# Network Effects on Strategic Interactions: A Laboratory Approach

Yang Zhang[1], Xiaomin Du
Dept of Industrial Engineering, Tsinghua University
yangzhanguser@mail.tsinghua.edu.cn, dxm16@mails.tsinghua.edu.cn

*We examine the effect of interaction structure (network) on two classes of collective activities, herding and shirking, respectively referring to the situation where a player's incentive to take a certain action increases and decreases if more of her network neighbors follow the same action. In our experiment, we find that subjects do not act according to theoretical equilibrium, and their frequencies of making the socially beneficial choice in herding and shirking games are inversely influenced by the number of network neighbors they have. Moreover, the observed local network effect is stronger in shirking games, while the global network effect is more significantly present in herding games. We explain the behavioral regularities through a hybrid learning model, which extends SEWA learning into a network context. As such, our learning model provides a foundation for the observed dynamics, disequilibrium behavior, as well as the local and global network effects.*

*Keywords: social networks, game theory, strategic complement, strategic substitute, behavioral experiment, learning.*

[1] This work is funded by National Natural Science Foundation of China (NSFC) under Grant No. 71501108, and Beijing Natural Science Foundation under Grant No. 9164030. Yang Zhang is the corresponding author.

## 1. Introduction

Social and economic networks serve as platforms for trade and information exchange, and correlate the activities of separate individuals (Sundararajan et al. 2013). Activities of users of online networking sites, such as liking and following a movie or subscribing to a club, can generate considerable influences on their peers. This gives rise to positive correlations of demands of similar products or services across friends (Aral and Walker 2011, Oestreicher-Singer and Sundararajan 2012a, 2012b). In other occasions, social networks help coordinate the individual choices, such as vaccination against viruses (Rao et al. 2007), the study of foreign languages (Galeotti et al. 2010), and software compatibility between coworkers (Lee et al. 2006). By means of laboratory experiments, this paper studies how the decentralized interaction of networked agents is affected by the network structure, which indicates "who interacts with whom". While the problem is generally difficult to tackle in both theory and empirical research, we are able to decompose multiple sources of network effects on agent activity in our experiment due to the incomplete information setting we consider. Our research goes beyond the linkage between aggregate network configurations and collective economic consequences, and inspect how the network affects activities at individual level (rather than group level). In this line, our study reveals how boundedly rational agents make use of the network information and adjust their strategies through learning, which shapes the aggregate observations.

In this paper, we represent decentralized activities in networks by two classes of games, *herding* / *shirking*, referring to the case where a player's incentive to take a certain action increases / decreases if more players in her network neighborhood follow the same action. Typical applications include technology adoptions under network effects (herding), and voluntary reduction of pollution (shirking). These concepts are illustrated by the examples that follow this paragraph. From the viewpoint of management and social planning, it is crucial to understand: *How does the network structure leverage the outcomes of herding and shirking, and foster socially beneficial behavior in strategic contexts? Why does the network exhibit the observed effect on behavior?*

**Example 1. Product adoption under network effects (herding)**. In the classical models for products with network effect (e.g. Katz and Shapiro. 1992, 1986, 1985), it is assumed that user adoptions of the product are affected by the entire user network size. Sometimes, however, a user's benefit from buying a product depends on the adoption decisions from a subset of agents in the user network, with whom he/she/it needs that product to interact. For example, suppose a firm decides whether to adopt the RFID (Radio Frequency Identification) system which carries readers and tags only working for the same type of system. For the system to function, the firm has to coordinate its installation decision with its business partners, but obviously it does not need to coordinate its adoption with every other firm. If we connect each company with the partners it needs to exchange data with RFID, we have obtained a network. In this language, Katz-Shapiro models studied the case where the network is *complete* (everyone connected to everyone else), whereas we concentrate on situations where each firm or customer only coordinates with a subset of the user population.

**Example 2. Pollution reduction (shirking)**. In many cases, pollution in one region yields a negative impact on the environment of all adjacent regions. Examples include water contamination, air pollution, among others. As such, the *costly* effort of reducing pollution made by the administration of one region can also benefit neighboring regions, giving rise to a shirking problem: Every region aspires to clean environment, but at the same time wants to free ride on neighbor's pollution reduction. The collective outcome of pollution reduction and its environmental consequence will then depend on the connection structure – "whose neighbor is whom".

In our experiment, subjects have incomplete information of the network structure. To be specific, subjects know about the number of their network neighbors, but have only distributional knowledge about the network structure outside their neighborhood. The setting of incomplete network information captures agents' cognitive limitation in large social networks, and gives rise to monotonic network effects in theory for both herding and shirking games (Galeotti et al. 2010, Sundararajan 2007). In our experiment we observe some of the network effects inspired by the theory, and suggest a set of simple topological measures (e.g. degree, density) that provide sufficient predictions for strategic behavior in networks.

A major contribution of this paper lies in the behavioral model established to answer the research question. The model extends the SEWA learning paradigm (Chong et al. 2006, Ho et al. 2007) into a network context, and explains the emergence of network effects, deviation from equilibrium, and behavioral dynamics as a consequence of individual level learning. We show that it is the same type of learning that governs both herding and shirking in networks, despite the opposite appearances of the two games. The effectiveness of the proposed behavioral model is demonstrated through comparisons with alternative models of adaptive learning and learning towards equilibrium.

In Section 2, we review the relevant literature. Section 3 introduces the experimental design and empirical hypotheses. In Section 4 we outline the basic findings from our data. Then in Section 5, the learning model is developed to characterize the individual-level behavior. Section 6 concludes with further discussions.

## 2. Literature Background

Decentralized interactions in networks have been gaining attention in economics and operations research in recent years. Theoretical models can be classified by the nature of games being studied and the settings of information. Assuming agents have full information about the network layout, Ballester et al. (2006), Bramoullé and Kranton (2007) and Bramoullé et al. (2014) investigate games with linear best replies. The games they explore exhibit strategic complementarity or strategic substitutability, which correspond to the notion of herding or shirking referred in our paper. Another stream of theoretical literature embraces incomplete information of the network structure, including Sundararajan (2007) on strategic complementary games, and Galeotti et al. (2010) on both games of strategic complements and strategic substitutes. Galeotti and Goyal (2010), Cho (2010), Hojman and Szeidl (2008) examine games of submodular nature, but with networks endogenously formed by the players. For more information on the literature of network games, we refer the readers to the recent surveys: Jackson (2008), Ioannides (2012), Jackson and Zenou (2015), and Jackson et al. (2017). Most existing works in this area draw upon standard game theory, assuming players are perfectly rational. In contrast, our paper takes an experimental approach which allows for the examination of bounded rationality.

In terms of experimental research, traditional laboratory games on collective activities have been focusing on the case where subjects interact globally, e.g. Van Huyck et al (1990) and Weber (2006) for coordination (herding) game, and Ledyard (1997) for public goods (shirking) game. In more recent works, the concept of social network is introduced into the experiment (see Kosfeld (2004) for a review). Keser et al. (1998), Berninghaus et al. (2002), Cassar (2007), Corbae and Duffy (2008) study the role of interaction structures in shaping the outcomes of coordination, cooperation, and equilibrium selection. Due to complexity of network structures, these works above only consider bilateral games played between each pair of connected players. Our game adopts a *multilateral* setup, which allows us to embed network configurations into the equilibrium predictions (see Section 3).[2] Closer to our work, Judd et al. (2010) and Rosenkranz and Weitzel (2012) analyze network experiments of coordination game and public goods game respectively, and Charness et al. (2014) studies the network game experiments under various informational

[2] With bilateral games, the player collects payoff from separate games played with each of her connected partners (referred as neighbors) in the network. In our game however, the player reaps profit from one unified game played multilaterally with all her neighbors. While in both cases (bilateral and multilateral games) the player has to take into account all neighbors' actions simultaneously, our setup allows us to decompose the effect of one's local connectivity and that of global network density on subject behavior under incomplete information.

conditions. Compared to those works, our paper is unique in the proposed learning model, which bridges the individual-level behavior and aggregate experimental observations. Our model embeds SEWA learning theory into the network context. In that way, our paper draws connections between network games and the experimental literature on learning (see Fudenberg and Levine 1998, Erev and Roth 1998, Camerer and Ho 1999, Camerer et al. 2002, Chong et al. 2006, Ho et al. 2007, among others).

While individual learning consists of an integral part of our analysis, we are aware of the separate empirical literature on social learning in network environments, such as Foster and Rosenzweig (1995), Conley and Udry (2001), Miguel and Kremer (2003), Munshi (2004), Mobius et al. (2007), Rao et al. (2007), etc. The main body of this research focuses on how the agent learns the values of her choices (e.g. receiving a vaccine, adopting an agricultural technique) through social networks, rather than the coordination of agent choices in networks, which is what we choose to study. Also, most empirical works in this area lack the analysis of comparative statics that describe the way equilibria change with the network configuration, which we are able to test in our experiment.

Our paper is practically motivated by the management science literature on strategic customer behavior. This stream of research broadly inspects the gaming between consumers and the selling firm, or between consumers themselves, in the purchase of certain products or services. In a broad sense, the afore-mentioned examples of decentralized product adoption under network effect (Katz and Shapiro 1992, 1986, 1985, Riggins et al. 1994, Lee and Mendelson 2008, etc.) can be classified as studies of strategic customer behavior. In the domain of revenue management, there is research that concentrates on the inventory and pricing policies of the firm when selling to strategic customers: Elmaghraby et al. (2008), Aviv and Pazgal (2008), Su and Zhang (2008, 2009), Cachon and Swinney (2009, 2011), among others. As pointed out by Liu and van Ryzin (2008), however, strategic interactions among customers may be constrained by the physical or social locations of individuals. As a result, an individual consumer in the marketplace might base her decision on the acts of people who hold some certain ties with her (i.e. being physically or socially around her). In that sense, our research has potential implications on how strategic customers coordinate their purchases in networked environments, which could inspire further studies on the resultant sales strategy of the firm.

## 3. Experimental Design and Hypotheses

### 3.1. Network

In our experiment, subjects are connected according to the network structures $G_l, G_h$ in Figure 1. Directly connected players are *neighbor*s. The number of one's neighbors is her *degree*. Subjects do not know the structure of network they live in (i.e. either network in Figure 1). Instead, each subject knows her own degree, and perceives neighbor degrees as random variables, for which the distribution is known to the subject. As discussed in Section 1, the setting of incomplete network information reflects the potential cognitive limitation that players have regarding the network institution (e.g. I know how many friends I have, but am uncertain about the number of friends of my friends).

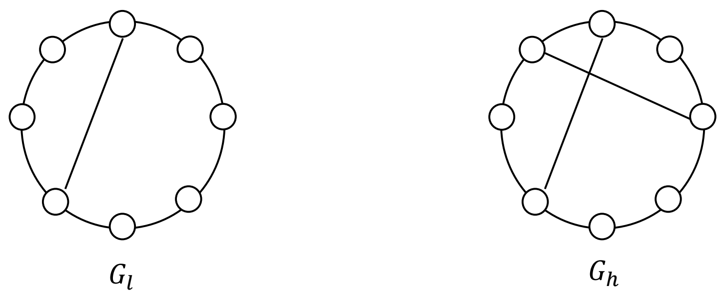


Figure 1. Network structures used in experiments

That said, what are the distributions of neighbor degrees in our laboratory networks? In the experiment, the subjects are randomly assigned over the network positions conditional on their degrees in an equally likely manner. This leads to the following neighbor degree distributions conditional on one's own degree (Table 1).

Table 1. The neighbor degree distributions for $G_h$ and $G_l$

| $G_h$ | | | $G_l$ | | |
|---|---|---|---|---|---|
| If the player is degree-2, | | | | | |
| probability | No. degree-3 neighbors | No. degree-2 neighbors | probability | No. degree-3 neighbors | No. degree-2 neighbors |
| 1/2 | 1 | 1 | 2/3 | 1 | 1 |
| 1/2 | 2 | 0 | 1/3 | 0 | 2 |
| If the player is degree-3, | | | | | |
| probability | No. degree-3 neighbors | No. degree-2 neighbors | probability | No. degree-3 neighbors | No. degree-2 neighbors |
| 1/2 | 2 | 1 | 1 | 1 | 2 |
| 1/2 | 1 | 2 | | | |

Under incomplete network information, one's own degree characterizes her *local network*, while her neighbor degree distribution (conditional on her own degree) measures the network outside her neighborhood. Specifically, the set of conditional neighbor degree distributions defines *network density* in a comparative sense: If every degree type in the network faces stochastically higher conditional neighbor degree distribution, then the overall *network density* is higher (Galeotti et al. 2010). In this paper we use network density as a measure of the *global network*. It is easy to show in Figure 1 (and confirm our intuition) that $G_h$ has higher density than $G_l$ (see Remark A-1, Appendix A). Formal definitions of neighbor degree distribution and network density are found in Appendix A.

It is worth pointing out that, the design of networks in our experiment is simple yet not restrictive. Note that the topological measures in our design, degree and network density, are independent of the network size and connection details. Under incomplete network information, subjects are informed of degree and network density, rather than the network size or specific patterns of connectivity (i.e. $G_h$ or $G_l$). Therefore, their behavior (as well as formal equilibrium) should not be sensitive to the specific network size or connectivity being used in the laboratory. Proposition A-1 and A-2 in Appendix A elaborate how the equilibria of the game are determined by degree and network density, which forms the basis of our hypotheses in Section 3.3. Alternative network designs, such as using random graphs (c.f. Bollobas 1985, Erdos and Renyi 1960) or scale free networks (Barabási and Albert, 1999) as the underlying network structures, would not provide the same degree of control and conciseness as does the incumbent design (which features fixed degree support and intuitively comparable density).

### 3.2. Game

In our experiment, subjects play the games of herding and shirking as shown in Table 2 and 3. In any game the subject faces a binary choice (0 as *low action*, or 1 as *high action*), and her profit depending on her own choice and choices of neighbors. As shown in Table 2 and 3, as the number of neighbors who select 1 increases, the best response of the player in question switches from 0 to 1 (from 1 to 0), reflecting the nature of herding (shirking). Notice that, in the context of this paper, the incentives for herding and shirking are respectively similar to those present in tacit coordination games (Van Huyck et al 1990) and public goods game of *threshold* structure (Cadsby and Maynes, 1999), with which we will frequently draw analogy when discussing our results. In both settings of herding and shirking, Choice 1 leads to higher social welfare than Choice 0 (see Remark 1 below). Therefore, Choice 1 (or Action 1) is called *socially beneficial choice*. The existence of socially beneficial choice is common in many collective activities, such as coordination (herding) games in Van Huyck et al. (1990), Harsanyi and Selten (1988) and public goods (shirking) games in Fehr and Gächter (2000). It also allows us to conveniently evaluate the system performance by the society-wise rate of adopting the socially beneficial choice (see Section 4). Also to be noted is our control for economic returns in the experiment: The rows of payoffs in Table 2 and 3 are exactly inversed, so that the two games under examination only differ in the direction, rather than the amount, of payoff difference between actions associated with each level of neighbor actions.

Table 2. The herding game

| Your Profit | | Number of Your Neighbors who Choose 1 | | | |
|---|---|---|---|---|---|
| | | 0 | 1 | 2 | 3 |
| Your Choice | 1 | 0 | 100 | 275 | 335 |
| | 0 | 100 | 175 | 225 | 260 |

Note: The payoff table presented here is used for degree-3 subjects. Degree-2 subjects shall see the same payoff table except for the last column excluded. See the instruction in Appendix C.

Table 3. The shirking game

| Your Profit | | Number of Your Neighbors who Choose 1 | | | |
|---|---|---|---|---|---|
| | | 0 | 1 | 2 | 3 |
| Your Choice | 1 | 100 | 175 | 225 | 260 |
| | 0 | 0 | 100 | 275 | 335 |

Note: The payoff table presented here is used for degree-3 subjects. Degree-2 subjects shall see the same payoff table except for the last column excluded. See the instruction in Appendix C.

To summarize, our experiment manipulates the *network structure* and the *type of game* that happens on the network. Each factor takes on two levels in a fully crossed 2×2 factorial design (high and low network density, herding and shirking games), yielding a total of four treatments, *Gl_Herd*, *Gh_Herd*, *Gl_Shirk*, *Gh_Shirk*, as recorded in Table 4. Each treatment consists of 5 *cohort*s, each of which has 8 subjects connected into $G_l$ or $G_h$ network (depending on the treatment). Each cohort is a statistically independent observation[3]. The total 20 cohorts are labeled from *N1*, *N2…N20*.

[3] To see, notice we use a between-subjects design, and subjects across different cohorts do not interact (The afore-mentioned mechanism of randomizing subjects over network positions is implemented *within* each cohort). Therefore, the independence of data across different cohorts is maintained.

Table 4. Summary of experimental design

| Treatments | | Network Structure | |
|---|---|---|---|
| | | $G_l$ | $G_h$ |
| Game | *Herding* | *Gl_Herd* | *Gh_Herd* |
| | *Shirking* | *Gl_Shirk* | *Gh_Shirk* |

Note: between-subjects design

**REMARK 1.** *In all treatments of the experiment, the efficient outcomes (which maximize the total expected payoff of players) are that both degree types take action 1 with probability 1.* [4]

The experiment is constituted by a repetition of the herding / shirking game for 20 rounds. The degree of each subject is predetermined (either 2 or 3) and kept unchanged throughout the experiment. In each round we randomize each cohort of 8 subjects over the network positions given their degrees.[5] On the one hand, this ensures that the partnerships constantly change and every round resembles a one-shot game. On the other hand, repeating the play will allow subjects to learn about the (one-shot) game, while fixing one's degree facilitates more effective learning and convergence of behavior. [6] Meanwhile, it seems implausible that subjects under incomplete information update their prior belief on network structures during the game, given the limited information provided to them.

In the herding (shirking) game described in Table 2 (3), we explicitly instructed subjects that their optimal choice is to choose 1 when there are more than or equal to (less than) 2 neighbors that execute the option 1. This eliminates the noise in the play attributed to not understanding the base game, so that the remaining variation in behavior is explained by the design factors (social network structure, type of game). After each round, players are informed of neighbors' decisions, while anonymity is maintained for all players. Sample instructions are found in the Appendix C.

In total, the data was collected from 160 subjects from March through April 2016, at a major public university in China. Subjects are primarily undergraduate students, recruited from an online

---

[4] The proof of Remark 1 is found in Appendix B.

[5] For illustration, denote the set of degree-3 nodes and that of degree-2 nodes in network $G_h$ by $H_{G_h}, L_{G_h}$ respectively. Then for each cohort of 8 subjects on $G_h$, there will be a half of people randomized equally likely over the nodes in $H_{G_h}$ and the other half randomized equally likely over the nodes in $L_{G_h}$, in each round of the game.

[6] In our game, if the subject's own type changes, the type distribution of her opponents will also change (see Table 1). Therefore, fixing one's degree substantially reduces one's mental burden in the game play. We thank an anonymous editor in suggesting this design feature to us.

information system. Cash is the only motivation for experimental participation. Subjects are paid proportional to their performance in the game. The software is programmed in zTree (Fischbacher 2007). Sample software screenshots are attached in Appendix D.

### 3.3. Hypothesis

The experimental game is of incomplete information, with player's degree as her private type and neighbor degree distribution as distribution of partner types. For such a Bayesian game, we focus on symmetric strategy $\sigma$, mapping from player's degree to her probability distribution over actions: $\sigma: k \rightarrow (p_k, 1-p_k)$, where $p_k$ is the probability that the degree-$k$ player assigns to action 1 (hence $1-p_k$ the probability of taking action 0). A strategy is *non-decreasing (non-increasing)* if $p_k$ weakly increases (weakly decreases) with $k$ ($p_k \leq (\geq) p_{k+1} \forall k$). The solution concept to our game is *Bayesian Nash equilibrium* (Gibbons 1992, Osborne and Rubinstein 1994, etc.), which we simply refer as *equilibrium*. Attributed to Galeotti et al. (2010), the background theory for our game in Appendix A yields a series of predictions under full rationality, from which we draw the hypotheses for our experiment. Specifically, our hypotheses inspect the following aspects of subject behavior: 1) *pointwise prediction*: Does the actual play of the game converge to its equilibrium? 2) *directional prediction*: Does the actual play of the game change with the social network structure in line with the predicted network effects?

**HYPOTHESIS 1. *Equilibrium play.*** *In each treatment, the actual pattern of play is consistent with one of the theoretical equilibria prescribed in Table 5 below.*

Table 5. Equilibrium structure[7]

| Treatment | *equilibrium* $(p_2, p_3)$ |
|---|---|
| *Gh_Herd* | (0, 0), (0.2, 1), (1, 1) |
| *Gl_Herd* | (0, 0), (0.68, 1), (1, 1) |
| *Gh_Shirk* | (1, 0.094) |
| *Gl_Shirk* | (1, 0) |

[7] The derivation of these equilibria is found in Appendix B. It can be shown in general that the Bayesian Nash equilibrium is unique in the shirking game defined in this paper, but not so in the herding game (as observed in Table 5). We will not explore this aspect further, and refer any interested reader to Galeotti et al (2010).

**HYPOTHESIS 2.** ***Network effects.*** **a. local network effect.** *In any given network, players with more neighbors choose 1 with higher (lower) probability in the herding (shirking) game.* **b. global network effect.** *For both herding and shirking games, players of any given degree type choose 1 with higher probability in denser networks.*

Both our hypotheses are made according to the theoretical predictions (Proposition A-1 and A-2 in Appendix A). In Hypothesis 2, we refer to *local network effect* as the pattern that one's action changes with her degree, and *global network effect* as the manner that the overall network density influences the choice of individuals. As implied by Hypothesis 2, we disentangle the two sources of network effects in our laboratory experiment: (Other things equal) 1) We compare the game plays of different degree types within the same network, so as to separate out the effect of local network while fixing the global network. 2) We analyze the data from the same degree type but in different networks, so as to identify the effect of global network while controlling that of local network.

## 4. Aggregate-level Results

In this section, we report the basic findings from our data. Define *1-rate* as the percentage of subjects who chose action 1 in a given setting (e.g. a treatment). Due to Remark 1 in Section 3, we can conveniently measure the system performance by 1-rate.[8] The notations in Table 6 below as well as those in previous sections apply to our dataset. In order to support the dataset decomposed in multiple ways, we flexibly combine the notations to designate any particular subset of data: For instance, *d3_Gl_Shirk* denotes the data generated by degree-3 players (*d3*) in low density network (*Gl*) for the shirking game (*Shirk*). The level of statistical significance is 5% unless specified otherwise.

[8] There are two primary reasons to use 1-rate as the indicator of system performance in our game: First, 1-rate is a simple measure as it is single-dimension. Second, 1-rate could be understood as the aggregated probability of individuals choosing 1, and thus has a link to the player strategy on the individual level.

Table 6. Notations for our dataset

| Notation | Interpretation |
|---|---|
| *Gh,Gl* | Network $G_h, G_l$ |
| *d3, d2* | degree-3, -2 |
| *Herd,Shirk* | herding, shirking |
| *round* | round of the game, taking values of 1,2… 20. |

**OBSERVATION 1.** *In none of the cases the exact pattern of play is consistent with the equilibrium prediction.*

Table 7. $\chi^2$ tests on equilibria as null hypotheses

| Treatment | $(p_2, p_3)$** | |
|---|---|---|
| | *observed 1-rate* | *equilibrium 1-rate* |
| *Gh_Herd* | (0.905, 0.92) | (0, 0)*<br>(0.2, 1)*<br>(1, 1)* |
| *Gl_Herd* | (0.607, 0.74) | (0, 0)*<br>(0.68, 1)*<br>(1, 1)* |
| *Gh_Shirk* | (0.748, 0.353) | (1, 0.094)* |
| *Gl_Shirk* | (0.667, 0.355) | (1, 0)* |

* p-value<0.05, two-tailed.
** Recall that $p_k$ denotes the probability that degree-$k$ type chooses 1, or empirically, the 1-rate for degree-$k$ players.

We conduct $\chi^2$ goodness of fit tests on whether the observed frequencies of choosing 1 by degree types in each treatment are consistent with the equilibrium strategies. While the details of the tests are found in Appendix E, all the equilibria in Table 5 taken as null hypotheses are turned down by the data at 5% significance level (see Table 7). Hypothesis 1 is therefore rejected. This outcome could be viewed as a confirmation of some account on behavioral game theory which indicated that normative equilibrium, especially in mixed strategy, may not be the exact predictor of the game (e.g. Brown and Rosenthal 1990). That said however, the failure of *pointwise* equilibria does not entirely dismiss the value of normative theory. As we will see, the *directional* predictions regarding some of the network effects inspired by the theory remain to hold.

**OBSERVATION 2. LOCAL NETWORK EFFECT** *The frequency of implementing socially beneficial choices significantly increases (decreases) with player degree in the herding (shirking) games except for Gh_Herd, where different degree types play a uniformly high level of action. The observed local network effect is stronger in shirking games than in herding games.*

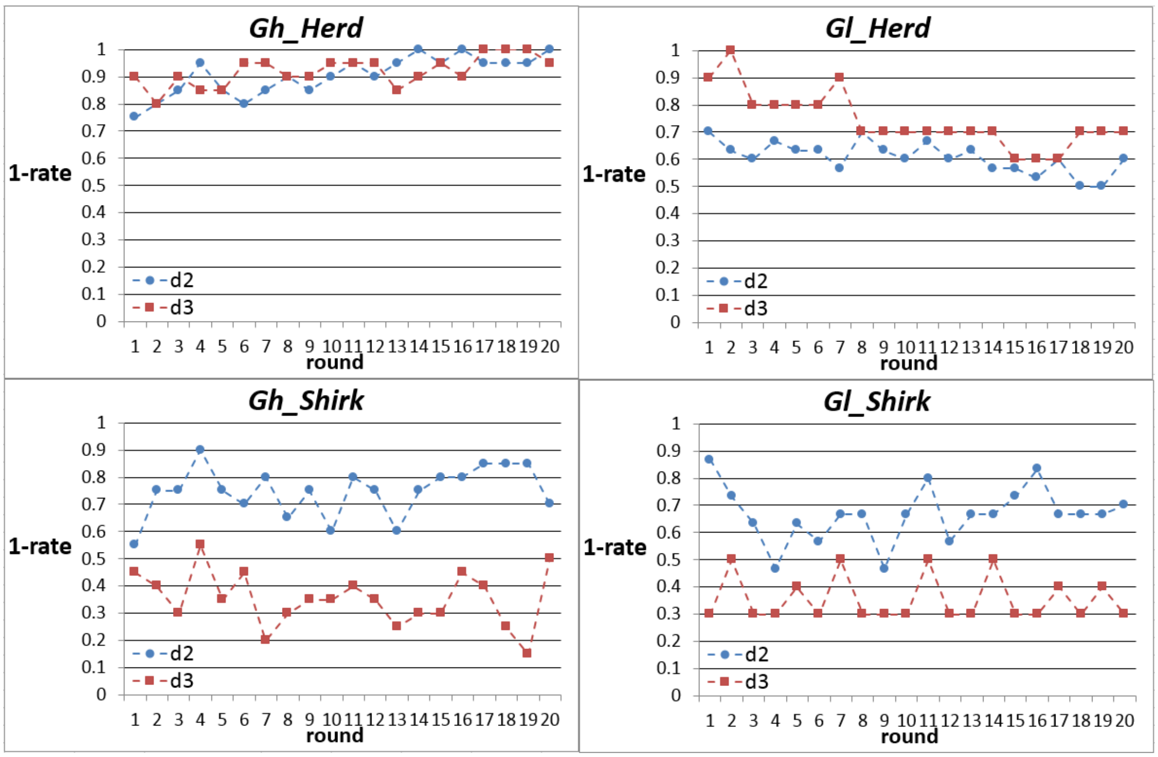


Figure 2. Local network effect

In all treatments except *Gh_Herd*, we observe a separation of actions by subjects' degree types. Furthermore, the player actions are separated in opposite directions in herding and shirking games – a pattern consistent with the local network effects in theory. See Figure 2 for illustration. However, the play is less separated in herding games than that in shirking games. The average difference in 1-rate between two degree types is 13.3% for treatment *Gl_Herd*, against those being 39.5% [31.2%] for *Gh_Shirk* [*Gl_Shirk*]. In other words, the pattern of behavior is more uniform across different degree types in herding games than in shirking games. Moreover, the separation tends to vanish in *Gh_Herd*. The average difference in 1-rate between two degree types is 1.5% for treatment *Gh_Herd*, against a base 1-rate of 90.5% for low-degree players in that treatment. Binomial tests using individual level data on the equality of probability of choosing 1 by degree-2 and degree-3 subjects located in the same network yield a $p$=0.155 for *Gh_Herd*, and 0.000 for the other treatments.[9]

It is also worth a note on the linkage between the observed local network effect in our study and the findings in Judd et al. (2010), which suggest that some aspect of network structure, such as long distance connectivity, exerts opposite influences on behavior in consensus and coloring (which correspond to herding and shirking in our paper). Different from the approach of Judd et al. (2010), we concentrate on the roles of one's *degree* in altering her choices in network games.

**OBSERVATION 3. GLOBAL NETWORK EFFECT**. *In all cases except for d3_Shirk, players implement the socially beneficial choice with significantly higher frequency in higher density networks. The observed global network effect is stronger in herding games than in shirking games.*

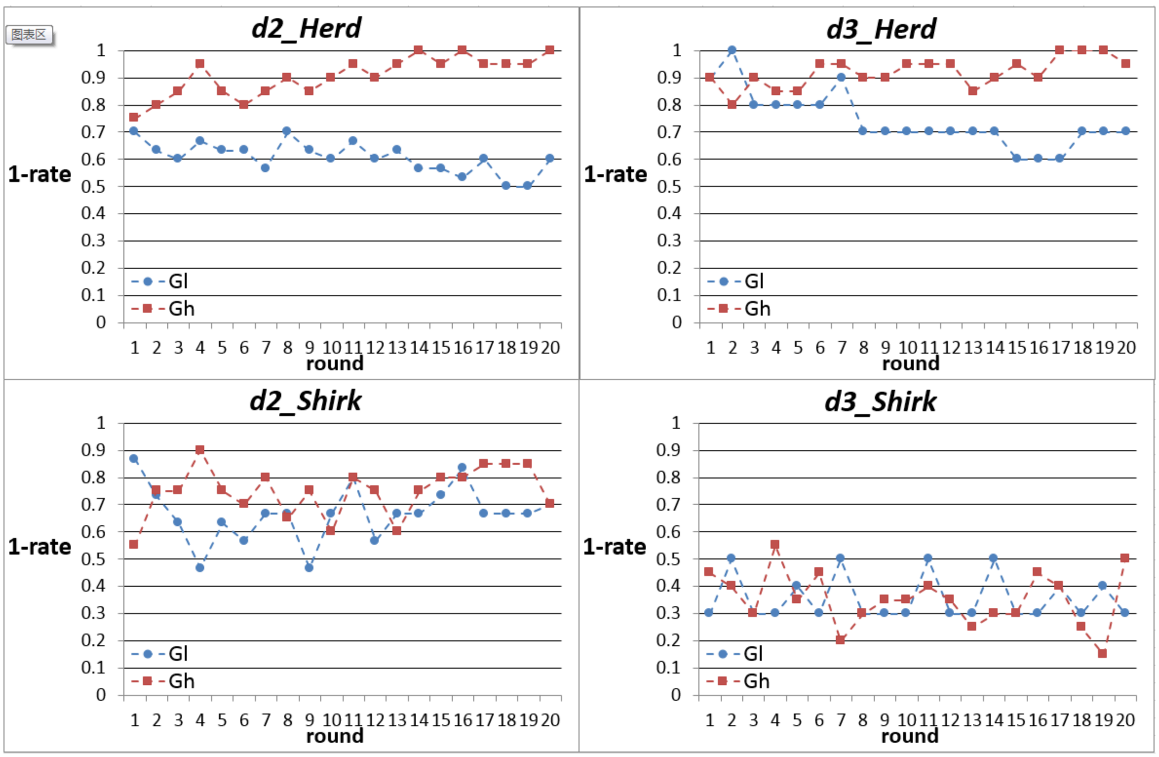


Figure 3. Global network effect

The global network effect in line with the theory is observed in herding treatments, where subjects in high density networks excel those under low network density in their frequency of

[9] The Binomial tests we used for Observation 2 [3] are one-sided. The null hypotheses of the tests are that, other things equal, subjects in different degree types [under different levels of network density] choose 1 with equal probability; and the local [global] network effect in Hypothesis 2a [2b] are taken as the alternative hypotheses.

choosing the socially beneficial option. The average difference in 1-rates by network density is 0.298 for *d2_Herd* and 0.18 for *d3_Herd*. In shirking games however, less evidence is found to support the global network effect that exists in principle. The average difference in 1-rates by network density is 0.081 [0.002] for *d2_Shirk* [*d3_Shirk*]. Binomial tests using individual level data on the equality of probability of choosing 1 by players located in networks with different density but having the same degree type produce $p$ =0.000 for both *d2_Herd* and *d3_Herd*, and 0.000 [0.562] for *d2_Shirk* [*d3_Shirk*].

**OBSERVATION 4. DYNAMICS OF BEHAVIOR** *In herding games, the direction of coordination is shifted by network density. The levels of action of both degree types increase [decrease] over time, under higher [lower] network density. In shirking games however, the behavior of both types exhibits an oscillating pattern over time, regardless of network density.*

At treatment level, Figure 2 presents the evolution of subject behavior over time. In herding games, the 1-rates of both high and low degree types move upward [downward] in high [low] density networks as the game proceeds. At the aggregate level, the increase of connectivity boosts the level of successful coordination in herding. In shirking games, the behavior is stable on average, but exhibiting significant variation. Because the incentive in shirking is to differentiate actions with each other, the observed behavior might be driven by "avoiding" the choices of opponents. This point is made more apparent with the individual level analysis in Section 5, which suggests adaptation to neighbor activities as a primary module of subject behavior. Finally, note that the current observation on behavioral dynamics is drawn on the aggregate level, and that an analysis of individual level behavior will follow in Section 5.

To summarize the results, our experiment finds that both herding and shirking in social networks are leveraged by the underlying network structure. Specifically, subject choices in both games are affected by their local network (degree), yet in opposite directions, and this local network effect stronger in shirking games than in herding games. On the contrast, the global network effect, which predicts higher 1-rates in denser networks for both degree types, is more visible in herding games than in shirking games. The subject choices do not converge to pointwise equilibria in either herding or shirking games. These findings partially support the set of experimental hypotheses, which are inspired by the formal theory. In order to explain the places where the standard theory fails, we develop a behavioral model in the next section, which generalizes the concept of experience-weighted attraction (EWA) learning under incomplete information into a network environment.

## 5. Behavioral Model & Individual-level Results

In testing the foregoing hypotheses on equilibrium and network effects, we find systematic deviations from rational play on the aggregate level. In this section, we shall investigate behavior at the individual level, which serves as a "root cause" of aggregate anomalies. More specifically, we shall characterize the kind of bounded rationality that underlies individual decision making in both herding and shirking through a unified, learning-based model. Inspired by Chong et al. (2006), we capture subject behavior by a *network-embedded* SEWA (Self-tuning Experience-Weighted Attraction) model under incomplete information, which nests two types of behavior as special cases: 1) *adaptive learning*, meaning that subjects evaluate the attractiveness of options upon past *realized* neighbor actions, and 2) *equilibrium learning*, in which subjects act upon the *expectation* of neighbor actions (and eventually reach the equilibrium).[10] Subjects update their beliefs (in a Bayesian manner) throughout the game regarding the extent to which their opponents learn towards equilibrium, and that to which they learn adaptively. We also relate the results of the learning models to the aggregate-level observations discussed in the previous section.

In this section, we construct our behavioral model based on Chong et al. (2006), which develops the standard SEWA framework (Ho et al. 2007) into a context of incomplete information. We further integrate the model into a network context. For purposes of exposition and comparison, we will introduce models of adaptive learning and equilibrium learning as the building blocks to the general model that follows. The fit of our behavioral model to the data will be examined in the subsequent section.

### 5.1. Adaptive learning

Consider bounded rational players who optimize noisily. The probability perceived by player $i$ at round $t+1$ that a generic degree-$k$ player chooses Action 1 is given by the logit function below.

$$p_{i,t+1,k}^{a} = \frac{e^{\lambda_i \; A_{itk}^{a}(1)}}{e^{\lambda_i \; A_{itk}^{a}(1)} + e^{\lambda_i \; A_{itk}^{a}(0)}}, \tag{1}$$

where $A_{itk}^{a}(x)$ is the *attraction* of action $x \in \{0,1\}$ to a generic degree-$k$ player, as perceived by player $i$ at round $t$ under adaptive learning.[11] As in (1), options with higher attraction shall be

---

[10] The mix of learning and equilibrium is essential in the construction of the SEWA model in Chong et al (2006). We extend this framework into a network setting, which allows us to study hybrid learning behavior in networks.

[11] We focus on beliefs consistent with one's own action. That is, if player $i$ happens to have a degree of $k$, she herself will choose 1 with probability $p_{itk}^{a}$ at round $t$ under adaptive learning.

picked up with higher probability. In addition, larger the value of $\lambda_i$ , more concentrated is the choice distribution around the exact option that is most attractive. In that way, the parameter $\lambda_i$ captures the degree of rationality of player $i$ in decision making.

Determining the attraction of options is a central issue to the family of experience-weighted attraction (EWA) learning, which we will illustrate below. Let $x_{it}, k_{it}, N_{it}$ respectively be the action, degree, and the set of neighbors of player $i$ at round $t$. $y_{it} \coloneqq \Sigma_{j \in N_{it}} x_{jt}$ represents the number of $i$'s neighbors choosing 1 at round $t$. Then the attraction $A_{itk}^{a}(x)$ under adaptive learning is determined as follows.

$$A_{itk}^{a}(x) = \left(1 - \frac{1}{n_{it}}\right) A_{i,t-1,k}^{a}(x) + \frac{1}{n_{it}} \Delta_{it}^{a}(x) v_i(x, y_{it}), \quad (2)$$

where $v_i(x, y)$ is the payoff of player $i$ if she had chosen $x \in \{0,1\}$ in a certain round, given $y$ as the number of neighbors who played 1 in that round; $\Delta_{it}^{a}(x) = 1$ if *either* of the following two conditions is met: 1) $x$ was chosen by player $i$ in round $t$, or 2) $x$ was not chosen but player $i$ could be better off choosing it in round $t$. If neither condition was satisfied, $\Delta_{it}^{a}(x) = 0$. The parameter $n_{it}$ ($n_{it} \geq 1$) controls the relative weight that the player assigns to the latest experience against that assigned to the historical attraction in the updating procedure. For brevity, we leave the update of $n_{it}$ in Appendix E. To initialize the learning process, notice that $p_{i1k}^{a} = \frac{1}{1+e^{-\lambda_i \ \Delta A_{i0k}^{a}}}$ , where $\Delta A_{i0k}^{a} \coloneqq A_{i0k}^{a}(1) - A_{i0k}^{a}(0)$ is left as a variable for estimation.

The adaptive learning has its root in the classical reinforcement dynamic based on the past history of game. At the same time, by assigning reinforcement to better forgone options, the adaptive learning also entails a local best responding dynamic. Noticeably, the adaptive learning uses only information *within one's neighborhood*, as adaptive learners only adjust their strategies against immediate neighbors.

### 5.2. Equilibrium learning

Under the framework of equilibrium learning, the player iterates the quantal response to neighbor strategy over rounds until it reinforces the belief. Similar to adaptive learning, the player choice involves random noise, which weakens the rationality required for the standard game-theoretical model. Specifically, the probability perceived by player $i$ at round $t+1$ that a generic degree-$k$ player chooses Action 1 in equilibrium learning is given by

$$p_{i,t+1,k}^{q} = \frac{e^{\lambda_i \ A_{itk}^{q}(1)}}{e^{\lambda_i \ A_{itk}^{q}(1)} + e^{\lambda_i \ A_{itk}^{q}(0)}}. \quad (3)$$

where $\lambda_i$ captures the degree of rationality in discerning payoff differences across multiple options. In contrast to adaptive learning, players under equilibrium learning base their decisions on the *expected* payoff, which is determined from the network density joint with the opponent's strategy. As such, both one's local network and the global network have influences on the equilibrium learning. The specific form of the attraction of option $x \in \{0,1\}$ to a generic degree-$k$ player as perceived by player $i$ at round $t$ is given by

$$A_{itk}^{q}(x) = U(x, \sigma_{it}, k, G_i)\,, \tag{4}$$

in which $G_i = 1(0)$ if the network that player $i$ resides in has a high (low) density[12], and $U(x, \sigma_{it}, k, G_i)$ is the expected payoff of a degree-$k$ individual who plays $x$ under network density $G_i$ given perceived neighbors strategy $\sigma_{it}: k \to \left(p_{itk}^{q}, 1 - p_{itk}^{q}\right)$. The expression of $U(x, \sigma, k, G)$ is provided as (A-5) in Appendix A. Denote $\Delta U(\sigma_{it}, k, G_i) \coloneqq U(1, \sigma_{it}, k, G_i) - U(0, \sigma_{it}, k, G_i)$. To initialize the learning, let $\Delta A_{i0k}^{q} \coloneqq A_{i0k}^{q}(1) - A_{i0k}^{q}(0)$, whose value will be left for empirical estimation. To close the loop, we assume that the strategy implemented by player $i$ is consistent with her belief. That is, player $i$ herself in round $t$ will execute Action 1 with probability $p_{itk_i}^{q}$. If the beliefs converge, this process will ultimately lead to a logit quantal response equilibrium as defined in McKelvey and Palfrey (1995) – see Proposition 1 below, whose proof is found in Appendix E. As such, equilibrium learning elaborates how individual activities approach a prescribed equilibrium over time.

**PROPOSITION 1.** *The equilibrium learning converges to a quantal response equilibrium if* $\lambda_i \sum_k \left| \frac{\partial \Delta U(\sigma_{it}, k_i, G_i)}{\partial p_{itk}^{q}} \right| < 4, \forall i, t.$

### 5.3. Hybrid learning

Equilibrium models are standard arguments of game theory, and have the advantage of self-reinforcement. Learning models differentiate from equilibrium models in that they allow players to accustom themselves to the game by a sequence of suboptimal moves. The hybrid learning model in this section combines the features of equilibrium and learning, so as to capture the behavior in networks that falls between the two paradigms. More specifically, the probability perceived by player $i$ at round $t$ that a generic degree-$k$ player chooses Action 1 under hybrid learning is

[12] Recall that in our experiment, each subject plays in a fixed cohort through all rounds. Therefore, $G_i$ does not vary with $t$.

$$p_{itk}^{h} = r_{itk} p_{itk}^{q} + (1 - r_{itk}) p_{itk}^{a}, \tag{5}$$

where $r_{itk}$ denotes the probability that the generic degree-$k$ player follows the equilibrium learning ($1 - r_{itk}$ the probability that she adopts adaptive learning), as perceived by player $i$ in round $t$.[13] [14] Given her own degree $k_i$, player $i$ updates her belief on $r_{i,t+1,k_i}$ in a Bayesian manner (shown as (6) below). As one's degree is fixed in the experiment (a design feature discussed before), she does not have experience in playing a strategy upon another degree; thus we assume the player's belief on $r_{itk}$ for $k \neq k_i$ sticks to its initial values $r_{i0k}$.

$$r_{i,t+1,k_i} = \begin{cases} \dfrac{p_{i,t+1,k_i}^{q} r_{itk_i}}{p_{i,t+1,k_i}^{q} r_{itk_i} + p_{i,t+1,k_i}^{a}\left(1 - r_{itk_i}\right)}, & x_{it} = 1 \\ \dfrac{\left(1 - p_{i,t+1,k_i}^{q}\right) r_{itk_i}}{\left(1 - p_{i,t+1,k_i}^{q}\right) r_{itk_i} + \left(1 - p_{i,t+1,k_i}^{a}\right)\left(1 - r_{itk_i}\right)}, & x_{it} = 0 \end{cases} \tag{6}$$

The model of hybrid learning entails the foregoing models as special cases: If $r_{itk} = 0$, it reduces to adaptive learning. If $r_{itk} = 1$, the model degenerates into equilibrium learning, which may converge to quantal response equilibrium per Proposition 1 (further to Bayesian Nash equilibrium if $\lambda_i \to \infty$ [15]). Next, we shall evaluate the empirical fit of the proposed learning models, and develop further ties between the results of learning analysis and the aggregate-level findings in Section 4.

### 5.4. Individual behavior

In this section we fit the learning models with our data. In particular, we will identify the most plausible model for the individual-level behavior based on the empirical fit. For that purpose, we will estimate the parameters that characterize the learning models, i.e. $\lambda_i, \Delta A_{i0k}^{a}$ for adaptive learning, $\lambda_i, \Delta A_{i0k}^{q}$ for equilibrium learning, and $\lambda_i, r_{i0k}, \Delta A_{i0k}^{a}, \Delta A_{i0k}^{q}$ for hybrid learning. To maintain the frugality of the model, we restrict to the case where different degree types do not *a priori* differ in the preference over options, i.e. $\Delta A_{i02}^{x} = \Delta A_{i03}^{x}, x \in \{a, q\}$. We consider the "cohort-specific" version of each model with each parameter above taking the same value across

---

[13] Alternatively, $r_{itk}$ could be interpreted as the proportion of equilibrium learners in the population of degree-$k$ players (the rest of degree-$k$ population being adaptive learners), as perceived by player $i$ at round $t$.

[14] We focus on beliefs reinforced by one's own action. That means, player $i$ at round $t$ will choose 1 with probability $p_{itk_i}^{h}$.

[15] In McKelvey and Palfrey (1995), the definition of QRE bears some relation with Bayesian Nash equilibrium. However, the two equilibrium notions are not structurally equivalent in our paper, as the payoff function of players cannot be rearranged to separate out a private disturbance.

individuals within the same cohort, which improves both parsimony and consistency (in terms of agent beliefs) [16] as elaborated in Appendix E.3. The estimation approach is illustrated in detail in Appendix E.4.

Table 8. Comparison of behavioral models†

| cohort-specific model | LL* | $\chi^2$ (d.f.)** | $p$-value** |
|---|---|---|---|
| hybrid learning | -466.46 | - | - |
| adaptive learning | -565.28 | 197.64(60) | 0.000 |
| equilibrium learning | -881.21 | 829.50 (60) | 0.000 |

† based on the entire dataset that includes all treatments
* log likelihood (LL) produced by maximum likelihood estimation (MLE).
** null [alternative] hypothesis of likelihood ratio test: The reduced model (adaptive learning or equilibrium learning) is more [less] efficient than the full model (hybrid learning) in fitting the data.

As suggested by Table 8, the hybrid model outperforms its special cases in fitting the data, with both P-values of 0.000 from likelihood ratio tests against equilibrium learning and adaptive learning (thereby rejecting the null hypotheses which favor the respective special cases). As a result, we conclude that the hybrid model is the most appropriate one in explaining our observations.

Table 9. Results of hybrid learning*

| treatment | $r_{itk}$ averaged over all periods** | | $r_{itk}$ averaged over the last five periods** | |
|---|---|---|---|---|
| | $\overline{r_3}$ | $\overline{r_2}$ | $\overline{r_3}$ | $\overline{r_2}$ |
| *Gh_Shirk* | 0.17<br>(0.108) | 0.39<br>(0.175) | 0.19<br>(0.108) | 0.40<br>(0.206) |
| *Gl_Shirk* | 0.35<br>(0.136) | 0.45<br>(0.127) | 0.36<br>(0.134) | 0.49<br>(0.150) |
| *Gh_Herd* | 0.25<br>(0.173) | 0.27<br>(0.215) | 0.26<br>(0.166) | 0.25<br>(0.206) |
| *Gl_Herd* | 0.28<br>(0.414) | 0.31<br>(0.269) | 0.29<br>(0.413) | 0.30<br>(0.292) |

* cohort-specific model, estimated by MLE
** calculated based on the estimated $r_{i0k}$, $\Delta A^{a}_{i0k}$, $\Delta A^{q}_{i0k}$, and $\lambda_i$
The statistics in the table are mean and standard deviation (the latter shown in parentheses).

[16] For the hybrid model, recall that $r_{itk}$ is defined upon the perception of the player in question regarding her neighbors' types of learning. Therefore, a model that imposes the same $r_{itk}$ within the same cohort ensures that the perceptions of players in the same cohort are consistent. We do not need to seek such consistency *across* cohorts though, since cohorts are independent with each other by the experimental design.

The parameter estimation results for hybrid learning can be found in Table E-3, Appendix E. Here we shall center our discussion on $r_{itk}$ [17]. In order to get a sense of stabilized beliefs, we calculate the time-averages of $r_{itk}$ (denoted by $\overline{r_k}$) over the last quarter (Period 16-20) of the experiment, and those over the whole experiment, for each degree type $k = 2{,}3$ in each cohort. Table 9 provides the mean and standard deviation of $\overline{r_k}$ in each treatment (which consists of five cohorts, thus five data points for $\overline{r_k}$) under the estimated parameters ($r_{i0k}, \Delta A^a_{i0k}, \Delta A^q_{i0k}, \lambda_i$). From Table 9, one can see that there is no noteworthy difference between $\overline{r_k}$ over all experimental rounds and that over the last five rounds, indicating stable beliefs over time. Furthermore, the treatment-averages of $\overline{r_k}$ never exceed 50%, and in five out of eight cases are less than 1/3 (in either case with all-rounds-average or last-quarter-average). The values of $\overline{r_k}$ being low implies that the hybrid learning is primarily led by its adaptive component. This suggests that subjects in all treatments have primarily focused on their immediate neighborhoods, thus giving rise to the observed local network effect (Observation 2, Section 4). Furthermore, the local adaptation may drive the herding in different directions, under different network density: In the high [low] density networks, since players are more likely to have high [low] degree neighbors. Thus they tend to observe higher [lower] action played by their neighbors. As a result, the focal player is induced to raise [reduce] her action level over time, which displays an inclining [declining] level of action for both degree types (Observation 4, Section 4). In the extreme case, the difference in action vanishes between different degree types, and the local network effect is replaced by pooled coordination (as in the case of *Gh_Herd*, Observation 2, Section 4). In the shirking games, players under high [low] network density tend to observe lower [higher] neighbor actions; then local adaptation makes them raise [reduce] their own action. As such, agents in shirking games have a tendency to invert the previous play, which leads to the oscillating behavioral pattern (Observation 4, Section 4) and dampens the effect of global network on the behavior (Observation 3, Section 4). Finally, the equilibrium component of hybrid learning appears to be weak (evidenced by low $\overline{r_k}$ values), which explains why the play does not converge to the equilibrium (Observation 1, Section 4).

We have three additional remarks for Table 9 regarding the learning style of subjects. In shirking treatments: First, the estimated $\overline{r_k}$ is higher for lower degree types, implying that lower degree players execute more equilibrium learning – the kind of learning that is based on expectation. To understand why, note that fewer neighbors mean fewer sources of strategic variability for a

[17] $r_{itk}$ calibrates the nature of learning and is thus the focus of our parameter estimation. The rest of parameters, $\lambda_i$ and the initial values, are useful mainly in the sense of calculating stabilized beliefs (time-averages of $r_{itk}$ or $\overline{r_k}$). See Table 9 for illustration.

player to cope with. So plausibly, players with fewer connections are able to form more accurate expectations about neighbor actions and use these expectations to make decisions. Second, $\overline{r_k}$ is higher in lower density treatments. That implies, other things equal, that the degree of equilibrium learning is higher under lower network density. To understand this, notice that in low density networks one's neighbors are more likely to have low degree, and that these neighbors will make more predictable, expectation-based decisions (see the first point). That, in turn, would allow the player in question to form more accurate expectation in his or her own learning under lower network density. Third, the estimates of $\overline{r_k}$ for herding treatments display the same patterns as do they in shirking (i.e. decreasing in density, decreasing in degree), but less significantly so – The respective differences are much smaller in herding than in shirking. In other words, individuals demonstrate more uniform behavior in herding than in shirking. This uniformity of behavior may stem from the alignment of objectives and the existence of win-win improvement in herding games.

To sum up, we find that the individual level behavior in the network games is primarily driven by the adaptation to realized neighbor actions rather than the expectation of opponent activities (whereas the latter leads to equilibrium). Moreover in shirking games, the extent of equilibrium learning decreases in subject degree, and decreases in network density. In herding games on the contrast, the learning styles are more homogeneous across degree types and networks. That accounts for the aggregate-level observations in Section 4, including the patterns of local and global network effects, the lack of equilibrium play, and the heterogeneity and dynamics of subject behavior.

## 6. Conclusion & Discussion

We examine in the laboratory two types of collective activities in social networks, herding and shirking, respectively referring to the case where a player's incentive to take a certain action increases and decreases if more of her network neighbors choose the same action. The set of relationships indicating "whose action affects whose payoff" defines the social network. Our experiments shed light on how the network structure leverages the outcome of interactions and facilitates socially beneficial behavior. Our experiment decomposes and tests the local and global network effects inspired by the theory – Players with more social ties make more [fewer] socially beneficial choices in herding [shirking] games, and the rate of making socially beneficial choices increases with the density of network. We find that the subject choices in herding and shirking games are not consistent with theoretical equilibria; the local network effect is stronger in shirking games than in herding games. Global network effect, in contrast, is considerably present in herding but less so in shirking. On the individual level, we explain the behavioral rationale behind both herding and shirking by a unified learning-based model, which is embedded into the underlying network structure. We find through our learning model that subjects primarily adapt to their *local* neighbor activities rather than acting upon expectation (while the latter is essential in making the equilibrium). That leads to the observed dynamics, off-equilibrium behavior, as well as the local and global network effects.

In terms of implications for practice, our findings suggest the role of network structure that a manager or administrator should understand when dealing with strategic agents in social networks. For instance, the observed local network effect implies that those who are well connected are able to generate considerable externalities by their decisions and thereby exert substantial influence on other's decisions and social welfare. Thus for a firm which provides technologies or devices whose demand is subject to the network effect, it is important to promote the product to the appropriate "social hubs" in the network, in order to achieve the desired outcome of consumer herding. For state governments interested in improving the environment in the presence of externality, the geographic proximity of countries or regions might help predict the probable outcome of pollution reduction and identify the key contributors to the social surplus. In addition, the pattern of learning and adaptation displayed in the game suggests to the firm / social planner the importance of one's past experience in shaping her strategy. That is to say, for example, the access to the historical records of customers in similar network activities can be helpful in predicting the likelihood of their consumption in the network. As such, it would be an important avenue for future research to incorporate the above consideration into the marketing policy of firms that sell to strategic customers in network settings.

Social and economic networks in practice usually have large sizes. Will the same conclusion we draw upon laboratory networks hold in networks of much larger scale? Note when the network is large, the information that individuals have about the network structure is likely to be incomplete. In this case, it seems one could anticipate the same network effects as in Hypothesis 2 (which are independent of network size) for larger networks. If the information setting is altered (e.g. more network information becoming available), it will then be important to examine the robustness of the conclusions in this paper with varying network sizes.

The network effects explored in this paper are related to those studied in sociology: The *egocentric* analysis focuses on one's personal ties and its effect on the individual behavior, while *sociocentric* analysis examines how the performance of a group is affected by the interaction architecture within that group (Chung et al. 2005). Commonly used network measures for egocentric and sociocentric analyses include degree, closeness centrality and betweenness centrality (Marsden, 2002). In our paper, however, we characterize one's local network and global network by degree and network density, correspondingly, and investigate the network effects in a *strategic* context (under *incomplete information*). Despite these differences, it remains to be interesting to study the efficacy of other sociological measures of network effects in our settings of collective activities.

*Due to the page limit, Appendix will be separated from the paper and posted online. The notations in the Appendix inherit from those the main text, unless otherwise clarified.*

## Appendix A.
## The Background Theory

This section presents the theoretical results of our game that are relevant to the establishment of experimental hypotheses in Section 3, the main text. Unless otherwise noted, the theoretical work in this section is attributed to Galeotti et al. (2010).

### A.1. Network

Let $N_i$ be the set of player $i$'s neighbors and $k_i$ be $i$'s degree (number of neighbors), the latter being the private information of player $i$ (or her *type*). The set of feasible degree values in a size $N$ network is denoted by $\kappa := \{1,2, \ldots N-1\}$. $\underline{k}$ ($\overline{k}$) is the lowest (highest) degree value in a given network. Let $G(\mathbf{k}|k)$ be the probability that neighbor degrees are ($k$-dimensional vector) **k,** conditional on one's own degree being $k$. Describe *network density* $F$ as the collection of conditional neighbor degree distributions $G(\cdot|k)$ for every degree type $k$, i.e. $F \coloneqq \left\{[G(\mathbf{k}|k)]_{\mathbf{k}\in\kappa^k}\right\}_{k\in\kappa}$. Let $E_{G(\cdot|k)}[f] \coloneqq \Sigma_{\mathbf{k}\in\kappa^k} G(\mathbf{k}|k) f(\mathbf{k})$, where $f: \kappa^k \to \mathbb{R}$ is a non-decreasing mapping. For comparison consider another network $F'$, $F' \coloneqq \left\{[G'(\mathbf{k}|k)]_{\mathbf{k}\in\kappa^k}\right\}_{k\in\kappa}$. $F'$ is said to have *higher* density (or *denser*) than $F$, if for any non-decreasing $f$

$$E_{G'(\cdot|k)}[f] > E_{G(\cdot|k)}[f], \forall k. \qquad \text{(A-1)}$$

That is, if one network allows, for every degree type, higher neighbor degree distribution than does the other, then the former network is denser. That said, one can readily compare the density between the two networks used in our experiment (Figure A-1):

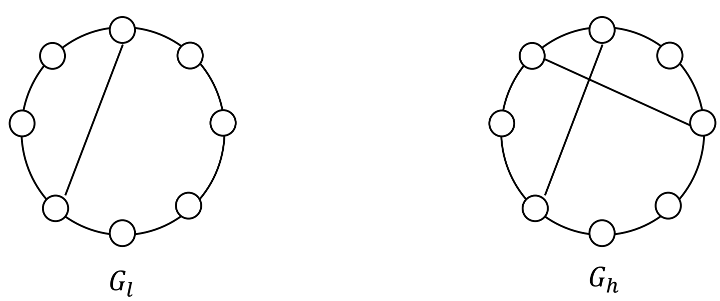


Figure A-1. Network structures used in experiments (reproduced from Figure 1)

**REMARK A-1**. *$G_h$ is denser than $G_l$.*

**Proof.** For the networks $G_h$ and $G_l$, neighbor degree distributions conditional on one's own degree, $G_h(\cdot\,|k)$ and $G_l(\cdot\,|k)$, are shown in Table A-1. Therefore we have

$$E_{G_h(\cdot|2)}[f_2] = \tfrac{1}{2}f_2(2,3) + \tfrac{1}{2}f_2(3,3) > \tfrac{2}{3}f_2(2,3) + \tfrac{1}{3}f_2(3,3) > \tfrac{2}{3}f_2(2,3) + \tfrac{1}{3}f_2(2,2) = E_{G_l(\cdot|2)}[f_2],$$

$$E_{G_h(\cdot|3)}[f_3] = \tfrac{1}{2}f_3(2,3,3) + \tfrac{1}{2}f_3(2,2,3) > f_3(2,2,3) = E_{G_l(\cdot|3)}[f_3],$$

for any $f_k$ ($k = 2,3$) that is non-decreasing in its argument. Applying the definition of network density (A-1) completes the proof. ■

Table A-1. The neighbor degree distributions for $G_h$ and $G_l$ (reproduced from Table 1)

| $G_h$ | | | $G_l$ | | |
|---|---|---|---|---|---|
| If the player is degree-2, | | | | | |
| probability | No. degree-3 neighbors | No. degree-2 neighbors | probability | No. degree-3 neighbors | No. degree-2 neighbors |
| 1/2 | 1 | 1 | 2/3 | 1 | 1 |
| 1/2 | 2 | 0 | 1/3 | 0 | 2 |
| If the player is degree-3, | | | | | |
| probability | No. degree-3 neighbors | No. degree-2 neighbors | probability | No. degree-3 neighbors | No. degree-2 neighbors |
| 1/2 | 2 | 1 | 1 | 1 | 2 |
| 1/2 | 1 | 2 | | | |

### A.2. Game

Let $x_i$ be player $i$'s (binary) decision, $x_i \in \{0,1\}$. Player $i$'s payoff, denoted by $v_i(x_i; \boldsymbol{x}_{N_i})$, is determined by her own decision $x_i$ and actions of her neighbors, $\boldsymbol{x}_{N_i} \coloneqq (x_j)_{j\in N_i}$. The player payoff exhibits a *herding* [*shirking*] nature, if for any $i$ and any $k_i$-dimension vector $\boldsymbol{x} \geq \boldsymbol{x}'$

$$v_i(1;\boldsymbol{x}) - v_i(0;\boldsymbol{x}) \geq [\leq] v_i(1;\boldsymbol{x}') - v_i(0;\boldsymbol{x}'). \tag{A-2}$$

That is, action 1 becomes more [less] appealing to the player if more of her neighbors take action 1. In other words, the player is always incentivized to follow [avoid] the choice of majority. The resultant game is called a *herding* [*shirking*] game. By (A-2), it is straightforward to verify the herding and shirking nature of games in Table 2 and 3 in the main text, respectively. The following properties are assumed.

**ASSUMPTION A-1**. $v_i(x_i; (\boldsymbol{x}, 0)) = v_i(x_i; \boldsymbol{x})\ \forall i.$ (A-3)

**ASSUMPTION A-2**. In the herding [shirking] game,

$$v_i(1;\mathbf{1}) - v_i(0;\mathbf{1}) > [<]0, v_i(1;\mathbf{0}) - v_i(0;\mathbf{0}) < [>]0 \ \forall i \in N \tag{A-4}$$

Assumption A-1 indicates that having a neighbor who takes action 0 is payoff-equivalent to not having that neighbor. As discussed in Galeotti, et al. (2010), Assumption A-1 is satisfied in many economic models including those where one's profit depends on the sum of neighbor actions (which is the case we focus on in the main text). Assumption A-2 sets the boundary conditions for the games to be well-behaving: In the extreme case where all neighbors follow action 1 (0), the player in question should optimally choose 1 (0) [0 (1)] in a herding [shirking] game. Assumptions A-1 and A-2 are both met in our experiment (Section 3, main text).

In our game, degree is the player's private type, neighbor degree distribution is the distribution of the player's partner types. For such a Bayesian game, we focus on symmetric player strategy $\sigma$, mapping from player's degree to her probability distribution over actions, $\sigma: k \rightarrow (p_k, 1 - p_k)$, where $p_k$ is the probability that the degree-$k$ player assigns to action 1 (hence $1 - p_k$ the probability of taking action 0). A strategy is *non-decreasing (non-increasing)* if $p_k$ weakly increases (weakly decreases) with $k$ ($p_k \leq (\geq) p_{k+1} \forall k$). The solution concept to our games is *Bayesian Nash equilibrium* (Gibbons 1992, Osborne and Rubinstein 1998, etc.), which we simply refer as *equilibrium*. Galeotti et al. (2010) has extensively studied the existence of monotonic equilibrium in herding and shirking games, and their results rely on some structural properties of the network game, which we will introduce below.

Let

$$U(x_i, \sigma, k_i, G) \coloneqq \int_{\boldsymbol{x}_{N_i}} v_i(x_i; \boldsymbol{x}_{N_i}) \mathrm{d}\phi(\boldsymbol{x}_{N_i}, \sigma, k_i, G) \tag{A-5}$$

be the expected payoff of degree type $k_i$ choosing $x_i$ in network $G$ when neighbors are playing the strategy $\sigma$, where $\phi(\cdot, \sigma, k_i, G)$ is the probability distribution over neighbor actions induced by the neighbor degree distribution and the neighbor strategy. The payoff exhibits *degree substitution* [*degree complementarity*] if for any non-increasing [non-decreasing] $\sigma$,

$$U(1, \sigma, k, G) - U(0, \sigma, k, G) \leq [\geq] U(1, \sigma, k', G) - U(0, \sigma, k', G), \forall k > k'. \tag{A-6}$$

Degree substitution (complementarity) implies that a higher degree type will have less (more) incentive to choose 1 under non-increasing (non-decreasing) strategy, which essentially extends the notion of shirking (herding) into an *incomplete information* scenario.

**PROPOSITION A-1 (Galeotti et al. 2010)**. *There exists a non-decreasing (non-increasing) equilibrium in the game that exhibits degree complementarity (degree substitution).*

The proof of Proposition A-1 can be found in Galeotti et al. (2010) (as Proposition 1). To understand the intuition behind the proposition: Under Assumption A-1, having one more neighbor who plays action 0 yields the same payoff to the player in question as not having this neighbor. Thus, a lower degree player is essentially dealing with the same number of neighbors as a higher degree player does, but with some of them playing action 0. Consequently, in a herding [shirking] game, the lower degree type would become more prone to action 0 [1], relative to a higher degree type. This gives rise to the monotonicity in the equilibrium. As shown in Remark A-2 below, our experimental games admit degree complementarity or degree substitution, and therefore possess the respective monotonic equilibria (which can be found in Table 5 in the main text).

**REMARK A-2**. *The games in treatments Gh_Herd and Gl_Herd (Gh_Shirk and Gl_Shirk) exhibit degree complementarity (degree substitution).*

**Proof**. For each treatment, define $\Delta U(\sigma, \text{treatment}) \coloneqq U(1,\sigma,3,G) - U(0,\sigma,3,G) - \big(U(1,\sigma,2,G) - U(0,\sigma,2,G)\big)$, where $G$ represents the network structure used in that treatment. In order to obtain degree complementarity [substitution] in respective treatments, we need to show the following:

$$\begin{cases} \Delta U(\sigma, Gh\_Herd) \geq 0 \\ p_2 \leq p_3 \\ p_2, p_3 \in [0,1] \end{cases}, \begin{cases} \Delta U(\sigma, Gl\_Herd) \geq 0 \\ p_2 \leq p_3 \\ p_2, p_3 \in [0,1] \end{cases}, \begin{cases} \Delta U(\sigma, Gh\_Shirk) \leq 0 \\ p_2 \geq p_3 \\ p_2, p_3 \in [0,1] \end{cases}, \begin{cases} \Delta U(\sigma, Gl\_Shirk) \leq 0 \\ p_2 \geq p_3 \\ p_2, p_3 \in [0,1] \end{cases}.$$

Because the payoff matrices for herding and shirking games in our experiment (Table 2 and 3) are exactly reversed from each other, it easily holds that $\Delta U(\sigma, Gh\_Herd) = -\Delta U(\sigma, Gh\_Shirk)$, and $\Delta U(\sigma, Gl\_Herd) = -\Delta U(\sigma, Gl\_Shirk)$. Hence, it suffices to show that

$$\begin{cases} \Delta U(\sigma, Gh\_Shirk) \leq 0 \\ \\ p_2, p_3 \in [0,1] \end{cases}, \begin{cases} \Delta U(\sigma, Gl\_Shirk) \leq 0 \\ \\ p_2, p_3 \in [0,1] \end{cases},$$

which can be easily verified to hold in our experimental setting. Take *Gh_Shirk* for example: Applying the expressions of $U(\cdot)$ elaborated in (B-1) and (B-2) in Appendix B, one can obtain

$$\Delta U(\sigma, Gh\_Shirk) = 25p_2\left(-1 - 6p_3 + 4p_3^2 + p_2(-2 + 4p_3)\right),$$

which is non-positive for any $p_2, p_3 \in [0,1]$. Similar derivations also suggest that $\Delta U(\sigma, Gl\_Shirk) \leq 0$ for $p_2, p_3 \in [0,1]$. ∎

Proposition A-1 establishes the equilibrium structure on the class of games we study. In order to seek for more insights, we shall next investigate the *comparative statics* of equilibrium with respect to the underlying network structure.

**PROPOSITION A-2**. *Consider the herding (shirking) game played in two networks $G, G'$. $G$ is denser than $G'$. Then compared to any non-decreasing (non-increasing) herding (shirking) equilibrium $\sigma'$ that exists for $G'$, there exists a non-decreasing (non-increasing) equilibrium $\sigma$ in $G$, where the probability of choosing 1 is weakly higher for each degree type.*

**Proof.** The shirking game part of the result is proved by Proposition 5 of Galeotti et al. (2010). For the herding game part, first observe that there exist a greatest and a least equilibria in the order of stochastic dominance (Theorem 14, Van Zandt and Vives 2007). Moreover, viewing neighbor degree distribution as the (joint) type distribution of opponents, one can readily apply Proposition 16 of Van Zandt and Vives (2007) to show that the greatest equilibrium weakly increases (in the sense of stochastic dominance) with the network density. In other words, for each non-decreasing herding equilibrium in the original network, one can find a greater equilibrium of the same monotonicity in a denser network. Then the desired result follows. ■

For Proposition A-2, it is easy to see its intuition in case of the herding game: By definition, higher network density means stochastically higher neighbor degrees (in the sense of stochastic dominance in neighbor degree distribution) for every degree type involved. If we conjecture that the level of action of each degree type increases with the network density, then the neighbor actions will *ex ante* increase under the maintained conjecture joint with the non-decreasing strategy used by neighbors. The player in question should thus increase the level of her own action, as a result of herding. This confirms our conjecture, implying the existence of the equilibrium in the denser network with higher action for every degree type. The theoretical arguments behind Proposition A-2 for the shirking game are trickier and involve a contradiction. First note in the binary action context, non-increasing strategy is characterized by a threshold $\tau$, where players have more [less] degree than $\tau$ play pure action 0 [1] and the degree type-$\tau$ randomizes. Start with the hypothesis that the equilibrium threshold is strictly lowered in the denser network. Since denser network indicates that any player's neighbor will *ex ante* have more neighbors, the neighbors of the player in question should lower their actions (given more neighbors they have, the maintained hypothesis, and the non-increasing-ness of their strategy). In response, the player in question should increase her action as a result of shirking, which contradicts with the maintained hypothesis. Therefore, we conclude the opposite of the hypothesis, which suggests weakly higher equilibrium threshold in denser network. In other words, every degree type weakly increases her action in a shirking equilibrium under higher network density.

We interpret the above theoretical results as the effects of network structure on agent behavior. To see, notice that Proposition A-1 specifies how equilibrium action changes with one's degree and Proposition A-2 is about how equilibrium as a whole is shifted by the underlying network topology. In the main text, Hypothesis 2a and 2b are respectively built on Proposition A-1 and A-2, but differ from the theory in that they are constructed upon general, not-necessarily-equilibrium states [18]. In that sense, Hypothesis 1 and 2 can be tested unconfoundedly, without any joint condition such as equilibrium.

## Appendix B.

**THE DERIVATION OF EQUILIBRIA FOR TABLE 5.** Table 5 in the main text contains equilibria of games in all treatments. This section illustrates how these equilibria are derived. In what follows, we calculate the equilibrium for treatment *Gh_Shirk* (c.f. Table 4). The equilibria of games of other treatments can be worked out similarly.

Represent the player payoff in the experimental shirking game (Table 3) by the following matrix, $V \coloneqq \begin{bmatrix} 100 & 175 & 225 & 260 \\ 0 & 100 & 275 & 335 \end{bmatrix}$, whose components are denoted by $V_{i,j}$ $(i = 1,2; j = 1,2,3,4)$. Applying the neighbor degree distributions shown in Table A-1, and the $U$-notations defined in (A-5), we have:

$$U(1,\sigma,2,G_h) - U(0,\sigma,2,G_h) = 1/2\left(p_3^2(V_{1,3} - V_{2,3}) + 2p_3(1-p_3)(V_{1,2} - V_{2,2}) + (1-p_3)^2(V_{1,1} - V_{2,1})\right) + 1/2\left(p_2p_3(V_{1,3} - V_{2,3}) + ((1-p_2)p_3 + p_2(1-p_3))(V_{1,2} - V_{2,2}) + (1-p_2)(1-p_3)(V_{1,1} - V_{2,1})\right), \tag{B-1}$$

$$U(1,\sigma,3,G_h) - U(0,\sigma,3,G_h) = 1/2\,(p_2(p_3^2(V_{1,4} - V_{2,4}) + 2p_3(1-p_3)(V_{1,3} - V_{2,3}) + (1-p_3)^2(V_{1,2} - V_{2,2})) + (1-p_2)(p_3^2(V_{1,3} - V_{2,3}) + 2p_3(1-p_3)(V_{1,2} - V_{2,2}) + (1-p_3)^2(V_{1,1} - V_{2,1}))) + 1/2\,(p_3(p_2^2(V_{1,4} - V_{2,4}) + 2p_2(1-p_2)(V_{1,3} - V_{2,3}) + (1-p_2)^2(V_{1,2} - V_{2,2})) + (1-p_3)(p_2^2(V_{1,3} - V_{2,3}) + 2p_2(1-p_2)(V_{1,2} - V_{2,2}) + (1-p_2)^2(V_{1,1} - V_{2,1}))). \tag{B-2}$$

Based on that, we now check for potential equilibria in different scenarios. For example, if there exists an equilibrium where the low degree type chooses 1 and the high type is indifferent between the choices, then the equilibrium must be the solution to the following system:

[18] To see that, notice that equilibrium is a prerequisite to Proposition A-1 and A-2, but it is not required in the statement of Hypothesis 2.

$$\begin{cases} U(1,\sigma,2,G_h) - U(0,\sigma,2,G_h) > 0 \\ U(1,\sigma,3,G_h) - U(0,\sigma,3,G_h) = 0 \, , \\ \quad\quad p_2 = 1, 0 < p_3 < 1 \end{cases} \quad \text{(B-3)}$$

Solving (B-3) gives us an equilibrium $p_2 = 1, p_3 = 0.094$. Furthermore, it can be checked that *Gh_Shirk* does not have any other equilibrium.

**PROOF OF REMARK 1**. We compute the efficient allocation, denoted by $(p_2^*, p_3^*)$ (which maximizes the *ex ante* social welfare), for treatment *Gh_Shirk*. The efficient allocation for other treatments can be analogously worked out. Let $U_{k,G_h}^{\sigma} = p_k U(1,\sigma,k,G_h) + (1-p_k)U(0,\sigma,k,G_h)$. Then $(p_2^*, p_3^*)$ is the solution to the following problem.

$$\max_{p_2,p_3} \frac{1}{2} U_{2,G_h}^{\sigma} + \frac{1}{2} U_{3,G_h}^{\sigma}, \, s.t. \, p_2 \in [0,1], p_3 \in [0,1]. \quad \text{(B-4)}$$

Then it is easy to find that $p_2^* = 1, p_3^* = 1$ for treatment *Gh_Shirk*. ■

## Appendix C.
## Experimental Instruction

We present the experimental instruction for treatment *Gh_Shirk* in this appendix. The instruction can be straightforwardly adapted to *Gl_Shirk* with only change of the neighbor degree distribution, and to the *herding* treatments by the change of payoff values and labels.

****************************** Instruction: *Gh_Shirk* ******************************

***General.*** Welcome and thank you for participating in this experiment. In this experiment you will earn money. From now on until the end of the experiment, please do not communicate with other participants. If you have any question, please raise your hand. An experimenter will come to your place and answer your question privately.

***The Game.*** In the Experiment we use ECU (Experimental Currency Unit) as the monetary unit. The profits you make during the experiment will be added to this account in ECU. At the end of the experiment, the balance of the account will be converted from ECUs into Chinese yuan according to the conversion rate stated below, and paid out in cash after the experiment.

The experiment lasts for 20 rounds. In each round, participants will be organized in a network. In this network, you are connected to either two or three people, who are your neighbors. Every round your neighbors will be different people, but you will have the same number of neighbors. You know the number of neighbors whom you are connected, but you will not know their identity. Specifically,

*If you have two neighbors,*

• With 1/2 chance, one of your neighbors has 2 neighbors (including yourself) while the other has 3 neighbors (including yourself).

• With 1/2 chance, each of your neighbors has 3 neighbors (including yourself).

*If you have three neighbors,*

• With 1/2 chance, one of your neighbors has 2 neighbors (including yourself) while the other two neighbors have 3 neighbors each (including yourself).

• With 1/2 chance, one of your neighbors has 3 neighbors (including yourself) while the other two neighbors have 2 neighbors each (including yourself).

At the beginning of each round, each person in the network chooses one of the two options: A, B. Each does so without any knowledge of what any other person decides. The profit you earn depends on how many neighbors you have, the option you choose, and how many of your neighbors choose A. Specifically,

*If you choose A and you have two neighbors then*

• If both neighbors choose A, your profit is 225 ECU.

• If exactly one neighbor chooses A your profit is 175 ECU.

• If neither neighbor chooses A your profit is 100 ECU.

*If you choose B and you have two neighbors then*

• If both neighbors choose A, your profit is 275 ECU.

• If exactly one neighbor chooses A your profit is 100 ECU.

• If neither neighbor chooses A your profit is 0 ECU.

*This is summarized in Table 1 as below, also shown on your computer screen during your play.*

| Your Profit | | Number of Your Neighbors who Choose A | | |
|---|---|---|---|---|
| | | 2 | 1 | 0 |
| Your Choice | A | 225 | 175 | 100 |
| | B | 275 | 100 | 0 |

Table 1. *Two-Neighbor Payoff Table*

*If you choose A and you have three neighbors then*

• If all three neighbors choose A your profit is 260 ECU.

• If exactly two neighbors choose A, your profit is 225 ECU.

• If exactly one neighbor chooses A your profit is 175 ECU.

• If no neighbor chooses A your profit is 100 ECU.

*If you choose B and you have three neighbors then*

• If all three neighbors choose A your profit is 335 ECU.

• If exactly two neighbors choose A, your profit is 275 ECU.

• If exactly one neighbor chooses A your profit is 100 ECU.

• If no neighbor chooses A your profit is 0 ECU.

*This is summarized in Table 2 as below, also shown on your computer screen during your play.*

| Your Profit | | Number of Your Neighbors who Choose A | | | |
|---|---|---|---|---|---|
| | | 3 | 2 | 1 | 0 |
| Your Choice | A | 260 | 225 | 175 | 100 |
| | B | 335 | 275 | 100 | 0 |

Table 2. *Three-Neighbor Payoff Table*

There are a few tips to keep in mind, which will help you earn more profits:

- I earn more profit by choosing B if two or more than two of my neighbors choose A.
- I earn more profit by choosing A if none or one of my neighbor chooses A.

The conversion rate is 1 ECU=0.01 Chinese Yuan. Before the game starts, you are required to complete a quiz, which covers the important knowledge about the game. The quiz will be shown on your computer screen. You will start to play the game only after you correctly answer all the questions in the quiz.

***Consent Forms***. Please read the consent form that is delivered to you before the start of the experiment.

**************************************************************************************

## Appendix D.

## The Experimental Software Interface

This section provides snapshots of the experimental software. The software is programmed with zTree (Fischbacher, 2007). Subjects began with a quiz testing their understanding of the game, with no earning accumulated to the game. The quiz screens are shown in Figure D-1. The actual decision making interfaces are found in Figure D-2. For conciseness, we only include the interfaces for *d2_Gl_Shirk* (which means the screens shown to degree-2 players in treatment *Gl_Shirk*). The other cases not covered in the screenshots were presented to the subjects in an analogous way.

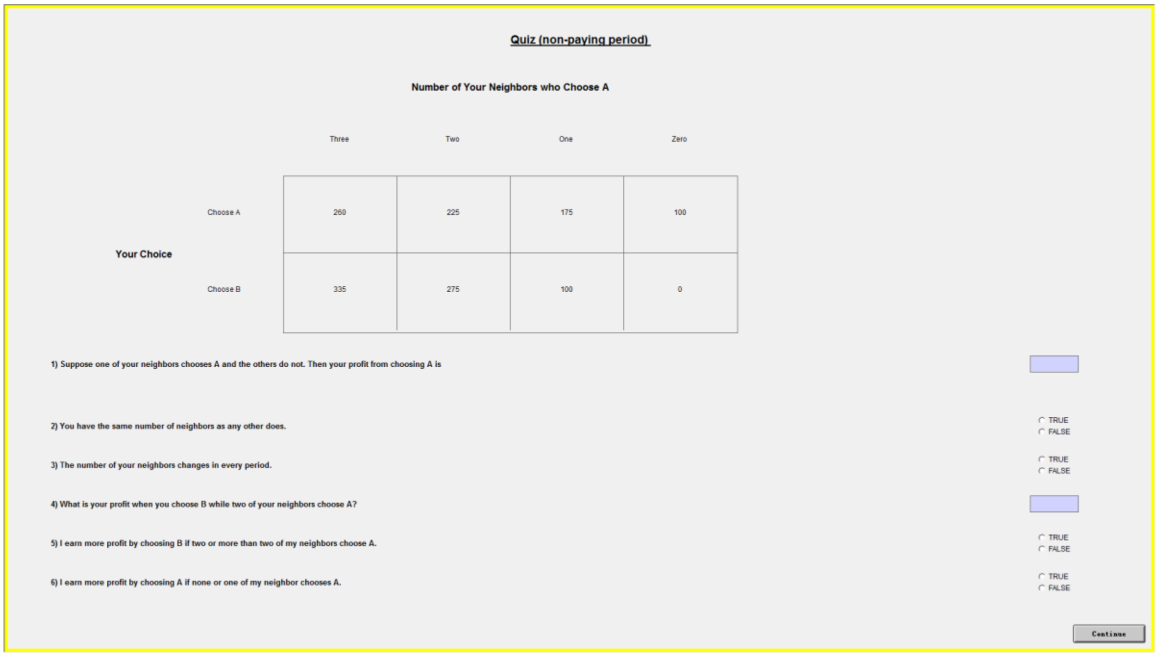


Figure D-1. Quiz

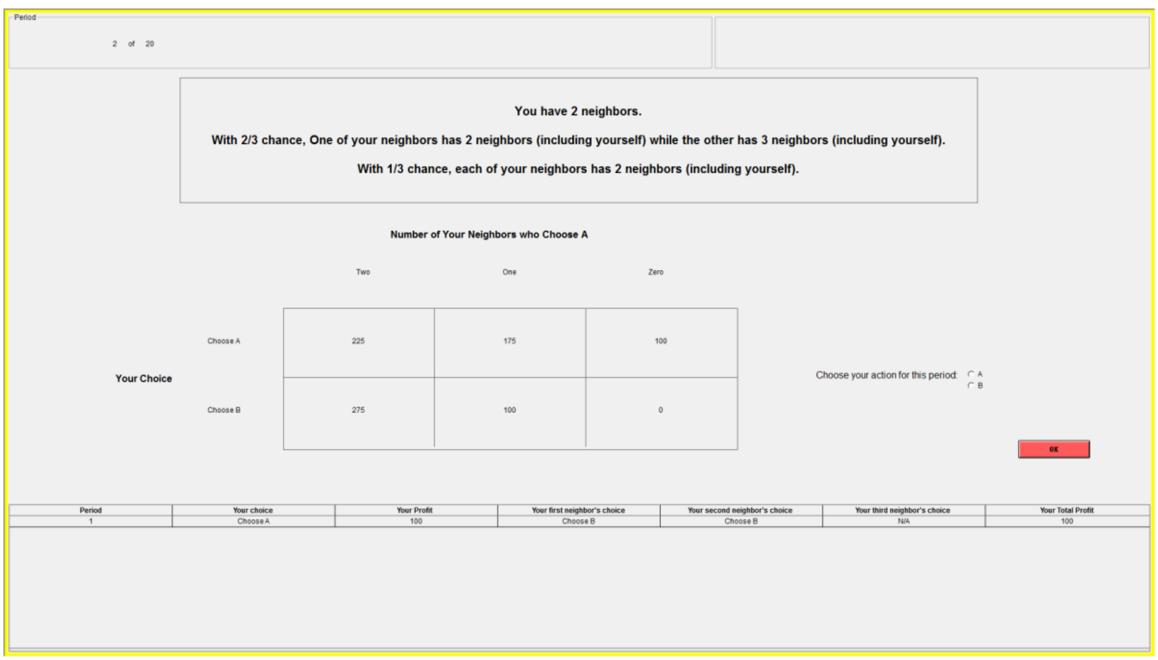


Figure D-2. Decision making

## Appendix E.

## Data Analysis

### E.1. Test of the equilibrium hypothesis

In Section 4, we have noted the ineffectiveness of formal equilibrium in representing the subject behavior (Observation 1). This section will elaborate in detail the $\chi^2$ goodness of fit test we used to reach the above conclusion. Denote by $I_{d2}$, $I_{d3}$ the numbers of degree-2 and degree-3 players who choose 1 in a cohort, respectively, and $N_{di}$ the number of degree-$i$ players in the cohort[19]. Then the *category* of the test consists of the dyad $(I_{d2}, I_{d3})$. The *expected* frequency at each category $(I_{d2}, I_{d3})$ is given by $\mathbb{N}\prod_{i=2,3}\binom{N_{di}}{I_{di}}p_i{}^{I_{di}}(1-p_i)^{N_{di}-I_{di}}$, under the equilibrium strategy $(p_2, p_3)$[20] and the sample size $\mathbb{N}$. The sample size for the test in each treatment is 5 cohorts $\times$ 20 periods = 100 (realized outcomes of the treatment game). That is, $\mathbb{N} = 100$. Then collecting the observed frequency at each category and comparing it with the expected one will yield the $\chi^2$ test statistics

$$\chi^2 = \sum_{\text{categories}} \frac{(\text{observed frequency} - \text{expected frequency})^2}{\text{expected frequency}}, \tag{E-1}$$

with the degree of freedom = $(N_{d2} - 1)(N_{d3} - 1)$. As such, the test is methodologically consistent with that used in Brown and Rosenthal (1990) for mixed strategy equilibria, except that the players in our game are identified by their degree types. For illustration purpose, Table E-1 details the test for treatment *Gh Shirk*, where the unique equilibrium is that $p_2 = 1, p_3 = 0.094$. The tests of other equilibria can be done in a similar fashion. As a result, all the equilibria are rejected with *p*-value of 0.000.

---

[19] In our experiment, $N_{di}$ differs by the network density. In high density network $(G_h)$, $N_{d2} = N_{d3} = 4$, while in low density network $(G_l)$, $N_{d2} = 6, N_{d3} = 2$ (Refer to Figure 1 in the paper).

[20] Recall that $p_k$ is the probability of degree-$k$ players choosing 1 at the equilibrium.

Table E-1. $\chi^2$ goodness of fit test on the *Gh Shirk* equilibrium

| Ho: $p_2 = 1$, $p_3 = 0.094$ | $I_{d3}$ | | | | |
|---|---|---|---|---|---|
| $I_{d2}$ | 0 | 1 | 2 | 3 | 4 |
| 0 | 1(0) | 0(0) | 0(0) | 0(0) | 0(0) |
| 1 | 0(0) | 2(0) | 0(0) | 0(0) | 2(0) |
| 2 | 3(0) | 5(0) | 11(0) | 2(0) | 1(0) |
| 3 | 5(0) | 17(0) | 7(0) | 6(0) | 6(0) |
| 4 | 8(67.38) | 1(27.96) | 14(4.35) | 4(0.30) | 5(0.01) |
| | | | $\chi^2(9) = \infty$ | 0.000* | |

The statistics in the table are the observed frequency and the expected frequency by equilibrium (the latter shown in parentheses)
sample size=100
* *p*-value

The exact testing of equilibria in our treatment games tends to be challenging, since every equilibrium involves at least one degree type choosing some action with probability 1 (c.f. Table 5, main text). That leaves some category for which the expected frequency is 0. If the observed frequency in that category is not 0 (such as the many cases in Table E-1), it will result in some infinitely large $\chi^2$ test statistic and reject the null hypothesis unambiguously. Practically speaking, the extreme predictions of formal equilibria are very unlikely to survive in reality; thus the rejection of the equilibrium hypothesis by our data gives no surprise.

In Section 5, we also attempt an individual-level explanation on why subjects fail to reach equilibrium – It happens because subjects do not very much follow the equilibrium learning. In fact, we depict their pattern of learning as a combination of equilibrium learning and adaptive learning, and the latter being a stronger component as suggested by the data. This learning-based approach accounts for not only the dismissal of formal equilibrium, but also the findings with regard to the network effects and dynamics on the aggregate level. See Section 5 for details.

**E.2. Learning: Model**

**Adaptive learning**

Section 5 presents the learning model we used to analyze individual behavior. The model resembles features of SEWA (Chong et al. 2006, Ho et al. 2007) model, and is extended by us into the network setting. While the main construct of the model is provided in Section 5, we are left with details regarding the functional parameter ($n_{it}$) that coevolves with the game. To proceed, let

$$n_{it} = \phi_{it} n_{i,t-1} + 1, \tag{E-2}$$

where the initial value $n_{i0} = 0$, and $\phi_{it}$ is computed as follows (in the new index):

$$\phi_{it} = 1 - \frac{1}{2}\left(\sum_x \left[\frac{\sum_{\sigma=t-1}^{t}\sum_{j\in N_{i\sigma}} I(x,x_{j\sigma})}{2k_i} - \frac{\sum_{\sigma=1}^{t}\sum_{j\in N_{i\sigma}} I(x,x_{j\sigma})}{tk_i}\right]^2\right), \qquad \text{(E-3)}$$

in which $I(x,x_{it}) = \begin{cases} 1, & x = x_{it} \\ 0, & x \neq x_{it} \end{cases}$.

The definition of $\phi_{it}$ involves comparing the rate of each action being chosen within one's neighborhood averaged over the most recent two periods, to that averaged over all periods until now. A lower value of $\phi_{it}$ hence indicates higher variability of the learning environment, which reduces the value of $n_{it}$ and means that the update of choice attraction (c.f. (2) in the main text) will bend towards the most recent experience.

**Equilibrium learning**

The rest of Appendix E.2 deals with equilibrium learning, which is introduced in Section 5.2 in the main text. Specifically, we shall prove the result that the equilibrium learning indeed leads to equilibrium (Proposition 1, Section 5.2).

**Proof of Proposition 1.** Define a vector-valued function $p_{it}^q \coloneqq \left(p_{itk}^q\right)_{k\in\kappa}$. Then the equilibrium learning implies the following quantal response iteration, $p_{i,t+1}^q \coloneqq p_{i,t+1}^q\left(p_{it}^q\right)$, where each component $p_{i,t+1,k}^q = \frac{e^{\lambda_i\ U(1,\sigma_{it},k,G_i)}}{e^{\lambda_i\ U(1,\sigma_{it},k,G_i)} + e^{\lambda_i\ U(0,\sigma_{it},k,G_i)}}, \forall k$ (c.f. (3) and (4) in Section 5). Then each entry of the Jacobian matrix $\frac{\partial p_{i,t+1}^q}{\partial p_{it}^q}$ takes the following form:

$$\frac{\partial p_{i,t+1,k}^q}{\partial p_{itk'}^q} = \frac{-e^{\lambda_i\ \left(U(0,\sigma_{it},k,G_i)-U(1,\sigma_{it},k,G_i)\right)}\lambda_i}{\left(1+e^{\lambda_i\ \left(U(0,\sigma_{it},k,G_i)-U(1,\sigma_{it},k,G_i)\right)}\right)^2}\frac{\partial\left(U(0,\sigma_{it},k,G_i)-U(1,\sigma_{it},k,G_i)\right)}{\partial p_{itk'}^q}, \forall k,k'. \qquad \text{(E-4)}$$

Applying the ∞-norm (denoted by ‖·‖), one obtains

$$\left\|\frac{\partial p_{i,t+1}^q}{\partial p_{it}^q}\right\|$$

$$= \max_k \left\{\sum_{k'}\left|\frac{\partial p_{i,t+1,k}^q}{\partial p_{itk'}^q}\right|\right\}$$

$$\underbrace{=}_{\text{(E-4)}} \max_k \left\{\sum_{k'}\left|\frac{-e^{\lambda_i\ \left(U(0,\sigma_{it},k,G_i)-U(1,\sigma_{it},k,G_i)\right)}\lambda_i}{\left(1+e^{\lambda_i\ \left(U(0,\sigma_{it},k,G_i)-U(1,\sigma_{it},k,G_i)\right)}\right)^2}\frac{\partial\left(U(0,\sigma_{it},k,G_i)-U(1,\sigma_{it},k,G_i)\right)}{\partial p_{itk'}^q}\right|\right\}$$

$$= \max_k \left\{ \frac{e^{\lambda_i \left(U(0,\sigma_{it},k,G_i)-U(1,\sigma_{it},k,G_i)\right)}}{\left(1+e^{\lambda_i \left(U(0,\sigma_{it},k,G_i)-U(1,\sigma_{it},k,G_i)\right)}\right)^2} \lambda_i \sum_{k'} \left| \frac{\partial\left(U(0,\sigma_{it},k,G_i)-U(1,\sigma_{it},k,G_i)\right)}{\partial p^q_{itk'}} \right| \right\}$$

In order for the quantal response to be a contraction mapping, we need to have the norm of the Jacobian sufficiently bounded. For that purpose, observe $0 < \frac{e^{\lambda_i \left(U(0,\sigma_{it},k,G_i)-U(1,\sigma_{it},k,G_i)\right)}}{\left(1+e^{\lambda_i \left(U(0,\sigma_{it},k,G_i)-U(1,\sigma_{it},k,G_i)\right)}\right)^2} \leq \frac{1}{4}$. Thus, we will have $\left\| \frac{\partial p^q_{i,t+1}}{\partial p^q_{it}} \right\| < 1$if $\lambda_i \sum_{k'} \left| \frac{\partial\left(U(0,\sigma_{it},k,G_i)-U(1,\sigma_{it},k,G_i)\right)}{\partial p^q_{itk'}} \right| < 4$. It then follows that $p^q_{it}(\cdot)$ constitutes a contraction mapping, of which the iteration over $t$ converges to a quantal response equilibrium $\left(p^q_{i,\cdot}\right)_{i=1,2\ldots N}$ by Banach fixed point theorem. ■

### E.3. Learning: Results

In Section 5, we have examined the performance of learning models with cohort-specific parameterization. While the cohort-specific version of the models imposes the desired consistency on agent beliefs (see footnote 15 in the main text), this section will demonstrate its efficiency in fitting the data against the individually parameterized model (referred as *individual-specific* model). As suggested in Table E-2, the cohort-specific model is favored by the likelihood ratio test over the individual-specific model in all cases, where the $p$-value of the test is 1.000 for hybrid learning, and 1.000, 0.944 for equilibrium learning and adaptive learning, correspondingly.[21] Thus, our further comparisons of the behavioral models in Section 5 are based on their cohort-specific form.

---

[21] Our estimations for cohort- (individual-) specific models are conducted independently across cohorts (subjects). This is the natural approach for estimating the cohort-specific model, given the cohorts are independent by design (Section 3.2). To see why this approach also works appropriately for the individual-specific model, note that the estimation may not be done in simultaneous equations across subjects because the network information is incomplete and neighborhoods vary all the time, which largely removes the associations of player identities in the game (i.e. Players are recognized by their degree types rather than individual identities). Hence, it would be proper to estimate the individual parameters independently. That said, the independence of parameter estimation substantially reduces the computational effort required for fitting our models, which involve large numbers of parameters – To see, there are 100 (800), 40 (320), and 40 (320) parameters for the cohort- (individual-)specific version of hybrid learning, adaptive learning, and equilibrium learning, correspondingly.

Table E-2. Comparison of behavioral models†

| | hybrid learning | adaptive learning | equilibrium learning |
|---|---|---|---|
| LL (individual-specific)* | -322.29 | -443.59 | -829.48 |
| likelihood ratio test ($p$-value)** | $\chi^2(700)$=288.34 (1.000) | $\chi^2(280)$=243.38 (0.944) | $\chi^2(280)$=103.46 (1.000) |

† based on the entire dataset that includes all treatments
* log likelihood (LL) produced by maximum likelihood estimation (MLE).
** null [alternative] hypothesis: The reduced (cohort-specific) model is more [less] efficient than the full (individual-specific) model in fitting the data.

In the main text we focused on examining the estimation of $r_{itk}$ and its associated derivations ($\overline{r_k}$), as they are crucial to the interpretation of learning behavior. For the sake of completeness, we shall in Table E-3 below provide the estimation results of the original parameters of the hybrid model.

Table E-3. Results of hybrid learning (continued)

| treatment | cohort-specific model, estimated by MLE | | | | |
|---|---|---|---|---|---|
| | $\lambda_i$[22] | $r_{i03}$ | $r_{i02}$ | $\Delta A_{i0k}^{q}$ | $\Delta A_{i0k}^{a}$ |
| *Gh_Shirk* | 0.037 (0.008) | 0.09 (0.117) | 0.26 (0.149) | 116.3 (74.9) | 292.5 (21.5) |
| *Gl_Shirk* | 0.038 (0.020) | 0.32 (0.174) | 0.23 (0.115) | 85.1 (67.3) | 246.9 (97.0) |
| *Gh_Herd* | 0.427 (0.879) | 0.20 (0.238) | 0.34 (0.255) | 87.6 (75.6) | 72.6 (144.5) |
| *Gl_Herd* | 0.040 (0.018) | 0.276 (0.418) | 0.34 (0.291) | 20.2 (31.2) | 335.1 (76.6) |

mean (standard deviation)

## E.4. Learning: Estimation Approach

This appendix will elaborate in detail the estimation strategy in Section 5.4, with the hybrid learning model as an example. The estimations of other learning models in our paper work out in a similar way. For the ease of understanding, we decompose hybrid learning in Figure E-1 below, and we massage the notes and references into the figure to make it self explanatory.

[22] It has to be noted that the seemingly low estimated values of $\lambda_i$ *per se* do not automatically imply high irrationality of the subjects involved, for the following reason: In determining the choice probabilities in the hybrid model, $\lambda_i$ is not directly linked to the attraction of any action (c.f. (5) in the main text). Therefore, unlike in the models of equilibrium learning and adaptive learning, the estimate of $\lambda_i$ in Table E-3 cannot be readily interpreted as the degree of rationality in hybrid learning, and thus mainly serves to calculate other measures of interests (e.g. $\overline{r_k}$ in the main text).

$$p_{itk}^{h} = r_{itk} p_{itk}^{q} + (1 - r_{itk}) p_{itk}^{a}$$

$$r_{i,t+1,k_i} = \begin{cases} \dfrac{p_{i,t+1,k_i}^{q} r_{itk_i}}{p_{i,t+1,k_i}^{q} r_{itk_i} + p_{i,t+1,k_i}^{a} (1 - r_{itk_i})}, & x_{it} = 1 \\ \dfrac{(1 - p_{i,t+1,k_i}^{q}) r_{itk_i}}{(1 - p_{i,t+1,k_i}^{q}) r_{itk_i} + (1 - p_{i,t+1,k_i}^{a})(1 - r_{itk_i})}, & x_{it} = 0 \end{cases}$$

■ For $k \neq k_i$, $r_{itk}$ sticks to its initial value $r_{i0k}$.

■ The initial values $r_{i02}, r_{i03}$ are to be estimated.

$$p_{i,t+1,k}^{q} = \frac{1}{1 + e^{-\lambda_i \Delta A_{itk}^{q}}}$$

■ $\Delta A_{itk}^{q} \coloneqq A_{itk}^{q}(1) - A_{itk}^{q}(0)$

$$p_{i,t+1,k}^{a} = \frac{1}{1 + e^{-\lambda_i \Delta A_{itk}^{a}}}$$

■ $\Delta A_{itk}^{a} \coloneqq A_{itk}^{a}(1) - A_{itk}^{a}(0)$

■ $\lambda_i$ is to be estimated.

$$\Delta A_{itk}^{q} = \Delta U(\sigma_{it}, k, G_i)$$

$$A_{itk}^{a}(x) = \left(1 - \frac{1}{n_{it}}\right) A_{i,t-1,k}^{a}(x) + \frac{1}{n_{it}} \Delta_{it}^{a}(x) v_i(x, y_{it})$$

■ $\Delta U(\sigma_{it}, k, G_i)$ (see equations (B-1) and (B-2) above) is a function of $p_{it2}, p_{it3}$.

■ $\Delta A_{i0k}^{q}$ is to be estimated.

■ $n_{it}, v_i(x, y_{it}), \Delta_{it}^{a}(x)$ are solely determined by the realized data. For $n_{it}$, see equations (E-2) and (E-3) above. For $v_i(x, y_{it})$ and $\Delta_{it}^{a}(x)$, see Section 5.1.

■ $\Delta A_{i0k}^{a}$ is to be estimated.

Figure E-1. Estimation approach of hybrid learning: A graphical illustration

Overall, the hybrid learning is composed of equilibrium learning and adaptive learning, each of these components evolving in a recursive manner. As shown in Figure E-1, we estimate the initial values $r_{i02}, r_{i03}, \Delta A_{i0k}^{q}, \Delta A_{i0k}^{a}$ (which kick off the dynamics), and the sensitivity parameter $\lambda_i$ (which controls the convergence). The maximal likelihood estimation then produces the results we presented in Table E-3 and in Table 9.

**References**


Aral, S., Walker, D., 2011. Creating social contagion through viral product design: A randomized trial of peer influence in networks. Management Science, 57(9), 1623-1639.

Aviv, Y., Pazgal, A., 2008. Optimal pricing of seasonal products in the presence of forward-looking consumers. Manufacturing & Service Operations Management, 10(3), 339-359.

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

Barabási, A. L., Albert, R., 1999. Emergence of scaling in random networks. Science, 286(5439), 509-512.

Berninghaus, S. K., Ehrhart, K. M., Keser, C., 2002. Conventions and local interaction structures: experimental evidence. Games and Economic Behavior, 39(2), 177-205.

Bollobás B., 1985. Random Graphs. London: Academic Press.

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. The American Economic Review, 104(3), 898-930.

Brown, J. N., Rosenthal, R. W., 1990. Testing the minimax hypothesis: a re-examination of O'Neill's game experiment. Econometrica, 1065-1081.

Cachon, G. P., Swinney, R., 2009. Purchasing, pricing, and quick response in the presence of strategic consumers. Management Science, 55(3), 497-511.

Cachon, G. P., Swinney, R., 2011. The value of fast fashion: Quick response, enhanced design, and strategic consumer behavior. Management Science, 57(4), 778-795.

Cadsby, C. B., Maynes, E., 1999. Voluntary provision of threshold public goods with continuous contributions: experimental evidence. Journal of Public Economics, 71(1), 53-73.

Camerer, C., Hua Ho, T., 1999. Experience-weighted attraction learning in normal form games. Econometrica, 67(4), 827-874.

Camerer, C. F., Ho, T. H., Chong, J. K., 2002. Sophisticated experience-weighted attraction learning and strategic teaching in repeated games. Journal of Economic theory, 104(1), 137-188.

Cassar, A., 2007. Coordination and cooperation in local, random and small world networks: Experimental evidence. Games and Economic Behavior, 58(2), 209-230.

Charness, G., Feri, F., Meléndez‐Jiménez, M. A., Sutter, M., 2014. Experimental games on networks: Underpinnings of behavior and equilibrium selection. Econometrica, 82(5), 1615-1670.

Cho, M., 2010. Endogenous formation of networks for local public goods. International Journal of Game Theory, 39(4), 529-562.

Chong, J. K., Camerer, C. F., Ho, T. H., 2006. A learning-based model of repeated games with incomplete information. Games and Economic Behavior, 55(2), 340-371.

Chung, K. K., Hossain, L., Davis, J., 2005. Exploring sociocentric and egocentric approaches for social network analysis. KMAP 2005: Second International Conference on Knowledge Management in Asia Pacific: 1– 8. Wellington: Victoria University of Wellington.

Conley, T., Udry, C., 2001. Social learning through networks: The adoption of new agricultural technologies in Ghana. American Journal of Agricultural Economics, 83(3), 668-673.

Corbae, D., Duffy, J., 2008. Experiments with network formation. Games and Economic Behavior, 64(1), 81-120.

Elmaghraby, W., Gülcü, A., Keskinocak, P., 2008. Designing optimal preannounced markdowns in the presence of rational customers with multiunit demands. Manufacturing & Service Operations Management, 10(1), 126-148.
Erdos, P., Rényi, A., 1960. On the evolution of random graphs. Publications of the Mathematical Institute of the Hungarian Academy of Sciences, 5(1), 17-60.
Erev, I., Roth, A. E., 1998. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. The American Economic Review, 88(4), 848-881.
Fehr, E., Gächter, S., 2000. Cooperation and punishment in public goods experiments. American Economic Review, 90(4), 980-994.
Fischbacher, U., 2007. z-Tree: Zurich toolbox for ready-made economic experiments. Experimental Economics, 10(2), 171-178.
Foster, A. D., Rosenzweig, M. R., 1995. Learning by doing and learning from others: Human capital and technical change in agriculture. Journal of political Economy, 103(6), 1176-1209.
Fudenberg, D., Levine, D. K., 1998. The theory of learning in games. Cambridge, MA: MIT Press.
Galeotti, A., Goyal, S., 2010. The law of the few. The American Economic Review, 100(4), 1468-1492.
Galeotti, A., Goyal, S., Jackson, M. O., Vega-Redondo, F., Yariv, L., 2010. Network games. The Review of Economic Studies, 77(1), 218-244.
Gibbons, R., 1992. Game theory for applied economists. Princeton, NJ: Princeton University Press.
Harsanyi, J. C., Selten, R., 1988. A general theory of equilibrium selection in games. Cambridge, MA: MIT Press.
Ho, T. H., Camerer, C. F., Chong, J. K., 2007. Self-tuning experience weighted attraction learning in games. Journal of Economic Theory, 133(1), 177-198.
Hojman, D. A., Szeidl, A., 2008. Core and periphery in networks. Journal of Economic Theory, 139(1), 295-309.
Ioannides, Y., 2012. From Neighborhoods to Nations: The Economics of Social Interactions. Princeton, NJ: Princeton University Press.
Jackson, M., 2008. Social and Economic Networks. Princeton, NJ: Princeton University Press.
Jackson, M., Zenou, Y., 2015. Games on networks. In: Young, P., Zamir, S. (Eds.). Handbook of Game Theory with Economic Applications, Vol. 4. Amsterdam: Elsevier Publisher, 95-163.
Jackson, M., Rogers, B., Zenou, Y., 2017. The economic consequences of social network structure. Journal of Economic Literature, 55(1), 49-95.
Judd, S., Kearns, M., Vorobeychik, Y., 2010. Behavioral dynamics and influence in networked coloring and consensus. Proceedings of the National Academy of Sciences, 107(34), 14978-14982.
Katz, M. L., Shapiro, C., 1985. Network externalities, competition, and compatibility. The American Economic Review, 75(3), 424-440.
Katz, M. L., Shapiro, C., 1986. Technology adoption in the presence of network externalities. The Journal of Political Economy, 94(4), 822-841.
Katz, M. L., Shapiro, C., 1992. Product introduction with network externalities. The Journal of Industrial Economics, 40(1), 55-83.
Keser, C., Ehrhart, K. M., Berninghaus, S. K., 1998. Coordination and local interaction: experimental evidence. Economics Letters, 58(3), 269-275.
Kosfeld, M., 2004. Economic networks in the laboratory: A survey. Review of Network Economics, 3(1).
Ledyard, J. O., 1997. Public goods: a survey of experimental research. In: Kagel J. O., Roth, A. E. (Eds.). Handbook of Experimental Economics, Princeton, NJ: Princeton University Press, 111-194.

Lee, E., Lee, J., Lee, J., 2006. Reconsideration of the winner-take-all hypothesis: Complex networks and local bias. Management Science, 52(12), 1838-1848.
Lee, D., Mendelson, H., 2008. Divide and conquer: Competing with free technology under network effects. Production and Operations Management, 17(1), 12-28.
Liu, Q., Van Ryzin, G. J., 2008. Strategic capacity rationing to induce early purchases. Management Science, 54(6), 1115-1131.
Marsden, P. V., 2002. Egocentric and sociocentric measures of network centrality. Social Networks, 24(4), 407-422.
McKelvey, R. D., Palfrey, T. R., 1995. Quantal response equilibria for normal form games. Games and Economic Behavior, 10(1), 6-38.
Miguel, E., Kremer, M., 2003. Networks, Social Learning and Technology Adoption: The Case of Deworming Drugs in Kenya. working paper.
Mobius, M. M., Niehaus, P., Rosenblat, T. S., 2007. Social learning and consumer demand. Meeting Papers. Society for Economic Dynamics.
Munshi, K., 2004. Social learning in a heterogeneous population: technology diffusion in the Indian Green Revolution. Journal of Development Economics, 73(1), 185-213.
Oestreicher-Singer, G., Sundararajan, A., 2012a. The visible hand? Demand effects of recommendation networks in electronic markets. Management Science, 58(11), 1963-1981.
Oestreicher-Singer, G., Sundararajan, A., 2012b. Recommendation networks and the long tail of electronic commerce. MIS Quarterly, 36(1), 65-83.
Osborne, M. J., Rubinstein, A., 1994. A course in game theory. Cambridge, MA: The MIT Press
Rao, N., Mobius, M. M., Rosenblat, T., 2007. Social networks and vaccination decisions. working paper.
Riggins, F. J., Kriebel, C. H., Mukhopadhyay, T., 1994. The growth of interorganizational systems in the presence of network externalities. Management Science, 40(8), 984-998.
Rosenkranz, S., Weitzel, U., 2012. Network structure and strategic investments: An experimental analysis. Games and Economic Behavior, 75(2), 898-920.
Su, X., Zhang, F., 2008. Strategic customer behavior, commitment, and supply chain performance. Management Science, 54(10), 1759-1773.
Su, X., Zhang, F., 2009. On the value of commitment and availability guarantees when selling to strategic consumers. Management Science. 55 (5), 713–726.
Sundararajan, A., 2007. Local network effects and complex network structure. The B.E. Journal of Theoretical Economics, 7(1), 1-35.
Sundararajan, A., Provost, F., Oestreicher-Singer, G., Aral, S., 2013. Information in digital, economic, and social networks. Information Systems Research, 24(4), 883-905.
Van Huyck, J. B., Battalio, R. C., Beil, R. O., 1990. Tacit coordination games, strategic uncertainty, and coordination failure. The American Economic Review, 80(1), 234-248.
Van Zandt, T., Vives, X., 2007. Monotone equilibria in Bayesian games of strategic complementarities. Journal of Economic Theory, 134(1), 339-360.
Weber, R. A., 2006. Managing growth to achieve efficient coordination in large groups. The American Economic Review, 96(1), 114-126.