Statistica Sinica Preprint No: SS-2024-0283

Title	High-Dimensional-Responses-Assisted Heterogeneous Nodal Influence Analysis
Manuscript ID	SS-2024-0283
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202024.0283
Complete List of Authors	Dongxue Zhang, Wei Lan, Danyang Huang and Huazhen Lin
Corresponding Authors	Danyang Huang
E-mails	dyhuang89@126.com

Notice: Accepted author version.

Statistica Sinica

HIGH-DIMENSIONAL-RESPONSES-ASSISTED HETEROGENEOUS NODAL INFLUENCE ANALYSIS

Dongxue Zhang, Wei Lan, Danyang Huang and Huazhen Lin

Southwestern University of Finance and Economics, Renmin University of China

Abstract: We consider an $m \times n$ matrix network data with m network nodes and n-dimensional responses for each node, where both m and n can diverge to infinity. The heterogeneity of network nodal influence is addressed by different influence parameters of each node, which are expressed through high-dimensional responses using a specific link function. By allowing heterogeneous error variances, we propose a response-assisted network influence model to integrate information of the matrix response variable and network structures across both m network nodes and n dimensions of responses. Since the traditional maximum likelihood estimator is invalid in this case, we build an "optimal" generalized method of moments estimator, to avoid estimating unknown error variances by restricting the diagonal of weighting matrix in quadratic moments. The consistency and asymptotic normality of the estimator are established. We have also developed a homogeneity test to examine the influence heterogeneity and presented simulations and an empirical study of fund and stock to demonstrate the model's utility.

Key words and phrases: Generalized method of moments estimation, heteroscedastic errors, matrix network data, matrix network influence model.

1. Introduction

Network data is commonly encountered in real practice, such as economics, sociology, and politics (Bramoullé et al., 2009; Ozsoylev and Walden, 2011; Fracassi, 2017). Due to the existence of peer effect, behavior pattern of an individual or an organization could influence its related neighbors in the network. In practice, influential powers or the magnitude of the influence are usually unbalanced across individuals. Some individuals, such as celebrities or opinion leaders, always have strong influence, while others tend to have weak influence (Li et al., 2013; Ma and Liu, 2014). Identifying the influential nodes in a network has important applications. For example, locating influential nodes can help companies determine how best to promote coupons to potential customers in the network to attract more sales (Anagnostopoulos et al., 2008).

There are many methods can be used to identify influential nodes. For example, some methods make use of topology information in the network, such as computing nodal degrees and centrality measurements (Newman, 2001; Weng et al., 2010; Ozsoylev and Walden, 2011). Recently, Fang and Hu (2018) integrated three network characteristics, which are social influence, entity similarity, and structural equivalence, to predict the top persuaders in a social network. However, these methods solely focus on network structure and do not take into account the response information from individual nodes. In fact, in network analysis, a node's influence is strongly correlated with its individual responses. For instance, consider the international import and export network. If the response variable is the trading volume of consumer goods, China's influence would be more significant due to its large population and rapid economic development. In contrast, if the response variable is the trading volume of oil, countries in the Middle East, like Saudi Arabia, with abundant oil resources, may wield greater influence. It is important to note that countries with higher nodal degrees or centralities may not necessarily have greater influence on consumer goods or oil trade, because they are not closely associated with the trading objects. Consequently, relying solely on network characteristics is limited and can yield misleading results. As a result, it is crucial to integrate information from the response variables and network structure for accurately identifying individual influence.

There exist some works identifying the influential nodes in a network in light of the nodes' responses. For example, Zou et al. (2021) investigated the spillover effects from one mutual fund to another via cash flows from a network perspective. Meanwhile, Zhu et al. (2019) identified the influential users via post length in a Sina Weibo dataset. Similarly, based on the Sina

Weibo data, Wu et al. (2022) analyzed the inward and outward influences of users via post numbers. These works focus on single response, that is *vector network data*. They can not be used to analyze *matrix network data* of large scale, represented by $\mathbf{Y} = (Y_{ij}) \in \mathbb{R}^{m\times n}$. The data consists of m network nodes with n-dimensional responses for each node with both m and n being allowed to go to infinity. As a result, the matrix network data considers possibly high-dimensional responses with diverging n, which include vector network data as a special case with $n = 1$. An example for the structure of matrix network data is presented in Figure 1. For the matrix network data, we need to model the correlations not only between network nodes, but also among high-dimensional responses. Consequently, the aforementioned models that designed for single response are not applicable. This motivates us to propose a high-dimensional response-assisted method to identify the influential nodes for the matrix network data.

Matrix network data are typically encountered in real practice. For example, we may consider fund and stock matrix network data. To capture various opportunities in the Chinese stock market, fund managers often select stocks in the WIND A50 Index for their portfolios. For each fund, the proportion of fund's portfolio invested in the 50 stocks involved in the index could be treated as a high-dimensional response, where the dimension is the

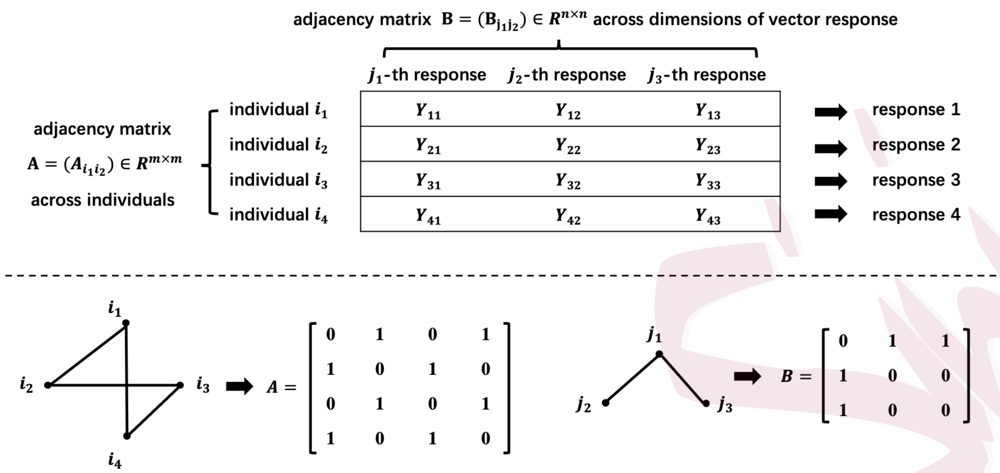

Figure 1: The example of structure for the matrix network data and adjacency matrices with m=4 and n=3. In the upper panel, the observation $\mathbb{Y} = (Y_{ij})$ is an $m \times n$ matrix, which involves m individuals with n-dimensional responses. In the lower panel, the two types of relationships among both m individuals and n dimensions of response are described as adjacency matrices $A = (A_{i_1 i_2}) \in \mathbb{R}^{m\times m}$ and $B = (B_{j_1 j_2}) \in \mathbb{R}^{n\times n}$, respectively. Specifically, for $1 \leq i_1, i_2 \leq m$, we define $A_{i_1 i_2} = 1$ if there is a connection from individual i_1 to individual i_2 and $A_{i_1 i_2} = 0$ otherwise. For the sake of completeness, we define $A_{ii} = 0$ for $1 \leq i \leq m$. We similarly define the adjacency matrix B.

number of stocks. In this way, the vectors of different funds form a matrix, where Y_{ij} describes the investment proportion of fund i's portfolio in stock j. We may also consider critic and film matrix data. Good reviews of the films can attract more viewers, while poor reviews can discourage viewers, suggesting that critics play a dual role in box office performance. For each critic, the reviews of different films can be regarded as a high-dimensional response. Hereafter, the vectors of different critics form a matrix, where Y_{ij} describes the review of critic i on film j. In the above examples, it is of great value to identify influential individuals (e.g., funds, critics) via the high-dimensional response (e.g., investment proportion, reviews). By doing so, fund managers and investors can consider the strategies of the influential funds for reference to reap profit; the film studios can strategically manage the influential critics' reviews to enhance box office revenue.

To quantify the heterogeneous nodal influence with matrix network data, we propose a novel high-dimensional-responses-assisted matrix network influence model (MNIM) to integrate the information from both high dimensional responses and network structures. In this scenario, the variance of the response variable will differ at each observation, depending on the strength and structure of the interactions (see e.g., Glaeser et al. (1996); Guiso et al. (2004); Anselin (2013)). To address this, the proposed model is considered as particular heteroscedastic formulations, in which we allow the variances of the error terms to be different across i and j. As a result, we need to estimate $m \times n$ unknown error variances, which is infeasible with the traditional maximum likelihood (ML) estimation method, shown in Section S1 of Supplementary Material. To tackle this, with certain designed moment conditions, we build an "optimal" generalized method of moments estimator (GMME) that does not require estimating unknown error variances. The consistency and asymptotic normality of the "optimal" GMME are further developed. Specifically, our theoretical framework is established under $mn \to \infty$ with a convergence rate of $\sqrt{mn}$ for unknown parameters. Hence, comparing with the convergence rate of $\sqrt{m}$ for uni-dimensional response models, the estimation efficiency is greatly improved with the assistance of the high-dimensional response variables. It is worth noting that our theoretical framework remains valid when considering fixed number of network nodes m. This sets our approach apart from the methodologies proposed in the works of Zhu et al. (2019), Zou et al. (2021) and Wu et al. (2022), all of which require m tends to infinity. Finally, we also develop homogeneity and link function tests to examine the influence pattern and to assess the adequacy of the pre-specified function.

The rest of this article is organized as follows. Section 2 describes the model and methodology. Specifically, we develop the GMME and apply it to two sets of test statistics, and further investigate their theoretical properties. Section 3 presents some simulations. Section 4 introduces a case study with fund and stock matrix data. Section 5 discusses the results and concludes the paper. All technique details are relegated to Supplementary Material.

2. Models and Methodology

2.1 Model setup

Let i $(1 \leq i \leq m)$ be the individual index, and j $(1 \leq j \leq n)$ be the dimension index of vector response, where m is the number of nodes in the network and n is the dimension of the response. For the j-th response of individual i, we have a response variable Y_{ij}. Denote $\mathbb{Y} = (Y_{ij}) \in \mathbb{R}^{m \times n}$, where both m and n can be fixed or tend to infinity, as long as mn goes to infinity. Next, we define $X_i^{(1)} = (X_{i1}^{(1)}, \cdots, X_{ip_1}^{(1)}) \in \mathbb{R}^{1 \times p_1}$ as the covariates for i-th individual, and $X_j^{(2)} = (X_{j1}^{(2)}, \cdots, X_{jp_2}^{(2)}) \in \mathbb{R}^{1 \times p_2}$ as the covariates for j-th response. To consider the network structure across both individuals and responses, we construct two adjacency matrices, $A = (A_{i_1 i_2}) \in \mathbb{R}^{m \times m}$ and $B = (B_{j_1 j_2}) \in \mathbb{R}^{n \times n}$ for $1 \leq i_1, i_2 \leq m$ and $1 \leq j_1, j_2 \leq n$, as in Figure 1 of the Introduction part. Correspondingly, the weighting matrices $U = (U_{i_1 i_2}) \in \mathbb{R}^{m \times m}$ and $V = (V_{j_1 j_2}) \in \mathbb{R}^{n \times n}$ can be obtained by row normalization with $U_{i_1 i_2} = A_{i_1 i_2} / \sum_{i_2=1}^{m} A_{i_1 i_2}$ and $V_{j_1 j_2} = B_{j_1 j_2} / \sum_{j_2=1}^{n} B_{j_1 j_2}$, where the row normalization is commonly used to reflect the node-level and response-level weighted average outcomes (see e.g., Lin (2010); Boucher et al. (2014); Liu et al. (2014)).

To model the relationship between Y_{ij}s using the network structure across both individuals and responses, we consider the following matrix network influence model (MNIM):

$$Y_{ij} = \sum_{i_1, j_1} \lambda_{i_1} U_{ii_1} V_{jj_1} Y_{i_1 j_1} + X_i^{(1)} \beta_1 + X_j^{(2)} \beta_2 + \epsilon_{ij}, \tag{2.1}$$

where λ_{i_1}, $\beta_1 = (\beta_{1,1}, \cdots, \beta_{1,p_1})^\top$, and $\beta_2 = (\beta_{2,1}, \cdots, \beta_{2,p_2})^\top$ are unknown parameters. We assume the random error terms $\epsilon_{ij}, 1 \le i \le m, 1 \le j \le n$ are independent with mean zero and variance σ_{ij}^2, where σ_{ij}^2 is unknown and changes with i and j. By Model (2.1), the influence effect of $Y_{i_1 j_1}$ on Y_{ij} is expressed by $\lambda_{i_1} U_{ii_1} V_{jj_1}$, which includes three components: (i) λ_{i_1}, the influential power of individual i_1 (e.g., see Zou et al. (2021)); (ii) U_{ii_1}, the interaction between individuals i and i_1; and (iii) V_{jj_1}, the interaction between responses j and j_1. Obviously, the larger λ_{i_1}, the larger the influence magnitude. Therefore, we refer to the vector $(\lambda_1, \cdots, \lambda_m)^\top \in \mathbb{R}^m$ as the matrix network influence index, through which we identify influential individuals for matrix network data formed by multi-dimensional responses with a potential high dimension going to infinity.

Remark 1. (Network Convolutional Covariates) Note that only covariates of node i and response j are included in Model (2.1). In practice, covariates of node i_1 and response j_1 that are connected with i and j, respectively, could also affect Y_{ij}. Inspired by the work of Bramoullé et al.

(2009) and Cohen-Cole et al. (2018), we consider to introduce more covariates in node and response level. This yields a generalized form of Model (2.1), which is,

$$Y_{ij} = \sum_{i_1,j_1} \lambda_{i_1} U_{ii_1} V_{jj_1} Y_{i_1j_1} + X_i^{(1)}\beta_1 + X_j^{(2)}\beta_2 + \sum_{i_1} U_{ii_1} X_{i_1}^{(1)}\beta_3 + \sum_{j_1} V_{jj_1} X_{j_1}^{(2)}\beta_4 + \epsilon_{ij},$$

where $\beta_3 = (\beta_{3,1}, \cdots, \beta_{3,p_1})^\top$ and $\beta_4 = (\beta_{4,1}, \cdots, \beta_{4,p_2})^\top$ reflect contextual effects of node i and response j, respectively. This makes the model more flexible. Since $\sum_{i_1} U_{ii_1} X_{i_1}^{(1)}$ and $\sum_{j_1} V_{jj_1} X_{j_1}^{(2)}$ can be calculated, they are treated as exogenous variables similar with $X_i^{(1)}$ and $X_j^{(2)}$. Therefore, the generalized model could be estimated through similar estimation technique. For the sake of simplicity, we focus on Model (2.1) to discuss the theoretical properties of estimators in the rest of this paper.

Remark 2. (Heterogeneous Covariate Parameters) Given that assuming homogeneous covariate parameters may be restrictive in some applications. To address this concern, a natural generalization is to allow the covariate parameters β_1 and β_2 to vary across i and j, respectively. Specifically, we consider the formulation

$$\beta_1^{(i)} = \Phi_1 \mathbf{u}_i + \alpha_1 \text{ and } \beta_2^{(j)} = \Phi_2 \mathbf{v}_j + \alpha_2,$$

as discussed in Fan et al. (2025). Here, $\mathbf{u}_i \in \mathbb{R}^m$ and $\mathbf{v}_j \in \mathbb{R}^n$ are the latent vectors that capture the heterogeneity of i and j, respectively, within the network structures implied by the similarity matrices constructed from covariates $\mathbb{X}_1$ and $\mathbb{X}_2$. The matrices $\Phi_1 \in \mathbb{R}^{p_1 \times m}$ and $\Phi_2 \in \mathbb{R}^{p_2 \times n}$ serve as loading matrices, while $\alpha_1 \in \mathbb{R}^{p_1}$ and $\alpha_2 \in \mathbb{R}^{p_2}$ represent baseline effects. This flexible specification allows each covariate to have its own parameter depending on the latent position of the corresponding individual and response. A special case of this framework assumes discrete group structures, in which $\mathbf{u}_i$ and $\mathbf{v}_j$ encode group memberships. In this setting, covariates within the same group share the same parameter, while those in different groups have distinct parameters, which is the approach we examined in the empirical analysis in Section S5.4 of the Supplementary Material.

Remark 3. (Examining Endogenous Peer Effects) In addition to identifying influential nodes, Model (2.1) can be used to examine heterogeneous endogenous peer effects. Within a social network, individuals engage in interactions that often result in their behavior correlating with the prevalence of that behavior within a reference group. Such influences are commonly conceptualized as "peer effects" in sociological discourse (Manski, 1993). For instance, Lalive and Cattaneo (2009) demonstrated the role of peer effects in children's schooling decisions; Arduini et al. (2020) further studied the heterogeneous peer effects across population subgroups. In contrast to these studies, our model accounts for heterogeneous peer effects at the individual level. In particular, in Model (2.1), the term $U_{ii_1}V_{jj_1}Y_{i_1j_1}$ represents the contribution of connected peers to the outcome Y_{ij}, and the scalar coefficient λ_{i_1} quantifies the strength of the peer effect exerted by individual i_1.

In Model (2.1), besides the $m \times n$ unknown variances, one also needs to estimate m parameters of λ. Inspired by Trusov, Bodapati, and Bucklin (2010), Zou et al. (2021), and Wu et al. (2022), we consider λ_i as a function of the associated attributes of individual i. That is, we parameterize λ_i with $\lambda_i(\gamma) = F(Z_i^\top \gamma)$, where $F(\cdot)$ is a strictly monotone and known function, $Z_i = (z_{i1}, \cdots, z_{id})^\top \in \mathbb{R}^d$ is the attribute vector with $z_{i1} \equiv 1$, and $\gamma = (\gamma_1, \cdots, \gamma_d)^\top \in \mathbb{R}^d$ is the corresponding unknown regression coefficient. Notably, this parameterization is reasonable in practice. For instance, a large fund always has greater influence than smaller funds in the fund network; similarly, a movie star in the Weibo network often has more influence than normal users.

To make the proposed Model (2.1) practically useful, it is necessary to specify the function $F(\cdot)$ that links attributes (i.e., Z_i) to influence parameters (i.e., λ_i). Motivated by Zhou et al. (2017), there are different link functions could be considered, for example: LINK I (logistic), $F(Z_i^\top \gamma) = e^{Z_i^\top \gamma}/(1 + e^{Z_i^\top \gamma})$; LINK II (inverse of log-log), $F(Z_i^\top \gamma) = 1 - e^{-e^{Z_i^\top \gamma}}$; and

LINK III (inverse of probit), $F(Z_i^\top\gamma) = \Phi(Z_i^\top\gamma)$, where $\Phi(\cdot)$ is the distribution function of the standard normal distribution.

Remark 4. (Link Function Selection) To compare the performances of candidate link functions, inspired by Hall and Pelletier (2011) and Rivers and Vuong (2002), we propose a novel link function testing statistic and establish its asymptotic distribution. Detailed theoretical results and numerical results of the test are shown in Section S4 and S5.1 of Supplementary Material, respectively.

Before concluding this section, we introduce the matrix and vector form of Model (2.1). Let $\epsilon = (\epsilon_{ij}) \in \mathbb{R}^{m\times n}$, $\mathbb{X}_1 = (X_1^{(1)\top}, \cdots, X_m^{(1)\top})^\top \in \mathbb{R}^{m\times p_1}$, $\mathbb{X}_2 = (X_1^{(2)\top}, \cdots, X_n^{(2)\top})^\top \in \mathbb{R}^{n\times p_2}$, $\mathcal{B}_1 = \beta_1 \otimes \mathbf{1}_{1\times n} \in \mathbb{R}^{p_1\times n}$, and $\mathcal{B}_2 = \beta_2 \otimes \mathbf{1}_{1\times m} \in \mathbb{R}^{p_2\times m}$, we further rewrite (2.1) in the matrix form as

$$\mathbb{Y} = U\Lambda(\gamma)\mathbb{Y}V + \mathbb{X}_1\mathcal{B}_1 + \mathcal{B}_2^\top\mathbb{X}_2^\top + \epsilon, \tag{2.2}$$

where $\Lambda(\gamma) = \text{diag}\{\lambda_1(\gamma), \cdots, \lambda_m(\gamma)\}$, and the variance structure of ϵ can be represented as $\Sigma := \text{Var}\{\text{vec}\,(\epsilon)\} = \text{diag}\{\sigma_{11}^2, \cdots, \sigma_{m1}^2, \cdots, \sigma_{1n}^2, \cdots, \sigma_{mn}^2\}$. Moreover, we define $y = \text{vec}(\mathbb{Y}) \in \mathbb{R}^{mn}$ and $e = \text{vec}(\epsilon) \in \mathbb{R}^{mn}$ as the vectorization of $\mathbb{Y}$ and ϵ, respectively. Denote $x_1 = \mathbf{1}_{n\times 1} \otimes \mathbb{X}_1 \in \mathbb{R}^{mn\times p_1}$, $x_2 = \mathbb{X}_2 \otimes \mathbf{1}_{m\times 1} \in \mathbb{R}^{mn\times p_2}$, $x = (x_1, x_2)$, and $\beta = (\beta_1^\top, \beta_2^\top)^\top$. Then, Model

(2.2) can be re-expressed in a vector form as

$$y = \left\{V^\top \otimes \bar{U}\left(\gamma\right)\right\} y + x\beta + e, \tag{2.3}$$

where $\otimes$ is the Kronecker product and $\bar{U}(\gamma) = U\Lambda(\gamma)$.

2.2 GMM estimation method

We design some moment conditions and consider to investigate the generalized method of moments (GMM) for parameter estimation (see e.g., Hansen (1982); Kelejian and Prucha (1999); Lee (2001)). Define $\theta = (\gamma^\top, \beta^\top)^\top \in \mathbb{R}^{d+p}$, $p = p_1 + p_2$, $S(\gamma) = I_{mn} - \{V^\top \otimes \bar{U}(\gamma)\}$, and $e(\theta) = S(\gamma)y - x\beta$ Denote the true parameter $\theta_0 = (\gamma_0^\top, \beta_0^\top)^\top$ and $e = S(\gamma_0)y - x\beta_0$. In the presence of heterogeneous error variances, the consistency of GMME with matrices $P_1, \cdots, P_L$ from $\mathcal{P}$, a class of matrices with $\mathrm{Diag}(P_l) = 0$, is based on the fundamental moment conditions $\mathrm{E}(e^\top P_l e) = 0$ with the results of $\mathrm{E}(e^\top P_l e) = \mathrm{tr}\{P_l \mathrm{E}(ee^\top)\} = \mathrm{tr}\{\Sigma \mathrm{Diag}(P_l)\} = 0$, for $l = 1, \cdots, L$. However, the fundamental moment conditions hold even if parameter $\theta \neq \theta_0$. As a result, another moment condition $\mathrm{E}(Q^\top e) = 0$ which holds if and only if parameter $\theta = \theta_0$ is introduced, where $Q \in \mathbb{R}^{mn \times k^\star}$ is an IV matrix constructed from x and $V^\top \otimes U$, and $k^\star$ is no smaller than the dimension of θ. Subsequently, we have $\mathrm{E}(e^\top P_1 e, \cdots, e^\top P_L e, e^\top Q)^\top = 0$ if and only if $\theta = \theta_0$, where P_l is from $\mathcal{P}$ for $l = 1, \cdots, L$. The set of moment functions for the GMM estimation hence is taken as

$$g(\theta) = \left(e(\theta)^\top P_1 e(\theta), \cdots, e(\theta)^\top P_L e(\theta), e(\theta)^\top Q\right)^\top. \tag{2.4}$$

It is noteworthy that the constraint $\mathrm{Diag}(P_l) = 0$ guarantees the correlation between each components of $P_l e$ and the corresponding component of e precisely canceled out. This helps avoid the need to estimate the unknown error variance.

Just like in the general GMM estimation framework (e.g., see Hansen (1982)), the moment functions (2.4) can be combined into a smaller set of equations $ag(\theta)$, where a is a matrix with a full row rank no smaller than the dimension of θ. In addition, following the GMM framework, the identification condition requires the unique solution of the limiting equations, $\lim_{mn\to\infty} \frac{1}{mn} a\mathrm{E}\{g(\theta)\} = 0$ at θ_0. Thus, the resulting GMME $\hat{\theta}$ can be derived from $\min_{\theta\in\Theta} g(\theta)^\top a^\top a g(\theta)$. Notably, $a^\top a$ is a non-negative definite matrix, which represents a weighting matrix in this minimization function. As a result, the estimation of a becomes the problem of selecting the optimum weighting matrix, for which we first uncover the asymptotic properties of the resulting estimators. Denote $B^s = B + B^\top$ for any square matrix B.

Define

$$\Omega = \mathrm{Var}\left\{g\left(\theta_0\right)\right\} = \begin{pmatrix} \mathrm{tr}\left\{\Sigma P_1\left(\Sigma P_1\right)^s\right\} & \mathrm{tr}\left\{\Sigma P_1\left(\Sigma P_2\right)^s\right\} & \cdots & 0 \\ \mathrm{tr}\left\{\Sigma P_2\left(\Sigma P_1\right)^s\right\} & \mathrm{tr}\left\{\Sigma P_2\left(\Sigma P_2\right)^s\right\} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & Q^\top \Sigma Q \end{pmatrix},$$

and

$$D = -\frac{\partial \mathrm{E}\left\{g\left(\theta_0\right)\right\}}{\partial \theta^\top} = -\begin{pmatrix} \mathrm{tr}\left\{\Sigma P_1^s G_{\gamma_1}\right\} & \cdots & \mathrm{tr}\left\{\Sigma P_1^s G_{\gamma_d}\right\} & 0_{1\times p} \\ \vdots & & \vdots & \vdots \\ \mathrm{tr}\left\{\Sigma P_L^s G_{\gamma_1}\right\} & \cdots & \mathrm{tr}\left\{\Sigma P_L^s G_{\gamma_d}\right\} & 0_{1\times p} \\ Q^\top G_{\gamma_1} x \beta_0 & \cdots & Q^\top G_{\gamma_d} x \beta_0 & -Q^\top x \end{pmatrix},$$

where $G_{\gamma_k} = G_{\gamma_k}(\gamma_0)$ and $S_{\gamma_k} = S_{\gamma_k}(\gamma_0)$. Here, $G_{\gamma_k}(\gamma) = S_{\gamma_k}(\gamma)S(\gamma)^{-1}$, $S_{\gamma_k}(\gamma) := \partial S(\gamma)/\partial\gamma_k = -V^\top \otimes \bar{U}_{\gamma_k}$, $\bar{U}_{\gamma_k} = U\Lambda_{\gamma_k}$, and $\Lambda_{\gamma_k} := \partial\Lambda(\gamma)/\partial\gamma_k = \mathrm{diag}\{z_{1k}F'(Z_1^\top\gamma), \cdots, z_{mk}F'(Z_m^\top\gamma)\}$ with $F'(\cdot)$ being the first order derivative of F. The following technical conditions are required to ensure the theoretical properties of the estimators.

(C1) (Error Term) The random errors ϵ_{ij}'s are independent with a mean of 0 and a variance of σ_{ij}^2. In addition, for any $1 \le i \le m$ and $1 \le j \le n$, assume that $E|\varepsilon_{ij}|^{4+\tau} < \infty$ for some $\tau > 0$.

(C2) (Covariates) The elements of the $mn \times p$ regressor matrix x are uniformly bounded, where x has the full rank p and $\lim_{mn\to\infty} \frac{1}{mn} x^\top x$

exists and is nonsingular. In addition, the $m \times d$ attribute matrix $(Z_1, \cdots, Z_m)^\top$ has the full column rank d.

(C3) (Parameter Space and Network Structure) Suppose $S(\gamma) = I_{mn} - \{V^\top \otimes \bar{U}(\gamma)\}$ is uniformly nonsingular over γ in the compact parameter space Θ. Additionally, assume that the matrix $S^{-1}(\gamma)$ and $V^\top \otimes U$ are uniformly bounded in both row and column sums in absolute value.

(C4) (Identification Condition)

(C4.1) The matrices P_l with $\text{Diag}(P_l) = 0$ are uniformly bounded in both row and column sums in absolute value, and the elements of Q are uniformly bounded.

(C4.2) Assume that either (a) $\lim_{mn\to\infty} \frac{1}{mn} Q^\top (G_{\gamma_1} x\beta, \cdots, G_{\gamma_d} x\beta, x)$ has the full rank $d{+}p$; or (b) $\lim_{mn\to\infty} \frac{1}{mn} Q^\top x$ has the full rank p, and $\lim_{mn\to\infty} \frac{1}{mn} (\text{tr}\{\Sigma P_l^s G_{\gamma_1}\}, \cdots, \text{tr}\{\Sigma P_l^s G_{\gamma_d}\})^\top$ has the full rank d for some l.

(C4.3) The true parameter θ_0 is in the interior of Θ.

Notably, the existence of the fourth or higher moments of ϵ_{ij} in Condition (C1) guarantees the variances of the quadratic forms of $e = \text{vec}(\epsilon)$ in the moment functions can be controlled. Condition (C2) assumes a nonstochastic x and its uniform boundedness for technically simplicity; otherwise, it can be replaced by finite moment conditions. Particularly, the matrix $A = (a_{ij})$ is uniformly bounded in row sums (column sums) in absolute value if the row sum matrix norm $\|A\|_\infty = \max_{1\leq i\leq n}\sum_{j=1}^{n}|a_{ij}|$ (column sum matrix norm $\|A\|_1 = \max_{1\leq j\leq n}\sum_{i=1}^{n}|a_{ij}|$) is bounded. The invertibility of $S(\gamma)$ in Condition (C3) ensures the parameters estimable, which is similar to Assumption 3 in Gupta and Robinson (2018). Condition (C3) also guarantees the spatial dependence between units can be tractable, similarly with the assumption in Kelejian and Prucha (1998). Furthermore, Condition (C4) guarantees the identification of θ_0 from the moment equations $\mathrm{E}\{g(\theta)\} = 0$ for a sufficiently large mn. In particular, a detailed interpretation of (C4.2) is carefully discussed in Section S2 of Supplementary Material, where either part (a) or (b) alone is sufficient for identification. Based on the above conditions, the asymptotic property of the GMME with a general weighting matrix $a^\top a$ is given in Theorem 1.

Theorem 1. *Under Conditions (C1)–(C4), as $mn \to \infty$, the GMME $\hat{\theta}$, derived from $min_{\theta\in\Theta} g(\theta)^\top a^\top a g(\theta)$, has an asymptotic distribution of*

$$\sqrt{mn}(\hat{\theta} - \theta_0) \xrightarrow{d} N(0, \mathcal{O}),$$

where $\mathcal{O} = \lim_{mn\to\infty}\left(\frac{1}{mn}D^\top a^\top \frac{1}{mn}aD\right)^{-1}\frac{1}{mn}D^\top a^\top \frac{1}{mn}a\Omega a^\top \frac{1}{mn}aD\left(\frac{1}{mn}D^\top a^\top \frac{1}{mn}aD\right)^{-1}$, *with D satisfying the assumption that* $\lim_{mn\to\infty}\frac{1}{mn}aD$ *exists and has the full rank $d + p$.*

Section S3.1 of Supplementary Material provides the proof for Theorem 1. To make asymptotically valid inferences from the GMME, we provide the consistent estimators of the matrices D and Ω. Let $\frac{1}{mn}\hat{D}$ and $\frac{1}{mn}\hat{\Omega}$ be estimators of $\frac{1}{mn}D$ and $\frac{1}{mn}\Omega$, respectively, with θ_0 replaced by a consistent initial estimator $\hat{\theta}$ and Σ by $\hat{\Sigma}$, where $\hat{\Sigma} = \mathrm{diag}\{\hat{\epsilon}_{11}^2, \cdots, \hat{\epsilon}_{m1}^2, \cdots, \hat{\epsilon}_{1n}^2, \cdots, \hat{\epsilon}_{mn}^2\}$, and $\hat{\epsilon}_{ij}$ are the residuals of the model, with θ_0 estimated by $\hat{\theta}$. We can show that $\frac{1}{mn}(\hat{D} - D) = o_p(1)$ and $\frac{1}{mn}(\hat{\Omega} - \Omega) = o_p(1)$.

From the proof for Theorem 1, we can conclude that, while the consistency of the GMME is not affected by the choice of the weighting matrix, its asymptotic variance is affected. Furthermore, based on the generalized Schwartz inequality, the optimal weighting matrix for the GMM estimation is $(\frac{1}{mn}\Omega)^{-1}$, recalling that Ω is the covariance matrix for the moment functions $g(\theta_0)$. To obtain the "optimal" GMME, we introduce the following regularity condition.

(C5) (Weighting Matrix) $\lim_{mn\to\infty} \frac{1}{mn}\Omega$ exists and is nonsingular.

Then, the asymptotic property of the "optimal" GMME with the optimal weighting matrix is given in Theorem 2.

Theorem 2. *Under Conditions (C1)–(C5), as $mn \to \infty$, the "optimal" GMME $\hat{\theta}_{opt}$ derived from $\min_{\theta\in\Theta} g(\theta)^\top \hat{\Omega}^{-1} g(\theta)$ has an asymptotic distribu-*

tion of

$$\sqrt{mn}(\hat{\theta}_{opt} - \theta_0) \xrightarrow{d} N(0, \mathcal{O}^{opt}),$$

where $\mathcal{O}^{opt} = \left(\lim_{mn\to\infty} \frac{1}{mn} D^\top \Omega^{-1} D\right)^{-1}$.

Section S3.2 of Supplementary Material presents the proof for Theorem 2. To employ Theorem 2 in practice, we show that $(\frac{1}{mn}\hat{D}^\top\hat{\Omega}^{-1}\hat{D})^{-1}$ is a consistent estimator of the asymptotic covariance matrix $\mathcal{O}^{opt}$ from the results of $\frac{1}{mn}(\hat{D} - D) = o_p(1)$ and $\frac{1}{mn}(\hat{\Omega} - \Omega) = o_p(1)$. When there is no special statement, we use the notation $\hat{\theta}$ as the feasible "optimal" GMME for convenience, as below.

Remark 5. (Convergence Rate) Notably, the above theorem indicates $\hat{\theta}$ is $\sqrt{mn}$-consistent, implying increasing either the network scale m or the number of responses n has the potential to improve the convergence rate. That is, either large scale network or high-dimensional responses could contribute to the efficiency of estimation. In addition, our theoretical framework is established under $mn \to \infty$, allowing for both fixed values and infinite tendencies of m and n. As a result, our theoretical results encompass the most existing works, such as the convergence rate of $\sqrt{m}$ for uni-dimensional response models with $n = 1$ and $m \to \infty$ (see e.g., Zhu et al. (2019); Zou et al. (2021); Wu et al. (2022)). To our knowledge, the study in the paper is the first attempt to address the theory and computation of network data where the network size or the dimension of the responses either remains fixed or extends towards infinity.

2.3 Homogeneity testing for GMME

After obtaining the parameter estimator of θ, we analyze the diverse effects across the influence indices $(\lambda_1, \cdots, \lambda_m)$ by assessing the homogeneity of influence in Model (2.3). Given that the associated parameterized attributes of different individual is, one may be interested in whether there exists heterogeneity among λ_is. The existing heterogeneity indicates that our proposed GMM estimation method is feasible and efficient. While the homogeneity means that all individuals share the common influence parameter in Model (2.3), in which only one parameter of λ needs to be estimated and our estimation method is no longer valid. Particularly, we consider the following test:

$$H_{0,\lambda} : \lambda_1 = \lambda_2 = \cdots = \lambda_m \text{ vs. } H_{1,\lambda} : \lambda_{i_1} \neq \lambda_{i_2} \text{ for some } i_1 \neq i_2.$$

Recalling that $\lambda_i = F(Z_i^\top \gamma)$, where $Z_i = (z_{i1}, \cdots, z_{id})^\top \in \mathbb{R}^d$ with $z_{i1} \equiv 1$ and $\gamma = (\gamma_1, \cdots, \gamma_d)^\top \in \mathbb{R}^d$ is the corresponding unknown coefficient vector. If $\gamma_2 = \cdots = \gamma_d = 0$, then $\lambda_i = F(\gamma_1)$ for $1 \leq i \leq m$; that is, λ_is are all equal. Therefore, the above hypotheses are equivalent to

$$H_{0,\gamma} : \gamma_2 = \cdots = \gamma_d = 0 \text{ vs. } H_{1,\gamma} : \gamma_i \neq 0 \text{ for some } i > 1. \tag{2.5}$$

To test the null hypothesis $H_{0,\gamma}$, in the rest of the paper, we reset $\theta = (\gamma^\top, \beta^\top)^\top$ as $\theta = (\theta_1^\top, \theta_2^\top)^\top$, where $\theta_1 = (\gamma_1, \beta^\top)^\top$ and $\theta_2 = (\gamma_2, \cdots, \gamma_d)^\top$. Denote $\hat{\theta}^{(r)}$ be the constrained "optimal" GMME of θ. The superscript (r) represents that the estimator is obtained by minimizing the function $g(\theta)^\top \hat{\Omega}^{-1} g(\theta)$ with the constraint of $\theta_2 = 0$.

Under the null hypothesis $H_{0,\gamma}$, there is no difference across the influence index $(\lambda_1, \cdots, \lambda_m)$, which is equivalent to the constraint of $\theta_2 = 0$. Denote $\hat{\Omega}^{(r)} = \text{Var}\{g(\hat{\theta}^{(r)})\}$, then the homogeneity test statistic for (2.5) can be defined as

$$T_\gamma = \left\{g(\hat{\theta}^{(r)})\right\}^\top \hat{\Omega}^{(r)^{-1}} \left\{g(\hat{\theta}^{(r)})\right\},$$

which theoretical property is obtained as below.

Theorem 3. *Under Conditions (C1)–(C5) and the null hypothesis $H_{0,\gamma}$, as $mn \to \infty$, we have*

$$T_\gamma \xrightarrow{d} \chi^2(L + k^\star - 1 - p).$$

Section S3.3 of Supplementary Material provides the proof for Theorem 3. Ultimately, Theorem 3 indicates that, under the null hypothesis $H_{0,\gamma}$, T_γ asymptotically follows a chi-squared distribution with degree $L + k^\star - 1 - p$.

In practice, $\hat{\Omega}^{(r)}$ is unknown, and it can be estimated by Ω, with θ_0 replaced by $\hat{\theta}^{(r)}$ and Σ by $\hat{\Sigma}^{(r)}$, where $\hat{\Sigma}^{(r)} = \text{diag}\{\hat{\epsilon}_{11}^{(r)2}, \cdots, \hat{\epsilon}_{m1}^{(r)2}, \cdots, \hat{\epsilon}_{1n}^{(r)2}, \cdots, \hat{\epsilon}_{mn}^{(r)2}\}$, and $\hat{\epsilon}_{ij}^{(r)}$ are the residuals of the model, with θ_0 estimated by $\hat{\theta}^{(r)}$. At the given significance level α, we reject $H_{0,\gamma}$ if $T_\gamma > \chi^2_\alpha$, where χ^2_α is the α-th upper quantile of the chi-squared distribution with degree $L + k^\star - 1 - p$. The rejection of $H_{0,\gamma}$ indicates that there is heterogeneity among λ_is.

3. Numerical Studies

3.1 Simulation models

In this section, we design several numerical experiments to demonstrate the finite sample performance of the "optimal" GMME and homogeneity test statistic. We consider the generation of weighting matrices, covariates, and heteroscedastic disturbances, respectively.

Weighting matrices. For adjacency matrices A and B, let the off-diagonals $A_{i_1 i_2}$ and $B_{j_1 j_2}$ be independent and identically generated from the Bernoulli distribution with probabilities of $0.5/m$ and $0.5/n$, respectively (e.g., see Erdös and Rényi (1959)). The diagonals A_{ii} and B_{jj} are zeros for any $1 \leq i \leq m$ and $1 \leq j \leq n$. Then, we obtain the weighted adjacency matrices U and V by normalizing the rows of A and B, respectively.

Covariates and parameters. The elements of $Z_i = (z_{i1}, z_{i2}, z_{i3})^\top$,

$X_i^{(1)} = (x_{i1}^{(1)}, x_{i2}^{(1)})$, and $X_j^{(2)} = (x_{j1}^{(2)}, x_{j2}^{(2)})$ are independent and identically generated from $N(0,1)$ except $z_{i1} \equiv 1$. Their corresponding coefficients are $\gamma = (\gamma_1, \gamma_2, \gamma_3)^\top = (0.3, 0.5, 0.4)^\top$, $\beta_1 = (\beta_{1,1}, \beta_{1,2})^\top = (0.3, 0.5)^\top$, and $\beta_2 = (\beta_{2,1}, \beta_{2,2})^\top = (0.2, 0.7)^\top$. Consequently, $\Lambda(\gamma) = \text{diag}\{F(Z_1^\top\gamma), \cdots, F(Z_m^\top\gamma)\}$, where $F(\cdot)$s are LINKs I–III from Section 2.1.

Heteroscedastic disturbances. To generate heteroscedastic disturbances, we consider group structures (e.g., see Lin and Lee (2010)). Specifically, the error terms ϵ_{ij} are generated from the normal distribution with means of 0 and different variances across groups. Based on the group structure, we consider two variance designs, V1 and V2. The details of Designs V1 and V2 are illustrated in Section S6 of Supplementary Material.

Finally, given matrices U, V, $\mathbb{X}_1 = (X_1^{(1)\top}, \cdots, X_m^{(1)\top})^\top$, $\mathbb{X}_2 = (X_1^{(2)\top}, \cdots, X_n^{(2)\top})^\top$, and the error term vector e, we generate the response vector y according to Model (2.3); that is $y = S^{-1}(\gamma)(x\beta + e)$, where $S(\gamma) = I_{mn} - \{V^\top \otimes [U\Lambda(\gamma)]\}$, $x = (\mathbf{1}_{n\times 1} \otimes \mathbb{X}_1, \mathbb{X}_2 \otimes \mathbf{1}_{m\times 1})$. In addition, let $L = 1$, $P_1 = V^\top \otimes U$ and $Q = (x, \{V^\top \otimes U\}x, \{V^\top \otimes U\}^2 x)$. All of the above settings satisfy Conditions (C1)–(C5).

3.2 Performance of GMM estimation

According to Theorem 2, we know that the "optimal" GMME $\hat{\theta}$ is consistent and asymptotically normal. To examine the asymptotic performance of the estimator, we consider different sample sizes; specifically: $(m,n) = (50,70)$, $(50,90)$, and $(100,90)$. To ensure a reliable evaluation, we consider all the results in the simulation on K=500 replications. In the k-th $(1 \le k \le K)$ realization, we denote the parameter estimator of θ_0 as $\hat{\theta}^{(k)} = (\hat{\gamma}_1^{(k)}, \hat{\gamma}_2^{(k)}, \hat{\gamma}_3^{(k)}, \hat{\beta}_{1,1}^{(k)}, \hat{\beta}_{1,2}^{(k)}, \hat{\beta}_{2,1}^{(k)}, \hat{\beta}_{2,2}^{(k)})^\top$ and its j-th component as $\hat{\theta}_j^{(k)}$. To evaluate the performance of the estimator $\hat{\theta}_j^{(k)}$, we calculate the bias as $\text{BIAS}_j = K^{-1}\sum_k(\hat{\theta}_j^{(k)} - \theta_{0j})$, and the empirical standard error as $\text{SE}_j = \{K^{-1}\sum_k(\hat{\theta}_j^{(k)} - K^{-1}\sum_k \hat{\theta}_j^{(k)})^2\}^{1/2}$. In addition, we define the average of the theoretical standard errors as $\text{tSE}_j = K^{-1}\sum_k \text{tSE}_j^{(k)}$, where $\text{tSE}_j^{(k)}$ is the theoretical standard error of θ_{0j} obtained in Theorem 2.

Table 3 presents the detailed simulation results under variance designs V1 and V2 for three link functions, respectively. We can conclude that the absolute values of BIAS, SE, and tSE generally decrease for all parameter estimates and all three link functions as the sample size mn increases. Moreover, the values of SE and tSE generally become closer as mn increases, which suggests that empirical standard error can be effectively approximated by the theoretical standard error. The above findings under both two variance designs corroborate with the results of Theorem 2.

3.3 Performance of homogeneity test statistic

Theorem 3 shows that the homogeneity test statistic T_γ is asymptotically chi-squared distributed under the null hypothesis. To investigate the performance of T_γ, we examine the empirical sizes and powers of the test statistic with a significance level of 0.05. Specifically, we set up the hypotheses $\gamma_\delta = (0.3, 0.5\delta, 0.4\delta)^\top$, with signal strength $\delta = 0, 0.1, 0.2, 0.4, 0.6, 0.8, 1$. Here, $\delta = 0$ evaluates the empirical sizes of the test, while the other values of δ are used to assess the empirical powers. Thus, the empirical size and power of the statistic can be calculated by the percentages of rejection in K replications under the hypotheses.

Figure 2 depicts the empirical sizes and powers of the homogeneity test under variance designs V1 and V2 for three link functions, respectively, with a significance level of 0.05. We can draw the following conclusions from the figures. First, the empirical sizes of T_γ are generally close to the nominal level of 0.05 for all three link functions; this demonstrates the validity of the testing procedures for the statistic. In addition, the empirical powers of T_γ increase and approach 100% when the sample size or signal strength increases. Accordingly, the proposed homogeneity test performs well in detecting heterogeneity among λ_is.

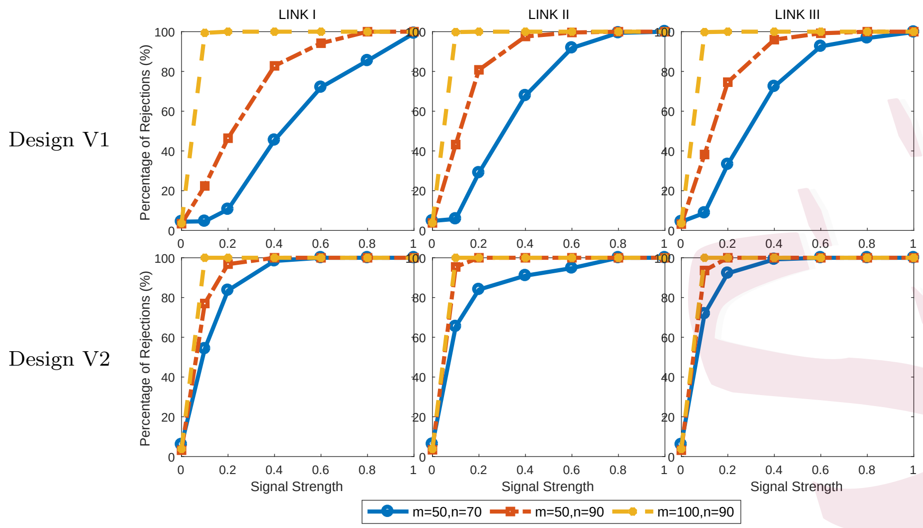

Figure 2: The empirical sizes and powers of the homogeneity test T_γ under variance designs V1 and V2 for three link functions, respectively, with a significance level of 0.05.

4. Empirical Studies

4.1 Networks and variables

To illustrate the usefulness of MNIM, we apply it to conduct a real data analysis on the spillover effect of funds' investment in stocks; that is, we explore the influence of funds via investment from a network perspective. Exploring the mechanism of the investment is crucial for both fund managers and general investors to establish profitable portfolios. While existing studies pay more attention to the characteristics of funds and stocks (see e.g., Daniel et al. (1997); Kacperczyk and Seru (2007)), they seldom consider the influence of investment from the network perspective. Combining the characteristics and network structures of funds and stocks, the proposed MNIM enables us to analyze the spillover effect flowing from one fund to another and identify influential funds via stock investment.

We focus on equity listed open-ended funds (LOF) and stocks in China's A-share market. To this end, we collect financial data on LOF from the second and third quarters of 2022 from the WIND financial database. After removing funds investing in the Hong Kong stock market and keeping only those appearing in all sample periods, there are m=74 funds in the sample. Additionally, as the WIND A50 Index comprehensively reflects the overall performance of the listed stock prices in China's A-share market, we collect data from the same period for the n=50 stocks involved in this index.

For each quarter, we define the response variable and weighted matrices as follows. Consider the dataset of the third quarter of 2022, for example. For $1 \leq i \leq m$ and $1 \leq j \leq n$, we define the response variable Y_{ij} as the investment proportion of fund i's portfolio in stock j. To assess the network influence of funds and stocks, we construct the matrices U and V as follows. For funds $1 \leq i_1, i_2 \leq m$, we define $A_{i_1 i_2} = 1$ if funds i_1 and i_2 hold at least one same stock in top ten stocks; otherwise, we define $A_{i_1 i_2} = 0$. For stocks $1 \leq j_1, j_2 \leq n$, we define $B_{j_1 j_2} = 1$ if stocks j_1 and j_2 belong to the same industry; otherwise, we define $B_{j_1j_2} = 0$. In addition, we define the diagonals A_{ii} and B_{jj} as zeros for any $1 \le i \le m$ and $1 \le j \le n$. Thus, we obtain the weighting matrices $U = (U_{i_1i_2}) \in \mathbb{R}^{m\times m}$ and $V = (V_{j_1j_2}) \in \mathbb{R}^{n\times n}$ by calculating $U_{i_1i_2} = A_{i_1i_2}/\sum_{i_2} A_{i_1i_2}$ and $V_{j_1j_2} = B_{j_1j_2}/\sum_{j_2} B_{j_1j_2}$.

Next, we evaluate covariates $X_i^{(1)}$ and $X_j^{(2)}$ based on the preceding one quarterly dataset to determine the characteristics of fund i and stock j, respectively. For $X_i^{(1)}$, we consider the following covariates: (i) *size*, the logarithm of the total market value of fund i; (ii) *volatility*, the standard deviation of the weekly returns of fund i; (iii) *return*, the return rate of the close price of fund i; and (iv) *age*, the logarithm of the age of fund i. For $X_j^{(2)}$, we consider the following covariates: (i) *size*, the logarithm of the total market value of stock j; (ii) *volatility*, the standard deviation of the weekly returns of stock j; and (iii) *return*, the return rate of the close price of stock j. Further, to measure the influential power of fund i, we use the following three attributes as covariates Z_i: (i) *size*, as defined above for fund i; (ii) *volatility*, as defined above for fund i; and (iii) *degree*, the number of funds connected to fund i. This is consistent with our sense that both size and volatility can affect the influence parameter. In addition, the influential power of nodes can also be characterized by the nodal degrees (see e.g., Carrington et al. (2005)).

4.2 Empirical results

We first use data from the previous quarter (i.e., the second quarter of 2022) to conduct the link function test to each pair of three link functions in Section 2.1, where the detailed testing procedures are presented in Section S5.3 of Supplementary Material. Next, we adopt the selected optimal link function LINK I (logistic) to identify the influential funds in the recent dataset from the third quarter of 2022.

Table 1 reports the parameter estimation and hypothesis testing results for LINK I based on the dataset from the third quarter of 2022. For covariates $X^{(1)}$ and $X^{(2)}$, the investment proportion is positively and significantly related to the size of funds and stocks with a significance level of 0.05, while the age of funds and the volatility of stocks is significantly negative. These results indicate that large funds and young funds tend to invest in stocks in the WIND A50 Index; meanwhile, the stocks in the index with large market capitalization and low volatility are more likely to be invested by the funds. For covariates Z, we first conduct the homogeneity test to assess the influential effect, leading to a p-value of 0.000. This suggests that there exists heterogeneity among influential powers of funds; that is, the spillover effect of one fund's investment on another depends on the fund's characteristics. Specifically, the coefficients of size and degree are positive and significant at a level of 0.05. This finding is consistent with common sense that a large and connected fund has more influence on other funds.

Table 1: The regression results of MNIM for LINK I with coefficient estimates, estimated standard errors, and p-values.

	Variables	Estimate	Standard Error	p-value
X_1	size	0.1543	0.0421	0.0002
	volatility	-0.2009	1.5509	0.8970
	return	0.4022	0.6964	0.5636
	age	-0.5565	0.1045	$\ll$0.0001
X_2	size	0.0002	0.0000	$\ll$0.0001
	volatility	-0.0062	0.0005	$\ll$0.0001
	return	0.2021	0.2555	0.4289
Z	intercept	0.2996	0.8299	0.7181
	size	0.4982	0.0817	$\ll$0.0001
	volatility	0.4011	1.5048	0.7899
	degree	0.2995	0.0468	$\ll$0.0001

Furthermore, to illustrate the usefulness of the matrix network influence index, we calculate $\hat{\lambda}_i = F(Z_i^\top \hat{\gamma})$ for each fund $1 \leq i \leq m$ and sort these estimators as $\hat{\lambda}_{(1)} \geq \hat{\lambda}_{(2)} \geq \cdots \geq \hat{\lambda}_{(m)}$. By applying the k-means clustering method to the sorted indices $\hat{\lambda}_{(i)}$s, we can categorize these funds into the following three clusters. Cluster I contains 56 funds with large influence effects. Cluster II consists of 15 funds with moderate influence indices. All other funds are categorized into Cluster III. To further illustrate the common characteristics within each cluster, we present the summary statistics of the two significant covariates Z, size and degree, in Table 2. It shows that the funds with great influence effects in Cluster I have largest degrees, followed by the funds in Cluster II, and finally the Cluster III. While the sizes in three clusters are almost the same. This indicates that a larger size does not necessarily lead to greater influence. This is because degree also plays a significant role in constructing the influence index. According to these classification results, fund managers and general investors may consider the strategies of the influential funds for reference to reap profits.

Table 2: The summary statistics of the covariates (size and degree) for Clusters I-III.

	Variables	Mean	Std.Dev.	Min	Median	Max
Cluster I	size	8.44	0.80	6.74	8.42	10.04
	degree	25.32	11.20	7.00	27.00	44.00
Cluster II	size	8.73	0.54	7.24	8.85	9.39
	degree	5.47	1.36	3.00	5.00	8.00
Cluster III	size	8.43	0.75	7.63	8.55	9.12
	degree	2.67	1.15	2.00	2.00	4.00

5. Concluding Remarks

In this paper, we proposed a novel MNIM to identify influential individuals with a high-dimensional response in matrix network data. We consider all potential sources of heterogeneity among individual influences and error variances. Specifically, we assigned different influence parameters to each individual, which was then parameterized with its associated attributes using a specified link function to reduce the parameter dimension. We allowed the variances of the error terms to be different across all of the observations, and employed the GMM by restricting the diagonal of weighting matrix in quadratic moments. The asymptotic properties of the estimators were investigated. Additionally, we conducted a homogeneity test to examine the heterogeneity of the influence parameters.

Considering the generalizability of the proposed model yields several interesting directions for future research. First, although we only present results for models with matrix network data, the framework can be further applied and adjusted for tensor network data, where each individual has a matrix observation. Second, our model assumes that the adjacency matrices were observed and fixed; one may consider allowing the adjacency matrix to be random or partially observed. Third, our current framework assumes a fixed number of parameters, but it may be further extended to high-dimensional cases by integrating regularization techniques to manage diverging parameters (Caner, 2009; Zhu et al., 2019). Fourth, researchers may also generalize the MNIM to a dynamic matrix network data setting. However, these extensions may involve several challenges brought by high dimensionality of the data. Nevertheless, we believe that these efforts would broaden the utility of the proposed model.

Supplementary Material

The online Supplementary Material includes six sections. Section S1 discusses the invalidity of ML estimation method. Section S2 gives the detailed interpretation of Condition (C4). Section S3 presents the proofs of Theorems. Section S4 introduces a novel link function test to check the adequacy of the pre-specified link function. Section S5 provides additional simulation studies and empirical studies. Section S6 illustrates the details of variance designs V1 and V2.

Acknowledgements

Lan's research was supported by the National Key R&D Program of China (2022YFA1003702), the National Natural Science Foundation of China (72422020, 72333001, 12171395 and 71991472) and the Joint Lab of Data Science and Business Intelligence at Southwestern University of Finance and Economics. Lin's research was supported by the National Key R&D Program of China (2022YFA1003702) and the National Natural Science Foundation of China (12426309). Huang's research was supported by the National Natural Science Foundation of China (72471230), the National Key R&D Program of China (2023YFC3304701), the Beijing Social Science Fund(23GLA008), the MOE Project of Key Research Institute of Humanities and Social Sciences (grant 22JJD110001) and Public Computing Cloud, Big Data and Responsible Artificial Intelligence for National Governance, Renmin University of China. Zhang's Research was supported by the Graduate Representative Achievement Cultivation Project of Southwestern University of Finance and Economics (JGS2024106).

References

Anagnostopoulos, A., Kumar, R. and Mahdian, M. (2008). Influence and correlation in social networks. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 7–15.

Anselin, L. (2013). *Spatial econometrics: methods and models (Vol. 4)*. Springer Science & Business Media.

Arduini, T., Patacchini, E. and Rainone, E. (2020). Treatment effects with heterogeneous externalities. *Journal of Business & Economic Statistics* **38**, 826–838.

Boucher, V., Bramoullé, Y., Djebbari, H. and Fortin, B. (2014). Do peers affect student achievement? Evidence from Canada using group size variation. *Journal of Applied Econometrics* **29**, 91–109.

Bramoullé, Y., Djebbari, H. and Fortin, B. (2009). Identification of peer effects through social networks. *Journal of Econometrics* **150**, 41–55.

Brown, S. J. and Goetzmann, W. N. (1995). Performance persistence. *The Journal of Finance* **50**, 679–698.

Caner, M. (2009). Lasso-type GMM estimator. *Econometric Theory* 25, 270–290.

Carrington, P. J., Scott, J. and Stanley Wasserman, S. (2005). *Models and methods in social network analysis*. Cambridge University Press, Cambridge.

Cohen-Cole, E., Liu, X. and Zenou, Y. (2018). Multivariate choices and identification of social interactions. *Journal of Applied Econometrics* **33**, 165–178.

Daniel, K., Grinblatt, M., Titman, S. and Wermers, R. (1997). Measuring mutual fund performance with characteristic-based benchmarks. *The Journal of Finance* **52**, 1035–1058.

Erdös, P. and Rényi, A. (1959). On Random Graphs I. *Publicationes Mathematicae* **6**, 290–297.

Fan, X., Fang, K., Lan, W. and Tsai, C.L. (2025). Network Varying Coefficient Model. *Journal of the American Statistical Association* (just-accepted), 1-32.

Fang, X. and Hu, P. J. H. (2018). Top persuader prediction for social networks. *MIS Quarterly* **42**, 63–82.

Fracassi, C. (2017). Corporate finance policies and social networks. *Management Science* **63**, 2420–2438.

Glaeser, E. L., Sacerdote, B. and Scheinkman, J. A. (1996). Crime and social interactions. *The*

Quarterly Journal of Economics **111**, 507–548.

Guiso, L., Sapienza, P. and Zingales, L. (2004). The role of social capital in financial development. *American Economic Review* **94**, 526–556.

Gupta, A. and Robinson, P. M. (2018). Pseudo maximum likelihood estimation of spatial autoregressive models with increasing dimension. *Journal of Econometrics* **202**, 92–107.

Hall, A. R. and Pelletier, D. (2011). Nonnested testing in models estimated via generalized method of moments. *Econometric Theory* **27**, 443–456.

Hansen, L. P. (1982). Large-sample properties of generalized method of moment estimators. *Econometrica* **50**, 1029–1054.

Kacperczyk, M. and Seru, A. (2007). Fund manager use of public information: New evidence on managerial skills. *The Journal of Finance* **62**, 485–528.

Kelejian, H. H. and Prucha, I. R. (1998). A generalized spatial two-stage least squares procedures for estimating a spatial autoregressive model with autoregressive disturbance. *Journal of Real Estate Finance and Economics* **17**, 99–121.

Kelejian, H. H. and Prucha, I. R. (1999). A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review* **40**, 509–533.

Lalive, R. and Cattaneo, M.A. (2009). Social interactions and schooling decisions. *The Review of Economics and Statistics* **91**, 457–477.

Lee, L. F. (2001). Generalized method of moments estimation of spatial autoregressive processes.

Unpublished manuscript, Department of Economics, Ohio State University, Columbus, OH.

Li, Y., Ma, S., Zhang, Y. and Huang, R. (2013). An improved mix framework for opinion leader identification in online learning communities. *Knowledge-Based Systems* **43**, 43–51.

Lin, X. (2010). Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservables. *Journal of Labor Economics* **28**, 825–860.

Lin, X. and Lee, L. f. (2010). GMM estimation of spatial autoregressive models with unknown heteroscedasticity. *Journal of Econometrics* **157**, 34–52.

Liu, X., Patacchini, E. and Zenou, Y. (2014). Endogenous peer effects: local aggregate or local average? *Journal of Economic Behavior & Organization* **103**, 39–59.

Ma, N. and Liu, Y. (2014). SuperedgeRank algorithm and its application in identifying opinion leader of online public opinion supernetwork. *Expert Systems With Applications* **41**, 1357–1368.

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

Newman, M. (2001). The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences* **98**, 404–409.

Ozsoylev, H. N. and Walden, J. (2011). Asset pricing in large information networks. *Journal of Economic Theory* **146**, 2252–2280.

Rivers, D. and Vuong, Q. (2002). Model selection tests for nonlinear dynamic models. *Econo-*

metrics Journal **5**, 1–39.

Trusov, M., Bodapati, A. and Bucklin, R. (2010). Determining influential users in internet social networks. *Journal of Marketing Research* **47**, 643–658.

Weng, J., Lim, E., Jiang, J. and He, Q. (2010). TwitterRank: finding topic-sensitive influential twitterers. In *Proceedings of the third ACM international conference on Web search and data mining*, 261–270.

Wu, Y., Lan, W., Zou, T. and Tsai, C. L. (2022). Inward and outward network influence analysis. *Journal of Business & Economics Statistics* **40**, 1617–1628.

Zhou, J., Tu, Y., Chen, Y. and Wang, H. (2017). Estimating spatial autocorrelation with sampled network data. *Journal of Business & Economics Statistics* **35**, 130–138.

Zhu, X., Chang, X., Li, R. and Wang, H. (2019). Portal nodes screening for large scale social networks. *Journal of Econometrics* **209**, 145–157.

Zou, T., Luo, R., Lan, W. and Tsai, C. L. (2021). Network influence analysis. *Statistica Sinica* **31**, 1727–1748.

Dongxue Zhang

Southwestern University of Finance and Economics, Chengdu, Sichuan 611130, China.

E-mail: dongxue.win@gmail.com

Wei Lan

Southwestern University of Finance and Economics, Chengdu, Sichuan 611130, China.

E-mail: lanwei@swufe.edu.cn

Danyang Huang

Renmin University of China, Beijing 100872, China.

E-mail: dyhuang89@126.com

Huazhen Lin

Southwestern University of Finance and Economics, Chengdu, Sichuan 611130, China.

E-mail: linhz@swufe.edu.cn

Table 3: The performance of the GMMEs of parameter θ under variance designs V1 and V2 for three link functions, respectively. The BIAS, SE and tSE values ($\times 10^2$) are reported for θ estimates.

			Design V1							Design V2						
	(m, n)	Measures	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\gamma}_3$	$\hat{\beta}_{1,1}$	$\hat{\beta}_{1,2}$	$\hat{\beta}_{2,1}$	$\hat{\beta}_{2,2}$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\gamma}_3$	$\hat{\beta}_{1,1}$	$\hat{\beta}_{1,2}$	$\hat{\beta}_{2,1}$	$\hat{\beta}_{2,2}$
LINK I	(50,70)	BIAS	0.35	0.62	0.33	0.00	-0.01	0.00	0.00	0.04	0.04	0.12	0.00	0.00	0.00	0.00
		SE	3.17	7.78	12.67	0.09	0.10	0.10	0.08	0.74	2.01	2.64	0.02	0.03	0.02	0.02
		tSE	3.42	8.23	12.37	0.09	0.10	0.09	0.08	0.78	1.93	2.60	0.02	0.03	0.02	0.02
	(50,90)	BIAS	-0.04	0.07	-0.12	0.00	0.00	0.00	0.00	-0.02	0.12	-0.02	0.00	0.00	0.00	0.00
		SE	1.03	2.96	2.53	0.06	0.07	0.07	0.06	0.38	1.25	0.87	0.02	0.02	0.02	0.02
		tSE	1.04	2.99	2.50	0.06	0.07	0.06	0.06	0.36	1.13	0.86	0.02	0.02	0.02	0.02
	(100,90)	BIAS	0.00	0.02	-0.04	0.00	0.00	0.00	0.00	0.00	0.02	0.03	0.00	0.00	0.00	0.01
		SE	0.53	1.20	1.18	0.03	0.03	0.03	0.02	0.33	0.73	0.81	0.01	0.02	0.02	0.01
		tSE	0.51	1.11	1.24	0.02	0.03	0.03	0.02	0.32	0.69	0.81	0.01	0.01	0.02	0.01
LINK II	(50,70)	BIAS	1.44	1.99	-0.07	0.00	0.00	0.00	0.01	0.16	0.22	0.10	0.00	0.00	0.00	0.00
		SE	6.65	10.24	11.06	0.09	0.11	0.10	0.08	1.28	2.26	2.53	0.02	0.03	0.02	0.02
		tSE	6.50	10.81	11.23	0.09	0.11	0.09	0.09	1.50	2.60	2.55	0.02	0.03	0.02	0.02
	(50,90)	BIAS	-0.03	0.03	-0.08	0.00	0.00	0.00	0.00	0.00	0.08	-0.02	0.00	0.00	0.00	0.00
		SE	0.73	2.17	1.99	0.06	0.07	0.07	0.06	0.23	0.90	0.66	0.02	0.02	0.02	0.02
		tSE	0.73	2.17	1.96	0.06	0.07	0.06	0.06	0.24	0.81	0.66	0.02	0.02	0.02	0.02
	(100,90)	BIAS	0.00	0.02	-0.03	0.00	0.00	0.00	0.00	0.01	0.03	0.04	0.00	0.00	0.00	0.00
		SE	0.45	0.99	0.96	0.03	0.03	0.03	0.02	0.30	0.61	0.69	0.01	0.02	0.02	0.01
		tSE	0.46	0.93	1.00	0.02	0.03	0.03	0.02	0.30	0.59	0.68	0.01	0.01	0.02	0.01
LINK III	(50,70)	BIAS	0.48	0.77	0.27	0.00	0.00	0.00	0.00	0.04	0.05	0.08	0.00	0.00	0.00	0.00
		SE	3.05	6.33	9.26	0.09	0.10	0.10	0.08	0.61	1.49	2.08	0.02	0.03	0.02	0.02
		tSE	3.10	6.65	9.40	0.09	0.11	0.09	0.08	0.69	1.55	2.03	0.02	0.03	0.02	0.02
	(50,90)	BIAS	-0.03	0.04	-0.09	0.00	0.00	0.00	0.00	0.00	0.08	-0.02	0.00	0.00	0.00	0.00
		SE	0.59	1.96	1.76	0.06	0.07	0.07	0.06	0.19	0.82	0.61	0.02	0.02	0.02	0.02
		tSE	0.59	1.98	1.74	0.06	0.07	0.06	0.06	0.19	0.74	0.60	0.02	0.02	0.02	0.02
	(100,90)	BIAS	0.00	0.01	-0.02	0.00	0.00	0.00	0.00	0.00	0.02	0.03	0.00	0.00	0.00	0.00
		SE	0.38	0.85	0.85	0.03	0.03	0.03	0.02	0.24	0.52	0.60	0.01	0.02	0.02	0.01
		tSE	0.37	0.80	0.89	0.02	0.03	0.03	0.02	0.24	0.50	0.59	0.01	0.01	0.02	0.01