



Article

Estimation and Testing of Random Effects Semiparametric Regression Model with Separable Space-Time Filters

Shuangshuang Li, Jianbao Chen * and Bogui Li

College of Mathematics and Statistics, Fujian Normal University, Fuzhou 350117, China

* Correspondence: jbchen@fjnu.edu.cn

Abstract: This paper focuses on studying a random effects semiparametric regression model (RESPRM) with separable space-time filters. The model cannot only capture the linearity and nonlinearity existing in a space-time dataset, but also avoid the inefficient estimators caused by ignoring spatial correlation and serial correlation in the error term of a space-time data regression model. Its profile quasi-maximum likelihood estimators (PQMLE) for parameters and nonparametric functions, and a generalized F-test statistic for checking the existence of nonlinear relationships are constructed. The asymptotic properties of estimators and asymptotic distribution of test statistic are derived. Monte Carlo simulations imply that our estimators and test statistic have good finite sample performance. The Indonesian rice farming data are used to illustrate our methods.

Keywords: RESPRM with separable space-time filters; PQMLE; generalized F-test statistic; asymptotic property; Monte Carlo simulation

**Citation:** 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. <https://doi.org/10.3390/fractalfract6120735>

Academic Editors: Viorel-Puiu Paun, Francesco Vilecco and Wanqing Song

Received: 25 October 2022

Accepted: 6 December 2022

Published: 11 December 2022

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



Copyright: © 2022 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

A space-time panel dataset is a sample collected from a number of spatial units over time periods. Such data widely exist in various research fields. It is of great theoretical significance and practical value to conduct statistical modeling and analysis on space-time panel data. There are three branches of panel data which are defined as ordinary panel data, spatial panel data, and space-time panel data given the variety heterogeneities of existing panel datasets. The regression models established on basis of ordinary panel data are called ordinary panel data regression models. Ordinary panel data regression models date back to the 1950s when the pioneering paper (Bates et al. [1]) was published. After that, the theories and methods of ordinary panel data regression models have been greatly enriched and they are widely available in many fields, such as economics, management and environmental science, see Chamberlain [2], Baltagi [3], Arellano [4], Baldev and Baltagi [5], and Hsiao [6], among others. Panel data spatial regression models generated from the turn of this century when spatial econometrics literature has exhibited a particular interest in the specification and estimation of econometric relationships based on spatial panel data (Elhorst [7]). Panel data spatial regression models deal with the spatial correlations between different individual units in panel data models and their theories have been well developed and widely used, see Druska and Horrace [8]; Egger, Pfaffermayr, and Winner [9]; Kapoor, Kelejjan, and Prucha [10]; Lee and Yu [11], etc.

Two problems hampering when modeling with space-time panel data are serial correlation and spatial correlation. Serial correlation lies between the observations of each spatial unit over time, and it is the domain of the voluminous time-series econometrics literature. Spatial correlation lies between the observations of the spatial units at each time period, and it is of the spatial econometrics literature, a subfield of econometrics for dealing with spatial interaction effects among geographical units such as individuals, firms, governments, etc. In order to overcome the defects that the panel data spatial regression models do not account for serial correlation and the ordinary panel data regression models

account for neither spatial correlation nor serial correlation, people have tried to combine serial correlation and spatial correlations, and established panel data parametric regression models with separable/nonseparable space-time filters, which are used to do researches on estimation, testing and empirical analysis. Elhorst [12] developed ML method of a panel data linear regression model with nonseparable space-time filters, while did not establish asymptotic properties of the estimators. Parent and LeSage [13] explored Markov Chain Monte Carlo methods of random effects panel data linear regression models with separable space-time filters, and the performance of the method was demonstrated by both Monte Carlo simulations and an applied illustration. Lee and Yu [14] studied MLE and its asymptotic properties of fixed and random effects panel data parametric regression models with separable space-time filters. Lee and Yu [15] provided QMLE and its asymptotic properties for a fixed effects panel data linear regression model with disturbances contained both separable space-time filters and nonseparable space-time filters. Baltagi et al. [16] derived several Lagrange Multiplier tests for panel data linear regression model with separable space-time filters containing random effects. Cohen and Paul [17] analysed the influencing factors of public infrastructure investment on the costs and productivity of private enterprises on basis of 1982–1996 state-level U.S. manufacturing data by panel data parametric regression model with separable space-time filters.

All literature on panel data regression models with separable/nonseparable space-time filters mentioned above focuses on panel data parametric regression models. Although the theories and applications of these models have been well developed, they are often unrealistic in real situations for the reason that they fail to capture complex structure (e.g., nonlinearity) owing to lacking of flexibility. Moreover, the form of parametric regression models may be misspecified, and estimators based on misspecified models are able to cause inconsistency and even erroneous conclusions. Driven by these reasons, Zhao et al. [18] constructed semiparametric minimum average variance estimation method, and proposed a F-test statistic of partially linear single-index panel regression model with separable space-time filters, then proved the asymptotic properties of estimators and test statistic. Bai, Hu, and You [19] established weighted semiparametric least squares and weighted polynomial spline series estimation method for parametric and nonparametric component, respectively, of panel data partially linear varying-coefficient regression model with separable space-time filters, and then proved their asymptotic normalities.

In this paper, we study estimation and testing of random effects semiparametric regression model (RESPRM) with separable space-time filters. By allowing a nonparametric component in parametric regression model with separable space-time filters, it can simultaneously capture the linear and nonlinear effects of covariates, spatial correlation of error structure, serial correlation of remainder error structure, and individual random effects. To the best of our knowledge, there is no related literature on this model. In this paper, we aim to study its profile quasi-maximum likelihood estimation (PQMLE) and hypothesis test methods, and then conduct systematic studies of the asymptotic properties and small sample performance for estimators and test statistic. Furthermore, we illustrate the proposed estimation and testing methods by using a real dataset.

The remainder of this paper is organized as follows. Section 2 introduces the RESPRM with separable space-time filters, and estimators for the model and test statistic for nonparametric component are constructed. Section 3 mainly provides asymptotic properties of estimators, asymptotic distribution of test statistic and conditional assumptions. Section 4 presents the finite sample performance of estimates and F-test statistic through Monte Carlo simulations. Section 5 illustrates the proposed methods by an application of Indonesian rice farming data. Conclusions are summarized in Section 6. The proofs of some important lemmas and theorems are given in Appendix A.

2. The Model, Estimation, and Testing

2.1. The Model

The RESPRM with separable space-time filters can be specified as:

$$y_{it} = x'_{it}\beta + m(z_{it}) + b_i + \varepsilon_{it}, \tag{1}$$

$$\varepsilon_{it} = \rho \sum_{j=1}^N w_{ij}\varepsilon_{jt} + v_{it}, \tag{2}$$

$$v_{it} = \lambda v_{i,t-1} + e_{it}, \quad 1 \leq i \leq N, \quad 1 < t \leq T, \tag{3}$$

where y_{it} is an observation of the response variable for the i th spatial unit at the t th time period, $x_{it} = (x_{it1}, x_{it2}, \dots, x_{itp})'$, and $z_{it} = (z_{it1}, z_{it2}, \dots, z_{itq})'$ are observations of $p \times 1$ and $q \times 1$ vectors of covariates, respectively, b_i is an individual random effect which is assumed to be $IID(0, \sigma_b^2)$, ε_{it} is an error term, v_{it} is a remainder error term which is weakly stationary sequence, e_{it} is a random error term which is assumed to be $IID(0, \sigma_e^2)$, $\beta = (\beta_1, \beta_2, \dots, \beta_p)'$ is a regression coefficient vector of x_{it} , $m(\cdot)$ is an unknown link function, w_{ij} is the (i, j) th element of spatial weights matrix, ρ and λ are spatial correlation and series correlation coefficients, respectively. Furthermore, we denote $\theta_0 = (\zeta'_0, \beta'_0)'$, $\zeta_0 = (\rho_0, \lambda_0, \sigma_{b0}^2, \sigma_{e0}^2)$ as the true parameter vector of $\theta = (\zeta', \beta)'$ and $\zeta = (\rho, \lambda, \sigma_b^2, \sigma_e^2)$, respectively, and $m_0(z)$ as the true value of nonparametric function of $m(z)$.

We find that separable space-time filters in our model result in a trade-off between the spatial correlation and serial correlation coefficients. According to Shi and Lee [20], it is not hard to obtain that the stationarity requires not only $|\rho| < 1, |\lambda| < 1$, but also $|\lambda| + |\rho| + |\rho\lambda| < 1$, i.e., parameter space Λ of λ and ρ , see Figure 1.

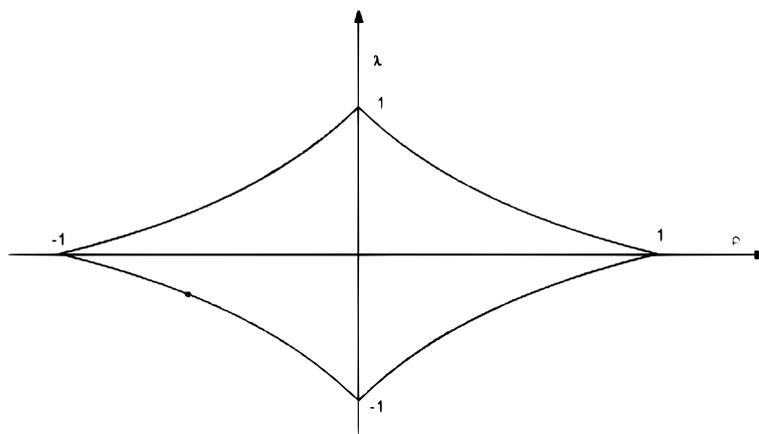


Figure 1. The parameter space Λ of λ and ρ .

2.2. Estimation

Because that random effect b_i could be regarded as time-invariant permanent spillovers (see Baltagi, Fingleton, and Pirotte [21]), the model (1) can be written in matrix form

$$Y = X\beta + M + \eta, \tag{4}$$

where $Y = (y_{11}, \dots, y_{NT})'$, $X = (x_{11}, \dots, x_{NT})'$, $M = (m(z_{11}), \dots, m(z_{NT}))'$, $b = (b_1, \dots, b_N)'$, $v = (v_{11}, \dots, v_{NT})'$, $\eta = l_T \otimes b + (I_T \otimes B^{-1})v$, l_T is a $T \times 1$ vector of ones, $B = I_N - \rho W$, I_N is an $N \times N$ identity matrix, W is an $N \times N$ spatial weights matrix. Then we obtain the variance matrix of transformed error structure η as follows

$$\Omega = E(\eta\eta') = \sigma_b^2(l_T l_T' \otimes I_N) + \sigma_e^2[V_\lambda \otimes (B^{-1} B^{-1}')],$$

$$\Omega^{-1} = \frac{1}{d^2(1 - \lambda^2)} V_\lambda^{-1} l_T l_T' V_\lambda^{-1} \otimes Z_0 + [V_\lambda^{-1} - \frac{1}{d^2(1 - \lambda^2)} V_\lambda^{-1} l_T l_T' V_\lambda^{-1}] \otimes (\frac{1}{\sigma_b^2} B B'),$$

$$|\Omega| = |V_\lambda|^N |\sigma_e^2 B^{-1} B'^{-1}|^{T-1} |Z_0^{-1}|,$$

where

$$V_\lambda = \frac{1}{1-\lambda^2} \begin{pmatrix} 1 & \lambda & \dots & \lambda^{T-1} \\ \lambda & 1 & \dots & \lambda^{T-2} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda^{T-1} & \lambda^{T-2} & \dots & 1 \end{pmatrix},$$

$Z_0 = [d^2(1-\lambda^2)\sigma_b^2 I_N + \sigma_e^2 B^{-1} B'^{-1}]^{-1}$, $d^2 = l_T^\lambda l_T^\lambda = \frac{1+\lambda}{1-\lambda} + (T-1)$ and $l_T^\lambda = (\frac{\sqrt{1+\lambda}}{\sqrt{1-\lambda}}, 1, \dots, 1)'$. Therefore, the log-likelihood function of the new transformed model (4) is defined in (5)

$$\ln L(\theta) = -\frac{NT}{2} \ln(2\pi) - \frac{1}{2} \ln|\Omega| - \frac{1}{2} \zeta'(\theta) \Omega^{-1} \zeta(\theta), \quad (5)$$

where $\zeta(\theta) = Y - M - X\beta$.

It is difficult to obtain the QMLE by maximizing (5) because that M^* is unknown. Thus, we will combine PQML method and Working Independence theory (Cai [22]) to estimate unknown parameters and nonparametric function of the model. Working independence assumption is often used to deal with nonparametric/semiparametric regression models which exist correlation structures. By ignoring the correlation structure entirely, and pretending as if the data within a cluster were independent, the asymptotically most efficient estimator of nonparametric function is obtained (Lin and Carroll [23]). It has been widely used in various correlation structures of nonparametric/semiparametric models, such as Fan et al. [24], Fan and Zhang [25], and Tang and Liu [26], among others.

The estimation steps are as follows:

Step 1. By supposing that β is known, the initial estimator of $m(z)$ can be obtained. Let

$$K_H(z) = \text{diag}(k_H(z_{11}-z), \dots, k_H(z_{1T}-z), \dots, k_H(z_{N1}-z), \dots, k_H(z_{NT}-z)),$$

$$\vec{Z}(z) = (Z_{11}(z), \dots, Z_{1T}(z), \dots, Z_{N1}(z), \dots, Z_{NT}(z))',$$

where $k_H(z) = |H|^{-1} k(H^{-1}z)$, $k(\cdot)$ is $q \times 1$ kernel function, $|H|$ is the determinant of $H = \text{diag}(h_1, \dots, h_q)$, $Z_{it}(z) = [1, H^{-1}(z_{it}-z)]'$. In addition, denote $\hat{m}_{IN}(z) = \hat{a}_1$, then we have $\hat{\delta} = (\hat{a}_1, (H\hat{a}_2)')$ and

$$\hat{\delta} = \arg \min_{a_1, a_2} \frac{1}{NT} [\tilde{Y} - \vec{Z}(z)\delta]' K_H(z) [\tilde{Y} - \vec{Z}(z)\delta],$$

where $\tilde{Y} = (\tilde{y}_{11}, \dots, \tilde{y}_{1T}, \dots, \tilde{y}_{N1}, \dots, \tilde{y}_{NT})'$, $\tilde{y}_{it} = y_{it} - x'_{it}\beta$, a_2 is the first derivative of $m(z)$. Let $S(z) = [\vec{Z}'(z)K_H(z)\vec{Z}(z)]^{-1}\vec{Z}'(z)K_H(z)$, then $\hat{\delta} = S(z)\tilde{Y}$ and

$$\hat{m}_{IN}(z) = s'(z)\tilde{Y}, \quad (6)$$

where $s(z) = e'_0 S(z)$, $e_0 = (1, 0, \dots, 0)'$ is a $(q+1) \times 1$ vector.

Step 2. Substituting (6) into (5), the approximate log-likelihood function is given by

$$\ln L(\theta, \hat{m}_{IN}(z)) = -\frac{NT}{2} \ln(2\pi) - \frac{1}{2} \ln|\Omega| - \frac{1}{2} (Y - X\beta)' (I - S)' \Omega^{-1} (I - S) (Y - X\beta). \quad (7)$$

By maximizing (7), the initial estimators of β can be obtained as follows

$$\hat{\beta}_{IN} = [X'(I - S)' \Omega^{-1} (I - S) X]^{-1} X' (I - S)' \Omega^{-1} (I - S) Y, \quad (8)$$

Step 3. Substituting $\hat{\beta}_{IN}$ into (7), we get the concentrated log-likelihood function of ρ and λ as

$$\hat{\zeta} = \arg \max_{\zeta} \ln L(\zeta). \quad (9)$$

Step 4. Substituting (9) into (8), the ultimate estimator of $\hat{\beta}$ is obtained, and further substituting $\hat{\beta}$ into (6), the ultimate estimator of $\hat{m}(z)$ is obtained.

2.3. Testing

By allowing nonparametric component in a parametric regression model with separable space-time filters, it provides us the ability to explore potential nonlinear relationship lies between covariates and response variable. In order to check whether the parametric specification is correct, it is necessary to do the following hypothesis test

$$H_0 : m(z_{it}) = \gamma_0 + \sum_{j=1}^q \gamma_j z_{itj} \quad \leftrightarrow \quad H_1 : m(z_{it}) \neq \gamma_0 + \sum_{j=1}^q \gamma_j z_{itj},$$

where $\gamma_0, \dots, \gamma_q$ are unknown parameters. Following Fan, Zhang, and Zhang [27], we obtain $\hat{m}(z_{it})$ by PQML method, and $\tilde{m}(z_{it}) = \tilde{\gamma}_0 + \sum_{j=1}^q \tilde{\gamma}_j z_{itj}$ is OLS estimator of the intercept and slopes in the linear regression of $y_{it} - x'_{it}\hat{\beta}$, $\hat{\beta}$ is also obtained by PQML method. The resulting residual sums of squares under the null and alternative hypotheses, respectively, are as follows

$$RSS(H_0) = \sum_i^N \sum_{t=1}^T [y_{it} - \tilde{m}(z_{it}) - x'_{it}\hat{\beta}]^2,$$

$$RSS(H_1) = \sum_i^N \sum_{t=1}^T [y_{it} - \hat{m}(z_{it}) - x'_{it}\hat{\beta}]^2.$$

To test the null hypothesis, we define generalized F-test statistic as follows:

$$F_{NT} = \frac{NT}{2} \frac{RSS(H_0) - RSS(H_1)}{RSS(H_1)}. \quad (10)$$

The asymptotic distribution of F_{NT} is given in Section 3.

3. Asymptotic Property

To prove the asymptotic properties of PQMLE and asymptotic distribution of generalized F-test statistic proposed in Section 2, we need to establish some useful assumptions and lemmas.

3.1. Assumptions

In order to provide a rigorous theoretical analysis, we make the following assumptions.

Assumption 1. (i) $\{x_i, z_i\}_{i=1, t=1}^{N, T}$ are nonstochastic variables and uniformly bounded (UB) in $\mathcal{X} \times \mathcal{Z}$.

(ii) There is a positive density function $f_t(\cdot)$ such that $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N v(z_i) = \int v(z) f_t(z) dz$ for any bounded continuous function $v(\cdot)$, and $f_t(\cdot)$ is bounded away from zero on \mathcal{Z} .

(iii) e_{it} are iid and $E(|e_{it}|^4) < \infty$.

Assumption 2. (i) W is a row normalized and prespecified spatial weights matrix.

(ii) B is nonsingular.

(iii) B and B^{-1} are UB in both row and column sums in absolute value.

Assumption 3. As $N \rightarrow \infty$, $\|H\| \rightarrow 0$, $N|H|^2 \rightarrow \infty$, $\|H\|^4 |H|^{-1} \rightarrow 0$ and $N|H| \|H\|^4 \rightarrow c \in [0, \infty)$, where $\|H\| = \text{tr}(H'H)^{\frac{1}{2}} = \{\sum_{ij} |h_{ij}|^2\}^{\frac{1}{2}}$ is the Euclidean norm of matrix H , and $|H|$ is the determinant of matrix H .

Assumption 4. The kernel function $K(\cdot)$ is a continuous density with compact support on R^q , all odd order moments of K vanish. Further, $K(\cdot)$ is bounded, $u^3K(u)$ and $u^3K'(u)$ are bounded and $\int u^4K(u)du < \infty$.

Assumption 5. For any $(\rho, \lambda) \in \Lambda$, we require that

$$\lim_{N, T \rightarrow \infty} \left[\frac{1}{NT} \ln |\Omega_0| + 1 - \frac{1}{NT} \ln |\Omega| - \text{tr}(\Omega^{-1} \Omega_0) \right] \neq 0,$$

where

$$|\Omega_0| = |V_{\lambda_0}|^N |\sigma_{\varepsilon_0}^2 B_0^{-1} B_0'^{-1}|^{T-1} |Z_{00}^{-1}|.$$

$$V_{\lambda_0} = \frac{1}{1 - \lambda_0^2} \begin{pmatrix} 1 & \lambda_0 & \dots & \lambda_0^{T-1} \\ \lambda_0 & 1 & \dots & \lambda_0^{T-2} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_0^{T-1} & \lambda_0^{T-2} & \dots & 1 \end{pmatrix},$$

$$Z_{00} = [d^2(1 - \lambda_0^2)\sigma_{\varepsilon_0}^2 I_N + \sigma_{\varepsilon_0}^2 B_0^{-1} B_0'^{-1}]^{-1}, \quad d_0^2 = l_T^{\lambda_0} l_T^{\lambda_0} = \frac{1 + \lambda_0}{1 - \lambda_0} + (T - 1) \text{ and}$$

$$l_T^{\lambda_0} = \left(\frac{\sqrt{1 + \lambda_0}}{\sqrt{1 - \lambda_0}}, 1, \dots, 1 \right)'.$$

Assumption 6. $L_{11}(L_{44}L_{22} - L_{24}^2) - L_{14}(L_{14}L_{22} + 2L_{12}L_{24}) > 0$, where $L_{11} = \frac{2\text{tr}[(B^{-1}(\rho_0)W)^2]}{N}$, $L_{44} = \frac{1}{2\sigma_{\varepsilon_0}^4}$, $L_{22} = -\frac{\text{tr}(\ddot{C}_0 C_0^{-1}) + \text{tr}(C_0^{-1} \dot{C}_0' \dot{C}_0 C_0^{-1})}{T} + \frac{1 - 3\lambda_0^2}{T(1 - \lambda_0^2)}$, $L_{14} = \frac{\text{tr}(WB_0^{-1})}{N\sigma_{\varepsilon_0}^2}$, $L_{24} = \frac{\text{tr}(C_0^{-1} \dot{C}_0')}{T\sigma_{\varepsilon_0}^2}$, $L_{12} = -\frac{2}{NT} \text{tr}(C_0^{-1} \dot{C} \otimes WB_0^{-1})$.

Remark: Assumption 1 concerns the features of covariates, random error term, and density function for the model (1). Assumption 2 provides the basic features of spatial weight matrix, it parallels Assumption 3 of Su and Jin [28] and Assumption 5 of Lee [29]. The assumption that B and B^{-1} are UB limits the spatial correlation to a manageable degree. Assumption 3 concerns the bandwidth sequence, it parallels Assumption 4 of Hamilton and Truong [30], and Assumption 7 of Ullah and Su [31]. Assumption 4 concerns the kernel function. Assumptions 5 and 6 are necessary for the consistency and asymptotic properties of the estimators.

3.2. Asymptotic Properties

From the model (4), we obtain the reduced form equation of Y as

$$Y = X\beta_0 + M_0 + \eta_0. \quad (11)$$

Let $S = (s_{11}, \dots, s_{NT})$, where $s_{it} = s(z_{it})$. Denote a typical entry of $s(z)$ by $s(z_{it}, z) = e'[\bar{Z}'(z)K_H(z)\bar{Z}(z)]^{-1}\bar{Z}'_{it}(z)K_H(z_{it} - z)$, where $\bar{Z}'_{it}(z)$ is a typical column of $\bar{Z}'(z)$.

Lemma 1. Suppose that Assumptions 1–5 hold, then we have

$$\|S\| = O_p(N^{-1/2}|H|^{-1/2}).$$

The proof is given by Hamilton and Truong [30].

Lemma 2. Suppose that Assumptions 1–5 hold, then we have

- (i) $s(z_{it}, z) = N^{-1}K_H(z_{it} - z)\bar{f}(z)\{1 + o_p(1)\}$, where $\bar{f}(z) = \sum_{i=1}^T f_i(z)$.
- (ii) There is a positive constant c such that $\lim_{N \rightarrow \infty} P\{N^{-1}[\bar{Z}'(z)K_H(z)\bar{Z}(z)]_{ij} \leq c \text{ for } z \in c_f, i, j = 1, \dots, q + 1\} = 1$.

(iii) There exists a positive constant c such that

$$\lim_{N \rightarrow \infty} P\{\sup_z \max_{1 \leq i \leq N, 1 \leq t \leq T} |s(z_{it}, z)| \leq c(NT)^{-1}|H|^{-1}\} = 1.$$

The proofs are given by Ullah and Su [31].

Lemma 3. Suppose that Assumptions 1–5 hold, then we have

$$\|(I - S)M\| = Op(\|H\|^2).$$

The proof is given by Hamilton and Truong [30].

With the above lemmas, we then state main results as follows. Their detail proofs can be found in the Appendix A.

Theorem 1. Suppose that Assumptions 1–5 hold, then we have $\hat{\theta} - \theta_0 = o_p(1)$.

Theorem 2. Suppose that Assumptions 1–6 hold, then 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} ” means convergence in distribution, $\Sigma_{\theta_0} = -E[\frac{1}{NT} \frac{\partial^2 \ln L(\theta_0)}{\partial \theta \partial \theta'}]$ is the average Hessian matrix (information matrix when e and b obey normal distributions), $\Omega_{\theta} = E(\Omega_{\theta_{NT}})$ and $\Omega_{\theta_{NT}} = \frac{1}{\sqrt{NT}} \frac{\partial \ln L(\theta_0)}{\partial \theta} \cdot \frac{1}{\sqrt{NT}} \frac{\partial \ln L(\theta_0)}{\partial \theta'} + \frac{1}{NT} \frac{\partial^2 \ln L(\theta_0)}{\partial \theta \partial \theta'}$.

Theorem 3. Suppose that Assumptions 1–4 hold, then we have

$$\sqrt{N|H|} \left(\hat{\delta} - \delta - Q^{-1} \begin{pmatrix} \bar{f}(z) (\int uu'K(u)du H \ddot{m}(z) H) \\ 0 \end{pmatrix} \right) \xrightarrow{D} N(0, \sigma^2(\delta)),$$

where $\ddot{m}(z)$ is the second derivative of $m(z)$, $\sigma^2(\delta) = \sigma_e^2 \{S(z)D_I(C_0' C_0 \otimes B_0' B_0)^{-1} D_I' S(z)'\}$, $D_I = I - D_{\hat{\theta}}^{-1} \hat{U}(\hat{C} \otimes \hat{B}' \hat{B})(I - S)$ and $D_{\hat{\theta}} = \hat{U}'(I_T \otimes \hat{B}' \hat{B}) \hat{U}$, $Q = \bar{f}(z) \begin{pmatrix} 1 & 0' \\ 0 & \int uu'K(u)du \end{pmatrix}$, $\bar{f}(z) = \sum_{t=1}^T f_t(z)$.

Theorem 4. Suppose that Assumptions 1–4 hold, then we have

$$r_k F_{NT} \xrightarrow{D} \chi_{df}^2,$$

where $r_k = \frac{K(0) - \frac{1}{2} \int K^2(z) du}{\int (K(z) - \frac{1}{2} K^*(z))^2 du}$, $c_k = K(0) - \frac{1}{2} \int K^2(z) du$, $K(z) = \text{diag}(H^{-1}k(H^{-1}(z_{11} - z)), \dots, H^{-1}k(H^{-1}(z_{NT} - z)))$, $df = r_k c_k |\mathcal{Z}|/|H|$, \mathcal{Z} is the support of z_{it} and H is a bandwidth sequence, besides, note that when $K(0) = \max K(z)$, we get $K(0) > \|K\|_2^2$, $c_k > \frac{1}{2}K(0)$ and hence $r_k > 0$.

4. Monte Carlo Simulation

4.1. Performance of the Estimates

In this section, we report the results of a small-scale Monte Carlo study which is aimed to evaluate the finite sample performance of PQMLEs. We use mean square error (MSE), standard deviation (SD), and mean absolute deviation error (MADE) to measure the parametric and nonparametric estimation performance, respectively, where

$$\text{MADE}_j = Q^{-1} \sum_{q=1}^Q |\hat{m}(z_q) - m(z_q)|, \quad j = 1, 2, \dots, mcn,$$

$mcn = 2000$, $\{z_q\}_{q=1}^Q$ are Q fixed grid points in support set of z . Furthermore, we apply the rule of thumb method of Mack and Silverman [32] to choose the optimal width and let $K(u) = (\frac{3}{4\sqrt{5}})(1 - \frac{1}{5}u^2)1(u^2 \leq 5)$ (see Su [33]).

We use data generated from the following model

$$\begin{aligned}
 y_{it} &= x_{it}\beta + m(z_{it}) + b_i + \varepsilon_{it}, \\
 \varepsilon_{it} &= \rho \sum_{j=1}^N w_{ij}\varepsilon_{jt} + v_{it}, \\
 v_{it} &= \lambda v_{i,t-1} + e_{it},
 \end{aligned}
 \tag{12}$$

where $x_{it} \sim U[-2, 2]$, $z_{it} \sim U[-3, 3]$, $e_{it} \sim i.i.d.N(0, 1)$, $b_i \sim i.i.d.N(0, 1)$, $m(z_{it}) = 0.3z_{it}^2$, $\beta = 1$, $\rho = 0.2, 0.4$, and $\lambda = 0.2, 0.4$ respectively. In addition, we defined weights matrix as Rook matrix (see Anselin [34]). The simulation results of cases $N = 50, 100, 200$, and $T = 10, 20$ are presented in Tables 1–4.

It is obvious from Tables 1–4 that: (1) the *MSEs* of all parametric estimators are fairly small, which indicates the parametric estimators are approximately unbiased; (2) for fixed T , when N increases, the *MSE* of $\hat{\rho}$, $\hat{\lambda}$, $\hat{\beta}$, $\hat{\sigma}_b^2$ and $\hat{\sigma}_e^2$ decrease rapidly. For fixed N , when T increases, the behavior of parametric estimates is similar to the case where N changes with the fixed T . (3) The *SDs* and *MADs* of $\hat{m}(z)$ decrease as T or N increases. It presents that the nonparametric estimates are convergent.

Table 1. Results of parametric estimates under the model (12) with $T = 10$.

		N = 50			N = 100			N = 200		
		True	Mean	MSE	True	Mean	MSE	True	Mean	MSE
$\rho = 0.2$ $\lambda = 0.2$	$\hat{\beta}$	1.0000	0.9854	0.0070	1.0000	1.0096	0.0032	1.0000	1.0060	0.0018
	$\hat{\rho}$	0.2000	0.2034	0.0108	0.2000	0.2027	0.0089	0.2000	0.2017	0.0044
	$\hat{\lambda}$	0.2000	0.2079	0.0036	0.2000	0.2074	0.0032	0.2000	0.2027	0.0027
	$\hat{\sigma}_e^2$	1.0000	1.0335	0.0013	1.0000	1.0245	0.0011	1.0000	1.0064	0.0008
	$\hat{\sigma}_b^2$	1.0000	1.0356	0.0045	1.0000	1.0297	0.0025	1.0000	1.0160	0.0013
	$\rho = 0.2$ $\lambda = 0.4$	$\hat{\beta}$	1.0000	1.0130	0.0049	1.0000	1.0041	0.0041	1.0000	1.0004
$\hat{\rho}$		0.2000	0.2059	0.0113	0.2000	0.2018	0.0103	0.2000	0.2041	0.0010
$\hat{\lambda}$		0.4000	0.4120	0.0056	0.4000	0.4046	0.0038	0.4000	0.4039	0.0013
$\hat{\sigma}_e^2$		1.0000	1.0567	0.0088	1.0000	1.0302	0.0045	1.0000	1.0290	0.0043
$\hat{\sigma}_b^2$		1.0000	1.0536	0.0057	1.0000	1.0145	0.0033	1.0000	1.0022	0.0025
$\rho = 0.4$ $\lambda = 0.2$		$\hat{\beta}$	1.0000	1.0349	0.0782	1.0000	1.0311	0.0601	1.0000	0.9774
	$\hat{\rho}$	0.4000	0.3846	0.0344	0.4000	0.4026	0.0409	0.4000	0.3929	0.0311
	$\hat{\lambda}$	0.2000	0.2069	0.0018	0.2000	0.2055	0.0015	0.2000	0.2044	0.0013
	$\hat{\sigma}_e^2$	1.0000	1.0429	0.0027	1.0000	1.0159	0.0023	1.0000	1.0083	0.0017
	$\hat{\sigma}_b^2$	1.0000	1.0378	0.0025	1.0000	1.0083	0.0018	1.0000	1.0055	0.0013
	$\rho = 0.4$ $\lambda = 0.4$	$\hat{\beta}$	1.0000	1.0356	0.0637	1.0000	1.0116	0.0441	1.0000	1.0071
$\hat{\rho}$		0.4000	0.4083	0.0011	0.4000	0.4014	0.0010	0.4000	0.4008	0.0005
$\hat{\lambda}$		0.4000	0.4109	0.0008	0.4000	0.4107	0.0008	0.4000	0.4009	0.0006
$\hat{\sigma}_e^2$		1.0000	1.0411	0.0019	1.0000	1.0212	0.0005	1.0000	1.0041	0.0004
$\hat{\sigma}_b^2$		1.0000	1.0399	0.0021	1.0000	1.0138	0.0020	1.0000	1.0022	0.0008

Table 2. Results of parametric estimates under the model (12) with $T = 20$.

		$N = 50$			$N = 100$			$N = 200$		
		True	Mean	MSE	True	Mean	MSE	True	Mean	MSE
$\lambda = 0.2$ $\rho = 0.2$	$\hat{\beta}$	1.0000	1.0068	0.0061	1.0000	0.9932	0.0025	1.0000	0.9945	0.0010
	$\hat{\rho}$	0.2000	0.1976	0.0012	0.2000	0.1990	0.0010	0.2000	0.2003	0.0005
	$\hat{\lambda}$	0.2000	0.2051	0.0022	0.2000	0.2019	0.0015	0.2000	0.2008	0.0008
	$\hat{\sigma}_e^2$	1.0000	1.0102	0.0009	1.0000	1.0098	0.0007	1.0000	1.0005	0.0005
	$\hat{\sigma}_b^2$	1.0000	1.0242	0.0034	1.0000	1.0145	0.0013	1.0000	1.0008	0.0007
$\rho = 0.2$ $\lambda = 0.4$	$\hat{\beta}$	1.0000	0.9879	0.0025	1.0000	1.0006	0.0008	1.0000	1.0003	0.0003
	$\hat{\rho}$	0.2000	0.2028	0.0097	0.2000	0.2011	0.0059	0.2000	0.2001	0.0001
	$\hat{\lambda}$	0.4000	0.4050	0.0043	0.4000	0.3972	0.0019	0.4000	0.4008	0.0011
	$\hat{\sigma}_e^2$	1.0000	1.0082	0.0037	1.0000	1.0056	0.0021	1.0000	1.0014	0.0014
	$\hat{\sigma}_b^2$	1.0000	1.0062	0.0026	1.0000	1.0047	0.0024	1.0000	1.0016	0.0015
$\rho = 0.4$ $\lambda = 0.2$	$\hat{\beta}$	1.0000	1.0157	0.0369	1.0000	0.9915	0.0282	1.0000	0.9941	0.0008
	$\hat{\rho}$	0.4000	0.3971	0.0312	0.4000	0.4010	0.0033	0.4000	0.3990	0.0005
	$\hat{\lambda}$	0.2000	0.2025	0.0011	0.2000	0.2013	0.0007	0.2000	0.2001	0.0002
	$\hat{\sigma}_e^2$	1.0000	1.0096	0.0013	1.0000	1.0068	0.0011	1.0000	1.0041	0.0008
	$\hat{\sigma}_b^2$	1.0000	1.0071	0.0020	1.0000	1.0048	0.0015	1.0000	1.0005	0.0007
$\rho = 0.4$ $\lambda = 0.4$	$\hat{\beta}$	1.0000	1.0276	0.0063	1.0000	1.0041	0.0043	1.0000	1.0008	0.0003
	$\hat{\rho}$	0.4000	0.4005	0.0004	0.4000	0.4004	0.0001	0.4000	0.4000	0.0000
	$\hat{\lambda}$	0.4000	0.4004	0.0004	0.4000	0.4001	0.0001	0.4000	0.4000	0.0000
	$\hat{\sigma}_e^2$	1.0000	1.0025	0.0012	1.0000	1.0012	0.0004	1.0000	1.0004	0.0002
	$\hat{\sigma}_b^2$	1.0000	1.0009	0.0016	1.0000	1.0008	0.0009	1.0000	1.0002	0.0005

Table 3. The medians and SDs of MADEs for nonparametric estimates under the model (12) with $T = 10$.

			$N = 50$	$N = 100$	$N = 200$
$\rho = 0.2$	$\lambda = 0.2$	Median	0.3364	0.3248	0.3179
		SD	0.2750	0.2203	0.2145
	$\lambda = 0.4$	Median	0.3108	0.3100	0.2927
		SD	0.2709	0.2467	0.2247
$\rho = 0.4$	$\lambda = 0.2$	Median	0.3060	0.3046	0.2933
		SD	0.2464	0.2156	0.2077
	$\lambda = 0.4$	Median	0.3275	0.3211	0.3031
		SD	0.2540	0.1924	0.1756

Table 4. The medians and SDs of MADEs for nonparametric estimates under the model (12) with $T = 20$.

			$N = 50$	$N = 100$	$N = 200$
$\rho = 0.2$	$\lambda = 0.2$	Median	0.3355	0.3174	0.2687
		SD	0.2429	0.1948	0.1530
	$\lambda = 0.4$	Median	0.3042	0.2902	0.2883
		SD	0.2113	0.1883	0.1858
$\rho = 0.4$	$\lambda = 0.2$	Median	0.3020	0.3007	0.2903
		SD	0.2225	0.2069	0.1908
	$\lambda = 0.4$	Median	0.3244	0.3001	0.2873
		SD	0.2040	0.1824	0.1722

Figure 2 shows the fitting image and 95% confidence intervals of $\hat{m}(z)$ under $\rho = 0.2$ and $\lambda = 0.2$, respectively, where the black short dashed line is the average fit over 2000

simulations $\hat{m}(z)$ by PQML method, the red solid line is the true values of $m(z)$, and the two black long dashed lines are the corresponding 95% confidence bands. By observing subgraphs of Figure 2, we know that shorter dotted line is fairly close to solid line and confidence band is narrow. These mean that nonparametric estimation procedure performances well under finite sample. To save space, we do not show the others cases under $\rho = 0.2$ with $\lambda = 0.4$, $\rho = 0.4$ with $\lambda = 0.2$ and $\rho = 0.4$ with $\lambda = 0.4$ of different N and T because they are similar with the case $\rho = 0.2$ with $\lambda = 0.2$.

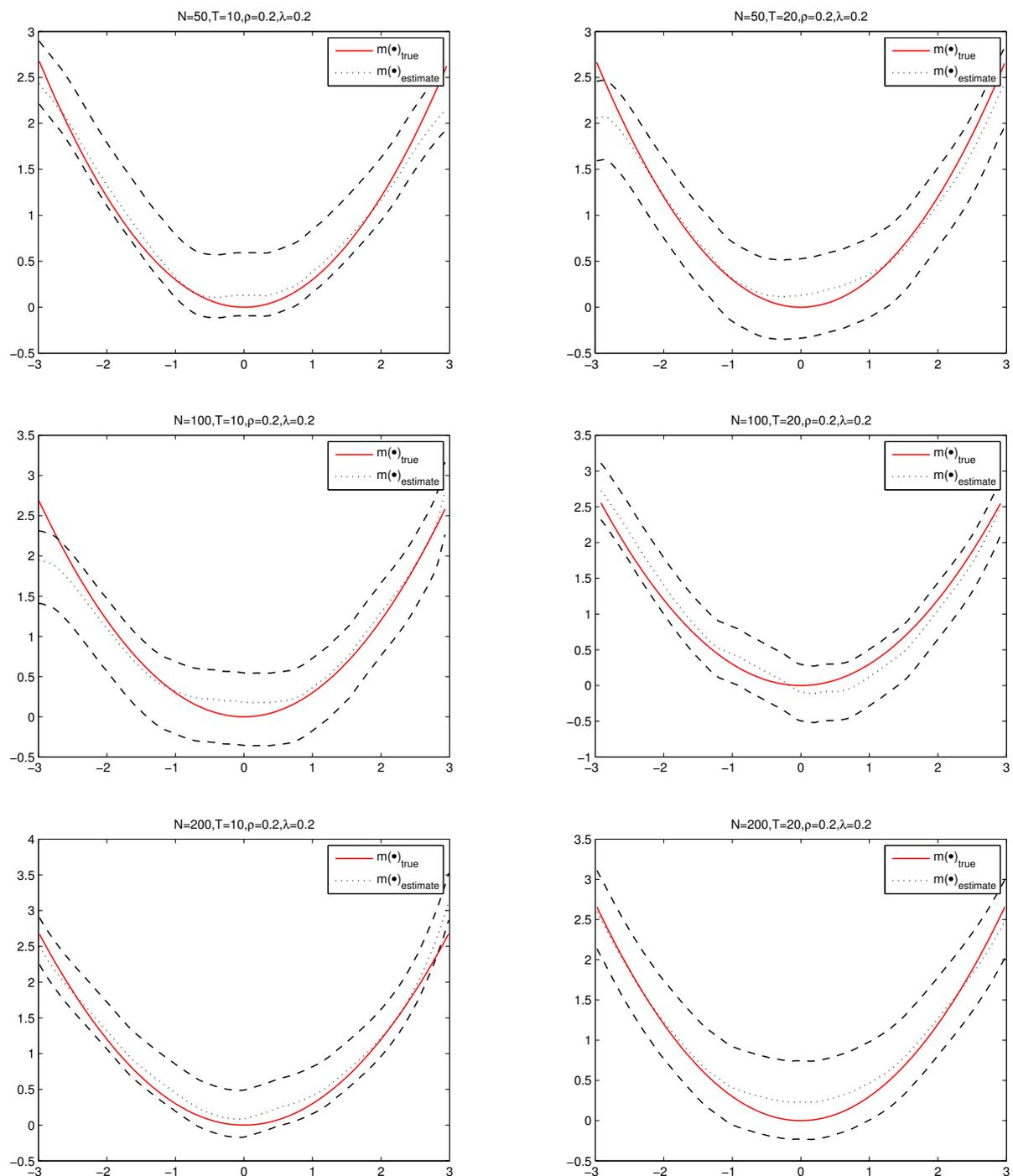


Figure 2. The fitting images and 95% confidence intervals of $\hat{m}(z)$ with $\rho = 0.2$ and $\lambda = 0.2$.

4.2. Comparison of Results of Different Models

In this section, we will investigate the necessity of including nonparametric components, spatial correlation, and serial correlation if the real model is model (12). By deleting spatial correlation, serial correlation, and spatiotemporal correlations, respectively, in the model (12), we get the following three deleted models:

$$\begin{cases} y_{it} = x'_{it}\beta + m(z_{it}) + b_i + v_{it}, \\ v_{it} = \lambda v_{i,t-1} + e_{it}, \end{cases} \tag{13}$$

$$\begin{cases} y_{it} = x'_{it}\beta + m(z_{it}) + b_i + \varepsilon_{it}, \\ \varepsilon_{it} = \rho \sum_{j=1}^N w_{ij}\varepsilon_{jt} + e_{it}, \end{cases} \tag{14}$$

$$y_{it} = x'_{it}\beta + m(z_{it}) + b_i + e_{it}, \tag{15}$$

where all variables in above models are the same as the model (12). To save space, we only study the case that $\rho = 0.4$ and $\lambda = 0.4$. In addition, we set $N = 50, 100, T = 5$ and $mcn = 2000$. By using PQML method in Section 2, the experimental results are presented in Table 5.

Table 5. Results of parametric estimates under the models (12)–(15).

		N = 50, T = 5					N = 100, T = 5				
		$\hat{\beta}$	$\hat{\rho}$	$\hat{\lambda}$	$\hat{\sigma}_e^2$	$\hat{\sigma}_b^2$	$\hat{\beta}$	$\hat{\rho}$	$\hat{\lambda}$	$\hat{\sigma}_e^2$	$\hat{\sigma}_b^2$
model (12)	Mean	0.9848	0.4135	0.4092	1.0211	1.0576	1.0067	0.4063	0.4014	1.0128	1.0266
	MSE	0.0048	0.0014	0.0013	0.0015	0.0165	0.0023	0.0011	0.0009	0.0013	0.0125
model (13)	Mean	0.4711	-	0.5372	1.0221	1.1063	0.8424	-	0.5043	1.0331	1.1221
	MS	5.7078	-	0.0168	0.1057	0.1244	0.9135	-	0.0283	0.0095	0.1365
	GR ₁ (%)	118,812.5	-	1192.3	6946.7	653.9	39,617.3	-	3044.4	630.7	992
model (14)	Mean	0.8372	0.6462	-	1.0234	1.1322	0.9010	0.5968	-	1.0928	1.1012
	MSE	6.2135	0.1281	-	0.0171	0.1211	0.8845	0.1330	-	0.1013	0.1403
	GR ₁ (%)	129,347.9	9050	-	1040	633.9	38,356.5	11,990.9	-	7692.3	1022.4
model (15)	Mean	0.0366	-	-	1.0225	1.1123	0.9828	-	-	1.0948	1.1044
	MSE	5.8645	-	-	0.0160	0.1781	0.7981	-	-	0.1036	0.1474
	GR ₁ (%)	122,077.1	-	-	966.7	979.4	34,600	-	-	7869.2	1079.2

Table 5 reports Means, MSEs and GR₁s of parametric estimates in the models (12)–(15), where GR₁ is growth rate of MSE on the basis of that in model (12). From Table 5, we get the following conclusions: (1) the means of parametric estimates in the real model get closer to the true values as the sample size increases. However, we find that the means of estimates of $\hat{\sigma}_e^2$ and $\hat{\sigma}_b^2$ in model (13), $\hat{\sigma}_e^2$ in the models (14) and (15) are very unstable, they do not converge to true values, respectively. Although the means of other parametric estimates get close to true values as the sample size increases, they still far away from true values and do not converge to true values. (2) In comparison with the real model, the MSEs of each parameter estimate in the deleted models increase. For most parametric estimates, they decrease with the increase in sample size except for $\hat{\lambda}$ and $\hat{\sigma}_b^2$ in the model (13), $\hat{\rho}$, $\hat{\sigma}_e^2$ and $\hat{\sigma}_b^2$ in the model (14), $\hat{\sigma}_e^2$ in the model (15). (3) The GR₁s of all parametric estimates are large, they do not decrease with the increase in sample size for some parametric estimates except for $\hat{\beta}$ in the model (13)–(15) and $\hat{\sigma}_e^2$ in the model (13).

Table 6 reports the medians and SDs of MADEs, GR₂s, and GR₃s for nonparametric estimates of the models (12)–(15), where GR₂ and GR₃ are growth rates of median and SD, respectively, on the basis of that in the model (12). It can be seen from Table 6 that, in comparison with the real model, the values of medians and SDs of MADEs for nonparametric estimates of the models (13)–(15) increase rapidly, and do not decrease with the increase in sample size.

Table 6. The medians and SDs of MADEs for nonparametric estimates under the models (12)–(15).

		Model (12)	Model (13)	Model (14)	Model (15)
N = 50, T = 5	Median	0.2232	0.7632	0.7259	0.7983
	GR ₂ (%)	(-)	(241.9)	(225.2)	(257.6)
	SD	0.1461	0.2857	0.3078	0.4305
	GR ₃ (%)	(-)	(95.5)	(1106.7)	(194.7)
N = 100, T = 5	Median	0.2042	1.2397	1.3323	1.2387
	GR ₂ (%)	(-)	(507.1)	(552.4)	(506.6)
	SD	0.1453	0.5595	0.4704	0.4810
	GR ₃ (%)	(-)	(285.1)	(223.7)	(231.0)

The above facts reveal that the efficiency of both parametric and nonparametric estimates depends on the correct specification of errors in (1)–(3), and violation of which can lead to inconsistency.

4.3. Performance of the Test

In this section, we aim to evaluate the power of hypothesis test proposed in Section 2.3. Consider the model (12), and the settings of x_{it}, b_i, w_{ij} and e_{it} are the same as those in model (12). In addition, we only study the case that $\rho = 0.2$ and $\lambda = 0.2$. Then, the corresponding hypothesis is obtained as follows:

$$H_0 : m(z_{it}) = \gamma_0 + \gamma_1 z_{it} \quad \text{versus} \quad H_1 : m(z_{it}) \neq \gamma_0 + \gamma_1 z_{it},$$

where $m(z_{it}) = \eta \sin(\pi(z_{it} - 0.5)) + 0.8z_{it} + 0.6$, η are values consisting of (0, 0.05, 0.1, 0.2, 0.5) and $z_{it} \sim U[0.2, 1.6]$. In particular, when $\eta = 0$, the above null hypothesis is exactly the true nonparametric function. As the value of η increases, the true nonparametric function is increasingly farther from the null hypothesis. Figure 3 presents the true nonparametric functions with different η -values. In this section, we conduct simulations that $N = 100, T = 8, 16$ under significant levels $\alpha = 0.01$ and 0.05 , respectively. For each case, there are 1000 repetitions. Then, the powers of the hypothesis test are shown in Figure 4.

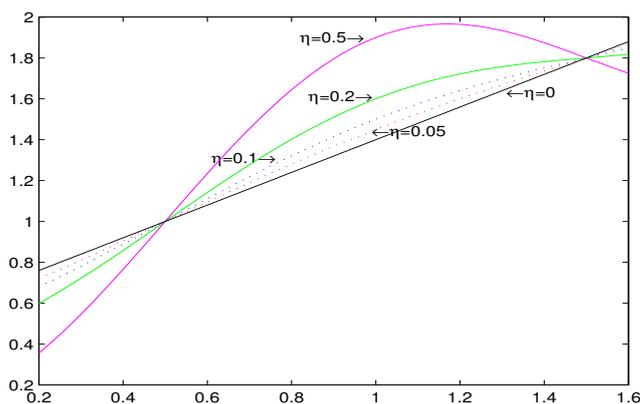


Figure 3. Trajectories of $m(z_{it})$ with $\eta = 0, 0.05, 0.1, 0.2,$ and $0.5,$ respectively.

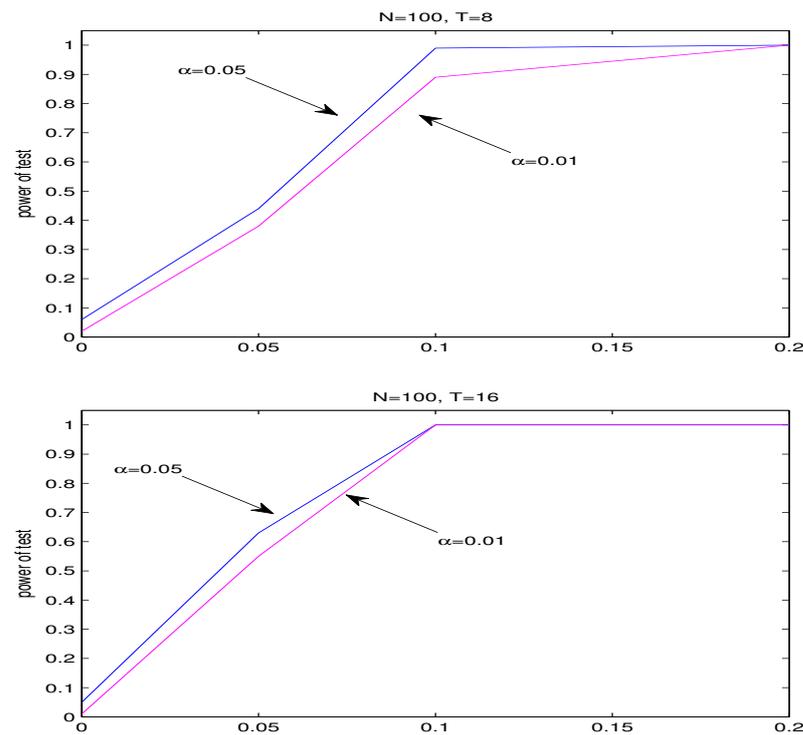


Figure 4. Powers of the test with $\alpha = 0.01$ and 0.05 , respectively, with sample size $N * T$.

We get from the Figure 4 and Table 7 that: (1) while $\eta = 0$, the power of the test is very close to the significance level. As the value of η increases, the power of the test statistic increases considerably, which indicates that our test statistic is sensitive to the alternative hypothesis of the proposed test problem. (2) While the number of spatial units is the same, the power of F-test with $T = 16$ is higher than that with $T = 8$. This shows that the F-test statistic performance get better as sample size increases.

Table 7. Powers of test with significance level α .

N	T	α	$\eta = 0$	$\eta = 0.05$	$\eta = 0.1$	$\eta = 0.2$	$\eta = 0.5$
100	8	0.05	0.06	0.44	0.99	1	1
		0.01	0.02	0.38	0.89	1	1
100	16	0.05	0.05	0.63	1	1	1
		0.01	0.01	0.55	1	1	1

5. Real Data Analysis

In this section, we will illustrate the prescribed estimation and testing methods by Indonesian rice farming dataset with $N = 171$ and $T = 6$ which is a quintessential example of large N and small T in the stochastic frontier literature (see detail for Feng and Horrace [35]). This dataset has five variables, one response variable and four covariates. The response variable is natural logarithm of output of Indonesian rice farming, and the covariates include high, mixed, seed, and land, which are defined in Table 8. The dataset is from the agricultural economic research center of the Ministry of agriculture of Indonesia and compiled by the Agro Economic Survey. Based on the panel data of related variables of 171 farms over 6 growing seasons (three wet and three dry seasons), we will explore the influencing factors of output of Indonesian rice farming by RESPRM with separable space-time filters.

Table 8. Variables and their definitions.

Variable	Definition
high	Dummy variable for high yield varieties ($high = 1$ if the i th form at t th growing reason used high yield varieties, otherwise $high = 0$)
mixed	Dummy variable for mixed yield varieties ($mixed = 1$ if the i th form at t th growing reason used mixed varieties, otherwise $mixed = 0$)
seed	Seed weight used in t th season on i th farm(measured in kilogram)
land	Total area that farmers cultivated with rice(measured in hectares)

In order to take the proposed RESPRM with separable space-time filters to Indonesian rice farming dataset, we apply the F-test statistic proposed in Section 2.3 to verify whether or not a nonlinear relationship between covariates and response variable exists. The test results are given in Table 9. From Table 9, we find that land (other covariates) has (have) significant nonlinear (linear) relationship with natural logarithm of output of Indonesian rice farming at significant level $\alpha = 0.01$.

Therefore, we consider the model (1)–(3), and set $1 \leq i \leq 171$, $1 \leq t \leq 6$, y_{it} is the i th observation of $\ln(rice)$ at t th growing season, x_{it} are the i th observation of high, mixed, and seed at t th growing season, respectively, z_{it} is land for the i th observation at t th growing season. w_{ij} equals 1 if farms i and j are in the same village, equals 0 otherwise (see detail for Druska and Horrace [8]).

Table 9. Results of F-test under significant level $\alpha = 0.01$.

Variable	High	Mixed	Seed	Land
F-test value	2.33326	3.4298	3.7706	185.0233 ***

Note: *** means that it is significant under test level 1%.

The parametric estimates results of the model (1)–(3) for fitting Indonesia rice farming data are described in Table 10. Via Table 10, we find that: (1) $\beta_i (i = 1, 2, 3) > 0$ show expectation of dummy variables for high, mixed, and seed have promotional effects on output of rice. (2) $\sigma_e^2 = 2.4049$ presents that the output of rice in different farm is relatively unstable and is affected by external fluctuations.

Table 10. Results of parametric estimates of model (1)–(3) for fitting Indonesia rice farming data.

Parameter	$\hat{\rho}$	$\hat{\lambda}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\sigma}_e^2$	$\hat{\sigma}_b^2$
Estimate	−0.7227 ***	0.3891 ***	3.2123 ***	2.6885 ***	0.0172 **	2.4049 ***	0.0023 ***

Note: *** and ** mean that they are significant under test level 1% and 5% respectively.

Figure 5 shows the results of estimation and corresponding 95% confidence intervals of $m(z)$, where the blue solid line is the average fit over 500 simulations, the red and green dotted lines are the corresponding 95% confidence bands. From Figure 5, we know that land has obvious nonlinear effects on output of rice, and output of rice increases with the increase in land.

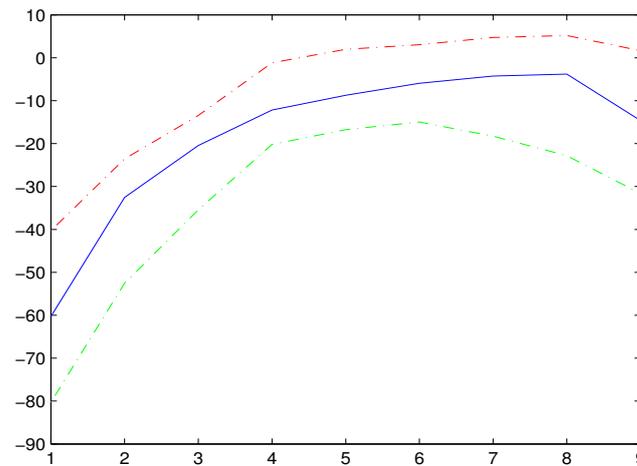


Figure 5. $\hat{m}(z)$ and its 95% confidence intervals for the dataset of Indonesian rice farming in China.

6. Conclusions

In this paper, we study PQMLE and hypothesis test for RESPRM with separable space-time filters. The proposed model can simultaneously capture linear and nonlinear effects of covariates, spatial correlation of error structure, serial correlation of remainder error structure, and individual random effects. With the given conditional assumptions, asymptotic properties of PQMLE and asymptotic distribution of nonparametric components are proved. Monte Carlo simulations are applied to investigate the performance of estimators and test statistic under finite samples, and the results show that proposed estimators and test statistic are well-behaved under finite samples, and consistency of parametric estimates is hard to guarantee if spatial correlation and series correlation in the real model are ignored in some cases. In addition, the practicability of the proposed methods are also assessed by a real dataset.

This paper focuses only on RESPRM with separable space-time filters but does not account for spatially lagged response variables. How to extend our proposed estimation and testing methods to a random effects spatial lag semiparametric regression model with separable space-time filters is an open problem. In addition, we may combine generalized moment estimation method, quantile regression method and local linear method to estimate parametric and nonparametric components of the random effects spatial lag semiparametric regression model with separable space-time filters and RESPRM with separable space-time filters.

Author Contributions: Formal analysis, S.L. and B.L.; Methodology, S.L. and J.C.; Software and Writing—original draft, S.L.; Supervision, Writing—review & editing and Funding acquisition, J.C.; Data curation, B.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).

Data Availability Statement: Not applicable.

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

Appendix A

This section aims to prove some lemmas and theorems given by Section 3, before that, we provide two frequently used evident facts.

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

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

Proof of Theorem 1. Define $Q_{NT} = E[\ln L(\theta)]$. By direct calculation, we have

$$Q_{NT}(\theta) = -\frac{1}{2NT} \ln[(2\pi)] - \frac{1}{2} \ln|\Omega| - \frac{1}{2} E[\zeta'(\theta)\Omega^{-1}\zeta(\theta)], \tag{A1}$$

and

$$\ln L(\theta) = -\frac{1}{2NT} \ln[(2\pi)] - \frac{1}{2} \ln|\Omega| - \frac{1}{2} \zeta'(\theta)\Omega^{-1}\zeta(\theta). \tag{A2}$$

To show the consistency of $\hat{\theta}$, we follow Lee [29] by showing the uniform convergence of $\frac{1}{NT}[\ln L(\theta) - Q_{NT}(\theta)]$ to zero on a compact convex parameter space Λ . By comparison with Lee [29], the difference is that our model setting is more complex. By White [36] (Theorem 3.4), it needs only to show

$$\frac{1}{NT}[\ln L(\theta) - Q_{NT}(\theta)] = op(1). \tag{A3}$$

and

$$\lim_{N \rightarrow \infty} \frac{1}{NT} [Q_{NT}(\theta) - Q_{NT}(\theta_0)] < 0. \tag{A4}$$

We first prove (A3). By direct calculation, we know that

$$\zeta'(\theta) = X\beta_0 + M_0 - (X\beta + M) + \zeta,$$

and

$$\begin{aligned} \frac{1}{NT}[\ln L(\theta) - Q_{NT}(\theta)] &= \frac{1}{2NT} \{ \zeta'(\theta)\Omega^{-1}\zeta(\theta) - E[\zeta'(\theta)\Omega^{-1}\zeta(\theta)] \} \\ &= \frac{1}{2NT} \zeta'\Omega^{-1}\zeta - p(\zeta) + \frac{1}{NT} \zeta'\Omega^{-1}(X\beta_0 + M_0), \end{aligned}$$

where $p(\zeta) = \frac{1}{NT} \text{tr}(\Omega^{-1}\Omega_0)$. As Ω^{-1} is UB, we know $\frac{1}{NT} \zeta'\Omega^{-1}\zeta - p(\zeta) = op(1)$, and as $\frac{1}{NT} \zeta'\Omega^{-1}(X\beta_0 + M_0) \xrightarrow{P} 0$, that is uniform convergence is proved.

We then prove (A4). Recall (A1) and $Q_{NT}(\theta_0) = -\frac{1}{2NT} \ln[(2\pi)] - \frac{1}{2} \ln|\Omega_0| - \frac{1}{2}$. By direct calculation, we know

$$E[\frac{1}{NT} \zeta'(\theta)\Omega^{-1}\zeta(\theta)] = \frac{1}{NT} [X\beta_0 + M_0 - (X\beta + M)]'\Omega^{-1} [X\beta_0 + M_0 - (X\beta + M)] + \frac{1}{NT} \text{tr}(\Omega^{-1}\Omega_0),$$

then

$$\begin{aligned} Q_{NT}(\theta) - Q_{NT}(\theta_0) &= -\frac{1}{2NT} \ln|\Omega| - \frac{1}{2NT} \text{tr}(\Omega^{-1}\Omega_0) - \frac{1}{2NT} H_{NT} - (\frac{1}{2NT} \ln|\Omega_0| - \frac{1}{2}) \\ &= T_{1NT} - T_{2NT}, \end{aligned}$$

where $T_{1NT} = -\frac{1}{2NT} \ln|\Omega| - \frac{1}{2NT} \text{tr}(\Omega^{-1}\Omega_0) - [\frac{1}{2NT} \ln|\Omega_0| - \frac{1}{2}]$, $T_{2NT} = \frac{1}{2NT} H_{NT}$. Conduct a random effects model $Y = X\beta + Ub + \varepsilon$, we know its log-likelihood function is as follows

$$\ln L_{NT}^p(\tilde{\theta}) = -\frac{1}{2NT} \ln[(2\pi)] - \frac{1}{2} \ln|\Omega| - \frac{1}{2NT} \zeta'(\tilde{\theta})\Omega^{-1}\zeta(\tilde{\theta}),$$

where $\zeta'(\tilde{\theta}) = Y - X\beta$. Using the information inequality of above model, $T_{1NT} \leq 0$. In addition, T_{2NT} is a quadratic function of $[X\beta_0 + M_0 - (X\beta + M)]$ and $T_{2NT} > 0$ for given any $\zeta \neq \zeta_0$. Next, we prove that $T_{1NT} < 0$ strictly, i.e.,

$$\frac{1}{NT} \ln|\Omega_0| + 1 - \frac{1}{NT} \ln|\Omega| - \text{tr}(\Omega^{-1}\Omega_0) \neq 0.$$

Combined Assumption 5 and uniform convergence, $\hat{\theta} - \theta_0 = op(1)$ is proved. \square

Proof of Theorem 2. By Taylor expansion of (5) at point θ_0 , we obtain

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

where $\tilde{\theta} = (\tilde{\zeta}', \tilde{\beta}')'$, $\zeta = (\tilde{\rho}, \tilde{\lambda}, \tilde{\sigma}_b^2, \tilde{\sigma}_e^2)$ and $\tilde{\theta}$ lies in between $\hat{\theta}$ and θ_0 . According to Theorem 1, we have $\tilde{\theta} \xrightarrow{P} \theta_0$. Denote

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

Therefore,

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

Next, we aim to prove

$$\frac{1}{NT} \frac{\partial^2 \ln L(\tilde{\theta})}{\partial \theta \partial \theta'} - \frac{1}{NT} \frac{\partial^2 \ln L(\theta_0)}{\partial \theta \partial \theta'} = op(1), \quad (\text{A5})$$

and

$$\frac{1}{\sqrt{NT}} \frac{\partial \ln L(\theta_0)}{\partial \theta} \xrightarrow{D} N(0, \Sigma_{\theta_0} + \Phi_{\theta}). \quad (\text{A6})$$

In order to prove (A5), we will show that all elements of $\frac{1}{NT} \frac{\partial^2 \ln L(\tilde{\theta})}{\partial \theta \partial \theta'} - \frac{1}{NT} \frac{\partial^2 \ln L(\theta_0)}{\partial \theta \partial \theta'}$ converge to 0 in probability. It can be calculated that

$$\begin{aligned} \frac{1}{\sqrt{NT}} \frac{\partial \ln L(\theta_0)}{\partial \beta'} &= \frac{1}{\sqrt{NT}} - X' \Omega_0^{-1} \bar{\zeta}(\theta), \\ \frac{1}{\sqrt{NT}} \frac{\partial \ln L(\theta_0)}{\partial \delta_i} &= \frac{1}{2} [\zeta'(\theta) \frac{\partial \Omega_0^{-1}}{\partial \delta_i} \zeta(\theta) - \text{tr}(\frac{\partial \Omega_0}{\partial \delta_i'} \Omega_0^{-1})] (i = 1, 2, 3, 4). \end{aligned}$$

where

$$\begin{aligned} \frac{\partial \Omega}{\partial \rho} &= \sigma_e^2 \{ V_{\lambda} \otimes B^{-1} [(WB^{-1})' + WB^{-1}] B^{-1'} \}, \\ \frac{\partial \Omega}{\partial \lambda} &= \sigma_e^2 \{ [\frac{2\lambda}{1-\lambda^2} V_{\lambda} + \frac{1}{1-\lambda^2} F_{\lambda}] \otimes B^{-1} B^{-1'} \}, \\ \frac{\partial \Omega}{\partial \sigma_e^2} &= V_{\lambda} \otimes (B^{-1} B^{-1'}), \\ \frac{\partial \Omega}{\partial \sigma_b^2} &= (I_T I_T') \otimes I_N, \end{aligned}$$

Ω_0 is the true value of Ω , F_{λ} is the first derivative of V_{λ} with respect to λ . Consequently, we get

$$\begin{aligned} \frac{1}{NT} \frac{\partial^2 \ln L(\theta_0)}{\partial \beta \partial \beta'} &= \frac{1}{NT} X' \Omega_0^{-1} X, \\ \frac{1}{NT} \frac{\partial^2 \ln L(\theta_0)}{\partial \beta \partial \delta_i} &= 0, \\ \frac{1}{NT} \frac{\partial^2 \ln L(\theta_0)}{\partial \delta_i^2} &= - \frac{1}{2NT} \text{tr} \left(\frac{\partial \Omega_0}{\partial \delta_i} \frac{\partial \Omega_0^{-1}}{\partial \delta_j} \right) (i, j = 1, 2, 3, 4). \end{aligned}$$

Then, by mean value theorem, Assumption 2 (iv) and Fact 2, we easily get

$$\begin{aligned} \frac{1}{NT} \frac{\partial^2 \ln L_1(\theta_0)}{\partial \beta \partial \beta'} - \frac{1}{NT} \frac{\partial^2 \ln L_1(\tilde{\theta})}{\partial \beta \partial \beta'} &= X'(\Omega_0^{-1} - \tilde{\Omega}^{-1})X \\ &= X' \left[\frac{V_\lambda^{-1} l_T l_T' V_\lambda^{-1}}{d^2(1-\lambda^2)} \otimes Z_0 - \frac{V_{\tilde{\lambda}}^{-1} l_T l_T' V_{\tilde{\lambda}}^{-1}}{\tilde{d}^2(1-\tilde{\lambda}^2)} \otimes \tilde{Z}_0 \right] X \\ &\quad + X'(V_\lambda^{-1} \otimes \frac{BB'}{\sigma_b^2} - V_{\tilde{\lambda}}^{-1} \otimes \frac{\tilde{B}\tilde{B}'}{\tilde{\sigma}_b^2})X \\ &\quad - X' \left[\frac{V_\lambda^{-1} l_T l_T' V_\lambda^{-1}}{d^2(1-\lambda^2)} \otimes \frac{BB'}{\sigma_b^2} - \frac{V_{\tilde{\lambda}}^{-1} l_T l_T' V_{\tilde{\lambda}}^{-1}}{\tilde{d}^2(1-\tilde{\lambda}^2)} \otimes \frac{\tilde{B}\tilde{B}'}{\tilde{\sigma}_b^2} \right] X \\ &= op(1). \end{aligned}$$

Then we can similarly prove that (A5) holds.

In order to prove (A6), we follow the idea of Lee [29]. 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 e and their means are all $op(1)$. With Assumption 1, we have that $\frac{1}{\sqrt{NT}} \frac{\partial \ln L(\theta_0)}{\partial \theta}$ is asymptotically normal distributed with 0 means by using the CLT for linear-quadratic forms of Theorem 1 in Kelejian and Prucha [37]. In the next step, we calculate the variance. Denote $\mu_{\eta 03}$ and $\mu_{\eta 04}$ as the third and fourth moments of η , respectively, A_{ii} be the i -th element of A . By using the facts that $E(U_n' A_n U_n) = \sigma_0^2 tr(A_n)$, $E(U_n' A_n U_n U_n' B_n U_n) = (\mu_{04} - 3\sigma_0^4) \sum_{i=1}^n a_{n,ii} b_{n,ii} + \sigma_0^4 [tr(A_n B_n) + tr(A_n B_n) + tr(A_n B_n')]$, it follows by direct calculation that

$$Var\left(\frac{1}{\sqrt{NT}} \frac{\partial \ln L(\theta_0)}{\partial \theta}\right) = E\left(\frac{1}{\sqrt{NT}} \frac{\partial \ln L(\theta_0)}{\partial \theta} \cdot \frac{1}{\sqrt{NT}} \frac{\partial \ln L(\theta_0)}{\partial \theta'}\right) = -E\left(\frac{1}{NT} \frac{\partial^2 \ln L(\theta_0)}{\partial \theta \partial \theta'}\right) + E(\Phi_{\theta_{NT}}),$$

where

$$\begin{aligned} -\frac{1}{NT} \frac{\partial^2 \ln L(\theta_0)}{\partial \theta \partial \theta'} &= \begin{pmatrix} -\frac{1}{NT} (X' \Omega_0 X)_{p \times p} & \mathbf{0}_{p \times 4} \\ \mathbf{0}_{4 \times p} & \frac{1}{2NT} tr\left(\frac{\partial \Omega_0}{\partial \delta_i'} \frac{\partial \Omega_0^{-1}}{\partial \delta_i'}\right) \end{pmatrix} + op(1) \equiv \Sigma_{\theta_0} + op(1), \\ \Phi_{\theta} &= \begin{pmatrix} \mathbf{0}_{p \times p} & -\frac{1}{2NT} \mu_{\eta 03} \frac{\partial \Omega_0}{\partial \delta_i'} diag(X' \Omega_0^{-1})_{p \times 4} \\ * & \left(\frac{\mu_{\eta 04} - 3\sigma_{\eta 0}^2}{4NT} \sum_{i=1}^{NT} \left(\frac{\partial \Omega_0^{-1}}{\partial \delta_i'}\right)_{ii}\right)_{4 \times 4} \end{pmatrix}. \end{aligned}$$

In particular, it is not difficult to know that $\Phi_{\theta} = o(1)$ when e and b obey normal distribution. This completes the proof. \square

Proof of Theorem 3. We follow the idea of Ullah and Su [31] to prove the theorem. The difference is that our error structure setting is more complex. Recall

$$\hat{\delta} = S(z)(Y - X\hat{\beta}),$$

then

$$\hat{\delta} = S(z)(Y - X\hat{\beta}) = S(z)[m(z) + \eta - X(\hat{\beta} - \beta_0)]. \tag{A7}$$

Write $\delta = \begin{pmatrix} m(z) \\ H\dot{m}(z) \end{pmatrix}$, where $\dot{m}(z)$ is the first derivative of $m(z)$. By the second order Taylor expression, we obtain $m(z_{it}) = \bar{Z}'_{it}(z)\delta + \frac{1}{2}(z_{it} - z)' \ddot{m}(z)(z_{it} - z) + op(\|H\|^2)$. Thus, we have

$$\begin{aligned} S(z)m(z_{it}) &= [\bar{Z}'(z)K_H(z)\bar{Z}(z)]^{-1} \bar{Z}'(z)K_H(z)\bar{Z}(z)\delta + \frac{1}{2}S(z_{it})(z_{it} - z)' \ddot{m}(z)(z_{it} - z) \\ &= \delta + \frac{1}{2}S(z_{it})(z_{it} - z)' \ddot{m}(z)(z_{it} - z). \end{aligned} \tag{A8}$$

Thus, we get

$$\hat{\delta} - \delta = \frac{1}{2} \sum_{i=1}^N \sum_{t=1}^T S(z_{it})(z_{it} - z)\dot{m}(z)(z_{it} - z) + S(z)\eta - S(z)X(\hat{\beta} - \beta)$$

and

$$\sqrt{N|H|}(\hat{\delta} - \delta) = B_{11} + B_{12} - B_{13}, \tag{A9}$$

where

$$B_{11} = \frac{\sqrt{N|H|}}{2} \sum_{i=1}^N \sum_{t=1}^T S(z_{it})(z_{it} - z)' \dot{m}(z)(z_{it} - z),$$

$$B_{12} = \sqrt{N|H|}S(z)\eta,$$

$$B_{13} = \sqrt{N|H|}S(z)X(\hat{\beta} - \beta).$$

For B_{11} , write

$$\begin{aligned} & \frac{\sqrt{N|H|}}{2} \sum_{i=1}^N \sum_{t=1}^T S(z_{it})(z_{it} - z)\dot{m}(z)(z_{it} - z) \\ &= \frac{\sqrt{N|H|}}{2} Q^{-1} \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \bar{Z}'_{it}(z)K_H(z_{it} - z)(z_{it} - z)' \dot{m}(z)(z_{it} - z) + op(1) \\ &= Q^{-1} \frac{\sqrt{N|H|}}{2} \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T [1(z_{it} - z)'H^{-1}]' K_H(z_{it} - z)(z_{it} - z)' \dot{m}(z)(z_{it} - z) + op(1) \\ &= Q^{-1} \frac{\sqrt{N|H|}}{2} \frac{1}{NT} \left(\frac{\sum_{i=1}^N \sum_{t=1}^T K_H(z_{it} - z)(z_{it} - z)' \dot{m}(z)(z_{it} - z)}{\sum_{i=1}^N \sum_{t=1}^T (z_{it} - z)'H^{-1} K_H(z_{it} - z)(z_{it} - z)' \dot{m}(z)(z_{it} - z)} \right) + op(1) \\ &= Q^{-1} \frac{\sqrt{N|H|}}{2} \left(\frac{\bar{f}(z) \int u'uK(u)duH\dot{m}(z)H}{0} \right) + op(1). \end{aligned}$$

Thus, we obtain

$$B_{11} = \frac{\bar{f}(z)\sqrt{N|H|}}{2} Q^{-1} \left(\frac{\int u'uK(u)duH\dot{m}(z)H}{0} \right) + op(1),$$

where $Q = \bar{f}(z) \begin{pmatrix} 1 & 0' \\ 0 & \int uu'K(u)du \end{pmatrix}$, $\bar{f}(z) = \sum_{t=1}^T f_t(z)$.

For B_{12} , write $B_{12} = \sqrt{N|H|}S(z)\eta$, we can show that $E[B_{12}] = 0$ and

$$Var(B_{12}) = E[B_{12}^2] \leq \sigma_\eta^2 = O(1).$$

Combined (A9), Slutsky's theorem and central limit theorem, we get

$$\sqrt{N|H|} \left(\hat{\delta} - \delta - Q^{-1} \left(\frac{\bar{f}(z)}{2} \left(\int uu'K(u)duH\dot{m}(z)H \right) \right) \right) \xrightarrow{D} N(0, \sigma^2(\eta)).$$

where $\sigma^2(\delta) = \Omega$, $Q = \bar{f}(z) \begin{pmatrix} 1 & 0' \\ 0 & \int uu'K(u)du \end{pmatrix}$, and $\bar{f}(z) = \sum_{t=1}^T f_t(z)$. \square

Proof of Theorem 4. It can be seen that

$$\begin{aligned}
 RSS(H1) &= \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \hat{m}(z_{it}) - x'_{it}\hat{\beta}]^2 \\
 &= \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \hat{m}(z_{it}) - x'_{it}\beta]^2 \\
 &\quad + \left\{ \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \hat{m}(z_{it}) - x'_{it}\hat{\beta}]^2 - \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \hat{m}(z_{it}) - x'_{it}\beta]^2 \right\} \\
 &= RSS^*(H1) + T_1,
 \end{aligned}$$

where $RSS^*(H1) = \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \hat{m}(z_{it}) - x'_{it}\beta]^2$ and $T_1 = \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \hat{m}(z_{it}) - x'_{it}\hat{\beta}]^2 - \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \hat{m}(z_{it}) - x'_{it}\beta]^2$. For T_1 , it is obtained by Theorems 1–3 and direct calculation that

$$\begin{aligned}
 T_1 &= \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \hat{m}(z_{it}) - x'_{it}\hat{\beta}]^2 - \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \hat{m}(z_{it}) - x'_{it}\beta]^2 \\
 &= \sum_{i=1}^N \sum_{t=1}^T [x'_{it}\beta + m(z_{it}) + \varepsilon_{it} - \hat{m}(z_{it}) - x'_{it}\hat{\beta}]^2 \\
 &\quad - \sum_{i=1}^N \sum_{t=1}^T [x'_{it}\beta + m(z_{it}) + \varepsilon_{it} - \hat{m}(z_{it}) - x'_{it}\beta]^2 \\
 &= \sum_{i=1}^N \sum_{t=1}^T [Op\left(\frac{1}{\sqrt{NT}}\right) + O(|H|^2) + |H|^2Op\left(\frac{1}{\sqrt{N}|H|}\right) + \eta]^2 \\
 &\quad - \sum_{i=1}^N \sum_{t=1}^T [O(|H|^2) + |H|^2Op\left(\frac{1}{\sqrt{N}|H|}\right) + \eta]^2 \\
 &= O(1).
 \end{aligned}$$

Similarly, we know that

$$\begin{aligned}
 RSS(H_0) &= \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \tilde{m}(z_{it}) - x'_{it}\hat{\beta}]^2 \\
 &= \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \tilde{m}(z_{it}) - x'_{it}\beta]^2 \\
 &\quad + \left\{ \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \tilde{m}(z_{it}) - x'_{it}\hat{\beta}]^2 - \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \tilde{m}(z_{it}) - x'_{it}\beta]^2 \right\} \\
 &= RSS^*(H_0) + T_0,
 \end{aligned}$$

where $RSS^*(H_0) = \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \tilde{m}(z_{it}) - x'_{it}\beta]^2$, and $T_0 = \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \tilde{m}(z_{it}) - x'_{it}\hat{\beta}]^2 - \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \tilde{m}(z_{it}) - x'_{it}\beta]^2$. For T_0 , it is obtained by Theorems 1–3 and calculation that

$$\begin{aligned}
T_0 &= \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \tilde{m}(z_{it}) - x'_{it}\hat{\beta}]^2 - \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \tilde{m}(z_{it}) - x'_{it}\beta]^2 \\
&= \sum_{i=1}^N \sum_{t=1}^T [x'_{it}\beta + m(z_{it}) + \varepsilon_{it} - \tilde{m}(z_{it}) - x'_{it}\hat{\beta}]^2 \\
&\quad - \sum_{i=1}^N \sum_{t=1}^T [x'_{it}\beta + m(z_{it}) + \varepsilon_{it} - \tilde{m}(z_{it}) - x'_{it}\beta]^2 \\
&= \sum_{i=1}^N \sum_{t=1}^T [Op(\frac{1}{\sqrt{NT}}) + \tilde{m}(z_{it}) - x'_{it}\hat{\beta}]^2 - \sum_{i=1}^N \sum_{t=1}^T [\tilde{m}(z_{it}) - x'_{it}\beta]^2 \\
&= Op(1).
\end{aligned}$$

In addition, we obtain that

$$F_{NT} = \frac{NT}{2} \frac{RSS^*(H_0) - RSS^*(H_1)}{RSS^*(H_1)} (1 + op(1)) + Op(1).$$

By Remark 3.4 in [27], under H_0 , we obtain

$$r_k F_{NT} \xrightarrow{D} \chi_{r_k c_k | \mathcal{Z} | / |H|}^2,$$

where $r_k = \frac{K(0) - \frac{1}{2} \int K^2(z) du}{\int (K(z) - \frac{1}{2} K * K(z))^2 du}$, $K(z) = \text{diag}(H^{-1}k(H^{-1}(z_{11} - z)), \dots, H^{-1}k(H^{-1}(z_{NT} - z)))$, $c_k = K(0) - \frac{1}{2} \int K^2(z) du$, \mathcal{Z} is the support of z_{it} and H is a bandwidth sequence. \square

References

- Bates, G.E.; Neyman, J. Contributions to the theory of accident proneness II: True of false contagion. *Univ. Calif. Publ. Statist.* **1952**, *1*, 255–276.
- Chamberlain, G. Multivariate regression models for panel data. *J. Econom.* **1982**, *18*, 5–46. [\[CrossRef\]](#)
- Baltagi, B. *Econometrics Analysis of Panel Data*; Wiley Press: New York, NY, USA, 2001.
- Arellano, M. *Panel Data Econometrics*; Oxford University Press: New York, NY, USA, 2003.
- Baldev, R.; Baltagi, B. *Panel Data Analysis*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2012.
- Hsiao, C. *Analysis of Panel Data*; Cambridge University Press: Cambridge, UK, 2014.
- Elhorst, J. *Spatial Econometrics: From Cross-Sectional Data to Spatial Panels*; Springer Press: Berlin/Heidelberg, Germany, 2014.
- Druska, V.; Horrace, W. Generalized moments estimation for spatial panel data: Indonesian rice farming. *Am. J. Agric. Econ.* **2004**, *86*, 185–198. [\[CrossRef\]](#)
- Egger, P.; Pfaffermayr, M.; Winner, H. An unbalanced spatial panel data approach to US state tax competition. *Econ. Lett.* **2005**, *88*, 329–335. [\[CrossRef\]](#)
- Kapoor, M.; Kelejian, H.; Prucha, I. Panel data models with spatially correlated error components. *J. Econ.* **2007**, *140*, 97–130. [\[CrossRef\]](#)
- Lee, L.; Yu, J. Estimation of spatial autoregressive panel data models with fixed effects. *J. Econ.* **2010**, *154*, 165–185. [\[CrossRef\]](#)
- Elhorst, J. Serial and spatial autocorrelation. *Econ. Lett.* **2008**, *100*, 422–424. [\[CrossRef\]](#)
- Parent, O.; LeSage, J. A space-time filter for panel data models containing random effects. *Comput. Stat. Data Anal.* **2011**, *55*, 475–490. [\[CrossRef\]](#)
- Lee, L.; Yu, J. Spatial panels: Random components versus fixed effects. *Int. Econ. Rev.* **2012**, *53*, 1369–1412. [\[CrossRef\]](#)
- Lee, L.; Yu, J. Estimation of fixed effects panel regression models with separable and nonseparable space-time filters. *J. Econ.* **2015**, *184*, 174–192. [\[CrossRef\]](#)
- Baltagi, B.; Song, S.; Jung, B.; Koh, W. Testing for serial correlation, spatial autocorrelation and random effects using panel data. *J. Econ.* **2007**, *140*, 5–51. [\[CrossRef\]](#)
- Cohen, J.; Morrison Paul, C. Public infrastructure investment, interstate spatial spillovers and manufacturing costs. *Rev. Econ. Stat.* **2004**, *86*, 551–560. [\[CrossRef\]](#)
- Zhao, J.; Zhao, Y.; Lin, J.; Miao, Z.-X.; Khaled, W. Estimation and testing for panel data partially linear single-index models with errors correlated in space and time. *Random Matrices* **2020**, *9*, 2150005. [\[CrossRef\]](#)
- Bai, Y.; Hu, J.; You, J. Panel data partially linear varying-coefficient model with both spatially and time-wise correlated errors. *Stat. Sinica* **2015**, *35*, 275–294. [\[CrossRef\]](#)
- Shi, W.; Lee, L. Spatial dynamic panel data models with interactive fixed effects. *J. Econ.* **2017**, *197*, 323–347. [\[CrossRef\]](#)

21. Baltagi, B.; Fingleton, B.; Pirotte, A. Estimating and forecasting with a dynamic spatial panel data model. *Oxford. B Econ. Stat.* **2014**, *76*, 112–138. [[CrossRef](#)]
22. Cai, Z. Trending time-varying coefficient time series models with serially correlated errors. *J. Econ.* **2007**, *136*, 163–188. [[CrossRef](#)]
23. 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. [[CrossRef](#)]
24. Fan, J.; Huang, T.; Li, R.Z. Analysis of longitudinal data with semi-parametric estimation of covariance function. *J. Am. Stat. Assoc.* **2007**, *102*, 632–641. [[CrossRef](#)]
25. Fan, J.; Zhang, W. Statistical methods with varying coefficient models. *Stat. Inference* **2008**, *1*, 179–195. [[CrossRef](#)]
26. Tang, L.; Liu, Y. Estimation of generalized spatial lag semi-parametric varying-coefficient panel model with random effects. *Stat. Res.* **2018**, *35*, 119–128.
27. Fan, J.; Zhang, C.; Zhang, J. Generalized likelihood ratio statistics and Wilks phenomenon. *Ann. Stat.* **2001**, *29*, 153–193. [[CrossRef](#)]
28. Su, L.; Jin, S. Profile quasi-maximum likelihood estimation of partially linear spatial autoregressive models. *J. Econ.* **2010**, *157*, 18–33. [[CrossRef](#)]
29. Lee, L. Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica* **2004**, *72*, 1899–1925. [[CrossRef](#)]
30. Hamilton, S.; Truong, Y. Local linear estimation in partly linear models. *J. Multivar. Anal.* **1997**, *60*, 1–19. [[CrossRef](#)]
31. Ullah, A.; Su, L. Profile likelihood estimation of partially linear panel data models with fixed effects. *Econ. Lett.* **2006**, *92*, 75–81.
32. Mack, Y.; Silverman, B. Weak and strong uniform consistency of kernel regression estimates. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **1982**, *61*, 405–415. [[CrossRef](#)]
33. Su, L. Semiparametric GMM estimation of spatial autoregressive models. *J. Econ.* **2012**, *167*, 543–560. [[CrossRef](#)]
34. Anselin, L. *Spatial Econometrics: Methods and Models*; Kluwer Academic Press: Dordrecht, The Netherlands, 1988.
35. Feng, Q.; Horrace, W. Alternative technical efficiency measures: Skew, bias and scale. *J. Appl. Economet.* **2012**, *27*, 253–268. [[CrossRef](#)]
36. White, H. *Estimation, Inference and Specification Analysis*; Cambridge University Press: New York, NY, USA, 1994.
37. Kelejian, H.; Prucha, I. On the asymptotic distribution of the Moran I test statistic with applications. *J. Econ.* **2001**, *104*, 219–257. [[CrossRef](#)]