Discussion Paper Series – CRC TR 224

Discussion Paper No. 420
Project B 05

Network Effects on Information Acquisition by DeGroot Updaters

Miguel Risco [1]

November 2023
(*First version: April 2023*)

[1] University of Bonn, Email: miguelriscc@gmail.com

Support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 is gratefully acknowledged.

Collaborative Research Center Transregio 224 - www.crctr224.de
Rheinische Friedrich-Wilhelms-Universität Bonn - Universität Mannheim

Network Effects on Information Acquisition by DeGroot Updaters*

Miguel Risco†

November 14, 2023

Abstract

In today's world, social networks have a significant impact on information processes, shaping individuals' beliefs and influencing their decisions. This paper proposes a model to understand how boundedly rational (DeGroot) individuals behave when seeking information to make decisions in situations where both social communication and private learning take place. The model assumes that information is a local public good, and individuals must decide how much effort to invest in costly information sources to improve their knowledge of the state of the world. Depending on the network structure and agents' positions, some individuals will invest in private learning, while others will free-ride on the social supply of information. The model shows that multiple equilibria can arise, and uniqueness is controlled by the lowest eigenvalue of a matrix determined by the network. The lowest eigenvalue roughly captures how two-sided a network is. Two-sided networks feature multiple equilibria. Under a utilitarian perspective, agents would be more informed than they are in equilibrium. Social welfare would be improved if influential agents increased their information acquisition levels.

Keywords: Information Acquisition, Learning, Public Goods, Network Effects, Information Diffusion, Bounded Rationality.

JEL Codes: D61, D83, D85, H41.

*Support by the German Research Foundation (DFG) through CRC TR 224 (Project B05) is gratefully acknowledged. I thank Sven Rady, Francesc Dilmé, Alexander Frug, Alexander Winter, Simon Block, Justus Preusser, Axel Niemeyer, an anonymous referee and participants at the seminars at University of Bonn and Universitat Pompeu Fabra for their helpful comments and discussions. Declarations of interest: none.

†Bonn Graduate School of Economics, University of Bonn. miguelriscc@gmail.com.

1 Introduction

Information is key to making decisions. Nowadays, social networks have a significant impact on information processes. We discuss various issues with family, friends, and colleagues, affecting their opinions and shaping our own. Random conversations about an upcoming election, the job market, or stock market performances can influence our beliefs. Breakthrough news spreads rapidly, and individuals are constantly updating their opinions. Apart from this social supply of information, individuals can learn privately, such as by searching on the internet or consulting a book. Therefore, it is essential to understand how individuals behave when they seek to acquire information to make decisions in situations where both social communication and private learning take place. To what extent do people exert effort themselves, and to what extent do they rely on others?

In this paper, we propose a model of information acquisition in networks in which individuals are boundedly rational, behaving as mechanical updaters when it comes to learning. With this in hand, they decide how much to invest in a costly information source to improve their knowledge of the state of the world. Mechanical updating here consists of agents merely taking weighted averages of the signals received—the so-called DeGroot updating rule from DeGroot (1974).

Despite considering boundedly-rational agents, we still apply the concept of Nash equilibrium at the stage where they determine their information acquisition. This is done in the spirit of an evolutionary concept of Nash equilibrium. An evolutionary model consists of a large population of boundedly-rational players playing some game repeatedly over time (Mailath, 1998). Evolutionary theory shows that such players eventually learn to play Nash equilibrium,[1] even in the absence of perfect rationality. The crucial assumption is that more successful behaviors become more prevalent due to a combination of imitation and the failure of unsuccessful behaviors.[2] In our model, boundedly-rational agents that update mechanically face the problem of provision of a local public good. The task of gathering information for subsequent decision-making recurs numerous times throughout an individual's life. In the spirit of evolutionary theory, we think of boundedly-rational agents who, despite their cognitive constraints, have learned to reach Nash equilibrium outcomes through their choices.

To provide an intuition for the formal model, consider an agent who wishes to become

1 In particular, the central notion in evolutionary game theory is that of *evolutionary stable strategy*, and theory shows that any symmetric strict Nash equilibrium is indeed an evolutionary stable strategy. See Mailath (1998) or Samuelson (2002) for an overview.

2 This is also discussed by Aumann (1997), asserting that ordinary people, in their daily activities, do not consciously adhere to rationality but evolve "rules of thumb" through evolution. If these rules prove effective, they proliferate and multiply, eventually reaching the equilibrium that strict rationality would have predicted.

more informed about a particular issue. We assume that she has some prior knowledge, and that informative conversations take place in her neighborhood—for example, at the office. There, each colleague exerts a different and fixed influence that shapes the agent's final opinion. Anticipating this, she decides how much effort to spend on private learning. Given that learning tools are similar among neighbors, we assume a positive correlation when it comes to private learning signals. Hence, each agent faces a problem of information acquisition in which information is a local public good. Individuals have to decide how much to invest in private learning, knowing that free social learning will take place later. Depending on the substitutability between information acquired personally and information acquired by others, but also on the neighbors' choices, agents will raise or lower their learning efforts. Free-ride behaviors will arise.

This paper provides three main contributions. First, we analyze and characterize the equilibria arising in the model. Depending on the network structure and their positions, agents will contribute with some learning or completely free-ride. In principle, there are multiple equilibria, and all of them can be calculated. A sufficient condition for equilibrium uniqueness is our second contribution. If this condition does not hold, the equilibria computations run in exponential time. Equilibrium uniqueness is controlled by the lowest eigenvalue of a matrix given by the network. This eigenvalue essentially captures how two-sided the corresponding network is, that is, whether agents can be divided into two sets with many links between them but just a few within. In a game of strategic substitutabilities, when an agent increases her effort, her neighbors decrease theirs in response, so that the neighbors' neighbors have to increase, and so on. When the network is two-sided, these direct effects accumulate and lead to several distinct equilibria. However, if the lowest eigenvalue is sufficiently large, the network will not be two-sided enough for the actions to rebound, and the equilibrium will be unique. Finally, we provide a welfare analysis. From a social (utilitarian) perspective, every agent would be more informed than she is in equilibrium. To satisfy this demand, at least the influential agents (those agents from which the others get the majority of information) have to increase their contribution. If the network is *too unbalanced*, this becomes a burden and the welfare of the influentials decreases. In general, the utilitarian optimum does not Pareto-dominate the equilibrium outcome.

The choice of the updating rule, i.e., how individuals process and incorporate the information received, is a relevant decision when trying to model social learning. One has to decide whether to employ the fully rational Bayesian focus or the naive, boundedly-rational approach, mainly represented by the already mentioned DeGroot rule. Quoting Acemoglu and Ozdaglar (2011), "which type of approach is appropriate is likely to depend on the specific question being investigated". We argue here that the DeGroot updating rule fits best within our context.

Bayesian updating requires an unrealistic cognitive demand for learning in large networks. However, the DeGroot rule provides a convenient alternative, given its simplicity and lack of restrictive requirements. In a simultaneous setting, Bayesian agents who receive Gaussian private signals behave like DeGroot updaters when subject to persuasion bias, as shown by DeMarzo, Vayanos, and Zwiebel (2003). In fact, if the game is one-shot (as it is in this paper), persuasion bias is not even necessary for such a result to hold. Still, their model deviates slightly from standard rational assumptions, as neighbors' signal precision is unknown. The relationship between Bayesian and DeGroot rules, especially for one-shot games, supports our model and is further analyzed in the Appendix. Nonetheless, after the first period, a pure Bayesian (not suffering from persuasion bias) would adjust for the information buried in the network, while DeGrootian agents would not.[3] The literature on networks has widely used the DeGroot rule in different settings. Golub and Jackson (2010) show that under some mild conditions on connectedness and influence, DeGroot agents converge to the belief that would result from the full aggregation of everyone's signal. Both Golub and Jackson (2012), devoted to study homophily, and Acemoglu, Ozdaglar, and ParandehGheibi (2010), which analyzes the tension between the spread of misinformation and information aggregation, also reflect how convenient DeGroot updating is for large networks analysis. However, the major drawback of the rule is that the choice of weights might seem arbitrary, particularly when communication lasts longer than one period. Furthermore, the assumption that everyone is informed at the outset may be too demanding. Banerjee, Breza, Chandrasekhar, and Mobius (2021) adapted the rule to sparse signals to address this issue.

This having been said, the empirical evidence heavily supports DeGroot updating. Various papers confront it against Bayesian learning in an experimental setting, concluding that it approximates better people's information aggregation rules (see Corazzini, Pavesi, Petrovich, and Stanca (2012), Mueller-Frank and Neri (2013), Grimm and Mengel (2020), Chandrasekhar, Larreguy, and Xandri (2020)). Although there is no definitive approach, many recent papers tend to use a boundedly rational model for both sequential and simultaneous settings. For example, Dasaratha and He (2020) assume that agents neglect redundancies of information and then aggregate heuristically, and Mueller-Frank and Neri (2021) consider a large class of boundedly rational or quasi-Bayesian rules, respectively.

Although modelling learning through a mechanical updating rule is overly simplistic, it allows us to isolate the network effects, which is the primary concern of this paper. Furthermore, we argue that assuming exogenous and fixed weights reflects human behavior. The influence that our neighbors exert on a concrete topic is almost predetermined. A wide range of factors such as past interactions, trustworthiness, and expected level of knowledge defines an influence level before communication occurs. Similarly, it seems

3 See Molavi, Tahbaz-Salehi, and Jadbabaie (2018) for an axiomatic foundation of the DeGroot rule under imperfect recall.

sensible that agents can endogenously set the influence of their own private learning on their views: the more time devoted to researching, the more reliable the agent perceives it to be. Thus, the expenditure of costly attention will reduce player-specific noise.

Galeotti and Goyal (2010) is a key paper in the literature on information acquisition in networks. In this paper, network-placed agents strategically select their links to access the information held by their neighbors. Every equilibrium displays the so-called "law of the few": the majority of individuals tend to get most of their information from a tiny subset of the group, the influentials. Our model shows that this result holds true for networks where a subset of agents, the populars, has a significantly higher weight than the rest, such as the core-periphery network. In such networks, popular agents become influential and acquire most of the information, while the others free-ride. This finding contrasts with Banerjee, Breza, Chandrasekhar, and Mobius (2021), where the sparse-signals structure indicates that being popular alone is insufficient for being influential. However, two assumptions in Galeotti and Goyal (2010) differ from our model: links are endogenous, allowing an agent to reach any other individual in a potentially large network, and homogeneous, meaning perfect substitutability. Network effects on information acquisition have also been analyzed from a Bayesian perspective. In Myatt and Wallace (2019), rational agents acquire information about the state of the world from sources that provide noisy signals. Paying costly attention reduces noise, and signals are possibly correlated across players, similar to our model. However, incentives differ as agents not only want to match the state of the world but also care about coordination asymmetrically. Furthermore, there is no communication stage. The player's centrality (in the sense of Bonacich) and correlations determine information acquisition, but centrality entails less expenditure, in contrast to Galeotti and Goyal (2010) and our paper. Finally, Denti (2017) introduces the concept of entropy reduction to study how players endogenously acquire costly information to decrease their uncertainty about fundamentals. Network effects induce externalities in information acquisition and are a source of multiple equilibria.

Regarding equilibrium analysis, our work closely follows that of Bramoullé, Kranton, and D'Amours (2014). Following previous attempts in the literature to characterize conditions for equilibrium in linear games of strategic complements (cf. Ballester, Calvó-Armengol, and Zenou (2006)) and strategic substitutes (especially in public goods, cf. Bramoullé, Kranton, et al. (2007)), the authors showed that equilibrium uniqueness and stability depend on the lowest eigenvalue of the network matrix. This is dependent on the two-sidedness of the network. Although our paper differs in setting and motivation, the best response function derived from our model is similar to that of Bramoullé, Kranton, and D'Amours (2014). Consequently, the result regarding the lowest eigenvalue characterizing equilibrium uniqueness is also similar. However, their model assumes that agents' contributions are reciprocal and weighted equally, which differs from our assumptions. This has two consequences. First, the potential theory introduced in Monderer and Shapley (1996), on which Bramoullé et al. base their results, cannot be applied here;

second, a wider range of networks can be analyzed. Nevertheless, if we restrict our setting to symmetric, homogeneous networks, an almost equivalent condition arises. Finally, our paper is also related to Bramoullé, Kranton, et al. (2007) model of pure public goods in exogenous networks, where again all contributions are weighted equally and there is perfect substitutability. In that model, the authors find that multiple equilibria typically exist, and there is always one in which some individuals contribute while others free-ride. This equilibrium is typically unique.

The rest of the paper is organized as follows. Section 2 describes and analyzes our model, Section 3 studies the equilibria, provides a uniqueness condition and presents some examples, and Section 4 analyzes the model from a social planner perspective. Section 5 briefly introduces a dynamic version of the model, and Section 6 concludes.

2 Model

We consider a finite set of n agents interacting via a social network represented by an $n \times n$ matrix $\mathbf{G} = (g_{ij})$, which is predetermined and stochastic: the entries in each row are non-negative and sum to one. Interactions need not be symmetric or two-sided, so in general $g_{ij} \neq g_{ji}$ and $g_{ij} > 0$ does not imply $g_{ji} > 0$.

Each agent holds a private signal s_i about a common underlying state of the world $\mu \in \mathbb{R}$, drawn independently from a normal distribution with mean μ and variance $\sigma^2 > 0$. There are two learning resources available to improve this signal, presented in the order in which they become accessible to the agent: active private learning from a more informative but costly source, and social learning from neighbors. The first resource involves drawing a signal $\mathcal{I}_i$ from a normal distribution with mean μ and variance $\tilde{\sigma}^2 < \sigma^2$, while the second resource involves aggregating the signals of the agent's direct neighbors in the network.

Both types of learning take the form of DeGroot updating of signals, following DeGroot (1974). Agents take a weighted average of their signals, i.e., they aggregate several indicators in just one. In the case of private learning, agent i decides the weights in the convex combination between s_i and $\mathcal{I}_i$. The costly signal $\mathcal{I}_i$ receives weight $x_i \in [0, 1]$ at linear cost $x_i c$ with $c > 0$. Costly signals are positively correlated across agents, $\text{Cov}(\mathcal{I}_i, \mathcal{I}_j) = \alpha > 0$ for all i, j. The original private signals are independent across agents and independent of all costly signals. Regarding social learning, agent i takes the weighted average of her neighbors' signals and her own. Weights are exogenously[4] given by the network matrix, and they represent influence or trust: agent i listens to agent j precisely at *intensity* g_{ij}.

4 Rational learners might adjust the weights based on neighbors' information levels, as discussed in Galeotti and Goyal (2010). However, in this case, we want to focus on situations where weights are pre-determined for a naïve learner due to past interactions, influence, or reputation, and cannot be modified.

The mechanical updating process described can be viewed as active learning with attention costs for boundedly rational agents. In addition to normal signals, it can also be interpreted from a Bayesian perspective, as demonstrated in DeMarzo, Vayanos, and Zwiebel (2003). Agents assign subjective precisions π_{ij} to each other, attempting to estimate the true precision of their signals. If the signals are independent and normal, Bayesian updating is equivalent to DeGroot updating, with weights given by $\frac{\pi_{ij}}{\sum_{j=1}^{n} \pi_{ij}}$ for social learning.[5] A similar argument applies to the active learning process; see the Appendix for a motivation of the present framework in terms of quasi-Bayesian updating as defined in DeMarzo, Vayanos, and Zwiebel (2003).

In the following, we use the term "beliefs" to refer to the most recently updated signal an agent holds about μ. A precise description of the learning process is as follows: The agent receives $s_i \sim \mathcal{N}(\mu, \sigma^2)$ and decides how much to spend on learning $\mathcal{I}_i \sim \mathcal{N}(\mu, \tilde{\sigma}^2)$. Once x_i is selected, the belief becomes $p_i = (1-x_i)s_i + x_i\mathcal{I}_i$ at cost $x_i c$. Finally, the social communication stage yields beliefs

$$\sum_{j=1}^{n} g_{ij} p_j = \sum_{j=1}^{n} g_{ij} \left((1-x_j)s_j + x_j \mathcal{I}_j \right).$$

Note that if i and j are not neighbors, $g_{ij} = 0$, so summing over i's neighbors is equivalent to summing over all n individuals. At this point, only one communication stage is assumed, but considering t stages would imply the substitution of $\mathbf{G}$ by $\mathbf{G}^t$, as shown in Section 5.

Agent i aims to obtain the most precise belief about μ at minimum cost, as deviations are penalized through a quadratic loss function. This is specified in the payoff function

$$-\left(\mu - \sum_{j=1}^{n} g_{ij} p_j\right)^2 - x_i c = -\left(\mu - \sum_{j=1}^{n} g_{ij}((1-x_j)s_j + x_j \mathcal{I}_j)\right)^2 - x_i c.$$

Although agent i is a naive, mechanical learner, we assume, based on evolutionary theory, that she is capable of reaching Nash equilibrium outcomes. Specifically, deciding how much to contribute to a public good is a typical example of a process in which boundedly-rational agents evolve toward Nash equilibrium outcomes in the long run (Mailath, 1998). Hence, we allow agent i to form expectations and best respond to others' choices, *as if* she were rational at this stage. She chooses the amount x_i that maximizes her expected utility:

$$\max_{x_i \in [0,1]} \left\{ \mathbb{E}\left[-\left(\mu - \sum_{j=1}^{n} g_{ij}((1-x_j)s_j + x_j \mathcal{I}_j)\right)^2 \right] - x_i c \right\}. \tag{1}$$

5 If $g_{ij} = 0$, then $\pi_{ij} = 0$.

3 Equilibrium

The equilibrium concept used in this model is Nash equilibrium, where each agent i chooses her information level x_i by best responding to others' equilibrium choices. It is important to note that $\mathbb{E}[s_i] = \mathbb{E}[\mathcal{I}_i] = \mu$ for all i. Additionally, every pair of signals is independent except for $\mathcal{I}_i$ and $\mathcal{I}_j$. As a result, $\mathbb{E}\left[\sum_{j=1}^n g_{ij}(x_j\mathcal{I}_j + (1-x_j)s_j)\right] = \mu$, while $\mathrm{Var}(x_j\mathcal{I}_j+(1-x_j)s_j) = x_j^2\tilde{\sigma}^2+(1-x_j)^2\sigma^2$ and $\mathrm{Cov}(x_j\mathcal{I}_j+(1-x_j)s_j, x_k\mathcal{I}_k+(1-x_k)s_k) = \alpha x_j x_k$. These equalities, along with the payoff equation, imply that only second moments matter. In fact,

$$\mathbb{E}\left[-\left(\mu - \sum_{j=1}^n g_{ij}(x_j\mathcal{I}_j + (1-x_j)s_j)\right)^2\right] = -\mathrm{Var}\left[\sum_{j=1}^n g_{ij}(x_j\mathcal{I}_j + (1-x_j)s_j)\right].$$

Using that for any sequence of random variables $\{\tilde{X}_j\}_{j=1}^n$ it holds that $\mathrm{Var}(\sum_{j=1}^n \tilde{X}_j) = \sum_{j=1}^n \mathrm{Var}(\tilde{X}_j) + 2\sum_{j=1}^n\sum_{k=1}^{j-1}\mathrm{Cov}(\tilde{X}_j,\tilde{X}_k)$, the maximization problem for agent i can be simplified as follows:

$$\max_{x_i\in[0,1]}\left\{-\tilde{\sigma}^2\sum_{j=1}^n g_{ij}^2x_j^2 - \sigma^2\sum_{j=1}^n g_{ij}^2(1-x_j)^2 - 2\alpha\sum_{j=1}^n\sum_{k=1}^{j-1} g_{ij}g_{ik}x_jx_k - cx_i\right\}. \tag{2}$$

The objective is strongly concave in the choice variable. The first order condition for an interior solution yields

$$x_i = \frac{2\sigma^2 - c/g_{ii}^2}{2(\tilde{\sigma}^2+\sigma^2)} - \frac{\alpha}{g_{ii}(\tilde{\sigma}^2+\sigma^2)}\sum_{j\neq i} g_{ij}x_j.$$

Note that this expression is bounded above by 1 but could be negative. As $x_i \in [0,1]$ by assumption, the optimal choice of active learning for agent i given others' choices x_{-i} is

$$x_i^* = \max\left\{0, \frac{2\sigma^2 - c/g_{ii}^2}{2(\tilde{\sigma}^2+\sigma^2)} - \frac{\alpha}{g_{ii}(\tilde{\sigma}^2+\sigma^2)}\sum_{j\neq i} g_{ij}x_j\right\}.$$

This best response function is similar to the one obtained when solving a maximization problem in a local public goods setting. Games of negative externalities or Cournot competition also yield similar forms. The only difference is that here, δ is divided by g_{ii}, a parameter that varies across agents. In the other cases, the substitutability factor is the same for all agents.

Note that only the weighted out-degree matters for information acquisition, but not the weighted in-degree.[6] In other words, agents care about who they are listening to (the g_{ij}s), but not who listens to them (the g_{ji}s). Furthermore, if $g_{ii} = 0$, then $x_i^* = 0$ trivially,

6 The out-degree of agent i is the total weight of links directed away from her. The in-degree is the total weight of links directed to her.

as active learning is a waste of resources for someone who does not assign positive weight to herself. Therefore, without loss of generality we can assume $g_{ii} > 0$. By setting

$$\bar{x}_i = \frac{2\sigma^2 - c/g_{ii}^2}{2(\tilde{\sigma}^2 + \sigma^2)},$$

and

$$\delta = \frac{\alpha}{(\tilde{\sigma}^2 + \sigma^2)},$$

we obtain

$$x_i^* = \max\left\{0, \bar{x}_i - \frac{\delta}{g_{ii}} \sum_{j \neq i} g_{ij} x_j\right\}.$$

Here, information refers to individuals' costly learnt signals and is a local public good. Each agent benefits from others' private learning via network communication. The quotient $\frac{\delta}{g_{ii}}$ scales the benefit and indicates the substitutability between an agent's and her neighbors' information. Agent i seeks to reach at least the *information target* $\bar{x}_i$ through a combination of her own information and her neighbors'. If the weighted contributions of the others are enough ($\frac{\delta}{g_{ii}} \sum_{j \neq i} g_{ij} x_j > \bar{x}_i$), then she will not spend on private learning, and $x_i^* = 0$. If not, she will make up the difference, and $x_i^* > 0$.

Let us analyze the scale factor $\frac{\delta}{g_{ii}}$, which measures the substitutability between the information purchased by an agent and her neighbors. On the one hand, $\frac{1}{g_{ii}}$ reflects how important others' contributions are to the particular agent i. If g_{ii} is small, almost all attention is paid to the neighbors, so their information matters considerably. In contrast, if g_{ii} is close to one, agent i essentially listens to herself. On the other hand, $\delta = \frac{\alpha}{\sigma^2 + \tilde{\sigma}^2}$ reflects the quality of the neighbors' information. The parameter α indicates how much information one can extract from others. Consequently, the higher α, the less information is purchased. The sum $\sigma^2 + \tilde{\sigma}^2$ expresses the overall level of uncertainty. If it grows, the incentives for an agent to increase her information level also grow. Note also that $\delta \in \left[0, \frac{1}{2}\right]$ by the Cauchy-Schwarz inequality.[7] Therefore, for any $g_{ii} \geq 1/2$, the scale factor is always lower than one. This is not surprising: if an agent listens to herself more than to others, the information that she acquires matters more.

The information target $\bar{x}_i$ indicates how well-informed each agent aims to be. The more precise $\mathcal{I}_i$ is in expectation terms—the lower $\tilde{\sigma}^2$—, the higher $\bar{x}_i$ the agent wants to attain. Additionally, the degree of an agent's own attention, represented by g_{ii}, matters: acquiring information through costly learning is more profitable if the agent puts a higher weight on herself when updating. An increase in costs makes information acquisition less

7 In fact, this inequality implies $0 \leq \alpha \leq \tilde{\sigma}^2$, and hence

$$0 \leq \delta \leq \frac{\tilde{\sigma}^2}{\sigma^2 + \tilde{\sigma}^2} = \frac{1}{2}.$$

attractive. It is worth noting that $\bar{x}_i = 1$ only if $\tilde{\sigma}^2 = 0$ (i.e., $\mathcal{I}_i = \mu$ and it is a perfect signal) and $c = 0$. In all other cases, a convex combination of s_i and $\mathcal{I}_i$ is preferred. It is more convenient for the agent to have two independent signals, even if one is much more informative than the other, than just one. Therefore, she does not want to get rid of s_i entirely and sets $x_i^* < 1$.

It has been shown that $x_i^* \in [0, 1]$. The combined best response function of the players maps $[0, 1]^n$ to itself and is continuous. Brouwer's fixed point theorem guarantees the existence of an equilibrium.

Proposition 3.1. *The game of information acquisition by DeGroot updaters has at least one Nash equilibrium.*

We should pay special attention to the limiting case of uncorrelated costly signals, i.e., $\alpha = 0$. Since the correlation between signals is zero, agents cannot extract any information from each other. The equilibrium analysis is then trivial: as $\delta > 0$, each agent chooses the information level

$$x_i^* = \max\{0, \bar{x}_i\}.$$

Since $\bar{x}_i \leq 1$, x_i^* is well-defined. Moreover, $\bar{x}_i > 0$ if and only if $g_{ii} > \sqrt{\frac{c}{2\sigma^2}}$. This is the target for active learning: every individual that weighs themselves more than $\sqrt{\frac{c}{2\sigma^2}}$ will choose a positive x_i^*, independently of the network. Due to the lack of strategic substitutability, equilibrium is unique.

3.1 Examples

With the existence of equilibrium proved, and before further general analysis, some prominent networks and their equilibria are discussed as illustrations.

The first class we consider is the **k-regular graphs with homogeneous weights**. This class consists of structures with n agents, each having k neighbors. All connections have identical weight, such that $g_{ij} = \frac{1}{k}$ for all i and j. Such a network could represent a small community where each member listens to everyone else, and influences are homogeneous. Complete networks in which every agent has the same degree are a subset of regular graphs. Other examples of symmetric structures represented by regular graphs are big societies of n individuals who cluster in small k-neighborhoods and the circle. The unique equilibrium of the k-regular graph with homogeneous weights is given by

$$x_i^* = \frac{2\sigma^2 - ck^2}{2(\sigma^2 + \tilde{\sigma}^2)(-\delta + \delta\, k + 1)} \quad \text{if } c \leq \frac{2\sigma^2}{k^2}; \quad x_i^* = 0 \text{ otherwise.}$$

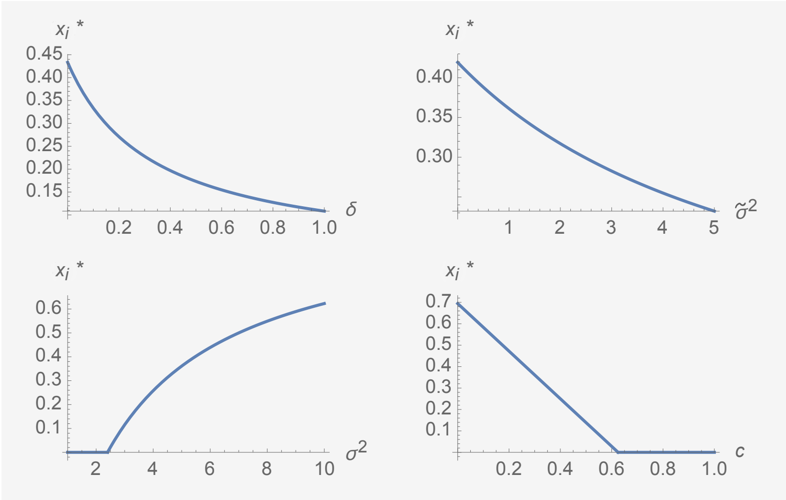


Figure 3: Comparative statics for the 4-regular graph.

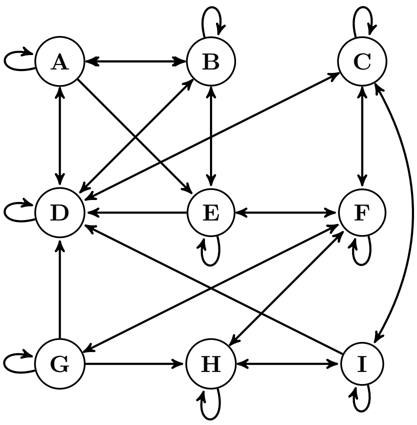


Figure 1: 4-regular graph.

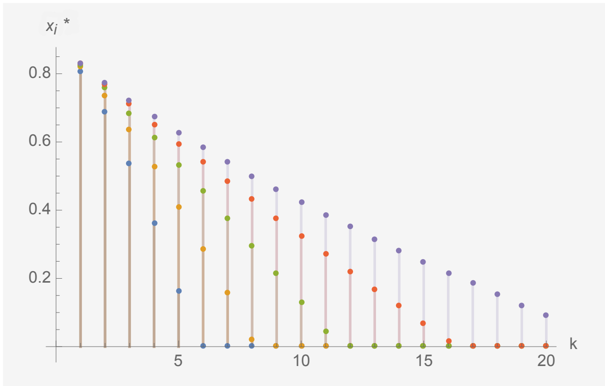


Figure 2: Information acquisition, k-regular graph.

As the number of neighbors grows, incentives to acquire information decrease. The more signals an agent listens to, the better, and active learning loses importance, as shown in Figure 2. The parameters are set to $\sigma^2 = 5$, $\tilde{\sigma}^2 = 1$, and $\delta = 0.07$. Each color corresponds to a different cost of $\mathcal{I}_i$, ranging from $c = 0.3$ to $c = 0$. Note that $x_i^* = 0$ as soon as $c \geq \frac{10}{k^2}$. The non-symmetric example displayed in Figure 1, which features nine agents of out-degree four, also has the above equilibrium. This is implied by the fact that only the acquisition choices of those whom the individual listens to matter. In other words, only the out-degree matters. Differences between networks with the same out-degree for all agents but different in-degrees will appear in the socially optimal allocation, where the in-degree also conditions behavior. This will be shown in Section 4. Comparative statics are presented in Figure 3 for a regular graph with four neighbors and weights $\frac{1}{4}$. The cost is $c = 0.3$ and the rest of the parameters are as above.

The second class is that of **stars**, where one prominent agent (the hub) is connected to every other agent (the spokes). The spokes, in turn, are connected only to the hub and themselves. A department in a firm, with a supervisor and some employees, is a leading example. An auction with an auctioneer in the center could be another example. Assume that the hub puts the same weight on everyone, and the spokes put weight ε on the hub. Therefore, the information levels in equilibrium for a society of n agents are given by the following equations:

$$x_H^* = \frac{2\sigma^2(1-\varepsilon)^2((n-1)\delta-1) - c(\delta(n-1)-(1-\varepsilon)^2n^2)}{2(\sigma^2+\tilde{\sigma}^2)(1-\varepsilon)((n-1)\varepsilon\delta^2+\varepsilon-1)} \text{ if } c \leq \frac{2\sigma^2(1-\varepsilon)^2((n-1)\delta-1)}{(\delta(n-1)-(1-\varepsilon)^2n^2)};$$

$x_H^* = 0$ otherwise,

$$x_S^* = \frac{2\sigma^2(1-\varepsilon)(1-\varepsilon-\delta\varepsilon) - c(1+\delta(\varepsilon-1)\varepsilon n^2)}{2(\sigma^2+\tilde{\sigma}^2)(1-\varepsilon)((n-1)\varepsilon\delta^2+\varepsilon-1)} \text{ if } c \leq \frac{2\sigma^2(1-\varepsilon)(1-\varepsilon-\delta\varepsilon)}{(1+\delta(\varepsilon-1)\varepsilon n^2)};$$

$x_S^* = 0$ otherwise.

The star graph for five agents is shown in Figure 4, where node A represents the hub (H) and nodes B, C, D, and E are the spokes (S). The blue links have a weight of $\frac{1}{5}$, while the bold links have a weight of ε. The hub shares her attention equally, while the spokes put weight ε on her. The information acquisition levels as a function of ε are also illustrated below, using the same parameter values as before: $c = 0.3$, $\sigma^2 = 5$, $\tilde{\sigma}^2 = 1$, and $\delta = 0.07$. When ε is small, the spokes need to invest significantly in learning. However, as ε grows, the hub's signal gains importance, and active learning becomes less valuable for the spokes. The hub responds to this behavior by adjusting her learning efforts. When ε is small, she can rely on the spokes to aggregate some signals, so only a modest amount of active learning is necessary. However, when ε grows large, the hub invests significantly more in learning to compensate for the drop in the spokes' contribution. Similarly, in a department of a firm, the supervisor reacts to employees' expertise, and the employees invest in learning only when it is useful. If they only follow the supervisor's orders, there are no incentives for them to learn independently.

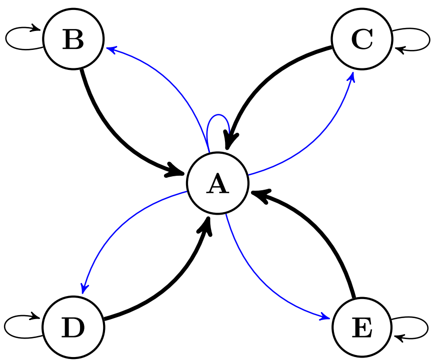


Figure 4: Star with five agents.

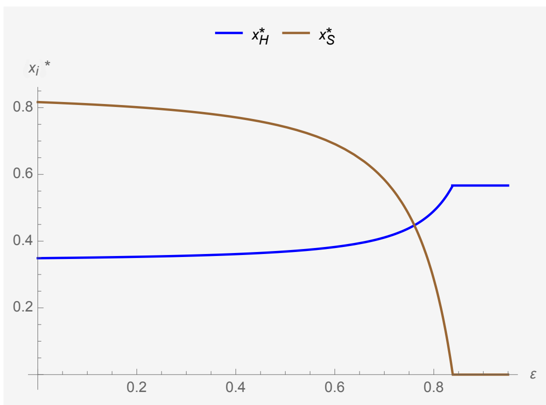


Figure 5: Information acquisition, star.

It is worth noting that the complete graph, in which each agent weights herself as ε and the other $n-1$ agents as $\frac{1-\varepsilon}{n-1}$, yields the same levels of information acquisition as the star with an own weight of ε.

We now move on to the class of **core-periphery** networks, which are characterized by a "dense, cohesive core and a sparse, unconnected periphery", as described in Borgatti and Everett (2000). Many relevant economic networks exhibit this structure, such as the lending behavior of banks (Fricke and Lux, 2015) or international trade networks (Fagiolo, Reyes, and Schiavo, 2010). Another example is the structure arising in Galeotti and Goyal (2010): the "few" constitute the core while the rest of the network (the periphery) free-rides on them. A particular case in which the core is formed by three individuals who share their attention equally and three periphery agents who listen to one core agent each is shown in Figure 6. The periphery agents put almost all of their weight (in this case, $\frac{9}{10}$) on the core agents, which causes their acquisition levels to rapidly drop to zero as δ increases. As soon as there is a minimal amount of substitutability, the periphery agents stop purchasing. On the other hand, the core agents acquire abundant information, as the core acts as a k-regular network for them and it is independent from the periphery. The results for k-regular networks hold and hence, the larger the core, the less information its agents acquire. The acquisition choices are given by:

$$x_C^* = \frac{2\sigma^2 - 9c}{2(\sigma^2 + \tilde{\sigma}^2)(2\delta + 1)} \text{ if } c \leq \frac{2\sigma^2}{9}; \quad x_C^* = 0 \text{ otherwise},$$
$$x_P^* = \max\left\{0, \frac{2\sigma^2 - 100c}{2(\sigma^2 + \tilde{\sigma}^2)} - \frac{9\delta(2\sigma^2 - 9c)}{2(\sigma^2 + \tilde{\sigma}^2)(2\delta + 1)}\right\}.$$

The cost c is set to $c = 0.06$ in this example to illustrate the decay in the periphery agents' acquisition levels. A value of $c = 0.3$ as above would lead to no acquisition even if $\delta = 0$, as periphery agents barely weight themselves.

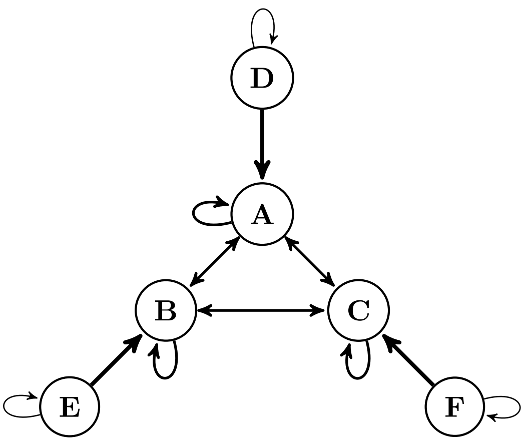


Figure 6: Core-periphery network.

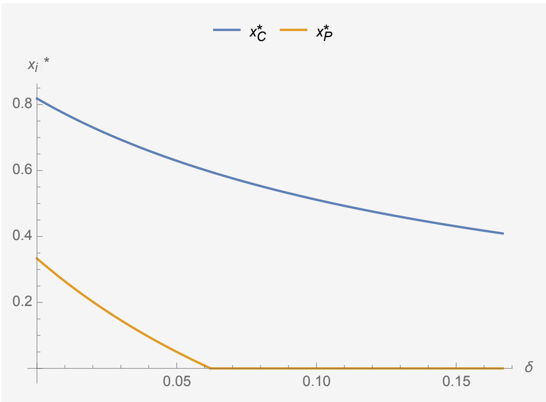


Figure 7: Information acquisition as a function of δ.

Next we discuss the **criminal network** from Ballester, Calvó-Armengol, and Zenou (2006). The authors use the network from Figure 8 to highlight the fact that influence is not necessarily equivalent to the number of connections (degree). They identify the key criminal as agent A, who, when removed, leads to the highest aggregate reduction in crime, despite not having the highest degree. Agent A plays a crucial role in bridging two fully interconnected communities of five criminals each. In terms of information acquisition, Figure 9 shows the active learning choices for this network. Assuming equal shares of attention (i.e. each agent listens equally to every neighbor), there are three different kinds of individuals: agent A, agents B, F, G and K (referred to as the B-class) and agents C, D, E, I, H and J (C-class). Apart from listening to themselves, agent A listens to the four B-class agents, B-class agents listen to agent A, one B-class individual and three C-class individuals, and C-class agents listen to two B-class individuals and two C-class individuals. Interestingly, the most influential agent for Ballester et al. (agent A) is also the most informed. The B-class agents tend to free-ride on the rest. The best reply functions are:

$$
\begin{aligned}
x_A^* &= \max\left\{0, \bar{x}_A - 4\delta x_B\right\}, \\
x_B^* &= \max\left\{0, \bar{x}_B - \delta(x_A + x_B + 3x_C)\right\}, \\
x_C^* &= \max\left\{0, \bar{x}_C - 2\delta(x_B + x_C)\right\}.
\end{aligned}
$$

Classes A and C respond to class B's choice, which is small in comparison. Information acquisition choices are shown in Figure 9 as a function of δ. The parameters are set as in the core-periphery example. Both A-class and C-class agents listen to five individuals each, so $\bar{x}_A = \bar{x}_C$. However, as B-class agents listen to six neighbors and weigh their own signal less, it holds that $\bar{x}_B < \bar{x}_A$. In the extreme case of $\delta = 0$, A-class agents and C-class agents purchase the same quantity, slightly more than B-class individuals. As soon as δ increases, B-class agents take advantage of substitutability and free-ride on A and the C-class agents. Agent A has only B-class neighbors, so although she extracts information from them, she has to make up the difference. In contrast, C-class agents have some C-class neighbors, so the free-riding behavior of B-class agents does not affect them as severely. Figure 9 shows that x_A^* is higher than x_C^* for all $\delta > 0$.

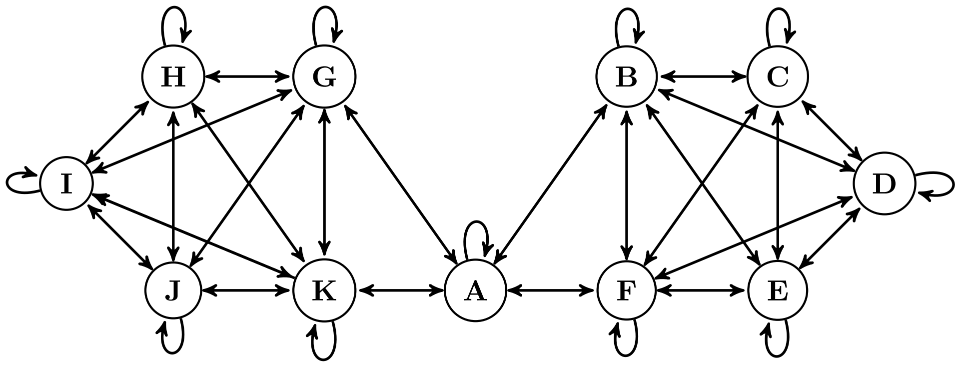


Figure 8: Criminal network.

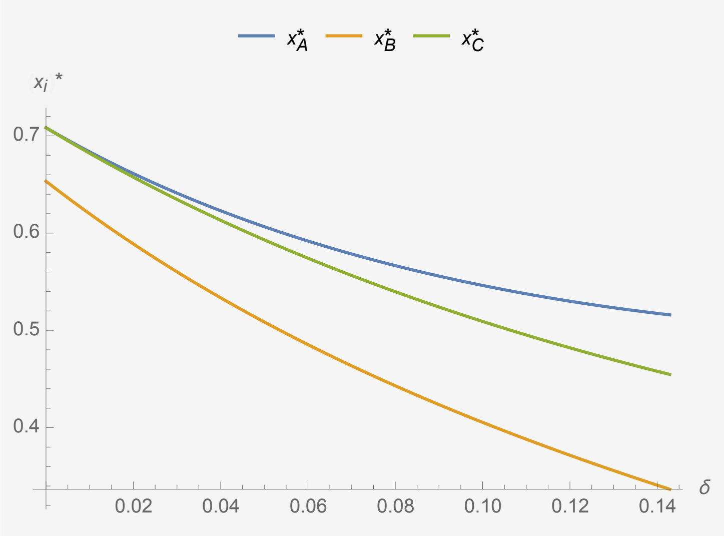


Figure 9: Information acquisition for the criminal network.

3.2 Equilibrium characterization

Let us divide the agents into two groups: *active* (A) agents, who are active learners ($x_i^* > 0$), and *passive* (P) agents. An equilibrium in which all agents belong to A is known as a *distributed equilibrium*, as effort is distributed among all agents. In contrast, a *specialized equilibrium* is such that only a few individuals (the specialists) learn, while the others free-ride.

This part mainly follows Bramoullé, Kranton, and D'Amours (2014). Without loss of generality, we can reorder the agents such that the first r are active and the last $n-r$ are passive. As $x_j = 0$ for all $j \in P$, for any individual i, we have $\sum_{j \neq i} g_{ij} x_j = \sum_{j \in A \setminus \{i\}} g_{ij} x_j$. Thus, for $i \in 1, ..., r$, an equilibrium requires that:

$$x_i^* = \bar{x}_i - \frac{\delta}{g_{ii}} \sum_{j \in A \setminus \{i\}} g_{ij} x_j^* > 0.$$

For $i \in \{r+1, ..., n\}$, an equilibrium requires that:

$$\bar{x}_i - \frac{\delta}{g_{ii}} \sum_{j \in A \setminus \{i\}} g_{ij} x_j^* \leq 0.$$

Let $\bar{\mathbf{x}}^A = (\bar{x}_1, ..., \bar{x}_r)$ and $\bar{\mathbf{x}}^P = (\bar{x}_{r+1}, ..., \bar{x}_n)$. The diagonal of a matrix A is denoted by d_A. Let $\mathbf{G}^A$ be the $r \times r$ minor corresponding to the active agents of the network, while $\mathbf{G}^P$ is the $(n-r-1) \times (n-r-1)$ minor of $\mathbf{G}$ corresponding to the passive agents. The $(n-r-1) \times r$ minor $\mathbf{G}^{P,A}$ of $\mathbf{G}$ is given by (g_{ij}) where $i \in P$ and $j \in A$. Rearranging the above expressions, we obtain the following result:

Proposition 3.2. *The profile of information levels* $\boldsymbol{x} = (x_1^*, ..., x_n^*) = (\boldsymbol{x}^A, 0)$ *with* $\boldsymbol{x}^A =$

$(x_1^*, ..., x_r^*) \in (0, 1]^r$ *constitutes an equilibrium if and only if*

$$\begin{cases} d_{\boldsymbol{G}^A} \bar{\boldsymbol{x}}^A & = & \left[(1-\delta) d_{\boldsymbol{G}^A} + \delta \, \boldsymbol{G}^A\right] \boldsymbol{x}^A, \\ d_{\boldsymbol{G}^P} \bar{\boldsymbol{x}}^P & \leq & \delta \, \boldsymbol{G}^{P,A} \boldsymbol{x}^A. \end{cases} \tag{3}$$

Note that given n agents, there are 2^n potential partitions. Obtaining all possible equilibria requires solving the system (3) for each partition. This can be done in two steps:

(i) First, solve for $\mathbf{x}^A$ in $d_{\mathbf{G}^A} \bar{\mathbf{x}}^A = \left[(1-\delta) d_{\mathbf{G}^A} + \delta \mathbf{G}^A\right] \mathbf{x}^A$. The solution is unique if and only if $\det[(1-\delta) d_{\mathbf{G}^A} + \delta \mathbf{G}^A] \neq 0$.

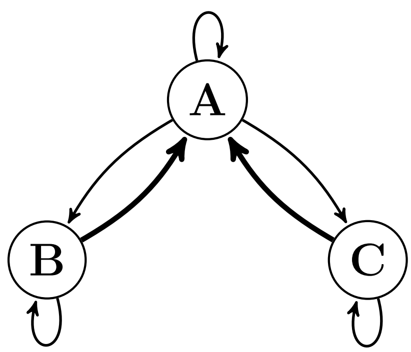
(ii) Then, check whether all components of x^A are strictly positive and $d_{\mathbf{G}^P} \bar{\mathbf{x}}^P \leq \delta \mathbf{G}^{P,A} \mathbf{x}^A$.

If the diagonal elements of $\mathbf{G}$ are identical, i.e., $g_{ii} = g_{jj}$ for all i, j, the number of equilibria is weakly lower than 2^n and can be computed in exponential time. In this case, the condition $\det[(1-\delta) d_{\mathbf{G}^A} + \delta \mathbf{G}^A] = 0$ simplifies to $\det\left[-\frac{(\delta-1) g_{ii}}{\delta} \mathbf{Id} + \mathbf{G}^A\right] = 0$, which holds if and only if $\mathbf{G}^A$ has an eigenvalue $\lambda = \frac{(\delta-1) g_{ii}}{\delta}$. Consequently, for almost all δ, the equation has a unique solution. While Bramoullé, Kranton, and D'Amours (2014) assume not only $d_{\mathbf{G}} = d_{\mathbf{Id}}$ but also matrix symmetry, we have shown that these assumptions are not necessary to obtain an explicit expression for equilibria.

3.3 Equilibrium uniqueness

In general, there might be multiple equilibria in this model. We present two examples.

The first example is a three-agent network that is incomplete and can also be visualized as a star. The weights of the connections between the agents are represented by the matrix $\mathbf{G}$. Figure 10 shows the graph corresponding to this network, with thicker arrows indicating larger weights.



$$\mathbf{G} = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 2/3 & 1/3 & 0 \\ 2/3 & 0 & 1/3 \end{pmatrix}$$

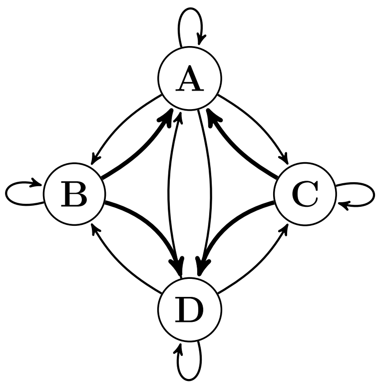
Figure 10: Incomplete three-agent network.

Since $g_{AA} = g_{BB} = g_{CC}$, it follows that $\bar{x}_A = \bar{x}_B = \bar{x}_C = \bar{x}$. Assuming $\delta = \frac{1}{2}$, the best reply functions become:

$$
\begin{aligned}
x_A &= \max\{0, \bar{x} - \frac{x_B + x_C}{2}\},\\
x_B &= \max\{0, \bar{x} - x_A\},\\
x_C &= \max\{0, \bar{x} - x_A\}.
\end{aligned}
$$

There are two distributed equilibria: $(x_A^*, x_B^*, x_C^*) = (\frac{\bar{x}}{2}, \frac{\bar{x}}{2}, \frac{\bar{x}}{2})$ and $(x_A^*, x_B^*, x_C^*) = (\frac{\bar{x}}{3}, \frac{2\bar{x}}{3}, \frac{2\bar{x}}{3})$. Additionally, there exist specialized equilibria where either agent A or both agents B and C purchase $\bar{x}$ while the others free-ride. Another similar example holds for $\delta = \frac{1}{k}$ and a star with k agents, demonstrating that the multiplicity of equilibria does not depend on the extreme assumption that $\delta = \frac{1}{2}$ (which is extreme in the sense that it implies $\tilde{\sigma}^2 = 0$).

The second example is a four-agents *eye*, shown in Figure 11. Weights are given by the matrix $\mathbf{G}'$.



$$
\mathbf{G}' = \begin{pmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 3/8 & 1/4 & 0 & 3/8 \\ 3/8 & 0 & 1/4 & 3/8 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix}
$$

Figure 11: Four-agents *eye*.

Once again, we assume $\delta = \frac{1}{2}$. Since $\bar{x}_i = \bar{x}$ for all i, the best replies are as follows:

$$
\begin{aligned}
x_A &= \max\left\{0, \bar{x} - \frac{x_B + x_C + x_D}{2}\right\},\\
x_B &= \max\left\{0, \bar{x} - \frac{3(x_A + x_D)}{4}\right\},\\
x_C &= \max\left\{0, \bar{x} - \frac{3(x_A + x_D)}{4}\right\},\\
x_D &= \max\left\{0, \bar{x} - \frac{x_A + x_B + x_C}{2}\right\}.
\end{aligned}
$$

There are two specialized equilibria: $\left(\frac{2}{3}\bar{x}, 0, 0, \frac{2}{3}\bar{x}\right)$ and $(0, \bar{x}, \bar{x}, 0)$. The rough idea behind multiplicity is that agents can be divided into two distinct groups so that active learning contributions vary between them. When one group learns more, the other decreases its effort, and vice versa. We will discuss this in detail in Subsection 3.3.2.

Next, we seek a structural condition on the network that guarantees uniqueness. It turns out that, given δ, the positive definiteness of a matrix that we denote $\mathbf{Q}$ ensures equilibrium uniqueness. This matrix $\mathbf{Q}$ can be determined from $\mathbf{G}$ in a one-to-one correspondence once δ is fixed.

Recall that agent i's expected payoffs are given by the following equation:

$$u_i(x_1, ..., x_n) = \mathbb{E}\left[-\left(\mu - \sum_{j=1}^{n} g_{ij}((1-x_j)s_j + x_j\mathcal{I}_j)\right)^2\right] - x_i c.$$

Proposition 3.3. *The profile of active learning choices $\boldsymbol{x}^* = (x_1^*, ..., x_n^*)$ is an equilibrium of the game if and only if*

$$(\boldsymbol{\theta} - \hat{\mathbf{Q}}\boldsymbol{x}^*)^T(\boldsymbol{x}^* - \boldsymbol{x}') \geq 0 \tag{4}$$

for any $\boldsymbol{x}' \in [0,1]^n$, with the matrix

$$\hat{\mathbf{Q}} = (\sigma^2 + \tilde{\sigma}^2)\begin{pmatrix} 2g_{11}^2 & 2\delta g_{11}g_{12} & ... & 2\delta g_{11}g_{1n} \\ 2\delta g_{22}g_{21} & 2g_{22}^2 & ... & 2\delta g_{22}g_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 2\delta g_{n1}g_{nn} & 2\delta g_{n2}g_{nn} & ... & 2g_{nn}^2 \end{pmatrix}$$

and the vector

$$\boldsymbol{\theta} = (2\sigma^2 g_{11}^2 - c, ..., 2\sigma^2 g_{nn}^2 - c).$$

Proof. First, the following equivalence is established: the profile $\boldsymbol{x}^*$ is an equilibrium if and only if

$$\frac{\partial}{\partial x_i}\left[u_i(x_i^*, x_{-i}^*)\right](x_i' - x_i^*) \leq 0$$

for all i and $x_i' \in [0,1]$.

Fixing a profile $\boldsymbol{x}^* \in [0,1]^n$ and an agent i, let us define

$$g(t) := u_i(x_i' + t(x_i^* - x_i'), x_{-i}^*)$$

for $0 \leq t \leq 1$ and $x_i' \in [0,1]$. The derivative with respect to t is given by $g'(t) = \frac{\partial}{\partial x_i}(u_i(x_i, x_{-i}^*))_{|x_i = x_i' + (x_i^* - x_i')}(x_i^* - x_i')$. If $\boldsymbol{x}^*$ is an equilibrium, $g(t)$ has a maximum at $t = 1$ and $g'(1) \geq 0$. Hence,

$$\frac{\partial}{\partial x_i}\left[u_i(x_i^*, x_{-i}^*)\right](x_i' - x_i^*) \leq 0.$$

Now, let us show the converse. Concavity of g follows from the concavity of u_i,[8] and then $g(t) \leq g(y) + g'(y)(t-y)$ for any $t, y \in [0,1]$. Choosing $t = 0$ and $y = 1$, we see that $g(0) \leq g(1) - g'(1)$. Moreover, $g'(1) \geq 0$ by assumption, so that $-g'(1) \leq 0$ and

8 The function u_i is clearly twice differentiable with respect to x_i and $\frac{\partial^2 u_i}{\partial x_i^2} = -2(-g_{ii}(\mathcal{I}_i - s_i))^2 < 0$.

$g(0) \leq g(1)$. This inequality implies that

$$u_i(x_i^*, x_{-i}^*) \geq u_i(x_i', x_{-i}^*)$$

for all $x_i' \in [0,1]$, and $\boldsymbol{x}^*$ is an equilibrium.

Summing up with respect to all agent yields

$$\sum_{i=1}^{n} \left(\frac{\partial}{\partial x_i} \left[u_i(x_i^*, x_{-i}^*) \right] (x_i' - x_i^*) \right) \leq 0.$$

Denoting by $\left(\frac{\partial}{\partial x_i} u_i(x_i^*, x_{-i}^*)\right)_i$ the vector given by stacking up all $\frac{\partial u_i}{\partial x_i}$, the previous inequality can be rewritten as

$$\left(\frac{\partial}{\partial x_i} [u_i(x_i^*, x_{-i}^*)] \right)_i^T (\boldsymbol{x}' - \boldsymbol{x}^*) \leq 0.$$

The profile of active learning choices $\boldsymbol{x}^*$ is an equilibrium if and only if this inequality holds for any $\boldsymbol{x}' \in [0,1]^n$.[9] It just remains to explicitely derive the vector components, which are given by

$$\frac{\partial}{\partial x_i} [u_i(x_i^*, x_{-i}^*)] = 2g_{ii}^2 \sigma^2 - 2g_{ii}^2 x_i^* (\sigma^2 + \tilde{\sigma}^2) - 2\alpha g_{ii} \sum_{j \neq i} g_{ij} x_j^* - c.$$

Finally, it is a mere verification to check that defining $\hat{\mathbf{Q}}$ and $\boldsymbol{\theta}$ as above, $\boldsymbol{x}^*$ is an equilibrium if and only if

$$(\boldsymbol{\theta} - \hat{\mathbf{Q}}\mathbf{x}^*)^T (\boldsymbol{x}^* - \boldsymbol{x}') \geq 0.$$

□

If $\hat{\mathbf{Q}}$ is positive definite, there is just one vector of information levels $\mathbf{x}^*$ that satisfies (4). This is the sufficient condition for equilibrium uniqueness.

Proposition 3.4. *If the matrix $\hat{\mathbf{Q}}$ is positive definite, the equilibrium is unique.*

Proof. Suppose $\mathbf{x}_1^*$ and $\mathbf{x}_2^*$ are two different equilibria. Then, $(\boldsymbol{\theta} - \hat{\mathbf{Q}}\mathbf{x}_1^*)^T (\mathbf{x}_1^* - \mathbf{x}_2^*) \geq 0$ and $(\boldsymbol{\theta} - \hat{\mathbf{Q}}\mathbf{x}_2^*)^T (\mathbf{x}_2^* - \mathbf{x}_1^*) \geq 0$. Summing up both inequalities yields $(\boldsymbol{\theta} - \hat{\mathbf{Q}}\mathbf{x}_1^*)^T (\mathbf{x}_1^* -$

9 If $\hat{\mathbf{x}}$ is not an equilibrium, then there is some agent j such that

$$\frac{\partial}{\partial x_j} (u_j(\hat{x}_j, \hat{x}_{-j})) (x_j' - \hat{x}_j) > 0$$

for some $x_j' \in [0,1]$. Hence, defining the profile $\tilde{\mathbf{x}}$ as $\tilde{x}_j = x_j'$ and $\tilde{x}_i = \hat{x}_i$ for $i \neq j$,

$$\sum_i \left(\frac{\partial}{\partial x_i} (u_i(\hat{x}_i, \hat{x}_{-i})) \right) (\tilde{x}_i - \hat{x}_i) = \frac{\partial}{\partial x_j} (u_j(\hat{x}_j, \hat{x}_{-j})) (x_j' - \hat{x}_j) > 0.$$

$\mathbf{x}_2^*) + (\boldsymbol{\theta} - \hat{\mathbf{Q}}\mathbf{x}_2^*)^T(\mathbf{x}_2^* - \mathbf{x}_1^*) \geq 0$, which holds if and only if

$$(\mathbf{x}_2^* - \mathbf{x}_1^*)^T\hat{\mathbf{Q}}(\mathbf{x}_2^* - \mathbf{x}_1^*) \leq 0.$$

But $\hat{\mathbf{Q}}$ is positive definite, i.e. $\mathbf{x}^T\hat{\mathbf{Q}}\mathbf{x} > 0$ for all $x \neq 0$. Consequently, $\mathbf{x}_1^* = \mathbf{x}_2^*$ and the equilibrium is unique. □

Dividing $\hat{\mathbf{Q}}$ by $\sigma^2 + \tilde{\sigma}^2$ does not change its definiteness and simplifies the expression—recall that $\sigma^2 + \tilde{\sigma}^2 > 0$.[10] Thus, $\mathbf{Q}$ is given by

$$\mathbf{Q} = \frac{1}{\sigma^2 + \tilde{\sigma}^2}\hat{\mathbf{Q}} = \begin{pmatrix} 2g_{11}^2 & 2\delta g_{11}g_{12} & \dots & 2\delta g_{11}g_{1n} \\ 2\delta g_{22}g_{21} & 2g_{22}^2 & \dots & 2\delta g_{22}g_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 2\delta g_{n1}g_{nn} & 2\delta g_{n2}g_{nn} & \dots & 2g_{nn}^2 \end{pmatrix}. \tag{5}$$

The following result shows that $\mathbf{Q}$ is completely determined by $\mathbf{G}$, once δ is fixed. Consequently, equilibrium uniqueness for this model depends solely on the influence network $\mathbf{G}$.

Proposition 3.5. *Given δ, there is a one-to-one correspondence between $\boldsymbol{Q}$ and $\boldsymbol{G}$.*

Proof. Given δ, the matrix $\mathbf{Q}$ is defined element-wise from $\mathbf{G}$ as in (5). Assume δ is fixed and denote this transformation by ϕ_δ. Let us show that it is possible to recover $\mathbf{G}$ from $\mathbf{Q}$. Denoting by q_{ij} the elements in $\mathbf{Q}$, let us define (element-wise) the transformation τ_δ by $\tau_\delta(q_{ii}) = \sqrt{q_{ii}}$ for all i and $\tau_\delta(q_{ij}) = \frac{q_{ij}}{\delta\sqrt{q_{ii}}}$ for all $i \neq j$. It is trivial to check that $\tau_\delta(\phi_\delta(\mathbf{G})) = \mathbf{G}$ and $\phi_\delta(\tau_\delta(\mathbf{Q})) = \mathbf{Q}$. □

In general, the matrix $\mathbf{Q}$ is an $n \times n$ matrix that need not be symmetric. Note that $\mathbf{x}^T\mathbf{Q}\mathbf{x} = \frac{1}{2}\mathbf{x}^T(\mathbf{Q} + \mathbf{Q}^T)\mathbf{x}$, and $\mathbf{Q}$ is positive definite if and only if $\mathbf{A} := \frac{1}{2}(\mathbf{Q} + \mathbf{Q}^T)$ is positive definite. Since $\mathbf{A}$ is symmetric, we can use the characterization of positive definiteness in terms of eigenvalues: a symmetric matrix is positive definite if and only if all of its eigenvalues are positive. Let $\lambda_1(\mathbf{A})$ denote the lowest eigenvalue of $\mathbf{A}$.

Corollary 3.6. *If $\lambda_1(\boldsymbol{A}) > 0$, then the equilibrium is unique.*

The explicit expression for $\mathbf{A}$ is given by:

$$\mathbf{A} = \begin{pmatrix} 2g_{11}^2 & \delta(g_{12}g_{11} + g_{21}g_{22}) & \dots & \delta(g_{1n}g_{11} + g_{n1}g_{nn}) \\ \delta(g_{21}g_{22} + g_{12}g_{11}) & 2g_{22}^2 & \dots & \delta(g_{2n}g_{22} + g_{n2}g_{nn}) \\ \vdots & \vdots & \ddots & \vdots \\ \delta(g_{n1}g_{nn} + g_{1n}g_{11}) & \delta(g_{n2}g_{nn} + g_{2n}g_{22}) & \dots & 2g_{nn}^2 \end{pmatrix}.$$

10 It would be possible to divide by $2(\sigma^2 + \tilde{\sigma}^2)$ instead, but keeping the factor 2 simplifies the expression for the matrix $\mathbf{A}$ later.

Note that this sufficient condition is independent of the cost c of active learning but depends on the influences between agents and the substitutability of information acquisition.

The scope of this condition is the focus of our subsequent analysis. We will make more restrictive assumptions on the model to explore particular cases of interest, which will eventually lead to a result similar to that of Bramoullé, Kranton, and D'Amours (2014). Later, we will apply the sufficient condition to the examples in Section 3.1. We first prove an auxiliary lemma.

Lemma 3.7. *Let $\boldsymbol{A}$ be a symmetric matrix and $\beta, \delta > 0$. The matrix $\beta \boldsymbol{Id} + \delta \boldsymbol{A}$ is positive definite if and only if $\lambda_1(\boldsymbol{A}) \geq -\frac{\beta}{\delta}$.*

Proof. The matrix $\beta\,\mathbf{Id} + \delta\mathbf{A}$ is positive definite if and only if all the solutions λ to $\det[\lambda\mathbf{Id} - (\beta\,\mathbf{Id} + \delta\mathbf{A})] = 0$ are strictly positive. The equation is equivalent to $\det[\frac{\lambda-\beta}{\delta}\,\mathbf{Id} - \mathbf{A}] = 0$. Note that the eigenvalues of $\mathbf{A}$ are the solutions t to the equation $\det[t\,\mathbf{Id} - \mathbf{A}] = 0$. Consequently, as $t = \frac{\lambda-\beta}{\delta}$, the condition $\lambda > 0$ can be translated into all eigenvalues t of $\mathbf{A}$ verifying $t > -\frac{\beta}{\delta}$. This is precisely the condition $\lambda_1(\mathbf{A}) > -\frac{\beta}{\delta}$. ◻

Next, we consider two particular cases that are worth exploring. Assuming that all agents pay the same attention to themselves, i.e., $g_{ii} = g_{jj}$ for all i, j, we can denote the diagonal terms of $\mathbf{G}$ by $\beta := g_{ii} > 0$. We define $\bar{\mathbf{A}}$ as

$$\bar{\mathbf{A}} = \begin{pmatrix} 0 & \frac{g_{12}+g_{21}}{2} & \dots & \frac{g_{1n}+g_{n1}}{2} \\ \frac{g_{21}+g_{12}}{2} & \ddots & \dots & \frac{g_{2n}+g_{n2}}{2} \\ \vdots & \dots & \ddots & \vdots \\ \frac{g_{n1}+g_{1n}}{2} & \frac{g_{n2}+g_{2n}}{2} & \dots & 0 \end{pmatrix}.$$

Using Lemma 3.7, we see that $\lambda_1(\mathbf{A}) > 0$ if and only if $\lambda_1(\bar{\mathbf{A}}) > -\frac{\beta}{\delta}$. Note that $\bar{\mathbf{A}}$ is simply $\bar{\mathbf{A}} = \frac{1}{2}(\mathbf{G} + \mathbf{G}^T) - \beta\mathbf{Id}$. Here, $\bar{\mathbf{A}}$ reflects the average flow of information between a pair of networks, or the undirected network associated with $\mathbf{G}$.

Now, assume that the network displays reciprocal relations, i.e., $g_{ij} = g_{ji}$, in addition to same self-importance across agents. This means that the influence of agent i on agent j is the same as that of agent j on agent i, and so the matrix $\mathbf{G}$ is symmetric and can be seen as undirected. Again, $\lambda_1(\mathbf{A}) > 0$ if and only if $\lambda_1(\bar{\mathbf{A}}) > -\frac{\beta}{\delta}$, but $\bar{\mathbf{A}}$ is now simply $\mathbf{G} - \beta\mathbf{Id}$. The matrix $\bar{\mathbf{A}} = \mathbf{G} - \beta\,\mathbf{Id}$ can be seen as a generalization of the matrix $\mathbf{G}$ in Bramoullé, Kranton, and D'Amours (2014), where $\bar{a}_{ij} \in [0, 1]$ instead of $g_{ij} \in \{0, 1\}$. The sufficient condition is equivalent to theirs. However, to derive such a result they use the potential theory developed by Monderer and Shapley (1996), which requires symmetry—this is why we cannot apply it to the general model.

Proposition 3.8. *The sufficient condition for the uniqueness of equilibrium can be specialized to two particular cases:*

- *If self-importance is equal across agents ($g_{ii} = g_{jj} = \beta$ for all i, j), the condition becomes $\lambda_1(\bar{\boldsymbol{A}}) > -\frac{\beta}{\delta}$ with $\bar{\boldsymbol{A}} = \frac{1}{2}(\boldsymbol{G} + \boldsymbol{G}^T) - \beta \boldsymbol{Id}$.*
- *If on top of that the influences are reciprocal ($g_{ij} = g_{ji}$ for all i, j), the condition becomes $\lambda_1(\bar{\boldsymbol{A}}) > -\frac{\beta}{\delta}$ with $\bar{\boldsymbol{A}} = \boldsymbol{G} - \beta \boldsymbol{Id}$.*

This proposition summarizes the results obtained so far, which show that the condition for the uniqueness of equilibrium can be specialized for two particular cases: when all agents have the same level of self-importance and when the network exhibits reciprocal relations between agents. In both cases, the condition involves the eigenvalue of a matrix $\bar{\mathbf{A}}$, which can be calculated based on the properties of the network. The precise definition of $\bar{\mathbf{A}}$ is given for each case.

3.3.1 Examples: Uniqueness

The networks analyzed in Section 3.1 are reviewed again to apply the equilibrium uniqueness condition.

First, we revisit the class of **k-regular graphs** with n agents that share their attention homogeneously. Proposition 3.8 applies, and the lowest eigenvalue of $\bar{\mathbf{A}}$ is $\lambda_1(\bar{\mathbf{A}}) = -\frac{1}{k}$. The equilibrium is unique if $\delta < 1$, which always holds. As an example of this class of networks, the matrix $\bar{\mathbf{A}}$ associated with the complete graph is given by

$$\bar{\mathbf{A}} = \begin{pmatrix} 0 & \frac{1}{n} & \dots & \frac{1}{n} \\ \frac{1}{n} & 0 & \dots & \vdots \\ \vdots & \dots & \ddots & \frac{1}{n} \\ \frac{1}{n} & \frac{1}{n} & \dots & 0 \end{pmatrix}.$$

Next, we consider the class of **stars**. Due to the asymmetry of $\mathbf{Q}$ and the different terms in the diagonal (self-importance is not equal across agents), only Corollary 3.6 applies. The equilibrium is unique if $\lambda_1(\mathbf{A}) > 0$, which depends on both δ and ε. Matrix $\mathbf{A}$ is given here by:

$$\mathbf{A} = \begin{pmatrix} \frac{2}{n^2} & \delta((1-\varepsilon)\varepsilon + \frac{1}{n^2}) & \dots & \delta((1-\varepsilon)\varepsilon + \frac{1}{n^2}) \\ \delta((1-\varepsilon)\varepsilon + \frac{1}{n^2}) & 2(1-\varepsilon)^2 & \dots & 0 \\ \vdots & \dots & \ddots & \vdots \\ \delta((1-\varepsilon)\varepsilon + \frac{1}{n^2}) & 0 & \dots & 2(1-\varepsilon)^2 \end{pmatrix}.$$

Figure 12 shows the values for which a unique equilibrium is ensured—every pair (δ, ε) such that the blue surface is above the orange plane.

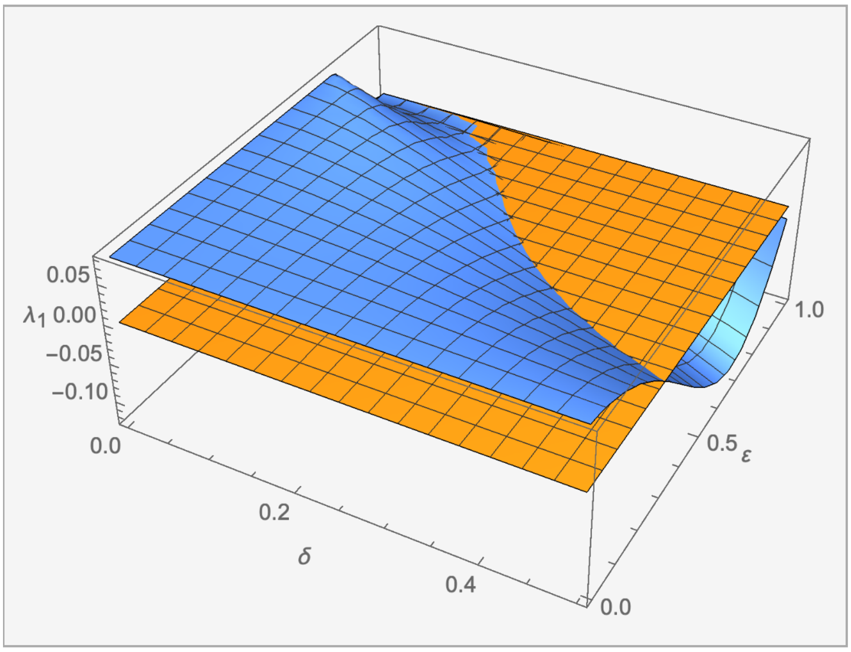


Figure 12: The lowest eigenvalue of the star.

A particular network structure belonging to the class of **core-periphery networks** was set in Figure 6. Here,

$$\mathbf{A} = \begin{pmatrix} \frac{2}{9} & \delta\frac{2}{81} & \delta\frac{2}{81} & \delta\frac{9}{100} & 0 & 0 \\ \delta\frac{2}{81} & \frac{2}{9} & \delta\frac{2}{81} & 0 & \delta\frac{9}{100} & 0 \\ \delta\frac{2}{81} & \delta\frac{2}{81} & \frac{2}{9} & 0 & 0 & \delta\frac{9}{100} \\ \delta\frac{9}{100} & 0 & 0 & \frac{2}{100} & 0 & 0 \\ 0 & \delta\frac{9}{100} & 0 & 0 & \frac{2}{100} & 0 \\ 0 & 0 & \delta\frac{9}{100} & 0 & 0 & \frac{2}{100} \end{pmatrix}.$$

and we apply Corollary 3.6. It turns out that $\lambda_1(\mathbf{A}) > 0$ for all $\delta \in [0, \frac{1}{2}]$, so the equilibrium is always unique.[11]

The **criminal network** from Ballester, Calvó-Armengol, and Zenou (2006) was represented in Figure 8. Proceeding as before, we calculate the lowest eigenvalue of $\mathbf{A}$.[12] We find that $\lambda_1(\mathbf{A}) > 0$ for all $\delta < 0.45011$, which guarantees a unique equilibrium for such values.

Finally, we consider the **incomplete network** depicted in Figure 10. Recall that

11 The explicit expression for the lowest eigenvalue of $\mathbf{A}$ is $\lambda_1(\mathbf{A}) = \frac{981-100\delta-\sqrt{670761-163800\delta+541441\delta^2}}{8100}$. We see that $\lambda_1(\mathbf{A})$ is a decreasing function of δ in $[0, 1/2]$. As it is strictly positive at $\delta = 1/2$, $\lambda_1(\mathbf{A}) > 0$ for all $\delta \in \left[0, \frac{1}{2}\right]$.

12 Let $y_1(\delta)$, $y_2(\delta)$ and $y_3(\delta)$ be the three roots of $-32400 - 97200\delta + 259200\delta^3 + (3096 + 6192\delta - 7200\delta^2)y + (-97 - 97\delta)y^2 + y^3$. Let $y_1(\delta)$ be the smallest root in $\delta \in \left[0, \frac{1}{2}\right]$. Then, $\lambda_1(\mathbf{A}) = \frac{1}{450}y_1(\delta)$ and $\lambda_1(\mathbf{A}) > 0 \Leftrightarrow \delta < 0.45011$.

$\delta = \frac{1}{2}$ and all diagonal terms are equal: $\beta = \frac{1}{3}$. To apply Proposition 3.8, we compute

$$\bar{\mathbf{A}} = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 \end{pmatrix}.$$

The uniqueness condition $\lambda_1(\bar{\mathbf{A}}) > -\frac{\beta}{\delta}$ is not satisfied because $\lambda_1(\bar{\mathbf{A}}) = -\frac{1}{\sqrt{2}} < -\frac{2}{3}$. This was expected, as we had already obtained two different equilibria for this particular network.

3.3.2 The lowest eigenvalue

The present subsection explores the meaning of the uniqueness condition and provides an intuition. A network is bipartite if agents can be divided into two sets, say R and S, such that if $i \in R$, i is not connected to any $j \in R$ except for herself. The network is completely bipartite if every $i \in R$ is connected to all $j \in S$. Bipartite networks represent disjoint or independent communities. An affiliation network is a classic example. Another bipartite network might be found when representing supervisor-candidate communication. A complete bipartite network represents one extreme of two-sidedness. The other extreme is the complete regular graph. In this subsection, we talk about two-sidedness as an intuitive measure of how close a network is to the complete bipartite graph.

First, let us briefly characterize $\lambda_1(\mathbf{A})$.[13] By definition, $\lambda_1(\mathbf{A}) = \min\{\lambda \in \mathbb{R} : \exists \epsilon \in \mathbb{R}^n$ satisfying $\lambda\epsilon = \mathbf{A}\epsilon\}$. Assuming $\epsilon \neq 0$, $\lambda\epsilon = \mathbf{A}\epsilon$ implies $\epsilon^T\lambda\epsilon = \epsilon^T\mathbf{A}\epsilon$, which leads to $\lambda\epsilon^T\epsilon = \epsilon^T\mathbf{A}\epsilon$, and finally to $\lambda||\epsilon||^2 = \epsilon^T\mathbf{A}\epsilon$. So, if $||\epsilon|| = 1$, then $\lambda = \epsilon^T\mathbf{A}\epsilon$. Hence, $\lambda_1(\mathbf{A}) = \min\Big\{\lambda \in \mathbb{R} : \lambda = \epsilon^T\mathbf{A}\epsilon \text{ and } ||\epsilon|| = 1\Big\}$.

Following Bramoullé, Kranton, and D'Amours (2014), we can use an eigenvector ϵ associated to $\lambda_1(\mathbf{A})$ to separate the agents into two groups. If $\epsilon_i \geq 0$, agent i belongs to set R. Otherwise, she belongs to set S. This leads to the decomposition

$$\lambda_1(\mathbf{A}) = \varepsilon^T\mathbf{A}\varepsilon = \overbrace{\sum_{i,j\in R} \epsilon_i\epsilon_j q_{ij}}^{>0} + \overbrace{\sum_{i,j\in S} \epsilon_i\epsilon_j q_{ij}}^{>0} + \overbrace{2\sum_{i\in R, j\in S} \epsilon_i\epsilon_j q_{ij}}^{<0}.$$

The greater the lowest eigenvalue, the more weight the network puts within sets and the less it puts between sets. Hence, the size of $\lambda_1(\mathbf{A})$ is related to the two-sidedness of the graph $\mathbf{A}$. The closer the network is to the complete bipartite graph, the lower $\lambda_1(\mathbf{A})$. This is because transferring weight from links within R or S to links between both sets decreases $\lambda_1(\mathbf{A})$. Creating new links between sets or removing links within R or S belong to that kind of weight transfer. Thus, making the graph more two-sided decreases the lowest eigenvalue.

13 Remember that when $\mathbf{Q}$ is symmetric, then $\mathbf{A} = \mathbf{Q}$.

Let us show how the division of agents into the two groups is induced by agents' listening structures. We have $\lambda_1(\mathbf{A}) = \epsilon^t \mathbf{A} \epsilon = \sum_{i,j} q_{ij}\epsilon_i\epsilon_j$ with $||\epsilon|| = 1$. Without loss of generality, let us assume that $\lambda_1(\mathbf{A}) > 0$ (if not, a similar reasoning holds). Then, agent i belongs to R if and only if $\lambda_1(\mathbf{A})\epsilon_i > 0$. Since $\lambda_1(\mathbf{A})\epsilon_i = \mathbf{A}\epsilon_i$, we see that $i \in R$ if $\sum_{j=1}^n q_{ij}\epsilon_j \geq 0$. Consequently, if the listening structures of two agents are similar, they will belong to the same set. For example, if $(q_{ij})_j$ and $(q_{kj})_j$ are similar, then $\sum_{j=1}^n q_{ij}\epsilon_j > 0$, $\sum_{j=1}^n q_{kj}\epsilon_j > 0$, and both i and k belong to R. Hence, the division of agents into two groups induced by $\lambda_1(\mathbf{A})$ responds to their listening structures.

Recall that $\lambda_1(\mathbf{A}) > 0$ ensures uniqueness. The less two-sided the network is, the higher the chances of a unique equilibrium. Roughly, two-sided networks allow the agents from R and S to switch contributions in different equilibria. This occurs because the effects of substitutability (namely, the fact that if an agent contributes more, her neighbors contribute less and so on) accumulate and lead to several equilibrium configurations. When the network is not two-sided, this rebounding effect collapses, and there is only one equilibrium.

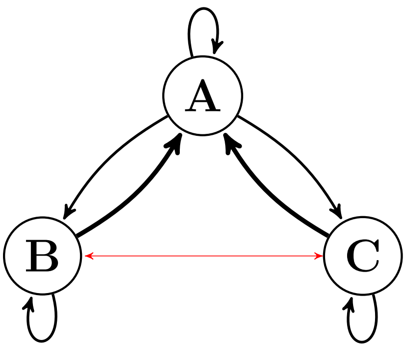


Figure 13: The extra link makes the network complete.

$$\mathbf{G}' = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 2/3 - \varepsilon & 1/3 & \varepsilon \\ 2/3 - \varepsilon & \varepsilon & 1/3 \end{pmatrix}$$

As an example, consider the incomplete network with three agents from Figure 10. Recall that for $\delta = \frac{1}{2}$ the network features multiple equilibria. The lowest eigenvalue is $\lambda_1(\mathbf{Q}) = \frac{2-3\sqrt{2}\delta}{9}$, and an associated eigenvector is $(-\sqrt{2}, 1, 1)$. Thus, the partition is given by $R = \{\text{A}\}$ and $S = \{\text{B,C}\}$.[14] The network is considerably two-sided. What would happen if we add a link between agents B and C, slightly decreasing the two-sidedness of the network according to ε? The resulting network, shown in Figure 13, would be less similar to the bipartite network of three agents. In fact, it turns out that for all $\varepsilon > 0.057$, $\lambda_1(\mathbf{Q}') > 0$ and the equilibrium is unique, where $\mathbf{Q}'$ is the matrix induced by $\mathbf{G}'$. The network is less two-sided, and multiplicity disappears.

Furthermore, it is worth noting that $g_{ii} > 0$ for all i and self-importance (which

14 Note that $(\sqrt{2}, -1, -1)$ is also an eigenvector associated to the eigenvalue $\frac{2-3\sqrt{2}\delta}{9}$. The partition induced by it is the equivalent to the one above: $R = \{B, C\}$ and $S = \{A\}$.

represent the influence an agent exert on herself) contribute to the positivity of $\lambda_1(\mathbf{Q})$. For example, consider the case $g_{ii} = g_{jj} = \beta$ for all i, j. In this case, Proposition 3.8 simplifies the uniqueness condition to $\lambda_1(\bar{\mathbf{A}}) > -\frac{\beta}{\delta}$, where $\bar{\mathbf{A}} = \frac{1}{2}(\mathbf{G} + \mathbf{G}^T) - \beta\mathbf{Id}$. Then, $\lambda_1(\bar{\mathbf{A}}) = \epsilon^T \bar{\mathbf{A}}\epsilon = \sum_{i\neq j}\left(\frac{g_{ij}+g_{ji}}{2}\right)\epsilon_i\epsilon_j$. The equilibrium is unique if

$$\sum_{i\neq j\in R}(g_{ij}+g_{ji})\epsilon_i\epsilon_j + \sum_{i\neq j\in S}(g_{ij}+g_{ji})\epsilon_i\epsilon_j + \sum_{i\in S, j\in S}(g_{ij}+g_{ji})\epsilon_i\epsilon_j > -\frac{\beta}{\delta}.$$

As agents put more weight on their own signals (i.e., as β grows), the network becomes less bipartite, which contributes to potential equilibrium uniqueness.

4 Social Welfare

So far, agents have behaved individually. Now, the focus is shifted to a social perspective that maximizes aggregated welfare in the network. We can think of a utilitarian social planner who decides on the levels of active learning to pursue this goal.

The vector of learning levels $\mathbf{x}^{UO} = (x_1^{UO}, ..., x_n^{UO})$ that maximizes the sum of agents' utilities (the utilitarian optimum) is given by

$$x_i^{UO} = \max\left\{0, \frac{2\sigma^2 - c/(\sum_{j=1}^n g_{ji}^2)}{2(\sigma^2+\tilde{\sigma}^2)} - \delta\frac{\sum_{j=1}^n g_{ji}}{\sum_{j=1}^n g_{ji}^2}\sum_{j\neq i} g_{ij}x_j\right\}. \tag{6}$$

Agent i's learning target in the utilitarian optimum is

$$\tilde{x}_i = \frac{2\sigma^2 - c/(\sum_{j=1}^n g_{ji}^2)}{2(\sigma^2+\tilde{\sigma}^2)}.$$

Now, we compare the target $\tilde{x}_i$ to the target in equilibrium, $\bar{x}_i$. The sum of the squares of i's influences is greater than the square of her self-influence: $\sum_{j=1}^n g_{ji}^2 \geq g_{ii}^2$. Then, $\frac{c}{\sum_{j=1}^n g_{ji}^2} \leq \frac{c}{g_{ii}^2}$, and directly from the definitions of targets, we get

$$\tilde{x}_i \geq \bar{x}_i.$$

Thus, in the utilitarian optimum, each agent would like to learn strictly more, except in the trivial case where she is isolated. This effect is due to $\sum_{j=1}^n g_{ji}^2$, which substitutes g_{ii}^2 in the target expression. Before, each agent just cared about self-benefit: the more she listened to her signal, the more information she needed. Now, the goal is shifted, and individuals must care about the influence they have on others. The term $\sum_{j=1}^n g_{ji}^2$ is a measure of i's total impact on the network. The larger the influence, the higher the target $\tilde{x}_i$.

However, the utilitarian level of active learning x_i^{UO} need not be higher than the equilibrium choice x_i^*. The last term in (6) indicates the amount of information agent i

does not need to purchase because of the substitutability effect. Substitutability is driven here by $\delta\frac{\sum_{j=1}^n g_{ji}}{\sum_{j=1}^n g_{ji}^2}$, whereas it was driven by the factor $\delta\frac{1}{g_{ii}}$ in equilibrium. Hence, it might be the case that an agent who engages in high levels of active learning in equilibrium is not influential at all (i.e., $\sum_{j=1}^n g_{ji}$ is small), and the planner asks her to decrease her effort. Even though every agent desires to become more informed (the target is higher), utilitarian maximization implies a more efficient share of effort in global terms. Thus, in general, there is no ranking regarding acquisition decisions. Formally,

$$x_i^{UO} \geq x_i^* \Leftrightarrow \frac{c}{2(\sigma^2+\tilde{\sigma}^2)}\left(\frac{1}{g_{ii}^2} - \frac{1}{\sum_{j=1}^n g_{ji}^2}\right) \geq \delta\left(\frac{\sum_{j=1}^n g_{ji}}{\sum_{j=1}^n g_{ji}^2}\sum_{j=1}^n g_{ji}x_j^{UO} - \frac{\sum_{j=1}^n g_{ij}x_j^*}{g_{ii}}\right).$$

On the one hand, we observe that the inequality would hold for networks in which attention is homogenously shared. On the other hand, it would also hold for networks with low levels of substitutability. This condition can be formalized.

Proposition 4.1. *For every network structure* $\boldsymbol{G}$ *there exists some* $\bar{\delta} \in (0,1)$ *such that if* $\delta \leq \bar{\delta}$*, then* $x_i^{UO} \geq x_i^*$ *for every agent* i.

The intuition behind this result is simple: for low levels of substitutability, every agent relies on her information target, which is always higher under the utilitarian planner. To illustrate the relation between network balance and the ranking in acquisition choices we provide an example. Suppose we have the network shown in Figure 10, and let the parameter values be $\delta = 0.2$, $c = 0.1$, $\sigma^2 = 3$, and $\tilde{\sigma}^2 = 1$. The network matrix and the equilibrium and utilitarian optimal choices are

$$\mathbf{G} = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}, \qquad \begin{array}{c} \mathbf{x}^* = (0.23, 0.23, 0.23), \\ \mathbf{x}^{UO} = (0.50, 0.51, 0.51). \end{array}$$

Here, the network is balanced, meaning weights are shared similarly among agents, and the utilitarian optimal choices are larger. However, if agent 1 becomes stubborn (i.e., g_{11} is close to 1), the network becomes unbalanced:

$$\mathbf{G}' = \begin{pmatrix} 8/10 & 1/10 & 1/10 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}, \qquad \begin{array}{c} (\mathbf{x}^*)' = (0.69, 0, 0), \\ (\mathbf{x}^{UO})' = (0.57, 0.39, 0.39). \end{array}$$

In this case, agent 1 is the only one exerting effort in equilibrium, while the others free-ride. From a social point of view, this is not efficient, and agent 1 has to decrease her contribution while agents 2 and 3 increase theirs significantly.

Finally, we show by example that, in general, there is no Pareto dominance between the utilitarian optimum and equilibria—not even for low values of δ. Consider the above networks $\mathbf{G}$ and $\mathbf{G}'$ for the same parameter values again, focusing on agent 1. In $\mathbf{G}$, agent

1's utilities in the utilitarian optimum and the unique equilibrium are

$$U_1(\mathbf{x}^*) = -2.322,$$
$$U_1(\mathbf{x}^{UO}) = 0.017,$$

while in $\mathbf{G}'$, agent 1 utilities are given by

$$U_1((\mathbf{x}^*)') = 0.025,$$
$$U_1((\mathbf{x}^{UO})') = 0.008.$$

In such an unbalanced network, agent 1 strictly prefers the equilibrium allocation: as her self-importance is large, the increase in agent 2 and 3 information purchases does not make up for the decrease in hers under the utilitarian optimum. Thus, $\mathbf{x}^{UO}$ maximizes the sum of utilities, but in general it does not improve the well-being of every agent.

5 Extension to multiple periods

So far, we have considered a scenario where agents communicate only once. However, if we introduce multiple communication periods, agents can obtain information not only from their immediate neighbors but also from neighbors' neighbors. As time progresses, the DeGrootian posterior signal incorporates signals from individuals located at increasing distances. After t periods, each agent holds a belief containing signals from all individuals who live within t degrees of separation. One significant advantage of DeGroot updating is that the weights of period t are simply given by the stochastic matrix $\mathbf{G}^t$. This implies that the information acquisition problem for t periods is identical to the one considered so far, except that the matrix $\mathbf{G}$ is now replaced by $\mathbf{G}^t$.

The limiting case $t \to \infty$ corresponds to long-run communication. There, each agent's posterior signal aggregates information from everyone in the network. Under very mild conditions there is *convergence*, meaning that different agents' posterior signals coincide. The $n \times n$ stochastic matrix $\mathbf{G}$ is said to be convergent if $\lim_{t\to\infty} \mathbf{G}^t\mathbf{v}$ exists for all $\mathbf{v} \in \mathbb{R}^n$. In this case, there exists a unique left eigenvector $\boldsymbol{\pi} = (\pi_1, ..., \pi_n)$ of $\mathbf{G}$ whose entries sum to 1 such that $(\lim_{t\to\infty} \mathbf{G}^t\mathbf{v})_i = \boldsymbol{\pi}^t\mathbf{v}$ for every i and all $\mathbf{v} \in \mathbb{R}^n$,[15] that is:

$$\lim_{t\to\infty} \mathbf{G}^t = \begin{pmatrix} \boldsymbol{\pi}^t \\ \vdots \\ \boldsymbol{\pi}^t \end{pmatrix}.$$

The components of $\boldsymbol{\pi}$ indicate how much each agent is listened to in the long-run. Again, this is equivalent to a public goods game, as there are n agents privately deciding how

15 This result is taken from Golub and Jackson (2010).

much to collaborate towards a common payoff. Thus, an influential individual, i.e., an individual with a large π_i, will purchase a significant amount of information, while another whose influence vanishes will just free-ride.

Requiring one agent to put positive weight on her belief (i.e., at least one $g_{ii} > 0$) is enough to ensure convergence for a stochastic matrix.[16] Hence, every network matrix analyzed in this paper is convergent. Equilibrium efforts for $t \to \infty$ are given by

$$x_i^* = \max\left\{0, \frac{2\sigma^2 - c/\pi_i^2}{2(\sigma^2 + \tilde{\sigma}^2)} - \frac{\delta}{\pi_i}\sum_{j \neq i} \pi_j x_j\right\}.$$

All results shown so far hold for the long run with the corresponding matrix $\lim_{t\to\infty} \mathbf{G}^t$.

The utilitarian optimum is given by

$$x_i^{UO} = \max\left\{0, \frac{2\sigma^2 - c/(n\pi_i^2)}{2(\sigma^2 + \tilde{\sigma}^2)} - \frac{\delta}{\pi_i}\sum_{j \neq i} \pi_j x_j\right\}.$$

It is worth noting that $x_i^* \leq x_i^{UO}$ for all agents i and every network, in stark contrast to the one-shot game. In the limit, neighborhoods disappear and each agent i influences every other agent, including herself, in the same manner: π_i. Hence, the substitutability of information is identical under both the utilitarian optimum and the equilibrium allocation. However, as the information target is always higher under the utilitarian optimum, the levels of information acquisition are also higher. However, the utilitarian optimum is not always a Pareto improvement. In a setting with very low (almost negligible) substitutability levels, for example, agents would prefer the equilibrium allocation to the utilitarian optimum.

5.1 Examples: Infinitely many communication periods

The networks analyzed in Section 3.1 are reviewed again assuming that agents communicate for infinitely many periods before acquisition decisions are made. Here, the network matrix $\mathbf{G}$ is substituted with the matrix $\lim_{t\to\infty} \mathbf{G}^t$, which is well-defined since every network matrix analyzed in this paper converges.

First, let us revisit the case of k**-regular graphs** with n agents that share their attention homogeneously. Given a fixed n, the specific limiting matrix $\lim_{t\to\infty} \mathbf{G}^t$ depends not only on $k \leq n$ but also on the network configuration. Agents not belonging to a cycle in the graph (excluding loops) will not be listened to in the long run, resulting in $\pi_j = 0$ for such agents j. The remaining agents (whose number we denote by $\tilde{k}$) share attention

16 If there is one agent i such that $g_{ii} > 0$, then the matrix is aperiodic. For strongly connected matrices, aperiodicity is necessary and sufficient for convergence; see Golub and Jackson (2010).

homogeneously. It holds that $\tilde{k} \geq k$. Then,

$$\pi_j = \begin{cases} 0 & \text{if there is no cycle to which } j \text{ belongs,} \\ \frac{1}{\tilde{k}} & \text{otherwise.} \end{cases}$$

And hence,

$$x_j^* = \begin{cases} 0 & \text{if there is no cycle to which } j \text{ belongs,} \\ \frac{2\sigma^2 - \tilde{k}^2 c}{2(\sigma^2+\tilde{\sigma}^2)(-\delta+\delta\tilde{k}+1)} & \text{if } j \text{ belongs to a cycle and} 2\sigma^2 - \tilde{k}c \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Comparing this with the one-shot communication version, we observe that fewer agents acquire information in the long run. Additionally, as $\tilde{k} \geq k$, the information acquisition levels decrease. This happens because, over time, all agents become connected to those who acquire information, and thus there is no need to acquire as much as before.

Now, we move to the class of **stars**. There, the hub pays homogeneous attention to the spokes, who, in turn, pay her attention ε. In the limit,

$$\boldsymbol{\pi} = \left(\frac{n\varepsilon}{(n-1)+n\varepsilon}, \frac{1}{(n-1)+n\varepsilon}, ..., \frac{1}{(n-1)+n\varepsilon} \right).$$

As the hub pays attention to the spokes, the star maintains the importance of all its members in the long run. Therefore, everyone is listened to, and, in principle, everyone acquires information—although the extent of information acquisition will also depend on the specific parameters involved. The hub is still the most influencial if ε is not too small. The acquisition levels in equilibrium are given by:

$$x_H^* = \frac{2\sigma^2 + \frac{1}{\varepsilon^2 n^2}(2\delta\varepsilon n(1+\varepsilon(n-2))n)\sigma^2 - c(\varepsilon n + n - 1)^2(1-\delta(2+(-1+\varepsilon(n-1))n))}{2(\sigma^2+\tilde{\sigma}^2)(\varepsilon-1)((n-1)\varepsilon\delta^2+\varepsilon-1)}$$

$$\text{if } c \leq \frac{(2\delta\varepsilon n(1+\varepsilon(n-2))n)\sigma^2}{(\varepsilon n+n-1)^2(1-\delta(2+(-1+\varepsilon(n-1))n))}; \quad x_H^* = 0 \text{ otherwise,}$$

$$x_S^* = \frac{2\sigma^2\varepsilon n(\delta\varepsilon n - 1) - c(\delta - \varepsilon n)(\varepsilon n + n - 1)^2}{2(\sigma^2+\tilde{\sigma}^2)(\varepsilon-1)((n-1)\varepsilon\delta^2+\varepsilon-1)\varepsilon n} \text{ if } c \leq \frac{2\sigma^2\varepsilon n(\delta\varepsilon n-1)}{(\delta-\varepsilon n)(\varepsilon n+n-1)^2};$$

$x_S^* = 0$ otherwise.

Now, revisiting the **core-periphery** network, it is important to note that agents in the core do not listen to peripheral agents, rendering the latter with no weight in the limit vector $\boldsymbol{\pi}$. If the core is composed by k agents, $\pi_j = \frac{1}{k}$ if j belongs to the core and $\pi_j = 0$ if j is peripheral. For the specific configuration from Section 3.1, acquisition levels are:

$$x_C^* = \frac{2\sigma^2 - 9c}{2(\sigma^2+\tilde{\sigma}^2)(2\delta+1)} \text{ if } c \leq \frac{2\sigma^2}{9}; \quad x_C^* = 0 \text{ otherwise,}$$

$$x_P^* = 0.$$

Core agents acquire the same information, but peripheral agents do not. Long-run communication does not affect core agents because they were already sharing information homogeneously (the restriction of $\mathbf{G}$ to the core is invariant under exponentiation to the power of t). Peripheral agents, in contrast, take into account their information in the one-shot game, but, with multiple communication rounds, their weight vanishes. Thus, they find it unprofitable to privately acquire information.

Finally, let us reexamine the **criminal network** from Ballester, Calvó-Armengol, and Zenou (2006). The limiting vector $\boldsymbol{\pi}$ is given by

$$\boldsymbol{\pi} = \left(\frac{5}{59}, \frac{6}{59}, \frac{5}{59}, \frac{5}{59}, \frac{5}{59}, \frac{6}{59}, \frac{6}{59}, \frac{5}{59}, \frac{5}{59}, \frac{5}{59}, \frac{6}{59}\right).$$

In the long run, only in-degree matters, but not network position. Hence, agent A no longer has a distinct role and there are just two classes of agents: B-class and C-class (to which agent A belongs now). B-class agents have in-degree 6, and $\pi_j = \frac{6}{59}$, and C-class agents have in-degree five, so $\pi_j = \frac{5}{59}$. Best reply functions are given by:

$$x_B^* = \max\left\{0, \bar{x}_B - \delta\frac{59}{6}(3x_B + 7x_C)\right\},$$
$$x_C^* = \max\left\{0, \bar{x}_C - \delta\frac{59}{5}(4x_B + 6x_C)\right\}.$$

Similar to the one-period communication game, B-class agents acquire less information. In particular, for the configuration of parameters used in Section 3.1, we can observe in Figure 14 that B-class agents completely free-ride on C-class agents. This happens because B-class agents consider acquiring private information too costly, relying instead on the information obtained from the seven C-class agents.

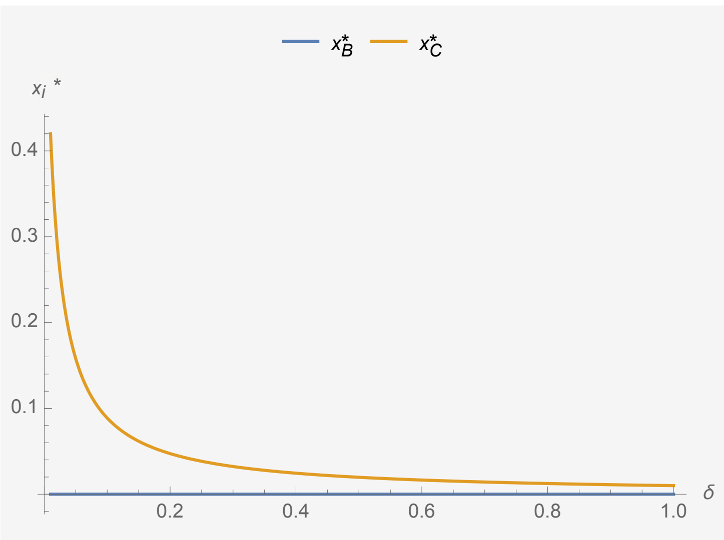


Figure 14: Acquisition for the criminal network in the long-run.

Finally, the examples from Section 3.3 are trivial in the long run. The incomplete three-agent network converges to a matrix characterized by the limit vector $\boldsymbol{\pi} = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right)$, while the four-agents *eye* converges to a matrix characterized by the limit vector $\boldsymbol{\pi} = \left(\frac{3}{10}, \frac{1}{5}, \frac{1}{5}, \frac{3}{10}\right)$. Both cases lead to unique equilibrium configurations.

6 Conclusion

This paper has analyzed the behavior of DeGroot updaters in a networked environment and studied the impact of substitutability and network structure on information acquisition and welfare. We have shown that the substitutability of agents' active learning efforts induces free-riding behavior and can lead to multiple equilibria. We have also provided a sufficient condition for equilibrium uniqueness in terms of the lowest eigenvalue of the matrix $\mathbf{A}$, which is determined by $\mathbf{G}$ and the parameter of substitutability δ. When this eigenvalue is positive, the equibrium is unique. Even if there are multiple equilibria, we have proposed a procedure for calculating them.

In terms of welfare, we have found that the information target is lower in equilibria than under the utilitarian paradigm. This is significant since the target is precisely the level of information an agent will have at the end of the game. We have shown that it is socially desirable to increase the information level of every agent. While increasing agents' active learning may seem like a solution, we show that in the one-shot game it is not. Not only the ranking in targets does not imply a ranking in acquisition levels, but the utilitarian optimum does not Pareto dominate the equilibrium allocation. Nevertheless, over the long run, neighborhood frictions are eliminated and the utilitarian allocation always exceeds the equilibrium allocation.

An interesting avenue for further research would be the implementation problem of a planner trying to incentivize DeGroot updaters to move from equilibrium levels of active learning to the utilitarian optimum. Public information policies, such as subsidizing external information sources, rewarding learning contributions, or creating new links to foster communication, could also be explored.

References

Acemoglu, D., and A. Ozdaglar (2011): "Opinion dynamics and learning in social networks," *Dynamic Games and Applications*, 1(1), 3–49.

Acemoglu, D., A. Ozdaglar, and A. ParandehGheibi (2010): "Spread of (mis) information in social networks," *Games and Economic Behavior*, 70(2), 194–227.

Aumann, R. J. (1997): "Rationality and bounded rationality," *Games and Economic behavior*, 21(1-2), 2–14.

Ballester, C., A. Calvó-Armengol, and Y. Zenou (2006): "Who's who in networks. Wanted: The key player," *Econometrica*, 74(5), 1403–1417.

Banerjee, A., E. Breza, A. G. Chandrasekhar, and M. Mobius (2021): "Naïve Learning with Uninformed Agents," *American Economic Review*, 111(11), 3540–74.

Borgatti, S. P., and M. G. Everett (2000): "Models of core/periphery structures," *Social Networks*, 21(4), 375–395.

Bramoullé, Y., R. Kranton, and M. D'Amours (2014): "Strategic interaction and networks," *American Economic Review*, 104(3,b), 898–930.

Bramoullé, Y., R. Kranton, et al. (2007): "Public goods in networks," *Journal of Economic Theory*, 135(1), 478–494.

Chandrasekhar, A. G., H. Larreguy, and J. P. Xandri (2020): "Testing models of social learning on networks: Evidence from two experiments," *Econometrica*, 88(1), 1–32.

Corazzini, L., F. Pavesi, B. Petrovich, and L. Stanca (2012): "Influential listeners: An experiment on persuasion bias in social networks," *European Economic Review*, 56(6), 1276–1288.

Dasaratha, K., and K. He (2020): "Network structure and naive sequential learning," *Theoretical Economics*, 15(2), 415–444.

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

DeMarzo, P. M., D. Vayanos, and J. Zwiebel (2003): "Persuasion bias, social influence, and unidimensional opinions," *Quarterly Journal of Economics*, 118(3), 909–968.

Denti, T. (2017): "Network effects in information acquisition," working paper, available at https://drive.google.com/file/d/1iAEILmFa5F_ihdaZSVQMnIWX235b3Nfs/view?usp=sharing.

Fagiolo, G., J. Reyes, and S. Schiavo (2010): "The evolution of the world trade web: a weighted-network analysis," *Journal of Evolutionary Economics*, 20(4), 479–514.

Fricke, D., and T. Lux (2015): "Core–periphery structure in the overnight money market: evidence from the e-mid trading platform," *Computational Economics*, 45(3), 359–395.

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

Golub, B., and M. O. Jackson (2010): "Naive learning in social networks and the wisdom of crowds," *American Economic Journal: Microeconomics*, 2(1), 112–49.

Golub, B., and M. O. Jackson (2012): "How homophily affects the speed of learning and best-response dynamics," *Quarterly Journal of Economics*, 127(3), 1287–1338.

Grimm, V., and F. Mengel (2020): "Experiments on belief formation in networks," *Journal of the European Economic Association*, 18(1), 49–82.

Mailath, G. J. (1998): "Do people play Nash equilibrium? Lessons from evolutionary game theory," *Journal of Economic Literature*, 36(3), 1347–1374.

Molavi, P., A. Tahbaz-Salehi, and A. Jadbabaie (2018): "A theory of non-Bayesian social learning," *Econometrica*, 86(2), 445–490.

Monderer, D., and L. S. Shapley (1996): "Potential games," *Games and Economic Behavior*, 14(1), 124–143.

Mueller-Frank, M., and C. Neri (2013): "Social learning in networks: Theory and experiments," working paper, available at SSRN: *https://ssrn.com/abstract=2328281*.

——— (2021): "A general analysis of boundedly rational learning in social networks," *Theoretical Economics*, 16(1), 317–357.

Myatt, D. P., and C. Wallace (2019): "Information acquisition and use by networked players," *Journal of Economic Theory*, 182, 360–401.

Samuelson, L. (2002): "Evolution and game theory," *Journal of Economic Perspectives*, 16(2), 47–66.

Appendix: Quasi-Bayesian Foundation

Regarding agents' cognitive sophistication, this paper follows the boundedly rational approach, which assumes that agents have limited cognitive resources and do not possess precise knowledge of their environment. Nonetheless, it is useful to connect the assumptions of this paper to the standard Bayesian framework. In this appendix, we provide a pure theoretical motivation for DeGroot updating in networks, following DeMarzo, Vayanos, and Zwiebel (2003). DeGroot updating can be viewed as a Bayesian updating process for agents that receive normally distributed signals but do not know the true variances of their neighbors' signals.

Consider n agents who want to estimate some unknown parameter $\mu \in \mathbb{R}$. Agent i receives an independent signal $x_i^0 \sim \mathcal{N}(\mu, \sigma_i^2)$, and she assigns some precision $\pi_{ij} = \frac{1}{\mathrm{Var}_i(x_j^0)}$ to agent j's signal, which may or may not be the true precision. Note that this assumption does not align with the standard Bayesian approach, which assumes that agents have precise knowledge of the signal structure. Agents communicate according to a social network $\tilde{\mathbf{G}}$, which is a directed graph that indicates whether agent i listens to agent j; $\tilde{g}_{ij} = 1$ if agent i listens to agent j, and $\tilde{g}_{ij} = 0$ otherwise. Each agent knows her own information, so $\tilde{g}_{ii} = 1$. Truthful reporting is assumed. Given normality and the assigned precisions, a sufficient statistic for the signals is their weighted average, with weights given by the precisions. DeMarzo, Vayanos, and Zwiebel (2003) denote such a statistic by x_i^1, and refer to it as agent i's *belief* after communication:

$$x_i^1 = \sum_{j=1}^{n} \frac{\tilde{g}_{ij}\pi_{ij}}{\sum_{j=1}^{n} \tilde{g}_{ij}\pi_{ij}} x_j^0.$$

The sufficiency of the statistic x_i^1 comes from the application of the Fisher-Neyman factorization theorem. Defining $g_{ij} := \sum_{j=1}^{n} \frac{\tilde{g}_{ij}\pi_{ij}}{\sum_{j=1}^{n} \tilde{g}_{ij}\pi_{ij}}$, we obtain the stochastic matrix $\mathbf{G} = (g_{ij})$. A DeGrootian population communicating according to $\mathbf{G}$ holds the same beliefs as the quasi-Bayesian population from DeMarzo, Vayanos, and Zwiebel (2003).[17] This insight provides additional motivation for the model described in this paper.

17 We say quasi-Bayesian because the critical assumption of potentially misperceived variances is not standard Bayesian. In a fully Bayesian world, agents would know the true precisions, and hence $\pi_{ij} = \pi_{kj} = \frac{1}{\mathrm{Var}(x_j^0)}$ for all i, k. This would imply $g_{ij} = g_{kj}$ in the equivalent DeGrootian network.