Naive Learning and Game Play in a Dual Social Network Framework*

Zhengzheng Pan† Robert P. Gilles‡

March 2010

Abstract

We observe that people perform economic activities within the social setting of a small group, while they obtain relevant information from a broader source. We capture this feature with a dynamic interaction model based on two separate social networks. Individuals play a coordination game in an *interaction network*. Meanwhile, all individuals update their strategies via a naive learning process using information from a separate *influence network* through which information is disseminated. In each time period, the interaction and influence networks co-evolve, and the individuals' strategies are updated through a modified French-DeGroot updating process. We show that through this updating process both network structures and players' mixed strategies always reach a steady state. In particular, conformity occurs in the long run when the interaction cost is sufficiently low. We also analyze the influence exerted by a minority group on these outcomes.

JEL: D70, D83, D85, C63, A14, L14

Keywords: social learning, social network, diffusion, coordination game

*Most of the work was done at the Economics Department of Virginia Tech. We thank Hans Haller for his valuable comments and useful suggestions on previous drafts of this paper.

†Corresponding author. Network Dynamics of Simulation Science Laboratory, Virginia Bioinformatics Institute, Virginia Tech, Blacksburg, VA 24061, USA. Email: zzpan@vt.edu

‡School of Management, Queen's University, Belfast BT7 1NN, UK. Email: r.gilles@qub.ac.uk

1 Introduction

The Nash equilibrium concept is a widely used and studied tool in game theory. How players get to those equilibrium states or actually play games, however, remain questions that motivate many of us. We know that Nash equilibria can be justified by assuming common knowledge (Aumann and Brandenburger, 1995), but such a strong assumption seems unreasonable when we try to understand game play in the cases involving less than perfectly informed players (Dekel and Fudenberg, 1990; Borgers, 1994; Canning, 1995), less than perfectly rational players (Simon, 1955; Gabaix, Laibson, Moloche, and Weinberg, 2006), or consider the complexity of the environment where games take place (Vega-Redondo, 2007).

What all learning models share in common is the accessibility to or observation of certain information. The information—which can concern strategies, payoffs, or signals from other players—is collected and used to adjust one's strategy, i.e., a player learns from the obtained information based on some rules. As Griliches (1957) already pointed out, one needs to have a connection in order to obtain such information. So, any learning process implicitly assumes an underlying social network which serves as a platform for the dissemination of information.

In this paper we introduce a framework in which game play and learning are explicitly separated. Hence, we present two separate social spheres: an interaction network that describes how individual players play a coordination game with each other and a social information sharing or "influence" network where players learn about the strategies and success of other players. The two networks are distinct, but correlated. We consider a learning process in which the interaction network, the influence network, and the selected strategies are all updated sequentially.

We extend the well-known French-DeGroot naive learning process to our dual network setting. The naive learning process seminally developed by French (1956) and DeGroot (1974) is in nature a simple Markovian approach to learning based on the principle that individuals use weighted averages of observed characteristics of other players.[1] We incorporate the French-DeGroot updating rule into our dual-network framework by having a player refer to the achieved payoffs by her endogenously chosen partners in game play when assigning weights on collected information. In other words, we assume complete availability of information on all aspects of game play, but then assume that players selectively process this information based on the performance of the observed players and with

[1]The French-DeGroot process has also been used in related work by Friedkin and Johnsen (1990, 1997); DeMarzo, Vayanos, and Zwiebel (2003); and Golub and Jackson (2007).

bounded rationality.

In each time period, *two* randomly selected players update their interaction neighborhood. This is followed by *all* players updating their influence weights as well as the mixed strategy they use in their game play. These updates are based on observations all players make about other players. Players take the costs of interaction into account in their decision-making. Only if expected payoffs exceed these costs, interaction (game play) with another player is initiated or maintained. Throughout we assume that observation of other players' strategies and payoffs is costless, although the game play activity is not.

We show that if interaction costs are sufficiently low, this naive learning process converges to a state of full conformity. This state is one in which all players play the *same* mixed strategy, interact with *all* other agents, and assign equal weights to observations made about *all* players. On the other hand, if the interaction costs are relatively high, no interaction takes place and all players remain autarkic. Finally, for cost levels in the mid-range, there emerges a myriad of outcomes in this naive learning process. Computer simulations show that there exist multiple steady states in these circumstances.

This observed conformism for low interaction costs coincides with the empirical studies on influence in social networks, which indicate that individuals' decision-making processes, opinions, and behavioral patterns are affected by their (social) neighbors. For instance, studies show similarity in investment patterns (Duflo and Saez, 2002) and behaviors of neighborhood peers appear to substantially affect youth behaviors (Case and Katz, 1991). However, the steady state under conformism is usually not a Nash equilibrium. This is as argued by Blume and Easley (2006): Naive learning is limited in nature and does not necessarily converge to a steady state satisfying rational expectations.

Finally, we consider the influence of minority groups of individuals who are interpreted to be *persistent*. Following a similar model developed in Pan (2008), we define a player to be persistent if she does not modify her initially assigned mixed strategy in the given coordination game. Persistent players, however, remain subject to the other aspects of the naive learning process. They update their interaction neighborhood as well as the influence weights that they assign to other players, which are not used to update the strategy.

We look at two types of persistency. First, we consider the case that all persistent players use exactly the same strategy in their game play. In this case of uniform persistency, for sufficiently low interaction costs the whole society converges to full conformity in which all players play the strategy adhered to these persistent players. As a corollary, a single persistent player can sway the whole community to select her initial strategy. Second, we consider a heterogeneous group of persistent players. In this case, for low enough interaction costs, the non-persistent players converge in their game play to a strategy that is a convex combination of the strategies used by members of the group of persistent players. Computer simulations show that with both types, increased interaction costs breaks down this conformity and a complex situation with multiple steady states emerges.

Relation to the existing literature

Within our novel dual network setting, we use three principles in our updating rules that distinguishes our model from previous studies. First, the nature of the French-DeGroot learning process requires the use of a continuous state space, thus restricting us to the use of mixed strategies in a 2×2 coordination game. While many have used coordination games in studying evolutionary learning process (Foster and Young, 1990; Kandori, Mailath, and Rob, 1993; Ellison, 1993), these studies mainly focus which pure strategies are used by the players in the population.

Second, the interaction network where game play takes place evolves endogenously over time, determined by cost-benefit evaluation and consent on new and existing links. Alternatively, Galeotti, Goyal, Jackson, Vega-Redondo, and Yariv (2009) introduced a model of game play in networks with learning, but in their model game play and learning are not differentiated and there is no endogenous mechanism for changing the network structure. Network dynamics in Jackson and Watts (2002) is similar to ours, but based on costless interaction and shows that there emerge fully connected networks. We consider positive interaction costs and analyze the outcomes based on cost levels, which do not always converge to a fully connected network.

Third, our learning process takes into account the network structure and results of game play when updating influence weights. Weights placed on information are the main factor in naive learning and are in many cases assumed to be constant over time (French, 1956; Bala and Goyal, 1998; DeMarzo, Vayanos, and Zwiebel, 2003; Friedkin and Johnsen, 1997; Golub and Jackson, 2007). Pan (2008), Hegselmann and Krause (2002) and Weisbuch, Deffuant, Amblard, and Nadal (2002) presented several naive learning models with time-varying weights to capture the change in players' attitudes in social communication. Without the layer of game play, the weight updating in these models certainly differs from ours. Our mechanism also ties the two social spheres together seamlessly.

The remainder of this paper is structured as follows. The next section introduces the formal setup of the model and the updating process. Section 3 analyzes the outcomes from the standard setting, while Section 4 discusses the influence of persistent players in our framework. Finally, Section 5 draws some conclusions for future directions of research. All proofs are relegated to the appendix.

2 A network model of social interaction

We consider a finite set of players $N = \{1, 2, \ldots, n\}$ who engage in binary value-generating activities with their neighbors in a social network. Their engagement is assumed to consist of playing a specified coordination game. Standard hypotheses usually impose that this association is based on some form of rational behavior. Here we explicitly restrict these players' rationality in that they do not optimize over their strategy set; instead, the players select a strategy by weighing the information collected on the success of other players and the strategies that they use.

Furthermore, we assume that these individuals collect this information through a separate information sharing network. Within the information collection process, players observe all other players' actions and payoffs, and assign a certain weight to the observation for each player. Based on these weights and observations, a player then determines her own strategy in the coordination games she plays in the social interaction network.

We first introduce a description of the social interaction network, the information sharing network, and subsequently formulate the dynamic interaction process.

2.1 The social interaction network

Each player in $N = \{1, 2, \ldots, n\}$ selectively builds social relationships with other players. The resulting *interaction network* at time $t \in \mathbb{N}$ is represented by an $n \times n$ adjacency matrix $\mathbf{G}^t$ with

$$\mathbf{G}^t_{ij} = \begin{cases} 1 & \text{if } i, j \text{ are connected,} \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Define $L^t_i = \{j \in N \mid \mathbf{G}^t_{ij} = 1\}$ as the set of player i's neighbors at time t in the interaction network $\mathbf{G}^t$. For technical convenience, we assume that each player is always connected with herself, i.e., $\mathbf{G}^t_{ii} = 1$, for all $i \in N$, for all t. Also, $\mathbf{G}^t_{ij} = \mathbf{G}^t_{ji}$ for all $i, j \in N$, for all t, implying that the interaction network $\mathbf{G}^t$ is symmetric.

Throughout we assume that every connection in $\mathbf{G}^t$ is consent-based, which means that permission from both players is required for a link to be formed. On the other hand, a single player can always sever any link under her control.[2]

Denote g^t as the network representation of the interaction network $\mathbf{G}^t$ defined by

$$g^t = \{ij \mid \mathbf{G}^t_{ij} = \mathbf{G}^t_{ji} = 1\}. \tag{2}$$

[2]This is akin to the concept of pairwise stability seminally introduced in Jackson and Wolinsky (1996) and the standard stability concept in matching markets (Roth and Sotomayor, 1990).

From the above, $ii \in g^t$ for every player $i \in N$ and every $t \in \mathbb{N}$. We use $g_0 = \{ii \mid i \in N\}$ to denote the sparsest possible network and $g_N = \{ij \mid i, j \in N\}$ to denote the complete network. The process of adding and deleting a link between players i and j at time t can be written as $g^t + ij$ and $g^t - ij$, respectively.

We assume that interaction is costly, i.e., both the initiation and the maintenance of a link between two players imposes the same cost $c \geqslant 0$ on both interacting parties. This implies that, when a link is initiated, both players pay the common interaction cost c. Also, each player pays the common interaction cost c for the maintenance of every existing link $ij \in g^t$, $j \neq i$, during each time period t. We emphasize that here we assume that each player $i \in N$ has no costs of interacting with herself.

A player $i \in N$ only interacts with her neighbors $j \in L_i^t$ at time $t \in \mathbb{N}$. The association between each pair of connected players is modelled as a 2×2 symmetric coordination game shown in Table 1.

	A	B
A	a, a	0, 0
B	0, 0	1, 1

Table 1: The 2×2 coordination game played between linked players

In this coordination game, we have two pure strategy Nash equilibria: (A, A) and (B, B). It is assumed that $a \geqslant 1$, therefore (A, A) is the Pareto optimal equilibrium. (A, A) is also the risk-dominant equilibrium (Harsanyi and Selten, 1988), which is the pure strategy equilibrium with a larger basin of attraction than (B, B).[3] Although from a strategic point of view, A is in all respects a superior convention, we study a social learning process in which the population will usually not settle on the convention to play A.

All players' actions at time t can be represented by an n-dimensional *mixed strategy* vector $\mathbf{p}^t = (p_1^t, \dots, p_n^t)^{\mathrm{T}} \in [0, 1]^n$, where $p_i^t \in [0, 1]$ is the probability that player i chooses A. The strategy vector is time-dependent as players modify their actions over time based on the information they collect. Now, at time t player i receives a payoff of π_{ij}^t by interacting with player j with

$$\pi_{ij}^t = a p_i^t p_j^t + (1 - p_i^t)(1 - p_j^t). \tag{3}$$

Due to the symmetric nature of the coordination game, each pair of players $i, j \in N$ re-

[3] In other words, A is the strategy that is a best response to the largest set of beliefs over possible plays of the opponent. Specifically, playing A is a player's best response if the fraction of her opponents who play A is greater than or equal to $\frac{1}{a+1}$ and less than $\frac{1}{2}$.

ceives identical payoffs from such coordination, i.e., $\pi^t_{ij} = \pi^t_{ji}$, for all $i, j \in N$.

2.2 Information dissemination

In our approach information dissemination is separated from actual game play that takes place in the interaction network $\mathbf{G}^t$. As mentioned in the introduction, every individual uses a naive updating process based on the selection of a weighted average of the mixed strategies of the players that influence her.

The hypothesis that a player can observe another player's *mixed* strategy is a very strong assumption, in particular since players actually only execute the two pure strategies A or B. This hypothesis, therefore, requires justification. We consider two possible justifications.

First, we can interpret this hypothesis as that players actually communicate their mixed strategies within the prevailing information sharing network. This is akin to the assumption that mixed strategies are deliberately selected objects that are intentionally executed by the players in the set N. This implies in turn that the naive learning process assumed here is actually a consequence of imperfections in the process of communication and information sharing between players that deliberately share full information about their selected mixed strategies.

Second, we can assume that players are actually more boundedly rational and observe each other's game play over a sufficiently long period of interaction. This is equivalent to the assumption that each time period t consists of multiple playing rounds in the given interaction network $\mathbf{G}^t$. Only after observing sufficiently many playing rounds, the updating process ensues based on the collected information, which for player j observing player i is assumed to approximate the true value of p^t_i. We leave such a playing process in time period t as unspecified in our model. However, such a specification could be incorporated into our framework. It is also clear that our main conclusions would not change with such an extension of our model. Such a modification would therefore just unnecessarily complicate the exposition.

Within our framework, information sharing is equivalent to observing other players' actions and attaching a weight to these observations. Such a weight can be interpreted as the level of "trust" that a player puts on another player's decisions. As such this weight indicates how much influence one player has over another. Therefore, we formally model the information sharing network as a set of influence weights.

Formally, for every time period $t \in \mathbb{N}$ we introduce an $n \times n$ nonnegative matrix $\mathbf{T}^t$ which we refer to as the *influence matrix* at time t. For all $i, j \in N$, the number $\mathbf{T}^t_{ij} \in [0, 1]$

indicates the weight that player i places on player j's strategic choice at time t and a higher weight indicates that one player weighs the other more on the choice of her strategy. Thus, the influence matrix $\mathbf{T}^t$ captures the information collection process at time t.

We assume that for every $t \in \mathbb{N}$ the influence matrix $\mathbf{T}^t$ is row-stochastic, i.e., the influence weights sum up to unity for each player $i \in N$:

$$\sum_{j=1}^{n} \mathbf{T}^t_{ij} = 1 \text{ and } \mathbf{T}^t_{ij} \geqslant 0, \text{ for all } i, j \in N, \text{ for all } t \in \mathbb{N}. \tag{4}$$

Unlike the interaction network $\mathbf{G}^t$, $\mathbf{T}^t$ may be asymmetric, so that $\mathbf{T}^t_{ij} \neq \mathbf{T}^t_{ji}$ for some i, j. Moreover, one's information collection is not restricted to one's neighbors. That is, for some i, j, $\mathbf{T}^t_{ij} > 0$ while $\mathbf{G}^t_{ij} = 0$. On the other hand, the two matrices are correlated through a dynamic updating process, as will be discussed in the next subsection.

2.3 The updating process

We first assign initial states for both social networks and players' strategies. Then, after initialization, two players are randomply selected and each selected player updates her interaction network using rules based on pairwise stability; subsequently, all players update their influence weights based on observations in the information-sharing network; and, finally, all players update their mixed strategies and play the game with their partners in the interaction network. Schematically, this updating process can be represented in a flow diagram in Figure 1.

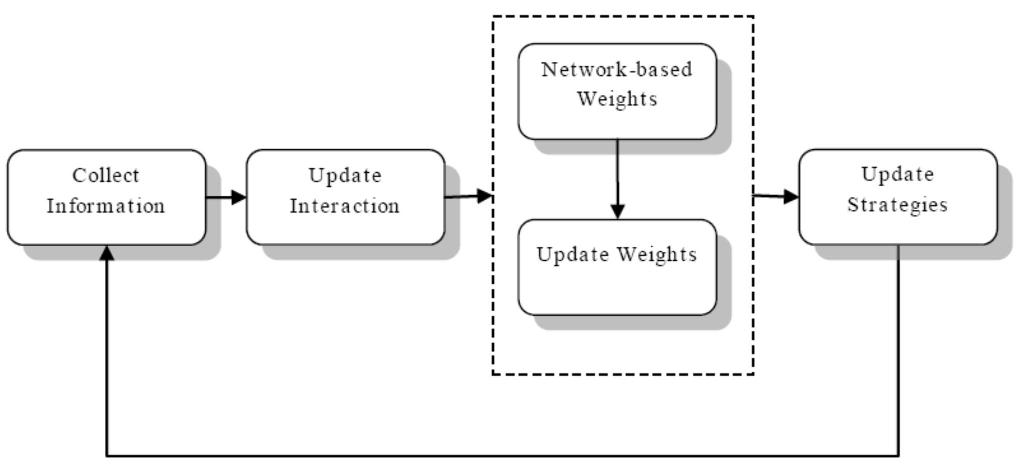


Figure 1: Updating process in our dual network framework

Initialization

At $t = 0$ we have an initial coordination structure $\mathbf{G}^0$ and an initial influence matrix $\mathbf{T}^0$. The initial coordination structure is assumed to be autarkic, i.e., $g^0 = g_0$ and given by

$$\mathbf{G}^0_{ii} = 1, \text{ for all } i \in N \text{ and } \mathbf{G}^0_{ij} = 0 \text{ for all } i \neq j. \tag{5}$$

As for the initialization of the information network and the strategies, we assume that players initially are assigned arbitrary strategies and an arbitrary influence distribution. That is,

$$p^0_i \in [0,1], \text{ for all } i \in N. \tag{6}$$

$$\sum_{j=1}^{n} \mathbf{T}^0_{ij} = 1 \text{ for all } i \in N; \quad \mathbf{T}^0_{ij} \in [0,1], \text{ for all } i, j \in N. \tag{7}$$

Note that although $\mathbf{G}^0_{ii} = 1$ for all i, the case that $\mathbf{T}^0_{ii} = 0$ is not excluded. Thus, it is possible that one assigns zero weight on oneself during the initialization period even though one plays the coordination game with oneself according to the hypotheses made so far.

Updating the interaction network

During the updating process in period t, two players i and j are randomly selected to consider their connectivity. The link ij will be formed (if the two are not connected) or maintained (if they are connected already) if and only if for both of them, the payoffs from the link covers at least the interaction cost.[4] That is, the updating rule for the interaction network $\mathbf{G}^t$ is given by

$$\mathbf{G}^t_{ij} = \begin{cases} 1 & \text{if } \pi^{t-1}_{ij} \geqslant c \\ 0 & \text{if } \pi^{t-1}_{ij} < c \end{cases} \tag{8}$$

and

$$\mathbf{G}^t_{hk} = \mathbf{G}^{t-1}_{hk} \text{ for all } (h,k) \neq (i,j) \tag{9}$$

Updating the information-sharing network

After the two randomly selected players i and j update their interaction relationship $\mathbf{G}^t_{ij}$ described above, all players update their weight assignments in the influence network $\mathbf{T}^{t-1}$. We model the updating of the influence matrix $\mathbf{T}^{t-1}$ to be based on the observed payoffs

[4] Note here that we assume that player's mixed strategies (and therefore payoffs) are fully observable and available. This again, is derived from the principles of naive learning, as discussed earlier.

from the game play in $\mathbf{G}^{t-1}$. The principle is that a player's partners or "neighbors" act as effective filters for more beneficial links and higher payoffs. The reasoning is that links are formed and maintained based on a cost-benefit evaluation. Therefore for a player, the connectedness between one of her neighbors and another player implies a reasonable potential for collecting sufficiently high payoffs between that player and her neighbor's partner.

Namely, when an player decides on how much influence weight to place on another player, she calculates the total payoffs that her neighbors could obtain from associating with that player, given all players' past actions and connectivity. This observation procedure is carried on among all players, while the selected connectivity shows its impact on processing the collected information.

We recall that the influence matrix is row-stochastic. Thus, each player redistributes her influence weight assignment proportionally according to the total payoffs and then normalizes the weights to make sure that the row sum equals to unity. This implies that the redistribution of influence follows the rules below:

$$\mathbf{T}^t_{ij} = \frac{w^t_{ij}}{\sum_{k=1}^{n} w^t_{ik}}, \text{ for all } i, j \in N \text{ and } t \in \mathbb{N}, \tag{10}$$

where

$$w^t_{ij} = \sum_{l \in L^t_i} \mathbf{G}^t_{lj} \pi^{t-1}_{lj}.$$

Consider the weight $\mathbf{T}^t_{ij}$ assigned by i to j. If j's action does not guarantee a sufficiently high payoff, j would not be connected with any of i's neighbors. That is, $\mathbf{G}^t_{kj} = 0$ for all $k \in L^t_i$. Consequently, $w^t_{ij} = 0$, which results to zero weight $\mathbf{T}^t_{ij} = 0$. That is, a player does not place any weight on someone who does not provide the potential to generate high enough payoffs.

Also, we note that if $L^t_i = \emptyset$, we have a problem when applying the equations above in that all weights w^t_{ij} are 0 and the sum of the i-th row of the influence matrix $\mathbf{T}^t$ does not add up to 1. This problem is prevented by the assumption that each player is connected with herself during initialization at $t = 0$ and stays connected with herself during any subsequent period $t \in \mathbb{N}$ since it is assumed that player i has no costs related to her self-referential coordination $\mathbf{G}^t_{ii} = 1$ or $ii \in g^t$.

For those who get zero weights in the influence matrix, their actions and information from them do not count when player i updates her mixed strategy p^t_i. In other words, each player actually takes the weighted average among the beneficial or potentially beneficial actions during the updating process.

Strategy updating and game play

Finally, all players update their mixed strategies based on the information collected in period $t-1$. Using the updated influence matrix $\mathbf{T}^t$, all players determine their mixed strategy using the French-DeGroot updating rule. That is,

$$p_i^t = \sum_{j \in N} \mathbf{T}_{ij}^t p_j^{t-1} \qquad \text{for all } i \in N, t > 0. \tag{11}$$

So the updating process for all players can be conveniently written as:

$$\mathbf{p}^t = \mathbf{T}^t \mathbf{p}^{t-1}. \tag{12}$$

After updating her mixed strategy, each player $i \in N$ plays the given coordination game with her neighbors $j \in L_i^t$ in $\mathbf{G}^t$ and collects payoffs for the period t. Subtracting her interaction costs for all active links (except the one with herself) in period t we get payoffs

$$\pi_i^t = \sum_{h \in L_i^t} \pi_{ih}^t - (\#L_i^t - 1)c = \sum_{h \in L_i^t} [a p_i^t p_h^t + (1-p_i^t)(1-p_h^t)] - (\#L_i^t - 1)c, \tag{13}$$

where $\#L_i^t$ is the number of i's neighbors during time period t.

The time period t ends when game play is completed. The process repeats in the next period $t+1$.

3 Convergence of Behavior

In the process described above, players update their neighborhood structure, influence weights, and strategies in a myopic manner in that they do not consider the implications of updating to the future. Rather, they base their decisions on the success of the mixed strategies adopted by the players in the past period that they observe. Our examination of the convergence of this learning process follows studies such as Ellison and Fudenberg (1993, 1995); Bala and Goyal (1998); Banerjee and Fudenberg (2004); Lorenz (2005); and Golub and Jackson (2007).

The cost of coordination $c \geqslant 0$ is a critical factor that largely determines the interaction network structure, which in turn affects the updating of the influence weights and consequently the adjustment of strategies. The next proposition determines the relevant bounds for the interaction cost c.

Proposition 3.1 *Let p^0 be the initialized strategy tuple in the player set N. Denote by $\underline{p}^0 = \min\{p_1^0, \ldots, p_n^0\}$ its lower bound and by $\overline{p}^0 = \max\{p_1^0, \ldots, p_n^0\}$ its upper bound. Let*

$$\underline{\pi} = \min\{\frac{a}{a+1},\ \ a\underline{p}^0\overline{p}^0 + (1-\underline{p}^0)(1-\overline{p}^0)\} \geqslant 0, \tag{14}$$

$$\overline{\pi} = \max\{a(\overline{p}^0)^2 + (1-\overline{p}^0)^2,\ \ a(\underline{p}^0)^2 + (1-\underline{p}^0)^2\} > 0. \tag{15}$$

Then $\underline{\pi}$ and $\overline{\pi}$ are a lower bound and an upper bound, respectively, for the set of all payoffs $\{\pi_{ij}^t \mid i, j \in N$ and $t \in \mathbb{N}\}$.

A proof of Proposition 3.1 can be found in Appendix A of this paper.

Our first main result states that if the costs of formation of interaction links are sufficiently low, the interaction network converges to a fully connected network. Also, the steady state strategies show complete conformity.

Theorem 1 *If $c \leqslant \underline{\pi}$, then the updating process converges to a situation in which there emerge a fully connected interaction network, evenly distributed weights, and all players choose the same mixed strategy $p^* \in [0,1]$, i.e., for every $\epsilon > 0$ there exists some $t^* > 0$ such that for all $t > t^*$ it holds that $\mathbf{G}_{ij}^t = 1$ as well as $|\mathbf{T}_{ij}^t - \frac{1}{n}| < \epsilon$ for all $i, j \in N$ and $|p_i^t - p^*| < \epsilon$ for all players $i \in N$.*

The limit strategy p^* acts as a convention in the given society, taking the form of a mixed strategy. We remark that at this low cost level, the society converges to a steady state that is not necessarily a Nash equilibrium unless we have extreme initial conditions where $\mathbf{p}^0 = (0, \ldots, 0)^{\mathrm{T}}$ or $\mathbf{p}^0 = (1, \ldots, 1)^{\mathrm{T}}$. In particular, even if some of the players select a pure strategy initially, they will abandon that selection in favor of a purely mixed strategy through the influence of other players. This statement is formalized in Lemma 2 in Appendix B, where also the proofs for Lemma 2 and Theorem 1 are collected.

The next assertion is a counterpoint to the observations made in Theorem 1. The proof is rather straightforward and given here.

Theorem 2 *If $c \geqslant \overline{\pi}$, then the updating process converges to a situation in which there emerge a autarkic interaction network, the influence matrix is an identity matrix, and all players choose their initial strategy p_i^0.*

Proof. Assume that $c \geqslant \overline{\pi}$. This implies that no two players will choose to be connected since the payoffs can never exceed the interaction cost. Therefore, the network remains in its initial autarkic pattern, where each player is connected to herself only.
Also, at $t = 1$, $\mathbf{T}_{ii}^1 = 1$ and $\mathbf{T}_{ij}^1 = 0$ for all i and $j \neq i$ because $w_{ij}^1 = 0$. This implies that the influence matrix $\mathbf{T}^1$ is an $n \times n$ identity matrix $\mathbf{I}(n)$ and $p_i^1 = p_i^0$ for all i. Obviously this pattern does not change, since $\mathbf{G}^t$ is always equal to the autarkic pattern. ■

Sensitivity Analysis

Above we have described the outcomes of the updating process when the interaction cost c is very low or very high. With the arbitrary initial weights and strategies, as well as the randomness in selecting players to update the interaction network, when cost is in the medium range determined as $\underline{\pi} < c < \overline{\pi}$, the outcomes cannot be effectively analyzed. In this subsection we examine the outcomes for different cost levels with the help of computer simulations.

In these simulations we set the highest possible payoff which can be obtain from playing (A, A) at $a = 2$, implying that the interaction cost c effectively ranges from 0 to 2. We take $\delta = 0.05$ as the increment in the cost level. We observe that in each of the simulated cases the dynamic updating process converges to a steady state. The society size varies among 20, 40, 60, 80, and 100. Recall that during each period $t \in \mathbb{N}$, two players $i, j \in N$ are chosen randomly to update their interaction network. Thus, each configuration of society size, cost, and initial conditions (the arbitrary $\mathbf{p}^0$ and $\mathbf{T}^0$) is used to run the simulation 3 times, in order to examine different outcome patterns to capture any randomness in updating the interaction network.

The change in society size does not show any significant effect on the final outcome. Therefore, in this subsection we only show the results for $n = 20$. Results are recorded when the learning process researches a steady state.[5]

In Figure 2 the x-axis shows value of cost c. In panel (a), the y-axis indicates the standard deviation among all strategies at the steady state given by

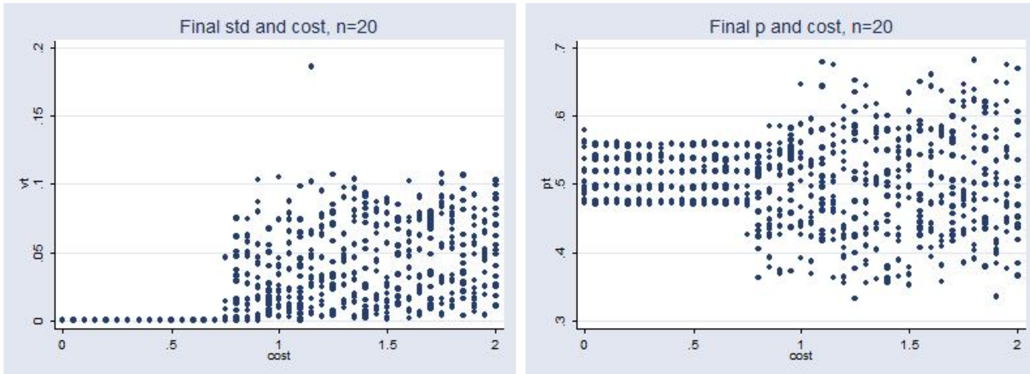
$$\sigma_p = \sqrt{\sum_{i \in N} (p_i^t - \mu_p)^2} \text{ where } \mu_p = \frac{1}{n} \sum_{i \in N} p_i^t.$$

In panel (b), the y-axis shows the mean value of all strategies at the steady state given by μ_p.

When standard deviation equals 0, it means that the strategies fully conform. In panel (a) we see that standard deviation is 0 when cost is relatively low. In these cases the mean values shown in panel (b) fall into a narrow range—represented by a thin horizontal bar showing some partially overlapped dots—due to the randomness in updating.

On the other hand, when the interaction costs increase the outcomes are quite random. The standard deviation could be anywhere between 0 and 0.1. In one case the standard deviation is 0.2, which suggests widely spread out distribution patterns of strategies. Also,

[5]Similar to Pan (2008), the computer program determines a steady state when $\| \Delta \mathbf{T}_x^t \| < \frac{1}{100n}$, where $\| \Delta \mathbf{T}_x^t \|$ is the norm of $\Delta \mathbf{T}_x^t = \mathbf{T}^t - \mathbf{T}^{t-x}$.



(a) Standard deviation

(b) Mean value

Figure 2: Cost sensitivity analysis of the basic model

the computed mean values form a cloud and exhibit obvious randomness as well. Note that in these cases the learning process still converges. However, it may or may not converge to a conformity pattern. In some cases where $\underline{\pi} < c < \overline{\pi}$, we have $\sigma_p = 0$, which implies conformity. We also observe cases with $\sigma_p \neq 0$, which implies convergence to a set of diverse steady states. That is, $\lim_{t\to\infty} p_i^t = p_i^*$ for all i with $p_i^* \neq p_j^*$ for some different players i and j.

In these cases, the outcome of the updating process depends on the order in which players are chosen in each updating round, as well as the interaction cost and the initial conditions. That is, during each period, the state of a link ij depends on the strategies of the randomly selected players i and j. The resulting interaction network might be fully or near fully connected if the two randomly chosen players happen to always or almost always have strategies that lead to high payoffs. Otherwise the network would be sparsely connected and we have closed groups, where a player only interacts with members in her group and members in each group conform to the same strategy.

To summarize, when the interaction cost exceeds $\underline{\pi}$, the interaction and influence networks as well as the strategies still converge. But the questions of whether the steady state strategies conform, or whether the patterns of the interaction network are complete, and the final influence weight distribution depend highly on the initializations as well as which players are chosen for updating the interaction network in each round.

4 Introducing Persistent Players

Founded on the notion of persistency introduced in the learning model of Pan (2008), we extend the basic model to include persistent players. A player is called "persistent" if she does not change her initially assigned mixed strategy over time. It was already shown in Pan (2008) that persistent players have a significant influence on the outcome of a French-DeGroot naive learning process. We confirm this insight in our more elaborate dual network framework as well.

The set of persistent players is introduced as a subset $S \subset N$, where we assume that $1 \leqslant |S| < n$. A persistent player $s \in S$ is characterized by the property that $p_s^t = p_s^0$ for all $t \in \mathbb{N}$. However, we assume that every persistent player updates her interaction network L_i^t as well as her influence weights $\{T_{ij}^t \mid j \in N\}$ according to the dynamic updating process introduced in the previous section. As such, a persistent player does not update her strategy in every time period as do the other (non-persistent) players in the population.

Our main insight is that the introduction of persistent players into the population alters the outcome of the social learning process significantly. We first consider the introduction of persistent players in the population with a common persistent strategy denoted by $p_\alpha \in [0,1]$. This models a *uniform* group of persistent players in the population.

Theorem 3 *Consider a situation in which $c \leqslant \underline{\pi}$ and there exists a set of persistent players $S \subset N$ such that $|S| \geqslant 1$ and all $s \in S$ have a common persistent strategy given by $p_s^0 = p_s^t = p_\alpha \in [0,1]$ for all $t \in \mathbb{N}$. Then the social learning process converges to a fully connected interaction network, evenly distributed weights, and all non-persistent players' strategies converge to p_α, i.e., for all $\epsilon > 0$ there exists some $t^* > 0$ such that for all $t > t^*$ it holds that $\mathbf{G}_{ij}^t = 1$ as well as $|\mathbf{T}_{ij}^t - \frac{1}{n}| < \epsilon$ for all $i, j \in N$ and $|p_{\overline{s}_i}^t - p_\alpha| < \epsilon$ for all non-persistent players $\overline{s}_i \notin S$.*

A proof of Theorem 3 is given in Appendix C.

Theorem 3 states that persistent players possess a form of widespread influence in determining all players' strategic choices. Namely, the final strategy of all players equals the persistent players' initial strategy p_α. So, when $p_\alpha = 1$ or $p_\alpha = 0$, the social learning process converges to the Nash equilibrium outcomes (A, A) and (B, B), respectively. Otherwise the strategy vector of the whole society reaches a steady state given by $\{p_\alpha \mid i \in N\}$ that is not necessarily a Nash equilibrium. Also, in this case where the persistent players have uniform initial (persistent) strategies, the total number of them $|S|$ only affects the speed of convergence, not the final outcome.

If we have *diverse* persistent players, the social learning process converges to a convex combination of the persistent strategies adhered to by members of the group of persistent players. In this case, the final strategy of normal players are significantly influenced by persistent players' initial strategies as well.

Theorem 4 *Consider a situation in which $c \leqslant \underline{\pi}$ and the subset of persistent players $S \subset N$ is characterized by $2 \leqslant |S| \leqslant n-1$ such that there are $s_i, s_j \in S$ with $p_{s_i}^t = p_{s_i}^0$ and $p_{s_j}^t = p_{s_j}^0$ for all $t \in \mathbb{N}$, where $p_{s_i}^0 \neq p_{s_j}^0$. Then the social learning process converges* *to a fully connected interaction network, identical weights, and all non-persistent players' strategies converge to some $p_\beta \in [0,1]$, i.e., for all $\epsilon > 0$ there exists some $t^* > 0$ such that for all $t > t^*$ it holds that $\mathbf{G}^t_{ij} = 1$ as well as $|\mathbf{T}^t_{ik} - \mathbf{T}^t_{jk}| < \epsilon$ for all $i, j, k \in N$ and $|\mathbf{T}^t_{\overline{s}_i\overline{s}_j} - \mathbf{T}^t_{\overline{s}_i\overline{s}_k}| < \epsilon$, $|p^t_{\overline{s}_i} - p_\beta| < \epsilon$ for all non-persistent players $\overline{s}_i, \overline{s}_j, \overline{s}_k \notin S$.*

A proof of Theorem 4 is given in Appendix D. Note the difference in the final $\mathbf{T}^t$ with the two types of persistency. When we have diverse persistent players, the final influence weights are not even. Namely, we do not have $\mathbf{T}^*_{ij} = \dfrac{1}{n}$ for all i, j. Instead, we have identical weights, i.e., $\mathbf{T}^*_{ik} = \mathbf{T}^*_{jk}$. In other words, each column shows elements of the same value. Besides, all non-persistent players assign the same weight to other non-persistent players. The key here is that the diverse persistent players do not receive the same weight in the steady state, because they do not generate the same payoffs (since their strategies are different).

The assertion of Theorem 4 leaves open the issue where exactly the social learning process leads the non-persistent players. Proposition 4.1 below partially solves this issue and states upper and lower bounds on the mixed strategy to which the non-persistent players converge.

Proposition 4.1 *Consider the situation as stated in Theorem 4. If there are $m = |S|$ diverse persistent players such that $\#\{p^0_s \mid s \in S\} = m$, then in the social learning process, there exists some $T' > 0$ such that for all $t > T'$ it holds that $\underline{p}_s \leqslant p^t_i \leqslant \overline{p}_s$ for every player $i \in N$, where $\underline{p}_s = \min_{s\in S} p^0_s$ and $\overline{p}_s = \max_{s\in S} p^0_s$.*

The proof of Proposition 4.1 is relegated to Appendix E of this paper.

Corollary 4.2 below follows immediately from the French-DeGroot strategy updating rule $p^t_i = \sum_{j\in N} \mathbf{T}^t_{ij} p^{t-1}_j$, and that

$$\lim_{t\to\infty} \mathbf{T}^t_{ik} = \lim_{t\to\infty} \mathbf{T}^t_{jk} = \frac{x\left[(a+1)p^*_k - 1\right] + n\left(1 - p^*_k\right)}{(a+1)x^2 - 2nx + n^2},$$

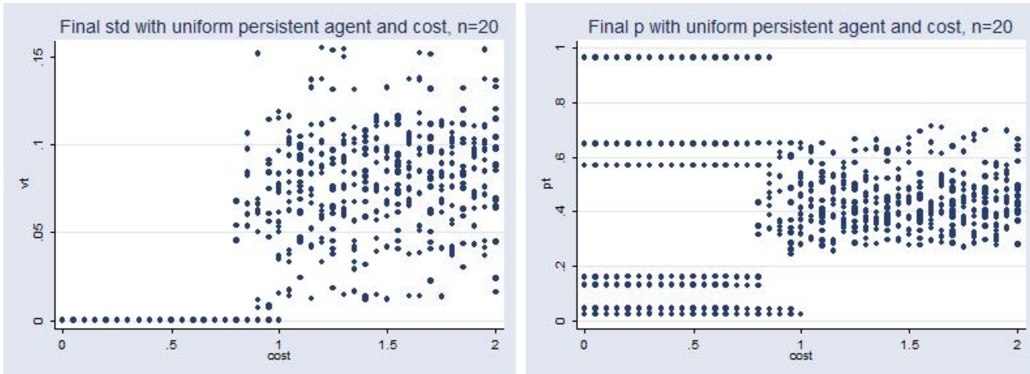
where $x = \sum_{i\in N} p^*_i = (n - |S|)p_\beta + \sum_{s\in S} p^0_s$.

Corollary 4.2 *Consider the situation with m diverse persistent players stated in Proposition 4.1, then all non-persistent players' strategies converge to p_β that satisfies the following properties*

$$\sum_{s\in S} (\mathbf{T}^*_{\cdot s} p^0_s) = p_\beta \sum_{s\in S} \mathbf{T}^*_{\cdot s}, \tag{16}$$

$$\frac{\mathbf{T}^*_{\cdot s_i}}{\mathbf{T}^*_{\cdot s_j}} = \frac{x\left[(a+1)p^0_{s_i} - 1\right] + n\left(1 - p^0_{s_i}\right)}{x\left[(a+1)p^0_{s_j} - 1\right] + n\left(1 - p^0_{s_j}\right)}, \text{ for all } s_i, s_j \in S, \tag{17}$$

where $\mathbf{T}^*$ is the limit influence matrix resulting from the social learning process.



(a) Standard deviation (b) Mean value

Figure 3: Cost sensitivity analysis with a single persistent player

The influence of persistent players is based on how non-persistent agents are influenced by them through the information-sharing network $\mathbf{T}^t$. Theorem 5 states that if costs are prohibitively high, no influence is exerted by any group of persistent players. The proof is very similar to that for Theorem 2 and is therefore omitted.

Theorem 5 *Consider any situation with persistent players. If $c \geqslant \overline{\pi}$, then the updating process converges to a situation in which there emerges an autarkic interaction network, the influence matrix is equal to the identity matrix, and all players choose their initial strategy p_i^0.*

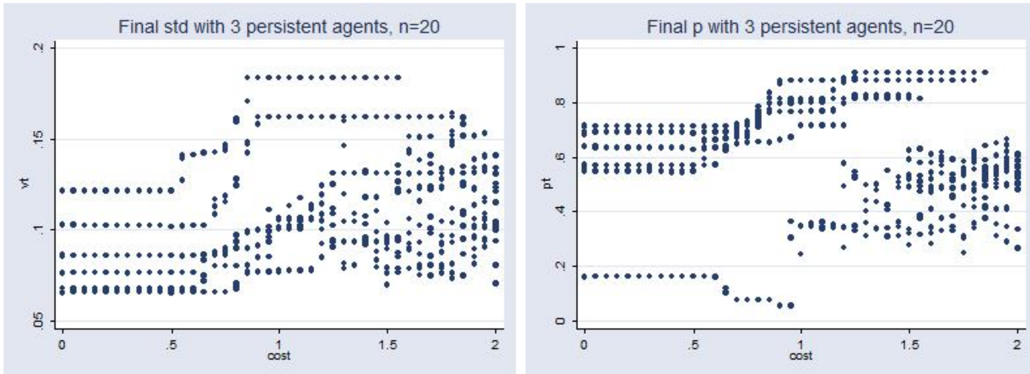
Sensitivity Analysis

We again use simulations to determine how the learning structure with persistent players behaves for medium interaction costs. We use the same settings as for the basic model, where $a = 2$ and c ranges from 0 to 2 in increments of 0.05. Persistent players are each assigned a random initial strategy. Each configuration is run 3 times.

Similar to the basic model, society size does not affect the final outcomes. Thus, for both uniform and diverse persistent players, we only show the cases where $n = 20$ with the x-axis showing value of cost and y-axis showing standard deviation and mean value of the final strategies. Figure 3 illustrates the case where we have a single persistent player which represents the uniform persistent model.

In Figure 4 we have 3 diverse persistent players, assigned 3 different initial strategies. Simulation results show that learning with different number of diverse persistent strategies exhibits similar outcome patterns.

Again, when cost is low enough, we observe conformism. Unlike the basic model, now the final strategy is fixed once initial strategies (of the persistent players) and influence weights (in case of diverse persistent players) are given. Hence, the mean value of p_i^t



(a) Standard deviation

(b) Mean value

Figure 4: Cost sensitivity analysis with 3 diverse persistent players

forms a line instead of a thin bar as in the basic model (Figure 2). The standard deviation is never 0 when we have diverse persistent players. Indeed, by Proposition 4.1, we know that p^* for non-persistent players is in the range of $(\underline{p}_s, \overline{p}_s)$. Thus, the standard deviation is at least $\sqrt{(\overline{p}_s - p^*)^2 + (\underline{p}_s - p^*)^2} > 0$. In this case, a constant standard deviation and mean value indicate conformism.

When cost exceeds $\underline{\pi}$, the outcomes are random and determined by the order of players chosen to update the interaction network, as well as the initial conditions. Same as the basic model, with higher cost we may have close groups with conformism in each group, and diverse strategies across groups. When cost is higher than $\overline{\pi}^0$, agents choose to stay isolated.

5 Concluding remarks

The social learning model discussed in this paper has a dynamic double-layer network structure. Namely, players play a coordination game with selected partners in an interaction network; on the other hand, they collect information about the resulting payoffs and the executed strategies in an influence network. Subsequently they update their interaction network, their influence weights, as well as the mixed strategy according to an extended variation of the naive social learning rule developed by French (1956) and DeGroot (1974).

It is a novel idea to separate the interaction network from the information collection framework, with the consideration that individuals first tend to collect abundant information and then process the information, before making a decision on a task or activity that is affiliated with chosen partners. Previous work with similar setting to either social sphere often assumed that the networks are exogenous and/or time-invariant; whereas in our framework both networks are endogenous and change over time. There is also a clear correlation between the two networks that ties them together in a sensible way.

In this study our focus is only on the cases where the interaction structure is completely open and freely determined. However, in realistic settings such open structures are rare; usually interaction is restricted by geographical and social distances. Many network structures have been categorized and investigated, in particular small-world and scale-free networks. In Pan (2009), different network structures are imposed on the learning population, with a network-distance combination learning rule in a similar dual network framework. Further research is necessary to delineate the various influences that affect boundedly rational or "naive" decision makers.

References

Aumann, R., and A. Brandenburger (1995): "Epistemic conditions for Nash equilibrium," *Econometrica*, 63.

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

Banerjee, A., and D. Fudenberg (2004): "Word-of-Mouth Learning," *Games and Economic Behavior*, 46, 1–22.

Blume, L., and D. Easley (2006): "If you're so smart, why aren't you rich? Belief selection in complete and incomplete markets," *Econometrica*, 74(4), 929–966.

Borgers, T. (1994): "Weak dominance and approximate common knowledge," *Journal of Economic Theory*, 64, 265–276.

Canning, D. (1995): "Learning and social equilibrium in large populations," *Learning and Rationality in Economics (A. Kirman and M. Salmon, Eds.)*.

Case, A., and L. Katz (1991): "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths," NBER Working Paper.

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

Dekel, E., and D. Fudenberg (1990): "Rational behavior with payoff uncertainty," *Journal of Economic Theory*, 52(2), 243–267.

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

Duflo, E., and E. Saez (2002): "Participation and Investment Decisions in a Retirement Plan: The Influence of Colleagues' Choices," *Journal of Public Economics*, 85, 121 – 148.

Ellison, G. (1993): "Learning, local interaction, and coordination," *Econometrica*, 61(5), 1047–1071.

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

——— (1995): "Word-of-mouth communication and social learning," *The Quarterly Journal of Economics*, 110(1), 93–125.

Foster, D., and P. Young (1990): "Stochastic evolutionary game dynamics.," *Theoretical Population Biology*, 38(2), 219–232.

French, J. R. P. (1956): "A formal theory of social power," *Psychological Review*, 63(3), 181–194.

Friedkin, N. E., and E. C. Johnsen (1990): "Social influence and opinions," *Journal of Mathematical Sociology*, 15, 193–206.

——— (1997): "Social positions in influence networks," *Social networks*, 19, 209–222.

Gabaix, X., D. Laibson, G. Moloche, and S. Weinberg (2006): "Costly information acquisition: Experimental analysis of a boundedly rational model," *American Economic Review*, Sept., 1043–1068.

Galeotti, A., S. Goyal, M. O. Jackson, F. Vega-Redondo, and L. Yariv (2009): "Network games," Forthcoming: Review of Economic Studies.

Golub, B., and M. O. Jackson (2007): "Naive learning in social networks: convergence, influence, and the wisdom of crowds," Revised: May 29th, 2007.

Griliches, G. (1957): "Hybrid corn: an exploration in the economics of technical change," *Econometrica*, 25, 501–522.

Harsanyi, J. C., and R. Selten (1988): *A general theory of equilibrium selection in games*. MIT Press.

Hegselmann, R., and U. Krause (2002): "Opinion dynamics and bounded confidence, models, analysis and simulation," *Journal of Artificial Societies and Social Simulation*, 5(3), http://jasss.soc.surrey.ac.uk/5/3/2.html.

Jackson, M., and A. Watts (2002): "On the Formation of Interaction Networks in Social Coordination Games," *Games and Economic Behavior*, 41, 265–291.

Jackson, M. O., and A. Wolinsky (1996): "A Strategic Model of Social and Economic Networks," *Journal of Economic Theory*, 71, 44–74.

KANDORI, M., G. MAILATH, AND R. ROB (1993): "Learning, mutation, and long run equilibria in games," *Econometrica*, 61(1), 29–56.

LORENZ, J. (2005): "A stabilization theorem for dynamics of continuous opinions," *Physica A: Statistical Mechanics and its Applications*, 355(1), 217–223.

PAN, Z. (2008): "Trust, influence, and convergence of behavior in social networks," Working paper.

——— (2009): "Neighborhood structure and learning in evolving social networks," Working paper.

ROTH, A., AND M. SOTOMAYOR (1990): *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Cambridge University Press, Cambridge, Massachusetts.

SENETA, E. (1981): *Non-negative matrices and Markov chains*. Springer-Verlag New York Inc.

SIMON, H. (1955): "A behavioral model of rational choice," *The Quarterly Journal of Economics*, 69(1), 99–118.

VEGA-REDONDO, F. (2007): *Complex Social Networks*. Cambridge University Press.

WEISBUCH, G., G. DEFFUANT, F. AMBLARD, AND J. P. NADAL (2002): "Meet, discuss, and segregate!," *Complexity*, 7(3), 55–63.

Appendices

A Proof of Proposition 3.1

Denote for each t,

$$\underline{p}^t = \min\{p_1^t, \ldots, p_n^t\}, \quad \overline{p}^t = \max\{p_1^t, \ldots, p_n^t\}.$$

Thus $p_i^t \in [\underline{p}^t, \overline{p}^t]$ for all i. Moreover, $p_i^{t+1} \in [\underline{p}^t, \overline{p}^t]$ for all i since p_i^{t+1} is a convex combination of p_i^t. That is, $\{\mathbf{p}^t\}_{t=0}^{\infty}$ is a sequence in a compact set $[\underline{p}^0, \overline{p}^0]^n$.

Recall that $\pi_{ij}^t = ap_i^t p_j^t + (1-p_i^t)(1-p_j^t)$. We have $\frac{\partial \pi_{ij}^t}{\partial p_j^t} = (a+1)p_i^t - 1$. So π_{ij}^t increases with p_j^t when $p_i^t \geqslant \frac{1}{a+1}$ and decreases with p_j^t when $p_i^t < \frac{1}{a+1}$.

1. $\frac{1}{a+1} \leqslant \underline{p}^0 \leqslant \overline{p}^0$.

   In this case, define $\pi_{min}^1 = a(\underline{p}^0)^2 + (1-\underline{p}^0)^2$, $\pi_{max}^1 = a(\overline{p}^0)^2 + (1-\overline{p}^0)^2$. Then for all i, t, $p_i^t \geqslant \underline{p} \geqslant \frac{1}{a+1}$, $\frac{\partial \pi_{ij}^t}{\partial p_i^t} \geqslant 0$. Thus, $\pi_{ij}^t \geqslant \pi_{min}^1$, for all i, j, t. Similarly, $\pi_{ij}^t \leqslant \pi_{max}^1$, for all i, j, t

2. $\underline{p}^0 \leqslant \overline{p}^0 < \frac{1}{a+1}$.

   In this case, define $\pi_{min}^2 = a(\overline{p}^0)^2 + (1-\overline{p}^0)^2$, $\pi_{max}^2 = a(\underline{p}^0)^2 + (1-\underline{p}^0)^2$. Then for all i, t, $p_i^t \leqslant \overline{p} < \frac{1}{a+1}$, $\frac{\partial \pi_{ij}^t}{\partial p_i^t} < 0$. Thus, $\pi_{ij}^t \geqslant \pi_{min}^2, \pi_{ij}^t \leqslant \pi_{max}^2$, for all i, j, t.

3. $\underline{p}^0 < \frac{1}{a+1} \leqslant \overline{p}^0$.

   In this case, define $\pi_{min}^3 = a\underline{p}^0\overline{p}^0 + (1-\underline{p}^0)(1-\overline{p}^0)$. For arbitrary i, j, t, without loss of generality, assume that $p_i^t \leqslant p_j^t$. We have 2 possible scenarios in this case.

   (i) $p_i^t < \frac{1}{a+1} \leqslant \overline{p}^0$. Then since $\frac{\partial \pi_{ij}^t}{\partial p_j^t} < 0$, $\overline{p}^0 \geqslant p_j^t$, we have $\pi_{ij}^t \geqslant ap_i^t\overline{p}^0 + (1 - p_i^t)(1-\overline{p}^0)$. And $\frac{\partial \pi_{ij}^t}{p_i^t} > 0$, $\underline{p}^0 \leqslant p_i^t$, so

   $$ap_i^t\overline{p}^0 + (1-p_i^t)(1-\overline{p}^0) \geqslant a\underline{p}^0\overline{p}^0 + (1-\underline{p}^0)(1-\overline{p}^0) = \pi_{min}^3,$$

   which implies that $\pi_{ij}^t \geqslant \pi_{min}^3$.
   Similarly, $\pi_{ij}^t \geqslant ap_i^t\underline{p}^0 + (1-p_i^t)(1-\underline{p}^0) \geqslant a(\underline{p}^0)^2 + (1-\underline{p}^0)^2 = \pi_{max}^2$

   (ii) $\frac{1}{a+1} \leqslant p_i^t \leqslant \overline{p}^0$. Then similar to the previous case,

   $$\pi_{ij}^t \geqslant ap_i^t\underline{p}^0 + (1-p_i^t)(1-\underline{p}^0) \geqslant a\underline{p}^0\overline{p}^0 + (1-\underline{p}^0)(1-\overline{p}^0) = \pi_{min}^3,$$

   $$\pi_{ij}^t \leqslant ap_i^t\overline{p}^0 + (1-p_i^t)(1-\overline{p}^0) \leqslant a(\overline{p}^0)^2 + (1-\overline{p}^0)^2 = \pi_{max}^1.$$

That is, in all cases, we can find a π^k_{min} such that $\pi^t_{ij} \geqslant \pi^k_{min}$ for all i, j, t. In particular, in the first 2 cases, $\pi^k_{min} = a\rho^2 + (1-\rho)^2 \geqslant \frac{a}{a+1}$, for $0 \leqslant \rho \leqslant 1$. So $\pi^t_{ij} \geqslant \frac{a}{a+1}$ in these 2 cases. Otherwise, $\pi^t_{ij} \geqslant a\underline{p}^0\overline{p}^0 + (1-\underline{p}^0)(1-\overline{p}^0)$.Define $\underline{\pi} = \min\{\frac{a}{a+1},\ a\underline{p}^0\overline{p}^0 + (1-\underline{p}^0)(1-\overline{p}^0)\}$, then it holds that $\pi^t_{ij} \geqslant \underline{\pi}$ for all i, j, t in all cases. Since $0 \leqslant \underline{p}^0 \leqslant \overline{p}^0 \leqslant 1$, $a\underline{p}^0\overline{p}^0 + (1-\underline{p}^0)(1-\overline{p}^0) \geqslant 0$. Also $\frac{a}{a+1} > 0$. Thus $\underline{\pi} \geqslant 0$; the equality holds only when $\underline{p}^0 = 0$ and $\overline{p}^0 = 1$.

Also, in all cases, we can find a π_{max} such that $\pi^t_{ij} \geqslant \pi_{max}$ for all i, j, t. We have 2 candidates for π_{max}, namely π^1_{max} and π^2_{max}. Therefore, define $\overline{\pi} = \max\{\pi^1_{max}, \pi^2_{max}\} > \frac{a}{a+1} > 0$, we have $\pi^t_{ij} \leqslant \pi_{max}$ for all i, j, t.

B Proof of Theorem 1

B.1 Preliminaries

We first state and prove a set of required intermediate results.

Lemma 1 *If $c \leqslant \underline{\pi}$, then with the updating process the interaction network g^t converges to be fully connected, i.e., $g^t \to g_N$ as $t \to \infty$.*

Proof. From Proposition 3.1 it follows that $\pi^t_{ij} \geqslant \underline{\pi} \geqslant c$ for all i, j, t. Every pair $i, j \in N$ is selected randomly, and, thus, in the long term each pair is selected with probability 1. So in this case, $\pi^t_{ij} \geqslant c$ is always true and any randomly chosen link ij at time t is always formed if $ij \notin g^{t-1}$ or stays formed if $ij \in g^{t-1}$. Thus, $g^t \to g_N$ in probability 1. Hence, there exists some $\bar{t} > 0$, such that for $t > \bar{t}$ we have $g^t = g^N$. ■

Lemma 2 *If $c \leqslant \underline{\pi}$ and there are $i, j \in N$ such that $p^0_i \neq p^0_j$, then with the updating process defined for the basic model, there exists $\hat{t} > 0$, s.t. for all $t > \hat{t}, 0 < p^t_k < 1$ for all $k \in N$.*

Proof. First, the conditions stated in the assertion indicate that the elements of $\mathbf{p}^0$ cannot be all 0s or all 1s, in which case we won't have $i, j \in N$ such that $p^0_i \neq p^0_j$.

As shown in the proof of Proposition 3.1, $\{\mathbf{p}^t\}^\infty_{t=0}$ is a sequence in a compact set $[\underline{p}^0, \overline{p}^0]^n$. So if $0 < p^0_i < 1$ for all i, then the assertion of Lemma 2 is true.

Next, consider the case where there exists $\gamma \in N$ such that p^0_γ is either 0 or 1. Then in order to have $p^t_\gamma = p^0_\gamma$, it must hold that $\mathbf{T}^t_{\gamma j} > 0$ if $p^0_\gamma = p^0_j$ and $\mathbf{T}^t_{\gamma j} = 0$ otherwise. Suppose that $\mathbf{T}^0$ satisfies that condition. Recall that

$$\mathbf{T}^t_{ij} = \frac{w^t_{ij}}{\sum_{k=1}^n w^t_{ik}}, \text{ for all } j \in N, t > 0,$$

$$\text{where } w^t_{ij} = \sum_{l \in L^t_i} \mathbf{G}^t_{lj} \pi^{t-1}_{lj}.$$

Consider two players γ, j such that $p^0_\gamma \neq p^0_j$. With Lemma 1, there exists $\hat{t}_\gamma$, s.t. for $t > \hat{t}_\gamma$, $j \in L(\gamma)^t$ and $j \in L(j)^t$. $\pi^t_{jj} = k(p^t_j)^2 + (1-p^t_j)^2 > 0$, which means that $w^t_{\gamma j} > 0$. Thus,

$\mathbf{T}^t_{\gamma j} > 0$ even though $p^0_\gamma \neq p^0_j$. In other words, player γ cannot remain her initial strategy of 0 or 1. We can repeat this process for all $\{\lambda \mid \lambda \in N, p^0_\lambda = 0 \text{ or } 1\}$. Thus $\exists \hat{t} = \max_\lambda \hat{t}_\lambda$, s.t. $\forall t > \hat{t}, 0 < p^t_k < 1$ for all $k \in N$. ∎

B.2 Proof of Theorem 1

With Lemma 1, we know that $\mathbf{G}^t$ converges to a fully connected network. As for the final strategy patterns, first consider the special cases where $\mathbf{p}^0 = (0, \ldots, 0)^{\mathrm{T}}$ and $\mathbf{p}^0 = (1, \ldots, 1)^{\mathrm{T}}$. Obviously, in these cases, strategy vector never changes via updating. So $p^* = p^0_i$.

Then, consider other cases. First, we show that $\sum_{k=1}^n |\mathbf{T}^t_{ik} - \mathbf{T}^t_{jk}| < 2 - \tilde{\tau}$, where $\tilde{\tau}$ is a positive number that has a lower bound.

By Lemma 2, we know that there exists $\hat{t} > 0$, s.t. for all $t > \hat{t}$, we have $0 < p^t_i < 1$ for all $i \in N$. Denote

$$\underline{p}^{\hat{t}} = \min_{i \in N} p^{\hat{t}}_i > 0, \quad \overline{p}^{\hat{t}} = \max_{i \in N} p^{\hat{t}}_i < 1.$$

Then it holds that for all $t > \hat{t} + 1$, for all i, $p^t_i \in [\underline{p}^{\hat{t}}, \overline{p}^{\hat{t}}]$. Define $\pi^{\hat{t}}_{min} = \min\{\frac{a}{a+1}, a\underline{p}^{\hat{t}}\overline{p}^{\hat{t}} + (1 - \underline{p}^{\hat{t}})(1 - \overline{p}^{\hat{t}})\}$. We have $\pi^{\hat{t}}_{min} > 0$ since $0 < \underline{p}^{\hat{t}} \leqslant \overline{p}^{\hat{t}} < 1$. Then we can mimic the proof of Proposition 3.1 and prove that that $\pi^t_{ij} \geqslant \pi^{\hat{t}}_{min}$, for all i, j and $t > \hat{t}$.[6] Also, $\pi^t_{ij} \leqslant a$ for all i, j, t. Then $\mathbf{T}^t_{ij} > \frac{\pi^{\hat{t}}_{min}}{an^2} > 0$.[7] Denote $\tilde{\tau} = \frac{\pi^{\hat{t}}_{min}}{an^2} > 0$. Then we have

$$\sum_{k=1}^n |\mathbf{T}^t_{ik} - \mathbf{T}^t_{jk}| < 2 - 2\tilde{\tau}.$$

Theorem 3.1 in Seneta (1981) states that,

$$\max_{i,j} |p^{t+1}_i - p^{t+1}_j| \leqslant \mu_t(\mathbf{T})\{\max_{i,j}] \mid p^t_i - p^t_j|\},$$

where $\mu_t(\mathbf{T}) = \frac{1}{2} \max_{i,j} \sum_{k=1}^n |\mathbf{T}^t_{ik} - \mathbf{T}^t_{jk}|$.
Since $\sum_{k=1}^n |\mathbf{T}^t_{ik} - \mathbf{T}^t_{jk}| < 2 - 2\tilde{\tau}$ for $t > \hat{t}$,

$$\mu_t(\mathbf{T}) = \frac{1}{2} \max_{i,j} \sum_{k=1}^n |\mathbf{T}^t_{ik} - \mathbf{T}^t_{jk}| < 1 - \tilde{\tau}, \text{ for } t > \hat{t}.$$

That is, $\mu_t(\mathbf{T}) \leqslant 1$ for all t and it is strictly less than and bounded away from 1 for $t > \hat{t}$. Since $\max_{i,j} |p^{t+1}_i - p^{t+1}_j| \leqslant \prod_{\tau=1}^t \mu_\tau(\mathbf{T}) \max_{i,j} |p^0_i - p^0_j|$, we have $\lim_{t\to\infty} |p^{t+1}_i - p^{t+1}_j| = 0$

As shown in the proof of Lemma 1, $\{\mathbf{p}^t\}_{t=0}^\infty$ is a sequence in a compact set $[\underline{p}^0, \overline{p}^0]^n$. Therefore, p^t_i is conforming to the same value p^*.

[6] Proof omitted here.

[7] This is induced from $w^t_{ij} \geq \pi^t_{ij} \geq \pi^{\hat{t}}_{min}$ and $\sum_k w^t_{ik} = \sum_k \sum_{l \in L^t_i} \mathbf{G}^t_{lk}\pi^{t-1}_{lk} \leq an^2$. The equality of the latter formula holds if and only if: $L^t_i = N$; and $\mathbf{G}^t_{lk} = 1$ for all $l \in L^t_i = N$; and $\pi^{t-1}_{lk} = a$ for all $l \in L^t_i = N$.

As for the influence matrix $\mathbf{T}^t$. Recall equation (10) that

$$\mathbf{T}_{ij}^t = \frac{w_{ij}^t}{\sum_{k=1}^n w_{ik}^t}, \text{ for all } j \in N, t > 0,$$

$$\text{where } w_{ij}^t = \sum_{l \in L_i^t} \mathbf{G}_{lj}^t \pi_{lj}^{t-1}.$$

Then if p_i^t converges to p_i^* for all i, the fully connected network results in such a $\mathbf{T}^t$ that

$$\lim_{t\to\infty} \mathbf{T}_{ij}^t = \mathbf{T}_{ij}^* = \frac{\sum_{k\in N}[ap_k^* p_j^* + (1-p_k^*)(1-p_j^*)]}{\sum_{l\in N}\sum_{k\in N}[ap_k^* p_l^* + (1-p_k^*)(1-p_l^*)]}, \forall i, j \in N.$$

The expression of $\mathbf{T}_{ij}^*$ can be simplified as

$$\mathbf{T}_{ij}^* = \frac{x[(a+1)p_j^* - 1] + n(1-p_j^*)}{(a+1)x^2 - 2nx + n^2}, \text{ where } x = \sum_{k\in N} p_k^*.$$

The simplified function shows that all the elements in each column j converges to the same value $\mathbf{T}_{\cdot j}^* = \frac{x[(a+1)p_j^*-1]+n(1-p_j^*)}{(a+1)x^2-2nx+n^2}$ which depends on p_j^*. Since $\lim_{t\to\infty} p_i^t = p^*$, for all $i \in N$, $\mathbf{T}_{\cdot i}^* = \mathbf{T}_{\cdot j}^*$ for all $i, j \in N$, which implies that $\lim_{t\to\infty} \mathbf{T}_{ij}^t = \frac{1}{n}$ for all i, j.

C Proof of Theorem 3

With Lemma 1, we know that $\mathbf{G}^t$ converges to a fully connected network. Denote $m = |S|$ as the total number of persistent players. Rearrange the players in such an order that number 1 to $n-m$ are normal players and the last m players are persistent. Essentially, the only step during updating that differs from the basic model is when players modify their strategies by taking weighted averages. Since persistent players do not change their strategies, their influence weight assignments do not affect their choices or the final outcome. Thus the strategy updating process can be rewritten as:

$$\mathbf{p}^t = \widetilde{\mathbf{T}}^t \mathbf{p}^{t-1},$$

where

$$\widetilde{\mathbf{T}}^t = \begin{pmatrix} \mathbf{T}_{11}^t & \dots & \mathbf{T}_{1j}^t & \dots & \dots & \mathbf{T}_{1n}^t \\ \vdots & & \vdots & & & \vdots \\ \mathbf{T}_{n-m,1}^t & \dots & \mathbf{T}_{n-m,j}^t & \dots & \dots & \mathbf{T}_{n-m,n}^t \\ 0 & \dots & 0 & \frac{1}{m} & \dots & \frac{1}{m} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & \frac{1}{m} & \dots & \frac{1}{m} \end{pmatrix}$$

In this case, define $\hat{\tau} = \min\{\frac{\pi_{min}^{\hat{t}}}{an^2}, \frac{1}{m}\}$, then $\mu_t(\widetilde{\mathbf{T}}) = \frac{1}{2}\max_{i,j}\sum_{s=1}^{n}|\widetilde{\mathbf{T}}_{is}^t - \widetilde{\mathbf{T}}_{js}^t| < 1 - \hat{\tau}$ still holds. Thus from the proof of Proposition 1, we have that $\mathbf{T}_{ij}^t \to \frac{1}{n}$, and that $p_i^t \to p^*$ for all i.

Next, we need to prove that $p_i^t \to p_\alpha$. Suppose that $p_i^t \to p^*$ and $p^* \neq p_\alpha$. Then for $\epsilon^* = \frac{1}{2}|p_\alpha - p^*|$, since for all $s \in S$, $p_s^t = p_\alpha$, for all t, $|p_\alpha - p^*| > \epsilon^*$ for all t. However, it is assumed that $p_i^t \to p^*$, which implies that $\forall \epsilon > 0, \exists t > 0$, s.t. $|p_i^t - p^*| < \epsilon$ for all i. Thus we have a contradiction, which means that $p^* = p_\alpha$. In other words, $\lim_{t\to\infty} p_i^t = p_\alpha$, for all i.

Similar to the basic model, with fully connected interaction network and conforming strategies, $\mathbf{T}_{ij}^t \to \frac{1}{n}$, for all i, j. The influence matrix exhibits equal distribution patterns at the stable state.

D Proof of Theorem 4

With Lemma 1, we know that $\mathbf{G}^t$ converges to a fully connected network. Similar to proof of Theorem 2, denote $m = |S|$ as the total number of persistent players and rearrange the players in such an order that number 1 to $n - m$ are normal players and the last m players are persistent.

Define

$$\tilde{\mathbf{T}}^t = \begin{pmatrix} \mathbf{T}_{11}^t & \dots & \mathbf{T}_{1j}^t & \dots & \mathbf{T}_{1n}^t \\ \vdots & & \vdots & & \vdots \\ \mathbf{T}_{n-m,1}^t & \dots & \mathbf{T}_{n-m,j}^t & \dots & \mathbf{T}_{n-m,n}^t \\ \underbrace{0,\dots,0}_{n-m-1} & & 0 & 1 & 0,\dots,\underbrace{0,\dots,0}_{m-1} \\ \vdots & & \vdots & & \vdots \\ 0 & \dots & 0 & \underbrace{0,\dots,0}_{m-1} & 1 \end{pmatrix}.$$

That is, a persistent player places weight 1 on herself and 0 on everybody else.

$$\mathbf{p}^t = \tilde{\mathbf{T}}^t \mathbf{p}^{t-1}. \tag{18}$$

Next, for normal players $i \leqslant n - m$, define

$$\mathbf{C}_i^t = \check{\mathbf{T}}^t \mathbf{C}_i^{t-1} = (\prod_{\theta=1}^{t} \check{\mathbf{T}}^\theta)\mathbf{C}_i^0, \tag{19}$$

where $\mathbf{C}_i^0$ is the i-th column of $\mathbf{T}^0$ and

$$\check{\mathbf{T}}^t = \begin{pmatrix} \mathbf{T}_{11}^t & \dots & \mathbf{T}_{1,n-m}^t & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ \mathbf{T}_{n,1}^t & \dots & \mathbf{T}_{n,n-m}^t & 0 & \dots & 0 \end{pmatrix}.$$

Then the j-th element of $\mathbf{C}_i^t$ indicate the weight that player j places on the player i during time t for $j \leqslant n - m$. It can be interpreted as player i's contribution to $\mathbf{p}^t$ received by j. Note that the last m elements of $\mathbf{C}_i^t$ is always 0, as it is indicated by equation (18).

For $\mathbf{C}_i^t = (C_{i1}^t, \ldots, C_{in}^t)^{\mathrm{T}}$, we have

$$\begin{aligned} |C_{ij}^t - C_{ik}^t| &= |\textstyle\sum_{l\in N}(\check{\mathbf{T}}_{jl}^t - \check{\mathbf{T}}_{kl}^t)C_l^{t-1}| \\ &\leq |\textstyle\sum_{l=1}^{n-m}(\mathbf{T}_{jl}^t - \mathbf{T}_{kl}^t)C_l^{t-1}| + \textstyle\sum_{l=n-m+1}^{n}|\mathbf{T}_{jl}^t - \mathbf{T}_{kl}^t|(\overline{C}_i^{t-1} - \underline{C}_i^{t-1}) \\ &\leqslant \mu_t(\mathbf{T})\{\max_{j,k}|C_{ij}^{t-1} - C_{ik}^{t-1}|\}. \end{aligned}$$

That is, $\max_{j,k}|C_{ij}^t - C_{ik}^t| \leqslant \mu_t(\mathbf{T})\{\max_{j,k}|C_{ij}^{t-1} - C_{ik}^{t-1}|\}$. The trick here is to use the original row-stochastic matrix $\mathbf{T}^t$. Then for the last m elements, for j, k such that $\mathbf{T}_{jl}^t - \mathbf{T}_{kl}^t \geqslant 0$ we replace C_{ij}^{t-1} with $\overline{C}_i^{t-1} = \max_{l\in N} C_{il}^{t-1}$ and C_{ik}^{t-1} with $\underline{C}_i^{t-1} = \min_{l\in N} C_{il}^{t-1}$. Vice versa. This allows us to utilize Seneta's theorem, which requires a row-stochastic matrix ($\check{\mathbf{T}}^t$ is not and $\mu_t(\tilde{\mathbf{T}}) = 1$).

From the proof of Theorem 1 we know that $|C_{ij}^t - C_{ik}^t| \rightarrow 0$. That is, for players $i, j, k \leqslant n - m$, we have $\lim_{t\to\infty}\mathbf{T}_{ik}^t = \lim_{t\to\infty}\mathbf{T}_{jk}^t = \mathbf{T}_{\cdot k}^*$. That is, the normal players converge to have same weight assignment on other normal players.

Then, for each persistent player s, define

$$\tilde{\mathbf{T}}_s^t = \begin{pmatrix} \mathbf{T}_{11}^t, \ldots, \mathbf{T}_{1,n-m}^t & 0, \ldots, 0 & \mathbf{T}_{1s}^t & 0, \ldots, 0 \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{T}_{s-1,1}^t, \ldots, \mathbf{T}_{s-1,n-m}^t & 0, \ldots, 0 & \mathbf{T}_{s-1,s}^t & 0, \ldots, 0 \\ 0, \ldots, 0 & 0, \ldots, 0 & 1 & 0, \ldots, 0 \\ \mathbf{T}_{s+1,1}^t, \ldots, \mathbf{T}_{s+1,n-m}^t & 0, \ldots, 0 & \mathbf{T}_{s+1,s}^t & 0, \ldots, 0 \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{T}_{n,1}^t, \ldots, \mathbf{T}_{n,n-m}^t & 0, \ldots, 0 & \mathbf{T}_{n,s}^t & 0, \ldots, 0 \end{pmatrix}.$$

That is, we set the s-th row of the (original) influence matrix to $\mathbf{1}_s$, where the s-th element is 1 and the rest are all 0. Also, the value of the i-th element in the s-th column is $\mathbf{T}_{is}^t$ (except $\mathbf{T}_{ss}^t$) instead of 0. Whereas the other $m - 1$ columns numbered after $n - m$ are all 0s.

Then we can now apply equation (19) to persistent players. Namely, $\mathbf{C}_s^t = \tilde{\mathbf{T}}_s^t\mathbf{C}_s^{t-1}$. Note here that for each s we have a specified matrix $\tilde{\mathbf{T}}_s^t$, unlike the case with normal players where we use the same matrix for them all.

Next, we use the same technic shown above and have that $|C_{sj}^t - C_{sk}^t| \rightarrow 0$, which means that normal players assign the same weight to the same persistent player as well.

Thus, for all $i, j \notin S$, $\mathbf{T}_{ik}^t = \mathbf{T}_{jk}^t$ for all k, which implies that $|p_i^t - p_i^{t-1}| = \sum_{k\in N}(\mathbf{T}_{ik}^t - \mathbf{T}_{jk}^t)p_k^{t-1} \rightarrow 0$. That is, $\mathbf{p}^t$ converges and all normal players' strategies conform:

$$\lim_{t\to\infty} p_i^t = \lim_{t\to\infty} p_j^t = p^* \text{ for all } i, j \notin S.$$

Also, $\mathbf{G}^t \to \mathbf{G}^N$, which means that $\mathbf{T}^t$ converges to $\mathbf{T}^*$ where

$$\lim_{t\to\infty} \mathbf{T}^t_{ik} = \lim_{t\to\infty} \mathbf{T}^t_{jk} = \frac{x[(a+1)p^*_k - 1] + n(1-p^*_k)}{(a+1)x^2 - 2nx + n^2},$$

where

$$x = \sum_{i\in N} p^*_i = (n - |S|)p_\beta + \sum_{s\in S} p^0_s, \text{ for all } i, j \notin S, \text{ for all } k \in N.$$

Since $\lim_{t\to\infty} p^t_i = \lim_{t\to\infty} p^t_j$ for $i \notin S$, actually $\lim_{t\to\infty} \mathbf{T}^t_{ij} = \lim_{t\to\infty} \mathbf{T}^t_{ik}$ as well. That is, elements in $(\mathbf{T}^*_{ij})_{(n-m)\times(n-m)}$ all have the same value.

Finally, we let $p_\beta = p^*$, which completes the proof of Theorem 4.

E Proof of Proposition 4.1

Suppose that $\underline{p}^t < \underline{p}_s$ for all t, that is, the lowest value of persistent players' initial strategies is never a lower bound. From the proof of Theorem 1 we know that there exists $\hat{t} > 0$ such that for $t > \hat{t}$, for all i, j $\mathbf{T}^t_{ij} > \tau$, where $\tau = \frac{\pi^{\hat{t}}_{min}}{an^2} \in (0, 1]$ is bounded away from 0. Then we have that:

$$\begin{aligned} \underline{p}^t &\geq (1 - m\tau)\underline{p}^{t-1} + \tau \textstyle\sum_{s\in S} p^0_s \\ &\geq (1 - m\tau)[(1 - m\tau)\underline{p}^{t-2} + \tau \textstyle\sum_{s\in S} p^0_s] + \tau \textstyle\sum_{s\in S} p^0_s \\ &\vdots \\ &\geq (1 - m\tau)^{t-\hat{t}}\underline{p}^{\hat{t}} + \tau(\textstyle\sum_{s\in S} p^0_s) \textstyle\sum_{\eta=0}^{t-\hat{t}}(1 - m\tau)^\eta. \end{aligned}$$

We have $\lim_{t\to\infty}(1 - m\tau)^{t-\hat{t}}\underline{p}^{\hat{t}} + \tau(\sum_{s\in S} p^0_s) \sum_{\eta=0}^{t-\hat{t}}(1 - m\tau)^\eta = \frac{1}{m}\sum_{s\in S} p^0_s > \underline{p}_s$,[8] which is a contradiction to the assumption that $\underline{p}^t < \underline{p}_s$. Thus there exists $t' > 0$ such that $\underline{p}_s$ is a lower bound of p^t_i for all $t > t'$.

Similarly, suppose that $\overline{p}_s$ is never an upper bound of p^t_i. Then we have that $\overline{p}^t \geq (1-m\tau)^{t-\hat{t}}\overline{p}^{\hat{t}} + \tau(\sum_{s\in S} p^0_s) \sum_{\eta=0}^{t-\hat{t}}(1-m\tau)^\eta \to \frac{1}{m}\sum_{s\in S} p^0_s < \overline{p}_s$, which is a contradiction to the assumption. Thus $\overline{p}_s$ is an upper bound of p^t_i for all $t > t'$.

Note that both $\underline{p}^t \geq \frac{1}{m}\sum_{s\in S} p^0_s$ and $\overline{p}^t \leq \frac{1}{m}\sum_{s\in S} p^0_s$ are induced from the invalid assumptions that $\underline{p}^t < \underline{p}_s$ and $\overline{p}^t > \overline{p}_s$. Therefore, they cannot be used to conclude that $p^t_i \to \frac{1}{m}\sum_{s\in S} p^0_s$. In fact, we see from Corollary 4.2 that it is not true.

[8]Here $\frac{1}{m}\sum_{s\in S} p^0_s \neq \underline{p}_s$ because we have diverse persistent players, which implies that there exists at least one persistent player whose strategy does not equal to $\underline{p}_s$. This reasoning also applies to the statement that $\frac{1}{m}\sum_{s\in S} p^0_s < \overline{p}_s$ below.