

# Applications of Functional Dependence to Spatial and Network Econometrics

Zeqi Wu

Institute of Statistics and Big Data, Renmin University of China

Talk@Clubear

November 28<sup>th</sup>, 2025

# References

- Wu, Zeqi, Wen Jiang, and Xingbai Xu (2024): “Applications of Functional Dependence to Spatial Econometrics,” *Econometric Theory*, published online, 1–36.
  - the individuals are located in  $\mathbb{R}^d$ ;
- Jiang, Wen, Yachen Wang, Zeqi Wu, and Xingbai Xu (2025): “Limit Theorems for Network Data without Metric Structure,” *arXiv:2511.17928*.
  - Individuals **need not** locate in a metric space.

- 1 Motivation
- 2 Spatial Functional Dependence
- 3 Properties of Functional Dependence
- 4 Conditional Functional Dependence
- 5 Spatial Panel Data Models
- 6 Functional Dependence under Nonlinear Transformations
- 7 Comparison of Functional Dependence and NED
- 8 Network Function Dependence
- 9 Conclusion

# Motivation

- ▶ Law of large numbers (LLN) and central limit theorems (CLT) are indispensable for econometrics and statistics
- ▶ Spatial econometrics: LLN/CLT for spatially correlated data
- ▶ there are some LLN/CLT in the literature: linear-quadratic form, mixing, near-epoch dependence (NED)
- ▶ they are **not convenient enough** or **some strong conditions are needed** for some applications
- ▶ We aim to develop a weak spatial dependence concept that is **more convenient to use** than above concepts, especially NED

# Tools for Spatial Econometrics: Linear-Quadratic Forms

- ▶ linear-quadratic forms: Kelejian and Prucha (1998, 2001); Lee (2004, 2007); Yu et al. (2008); ....

- ▶  $\epsilon_{i,n}$ 's are independent,  $\epsilon_n = (\epsilon_{1,n}, \dots, \epsilon_{n,n})'$ :  $\epsilon_n' A \epsilon_n + b' \epsilon_n$

- ▶ useful for linear models (Spatial autoregressive (SAR) model)

$$Y_n = \lambda W_n Y_n + X_n \beta + \epsilon_n \Rightarrow Y_n = (I_n - \lambda W_n)^{-1} (X_n \beta + \epsilon_n)$$

- ▶ appears in the log-likelihood function, GMM, Moran's I test statistics. e.g.,  $(Y_n - \lambda W_n Y_n - X_n \beta)' P_n (Y_n - \lambda W_n Y_n - X_n \beta)$

- ▶ inconvenient for

- many nonlinear estimators (quantile estimator, Huber estimator)
- nonlinear spatial models (Tobit model)

# Tools for Spatial Econometrics: Mixing

- mixing is widely used in time series and panel data
- however, spatial mixing (Jenish and Prucha, 2009) is not widely used in spatial econometrics, due to:
  - **hard to establish**, as it involves supremum over two  $\sigma$ -fields:

$$\alpha(\mathcal{A}, \mathcal{B}) \equiv \sup_{A \in \mathcal{A}, B \in \mathcal{B}} |P(A \cap B) - P(A)P(B)|$$

- not preserved under infinity summations (e.g.,  $\sum_{j=1}^n w_{ij} y_j$ )
- Xu and Lee (2023): the mixing property of linear spatial processes.
- For nonlinear processes, as far as we know, no work so far.

# Tools for Spatial Econometrics: Near-Epoch Dependence

**An Example:**  $y_t = \epsilon_t + \rho\epsilon_{t-1} + \rho^2\epsilon_{t-2} + \dots$ :  $y_t$  is mainly affected by  $\epsilon_{t-s}$  for small  $s$ , and the contribution of all  $\epsilon_{t-s}$  (large  $s$ ) is small.

**Definition.**  $\{Z_{i,n}, i \in D_n, n \geq 1\}$  is generated by  $\{v_{i,n}, i \in D_n, n \geq 1\}$ .  $\{Z_{i,n}\}$  is said to be  $L^p$ -near-epoch dependent (NED) on  $\{v_{i,n}\}$  if

$$\|Z_{i,n} - \mathbb{E}[Z_{i,n} | v_{j,n}, d_{ij} \leq s]\|_{L^p} \leq \psi(s) \rightarrow 0 \text{ as } s \rightarrow \infty.$$

# Tools for Spatial Econometrics: Near-Epoch Dependence

- (spatial) near-epoch dependence (NED): Jenish and Prucha (2012), Xu and Lee (2015a, 2018), Qu and Lee (2015), Qu et al. (2017), Liu et al. (2022), Xu et al. (2022, quantile regression).
- Its applications are wide, but
  - sometimes **strong moment conditions are needed** to preserve NED
  - **Limited to  $L^2$ -NED**, as conditional expectation might not be easy to calculate



# What's our work?

- We aim to develop a **weak spatial dependence concept** that is **more convenient to use** than NED
- We generalize the concept of functional dependence (Wu, W.B., 2005, PNAS) to the settings of spatial econometrics:
  - irregular lattice in  $\mathbb{R}^d$ , triangular arrays.
- We establish a set of theoretical tools for functional dependent data:
  - **moment inequality, exponential inequality, LLN, CLT**
  - properties of functional dependence under various transformations.

- 1 Motivation
- 2 Spatial Functional Dependence**
- 3 Properties of Functional Dependence
- 4 Conditional Functional Dependence
- 5 Spatial Panel Data Models
- 6 Functional Dependence under Nonlinear Transformations
- 7 Comparison of Functional Dependence and NED
- 8 Network Function Dependence
- 9 Conclusion

# Functional Dependence Measure

- Suppose there are some individuals (persons, cities, countries, etc, also called spatial units, nodes), and they are located in a lattice  $D_n \subset \mathbb{R}^d$ .
- For a cross-sectional data with  $n$  individuals,  $|D_n| = n$ .
- Let  $\epsilon = \{\epsilon_{i,n}, i \in D_n, n \geq 1\}$  be **independent**  $\mathbb{R}^{p_\epsilon}$ -valued triangular array

$$Y_{i,n} = g_{i,n}(\epsilon_{1,n}, \dots, \epsilon_{n,n}) = g_{i,n}(\epsilon_n), \quad (1)$$

# Functional Dependence Measure

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

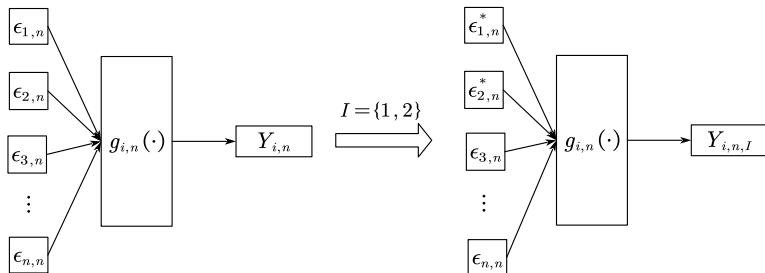
- ▶  $\forall i \in D_n$ , let  $\epsilon_{i,n}^*$  be an i.i.d. copy of  $\epsilon_{i,n}$ , and  $\epsilon_{i,n}^*$  is independent to all  $\epsilon_{j,n}$ ,  $j \in D_n$ .
- ▶  $\forall I \subset D_n$ , define  $\epsilon_{i,n,I} \equiv \epsilon_{i,n}^*$  if  $i \in I$  and  $\epsilon_{i,n,I} \equiv \epsilon_{i,n}$  if  $i \notin I$ ; denote  $\epsilon_{n,I} = (\epsilon_{i,n,I})_{i \in D_n}$ .
- ▶  $Y_{i,n,I} = g_{i,n}(\epsilon_{n,I})$  denotes a coupled version of  $Y_{i,n}$  on  $I$ .

**Definition.** Let  $p \geq 1$  be a constant.  $\forall n \geq 1$  and  $I \subset D_n$ , the  $L^p$  spatial FDM is

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

# Functional Dependence Measure

The  $\delta_p(i, j, n) \equiv \|Y_{i,n} - Y_{i,n,I}\|_{L^p}$  measures the influence of  $\{\epsilon_{j,n} : j \in I\}$  on  $Y_{i,n}$ : if  $\{\epsilon_{j,n} : j \in I\}$  is replaced by its i.i.d. copy  $\{\epsilon_{j,n}^* : j \in I\}$ , how much  $Y_{i,n}$  will change under  $L^p$ -norm.

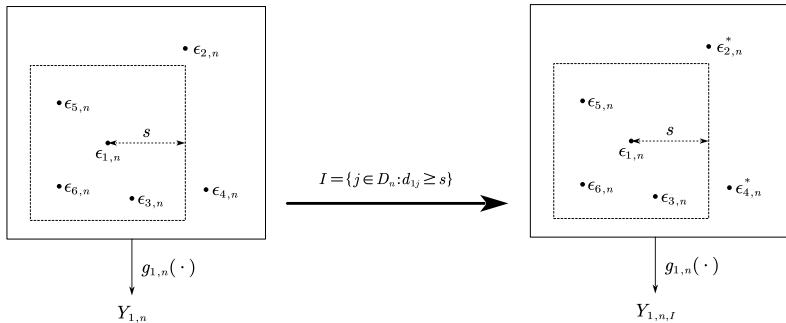


# Functional Dependence Coefficient

**Definition.**  $\{Y_{i,n}, i \in D_n, n \geq 1\}$  is said to be  $L^p$ -functionally dependent on  $\epsilon = \{\epsilon_{i,n}, i \in D_n, n \geq 1\}$  if and only if the  $L^p$ -functionally dependent coefficient

$$\Delta_p(s) \equiv \sup_n \sup_{i \in D_n} \delta_p \left( i, \{j \in D_n : d_{ij} \geq s\}, n \right) \rightarrow 0 \quad \text{as } s \rightarrow \infty. \quad (2)$$

# Functional Dependence Coefficient



# Functional Dependence Coefficient

$$\Delta_p(s) \equiv \sup_n \sup_{i \in D_n} \delta_p \left( i, \{j \in D_n : d_{ij} \geq s\}, n \right) \rightarrow 0 \quad \text{as } s \rightarrow \infty.$$

- ▶  $\lim_{s \rightarrow \infty} \Delta_p(s) = 0$ : the **total impacts from individuals far away can be arbitrarily small** uniformly in  $i$  and  $n$ .
- ▶ By Lyapunov's inequality, if  $Y_{i,n}$  is  $L^p$ -FD, it is also  $L^q$ -FD for all  $q \in [1, p]$ .
- ▶ “monotonicity”:  $\Delta_p(s) \leq 3\Delta_p(\tilde{s})$  for any  $s \geq \tilde{s}$ .



## Example: SAR Model

$$\begin{pmatrix} Y_{1,n} \\ \vdots \\ Y_{n,n} \end{pmatrix} = Y_n = F(\lambda W_n Y_n + \epsilon_n) = \begin{pmatrix} F(\lambda w_{1,n} Y_n + \epsilon_{1,n}) \\ \vdots \\ F(\lambda w_{n,n} Y_n + \epsilon_{n,n}) \end{pmatrix}, \quad (3)$$

- ▶  $W_n = (w_{ij,n})_{n \times n}$  is non-stochastic spatial weights matrix
  - $w_{i,\cdot}$ : the  $i$ th row of  $W_n$
- ▶  $F$  is a Lipschitz function:  $|F(e^\bullet) - F(e)| \leq L|e^\bullet - e|$
- ▶ Assume  $\zeta = L|\lambda| \sup_n \|W_n\|_\infty < 1$ :
  - Eq.(3) has a unique solution  $\Rightarrow Y_{i,n} = Y_{i,n}(\epsilon_n)$

## Example: SAR Model

$$Y_n = F(\lambda W_n Y_n + \epsilon_n).$$

► Let  $M_n \equiv L(I_n - L|\lambda| |W_n|)^{-1}$ , where  $|W_n| \equiv (|w_{ij,n}|)_{n \times n}$ .

► Then

$$|Y_{i,n}(\epsilon_n^{(1)}) - Y_{i,n}(\epsilon_n^{(2)})| \leq \sum_{j=1}^n M_{ij,n} |\epsilon_{j,n}^{(1)} - \epsilon_{j,n}^{(2)}|.$$

## Example: SAR Model

$$|Y_{i,n}(\epsilon_n^{(1)}) - Y_{i,n}(\epsilon_n^{(2)})| \leq \sum_{j=1}^n M_{ij,n} |\epsilon_{j,n}^{(1)} - \epsilon_{j,n}^{(2)}|.$$

► the  $L^p$ -FD coefficient

$$\Delta_p(s) \leq \left( 2 \sup_{j,n} \|\epsilon_{j,n}\|_p \right) \sup_{i,n} \sum_{j \in D_n: d_{ij} \geq s} M_{ij,n}.$$

► Suppose  $\epsilon_{i,n}$ 's are independent over  $i$  and uniformly  $L^p$ -bounded.

1  $w_{ij,n} \neq 0$  only if  $d_{ij} < \bar{d}_0$ :

$$\Delta_p(s) \leq C \zeta^{s/\bar{d}_0} \rightarrow 0 \quad \text{as } s \rightarrow \infty;$$

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

$$\Delta_p(s) \leq O\left(s^{-(\alpha-d)} (\log s)^{\alpha-d}\right) \rightarrow 0 \quad \text{as } s \rightarrow \infty.$$

# SAR Models: Extensions

- ▶ In most of cases:  $\epsilon_{i,n} = X'_{i,n}\beta + e_{i,n}$ . We can allow  $\{\epsilon_{i,n}\}$  to be dependent by
  - assuming that  $\{\epsilon_{i,n}\}$  is functionally dependent on another random field  $\{\eta_{i,n}\}$ ;
  - $\Delta_p(s) \leq O(s^{-(\alpha-d)}(\log s)^{\alpha-d}) + C \times \Delta_p^\epsilon(s/2)$ .
- ▶ We also calculate the FDM for
  - the threshold SAR models
  - functional-coefficient SAR models
  - smooth-coefficient SAR models
  - SAR models with **stochastic** weights matrices
  - ...

- 1 Motivation
- 2 Spatial Functional Dependence
- 3 Properties of Functional Dependence**
- 4 Conditional Functional Dependence
- 5 Spatial Panel Data Models
- 6 Functional Dependence under Nonlinear Transformations
- 7 Comparison of Functional Dependence and NED
- 8 Network Function Dependence
- 9 Conclusion

# A Moment Inequality

Let  $S_n \equiv \sum_{i \in D_n} Y_{i,n}$ .

**Theorem.** For  $p \geq 2$ , if the  $L^p$ -FD coefficient  $\Delta_p(s) = O(s^{-\kappa})$  for some  $\kappa > \frac{d}{2}$  as  $s \rightarrow \infty$ ,

$$\|S_n - \mathbb{E}S_n\|_{L^p} \leq C\sqrt{n}.$$

- ▶ The rate is the same as the i.i.d. case;
- ▶ For SAR models, when  $|w_{ij,n}| \leq cd_{ij}^{-\alpha}$ , we have  $\Delta_p(s) \leq O(s^{-(\alpha-d)} (\log s)^{\alpha-d})$
- ▶ If  $\alpha > 1.5d$ , then  $\Delta_p(s) = O(s^{-\kappa})$

# An Exponential Inequality

- Exponential inequality is useful in semi/nonparametric econometrics, high-dimensional statistics, machine learning.

**Theorem.** Assume  $\{Y_{i,n}\}$  is  $L^p$ -functionally dependent on  $\{\epsilon_{i,n}\}$  with  $\Delta_p(s) \leq O(p^\nu) O(s^{-\kappa})$  for some  $\kappa > \frac{d}{2}$  and  $\nu \geq 0$ . Denote  $\alpha = \frac{2}{1+2\nu}$ . Then  $\forall \delta > 0$ ,

$$\mathbb{P}(|S_n - \mathbb{E}S_n| \geq n\delta) \leq C_1 \exp(-C_2 n^{\alpha/2} \delta^\alpha).$$

- SAR:  $\Delta_p(s) \leq (2 \sup_{j,n} \|\epsilon_{j,n}\|_p) \sup_{i,n} \sum_{j \in D_n: d_{ij} \geq s} M_{ij,n};$

- $O(s^{-\kappa})$ :  $|w_{ij,n}| \leq cd_{ij}^{-\alpha}$  for some  $\alpha > 1.5d$ ;

- $O(p^\nu)$ :

- 1 when  $\epsilon_{i,n}$  is subexponential,  $\nu = 1$ ,  $n^{\alpha/2} = n^{1/3}$
- 2 when  $\epsilon_{i,n}$  is sub-Gaussian,  $\nu = \frac{1}{2}$ ,  $n^{\alpha/2} = n^{1/2}$
- 3 when  $\epsilon_{i,n}$  is uniformly bounded,  $\nu = 0$ ,  $n^{\alpha/2} = n$

# An Exponential Inequality

$$\mathbb{P}(|S_n - \mathbb{E}S_n| \geq n\delta) \leq C_1 \exp(-C_2 n^{\alpha/2} \delta^\alpha).$$

- The proof is based on the moment inequality and Taylor formula for  $e^x$
- Our result:  $n^{\alpha/2} = n^{1/(1+2\nu)}$ ;
  - rate in Xu and Lee (2018):  $n^{1/(2d+2+\nu)}$
- spatial FD: allow  $w_{ij} \lesssim d_{ij}^{-\alpha}$ ;
  - rate in Xu and Lee (2018) requires:  $w_{ij} \lesssim \exp(-d_{ij}^\alpha)$
- When  $\epsilon_{i,n}$  is uniformly bounded,  $\nu = 0$ .
  - the decaying rate with respect to  $n$  is the same as the standard Hoeffding's inequality



# Law of Large Numbers

**Theorem.** If (i)  $\{Y_{i,n}\}$  is uniformly  $L^p$ -bounded for some  $p > 1$ , and (ii)  $\{Y_{i,n}\}$  is  $L^1$ -FD on  $\{\epsilon_{i,n}\}$ , i.e.,  $\lim_{s \rightarrow \infty} \Delta_1(s) = 0$ , then

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

# Central Limit Theorem

**Theorem.** Denote  $\sigma_n^2 = \text{Var}(S_n)$ , if

(i)  $Y_{i,n}$  is uniformly  $L^p$ -bounded for some  $p > 2$ ,

(ii)  $\liminf_{n \rightarrow \infty} \sigma_n^2/n > 0$ ,

(iii)  $\Delta_2(s) = O(s^{-\kappa})$  for some  $\kappa > \frac{d}{2}$  as  $s \rightarrow \infty$ ,

then

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

► For SAR models with  $w_{ij} \lesssim d_{ij}^{-\alpha}$ :

■ FD CLT requires:  $\alpha > 1.5d$ ;

■ the NED CLT requires:  $\alpha > 2d$ ;

► By the Cramér–Wold device, we can generalize the CLT to the multivariate case.

- 1 Motivation
- 2 Spatial Functional Dependence
- 3 Properties of Functional Dependence
- 4 Conditional Functional Dependence**
- 5 Spatial Panel Data Models
- 6 Functional Dependence under Nonlinear Transformations
- 7 Comparison of Functional Dependence and NED
- 8 Network Function Dependence
- 9 Conclusion

# Conditional Spatial Functional Dependence

- $(\Omega, \mathcal{F}, \mathbb{P})$ : the underlying probability space;  $\mathcal{C}$  is a sub- $\sigma$ -field of  $\mathcal{F}$ .
- $\epsilon_{i,n}$ 's conditionally independent given  $\mathcal{C}$ .
- 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  $i \notin I$ .
- $Y_{i,n} = g_{i,n}(\epsilon_n)$ ;  $Y_{i,n,I} = g_{i,n}(\epsilon_{n,I})$  is a coupled version of  $Y_{i,n}$  on  $I$ .

**Definition.** For  $p \geq 1, n \geq 1$  and  $I \subset D_n$ , define the conditional FDM

$$\delta_p^{\mathcal{C}}(i, I, n) \equiv \|Y_{i,n} - Y_{i,n,I}\|_{L^p, \mathcal{C}} \equiv (\mathbb{E}|Y_{i,n} - Y_{i,n,I}|^p | \mathcal{C})^{1/p}.$$

And  $\{Y_{i,n}, i \in D_n\}$  is  $\mathcal{C}$ - $L^p$ -functionally dependent on  $\epsilon = \{\epsilon_{i,n}, i \in D_n\}$  if and only if

$$\Delta_p^{\mathcal{C}}(s) \equiv \sup_{n \geq 1} \sup_{i \in D_n} \delta_p^{\mathcal{C}}(i, \{j : d_{ij} \geq s\}, n) \xrightarrow{a.s.} 0 \quad \text{as } s \rightarrow \infty.$$

- 1 Motivation
- 2 Spatial Functional Dependence
- 3 Properties of Functional Dependence
- 4 Conditional Functional Dependence
- 5 Spatial Panel Data Models**
- 6 Functional Dependence under Nonlinear Transformations
- 7 Comparison of Functional Dependence and NED
- 8 Network Function Dependence
- 9 Conclusion

# Spatial Panel Data Model

- ▶  $T$  periods,  $N$  spatial units in  $D_N \subset \mathbb{R}^d$ .
- ▶ for individual  $i$  at time  $t$ , we regard  $(i, t)$  as a point in  $\mathbb{R}^{d+1}$ :  
 $(i, t) \in D_{NT} \equiv \{(i, t) \in \mathbb{R}^{d+1} : i \in D_N, t = T, T-1, \dots\}$

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

- ▶ Spatial dynamic panel data (SDPD) model:

$$Y_{Nt} = \lambda W_N Y_{Nt} + \gamma Y_{N,t-1} + \rho W_N Y_{N,t-1} + X_{Nt} \beta + \mu_t l_N + \nu_N + V_{Nt}$$

- ▶ We calculate the spatial FDM conditional on  $(\mu_t, \nu_N)$ :  
 $\mathcal{C} \equiv \bigvee_{t=-\infty}^{\infty} \bigvee_{N=1}^{\infty} \sigma(\mu_t, \nu_N)$

# Spatial Panel Data Model

Assumptions:

- 1  $|w_{ij,N}| \leq cd_{ij}^{-\alpha}$  for some constants  $c > 0$  and  $\alpha > d$ .
- 2  $\sup_N \|W_N\|_\infty \leq 1$  and  $|\lambda| + |\gamma| + |\rho| < 1$ . And  $\zeta \equiv \frac{|\gamma|+|\rho|}{1-|\lambda|} < 1$ .
- 3  $\sup_{N,T} \sup_{i,t} \|\epsilon_{it}\|_{L^p,C} < \infty$  a.s. for some  $p \geq 1$ .
- 4 Conditional on  $C$ ,  $(x'_{it}, v_{it})$ 's are independent over  $i$  and  $t$ .

**Proposition.**  $\{y_{it} : (i, t) \in D_{NT}\}$  is  $C$ - $L^p$ -FD on  $\{\epsilon_{it}\}$  with  $\Delta_p^C(s) = O\left(s^{-(\alpha-d)} (\log s)^{\alpha-d}\right)$  a.s. as  $s \rightarrow \infty$ .

- 1 Motivation
- 2 Spatial Functional Dependence
- 3 Properties of Functional Dependence
- 4 Conditional Functional Dependence
- 5 Spatial Panel Data Models
- 6 Functional Dependence under Nonlinear Transformations**
- 7 Comparison of Functional Dependence and NED
- 8 Network Function Dependence
- 9 Conclusion



# Nonlinear Transformations in Spatial Econometrics

In spatial econometrics, we need to deal with a lot of nonlinear transformations about spatial random variables:

- In the log-likelihood function based on normal distribution:  $y^2$
- censored data, binary data, quantile regression:  $1(y > 0)$
- Least absolute deviation:  $|y|$
- Spatial Tobit model:  $\Phi(y)$ ,  $\ln \Phi(y)$ ,  $\frac{\phi(y)}{\Phi(y)}$ , ...
- Huber estimator:  $\max(0, x)$ , or  $\min(1, \max(0, x))$
- variance estimator:  $y^2$  or  $yz$

# FDM under Nonlinear transformations

$H_{i,n}$  is a function:

$$\|H_{i,n}(y) - H_{i,n}(y^\bullet)\| \leq B_{i,n}(y, y^\bullet) \|y - y^\bullet\|$$

Denote  $Z_{i,n} \equiv H_{i,n}(Y_{i,n})$ .

**Proposition.** If  $B_{i,n}(y, y^\bullet) \leq C < \infty$ , then  $\delta_{Z,p}(i, j, n) \leq C\delta_{Y,p}(i, j, n)$  and  $\Delta_{Z,p}(s) \leq C\Delta_{Y,p}(s)$ .

► **Examples.**  $H(x) = \phi(x)$ ,  $H(x) = \max(x, 0)$ ,  $H(x) = |x|$ ,  
 $H(x) = \min(1, \max(0, x))$

# FDM under Nonlinear transformations

Denote  $Z_{i,n} \equiv H_{i,n}(Y_{i,n})$

$$\|H_{i,n}(y) - H_{i,n}(y^\bullet)\| \leq B_{i,n}(y, y^\bullet) \|y - y^\bullet\|$$

**Proposition.** Suppose  $B_{i,n}(y, y^\bullet) \leq C(\|y\|^a + \|y^\bullet\|^a + 1)$  for some  $a \geq 1$ . The constants  $p, q, r \geq 1$  satisfy  $p^{-1} = q^{-1} + r^{-1}$ . If  $\{Y_{i,n}\}$  is  $L^q$ -FD on  $\{\epsilon_{i,n}\}$  and  $\sup_{n,i \in D_n} \|Y_{i,n}\|_{L^{ar}} < \infty$ , then  $\Delta_p^Z(s) \leq C_1 \Delta_q^Y(s)$ .

- Since  $q > p$ , this prop allows us to establish FD of lower order using FD of higher order. The order of  $\Delta_p^Z(s)$  and  $\Delta_q^Y(s)$  are the same.
- **Examples.**  $H(x) = x^2$ ,  
 $|\ln \Phi(x_1) - \ln \Phi(x_2)| \leq C(|x_1| + |x_2| + 1)|x_1 - x_2|$

# FDM under Nonlinear transformations

Denote  $Z_{i,n} \equiv H_{i,n}(Y_{i,n})$

$$\|H_{i,n}(y) - H_{i,n}(y^\bullet)\| \leq B_{i,n}(y, y^\bullet) \|y - y^\bullet\|$$

**Proposition.** Suppose  $B_{i,n}(y, y^\bullet) \leq C(\|y\|^a + \|y^\bullet\|^a + 1)$  for some  $a \geq 1$  and  $\sup_{n,i \in D_n} \|Y_{i,n}\|_{L^q} < \infty$  for some  $q$  satisfying  $q > \max\{(a+1)p, \frac{ap}{p-1}\}$ , where  $p > 1$ . Then

$$\Delta_p^Z(s) \leq C_2 \Delta_p^Y(s)^{(q-ap-p)/(pq-ap-p)}.$$

- This proposition allows us to calculate  $L^p$  FDM of  $Z$  using  $L^p$  FDM of  $Y$ , but the **decreasing rate is slower**, as  $\frac{q-ap-p}{pq-ap-p} < 1$
- NED has a similar property

**Proposition.** Denote  $Z_{i,n} \equiv 1 (Y_{i,n} > 0)$ . Suppose the probability density functions of  $\{Y_{i,n}\}$  are uniformly bounded in  $i$  and  $n$ . Then there exists a constant  $C > 0$  such that

$$\Delta_p^Z(s) \leq C \Delta_p^Y(s)^{1/(p+1)}.$$

- 1 Motivation
- 2 Spatial Functional Dependence
- 3 Properties of Functional Dependence
- 4 Conditional Functional Dependence
- 5 Spatial Panel Data Models
- 6 Functional Dependence under Nonlinear Transformations
- 7 Comparison of Functional Dependence and NED**
- 8 Network Function Dependence
- 9 Conclusion

# Relationship between Spatial FD and NED

**Definition.** For some  $p \geq 1$ ,  $Z = \{Z_{i,n}, i \in D_n, n \geq 1\}$  said to be  $L^p$ -near-epoch dependent on  $\epsilon = \{\epsilon_{i,n}, i \in D_n, n \geq 1\}$  if

$$\left\| Z_{i,n} - \mathbb{E}(Z_{i,n} | \epsilon_{j,n}, d_{ij} \leq s) \right\|_{L^p} \leq \psi(s)$$

with  $\lim_{s \rightarrow \infty} \psi(s) = 0$ .

- NED: every spatial unit is mainly affected by the  $\epsilon_{j,n}$  of its close neighbors
- spatial FD: the impacts of the  $\epsilon_{j,n}$ 's faraway are small
- The ideas of these two concepts are similar.
- What's the relationship between FD and NED?

# Relationship between Spatial FD and NED

**Theorem.** (1) When  $\epsilon_{i,n}$ 's independent,  $L^p$ -NED coefficient  $\leq L^p$ -FD coefficient:

$$\psi_p(s) = \sup_{i,n} \left\| Z_{i,n} - \mathbb{E}(Z_{i,n} | \epsilon_{j,n}, d_{ij} \leq s) \right\|_{L^p} \leq \Delta_p(s).$$

(2) If  $Y_{i,n} = \sum_{j \in D_n} w_{ij,n} \epsilon_{j,n}$ , where  $w_{ij,n}$ 's are non-stochastic coefficients and  $\epsilon_{i,n}$ 's are independent, then

$$\Delta_p(s) \leq 2\psi_p(s).$$

- When  $\epsilon_{i,n}$ 's are **NOT** independent, the above conclusion might not hold.



# Relationship between Spatial FD and NED

- Spatial FDM is **more convenient to calculate** than spatial NED, especially under nonlinear transformations and  $p > 2$ .
- It usually requires **weaker conditions** to establish a CLT and an exponential inequality by using FDM.
  - For CLT, it only requires  $\alpha > 1.5d$  under FDM instead of  $\alpha > 2d$  under NED.
  - The exponential inequality under FDM enjoys both less restrictive conditions and faster decaying rate.

- 1 Motivation
- 2 Spatial Functional Dependence
- 3 Properties of Functional Dependence
- 4 Conditional Functional Dependence
- 5 Spatial Panel Data Models
- 6 Functional Dependence under Nonlinear Transformations
- 7 Comparison of Functional Dependence and NED
- 8 Network Function Dependence**
- 9 Conclusion

# Function Dependence in Network Data

- $1, 2, \dots, n$  are  $n$  nodes in a network;
- $i$ 's **might not** be located in a Euclidean space  $\mathbb{R}^d$ ;
  - e.g., financial and social networks
- the order of  $1, 2, \dots, n$  can be arbitrary;
- Let  $\epsilon = \{\epsilon_{i,n} : 1 \leq i \leq n, n \geq 1\}$  be **independent**  $\mathbb{R}^{p_\epsilon}$ -valued triangular array

$$Y_{i,n} = g_{i,n}(\epsilon_{1,n}, \dots, \epsilon_{n,n}) = g_{i,n}(\epsilon_n). \quad (4)$$

# Definition

- ▶ Let  $\epsilon_{j,n}^*$  be an i.i.d. copy of  $\epsilon_{j,n}$ , and  $\epsilon_{j,n}^*$  is independent to  $\epsilon_{i,n}$  for any  $i \neq j$ .
- ▶ Denote  $Y_{i,n,j}$  as the coupled version of  $Y_{i,n}$  with  $\epsilon_{j,n}$  replaced by its i.i.d. copy  $\epsilon_{j,n}^*$ , i.e.,

$$Y_{i,n,j} \equiv g_{i,n} \left( \epsilon_{1,n}, \dots, \epsilon_{j-1,n}, \epsilon_{j,n}^*, \epsilon_{j+1,n}, \dots, \epsilon_{n,n} \right).$$

**Definition.** Let  $p \geq 1$  be a constant. Define the functional dependence measure as

$$\delta_{p,n}(i, j) \equiv \|Y_{i,n} - Y_{i,n,j}\|_{L^p}. \quad (5)$$

It measures the impact of  $\epsilon_{j,n}$  on  $Y_{i,n}$ .

# Definition

► Recall

$$\delta_{p,n}(i, j) \equiv \|Y_{i,n} - Y_{i,n,j}\|_{L^p}.$$

**Definition.** For  $p \geq 1$  and  $q \geq 1$ ,  $\{Y_{j,n}\}$  is said to be  $(L^p, q)$ -functionally dependent on  $\{\epsilon_{i,n}\}$  if

$$\Delta_{p,q} \equiv \frac{1}{n^q} \sum_{j=1}^n \left[ \sum_{i=1}^n \delta_{p,n}(i, j) \right]^q = o(1).$$

► A sufficient condition is

$$\frac{1}{n} \sum_{j=1}^n \left[ \sum_{i=1}^n \delta_{p,n}(i, j) \right]^q = o(1)$$

for some  $q > 1$ .

# Definition

- ▶ For SAR model,  $\delta_{p,n}(i, j) \propto M_{ij,n}$ , where  $M_n \equiv L(I_n - L|\lambda W_n|)^{-1}$
- ▶  $\sum_{i=1}^n \delta_{p,n}(i, j) \propto \sum_{i=1}^n M_{ij,n}$  is the  $j^{th}$  column sum of  $M_n$ :
  - the total impact of  $e_{j,n}$  on all  $Y_{i,n}$ 's
  - can be regarded as the “influence power” of  $j$
- ▶  $\Delta_{p,q} = o(1)$  generalizes the condition  $\sup_n \|M_n\|_1 < \infty$  in many spatial econometric papers.
  - We allow  $\sup_{j,n} \sum_{i=1}^n \delta_{p,n}(i, j) = \infty$ .
- ▶ However,  $\Delta_{p,q} = o(1)$  excludes the case that all  $Y_{i,n}$ 's are mainly affected by the same very few  $e_{j,n}$ 's.
  - Consider:  $Y_{i,n} = e_{1,n} \forall i = 1, \dots, n$ . Then  $\sum_{i=1}^n \delta_{p,n}(i, 1) \propto n$ , and thus  $\Delta_{p,q} \propto 1$ .

# Moment Inequality

**Theorem.** We have

$$\left\| \frac{1}{n} \sum_{i=1}^n (Y_{i,n} - \mathbb{E}Y_{i,n}) \right\|_{L^p} \leq C_p \left\{ \Delta_{p, \min\{p, 2\}} \right\}^{1/\min\{p, 2\}}.$$

► If  $\sum_{i=1}^n \delta_{p,n}(i, j) < \infty$  for some  $p \geq 2$ , then

$$\Delta_{p,2} \equiv \frac{1}{n^2} \sum_{j=1}^n \left[ \sum_{i=1}^n \delta_{p,n}(i, j) \right]^2 = O(n^{-1}).$$

► As a result,  $\left\| \frac{1}{n} \sum_{i=1}^n (Y_{i,n} - \mathbb{E}Y_{i,n}) \right\|_{L^p} = O(\sqrt{n})$ , same rate as the i.i.d. case.

# Exponential Inequality

**Theorem.** Assume  $\gamma_0 \equiv \sup_{p \geq 2} \sup_{n \geq 1} p^{-\nu} \sqrt{n} \Delta_{p,2}^{1/2} < \infty$  for some  $\nu \geq 0$ . Then  $\forall t > 0$ ,

$$\mathbb{P} \left( \left| \frac{1}{n} \sum_{i=1}^n (Y_{i,n} - \mathbb{E} Y_{i,n}) \right| \geq t \right) \leq C_1 \exp \left( -C_2 n^{\frac{1}{1+2\nu}} t^{\frac{2}{1+2\nu}} \right).$$

► For SAR, recall  $\delta_{p,n}(i, j) \leq 2\|\epsilon\|_{L^p} M_{ij,n}$ . If 
$$\frac{1}{n} \sum_{j=1}^n \left[ \sum_{i=1}^n M_{ij,n} \right]^2 < \infty,$$

then

$$\Delta_{p,2}^{1/2} \equiv \sqrt{\frac{1}{n^2} \sum_{j=1}^n \left[ \sum_{i=1}^n \delta_{p,n}(i, j) \right]^2} \leq 2\|\epsilon\|_{L^p} \frac{1}{\sqrt{n}};$$

- 1 when  $\epsilon_{i,n}$  is subexponential:  $\nu = 1$ ,  $n^{\frac{1}{1+2\nu}} = n^{1/3}$
- 2 when  $\epsilon_{i,n}$  is sub-Gaussian:  $\nu = \frac{1}{2}$ ,  $n^{\frac{1}{1+2\nu}} = n^{1/2}$
- 3 when  $\epsilon_{i,n}$  is uniformly bounded:  $\nu = 0$ ,  $n^{\frac{1}{1+2\nu}} = n$



# Central Limit Theorem

**Theorem.** Let  $p > 2$  be a constant and  $\sigma_n^2 \equiv \text{Var}(\sum_{i=1}^n Y_{i,n})$ . Suppose that (i)  $\liminf_{n \rightarrow \infty} \frac{\sigma_n^2}{n} > 0$  (ii)

$$\sup_{n,j} \sum_{i=1}^n \delta_{p,n}(i,j) < \infty, \quad (6)$$

and (iii)

$$\frac{1}{n^{\min\{2,p/2\}}} \sum_{i=1}^n \left\{ \sum_{j=1}^n \sum_{k=1}^n \min\{\delta_{p,n}(k,i), \delta_{p,n}(k,j)\} \right\}^{\min\{2,p/2\}} = o(1) \quad (7)$$

as  $n \rightarrow \infty$ . Then

$$\frac{\sum_{i=1}^n (Y_{i,n} - \mathbb{E}Y_{i,n})}{\sigma_n} \xrightarrow{d} N(0, 1).$$

# Central Limit Theorem

- ▶  $j(i)$ : the **index** of the  $j$ th largest value of  $\{\delta_{p,n}(i, j, C_n) : j = 1, \dots, n\}$ .

- ▶ When  $p \geq 4$ , a sufficient condition for (7) is

$$\delta_{p,n}(i, j(i)) \leq C\{j(i)\}^{-\alpha}, \quad \forall 1 \leq i \leq n, \quad (8)$$

where  $\alpha > 2$  and  $C > 0$  are some constants.

- ▶ Existing weak dependence concepts:  $\theta_s$  decreases as **distance**  $s$  increases;
  - $s \rightarrow \infty$  excludes networks with small diameters;
- ▶ Ours: **no distance**; applies to networks **with small diameters**.

- 1 Motivation
- 2 Spatial Functional Dependence
- 3 Properties of Functional Dependence
- 4 Conditional Functional Dependence
- 5 Spatial Panel Data Models
- 6 Functional Dependence under Nonlinear Transformations
- 7 Comparison of Functional Dependence and NED
- 8 Network Function Dependence
- 9 Conclusion**

# Conclusion

- ▶ We generalize the concept of functional dependence proposed in Wu (2005) to spatial and network econometric settings: **easy to verify, convenient to use**
- ▶ We establish a moment inequality, an exponential inequality, LLN, and CLT
- ▶ We calculate the FDM for some models;
- ▶ spatial FDM: individuals are in  $\mathbb{R}^d$ ;
- ▶ networks FDM: no metric space.

# *Thank You!*

- Email: `wuzeqi@ruc.edu.cn`
- Homepage: <https://zeqiwu1202.github.io>