Statistica Sinica **31** (2021), 1727-1748
doi:https://doi.org/10.5705/ss.202019.0242

NETWORK INFLUENCE ANALYSIS

Tao Zou, Ronghua Luo, Wei Lan and Chih-Ling Tsai

The Australian National University, Southwestern University of Finance and Economics and University of California, Davis

Abstract: Owing to the rapid development of social networking sites, the spatial autoregressive (SAR) model plays an important role in social network studies. However, the underlying structure of the SAR model implicitly assumes that all nodes (or actors or users) within the network have the same influential power, measured by the common autocorrelation parameter. Hence, the classical SAR model is unable to identify influential nodes. Therefore, we propose an adaptive SAR model that incorporates a network influence index, which includes the classical SAR model as a special case. Using the proposed model without imposing a specific error distribution, we apply the quasi-maximum likelihood approach to estimate the unknown parameters of the index. Then, we use these parameters to characterize the influential power of each node. We establish the asymptotic properties of the parameter estimates, and present three test statistics that we use to assess the homogeneity of the network influence indices. The usefulness of the adaptive SAR model and its associated network index is illustrated using simulation studies and an empirical investigation of the spillover effects in Chinese mutual fund cash flows.

Key words and phrases: Network influence, quasi-maximum likelihood estimation, spatial autoregressive model, weighted chi-squared test.

1. Introduction

In the last three decades, online social network sites (SNSs) have developed rapidly across different disciplines and professions. As a result, many SNSs, such as Facebook, Twitter, and Weibo, have gathered large amounts of data encompassing both users' personal information and network relationships. These important and valuable types of data have attracted considerable attention from both industry practitioners and academic researchers. For example, Wang, Yu and Wei (2012) demonstrated that advertising agencies can effectively promote new products through SNSs. Kass-Hout and Alhinnawi (2013) found that SNSs allow researchers to investigate the person-to-person spread of communicable diseases and behaviors. Ozsoylev et al. (2014) used network information to study

Corresponding author: Wei Lan, Southwestern University of Finance and Economics, Chengdu, Shichuan 611130, China. E-mail: lanwei@swufe.edu.cn.

the trading behavior of investors, and found that central investors earn higher returns. Fracassi (2017) indicated that managers' social networks can affect their corporate policy decisions. The above examples indicate how extensively social networks have been applied in practice.

To understand the network structure, we construct a network with n nodes, and set $a_{ij} = 1$ if a direct connection leads from node i to node j, and $a_{ij} = 0$ otherwise. For the sake of completeness, define $a_{jj} = 0$, for any $1 \leq j \leq n$. Accordingly, the matrix $A = (a_{ij})_{n\times n} \in \mathbb{R}^{n\times n}$, for $i, j = 1, \ldots, n$, describes the network relationships among the n nodes. In social network studies, A is called an adjacency matrix, presenting useful information on any two adjacent nodes (see, e.g., Zhu et al. (2017); Zou et al. (2017)). For node i, let Y_i be its associated response variable. To assess the influential power of each node, we can use the network structure to understand the relationships between the Y_is. Hence, we first consider the following spatial autoregressive (SAR) process, which is commonly used to model social network information:

$$Y_i = \lambda \sum_{j=1}^{n} w_{ij} Y_j + \varepsilon_i, \tag{1.1}$$

where $\lambda > 0$ is the autocorrelation (or influence) parameter, $w_{ij} = a_{ij}/\sum_{j=1}^{n} a_{ij}$, and ε_i is the random error, for $i = 1, \ldots, n$. Useful references for model (1.1) can be found in Whittle (1954); Ord (1975); LeSage and Pace (2009); Zhou et al. (2017).

Model (1.1) basically decomposes Y_i into two parts: (i) the total amount of information allocated to node i from nodes $j \neq i$ in the network, which is $\sum_{j=1}^{n} w_{ij} Y_j$, together with the influence parameter λ; and (ii) information from the outside of the network, denoted by ε_i.

Although model (1.1) is widely used to characterize the relationships between the Y_is, it is unable to identify influential nodes. This is because model (1.1) simply assumes all nodes have the same influential power, measured by the parameter λ. In practice, however, node i can have more (or less) influence than node j, for any two connected nodes i and j. Accordingly, the influence parameter can vary across nodes. To this end, let λ_j be the influence measure of node j, for $j = 1, \ldots, n$, in the network. Then, the information of node i received from node j is $Y_j w_{ij} \lambda_j$. Accordingly, we propose the following model:

$$Y_i = \sum_{j=1}^{n} Y_j \lambda_j w_{ij} + \varepsilon_i. \tag{1.2}$$

This model allows us to identify influential nodes via their associated influence measures λ_j, which has interesting real-world applications. For example, Anagnostopoulos, Kumar and Mahdian (2008) stated that, "A marketing firm, for example, can use this information to design viral marketing campaigns or give out coupons to influential nodes in the network."

From model (1.2), the influence of Y_j on Y_i is $\lambda_j w_{ij}$. Accordingly, it includes two components: (i) λ_j, which characterizes the influential power of node j; and (ii) w_{ij}, which describes the interaction between nodes i and j. When all λ_i are equal, model (1.2) reduces to the classical SAR model (1.1) (e.g., see Lee (2004); LeSage and Pace (2009)). Because model (1.2) is able to characterize the influential power of each node, we refer to it as the adaptive SAR model, and refer to its associated vector $(\lambda_1, \ldots, \lambda_n)^\top \in \mathbb{R}^n$ as the network influence index.

Note that Dou, Parrellab and Yao (2016) proposed the model $Y_i = \lambda_i \sum_{j=1}^n w_{ij} Y_j + \varepsilon_i$, and also studied influential effects. However, the λ_i in their model measures the magnitude of node i being influenced by its connected nodes. In contrast, λ_j in model (1.2) denotes node j's own influential power, which can affect its connected nodes.

In this paper, we demonstrate the novelty and usefulness of the proposed adaptive SAR model. To this end, we study the parameter estimators and their properties in the proposed model without imposing a specific error distribution. Then we make inferences on the influence index and illustrate its usefulness. We find that the adaptive SAR model can play an important role in identifying the most influential nodes, which is a key problem in social network analysis.

The rest of this paper is organized as follows. Section 2 presents the detailed adaptive SAR model structure, applies the quasi-maximum likelihood approach of Lee (2004) to estimate the unknown parameters, and explores the model's asymptotic properties. In addition, Section 2 provides three test statistics (quasi-likelihood ratio test, quasi-score test, and quasi-Wald test), which we use to compare the adaptive SAR model and the classical SAR model. This allows us to determine the contribution of the influence index. Monte Carlo studies and an empirical analysis of the Chinese mutual fund market are given in Sections 3 and 4, respectively. A short discussion and some concluding remarks are presented in Section 5. The Appendix presents five useful conditions to establish the theoretical results. The technical material, additional simulation studies, and empirical results are relegated to the Supplementary Material.

2. Models and Methodology

2.1. Models with parametrization

In addition to the network effect in model (1.2), the response Y_i can be affected by node i's own attributes. Accordingly, we extend model (1.2) as follows:

$$Y_i = \sum_{j=1}^{n} Y_j \lambda_j w_{ij} + X_i^\top \alpha + \varepsilon_i \text{ (i.e., } \mathbb{Y} = W\Lambda\mathbb{Y} + \mathbb{X}\alpha + \mathcal{E}), \tag{2.1}$$

where $X_i = (x_{i1}, \ldots, x_{ip})^\top \in \mathbb{R}^p$ represents the p-dimensional covariates associated with their corresponding attributes, $\alpha = (\alpha_1, \ldots, \alpha_p)^\top$ is a $p \times 1$ unknown regression vector, $\mathbb{Y} = (Y_1, \ldots, Y_n)^\top$, $W = (w_{ij})_{n\times n} \in \mathbb{R}^{n\times n}$, $\mathcal{E} = (\varepsilon_1, \ldots, \varepsilon_n)^\top$, $\mathbb{X} = (X_1, \ldots, X_n)^\top$, and $\Lambda = \text{diag}\{\lambda_1, \ldots, \lambda_n\}$ denotes a diagonal matrix with $\lambda_1, \ldots, \lambda_n$ as its diagonal entries. The error components ε_i of $\mathcal{E}$ are assumed to be independent and identically distributed (i.i.d.) with mean zero and finite variance σ^2.

In the adaptive SAR model (2.1), one needs to estimate n parameters of λ and p parameters of α, which is infeasible with only n observations. Note that λ_i measures node i's influential power, which should be affected by its own attributes. For example, a movie star in the Weibo network often has larger influential power than normal users. That is, the influential power of node i is affected by its vocation. To this end, let $Z_i = (z_{i1}, \ldots, z_{id})^\top \in \mathbb{R}^{d\times 1}$, $z_{i1} \equiv 1$, and $Z_{-1,i} = (z_{i2}, \ldots, z_{id})^\top$ be the $d-1$ possible attributes that may affect the influential power of node i. In addition, we assume that $\mathbb{Z}_{-1} = (Z_{-1,1}, \ldots, Z_{-1,n})^\top \in \mathbb{R}^{n\times(d-1)}$ is of full rank. Then, we parameterize the network influence index λ_i by $\lambda_i(\beta) = F(Z_i^\top \beta)$, where $F(\cdot)$ is a strictly monotone and known function, and $\beta = (\beta_1, \ldots, \beta_d)^\top \in \mathbb{R}^{d\times 1}$ is an unknown influence coefficient vector. Accordingly, $z_{i1} \equiv 1$ is associated with the intercept β_1, for $i = 1, \ldots, n$. If $\beta_2 = \cdots = \beta_d = 0$, then $\lambda_i = F(\beta_1)$; that is, the λ_i are all equal. This implies that the classical SAR model is a special case of the adaptive SAR model. Because Λ is a function of β, we further express (2.1) as

$$\mathbb{Y} = W\Lambda(\beta)\mathbb{Y} + \mathbb{X}\alpha + \mathcal{E}. \tag{2.2}$$

In the above equation, the parameter vector α is associated with the covariate matrix $\mathbb{X}$. Analogously to classical regression models, α can be interpreted as the effect of the covariate matrix $\mathbb{X}$ on the mean of the vector $\{I_n - W\Lambda(\beta)\}\mathbb{Y}$. On the other hand, the vector β is the effect of the attributes $\mathbb{Z}$ on the influence indices, $\lambda_1, \ldots, \lambda_n$.

To make the proposed model (2.2) practically useful, one needs to specify the link function $F(\cdot)$. One often assumes the influence parameter λ satisfies $|\lambda| < 1$ in the SAR model setting to ensure the invertibility of $I_n - \lambda W$ for any weighting matrix W (see, e.g., LeSage and Pace (2009)), where I_n is the $n \times n$ identity matrix. Recently, Zhou et al. (2017) indicated that nonnegative λ could provide a more precise interpretation in social network analysis. This motivates us to consider the following three link functions, which are often used in binary regression models: logistic, inverse of the probit, and inverse of the log-log. In fact, the parameter λ in the SAR model can be any value, as long as $I_n - \lambda W$ is invertible, as mentioned in Lee (2004). Hence, we adopt the inverse of the canonical link function from the Poisson regression model, and propose an exponential link function, which can be larger than one in our adaptive SAR model by requiring instead that $I_n - W\Lambda(\beta)$ in (2.2) be invertible.

The four link functions mentioned above can be summarized as follows: LINK I (logistic), $F(Z_i^\top\beta) = e^{Z_i^\top\beta}/(1+e^{Z_i^\top\beta})$; LINK II (inverse of the probit), $F(Z_i^\top\beta) = \Phi(Z_i^\top\beta)$, where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution; LINK III (inverse of the log-log), $F(Z_i^\top\beta) = 1 - e^{-e^{Z_i^\top\beta}}$; and LINK IV (exponential), $F(Z_i^\top\beta) = e^{Z_i^\top\beta}$. We next study parameter estimators for model (2.2) under a given link function.

2.2. Quasi-maximum likelihood estimation

We follow Lee (2004) approach and employ the quasi-maximum likelihood estimation (QMLE) method to estimate the unknown parameters in model (2.2). Specifically, the estimator is derived from a normal likelihood, but the random errors in model (2.2) are not required to be normally distributed, and the corresponding assumptions are stated below equation (2.1).

Define $S(\beta) = I_n - W\Lambda(\beta)$. We then have $\mathcal{E} = S(\beta)\mathbb{Y} - \mathbb{X}\alpha$. Based on the Jacobian transformation, the normal log-likelihood function of (2.2) is

$$\ell(\theta) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\{S(\beta)\mathbb{Y} - \mathbb{X}\alpha\}^\top\{S(\beta)\mathbb{Y} - \mathbb{X}\alpha\} + \log|\det\{S(\beta)\}|,$$

where $\theta = (\alpha^\top, \beta^\top, \sigma^2)^\top$. Define $\mathcal{E}(\alpha,\beta) = S(\beta)\mathbb{Y} - \mathbb{X}\alpha$, which is a function of α and β. Note that $\mathcal{E}$ is $\mathcal{E}(\alpha,\beta)$ evaluated at the true parameter values of α and β. We then adopt Lee (2004) concentrated quasi-likelihood approach to estimate the parameters. Specifically, given β, we maximize $\ell(\theta)$ with respect to α and

σ^2, which leads to

$$\hat{\alpha}(\beta) = \left(\mathbb{X}^\top \mathbb{X}\right)^{-1} \mathbb{X}^\top S(\beta)\mathbb{Y}, \text{ and}$$

$$\hat{\sigma}^2(\hat{\alpha}(\beta), \beta) = \frac{1}{n}\mathcal{E}(\hat{\alpha}(\beta), \beta)^\top \mathcal{E}(\hat{\alpha}(\beta), \beta) = \frac{1}{n}\mathbb{Y}^\top S(\beta)^\top \mathcal{M}_{\mathbb{X}} S(\beta)\mathbb{Y},$$

where $\mathcal{E}(\hat{\alpha}(\beta), \beta) = \mathcal{M}_{\mathbb{X}} S(\beta)\mathbb{Y}$ and $\mathcal{M}_{\mathbb{X}} = I_n - \mathbb{X}\left(\mathbb{X}^\top \mathbb{X}\right)^{-1} \mathbb{X}^\top$. Accordingly, the resulting concentrated quasi-log-likelihood is

$$\begin{aligned} \ell_c(\beta) &= \ell(\hat{\alpha}(\beta), \beta, \hat{\sigma}^2(\hat{\alpha}(\beta), \beta)) \\ &= -\frac{n}{2}\log(2\pi) - \frac{n}{2} - \frac{n}{2}\log \hat{\sigma}^2(\hat{\alpha}(\beta), \beta) + \log|\det\{S(\beta)\}|. \end{aligned}$$

Next, we maximize the above equation with respect to β, which yields the QMLE $\hat{\beta} = \operatorname{argmax}_\beta \ell_c(\beta)$. We then obtain the QMLEs of α and σ^2, as $\hat{\alpha} = \hat{\alpha}(\hat{\beta})$ and $\hat{\sigma}^2 = \hat{\sigma}^2(\hat{\alpha}, \hat{\beta})$, respectively. We next introduce the notation and equations used to develop the asymptotic distribution of $\hat{\theta} = (\hat{\alpha}^\top, \hat{\beta}^\top, \hat{\sigma}^2)^\top$.

Let $\Lambda_{\beta_k}(\beta) := \partial\Lambda(\beta)/\partial\beta_k = \operatorname{diag}\{z_{1k}F'(Z_1^\top\beta), \ldots, z_{nk}F'(Z_n^\top\beta)\}$, for $k = 1, \ldots, d$. In the following, we use the generic notation $(g_{k_1k_2})_{K_1\times K_2}$ to denote a matrix that has dimension $K_1 \times K_2$ and whose (k_1, k_2)th element is $g_{k_1k_2}$, for $k_1 = 1, \ldots, K_1$ and $k_2 = 1, \ldots, K_2$. After algebraic simplification, the Fisher information matrix of the quasi-log-likelihood $\ell(\theta)$ is

$$\mathcal{I}_n(\theta) := -n^{-1}\mathrm{E}\left\{\frac{\partial\ell^2(\theta)}{\partial\theta\partial\theta^\top}\right\} = \begin{pmatrix} \sigma^{-2}n^{-1}\mathbb{X}^\top\mathbb{X} & \mathcal{I}_{\alpha\beta,n} & 0_{p\times 1} \\ \mathcal{I}_{\beta\alpha,n} & \mathcal{I}_{\beta\beta,n} & \mathcal{I}_{\beta\sigma^2,n} \\ 0_{1\times p} & \mathcal{I}_{\sigma^2\beta,n} & 2^{-1}\sigma^{-4} \end{pmatrix}, \text{ where} \quad (2.3)$$

$$\begin{aligned} \mathcal{I}_{\alpha\beta,n} &= \frac{1}{n\sigma^2}\left(\mathbb{X}^\top W\Lambda_{\beta_1}(\beta)S^{-1}(\beta)\mathbb{X}\alpha, \ldots, \mathbb{X}^\top W\Lambda_{\beta_d}(\beta)S^{-1}(\beta)\mathbb{X}\alpha\right), \\ \mathcal{I}_{\beta\beta,n} &= n^{-1}\Bigg(\operatorname{tr}\left\{W\Lambda_{\beta_{k_1}}(\beta)S^{-1}(\beta)W\Lambda_{\beta_{k_2}}(\beta)S^{-1}(\beta)\right\} \\ &\quad +\operatorname{tr}\left\{W\Lambda_{\beta_{k_1}}(\beta)S^{-1}(\beta)S^{-1}(\beta)^\top\Lambda_{\beta_{k_2}}(\beta)W^\top\right\} \\ &\quad +\frac{1}{\sigma^2}\alpha^\top\mathbb{X}^\top S^{-1}(\beta)^\top\Lambda_{\beta_{k_1}}(\beta)W^\top W\Lambda_{\beta_{k_2}}(\beta)S^{-1}(\beta)\mathbb{X}\alpha\Bigg)_{d\times d}, \\ \mathcal{I}_{\beta\sigma^2,n} &= \frac{1}{n\sigma^2}\left(\operatorname{tr}\left\{W\Lambda_{\beta_1}(\beta)S^{-1}(\beta)\right\}, \ldots, \operatorname{tr}\left\{W\Lambda_{\beta_d}(\beta)S^{-1}(\beta)\right\}\right)^\top, \end{aligned}$$

$\mathcal{I}_{\beta\alpha,n} = \mathcal{I}_{\alpha\beta,n}^\top$ and $\mathcal{I}_{\sigma^2\beta,n} = \mathcal{I}_{\beta\sigma^2,n}^\top$.

Let $\circ$ be the Hadamard product of matrices, $l_n = (1, \ldots, 1)^\top \in \mathbb{R}^{n\times 1}$, and $\mathbb{X}_j = (x_{1j}, \ldots, x_{nj})^\top \in \mathbb{R}^n$, for $j = 1, \ldots, p$. Because the random error vector $\mathcal{E} = (\varepsilon_1, \ldots, \varepsilon_n)^\top$ in model (2.2) is not required to be normally distributed, the third' order moment $\mu^{(3)} = \mathrm{E}(\varepsilon_i^3)$ and the fourth' order moment $\mu^{(4)} = \mathrm{E}(\varepsilon_i^4)$

are involved in the asymptotic distribution of $\hat{\theta}$. We then denote the matrix $\mathcal{J}_n(\theta, \mu^{(3)}, \mu^{(4)})$ as follows:

$$\mathcal{J}_n(\theta, \mu^{(3)}, \mu^{(4)}) = \begin{pmatrix} 0_{p\times p} & \mathcal{J}_{\alpha\beta,n} & \dfrac{\mu^{(3)}\mathbb{X}^\top l_n}{2n\sigma^6} \\ \mathcal{J}_{\beta\alpha,n} & \mathcal{J}_{\beta\beta,n} & \mathcal{J}_{\beta\sigma^2,n} \\ \dfrac{\mu^{(3)} l_n^\top \mathbb{X}}{2n\sigma^6} & \mathcal{J}_{\sigma^2\beta,n} & \dfrac{\mu^{(4)}-3\sigma^4}{4\sigma^8} \end{pmatrix}, \text{ where}$$

$$\mathcal{J}_{\alpha\beta,n} = \frac{\mu^{(3)}}{n\sigma^4}\left(\operatorname{tr}\left[\left(\mathbb{X}_j l_n^\top\right) \circ \left\{W\Lambda_{\beta_k}(\beta)S^{-1}(\beta)\right\}\right]\right)_{p\times d}, \ \mathcal{J}_{\beta\alpha,n} = \mathcal{J}_{\alpha\beta,n}{}^\top,$$

$$\begin{aligned}\mathcal{J}_{\beta\beta,n} = {} & \frac{\mu^{(4)}-3\sigma^4}{n\sigma^4}\left(\operatorname{tr}\left[\left\{W\Lambda_{\beta_{k_1}}(\beta)S^{-1}(\beta)\right\} \circ \left\{W\Lambda_{\beta_{k_2}}(\beta)S^{-1}(\beta)\right\}\right]\right)_{d\times d} \\ & + \frac{\mu^{(3)}}{n\sigma^4}\left(\operatorname{tr}\left[\left\{W\Lambda_{\beta_{k_1}}(\beta)S^{-1}(\beta)\mathbb{X}\alpha l_n^\top\right\} \circ \left\{W\Lambda_{\beta_{k_2}}(\beta)S^{-1}(\beta)\right\}\right]\right)_{d\times d} \\ & + \frac{\mu^{(3)}}{n\sigma^4}\left(\operatorname{tr}\left[\left\{W\Lambda_{\beta_{k_2}}(\beta)S^{-1}(\beta)\mathbb{X}\alpha l_n^\top\right\} \circ \left\{W\Lambda_{\beta_{k_1}}(\beta)S^{-1}(\beta)\right\}\right]\right)_{d\times d},\end{aligned}$$

$$\begin{aligned}\mathcal{J}_{\beta\sigma^2,n} = {} & \frac{\mu^{(4)}-3\sigma^4}{2n\sigma^6}\left(\operatorname{tr}\left\{W\Lambda_{\beta_k}(\beta)S^{-1}(\beta)\right\}\right)_{d\times 1} \\ & + \frac{\mu^{(3)}}{2n\sigma^6}\left(l_n^\top W\Lambda_{\beta_k}(\beta)S^{-1}(\beta)\mathbb{X}\alpha\right)_{d\times 1},\end{aligned}$$

$\mathcal{J}_{\sigma^2\beta,n} = \mathcal{J}_{\beta\sigma^2,n}{}^\top$, and $l_n = (1, \ldots, 1)^\top \in \mathbb{R}^n$. The asymptotic distribution of $\hat{\theta}$ is given in the following theorem.

Theorem 1. *Under Conditions* (C1)–(C5) *in the Appendix,* $\sqrt{n}(\hat{\theta}-\theta)$ *is asymptotic normal with mean zero and covariance matrix* $\mathcal{I}^{-1}(\theta)+\mathcal{I}^{-1}(\theta)\mathcal{J}(\theta, \mu^{(3)}, \mu^{(4)})\,\mathcal{I}^{-1}(\theta)$*, where* $\mathcal{I}(\theta)$ *and* $\mathcal{J}(\theta, \mu^{(3)}, \mu^{(4)})$ *are stated in Condition* (C5)*, and are the convergences of matrices* $\mathcal{I}_n(\theta)$ *and* $\mathcal{J}_n(\theta, \mu^{(3)}, \mu^{(4)})$*, respectively.*

In practice, both $\mathcal{I}(\theta)$ and $\mathcal{J}(\theta, \mu^{(3)}, \mu^{(4)})$ are unknown. To make the above theorem practically useful, one needs to find their consistent estimators. Using the fact that $\mathcal{I}_n(\theta) \to \mathcal{I}(\theta)$ and $\mathcal{J}_n(\theta, \mu^{(3)}, \mu^{(4)}) \to \mathcal{J}(\theta, \mu^{(3)}, \mu^{(4)})$, we can show that the asymptotic covariance matrix $\mathcal{I}^{-1}(\theta)+\mathcal{I}^{-1}(\theta)\mathcal{J}(\theta, \mu^{(3)}, \mu^{(4)})\,\mathcal{I}^{-1}(\theta)$ can be consistently estimated by $\mathcal{I}_n^{-1}(\hat{\theta})+\mathcal{I}_n^{-1}(\hat{\theta})\mathcal{J}_n(\hat{\theta}, \hat{\mu}^{(3)}, \hat{\mu}^{(4)})\mathcal{I}_n^{-1}(\hat{\theta})$, where $\hat{\mu}^{(s)} = n^{-1}\sum_{i=1}^n \hat{\varepsilon}_i^s$, for $s = 3, 4$, and $(\hat{\varepsilon}_1, \ldots, \hat{\varepsilon}_n)^\top = \mathcal{E}(\hat{\alpha}, \hat{\beta})$.

2.3. Homogeneous influence test

After obtaining the parameter estimator $\hat{\theta}$ and its asymptotic property, we next assess the homogeneity of the influence in (2.2) by testing the effect of different influence indices λ_i. To this end, we consider the following null and alternative hypotheses:

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

According to the definition $\lambda_i(\beta) = F(Z_i^\top \beta)$, for $i = 1, \ldots, n$, the above hypotheses are equivalent to

$$H_0 : \beta_2 = \cdots = \beta_d = 0 \ \text{ vs. } \ H_1 : \text{ at least one of } \beta_2, \ldots, \beta_d \text{ is not zero}, \quad (2.4)$$

under the assumptions that the link function $F(\cdot)$ is strictly monotone and the covariate matrix $\mathbb{Z}_{-1}$ is of full rank. If one does not reject the null hypothesis, then the SAR model and its associated estimators and properties can be considered (e.g., see Lee (2004)).

Within the maximum likelihood framework, there are three commonly used tests for making inferences about β. They are the likelihood ratio test, Wald test, and score (i.e., Lagrange multiplier) test. Therefore, we use these to test (2.4). Because we consider the quasi-likelihood function and QMLE, we refer to them as the quasi-likelihood ratio test, quasi-Wald test, and quasi-score test. We first consider the quasi-likelihood ratio test. Given $\hat{\theta} = (\hat{\alpha}^\top, \hat{\beta}^\top, \hat{\sigma}^2)^\top$, we obtain the estimated quasi-log-likelihood function $\ell(\hat{\theta}) = \ell(\hat{\alpha}, \hat{\beta}, \hat{\sigma}^2)$. Under the null hypothesis of $H_0 : \beta_2 = \cdots = \beta_d = 0$, we can also obtain the constrained QMLE $\hat{\theta}^{(r)}$ and its associated quasi-log-likelihood function $\ell(\hat{\theta}^{(r)})$. Accordingly, the quasi-likelihood ratio test statistic is

$$T_{lr} = -2 \left\{ \ell(\hat{\theta}^{(r)}) - \ell(\hat{\theta}) \right\}.$$

To show the theoretical properties of T_{lr}, we introduce additional notation and equations. Let

$$\Delta_c = \begin{pmatrix} I_p & 0_{p\times 1} & 0_{p\times(d-1)} & 0_{p\times 1} \\ 0_{1\times p} & 1 & 0_{1\times(d-1)} & 0 \\ 0_{1\times p} & 0 & 0_{1\times(d-1)} & 1 \end{pmatrix} \in \mathbb{R}^{(p+2)\times(p+d+1)},$$

where $0_{K_1\times K_2}$ denotes a $K_1 \times K_2$ matrix with all its elements zeros. Let $\mathcal{I}_{11}(\theta) = \Delta_c \mathcal{I}(\theta) \Delta_c^\top$ and

$$\mathcal{I}_{11}^{-1}(\theta) = \left(\Delta_c \mathcal{I}(\theta)\Delta_c^\top\right)^{-1} =: \begin{pmatrix} \iota_{11}(\theta) & \iota_{12}(\theta) \\ \iota_{21}(\theta) & \iota_{22}(\theta) \end{pmatrix},$$

where $\iota_{11}(\theta) \in \mathbb{R}^{(p+1)\times(p+1)}$, $\iota_{12}(\theta) \in \mathbb{R}^{(p+1)\times 1}$, $\iota_{21}(\theta) \in \mathbb{R}^{1\times(p+1)}$, and $\iota_{22}(\theta) \in \mathbb{R}^{1\times 1}$. In addition, let

$$\mathcal{I}_1(\theta) = \begin{pmatrix} \iota_{11}(\theta) & 0_{(p+1)\times(d-1)} & \iota_{12}(\theta) \\ 0_{(d-1)\times(p+1)} & 0_{(d-1)\times(d-1)} & 0_{(d-1)\times 1} \\ \iota_{21}(\theta) & 0_{1\times(d-1)} & \iota_{22}(\theta) \end{pmatrix}, \tag{2.5}$$

and denote $\mathcal{K}(\theta, \mu^{(3)}, \mu^{(4)}) = \mathcal{I}(\theta) + \mathcal{J}(\theta, \mu^{(3)}, \mu^{(4)})$. Then, the asymptotic distribution of T_{lr} is given below.

Theorem 2. *Assume Conditions* (C1)–(C5) *in the Appendix hold. Under the null hypothesis H_0, the quasi-likelihood ratio test statistic T_{lr} is asymptotically distributed as $\sum_{l=1}^{p+d+1} \lambda_l(\theta, \mu^{(3)}, \mu^{(4)})\chi_l^2(1)$ as $n \to \infty$, where $\lambda_l(\theta, \mu^{(3)}, \mu^{(4)})$ is the lth largest eigenvalue of the matrix $\mathcal{K}^{1/2}(\theta, \mu^{(3)}, \mu^{(4)})\{\mathcal{I}^{-1}(\theta) - \mathcal{I}_1(\theta)\}\mathcal{K}^{1/2}(\theta, \mu^{(3)}, \mu^{(4)})$, and $\chi_l^2(1)$ are independent chi-squared random variables with one degree of freedom, for $l = 1, \ldots, (p + d + 1)$. Furthermore, under the normal assumption of $\mathcal{E}$, T_{lr} is asymptotically $\chi^2(d-1)$.*

In practice, $\lambda_l(\theta, \mu^{(3)}, \mu^{(4)})$ is unknown, and can be estimated using $\lambda_{n,l}(\hat{\theta}^{(r)}, \hat{\mu}^{(3,r)}, \hat{\mu}^{(4,r)})$, where $\lambda_{n,l}(\hat{\theta}^{(r)}, \hat{\mu}^{(3,r)}, \hat{\mu}^{(4,r)})$ is the lth largest eigenvalue of the $(p+d+1)\times(p+d+1)$ matrix $\mathcal{K}_n^{1/2}(\hat{\theta}^{(r)}, \hat{\mu}^{(3,r)}, \hat{\mu}^{(4,r)})\{\mathcal{I}_n^{-1}(\hat{\theta}^{(r)}) - \mathcal{I}_{n,1}(\hat{\theta}^{(r)})\}\mathcal{K}_n^{1/2}(\hat{\theta}^{(r)}, \hat{\mu}^{(3,r)}, \hat{\mu}^{(4,r)})$. Note first that $\mathcal{K}_n(\hat{\theta}^{(r)}, \hat{\mu}^{(3,r)}, \hat{\mu}^{(4,r)}) = \mathcal{I}_n(\hat{\theta}^{(r)}) + \mathcal{J}_n(\hat{\theta}^{(r)}, \hat{\mu}^{(3,r)}, \hat{\mu}^{(4,r)})$ is a consistent estimator of $\mathcal{K}(\theta, \mu^{(3)}, \mu^{(4)})$, second that $\mathcal{I}_{n,1}(\hat{\theta}^{(r)})$ is a consistent estimator of $\mathcal{I}_1(\theta)$, $\hat{\theta}^{(r)} = (\hat{\alpha}^{(r)\top}, \hat{\beta}^{(r)\top}, \{\hat{\sigma}^{(r)}\}^2)^\top$, and finally that $\hat{\mu}^{(s,r)} = n^{-1}\sum_{i=1}^n \{\hat{\varepsilon}_i^{(r)}\}^s$, for $s = 3, 4$, with $(\hat{\varepsilon}_1^{(r)}, \ldots, \hat{\varepsilon}_n^{(r)})^\top = \mathcal{E}(\hat{\alpha}^{(r)}, \hat{\beta}^{(r)})$.

An alternative approach to testing H_0 is the quasi-Wald test. Let

$$\Delta = \left(0_{(d-1)\times p}\ 0_{(d-1)\times 1}\ I_{d-1}\ 0_{(d-1)\times 1}\right) \in \mathbb{R}^{(d-1)\times(p+d+1)}.$$

Then, the quasi-Wald test statistic for testing H_0 can be constructed as follows:

$$T_w = (\Delta\hat{\theta})^\top \left[\Delta \left\{n^{-1}\mathcal{I}_n^{-1}(\hat{\theta})\mathcal{K}_n(\hat{\theta}, \hat{\mu}^{(3)}, \hat{\mu}^{(4)})\mathcal{I}_n^{-1}(\hat{\theta})\right\} \Delta^\top\right]^{-1} \Delta\hat{\theta},$$

and its asymptotic distribution is given below.

Corollary 1. *Assume Conditions* (C1)–(C5) *in the Appendix hold. Then, under the null hypothesis H_0, we have $T_w \xrightarrow{d} \chi^2(d-1)$ as $n \to \infty$.*

We lastly consider the quasi-score test. The advantage of this test is that we need only obtain the constrained estimator $\hat{\theta}^{(r)}$ under the null hypothesis of $H_0 : \beta_2 = \cdots = \beta_d = 0$. Specifically, the quasi-score test can be constructed as

$$T_s = n^{-1} \left\{ \frac{\partial \ell(\hat{\theta}^{(r)})}{\partial \theta} \right\}^\top \mathcal{I}_n^{-1}(\hat{\theta}^{(r)}) \frac{\partial \ell(\hat{\theta}^{(r)})}{\partial \theta}.$$

A detailed expression of $\partial \ell(\theta)/\partial \theta$ can be found in (S1.6) of the Supplementary Material. The asymptotic distribution of T_s is given in the following corollary.

Corollary 2. *Assume Conditions* (C1)–(C5) *in the Appendix hold. Under the null hypothesis H_0, the test statistic $T_s = T_{lr} + o_{\mathrm{P}}(1)$ as $n \to \infty$.*

The above corollary indicates that the quasi-score test and the quasi-likelihood ratio test are asymptotically equivalent to the weighted chi-squared distribution. To the best of our knowledge, such an asymptotic result, obtained without imposing a specific error distribution, has not been rigorously discussed in the SAR literature. It is also known that under the normal assumption, the quasi-likelihood ratio test, quasi-Wald test, and quasi-score test are all asymptotically equivalent as $n \to \infty$, whereas this may not be true under a nonnormal assumption. A good review of these three tests can be found in Rao (2005). Because the three tests can vary in terms of their finite-sample performance, we evaluate them in the following simulation studies.

3. Simulation Studies

To demonstrate the finite-sample performance of our proposed adaptive SAR model, we conduct simulation studies with various settings. Let the diagonals of the adjacency matrix A be zeros, and the off-diagonals of A be independent and identically generated from the Bernoulli distribution with probability $5/n$. Then, the weighting matrix is set to $W = (w_{ij})_{n\times n} \in \mathbb{R}^{n\times n}$, with $w_{ij} = a_{ij}/\sum_{j=1}^n a_{ij}$, for $i, j = 1, \ldots, n$. Consider the 2×1 covariate vector $X_i = (x_{i1}, x_{i2})^\top$, with $x_{i1} \equiv 1$ and x_{i2} being independent and identically generated from the standard normal distribution $N(0, 1)$, and their corresponding regression parameters being $\alpha = (\alpha_1, \alpha_2)^\top = (2, 1)^\top$. In addition, consider the 3×1 influential covariates $Z_i = (z_{i1}, z_{i2}, z_{i3})^\top$, where $z_{i1} \equiv 1$, and z_{i2} and z_{i3} are independent and identically generated from the uniform distribution $U(-0.25, 0.25)$ and the normal distribution $N(0, 0.2^2)$, respectively. Six sets of parameters $\beta = (\beta_1, \beta_2, \beta_3)^\top \in \mathbb{R}^3$ are associated with the influential covariates Z_i: $(\beta_1, \beta_2, \beta_3) = (-1, 5\varrho, -2\varrho)$, where $\varrho = 0.0$, 0.2, 0.4, 0.6, 0.8, and 1.0 measure the signal strengths of the covariates, and $\varrho = 0.0$ corresponds to the classical SAR model. As a result, the network influence matrix is $\Lambda = \text{diag}\{F(Z_1^\top \beta), \ldots, F(Z_n^\top \beta)\}$, where the link functions $F(\cdot)$ are LINKs I–IV presented in Section 2.1. Note that the above model settings satisfy the technical Conditions (C1)–(C5) in the Appendix. Finally, the response vector $\mathbb{Y}$ is generated from model (2.2) with the above setting, and its associated random error terms ε_i $(i = 1, \ldots, n)$ are independent and identically generated from four distributions: the normal distribution $N(0, \sigma^2)$; $\sigma\zeta$, where ζ follows a mixture normal distribution $0.9N(0, 5/9) + 0.1N(0, 5)$; a standardized t_3 distribution; and a standardized exponential distribution, with $\sigma^2 = 1$. This allows us to examine the robustness of the parameter estimates with respect to the error distributions.

For each setting, we consider three sample sizes: n=200, 500, and 1,000. In addition, all simulations are performed 1,000 times. To assess the performance of the parameter estimators, we define $\hat{\theta}^{(k)} = (\hat{\alpha}_1^{(k)}, \hat{\alpha}_2^{(k)}, \hat{\beta}_1^{(k)}, \hat{\beta}_2^{(k)}, \hat{\beta}_3^{(k)}, \hat{\sigma}^{2(k)})^\top \in \mathbb{R}^6$ as the vector estimate of θ obtained using the QMLE approach in the kth realization. For each component of θ, say θ_j, the averaged bias of $\hat{\theta}_j^{(k)}$, for $k = 1, \ldots, 1000$, is $\text{BIAS} = 1000^{-1} \sum_k (\hat{\theta}_j^{(k)} - \theta_j)$, and the standard deviation of $\hat{\theta}_j^{(k)}$ is $\text{SD} = \big\{1000^{-1} \sum_k (\hat{\theta}_j^{(k)} - 1000^{-1} \sum_k \hat{\theta}_j^{(k)})^2\big\}^{1/2}$. Thus, the root mean squared error is $\text{RMSE} = \sqrt{\text{SD}^2 + \text{BIAS}^2}$.

For normal random errors, Table S.1 in the Supplementary Material reports the BIAS, SD, and RMSE of the QMLEs across 1,000 realizations for the four link functions and three sample sizes. To save space, we present only the results for the setting with coefficients $(\beta_1, \beta_2, \beta_3) = (-1, 5, -2)$, because the setting with coefficients $(\beta_1, \beta_2, \beta_3) = (-1, 5\varrho, -2\varrho)$ yields similar findings for ϱ = 0.0, 0.2, 0.4, 0.6, and 0.8. From Table S.1, we find that, in general, the absolute values of BIAS and SD become smaller for all parameter estimates and for all four link functions when n gets large. It is not surprising that the RMSE shows the same pattern.

We further study the performance of the QMLE when the random errors are mixture normal, standardized t_3, and standardized exponential. Tables S.2–S.4 in the Supplementary Material indicate that the resulting estimators yield qualitatively similar conclusions to those obtained from the Gaussian errors. Hence, our estimates still exhibit nice properties under these three nonnormal cases. The above findings support our theoretical result that the QMLEs are consistent and asymptotically normal.

We next assess the finite-sample performance of the quasi-likelihood ratio test, quasi-Wald test, and quasi-score test. Note that both the quasi-likelihood ratio test statistic T_{lr} and the quasi-score test statistic T_s are asymptotically weighted chi-squared distributed with weights $\lambda_l(\theta, \mu^{(3)}, \mu^{(4)})$ under the null hypothesis. In order to conduct these two tests, we independently and identically generate $\{\chi^2_{l,m} : l = 1, \ldots, (p + d + 1)$, and $m = 1, \ldots, 10000\}$ from the chi-squared distribution with one degree of freedom. Let T be either of these two test statistics, T_{lr} or T_s. We can compute the p-values of the quasi-likelihood ratio test and the quasi-score test approximately using $p\text{-value}_1 = 10000^{-1} \sum_m I\{T > \sum_{l=1}^{p+d+1} \lambda_l(\theta, \mu^{(3)}, \mu^{(4)})\chi^2_{l,m}\}$ and $p\text{-value}_2 = 10000^{-1} \sum_m I\{T > \sum_{l=1}^{p+d+1} \lambda_{n,l} (\hat{\theta}^{(r)}, \hat{\mu}^{(3,r)}, \hat{\mu}^{(4,r)})\chi^2_{l,m}\}$, respectively. Here $\lambda_{n,l}(\hat{\theta}^{(r)}, \hat{\mu}^{(3,r)}, \hat{\mu}^{(4,r)})$ is a consistent estimator of $\lambda_l(\theta, \mu^{(3)}, \mu^{(4)})$ under the null hypothesis stated below Theorem 2, and $I\{\cdot\}$ is the indicator function. Based on our simulation results, we find that p-value$_1$ and p-value$_2$ yield very similar results under the null hypothesis. In addition, p-value$_1$ is not applicable because θ, $\mu^{(3)}$, and $\mu^{(4)}$ are unknown. As a result, we use p-value$_2$ to assess the performance of the quasi-likelihood ratio test and the quasi-score test.

We evaluate the empirical sizes of the quasi-likelihood ratio test, quasi-Wald test, and quasi-score test with significance levels ranging from 0.01 to 0.30, and examine their empirical power using the significance level 0.05. For the exponential link function under the mixture normal distribution, Figures 1 and 2 depict the size and power, respectively, when $n = 200$, 500, and 1,000. The other three link functions under the mixture normal and the other three random error distributions yield similar findings; thus, they are not presented here. The empirical size and power are the percentages of rejections under H_0 and H_1, respectively, using the hypothesis test (2.4) with 1,000 realizations. Specifically, the empirical size is the percentage of rejections under the setting of $(\beta_1, \beta_2, \beta_3) = (-1, 0, 0)$, whereas the empirical power is the percentage of rejections under the setting of $(\beta_1, \beta_2, \beta_3) = (-1, 5\varrho, -2\varrho)$, where the signal strength is $\varrho > 0$.

From Figures 1 and 2, we obtain four interesting findings. The first is that the empirical sizes of the three tests are almost identical to the predetermined significance levels as $n = 1{,}000$. The second is that the empirical power of the three tests tends to 100% when the sample size n or the signal strength ϱ increases. These two findings indicate that the three homogeneous influence tests perform well when n is large. The third is that the quasi-likelihood ratio test is sometimes oversized (anticonservative) and the quasi-Wald test is undersized (conservative) when n is not sufficiently large (see Figure 1). In contrast, the quasi-score test enables us to control the size reasonably well, especially at the significance level 0.05. Lastly, Figure 2 shows that the quasi-score and quasi-likelihood ratio tests are very similar in terms of power for different n and ϱ. However, the quasi-Wald test is not powerful when the signal strength ϱ is small. Based on the above four

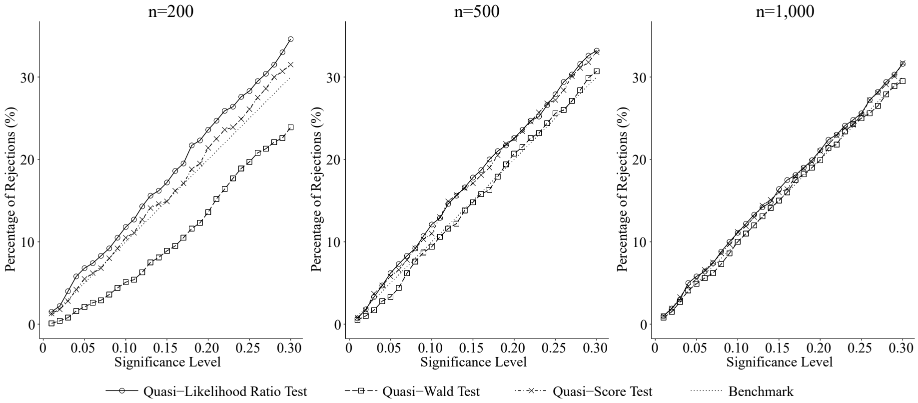

Figure 1. The empirical sizes of the three homogeneous influence tests for significance levels ranging from 0.01 to 0.30 under the setting of the exponential link function. The benchmark represents the ideal case when the percentage of rejections from 1,000 realizations is equal to the significance level. The i.i.d. random errors are simulated from $\sigma\zeta$, where ζ follows a mixture normal distribution $0.9N(0, 5/9) + 0.1N(0, 5)$.

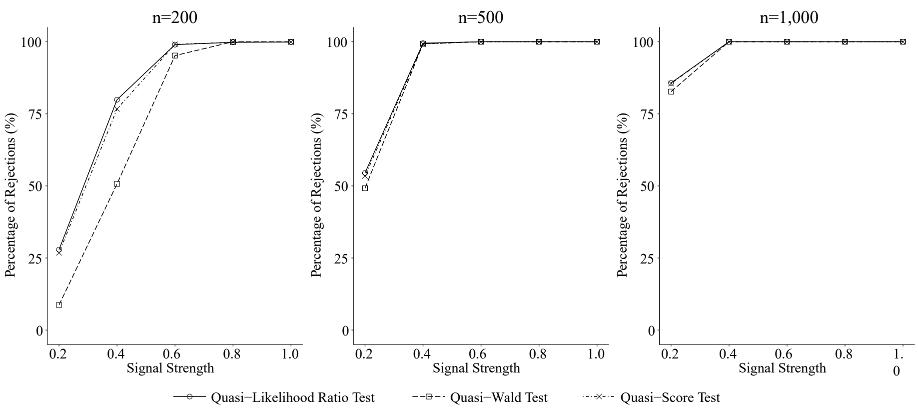

Figure 2. The empirical power of the three homogeneous influence tests at a nominal level of 0.05 under the exponential link function with 1,000 realizations. The signal strengths ϱ = 0.0, 0.2, 0.4, 0.6, 0.8, and 1.0, which correspond to the settings of $(\beta_1, \beta_2, \beta_3) = (-1, 5\varrho, -2\varrho)$. The i.i.d. random errors are simulated from $\sigma\zeta$, where ζ follows a mixture normal distribution $0.9N(0, 5/9) + 0.1N(0, 5)$.

findings, we conclude that the quasi-score test performs best at the significance level 0.05. In addition, the calculation of the quasi-score test only involves the constrained QMLE under H_0, which is easier to compute than the other two tests. Consequently, we recommend using the quasi-score test to compare the SAR and adaptive SAR models in practice, particularly when the sample size is not sufficiently large.

4. Real-Data Analysis

4.1. Network and covariates

To demonstrate the usefulness of the proposed adaptive SAR model, we present a real example of the spillover effect using Chinese mutual fund cash flows, where this effect is crucial for both fund managers and general investors (Spitz (1970); Sirri and Tufano (1998); Zheng (1999); Nanda, Wang and Zheng (2004)). For example, for fund managers, these cash flows are usually compensated for from the management fees that are charged as a fixed proportion of the total net assets under management. To explore the mechanism of cash flows, prior studies (see e.g., Spitz (1970); Sirri and Tufano (1998); Zheng (1999); Nanda, Wang and Zheng (2004); Brown and Wu (2016)) have addressed the characteristics of the mutual funds themselves, but do not consider the influence of mutual funds on cash flows from a network perspective, that is, the spillover effect. The proposed adaptive SAR model enables us to discuss this mechanism of influence from one mutual fund to another via cash flows by combining the characteristics of the fund itself and the network structure among the mutual funds.

To this end, we collect data on actively managed open-ended mutual funds in the second semi-annual period of 2015 from the WIND financial database, which is one of the most authoritative databases on the Chinese financial market. After removing funds in existence for less than one year, there are 420 mutual funds in our sample. To assess the network influence of mutual funds, we construct the network as follows. Define funds i and j as being connected (i.e., $a_{ij} = a_{ji} = 1$) if the two funds allocate at least 2.5% of their portfolios to the same stock (see Pareek (2012)). Otherwise, we consider the two funds to be disconnected; that is, $a_{ij} = 0$. As a robustness check, the allocations of funds at 1% and 5% are also considered, and yield similar results.

We next define the response variable, cash flow, as follows. The cash flow of fund i at time t, C_{it}, is calculated from the equation (Zheng (1999); Nanda, Wang and Zheng (2004))

$$C_{it} = \frac{TA_{it} - TA_{i,t-1}(1 + r_{it})}{TA_{it}},$$

where TA_{it} is the total net assets of fund i at time t, and r_{it} is the fund return at time t. To avoid the impact of outliers induced by the cash flow, we remove the top 2.5% of the observations (11 funds) from the data set. Therefore, we are left with 409 observations in our sample, and the resulting network density for these 409 funds is 20.9%. Removing the top percentage of observations is not uncommon in finance applications; for instance, Choi, Kahraman and Mukherjee (2016) proposed removing the top 2.5% of mutual funds by cash flow, and Li and Schürhoff (2019) suggested eliminating the top percentage of observations when studying financial networks. In addition, after removing those observations, the distribution of the remaining cash flow is not heavy-tailed. Thus, the moment assumption in Condition (C1) is satisfied.

In the spirit of the pioneering work of Spitz (1970), we include four control variables as X covariates to account for their effect on cash flow: (i) Size: the logarithm of the total net assets of fund i at time $t-1$; (ii) Age: the logarithm of the age of fund i at the end of $t-1$; (iii) Return: the raw return of fund i at time $t-1$; and (iv) Alpha: the risk-adjusted return of fund i, measured using the intercept of Carhart (1997) four-factor model. To quantify the influential power of the spillover effect on cash flow, we include three variables as Z covariates: (1) Size, defined above; (2) Volatility: the standard deviation of the weekly returns of fund i at time $t-1$; and (3) Degree: the number of funds connected to fund i. It seems natural that both volatility and size can be influential. We also include the degree in Z covariates. This is motivated by the empirical work of Ozsoylev et al. (2014), who found that the central investor not only performs better, but also has a larger impact on its neighbor investors. Finally, both the X and the Z covariates are standardized to have a zero mean and a unit standard deviation.

4.2. Empirical results

We fit the data using the proposed adaptive SAR model under four different link functions: exponential, logistic, inverse of the probit, and inverse of the log-log link. Their corresponding quasi-loglikelihood values evaluated at their associated QMLEs are -460.731, -463.549, -463.674, and -463.549. Motivated by Vuong (1989), we apply the exponential link function, because it has the largest estimated quasi-loglikelihood value. Based on this link function, Table 1 reports the resulting parameter estimators and their associated standard errors and t-statistics, as well as the p-values of the three homogeneous influence tests.

For the X covariates, we find that the cash flow after adjusting for influential effects is positively and significantly related to past size and raw return at the 5% significance level. For example, $\hat{\alpha}_{\text{Size}} = 0.0180$ in Table 3 implies that fund i's size has a significant and positive effect on the corresponding response Y_i (cash flow), after removing the effects of other connected cash flows. The coefficients of fund age and alpha are positive, but not significant. The above findings are consistent with those of existing research (see, e.g., Brown, Harlow and Starks (1996); Sirri and Tufano (1998); Zheng (1999)). This implies that investors tend to invest in

Table 1. The regression results of the adaptive SAR model with the exponential link function.

		Estimate	Standard-Error	t-statistic	p-value
X	Intercept	-0.1584	0.0111	-14.3178	0.0000
	Size	0.0180	0.0083	2.1717	0.0299
	Age	0.0101	0.0084	1.2053	0.2281
	Return	0.0633	0.0083	7.5817	0.0000
	Alpha	0.0107	0.0094	1.1371	0.2555
Z	Intercept	-15.2207	8.7928	-1.7310	0.0835
	Degree	4.4926	2.1233	2.1158	0.0344
	Size	0.6434	0.7159	0.8987	0.3688
	Volatility	-5.4266	2.7665	-1.9616	0.0498
	σ^2	0.0233	0.0030	7.8225	0.0000

big funds. In addition, investors pay more attention to the raw return than they do to the risk-adjusted return, because the former is more easily observed.

For the Z covariates, we employ the quasi-score test to assess the influential effect. The resulting p-value is 0.019, which indicates that the influential power of the spillover on cash flows among mutual funds indeed depend on the funds' influential characteristics. Table 1 shows three interesting findings. First, the influential power is positively and significantly related to the degree at the 5% significance level. Specifically, $\hat{\beta}_{\text{Degree}} = 4.4926$ in Table 3 indicates a significant and positive effect of fund i's degree on its influential power λ_i. This finding is not surprising because having more connections yields a bigger influential power, after controlling the size and volatility. It is also consistent with the results of Ozsoylev and Walden (2011) and Sirri and Tufano (1998). Second, the coefficient of size is positive, but not significant. Therefore, the influential power may not depend strongly on a fund's size. Lastly, the coefficient of volatility is negative and significant at the 5% significance level. This finding is consistent with the intuition that a stable fund has a larger impact on other funds.

To further illustrate the usefulness of the adaptive SAR model, we compute the estimated influence index $\hat{\lambda}_i = \exp(Z_i^\top \hat{\beta})$, for $i = 1, \ldots, n$. We then sort the $\hat{\lambda}_i$ and obtain $\hat{\lambda}_{(1)} \geq \cdots \geq \hat{\lambda}_{(n)}$. Figure 3 depicts the sorted influence indices. We next conduct a k-means clustering analysis based on the sorted $\hat{\lambda}_i$ using the R package `NbClust`. Accordingly, the best number of clusters is four, as shown in Figure 3. Cluster I includes only the mutual fund with the largest influence index. Cluster II consists of the mutual funds with the second and third largest influence indices. Cluster III consists of the mutual funds with the fourth, fifth, sixth, seventh, and eighth largest influence indices. The other mutual funds,

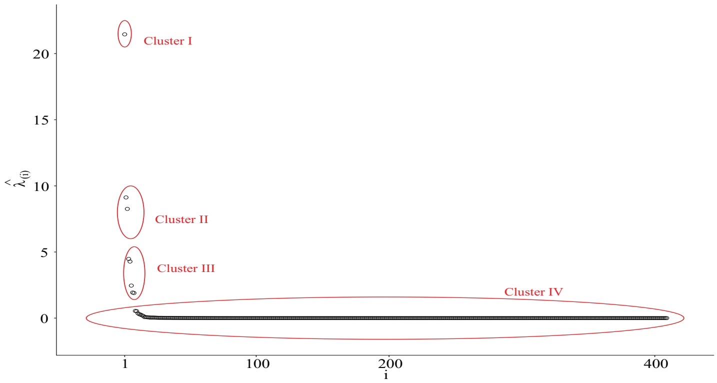

Figure 3. The sorted influence indices ($\hat{\lambda}_{(i)}$) of the $i = 1, \ldots, 409$ mutual funds.

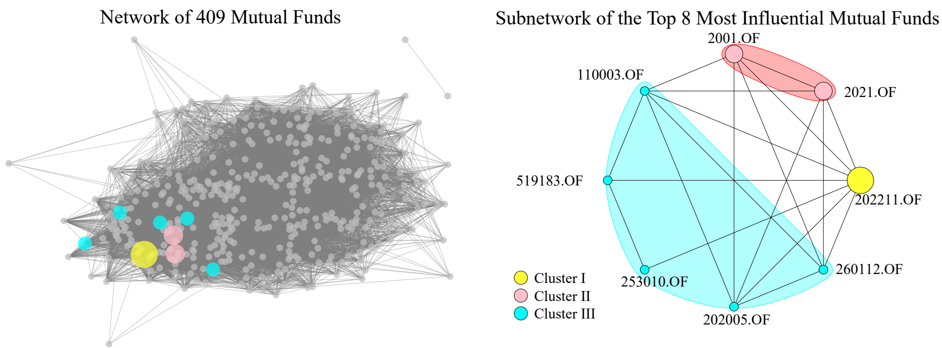

Figure 4. The network of 409 mutual funds and the subnetwork of the top eight most influential mutual funds, with their associated codes.

whose influence indices are all close to zero, are categorized into Cluster IV.

To visualize the influential power, Figure 4 depicts the four clusters in the network of 409 mutual funds. Each node in Figure 4 is a mutual fund, and we configure the node sizes from large to small to represent Clusters I–IV, respectively. Specifically, the left panel of Figure 4 reveals the whole network structure of the 409 mutual funds, and the top eight most influential mutual funds in Clusters I–III are marked in colors. The detailed subnetwork structure among these eight most influential mutual funds is presented in the right panel of Figure 4. Note that the location of each node in the left panel is constructed based on the number of the node's connections (i.e., the degree of each node). As a result, the more connections a node has, the closer it is to the center of the network. However, none of the top eight most influential mutual funds is located in the center. This indicates that a larger degree does not necessarily lead to greater influence. This is because volatility also plays a significant role in constructing the influence index. For the sake of illustration, we present the eight largest influence indices along with their two significant covariates, degree and volatility, in Table S.9 of the Supplementary Material. It shows that although the second, third, fourth, and seventh influential funds have more connections than the most influential fund, their volatilities are higher. Accordingly, the fund with the largest influence index does not have the highest number of connections. Note too that the right panel of Figure 4 indicates that the fund 202211.OF, with the largest influential power, is connected to the other top seven influential funds. We also observe that these top eight influential funds are almost all connected to each other within the network constructed by the 409 Chinese mutual funds. In summary, we have used the adaptive SAR model to effectively identify influential funds, with valuable findings.

5. Conclusion

We have proposed the adaptive SAR model and introduced an influence index for identifying influential nodes in a large network. In addition, we obtained the asymptotic properties of the parameter estimates, which allow us to make inferences on the network influence index. The usefulness of the adaptive SAR model and its associated network influence index was demonstrated using Monte Carlo studies and an application from the Chinese mutual fund market. We believe empirical finance researchers can apply the proposed model to investigate other possible factors (e.g., centrality) that may determine the influential power of individual mutual funds.

In practical applications, one usually considers positive influence parameters (e.g., Zhou et al. (2017)). However, using the fact that $F(Z_i^\top \beta) \in (0,1)$ for LINKs I–III, the transformation $G(Z_i^\top \beta) = 2F(Z_i^\top \beta) - 1$ can lead to $\lambda_i(\beta) \in (-1,1)$ if we specify $\lambda_i(\beta) = G(Z_i^\top \beta)$ in model (2.2). Thus, one can assume a negative influence index if it is needed to broaden the application of the adaptive SAR model. We next identify four avenues for future research. The first is to employ a nonparametric approach to constructing the network influence indices. The second is to use the screening or regularization method to obtain the sparse solution for constructing n influence indices λ_i (e.g., see Zhu et al. 2019a), or to develop a test statistic to test whether a subset of λ_i are equal. The third avenue is to propose a computationally feasible estimation approach (e.g., the least squares method in Huang et al. (2019) and Zhu et al. (2019b)) to overcome the computational challenge of the QMLE under large-scale networks (see the numerical illustrations in Section S.4 of the Supplementary Material). The fourth avenue is to extend the adaptive SAR model (1.2) to $Y_i = \sum_{j=1}^{n} \lambda_{ij} Y_j + \varepsilon_i$ so that the closeness between node i and node j can be characterized by the influence parameter λ_{ij}. We believe that these efforts would further increase the application of the adaptive SAR model.

Supplementary Material

The online Supplementary Material consists of four parts (Sections S.1–S.4). Section S.1 introduces five technical lemmas and their proofs. Section S.2 discusses the technical conditions in the Appendix. Section S.3 presents the proofs of the theorems and corollaries. Section S.4 provides simulation studies that assess the robustness of our proposed parameter estimate, as well as additional empirical results.

Acknowledgements

The authors sincerely thank the Editor, Associate Editor, and three anonymous referees for their constructive suggestions and comments. This research was supported by the National Natural Science Foundation of China (NSFC, 71991472, 11931014, 71873110, 71532001), National Social Science Foundation of China (19ZDA074), ANU College of Business and Economics Early Career Researcher Grant, the RSFAS Cross-Disciplinary Grant, the Joint Lab of Data Science and Business Intelligence at Southwestern University of Finance and Economics, and the UC Davis endowment fund. This research was undertaken with the assistance of computational resources provided by the Australian Government through the National Computational Infrastructure (NCI) under the ANU Merit Allocation Scheme (ANUMAS).

Appendix

A. Appendix

This Appendix introduces five useful conditions. As defined in details in Section S.1 of the Supplementary Material, $\|\cdot\|_s$ denotes the vector s-norm or the matrix s-norm for $1 \le s \le \infty$ and $|G|_\infty = \|\text{vec}(G)\|_\infty$ denotes the element-wise ℓ_∞ norm for any generic matrix G. The discussions of the following conditions are presented in Section S.2 of the Supplementary Material.

(C1) Assume that the random errors ε_i are i.i.d. with mean zero, and there exists some $\eta > 0$ such that $\mathrm{E}|\varepsilon_i|^{4+\eta} < \infty$.

(C2) Assume $\sup_{n\geq 1} \|W\|_1 < \infty$, $\sup_{n\geq 1} \|W\|_\infty < \infty$ and $\sup_{n\geq 1} |\mathbb{X}|_\infty < \infty$.

(C3) Assume that $S(\beta) = I_n - W\Lambda(\beta)$ is nonsingular uniformly over β in a compact parameter space $\mathcal{B}$ and the true parameter β is in the interior of $\mathcal{B}$. In addition, assume $\sup_{\beta\in\mathcal{B}} \sup_{n\geq 1} \|S^{-1}(\beta)\|_1 < \infty$ and $\sup_{\beta\in\mathcal{B}} \sup_{n\geq 1} \|S^{-1}(\beta)\|_\infty < \infty$.

(C4) Assume, for the true parameter β,

$$\sup_{n\geq 1} \max_{1\leq i\leq n} |z_{ik_1} F'(Z_i^\top \beta)| < \infty, \sup_{n\geq 1} \max_{1\leq i\leq n} |z_{ik_1} z_{ik_2} F''(Z_i^\top \beta)| < \infty, \text{and}$$
$$\sup_{\beta\in\mathcal{B}} \sup_{n\geq 1} \max_{1\leq i\leq n} |z_{ik_1} z_{ik_2} z_{ik_3} F'''(Z_i^\top \beta)| < \infty$$

for any $k_1, k_2, k_3 \in \{1, \ldots, d\}$, where the link function F is assumed to be three times differentiable.

(C5) Assume $\mathcal{I}_n(\theta) \to \mathcal{I}(\theta)$ and $\mathcal{J}_n(\theta, \mu^{(3)}, \mu^{(4)}) \to \mathcal{J}(\theta, \mu^{(3)}, \mu^{(4)})$ as $n \to \infty$, where $\mathcal{I}_n(\theta)$ and $\mathcal{J}_n(\theta, \mu^{(3)}, \mu^{(4)})$ are defined above Theorem 1. We further assume $\mathcal{I}(\theta)$ and $\mathcal{I}(\theta) + \mathcal{J}(\theta, \mu^{(3)}, \mu^{(4)})$ are finite and positive definite.

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.

Brown, C., Harlow, V. and Starks, T. (1996). Of tournaments and temptations: An analysis of managerial incentives in the mutual fund industry. *The Journal of Finance* **51**, 85–110.

Brown, D. P. and Wu, Y. (2016). Mutual fund flows and cross-fund learning within families. *The Journal of Finance* **71**, 383-424.

Carhart, M. (1997). On persistence in mutual fund performance. *The Journal of Finance* **52**, 57–82.

Choi, D., Kahraman, B. and Mukherjee, A. (2016). Learning about mutual fund managers. *The Journal of Finance* **71**, 2809–2860.

Dou, B., Parrellab, M. L. and Yao, Q. (2016). Generalized Yule-Walker estimation for spatio-temporal models with unknown diagonal coefficients. *Journal of Econometrics* **194**, 369–382.

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

Huang, D., Lan, W., Zhang, H. and Wang, H. (2019). Least squares estimation of spatial autoregressive models for large-scale social network. *Electronic Journal of Statistics* **13**, 1135-1165.

Kass-Hout, T. A. and Alhinnawi, H. (2013). Social media in public health. *British Medical Bulletin* **108**, 5–24.

Lee, L. F. (2004). Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica* **72**, 1899–1925.

LeSage, J. and Pace, R. K. (2009). *Introduction to Spatial Econometrics.* Chapman & Hall, New York.

Li, D. and Schürhoff, N. (2019). Dealer Networks. *The Journal of Finance* **74**, 91–144.

Nanda, N., Wang, J. and Zheng, L. (2004). Family values and the star phenomenon: Strategies of mutual fund families. *Review of Financial Studies* **17**, 667–698.

Ord, J. (1975). Estimation methods for models of spatial interaction. *Journal of the American Statistical Association* **70**, 120–126.

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

Ozsoylev, N., Walden, J., Yavuz, D. and Bildik, R. (2014). Investor networks in the stock market. *Review of Financial Studies* **27**, 1323–1366.

Pareek, A. (2012). Information networks: Implications for mutual fund trading behavior and stock returns. *SSRN Electronic Journal.* DOI: 10.2139/ssrn.1361779.

Rao, C. R. (2005). Score test: Historical review and recent developments. In *Advances in Ranking and Selection, Multiple Comparisons, and Reliability* (Edited by N. Balakrishnan, H. N. Nagaraja and N. Kannan), 3–20. Birkhäuser Boston, Cambridge.

Sirri, R. and Tufano, P. (1998). Costly search and mutual fund flows. *Journal of Economic Theory* **53**, 1589–1622.

Spitz, E. (1970). Mutual fund performance and cash inflows. *Applied Economics* **2**, 141–145.

Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica* **57**, 307–333.

Wang, X., Yu, C. and Wei, Y. (2012). Social media peer communication and impacts on purchase intentions: A consumer socialization framework. *Journal of Interactive Marketing* **16**, 198–208.

Whittle, P. (1954). On stationary processes in the plane. *Biometrika* **41**, 434–449.

Zheng, L. (1999). Is money smart? A study of mutual fund investors' fund selection ability. *The Journal of Finance* **54**, 901-933.

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

Zhu, X., Pan, R., Li, G., Liu, Y. and Wang, H. (2017). Network vector autoregression. *The Annals of Statistics* **45**, 1096–1123.

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

Zhu, X., Huang, D., Pan, R. and Wang, H. (2019b). Multivariate spatial autoregression for large scale social networks. *Journal of Econometrics* **215**, 591–606.

Zou, T., Lan, W., Wang, H. and Tsai, C.-L. (2017). Covariance regression analysis. *Journal of the American Statistical Association* **112**, 266–281.

Tao Zou

The Australian National University, Canberra ACT 0200, Australia.

E-mail: tao.zou@anu.edu.au

Ronghua Luo

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

E-mail: ronghua@swufe.edu.cn

Wei Lan
Southwestern University of Finance and Economics, Chengdu, Shichuan 611130, China.
E-mail: lanwei@swufe.edu.cn
Chih-Ling Tsai
University of California, Davis, CA 95616, USA.
E-mail: cltucd@gmail.com

(Received January 2019; accepted December 2019)