symmetry

MDPI

Article

PQMLE of a Partially Linear Varying Coefficient Spatial Autoregressive Panel Model with Random Effects

Shuangshuang Li [1], Jianbao Chen [1,*] and Danqing Chen [1,2]

[1] College of Mathematics and Statistics, Fujian Normal University, Fuzhou 350117, China; qbx20190068@yjs.fjnu.edu.cn (S.L.); 19851523@fjut.edu.cn (D.C.)
[2] School of Computer Science and Mathematics, Fujian University of Technology, Fuzhou 350118, China
* Correspondence: jbchen@fjnu.edu.cn

Abstract: This article deals with asymmetrical spatial data which can be modeled by a partially linear varying coefficient spatial autoregressive panel model (PLVCSARPM) with random effects. We constructed its profile quasi-maximum likelihood estimators (PQMLE). The consistency and asymptotic normality of the estimators were proved under some regular conditions. Monte Carlo simulations implied our estimators have good finite sample performance. Finally, a set of asymmetric real data applications was analyzed for illustrating the performance of the provided method.

Keywords: PLVCSARPM; PQMLE; consistency asymptotic normality; Monte Carlo simulation; asymmetric data

check for updates

Citation: Li, S.; Chen, J.; Chen, D. PQMLE of a Partially Linear, Varying Coefficient Spatial Autoregressive Panel Model with Random Effects. *Symmetry* **2021**, *13*, 2057. https://doi.org/10.3390/sym13112057

Academic Editors: Célia Nunes, Miguel Fonseca and Elsa Moreira

Received: 26 September 2021
Accepted: 19 October 2021
Published: 1 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Spatial regression models are used to deal with spatially dependent data which widely exist in many fields, such as economics, environmental science and geography. According to the different types of spatial interaction effects, spatial regression models can be sorted into three basic categories (see [1]). The first category is spatial autoregressive (SAR) models, which include endogenous interaction effects among observations of response variables at different spatial scales (see [2]). The second category is spatial durbin models (SDM), which include exogenous interaction effects among observations of covariates and endogenous interaction effects among observations of response variables at different spatial scales (see [3,4]). The third category is spatial error models (SEM), which include interaction effects among disturbance terms of different spatial scales (see [5]). Among them, the SAR models proposed by [6] may be the most popular. The developments in the testing and estimation of SAR models for sectional data were summarized in the books by [7–9] and the surveys by [10–16], among others. Compared to sectional data models, panel data models exhibit an excellent capacity for capturing the complex situations by using abundant datasets built up over time and adding individual-specific or time-specific fixed or random effects. Their theories, methods and applications can be found in the books by [17–20] and the surveys by [21–26], among others.

The above mentioned research literature mainly focuses on linear parametric models. Although the estimations and properties of these models have been well established, they are often unrealistic in application, for the reason that they are unable to accommodate sufficient flexibility to accommodate complex structures (e.g., nonlinearity). Moreover, mis-specification of the data generation mechanism by a linear parametric model could lead to excessive modeling biases or even erroneous conclusions.

Ref. [27] pointed out that the relationship between variables in space usually exhibits highly complexity in reality. Therefore, researches have proposed a number of solution methods. In video coding systems, transmission problems are usually dealt with wavelet-based methods. Ref. [28] proposed a low-band-shift method which avoids the shift-variant property of the wavelet transform and performs the motion compensation more precisely and efficiently. Ref. [29] obtained a wavelet-based lossless video coding scheme that switches between two operations based on the different amount of motion activities between two consecutive frames. More theories and applications can be found in [30–32]. In econometrics, some nonparametric and semiparametric spatial regression models have been developed to relax the linear parametric model settings. For sectional spatial data, ref. [33] studied the GMM estimation of a nonparametric SAR model; Ref. [34] investigated the PQMLE of a partially linear nonparametric SAR model. However, such nonparametric SAR models may cause the "curse of dimensionality" when the dimension of covariates is higher. In order to overcome this drawback, ref. [35] studied two-step SGMM estimators and their asymptotic properties for spatial models with space-varying coefficients. Ref. [36] proposed a semiparametric series-based least squares estimating procedure for a semiparametric varying coefficient mixed regressive SAR model and derived the asymptotical normality of the estimators. Some other related research works can be found in [37–42]. For panel spatial data, ref. [43] obtained PMLE and its asymptotical normality of varying the coefficient SAR panel model with random effects; Ref. [44] applied instrumental variable estimation to a semiparametric varying coefficient spatial panel data model with random effects and the investigated asymptotical normality of the estimators.

In this paper, we extend the varying coefficient spatial panel model with random effects given in [43] to a PLVCSARPM with random effects. By adding a linear part to the model of [43], we can simultaneously capture linearity, non-linearity and the spatial correlation relationships of exogenous variables in a response variable. By using the profile quasimaximum likelihood method to estimate PLVCSARPM with random effects, we proved the consistency and asymptotic normality of the estimators under some regular conditions. Monte Carlo simulations and real data analysis show that our estimators perform well.

This paper is organized as follows: Section 2 introduces the PLVCSARPM with random effects and constructs its PQMLE. Section 3 proves the asymptotic properties of estimators. Section 4 presents the small sample estimates using Monte Carlo simulations. Section 5 analyzes a set of asymmetric real data applications for illustrating the performance of the proposed method. A summary is given in Section 6. The proofs of some important theorems and lemmas are given in Appendix A.

2. The Model and Estimators

Consider the following partially linear varying coefficient spatial autoregressive panel model (PLVCSARPM) with random effects:

$$y_{it} = \rho(W_0 Y_t)_i + x_{it}^{'}\beta(u_{it}) + z_{it}^{'}\alpha + b_i + \varepsilon_{it}, 1 \le i \le N, 1 \le t \le T, \tag{1}$$

where i refers to a spatial unit; t refers to a given time period; y_{it} are observations of a response variable; $Y_t = (y_{1t}, y_{2t}, \cdots, y_{Nt})^{'}$; $x_{it} = (x_{it1}, x_{it2}, \cdots, x_{itp})^{'}$ and $z_{it} = (z_{it1}, z_{it2}, \cdots, z_{itq})^{'}$ are observations of p-dimensional and q-dimensional covariates, respectively; $\beta(u_{it}) = (\beta_1(u_{it}), \beta_2(u_{it}), \cdots, \beta_p(u_{it}))'$ is an unknown univariate varying coefficient function vector; $\beta_s(u)$ $(s = 1, 2, \cdots, p)$ are unknown smoothing functions of u; ρ $(|\rho| < 1)$ is an unknown spatial correlation coefficient; $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_q)'$ is a regression coefficient vector of z_{it}; W_0 is an $N \times N$ predetermined spatial weight matrix; $(W_0 Y_t)_i$ is the ith component of $W_0 Y_t$; ε_{it} are i.i.d. error terms with zero means and variance σ_ε^2; b_i are i.i.d. variables with zero means and variance σ_b^2, ε_{it} are independent of b_i. Let $\theta_0 = (\rho_0, \alpha_0', \sigma_{b0}^2, \sigma_{\varepsilon 0}^2)'$ be the true parameter vector of $\theta = (\rho, \alpha', \sigma_b^2, \sigma_\varepsilon^2)'$ and $\beta_0(u_{it})$ be the true varying coefficient function of $\beta(u_{it})$.

The model (1) can be simplified as the following matrix form:

$$Y = \rho W Y + X\beta(u) + Z\alpha + Ub + \varepsilon, \tag{2}$$

where $Y = (y_{11}, y_{12}, \cdots, y_{NT})'$; $X = (x'_{11}, x'_{12}, \cdots, x'_{NT})'$; $Z = (z'_{11}, z'_{12}, \cdots, z'_{NT})'$; $\varepsilon = (\varepsilon_{11}, \varepsilon_{12}, \cdots, \varepsilon_{NT})'$; $W = W_0 \otimes I_T$; $b = (b_1, b_2, \cdots, b_N)'$; $U = I_N \otimes e_T$; I_N is an $N \times N$ identity matrix; I_T is a $T \times T$ identity matrix; $\otimes$ denotes the Kronecker product; e_T is a $T \times 1$ vector consisting of 1.

Define $A(\rho) = I - \rho W$; then the model (2) can be rewritten as:

$$A(\rho)Y = X\beta(u) + Z\alpha + Ub + \varepsilon, \tag{3}$$

where I is an $NT \times NT$ unit matrix. For the model (3), it is easy to get the following facts:

$$\partial(Ub + \varepsilon)/\partial(Y) = A(\rho), \Sigma = Var(Ub + \varepsilon) = \sigma_\varepsilon^2 I + \sigma_b^2 I_N \otimes (e_T e'_T),$$

$$|\Sigma| = \sigma_\varepsilon^{2N(T-1)}(\sigma_b^2 + T\sigma_\varepsilon^2)^N, \Sigma^{-1} = \frac{1}{\sigma_\varepsilon^2} I + (\frac{1}{\sigma_\varepsilon^2 + T\sigma_b^2} - \frac{1}{T\sigma_b^2}) I_N \otimes (\frac{1}{T} e_T e'_T).$$

According to [12], the quasi-log-likelihood function of the model (3) can be written as follows:

$$\begin{aligned} lnL(\theta) = & -\frac{N(T-1)}{2} ln\sigma_\varepsilon^2 - \frac{N}{2} ln(\sigma_\varepsilon^2 + T\sigma_b^2) \\ & - \frac{1}{2(\sigma_\varepsilon^2 + T\sigma_b^2)} [A(\rho)Y - X\beta(u) - Z\alpha]' H [A(\rho)Y - X\beta(u) - Z\alpha] \\ & - \frac{1}{2\sigma_\varepsilon^2} [A(\rho)Y - X\beta(u) - Z\alpha]'(I - H)[A(\rho)Y - X\beta(u) - Z\alpha] + ln|A(\rho)| + c, \end{aligned} \tag{4}$$

where $H = I_N \otimes (T^{-1} e_T e'_T)$, and c is a constant.

By maximizing the above quasi-log-likelihood function with respect to θ, the quasi-maximum likelihood estimators of α, σ_b^2 and σ_ε^2 can be easily obtained as

$$\tilde{\alpha} = [Z'(I - S)(\frac{H}{\sigma_\varepsilon^2 + T\sigma_b^2} + \frac{I - H}{\sigma_\varepsilon^2})(I - S)Z]^{-1} Z'(I - S)(\frac{H}{\sigma_\varepsilon^2 + T\sigma_b^2} + \frac{I - H}{\sigma_\varepsilon^2}) A(\rho)Y, \tag{5}$$

$$\tilde{\sigma}_\varepsilon^2 = [N(T-1)]^{-1}[A(\rho)Y - X\beta(u) - Z\alpha]'(I - H)[A(\rho)Y - X\beta(u) - Z\alpha], \tag{6}$$

$$\tilde{\sigma}_b^2 = (NT)^{-1}[A(\rho)Y - X\beta(u) - Z\alpha]' H [A(\rho)Y - X\beta(u) - Z\alpha] - T^{-1}\tilde{\sigma}_\varepsilon^2. \tag{7}$$

By substituting (5)–(7) into (4), we have the concentrated quasi-log-likelihood function of ρ as

$$lnL(\rho) = \frac{-N(T-1)}{2} ln\tilde{\sigma}_\varepsilon^2 - \frac{N}{2} ln(\tilde{\sigma}_\varepsilon^2 + T\tilde{\sigma}_b^2) - \frac{NT}{2} + ln|A(\rho)|.$$

It is obvious that we cannot directly obtain the quasi-maximum likelihood estimator of ρ by maximizing the above formula because $\beta(u_{it})$ is an unknown function. In order to overcome this problem, we use the PQMLE method and working independence theory ([45,46]) to estimate the unknown parameters and varying coefficient functions of the model (1).

The main steps are as follows:

Step 1 Suppose that θ is known. $\beta_s(\cdot)$ can be approximated by the first-order Taylor expansion

$$\beta_s(u_{it}) \approx \beta_s(u) + (u_{it} - u)\dot{\beta}_s(u) \hat{=} a_{1s} + a_{2s}(u_{it} - u), s = 1, 2, \cdots, p$$

for u_{it} in a neighborhood of u, where a_{2s} is the first order derivative of $\beta(u)$. Let $a_1 = (a_{11}, a_{12}, \cdots, a_{1p})'$ and $a_2 = (a_{21}, a_{22}, \cdots, a_{2p})'$; then estimators of a_1 and a_2 can be obtained by

$$(\hat{a}_1, \hat{a}_2) = \arg\min_{a_1, a_2} \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} [\tilde{y}_{it} - \sum_{s=1}^{p} (a_{1s} + a_{2s}(u_{it} - u)) x_{its}]^2 K_h(u_{it} - u),$$

where $\tilde{y}_{it} = y_{it} - \rho(W_0 Y_t)_i - z'_{it}\alpha$, $K_h(u_{it} - u) = h^{-1}K((u_{it} - u)/h)$, $K(\cdot)$ is a kernel function and h is the bandwidth. Therefore, the feasible initial estimator of $\beta(u)$ can be obtained by $\hat{\beta}_{IN}(u) = \hat{a}_1 = (\hat{a}_{11}, \hat{a}_{12}, \cdots, \hat{a}_{1p})'$.

Denote $W_u = diag(K_h(u_{11} - u), K_h(u_{12} - u), \cdots, K_h(u_{NT} - u))$, $\tilde{Y} = (\tilde{y}_{11}, \tilde{y}_{12}, \cdots, \tilde{y}_{NT})'$, $\delta = (a_{11}, \cdots, a_{1p}, ha_{21}, \cdots, ha_{2p})'$ and $D_u = \begin{pmatrix} x_{11} & \cdots & x_{NT} \\ \frac{u_{11}-u}{h} x_{11} & \cdots & \frac{u_{NT}-u}{h} x_{NT} \end{pmatrix}'$, then we have

$$\hat{\delta} = \arg\min_{\delta} \frac{1}{NT} (\tilde{Y} - D_u\delta)' W_u (\tilde{Y} - D_u\delta).$$

Therefore, we obtain

$$\hat{\delta} = (D'_u W_u D_u)^{-1} D'_u W_u \tilde{Y}.$$

Let $S_u = \frac{1}{NT} D'_u W_u D_u$, $T_u = \frac{1}{NT} D'_u W_u \tilde{Y}$ and $e_0 = (I_p, 0_p)$. It is easy to know that

$$\hat{\beta}_{IN}(u) = (\hat{a}_{11}, \hat{a}_{12}, \cdots, \hat{a}_{1p})' = e_0' S_u^{-1} T_u, \tag{8}$$

where $e_0' S_u = s(u)$. Consequently, the initial estimator of $X\beta(u)$ is given by

$$X\hat{\beta}_{IN}(u) = \begin{pmatrix} (x'_{11} \quad 0)(D'_{u_{11}} W_{u_{11}} D_{u_{11}})^{-1} D'_{u_{11}} W_{u_{11}} \\ (x'_{12} \quad 0)(D'_{u_{12}} W_{u_{12}} D_{u_{12}})^{-1} D'_{u_{12}} W_{u_{12}} \\ \vdots \\ (x'_{NT} \quad 0)(D'_{u_{NT}} W_{u_{NT}} D_{u_{NT}})^{-1} D'_{u_{NT}} W_{u_{NT}} \end{pmatrix} \tilde{Y} \triangleq S\tilde{Y},$$

where

$$S = \begin{pmatrix} (x'_{11} \quad 0)(D'_{u_{11}} W_{u_{11}} D_{u_{11}})^{-1} D'_{u_{11}} W_{u_{11}} \\ (x'_{12} \quad 0)(D'_{u_{12}} W_{u_{12}} D_{u_{12}})^{-1} D'_{u_{12}} W_{u_{12}} \\ \vdots \\ (x'_{NT} \quad 0)(D'_{u_{NT}} W_{u_{NT}} D_{u_{NT}})^{-1} D'_{u_{NT}} W_{u_{NT}} \end{pmatrix}.$$

Step 2 Replacing $\beta(u)$ of (4) with $\hat{\beta}_{IN}(u)$, the estimator $\hat{\theta}$ of θ can be obtained by maximizing approximate quasi-log-likelihood function:

$$\begin{aligned} ln\hat{L}(\theta) = &-\frac{N(T-1)}{2} ln\sigma_\varepsilon^2 - \frac{1}{2(\sigma_\varepsilon^2 + T\sigma_b^2)} [A(\rho)Y - Z\alpha]'(I-S)'H(I-S)[A(\rho)Y - Z\alpha] \\ &- \frac{N}{2} ln(\sigma_\varepsilon^2 + T\sigma_b^2) - \frac{1}{2\sigma_\varepsilon^2} [A(\rho)Y - Z\alpha]'(I-S)'(I-H)(I-S)[A(\rho)Y - Z\alpha] \\ &+ ln|A(\rho)| + c. \end{aligned} \tag{9}$$

In the real estimation of θ, the procedure is realized by following steps:

Firstly, assume ρ is known. The initial estimators of σ_b^2, σ_ε^2 and α are obtained by maximizing (9). Then, we find

$$\hat{\sigma}^2_{\varepsilon IN} = [N(T-1)]^{-1} [\tilde{Y} - X\hat{\beta}_{IN}(u)]'(I-H)[\tilde{Y} - X\hat{\beta}_{IN}(u)], \tag{10}$$

$$\hat{\sigma}^2_{bIN} = (NT)^{-1} [\tilde{Y} - X\hat{\beta}_{IN}(u)]' H [\tilde{Y} - X\hat{\beta}_{IN}(u)] - T^{-1}\hat{\sigma}^2_{\varepsilon IN}, \tag{11}$$

$$\hat{\alpha}_{IN} = [Z'(I-S)(\frac{H}{\hat{\sigma}^2_{\varepsilon IN} + T\hat{\sigma}^2_{bIN}} + \frac{I-H}{\hat{\sigma}^2_{\varepsilon IN}})(I-S)Z]^{-1}Z'(I-S)(\frac{H}{\hat{\sigma}^2_{\varepsilon IN} + T\hat{\sigma}^2_{bIN}} + \frac{I-H}{\hat{\sigma}^2_{\varepsilon IN}})A(\rho)Y. \quad (12)$$

Secondly, with the estimated $\hat{\sigma}^2_{\varepsilon IN}$, $\hat{\sigma}^2_{bIN}$ and $\hat{\alpha}_{IN}$, update $\hat{\rho}$ by maximizing the concentrated quasi-log-likelihood function of ρ:

$$ln\tilde{L}(\rho) = -\frac{N(T-1)}{2}ln\hat{\sigma}^2_{\varepsilon IN} - \frac{N}{2}ln(\hat{\sigma}^2_{\varepsilon IN} + T\hat{\sigma}^2_{bIN}) + ln|A(\rho)| + c.$$

Therefore, the estimator of ρ is obtained by:

$$\hat{\rho} = \arg\max_{\rho} ln\tilde{L}(\rho).$$

Step 3 By substituting $\hat{\rho}$ into (10), (11) and (12), respectively, the final estimators of σ^2_b, σ^2_ε and α are computed as follows:

$$\hat{\sigma}^2_\varepsilon = [N(T-1)]^{-1}\hat{\hat{Y}}'(I-S)'(I-H)(I-S)\hat{\hat{Y}}, \quad (13)$$

$$\hat{\sigma}^2_b = (NT)^{-1}\hat{\hat{Y}}'(I-S)H(I-S)\hat{\hat{Y}} - T^{-1}\hat{\sigma}^2_\varepsilon,$$

$$\hat{\alpha} = [Z'(I-S)(\frac{H}{\hat{\sigma}^2_\varepsilon + T\hat{\sigma}^2_b} + \frac{I-H}{\hat{\sigma}^2_\varepsilon})(I-S)Z]^{-1}Z'(I-S)(\frac{H}{\hat{\sigma}^2_\varepsilon + T\hat{\sigma}^2_b} + \frac{I-H}{\hat{\sigma}^2_\varepsilon})A(\hat{\rho})Y.$$

Step 4 By replacing ρ with $\hat{\rho}$ in (8), we get the ultimate estimator of $\beta(u)$:

$$\hat{\beta}(u) = e_0'[D_u'W_uD_u]^{-1}D_u'W_u\hat{\hat{Y}}.$$

where $\hat{\hat{Y}} = A(\hat{\rho})Y - Z\hat{\alpha}$, $A(\hat{\rho}) = I - \hat{\rho}W$.

3. Asymptotic Properties for the Estimators

In this section, we focus on studying consistency and asymptotic normality of the PQMLEs given in Section 2. To prove these asymptotic properties, we need the following assumptions to hold.

To provide a rigorous analysis, we make the following assumptions.

Assumption 1. *(i) ε_{it} and b_i are uncorrelated to x_{it} and z_{it}, and satisfy $E(b_i|x_{it}, z_{it}) = 0$, $Var(b_i|x_{it}, z_{it}) = \sigma^2_b < \infty$, $E(\|b_i x'_{it}\|) < \infty$, $E(\|b_i z'_{it}\|) < \infty$, $E(\varepsilon_i|x_{it}, z_{it}) = 0$, $E(\|\varepsilon_i x'_{it}\|) < \infty$, $Var(\varepsilon_i|x_{it}, z_{it}) = \sigma^2_\varepsilon < \infty$ and $E(\|\varepsilon_i z'_{it}\|) < \infty$, where $\|\cdot\|$ represents the Euclidean norm. Moreover, $E(|\varepsilon|^r) < \infty$ and $E(|b|^r) < \infty$ for some $r > 4$.*

(ii) $\{(u_{it}, \varepsilon_{it})\}_{i=1,t=1}^{N,T}$ are i.i.d random sequences. The density function $f_t(u)$ of u_{it} is non-zero, uniformly bounded and second-order continuously differentiable on $\mathbb{U}$, where $\mathbb{U}$ is the supporting set of $K(u)$. $\Omega_{11}(u) = E(x_{it}x'_{it}|u_{it} = u) < \infty$ exists, and it has a second-order continuous derivative; $\Omega_{12}(u) = E(x_{it}z'_{it}|u_{it} = u)$ and $\Omega_{22}(u) = E(z_{it}z'_{it}|u_{it} = u)$ are second-order continuously differentiable on $\mathbb{U}$; $E(x_{it}x'_{is}|u_{it} = u) = 0$, $E(z_{it}z'_{is}|u_{it} = u) = 0$ and $E(x_{it}z'_{is}|u_{it} = u) = 0$ for $t \neq s$; $E(x_{it}x'_{it}|u_{it} = u) \neq 0$, $E(x_{it}\varepsilon_{it}|u_{it}) = 0$ and $E(x_{it}b_{it}|u_{it}) = 0$.

(iii) Real valued function $\beta_s(u)(s = 1, 2, \cdots, p)$ is second-order, continuously differentiable and satisfies the first order Lipschitz condition, $|x'_{it}\beta_s(u)| \leq m_\beta$ at any u, $|z'_{it}\alpha| \leq m_\alpha$, where m_β and m_α are positive and constant.

(iv) $\{x_{it}, z_{it}\}_{i=1,t=1}^{N,T}$ are i.i.d. random variables from the population. Moreover, $\max_{i,t} E(|x_{it}|^r) < \infty$ and $\max_{i,t} E(|z_{it}|^r) < \infty$ for $r \geq 4$.

Assumption 2. *(i) As a normalization, the diagonal elements ϖ_{ii} of W_0 are 0 for all i and ϖ_{ij} are at most of order l_N^{-1}, denoted by $O(1/l_N)$.*

(ii) The ratio $l_N/N \to 0$ as N goes to infinity.

(iii) $A(\rho)$ is nonsingular for any $|\rho| < 1$.

(iv) The matrices W_0 and $A^{-1}(\rho)$ are uniformly bounded in both row and column sums in absolute value.

Assumption 3. *$K(\cdot)$ is a nonnegative continuous even function. Let $\mu_l = \int k(v)v^l dv$, $v_l = \int k^2(v)v^l dv$; then $\mu_l = v_l = 0$ for any positive odd number. Meanwhile, $\mu_0 = 1$, $\mu_2 \neq 0$.*

Assumption 4. *If $N \to \infty$, $T \to \infty$ and $h \to 0$, then $NTh \to \infty$.*

Assumption 5. *There is an unique θ_0 to make the model (1) tenable.*

Assumption 6. *$\phi_{Q_0Z}\phi_{ZZ}^{-1}\phi'_{Q_0Z} + \phi_{Q_0Q_0} \geq 0$, where*

$$\phi_{Q_0Z} = \lim_{N\to\infty}\lim_{T\to\infty}\frac{1}{NT}Q_0'(I-S)'(I-S)Z, \quad \phi_{ZZ} = \lim_{N\to\infty}\lim_{T\to\infty}\frac{1}{NT}Z'(I-S)'(I-S)Z,$$

$$\phi_{Q_0Q_0} = \lim_{N\to\infty}\lim_{T\to\infty}\frac{1}{NT}Q_0'(I-S)'(I-S)Q_0, \quad Q_0 = G_0(Z\alpha_0 + X\beta_0), G_0 = WA^{-1}(\rho_0).$$

Remark 1. *Assumption 1 provides the essential features of the regressors and disturbances for the model (see [47]). Assumption 2 concerns the basic features of the spatial weights matrix and the parallel Assumptions 2–5 of [12]. Assumption 2(i) is always satisfied if l_N is a bounded sequence. We allow l_N to be divergent, but at a rate smaller than N, as specified in Assumption 2(ii). Assumption 2(iii) guarantees that model (2) has an equilibrium given by (3). Assumption 2(iv) is also assumed in [12], and it limits the spatial correlation to some degree but facilitates the study of the asymptotic properties of the spatial parameter estimators. Assumptions 3 and 4 concern the kernel function and bandwidth sequence. Assumption 5 offers a unique identification condition, and Assumption 6 is necessary for proof of asymptotic normality.*

In order to prove large sample properties of estimators, we need introduce the following useful lemmas. Before that, we simplify the model (2) and obtain reduced form equation of Y as follows:

$$Y = X\beta_0(u) + Z\alpha_0 + \rho G(Z\alpha_0 + X\beta_0(u)) + A^{-1}(\rho_0)(Ub + \varepsilon),$$

where $A^{-1}(\rho_0) = I + \rho_0 G_0$; $G_0 = WA^{-1}(\rho_0)$. The above equations are frequently used in a later derivation.

Lemma 1. *Let $\{(X_1, Y_1), (X_2, Y_2), \cdots, (X_n, Y_n)\}$ be i.i.d. random vectors, where Y_i are scalar random variables. Further, assume that $E|y|^r f(x,y)dy < \infty$, where f denotes the joint density of (X,Y). Let K be a bounded positive function with a bounded support, satisfying a Lipschitz condition. Given that $n^{2\tau-1}h \to \infty$ for some $\tau < 1 - r^{-1}$,*

$$\sup_x \left|\frac{1}{n}\sum_{i=1}^{n}[K_h(X_i - x)Y_i - EK_h(X_i - x)Y_i]\right| = O_p(\{\frac{\log(1/h)}{nh}\}^{1/2}).$$

The proof can be found in [48].

Lemma 2. *Under Assumptions 1–4, we have*

$$S_u \xrightarrow{p} F(u)\begin{pmatrix} \Omega_{11}(u) & 0 \\ 0 & \mu_2\Omega_{11}(u) \end{pmatrix},$$

where $F(u) = \lim_{T\to\infty} \frac{1}{T}\sum_{i=1}^{T} f_t(u)$ *and* $\Omega_{11} = E(x_{it}x_{it}'|u_{it} = u)$*,* $i = 1, 2, \cdots, N, t = 1, 2, \cdots, T$.

Proof See the Appendix A.

Lemma 3. *Under Assumptions 1–4, we have* $\hat{\beta}_{IN}(u) \xrightarrow{p} \beta_0(u)$.

Proof See the Appendix A.

Lemma 4. *Under Assumptions 1–4, we have*

(i) $\frac{1}{NT}Z'(I-S)'(I-H)(I-S)Z \xrightarrow{p} \Omega_{22}(u)\Omega_{11}^{-1}(u)\Omega_{12}(u)$;
(ii) $(I-S)X\beta(u) = op(1)$, $(I-S)Z\alpha = o_p(1)$;
(iii) $(I-S)GX\beta(u) = O_p(1)$, $(I-S)GZ\alpha = O_p(1)$;
(iv) $S(Ub+\varepsilon) = o_p(1)$, $SG(Ub+\varepsilon) = o_p(1)$.

The proof can be found in [37].

Lemma 5. *Under Assumptions 1–6,* $\Sigma_{\theta_0} = E[\frac{1}{\sqrt{NT}}\frac{\partial^2 lnL(\theta_0)}{\partial\theta\partial\theta'}]$ *is a positive definite matrix and*

$$\frac{1}{\sqrt{NT}}\frac{\partial^2 lnL(\theta_0)}{\partial\theta\partial\theta'} = \begin{pmatrix} \frac{\sigma_b^2 E[Q_0'PQ_0]}{N(\sigma_{\varepsilon0}^2+T\sigma_{b0}^2)\sigma_{\varepsilon0}^2} + \frac{\sigma_{b0}^2[tr(G_0'G_0)+tr(G_0'HG_0)]}{N\sigma_{\varepsilon0}^2} + \frac{tr(G_0'G_0)-tr(G_0^2)}{NT} & * & * & * \\ \frac{tr(G_0'H)}{NT} + \frac{(\sigma_{\varepsilon0}^2+T\sigma_{b0}^2)[tr(G_0')-tr(G_0'H)]}{NT\sigma_{\varepsilon0}^4} & -\frac{1}{2\sigma_{\varepsilon0}^4} & * & * \\ \frac{tr(G_0'H)}{NT(\sigma_{\varepsilon0}^2+T\sigma_{b0}^2)} & -\frac{1}{2(\sigma_{\varepsilon0}^2+T\sigma_{b0}^2)^2} & -\frac{T}{2(\sigma_{\varepsilon0}^2+T\sigma_{b0}^2)^2} & * \\ \frac{\sigma_{b0}^2 E[Q_0'PZ]}{N(\sigma_{\varepsilon0}^2+T\sigma_{b0}^2)\sigma_{\varepsilon0}^2} & 0 & 0 & -\frac{E[Z'PZ]}{NT\sigma_{\varepsilon0}^2(\sigma_{\varepsilon0}^2+T\sigma_{b0}^2)} \end{pmatrix} + o_p(1).$$

Proof See the Appendix A.

With the above lemmas, we state main results as follows. Their detailed proofs are given in the Appendix A.

Theorem 1. *Under Assumptions 1–5, we have* $\hat{\rho} - \rho_0 = op(1)$ *and* $\hat{\alpha} - \alpha_0 = op(1)$.

Theorem 2. *Under Assumptions 1–5, we have* $\hat{\beta}(u) - \beta_0(u) = op(1)$.

Theorem 3. *Under Assumptions 1–5, we have* $\hat{\sigma}_\varepsilon^2 - \sigma_{\varepsilon0}^2 = op(1)$ *and* $\hat{\sigma}_b^2 - \sigma_{b0}^2 = op(1)$.

Theorem 4. *Under Assumptions 1–6, we have*

$$\sqrt{NT}(\hat{\theta} - \theta_0) \xrightarrow{D} N(0, \Sigma_{\theta_0}^{-1} + \Sigma_{\theta_0}^{-1}\Omega_\theta\Sigma_{\theta_0}^{-1}).$$

where "$\xrightarrow{D}$*" represents convergence in distribution,* $\Sigma_{\theta_0} = -E[\frac{1}{\sqrt{NT}}\frac{\partial^2 lnL(\theta_0)}{\partial\theta\partial\theta'}]$ *is the average Hessian matrix (information matrix when ε and b are normally distributed, respectively),* $\Omega_\theta = E(\Omega_{\theta_{NT}})$ *and* $\Omega_{\theta_{NT}} = \frac{1}{\sqrt{NT}}\frac{\partial lnL(\theta_0)}{\partial\theta} \cdot \frac{1}{\sqrt{NT}}\frac{\partial lnL(\theta_0)}{\partial\theta'} + \frac{1}{NT}\frac{\partial^2 lnL(\theta_0)}{\partial\theta\partial\theta'}$.

Theorem 5. *Under Assumptions 1–5, we have*

$$\sqrt{NTh}(\hat{\beta}(u) - \beta_0(u) - \varphi(u)) \xrightarrow{D} N(0, \gamma^2(u)),$$

where $\gamma^2(u)$ *and* $\varphi(u)$ *satisfy* $\varphi(u) = \frac{1}{2}h^2\mu_2\ddot{\beta}(u)$ *and* $\gamma^2(u) = v_0(\sigma^2_{\varepsilon,0} + \sigma^2_{b,0})[F(u)\Omega_{11}(u)]^{-1}$, $F(u) = \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} f_t(u)$, $\Omega_{11}(u) = E(x_{it}x_{it}^{'}|u_{it} = u)$ *and* $\ddot{\beta}(u)$ *is the second-order derivative of* $\beta(u)$, $i = 1, 2, \cdots, N, t = 1, 2, \cdots, T$. *Furthermore, if* $NTh^5 \to 0$, *then we have*

$$\sqrt{NTh}(\hat{\beta}(u) - \beta_0(u)) \xrightarrow{D} N(0, \gamma^2(u)).$$

4. Monte Carlo Simulation

In this section, Monte Carlo simulations are presented, which were carried out to investigate the finite sample performance of PQMLEs. The sample standard deviation (SD) and two root mean square errors ($RMSE1$ and $RMSE2$) were used to measure the estimation performance. $RMSE1$ and $RMSE2$ are defined as:

$$\text{RMSE1} = [\frac{1}{mcn}\sum_{i=1}^{mcn}(\hat{\theta}_i - \theta_0)^2]^{\frac{1}{2}}, \quad \text{RMSE2} = (\hat{\theta}_{0.5} - \theta_0)^2 + \frac{\hat{\theta}_{0.75} - \hat{\theta}_{0.25}}{1.35},$$

where mcn is the number of iterations; $\hat{\theta}_i (i = 1, 2, \cdots, mcn)$ are estimates of θ for each iteration, θ_0 is the true value of θ; $\hat{\theta}_{0.25}$, $\hat{\theta}_{0.5}$ and $\hat{\theta}_{0.75}$ are the upper quartile, median and lower quartile of parametric estimates, respectively. For the nonparametric estimates, we took the mean absolute deviation error ($MADE$) as the evaluation criterion:

$$\text{MADE}_j = Q^{-1}\sum_{q=1}^{Q}|\hat{\beta}_j(u_q) - \beta_j(u_q)|, \ \ j = 1, 2, \cdots, mcn,$$

where $\{u_q\}_{q=1}^{Q}$ are Q fixed grid points in the support set of u. In the simulations, we applied the rule of thumb method of [48] to choose the optimal width and let kernel function be an Epanechnikov kernel $K(u) = (\frac{3}{4\sqrt{5}})(1 - \frac{1}{5}u^2)1(u^2 \le 5)$ (see [33]).

We ran a small simulation experiment with $mcn = 500$ and generated the simulation data from following model:

$$y_{it} = \rho(W_0Y_t)_i + x_{it}^{'}\beta(u_{it}) + z_{it}^{'}\alpha + b_i + \varepsilon_{it}, i = 1, 2, \cdots, N, t = 1, 2, \cdots, T,$$

where we assume that $x_{it} \sim U[-2, 2]$, $z_{it} \sim U[-2, 2]$, $u_{it} \sim U[-3, 3]$, $\varepsilon_{it} \sim i.i.d.N(0, 0.5)$, $b_i \sim i.i.d.N(0, 1)$, $\beta_1(u_{it}) = 0.5u_{it} + sin(1.5u_{it})$, $\beta_2(u_{it}) = u_{it}^2 + 0.5u_{it}$, $(\alpha_1, \alpha_2) = (1, 1.5)$ and $\rho = 0.25, 0.5$ and 0.75, respectively. Furthermore, we chose the Rook weight matrix (see [7]) to investigate the influence of the spatial weight matrix on the estimates. Our simulation results for both cases $T = 10$ and $T = 15$ are presented in Tables 1–6.

Table 1. The medians and SDs of MADE values for $\beta_1(u)$.

Parameter		**T = 10**				**T = 15**			
		N = 49	**N = 64**	**N = 81**	**N = 100**	**N = 49**	**N = 64**	**N = 81**	**N = 100**
	Median	0.1681	0.1519	0.1415	0.1326	0.1472	0.1340	0.1337	0.1301
$\rho = 0.25$	SD	0.0352	0.0337	0.0283	0.0247	0.0289	0.0248	0.0240	0.0203
	Median	0.1769	0.1617	0.1498	0.1424	0.1587	0.1475	0.1400	0.1385
$\rho = 0.5$	SD	0.0377	0.0338	0.0295	0.0256	0.0298	0.0296	0.0243	0.0220
	Median	0.1553	0.1421	0.1316	0.1219	0.1384	0.1205	0.1137	0.1048
$\rho = 0.75$	SD	0.0335	0.0280	0.0254	0.0239	0.0256	0.0226	0.0211	0.0200

Table 2. The medians and SDs of MADE values for $\beta_2(u)$.

Parameter		T = 10				T = 15			
		N = 49	N = 64	N = 81	N = 100	N = 49	N = 64	N = 81	N = 100
$\rho = 0.25$	Median	0.0944	0.0864	0.0777	0.0677	0.084	0.0724	0.0704	0.0677
	SD	0.0326	0.0313	0.0249	0.0224	0.0274	0.0230	0.0201	0.0178
$\rho = 0.5$	Median	0.0985	0.0891	0.0836	0.0768	0.0868	0.0785	0.074	0.0724
	SD	0.0322	0.0304	0.0258	0.0238	0.0276	0.0234	0.0206	0.0186
$\rho = 0.75$	Median	0.0957	0.0863	0.0795	0.0729	0.0807	0.0743	0.0648	0.0608
	SD	0.0314	0.0257	0.0211	0.0213	0.0254	0.0223	0.0191	0.0173

Table 3. The results of parametric estimates with T = 10(1).

Parameter	True Value	N = 49				N = 64			
		MEAN	SD	RMSE1	RMSE2	MEAN	SD	RMSE1	RMSE2
ρ	0.2500	0.2836	0.0685	0.0859	0.0657	0.2720	0.0662	0.0851	0.0622
σ_ε^2	0.5000	0.5949	0.0625	0.1136	0.0608	0.5844	0.0619	0.1046	0.0592
σ_b^2	1.0000	0.9070	0.1971	0.2176	0.1979	0.9226	0.1786	0.1944	0.1874
α_1	1.0000	0.8755	0.0511	0.0072	0.0057	0.8879	0.0439	0.0065	0.0042
α_2	1.5000	1.5721	0.0111	0.0155	0.0097	1.5983	0.0039	0.0126	0.0052
ρ	0.5000	0.5236	0.0532	0.0896	0.0599	0.5222	0.0480	0.0878	0.0544
σ_ε^2	0.5000	0.6026	0.0729	0.1258	0.0688	0.5913	0.0644	0.1116	0.0639
σ_b^2	1.0000	0.8785	0.2026	0.2360	0.2107	0.9059	0.1819	0.2046	0.1999
α_1	1.0000	0.8690	0.0058	0.0076	0.0098	0.9145	0.0036	0.0049	0.0044
α_2	1.5000	1.5764	0.0079	0.0172	0.0286	1.6692	0.0056	0.0073	0.0058
ρ	0.7500	0.8083	0.0253	0.0635	0.0295	0.8110	0.0246	0.0657	0.0294
σ_ε^2	0.5000	0.6059	0.0702	0.1270	0.0734	0.6048	0.0659	0.1237	0.0679
σ_b^2	1.0000	0.8864	0.2063	0.2352	0.2212	0.8623	0.1694	0.2181	0.1944
α_1	1.0000	0.8421	0.0472	0.0091	0.0009	0.8935	0.0442	0.0061	0.0007
α_2	1.5000	1.5260	0.0023	0.0249	0.0002	1.5152	0.0012	0.0113	0.0001

Table 4. The results of parametric estimates with T = 10(2).

Parameter	True Value	N = 81				N = 100			
		MEAN	SD	RMSE1	RMSE2	MEAN	SD	RMSE1	RMSE2
ρ	0.2500	0.3033	0.0578	0.0785	0.0620	0.3051	0.0527	0.0762	0.0531
σ_ε^2	0.5000	0.5585	0.0456	0.0741	0.0432	0.5530	0.0389	0.0657	0.0369
σ_b^2	1.0000	0.9289	0.1596	0.1745	0.1691	0.9616	0.1524	0.1569	0.1477
α_1	1.0000	0.9468	0.0364	0.0032	0.0075	1.0757	0.0363	0.0044	0.0036
α_2	1.5000	1.5410	0.0402	0.0476	0.0089	1.5246	0.0373	0.0057	0.0039
ρ	0.5000	0.5112	0.0396	0.0814	0.0465	0.5033	0.0339	0.0808	0.0414
σ_ε^2	0.5000	0.5765	0.0505	0.0916	0.0537	0.5679	0.0436	0.0806	0.0457
σ_b^2	1.0000	0.9203	0.1575	0.1763	0.1481	0.9352	0.1480	0.1613	0.1565
α_1	1.0000	0.8403	0.0082	0.0092	0.0017	0.9422	0.0014	0.0033	0.0015
α_2	1.5000	1.4713	0.0034	0.0255	0.0008	1.4743	0.0032	0.0033	0.0007
ρ	0.7500	0.8094	0.0208	0.0629	0.0218	0.8101	0.0178	0.0627	0.0215
σ_ε^2	0.5000	0.5823	0.0548	0.1057	0.0600	0.5821	0.0522	0.0933	0.052
σ_b^2	1.0000	0.8966	0.1520	0.1837	0.1669	0.9054	0.1464	0.1741	0.1610
α_1	1.0000	0.8568	0.0061	0.0083	0.0088	0.9889	0.0035	0.0006	0.0005
α_2	1.5000	1.6519	0.0022	0.0205	0.0231	1.4913	0.0017	0.0001	0.0001

Table 5. The results of parametric estimates with T = 15(1).

Parameter	True	N = 49				N = 64			
	Value	MEAN	SD	RMSE1	RMSE2	MEAN	SD	RMSE1	RMSE2
ρ	0.2500	0.2751	0.0670	0.0801	0.0608	0.2740	0.0637	0.0780	0.0614
σ_ε^2	0.5000	0.5761	0.0556	0.0942	0.0548	0.5626	0.0523	0.0815	0.0385
σ_b^2	1.0000	0.9374	0.2005	0.1989	0.1046	0.9363	0.1752	0.1861	0.1872
α_1	1.0000	1.1449	0.0356	0.0054	0.0049	0.9441	0.0335	0.0032	0.0025
α_2	1.5000	1.4568	0.0052	0.0210	0.0073	1.5855	0.0034	0.0031	0.0019
ρ	0.5000	0.5749	0.0517	0.0910	0.0613	0.5766	0.0433	0.0879	0.0439
σ_ε^2	0.5000	0.5922	0.0596	0.1098	0.0614	0.5779	0.0598	0.0982	0.0515
σ_b^2	1.0000	0.9193	0.2160	0.2303	0.2313	0.9126	0.1729	0.1935	0.1814
α_1	1.0000	1.0623	0.0202	0.0036	0.0101	1.0652	0.0021	0.0038	0.0038
α_2	1.5000	1.3248	0.0026	0.0039	0.0307	1.3651	0.0023	0.0043	0.0182
ρ	0.7500	0.8118	0.0270	0.0674	0.0331	0.8146	0.0226	0.0684	0.0264
σ_ε^2	0.5000	0.6058	0.0707	0.1272	0.0671	0.5929	0.0611	0.1112	0.0622
σ_b^2	1.0000	0.8746	0.2018	0.2173	0.2024	0.8788	0.1688	0.2076	0.1708
α_1	1.0000	1.0408	0.0093	0.0024	0.0008	1.0420	0.0046	0.0024	0.0016
α_2	1.5000	1.4716	0.0006	0.0017	0.0002	1.4956	0.0004	0.0015	0.0008

Table 6. The results of parametric estimates with T = 15(2).

Parameter	True	N = 81				N = 100			
	Value	MEAN	SD	RMSE1	RMSE2	MEAN	SD	RMSE1	RMSE2
ρ	0.2500	0.2662	0.0545	0.0714	0.0576	0.2646	0.0493	0.0664	0.0525
σ_ε^2	0.5000	0.5566	0.0441	0.0717	0.0397	0.5490	0.0392	0.0627	0.0330
σ_b^2	1.0000	0.9508	0.1555	0.1629	0.1524	0.9772	0.1507	0.1522	0.1565
α_1	1.0000	0.9450	0.0202	0.0031	0.0055	0.9862	0.0008	0.0106	0.0024
α_2	1.5000	1.5622	0.0413	0.0030	0.0169	1.3702	0.0332	0.0028	0.0017
ρ	0.5000	0.5750	0.0343	0.0733	0.0448	0.4930	0.0297	0.0063	0.0204
σ_ε^2	0.5000	0.5633	0.0488	0.0800	0.0476	0.5364	0.0366	0.0602	0.0433
σ_b^2	1.0000	0.9311	0.1517	0.1634	0.1525	0.9532	0.0880	0.1337	0.0935
α_1	1.0000	0.9822	0.0351	0.0026	0.0018	0.9673	0.0283	0.0013	0.0013
α_2	1.5000	1.4033	0.0029	0.0020	0.0005	1.4980	0.0203	0.0011	0.0007
ρ	0.7500	0.8156	0.0205	0.0687	0.0235	0.8066	0.0153	0.0543	0.0137
σ_ε^2	0.5000	0.5904	0.0522	0.0974	0.0564	0.5806	0.0501	0.0949	0.0512
σ_b^2	1.0000	0.8806	0.1486	0.1783	0.1680	0.9254	0.1369	0.1635	0.1597
α_1	1.0000	1.0338	0.0049	0.0023	0.0077	0.9856	0.0026	0.0002	0.0005
α_2	1.5000	1.3659	0.0326	0.0018	0.0180	1.4977	0.0008	0.0001	0.0001

By observing the simulation results in Tables 1–6, one can obtain the following findings: (1) The $RMSE1s$ and $RMSE2s$ for $\hat{\rho}$, $\hat{\sigma}_\varepsilon^2$, $\hat{\alpha}_1$ and $\hat{\alpha}_2$ were fairly small for almost all cases, and they decreased as N increased. The SDs and $RMSEs$ for $\hat{\sigma}_b^2$ are not negligible for small sample sizes, but they decreased as N increased. (2) For fixed T, as N increased, the SDs for $\hat{\rho}$, $\hat{\sigma}_\varepsilon^2$, $\hat{\sigma}_b^2$, $\hat{\alpha}_1$ and $\hat{\alpha}_2$ decreased for all cases. $RMSE1s$ and $RMSE2s$ for $\hat{\alpha}_1$ and $\hat{\alpha}_2$ decreased rapidly, whereas the $RMSEs$ for estimates of the others parameters did not change much. For fixed N, as T increased, the behavior of the estimates of parameters was similar to the case where N changed under the fixed T. (3) The SDs for $MADEs$ of varying coefficient functions $\beta_1(u)$ and $\beta_2(u)$ decreased as T or N increased. Combined with the above three findings, we conclude that the estimates of the parameter and unknown varying coefficient functions were convergent.

Figures 1 and 2 present the fitting results and 95% confidence intervals of $\beta_1(u)$ and $\beta_2(u)$ under $\rho = 0.5$, respectively, where the short dashed curves are the average fits over 500 simulations $\hat{\beta}_s(u)$ $(s = 1, 2)$ by PQMLE, the solid curves are the true values of $\beta_s(u)$ $(s = 1, 2)$ and the two long dashed curves are the corresponding 95% confidence bands. By observing every subgraph in Figures 1 and 2, we can see that the short dashed curve is fairly close to solid curve and the corresponding confidence bandwidth is narrow. This illustrates that the nonparametric estimation procedure works well for small samples. To save space, we do not present the cases $\rho = 0.25$ and $\rho = 0.75$ because they had similar results as the case $\rho = 0.5$.

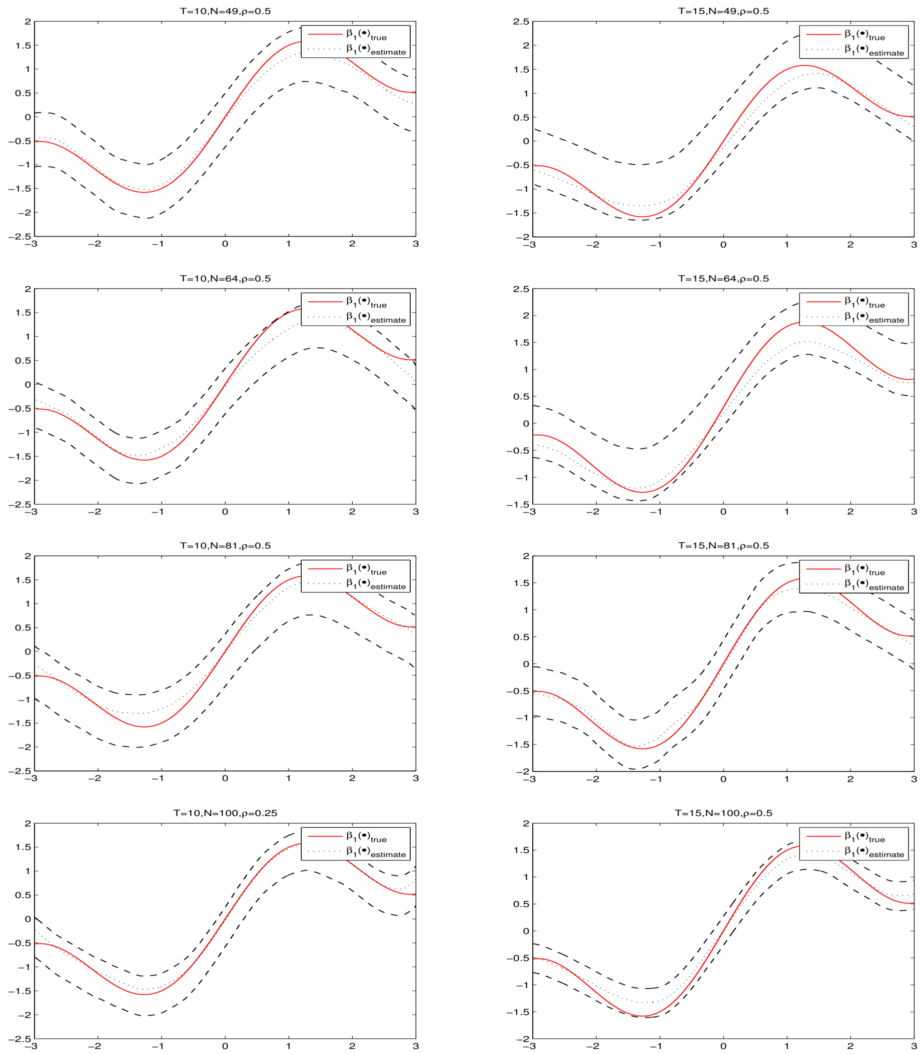

Figure 1. The fitting results and 95% confidence intervals of $\beta_1(u)$ under $\rho = 0.5$.

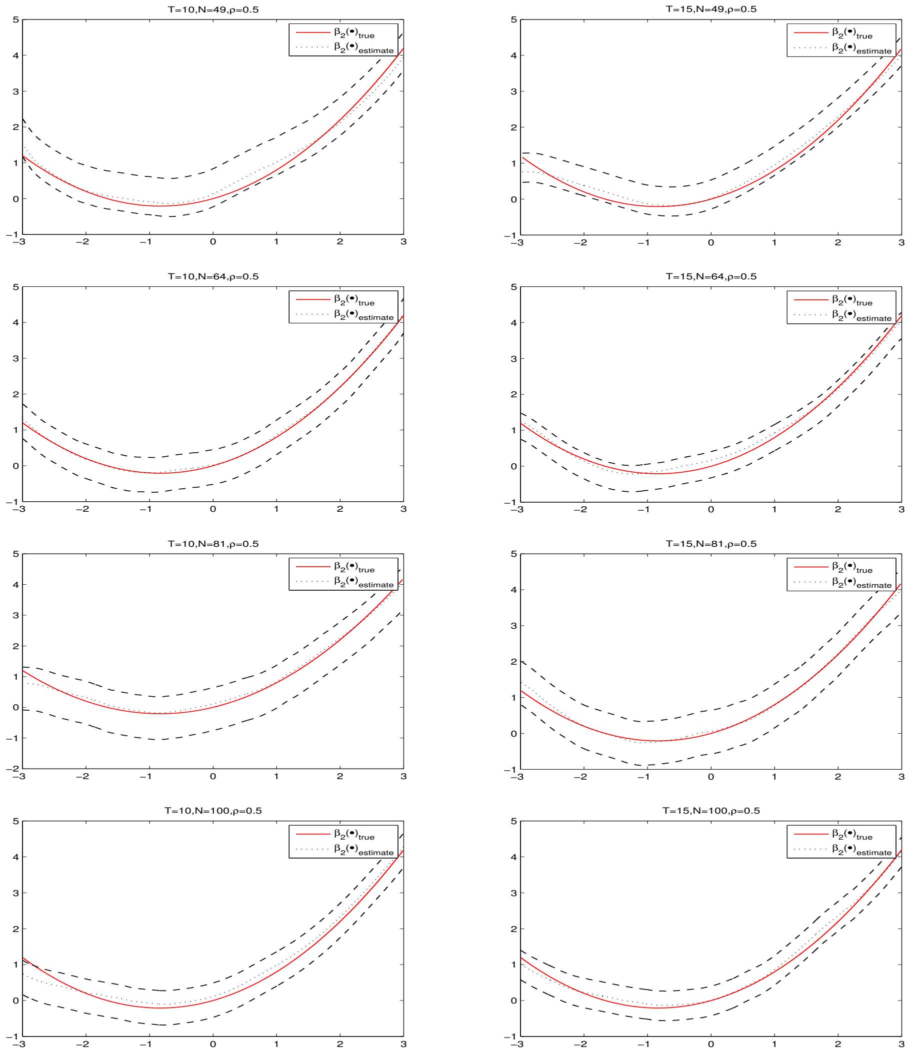

Figure 2. The fitting results and 95% confidence intervals of $\beta_2(u)$ under $\rho = 0.5$.

5. Real Data Analysis

We applied the housing prices of Chinese city data to analyze the proposed model with real data. The data were obtained from the China Statistical Yearbook, the China City Statistical Yearbook and the China Statistical Yearbook for Regional Economies. Based on the panel data of related variables of 287 cities at/above the prefecture level (except the cities in Taiwan, Hong Kong and Macau) in China from 2011 to 2018, we explored the influencing factors of housing prices of Chinese cities by PLVCSARPM with random effects.

Taking [49–51] as references, we collected nine variables related to housing prices of China cities, including each city's average selling price of residential houses (denoted by HP, yuan/sq.m), the expectation of housing price trends (EHP, %), population density (POD, person), annual per capita disposable income of urban households (ADI, yuan), loan–to–GDP ratio (MON, %), natural growth rate of population (NGR,), sulphur dioxide emission (SDE, 10,000 tons), area of grassland (AOG, 10,000 hectares) and the value of buildings completed (VBC, yuan/sq.m). According to result of non-linear regression, the established model is given by

$$y_{it} = \rho(W_0 Y_t)_i + x_{it}'\beta(u_{it}) + z_{it}'\alpha + b_i + \varepsilon_{it}, 1 \le i \le 287, 1 \le t \le 8, \tag{14}$$

where y_{it} represents the ith observation of ln(HP) at time t; z_{it} represents the ith observation of EHP at time t; x_{it} means the ith observation of POD, ln(ADI), MON, NGR, SDE and AOG at time t, respectively, u_{it} represents the ith observation of ln(VBC) at time t.

In order to transfer the asymmetric distribution of POD to nearly uniform distribution on (0,1), we set

$$POD = \big(POD^{1/3} - \min(POD^{1/3})\big) / \big(\max(POD^{1/3}) - \min(POD^{1/3})\big).$$

The spatial weight matrix W_0 we adopted is calculated as follows:

$$w_{ij} = exp\big(-\|s_i - s_j\|\big) / \sum_{k=1}^{264} exp\big(-\|s_i - s_k\|\big),$$

where $s_i = (longitude_i, latitude_i)$.

Figure 3 presents the estimation results and corresponding 95% confidence intervals of varying coefficient functions $\beta_s(s = 1, 2, \cdots, 6)$, where the black dashed line curves are the average fits over 500 simulations and the red solid lines are the corresponding 95% confidence bands. It can be seen from Figure 3 that all covariates variables have obvious non-linear effects on housing prices of Chinese cities.

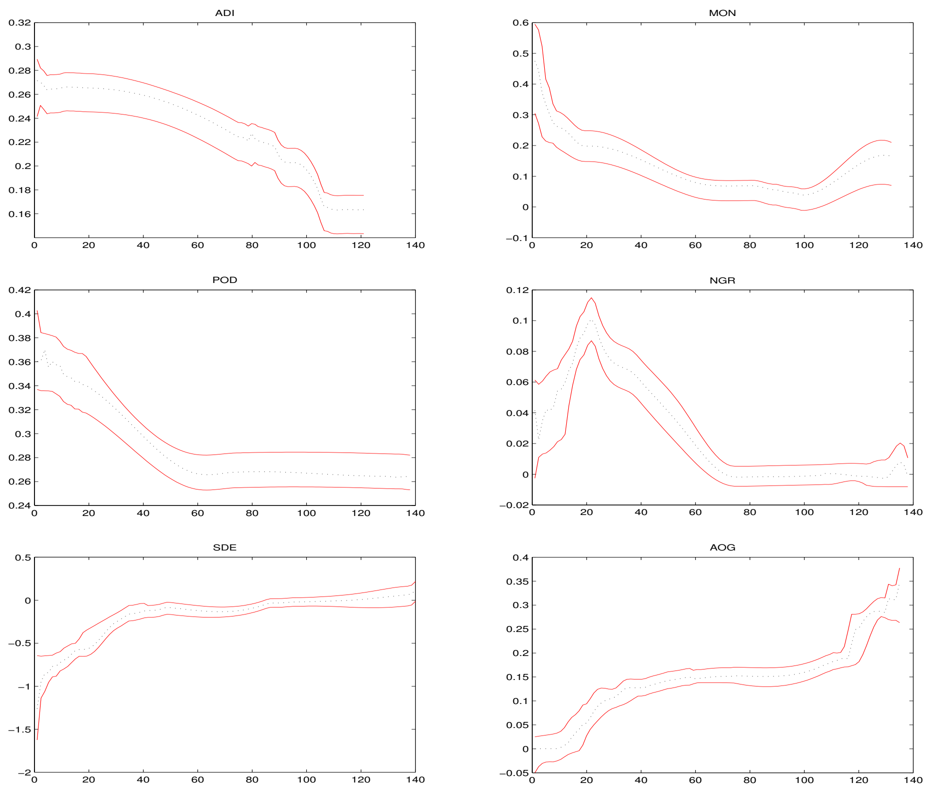

Figure 3. The estimated functions $\beta_s(u)$ and corresponding 95% confidence intervals for the housing prices of Chinese cities.

The estimation results of parameters in the model (14) are reported in Table 7. It can be seen from Table 7: (1) all estimates of parameters are all significant; (2) spatial correlation coefficient $\hat{\rho} = 0.3820 > 0$, which means there exists a spatial spill over effect for housing prices of Chinese cities; (3) $\hat{\alpha} = 0.9869 > 0$ indicates that the expectation of housing price trends (EHP) has a promotional effect on housing prices of Chinese cities; (4) $\hat{\sigma}_{\varepsilon}^2 = 0.0163$ shows that the growth of housing prices in different regions is relatively stable and is less affected by external fluctuations.

Table 7. The estimation results of unknown parameters in the model (14).

	$\hat{\rho}$	$\hat{\alpha}$	$\hat{\sigma}_b^2$	$\hat{\sigma}_\varepsilon^2$
Estimators	0.3820 ***	0.9869 ***	0.1375 **	0.0163 **
SD	0.0243	0.0397	5.0344	0.3687

Notes: ** and *** are significant at the significance levels of 5% and 1%, respectively.

6. Conclusions

In this paper, we proposed PQMLE of PLVCSARPM with random effects. Our model has the following advantages: (1) It can overcome the "curse of dimensionality" in the nonparametric spatial regression model effectively. (2) It can simultaneously study the linear and non-linear effects of coveriates. (3) It can investigate the spatial correlation effect of response variables. Under some regular conditions, consistency and asymptotic normality of the estimators for parameters and varying coefficient functions were derived. Monte Carlo simulations showed the proposed estimators are well behaved with finite samples. Furthermore, the performance of the proposed method was also assessed on a set of asymmetric real data.

This paper only focused on the PQMLE of PLVCSARPM with random effects. In future research, we may try to extend our method to more general models, such as a partially linear varying coefficient spatial autoregressive model with autoregressive disturbances. In addition, we also need study the issues of Bayesian analysis, variable selection and quantile regression in these models.

Author Contributions: Supervision, S.L. and J.C.; software, S.L.; methodology, S.L.; writing—original draft preparation, S.L.; writing—review and editing, S.L., J.C. and D.C. All authors have read and agreed to the published version of manuscript.

Funding: This work was supported by the Natural Science Foundation of Fujian Province (2020J01170, 2021J01662); the Natural Science Foundation of China (12001105); the Program for Innovative Research Team in Science and Technology in Fujian Province University; and the Fujian Normal University Innovation Team Foundation "Probability and Statistics: Theory and Application" (IRTL1704).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are openly available in Reference [50].

Acknowledgments: The authors are deeply grateful to editors and anonymous referees for their careful reading and insightful comments. Their comments led us to significantly improve the paper.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

To proceed with the proofs of main lemmas and theorems, we first provide two frequently used evident facts (see [14]).

Fact 1. If the row and column sums of the $n \times n$ matrices $\mathbf{B_{1n}}$ and $\mathbf{B_{2n}}$ are uniformly bounded in absolute value, then the row and column sums of $\mathbf{B_{1n}B_{2n}}$ are also uniformly bounded in absolute value.

Fact 2. If the row (resp. column) sums of $\mathbf{B_{1n}}$ are uniformly bounded in absolute value and $\mathbf{B_{2n}}$ is a conformable matrix whose elements are uniformly $O(o_n)$, then so are the elements of $\mathbf{B_{1n}B_{2n}}$ (resp. $\mathbf{B_{2n}B_{1n}}$).

Proof of Lemma 2. Recall $S_u = \frac{1}{NT} D_u' W_u D_u$. By straightforward calculation, it is not difficult to get

$$S_u = \begin{pmatrix} S_{u0} & S_{u1} \\ S_{u1} & S_{u2} \end{pmatrix},$$

where $S_{ul} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} x_{it}x_{it}(\frac{u_{it}-u}{h})^l k_h(u_{it}-u), l = 0,1,2$. Denote $D_{itl} = x_{it}x'_{it}k_h(u_{it} - u)(\frac{u_{it}-u}{h})^l$, $C_{tl} = \frac{1}{N}\sum_{i=1}^{N} D_{itl}$, then we have

$$S_{ul} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} D_{itl} = \frac{1}{T}\sum_{t=1}^{T} C_{tl}.$$

By Assumptions 1–3, we know

$$\begin{aligned}
EC_{tl} =& \frac{1}{N}\sum_{i=1}^{N} ED_{itl} = \frac{1}{N}\sum_{i=1}^{N} E[x_{it}x'_{it}k_h(u_{it}-u)(\frac{u_{it}-u}{h})^l] \\
=& E\{E[x_{it}x'_{it}k_h(u_{it}-u)(\frac{u_{it-u}}{h})^l|_{u_{it}}]\} \\
=& E\{k_h(u_{it}-u)(\frac{u_{it}-u}{h})^l E(x_{it}x'_{it}|u_{it})\} \\
=& E\{k_h(u_{it}-u)(\frac{u_{it}-u}{h})^l \Omega_{11}(u_{it})\} \\
=& \int h^{-1}k((u_{it}-u)/h)f_t(u_{it})(\frac{u_{it}-u}{h})^l \Omega_{11}(u_{it})du_{it} \\
=& \int k(v)f_t(u+hv)v^l\Omega_{11}(u+hv)dv \\
=& \int k(v)[f_t(u) + hf'_t(u)v + \frac{h^2}{2}f''_t(\xi)v^2]v^l\Omega_{11}(u+hv)dv \\
=& (1+O(h^2))f_t(u)\mu_l\Omega_{11}(u),
\end{aligned}$$

where $\mu_l = \int v^l k(v)dv$ and $v = \frac{u_{it}-u}{h}$, $D_{it}(i = 1,2,\cdots,N)$ are *i.i.d* for any fixed t and l. According to Khinchine law of large numbers, we have $C_{tl} \xrightarrow{p} EC_{tl}$. Therefore, it is not hard to get that

$$\begin{aligned}
S_{ul} =& \frac{1}{T}\sum_{t=1}^{T} C_{tl} = \frac{1}{T}\sum_{t=1}^{T}[(1+op(1))EC_{tl}] \\
=& \frac{1}{T}\sum_{t=1}^{T}[f_t(u)\mu_l\Omega_{11}(u) + O(h^2)](1+op(1)) \\
=& \frac{1}{T}\sum_{t=1}^{T}[f_t(u)\mu_l\Omega_{11}(u)] + O(h^2).
\end{aligned}$$

Let $F(u) = \underset{T\to\infty}{lim}\frac{1}{T}\sum_{t=1}^{T} f_t(u)$, then $S_{ul} \xrightarrow{p} \mu_l F(u)\Omega_{11}(u)$. Hence,

$$S_u = \begin{pmatrix} S_{u0} & S_{u1} \\ S_{u1} & S_{u2} \end{pmatrix} \xrightarrow{p} F(u)\begin{pmatrix} \Omega_{11}(u) & 0 \\ 0 & \mu_2\Omega_{11}(u) \end{pmatrix}.$$

□

Proof of Lemma 3. Recall $T_u = \frac{1}{NT}D'_u W_u \tilde{Y}$. By straightforward calculation, we obtain

$$T_u = \begin{pmatrix} \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} x_{it}\tilde{y}_{it}k_h(u_{it}-u) \\ \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} x_{it}\tilde{y}_{it}(\frac{u_{it}-u}{h})^1 k_h(u_{it}-u) \end{pmatrix} \triangleq \begin{pmatrix} T_{u,0} \\ T_{u,1} \end{pmatrix}$$

where $T_{ul} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} x_{it}\tilde{y}_{it}(\frac{u_{it}-u}{h})^l k_h(u_{it}-u)$ for $l = 0, 1$. Then, it is not hard to obtain

$$\begin{aligned} T_{ul} =& \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} x_{it}\tilde{y}_{it}(\frac{u_{it}-u}{h})^l k_h(u_{it}-u) \\ =& \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} x_{it}x_{it}'\beta_0(u_{it})(\frac{u_{it}-u}{h})^l k_h(u_{it}-u) \\ &+ \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} x_{it}(b_i+\varepsilon_{it})(\frac{u_{it}-u}{h})^l k_h(u_{it}-u). \end{aligned} \tag{A1}$$

For the first term of above equality in (A1), by similar proof procedures of Lemma 2, it is easily to get that

$$\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} x_{it}x_{it}'\beta_0(u_{it})(\frac{u_{it}-u}{h})^l k_h(u_{it}-u) \xrightarrow{p} F(u)\mu_l\Omega_{11}(u)\beta_0(u).$$

For the second term of above equality in (A1), we obtain

$$\begin{aligned} &E[\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} x_{it}(b_i+\varepsilon_{it})(\frac{u_{it}-u}{h})^l k_h(u_{it}-u)] \\ &= E[E[\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} x_{it}(b_i+\varepsilon_{it})(\frac{u_{it}-u}{h})^l k_h(u_{it}-u)|u_{it}]] \\ &= E[\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} E[x_{it}(b_i+\varepsilon_{it})|u_{it}](\frac{u_{it}-u}{h})^l k_h(u_{it}-u)] \\ &= 0, \end{aligned}$$

and

$$\begin{aligned} &Var(\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} x_{it}(b_i+\varepsilon_{it})(\frac{u_{it}-u}{h})^l k_h(u_{it}-u)) \\ &= \frac{1}{(NT)^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\{E[k_h^2(u_{it}-u)(\frac{u_{it}-u}{h})^{2l} x_{it}x_{it}']E[(b_i+\varepsilon_{it})^2]\} \\ &= o(1). \end{aligned}$$

Therefore,

$$T_u = \begin{pmatrix} T_{u0} \\ T_{u1} \end{pmatrix} \xrightarrow{p} F(u)\begin{pmatrix} \mu_l\Omega_{11}(u)\beta_0(u) \\ 0 \end{pmatrix}$$

Furthermore, by using (8) and Lemma 2, we easily know $\hat{\beta}_{IN}(u) \xrightarrow{p} \beta_0(u)$. □

Proof of Lemma 5. Its proof is analogous to Lemma 4.2 in [34] and Theorem 3.2 in [12]. The major difference is that our model is panel model with random effects instead of sectional model.

In order to prove Σ_{θ_0} is positive definite, we need only prove that $\kappa = 0$ according to $\Sigma_{\theta_0}\kappa = 0$, where $\kappa = (\kappa_1, \kappa_2, \kappa_3, \kappa_4')'$, $\kappa_1, \kappa_2, \kappa_3$ are constants, and κ_4 is a q dimension column vector.

It follows from $\Sigma_{\theta_0}\kappa = 0$ that:

$$\kappa_2 = \frac{2(\sigma_{\varepsilon 0}^2 + T\sigma_{b0}^2)tr(G_0'H)[T(\sigma_{\varepsilon 0}^2 + T\sigma_{b0}^2)\sigma_\varepsilon^4 + T(\sigma_{\varepsilon 0}^2 + T\sigma_{b0}^2) - \sigma_{\varepsilon 0}^4]}{NT(\sigma_{\varepsilon 0}^2 + T\sigma_{b0}^2)^2 + N\sigma_{\varepsilon 0}^4}\kappa_1,$$

$$\kappa_3 = -\frac{tr(G_0'H)(\sigma_{\varepsilon 0}^2 + T\sigma_{b0}^2)[2NT(\sigma_{\varepsilon 0}^2 + T\sigma_{b0}^2) + 2N\sigma_{\varepsilon 0}^4 - T^2\sigma_{\varepsilon 0}^4(\sigma_{\varepsilon 0}^2 + T\sigma_{b0}^2) + T^2(\sigma_{\varepsilon 0}^2 + T\sigma_{b0}^2) - T\sigma_{\varepsilon 0}^4]}{NT^2(\sigma_{\varepsilon 0}^2 + T\sigma_{b0}^2) + NT\sigma_{\varepsilon 0}^4}\kappa_1$$

and

$$\kappa_4 = T\sigma_{b0}^2 \phi_{ZZ}^{-1} \phi_{Q_0 Z} \kappa_1.$$

By straightforward calculations, it can be simplified to prove

$$(T\sigma_{b0}^2 \phi_{Q_0 Z} \phi_{ZZ}^{-1} \phi'_{Q_0 Z} + \phi_{Q_0 Q_0})\kappa_1 + c_3\kappa_1 + [(\sigma_{\varepsilon 0}^2 + T\sigma_{b0}^2)^2 tr(G_0' G_0) - (\sigma_{\varepsilon 0}^2 + T\sigma_b^2)\sigma_{\varepsilon 0}^2 tr(G_0^2)]\kappa_1 = 0,$$

where $c_3 = T\sigma_{\varepsilon 0}^4 (\sigma_{\varepsilon 0}^2 + \sigma_{b0}^2) + T(\sigma_{\varepsilon 0}^2 + \sigma_{b0}^2) - \sigma_{\varepsilon 0}^4 > 0$.

By Lemma 4 in [43], it is easy to obtain that $(\sigma_{\varepsilon 0}^2 + T\sigma_{b0}^2)^2 tr(G_0' G_0) - (\sigma_{\varepsilon 0}^2 + T\sigma_{b0}^2)\sigma_\varepsilon^2 tr(G_0^2) > 0$. Furthermore, according to Assumption 6, $T\sigma_{b0}^2 \phi_{Q_0 Z} \phi_{ZZ}^{-1} \phi'_{Q_0 Z} + \phi_{Q_0 Q_0} \geq 0$. Thus, we have $\kappa_1 = 0$ and so $\kappa = 0$. This completes the proof. □

Proof of Theorem 1. Let $\{\hat{\alpha}', \hat{\rho}\} = \arg\max_{\alpha',\rho} ln\tilde{L}(\alpha', \rho)$ and $\{\alpha_0', \rho_0\} = \arg\max_{\alpha',\rho} lnL(\alpha', \rho)$. We adopt the idea of Theorem 3.2 in [12] to prove the consistency of $\hat{\rho}$ and $\hat{\alpha}$. By the Consistency of Extrema Estimators in [52], it suffices to show that

$$\frac{1}{NT}[ln\tilde{L}(\alpha', \rho) - lnL(\alpha', \rho)] = op(1),$$

where

$$ln\tilde{L}(\alpha', \rho) = -\frac{N(T-1)}{2} ln\hat{\sigma}_{\varepsilon_{IN}}^2 - \frac{N}{2} ln(\hat{\sigma}_{\varepsilon_{IN}}^2 + T\hat{\sigma}_{b_{IN}}^2) + ln|A(\rho)|,$$

and

$$lnL(\alpha', \rho) = -\frac{N(T-1)}{2} ln\tilde{\sigma}_\varepsilon^2 - \frac{N}{2} ln(\tilde{\sigma}_\varepsilon^2 + T\tilde{\sigma}_b^2) + ln|A(\rho)|.$$

By straightforward calculation, we have

$$\frac{1}{NT}[ln\tilde{L}(\alpha', \rho) - lnL(\alpha', \rho)] = \frac{T-1}{2T}(ln\tilde{\sigma}_\varepsilon^2 - ln\hat{\sigma}_{\varepsilon IN}^2) + \frac{1}{2T}[ln(\tilde{\sigma}_\varepsilon^2 + T\tilde{\sigma}_b^2) - ln(\hat{\sigma}_{\varepsilon IN}^2 + T\hat{\sigma}_{bIN}^2)].$$

Therefore, it only needs to prove that

$$\hat{\sigma}_{\varepsilon IN}^2 - \tilde{\sigma}_\varepsilon^2 = op(1) \tag{A2}$$

and

$$\hat{\sigma}_{bIN}^2 - \tilde{\sigma}_b^2 = op(1). \tag{A3}$$

We first prove (A2) holds.
it is obvious that

$$\begin{aligned}
\hat{\sigma}_{\varepsilon IN}^2 =& \frac{1}{N(T-1)}[\tilde{Y} - X\hat{\beta}_{IN}(u)]'(I-H)[\tilde{Y} - X\hat{\beta}_{IN}(u)] \\
=& \frac{1}{N(T-1)}[\tilde{Y} - X\beta(u)]'(I-H)[\tilde{Y} - X\beta(u)] \\
&+ \frac{2}{N(T-1)}[\tilde{Y} - X\beta(u)]'(I-H)(X\beta(u) - X\hat{\beta}_{IN}(u)) \\
&+ \frac{1}{N(T-1)}(X\beta(u) - X\hat{\beta}_{IN}(u))'(I-H)(X\beta(u) - X\hat{\beta}_{IN}(u)) \\
=& \tilde{\sigma}_\varepsilon^2 + \frac{2}{N(T-1)}[\tilde{Y} - X\beta(u)]'(I-H)(X\beta(u) - X\hat{\beta}_{IN}(u)) \\
&+ \frac{1}{N(T-1)}(X\beta(u) - X\hat{\beta}_{IN}(u))'(I-H)(X\beta(u) - X\hat{\beta}_{IN}(u)).
\end{aligned}$$

Thus, it is necessary to prove the last two terms of above equality converge to 0 in probability. By Lemma 3 and Assumption 1(iv), we have $X\beta(u) - X\hat{\beta}_{IN}(u) = op(1)$. Moreover,

$$
\begin{aligned}
& [\tilde{Y} - X\beta(u)]'(I - H) = (Ub + \varepsilon)'(I - H) \\
= \quad & (\varepsilon_{11} - T^{-1}\sum_{t=1}^{T}\varepsilon_{1t}, \varepsilon_{12} - T^{-1}\sum_{t=1}^{T}\varepsilon_{1t}, \cdots, \varepsilon_{NT} - T^{-1}\sum_{t=1}^{T}\varepsilon_{Nt}).
\end{aligned}
$$

By weak law of large numbers,

$$
\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}|\varepsilon_{it} - T^{-1}\sum_{t=1}^{T}\varepsilon_{it}| \leq \frac{2}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}|\varepsilon_{it}| \leq 2(\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\varepsilon_{it}^{2})^{\frac{1}{2}} \xrightarrow{p} 2\sigma_{\varepsilon}.
$$

Thus

$$
\frac{2}{N(T-1)}[\tilde{Y} - X\beta(u)]'(I - H)[X\beta(u) - X\hat{\beta}_{IN}(u)] = op(1) \tag{A4}
$$

It follows from $X\beta(u) - X\hat{\beta}_{IN}(u) = op(1)$ that

$$
\begin{aligned}
& \frac{1}{N(T-1)}[X\beta(u) - X\hat{\beta}_{IN}(u)]'(I - H)[X\beta(u) - X\hat{\beta}_{IN}(u)] \\
& \leq \frac{1}{N(T-1)}[X\beta(u) - X\hat{\beta}_{IN}(u)]'[X\beta(u) - X\hat{\beta}_{IN}(u)] \\
& = op(1)
\end{aligned} \tag{A5}
$$

By (A4) and (A5), we have

$$
\hat{\sigma}_{\varepsilon IN}^{2} - \tilde{\sigma}_{\varepsilon}^{2} = op(1).
$$

Similar to the proof of (A2), it is not difficult to verify that $\hat{\sigma}_{bIN}^{2} - \tilde{\sigma}_{b}^{2} = op(1)$. According to the continuity of $ln(\cdot)$, we have

$$
\frac{1}{NT}[ln\tilde{L}(\alpha', \rho) - lnL(\alpha', \rho)] = op(1).
$$

Consequently, the consistency of $\hat{\rho}$ and $\hat{\alpha}$ are proved. □

Proof of Theorem 2. It follows by straightforward calculation that

$$
T_u = \frac{1}{NT}D_u'W_u\tilde{Y} = \begin{pmatrix} \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}x_{it}\tilde{y}_{it}k_h(u_{it} - u) \\ \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}x_{it}\tilde{y}_{it}(\frac{u_{it}-u}{h})^{1}k_h(u_{it} - u) \end{pmatrix} \hat{=} \begin{pmatrix} T_{u,0} \\ T_{u,1} \end{pmatrix},
$$

where $T_{u,l} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}x_{it}\tilde{y}_{it}(\frac{u_{it}-u}{h})^{l}k_h(u_{it} - u), l = 0, 1$.

Moreover, $\hat{\beta}(u) = e_0'[D_u'W_uD_u]^{-1}D_u'W_u\hat{\tilde{Y}} = e_0'S_u^{-1}\hat{T}_u$, where $\hat{T}_u = D_u'W_u\hat{\tilde{Y}} = \begin{pmatrix} \hat{T}_{u,0} \\ \hat{T}_{u,1} \end{pmatrix}$ and $\hat{T}_{u,l} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}x_{it}\hat{\tilde{y}}_{it}(\frac{u_{it}-u}{h})^{l}k_h(u_{it} - u), l = 0, 1$. By Theorem 1, it is easy to obtain that

$$
\begin{aligned}
\hat{\tilde{y}}_{it} &= y_{it} - \hat{\rho}(W_0Y_t)_i - z_{it}'\hat{\alpha} \\
&= y_{it} - (\rho + \hat{\rho} - \rho)(W_0Y_t)_i - z_{it}'(\alpha + \hat{\alpha} - \alpha) \\
&= \tilde{y}(1 + op(1)).
\end{aligned}
$$

Thus,

$$
\begin{aligned}
\hat{T}_{u,l} &= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}x_{it}\hat{\tilde{y}}_{it}(\frac{u_{it}-u}{h})^{l}k_h(u_{it} - u) \\
&= (1 + op(1))\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}x_{it}\tilde{y}_{it}(\frac{u_{it}-u}{h})^{l}k_h(u_{it} - u) \\
&= (1 + op(1))T_{u,l}.
\end{aligned}
$$

Therefore, $\hat{\beta}(u) = e_0' S_u^{-1} \hat{T}_u = (1 + op(1)) e_0' S_u^{-1} T_u = (1 + op(1)) \hat{\beta}_{IN}(u)$. Furthermore, we have $\hat{\beta}(u) - \beta_0(u) = \hat{\beta}(u) - \hat{\beta}_{IN}(u) + \hat{\beta}_{IN}(u) - \beta_0(u) = op(1)$ by using Lemma 3. □

Proof of Theorem 3. By Theorem 1, we have

$$\tilde{\sigma}_\varepsilon^2 - \hat{\sigma}_{\varepsilon_{IN}}^2 = op(1), \quad \tilde{\sigma}_b^2 - \hat{\sigma}_{b_{IN}}^2 = op(1).$$

To prove our conclusions, we only need to verify that

$$\tilde{\sigma}_\varepsilon^2 - \sigma_{\varepsilon_0}^2 = op(1), \quad \tilde{\sigma}_b^2 - \sigma_{b_0}^2 = op(1) \tag{A6}$$

and

$$\hat{\sigma}_\varepsilon^2 - \hat{\sigma}_{\varepsilon_{IN}}^2 = op(1), \quad \hat{\sigma}_b^2 - \hat{\sigma}_{b_{IN}}^2 = op(1). \tag{A7}$$

Firstly, we prove (A6) holds. It is easy to know

$$\begin{aligned}
\tilde{\sigma}_\varepsilon^2 &= \frac{1}{N(T-1)}[A(\rho)Y - X\beta(u) - Z\alpha]'(I-H)[A(\rho)Y - X\beta(u) - Z\alpha] \\
&= \frac{1}{N(T-1)}(Ub+\varepsilon)'(I-H)(Ub+\varepsilon) = \frac{1}{N(T-1)}[\sum_{i=1}^{N}\sum_{t=1}^{T}\varepsilon_{it}^2 - \frac{1}{T}\sum_{i=1}^{N}(\sum_{t=1}^{T}\varepsilon_{it}^2)^2] \\
&= \frac{1}{N(T-1)}\sum_{t=1}^{T}\sum_{i=1}^{N}\varepsilon_{it}^2 - \frac{1}{(T-1)}[\frac{1}{NT}\sum_{i=1}^{N}(\sum_{t=1}^{T}\varepsilon_{it})^2].
\end{aligned}$$

According to Khinchine law of large numbers, we have

$$\frac{1}{N(T-1)}\sum_{i=1}^{N}\sum_{t=1}^{T}\varepsilon_{it}^2 \xrightarrow{P} E(\varepsilon_{it}^2) = \sigma_{\varepsilon 0}^2$$

and

$$\frac{1}{NT}\sum_{i=1}^{N}(\sum_{t=1}^{T}\varepsilon_{it})^2 = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\varepsilon_{it}^2 + \frac{2}{NT}\sum_{i=1}^{N}\sum_{j\neq k}^{T}\varepsilon_{ij}\varepsilon_{ik} \xrightarrow{P} E(\varepsilon_{it}^2) = \sigma_{\varepsilon 0}^2.$$

Thus, $\frac{1}{(T-1)}[\frac{1}{NT}\sum_{i=1}^{N}(\sum_{t=1}^{T}\varepsilon_{it})^2] = op(1)$. Therefore, $\tilde{\sigma}_\varepsilon^2 - \sigma_{\varepsilon_0}^2 = op(1)$. By similar way, it is not hard to get $\tilde{\sigma}_b^2 - \hat{\sigma}_{b_0}^2 = op(1)$.

Secondly, we prove (A7). By (13), Theorem 1 and Theorem 2, a simple calculation yields

$$\begin{aligned}
\hat{\sigma}_\varepsilon^2 &= \frac{1}{N(T-1)}[A(\hat{\rho})Y - X\hat{\beta}(u) - Z\hat{\alpha}]'(I-H)[A(\hat{\rho})Y - X\hat{\beta}(u) - Z\hat{\alpha}] \\
&= \frac{1}{N(T-1)}(\hat{\tilde{Y}} - X\hat{\beta}(u))'(I-H)(\hat{\tilde{Y}} - X\hat{\beta}(u)) \\
&= \frac{1}{N(T-1)}(1+op(1))(\tilde{Y} - X\hat{\beta}_{IN}(u))'(I-H)(1+op(1))(\tilde{Y} - X\hat{\beta}_{IN}(u)) \\
&= (1+op(1))\hat{\sigma}_{\varepsilon_{IN}}^2.
\end{aligned}$$

Thus, $\hat{\sigma}_\varepsilon^2 - \hat{\sigma}_{\varepsilon_{IN}}^2 = op(1)$. By a similar way, we can obtain $\hat{\sigma}_b^2 - \hat{\sigma}_{b_{IN}}^2 = op(1)$. This completes the proof. □

Proof of Theorem 4. By Taylor expansion of $\frac{\partial ln\tilde{L}(\theta)}{\partial\theta}|_{\theta=\hat{\theta}} = 0$ at θ_0, we have

$$\frac{\partial ln\tilde{L}(\theta)}{\partial\theta}|_{\theta=\theta_0} + \frac{\partial^2 ln\tilde{L}(\theta)}{\partial\theta\partial\theta'}|_{\theta=\tilde{\theta}}(\hat{\theta} - \theta_0) = 0,$$

where $\tilde{\theta} = (\tilde{\rho}, \tilde{\alpha}', \tilde{\sigma}_b^2, \tilde{\sigma}_\varepsilon^2)'$ and $\tilde{\theta}$ lies between $\hat{\theta}$ and θ_0. By Theorem 1 and Theorem 3, we easily know that $\tilde{\theta}$ converges to θ_0 in probability.

Denote

$$\frac{\partial ln\tilde{L}(\theta_0)}{\partial\theta} \hat{=} \frac{\partial ln\tilde{L}(\theta)}{\partial\theta}|_{\theta=\theta_0}, \quad \frac{\partial^2 ln\tilde{L}(\theta_0)}{\partial\theta\partial\theta'} \hat{=} \frac{\partial^2 ln\tilde{L}(\theta)}{\partial\theta\partial\theta'}|_{\theta=\theta_0}, \quad \frac{\partial^2 ln\tilde{L}(\tilde{\theta})}{\partial\theta\partial\theta'} \hat{=} \frac{\partial^2 ln\tilde{L}(\theta)}{\partial\theta\partial\theta'}|_{\theta=\tilde{\theta}}.$$

Thus,

$$\sqrt{NT}(\hat{\theta}-\theta_0) = -(\frac{1}{NT}\frac{\partial^2 ln\tilde{L}(\tilde{\theta})}{\partial\theta\partial\theta'})^{-1}\frac{1}{\sqrt{NT}}\frac{\partial ln\tilde{L}(\theta_0)}{\partial\theta}.$$

Next, we prove

$$\frac{1}{NT}\frac{\partial^2 ln\tilde{L}(\tilde{\theta})}{\partial\theta\partial\theta'} - \frac{1}{NT}\frac{\partial^2 lnL(\theta_0)}{\partial\theta\partial\theta'} = op(1) \tag{A8}$$

and

$$\frac{1}{\sqrt{NT}}\frac{\partial ln\tilde{L}(\theta_0)}{\partial\theta} \xrightarrow{D} N(0, \Sigma_{\theta_0} + \Omega_\theta). \tag{A9}$$

To prove (A8), we need to show that each element of $\frac{1}{NT}\frac{\partial^2 ln\tilde{L}(\tilde{\theta})}{\partial\theta\partial\theta'} - \frac{1}{NT}\frac{\partial^2 lnL(\theta_0)}{\partial\theta\partial\theta'}$ converges to 0 in probability. It is not difficult to get

$$\begin{aligned}
\frac{1}{NT}\frac{\partial^2 lnL(\theta_0)}{\partial\rho\partial\rho'} =& -\frac{1}{NT}[\frac{1}{\sigma_{\varepsilon0}^2+T\sigma_{b0}^2}(WY)'(I-S)'H(I-S)WY \\
&+\frac{1}{\sigma_{\varepsilon0}^2}(WY)'(I-S)'(I-H)(I-S)WY + tr(G_0^2)],
\end{aligned}$$

$$\frac{1}{NT}\frac{\partial^2 lnL(\theta_0)}{\partial\rho\partial\sigma_\varepsilon^2} = -\frac{1}{NT}Y'W'(I-S)'[\frac{H}{(\sigma_{\varepsilon0}^2+T\sigma_{b0}^2)^2} + \frac{I-H}{\sigma_{\varepsilon0}^4}](I-S)[A(\rho_0)Y - Z\alpha_0],$$

$$\frac{1}{NT}\frac{\partial^2 lnL(\theta_0)}{\partial\rho\partial\sigma_b^2} = -\frac{1}{NT}Y'W'(I-S)'\frac{H}{\sigma_{\varepsilon0}^2+T\sigma_{b0}^2}(I-S)[A(\rho_0)Y - Z\alpha_0],$$

$$\frac{1}{NT}\frac{\partial^2 lnL(\theta_0)}{\partial\rho\partial\alpha} = -\frac{1}{NT}Y'W'(I-S)'(\frac{H}{\sigma_{\varepsilon0}^2+T\sigma_{b0}^2} + \frac{I-H}{\sigma_{\varepsilon0}^2})(I-S)Z,$$

$$\begin{aligned}
\frac{1}{NT}\frac{\partial^2 lnL(\theta_0)}{\partial\sigma_\varepsilon^2\partial\sigma_\varepsilon^2} =& -\frac{1}{NT}[A(\rho_0)Y - Z\alpha]'(I-S)'[\frac{H}{(\sigma_{\varepsilon0}^2+T\sigma_{b0}^2)^3} + \frac{I-H}{\sigma_{\varepsilon0}^6}](I-S)[A(\rho_0)Y - Z\alpha_0] \\
&+\frac{T-1}{T\sigma_{\varepsilon0}^4} + \frac{1}{2T(\sigma_{\varepsilon0}^2+T\sigma_{b0}^2)^2},
\end{aligned}$$

$$\frac{1}{NT}\frac{\partial^2 lnL(\theta_0)}{\partial\sigma_\varepsilon^2\partial\sigma_b^2} = \frac{1}{2(\sigma_{\varepsilon0}^2+T\sigma_{b0}^2)^2} - \frac{1}{2N(\sigma_{\varepsilon0}^2+T\sigma_{b0}^2)^3}[A(\rho_0)Y - Z\alpha_0]'(I-S)'H(I-S)[A(\rho_0)Y - Z\alpha_0],$$

$$\frac{1}{NT}\frac{\partial^2 lnL(\theta_0)}{\partial\sigma_\varepsilon^2\partial\alpha} = -\frac{1}{NT}Z'(I-S)'[\frac{H}{(\sigma_{\varepsilon0}^2+T\sigma_{b0}^2)^2} + \frac{I-H}{\sigma_{\varepsilon0}^4}](I-S)[A(\rho_0)Y - Z\alpha_0],$$

$$\frac{1}{NT}\frac{\partial^2 lnL(\theta_0)}{\partial\sigma_b^2\partial\sigma_b^{2\prime}} = \frac{T}{2(\sigma_{\varepsilon0}^2+T\sigma_{b0}^2)^2} - \frac{T}{N(\sigma_{\varepsilon0}^2+T\sigma_{b0}^2)^3}[A(\rho_0)Y - Z\alpha_0]'(I-S)'H(I-S)[A(\rho_0)Y - Z\alpha_0],$$

$$\frac{1}{NT}\frac{\partial^2 lnL(\theta_0)}{\partial\sigma_b^2\partial\alpha'} = \frac{1}{N(\sigma_{\varepsilon0}^2+T\sigma_{b0}^2)^2}Z'(I-S)'H(I-S)[A(\rho_0)Y - Z\alpha_0],$$

$$\frac{1}{NT}\frac{\partial^2 lnL(\theta_0)}{\partial\alpha\partial\alpha'} = -\frac{1}{NT(\sigma_{\varepsilon0}^2+T\sigma_{b0}^2)}Z'(I-S)'(I-S)Z.$$

By mean value theorem, Assumption 2 (iv) and Lemma 5, we have

$$\begin{aligned}\frac{1}{NT}\Big(\frac{\partial^2 lnL(\tilde{\theta})}{\partial\rho^2}-\frac{\partial^2 lnL(\theta_0)}{\partial\rho^2}\Big) =& \frac{1}{NT}\Big(\frac{1}{\tilde{\sigma}_\varepsilon^2+T\tilde{\sigma}_b^2}-\frac{1}{\sigma_{\varepsilon 0}^2+T\sigma_{b0}^2}\Big)(WY)'(I-S)'H(I-S)WY \\ &+\frac{1}{NT}\Big(\frac{1}{\tilde{\sigma}_\varepsilon^2}-\frac{1}{\tilde{\sigma}_{\varepsilon 0}^2}\Big)(WY)'(I-S)'(I-H)(I-S)WY \\ &+\frac{1}{NT}[tr(G^2(\tilde{\rho}))-tr(G^2(\rho_0))] \\ =&\frac{2}{NT}tr(G^3(\tilde{\rho}^*))(\tilde{\rho}-\rho_0)+op(1)=op(1),\end{aligned}$$

where $\tilde{\rho}^*$ lies between $\tilde{\rho}$ and ρ_0. Similarly, we have

$$\frac{1}{NT}\frac{\partial^2 ln\tilde{L}(\tilde{\theta})}{\partial\theta\partial\theta'}-\frac{1}{NT}\frac{\partial^2 lnL(\theta_0)}{\partial\theta\partial\theta'}=op(1).$$

To prove (A9), we adopt the idea of [12]. It is easy to obtain that the components of $\frac{1}{\sqrt{NT}}\frac{\partial ln\tilde{L}(\theta_0)}{\partial\theta}$ are linear or quadratic functions of $Ub+\varepsilon$ and their means are all $op(1)$. Under Assumption 1, we have that $\frac{1}{\sqrt{NT}}\frac{\partial ln\tilde{L}(\theta_0)}{\partial\theta}$ is asymptotically normal distributed with 0 means by using the CLT for linear-quadratic forms of Theorem 1 in [53]. In the following, we calculate its variance. According to the structure of the Fisher information matrix, we know that

$$Var\Big(\frac{1}{\sqrt{NT}}\frac{\partial ln\tilde{L}(\theta_0)}{\partial\theta}\Big)=E\Big(\frac{1}{\sqrt{NT}}\frac{\partial ln\tilde{L}(\theta_0)}{\partial\theta}\cdot\frac{1}{\sqrt{NT}}\frac{\partial ln\tilde{L}(\theta_0)}{\partial\theta'}\Big)=-E\Big(\frac{1}{NT}\frac{\partial^2 lnL(\theta_0)}{\partial\theta\partial\theta'}\Big)+E(\Omega_{\theta_{NT}}),$$

where

$$\begin{aligned}&\frac{1}{NT}\frac{\partial^2 lnL(\theta_0)}{\partial\theta\partial\theta'}= \\ &\begin{pmatrix} -\{\frac{\sigma_b^2E[Q_0'PQ_0]}{N(\sigma_{\varepsilon 0}^2+T\sigma_{b0}^2)\sigma_{\varepsilon 0}^2}+\frac{\sigma_{b0}^2[tr(G_0'G_0)+tr(G_0'HG_0)]}{N\sigma_{\varepsilon 0}^2}+\frac{tr(G_0'G_0)-tr(G_0^2)}{NT}\} & * & * & * \\ -\frac{tr(G_0'H)}{NT}-\frac{(\sigma_{\varepsilon 0}^2+T\sigma_{b0}^2)[tr(G_0')-tr(G_0'H)]}{NT\sigma_{\varepsilon 0}^4} & -\frac{1}{2\sigma_{\varepsilon 0}^4} & * & * \\ \frac{tr(G_0'H)}{NT(\sigma_{\varepsilon 0}^2+T\sigma_{b0}^2)} & -\frac{1}{2(\sigma_{\varepsilon 0}^2+T\sigma_{b0}^2)^2} & -\frac{T}{2(\sigma_{\varepsilon 0}^2+T\sigma_{b0}^2)^2} & * \\ \frac{\sigma_b^2E[Q_0'PZ]}{N(\sigma_{\varepsilon 0}^2+T\sigma_{b0}^2)\sigma_{\varepsilon 0}^2} & 0 & 0 & -\frac{E[Z'PZ]}{NT(\sigma_{\varepsilon 0}^2+T\sigma_{b0}^2)} \end{pmatrix} \\ &+op(1),\end{aligned}$$

$$\Omega_{\theta_{NT}}=\begin{pmatrix} \Omega_{\theta_{NT}11} & \Omega_{\theta_{NT}12} & \Omega_{\theta_{NT}13} & \Omega_{\theta_{NT}14} \\ * & \Omega_{\theta_{NT}22} & \Omega_{\theta_{NT}23} & \Omega_{\theta_{NT}24} \\ * & * & \Omega_{\theta_{NT}33} & \Omega_{\theta_{NT}34} \\ * & * & * & \Omega_{\theta_{NT}44} \end{pmatrix}.$$

Let μ_{b03} and $\mu_{\varepsilon 03}$ be the third moments of b and ε, respectively; μ_{b04} and $\mu_{\varepsilon 04}$ be the forth moments of b and ε, respectively; and A_{ii} be the ii-th element of A. By using the Lemma 5 and the facts that $E(U_n'A_nU_n)=\sigma_0^2tr(A_n)$, $E(U_n'A_nU_nU_n'B_nU_n)=(\mu_4-3\sigma_0^4)\sum_{i=1}^n a_{n,ii}b_{n,ii}+\sigma_0^4[tr(A_nB_n)+tr(A_nB_n)+tr(A_nB_n')]$, it follows by straightforward calculation that

$$\Omega_{\theta_{NT}11} = \frac{(\mu_{b04} - 3\sigma_{b0}^4)\sum_{i=1}^{NT}(U'HU)_{ii}[U'U - U'HU]_{ii} + (\mu_{\varepsilon 04} - 3\sigma_{\varepsilon 0}^4)\sum_{i=1}^{NT} H_{ii}(I-H)_{ii}}{NT}$$
$$+ \frac{\mu_{b3}Q_0'HUdiag(U'G_0U - U'HG_0U) + \mu_{\varepsilon 03}Q_0'Hdiag(G_0 - HG_0)}{NT^2\sigma_{\varepsilon 0}^2}$$
$$+ \frac{(\mu_{b04} - 3\sigma_{b0}^4)\sum_{i=1}^{NT}(UHU)_{ii}^2 + (\mu_{\varepsilon 04} - 3\sigma_{\varepsilon 0}^4)\sum_{i=1}^{NT}(I-H)_{ii}}{NT^2}$$
$$+ \frac{\mu_{b03}Q_0'HUdiag(U'HG_0U) + \mu_{\varepsilon 03}Q_0'Hdiag(HG_0)}{NT^2(\sigma_{\varepsilon 0}^2 + T\sigma_b^2)},$$

$$\Omega_{\theta_{NT}12} = \frac{(\mu_{b04} - 3\sigma_{b0}^4)\sum_{i=1}^{NT}[(U'G_0U - U'G_0HU)_{ii}[U'U - U'HU]_{ii}]}{2NT\sigma_{\varepsilon 0}^6}$$
$$+ \frac{+(\mu_{\varepsilon 04} - 3\sigma_{\varepsilon 0}^4)\sum_{i=1}^{NT}[(G_0 - G_0H)_{ii}(I-H)_{ii}]}{2NT\sigma_{\varepsilon 0}^6}$$
$$+ \frac{(\mu_{b04} - 3\sigma_{b0}^4)\sum_{i=1}^{NT}[(U'G_0HU)_{ii}(U'U - U'HU)_{ii}] + (\mu_{\varepsilon 04} - 3\sigma_{\varepsilon 0}^4)\sum_{i=1}^{NT}[(G_0'H)_{ii}(I-H)_{ii}]}{2NT\sigma_{\varepsilon 0}^4(\sigma_{\varepsilon 0}^2 + T\sigma_{b0}^2)}$$
$$+ \frac{(\mu_{b04} - 3\sigma_{b0}^4)\sum_{i=1}^{NT}[(U'G_0HU)_{ii}(U'U - U'HU)_{ii}] + (\mu_{\varepsilon 04} - 3\sigma_{\varepsilon 0}^4)\sum_{i=1}^{NT}[(G_0'H)_{ii}(I-H)_{ii}]}{2NT\sigma_{\varepsilon 0}^2(\sigma_{\varepsilon 0}^2 + T\sigma_{b0}^2)^2}$$
$$+ \frac{(\mu_{b04} - 3\sigma_{b0}^4)\sum_{i=1}^{NT}(U'G_0HU)_{ii} + (\mu_{\varepsilon 04} - 3\sigma_{\varepsilon 0}^4)\sum_{i=1}^{NT}[(G_0H)_{ii}H_{ii}]}{2NT(\sigma_{\varepsilon 0}^2 + T\sigma_{b0}^2)^3}$$
$$+ \frac{\mu_{b03}Q_0'HUdiag[U'U - U'HU] + \mu_{\varepsilon 03}Q_0'Hdiag(I-H)}{2NT\sigma_{\varepsilon 0}^4(\sigma_{\varepsilon 0}^2 + T\sigma_{b0}^2)}$$
$$+ \frac{\mu_{b03}Q_0'(I-H)Udiag(U'HU) + \mu_{\varepsilon 03}Q_0'(I-H)diag(H)}{2NT\sigma_{\varepsilon 0}^2(\sigma_{\varepsilon 0}^2 + T\sigma_{b0}^2)^2}$$
$$+ \frac{\mu_{b03}Q_0'HUdiag(U'HU) + \mu_{\varepsilon 03}Q_0'Hdiag(H)}{2NT(\sigma_{\varepsilon 0}^2 + T\sigma_{b0}^2)^3},$$

$$\Omega_{\theta_{NT}13} = \frac{\mu_{b03}Q_0'HUdiag(U'HU) + \mu_{\varepsilon 03}Q_0'Hdiag(H)}{2N(\sigma_{\varepsilon 0}^2 + T\sigma_{b0}^2)^3}$$
$$+ \frac{\mu_{b03}Q_0'(I-H)Udiag(U'HU) + \mu_{\varepsilon 03}Q_0'(I-H)diag(H)}{2N\sigma_{\varepsilon 0}^2(\sigma_{\varepsilon 0}^2 + T\sigma_{b0}^2)^2}$$
$$+ \frac{(\mu_{b04} - 3\sigma_{b0}^4)\sum_{i=1}^{NT}[(U'G_0HU)_{ii}(U'U - U'HU)_{ii}] + (\mu_{\varepsilon 04} - 3\sigma_{\varepsilon 0}^4)\sum_{i=1}^{NT}[(G_0'H)_{ii}(I-H)_{ii}]}{2N\sigma_{\varepsilon 0}^2(\sigma_{\varepsilon 0}^2 + T\sigma_{b0}^2)^2}$$
$$+ \frac{(\mu_{b04} - 3\sigma_{b0}^4)\sum_{i=1}^{NT}[(U'G_0HU)_{ii}(U'HU)_{ii}] + (\mu_{\varepsilon 04} - 3\sigma_{\varepsilon 0}^4)\sum_{i=1}^{NT}[(G_0'H)_{ii}H_{ii}]}{2N(\sigma_{\varepsilon 0}^2 + T\sigma_{b0}^2)^3},$$

$$\Omega_{\theta_{NT}14} = 0,$$

$$\Omega_{\theta_{NT}22} = \frac{(\mu_{b04} - 3\sigma_{b0}^4)\sum_{i=1}^{NT}[(U'HU)_{ii}(U'U - U'HU)_{ii}] + (\mu_{\varepsilon 04} - 3\sigma_{\varepsilon 0}^4)\sum_{i=1}^{NT}[H_{ii}(I-H)_{ii}]}{\sigma_{\varepsilon 0}^4(\sigma_{\varepsilon 0}^2 + T\sigma_{b0}^2)^2}$$
$$+ \frac{(\mu_{b04} - 3\sigma_{b0}^4)\sum_{i=1}^{NT}(U'U - U'HU)_{ii}^2 + (\mu_{\varepsilon 04} - 3\sigma_{\varepsilon 0}^4)\sum_{i=1}^{NT}(I-H)_{ii}^2}{\sigma_{\varepsilon 0}^8}$$
$$+ \frac{(\mu_{b04} - 3\sigma_{b0}^4)\sum_{i=1}^{NT}(UHU)_{ii}^2 + (\mu_{\varepsilon 04} - 3\sigma_{\varepsilon 0}^4)\sum_{i=1}^{NT} H_{ii}^2}{NT(\sigma_{\varepsilon 0}^2 + T\sigma_{b0}^2)^4},$$

$$
\begin{aligned}
\Omega_{\theta_{NT}23} =& \frac{(\mu_{b04}-3\sigma_{b0}^4)\sum_{i=1}^{NT}[(U'HU)_{ii}(U'U-U'HU)_{ii}]+(\mu_{\varepsilon04}-3\sigma_{\varepsilon0}^4)\sum_{i=1}^{NT}[H_{ii}(I-H)_{ii}]}{4N\sigma_{\varepsilon0}^4(\sigma_{\varepsilon0}^2+T\sigma_{b0}^2)^2}\\
&+\frac{(\mu_{b04}-3\sigma_{b0}^4)\sum_{i=1}^{NT}(U'HU)_{ii}^2+(\mu_{\varepsilon04}-3\sigma_{\varepsilon0}^4)\sum_{i=1}^{NT}H_{ii}^2}{4N(\sigma_{\varepsilon0}^2+T\sigma_{b0}^2)^3},\\
\Omega_{\theta_{NT}24} =& \frac{\mu_{b03}Z'H'U'diag(U'HU)+\mu_{\varepsilon03}Z'Hdiag(H)}{2NT(\sigma_{\varepsilon0}^2+T\sigma_{b0}^2)^3}\\
&+\frac{\mu_{b03}Z'(I-H)'U'diag(U'HU)+\mu_{\varepsilon03}Z'(I-H)'diag(H)}{2NT\sigma_{\varepsilon0}^2(\sigma_{\varepsilon0}^2+T\sigma_{b0}^2)^2}\\
&+\frac{\mu_{b03}Z'(I-H)'U'diag(U'U-U'HU)+\mu_{\varepsilon03}Z'(I-H)'diag(I-H)}{2NT^2\sigma_{\varepsilon0}^6}\\
&+\frac{\mu_{b03}Z'H'U'diag(U'U-U'HU)+\mu_{\varepsilon03}Z'H'diag(I-H)}{2NT\sigma_{\varepsilon0}^4(\sigma_{\varepsilon0}^2+T\sigma_{b0}^2)},\\
\Omega_{\theta_{NT}33} =& \frac{T[(\mu_{b04}-3\sigma_{b0}^4)\sum_{i=1}^{NT}(U'HU)_{ii}^2+(\mu_{\varepsilon04}-3\sigma_{\varepsilon0}^4)\sum_{i=1}^{NT}H_{ii}^2]}{4N(\sigma_{\varepsilon0}^2+T\sigma_{b0}^2)^4}\\
\Omega_{\theta_{NT}34} =& \frac{\mu_{b03}Z'(I-H)'U'diag(U'HU)+\mu_{\varepsilon03}Z'(I-H)'diag(H)}{2NT\sigma_{\varepsilon0}^2(\sigma_{\varepsilon0}^2+T\sigma_{b0}^2)}\\
&+\frac{\mu_{b03}Z'H'U'diag(U'HU)+\mu_{\varepsilon03}Z'H'diag(H)}{2N(\sigma_{\varepsilon0}^2+T\sigma_{b0}^2)^2},\\
\Omega_{\theta_{NT}44} =& 0.
\end{aligned}
$$

Particularly, it is not difficult to know that $\Omega_\theta = O(\frac{1}{T}) = o(1)$ when b_i and ε_{it} are normally distributed, respectively. This completes the proof. □

Proof of Theorem 5. According to Lemma 2, we know

$$S_u = \Lambda(u) + op(1),$$

where $\Lambda(u) = F_u diag(\Omega_{11}(u), \mu_2\Omega_{11}(u))$. By using the fact that $(A+B)^{-1} = A^{-1} - A^{-1}(I+BA^{-1})^{-1}BA^{-1}$, we have

$$\hat{\delta} = \begin{pmatrix} \hat{\beta}_{IN}(u) \\ h\dot{\hat{\beta}}_{IN}(u) \end{pmatrix} = S_u^{-1}T_u \xrightarrow{P} \Lambda^{-1}(u)T_u.$$

Denote $R_u = \frac{1}{NT}D_u'W_u(Ub+\varepsilon)$. By using Lemma 1 and Taylor expansion of $\beta_0(u_{it})$ at u, it is not difficult to get

$$T_u - R_u = S_u\begin{pmatrix} \beta_0(u) \\ h\dot{\beta}_0(u) \end{pmatrix} + \frac{1}{2}h^2\begin{pmatrix} \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T k_h(u_{it}-u)x_{it}x_{it}'(\frac{u_{it}-u}{h})^2\ddot{\beta}_0(u) \\ \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T k_h(u_{it}-u)x_{it}x_{it}'(\frac{u_{it}-u}{h})^3\ddot{\beta}_0(u) \end{pmatrix} + o(h^2).$$

According to the proof procedure of Lemma 2, we know

$$E[k_h(u_{it}-u)x_{it}x_{it}'(\frac{u_{it}-u}{h})^2\ddot{\beta}_0(u)] = \mu_2 f(u)\Omega_{11}(u)\ddot{\beta}_0(u).$$

Thus

$$\hat{\beta}_{IN}(u)-\beta_0(u)=e_0'\Lambda^{-1}(u)R_u+e_0'\frac{1}{2}h^2\Lambda^{-1}(u)\begin{pmatrix}\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T k_h(u_{it}-u)x_{it}x_{it}'(\frac{u_{it}-u}{h})^2\ddot{\beta}_0(u)\\ \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T k_h(u_{it}-u)x_{it}x_{it}'(\frac{u_{it}-u}{h})^3\ddot{\beta}_0(u)\end{pmatrix}+o(h^2)$$
$$=\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T k_h(u_{it}-u)x_{it}(b_i+\varepsilon_{it})[F(u)\Omega(u)]^{-1}+\frac{1}{2}h^2\mu_2\ddot{\beta}_0(u)+o(h^2).$$

Define $\varsigma_{it}=\sqrt{NTh}\sum_{i=1}^N\sum_{t=1}^T k_h(u_{it}-u)x_{it}(b_i+\varepsilon_{it})$, then

$$E(\varsigma_{it})=0$$

and

$$E(\varsigma_{it}^2)=hE[k_h^2(u_{it}-u)x_{it}x_{it}'(b_i+\varepsilon_{it})^2]$$
$$=v_0(\sigma_{\varepsilon,0}^2+\sigma_{b,0}^2)F(u)\Omega_{11}(u)+op(1).$$

According to CLT, we have

$$\sqrt{NTh}\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T k_h(u_{it}-u)x_{it}(b_i+\varepsilon_{it})\xrightarrow{D}N(0,v_0(\sigma_{\varepsilon,0}^2+\sigma_{b,0}^2)F(u)\Omega_{11}(u))$$

and

$$\sqrt{NTh}(\hat{\beta}_{IN}(u)-\beta_0(u)-\varphi(u))\xrightarrow{D}N(0,\gamma^2(u)),$$

where $\varphi(u)=\frac{1}{2}h^2\mu_2\ddot{\beta}_0(u)$ and $\gamma^2(u)=v_0(\sigma_{\varepsilon,0}^2+\sigma_{b,0}^2)[F(u)\Omega_{11}(u)]^{-1}$. Furthermore, if $NTh^5\to 0$, then

$$\sqrt{NTh}(\hat{\beta}_{IN}(u)-\beta_0(u))\xrightarrow{D}N(0,\gamma^2(u)).$$

By Lemma 3 and Theorem 2, we obtain that

$$\sqrt{NTh}(\hat{\beta}(u)-\beta_0(u)-\varphi(u))\xrightarrow{D}N(0,\gamma^2(u)).$$

In particular, when $NTh^5\to 0$, $\sqrt{NTh}(\hat{\beta}(u)-\beta_0(u))\xrightarrow{D}N(0,\gamma^2(u))$ holds. □

References

1. Elhorst, J.P. *Spatial Econometrics: From Cross-sectional Data to Spatial Panels*; Springer: Berlin, Germany, 2014.
2. Brueckner, J.K. Strategic interaction among local governments: An overview of empirical studies. *Int. Reg. Sci. Rev.* **2003**, *26*, 175–188.
3. LeSage, J.P.; Pace, R.K. *Introduction to Spatial Econometrics*; CRC Press, Taylor & Francis Group: Boca Raton, MA, USA, 2009.
4. Elhorst, J.P. Applied spatial econometrics: raising the bar. *Spat. Econ. Anal.* **2010**, *5*, 9–28.
5. Allers, M.A.; Elhorst, J.P. Tax mimicking and yardstick competition among governments in the Netherlands. *Int. Tax Public Financ.* **2005**, *12*, 493–513.
6. Cliff, A.D.; Ord, J.K. *Spatial Autocorrelation*; Pion Ltd.: London, UK, 1973.
7. Anselin, L. *Spatial Econometrics: Methods and Models*; Kluwer Academic: Dordrecht, The Netherlands, 1988.
8. Anselin, L.; Bera, A.K. Patial dependence in linear regression models with an introduction to spatial econometrics. In *Handbook of Applied Economics Statistics*; Ullah, A., Giles, D.E.A., Eds.; Marcel Dekker: New York, NY, USA, 1998.
9. Cressie, N. *Statistics for Spatial Data*; John Wiley & Sons: New York, NY, USA, 1993.
10. Ord, J.K. Estimation methods for models of spatial interaction. *J. Am. Stat. Assoc.* **2004**, *70*, 120–126.
11. Anselin, L. Spatial econometrics. In *A Companion to Theoretical Econometrics*; Baltagi, B., Ed.; Blackwell Publishers: Hoboken, NJ, USA, 2001.
12. Lee, L.F. Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica* **2004**, *72*, 1899–1925.
13. Kelejian, H.H.; Prucha, I.R. A generalized spatial two stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbance. *J. Real Estate Financ. Econ.* **1998**, *17*, 99–121.
14. Kelejian, H.H.; Prucha, I.R. A generalized moments estimator for the autoregressive parameter in a spatial model. *Int. Econ. Rev.* **1999**, *40*, 509–533.

15. Lee, L. F. Best spatial two-stage least squares estimators for a spatial autoregressive model with autoregressive disturbances. *Econom. Rev.* **2003**, *22*, 307–335.
16. Lee, L.F. GMM and 2SLS estimation of mixed regressive, spatial autoregressive models. *J. Econom.* **2007**, *137*, 489–514.
17. Baltagi. B. *Econometric Analysis of Panel Data*; John Wiley & Sons: New York, NY, USA, 2008.
18. Arellano, M. *Panel Data Econometrics*; Oxford University Press: Oxford, UK, 2003.
19. Baldev, R.; Baltagi, B. *Panel Data Analysis*; Springer Science and Business Media: Berlin, Germany, 2012.
20. Hsiao, C. *Analysis of Panel Data*; Cambridge University Press: Cambridge, UK, 2014.
21. Baltagi, B.; Song, S.H.; Koh, W. Testing panel data regression models with spatial error correlation. *J. Econom.* **2003**, *117*, 123–150.
22. Baltagi, B.; Egger, P.; Pfaffermayr, M. *A Generalized Spatial Panel Data Model with Random Effects*; Working Paper; Syracuse University: New York, NY, USA, 2007.
23. Elhorst, J.P. Specification and estimation of spatial panel data models. *Int. Reg. Sci. Rev.* **2003**, *26*, 244–268.
24. Kapoor, M.; Kelejian, H.H.; Prucha, I.R. Panel data models with spatially correlated error components. *J. Econom.* **2007**, *140*, 97–130.
25. Yu. J.; Jong, R.D.; Lee, L.F. *Quasi-Maximum Likelihood Estimators for Spatial Dynamic Panel Data with Fixed Effects When Both n and T are Large: A Nonstationary Case*; Working Paper; Ohio State University: Columbus, OH, USA, 2007.
26. Yu, J.; Jong, R.D.; Lee, L.F. Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and T are large. *J. Econom.* **2008**, *146*, 118–134.
27. Paelinck, J.H.P.; Klaassen, L.H. *Spatial Econometrics*; Gower: Farnborough, UK, 1979.
28. Park, H.W.; Kim, H.S. Motion estimation using low-band-shift method for wavelet-based moving-picture coding. *IEEE Trans. Image Process.* **2000**, *9*, 577–587.
29. Gong, Y.; Pullalarevu, S.; Sheikh, S. A wavelet-based lossless video coding scheme. In Proceedings of the 7th International Conference on Signal Processing, ICSP'04, Beijing, China, 31 August–4 September 2004; Volume 2, pp. 1123–1126.
30. Guariglia, E.; Silvestrov, S. *Fractional-Wavelet Analysis of Positive definite Distributions and Wavelets on $\mathscr{D}'(\mathbb{C})$: Engineering Mathematics II*; Springer: Cham, Switzerland, 2016; pp. 337–353.
31. Guariglia, E. Primality, Fractality, and Image Analysis. *Entropy* **2019**, *21*, 304–316.
32. Aoun, N.; El'Arbi, B.M.; Amar, C.B. Multiresolution motion estimation and compensation for video coding. In Proceedings of the IEEE International Conference on Signal Processing, Beijing, China, 3–7 December 2010.
33. Su, L. Semiparametric GMM estimation of spatial autoregressive models. *J. Econom.* **2012**, *167*, 543–560.
34. Su, L.; Jin, S. Profile quasi-maximum likelihood estimation of partially linear spatial autoregressive models. *J. Econom.* **2010**, *157*, 18–33.
35. Wei, H.; Sun, Y. Heteroskedasticity-robust semi-parametric GMM estimation of a spatial model with space-varying coefficients. *Spat. Econ. Anal.* **2017**, *12*, 113–128.
36. Sun, Y.; Zhang, Y.; Huang, J.Z. Estimation of a semiparametric varying-coefficient mixed regressive spatial autoregressive model. *Econom. Stat.* **2019**, *9*, 140–155.
37. Chen, J.; Sun, L. Spatial lag single-index panel model with random effects. *Stat. Res.* **2015**, *32*, 95–101.
38. Li, K.; Chen, J. Profile maximum likelihood estimation of semi-parametric varying coefficient spatial lag model. *J. Quant. Tech. Econ.* **2013**, *4*, 85–98.
39. Du, J.; Sun, X.; Cao, R.; Zhang, Z. Statistical inference for partially linear additive spatial autoregressive models. *Spat. Stat.* **2018**, *25*, 52–67.
40. Hu, X.; Liu, F.; Wang, Z. Testing serial correlation in semiparametric varying coefficient partially linear errors-in-variables model. *J. Syst. Sci. Complex* **2009**, *3*, 483–494.
41. Hardle, W.; Hall, P.; Ichimura, H. Optimal smoothing in single-index models. *Ann. Stat.* **1993**, *21*, 157–178.
42. Opsomer, J.D.; Ruppert, D. Fitting a bivariate additive model by local polynomial regression. *Ann. Stat.* **1997**, *25*, 186–211.
43. Chen, J.; Sun, L. Estimation of varying coefficient spatial autoregression panel model with random effects. *Stat. Res.* **2017**, *34*, 118–128.
44. Zhang, Y.Q.; Shen, D.M. Estimation of semi-parametric varying coefficient spatial panel data models with random-effects. *J. Stat. Plan. Inference* **2015**, *159*, 64–80.
45. Cai, Z. Trending time-varying coefficient time series models with serially correlated errors. *J. Econom.* **2007**, *136*, 163–188.
46. Lin, X.; Carroll, R. Nonparametric function estimation for clustered data when the predictor is measured without (with) error. *J. Am. Stat. Assoc.* **2000**, *95*, 520–534.
47. Fan, J.; Huang, T. Profile likelihood inferences on semiparametric varying coefficient partially linear models. *Bernoulli* **2005**, *11*, 1031–1057.
48. Mack, P.Y.; Silverman, W.B. Weak and strong uniform consistency of kernel regression estimations. *Z. Wahrscheinlichkeitstheorie Verwandte Geb.* **1982**, *61*, 405–415.
49. Ding, R.; Ni, P. Regional spatial linkage and spillover effect of house prices of Chinas cities. *Financ. Trade Econ.* **2015**, *6*, 136–150.
50. Sun, Y.; Wu, Y. Estimation and testing for a partially linear single-index spatial regression model. *Spat. Econ. Anal.* **2018**, *13*, 1–17.

51. Liang, Y.; Gao, T. An empirical analysis on the causes of price fluctuation of commercial housing in China. *Manag. World* **2006**, *8*, 76–82.
52. White, H. *Estimation, Inference and Specification Analysis*; Cambridge University Press: New York, NY, USA, 1994.
53. Kelejian, H.H.; Prucha, I.R. On the asymptotic distribution of the Moran I test statistic with applications. *J. Econom.* **2001**, *104*, 219–257.