Dispersed Information, Social Networks and Aggregate Behavior

Jakob Grazzini*

Domenico Massaro§

March 1, 2021

Abstract

This paper argues that, in the presence of dispersed information, individual-level idiosyncratic noise may propagate at the aggregate level when agents are connected through a social network. When information about a common fundamental is incomplete and heterogeneous across agents, it is beneficial to consider the actions of other agents because of the additional information conveyed by these actions. We refer to the act of using other agents' actions in the individual decision process as social learning. This paper shows that social learning aimed at reducing the error of individual actions with respect to the fundamental may increase the error of the aggregate action depending on the network topology. Moreover, if the network is very asymmetric, the error of the aggregate action does not decay as predicted by the law of large numbers.

JEL codes: E70, D83, D84, D85.
Keywords: Dispersed Information, Expectations, Social Networks.

Declarations of Interest: None.

Fundings: Jakob Grazzini acknowledges financial support from the European Union, Seventh Framework Programme FP7/2007-2013 Socio-economic Sciences and Humanities under Grant Agreement No. 612796 - MACFINROBODS. Domenico Massaro acknowledges financial support from the Ministry of Education, Universities and Research of Italy (MIUR), program *SIR* (grant n. RBSI144KWH). None of the above are responsible for errors in this paper.

*Assistant professor, Department of Economics and Management, University of Pavia, Via San Felice 5, 27100 Pavia, Italy; Complexity Lab in Economics. Email: `jakob.grazzini@unipv.it`.

§Assistant professor, Department of Economics and Finance, Catholic University of Milan, Via Necchi 5, 20123 Milan, Italy; Complexity Lab in Economics. Email: `domenico.massaro@unicatt.it`.

1 Introduction

Human beings are social animals. The characteristic of social animals is that they live in groups and interact with other members of the group to perform vital tasks, such as defense or feeding. Several studies in biology document forms of interaction in which the group allows individuals to take advantage of the information gathered by others. Pulliam (1973) for example studies the flocking behavior of finches and shows that an advantage of feeding in groups is to increase the probability to detect a predator ("many eyes effect"). If one finch spots a predator and decides to fly off in alarm, the other finches observe this action and follow without having actually seen the danger. Moreover, assuming that anti-predatory vigilance is costly, for example because it is time consuming and alternative to feeding, the group provides a simple and effective cost sharing mechanism (Fernandez et al., 2003). In fact, the propensity to live in groups and learn from others' behavior can be considered an evolutionary response that promotes survival in complex environments for both animals and human beings (Henrich, 2015).

There are countless social and economic situations in which human beings are influenced by what others around them are doing when deciding upon an action. The decisions of others can be relevant for individual decision making for a variety of reasons. First, in the presence of *payoff externalities*, the actions of others can directly influence the utility function. Typical examples are market environments with strategic complementarities. Second, the decisions of other agents may matter for individual decision making in the presence of *informational externalities*. In such environments individual decisions reflect relevant information, hence observing the actions of others allows to exploit such information.

We focus on the second scenario and consider the case in which informational externalities arise due to *dispersed information*, i.e., information which is incomplete and heterogeneous across agents. Following Bandura and McClelland (1977), we call the act of observing and learning from other agents' actions *social learning*.

The goal of this paper is to show that, although social learning increases the accuracy of individual actions when information is dispersed, at the aggregate level accuracy can decrease depending on the structure of the social network shaping the patterns of interaction and learning among individuals. To this end, we consider a setting where agents receive independent noisy signals about the true value of a common variable of interest and social learning occurs through an arbitrary social network. Although in our setting individual payoffs do not depend on the decisions of others, agents have an incentive to pay attention to other individuals because their decisions reflect valuable additional information. This simple setup, in which there are no strategic complementarities and the interaction among agents is purely informational, allows us to isolate the impact of different network topologies on the transmission of idiosyncratic noise from the individual to the aggregate level.[1]

The essence of the model is as follows. A set of agents must take an action, receive a private signal on a common payoff-relevant fundamental state variable, and can observe the actions of a subset of other agents. The subset of observable agents is defined by a social network. The "observational structure" defined by the network is exogenous and can correspond to e.g., geographical proximity or social relationships. The setup of the model is very general and can be applied to a variety of economic environments. For the sake of concreteness, we frame the analysis in a setup where a group of forecasters make predictions about an economic fundamental of interest. This is a relevant example as expectations play a central role in every segment of economics. In this setting we show that social learning aimed at increasing the accuracy of individual forecasts may reduce accuracy at the aggregate level. Our main results, contained in Propositions 1 – 3 below, can be summarized as follows.

First, the case of isolated forecasters, i.e., forecasters acting only in reaction to their own signal, represents an upper bound for the error of individual forecasts measured as expected squared forecast error. In fact, when agents can observe each other's predictions, they are able to increase the accuracy of their forecasts. This is quite intuitive as, with social learning, each forecaster is able to exploit the additional information embedded in other agents' predictions. Moreover, this result is independent on the topology of the network.

Second, while reducing the error at the individual level, social learning may, on the other hand, lead to an increase of the error at the aggregate level defined as expected squared error of the average forecast. In fact, the case of isolated agents represents a lower bound for the squared error of aggregate forecasts and, depending on the

[1]Models with noisy signals and strategic complementarities have been considered by Morris and Shin (2002), Angeletos and Pavan (2007), Angeletos and La'O (2013), Colombo et al. (2014), Benhabib et al. (2015), Chahrour and Gaballo (2015) and Angeletos et al. (2016) among others. All papers mentioned above abstract from considerations about the impact of social learning at the aggregate level for different social networks topologies, which is instead the focus of this paper.

properties of the network topology, learning by observing other agents' predictions may reduce aggregate accuracy.

Finally, our paper shows that imperfect information may have an impact on the squared error of the average forecast even when the number of forecasters is large. The effect of idiosyncratic noise on aggregate outcomes has traditionally been discarded following the standard diversification argument. According to the latter, in an economy composed by N agents, the aggregate effect of independent idiosyncratic disturbances should decay at a rate $1/\sqrt{N}$ and it is therefore negligible in large economies (see e.g., Lucas, 1977). We show that the diversification argument may not hold in our framework in the presence of social networks. In fact, social networks may translate imperfect information at the individual level into forecast errors at the aggregate level. In particular, our results demonstrate that when the informational network is very asymmetric, i.e., a small set of agents have relatively high in-degrees, then the aggregate squared error decays at a rate much slower than $1/\sqrt{N}$. In other words when many agents look at the forecasts of the same small number of agents, then the influence exerted by agents that are very *central* in the informational network decays very slowly as the number of agents in the economy increases.

Results on aggregate forecast errors have important implications for economic outcomes. Expectations play a key role in economics and finance since decisions of economic agents crucially depend on their beliefs about the state of the economy. Households' consumption decisions depend on their expected income and real interest rates. In the presence of nominal rigidities, firms' pricing decisions depend on their expectations about marginal costs and inflation. Trading behavior in financial markets is affected by beliefs about fundamentals. Moreover, policy decisions may react to markets' expectations. For example, central banks increasingly monitor private sector expectations when choosing monetary policy. As a consequence, economic performance and welfare are strongly affected by aggregate expectations. Consider as a first example the outcome of court trials. In situations where court decisions depend on deliberations of a jury, errors in aggregate beliefs may result in biased sentences. Moreover, it is easy to show that aggregate forecast errors are positively related to the volatility of aggregate forecasts. In models where aggregate consumption depends on aggregate expectations (see e.g. Branch and McGough, 2009), welfare losses are proportional to the volatility of aggregate expectations as shown by Lucas (2003). Finally, prices in financial markets typically depend on aggregate forecasts about future prices and economic fundamentals. Volatility of aggregate beliefs results in volatility of market prices, which in turn affects welfare as in the case of investors that are mean-variance optimizers.

Our work relates to several strands of research. Banerjee (1992), Bikhchandani et al. (1992) and Smith and Sørensen (2000) among others study learning models in which agents can observe the actions of other agents. These papers investigate whether sequential learning mechanisms, defined as observational learning, lead to informational cascades, and to an inefficient aggregation of private information. Informational cascades are defined in Bikhchandani et al. (1992) as situations in which agents follow the actions of the preceding agents, disregarding their own private information. Informational cascades emerge when the set of possible actions is discrete (Bikhchandani et al., 1998). In our analysis we consider simultaneous decisions as in Gale and Kariv (2003), and a continuous set of possible actions, ruling out the possibility of informational cascades. Moreover, we introduce a network structure defining the patterns of social learning and show that it plays an important role for aggregate outcomes. Ellison and Fudenberg (1993, 1995) study private information aggregation when agents can observe choices and payoffs of other agents and use rule of thumb heuristics to decide their own action. They show that even with simple heuristic behaviors, social learning can lead to efficient outcomes. We analyze instead a framework in which optimizing agents do not observe other agents' payoffs and we focus on the impact of the observational network on individual and aggregate deviations from the fundamental.

The papers stemming from the seminal contribution of DeGroot (1974), e.g., Bala and Goyal (1998), Golub and Jackson (2010), Acemoglu et al. (2011) and DeMarzo et al. (2003), analyze instead richer network structures. These papers focus on network topologies ensuring convergence to the true underlying fundamental (or unidimensional opinions) both under Bayesian and non-Bayesian learning/updating of beliefs. We consider a setting in which agents aggregate information optimally conditional on their knowledge of the network structure as described in Section 2, and we focus on the accuracy of aggregate actions.

Another important stream of literature related to our work concerns the social value of public information in presence of imperfect private information (see e.g., Morris and Shin, 2002; Angeletos and Pavan, 2004; Colombo et al., 2014, among others). The model in which we frame our analysis presents several substantial differences with this literature. First, the focus of our paper is on the impact of social learning on the accuracy of aggregate forecasts rather then on social welfare. Second, in our framework agents do not have access to public information, but they can observe the actions of different subsets of other agents. Finally, to isolate the network effect on the individual actions, we assume that there are no strategic complementarities.

Our work is also related to the literature on the impact of idiosyncratic shocks at the aggregate level. Dupor (1999) and Horvath (1998, 2000) debated about the diversification argument mentioned above. Gabaix (2011) shows that the $1/\sqrt{N}$ diversification argument does not apply when the firm size distribution is sufficiently fat-tailed, while Acemoglu et al. (2012) show that the argument is not valid in the presence of asymmetric input-output links between sectors.[2] We show that, even in the presence of economic agents with identical size and without input-output relations between different sectors, the diversification argument may fail in the presence of dispersed information and social learning.

Finally, this paper is similar in spirit to Barrdear (2014), who studies the influence of social learning over an opaque observational network. Barrdear (2014) considers agents interacting in an environment featuring strategic complementarities and focuses on the impact of higher-order beliefs. In such setup, in order to make the problem tractable, it is necessary to impose some restrictions on the network structure and on agents' knowledge of the latter. We consider instead a simpler setup without strategic complementarities, which allows us to isolate the effect of arbitrary network topologies, without imposing any restriction on agents' knowledge of the network.

The outline of the paper is as follows. Section 2 presents a model with dispersed information and social learning. Section 3 analyzes the effect of different network topologies at the individual and aggregate level. Section 4 concludes.

[2]Earlier contributions on the topic include Jovanovic (1987) and Durlauf (1993) who show that strategic complementarities and local firms' interactions may translate shocks occurring at the firm-level into aggregate volatility. Moreover, Bak et al. (1993) focus on the role of supply chains in aggregate fluctuations.

2 Model

The economy is populated by a finite set of expected utility maximizing agents, $\mathfrak{N} = \{1, 2, \ldots, N\}$, indexed by $i = 1 \ldots N$. Individual utility is given by

$$u_i = -(k_i - \theta)^2 \ ,$$

where $k_i \in \mathbb{R}$ denotes the action of agent i and $\theta \in \mathbb{R}$ is an unknown exogenous fundamental. To frame our analysis in a concrete economic example, consider a set of forecasters trying to predict a relevant economic variable. In this setup k_i can be thought of as forecaster i's prediction of the variable of interest denoted by θ. The optimal forecast is therefore given by

$$k_i = \mathrm{E}_i[\theta] = \mathrm{E}[\theta|\mathcal{I}_i] \ , \tag{1}$$

where $\mathcal{I}_i$ is the information set available to agent i. For the sake of simplicity, suppose that θ is drawn independently in each time period. This allows us to drop the time subscript and consider forecast decisions in an arbitrary time period. Moreover, we consider an improper uniform prior distribution for notational convenience. Forecasters receive a private signal s_i about the fundamental

$$s_i = \theta + \sigma\varepsilon_i \ , \tag{2}$$

where σ is the standard deviation of the private signal and $\varepsilon_i \sim \mathcal{N}(0, 1)$ is an i.i.d. idiosyncratic disturbance.[3] Notice that Eqs. (1) - (2) ensure that forecasters' payoffs are independent from each other and that interaction among forecasters is purely informational.

When agents cannot observe forecast decisions of other agents in the economy, the information set of each forecaster consists only of her private signal. Therefore, since all stochastic elements are normally distributed, in isolation we have that

$$k_i = \mathrm{E}[\theta|\mathcal{I}_i] = s_i \ . \tag{3}$$

[3] In order to simplify the analysis, we consider uncorrelated private signals with homogeneous variance σ^2. In Online Appendix 3 we explore the implications of private signals with heterogeneous variance σ_i^2, while in Online Appendix 4 we analyse the implications of correlated private signals.

Now consider the case in which forecasters can observe the predictions of a subset of other forecasters (as in e.g., Banerjee, 1992; Bikhchandani et al., 1992) through an exogenous directed social network. The network is described by an $N \times N$ matrix Ψ, whose elements are $\psi_{ij} \in \{0, 1\}$. If forecaster i observes forecaster j, then $\psi_{ij} = 1$, otherwise $\psi_{ij} = 0$. Matrix Ψ can be asymmetric and links can be one-sided, so that we may have $\psi_{ij} = 1$ and $\psi_{ji} = 0$. The network topology determines the observational structure, i.e., the subset of other forecasters observable by each forecaster. We denote the subset of forecasters observed by forecaster i as $\Psi(i) = \{j \in \mathfrak{N} | \psi_{ij} = 1\}$ and the information set of forecaster i as $\mathcal{I}_i = \{s_i, k_{j \in \Psi(i)}\}$. Given that private signals are normally distributed and, as shown below in Eq. (5), equilibrium decisions are linear aggregation of signals, optimal information weighting strategies are linear. Therefore, forecaster i's expected value of θ and optimal forecast decision can be written as:

$$k_i = \alpha_i s_i + (1 - \alpha_i) \sum_j w_{ij} k_j \ , \tag{4}$$

where α_i denotes the weight given to private information relative to the information obtained from the network, and the set $\{w_{ij}\}$ collects the relative weights assigned to the decisions of observed forecasters. Forecasters set the values α_i and $\{w_{ij}\}$ optimally using available information, therefore the following necessary conditions for optimality must hold:

a) $0 < \alpha_i \leq 1 \ \forall i$;

b) $w_{ii} = 0$ and $w_{ij} \geq 0$ with $j \neq i, j \in \Psi(i) \ \forall i$;

c) $\sum_{j \in \Psi(i)} w_{ij} = 1 \ \forall i$.

The first condition simply states that optimizing forecasters do not disregard their private signal. The second condition means that forecasters must assign weakly positive weights to other forecasters' decisions. The reason why some weights can be zero is that observation of different agents in the network may convey the same information. This means that optimal information weighting should avoid the *persuasion bias*, i.e., failure to account for possible repetitions in received information as defined in DeMarzo et al. (2003).[4] The last condition ensures the unbiasedness of equilibrium

[4]In general, given the network structure, it is not always possible to completely eliminate the persuasion bias. As argued below, our results are valid even if agents fail to properly account

signal aggregation. The specific values of weights $\{\alpha_i, w_{ij}\}$ depend on the exogenous network structure. In order to encompass any possible network topology, in what follows we derive the model's equilibrium for generic weights $\{\alpha_i, w_{ij}\}$ subject only to the general optimality conditions a), b) and c). Although, as argued below, our results do not depend on the specific set of weights $\{\alpha_i, w_{ij}\}$, for the sake of completeness in Section 2.1 we offer a characterization of optimal information weighting for specific network structures.

In each period forecasters choose predictions simultaneously. Let k denote the $N \times 1$ vector of individual forecast decisions, s the $N \times 1$ vector of private signals, D the diagonal matrix defined as $[\mathrm{D}]_{ii} = \alpha_i$, $\forall i$, and W the $N \times N$ stochastic matrix of weights w_{ij} (i.e., $\sum_j w_{ij} = 1$). We can then write Eq. (4) in matrix form as

$$k = \mathrm{D}s + (\mathrm{I} - \mathrm{D})\mathrm{W}k \ ,$$

where the entries of matrices D and W depend on the exogenous network structure and on the weights chosen by firms. The *rational expectations equilibrium* is then given by

$$k = [\mathrm{I} - (\mathrm{I} - \mathrm{D})\mathrm{W}]^{-1}\mathrm{D}s \ . \tag{5}$$

Eq. (5) shows that, when predictions are observable through the informational network, the equilibrium forecast decisions are linear combinations of private signals. The extreme case in which $\alpha_i = 1 \ \forall i$ is equivalent to the case of isolation in which agents only consider their private information and the network does not play any role.

The equilibrium in Eq. (5) can be rationalized as follows. Forecasts are submitted at the end of each time interval. For the sake of concreteness, one can think about respondents to surveys of forecasts (e.g., Survey of Professional Forecasters), who must post a forecast every quarter. Within the quarter, forecasters receive their signal, formulate their provisional forecast and may communicate with other respondents to the survey. Communicated provisional forecasts are then incorporated in the information set to improve forecasts' precisions. The equilibrium in Eq. (5) can be thought of as the fixed point of the dynamic process

$$k_\tau = \mathrm{D}s + (\mathrm{I} - \mathrm{D})\mathrm{W}k_{\tau-1} \ , \tag{6}$$

for possible repetitions of the information they receive and weight information according to their *subjective* relative precision.

occurring in *notional* time τ within the quarter, in which each forecaster observes the provisional predictions of other agents in the social network and updates her beliefs according to Eq. (4). The following lemma establishes conditions for convergence of the dynamic process in Eq. (6) and thus for the the existence of the equilibrium in Eq. (5).

Lemma 1. *When conditions a), b) and c) are satisfied, the rational expectations equilibrium in* (5) *exists for any network topology.*

The proof is in Appendix A, where we show that the equilibrium in Eq. (5) exists also under more general conditions regarding the weights α_i, i.e., when some $\alpha_i = 0$, provided that each firm in the network is reached, directly or indirectly, by at least one signal. At this point it is also worth emphasizing the differences between the information weighting *à la* DeGroot (1974) (see e.g., DeMarzo et al., 2003; Golub and Jackson, 2010, among others) and the information weighting in Eq. (4). In DeGroot (1974) agents take their decisions in each step τ using a weighted average of decisions in $\tau - 1$, including their own. In our setting, agents take their forecast decisions using a weighted average of other agents' predictions and their own private signal. The former mechanism may lead to solutions in which agents do not take into account their own signals in equilibrium. On the other hand, the mechanism described in Eq. (4), provided that $\alpha_i > 0 \; \forall i$, ensures that the information contained in private signals is not disregarded in equilibrium decisions.

The goal of the paper is to investigate the impact of social learning on the expected quadratic deviation of individual and aggregate forecasts from the fundamental. In the remainder we will refer to these measures respectively as *individual* and *aggregate error*. The error of individual forecasts is therefore defined as

$$\mathrm{v}(k_i) \equiv \mathrm{E}(k_i - \theta)^2 \; .$$

Moreover, let us define the aggregate action as the sum of individual forecasts, normalized by the number of agents in the economic system

$$K = \frac{1}{N} \sum_i k_i \; . \tag{7}$$

The error of the aggregate forecast is thus given by

$$\mathrm{v}(K) \equiv \mathrm{E}(K-\theta)^2 \ .$$

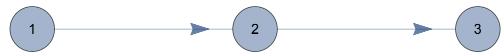
Before proceeding with the analysis, we remark that considering a diffuse prior for θ does not affect our analysis. We are in fact interested in comparing expected squared deviations of individual and aggregate forecast from the fundamental in case of isolation and in the presence of social learning. Results on the comparison between the two cases do not depend on the assumed variance of θ (see Online Appendix 1).

2.1 Equilibrium and Optimal Information Weighting

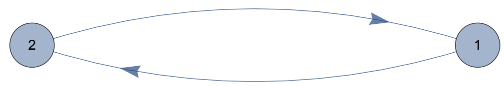
The specific values of $\{\alpha_i, w_{ij}\}$ are endogenous to the optimal information weighting problem and depend on the exogenous network structure. In order to assess the *objective* informational content of each observed decision, and set α_i and w_{ij} accordingly, forecasters must know the exact structure of the network. The reason is that each individual decision is an aggregation of signals, as described in Eq. (5). In fact, the information conveyed by the forecast decision of each agent j, depends not only on its signal s_j, but also on the information (i.e., signals) contained in the decisions observed by agent j and so on. Hence, to weight their private information and the decisions of others according to *objective* informational content, forecasters need to know the source of all the information that influenced, both directly and indirectly, the decisions of other agents. This is consistent with the behavior of optimizing agents setting weights on different sources of information knowing the equilibrium in Eq. (5). In the remainder of the section we characterize optimal weights for given simple network structures represented in Fig. 1.[5] Although extremely simple, such structures can be considered as building blocks for more complicated network topologies.

The network topologies depicted in Fig. 1 are representative of qualitatively different types of informational structures. The topology displayed in Fig. 1(a) refers to an *acyclic* network, i.e., a case in which the information embedded in individual signals is not present in observed actions. On the contrary, the topology in Fig. 1(b) depicts a *cyclic* network, i.e., a scenario in which the information contained in individual signals is present in observed actions. As shown below, these settings correspond to qualitatively different scenarios for information weighting. In both examples each

[5] Details on the computation of optimal weights are provided in Online Appendix 2.



(a) Acyclic network



(b) Cyclic network

Figure 1: Examples of different network types (acyclic vs. cyclic).

forecaster can observe the prediction of only another agent in the economy, therefore, by construction, $w_{ij} = 1$ for $j \in \Psi(i)$ for all i. We can thus focus on α_i, i.e., how each forecaster i weights its private information relative to information in the social network.

Consider the acyclic network in Fig. 1(a). Forecaster 3 observes only her private signal, so she will set $\alpha_3 = 1$ and predict $k_3 = s_3$. In this case the error $\mathrm{v}(k_3)$ coincides with the variance of signal σ^2. Forecaster 2 observes both her private signal s_2 and the prediction k_3 of forecaster 3, so that her forecast is given by $k_2 = \alpha_2 s_2 + (1 - \alpha_2)k_3$. Optimal weighting of information requires forecaster 2 to set α_2 according to the relative precision of s_2, i.e., $\mathrm{v}(s_2)^{-1} = \sigma^{-2}$, compared to precision of the informational content of k_3, namely

$$\alpha_2 = \frac{\sigma^{-2}}{\sigma^{-2} + \sigma^{-2}} = \frac{1}{2} \, ,$$

where we have used the fact that $k_3 = s_3$. Therefore, the error of prediction k_2 is given by

$$\mathrm{v}(k_2) = \left(\alpha_2^2 + (1 - \alpha_2)^2\right) \sigma^2 = \frac{1}{2}\sigma^2 \; .$$

Forecaster 1 observes signal s_1 and the prediction k_2, so that $k_1 = \alpha_1 s_1 + (1 - \alpha_1)k_2$. Therefore weight α_1 is given by

$$\alpha_1 = \frac{\mathrm{v}(s_1)^{-1}}{\mathrm{v}(s_1)^{-1} + \mathrm{v}(k_2)^{-1}} = \frac{\sigma^{-2}}{\sigma^{-2} + 2\sigma^{-2}} = \frac{1}{3} \; .$$

Therefore the error of forecast k_1 is given by

$$\mathrm{v}(k_1) = \big(\alpha_1^2 + (1-\alpha_1)^2(\alpha_2^2 + (1-\alpha_2)^2)\big)\,\sigma^2 = \frac{1}{3}\sigma^2\ .$$

Consider now the cyclic network in Fig. 1(b), where forecaster 1 observes forecaster 2, and viceversa. The difference with the acyclic network is that in this case the information contained in the signal of e.g., forecaster 1 enters her decision both directly and indirectly, through observation of the prediction of forecaster 2. Therefore, for each $\alpha_1 > 0$, we have that forecaster 1 suffers from persuasion bias, in the sense that she fails to account for the repetition of information contained in s_1. The same reasoning applies to forecaster 2. If the information conveyed by individual signal s_i is already present in her observational network, forecaster i should set $\alpha_i \to 0$ in order to minimize the persuasion bias and avoid overweighting the information contained in s_i. Notice however that α_i should be strictly positive, otherwise the information contained in s_i would not enter the network at all, violating optimality condition a) above. To formalize the argument above, consider the forecasts of agents 1 and 2 in the simple network described in Fig. 1(b):

$$\begin{aligned} k_1 &= \alpha_1 s_1 + (1-\alpha_1)k_2 \\ k_2 &= \alpha_2 s_2 + (1-\alpha_2)k_1\ . \end{aligned}$$

Equilibrium forecasts are therefore given by

$$\begin{aligned} k_1 &= \frac{\alpha_1}{\alpha_1+\alpha_2-\alpha_1\alpha_2}s_1 + \frac{(1-\alpha_1)\alpha_2}{\alpha_1+\alpha_2-\alpha_1\alpha_2}s_2 \\ k_2 &= \frac{\alpha_2}{\alpha_1+\alpha_2-\alpha_1\alpha_2}s_2 + \frac{(1-\alpha_2)\alpha_1}{\alpha_1+\alpha_2-\alpha_1\alpha_2}s_1\ . \end{aligned}$$

From the equations above it is clear that each equilibrium forecast is a combination of both s_1 and s_2. Since signals have the same precision, optimal information weighting requires the equilibrium weights of both signals to be 1/2. Consider for example forecast k_1. When $(\alpha_1, \alpha_2) \to (0,0)$ along the same line, we indeed have that

$$\lim_{(\alpha_1,\alpha_2)\to(0,0)} \frac{\alpha_1}{\alpha_1+\alpha_2-\alpha_1\alpha_2} = \frac{1}{2}\ .$$

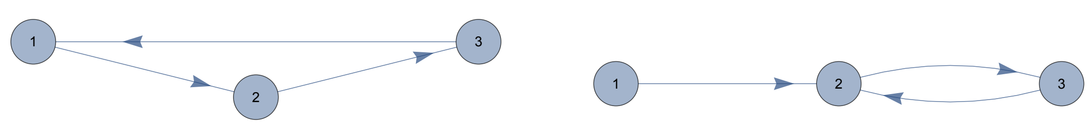
A symmetric argument applies to equilibrium weights in forecast k_2.[6] The results derived for these simple examples can be applied to more complex network structures.

Finally, we remark that the optimal weighting described above requires complete knowledge of the network structure, as well as common knowledge of rationality. However, all the results derived in the paper are valid for *any* set of weights $\{\alpha_i, w_{ij}\}$ provided that the necessary conditions a), b), c) listed in Section 2 are satisfied. In particular, our findings also apply to cases in which forecasters are boundedly rational and weight information according to *subjective* (instead of *objective*) relative precisions, which can be thought of as the outcome of any estimation heuristic.

3 Network Topology, Individual and Aggregate Error

In this section we focus on the impact of social learning on the accuracy of both individual and aggregate forecasts, and relate it to the structure of the network characterizing the interaction patterns among forecasters.

To help the intuition we will complement the exposition of our results with simple examples. Fig. 2 displays two different network configurations. In both cases we have $N = 3$ and all forecasters observe the prediction of another agent. In the network depicted in Fig. 2(a) each forecaster observes the prediction of a different agent, so that all nodes in the network have the same in-degree. In the network depicted in Fig. 2(b) all forecasters observe the same agent $i = 2$ and the latter observes another forecaster $j = 3$, so that in-degrees are heterogeneous across nodes in the network.



(a) Homogeneous in-degree network (b) Heterogeneous in-degree network

Figure 2: Examples of different network types (homogeneous vs. heterogeneous in-degree).

[6] It is reasonable to consider $(\alpha_1, \alpha_2) \to (0, 0)$ along the same line since signals have the same precision and there is common knowledge of rationality.

The corresponding adjacency matrices are shown in Eq. (8)

$$\mathrm{W}_{hom} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \qquad \mathrm{W}_{het} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} , \tag{8}$$

where W_{hom} refers to the homogeneous in-degree network in Fig. 2(a) and W_{het} refers to the heterogeneous in-degree network in Fig. 2(b). In both cases, given that each forecaster can observe only another forecaster in the network, the weight w_{ij} assigned to that observation is equal to 1 by construction. In the following sections we will use these simple network configurations as examples to illustrate the impact of different topologies on forecast errors at the individual and aggregate level, but we remark that our results are derived for arbitrary network structures.

3.1 Accuracy of Individual Forecasts

In what follows we show that social learning leads to higher accuracy of individual forecasts. Using Eq. (5), individual forecasts in equilibrium can be written as

$$k_i = \sum_j \hat{w}_{ij} \alpha_j s_j ,$$

where $\hat{w}_{ij}$ denotes the element (i, j) of matrix $\hat{\mathrm{W}}$ defined as $\hat{\mathrm{W}} \equiv [\mathrm{I} - (\mathrm{I} - \mathrm{D})\mathrm{W}]^{-1}$.

Before proceeding with the analysis, we describe an important property of the matrix $\mathrm{C} \equiv \hat{\mathrm{W}}\mathrm{D}$, which maps private signals into equilibrium forecasts according to Eq. (5), in the following lemma.

Lemma 2. *Matrix C is a stochastic matrix, i.e.,* $\sum_j \hat{w}_{ij}\alpha_j = 1 \ \forall i$.

The proof is in Appendix A. The error of individual forecasts is then given by

$$\begin{aligned} \mathrm{v}(k_i) &= \mathrm{E}(k_i - \theta)^2 \\ &= \mathrm{E}\left(\sum_j \hat{w}_{ij}\alpha_j s_j - \theta\right)^2 \\ &= \mathrm{E}\left(\theta \sum_j \hat{w}_{ij}\alpha_j + \sum_j \hat{w}_{ij}\alpha_j \varepsilon_j - \theta\right)^2 \\ &= \sum_j \hat{w}_{ij}^2 \alpha_j^2 \sigma^2 , \end{aligned} \tag{9}$$

where the last equality follows from Lemma 2 and from the fact that idiosyncratic disturbances are independent across agents. In the case of isolated forecasters there are no informational links between them, i.e., W is a matrix of zeros (hence $\hat{\mathrm{W}}$ is an identity matrix), and $\alpha_i = 1 \; \forall i$, meaning that agents only consider their private signal and therefore

$$\mathrm{v}(k_i) = \sigma^2 \; . \tag{10}$$

Social learning happens when $\alpha_i < 1$ and $w_{ij} > 0$ for at least one $j \neq i$, which means that forecaster i has at least one informational link with another forecaster j, and she uses the information embedded in the prediction made by forecaster j. The impact of social learning on individual forecast errors is stated in the following proposition:

Proposition 1. *The individual expected squared forecast error in the case of social learning is always less than the individual expected squared forecast error in case of isolated agents, that is*

$$\sum_j \hat{w}_{ij}^2 \alpha_j^2 < 1 \; . \tag{11}$$

The proof is in Appendix A. Proposition 1 shows that social learning is beneficial from forecasters' point of view. In this way agents are in fact able to increase the precision of their individual forecasts of the fundamental. The intuition for this result is that the information about fundamentals contained in individual signals is spread through the network and in equilibrium forecasters are able to exploit this additional information by observing the predictions of other agents.

The following examples illustrate the information spreading mechanism in the networks described in Fig. 2. Before proceeding, we remark that the reasoning followed in the simple examples below to set weights $\{\alpha_i, w_{ij}\}$ reflects information weighting according to objective relative precisions along the lines described in Section 2.1. As discussed in Section 2.1, qualitative results do not change when information is weighted according to generic subjective relative precisions.

Example 1.

Consider the informational structure described in Fig. 2(a). In this case the prediction of e.g., forecaster 1 is given by $k_1 = \alpha_1 s_1 + (1 - \alpha_1) w_{12} k_2$, i.e., using both the private signal s_1 and the prediction of forecaster 2, which in turn is formulated using her private signal s_2 and the prediction of forecaster 3. The network structure allows to incorporate information from other forecasters' signals in individual predictions. In fact, in equilibrium we have that forecasts are determined according to Eq. (5) and therefore the way in which private signals are spread through the network depends on the matrix $\mathrm{C} \equiv [\mathrm{I} - (\mathrm{I} - \mathrm{D})\,\mathrm{W}]^{-1}\,\mathrm{D}$. Given the structure of the network, each forecaster observes directly or indirectly the predictions of all other forecasters in the network. Therefore, for each forecaster i, the information conveyed by the network already includes her own signal s_i. Consider for example forecaster 1 who is observing the prediction of forecaster 2, which in turn observes the prediction of forecaster 3. The structure of the network implies that forecaster 1 is observing indirectly forecaster 3, which is in turn observing forecaster 1 herself. The network is therefore cyclic and each forecaster is facing the optimal weighting problem discussed in Section 2.1 for the topology in Fig. 1(b). In fact, the information contained in the signal of forecaster i is included in her prediction both directly, with weight α_i, and indirectly, with weight $1 - \alpha_i$. When all forecasters set $\alpha_i \to 0$ (strictly positive), matrix $\mathrm{C}_{hom} \equiv [\mathrm{I} - (\mathrm{I} - \mathrm{D})\,\mathrm{W}_{hom}]^{-1}\,\mathrm{D}$ is given by

$$\mathrm{C}_{hom} = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix} \qquad .$$

The expression above shows that, due to the network of observational links, the individual prediction of each forecaster is influenced by the private signals of all forecasters. For example, prediction k_1 is given by

$$k_1 = \sum_{j=1}^{3} c_{1j} s_j = \frac{1}{3} s_1 + \frac{1}{3} s_2 + \frac{1}{3} s_3 \ .$$

The expression above makes clear that matrix C_{hom} reflects optimal information weighting since in equilibrium each signal is weighted according to its objective relative precision. Given the topology of the network in the example we have that, for each forecaster, the error of individual predictions with social learning is given by $0.33\sigma^2$ and thus smaller than the case of isolation. ♣

The reduction of individual forecast errors in the presence of social learning with respect to the case of isolation is independent of the network structure. In fact, as long as a forecaster has at least one informational link and therefore looks at the prediction of at least one additional forecaster (on top of reacting to its own signal), the deviation of its forecast from the fundamental will be lower than the case of isolation.

Nevertheless, different network topologies may imply different levels of individual squared errors, as shown in the following example.

Example 2.

Consider the informational structure described in Fig. 2(b). Forecasters 2 and 3 are solving the same information weighting problem described in Section 2.1 for the topology in Fig. 1(b). Therefore we have that $\alpha_2 = \alpha_3 \to 0$ (strictly positive). Forecaster 1 observes the prediction of forecaster 2, which contains information from s_2 and s_3, while her prediction is not observed by any other forecaster in the network. Therefore she faces an information weighting problem similar to that of forecaster 1 in the example in Fig. 1(a). In fact, to reflect objective relative precision, the weight assigned by forecaster 1 to her own signal must be equal to the inverse of the total number of signals affecting her equilibrium forecast, i.e., $\alpha_1 = 1/3$. Matrix C_{het} is therefore equal to

$$\mathrm{C}_{het} = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix} \qquad .$$

The expression above shows that the prediction of forecaster 1 is based on information contained in the individual signals of two additional forecasters, while the predictions of forecasters 2 and 3 only exploit information from one additional agent. We have therefore that $\mathrm{v}(k_1) = 0.33\sigma^2$, while $\mathrm{v}(k_2) = \mathrm{v}(k_3) = 0.5\sigma^2$. In the presence of heterogeneity in agents' in-degrees, individual prediction errors may differ across forecasters, but they are still lower than the case of isolated forecasters. ♣

3.2 Accuracy of Aggregate Forecast

In this section we show that when forecasters look at other agents' predictions to make their own forecasts, the error of the aggregate forecast may increase with respect to the case of isolation depending on the topology of the informational network. Using Eqs. (5) and (7), we can write aggregate forecast as

$$K = \frac{1}{N} \sum_i \sum_j \hat{w}_{ij} \alpha_j s_j \ , \tag{12}$$

where again $\hat{w}_{ij}$ denotes the element (i, j) of the matrix $\hat{\mathrm{W}}$ defined as $\hat{\mathrm{W}} \equiv [\mathrm{I} - (\mathrm{I} - \mathrm{D})\mathrm{W}]^{-1}$. Using matrix $\mathrm{C} \equiv \hat{\mathrm{W}}\mathrm{D}$ we can define the $1 \times N$ vector v' as

$$v' \equiv e'\mathrm{C} \ , \tag{13}$$

where $e' \equiv [1, \ldots, 1]$, so that $v_j = \sum_i c_{ij} = \alpha_j \sum_i \hat{w}_{ij}$ and

$$K = \frac{1}{N} \sum_j v_j s_j \ .$$

Vector v can be defined as an *influence vector*, since each element v_j determines the influence of signal s_j on aggregate forecast. The influence vector v is related to the Bonacich (in-degree) centrality measure (Bonacich, 1987). If the Bonacich centrality of forecaster j (summarized by the term $\sum_i \hat{w}_{ij}$) in the observational network increases, the influence of forecaster j's signal will increase. But there is also a second effect. Given the observational network, increasing α_j (holding constant all $\alpha_{i \neq j}$), will increase the influence of forecaster j's signal. The intuition is that, if α_j is relatively high, forecaster j's signal will be largely reflected in its prediction, and therefore it will have relatively higher influence on the predictions of agents observing j. The squared error of K can be written as

$$\begin{aligned} \mathrm{v}(K) &= \mathrm{E}(K-\theta)^2 \\ &= \mathrm{E}\left(\frac{1}{N}\sum_i\sum_j \hat{w}_{ij}\alpha_j s_j - \theta\right)^2 \\ &= \mathrm{E}\left(\frac{1}{N}\sum_i \theta \sum_j \hat{w}_{ij}\alpha_j + \frac{1}{N}\sum_i\sum_j \hat{w}_{ij}\alpha_j\varepsilon_j - \theta\right)^2 \\ &= \frac{1}{N^2}\sum_j v_j^2\sigma^2 \ , \end{aligned} \tag{14}$$

where the last equality follows from Lemma 2 and from the fact that idiosyncratic disturbances are independent across agents. In the absence of social learning we have that $v_j = 1$ for all j, and therefore the error of the aggregate forecast is

$$\mathrm{v}(K) = \frac{\sigma^2}{N} \ . \tag{15}$$

The impact of social learning on the accuracy of the aggregate forecast is described in the following proposition:

Proposition 2. *The expected squared error of the aggregate forecast in the case of social learning is always greater than, or equal to, the expected squared error of the aggregate forecast in the case of isolated agents, that is*

$$\frac{1}{N}\sum_j v_j^2 \geq 1 \ . \tag{16}$$

The proof is in Appendix A. According to Proposition 2, the case of isolated forecasters represents a lower bound for the aggregate error. The intuition for this result is that social learning introduces correlation among individual forecasts. In the absence of social learning we have that individual decisions are independent from each other and therefore the aggregate error depends on the sum of individual errors, that is

$$\mathrm{v}(K) = \frac{1}{N^2}\sum_i \mathrm{v}(k_i) \ . \tag{17}$$

In the presence of social learning, individual decisions are not independent and therefore the aggregate error depends also on the covariance among individual decisions, as dictated by the network structure, so that

$$\text{v}(K) = \frac{1}{N^2} \sum_i \text{v}(k_i) + \frac{2}{N^2} \sum_{i \neq j} \text{cov}(k_i, k_j) \; . \tag{18}$$

The impact of social learning is twofold. First, as shown in Proposition 1, The first term of Eq. (18), i.e., the sum of individual forecast errors is lower when compared to the case of isolated forecasters in Eq. (17). Second, social learning introduces a covariance element given by the second term in Eq. (18). Proposition 2 shows that the net effect depends on the structure of the observational network and on the weights attached by each forecaster to the different sources of information. In particular, if vector v has heterogeneous entries, the error of the aggregate forecast increases. The only case in which the error of the aggregate forecast under social learning is equal to the case of isolation is when the signal of each forecaster in the network has exactly the same influence on aggregate forecast, i.e., when $v_j = 1$ for all j. This scenario is verified when all forecasters in the network have the same weighted in-degree and out-degree, as in Fig. 2(a), and weights α_i are homogeneous. Any other case results in an influence vector with heterogeneous elements and thus the error of the aggregate forecast increases with respect to the case of isolation.

The following example illustrates the impact of heterogeneous centralities among forecasters on aggregate accuracy.

Example 3.

Consider the networks described in Figs. 2(a) and 2(b) and the matrices of equilibrium weights derived in Examples 1 and 2. Given matrices C_{hom} and C_{het}, we can compute the influence vectors associated to each network using Eq. (13):

$$v_{hom} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \qquad v_{het} = \begin{pmatrix} 1/3 \\ 4/3 \\ 4/3 \end{pmatrix} \; .$$

Each element v_i of the influence vector describes the sum of equilibrium weights assigned to signal s_i. The higher v_i, the higher the *centrality* of forecaster i. In the network described in Fig. 2(b), forecaster 2 is highly central since all other forecasters look at her prediction when making their own choices. Note that forecaster 3 is also highly central. The reason is that the prediction of forecaster 3 is used by forecaster 2. The centrality of each forecaster i, as measured by the influence vector, depends not only on the number of agents directly observing forecaster i, but also on the number of other forecasters observing agents who are looking at i. This means that the influence of each forecaster is recursively related to the influence of other forecasters who observe her prediction (see Jackson et al., 2015). On the contrary, in the network described in Fig. 2(a), all firms have the same centrality. Using Eq. (14) we can compute the error of aggregate forecasts for both networks

$$\mathrm{v}(K_{hom}) = 0.33\sigma^2 \qquad \mathrm{v}(K_{het}) = 0.41\sigma^2 \ .$$

For the network in Fig. 2(a), the error of the aggregate forecast is equal to the case of isolated forecasters, i.e., $\sigma^2/3$, meaning that the reduction in the sum of individual errors and the positive covariances in individual decisions balance each other out (see Eq. (18)). On the opposite, heterogeneity in the centrality of forecasters, as in the case of the network in Fig. 2(b), leads to an increase in the error of the aggregate forecast. ♣

3.3 Decay Rate of the Aggregate Forecast Error

In this section we characterize the decay rate of the aggregate forecast error as the number of forecasters in the network grows large. In particular, we show how the limiting behavior of the aggregate error as $N \to \infty$ depends on the distribution of the influence vector's elements and ultimately on the structural properties of the informational network.

Consider a sequence of networks indexed by the number of forecasters $N \geq 1$, with the structure of informational links described by W_N. The corresponding sequences of aggregate forecasts and influence vectors are denoted respectively by $\{K_N\}$ and $\{v_N\}$. Assuming that the variance of idiosyncratic signals is independent of the size of the network N, i.e., $\sigma_N^2 = \sigma^2 \ \forall N$, we have that

$$\mathrm{v}(K_N) = \frac{\sigma^2}{N^2} \sum_j v_{j,N}^2 \ .$$

Let us define the following notation. Given two series of positive real numbers $\{x_N\}$

and $\{a_N\}$, we write $x_N \sim a_N$ if there exist $M \geq 0$ and finite constants $A > 0$ and $A \leq B < \infty$, such that $\inf_{N \geq M} x_N/a_N \geq A$ and $\sup_{N \geq M} x_N/a_N \leq B$. In other words, $x_N \sim a_N$ means that for $N \geq M$ the sequences $\{x_N\}$ and $\{a_N\}$ grow at the same rate. Moreover, we write $x_N = \Omega(a_N)$ if there exist $M \geq 0$ and finite constant $A > 0$, such that $\inf_{N \geq M} x_N/a_N \geq A$. This means that, for decreasing sequences, x_N decreases more slowly than a_N.

Therefore we can write that

$$\sqrt{\mathrm{v}(K_N)} \sim \frac{1}{N} \|v_N\|_2 \ . \tag{19}$$

Eq. (19) implies that the error of the aggregate forecast may decay with a rate different from $1/\sqrt{N}$, i.e., the rate implied by the law of large numbers, according to the properties of the influence vector v_N. The limiting behavior of $\|v_N\|_2$ as $N \to \infty$ depends on the distribution of $v_{j,N}$. In the following proposition we characterize the decay rate of the aggregate error for both cases of fat-tailed and thin-tailed distribution of $v_{j,N}$.

Proposition 3. *Consider a series of forecasters' networks indexed by $N \geq 1$. Assume that elements in the sequence of influence vectors $v_1, \dots, v_N$ have a power law distribution, so that*

$$P_N\left(v_{j,N} > x\right) = c_N L(x) x^{-\zeta}$$

where $c_N \sim 1$ is a sequence of real positive numbers, $L(x)$ is slowly varying function, meaning that $\lim_{x\to\infty} L(x)x^{\epsilon} = \infty$ and $\lim_{x\to\infty} L(x)x^{-\epsilon} = 0$ for all $\epsilon > 0$, and $\max_j v_{j,N} \sim N^{1/\zeta}$. Then, as $N \to \infty$, the aggregate error follows

a) $\sqrt{\mathrm{v}(K_N)} = \Omega\left(N^{(1-\zeta)/\zeta - \epsilon'}\right)$ *for* $1 < \zeta \leq 2$

b) $\sqrt{\mathrm{v}(K_N)} \sim N^{-1/2}$ *for* $\zeta > 2$

where $\epsilon' = \epsilon/(2\zeta)$.

The proof is in Appendix A. Proposition 3 implies that when the distribution of forecasters signals' influence v has thin tails ($\zeta > 2$), then the (squared root of the) aggregate forecast error decays at rate $1/\sqrt{N}$. On the contrary, when the distribution has fat tails ($1 < \zeta \leq 2$) the decay rate is much slower. A fat-tailed distribution of v implies a greater heterogeneity in the influences of forecasters' signals, corresponding to the case in which many forecasters look at the prediction of the same small number of forecasters. When a small number of forecasters are highly central, hence very influential in the network, the impact of their signals decays slowly as the number of agent N increases. This implies that, in the presence of dispersed information, social learning and very asymmetric network structures, significant aggregate errors may result from idiosyncratic noise at the micro-level, even if the number of firms is very large.

Influence vector and network primitives

We can relate forecasters' influence to the structural properties of the observational network. Using Lemma 1 (and results in Appendix A.1) we can write matrix C, mapping individual signals into equilibrium forecasts, as

$$\mathrm{C} = \sum_{z=0}^{\infty} [(\mathrm{I}-\mathrm{D})\mathrm{W}]^z \, \mathrm{D} \ .$$

From the expression above and the definition of influence vector in Eq. (13), it follows that

$$v' = e' \left(\sum_{z=0}^{\infty} [(\mathrm{I}-\mathrm{D})\mathrm{W}]^z \, \mathrm{D} \right) \ ,$$

which implies that the influence vector satisfies the following inequality

$$v' \geq e'\mathrm{D} + e'(\mathrm{I}-\mathrm{D})\mathrm{W}\mathrm{D} \ .$$

Each element v_j of the influence vector thus satisfies

$$v_j \geq \alpha_j + \alpha_j \sum_i (1-\alpha_i) w_{ij} \ . \tag{20}$$

From Eq. (20) it straightforward to see that the influence of each signal j on aggregate forecast is directly related to the weighted in-degree of forecaster j, i.e., $\sum_i (1-\alpha_i) w_{ij}$. Since the sum of signal influences is bounded from above, i.e., $\sum_j v_j = N$, it is easy to see that we can only observe an asymmetric distribution of v_j if some forecasters have a large in-degree, i.e., they are observed by a large number of other forecasters in the network. In informational structures characterized by heterogeneous in-degrees, forecasters with a relatively large in-degree have high centrality in the network and therefore their signals have large influence on aggregate forecast.

4 Conclusions

The behavior of peers, friends or in general other members of a social or economic group, represents a valuable source of information for the *homo oeconomicus*. Observation of others' behavior is deeply rooted in human nature as a consequence of the adaptation to complex environments, where it is difficult to collect and process all available information. This paper shows that this micro-behavior, which we have called social learning, can have relevant consequences at the aggregate level. The aggregate effect of social learning depends on the topology of the network describing the links between agents. We show that, according to the network structure, social learning in the presence of dispersed information can propagate idiosyncratic noise at the aggregate level.

If the network is symmetric, in the sense that all agents have the same influence in equilibrium, then the aggregate forecast error with social learning is at its minimum level and coincides with the aggregate error when agents are isolated. For any other network configuration, social learning leads to an increase in the aggregate error. The accuracy of the aggregate forecast is negatively related to the concentration of influence in the network.

Moreover, we show that the diversification argument does not always apply in the presence of social learning. If the influence vector is sufficiently asymmetric, i.e., there exist few very influential agents in the network, then the aggregate impact of independent individual-level shocks does not decay at a rate equal to $1/\sqrt{N}$, resulting therefore in a significant error at the aggregate level.

As previously argued, errors in aggregate beliefs may affect economic performance and welfare whenever decisions and economic outcomes depend on aggregate expectations. Examples range from the outcomes of court trials to welfare losses related to volatility in consumption and asset prices.

An important question regards the empirical relevance of social learning. From a qualitatively point of view, the impact of social learning crucially depends, as argued above, on the topology of the observational network. Galeotti and Goyal (2010) list a series of empirical works suggesting asymmetric topologies in informational networks. In accordance to the empirical evidence, they propose a network formation game, in which asymmetric topologies emerge as the equilibrium outcome. Moreover, Goyal et al. (2016) present experimental evidence supporting the emergence of asymmetric informational networks. Bikhchandani et al. (1992, 1998) argue in favor of the presence of *fashion leaders*, i.e., "expert" agents observed by many other agents, and Gilbert and Lieberman (1987) show that "smaller firms tend to imitate the investment activity of others". Therefore, we conclude that it is highly plausible that observational networks are asymmetric, and consequently that dispersed information and social learning play an important role for the accuracy of aggregate forecasts.

Acknowledgments: We thank Carlo Lucibello, Nicolò Pecora, Paolo Pin and the participants to the CESifo Area Conference on Macro, Money and International Finance (Munich, 2016), JAFEE20 International Conference (Tokyo, 2016), Computing in Economics and Finance (2016), Computing in Economics and Finance (2018) for stimulating discussions and helpful comments.

Appendix A Proofs

A.1 Proof of Lemma 1

Proof. Iterating Eq. (6) we get

$$k_\tau = \sum_{z=1}^{\tau-1} [(\mathrm{I}-\mathrm{D})\mathrm{W}]^z \, \mathrm{D}s + [(\mathrm{I}-\mathrm{D})\mathrm{W}]^\tau \, k_0 \ .$$

where k_0 is the vector of initial conditions, i.e., the initial provisional forecasts s. Within each period forecasters observe the *temporary* predictions of forecasting agents in their information network and update their decision before the true state of the world θ is revealed, resulting in a convergence process *à la* DeGroot (1974) and subsequent literature (e.g., DeMarzo et al., 2003; Golub and Jackson, 2010).

Define $\mathrm{A} \equiv (\mathrm{I}-\mathrm{D})\mathrm{W}$. We then have that $\lim_{\tau\to\infty} \mathrm{A}^\tau = 0$ and $\sum_{z=0}^{\infty} \mathrm{A}^z = [\mathrm{I}-\mathrm{A}]^{-1}$ when the spectral radius of A, defined as $\rho(\mathrm{A}) = \max_{1\le i\le N} |\lambda_i|$ where λ_i is the i-th eigenvalue of A, is strictly smaller than one. Then notice that

$$\|\mathrm{A}\|_\infty = \max\left\{\sum_j |(1-\alpha_i)w_{ij}| \mid 1 \le i \le N\right\} = \max\left\{|(1-\alpha_i)| \mid 1 \le i \le N\right\} \ ,$$

given that matrix W is stochastic. Moreover, for a generic eigenvector-eigenvalue pair (x, λ) with $x \neq 0$, we have that $\lambda x = \mathrm{A}x$ and therefore

$$\|\lambda x\|_\infty = |\lambda| \, \|x\|_\infty = \|\mathrm{A}x\|_\infty \le \|\mathrm{A}\|_\infty \|x\|_\infty \ \Rightarrow |\lambda| \le \|\mathrm{A}\|_\infty \ ,$$

where the inequality follows from the submultiplicativity property of the matrix norm. When $0 < \alpha_i \le 1 \ \forall i \in [1, N]$, we have that $\|\mathrm{A}\|_\infty < 1$ implying that $\rho(\mathrm{A}) < 1$. □

In the following we show that the equilibrium in Eq. (5) exists also under more general conditions. Before identifying such conditions let us introduce the following definitions. Following Golub and Jackson (2010), we define a group of nodes $Z \subset N$ as *closed* relative to a generic adjacency matrix Ω if $i \in Z$ and $\omega_{ij} > 0$ imply that $j \in Z$. A closed group of nodes is *minimally closed* relative to Ω if it is closed and no nonempty strict subset is closed.

The equilibrium in Eq. (5) exists when

a) $0 \leq \alpha_i \leq 1 \; \forall i \in [1, N]$ *with at least one* $\alpha_i > 0$*, and matrix* W *is irreducible.*

b) $0 \leq \alpha_i \leq 1$ and for each minimally closed group $\mathcal{Z}_j$ relative to (I − D)W there exists at least one $i \in \mathcal{Z}_j$ such that $\alpha_i > 0$.

Proof. We start by proving that if $\mathrm{A} = (\mathrm{I} - \mathrm{D})\mathrm{W}$ is irreducible and at least one $\alpha_i > 0$, then it must be that $\rho(\mathrm{A}) < 1$. Notice that when at least one $\alpha_i > 0$, then matrix A is substochastic. Denoting by $e' = [1 \dots 1]$, this implies that $\mathrm{A}e \leq e$ and $\mathrm{A}e \neq e$. When matrix A is irreducible, from the Perron-Frobenius theorem it follows that $\rho(\mathrm{A}) = 1$ would imply $\mathrm{A}e = e$, which is impossible by construction. Therefore $\rho(\mathrm{A}) < 1$ follows from the result $\rho(\mathrm{A}) \leq \|\mathrm{A}\|_\infty$ derived above.

Let's now consider the case in which A is reducible. In general, if $0 \leq \alpha_i < 1$, the reducibility of matrix A follows from the reducibility of matrix W. If instead $\alpha_i = 1$ for some agents $i \in [1, N]$, then matrix A is reducible even if matrix W is irreducible. In what follows we define conditions such that $\rho(\mathrm{A}) < 1$ when A is reducible. If A is reducible, then following Meyer (2000, page 694), it is possible to write matrix A in the canonical form for reducible matrices

$$\mathrm{A} \sim \left(\begin{array}{cccc|cccc} \mathrm{A}_{11} & \mathrm{A}_{12} & \cdots & \mathrm{A}_{1r} & \mathrm{A}_{1,r+1} & \mathrm{A}_{1,r+2} & \cdots & \mathrm{A}_{1,m} \\ 0 & \mathrm{A}_{22} & \cdots & \mathrm{A}_{2r} & \mathrm{A}_{2,r+1} & \mathrm{A}_{2,r+2} & \cdots & \mathrm{A}_{2,m} \\ \vdots & & \ddots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \mathrm{A}_{rr} & \mathrm{A}_{r,r+1} & \mathrm{A}_{r,r+2} & \cdots & \mathrm{A}_{r,m} \\ \hline 0 & 0 & \cdots & 0 & \mathrm{A}_{r+1,r+1} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \mathrm{A}_{r+2,r+2} & \cdots & 0 \\ \vdots & & \cdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & \mathrm{A}_{m,m} \end{array}\right),$$

where each $\mathrm{A}_{11}, \dots, \mathrm{A}_{rr}$ is either irreducible or $[0]_{1\times 1}$, and $\mathrm{A}_{r+1,r+1}, \dots \mathrm{A}_{mm}$ are irreducible. As noted in Meyer (2000), the effect of such a symmetric permutation is simply to relabel the nodes in the original network. Therefore a generic A_{sz} denotes the sub-network describing the connections from agents in rows s to agents in columns z. Define $\alpha(s)$ as the vector containing the α_i set by each agent i belonging to the set described by rows s.

As a first step, consider the matrices A_{kk} for $k = 1, 2, \dots, r$ and observe that $\rho(\mathrm{A}_{kk}) < 1$ for each $k = 1, 2, \dots, r$, for any possible value of entry α_i in vector $\alpha(k)$. This is certainly true when $\mathrm{A}_{kk} = [0]_{1\times 1}$, so consider the case in which A_{kk} is irreducible and notice that A_{kk} is substochastic by construction because there must be blocks A_{kj}, $j \neq k$ that have nonzero entries. From the previous result we know that irreducible substochastic matrices are characterized by a spectral radius strictly smaller than one.

Consider now the matrices A_{kk} for $k = r+1, \ldots, m$, which refer to the minimally closed groups relative to A. These matrices are substochastic if and only if at least one α_i in vector $\alpha(k)$ is positive. Once again, these irreducible stochastic matrices are characterized by a spectral radius strictly smaller than one.

Therefore, when at least one agent i in each minimally closed group of A sets $\alpha_i > 0$, we conclude that $\rho(\mathrm{A}) < 1$. □

In other words, for equilibrium in Eq. (5) to exist, each agent must receive at least one signal directly and/or indirectly. If $\alpha_i > 0 \; \forall i$, this condition is satisfied for any possible network topology. If instead $\alpha_i = 0$ for some agent i, then for equilibrium in Eq. (5) to exist, at least one agent i in any minimally closed group must have $\alpha_i > 0$ guaranteeing that all agents in the network are reached by at least one signal.

A.2 Proof of Lemma 2

Proof. Define the vector $e' = [1 \ldots 1]$. Proving that $\sum_j \hat{w}_{ij}\alpha_j = 1 \; \forall i$ is equivalent to prove that $\mathrm{C}e = e$ or equivalently that $\mathrm{C}^{-1}e = e$, since C is invertible. Start from

$$\mathrm{C} = [\mathrm{I} - (\mathrm{I} - \mathrm{D})\mathrm{W}]^{-1}\mathrm{D} \; ,$$

and pre-multiply both sides by C^{-1} to get

$$\mathrm{I} = \mathrm{C}^{-1}[\mathrm{I} - (\mathrm{I} - \mathrm{D})\mathrm{W}]^{-1}\mathrm{D} \; .$$

Post-multiplying by D^{-1}

$$\mathrm{D}^{-1} = \mathrm{C}^{-1}[\mathrm{I} - (\mathrm{I} - \mathrm{D})\mathrm{W}]^{-1} \; ,$$

and by $[\mathrm{I} - (\mathrm{I} - \mathrm{D})\mathrm{W}]$ we get

$$\mathrm{D}^{-1}[\mathrm{I} - (\mathrm{I} - \mathrm{D})\mathrm{W}] = \mathrm{C}^{-1} \; .$$

Post-multiplying both sides by e, we have

$$\begin{aligned}
\mathrm{C}^{-1}e &= \mathrm{D}^{-1}[\mathrm{I}-(\mathrm{I}-\mathrm{D})\mathrm{W}]e\\
\mathrm{C}^{-1}e &= \mathrm{D}^{-1}[e-(\mathrm{I}-\mathrm{D})\mathrm{W}e]\\
\mathrm{C}^{-1}e &= \mathrm{D}^{-1}[e-(\mathrm{I}-\mathrm{D})e]\\
\mathrm{C}^{-1}e &= \mathrm{D}^{-1}[e-e+\mathrm{D}e]\\
\mathrm{C}^{-1}e &= \mathrm{D}^{-1}\mathrm{D}e\\
\mathrm{C}^{-1}e &= e\,,
\end{aligned}$$

where the third equality follows from the fact that W is stochastic. □

A.3 Proof of Proposition 1

Proof. Having established the result in Lemma 2 we can proceed to prove Proposition 1 as follows. Eq. (11) follows from the comparison of Eqs. (9) and (10). We start by defining an M-matrix (Plemmons, 1977):

> **Definition.** *An $N \times N$ matrix C that can be expressed in the form $C = sI - A$, where $a_{ij} \geq 0$ is the (ij)-th element of matrix A, $1 \leq i, j \leq N$ and $s \geq \rho(A)$, the maximum of the moduli of the eigenvalues of A, is called an M-matrix.*

It is straightforward to show that matrix $\mathrm{I}-(\mathrm{I}-\mathrm{D})\mathrm{W}$ is an M-matrix. Define $s = 1$ and $\mathrm{A} = (\mathrm{I}-\mathrm{D})\mathrm{W}$. By construction we know that $a_{ij} \geq 0$, while we showed in the proof of Theorem 1 that $\rho(\mathrm{A}) \leq 1$.

Since $\mathrm{I}-(\mathrm{I}-\mathrm{D})\mathrm{W}$ is an M-matrix, we know that it is inverse-positive (Plemmons, 1977), i.e., each element $\hat{w}_{ij}$ of $\hat{\mathrm{W}} = [\mathrm{I}-(\mathrm{I}-\mathrm{D})\mathrm{W}]^{-1}$ is non-negative. From Lemma 2 we know that $\sum_j \hat{w}_{ij}\alpha_j = 1$ and therefore, given that $0 < \alpha_j \leq 1\ \forall j$, we have that $0 \leq \hat{w}_{ij}\alpha_j < 1\ \forall j$. Therefore, defining $f(x) = x^2$, we have that

$$\sum_j f(\hat{w}_{ij}\alpha_j) \leq f\left(\sum_j \hat{w}_{ij}\alpha_j\right) = 1\,,$$

where the inequality follows from the fact that f is a superadditive function for non-negative real numbers. The only case in which the above expression holds as an equality is when there is no social learning, i.e., when W is a zero matrix (meaning that $\hat{\mathrm{W}}$ is an identity matrix) and $\alpha_j = 1\ \forall j$. □

A.4 Proof of Proposition 2

Proof. Eq. (16) follows from the comparison of Eqs. (14) and (15). We can rewrite Eq. (16) as

$$\sum_j v_j^2 \geq N \Rightarrow \|v\|_2 \geq \sqrt{N} ,$$

and prove it using the Cauchy-Schwarz inequality. In fact, noticing from the results in Lemma 2 that $\sum_j v_j = N$, we can write

$$\begin{aligned} \left(\sum_j v_j^2\right) \cdot N &\geq \left(\sum_j v_j\right)^2 \\ \|v\|_2 \sqrt{N} &\geq N \\ \|v\|_2 &\geq \sqrt{N} . \end{aligned}$$

□

A.5 Proof of Proposition 3

Proof. We start by considering case *a)* in which $1 < \zeta \leq 2$. Recall from Eq. (19) that

$$\operatorname{std}(K_N) \sim \frac{1}{N} \|v_N\|_2 \; .$$

For $N \to \infty$ we have that

$$\|v_N\|_2 = \sqrt{\sum_j v_{j,N}^2} = \sqrt{N \mathrm{E}\left[v_{j,N}^2\right]} \; .$$

Defining $v_{\max,N} \equiv \max_j v_{j,N}$ we can write, since v^2 is a non-negative random variable, that

$$\mathrm{E}\left[v_{j,N}^2\right] = \int_0^{v_{\max,N}} 2x P_N(v_{j,N} > x) dx = 2c_N \int_0^{v_{\max,N}} x L(x) x^{-\zeta} dx \; .$$

Given that $L(x)$ is a slowly varying function such that $\lim_{x\to\infty} L(x) x^{-\epsilon} = 0$ for $\epsilon > 0$, we have that

$$\mathrm{E}\left[v_{j,N}^2\right] = 2c_N \int_0^{v_{\max,N}} x L(x) x^{-\zeta} dx \geq 2c_N \int_0^{v_{\max,N}} x^{1-\zeta-\epsilon} dx \; .$$

Given that $v_{\max,N} \sim N^{1/\zeta}$, we can compute the integral to get

$$\mathrm{E}\left[v_{j,N}^2\right] \geq 2c_N(2-\zeta-\epsilon)^{-1}\left[x^{2-\zeta-\epsilon}\right]_0^{N^{1/\zeta}} = \bar{c}_N N^{\frac{2-\zeta-\epsilon}{\zeta}} ,$$

where $\bar{c}_N \equiv 2c_N(2-\zeta-\epsilon)^{-1}$. This implies that

$$\frac{1}{N}\left\|v_N\right\|_2 = \frac{1}{N}\sqrt{N\mathrm{E}\left[v_{j,N}^2\right]} \geq \bar{c}_N N^{\frac{1-\zeta}{\zeta}-\frac{\epsilon}{2\zeta}} .$$

From the last equation it follows that

$$\mathrm{std}\left(K_N\right) = \Omega(N^{\frac{1-\zeta}{\zeta}-\epsilon'}) ,$$

where $\epsilon' \equiv \frac{\epsilon}{2\zeta}$.

We now consider case *b)*. We start by noticing that when $\zeta > 2$ we have that $\mathrm{E}\left[v_{j,N}^2\right] = V$, where $0 < V < \infty$. Therefore

$$\left\|v_N\right\|_2 = \sqrt{N\mathrm{E}\left[v_{j,N}^2\right]} = \sqrt{NV} ,$$

from which it follows that

$$\mathrm{std}\left(K_N\right) \sim N^{-1/2} .$$

□

References

Acemoglu, D., Carvalho, V. M., Ozdaglar, A., and Tahbaz-Salehi, A. (2012). The network origins of aggregate fluctuations. *Econometrica*, 80(5):1977–2016.

Acemoglu, D., Dahleh, M. A., Lobel, I., and Ozdaglar, A. (2011). Bayesian learning in social networks. *The Review of Economic Studies*, 78(4):1201–1236.

Angeletos, G. M., Iovino, L., and La'O, J. (2016). Real rigidity, nominal rigidity, and the social value of information. *American Economic Review*, 106(1):200–227.

Angeletos, G. M. and La'O, J. (2013). Sentiments. *Econometrica*, 81(2):739–779.

Angeletos, G. M. and Pavan, A. (2004). Transparency of information and coordination in economies with investment complementarities. *American Economic Review*, 94(2):91–98.

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

Bak, P., Chen, K., Scheinkman, J., and Woodford, M. (1993). Aggregate fluctuations from independent sectoral shocks: self-organized criticality in a model of production and inventory dynamics. *Ricerche Economiche*, 47(1):3–30.

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

Bandura, A. and McClelland, D. C. (1977). *Social learning theory.* Prentice-Hall Englewood Cliffs, NJ.

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

Barrdear, J. (2014). Peering into the mist: social learning over an opaque observation network. Working paper, Bank of England.

Benhabib, J., Wang, P., and Wen, Y. (2015). Sentiments and aggregate demand fluctuations. *Econometrica*, 83(2):549–585.

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

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1998). Learning from the behavior of others: Conformity, fads, and informational cascades. *The Journal of Economic Perspectives*, pages 151–170.

Bonacich, P. (1987). Power and centrality: A family of measures. *American journal of sociology*, pages 1170–1182.

Branch, W. A. and McGough, B. (2009). A new keynesian model with heterogeneous expectations. *Journal of Economic Dynamics and Control*, 33(5):1036 – 1051.

Chahrour, R. and Gaballo, G. (2015). On the nature and stability of sentiments. Working paper, Boston College Department of Economics.

Colombo, L., Femminis, G., and Pavan, A. (2014). Information acquisition and welfare. *The Review of Economic Studies*, 81(4):1438–1483.

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

DeMarzo, P. M., Vayanos, D., and Zwiebel, J. (2003). Persuasion bias, social influence, and unidimensional opinions. *The Quarterly Journal of Economics*, 118(3):909 – 928.

Dupor, B. (1999). Aggregation and irrelevance in multi-sector models. *Journal of Monetary Economics*, 43(2):391–409.

Durlauf, S. N. (1993). Nonergodic economic growth. *The Review of Economic Studies*, 60(2):349–366.

Ellison, G. and Fudenberg, D. (1993). Rules of thumb for social learning. *Journal of Political Economy*, 101(4):612–643.

Ellison, G. and Fudenberg, D. (1995). Word-of-mouth communication and social learning. *Quarterly Journal of Economics*, 110(1):93–125.

Fernandez, G. J., Capurro, A. F., and Reboreda, J. C. (2003). Effect of group size on individual and collective vigilance in greater rheas. *Ethology*, 109(5):413–425.

Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica*, 79(3):733–772.

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

Galeotti, A. and Goyal, S. (2010). The law of the few. *American Economic Review*, 100(4):1468–92.

Gilbert, R. J. and Lieberman, M. (1987). Investment and coordination in oligopolistic industries. *The Rand Journal of Economics*, pages 17–33.

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

Goyal, S., Rosenkranz, S., Weitzel, U., and Buskens, V. (2016). Information acquisition and exchange in social networks. *The Economic Journal*. Forthcoming.

Henrich, J. (2015). *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter*. Princeton University Press.

Horvath, M. (1998). Cyclicality and sectoral linkages: Aggregate fluctuations from independent sectoral shocks. *Review of Economic Dynamics*, 1(4):781–808.

Horvath, M. (2000). Sectoral shocks and aggregate fluctuations. *Journal of Monetary Economics*, 45(1):69–106.

Jackson, M. O., Rogers, B. W., and Zenou, Y. (2015). The economic consequences of social network structure. *Available at SSRN*.

Jovanovic, B. (1987). Micro Shocks and Aggregate Risk. *The Quarterly Journal of Economics*, 102(2):395–409.

Lucas, Robert E., J. (2003). Macroeconomic priorities. *American Economic Review*, 93(1):1–14.

Lucas, R. E. (1977). Understanding business cycles. *Carnegie-Rochester Conference Series on Public Policy*, 5(1):7–29.

Meyer, C. D. (2000). *Matrix analysis and applied linear algebra*. Society for industrial and applied mathematics.

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

Plemmons, R. (1977). M-matrix characterizations . i – nonsingular m-matrices. *Linear Algebra and its Applications*, 18(2):175 – 188.

Pulliam, H. R. (1973). On the advantages of flocking. *Journal of Theoretical Biology*, 38(2):419–422.

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