mathematics

Article

Estimation of Fixed Effects Partially Linear Varying Coefficient Panel Data Regression Model with Nonseparable Space-Time Filters

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

[1] School of Mathematics and Statistics, Fujian Normal University, Fuzhou 350117, China
[2] School of Mathematics and Statistics, Henan University of Science and Technology, Luoyang 471000, China
* Correspondence: jbchen@fjnu.edu.cn

Abstract: Space-time panel data widely exist in many research fields such as economics, management, geography and environmental science. It is of interest to study the relationship between response variable and regressors which come from above fields by establishing regression models. This paper introduces a new fixed effects partially linear varying coefficient panel data regression model with nonseparable space-time filters. On the basis of approximating the varying coefficient functions with a powerful B-spline method, the profile quasi-maximum likelihood estimators of parameters and varying coefficient functions are constructed. Under some regular conditions, we derive their consistency and asymptotic normality. Monte Carlo simulation shows that our estimates have good finite performance and ignoring spatial and serial correlations may lead to inefficiency of estimates. Finally, the driving forces of Chinese resident consumption rate are studied using our estimation method.

Keywords: partially linear varying coefficient panel data regression model; profile quasi-maximum likelihood estimation; nonseparable space-time filters; asymptotic property; Monte Carlo simulation

MSC: 62F10; 62F12; 62G05; 62G20; 62P20

check for updates

Citation: Li, B.; Chen, J.; Li, S. Estimation of Fixed Effects Partially Linear Varying Coefficient Panel Data Regression Model with Nonseparable Space-Time Filters. *Mathematics* **2023**, *11*, 1531. https://doi.org/10.3390/math11061531

Academic Editor: Stefano Bonnini

Received: 24 February 2023
Revised: 17 March 2023
Accepted: 20 March 2023
Published: 21 March 2023

1. Introduction

A space-time panel dataset is one sample collected from a number of spatial units over time periods (Li et al. [1]). Such datasets widely exist in economics, management, geography, environmental science and other research fields. How to effectively analyze space-time panel datasets and construct space-time panel data regression models has great theoretical and empirical significance. The space-time panel data regression models are a natural extension of panel data regression models. In the early 19th century, “regression” was first mentioned in the works of Legendre and Gauss. Later, at the turn of the 19th and 20th centuries, Galton and Pearson conceptualized regression, there were a number of regression models for analyzing panel data and exploring the association between dependent variable and regressors (Hsiao [2]; Baltagi [3]; Porter et al. [4]; Zamanzade [5]; Imai and Kim [6]). Among them, parametric panel data regression models have been widely used to study linear influence of regressors. Since the 1990s, nonparametric methods have been gradually applied into regression analysis (Fan and Gijbels [7]; Luo et al. [8]; Ullah et al. [9]; Dai et al. [10]), Li and Stengos [11] first proposed nonparametric panel data regression models to explore nonlinear influence of regressors. However, such models have their drawbacks. Parametric panel data regression models need to be precisely pre-specified, misspecified model forms can lead to inconsistent estimates as well as incorrect policy prescriptions. Although nonparametric panel data regression models are useful whenever we are not certain what the correct functional forms are, they may face the “curse of dimensionality” when the dimension of regressors is higher (Fan and Gijbels [7]), namely, the estimation accuracy decreases rapidly with the number of regressors increasing. Therefore, scholars proposed a number of non/semiparametric panel data regression models with a dimension reduction function to more flexibly overcome the "curse of dimensionality" encountered in practice, for example, partially linear additive panel data regression model, partially linear single-index panel data regression model and partially linear varying coefficient panel data regression model. In recent years, a series of their estimation methods have been also developed, including profile least squares estimation (Baltagi et al. [12]; Chen et al. [13]; Huang et al. [14]; Yong et al. [15]; Zhou et al. [16]; Zhang and Shen [17]), profile quasi-maximum likelihood estimation (Li et al. [18]; Su and Ullah [19]; Wu et al. [20]; Hu [21]), generalized method of moment estimation (GMM) (Tran and Tsionas [22]; Su and Ullah [23]), and others (Liu and Zhuang [24]).

All those modeling techniques and corresponding statistical inference methods for the above-mentioned semiparametric panel data regression models need the assumption that there is no correlation among the individuals or time periods. Elhorst [25] pointed out that two problems hampering the modeling of space-time panel data are serial correlation between the observations on each spatial unit over time and spatial correlation between the observations on the spatial units at each point in time. Furthermore, Baltagi et al. [12] mentioned that ignoring the serial correlation in the errors will result in consistent, but inefficient estimates of the regression coefficients and biased standard errors. Therefore, some scholars added nonseparable space-time filters, that is, space-time error correlation are modeled jointly, or separable space-time filters, that is, space-time error correlation are modeled independently from one another, under the framework of semi/parametric panel data regression models. The estimation, testing and empirical analysis of these models have been studied in recent years. Baltagi et al. [26] derived joint and conditional Lagrange Multiplier (LM) and Likelihood Ratio (LR) test statistics of random effects parametric panel regression model with separable space-time correlations and presented their small sample performance using Monte Carlo experiments. Elhorst [25] constructed a random effects parametric panel regression model with nonseparable space-time filters and presented its maximum likelihood estimation. Parent and LeSage [27] explored the Markov Chain Monte Carlo method of random effects parametric panel regression model with separable space-time filters—both Monte Carlo simulation and an application were used to illustrate the method. Lee and Yu [28] investigated quasi-maximum likelihood estimation for fixed effects parametric panel regression model with separable or nonseparable space-time filters, which might be spatially stable or unstable. They also derived consistency and asymptotic normality of the estimators under some regular conditions. Bai et al. [29] proposed a random effects partially linear varying coefficient panel model with separable space-time filters and derived consistency and asymptotic normality of weighted semiparametric least squares estimators. Zhao et al. [30] constructed weighted semiparametric least squares estimators and generalized F-type test statistic for random effects partially linear single-index panel model with separable space-time filters. They also derived the asymptotic properties of estimators and the asymptotic distribution of F-type test statistic. Li et al. [1] studied profile quasi-maximum likelihood estimation and generalized F-type test of random effects partially linear nonparametric panel model with separable space-time filters and obtained the consistency and asymptotic normality of parametric and nonparametric estimators as well as asymptotic distribution of generalized F-type test statistic. Monte Carlo simulation and Indonesian rice farming data were used to illustrate their methods.

To the best of our knowledge, there are no non/semiparametric spatiotemporal econometric models that study both fixed effects and nonseparable space-time filters in the existing literature. In this paper, we attempt to propose a fixed effects partially linear varying coefficient panel data regression model (PLVCPDRM) with nonseparable space-time filters. It can simultaneously capture the linear and nonlinear effects of regressors, spatial and serial correlations of error structure, and individual fixed effects. Our aim is to construct profile quasi-maximum likelihood estimators (PQMLE) of this model and systematically study their asymptotic properties and finite sample performance. Furthermore, the proposed estimation method is illustrated by using a real dataset.

The rest of this paper is organized as follows: Section 2 presents a fixed effects PLVCPDRM with nonseparable space-time filters and its PQMLEs. Section 3 lays out some regular assumptions and asymptotic properties. Section 4 reports simulation results for examining the finite sample performance of the proposed estimators. Section 5 shows the empirical study for illustrating the proposed methodology. Conclusions are summarized in Section 6. Appendix A presents a lemma and proofs of the main theorems.

2. Model and Estimation

Consider a fixed effects PLVCPDRM with nonseparable space-time filters:

$$Y_{Nt} = X_{Nt}\beta + Z_{\alpha,Nt} + b + \varepsilon_{Nt}, \quad t = 1, \ldots, T, \tag{1}$$

$$\varepsilon_{Nt} = \rho W \varepsilon_{Nt} + \lambda \varepsilon_{N,t-1} + e_{Nt}, \tag{2}$$

where $Y_{Nt} = (y_{1t}, y_{2t}, \ldots, y_{Nt})'$, y_{it} are observations of a response variable, $i = 1, \ldots, N$; $X_{Nt} = (x_{1t}, x_{2t}, \ldots, x_{Nt})'$, $Z_{\alpha,Nt} = (z'_{1t}\alpha(u_{1t}), \ldots, z'_{Nt}\alpha(u_{Nt}))'$, $x_{it} = (x_{it1}, x_{it2}, \ldots, x_{itp})'$ and $z_{it} = (z_{it1}, z_{it2}, \ldots, z_{itq})'$ are observations of p-dimensional and q-dimensional exogenous regressors, respectively; β is a regression coefficient vector of x_{it}, $\alpha(u_{it}) = (\alpha_1(u_{it}), \alpha_2(u_{it}), \ldots, \alpha_q(u_{it}))'$ is an unknown univariate varying coefficient function vector, $\alpha_l(u)(l = 1, \ldots, q)$ are smoothing functions of u, u is an intermediate univariate variable; $b = (b_1, \ldots, b_N)'$ are fixed effects satisfying $\sum_{i=1}^N b_i = 0$ for identification purpose; W is an $N \times N$ row-normalized non-negative spatial weights matrix with zero diagonals; ε_{Nt} is an $N \times 1$ vector of disturbance term, e_{Nt} is an $N \times 1$ vector of random error term which is assumed to be $i.i.d.(0, \sigma_e^2)$. In order to keep the stationarity of the model (1)–(2), serial correlation coefficient λ and spatial correlation coefficient ρ should belong to parameter space $\Theta = \{(\lambda, \rho) : \lambda + \rho < 1, \lambda + \rho > -1, \lambda - \rho > -1, \lambda - \rho < 1\}$ (Elhorst [25]; Lee and Yu [28]), see Figure 1.

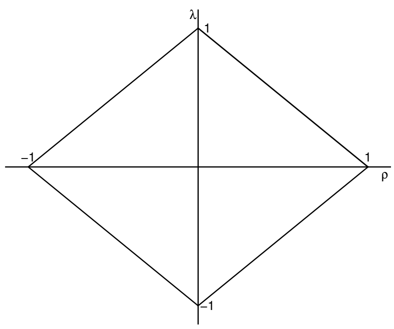

Figure 1. The parameter space Θ of ρ and λ.

For the model (1)–(2), it is necessary to identify an appropriate estimation method to obtain estimators of the unknown parameter vector $\theta = (\beta', \gamma', \rho, \lambda, \sigma_e^2)'$ and varying coefficient functions $\alpha_l(\cdot)(l = 1, \ldots, q)$.

Before proceeding to the estimation procedure, the fitting problem of the varying coefficient functions needs to be solved priority. Polynomial spline method is efficient in function approximation and numerical computation. Polynomial splines are piecewise polynomials with the polynomial pieces joining together smoothly at a set of interior knot points (see De Boor [31]; Huang and Shen [32]; Zou and Zhu [33]). B-spline is a special form of polynomial spline. Considering that the B-spline basis has better numerical properties than other basis functions, we use the B-spline method to approximate the varying coefficient functions $\alpha_l(u)(l = 1, \ldots, q)$ in the model (1). To be precise, let $a = \min\{u_{11}, \ldots, u_{NT}\}$, $d = \max\{u_{11}, \ldots, u_{NT}\}$ and $a = \xi_0 < \xi_1 < \cdots < \xi_{k_l} = d(l = 1, \ldots, q)$ be a partition of interval $[a, d]$. Using the ξ_i as knots, we have $\kappa_l = k_l + k_0$ normalized B-spline basis function of order $(k_0 - 1)$ that forms a basis function for the linear spline space $S_{k_l}^{k_0}$ on $U = \{u_{it} \in \mathbb{R}\}$. Denote B-spline basis function $\zeta_l^{\kappa_l}(u) = (\zeta_{l1}(u), \ldots, \zeta_{l\kappa_l}(u))'$, we can approximate $\alpha_l(u)$ by some spline function in $S_{k_l}^{k_0}$: $\alpha_l(u) \approx \zeta_l^{\kappa_l \prime}(u)\gamma_l$, where $\gamma_l = (\gamma_{l1}, \ldots, \gamma_{l\kappa_l})'$ is an unknown $\kappa_l \times 1$ spline coefficient vector. Thus, the model (1) can be written as

$$Y_{Nt} = X_{Nt}\beta + \widetilde{Z}_{Nt}\gamma + b + \varepsilon_{Nt}, \tag{3}$$

where $\gamma = (\gamma_1', \ldots, \gamma_q')'$, $\widetilde{Z}_{Nt} = (\widetilde{z}_{1t}, \ldots, \widetilde{z}_{Nt})'$, $\widetilde{z}_{it}' = z_{it}'\zeta_{q,K}(u_{it})$ and

$$\zeta_{q,K}(u) = \begin{pmatrix} \zeta_{11}(u) & \ldots & \zeta_{1\kappa_1}(u) & 0 & \ldots & 0 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & 0 & 0 & \ldots & 0 & \zeta_{q1}(u) & \ldots & \zeta_{q\kappa_q}(u) \end{pmatrix}$$

is a $q \times K$ matrix, $K = \sum_{s=1}^{q} \kappa_s$.

For any $N \times 1$ vector $\mathbb{A}_{Nt}$, denote $\Delta\mathbb{A}_{Nt} = \mathbb{A}_{Nt} - \mathbb{A}_{N,t-1}$ as the first order difference. By first difference of the model (2)–(3) to eliminate the fixed effects, we have

$$\Delta Y_{Nt} = \Delta X_{Nt}\beta + \Delta\widetilde{Z}_{Nt}\gamma + \Delta\varepsilon_{Nt}, \quad t = 2, 3, \ldots, T, \tag{4}$$

$$\Delta\varepsilon_{Nt} = \rho W \Delta\varepsilon_{Nt} + \lambda\Delta\varepsilon_{N,t-1} + \Delta e_{Nt}. \tag{5}$$

Note that $\Delta Y_{Nt} = Y_{Nt} - Y_{N,t-1}$ is observable for $t = 2, 3, \ldots, T$, $\Delta\varepsilon_{N1}$ can't be observed. Let $\eta = (\rho, \lambda)'$, $S_N(\rho) = I_N - \rho W$, $R_N(\lambda) = \lambda I_N$, $S_N = S(\rho_0)$, $R_N = R(\lambda_0)$ and I_N is an $N \times N$ identity matrix. The Equation (5) can be rewritten as $S_N(\rho)\Delta\varepsilon_{Nt} = R_N(\lambda)\Delta\varepsilon_{N,t-1} + \Delta e_{Nt}$ for all t. With backward substitution, we have $S_N(\rho)\Delta\varepsilon_{N,2} = \sum_{j=0}^{\infty} A_N^j(\eta)\Delta e_{N,2-j}$, where $A_N(\eta) = R_N(\lambda)S_N(\rho)^{-1}$. By denoting $\Delta\varepsilon_{N,T-1} = \left(\Delta\varepsilon_{N2}', \ldots, \Delta\varepsilon_{NT}'\right)'$ and

$$B_{N,T-1}(\eta) = \begin{pmatrix} S_N(\rho) & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} \\ -R_N(\lambda) & S_N(\rho) & \ldots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -R_N(\lambda) & \ldots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & -R_N(\lambda) & S_N(\rho) \end{pmatrix},$$

the matrix form of the Equation (5) can be simply expressed as $B_{N,T-1}(\eta)\Delta\varepsilon_{N,T-1} = ((S_N(\rho)\Delta\varepsilon_{N2})', \Delta e_{N3}', \ldots, \Delta e_{NT}')'$. As $\mathrm{Var}\left[\sum_{j=0}^{\infty} A_N^j(\eta)\Delta e_{N,2-j}\right] = \sigma_e^2 K_N(\eta)$, where

$$K_N(\eta) \equiv I_N + \sum_{j=0}^{\infty} A^j(\eta)(A_N(\eta) - I_N)(A_N(\eta) - I_N)' A_N'^{j}(\eta),$$

and $K_N = K_N(\eta_0)$, we can obtain $\mathrm{Var}(B_{N,T-1}(\eta)\Delta\varepsilon_{N,T-1}) = \sigma_e^2\Omega_{N,T-1}(\eta)$ with

$$\Omega_{N,T-1}(\eta) = \begin{pmatrix} K_N(\eta) & -I_N & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} \\ -I_N & 2I_N & -I_N & \ldots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -I_N & 2I_N & \ldots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & 2I_N & -I_N \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & -I_N & 2I_N \end{pmatrix}.$$

Note that the only unknown element of $\Omega_{N,T-1}(\eta)$ is $K_N(\eta)$. In order to obtain determinant and inverse of $\Omega_{N,T-1}(\eta)$, we define a confirmable block matrix (Hsiao et al [34]; Lee and Yu [28]) as

$$P_{N,T-1}(\eta)=\begin{pmatrix} I_N & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ I_N & K_N(\eta) & \mathbf{0} & \dots & \mathbf{0} \\ I_N & K_N(\eta) & 2K_N(\eta)-I_N & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ I_N & K_N(\eta) & 2K_N(\eta)-I_N & \dots & (T-2)K_N(\eta)-(T-3)I_N \end{pmatrix}.$$

From straight calculation, we know that

$$\begin{aligned} D_{N,T-1}(\eta) &\equiv P_{N,T-1}(\eta)\Omega_{N,T-1}(\eta)P'_{N,T-1}(\eta) \\ &= \mathrm{diag}\{K_N(\eta),(2K_N(\eta)-I_N)K_N(\eta),(3K_N(\eta)-2I_N)(2K_N(\eta)-I_N),\dots, \\ &[(T-1)K_N(\eta)-(T-2)I_N][(T-2)K_N(\eta)-(T-3)I_N]\}. \end{aligned}$$

Thus, the determinant $|\Omega_{N,T-1}(\eta)| = |D_{N,T-1}(\eta)|/|P_{N,T-1}(\eta)|^2 = |(T-1)K_N(\eta)-(T-2)I_N|$ and the inverse $\Omega^{-1}_{N,T-1}(\eta) = P_{N,T-1}(\eta)'D_{N,T-1}(\eta)^{-1}P_{N,T-1}(\eta)$. Therefore, the quasi-log-likelihood function can be written as

$$\begin{aligned} \log L_{N,T}(\theta) = &-\frac{N(T-1)}{2}\log\left(2\pi\sigma_e^2\right) - \frac{1}{2}\log|I_N+(T-1)(K_N(\eta)-I_N)| \\ &+(T-1)\log|S_N(\rho)| - \frac{1}{2\sigma_e^2}\left[Y-X\beta-\widetilde{Z}\gamma\right]'\mathbb{J}_{NT}(\eta)\left[Y-X\beta-\widetilde{Z}\gamma\right], \end{aligned} \tag{6}$$

where $Y=(Y'_{N1},\dots,Y'_{NT})'$, $X=(X'_{N1},\dots,X'_{NT})'$, $\widetilde{Z}=(\widetilde{z}_{11},\dots,\widetilde{z}_{NT})'$, $\mathbb{J}_{NT}(\eta)=L'_{N,(T-1)T}B'_{N,T-1}(\eta)\Omega^{-1}_{N,T-1}(\eta)B_{N,T-1}(\eta)L_{N,(T-1)T}$, $L_{N,(T-1)T}=L\otimes I_N$ with

$$L=\begin{pmatrix} -1 & 1 & 0 & \dots & 0 \\ 0 & -1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & -1 & 1 \end{pmatrix}$$

as the first order difference transformation matrix of dimension $(T-1)\times T$.

Motivated by Su and Jin [35], we obtain PQMLEs of parameter vector θ and varying coefficient functions $\alpha_l(\cdot)(l=1,\dots,q)$ by the following the two-step estimation procedure:

Step 1: Assuming the parameter η is known, the initial estimators of $(\beta',\gamma',\sigma_e^2)'$ can be obtained by maximizing quasi-log-likelihood function (6):

$$\hat{\beta}_{IN} = \left[X'\mathbb{J}_{NT}(\eta)X\right]^{-1}X'\mathbb{J}_{NT}(\eta)\left[Y-\widetilde{Z}\hat{\gamma}_{IN}\right],$$

$$\hat{\gamma}_{IN} = \left[\widetilde{Z}'\mathbb{J}_{NT}(\eta)\widetilde{Z}\right]^{-1}\widetilde{Z}'\mathbb{J}_{NT}(\eta)\left[Y-X\hat{\beta}_{IN}\right],$$

$$\hat{\sigma}^2_{eIN} = \frac{1}{N(T-1)}\left[Y-X\hat{\beta}_{IN}-\widetilde{Z}\hat{\gamma}_{IN}\right]'\mathbb{J}_{NT}(\eta)\left[Y-X\hat{\beta}_{IN}-\widetilde{Z}\hat{\gamma}_{IN}\right].$$

Step 2: With the estimated $\hat{\beta}_{IN}$, $\hat{\gamma}_{IN}$ and $\hat{\sigma}^2_{eIN}$, PQMLE of η can be obtained by maximizing the concentrated quasi-log-likelihood function of η:

$$\begin{aligned} \log L_{N,T}(\eta) = &-\frac{N(T-1)}{2}\log(2\pi) - \frac{N(T-1)}{2}(\log\hat{\sigma}^2_{eIN}+1) \\ &-\frac{1}{2}\log|I_N+(T-1)(K_N(\eta)-I_N)| + (T-1)\log|S_N(\rho)|. \end{aligned}$$

The final estimator of η is given by $\hat{\eta} = \arg\max_\eta \log L_{N,T}(\eta)$. With the estimated $\hat{\eta}$, update

$\hat{\beta}'_{IN}$, $\hat{\gamma}'_{IN}$ and $\hat{\sigma}^2_{eIN}$, we can obtain the final PQMLEs as

$$\hat{\beta} = \left[X'(I-\mathbb{S}_{\hat{\eta}})'\mathbb{J}_{NT}(\hat{\eta})(I-\mathbb{S}_{\hat{\eta}})X\right]^{-1}X'(I-\mathbb{S}_{\hat{\eta}})'\mathbb{J}_{NT}(\hat{\eta})(I-\mathbb{S}_{\hat{\eta}})Y,$$

$$\hat{\gamma} = \left[\widetilde{Z}'\mathbb{J}_{NT}(\hat{\eta})\widetilde{Z}\right]^{-1}\widetilde{Z}'\mathbb{J}_{NT}(\hat{\eta})\left[Y - X\hat{\beta}\right], \tag{7}$$

$$\hat{\sigma}^2_e = \frac{1}{N(T-1)}\left[Y - X\hat{\beta}\right]'(I-\mathbb{S}_{\hat{\eta}})'\mathbb{J}_{NT}(\hat{\eta})(I-\mathbb{S}_{\hat{\eta}})\left[Y - X\hat{\beta}\right],$$

where I is an identity matrix of dimension NT, $\mathbb{S}_\eta = \widetilde{Z}\left[\widetilde{Z}'\mathbb{J}_{NT}(\eta)\widetilde{Z}\right]^{-1}\widetilde{Z}'\mathbb{J}_{NT}(\eta)$. Then, the estimator of the nonparametric function $\alpha(u)$ can be written as

$$\hat{\alpha}(u) = \zeta_{q,K}(u)\hat{\gamma}. \tag{8}$$

3. Asymptotic Properties

To derive the asymptotic properties of the estimators, we first introduce some regular assumptions. For clear exposition, denote $\theta_0 = (\beta'_0, \gamma'_0, \eta'_0, \sigma^2_{e0})'$, $\theta^*_0 = (\beta'_0, \eta'_0, \sigma^2_{e0})'$ and $\eta_0 = (\rho_0, \lambda_0)'$ as the true parameter vector of θ, θ^* and η, respectively, and $\alpha_0(u)$ as the true varying coefficient function vector of $\alpha(u)$.

Assumption 1. *(i) The sequences $\{x_{it}\}_{i=1,t=1}^{N,T}$, $\{z_{it}\}_{i=1,t=1}^{N,T}$ and $\{u_{it}\}_{i=1,t=1}^{N,T}$ are nonstochastic, and they have bounded support set on $\mathbb{R}^p$, $\mathbb{R}^q$ and $\mathbb{R}^1$ respectively. In addition, u_{it} forms a sequence of designs such that they are analogous to a positive and bounded "design density" $f_U(u)$ (Su and Jin [35]).*

(ii) For any bounded continuous function $h(\cdot)$, it holds that

$$\lim_{N\to\infty}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}h(u_{it}) = \int_U h(u)f_U(u)du. \tag{9}$$

(iii) The parameter $\beta \in \mathbb{R}^p$ in a neighborhood of β_0 satisfies $|x'_{it}\beta| \le m_x$, where m_x is a positive constant.

Assumption 2. *The disturbances $\{e_{it}\}_{i=1,t=2}^{N,T}$ are i.i.d. with zero mean, variance σ^2_{e0} and $E|e_{it}|^{4+\epsilon} < \infty$ for some $\epsilon > 0$.*

Assumption 3. *(i) For every K, the smallest eigenvalue of $\widetilde{Z}'\widetilde{Z}/NT$ and $\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}/NT$ are bounded away from zero uniformly in K.*

(ii) There is a sequence of constants $\zeta_0(K)$ satisfying $\sup_{u\in U}\|\zeta_{q,K}(u)\| \le \zeta_0(K)$ such that $\zeta_0^2(K)K/N \to 0$ as $N \to \infty$.

(iii) For any r_1-th$(r_1 \ge 2)$ continuously differentiable bounded function $\alpha(\cdot)$ satisfying the normalization of $\alpha_0(\cdot)$, there exist some $r_2 > 0$ such that $\sup_{u\in U}|z'_{it}\alpha(u_{it}) - \tilde{z}'_{it}\gamma| = O(K^{-r_2})$ as $K \to \infty$ and $\sqrt{N}K^{-r_2} \to 0$ as $N \to \infty$.

Assumption 4. *(i) W is a row-normalized and prespecified spatial weights matrix.*

(ii) Row and column sums of W in absolute value are uniformly bounded (i.e., UB).

(iii) $S_N(\rho)$ is invertible for all $\rho \in \mathbb{P}$, where $\mathbb{P}$ is compact and the true parameter ρ_0 is in the interior of $\mathbb{P}$. Additionally, $S_N^{-1}(\rho)$ is UB for $\rho \in \mathbb{P}$.

Assumption 5. *(i) $\sum_{h=1}^{\infty} abs(A_N^h)$ is UB, where $[abs(A_N)]_{ij} = |A_{N,ij}|$ and $A_N = A_N(\eta_0)$.*

(ii) $\mathbb{J}_{NT}(\eta)$ is UB.

(iii) The limit of the information matrix (A4) in Appendix A is nonsingular.

(iv) $\lim_{N\to\infty}\frac{1}{N(T-1)}[X, \widetilde{Z}]'\mathbb{J}_{NT}(\eta)[X, \widetilde{Z}]$ is nonsingular.

Assumption 6. *$\lim_{N\to\infty} T_{NT,1}(\eta,\sigma_e^2) \neq 0$ for $(\eta,\sigma_e^2)' \neq (\eta_0,\sigma_{e0}^2)'$, where $T_{NT,1}$ is defined in (A1).*

Remark 1. *The fixed bounded design in Assumption 1 is typically assumed in spatial econometric literature, see Kelejian and Prucha [36], Kelejian and Prucha [37], Su and Jin [35] and Cheng and Chen [38]. Assumption 1 (ii) parallels Assumption 1 of Su and Jin [35] and Assumption 2.1 (iv) of Hu et al. [39]. It means that if $\{u_{it}\}_{i=1,t=1}^{N,T}$ are i.i.d. with the density $f_U(\cdot)$, the Equation (9) holds with probability 1. Assumption 2 presents regularity assumptions for error terms e_{it}. Assumption 3 is a set of mild conditions on the B-spline method (see Newey [40]; Hu et al. [39]; Yong et al. [15]; Zhang [41]). Assumption 3(i) ensures that $\widetilde{Z}'\widetilde{Z}$ and $\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}$ are asymptotically nonsingular, which parallels Assumption 3 of Zhang [41] and Assumption 2(i) of Newey [40]. Newey [40] gave some primitive conditions for power series and splines such that Assumption 3(ii)–(iii) hold. In addition, Assumption 3(iii) is the counterpart assumption in the kernel method. Assumption 4 provides the basic features of the spatial weight matrix. The uniform boundedness of W and $S_N^{-1}(\rho)$ limits the spatial correlation to a manageable degree in Assumptions 4(ii)–(iii). Assumption 5(i) is the absolute summability condition and row/column sum boundedness condition for disturbances, which will play an important role for the proofs of asymptotic properties. To prove the absolute summability of A_N, a sufficient condition is $\|A_N\| < 1$ for any matrix norm (see Corollary 5.6.16 in Horn and Johnson [42]) that satisfies $\|A_N\| = \|abs(A_N)\|$. When $\|A_N\| < 1$, $\sum_{h=0}^{\infty} A_N^h$ exists and can be defined as $(I_N - A_N)^{-1}$. Under the condition that the inverse of the variance matrix of $(1-\phi)^{1/2}e_{Nt} + (A_N - I_N)(e_{N,t-1} + A_N e_{N,t-2} + A_N^2 e_{N,t-3} + \dots)$ is UB for $\phi = 0, 1$ and $\frac{T-2}{T-1}$, Assumption 5(ii) can be certified. Assumption 5(iii)–(iv) is used for establishing the uniqueness identification and asymptotic normality of the proposed estimators. Assumption 6 specifies an identification condition for the estimators of parameters when Assumption 5(iv) is not satisfied.*

In order to prove consistency of the parametric estimators, we need to obtain the expected value function for the quasi-log-likelihood function (6) divided by the effective sample size $N(T-1)$. The relationship $B_{NT}^*\varepsilon = e$ between $\varepsilon = (\varepsilon'_{N1}, \dots, \varepsilon'_{NT})'$ and $e = (e_{N1}^{\dagger\prime}, \dots, e'_{NT})'$ (the first block of N in e are not exactly the original e_{N1} and all the entries are i.i.d. under normality) would be used frequently, where

$$B_{NT}^* = \begin{pmatrix} Q_N^* & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ -R_N & S_N & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -R_N & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & -R_N & S_N \end{pmatrix}_{NT\times NT}$$

and $Q_N^* = \left(\sum_{j=0}^{\infty} A_N^j A_N^{j\prime}\right)^{-1/2} S_N$. Thus, $e_{N1}^{\dagger} = Q_N^* S_N^{-1} \sum_{j=0}^{\infty} A_N^j e_{N,1-j}$ and $Q_N^* \operatorname{Var}(\varepsilon_{N1}) Q_N^{*\prime} = \sigma_{e0}^2 I_N$ under the normality of disturbances. Split B_{NT}^* into four block matrices, one of which is Q_N^*. Utilizing the formula $\begin{pmatrix} A & \mathbf{0} \\ B & C \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} & \mathbf{0} \\ -C^{-1}BA^{-1} & C^{-1} \end{pmatrix}$ for inversion of a block matrix, we have that

$$B_{NT}^{*-1} = \begin{pmatrix} Q_N^{*-1} & & & & & & \\ A_N Q_N^{*-1} & S_N^{-1} & & & & & \\ A_N^2 Q_N^{*-1} & A_N S_N^{-1} & S_N^{-1} & & & & \\ \vdots & \vdots & \vdots & \ddots & & & \\ \vdots & \vdots & \vdots & \ddots & \ddots & & \\ A_N^{T-1} Q_N^{*-1} & A_N^{T-2} S_N^{-1} & A_N^{T-3} S_N^{-1} & \cdots & \cdots & A_N S_N^{-1} & S_N^{-1} \end{pmatrix}.$$

Define $Q_{N,T}(\theta) = \mathrm{E}(\log L_{N,T}(\theta)/N(T-1))$, then

$$\begin{aligned} Q_{N,T}(\theta) = &-\frac{1}{2}\log\left(2\pi\sigma_e^2\right) + \frac{1}{N}\log|S(\rho)| - \frac{1}{2N(T-1)}\log|I_N + (T-1)(K_N - I_N)| \\ &- \frac{1}{2\sigma_e^2 N(T-1)} e' B_{NT}^{*-1\prime} \mathbb{J}_{NT}(\eta) B_{NT}^{*-1} e. \end{aligned} \tag{10}$$

However, when $e^{\dagger} = (e_{N1}^{\dagger\prime}, \ldots, e_{NT}')'$ are not normally distributed, elements in $e_{N1}^{\dagger}$ are uncorrelated but not necessarily independent of each other even though they are independent with $e_{Nt}(t = 2, \ldots, T)$. Consider the case that the process starts at a finite past period, such as $t = -m$. Denote $e_{N,T+m} = (e_{N,1-m}', e_{N,1-(m-1)}', \ldots, e_{N0}', e_{N1}', \ldots, e_{NT}')'$, which includes the original *i.i.d.* disturbances vectors, we have $e^{\dagger} = F_{NT,N(T+m)} e_{N,T+m}$, where

$$F_{NT,N(T+m)} = \begin{pmatrix} \left(\sum_{j=0}^{m} A_N^j A_N^{j\prime}\right)^{-\frac{1}{2}} \cdot A_N^m & \ldots & \left(\sum_{j=0}^{m} A_N^j A_N^{j\prime}\right)^{-\frac{1}{2}} \cdot A_N & \left(\sum_{j=0}^{m} A_N^j A_N^{j\prime}\right)^{-\frac{1}{2}} \cdot I_N & & & \\ & & & & I_N & & \\ & & & & & \ddots & \\ & & & & & & I_N \end{pmatrix}$$

is UB. Under non-normality, we can obtain

$$\begin{aligned} Q_{N,T}(\theta) = &-\frac{1}{2}\log\left(2\pi\sigma_e^2\right) + \frac{1}{N}\log|S(\rho)| - \frac{1}{2N(T-1)}\log|I_N + (T-1)(K_N - I_N)| \\ &- \frac{1}{2\sigma_e^2 N(T-1)} e_{N,T+m}' F_{NT,N(T+m)}' B_{NT}^{*-1\prime} \mathbb{J}_{NT}(\eta) B_{NT}^{*-1} F_{NT,N(T+m)} e_{N,T+m}. \end{aligned} \tag{11}$$

To show the consistency of $\hat{\theta}$, we follow Lee [43] by identifying θ_0 based upon the maximum value of $Q_{N,T}(\theta)$ and showing the uniform convergence of $\frac{1}{N(T-1)}\log L_{N,T}(\theta) - Q_{N,T}(\theta)$ to zero, consistency of $\hat{\theta}$ follows.

Theorem 1. *Suppose Assumptions 1–6 hold, θ_0 is globally identifiable and $\hat{\theta}$ is consistent with θ_0.*

Theorem 2. *Suppose Assumptions 1–6 hold, as $N \to \infty$ simultaneously, we have*

$$\sqrt{NT}(\hat{\theta}^* - \theta_0^*) \xrightarrow{L} \mathrm{N}\left(0, \Sigma_{\theta_0^*}^{-1} + \Sigma_{\theta_0^*}^{-1}\Omega_{\theta_0^*}\Sigma_{\theta_0^*}^{-1}\right).$$

where "$\xrightarrow{L}$" means convergence in distribution, $\Sigma_{\theta_0^*} = -\lim_{N,T\to\infty} E\left(\frac{1}{N(T-1)}\frac{\partial^2 \log L_{N,T}(\theta_0^*)}{\partial\theta^*\partial\theta^{*\prime}}\right)$ is an expected Hessian matrix showed in (A5) and $E\left(\frac{1}{N(T-1)}\frac{\partial \log L_{N,T}(\theta_0^*)}{\partial\theta^*}\frac{\partial \log L_{N,T}(\theta_0^*)}{\partial\theta^{*\prime}}\right) = \Sigma_{\theta_0^*} + \Omega_{\theta_0^*} + o_P(1)$ with $\Omega_{\theta_0^*}$ defined in (A6).

Theorem 3. *Suppose Assumptions 1–5 hold, we have*

$$|\hat{\alpha}(u_{it}) - \hat{\alpha}_0(u_{it})| = O_P\left(\zeta_0(K)(\sqrt{K}/\sqrt{N} + K^{-r_2})\right).$$

Remark 2. *The term K/N essentially corresponds to a variance term and K^{-2r_2} to a bias term. When K is chosen as $N^{\frac{1}{1+2r_2}}$ so that these two terms go to zero at the same rate, which occurs when K goes to infinity at the same rate as $N^{\frac{1}{1+2r_2}}$ (and the side condition $\zeta_0(K)^2K/N \to 0$ is satisfied), the convergence rate will be $N^{-\frac{r_2}{1+2r_2}}$.*

Theorem 4. *Suppose Assumptions 1–5 hold, as $N \to \infty$ simultaneously, we have*

$$\Lambda_u^{-1/2}(\hat{\alpha}(u) - \alpha^*(u)) \xrightarrow{L} \mathrm{N}(0, \sigma_{e0}^2 I_K),$$

where $\Lambda_u = \zeta_{q,K}(u)[\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}]^{-1}\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)B_{NT}^{*-1}B_{NT}^{*\prime-1}\ \mathbb{J}_{NT}'(\eta_0)\widetilde{Z}[\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}]^{-1}\zeta_{q,K}'(u)$, $\alpha^*(u) = \zeta_{q,K}(u)\gamma_0$, *$I_K$ is an identity matrix of dimension K.*

4. Simulation Studies

In this section, we report the results of Monte Carlo simulation experiments to examine the finite sample performance of the proposed estimation method. In order to illustrate the estimation accuracy of parameters, we use the sample mean (Mean), the sample standard deviation (SD) and the root mean square error (RMSE) as the evaluation criteria. Here,

$$\mathrm{RMSE} = \left(\frac{1}{mcn}\sum_{i=1}^{mcn}(\hat{\theta}_i - \theta_0)^2\right)^{\frac{1}{2}},$$

where mcn is the simulation times, $\hat{\theta}_i(i = 1, 2, \ldots, mcn)$ are the parametric estimates of each simulation and θ_0 is the true value. For the nonparametric estimates, we consider the mean absolute deviation error (MADE) as the evaluation criterion which is defined as

$$\mathrm{MADE}_j = Q^{-1}\sum_{q=1}^{Q}|\hat{g}_j(u_q) - g_j(u_q)|, \quad j = 1, 2, \ldots, mcn,$$

where $\{u_q\}_{q=1}^{Q}$ are Q fixed grid points at support set of u.

Example 1. *The first example is to evaluate the performance of the estimation procedure. Consider the following data-generated processes:*

$$\begin{cases} y_{it} = x_{it1}\beta_1 + x_{it2}\beta_2 + z_{it1}\alpha_1(u_{it}) + z_{it2}\alpha_2(u_{it}) + b_i + \varepsilon_{it}, \\ \varepsilon_{it} = \rho\sum_{j=1}^{N} w_{ij}\varepsilon_{jt} + \lambda\varepsilon_{i,t-1} + e_{it}, \end{cases} \tag{12}$$

where $x_{itp} \sim U[-2, 2](p = 1, 2)$, $z_{itq} \sim U[-2, 2](q = 1, 2)$, $u_{it} \sim U[-3, 3]$, $b_i \sim i.i.d.N(0, 1)$, $e_{it} \sim i.i.d.N(0, 0.5)$, the link functions $\alpha_1(u_{it}) = 0.5u_{it} + sin(1.5u_{it})$ and $\alpha_2(u_{it}) = u_{it}^2 + 0.5u_{it}$, $(\beta_1, \beta_2) = (1, 1.5)$. As in Su [44], the spatial weighting matrix is set to the Rook weight matrix. Sample size is $T = 10, 15, 20$ and $N = 25, 49, 81$. For each case, we ran 500 simulations. The R software was used.

Table 1 summarizes Means, Medians, SDs and RMSEs for parametric estimates of $\hat{\beta}_1, \hat{\beta}_2, \hat{\rho}, \hat{\lambda}$ and $\hat{\sigma}_e^2$ when the true values of spatial correlation coefficient and serial correlation coefficient are set as $(\rho, \lambda) = (0.4, 0.4), (0.2, 0.7)$ and $(0.7, 0.2)$, respectively. Tables 2 and 3 give the median and SD of MADE values of $\hat{\alpha}_1(u)$ and $\hat{\alpha}_2(u)$ at 20 fixed grid points in all cases, respectively. We have the following finds: (1) The estimates of $\beta_1, \beta_2, \rho, \lambda, \sigma_e^2$ are close to true values for all cases; (2) SDs and RMSEs for $\hat{\beta}_1, \hat{\beta}_2, \hat{\rho}, \hat{\lambda}, \hat{\sigma}_e^2$ are fairly small for all cases; (3) For fixed T(or N), as N(or T) increased, the SDs and RMSEs for estimates of all parameters decrease; (4) The SDs and Medians for 500 MADEs of $\hat{\alpha}_1(u)$ and $\hat{\alpha}_2(u)$ at 20 fixed grid points decrease as T or N is increased. Based on these findings, we conclude that the estimates of all parameters and varying coefficient functions are fairly close to their true values, and the deviations decrease with increasing of sample size. Overall, our proposed estimators for the model (12) perform well in finite sample cases.

Table 1. Simulation results of parametric estimates $\hat{\beta}_1, \hat{\beta}_2, \hat{\rho}, \hat{\lambda}$ and $\hat{\sigma}_e^2$.

			***T* = 10**			***T* = 15**			***T* = 20**		
N	**Parameter**	(ρ, λ)	$\mathbf{(0.2, 0.7)}$	$\mathbf{(0.4, 0.4)}$	$\mathbf{(0.7, 0.2)}$	$\mathbf{(0.2, 0.7)}$	$\mathbf{(0.4, 0.4)}$	$\mathbf{(0.7, 0.2)}$	$\mathbf{(0.2, 0.7)}$	$\mathbf{(0.4, 0.4)}$	$\mathbf{(0.7, 0.2)}$
25	β_1	Mean	0.9998	0.9993	0.9990	0.9994	0.9978	0.9968	0.9999	0.9998	0.9999
		Median	0.9985	0.9962	1.0002	1.0004	0.9987	0.9989	1.0002	1.0003	1.0003
		SD	0.0344	0.0391	0.0399	0.0277	0.0310	0.0309	0.0235	0.0261	0.0259
		RMSE	0.0344	0.0390	0.0398	0.0277	0.0310	0.0310	0.0235	0.0260	0.0259
	β_2	Mean	1.4971	1.4972	1.4963	1.4976	1.4978	1.4983	1.5015	1.5011	1.5004
		Median	1.4970	1.4940	1.4960	1.4984	1.4978	1.4975	1.5018	1.5017	1.5007
		SD	0.0357	0.0394	0.0380	0.0273	0.0300	0.0296	0.0231	0.0257	0.0257
		RMSE	0.0357	0.0395	0.0381	0.0273	0.0300	0.0296	0.0231	0.0257	0.0257
	ρ	Mean	0.2513	0.4282	0.7199	0.2354	0.4139	0.7103	0.2303	0.4125	0.7089
		Median	0.2489	0.4335	0.7214	0.2385	0.4162	0.7130	0.2317	0.4115	0.7060
		SD	0.0948	0.0640	0.0396	0.0713	0.0465	0.0279	0.0601	0.0417	0.0236
		RMSE	0.1076	0.0698	0.0443	0.0795	0.0484	0.0297	0.0671	0.0435	0.0252
	λ	Mean	0.7598	0.4237	0.2022	0.7423	0.4153	0.2029	0.7339	0.4091	0.1984
		Median	0.7625	0.4368	0.1995	0.7484	0.4101	0.1965	0.7339	0.4054	0.2000
		SD	0.0921	0.0955	0.0592	0.0723	0.0616	0.0409	0.0626	0.0526	0.0316
		RMSE	0.1096	0.0982	0.0591	0.0837	0.0633	0.0410	0.0711	0.0533	0.0316
	σ_e^2	Mean	0.4196	0.4020	0.3950	0.4483	0.4366	0.4334	0.4608	0.4507	0.4473
		Median	0.4172	0.4139	0.3944	0.4490	0.4388	0.4368	0.4606	0.4501	0.4496
		SD	0.0438	0.0443	0.0470	0.0385	0.0361	0.0383	0.0349	0.0313	0.0328
		RMSE	0.0915	0.1075	0.1149	0.0644	0.0729	0.0768	0.0524	0.0584	0.0621
49	β_1	Mean	1.0005	1.0012	1.0017	0.9984	0.9984	0.9994	0.9993	0.9991	1.0006
		Median	0.9994	1.0026	1.0036	0.9991	0.9998	0.9980	1.0003	0.9989	0.9999
		SD	0.0267	0.0293	0.0286	0.0185	0.0207	0.0209	0.0176	0.0195	0.0175
		RMSE	0.0267	0.0293	0.0286	0.0185	0.0207	0.0209	0.0176	0.0194	0.0175
	β_2	Mean	1.5010	1.5006	1.5002	1.5014	1.5011	1.5008	1.4982	1.4977	1.5027
		Median	1.5009	1.4944	1.5011	1.5009	1.5015	1.5022	1.4981	1.4970	1.5007
		SD	0.0264	0.0290	0.0283	0.0187	0.0204	0.0215	0.0159	0.0179	0.0181
		RMSE	0.0264	0.0290	0.0283	0.0187	0.0204	0.0216	0.0159	0.0180	0.0181
	ρ	Mean	0.2545	0.4213	0.7109	0.2456	0.4135	0.7064	0.2347	0.4042	0.7059
		Median	0.2538	0.4209	0.7133	0.2446	0.4110	0.7079	0.2297	0.4023	0.7064
		SD	0.0728	0.0466	0.0321	0.0645	0.0403	0.0233	0.0556	0.0282	0.0168
		RMSE	0.0908	0.0512	0.0338	0.0789	0.0424	0.0241	0.0655	0.0285	0.0177
	λ	Mean	0.7540	0.4207	0.2089	0.7482	0.4079	0.1997	0.7410	0.4066	0.1961
		Median	0.7533	0.4122	0.1995	0.7446	0.3983	0.1944	0.7468	0.4042	0.1959
		SD	0.0780	0.0847	0.0512	0.0661	0.0481	0.0353	0.0570	0.0388	0.0272
		RMSE	0.0947	0.0900	0.0519	0.0817	0.0486	0.0354	0.0701	0.0393	0.0275
	σ_e^2	Mean	0.4412	0.4291	0.4238	0.4680	0.4526	0.4481	0.4759	0.4650	0.4582
		Median	0.4378	0.4205	0.4179	0.4649	0.4536	0.4452	0.4714	0.4614	0.4590
		SD	0.0378	0.0350	0.0393	0.0319	0.0277	0.0311	0.0290	0.0261	0.0252
		RMSE	0.0692	0.0790	0.0857	0.0551	0.0549	0.0605	0.0417	0.0436	0.0488
81	β_1	Mean	1.0013	1.0016	1.0018	0.9995	0.9996	1.0000	1.0001	1.0001	0.9997
		Median	1.0004	1.0029	1.0036	1.0002	0.9999	1.0003	1.0002	1.0003	1.0003
		SD	0.0205	0.0223	0.0216	0.0156	0.0175	0.0175	0.0126	0.0146	0.0145
		RMSE	0.0205	0.0223	0.0216	0.0156	0.0175	0.0175	0.0126	0.0146	0.0145
	β_2	Mean	1.5012	1.5017	1.5019	1.4996	1.4993	1.4991	1.5002	1.5006	1.5005
		Median	1.5023	1.5023	1.5022	1.5000	1.5007	1.5009	1.5009	1.5014	1.5007
		SD	0.0197	0.0219	0.0215	0.0147	0.0160	0.0158	0.0127	0.0141	0.0139
		RMSE	0.0197	0.0219	0.0215	0.0147	0.0160	0.0157	0.0127	0.0141	0.0139
	ρ	Mean	0.2654	0.4161	0.7087	0.2489	0.4096	0.7049	0.2463	0.4072	0.7042
		Median	0.2458	0.4184	0.7099	0.2497	0.4079	0.7058	0.2455	0.4071	0.7060
		SD	0.0624	0.0333	0.0252	0.0522	0.0259	0.0163	0.0492	0.0214	0.0161
		RMSE	0.0767	0.0370	0.0266	0.0714	0.0276	0.0170	0.0551	0.0225	0.0166
	λ	Mean	0.7556	0.4190	0.2088	0.7519	0.4076	0.2016	0.7427	0.4044	0.2013
		Median	0.7505	0.4002	0.2006	0.7507	0.4025	0.1973	0.7453	0.4037	0.2000
		SD	0.0723	0.0688	0.0442	0.0531	0.0425	0.0274	0.0443	0.0289	0.0231
		RMSE	0.0911	0.0712	0.0450	0.0742	0.0430	0.0274	0.0598	0.0292	0.0231
	σ_e^2	Mean	0.4518	0.4362	0.4340	0.4724	0.4577	0.4565	0.4779	0.4654	0.4649
		Median	0.4485	0.4333	0.4287	0.4726	0.4570	0.4558	0.4753	0.4656	0.4650
		SD	0.0366	0.0305	0.0338	0.0254	0.0223	0.0236	0.0202	0.0174	0.0220
		RMSE	0.0604	0.0707	0.0742	0.0475	0.0478	0.0495	0.0335	0.0387	0.0414

Figures 2 and 3 present the fitting results and 95% confidence intervals of $\alpha_1(u)$ and $\alpha_2(u)$ under $N = 49$ (or 81) and $T = 15$ (or 20), where the short dashed curves in black are the average fits over 500 simulations by PQMLE, the solid curves in red are the true values of nonparametric functions and the two long dashed curves in black are the corresponding 95% confidence bands. We can see that the short dashed curve is close to the solid curve, and the confidence bandwidth gradually becomes narrow with the increase of the sample size. They indicate that the nonparametric estimation procedure is feasible in the case of small samples.

Table 2. The Medians and SDs of MADE values for $\hat{\alpha}_1(u)$.

		***T* = 10**			***T* = 15**			***T* = 20**		
(ρ, λ)		***N* = 25**	***N* = 49**	***N* = 81**	***N* = 25**	***N* = 49**	***N* = 81**	***N* = 25**	***N* = 49**	***N* = 81**
(0.2, 0.7)	Median	0.0782	0.0553	0.0443	0.0622	0.0426	0.0347	0.0553	0.0405	0.0310
	SD	0.0203	0.0148	0.0113	0.0174	0.0105	0.0096	0.0157	0.0101	0.0080
(0.4, 0.4)	Median	0.0853	0.0604	0.0488	0.0693	0.0473	0.0379	0.0628	0.0445	0.0337
	SD	0.0219	0.0159	0.0121	0.0193	0.0116	0.0110	0.0177	0.0112	0.0092
(0.7, 0.2)	Median	0.0835	0.0588	0.0495	0.0670	0.0491	0.0383	0.0627	0.0424	0.0334
	SD	0.0217	0.0150	0.0117	0.0187	0.0128	0.0108	0.0177	0.0107	0.0090

Table 3. The Medians and SDs of MADE values for $\hat{\alpha}_2(u)$.

		***T* = 10**			***T* = 15**			***T* = 20**		
(ρ, λ)		***N* = 25**	***N* = 49**	***N* = 81**	***N* = 25**	***N* = 49**	***N* = 81**	***N* = 25**	***N* = 49**	***N* = 81**
(0.2, 0.7)	Median	0.0776	0.0532	0.0444	0.0683	0.0442	0.0350	0.0563	0.0378	0.0306
	SD	0.0213	0.0152	0.0115	0.0183	0.0115	0.0091	0.0143	0.0112	0.0078
(0.4, 0.4)	Median	0.0856	0.0599	0.0480	0.0672	0.0498	0.0384	0.0618	0.0426	0.0336
	SD	0.0231	0.0166	0.0125	0.0198	0.0128	0.0099	0.0153	0.0123	0.0088
(0.7, 0.2)	Median	0.0846	0.0606	0.0476	0.0653	0.0460	0.0375	0.0607	0.0413	0.0329
	SD	0.0226	0.0159	0.0129	0.0191	0.0113	0.0097	0.0153	0.0115	0.0090

Example 2. *The second example is used to show that misspecification for the model (12) will lead to inconsistent parameter estimates. Here are the three most likely misspecified models, which ignore the spatial correlation, serial correlation and spatio-temporal correlations in the model (12), respectively:*

$$\begin{cases} y_{it} = x_{it1}\beta_1 + x_{it2}\beta_2 + z_{it1}\alpha_1(u_{it}) + z_{it2}\alpha_2(u_{it}) + b_i + \varepsilon_{it}, \\ \varepsilon_{it} = \lambda\varepsilon_{i,t-1} + e_{it}, \end{cases} \tag{13}$$

$$\begin{cases} y_{it} = x_{it1}\beta_1 + x_{it2}\beta_2 + z_{it1}\alpha_1(u_{it}) + z_{it2}\alpha_2(u_{it}) + b_i + \varepsilon_{it}, \\ \varepsilon_{it} = \rho \sum\limits_{j=1}^{N} w_{ij}\varepsilon_{jt} + e_{it}, \end{cases} \tag{14}$$

$$y_{it} = x_{it1}\beta_1 + x_{it2}\beta_2 + z_{it1}\alpha_1(u_{it}) + z_{it2}\alpha_2(u_{it}) + b_i + e_{it}, \tag{15}$$

where all variables in the above models are the same as the model (12). No loss of generality, we only study the case that $\rho = 0.4$ and $\lambda = 0.4$. Additionally, we set $N = 25, 49$, $T = 10$ and $mcn = 500$. The experimental results are presented in Table 4.

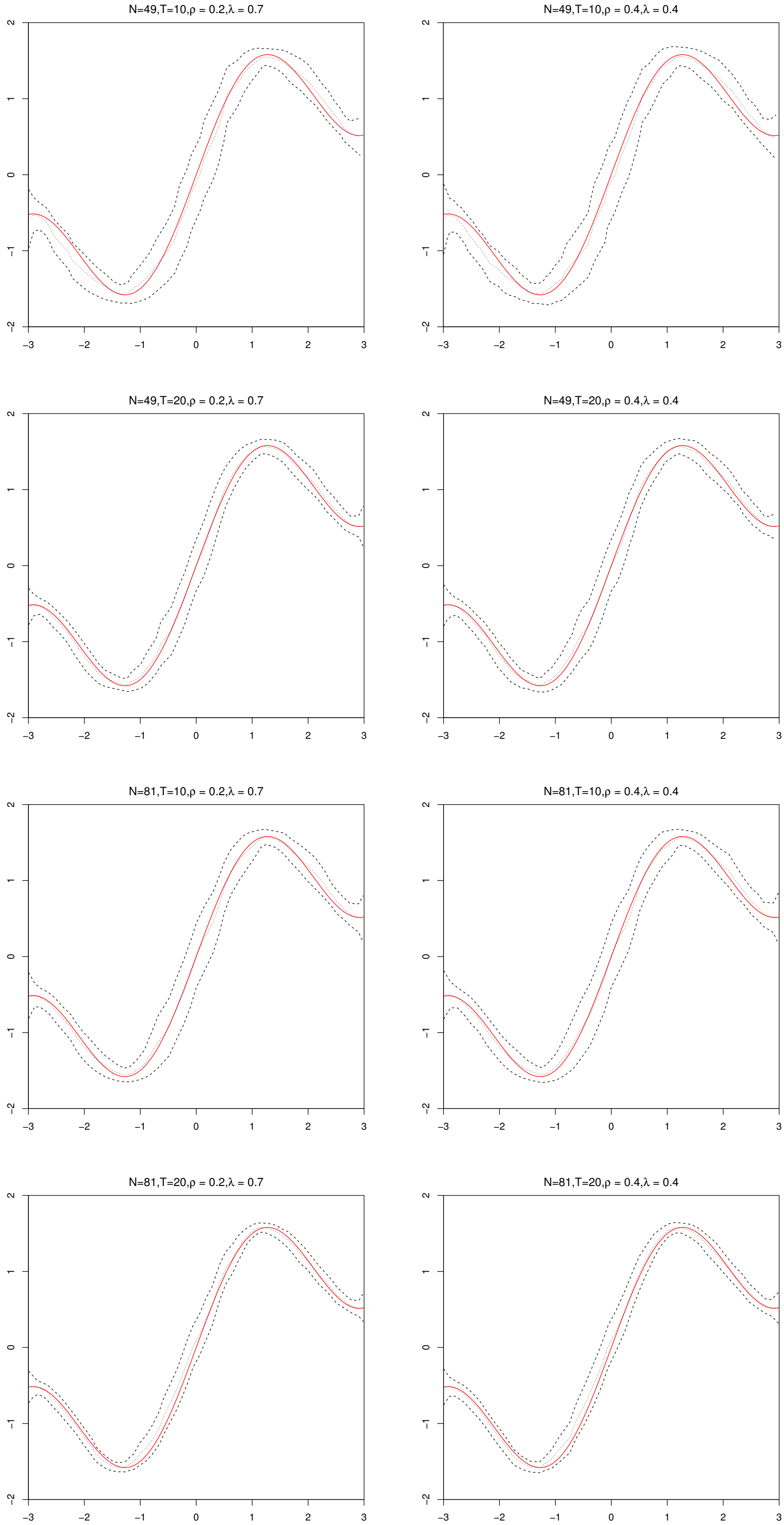

Figure 2. The fitting results and 95% confidence intervals of $\alpha_1(u)$ in the model (12).

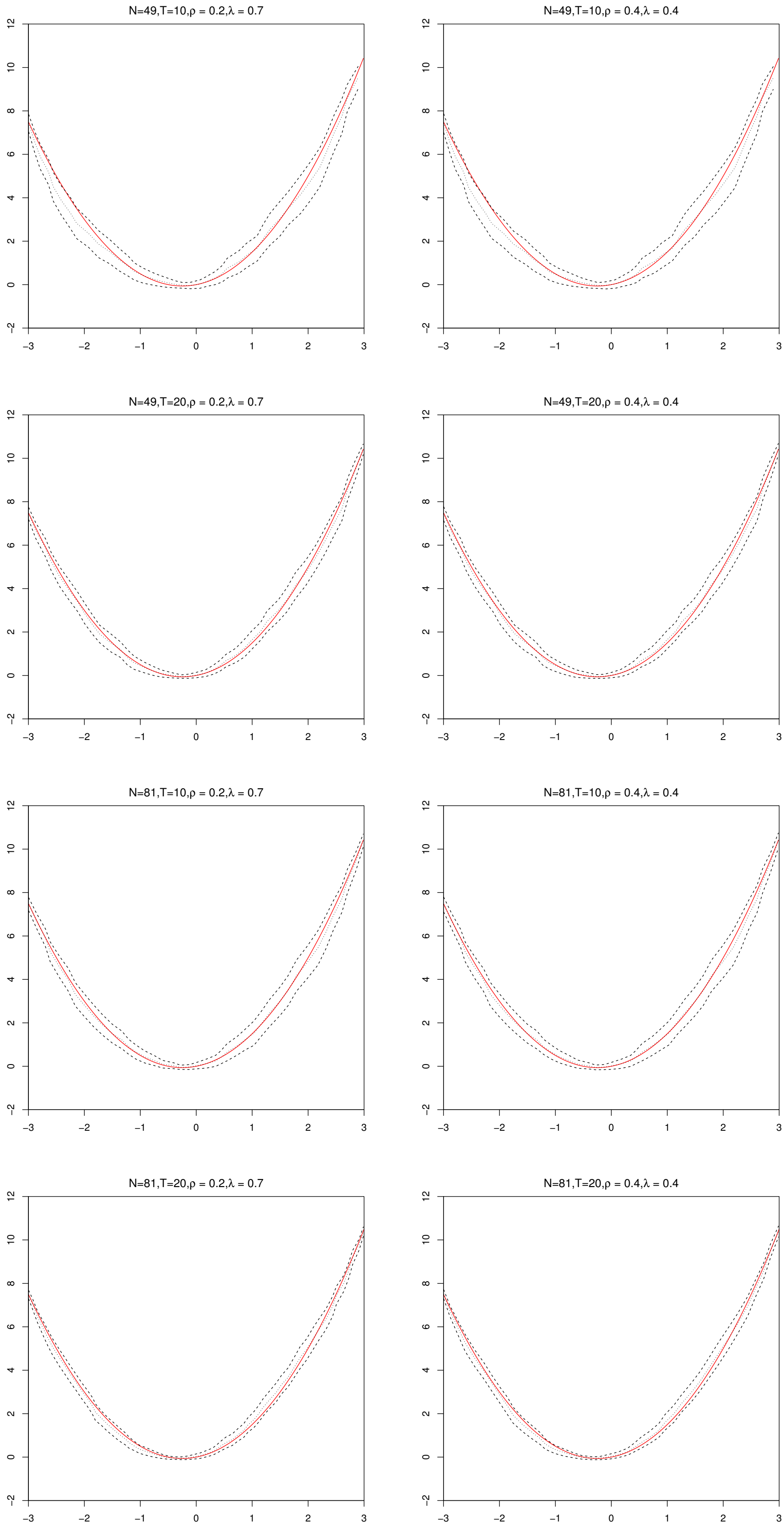

Figure 3. The fitting results and 95% confidence intervals of $\alpha_2(u)$ in the model (12).

Table 4 lists the Means, Medians, SDs, RMSEs and MRs of parameter estimates in the models (12)–(15), where MR is the growth rate of RMSE on the basis of that in the model (12).

From Table 4, we can see that: (1) The Means and Medians for the estimates of all parameters in the model (12) are closer to true values as sample size increases. However, it is easy to see that the Means and Medians of $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\rho}$, $\hat{\lambda}$ and $\hat{\sigma}_e^2$ in the models (13)–(15) do not converge with the increasing of N, indicating that they are not stable. (2) The SDs and RMSEs of almost all parameter estimators in the models (13)–(15) are larger than that in the model (12). In particular, the SDs and RMSEs of $\hat{\sigma}_e^2$ do not decrease with the increasing of sample size. (3) MRs of most parameter estimators are greater than 0% and increase with the increasing of sample size, especially for $\hat{\beta}_2$, $\hat{\rho}$ and $\hat{\sigma}_e^2$. In addition, MRs of $\hat{\sigma}_e^2$ in the models (13) and (14) are less than 0%, which again indicates that the estimator of σ_e^2 is unstable. It can be concluded that model misspecification would result in inconsistent parameter estimators. It further indicates that our proposed model is more effective and reliable.

Table 4. Simulation results of parametric estimates for the models (12)–(15).

Model		**$N = 25, T=10$**					**$N = 49, T=10$**				
		β_1	β_2	ρ	λ	σ_e^2	β_1	β_2	ρ	λ	σ_e^2
Model (12)	Mean	1.9993	1.4972	0.4282	0.4237	0.4020	1.0012	1.5006	0.4213	0.4207	0.4291
	Median	0.9962	1.4940	0.4335	0.4368	0.4139	1.0016	1.4944	0.4209	0.4122	0.4205
	SD	0.0391	0.0394	0.0640	0.0955	0.0443	0.0293	0.0290	0.0466	0.0847	0.0350
	RMSE	0.0390	0.0395	0.0698	0.0982	0.1075	0.0293	0.0290	0.0512	0.0900	0.0790
Model (13)	Mean	0.9989	1.4996	-	0.5610	0.5508	1.0011	1.5012	-	0.5456	0.5735
	Median	0.9981	1.4985	-	0.5649	0.5485	1.0026	1.4924	-	0.5689	0.5778
	SD	0.0429	0.0464	-	0.1066	0.0671	0.0324	0.0328	-	0.0694	0.0499
	RMSE	0.0428	0.0463	-	0.1930	0.0841	0.0324	0.0328	-	0.1612	0.0888
	MR	97.4%	17.21%	-	96.54%	−21.77%	10.58%	13.10%	-	79.11%	12.41%
Model (14)	Mean	0.9984	1.4956	0.4145	-	0.4145	1.0009	1.5003	0.4542	-	0.4410
	Median	0.9961	1.4939	0.4667	-	0.4139	1.0006	1.5003	0.4567	-	0.4379
	SD	0.0440	0.0426	0.0670	-	0.0499	0.0310	0.0328	0.0448	-	0.0351
	RMSE	0.0439	0.0427	0.0942	-	0.0989	0.0310	0.0328	0.0703	-	0.0686
	MR	12.56%	8.10%	34.96%	-	−8.00%	5.80%	13.10%	37.30%	-	−13.16%
Model (15)	Mean	0.9969	1.4991	-	-	0.6073	1.0000	1.5007	-	-	0.6309
	Median	0.9964	1.4982	-	-	0.5955	0.9987	1.5032	-	-	0.6278
	SD	0.0550	0.0544	-	-	0.0953	0.0368	0.0408	-	-	0.0677
	RMSE	0.0549	0.0543	-	-	0.1434	0.0367	0.0407	-	-	0.1473
	MR	40.76%	37.46%	-	-	33.39%	25.25%	40.34%	-	-	86.45%

Note: True values $(\beta_{10}, \beta_{20}, \sigma_{e0}^2)' = (1, 1.5, 0.5)'$ for the models (12)–(15), $\lambda_0 = 0.4$ for the model (13) and $\rho_0 = 0.4$ for the model (14).

5. Real Data Analysis

In this section, we employ the proposed model and its estimation method to study the driving forces of Chinese resident consumption rate. This dataset was collected on 1 August 2022) from the *China Statistical Yearbook* (http://www.stats.gov.cn/sj/ndsj/) for 2008 to 2020 and covers 30 provincial administrative regions (except Tibet, Taiwan, Hong Kong and Macau). Based on the research results drawn by Ding and Chen [45] and Ding [46], let YC be response variable and LR, CR, ER, GR and TR be regressors. There is no doubt that per capita disposable income has an important impact on the resident consumption rate. Therefore, we assume that the impacts of the above regressors on resident consumption rate may be realized through per capita disposable income and IR is selected as their intermediate variable. The definitions of these variables and their meanings are given in Table 5.

Table 5. Variable definitions and their meanings.

Response Variable	Definition
YC	The ratio of resident consumption to GDP
Regressors	**Definition**
LR	The ratio of the population over 65 to the population between 14 and 65
CR	The ratio of the population under 14 to the population between 14 and 65
ER	The ratio of the population with junior college degree or above to total population
GR	The ratio of the male population to total population
TR	The ratio of the tertiary industry to GDP
Intermediate Variable	**Definition**
IR	Growth rate of per capita disposable income

Firstly, Table 6 and Figure 4 show the descriptive statistics of the response variable, five regressors and intermediate variable. From observing Table 6, we can draw the conclusion that LR, CR, ER, GR, TR and IR are steady, as well as concluding that GR has a small fluctuation range. In addition, Figure 4 presents the scatter plots between YC versus each regressor (LR, CR, ER, GR and TR). It can be found that the regressor LR has a linear effect on the response variable YC. The rest of the regressors have nonlinear effects on the response variable YC.

Table 6. The descriptive statistics of the response variable, five regressors and intermediate variable.

	Min	1st Qu.	Median	Mean	3rd Qu.	Max	SD	Range
YC	0.2160	0.3260	0.3609	0.3642	0.4001	0.5811	0.0579	0.3651
LR	0.0744	0.1132	0.1361	0.1407	0.1612	0.2548	0.0351	0.1804
CR	0.0965	0.1791	0.2311	0.2289	0.2767	0.3981	0.0639	0.3016
ER	0.0285	0.0806	0.1063	0.1219	0.1398	0.4769	0.0698	0.4484
GR	0.4873	0.5057	0.5108	0.5117	0.5168	0.5519	0.0091	0.0645
TR	0.2830	0.3875	0.4461	0.4581	0.5100	0.8400	0.0990	0.5570
IR	−0.0285	0.0795	0.0890	0.0946	0.1162	0.2038	0.0355	0.2323

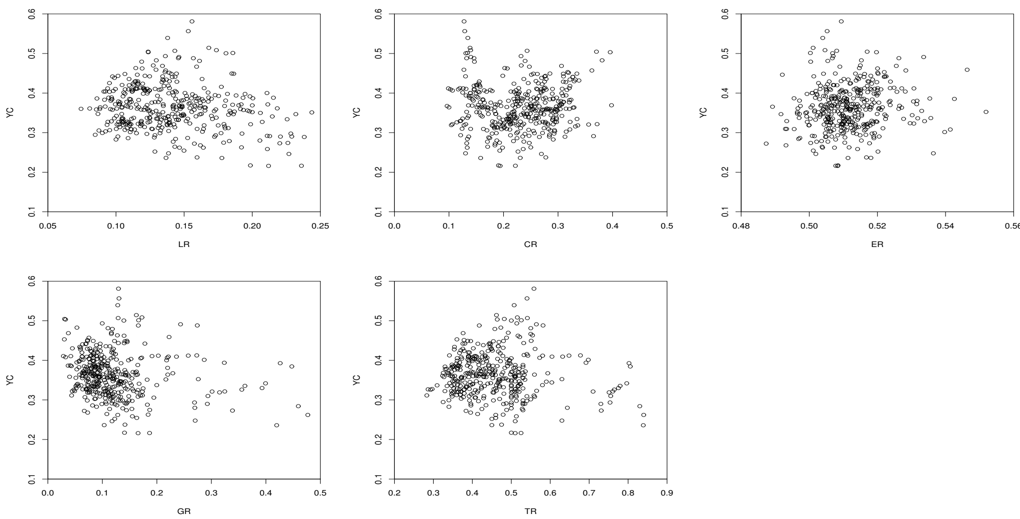

Figure 4. Scatter plots of the response variable versus five regressors, respectively.

Based on the above comprehensive analysis, the study on driving forces of Chinese resident consumption rate can be analyzed by establishing the following PLVCPDRM with nonseparable space-time filters:

$$\begin{aligned} YC_{it} &= LR_{it}\beta + CR_{it}\alpha_1(IR_{it}) + ER_{it}\alpha_2(IR_{it}) + GR_{it}\alpha_3(IR_{it}) + TR_{it}\alpha_4(IR_{it}) + b_i + \varepsilon_{it} \\ \varepsilon_{it} &= \rho\sum_{j=1}^{N} w_{ij}\varepsilon_{jt} + \lambda\varepsilon_{i,t-1} + e_{it}, \quad i = 1,\ldots,30, t = 1,\ldots,13, \end{aligned} \tag{16}$$

where $W = (w_{ij})_{30\times30}$ is a normalized spatial weight matrix calculated by the Euclidean distance in the light of the longitude and latitude coordinates of any two provinces.

Table 7 reports the estimation results of parameters in the model (16). It can be seen that $\hat{\rho} = 0.6126$ and $\hat{\lambda} = 0.3674$ are significant. Namely, it indicates that there exist strong and positive spatial and serial correlations among the disturbance terms in the model (16). Furthermore, $\hat{\beta} = -0.1751 < 0$ is significant, which means that the linear effect of LR on the resident consumption rate is negative. Figure 5 shows the varying coefficient effects of CR, ER, GR and TR to YC and their 95% confidence intervals. It can be seen that CR, ER, GR and TR have obvious nonlinear effects on resident consumption rate with IR.

Table 7. Estimation results of parameters in the model (16).

	β	ρ	λ	σ_e^2
Estimator	−0.1751 ***	0.6126 ***	0.3674 ***	0.0005 ***
SD	0.1072	0.1432	0.0329	9.6963×10^{-5}

Notes: *** represents that the regressor is significant under the significance level 1%.

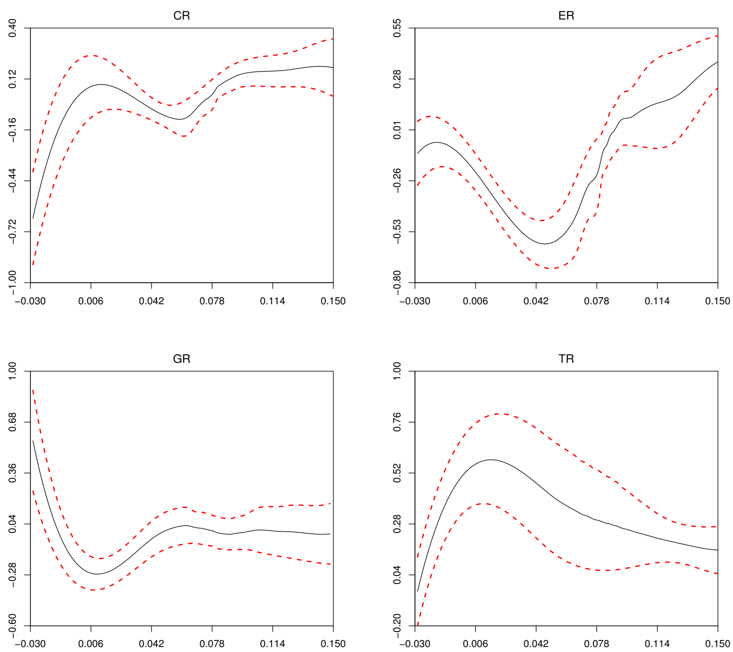

Figure 5. Varying coefficient effects of CR, ER, GR and TR to YC and their 95% confidence intervals, respectively.

6. Concluding Remarks

In order to sufficiently use the information of spatial and serial correlations in the disturbances when modeling space-time data by regression models, we propose a fixed effects PLVCPDRM with nonseparable space-time filters. It can not only simultaneously capture non/linear effects of regressors and space-time correlations of error structure, but also overcome the “curse of dimensionality” in multivariate nonparametric regression models.

In this paper, the PQMLEs of unknown parameters and varying coefficient functions for this model are constructed. Under the regular assumptions, we prove that the estimators satisfy consistency and asymptotic normality. Monte Carlo simulations show that the proposed estimators have good finite sample performances. In addition, ignoring spatial and serial correlations in errors of the model would result in inconsistent and inefficient estimators. Finally, a Chinese resident consumption rate dataset is used to illustrate our estimation method.

This paper mainly focuses on the estimation of a fixed effects PLVCPDRM with nonseparable space-time filters. In the future, we may study the methods of variable selection, Bayesian estimation and quantile regression for the proposed model in our paper; we can also use the proposed method to study similar semiparametric panel data regression models with space-time filters.

Author Contributions: Formal analysis, B.L. and S.L.; Methodology, B.L. and J.C.; Software and writing—original draft, B.L.; Supervision, writing—review and editing and funding acquisition, J.C.; Data curation, S.L. All authors have read and agreed to the published version of manuscript.

Funding: This work is supported by National Social Science Fund of China (22BTJ024) and Natural Science Foundation of Fujian Province (2020J01170, 2022J01193).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

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

Appendix A

To prove the theoretical results, the following facts and lemma will be used frequently in the sequel.

Fact 1: If $A_{1,NT}$ and $A_{2,NT}$ are $NT \times NT$ matrices that are uniformly bounded in row sums (resp., column sums), then $A_{1,NT}A_{2,NT}$ is also uniformly bounded in row sums (resp., column sums).

Fact 2: If $A_{1,NT}$ is uniformly bounded in row sums (resp., column sums) and $A_{2,NT}$ is a conformable matrix whose elements are uniformly $O(o_{NT})$, then so are the elements of $A_{1,NT}A_{2,NT}$ (resp. $A_{2,NT}A_{1,NT}$).

The above two Facts can be found in Su and Jin [35].

Lemma A1. *Under Assumptions 1–3, we have that*

(i) $\widetilde{Z}'\widetilde{Z}/NT - I_K = O_P\left(\zeta_0(K)\sqrt{K}/\sqrt{N}\right)$.

(ii) $\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}/NT - I_K = O_P\left(\zeta_0(K)\sqrt{K}/\sqrt{N}\right)$.

Proof. (i) See the proof of Theorem 1 in Newey [40] (pp. 161–162); (ii) It follows from Assumption 3(i) by similar proof of (i). □

Proof of Theorem 1. Substituting $\varepsilon(\beta,\gamma) = \varepsilon + X(\beta_0 - \beta) + \widetilde{Z}(\gamma_0 - \gamma)$ into the quasi-log-likelihood function (6), we have that

$$\begin{aligned}\log L_{N,T}(\theta) =& -\frac{N(T-1)}{2}\log\left(2\pi\sigma_e^2\right) + (T-1)\log|S_N(\rho)| - \frac{1}{2}\log|\Omega_{N,T-1}(\eta)| \\ & - \frac{1}{2\sigma_e^2}\varepsilon'\mathbb{J}_{NT}(\eta)\varepsilon - \frac{1}{\sigma_e^2}\varepsilon'\mathbb{J}_{NT}(\eta)[X(\beta_0-\beta)+\widetilde{Z}(\gamma_0-\gamma)] \\ & - \frac{1}{2\sigma_e^2}[X(\beta_0-\beta)+\widetilde{Z}(\gamma_0-\gamma)]'\mathbb{J}_{NT}(\eta)[X(\beta_0-\beta)+\widetilde{Z}(\gamma_0-\gamma)] \\ =& \log L_{NT,1}(\theta) + \log L_{NT,2}(\theta) + \log L_{NT,3}(\theta),\end{aligned}$$

where

$$\begin{aligned}\log L_{NT,1}(\theta) =& -\frac{N(T-1)}{2}\log\left(2\pi\sigma_e^2\right) + (T-1)\log|S_N(\rho)| - \frac{1}{2}\log|\Omega_{N,T-1}(\eta)| \\ & - \frac{1}{2\sigma_e^2}\varepsilon'\mathbb{J}_{NT}(\eta)\varepsilon,\end{aligned}$$

$$\begin{aligned}\log L_{NT,2}(\theta) &= -\frac{1}{\sigma_e^2}\varepsilon'\mathbb{J}_{NT}(\eta)[X(\beta_0-\beta)+\widetilde{Z}(\gamma_0-\gamma)], \\ \log L_{NT,3}(\theta) &= -\frac{1}{2\sigma_e^2}[X(\beta_0-\beta)+\widetilde{Z}(\gamma_0-\gamma)]'\mathbb{J}_{NT}(\eta)[X(\beta_0-\beta)+\widetilde{Z}(\gamma_0-\gamma)].\end{aligned}$$

In order to prove that $\frac{1}{N(T-1)}\log L_{N,T}(\theta) - Q_{N,T}(\theta) \xrightarrow{P} 0$ uniformly for θ, it is sufficient to prove that $\frac{1}{N(T-1)}\log L_{NT,j}(\theta) - Q_{NT,j}(\theta) \xrightarrow{P} 0$ uniformly for θ according to that $\log L_{NT,3}(\theta)$ is deterministic by Assumption 4, where $Q_{NT,j}(\theta) = E\frac{1}{N(T-1)}\log L_{NT,j}(\theta)$, $j = 1,2$. For case $j = 1$, as $\varepsilon = B_{NT}^{*-1}F_{NT,N(T+m)}e_{N,T+m}$, we have that

$$\begin{aligned}&\frac{1}{N(T-1)}\log L_{NT,1}(\theta) - Q_{NT,1}(\theta) \\ =& -\frac{1}{2\sigma_e^2}\Bigg[\frac{1}{N(T-1)}e'_{N,T+m}F'_{NT,N(T+m)}B_{NT}^{*-1\prime}\mathbb{J}_{NT}(\eta)B_{NT}^{*-1}F_{NT,N(T+m)}e_{N,T+m} \\ & - E\frac{1}{N(T-1)}e'_{N,T+m}F'_{NT,N(T+m)}B_{NT}^{*-1\prime}\mathbb{J}_{NT}(\eta)B_{NT}^{*-1}F_{NT,N(T+m)}e_{N,T+m}\Bigg].\end{aligned}$$

By using Lemma 7 in Yu et al. [47], we have that $\frac{1}{N(T-1)}\log L_{NT,1}(\theta) - Q_{NT,1}(\theta) \xrightarrow{P} 0$ uniformly for θ when T is fixed due to the explicit forms of $F_{NT,N(T+m)}$, B_{NT}^{*-1} and $\mathbb{J}_{NT}(\eta)$ which are UB from Assumption 4. For case $j = 2$, similarly, as B_{NT}^{*-1} and $\mathbb{J}_{NT}(\eta)$ are UB, using Lemma 8 in Yu et al. [47], we have that $\frac{1}{N(T-1)}e'_{N,T+m}F'_{NT,N(T+m)}B_{NT}^{*-1\prime}\mathbb{J}_{NT}(\theta)[X(\beta_0-\beta) + \widetilde{Z}(\gamma_0-\gamma)] \xrightarrow{P} 0$ when T is fixed.

To prove that $Q_{N,T}(\theta)$ is uniformly equicontinuous, we just need to investigate $Q_{NT,1}(\theta)$ and $Q_{NT,3}(\theta)$ according to that $Q_{NT,2}(\theta) = 0$. It is easy to know that

$$\begin{aligned}Q_{NT,1}(\theta) =& -\frac{1}{2}\log\left(2\pi\sigma_e^2\right) + \frac{1}{N}\log|S_N(\rho)| - \frac{1}{2N(T-1)}\log|\Omega_{N,T-1}(\eta)| \\ & - \frac{1}{2\sigma_e^2 N(T-1)}E\left[e'_{N,T+m}F'_{NT,N(T+m)}B_{NT}^{*-1\prime}\mathbb{J}_{NT}(\eta)B_{NT}^{*-1}F_{NT,N(T+m)}e_{N,T+m}\right].\end{aligned}$$

It is obvious that $-\frac{1}{2}\log(2\pi\sigma_e^2) + \frac{1}{N}\log|S_N(\rho)|$ is uniformly equicontinuous for η and σ_e^2, so is $\log|\Omega_{N,T-1}(\eta)|$. Furthermore, we know that

$$E[e'_{N,T+m}F'_{NT,N(T+m)}B_{NT}^{*-1\prime}\mathbb{J}_{NT}(\eta)B_{NT}^{*-1}F_{NT,N(T+m)}e_{N,T+m}] = \sigma_{e0}^2\,\mathrm{tr}[B_{NT}^{*-1\prime}\mathbb{J}_{NT}(\eta)B_{NT}^{*-1}]$$

by $F_{NT,N(T+m)}F'_{NT,N(T+m)} = I$. With the explicit form of $\mathbb{J}_{NT}(\eta)$, $\frac{1}{\sigma_e^2}\operatorname{tr}[B_{NT}^{*-1\prime}\mathbb{J}_{NT}(\eta)B_{NT}^{*-1}]$ is uniformly equicontinuous for η and σ_e^2. Thus, $Q_{NT,1}(\theta)$ is uniformly equicontinuous for θ and σ_e^2. For $Q_{NT,3}(\theta)$, we find that $Q_{NT,3}(\theta)$ is a linear quadratic form of parameters β and γ, and a function of $\mathbb{J}_{NT}(\eta)$. Thus, it is uniformly equicontinuous for θ.

To prove identification uniqueness of θ_0, note that

$$E\frac{1}{N(T-1)}\log L_{N,T}(\theta) - E\frac{1}{N(T-1)}\log L_{N,T}(\theta_0) \equiv T_{NT,1}\left(\eta,\sigma_e^2\right) + T_{NT,2}(\theta), \tag{A1}$$

where

$$\begin{aligned} &T_{NT,1}\left(\eta,\sigma_e^2\right) \\ &= -\frac{1}{2N(T-1)}\log\left|\sigma_e^2\Omega_{N,T-1}(\eta)\right| + \frac{1}{N}\log|S_N(\rho)| + \frac{1}{2N(T-1)}\log\left|\sigma_{e0}^2\Omega_{N,T-1}\right| - \frac{1}{N}\ln|S_N| \\ &\quad - \frac{1}{2\sigma_e^2 N(T-1)}E\left[e'_{N,T+m}F'_{NT,N(T+m)}B_{NT}^{*-1\prime}\mathbb{J}_{NT}(\eta)B_{NT}^{*-1}F_{NT,N(T+m)}e_{N,T+m}\right] + \frac{1}{2} \end{aligned}$$

and

$$T_{NT,2}(\theta) = -\frac{1}{2\sigma_e^2 N(T-1)}\left[X(\beta_0-\beta) + \widetilde{Z}(\gamma_0-\gamma)\right]'\mathbb{J}_{NT}(\eta)\left[X(\beta_0-\beta) + \widetilde{Z}(\gamma_0-\gamma)\right].$$

Consider an auxiliary nonseparable space-time disturbance process: $\varepsilon_{Nt} = \rho W\varepsilon_{Nt} + \lambda\varepsilon_{N,t-1} + e_{Nt}(t = 1,\ldots,T)$, where its quasi-log-likelihood function is

$$\begin{aligned} \log L_{p,NT}(\eta,\sigma_e^2) = &-\frac{N(T-1)}{2}\log\left(2\pi\sigma_e^2\right) + (T-1)\log|S_N(\rho)| - \frac{1}{2}\log|\Omega_{N,T-1}(\eta)| \\ &- \frac{1}{2\sigma_e^2}\varepsilon'\mathbb{J}_{NT}(\eta)\varepsilon. \end{aligned}$$

According to the information inequality for the auxiliary nonseparable space-time disturbance process, we know that $T_{NT,1}(\eta,\sigma_e^2) \le 0$ for any η and σ_e^2. Additionally, $T_{NT,2}(\theta)$ is a quadratic function of β and α with a negative semidefinite matrix given θ. We can find that identification uniqueness of β_0 and γ_0 would be possible when

$$\lim_{N\to\infty}\frac{1}{N(T-1)}[X,\widetilde{Z}]'\mathbb{J}_{NT}(\eta)[X,\widetilde{Z}]$$

is nonsingular given any value of η in Assumption 5 (iv), then $T_{NT,2}(\theta) < 0$ for any $\beta \ne \beta_0$ and $\gamma \ne \gamma_0$. In addition, when

$$\lim_{N\to\infty} T_{NT,1}\left(\eta,\sigma_e^2\right) \ne 0$$

for $\left(\eta,\sigma_e^2\right)' \ne \left(\eta_0,\sigma_{e0}^2\right)'$ in Assumption 6 is satisfied, the identification uniqueness of η_0 and σ_{e0}^2 is obtained. This completes the proof. □

Proof of Theorem 2. Denote $\theta^* = (\beta',\rho,\lambda,\sigma_e^2)'$ and $\theta_0^* = (\beta_0',\rho_0,\lambda_0,\sigma_{e0}^2)'$. According to the Taylor expansion of the first-order condition from maximizing the quasi-log-likelihood function

$$\begin{aligned} \log L_{N,T}(\theta^*) = &-\frac{N(T-1)}{2}\log\left(2\pi\sigma_e^2\right) - \frac{1}{2}\log|I_N + (T-1)(K_N(\eta) - I_N)| \\ &+ (T-1)\log|S_N(\rho)| - \frac{1}{2\sigma_e^2}[Y - X\beta]'(I - \mathbb{S}_\eta)'\mathbb{J}_{NT}(\eta)(I - \mathbb{S}_\eta)[Y - X\beta], \end{aligned}$$

we have

$$\sqrt{N(T-1)}(\hat{\theta}^* - \theta_0^*) = -\left(\frac{1}{N(T-1)}\frac{\partial^2 \log L_{N,T}(\tilde{\theta}^*_{NT})}{\partial\theta^*\partial\theta^{*\prime}}\right)^{-1}\frac{1}{\sqrt{N(T-1)}}\frac{\partial \log L_{N,T}(\theta_0^*)}{\partial\theta^*},$$

where $\tilde{\theta}^*_{NT}$ lies between $\hat{\theta}^*$ and θ_0^* and converges to θ_0^* in probability by Theorem 1. The proof is completed if we can show that

$$\frac{1}{\sqrt{N(T-1)}}\frac{\partial \log L_{N,T}(\theta_0^*)}{\partial\theta^*} \xrightarrow{L} \mathrm{N}\left(0, \Sigma_{\theta_0^*} + \Omega_{\theta_0^*}\right), \tag{A2}$$

$$\frac{1}{N(T-1)}\frac{\partial^2 \log L_{N,T}(\theta_0^*)}{\partial\theta^*\partial\theta^{*\prime}} - \Sigma_{\theta_0^*} = o_P(1) \tag{A3}$$

and

$$\frac{1}{N(T-1)}\frac{\partial^2 \log L_{N,T}(\tilde{\theta}^*_{NT})}{\partial\theta^*\partial\theta^{*\prime}} - \frac{1}{N(T-1)}\frac{\partial^2 \log L_{N,T}(\theta_0^*)}{\partial\theta^*\partial\theta^{*\prime}} = o_P(1) \text{ uniformly in } \tilde{\theta}^*_{NT}. \tag{A4}$$

To prove that (A2)–(A4) hold, we need to compute the following scores under the non-normality of errors:

$$\begin{aligned}
\frac{\partial \log L_{N,T}(\theta_0^*)}{\partial\beta} =& \frac{1}{\sigma_{e0}^2}X'(I-\mathbb{S}_{\eta_0})'\mathbb{J}_{NT}(I-\mathbb{S}_{\eta_0})B^{*-1}_{NT}F_{NT,N(T+m)}e_{N,T+m} + o_P(1),\\
\frac{\partial \log L_{N,T}(\theta_0^*)}{\partial\rho} =& -\frac{1}{2\sigma_{e0}^2}e'_{N,T+m}F'_{NT,N(T+m)}B^{*-1\prime}_{NT}(I-\mathbb{S}_{\eta_0})'\frac{\partial\mathbb{J}_{NT}}{\partial\rho}(I-\mathbb{S}_{\eta_0})B^{*-1}_{NT}F_{NT,N(T+m)}e_{N,T+m}\\
&-(T-1)\operatorname{tr}\left(WS_N^{-1}\right) - \frac{1}{2}\operatorname{tr}\left(K_N^{-1}\frac{\partial K_N}{\partial\rho}\right) + o_P(1),\\
\frac{\partial \log L_{N,T}(\theta_0^*)}{\partial\lambda} =& -\frac{1}{2\sigma_{e0}^2}e'_{N,T+m}F'_{NT,N(T+m)}B^{*-1\prime}_{NT}(I-\mathbb{S}_{\eta_0})'\frac{\partial\mathbb{J}_{NT}}{\partial\lambda}(I-\mathbb{S}_{\delta_0})B^{*-1}_{NT}F_{NT,N(T+m)}e_{N,T+m}\\
&-\frac{1}{2}\operatorname{tr}\left(K_N^{-1}\frac{\partial K_N}{\partial\lambda}\right) + o_P(1),\\
\frac{\partial \log L_{N,T}(\theta_0^*)}{\partial\sigma_e^2} =& \frac{1}{2\sigma_{e0}^4}e'_{N,T+m}F'_{NT,N(T+m)}B^{*-1\prime}_{NT}(I-\mathbb{S}_{\eta_0})'\mathbb{J}_{NT}(I-\mathbb{S}_{\eta_0})B^{*-1}_{NT}F_{NT,N(T+m)}e_{N,T+m}\\
&-\frac{N(T-1)}{2\sigma_{e0}^2} + o_P(1).
\end{aligned}$$

Defining

$$\begin{aligned}
\Delta_{NT} \equiv& \Bigg[\operatorname{vec}(F'_{NT,N(T+m)}B^{*-1\prime}_{NT}(I-\mathbb{S}_{\eta_0})'\frac{\partial\mathbb{J}_{NT}}{\partial\rho}(I-\mathbb{S}_{\eta_0})B^{*-1}_{NT}F_{NT,N(T+m)}),\\
&\operatorname{vec}(F'_{NT,N(T+m)}B^{*-1\prime}_{NT}(I-\mathbb{S}_{\eta_0})'\frac{\partial\mathbb{J}_{NT}}{\partial\lambda}(I-\mathbb{S}_{\eta_0})B^{*-1}_{NT}F_{NT,N(T+m)}),\\
&-\frac{1}{\sigma_{e0}^2}\operatorname{vec}(F'_{NT,N(T+m)}B^{*-1\prime}_{NT}(I-\mathbb{S}_{\eta_0})'\mathbb{J}_{NT}(I-\mathbb{S}_{\eta_0})B^{*-1}_{NT}F_{NT,N(T+m)})\Bigg]'\\
&\times\Bigg[\operatorname{vec}((F'_{NT,N(T+m)}B^{*-1\prime}_{NT}(I-\mathbb{S}_{\eta_0})'\frac{\partial\mathbb{J}_{NT}}{\partial\rho}(I-\mathbb{S}_{\eta_0})B^{*-1}_{NT}F_{NT,N(T+m)})^s),\\
&\operatorname{vec}((F'_{NT,N(T+m)}B^{*-1\prime}_{NT}(I-\mathbb{S}_{\eta_0})'\frac{\partial\mathbb{J}_{NT}}{\partial\lambda}(I-\mathbb{S}_{\eta_0})B^{*-1}_{NT}F_{NT,N(T+m)})^s),\\
&-\frac{1}{\sigma_{e0}^2}\operatorname{vec}((F'_{NT,N(T+m)}B^{*-1\prime}_{NT}(I-\mathbb{S}_{\eta_0})'\mathbb{J}_{NT}(I-\mathbb{S}_{\eta_0})B^{*-1}_{NT}F_{NT,N(T+m)})^s)\Bigg]
\end{aligned}$$

and $\mathcal{A}^s = \mathcal{A} + \mathcal{A}'$ for any square matrix $\mathcal{A}$, we can obtain the expected Hessian matrix

$$\Sigma_{\theta_0^*} = \frac{1}{N(T-1)} \begin{pmatrix} \frac{1}{\sigma_{e0}^2} X'(I-\mathbb{S}_{\eta_0})'\mathbb{J}_{NT}(I-\mathbb{S}_{\eta_0})X & \mathbf{0}_{p\times 3} \\ * & \frac{1}{4}\Delta_{NT} \end{pmatrix} \tag{A5}$$

which is a symmetric matrix.

According to the above results, it is not hard to obtain that

$$E\left[\frac{1}{N(T-1)} \frac{\partial \log L_{N,T}(\theta_0^*)}{\partial \theta^*} \frac{\partial \log L_{N,T}(\theta_0^*)}{\partial \theta^{*\prime}}\right] = \Sigma_{\theta_0^*} + \Omega_{\theta_0^*} + o_P(1),$$

where $\Omega_{\theta_0^*}$ is related to the third and fourth moments of e_{it}. The expression of $\Omega_{\theta_0^*}$ is as follows

$$\Omega_{\theta_0^*} = \frac{1}{N(T-1)} \begin{pmatrix} \mathbf{0}_{p\times p} & \frac{\mu_3}{\sigma_{e0}^4} X'(I-\mathbb{S}_{\eta_0})'\mathbb{J}_{NT} B_{NT}^{*-1} F_{NT,N(T+m)} \mathcal{P} \\ * & \frac{(\mu_4 - 3\sigma_{e0}^4)}{\sigma_{e0}^4} \mathcal{P}'\mathcal{P} \end{pmatrix}, \tag{A6}$$

where

$$\begin{aligned} \mathcal{P} = \Big[& -\frac{1}{2}\operatorname{vec}_D(F'_{NT,N(T+m)} B_{NT}^{*-1\prime}(I-\mathbb{S}_{\delta_0})' \frac{\partial \mathbb{J}_{NT}}{\partial \rho}(I-\mathbb{S}_{\delta_0}) B_{NT}^{*-1} F_{NT,N(T+m)}), \\ & -\frac{1}{2}\operatorname{vec}_D(F'_{NT,N(T+m)} B_{NT}^{*-1\prime}(I-\mathbb{S}_{\delta_0})' \frac{\partial \mathbb{J}_{NT}}{\partial \lambda}(I-\mathbb{S}_{\delta_0}) B_{NT}^{*-1} F_{NT,N(T+m)}), \\ & \frac{1}{2\sigma_{e0}^2}\operatorname{vec}_D(F'_{NT,N(T+m)} B_{NT}^{*-1\prime}(I-\mathbb{S}_{\delta_0})' \mathbb{J}_{NT}(I-\mathbb{S}_{\delta_0}) B_{NT}^{*-1} F_{NT,N(T+m)}) \Big] \end{aligned}$$

and $\operatorname{vec}_D(\mathcal{A})$ is the column vector formed by diagonal elements of a square matrix $\mathcal{A}$.

The components of $\frac{1}{\sqrt{N(T-1)}} \frac{\partial \log L_{N,T}(\theta_0^*)}{\partial \theta^*}$ are linear or quadratic functions of $e_{N,T+m}$. (A2) can be proved by the central limit theorem for linear quadratic forms of Theorem 1 in Kelejian and Prucha [48]. (A3) and (A4) can be proved by applying (38)–(41) in Yu et al. [47]. This completes the proof. □

Proof of Theorem 3. Note that $\hat{\eta}$ is consistent with η_0 in Theorem 1, from the Equation (7), it holds that

$$\begin{aligned} \hat{\gamma} =& [\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}]^{-1}\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)[Y - X\hat{\beta}] \\ =& [\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}]^{-1}\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)[X(\beta_0-\hat{\beta}) + (Z_{\alpha_0} - \widetilde{Z}\gamma_0) + \widetilde{Z}\gamma_0 + \varepsilon] \\ =& [\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}]^{-1}\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)X(\beta_0-\hat{\beta}) \\ & + [\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}]^{-1}\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)(Z_{\alpha_0} - \widetilde{Z}\gamma_0) \\ & + [\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}]^{-1}\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\varepsilon + \gamma_0, \end{aligned} \tag{A7}$$

where $Z_{\alpha_0} = (z'_{11}\alpha_0(u_{11}), \ldots, z'_{NT}\alpha_0(u_{NT}))'$. Consider the first term of the last equation in (A7), let 1_N be the indicator function for the smallest eigenvalue of $\widetilde{Z}'\widetilde{Z}/NT$ and $\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}/NT$ being greater than $1/2$. Then, $\lim_{N\to\infty} P(1_N = 1) = 1$. By Assumption 3, Lemma A1 and Fact 2, we have that

$$\begin{aligned} & 1_N \left\| [\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}]^{-1}\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)X(\beta_0-\hat{\beta}) \right\|^2 \\ =& 1_N(\beta_0-\hat{\beta})'X'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}[\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}]^{-2}\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)X(\beta_0-\hat{\beta}) \\ =& O_P(\frac{1}{NT})1_N(\beta_0-\hat{\beta})'X'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}[\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}/NT]^{-1}\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)X(\beta_0-\hat{\beta})/(NT) \\ \leq& O_P(\frac{1}{NT}). \end{aligned} \tag{A8}$$

For the second term of last equation in (A7), note that $\left\|Z_{\alpha_0} - \widetilde{Z}\gamma_0\right\| = O_P(K^{-r_2})$ by Assumption 3 (iii), then

$$\begin{aligned}
&1_N\left\|[\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}/NT]^{-1}\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)(Z_{\alpha_0} - \widetilde{Z}\gamma_0)/NT\right\| \\
=&1_N\Big\{(Z_{\alpha_0} - \widetilde{Z}\gamma_0)'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}[\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}/NT]^{-\frac{1}{2}}[\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}/NT]^{-1} \\
&\cdot[\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}/NT]^{-\frac{1}{2}}\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)(Z_{\alpha_0} - \widetilde{Z}\gamma_0)/NT\Big\}^{\frac{1}{2}} \\
\le&O_P(1)1_N\left\|\left[\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}/NT\right]^{-\frac{1}{2}}\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)(Z_{\alpha_0} - \widetilde{Z}\gamma_0)/NT\right\| \\
=&O_P(K^{-r_2}).
\end{aligned} \tag{A9}$$

For the third term of the last equation in (A7), it suffices to prove

$$\begin{aligned}
&E\left\{1_N\left\|[\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}/NT]^{-\frac{1}{2}}\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\varepsilon/NT\right\|^2\right\} \\
=&1_N E\left\{\varepsilon'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}[\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}/NT]^{-1}\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\varepsilon\right\}/(NT)^2 \\
=&1_N\,\mathrm{tr}\Big\{F'_{NT,N(T+m)}B^{*-1\prime}_{NT}\mathbb{J}_{NT}(\eta_0)\widetilde{Z}[\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}/NT]^{-1} \\
&\cdot\,\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)B^{*-1}_{NT}F_{NT,N(T+m)}\Big\}/(NT)^2 \\
\le&K/NT
\end{aligned}$$

by $F_{NT,N(T+m)}F'_{NT,N(T+m)} = I$ and $\mathbb{J}_{NT}(\eta_0)B^{*-1}_{NT}B^{*-1\prime}_{NT}\mathbb{J}_{NT}(\eta_0) = \mathbb{J}_{NT}(\eta_0)$. According to the Markov inequality, it follows that $1_N\left\|[\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}/NT]^{-\frac{1}{2}}\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\varepsilon/NT\right\| = O_P(\sqrt{K}/\sqrt{N})$.

Hence, we have that

$$\begin{aligned}
&1_N\left\|[\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}/NT]^{-1}\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\varepsilon/NT\right\| \\
\le&O_P(1)1_N\left\|[\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}/NT]^{-\frac{1}{2}}\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\varepsilon/NT\right\| \\
=&O_P(\sqrt{K}/\sqrt{N}).
\end{aligned} \tag{A10}$$

Based on (A8)–(A10), the formula (A7) can be written as

$$\|\hat{\gamma} - \gamma_0\| = O_P(\sqrt{K}/\sqrt{N}) + O_P(K^{-r_2}). \tag{A11}$$

By Assumption 3, (A11) and Theorem 2, it is easy to obtain that

$$\begin{aligned}
1_N|\hat{\alpha}(u_{it}) - \alpha_0(u_{it})| &= 1_N|\zeta_{q,K}(u_{it})(\hat{\gamma} - \gamma_0) + (\zeta_{q,K}(u_{it})\gamma_0 - \alpha_0(u_{it}))| \\
&\le 1_N|\zeta_{q,K}(u_{it})(\hat{\gamma} - \gamma_0)| + |(\zeta_{q,K}(u_{it})\gamma_0 - \alpha_0(u_{it}))| \\
&= O_P\left(\zeta_0(K)\left(\sqrt{K}/\sqrt{N} + K^{-r_2}\right)\right).
\end{aligned}$$

□

Proof of Theorem 4. According to (A7), we have

$$\hat{\gamma} - \gamma_0 = [\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}]^{-1}\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)B^{*-1}_{NT}e + O_P\left(\sqrt{1/N} + K^{-r_2}\right).$$

Denote $\alpha^*(u) = \zeta_{q,K}(u)\gamma_0$, we know

$$\hat{\alpha}(u) - \alpha^*(u) = \zeta_{q,K}(u)[\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}]^{-1}\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)B_{NT}^{*-1}e + O_P\left(\|\zeta_{q,K}(u)\|\left(\sqrt{1/N} + K^{-r_2}\right)\right).$$

For any fixed point $u \in (a, d)$, as $N \to \infty$, applying the central limit theorem, we can obtain that

$$\Lambda_u^{-1/2}(\hat{\alpha}(u) - \alpha^*(u)) \xrightarrow{L} \mathrm{N}(0, \sigma_{e0}^2),$$

where $\Lambda_u = \zeta_{q,K}(u)[\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}]^{-1}\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)B_{NT}^{*-1}B_{NT}^{*\prime-1}\,\mathbb{J}'_{NT}(\eta_0)\widetilde{Z}[\widetilde{Z}'\mathbb{J}_{NT}(\eta_0)\widetilde{Z}]^{-1}\zeta'_{q,K}(u)$. This completes the proof of Theorem 4. □

References

1. Li, S.; Chen, J.; Li, B. Estimation and testing of random effects semiparametric regression model with separable space-time filters. *Fractal Fract.* **2022**, *6*, 735. [CrossRef]
2. Hsiao, C. Benefits and limitations of panel data. *Economet. Rev.* **1985**, *4*, 121–174. [CrossRef]
3. Baltagi, B.H. Panel data methods. In *Handbook of Applied Economic Statistics*; CRC Press: Boca Raton, FL, USA, 1998; pp. 311–323.
4. Porter, C.O.; Outlaw, R.; Gale, J.P.; Cho, T.S. The use of online panel data in management research: A review and recommendations. *J. Manag.* **2019**, *45*, 319–344. [CrossRef]
5. Zamanzade, E. EDF-based tests of exponentiality in pair ranked set sampling. *Stat. Pap.* **2019**, *60*, 2141–2159. [CrossRef]
6. Imai, K.; Kim, I.S. On the use of two-way fixed effects regression models for causal inference with panel data. *Polit. Anal.* **2021**, *29*, 405–415. [CrossRef]
7. Fan, J.; Gijbels, I. *Local Polynomial Modelling and Its Applications*; Chapman and Hall: London, UK, 1996.
8. Luo, G.; Wu, M.; Pang, Z. Empirical likelihood inference for the semiparametric varying-coefficient spatial autoregressive model. *J. Syst. Sci. Complex.* **2021**, *34*, 2310–2333. [CrossRef]
9. Ullah, A.; Wang, T.; Yao, W. Semiparametric partially linear varying coefficient modal regression. *J. Econom.* **2022**. [CrossRef]
10. Dai, X.; Li, S.; Jin, L.; Tian, M. Quantile regression for partially linear varying coefficient spatial autoregressive models. *Commun. Stat-Simul.* **2022**, *1–16*. [CrossRef]
11. Li, Q.; Stengos, T. Semiparametric estimation of partially linear panel data models. *J. Econom.* **1996**, *71*, 389–397. [CrossRef]
12. Baltagi, B.H.; Li, D. Series estimation of partially linear panel data models with fixed effects. *Ann. Econ. Financ.* **2002**, *3*, 103–116.
13. Chen, J.; Gao, J.; Li, D. Estimation in partially linear single-index panel data models with fixed effects. *J. Bus. Econ. Stat.* **2013**, *31*, 315–330. [CrossRef]
14. Huang, B.; Sun, Y.; Wang, S. Estimation of partially linear panel data models with cross-sectional dependence. *J. Syst. Sci. Complex.* **2021**, *34*, 2219–2230. [CrossRef]
15. Yong, A.; Cheng, H.; Dong, L. Semiparametric estimation of partially linear varying coefficient panel data models. *Adv. Econom.* **2016**, *36*, 47–65.
16. Zhou, B.; You, J.; Xu, Q.; Chen, G. Weighted profile least squares estimation for a panel data varying-coefficient partially linear model. *Chin. Ann. Math. B* **2010**, *31*, 247–272. [CrossRef]
17. Zhang, Y.; Shen, D. Estimation of semi-parametric varying-coefficient spatial panel data models with random-effects. *J. Stat. Plan. Infer.* **2015**, *159*, 64–80. [CrossRef]
18. Li, S.; Chen, J.; Chen, D. PQMLE of a partially linear varying coefficient spatial autoregressive panel model with random effects. *Symmetry* **2021**, *13*, 2057. [CrossRef]
19. Su, L.; Ullah, A. Profile likelihood estimation of partially linear panel data models with fixed effects. *Econ. Lett.* **2006**, *92*, 75–81. [CrossRef]
20. Wu, Q.; Luo, X.; Li, Y. Partially linear varying-coefficient panel data models with fixed effects. *Far East J. Theor. Stat.* **2008**, *25*, 229–238.
21. Hu, X. Estimation in a semi-varying coefficient model for panel data with fixed effects. *J. Syst. Sci. Complex.* **2014**, *27*, 594–604. [CrossRef]
22. Tran, K.C.; Tsionas, E.G. Local GMM estimation of semiparametric panel data with smooth coefficient models. *Economet. Rev.* **2009**, *29*, 39–61. [CrossRef]
23. Su, L.; Ullah, A. Nonparametric and semiparametric panel econometric models: Estimation and testing. In *Handbook of Empirical Economics and Finance*; CRC Press: Boca Raton, FL, USA, 2011; pp. 455–497.
24. Liu, Y.; Zhuang, X. Shrinkage estimation of semi-parametric spatial autoregressive panel data model with fixed effects. *Stat. Probabil. Lett.* **2023**, *194*, 109746. [CrossRef]
25. Elhorst, J.P. Serial and spatial error correlation. *Econ. Lett.* **2008**, *99*, 422–424. [CrossRef]
26. Baltagi, B.H.; Song, S.H.; Jung, B.C.; Koh, W. Testing for serial correlation, spatial autocorrelation and random effects using panel data. *J. Econ.* **2007**, *140*, 5–51. [CrossRef]
27. Parent, O.; LeSage, J.P. A space-time filter for panel data models containing random effects. *Comput. Stat. Data Anal.* **2011**, *55*, 475–490. [CrossRef]

28. Lee, L.; Yu, J. Estimation of fixed effects panel regression models with separable and nonseparable space-time filters. *J. Econ.* **2015**, *184*, 174–192. [CrossRef]
29. Bai, Y.; Hu, J.; You, J. Panel data partially linear varying-coefficient model with both spatially and time-wise correlated errors. *Stat. Sin.* **2015**, *35*, 275–294. [CrossRef]
30. Zhao, J.Q.; Zhao, Y.; Lin, J.; Miao, Z.; Khaled, W. Estimation and testing for panel data partially linear single-index models with errors correlated in space and time. *Random Matrices Theory Appl.* **2020**, *9*, 2150005. [CrossRef]
31. De Boor, C. *A Practical Guide to Splines*; Springer: New York, NY, USA, 1978.
32. Huang, J.Z.; Shen, H. Functional coefficient regression models for non-linear time series: A polynomial spline approach. *Scand. J. Stat.* **2004**, *31*, 515–534. [CrossRef]
33. Zou, Q.; Zhu, Z. M-estimators for single-index model using B-spline. *Metrika* **2014**, *77*, 225–246. [CrossRef]
34. Hsiao, C.; Pesaran, H.; Tahmiscioglu, K. Maximum likelihood estimation of fixed effects dynamic panel data models covering short time periods. *J. Econ.* **2002**, *109*, 107–150. [CrossRef]
35. Su, L.; Jin, S. Profile quasi-maximum likelihood estimation of partially linear spatial autoregressive models. *J. Econ.* **2010**, *157*, 18–33. [CrossRef]
36. Kelejian, H.H.; Prucha, I.R. A generalized spatial two-stage least squares procedure for estimation a spatial autoregressive model with autoregressive disturbances. *J. Real Estate Financ. Econ.* **1998**, *17*, 99–121. [CrossRef]
37. Kelejian, H.H.; Prucha, I.R. A generalized moments estimator for the autoregressive parameter in a spatial model. *Int. ECcon. Rev.* **1999**, *40*, 509–533. [CrossRef]
38. Cheng, S.; Chen, J. Estimation of partially linear single-index spatial autoregressive model. *Stat. Pap.* **2021**, *62*, 495–531. [CrossRef]
39. Hu, J.; Liu, F.; You, J. Panel data partially linear model with fixed effects, spatial autoregressive error components and unspecified intertemporal correlation. *J. Multivar. Anal.* **2014**, *130*, 64–89. [CrossRef]
40. Newey, W.K. Convergence rates and asymptotic normality for series estimators. *J. Econ.* **1997**, *79*, 147–168. [CrossRef]
41. Zhang, Y. Estimation of partially specified spatial panel data models with random-effects and spatially correlated error components. *Commun. Stat.-Theory Methods* **2017**, *46*, 1056–1079. [CrossRef]
42. Horn, R.; Johnson, C. *Matrix Algebra*; Cambridge University Press: Cambridge, UK, 1985.
43. Lee, L. Asymptotic distributions of quasi-maximum likelihood estimators for spatial econometric models. *Econometrica* **2004**, *72*, 1899–1925. [CrossRef]
44. Su, L. Semiparametric GMM estimation of spatial autoregressive models. *J. Econ.* **2012**, *167*, 543–560. [CrossRef]
45. Ding, F.; Chen, J. Fast efficient estimators of partially linear varying coefficient panel model with fixed effects. *Chin. J. Appl. Probab. Statist.* **2019**, *36*, 11.
46. Ding, F. Penalized empirical likelihood estimation for partially linear single index panel model with fixed effects. *Chin. J. Appl. Probab. Statist.* **2019**, *35*, 573–593.
47. Yu, J.; De Jong, R.; Lee, L. Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and T are large. *J. Econ.* **2008**, *146*, 118–134. [CrossRef]
48. Kelejian, H.H.; Prucha, I.R. On the asymptotic distribution of the Moran I test statistic with applications. *J. Econ.* **2001**, *104*, 219–257. [CrossRef]