Federal Reserve Bank of Boston®

Supervisory Research and Analysis Unit

Working Paper | SRA 21-02 | August 10, 2021

A Spatial Panel Quantile Model with Unobserved Heterogeneity

Tomohiro Ando, Kunpeng Li and Lina Lu

Supervisory Research and Analysis (SRA) Working Papers present economic, financial and policy-related research conducted by staff in the Federal Reserve Bank of Boston's Supervisory Research and Analysis Unit. SRA Working Papers can be downloaded without charge at: http://www.bostonfed.org/publications/sra/

A Spatial Panel Quantile Model with Unobserved Heterogeneity*

Tomohiro Ando[†], Kunpeng Li[‡], Lina Lu[§¶]

August 6, 2021

Abstract

This paper introduces a spatial panel quantile model with unobserved heterogeneity. The proposed model is capable of capturing high-dimensional cross-sectional dependence and allows heterogeneous regression coefficients. For estimating model parameters, a new estimation procedure is proposed. When both the time and cross-sectional dimensions of the panel go to infinity, the uniform consistency and the asymptotic normality of the estimated parameters are established. In order to determine the dimension of the interactive fixed effects, we propose a new information criterion. It is shown that the criterion asymptotically selects the true dimension. Monte Carlo simulations document the satisfactory performance of the proposed method. Finally, the method is applied to study the quantile co-movement structure of the U.S. stock market by taking into account the input-output linkages as firms are connected through the input-output production network.

Keywords: Endogeneity; Heterogeneous panel; Quantile factor structure; Serial and cross-sectional correlations; Spatial effects

JEL classification: C31, C33, E44

*We would like to thank Co-editor, Associate editor and two anonymous referees for their insightful comments that greatly improved our paper. We also thank Jushan Bai for his constructive comments on an early version of this paper. We also would like to thank Mark Manion and Xiyu Wang for their outstanding research assistance.

†Melbourne Business School, Melbourne University, 200 Leicester Street, Carlton, Victoria 3053, Australia. Email: T.Ando@mbs.edu

‡International School of Economics and Management, Capital University of Economics and Business, Fengtai district, Beijing, 100070, China. Email: likp.07@gmail.com

§Federal Reserve Bank of Boston, the Supervisory Research and Analysis Unit, 600 Atlantic Avenue, Boston, MA 02210, USA. Email: lina.lu@bos.frb.org

¶The views expressed in the paper are those of the authors and should not be interpreted as reflecting the views of the Federal Reserve Bank of Boston, or any other person associated with the Federal Reserve System.

1 Introduction

The goal of this paper is to develop a new statistical method for analyzing the quantile co-movement of a large number of time series with explicit allowance of spatial interactions among them, and to empirically study the quantile co-movement structure of the U.S. stock market by taking into account the linkages among firms through the input-output production network.

Studies of searching for factors to explain the co-movement of expected returns in stock markets can date back at least to Fama and French (1993), and have experienced rapid growth in the last three decades, see Griffin (2002), Hou et al. (2011), and so forth. The ability of quantile models to capture the heterogeneous impact of explanatory variable on different distribution points of the outcome make them appealing for stock return studies. Recently, Ando and Bai (2020) further introduce the interactive fixed effects to the panel quantile model to deal with the unobserved endogenous heterogeneity within stock data. In this paper, we emphasize the linkages through firms' input-output production network in the study of the co-movement of expected returns in stock markets.

The production network provides an important transmission channel for the propagation of various economic shocks in the economy, such as idiosyncratic shocks (Gabaix, 2011; Acemoglu et al., 2012), macroeconomic shocks (Ozdagli and Weber, 2020; Pasten et al., 2020), among others. Ozdagli and Weber (2020) emphasize the input-output production linkages across firms in the study of the impacts of monetary policy on the stock market. Ahern and Harford (2014) find that stronger production networks lead to a greater incidence of cross-industry mergers. Acemoglu et al. (2016) and Barrot and Sauvagnat (2016) show that the production network is important for the transmission of federal spending, trade, technology and knowledge shocks. Luo (2016) investigates how these production linkages, together with financial linkages, lead to the propagation of financial shocks.

In the study of expected returns in stock markets, financial crises or monetary policy changes can have substantial impacts on the stock market, and their impact on one particular industry, through the input-output production network, can further affect that industry's suppliers by reducing its demand for the goods and services of its suppliers. This assertion is supported by the observed fact that the U.S. subprime mortgage crisis of 2007 led to the automotive industry crisis during 2008-2010 and further affected the suppliers of the automotive industry, such as the rubber and oil industries.[1] Monetary policies can also have a substantial influence on asset prices, see Bernanke and Kuttner (2005), Nakamura and Steinsson (2018), Gorodnichenko and Weber (2016), among others, and such impacts can be further decomposed into direct effects and network effects, which originate from the production network (see Ozdagli and Weber (2020)). Therefore, from the perspective of governmental policy makers, regulators and asset managers, it is important and worthwhile to investigate the production linkages and the quantile co-movement of the stock market together.

By studying the quantile co-movement of a large number of stock returns in the U.S. market and, at the same time, taking into account the production linkages (or the spillover effect), this paper aims to answer the following empirical questions:

1. What is the strength of the spillover effect through the production network? Does it change over time?

2. How does monetary policy (specifically, interest rates) affect the stock market, directly and indirectly, through the production network? How much of its total effect is due to direct effects and how much is due to network effects?

3. In the presence of the production network, do the quantile common-factor structures that explain the asset return distribution vary across quantiles? Are they symmetric across quantiles (in other words, are they identical in the lower and upper tails)?

4. Is there any special economic meaning for the unobservable common factors?

5. Are the co-movements of quantiles captured by the stocks' sector-level classification characteristic?

Recently, Ando and Bai (2020) use a panel quantile model with interactive effects to study the financial stocks in the U.S. market. They find that the common factor structures in the

[1] For example, as mentioned in Luo (2016), during the automotive crisis period, General Motors Co. significantly reduced its demand owing to a severe liquidity problem and consequently, American Axle & Manufacturing Holdings Inc., one of the major suppliers of GM, experienced a net loss of $112.1 million in the fourth quarter of 2008.

tails and those at the mean are not always identical. Moreover, this unobservable structure varies over quantiles. However, their approach fails to account for the network effects, a point as emphasized before, that should be carefully addressed.

In this paper, we propose a spatial panel quantile model with interactive effects to address the above five important issues. In our model, we use a spatial term to capture the network effects, and use the factor structure to deal with the unobservable heterogeneity in the data. Our model is related to two popular models, the spatial panel data model and the panel model with unobserved factor structure. Previous studies on spatial panel data models include Kelejian and Prucha (2004), Lee (2004), Baltagi (2011), Reich et al. (2011), Bai and Li (2013, 2021), etc. Previous studies on large-dimensional linear panel models with unobserved factor structures include Ando and Bai (2017), Bai (2009), Bai and Li (2012), Bai and Liao (2016), Bai and Ng (2002, 2013), Hallin and Liška (2007), Moon and Weidner (2015), Pesaran (2006), Stock and Watson (2002), Lu and Su (2016), among others. The existing studies focus primarily on the panel mean model, studies on panel quantile models with unobserved factor structures are scant. Ando and Bai (2020) and Harding et al. (2020) study quantile panel models with unobserved factor structures.[2] But spatial effects are absent in these two studies.

This paper makes a thorough investigation on the estimation and inferential theory of the proposed model. The parameter estimation is a challenging issue and we develop a new parameter estimation procedure. We establish the asymptotic theory of the estimators, including consistency, convergence rates and limiting distributions. To show the asymptotic properties, we will encounter several theoretical challenges, such as large dimensional incidental parameters due to the loadings and factors, the nonsmooth objective function for the quantile regression, the nonlinearity arising from the spatial term, the heterogeneous regression coefficients, as well as separately identifying the factor loadings and the factors because of the rotational indeterminacy. Partial theoretical challenges have been emphasized by some recent studies, such as Ando and Bai (2020), Chen et al. (2021) and He et al. (2020). Such studies provide some useful tools to analyze the current model but we note that these tools are far from enough to support our theoretical analysis. More new arguments are developed

[2] For reference, Chernozhukov and Hansen (2006) proposes an instrumental quantile regression method for structural and treatment effect models.

in this paper.

Our contributions are summarized as follows. First, a spatial heterogeneous panel quantile model with interactive effects is introduced. Second, a new parameter estimation procedure and a new model selection criterion are proposed. Third, some new arguments are developed and the asymptotic properties of our estimator are established. Furthermore, we show that the proposed model selection criterion is able to consistently estimate the true dimension of interactive effects that may vary across quantiles. Finally, we apply the proposed model and the estimation method to study the U.S. stock market. We find that the number of common factors and the strength of the production linkage vary across quantiles. Furthermore, the number of common factors is smaller at the lower tail than at the upper tail, indicating that when the market is in stress, there are fewer dominant factors driving the co-movement of the market. Regarding the production linkage, it is stronger at the tails than at the median. These findings are new, and are helpful to understand the U.S. stock market.

The paper is organized as follows. Section 2 introduces a new spatial panel quantile model with interactive fixed effects. In Section 3, we introduce the estimation and model selection method. Section 4 presents a set of assumptions and Section 5 investigates some asymptotic properties of the proposed method. Section 6 contains Monte Carlo simulation results which indicate that the proposed estimation procedure works well. In Section 7, the proposed method is applied to the U.S. stock market data. Section 8 provides our concluding remarks. To save space, all technical proofs are provided in the online supplementary document.

Notations Let $\|A\| = [\mathrm{tr}(A'A)]^{1/2}$ be the usual norm of the matrix A, where "tr" denotes the trace of a square matrix. In addition, for any $N \times N$ matrix $\|A\|_1$ is defined as $\|A\|_1 = \max_{1\le j\le N} \sum_{i=1}^{N} |a_{ij}|$ where a_{ij} is the (i,j)-th element of A. Similarly, $\|A\|_\infty = \max_{1\le i\le N} \sum_{j=1}^{N} |a_{ij}|$. The equation $a_n = O(b_n)$ states that the deterministic sequence a_n is at most of order b_n; $c_n = O_p(d_n)$ states that the random variable c_n is at most of order d_n in terms of probability and $c_n = o_p(d_n)$ is of a smaller order in terms of probability.

2 The model

Suppose that, for the i-th unit $(i = 1, ..., N)$ at time t $(t = 1, ..., T)$, its response y_{it} is observed together with a set of p explanatory variables $\{x_{it,1}, ..., x_{it,p}\}$. Our objective is to estimate the τ-th quantile function of y_{it} by jointly modeling spatial interactions and common shocks. More specifically, we define the τ-th quantile function of y_{it} as

$$
\begin{aligned}
Q_{y_{it}}\left(\tau|\boldsymbol{Q}_{(-i)t}(\tau), \boldsymbol{x}_{it}, \rho_{i,\tau}, \boldsymbol{b}_{i,\tau}, \boldsymbol{f}_{t,\tau}, \boldsymbol{\lambda}_{i,\tau}\right) &\equiv \rho_{i,\tau}\sum_{j=1}^{N} w_{ij} Q_{y_{jt}}\left(\tau|\boldsymbol{Q}_{(-j)t}(\tau), \boldsymbol{x}_{jt}, \rho_{j,\tau}, \boldsymbol{b}_{j,\tau}, \boldsymbol{f}_{t,\tau}, \boldsymbol{\lambda}_{j,\tau}\right) \\
&\quad + \sum_{k=1}^{p} x_{it,k} b_{ik,\tau} + \sum_{k=1}^{r} f_{tk,\tau}\lambda_{ik,\tau} + G^{-1}_{i,e_{it}}(\tau) \\
&\equiv \rho_{i,\tau}\sum_{j=1}^{N} w_{ij} Q_{y_{jt}}\left(\tau|\boldsymbol{Q}_{(-j)t}(\tau), \boldsymbol{x}_{jt}, \rho_{j,\tau}, \boldsymbol{b}_{j,\tau}, \boldsymbol{f}_{t,\tau}, \boldsymbol{\lambda}_{j,\tau}\right) \\
&\quad + \boldsymbol{x}'_{it}\boldsymbol{b}_{i,\tau} + \boldsymbol{f}'_{t,\tau}\boldsymbol{\lambda}_{i,\tau} \qquad (1)
\end{aligned}
$$

for $i = 1, \ldots, N$ and $t = 1, \ldots, T$. Here w_{ij} $(i = 1, 2, \cdots, N; j = 1, 2, \cdots, N)$ are pre-specified spatial weights; $\boldsymbol{Q}_{(-i)t}(\tau) \equiv \{Q_{y_{jt}}(\tau|\boldsymbol{Q}_{(-j),t}(\tau), \boldsymbol{x}_{jt}, \rho_{j,\tau}, \boldsymbol{b}_{j,\tau}, \boldsymbol{f}_{t,\tau}, \boldsymbol{\lambda}_{j,\tau}); j \neq i\}$ is a collection of quantile functions; $\rho_{i,\tau}$ is the spatial parameter capturing the strength of the spillover effect, which may depend on the quantile τ; $\boldsymbol{x}_{it} = (1, x_{it,1}, ..., x_{it,p})'$ is $(p+1)$ -dimensional vector of explanatory variables; $\boldsymbol{b}_{i,\tau} = (b_{i,0,\tau}, b_{i,1,\tau}, ..., b_{i,p,\tau})'$ is a $(p+1)$-dimensional vector of regression coefficients; e_{it} is the error term and $G^{-1}_{i,e_{it}}(\tau)$ is the τ-th quantile point of e_{it} with $G_{i,e_{it}}(\cdot)$ being the cumulative distribution function of e_{it}. Our definition implicitly assumes that e_{it} is identically distributed over t while its distribution may vary over i. Notice that the τ-th quantile of the idiosyncratic error $G^{-1}_{i,e_{it}}(\tau)$, which depends only on i and τ, is absorbed by the term $\boldsymbol{x}'_{it}\boldsymbol{b}_{i,\tau}$ since the first element of $\boldsymbol{x}_{it}$ is 1.

In this paper, we allow the regression coefficients $(\rho_{i,\tau}, \boldsymbol{b}'_{i,\tau})'$ to be heterogeneous across i. The econometric literature has paid much attention to the heterogeneous coefficients in interactive-effect models (see, e.g., Pesaran (2006), Ando and Bai (2020), Li, Cui and Lu (2020)), and recently witnesses growing interests on heterogeneous coefficients in spatial models (see, e.g., Aquaro et al. (2021), Koopman et al. (2021), Zhu et al. (2020)). So far, the existing studies focus mostly on the mean regression and the current paper complements the literature with the quantile regression. An advantage of heterogeneous specification is that the model is more flexible to fit real data and more robust to misspecification. $\boldsymbol{f}_{t,\tau}$ is an r_τ -dimensional vector of common factors and $\boldsymbol{\lambda}_{i,\tau}$ is the corresponding r_τ -dimensional vector of unobservable heterogeneous responses to the common factors, termed the factor loadings. In the literature, the part $\boldsymbol{f}'_{t,\tau}\boldsymbol{\lambda}_{i,\tau}$ is also known as the interactive fixed effects, after Bai (2009). In our panel quantile model, we allow that the loadings, the factors and the dimensionality of interactive effects all depend on the quantile τ. By assuming that the conditional quantile function in (1) is monotone function of τ, we have the following equivalent expression:

$$\begin{aligned} y_{it,u_{it}} = \rho_{i,u_{it}} \sum_{j=1}^{N} w_{ij} Q_{y_{jt}}(u_{it}|\boldsymbol{Q}_{(-j)t}(u_{it}), \boldsymbol{x}_{jt}, \rho_{j,u_{it}}, \boldsymbol{b}_{j,u_{it}}, \boldsymbol{f}_{t,u_{it}}, \boldsymbol{\lambda}_{j,u_{it}}) \\ + \sum_{k=1}^{p} x_{it,k} b_{ik,u_{it}} + \sum_{l=1}^{r} f_{tl,u_{it}} \lambda_{il,u_{it}} + e_{it,u_{it}}, \end{aligned} \tag{2}$$

for $i = 1, \ldots, N$ and $t = 1, \ldots, T$. Here u_{it} are independent and identically distributed following the standard uniform distribution $U(0,1)$. Throughout the paper, we treat the true unobservable $\boldsymbol{f}_{t,\tau}$ and $\boldsymbol{\lambda}_{i,\tau}$ as fixed parameters. To address the five empirical questions proposed in the Introduction section, we integrate the spatial term, heterogeneous coefficients, and factor structure into model (1). In the input-output production application in Section 7, we see that the proposed model fits the real data well in the sense that the estimated parameters vary over a wide range across quantiles and individuals.

Throughout the paper, we treat the weights matrix $W = [w_{ij}]_{N\times N}$ as exogenous. This treatment is standard in spatial econometrics, see Kelejian and Prucha (2004), Lee (2004), Yu et al. (2008) and so forth. An important recent development is to allow the spatial weights matrix to be endogenous, see Qu and Lee (2015). However, we do not pursue this interesting extension and leave it as the future work. In empirical applications, the spatial weights matrix can be constructed by geographic distance, economic distance or social distance, depending on various research interests. We will discuss the choices of the weights matrix in Remark 1.

It is well documented in the factor literature that the factors and the factor loadings can only be identified up to a rotation. To eliminate this rotational indeterminacy, we need to impose normalization restrictions to achieve a full identification. In this paper, we follow Bai and Li (2013) to impose the following restrictions on $F_\tau = (\boldsymbol{f}_{1,\tau}, ..., \boldsymbol{f}_{T,\tau})'$ and $\Lambda_\tau = (\boldsymbol{\lambda}_{1,\tau}, ..., \boldsymbol{\lambda}_{N,\tau})'$:

$$\textbf{Normalization Conditions}: \quad \frac{1}{T}F'_\tau F_\tau = I_{r_\tau} \quad \text{and} \quad \frac{1}{N}\Lambda'_\tau \Lambda_\tau = D, \tag{3}$$

where I_{r_τ} is an $r_\tau \times r_\tau$ identity matrix, and D is a diagonal matrix whose diagonal elements are distinct and are arranged in a descending order. There are alternative choices for full identification, and readers are referred to Bai and Li (2012), Bai and Li (2013) and others, for other identification strategies.

Similar to Belloni and Chernozhukov (2011), Tang et al. (2013), Sherwood and Wang (2016), Ando and Bai (2020) and Chen et al. (2021), this paper focuses on the quantile function under a particular τ value instead of the entire quantile function over some interval. As emphasized by one referee, the monotonicity of the quantile function should be guaranteed. Suppose that the $N \times N$ matrix $P(\boldsymbol{\rho}_\tau) \equiv (I - \boldsymbol{\rho}_\tau W)^{-1}$ exists, where $\boldsymbol{\rho}_\tau = \mathrm{diag}(\rho_{1,\tau}, \ldots, \rho_{N,\tau})$, and $W = [w_{ij}]$ is the $N \times N$ spatial weights matrix. We denote the (i,j)th element of $P(\boldsymbol{\rho}_\tau)$ as $p_{ij}(\boldsymbol{\rho}_\tau)$. Then, model (1) can be rewritten as

$$\begin{bmatrix} Q_{y_{1t}}(\tau | X_t, \boldsymbol{\rho}_\tau, B_\tau, \boldsymbol{f}_{t,\tau}, \Lambda_\tau) \\ \vdots \\ Q_{y_{Nt}}(\tau | X_t, \boldsymbol{\rho}_\tau, B_\tau, \boldsymbol{f}_{t,\tau}, \Lambda_\tau) \end{bmatrix} \equiv \begin{bmatrix} \sum_{k=1}^N p_{1k}(\boldsymbol{\rho}_\tau)(\boldsymbol{x}_{kt}'\boldsymbol{b}_{k,\tau} + \boldsymbol{f}_{t,\tau}'\boldsymbol{\lambda}_{k,\tau}) \\ \vdots \\ \sum_{k=1}^N p_{Nk}(\boldsymbol{\rho}_\tau)(\boldsymbol{x}_{kt}'\boldsymbol{b}_{k,\tau} + \boldsymbol{f}_{t,\tau}'\boldsymbol{\lambda}_{k,\tau}) \end{bmatrix}. \tag{4}$$

or, for each j,

$$Q_{y_{jt}}\Big(\tau | X_t, \boldsymbol{\rho}_\tau, B_\tau, \boldsymbol{f}_{t,\tau}, \Lambda_\tau\Big) = \sum_{k=1}^N p_{jk}(\boldsymbol{\rho}_\tau)\Big(\boldsymbol{x}_{kt}'\boldsymbol{b}_{k,\tau} + \boldsymbol{f}_{t,\tau}'\boldsymbol{\lambda}_{k,\tau}\Big),$$

where $B_\tau = (\boldsymbol{b}_{1,\tau}, \boldsymbol{b}_{2,\tau}, \cdots, \boldsymbol{b}_{N,\tau})'$. Taking the first derivative of $Q_{y_{jt}}(\tau | X_t, \boldsymbol{\rho}_\tau, B_\tau, \boldsymbol{f}_{t,\tau}, \Lambda_\tau)$ in (4), the monotonicity of the quantile function is equivalent to the following first order condition

$$\begin{aligned} &\frac{\partial Q_{y_{jt}}(\tau | X_t, \boldsymbol{\rho}_\tau, B_\tau, \boldsymbol{f}_{t,\tau}, \Lambda_\tau)}{\partial \tau} \\ &\qquad = \sum_{k=1}^N \frac{\partial p_{jk}(\boldsymbol{\rho}_\tau)}{\partial \tau}(\boldsymbol{x}_{kt}'\boldsymbol{b}_{k,\tau} + \boldsymbol{f}_{t,\tau}'\boldsymbol{\lambda}_{k,\tau}) + \sum_{k=1}^N p_{jk}(\boldsymbol{\rho}_\tau)\frac{\partial(\boldsymbol{x}_{kt}'\boldsymbol{b}_{k,\tau} + \boldsymbol{f}_{t,\tau}'\boldsymbol{\lambda}_{k,\tau})}{\partial \tau} > 0 \end{aligned} \tag{5}$$

for $\tau \in (0,1)$, $j = 1, ..., N$ and $t = 1, ..., T$. Although there are some previous studies that focus on the entire quantile function under the monotonicity of the quantile function (for e.g., He (1997), Bondell et al. (2010), Chernozhukov et al. (2010), Dette and Volgushev (2008), Yuan et al. (2017)), these studies are conducted for cross-sectional regression, not for a quantile factor model. Even for a pure quantile factor model, $y_{it,\tau} = \boldsymbol{f}_{t,\tau}'\boldsymbol{\lambda}_{i,\tau} + \varepsilon_{it,\tau}$, it is not trivial how to impose both a monotonicity restriction and an identification condition on the factor structure $\boldsymbol{f}_{t,\tau}'\boldsymbol{\lambda}_{i,\tau}$ simultaneously for all τ in a satisfactory manner. The previous studies on quantile factor models such as Ando and Bai (2020) and Chen et al. (2021) focused on a specific quantile and thus the monotonicity issue is ignored. In the supplement, we provide some discussions on the sufficient conditions that guarantee (5).

Remark 1 We present some typical examples of weight choices that are widely adopted in the literature. In geographic spatial models, Lin and Lee (2010) and Shi and Lee (2017) use a binary weight matrix, where the weight w_{ij} is one if i and j are neighbours, and zero otherwise. Jeanty et al. (2010) consider two choices of weights based on the geographic distance. One defines w_{ij} as a binary distance-based weight, as w_{ij} equals one only if the distance is smaller than a certain distance threshold, and zero otherwise. The other defines the weight w_{ij} as an inverse distance function d_{ij}^{-a}, where d_{ij} measures the distance, and a is a dampening coefficient indicating how fast the weight decreases with distance. Furthermore, Cohen-Cole et al. (2018) and Liu (2014) consider the weights in the multi-choice game framework of a social network model, where w_{ij} is equal to one if individuals i and j are connected; and zero otherwise. In macroeconomics, the input-output production network has been used when studying the transmission of various economic shocks in the economy (see Acemoglu et al. (2016); Ozdagli and Weber (2020)), where w_{ij} is defined as the output amount of industry i used by industry j as input. In international economics, the weight matrix can be constructed according to the trade flows to capture the trade network, see Lu (2017); Lu and Luo (2020).

3 Estimation and model selection

3.1 Estimation

Because of the spatial term, the conditional quantile function of y_{it} depends on the conditional quantile functions of other y_{jt}s. This spatial dependence, together with the existence of large dimensional incidental parameters $\boldsymbol{\lambda}_{i,\tau}$ and $\boldsymbol{f}_{t,\tau}$, poses challenges to estimate the current model. Here we give a detailed description on the approach to obtaining the numerical values.

We estimate the unknown parameters $\boldsymbol{\rho}_\tau$, B_τ, Λ_τ, and F_τ simultaneously by minimizing the following objective function

$$\ell_\tau(Y|X, \boldsymbol{\rho}_\tau, B_\tau, F_\tau, \Lambda_\tau) = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} q_\tau\Big(y_{it} - Q_{y_{it}}(\tau|X_t, \boldsymbol{\rho}_\tau, B_\tau, \boldsymbol{f}_{t,\tau}, \Lambda_\tau)\Big) \tag{6}$$

where $Q_{y_{it}}(\tau|X_t, \boldsymbol{\rho}_\tau, B_\tau, \boldsymbol{f}_{t,\tau}, \Lambda_\tau)$ is defined in (4), $q_\tau(u) = u(\tau - I(u \le 0))$ is the quantile loss function, $Y \equiv \{y_{it}|i = 1, ..., N, t = 1, ..., T\}$ and $X \equiv \{\boldsymbol{x}_{it}|i = 1, ..., N, t = 1, ..., T\}$.

Because the objective function is nonlinear in unknown parameters, we update them sequentially. As in below, the expression (4) will be useful for developing our estimation algorithm. At this moment, we assume that the number of common factors r_τ is known. An information criterion that can consistently estimate the number of common factors is provided in Section 3.2. Here we use $A_{\tau,(-k)}$ to denote A_τ excluding the kth unit, and the symbols with bar to denote the latest values of parameters. Our estimation is an iterative procedure and each iteration consists of the following three steps.

Step One. For given i and the latest values $\bar{B}_\tau$, $\bar{F}_\tau$, $\bar{\Lambda}_\tau$, $\bar{\boldsymbol{\rho}}_{\tau,(-i)}$, we update $\rho_{i,\tau}$ as the minimizer of

$$\ell_\tau(\boldsymbol{\rho}_\tau) = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} q_\tau\Big(y_{it} - \sum_{k=1}^{N} p_{ik}(\check{\boldsymbol{\rho}}_\tau)(\boldsymbol{x}'_{kt}\bar{\boldsymbol{b}}_{k,\tau} + \bar{\boldsymbol{f}}'_{t,\tau}\bar{\boldsymbol{\lambda}}_{k,\tau})\Big). \tag{7}$$

where $\check{\boldsymbol{\rho}}_\tau = \mathrm{diag}(\bar{\rho}_{1,\tau}, \cdots, \bar{\rho}_{i-1,\tau}, \rho_{i,\tau}, \bar{\rho}_{i+1,\tau}, \cdots, \bar{\rho}_{N,\tau})$. We update the diagonal elements of $\boldsymbol{\rho}_\tau$ one by one from 1 to N.

Step two. For notational simplicity, we use $Q_{y_{jt}}(\tau)$ to denote the conditional quantile function of y_{jt}, $Q_{y_{jt}}(\tau|X_t, \boldsymbol{\rho}_\tau, B_\tau, \boldsymbol{f}_{t,\tau}, \Lambda_\tau)$ for $j = 1, ..., N$ and $t = 1, ..., T$. Given i and the latest values $\bar{\boldsymbol{\rho}}_\tau$, $\bar{B}_{\tau,(-i)}$, $\bar{\Lambda}_{\tau,(-i)}$ and $\bar{F}_\tau$, we update $\boldsymbol{b}_{i,\tau}$ and $\boldsymbol{\lambda}_{i,\tau}$ by minimizing

$$\ell_\tau(\boldsymbol{b}_{i,\tau}, \boldsymbol{\lambda}_{i,\tau}) = \frac{1}{T}\sum_{t=1}^{T} q_\tau\Big(y_{it} - \bar{\rho}_{i,\tau}\sum_{j=1}^{N} w_{ij}\bar{Q}_{y_{jt}}(\tau) - \boldsymbol{x}'_{it}\boldsymbol{b}_{i,\tau} - \bar{\boldsymbol{f}}'_{t,\tau}\boldsymbol{\lambda}_{i,\tau}\Big) \tag{8}$$

and we update B_τ and Λ_τ rows by rows. Notice that, in the above formula, $\bar{Q}_{y_{jt}}(\tau)$ is defined as in (4), which depends on $\bar{\boldsymbol{\rho}}_\tau, \bar{B}_{\tau,(-i)}, \bar{\Lambda}_{\tau,(-i)}, \bar{F}_\tau, \boldsymbol{b}_{i,\tau}$ and $\boldsymbol{\lambda}_{i,\tau}$.

Step three. Given t and the latest values $\bar{\boldsymbol{\rho}}_\tau$, $\bar{B}_\tau$ and $\bar{\Lambda}_\tau$, we update $\boldsymbol{f}_{t,\tau}$ as the minimizer of

$$\ell_\tau(\boldsymbol{f}_{t,\tau}) = \frac{1}{N}\sum_{i=1}^{N} q_\tau\Big(y_{it} - \bar{\rho}_{i,\tau}\sum_{j=1}^{N} w_{ij}\bar{Q}_{y_{jt}}(\tau) - \boldsymbol{x}'_{it}\bar{\boldsymbol{b}}_{i,\tau} - \boldsymbol{f}'_{t,\tau}\bar{\boldsymbol{\lambda}}_{i,\tau}\Big) \tag{9}$$

for $t = 1, ..., T$. Again, notice that in the above formula, $\bar{Q}_{y_{jt}}(\tau)$ is defined as in (4), which depends on $\bar{\boldsymbol{\rho}}_\tau, \bar{B}_\tau, \bar{\Lambda}_\tau$ and $\boldsymbol{f}_{t,\tau}$.

Since we update Λ_τ given the current F_τ and update F_τ given the current Λ_τ, the same rotational matrix passes down from the current iteration to the next one. This implies that the rotational indeterminacy issue pointed out in Section 2 can be ignored in our iteration estimation procedure. However, when the tolerance condition is satisfied and the iteration ceases, we need to normalize the estimator to guarantee that the normalization conditions are satisfied. Here are the concrete steps. Let $(\tilde{\Lambda}_\tau, \tilde{F}_\tau)$ be the estimator of the last iteration. First, compute the covariance matrix of $\tilde{M}_{ff} = \frac{1}{T}\tilde{F}_\tau'\tilde{F}_\tau$. Second, compute the matrix $\tilde{M}_{ff}^{1/2}\frac{1}{N}\tilde{\Lambda}_\tau'\tilde{\Lambda}_\tau\tilde{M}_{ff}^{1/2}$ and let $\mathcal{R}$ denote the eigenvector matrix of this matrix whose corresponding values are arranged in the descending order. Third, calculate

$$\hat{\Lambda}_\tau = \tilde{\Lambda}_\tau\tilde{M}_{ff}^{1/2}\mathcal{R}, \qquad \hat{F}_\tau = \tilde{F}_\tau\tilde{M}_{ff}^{-1/2}\mathcal{R}. \tag{10}$$

One can readily check that the estimator $(\hat{\Lambda}_\tau, \hat{F}_\tau)$ satisfies the normalization conditions. Now we summarize the above estimation procedures into the following algorithm.

Algorithm 1 Algorithm for the spatial quantile panel model with unobserved heterogeneity

Step 1. Initialize $\hat{\boldsymbol{\rho}}_\tau$, $\hat{B}_\tau$, $\hat{F}_\tau$ and $\hat{\Lambda}_\tau$.

Step 2. Given $\hat{B}_\tau$, $\hat{F}_\tau$, $\hat{\Lambda}_\tau$, and $\hat{\boldsymbol{\rho}}_{\tau,(-i)}$, update $\rho_{i,\tau}$ by using (7) for $i = 1, 2, \ldots, N$.

Step 3. Given $\hat{\boldsymbol{\rho}}_\tau$, $\hat{B}_{\tau,(-i)}$, $\hat{\Lambda}_{\tau,(-i)}$ and $\hat{F}_\tau$, update $\hat{\boldsymbol{b}}_{i,\tau}$ and $\hat{\boldsymbol{\lambda}}_{i,\tau}$ by using (8) for $i = 1, 2, \ldots, N$.

Step 4. Given $\hat{\boldsymbol{\rho}}_\tau$, $\hat{B}_\tau$ and $\hat{\Lambda}_\tau$, update $\hat{\boldsymbol{f}}_{t,\tau}$ by using (9) for $t = 1, 2, \ldots T$.

Step 5 . Repeat Step 2 $\sim$ Step 4 until convergence.

Step 6. Apply (10) to ensure that $\hat{F}_\tau$ and $\hat{\Lambda}_\tau$ satisfy the normalization conditions (3).

The tolerance condition to stop the iteration can be based on the difference of the estimators in the last iteration and the current iteration, that is, the tolerance condition can be formulated as

$$N^{-1}\sum_{i=1}^{N}(\hat{\rho}_{i,\tau}^{new} - \hat{\rho}_{i,\tau}^{old})^2 + N^{-1}\sum_{i=1}^{N}\|\hat{\boldsymbol{b}}_{i,\tau}^{new} - \hat{\boldsymbol{b}}_{i,\tau}^{old}\|^2 + (NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[(\hat{\boldsymbol{f}}_{t,\tau}'\hat{\boldsymbol{\lambda}}_{i,\tau})^{new} - (\hat{\boldsymbol{f}}_{t,\tau}'\hat{\boldsymbol{\lambda}}_{i,\tau})^{old}\right]^2 < \delta^2,$$

where δ^2 is some pre-specified value, which can be, say, 10^{-3}.

Remark 2 The initial values of parameters in Step 1 can be delivered as follows. First, regarding the spatial parameter, for each t, we can treat y_{it} as a cross-sectional spatial autoregressive model ignoring the explanatory variable and the common factor structure to get an estimate of $\rho_{i,\tau}$ by simple OLS. Next, for each i, we directly apply a cross-sectional quantile regression to the transformed data $(y_{it} - \hat{\rho}_{i,\tau,ini} \sum_{j=1}^{N} w_{ij} y_{jt})$ to obtain $\hat{\boldsymbol{b}}_{i,\tau,ini}$, ignoring the common factor structure. Third, we obtain an estimate of $\hat{F}_\tau = (\hat{\boldsymbol{f}}_{1,\tau}, ..., \hat{\boldsymbol{f}}_{T,\tau})'$ by the Principal Component Method based on dataset Z_τ with its (i,t)-th element being $z_{it} = y_{it} - \hat{\rho}_{i,\tau,ini} \sum_{j=1}^{N} w_{ij} y_{jt} - \boldsymbol{x}_{it}' \hat{\boldsymbol{b}}_{i,\tau,ini}$, subject to the normalization condition $F_\tau' F_\tau / T = I_{r_\tau}$. Finally, we set $\hat{\Lambda}_{\tau,ini} = (\hat{\boldsymbol{\lambda}}_{1,\tau,ini}, \cdots, \hat{\boldsymbol{\lambda}}_{N,\tau,ini})'$ as $Z_\tau \hat{F}_\tau (\hat{F}_\tau' \hat{F}_\tau)^{-1}$.

It can be shown that the loss function defined in (6) does not increase in Steps 2-4 of each iteration due to the fact that the parameters are updated as the local minimizers of the loss function given the current values of other parameters. So the resultant estimators converge at least to a local minimizer, although the global minimality cannot be guaranteed because of the non-convexity of the loss function arising from the interactive-effects term $\lambda_{i,\tau}' f_{t,\tau}$ and the spatial term. In the simulations of Section 6, we find that the estimators from the proposed algorithm converge quickly under large N and large T, and are robust to the initial values. So it looks like the proposed algorithm does not suffer from severe local minimizer issues.

3.2 Model selection

Let $\hat{\boldsymbol{\rho}}_\tau(r)$, $\hat{\boldsymbol{b}}_{k,\tau}(r)$, $\hat{\boldsymbol{f}}_{t,\tau}(r)$ and $\hat{\boldsymbol{\lambda}}_{k,\tau}(r)$ be the estimators when the number of common factors is set to r. According to (4), one can readily compute the conditional quantile function of y_{it},

$$\widehat{Q}_{y_{it}}^{(r)}(\tau) \equiv \sum_{k=1}^{N} p_{ik}\big(\hat{\boldsymbol{\rho}}_\tau(r)\big) \Big[\boldsymbol{x}_{kt}' \hat{\boldsymbol{b}}_{k,\tau}(r) + \hat{\boldsymbol{f}}_{t,\tau}(r)' \hat{\boldsymbol{\lambda}}_{k,\tau}(r)\Big].$$

The number of factors can be estimated by minimizing the following information criterion

$$IC_\tau(r) = \log\Big[\frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} q_\tau\Big(y_{it} - \widehat{Q}_{y_{it}}^{(r)}(\tau)\Big)\Big] + r \times q(N,T) \tag{11}$$

where $q_\tau(\cdot)$ is the quantile loss function defined in (6), and $q(N,T)$ is a penalty on overfitting of the interactive effects. In the numerical study, we specify this function as

$$q(N,T) = \log\Big(\frac{NT}{N+T}\Big)\Big(\frac{N+T}{NT}\Big). \tag{12}$$

The asymptotic validity of $IC_\tau(r)$ is justified by Theorem 4 in Section 5.

4 Assumptions

We denote the true spatial parameter and the true regression coefficient as $\rho_{i,0,\tau}$ and $\boldsymbol{b}_{i,0,\tau}$, respectively. Similarly, we denote $F_{0,\tau} = (\boldsymbol{f}_{1,0,\tau}, ..., \boldsymbol{f}_{T,0,\tau})'$ and $\Lambda_{0,\tau} = (\boldsymbol{\lambda}_{1,0,\tau}, ..., \boldsymbol{\lambda}_{N,0,\tau})'$ as the true factors and loadings. A set of regularity conditions that are needed for theoretical analysis are given as follows.

Assumption A: Common factors

Let $\mathcal{F}$ be a compact subset of R^{r_τ}. The common factors $\boldsymbol{f}_{t,0,\tau} \in \mathcal{F}$ satisfy $T^{-1}\sum_{t=1}^{T} \boldsymbol{f}_{t,0,\tau}\boldsymbol{f}'_{t,0,\tau} = I_{r_\tau}$.

Assumption B: Factor loadings and regression coefficients

(B1) Let $\mathcal{P}$, $\mathcal{B}$ and $\mathcal{L}$ be compact subsets of $\mathbb{R}$, $\mathbb{R}^{p+1}$ and $\mathbb{R}^{r_\tau}$, respectively. The spatial parameter $\rho_{i,0,\tau}$, the regression coefficient $\boldsymbol{b}_{i,0,\tau}$, and the factor-loading $\boldsymbol{\lambda}_{i,0,\tau}$ satisfy that $\rho_{i,0,\tau} \in \mathcal{P}$, $\boldsymbol{b}_{i,0,\tau} \in \mathcal{B}$ and $\boldsymbol{\lambda}_{i,0,\tau} \in \mathcal{L}$ for each i.

(B2) The factor-loading matrix $\Lambda_{0,\tau} = (\boldsymbol{\lambda}_{1,0,\tau}, \ldots, \boldsymbol{\lambda}_{N,0,\tau})'$ satisfies $N^{-1}\sum_{i=1}^{N} \boldsymbol{\lambda}_{i,0,\tau}\boldsymbol{\lambda}'_{i,0,\tau} \xrightarrow{p} \Sigma_{\Lambda_\tau}$, where Σ_{Λ_τ} is an $r_\tau \times r_\tau$ positive definite diagonal matrix with diagonal elements distinct and arranged in the descending order. In addition, the eigenvalues of Σ_{Λ_τ} are distinct.

Assumption C: Idiosyncratic error terms

(C1): The random variable

$$\varepsilon_{it,\tau} = y_{it} - \sum_{k=1}^{N} p_{ik}(\boldsymbol{\rho}_{0,\tau})\Big(\boldsymbol{x}'_{kt}\boldsymbol{b}_{k,0,\tau} + \boldsymbol{f}'_{t,0,\tau}\boldsymbol{\lambda}_{k,0,\tau}\Big)$$

satisfies $P(\varepsilon_{it,\tau} \leq 0) = \tau$, and is independently distributed over i and t, conditional on X_t, $\boldsymbol{\rho}_{0,\tau}$, $B_{0,\tau}$, $F_{0,\tau}$ and $\Lambda_{0,\tau}$.

(C2): The conditional density function of $\varepsilon_{it,\tau}$ given $(X_t, \boldsymbol{\rho}_{0,\tau}, B_{0,\tau}, F_{0,\tau}, \Lambda_{0,\tau})$, denoted as $g_{it}(\varepsilon_{it,\tau})$, is continuous. In addition, for any compact set $\mathcal{C}$, there exists a positive constant $\underline{g} > 0$ (depending on $\mathcal{C}$) such that $\inf_{c\in\mathcal{C}} g_{it}(c) \geq \underline{g}$ for all i and t.

Assumption D: Weight matrix

(D1): W is an exogenous spatial weights matrix whose diagonal elements of W are all zeros. In addition, W is bounded by some constant C for all N under $\|\cdot\|_1$ and $\|\cdot\|_\infty$.

(D2): Define $P(\boldsymbol{\rho}_\tau) = (I - \boldsymbol{\rho}_\tau W)^{-1}$ for $\boldsymbol{\rho}_\tau = \mathrm{diag}(\rho_{1,\tau}, \rho_{2,\tau}, \ldots, \rho_{N,\tau})$, an N-dimensional diagonal matrix with $\rho_{i,\tau}$ being an interior point of $\mathcal{P}$ for each i. The matrix $P(\boldsymbol{\rho}_\tau)$ is invertible over $\mathcal{P}^N$ and satisfies

$$\limsup_{N\to\infty} \|P(\boldsymbol{\rho}_\tau)\|_1 \vee \|P(\boldsymbol{\rho}_\tau)\|_\infty \leq C,$$

for each $\mathrm{diag}(\boldsymbol{\rho}_\tau) \in \mathcal{P}^N$, where C is some positive constant large enough.

Assumption E: Explanatory variables and design matrix

(E1): For a positive constant C_x, explanatory variables satisfy $\sup_{it} \|\boldsymbol{x}_{it}\| \leq C_x$ almost surely.

(E2): Let $\mathcal{X}(B)$ be an $N \times T$ matrix with its (i,t)-th entry equal to $\boldsymbol{x}_{it}'\boldsymbol{b}_i$, where $B = (\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_N)'$. Define $d_{it,0,\tau}$ to be the (i,t)th element of $W(I_N - \boldsymbol{\rho}_{0,\tau}W)^{-1}(\mathcal{X}(B_{0,\tau}) + \Lambda_{0,\tau}F_{0,\tau}')$. Let $\boldsymbol{v}_{it,\tau} = (d_{it,0,\tau}, \boldsymbol{x}_{it}')'$ and $\boldsymbol{V}_{i,\tau} = (\boldsymbol{v}_{i1,\tau}, \boldsymbol{v}_{i2,\tau}, \ldots, \boldsymbol{v}_{iT,\tau})'$. Define $A_{i,\tau} = \frac{1}{T}\boldsymbol{V}_{i,\tau}' M_{F_\tau} \boldsymbol{V}_{i,\tau}$, $B_{i,\tau} = (\boldsymbol{\lambda}_{i,0,\tau}\boldsymbol{\lambda}_{i,0,\tau}') \otimes I_T$, $C_i = \frac{1}{\sqrt{T}}[\boldsymbol{\lambda}_{i,0,\tau} \otimes (M_{F_\tau}\boldsymbol{V}_{i,\tau})]'$, $\boldsymbol{\eta} = \frac{1}{\sqrt{T}}\mathrm{vec}(M_{F_\tau}F_{0,\tau})$, and $M_{F_\tau} = I - F_\tau(F_\tau'F_\tau)^{-1}F_\tau'$. Let $\mathscr{F}_\tau$ be the collection of F_τ such that $\mathscr{F}_\tau = \{F_\tau : F_\tau'F_\tau/T = I_{r_\tau}\}$. We assume that with probability approaching one,

$$\inf_{F_\tau \in \mathscr{F}_\tau} \lambda_{\min}\left[\frac{1}{N}\sum_{i=1}^N E_{i,\tau}(F_\tau)\right] > 0,$$

where $\lambda_{\min}(\mathbb{M})$ denotes the smallest eigenvalue of matrix $\mathbb{M}$, and $E_{i,\tau}(F_\tau) = B_{i,\tau} - C_{i,\tau}'A_{i,\tau}^{-1}C_{i,\tau}$.

(E3): Let $\mathcal{V}_\tau(\boldsymbol{\phi}_\tau)$ be the $N \times T$ matrix with its (i,t)th entry equal to $d_{it,0,\tau}\rho_{i,\tau} + \boldsymbol{x}_{it}'\boldsymbol{b}_{i,\tau}$, where $\boldsymbol{\phi}_\tau = (\boldsymbol{\phi}_{1,\tau}, \boldsymbol{\phi}_{2,\tau}, \ldots, \boldsymbol{\phi}_{N,\tau})'$ and $\boldsymbol{\phi}_{i,\tau} = (\rho_{i,\tau}, \boldsymbol{b}_{i,\tau}')'$. For any nonzero $\boldsymbol{\phi}_\tau$, there exists a positive constant $\check{c} > 0$ such that with probability approaching one,

$$\frac{1}{NT}\|M_{\Lambda_{0,\tau}}\mathcal{V}_\tau(\boldsymbol{\phi}_\tau)M_{F_{0,\tau}}\|^2 \geq \check{c}\frac{1}{N}\sum_{i=1}^N \|\boldsymbol{\phi}_{i,\tau}\|^2,$$

where $M_{\Lambda_{0,\tau}} = I - \Lambda_{0,\tau}(\Lambda_{0,\tau}'\Lambda_{0,\tau})^{-1}\Lambda_{0,\tau}'$.

(E4): For each i, we assume that there exists a constant $c > 0$ such that for each i, with probability approaching one,

$$\liminf_{T\to\infty} \lambda_{\min}\left(\frac{1}{T}\boldsymbol{V}_{i,\tau}'M_{F_{0,\tau}}\boldsymbol{V}_{i,\tau}\right) \geq c.$$

Remark 3 Several comments on our assumptions are provided here. Assumptions A and B are consistent with our normalization conditions (3). In the factor literature, it is more common to assume that (i) $T^{-1}\sum_{t=1}^{T}\boldsymbol{f}_{t,0,\tau}\boldsymbol{f}'_{t,0,\tau}\xrightarrow{p}\Sigma_{f,\tau}$ and (ii) $\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\lambda}_{i,0,\tau}\boldsymbol{\lambda}'_{i,0,\tau}\xrightarrow{p}\Sigma_{\lambda,\tau}$, where $\Sigma_{f,\tau}$ and $\Sigma_{\lambda,\tau}$ are both positive definite matrices. However, as pointed out in the discussions related with (10), we can always transform any set that satisfies (i) and (ii) into one that satisfies (3). For this reason, we directly impose the conditions that feature (3).

Assumption C1 may be relaxed to allow the serial dependency but we do not pursue this direction in this paper. Assumption C2 is standard in the quantile models. Assumption D1 is a standard assumption for the spatial weights matrix and commonly used in the spatial modeling literature, such as Kelejian and Prucha (2004), Lee (2004), Yu et al. (2008), Bai and Li (2013, 2021) and Shi and Lee (2017), among others. Zero values of diagonal element of W entails no loss of generality. To see this, note that

$$
\begin{aligned}
&Q_{y_{it}}\left(\tau|\boldsymbol{Q}_{(-i)t}(\tau),\boldsymbol{x}_{it},\rho_{i,0,\tau},\boldsymbol{b}_{i,0,\tau},\boldsymbol{f}_{t,0,\tau},\boldsymbol{\lambda}_{i,0,\tau}\right)\\
&\quad=\rho_{i,0,\tau}\sum_{j=1}^{N}w_{ij}Q_{y_{jt}}\left(\tau|\boldsymbol{Q}_{(-j)t}(\tau),\boldsymbol{x}_{jt},\rho_{j,0,\tau},\boldsymbol{b}_{j,0,\tau},\boldsymbol{f}_{t,0,\tau},\boldsymbol{\lambda}_{j,0,\tau}\right)+\boldsymbol{x}'_{it}\boldsymbol{b}_{i,0,\tau}+\boldsymbol{f}'_{t,0,\tau}\boldsymbol{\lambda}_{i,0,\tau},
\end{aligned}
$$

if $w_{ii}\neq 0$, the model can be alternatively written as

$$
\begin{aligned}
&(1-\rho_{i,0,\tau}w_{ii})Q_{y_{it}}\left(\tau|\boldsymbol{Q}_{(-i)t}(\tau),\boldsymbol{x}_{it},\rho_{i,0,\tau},\boldsymbol{b}_{i,0,\tau},\boldsymbol{f}_{t,0,\tau},\boldsymbol{\lambda}_{i,0,\tau}\right)\\
&=\rho_{i,0,\tau}\sum_{j=1,j\neq i}^{N}w_{ij}Q_{y_{jt}}\left(\tau|\boldsymbol{Q}_{(-j)t}(\tau),\boldsymbol{x}_{jt},\rho_{j,0,\tau},\boldsymbol{b}_{j,0,\tau},\boldsymbol{f}_{t,0,\tau},\boldsymbol{\lambda}_{j,0,\tau}\right)+\boldsymbol{x}'_{it,\tau}\boldsymbol{b}_{i,0,\tau}+\boldsymbol{\lambda}'_{i,0,\tau}\boldsymbol{f}_{t,0,\tau}.
\end{aligned}
$$

Let

$$
\rho^{\dagger}_{i,0,\tau}=\frac{\rho_{i,0,\tau}}{1-\rho_{i,0,\tau}w_{ii}},\quad \boldsymbol{b}^{\dagger}_{i,0,\tau}=\frac{\boldsymbol{b}_{i,0,\tau}}{1-\rho_{i,0,\tau}w_{ii}},\quad\text{and}\quad \boldsymbol{\lambda}^{\dagger}_{i,0,\tau}=\frac{\boldsymbol{\lambda}_{i,0,\tau}}{1-\rho_{i,\tau,0}w_{ii}},
$$

We obtain an equivalent model with the diagonal elements of W being zeros. So this condition is indispensable for the identification of $\boldsymbol{\rho}_{0,\tau}$. Assumption D2 is imposed for the purpose of theoretical analysis. As seen in the online supplement, the matrix $P(\boldsymbol{\rho}_\tau)$ plays an important role in our theoretical analysis. This assumption is widely-adopted in the spatial econometrics.

Assumption E is imposed for the identification of parameters. More specifically, Assumption E.2 is imposed for the identification of $\Lambda_\tau F'_\tau$. It has been made in several previous studies, such as Ando and Bai (2020). Our assumption E.2 is slightly different from those in the previous studies because it needs to accommodate the presence of the spatial term, but the mechanism behind this assumption is the same. This assumption, together with the normalization conditions (3) which are used to preclude the rotational indeterminacy, can give the identification of Λ_τ and F_τ (up to column signs).

Assumption E.3 can be viewed as an extended version of the condition $\inf_{F \in \mathscr{F}} D(F) > 0$ in Assumption A of Bai (2009). Note that since we have already identified $F_{0,\tau}$ in Assumption E.2, we only need $D(F_{0,\tau}) > 0$ instead of $\inf_{F \in \mathscr{F}} D(F) > 0$. Because of the heterogeneous coefficients and introduction of the spatial term, we generalize $D(F_{0,\tau}) > 0$ to Assumption E.3. Basically speaking, Assumption E.3 is imposed for the identification of $\phi_{i,0,\tau} = (\rho_{i,0,\tau}, \boldsymbol{b}'_{i,0,\tau})'$ under the average Frobenius norm. In other words, for any $\phi_a = [\phi_{a,1}, \ldots, \phi_{a,N}]$ and $\phi_b = [\phi_{b,1}, \ldots, \phi_{b,N}]$ such that

$$\liminf_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} \|\phi_{a,i} - \phi_{b,i}\|^2 \geq c$$

for some constant $c > 0$, ϕ_a can be separated from ϕ_b in our model. This implies that for the unknown true values $\phi_{0,\tau} = (\phi_{1,0,\tau}, \phi_{2,0,\tau}, \ldots, \phi_{N,0,\tau})$, it can be identified from those points ϕ such that

$$\liminf_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} \|\phi_i - \phi_{i,0,\tau}\|^2 \geq c.$$

By Assumption E.3 we can establish the global properties of the estimators, but this assumption is still insufficient to achieve the full identification. Consider $\phi^* = (\phi_1^*, \phi_{2,0,\tau}, \ldots, \phi_{N,0,\tau})$ with $\phi_1^* \neq \phi_{1,0,\tau}$. Assumption E.3 would fail to separate ϕ^* from $\phi_{0,\tau}$. For this reason, we further impose Assumption E.4 to achieve the individual (or full) identification.

To explain the intuitions of E.4, we start our discussions with the traditional spatial autoregressive (SAR) model considered in Lee (2004)

$$Y = \rho_0 WY + X\beta_0 + e.$$

We treat X as non-random for simplicity. Taking expectation on both sides, we obtain

$$\begin{aligned} E(Y) &= (I_N - \rho_0 W)^{-1} X\beta_0 = (I_N - \rho_0 W + \rho_0 W)(I_N - \rho_0 W)^{-1} X\beta_0 \\ &= \underbrace{\left[W(I_N - \rho_0 W)^{-1} X\beta_0\right]}_{D} \rho_0 + X\beta_0. \end{aligned}$$

Like the linear regression $E(Y) = X\beta_0$, the spatial model can be viewed as a linear regression model with a generated regressor D. For this reason, Lee (2004) imposes the following condition (Assumption 8 in Lee (2004)) to identify ρ_0 and β_0:

Identification condition for the SAR model (ICsar): The matrix $V = [D, X]$ is of full column rank.

In our paper, Assumption E.4 plays the same role as the above ICsar. To see this, first note that in our paper we also have a generated regressor

$$\boldsymbol{D}_{0,\tau} = W(I_N - \boldsymbol{\rho}_{0,\tau} W)^{-1}\Big(\mathcal{X}(B_{0,\tau}) + \Lambda_{0,\tau} F'_{0,\tau}\Big),$$

which has the same structure as D in the SAR model. Therefore the matrix $\boldsymbol{V}_{i,\tau} = (\boldsymbol{v}_{i1,\tau}, \boldsymbol{v}_{i2,\tau}, \ldots, \boldsymbol{v}_{iT,\tau})'$ with $\boldsymbol{v}_{it,\tau} = (d_{it,0,\tau}, \boldsymbol{x}'_{it})'$ in our paper is the counterpart of V in the SAR model. Because of the presence of interactive effects, the model now is

$$Q_{y_{it}}(\tau | X, \Lambda_{0,\tau}, F_{0,\tau}, \boldsymbol{\rho}_{0,\tau}, B_{0,\tau}) = d_{it,0,\tau}\rho_{i,0,\tau} + \boldsymbol{x}'_{it}\boldsymbol{b}_{i,0,\tau} + \boldsymbol{f}'_{t,0,\tau}\boldsymbol{\lambda}_{i,0,\tau}$$

Suppose that we observe $\boldsymbol{f}_{t,0,\tau}$. Now the identification condition should be that $[\boldsymbol{V}_{i,\tau}, F_{0,\tau}]$ is of the full column rank. Since the full column rank of $F_{0,\tau}$ is guaranteed by Assumption A, one only needs to impose

$$\liminf_{T\to\infty} \lambda_{\min}\Big(\frac{1}{T}\boldsymbol{V}'_{i,\tau} M_{F_{0,\tau}} \boldsymbol{V}_{i,\tau}\Big) \geq c,$$

as required by Assumption E.4. Note that the ICsar would break down if $\beta_0 = 0$. The similar situation also appears in Assumption E.4, that is, if $\boldsymbol{b}_{i,0,\tau} = 0$ for all i, the above condition would break down.

Unfortunately, both $\boldsymbol{f}_{t,0,\tau}$ and $\boldsymbol{\lambda}_{i,0,\tau}$ are unobserved in our model. For this reason, Assumption E.4 alone is not enough to achieve identification. But with assistance of Assumptions E.2 and E.3, the identification conditions in the interactive-effects models, full identification is obtained.

5 Asymptotic theory

This section presents the asymptotic properties of the estimators. Recent literature pays much attention on quantile factor models, see Chen et al. (2021) and He et al. (2020). The current model is an extension of the quantile factor model because we explicitly allow the spatial term and exogeneous regressors. Because of this feature, the analysis in this paper is more complicated. Some difficulties are illustrated below. We first establish the consistency of the estimators, which is given in the following theorem.

Theorem 1 *Under Assumptions A–E, if* $\log N/T \to 0$*, the estimators are uniformly consistent, i.e.,*

$$\max_{1\le i\le N} |\hat{\rho}_{i,\tau} - \rho_{i,0,\tau}| = o_p(1), \qquad \max_{1\le i\le N} \|\hat{\boldsymbol{b}}_{i,\tau} - \boldsymbol{b}_{i,0,\tau}\| = o_p(1), \qquad \max_{1\le i\le N} \|\hat{\boldsymbol{\lambda}}_{i,\tau} - \boldsymbol{\lambda}_{i,0,\tau}\| = o_p(1).$$

In addition, let $\boldsymbol{\psi}_{k,0,\tau} = (\rho_{k,0,\tau}, \boldsymbol{b}'_{k,0,\tau}, \boldsymbol{\lambda}'_{k,0,\tau})'$ *and* $\hat{\boldsymbol{\psi}}_{k,\tau}$ *be its estimator. We have the following average consistency result*

$$\frac{1}{N}\sum_{i=1}^{N} \|\hat{\boldsymbol{\psi}}_{i,\tau} - \boldsymbol{\psi}_{i,0,\tau}\|^2 = o_p(1), \qquad \frac{1}{T}\sum_{t=1}^{T} \|\hat{\boldsymbol{f}}_{t,\tau} - \boldsymbol{f}_{t,0,\tau}\|^2 = o_p(1).$$

The proof of Theorem 1 requires considerable amount of work. Apart from the theoretical difficulties in the quantile factor model pointed out by Chen et al. (2021), such as no closed form for the estimators, the presence of the spatial term and the exogenous regressors poses new theoretical challenges to our analysis. For example, we would like to show that the non-random part of the objective function behaves like a quadratic function in the neighborhood of the true parameters. To this end, we need to separate the regression coefficients $\boldsymbol{\rho}_\tau$ and B_τ from the incidental parameters Λ_τ and F_τ because the latter suffers rotational indeterminacy. This separation is non-trivial and requires some new arguments.

Given the above results, we can further derive the following average convergence rate of the estimators. The results of the convergence rates are very useful to give the consistent information criterion.

Theorem 2 *Let* $\delta_{NT} = \max(\frac{1}{\sqrt{N}}, \frac{1}{\sqrt{T}})$*. Under the assumptions of Theorem 1, we have*

$$\frac{1}{N}\sum_{i=1}^{N} \|\hat{\boldsymbol{\psi}}_{i,\tau} - \boldsymbol{\psi}_{i,0,\tau}\|^2 = O_p(\delta_{NT}^2), \qquad \frac{1}{T}\sum_{t=1}^{T} \|\hat{\boldsymbol{f}}_{t,\tau} - \boldsymbol{f}_{t,0,\tau}\|^2 = O_p(\delta_{NT}^2).$$

The above theorem indicates that the quantile estimators have the same average convergence rates as in the standard factor model and the mean regression with unobserved heterogeneity. This partially explains the results in Theorem 4, which proves that the penalty in the standard factor model continues to work in the quantile background when determining the number of factors.

Now we present the asymptotic normality of the estimators, which is summarized in the following theorem.

Theorem 3 *Under Assumptions A–E, if $N/T \to C^\diamond$ for some $C^\diamond \in (0,\infty)$, we have, for each k,*

$$\sqrt{T}(\hat{\boldsymbol{\psi}}_{k,\tau} - \boldsymbol{\psi}_{k,0,\tau}) \xrightarrow{d} N(0, \tilde{\boldsymbol{\Omega}}_{k,\tau}^{-1}\boldsymbol{\Omega}_{k,\tau}\tilde{\boldsymbol{\Omega}}_{k,\tau}^{-1}),$$

where

$$\tilde{\boldsymbol{\Omega}}_{k,\tau} = \lim_{T\to\infty}\frac{1}{T}\sum_{i=1}^{N}\sum_{t=1}^{T} g_{it,0,\tau}p_{ik,0,\tau}^2 \boldsymbol{w}_{kt,0,\tau}\boldsymbol{w}'_{kt,0,\tau},$$
$$\boldsymbol{\Omega}_{k,\tau} = \lim_{T\to\infty}\tau(1-\tau)\frac{1}{T}\sum_{i=1}^{N}\sum_{t=1}^{T} p_{ik,0,\tau}^2 \boldsymbol{w}_{kt,0,\tau}\boldsymbol{w}'_{kt,0,\tau};$$

with $\boldsymbol{w}_{it,\tau} = (d_{it,0,\tau}, \boldsymbol{x}'_{it}, \boldsymbol{f}'_{t,0,\tau})'$, $g_{it,0,\tau} = g_{it}(G_{it}^{-1}(\tau))$ and $p_{ik,0,\tau}$ is the (i,k)-th element of the matrix $(I_N - \boldsymbol{\rho}_{0,\tau}W)^{-1}$, where $d_{it,0,\tau}$ is defined in Assumption E.2, $g_{it}(\cdot)$ and $G_{it}(\cdot)$ are the density and cumulative distribution function of ε_{it}, respectively.

For each t,

$$\sqrt{N}(\hat{\boldsymbol{f}}_{t,\tau} - \boldsymbol{f}_{t,0,\tau}) \xrightarrow{d} N(0, \tilde{\boldsymbol{\Omega}}_{t,\tau}^{*-1}\boldsymbol{\Omega}_{\tau}^{*}\tilde{\boldsymbol{\Omega}}_{t,\tau}^{*-1}).$$

where

$$\tilde{\boldsymbol{\Omega}}_{t,\tau}^{*} = \operatorname*{plim}_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} g_{it,0,\tau}\boldsymbol{\lambda}_{i,0,\tau}^{*}\boldsymbol{\lambda}_{i,0,\tau}^{*\prime}, \qquad \boldsymbol{\Omega}_{\tau}^{*} = \operatorname*{plim}_{N\to\infty}\tau(1-\tau)\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\lambda}_{i,0,\tau}^{*}\boldsymbol{\lambda}_{i,0,\tau}^{*\prime},$$

where $\boldsymbol{\lambda}_{i,0,\tau}^{*} = \sum_{j=1}^{N} p_{ij,0,\tau}\boldsymbol{\lambda}_{j,0,\tau}$.

Theorem 3 implies that the estimator $\hat{\boldsymbol{\psi}}_{k,\tau}$ possesses the oracle properties in the sense that the limiting distribution of $\hat{\boldsymbol{\psi}}_{k,\tau}$ is the same as that of the infeasible estimator that is obtained when assuming the factors are observed a priori. The simulation results in Section 6 support this result.

In the following theorem, we show that our proposed information criterion is capable of estimating the true dimension of the interactive effects with probability approaching one. To this end, we additionally impose Assumption F to ensure that the parameters $(\rho_{i,0,\tau}, \boldsymbol{b}'_{i,0,\tau})'$ $(i = 1,\ldots,N)$ are identified in the over-fitted model.

Assumption F: Identification of b for the over-fitted model

For each k, there exists a constant $\check{c}_k > 0$ such that, with probability approaching one,

$$\inf_{F^k, \frac{1}{T}F^{k\prime}F^k = I_k}\frac{1}{NT}\|M_{\Lambda_{0,\tau}}\mathcal{V}_\tau(\boldsymbol{\phi}_\tau)M_{F^k}\|^2 \geq \check{c}_k\frac{1}{N}\sum_{i=1}^{N}\|\boldsymbol{\phi}_{i,\tau}\|^2,$$

where $\mathcal{V}_\tau(\phi_\tau)$ is defined in Assumption E.3.

To see the implication of the above assumption, one may consider the following linear model

$$Y_i = \boldsymbol{V}_i'\phi_{i,0} + F_0\boldsymbol{\lambda}_{i,0} + \varepsilon_i, \quad \text{or equivalently} \quad Y = \mathcal{V}(\phi_0) + \Lambda_0' F_0 + \varepsilon.$$

If there exists F^k such that $\frac{1}{NT}\|\mathcal{V}(\phi)M_{F^k}\|^2 = o_p(1)\frac{1}{N}\sum_{i=1}^N \|\phi_i\|^2$, we can simply post-multiply $M_{[F^k,F_0]}$ on both sides, and it is seen that all the useful information (i.e., $\mathcal{X}(B_0) + \Lambda_0' F_0$) are projected out. So $\phi_{i,0}$ would have an identification issue. Therefore it is plausible to impose $\frac{1}{NT}\|\mathcal{V}(\phi)M_{F^k}\|^2 \geq c\frac{1}{N}\sum_{i=1}^N \|\phi_i\|^2$ for all k. The above assumption is a little stronger than this requirement because it still needs to preclude the effect from the loadings Λ_0, but the intuition is the same.

Theorem 4 *Under Assumptions A-F, if $N/T \to c^\diamond$ for some $c^\diamond \in (0,\infty)$, the model selection criterion (11) would correctly estimate the dimensionality of the interactive effects with probability approaching one if the penalty function $q(N,T)$ satisfies*

$$q(N,T) \to 0 \;\text{ and }\; C_{NT} \times q(N,T) \to \infty,$$

where $C_{NT} = \min\{N,T\}$.

6 Monte Carlo simulation

We run simulations to examine the finite sample performance of the quantile estimators and the proposed information criterion. We separately evaluate the performance of the quantile estimators (assuming that the number of factors is known) and the performance of the proposed information criterion because this allows us to see each performance more clearly. In the appendix, we present the simulation results with the estimated number of factors. Overall, it can be seen there that the quantile estimators with the estimated number of factors have very similar performance as in the known-number case.

6.1 Performance of the quantile estimators

This subsection investigates the finite sample performance of the quantile estimation. Our data generating process (DGP) is essentially based on the quantile function (1), or equivalently its reduced form (4). Concretely, we first draw u_{it} independently from the uniform distribution $U[0,1]$ as our quantiles for each i and t. Fixing (i,t), based on u_{it} together with $(\rho_{j,u_{it}}, \boldsymbol{b}_{j,u_{it}}, \boldsymbol{\lambda}_{j,u_{it}})$ and $\boldsymbol{f}_{t,u_{it}}$ which are given below, we calculate $y_{it,u_{it}}$, the (i,t)th element of the $N \times T$ observed data matrix of dependent variable, via the expression

$$y_{it,u_{it}} = \sum_{k=1}^{N} p_{ik}(\boldsymbol{\rho}_{u_{it}})\left(\boldsymbol{x}_{kt}' \boldsymbol{b}_{k,u_{it}} + \boldsymbol{f}_{t,u_{it}}' \boldsymbol{\lambda}_{k,u_{it}}\right),$$

where the above expression implicitly uses the spatial weights matrix, which is generated as a decreasing exponential function of the distance between i and j according to $w_{ij} = 0.3^{|i-j|}$ if $i \neq j$, otherwise 0. Following the tradition of spatial econometrics, we row-normalize the original spatial weights matrix.

We consider the two types of DGPs for the regressors $\boldsymbol{x}_{it}$ and the remaining parameters $\rho_{j,u_{it}}, \boldsymbol{b}_{j,u_{it}}, \boldsymbol{\lambda}_{j,u_{it}}$ and $\boldsymbol{f}_{t,u_{it}}$. In DGP1, all parameters are constant across quantiles, i.e., they do not depend on τ. In DGP2, all parameters are extended to allow variation across quantiles, which is more plausible in viewpoint of applications.

DGP1. The spatial coefficient that varies across individuals is generated via $\rho_{k,u_{it}} = 0.5 + k/(100 * N)$. We set the number of common factors to be one and fixed across quantiles. The common factor $\boldsymbol{f}_{t,u_{it}} = f_{t1}$ and factor loading $\boldsymbol{\lambda}_{k,u_{it}} = \xi_{k1}$, where f_{t1} and ξ_{k1} are generated independently from the uniform distribution over $[0,2]$ and $[-2,2]$ respectively. The regressors $\boldsymbol{x}_{kt} = (x_{kt,1}, x_{kt,2}, x_{kt,3})'$, a three dimensional vector, is generated to allow correlations with the factor structure

$$x_{kt,1} = 1, \qquad x_{kt,2} = v_{kt,1} + 0.01 f_{t1}^2 + 0.01 \xi_{k1}^2, \qquad x_{kt,3} = v_{kt,2},$$

where $v_{kt,1}$ and $v_{kt,2}$ are both generated independently from a uniform distribution over $[0,1]$. The coefficient $\boldsymbol{b}_{k,u_{it}} = (b_{k1,u_{it}}, b_{k2,u_{it}}, b_{k3,u_{it}})'$ is generated by

$$b_{k1,u_{it}} = \Phi^{-1}(u_{it}), \qquad b_{k2,u_{it}} = -2 + \frac{k}{N}, \qquad b_{k3} = 2 + \frac{k}{N}$$

where $\Phi(\cdot)$ is the standard accumulative normal distribution. Putting all generated quantities into (4), we obtain y_{it}, which is the ith element of N-dimensional vector in (4).

DGP2. The regressors $\boldsymbol{x}_{kt}$ are generated by the same way as in DGP1. The spatial coefficient is generated via $\rho_{k,u_{it}} = 0.5 + u_{it}/100 + k/(100 * N)$. The generated $\boldsymbol{\rho}_\tau$ in this subsection has a similar magnitude as the estimated spatial coefficients in the empirical application below. We generate a $T \times 3$ common factor matrix $F = (f_{tl})$ ($t = 1, 2, \cdots, T$ and $k = 1, 2, 3$) such that each element follows a uniform distribution over $[0, 2]$. Then using the generated u_{it}, we create the common factors for the i-th unit at time t as

$$\boldsymbol{f}_{t,u_{it}} = \begin{cases} (f_{t1}) & \text{if } u_{it} \leq 0.2 \\ (f_{t1}, f_{t2})' & \text{if } 0.2 < u_{it} \leq 0.8 \\ (f_{t1}, f_{t2}, f_{t3})' & \text{if } 0.8 < u_{it} \end{cases} .$$

The dimension of the common factor structure varies over i and t since $u_{it} \sim U[0, 1]$. We also specify the factor loadings to be dependent on u_{it}, according to:

$$\boldsymbol{\lambda}_{k,u_{it}} = \begin{cases} (\xi_{k1} + 0.01u_{it}) & \text{if } u_{it} \leq 0.2 \\ (\xi_{k1} + 0.01u_{it}, \xi_{k2} + 0.01u_{it})' & \text{if } 0.2 < u_{it} \leq 0.8 \\ (\xi_{k1} + 0.01u_{it}, \xi_{k2} + 0.01u_{it}, \xi_{k3} + 0.01u_{it})' & \text{if } 0.8 < u_{it} \end{cases} ,$$

where ξ_{k1}, ξ_{k2} and ξ_{k3} are generated from a uniform distribution over $[-2, 2]$, independently. The coefficient $\boldsymbol{b}_{k,u_{it}} = (b_{k1,u_{it}}, b_{k2,u_{it}}, b_{k3,u_{it}})'$ is generated by:

$$b_{k1,u_{it}} = \Phi^{-1}(u_{it}), \qquad b_{k2,u_{it}} = -2 + \frac{k}{N} + 0.01u_{it}, \qquad b_{k3,u_{it}} = 2 + \frac{k}{N} + 0.01u_{it}$$

Putting all generated quantities into (4), we obtain y_{it}, which is the ith element of N-dimensional vector in (4).

Simulation result. We measure the performance of the estimator by reporting the average mean squared error (MSE) between the true value and the estimates, defined as

$$MSE_b = \frac{1}{SNp} \sum_{s=1}^{S} \sum_{i=1}^{N} \| \boldsymbol{b}_{i,\tau,0}^{(s)} - \hat{\boldsymbol{b}}_{i,\tau}^{(s)} \|^2 \tag{13}$$

$$MSE_\rho = \frac{1}{SN} \sum_{s=1}^{S} \sum_{i=1}^{N} (\rho_{i,\tau,0}^{(s)} - \hat{\rho}_{i,\tau}^{(s)})^2 \tag{14}$$

where $\boldsymbol{b}_{i,\tau,0}^{(s)}$ and $\rho_{i,\tau,0}^{(s)}$ are the true values in the sth repetition, and $\hat{\boldsymbol{b}}_{i,\tau}^{(s)}$ and $\hat{\rho}_{i,\tau}^{(s)}$ are the corresponding estimators; S represents the total number of repetitions and is set to 1000 in this paper. In our simulations, we are more interested in estimation precision of the coefficients of the generated regressors, i.e., $b_{i2,\tau}$ and $b_{i3,\tau}$ for each i. So $\boldsymbol{b}_{i,\tau,0}^{(s)}$ and $\hat{\boldsymbol{b}}_{i,\tau}^{(s)}$ are actually two dimensional vectors obtained by deleting the first element of the original estimates and $p = 2$ in (13).

Tables 1 and 2 report the MSEs for $\boldsymbol{b}_\tau$ and $\boldsymbol{\rho}_\tau$ under different combinations of N and T, together with the corresponding standard deviations (SE) of the MSEs over 1000 repetitions. We also report the performance of the infeasible estimates in the case where the factors are observed a priori. The results of the infeasible estimates serve as a benchmark for comparison. There are two main findings from the results. First, across quantiles and under both DGP1 and DGP2, the MSEs for both $\boldsymbol{b}$ and $\boldsymbol{\rho}$ do not change much when N increases, but decrease significantly when T increases. This is consistent with the result in Theorem 3 that the convergence rate of $\boldsymbol{b}_{i,\tau}$ and $\rho_{i,\tau}$ is $\sqrt{T}$ when N and T are comparable. Second, it can be seen from the two tables that both MSE and SE in the case when the factors are unknown are close to the results in the infeasible case when the factors are assumed known, which collaborates our theoretical finding that the proposed estimator when the factors are unknown have the same asymptotic property as the infeasible estimator when the factors are known.

Moreover, we compare the performance of our estimator with other two estimators, one is the estimator without considering the spatial effect (denoted as $\hat{\boldsymbol{b}}_{i,\tau}^{(NoSpatial)}$), and the other is the estimator without considering the factor structure (denoted as $\hat{\boldsymbol{b}}_{i,\tau}^{(NoFactor)}$ and $\hat{\rho}_{i,\tau}^{(NoFactor)}$). More specifically, $\hat{\boldsymbol{b}}_{i,\tau}^{(NoSpatial)}$ is computed by minimizing the following loss function (similar to Ando and Bai (2020))

$$\ell_\tau(Y|X, B_\tau, F_\tau, \Lambda_\tau) = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} q_\tau\left(y_{it} - \boldsymbol{x}_{it}'\boldsymbol{b}_{i,\tau} - \boldsymbol{\lambda}_{i,\tau}'\boldsymbol{f}_{t,\tau}\right).$$

Also, $\hat{\boldsymbol{b}}_{i,\tau}^{(\text{NoFactor})}$ and $\hat{\rho}_{i,\tau}^{(\text{NoFactor})}$ are computed by minimizing the following loss function

$$\ell_\tau(Y|X, \boldsymbol{\rho}_\tau, B_\tau) = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} q_\tau\Big(y_{it} - Q_{y_{it}}(\tau|X_t, \boldsymbol{\rho}_\tau, B_\tau)\Big),$$

where

$$Q_{y_{it}}\Big(\tau|X_t, \boldsymbol{\rho}_\tau, B_\tau\Big) \equiv \sum_{k=1}^{N} p_{ik}(\boldsymbol{\rho}_\tau)\boldsymbol{x}_{kt}'\boldsymbol{b}_{k,\tau}$$

with $p_{ik}(\boldsymbol{\rho}_\tau)$ being defined as the (i,k)th entry of the matrix $(I - \boldsymbol{\rho}_\tau W)^{-1}$. The simulation results of MSEs for the above two estimators are reported in Table 3. The simulation results imply that the estimator accounting for the spatial effect and the factor structure has pronounced better performance than both the estimator ignoring the spatial effect, and the estimator ignoring the factor structure.

6.2 Performance of the model selection criterion

To investigate the performance of our model selection criterion, we use the same DGPs as in the previous section. By computing the scores of the $IC(r)$ over all possible r ranging from 0 to 7, the number of common factors is estimated as the minimizer of the $IC(r)$.

The simulation results are presented in Table 4 for DGP1 and Table 5 for DGP2, where both report the percentages of the estimated number of common factors $\hat{r}_\tau$ over 1000 repetitions. As shown in both tables, our information criterion can correctly select the number of common factors at a high percentage, close to 100%, when T and N are large.

Additionally, we have run simulations to investigate the performance of the quantile estimators using estimated $\hat{r}_\tau$ rather than the true r_τ. The corresponding simulation results are provided in Table 1 in the online supplement. The results indicate that using the estimated number of common factors does not change the performance of the proposed estimator much. A plausible explanation is that the proposed estimators are not sensitive to overestimation of the number of factors. This nice feature is proved by Moon and Weidner (2015) in the linear interactive-effects model. Our simulation results indicate that the same feature also holds in our model.

7 Empirical application

7.1 Data and model

We apply the proposed model to study the quantile co-movement of the US stock market and its determinants with explicit consideration of the production linkages. We consider three important periods: the period of the U.S. subprime mortgage crisis, the period of the European sovereign debt crisis, and the recent period of monetary policy tightening when the Fed began to increase the Federal Funds rate starting in December 2015. The three periods correspond to

Period I: January 1, 2007, to April 30, 2009

Period II: September 1, 2009, to December 31, 2012

Period III: December 1, 2015, to June 30, 2018

We look at all firms listed in the U.S. stock market at daily frequency. Our data source is the Center for Research in Security Prices from the Wharton Research Data Services. Stocks with no variation are excluded from the sample. To capture the production linkages, we first aggregate daily firm level stock returns to industry level returns, and next construct the industry level production network. The industry classification is implemented using the Input-Output (IO) account data from the Bureau of Economic Analysis (BEA). More details are provided in Section 7.2. We focus on the industry average returns instead of firms' returns for two reasons. First, the input-output linkages among industries are much more stable over time and are determined by technology rather than by choice. Second, sometimes large and financially-unconstrained firms dominate industry returns. With industry average returns, we can isolate the demand effects from other potential effects that might work through financial frictions. The industry level production network is widely-adopted in literature, see for instance Ahern and Harford (2014), Pasten et al. (2020), and Ozdagli and Weber (2020).

We consider using the following equation to model the U.S. stock market:

$$Q_{y_{it}}\left(\tau|\boldsymbol{Q}_{(-i)t}(\tau), \boldsymbol{x}_{it}, \rho_{i,\tau}, b_{i,\tau}, \boldsymbol{f}_{t,\tau}, \boldsymbol{\lambda}_{i,\tau}\right) = \alpha_{i,\tau} + \rho_{i,\tau}\sum_{j=1}^{N} w_{ij} Q_{y_{jt}}\left(\tau|\boldsymbol{Q}_{(-j)t}(\tau), \boldsymbol{x}_{jt}, \rho_{j,\tau}, b_{j,\tau}\boldsymbol{f}_{t,\tau}, \boldsymbol{\lambda}_{j,\tau}\right)$$
$$+ \textbf{InterestRate}_t * b_{i,\tau} + \boldsymbol{f}'_{t,\tau}\boldsymbol{\lambda}_{i,\tau}, \tag{15}$$

where the weight w_{ij} is the share of industry i's output being used by industry j as inputs, computed using the IO tables from BEA. More details about the weights are given in Section 7.2. The weights measure the intensity of the production linkage.

We include the interest rate change as the regressor, because we would like to see how the interest rate change affects the stock market directly and indirectly through the production network. Throughout the application, we use the U.S. 10-year Treasury yield change to represent the interest rate change. The 10-year Treasury yield change is at daily frequency.

In the following, Section 7.2 provides details about industry aggregation and weights computation and Section 7.3 presents the empirical results.

7.2 Industry aggregation and weights computation

To compute the industry average stock return, we follow the BEA's industry classification at the detailed accounts level, which is the 6-digit IO code level. This detailed classification has a total of 405 industries for the U.S. market.[3] Figures 1 and 2 in the online supplementary document plots the empirical production (input-output) network corresponding to U.S. Input-Output Data. Table 3 in the online supplementary document provides a list of industries.

We follow the same approach used in Ahern and Harford (2014), Pasten et al. (2020) and Ozdagli and Weber (2020) to construct the weights. From the BEA, we get the Supply Table ($SUPPLY$) and the Use Table (USE), where both are an industry-by-commodity matrix of dollar trade flows. The (i, c) entry in the Supply Table, denoted as $SUPPLY_{ic}$, represents the dollar amount of commodity c that industry i produces; the (j, c) entry in the Use Table, denoted as USE_{jc}, represents the dollar amount of commodity c that industry j purchases as inputs. We first use $SUPPLY$ to compute the share of each commodity c that each industry i produces. We call this the $SHARE$ matrix, which is industry-by-commodity, and each entry is computed by $SHARE_{ic} = \frac{SUPPLY_{ic}}{\sum_{c \in \mathbf{C}} SUPPLY_{ic}}$ with $\mathbf{C}$ being the collection of all commodities. Then we compute the dollar amount of commodities that industry j purchases from industry i, used as j's inputs. We call this the $FLOW$ matrix, which is industry-by-industry, and each entry is computed by $FLOW_{ij} = \sum_{c \in \mathbf{C}} SHARE_{ic} * USE_{cj}$. Finally, the weight is computed according to $w_{ij} = \frac{FLOW_{ij}}{\sum_{c \in \mathbf{C}} SUPPLY_{ic}}$.

In our empirical analysis, since the IO accounts data (Supply and Use Tables) at the most disaggregated level (i.e., the 6-digit IO code level) are published by BEA every five years, we use the IO tables in year 2007 for Period I and use the IO tables in year 2012 for Periods II and III, as they are the latest. To assess whether the IO tables in 2012 are similar to those in 2007, we implement the Mantel test (see Mantel (1967)) which is a popular statistical test of the correlation between two matrices. The result indicates the IO tables are statistically similar between 2007 and 2012.[4]

[3]Source: https://www.bea.gov/industry/input-output-accounts-data. IO tables provide IO code and name together with the corresponding NAICS code, where NAICS code information is also included in the CRSP stock data. So we aggregate firm level stock return to industry level stock return by matching NAICS with IO code.

[4] Additionally, we also compute similarity measures between the two versions, applying existing methods (for example, the Jaccard metric which is measured as the inverse of the number of links belonging to both networks divided by the number of links that belong to at least one network, see Anand et al. (2018) for more details), that give us a measure very close to one, implying the IO tables are very similar between 2007 and 2012.

7.3 Empirical results

7.3.1 Number of common factors

We consider five different quantiles: $\tau = 0.05, 0.25, 0.50, 0.75, 0.95$. The number of common factors is selected by the IC method in (11) with the maximum number set to ten. For each period and each quantile, the number of common factors is determined by the IC method.

Table 6 represents the estimated number of common factors for each period and each quantile. As we can see, the number of common factors for $\tau = 0.05$ in Period I (the U.S. subprime mortgage crisis period) is determined as 2. During Period I, the number of common factors is small and stable across quantiles, while in Period III (after the crisis, the recent period of monetary policy tightening), there are fewer detected common factors in the lower tail than in the upper tail. Basically speaking, our results indicate that during the crisis period, the degree of financial market heterogeneity after controlling the input-output linkages is relatively low and stable, as there are only one or two common factors driving the co-movement in the market across the quantiles except for $\tau = 0.95$, while during non-crisis period (Period III), there is a higher degree of market heterogeneity at the upper tail.

Overall, the common factor structures that explain the asset return distributions vary across quantiles and across periods. In addition, the common factor structures are not symmetric as their structures are different between the lower and upper tails.

7.3.2 The production network and decomposition of effects

The weighted average term on the right hand side of (15) captures the input-output linkages, where the parameter $\rho_{i,\tau}$ indicates the strength of such spillover effects or the interconnectedness of the market. Table 7 summarizes the average of the estimated value of the spatial parameter ($\hat{\rho}_{avg,\tau} = \frac{1}{N}\sum_{i=1}^{N}\hat{\rho}_{i,\tau}$) for different periods and different quantiles, with standard errors computed from the bootstrap method. Specifically, we consider two bootstrap approaches: (a) The bootstrapped residual at time t is drawn from the pool $(\hat{\varepsilon}_1, \hat{\varepsilon}_2, \cdots, \hat{\varepsilon}_T)$ where we treat the N-dimensional vector, $\hat{\varepsilon}_t = (\hat{\varepsilon}_{1t}, \cdots, \hat{\varepsilon}_{Nt})'$, as one point; (b) The next b-period bootstrapped residual at time t is drawn from the same pool by treating $N \times b$ residuals matrix $\hat{\varepsilon}_t^* = (\hat{\varepsilon}_t, \hat{\varepsilon}_{t+1}, \cdots, \hat{\varepsilon}_{t+b-1})$ as one point. Following Hou et al. (2021), we set $b = 3$. The standard error results are similar under the above two approaches. Therefore, we report the results under approach (a) in Table 7 in the main text, and present the results under approach (b) in Table 3 in the appendix.

From the table, we can see that, overall, the magnitude of spillover effects decreases over time (as $\hat{\rho}_{avg,\tau}$ decreases over periods from Period I to Period III). This indicates that the spillover effects are stronger during the U.S. subprime mortgage crisis and weaker during the recent period of monetary policy tightening. Across quantiles, the magnitude of spillover effects is generally higher at the tails than at the median with more pronounced disparity in Periods II and III than Period I. This implies that the market is more interconnected during crisis periods, and is also more interconnected at the tails than at the median. Our finding of the declining spillover effects is consistent with the existing findings in the related literature. For example, Diebold and Yilmaz (2014) show that the interconnectedness or spillover effect in major U.S. financial institutions' stock return volatilities is higher during the crisis periods, whereas they measure interconnectedness in terms of the variance decomposition of stock return volatilities.

To further analyze the production network interaction, we decompose the total effect of interest rate changes on stock returns into direct effects and network effects. The three effects are defined in the same spirit of those in LeSage and Pace (2009), Li (2017) and Ozdagli and Weber (2020). Here are details. Let $Q_{y_{it}} = Q_{y_{it}}(\tau|\boldsymbol{Q}_{(-i)t}(\tau), \boldsymbol{x}_{it}, \rho_{i,\tau}, b_{i,\tau}, \boldsymbol{f}_{t,\tau}, \boldsymbol{\lambda}_{i,\tau})$ and $Q_{y_t} = (Q_{y_{1t}}, Q_{y_{2t}}, \cdots, Q_{y_{Nt}})'$. Model (15) can be rewritten as:

$$Q_{y_t} = (I - \boldsymbol{\rho}_\tau W)^{-1}\alpha_\tau + (I - \boldsymbol{\rho}_\tau W)^{-1}\mathcal{B}_\tau * \mathbf{InterestRate}_t + (I - \boldsymbol{\rho}_\tau W)^{-1}\Lambda_\tau * \boldsymbol{f}_{t,\tau} \tag{16}$$

where I is an N-by-N identity matrix; $\boldsymbol{\alpha}_\tau = (\alpha_{1,\tau}, \alpha_{2,\tau}, \cdots, \alpha_{N,\tau})'$; $\Lambda_\tau = (\boldsymbol{\lambda}_{1,\tau}, \boldsymbol{\lambda}_{2,\tau}, \cdots, \boldsymbol{\lambda}_{N,\tau})'$; $\boldsymbol{\rho}_\tau = \text{diag}(\rho_{1,\tau}, \rho_{2,\tau}, \ldots, \rho_{N,\tau})$ is an N-by-N matrix with its i-th diagonal entry being $\rho_{i,\tau}$; similarly, $\mathcal{B}_\tau = \text{diag}(b_{1,\tau}, b_{2,\tau}, \cdots, b_{N,\tau})$ is an N-by-N matrix with its i-th diagonal entry being $b_{i,\tau}$; $\mathbf{InterestRate}_t$ is an N-by-1 vector with each element being $InterestRate_t$. From the above expression, we see that the partial derivative of the conditional quantile with respect to the interest rate change is $(I - \boldsymbol{\rho}_\tau W)^{-1}\mathcal{B}_\tau$, which is an N-by-N matrix. Let

$$Effect_\tau = (I - \boldsymbol{\rho}_\tau W)^{-1}\mathcal{B}_\tau.$$

The diagonal elements of *Effect* represent the direct effect (also including the feedback effect, where an industry's response to the interest rate can travel back to itself through the production network); the off-diagonal elements represent the network effect (including high order network effects). Then the average total effect(ATE), average direct effect (ADE) and average network effect (ANE) are defined as

$$\mathrm{ATE}_\tau = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N} \mathit{Effect}_{ij,\tau}, \quad \mathrm{ADE}_\tau = \frac{1}{N}\sum_{i=1}^{N} \mathit{Effect}_{ii,\tau}, \quad \mathrm{ANE}_\tau = \frac{1}{N}\sum_{i=1}^{N}\sum_{j\neq i}^{N} \mathit{Effect}_{ij,\tau}.$$

We compute the percentage of the average direct effect and the percentage of the average network effect, in order to compare their size. The results of this decomposition in both values and percentages are provided in Table 8 and also plotted in Figure 1. The standard errors reported in Table 8 are computed via the Monte Carlo method. More concretely, we first compute the standard errors for both $\hat{\rho}_{i,\tau}$ and $\hat{\boldsymbol{b}}_{i,\tau}$, then generate new $\rho^s_{i,\tau}$ and $\boldsymbol{b}^s_{i,\tau}$ based on their distributions given in Theorem 3 and calculate the corresponding total, direct and network effects. Finally we compute the standard errors for these three effects based on 1000 repetitions ($s = 1, 2, \cdots, 1000$).

There are three main findings from the results. First, the magnitude of the effects vary across periods. As shown in Figure 1, the total effect is higher in Periods I and II than in III, and its size is similar between Periods I and II. This is also true for the direct and network effects. Second, overall, the direct effects are higher than the network effects across all three periods, and their difference is smallest in Period I while biggest in Period III (as shown in Table 8, the percentage of the direct effect out of the total effect is highest in Period III while lowest in Period I). Third, across quantiles, the direct effect contributes more to the total effect at the upper tail than at the lower tail during Period III, but exhibits a different pattern in Period I. For comparison, Ozdagli and Weber (2020) focus on how the unexpected changes in the target federal funds rate affect stock prices, and find that, at the mean, there is a significant network effect which contributes more than 73 percent to the overall effect, based on a sample of scheduled FOMC meetings from 1994 to 2008.

7.3.3 Meaning of common factors

In factor models, the extracted common factors generally do not have an immediate economic interpretation. So we need an independent step to explore the economic meanings of the extracted common factors. In this subsection, we fulfill this work by regressing the estimated factors on some observable variables which are believed to explain stock market returns. We choose Fama-French North American 5 factors (hereafter, Fama-French 5 factors for short) and crude oil price return, since both have explanatory powers on stock market returns as documented in many studies.[5] For each period and each quantile, we regress the extracted factors on both Fama-French 5 Factors and crude oil price return:

$$\hat{f}_{jt,\tau} = \mathrm{FF}_t' * \delta_j + \mathrm{OilReturn}_t * \eta_j + e_{jt}, \qquad j = 1, 2, \cdots, \hat{r}_\tau \tag{17}$$

where $\hat{f}_{jt,\tau}$ is the estimated j-th common factor at time t and at quantile τ; $\hat{r}_\tau$ is the estimated number of common factors; $OilReturn_t$ is the crude oil price return at time t; $FF_t = (Mkt_t, HML_t, SMB_t, RMW_t, CMA_t)'$ is the vector of the Fama-French factors at time t. Since the Fama-French factors are well-known in the literature, we omit the detailed definition and refer readers to Fama and French (2015) for a complete description of these five factors. The historical data of these five factors can be downloaded from the publicly available Fama-French data library.[6]

Based on the above regression, we compute the ratio of the number of significant regression coefficients relative to the total Fama-French factors for all the $\hat{r}_\tau$ regressions. Likewise, we compute the ratio for the crude oil price return. The results are summarized in Table 9. From the results, we can see that for all periods, the estimated common factors are more significantly related to the $OilReturn$ than to the Fama-French factors. Regarding the $OilReturn$, at the lower tail, the estimated common factors are significantly related to it during all three periods, while at the upper tail, this relation is weaker in Periods II and III. Regarding the Fama-French five factors, different periods have different relation patterns, as in Period III,

[5] In the literature of asset pricing in finance, the Fama-French 5 factors are commonly used to explain the mean of returns on stocks (for references, see Fama and French (1993, 2015) and etc). Regarding oil price, many studies in the literature on analyzing stock market returns find, oil price changes have a significant impact on stock market returns (for references, see Kilian and Park (2009); Broadstock and Filis (2014); Ready (2018)). In this paper, we found that the extracted common factors are correlated with both the Fama-French NA 5 factors and oil price across different quantiles.

[6] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

the proportion of significant coefficients varies a lot more across quantiles than it does in Periods I and II. For illustrative purpose, we plot these estimated common factors together with Fama-French factors and *OilReturn* in Figure 7 in the supplement, where we can see both of them have a similar pattern of volatility starting in late 2008. Moreover, we can associate the latent common factors with some economic activities and market events (please see the last paragraph in Section E in the supplement for more details).

7.3.4 Sector-level classification

One advantage of the model (15) is its heterogeneous and quantile-dependent regression coefficient $b_{i,\tau}$, which allows us to study how the sensitivity of industry stock returns to yield changes varies across quantiles, and furthermore, how the sensitivity changes across industries. In order to see whether the sensitivity varies across quantiles and across industries, we plot both the point estimate and the 95% confidence interval for $\hat{b}_{i,\tau}$ (standard errors are computed based on the variance formula derived in Theorem 3) for a number of industries selected from different sectors, presented in Figures 5 and 6 in the supplement. The results indicate that $\hat{b}_{i,\tau}$ indeed varies across different quantiles.

In the following, we study how the sensitivity changes across industries and whether sector-level information could be helpful to explain it or not. So far, our analysis is based on detailed industry accounts at the 6-digit IO code level, which could further be classified into main sectors at the 2-digit IO code level, with a total of 20 sectors (Table 10 provides a list of sectors).[7] It is interesting to investigate the effect of sectors on the industry stock returns, by applying a clustering approach to both of the estimated regression coefficients and factor loadings. Our clustering approach aims to assign membership based on the similarities in the sensitivity to the common factors (both observable and unobservable factors). The sensitivity to the observable factors ($InterestRate_t$) is captured by $\hat{b}_{i,\tau}$ and the sensitivity to the unobservable factors (i.e., the extracted common factors $\hat{f}_{t,\tau}$) is captured by $\hat{\lambda}_{i,\tau}$. Note that the industry-to-sector classification is known, so we can create a two-way table of the assigned membership against the industry-to-sector classification.

More specifically, because of the resultant 20 sectors, we set the number of groups as 20,

[7]Table 2 in the online supplementary document presents a list of industries and their corresponding sectors.

and do clustering based on $\{(\hat{b}'_{i,\tau}, \hat{\lambda}'_{i,\tau}); i = 1, 2, \cdots, N\}$, for each period and each quantile. The results for quantiles $\tau = 0.05$ and 0.25 are shown in Figure 2. The row presents sector and the column represents the assigned group. An (i, j)-th element denotes the percentage of industries included in sector i such that they are assigned to the j-th group. Each row denotes the distribution of industries included in one particular sector that belong to different assigned groups. For example, from those tables, we can see that at the sector level, "DURABLE GOODS" and "NONDURABLE GOODS" are similar to each other in terms of the sensitivity to the common factors. Moreover, the similarity is stable over different periods and at $\tau = 0.05$ and $\tau = 0.95$ quantiles.

Theoretically, if the source of the sensitivity to the common factors is from sector level classification, it is expected that the two-way table of the assigned group membership from the clustering approach against the sector classification would be diagonal. However, as shown in Figure 2, there is not enough evidence to support the diagonality. Similar results are found for other quantiles $\tau = 0.25, 0.50$ and 0.75. This indicates that the observable sector level classification characteristic is insufficient to explain the heterogeneous sensitivities to the factors.

7.3.5 Concluding discussions

We are now in a position to answer the five questions proposed in the introduction. For the first question, our analysis shows that there are significant spillover effects in the U.S. stock market, and their strength is decreasing over periods and is slightly higher at the tails than at the median. For the second question, by decomposing the total effect of the interest rate change on the stock market into direct and network effects, we find that the magnitude of effects is higher in the crisis period, and the direct effect generally contributes more to the total effect over periods, while the composition pattern across quantiles is different from that over periods. For the third question, our model selection criterion results indicate that overall, the number of detected common factors differs across quantiles and periods. Moreover, the quantile factor structures are not symmetric between the lower and upper tails. For the fourth question, our analysis indicates that the extracted common factors are more related to the crude oil price return than the well-known Fama-French factors. For the last question, we implement a clustering approach based on the similarity of the sensitivity. Our results imply that the sector level classification is not sufficient in explaining the sensitivity of the stock market to the common factors, concluding that a diversified investment strategy only based on the sector classification is inadequate.

8 Conclusion

In this paper, we introduce a new spatial panel quantile model with interactive fixed effects. The model allows us to accommodate three interesting features simultaneously, that is spatial effects (spillover effects), heterogeneous regression quantile coefficients and unobservable common factors that may vary across quantiles.

To estimate the model, a new parameter estimation procedure and a new model selection criterion are proposed. We establish the asymptotic properties of our estimator including the consistency, convergence rates and limiting distributions. Moreover, we prove that the proposed model selection criterion can estimate the dimension of interactive fixed effects with probability approaching one. We apply the proposed model and estimation method to the U.S. stock market. Some new findings are uncovered by our new model.

There are still some open problems. First, this paper considers the quantile function at a particular τ. If the entire quantile function is considered instead, more strict restrictions are required to ensure monotonicity, the tools to investigate the asymptotic theory in this general case are still lacking and need to be developed. Second, we assume that the weight matrix W is exogenous. But it is possible that W is endogenous in real data applications. Further efforts should be made to address this concern. We would like to investigate these issues in the future.

References

Anand, K., et al. (2018). "The missing links: A global study on uncovering financial network structures from partial data," *Journal of Financial Stability*, **35**, 107–119.

Acemoglu, D., Akcigit, U., and Kerr, W. (2016). "Networks and the macroeconomy: an empirical exploration," *NBER Macroeconomics Annual*, **30(1)**, 273–353.

Acemoglu, D., Carvalho, V.M., Ozdaglar, A. and Tahbaz-Salehi, A. (2012). "The network origins of aggregate fluctuations," *Econometrica*, **80(5)**, 1977–2016.

Ahern, K.R., and Harford, J. (2014). "The importance of industry links in merger waves," *Journal of Finance*, **69(2)**, 527–576.

Ando, T. and Bai, J. (2017). "Clustering huge number of time series: A panel data approach with high-dimensional predictors and factor structures," *Journal of the American Statistical Association*, **112**, 1182–1198.

Ando, T. and Bai, J. (2020). "Quantile co-movement in financial markets: A panel quantile model with unobserved heterogeneity," *Journal of the American Statistical Association*, **115**, 266–279.

Aquaro, M., Bailey, N., and Pesaran, M. H. (2021). Estimation and inference for spatial models with heterogeneous coefficients: an application to US house prices. *Journal of Applied Econometrics*, 36(1), 18-44.

Bai, J. (2009). "Panel data models with interactive fixed effects," *Econometrica*, **77**, 1229–1279.

Bai, J. and Li, K. (2012). "Statistical analysis of factor models of high dimension," *The Annals of Statistics*, **40(1)**, 436–465.

Bai, J. and Li, K. (2013). "Spatial panel data models with common shocks," *manuscript*.

Bai, J. and Li, K. (2021). "Dynamic spatial panel data models with common shocks," Forthcoming in *Journal of Econometrics*.

Bai, J. and Liao, Y. (2016). "Efficient estimation of approximate factor models via penalized maximum likelihood," *Journal of Econometrics*, **191(1)**, 1–18.

Bai, J. and Ng, S. (2002). "Determining the number of factors in approximate factor models," *Econometrica*, **70**, 191–221.

Bai, J. and Ng, S. (2013). "Principal components estimation and identification of static factors," *Journal of Econometrics*, **176**, 18–29.

Baltagi, B.H. (2011). "Spatial Panels," Chapter 15 in The Handbook of Empirical Economics and Finance, Aman Ullah and David E.A. Giles, editors, Chapman and Hall, pp.435–454.

Barrot, J.N. and Sauvagnat, J. (2016). "Input specificity and the propagation of idiosyncratic shocks in production networks," *Quarterly Journal of Economics*, **131(3)**, 1543–1592.

Belloni, A., and Chernozhukov, V. (2011). "ℓ_1-penalized quantile regression in high-dimensional sparse models," *Annals of Statistics*, **39**, 82–130.

Bernanke, B.S. and Kuttner, K.N. (2005). "What explains the stock market's reaction to Federal Reserve policy?" *Journal of Finance*, **60(3)**, 1221–1257.

Bondell, H.D., Reich, B.J. and Wang, H. (2010). "Noncrossing quantile regression curve estimation," *Biometrika*, **97**, 825–838.

Broadstock, D.C., and Filis, G. (2014) "Oil price shocks and stock market returns: New evidence from the United States and China," *Journal of International Financial Markets, Institutions and Money*, **33**, 417–433.

Chen, L., Gonalo, J. and Dolado, J. (2021), Quantile factor models, *Econometrica*, **89**, 875–910.

Chernozhukov, V., Fernández-Val, I. and Galichon, A. (2010). "Quantile and Probability Curves Without Crossing," *Econometrica*, **78**, 1093–1125.

Chernozhukov, V., and Hansen, C. (2006) "Instrumental quantile regression inference for structural and treatment effect models," *Journal of Econometrics*, **127**, 491–525.

Cohen-Cole, E., Liu, X. and Zenou, Y. (2018). "Multivariate choice and identification of social interactions," *Journal of Applied Econometrics*, **33(2)**, 165–178.

Dette, H. and Volgushev, S. (2008). "Non-crossing nonparametric estimates of quantile curves," *Journal of the Royal Statistical Society*, **B70**, 609–627.

Diebold, F.X., and Yilmaz, K. (2014) "On the Network Topology of Variance Decompositions: Measuring the Connectedness of Financial Firms," *Journal of Econometrics*, **182**, 119–134.

Fama, E.F. and French, K.R. (1993). "Common factors in the returns on stocks and bonds," *Journal of Financial Economics*, **33**, 3–56.

Fama, E.F. and French, K.R. (2015). "A five-factor asset pricing model," *Journal of Financial Economics*, **116(1)**, 1–22.

Gabaix, X. (2011). "The granular origins of aggregate fluctuations," *Econometrica*, **79(3)**, 733–772.

Gorodnichenko, Y. and Weber, M.(2016). "Are stock prices costly? Evidence from the stock market," *American Economic Review*, **106(1)**, 165–199.

Griffin, J. M. (2002). "Are the Fama and French Factors Global or Country Specific," *Review of Financial Studies*, **15**, 783–803.

Hallin, M., and Liška, R. (2007) "The generalized dynamic factor model: determining the number of factors," *Journal of the American Statistical Association*, **102**, 603–617.

Harding, M., Lamarche, C., and Pesaran, M.H. (2020) "Common Correlated Effects Estimation of Heterogeneous Dynamic Panel Quantile Regression Models," *Journal of Applied Econometrics*, **35(3)**, 294–314.

He, X. (1997)"Quantile Curves Without Crossing," *American Statistician*, **51**, 186–192.

He, Y., Kong X., Wang X., Yu L. (2020), Factor Analysis without Moment Constraint, arXiv:2006.08214.

Hou, K., Karolyi, G.A., and Kho, B.-C. (2011). "What Factors Drive Global Stock Returns," *Review of Financial Studies*, **24**, 2527–2574.

Hou, L., Li, K., Li, Q. and Ouyang, M., (2021). Revisiting the location of FDI in China: A panel data approach with heterogeneous shocks. *Journal of Econometrics*, 221(2), 483–509.

Jeanty, P.W., Partridge, M., and Irwin, E. (2010). "Estimation of a spatial simultaneous equation model of population migration and housing price dynamics," *Regional Science and Urban Economics*, **40**, 343–352.

Kelejian, H. H., and Prucha, I. R. (2004). "Estimation of simultaneous systems of spatially interrelated cross sectional equations," *Journal of Econometrics*, **118**, 27–50.

Kilian, L., and Park, C. (2009) "The impact of oil price shocks on the U.S. stock market," *International Economic Review*, **50(4)**, 1267–1287.

Koopman, S. J., Schaumburg, J., and Wiersma, Q. (2021). Joint Modelling and Estimation of Global and Local Cross-Sectional Dependence in Large Panels. SSRN No.3770875.

Lee, L. (2004) "Asymptotic distributions of quasi-maximum likelihood estimator for spatial autoregressive models," *Econometrica*, **72(6)**, 1899–1925.

LeSage, J.P., and R.K., Pace (2009). "Introduction to spatial econometrics," Boca Raton, FL: CRC Press.

Li, K., (2017). Fixed-effects dynamic spatial panel data models and impulse response analysis. *Journal of Econometrics*, 198(1), pp.102-121.

Li, K., Cui, G. and Lu, L., (2020). Efficient estimation of heterogeneous coefficients in panel data models with common shocks. *Journal of Econometrics*, 216(2), pp.327-353.

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

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

Lu, L. (2017). "Simultaneous spatial panel data models with common shocks," *Federal Reserve Bank of Boston, RPA working paper*, RPA17–03.

Lu, L., and Luo, S. (2020). "The global effects of U.S. monetary policy on equity and bond markets: a spatial panel data model approach," *manuscript*.

Lu, X., and Su, L. (2016). "Shrinkage estimation of dynamic panel data models with interactive fixed effecs," *Journal of Econometrics*, **190(1)**, 148–175.

Luo, S. (2020). "Propagation of financial shocks in an input-output economy with trade and financial linkages of firms," *Review of Economic Dynamics*, **36**, 246–269.

Mantel, N. (1967). "The detection of disease clustering and a generalized regression approach," *Cancer Research*, **27(2**, 209–220. PMID 6018555.

Moon, H. R., and Weidner, M. (2015) "Linear regression for panel with unknown number of factors as interactive fixed effects," *Econometrica*, **83**, 1543–1579.

Nakamura, A., and Steinsson, J. (2018) "High frequency identification of monetary non-neutrality: the information effect," *Quarterly Journal of Economics*, **133(3)**, 1283–1330.

Ozdagli, A., and Weber, M. (2020). "Monetary policy through production networks: evidence from the stock market," *manuscript*.

Pasten, E., Schoenle, R., and Weber, W. (2020). "The propagation of monetary policy shocks in a heterogeneous production economy," forthcoming at *Journal of Monetary Economics*.

Pesaran, M.H. (2006) "Estimation and inference in large heterogeneous panels with a multifactor error structure," *Econometrica*, **74(4)**, 967–1012.

Qu, X. and Lee, L.F., (2015). Estimating a spatial autoregressive model with an endogenous spatial weight matrix. *Journal of Econometrics*, 184(2), pp.209-232.

Ready, R.C. (2018) "Oil prices and the stock market," *Review of Finance*, **22(1)**, 155–176.

Reich, B. J., M. Fuentes, and D. B. Dunson (2011). "Bayesian spatial quantile regression," *Journal of the American Statistical Association*, **106 (493)**, 6–20

Sherwood, B. and Wang, L. (2016). "Partially linear additive quantile regression in ultra-high dimension," *Annals of Statistics*, **44**, 288–317.

Shi, W. and Lee, L. (2017). "Spatial dynamic panel data models with interactive fixed effects," *Journal of Econometrics*, **197**, 323–347.

Stock, J.H., and Watson, M.W. (2002) "Forecasting using principal components from a large number of predictors," *Journal of the American Statistical Association*, **97**, 1167–1179.

Tang, Y., Song, X., Wang, H.J. and Zhu, Z. (2013). “Variable selection in high-dimensional quantile varying coefficient models,” *Journal of Multivariate Analysis*, **122**, 115–132.

Yu, J., de Jong, R., and Lee, L. (2008). “Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and T are large,” *Journal of Econometrics*, **146(1)**, 118–134.

Yuan, Y., Chen, N. and Zhou, S. (2017). “Modelling Regression Quantile Process using Monotone B-splines,” *Technometrics*, **59**, 338–350.

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.

Table 1: Simulation results under DGP1

			DGP1 (F is unknown)				DGP1 (F is known)			
τ	N	T	MSE_b	SE_b	MSE_ρ	SE_ρ	MSE_b	SE_b	MSE_ρ	SE_ρ
0.05	100	100	0.1678	0.0183	0.0109	0.0021	0.1545	0.0165	0.0094	0.0017
	200	100	0.1692	0.0135	0.0109	0.0016	0.1556	0.0096	0.0096	0.0014
	100	200	0.0889	0.0096	0.0085	0.0018	0.0757	0.0088	0.0070	0.0016
	200	200	0.0900	0.0086	0.0087	0.0013	0.0734	0.0070	0.0072	0.0015
	300	300	0.0641	0.0049	0.0077	0.0011	0.0568	0.0037	0.0065	0.0010
	300	400	0.0515	0.0039	0.0074	0.0009	0.0414	0.0031	0.0060	0.0007
	300	500	0.0439	0.0034	0.0072	0.0007	0.0371	0.0027	0.0055	0.0005
0.5	100	100	0.0770	0.0108	0.0103	0.0020	0.0640	0.0086	0.0079	0.0019
	200	100	0.0765	0.0071	0.0102	0.0016	0.0611	0.0063	0.0074	0.0013
	100	200	0.0452	0.0073	0.0087	0.0019	0.0347	0.0066	0.0070	0.0017
	200	200	0.0447	0.0050	0.0084	0.0012	0.0329	0.0041	0.0069	0.0011
	300	300	0.0338	0.0037	0.0078	0.0009	0.0259	0.0034	0.0066	0.0007
	300	400	0.0288	0.0035	0.0076	0.0007	0.0196	0.0031	0.0063	0.0004
	300	500	0.0246	0.0032	0.0072	0.0003	0.0192	0.0025	0.0059	0.0003
0.95	100	100	0.1634	0.0180	0.0135	0.0027	0.1443	0.0155	0.0112	0.0020
	200	100	0.1668	0.0144	0.0136	0.0019	0.1454	0.0115	0.0115	0.0017
	100	200	0.0855	0.0095	0.0084	0.0017	0.0604	0.0083	0.0076	0.0018
	200	200	0.0868	0.0083	0.0085	0.0014	0.0572	0.0063	0.0078	0.0013
	300	300	0.0624	0.0045	0.0076	0.0012	0.0432	0.0036	0.0068	0.0011
	300	400	0.0499	0.0038	0.0073	0.0010	0.0391	0.0029	0.0062	0.0009
	300	500	0.0429	0.0034	0.0071	0.0009	0.0325	0.0023	0.0056	0.0007

Note: The results in the left panel when F is known denote the estimation results of our proposed estimator when all parameters (including factors) are unknown and need to be estimated. With comparison, the results in the right panel when F is known denotes the infeasible estimation results of our proposed estimator when assuming the factors are observed and do not need to be estimated. MSE is the averaged mean squared error computed over 1000 repetitions, defined as in (14). SE represents the standard deviations of the mean squared errors.

Table 2: Simulation results under DGP2

			DGP2 (F is unknown)				DGP2 (F is known)			
τ	N	T	MSE_b	SE_b	MSE_ρ	SE_ρ	MSE_b	SE_b	MSE_ρ	SE_ρ
0.05	100	100	0.1666	0.0206	0.0116	0.0023	0.1591	0.0184	0.0105	0.0020
	200	100	0.1694	0.0143	0.0111	0.0013	0.1540	0.0121	0.0107	0.0018
	100	200	0.0886	0.0098	0.0094	0.0015	0.0800	0.0088	0.0082	0.0019
	200	200	0.0916	0.0086	0.0093	0.0014	0.0781	0.0078	0.0085	0.0014
	300	300	0.0655	0.0066	0.0088	0.0012	0.0589	0.0060	0.0078	0.0011
	300	400	0.0528	0.0050	0.0078	0.0013	0.0451	0.0045	0.0070	0.0009
	300	500	0.0455	0.0039	0.0074	0.0010	0.0432	0.0032	0.0065	0.0007
0.5	100	100	0.0829	0.0114	0.0129	0.0024	0.0682	0.0081	0.0110	0.0019
	200	100	0.0814	0.0083	0.0122	0.0017	0.0655	0.0065	0.0113	0.0017
	100	200	0.0479	0.0077	0.0117	0.0020	0.0400	0.0070	0.0095	0.0018
	200	200	0.0495	0.0062	0.0115	0.0016	0.0371	0.0051	0.0097	0.0015
	300	300	0.0393	0.0058	0.0111	0.0014	0.0301	0.0046	0.0090	0.0013
	300	400	0.0344	0.0055	0.0111	0.0014	0.0270	0.0040	0.0084	0.0011
	300	500	0.0313	0.0052	0.0102	0.0013	0.0267	0.0035	0.0079	0.0010
0.95	100	100	0.1706	0.0177	0.0159	0.0023	0.1664	0.0159	0.0135	0.0020
	200	100	0.1731	0.0137	0.0157	0.0018	0.1633	0.0128	0.0136	0.0017
	100	200	0.0885	0.0103	0.0135	0.0021	0.0769	0.0096	0.0110	0.0019
	200	200	0.0927	0.0070	0.0133	0.0016	0.0790	0.0062	0.0114	0.0015
	300	300	0.0694	0.0057	0.0126	0.0014	0.0576	0.0050	0.0098	0.0011
	300	400	0.0575	0.0054	0.0120	0.0013	0.0533	0.0049	0.0090	0.0009
	300	500	0.0504	0.0048	0.0115	0.0012	0.0463	0.0038	0.0084	0.0007

Note: The results in the left panel when F is known denote the estimation results of our proposed estimator when all parameters (including factors) are unknown and need to be estimated. With comparison, the results in the right panel when F is known denotes the infeasible estimation results of our proposed estimator when assuming the factors are observed and do not need to be estimated. MSE is the averaged mean squared error computed over 1000 repetitions, defined as in (14). SE represents the standard deviations of the mean squared errors.

Table 3: Simulation results for two other estimators (NoSpatial and NoFactor)

			DGP1			DGP2		
			NoSpatial	NoFactor		NoSpatial	NoFactor	
τ	N	T	MSE_b	MSE_b	MSE_ρ	MSE_b	MSE_b	MSE_ρ
0.05	100	100	0.2892	0.2658	0.0214	0.2970	0.2679	0.0215
	200	100	0.3037	0.2677	0.0209	0.2942	0.2729	0.0213
	100	200	0.1529	0.1593	0.0171	0.1543	0.1577	0.0173
	200	200	0.1552	0.1602	0.0167	0.1625	0.1593	0.0173
	300	300	0.1138	0.1146	0.0138	0.1197	0.1142	0.0146
	300	400	0.0939	0.1011	0.0126	0.0979	0.1032	0.0141
	300	500	0.0803	0.0839	0.0112	0.0856	0.0835	0.0127
0.5	100	100	0.1274	0.1772	0.0204	0.1285	0.1816	0.0231
	200	100	0.1307	0.1772	0.0200	0.1249	0.1831	0.0224
	100	200	0.0793	0.1155	0.0167	0.0815	0.1181	0.0195
	200	200	0.0818	0.1159	0.0167	0.0824	0.1192	0.0193
	300	300	0.0674	0.0946	0.0139	0.0687	0.0972	0.0168
	300	400	0.0595	0.0893	0.0127	0.0613	0.0931	0.0159
	300	500	0.0554	0.0746	0.0112	0.0576	0.0826	0.0155
0.95	100	100	0.2900	0.3641	0.0231	0.2966	0.3725	0.0261
	200	100	0.3114	0.3687	0.0234	0.3058	0.3787	0.0257
	100	200	0.1499	0.1875	0.0170	0.1578	0.1894	0.0218
	200	200	0.1569	0.1883	0.0165	0.1603	0.1926	0.0212
	300	300	0.1118	0.1428	0.0137	0.1185	0.1480	0.0191
	300	400	0.0907	0.1302	0.0125	0.0970	0.1362	0.0179
	300	500	0.0780	0.1023	0.0110	0.0852	0.1076	0.0165

Note: The results in the left panel when F is known denote the estimation results of our proposed estimator when all parameters (including factors) are unknown and need to be estimated. With comparison, the results in the right panel when F is known denotes the infeasible estimation results of our proposed estimator when assuming the factors are observed and do not need to be estimated. MSE is the averaged mean squared error computed over 1000 repetitions, defined as in (14).

Table 4: Simulation results for determining the number of common factors $\hat{r}_\tau$ under DGP1

		$\hat{r}_\tau$	0	1	2	3	4	5	6	7	R_{true}
τ	T	N									
0.05	100	100	0	47.6%	52.4%	0	0	0	0	0	1
	200	100	0	60.8%	39.2%	0	0	0	0	0	1
	100	200	0	69.0%	31.0%	0	0	0	0	0	1
	200	200	0	78.8%	21.2%	0	0	0	0	0	1
	300	300	0	85.2%	14.8%	0	0	0	0	0	1
	300	400	0	90.2%	9.8%	0	0	0	0	0	1
	300	500	0	95.4%	4.6%	0	0	0	0	0	1
0.50	100	100	0	100%	0	0	0	0	0	0	1
	200	100	0	100%	0	0	0	0	0	0	1
	100	200	0	100%	0	0	0	0	0	0	1
	200	200	0	100%	0	0	0	0	0	0	1
	300	300	0	100%	0	0	0	0	0	0	1
	300	400	0	100%	0	0	0	0	0	0	1
	300	500	0	100%	0	0	0	0	0	0	1
0.95	100	100	0	42.2%	57.8%	0	0	0	0	0	1
	200	100	0	45.8%	54.2%	0	0	0	0	0	1
	100	200	0	66.6%	33.4%	0	0	0	0	0	1
	200	200	0	77.4%	22.6%	0	0	0	0	0	1
	300	300	0	84.0%	16.0%	0	0	0	0	0	1
	300	400	0	89.4%	10.6%	0	0	0	0	0	1
	300	500	0	94.6%	5.4%	0	0	0	0	0	1

Note: This table reports the percentage of the selected number of common factors $\hat{r}_\tau$ under DGP1 over 1000 simulation runs. The results are for $\tau = 0.05$, $\tau = 0.50$ and $\tau = 0.95$ quantile points. The true number of common factors is one for all the three quantile points as indicated in the last column.

Table 5: Simulation results for determining the number of common factors $\hat{r}_\tau$ under DGP2

		$\hat{r}_\tau$	0	1	2	3	4	5	6	7	R_{true}
τ	T	N									
0.05	100	100	0	40.8%	59.2%	0	0	0	0	0	1
	200	100	0	50.6%	49.4%	0	0	0	0	0	1
	100	200	0	58.6%	41.4%	0	0	0	0	0	1
	200	200	0	71.0%	29.0%	0	0	0	0	0	1
	300	300	0	79.8%	20.2%	0	0	0	0	0	1
	300	400	0	86.2%	13.8%	0	0	0	0	0	1
	300	500	0	93.4%	6.6%	0	0	0	0	0	1
0.50	100	100	0	0	100%	0	0	0	0	0	2
	200	100	0	0	100%	0	0	0	0	0	2
	100	200	0	0	100%	0	0	0	0	0	2
	200	200	0	0	100%	0	0	0	0	0	2
	300	300	0	0	100%	0	0	0	0	0	2
	300	400	0	0	100%	0	0	0	0	0	2
	300	500	0	0	100%	0	0	0	0	0	2
0.95	100	100	0	0	1.6%	34.6%	63.8%	0	0	0	3
	200	100	0	0	0.8%	40.4%	58.8%	0	0	0	3
	100	200	0	0	0%	51.0%	49.0%	0	0	0	3
	200	200	0	0	0%	63.8%	36.2%	0	0	0	3
	300	300	0	0	0%	74.2%	25.8%	0	0	0	3
	300	400	0	0	0%	83.6%	16.4%	0	0	0	3
	300	500	0	0	0%	92.0%	8.0%	0	0	0	3

Note: This table reports the percentage of the selected number of common factors $\hat{r}_\tau$ under DGP2 over 1000 simulation runs. The results are for $\tau = 0.05$, $\tau = 0.50$ and $\tau = 0.95$ quantile points. The true number of common factors are $r_{(\tau=0.05)} = 1$, $r_{(\tau=0.50)} = 2$ and $r_{(\tau=0.95)} = 3$ as indicated in the last column.

Table 6: Estimated number of common factors $\hat{r}_\tau$

τ	Period I	Period II	Period III
0.05	2	1	1
0.25	2	1	1
0.50	1	2	5
0.75	2	2	4
0.95	3	5	6

Period I: 01/01/2007 - 04/30/2009, $N = 324$ and $T = 582$;
Period II: 09/01/2009 - 12/31/2012, $N = 323$ and $T = 831$;
Period III: 12/01/2015 - 06/30/2018, $N = 307$ and $T = 647$.

Table 7: Average of the estimated spatial parameter $\hat{\rho}_{avg,\tau} = \frac{1}{N}\sum_{i=1}^{N}\hat{\rho}_{i,\tau}$

τ	Period I	Period II	Period III
0.05	0.615	0.537	0.355
	(0.009)	(0.039)	(0.012)
0.25	0.609	0.532	0.353
	(0.009)	(0.015)	(0.036)
0.5	0.610	0.498	0.321
	(0.009)	(0.017)	(0.016)
0.75	0.612	0.555	0.358
	(0.007)	(0.016)	(0.031)
0.95	0.614	0.548	0.351
	(0.008)	(0.012)	(0.026)

Note: Standard errors are in parenthesis.
All estimates are significant at 1% level.
Period I: 01/01/2007 - 04/30/2009, $N = 324$ and $T = 582$;
Period II: 09/01/2009 - 12/31/2012, $N = 323$ and $T = 831$;
Period III: 12/01/2015 - 06/30/2018, $N = 307$ and $T = 647$.

Table 8: Effect Decomposition: Total, Direct and Network Effects

	(1)	(2)	(3)	(4)	(5)
		Effects in Value		Effects in Percentage	
	Total Effect	Direct Effect	Network Effect	Direct Effect	Network Effect
			Period I		
0.05	0.1211	0.0655	0.0556	54.08%	45.92%
	(0.0081)	(0.0046)	(0.0043)		
0.25	0.1056	0.0572	0.0484	54.20%	45.80%
	(0.0049)	(0.0028)	(0.0027)		
0.5	0.0964	0.0516	0.0448	53.50%	46.50%
	(0.0046)	(0.0025)	(0.0028)		
0.75	0.0982	0.0521	0.0461	53.03%	46.97%
	(0.0053)	(0.0030)	(0.0029)		
0.95	0.1050	0.0504	0.0546	48.30%	51.70%
	(0.0045)	(0.0027)	(0.0023)		
			Period II		
0.05	0.1096	0.0608	0.0489	55.43%	44.57%
	(0.0010)	(0.0007)	(0.0005)		
0.25	0.0984	0.0589	0.0395	59.88%	40.12%
	(0.0014)	(0.0009)	(0.0007)		
0.5	0.0918	0.0595	0.0323	64.80%	35.20%
	(0.0019)	(0.0009)	(0.0014)		
0.75	0.1076	0.0655	0.0421	60.90%	39.10%
	(0.0019)	(0.0012)	(0.0009)		
0.95	0.1017	0.0624	0.0393	61.31%	38.69%
	(0.0015)	(0.0009)	(0.0007)		
			Period III		
0.05	0.0770	0.0536	0.0234	69.56%	30.44%
	(0.0015)	(0.0011)	(0.0005)		
0.25	0.0731	0.0510	0.0221	69.80%	30.20%
	(0.0022)	(0.0015)	(0.0010)		
0.5	0.0682	0.0489	0.0194	71.59%	28.41%
	(0.0039)	(0.0022)	(0.0030)		
0.75	0.0778	0.0545	0.0233	70.02%	29.98%
	(0.0032)	(0.0021)	(0.0016)		
0.95	0.0746	0.0570	0.0176	76.45%	23.55%
	(0.0014)	(0.0012)	(0.0004)		

Note: Standard errors are in parenthesis. All the three effects are significant at 1% level. Period I: 01/01/2007 - 04/30/2009; Period II: 09/01/2009 - 12/31/2012; Period III: 12/01/2015 - 06/30/2018.

Table 9: Link the common factors to Fama-French 5 Factors and Oil Return

	(1)	(2)	(3)	(4)	(5)	(6)
	Period I		Period II		Period III	
τ	FF 5Factors	Oil Return	FF 5Factors	Oil Return	FF 5Factors	Oil Return
0.05	0.4	1	0.4	1	0.8	1
0.25	0.3	1	0.4	1	0.8	1
0.5	0.4	1	0.2	0.5	0.2	0.4
0.75	0.6	1	0.2	0.5	0.15	0.5
0.95	0.27	0.67	0.16	0.4	0.2	0.33

Note: See Section 7.3.3 for more details. Period I: 01/01/2007 - 04/30/2009; Period II: 09/01/2009 - 12/31/2012; Period III: 12/01/2015 - 06/30/2018.

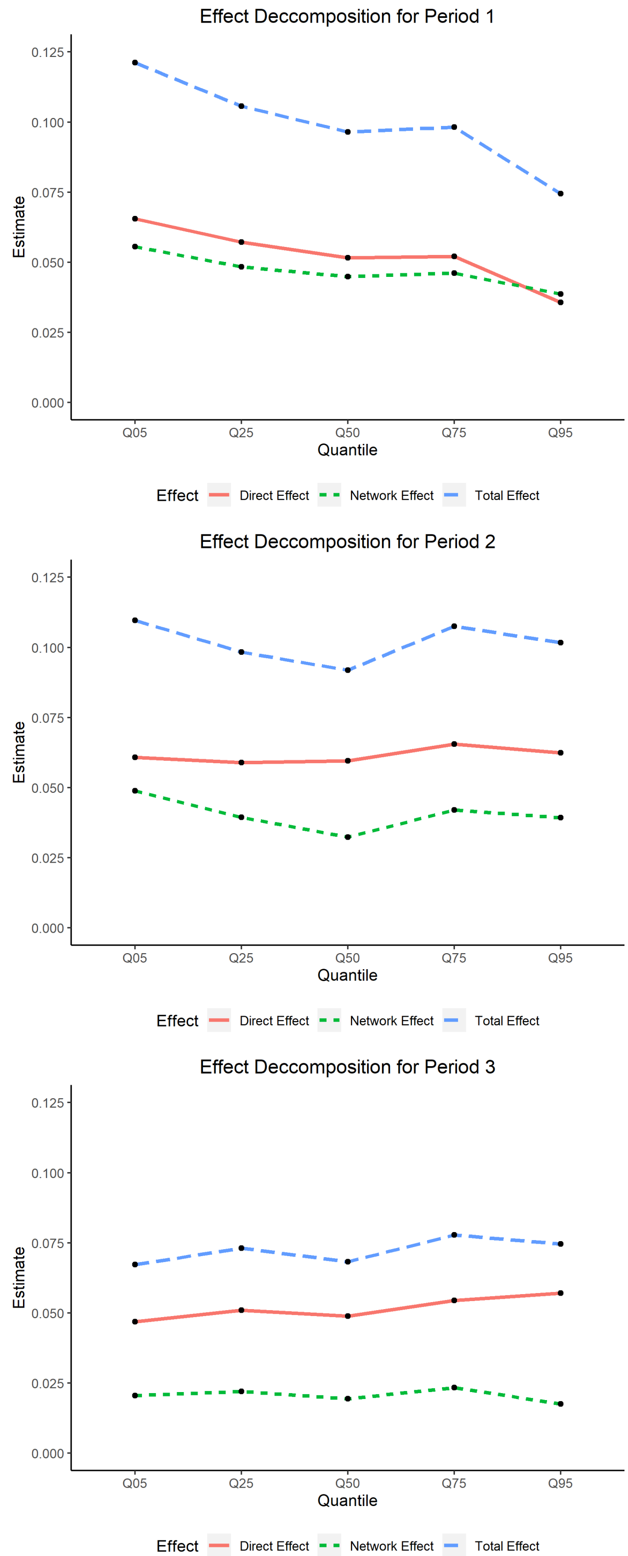

Figure 1: Effect Decomposition in Value

Note: See Section 7.3.3 and Table 8 for more details. Period I: 01/01/2007 - 04/30/2009; Period II: 09/01/2009 - 12/31/2012; Period III: 12/01/2015 - 06/30/2018.

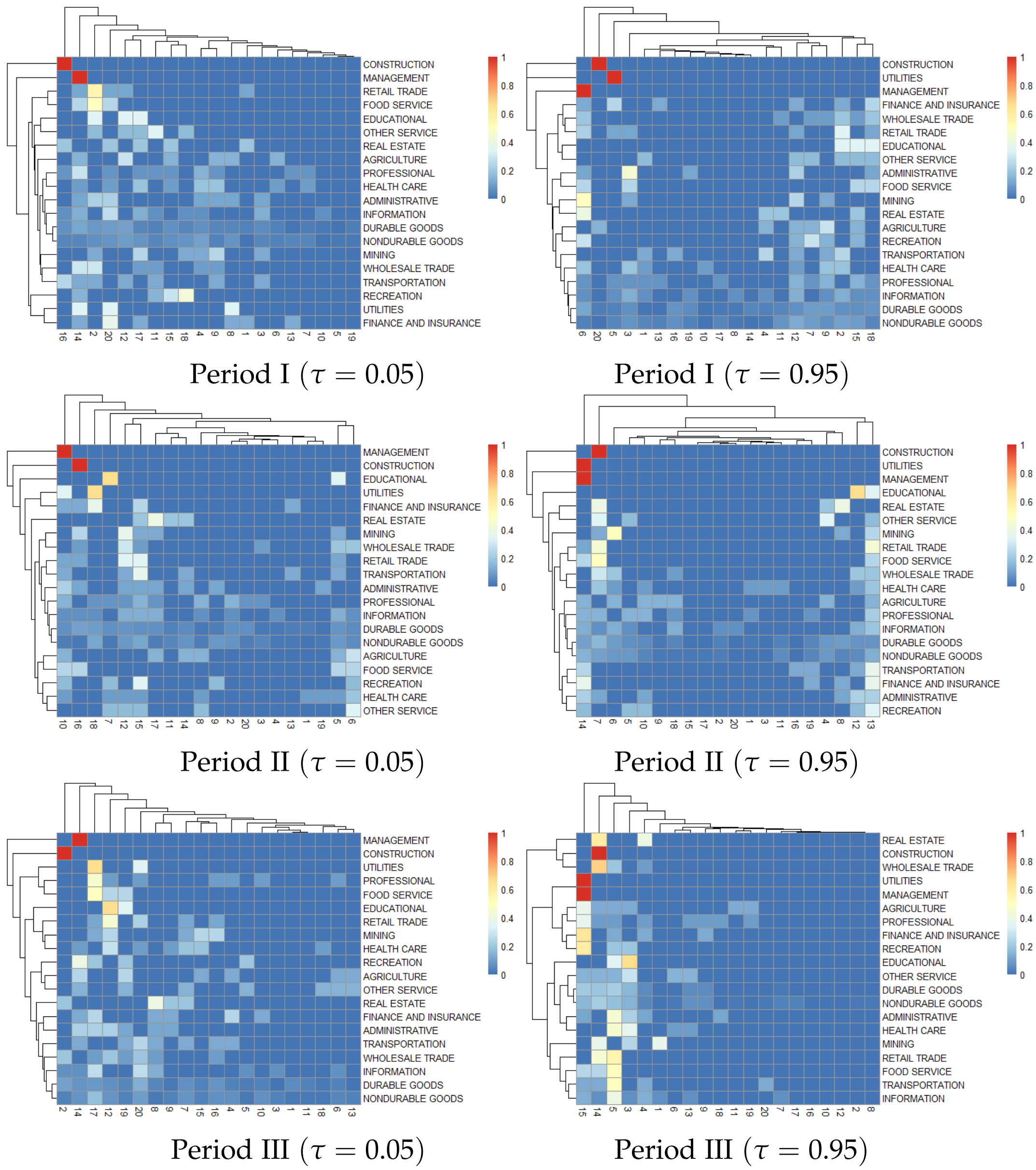

Figure 2: Distribution of detailed industries in each of the sectors

Note: See Section 7.3.4 for more details. An (i, j)-th element represents the percentage of industries in sector i such that they belong to the j-th group. Period I: 01/01/2007 - 04/30/2009; Period II: 09/01/2009 - 12/31/2012; Period III: 12/01/2015 - 06/30/2018.

Table 10: List of Sectors (their short names are used in Figure 2)

Sector ID	Sector Name	Sector Name Short
11	AGRICULTURE,FORESTRY, FISHING, AND HUNTING	AGRICULTURE
21	MINING	MINING
22	UTILITIES	UTILITIES
23	CONSTRUCTION	CONSTRUCTION
33DG	DURABLE GOODS	DURABLE GOODS
31ND	NONDURABLE GOODS	NONDURABLE GOODS
42	WHOLESALE TRADE	WHOLESALE TRADE
44RT	RETAIL TRADE	RETAIL TRADE
48TW	TRANSPORTATION AND WAREHOUSING, EXCLUDING POSTAL SERVICE	TRANSPORTATION
51	INFORMATION	INFORMATION
52	FINANCE AND INSURANCE	FINANCE AND INSURANCE
53	REAL ESTATE AND RENTAL AND LEASING	REAL ESTATE
54	PROFESSIONAL AND TECHNICAL SERVICES	PROFESSIONAL
55	MANAGEMENT OF COMPANIES AND ENTREPRISES	MANAGEMENT
56	ADMINISTRATIVE AND WASTE SERVICES	ADMINISTRATIVE
61	EDUCATIONAL SERVICES	EDUCATIONAL
62	HEALTH CARE AND SOCIAL ASSISTANCE	HEALTH CARE
71	ARTS, ENTERTAINMENT, AND RECREATION	RECREATION
72	ACCOMMODATION AND FOOD SERVICES	FOOD SERVICE
81	OTHER SERVICES, EXCEPT GOVERNMENT	OTHER SERVICE