GOVERNANCE AND THE EFFICIENCY OF ECONOMIC SYSTEMS
GESY

Discussion Paper No. 485

Learning faster or more precisely? Strategic experimentation in networks

Mirjam Wuggenig*

* Bonn Graduate School of Economics

December 2014

Financial support from the Deutsche Forschungsgemeinschaft through SFB/TR 15 is gratefully acknowledged.

Sonderforschungsbereich/Transregio 15 · www.sfbtr15.de
Universität Mannheim · Freie Universität Berlin · Humboldt-Universität zu Berlin · Ludwig-Maximilians-Universität München
Rheinische Friedrich-Wilhelms-Universität Bonn · Zentrum für Europäische Wirtschaftsforschung Mannheim

Speaker: Prof. Dr. Klaus M. Schmidt · Department of Economics · University of Munich · D-80539 Munich,
Phone: +49(89)2180 2250 · Fax: +49(89)2180 3510

Learning faster or more precisely? Strategic experimentation in networks*

Mirjam Wuggenig†

December 4, 2014

Abstract

The paper analyzes a dynamic model of rational strategic learning in a network. It complements existing literature by providing a detailed picture of short-run dynamics in a game of strategic experimentation where agents are located in a social network. We show that the delay in information transmission caused by incomplete network structures may induce players to increase own experimentation efforts. As a consequence a complete network can fail to be optimal even if there are no costs for links. This means that in the design of networks there exists a trade-off between the speed of learning and accuracy.

Key Words: Strategic Experimentation, Networks, Learning

JEL codes: C73, D83, D85

*I would like to thank Sven Rady for helpful insights and discussions. Further, I would like to thank the participants of the SFB/TR 15 Young Researcher Workshop Spring 2013, the Bonn Micro Workshop as well as the Spring Meeting of Young Economists 2014 for valuable comments. Financial support from the Deutsche Forschungsgemeinschaft through SFB/TR 15 is gratefully acknowledged.

†Bonn Graduate School of Economics, Kaiserstraße 1, 53113 Bonn, Germany. Email: wuggenig@uni-bonn.de

"Some people will never learn anything, for this reason, because they understand everything too soon." Alexander Pope

1 Introduction

The experience of others plays an important role when individuals have to take decisions about alternatives that they cannot perfectly evaluate themselves. For example, in situations of product choice, a person will base his or her decision on own past experiences, ask friends and coworkers about their opinions, and additionally collect information via other sources as for instance, customer reviews on the internet. One way to model learning situations where people have to take decisions under uncertainty is by so called bandit models (see e.g., Bolton & Harris, 1999 or Keller, Rady and Cripps, 2005 [KRC, hereafter]). The idea of these models is that players have to choose between different options (different arms of a bandit machine) under imperfect knowledge of their relative advantage, that is, the outcomes of the arms are uncertain. By playing repeatedly, the agents can learn about the type of the arm, however, this learning or experimentation is costly as future payoffs are discounted. Such bandit models can provide a framework to discuss different (economic) situations as e.g., specific problems of product choice, research or innovation.

So far, most models of strategic experimentation assume that agents interact with everyone else in society. That is, each agent can observe or communicate with the same set of other individuals and as actions and payoffs are publicly observable, a common belief about the state of the world prevails. This assumption will be relaxed by letting agents interact directly only with a subset of agents that is determined by the structure of connections in a (social) network. This extension of the model is thought to better reflect interaction patterns in reality, where without doubt the structure of relationships in shaping beliefs and opinions plays an important role. Empirical work in economics highlights the impact of network structures in labor markets (e.g., regarding information about job vacancies (see e.g., Calvo-Armengol and Jackson, 2004)) or finds evidence of the importance of interaction patterns in learning about a new technology (see e.g., Conley and Udry, 2010). Learning and innovation are influenced by the structures of communication and information exchange between different sources. In the field of research, workshops and conferences bring together researchers from dispersed geographical regions and different fields of specialization to enable exchange of ideas. Similarly, innovation plays an important role for firms to secure competitiveness, and the structure of information exchange between subsidiaries of multinational organizations might be a key to success. For example, Nobel and Birkinshaw (1998) analyze communication patterns between subsidiaries of multinational corporations in international R&D operations and find that innovation is associated with comparatively high levels of communication within the firm and outside. Teece (1994) emphasizes the importance of organizational structures that enable an easy flow of communication between business units and guarantee a high speed of learning.

The aim of our model is to provide insight into how the structure of relations influences the evolution of beliefs, decisions and incentives of rational agents who need to acquire information. More precisely, we consider a dynamic model of strategic learning in which individuals can generate own information (through experimentation) and obtain information through interaction (observation and verifiable message exchange) with others. The interaction possibilities are determined by specific social interaction structures. As a basis for the model, a discrete time version of the exponential bandit model by KRC as in Heidhues, Rady and Strack (2012) [HRS hereafter] will be used. In this model agents can choose between a safe option and a risky option. The payoff of the risky alternative depends on the state of the world which can be either good (i.e., generate higher payoffs than the safe option) or bad. Agents base their decision on their belief, i.e., the probability attached to the good state of the world and their beliefs depend on their observations and hence the interaction structure. This interaction structure will be fixed and imposed on the agents before the game starts.

First, we characterize symmetric equilibria in Markovian strategies of the strategic experimentation game in three different network structures, the complete network, the ring and the star network. Further, experimentation intensities in equilibrium are compared across these structures. It is shown that in a network structure in which agents learn from unobserved players (neighbors of neighbors) with a delay, players increase their experimentation intensity or effort to compensate for the slightly worse possibility to learn from others. Depending on the structure and the belief, agents are able to fully outweigh this loss and thereby keep expected utilities unaltered compared to interaction in a complete network. The agents' strategies depend on their beliefs and there exists an upper cut-off belief above which agents experiment with full intensity and a lower cut-off below which experimentation ceases. These cut-off beliefs depend on the network structure as well as time and take into account whether agents still expect information that was generated by unobserved individuals to arrive. Specialization, where one player does not experiment while others do, arises in networks where agents are not symmetric with respect to their position as in the star network. While for some beliefs specialization can be beneficial for society, it is detrimental to welfare for others.

The obtained results offer insights into the incentives that drive the behavior of rational agents. Taking research or innovation as examples, a welfare analysis of the model provides insights which might be relevant to government authorities or companies for structuring and subsidizing research projects. Objectives of decision makers can be manifold as for instance, cost minimization, utility maximization, the maximization of the speed of learning through fast information transmission or completeness (that is, more precise or accurate learning which implies that the probability of mistakenly abandoning a good risky arm is minimized). Interestingly, even if there are no costs for links there exists an interval of beliefs for which the complete network does not generate highest payoffs. As part of a welfare comparison, we observe a trade-off between interaction structures that enable a fast speed of learning and structures in which learning is more precise. How this trade-off is resolved depends on the discount factor.

In strategic experimentation models, where agents can observe the outcomes and actions of others, strong incentives to free-ride on the experimentation effort of others exist and prevent the socially optimal outcome (see e.g., Bolton and Harris, 1999 or KRC). Bimpikis and Drakopoulos (2014) show that (full) efficiency can be obtained in the model of KRC if information is aggregated and released with an optimal delay. As we will see, a network structure also causes time lags in the information transmission that can increase experimentation efforts and mitigate free-riding. The structure of the model further allows us to analyze the results in the context of organizational design thereby offering insights about optimal organization or communication structures.

The paper contributes to the theory of rational strategic learning in networks and aims to fill the gap between static models and dynamic models that focus on long-run results and conditions for complete learning. Due to the complexity that network settings can create, attention was often restricted to the behavior of myopic or boundedly rational agents to ensure tractability.[1] In a recent contribution Sadler (2014) analyzes a strategic experimentation problem as in Bolton and Harris (1999) in a network setting with boundedly rational agents. In this model, each player assumes that her neighbors have the same belief as she does and players do not learn anything from the actions of neighbors (and consequently agents do not draw inferences about actions or outcomes of neighbors of neighbors).

[1] See e.g. Jackson (2008), Chapter 8 or Goyal (2009), Chapter 5 for different types of learning models in a network setting.

In the field of rational learning, there are several examples of Bayesian learning models focusing on asymptotic long-run results and conditions for complete learning or convergence of actions or payoffs (see e.g. Gale and Kariv, 2003, Acemoglu and Ozdaglar, 2011, Rosenberg, Solan and Vieille, 2009, Acemoglu, Bimpikis and Ozdaglar, 2014). These results, however, offer little insight into how social relations shape incentives in early stages of the learning process.

The paper most closely related to our model is Bramoullé and Kranton (2007) [BK hereafter] who investigate a public goods game in a network and are able to draw conclusions about short run effects in a static framework. The authors show that networks can lead to specialization and that this specialization can have welfare benefits, features that are confirmed within our framework. One difference between the model of BK and our model is that in games of strategic experimentation only informational externalities are present and no payoff externalities. Moreover, in BK's framework an agent's effort is a substitute for the effort of neighbors but not neighbors of neighbors meaning that they focus on local public goods and most importantly, BK consider a static setting. As learning, innovation and research have a dynamic character, a dynamic perspective might be better suited to analyze these processes. Such a perspective yields additional insights concerning the updating rules agents use, the effects of different beliefs within a society that are a consequence of asymmetric positions, and the impact of network structures on the speed and accuracy of learning.

The paper is structured as follows: Section 2 introduces the basic model. In Section 3 the complete network is analyzed to set up a benchmark case for future comparison. Section 4 analyzes a simple incomplete interaction structure, namely a ring, to see how spatial structures change the problem at hand. In Section 5, the star network as the simplest irregular network is considered, before a welfare analysis is conducted in Section 6. Section 7 contains a discussion and conclusion. All proofs are relegated to the Appendix.

2 Model

First, we describe the underlying bandit model. After that, main concepts of the network structure are outlined and the timing and information structure are specified. With the help of a short example we briefly show how a network structure affects updating rules. Finally, strategies as well as the equilibrium concept are discussed.

2.1 A two-armed bandit model

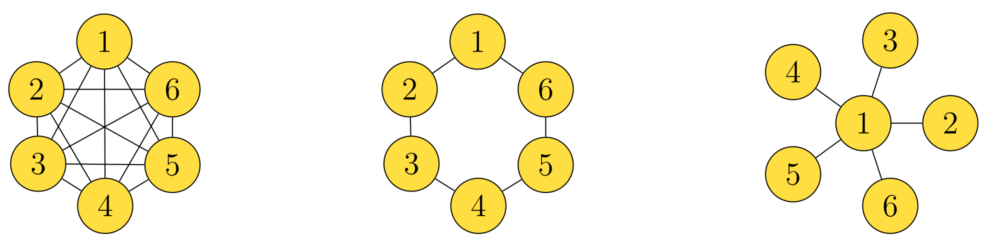
The model is based on the two-armed exponential bandit model as described by KRC or more specifically the discrete time version thereof by HRS. There are agents $i \in N$ and we denote the cardinality of N by n. Time is discrete, $t = 1, 2, ...$ and players discount future payoffs by a common discount factor $\delta \in (0, 1)$. Agents can decide how much effort to invest in each of two projects, which correspond to two different arms of a bandit machine. The safe arm yields a fixed deterministic payoff normalized to 0. The second arm is risky (denoted by R) with an uncertain payoff $X_i(t)$.

The distribution of the risky payoffs is independent across players and time and only depends on the state of the world, which is either good ($\theta = 1$) or bad ($\theta = 0$). If it is bad, then R yields a low payoff X_L, if it is good, then it yields either a high payoff X_H or a low payoff X_L, where $X_L < 0 < X_H$. The probability of receiving a high payoff is zero if the arm is bad and $P(X_H|\theta = 1) = \pi > 0$ if it is good. Consequently, the first high payoff realization (also called a breakthrough) perfectly reveals that the risky arm is good. The conditional expectation $E[X_i(t)|\theta]$ of the risky payoff in any given period is denoted by E_θ and additionally to the fact that $E_0 < 0$ we assume that $E_1 > 0$, which means that it is optimal for the players to use the risky arm if $\theta = 1$ and the safe arm if $\theta = 0$.

Players hold a belief about the risky arm being good, and it is assumed that everyone starts with a common prior. The belief, denoted by p, depends on the arrival of a breakthrough and is therefore a random variable. The agents influence each other only through the impact of their action on the belief of others, meaning there are only informational externalities and no payoff externalities. In the model of HRS and KRC, all players interact with everyone else in the society and hence all agents hold a common posterior belief. This will no longer be true in our model, where players interact only with a subset of society. A player's belief about the risky arm being good depends on whether she learns about a breakthrough or not. Once an agent learns about a breakthrough, her uncertainty about the type of the arm is fully resolved and the posterior belief jumps to 1. As long as agents experiment without learning about a breakthrough, they update their belief according to Bayes' rule and the belief about the risky project being good decreases. Players are said to experiment if they use the risky arm before knowing its type.

2.2 Introducing a network structure

Given a set of nodes N (representing individuals), a network or graph g is an $n \times n$ interaction matrix that represents the relationships in the society. The typical element is denoted by $g_{ij} \in \{0, 1\}$. If $g_{ij} = 1$, a link between i and j exists and implies that these two individuals can interact with each other, i.e., exchange information about actions and outcomes. The matrix is symmetric ($g_{ij} = g_{ji}$), meaning links are undirected, and always has 1 on the main diagonal (every individual can observe his or her own actions and outcomes, i.e., $g_{ii} = 1$ for all i). The structure of relations is assumed to be common knowledge. If a link between two individuals exists, those agents are considered as neighbors. The *neighborhood* of agent i is denoted by N_i and defined as $N_i(g) = \{j \neq i : g_{ij} = 1\}$. The number of direct neighbors, $\#N_i$, is called the *degree* d_i of agent i. Subsequently a fixed interaction structure g will be assumed. The game is analyzed in three different network structures: the complete network[2] as a benchmark case; the ring, an incomplete but regular[3] structure; and the star network with one player in the center and all other $n - 1$ players only connected to the central player (see Figure 1).



(a) Complete Network (b) Ring Network (c) Star Network

Figure 1: Network structures for $n = 6$.

2.3 Actions, information structure and timing

Agents are endowed with one unit of (perfectly divisible) effort each time period $t = 1, 2, ...$ that can be allocated between the two projects. The experimentation effort $\phi_i(t) \in [0, 1]$ corresponds to the fraction of the unit resource that is allocated to the risky arm and $1 - \phi_i(t)$ is allocated to the safe arm. That is, we write $\phi_i(t) = 1$ if the agent uses the risky arm exclusively and $\phi_i(t) = 0$ if the safe arm is used exclusively. To provide a distinction between information generated in the first and in the second round of the game we will also denote first round experimentation effort by $\alpha_i \in [0, 1]$, that is $\phi_i(1) = \alpha_i$, and second round experimentation effort by $\beta_i \in [0, 1]$. Here β_i is the experimentation effort conditional on not having observed a breakthrough in the first round.

The game can be divided into two stages. In the first stage, which consists of periods

[2] A complete network is a network in which every agent is connected to everyone else.

[3] Regular networks are networks where all players have the same number of neighbors, i.e. $d_i = d_j$ for all $i, j \in N$.

1 and 2, agents can experiment and transmit information. In the second stage starting from $t = 3$, they cannot experiment anymore but still exchange information about what they observed. More precisely, for $t \geq 3$ we have $\phi_i(t) = 0$ if agent i did not learn about a success up to time t and $\phi_i(t) = 1$ otherwise. Restricting the time horizon for experimentation to two periods is sufficient to capture the dynamic perspective relevant to the learning process. Further, it reflects the assumption that agents cannot engage in possibly infinitely long experimentation; research funds for example, are typically granted over a period of a few years. Information that reaches agents later is still relevant as they can switch back from the safe to the risky option in case they learn or know this is optimal.

The timing is as follows: agents start in $t = 1$ with a common prior belief $p(1)$. Each agent chooses an experimentation intensity or effort for the first round $\alpha_i \in [0, 1]$, determining whether the risky or safe option is chosen. At the end of the first round players observe their own outcomes as well as actions and outcomes of their neighbors and update their prior accordingly to $p_i(2)$. Those agents who have not observed a success choose $\beta_i \in [0, 1]$ for the second round, observe outcomes and actions in their neighborhood and exchange verifiable reports about previous experiments by unobserved agents, i.e., in $t = 2$ agent i knows α_h as well as $X_h(1)$ for all agents $h \in N_j \backslash N_i$ where $j \in N_i$.[4] From $t = 3$ onwards agents cannot experiment anymore, but still exchange reports.

Formally, agent i's information at a given point in time consists of

$$\mathcal{I}_i(t) = \{\mathbf{H}_i(t), \mathbf{r}_i(t)\},$$

where

$$\mathbf{H}_i(t) = \{\phi_i(1), X_i(1), ..., \phi_i(t), X_i(t)\}$$

is the complete history of actions and outcomes for agent i up to time t and $\mathbf{r}_i(t) = (r_i(1), ..., r_i(t))$ is the history of reports agent i received. Each element $r_i(t)$ is a vector that contains for each agent $j \in N_i$ the history $\mathbf{H}_j(t)$ up to this point in time as well as the reports j received up to $t - 1$, i.e., $\mathbf{r}_j(t - 1)$.

The main difference between complete and incomplete network structures lies in the fact that as soon as the network is incomplete at least some of the agents do not possess

[4] Reports are assumed to be verifiable so that agents have no possibility to lie. If agents are allowed to freely choose any message, they may find it optimal to report a breakthrough although there was none in order to induce additional experiments. See HRS for a strategic experimentation game in which payoffs are privately observed and agents can exchange cheap talk messages.

complete information about (past) actions and payoffs of others. Consequently, when interacting with their neighbors, agents obtain information from them about (past) actions and payoffs of unobserved agents and use this information to make inferences about the true state of the world. In the complete network the probability that agent i observes a breakthrough is equal to the probability that there is a breakthrough. In incomplete networks the probability of learning about a breakthrough at a given point in time depends on the entire structure of relations, and information about a breakthrough will travel along the paths in the network. This implies that players will not necessarily hold a common belief about the state of the world. We will illustrate the impact of the network structure on the updating of beliefs with the help of a short example.

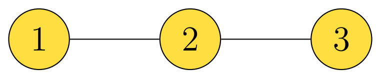


Figure 2: The star network, $n = 3$.

Example 1 There are three agents $i = 1, 2, 3$, whose connections can be described by the following interaction matrix (see Figure 2)

$$g = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

As $g_{2j} = 1$ for all $j \in N$, agent 2 has complete information and can observe all actions and payoffs at any point in time. The other two agents only observe agent 2 and their own actions and payoffs, and they receive information through agent 2. When agents experiment, they either generate a breakthrough or not. In our example a successful discovery by agent 2 immediately reveals to everyone that the risky arm is good. If agent 1 has a breakthrough, only 1 and 2 know about it. However, agent 2 informs agent 3 about the breakthrough so that agent 3 knows about it one period later. As long as there is no breakthrough, all agents update their beliefs depending on how many unsuccessful experiments they learn about. If the risky arm is good, the probability that player i observes a breakthrough if she experiments with intensity $\phi_i(t)$ is given by $\phi_i(t)\pi$. Taking the experimentation effort of others into account, player 2 updates her belief according to

$$p_2(t+1) = \frac{p_2(t)\prod_{i=1}^{3}(1 - \phi_i(t)\pi)}{p_2(t)\prod_{i=1}^{3}(1 - \phi_i(t)\pi) + 1 - p_2(t)}$$

if no breakthrough occurs, where $\prod_{i=1}^{3}(1-\phi_i(t)\pi)$ reflects the experiments conducted by 2 and her neighbors. The numerator is the probability of not observing a breakthrough at a good risky arm and the denominator gives the total probability of not observing a breakthrough. In case of a breakthrough the posterior jumps to 1. Player 1 updates her belief according to

$$p_1(t+1) = \frac{p_1(t)\prod_{i=1}^{2}(1-\phi_i(t)\pi)(1-\phi_3(t-1)\pi)}{p_1(t)\prod_{i=1}^{2}(1-\phi_i(t)\pi)(1-\phi_3(t-1)\pi)+1-p_1(t)},$$

if there is no breakthrough. At time t agent 1 observes the outcome of her own experiment, as well as agent 2's experiment. While agent 1 does not observe agent 3's current experiment she gets informed about the experiment performed in $t-1$, which is captured by the term $(1-\phi_3(t-1)\pi)$. Agent 3's belief at time t is derived analogously.

In the example player 1 and player 3 are only connected via player 2. In general, more complex interaction structures (for example, large irregular structures with overlapping neighborhoods) are possible, so that the computations agents have to make are getting considerably more complicated. How many experiments agents learn about and hence the updating of beliefs, depends on time and on location.

2.4 Strategies and equilibrium concept

At $t = 1, 2$ players are restricted to (pure) Markovian strategies in that $\phi_i(t)$ can depend on the belief $p_i(t)$ and time t only. For $t \geq 3$, behavior is restricted as already described in Section 2.3. In round t, agent i obtains a payoff of $\phi_i(t)X_i(t)$ and player i's total expected (normalized and discounted) payoff is given by

$$(1-\delta)E\left[\sum_{t=1}^{\infty}\delta^{t-1}\phi_i(t)X_i(t)\right], \tag{1}$$

where the expectation is taken with respect to $p_i(t)$ and $\phi_i(t)$. The reason why the sum over all payoffs goes from 1 to ∞ (and not only until 2) is that agents are allowed to switch to the risky arm if information about a breakthrough reaches them at a later date (see Section 2.3). The solution concept is *Markov perfect equilibrium.* In what follows we restrict attention to equilibria in which agents who are symmetric with respect to their position in a network use symmetric strategies.

3 The Single Agent and the Complete Network

Before we analyze the generation and diffusion of information in incomplete networks we first explain how the experimentation problem is solved by a single agent. After that, we look at the model with n agents, where each individual can observe everyone else. Expressed in terms of networks this corresponds to the empty and the complete network.

3.1 The single agent problem

This section outlines the results that are obtained if a single agent faces the two armed bandit problem described above. For a single player posterior beliefs are determined by Bayes' rule according to

$$p_i(t+1) = \frac{p_i(t)(1-\phi_i(t)\pi)}{1-p_i(t)\phi_i(t)\pi} \tag{2}$$

if she does not observe a breakthrough. The posterior jumps to 1 after a success. The single agent's maximization problem for two rounds of experimentation is

$$\max_{\alpha_i,\beta_i\in[0,1]} U(p(1),\alpha_i,\beta_i),$$

where

$$\begin{aligned} U(p(1),\alpha_i,\beta_i) &= \alpha_i[(1-\delta)E_{p(1)} + \delta E_1 p(1)\pi] + \\ &\quad \delta(1-p(1)\alpha_i\pi)\beta_i[(1-\delta)E_{p(2)} + \delta E_1 p(2)\pi], \end{aligned} \tag{3}$$

with $E_p = E_1 p + (1-p)E_0$ and $p(2)$ given by (2). In expression (3) the first part, $\alpha_i(1-\delta)E_{p(1)}$, is the expected and normalized current payoff the agent obtains at $t=1$ by exerting effort α_i. A good risky arm generates a payoff of E_1, while a bad risky arm gives E_0. The remaining terms represent the discounted expected continuation payoff. The continuation payoff depends on the occurrence of a breakthrough and is E_1 with the probability $\alpha_i p(1)\pi$ that the risky arm is good and a breakthrough occurs. After a breakthrough the agent knows that the risky arm dominates the safe arm and hence she will continue to play risky forever. If the agent does not observe a success she can experiment one more time in $t=2$. The probability of not observing a breakthrough consists of the probability that the risky arm is bad, $1-p(1)$, and the probability that it is good, but the agent nevertheless did not have a breakthrough, $p(1)(1-\alpha_i\pi)$. As there are only two rounds of experimentation, the agents' continuation payoff after a second unsuccessful experiment is zero forever.

The agent solves

$$\max_{\alpha_i,\beta_i\in[0,1]} U(p(1),\alpha_i,\beta_i),$$

where $U(p(1),\alpha_i,\beta_i)$ is given by (3). Replacing $p(2)$ by $\frac{p(1)(1-\alpha_i\pi)}{1-\alpha_i p(1)\pi}$ in (3) we can see that the objective function is linear in both choice variables α_i and β_i. This implies that the solution to the maximization problem is located on the boundaries of the $[0,1]\times[0,1]$ square. Actually, for the case considered here only three points for (α,β), namely $(0,0)$, $(1,0)$ and $(1,1)$, can be optimal. It is easy to show that it is never optimal for an agent to delay experimenting, that is, to play $(0,1)$. The reason for this is simply that future payoffs are discounted, which implies that if an agent wants to experiment, then it is optimal to do so earlier rather than later. Which strategy is optimal depends on $p(1)$ and can be easily found by comparing expected utilities for each pair of actions. The action profile $(0,0)$ is optimal for any $p(1)$ smaller than

$$p^a = \frac{(1-\delta)|E_0|}{(1-\delta)[|E_0|+E_1]+\delta E_1\pi}, \tag{4}$$

where a stands for autarky. As long as $p(1)\geq p^a$, the expected payoff from experimenting is positive which means that the risky arm is preferred. If an agent experiments without success her belief declines and as expected payoffs are increasing in beliefs, per period payoffs from experimenting also decrease over time if there is no breakthrough. Agents stop experimenting at a belief $p^a>0$, which means that it is possible that they abandon the risky project although it is good. The cut-off belief p^a decreases in δ, which means that as agents are getting more patient, complete learning becomes more likely. That is, the final posterior belief is smaller and hence the probability of mistakenly switching from the risky to the safe arm although the risky arm is good, decreases. Finally, the prior belief above which it is optimal to experiment in both rounds is

$$p(1) = \frac{(1-\delta)|E_0|}{(1-\delta)|E_0|+E_1[1-\delta-\pi+\delta\pi(2-\pi)]},$$

for which the corresponding posterior after one failed experiment equals p^a. Figure 3 depicts the relationship between belief and optimal experimentation effort.

Each round a single agent either uses the safe arm exclusively or the risky arm exclusively, and her actions only depend on the belief in the given round, that is, in $t=1,2$

$$\phi_i^a(t) = \begin{cases} 1 & \text{for } p(t)\geq p^a, \\ 0 & \text{otherwise.} \end{cases}$$

The single agent's strategy does not depend on time t for two reasons. First, even though experimentation is not possible for $t\geq 3$, agents can still use the risky arm (and obtain

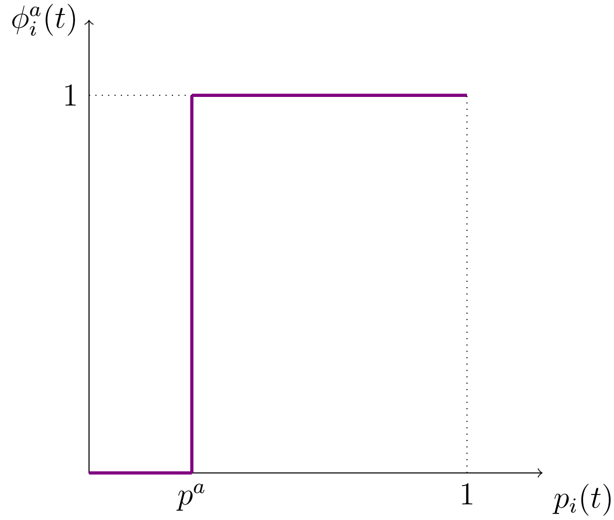


Figure 3: Optimal experimentation effort of a single agent ($\pi = 0.2$, $\delta = 0.9$, $E_1 = 1$, $E_0 = -1$ and $p^a \approx 0.26$).

high payoffs) in case this is known to be optimal (i.e., the game does not end after $t = 2$). Second, for a single agent all payoff-relevant information is captured by the current belief. We will see that time only matters when agents still expect information that was generated by unobserved agents to reach them at a later date.

3.2 Strategic experimentation in a complete network

Let us now consider agents interacting strategically in a complete network. Each player maximizes her expected utility over α_i and β_i given her belief and the strategies of the other players, α_{-i}, β_{-i}. We denote $\prod_{i=1}^{n}(1-\alpha_i\pi)$ by $\widetilde{\alpha}$ and $\prod_{i=1}^{n}(1-\beta_i\pi)$ by $\widetilde{\beta}$. Each player solves

$$\max_{\alpha_i,\beta_i\in[0,1]} U(p(1),\alpha_i,\beta_i,\alpha_{-i},\beta_{-i})$$

where

$$\begin{aligned} U(p(1),\alpha_i,\beta_i,\alpha_{-i},\beta_{-i}) &= \alpha_i(1-\delta)E_{p(1)} + \delta E_1 p(1)(1-\widetilde{\alpha}) + \\ &\quad \delta\left(1-p(1)(1-\widetilde{\alpha})\right)\left(\beta_i(1-\delta)E_{p(2)} + \delta E_1 p(2)(1-\widetilde{\beta})\right) \quad (5) \end{aligned}$$

and

$$p(2) = \frac{p(1)\widetilde{\alpha}}{p(1)\widetilde{\alpha}+1-p(1)}.$$

The difference to the problem of a single agent is that the continuation payoff of agent i now also depends on the actions of the other players. Proposition 1 describes the optimal experimentation effort in a symmetric equilibrium.

Proposition 1. *In a symmetric equilibrium in a complete network with n agents, the common strategy in $t = 1, 2$ is given by*

$$\phi^c(t) = \begin{cases} 1 & \text{for } p(t) \in [\overline{p}^c, 1], \\ \frac{1}{\pi} - \frac{1}{\pi}\left(\frac{(1-\delta)|E_0|}{\delta E_1 \pi p(t)} - \frac{(1-\delta)(|E_0|+E_1)}{\delta E_1 \pi}\right)^{\frac{1}{n-1}} & \text{for } p(t) \in (p^a, \overline{p}^c), \\ 0 & \text{for } p(t) \in [0, p^a], \end{cases} \tag{6}$$

where

$$\overline{p}^c = \frac{(1-\delta)|E_0|}{(1-\delta)[|E_0| + E_1] + \delta E_1 \pi (1-\pi)^{n-1}}.$$

We can see that there exists an interval of beliefs such that in a symmetric equilibrium players want to simultaneously use both arms. In this interval α^c is chosen such that agents are indifferent between the risky and the safe arm. There now exists an upper cut-off belief, $\overline{p}^c$, which is the belief above which agent i experiments with intensity 1 even if all others also experiment with full intensity. Starting from $\overline{p}^c$ agents decrease their experimentation intensity as the belief decreases, up to the point where $\phi^c(t) = 0$, which holds for any belief below p^a. Figure 4 depicts the equilibrium strategy.[5]

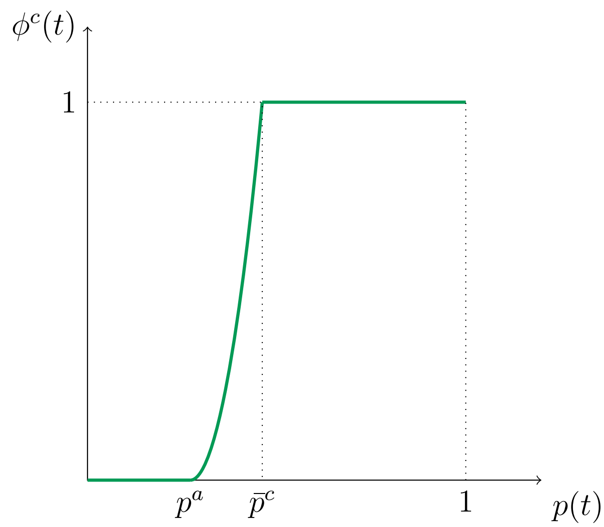


Figure 4: Equilibrium experimentation effort in a complete network ($\pi = 0.2$, $\delta = 0.9$, $n = 12$, $E_1 = 1$, $E_0 = -1$, $p^a \approx 0.26$ and $\overline{p}^c \approx 0.46$).

Several features of the equilibrium experimentation strategy are worth noting. First, there is at most one round in which agents simultaneously want to use both arms. In fact, n failed experiments from $\overline{p}^c$ generate a posterior belief below p^a, and the effort $\phi^c(p(t))$

[5]Note that depending on the parameters of the model the relationship between effort and belief can be convex, concave or both. Effort is, however, monotonically increasing in the agents' belief.

of beliefs $p(t) \in (p^a, \bar{p}^c)$ causes the posterior to fall below the single agent cut-off if there is no success. Second, as in the single agent case, agents do not have an incentive to delay any experiments. This implies that in case $\beta^c > 0$, we know that $\alpha^c = 1$. Third, the upper cut-off $\overline{p}^c$ is increasing in n, whereas the lower threshold is given by p^a. Social optimality requires experimentation beyond the single agent cut-off since agents benefit from each other's experimentation effort. However, agents do not experiment below p^a and even stop experimenting with full intensity earlier with an increasing number of agents, which is a particularly stark manifestation of the free-riding effect (see HRS or KRC).

4 The Ring Network

Having set up the complete network as the benchmark case, we now turn to the strategic experimentation problem when agents are located on a ring. In the ring network every agent has two direct neighbors. As players are symmetric there again exists a symmetric equilibrium. Individuals experiment and communicate, i.e., exchange messages about observations made in their neighborhood as well as messages received from their neighbors. The underlying structure is illustrated in Figure 1b for $n = 6$ and implies that agent 1 is informed about the outcomes of the experiments of agent 3 with one period delay through communication with agent 2 and so on.

In the ring network we have to distinguish between an even and an odd number of players, as this determines how much information arrives in the last round where new information reaches agent i. As the results are similar in both cases we will only discuss the case where n is odd here. Expected payoffs for a given prior and strategy profile are

$$\begin{aligned} U(p(1), \alpha_i, \beta_i, \alpha_{-i}, \beta_{-i}) &= \alpha_i(1-\delta)E_{p(1)} + \delta E_1 p(1)[1-(1-\alpha_j\pi)^2(1-\alpha_i\pi)] \\ &\quad +\delta[1-p(1)+p(1)(1-\alpha_j\pi)^2(1-\alpha_i\pi)]u(p(2)); \end{aligned}$$

as we are solving for symmetric equilibria, we are assuming here that all agents $j \neq i$ use the same strategy. The term

$$\begin{aligned} u(p(2)) &= \beta_i(1-\delta)E_{p(2)} + \delta E_1 p(2)[1-(1-\beta_j\pi)^2(1-\beta_i\pi)(1-\alpha_j\pi)^2] \\ &\quad +\delta[1-p(2)+p(2)(1-\beta_j\pi)^2(1-\beta_i\pi)(1-\alpha_j\pi)^2]u(p(3)), \end{aligned}$$

as well as $u(p(3))$, $u(p(4))$ and so on, is determined by the information about past experiments traveling through the network. That is, with the probability that at least one experiment agent i learns about in period 1 is successful, $p(1)[1-(1-\alpha_j\pi)^2(1-\alpha_i\pi)]$, she expects to get a continuation payoff of E_1 from the next period onwards. These are the two experiments of the neighbors as well as the own experiment. In case all these experiments were unsuccessful, in the next round she and her neighbors can again experiment and further there is the chance that neighbors of neighbors had a breakthrough in the first round about which agent i will learn now. In $u(p(2))$, the factor $(1-\alpha_j\pi)^2$ in $1-(1-\beta_j^{\pi})^2(1-\beta_i^{\pi})(1-\alpha_j^{\pi})^2$ represents the experiments of neighbors of neighbors in $t=1$, $(1-\beta_j\pi)^2$ the two experiments of the direct neighbors in $t=2$, and $1-\beta_i\pi$ the own experiment in $t=2$.

Figure 5 illustrates how the game evolves for $n=4$. At the beginning the agents start with a prior belief $p(1)$, choose their optimal experimentation intensity α and receive the expected payoff for this round. Then the agent either knows that the state of the world is good if there was a breakthrough in her neighborhood, or she chooses her optimal experimentation intensity β based on her updated belief $p(2)$. The difference between the first and the second period is that in $t=2$ the agent does not only observe her own and her neighbors' experiments, but also receives information about the outcome of the first round experiment of the unobserved agent. In the round afterwards experimenting is no longer possible. However, there still might arrive information about a breakthrough of the unobserved agents. This process continues until all information has reached agent i which takes the longer the more players there are.

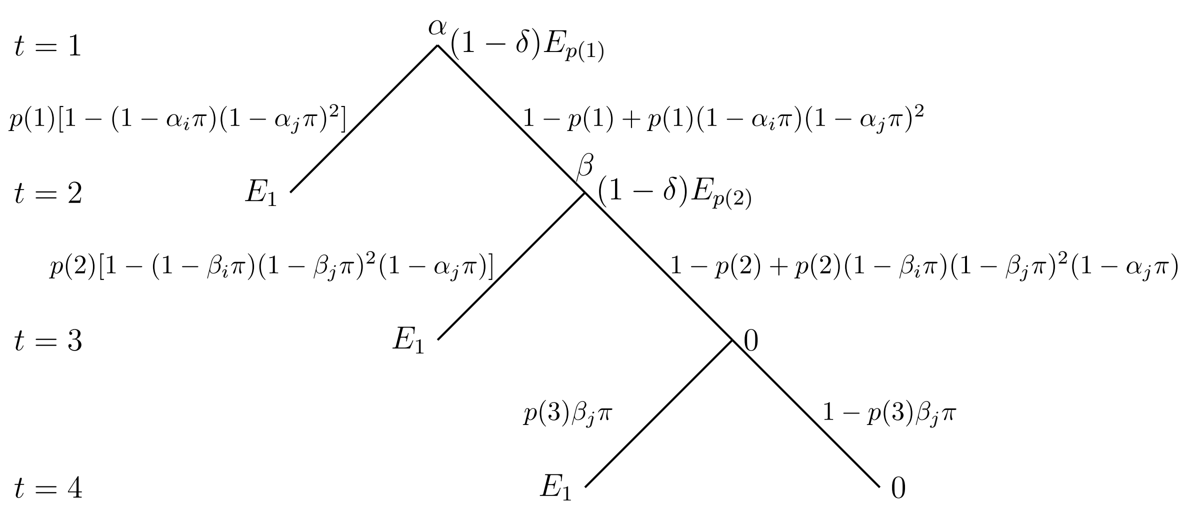


Figure 5: The strategic experimentation game in a ring network, $n=4$.

In contrast to the complete network, it is important to highlight that first period and second period equilibrium cut-offs will be distinct. This can be ascribed to the fact that after one round of experimentation, information is traveling through the network and agents anticipate that this information will reach them. Apart from this, the problem is similar to the complete network. Expected payoffs are again linear in α_i and β_i and best responses and equilibrium cut-offs can be found by the same arguments as in the previous section. In order to analyze the equilibrium behavior of the agents we introduce the expressions I_1^r and I_2^r. I_1^r represents the difference in expected payoffs of action profiles $(1,0)$ and $(0,0)$ for symmetric actions of the other players. This means that $I_1^r > 0$ implies that payoffs from experimenting are higher than payoffs from not experimenting in $t = 1$ and at $I_1^r = 0$ agents are indifferent. I_2^r refers to the difference in expected payoffs between $(1,1)$ and $(1,0)$. The expressions for the I_1^r, I_2^r and equilibrium cut-off beliefs can be found in the Appendix. Proposition 2 summarizes the main points.

Proposition 2. *In a symmetric equilibrium in a ring network with an odd number of players, each player chooses the following action: at* $t = 1$

- $\alpha^r = 1$ *for* $p(1) \in [\overline{p}^r(1), 1]$,
- $\alpha^r = 0$ *for* $p(1) \in [0, p^a]$,
- $\alpha^r \in (0,1)$ *is defined uniquely by the root of* I_1^r *on* $[0,1]$ *for* $p(1) \in (p^a, \overline{p}^r(1))$.

At $t = 2$,

- $\beta^r = 1$ *for* $p(2) \in [\overline{p}^r(2), 1]$,
- $\beta^r = 0$ *for* $p(2) \in [0, \underline{p}^r(2)]$,
- $\beta^r \in (0,1)$ *is defined uniquely by the root of* I_2^r *on* $[0,1]$ *for* $p(2) \in (\underline{p}^r(2), \overline{p}^r(2))$.

As can be seen in Proposition 2 the lower cut-off below which experimentation ceases in the first period, is equal to the single agent cut-off. The upper equilibrium cut-off in the ring differs from the one in the complete network and it is easy to verify that $\overline{p}^r(1) < \overline{p}^c$ with the difference $\overline{p}^c - \overline{p}^r(1)$ monotonically increasing in n. This difference increases in the number of players because information needs longer to be transmitted and agents in the ring might experiment themselves instead of waiting for information of unobserved players. In addition, we have $\overline{p}^r(2) > \overline{p}^r(1)$ and the differences between $\overline{p}^r(1)$ and $\overline{p}^r(2)$ (as well as between p^a and $\underline{p}^r(2)$) stem from the two experiments of the neighbors of neighbors that agent i learns about in the second round. Comparing the second round cut-off in the ring, $\overline{p}^r(2)$, to the upper cut-off of the complete network, we can see that the difference between $\overline{p}^r(2)$ and $\overline{p}^c$ is positive for small n. However, as n increases this difference decreases and turns negative such that for a larger number of agents $\overline{p}^r(2) < \overline{p}^c$.

The longer agents have to wait for information, the more likely they will find it optimal to experiment themselves in the meantime.

We are interested in the difference between a complete network and the ring in terms of experimentation effort in equilibrium. Proposition 3 below shows that in the ring network the number of experiments is never smaller than in the complete network. In $t = 1$ it is easy to show that for high priors in both networks all agents experiment, for very pessimistic priors no one experiments and for intermediate values where agents use both options, the experimentation intensity is higher in the ring. This shows that agents compensate a worse possibility to learn from others through increased own effort. Figure 6 illustrates this finding.

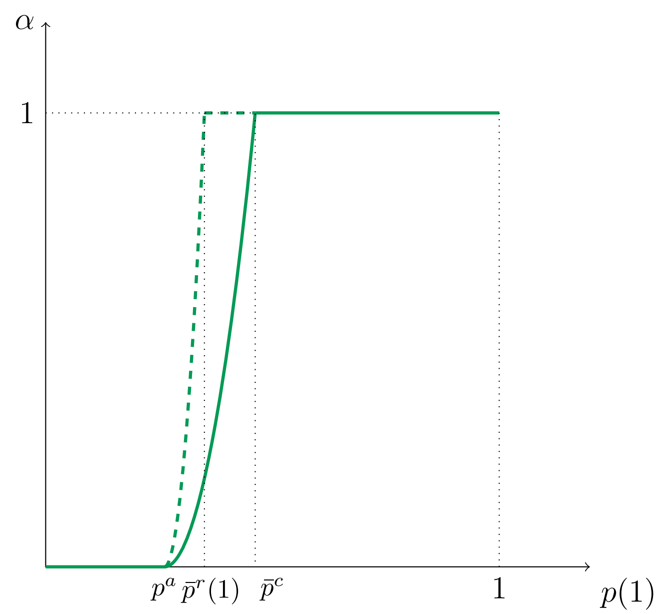


Figure 6: Equilibrium experimentation effort in a ring network (dashed line) and complete network (solid line) at $t = 1$ ($\pi = 0.2$, $\delta = 0.9$, $n = 12$, $E_1 = 1$, $E_0 = -1$, $p^a \approx 0.26$, $\bar{p}^c \approx 0.46$ and $\bar{p}^r(1) \approx 0.34$).

If all first round experiments fail, beliefs in the two networks in the second round are different as agents are already more pessimistic in the complete network. Taking the difference in posterior beliefs into account, it can be shown that the number of second round experiments in the ring and the complete network will either be the same, or that experimentation intensities will be higher in the ring. In order to be able to compare efforts in $t = 2$ across different networks, we express beliefs in terms of $p^c(2)$. This means that we make use of the fact that in equilibrium the relationship between posterior beliefs in the two networks after one round of failed experimentation is given by

$$p^c(2) = \frac{p^r(2)(1-\pi)^{n-3}}{p^r(2)(1-\pi)^{n-3} + 1 - p^r(2)}.$$

Proposition 3. *Experimentation intensities in the ring network are at least as high as in the complete network. More precisely,*

$$\alpha^r > \alpha^c \quad \text{for } p(1) \in (p^a, \overline{p}^c),$$
$$\alpha^r = \alpha^c \quad \text{for } p(1) \in [0, p^a] \cup [\overline{p}^c, 1]$$

and

$$\beta^r > \beta^c \quad \text{for } p^c(2) \in (\widetilde{\underline{p}}^r(2), \overline{p}^c),$$
$$\beta^r = \beta^c \quad \text{for } p^c(2) \in [0, \widetilde{\underline{p}}^r(2)] \cup [\overline{p}^c, 1]$$

where

$$\widetilde{\underline{p}}^r(2) = \frac{\underline{p}^r(2)(1-\pi)^{n-3}}{\underline{p}^r(2)(1-\pi)^{n-3} + 1 - \underline{p}^r(2)} < p^a.$$

The first thing we note about Proposition 3 is that agents in the ring network exert higher effort than agents in the complete network on certain intervals of beliefs. Further, agents in the ring network experiment in $t = 2$ at beliefs for which the posterior after n failed experiments is already below p^a. If information arrives with delay, agents might be better off experimenting themselves instead of waiting for information generated by others. However, as this information will eventually reach them, the final posterior belief in the ring network can be more pessimistic than in the complete network. That is, the probability of mistakenly abandoning a good risky project decreases and learning is more accurate. This is in line with the finding of Bimpikis and Drakopoulos (2014) that delaying information revelation might increase the amount of experimentation. The speed of learning, measured by the number of time periods until information has traveled to every node in the network, decreases due to the incomplete network structure. However, free-riding is reduced as players increase their effort over certain intervals of beliefs.

The analysis of the ring network shows that agents adjust their experimentation effort in equilibrium when information arrives with a delay. To obtain a better understanding of the role of different interaction structures, we now turn to the star network to explore the impact of asymmetric positions. We will see that the irregular structure of the star network has interesting consequences for experimentation efforts (as well as payoffs) in equilibrium.

5 The Star Network

In the star network one player, called the hub, is located in the center and has a link to each of the other $n-1$ players. The players at exterior positions, also called peripheral players, are only connected to the hub. Players are no longer symmetric and hence an equilibrium in which all players use the same strategy does not exist. In Proposition 4 below we construct an equilibrium where peripheral players use symmetric strategies and the hub exerts less effort than agents in a symmetric equilibrium in the complete network. More precisely, the hub exerts full effort until $\overline{p}^c$ and for beliefs below does not experiment at all. The peripheral players on the other hand use higher effort compared to the complete network. Proposition 4 describes the structure of the equilibrium where I_1^s and I_2^s are the respective counterparts to I_1^r and I_2^r for the peripheral players in the star network. The expressions for the various cut-offs beliefs as well as I_1^s and I_2^s can again be found in the Appendix.

Proposition 4. *An equilibrium of the strategic experimentation game in the star network where peripheral agents use symmetric strategies can be described as follows. The experimentation intensity at $t = 1, 2$ for the hub satisfies*

$$\phi^h(t) = \begin{cases} 1 & \text{for } p(t) \in [\overline{p}^c, 1], \\ 0 & \text{otherwise.} \end{cases}$$

For the peripheral players equilibrium actions in the first round are

- $\alpha^s = 1$ *for* $p(1) \in [\overline{p}^s(1), 1]$,
- $\alpha^s = 0$ *for* $p(1) \in [0, p^a]$,
- $\alpha^s \in (0, 1)$ *is defined uniquely for* $p(1) \in (p^a, \overline{p}^s(1))$ *by the root of* I_1^s *on* $[0, 1]$.

Second round experimentation intensities are

- $\beta^s = 1$ *for* $p(2) \in [\overline{p}^s(2), 1]$,
- $\beta^s = 0$ *for* $p(2) \in [0, \underline{p}^s(2)]$,
- $\beta^s \in (0, 1)$ *is defined uniquely for* $p(2) \in (\underline{p}^s(2), \overline{p}^s(2))$ *by the root of* I_2^s *on* $[0, 1]$.

Agents are now no longer in symmetric positions and the hub faces a different problem than the peripheral players. In particular, the central player is completely informed about all experiments like in a complete network. Hence, it is optimal for him to experiment with full intensity for any belief above $\bar{p}^c$. For beliefs between $(p^a, \bar{p}^c)$ the central player would follow the complete network equilibrium path if the peripheral players did so as well. If the peripheral players exert lower effort, that is if $\alpha^s < \alpha^c$, the hub finds it optimal to increase his effort compared to the complete network. On the other hand, if $\alpha^s > \alpha^c$, the best response for the hub is not to experiment at all. As it can be shown that $\bar{p}^s(1) < \bar{p}^c$, we know that in the interval $[\bar{p}^s(1), \bar{p}^c)$ the hub does not experiment. We can then show that, if the hub does not experiment at all for beliefs below $\bar{p}^c$, the best response for the peripheral players in $(p^a, \bar{p}^c)$ is to exert higher effort than in the complete network, that is, $\alpha^s > \alpha^c$. For priors above or below the interval $(p^a, \bar{p}^c)$ there will be full or no experimentation, respectively. This is illustrated in Figure 7.

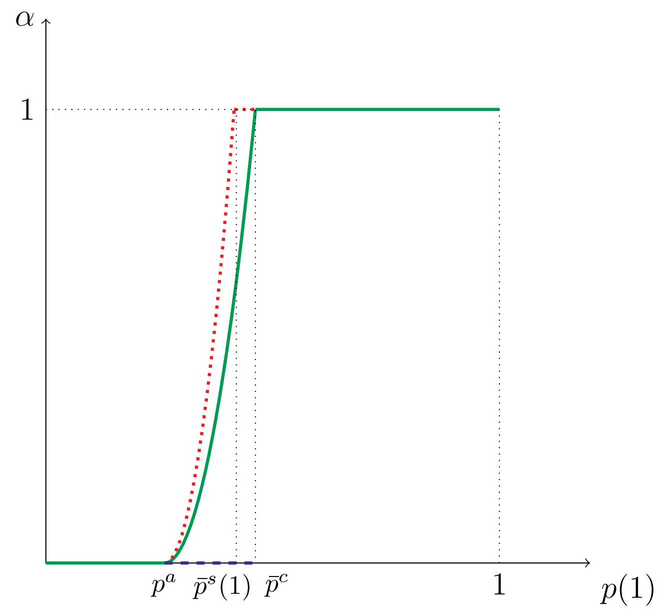


Figure 7: Equilibrium experimentation effort of the peripheral players in the star network (bold dotted line), the central player in the star network (dashed line) and in a complete network (solid line) at $t = 1$ ($\pi = 0.2$, $\delta = 0.9$, $n = 12$, $E_1 = 1$, $E_0 = -1$, $p^a \approx 0.26$, $\bar{p}^c \approx 0.46$ and $\bar{p}^s(1) \approx 0.42$).

Remark 1. *For some values of the parameters there exists a second equilibrium in which the peripheral players use symmetric strategies. In this equilibrium the hub exerts full effort for beliefs* $p(1) \in [p^a, \overline{p}^s(1)) \cup [\overline{p}^c, 1]$ *and no effort for* $p(1) \in [0, p^a) \cup [\overline{p}^s(1), \overline{p}^c)$*. This means that the effort of the hub is non-monotonic in the belief. The peripheral agents exert full effort for beliefs above* $\overline{p}^s(1)$*. For any belief in* $[p^a, \overline{p}^s(1))$ *their experimentation intensity is lower than the experimentation intensity that would make the hub indifferent,* *that is, $\alpha^s < \alpha^c$. The second equilibrium only exists if n is small and δ and π are large. For this reason we will subsequently restrict attention to the equilibrium of Proposition 4.*

Whether there will be more experiments in the star or the complete network depends on the possibility of the peripheral agents to counterbalance the decreased experimentation intensity of the hub. In the equilibrium described in Proposition 4 the experimentation effort of the hub is below or equal to the effort level of the peripheral players, that is

$$\alpha^h + \beta^h \leq \alpha^s + \beta^s.$$

As will be shown in Proposition 5, except for a combination of parameter values where n is small and δ and π are rather large, overall experimentation intensities in the star network are higher or equal to experimentation effort in the complete network.

Proposition 5. *Comparing effort in the complete network and the star network we obtain*

$$(n-1)\alpha^s + \alpha^h = n\alpha^c \text{ for all } p(1) \in [0, p^a] \cup [\overline{p}^c, 1]$$

and

$$(n-1)\beta^s + \beta^h = n\beta^c \text{ for all } p^c(2) \in [0, \widetilde{\underline{p}}^s(2)] \cup [\overline{p}^c, 1]$$

where

$$\widetilde{\underline{p}}^s(2) = \frac{\underline{p}^s(2)(1-\pi)^{n-2}}{\underline{p}^s(2)(1-\pi)^{n-2} + 1 - \underline{p}^s(2)} < p^a.$$

For $p(1) \in (p^a, \overline{p}^c)$ there exists a strict subset $S_n(p(1))$ of $[0,1]^2$ such that

$$(n-1)\alpha^s + \alpha^h > n\alpha^c \text{ if and only if } (\delta, \pi) \in S_n(p(1)).$$

Moreover, $\lambda(S_n(p(1))) \to 1$ as $n \to \infty$ with λ denoting the Lebesgue measure on $\mathbb{R}^2$. Similarly, for $p^c(2) \in (\widetilde{\underline{p}}^s(2), \overline{p}^c)$ there exists a strict subset $S_n(p(2))$ of $[0,1]^2$ such that

$$(n-1)\beta^s + \beta^h > n\beta^c \text{ if and only if } (\delta, \pi) \in S_n(p(2))$$

and $\lambda(S_n(p(2))) \to 1$ as $n \to \infty$.

The first part of Proposition 5 states the intervals of beliefs in which experimentation effort in equilibrium in the complete network is equal to the star network, because there is either no experimentation or all agents exert full effort. For beliefs outside these intervals (that is, $p(1) \in (p^a, \overline{p}^c)$) and $p^c(2) \in (\widetilde{\underline{p}}^s(2), \overline{p}^c)$ we know that $\phi^h(t) = 0$. The region $S_n(p(1))$ is then defined as all combinations of δ and π for which total effort in $t = 1$ in the complete network is strictly smaller than in the star network. By analyzing this expression (see Appendix) numerically, one can see that the value for δ below which

$(n-1)\alpha^s \geq n\alpha^c$ is in general "quite close" to 1. For example, for $n = 3$, $(n-1)\alpha^s \geq n\alpha^c$ as long as $\delta \leq \frac{8}{9}$ even if π takes values arbitrarily close to 1. As n increases, the threshold value for δ increases and already for relatively small n ($n = 6$) $\delta \leq 0.99$ suffices to guarantee that $(n-1)\alpha^s \geq n\alpha^c$ again assuming values of π close to 1. The lower π, the higher is δ below which $(n-1)\alpha^s \geq n\alpha^c$.

The total experimentation intensity in the interval where agents use both arms is higher in the star network except for a combination of parameter values with high δ, high π and small n. That is, unless agents are very patient, effort in the star is higher even though the hub does not experiment. This indicates that the peripheral agents increase own efforts accordingly to outweigh the missing experimentation of the hub as well as the payoff disadvantage that arises from delayed information transmission. The higher δ, the closer are α^s and α^c and consequently the more likely is $(n-1)\alpha^s < n\alpha^c$. If the number of players is small, more weight is placed on the hub and it is more difficult for the peripheral players to compensate for the missing experiment of the central player.

We now turn to a comparison of experimentation intensities in the ring and the star network. As the number of agents increases, more information arrives with a greater number of time lags in the ring. In the star network on the other hand, the delay does not change if the number of players changes. Restricting attention to intervals of beliefs in which neither $\phi^r(t) = \phi^s(t) = 0$ nor $\phi^r(t) = \phi^s(t) = 1$, we show in Proposition 6 below that for small n, experimentation intensities in the star network are no smaller than those in the ring while for a large number of players it depends on δ and π.

Proposition 6. *Comparing $\phi^s(t)$ to $\phi^r(t)$ for all $p(t)$ from the interval in which at least in one of the two networks agents are indifferent between experimenting and using the safe option (that is, $I_t^r = 0$, or $I_t^s = 0$, or both), we have for $t = 1, 2$ that*

(i) there exists $n_t \in \mathbb{N}$ such that for all $n < n_t$, $\phi^s(t) \geq \phi^r(t)$ for all $(\delta, \pi) \in [0,1]^2$ and

(ii) as $n \to \infty$ the region of (δ, π) in which $\phi^s(t) \geq \phi^r(t)$ is a strict subset of $[0,1]^2$.

The first point of Proposition 6 tells us that for a small number of players efforts in the star are greater (in the interval of beliefs where agents use both arms) or equal to effort in the ring. For a larger number of agents, this is no longer true in general. Part (*ii*) of the proposition says that as n becomes large, the region for which effort in the star network is higher becomes a strict subset of $[0,1]^2$. As a consequence of part (*ii*) and the fact that $\phi^r(t)$ and $\phi^s(t)$ intersect only at one belief (e.g., in $t = 1$ at p^a), we can also conclude that there exists some finite natural number such that for all n above this number there exists a non-empty set of parameters (δ, π) for which $\phi^s(t) < \phi^r(t)$.

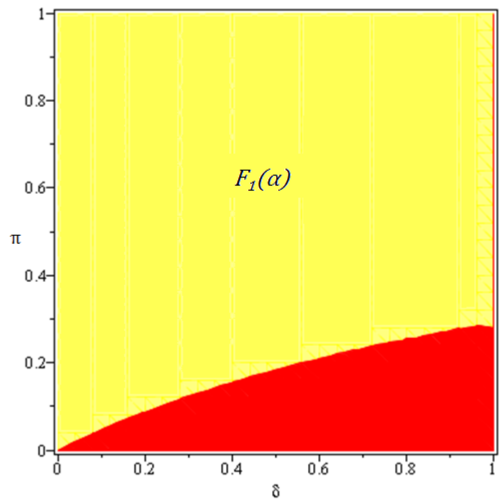


Figure 8: Equilibrium experimentation effort in $t = 1$ in the ring network compared to the star network for $n \to \infty$. In the light region (denoted by $F_1(\alpha)$) $\alpha^s \geq \alpha^r$ and in the dark region $\alpha^s < \alpha^r$.

From Proposition 6 we can infer that the effect of incomplete network structures on experimentation intensities heavily depends on the discount factor δ and the success rate π. Suppose in both networks all agents experiment with full intensity at time $t = 1$. Then, at the end of this round, peripheral agents in the star network learn about one experiment fewer than agents in the ring network. For δ close to 1 this does not matter to these agents as they are almost perfectly patient and it makes little difference to them at which point in time information arrives. On the other hand, the closer δ is to zero, the more agents in the star network care about this missing experiment, making them increase own effort. Figure 8 shows the region in which $\alpha^s > \alpha^r$ for $n \to \infty$. In the light region we have $\alpha^s \geq \alpha^r$ and vice versa in the dark region. For example, we can see that $\alpha^r \geq \alpha^s$ only if π is not too large. If the probability of a breakthrough is low, it is relatively more likely that agents will learn about a breakthrough later in the ring than in the star network. Thus, agents in the ring increase their experimentation effort to balance this effect. Note that Proposition 6 does not tell us in which network total effort is higher, which matters in the interval where the hub does not experiment. Similarly to the complete network, in case we have $\phi^s(t) \geq \phi^r(t)$, there will be a combination of parameter values for which total experimentation effort is higher in the ring network, because peripheral players cannot compensate for the missing experiment of the hub.

Before turning to the question which network generates the highest welfare among the three structures considered, let us briefly repeat the main findings of the previous sections. First, we showed that agents increase own effort if information arrives with delay as it is better for them to experiment themselves instead of waiting for information generated by others. Second, in irregular structures there can be specialization where some agents experiment while others free-ride. Experimenting agents increase their effort to outweigh the missing experiments as well as the delay in the arrival of information.

6 Welfare Analysis

In the preceding sections it was shown that effort exerted in equilibrium varies with the interaction structure. In this section we want to analyze the implications of these differences for expected payoffs in equilibrium. Assuming that it is costly to establish a communication or interaction structure, we are now interested in which of the three networks would be chosen (before the agents engage in the experimentation game) by a social planner that aims to maximize welfare given the strategic behavior of the players.

There are fixed costs $k \geq 0$ per link that have to be paid ex ante. The total number of links in network g depends on the network structure and is $n(n-1)/2$ in the complete network, n in the ring and $n-1$ in the star network. The main criterion to measure the performance of different structures are equilibrium payoffs. The explicit expressions for expected equilibrium payoffs in the different networks as well as all technical details of the subsequent analysis can be found in Appendix B. Welfare is defined as the total expected payoff in equilibrium minus total costs for building the infrastructure. For the complete network this is

$$W^c(p(1)) = nU^c(p(1)) - \frac{n(n-1)}{2}k.$$

For the other networks it is defined in an analogous way, that is $W^r(p(1)) = nU^r(p(1)) - nk$ and $W^s(p(1)) = (n-1)U^s(p(1)) + U^h(p(1)) - (n-1)k$. A network $g \in \{c, r, s\}$ is optimal for a given prior belief $p(1)$ and set of parameters (δ, π, k, n) if and only if

$$W^g(p(1)) \geq W^{g'}(p(1)), \text{ for all } g' \in \{c, r, s\}.$$

We write $g \succ g'$ if network g generates strictly higher welfare than network g' and $g \sim g'$ if $W^g(p(1)) = W^{g'}(p(1))$.

Proposition 7 below states which network is optimal when $k = 0$ and the prior belief $p(1)$ is such that in case all experiments in $t = 1$ fail, there are no experiments in $t = 2$, that is, $p(2) \leq \underline{p}^g(2)$ for all g.[6] For simplicity of exposition, in the subsequent analysis we impose $E_1 = 1$ and $E_0 = -1$. Note that we do not include the empty network in our analysis, which would of course be optimal for very pessimistic priors or high costs. Without the empty network, clearly, the star network is optimal for sufficiently high costs. What is more interesting, however, is that the star network is strictly preferred to the complete network over a certain interval of priors even if links do not incur any costs.

Proposition 7. *The following conditions determine which network is optimal for $k = 0$ and for $p(1)$ such that in case all experiments in $t = 1$ fail, $p(2) \leq \underline{p}^g(2)$ for all g :*
(i) For $p(1) \in [0, p^a]$: $c \sim r \sim s$;
(ii) for $p(1) \in (p^a, \overline{p}^s(1)]$: $s \succ c, r$ and the relation between c and r is given in (iii);
(iii) $c \sim r$ for $p(1) \in [0, \overline{p}^r(1)]$ and $c \succ r$ for $p(1) \in (\overline{p}^r(1), 1]$;
(iv) for $p(1) \in (\overline{p}^s(1), \overline{p}^c]$: $c \succ s$ if and only if

$$(1-\delta)(2p(1)-1)+\delta p(1)[(1-\pi)^{n-1}[1+\delta(n-1)]+(1-\delta)(n-1)(1-\pi)-n(1-\alpha^c\pi)^n] > 0.$$

(v) For $p(1) \in (\overline{p}^c(1), 1]$: $c \succ r, s$.

For $p(1) \in (p^a, \overline{p}^s(1)]$ the complete network is never optimal even if costs for links are zero. This result is somewhat surprising as one might think that it is optimal to have as many links as possible if they are costless to allow a fast flow of information. However, in this interval of beliefs the star network is strictly optimal for two reasons. First, average expected payoffs in the star (where the hub does not experiment) are higher than in the complete network or the ring, because the hub does not bear the costs of experimentation but receives the informational benefits. Second, up to $\overline{p}^s(1)$ the peripheral players can increase their experimentation effort so as to fully compensate for both the lack of experimentation of the hub as well as the delay in the information transmission. Up to this threshold, therefore, welfare in the star network is strictly higher than in the ring or complete network. At some belief above this threshold this result is reversed and the missing experiment of the central player implies that average expected payoffs are lower in the star network than in the other networks. Corollary 1 summarizes this result.

Corollary 1. *Specialization in the star network, where $\alpha^h = 0$ and $\alpha^s > 0$, can be beneficial as well as detrimental to overall welfare.*

Another interesting observation can be made by comparing the complete network to the ring. As pointed out in Section 4, for beliefs in the interval $(p^a, \overline{p}^r(1)]$ agents

[6] Appendix B contains a generalization to two rounds of experimenting as well as complete description of the case $n = 4$ for any cost level $k \geq 0$.

exert higher effort in the ring network than in the complete network. More precisely, agents increase their effort to exactly offset the payoff disadvantage resulting from a slower speed of learning (measured as the number of time periods until all information has reached every node in the network). This means that expected payoffs in the ring and the complete network are identical for beliefs in which the players in the ring use interior experimentation intensities. If all agents in both networks experiment with full intensity agents learn faster in the complete network and are better off. This implies that, as stated in Corollary 2 below, there exists a trade-off between network structures that enable a high speed of learning and structures in which the final posterior in case all experiments were unsuccessful is more pessimistic.

Corollary 2. *In the selection of the optimal network structure there exists a trade-off between structures associated with high speed of learning, and structures that lead to higher accuracy of learning.*

This trade-off is also apparent when looking at a situation where some agents experiment in $t = 2$ after a round of failed experimentation in $t = 1$. It is possible that in equilibrium in $t = 2$ only the peripheral players in the star experiment. This situation can be used to illustrate how the discount factor δ influences which network is optimal. One main advantage of the complete network compared to incomplete structures lies in the speed of learning, making it increasingly attractive the stronger future payoffs are discounted. In the interval of beliefs in which only the peripheral players experiment in both rounds, whereas all agents in other networks experiment only in $t = 1$, it can be shown that for high values of the discount factor δ close to 1, the star network is always preferred (see Appendix B). On the other hand, for δ close to 0, the complete network is preferred for $k = 0$. This comparison stresses again the existing trade-off between faster learning and more complete learning. How this trade-off is resolved depends on the discount factor.

In the course of this section we observed that whether a certain network is optimal, depends on the agents' possibility to increase their experimentation effort in order to compensate for the disadvantage of delayed information exchange in incomplete structures. For costs of links equal to zero, c can only be optimal for prior beliefs $p(1)$ such that $\alpha^s = \alpha^r = 1$, as otherwise agents can increase their experimentation effort in order to outweigh the delayed arrival of information. At some belief in the interval $(\overline{p}^s(1), \overline{p}^c(1)]$ the peripheral players in the star can no longer compensate for the nonexperimenting hub and total experimentation effort is lower than optimal. Generally, as long as players can increase their efforts to ensure that expected equilibrium payoffs are the same as in other network structures, only the number of links determines which network is optimal.

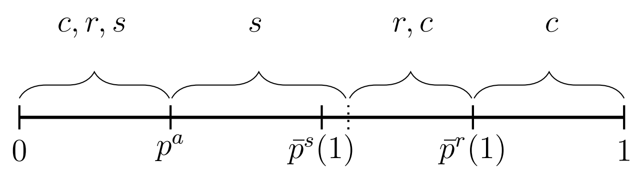


Figure 9: Optimal networks for $k = 0$, $n = 4$ and $p(1)$ such that $p(2) \leq \underline{p}^g(2)$ for all g if all experiments in $t = 1$ fail, for different intervals of the prior. The dotted line between $\bar{p}^s(1)$ and $\bar{p}^c(1)$ indicates the belief at which peripheral players in the star network can no longer compensate for the missing experiment of the hub and the delay in information transmission by increasing own effort.

In the star network an additional effect comes into play, namely the payoff advantage of the non-experimenting hub, which explains why even for zero costs the star is strictly preferred for low priors. Figure 9 graphically illustrates for $n = 4$ which of the three networks is optimal on different intervals of priors.

Our analysis confirms two results of BK. First, it shows that under certain circumstances specialization (that is, some agents exert effort while others free-ride) might benefit society, and second, welfare can be higher in incomplete interaction structures. However, we can also show the opposite effect, namely that for certain beliefs specialization can have a negative impact on overall welfare.

As mentioned in the introduction, network structures can also be interpreted as organizational structures that determine the flow of information within an organization. When deciding on the optimal organizational structure (for example, centralized vs. decentralized structures), decision-makers might pursue various objectives. For instance, if the objective is to minimize the costs of information transmission, a centralized structure such as the star network is optimal. Centralization enables a comparatively fast flow of information at lowest possible costs. From the perspective of the management of a firm centralization additionally offers the advantage that a central authority can accumulate and disseminate information.

Interestingly, we show that a fast flow of information does not necessarily maximize welfare even if information can be distributed to all players immediately at no cost. However, what needs be stressed in this context is that the star network is strictly optimal for low prior beliefs only because of specialization. This implies that asymmetric equilibria in the complete (or ring) network will most likely generate higher welfare than the symmetric one (see KRC or Bramoullè, Kranton, D'Amours, 2014). Thus, it is not clear whether the star is still strictly preferred once agents in the complete network are allowed to use asymmetric strategies. However, if we compare the symmetric equilibrium in the ring network to the complete network we see that for beliefs where agents in the ring network balance the delay in information transmission by increasing effort, expected equilibrium payoffs are identical. This contradicts the findings of Teece (1994) that innovation has to be associated with a fast transmission of information.

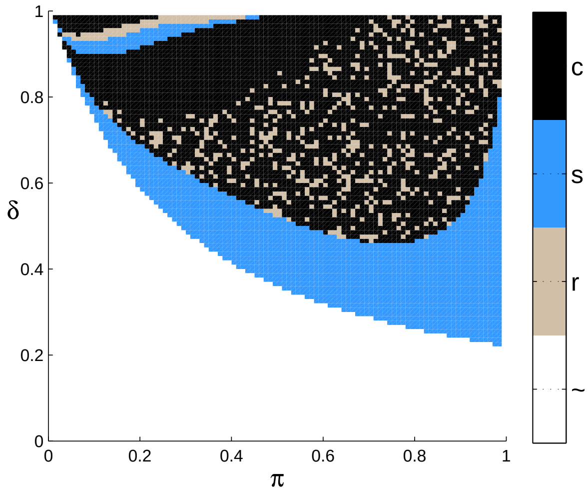
6.1 Numerical example

In this section we present numerical results that complement the preceding analytical discussion. While up to this point we focused on the role of the prior, we now want to obtain a better understanding of the role of different parameters. In our numerical example we show which network is optimal in a (π, δ)-grid given fixed values of the other parameters.

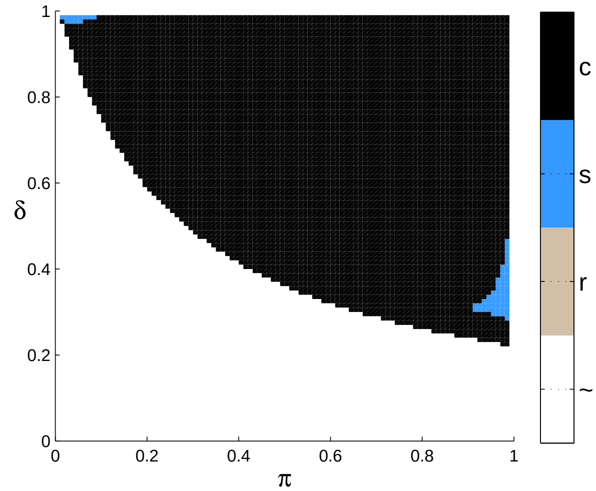
Figure 10 illustrates the results. It shows which network is optimal for $E_0 = -1$ and $E_1 = 1$. The results are calculated for $\pi \in [0.01, 0.99]$ and $\delta \in [0.01, 0.99]$ both in steps of 0.01. In the white region no network is strictly optimal as in this region there is no experimentation (that is, we have indifference). Light grey areas indicate all combinations of δ and π in which the ring network is optimal, dark grey represents optimality of the star network, and black means that the complete network is optimal. The three panels on the left display the results for $n = 5$, while those on the right have $n = 25$. In the first row $p(1) = 0.45$ and $k = 0$, in the second row the prior belief is increased to $p(1) = 0.96$ while $k = 0$, and in the last row we look at $p(1) = 0.96$ for costs $k = 0.001$.

In Figure 10a we see that for low values of δ and π no network is strictly optimal, as no agent experiments. For medium values of δ, e.g., $\delta = 0.4$, we have indifference for low values of π and the star network dominates for high π. As δ and π increase, the star network is less often optimal and expected welfare is highest in the complete network. More precisely, in 10a the complete network is optimal in 37.4% of the cases, the star in 17.6%, the ring network in 7.4% and in 37.6% of the cases we have indifference. If we increase n to 25 (see Figure 10b) the ring network is never optimal and the complete network is optimal for values of the parameters where for $n = 5$ the star is optimal. The percentages change to 61.2% for the complete network, 1.3% for the star and 37.5% for indifference.

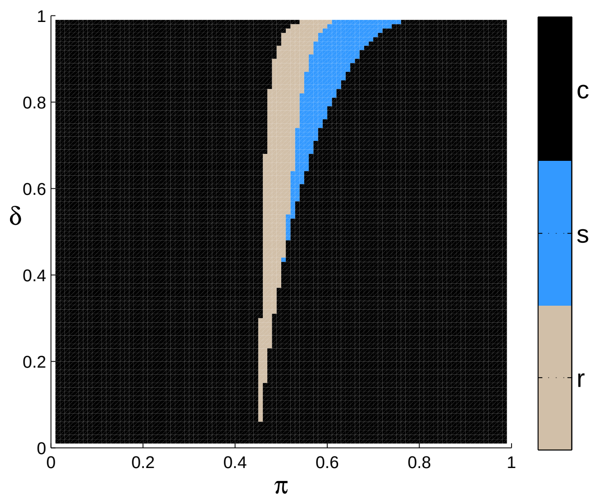
In Figures 10c and 10d agents are very optimistic and experiment for sure. That is, there is no region of indifference. As expected, the complete network is optimal in this



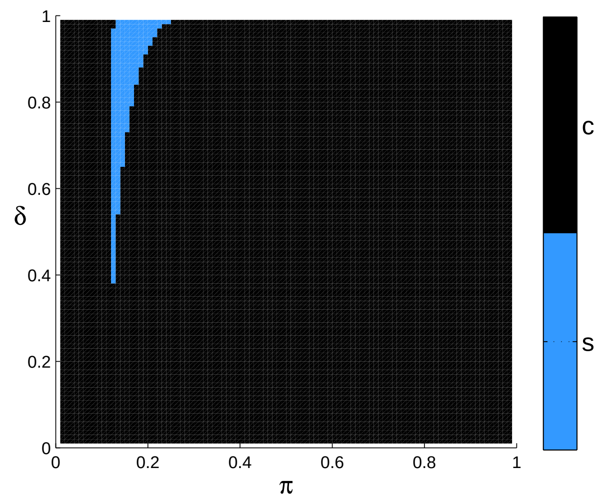
(a) $p(1) = 0.45$, $k = 0$, $n = 5$.



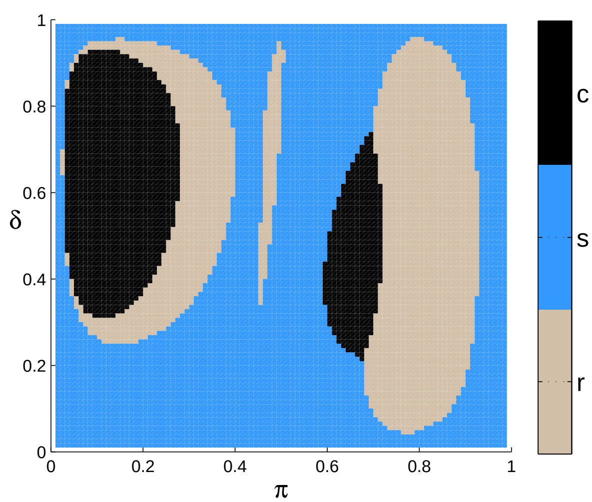
(b) $p(1) = 0.45$, $k = 0$, $n = 25$.



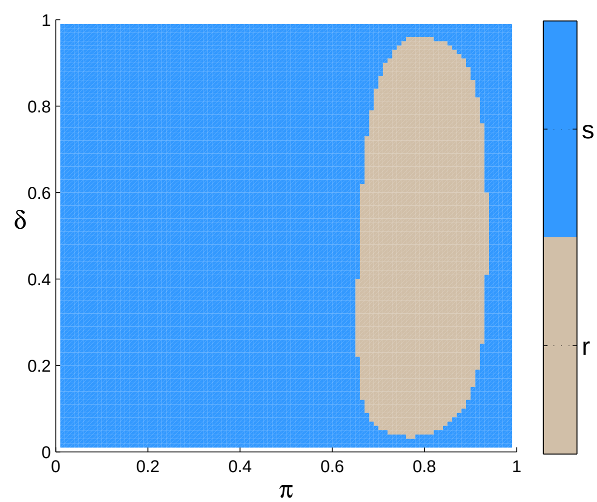
(c) $p(1) = 0.96$, $k = 0$, $n = 5$.



(d) $p(1) = 0.96$, $k = 0$, $n = 25$.



(e) $p(1) = 0.96$, $k = 0.001$, $n = 5$.



(f) $p(1) = 0.96$, $k = 0.001$, $n = 25$.

Figure 10: Optimal networks for $E_0 = -1$ and $E_1 = 1$.

case for a large combination of parameters (92.4% in 10c and 97.6% in 10d). However, for intermediate values of π there exists an area in which the star network or the ring generate higher welfare. Increasing the number of players to $n = 25$ shifts the region in which the star network dominates to the left, that is, to lower values of π. Moreover, the ring network is never optimal.

In the last row in Figures 10e and 10f we introduce positive costs for links. Naturally, the region in which the complete network is optimal shrinks for $n = 5$ and completely disappears for $n = 25$. In fact, for $n = 25$ and $k \geq 0.001$ the complete network is suboptimal for all δ, π, and $p(1)$. Moreover, as soon as $k > 0$, there is no combination of parameters in which agents are indifferent between different network structures.

7 Discussion

In the preceding sections we analyzed a game of strategic experimentation in three different network structures. First, as a benchmark the complete network was considered. Second, the ring network was analyzed and we showed that relative to the complete network agents increase their effort when information arrives with a delay. Consequently, for a given prior belief the total amount of experimentation in a ring is higher or equal to the amount of experimentation in a complete network. Agents increase their experimentation effort to exactly balance the payoff disadvantage resulting from the delay in information transmission. Third, by analyzing the strategic experimentation game in the star network we showed that the hub experiments with full intensity up to a threshold belief and then stops completely. Although the peripheral players increase their effort relative to the complete network in the interval where the hub stops "too early", they are not always able to fully compensate for the non-experimenting hub. Depending on the belief this specialization in the star network can be beneficial as well as detrimental for society.

Generally, there exists a trade-off between faster learning and more accurate learning. Different network structures have different effects on the outcome of the experimentation game and consequently on welfare. While the star network minimizes the costs for links, the complete network maximizes the speed of learning. In which of the three networks learning will be most accurate depends on the prior belief as well as the parameters of the model. Even though our model differs in several features we confirm the finding of BK that equilibria in the star are specialized. Compared to the static framework of BK the dynamic perspective allows us to show how cut-off beliefs depend on the network structure. Further, agents hold different posterior beliefs depending on their position. We find like Bimpikis and Drakopoulos (2014) that if information arrives with delay, effort might increase and free-riding decrease.

Our analysis showed that it is possible to investigate details of rational learning processes in a network without being restricted to focus on asymptotic results or introduce some form of myopia or bounded rationality. Nevertheless, the model considered here captures very particular learning situations due to its special structure with fully revealing breakthroughs. This implies that our results cannot be easily generalized to other payoff generating processes. Another shortcoming of the analysis is the restriction to symmetric equilibria which may not be without loss of generality.[7]

Despite the complexity network structures can create, we showed that they affect the behavior of agents in an intuitive way. This offers some suggestions as to how equilibrium outcomes and strategies could be characterized in other settings (e.g., for other payoff generating processes) as well. Further, the network structures considered in this paper can be understood as specific monitoring structures, and it would be possible to analyze the strategic experimentation game for monitoring structures which are not derived from networks. What should be clear, however, is that the empty and the complete network are two opposite ends of the spectrum. Consequently, for symmetric monitoring structures we expect the main conclusions of the network case to remain valid. Of course, it would be desirable to obtain a generalization of the results for irregular structures as well, which seems to be considerably more involved and will most likely imply specialization as in the star network.

[7]See KRC or Bramoullè, Kranton and D'Amours (2014).

A Appendix

Proof of Proposition 1. The objective function is linear in the choice variables and either $(\alpha_i^c, 0)$ or $(1, \beta_i^c)$ is optimal where $\alpha_i^c, \beta_i^c \in [0,1]$. That is, an optimal strategy can never require that agents choose interior experimentation intensities in both rounds. Comparing expected payoffs for α_i^c at $(0,0)$ and $(1,0)$ we find that agent i is indifferent between experimenting in the first round and not experimenting as long as

$$I_1^c = (1-\delta)E_{p(1)} + \delta E_1 p(1)\pi(1-\alpha_j\pi)$$

equals zero. If $I_1^c > 0$ it is optimal to choose $\alpha^c = 1$ and if $I_1^c < 0$ the safe arm is optimal. From I_1^c we can derive the optimal experimentation effort in a symmetric equilibrium with no experimentation at $t = 2$ as

$$\alpha^c = \frac{1}{\pi} - \frac{1}{\pi}\left(\frac{(1-\delta)|E_0|}{\delta E_1 \pi p(t)} - \frac{(1-\delta)(|E_0| + E_1)}{\delta E_1 \pi}\right)^{\frac{1}{n-1}}$$

which has $\alpha^c = 0$ as optimal solution for beliefs $p(1)$ below p^a and $\alpha^c = 1$ for any beliefs $p(1)$ above $\overline{p}^c$. The optimal experimentation intensity β^c can be found along the same lines by comparing utilities from $(1,0)$ and $(1,1)$. The agent is indifferent if

$$I_2^c = (1-\delta)E_{p(2)} + \delta E_1 p(2)\pi(1-\beta_j\pi)$$

equals zero. This expression is the same as I_1^c with β instead of α and $p(2)$ instead of $p(1)$. It can be shown that $\beta^c > 0$ implies $\alpha^c = 1$. Thus, in a symmetric equilibrium,

$$\phi_i^c(t) = \begin{cases} 1 & \text{for } p(t) \in [\overline{p}^c, 1), \\ \frac{1}{\pi} - \frac{1}{\pi}\left(\frac{(1-\delta)|E_0|}{\delta E_1 \pi p(t)} - \frac{(1-\delta)(|E_0|+E_1)}{\delta E_1 \pi}\right)^{\frac{1}{n-1}} & \text{for } p(t) \in (p^a, \overline{p}^c), \\ 0 & \text{for } p(t) \in [0, p^a], \end{cases}$$

for all $i \in N$. The strategy profile $\alpha_i^c = \beta_i^c = 1$ is optimal for any prior belief $p(1)$ above

$$\frac{(1-\delta)|E_0|}{(1-\delta)|E_0| + E_1(1-\pi)^n[1-\delta+\delta\pi(1-\pi)^{n-1}]},$$

which is updated to a posterior $p(2) = \overline{p}^c$ after n failed experiments. ■

Proof of Proposition 2. The cut-off beliefs and corresponding intensities can be found by solving for α^r at which

$$I_1^r = (1-\delta)E_{p(1)} + \delta E_1 p(1)\pi\left((1-\alpha^r\pi)^2 - [1-(1-\alpha^r\pi)^2]\sum_{t=1}^{\frac{n-1}{2}-1} \delta^t (1-\alpha^r\pi)^{2t}\right)$$

equals zero. I_1^r is the difference in expected utility between action profiles $(0,0)$ and $(1,0)$. If $I_1^r = 0$ agents are indifferent between experimenting and not experimenting. The lower cut-off belief below which $\alpha^r = 0$ is given by $\underline{p}^r(1) = p^a$, while the belief above which $\alpha^r = 1$ is

$$\overline{p}^r(1) = \frac{(1-\delta)|E_0|}{(1-\delta)[|E_0| + E_1] + \delta E_1 \pi \left[(1-\pi)^2 - \pi(2-\pi)\sum_{t=1}^{\frac{n-1}{2}-1} \delta^t (1-\pi)^{2t}\right]}.$$

In between these two cut-offs agents use interior intensities that are increasing in the belief. In contrast to the complete network an explicit simple expression for α^r cannot be derived from the above equation. A discussion of this expression together with a proof that the root on $[0,1]$ exists and is unique can be found below. If $\alpha_i^r = 1$ for all i, beliefs in $t=2$ are given by

$$p^r(2) = \frac{p(1)(1-\pi)^3}{p(1)(1-\pi)^3 + 1 - p(1)}.$$

Taking into account that information is still traveling along the paths in the network, similar to one round we define

$$I_2^r = (1-\delta)E_{p^r(2)} + \delta E_1 p^r(2)\pi(1-\pi)^2(1-\beta^r\pi)^2 - \delta E_1 p^r(2)\pi B,$$

where

$$\begin{aligned} B &= \delta^{\frac{n-3}{2}}(1-\pi)^n(1-\beta^r\pi)^{n-3}[1-(1-\beta^r\pi)^2] \\ &\quad + [1-(1-\pi)^2(1-\beta^r\pi)^2]\sum_{t=1}^{\frac{n-5}{2}} \delta^t(1-\pi)^{2t}(1-\beta^r\pi)^{2t}. \end{aligned}$$

Now $I_2^r = 0$ implies indifference between $(1,0)$ and $(1,1)$ given the other players' strategies. From this we can derive β^r which attains its lower bound 0 at

$$\underline{p}^r(2) = \frac{(1-\delta)|E_0|}{(1-\delta)[|E_0| + E_1] + \delta E_1 \pi \left[(1-\pi)^2 - \pi(2-\pi)\sum_{t=1}^{\frac{n-5}{2}} \delta^t (1-\pi)^{2t}\right]},$$

and its upper bound 1 at

$$\overline{p}^r(2) = \frac{(1-\delta)|E_0|}{(1-\delta)[|E_0| + E_1] + \delta E_1 \pi D},$$

where $D = (1-\pi)^4 - \delta^{\frac{n-3}{2}}(1-\pi)^{2n-3}\pi(2-\pi) - [1-(1-\pi)^4]\sum_{t=1}^{\frac{n-5}{2}} \delta^t(1-\pi)^{4t}$.

Existence and uniqueness of α^r, β^r : The expression for α can be found by analyzing $I_1^r = 0$ and for β we have $I_2^r = 0$. $I_1^r = 0$ can be rewritten as

$$\frac{(1-\delta)E_{p(1)}}{\delta E_1 p(1)\pi} + (1-\alpha\pi)^2 - [1-(1-\alpha\pi)^2]\sum_{t=1}^{\frac{n-1}{2}-1} \delta^t (1-\alpha\pi)^{2t} = 0, \qquad \text{(A.7)}$$

where the expression on the l.h.s is a polynomial of order $n-1$ in α. To show that the root on $[0,1]$ is unique, it is enough to show that $(A.7)$ is strictly monotonically decreasing for $\alpha \in [0,1]$. We rewrite $I_1^r = 0$ as

$$
\begin{aligned}
0 &= (1-\alpha\pi)^2 - \delta(1-\alpha\pi)^2 - \delta^2(1-\alpha\pi)^4 - .... - \delta^{\frac{n-1}{2}-1}(1-\alpha\pi)^{n-3} + \\
&\quad \delta(1-\alpha\pi)^4 + \delta^2(1-\alpha\pi)^6 + ... + \delta^{\frac{n-1}{2}-1}(1-\alpha\pi)^{n-1} + \frac{(1-\delta)E_{p(1)}}{\delta E_1 p(1)\pi} \\
&= (1-\delta)\left[(1-\alpha\pi)^2 + \delta(1-\alpha\pi)^4 + \delta^2(1-\alpha\pi)^6 + ... + \delta^{\frac{n-1}{2}-2}(1-\alpha\pi)^{n-3}\right] - \\
&\quad \delta^{\frac{n-1}{2}-1}(1-\alpha\pi)^{n-1} + \frac{(1-\delta)E_{p(1)}}{\delta E_1 p(1)\pi}.
\end{aligned}
$$

Taking the derivative w.r.t. α we obtain

$$
(1-\delta)[-2\pi(1-\alpha\pi) - 4\delta\pi(1-\alpha\pi)^3 - ...] - (n-1)\delta^{\frac{n-1}{2}-1}\pi(1-\alpha\pi)^{n-2},
$$

which is clearly negative for $\alpha \in [0,1]$. A similar analysis can be carried out for β. ∎

Proof of Proposition 3. The proof proceeds in two steps and separates the problem in $t = 1$ from the one in $t = 2$. If $p(1) \geq \overline{p}^c$, all agents experiment with intensity 1 and if $p(1) \leq p^a$, no agent experiments. For $p(1) \in (p^a, \overline{p}^c)$ optimal experimentation intensities are higher in the ring. For beliefs in $[\overline{p}^r(1), \overline{p}^c)$ agents in the ring play with full intensity while players in the complete network have effort levels below 1. The difference $\overline{p}^c - \overline{p}^r(1)$ is given by

$$
(1-\pi)^2\left([1-(1-\pi)^{n-3}] - \delta\pi(2-\pi)\frac{1-\delta^{\frac{n-3}{2}}(1-\pi)^{n-3}}{1-\delta(1-\pi)^2}\right).
$$

For beliefs in $(p^a, \overline{p}^r(1))$ we know that in equilibrium $I_1^r = 0$ and $I_1^c = 0$. As prior beliefs are assumed to be identical it follows that

$$
(1-\alpha^c\pi)^{n-1} - (1-\alpha^r\pi)^2\left(\frac{1-\delta+\delta^{\frac{n-1}{2}}(1-\alpha^r\pi)^{n-3}[1-(1-\alpha^r\pi)^2]}{1-\delta(1-\alpha^r\pi)^2}\right) = 0.
$$

The equality holds for $\alpha^r = \alpha^c = 0$ and in case we set $\alpha^r = \alpha^c = \alpha$ this term is monotonically decreasing in α and negative for any $\alpha > 0$. Consequently, for the equality to hold we need

$$
\alpha^r > \alpha^c.
$$

Further, the term is decreasing in n and consequently the difference $\alpha^r - \alpha^c$ increases in n.

In $t = 2$ it has to be shown that for any prior $p(1) \geq \overline{p}^c$ (which implies $\alpha^r = \alpha^c = 1$), the second round experimentation intensity in the ring, β^r, is at least as high as its counterpart in the complete network, β^c. A direct comparison is not possible, as agents hold different posteriors. The posterior in the complete network after one round of failed experimentation is given by

$$p^c(2) = \frac{p(1)(1-\pi)^n}{p(1)(1-\pi)^n + 1 - p(1)},$$

and the posterior in the ring is given by

$$p^r(2) = \frac{p(1)(1-\pi)^3}{p(1)(1-\pi)^3 + 1 - p(1)}.$$

Let us first consider the beliefs for which efforts are the same in both networks. If $p^c(2) \geq \overline{p}^c$ we know that $p^r(2) > \overline{p}^r(2)$ and $\beta^c = \beta^r = 1$. On the other hand if $p^r(2) \leq \underline{p}^r(2)$, then $p^c(2) < p^a$ and $\beta^c = \beta^r = 0$. This implies that experimentation intensities in the two networks are the same for any $p^c(2) \in [0, \widetilde{\underline{p}}^r(2)] \cup [\overline{p}^c, 1]$, where

$$\widetilde{\underline{p}}^r(2) = \frac{\underline{p}^r(2)(1-\pi)^{n-3}}{\underline{p}^r(2)(1-\pi)^{n-3} + 1 - \underline{p}^r(2)}.$$

For β^r and β^c that maximize the agents' utility in the corresponding network in the interval where agents use both arms, the corresponding beliefs are given by

$$\begin{aligned} p^r(2) \;=\; & (1-\delta)|E_0| / \{(1-\delta)[|E_0| + E_1] + \delta E_1 \pi [(1-\pi)^2(1-\beta^r\pi)^2 \\ & -\delta^{\frac{n-3}{2}}(1-\pi)^n(1-\beta^r\pi)^{n-3}\beta^r\pi(2-\beta^r\pi) \\ & -[1-(1-\pi)^2(1-\beta^r\pi)^2] \sum_{t=1}^{\frac{n-1}{2}-2} \delta^t(1-\pi)^{2t}(1-\beta^r\pi)^{2t}]\} \end{aligned} \tag{A.8}$$

for the ring and

$$p^c(2) = \frac{(1-\delta)|E_0|}{(1-\delta)[|E_0| + E_1] + \delta E_1 \pi (1-\beta^c\pi)^{n-1}} \tag{A.9}$$

for the complete network. Further,

$$p^c(2) = \frac{p^r(2)(1-\pi)^{n-3}}{p^r(2)(1-\pi)^{n-3} + 1 - p^r(2)}. \tag{A.10}$$

Replacing $p^r(2)$ in Equation $(A.10)$ by $(A.8)$ and then solving for $(A.9) = (A.10)$ implies that

$$\begin{aligned} 0 \;=\; & (1-\beta^c\pi)^{n-1}(1-\pi)^3 - (1-\delta)[1-(1-\pi)^3] \\ & -\delta\pi[(1-\pi)^2(1-\beta^r\pi)^2 - \delta^{\frac{n-3}{2}}(1-\pi)^n(1-\beta^r\pi)^{n-3}\beta^r\pi(2-\beta^r\pi) \\ & -[1-(1-\pi)^2(1-\beta^r\pi)^2] \sum_{t=1}^{\frac{n-1}{2}-2} \delta^t(1-\pi)^{2t}(1-\beta^r\pi)^{2t}] \end{aligned}$$

which only holds if

$$\beta^r > \beta^c.$$

■

Proof of Proposition 4. Let us start by considering the central player. Comparing expected utility from experimenting with intensity α^h to not experimenting at all, it follows that using the risky arm is optimal for the hub as long as $(1-\delta)E_{p(1)} + \delta E_1 p\pi(1-\alpha^s\pi)^{n-1} \geq 0$. The cut-off belief above which an experimentation intensity of 1 is optimal for the hub is given by $\overline{p}^h = \overline{p}^c$, and the lower cut-off by $\underline{p}^h = p^a$. If $(1-\delta)E_{p(1)} + \delta E_1 p\pi(1-\alpha^s\pi)^{n-1} = 0$, then α^s is given by (6). This implies that the hub is indifferent between experimenting and not experimenting on the interval $[p^a, \overline{p}^c]$ if the peripheral players choose $\alpha^s = \alpha^c$. If $\alpha^s > \alpha^c$ for a given belief then $(1-\delta)E_{p(1)} + \delta E_1 p\pi(1-\alpha^s\pi)^{n-1} < 0$ and consequently the hub stops experimenting immediately. On the other hand if $\alpha^s < \alpha^c$, then $(1-\delta)E_p + \delta E_1 p\pi(1-\alpha^s\pi)^{n-1} > 0$ and the hub will exclusively use the risky option.

Peripheral players are symmetric and receive all their information from the hub. In $t = 1$ they are indifferent between the risky and the safe arm as long as $I_1^s = 0$, where

$$I_1^s = (1-\delta)E_{p(1)} + \delta E_1 p(1)\pi(1-\alpha^h\pi)(1-\delta+\delta(1-\alpha^s\pi)^{n-2}).$$

From this we can derive α^s and the corresponding cut-off beliefs $\underline{p}^s(1) = p^a$, and

$$\overline{p}^s(1) = \frac{(1-\delta)|E_0|}{(1-\delta)[|E_0| + E_1] + E_1\delta\pi(1-\alpha^h\pi)(1-\delta+\delta(1-\pi)^{n-2})}.$$

Existence and uniqueness of α^s can be easily verified by analyzing the expression I_1^s. The minimum of this function is at $\frac{1}{\pi}$ which implies that there is only one root on $[0, 1]$ due to the parabolic shape of the function. In the first round it can be shown that $\underline{p}^h = p^a = \underline{p}^s(1)$ and $\overline{p}^s(1) < \overline{p}^h = \overline{p}^c$, where the last inequality holds for all $\alpha^h \in [0, 1]$. Consequently, we can see that in the interval $[\overline{p}^s(1), \overline{p}^h)$, the peripheral players experiment with an effort level that violates the indifference condition of the hub (i.e., $\alpha^s > \alpha^c$) which implies that the central player will stop experimenting immediately for any belief slightly below $\overline{p}^h = \overline{p}^c$. For beliefs in $(p^a, \overline{p}^s(1))$, if the hub does not experiment, the experimentation intensity of the peripheral players is higher than it would be in a symmetric equilibrium in the complete network. Consequently, the hub does not experiment in this region either. More precisely, from comparing I_1^s and I_1^c we obtain that $\alpha^s > \alpha^c$ for all $\alpha^h \in [0, \alpha^c]$.

Let us turn to the problem in $t = 2$. As before, an optimal strategy requires either $(\alpha_i^s, 0)$ or $(1, \beta_i^s)$. After a first round where $p(1) \in [\overline{p}^c, 1]$ and hence $\alpha^h = \alpha^s = 1$, the posterior beliefs of the agents are $p^s(2) = \frac{p(1)(1-\pi)^2}{p(1)(1-\pi)^2+1-p(1)}$ for the peripheral players and

$p^h(2) = \frac{p^s(2)(1-\pi)^{n-2}}{p^s(2)(1-\pi)^{n-2}+1-p^s(2)}$ for the hub. Not only do agents now hold different beliefs, also the upper and lower cut-offs for the peripheral players are different due to the first round information that will still reach them. In the second round experimentation intensities of the peripheral players can be derived from

$$I_2^s = (1-\delta)E_{p^s(2)} + \delta E_1 p^s(2)\pi(1-\pi)^{n-2}(1-\beta^h\pi)(1-\delta+\delta(1-\beta^s\pi)^{n-2})$$

where, by imposing $I_2^s = 0$, we obtain

$$\overline{p}^s(2) = \frac{(1-\delta)|E_0|}{(1-\delta)[|E_0|+E_1] + E_1\delta\pi(1-\pi)^{n-2}(1-\beta^h\pi)(1-\delta+\delta(1-\pi)^{n-2}))}$$

and

$$\underline{p}^s(2) = \frac{(1-\delta)|E_0|}{(1-\delta)[|E_0|+E_1] + E_1\delta\pi(1-\pi)^{n-2}}.$$

We have $\overline{p}^h(2) = \overline{p}^c > \underline{p}^s(2) > p^a = \underline{p}^h$. Further $\overline{p}^s(2) > \overline{p}^c$ for $\beta^h = 1$. For $\beta^h \in [0,1)$ the ordering of $\overline{p}^s(2)$ and $\overline{p}^c$ depends on the relative magnitude of δ and π as well as on n. Now we want to show that it is still optimal that either all agents choose effort level 1 (for high beliefs), effort level 0 (for pessimistic beliefs) or the peripheral players choose $\beta^s \in (0,1)$ while the hub does not experiment. If agents in the complete network and the peripheral players have the same β as optimal effort level, then their beliefs are less than $n-2$ failed experiments apart from each other. This means that if the distance (measured in experiments) is $n-2$, the belief and effort level of the peripheral players is higher than for the complete network in the interval where agents use both arms. Then it is optimal for the hub to stop experimenting completely below $\overline{p}^c$. For beliefs slightly below $\overline{p}^c$, the peripheral players still experiment with full intensity whereas the hub wants to lower his intensity and hence chooses $\beta^h = 0$. It is straightforward to show that

$$\frac{\overline{p}^s(2)(1-\pi)^{n-2}}{\overline{p}^s(2)(1-\pi)^{n-2}+1-\overline{p}^s(2)} < \overline{p}^c.$$

Note that existence and uniqueness of β^s can be shown by analyzing the expression I_2^s based on the same arguments as for α^s. ■

Proof of Proposition 5. The proof consists of two parts. Part 1 is for beliefs such that in case all experiments in $t=1$ fail, there will be no experimentation in $t=2$. Part 2 describes the proof for beliefs where agents experiment in $t=1,2$. First, we know that for prior beliefs in $[0,p^a]$ and $[\overline{p}^c, 1]$ the first round experimentation intensity is the same in both networks. Hence, the interesting interval in the first round is $(p^a, \overline{p}^c)$ in which the hub does not experiment. Therefore, $n\alpha^c$ has to be compared to $(n-1)\alpha^s$. Note that we have in this interval along the equilibrium path

$$p^c(1) = \frac{(1-\delta)|E_0|}{(1-\delta)[|E_0|+E_1] + \delta E_1\pi(1-\alpha^c\pi)^{n-1}},$$

for the complete network and

$$p^s(1) = \frac{(1-\delta)|E_0|}{(1-\delta)[|E_0| + E_1] + \delta E_1 \pi (1-\delta + \delta(1-\alpha^s \pi)^{n-2})},$$

for the star. This implies that for a given fixed belief the relation between α^c and α^s can be found through these expressions and is given by

$$\alpha^c = \frac{1}{\pi} - \frac{1}{\pi}(1-\delta+\delta(1-\alpha^s\pi)^{n-2})^{\frac{1}{n-1}}.$$

Now the difference between $\frac{n-1}{n}\alpha^s - \alpha^c$ can be defined as

$$\Gamma_n(\delta, \pi, p(1)) := 1 - \delta - \left(1 - \frac{n-1}{n}\alpha^s \pi\right)^{n-1} + \delta(1-\alpha^s\pi)^{n-2}.$$

Based on the expression for $\Gamma_n(\delta, \pi, p(1))$ we can then define the region $S_n(p(1)) \subset [0,1]^2$ for $p(1) \in (p^a, \bar{p}^c)$ as

$$S_n(p(1)) := \{\delta, \pi \in [0,1]^2 : \ \Gamma_n(\delta, \pi, p(1)) > 0\}.$$

That is, $S_n(p(1))$ is the set of all combinations of δ and π for which $n\alpha^c < (n-1)\alpha^s$. Clearly, $\Gamma_n(\delta, \pi, p(1)) \to 1-\delta > 0$ as $n \to \infty$ for all $\delta, \pi \in [0,1]^2$ and hence $\lambda(S_n(p(1)) \to 1$ as $n \to \infty$.

If agents experiment as well in $t = 2$ a similar argument as above can be used with additionally making use of the fact that $p^c(2) = \frac{p^s(2)(1-\pi)^2}{p^s(2)(1-\pi)^2 + 1 - p^s(2)}$. That is, setting

$$p^c(2) = \frac{(1-\delta)|E_0|}{(1-\delta)[|E_0| + E_1] + \delta E_1 \pi (1-\beta^c\pi)^{n-1}},$$

it can be solved for $p^s(2)$, which has to be equal to

$$\frac{(1-\delta)|E_0|}{(1-\delta)[|E_0| + E_1] + \delta E_1 \pi (1-\pi)^{n-2}(1-\delta+\delta(1-\beta^s\pi)^{n-2})}.$$

Expressing β^c in terms of β^s, to find out whether $\beta^c \leq \frac{n-1}{n}\beta^s$ we analyze

$$\Gamma_n(\delta, \pi, p(2)) := 1 - \delta - \left(1 - \frac{n-1}{n}\beta^s\pi\right)^{n-1} + \delta(1-\beta^s\pi)^{n-2}) + \frac{(1-\delta)[1-(1-\pi)^{n-2}]}{\delta\pi(1-\pi)^{n-2}},$$

by the same arguments as for $\Gamma_n(\delta, \pi, p(1))$. $\Gamma_n(\delta, \pi, p(2))$ is equivalent to $\Gamma_n(\delta, \pi, p(1))$ up to replacing α^s by β^s and adding a positive constant. $S_n(p(2)) \subset [0,1]^2$ can be defined in an analogous way for $p^c(2) \in (\underline{\tilde{p}}^s(2), \bar{p}^c)$ as

$$S_n(p(2)) = \{\delta, \pi \in [0,1]^2 : \ \Gamma_n(\delta, \pi, p(2)) > 0\}.$$

■

Proof of Proposition 6. We start the proof by defining $F_{n,1}(\alpha)$, which is derived by considering the difference between I_1^s and I_1^r and imposing $\alpha^r = \alpha^s = \alpha$, i.e.,

$$F_{n,1}(\alpha) := \{\delta, \pi \in [0,1]^2 : \; I_1^r - I_1^s \geq 0\},$$

where $I_1^r - I_1^s$ for $\alpha^r = \alpha^s = \alpha$ is given by

$$[1-\delta(1-\alpha\pi)^2][\delta - 1 - \delta(1-\alpha\pi)^{n-2} + (1-\alpha\pi)^2] - [1-(1-\alpha\pi)^2]\delta^{\frac{n-1}{2}}(1-(1-\alpha\pi)^{n-1})].$$

This means $F_{n,1}(\alpha)$ represents all combinations of δ and π such that the stated inequality is satisfied, which in turn implies $\alpha^s \geq \alpha^r$ along the equilibrium path. For proving part (i) of the proposition it is easy to verify that for small n (e.g., $n = 3$) the inequality is satisfied for all $\delta, \pi \in [0,1]$. This suffices to conclude that there exists some finite $n_1 \in \mathbb{N}$ such that for all $n < n_1$ we have $F_{n,1}(\alpha) = [0,1]^2$. For the second part we explore the behavior of $F_{n,1}(\alpha)$ in the limit as $n \to \infty$ and obtain

$$F_1(\alpha) := \{\delta, \pi \in [0,1]^2 : \; [1-\delta(1-\alpha\pi)^2][\delta - 1 + (1-\alpha\pi)^2] \geq 0\},$$

where it can be shown that the inequality fails to hold for some values of δ and π implying that $F_1(\alpha)$ is a strict subset of $[0,1]^2$.

If agents experiment in $t = 2$ as well, we proceed in an analogous way replacing I_1^s and I_1^r with I_2^s and I_2^r and additionally making use of the fact that

$$p^r(2) = \frac{p^s(2)(1-\pi)}{p^s(2)(1-\pi) + 1 - p^s(2)}.$$

■

B Welfare Analysis

Before we start comparing welfare in the three networks for different beliefs, we state per capita expected equilibrium payoffs for each network. In the complete network we have

$$\begin{aligned} U^c(p(1)) &= \alpha^c(1-\delta)E_{p(1)} - \delta(1-\delta)|E_0|(1-p(1))\beta^c \\ &\quad + \delta E_1 p(1)[1-(1-\alpha^c\pi)^n[1-\delta-\beta^c(1-\delta)+\delta(1-\beta^c\pi)^n]]. \end{aligned}$$

For the ring we have

$$\begin{aligned} U^r(p(1)) &= \alpha^r(1-\delta)E_{p(1)} - \delta(1-\delta)|E_0|(1-p(1))\beta^r + \delta(1-\delta)\beta^r E_1 p(1)(1-\alpha^r\pi)^3 \\ &\quad + \delta E_1 p(1)[1-(1-\alpha^r\pi)^3 + \delta(1-\alpha^r\pi)^3[1-(1-\beta^r\pi)^3(1-\alpha^r\pi)^2]] \\ &\quad + p(1)E_1(1-\alpha^r\pi)(1-\beta^r\pi)[1-(1-\beta^r\pi)^2(1-\alpha^r\pi)^2] \times \\ &\quad \sum_{t=3}^{\frac{n-1}{2}} \delta^t (1-\alpha^r\pi)^{2(t-1)}(1-\beta^r\pi)^{2(t-1)} \\ &\quad + \delta^{\frac{n+1}{2}} p(1)E_1(1-\alpha^r\pi)^n(1-\beta^r\pi)^{n-2}[1-(1-\beta^r\pi)^2]. \end{aligned}$$

For the peripheral players and the hub in the star network, we have

$$\begin{aligned} U^s(p(1)) &= \alpha^s(1-\delta)E_{p(1)} + \delta E_1 p(1)[1-(1-\alpha^s\pi)(1-\alpha^h\pi)] \\ &\quad +\delta(1-p(1)+p(1)(1-\alpha^s\pi)(1-\alpha^h\pi))[\beta^s(1-\delta)E_{p(2)} \\ &\quad +\delta E_1 p(2)[1-(1-\alpha^s\pi)^{n-2}(1-\beta^s\pi)(1-\beta^h\pi)[1-\delta+\delta(1-\beta^s\pi)^{n-2}]], \end{aligned}$$

and

$$U^h(p(1)) = \begin{array}{l} \alpha^h(1-\delta)E_{p(1)} + \delta E_1 p(1)[1-(1-\alpha^s\pi)^{n-1}(1-\alpha^h\pi)] \\ +\delta(1-p(1)+p(1)(1-\alpha^s\pi)^{n-1}(1-\alpha^h\pi))[\beta^h(1-\delta)E_{p^h(2)} \\ +\delta E_1 p^h(2)[1-(1-\beta^s\pi)^{n-1}(1-\beta^h\pi)]], \end{array}$$

respectively. The expressions for the respective α and β can be found in Propositions 1, 2 and 4.

Proof of Proposition 7. Part (i) is obvious as for $p(1) \in [0, p^a]$ no one experiments in any network and hence expected payoffs are zero for all networks.
To show (ii) we compare $W^c(p(1))$ with $W^s(p(1))$ making use of the fact that in the relevant interval $I_1^c = 0$ and $I_1^s = 0$ and further

$$1-\delta+\delta(1-\alpha^s\pi)^{n-2} = (1-\alpha^c\pi)^{n-1}.$$

The result then follows from the fact that $\alpha^s > \alpha^c$ in equilibrium.
For (iii) we obtain the following. By comparing $W^c(p(1))$ and $W^r(p(1))$ for $p(1) \in [0, \bar{p}^r(1)]$ it is straightforward to show that $c \sim r$, as in this interval $I_1^c = 0$ and $I_1^r = 0$. $c \succ r$ for $p(1) \in (\bar{p}^r(1), 1]$ follows from discounting, i.e., the fact that $\delta < 1$.
For (iv), the condition

$$(1-\delta)(2p(1)-1)+\delta p(1)[(1-\pi)^{n-1}[1+\delta(n-1)]+(1-\delta)(n-1)(1-\pi)-n(1-\alpha^c\pi)^n] > 0.$$

is derived from $W^c(p(1)) - W^s(p(1))$. That is, if this inequality is satisfied, we have $W^c(p(1)) - W^s(p(1)) > 0$ for $p(1) \in (\bar{p}^s(1), \bar{p}^c(1))$.
Finally, a comparison of $W^c(p(1))$ and $W^s(p(1))$ for the case when all agents in both networks experiment with full intensity, shows that due to discounting, expected payoffs are higher in the complete network, which proves part (v). ■

Network optimality for arbitrary costs. In Proposition 7 we set $k = 0$ and did not consider experimentation in $t = 2$. We will now discuss network optimality for any arbitrary $k \geq 0$ allowing experimentation in $t = 2$. As before payoffs are compared for different intervals of the prior. We use the cut-off beliefs in order to specify intervals for $p(1)$ such that within each interval equilibrium actions do not change and hence a comparison across different structures is possible. More precisely, ordering the cut-offs accordingly (e.g., $p^a, \overline{p}^s(1), \overline{p}^r(1), \overline{p}^c$) we separate the $[0,1]$ interval into $M+1$ subsections that are defined by the cut-off beliefs. Each interval is denoted by τ_m, where $m = 0, ..., M$, so that τ_0 denotes the interval $[0, p^a]$ and so on. Lemma 1 states how to explicitly calculate welfare for different beliefs in the three networks and derive conditions on the cost parameter k. In particular, the terms $K^{\tau_m}_{g,g'}$ represent indifference conditions where $k = K^{\tau_m}_{g,g'}$ implies that $W^g(p(1)) = W^{g'}(p(1))$ in the given interval so that the social planner is indifferent between network g and g'.

Lemma 1. *The following conditions characterize optimal networks: For a given* τ_m

- c *is preferred to* s *and* r *if* $k < \min\{K^{\tau_m}_{c,r}, K^{\tau_m}_{c,s}\}$
- s *is preferred to* c *and* r *if* $k > \max\{K^{\tau_m}_{c,s}, K^{\tau_m}_{r,s}\}$,
- r *is preferred to* c *and* s *if* $\max\{K^{\tau_m}_{c,s}, K^{\tau_m}_{r,s}\} > k > \min\{K^{\tau_m}_{c,r}, K^{\tau_m}_{c,s}\}$,

where $\{\tau_m\}^M_{m=0}$ *specifies the relevant interval for* $p(1)$ *and* $K^{\tau_m}_{g,g'}$ *represents the cost level* k *for which* $W^g(p(1)) = W^{g'}(p(1))$ *in the interval* τ_m.

Proof of Lemma 1. First, we derive our intervals of interest and denote them for convenience by $\tau_0, \tau_1, ...$ We have $\tau_0 = [0, p^a]$, $\tau_1 = (p^a, \min\{\overline{p}^r(1), \overline{p}^s(1)\}]$, $\tau_2 = (\min\{\overline{p}^r(1), \overline{p}^s(1)\}, \max\{\overline{p}^r(1), \overline{p}^s(1)\}]$ and $\tau_3 = (\max\{\overline{p}^r(1), \overline{p}^s(1)\}, \overline{p}^c]$. If agents experiment in $t = 2$ as well, the intervals are defined in a similar way taking the difference in posterior beliefs into account. Based on second round cut-offs, p^a, $\underline{p}^r(2)$, $\underline{p}^s(2)$, $\overline{p}^r(2)$, $\overline{p}^s(2)$ and $\overline{p}^c$, we can derive the corresponding intervals for prior beliefs. Given the expressions for equilibrium payoffs and the definition of welfare above, we then calculate for every possible pair of networks g and g' and every interval $\{\tau_m\}^M_{m=0}$, the cost level k for which the equality $W^g(p(1)) = W^{g'}(p(1))$ is satisfied. We denote the cost level k such that that $W^c(p(1)) = W^s(p(1))$ at τ_0 by $K^{\tau_0}_{c,s}$. This means that at $k = K^{\tau_0}_{c,s}$ we obtain indifference between c and s. These values can be calculated for every interval and every network.

■

B.1 Example: Optimal networks for four players

To provide a better understanding of the conditions determining which network is optimal, we provide here the details for $n = 4$. First, we analyze which network is optimal for an interval of prior beliefs such that there will be at most one round of experimentation in each network. That is, the posterior belief after the first round is for every agent in every network below the lower cut-off: $p(2) \leq \underline{p}^g(2)$, for all g.

Part (1) - Optimal networks, One Round. The following conditions characterize optimal networks:

- $p(1) \in [0, \overline{p}^s(1)]$: s is preferred to r and c for all $k \geq 0$;
- $p(1) \in (\overline{p}^s(1), \overline{p}^r(1)]$: c is optimal only if $k = 0$. s is optimal if $k \geq K^{\tau_1}_{s,r}$, and r if $k \leq K^{\tau_1}_{s,r}$, where

$$K^{\tau_1}_{s,r} = (1-\delta)(1-2p(1))\left(3 - \frac{4}{\pi}\right) + \delta p(1)(1-\pi)[(2-\pi)(2-3\delta\pi) + \pi^2];$$

- $p(1) \in (\overline{p}^r(1), \overline{p}^c]$: c is optimal for $k \leq K^{\tau_2}_{c,r}$, with

$$K^{\tau_2}_{c,r} = 2(1-\delta)(1-2p(1))\left(1 - \frac{1}{\pi}\right) + 2\delta p(1)(1-\pi)^3(1-\delta\pi),$$

  s for $k \geq K^{\tau_2}_{r,s}$, where

$$K^{\tau_2}_{r,s} = (1-\delta)(2p(1)-1) + \delta p(1)\pi(1-\pi)[3(2-\pi)(1-\delta) + 4\delta(1-\pi)^2],$$

  and r for $K^{\tau_2}_{c,r} \geq k \geq K^{\tau_2}_{r,s}$;

- $p(1) \in (\overline{p}^c, 1]$ and $p(2) \leq \underline{p}^g(2)$, for all $g \in G$: c is optimal for $k \leq K^{\tau_3}_{c,r}$, where

$$K^{\tau_3}_{c,r} = \delta p(1)\pi(1-\pi)^2(1-\delta)(2-2\pi),$$

  s for $k \geq K^{\tau_3}_{s,r}$,

$$K^{\tau_3}_{s,r} = \delta p(1)\pi(1-\pi)^2(1-\delta)(2+\pi),$$

  and r for $K^{\tau_3}_{s,r} \geq k \geq K^{\tau_3}_{c,r}$.

Proof. In the proof we compare payoffs for different beliefs and networks. For $p(1) \in [0, p^a]$, payoffs are zero. Consequently, the structure with the lowest costs, i.e., the star network, is optimal. For $p(1) \in (p^a, \overline{p}^s(1)]$, optimal experimentation intensities are $\alpha^h = 0$, $\alpha^r, \alpha^c \in (0,1)$ and $\alpha^s \in (0,1]$. For the complete network we obtain

$$W^c = 4\alpha^c(1-\delta)(2p(1)-1) + 4\delta p(1)[1-(1-\alpha^c\pi)^4] - 6k,$$

where we know that in this interval of beliefs $(1-\delta)(2p(1)-1) + \delta p(1)\pi(1-\alpha^c\pi)^3 = 0$ by $I^c_1 = 0$ which implies that

$$W^c = 4\delta p(1)[1-(1-\alpha^c\pi)^3] - 6k.$$

Further, we obtain $W^r = 4\delta p(1)[1-(1-\alpha^r\pi)^2(1-\delta\alpha^r\pi)] - 4k$, and $W^s = \delta p(1)[1-(1-\alpha^s\pi)^3 + 3\delta\alpha^s\pi(2-\alpha^s\pi)] - 3k$. Further we know from I_1^s, I_1^r and I_1^c that $(1-\alpha^r\pi)^2(1-\delta\alpha^r\pi) = (1-\alpha^c\pi)^3 = 1-\delta\alpha^s\pi(2-\alpha^s\pi)$. This implies that $W^c - W^r = -2k$, and $W^r - W^s = \delta p(1)[(1-\alpha^s\pi)^3 - (1-\alpha^r\pi)^2(1-\delta\alpha^r\pi)] - k$, where the term in square brackets is negative as $\alpha^s > \alpha^r$. Further, $W^c - W^s \leq W^r - W^s$ for $k \geq 0$.

For $p(1) \in [\overline{p}^s(1), \overline{p}^r(1))$ peripheral players exert full effort, while in the other networks the agents exert effort below 1. This changes welfare in the star network and leaves the ring and the complete network unaltered. As the ring is still (weakly) preferred to the complete network for $k \geq 0$ we restrict attention to the condition under which $r \succ s$. From $W^r - W^s$, where

$$W^s = 3(1-\delta)(2p(1)-1) + \delta p(1)[4-(1-\pi)^3 - 3(1-\pi)[1-\delta\pi(2-\pi)]] - 3k,$$

and further from $I_1^s = 0$ we have that $(1-\alpha^r\pi)^2(1-\delta\alpha^r\pi) = \frac{(1-\delta)(1-2p(1))}{\delta p_1 \pi}$. It follows that for $k \leq K_{s,r}^{\tau_1}$ the ring is optimal and for $k \geq K_{s,r}^{\tau_1}$, the star. $K_{s,r}^{\tau_1}$ becomes positive in the interval $[\overline{p}^s(1), \overline{p}^r(1))$.

For $p(1) \in (\overline{p}^r(1), \overline{p}^c]$, the comparison with the ring network changes, where now we have $W^r = 4(1-\delta)(2p(1)-1) + 4\delta p(1)[1-(1-\pi)^3(1-\delta\pi)] - 4k$. From $W^c - W^r$ we obtain $K_{c,r}^{\tau_2}$, which equals zero at $\overline{p}^r(1)$ and is above zero at $\overline{p}^c$ for all $\delta, \pi \in (0,1)$. By finding k such that the difference between $W^r - W^s = 0$ we derive $K_{r,s}^{\tau_2}$ and the difference between W^c and W^s yields

$$K_{c,s}^{\tau_2} = \delta p(1)(1-\pi)\left[\frac{1}{3}(1-\pi)^2 + 1 - \delta\pi(2-\pi)\right] + (1-\delta)(1-2p(1))\left(1-\frac{4}{3\pi}\right).$$

The complete network is optimal if $k \leq \min\{K_{c,s}^{\tau_2}, K_{c,r}^{\tau_2}\}$, where the minimum of the two is given by $K_{c,r}^{\tau_2}$ in the interval of interest. This can be seen by analyzing $K_{c,r}^{\tau_2} - K_{c,s}^{\tau_2}$, which is negative at the interval boundaries $\overline{p}^r(1)$ and $\overline{p}^c$. As $K_{c,r}^{\tau_2} - K_{c,s}^{\tau_2}$ is linear in $p(1)$ this implies that it is negative for all $p(1) \in [\overline{p}^r(1), \overline{p}^c]$. The star is optimal if $k \geq \max\{K_{c,s}^{\tau_2}, K_{r,s}^{\tau_2}\}$ where $K_{r,s}^{\tau_2}$ is the maximum in this interval and consequently the ring is optimal for $K_{r,s}^{\tau_2} \geq k \geq K_{c,r}^{\tau_2}$. $K_{c,r}^{\tau_2}$ equals zero at $\overline{p}^r(1)$ and is positive at $\overline{p}^c$ for all $\delta, \pi \in [0,1]$.

For $p(1) \in (\overline{p}^c, 1]$ and $p(2) \leq \underline{p}^g(2)$ for all g in all networks, all agents experiment one time with full effort. We obtain $K_{c,r}^{\tau_3}$ from $W^c - W^r$, $K_{r,s}^{\tau_3}$ from $W^r - W^s$ and $K_{r,s}^{\tau_3} > \delta p(1)\pi(1-\pi)^2(2-\pi) > K_{c,r}^{\tau_3}$, where $\delta p(1)\pi(1-\pi)^2(2-\pi) = K_{c,s}^{\tau_3}$ is obtained from $W^c - W^s$.

■

Part (2) - Optimal networks, Two Rounds. For two rounds of experimentation, that is, $\alpha^g = 1$, for all g and $\beta^g > 0$ for at least one g we have for

- $p^c(2) \in [0, p^a]$, $p^r(2) \in [0, \underline{p}^r(2)]$, $p^s(2) \in (\underline{p}^s(2), \overline{p}^s(2)]$: c is optimal if $k \leq \min\{K^{\tau_4}_{c,r}, K^{\tau_4}_{c,s}\}$, where $K^{\tau_4}_{c,r} = K^{\tau_3}_{c,r}$,

$$\begin{aligned} K^{\tau_4}_{c,s} &= \delta(1-\pi)^2 p(1)\left[(1-\delta)\left(1-\frac{1}{\pi}\right) - (1-\pi)^2\left(\frac{\delta}{3} + 1 - \frac{\delta}{3}(1-\beta^s\pi)^3\right)\right] \\ &\quad + \delta(1-\delta)\frac{1}{\pi}(1-p(1)), \end{aligned}$$

and β^s from Proposition (4). s is optimal for $k \geq \max\{K^{\tau_4}_{c,s}, K^{\tau_4}_{r,s}\}$, where

$$\begin{aligned} K^{\tau_4}_{r,s} &= \delta(1-\pi)^2 p(1)\left[(1-\pi)[(1-\delta)(1-\pi) - 4(1-\delta\pi)] + 3\left(1-\frac{1}{\pi}\right)(1-\delta)\right] \\ &\quad + 3\delta\frac{1}{\pi}(1-\delta)(1-p(1)) + \delta^2 p(1)(1-\pi)^4(1-\beta^s\pi)^3, \end{aligned}$$

and r for $\max\{K^{\tau_4}_{c,s}, K^{\tau_4}_{r,s}\} \geq k \geq \min\{C_3, K^{\tau_4}_{c,s}\}$;

- $p^c(2) \in [0, \overline{p}^c]$, $p^r(2) \in (\underline{p}^r(2), \overline{p}^r(2)]$, $p^s(2) \in (\underline{p}^s(2), \overline{p}^s(2)]$: $s \succ c, r$ for all $k \geq 0$.

- $p^c(2) \in (p^a, \overline{p}^c]$, $p^r(2) \in (\underline{p}^r(2), \overline{p}^r(2)]$, $p^s(2) \in (\overline{p}^s(2), 1]$: c is never optimal as soon as $k \geq 0$. s is optimal if $k \geq K^{\tau_6}_{r,s}$, where

$$\begin{aligned} K^{\tau_6}_{r,s} &= \delta^2 p(1)(1-\pi)^5[(1-\pi)^2 + 3 - 3\delta\pi(2-\pi)] \\ &\quad + \delta(1-\delta)(1-p(1)-p(1)(1-\pi)^3)\left(3-\frac{4}{\pi}\right) + \delta(1-\delta)p(1)\pi(1-\pi)^3, \end{aligned}$$

and r is optimal for $k \leq K^{\tau_6}_{r,s}$.

- $p^c(2) \in (p^a, \overline{p}^c]$, $p^r(2) \in (\overline{p}^r(2), 1]$, $p^s(2) \in (\overline{p}^s(2), 1]$: $c \succ s, r$ for $k \leq \min\{K^{\tau_7}_{c,r}, K^{\tau_7}_{c,s}\}$, where

$$K^{\tau_7}_{c,r} = 2\delta\left[p(1)(1-\pi)^7\delta(1-\delta\pi) + (1-\delta)\left(1-\frac{1}{\pi}\right)[1-p(1)-p(1)(1-\pi)^4]\right],$$

and

$$K^{\tau_7}_{c,s} = \delta\left[\begin{array}{c} (1-\delta)\left(\frac{4}{3\pi}-1\right)(p(1)(1-\pi)^4 - 1 + p(1)) \\ +\delta p(1)(1-\pi)^5[1-\delta\pi(2-\pi)+\frac{1}{3}(1-\pi)^2] \end{array}\right],$$

s for $k \geq \max\{K^{\tau_7}_{c,s}, K^{\tau_7}_{r,s}\}$, where $K^{\tau_7}_{r,s}$ is given by

$$\delta^2 p(1)(1-\pi)^5\pi[3(2-\pi)(1-\delta) + 4\delta(1-\pi)^2]]\delta(1-\delta)(1-p(1)-p(1)(1-\pi)^4),$$

and the ring if $\max\{K^{\tau_7}_{c,s}, K^{\tau_7}_{r,s}\} \geq k \geq \min\{K^{\tau_7}_{c,r}, K^{\tau_7}_{c,s}\}$;

- $p^c(2) \in (\overline{p}^c, 1]$, $p^r(2) \in (\overline{p}^r(2), 1]$, $p^s(2) \in (\overline{p}^s(2), 1]$: c is optimal for $k \leq K^{\tau_8}_{c,r} = \delta^2 p(1)\pi(1-\pi)^6(1-\delta)(2-2\pi)$, s is optimal for $k \geq K^{\tau_8}_{r,s} = \delta^2 p(1)\pi(1-\pi)^6(1-\delta)(2+\pi)$, and r for $K^{\tau_8}_{r,s} \geq c \geq K^{\tau_8}_{c,r}$.

Proof. The proof is analogous to the proof for one round, simply taking the differences in posterior beliefs into account. Details are available upon request. ■

References

Acemoglu, D., Bimpikis, K., Ozdaglar, A. (2014), Dynamics of information exchange in endogenous social networks, *Theoretical Economics,* Vol. 9: 41-97.

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

Bimpikis, K., Drakopoulos, K., (2014) Disclosing information in strategic experimentation, *working paper*, Stanford University and Massachusetts Institute of Technology.

Bolton, P., Harris, C. (1999), Strategic experimentation, *Econometrica 67:* 349-374.

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

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

Calvó-Armengol, A., Jackson, M. (2004), The effects of social networks on employment and inequality, *American Economic Review*, 94: 426-454.

Chatterjee, K., Dutta, B. (2012), Creditability and strategic learning in networks, *The Warwick Economic Research Papers Series (TWERPS)* No. 972, University of Warwick.

Conley, T.G., and C. Udry, (2010), Learning about a new technology: Pineapple in Ghana", *American Economic Review* 100: 35-69.

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

Goyal, S. (2009), *Connections: An Introduction to the Economics of Networks,* Princeton University Press.

Heidhues, P., Rady, S., Strack, P. (2012), Strategic experimentation with private payoffs, *SFB/TR 15 Discussion Paper Series* No. 387.

Jackson, M. (2008), *Social and Economic Networks*, Princeton University Press.

Keller, G., Rady, S., Cripps, M. (2005), Strategic experimentation with exponential bandits, *Econometrica* 73: 39-68.

Nobel, R., Birkinshaw, J. (1998), Innovation in multinational corporations: Control and communication patterns in international R&D operations, *Strategic Management Journal* 19: 479-496.

Rosenberg, D., Solan, E., Vieille, N. (2009), Information externalities and emergence of consensus, *Games and Economic Behavior* 66: 979-994.

Sadler, E. (2014), Innovation adoption and collective experimentation, *working paper,* New York University.

Teece, D. (1994) Firm organization, industrial structure, and technological innovation, *Journal of Economic Behavior & Organization* 31(2): 193-224