Statistica Sinica **30** (2020), 1723-1740
doi:https://doi.org/10.5705/ss.202018.0234

NETWORK GARCH MODEL

Jing Zhou, Dong Li, Rui Pan and Hansheng Wang

Renmin University of China, Tsinghua University,
Central University of Finance and Economics and Peking University

Abstract: The multivariate GARCH (MGARCH) model is popular for analyzing financial time series data. However, statistical inferences for MGARCH models are quite challenging, owing to the high dimension issue. To overcome this difficulty, we propose a network GARCH model that uses information derived from an appropriately defined network structure. This decreases the number of unknown parameters and reduces the computational complexity substantially. We also rigorously establish the strict and weak stationarity of the network GARCH model. In order to estimate the model, a quasi-maximum likelihood estimator (QMLE) is developed, and its asymptotic properties are investigated. Simulation studies are carried out to assess the performance of the QMLE in finite samples, and empirical examples are analyzed to illustrate the usefulness of network GARCH models.

Key words and phrases: GARCH model, multivariate GARCH Model, network structure, quasi-maximum likelihood estimator.

1. Introduction

Financial time series data are becoming increasingly available. A particular focus when using such data is modeling the conditional variance, which can be achieved using a number of statistical models, among which the ARCH model (Engle (1982)) and the GARCH model (Bollerslev (1986)) are popular. These models have proved to be important and powerful tools (Tsay (2003); Fan and Yao (2017)), and have been extended in numerous ways, including, but not limited to, the EGARCH, fGARCH, GARCH-M, GJR-GARCH, IGARCH, NGARCH, QGARCH, TGARCH, and ARFIMA-GARCH models. For related works on GARCH models, see Lee and Hansen (1994), Lumsdaine (1996), Hall and Yao (2003),Polonik and Yao (2008), Zhu and Ling (2011) and Conrad and Mammen (2016), among others.

The aforementioned GARCH models apply mainly to univariate time series data. However, portfolio optimization and risk management in a real capital market often require cross-sections of hundreds of different stocks. This led to

the development of multivariate GARCH (MGARCH) models (Bollerslev and Wooldridge (1988); Bollerslev (1990); Engle and Kroner (1995); Tse and Tsui (2002)), which can be used to study the dynamic relations between several stocks simultaneously. On the one hand, an MGARCH model considers one particular stock's historical returns and its conditional variance. On the other hand, it takes into account information on other stocks (e.g., the correlation or covariance). Conceptually, it is easy to extend a univariate GARCH model to the multivariate case. However, statistical inferences for MGARCH models are quite challenging, because the parameters of interest are all included in the inverse conditional covariance matrix during the estimation procedure. In particular, the number of parameters increases rapidly with the number of stocks, making the resulting estimates highly unstable. Furthermore, the positive definiteness of the conditional covariance matrix has to be ensured. To make the model more applicable in practice, additional structures need to be imposed on the covariance matrix.

The main challenge in MGARCH modeling is to impose a realistic, but parsimonious specification to make the covariance matrix positive definite. Popular methods of doing so include imposing certain structures on the conditional covariance matrix (Bollerslev (1990); Engle and Kroner (1995); Tse and Tsui (2002)) or applying dimension-reduction techniques (Pan, Polonik and Yao (2010); Lam and Yao (2012); Li et al. (2016)). Popular models in the former group include the CCC-GARCH models (Bollerslev (1990); Tse (2000)) and DCC-GARCH models (Tse and Tsui (2002)). In the latter group, factor modeling has been used extensively by introducing both observed factors (Engle, Ng and Rotschild (1990); Bollerslev and Engle (1993); Tao et al. (2012)) and unobserved factors (Pan, Polonik and Yao (2010); Hu and Tsay (2013); Li et al. (2016)). In this study, we develop a new method for an MGARCH model, based on network structure data. We call the proposed model the network GARCH model.

Network analyses have been incorporated successfully in many fields, including sociology, marketing, and organization behavior, among others. Empirical findings show that model performance can be improved significantly by incorporating network structure information (Goel and Goldstein (2014); Nitzan and Libai (2011); Wei et al. (2014)). In a real stock market, investors can select different stocks to diversify their investment risk. However, stocks in which investors have a particularly large holding always indicate unique information (Livingston (2010); Pareek (2012); Chou and Lee (2012); Bajo, Crociand Marinelli (2017)). Often, if two stocks have a common shareholder, their financial performance (i.e.,

stock returns) is highly correlated. Therefore, common stock shareholdings can be used to construct an information network. Furthermore, the network structure will be reflected in the conditional covariance matrix. This will lead to a highly parsimonious model and a parameter dimension reduction.

The proposed model is distinct from existing models in the following ways. First, in contrast to the traditional GARCH model, the network GARCH model uses information on other stocks by defining an appropriate network structure. Second, the network GARCH model can process large numbers of stocks in terms of volatility forecasting. Third, the number of unknown parameters in the network GARCH model is substantially lower, owing to the network structure. Fourth, the computational complexity drops from $O(N^2)$ to $O(N)$. Fifth, the strictly stationary solution for the new model can be established rigorously under certain conditions. Lastly, in order to estimate network GARCH model, we propose a quasi-maximum likelihood estimator (QMLE) and investigate its asymptotics.

The rest of the article is organized as follows. Section 2 presents the network GARCH model, including its strict stationarity solution and asymptotic properties. Section 3 shows the finite-sample performance of the network GARCH model using extensive numerical studies of both simulated and real data sets. Section 4 concludes the paper. All proofs are relegated to the online Supplementary Material.

2. The Network GARCH Model

2.1. Model setup

Let i be the stock index in the stock market, where $1 \leq i \leq N$. Here, N denotes the number of stocks, and is fixed in our model. For stock i, we assume a continuous response variable $y_{it} \in \mathbb{R}^1$ is observed for $t = 1, \ldots, T$. For example, y_{it} may denote the return on stock i at time t in the stock market. Following Engle (1982) and Bollerslev (1986), a classical GARCH(1,1) model is defined as

$$y_{it} = \varepsilon_{it}\sqrt{h_{it}}, \quad h_{it} = \omega_0 + \alpha_0 y_{i,t-1}^2 + \beta_0 h_{i,t-1}, \tag{2.1}$$

where $\{\varepsilon_{it}\}$ is a sequence of independent and identically distributed (i.i.d.) random variables, with $E(\varepsilon_{it}) = 0$ and $\text{var}(\varepsilon_{it}) = 1$. In addition, $\omega_0 > 0$, $\alpha_0 \geq 0$, and $\beta_0 \geq 0$ are unknown parameters, and are interpreted as volatility parameters. Lastly, h_{it} is the conditional variance. Bollerslev (1986) showed that model (2.1) defines a second-order stationary solution if and only if $\alpha_0 + \beta_0 < 1$. Later,

Nelson (1990) proved that there exists a unique strictly stationary solution to model (2.1) if and only if $E\log(\beta_0 + \alpha_0\varepsilon_{it}^2) < 0$.

To take the network structure into consideration, we first define an adjacency matrix $A = (a_{ij}) \in \mathbb{R}^{N\times N}$, where $a_{ij} = 1$ if stock i is connected to stock j, and $a_{ij} = 0$ otherwise. Let $a_{ii} = 0$ for $1 \le i \le N$. In our empirical example, we assume that two stocks are connected (i.e., $a_{ij} = a_{ji} = 1$) if they share at least one common shareholder among their respective top 10 shareholders. In order to model this network dependence structure, we incorporate a new term in model (2.1). This yields the network GARCH model

$$y_{it} = \varepsilon_{it}\sqrt{h_{it}}, \quad h_{it} = \omega_0 + \alpha_0 y_{i,t-1}^2 + \lambda_0\, d_i^{-1}\sum_{j\neq i} a_{ij}y_{j,t-1}^2 + \beta_0 h_{i,t-1}, \tag{2.2}$$

where $\sum_{j\neq i}$ represents $\sum_{j=1,j\neq i}^{N}$, and $d_i = \sum_{j=1}^{N} a_{ij}$ is the total number of stocks to which i connects, which is the out-degree. The coefficient associated with the out-degree is λ_0. If $d_i = 0$ for some i, stock i does not have an out-degree, and is regarded as isolated. In this case, we define $d_i^{-1}\sum_{j\neq i} a_{ij}y_{j,t-1}^2 = 0$, following convention. This idea is similar to, but not the same as that in the network VAR model (Zhu et al. (2017)). Although the latter model includes network structure information in its specification, the network VAR fits the conditional mean, whereas the proposed model fits the conditional variance. In contrast to the traditional MGARCH model, we implicitly assume that stock i is affected only by its direct connected neighbors (i.e., $a_{ij} = 1$). This is typically true in practice, because the activities of j with $a_{ij} = 0$ cannot be observed by i. Thus, λ_0 captures the average influence of other stocks on stock i, which we interpret as the *network effect*. The assumption that a stock experiences only an average effect of all its connected neighbors may have some limitations, thus limiting the applicability of the proposed network GARCH model. For example, the model is particularly suitable when the stocks belong to the same industry. In the next subsection, we derive a strict stationarity solution for model (2.2).

2.2. Strict stationarity

Here, we derive a stationary solution for model (2.2). For simplicity, we define $\mathbf{y}_t = (y_{1t},\ldots,y_{Nt})'$, $\mathbf{h}_t = (h_{1t},\ldots,h_{Nt})'$, $\mathbf{D} = \mathrm{diag}(d_1,\ldots,d_N)$, $\mathcal{E}_t = \mathrm{diag}(\varepsilon_{1t}^2,\ldots,\varepsilon_{Nt}^2)$, and $\mathbf{B}_t = \beta_0\mathbf{I}_N + \alpha_0\mathcal{E}_t + \lambda_0\mathbf{D}^{-1}A\mathcal{E}_t$, where $\mathbf{I}_N$ denotes an $N\times N$ identity matrix. Then, model (2.2) can be rewritten in vector form as

$$\mathbf{h}_t = \omega_0\mathbf{1}_N + \mathbf{B}_{t-1}\mathbf{h}_{t-1}, \tag{2.3}$$

where $\mathbf{1}_N = (1, \ldots, 1)'$ is a vector with a compatible dimension. The top Lyapunov exponent associated with $\mathbf{h}_t$ in (2.3) is defined as

$$\gamma_0 = \inf \left\{ \frac{1}{n} E \log \|\mathbf{B}_n \mathbf{B}_{n-1} \cdots \mathbf{B}_1\|_*, \quad n \in \mathbb{N} \right\},$$

where $\mathbb{N}$ is the set of natural numbers and $\|\cdot\|_*$ defines the operator norm of $N \times N$ matrices. By Theorem 3.2 in Bougerol and Picard (1992), $\mathbf{h}_t$ converges almost surely (a.s.), and the process (y_{it}) in model (2.2) has a unique strictly stationary solution if and only if $\gamma_0 < 0$. Under this condition, we have the following expression:

$$\mathbf{h}_t = \omega_0 \mathbf{1}_N + \omega_0 \sum_{j=1}^{\infty} \left\{ \prod_{i=1}^{j} \mathbf{B}_{t-i} \right\} \mathbf{1}_N. \tag{2.4}$$

However, γ_0 is closely related to the distribution of ε_{it}. Although it can be simulated using the Monte Carlo method, it is not easy to calculate in practice. In what follows, we give a sufficient condition to ensure $\gamma_0 < 0$; see Theorem 1 and the proof in Appendix A.

Theorem 1. *If $\alpha_0 + \lambda_0 + \beta_0 < 1$, there exists a unique strictly stationary solution with finite second moments for model (2.2); that is, $E\|\mathbf{y}_t\|^2 < \infty$, where $\|\cdot\|$ is the Euclidean norm. In particular, under Assumption 1 below, the long-run variance-covariance matrix of $\mathbf{y}_t$ is*

$$\Sigma_{\mathbf{y}} = \operatorname{diag}(\omega_0[(1 - \alpha_0 - \beta_0)\mathbf{I}_N - \lambda_0 \mathbf{D}^{-1} A]^{-1} \mathbf{1}_N).$$

Remark 1. Theorem 1 is constructed with a fixed N, which will break if N is diverging because the dimension of $\mathbf{y}_t$ is increasing. In this case, it seems a meaningful concept of stationarity is highly questionable. As a result, if N is diverging, Theorem 1 not only breaks, but is also not fixable. However, as shown in the following real-data analysis, our model can cope with a dimension much higher than those of conventional models.

2.3. Quasi-maximum likelihood estimator

Assume that observations $(\mathbf{y}_1, \ldots, \mathbf{y}_T)$ are from model (2.2) with true value $\theta_0 = (\omega_0, \alpha_0, \lambda_0, \beta_0)' \in \mathbb{R}^4$. Here, T is the sample size. Let $\theta = (\omega, \alpha, \lambda, \beta)' \in \mathbb{R}^4$ be the parameter. The quasi-log-likelihood function (ignoring a constant) is given by

$$\tilde{L}(\theta) = \frac{1}{T}\sum_{t=1}^{T} \tilde{\ell}_t(\theta), \quad \tilde{\ell}_t(\theta) = \frac{1}{N}\sum_{i=1}^{N} \left\{ \log \tilde{\sigma}_{it}^2(\theta) + \frac{y_{it}^2}{\tilde{\sigma}_{it}^2(\theta)} \right\},$$

where $\tilde{\sigma}_{it}^2(\theta)$ is defined recursively for $t \geq 1$ by

$$\tilde{\sigma}_{it}^2(\theta) = \omega + \alpha y_{i,t-1}^2 + \lambda\, d_i^{-1} \sum_{j \neq i} a_{ij} y_{j,t-1}^2 + \beta \tilde{\sigma}_{i,t-1}^2(\theta),$$

with $\tilde{\sigma}_{i0}^2(\theta) \equiv 0$.

The QMLE is defined as

$$\widehat{\theta} = (\widehat{\omega}, \widehat{\alpha}, \widehat{\lambda}, \widehat{\beta})' = \arg\min_{\theta \in \Theta} \tilde{L}(\theta), \tag{2.5}$$

where Θ is the parameter space.

To discuss the asymptotic properties of $\widehat{\theta}$, it is convenient to approximate the sequence $\{\tilde{\sigma}_{it}^2(\theta)\}$ by an ergodic stationary sequence $\{\sigma_{it}^2(\theta)\}$, which is defined as

$$\sigma_{it}^2(\theta) = \omega + \alpha y_{i,t-1}^2 + \lambda\, d_i^{-1} \sum_{j \neq i} a_{ij} y_{j,t-1}^2 + \beta \sigma_{i,t-1}^2(\theta), \tag{2.6}$$

for any t and each i. In addition, $\sigma_{it}^2(\theta_0) = h_{it}$. Similarly to the definitions of $\tilde{L}(\theta)$ and $\tilde{\ell}_t(\theta)$, we can define

$$L(\theta) = \frac{1}{T}\sum_{t=1}^{T} \ell_t(\theta), \quad \ell_t(\theta) = \frac{1}{N}\sum_{i=1}^{N} \left\{ \log \sigma_{it}^2(\theta) + \frac{y_{it}^2}{\sigma_{it}^2(\theta)} \right\}. \tag{2.7}$$

Before stating our main results, we give two assumptions that are standard in studies of GARCH-type models.

Assumption 1. $\{\varepsilon_{it}\}$ is i.i.d. across i and t with zero mean and unit variance. Furthermore, assume ε_{it}^2 is nondegenerate.

Assumption 2. The parameter space Θ is a compact subset of $\{\theta : \omega > 0, \alpha > 0, \lambda > 0, \beta > 0, \alpha + \lambda + \beta < 1\}$ and $\theta_0 \in \Theta$.

The following theorem states the strong consistency and asymptotic normality of the QMLE $\widehat{\theta}$.

Theorem 2. *If Assumptions* 1–2 *hold, then* $\widehat{\theta} \xrightarrow{a.s.} \theta_0$ *as* $T \to \infty$. *Furthermore, if* $\kappa_4 = E\varepsilon_{it}^4 < \infty$ *and* θ_0 *is an interior point of* Θ, *then as* $T \to \infty$,

$$\sqrt{NT}\big(\widehat{\theta}-\theta_0\big)\to_d \mathcal{N}(0,(\kappa_4-1)\Sigma^{-1}),$$

where

$$\Sigma=\frac{1}{N}\sum_{i=1}^{N}E\left(\frac{1}{h_{it}^2}\frac{\partial\sigma_{it}^2(\theta_0)}{\partial\theta}\frac{\partial\sigma_{it}^2(\theta_0)}{\partial\theta'}\right),$$

$$\frac{\partial\sigma_{it}^2(\theta_0)}{\partial\theta}=\left(1,y_{i,t-1}^2,d_i^{-1}\sum_{j\neq i}a_{ij}y_{j,t-1}^2,\sigma_{i,t-1}^2\right)'+\beta_0\frac{\partial\sigma_{i,t-1}^2(\theta_0)}{\partial\theta}.$$

The adjacency matrix is used to calculate the parameter derivative. Given an initial guess of θ_0, we repeat the above process until we obtain a convergent θ. To make a statistical inference for θ_0, we need to estimate κ_4 and Σ. In practice, κ_4 can be consistently estimated as $\widehat{\kappa}_4=(NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{\varepsilon}_{it}^4$, where $\hat{\varepsilon}_{it}=(\hat{h}_{it})^{-1/2}y_{it}$. A consistent estimator of Σ is its sample counterpart:

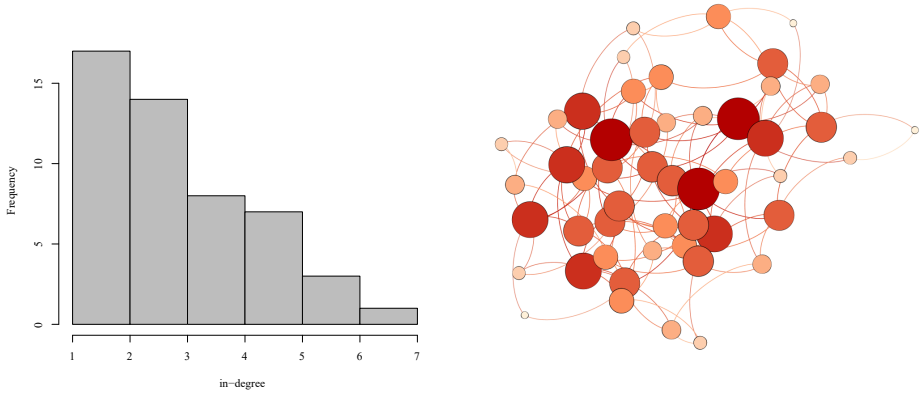
$$\widehat{\Sigma}=\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\Big(\frac{1}{\tilde{\sigma}_{it}^4(\widehat{\theta})}\frac{\partial\tilde{\sigma}_{it}^2(\widehat{\theta})}{\partial\theta}\frac{\partial\tilde{\sigma}_{it}^2(\widehat{\theta})}{\partial\theta'}\Big).$$

Remark 2. MGARCH models have been well studied. Therefore, our work provides little contribution to the parameter consistency theory of a "general" MGARCH model. Instead, we propose a special form of MGARCH model that takes into account an observed network structure. As a result, the number of unknown parameters is substantially reduced, enabling us to cope with a dimension much higher than those of conventional models. However, the drawback of our model is that it cannot be used to fit a general multivariate time series if no meaningful network structure is defined.

3. Numerical Studies

3.1. Simulated data

To demonstrate the finite-sample performance of the proposed model, we present three simulation examples. These three examples are similar, but vary in terms of the mechanism used to generate the network structure A. Once A is simulated, it is fixed throughout the remaining simulation studies. For a given network structure A, the response variable y_{it} is generated according to model (2.2). Then, for a given sample size T, we first simulate a time series sequence



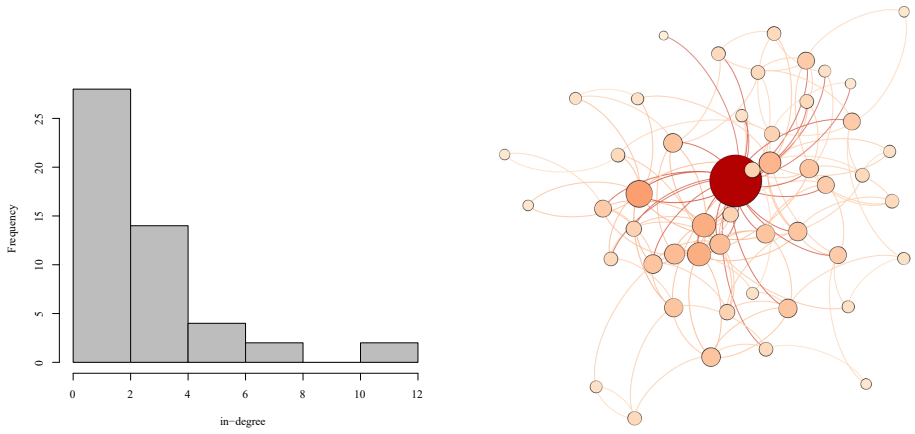
(a) (b)

Figure 1. Left panel: in-degree distribution for Example 1 with $N = 50$. Right panel: network visualization. A dot denotes a node and a line represents an edge. A deeper color and a larger dot indicate a larger in-degree.

of y_{it} of length $T_0 + T$. The first T_0 observations are dropped to eliminate the effect of the initial values. In the simulation study, we set $T_0 = 5,000$ for all three experiments. Finally, a sequence of $\{y_{it}\}$ is simulated.

Example 1. (Random Distributed Network Structure) We present a simple network structure with an in-degree (i.e., $q_i = \sum_{j=1}^{N} a_{ji}$) that follows a random distribution. This means there are no influential nodes (i.e., with a relatively large in-degree) in the network. We first generate N i.i.d. random variables from a uniform distribution between 0 and 5. Denote these variables by U_i, with $1 \leq i \leq N$. For each node i, we randomly select a sample size of $[U_i]$ from $\mathcal{S}_F = \{1, 2, \ldots, N\}$, without replacement, where $[U_i]$ denotes the smallest integer not less than U_i. Denote the sample by $\mathcal{S}_i$. Define $a_{ij} = 1$ if $j \in \mathcal{S}_i$, and $a_{ij} = 0$ otherwise. This leads to the adjacency matrix A. A histogram of the in-degree and a visualization of this network structure are shown Figure 1. From this figure, we can see that the distribution of the in-degree is almost random, and that there are no influential nodes.

Example 2. (Power-Law Distributed Network Structure) We next consider the power-law distributed network structure (Clauset, Shalizi and Newman (2009)). This network structure reflects a general phenomenon in which the majority of nodes have very few connections, while a small number have many connections.

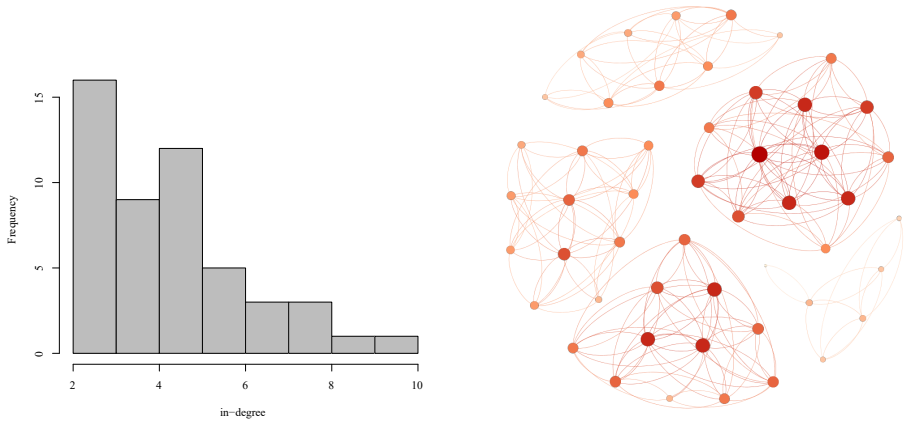


(a) (b)

Figure 2. Left panel: in-degree distribution for Example 2 with $N = 50$. Right panel: network visualization. A dot denotes a node and a line represents an edge. A deeper color and a larger dot indicate a larger in-degree.

In such networks, there always exist some influential nodes with a very large in-degree (e.g., celebrities). To mimic this network structure, we follow Clauset, Shalizi and Newman (2009), generating A as follows. First, we simulate each node's out-degree in the same way as in Example 1. Next, we generate a further N i.i.d. random variables (e.g., denoted as r_i, for $i = 1, \ldots, N$) according to the discrete power-law distribution; that is, $P(r_i) = ck^{-\alpha}$ for a normalizing constant c and exponent parameter $s = 2.5$. A smaller value of s implies a heavier distribution tail. We then normalize each r_i to its corresponding probability, $p_i = r_i / \sum_{i=1}^{N} r_i$. For each node i, we select a sample size of $[U_i]$ according to the probability of p_i from $\mathcal{S}_F = \{1, 2, \ldots, N\}$, without replacement. Denote the sample by $\mathcal{S}_i$. Define $a_{ij} = 1$ if $j \in \mathcal{S}_i$, and $a_{ij} = 0$ otherwise. A histogram of the in-degree and a visualization of this network structure are shown Figure 2. We can see that there is at least one node with a very large degree, which indicates that it could be an influential node.

Example 3. (Stochastic Block Network Structure) The stochastic block model (Nowicki and Snijders (2011)) is another popular network topology in the literature. For example, in a stock market, stocks may belong to different industries. The performance of a stock is very likely to be influenced by its neighbors in the same industry block. Following Nowicki and Snijders (2011), let $K = N/10$



(a) (b)

Figure 3. Left panel: in-degree distribution for Example 3 with $N = 50$. Right panel: network visualization with $K = 5$. A dot denotes a node and a line represents an edge.

be the total number of blocks in this case. For simplicity, we randomly assign a block label ($k = 1, 2, \ldots, K$) to each node, with equal probability. Next, set $P(a_{ij} = 1) = 0.5$ if i and j belong to the same block, and $P(a_{ij} = 1) = 0.001/N$ otherwise. This means nodes in the same block are more likely to be connected to each other than are nodes from different blocks. A histogram of the in-degree and a visualization of this network structure are shown Figure 3. From Figure 3, we can clearly see there are five blocks in this simulated network structure..

3.2. Simulation results

For each example, various combinations of network size (i.e., $N = 50, 100, 200$) and sample size (i.e., $T = 100, 200, 400$) are investigated. To make the simulation results more stable, we consider different setups of the true parameters $\theta_0 = (\omega_0, \alpha_0, \lambda_0, \beta_0)'$. In the first example, θ_0 is fixed as $(0.005, 0.1, 0.1, 0.7)'$. In the second example, θ_0 is fixed as $(0.01, 0.1, 0.2, 0.6)'$. Lastly, in the block case, θ_0 is fixed as $(0.02, 0.1, 0.3, 0.5)'$. Each simulation is randomly replicated $M = 1,000$ times. Let $\hat{\theta}^{(m)} = (\hat{\theta}_k^{(m)})' = (\hat{\omega}_0^{(m)}, \hat{\alpha}_0^{(m)}, \hat{\lambda}_0^{(m)}, \hat{\beta}_0^{(m)})'$ be the estimators obtained in the mth ($1 \le m \le M$) replication. We consider two measures to evaluate the finite-sample performance of the proposed method. First, for a given parameter θ_k, with $1 \le k \le 4$, the root-mean-square error is evaluated by $\text{RMSE}_k = \{M^{-1} \sum_{m=1}^{M} (\hat{\theta}_k^{(m)} - \theta_k)^2\}^{1/2}$. Second, for each $1 \le k \le 4$, a 95% confidence

Table 1. Simulation results for Example 1 with 1,000 replications. The RMSE values ($\times 10^{-2}$) are reported for each estimate. The corresponding CP (in %) is given in parentheses. Network density (ND) is also reported.

	T	ω_0	α_0	λ_0	β_0	ND(%)
$N = 50$	100	0.14 (94.9)	1.43 (94.9)	1.66 (95.1)	4.23 (93.7)	6.41
	200	0.08 (95.7)	0.99 (94.9)	1.12 (94.7)	2.79 (94.8)	
	400	0.06 (95.2)	0.70 (95.3)	0.79 (94.8)	1.87 (95.2)	
$N = 100$	100	0.11 (95.0)	1.00 (95.4)	1.18 (94.5)	3.02 (94.6)	3.20
	200	0.07 (95.0)	0.71 (94.6)	0.78 (96.5)	2.04 (94.1)	
	400	0.05 (95.2)	0.49 (95.5)	0.58 (94.2)	1.38 (94.7)	
$N = 200$	100	0.08 (94.0)	0.72 (94.2)	0.81 (94.6)	2.25 (92.0)	1.52
	200	0.05 (93.0)	0.52 (94.4)	0.55 (95.3)	1.47 (94.2)	
	400	0.03 (95.0)	0.36 (93.9)	0.40 (94.6)	1.00 (93.3)	

Table 2. Simulation results for Example 2 with 1,000 replications. The RMSE values ($\times 10^{-2}$) are reported for each estimate. The corresponding CP (in %) is given in parentheses. Network density (ND) is also reported.

	T	ω_0	α_0	λ_0	β_0	ND(%)
$N = 50$	100	0.47 (91.7)	1.56 (94.3)	2.00 (92.2)	3.93 (88.7)	5.51
	200	0.16 (93.5)	1.04 (93.6)	1.27 (93.6)	2.35 (91.9)	
	400	0.11 (93.5)	0.73 (94.3)	0.85 (94.0)	1.61 (92.0)	
$N = 100$	100	0.20 (94.5)	1.04 (95.4)	1.44 (94.3)	2.82 (93.9)	3.09
	200	0.13 (94.0)	0.74 (93.8)	0.95 (95.9)	1.97 (93.1)	
	400	0.09 (95.0)	0.50 (95.4)	0.70 (95.5)	1.33 (93.6)	
$N = 200$	100	0.14 (93.3)	0.73 (94.8)	1.01 (94.4)	2.13 (93.5)	1.50
	200	0.09 (93.6)	0.54 (94.6)	0.72 (94.0)	1.40 (93.5)	
	400	0.06 (95.3)	0.37 (94.7)	0.50 (94.7)	0.95 (94.6)	

interval is constructed for θ_k as $\mathrm{CI}_k^{(m)} = (\hat{\theta}_k^{(m)} - z_{0.975}\widehat{\mathrm{SE}}_k^{(m)}, \hat{\theta}_k^{(m)} + z_{0.975}\widehat{\mathrm{SE}}_k^{(m)})$, where $\widehat{\mathrm{SE}}_k^{(m)}$ is the square root of the jth diagonal element of $\widehat{\Sigma}_T$, and z_α is the αth quantile of a standard normal distribution. Then, the coverage probability is computed as $\mathrm{CP}_k = M^{-1}\sum_{m=1}^{M} I(\theta_k \in \mathrm{CI}_k^{(m)})$, where $I(\cdot)$ is the indicator function. Lastly, the network density (i.e., $\{N(N-1)\}^{-1}\sum_{i,j} a_{ij}$) is also reported.

The simulation results are summarized in Tables 1–3.

For the first example in Table 1, the estimators are consistent, with RMSE

Table 3. Simulation results for Example 3 with 1,000 replications. The RMSE values ($\times 10^{-2}$) are reported for each estimate. The corresponding CP (in %) is given in parentheses. Network density (ND) is also reported.

	T	ω_0	α_0	λ_0	β_0	ND(%)
$N = 50$	100	0.45 (95.4)	1.61 (95.1)	3.50 (95.1)	5.42 (94.4)	9.43
	200	0.29 (95.3)	1.13 (94.5)	2.37 (95.6)	3.64 (94.4)	
	400	0.21 (95.4)	0.79 (95.5)	1.75 (95.4)	2.61 (94.4)	
$N = 100$	100	0.30 (95.9)	1.15 (94.8)	2.53 (93.8)	3.79 (94.2)	4.65
	200	0.20 (95.3)	0.79 (95.2)	1.73 (95.1)	2.57 (94.7)	
	400	0.14 (95.4)	0.55 (95.2)	1.20 (94.6)	1.71 (95.8)	
$N = 200$	100	0.20 (94.3)	0.79 (95.6)	1.79 (93.8)	2.75 (93.3)	2.41
	200	0.14 (94.2)	0.58 (95.4)	1.22 (95.7)	1.86 (94.6)	
	400	0.09 (94.7)	0.40 (95.3)	0.87 (94.7)	1.27 (95.4)	

values that decrease toward zero as $T \to \infty$. For example, consider λ_0 (i.e., the estimated network effect) with $N = 100$. The RMSE value drops from 1.18% to 0.58% as T increases from 100 to 400. Furthermore, the reported coverage probabilities (i.e., CP) for each parameter (θ_k) are all fairly close to the nominal level of 95%. This suggests that the estimated standard error (i.e., $\widehat{\text{SE}}$) approximates the true SE well. Quantitatively similar results are obtained for Example 2 (see Table 2) and Example 3 (see Table 3). Together, these findings confirm that the proposed estimator $\hat{\theta}$ is consistent and asymptotically normal.

3.3. Real-data analysis

In this subsection, we use the proposed model to analyze real stock market data sets. The stock data are taken from the Chinese A share market traded on the Shanghai Stock Exchange and the Shenzhen Stock Exchange in 2014. Specifically, the response variable y_{it} is the log of the daily return. According to the industry classification criteria provided by the China Securities Regulatory Commission, each stock is placed into one of 18 categories. We use the following four categories to assess the performance of the proposed model: the *Mining Industry* with 68 stocks, *Real Estate* with 127 stocks, *Wholesale and Retail* with 139 stocks, and *Manufacturing* with 1,515 stocks. For each category, the network structure is constructed based on common shared ownership information. Specifically, we try three different choices of adjacency matrix to compare the estimation results. First, we collect information on the top 10 shareholders for each

Table 4. Estimate results for the stock data set.

		Case 1		Case 2		Case 3	
Category	Parameter	Estimate	p-value	Estimate	p-value	Estimate	p-value
Mining Industry	ω_0	0.0001	<0.001	0.0001	<0.001	0.0001	<0.001
	α_0	0.1933	<0.001	0.1992	<0.001	0.1978	<0.001
	λ_0	0.0571	0.0015	0.0712	0.0047	0.0451	0.0264
	β_0	0.5614	<0.001	0.5578	<0.001	0.5628	<0.001
Real Estate	ω_0	0.0001	<0.001	0.0001	<0.001	0.0001	<0.001
	α_0	0.1869	<0.001	0.1868	<0.001	0.1873	<0.001
	λ_0	0.0311	0.0070	0.0468	0.0265	0.0274	0.0527
	β_0	0.6605	<0.001	0.6756	<0.001	0.6772	<0.001
Wholesale and Retail	ω_0	0.0001	<0.001	0.0001	<0.001	0.0001	<0.001
	α_0	0.1855	<0.001	0.1907	<0.001	0.1897	<0.001
	λ_0	0.0447	0.0001	0.0199	0.3281	0.0264	0.0399
	β_0	0.6867	<0.001	0.6937	<0.001	0.6889	<0.001
Manufacturing	ω_0	0.5×10^{-4}	0.0091	0.0001	<0.001	0.0001	<0.001
	α_0	0.1417	<0.001	0.1595	<0.001	0.1499	<0.001
	λ_0	0.0916	0.0027	0.0196	0.2982	0.0587	0.0109
	β_0	0.7101	<0.001	0.6982	<0.001	0.7104	<0.001

stock. The network structure (i.e., adjacency matrix) is constructed as follows. For any two arbitrary stocks i and j, $a_{ij} = 1$ if they share at least one common shareholder, otherwise $a_{ij} = 0$. The second matrix is similar to the first except that for stocks i and j, $a_{ij} = 1$ if they share at least two common shareholders, otherwise $a_{ij} = 0$. Lastly, we collect information on the top 5 shareholders for each stock. For stocks i and j, $a_{ij} = 1$ if they share at least one common shareholder, otherwise $a_{ij} = 0$.

To provide a descriptive analysis for these four categories, we plot the daily averaged stock return for each industry in Figure 4. We also display their network structure in Figure 5. Clearly, the network structures of the four industries differ from each other. Then, we apply the network GARCH model to each data set. The estimation results are given in Table 4.

Table 4 shows that the results are consistent across different choices of adjacency matrix. Nearly all estimates are statistically significant at the 1% or 5% levels, except for the second case for manufacturing and retail. This is understandable, because the network in this case is too sparse. This leads to a nonsignificant effect of the network structure. For example, the estimated net-

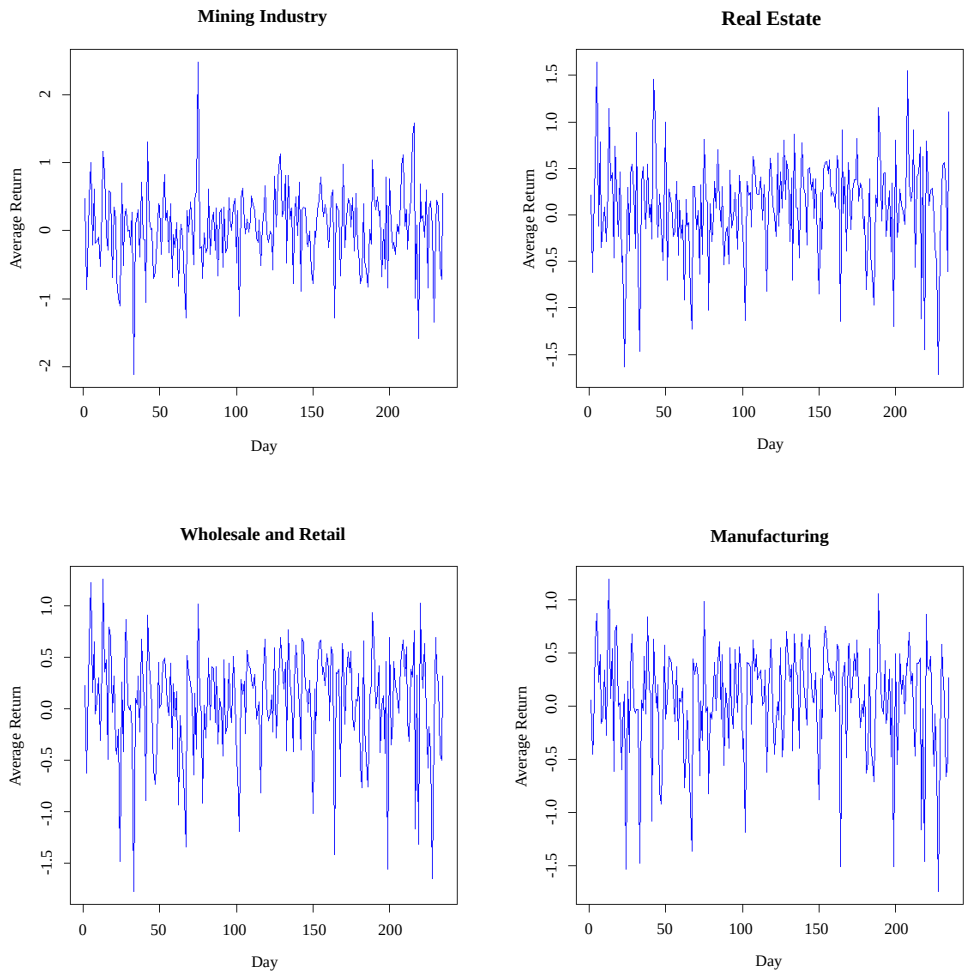
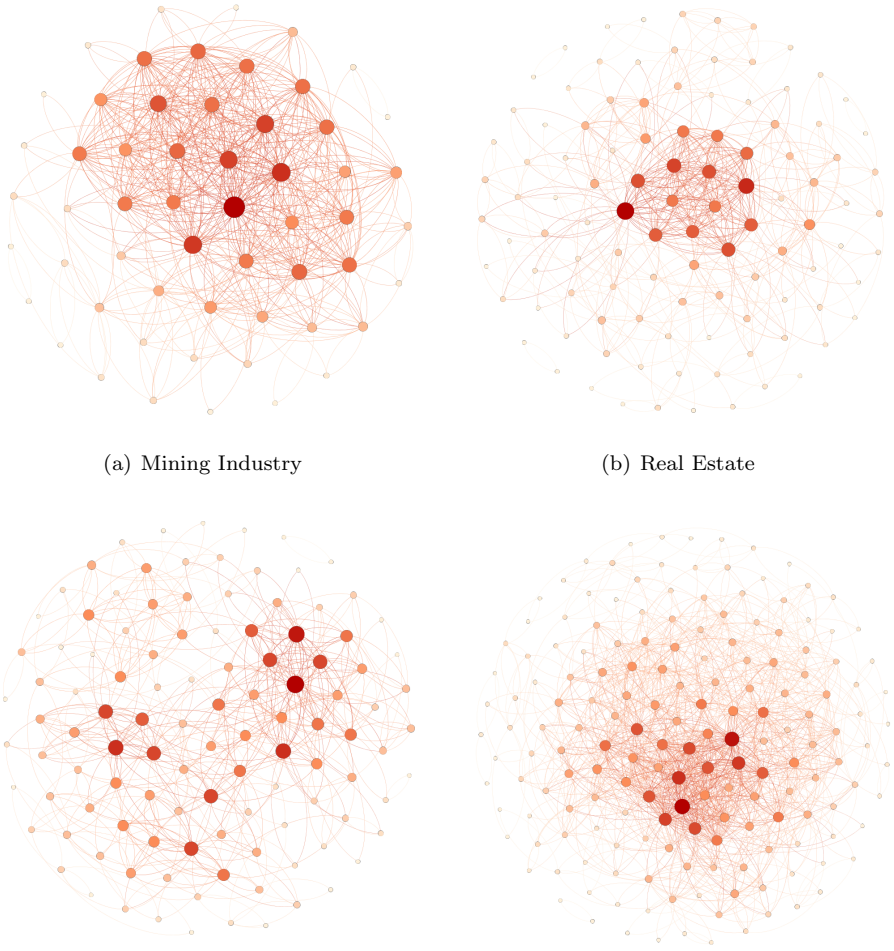


Figure 4. Average stock return for the four categories in 2014.

work effect λ_0 (0.0571) for the mining industry suggests that the return of a stock is positively related to the performance of its connected neighbors. The estimated α_0 (0.1933) confirms that a stock with a higher (lower) return in the past is likely to exhibit a higher (lower) performance in the future. Finally, the estimated β_0 (0.5614) is very strong compared with the other two effects, indicating that the variance of a stock could be very large in practice. Similar results are found in the other three categories.



(a) Mining Industry

(b) Real Estate

(c) Wholesale and Retail

(d) Manufacturing

Figure 5. Network structure for the four categories.

4. Conclusion

We propose a network GARCH model that takes network structure information into consideration. To capture the impact of connected neighbors, we introduce a network structure term to the traditional GARCH (1,1) model. The proposed model decreases the computational complexity substantially from $O(N^2)$ to $O(N)$. The resulting estimators enjoy asymptotic properties, and these findings are confirmed by extensive numerical studies. We further illustrate our model using a real data set from the Chinese stock market. A significant network

structure term is detected.

To conclude this article, we present a number of interesting topics for future study. First, the network structure term in model (2.2) is added on $y_{j,t-1}^2$. However, the adjacency matrix can also be added on $h_{j,t-1}$. This makes the resulting model considerably more complicated, and the associated theoretical development more challenging. Nevertheless, linking $h_{j,t-1}$ would be a worthwhile research topic. Second, the parameter θ is assumed to be the same across stocks. However, in reality, stock heterogeneity may mean that stocks have different reactions to an effect. Thus, investigating a varying-coefficient network GARCH model is another interesting problem worth pursuing. Lastly, the network structure discussed in this paper is simple and straightforward. The network term is driven by one parameter, λ. This may limit our model in some contexts. However, our empirical results show that, for instance, the model performs well when the stocks are from the same category. In future research, we will consider network structures that are more flexible, such as a sub-block structure.

Supplementary Material

All technical details can be found in the online Supplementary Material.

Acknowledgements

We express our sincere thanks to the Co-editor, Associate Editor, and anonymous referee for their constructive suggestions. Zhou's research was supported in part by the NSFC (No.71702185), Beijing Municipal Social Science Foundation (No.19GLC052) and the fund for building world-class universities (disciplines) of Renmin University of China, Center for Applied Statistics of Renmin University of China. Li's research was supported in part by the NSFC (No.11571348, No.11771239) and the Tsinghua University Initiative Scientific Research Program (No.2019Z07L01009). Pan's research was supported in part by the NSFC (No.11601539, No.11631003, No.71771224), the Fundamental Research Funds for the Central Universities (QL18010), the Program for Innovation Research in Central University of Finance and Economics, and Youth Talent Development Support Program (QYP1911). Wang's research was partially supported by the NSFC (No. 11831008, No.11525101, No.71532001) and China's National Key Research Special Program (No.2016YFC0207704).

References

Bajo, E., Croci, E. and Marinelli, N. (2017). Institutional investors networks and firm value. *Available at SSRN:* `https://ssrn.com/abstract=2633541`.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *J. Econometrics* **31**, 307-327.

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model. *Rev. Econ. Stat.* **72**, 498-505.

Bollerslev, T. and Engle R. F.(1993). Common persistence in conditional variances. *Econometrica* **61**, 167-186.

Bollerslev, T. and Wooldridge, J. M. (1988). A capital asset pricing model with time varying covariances. *J. Political Econ.* **96**, 116-31.

Bougerol, P. and Picard, N. (1992). Stationarity of GARCH processes and of some nonnegative time series. *J. Econometrics* **52**, 115-127.

Clauset, A, Shalizi, C. R. and Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Rev.* **51**, 661-703.

Chou, P. I. and Lee, C. H. (2012). Is concentration a good idea? evidence from active fund management. *Asia-Pacific Financial Markets* **19**, 23-41.

Conrad, C. and Mammen, E. (2016). Asymptotics for parametric GARCH-in-Mean models. *J. Econometrics* **194**, 319-329.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica* **50**, 987-1007.

Engle, R. F. and Kroner, K. F. (1995). Multivariate simultaneous GARCH. *Econometric Theory* **11**, 122-150.

Engle, R. F., Ng, V. K. and Rotschild, M. (1990). Asset pricing with a factor ARCH covariance structure: empirical estimates for treasury bills. *J. Econometrics* **45**, 213-238.

Fan, J. and Yao, Q. (2017). *The Elements of Financial Econometrics*. Cambridge University Press, Cambridge.

Goel, S. and Goldstein, D. G. (2014). Predicting individual behavior with social networks. *Marketing Science* **33**, 82-93.

Hall, P. and Yao, Q. (2003). Inference in ARCH and GARCH models with heavy-tailed errors. *Econometrica* **71**, 285-317.

Hu, Y. P. and Tsay, R. S. (2013). Principal volatility component analysis. *J. Bus. Econom. Statist.* **32**, 153-164.

Lam, C. and Yao, Q. (2012). Factor modelling for high-dimensional time series: inference for the number of factors. *Ann. Statist.* **40**, 694-726.

Lee, S. W. and Hansen, B. E. (1994). Asymptotic theory for the GARCH(1,1) quasi maximum likelihood estimator. *Econometric Theory* **10**, 29-52.

Li, W., Gao, J., Li, K. and Yao, Q. (2016). Modelling multivariate volatilities via latent common factors. *J. Bus. Econom. Statist.* **34**, 564-573.

Livingston, L. (2010). Evaluating alternative weighting schemes for stocks in a "best ideas" portfolio. *International Journal of Business and Finance Research* **4**, 117-136.

Lumsdaine, R. L. (1996). Consistency and asymptotic normality of the quasi-maximum likelihood estimator in IGARCH(1,1) and covariance stationary GARCH(1,1) models. *Econometrica* **64**, 575-96.

Nelson, D. B. (1990). Stationarity and persistence in the GARCH(1,1) model. *Econometric Theory* **6**, 318–334.

Nitzan, I. and Libai, B. (2011). Social effects on customer retention. *J. Mark.* **75**, 24-38.

Nowicki, K. and Snijders, T. A. B. (2011). Estimation and prediction for stochastic block structures. *J. Amer. Statist. Assoc.* **96**, 1077-1087.

Pan, J., Polonik, W. and Yao, Q. (2010). Estimating factor models for multivariate volatilities: An innovation expansion method. In *Proceedings of COMPSTAT'2010*, 305–314. Physica-Verlag, Heidelberg.

Pareek, A. (2012). Information networks: implications for mutual fund trading behavior and stock returns. *AFA 2010 Atlanta Meetings Paper*. Available at SSRN: `https://ssrn.com/abstract=1361779`.

Polonik, W. and Yao, Q. (2008). Testing for multivariate volatility functions using minimum volume sets and inverse regression. *J. Econometrics* **147**, 151-162.

Tao, M., Wang, Y., Yao, Q. and Zou, J.(2012). Large volatility matrix inference via combining low-frequency and high-frequency approaches. *J. Amer. Statist. Assoc.* **106**, 1025-1040.

Tsay, R. S. (2003). *Analysis of Financial Time Series*. Wiley.

Tse, Y. K. (2000). A test for constant correlations in a multivariate garch model. *J. Econometrics* **98**, 107-127.

Tse, Y. K. and Tsui, A. K. C. (2002). A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *J. Bus. Econom. Statist.* **20**, 351-62.

Wei, Y., Yildirim, P., Christophe, V. D. B. and Dellarocas, C. (2014). Credit scoring with social network data. *Social Science Electronic Publishing* **35**, 234-258.

Zhu, K. and Ling, S. (2011). Global self-weighted and local quasi-maximum exponential likelihood estimators for ARMA-GARCH/IGARCH models. *Ann. Statist.* **39**, 849-873.

Zhu, X. N., Pan, R., Li, G. D., Liu, Y. W. and Wang, H. S. (2017). Network vector autoregression. *Ann. Statist.* **45**, 1096-1123.

School of Statistics, Renmin University of China, Beijing 100872, China.

E-mail: zhoujing_89@126.com

Center for Statistical Science and Department of Industrial Engineering, Tsinghua University, Beijing 100084, China.

E-mail: malidong@tsinghua.edu.cn

School of Statistics and Mathematics, Central University of Finance and Economics, Beijing 100081, China.

E-mail: panrui_cufe@126.com

Guanghua School of Management, Peking University, Beijing 100871, China.

E-mail: hansheng@pku.edu.cn

(Received January 2018; accepted October 2018)