices. One main contribution of our paper is to introduce repeated and bidirectional interaction between nodes and reexamine classical social learning questions in light of this. Substantively new insights come out: for example, with bidirectional or "egalitarian" observation structures, information aggregation benchmarks cannot be achieved on any network, whereas with directed links, these benchmarks can be attained at least by some agents. Beyond particular substantive insights, models that hinge on a sequential structure are difficult to take to data, while Section 4 shows that our model is amenable to structural estimation.

A main implication of the sequential social learning models is that cascades of incorrect decisions are typical. These inefficiencies are related to the ones identified by our Theorem 1: both occur because agents place less weight on their private signals than is socially

[13] Golub and Sadler (2016) survey this literature.

optimal. Our inefficiency manifests in steady state. Because the learning weights are described by fixed points, we can offer new insights on the structural features of networks that affect the inefficiency, as illustrated by Section 5.2.

Another robust aspect of rational learning in sequential networks is anti-imitation. Eyster and Rabin (2014) give general conditions for agents to anti-imitate in the sequential model. Anti-imitation can occur at equilibrium in our model, as in the examples in Section 5.2 and 5.3, but we also find agents may not anti-imitate in the rational model, and offer a structural approach to determining whether the lack of anti-imitation is evidence of suboptimal weights (Section 5.2). Intuitively, anti-imitation is especially natural when an agent observes a large number of pieces of information, some of which provide much of the information common to many of the others. (These signals then should be "subtracted off" to avoid overweighting them.) The ordering of agents in sequential models makes such asymmetries difficult to avoid: early agents are usually confounds. In our model, whether confounds are present is more flexible, and this sheds new light on the network structures behind anti-imitation.

7.4.2 DeGroot Learning

Stationary play at equilibrium in our model resembles the DeGroot heuristic in that agents place constant, linear weights on their neighbors. Other studies on the DeGroot model analyze the same heuristic when information arrives only once. DeMarzo, Vayanos, and Zweibel (2003) offer a theory as to why that heuristic may be used: agents in such an environment neglect the changes in the distribution of their observations over time. We give an alternative, Bayesian microfoundation for playing linear combinations of observed actions with time-invariant weights, coming from the fact that the environment is stationary and so, in fact, the distributions of agents' behavior does not change. Indeed, it may be interesting to consider the possibility that agents behaving according to the DeGroot heuristic even when it is not appropriate may have to do with their experiences in stationary environments where it is closer to optimal.

Golub and Jackson (2010) study learning under DeGroot and give conditions for agents to learn well on large networks. These results contrast with our results on asymptotic learning, which show that a changing state creates inefficiencies that persist on large networks, as long as there is some bilateral observation.

Jadbabaie, Molavi, Sandroni, and Tahbaz-Salehi (2012) look at DeGroot when agents receive new signals in each period. Because the state remains fixed in their model, stationary strategies are still not Bayesian best responses—the deviation from rationality has recently been axiomatized in Molavi, Tahbaz-Salehi, and Jadbabaie (2016). However, eventually receive a large number of observations about a fixed state, and the authors show even naive agents eventually learn the true state under mild conditions. Making the state dynamic removes the possibility of a simple heuristic that achieves perfect learning, and allows us to study how the network mediates inefficiencies.

7.4.3 Recent Models with Evolving States

Perhaps most related to our work, several recent computer science and engineering papers study similar environments involving social learning about a changing state. Frongillo, Schoenebeck, and Tamuz (2011) study best-response dynamics, which are the same as our dynamics with optimizing agents, in a closely related model. Their results on best-response dynamics focus on the complete network case where all agents observe all of yesterday's actions. Thus, several of our results generalize computations in Frongillo, Schoenebeck, and Tamuz (2011) from the complete network to arbitrary networks. The authors also characterize the steady-state distribution of behavior with arbitrary (non-equilibrium) fixed weights on any network.

The stochastic process and information structure in Shahrampour, Rakhlin, and Jadbabaie (2013) are also the same as ours, though their analysis does not include optimizing agents. The authors consider a class of fixed weights and study heuristics, computing or bounding various measures of welfare. We compare welfare under such fixed exogenous weights with the welfare obtained by optimizing agents at equilibrium. Because our model endogenously selects weights, we can consider counterfactual scenarios as the network changes.

From an engineering perspective, to our knowledge, this paper is the first to study the Kalman filtering problem with multiple agents non-cooperatively selecting endogenous weights. Thus, there is a tight connection to a large engineering literature on the Kalman filter.

In economics, the model in Alatas, Banerjee, Chandrasekhar, Hanna, and Olken (2016) most closely resembles ours. Their agents are not fully Bayesian, ignoring the correlation between social observations. The model is estimated using data on social learning in Indonesian villages, where the state variables are the wealth of villagers. Our Bayesian version may allow empirical researchers doing similar analyses to analyze the sophistication of a learning process.

Moscarini, Ottaviani, and Smith (1998) and van Oosten (2016) study network learning models where the binary state evolves as a two-state Markov chain. Their results focus largely on the frequency and dynamics of cascades, while we take a richer underlying stochastic process and study the resulting stationary equilibrium.

References

Acemoglu, D., M. Dahleh, I. Lobel, and A. Ozdaglar (2011): "Bayesian Learning in Social Networks," *The Review of Economic Studies*, 78, 1201–1236.

Alatas, V., A. Banerjee, A. G. Chandrasekhar, R. Hanna, and B. A. Olken (2016): "Network structure and the aggregation of information: Theory and evidence from Indonesia," *The American Economic Review*, 106, 1663–1704.

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

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, 992–1026.

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

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

Eyster, E. and M. Rabin (2010): "Naive herding in rich-information settings," *American Economic Journal: Microeconomics*, 2, 221–243.

——— (2014): "Extensive Imitation is Irrational and Harmful," *Quarterly Journal of Economics*, 129, 1861–1898.

Frongillo, R., G. Schoenebeck, and O. Tamuz (2011): "Social Learning in a Changing World," *Internet and Network Economics*, 146–157.

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

Geanakoplos, J. D. and H. M. Polemarchakis (1982): "We can't disagree forever," *Journal of Economic Theory*, 28, 192–200.

Golub, B. and M. Jackson (2010): "Naive learning in social networks and the wisdom of crowds," *American Economic Journal: Microeconomics*, 2, 112–149.

Golub, B. and E. Sadler (2016): "Learning in Social Networks," in *The Oxford Handbook of the Economics of Networks*, ed. by Y. Bramoullé, A. Galeotti, B. Rogers, and B. Rogers, Oxford University Press, chap. 19, 504–542.

Harel, M., E. Mossel, P. Strack, and O. Tamuz (2017): "The Speed of Social Learning," *arXiv preprint arXiv:1412.7172*.

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

Jensen, R. (2007): "The digital provide: Information (technology), market performance, and welfare in the South Indian fisheries sector," *The quarterly journal of economics*, 122, 879–924.

Kay, S. M. (1993): *Fundamentals of statistical signal processing*, Prentice Hall PTR.

Lobel, I. and E. Sadler (2015a): "Information diffusion in networks through social learning," *Theoretical Economics*, 10, 807–851.

——— (2015b): "Preferences, homophily, and social learning," *Operations Research*, 64, 564–584.

Maggio, M. D., F. Franzoni, A. Kermani, and C. Sommavilla (2017): "The relevance of broker networks for information diffusion in the stock market," *Swiss Finance Institute Research Paper No. 16-63; Harvard Business School Finance Working Paper. Available at SSRN: https://ssrn.com/abstract=2860118*.

Manea, M. (2011): "Bargaining in stationary networks," *The American Economic Review*, 101, 2042–2080.

Manresa, E. (2013): "Estimating the structure of social interactions using panel data," *Unpublished Manuscript. CEMFI, Madrid*.

Molavi, P., A. Tahbaz-Salehi, and A. Jadbabaie (2016): "Foundations of non-bayesian social learning," .

Moscarini, G., M. Ottaviani, and L. Smith (1998): "Social Learning in a Changing World," *Economic Theory*, 11, 657–665.

Mossel, E., A. Sly, and O. Tamuz (2015): "Strategic learning and the topology of social networks," *Econometrica*, 83, 1755–1794.

Rahimian, M. A. and A. Jadbabaie (2017): "Bayesian learning without recall," *IEEE Transactions on Signal and Information Processing over Networks*.

Shahrampour, S., S. Rakhlin, and A. Jadbabaie (2013): "Online learning of dynamic parameters in social networks," *Advances in Neural Information Processing Systems*.

Srinivasan, J. and J. Burrell (2013): "Revisiting the fishers of Kerala, India," in *Proceedings of the Sixth International Conference on Information and Communication Technologies and Development: Full Papers-Volume 1*, ACM, 56–66.

Talamàs, E. (2017): "Prices and Efficiency in Networked Markets," *Available at SSRN: https://ssrn.com/abstract=2882377*.

van Oosten, R. (2016): "Learning from Neighbors in a Changing World," *Master's Thesis*.

Weizsäcker, G. (2010): "Do we follow others when we should? A simple test of rational expectations," *The American Economic Review*, 100, 2340–2360.

A Proofs

Proof of Proposition 1. Recall from Section 3.3 the map Φ, which gives the next-period covariance matrix $\Phi(\boldsymbol{V}_t)$ for any $\boldsymbol{V}_t$. The expression given there for this map ensures that its entries are continuous functions of the entries of $\boldsymbol{V}_t$. Our strategy is to show that this function maps a compact set to itself, which, by Brouwer's fixed-point theorem, ensures that Φ has a fixed point $\widehat{\boldsymbol{V}}$.

We now define the compact set. Because memory is arbitrary, entries of $\boldsymbol{V}_t$ are covariances between pairs of observations from any periods available in memory. Let k, l be two indices of such observations, corresponding to actions taken by agents i and j respectively, and let $\bar{\sigma}_i = \max\{\sigma_i^2, \rho^{m-1}\sigma_i^2 + \frac{1-\rho^{m-1}}{1-\rho}\}$. Now let $\mathcal{K} \subset \mathcal{V}$ be the subset of symmetric positive semi-definite matrices $\boldsymbol{V}_t$ such that, for any such k, l,

$$\boldsymbol{V}_{kk,t} \in \left[\min\left\{\frac{1}{1+\sigma_i^{-2}}, \frac{\rho^{m-1}}{1+\sigma_i^{-2}} + \frac{1-\rho^{m-1}}{1-\rho}\right\}, \max\left\{\sigma_i^2, \rho^{m-1}\sigma_i^2 + \frac{1-\rho^{m-1}}{1-\rho}\right\}\right]$$
$$\boldsymbol{V}_{kl,t} \in [-\bar{\sigma}_i\bar{\sigma}_j, \bar{\sigma}_i\bar{\sigma}_j].$$

This set is closed and convex, and we claim that $\Phi(\mathcal{K}) \subset \mathcal{K}$.

To show this claim, we will first bound the variance of any agent's action (at any period in memory). Note that a Bayesian agent will not choose an action with a larger variance than her signal, which has variance σ_i^2.

For a lower bound on the variance of the agent's action, note that if she knew the previous period's state and her own signal, then the variance of her action would be $\frac{1}{1+\sigma_i^{-2}}$. Thus an agent observing only noisy estimates of θ_t and her own signal can do no better. By the same reasoning applied to i's self from m periods ago, the variance of the estimate of θ_{t-1} based on i's action from m periods ago cannot exceed $\rho^{m-1}\sigma_i^2 + \frac{1-\rho^{m-1}}{1-\rho}$ or be smaller than $\frac{\rho^{m-1}}{1+\sigma_i^{-2}} + \frac{1-\rho^{m-1}}{1-\rho}$. This establishes bounds on $\boldsymbol{V}_{kk,t}$ for observations k coming from either the most recent or the oldest available period. The corresponding bounds from the periods between $t-m+1$ and t are always weaker than at least one of the two bounds we have described, so we need only take minima and maxima over two terms.

This established the claimed bound on the variances. The bounds on covariances follow from Cauchy-Schwartz. □

Proof of Proposition 2. We will show that when the network satisfies the condition in the proposition statement, Φ induces a contraction on a suitable space. For each agent, we can consider the variance of the best estimator for yesterday's state based on observed actions. These variances are tractable because they satisfy the envelope theorem. Moreover, the space of these variances is a sufficient statistic for determining all agent strategies and action variances.

Let $r_{i,t}$ be i's *social signal*—the best estimator of θ_t based on the period $t-1$ actions of agents in N_i—and let $\kappa_{i,t}^2$ be the variance of $r_{i,t} - \theta_t$.

We claim that Φ induces a map $\widetilde{\Phi}$ on the space of variances $\kappa_{i,t}^2$, which we denote $\widetilde{\mathcal{V}}$. We must check the period t variances $(\kappa_{i,t}^2)_i$ uniquely determine all period $t+1$ variances $(\kappa_{i,t+1}^2)_i$: The variance $\boldsymbol{V}_{ii,t}$ of agent i's action, as well as the covariances $\boldsymbol{V}_{ii',t}$ of all pairs of agents i, i' with $N_i = N_{i'}$, are determined by $\kappa_{i,t}^2$. Moreover, by the condition on our network, these variances and covariances determine all agents' strategies in period $t+1$, and this is enough to pin down all period $t+1$ variances $\kappa_{i,t+1}^2$.

The proof proceeds by showing $\widetilde{\Phi}$ is a contraction on $\widetilde{\mathcal{V}}$ in the sup norm.

For each agent j, we have $N_i = N_{i'}$ for all $i, i' \in N_j$. So the period t actions of an agent i' in N_j are

$$a_{i',t} = \frac{(\rho^2\kappa_{i,t}^2 + 1)^{-1}}{\sigma_{i'}^{-2} + (\rho^2\kappa_{i,t}^2 + 1)^{-1}} \cdot r_{i,t} + \frac{\sigma_{i'}^{-2}}{\sigma_{i'}^{-2} + (\rho^2\kappa_{i,t}^2 + 1)^{-1}} \cdot s_{i',t} \tag{5}$$

where $s_{i',t}$ is agent (i')'s signal in period t and $r_{i,t}$ the social signal of i (the same one that i' has). It follows from this formula that each action observed by j is a linear combination of a private signal and a *common* estimator $r_{i,t}$, with positive coefficients which sum to one. For simplicity we write

$$a_{i',t} = b_0 \cdot r_{i,t} + b_{i'} \cdot s_{i',t} \tag{6}$$

(where b_0 and $b_{i'}$ depend on i' and t, but we omit these subscripts). We will use the facts $0 < b_0 < 1$ and $0 < b_{i'} < 1$.

We are interested in how $\kappa_{j,t}^2 = \text{Var}(r_{j,t} - \theta_t)$ depends on $\kappa_{i,t-1}^2 = \text{Var}(r_{i,t-1} - \theta_{t-1})$. The estimator $r_{j,t}$ is a linear combination of observed actions $a_{i',t}$, and therefore can be expanded as a linear combination of signals $s_{i',t}$ and the estimator $r_{i,t-1}$. We can write

$$r_{j,t} = c_0 \cdot (\rho r_{i,t-1}) + \sum_{i'} c_{i'} s_{i',t} \tag{7}$$

and therefore (taking variances of both sides)

$$\begin{aligned} \kappa_{j,t}^2 = \text{Var}(r_{j,t} - \theta_t) &= c_0 \text{Var}(\rho r_{i,t-1} - \theta_t) + \sum_{i'} c_{i'} \sigma_{i'}^2 \\ &= c_0(\kappa_{i,t-1}^2 + 1) + \sum_{i'} c_{i'} \sigma_{i'}^2 \end{aligned}$$

The desired result, that $\widetilde{\Phi}$ is a contraction, will follow if we can show that the derivative

$\frac{d\kappa_{j,t}^2}{d\kappa_{i,t-1}^2} \in [0, \delta]$ for some $\delta < 1$. By the envelope theorem, when calculating this derivative, we can assume that the weights placed on actions $a_{i',t-1}$ by the estimator $r_{j,t}$ do not change as we vary $\kappa_{i,t-1}^2$, and therefore c_0 and the $c_{i'}$ above do not change. So it is enough to show the coefficient c_0 on $\kappa_{i,t-1}^2$ is in $[0, \delta]$.

The intuition for the lower bound is that *anti-imitation* (agents placing negative weights on observed actions) only occurs if observed actions put too much weight on public information. But if $c_0 < 0$, then the weight on public information is actually negative so there is no reason to anti-imitate. This is formalized in the following lemma. □

Lemma 1. *Agent j's social signal places non-negative weight on agent i's social signal from the previous period, i.e., $c_0 \geq 0$.*

Proof. To check this formally, suppose that c_0 is negative. Then the social signal $r_{j,t}$ puts negative weight on some observed action—say the action $a_{k,t-1}$ of agent k. We want to check that the covariance of $r_{j,t} - \theta_t$ and $a_{k,t-1} - \theta_t$ is negative. Using (6) and (7), we compute that

$$\begin{aligned}\mathrm{Cov}(r_{j,t} - \theta_t, a_{k,t-1} - \theta_t) &= \mathrm{Cov}\left(c_0(\rho r_{i,t-1} - \theta_t) + \sum_{i' \in N_j} c_{i'}(s_{i',t} - \theta_t)), b_0(\rho r_{i,t-1} - \theta_t) + b_k(s_{k,t-1} - \theta_t)\right) \\ &= c_0 b_0 \mathrm{Var}(\rho r_{i,t-1} - \theta_t) + c_k b_k \mathrm{Var}(s_{k,t-1} - \theta_t)\end{aligned}$$

because all distinct summands above are mutually independent. We have $b_0, b_k > 0$, while $c_0 < 0$ by assumption and $c_k < 0$ because the estimator $r_{j,t}$ puts negative weight on $a_{k,t-1}$. So the expression above is negative. Therefore, it follows from the usual Gaussian Bayesian updating formula that the best estimator of θ_t given $r_{j,t}$ and $a_{k,t-1}$ puts positive weight on $a_{k,t-1}$. However, this is a contradiction: the best estimator of θ_t given $r_{j,t}$ and $a_{k,t-1}$ is simply $r_{j,t}$, because $r_{j,t}$ was defined as the best estimator of θ_t given observations that included $a_{k,t-1}$. □

Proof. Now, for the upper bound $c_0 \leq \delta$, the idea is that $r_{j,t}$ puts more weight on agents with better signals while these agents put little weight on public information, which keeps the overall weight on public information from growing too large.

Note that $r_{j,t}$ is a linear combination of actions $\rho a_{i',t-1}$ for $i \in N_j$, with coefficients summing to 1. The only way the coefficient on $\rho r_{i,t-1}$ in $r_{j,t}$ could be at least 1 would be if some of these coefficients on $\rho a_{i',t-1}$ were negative and the estimator $r_{j,t}$ placed greater weight on actions $a_{i',t-1}$ which placed more weight on $r_{j,t}$.

Applying the formula (5) for $a_{i',t-1}$, we see that the coefficient b_0 on $\rho r_{i,t-1}$ is less than 1 and increasing in $\sigma_{i'}$. On the other hand, it is clear that the weight on $a_{i',t-1}$ in the social signal $r_{j,t}$ is decreasing in $\sigma_{i'}$: more weight should be put on more precise individuals. So in fact the estimator $r_{j,t}$ places less weight on actions $a_{i',t-1}$ which placed more weight on $r_{i,t}$.

Moreover, the coefficients placed on private signals are bounded below by a positive constant when we restrict to covariances in the image of $\widetilde{\Phi}$ (because all covariances are bounded as in the proof of Proposition 1). Therefore, each agent $i' \in N_j$ places weight at most δ on the estimator $\rho r_{i,t-1}$ for some $\delta < 1$. Agent j's social signal $r_{j,t}$ is a sum of these agents' actions with coefficients summing to 1 and satisfying the monotonicity property above. We conclude that the coefficient on $\rho r_{i,t-1}$ in the expression for $r_{j,t}$ is at most δ. We conclude that the coefficient on $\rho r_{i,t-1}$ in $r_{j,t}$ is bounded above by some $\delta < 1$. □

Proof of Proposition 3. The proof is by induction. Reordering, we can assume there are agents $1, 2, ..., n$ such that i does not observe j whenever $i < j$. We claim that for each i and j (not necessarily distinct), agent i and j's action weights and the covariances $\boldsymbol{V}_{ij,t}$ are constant for $t \geq mi, mj$. The case $i = j = 1$ is clear, because the first agent has no neighbors and so puts weight one on her private signal for all $t \geq 1$.

Suppose the claim holds for $i, j \leq k$ and consider the claim for $i \leq j = k+1$. In all periods $t \geq mj + 1$, all agents observed by j have constant action weights and covariances by the inductive hypothesis. It follows from our formula for best-responses that agent j's action weights are constant as well. For any $i \leq j$, both i and j use constant weights and their observations have the same covariances, so $\boldsymbol{V}_{ij,t}$ is constant for $t \geq mj + 1$.

By induction we conclude that all covariances are constant after mn periods. As a consequence, there is a unique equilibrium covariance matrix $\widehat{\boldsymbol{V}}$ and $\Phi^{mn}(\boldsymbol{V}) = \widehat{\boldsymbol{V}}$ for any $\boldsymbol{V}$. □

Proof of Theorem 1. We provide the proof in the case $m = 1$ to simplify notation. The argument carries through with arbitrary finite memory.

Consider an agent l who places positive weight on an agent k and positive weight on at least one other agent. Define weights $\overline{W}$ by $\overline{W}_{ij} = \widehat{W}_{ij}$ and $\overline{w}_i^s = \widehat{w}_i^s$ for all $i \neq k$, $\overline{W}_{kj} = (1-\epsilon)\widehat{W}_{kj}$ for all $j \leq n$, and $\overline{w}_k^s = (1-\epsilon)\widehat{w}_k^s + \epsilon$. In words, agent k places weight $(1-\epsilon)$ on her equilibrium strategy and extra weight ϵ on her private signal. All other players use the same weights at equilibrium.

Suppose we are at equilibrium until time t, but in period t and all subsequent periods agents instead use weights $\overline{W}$. These weights give an alternate updating function $\overline{\Phi}$ on the space of covariance matrices. Because the weights $\overline{W}$ are positive and fixed, all coordinates of $\overline{\Phi}$ are increasing, linear functions of all previous period variances and covariances. Explicitly, the diagonal terms are

$$[\overline{\Phi}(\boldsymbol{V}_t)]_{ii} = (\overline{w}_i^s)^2\sigma_i^2 + \sum_{j,j'\leq n} \overline{W}_{ij}\overline{W}_{ij'}\boldsymbol{V}_{jj',t}$$

and the off-diagonal terms are

$$[\overline{\Phi}(\boldsymbol{V}_t)]_{ii'} = \sum_{j,j'\leq n} \overline{W}_{ij}\overline{W}_{i'j'}\boldsymbol{V}_{jj,t'}.$$

So it is sufficient to show the variances $\overline{\Phi}^h(\widehat{\boldsymbol{V}})$ after applying $\overline{\Phi}$ for h periods Pareto dominate the variances in $\widehat{\boldsymbol{V}}$ for some h.

In period t, the change in weights decreases the covariance $\boldsymbol{V}_{jk,t}$ of k and some other agent j, who l also observes, by $f(\epsilon)$ of order $\Theta(\epsilon)$. By the envelope theorem, the change in weights only increases the variance ν_k^2 by $O(\epsilon^2)$. Taking ϵ sufficiently small, we can ignore $O(\epsilon^2)$ terms.

There exists a constant $\delta > 0$ such that all equilibrium weights on observed neighbors are at least δ. Then each coordinate $[\Phi(\boldsymbol{V})]_{ii}$ is linear with coefficient at least δ^2 on each variance or covariance of agents observed by i.

Because agent l observes k and another agent, agent l's variance will decrease below its equilibrium level by at least $\delta^2 f(\epsilon)$ in period $t+1$. Because $\overline{\Phi}$ is increasing in all entries and we are only decreasing covariances, agent l's variance will also decrease below its equilibrium level by at least $\delta^2 f(\epsilon)$ in all periods $t' > t+1$.

Because the network is strongly connected and finite, the network has a diameter d. After $d+1$ periods, the variances of all agents have decreased by at least $\delta^{2d+2} f(\epsilon)$ from their equilibrium levels. This gives a Pareto improvement. $\square$

Proof of Proposition 4. The proof uses the network in Figure 1. Agent 1 has no signal, i.e. $\sigma_1 = \infty$, while agents $4, ..., n$ have identical signals with variance σ_4^2. First, suppose that $\sigma_3 = \sigma_4$, so that all agents except the first two have the same signal variance. Then agent 1 puts equal weight $\frac{1}{n-2}$ on each of her observations, agent 1 contibutes to the other agents by aggregating everyone's signals, and 1's previous action is informative to agents 3,...,n.

Taking σ_4 sufficiently large and σ_2 sufficiently small, we may assume that agents $3, ..., n$

put weight at least $1-\xi$ on agent 2's previous action, for some arbitrary $1 > \xi > 0$. As a result, agent 1 indirectly places weight at least $1-\xi$ on agent 2's action from two periods ago, and so $\nu_1^2 > (1-\xi)^2(1+\rho^2(1+\rho^2\sigma_2^2)) > (1-\xi)^2(\rho^2+1)$.

Now suppose instead that $\sigma_3 = \infty$, we will show now that ν_i^2 strictly decreases for all i except for 2 for which it stays the same ($\nu_2^2 = \sigma_2^2$).

Agent 3 loses their signal and hence puts weight only on 1's and 2's past actions (with the same proportions as the other agents), which allowes agent 1 to calculate the private signals of agents $4, ..., n$. Taking n sufficiently large, we have $\nu_1^2 < 1+\xi$. For ξ small and ρ large, agent 1's variance has decreased. Because agents $4, ..., n$ observe agent 1's action, agent 2's action and their private signal, and moreover agent 1 and 2's actions are conditionally independent, their variances decrease when ν_1^2 does.

It remains to check that ν_3^2 decreases when the signal variance σ_3^2 increases. We now fix σ_2 and ρ so that for each σ_4 sufficiently large, there exists n large enough such that increasing σ_3 from σ_4 to ∞ decreases ν_1^2 by at least ξ. This can be done by the same argument as in the previous paragraphs. When σ_4 is sufficiently large, improving ν_1^2 by ξ and decreasing σ_3^2 from σ_4^2 to ∞ helps agent 3. Therefore changing σ_3 from σ_4 to ∞ is Pareto improving. □

Proof of Proposition 5. Suppose agent 1 learns asymptotically. We can assume that agent 1 has at least one neighbor in each G_n.

If agent 1 learns asymptotically, then $\widehat{\nu}_i^2(n) < 1$ for n sufficiently large. Fix any n so that $\nu_1^2(n) < 1$ (to simplify notation we drop references to n for the remainder of the proof).

Then, at equilibrium, any agent i connected to 1 has a best estimator $r_{i,t}$ of θ_t based on observed actions with variance $\kappa_{i,t}^2 = \mathrm{Var}(r_{i,t} - \theta_t)$ less than $1+\rho^2 \leq 2$. So agent i's action

$$a_{i,t} = \frac{\kappa_{i,t}^{-2} r_{i,t} + \sigma_i^{-2} s_{i,t}}{\kappa_{i,t}^{-2} + \sigma_i^{-2}},$$

puts weight at least $\frac{2}{2+\underline{\sigma}^2}$ on observed actions.

Therefore, in period t, agent 1's best estimator $s_{1,\theta_{t-1}}$ of θ_{t-1} (indirectly) puts weight at least $\frac{2\rho}{2+\underline{\sigma}^2}$ on actions from period $t-2$. Because

$$\mathrm{Var}(\rho \sum b_j a_{j,t-2} - \theta_{t-1}) = \mathrm{Var}(\rho \sum b_j a_{j,t-2} - \theta_{t-2}) + \mathrm{Var}(\theta_{t-2} - \theta_{t-1}) \geq 1$$

for any coefficients b_j summing to 1, the variance of $r_{i,t} - \theta_{t-1}$ is at least $\frac{4\rho^2}{(2+\underline{\sigma}^2)^2}$. But then agent 1's action variance is bounded away from $(\sigma_i^{-2}+1)$ and this bound holds for all large enough n, which contradicts our assumption that agent 1 learns asymptotically. $\square$

Proof of Proposition 6. Suppose i learns asymptotically. As in the undirected case, it is sufficient to show the weights all agents observed by i place on previous actions are bounded away from zero (independent of n). Because i learns asymptotically, $\nu_i^2(n) \leq 1$ for all n sufficiently large. So all agents j observed by agent i have a path of length at most M to an agent with action variance $\nu_k^2 < \max(\overline{\sigma}^2, 1)$. Because signal variances are bounded below by $\underline{\sigma}^2$, this implies the result. $\square$