Dynamic network quantile regression model

Article

Accepted Version

Dynamic Network Quantile Regression Model

Xiu Xu * Weining Wang † Yongcheol Shin ‡ Chaowen Zheng§

Abstract

We propose a dynamic network quantile regression model to investigate the quantile connectedness using a predetermined network information. We extend the existing network quantile autoregression model of Zhu et al. (2019b) by explicitly allowing the contemporaneous network effects and controlling for the common factors across quantiles. To cope with the endogeneity issue due to simultaneous network spillovers, we adopt the instrumental variable quantile regression (IVQR) estimation and derive the consistency and asymptotic normality of the IVQR estimator using the near epoch dependence property of the network process. Via Monte Carlo simulations, we confirm the satisfactory performance of the IVQR estimator across different quantiles under the different network structures. Finally, we demonstrate the usefulness of our proposed approach with an application to the dataset on the stocks traded in NYSE and NASDAQ in 2016.

JEL classification: C32, C51, G17

Keywords: Dynamic Network Quantile Regression Model, Simultaneous Network Endogeneity, IVQR Estimator, Quantile Connectedness.

*Dongwu Business School, Soochow University, 50 Donghuan Road, Suzhou, Jiangsu 215021, PR China. Email: xiux@suda.edu.cn.

†Department of Economics and Related Studies, University of York, Heslington, York, YO10 5DD, UK. Email: weining.wang@york.ac.uk.

‡Department of Economics and Related Studies, University of York, Heslington, York, YO10 5DD, UK. Email: yongcheol.shin@york.ac.uk.

§Department of Economics, University of Reading, Reading, RG6 6EL, UK. Email: chaowenzheng27@gmail.com.

1 Introduction

The topology of financial networks is central to the study of financial contagion and systemic risk, see Fafchamps and Gubert (2007), Acemoglu et al. (2015), Hautsch et al. (2015) among others. Given the relevance of tail dependence for financial supervision and risk management (Betz et al., 2016), a joint analysis of network effect and tail dependence becomes more important because the implications derived from network models evaluated by conventional conditional mean estimators cannot necessarily be generalized to the tails. Ando et al. (2021) show that major adverse events are associated with an increase in average connectedness but that their effects on the tails significantly differ.

Quantile regression (QR) has been a powerful tool for characterizing the heterogeneous policy impacts and measuring tail-event driven risk (e.g. Härdle et al. (2016)). Following a seminal work by Bassett and Koenker (1978), QR can be used to evaluate the entire range of the conditional distribution. Recently, the literature on quantile time series regression has been rapidly growing. Koenker and Xiao (2006) propose a quantile autoregressive model while Galvao et al. (2013) study QR in an autoregressive dynamic framework with exogenous stationary covariates. Following the analysis of quantile cointegration in Xiao (2009), Cho et al. (2015) bring QR to the autoregressive distributed lag (ARDL) model literature. The quantile ARDL process captures both the long-run and short-run relationships at any desired location in the conditional distribution. Engle and Manganelli (2004) advance a conditional autoregressive value at risk model whilst White et al. (2015) propose a vector autoregressive (VAR) model for analyzing quantile dynamics.

However, in most financial systems, multiple entities are often intertwined and interacted with each other, which can be represented as networks (Hautsch et al., 2015; Härdle et al., 2016; Chen et al., 2019). In this context, Zhu et al. (2017) develop a network autoregression (NAR) model, which has gained great popularity in social network analysis. A number of extensions have been developed. Zhu et al. (2020) consider a multivariate spatial autoregression model. Zhu et al. (2019a) investigate a screening method to select influential nodes. Zhu (2020) studies nonconvex penalized estimation methods in high-dimensional VAR models while Zhu and Pan (2020) extend the network VAR model to allow for group-specific parameters.

In particular, Zhu et al. (2019b) extend the NAR model by Zhu et al. (2017) and propose a network quantile autoregression (NQAR) model in order to analyze tail dependency in a dynamic network with a large number of nodes. The NQAR model consists of a system of equations, in which a continuous response is related to its lagged connected nodes, the response of the same node in the previous time period as well as node specific characteristics in a quantile autoregression process. However, main weakness of the NQAR model lies in that it does not accommodate the contemporaneous impact of connected nodes, even though the simultaneous network/peer effects are pervasive in empirical studies (Liu, 2014; Forni and Gambetti, 2010). If they are statistically significant, the estimation of the NQAR model would become inconsistent and misleading.

In this paper, as a main contribution, we extend the NQAR model and propose a general dynamic network quantile regression model (DNQR) by explicitly incorporating contemporaneous and lagged network effects of connected nodes as well as the impacts of node-specific variables and observed common effects. Notice, however, that the simultaneous network effects are inherently endogenous to the system, which leads to inconsistent estimates at conditional mean regression as well as in QR, see Wüthrich (2019, 2020); Chernozhukov et al. (2020). To cope with the endogeneity issue in the different contexts, many studies have attempted to apply the instrumental variable quantile regression (IVQR) estimation advanced by Chernozhukov and Hansen (2006), e.g. Frölich and Melly (2013), Su and Hoshino (2016) and Machado and Silva (2019).

To deal with this challenging issue we adopt the IVQR approach. The social network data are similar to the spatial data, in the sense that observations from connected users are correlated. This makes the spatial autoregressive model good candidate for network data analysis. However, our work can be regarded as a nontrivial extension of Su and Yang (2011), who apply the IVQR approach to analyzing the cross-section data using a linear spatial autoregressive model, to a dynamic network quantile model with nodal heterogeneity and common factors, which can shed further lights on uncovering the complex tail dependency in dynamic networks with a large number of nodes and time periods. Our study is also closely related to the growing literature on the tail comovements in a complex financial system, see Diebold and Yilmaz (2014), Hautsch et al. (2014), White et al. (2015) and Ando et al. (2021).

More importantly, we follow Jenish and Prucha (2012) and Xu and Lee (2015), and develop the general asymptotic theory for the IVQR estimator by applying the spatial near epoch dependence (NED) of the underlying network processes. We derive the detailed conditions on the network processes in order to establish the consistency and asymptotic normality of the IVQR estimator. Via Monte Carlo simulations, we confirm that the finite sample performance of the IVQR estimator is satisfactory across quantiles and the different error distributions under the different network structures.

Finally, we demonstrate the utility of our approach with an application to the dataset on the stocks traded in NYSE and NASDAQ in 2016, through the common shareholder network constructed using the information on the common mutual fund ownership, e.g. Anton and Polk (2014). We find that the contemporary network effects (measured by returns of connected stocks in the same period) are positive and significant and dominate all the other effects across all quantiles. Furthermore, they are stronger at the lower tails than at the upper tails, suggesting that the contemporaneous network effects should be explicitly and carefully analysed in the dynamic network quantile model.

This paper proceeds as follows. In Section 2 we outline the DNQR model and derive the stationarity condition of the underlying network process. Section 3 introduces the IVQR estimation and develops its asymptotic properties using the spatial NED approach. In Section 4 we provide simulation results, showing that the IVQR estimation performs satisfactory. In Section 5 the DNQR model is applied to the US financial market data. Section 6 provides concluding remarks. The mathematical proofs and the additional simulation and empirical results are presented in the Online Appendix. The replication code can be found here on GitHub.

Notations: For a vector $v = (v_1, \ldots, v_m)^\top \in \mathbb{R}^m$, we denote $|v|_k = (\sum_{i=1}^m |v_i|^k)^{1/k}$, $\|v\|_k = (\sum_{i=1}^m \mathsf{E}\,|v_i|^k)^{1/k}$, and $|v|_\infty = \max_{i\leq m} |v_i|$, where k is a positive integer, and E is the expectation operator. For any $n \times m$ matrix $A = (a_{ij})_{1\leq i\leq n, 1\leq j\leq m}$, we define the two norm and the max norm by $|A|_2 = \sup_{\{v\in\mathbb{R}^m, |v|_2=1\}} |Av|_2$ and $|A|_{\max} = \max_{i,j} |a_{ij}|$, respectively. Define the column-sum and the row-sum by $\|A\|_1 = \max\limits_{1\leqslant j\leqslant m} \sum\limits_{i=1}^{n} |a_{ij}|$ and $\|A\|_\infty = \max\limits_{1\leqslant i\leqslant n} \sum\limits_{j=1}^{m} |a_{ij}|$. We write $a_n = O(b_n)$ or $a_n \lesssim b_n$ if there exists a positive constant C such that $a_n/b_n \leqslant C$ for all large n, and denote $a_n = o(b_n)$ (resp. $a_n \sim b_n$), if $a_n/b_n \to 0$

(resp. $a_n/b_n \to c$ for a positive constant c). For two sequences of random variables (X_n) and (Y_n), we write $X_n = o_p(Y_n)$ if $X_n/Y_n \to 0$ in probability. Let I_N be an $N \times N$ identity matrix, $\mathbf{I}(\cdot)$ the indicator function, $\mathbf{1}_N$ a $N \times 1$ vector with each element as one, and $\mathbb{N}$ the integer set.

2 The Model

Consider the large scale network time series data with N nodes for $1 \leqslant i \leqslant N$, and T time periods for $1 \leqslant t \leqslant T$, which is observationally equivalent to a regular panel data. To describe their relationship, we construct an adjacency matrix, $A = (a_{ij}) \in \mathbb{R}^{N\times N}$, where $a_{ij} = 1$ if the node i follows the node j, and $a_{ij} = 0$ otherwise. We do not allow the self-following relation, $a_{ii} = 0$. Define the row-normalised network matrix as $W = (w_{ij}) \in \mathbb{R}^{N\times N}$, where $w_{ij} = n_i^{-1} a_{ij}$ and $n_i = \sum_{j=1}^{N} a_{ij}$. Let $\mathbb{Y}_t = (Y_{1t}, \cdots, Y_{Nt})^\top \in \mathbb{R}^N$ be the continuous response (e.g., tweet length) collected at time t, and U_{it} be a sequence of *i.i.d.* uniform random variables on the set $[0, 1]$.

We then consider the following DNQR model:

$$\begin{aligned} Y_{it} &= \gamma_0^0(U_{it}) + \sum_{l=1}^{q} \alpha_l^0(U_{it}) Z_{il} + \gamma_1^0(U_{it}) \sum_{j=1}^{N} w_{ij} Y_{jt} \\ &\quad + \gamma_2^0(U_{it}) \sum_{j=1}^{N} w_{ij} Y_{j,t-1} + \gamma_3^0(U_{it}) Y_{i,t-1} + \sum_{k=0}^{p} F_{t-k}^\top \beta_k^0(U_{it}) \\ &\stackrel{\text{def}}{=} h(U_{it}, Z_{i1}, \ldots, Z_{iq}, \sum_{j=1}^{N} w_{ij} Y_{jt}, \sum_{j=1}^{N} w_{ij} Y_{j,t-1}, Y_{i,t-1}, F_t, \ldots, F_{t-p}), \end{aligned} \tag{1}$$

for $i = 1, \cdots, N$ and $t = 1, \cdots, T$, where $\gamma_j^0(\cdot)$ for $j = 0, 1, 2, 3$, α_l^0 for $l = 1, \cdots, q$, and each elements in $\beta_k^0 \in \mathbb{R}^m$ for $k = 0, 1, \cdots, p$ are unknown parameter functions from $[0, 1]$ to $\mathbb{R}$, and the superscript 0 is used to denote the true value of parameters. $Z_i = (Z_{i1}, \cdots, Z_{iq})^\top \in \mathbb{R}^q$ is a $q \times 1$ vector of time-invariant node-specific covariates, and $F_t = (F_{t1}, \cdots, F_{tm})^\top \in \mathbb{R}^m$ is an $m \times 1$ vector of time-varying common covariates that capture the systematic influences on response variable Y_{it}.

If the right hand side of the DNQR model (1), i.e. $h(u, \cdots)$, is monotonically increasing in u with other values fixed at any point, then we can write the τ-th conditional

quantile function of Y_{it} as

$$Q_{Y_{it}}(\tau|\mathcal{F}_t) = \gamma_0^0(\tau) + \sum_{l=1}^{q} \alpha_l^0(\tau) Z_{il} + \gamma_1^0(\tau) \sum_{j=1}^{N} w_{ij} Y_{jt} + \gamma_2^0(\tau) \sum_{j=1}^{N} w_{ij} Y_{j,t-1} + \gamma_3^0(\tau) Y_{i,t-1} + \sum_{k=0}^{p} F_{t-k}^\top \beta_k^0(\tau), \tag{2}$$

where $\mathcal{F}_t = \{Z_1, \cdots, Z_N, \mathbb{Y}_{t-1}, \mathbb{Y}_t, F_t, F_{t-1}, \cdots, F_{t-p}\}$ is the information set. The first component, $\gamma_0^0(\tau) + \sum_{l=1}^{q} \alpha_l^0(\tau) Z_{il}$ is the quantile-specific nodal impact of the node i, where $\gamma_0^0(\tau)$ is the baseline function and Z_{il}s are assumed to be independent from $U_{it}s$. Next, network interactions between nodes are captured via both contemporaneous and lagged network variables, $\sum_{j=1}^{N} w_{ij} Y_{jt}$ and $\sum_{j=1}^{N} w_{ij} Y_{j,t-1}$, with $\gamma_1^0(\tau)$ capturing the quantile-specific simultaneous network effects and $\gamma_2^0(\tau)$ measuring the lagged diffusion network effects. $\gamma_3^0(\tau)$ is the quantile-specific momentum function, capturing the temporal dynamics for the same node. Furthermore, we control for the dynamic impacts of the (observed) common macroeconomic and financial factors, F_t, which can mitigate any remaining common shock effect in the data.

Let $\mathbb{F}_t = (F_t^\top, \cdots, F_{t-p}^\top)^\top \in \mathbb{R}^{(p+1)m}$. Define $\mathbf{A}_{0t} = (\gamma_0^0(U_{it}) + \sum_{l=1}^{q} \alpha_l^0(U_{it}) Z_{il}, 1 \leqslant i \leqslant N)^\top \in \mathbb{R}^N$, $\mathbf{A}_{1t} = diag\{\gamma_1^0(U_{it}), 1 \leqslant i \leqslant N\} \in \mathbb{R}^{N \times N}$, $\mathbf{A}_{2t} = diag\{\gamma_2^0(U_{it}), 1 \leqslant i \leqslant N\} \in \mathbb{R}^{N \times N}$, $\mathbf{A}_{3t} = diag\{\gamma_3^0(U_{it}), 1 \leqslant i \leqslant N\} \in \mathbb{R}^{N \times N}$, and $\mathbf{B}_t = ((\beta_0^{0\top}(U_{it}), \cdots, \beta_p^{0\top}(U_{it}))^\top, 1 \leqslant i \leqslant N)^\top \in \mathbb{R}^{N \times (p+1)m}$. The DNQR model (1) can be expressed compactly in a matrix form:

$$\mathbb{Y}_t = \Gamma + \mathbf{A}_{1t} W \mathbb{Y}_t + \mathbf{H}_t \mathbb{Y}_{t-1} + \mathbf{B}_t \mathbb{F}_t + V_t, \tag{3}$$

where $\mathbf{H}_t = \mathbf{A}_{2t} W + \mathbf{A}_{3t} \in \mathbb{R}^{N \times N}$, $\Gamma = \mathsf{E}(\mathbf{A}_{0t})$, and $V_t = \mathbf{A}_{0t} - \Gamma \in \mathbb{R}^N$ is *i.i.d.* over t with mean $\mathbf{0}$ and variance-covariance matrix, $\Sigma_V = \sigma_V^2 I_N \in \mathbb{R}^{N \times N}$.

Notice that the DNQR model can be regarded as a substantial extension of Koenker and Xiao (2006), who provide a classic framework for the analysis of the random-coefficient model in the quantile autoregression. Moreover, the DNQR model encompasses the NAR model by Zhu et al. (2017) and the NQAR model by Zhu et al. (2019b), through jointly incorporating contemporaneous and lagged network effects of connected nodes as well as exogenous common effects.

We show that the DNQR model is subject to the endogeneity issue due to contempo-

raneous network spillovers across nodes. Consider a simple two-equation system:

$$Y_{1t} = \gamma_0^0(U_{1t}) + \gamma_1^0(U_{1t})a_{12}Y_{2t}, \tag{4}$$
$$Y_{2t} = \gamma_0^0(U_{2t}) + \gamma_1^0(U_{2t})a_{21}Y_{1t}. \tag{5}$$

Assuming that $1 - a_{21}a_{12}\gamma_1^0(U_{1t})\gamma_1^0(U_{2t}) \neq 0$, we obtain the following solutions:

$$Y_{1t} = (\gamma_0^0(U_{1t}) + \gamma_0^0(U_{2t})\gamma_1^0(U_{1t})a_{12})/(1 - a_{21}a_{12}\gamma_1^0(U_{1t})\gamma_1^0(U_{2t})), \tag{6}$$
$$Y_{2t} = (\gamma_0^0(U_{2t}) + \gamma_0^0(U_{1t})\gamma_1^0(U_{2t})a_{21})/(1 - a_{21}a_{12}\gamma_1^0(U_{1t})\gamma_1^0(U_{2t})). \tag{7}$$

As Y_{1t} is a function of U_{1t} and U_{2t}, the monotone argument cannot be applied because

$$\mathrm{P}\left(Y_{it} \leqslant \gamma_0^0(\tau) + \gamma_1^0(\tau)\overline{Y}_{it} | \overline{Y}_{it}\right) \neq \tau \quad a.s. \text{ for } i = 1, 2, \tag{8}$$

where $\overline{Y}_{1t} = a_{12}Y_{2t}$ and $\overline{Y}_{2t} = a_{21}Y_{1t}$. This shows that the endogeneity is caused by the contemporaneous network term, $\overline{Y}_{it}$.

The simultaneous network spillover would cause inconsistency. Consider the simple mean regression, $Y_{it} = \lambda \sum_{j \neq i} w_{ij}Y_{jt} + \varepsilon_{it}$. Let $w_i = (w_{i1}, ..., w_{ij}, ..., w_{iN})^\top \in \mathbb{R}^N$, and $\tilde{w}_{ij}$ as the (i, j) element of the matrix $(I_N - \lambda W)^{-1}$. Assuming that $\mathsf{E}(\varepsilon_{it}\varepsilon_{jt}) = 0$ if $i \neq j$ and $\mathsf{E}(\varepsilon_{it}^2) = \sigma_i^2$, then the bias term (the average correlation between the endogenous variable and the error term) will be of the order, $\lim_{N,T\to\infty}(NT)^{-1}\sum_i \sum_t \mathsf{E}(w_i^\top \mathbb{Y}_t \varepsilon_{it}) = \lim_{N\to\infty} N^{-1}\sum_i \sum_{j\neq i} w_{ij}\tilde{w}_{ji}\sigma_i^2 \lesssim c$, where c is a constant. This is not negligible to zero unless $\lim_{N\to\infty} N^{-1}\sum_i \sum_{j\neq i} w_{ij}\tilde{w}_{ji}\sigma_i^2 = o(1)$. Thus, the estimation is likely to be biased unless the link of the network is very weak. In the quantile case, the leading bias term will be of the order:

$$\lim_{N,T\to\infty}(NT)^{-1}\sum_i \sum_t \mathsf{E}((\tau - \mathbf{I}(\varepsilon_{it,\tau} \leqslant 0))w_i^\top \mathbb{Y}_t), \tag{9}$$

which does not tend to zero, where $\varepsilon_{it,\tau} = Y_{it} - Q_{Y_{it}}(\tau|\mathbb{Y}_t)$ is the τ-th QR error.

The nontrivial estimation issue for the DNQR model lies in that the endogeneity caused by contemporaneous network spillovers renders the ordinary QR estimator to be inconsistent. Chernozhukov and Hansen (2005, 2006, 2008) propose the IVQR approach

to estimating quantile treatment effects and develop the robust inference. Chernozhukov et al. (2020) develop a novel technique to constructing simultaneous confidence bands for quantile functions and quantile effects in nonlinear network panels.

We follow the IVQR approach to cope with the simultaneous network endogeneity. Notice that Su and Yang (2011) apply the IVQR approach to an analysis of the cross-section data using a linear spatial autoregressive model. However, our work can be regarded as a nontrivial extension of Su and Yang (2011) to a dynamic network quantile model with nodal heterogeneity and common factors, which can shed further lights on uncovering the complex tail dependency in dynamic networks with a large number of nodes and time periods. More importantly, we derive a general asymptotic theory by using the spatial NED property of the network process in Section 3.2.

It is worthy to note that one can possibly explore GMM method to estimate the DNQR model. As noted in Chernozhukov and Hansen (2006), the IVQR estimator is asymptotically equivalent to a GMM estimator and some researcher have also developed exact computation of the GMM estimates of the IVQR parameters, see Chen and Lee (2018) and Firpo et al. (2021). Under the GMM framework, we would also have more issues that deserve investigation, e.g. the efficiency of the estimation. We leave these for further research directions.

2.1 Stationarity

In this section we derive the stationarity conditions for $\mathbb{Y}_t$ in (3), and its asymptotic distribution. Notice that the DNQR model can easily produce predictions of quantiles, $\widehat{Q}_{Y_{it}}(\tau|\mathcal{F}_t)$, given the network structure and the data history, by plugging the estimated parameters into (2). To this end, it is important to derive the conditions under which the network process is stationary. Further, stationarity may be required to identify some parameters. For example, if we wish to uncover the variance structure of series of interest, it would be crucial to check whether it changes over time or not.

Define $S_t = I_N - \mathbf{A}_{1t}W$. Then, we make the following assumptions:

Assumption 2.1. *(1) Let* $\Upsilon = \max_i |\gamma_1^0(U_{it})| \leqslant c_1 < 1$ *and* $|W|_2 \leqslant 1$*, where* W *is a row*

normalized network matrix with $\sum_{j=1}^{N} w_{ij} = 1$, *and* c_1 *is a positive constant. Assume that* U_{it} *and* Z_i *are i.i.d. over* i *and* t, *and* F_t *are i.i.d. The* k*th moments of* F_t *and* Z_i *are finite,* $k > 2$.

(2) $\max_i |\gamma_2^0(U_{it})| + \max_i |\gamma_3^0(U_{it})| \leqslant c_{23} < 1$ *and* $c_1 + c_{23} < 1$, *where* c_{23} *is a positive constant.*

(3) $\max_i |\gamma_0^0(U_{it})| + \max_i \sum_{l=1}^{q} |\alpha_l^0(U_{it})||Z_{il}| \leqslant d_z$, *and* $|\mathbf{B}_t|_\infty |\mathbb{F}_t|_1 \leqslant d_f$, *where* d_z *and* d_f *are random variables with bounded moments. Let* $\mathbb{D}_t = S_t^{-1}(\mathbf{B}_t \mathbb{F}_t + \mathbf{A}_{0t})$ *with* $\mathbb{D} = \mathsf{E}\, \mathbb{D}_t$ *and the elementwise maximum value* $\mathbb{D}_{\max} < \infty$. *Then,* $\max_t |\, \mathsf{E}\{vec(\mathbb{D}_{t-l_1} \mathbb{D}_{t-l_1}^\top)\}|_\infty \leqslant \sigma_{d\max} < \infty$, *where* $l_1 = (0, 1, \cdots, t-1)$.

(4) The right hand side of the model (1), *i.e.* $h(u, \cdots)$, *is monotonically increasing in* $u \in [0, 1]$.

Assumption 2.1(1) assures the invertibility of S_t. The model (3) has a unique solution if and only if every principal minor of $I_N - \mathbf{A}_{1t} W$ is positive, which is met by Assumption 2.1(1), though it is only a sufficient condition. Assumption 2.1(2) is necessary to obtain the strict stationarity of $\{\mathbb{Y}_t\}_t$. Under Assumptions 2.1(2) and (3), the covariance stationarity can be achieved. Then, we have the following lemma.

Lemma 2.1. *Let* $\mathcal{C}_z \overset{\text{def}}{=} \sigma(Z_1, \cdots, Z_N)$, *where* $\sigma(\cdot)$ *denotes a sigma field. Then, under Assumption 2.1 and conditional on* $\mathcal{C}_z$, *the process* $\{\mathbb{Y}_t\}_t$ *is strictly stationary as well as covariance stationary.*

We introduce the NED concept in Section 3.2 to ensure that the dependency of the processes is decaying appropriately, which is the key in proving the consistency and asymptotic normality of the proposed IVQR estimator. In sum, stationarity is required for moment estimation and forecasting while the NED property is utilized to prove the parameter consistency and asymptotic normality.

Once $\mathbb{Y}_t$ is shown to be strictly stationary, $\mathbb{Y}_t$ is covariance stationary if $\mathsf{Var}(\mathbb{Y}_t)$ and $\Gamma_l = \mathrm{Cov}(\mathbb{Y}_t, \mathbb{Y}_{t-l})$ exist. Rewrite the model (3) as

$$\mathbb{Y}_t = S_t^{-1} \mathbf{H}_t \mathbb{Y}_{t-1} + S_t^{-1} \mathbf{B}_t \mathbb{F}_t + S_t^{-1} \mathbf{A}_{0t} \tag{10}$$

where $S_t = I_N - \mathbf{A}_{1t}W$. Then we have the following covariance stationary solution:

$$\mathbb{Y}_t = \sum_{l=0}^{\infty} \Pi_l \mathbb{D}_{t-l} = \sum_{l=0}^{\infty} \Pi_l S_{t-l}^{-1} \mathbf{B}_{t-l} \mathbb{F}_{t-l} + \sum_{l=0}^{\infty} \Pi_l S_{t-l}^{-1} \mathbf{A}_{0t}, \tag{11}$$

where $\mathbb{D}_t = S_t^{-1}(\mathbf{B}_t \mathbb{F}_t + \mathbf{A}_{0t})$, $M_t = S_t^{-1}\mathbf{H}_t$ and $\Pi_l = M_t \times \cdots \times M_{t-l+1}$ for $l > 1$ with $\Pi_0 = I_N$ and $\Pi_1 = M_t$. In the Online Appendix A.1 we prove that the covariance of $\mathbb{Y}_t$ exist under Assumption 2.1.

2.2 Asymptotic Stationary Distribution

Define any vector $a \in \mathbb{R}^N$ with $|a|_2 = 1$ and fixed d number of non zero elements. Let $\tilde{\mathbb{Y}}_t = \mathbb{Y}_t - \mu_{\mathbb{Y}}$, $L_T = \sum_{t=1}^{T} a^\top \tilde{\mathbb{Y}}_t$, and $L_t = L_{\lfloor t \rfloor} + (t - \lfloor t \rfloor) a^\top \tilde{\mathbb{Y}}_{\lfloor t \rfloor + 1}, t \geqslant 1$, where $\mu_{\mathbb{Y}} = \mathsf{E}(\mathbb{Y}_t)$ and $\lfloor t \rfloor = \max\{k \in \mathbb{Z} : k \geqslant t\}$ is the floor function. We then show that the average response is asymptotically normally distributed.

Theorem 1. *Consider any vector $a \in \mathbb{R}^N$ with $|a|_2 = 1$ and fixed $d < N$ number of nonzero elements. Under Assumption 2.1 and conditional on $\mathcal{C}_z$, then*

$$\frac{L_{T\nu}}{\sqrt{T}} \Rightarrow \sigma_{a\mathbb{Y}} \mathfrak{B}(\nu), \quad 0 \leqslant \nu \leqslant 1 \tag{12}$$

where $\sigma_{a\mathbb{Y}}^2 \overset{\text{def}}{=} \sum_{l \geqslant 0} a^\top \Gamma_l a$ is the long run variance of $a^\top \tilde{\mathbb{Y}}_t$ and $\mathfrak{B}(\nu)$ $(0 \leqslant \nu \leqslant 1)$ is a Brownian motion.

Remark For $\nu = 1$, Theorem 1 implies:

$$\sqrt{T}(a^\top(\overline{\mathbb{Y}} - \mu_{\mathbb{Y}})) \xrightarrow{\mathcal{L}} \mathsf{N}(0, \sigma_{a\mathbb{Y}}^2), \quad \text{as} \quad T \to \infty. \tag{13}$$

where $\overline{\mathbb{Y}} = T^{-1} \sum_{t=1}^{T} \mathbb{Y}_t$. Thus, the mean of $\mathbb{Y}_t$ converges in law to a normal distribution.

3 The IVQR Estimation

We first introduce the estimation algorithms of the IVQR approach. We then discuss the underlying assumptions and develop the asymptotic theories.

3.1 IVQR Estimator

Suppose that there exists an $N \times \ell$ matrix of instrumental variables (IV), denoted $\mathbf{R}_t = (R_{1t}, \cdots, R_{Nt})^\top \in \mathbb{R}^{N\times\ell}$, which is assumed to be independent of U_{it}. Then, we have the following quantile conditions:

$$\mathrm{P}\left(Y_{it} \leqslant \gamma_1^0(\tau)\overline{Y}_{it} + X_{it}^\top \phi^0(\tau) | X_{it}, R_{it}\right) = \tau \ a.s. \tag{14}$$

where $\overline{Y}_{it} = \sum_{j=1}^N w_{ij} Y_{jt}$ and $X_{it} = \left(1, Z_i^\top, \overline{Y}_{i,t-1}, Y_{i,t-1}, F_t^\top, \cdots, F_{t-p}^\top\right)^\top$ with $\phi^0(\tau) = [\gamma_0^0(\tau), \alpha_1^0(\tau), ..., \alpha_q^0(\tau), \gamma_2^0(\tau), \gamma_3^0(\tau), \beta_0^{0\top}(\tau), ..., \beta_p^{0\top}(\tau)]^\top \in \mathbb{R}^{3+q+(p+1)m}$. The above conditional probability is a measurable function of (X_{it}, R_{it}).

In general, the valid IVs should satisfy the quantile conditions in (14), and do not lead to collinearity among R_{it} and X_{it}. See Theorem 3 for the asymptotic formula of the variance matrix of the IVQR estimator. The estimation efficiency will be improved by choosing R_{it} appropriately. Following the literature (see e.g. Su and Yang (2011)), we may choose R_{it} to be the higher network orders of lagged dependent variables such as $e_i^\top W^2 \mathbb{Y}_{t-1}$, $[e_i^\top W^2 \mathbb{Y}_{t-1}, e_i^\top W \mathbb{Y}_{t-2}]$ and so on, where e_i is a vector with unity on the i-th element and zeros otherwise. Based on the satisfactory simulation evidence reported in Section 4, we suggest using $[e_i^\top W^2 \mathbb{Y}_{t-1}, e_i^\top W^3 \mathbb{Y}_{t-1}]$ as IVs.

To solve (14) we need to find the unknown true parameters $(\gamma_1^0(\tau), \phi^{0\top}(\tau))^\top$ such that $\mathbf{0}$ is a solution to the quantile estimation of $Y_{it} - \gamma_1^0(\tau)\overline{Y}_{it} - X_{it}^\top \phi^0(\tau)$ on R_{it}:

$$\mathbf{0} \in \arg\min_{g\in\mathscr{G}} \mathsf{E}\left[\rho_\tau\left\{Y_{it} - \gamma_1^0(\tau)\overline{Y}_{it} - X_{it}^\top \phi^0(\tau) - g\left(R_{it}\right)\right\}\right], \tag{15}$$

where $\mathscr{G}$ is the class of measurable functions of R_{it} and $\rho_\tau(u) = u\{\tau - \mathbf{I}(u < 0)\}$ is the check function with $\mathbf{I}(\cdot)$ the indicator function. We then restrict $\mathscr{G}$ to the class of linear-in-parameter functions:

$$\mathscr{G} = \{g\left(R_{it}\right) = R_{it}^\top \lambda(\tau) : \lambda \in \Lambda\}, \tag{16}$$

where Λ is a compact set in $\mathbb{R}^\ell$. Alternatively, we may construct the transformed IVs by the least squares projection of $\overline{Y}_{it}$ on R_{it} as in Chernozhukov and Hansen (2005, 2006).

Then, we obtain the sample analogue of the objective function:

$$Q(\gamma_1(\tau), \phi(\tau), \lambda(\tau)) = \sum_{i=1}^{N}\sum_{t=1}^{T}\left[\rho_\tau\left\{Y_{it} - \gamma_1(\tau)\overline{Y}_{it} - X_{it}^\top\phi(\tau) - R_{it}^\top\lambda(\tau)\right\}\right]. \quad (17)$$

Let $\theta(\tau) = (\gamma_1(\tau), \phi^\top(\tau))^\top$ and $\eta(\tau) = (\phi^\top(\tau), \lambda^\top(\tau))^\top$. The IVQR estimator, $(\widehat{\gamma}_1(\tau), \widehat{\phi}^\top(\tau), \widehat{\lambda}^\top(\tau))^\top$, obtained by minimizing (17), is expected to converge to the true parameters, $(\gamma_1^0(\tau), \phi^{0\top}(\tau), \mathbf{0}^\top)^\top$. For a given value of endogenous parameter, $\tilde{\gamma}_1(\tau)$, over a grid set of the interval $(-1, 1)$, we first run the ordinary QR of $Y_{it} - \tilde{\gamma}_1(\tau)\overline{Y}_{it}$ on (X_{it}, R_{it}) and obtain the corresponding estimator, denoted $\widehat{\eta}(\tilde{\gamma}_1(\tau), \tau) = \left[\widehat{\phi}^\top(\tilde{\gamma}_1(\tau), \tau), \widehat{\lambda}^\top(\tilde{\gamma}_1(\tau), \tau)\right]^\top$. Next, we select $\tilde{\gamma}_1(\tau)$ which minimizes $|\widehat{\lambda}(\tilde{\gamma}_1(\tau), \tau)|_2^2$ over the interval $(-1, 1)$, denoted as $\widehat{\gamma}_1(\tau)$. The IVQR estimator of $\theta(\tau)$ is then obtained by $\widehat{\theta}(\tau) = (\widehat{\gamma}_1(\tau), \widehat{\phi}^\top(\widehat{\gamma}_1(\tau), \tau))^\top$. For a given quantile index τ, the IVQR estimation can proceed as follows:

Step 1. For a given value of $\tilde{\gamma}_1(\tau)$, run the QR of $Y_{it} - \tilde{\gamma}_1(\tau)\overline{Y}_{it}$ on $(X_{it}^\top, R_{it}^\top)^\top$ and obtain:

$$\widehat{\eta}(\tilde{\gamma}_1(\tau), \tau) = \arg\min_{(\phi,\lambda)} Q(\tilde{\gamma}_1(\tau), \phi(\tau), \lambda(\tau)). \quad (18)$$

Step 2. Minimize a weighted norm of $\widehat{\lambda}(\tilde{\gamma}_1(\tau), \tau)$ over $\tilde{\gamma}_1(\tau)$ to obtain the IVQR estimator of $\gamma_1(\tau)$:

$$\widehat{\gamma}_1(\tau) = \arg\min_{\tilde{\gamma}_1 \in (-1,1)} \widehat{\lambda}^\top(\tilde{\gamma}_1(\tau), \tau)\ \mathrm{A}\ \widehat{\lambda}(\tilde{\gamma}_1(\tau), \tau), \quad (19)$$

where A is some positive definite matrix. Without loss of generality we set $\mathrm{A} = I$ throughout the paper.

Step 3. Run the QR of $Y_{it} - \widehat{\gamma}_1(\tau)\overline{Y}_{it}$ on X_{it}, and obtain the estimator of $\phi(\tau)$, denoted $\widehat{\phi}(\tau) = \widehat{\phi}(\widehat{\gamma}_1(\tau), \tau)$. Finally, we obtain the IVQR estimator by

$$\widehat{\theta}(\tau) = (\widehat{\gamma}_1(\tau), \widehat{\phi}^\top(\tau))^\top = (\widehat{\gamma}_1(\tau), \widehat{\phi}^\top(\widehat{\gamma}_1(\tau), \tau))^\top. \quad (20)$$

3.2 Asymptotic Theory

To develop the asymptotic theory for the IVQR estimator, we need to deal with some topological properties of $\mathbb{Y}_t$ that are spatially and temporally dependent. We follow Jenish and Prucha (2009, 2012) and utilize NED to address the spatial dependence of the statistics. The derivation of the asymptotic property follows from the standard M-estimation, including the quantile loss function as a special case. First, conditional on the common factors, we show in Section 3.2.1 that the elements of $\{\mathbb{Y}_t\}_t$ is an NED process. Then, in Section 3.2.2, we derive the asymptotic distribution of the IVQR estimator under certain regularity conditions. As we aim to apply the DNQR model to a network dataset with the large number of nodes and time periods, we mainly employ the large N and large T asymptotics, though the asymptotic theory can be equally developed for large T and fixed N or fixed N and large T (as pointed out by an anonymous referee.)

3.2.1 NED Properties of the Network Processes

The NED process definition adopted in this paper is developed by Jenish and Prucha (2012) that extends the notion of NED processes used in the time series to random fields. It is a more generalized dependence concept than the mixing dependence. The resulting process therefore can accommodate a wide range of models with weak spatial dependence. In particular, the NED property is preserved under general data transformations that does not necessarily hold for a mixing process. We first review some theories of NED random fields in Jenish and Prucha (2012).

The observations for each node can be modeled as a realization of a dependent heterogeneous process indexed by a point in $\mathbb{R}^d$ with $d \geqslant 1$. We consider a random field $D \subseteq \mathbb{R}^d$. The space $\mathbb{R}^d$ is endowed with the metric $\rho(j, j') = \max_{1 \leqslant l \leqslant d} |j_l - j'_l|$ with the corresponding norm, $|j| = \max_{1 \leqslant l \leqslant d} |j_l|$, where j_l is the l-th element of j. The distance between any subsets $U, V \subseteq D$ is defined as $\rho(U, V) = \inf\{\rho(j, j') : j \in U \text{ and } j' \in V\}$. Let $|U|$ denote the cardinality of a finite subset, U. In the two dimensional case with $d = 2$ and $j = j(i, t)$, we have: $\rho((i, t), (i', t')) = \max(|i - i'|, |t - t'|)$.

Assumption 3.1. *Let the lattice $D_{NT} \subseteq D \subseteq \mathbb{R}^d$ with $d = 2$, be countably infinite, where*

the cardinality of D_{NT} *satisfies* $\lim_{N,T\to\infty} |D_{NT}| \to \infty$. *Then,* $\rho(j,j') \geqslant \rho_0, \forall j,j' \in D$, *where* ρ_0 *is a constant. We set* $\rho_0 = 1$ *w.l.o.g.*

The minimum distance assumption in Assumption 3.1 is used for increasing domain asymptotics. It ensures the growth of the sample size as the sample regions D_{NT} expands. The setting is introduced in Jenish and Prucha (2012) for spatial mixing and NED processes. Note that the space D can be a space of socio-economic characteristics or geographical space, and the metric is not restricted to physical distance.

Definition 3.1 (NED). *Let* $\mathcal{Z} = \{\mathcal{Z}_{it}, (i,t) \in D_{NT}, NT \geqslant 1\}$ *and* $\zeta = \{\zeta_{it}, (i,t) \in D_{NT}, NT \geqslant 1\}$ *be random fields with* $\|\mathcal{Z}_{it}\|_{p'} < \infty$ *for* $p' \geqslant 1$, *where* $D_{NT} \subseteq D$ *with its cardinality given by* $|D_{NT}| = NT$. *Let* $\{d_{it}, (i,t) \in D_{NT}, NT \geqslant 1\}$ *be an array of positive constants. Then, the random field* $\mathcal{Z}$ *is* $L_{p'}$*-NED on the random field* ζ *if*

$$\|\mathcal{Z}_{it} - \mathsf{E}(\mathcal{Z}_{it}|\mathscr{F}_{it}(s))\|_{p'} < d_{it}\varphi(s),$$

for some sequence $\varphi(s) \geqslant 0$ *with* $\lim_{s\to\infty} \varphi(s) = 0$, *where* $\varphi(s)$ *is the NED coefficient,* d_{it} *is the NED scaling factor, and* $\mathscr{F}_{it}(s) = \sigma(\zeta_{i't'} : (i',t') \in D_{NT}, \rho((i',t'),(i,t)) \leqslant s)$ *is the* σ*-field generated by* $\zeta_{i't'}$ *within distance* s *from* (i,t). *If* $\sup_{N,T} \sup_{(i,t)\in D_{NT}} d_{it} < \infty$, *then* $\mathcal{Z}$ *is uniformly* $L_{p'}$*-NED on* ζ.

The above definition actually requires $\mathcal{Z}_{it}$ to have decayed dependence both in terms of the time and cross-sectional distance. Thus $\psi(s)$ can be made arbitrarily small by increasing the size of the neighborhood, i.e. s. In this paper, we consider the L_2-NED properties of random field $\mathcal{Z}$ on some α-mixing random field.

Define $\mathcal{C}_f \overset{\text{def}}{=} \sigma(\mathbb{F}_t, \cdots, \mathbb{F}_{t-p})$, $\mathcal{C}_z \overset{\text{def}}{=} \sigma(Z_1, \cdots, Z_N)$, where $\sigma(\cdot)$ denotes a sigma field, and $\mathcal{C} = \mathcal{C}_f \cup \mathcal{C}_z$. We now discuss the NED properties of $\{Y_{it}\}_{i,t}$ on the base $\{U_{it}\}_{i,t}$, where $\mathscr{F}_{it}(s) = \sigma(U_{i't'}, \mathcal{C} : (i',t') \in D_{NT}, \rho((i',t'),(i,t)) \leqslant s)$ is the σ-field generated by random vectors, $U_{i't'}$ located within distance s from (i,t).

Notice that the innovation U_{it} is assumed to be *i.i.d.* over i and t in Assumption 2.1(1), though it is well-known that *i.i.d.* is a special case of α-mixing. The above condition implies that $\{U_{it}\}_{i,t}$ is an α-mixing random field.

Following Xu and Lee (2015), we outline some conditions on NED properties of $\mathbb{Y}_t$.

Assumption 3.2. *The network matrix W is non-stochastic with zero diagonals and uniformly bounded for all N with absolute row and column sums such that the matrix $S_t = I_N - \mathbf{A}_{1t}W$ is nonsingular. We consider two cases for $w_{ij} \geqslant 0$ for any i, j.*

(1) Case 1: $|w_{ij}| \leqslant \pi_0 \rho(i,j)^{-c_w}$ with constants $\pi_0 \geqslant 0$ and $c_w > d$. In addition, there exists at most $K(\geqslant 1)$ number of columns in W, with $\min_u |\gamma_1^0(u)| \sum_{i=1}^n |w_{ij}| > \Upsilon$, where K is an integer and the positive constant Υ is defined in Assumption 2.1 (1).

(2) Case 2: Two nodes influence each other only if they are located sufficiently close; namely, $w_{ij} \neq 0$ if $\rho(i,j) \leqslant \bar{\rho}_0$ and $w_{ij} = 0$ otherwise, where we set the constant $\bar{\rho}_0 > 1$ w.l.o.g.

Assumption 3.2 is mainly used to restrict the NED coefficients, $\varphi(s) \to 0$ as $s \to \infty$. Assumption 3.2 (1) allows two individuals to have direct interaction even though their locations are far away from each other, with the requirement that the strength of interaction w_{ij} declines with the distance $\rho(i,j)$ in the power of c_w. This assumption is in line with Xu and Lee (2015). By excluding a limited number of nodes $K(\geqslant 1)$, the total effects on other units from each node should be bounded, i.e., we assume that $\sup|\gamma_1^0(u)| \sup_j \sum_{i=1}^N |w_{ij}| < \Upsilon$ or $\sup|\gamma_1^0(u)| \sup_j \sum_{i=1}^N |w_{ij}| < 1$ w.l.o.g. This corresponds to the existence of a narrow number of units with large aggregate effects on others even as the total number of nodes rises. Assumption 3.2(2) allows two individuals to have direct interaction only if they are located within a specific distance. Notice that this assumption does not allow star nodes, but one can see Pesaran and Yang (2020, 2021) and Kapetanios et al. (2021) for some extensions.

Let $u_{it} = u_{it}(\gamma_1, \phi, \lambda, \tau) = Y_{it} - \gamma_1(\tau)\overline{Y}_{it} - X_{it}^\top \phi(\tau) - R_{it}^\top \lambda(\tau)$ with the check function, $\rho_\tau(u) = (\tau - \mathbf{I}(u \leqslant 0))u$ and $\psi_\tau(u) = \tau - \mathbf{I}(u \leqslant 0)$ (the directional derivative of $\rho_\tau(u)$). Proposition 3.1 provides the NED properties of $\{Y_{it}\}_{i,t}$, and its transformations $\{\rho_\tau(u_{it})\}_{i,t}$, $\{\psi_\tau(u_{it})\}_{i,t}$ on the base $\{U_{it}\}_{i,t}$.

Proposition 3.1. *(1) Under Assumptions 2.1(1), 3.1 and 3.2(1), and conditional on $\mathcal{C}$, $\{Y_{it}\}_{i,t}$ is uniformly L_2-NED on $\{U_{it}\}_{i,t}$ such that $\|Y_{it} - \mathsf{E}(Y_{it}|\mathscr{F}_{it}(s))\|_2 < Cs^{-(c_w-d)}$ for $c_w > d$ and some $C > 0$ that does not depend on i and t. The same conclusion holds for $\{u_{it}\}_{i,t}$. The transformations $\{\psi_\tau(u_{it})\}_{i,t}$ and $\{\rho_\tau(u_{it})\}_{i,t}$ are also L_2-NED on $\{U_{it}\}_{i,t}$.*

(2) Under Assumptions 2.1(1), 3.1 and 3.2(2), and conditional on $\mathcal{C}$, $\{Y_{it}\}_{i,t}$ is uniformly L_2-NED on $\{U_{it}\}_{it}$ such that $\|Y_{it} - \mathsf{E}(Y_{it}|\mathscr{F}_{it}(s))\|_2 < C\Upsilon^{s/\bar{\rho}_0}$ $(\Upsilon < 1)$ for some $C > 0$ that does not depend on i and t. The same conclusion holds for $\{u_{it}\}_{i,t}$. The transformations, $\{\psi_\tau(u_{it})\}_{i,t}$ and $\{\rho_\tau(u_{it})\}_{i,t}$ are also L_2-NED on $\{U_{it}\}_{i,t}$.

Define $s_{it}(\gamma_1^0, \eta^0(\gamma_1^0,\tau),\tau) = \psi_\tau\left\{Y_{it} - \gamma_1^0(\tau)\overline{Y}_{it} - \Psi_{it}^\top \eta^0(\gamma_1^0,\tau)\right\}\Psi_{it}$ where $\Psi_{it} = (R_{it}^\top, X_{it}^\top)^\top$, $\check{s}_{it} = \check{s}_{it}(\gamma_1^0, \eta^0(\gamma_1^0,\tau),\tau) = s_{it}(\gamma_1^0, \eta^0(\gamma_1^0,\tau),\tau) - \mathsf{E}\, s_{it}(\gamma_1^0, \eta^0(\gamma_1^0,\tau),\tau)$ and $\mathsf{E}\, s_{it}(\gamma_1^0, \eta^0(\gamma_1^0,\tau),\tau) = 0$. Conditioning on $\mathcal{C}$, it is easily seen that the process $\{\check{s}_{it}\}_{i,t}$ is also uniformly L_2-NED on $\{U_{it}\}_{i,t}$. To derive the central limit theorem for $G_{NT}^0 = \frac{1}{\sqrt{NT}}\sum_{i=1}^N\sum_{t=1}^T \check{s}_{it} = \frac{1}{\sqrt{NT}}\sum_{i=1}^N\sum_{t=1}^T[s_{it}(\gamma_1^0, \eta^0(\gamma_1^0,\tau),\tau) - \mathsf{E}\, s_{it}(\gamma_1^0, \eta^0(\gamma_1^0,\tau),\tau)]$, where the variance of G_{NT}^0 is given by $\Omega_0 = \tau(1-\tau)\lim_{N,T\to\infty}(NT)^{-1}\sum_{i,t}\mathsf{E}(\Psi_{it}\Psi_{it}^\top|\mathcal{C})$, we make the following assumptions:

Assumption 3.3. *$\{\rho_\tau(u_{it})\}_{i,t}$ is uniformly $L_{p'}$-bounded for $p' > 1$, i.e., $\sup_{(\gamma_1,\phi,\lambda,\tau)}\sup_{N,T}\sup_{(i,t)\in D_{NT}}\mathsf{E}\,|\rho_\tau(u_{it})|^{p'} < \infty$, where p' is an integer.*

Assumption 3.4. *(Uniform $L_{2+\delta}$ Integrability)*

(1) $\{\check{s}_{it}\}_{i,t}$ is uniformly $L_{2+\delta}$ integrable for some $\delta > 0$, i.e., $\lim_{k\to\infty}\sup_{N,T}\sup_{(i,t)\in D_{NT}}\mathsf{E}\{|\check{s}_{it}|^{2+\delta}\mathbf{I}(|\check{s}_{it}| > k)\} = 0$.

(2) Ω_0 exists and is positive definite.

(3) NED coefficients of $\{\check{s}_{it}\}_{i,t}$ satisfy: $\sum_{h=1}^\infty h^{d-1}\varphi(h) < \infty$.

Assumption 3.3 imposes the moment conditions of $\rho_\tau(u_{it})$ and requires the existence of moments of order slightly greater than 1, which is employed in weak law of large numbers to achieve the uniform consistency. Assumption 3.4 sets the mixing coefficients of the underlying mixing fields and is utilized for the limit theory. Since U_{it} is assumed to be *i.i.d.*, these conditions are automatically satisfied.

3.2.2 Asymptotic Distribution of the IVQR Estimator

Assumption 3.5 (Conditions for identification and estimation)**.** *(1) (Compactness and Convexity) For all τ, $(\gamma_1(\tau), \phi(\tau)) \in \mathcal{A}\times\mathcal{B}$, where $\mathcal{A}\times\mathcal{B}$ is compact and convex.*

(2) (Full Rank and Continuity) $\mathbb{Y}_t$ *has bounded conditional density, a.s.* $\sup_{\mathbb{Y}_t\in\mathbb{R}^N} f_{\mathbb{Y}_t|\mathcal{F}_t}(y) < \infty$*, where* $\mathcal{F}_t = \{Z_1, \cdots, Z_N, \mathbb{Y}_{t-1}, \mathbb{Y}_t, F_t, F_{t-1}, \cdots, F_{t-p}\}$ *is the information set. Define*

$$
\begin{aligned}
S_{NT}(\pi,\tau) &= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[\psi_\tau\left\{Y_{it}-\gamma_1(\tau)\overline{Y}_{it}-X_{it}^\top\phi(\tau)-R_{it}^\top\lambda(\tau)\right\}\Psi_{it}\right], && (21)\\
S_\infty(\pi,\tau) &= \lim_{N,T\to\infty}\mathsf{E}\left[S_{NT}(\pi,\tau)|\mathcal{C}\right], \quad S_\infty^*(\pi,\tau)=\lim_{N,T\to\infty}\mathsf{E}\left[S_{NT}(\pi,\tau)\right], && (22)\\
S_{NT}(\theta,\tau) &= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[\psi_\tau\left\{Y_{it}-\gamma_1(\tau)\overline{Y}_{it}-X_{it}^\top\phi(\tau)\right\}\Psi_{it}\right], && (23)\\
S_\infty(\theta,\tau) &= \lim_{N,T\to\infty}\mathsf{E}\left[S_{NT}(\theta,\tau)|\mathcal{C}\right], \quad S_\infty^*(\theta,\tau)=\lim_{N,T\to\infty}\mathsf{E}\left[S_{NT}(\theta,\tau)\right], && (24)
\end{aligned}
$$

where $\pi \equiv (\gamma_1, \phi^\top, \lambda^\top)^\top$, $\theta \equiv (\gamma_1, \phi^\top)^\top$, $\psi_\tau(u) = \tau - \mathbf{I}(u<0)$, *and* $\Psi_{it} = (X_{it}^\top, R_{it}^\top)^\top$. *Then, the Jacobian matrices,* $\dfrac{\partial S_\infty(\theta,\tau)}{\partial(\gamma_1,\phi^\top)}$ *and* $\dfrac{\partial S_\infty(\pi,\tau)}{\partial(\phi^\top,\lambda^\top)}$ *are continuous and have full rank, uniformly over* $\mathcal{A}\times\mathcal{B}\times\mathcal{G}\times\mathcal{T}$*, where* $\mathcal{G}$ *is a compact set with* $\lambda(\tau)\in\mathcal{G}$*,* $\mathcal{T}$ *is a compact set with* $\tau\in\mathcal{T}$*, and the image of* $\mathcal{A}\times\mathcal{B}$ *under the mapping* $\theta\equiv(\gamma_1,\phi^\top)^\top \mapsto S_\infty(\theta,\tau)$ *is simply connected.*

(3) For a fixed $\tau\in\mathcal{T}$*, the unknown true parameter,* $\theta^0(\tau) = (\gamma_1^0(\tau), \phi^{0\top}(\tau))^\top$ *uniquely solves* $S_\infty(\theta,\tau)=0$ *over* $\mathcal{A}\times\mathcal{B}$*.*

The compactness of the parameter space in Assumption 3.5(1) is needed for $\gamma_1(\tau)$ due to the non-convex objective function. Assumption 3.5(2) implies the global identification of the parameters while the continuity condition is required for deriving the asymptotic normality. Assumption 3.5(3) requires that $\theta^0(\tau) = (\gamma_1^0(\tau), \phi^{0\top}(\tau))^\top$ to be the unique solution to $S_\infty(\theta,\tau)=0$, which is necessary for consistency of the estimator.

Let $\widehat{\theta}(\tau) = (\widehat{\gamma}_1(\tau), \widehat{\phi}^\top(\tau))^\top$ be the IVQR estimator of $\theta^0(\tau) = (\gamma_1^0(\tau), \phi^{0\top}(\tau))^\top$, where $\widehat{\phi}(\tau) = \widehat{\phi}(\widehat{\gamma}_1(\tau),\tau)$. Define the $(4+q+(p+1)m)\times(4+q+(p+1)m)$ matrices:

$$
J(\tau) = \left.\frac{\partial S_\infty(\pi,\tau)}{\partial(\gamma_1,\phi^\top)}\right|_{\gamma_1=\gamma_1^0,\phi=\phi^0,\lambda=\mathbf{0}}, \quad J^*(\tau) = \left.\frac{\partial S_\infty^*(\pi,\tau)}{\partial(\gamma_1,\phi^\top)}\right|_{\gamma_1=\gamma_1^0,\phi=\phi^0,\lambda=\mathbf{0}}. \tag{25}
$$

Theorem 2 (Linearization)**.** *Under Assumptions 3.1–3.3 and 3.5, as* $\min\{N,T\}\to\infty$*,*

$$
\sqrt{NT}\left\{\widehat{\theta}(\tau)-\theta^0(\tau)\right\} = -J^{-1}(\tau)G_{NT}^0(\theta^0,\tau)+o_p(1). \tag{26}
$$

Under Assumption 3.4, the NED process $\{\check{s}_{it}\}_{i,t}$ satisfies the central limit theorem:

$G^0_{NT}(\theta^0,\tau) = \frac{1}{\sqrt{NT}}\sum_{i=1}^N\sum_{t=1}^T \check{s}_{it}$ follows a zero mean Gaussian process with covariance function, $\Omega_0 = \tau(1-\tau)\lim_{N,T\to\infty}(NT)^{-1}\sum_{i,t}\mathsf{E}(\Psi_{it}\Psi_{it}^\top)$.

Theorem 3. *Under Assumption 3.1–3.5, we have* $\Omega_0^{-1}\Omega_0^* \to_p I$, *and* $J^{-1}(\tau)J^*(\tau) \to_p I$, *where* $\Omega_0 = \tau(1-\tau)\lim_{N,T\to\infty}(NT)^{-1}\sum_{i,t}\mathsf{E}(\Psi_{it}\Psi_{it}^\top|\mathcal{C})$, $\Omega_0^* = \tau(1-\tau)\lim_{N,T\to\infty}(NT)^{-1}\sum_{i,t}\mathsf{E}(\Psi_{it}\Psi_{it}^\top)$, *and* $J(\tau), J^*(\tau)$ *are defined in (25). Then, as* $\min\{N,T\}\to\infty$,

$$\sqrt{NT}\left\{\widehat{\theta}(\tau) - \theta^0(\tau)\right\} \xrightarrow{d} \mathsf{N}(0,\Sigma_\theta), \tag{27}$$

where $\Sigma_\theta = J^*(\tau)^{-1}\Omega_0^* J^*(\tau)^{-1}$.

We estimate Ω_0^* and $J^*(\tau)$ consistently by

$$\widehat{\Omega}_0^* = (NT)^{-1}\tau(1-\tau)\sum_{i=1}^N\sum_{t=1}^T \Psi_{it}\Psi_{it}^\top, \tag{28}$$

$$\widehat{J}^*(\tau) = (2NTh_b)^{-1}\sum_{i=1}^N\sum_{t=1}^T \mathbf{I}(\widehat{u}_{it}\leqslant h_b)\Psi_{it}(\overline{Y}_{it}, X_{it}^\top), \tag{29}$$

where $\widehat{u}_{it} = u_{it}(\widehat{\theta}(\tau),\tau)$, and h_b is the bandwidth (see (30) below for details). Notice that the use of valid IVs do not result in $\widehat{J}^*(\tau)$ and $\widehat{\Omega}_0^*$ having or being closer to singularities, which causes the variance of the estimator $\widehat{\theta}(\tau)$ to be unreliably large.

4 Monte Carlo Simulations

In this section, we examine finite sample properties of the IVQR estimator via a Monte Carlo simulation study using three different network structures.

4.1 The Setup

We construct the data generating process based on the DNQR model as follows: First, we generate the five nodal covariates, $Z_i = (Z_{i1},\cdots,Z_{i5})^\top \in \mathbb{R}^5, (q=5)$ from a multivariate normal distribution $\mathsf{N}(\mathbf{0},\Sigma_z)$, where $\Sigma_z = (\sigma_{j_1j_2})$ and $\sigma_{j_1j_2} = 0.5^{|j_1-j_2|}$. Then, we construct the two common covariates, $F_t = (F_{1t}, F_{2t})^\top \in \mathbb{R}^2, (m=2)$ from the i.i.d standard normal

distribution. Let the true parameters $\gamma^0_{j,it} = \gamma^0_j(U_{it})$ for $j = 0, 1, 2, 3$, $\alpha^0_{j,it} = \alpha^0_j(U_{it})$ for $j = 1, ..., 5$, and $\beta^0_{jk,it} = \beta^0_{jk}(U_{it})$ for $j = 1, 2$ and $k = 0, 1$, where we set the lag of two common covariates to 1 ($p = 1$). We then generate the random coefficients by

$$\begin{aligned}
\gamma^0_{0,it} =& \mathfrak{u}_{it},\ \gamma^0_{1,it} = 0.1\Phi(\mathfrak{u}_{it}),\ \gamma^0_{2,it} = 0.4\{1 + \exp(\mathfrak{u}_{it})\}^{-1}\exp(\mathfrak{u}_{it}),\ \gamma^0_{3,it} = 0.4\Phi(\mathfrak{u}_{it}),\\
\alpha^0_{1,it} =& 0.5\Phi(\mathfrak{u}_{it}),\ \alpha^0_{2,it} = 0.3\mathbf{G}(\mathfrak{u}_{it}, 1, 2),\ \alpha^0_{3,it} = 0.2\mathbf{G}(\mathfrak{u}_{it}, 2, 2),\\
\alpha^0_{4,it} =& 0.25\mathbf{G}(\mathfrak{u}_{it}, 3, 2), \alpha^0_{5,it} = 0.2\mathbf{G}(\mathfrak{u}_{it}, 2, 1),\\
\beta^0_{10,it} =& 0.1\Phi(\mathfrak{u}_{it}),\ \beta^0_{11,it} = 0.3\mathbf{G}(\mathfrak{u}_{it}, 2, 2), \beta^0_{20,it} = 0.2\mathbf{G}(\mathfrak{u}_{it}, 1, 2),\ \beta^0_{21,it} = 0.3\mathbf{G}(\mathfrak{u}_{it}, 2, 1),
\end{aligned}$$

where $\Phi(\cdot)$ is the standard normal distribution function, $\mathbf{G}(\cdot, a, b)$ is the Gamma distribution function with shape parameter a and scale parameter b, and $\mathfrak{u}_{it}$s are i.i.d random variables, generated either from (a) the standard normal distribution or from (b) the t-distribution with 5 degrees of freedom. Notice that U_{it} can be generated by $U_{it} = F(\mathfrak{u}_{it})$, where $F(\cdot)$ is cumulative distribution function of $\mathfrak{u}_{it}$. Finally, $\mathbb{Y}_t$s are generated by (1).

To check the robustness of the finite sample performance of the IVQR estimator, we consider the following three different adjacency matrices, e.g. Zhu et al. (2019b).

Type 1. (Dyad Independence Model) Holland and Leinhardt (1981) introduce this model with a dyad, $D_{ij} = (a_{ij}, a_{ji})$ for $1 \leqslant i < j \leqslant N$, where D_{ij}s are assumed to be independent. We set the probability of dyad being mutually connected to $\mathrm{P}(D_{ij} = (1, 1)) = 2N^{-1}$ to ensure the network sparsity. Then, we set $\mathrm{P}(D_{ij} = (1, 0)) = \mathrm{P}(D_{ij} = (0, 1)) = 0.5N^{-0.8}$, which implies that the expected degree for each node is $O(N^{0.2})$. Accordingly, we have $\mathrm{P}(D_{ij} = (0, 0)) = 1 - 2N^{-1} - N^{-0.8}$, which tends to 1 as $N \to \infty$.

Type 2. (Stochastic Block Model) We first consider the Stochastic Block Model with an important application in community detection by Zhao et al. (2012). We follow Nowicki and Snijders (2001) and randomly assign each node a block label index from 1 to L, where $L \in \{5, 10, 20\}$. We then set $\mathrm{P}(a_{ij} = 1) = 0.3N^{-0.3}$ if i and j are in the same block, and $\mathrm{P}(a_{ij} = 1) = 0.3N^{-1}$ otherwise. Thus, the nodes within the same block have higher probability of connecting with each other than the nodes between blocks.

Type 3. (Power-law Distribution Network) In practice, the majority of nodes in the network have a small number of links while a small number of nodes have a large

number of links, see Barabási and Albert (1999). In this case the degrees of nodes can be characterized by the power-law distribution. We simulate the adjacency matrix as follows: For each node, we generate the in-degree, $d_i = \sum_j a_{ji}$ according to the discrete power-law distribution such as $\mathrm{P}(d_i = \check{k}) = c\check{k}^{-\beta}$, where c is a normalizing constant and the exponent parameter β is set at 2.5 as in Clauset et al. (2009). Finally, for the i-th node, we randomly select d_i nodes as its followers.

To estimate the variance of the IVQR estimator, we follow Koenker and Xiao (2006) and select the bandwidth h_b in (28) as follows (Hall and Sheather (1988)):

$$h_b = (NT)^{-1/3} \varrho_\alpha^{2/3} \left[\frac{1.5\bar{\varphi}^2(\Phi^{-1}(\tau))}{2(\Phi^{-1}(\tau))^2 + 1} \right]^{1/3}, \tag{30}$$

where $\bar{\varphi}(\cdot)$ and $\Phi(\cdot)$ are the probability density and distribution function of standard normal distribution and ϱ_α satisfies $\Phi(\varrho_\alpha) = 1 - \alpha/2$ for the construction of $1 - \alpha$ confidence intervals. We have also considered an alternative selection criterion by Bofinger (1975), and obtained qualitatively similar results that are available upon request.

For the IVQR estimation, we suggest using $\mathbf{R}_t = [W^2\mathbb{Y}_{t-1}, W^3\mathbb{Y}_{t-1}]$, where $\mathbf{R}_t = (R_{1t}, \cdots, R_{Nt})^\top$ and W is the row-sum normalized network matrix. Although we may select the higher network orders such as $[W^2\mathbb{Y}_{t-1}, W^3\mathbb{Y}_{t-1}, W^2\mathbb{Y}_{t-2}, W^3\mathbb{Y}_{t-2}, ...]$, we find that these two instruments are often the best choice.

4.2 Simulation Results

Using 1000 replications, we evaluate the finite sample performance of the IVQR estimator at the different quantiles, $\tau = 0.1, 0.5, 0.9$ for the (N, T) pairs with $N, T = 100, 200, 500$. Table 1 presents the simulation results for Type 1 network in terms of RMSE. Overall, RMSEs of all the parameters decrease monotonically as N or T increases across the different quantiles and for the different distributions of $\mathfrak{u}_{it}$, which is in line with the asymptotic theory. But, RMSEs of γ_1 are larger than those of other parameters, especially in small samples $(N, T = 100)$, which mainly reflects uncertainty associated with the selection of the IVs. As the sample size grows ($N = 500$ or $T = 500$), all RMSEs decline sharply. Turning to the simulation results for Type 2 and Type 3 networks, reported

in B1 and B2 in the Online Appendix, we observe qualitatively similar findings. In the Online Appendix, we also provide the simulation results for biases, see Tables B5–B7. The biases of the IVQR estimators are mostly negligible across the different quantiles, for the different distributions of $\mathfrak{u}_{it}$ and for all the sample pairs of (N, T).

[Insert Table 1 here]

Next, in Table 2, we report the coverage probability for the TYPE 1 network by evaluating if the estimates fall into the 95% confidence interval at each replication. Overall, we find that the coverage probabilities of all the parameters are close to the nominal 95% level across the different quantiles and the different distributions of $\mathfrak{u}_{it}$, and for all the sample pairs (N, T). This implies the accuracy of the inference for the IVQR estimator. From the results for TYPE 2 and TYPE 3 networks, reported in Tables B3 and B4 in the Online Appendix, we also find that the coverage probabilities of all parameters are close to the nominal 95% level.

[Insert Table 2 here]

For comparison, we report the simulation results by applying the ordinary QR estimator in Tables B8–B16 in the Online Appendix B.2. We observe that RMSEs are larger. More importantly, RMSEs barely decrease with the sample size, especially for γ_1. Furthermore, coverage probabilities are well below the nominal 95% level, and the biases are large especially for γ_1, which remain substantial as the sample size increases. This clearly demonstrates the importance of using the IVQR estimator for the DNQR model.

5 Application

We explore the financial network quantile connectedness among the stock returns. Anton and Polk (2014) find that stock returns tend to display significant comovements due to common active mutual fund owners. In addition, Pirinsky and Wang (2006) document strong comovements in the stock returns of firms headquartered in the same geographic area. Garcia and Norli (2012) point out that the firms headquartered in the same geographic area have achieved uniformly excessive returns compared to geographically dis-

persed firms (the return local bias).

We consider the two different financial network structures: the common shareholder based network and the headquarter location based network. Notice that the order of network nodes can be randomly pre-determined and has no effect on the estimation results. For simplicity, we set up $\mathbb{Y}_t \in \mathbb{R}^N$ in which the row elements are arranged in alphabetical order by the unique trading code. Then the order of nodes in the pre-determined networks are the same as the individual order in $\mathbb{Y}_t$, that is alphabetically ordered by the unique stock trading code. The pre-determined networks are constructed by using information on the common mutual fund ownership and the uniform headquarter location. In particular, we let the stocks be connected if they are invested in by at least five common shareholders (W_{CS5}) while the companies with headquarters located in the same city are treated as connected (W_{HQ}).

We collect the data on all the stocks traded in NYSE and NASDAQ in 2016 from Datastream. The dataset on mutual fund holdings are downloaded from Thomson Reuters whilst the addresses of firms' headquarters are collected from COMPUSTAT. After merging these data from the databases according to the unique trading code and removing the stocks with missing values, we finally obtain 943 stocks ($N = 943$) over the whole time period $T = 252$. We then collect these stock return data from Datastream. We also obtain node-specific covariates such as market capitalization, book value per share, cash flow and price-earning ratio from Datastream, which are then standardized. Finally, we collect VIX from Datastream and the Fama-French three factors (excess market return, SMB, HML) from the website of French's homepage as the common covariates.

The network density is 3.24% for W_{CS5} and 0.63 % for W_{HQ}, respectively. In Figure 1 we display the topology of two networks for the top 100 market-value stocks only for visualization convenience. The larger nodes imply the higher market capitalization while the darker nodes present the higher connectedness especially for the network with W_{CS5}. Here we observe quite different network structures. There is a large connected group in Figure 1a, showing that the stocks are more centrally connected by common investors. On the contrary Figure 1b displays more small groups, implying that the stocks are more locally connected when the network is measured by uniform headquarter locations.

[Insert Figure 1 here]

Following the simulation results, we select IVs as $\mathbf{R}_t = [W^2\mathbb{Y}_{t-1}, W^3\mathbb{Y}_{t-1}]$. We present the estimation results for the proposed DNQR model together with the two alternative models: (i) the original NQAR model without contemporaneous network effects and common covariates, and (ii) the factor-augmented NQAR model without contemporaneous network effects, denoted NQARF. To compare the relative performance of the alternative models, we follow Koenker and Machado (1999) and evaluate the goodness of fits across the different quantiles. Consider a linear model for the conditional quantile function,

$$Q_{Y_{it}}(\tau|X_{it}) = X_{1it}^\top\theta_1(\tau) + X_{2it}^\top\theta_2(\tau) \tag{31}$$

where $X_{it} = (X_{1it}^\top, X_{2it}^\top)^\top$. Let $\widehat{\theta}(\tau) = (\widehat{\theta}_1^\top(\tau), \widehat{\theta}_2^\top(\tau))^\top$ be the unrestricted estimator, which is the minimizer of $\widehat{V}(\tau) = \min \sum_{i=1}^{N}\sum_{t=1}^{T} \rho_\tau\left\{Y_{it} - X_{it}^\top\theta\right\}$ while $\widetilde{\theta}(\tau) = (\widetilde{\theta}_1^\top(\tau), \mathbf{0}^\top)^\top$ denotes the minimizer for the constrained model, $\widetilde{V}(\tau) = \min \sum_{i=1}^{N}\sum_{t=1}^{T} \rho_\tau\left\{Y_{it} - X_{1it}^\top\theta_1\right\}$. Define the goodness-of-fit criterion as

$$R^2(\tau) = 1 - \widehat{V}(\tau)/\widetilde{V}(\tau), \tag{32}$$

which measures the overall decreased percentage of the quantile loss function of the unrestricted model with respect to the restricted model.

The estimation results for the network W_{CS5} are presented in Table 3. For convenience we present the coefficients and the standard errors multiplied by 10^2. We find that the estimated contemporaneous network effects (γ_1) by the DNQR model, are significantly positive and dominate all other effects across all quantiles. The lagged network coefficient (γ_2) and the dynamic coefficient (γ_3) are also significant across quantiles with relatively smaller magnitudes. Furthermore, the goodness of fit, $R^2(\tau)$ reported in the last row, shows that the overall loss function of the DNQR model drops about 7%–9.5% relative to the NQAR and 6.9%–9.5% relative to the NQARF, respectively, suggesting that the contemporaneous network effects should be explicitly accommodated in the dynamic network quantile model. Moreover, the effects of node-specific covariates are significant across quantiles (except Cash) while those of common covariates are all significant across

quantiles.

[Insert Table 3 here]

Next, we display the QR coefficients across the different quantiles in Figure 2. The dashed line is the QR coefficient while the grey area indicates a kernel density based confidence band advanced by Powell (1991). The contemporaneous network and dynamic coefficients, $\widehat{\gamma}_1(\cdot)$ and $\widehat{\gamma}_3(\cdot)$, are significant (as their bands exclude the null effect), while the lagged diffusion network coefficient, $\widehat{\gamma}_2(\cdot)$ tends to be insignificant only at the middle. $\widehat{\gamma}_1(\cdot)$ shows a downward trend with the quantile level, suggesting that the contemporaneous quantile connectedness is stronger at the lower tails (mainly characterised with the market turmoils). On the other hand, both $\widehat{\gamma}_2(\cdot)$ and $\widehat{\gamma}_3(\cdot)$ display the U-shaped pattern, implying that their effects are stronger at the tails than at the median. The QR effects of node-specific covariates mostly display a downward trend with the quantile level (except for insignificant Cash), suggesting their effects are stronger at the lower tails than at the upper tails (mainly characterised with the bulls market). Turning to the QR effects of common factors, we observe a mixed finding: the impacts of VIX and the market factor increase with quantiles whilst those of SMB and HML factors decrease with quantiles.

[Insert Figure 2 here]

Finally, as the robustness check, we provide the two additional estimation results in the Online Appendix. First, we reconstruct the network matrix by changing the number of common shareholders to $CS = 3$ (W_{CS3}) and $CS = 7$ (W_{CS7}), and the results are reported in Tables C1–C2 and Figures C1–C2. Notice that the network density of W_{CS3} is dense at 25.25% and relatively sparse at 0.41% for W_{CS7}. Overall, we find qualitatively similar results to those reported for W_{CS5}. One notable observation is that the contemporaneous network effects measured by γ_1 tend to decrease monotonically as the network becomes more sparse. For example, at $\tau = 0.1$, $\widehat{\gamma}_1$ is estimated at 0.69 for W_{CS3}, 0.54 for W_{CS5}, and 0.35 W_{CS7}, respectively. Still, we find that the patterns of the quantile specific coefficients reported in Figures C1–C2 are qualitatively similar to those displayed in Figure 2.

Next, we estimate the models using the headquarter location network W_{HQ}, and present the estimation results in Table C3 and Figure C3 in the Online Appendix. Again,

we find the qualitatively similar results, highlighting the importance of the contemporaneous network effect, which is also stronger at the lower tails than at the upper tails.

6 Conclusion

We develop a dynamic network quantile model that accommodates both temporal and cross-sectional dependence. Using the predetermined network information, we analyse the dynamic quantile connectedness within a network topology. The distinguishing feature of the DNQR model lies in that the behavior/response of a given node is not only influenced by its previous behavior/response, but also connected with a weighted average of contemporaneous and lagged behaviors/responses from peers.

The main challenge associated with the DNQR model is the presence of endogeneity stemming from the simultaneous network effect. In this regard, we develop the IVQR estimation, and derive the consistency and asymptotic normality of the IVQR estimator using the NED property of the network process. Monte Carlo exercises confirm the satisfactory performance of the IVQR estimator with the predetermined internal IVs across different quantiles under the different network structures.

Finally, we demonstrate the usefulness of our proposed approach with an application to the dataset on the stocks traded in NYSE and NASDAQ in 2016. In particular, we find that the contemporary network effects are significant and dominant across all quantiles. Furthermore, their effects display a downward trend with the quantile level, suggesting that the contemporaneous quantile connectedness is stronger at the lower tails.

Supplementary Materials

The online supplement contains all technical proofs, additional simulation and application results, as well as the codes and data for the simulation and application.

Acknowledgements

We are mostly grateful for the insightful comments by the editor, the associated editor and three anonymous referees, which substantially improved the article. All errors remains on our own.

Funding

Xiu Xu's research is supported by the Natural Science Foundation of China (Grant No. 71803140). Yongcheol Shin and Weining Wang's research is partially supported by the ESRC (Grant Reference: ES/T01573X/1). Chaowen Zheng's research is supported by the Royal Economic Society.

References

Acemoglu, D., Ozdaglar, A., and Tahbaz-Salehi, A. (2015). Systemic risk and stability in financial networks. *American Economic Review*, 105(2):564–608.

Ando, T., Greenwood-Nimmo, M., and Shin, Y. (2021). Quantile connectedness: Modeling tail behavior in the topology of financial networks. *Management Science,* Forthcoming.

Anton, M. and Polk, C. (2014). Connected stocks. *Journal of Finance*, 69(3):1099–1127.

Barabási, A. L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439):509–512.

Bassett, G. and Koenker, R. (1978). Asymptotic theory of least absolute error regression. *Journal of the American Statistical Association*, 73(363):618–622.

Betz, F., Hautsch, N., Peltonen, T. A., and Schienle, M. (2016). Systemic risk spillovers in the european banking and sovereign network. *Journal of Financial Stability*, 25:206–224.

Bofinger, E. (1975). Estimation of a density function using order statistics. *Australian Journal of Statistics*, 17(1):1–7.

Chen, C. Y.-H., Härdle, W. K., and Okhrin, Y. (2019). Tail event driven networks of SIFIs. *Journal of Econometrics*, 208(1):282–298.

Chen, L.-Y. and Lee, S. (2018). Exact computation of GMM estimators for instrumental variable quantile regression models. *Journal of Applied Econometrics*, 33(4):553–567.

Chernozhukov, V., Fernandez-Val, I., and Weidner, M. (2020). Network and panel quantile effects via distribution regression. *Journal of Econometrics,* Forthcoming.

Chernozhukov, V. and Hansen, C. (2005). An IV model of quantile treatment effects. *Econometrica*, 73(1):245–261.

Chernozhukov, V. and Hansen, C. (2006). Instrumental quantile regression inference for structural and treatment effect models. *Journal of Econometrics*, 132(2):491–525.

Chernozhukov, V. and Hansen, C. (2008). Instrumental variable quantile regression: A robust inference approach. *Journal of Econometrics*, 142(1):379–398.

Cho, J. S., Kim, T.-h., and Shin, Y. (2015). Quantile cointegration in the autoregressive distributed-lag modeling framework. *Journal of Econometrics*, 188(1):281–300.

Clauset, A., Shalizi, C. R., and Newman, M. E. (2009). Power-law distributions in empirical data. *SIAM review*, 51(4):661–703.

Diebold, F. X. and Yilmaz, K. (2014). On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics*, 182(1):119–134.

Engle, R. F. and Manganelli, S. (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics*, 22(4):367–381.

Fafchamps, M. and Gubert, F. (2007). Risk sharing and network formation. *American Economic Review*, 97(2):75–79.

Firpo, S., Galvao, A. F., Pinto, C., Poirier, A., and Sanroman, G. (2021). GMM quantile regression. *Journal of Econometrics,* Forthcoming.

Forni, M. and Gambetti, L. (2010). The dynamic effects of monetary policy: A structural factor model approach. *Journal of Monetary Economics*, 57(2):203–216.

Frölich, M. and Melly, B. (2013). Unconditional quantile treatment effects under endogeneity. *Journal of Business & Economic Statistics*, 31(3):346–357.

Galvao, A. F., Montes-Rojas, G., and Park, S. Y. (2013). Quantile autoregressive distributed lag model with an application to house price returns. *Oxford Bulletin of Economics and Statistics*, 75(2):307–321.

Garcia, D. and Norli, O. (2012). Geographic dispersion and stock returns. *Journal of Financial Economics*, 106(3):547–565.

Hall, P. and Sheather, S. J. (1988). On the distribution of a studentized quantile. *Journal of the Royal Statistical Society: Series B (Methodological)*, 50(3):381–391.

Härdle, W. K., Wang, W., and Yu, L. (2016). Tenet: Tail-event driven network risk. *Journal of Econometrics*, 192(2):499–513.

Hautsch, N., Schaumburg, J., and Schienle, M. (2014). Forecasting systemic impact in financial networks. *International Journal of Forecasting*, 30(3):781–794.

Hautsch, N., Schaumburg, J., and Schienle, M. (2015). Financial network systemic risk contributions. *Review of Finance*, 19(2):685–738.

Holland, P. W. and Leinhardt, S. (1981). An exponential family of probability distributions for directed graphs. *Journal of the American Statistical Association*, 76(373):33–50.

Jenish, N. and Prucha, I. R. (2009). Central limit theorems and uniform laws of large numbers for arrays of random fields. *Journal of Econometrics*, 150(1):86–98.

Jenish, N. and Prucha, I. R. (2012). On spatial processes and asymptotic inference under near-epoch dependence. *Journal of Econometrics*, 170(1):178–190.

Kapetanios, G., Pesaran, M. H., and Reese, S. (2021). Detection of units with pervasive effects in large panel data models. *Journal of Econometrics*, 221(2):510–541.

Koenker, R. and Machado, J. A. (1999). Goodness of fit and related inference processes for quantile regression. *Journal of the American Statistical Association*, 94(448):1296–1310.

Koenker, R. and Xiao, Z. (2006). Quantile autoregression. *Journal of the American Statistical Association*, 101(475):980–990.

Liu, X. (2014). Identification and efficient estimation of simultaneous equations network models. *Journal of Business & Economic Statistics*, 32(4):516–536.

Machado, J. A. and Silva, J. S. (2019). Quantiles via moments. *Journal of Econometrics*, 213(1):145–173.

Nowicki, K. and Snijders, T. A. B. (2001). Estimation and prediction for stochastic blockstructures. *Journal of the American Statistical Association*, 96(455):1077–1087.

Pesaran, M. H. and Yang, C. F. (2020). Econometric analysis of production networks with dominant units. *Journal of Econometrics*, 219(2):507–541.

Pesaran, M. H. and Yang, C. F. (2021). Estimation and inference in spatial models with dominant units. *Journal of Econometrics*, 221(2):591–615.

Pirinsky, C. and Wang, Q. (2006). Does corporate headquarters location matter for stock returns? *Journal of Finance*, 61(4):1991–2015.

Powell, J. L. (1991). Estimation of monotonic regression models under quantile restrictions. In Barnett, W. A., Powell, J. L., and Tauchen, G. E., editors, *Nonparametric and Semiparametric Methods in Econometrics and Statistics*. Cambridge University Press, Cambridge, UK.

Su, L. and Hoshino, T. (2016). Sieve instrumental variable quantile regression estimation of functional coefficient models. *Journal of Econometrics*, 191(1):231–254.

Su, L. and Yang, Z. (2011). Instrumental variable quantile estimation of spatial autoregression models. *Unpublished Manuscript,* Singapore Management University.

White, H., Kim, T.-H., and Manganelli, S. (2015). VAR for VaR: Measuring tail dependence using multivariate regression quantiles. *Journal of Econometrics*, 187(1):169–188.

Wüthrich, K. (2019). A closed-form estimator for quantile treatment effects with endogeneity. *Journal of Econometrics*, 210(2):219–235.

Wüthrich, K. (2020). A comparison of two quantile models with endogeneity. *Journal of Business & Economic Statistics*, 38(2):443–456.

Xiao, Z. (2009). Quantile cointegrating regression. *Journal of Econometrics*, 150(2):248–260.

Xu, X. and Lee, L.-f. (2015). Maximum likelihood estimation of a spatial autoregressive Tobit model. *Journal of Econometrics*, 188(1):264–280.

Zhao, Y., Levina, E., and Zhu, J. (2012). Consistency of community detection in networks under degree-corrected stochastic block models. *Annals of Statistics*, 40(4):2266–2292.

Zhu, X. (2020). Nonconcave penalized estimation in sparse vector autoregression model. *Electronic Journal of Statistics*, 14(1):1413–1448.

Zhu, X., Chang, X., Li, R., and Wang, H. (2019a). Portal nodes screening for large scale social networks. *Journal of Econometrics*, 209(2):145–157.

Zhu, X., Huang, D., Pan, R., and Wang, H. (2020). Multivariate spatial autoregressive model for large scale social networks. *Journal of Econometrics*, 215(2):591–606.

Zhu, X. and Pan, R. (2020). Grouped network vector autoregression. *Statistica Sinca*, 30(3):1437–1462.

Zhu, X., Pan, R., Li, G., Liu, Y., and Wang, H. (2017). Network vector autoregression. *Annals of Statistics*, 45(3):1096–1123.

Zhu, X., Wang, W., Wang, H., and Härdle, W. K. (2019b). Network quantile autoregression. *Journal of Econometrics*, 212(1):345–358.

Table 1: RMSE (×100) for TYPE 1 Network

	Dist.	τ	γ_0	γ_1	γ_2	γ_3	α_1	α_2	α_3	α_4	α_5	β_1	β_2	β_3	β_4
								$N = 100$							
$T = 100$	$N(0,1)$	0.1	1.64	5.35	1.41	3.04	1.75	1.87	1.77	1.83	1.61	1.47	1.64	1.42	1.42
		0.5	1.49	4.75	1.19	2.66	1.38	1.58	1.46	1.51	1.31	1.17	1.33	1.14	1.13
		0.9	1.71	5.18	1.39	2.95	1.63	1.74	1.73	1.74	1.52	1.32	1.56	1.31	1.32
	$t(5)$	0.1	1.95	4.98	1.27	2.82	1.98	2.17	2.17	2.22	1.92	1.67	1.79	1.66	1.68
		0.5	1.55	3.81	0.94	2.13	1.37	1.68	1.59	1.53	1.41	1.16	1.28	1.14	1.15
		0.9	2.00	4.84	1.23	2.72	1.94	2.08	2.07	2.05	1.89	1.57	1.75	1.53	1.54
$T = 200$	$N(0,1)$	0.1	1.18	4.28	1.01	2.38	1.21	1.29	1.34	1.27	1.14	1.03	1.14	1.00	0.95
		0.5	1.06	3.62	0.87	2.00	1.00	1.14	1.06	1.06	0.92	0.81	0.95	0.75	0.76
		0.9	1.24	4.28	1.00	2.38	1.18	1.27	1.26	1.25	1.06	0.97	1.09	0.89	0.88
	$t(5)$	0.1	1.38	4.15	0.92	2.19	1.43	1.51	1.51	1.57	1.42	1.21	1.29	1.15	1.17
		0.5	1.07	3.09	0.64	1.55	0.98	1.14	1.10	1.06	0.97	0.80	0.93	0.80	0.81
		0.9	1.33	4.07	0.88	2.18	1.37	1.48	1.57	1.42	1.30	1.09	1.28	1.02	1.10
$T = 500$	$N(0,1)$	0.1	1.01	3.85	0.82	2.10	0.97	1.03	1.06	1.11	0.97	0.82	1.00	0.84	0.79
		0.5	0.89	3.24	0.70	1.76	0.83	0.93	0.86	0.83	0.74	0.63	0.77	0.62	0.68
		0.9	1.00	3.64	0.81	2.05	0.99	1.02	1.01	0.98	0.89	0.80	0.89	0.74	0.77
	$t(5)$	0.1	1.14	3.50	0.72	1.81	1.12	1.22	1.27	1.28	1.13	0.98	1.03	0.94	0.93
		0.5	0.91	2.56	0.56	1.29	0.82	0.94	0.90	0.93	0.78	0.67	0.77	0.61	0.63
		0.9	1.13	3.22	0.71	1.76	1.18	1.17	1.23	1.16	1.05	0.90	0.99	0.89	0.87
								$N = 200$							
$T = 100$	$N(0,1)$	0.1	1.13	4.11	0.98	2.28	1.19	1.33	1.29	1.26	1.10	1.05	1.17	1.00	0.96
		0.5	1.00	3.39	0.84	1.92	0.91	1.11	1.00	1.03	0.92	0.82	0.91	0.79	0.78
		0.9	1.14	3.88	0.96	2.27	1.11	1.22	1.18	1.20	1.07	0.95	1.08	0.91	0.93
	$t(5)$	0.1	1.32	3.89	0.88	1.98	1.39	1.52	1.51	1.49	1.31	1.19	1.36	1.21	1.21
		0.5	1.02	2.81	0.65	1.44	1.00	1.10	1.03	1.08	0.94	0.83	0.92	0.83	0.82
		0.9	1.34	3.64	0.89	2.00	1.30	1.48	1.41	1.40	1.27	1.12	1.28	1.14	1.07
$T = 200$	$N(0,1)$	0.1	0.81	3.39	0.76	1.80	0.88	0.92	0.91	0.86	0.77	0.72	0.85	0.70	0.70
		0.5	0.71	2.68	0.64	1.52	0.66	0.76	0.64	0.67	0.64	0.59	0.70	0.53	0.52
		0.9	0.85	3.30	0.68	1.74	0.90	0.88	0.86	0.81	0.77	0.71	0.83	0.65	0.62
	$t(5)$	0.1	0.93	2.95	0.63	1.53	1.00	1.08	1.05	1.04	0.97	0.90	0.96	0.85	0.81
		0.5	0.72	2.17	0.47	1.08	0.64	0.83	0.74	0.73	0.67	0.60	0.63	0.55	0.55
		0.9	0.92	2.92	0.65	1.59	0.97	1.06	0.99	1.03	0.95	0.77	0.90	0.79	0.76
$T = 500$	$N(0,1)$	0.1	0.62	2.44	0.56	1.42	0.72	0.80	0.78	0.70	0.66	0.58	0.61	0.55	0.55
		0.5	0.60	1.97	0.50	1.17	0.49	0.69	0.58	0.56	0.52	0.47	0.56	0.43	0.40
		0.9	0.64	2.29	0.56	1.35	0.72	0.74	0.68	0.69	0.66	0.56	0.60	0.50	0.55
	$t(5)$	0.1	0.71	2.35	0.57	1.29	0.78	0.84	0.85	0.86	0.82	0.69	0.72	0.69	0.69
		0.5	0.61	1.49	0.38	0.79	0.55	0.65	0.63	0.62	0.51	0.47	0.49	0.43	0.45
		0.9	0.77	2.22	0.50	1.16	0.81	0.87	0.94	0.82	0.74	0.61	0.71	0.58	0.57
								$N = 500$							
$T = 100$	$N(0,1)$	0.1	0.92	3.59	0.82	2.01	0.95	1.08	1.07	0.99	0.90	0.85	0.97	0.81	0.81
		0.5	0.82	2.85	0.66	1.57	0.75	0.83	0.83	0.84	0.73	0.66	0.78	0.61	0.62
		0.9	0.94	3.35	0.82	1.84	0.97	1.02	1.02	1.00	0.88	0.76	0.86	0.75	0.74
	$t(5)$	0.1	1.08	3.33	0.75	1.69	1.06	1.25	1.23	1.22	1.06	0.99	1.10	0.94	0.91
		0.5	0.85	2.35	0.55	1.21	0.73	0.93	0.81	0.85	0.79	0.67	0.74	0.67	0.66
		0.9	1.05	3.17	0.72	1.69	1.07	1.11	1.15	1.17	0.99	0.89	1.01	0.90	0.93
$T = 200$	$N(0,1)$	0.1	0.70	2.70	0.60	1.49	0.72	0.72	0.68	0.66	0.65	0.61	0.68	0.57	0.56
		0.5	0.60	2.24	0.49	1.19	0.58	0.65	0.61	0.64	0.53	0.46	0.60	0.46	0.46
		0.9	0.63	2.69	0.62	1.43	0.65	0.74	0.71	0.66	0.63	0.57	0.64	0.49	0.51
	$t(5)$	0.1	0.70	2.42	0.51	1.26	0.86	0.89	0.95	0.89	0.78	0.70	0.78	0.65	0.67
		0.5	0.61	1.65	0.39	0.84	0.53	0.62	0.61	0.59	0.54	0.46	0.56	0.48	0.45
		0.9	0.81	2.29	0.52	1.26	0.73	0.87	0.82	0.83	0.72	0.69	0.76	0.63	0.66
$T = 500$	$N(0,1)$	0.1	0.51	2.33	0.52	1.25	0.59	0.57	0.58	0.64	0.60	0.45	0.56	0.46	0.41
		0.5	0.45	1.63	0.43	0.91	0.40	0.65	0.47	0.45	0.47	0.40	0.42	0.31	0.34
		0.9	0.61	2.10	0.49	1.11	0.61	0.56	0.56	0.57	0.50	0.37	0.53	0.41	0.45
	$t(5)$	0.1	0.56	2.10	0.45	1.11	0.70	0.59	0.71	0.82	0.66	0.60	0.68	0.63	0.51
		0.5	0.52	1.30	0.35	0.67	0.46	0.60	0.49	0.43	0.43	0.39	0.40	0.36	0.35
		0.9	0.64	2.10	0.46	1.10	0.72	0.77	0.64	0.62	0.56	0.47	0.59	0.48	0.58

Notes: The simulation results are based on the DGP in Section 4.1 with 1000 replications and reported across the three different quantiles, $\tau = (0.1, 0.5, 0.9)$ for the sample pairs, $(N, T) = 100, 200, 500$, where we generate $\mathfrak{u}_{it}$ from either a standard normal distribution, $N(0,1)$ or a t-distribution with 5 degrees of freedom, $t(5)$.

Table 2: Coverage Probability (×100) for Type 1 Network

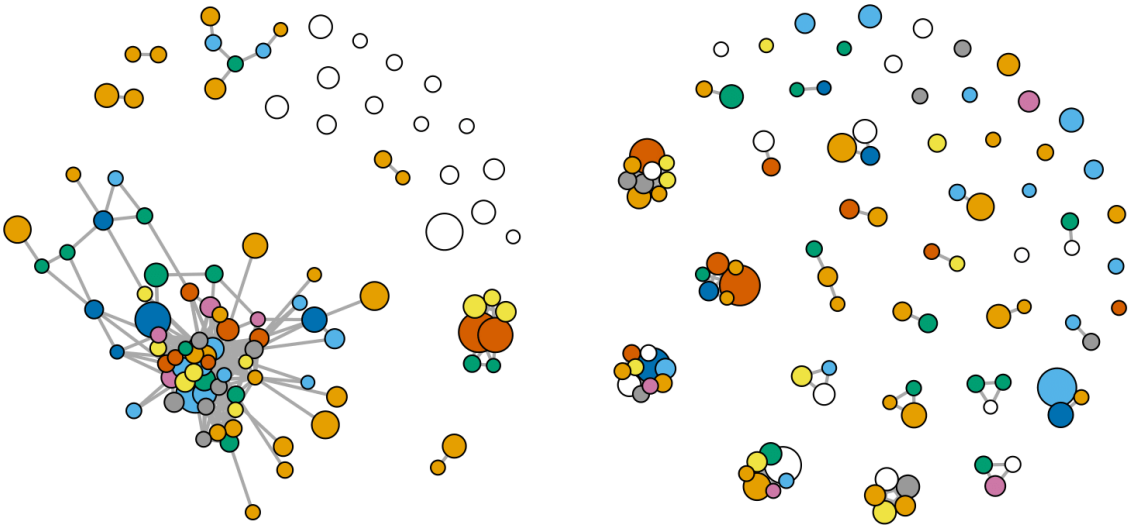
	Dist.	τ	γ_0	γ_1	γ_2	γ_3	α_1	α_2	α_3	α_4	α_5	β_1	β_2	β_3	β_4
								$N = 100$							
$T = 100$	$N(0,1)$	0.1	93.5	97.8	92.9	97.3	93.1	93.2	94.9	94.6	94.8	94.7	95.3	94.5	94.7
		0.5	93.0	97.2	93.5	95.3	94.2	93.0	95.4	95.0	94.8	94.2	96.1	94.7	94.4
		0.9	92.8	97.3	94.0	94.4	93.8	92.1	94.8	95.1	94.6	95.6	94.6	94.8	94.6
	$t(5)$	0.1	93.3	98.6	92.7	95.0	94.5	94.6	94.0	94.5	95.0	95.2	94.4	94.8	94.4
		0.5	92.1	96.8	93.0	96.7	95.0	94.3	94.3	95.1	95.6	96.2	95.0	96.2	95.8
		0.9	93.0	97.2	94.2	95.1	93.9	94.9	94.5	94.8	95.8	94.5	94.2	95.5	94.8
$T = 200$	$N(0,1)$	0.1	93.4	96.2	95.6	95.4	95.6	92.8	94.6	97.0	94.0	95.0	95.8	94.4	94.4
		0.5	95.3	93.6	93.9	94.7	96.4	94.8	96.5	94.6	94.6	94.5	94.0	96.0	95.3
		0.9	93.1	92.9	93.9	94.9	93.7	94.8	94.6	94.8	94.9	93.9	95.4	94.6	95.8
	$t(5)$	0.1	94.8	98.8	93.8	97.0	96.3	94.3	96.3	96.5	96.5	96.8	96.5	97.3	96.5
		0.5	94.2	92.8	94.2	91.4	95.4	94.8	95.8	95.2	95.0	96.0	95.0	94.8	96.8
		0.9	93.8	98.3	92.8	96.5	93.5	96.3	95.8	96.8	96.0	96.0	95.8	93.0	96.5
$T = 500$	$N(0,1)$	0.1	92.4	95.9	92.7	94.8	94.1	96.4	94.8	94.3	94.4	96.4	94.1	93.1	94.9
		0.5	90.9	92.6	91.5	92.9	93.5	93.5	95.7	95.7	96.0	96.9	95.2	96.7	94.7
		0.9	90.9	95.3	92.9	92.7	92.9	94.7	95.1	95.5	95.2	94.1	95.5	94.1	95.9
	$t(5)$	0.1	93.9	93.7	94.7	94.1	95.2	95.3	95.1	94.8	95.7	94.8	95.5	94.0	95.7
		0.5	92.3	93.2	92.7	93.3	96.1	94.0	95.6	94.1	95.6	96.0	95.6	97.1	96.8
		0.9	91.5	92.9	93.7	92.7	93.3	95.5	94.0	94.8	94.0	95.7	96.5	95.6	95.6
								$N = 200$							
$T = 100$	$N(0,1)$	0.1	93.3	96.5	95.5	95.8	95.5	94.3	96.0	96.8	95.5	94.3	94.8	95.0	94.8
		0.5	91.0	94.0	93.3	97.0	96.5	92.0	94.5	95.5	94.5	95.8	94.5	96.0	95.3
		0.9	92.5	95.5	91.5	94.8	94.5	93.8	95.3	93.5	96.3	95.0	94.0	93.3	94.0
	$t(5)$	0.1	94.3	97.6	92.9	97.2	93.9	94.1	94.7	95.4	94.9	94.8	94.1	94.1	93.7
		0.5	90.8	93.7	91.5	94.2	94.2	93.2	94.5	94.4	94.9	95.1	94.7	94.7	94.9
		0.9	91.1	97.3	91.9	94.5	94.8	93.5	94.3	94.6	94.5	94.7	94.1	93.7	95.3
$T = 200$	$N(0,1)$	0.1	94.0	93.5	91.0	92.0	94.5	93.5	95.5	93.5	94.5	93.0	95.5	96.0	94.0
		0.5	90.4	93.2	94.3	92.4	96.2	95.2	97.0	95.6	95.8	95.4	93.6	95.8	96.8
		0.9	91.3	94.2	92.4	93.4	90.4	93.6	94.2	96.4	94.4	93.0	93.8	95.2	96.0
	$t(5)$	0.1	94.6	93.0	93.4	94.0	94.2	95.4	95.6	95.4	94.2	93.6	92.4	95.6	95.4
		0.5	94.2	92.2	91.4	93.8	97.0	92.4	94.8	96.0	95.8	95.4	96.0	96.0	96.4
		0.9	93.5	95.0	95.5	94.5	93.5	96.0	95.0	97.5	94.5	95.5	96.0	96.0	95.0
$T = 500$	$N(0,1)$	0.1	95.2	92.8	92.8	93.6	93.2	94.4	93.6	95.6	94.0	96.4	96.4	95.6	95.6
		0.5	93.7	92.7	96.0	94.3	95.0	94.4	95.3	94.7	93.3	94.7	94.5	95.2	95.3
		0.9	92.7	93.1	94.6	93.3	94.6	94.0	94.7	96.0	94.7	95.3	94.7	94.7	94.9
	$t(5)$	0.1	94.2	95.3	96.0	94.6	95.0	94.7	94.6	94.9	95.3	95.5	94.3	95.2	95.6
		0.5	94.3	94.4	92.0	95.6	94.0	94.8	94.8	95.6	96.4	94.4	95.2	96.8	97.2
		0.9	92.6	95.3	94.6	94.2	95.7	95.1	95.3	94.6	95.6	96.3	94.00	94.3	96.0
								$N = 500$							
$T = 100$	$N(0,1)$	0.1	93.2	96.4	94.0	94.2	95.6	93.6	94.8	95.6	96.0	96.0	93.8	94.6	95.4
		0.5	94.3	94.6	93.8	94.0	95.6	95.0	95.0	95.4	95.2	97.0	96.2	96.6	96.2
		0.9	92.5	93.5	92.0	93.5	95.0	94.0	95.0	95.3	95.5	93.7	96.5	96.3	94.1
	$t(5)$	0.1	93.4	96.4	94.0	96.0	96.8	93.2	95.0	95.4	95.8	94.4	95.0	96.0	96.6
		0.5	93.0	92.5	93.0	93.5	95.0	90.0	96.0	94.0	95.0	95.5	93.5	95.5	95.5
		0.9	93.2	94.0	92.4	94.0	95.8	96.8	95.8	95.6	95.4	97.0	94.4	96.2	94.4
$T = 200$	$N(0,1)$	0.1	92.6	92.4	93.3	93.3	92.8	94.3	96.7	95.2	94.8	94.3	93.3	94.3	95.7
		0.5	91.7	92.5	93.0	92.8	94.5	95.9	94.5	95.5	94.8	95.7	95.3	94.5	95.2
		0.9	91.0	92.5	93.2	93.3	93.3	95.2	93.8	94.3	94.8	93.3	96.2	98.1	97.1
	$t(5)$	0.1	94.3	92.4	93.8	93.3	93.3	94.3	92.4	93.8	93.3	95.2	95.2	94.8	94.8
		0.5	91.6	94.8	96.2	95.7	96.7	94.8	96.2	95.2	96.2	97.1	94.3	97.1	97.1
		0.9	93.5	93.9	94.3	95.4	95.2	95.2	95.2	94.8	95.2	94.3	94.8	94.3	94.3
$T = 500$	$N(0,1)$	0.1	92.7	93.6	93.6	95.5	93.6	91.8	95.5	96.4	94.7	93.6	93.6	95.5	95.6
		0.5	91.8	93.6	93.6	92.7	95.5	95.2	95.5	95.5	94.3	96.4	94.6	94.6	94.6
		0.9	92.9	94.1	94.6	92.3	92.7	90.9	90.9	93.6	95.2	95.5	92.7	94.6	94.8
	$t(5)$	0.1	93.9	93.2	93.6	92.2	92.7	94.6	92.7	99.1	94.6	96.4	93.6	97.3	96.4
		0.5	93.1	92.4	92.7	90.0	95.5	93.1	97.3	96.4	96.4	95.5	92.7	95.5	96.4
		0.9	93.9	95.4	93.8	93.6	94.8	95.1	95.2	95.3	94.3	95.6	94.4	95.7	95.5

Notes: See the notes to Table 1.

Table 3: Estimation Results for the Network W_{CS5}

	DNQR			NQARF			NQAR		
	$\tau = 0.1$	$\tau = 0.5$	$\tau = 0.9$	$\tau = 0.1$	$\tau = 0.5$	$\tau = 0.9$	$\tau = 0.1$	$\tau = 0.5$	$\tau = 0.9$
$\widehat{\gamma}_0$	-2.28^{***}	0.02^{***}	2.33^{***}	-2.55^{***}	0.05^{***}	2.64^{***}	-2.55^{***}	0.04^{***}	2.65^{***}
	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)
$\widehat{\gamma}_1$	54.07^{***}	57.55^{***}	47.09^{***}	-	-	-	-	-	-
	(1.43)	(0.81)	(1.49)						
$\widehat{\gamma}_2$	3.39^{***}	0.55^{***}	2.56^{***}	3.96^{***}	-0.62^{***}	2.67^{***}	4.90^{***}	-0.55^{***}	2.47^{***}
	(0.48)	(0.33)	(0.53)	(0.66)	(0.21)	(0.63)	(0.65)	(0.20)	(0.62)
$\widehat{\gamma}_3$	-1.29^{***}	-2.41^{***}	-2.45^{***}	-0.41	-2.10	-2.17	-0.09	-1.59	-2.19
	(0.35)	(0.27)	(0.34)	(0.42)	(0.13)	(0.39)	(0.41)	(0.13)	(0.39)
SIZE	0.09^{***}	0.00^{***}	-0.09^{***}	0.09^{***}	0.00^{***}	-0.09^{***}	0.10^{***}	0.00^{***}	-0.09^{***}
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
BM	0.13^{***}	0.01^{***}	-0.11^{***}	0.14^{***}	0.02^{***}	-0.12^{***}	0.14^{***}	0.02^{***}	-0.12^{***}
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Cash	0.00	0.00	0.02	0.00	0.01	0.02	0.00	0.01	0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
PE	0.04^{***}	0.00^{***}	-0.02^{***}	0.03^{***}	0.01^{***}	-0.02^{***}	0.03^{***}	0.01^{***}	-0.01^{***}
	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)
VIX	-0.16^{***}	-0.04^{***}	-0.07^{***}	-0.23^{***}	-0.08^{***}	-0.12^{***}	-	-	-
	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)			
Rm - Rf	-0.07^{***}	0.02^{***}	0.10^{***}	-0.14^{***}	0.03^{***}	0.15^{***}	-	-	-
	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)			
SMB	0.02^{*}	0.00^{*}	-0.01^{*}	0.00	0.01	0.04	-	-	-
	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)			
HML	0.03^{**}	0.00^{**}	-0.04^{**}	0.01	-0.01	-0.10	-	-	-
	(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)			
Goodn.fit.	-	-	-	8.68	9.45	6.88	9.39	9.50	7.00

Notes: The dataset consists of $N = 943$ stocks with $T = 252$ time periods. The network matrix, W_{CS5} is constructed by checking if the stocks are invested in by at least five common shareholders with the network density, 3.24%. The estimates ($\times 10^2$) are reported across different quantiles $\tau = 0.1, 0.5, 0.9$, and the value in parentheses is the standard error ($\times 10^2$). DNQR denotes the proposed model, NQAR is the model without contemporaneous network effects and common factors, and NQARF is the factor-augmented NQAR model. Goodn.fit. ($\times 10^2$) represents the goodness of fit of DNQR model with respect to the other models. The 1%, 5% and 10% significance levels are denoted by ***, **, *, respectively.



(a) The Topology of the Network W_{CS5} (b) The Topology of the Network W_{HQ}

Figure 1: We depict the top 100 market value stocks out of 943 firms selected. (a): the common shareholder network W_{CS5}, constructed by checking if the stocks are invested in by at least five common shareholders. (b): the uniform headquarter location network W_{HQ}, constructed by checking if the headquarters of companies are located in the same city. The larger nodes imply higher market capitalization while the darker nodes present higher connectedness.

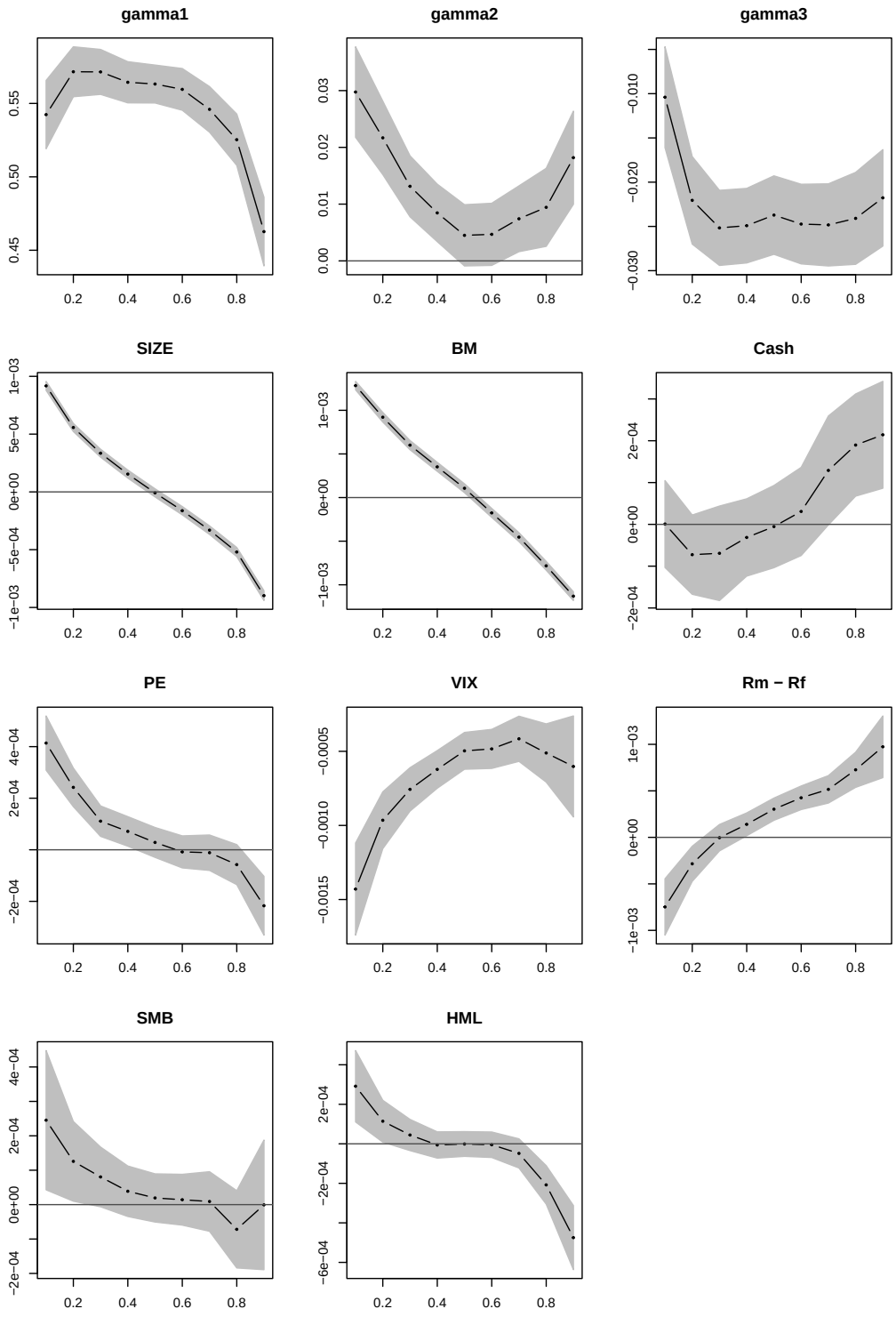


Figure 2: Quantile-specific Coefficients for the Network W_{CS5}

Notes: The dashed line is the QR coefficient while the grey area indicates a kernel density based confidence band advanced by Powell (1991). They are displayed across quantiles, $\tau = 0.1, 0.2, \cdots, 0.9$.