Estimation of peer pressure in dynamic homogeneous social networks

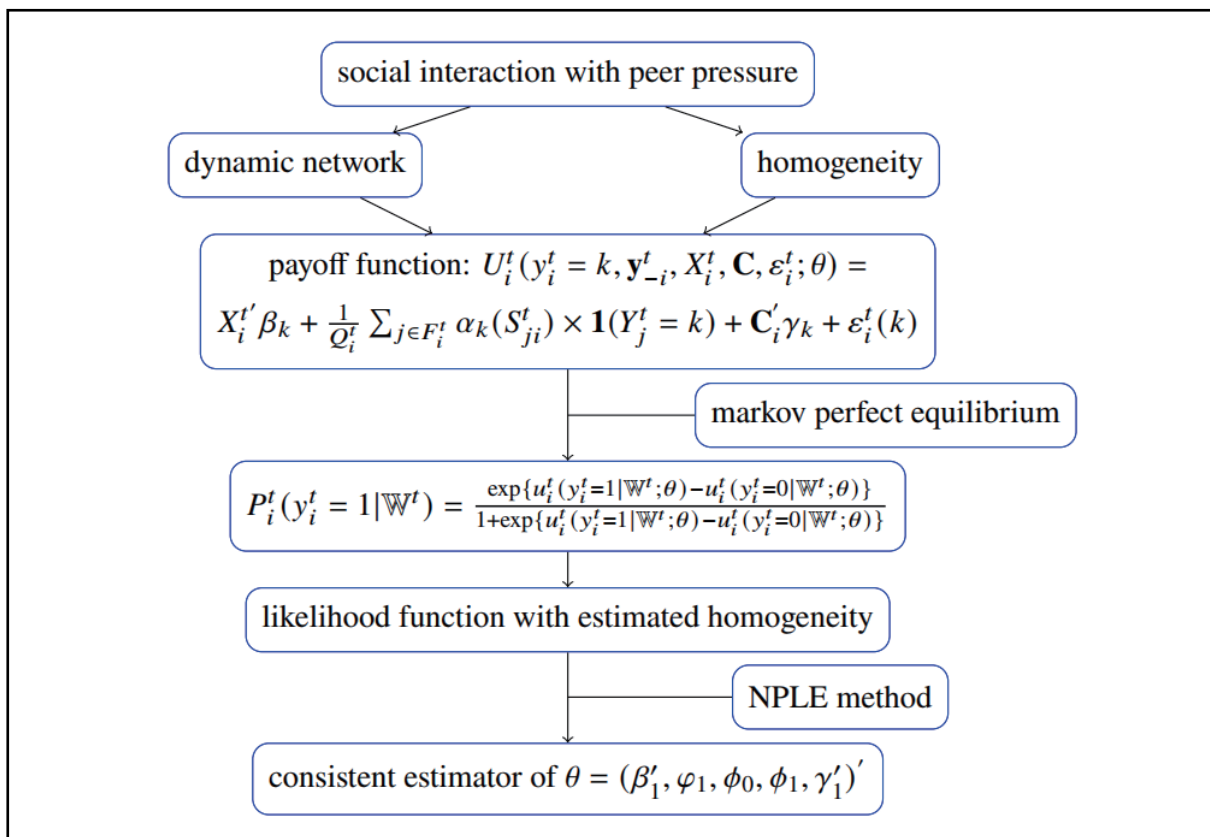Jie Liu[1], Pengyi Wang[1], Jiayang Zhao[1] ✉, and Yu Dong[2]

[1]International Institute of Finance, School of Management, University of Science and Technology of China, Hefei 230026, China;
[2]School of Economics and Management, Anhui University of Science and Technology, Huainan 232001, China

✉Correspondence: Jiayang Zhao, E-mail: zhaojy19@mail.ustc.edu.cn

Graphical abstract

A novel network game model is proposed to quantify social interaction with peer pressure in dynamic networks. New algorithms are designed to identify homogeneity. The NPLE method is introduced to estimate the model parameters.

Public summary

- This paper introduces the Generalized Constant Peer Effect (GCPE) model, a novel network game model that quantifies social interactions within dynamic networks.

- The proposed model can efficiently mitigate estimation inaccuracies related to peer pressure by integrating homogeneity.

- We design innovative algorithms to accurately identify homogeneity. Then we apply the Nested Pseudo-Likelihood Estimation (NPLE) method to obtain consistent estimators of parameters.

Estimation of peer pressure in dynamic homogeneous social networks

Jie Liu[1], Pengyi Wang[1], Jiayang Zhao[1] ✉, and Yu Dong[2]

[1]International Institute of Finance, School of Management, University of Science and Technology of China, Hefei 230026, China;
[2]School of Economics and Management, Anhui University of Science and Technology, Huainan 232001, China

✉Correspondence: Jiayang Zhao, E-mail: zhaojy19@mail.ustc.edu.cn

Abstract: Social interaction with peer pressure is widely studied in social network analysis. Game theory can be utilized to model dynamic social interaction and one class of game network models assumes that peopleos decision payoff functions hinge on individual covariates and the choices of their friends. However, peer pressure would be misidentified and induce a non-negligible bias when incomplete covariates are involved in the game model. For this reason, we develop a generalized constant peer effects model based on homogeneity structure in dynamic social networks. The new model can effectively avoid bias through homogeneity pursuit and can be applied to a wider range of scenarios. To estimate peer pressure in the model, we first present two algorithms based on the initialize expand merge method and the polynomial-time two-stage method to estimate homogeneity parameters. Then we apply the nested pseudo-likelihood method and obtain consistent estimators of peer pressure. Simulation evaluations show that our proposed methodology can achieve desirable and effective results in terms of the community misclassification rate and parameter estimation error. We also illustrate the advantages of our model in the empirical analysis when compared with a benchmark model.

Keywords: dynamic network; game theory; homogeneity; peer pressure; social interaction

CLC number: O212.1   **Document code:** A
2020 Mathematics Subject Classification: 62F10

1　Introduction

People's behavior based on individual interactions in social networks has been widely studied by researchers[1–5]. The individual's behavior needs to take into account the effect of peer pressure between connected nodes in dynamic social networks, i.e., the influence of partners in decision-making. For example, when conducting surveys on students' smoking behaviors, people usually only know whether the student has smoked during a particular time period. Peer pressure refers to the influence of the behaviors of friends on these students' smoking behavior. Game theory is commonly utilized to describe social interactions in networks[6–8]. Through observation, we can discover that those who make the same decision often have similar characteristics, which is homogeneity. The homogeneity structure is quite common in social network analysis[9–12]. Peer pressure is a type of social influence that is often confused with homogeneity when discussing individual behavior, see Refs. [12, 13]. Ref. [14] studies a practical problem that takes homogeneity into account. Therefore, to obtain more accurate estimates of peer pressure, it is necessary to consider the impact of homogeneity on peer pressure. However, the above study based on game theory on peer pressure does not consider the potential confusion between social influence and homogeneity, which may lead to bias in the estimation of social influence. There is some relevant work on

distinguishing homogeneity from social influence, see Ref. [13, 15, 16]. It is necessary to consider the impact of homogeneity in game models to obtain more accurate peer pressure. Therefore, combining the two is meaningful research.

Game theory can be used to statistically model dynamic social interaction with peer pressure, and one class of game network models usually assumes that people's decision payoff functions depend on individual covariates and their friends' decisions and then derives the corresponding conditional choice probability equations based on the payoff functions. In Refs. [17, 18], they discussed a dynamic game problem based on Markov games, but the focus of their research was not peer pressure. However, the deduction approach for dynamic games can be applied to other models. Ref. [19] proposed the concept of "individual behavior influenced by others". Ref. [20] proposed the concept of "local interaction" for the first time. Based on the above work, Ref. [7] proposed a static network game model with incomplete information to study the estimation of peer pressure. Ref. [8] expanded upon the model in Ref. [7], allowing for peer pressure to depend on social influence, but the model remains limited to static scenarios. Their studies did not take into account the effect of homogeneity on individual decision-making, which might lead to bias in the estimation of peer pressure confounded with homogeneity. In addition, they implied that the observed covariates

were complete. However, this is difficult in practice, which can also lead to estimation bias.

Homogeneity pursuit is an effective way to avoid possible bias caused by incomplete covariates, where incomplete covariates imply the existence of node attributes related to the formation of social relationships or to node behavior that have not been observed[16]. People tend to interact with people who are similar to them, which implies that homogeneity plays a crucial role in the formation of social network structures and it indeed has an influence on individual behaviors[11]. Ref. [13] introduced a continuous potential space setting to estimate homogeneity to correctly identify social influence. Ref. [15] used the potential community setting to estimate homogeneity. Both [13] and [15] noted that reducing the estimation bias of social influence can be effectively achieved by considering homogeneity. Based on the above work, Ref. [16] proposed a linear model for explaining social interactions in homogeneous social networks, and when the estimation of homogeneity meets certain conditions, the estimator of the model converges. However, people's behavior is discrete in many scenarios, and the property of consistent convergence of corresponding parameters no longer holds in our game network model. Then, the homogeneity estimation method is no longer applicable in our model's corresponding network, and we need to redesign the estimation method for homogeneity. Moreover, Ref. [16] assumed that different individuals have the same influence on each other, which does not align well with reality.

This study introduces a new game model with individual homogeneity to characterize social interactions with peer pressure in dynamic networks. The new model is called the generalized constant peer effect (GCPE) model. To construct such a dynamic network model, we adopt a Markov perfect equilibrium as the game to obtain an equilibrium solution and then give the conditional choice probability equation. The GCPE model assumes that discrete individual decisions are determined by the covariates unrelated to homogeneity, peer pressure, homogeneity, and random disturbance terms. It can be thought of as the extension of the game model in Ref. [8] and the model in Ref. [16]. We propose a simple, efficient, and accurate estimation procedure to estimate homogeneity and peer pressure. Specifically, we first utilize two algorithms based on the initialize expand merge method and the polynomial-time two-stage method to estimate homogeneity parameters because homogeneity cannot be directly observed. Then we apply the nested pseudo-likelihood (NPLE) method to efficiently obtain consistent estimators of peer pressure.

The remainder of this paper is organized as follows. In Section 2, we present a generalized constant peer effect (GCPE) model. Two algorithms are given to estimate homogeneity in Section 3. In Section 4, the NPLE method and the corresponding theoretical properties are represented. In Section 5, we demonstrate the estimation results of the homogeneity algorithm and validate the theoretical results of the NPLE estimator. Section 6 compares the results of our proposed model with a benchmark model in applications to real data. Conclusions are provided in Section 7. The proof details of theorems and some simulation results are given in the Appendix.

2   Generalized constant peer effects model

We consider a Markov game model in dynamic social networks. The dynamic network is a series of random graphs consisting of nodes with directed edges. Let $i \in \mathcal{I} = \{1, \cdots, n\}$ denote each node. The relationship between nodes at time $t \in \mathcal{T} = \{1, \cdots, m\}$ is represented by the adjacency matrix $\mathbb{G}^t \in \mathbb{R}^{n \times n}$, where the (i, j) entry $g_{ij}^t = 1$ if individual i thinks j is his/her best friend at time t, and $g_{ij}^t = 0$ otherwise. we follow the convention to let $g_{ii}^t = 0$ for $i \in \mathcal{I}$ and $t \in \mathcal{T}$. Denoting $F_i^t = \{j \in \mathcal{I} : g_{ij}^t = 1\}$ to be the group of i's best friends.

The latent community partition of each network is denoted by $D = (D_1, \cdots, D_M)$, where $M < n$ is the number of communities. Under the potential community setting, we refer to a class of multi-graph community identification methods, where a dynamic network is a kind of multi-graph network, such as Ref. [21, 22]. This type of research considers community labeling to be global, not local. Specifically, MIT Reality Mining data in Ref. [21] is a dynamic network dataset, but community identification is performed by identifying a global community label for it as a whole rather than a local community label for each moment of the network individually. Therefore, let C_i be the homogeneity parameter of node i that does not change over time and it takes the following value:

$$C_i = \sum_{j=1}^{M} d_j \mathbf{1}(i \in D_j),$$

where d_j is the community-specific parameter in community D_j.

In dynamic social networks, we record the social behavior of node i at time $t \in (-\infty, +\infty)$ as $Y_i^t \in \{0, 1\}$. Y_i^t can be the student's decision to smoke or not as mentioned above. It is worth noting that our model introduced later also allows for multiple values of Y_i^t without loss of generality. Then the utility function of node i is given by

$$U_i^t(y_i^t = k, \mathbf{y}_{-i}^t, X_i^t, \mathbf{C}, \varepsilon_i^t; \theta) =$$
$$X_i^{t'}\beta_k + \frac{1}{Q_i^t}\sum_{j \in F_i^t} \alpha_k(S_{ji}^t) \times \mathbf{1}(Y_j^t = k) + \mathbf{C}_i'\gamma_k + \varepsilon_i^t(k).$$

Here, $\mathbf{y}_{-i}^t$ is the behavior vector of all nodes except node i. $X_i^t \in \mathbb{R}^d$ is the covariate vector of individual i that is not relevant to homogeneity parameter $\mathbf{C}_i$. Q_i^t is the total number of i's best friends at time t, we set $Q_i^t = 1$ if $Q_i^t = 0$ to ensure that $1/Q_i^t$ makes sense. S_{ji}^t is a social-influence measure of node j on node i at time t; $\mathbf{1}(A)$ is an indicator function of a set A; $\varepsilon_i^t(k) \in \mathbb{R}$ are unobserved action-dependent utility shocks, where $k \in \{0, 1\}$. Let $\varepsilon_i^t = \{\varepsilon_i^t(0), \varepsilon_i^t(1)\}$. Moreover, β_k is the coefficient vectors to be estimated for X_i^t, γ_k is the coefficients vectors to be estimated for $\mathbf{C}_i$, and $\alpha_k(\cdot)$ are unknown functions that measure the peer pressure on player i from his/her friend j when they make the same decision for $k \in \{0, 1\}$.

Under the condition of maximizing the total expected discounted payoff, we use a Markov game to derive the payoff

function with a homogeneous structure as the following conditional choice probability equation, i.e., the GCPE model:

$$P_i^t(y_i^t = 1|\mathbb{W}^t) = \frac{\exp\{u_i^t(y_i^t = 1|\mathbb{W}^t;\theta) - u_i^t(y_i^t = 0|\mathbb{W}^t;\theta)\}}{1 + \exp\{u_i^t(y_i^t = 1|\mathbb{W}^t;\theta) - u_i^t(y_i^t = 0|\mathbb{W}^t;\theta)\}}, \quad (1)$$

where

$$u_i^t(y_i^t = k|\mathbb{W}^t;\theta) = X^{t'}\beta_k + \frac{1}{Q_i^t}\sum_{j\in F_i^t}\alpha_k(S_{ji}^t) \times P(y_j^t = k|\mathbb{W}^t) + C_i'\gamma_k,$$

$\mathbb{W}^t = (X^t, C, G^t)$, $C = (C_1, \cdots, C_n)$ is the set of homogeneity of nodes in the network, θ is the parameter to be estimated and will be introduced later in detail, and the total expected discounted payoff function will also be introduced later. The GCPE model describes how individual i's decision-making is influenced by self factors and his/her social friends. The detailed derivation is described in Section 2.1.

2.1 Model derivation

Markov perfect equilibrium

In dynamic games, the adjacency matrix $\mathbb{G}^t$ will change over time, and the state variables (X_i^t, ε_i^t) will change as players make decisions. Therefore, their change is described by an exogenous probability. To solve for Markov perfect equilibrium solutions, we further make the assumption of conditional independence of their changes:

Assumption 2.1. The exogenous probability function satisfies conditional independence, that is,

$$p(X^{t+1}, y_i^t, \varepsilon_i^{t+1}, \mathbb{G}^{t+1}|X^t, y_{-i}^t, y^{t-1}, C, \mathbb{G}^t, \varepsilon_i^t) =$$
$$p(X^{t+1}, y_i^t|X^t, y_{-i}^t, y^{t-1},)p(\mathbb{G}^{t+1}|C)p(\varepsilon_i^{t+1}). \quad (2)$$

This assumption is quite common for Markov games, such as Refs. [17, 18]. Player i pursues the highest expected return, usually demonstrating long-term consideration in the Markov game. Let $\eta \in (0, 1)$ denote the discount factor. Given the state (X_i^t, ε_i^t) at time t, player i will make a series of decisions to maximize its total expected discounted payoff, i.e.

$$\max_{y_i^t, y_i^{t+1}, \cdots} \mathbb{E}\left\{\sum_{\tau=t}^{\infty} \eta^{\tau-t} U_i^\tau(y_i^\tau, y_{-i}^\tau, X_i^\tau, C, \varepsilon_i^\tau; \theta)|(X^t, C, \mathbb{G}^t, \varepsilon_i^t)\right\},$$

where the expectation is taken over the state evolution given in Eq. (2).

Because the state transition is stationary, we adopt the Markov perfect equilibrium as the equilibrium concept, and thus the time index t can be omitted from the notation. Let $P_i(y_i|\mathbb{W})$ be the conditional selection probability of player i choosing decision $y_i \in \{0, 1\}$ in state $\mathbb{W} = (X, C, \mathbb{G})$. Given $P_j(y_j|\mathbb{W})$, $\forall j \neq i$, the current expected payoff of player i choosing decision y_i in state $\mathbb{W}$ is

$$u_i(y_i = k|\mathbb{W};\theta) =$$
$$\sum_{y_{-i}\in\{0,1\}^{n-1}} \{[\prod_{y_j\in y_{-i}} P_j(y_j|\mathbb{W})]U_i(y_i, y_{-i}, X, C, \theta)\} =$$
$$X'\beta_k + \frac{1}{Q_i}\sum_{j\in F_i}\alpha_k(S_{ji}) \times P(y_j = k|\mathbb{W}) + C_i'\gamma_k.$$

Bellman optimality. Let $V_i(x)$ be the conditional expectation function of player i in state x. For any $i \in \mathcal{I}$ and $\mathbb{W} \in \mathcal{W}$,

we have

$$V_i(\mathbb{W}) = \sum_{y_i\in Y} P_i(y_i|\mathbb{W})[u_i(y_i|\mathbb{W};\theta) + e_i^P(y_i, \mathbb{W})] +$$
$$\eta \int_{\mathcal{W}} V_i(\mathbb{W}')f_{\mathcal{W}}^P(\mathbb{W}'|\mathbb{W})d\mathbb{W}',$$

where $f_{\mathcal{W}}^P(\mathbb{W}'|\mathbb{W})$ is a conditional density function of state $\mathbb{W}'$ when given $\mathbb{W}$, and $e_i^P(y_i, \mathbb{W})$ is a function related to the distribution of $\varepsilon_i(k)$.

Conditional choice probability equation. Let $v_i(y_i|\mathbb{W})$ be the conditional choice-specific expected value function of player i, then

$$v_i(y_i|\mathbb{W}) = u_i(y_i|\mathbb{W};\theta) + \eta \int_{\mathcal{W}} V_i(\mathbb{W}')f_{\mathcal{W}}^P(\mathbb{W}'|\mathbb{W}, y_i)d\mathbb{W}'. \quad (3)$$

where $f_{\mathcal{W}}^P(\mathbb{W}'|\mathbb{W}, y_i)$ is a conditional density function.

When the private shocks ε_i are observed, player i makes decision $y_i = k$ if and only if $k \in \arg\max_{k\in\{0,1\}}\{v_i(y_i = k|\mathbb{W}) + \varepsilon_i(y_i = k)\}$. Therefore, we have

$$P_i(y_i = k|\mathbb{W}) = \Pr[\varepsilon_i|v_i(y_i = k|\mathbb{W}) + \varepsilon_i(y_i = k) >$$
$$\max_{j\in\{0,1\}-\{k\}}\{v_i(y_i = j|\mathbb{W}) + \varepsilon_i(y_i = j)\}]$$

Conditional choice probability equation

To derive the above equation more specifically, we give the following assumption about the distribution of private shocks $\varepsilon_i^t(k)$.

Assumption 2.2. The error terms $\varepsilon_i^t(k)$, $i \in \{1, \cdots, n\}$, $k \in \{0, 1\}$ are independent and identically distributed (i.i.d.) across both players and actions at time t. Furthermore, the error term follows the extreme value distribution (Type Ⅰ) with density

$$f(t) = \exp(-t)\exp(-\exp(-t)).$$

Assumption 2.2 is often used in game network models[7, 8, 17, 18]. Note that the Gaussian distribution and other distributions can also be applied to derive the conditional choice probability equation, but the corresponding formulation will be more sophisticated than that of the extreme value distribution (Type I). Under Assumption 2.2, we can derive a simple form of $P_i(y_i = k|\mathbb{W})$, that is

$$P_i(y_i = k|\mathbb{W}) = \frac{\exp\{v_i(y_i = k|\mathbb{W})\}}{\sum_{k=1}^{2}\exp\{v_i(y_i = k|\mathbb{W})\}},$$

i.e.,

$$P_i(y_i = 1|\mathbb{W}) = \frac{\exp\{v_i(y_i = 1|\mathbb{W}) - v_i(y_i = 0|\mathbb{W})\}}{1 + \exp\{v_i(y_i = 1|\mathbb{W}) - v_i(y_i = 0|\mathbb{W})\}}. \quad (4)$$

The following assumption is imposed to simplify Eq. (4).

Assumption 2.3. $f_{\mathcal{W}}^P(\mathbb{W}'|\mathbb{W}, y_i)$ is independent of y_i.

Since homogeneity replaces the attribute variables X in state $\mathbb{W}$, which are relevant to the players' decisions, thus Assumption 2.3 is reasonable. Under Assumption 2.3, we further simplify Eq. (4) and continue to use the symbol t. The final solution yields our proposed GCPE model:

$$P_i^t(y_i^t = 1|\mathbb{W}^t) = \frac{\exp\{u_i^t(y_i^t = 1|\mathbb{W}^t;\theta) - u_i^t(y_i^t = 0|\mathbb{W}^t;\theta)\}}{1 + \exp\{u_i^t(y_i^t = 1|\mathbb{W}^t;\theta) - u_i^t(y_i^t = 0|\mathbb{W}^t;\theta)\}}.$$

2.2 Model linearization

For the sake of tractability, we linearize $\alpha(\cdot)$ for empirical analysis, see Refs. [7, 8], let $\alpha_0(s) = \varphi_0 + \varphi_1 \times s$ and $\alpha_1(s) = \phi_0 + \phi_1 \times s$. For model identification, we set $\beta_0 = 0$, $\gamma_0 = 0$ and $\varphi_0 = 0$. By following Ref. [8], we can treat $P_i^t(y_i^t = 1|\mathbb{W}^t)$ as a known object. Let $T_i^t = \ln P_i^t(y_i^t = 1|\mathbb{W}^t) - \ln(1 - P_i^t(y_i^t = 1|\mathbb{W}^t))$, we can obtain that

$$T_i^t = X_i^{t'}\beta_1 + \varphi_1 \sum_{j \in F_i^t} \frac{S_{ji}^t(P_j^t(y_j^t = 1|\mathbb{W}^t) - 1)}{Q_i^t} +$$

$$\phi_0 \sum_{j \in F_i^t} \frac{P_j^t(y_j^t = 1|\mathbb{W}^t)}{Q_i^t} + \phi_1 \sum_{j \in F_i^t} \frac{S_{ji}^t P_j^t(y_j^t = 1|\mathbb{W}^t)}{Q_i^t} + \mathbf{C}_i^{'}\gamma_1, \quad (5)$$

where φ_1 represents the peer pressure effect exerted by node i's friends who are different from node i's decision when the social influence of nodes in the network is different. ϕ_0 quantifies the peer pressure effect of node i's friends, whose decisions are the same as node i when the social influence of all nodes is equal. ϕ_1 refers to the peer pressure effect of node i's friends whose decisions are the same as those of node i when nodal social influence remains fixed in the network.

Our target parameters can be denoted by $\theta = (\beta_1', \varphi_1, \phi_0, \phi_1, \gamma_1')' \in \Theta$, where Θ is the parameter space. Let $\theta_0 \in \Theta$ be the true parameter for the game.

3 Estimation of homogeneity

As stated in Ref. [16], we cannot directly observe the homogeneity parameter $\mathbf{C}_i$ of player i. Thus we turn to utilizing the attribute covariates of player i and its social relationships to approximately estimate the homogeneity parameter $\mathbf{C}_i$. $\hat{\mathbf{C}}$ denotes the estimated homogeneity vector of $\mathbf{C}$.

We follow the idea of Ref. [23] and define a weighted adjacency matrix W as follows:

$$W_{ij} = \alpha \frac{\sum_{t \in \mathcal{T}} \mathbb{G}_{ij}^t}{m} + (1 - \alpha)f(X_i, X_j),$$

where $f(X_i, X_j)$ is a function that takes values between 0 and 1, $0 \le \alpha \le 1$ is a constant. The number of edges in the weighted directed network is too large, thus we choose to delete the edges with smaller weights to ensure that our proposed algorithms can work. Specifically, we set $W_{ij} = 0$ if $W_{ij} < \omega_{0.625}$ where $\omega_{0.625}$ is the 0.625 quantile of W. Let $F_i^W = \{j \in \mathcal{I} : W_{ij} > 0\}$, and $F^W = \{F_1^W, \cdots, F_n^W\}$.

For community identification of homogeneity parameters, we first present a revised algorithm of the IEM algorithm in Ref. [24]. We use Sim_{ij}^{IEMC} to replace Sim_{ij}^{IEM}, where Sim_{ij}^{IEMC} follows the pattern of the directed network formation corresponding to the GCPE model. Then, the IEM algorithm pulls j into the community of i when $j = \text{argmax}_{j \neq i} Sim_{ij}^{\text{IEMC}}$, but we instead pull i into j's community, which is more consistent with our model and can effectively improve algorithm efficiency. In summary, our proposed Algorithm 1 (initialize-expand-merge-change method) can be applied to directed weight networks.

Although Algorithm 1 is able to estimate the communities of heterogeneous parameters very well, there is a more

---

Algorithm 1. Initialize-expand-merge-change (IEMC) method.

Input: Weighted adjacency matrix W, Node's best friend group F^W

Output: Community assignment σ^0, the number of communities k.

1: Select a random unlabeled node i and label node i.
2: Find a node $j \in \mathcal{I}$ to maximize weighted similarity based on transit node Sim_{ij}^{IEMC} and a labeled node j, where

$$Sim_{ij}^{\text{IEMC}} = \begin{cases} 0, \text{ if } W_{ij} = 0; \\ \dfrac{W_{ij}}{\sum_{k \in \mathcal{I}} W_{ik} + \sum_{k \in \mathcal{I}} W_{kj} - W_{ij}}, \text{ if } W_{ij} > 0 \text{ and } W_{ik} * W_{kj} = 0; \\ \sum_{h \in F_k^W} \dfrac{\sum_{h \in F_k^W} W_{kh} * W_{ik} + W_{kj}}{|F_k^W| \sum_{h \in \mathcal{I}} W_{ih} + \sum_{h \in \mathcal{I}} W_{hj} - W_{ij}}, \text{ if } k \in \{k : W_{ik} * W_{kj} > 0\}. \end{cases}$$

3: If node i and node j are not in the same community, merge the communities of node i and node j, then let $i = j$ and back to Step 2.
4: If node i and node j are in the same community, and there are unlabeled nodes in $\mathcal{I}$, back to Step 1.
5: Let C be the current community collection and calculate $Q(C)$, where

$$Q(C) = \frac{1}{2\sum_{ij} W_{ij}} \sum_{ij} (W_{ij} - \frac{\sum_k W_{ik} + \sum_k W_{kj}}{2\sum_{ij} W_{ij}})\mathbf{I}(C_i = C_j).$$

6: **for** $i = 1$ to $|C| - 1$ and $j = i + 1$ to $|C|$ **do**
7: Merge community i and community j and record this new community collection as C_{-ij}, then calculate $Q_{-ij} = Q(C_{-ij})$
8: **end for**
9: if $\max_{C_{-ij}}(Q_{-ij}) > Q(C)$, let $C = \text{argmax}_{C_{-ij}}(Q_{-ij})$ and back to Step 5; else return $\sigma^0(i) = C(i)$ for all $i \in \mathcal{I}$ and $k = |C|$.

---

charming algorithm in[25] that bounds the probability of misclassification, which is of order $e^{-O(n)}$. However, their algorithm is only suitable for undirected networks, thus it cannot be directly used in our GCPE model. To further improve the accuracy of community identification, we design a modified cross-validation version of Algorithm 1 and introduce Algorithm 2. Compared with the Algorithm 1, Algorithm 2 simplifies and changes the initial screening Steps 2–4 in Algorithm 1 at the cost of heavy computation. Based on the idea of Ref. [25], we present our Algorithm 2.

Algorithm 2. A refinement scheme for community detection.

Input: Weighted adjacency matrix W, number of communities k, initial community detection method σ^0 (i.e., Algorithm 1).

Output: Community assignment $\tilde{\sigma}$.

1: **for** $i = 1$ to n **do**
2: Apply σ^0 on W_{-i} for all $j \neq i$, obtain $\sigma_i^0(j)$ and $\sigma_i^0(i) = 0$, where W_{-i} denotes the $(n-1) \times (n-1)$ submatrix of W with its ith row and ith column removed.
3: Define $\tilde{\sigma}_i : [n] \to [k]$ by setting $\tilde{\sigma}_i(j) = \sigma_i^0(j)$ for all $i \neq j$ where $[n] = \{1, \cdots, n\}$, and

$$\tilde{\sigma}_i(i) = \text{argmax}_{l \in [k]} \sum_{\sigma_i^0(j) = l} W_{ij}.$$

4: **end for**
5: Define $\tilde{\sigma}(i) = \tilde{\sigma}_i(i)$. For $i = 2, \cdots, n$, define $\tilde{\sigma}(i) = \text{argmax}_{l \in [k]} |\{j : \tilde{\sigma}_1(j) = l\} \cap \{j : \tilde{\sigma}_i(j) = \tilde{\sigma}_i(i)\}|$.

Assumption 3.1. As $n \longrightarrow \infty$, $Pr(\hat{\mathbf{C}} \neq \mathbf{C}) \longrightarrow 0$.

Assumption 3.1 is reasonable for homogeneity estimation. Refs. [26, 27] pointed out that it is possible to consistently recover the communities for networks based on latent community setting when very mild regularity conditions are satisfied. The homogeneity structure in our proposed GCPE model is the latent community setting. In addition, after summarizing many relevant studies, Ref. [16] mentions that there are many community detection methods that satisfy this assumption. One such method is described in Ref. [25], and it provides us with a framework for constructing 2. In Section 5, we show numerically that the community misclassification rate of the two algorithms in the GCPE model tends to 0 when $n \to \infty$.

4 Estimation of peer pressure

Once we obtain the estimated homogeneity vector $\hat{\mathbf{C}}$, we can plug $\hat{\mathbf{C}}$ into the linearization form of the GCPE model in Eq. (5). Let $\tilde{\mathbb{W}}^t = (X, \hat{\mathbf{C}}, \mathbb{G})$ and $\tilde{T}_i^t = \ln P_i^t(y_i^t=1|\tilde{\mathbb{W}}^t) - \ln(1 - P_i^t(y_i^t=1|\tilde{\mathbb{W}}^t))$, we have

$$Z_i^t = \left(X_i^{t\prime}, \sum_{j\in F_i^t} \frac{S_{ji}^t(P_j^t(y_j^t=1|\mathbb{W}^t)-1)}{Q_i^t}, \sum_{j\in F_i^t} \frac{P_j^t(y_j^t=1|\mathbb{W}^t)}{Q_i^t}, \sum_{j\in F_i^t} \frac{S_{ji}^t P_j^t(y_j^t=1|\mathbb{W}^t)}{Q_i^t}, \mathbf{C}_i^\prime\right)',$$

$$\tilde{Z}_i^t = \left(X_i^{t\prime}, \sum_{j\in F_i^t} \frac{S_{ji}^t(P_j^t(y_j^t=1|\tilde{\mathbb{W}}^t)-1)}{Q_i^t}, \sum_{j\in F_i^t} \frac{P_j^t(y_j^t=1|\tilde{\mathbb{W}}^t)}{Q_i^t}, \sum_{j\in F_i^t} \frac{S_{ji}^t P_j^t(y_j^t=1|\tilde{\mathbb{W}}^t)}{Q_i^t}, \hat{\mathbf{C}}_i^\prime\right)',$$

$\mathfrak{P}^{t*}(\mathbb{W}^t) = (P_1^t(y_1^{t=1}|\mathbb{W}^t), \cdots, P_n^t(y_1^{t=1}|\mathbb{W}^t))'$ and $\mathfrak{P}^t = (p_1^t, \cdots, p_n^t) \in [0,1]^n$ be the Markov perfect equilibrium choice probability solution and a general probability distribution at time t. For any $\mathfrak{P}^t \in [0,1]^n$, we have

$$Z_i^t(\mathfrak{P}^t) = \left(X_i^{t\prime}, \sum_{j\in F_i^t} \frac{S_{ji}^t(p_j^t-1)}{Q_i^t}, \sum_{j\in F_i^t} \frac{p_j^t}{Q_i^t}, \sum_{j\in F_i^t} \frac{S_{ji}^t p_j^t}{Q_i^t}, \mathbf{C}_i^\prime\right)' \in \mathbf{R}^{d+3+k},$$

$$\tilde{Z}_i^t(\mathfrak{P}^t) = \left(X_i^{t\prime}, \sum_{j\in F_i^t} \frac{S_{ji}^t(p_j^t-1)}{Q_i^t}, \sum_{j\in F_i^t} \frac{p_j^t}{Q_i^t}, \sum_{j\in F_i^t} \frac{S_{ji}^t p_j^t}{Q_i^t}, \hat{\mathbf{C}}_i^\prime\right)' \in \mathbf{R}^{d+3+k}.$$

Then, the corresponding conditional choice probability equation is

$$\Gamma_i^t(\mathfrak{P}^t, \theta; \mathbb{W}^t) = \frac{\exp(Z_i^t(\mathfrak{P}^t)\theta)}{1 + \exp(Z_i^t(\mathfrak{P}^t)\theta)}.$$

We denote $\mathfrak{P}^t(\theta; \mathbb{W}^t)$ as the solution of the equation: $\Gamma_i^t(\mathfrak{P}^t, \theta; \mathbb{W}^t) = p_i^t$ for all $t \in \mathcal{T}$ and $i \in \mathcal{I}$. Note that we cannot directly determine $\mathfrak{P}^{t*}(\mathbb{W}^t) = \mathfrak{P}^t(\theta; \mathbb{W}^t)$, because it is difficult to guarantee the uniqueness of the structural parameters θ, that is, a solution may correspond to multiple structural parameters θ. Nevertheless, we can determine $\mathfrak{P}^{t*}(\mathbb{W}^t) = \mathfrak{P}^t(\theta_0; \mathbb{W}^t)$, where $\mathfrak{P}^t(\theta_0; \mathbb{W}^t)$ is unique. we further define a pseudo log-likelihood function:

$$\tilde{T}_i^t = X_i^{t\prime}\tilde{\beta}_1 + \tilde{\varphi}_1 \sum_{j\in F_i^t} \frac{S_{ji}^t(P_j^t(y_j^t=1|\tilde{\mathbb{W}}^t)-1)}{Q_i^t} +$$

$$\tilde{\phi}_0 \sum_{j\in F_i^t} \frac{P_j^t(y_j^t=1|\tilde{\mathbb{W}}^t)}{Q_i^t} + \tilde{\phi}_1 \sum_{j\in F_i^t} \frac{S_{ji}^t P_j^t(y_j^t=1|\tilde{\mathbb{W}}^t)}{Q_i^t} + \hat{\mathbf{C}}_i^\prime \tilde{\gamma}_1.$$

Our target parameters can be denoted by $\tilde{\theta} = (\tilde{\beta}_1', \tilde{\varphi}_1, \tilde{\phi}_0, \tilde{\phi}_1, \tilde{\gamma}_1')' \in \Theta$. Let $\tilde{\theta}_0$ be the true parameter given $\hat{\mathbf{C}}$. However, the estimation procedure of $\tilde{\theta}$ by directly using the maximum likelihood method may suffer from a huge computational burden caused by the game. Therefore, we introduce the nested pseudo-likelihood estimation (NPLE) method to address the problem of excessive computation. We further provide theoretical properties of the NPLE method.

4.1 NPLE method

In a large dynamic game network, it is obviously more difficult to calculate the equilibrium solution. The NPLE method can effectively alleviate the problem of excessive computational burden, and it is essentially an iterative algorithm consisting of a series of logit estimates that is easy to implement. In the later subsection we prove the convergence property of the NPLE method, which guarantees the accuracy of the method. To describe the NPLE method in our model, let

$$\tilde{L}_n(\tilde{\theta}, \mathfrak{P}) = \frac{1}{n} \sum_{t=1}^m \sum_{i=1}^n \left(Y_i^t \ln \Gamma_i^t(\mathfrak{P}^t, \tilde{\theta}; \tilde{\mathbb{W}}^t) + (1-Y_i^t)\ln(1-\Gamma_i^t(\mathfrak{P}^t, \tilde{\theta}; \tilde{\mathbb{W}}^t))\right),$$

where $\mathfrak{P} = (\mathfrak{P}^1, \cdots, \mathfrak{P}^m)$. Note that we are using the estimated homogeneity parameter $\hat{\mathbf{C}}$ instead of the true homogeneity parameter $\mathbf{C}$. If we want to obtain the pseudo log-likelihood function corresponding to the true parameter θ_0, we need to use the true homogeneity to construct the function. Therefore, we define the limiting form of this pseudo log-likelihood function as follows:

$$L(\theta, \mathfrak{P}) = \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^m \sum_{i=1}^n \mathbf{E}(Y_i^t \ln \Gamma_i^t(\mathfrak{P}^t, \theta; \mathbb{W}^t) + (1-Y_i^t)\ln(1-\Gamma_i^t(\mathfrak{P}^t, \theta; \mathbb{W}^t))),$$

Based on the pseudo log-likelihood function, we apply the NPLE method to estimate the parameters of interest, i.e.,

$$\tilde{\theta}_{\mathrm{NPLE}} = \arg\max_{\tilde{\theta}\in\Theta} \tilde{L}_n(\tilde{\theta}, \mathfrak{P}).$$

where $\mathfrak{P} = \Gamma(\mathfrak{P}, \tilde{\theta}; \tilde{\mathbb{W}})$.

First, we follow Ref. [7] and set the initial value of $\mathfrak{P}^{(0)}$ to be $(0, \cdots, 0) \in [0,1]^{n\times m}$. Second, we conduct an iterative procedure over the two steps described below:

Step 1: Given $\mathfrak{P}^{(j-1)}$, let

$$\tilde{\theta}^{(j)} = \arg\max_{\theta\in\Theta} \tilde{L}_n(\theta, \mathfrak{P}^{(j-1)}).$$

Step 2: Given $\tilde{\theta}^{(j)}$, let

$$\mathfrak{P}^{(j)} = \Gamma(\mathfrak{P}^{(j-1)}, \tilde{\theta}^{(j)}; \tilde{\mathbb{W}}),$$

where

$$\Gamma(\mathfrak{P}, \theta; \tilde{\mathbb{W}}) = (\Gamma_1^1(\mathfrak{P}, \theta; \tilde{\mathbb{W}}^1), \cdots, \Gamma_n^1(\mathfrak{P}, \theta; \tilde{\mathbb{W}}^1), \cdots,$$
$$\Gamma_1^m(\mathfrak{P}, \theta; \tilde{\mathbb{W}}^m), \cdots, \Gamma_n^m(\mathfrak{P}, \theta; \tilde{\mathbb{W}}^m)),$$

$$\tilde{\mathbb{W}} = (\tilde{\mathbb{W}}^1, \cdots, \tilde{\mathbb{W}}^m).$$

When $\left\| \tilde{\theta}^{(k)} - \tilde{\theta}^{(k-1)} \right\| < c$ or $\|\mathfrak{P}^{(k)} - \mathfrak{P}^{(k-1)}\| < c$ (c is small enough, e.g., 10^{-6}), we stop iterating. Then we have $\tilde{\theta}_{\mathrm{NPLE}} = \tilde{\theta}^{(k)}$.

Although there is an excellent condition to ensure the local convergence property of the NPLE algorithm in Ref. [28], it is still difficult to verify the uniqueness of the equilibrium solution in our model. However, this does not mean that we cannot use the NPLE method, because extensive experience through application shows that the NPLE method usually converges to the same solution, see Ref. [17].

4.2 Asymptotic analysis

We first define some notations of the network $\mathbb{G}^t$: let $N_{(i,h)}^t = \{j \in \mathcal{I} : (\mathbb{G}^{t^k})_{ji} \geqslant 1$ for $\forall k \leqslant h\}$ and $\mathbb{G}_{(i,h)}^t$ is a $N_{(i,h)}^t \times N_{(i,h)}^t$ submatrix of $\mathbb{G}^t$ for $i \in \mathcal{I}$, $\forall h \in \mathbb{N}$ and $t \in \mathcal{T}$. Then, we introduce the assumptions required by the theorem.

Assumption 4.1. Let $\sup_{s \in \mathcal{S}_0} |\alpha_0(s) + \alpha_1(s)| < 4$.

Assumption 4.1 ensures that Eq. (1) allows a unique equilibrium solution. This assumption is a sufficient assumption for equilibrium uniqueness in Bayesian games, see Ref. [8]. We found it is also a useful assumption for equilibrium uniqueness in our Markov games, so we still use it as a sufficient assumption. Similar assumptions can also be found in Ref. [7]. Under this assumption and previous assumptions, we can further obtain the following lemma.

Lemma 4.1. Under Assumptions 2.1, 2.2, 2.3 and 4.1 in Theorem 4.1, the games exist a unique pure strategy Markov perfect equilibria.

As mentioned in Ref. [8], Lemma 4.1 is an important theorem for statistical inference. We need it to avoid the problem of multiple equilibria. One can refer to Ref. [8] for more details.

Assumption 4.2. $E(Z_i^t Z_i^{t'})$ and $E(\tilde{Z}_i^t \tilde{Z}_i^{t'})$ are full rank.

Assumption 4.3. The parameter θ_0 is the only solution to the following equation

$$\theta = \arg\max_{c \in \Theta} L(c, \mathfrak{P}(\theta, \mathbb{W}))$$

Assumption 4.2 is used for model identification. Similar assumptions can be found in Refs. [7, 8]. As stated above and in Ref. [8], assumption 4.3 is an essential assumption for our theoretical results. Although it is easy to verify that θ_0 is a solution of the above equation, we cannot prove that θ_0 is the unique solution of the above equation. If the above equation has multiple solutions, then each solution is a convergence point of the NPLE method. Therefore, we need this assumption to guarantee that the NPLE method convergence point is unique. The NPLE method can still be used when the assumption is not satisfied. As mentioned in Ref. [17], we can use multiple NPLE initial points to verify whether the fixed point is unique, if not, consider the point that maximizes the pseudo

log-likelihood function among the multiple fixed points as the solution of the NPLE method.

Assumption 4.4. The parameter space Θ is compact and the support $\mathcal{S}_{xC}$ and $\mathcal{S}_{x\hat{C}}$ is bounded.

Assumption 4.5. There exists a positive integer constant c_0 such that

$$\max_{i \in \mathcal{I}, t \in \mathcal{T}} \sum_{j=1}^{n} g_{ij}^t \leqslant c_0$$

is always true.

Assumption 4.6. For any given $h \in \mathbb{N}$, the probability distribution of $\mathbb{G}_{(i,h)}^t$ converges to a limiting distribution as n tends to infinity for all $t \in \mathcal{T}$ and $i \in \mathcal{I}$, and if $N_{(i,h)}^t \cap N_{(j,h)}^t = \emptyset$, then $\mathbb{G}_{(i,h)}^t$ and $\mathbb{G}_{(j,h)}^t$ are independent. Moreover, the payoff covariates X_i^t are i.i.d. across players when the exogenous random network at time t is given.

Assumption 4.4 guarantees that our pseudo log-likelihood function is uniformly bounded, since it ensures that our probability selection function is uniformly bounded and far from zero, see Refs. [7, 8]. Assumption 4.5 aims to reduce the dependencies between the various players, please refer to Ref. [7] and the network decaying dependence condition (NDD condition) mentioned therein. It is also true that this is the case in some actual datasets. Assumption 4.6 is similar to that in Refs. [7, 8]. With regard to the assumption of payoff covariates X_i^t in assumption 4.6, we do not directly use the attribute variables of the players in the model (e.g., height, age, gender, etc.). Similarly, as described in Ref. [8], Under certain conditions, we can slightly relax this assumption.

Assumption 4.7. For any two players i and j, the central measure function S_{ji}^t between them needs to satisfy

$$S_{ji}^t = S(\mathbb{G}_{(i,q)}^t, \mathbb{G}_{(j,q)}^t, \{\hat{\mathbf{C}}_k \in \hat{\mathbf{C}} : k \in N_{(i,q)}^t \cup N_{(j,q)}^t\}),$$

where $S(\cdot, \cdot, \cdot)$ is a bounded function and q is a constant.

Assumption 4.7 is utilized to limit the dependence caused by social influence. In our model, we still want to keep the social influence setting in Ref. [8], but it will break some of the dependencies that were reduced in Refs. [7, 8]. Thus we need Assumption 4.7 to ensure that the dependencies in our model are not so strong.

Theorem 4.1. Under Assumptions 2.1, 2.2, 2.3 3.1, 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, and $m \ll n$, $h = \ln n/(4 \ln c)$, then,

$$\tilde{\theta}_{\mathrm{NPLE}} \xrightarrow{P} \theta_0$$

holds.

Theorem 4.1 guarantees the convergence property of the NPLE estimator, but it still has extensible aspects. It has been found that center metrics may not meet Assumption 4.7, we can modify this type of center metric to fit our model and avoid possible bias. Therefore, we make Assumption 4.8.

Assumption 4.8. For any two players i and j, the central measure function(S_{ji}^t) between them needs to satisfy

$$S_{ji}^{'t} = S(\mathbb{G}_{(i,\max(h,q))}^t, \mathbb{G}_{(j,\max(h,q))}^t, \{\hat{\mathbf{C}}_k : k \in N_{(i,\max(h,q))}^t \cup N_{(j,\max(h,q))}^t\}),$$

where $S(\cdot, \cdot, \cdot)$ is a bounded function, q is a constant and $\lim_{n \to \infty} S_{ji}^{'t} = \lim_{n \to \infty} S_{ji}^t$.

Theorem 4.2. Under assumptions 2.1, 2.2, 2.3 3.1, 4.1, 4.2,

4.3, 4.4, 4.5, 4.6, 4.8, and $m \ll n$, we let $h = \ln n / (4 \ln c)$, then,

$$\tilde{\theta}'_{NPLE} \xrightarrow{P} \theta_0$$

holds. Note that the h that we choose in this theorem is not unique and we use S'^t_{ji} instead of S^t_{ji} to estimate the parameters.

5 Monte Carlo experiments

In this section, we evaluate the performance of our proposed community identification algorithms for estimating homogeneity, and verify whether the NPLE estimation of peer pressure conforms to theoretical results. We first elaborate on the generation process of simulated data, and then present the simulation results.

5.1 Experimental setup

For a dynamic homogeneous network, we consider time $m = 2$ and assume that the number of directed edges of each node in the network has an upper bound, i.e., we set a minimum of 1 and a maximum of 10 directed edges for each node. Similar to Ref. [16], we further consider 4 potential communities for individual homogeneity parameters. The probability of having a directed edge between two nodes in the same community is 0.75, while the probability of having a directed edge between two nodes belonging to different communities is 0.25. The simulation results of other choices of link probabilities for the edges can be found in the Appendix. Furthermore, we generate four different groups of two-dimensional random attribute covariates for estimating the homogeneity parameter, that is, $D_1 : (N(2, 1/4), N(2, 1/4))$, $D_2 : (N(2, 1/4), N(-2, 1/4))$, $D_3 : (N(-2, 1/4), N(2, 1/4))$, $D_4 : (N(-2, 1/4), N(-2, 1/4))$, and set $f(X_i, X_j) = d(X_i, X_j)$, where $d(\cdot, \cdot)$ is Euclid distance.

Regarding social influence, we apply the KB centrality measure in Ref. [8], which is defined as follows:

$$S^t_i = \sum_{k=1}^{\infty} \sum_{j=1}^{n} \lambda^k ((\mathbb{G}^t)^k)_{ji},$$
$$S^t_{ji} = S^t_j - S^t_i.$$

It can be seen that using KB centrality as the true result in the simulation will lead to deviation in the estimation procedure, because it will violate the precondition of the theoretical result. We make the following changes to the KB centrality to conform to Theorem 4.2 for estimation.

$$S^t_i = \sum_{k=1}^{\infty} \sum_{j=1}^{n} \lambda^k ((\mathbb{G}^t_{(i,h)})^k)_{ji}, S^t_{ji} = S^t_j - S^t_i,$$

where $\mathbb{G}^t_{(i,h)}$ is a $N^t_{(i,h)} \times N^t_{(i,h)}$ submatrix of $\mathbb{G}^t$ for $i \in \mathcal{I}$, $h \in \mathbb{N}$ and $t \in \mathcal{T}$ and $N^t_{(i,h)} = \{j \in \mathcal{I} : (\mathbb{G}^{t^k})_{ji} \geqslant 1 \ for \ \forall k \leqslant h\}$. It is simple to verify that this modified KB centrality satisfies condition 2 of Theorem 4.2.

For nodal covariates $X_t \in \mathbb{R}$, we follow Ref. [7, 8] and let $X^t_i = S^t_i$, the corresponding coefficient is set to be $\beta_1 = -1$. With regard to homogeneity parameters, we let $\gamma'_1 = (-2, -1, 1, 2)'$. Then, we give each node a probability $1/n^{1.5}$ of community misidentification in each simulation to verify our theoretical results. The reason why we do not use

Fig. 1. Average community misclassification rate for Algorithm 1, which is based on the average of 50 replications.

our algorithm is because of the computational cost and we have difficulty in having enough time to perform simulations with large samples. Finally, the estimation results are obtained by simulating 100 replications.

5.2 Experimental results

We summarize the experimental results from the following three aspects. First, we evaluate the performance of the two proposed algorithms for estimating homogeneity parameters. Second, we analyze the effect of the NPLE estimation method of peer pressure parameters from the perspective of *mse*

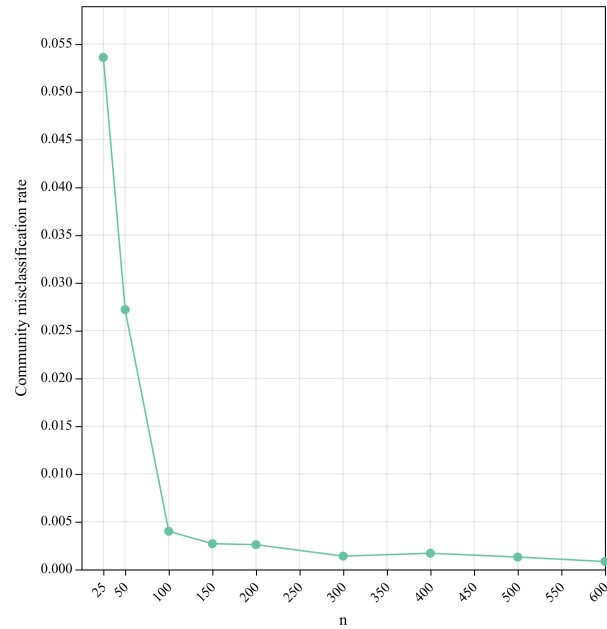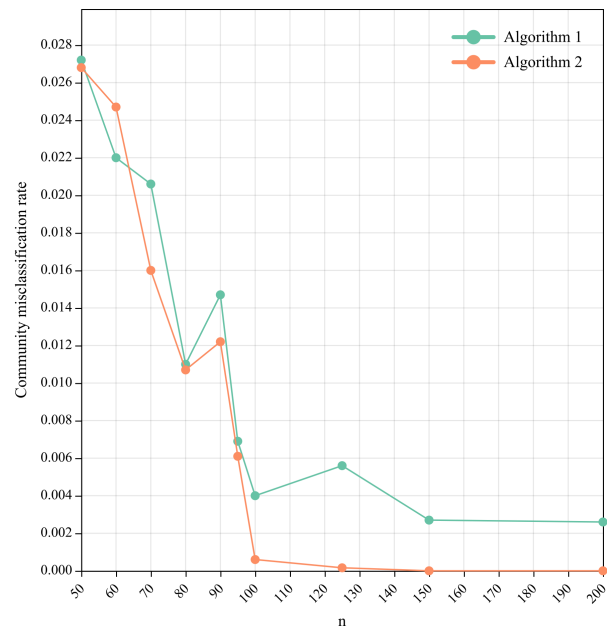**Fig. 2.** Average community misclassification rate for Algorithms 1 and 2, which is based on the average of 50 replications.

and *sd* of the estimated values, and compare the two cases of whether there is error in the estimation of homogeneity parameters. At the same time, we give the sensitivity analysis of the parameters. Finally, we compare the performance of our model and method in dynamic networks and static networks.

We can see from Fig. 1 that the average community misclassification rate for 50 repetitions by using Algorithm 1 has a significant decreasing trend as n increases. When the number of nodes n reaches 100, the community identification error of Algorithm 1 is able to reach within a desirable error range. Because the time cost of running Algorithm 2 once is equivalent to running Algorithm 1 n^2 times, we only compare the two algorithms when the network size n is much smaller. As shown in Fig. 2, when the network size n increases, the community misclassification rate of Algorithm 2 has a significant downward trend and even reaches 0 directly. Overall, the accuracy of Algorithm 2 is better than that of Algorithm 1 in the simulations. Therefore, we utilize Algorithm 2 for homogeneity identification when the network size n is moderate, and the accuracy of the community identification results is exchanged by the computational cost. When the network size n is large, Algorithm 1 can dramatically save computational cost while obtaining sufficient accuracy for homogeneity estimation.

In Table 1, as for peer pressure parameters $(\varphi_1, \phi_0, \phi_1)$, it can be found that with the increase of n, the estimated value $(\varphi_1, \phi_0, \phi_1)$ tends to approach the true value. In addition, when compared with Table 2, one can see that when n is small, the homogeneity identification error has a greater impact on the estimation of the result, but as n increases, this impact gradually decreases to disappear.

Moreover, it can be found from Table 1 that both the standard deviation (sd) and the mean square error (mse) decreases as n increases, and the speed of convergence to 0 is basically consistent with n. However, compared with Table 2, When n is not large enough, both the standard deviation and the mean square error are slightly inferior to the results estimated with the true homogeneity, which is still acceptable.

Finally, we compare the performance of our methodology in dynamic networks and static networks. From Table 3, we can see that although in a static network, with the increase in the number of nodes n, the estimation results still remain convergent, but all the estimation results are still worse than those in Table 2. Therefore, when the number of nodes is insufficient, the expansion of the time dimension can ensure that our estimation results are more accurate.

6　Empirical application: Teenage friends

Table 1. Peer pressure estimation with homogeneity identification error.

	Description of results	$\beta = -1$	$\varphi_1 = 1$	$\phi_0 = 1$	$\phi_1 = 1$	$\gamma_1'(1) = -2$	$\gamma_1'(2) = -1$	$\gamma_1'(3) = 1$	$\gamma_1'(4) = 2$
	mean	−1.0121	1.0211	0.9936	1.0046	−2.0063	−1.0064	1.0378	2.0324
n=500	sd	0.3844	0.5127	0.7130	0.6358	0.5607	0.5797	0.6780	0.7941
	mse	0.1464	0.2607	0.5033	0.4002	0.3113	0.3327	0.4566	0.6254
	mean	−1.0386	1.0209	0.9530	0.9677	−1.9985	−0.9785	1.0758	2.0837
n=1000	sd	0.2640	0.3237	0.4969	0.4245	0.4538	0.4053	0.5275	0.5704
	mse	0.0705	0.1041	0.2466	0.1794	0.2039	0.1631	0.2812	0.3291
	mean	−1.0100	0.9544	0.9995	0.9184	−2.0177	−1.0049	1.0023	2.0096
n=1500	sd	0.1739	0.2954	0.2616	0.3312	0.3824	0.2242	0.3916	0.3620
	mse	0.0300	0.0867	0.0678	0.1086	0.1449	0.0518	0.1518	0.1364

The first row of data represents the average of the predicted results, and the second row is the corresponding standard deviation, and the third row is mean squared error.

Table 2. Peer pressure estimation without homogeneity identification error.

	Description of results	$\beta = -1$	$\varphi_1 = 1$	$\phi_0 = 1$	$\phi_1 = 1$	$\gamma_1'(1) = -2$	$\gamma_1'(2) = -1$	$\gamma_1'(3) = 1$	$\gamma_1'(4) = 2$
	mean	−1.0662	0.9904	0.9811	0.9019	−2.0257	−0.9462	1.0923	2.1071
n=500	sd	0.3311	0.4596	0.7589	0.6382	0.5729	0.5127	0.6097	0.7203
	mse	0.1129	0.2093	0.5705	0.4128	0.3257	0.2632	0.3766	0.5252
	mean	−0.9984	0.9895	0.9884	0.9589	−2.0294	−1.0123	1.0075	2.0187
n=1000	sd	0.2322	0.3035	0.4612	0.4475	0.3812	0.3229	0.4173	0.4791
	mse	0.0534	0.0913	0.2107	0.1999	0.1448	0.1033	0.1724	0.2275
	mean	−1.0449	1.0200	0.9561	0.9710	−1.9413	−0.9538	1.0869	2.0915
n=1500	sd	0.1947	0.2463	0.4545	0.3837	0.3339	0.2932	0.4012	0.4564
	mse	0.0395	0.0605	0.2064	0.1466	0.1138	0.0872	0.1669	0.2146

The first row of data represents the average of the predicted results, and the second row is the corresponding standard deviation, and the third row is mean squared error.

Table 3. Peer pressure estimation without homogeneity identification error in static network.

	Description of results	$\beta = -1$	$\varphi_1 = 1$	$\phi_0 = 1$	$\phi_1 = 1$	$\gamma_1'(1) = -2$	$\gamma_1'(2) = -1$	$\gamma_1'(3) = 1$	$\gamma_1'(4) = 2$
	mean	−1.0142	0.9650	0.9821	0.9143	−2.1668	−1.0378	1.0212	2.0608
n=500	*sd*	0.4945	0.6656	1.0955	0.9485	0.8194	0.7533	0.8737	1.0097
	mse	0.2423	0.4398	1.1885	0.8981	0.6925	0.5632	0.7562	1.0130
	mean	−1.0496	1.0126	0.9703	0.9306	−1.9781	−0.9654	1.0754	2.1021
n=1000	*sd*	0.3532	0.4359	0.7097	0.6817	0.5820	0.5615	0.6576	0.7797
	mse	0.1260	0.1883	0.4995	0.4649	0.3358	0.3133	0.4338	0.6122
	mean	−1.0746	1.0175	0.8553	0.9018	−1.9112	−0.9108	1.1671	2.2148
n=1500	*sd*	0.3003	0.3659	0.6127	0.5475	0.5535	0.4602	0.5753	0.6518
	mse	0.0949	0.1328	0.3925	0.3064	0.3111	0.2177	0.3556	0.4668

The first row of data represents the average of the predicted results, the second row is the corresponding standard deviation, and the third row is the mean squared error.

Table 4. Statistical summary of the data.

Variable	Min	Max	Mean
smoking	0	1	0.47
KB centrality	0	1.43	0.33
money	−1	99.99	11.90
sex	0	1	0.57
age	12.8	14.3	13.32
smoke in home	−1	1	-0.13
parents smoke	−1	1	−0.02
brothers and sisters smoke	−1	1	0.74
num of friends	0	13	4.63

Note: In the table, sex: 0=girl, 1=boy; all family smoking: −1=yes 1=no.

Table 5. Estimation results.

Variable	GCPE model	KB-CPE model
KB centrality	−1.307	−2.768
Money	−	0.004
Sex	−	−0.221
Age	−	−0.094
Smoke in home	−	−0.812
Parents smoke	−	1.018
Brothers and sisters smoke	−	−0.559
Community1	−3.155	−
Community2	−1.623	−
Community3	−1.999	−
φ_1	0.581	0.597
ϕ_0	3.966	4.398
ϕ_1	9.326	−3.412
$\varphi_1 + \phi_1$	9.907	−2.816

and lifestyle study

6.1 Dataset description

The dataset is collected from the teenage friends and lifestyle study[①], which is designed to identify the influencing factors of adolescent smoking behavior during their growth. There are $n = 129$ teenagers without default attributes in the dataset and it also combines students' friendship networks and longitudinal survey data on their social and economic characteristics with time phases $m = 3$. Observed demographic characteristics included age, sex, pocket money, sports, music preferences and family smoking. The dependent variable Y is the smoking situation of students in the 3 phases. We record frequent smoking as $Y = 1$, and occasional smoking or non-smoking as $Y = 0$. Table 4 represents the descriptive statistics of the dataset.

6.2 Empirical results

We compare the performance of our GCPE model with a benchmark model in Ref. [8], which we call the KB-CPE model. The KB-CPE model does not consider homogeneity pursuit and is designed for static networks, thus we use the estimation results of the KB-CPE model based on the three

① https://www.stats.ox.ac.uk/snijders/siena/Glasgow_data.htm

periods' data for comparison. The initial value of Σ^0 in the NPLE method for both models is taken to be $\Sigma^0 = (0, \cdots, 0)$.

From Table 5, we can see that in our GCPE model, the three peer pressure parameters satisfy $0 < \varphi_1 < \phi_0 < \phi_1$. This implies that the adolescents who choose to smoke can be more affected by smoking friends than non-smoking friends, and whether there is a difference in the social influence of adolescents in social networks will not affect this conclusion. In addition, the individuals with higher KB centrality have greater influence on individuals with lower KB centrality, and non-smoking individuals with high KB centrality tend to avoid smoking when they are influenced by their non-smoking friends with low KB centrality.

For the KB-CPE model, $\varphi_1 > 0$ and $\phi_0 > 0$ but $\phi_1 < 0$, which can be problematic since when $S_{ji}^t > 0$ for all $j \in F_i^t$. Thus, the KB-CPE model cannot give a reasonable explanation for the estimation results of peer pressure parameters.

7 Conclusions

This study focuses on the estimation of peer pressure in dynamic homogeneous networks, which is extensively studied in social network analysis. We developed a generalized constant peer effects (GCPE) model based on the Markov perfect

Table 6. Overall change in the probability of establishing connections between communities.

	Description of results	$\beta=-1$	$\varphi_1=1$	$\phi_0=1$	$\phi_1=1$	$\gamma'_1(1)=-2$	$\gamma'_1(2)=-1$	$\gamma'_1(3)=1$	$\gamma'_1(4)=2$
	mean	−1.0662	1.0359	0.9494	0.8862	−1.914	−0.9346	1.1207	2.1141
P_{in}=0.6	sd	0.3581	0.4815	0.6219	0.6960	0.6108	0.5541	0.6104	0.6990
	mse	0.1313	0.2308	0.3854	0.4925	0.3768	0.3082	0.3834	0.4967
	mean	−1.0546	1.0218	0.9682	0.9100	−1.9207	−0.9497	1.0980	2.0856
P_{in}=0.65	sd	0.3931	0.5307	0.6349	0.6605	0.6320	0.5852	0.6490	0.7419
	mse	0.1560	0.2793	0.4000	0.4400	0.4018	0.3415	0.4266	0.5523
	mean	−1.0178	1.0000	0.9513	0.9401	−1.9788	−0.9903	1.0578	2.0440
P_{in}=0.7	sd	0.3773	0.4950	0.6759	0.6198	0.5705	0.5524	0.6202	0.7168
	mse	0.1413	0.2426	0.4547	0.3831	0.3227	0.3022	0.3842	0.5106
	mean	−1.0121	1.0211	0.9936	1.0046	−2.0063	−1.0064	1.0378	2.0324
P_{in}=0.75	sd	0.3844	0.5127	0.7130	0.6358	0.5607	0.5797	0.6780	0.7941
	mse	0.1464	0.2607	0.5033	0.4002	0.3113	0.3327	0.4566	0.6254
	mean	−1.0274	1.0232	1.0281	0.9985	−1.9805	−0.9958	1.0410	2.0283
P_{in}=0.8	sd	0.3613	0.5463	0.7509	0.5649	0.5510	0.5641	0.6808	0.7904
	mse	0.1300	0.2960	0.5590	0.3160	0.3009	0.3151	0.4606	0.6193
	mean	−1.0091	0.9622	1.0916	1.0139	−2.0263	−1.0278	0.9845	1.9485
P_{in}=0.85	sd	0.3913	0.5619	0.7502	0.5377	0.5805	0.6085	0.7539	0.8560
	mse	0.1517	0.3140	0.5655	0.2864	0.3343	0.3673	0.5630	0.7280
	mean	−1.0369	1.0129	1.0359	1.0453	−1.9987	−0.9779	1.0484	2.0308
P_{in}=0.9	sd	0.3681	0.4898	0.8767	0.5839	0.5721	0.5599	0.7943	0.9219
	mse	0.1355	0.2377	0.7623	0.3391	0.2340	0.3109	0.6269	0.8424

The first row of data represents the average of the predicted results, and the second row is the corresponding standard deviation, and the third row is mean squared error. P_{in} means the probability of establishing directed edges at nodes of the same community and P_{out} means the probability of establishing directed edges at nodes of different communities, where $P_{out} = 1 - P_{in}$. Network size $n = 400$. Homogeneity exists for identification errors.

Table 7. Different community link probabilities.

	Description of results	$\beta=-1$	$\varphi_1=1$	$\phi_0=1$	$\phi_1=1$	$\gamma'_1(1)=-2$	$\gamma'_1(2)=-1$	$\gamma'_1(3)=1$	$\gamma'_1(4)=2$
	mean	−0.9904	0.9583	0.9630	0.9222	−2.0339	−1.0596	1.0140	2.0494
n=500	sd	0.4742	0.7220	0.6022	0.7058	0.4392	0.4766	0.5805	0.6898
	mse	0.2227	0.5179	0.3604	0.4992	0.1921	0.2284	0.3338	0.4735
	mean	−0.9570	0.9952	1.0645	1.0291	−2.0320	−1.0426	0.9375	1.9346
n=1000	sd	0.3359	0.4383	0.5460	0.4611	0.2979	0.3241	0.4872	0.5672
	mse	0.1135	0.1893	0.2993	0.2113	0.0889	0.1058	0.2389	0.3228
	mean	−0.9642	0.9679	0.9679	1.0068	−2.0368	−1.0467	0.9911	1.9926
n=1500	sd	0.2883	0.3521	0.4112	0.4010	0.2749	0.2996	0.3734	0.4570
	mse	0.0836	0.1237	0.1684	0.1593	0.0761	0.0910	0.1381	0.2068

The first row of data represents the average of the predicted results, and the second row is the corresponding standard deviation, and the third row is mean squared error. The probability matrix $P = \begin{bmatrix} 0.75 & 0.25 & 0.35 & 0.15 \\ 0.35 & 0.80 & 0.05 & 0.20 \\ 0.30 & 0.35 & 0.70 & 0.25 \\ 0.15 & 0.25 & 0.15 & 0.75 \end{bmatrix}$ for establishing links between communities. Note that in a real simulation, the probability is not the above result; for example, $p_{(\mathbf{C}(1),\mathbf{C}(1))} = P_{11}$ and $p_{(\mathbf{C}(1),\mathbf{C}(2))} = (1 - P_{11})\dfrac{P_{12}}{P_{12} + P_{13} + P_{14}}$. Homogeneity exists for identification errors.

equilibrium, which can effectively avoid bias through homogeneity pursuit and can be applied to a wider range of networks. To estimate peer pressure in the GCPE model, we first present two algorithms to estimate homogeneity parameters. Then we introduce the nested pseudo-likelihood (NPLE) method and apply it to obtain consistent estimators of peer pressure. We further represent the theoretical properties of the NPLE estimators. Simulation evaluations show that our proposed methodology can achieve desirable and effective results in terms of the community misclassification rate and
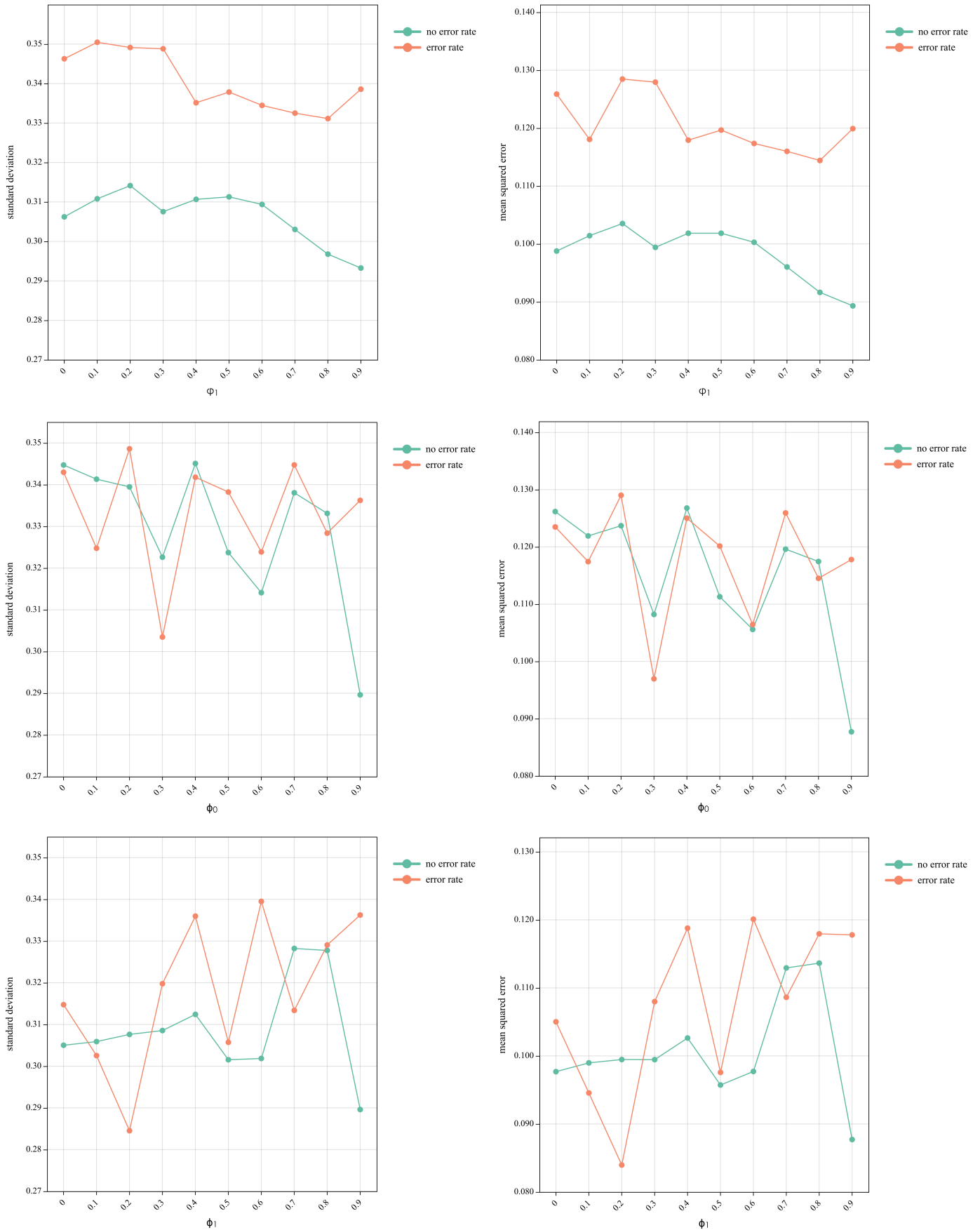
Fig. 3. The standard deviation and mean square error of the estimated peer pressure parameter.

parameter estimation error. We also illustrate the advantages of our model in empirical analysis when compared with the KB-CPE model.

Two interesting directions for future research are the plurality of decisions made by gamers and the multiple choices of homogeneous settings. In social networks, individuals do not behave in only two categories; thus, it is a natural idea to expand the decision dimensions of the players in the game. Although the homogeneity setting has been extensively studied and we have succeeded in proving the consistency of the NPLE estimators in the setting of homogeneity pursuit, the consistent properties of the NPLE estimators under other settings are also worth further study.

Acknowledgements

Conflict of interest

The authors declare that they have no conflict of interest.

Biographies

Jie Liu   is an Associate Professor at the University of Science and Technology of China (USTC). He received his Bachelor's degree and Ph.D. degree in Statistics from USTC in 2003 and 2008, respectively. His major research interests include random networks, statistical modelling, and network analysis.

Jiayang Zhao   is currently a Ph.D. student at the University of Science and Technology of China (USTC). His research interests focus on network data modelling and subgroup analysis.

References

[1] Durlauf S N, Ioannides Y M. Social interactions. *Annual Review of Economics,* **2010**, *2* (1): 451–478.

[2] Kim J, Kim M, Choi J, et al. Offline social interactions and online shopping demand: Does the degree of social interactions matter? *Journal of Business Research,* **2019**, *99*: 373–381.

[3] Poutvaara P, Siemers L R. Smoking and social interaction. *Journal of Health Economics,* **2008**, *27* (6): 1503–1515.

[4] Yin J, He X, Yang Y, et al. Outcome-based evaluations of social interaction valence in a contingent response context. *Frontiers in Psychology,* **2019**, *10*: 2557.

[5] Sirakaya S. Recidivism and social interactions. *Journal of the American Statistical Association,* **2006**, *101*: 863–877.

[6] Blume L E, Brock W A, Durlauf S N, et al. Linear social interactions models. *Journal of Political Economy,* **2015**, *123* (2): 444–496.

[7] Xu H. Social interactions in large networks: A game theoretic approach. *International Economic Review,* **2018**, *59* (1): 257–284.

[8] Lin Z, Xu H. Estimation of social-influence-dependent peer pressure in a large network game. *The Econometrics Journal,* **2017**, *20* (3):

S86–S102.

[9] Sun Z, Du Y, Chen X, et al. Implicit community discovery based on microblog theme homogeneit. *IOP Conference Series: Materials Science and Engineering,* **2020**, *790* (1): 012045.

[10] Favre G, Figeac J, Grossetti M, et al. Social distance in France: Evolution of homogeneity within personal networks from 2001 to 2017. *Social Networks,* **2022**, *68*: 70–83.

[11] Liu L, Wang X, Zheng Y, et al. Homogeneity trend on social networks changes evolutionary advantage in competitive information diffusion. *New Journal of Physics,* **2020**, *22* (1): 013019.

[12] Shalizi C R, Thomas A C. Homophily and contagion are generically confounded in observational social network studies. *Sociological Methods & Research,* **2011**, *40* (2): 211–239.

[13] Davin J P, Gupta S, Piskorski M J. Separating homophily and peer influence with latent space. Boston, MA: Harvard Business School, **2014**.

[14] Hill S, Provost F, Volinsky C. Network-based marketing: Identifying likely adopters via consumer networks. *Statistical Science,* **2006**, *21* (2): 256–276.

[15] Worrall H. Community detection as a method to control for homophily in social networks. Corpus ID: 15409339, **2014**.

[16] McFowland III E, Shalizi C R. Estimating causal peer influence in homophilous social networks by inferring latent locations. *Journal of the American Statistical Association,* **2023**, *118*: 707–718.

[17] Aguirregabiria V, Mira P. Sequential estimation of dynamic discrete games. *Econometrica,* **2007**, *75* (1): 1–53.

[18] Egesdal M, Lai Z, Su C-L. Estimating dynamic discrete-choice games of incomplete information. *Quantitative Economics,* **2015**, *6* (3): 567–597.

[19] Manski C F. Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies,* **1993**, *60* (3): 531–542.

[20] Seim K. An empirical model of firm entry with endogenous product-type choices. *The RAND Journal of Economics,* **2006**, *37* (3): 619–640.

[21] Han Q, Xu K, Airoldi E. Consistent estimation of dynamic and multilayer block models. In: Proceedings of the 32nd International Conference on Machine Learning. Lille, France: JMLR, **2015**, 37: 1511–1520.

[22] Pensky M, Zhang T. Spectral clustering in the dynamic stochastic block model. *Electronic Journal of Statistics,* **2019**, *13*: 678–709.

[23] Chunaev P. Community detection in node-attributed social networks: A survey. *Computer Science Review,* **2020**, *37*: 100286.

[24] Liu M, Guo J, Chen J. Community discovery in weighted networks based on the similarity of common neighbors. *Journal of Information Processing Systems,* **2019**, *15* (5): 1055–1067.

[25] Gao C, Ma Z, Zhang A Y, et al. Achieving optimal misclassification proportion in stochastic block models. *The Journal of Machine Learning Research,* **2017**, *18* (1): 1980–2024.

[26] Bickel P J, Chen A. A nonparametric view of network models and Newman–Girvan and other modularities. 2009. *Proceedings of the National Academy of Sciences,* **2009**, *106* (50): 21068–21073.

[27] Zhao Y, Levina E, Zhu J. Consistency of community detection in networks under degree-corrected stochastic block models. *Annals of Statistics,* **2012**, *40*: 2266–2292.

[28] Kasahara H, Shimotsu K. Sequential estimation of structural models with a fixed point constraint. *Econometrica,* **2012**, *80* (5): 2303–2319.

Appendix

A1   Proof of Theorem 4.1

To complete this proof, we need to use part of the notations from[7]. Let $\mathfrak{P}_i^{th}(\mathbb{W}^t)$ represent the Markov perfect equilibrium solution of player i in subnetwork $\mathbb{G}_{(i,h)}^t$ of $\mathbb{G}^t$ at time t. Let

$$L_n(\theta, \mathfrak{P}) = \frac{1}{n} \sum_{t=1}^{m} \sum_{i=1}^{n} \mathbf{E}[(Y_i^t \ln \Gamma_i^t(\mathfrak{P}^t, \theta; \mathbb{W}^t) + (1 - Y_i^t) \ln(1 - \Gamma_i^t(\mathfrak{P}^t, \theta; \mathbb{W}^t)))],$$

$$\Lambda_n = \{\theta \in \Theta : \theta = \mathrm{argmax}_{c \in \Theta} L_n(c, \mathfrak{P}(\theta, \mathbb{W}))\},$$

$$\hat{L}_n(\tilde{\theta}, \mathfrak{P}) = \frac{1}{n} \sum_{t=1}^{m} \sum_{i=1}^{n} \mathbf{E}[\left(Y_i^t \ln \Gamma_i^t\left(\mathfrak{P}^t, \tilde{\theta}; \tilde{\mathbb{W}}^t\right) + (1 - Y_i^t) \ln(1 - \Gamma_i^t\left(\mathfrak{P}^t, \tilde{\theta}; \tilde{\mathbb{W}}^t\right))\right)],$$

$$\tilde{\Lambda}_n = \{\tilde{\theta} \in \Theta : \tilde{\theta} = \mathrm{argmax}_{c \in \Theta} \hat{L}_n(c, \mathfrak{P}(\tilde{\theta}, \tilde{\mathbb{W}}))\}.$$

It can be seen from Ref. [8] that $\Lambda_n \to \theta_0$ as $n \to \infty$. Then, we need to complete the following two parts of the proof:

$$\sup_{\theta \in \Theta} |\tilde{L}_n(\theta, \mathfrak{P}) - L_n(\theta, \mathfrak{P})| \xrightarrow{P} 0,$$

$$d_{\mathcal{H}}(\Lambda_n, \tilde{\theta}_{NPLE}) \xrightarrow{P} 0,$$

where $d_{\mathcal{H}}$ is the Hausdorff distance measure. According to Ref. [7], it is sufficient to prove that $\tilde{L}_n(c, \mathfrak{P}) - L_n(c, \mathfrak{P}) \xrightarrow{P} 0$.

$$\tilde{L}_n(c, \mathfrak{P}) - L_n(c, \mathfrak{P}) = \frac{1}{n} \sum_{t=1}^{m} \sum_{i=1}^{n} \left(Y_i^t \ln \mathfrak{P}_i^{t*}(\tilde{\mathbb{W}}^t; c) - \mathbf{E} \mathfrak{P}_i^{t*}(\mathbb{W}^t; \theta) \ln \mathfrak{P}_i^{t*}(\mathbb{W}^t; c)\right) + \frac{1}{n} \sum_{t=1}^{m} \sum_{i=1}^{n} ((1 - Y_i^t) \ln(1 - \mathfrak{P}_i^{t*}(\tilde{\mathbb{W}}^t; c)) -$$

$$\mathbf{E}(1 - \mathfrak{P}_i^{t*}(\mathbb{W}^t; \theta)) \ln(1 - \mathfrak{P}_i^{t*}(\mathbb{W}^t; c))) = l_1 + l_2 + l_3 + l_3 + l_4 + l_5 + l_6 + l_7 + l_8,$$

where

$$l_1 := \frac{1}{n} \sum_{t=1}^{m} \sum_{i=1}^{n} \left((Y_i^t - \mathfrak{P}_i^{t*}(\tilde{\mathbb{W}}^t; \theta)) \ln \mathfrak{P}_i^{t*}(\tilde{\mathbb{W}}^t; c)\right),$$

$$l_2 := \frac{1}{n} \sum_{t=1}^{m} \sum_{i=1}^{n} \left(((1 - Y_i^t) - (1 - \mathfrak{P}_i^{t*}(\tilde{\mathbb{W}}^t; \theta))) \ln(1 - \mathfrak{P}_i^{t*}(\tilde{\mathbb{W}}^t; c))\right),$$

$$l_3 := \frac{1}{n} \sum_{t=1}^{m} \sum_{i=1}^{n} \left(\mathfrak{P}_i^{t*}(\tilde{\mathbb{W}}^t; \theta) \ln \mathfrak{P}_i^{t*}(\tilde{\mathbb{W}}^t; c) - \mathfrak{P}_i^{th}(\mathbb{W}^t; \theta) \ln \mathfrak{P}_i^{th}(\mathbb{W}^t; c)\right),$$

$$l_4 := \frac{1}{n} \sum_{t=1}^{m} \sum_{i=1}^{n} ((1 - \mathfrak{P}_i^{t*}(\tilde{\mathbb{W}}^t; \theta)) \ln(1 - \mathfrak{P}_i^{t*}(\tilde{\mathbb{W}}^t; c)) - (1 - \mathfrak{P}_i^{th}(\mathbb{W}^t; \theta)) \ln(1 - \mathfrak{P}_i^{th}(\mathbb{W}^t; c))),$$

$$l_5 := \frac{1}{n} \sum_{t=1}^{m} \sum_{i=1}^{n} (\mathfrak{P}_i^{th}(\mathbb{W}^t; \theta) \ln \mathfrak{P}_i^{th}(\mathbb{W}^t; c) - \mathbf{E} \mathfrak{P}_i^{th}(\mathbb{W}^t; \theta) \ln \mathfrak{P}_i^{th}(\mathbb{W}^t; c),)$$

$$l_6 := \frac{1}{n} \sum_{t=1}^{m} \sum_{i=1}^{n} ((1 - \mathfrak{P}_i^{th}(\mathbb{W}^t; \theta)) \ln(1 - \mathfrak{P}_i^{th}(\mathbb{W}^t; c)) - \mathbf{E}(1 - \mathfrak{P}_i^{th}(\mathbb{W}^t; \theta)) \ln(1 - \mathfrak{P}_i^{th}(\mathbb{W}^t; c))),$$

$$l_7 := \frac{1}{n} \sum_{t=1}^{m} \sum_{i=1}^{n} (\mathbf{E} \mathfrak{P}_i^{th}(\mathbb{W}^t; \theta) \ln \mathfrak{P}_i^{th}(\mathbb{W}^t; c) - \mathbf{E} \mathfrak{P}_i^{t*}(\mathbb{W}^t; \theta) \ln \mathfrak{P}_i^{t*}(\mathbb{W}^t; c)),$$

$$l_8 := \frac{1}{n} \sum_{t=1}^{m} \sum_{i=1}^{n} (\mathbf{E}(1 - \mathfrak{P}_i^{th}(\mathbb{W}^t; \theta)) \ln(1 - \mathfrak{P}_i^{th}(\mathbb{W}^t; c)) - \mathbf{E}(1 - \mathfrak{P}_i^{t*}(\mathbb{W}^t; \theta)) \ln(1 - \mathfrak{P}_i^{t*}(\mathbb{W}^t; c)).$$

We can obtain that $l_1 + l_2 \xrightarrow{P} 0$, $l_7 + l_8 \xrightarrow{P} 0$ and $min\{\mathfrak{P}_i^{t*}(\mathbb{W}^t; \theta), (1 - \mathfrak{P}_i^{t*}(\mathbb{W}^t; \theta))\} \geqslant \mathfrak{P}_0$.

$$\mathbf{E} \left|\mathfrak{P}_i^{t*}(\tilde{\mathbb{W}}^t; \theta) \ln \mathfrak{P}_i^{t*}(\tilde{\mathbb{W}}^t; c) - \mathfrak{P}_i^{th}(\mathbb{W}^t; \theta) \ln \mathfrak{P}_i^{th}(\mathbb{W}^t; c)\right| \leqslant$$

$$\mathbf{E} \left|(\mathfrak{P}_i^{t*}(\tilde{\mathbb{W}}^t; \theta) - \mathfrak{P}_i^{th}(\mathbb{W}^t; \theta)) \ln \mathfrak{P}_i^{t*}(\tilde{\mathbb{W}}^t; c)\right| + \mathbf{E} \left|\mathfrak{P}_i^{th}(\mathbb{W}^t; \theta)(\ln \mathfrak{P}_i^{t*}(\tilde{\mathbb{W}}^t; c) - \ln \mathfrak{P}_i^{th}(\mathbb{W}^t; c))\right| \leqslant$$

$$-\ln \mathfrak{P}_0 \mathbf{E} \left|(\mathfrak{P}_i^{t*}(\tilde{\mathbb{W}}^t; \theta) - \mathfrak{P}_i^{th}(\mathbb{W}^t; \theta))\right| + \frac{1}{\mathfrak{P}_0} \mathbf{E} \left|(\mathfrak{P}_i^{t*}(\tilde{\mathbb{W}}^t; c) - \mathfrak{P}_i^{th}(\mathbb{W}^t; c))\right|.$$

By Assumption 3.1, we can obtain

$$Pr\{\left|(\mathfrak{P}_i^{t*}(\tilde{\mathbb{W}}^t; \theta) - \mathfrak{P}_i^{th}(\mathbb{W}^t; \theta))\right| \neq \left|(\mathfrak{P}_i^{t*}(\mathbb{W}^t; \theta) - \mathfrak{P}_i^{th}(\mathbb{W}^t; \theta))\right|\} \to 0, \quad \text{as} \quad n \to \infty.$$

Then,

$$\mathbf{E} \left|(\mathfrak{P}_i^{t*}(\tilde{\mathbb{W}}^t; \theta) - \mathfrak{P}_i^{th}(\mathbb{W}^t; \theta))\right| \to 0, \quad \text{as} \quad n \to \infty.$$

Therefore, $l_3 + l_4 \xrightarrow{P} 0$. Next we need to prove that $l_5 + l_6 \xrightarrow{P} 0$.

$$\mathbf{E}\{\frac{1}{n}\sum_{t=1}^{m}\sum_{i=1}^{n}(\mathfrak{P}_i^{th}(\mathbb{W}^t;\theta)\ln\mathfrak{P}_i^{th}(\mathbb{W}^t;c) - \mathbf{E}\mathfrak{P}_i^{th}(\mathbb{W}^t;\theta)\ln\mathfrak{P}_i^{th}(\mathbb{W}^t;c))\}^2 =$$

$$\frac{1}{n^2}\sum_{t=1}^{m}\sum_{i,j=1}^{n}\mathbf{Cov}\left(\mathfrak{P}_i^{th}(\mathbb{W}^t;\theta)\ln\mathfrak{P}_i^{th}(\mathbb{W}^t;c), \mathfrak{P}_j^{th}(\mathbb{W}^t;\theta)\ln\mathfrak{P}_j^{th}(\mathbb{W}^t;c)\right).$$

From Ref. [7], we can know that

$$\mathbf{Cov}\left(\mathfrak{P}_i^{th}(\mathbb{W}^t;\theta)\ln\mathfrak{P}_i^{th}(\mathbb{W}^t;c), \mathfrak{P}_j^{th}(\mathbb{W}^t;\theta)\ln\mathfrak{P}_j^{th}(\mathbb{W}^t;c)\right) \leqslant (\ln\mathfrak{P}_0)^2.$$

Then, we will count how many related players each player has. By condition 7, we can know if there does not exist $k \in N_{(i,h)}^t \cap N_{(j,h)}^t$, then $\mathfrak{P}_i^{th}(\mathbb{W}^t;\theta)$ and $\mathfrak{P}_j^{th}(\mathbb{W}^t;\theta)$ are independent. By condition 8, we can easily find that as n increases, the correlation brought by S_{ji}^t becomes increasingly weaker. Therefore, by condition 6, we can determine that each player is associated with at most $(1 + c_0 + \cdots + c_0^h) \leqslant c_0^{h+1}$ players.

Last, we have

$$\mathbf{E}\{\frac{1}{n}\sum_{t=1}^{m}\sum_{i=1}^{n}(\mathfrak{P}_i^{th}(\mathbb{W}^t;\theta)\ln\mathfrak{P}_i^{th}(\mathbb{W}^t;c) - \mathbf{E}\mathfrak{P}_i^{th}(\mathbb{W}^t;\theta)\ln\mathfrak{P}_i^{th}(\mathbb{W}^t;c))\}^2 = \frac{1}{n^2}\sum_{t=1}^{m}\sum_{i,j=1}^{n}\mathbf{Cov}\left(\mathfrak{P}_i^{th}(\mathbb{W}^t;\theta)\ln\mathfrak{P}_i^{th}(\mathbb{W}^t;c), \mathfrak{P}_j^{th}(\mathbb{W}^t;\theta)\ln\mathfrak{P}_j^{th}(\mathbb{W}^t;c)\right) \leqslant$$

$$\frac{c_0^{h+1}}{n}(\ln\mathfrak{P}_0)^2 \to 0.$$

Therefore, $l_5 + l_6 \xrightarrow{P} 0$ and it follows that $\sup_{\theta\in\Theta}|\tilde{L}_n(\theta,\mathfrak{P}) - L_n(\theta,\mathfrak{P})| \xrightarrow{P} 0$.

According to Ref. [8], we can obtain $d_{\mathcal{H}}(\tilde{\Lambda}_n, \tilde{\theta}_{NPLE}) \xrightarrow{p} 0$. Therefore, it is sufficient to prove $d_{\mathcal{H}}(\Lambda_n, \tilde{\Lambda}_n) \xrightarrow{p} 0$. From the definitions of Λ_n and $\tilde{\Lambda}_n$, we can find that under Assumption 3.1, $Pr\{\Lambda_n \neq \tilde{\Lambda}_n\} \to 0$, as $n \to \infty$ holds. Then, $Pr\{d_{\mathcal{H}}(\Lambda_n, \tilde{\Lambda}_n) \to 0\} \to 1$, as $n \to \infty$ holds. Thus we can obtain $d_{\mathcal{H}}(\Lambda_n, \tilde{\Lambda}_n) \xrightarrow{P} 0$. So far, we have completed the proof of $d_{\mathcal{H}}(\Lambda_n, \tilde{\theta}_{NPLE}) \xrightarrow{P} 0$. In summary, $\tilde{\theta}_{NPLE} \xrightarrow{P} \theta_0$.

A2   Proof of Theorem 4.2

In addition to the symbols used in the proof of Theorem 4.1, we need to supplement the symbols:

$$L_n'(\theta,\mathfrak{P}) = \frac{1}{n}\sum_{t=1}^{m}\sum_{i=1}^{n}\mathbf{E}[(Y_i^t\ln\Gamma_i^{t'}(\mathfrak{P}^t,\theta;\mathbb{W}^t) + (1-Y_i^t)\ln(1-\Gamma_i^{t'}(\mathfrak{P}^t,\theta;\mathbb{W}^t)))],$$

and

$$\Lambda_n' = \{\theta \in \Theta : \theta = \mathrm{argmax}_{c\in\Theta}L_n'(c,\mathfrak{P}(\theta,\mathbb{W}))\},$$

where $L_n'(\theta,\mathfrak{P})$ represents the equation after replacing S_{ji}^t with $S_{ji}^{t'}$.

It is not difficult to verify that L_n is a continuous function with respect to S_{ji}^t. then, from the proof in Ref. [8], we can deduce that $\Lambda_n' \to \theta_0$ as $n \to \infty$. Next, Remove equations l_5 and l_6 in the proof of Theorem 4.1, the proofs are the same as Theorem 4.1.

Similarly, we will count how many related plays each play has. By condition 7 in Theorem 4.1, we know if there does not exist $k \in N_{(i,h)}^t \cap N_{(j,h)}^t$, then $\mathfrak{P}_i^{th}(\mathbb{W}^t;\theta)$ and $\mathfrak{P}_j^{th}(\mathbb{W}^t;\theta)$ are independent. By condition 1, the correlation $S_{ji}^{t'}$ is no longer a negligible quantity, but under our adjustment, it will be a manageable quantity. From the original need to only consider the correlation of a player with other players, we have changed to consider the correlation between players in a subnetwork and other players. Therefore, by condition 6, we can determine that each player is associated with at most $(1 + c_0 + \cdots + c_0^h)^2 \leqslant c_0^{2h+2}$ players. Therefore, $l_5 + l_6 \xrightarrow{P} 0$, which follows that $\sup_{\theta\in\Theta}|\tilde{L}_n(\theta,\mathfrak{P}) - L_n(\theta,\mathfrak{P})| \xrightarrow{P} 0$. To sum up, $\tilde{\theta}_{NPLE}' \xrightarrow{P} \theta_0$.