Applications of Functional Dependence to Spatial Econometrics

Zeqi Wu[a], Wen Jiang[b], and Xingbai Xu[*c]

[a]Institute of Statistics & Big Data, Renmin University of China

[b]New Huadu Business School, Minjiang University

[c]MOE Key Lab of Econometrics and Fujian Key Lab of Statistics, Wang Yanan Institute for Studies in Economics (WISE), Department of Statistics and Data Science, School of Economics, Xiamen University

Abstract

In this paper, we generalize the concept of functional dependence from time series (Wu, 2005) and stationary random fields (El Machkouri, Volný and Wu, 2013) to nonstationary spatial processes. Within conventional settings in spatial econometrics, we define the concept of spatial functional dependence measure and establish a moment inequality, an exponential inequality, a Nagaev-type inequality, a law of large numbers, and a central limit theorem. We show that the dependent variables generated by some common spatial econometric models, including spatial autoregressive models, threshold spatial autoregressive models and spatial panel data models, are functionally dependent under regular conditions. Furthermore, we investigate the properties of functional dependence measures under various transformations, which are useful in applications. Moreover, we compare spatial functional dependence with the spatial mixing and

[*]We sincerely thank the editor Peter C.B. Phillips, the co-editor Liangjun Su, and three anonymous referees for their helpful comments and suggestions on the earlier version of this paper. Xu gratefully acknowledges the financial support from NSFC (NO.72073110, NO.72333001), Basic Scientific Center Project 71988101 of NSFC, and the 111 Project (B13028). Address correspondence to: Xingbai Xu, Wang Yanan Institute for Studies in Economics, Xiamen University, Xiamen, China; e-mail address: xuxingbai@xmu.edu.cn. Zeqi Wu and Wen Jiang are the co-first authors.

spatial near-epoch dependence proposed in Jenish and Prucha (2009, 2012), and we illustrate its advantages.

Key words: spatial functional dependence, near-epoch dependence, law of large numbers, central limit theorem, spatial autoregressive model, spatial dynamic panel data model

JEL: C10, C19, C21, C23

1. Introduction

In recent years, spatial econometric models have been widely applied to various fields of economics, e.g., agricultural economics, international trade, climate economics, and regional and urban economics. Accordingly, various spatial econometric models and estimation methods are investigated in the literature. To study asymptotic theories for estimators of spatial econometric models, some limiting laws and dependence concepts are indispensable. Early development of spatial econometrics, especially linear spatial models, has relied mainly on the theories of linear-quadratic forms of independent variables. See, Kelejian and Prucha (1998, 2001), Lee (2004, 2007), and Yu, de Jong and Lee (2008), among many others. However, these theories are not applicable to some recent development in spatial econometrics, e.g., the spatial panel data model with endogenous spatial weights matrix (Qu, Lee and Yu, 2017), robust estimators (Liu, Xu, Lee and Mei, 2022), quantile estimators (Xu, Wang, Shin and Zheng, 2022), and nonlinear spatial econometric models (Xu and Lee, 2015a,b). In these papers, the authors employ weak spatial dependence concepts like spatial strong mixing or spatial near-epoch dependence (NED). Strong mixing and NED are widely used in time series (Davidson, 1994; Doukhan, 1994) and stationary random fields (Bolthausen, 1982; Dedecker, 1998)[1], and they are generalized to spatial econometric settings by Jenish and Prucha (2009, 2012).

However, strong mixing and NED have certain shortcomings. The strong mixing coefficient involves the calculation of supremum over two σ-fields and hence is quite complicated and inconvenient (Doukhan and Louhichi, 1999; Wu, 2005; Xu and Lee, 2024). Moreover, even some AR(1) processes do not satisfy the strong mixing condition (Andrews, 1984; Wu, 2005). For NED, its application is mainly restricted to L^2-NED, as L^p-NED ($p \neq 2$) is usually not easy to establish; and in some cases, some strong moment conditions are needed to preserve NED properties. Therefore,

[1]A random field $Y : \mathbb{R}^d(\text{ or } \mathbb{Z}^d) \to \mathbb{R}^{p_Y}$ is stationary means that the joint distribution of $(Y_{s_1}, Y_{s_2}, \dots, Y_{s_t})$ does not change under the translation of $(s_1, \dots, s_t)$, i.e., the joint probability density (or mass) function $f(Y_{s_1}, Y_{s_2}, \dots, Y_{s_t}) = f(Y_{s_1+r}, Y_{s_2+r}, \dots, Y_{s_t+r})$ for any $(s_1, \dots, s_t)$ and $r \in \mathbb{R}^d(\text{ or } \mathbb{Z}^d)$.

we aim to find a better notion of weak spatial dependence.

In Wu (2005), the concept of functional dependence (FD), also called physical dependence, is proposed. It is often easy to verify and has many good properties. Based on this concept, Liu, Xiao and Wu (2013) and Wu and Wu (2016) establish the Nagaev-type, Rosenthal-type, and exponential inequalities. El Machkouri et al. (2013) generalize the functional dependence from time series to stationary random fields located in $\mathbb{Z}^d$ and study its limit theorems. Functional dependence has been widely used in statistics to establish asymptotic theories of various statistics (Chen, Xu and Wu, 2013; Wu, 2011; Wu and Wu, 2016; Zhou and Wu, 2009).

However, the theory of FD on stationary random fields in $\mathbb{Z}^d$ in El Machkouri et al. (2013) does not apply to spatial econometrics directly. The reasons for this are two-fold: (1) the spatial units are located in $\mathbb{Z}^d$, which is seldom the setting in spatial econometrics; (2) the data-generating process is supposed to be homogeneous and the spatial process is required to be stationary. On the contrary, in spatial econometrics, the spatial units are usually unevenly spaced, and the spatial random variables are often nonstationary and heterogeneous triangular arrays. To fill this gap, we generalize the spatial functional dependence in El Machkouri et al. (2013). We allow (1) the spatial units to be located in an unevenly spaced lattice, (2) the spatial process to be nonstationary, and (3) the random variables to be a heterogeneous triangular array. Based on the spatial functional dependence measure (FDM), we establish a moment inequality, an exponential inequality, a Nagaev-type inequality, a law of large numbers (LLN), and a central limit theorem (CLT) that are sufficiently general to accommodate more applications of interest. We want to emphasize that the generalization from El Machkouri et al. (2013) to our paper is not trivial, because the techniques in their proofs do not apply to our setup due to heterogeneity and nonstationarity.

Our FDM concept overcomes the shortcomings of mixing and NED. (1) It is easy to calculate for many spatial econometric models, as it does not involve σ-field or conditional expectation. For convenience, Su, Wang and Xu (2023) apply the theory of spatial FDM developed in this paper to study a heterogeneous spatial dynamic panel data model. (2) Compared to NED, it can be conveniently established under L^p-norm for any $p \geq 1$, and is, therefore, more flexible, especially under nonlinear transformations. (3) Compared to those needed for NED, weaker conditions suffice for a CLT and an exponential inequality via spatial FDM.

This paper is organized as follows. In Section 2, we present the definitions of spatial FDM and spatial FD coefficient. In Section 3, we investigate their theoretical properties, including some inequalities, an LLN, a CLT, and a heteroskedasticity and autocorrelation consistent estimator for the variance term in the CLT. In Section 4, we calculate the FDM of a nonlinear spatial autoregressive (SAR) model, a threshold SAR model and a spatial panel data model. In Section 5, we investigate the properties of spatial FDM and the spatial FD coefficient under various common transformations. In Section 6, we compare spatial FDM with NED. Section 7 concludes this paper. The proofs for the LLN and the CLT are collected in the appendices, and all other proofs are provided in the supplementary material. All sections, lemmas and equations whose numberings begin with "S" (e.g., Lemma S.3) are in the supplementary material.

Notation: The set of positive integers is denoted by $\mathbb{N} \equiv \{1, 2, \ldots\}$. For any column vector $x = (x_1, x_2, \ldots, x_d)' \in \mathbb{R}^d$, where $\mathbb{R}^d$ is the d-dimensional Euclidean space, $\|x\| = (x'x)^{1/2}$ denotes its Euclidean norm, $\|x\|_\infty = \max_{1\leq k\leq d} |x_k|$ represents its infinity vector norm, and $\|x\|_1 = \sum_{k=1}^{d} |x_k|$ denotes its 1-norm. For any random vector $X \in \mathbb{R}^d$, its L^p-norm is defined as $\|X\|_{L^p} \equiv [\mathbb{E}(\|X\|^p)]^{1/p}$. For any square matrix $A = (a_{ij})_{n\times n}$, its maximum row sum norm is defined as $\|A\|_\infty = \max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij}|$, $A_{i\cdot}$ denotes its ith row, and $|A|$ is defined as $|A| = (|a_{ij}|)_{n\times n}$. For any real number a, $\lfloor a \rfloor$ denotes its integer part, i.e., $\lfloor a \rfloor = \max\{b \in \mathbb{Z} : b \leq a\}$, and $\lceil a \rceil \equiv \min\{b \in \mathbb{Z} : b \geq a\}$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. For any sub-σ-field $\mathcal{C}$ of $\mathcal{F}$, we write $\mathbb{P}_{\mathcal{C}}(\cdot) \equiv \mathbb{P}(\cdot \mid \mathcal{C})$, $\mathbb{E}_{\mathcal{C}}(\cdot) \equiv \mathbb{E}(\cdot \mid \mathcal{C})$, and $\mathrm{Var}_{\mathcal{C}}(\cdot) \equiv \mathrm{Var}(\cdot|\mathcal{C})$. For a random vector X, let $\|X\|_{L^p,\mathcal{C}} = [\mathbb{E}_{\mathcal{C}}(\|X\|^p)]^{1/p}$. Let $\xrightarrow{p}$, $\xrightarrow{L^p}$, $\xrightarrow{d}$, and $\xrightarrow{a.s.}$ denote convergence in probability, L^p convergence, convergence in distribution, and convergence almost surely, respectively. For any set D, $|D|$ denotes its cardinality. For any two nonnegative functions $f(x)$ and $g(x)$ defined on $[0, \infty)$, $f(x) = O(g(x))$ as $x \to \infty$ means that there exist constants $M > 0$ and x_0 such that $f(x) \leq Mg(x)$ whenever

$x \geq x_0$ and $f(x) < \infty$ for all $x \in [0, \infty)$[2]. For two sequences $\{a_n\}$ and $\{b_n\}$, $a_n \sim b_n$ if and only if (iff) $\lim_{n\to\infty} \frac{a_n}{b_n} = 1$.

2. Spatial Functional Dependence Measure and Spatial Functional Dependence Coefficient

In this section, we define the spatial FDM and the spatial functional dependence coefficient. First, we introduce some notation. Suppose there are some individuals (e.g., persons, cities, countries), also called spatial units in this paper, located in a lattice $D_n \subset \mathbb{R}^d$. Here, D_n can either be a finite set whose cardinality $|D_n| = n$, or be a countably infinite set. We focus on two settings: (1) for cross-sectional data with n individuals, $|D_n| = n$; (2) for spatial panel data, each spatial unit i is located in $\bar{D}_n \subset \mathbb{R}^d$, but we regard (i,t), the combination of spatial unit i and time t, as a point in $\mathbb{R}^{d+1}$ and define $D_n \equiv \bar{D}_n \times \{\ldots, T-1, T\} \equiv \{(i,t) : i \in \bar{D}_n, t \leq T\} \subset \mathbb{R}^{d+1}$. In setting (2), $|D_n| = \infty$. For any two individuals $i = (i_1, \ldots, i_d)$ and $j = (j_1, \ldots, j_d)$ in $\mathbb{R}^d$, $d_{ij} \equiv \max_{1\leq k\leq d} |i_k - j_k|$ denotes their distance.

Let $\epsilon = \{\epsilon_{i,n}, i \in D_n, n \geq 1\}$ be an $\mathbb{R}^{p_\epsilon}$-valued independent random field. Another random field $Y = \{Y_{i,n}, i \in D_n, n \geq 1\}$ is generated by

$$Y_{i,n} = g_{i,n}(\epsilon_n), \tag{2.1}$$

where $\{g_{i,n}, i \in D_n, n \geq 1\}$ is a set of $\mathbb{R}^{p_Y}$-valued Borel-measurable functions and $\epsilon_n = \left(\left(\epsilon'_{i,n}\right)_{i\in D_n}\right)'$. In some models, e.g., a linear SAR model, the explicit functional form of $g_{i,n}(\cdot)$ is known. However, in many nonlinear spatial econometric models, e.g., the SAR Tobit model in Xu and Lee (2015a), we do not know the explicit functional form of $g_{i,n}(\cdot)$, but it does not affect our analysis. See

[2] In standard big O notation, usually it is not required that $f(x) < \infty$, but we impose this for the convenience of our presentation. With this definition, we can safely claim that (i) $\sup_{x\in A} g(x)^{-1} O(g(x)) < \infty$ for any closed set $A \subset \{y \in [0,\infty) : g(y) > 0\}$, (ii) $\sum_{m=1}^{\infty} O(m^{-\delta}) < \infty$ for any $\delta > 1$, and (iii) $f(x) = O(x^{-\alpha})$ for some $\alpha > 0$ implies $f(x) \leq C(x+1)^{-\alpha}$ for any $x \in [0,\infty)$, where $C > 0$ is a constant not depending on x.

Section 4.2 for more details.

Let $\left(\left((\epsilon^*_{i,n})'\right)_{i\in D_n}\right)'$ be an independently and identically distributed (i.i.d.) copy of ϵ_n. For any set $I \subset D_n$, we define $\epsilon_{i,n,I} \equiv \epsilon^*_{i,n}$ if $i \in I$ and $\epsilon_{i,n,I} \equiv \epsilon_{i,n}$ if $i \notin I$; we write $\epsilon_{n,I} \equiv \left(\left(\epsilon'_{i,n,I}\right)_{i\in D_n}\right)'$. Furthermore, $Y_{i,n,I} = g_{i,n}\left(\epsilon_{n,I}\right)$ is called a coupled version of $Y_{i,n}$ on I and $Y_{n,I} = \left(\left(Y'_{i,n,I}\right)_{i\in D_n}\right)'$. All our discussion in Sections 2, 3, and 5 is based on (2.1).

Throughout the paper, we maintain these conventions on notation and the following assumption concerning the lattice D_n.

Assumption 1. *For all $i \neq j \in D_n$, $d_{ij} \geq 1$.*

Assumption 1 employs the increasing-domain asymptotics and rules out the scenario of infilled asymptotics (also called fixed domain asymptotics), and it is commonly used in the spatial econometrics literature (Jenish and Prucha, 2009, 2012; Liu et al., 2022; Qu and Lee, 2015; Xu and Lee, 2015a,b, 2018; Xu et al., 2022). As n increases to infinity, the diameter of D_n also tends to infinity. When we consider the geographical distance between two cities, Assumption 1 means that the centers of cities cannot be too close to each other. Although the diameter of the earth is finite, which restricts the diameter of D_n, when we apply our theory, we expect that the sample size is large enough such that our asymptotic theory can approximate well. In addition, the distance might be "social-economic distance"[3], which means that some coordinates might be economic or social characteristics of the spatial units. So even if the geographical distance between two individuals is small, their social-economic distance might be large. In spatial statistics and computer science, sometimes researchers consider infill asymptotics, e.g., when we study the image of someone's brain, but we do not consider infill asymptotics in this paper.

Now, we are ready to introduce our main concepts.

[3]See Conley and Topa (2002) and the paragraph below Assumption 1 in Qu and Lee (2015) for some discussion about social-economic distance.

2.1. Spatial Functional Dependence Measure

Definition 2.1. *For any $p \geq 1$, $n \geq 1$ and $I \subset D_n$, we define the (spatial) L^p-functional dependence measure as*

$$\delta_p(i, I, n) \equiv \|Y_{i,n} - Y_{i,n,I}\|_{L^p} = \|g_{i,n}(\epsilon_n) - g_{i,n}(\epsilon_{n,I})\|_{L^p}.$$

When $I = \{j\}$ is a singleton, we simplify the notation as $\delta_p(i, j, n) \equiv \delta_p(i, \{j\}, n)$.

When $I = \emptyset$, $\delta_p(i, \emptyset, n) = 0$, which causes no conflict. Definition 2.1 is a generalization of those in Wu (2005) and El Machkouri et al. (2013). The differences lie in three aspects. First, the index sets in Wu (2005) and El Machkouri et al. (2013) are respectively $\mathbb{Z}$ and $\mathbb{Z}^d$. Instead, we consider an unevenly spaced lattice D_n in $\mathbb{R}^d$, which is in line with the paradigm of spatial econometrics. Second, Wu (2005) and El Machkouri et al. (2013) require the nonlinear transformation g to be invariant over i and n, which is ruled out by almost all spatial econometric models, but we allow different $g_{i,n}$ for different i and n. Third, Wu (2005) and El Machkouri et al. (2013) set $\epsilon_{i,n}$'s to be i.i.d., but we allow the $\epsilon_{i,n}$'s to be non-identically distributed. Thus, $\{Y_{i,n}, i \in D_n\}$ might be nonstationary and heterogeneous in our setup.

Spatial statistics usually focuses on Gaussian processes and many results are based on correlation or covariance functions, and FDM is closely related to the correlation or covariance of $Y_{i,n}$ and $Y_{i,n,I}$.[4] Consider a linear process $Y_{i,n} = \sum_{j \in D_n} A_{ij,n}\epsilon_{j,n}$, where $|D_n| = n$, $A_{ij,n}$'s are constant and $\epsilon_{j,n}$'s are i.i.d. with expectation zero and unit variance. For any set $I \subset D_n$, direct calculations show that $\frac{1}{2}\delta_2(i, I, n)^2 + \mathrm{Cov}(Y_{i,n}, Y_{i,n,I}) = \mathrm{Var}(Y_{i,n})$. If we consider the case where $\mathrm{Var}(Y_{i,n}) = 1$, i.e., $\sum_{j \in D_n} A_{ij,n}^2 = 1$, the previous relationship becomes $\mathrm{Corr}(Y_{i,n}, Y_{i,n,I}) = 1 - \frac{1}{2}\delta_2(i, I, n)^2$. In addition, the FD property is also related to the covariance between two different $Y_{i,n}$'s and the estimation of the asymptotic variance of $n^{-1/2}\sum_{i \in D_n}(Y_{i,n} - \mathbb{E}Y_{i,n})$.[5] Since FDM does not require

[4] We thank an anonymous referee for sharing his/her deep insight into both the relationship between FDM and the correlation (or covariance) functions and the possible applications of FDM in studying non-Gaussian spatial processes.

[5] See Corollary 6.1 and Section 3.3 for details.

$\epsilon_{i,n}$'s to be normally distributed, a door is open to studying the important and challenging problem of inference of non-Gaussian spatial processes.

2.2. Functional Dependence Coefficient

Definition 2.2 (The L^p-functional dependence coefficient). *For any $p \geq 1$ and $s \geq 0$, we define the L^p-functional dependence (L^p-FD) coefficient, also called p-stability coefficient, as*

$$\Delta_p(s) \equiv \sup_n \sup_{i \in D_n} \delta_p\left(i, \{j \in D_n : d_{ij} \geq s\}, n\right). \tag{2.2}$$

When $\Delta_p(s) \to 0$ as $s \to \infty$, $\{Y_{i,n}\}$ is said to be L^p-functionally dependent (L^p-FD) or p-stable on the independent random field $\{\epsilon_{i,n}\}$.

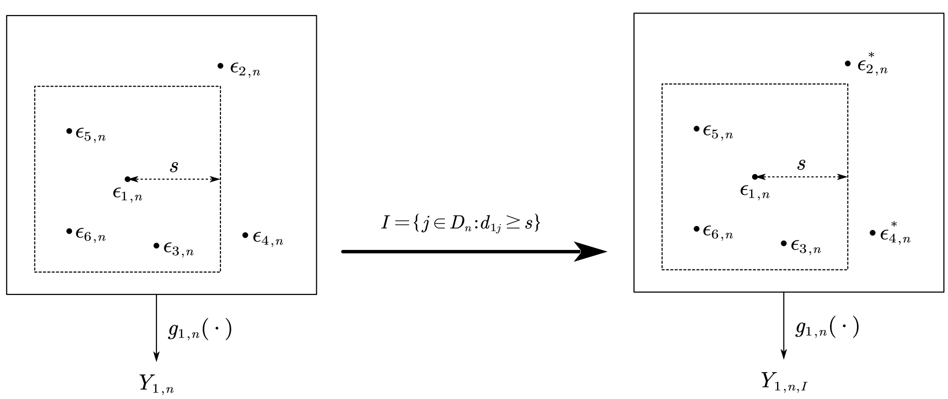

Figure 1: An illustration of $\delta_p\left(1, I \equiv \{j \in D_n : d_{1j} \geq s\}, n\right) \equiv \|Y_{1,n} - Y_{1,n,I}\|_{L^p}$

Figure 1 illustrates the definition of $\Delta_p(s)$. The L^p-FD coefficient defined above is easy to calculate and enjoys many desirable properties as shown in Sections 3-6. The $\delta_p\left(i, \{j \in D_n : d_{ij} \geq s\}, n\right)$ in (2.2) measures the total influence of $\epsilon_{j,n}$'s ($d_{ij} \geq s$) on $Y_{i,n}$, defined as the magnitude of the change of $Y_{i,n}$ under L^p-norm if $\epsilon_{j,n}$'s are replaced by their i.i.d. copy $\epsilon^*_{j,n}$'s simultaneously. Therefore, $\Delta_p(s) \to 0$ as $s \to \infty$ implies that the total impact from individuals far away can be arbitrarily small uniformly in both i and n. Note that by Lyapunov's inequality, if $\{Y_{i,n}\}$ is L^p-FD on $\{\epsilon_{i,n}\}$, it is also L^q-FD on $\{\epsilon_{i,n}\}$ for all $q \in [1, p]$. Although we do not know whether $\Delta_p(s)$ is weakly decreasing, $\Delta_p(s)$ has a property similar to monotonicity: $\Delta_p(s) \leq 3\Delta_p(\tilde{s})$ for any $s \geq \tilde{s}$ (Lemma S.3).

Compared with the functional dependence concepts in Wu (2005) and El Machkouri et al. (2013), ours is better suited for spatial econometric settings. Since $\{Y_{i,n}\}$ might be nonstationary and heterogeneous, we need to calculate the spatial FDM of every unit and take the supremum over all spatial units, but Wu (2005) and El Machkouri et al. (2013) do not need to do so, as they study stationary processes. In addition, our concept $\Delta_p(s)$ employs the information of distance s, while El Machkouri et al. (2013) define $\Delta_p(s)$ as $\Delta_p \equiv \sum_{j \in \mathbb{R}^d} \delta_p(i, j)$. So, our definition shares some similarity to spatial NED.

Compared with mixing (Jenish and Prucha, 2009) and NED (Jenish and Prucha, 2012), our functional dependence coefficient in Definition 2.2 is more convenient. The strong mixing coefficient is challenging to calculate since it involves complicated manipulation of taking the supremum over σ-fields. Calculating spatial NED coefficient involves conditional expectation, which sometimes is not easy. But calculating the L^p-FD coefficient is quite convenient because the construction of the coupled version $Y_{i,n,I}$ is explicit. The advantage of functional dependence over spatial NED is discussed in detail in Section 6.

Furthermore, we define the concept of *the second-type functional dependence coefficient*, which is mainly used in our proofs, in Appendix B. Moreover, we generalize the concept of L^p-FD to *the conditional L^p-FD,* which is particularly useful for spatial panel data models. See Appendix D for details.

3. Properties of Spatial Functional Dependence

In this section, we establish some useful inequalities, an LLN, and a CLT for the $Y_{i,n}$ generated by (2.1). Throughout this section, we let T_n be a finite subset of D_n, we assume $|T_n| \to \infty$ as $n \to \infty$, and write $S_n \equiv \sum_{i \in T_n} Y_{i,n}$.[6]

3.1. Inequalities under Spatial Functional Dependence

Moment and probability inequalities are crucial for developing limit theorems. In this subsection, we establish a moment inequality, an exponential inequality and a Nagaev-type inequality under spatial functional dependence.

3.1.1. A moment inequality

Theorem 3.1. *Under Assumption 1, if $\{Y_{i,n}\}$ is L^p-FD on $\{\epsilon_{i,n}\}$ for some $p \geq 2$ with the L^p-FD coefficient $\Delta_p(s) = O\left(s^{-\kappa}\right)$ for some $\kappa > \frac{d}{2}$ as $s \to \infty$, then*

$$\left\|S_n - \mathbb{E}S_n\right\|_{L^p} \leq C \left|T_n\right|^{1/2},$$

where $C > 0$ is a constant depending neither on T_n nor n.

Theorem 3.1 implies that $\left\|\sum_{i \in T_n}(Y_{i,n} - \mathbb{E}Y_{i,n})\right\|_{L^p} = O(\sqrt{|T_n|})$ as $n \to \infty$, the same order as the i.i.d. case. This inequality not only gives the convergence rate of the LLN but also plays an essential role in establishing the CLT and the exponential inequality. The constant C has an explicit form, see the proof of this theorem and Theorem B.1 for more information.

3.1.2. An exponential inequality

Exponential inequalities play an indispensable role in high-dimensional statistics, nonparametric and semiparametric econometrics. White and Wooldridge (1991) collect some exponential inequalities for time series. Xu and Lee (2018) establish an exponential inequality for spatial NED random fields. Wainwright (2019) focuses on the independent case for high-dimensional models.

[6] When we study a cross-sectional spatial econometric model, $T_n = D_n$ and $|D_n| = n$, and we do not need to introduce T_n. However, when we study a spatial panel data model, we usually assume that the underlying data originate from $t = -\infty$, and thus, $|D_n| = \infty$. In practice, the observable data are only a subset of all y_{it}'s. Thus, we introduce $T_n \subset D_n$, rather than use D_n.

Theorem 3.2. *Under Assumption 1, if (i) $\mathbb{E}Y_{i,n} = 0$ for all $n \geq 1$, $i \in T_n$, and (ii) $\{Y_{i,n}\}$ is L^p-FD on $\{\epsilon_{i,n}\}$ for any real number $p \geq 2$ with the L^p-FD coefficient $\Delta_p(s) \leq O(p^\nu) O(s^{-\kappa})$ for some $\kappa > \frac{d}{2}$ and $\nu \geq 0$ as $p \to \infty$ and $s \to \infty$, where $O(p^\nu)$ does not depend on s and $O(s^{-\kappa})$ does not depend on p, then for any $\epsilon > 0$,*

$$\mathbb{P}(|S_n| \geq |T_n| \epsilon) \leq C_1 \exp\left(-C_2 |T_n|^{1/(1+2\nu)} \epsilon^{2/(1+2\nu)}\right), \tag{3.1}$$

where the constants $C_1, C_2 > 0$ depend neither on T_n, n nor ϵ.

The condition $\Delta_p(s) \leq O(p^\nu) O(s^{-\kappa})$ restricts the speed at which $\Delta_p(s)$ decreases as $s \to \infty$ and the speed at which $\Delta_p(s)$ increases as $p \to \infty$, and requires that the effects of s and p on $\Delta_p(s)$ be separable. This condition can be easily satisfied for the SAR models discussed in Section 4.2. However, it is still possible that this condition is not satisfied, e.g., $\Delta_p(s) = \eta^{s/p}$ for some $0 < \eta < 1$. For such situations, one can refer to Theorem B.2 in Appendix B for a more general condition.

Next, we compare our exponential inequality with those in the literature.

1. Compared with the exponential inequality in Xu and Lee (2018), ours enjoys some desirable features. First, in our exponential inequality, the term $|T_n|^{1/(1+2\nu)}$ does not depend on d. Second, Xu and Lee (2018) require the NED coefficient to decrease exponentially fast, while we allow the L^p-FD coefficient to decrease at the speed of a power function. Third, the decay rate of our exponential inequality is faster than that of Xu and Lee (2018). Therefore, all the shortcomings of the exponential inequality in Xu and Lee (2018) pointed out by Yuan and Spindler (2022, p.4) have been overcome.

2. Yuan and Spindler (2022) also study the exponential inequality under NED. Compared with their results, our exponential inequality does not have a remainder term.

3. Compared with the standard Bernstein's inequality and Hoeffding's inequality (Wainwright,

2019), we see that when $\nu = 0$ in (3.1), the decay rate with respect to n is the same as the independent case; when $\nu > 0$, the decay rate is slower.

3.1.3. A Nagaev-type inequality

The condition $\Delta_p(s) \leq O(p^{\nu}) O(s^{-\kappa})$ in Theorem 3.2 usually requires that $Y_{i,n}$ have infinite order of moments, which might be restrictive in some applications in spatial econometrics. In fact, if only a finite order of moments exists, we have a Nagaev-type inequality. Nagaev (1979) establishes the Nagaev inequality for i.i.d. random variables. Liu et al. (2013) and Wu and Wu (2016) establish two Nagaev-type inequalities for functionally dependent time series. In this paper, we follow the idea in Wu and Wu (2016) to establish a Nagaev-type inequality for functionally dependent spatial variables. To begin with, we generalize the dependence-adjusted norm (DAN) concept given in Wu and Wu (2016), which plays the role of L^p-norm in the traditional Nagaev inequality (Lemma S.11). For any $\omega \geq 0$, we define the DAN as

$$\|Y.\|_{p,\omega} \equiv \sup_{s\geq 0} (s+1)^{\omega} \Delta_p(s) < \infty.$$

Theorem 3.3. *We assume* $\mathbb{E}Y_{i,n} = 0$ *for all* $i \in T_n$. *If* $\|Y.\|_{p,\omega} < \infty$ for some $\omega > d$ and $p > 2$. *Then, for all* $x > 0$ *and* $\kappa \geq 1$ *satisfying* $\kappa > \frac{3}{2(\omega-d)}$,

$$\mathbb{P}(|S_n| \geq 2x) \leq \frac{C_1 \|Y.\|_{p,\omega}^p |T_n|}{x^p} + C_2 |T_n|^{\kappa d} \exp\left(-\frac{C_3 x^2}{\|Y.\|_{2,\omega}^2 |T_n|}\right),$$

where $C_1, C_2, C_3 > 0$ *are constants depending neither on* x, n *nor* T_n.

3.2. Limit Theorems under Spatial Functional Dependence

Now, we establish an LLN and a CLT under spatial functional dependence, which are vital to establishing large sample properties of various estimators and test statistics in econometrics and statistics.

3.2.1. The law of large numbers

Theorem 3.4. *Under Assumption 1, if (i)* $\|Y\|_{L^p} = \sup_n \sup_{i \in D_n} \|Y_{i,n}\|_{L^p} < \infty$ *for some* $p > 1$*, and (ii)* $\{Y_{i,n}\}$ *is* L^1*-FD on* $\{\epsilon_{i,n}\}$*, i.e.,* $\lim_{s\to\infty} \Delta_1(s) = 0$*, then*

$$|T_n|^{-1} (S_n - \mathbb{E}S_n) \xrightarrow{L^1} 0.$$

We note that the moment inequality (Theorem 3.1) also implies an LLN. Theorem 3.1 requires some conditions on L^p-spatial FD coefficient ($p \geq 2$); however, Theorem 3.4 only imposes conditions on L^1-FD coefficient for the LLN, and it only requires that $\lim_{s\to\infty} \Delta_1(s) = 0$ without any specific decreasing rate.

3.2.2. The central limit theorem

Theorem 3.5. *When* $p_Y = 1$*, under Assumption 1, if (i)* $\sup_n \sup_{i \in D_n} \|Y_{i,n}\|_{L^p} < \infty$ *for some* $p > 2$ *, (ii)* $B \equiv \liminf_{n\to\infty} |T_n|^{-1} \sigma_n^2 > 0$*, where* $\sigma_n^2 \equiv \mathrm{Var}(S_n)$*, and (iii) the* L^2*-FD coefficient of* $\{Y_{i,n}\}$ *on* $\{\epsilon_{i,n}\}$ *satisfies* $\Delta_2(s) = O(s^{-\kappa})$ *as* $s \to \infty$ *for some* $\kappa > \frac{d}{2}$*, then*

$$\frac{S_n - \mathbb{E}S_n}{\sigma_n} \xrightarrow{d} N(0,1).$$

Conditions (i)-(ii) in Theorem 3.5 are standard in establishing CLTs (see Jenish and Prucha, 2012, etc.). Condition (iii) requires that the dependence among $Y_{i,n}$'s cannot be too strong.

The NED CLT in Jenish and Prucha (2012) requires that the L^2-NED coefficient $\psi(s)$ satisfies $\sum_{m=1}^{\infty} m^{d-1}\psi(m) < \infty$. Our spatial functional dependence CLT requires only that the L^2-FD coefficient $\Delta_2(s)$ decreases faster than $s^{-d/2}$, which is less restrictive.

By the Cramér-Wold device, we can generalize Theorem 3.5 to the multivariate case:

Corollary 3.1. *We write* $\Sigma_n \equiv \mathrm{Var}(S_n)$ *and* $\lambda_{\min}(\Sigma_n)$ *is the minimum eigenvalue of* Σ_n*. When* $p_Y \geq 1$*, under Assumption 1, if (i)* $\sup_n \sup_{i \in D_n} \|Y_{i,n}\|_{L^p} < \infty$ *for some* $p > 2$*, (ii)* $B \equiv$

$\liminf_{n\to\infty} |T_n|^{-1} \lambda_{\min}(\Sigma_n) > 0$, *and (iii) the* L^2*-FD coefficient of* $\{Y_{i,n}\}$ *on* $\{\epsilon_{i,n}\}$ *satisfies* $\Delta_2(s) = O\left(s^{-\kappa}\right)$ *for some* $\kappa > \frac{d}{2}$, *then*

$$\Sigma_n^{-1/2}\left(S_n - \mathbb{E}S_n\right) \xrightarrow{d} N\left(0, I_{p_Y}\right).$$

3.3. Heteroskedasticity and Autocorrelation Consistent Estimator

For inference, we propose a heteroskedasticity and autocorrelation consistent (HAC) estimator for the variance term $V_n \equiv \mathrm{Var}(|T_n|^{-1/2}S_n) = |T_n|^{-1}\Sigma_n$ in the CLT (Corollary 3.1). The idea is borrowed from Kojevnikov, Marmer and Song (2021). We assume $\mathbb{E}Y_{i,n} = 0$ for all $i \in T_n$. Then, the variance of $|T_n|^{-1/2}S_n$ is

$$V_n \equiv \mathrm{Var}(|T_n|^{-1/2}S_n) = \sum_{s\geq 0} v_n(s), \tag{3.2}$$

where $v_n(s) = |T_n|^{-1}\sum_{i\in T_n}\sum_{j\in T_n: d_{ij}\in[s,s+1)} \mathbb{E}(Y_{i,n}Y'_{j,n})$.

As in the time series literature, we employ a kernel function $k(\cdot) : \mathbb{R} \to [-1, 1]$ to assign weights to the auto-covariance terms $v_n(s)$ so that we can estimate V_n consistently. Let b_n be the bandwidth. Then, the HAC estimator of V_n is given by

$$\widehat{V}_n = \sum_{s\geq 0} k_n(s)\widehat{v}_n(s), \tag{3.3}$$

where $k_n(s) = k(s/b_n)$ and $\widehat{v}_n(s) = |T_n|^{-1}\sum_{i\in T_n}\sum_{j\in T_n: d_{ij}\in[s,s+1)} Y_{i,n}Y'_{j,n}$.

Next, we establish the consistency of the HAC estimator by imposing certain assumptions on the moment and weak dependence of $\{Y_{i,n}\}$, the bandwidth b_n, and the kernel function $k(\cdot)$.

Theorem 3.6. *Let* $2 < p_0 \leq q_0 \in \mathbb{R}$ *satisfy* $\frac{1}{p_0} + \frac{3}{q_0} = \frac{1}{2}$. *We suppose* $\mathbb{E}Y_{i,n} = 0$ *for all* $i \in T_n$. *If (i)* $\|Y\|_{L^{q_0}} \equiv \sup_{i,n}\|Y_{i,n}\|_{L^{q_0}} < \infty$, *(ii) the kernel function* $k(\cdot)$ *satisfies* $k(0) = 1$, $k(u) = 0$ *when* $|u| > 1$, $k(u) = k(-u)$ *for all* $u \in \mathbb{R}$, *and* $|k(u) - 1| \leq C_k|u|^{1+c_k}$ *for some constants*

$C_k, c_k > 0$, *(iii)* $\{Y_{i,n}\}$ *is* L^2*-FD on an independent random field* $\{\epsilon_{i,n}\}$ *with the* L^2*-FD coefficient satisfying* $\Delta_2(s) = C_\Delta s^{-c_\Delta}$*, where the constants* $C_\Delta > 0$ *and* $c_\Delta > \max\left\{c_k + d + 1, \frac{(2q_0-6)d}{q_0-6}\right\}$*, (iv)* $b_n = C_b|T_n|^{c_b}$ *for some constants* $C_b > 0$ *and* $c_b \in (0, \frac{1}{2d})$*, then as* $n \to \infty$*,*

$$\widehat{V}_n - V_n = o_p(1).$$

Remark 3.1. We note that $2 < p_0 \leq q_0$ and $\frac{1}{p_0} + \frac{3}{q_0} = \frac{1}{2}$ imply $q_0 \geq 8$. Condition (ii) is satisfied by most common kernel functions with compact supports, such as $k(u) = 1(|u| \leq 1)$. Condition (iii) requires the L^2-FD coefficient of $\{Y_{i,n}\}$ to decrease sufficiently fast. Condition (iv) assumes that the bandwidth b_n increases as a power function of $|T_n|$.

4. Examples of Spatial Stable Processes

In this section, we provide some primitive conditions to calculate the spatial FDM and the L^p-FD coefficient of $\{Y_{i,n}\}$ generated by an SAR model, a threshold SAR model, and a spatial panel data model. We also show how FD is employed to establish asymptotic distributions of estimators of the SAR Tobit model and a dynamic network quantile regression model.

4.1. A General Criterion

First, we provide a general criterion to establish the L^p-FD property. Let $X = \{X_{i,n}, i \in D_n, n \geq 1\}$ be an $\mathbb{R}^{p_X}$-valued triangular array random field and denote $X_n \equiv \left(\left(X_{i,n}'\right)_{i \in D_n}\right)'$. Suppose $\{Y_{i,n}, i \in D_n, n \geq 1\}$ is generated by[7]

$$Y_{i,n} = h_{i,n}\left(X_n\right), \tag{4.1}$$

[7]When $X_{i,n}$'s are independent, representation (4.1) is identical to (2.1). However, representation (4.1) allows $X_{i,n}$'s to be dependent, more specifically, allows $\{X_{i,n} = X_{i,n}(u_n)\}$ to be generated by an independent random field $\{u_{i,n}\}$. In this case, $h_{i,n}(X_n(\cdot))$ can be regarded as $g_{i,n}(\cdot)$ in (2.1).

where $h_{i,n} : (\mathbb{R}^{p_X})^n \to \mathbb{R}^{p_Y}$ satisfies the following condition: for all $x, x^\bullet \in (\mathbb{R}^{p_X})^n$,

$$\|h_{i,n}(x) - h_{i,n}(x^\bullet)\| \leq \sum_{j \in D_n} m_{ij,n} \left\|x_j - x_j^\bullet\right\|, \tag{4.2}$$

where x_j is the jth component of x. Denote $\phi(s) \equiv \sup_{n, i \in D_n} \sum_{j \in D_n : d_{ij} \geq s} m_{ij,n}$.

Proposition 4.1. *If (i) $\lim_{s \to \infty} \phi(s) = 0$, (ii) $X_{i,n}$'s are independent, and (iii) $\|X\|_{L^p} = \sup_{n, i \in D_n} \|X_{i,n}\|_{L^p} < \infty$ for some $p \geq 1$, then (i) for all $i, j \in D_n$, $\delta_p(i,j,n) \leq 2\|X\|_{L^p} m_{ij,n}$, and (ii) $\{Y_{i,n}\}$ is L^p-FD on $\{X_{i,n}\}$ with the L^p-FD coefficient $\Delta_p(s)$ satisfying $\Delta_p(s) \leq 2\|X\|_{L^p} \phi(s)$ for all $s \in [0, \infty)$.*

When $X_{i,n}$'s are dependent, we assume they are generated by another latent independent random field $\{u_{i,n} : i \in D_n, n \geq 1\}$, and the L^p-FD property of $\{Y_{i,n}\}$ is more complicated than that in Proposition 4.1. The result is presented in Proposition 4.2.

Proposition 4.2. *If (i) $\lim_{s \to \infty} \phi(s) = 0$ and $\phi(0) = \sup_{n, i \in D_n} \sum_{j \in D_n} m_{ij,n} < \infty$, (ii) for some $p \geq 1$, $\{X_{i,n}\}$ is L^p-FD on an independent random field $\{u_{i,n}\}$ with the FDM $\delta_{X,p}(i, I, n)$ and the L^p-FD coefficient $\Delta_{X,p}(s)$ satisfying $\lim_{s \to \infty} \Delta_{X,p}(s) = 0$ and $\Delta_{X,p}(0) < \infty$, then (i) for all $i, k \in D_n$, the FDM of $\{Y_{i,n}\}$ on $\{u_{i,n}\}$ satisfies $\delta_p(i,k,n) \leq \sum_{j \in D_n} m_{ij,n} \delta_{X,p}(j,k,n)$, and (ii) the L^p-FD coefficient $\Delta_p(s)$ satisfies $\Delta_p(s) \leq 3\Delta_{X,p}(0)\phi(\tilde{s}) + 3\phi(0)\Delta_{X,p}(s - \tilde{s})$ for all $s \in [0, \infty)$ and $\tilde{s} \in [0, s]$. In particular, $\Delta_p(s) \leq 3\Delta_{X,p}(0)\phi\left(\frac{s}{2}\right) + 3\phi(0)\Delta_{X,p}\left(\frac{s}{2}\right)$.*

We note that the conditions on $m_{ij,n}$ for L^p-FD and L^2-NED (Proposition 1 in Jenish and Prucha, 2012) are almost identical. However, here the $p \geq 1$ can be an arbitrary number; Proposition 1 in Jenish and Prucha (2012) is applicable only to $p = 2$. With more choices for p, FD is more flexible and more convenient than NED in applications.

4.2. Spatial Autoregressive Models

In this subsection, we calculate the FDM and the L^p-FD coefficient for the SAR models. The individuals $1, 2, \ldots, n$ are located in some lattice $D_n \subset \mathbb{R}^d$ satisfying Assumption 1, and we identify each individual with its location in $\mathbb{R}^d$ for simplicity.

4.2.1. SAR model

The SAR model can be written as

$$\begin{pmatrix} Y_{1,n} \\ \vdots \\ Y_{n,n} \end{pmatrix} = Y_n = F\left(\lambda W_n Y_n + X_n\beta + \epsilon_n\right) = \begin{pmatrix} F\left(\lambda W_{1\cdot,n} Y_n + X_{1,n}'\beta + \epsilon_{1,n}\right) \\ \vdots \\ F\left(\lambda W_{n\cdot,n} Y_n + X_{n,n}'\beta + \epsilon_{n,n}\right) \end{pmatrix}, \tag{4.3}$$

where $W_n = (w_{ij,n})_{n\times n}$ is a nonstochastic and nonzero spatial weights matrix, $W_{i\cdot,n}$ is the ith row of W_n, $F : \mathbb{R} \to \mathbb{R}$ is a Borel-measurable function, $F(a) \equiv (F(a_1), \ldots, F(a_n))'$ for any column vector $a = (a_1, \ldots, a_n)' \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$ and $\beta \in \mathbb{R}^K$ are true model parameters, $X_n = (X_{1,n}, X_{2,n}, \ldots, X_{n,n})' \in \mathbb{R}^{n\times K}$ is the exogenous variable matrix, and $\epsilon_n = (\epsilon_{1,n}, \epsilon_{2,n}, \ldots, \epsilon_{n,n})' \in \mathbb{R}^n$ is the disturbance term. The SAR model and its variants have been widely used in applications. When $F(x) = x$, (4.3) becomes $Y_n = \lambda W_n Y_n + X_n\beta + \epsilon_n$, which is the standard (linear) SAR model; when $F(x) = \max(0, x)$, (4.3) becomes the SAR Tobit model studied in Xu and Lee (2015a).

We employ Propositions 4.1 and 4.2 to show that the $\{Y_{i,n}\}$ generated by (4.3) is FD under some weak conditions. To do so, we need to impose some assumptions on the function F, the spatial weights matrix W_n, $\{X_{i,n}\}$, and $\{\epsilon_{i,n}\}$.

Assumption 2. *F is a Lipschitz function with the Lipschitz constant $L > 0$, i.e., for any $e, e^\bullet \in \mathbb{R}$, $|F(e^\bullet) - F(e)| \le L\,|e^\bullet - e|$. And $\zeta \equiv L\,|\lambda| \sup_n \|W_n\|_\infty < 1$.*

Assumption 2 is a generalization of Assumption 2 in Xu and Lee (2015a) and Assumption 3 in Xu and Lee (2015b). It ensures the existence and uniqueness of the solution of (4.3). See Xu and Lee (2015a) for more discussion about it.

Assumption 3. *The weights $w_{ij,n}$'s in W_n satisfy one of the following conditions:*

(1) Only individuals whose distances are less than some specific constant $\bar{d}_0 > 1$ may affect each other directly, i.e., $w_{ij,n}$ can be nonzero only if $d_{ij} < \bar{d}_0$;

(2) $|w_{ij,n}| \le c d_{ij}^{-\alpha}$ *for some constants* $c > 0$ *and* $\alpha > d$.

Assumption 3 is intuitive. From our definition, FD property implies that when the distance d_{ij} is large, $\epsilon_{i,n}$ has a negligible impact on $Y_{j,n}$. In the SAR model, $w_{ij,n}$ represents the direct impact of $Y_{j,n}$ on $Y_{i,n}$. Thus, intuitively, for an FD SAR process, $|w_{ij,n}|$ should decrease as d_{ij} increases. Assumption 3(1) implies that there is a threshold distance $\bar{d}_0$ such that when $d_{ij} \ge \bar{d}_0$, $w_{ij,n}$ will be zero. Assumption 3(2) allows $w_{ij,n}$ to decrease as a power function of the distance d_{ij}. In fact, we relax Assumption 3(2) in Xu and Lee (2015a) which requires the number of spatial units with strong impacts to be uniformly bounded. And if we impose a faster decreasing rate on $w_{ij,n}$, e.g., $|w_{ij,n}| \le c \exp(-\alpha d_{ij})$, we can obtain a stronger conclusion.

Assumption 4. *One of the following conditions is satisfied:*

(1) $(X_{i,n}', \epsilon_{i,n})$*'s are independent over* i*;*

(2) for some $p \ge 1$, $\left\{ \left(X_{i,n}', \epsilon_{i,n} \right)' : i \in D_n, n \ge 1 \right\}$ *is* L^p*-FD on an independent random field* $u = \{u_{i,n} : i \in D_n, n \ge 1\}$ *with the spatial FDM* $\delta_{X\epsilon,p}(i, I, n)$ *and the* L^p*-FD coefficient* $\Delta_{X\epsilon,p}(s)$ *satisfying* $\lim_{s\to\infty} \Delta_{X\epsilon,p}(s) = 0$ *and* $\Delta_{X\epsilon,p}(0) < \infty$.

Assumption 4 considers two cases: (1) $(X_{i,n}', \epsilon_{i,n})$'s are independent, and (2) they are spatially dependent. In both cases, we do not require that $X_{i,n}$ and $\epsilon_{i,n}$ be independent. So, conditional heteroskedasticity is allowed. When $(X_{i,n}', \epsilon_{i,n})$'s are spatially dependent, we suppose that they are generated by some independent underlying random vectors $u_{i,n}$'s. Similar ideas are widely employed. For example, in time series, we usually model a dependent process as a moving average process. Since Assumption 4(2) allows $\epsilon_{i,n}$'s to be dependent, the SAR model with an SAR disturbance (called the SARAR model) is a special case of (4.3), and we discuss it in Section 4.2.2. Moreover, Assumption 4(2) allows contextual effects. If $\{X_{i,n}\}$ is L^p-FD on some independent random field $\{u_{i,n}\}$, by Proposition 4.2, $\{W_{i\cdot,n} X_n\}$ is also L^p-FD on $\{u_{i,n}\}$ under some reasonable conditions. Thus, Assumption 4(2) allows $\{W_{i\cdot,n} X_n\}$ as a special term.

We recall that L is the Lipschitz constant of $F(\cdot)$. To present our main results, denote

$$M_n \equiv (m_{ij,n})_{n\times n} \equiv L\left(I_n - L\left|\lambda\right|\left|W_n\right|\right)^{-1},$$

where $|W_n| \equiv (|w_{ij,n}|)_{n\times n}$. Under Assumption 2, M_n is well-defined and Neumann's expansion is allowed.

Proposition 4.3. *We assume $C_{x\epsilon,p} \equiv \sup_{n,i}\left\|X'_{i,n}\beta + \epsilon_{i,n}\right\|_{L^p} < \infty$, where p is the same as that in Assumption 4. Let $\delta_p(i,j,n)$ and $\Delta_p(s)$ denote the FDM and the L^p-FD coefficient of $\{Y_{i,n}\}$, respectively.*

(1) Under Assumption 2, $\{Y_{i,n}\}$ is uniformly L^p-bounded.

(2) (i) Under Assumptions 2 and 4(1), $\delta_p(i,j,n) \le 2C_{x\epsilon,p}m_{ij,n}$ for all $i,j \in D_n$.

(ii) Under Assumptions 2 and 4(2), $\delta_p(i,k,n) \le \sum_{j=1}^{n}(\|\beta\|+1)m_{ij,n}\delta_{X\epsilon,p}(j,k,n)$ for all $i,k \in D_n$.

(3) Under Assumptions 1, 2, 3(1), and 4(1), $\{Y_{i,n}\}$ is L^p-FD on $\left\{\left(X'_{i,n},\epsilon_{i,n}\right)'\right\}$ and $\Delta_p(s) \le 2C_{x\epsilon,p}\phi(s)$ for all $s\in[0,\infty)$, where $\phi(s) \le \frac{L}{1-\zeta}\zeta^{s/\bar{d}_0}$ and ζ is defined in Assumption 2.

(4) Under Assumptions 1, 2, 3(2), and 4(1), $\{Y_{i,n}\}$ is L^p-FD on $\left\{\left(X'_{i,n},\epsilon_{i,n}\right)'\right\}$ and $\Delta_p(s) \le 2C_{x\epsilon,p}\phi(s)$ for all $s\in[0,\infty)$, where $\phi(s) = O\left(s^{-(\alpha-d)}(\log s)^{\alpha-d}\right)$ does not depend on p.

(5) Under Assumptions 1, 2, 3(1), and 4(2), $\{Y_{i,n}\}$ is L^p-FD on $\{u_{i,n}\}$ and

$$\Delta_p(s) \le 3\left(\|\beta\|+1\right)\Delta_{X\epsilon,p}(0)\phi\left(\frac{s}{2}\right) + 3\left(\|\beta\|+1\right)\phi\left(0\right)\Delta_{X\epsilon,p}\left(\frac{s}{2}\right)$$

for all $s\in[0,\infty)$, where $\phi(s) = \frac{L}{1-\zeta}\zeta^{s/\bar{d}_0}$. In particular, as $s\to\infty$,

(i) if $\Delta_{X\epsilon,p}(s) = O\left(s^{-\alpha_1}\right)$ for some $\alpha_1 > 0$, then $\Delta_p(s) = O\left(s^{-\alpha_1}\right)$;

(ii) if $\Delta_{X\epsilon,p}(s) = O\left(\eta^s\right)$ for some $0<\eta<1$, then $\Delta_p(s) = O\left(\xi^s\right)$, where $\xi = \max\left(\eta^{1/2},\zeta^{1/\left(2\bar{d}_0\right)}\right)$.

(6) Under Assumptions 1, 2, 3(2), and 4(2), $\{Y_{i,n}\}$ is L^p-FD on the random field $\{u_{i,n}\}$ and

$$\Delta_p(s) \leq 3\left(\|\beta\|+1\right)\Delta_{X\epsilon,p}(0)\phi\left(\frac{s}{2}\right)+3\left(\|\beta\|+1\right)\phi\left(0\right)\Delta_{X\epsilon,p}\left(\frac{s}{2}\right)$$

for all $s\in[0,\infty)$, where $\phi\left(s\right)=O\left(s^{-(\alpha-d)}\left(\log s\right)^{\alpha-d}\right)$ does not depend on p. In particular, as $s\to\infty$, if $\Delta_{X\epsilon,p}\left(s\right)=O\left(s^{-(\alpha-d)}\left(\log s\right)^{\alpha-d}\right)$, then $\Delta_p\left(s\right)=O\left(s^{-(\alpha-d)}\left(\log s\right)^{\alpha-d}\right)$.

From Proposition 4.3, under certain conditions, the $\{Y_{i,n}\}$ generated by the SAR model is L^p-FD. In Section S.2, we apply Proposition 4.3 to show that the score function of the SAR Tobit model studied in Xu and Lee (2015a) satisfies a CLT, and this is a critical step in establishing the asymptotic distribution of their estimator.

4.2.2. SARAR model

The SARAR model is a generalization of the SAR model and is widely used in applications. Thus, we explore its functional dependence properties. The form of the SARAR model is the same as (4.3), but $\epsilon_n=\rho M_n\epsilon_n+v_n$, where $v_n=(v_{1,n},\ldots,v_{n,n})'$, $v_{i,n}$'s are i.i.d. random variables, and $M_n=(m_{ij,n})_{n\times n}$ is a nonstochastic and nonzero spatial weights matrix. As mentioned previously, the SARAR model is just a special case of the previous SAR model in our setting. Thus, we can employ Proposition 4.3 to establish the L^p-FD property of the SARAR model by imposing some conditions on M_n to ensure that $\{\epsilon_{i,n}\}$ is L^p-FD on $\{v_{i,n}\}$.

Assumption 5. *(1) The Lipschitz constant of $F:\mathbb{R}\to\mathbb{R}$ is L, and $\zeta\equiv L\left|\lambda\right|\sup_n\left\|W_n\right\|_\infty<1$;*

(2) $|w_{ij,n}|\leq cd_{ij}^{-\alpha}$ and $|m_{ij,n}|\leq cd_{ij}^{-\alpha}$ for some constants $c>0$ and $\alpha>d$;

(3) for some $p\geq 1$, $\{X_{i,n}\}$ is L^p-FD on an independent random field $\{u_{i,n}:i\in D_n,n\geq 1\}$ with the L^p-FD coefficient $\Delta_{X,p}\left(s\right)=O\left(s^{-(\alpha-d)}\left(\log s\right)^{\alpha-d}\right)$ satisfying $\Delta_{X,p}\left(0\right)<\infty$; and $\left(u_{i,n}',v_{i,n}\right)$'s are independent over i;

(4) $\sup_n\left\|\rho M_n\right\|_\infty<1$;

(5) $\|v\|_{L^p} = \sup_{n,i} \|v_{i,n}\|_{L^p} < \infty, \|X\|_{L^p} = \sup_{n,i} \left\|X'_{i,n}\beta\right\|_{L^p} < \infty.$

Assumption 5 inherits the assumptions of Proposition 4.3(6) directly. Consequently, the L^p-FD coefficient of $\{Y_{i,n}\}$ is a direct result of Proposition 4.3(6).

Proposition 4.4. *Under Assumptions 1 and 5, $\{Y_{i,n}\}$ is L^p-FD on $\left\{\left(u'_{i,n}, v_{i,n}\right)'\right\}$ with the L^p-FD coefficient $\Delta_p(s) = O\left(s^{-(\alpha-d)} (\log s)^{\alpha-d}\right)$ as $s \to \infty$.*

4.2.3. SARMA model

The SAR model with moving average disturbances (SARMA model) is another generalization of SAR model (Doğan and Taşp1nar, 2013; Fingleton, 2008; Huang, 1984). The form of the SARMA model is the same as (4.3), except that $\epsilon_n = v_n - \rho M_n v_n$, where $v_n = (v_{1,n}, \ldots, v_{n,n})'$, $v_{i,n}$'s are i.i.d. random variables, and $M_n = (m_{ij,n})_{n\times n}$ is a nonstochastic and nonzero spatial weights matrix. Here are the assumptions needed to establish the FD properties of the SARMA model.

Assumption 6. *(1) The Lipschitz constant of $F : \mathbb{R} \to \mathbb{R}$ is L, and $\zeta \equiv L|\lambda| \sup_n \|W_n\|_\infty < 1$;*

(2) $|w_{ij,n}| \le c d_{ij}^{-\alpha}$ and $|m_{ij,n}| \le c d_{ij}^{-\alpha}$ for some constants $c > 0$ and $\alpha > d$;

(3) for some $p \ge 1$, $\{X_{i,n}\}$ is L^p-FD on an independent random field $\{u_{i,n} : i \in D_n, n \ge 1\}$ with the L^p-FD coefficient $\Delta_{X,p}(s)$ satisfying $\Delta_{X,p}(s) = O\left(s^{-(\alpha-d)} (\log s)^{\alpha-d}\right)$ as $s \to \infty$ and $\Delta_{X,p}(0) < \infty$; and $\left(u'_{i,n}, v_{i,n}\right)$'s are independent over i;

(4) $\|v\|_{L^p} = \sup_{n,i} \|v_{i,n}\|_{L^p} < \infty$ *and* $\|X\|_{L^p} = \sup_{n,i} \left\|X'_{i,n}\beta\right\|_{L^p} < \infty.$

Like the SARAR model, Assumption 6 also inherits the assumptions of Proposition 4.3(6). Thus the L^p-FD coefficient of $\{Y_{i,n}\}$ is a direct result of Proposition 4.3(6).

Proposition 4.5. *Under Assumptions 1 and 6, $\{Y_{i,n}\}$ is L^p-FD on $\left\{\left(u'_{i,n}, v_{i,n}\right)'\right\}$ with the L^p-FD coefficient $\Delta_p(s) = O\left(s^{-(\alpha-d)} (\log s)^{\alpha-d}\right)$ as $s \to \infty$.*

4.3. A Threshold Spatial Autoregressive Model

A threshold spatial autoregressive (TSAR) model, which combines a threshold model and an SAR model, has received increasing attention recently. Deng (2018) considers a TSAR model and proposes a two-stage least squares estimator for the model. Li (2022) studies the quasi-maximum likelihood estimation of a TSAR model. Here, we explore the functional dependence properties of the TSAR model in Li (2022), which can be written as

$$Y_n = \left(\lambda_1 D_\gamma + \lambda_2 \bar{D}_\gamma\right) W_n Y_n + D_\gamma X_n \beta_1 + \bar{D}_\gamma X_n \beta_2 + \epsilon_n, \tag{4.4}$$

where $Y_n = (Y_{1,n}, \ldots, Y_{n,n})'$, $D_\gamma = \operatorname{diag}\{1(q_{1,n} \le \gamma), \ldots, 1(q_{n,n} \le \gamma)\}$, $\bar{D}_\gamma = I_n - D_\gamma$, $\lambda_1, \lambda_2, \gamma \in \mathbb{R}$ and $\beta_1, \beta_2 \in \mathbb{R}^K$ are true model parameters, $X_n = (X_{1,n}, X_{2,n}, \ldots, X_{n,n})' \in \mathbb{R}^{n \times K}$ are exogenous variables, $q_{i,n}$'s are the exogenous threshold variables which might be part of $x_{i,n}$, $\epsilon_n = (\epsilon_{1,n}, \epsilon_{2,n}, \ldots, \epsilon_{n,n})' \in \mathbb{R}^n$ is the disturbance term, and $W_n = (w_{ij,n})_{n \times n}$ is a nonstochastic and nonzero spatial weights matrix. We first state some assumptions.

Assumption 7. *(1) $\lambda \sup_n \|W_n\|_\infty < 1$, where $\lambda \equiv \max\{|\lambda_1|, |\lambda_2|\}$;*

(2) $|w_{ij,n}| \le c d_{ij}^{-\alpha}$ for some constants $c > 0$ and $\alpha > d$;

(3) $q_{i,n}$'s are independent across i; for some $p \ge 1$, $\left\{\left(X_{i,n}', \epsilon_{i,n}\right)' : i \in D_n, n \ge 1\right\}$ is L^p-FD on an independent random field $u = \{u_{i,n} : i \in D_n, n \ge 1\}$ with the spatial FDM $\delta_{X\epsilon,p}(i, I, n)$ and the L^p-FD coefficient $\Delta_{X\epsilon,p}(s)$ satisfying $\Delta_{X\epsilon,p}(s) = O\left(s^{-(\alpha-d)} (\log s)^{\alpha-d}\right)$ as $s \to \infty$ and $\Delta_{X\epsilon,p}(0) < \infty$;

(4) $\|\epsilon\|_{L^p} = \sup_{n,i} \|\epsilon_{i,n}\|_{L^p} < \infty$ and $\|X\|_{L^p} = \sup_{n,i} \|X_{i,n}\|_{L^p} < \infty$.

Proposition 4.6. *Under Assumptions 1 and 7, the $\{Y_{i,n}\}$ generated by the model (4.4) is L^p-FD on $\left\{\left(u_{i,n}', q_{i,n}\right)'\right\}$ with the L^p-FD coefficient $\Delta_p(s) = O\left(s^{-(\alpha-d)} (\log s)^{\alpha-d}\right)$ as $s \to \infty$.*

4.4. Spatial Panel Data Models

In this section, we discuss the functional dependence of spatial panel data models. We suppose that in the panel data, there are N individuals named as $1, \ldots, N$ and they are located in $D_N \subset \mathbb{R}^d$, and the time periods originate from $-\infty$: $t = \ldots, -1, 0, 1, \ldots, T$. We regard each individual i at time t as a point in the $(d+1)$-dimensional spatial-temporal space $\mathbb{R}^{d+1}$: $(i,t) \in D_{NT} \equiv \left\{(i,t) \in \mathbb{R}^{d+1} : i \in D_N, t = T, T-1, \ldots\right\}$. We adopt the same metric as in Qu et al. (2017):

$$d_{it;j\tau} \equiv \|(i,t) - (j,\tau)\|_\infty = \max\{d_{ij}, |t-\tau|\} \equiv \max\left\{\max_{1\le k\le d}|i_k - j_k|, |t-\tau|\right\}.$$

We still consider the setting in Assumption 1, i.e., $d_{ij} \geq 1$ for any $i \neq j$. Consequently, for any different pair $(i,t), (j,\tau) \in D_{NT}$, $d_{it;j\tau} \geq 1$.

We suppose that the fixed effects in the spatial panel data model are random, which includes nonstochastic fixed effects as a special case. So, we employ the concept of conditional spatial FDM here.[8] Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the underlying probability space and $\mathcal{C}$ be a sub-σ-field of $\mathcal{F}$. We suppose ϵ_{it}'s ($(i,t) \in D_{NT}$) are conditionally independent on $\mathcal{C}$, and we write $\varepsilon_{Nt} \equiv (\epsilon'_{1t}, \epsilon'_{2t}, \ldots, \epsilon'_{Nt})'$. We suppose y_{it}'s ($(i,t) \in D_{NT}$) are generated by ε_{Nt}'s: $y_{it} = g_{it}(\varepsilon_{Nt}, \varepsilon_{N,t-1}, \ldots)$. We write $Y_{Nt} = (y_{1t}, \ldots, y_{Nt})'$ and $G_{Nt} = (g_{1t}, \ldots, g_{Nt})'$. We can also write the system as

$$Y_{Nt} = \begin{pmatrix} g_{1t}(\varepsilon_{Nt}, \varepsilon_{N,t-1}, \ldots) \\ \vdots \\ g_{Nt}(\varepsilon_{Nt}, \varepsilon_{N,t-1}, \ldots) \end{pmatrix} \equiv G_{Nt}(\varepsilon_{Nt}, \varepsilon_{N,t-1}, \ldots). \tag{4.5}$$

For all $(i,t) \in D_{NT}$, given $\mathcal{C}$, let ϵ^*_{it} be an i.i.d. copy of ϵ_{it}, and ϵ^*_{it} is independent of $\epsilon_{j\tau}$ for all $(j,\tau) \in D_{NT}$. For any set $I \subset D_{NT}$, we define $\epsilon_{it,I} \equiv \epsilon^*_{it}$ if $(i,t) \in I$ and $\epsilon_{it,I} \equiv \epsilon_{it}$ otherwise, and $\varepsilon_{Nt,I} \equiv \left(\epsilon'_{1t,I}, \epsilon'_{2t,I}, \ldots, \epsilon'_{Nt,I}\right)'$. Then $y_{it,I} = g_{it}(\varepsilon_{Nt,I}, \varepsilon_{N,t-1,I}, \ldots)$ is a coupled version of y_{it} on I and $Y_{Nt,I} = (y_{1t,I}, \ldots, y_{Nt,I})'$. Although the notation for spatial panel data is slightly different

[8] See Appendix D for details about the conditional spatial functional dependence.

from that in Section 2, e.g., D_{NT} here corresponds to D_n in Section 2, the setting in the spatial panel data is a special case of the general setting in Section 2. For clarity, we restate Definitions 2.1 and 2.2 in the spatial panel data setting.

Definition 4.1 (The FDM for spatial panel data)**.** *For $p \geq 1$, $(i,t) \in D_{NT}$ and $I \subset D_{NT}$, define the conditional functional dependence measure as $\delta_p^{\mathcal{C}}(it, I) \equiv \|y_{it} - y_{it,I}\|_{L^p,\mathcal{C}}$. When $I = \{(j,\tau)\}$ is a singleton, we simplify the notation as $\delta_p^{\mathcal{C}}(it, \{(j,\tau)\}) \equiv \delta_p^{\mathcal{C}}(it, j\tau)$.*

Definition 4.2 (The L^p-FD coefficient for spatial panel data)**.** *Let $p \geq 1$. For the system in (4.5), $\{y_{it}\}$ is said to be $\mathcal{C}$-conditionally L^p-functionally dependent (L^p-FD) or $\mathcal{C}$-conditionally p-stable on $\{\epsilon_{it}\}$ if the $\mathcal{C}$-conditional L^p-functional dependence (L^p-FD) coefficient satisfies*

$$\Delta_p^{\mathcal{C}}(s) \equiv \sup_{N,T} \sup_{(i,t) \in D_{NT}} \delta_p^{\mathcal{C}}(it, \{(j,\tau) \in D_{NT} : d_{it;j\tau} \geq s\}) \to 0 \quad \text{as } s \to \infty. \tag{4.6}$$

4.4.1. A general SDPD model

Next, we study the functional dependence properties of the spatial dynamic panel data (SDPD) model, which has been widely investigated in the literature. See, e.g., Yu et al. (2008) and Lee and Yu (2010), among many others. The SDPD model is specified as

$$Y_{Nt} = \lambda W_N Y_{Nt} + \gamma Y_{N,t-1} + \rho W_N Y_{N,t-1} + X_{Nt}\beta + \mu_t l_N + \boldsymbol{\nu}_N + V_{Nt}, \tag{4.7}$$

where $t = T, T-1, \ldots, i = 1, \ldots, N$, $Y_{Nt} = (y_{1t}, y_{2t}, \ldots, y_{Nt})'$, $W_N = (w_{ij,N})_{N\times N}$ is a nonstochastic spatial weights matrix and invariant as t changes, $X_{Nt} = (x_{1t}, \ldots, x_{Nt})' \in \mathbb{R}^{N\times p}$ is the regressor matrix, μ_t is the time fixed effect at period t, $l_N = (1, \ldots, 1)'$ is N-dimensional, $\boldsymbol{\nu}_N = (\nu_1, \ldots, \nu_N)'$ is an $N \times 1$ column vector of individual fixed effects, $V_{Nt} = (v_{1t}, \ldots, v_{Nt})'$ is the disturbance term, and $\lambda, \gamma, \rho, \beta \in \mathbb{R}$ are true model parameters. Denote $S_N \equiv I_N - \lambda W_N$, $A_N \equiv S_N^{-1}(\gamma I_N + \rho W_N)$, $\varepsilon_{Nt} \equiv X_{Nt}\beta + \mu_t l_N + \boldsymbol{\nu}_N + V_{Nt}$ and $\varepsilon_{Nt} \equiv (\epsilon_{1t}, \ldots, \epsilon_{Nt})'$. Then (4.7) can be written as $Y_{Nt} =$

$A_N Y_{N,t-1} + S_N^{-1}\varepsilon_{Nt}$. Under some suitable conditions, by iterating the above equation, we have

$$Y_{Nt} = \sum_{h=0}^{\infty} A_N^h S_N^{-1} \varepsilon_{N,t-h}. \tag{4.8}$$

For this example, we define $\mathcal{C} \equiv \vee_{t=-\infty}^{\infty} \vee_{N=1}^{\infty} \sigma(\mu_t, \boldsymbol{\nu}_N)$ as the sub-σ-field generated by all fixed effects. To obtain the conditional L^p-FD coefficient for the SDPD model, the following assumptions are needed.

Assumption 8. $|w_{ij,N}| \leq c d_{ij}^{-\alpha}$ *for some constants* $c > 0$ *and* $\alpha > d$.

Assumption 9. $\sup_N \|W_N\|_\infty \leq 1$ *and* $|\lambda| + |\gamma| + |\rho| < 1$. *Denote* $\zeta \equiv \frac{|\gamma|+|\rho|}{1-|\lambda|} < 1$.

Assumption 10. $\|\epsilon\|_{L^p,\mathcal{C}} \equiv \sup_{N,T} \sup_{i,t} \|\epsilon_{it}\|_{L^p,\mathcal{C}} < \infty$ *a.s. for some* $p \geq 1$.

Assumption 11. *Conditional on* $\mathcal{C}$, (x_{it}', v_{it})*'s are independent over* i *and* t.

These assumptions are like those for the SAR model, but all the statements here are conditional on $\mathcal{C}$.

Proposition 4.7. *For model (4.7), under Assumptions 1 and 8-11, (1)* $\{y_{it} : (i,t) \in D_{NT}\}$ *is* $\mathcal{C}$*-conditionally* L^p*-FD on* $\{\epsilon_{it}\}$ *with the* $\mathcal{C}$*-conditional* L^p*-FD coefficient* $\Delta_p^{\mathcal{C}}(s) = \|\epsilon\|_{L^p,\mathcal{C}} O\left(s^{-(\alpha-d)} (\log s)^{\alpha-d}\right)$ *almost surely as* $s \to \infty$*; (2) the same conclusion also holds for* $\{W_{i\cdot,N} Y_{Nt} : (i,t) \in D_{NT}\}$.

Remark 4.1. In Proposition 4.7, we require x_{it}'s and v_{it}'s to be conditionally independent on $\mathcal{C}$. In Section S.8.4 in the online supplement, we provide an example where v_{it}'s are correlated. Our conclusion can also be generalized to allow x_{it}'s to be correlated in both the spatial and time dimension. For instance, we consider $X_{Nt} = \sum_{\tau=0}^{\infty} D_\tau \tilde{X}_{N,t-\tau}$, where the $p_x \times 1$ random vectors $\tilde{x}_{it}$'s (the transpose of the ith row of $\tilde{X}_{Nt}$) are identically distributed and conditionally independent on $\mathcal{C}$ over i and t, D_τ is an $N \times N$ nonstochastic matrix whose row-sum-norm decreases exponentially as $\tau \to \infty$, i.e., $\|D_\tau\|_\infty \leq C_0 \exp(-C_1 \tau)$ for some constants $C_0, C_1 > 0$ and $\sup_i \sum_{j:d_{ij} \geq s} (D_\tau)_{ij} =$

$O\left(s^{-(\alpha-d)}(\log s)^{\alpha-d}\right)$ for any $\tau \geq 0$, where $(D_\tau)_{ij}$ is the (i,j)th entry of D_τ. Then

$$\sum_{h=0}^{\infty} A_N^h S_N^{-1} X_{N,t-h}\beta = \sum_{h=0}^{\infty} A_N^h S_N^{-1}\left(\sum_{\tau=0}^{\infty} D_\tau \tilde{X}_{N,t-h-\tau}\right)\beta = \sum_{k=0}^{\infty}\left(\sum_{\tau=0}^{k} A_N^{k-\tau} S_N^{-1} D_\tau\right)\tilde{X}_{N,t-k}\beta.$$

From the proof of Lemma A.8 in Su et al. (2023), $\left\|\sum_{\tau=0}^{k} A_N^{k-\tau} S_N^{-1} D_\tau\right\|_\infty$ decreases exponentially as $k \to \infty$. This fact can be used to replace the fact that $\left\|A_N^h S_N^{-1}\right\|_\infty$ decreases exponentially as $h \to \infty$ in the proof of Proposition 4.7. Further, by (1) replacing $t_1 - t_2 \geq s$ and $0 \leq t_1 - t_2 < s$ in the last inequality of (S.37) by $t_1 - t_2 \geq \tilde{s}$ and $0 \leq t_1 - t_2 < \tilde{s}$ respectively, where $\tilde{s}$ depends on s and $\tilde{s} \leq s$, and (2) selecting $\tilde{s}$ appropriately, we can show that $\{y_{it} : (i,t) \in D_{NT}\}$ is $\mathcal{C}$-conditionally L^p-FD on $\left\{(\tilde{x}_{it}', v_{it})'\right\}$ with the $\mathcal{C}$-conditional L^p-FD coefficient $\Delta_p^{\mathcal{C}}(s) = O\left(s^{-(\alpha-d)}(\log s)^{\alpha-d+1}\right)$ a.s. as $s \to \infty$.

Remark 4.2. We can allow the slope coefficients of different individuals to be different, as long as their upper bounds satisfy Assumption 9. For example, we denote the spatial coefficient for the ith row of $W_N Y_{Nt}$ by λ_i, $\lambda \equiv \sup_{N\in\mathbb{N}} \sup_{i=1,\ldots,N} |\lambda_i|$, and $\Lambda_N \equiv \operatorname{diag}\{\lambda_1, \ldots, \lambda_N\}$. Then

$$\Lambda_N W_N = \lambda \operatorname{diag}\left\{\frac{\lambda_1}{\lambda}, \ldots, \frac{\lambda_N}{\lambda}\right\} W_N \equiv \lambda \check{W}_N.$$

where $\check{W}_N \equiv \operatorname{diag}\left\{\frac{\lambda_1}{\lambda}, \ldots, \frac{\lambda_N}{\lambda}\right\} W_N$ can be regarded as the new spatial weights matrix. Then Proposition 4.7 remains applicable.

4.4.2. A DNQR model

Next, we give an example to illustrate how functional dependence is used to derive a CLT, which is an important step in deriving the asymptotic distribution of the estimator for the dynamic network quantile regression (DNQR) model in Xu et al. (2022). Their model can be specified as[9]

$$Y_{Nt} = \gamma_{1\tau} W_N Y_{Nt} + \gamma_{\tau 2} W_N Y_{N,t-1} + \gamma_{\tau 3} Y_{N,t-1} + \gamma_{0\tau} l_N + Z_{Nt}\alpha_\tau + l_N B_\tau' F_t + V_{Nt}, \tag{4.9}$$

where $t = T, T-1, \ldots, i = 1, \ldots, N$, $Y_{Nt} = (y_{1t}, y_{2t}, \ldots, y_{Nt})'$, $W_N = (w_{ij,N})_{N\times N}$ is a nonstochastic time-invariant spatial weights matrix, $Z_{Nt} = (z_{1t}, \ldots, z_{Nt})' \in \mathbb{R}^{N\times p_z}$ is the exogenous regressor matrix, $\gamma_{0\tau}$ is the intercept term, $l_N = (1, \ldots, 1)'$ is N-dimensional, $f_t = (f_{t1}, \ldots, f_{tm})' \in \mathbb{R}^{m\times 1}$ is a vector of time-varying common factors and $F_t = \left(f_t', \ldots, f_{t-k}'\right)' \in \mathbb{R}^{(k+1)m\times 1}$, $V_{Nt} = (v_{1t}, \ldots, v_{Nt})'$ is the disturbance term and v_{it}'s are independent, and $\gamma_{0\tau}, \gamma_{1\tau}, \gamma_{2\tau}, \gamma_{3\tau} \in \mathbb{R}$, $\alpha_\tau \in \mathbb{R}^{p_z}$, and $B_\tau = (\beta_{0\tau}', \ldots, \beta_{k\tau}')' \in \mathbb{R}^{(k+1)m\times 1}$ are true model parameters. We write $x_{it} = (1, z_{i,t}', \overline{Y}_{i,t-1}, Y_{i,t-1}, F_t')'$ and $\phi_\tau = (\gamma_{0\tau}, \alpha_\tau', \gamma_{\tau 2}, \gamma_{\tau 3}, B_\tau')'$, where $\overline{Y}_{i,t-1} = W_{i\cdot,N} Y_{N,t-1}$. We take the instrumental variable as $r_{it} = (e_i' W_N^2 Y_{N,t-1}, e_i' W_N^3 Y_{N,t-1})' \in \mathbb{R}^2$, where $e_i \in \mathbb{R}^N$ is a column vector with unity on the ith entry and zeros otherwise. We write $\Psi_{it} \equiv (x_{it}', r_{it}')'$, $u_{it} \equiv y_{it} - \gamma_{1\tau}\overline{Y}_{it} - x_{it}'\phi_\tau$ and $s_{it} \equiv \psi_\tau(u_{it}) \cdot \Psi_{it}$, where $\overline{Y}_{it} \equiv W_{i\cdot,N} Y_{Nt}$ and $\psi_\tau(\cdot) \equiv \tau - 1(\cdot \le 0)$. In Xu et al. (2022), they explore the asymptotic theory for the instrumental variable quantile regression (IVQR) estimator by establishing the NED property of $\{y_{it}\}$. Here, we establish the CLT for $\{s_{it}\}$, a crucial step in establishing the asymptotic normality of the IVQR estimator by using FD.

Assumption 12. *Let $\mathcal{C} \equiv \vee_{t=-\infty}^{\infty} \sigma(z_{1t}', \ldots z_{Nt}', F_t')$.*

(1) $\sup_N \|W_N\|_\infty = 1$ *and* $|\gamma_{1\tau}| + |\gamma_{2\tau}| + |\gamma_{3\tau}| < 1$*. We write* $\zeta \equiv \frac{|\gamma_{2\tau}| + |\gamma_{3\tau}|}{1 - |\gamma_{1\tau}|} < 1$*.*

(2) $\|v\|_{L^p,\mathcal{C}} \equiv \sup_{N,T,i,t} \|v_{it}\|_{L^p,\mathcal{C}} < \infty$ *a.s. for some* $p > 2$*;* $|\gamma_{0\tau}| + \sup_{N,T,i,t} |z_{it}'\alpha_\tau| \le d_z < \infty$*; and* $\|B_\tau\|_\infty \sup_{T,t} \|F_t\|_1 \le d_f < \infty$*.*

(3) $|w_{ij,N}| \le c d_{ij}^{-\alpha}$ *for some constants* $c > 0$ *and* $\alpha > \frac{3}{2}d + \frac{1}{2}$*.*

(4) Conditional on $\mathcal{C}$*,* v_{it}*'s are independent over* i *and* t*.*

(5) $\Omega \equiv \tau(1-\tau) \lim_{N,T\to\infty} (NT)^{-1} \sum_{i=1}^N \sum_{t=1}^T \mathbb{E}\left(\Psi_{it}\Psi_{it}' \mid \mathcal{C}\right)$ *is nonsingular a.s.*

[9] Here, we use a slightly different form, but it is equivalent to the one in Xu et al. (2022).

(6) $\mathbb{P}(v_{it} \leq 0 \mid \mathcal{C}, x_{it}, r_{it}) = \tau$ *a.s.*

Assumptions 12(1)-(4) directly inherit Assumptions 8-11. Similar assumptions as those in Assumption 12 are also employed in Assumptions 2.1 and 3.2 and Theorem 3 in Xu et al. (2022). Xu et al. (2022) require $\alpha > 2d + 1$ in Assumption 12(3) to establish NED, so our assumption is less restrictive. Then we have the following CLT.

Proposition 4.8. *For model (4.9), let* $G_{NT} = \sum_{t=1}^{T} \sum_{i=1}^{N} s_{it}$*. Under Assumptions 1 and 12,*

$$\Omega^{-1/2} \frac{G_{NT} - \mathbb{E}_{\mathcal{C}} G_{NT}}{\sqrt{NT}} \xrightarrow{d} N(0, I).$$

To conduct inference, one must estimate Ω consistently. A natural estimator is

$$\hat{\Omega} = \frac{\tau(1-\tau)}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \Psi_{it} \Psi_{it}',$$

and it is a consistent estimator for Ω by the conditional LLN under functional dependence (Theorem D.1).

Proposition 4.9. *For model (4.9), under Assumptions 1 and 12,* $\hat{\Omega} \xrightarrow{p} \Omega$.

Remark. Though the spatial weights matrices (W_n or W_N) considered in this section are nonstochastic, our theory can also accommodate stochastic matrices. In Section S.8 in the online supplement, we discuss the spatial FD properties of more examples, including the models with a stochastic (or even possibly endogenous) spatial weights matrix. All the examples in this section and Section S.8 share a similar structure: the right hand sides of these data generating processes (e.g., (4.3), (4.4), and (4.7)) are all Lipschitz functions of the spatial interaction term ($W_n Y_n$ or $W_N Y_{Nt}$) and the right-hand-side function is a contraction mapping of Y_n or Y_{Nt}, which is preserved under our assumptions (e.g., Assumption 2). The Lipschitz and contraction mapping properties are vital for condition (4.2) to hold such that the general criteria (Propositions 4.1 and 4.2) are applicable to establish the spatial FD property. When the Lipschitz and contraction mapping properties do not hold, e.g., the right hand side is an indicator function (when $Y_{i,n}$ is discrete), the general criteria are not applicable and we have to resort to other methods to establish the spatial FD property of $Y_{i,n}$. Finally, the establishment of the FD property does not require the coefficients in the models to be homogeneous (i.e., identical for all individuals). We allow for individual heterogeneity in the models (e.g., the functional-coefficient SAR model in Sun (2016), the smooth-coefficient SAR model in Malikov and Sun (2017) and the heterogeneous SDPD model in Su et al. (2023)). See Section S.8 in the online supplement for more information.

5. Transformations of Spatial Stable Processes

In this section, we investigate the FDM and the FD coefficient under various transformations. In applications, estimators and testing statistics are certain functions of the data. Thus, one needs to calculate the FD coefficients of those estimators and testing statistics to employ the tools in Section 3. Since $\Delta_p(s) \equiv \sup_n \sup_{i\in D_n} \delta_p\left(i, \{j \in D_n : d_{ij} \geq s\}, n\right)$, it suffices to consider the properties of L^p-FDM $\delta_p(i, I, n)$ under various transformations.[10] Throughout this section, denote the L^p-FDM ($p \geq 1$) of the random field $\{Y_{i,n}\}$ ($\{Z_{i,n}\}$ or $\{X_{i,n}\}$) over an independent random field $\{\epsilon_{i,n}\}$ by $\delta_{Y,p}(i, I, n)$ ($\delta_{Z,p}(i, I, n)$ or $\delta_{X,p}(i, I, n)$).

First, we consider a family of functions $H_{i,n} : \mathbb{R}^{p_Y} \to \mathbb{R}^{p_Z}$ satisfying the following condition: for all $y, y^\bullet \in \mathbb{R}^{p_Y}$,

$$\left\| H_{i,n}\left(y\right) - H_{i,n}\left(y^\bullet\right)\right\| \leq B_{i,n}\left(y, y^\bullet\right) \left\|y - y^\bullet\right\|. \tag{5.1}$$

We write $Z_{i,n} \equiv H_{i,n}\left(Y_{i,n}\right)$ in Propositions 5.1-5.3. When $B_{i,n}$ is bounded by a constant C, from the following proposition, we have $\delta_{Z,p}(i, I, n) \leq C\delta_{Y,p}(i, I, n)$. And NED shares a similar property.

Proposition 5.1. *If $\sup_{n,i} \sup_{y,y^\bullet} B_{i,n}\left(y, y^\bullet\right) \leq C < \infty$ in (5.1) for some constant C, then $\delta_{Z,p}\left(i, I, n\right) \leq C\delta_{Y,p}\left(i, I, n\right)$ for any $p \geq 1$, $i \in D_n$, $I \subset D_n$, $n \geq 1$.*

[10] The $\theta_{m,p,\iota}$ defined in Appendix B is also a special case of $\delta_p(i, I, n)$ with $I = I_{i,m,\iota}$. Thus, the properties of FDM $\delta_p(i, I, n)$ under transformations are also applicable for $\theta_{m,p,\iota}$.

When $B_{i,n}(y, y^\bullet)$ is unbounded, e.g., $H_{i,n}(x) = x^2$, the following two propositions summarize the corresponding results.

Proposition 5.2. *We suppose $B_{i,n}(y, y^\bullet) \leq C_1 (\|y\|^a + \|y^\bullet\|^a + 1)$ in (5.1) for some constants $a > 0$ and $C_1 < \infty$. The constants $p, q, r \geq 1$ satisfy $p^{-1} = q^{-1} + r^{-1}$. If $\|Y\|_{L^{ar}} \equiv \sup_{n, i \in D_n} \|Y_{i,n}\|_{L^{ar}} < \infty$, then $\delta_{Z,p}(i, I, n) \leq C_1 (2 \|Y\|^a_{L^{ar}} + 1) \delta_{Y,q}(i, I, n)$ for all $i \in D_n$, $I \subset D_n$, and $n \geq 1$.*

We note that there is a trade-off between p, q, and r. If we want a larger p, then a larger q or a larger r is needed. In the NED case (Lemma A.4, Xu and Lee, 2015a), p is restricted to be 2, but here p can be any number greater than or equal to 1. In Proposition 5.2, the L^p-FD of $\{Z_{i,n}\}$ is preserved by the L^q-FD of $\{Y_{i,n}\}$ for some $q > p$. In fact, when $\{Y_{i,n}\}$ is L^p-FD, $\{Z_{i,n}\}$ might also be L^p-FD, as can be seen in the following proposition.

Proposition 5.3. *We suppose $B_{i,n}(y, y^\bullet) \leq C_1 (\|y\|^a + \|y^\bullet\|^a + 1)$ in (5.1) for some constants $a \geq 1$ and $C_1 < \infty$, and $\|Y\|_{L^q} \equiv \sup_{n, i \in D_n} \|Y_{i,n}\|_{L^q} < \infty$ for some q satisfying $q > (a+1)p$ and $q \geq \frac{ap}{p-1}$, where $p > 1$ is a constant. Then for any $i \in D_n$ and $I \subset D_n$, there exists a constant $C_2 > 0$[11] such that*

$$\delta_{Z,p}(i, I, n) \leq C_2 \{\delta_{Y,p}(i, I, n)\}^{(q-ap-p)/(pq-ap-p)}. \tag{5.2}$$

Let us compare Propositions 5.2 and 5.3. Suppose that we want to establish a CLT for $\{Z_{i,n}\}$. If we employ Proposition 5.2, by Theorem 3.5, we need $\Delta_{Y,q}(s) = O(s^{-d/2})$ for some $q > 2$. If instead we employ Proposition 5.3, we need $\Delta_{Y,2}(s) = O(s^{-\kappa})$ for some $\kappa > \frac{d(pq-ap-p)}{2(q-ap-p)}$ as $s \to \infty$. Since $\frac{pq-ap-p}{q-ap-p} > 1$, we require a faster decreasing rate of $\Delta_{Y,2}(s)$ than that of $\Delta_{Y,q}(s)$ when we use Proposition 5.2. The price to employ Proposition 5.2 is a higher order FD coefficient, i.e, $q > 2$.

Next, we consider a discontinuous nonlinear transformation, $1(\cdot > 0)$, which is widely used to study binary data and censored data.

[11]Here, C_2 might depend on p. If one wants to establish an exponential inequality for $Z_{i,n}$, one must refer to the proof of this proposition to determine how C_2 depends on p. This is also the case for Proposition 5.7.

Proposition 5.4. *We write* $Z_{i,n} \equiv 1(Y_{i,n} > 0)$ *and suppose the probability density functions of* $\{Y_{i,n}\}$ *are uniformly bounded in* i *and* n*. Then, for any* $p \geq 1$*,* $i \in D_n$*, and* $I \subset D_n$*, there exists a constant* $C > 0$ *not depending on* p*,* i*,* I*, or* n*, such that*

$$\delta_{Z,p}(i, I, n) \leq C \{\delta_{Y,p}(i, I, n)\}^{1/(p+1)}.$$

We suppose $Y_{i,n}$ and $Z_{i,n}$ are real-valued in the following. In applications, one usually needs to deal with the summation or product of $Y_{i,n}$ and $Z_{i,n}$. The case of summation is a direct result of Minkowski's inequality, and thus we omit its proof.

Proposition 5.5. *The* L^p*-FDM of* $\{Y_{i,n} + Z_{i,n}\}$ *satisfies* $\delta_{Y+Z,p}(i, I, n) \leq \delta_{Y,p}(i, I, n) + \delta_{Z,p}(i, I, n)$ *for any* $i \in D_n$ *and* $I \subset D_n$ *and* $p \geq 1$*.*

The case of product is more complicated. Like Propositions 5.2-5.3, we also have two results. We write $X_{i,n} \equiv Y_{i,n} Z_{i,n}$ in the following two propositions.

Proposition 5.6. *We suppose* $\{Y_{i,n} \in \mathbb{R}\}$ *and* $\{Z_{i,n} \in \mathbb{R}\}$ *are two random fields on the independent random field* $\{\epsilon_{i,n}\}$ *with* $\|Y\|_{L^{r_2}} = \sup_{n,i} \|Y_{i,n}\|_{L^{r_2}} < \infty$ *and* $\|Z\|_{L^{r_1}} = \sup_{n,i} \|Z_{i,n}\|_{L^{r_1}} < \infty$*, where* $r_1, r_2 > 1$*. Let* $p, q_1, q_2 > 1$ *be constants and* $p^{-1} = q_1^{-1} + r_1^{-1} = q_2^{-1} + r_2^{-1}$*. Then, the FDM of* $\{X_{i,n}\}$ *on* $\{\epsilon_{i,n}\}$ *satisfies* $\delta_{X,p}(i, I, n) \leq \|Z\|_{L^{r_1}} \delta_{Y,q_1}(i, I, n) + \|Y\|_{L^{r_2}} \delta_{Z,q_2}(i, I, n)$ *for any* $i \in D_n$ *and* $I \subset D_n$*.*

As Proposition 5.2, Proposition 5.6 employs higher order FDMs to calculate L^p-FDM, i.e., q_1 and q_2 are both greater than p. We can avoid higher order FDMs with the help of the following proposition. But the price is that the decay rate of the L^p-FDM is slower.

Proposition 5.7. *We suppose* $\sup_{n,i} \|Y_{i,n}\|_{L^q} < \infty$ *and* $\sup_{n,i} \|Z_{i,n}\|_{L^q} < \infty$ *for some* q *satisfying* $q > 2p$ *and* $q \geq \frac{p}{p-1}$*, where* $p > 1$ *is a constant. Then, for any* $i \in D_n$ *and* $I \subset D_n$*, there exist constants* $C_1, C_2 > 0$ *such that*

$$\delta_{X,p}(i, I, n) \leq C_1 \{\delta_{Y,p}(i, I, n)\}^{(q-2p)/(pq-2p)} + C_2 \{\delta_{Z,p}(i, I, n)\}^{(q-2p)/(pq-2p)}.$$

In practice, we can use either Proposition 5.6 or 5.7 depending on different conditions. An advantage of Proposition 5.7 is that the p in the $\delta_p(i, I, n)$ coefficient of three or even more random fields could be kept unchanged. However, since $\frac{q-2p}{pq-2p} < 1$, compared to using Proposition 5.6, we need faster rates for $\delta_{Y,p}(i, I, n)$ and $\delta_{Z,p}(i, I, n)$ ($\Delta_{Y,p}(s)$ and $\Delta_{Z,p}(s)$) to establish the inequalities and limit theorems in Section 3 for $\{X_{i,n}\}$ when we apply Proposition 5.7.

6. Comparison of Functional Dependence and NED

In this section, we compare spatial FD and spatial NED thoroughly. Spatial NED was proposed by Jenish and Prucha (2012). For the convenience of reference, we review its definition first.

Definition. *For some $p \geq 1$, let $Z = \{Z_{i,n}, i \in D_n, n \geq 1\}$ and $\epsilon = \{\epsilon_{i,n}, i \in D_n, n \geq 1\}$ be two random fields with $\|Z_{i,n}\|_{L^p} < \infty$, and D_n satisfies Assumption 1. The random field Z is said to be uniformly L^p-NED on ϵ if $\|Z_{i,n} - \mathbb{E}(Z_{i,n}|\mathcal{F}_{i,n}(s))\|_{L^p} \leq C\psi(s)$ for some constant C and some sequence $\psi(s) \geq 0$ with $\lim_{s\to\infty}\psi(s) = 0$, where $\mathcal{F}_{i,n}(s) \equiv \sigma(\epsilon_{j,n} : d_{ij} < s)$ denotes the sub-σ-field generated by the $\epsilon_{j,n}$'s located within the open ball centering at $i \in D_n$ and of radius s. The C is called the NED scaling factor. The $\psi(s)$ is called the NED coefficient and can be without loss of generality (w.l.o.g.) assumed to be nonincreasing.*

The idea of NED is that if every spatial unit is mainly affected by its close neighbors, while spatial functional dependence means that the effects of faraway spatial units are negligible. The ideas of these two concepts are similar. Hence, it is natural to ask whether there is any relationship between them. We answer this question in the following theorem.

Theorem 6.1. *(1) If $\{Y_{i,n}\}$ is L^p-FD on an independent random field $\{\epsilon_{i,n}\}$, i.e., $\lim_{s\to\infty}\Delta_p(s) = 0$, then $\{Y_{i,n}\}$ is uniformly L^p-NED on $\{\epsilon_{i,n}\}$ with the NED scaling factor $C = 1$ and the NED coefficient $\psi(s) \leq \Delta_p(s)$.*

(2) If $\{Y_{i,n}\}$ is L^p-NED on an independent random field $\{\epsilon_{i,n}\}$ with the NED scaling factor $C = 1$, i.e., $\lim_{s\to\infty}\psi(s) = 0$, and in addition $Y_{i,n} = \sum_{j\in D_n} w_{ij,n}\epsilon_{j,n}$ for any $i \in D_n$,

where $w_{ij,n}$'s are nonstochastic coefficients, then $\{Y_{i,n}\}$ is L^p-FD on $\{\epsilon_{i,n}\}$ with the L^p-FD coefficient $\Delta_p(s) \leq 2\psi(s)$.

Therefore, L^p-FD implies L^p-NED and they are equivalent when $Y_{i,n}$ is a linear process in $\epsilon_{j,n}$'s. Now, all properties of a NED random field on an independent random field $\{\epsilon_{i,n}\}$ also hold for spatial functional dependent processes, e.g., the following covariance inequality. This implies that $|\mathrm{Cov}(Y_{i,n}, Y_{j,n})|$ decreases to 0 as d_{ij} increases to ∞. In other words, $Y_{i,n}$ is mainly correlated with those $Y_{j,n}$'s of close neighbors.

Corollary 6.1. *Under Assumption 1, if (i) $\|Y\|_{L^2} \equiv \sup_{n,i} \|Y_{i,n}\|_{L^2} < \infty$, and (ii) $\{Y_{i,n}\}$ is L^2-FD on an independent random field $\{\epsilon_{i,n}\}$, i.e., $\lim_{s\to\infty} \Delta_2(s) = 0$, then for all $i \neq j \in D_n$ and $0 < s \leq \frac{d_{ij}}{2}$, $|\mathrm{Cov}(Y_{i,n}, Y_{j,n})| \leq 2\|Y\|_{L^2} \Delta_2(s)$.*

Though FD implies NED when $\epsilon_{i,n}$'s are independent, we note that FD is not only a special case of NED, but a more powerful and convenient weak dependence concept. Here, we summarize the advantages of FD over NED.

1. Spatial FD is more convenient to calculate than spatial NED, especially when we need to deal with nonlinear transformations. When we need to deal with various nonlinear transformations, in many cases, only L^2-NED is convenient. This is because the definition of NED involves a conditional expectation, and the conditional expectation is the best predictor under L^2-distance. This property is widely used in the proofs about NED under nonlinear transformations. See, e.g., Lemmas A.2 and A.4 in Xu and Lee (2015a). However, the conditional expectation is not needed to calculate L^p-FDMs or L^p-FD coefficients. So, we can usually obtain the L^p-FDM conveniently for any $p \geq 1$ under suitable conditions; and the L^p-FD property can be preserved under various transformations, as can be seen from Section 5.

2. As shown in Theorems 3.2 and 3.5, compared to using NED, it usually requires weaker conditions to establish a CLT and an exponential inequality by using FDM. For CLT, it only

requires the L^2-FD coefficient to decrease slightly faster than $s^{-d/2}$; however, it requires L^2-NED coefficient to decrease slightly faster than s^{-d}. The exponential inequality under FDM enjoys both less restrictive conditions and a faster decay rate, as discussed in Section 3.1.2.

3. Compared to NED, weaker conditions are needed to establish FD properties. For example, in Case 1 of Assumption 3.2 in Xu et al. (2022), in addition to the condition that $|w_{ij}| \leq c d_{ij}^{-\alpha}$, another condition about the column sums of W_n is needed, which is not needed in our paper (see Assumption 12(3)).

Due to these reasons, we believe that spatial FDM is a more powerful and convenient weak dependence concept than NED for theoretical studies in spatial econometrics.

7. Conclusion

In this paper, we generalize the concept of functional dependence proposed in Wu (2005) to the spatial FD to fit the common settings in spatial econometrics. We establish a moment inequality, an exponential inequality, a Nagaev-type inequality, a law of large numbers, and a central limit theorem such that they can be employed in future studies in spatial econometrics. We verify the concepts for a nonlinear SAR model, a threshold SAR model and an SDPD model. Furthermore, we establish different conditions to preserve the spatial FD property under various transformations. We compare spatial FD with the spatial NED proposed by Jenish and Prucha (2012), and illustrate its advantages over the spatial NED.

There are some future research directions. (1) If a better strategy can be found to prove Theorem B.1 such that the term $\iota_m^{d/2}$ in the definition of the second-type L^p-FD coefficient $\Theta_{s,p,\iota} \equiv \sum_{m=s}^{\infty} \iota_m^{d/2} \theta_{m,p,\iota}$ can be dropped, and some conditions in our theoretical results can be relaxed. (2) We are working on relaxing the assumption that individuals are located in a Euclidean space such that the FD theory can be applied to more general network data. (3) We are applying the tools developed in this paper to study the quantile regression of spatial econometric models.

Appendices

A. Two Lemmas for CLT

Lemma A.1. *(CLT for spatially m-dependent triangular array). Let $m \geq 0$ be fixed. $\{X_{i,n}, i \in T_n, n \geq 1\}$ is a spatially m-dependent zero-mean triangular array (i.e., $X_{i,n}$ and $X_{j,n}$ are independent when $d_{ij} \geq m$), where T_n satisfies Assumption 1. And $\lim_{k\to\infty} \sup_{n,i\in T_n} \mathbb{E}[X_{i,n}^2 1(|X_{i,n}| > k)] = 0$, i.e., $X_{i,n}$'s are uniformly L^2-integrable. Denote $S_n \equiv \sum_{i\in T_n} X_{i,n}$ and $\sigma_n^2 \equiv \operatorname{Var}(S_n)$. Assume $B \equiv \liminf_{n\to\infty} |T_n|^{-1} \sigma_n^2 > 0$. Then*

$$\frac{S_n}{\sigma_n} \xrightarrow{d} N(0,1).$$

Proof. Since $\{X_{i,n}, i \in T_n, n \geq 1\}$ is spatially m-dependent, its ϕ-mixing coefficients $\bar{\phi}_{k,l}(r) = 0$ for all $k, l \in \mathbb{N}$ when $r > m$. Thus, Assumption 4 in Jenish and Prucha (2009) is satisfied. Since Assumptions 1, 2 and 5 in Jenish and Prucha (2009) are also satisfied, by Theorem 1(b) in Jenish and Prucha (2009), $\frac{S_n}{\sigma_n} \xrightarrow{d} N(0,1)$. ■

Lemma A.2. *(Proposition 6.3.9 in Brockwell and Davis, 1991). Let W_n, $n = 1, 2, \ldots$ and U_{ns}, $s = 1, 2, \ldots$, be random vectors such that (1) $U_{ns} \xrightarrow{d} U_s$ as $n \to \infty$ for each $s = 1, 2, \ldots$; (2) $U_s \xrightarrow{d} U$ as $s \to \infty$; (3) $\lim_{s\to\infty} \limsup_{n\to\infty} \mathbb{P}(|W_n - U_{ns}| > \epsilon) = 0$ for every $\epsilon > 0$. Then $W_n \xrightarrow{d} U$ as $n \to \infty$.*

B. Second-type Functional Dependence Coefficient

In this section, we introduce the second-type L^p-functional dependence coefficient, which is mainly used to develop our theory. In this paper, when we mention an L^p-FD coefficient without "second-type", we refer to the L^p-FD coefficient in Definition 2.2. To begin with, we define $\mathscr{I} \equiv \{\iota = (\iota_0, \iota_1, \ldots) : \iota_0 = 0, \iota_m > \iota_{m-1}, \iota_m \in \mathbb{N} \text{ for all } m \geq 1\}$ to be the set of all strictly increasing integer-valued sequences ι starting at $\iota_0 = 0$. The proofs of this section are collected in Section S.3.

Definition B.1 (The second-type L^p-functional dependence coefficient). *For any $p \geq 1$, $m \in \mathbb{N}$, and $\iota \in \mathscr{I}$, denote $I_{i,m,\iota} = \{j \in D_n : d_{ij} \in [\iota_{m-1}, \iota_m)\}$ and*

$$\theta_{m,p,\iota} \equiv \sup_n \sup_{i \in D_n} \delta_p\left(i, I_{i,m,\iota}, n\right) = \sup_n \sup_{i \in D_n} \left\|Y_{i,n} - Y_{i,n,I_{i,m,\iota}}\right\|_{L^p}.$$

For any $s \in \mathbb{N}$, the second-type L^p-functional dependence coefficient is defined as

$$\Theta_{s,p,\iota} \equiv \sum_{m=s}^{\infty} \iota_m^{d/2} \theta_{m,p,\iota},$$

and denote $\Theta_{p,\iota} \equiv \Theta_{1,p,\iota} = \sum_{m=1}^{\infty} \iota_m^{d/2} \theta_{m,p,\iota}$.

In Definition B.1, $I_{i,m,\iota}$ is the set of individuals whose distance to i is within $[\iota_{m-1}, \iota_m)$, which can be regarded as a ring in $\mathbb{R}^d$. Therefore, $\delta_p\left(i, I_{i,m,\iota}, n\right)$ measures the impact of $\epsilon_{j,n}$'s in this ring on $Y_{i,n}$, and $\theta_{m,p,\iota}$ is its supremum over i and n. $\Theta_{s,p,\iota} \equiv \sum_{m=s}^{\infty} \iota_m^{d/2} \theta_{m,p,\iota}$ is a weighted sum of $\theta_{m,p,\iota}$ with $m \geq s$, measuring the total impact of $\epsilon_{j,n}$'s with distance $d_{ij} \geq \iota_{m-1}$. Thus, $\Theta_{s,p,\iota}$ decreases as the distance s increases.

Definition B.1 is motivated by Wu (2005), El Machkouri et al. (2013), Liu et al. (2013), and Wu and Wu (2016). They define $\Theta_p \equiv \sum_{m=1}^{\infty} \theta_{m,p,\tilde{\iota}}$, where $\tilde{\iota} = (0, 1, 2, \ldots)$. Their $\tilde{\iota}$ is a special case of ours. Using various ι's, we can improve some of our theoretical results.[12] Notice that they do not have the term $\iota_m^{d/2}$, but this term is essential to establish the moment inequality in our setup (see the proof of Theorem B.1).

We now employ $\theta_{m,p,\iota}$ and $\Theta_{s,p,\iota}$ to establish a moment inequality and an exponential inequality, which will lead to Theorems 3.1 and 3.2. To start with, we first give a crucial lemma.

Lemma B.1. *For system (2.1), let $\mathcal{F}_{i,n}(s) \equiv \sigma\left(\epsilon_{j,n} : d_{ij} < s\right)$ denote the sub-σ-field generated by the $\epsilon_{j,n}$'s located within the open ball centering at $i \in D_n$ and of radius s. Denote $V_{i,n,\iota}(m) \equiv$*

[12] We will elaborate on this at the end of this section.

$\mathbb{E}\left(Y_{i,n}|\mathcal{F}_{i,n}\left(\iota_m\right)\right) - \mathbb{E}\left(Y_{i,n}|\mathcal{F}_{i,n}\left(\iota_{m-1}\right)\right)$. *Then for any $i \in D_n$, $m \in \mathbb{N}$, $p \geq 1$ and $\iota \in \mathscr{I}$, we have*

$$\|V_{i,n,\iota}(m)\|_{L^p} \leq \theta_{m,p,\iota}. \tag{B.1}$$

In the following of this section, let T_n be a finite subset of D_n such that $|T_n| \to \infty$ as $n \to \infty$, $S_n \equiv \sum_{i\in T_n} Y_{i,n}$ and $Z_n \equiv S_n/\sqrt{|T_n|}$. The moment inequality is stated as follows.

Theorem B.1. *Under Assumption 1, if $\Theta_{p,\iota} < \infty$ for some $p \geq 2$ and $\iota \in \mathscr{I}$, then*

$$\left\|\sum_{i\in T_n}\left(Y_{i,n} - \mathbb{E}Y_{i,n}\right)\right\|_{L^p} \leq 2^d\sqrt{p-1}\,\Theta_{p,\iota}\left|T_n\right|^{1/2}. \tag{B.2}$$

The main strategy to prove Theorem B.1 is to decompose every $(Y_{i,n} - \mathbb{E}Y_{i,n})$ as a summation of a martingale difference array $\{V_{i,n,\iota}(m)\}_{m=1}^{\infty}$ and apply Lemma B.1 to bound the L^p-norm of the mth element of the martingale difference array by $\theta_{m,p,\iota}$. An application of Theorem B.1 is the following exponential inequality.

Theorem B.2. *Under Assumption 1, if (i) $\mathbb{E}Y_{i,n} = 0$ for any $i \in T_n$, (ii) for any real number $p \geq 2$, there exists a sequence $\iota^{(p)} \in \mathscr{I}$ such that $\Theta_{p,\iota^{(p)}} < \infty$, and (iii)*

$$\gamma_0 \equiv \sup_{p\geq 2} p^{-\nu}\Theta_{p,\iota^{(p)}} < \infty, \tag{B.3}$$

then for $\alpha = \frac{2}{1+2\nu}$ and for all $t \in [0, t_0)$, we have

$$m\left(t\right) \equiv \mathbb{E}\left[\exp\left(t\left|Z_n\right|^{\alpha}\right)\right] \leq 1 + c_\alpha\left(1 - \frac{t}{t_0}\right)^{-1/2}\frac{t}{t_0},$$

where $t_0 = \left(e\alpha\gamma_0^\alpha 2^{\alpha d}\right)^{-1}$ and c_α is a constant depending only on α. Hence, for any $\epsilon > 0$, by taking $t = t_0/2$, we have

$$\mathbb{P}\left(\left|S_n\right| \geq \left|T_n\right|\epsilon\right) \leq \left(1 + \frac{\sqrt{2}c_\alpha}{2}\right)\exp\left(-\frac{\left|T_n\right|^{1/(1+2\nu)}\epsilon^{2/(1+2\nu)}}{2^{\alpha d+1}e\alpha\gamma_0^\alpha}\right).$$

Condition (B.3) is similar to (2.21) in Wu and Wu (2016). It assumes that $\Theta_{p,\iota^{(p)}}$ increases slower than Cp^{ν} for some $\nu \geq 0$ as $p \to \infty$. As mentioned in Wu and Wu (2016), γ_0 can be regarded as a dependence-adjusted norm.

Next, we summarize the relations between the two types of L^p-FD coefficients in Lemmas B.2-B.5. Lemmas B.3-B.5 are the keys to transfer the properties of $\theta_{m,p,\iota}$ and $\Theta_{s,p,\iota}$ to the properties of $\Delta_p(s)$ in Section 3 and they will be used in the proofs of Theorems 3.1-3.5. In the following lemmas, $\Delta_p(s)$ denotes the L^p-FD coefficient of $\{Y_{i,n}\}$ on $\{\epsilon_{i,n}\}$.

Lemma B.2. *For any $p \geq 1$, $m \geq 1$, and $\iota \in \mathscr{I}$, we have $\theta_{m,p,\iota} \leq 3\Delta_p(\iota_{m-1})$. Immediately,*

$$\Theta_{s,p,\iota} \leq 3 \sum_{m=s}^{\infty} \iota_m^{d/2} \Delta_p(\iota_{m-1}) \quad \text{and} \quad \Theta_{p,\iota} \leq 3 \sum_{m=1}^{\infty} \iota_m^{d/2} \Delta_p(\iota_{m-1}).$$

Lemma B.3. *If $\{Y_{i,n}\}$ is L^1-FD on $\{\epsilon_{i,n}\}$, then $\lim_{s\to\infty} \sum_{m=s}^{\infty} \theta_{m,1,\iota} = 0$ for some $\iota \in \mathscr{I}$.*

Lemma B.4. *For any $p \geq 1$, if $\Delta_p(0) < \infty$ and $\Delta_p(s) = O\left(s^{-\kappa}\right)$ as $s \to \infty$ for some $\kappa > \frac{d}{2}$, then $\Theta_{p,\iota} < \infty$ and $\Theta_{s,p,\iota} = o\left(s^{-1}\right)$ as $s \to \infty$ for some $\iota \in \mathscr{I}$.*

Lemma B.5. *If $\{Y_{i,n}\}$ is L^p-FD on $\{\epsilon_{i,n}\}$ for any $p \geq 2$ with $\Delta_p(s) \leq O\left(p^{\nu}\right) O\left(s^{-\kappa}\right)$ for some $\kappa > \frac{d}{2}$ and $\nu \geq 0$ as $p \to \infty$ and $s \to \infty$, where $O\left(p^{\nu}\right)$ does not depend on s and $O\left(s^{-\kappa}\right)$ does not depend on p, then $\gamma_0 \equiv \sup_{p\geq 2} p^{-\nu} \Theta_{p,\iota} < \infty$ for some $\iota \in \mathscr{I}$.*

Finally, we illustrate how various ι's can improve our results. Take Theorem B.1 as an example. If we fix ι as $\iota^* = (0, 1, 2, \ldots)$, then $\Theta_{p,\iota^*} = \sum_{m=1}^{\infty} m^{d/2} \theta_{m,p,\iota^*} \leq \sum_{m=1}^{\infty} m^{d/2} \Delta_p(m-1)$ by Lemma B.2. To establish the moment inequality, we need the condition $\Delta_p(s) = O(s^{-d/2-1-\delta})$ for some $\delta > 0$ to ensure $\Theta_{p,\iota^*} < \infty$. However, from Lemma B.4, whenever $\Delta_p(s) = O(s^{-d/2-\delta})$ for some $\delta > 0$, we have $\Theta_{p,\iota} < \infty$ for some ι. Similar improvement also appears in the proofs of the LLN, the CLT and the exponential inequality.

C. Proofs for Section 3.2

The proofs in this section rely heavily on the theory of the second-type L^p-FD coefficient in Appendix B. Recall $\mathscr{I} \equiv \{\iota = (\iota_0, \iota_1, \ldots) : \iota_0 = 0, \iota_m > \iota_{m-1}, \iota_m \in \mathbb{N} \text{ for all } m \geq 1\}$.

Proof of Theorem 3.4. The idea of the proof is borrowed from that for the LLN in Jenish and Prucha (2012). By Condition (ii) in this theorem and Lemma B.3, there exists a sequence $\iota \in \mathscr{I}$ such that

$$\lim_{s \to \infty} \sum_{m=s}^{\infty} \theta_{m,1,\iota} = 0. \tag{C.1}$$

Recall that $\mathcal{F}_{i,n}(m) = \sigma\left(\epsilon_{j,n} : d_{ij} < m\right)$. For any fixed $s \in \mathbb{N}$, we decompose $Y_{i,n} - \mathbb{E}Y_{i,n}$ as

$$Y_{i,n} - \mathbb{E}Y_{i,n} = \xi_{i,n}^s + \eta_{i,n}^s,$$

where $\xi_{i,n}^s = \mathbb{E}\left(Y_{i,n} | \mathcal{F}_{i,n}\left(\iota_s\right)\right) - \mathbb{E}Y_{i,n}$ and $\eta_{i,n}^s = Y_{i,n} - \mathbb{E}(Y_{i,n} | \mathcal{F}_{i,n}\left(\iota_s\right))$. Therefore, it suffices to show that both $\xi_{i,n}^s$ and $\eta_{i,n}^s$ satisfy an LLN.

(1) Consider $\xi_{i,n}^s$ first. It suffices to show that $\xi_{i,n}^s$ satisfies the assumptions of Theorem 3 in Jenish and Prucha (2009). First, for all $s \geq 1$, $i \in T_n$, and $n \geq 1$, by conditional Jensen's inequality,

$$\sup_{n, i \in T_n} \left\|\xi_{i,n}^s\right\|_{L^p} \leq 2 \sup_{n, i \in T_n} \left\|Y_{i,n}\right\|_{L^p} < \infty.$$

So, $\{\xi_{i,n}^s, i \in T_n, n \in \mathbb{N}\}$ is uniformly L^p-bounded for $p > 1$, and as a result, it is uniformly L^1-integrable. Second, since $\xi_{i,n}^s$ is measurable with respect to $\mathcal{F}_{i,n}\left(\iota_s\right)$ and $\epsilon_{i,n}$'s are independent, $\xi_{i,n}^s$ and $\xi_{j,n}^s$ are independent when $d_{ij} \geq 2\iota_s$. Thus, the α-mixing coefficient $\bar{\alpha}_{\xi^s}(1,1,r)$ of $\xi_{i,n}^s$ will become zero when $r \geq 2\iota_s$, which indicates that $\sum_{m=1}^{\infty} m^{d-1} \bar{\alpha}_{\xi^s}(1,1,m) < \infty$. Therefore, all the conditions in Theorem 3 in Jenish and Prucha (2009) are satisfied for $\xi_{i,n}^s$. So, for each $s \geq 1$,

$$\frac{1}{|T_n|} \left\| \sum_{i \in T_n} \xi_{i,n}^s \right\|_{L^1} \to 0 \text{ as } n \to \infty. \tag{C.2}$$

(2) Next, we will investigate $\eta_{i,n}^s$. Recall $V_{i,n,\iota}(k) \equiv \mathbb{E}\left(Y_{i,n}|\mathcal{F}_{i,n}(\iota_k)\right) - \mathbb{E}\left(Y_{i,n}|\mathcal{F}_{i,n}(\iota_{k-1})\right)$ and note that $\eta_{i,n}^s = \sum_{k=s+1}^{\infty} V_{i,n,\iota}(k)$. Thus,

$$\begin{aligned}
&\frac{1}{|T_n|}\left\|\sum_{i\in T_n}\eta_{i,n}^s\right\|_{L^1} = \frac{1}{|T_n|}\left\|\sum_{i\in T_n}\sum_{k=s+1}^{\infty}V_{i,n,\iota}(k)\right\|_{L^1} \le \frac{1}{|T_n|}\sum_{k=s+1}^{\infty}\sum_{i\in T_n}\|V_{i,n,\iota}(k)\|_{L^1} \\
\le&\frac{1}{|T_n|}\sum_{k=s+1}^{\infty}\sum_{i\in T_n}\theta_{k,1,\iota} = \sum_{k=s+1}^{\infty}\theta_{k,1,\iota} \to 0 \text{ as } s\to\infty,
\end{aligned} \tag{C.3}$$

where the last inequality follows from (B.1) and the last limit follows from (C.1).

Combining (C.2) and (C.3), for all $s \ge 1$, we have

$$\limsup_{n\to\infty}\frac{1}{|T_n|}\left\|\sum_{i\in T_n}(Y_{i,n}-\mathbb{E}Y_{i,n})\right\| \le \limsup_{n\to\infty}\frac{1}{|T_n|}\left\|\sum_{i\in T_n}\xi_{i,n}^s\right\| + \limsup_{n\to\infty}\frac{1}{|T_n|}\left\|\sum_{i\in T_n}\eta_{i,n}^s\right\| \le \sum_{k=s+1}^{\infty}\theta_{k,1,\iota}.$$

By letting $s\to\infty$, we complete the proof. ■

Proof of Theorem 3.5. This proof adopts the strategy employed by Jenish and Prucha (2012) in proving their NED CLT. As this proof is lengthy, we break it up into several parts.

Step 1. Decomposition of $Y_{i,n}$. By Condition (i) in this theorem and Lyapunov's inequality,

$$\Delta_2(0) = \sup_n \sup_{i\in D_n}\left\|Y_{i,n} - Y_{i,n,\{j:d_{ij}\ge 0\}}\right\|_{L^2} \le 2\sup_n \sup_{i\in D_n}\|Y_{i,n}\|_{L^p} < \infty.$$

Together with Condition (iii) in this theorem, by Lemma B.4, there exists a sequence $\iota \in \mathscr{I}$ such that

$$\Theta_{2,\iota} < \infty \text{ and } \Theta_{s,2,\iota} = o(s^{-1}) \tag{C.4}$$

as $s\to\infty$. Now, recall that $\mathcal{F}_{i,n}(m) = \sigma\left(\epsilon_{j,n} : d_{ij} < m\right)$. For any fixed $s\in\mathbb{N}$, we decompose $Y_{i,n} - \mathbb{E}Y_{i,n}$ as follows,

$$Y_{i,n} - \mathbb{E}Y_{i,n} = \xi_{i,n}^s + \eta_{i,n}^s,$$

where $\xi_{i,n}^{s}=\mathbb{E}\left(Y_{i,n}|\mathcal{F}_{i,n}\left(\iota_{s}\right)\right)-\mathbb{E}Y_{i,n}$ and $\eta_{i,n}^{s}=Y_{i,n}-\mathbb{E}(Y_{i,n}|\mathcal{F}_{i,n}\left(\iota_{s}\right))$. Let

$$S_{n,s}=\sum_{i\in T_{n}}\xi_{i,n}^{s},\qquad \tilde{S}_{n,s}=\sum_{i\in T_{n}}\eta_{i,n}^{s},\qquad \sigma_{n,s}^{2}=\operatorname{Var}\left(S_{n,s}\right),\qquad \tilde{\sigma}_{n,s}^{2}=\operatorname{Var}\left(\tilde{S}_{n,s}\right).$$

By the Minkowski inequality and $S_{n}-\mathbb{E}S_{n}=S_{n,s}+\tilde{S}_{n,s}$, we have

$$\sigma_{n}=\left\|S_{n}-\mathbb{E}S_{n}\right\|_{L^{2}}\leq\left\|S_{n,s}\right\|_{L^{2}}+\left\|\tilde{S}_{n,s}\right\|_{L^{2}}=\sigma_{n,s}+\tilde{\sigma}_{n,s}.$$

Similar inequalities hold if we exchange the locations of $\sigma_{n},\sigma_{n,s},\tilde{\sigma}_{n,s}$ in the above inequality, which leads to

$$\left|\sigma_{n}-\sigma_{n,s}\right|\leq\tilde{\sigma}_{n,s}\quad\text{and}\quad\left|\sigma_{n}-\tilde{\sigma}_{n,s}\right|\leq\sigma_{n,s}.\tag{C.5}$$

Now we consider the spatial FDM of $\left\{\eta_{i,n}^{s}:i\in D_{n},n\geq1\right\}$ on $\{\epsilon_{i,n}:i\in D_{n},n\geq1\}$. Recall that $I_{i,m,\iota}=\{j\in D_{n}:d_{ij}\in[\iota_{m-1},\iota_{m})\}$ and denote $\check{\mathcal{F}}_{i,n}\left(\iota_{s}\right)=\sigma\left\{\epsilon_{j,n}:j\in\{d_{ij}<\iota_{s}\}\backslash I_{i,m,\iota}\right\}$, $\check{\mathcal{F}}_{i,m,\iota}=\sigma\left\{\epsilon_{j,n}^{*}:j\in I_{i,m,\iota}\right\}$. Then

$$\begin{cases}\eta_{i,n}^{s}-\eta_{i,n,I_{i,m,\iota}}^{s}=Y_{i,n}-\mathbb{E}\left(Y_{i,n}\left|\mathcal{F}_{i,n}\left(\iota_{s}\right)\right.\right)-Y_{i,n,I_{i,m,\iota}}+\mathbb{E}\left(Y_{i,n,I_{i,m,\iota}}\middle|\check{\mathcal{F}}_{i,n}\left(\iota_{s}\right)\vee\check{\mathcal{F}}_{i,m,\iota}\right), & \iota_{m}\leq\iota_{s},\\ \eta_{i,n}^{s}-\eta_{i,n,I_{i,m,\iota}}^{s}=Y_{i,n}-Y_{i,n,I_{i,m,\iota}} & \iota_{m}>\iota_{s}.\end{cases}$$

Let $\theta_{m,2,\iota}^{s}\equiv\sup_{n}\sup_{i\in D_{n}}\left\|\eta_{i,n}^{s}-\eta_{i,n,I_{i,m,\iota}}^{s}\right\|_{L^{2}}$ and $\Theta_{2,\iota}^{s}\equiv\sum_{m=1}^{\infty}\iota_{m}^{d/2}\theta_{m,2,\iota}^{s}$. When $m\leq s$, i.e., $\iota_{m}\leq\iota_{s}$, because $V_{i,n,\iota}\left(k\right)=\mathbb{E}\left(Y_{i,n}|\mathcal{F}_{i,n}\left(\iota_{k}\right)\right)-\mathbb{E}\left(Y_{i,n}|\mathcal{F}_{i,n}\left(\iota_{k-1}\right)\right)$ and $\eta_{i,n}^{s}=\sum_{k=s+1}^{\infty}V_{i,n,\iota}\left(k\right)$, by Minkowski's inequality,

$$\theta_{m,2,\iota}^{s}\leq\sup_{n}\sup_{i\in D_{n}}\left\{\left\|\eta_{i,n}^{s}\right\|_{L^{2}}+\left\|\eta_{i,n,I_{i,m,\iota}}^{s}\right\|_{L^{2}}\right\}\leq2\sup_{n}\sup_{i\in D_{n}}\left\|\eta_{i,n}^{s}\right\|_{L^{2}}=2\sup_{n}\sup_{i\in D_{n}}\left\|\sum_{k=s+1}^{\infty}V_{i,n,\iota}\left(k\right)\right\|_{L^{2}}$$
$$\leq2\sum_{k=s+1}^{\infty}\sup_{n}\sup_{i\in D_{n}}\left\|V_{i,n,\iota}\left(k\right)\right\|_{L^{2}}\leq2\sum_{k=s+1}^{\infty}\theta_{k,2,\iota},$$

where the last inequality follows from (B.1). When $m > s$, i.e., $\iota_m > \iota_s$, we have

$$\theta^s_{m,2,\iota} = \sup_n \sup_{i \in D_n} \left\| Y_{i,n} - Y_{i,n,I_{i,m,\iota}} \right\|_{L^2} = \theta_{m,2,\iota}.$$

Therefore, by the above two results,

$$\Theta^s_{2,\iota} \equiv \sum_{m=1}^{\infty} \iota_m^{d/2} \theta^s_{m,2,\iota} = \sum_{m=1}^{s} \iota_m^{d/2} \theta^s_{m,2,\iota} + \sum_{m=s+1}^{\infty} \iota_m^{d/2} \theta^s_{m,2,\iota} \leq 2 \sum_{m=1}^{s} \iota_m^{d/2} \sum_{k=s+1}^{\infty} \theta_{k,2,\iota} + \sum_{m=s+1}^{\infty} \iota_m^{d/2} \theta_{m,2,\iota}$$
$$\leq 2 \sum_{m=1}^{s} \sum_{k=s+1}^{\infty} \iota_k^{d/2} \theta_{k,2,\iota} + 2 \sum_{m=s+1}^{\infty} \iota_m^{d/2} \theta_{m,2,\iota} = 2\left(s+1\right) \Theta_{s+1,2,\iota} \to 0 \text{ as } s \to \infty,$$

where the last limit follows from $\Theta_{s,2,\iota} = o(s^{-1})$ in (C.4). Next, from (C.4), Theorem B.1 implies that $\sigma_n = \|S_n - \mathbb{E}S_n\|_{L^2} \leq 2^d \Theta_{2,\iota} \sqrt{|T_n|}$ and

$$\tilde{\sigma}_{n,s} = \left\| \tilde{S}_{n,s} \right\|_{L^2} = \left\| \sum_{i \in T_n} \eta^s_{i,n} \right\|_{L^2} \leq 2^d \Theta^s_{2,\iota} \sqrt{|T_n|}. \tag{C.6}$$

From Condition (ii) in this theorem, $\sigma_n \geq \sqrt{B\left|T_n\right|}$ for all $n \geq N$ (w.l.o.g. set $N = 1$). Consequently,

$$\lim_{s \to \infty} \sup_{n \geq 1} \frac{\tilde{\sigma}_{n,s}}{\sigma_n} \leq \lim_{s \to \infty} \frac{2^d \Theta^s_{2,\iota}}{\sqrt{B}} = 0. \tag{C.7}$$

By (C.5) and (C.7),

$$\lim_{s \to \infty} \limsup_{n \to \infty} \left| 1 - \frac{\sigma_{n,s}}{\sigma_n} \right| \leq \lim_{s \to \infty} \sup_{n \geq 1} \frac{\tilde{\sigma}_{n,s}}{\sigma_n} = 0, \tag{C.8}$$

and

$$C \equiv \sup_{n \geq 1} \sup_{s \in \mathbb{N}} \frac{\sigma_{n,s}}{\sigma_n} < \infty. \tag{C.9}$$

Step 2. Establish CLT for $S_{n,s} = \sum_{i \in T_n} \xi^s_{i,n}$. To do so, we need to show that for any fixed s, $\xi^s_{i,n}$ satisfies the conditions of Lemma A.1. First, $\left\{ \xi^s_{i,n} \right\}$ is $2\iota_s$-dependent because $\xi^s_{i,n}$ is measurable with respect to $\mathcal{F}_{i,n}(\iota_s)$ and $\epsilon_{i,n}$'s are independent. Second, by the conditional Jensen inequality,

$$\sup_{n,i\in T_n}\left\|\xi_{i,n}^s\right\|_{L^p}=\sup_{n,i\in T_n}\left\|\mathbb{E}\left(Y_{i,n}\left|\mathcal{F}_{i,n}\left(\iota_s\right)\right.\right)-\mathbb{E}Y_{i,n}\right\|_{L^p}\leq 2\sup_{n,i\in D_n}\left\|Y_{i,n}\right\|_{L^p}<\infty.$$

So, $\{\xi_{i,n}^s\}$ is uniformly L^p-bounded. Since $p>2$, $\{\xi_{i,n}^s\}$ is also uniformly L^2 integrable. Third, by (C.6), we have $\frac{\tilde{\sigma}_{n,s}}{\sqrt{|T_n|}}\leq 2^d\Theta_{2,\iota}^s$. Since $\lim_{s\to\infty}\Theta_{2,\iota}^s=0$, there exists s_0 such that whenever $s\geq s_0$, $\frac{\tilde{\sigma}_{n,s}}{\sqrt{|T_n|}}\leq 2^d\Theta_{2,\iota}^s\leq\frac{\sqrt{B}}{2}$. Therefore, it follows from (C.5) that for all $s\geq s_0$, $\left(\sigma_n-\tilde{\sigma}_{n,s}\right)/\sqrt{|T_n|}\leq\sigma_{n,s}/\sqrt{|T_n|}$. Hence,

$$\liminf_{n\to\infty}\frac{\sigma_{n,s}}{\sqrt{|T_n|}}\geq\liminf_{n\to\infty}\frac{\sigma_n}{\sqrt{|T_n|}}-\limsup_{n\to\infty}\frac{\tilde{\sigma}_{n,s}}{\sqrt{|T_n|}}\geq\sqrt{B}-\frac{\sqrt{B}}{2}=\frac{\sqrt{B}}{2}>0.$$

Thus, by Lemma A.1, when $s\geq s_0$,

$$\frac{S_{n,s}}{\sigma_{n,s}}\xrightarrow{d}N(0,1)\quad\text{as } n\to\infty. \tag{C.10}$$

Since the value of s_0 does not affect the later analysis, suppose $s_0=1$ in the following w.l.o.g.

Step 3. CLT for $\sigma_n^{-1}\sum_{i\in T_n}\left(Y_{i,n}-\mathbb{E}Y_{i,n}\right)$. Next, we will show that the just established CLT for $\left\{\xi_{i,n}^s\right\}$ can be carried over to $\{Y_{i,n}\}$ by the same argument as in Jenish and Prucha (2012). Denote $W_n=\sigma_n^{-1}\left(S_n-\mathbb{E}S_n\right)$ and $U_{ns}=\sigma_n^{-1}S_{n,s}$. Then $W_n-U_{ns}=\sigma_n^{-1}\tilde{S}_{n,s}$. Condition (3) of Lemma A.2 holds because

$$\lim_{s\to\infty}\limsup_{n\to\infty}\mathbb{P}\left(\left|W_n-U_{ns}\right|>\epsilon\right)=\lim_{s\to\infty}\limsup_{n\to\infty}\mathbb{P}\left(\left|\sigma_n^{-1}\tilde{S}_{n,s}\right|^2>\epsilon^2\right)\leq\lim_{s\to\infty}\limsup_{n\to\infty}\frac{\tilde{\sigma}_{n,s}^2}{\sigma_n^2\epsilon^2}=0, \tag{C.11}$$

where the inequality follows from Markov's inequality and the last limit is due to (C.7). Next, we proceed to show $W_n\xrightarrow{d}U\sim N(0,1)$ by contradiction. In order to do that, let $\mathscr{M}$ be the set of all probability measures on $(\mathbb{R},\mathcal{B})$, and observe that we can metricize $\mathscr{M}$ by, e.g., the Prokhorov distance $d(\cdot,\cdot)$. Let μ_n and μ be the probability measure corresponding to W_n

and U, respectively. Then $W_n \xrightarrow{d} U \iff \mu_n \to \mu \iff d(\mu_n, \mu) \to 0$ as $n \to \infty$. Now, we suppose μ_n does not converge to μ, i.e., for some $\epsilon > 0$, there exists a subsequence $\{n(m)\}_{m=1}^{\infty}$ such that $d\left(\mu_{n(m)}, \mu\right) > \epsilon$ for all $n(m)$. From (C.9), $0 < \frac{\sigma_{n,s}}{\sigma_n} \leq C < \infty$, i.e., $\left\{\frac{\sigma_{n,s}}{\sigma_n}\right\}_{n=1}^{\infty}$ is a uniformly bounded sequence over $s \in \mathbb{N}$. Especially, for $s = 1$, $\left\{\sigma_{n(m),1}/\sigma_{n(m)}\right\}_{m=1}^{\infty}$ is a bounded sequence. By the Bolzano-Weierstrass Theorem, it has a convergent subsequence $\left\{\sigma_{n(m(k_1)),1}/\sigma_{n(m(k_1))}\right\}_{k_1=1}^{\infty}$ such that $\sigma_{n(m(k_1)),1}/\sigma_{n(m(k_1))} \to p(1)$ as $k_1 \to \infty$. For $s = 2$, consider $\left\{\sigma_{n(m(k_1)),2}/\sigma_{n(m(k_1))}\right\}$. By the same argument, there exists a further subsequence $\{n(m(k_1(k_2)))\}$ such that $\sigma_{n(m(k_1(k_2))),2}/\sigma_{n(m(k_1(k_2)))} \to p(2)$. Repeating this argument, we can construct a subsequence $\{n(m(k_1(k_2(\cdots(k_s)))))\}$ for all $s \geq 1$ and $\sigma_{n(m(k_1(k_2(\cdots(k_s))))),s}/\sigma_{n(m(k_1(k_2(\cdots(k_s)))))} \to p(s)$ as $k_s \to \infty$. Now construct a subsequence $\{n_l\}$: n_1 is the first element of $\{n(m(k_1))\}$, n_2 is the second element of $\{n(m(k_1(k_2)))\}$, and so on. Then for all $s \geq 1$,

$$\lim_{l\to\infty} \frac{\sigma_{n_l,s}}{\sigma_{n_l}} = p(s).$$

It follows from (C.10), $U_{ns} = \frac{\sigma_{n,s}}{\sigma_n}\left[\sigma_{n,s}^{-1}\sum_{i\in T_n}\xi_{i,n}^s\right]$, and Slutsky's theorem that $U_{n_l s} \xrightarrow{d} U_s \sim N\left(0, p^2(s)\right)$ as $l \to \infty$. Since $|p(s) - 1| \leq \left|p(s) - \frac{\sigma_{n_l,s}}{\sigma_{n_l}}\right| + \left|\frac{\sigma_{n_l,s}}{\sigma_{n_l}} - 1\right|$,

$$\lim_{s\to\infty} |p(s) - 1| \leq \lim_{s\to\infty} \limsup_{l\to\infty} \left|p(s) - \frac{\sigma_{n_l,s}}{\sigma_{n_l}}\right| + \lim_{s\to\infty} \limsup_{l\to\infty} \left|\frac{\sigma_{n,s}}{\sigma_n} - 1\right| = 0,$$

where the last limit follows from (C.8). Therefore, $U_s \xrightarrow{d} U$. And by (C.11),

$$\lim_{s\to\infty} \limsup_{l\to\infty} \mathbb{P}\left(|W_{n_l} - U_{n_l s}| > \epsilon\right) \leq \lim_{s\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(|W_n - U_{ns}| > \epsilon\right) = 0.$$

Then by Lemma A.2, $W_{n_l} \xrightarrow{d} U \sim N(0,1)$ as $l \to \infty$. So, $d(W_{n_l}, U) \to 0$. Since $\{n_l\} \subset \{n(m)\}$, $d(W_{n_l}, U) \to 0$ contradicts the assumption that $d\left(\mu_{n(m)}, \mu\right) > \epsilon$ for all $n(m)$. Hence, $\sigma_n^{-1}(S_n - \mathbb{E}S_n) = W_n \xrightarrow{d} U$. ■

D. Conditional Spatial Functional Dependence

We generalize the concept of spatial FDM to the conditional spatial FDM. The only difference to the original spatial functional dependence is that now the underlying random field becomes conditionally independent (see, e.g., Chow and Teicher, 2003) and all expectations are taken conditionally. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the underlying probability space and $\mathcal{C}$ be a sub-σ-field of $\mathcal{F}$.

Let Y_n be defined as in (2.1), where $\epsilon_{i,n}$'s are conditionally independent given $\mathcal{C}$. So, $\epsilon_{i,n}$'s might be dependent on each other unconditionally. Suppose that conditional on $\mathcal{C}$, $\epsilon_{i,n}^*$ is an i.i.d. copy of $\epsilon_{i,n}$. For a set $I \subset D_n$, define $\epsilon_{i,n,I} \equiv \epsilon_{i,n}^*$ if $i \in I$ and $\epsilon_{i,n,I} \equiv \epsilon_{i,n}$ if $i \notin I$; we denote $\epsilon_{n,I} = \left(\left(\epsilon_{i,n,I}'\right)_{i \in D_n}\right)'$. Then $Y_{i,n,I} = g_{i,n}\left(\epsilon_{n,I}\right)$ is a coupled version of $Y_{i,n}$ on I and $Y_{n,I} = \left(Y_{1,n,I}, \ldots, Y_{n,n,I}\right)'$.

Definition D.1 (Conditional spatial functional dependence)**.** *Let Y_n and ϵ_n be defined as above. For $p \geq 1, n \geq 1$ and $I \subset D_n$, define $\delta_p^{\mathcal{C}}(i, I, n) \equiv \left\| Y_{i,n} - Y_{i,n,I} \right\|_{L^p, \mathcal{C}}$. And we say that $Y = \{Y_{i,n}, i \in D_n, n \geq 1\}$ is $\mathcal{C}$-conditionally L^p-functionally dependent or $\mathcal{C}$-conditionally p-stable on $\epsilon = \{\epsilon_{i,n}, i \in D_n, n \geq 1\}$ if the $\mathcal{C}$-conditional L^p-functional dependence coefficient*

$$\Delta_p^{\mathcal{C}}(s) \equiv \sup_{n \geq 1} \sup_{i \in D_n} \delta_p^{\mathcal{C}}\left(i, \{j : d_{ij} \geq s\}, n\right) \rightarrow 0 \textit{ almost surely (a.s.)} \quad \textit{as } s \rightarrow \infty. \tag{D.1}$$

The conditional spatial functional dependence inherits the properties of the unconditional version. This is because the theorems used in the proofs of the unconditional theorems can be generalized to the corresponding conditional versions (see Prakasa Rao, 2009; Roussas, 2008; Yuan, Wei and Lei, 2014 and the supplementary document of Forchini, Jiang and Peng, 2018). Now, we state our LLN and CLT under conditional spatial functional dependence. In the following, suppose that $\{Y_{i,n}, i \in D_n, n \geq 1\}$ is generated by $\epsilon = \{\epsilon_{i,n}, i \in D_n, n \geq 1\}$, and $\epsilon_{i,n}$'s are conditionally independent given $\mathcal{C}$. T_n is a finite subset of D_n satisfying $|T_n| \rightarrow \infty$ as $n \rightarrow \infty$, and we write $S_n \equiv \sum_{i \in T_n} Y_{i,n}$ and $\sigma_n^2 \equiv \mathrm{Var}_{\mathcal{C}}\left(S_n\right)$.

Theorem D.1 (Law of large numbers)**.** *Under Assumption 1, suppose that* $\sup_{n\geq 1}\sup_{i\in D_n}\|Y_{i,n}\|_{L^p,\mathcal{C}} < \infty$ *a.s. for some* $p > 1$ *and* $\{Y_{i,n}\}$ *is* $\mathcal{C}$*-conditionally* L^1*-FD on* $\{\epsilon_{i,n}\}$*, i.e.,* $\lim_{s\to\infty}\Delta_1(s) = 0$ *a.s. as* $s \to \infty$*. Then*

$$|T_n|^{-1}\left(S_n - \mathbb{E}_{\mathcal{C}}S_n\right) \xrightarrow{p} 0.$$

Proof. From Theorem 3.4, $|T_n|^{-1}\|S_n - \mathbb{E}_{\mathcal{C}}S_n\|_{L^1,\mathcal{C}} \xrightarrow{a.s.} 0$ as $n \to \infty$. Thus, by the Markov inequality, for any $\epsilon > 0$,

$$\mathbb{E}_{\mathcal{C}}\left\{1\left[|T_n|^{-1}\|S_n - \mathbb{E}_{\mathcal{C}}S_n\| > \epsilon\right]\right\} \xrightarrow{a.s.} 0,$$

as $n \to \infty$. Since the indicator function $1(\cdot)$ is always bounded by 1, by the law of iterated expectation and the bounded convergence theorem,

$$\mathbb{E}\left\{1\left[|T_n|^{-1}\|S_n - \mathbb{E}_{\mathcal{C}}S_n\| > \epsilon\right]\right\} = \mathbb{E}\mathbb{E}_{\mathcal{C}}\left\{1\left[|T_n|^{-1}\|S_n - \mathbb{E}_{\mathcal{C}}S_n\| > \epsilon\right]\right\} \to 0,$$

i.e., $|T_n|^{-1}\left(S_n - \mathbb{E}_{\mathcal{C}}S_n\right) \xrightarrow{p} 0$. ∎

Theorem D.2 (Central limit theorem)**.** *Under Assumption 1, suppose the following conditions hold: (1)* $\sup_{n\geq 1}\sup_{i\in D_n}\|Y_{i,n}\|_{L^p,\mathcal{C}} < \infty$ *a.s. for some* $p > 2$*; (2)* $\liminf_{n\to\infty}|T_n|^{-1}\sigma_n^2 > 0$ *a.s.; (3) the* $\mathcal{C}$*-conditional* L^2*-FD coefficient of* $\{Y_{i,n}\}$ *on* $\{\epsilon_{i,n}\}$ *satisfies* $\Delta_2(s) = O\left(s^{-\kappa}\right)$ *a.s. as* $s \to \infty$ *for some* $\kappa > \frac{d}{2}$*. Then*

$$\frac{S_n - \mathbb{E}_{\mathcal{C}}S_n}{\sigma_n} \xrightarrow{d} N(0,1).$$

Proof. From Theorem 3.5, for all $x \in \mathbb{R}$, $\mathbb{P}_{\mathcal{C}}\left(\frac{S_n - \mathbb{E}_{\mathcal{C}}S_n}{\sigma_n} \leq x\right) \xrightarrow{a.s.} \Phi(x)$ as $n \to \infty$, where $\Phi(\cdot)$ is the cumulative distribution function of $N(0,1)$. Since $\mathbb{P}_{\mathcal{C}}(\cdot)$ is always bounded by 1, by the law of iterated expectation and the bounded convergence theorem,

$$\mathbb{P}\left(\frac{S_n - \mathbb{E}_{\mathcal{C}}S_n}{\sigma_n} \leq x\right) = \mathbb{E}\mathbb{P}_{\mathcal{C}}\left(\frac{S_n - \mathbb{E}_{\mathcal{C}}S_n}{\sigma_n} \leq x\right) \to \Phi(x),$$

as $n \to \infty$ for all $x \in \mathbb{R}$, i.e., $\frac{S_n - \mathbb{E}_{\mathcal{C}}S_n}{\sigma_n} \xrightarrow{d} N(0,1)$. ∎

Supplementary Material

Wu, Z., Jiang, W., & Xu, X. (2024). Supplement to "Applications of Functional Dependence to Spatial Econometrics", Econometric Theory Supplementary Material. To view, please visit: https://doi.org/10.1017/S026646662400015X.

References

Andrews, D. W. (1984): "Non-strong mixing autoregressive processes," *Journal of Applied Probability*, 21, 930–934.

Bolthausen, E. (1982): "On the Central Limit Theorem for Stationary Mixing Random Fields," *The Annals of Probability*, 10, 1047–1050.

Brockwell, P. J., and R. A. Davis (1991): *Time Series: Theory and Methods*, New York: Springer-Verlag.

Chen, X., M. Xu, and W. B. Wu (2013): "Covariance and Precision Matrix Estimation for High-Dimensional Time Series," *The Annals of Statistics*, 41, 2994–3021.

Chow, Y. S., and H. Teicher (2003): *Probability Theory: Independence, Interchangeability, Martingales*, New York: Springer-Verlag, 3rd edition.

Conley, T. G., and G. Topa (2002): "Socio-economic distance and spatial patterns in unemployment," *Journal of Applied Econometrics*, 17, 303–327.

Davidson, J. (1994): *Stochastic Limit Theory: An Introduction for Econometricians*, Oxford: Oxford University Press.

Dedecker, J. (1998): "A Central Limit Theorem for Stationary Random Fields," *Probability Theory and Related Fields*, 110, 397–426.

Deng, Y. (2018): "Estimation for the spatial autoregressive threshold model," *Economics Letters*, 171, 172–175.

Doğan, O., and S. Taşp1nar (2013): "GMM estimation of spatial autoregressive models with moving average disturbances," *Regional Science and Urban Economics*, 43, 903–926.

Doukhan, P. (1994): *Mixing: Properties and Examples*, New York: Springer.

Doukhan, P., and S. Louhichi (1999): "A New Weak Dependence Condition and Applications to Moment Inequalities," *Stochastic Processes and their Applications*, 84, 313–342.

El Machkouri, M., D. Volný, and W. B. Wu (2013): "A Central Limit Theorem for Stationary Random Fields," *Stochastic Processes and their Applications*, 123, 1–14.

Fingleton, B. (2008): "A generalized method of moments estimator for a spatial model with moving average errors, with application to real estate prices," *Empirical Economics*, 34, 35–57.

Forchini, G., B. Jiang, and B. Peng (2018): "TSLS and LIML estimators in panels with unobserved shocks," *Econometrics*, 6, 19.

Huang, J. S. (1984): "The autoregressive moving average model for spatial analysis," *Australian Journal of Statistics*, 26, 169–178.

Jenish, N., and I. R. Prucha (2009): "Central Limit Theorems and Uniform Laws of Large Numbers for Arrays of Random Fields," *Journal of Econometrics*, 150, 86–98.

——— (2012): "On Spatial Processes and Asymptotic Inference under Near-Epoch Dependence," *Journal of Econometrics*, 170, 178–190.

Kelejian, H. H., and I. R. Prucha (1998): "A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances," *The Journal of Real Estate Finance and Economics*, 17, 99–121.

——— (2001): "On the Asymptotic Distribution of the Moran I Test Statistic with Applications," *Journal of Econometrics*, 104, 219–257.

Kojevnikov, D., V. Marmer, and K. Song (2021): "Limit Theorems for Network Dependent Random Variables," *Journal of Econometrics*, 222, 882–908.

Lee, L.-f. (2004): "Asymptotic Distributions of Quasi-Maximum Likelihood Estimators for Spatial Autoregressive Models," *Econometrica*, 72, 1899–1925.

——— (2007): "GMM and 2SLS Estimation of Mixed Regressive, Spatial Autoregressive Models," *Journal of Econometrics*, 137, 489–514.

Lee, L.-f., and J. Yu (2010): "Estimation of Spatial Autoregressive Panel Data Models with Fixed Effects," *Journal of Econometrics*, 154, 165–185.

Li, K. (2022): "Threshold spatial autoregressive model," *MPRA working paper No.113568.*

Liu, T., X. Xu, L.-f. Lee, and Y. Mei (2022): "Robust Estimation for the Spatial Autoregressive Model," *Working Paper,*, Xiamen University.

Liu, W., H. Xiao, and W. B. Wu (2013): "Probability and Moment Inequalities under Dependence," *Statistica Sinica*, 23, 1257–1272.

Malikov, E., and Y. Sun (2017): "Semiparametric estimation and testing of smooth coefficient spatial autoregressive models," *Journal of Econometrics*, 199, 12–34.

Nagaev, S. V. (1979): "Large Deviations of Sums of Independent Random Variables," *The Annals of Probability*, 7, 745–789.

Prakasa Rao, B. L. S. (2009): "Conditional Independence, Conditional Mixing and Conditional Association," *Annals of the Institute of Statistical Mathematics*, 61, 441–460.

Qu, X., and L.-f. Lee (2015): "Estimating a Spatial Autoregressive Model with an Endogenous Spatial Weight Matrix," *Journal of Econometrics*, 184, 209–232.

Qu, X., L.-f. Lee, and J. Yu (2017): "QML Estimation of Spatial Dynamic Panel Data Models with Endogenous Time Varying Spatial Weights Matrices," *Journal of Econometrics*, 197, 173–201.

Roussas, G. G. (2008): "On Conditional Independence, Mixing, and Association," *Stochastic Analysis and Applications*, 26, 1274–1309.

Su, L., W. Wang, and X. Xu (2023): "Identifying Latent Group Structures in Spatial Dynamic Panels," *Journal of Econometrics*, 235, 1955–1980.

Sun, Y. (2016): "Functional-coefficient spatial autoregressive models with nonparametric spatial weights," *Journal of Econometrics*, 195, 134–153.

Wainwright, M. J. (2019): *High-Dimensional Statistics: A Non-Asymptotic Viewpoint*, Cambridge: Cambridge University Press.

White, H., and J. M. Wooldridge (1991): "Some Results for Sieve Estimation with Dependent Observations," in *Non-Parametric and Semi-Parametric Methods in Econometrics and Statistics* ed. by Barnett, W. A., Powell, J., and Tauchen, G. E. Cambridge: Cambridge University Press, 459–493.

Wu, W. B. (2005): "Nonlinear System Theory: Another Look at Dependence," *Proceedings of the National Academy of Sciences*, 102, 14150–14154.

——— (2011): "Asymptotic Theory for Stationary Processes," *Statistics and its Interface*, 4, 207–226.

Wu, W. B., and Y. N. Wu (2016): "Performance Bounds for Parameter Estimates of High-

Dimensional Linear Models with Correlated Errors," *Electronic Journal of Statistics*, 10, 352–379.

Xu, X., and L.-f. Lee (2015a): "Maximum Likelihood Estimation of a Spatial Autoregressive Tobit Model," *Journal of Econometrics*, 188, 264–280.

——— (2015b): "A Spatial Autoregressive Model with a Nonlinear Transformation of the Dependent Variable," *Journal of Econometrics*, 186, 1–18.

——— (2018): "Sieve Maximum Likelihood Estimation of the Spatial Autoregressive Tobit Model," *Journal of Econometrics*, 203, 96–112.

——— (2024): "Conditions for α-mixing Linear Random Fields: With Applications to Spatial Econometric Models," *China Journal of Econmetrics,*, 4, 26–57.

Xu, X., W. Wang, Y. Shin, and C. Zheng (2022): "Dynamic Network Quantile Regression Model," *Journal of Business & Economic Statistics*, 0, 1–15.

Yu, J., R. de Jong, and L.-f. Lee (2008): "Quasi-Maximum Likelihood Estimators for Spatial Dynamic Panel Data with Fixed Effects When Both n and T Are Large," *Journal of Econometrics*, 146, 118–134.

Yuan, D.-M., L.-R. Wei, and L. Lei (2014): "Conditional Central Limit Theorems for a Sequence of Conditional Independent Random Variables," *Journal of the Korean Mathematical Society*, 51, 1–15.

Yuan, Z., and M. Spindler (2022): "Bernstein-Type Inequalities and Nonparametric Estimation under near-Epoch Dependence," *Working Paper,*, University of Hamburg.

Zhou, Z., and W. B. Wu (2009): "Local Linear Quantile Estimation for Nonstationary Time Series," *The Annals of Statistics*, 37, 2696–2729.