



Article

PQMLE and Generalized F-Test of Random Effects Semiparametric Model with Serially and Spatially Correlated Nonseparable Error

Shuangshuang Li ¹, Jianbao Chen ^{2,*} and Danqing Chen ³

¹ School of Mathematics and Statistics, Henan University of Science and Technology, Luoyang 471000, China; 9906436@haust.edu.cn

² School of Mathematics and Statistics & Fujian Provincial Key Laboratory of Statistics and Artificial Intelligence, Fujian Normal University, Fuzhou 350117, China

³ School of Computer Science and Mathematics, Fujian University of Technology, Fuzhou 350117, China; 19851523@fjut.edu.cn

* Correspondence: jbjy2006@126.com

Abstract: Semiparametric panel data models are powerful tools for analyzing data with complex characteristics such as linearity and nonlinearity of covariates. This study aims to investigate the estimation and testing of a random effects semiparametric model (RESPM) with serially and spatially correlated nonseparable error, utilizing a combination of profile quasi-maximum likelihood estimation and local linear approximation. Profile quasi-maximum likelihood estimators (PQMLEs) for unknowns and a generalized F-test statistic F_{NT} are built to determine the beingness of nonlinear relationships. The asymptotic properties of PQMLEs and F_{NT} are proven under regular assumptions. The Monte Carlo results imply that the PQMLEs and F_{NT} performances are excellent on finite samples; however, missing the spatially and serially correlated error leads to estimator inefficiency and bias. Indonesian rice-farming data is used to illustrate the proposed approach, and indicates that *land area* exhibits a significant nonlinear relationship with *rice yield*, in addition, *high-yield varieties*, *mixed-yield varieties*, and *seed weight* have significant positive impacts on rice yield.

Keywords: RESPM with serially and spatially nonseparable error; PQMLE; generalized F-test statistic; asymptotic property; Monte Carlo; Indonesian rice-farming data



Citation: Li, S.; Chen, J.; Chen, D. PQMLE and Generalized F-Test of random Effects Semiparametric Model with Serially and Spatially Correlated Nonseparable Error. *Fractal Fract.* **2024**, *8*, 386. <https://doi.org/10.3390/fractalfract8070386>

Academic Editors: Yuliya Mishura, Kostiantyn Ralchenko and Anton Yurchenko-Tytarenko

Received: 7 May 2024
Revised: 24 June 2024
Accepted: 27 June 2024
Published: 28 June 2024



Copyright: © 2024 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

By tracking each individual in a predetermined sample over time, a panel dataset with multiple observations can be collected. In contrast to cross-sectional data (Cheng and Chen [1]), the panel data structure is sufficiently abundant to permit estimation and testing of regression models that include not only individual characteristics observed by the researchers but also unobserved characteristics that vary across individuals or/and time periods, simply referred to as “individual or/and time effects”. In recent decades, many studies have focused on parametric panel data models (Chamberlain [2], Chamberlain [3], Hsiao [4], Baltagi [5], Arellano [6], Wooldridge [7], Thapa et al. [8], and Rehfeldt and Weiss [9]).

In practice, a lot of spatial panel datasets are collected from “locations” and require further studying. Therefore, based on classic parametric panel data models, different types of parametric panel data models with spatial correlations at different locations are presented. There are three fundamental types of spatial parametric panel data model (Elhorst [10]): spatial autoregressive (SAR) model (Brueckner [11]), spatial Durbin model (SDM) (LeSage [12], Elhorst [13]), and spatial error model (SEM) (Allers and Elhorst [14]). A more detailed discussion can be found in Elhorst [15], Baltagi and Li [16], Pesaran [17], Kapoor et al. [18], Lee and Yu [19], Mutl and Pfaffermayr [20], and Baltagi et al. [21]. Among these, the SEM introduced by Anselin [22] accounts for the interaction effects

in spatially correlated determinants of the dependent variable that is excluded from the model or in unobserved shocks with a spatial mode (Elhorst [13]). The development of testing and estimation of SEM has been summarized in books by Baltagi [5], Elhorst [10], and Anselin [22] as well as in surveys by Baltagi et al. [23], Bordonon et al. [24], and Kelejian and Prucha [25], among others. However, the aforementioned SEM assumes that the only correlation over time results from the same regional impact presented across the panel. In spatial panel data analysis, this assumption may be too restrictive for real situations. For example, if an undetected shock influences investments across regions, behavioral relationships are affected, at least in the next few stages. Thus, ignoring serially correlated errors may lead to inefficient regression coefficient estimators (Baltagi [5]).

By adding serially correlated errors to SEM, parametric panel data models with serially and spatially correlated errors can be established. The extended models consider not only serial correlations through spatial units but also spatial correlations at every point temporally. Researchers have established two types of parametric panel data models with serially and spatially correlated separable and nonseparable errors. The model with the separable error was used to analyze the effects of public infrastructure investment on the costs and productivity of private enterprises (Cohen and Paul [26]). The MLE and corresponding asymptotic properties of spatial panel data model with serially and spatially correlated separable error were investigated by Lee and Yu [27]. A random effects parametric panel data model with serially and spatially correlated error was introduced by Parent and LeSage [28], and the estimators were obtained using the Markov chain Monte Carlo (MCMC) method. Elhorst [29] constructed the MLE of a parametric panel data model with serially and spatially correlated nonseparable error without establishing the asymptotic properties of the estimators. Lee and Yu [30] established the QMLE and corresponding asymptotic properties for both linear panel data models with serially and spatially correlated separable and nonseparable errors.

Although the theories and applications of the aforementioned linear parametric panel data models with serially and spatially correlated separable/nonseparable errors have been thoroughly explored, their usefulness in applications is often impractical because they lack flexibility and are limited in their ability to detect complicated structures (e.g., nonlinearity). Furthermore, a misspecification of the data generating process by the linear parametric panel data model may result in significant modeling bias and can lead to incorrect findings. For these reasons, Zhao et al. [31] investigated the estimation and testing issues of a partially linear single-index panel data model with serially and spatially correlated separable errors using the semiparametric minimum average variance method and a F-type test statistic. Li et al. [32] obtained the PQMLE, the generalized F-test, and the asymptotic properties for a partially linear nonparametric regression model with serially and spatially correlated separable error. Li et al. [33] studied the PQMLEs of the unknowns for the fixed effects partially linear varying coefficient panel data model with serially and spatially correlated nonseparable error, thereby deriving the consistency and asymptotic normality. Bai et al. [34] constructed the estimators of the parametric and nonparametric components of a partially linear varying-coefficient panel data model with serially and spatially correlated separable errors using weighted semiparametric least squares and weighted polynomial spline series methods, respectively, thereby proving their asymptotic normality.

By adding a nonparametric component to a random effects parametric model with serially and spatially correlated nonseparable errors, a new random effects semiparametric model (RESPM) can be established to concurrently capture the linearity and nonlinearity of the covariates, spatial correlation and serial correlation of errors, and individual random effects. This study explores its PQMLE and hypothesis testing issues from theoretical, simulation, and application perspectives. Our proposed model is different from the model proposed by Li et al. [32], which assumes spatially and serially correlated errors are separable. The difference in model structure leads to differences in the following aspects: (1) The parameter spaces (i.e., stationarity conditions) of the two models are different. (2) There are differences in the estimation process, assumption conditions and theorem proofs.

The remainder of this paper is organized as follows: In Section 2, a RESPM with serially and spatially correlated nonseparable error is introduced, and the estimation and testing procedures are established. Section 3 lists several regular assumptions and theorems for the estimators and test statistic. In Sections 4 and 5, numerical simulations and real data analyses are presented. Finally, Section 6 summarizes the results. The Appendix contains the proofs of the lemmas and theorems.

Notation. Some important symbols and their definitions that appear throughout the paper are given in Table 1.

Table 1. Some important symbols and their definitions.

Symbol	Definition
ω_{max}	Maximum eigenvalue of matrix W
ω_{min}	Minimum eigenvalue of matrix W
$ \rho $	The absolute value of ρ
$ \Sigma $	The determinant of matrix Σ
Σ'	The transpose of matrix Σ
Σ^{-1}	The inverse of matrix Σ
$tr(\Sigma)$	The Trace of matrix Σ
$diag(a_1, \dots, a_n)$	Diagonal matrix composed of elements $a_1, \dots, a_n$
$\ S\ $	$tr(S'S)^{\frac{1}{2}} = \{\sum_{ij} s_{ij} ^2\}^{\frac{1}{2}}$
$\xrightarrow{D}$	Convergence in distribution
$\xrightarrow{P}$	Convergence in probability
<i>iid</i>	Independent identically distributed
$K * K$	The convolution of function K
$\otimes$	Kronecker product
$\equiv$	“Defined as” or “marked as”
$\mathcal{N}(0, \Sigma)$	Normal distribution with mean 0 and variance Σ
χ_{df}^2	χ^2 distribution with degree of freedom df

2. Estimation and Testing

2.1. The Model

The form of RESPM with serially and spatially correlated nonseparable errors is as follows:

$$y_{it} = z'_{it}\gamma + f(l_{it}) + b_i + v_{it}, \quad (1)$$

where i ($1 \leq i \leq N$) denotes the i -th spatial unit; t ($1 \leq t \leq T$) denotes the t -th period; y_{it} denotes the observation of the response variable; $z_{it} = (z_{it1}, z_{it2}, \dots, z_{itp})'$ and $l_{it} = (l_{it1}, l_{it2}, \dots, l_{itq})'$ denote observations of p - and q -dimensional covariates, respectively; $\gamma = (\gamma_1, \gamma_2, \dots, \gamma_p)'$ is a p -dimensional regression coefficient vector of z_{it} ; $f(\cdot)$ is an unknown nonparametric function; b_i represents the i -th individual random effect; and v_{it} represents the it -th error term and is independent of b_i .

Let $Y_t = (y_{1t}, \dots, y_{Nt})'$, $Z_t = (z_{1t}, \dots, z_{Nt})'$, $F_t = (f(l_{1t}), \dots, f(l_{Nt}))'$, and $b = (b_1, \dots, b_N)'$. Then, model (1) can be expressed in the following matrix form

$$Y_t = Z_t\gamma + F_t + b + v_t, \quad (2)$$

and its error term follows a serially and spatially correlated nonseparable error structure.

$$v_t = \rho_1 W v_t + \rho_2 v_{t-1} + \epsilon_t, \quad (3)$$

where $v_t = (v_{1t}, \dots, v_{Nt})'$ and $W = (w_{ij})_{N \times N}$ is predetermined spatial weights matrix; $\epsilon_t = (\epsilon_{1t}, \dots, \epsilon_{Nt})'$, $\epsilon_{it} \sim IID(0, \xi_{\epsilon}^2)$, and $b_i \sim IID(0, \xi_b^2)$, ξ_{ϵ}^2 , ξ_b^2 are the variances of ϵ_{it} and b_i , respectively; and ρ_1 and ρ_2 denote the spatial and serial correlation coefficients,

respectively. Similar to Elhorst [29], model stationarity requires not only $|\rho_2| < 1$ and $1/\omega_{min} < \rho_1 < 1/\omega_{max}$ but also

$$\begin{cases} |\rho_2| < 1 - \rho_1\omega_{max}, & \text{if } \rho_1 \geq 0 \\ |\rho_2| > 1 - \rho_1\omega_{min}, & \text{if } \rho_1 < 0 \end{cases}$$

where ω_{max} and ω_{min} denote the largest and smallest eigenvalues of W , respectively. Furthermore, let $\delta_0 = (\rho_{10}, \rho_{20}, \gamma'_0, \xi_{\epsilon 0}^2, \xi_{b 0}^2)'$ be the true values of $\delta = (\rho_1, \rho_2, \gamma', \xi_{\epsilon}^2, \xi_b^2)'$; $f_0(\iota)$ is the true value of $f(\iota)$.

2.2. PQMLE

To ensure homogeneity of variance, (3) can be rewritten as

$$Cv_t = \rho_2 v_{t-1} + v_t, \quad (4)$$

where $C = C(\rho_1) \equiv I_N - \rho_1 W$, I_N denotes an $N \times N$ identity matrix. By repeated substitution, we obtain the following equation: $Cv_t = v_t + \rho_2 C^{-1}v_{t-1} + \rho_2^2 C^{-2}v_{t-2} + \dots$ because C is nonsingular. Therefore, $E(Cv_t) = 0$, and the variance of Cv_t is

$$\begin{aligned} \text{Var}(Cv_t) &= E(v_t v_t') + \rho_2^2 C^{-1} E(v_{t-1} v_{t-1}') C'^{-1} + \rho_2^4 C^{-2} E(v_{t-2} v_{t-2}') C'^{-2} + \dots \\ &= \xi_{\epsilon}^2 (I_N + \rho_2^2 C^{-1} C'^{-1} + \rho_2^4 C^{-2} C'^{-2} + \dots) \\ &= \xi_{\epsilon}^2 [I_N - \rho_2^2 (C' C)^{-1}]^{-1}. \end{aligned} \quad (5)$$

Substituting $v_t + b = Y_t - F_t - Z_t \gamma$ and $v_{t-1} + b = Y_{t-1} - F_{t-1} - Z_{t-1} \gamma$ into (3), we obtain

$$CY_t = C(F_t + Z_t \gamma) + \rho_2(Y_{t-1} - F_{t-1} - Z_{t-1} \gamma) + \epsilon_t + b, \quad t > 1.$$

Because $E(v_t) = 0$ and $\text{Var}(v_t) = \xi_{\epsilon}^2 C^{-1} [I_N - \rho_2^2 (C' C)^{-1}]^{-1} C'^{-1}$, model (2) for the first period of Y_t can be written as

$$\Gamma^{1/2} CY_1 = \Gamma^{1/2} CF_1 + \Gamma^{1/2} CZ_1 \gamma + \Gamma^{1/2} CUb + \epsilon_1 + b,$$

where $\Gamma = I_N - \rho_2^2 (C' C)^{-1}$ and $\Gamma^{1/2}$ may potentially be the Cholesky or spectral decomposition of Γ ; $U = I_N \otimes I_T$, $\otimes$ denotes the Kronecker product. Letting $\eta_t = \epsilon_t + b$, then $\eta = Ub + \epsilon$, and the covariance matrix of η can be calculated as follows:

$$\Sigma = \text{Var}(Ub + \epsilon) = \xi_{\epsilon}^2 I + \xi_b^2 I_N \otimes (l_T l_T'), \quad |\Sigma| = \xi_{\epsilon}^{2N(T-1)} (\xi_b^2 + T\xi_{\epsilon}^2)^N,$$

$$\Sigma^{-1} = \frac{1}{\xi_{\epsilon}^2} I + \left(\frac{1}{\xi_{\epsilon}^2 + T\xi_b^2} - \frac{1}{T\xi_b^2} \right) I_N \otimes \left(\frac{1}{T} l_T l_T' \right),$$

where $\eta = (\eta_1, \dots, \eta_T)'$ and $\epsilon = (\epsilon_1, \dots, \epsilon_T)'$; I denotes the $NT \times NT$ identity matrix; and l_T is a T -dimensional vector consisting of ones. Hence, the log quasi-likelihood function of the proposed model is obtained as follows:

$$\begin{aligned} \log L(\delta) &= -\frac{NT}{2} \log(2\pi) + \frac{1}{2} \log |\Gamma| + T \log |C| - \frac{N(T-1)}{2} \log \xi_{\epsilon}^2 - \frac{N}{2} \log (\xi_{\epsilon}^2 + T\xi_b^2) \\ &\quad - \frac{\sum_{t=2}^T [C(Y_t - Z_t \gamma - F_t) - \rho_2(Y_{t-1} - Z_{t-1} \gamma - F_{t-1})]' H [C(Y_t - Z_t \gamma - F_t) - \rho_2(Y_{t-1} - Z_{t-1} \gamma - F_{t-1})]}{2(\xi_{\epsilon}^2 + T\xi_b^2)} \\ &\quad - \frac{\sum_{t=2}^T [C(Y_t - Z_t \gamma - F_t) - \rho_2(Y_{t-1} - Z_{t-1} \gamma - F_{t-1})]' (I - H) [C(Y_t - Z_t \gamma - F_t) - \rho_2(Y_{t-1} - Z_{t-1} \gamma - F_{t-1})]}{2\xi_{\epsilon}^2} \\ &\quad - \frac{(Y_1 - Z_1 \gamma - F_1)' C \Gamma^{1/2'} (I - H) \Gamma^{1/2} C (Y_1 - Z_1 \gamma - F_1)}{2\xi_{\epsilon}^2} - \frac{(Y_1 - Z_1 \gamma - F_1)' C \Gamma^{1/2'} H \Gamma^{1/2} C (Y_1 - Z_1 \gamma - F_1)}{2(\xi_{\epsilon}^2 + T\xi_b^2)}, \end{aligned} \quad (6)$$

where $H = I_N \otimes (T^{-1} l_T l_T')$.

Because F_t is unknown, Equation (6) cannot be maximized to obtain the quasi-maximum likelihood estimators of the unknowns. To solve this issue, the unknowns

of the model are estimated by combining PQMLE and Working Independence Theory (Cai [35]). The primary steps are as follows:

Step 1. Let $\hat{f}_{IN}(\iota)$ be the initial estimator of $f(\iota)$. We further assume that γ is known and let $\hat{f}_{IN}(\iota) = \hat{a}_1$. Then, $\hat{\phi} = (\hat{a}_1, (S\hat{a}_2)')'$ can be obtained as

$$\hat{\phi} = \arg \min_{a_1, a_2} \frac{1}{NT} [\tilde{Y} - \tilde{\iota}(\iota)\phi]' K_S(\iota) [\tilde{Y} - \tilde{\iota}(\iota)\phi],$$

where a_2 is the first derivative of $f(\iota)$, $\tilde{Y} = (\tilde{y}_{11}, \dots, \tilde{y}_{1T}, \dots, \tilde{y}_{N1}, \dots, \tilde{y}_{NT})'$, $\tilde{y}_{it} = y_{it} - z'_{it}\gamma$, $\tilde{\iota}(\iota) = [\iota_{11}(\iota), \dots, \iota_{1T}(\iota), \dots, \iota_{N1}(\iota), \dots, \iota_{NT}(\iota)]'$, $\iota_{it}(\iota) = [1, S^{-1}(\iota_{it} - \iota)]'$; $|S|$ is the determinant of $S = \text{diag}(s_1, \dots, s_q)$, $K_S(\iota) = \text{diag}[k_S(\iota_{11} - \iota), \dots, k_S(\iota_{1T} - \iota), \dots, k_S(\iota_{N1} - \iota), \dots, k_S(\iota_{NT} - \iota)]$, $k_S(\iota) = |S|^{-1}k(S^{-1}\iota)$; and $k(\cdot)$ is a q -dimensional kernel function.

Let $\iota(\iota) = [\tilde{\iota}'(\iota)K_S(\iota)\tilde{\iota}(\iota)]^{-1}\tilde{\iota}'(\iota)K_S(\iota)$, then $\hat{\phi} = \iota(\iota)\tilde{Y}$ and

$$\hat{f}_{IN}(\iota) = \iota'_e(\iota)\tilde{Y}, \quad (7)$$

where $\iota_e(\iota) = e'_0\iota(\iota)$, $e_0 = (1, 0, \dots, 0)'$.

Step 2. Denote $Y_1^* = \Gamma^{1/2}C(Y_1 - \iota_{e,1}Y)$, $Z_1^* = \Gamma^{1/2}C(Z_1 - \iota_{e,1}Z)$, $Y_t^* = C(Y_t - \iota_{e,t}Y) - \rho_2(Y_{t-1} - \iota_{e,t-1}Y)$, $Z_t^* = C(Z_t - \iota_{e,t}Z) - \rho_2(Z_{t-1} - \iota_{e,t-1}Z)$, where $\iota_{e,t}$ is an $N \times NT$ block of $\iota_e(\iota)$ in Equation (7) and $t > 2$. Substituting $\hat{f}_{IN}(\iota)$ into (6), we obtain

$$\begin{aligned} \log L(\delta, \hat{f}_{IN}(\iota)) = & -\frac{1}{2} \log |\Gamma| + T \log |C| - \frac{N(T-1)}{2} \log \hat{\zeta}_\epsilon^2 - \frac{N}{2} \log (\hat{\zeta}_\epsilon^2 + T \hat{\zeta}_b^2) \\ & - \frac{(Y^* - Z^*\gamma)' H (Y^* - Z^*\gamma)}{2 \hat{\zeta}_\epsilon^2} - \frac{(Y^* - Z^*\gamma)' (I - H) (Y^* - Z^*\gamma)}{2 (\hat{\zeta}_\epsilon^2 + T \hat{\zeta}_b^2)}, \end{aligned} \quad (8)$$

where $Y^* = (Y_1^*, \dots, Y_T^*)'$, $Z^* = (Z_1^*, \dots, Z_T^*)'$. Therefore, the initial estimators of γ , $\hat{\zeta}_\epsilon^2$, and $\hat{\zeta}_b^2$ are:

$$\hat{\zeta}_{\epsilon IN}^2 = \frac{1}{N(T-1)} \tilde{Y}^{*'} (I - H) \tilde{Y}^*, \quad (9)$$

$$\hat{\zeta}_{b IN}^2 = \frac{1}{NT} \tilde{Y}^{*'} H \tilde{Y}^* - \frac{1}{T} \hat{\zeta}_{\epsilon IN}^2, \quad (10)$$

$$\hat{\gamma}_{IN} = [Z^{*'} \left(\frac{H}{\hat{\zeta}_{\epsilon IN}^2} + \frac{I - H}{\hat{\zeta}_{\epsilon IN}^2 + T \hat{\zeta}_{b IN}^2} \right) Z^*]^{-1} Z^{*'} \left(\frac{H}{\hat{\zeta}_{\epsilon IN}^2} + \frac{I - H}{\hat{\zeta}_{\epsilon IN}^2 + T \hat{\zeta}_{b IN}^2} \right) Y^*, \quad (11)$$

where $\tilde{Y}^* = (\tilde{Y}_1^*, \dots, \tilde{Y}_T^*)'$, and $\tilde{Y}_t^* = Y_t^* - Z_t^{*'} \gamma$.

Step 3. By substituting (9)–(11) into (8), the concentrated log quasi-likelihood function of ρ_1 and ρ_2 can be calculated as follows:

$$\log \tilde{L}(\rho_1, \rho_2) = -\frac{1}{2} \log |\Gamma| + T \log |C| - \frac{N(T-1)}{2} \log \hat{\zeta}_{\epsilon IN}^2 - \frac{N}{2} \log (\hat{\zeta}_{\epsilon IN}^2 + T \hat{\zeta}_{b IN}^2). \quad (12)$$

Because Equation (12) is a nonlinear function of ρ_1 and ρ_2 , the corresponding estimators obtained by the optimization algorithm are

$$(\hat{\rho}_1, \hat{\rho}_2) = \arg \max_{\rho_1, \rho_2} \log \tilde{L}(\rho_1, \rho_2). \quad (13)$$

Step 4. Substituting $\hat{\rho}_1$ and $\hat{\rho}_2$ into (9)–(11), we obtain the ultimate estimators of γ , $\hat{\zeta}_\epsilon^2$, and $\hat{\zeta}_b^2$ as $\hat{\gamma}$, $\hat{\zeta}_\epsilon^2$, and $\hat{\zeta}_b^2$, respectively. By replacing $\hat{\gamma}$, $\hat{\zeta}_\epsilon^2$, and $\hat{\zeta}_b^2$ in (7), the ultimate estimator $\hat{f}(\iota)$ of $f(\iota)$ are obtained. Hence,

$$\hat{\zeta}_\epsilon^2 = \frac{1}{N(T-1)} \hat{Y}^{*'} (I - H) \hat{Y}^*, \quad (14)$$

$$\hat{\zeta}_b^2 = \frac{1}{NT} \hat{Y}^{*'} H \hat{Y}^* - \frac{1}{T} \hat{\zeta}_\epsilon^2, \quad (15)$$

$$\hat{\gamma} = [\hat{Z}^{*'} (\frac{H}{\hat{\zeta}_\epsilon^2} + \frac{I-H}{\hat{\zeta}_\epsilon^2 + T\hat{\zeta}_b^2}) \hat{Z}^*]^{-1} \hat{Z}^{*'} (\frac{H}{\hat{\zeta}_\epsilon^2} + \frac{I-H}{\hat{\zeta}_\epsilon^2 + T\hat{\zeta}_b^2}) \hat{Y}^*, \quad (16)$$

$$\hat{f}(\iota) = \iota_e'(\iota) \hat{Y}, \quad (17)$$

where $\hat{Y}^* = (\hat{Y}_1^{*'}, \dots, \hat{Y}_T^{*'})'$, $\hat{Y}_t^* = \hat{Y}_t - \hat{Z}_t^{*'} \hat{\gamma}$, $\hat{Y}^* = (\hat{Y}_1^{*'}, \dots, \hat{Y}_T^{*'})'$, $\hat{Z}^* = (\hat{Z}_1^{*'}, \dots, \hat{Z}_T^{*'})'$, $\hat{Y}_1^* = \hat{\Gamma}^{1/2} \hat{C} (Y_1 - \iota_{e,1} Y)$, $\hat{Z}_1^* = \hat{\Gamma}^{1/2} \hat{C} (Z_1 - \iota_{e,1} Z)$, $\hat{Y}_t^* = \hat{C} (Y_t - \iota_{e,t} Y) - \hat{\rho}_2 (Y_{t-1} - \iota_{e,t-1} Y)$, $\hat{Z}_t^* = \hat{C} (Z_t - \iota_{e,t} Z) - \hat{\rho}_2 (Z_{t-1} - \iota_{e,t-1} Z)$, $\hat{\Gamma} = I_N - \hat{\rho}_2^2 (\hat{C}' \hat{C})^{-1}$, $\hat{C} = I_N - \hat{\rho}_1 W$, and $\hat{Y} = Y - Z' \hat{\gamma}$.

2.3. Hypothesis Test for Nonparametric Component

As outlined in Section 2.2, a hypothesis test is constructed to ensure the rationality of nonparametric model specification (1) as follows:

$$H_0 : f(\iota_{it}) = \zeta_0 + \sum_{j=1}^q \zeta_j \iota_{itj} \quad \text{versus} \quad H_1 : f(\iota_{it}) \neq \zeta_0 + \sum_{j=1}^q \zeta_j \iota_{itj},$$

where $\zeta_0, \zeta_1, \dots, \zeta_q$ are constant parameters. Following the generalized F-test statistic introduced by Fan et al. [36], let $\hat{f}(\iota_{it})$ be the estimator obtained by the PQML method in Section 2.2. Let $\tilde{f}(\iota_{it}) = \tilde{\zeta}_0 + \sum_{j=1}^q \tilde{\zeta}_j \iota_{itj}$ and $\tilde{\zeta}$ be the quasi-maximum likelihood estimators. The resulting residual sum of squares can be obtained using the null and alternative hypotheses, respectively:

$$\begin{aligned} RSS(H_0) &= \sum_i^N \sum_{t=1}^T [y_{it} - \tilde{f}(\iota_{it}) - z_{it}' \tilde{\gamma}]^2, \\ RSS(H_1) &= \sum_i^N \sum_{t=1}^T [y_{it} - \hat{f}(\iota_{it}) - z_{it}' \hat{\gamma}]^2. \end{aligned}$$

To test the null hypothesis, a test statistic is conducted as follows:

$$F_{NT} = \frac{NT}{2} \frac{RSS(H_0) - RSS(H_1)}{RSS(H_1)}. \quad (18)$$

3. Asymptotic Theory

3.1. Assumptions

Before presenting the theoretical analysis of the proposed PQMLEs and test statistic, we make some regular assumptions.

Assumption 1. (i) Covariates $\{z_{it}, \iota_{it}\}_{i=1, t=1}^{N, T}$ are nonstochastic regressors which are uniformly bounded (UB) in $\mathcal{Z} \times \Delta$, where $\mathcal{Z}$ and Δ are parameter spaces of z_{it} and ι_{it} , respectively.

(ii) There exist $m_t(\cdot)$ and a continuous function $\varepsilon(\cdot)$ that satisfy $m_t(\cdot) > 0$ and $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \varepsilon(\iota_i) = \int \varepsilon(\iota) m_t(\iota) d\iota$.

Assumption 2. (i) $w_{ii} = 0$, $w_{ij} = O(1/l_N)$, where $W = (w_{ij})_{N \times N}$.

(ii) When $N \rightarrow \infty$, $l_N/N \rightarrow 0$ holds.

(iii) C is nonsingular in parameter space Λ .

(iv) The row and column sums of W and C^{-1} are UB.

Assumption 3. As $N \rightarrow \infty$, $\|S\| \rightarrow 0$, $N\|S\|^2 \rightarrow \infty$, $\|S\|^4 \|S\|^{-1} \rightarrow 0$ and $N\|S\| \|S\|^4 \rightarrow M \in [0, \infty)$, where M is a constant.

Assumption 4. $K(\cdot)$ is a kernel function on R^q with 0 odd order moments.

Assumption 5. The models (1)–(3) are tenable with $\delta_0 = (\rho_{10}, \rho_{20}, \gamma'_0, \tilde{\zeta}_{\epsilon 0}^2, \tilde{\zeta}_{b 0}^2)'$.

Assumption 6. Σ_{δ_0} is a positive definite matrix, as defined in Theorem 2.

Remark 1. Assumption 1 provides the features of the covariates, random error term, and density function. Assumption 2 concerns the basic features of the spatial weights matrix parallels Assumption 3 in Su and Jin [37] and Assumption 5 in Lee [38]. Assumptions 3 and 4 are related to the bandwidth and kernel functions. Specifically, Assumption 3 is similar to Assumption 4 of Hamilton and Truong [39] and Assumption 7 of Su and Ullah [40] and is easily satisfied by considering $S = \text{diag}(s_1, \dots, s_q)$ with $s_i \propto N^{1/(4+q)}$ for $q < 4$. Assumptions 5 and 6 are important for proving the consistency and asymptotic properties of the estimators.

3.2. Asymptotic Properties

The reduced form of y_{it} is obtained as follows:

$$y_{it} = z'_{it}\gamma_0 + f_0(\iota_{it}) + b_i + v_{it}, \quad (19)$$

$$v_{it} = \rho_{10} \sum_{j=1}^N w_{ij}v_{jt} + \rho_{20}v_{i,t-1} + \epsilon_{it}. \quad (20)$$

Note that $\iota_e = (\iota_{e11}, \dots, \iota_{eNT})$, where $\iota_{eit} = \iota_e(\iota_{it})$. Denote a typical entry of $\iota_e(\iota)$ by $\iota_e(\iota_{it}, \iota) = e'[\tilde{\iota}'(\iota)K_S(\iota)\tilde{\iota}(\iota)]^{-1}\tilde{\iota}'_{it}(\iota)K_S(\iota_{it} - \iota)$, where $\tilde{\iota}_{it}(\iota)$ is a typical column of $\tilde{\iota}(\iota)$.

Lemma 1. Upon fulfilling Assumptions 1–5, we have $\|\iota\| = O_p(N^{-1/2}|S|^{-1/2})$.

Hamilton and Truong [39] provided the proof.

Lemma 2. Upon fulfilling Assumptions 1–5, we have

- (i) $\iota_e(\iota_{it}, \iota) = N^{-1}K_S(\iota_{it} - \iota)\bar{m}(\iota)[1 + o_p(1)]$, where $\bar{m}(\iota) = \sum_{t=1}^T m_t(\iota)$.
- (ii) $\lim_{N \rightarrow \infty} P\{N^{-1}[\tilde{\iota}'(\iota)K_S(\iota)\tilde{\iota}(\iota)]_{ij} \leq M\} = 1$ for $\iota \in \Delta, i, j = 1, \dots, q + 1$.
- (iii) $\lim_{N \rightarrow \infty} P\{\sup_{\iota} \max_{1 \leq i \leq N, 1 \leq t \leq T} |\iota_e(\iota_{it}, \iota)| \leq O(NT)^{-1}|S|^{-1}\} = 1$.

Su and Ullah [41] provided the proofs.

Lemma 3. Upon fulfilling Assumptions 1–5, we have $\|(I - \iota)(F_t)_{NT \times NT}\| = O_p(\|S\|^2)$.

Hamilton and Truong [39] provided the proof.

Using the lemmas listed above, we establish the following theorems:

Theorem 1. Fulfillment of Assumptions 1–5 leads to $\hat{\delta} - \delta_0 = o_p(1)$.

Theorem 2. Fulfillment of Assumptions 1–6 leads to

$$\sqrt{NT}(\hat{\delta} - \delta_0) \xrightarrow{D} \mathcal{N}(0, \lim_{N \rightarrow \infty} (\Sigma_{\delta_0}^{-1} + \Sigma_{\delta_0}^{-1}\Omega_{\delta_0}\Sigma_{\delta_0}^{-1})),$$

where $\Sigma_{\delta_0} = -E[\frac{1}{\sqrt{NT}} \frac{\partial^2 \log L(\delta_0)}{\partial \delta \partial \delta'}]$, Ω_{δ_0} is defined in (A7).

Theorem 3. Fulfillment of Assumptions 1–4 leads to

$$\sqrt{N|S|}(\hat{\phi} - \phi) \xrightarrow{D} \mathcal{N}(0, \sigma^2(\phi)),$$

where $\sigma^2(\phi) = \dot{G}'\iota'(\iota)\sigma_{\eta}^2\iota(\iota)\dot{G}$ and the form of $\dot{G}$ is provided in the Appendix.

Theorem 4. Fulfillment of Assumptions 1–4 leads to

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

where $df = r_k c_k |\Delta|/|S|$, $r_k = \frac{K(0) - \frac{1}{2} \int K^2(\iota) d\iota}{\int (K(\iota) - \frac{1}{2} K * K(\iota))^2 d\iota}$, $c_k = K(0) - \frac{1}{2} \int K^2(\iota) d\iota$, Δ is the support set of ι , $K(\iota) = \text{diag}(S^{-1}k(S^{-1}(\iota_{11} - \iota)), \dots, S^{-1}k(S^{-1}(\iota_{NT} - \iota)))$.

Remark 2. This paper focuses on the case with a large N and fixed T . In terms of estimation procedure, there is no restriction on N and T . Therefore, the estimation method can be applied to both fixed and large T cases. However, the asymptotic theory of estimators with both the fixed T and large T cases may be different.

4. Monte Carlo Simulation

4.1. Performance of PQMLEs

This section is devoted to use Monte Carlo simulation to examine the performance of PQMLEs in finite sample cases. The mean square error (MSE), mean absolute deviation error (MADE), and standard deviation (SD) are used as the evaluation criteria, where

$$\text{MADE}_j = Q^{-1} \sum_{q=1}^Q |\hat{f}(\iota_q) - f(\iota_q)|, \quad j = 1, 2, \dots, \text{srt},$$

$\{\iota_q\}_{q=1}^Q$ are fixed nodes in the support set of ι , and srt denotes the number of repetitions. In the subsequent simulations, $K(\iota) = \frac{3}{4\sqrt{5}}(1 - \frac{1}{5}\iota^2)1(\iota^2 \leq 5)$ (Su [42]) and the rule of thumb method (Mack and Silverman [43]) is utilized to select window width and srt is set as 1000.

The simulation data are generated using the following model:

$$\begin{cases} y_{it} = z'_{it}\gamma + f(\iota_{it}) + b_i + v_{it}, \\ v_{it} = \rho_1 \sum_{j=1}^N w_{ij}v_{jt} + \rho_2 v_{i,t-1} + \epsilon_{it}, \quad i = 1, \dots, N, t = 1, \dots, T, \end{cases} \quad (21)$$

where N, T are the numbers of spatial unit and time period, respectively. In order to study the performance of PQMLEs in different finite sample cases, we take $N = 40, 60, 80$ and $T = 10, 15, 20$ respectively. $z_{it} \sim U[-2, 2]$, $\iota_{it} \sim U[-3, 3]$, $b_i \sim i.i.d.\mathcal{N}(0, 0.5)$, $\epsilon_{it} \sim i.i.d.\mathcal{N}(0, 0.5)$, $f(\iota_{it}) = \sin(-\frac{1}{2}\iota_{it}^2) + 0.6\iota_{it} + 1$, $\gamma = 1$, respectively. According to the parameter space of ρ_1 and ρ_2 , we set $\rho_1 = 0.2, 0.7$, and $\rho_2 = 0.4, 0.5$, respectively. $W = (w_{ij})_{N \times N}$ is replaced with the Rook weights matrix (Anselin [22]). The simulation results are listed in Figure 1 and Tables 2–7.

Figure 1 shows the trajectories of $\hat{f}(\iota)$ and its 95% confidence interval for $\rho_1 = 0.2$ and $\rho_2 = 0.7$. In Figure 1, the green solid, red solid, and black dashed curves denote $\hat{f}(\iota)$, $f(\iota)$, and the corresponding 95% confidence intervals, respectively. Upon examining each curve and subgraph carefully, clearly, $\hat{f}(\iota)$ is quite close to $f(\iota)$ and its confidence bandwidth is tight, indicating that the PQMLE of nonparametric component works effectively on finite samples.

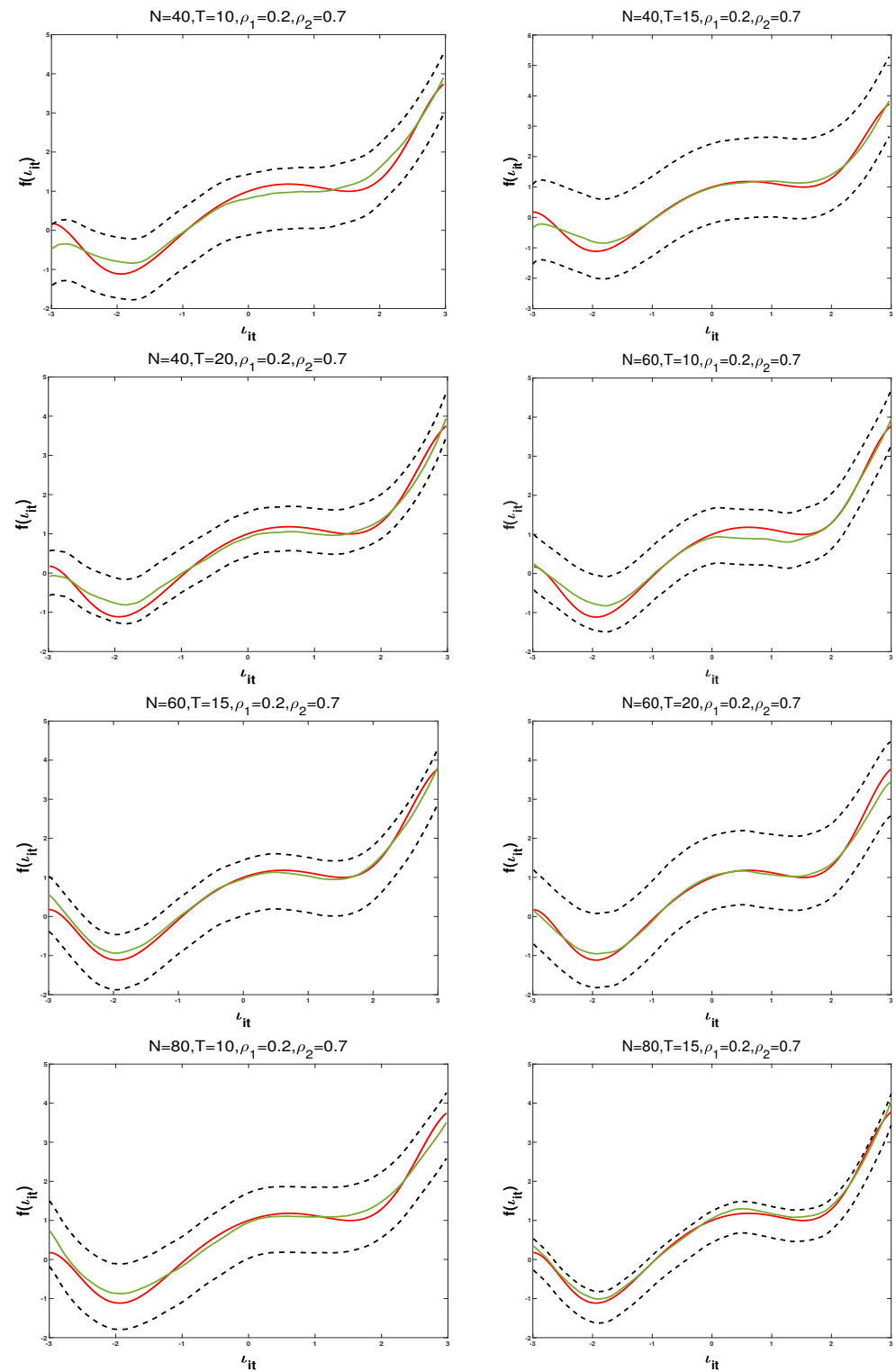


Figure 1. Cont.

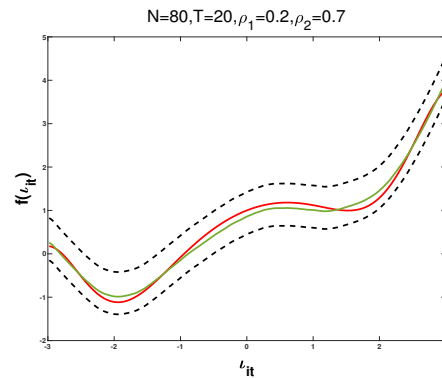


Figure 1. $\hat{f}(t)$ and its 95% confidence intervals with $\rho_1 = 0.2$ and $\rho_2 = 0.7$ under different N and T .

Table 2. Parametric simulation results with $N = 40$.

			T = 10		T = 15		T = 20	
	Parameter	True	Mean	MSE	Mean	MSE	Mean	MSE
$\rho_1 = 0.2$ $\rho_2 = 0.7$	$\hat{\rho}_1$	0.2000	0.1746	0.0017	0.2113	0.0015	0.2046	0.0012
	$\hat{\rho}_2$	0.7000	0.7646	0.0014	0.6786	0.0009	0.6895	0.0007
	$\hat{\gamma}$	1.6000	1.6034	0.0022	1.6010	0.0016	1.6008	0.0013
	$\hat{\zeta}_e^2$	0.5000	0.4746	0.0346	0.4932	0.0214	0.5035	0.0075
	$\hat{\zeta}_b^2$	0.5000	0.5512	0.0154	0.5313	0.0045	0.5054	0.0014
$\rho_1 = 0.4$ $\rho_2 = 0.5$	$\hat{\rho}_1$	0.4000	0.4279	0.0038	0.3870	0.0028	0.3889	0.0020
	$\hat{\rho}_2$	0.5000	0.5590	0.0029	0.5589	0.0025	0.5330	0.0019
	$\hat{\gamma}$	1.6000	1.6005	0.0009	1.6003	0.0007	1.5998	0.0004
	$\hat{\zeta}_e^2$	0.5000	0.5700	0.0242	0.5353	0.0212	0.5243	0.0205
	$\hat{\zeta}_b^2$	0.5000	0.5425	0.0351	0.5166	0.0138	0.5138	0.0108
$\rho_1 = 0.5$ $\rho_2 = 0.4$	$\hat{\rho}_1$	0.5000	0.5287	0.0037	0.4722	0.0032	0.4869	0.0031
	$\hat{\rho}_2$	0.4000	0.4640	0.0031	0.4264	0.0030	0.4113	0.0028
	$\hat{\gamma}$	1.6000	1.5981	0.0007	1.6005	0.0003	1.6004	0.0002
	$\hat{\zeta}_e^2$	0.5000	0.5390	0.0176	0.5184	0.0141	0.4935	0.0034
	$\hat{\zeta}_b^2$	0.5000	0.5615	0.0177	0.4820	0.0119	0.5015	0.0039
$\rho_1 = 0.7$ $\rho_2 = 0.2$	$\hat{\rho}_1$	0.7000	0.6667	0.0055	0.7092	0.0041	0.7054	0.0040
	$\hat{\rho}_2$	0.2000	0.2860	0.0052	0.2319	0.0043	0.2038	0.0037
	$\hat{\gamma}$	1.6000	1.6020	0.0008	1.5996	0.0003	1.5997	0.0002
	$\hat{\zeta}_e^2$	0.5000	0.5446	0.0051	0.5372	0.0024	0.5232	0.0014
	$\hat{\zeta}_b^2$	0.5000	0.5578	0.0152	0.5381	0.0091	0.5295	0.0027

Table 3. Parametric simulation results with $N = 60$.

			T = 10		T = 15		T = 20	
	Parameter	True	Mean	MSE	Mean	MSE	Mean	MSE
$\rho_1 = 0.2$ $\rho_2 = 0.7$	$\hat{\rho}_1$	0.2000	0.2130	0.0009	0.2081	0.0007	0.2007	0.0005
	$\hat{\rho}_2$	0.7000	0.7171	0.0016	0.6878	0.0014	0.6954	0.0004
	$\hat{\gamma}$	1.6000	1.5981	0.0011	1.6006	0.0010	1.6002	0.0008
	$\hat{\zeta}_e^2$	0.5000	0.4490	0.0287	0.4773	0.0213	0.5214	0.0057
	$\hat{\zeta}_b^2$	0.5000	0.5626	0.0074	0.5441	0.0071	0.4947	0.0069
$\rho_1 = 0.4$ $\rho_2 = 0.5$	$\hat{\rho}_1$	0.4000	0.3912	0.0017	0.4074	0.0013	0.3949	0.0011
	$\hat{\rho}_2$	0.5000	0.5082	0.0018	0.5045	0.0015	0.5039	0.0012
	$\hat{\gamma}$	1.6000	1.6007	0.0005	1.6004	0.0003	1.5999	0.0002
	$\hat{\zeta}_e^2$	0.5000	0.5594	0.0205	0.4900	0.0200	0.5099	0.0195
	$\hat{\zeta}_b^2$	0.5000	0.4714	0.0090	0.4847	0.0087	0.5077	0.0079

Table 3. Cont.

	Parameter	True	T = 10		T = 15		T = 20	
			Mean	MSE	Mean	MSE	Mean	MSE
$\rho_1 = 0.5$ $\rho_2 = 0.4$	$\hat{\rho}_1$	0.5000	0.4730	0.0024	0.4823	0.0018	0.4858	0.0016
	$\hat{\rho}_2$	0.4000	0.4322	0.0022	0.4273	0.0017	0.4139	0.0014
	$\hat{\gamma}$	1.6000	1.6005	0.0003	1.6003	0.0002	1.5998	0.0001
	$\hat{\zeta}_e^2$	0.5000	0.5591	0.0071	0.5125	0.0065	0.5072	0.0026
	$\hat{\zeta}_b^2$	0.5000	0.4851	0.0118	0.4856	0.0103	0.5100	0.0030
$\rho_1 = 0.7$ $\rho_2 = 0.2$	$\hat{\rho}_1$	0.7000	0.6808	0.0026	0.7130	0.0024	0.6876	0.0019
	$\hat{\rho}_2$	0.2000	0.2189	0.0025	0.2124	0.0023	0.1970	0.0016
	$\hat{\gamma}$	1.6000	1.6010	0.0003	1.6008	0.0002	1.6007	0.0001
	$\hat{\zeta}_e^2$	0.5000	0.5509	0.0035	0.5214	0.0020	0.5085	0.0012
	$\hat{\zeta}_b^2$	0.5000	0.5556	0.0086	0.5302	0.0084	0.5170	0.0019

Table 4. Parametric simulation results with N = 80.

	Parameter	True	T = 10		T = 15		T = 20	
			Mean	MSE	Mean	MSE	Mean	MSE
$\rho_1 = 0.2$ $\rho_2 = 0.7$	$\hat{\rho}_1$	0.2000	0.2066	0.0004	0.2048	0.0004	0.2003	0.0001
	$\hat{\rho}_2$	0.7000	0.6909	0.0006	0.6978	0.0003	0.6989	0.0002
	$\hat{\gamma}$	1.6000	1.5991	0.0013	1.5997	0.0007	1.6002	0.0006
	$\hat{\zeta}_e^2$	0.5000	0.4755	0.0079	0.5193	0.0047	0.4967	0.0046
	$\hat{\zeta}_b^2$	0.5000	0.5361	0.0063	0.5088	0.0061	0.5004	0.0039
$\rho_1 = 0.4$ $\rho_2 = 0.5$	$\hat{\rho}_1$	0.4000	0.4012	0.0011	0.4004	0.0009	0.3999	0.0007
	$\hat{\rho}_2$	0.5000	0.5092	0.0010	0.5081	0.0008	0.5055	0.0007
	$\hat{\gamma}$	1.6000	1.5986	0.0004	1.6010	0.0003	1.6003	0.0002
	$\hat{\zeta}_e^2$	0.5000	0.4763	0.0204	0.4857	0.0157	0.4961	0.0097
	$\hat{\zeta}_b^2$	0.5000	0.5595	0.0087	0.4813	0.0074	0.4899	0.0068
$\rho_1 = 0.5$ $\rho_2 = 0.4$	$\hat{\rho}_1$	0.5000	0.4933	0.0018	0.4997	0.0013	0.4998	0.0010
	$\hat{\rho}_2$	0.4000	0.4401	0.0017	0.4200	0.0014	0.4064	0.0013
	$\hat{\gamma}$	1.6000	1.5991	0.0002	1.5994	0.0002	1.6002	0.0001
	$\hat{\zeta}_e^2$	0.5000	0.5034	0.0067	0.5008	0.0048	0.5004	0.0021
	$\hat{\zeta}_b^2$	0.5000	0.4628	0.0066	0.4767	0.0053	0.4802	0.0025
$\rho_1 = 0.7$ $\rho_2 = 0.2$	$\hat{\rho}_1$	0.7000	0.7272	0.0028	0.7258	0.0019	0.7119	0.0012
	$\hat{\rho}_2$	0.2000	0.1840	0.0046	0.1879	0.0039	0.1927	0.0013
	$\hat{\gamma}$	1.6000	1.5993	0.0008	1.6004	0.0005	1.6003	0.0001
	$\hat{\zeta}_e^2$	0.5000	0.5043	0.0031	0.5005	0.0018	0.4997	0.0011
	$\hat{\zeta}_b^2$	0.5000	0.5191	0.0048	0.4835	0.0034	0.5127	0.0010

Table 5. Nonparametric simulation results for medians and SDs of MADEs with N = 40.

			T = 10	T = 15	T = 20
$\rho_1 = 0.2$	$\rho_2 = 0.7$	Median	0.3000	0.2635	0.2198
		SD	0.2826	0.2588	0.2518
$\rho_1 = 0.4$	$\rho_2 = 0.5$	Median	0.2430	0.2365	0.2313
		SD	0.2939	0.2584	0.1742
$\rho_1 = 0.5$	$\rho_2 = 0.4$	Median	0.3013	0.2779	0.2281
		SD	0.2089	0.2085	0.1833
$\rho_1 = 0.7$	$\rho_2 = 0.2$	Median	0.3635	0.3627	0.3502
		SD	0.3116	0.3012	0.2903

Table 6. Nonparametric simulation results for medians and SDs of MADEs with $N = 60$.

			T = 10	T = 15	T = 20
$\rho_1 = 0.2$	$\rho_2 = 0.7$	Median	0.2656	0.2532	0.2145
		SD	0.2088	0.2013	0.1466
$\rho_1 = 0.4$	$\rho_2 = 0.5$	Median	0.2099	0.2022	0.1768
		SD	0.1872	0.1666	0.1653
$\rho_1 = 0.5$	$\rho_2 = 0.4$	Median	0.2917	0.2723	0.2188
		SD	0.2044	0.2032	0.1809
$\rho_1 = 0.7$	$\rho_2 = 0.2$	Median	0.3550	0.3365	0.3244
		SD	0.2857	0.2781	0.2235

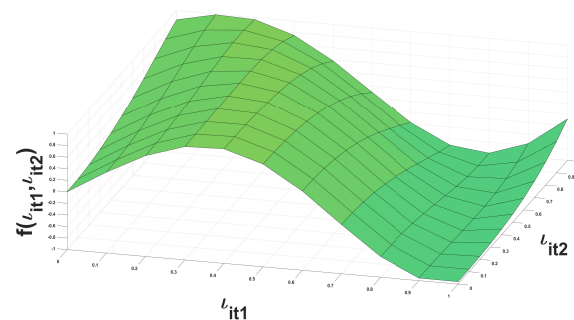
Table 7. Nonparametric simulation results for medians and SDs of MADEs with $N = 80$.

			T = 10	T = 15	T = 20
$\rho_1 = 0.2$	$\rho_2 = 0.7$	Median	0.2425	0.2415	0.1743
		SD	0.2049	0.1768	0.0952
$\rho_1 = 0.4$	$\rho_2 = 0.5$	Median	0.1991	0.1855	0.1701
		SD	0.1807	0.1594	0.1530
$\rho_1 = 0.5$	$\rho_2 = 0.4$	Median	0.2779	0.2392	0.2082
		SD	0.2008	0.1752	0.1747
$\rho_1 = 0.7$	$\rho_2 = 0.2$	Median	0.3074	0.3071	0.2835
		SD	0.2008	0.1976	0.1598

In Tables 2–7, we observe the following: First, the $MSEs$ of $\hat{\rho}_1$, $\hat{\rho}_2$, and $\hat{\gamma}$ are extremely small and decrease with an increase in N . The $MSEs$ of $\hat{\xi}_e^2$ and $\hat{\xi}_b^2$ are slightly larger than those of the other parameter estimates, but decrease with an increase in N or T . When N is fixed and T gradually increases, the behavior of the parameter estimates resembles that of N varying under a fixed T . Second, the *Medians* and *SDs* for *MADEs* of $\hat{f}(t)$ decrease as T or N increases. Thus, the PQMLEs are convergent.

4.2. Monte Carlo Simulation of Multivariate Nonparametric Estimates

A two-dimensional nonparametric function is selected for the simulation experiment. Considering model (21), we set $f(l_{it}) = f(l_{it1}, l_{it2}) = \sin(l_{it1}^2 + l_{it2}^2 + 4l_{it1})$, $l_{it1} \sim U[0, 1]$, $l_{it2} \sim U[0, 1]$, $\rho_1 = 0.2$, $\rho_2 = 0.7$, $N = 40$, and $T = 10$. The settings of the other variables in this simulation are the same as those of the model in Section 4.1. The simulation results indicate that the parametric estimation results are similar to those described in Section 4.1. To save space, we only present the nonparametric simulation results. Figures 2 and 3 show the trajectories of $f(t)$ and $\hat{f}(t)$, where Figure 2 is $f(t)$ and Figure 3 is $\hat{f}(t)$. The median and SD values of $\hat{f}(t)$ are 0.2987 and 0.1328, respectively. The trajectory of $\hat{f}(t)$ is similar to the real trajectory, and the fitting effect is acceptable.

**Figure 2.** The trajectory of $f(\cdot)$.

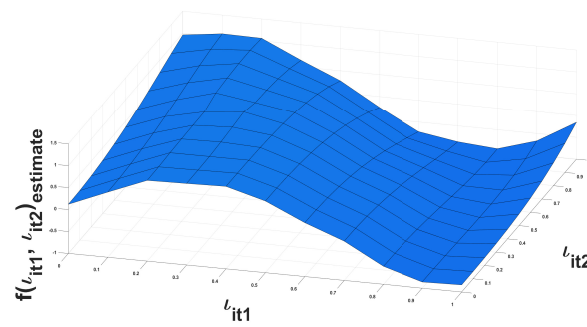


Figure 3. The trajectory of $\hat{f}(\cdot)$.

4.3. Simulation Results of Different Models

To examine the necessity of including $f(l)$, ρ_1 , and ρ_2 of the model (21), four new models are established by successively eliminating $f(l)$, ρ_1 , ρ_2 , ρ_1 , and ρ_2 from model (21) as follows:

$$\begin{cases} y_{it} = z'_{it}\gamma + b_i + v_{it}, \\ v_{it} = \rho_1 \sum_{j=1}^N w_{ij}v_{jt} + \rho_2 v_{i,t-1} + \epsilon_{it}, \end{cases} \quad (22)$$

$$\begin{cases} y_{it} = z'_{it}\gamma + f(l_{it}) + b_i + v_{it}, \\ v_{it} = \rho_2 v_{i,t-1} + \epsilon_{it}, \end{cases} \quad (23)$$

$$\begin{cases} y_{it} = z'_{it}\gamma + f(l_{it}) + b_i + v_{it}, \\ v_{it} = \rho_1 \sum_{j=1}^N w_{ij}v_{jt} + \epsilon_{it}, \end{cases} \quad (24)$$

$$y_{it} = z'_{it}\gamma + f(l_{it}) + b_i + \epsilon_{it}, \quad (25)$$

where all the variables of the aforementioned four models have the same setting as in model (21). To save space, the following simulation results only present the cases of $\rho_1 = 0.4$, $\rho_2 = 0.5$, $N = 40$, $T = 10, 15$, and $srt = 1000$: The PQMLE results for the parameters of models (21)–(25) are provided in Table 8.

Table 8 displays the parametric estimation results for models (21)–(25), where GRM represents the growth rate of MSE which is compared with the MSE of model (21). From Table 7, the following facts can be deduced: (1) The Means of the parameter estimates in model (21) are closer to the real value as sample size increases. The Means of parameter estimates in models (22)–(25) approach the real value with increasing sample size except for $\hat{\rho}_1$, $\hat{\rho}_2$, and $\hat{\xi}_b^2$ in model (22), $\hat{\gamma}$ and $\hat{\xi}_b^2$ in model (23), $\hat{\gamma}$ and $\hat{\xi}_b^2$ in model (24), and all the parameters in model (25). (2) Compared to model (21), the MSEs of each parameter in models (22)–(25) increase, and for almost all parametric estimates, they decrease with increasing sample size, except for $\hat{\rho}_2$ and $\hat{\xi}_b^2$ in model (22), $\hat{\xi}_b^2$ in model (23), and $\hat{\rho}_1$ in model (24). (3) The GRMs of all parametric estimates are large. However, GRMs do not decrease with increasing sample size for some parameter estimates, such as $\hat{\rho}_2$ and $\hat{\xi}_b^2$ in model (22), $\hat{\xi}_e^2$ and $\hat{\xi}_b^2$ in model (23), and $\hat{\rho}_1$ and $\hat{\xi}_b^2$ in model (24). This implies that consistency of parameter estimation is difficult to ensure in case of model misspecification.

Table 8. Parameter simulation results of the models (22)–(25) with $N = 40$.

		$T = 10$					$T = 15$				
		$\hat{\gamma}$	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\xi}_e^2$	$\hat{\xi}_b^2$	$\hat{\gamma}$	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\xi}_e^2$	$\hat{\xi}_b^2$
model (21)	Mean	1.6005	0.4279	0.5590	0.5700	0.5425	1.6003	0.3870	0.5589	0.5353	0.5166
	MSE	0.0009	0.0038	0.0029	0.0242	0.0351	0.0007	0.0028	0.0025	0.0212	0.0138
model (22)	Mean	1.6036	0.5221	0.3696	0.4879	0.6991	1.6011	0.5386	0.3611	0.4888	0.7302
	MSE	0.2011	0.5039	0.7034	0.0054	0.1184	0.0008	0.1081	0.7051	0.0022	0.9068
	GRM(%)	22244	13160	24155	4.96	237	14.28	3760	28104	3	6471

Table 8. Cont.

		$T = 10$					$T = 15$				
		$\hat{\gamma}$	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\xi}_\epsilon^2$	$\hat{\xi}_b^2$	$\hat{\gamma}$	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\xi}_\epsilon^2$	$\hat{\xi}_b^2$
model (23)	Mean	1.5978	-	0.3610	0.7530	0.3147	1.6024	-	0.3840	0.6282	0.2515
	MSE	0.0309	-	1.0052	1.0206	0.0467	0.0013	-	0.3026	1.0065	1.0143
	GRM(%)	3333	-	34562	4117	33	85.71	-	12004	4647	7250
model (24)	Mean	1.5980	0.5406	-	0.6458	0.4181	1.6024	0.5070	-	0.6282	0.3487
	MSE	0.0016	0.0083	-	0.1107	0.0513	0.0012	0.0124	-	0.0365	0.0294
	GRM(%)	77.77	118.4	-	357.4	46	71.42	342.8	-	72.2	113.0
model (25)	Mean	1.6005	-	-	0.6258	0.4167	1.6025	-	-	0.6477	0.2888
	MSE	0.0013	-	-	0.0472	0.0593	0.0007	-	-	0.0452	0.0153
	GRM(%)	44.44	-	-	1855	68.9	0	-	-	113	10.86

4.4. Performance of the Testing

In this section, to assess the performance of F_{NT} in Section 2.3, the power is calculated. Consider model (21), where the settings of z_{it} , b_i , w_{ij} and ϵ_{it} remain unchanged from those in Section 4.1, namely, $\rho_1 = 0.2$, $\rho_2 = 0.2$, $N = 80$, $T = 8, 16$, $srt = 1000$. Thus, we study the following hypotheses:

$$H_0 : f(\iota_{it}) = \zeta_0 + \zeta_1 \iota_{it} \quad \text{versus} \quad H_1 : f(\iota_{it}) \neq \zeta_0 + \zeta_1 \iota_{it},$$

where $f(\iota_{it})$ denotes a series of nonparametric functions with $f(\iota_{it}) = \tau \sin(3\iota_{it}^2) + 0.6\iota_{it} + 1$, $\iota_{it} \sim U[0.2, 1.6]$ and τ is set to $(0, 0.05, 0.1, 0.2, 0.5)$. $f(\iota_{it}) = \tau \sin(3\iota_{it}^2) + 0.6\iota_{it} + 1$ is a combination of a nonparametric function controlled by τ and a linear function. Moreover, when τ is 0, $f(\iota_{it})$ is a linear function, and as τ increases slightly, $f(\iota_{it})$ is a nonlinear function. All the trajectories of $f(\iota)$ at different τ values are shown in Figure 4, where clearly the trajectory of $f(\iota)$ at $\tau = 0$ is the same as that under the null hypothesis. As τ increases, the trajectory of $f(\iota)$ digresses from the null hypothesis. The powers of F_{NT} are shown in Figure 5.

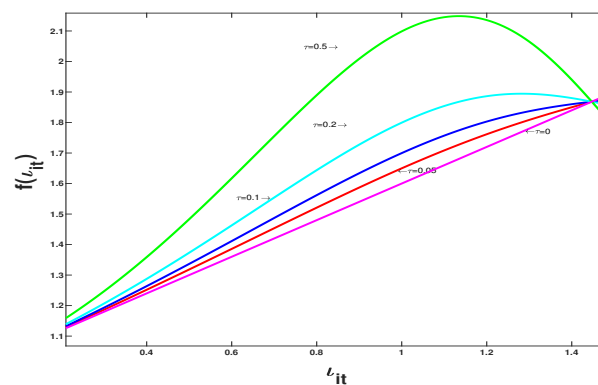


Figure 4. Trajectories of $f(\iota)$ under $\tau = 0, 0.05, 0.1, 0.2$, and 0.5 , respectively.

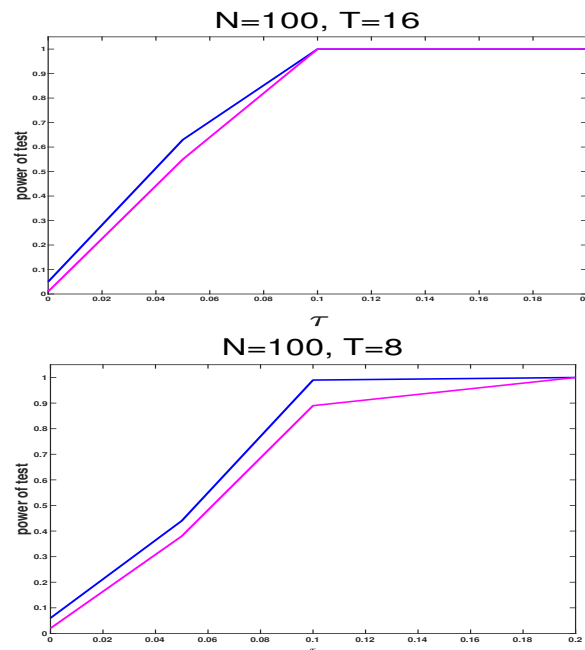


Figure 5. The test size ($\tau = 0$) and the powers ($\tau \neq 0$) of test statistic F_{NT} under $\alpha = 0.01, 0.05$, respectively, with sample size $N * T$.

In Figure 5, (1) when $\tau = 0$, the test size approaches its significance level and surges with an increase in τ . This implies that F_{NT} is susceptible to the alternative hypothesis of the given test problem. (2) For the same N , testing at $T = 16$ outperforms that of $T = 8$. This demonstrates that testing performance improves with increasing sample size.

5. Real Data Analysis

In this section, the proposed estimation and testing methods are demonstrated on the Indonesian rice farm dataset. This dataset is collected by the Agricultural Economic Research Center of the Ministry of Agriculture of Indonesia. It covers 171 farms over six growing seasons (three wet and three dry seasons) and has been widely used in random effects models (Feng and Horrace [44]).

The dataset includes observations of five variables which are rice yield, high-yield varieties, mixed-yield varieties, seed weight, and land area, respectively. Rice yield is taken as response variable and others are taken as covariates whose definitions are given in Table 9.

The testing method proposed in Section 2.3 is used to determine whether a nonlinear connection exists between the covariates and response variable. The outcomes of the F-test are listed in Table 10. Table 10 shows that *land area* (other covariates) exhibits a significant nonlinear (linear) relationship with *rice yield* at a significance level of $\alpha = 0.01$.

Table 9. The description of variables.

Variable	Description
high yield varieties	If high yield varieties are used, value of high yield varieties equals 1, else equals 0.
mixed yield varieties	If mixed yield varieties are used, value of mixed yield varieties equals 1, else equals 0.
seed weight	Seed weight used at t -th season on i -th farm (unit: kilogram)
land area	Number of hectares of rice

Table 10. F-test results with $\alpha = 0.01$.

Variable	High Yield Variety	Mixed Yield Varietie	Seed Weight	Land Area
F-test value	4.43	5.16	1.27	181.64 ***

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

Therefore, the following model is established for the data analysis:

$$\begin{cases} y_{it} = z'_{it}\gamma + f(l_{it}) + b_i + v_{it}, \\ v_{it} = \rho_1 \sum_{j=1}^N w_{ij}v_{jt} + \rho_2 v_{i,t-1} + \epsilon_{it}, \quad 1 \leq i \leq 171, 1 \leq t \leq 6, \end{cases} \quad (26)$$

where y_{it} , z_{it} , and l_{it} denote the i th observation of log(rice), high- and mixed-yield varieties, seed weight, and land area during the t th growing season, respectively. Furthermore, the w_{ij} is determined using the following method (Druska and Horrace [45]):

$$\begin{cases} w_{ij} = 1, & \text{if } i \neq j \\ w_{ij} = 0, & \text{if } i = j \end{cases}$$

where i, j denote the i -th and j -th village, respectively. Moreover, the weights matrix is normalized to ensure the elements of each row sum to 1. Additionally, the weights matrix is assumed time invariant, so the t subscript can be dropped.

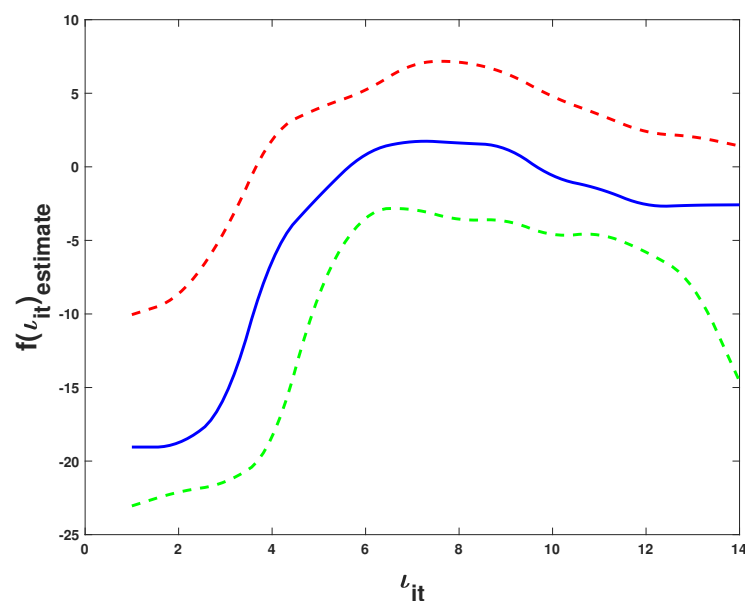
Table 11 shows the parametric estimation results, which reveals that (1) All linear co-variables $\hat{\gamma}_i (i = 1, 2, 3) > 0$ have a positive impact on rice yield. (2) $\hat{\zeta}_\epsilon^2 = 0.3140$ demonstrates that the rice yield is generally steady and less susceptible to exogenous perturbations.

Figure 6 presents $\hat{f}(l)$ and its 95% confidence interval, where the blue short dashed curve denotes $\hat{f}(l)$, and the red and green solid curves correspond to the 95% confidence bands. Clearly, *land area* has a nonlinear effect on rice yield.

Table 11. Parameter estimation results.

Parameter	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\gamma}_3$	$\hat{\zeta}_\epsilon^2$	$\hat{\zeta}_b^2$
Estimate	0.9018 ***	0.0124 ***	0.2818 ***	0.4881 ***	0.0020 ***	0.3140 ***	0.2907 ***

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

**Figure 6.** The nonlinear effect of l_{it} (land area) on y_{it} (log(rice)).

6. Summary

This study mainly explores the PQMLE and F-test of RESPM with serially and spatially correlated nonseparable error. Our model simultaneously captures the linear and nonlinear effects of the covariates of spatially and serially correlated errors and individual random effects. PQMLEs are constructed for unknowns along with test statistic F_{NT} to determine the beingness of nonlinear relationships. Their asymptotic properties are derived under regular conditions. The Monte Carlo results indicate that the proposed estimators and test statistic behave well on finite samples and that ignoring the spatial and serial correlations of errors may lead to inefficient and biased estimators. The proposed estimation and testing techniques are used to analyze Indonesian rice farming data.

The paper can be extended in several ways. We can apply spline estimation method or local polynomial method to approximate the nonparametric function, and combine GMM (Cheng and Chen [46]), quadratic inference function estimation (Qu [47]) or PQMLE to obtain the estimators of unknowns. Additionally, we can also consider extending these methods to other similar models, and the Bayesian analysis, variable selection, and quantile regression analysis of these models are also worth studying. In the future, we may use the proposed estimation and test methods to conduct empirical analysis, for example, investigating the drivers of CO_2 , $PM_{2.5}$ and energy efficiency.

Author Contributions: Formal analysis, S.L. and D.C.; methodology, S.L. and J.C.; software and writing—original draft, S.L.; supervision, writing—review and editing, and funding acquisition, J.C.; data curation, D.C. All authors have read and agreed to the published version of manuscript.

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

Data Availability Statement: Data is available upon request from the authors.

Acknowledgments: The authors thank the editors and reviewers for their hard work, who have made outstanding contributions to the improvement of the quality of this paper.

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

Appendix A

Proof of Theorem 1. Define $Q_{NT} = \max_{\gamma, \hat{\xi}_{\epsilon}^2} E[\log L(\delta, \hat{f}_{IN}(\iota))]$. A straightforward calculation provides

$$Q_{NT}(\rho_1, \rho_2) = -\frac{1}{2} \log |\Gamma| + T \log |C| - \frac{N(T-1)}{2} E[\log \hat{\xi}_{\epsilon IN}^2] - \frac{N}{2} E[\log(\hat{\xi}_{\epsilon IN}^2 + T \hat{\xi}_{b IN}^2)]$$

and

$$\log \tilde{L}(\rho_1, \rho_2) = -\frac{1}{2} \log |\Gamma| + T \log |C| - \frac{N(T-1)}{2} \log \hat{\xi}_{\epsilon IN}^2 - \frac{N}{2} \log(\hat{\xi}_{\epsilon IN}^2 + T \hat{\xi}_{b IN}^2),$$

The method for proving the consistency of $\hat{\delta}$ is adopted from Lee [38] and White [48] (Theorem 3.4), to show

$$\frac{1}{NT} [\log L(\rho_1, \rho_2) - Q_{NT}(\rho_1, \rho_2)] = op(1). \quad (A1)$$

By (11), we have $\frac{1}{NT} [\log L(\rho_1, \rho_2) - Q_n(\rho_1, \rho_2)] = \frac{T-1}{2T} [\log \hat{\xi}_{\epsilon IN}^2 - \log E(\hat{\xi}_{\epsilon IN}^2)] + \frac{1}{2T} [\log(\hat{\xi}_{\epsilon IN}^2 + T \hat{\xi}_{b IN}^2) - E(\log(\hat{\xi}_{\epsilon IN}^2 + T \hat{\xi}_{b IN}^2))]$.

Therefore, to prove (A1), we only need to prove that

$$\hat{\xi}_{\epsilon IN}^2 - E(\hat{\xi}_{\epsilon IN}^2) = op(1) \quad (A2)$$

and

$$\hat{\xi}_{b IN}^2 - E(\hat{\xi}_{b IN}^2) = op(1). \quad (A3)$$

First, we prove that (A2) holds. Recall

$$\hat{\xi}_{\varepsilon IN}^2 = \frac{1}{N(T-1)} \tilde{Y}^{*'} (I - H) \tilde{Y}^*,$$

where $\tilde{Y}^* = (\tilde{Y}_1^{*'}, \dots, \tilde{Y}_T^{*'})'$, $\tilde{Y}_1^* = \Gamma^{1/2} C(Y_1 - \iota_{e,1} Y) - \Gamma^{1/2} C(Z_1 - \iota_{e,1} Z)$, and $\tilde{Y}_t^* = C(Y_t - \iota_{e,t} Y) - \rho_2(Y_{t-1} - \iota_{e,t-1} Y) - [C(Z_t - \iota_{e,t} Z) - \rho_2(Z_{t-1} - \iota_{e,t-1} Z)]$ ($t = 2, \dots, T$). By a straightforward calculation, we obtain:

$$\hat{\xi}_{\varepsilon IN}^2 = D_1 - D_2,$$

where $D_1 = \frac{1}{N(T-1)} \tilde{Y}^{*'} \tilde{Y}^*$ and $D_2 = \frac{1}{N(T-1)} \tilde{Y}^{*'} H \tilde{Y}^*$.

For D_1 ,

$$D_1 = \frac{1}{N(T-1)} \tilde{Y}_1^{*'} \tilde{Y}_1^* + \frac{1}{N(T-1)} \sum_{t=2}^T \tilde{Y}_t^{*'} \tilde{Y}_t^*. \quad (\text{A4})$$

For the first term of (A4), we have

$$\begin{aligned} \frac{1}{N(T-1)} \tilde{Y}_1^{*'} \tilde{Y}_1^* &= \frac{[\Gamma^{1/2} C(Y_1 - \iota_{e,1} Y) - \Gamma^{1/2} C(Z_1 - \iota_{e,1} Z)]' [\Gamma^{1/2} C(Y_1 - \iota_{e,1} Y) - \Gamma^{1/2} C(Z_1 - \iota_{e,1} Z)]}{N(T-1)} \\ &= \frac{[\Gamma_0^{\frac{1}{2}} C_0 F_{10} + \eta_1 - \Gamma_0^{\frac{1}{2}} C_0 \iota_{e,1} (Y - Z\gamma)]'}{N(T-1)} \frac{[\Gamma_0^{\frac{1}{2}} C_0 F_{10} + \eta_1 - \Gamma_0^{\frac{1}{2}} C_0 \iota_{e,1} (Y - Z\gamma)]}{N(T-1)} \end{aligned}$$

and

$$\iota_{e,1} Y = \iota_{e,1} \begin{pmatrix} F_{10} + Z_1 \gamma_0 + C_0^{-1} \Gamma_0^{-\frac{1}{2}} \eta_{10} \\ F_{20} + Z_2 \gamma_0 + \rho_{20} C_0^{-2} \Gamma_0^{-\frac{1}{2}} \eta_{10} + C_0^{-1} \eta_{20} \\ F_{30} + Z_3 \gamma_0 + \rho_{20}^2 C_0^{-3} \Gamma_0^{-\frac{1}{2}} \eta_{10} + \rho_{20} C_0^{-2} \eta_{20} + C_0^{-1} \eta_{30} \\ \vdots \\ F_{T0} + Z_T \gamma_0 + \rho_{20}^{T-1} C_0^{-T} \Gamma_0^{-\frac{1}{2}} \eta_{10} + \rho_{20}^{T-2} C_0^{-T+1} \eta_{20} + \dots + C_0^{-1} \eta_{T0} \end{pmatrix},$$

then

$$\frac{1}{N(T-1)} \tilde{Y}_1^{*'} \tilde{Y}_1^* = D_{11} + D_{12} + D_{13},$$

where

$$\begin{aligned} D_{11} &= \frac{1}{N(T-1)} (\Gamma_0^{\frac{1}{2}} C_0 F_{10} - \Gamma_0^{\frac{1}{2}} C_0 \iota_{e,1} F_0)' (\Gamma_0^{\frac{1}{2}} C_0 F_{10} - \Gamma_0^{\frac{1}{2}} C_0 \iota_{e,1} F_0), \\ D_{12} &= \frac{1}{N(T-1)} (\eta_{10} - \begin{pmatrix} C_0^{-1} \Gamma_0^{-\frac{1}{2}} \eta_{10} \\ \rho_{20} C_0^{-2} \Gamma_0^{-\frac{1}{2}} \eta_{10} + C_0^{-1} \eta_{20} \\ \rho_{20}^2 C_0^{-3} \Gamma_0^{-\frac{1}{2}} \eta_{10} + \rho_{20} C_0^{-2} \eta_{20} + C_0^{-1} \eta_{30} \\ \vdots \\ \rho_{20}^{T-1} C_0^{-T} \Gamma_0^{-\frac{1}{2}} \eta_{10} + \rho_{20}^{T-2} C_0^{-T+1} \eta_{20} + \dots + C_0^{-1} \eta_{T0} \end{pmatrix})' \\ &\quad \times (\Gamma_0^{\frac{1}{2}} C_0 F_{10} - \Gamma_0^{\frac{1}{2}} C_0 \iota_{e,1} F_0), \\ D_{13} &= \frac{1}{N(T-1)} (\eta_{10} - \begin{pmatrix} B_0^{-1} \Gamma_0^{-\frac{1}{2}} \eta_{10} \\ \rho_{20} C_0^{-2} \Gamma_0^{-\frac{1}{2}} \eta_{10} + C_0^{-1} \eta_{20} \\ \rho_{20}^2 C_0^{-3} \Gamma_0^{-\frac{1}{2}} \eta_{10} + \rho_{20} C_0^{-2} \eta_{20} + C_0^{-1} \eta_{30} \\ \vdots \\ \rho_{20}^{T-1} C_0^{-T} \Gamma_0^{-\frac{1}{2}} \eta_{10} + \rho_{20}^{T-2} C_0^{-T+1} \eta_{20} + \dots + C_0^{-1} \eta_{T0} \end{pmatrix})' \end{aligned}$$

$$\times (\eta_{10} - \begin{pmatrix} C_0^{-1} \Gamma_0^{-\frac{1}{2}} \eta_{10} \\ \rho_{20} C_0^{-2} \Gamma_0^{-\frac{1}{2}} \eta_{10} + C_0^{-1} \eta_{20} \\ \rho_{20}^2 C_0^{-3} \Gamma_0^{-\frac{1}{2}} \eta_{10} + \rho_{20} C_0^{-2} \eta_{20} + C_0^{-1} \eta_{30} \\ \vdots \\ \rho_{20}^{T-1} C_0^{-T} \Gamma_0^{-\frac{1}{2}} \eta_{10} + \rho_{20}^{T-2} C_0^{-T+1} \eta_{20} + \dots + C_0^{-1} \eta_{T0} \end{pmatrix}),$$

and $F_0 = (F_{10}, \dots, F_{T0})'$.

For the above expression of D_{11} , we find that the first term converges to the probabilistic mean with certainty.

The above expression for D_{12} can be decomposed as follows:

$$D_{12} = D_{121} - D_{122} - D_{123} + D_{124},$$

$$\text{where } D_{121} = \frac{1}{N(T-1)} \eta'_{10} \Gamma_0^{\frac{1}{2}} C_0 F_{10}, D_{122} = \frac{1}{N(T-1)} \eta'_{10} \Gamma_0^{\frac{1}{2}} C_0 \iota_{e,1} F_0,$$

$$D_{123} = \frac{1}{N(T-1)} (\Gamma_0^{\frac{1}{2}} C_0 \iota_{e,1} \begin{pmatrix} C_0^{-1} \Gamma_0^{-\frac{1}{2}} \eta_{10} \\ \rho_{20} C_0^{-2} \Gamma_0^{-\frac{1}{2}} \eta_{10} + C_0^{-1} \eta_{20} \\ \rho_{20}^2 C_0^{-3} \Gamma_0^{-\frac{1}{2}} \eta_{10} + \rho_{20} C_0^{-2} \eta_{20} + C_0^{-1} \eta_{30} \\ \vdots \\ \rho_{20}^{T-1} C_0^{-T} \Gamma_0^{-\frac{1}{2}} \eta_{10} + \rho_{20}^{T-2} C_0^{-T+1} \eta_{20} + \dots + C_0^{-1} \eta_{T0} \end{pmatrix})' \Gamma_0^{\frac{1}{2}} C_0 F_{10},$$

$$D_{124} = \frac{1}{N(T-1)} (\Gamma_0^{\frac{1}{2}} C_0 \iota_{e,1} \begin{pmatrix} C_0^{-1} \Gamma_0^{-\frac{1}{2}} \eta_{10} \\ \rho_{20} C_0^{-2} \Gamma_0^{-\frac{1}{2}} \eta_{10} + C_0^{-1} \eta_{20} \\ \rho_{20}^2 C_0^{-3} \Gamma_0^{-\frac{1}{2}} \eta_{10} + \rho_{20} C_0^{-2} \eta_{20} + C_0^{-1} \eta_{30} \\ \vdots \\ \rho_{20}^{T-1} C_0^{-T} \Gamma_0^{-\frac{1}{2}} \eta_{10} + \rho_{20}^{T-2} C_0^{-T+1} \eta_{20} + \dots + C_0^{-1} \eta_{T0} \end{pmatrix})' \Gamma_0^{\frac{1}{2}} C_0 \iota_{e,1} F_0.$$

By simple calculations and Lemma 1, we obtain

$$E(D_{121}) = 0$$

and

$$\text{Var}(D_{121}) = \frac{1}{N^2(T-1)^2} \text{var}(\eta_{10}) (\Gamma_0^{\frac{1}{2}} C_0 F_{10})' (\Gamma_0^{\frac{1}{2}} C_0 F_{10}) = O(1).$$

According to Chebyshev's law of large numbers, $D_{121} - E[D_{121}] = op(1)$. Similarly, we obtain $D_{12i} - E[D_{12i}] = op(1)$ ($i = 2, 3, 4$). Therefore, $D_{12} - E[D_{12}] = op(1)$ is proven, and thus, D_1 . The proof of $D_2 - E[D_2] = op(1)$ is similar to that of $D_1 - E[D_1] = op(1)$. To avoid repetition, this is not described in this article. Combined with Assumption 5, the consistency of $\hat{\delta}$ is obtained.

□

Proof of Theorem 2. Taylor expanding (8) at δ_0 , we obtain

$$\frac{\partial \log L(\delta)}{\partial \delta} \Big|_{\delta=\delta_0} + \frac{\partial^2 \log L(\delta)}{\partial \delta \partial \delta} \Big|_{\delta=\tilde{\delta}} (\hat{\delta} - \delta_0) = 0,$$

where $\tilde{\delta} = (\tilde{\rho}_1, \tilde{\rho}_2, \tilde{\gamma}, \tilde{\zeta}_{\tilde{e}}, \tilde{\zeta}_{\tilde{b}})'$ and $\tilde{\delta}$ lies in between $\hat{\delta}$ and δ_0 . By Theorem 1, we have $\tilde{\delta} \xrightarrow{P} \delta_0$.

Denote

$$\frac{\partial \log L(\delta_0)}{\partial \delta} \triangleq \frac{\partial \log L(\delta)}{\partial \delta} \Big|_{\delta=\delta_0}, \quad \frac{\partial^2 \log L(\delta_0)}{\partial \delta \partial \delta'} \triangleq \frac{\partial^2 \log L(\delta)}{\partial \delta \partial \delta'} \Big|_{\delta=\delta_0}, \quad \frac{\partial^2 \log L(\tilde{\delta})}{\partial \delta \partial \delta'} \triangleq \frac{\partial^2 \log L(\delta)}{\partial \delta \partial \delta'} \Big|_{\delta=\tilde{\delta}}.$$

Thus,

$$\sqrt{NT}(\hat{\delta} - \delta_0) = -\left(\frac{1}{NT} \frac{\partial^2 \log L(\tilde{\delta})}{\partial \delta \partial \delta'}\right)^{-1} \frac{1}{\sqrt{NT}} \frac{\partial \log L(\delta_0)}{\partial \delta}.$$

Next, we need to prove

$$\frac{1}{NT} \frac{\partial^2 \log L(\tilde{\delta})}{\partial \delta \partial \delta'} - \frac{1}{NT} \frac{\partial^2 \log L(\delta_0)}{\partial \delta \partial \delta'} = op(1) \quad (\text{A5})$$

and

$$\frac{1}{\sqrt{NT}} \frac{\partial \log L(\delta_0)}{\partial \delta} \xrightarrow{D} \mathcal{N}(0, \Sigma_{\delta_0} + \Omega_{\delta}). \quad (\text{A6})$$

To prove (A5), we need to show that each element of $\frac{1}{NT} \frac{\partial^2 \log L(\tilde{\delta})}{\partial \delta \partial \delta'} - \frac{1}{NT} \frac{\partial^2 \log L(\delta_0)}{\partial \delta \partial \delta'}$ converges to 0 in probability. It can be shown that

$$\begin{aligned} \frac{1}{\sqrt{NT}} \frac{\partial \log L(\delta_0)}{\partial \rho_1} &= -\frac{\rho_{20}^2}{\sqrt{NT}} \text{tr}(\Gamma_0^{-1} C_0^{-3} W) - \frac{\sqrt{T}}{\sqrt{N}} \text{tr}(C_0^{-1} W) + \frac{\eta' A_4' W (I - H) A_1 \eta}{\sqrt{NT} \xi_{\epsilon 0}^2} \\ &\quad + \frac{\eta_1' W' \Gamma_0^{\frac{1}{2}'} H \Gamma_0^{\frac{1}{2}} C_0 \eta_1}{\sqrt{NT} (\xi_{\epsilon 0}^2 + T \xi_{b0}^2)} + \frac{\rho_{20}^2 \eta_1' C_0' [\Gamma_0^{-\frac{1}{2}'} (C_0' C_0)^{-2} W' C_0]' H \Gamma_0^{\frac{1}{2}} C_0 \eta_1 + \eta' A_4' W H A_1 \eta}{\sqrt{NT} (\xi_{\epsilon 0}^2 + T \xi_{b0}^2)} \\ &\quad + \frac{\eta_1' W' \Gamma_0^{\frac{1}{2}'} (I - H) \Gamma_0^{\frac{1}{2}} C_0 \eta_1 + \rho_{20}^2 \eta_1' C_0' [\Gamma_0^{-\frac{1}{2}'} (C_0' C_0)^{-2} W' C_0]' (I - H) \Gamma_0^{\frac{1}{2}} C_0 \eta_1}{\sqrt{NT} \xi_{\epsilon 0}^2}, \\ \frac{1}{\sqrt{NT}} \frac{\partial \log L(\delta_0)}{\partial \rho_2} &= \frac{-\rho_{20}}{\sqrt{NT}} \text{tr}(\Gamma_0^{-1} C_0^{-2}) + \frac{\rho_{20} \eta_1' C_0' \Gamma_0^{-\frac{1}{2}'} (C_0' C_0)^{-1} H \Gamma_0^{\frac{1}{2}} C_0 \eta_1 + \eta' A_2' A_3' H A_1 \eta}{\sqrt{NT} (\xi_{\epsilon 0}^2 + T \xi_{b0}^2)} \\ &\quad + \frac{\rho_{20} \eta_1' C_0' \Gamma_0^{-\frac{1}{2}'} (C_0' C_0)^{-1} (I - H) \Gamma_0^{\frac{1}{2}} C_0 \eta_1 + \eta' A_2' A_3' (I - H) A_1 \eta}{\sqrt{NT} \xi_{\epsilon 0}^2}, \\ \frac{1}{\sqrt{NT}} \frac{\partial \log L(\delta_0)}{\partial \gamma'} &= \frac{Z_1' C_0' \Gamma_0^{\frac{1}{2}'} H \Gamma_0^{\frac{1}{2}} C_0 \eta_1 + Z' (\tilde{C} A_1 - \rho_{20} A_2) H A_1 \eta}{\sqrt{NT} (\xi_{\epsilon 0}^2 + T \xi_{b0}^2)} \\ &\quad + \frac{Z' (\tilde{C} A_1 - \rho_{20} A_2) (I - H) A_1 \eta + Z_1' C_0' \Gamma_0^{\frac{1}{2}'} (I - H) \Gamma_0^{\frac{1}{2}} C_0 \eta_1}{\sqrt{NT} \xi_{\epsilon 0}^2}, \\ \frac{1}{\sqrt{NT}} \frac{\partial \log L(\delta_0)}{\partial \xi_{\epsilon}^2} &= \frac{-\sqrt{N}(T-1)}{2\sqrt{T} \xi_{\epsilon 0}^2} - \frac{\sqrt{N}}{2\sqrt{T} (\xi_{\epsilon 0}^2 + T \xi_{b0}^2)} + \frac{\eta_1' C_0' \Gamma_0^{\frac{1}{2}'} H \Gamma_0^{\frac{1}{2}} C_0 \eta_1}{2\sqrt{NT} (\xi_{\epsilon 0}^2 + T \xi_{b0}^2)^2} \\ &\quad + \frac{\eta_1' C_0' \Gamma_0^{\frac{1}{2}'} (I - H) \Gamma_0^{\frac{1}{2}} C_0 \eta_1}{2\sqrt{NT} \xi_{\epsilon 0}^2} + \frac{\eta' A_1' H A_1 \eta}{2\sqrt{NT} (\xi_{\epsilon 0}^2 + T \xi_{b0}^2)^2} + \frac{\eta' A_1' (I - H) A_1 \eta}{2\sqrt{NT} \xi_{\epsilon 0}^4}, \\ \frac{1}{\sqrt{NT}} \frac{\partial \log L(\delta_0)}{\partial \xi_b^2} &= \frac{\sqrt{T}}{2\sqrt{N} (\xi_{\epsilon 0}^2 + T \xi_{b0}^2)^2} + \frac{\sqrt{T} \eta_1' C_0' \Gamma_0^{\frac{1}{2}'} H \Gamma_0^{\frac{1}{2}} C_0 \eta_1}{2\sqrt{N} (\xi_{\epsilon 0}^2 + T \xi_{b0}^2)^2} + \frac{\sqrt{T} \eta A_1' H A_1 \eta}{2\sqrt{N} (\xi_{\epsilon 0}^2 + T \xi_{b0}^2)^2}, \end{aligned}$$

where $\tilde{C} = I_T \otimes C$ and I_T is $T \times T$ identity matrix.

Consequently, we have

$$\begin{aligned}
\frac{1}{NT} \frac{\partial^2 \log L(\delta_0)}{\partial \rho_1^2} &= -\left(\frac{1}{NT} + \frac{11\rho_{20}^2}{NT} + \frac{2\rho_{20}^2 \xi_{b0}^2}{N}\right) \text{tr}[(C_0^{-1}W)^2] + o(1), \\
\frac{1}{NT} \frac{\partial^2 \log L(\delta_0)}{\partial \rho_1 \partial \rho_2} &= \frac{\rho_{20}^2}{NT} \text{tr}(C_0^{-1}W) - \frac{2\rho_{20}}{NT} \text{tr}(C_0^{-1}W) + o(1), \quad \frac{1}{NT} \frac{\partial^2 \log L(\delta_0)}{\partial \rho_1 \partial \gamma'} = 0, \\
\frac{1}{NT} \frac{\partial^2 \log L(\delta_0)}{\partial \rho_1 \partial \xi_\epsilon^2} &= \frac{-\text{tr}(\Gamma_0 C_0 W)}{NT(\xi_{\epsilon 0}^2 + T\xi_{b0}^2)} - \frac{\rho_{20}^2}{NT} \text{tr}(C_0^{-1}W) + o(1), \\
\frac{1}{NT} \frac{\partial^2 \log L(\delta_0)}{\partial \rho_1 \partial \xi_b^2} &= \frac{-\text{tr}(\Gamma_0 C_0 W) + \rho_{20}^2 \text{tr}(C_0^{-1}W) + \text{tr}(A_4 \tilde{W})}{N(\xi_{\epsilon 0}^2 + T\xi_{b0}^2)} + o(1), \\
\frac{1}{NT} \frac{\partial^2 \log L(\delta_0)}{\partial \rho_2^2} &= \frac{(1-T)\xi_{b0}^2}{\xi_{\epsilon 0}^2} + o(1), \quad \frac{1}{NT} \frac{\partial^2 \log L(\delta_0)}{\partial \rho_2 \partial \gamma'} = 0, \\
\frac{1}{NT} \frac{\partial^2 \log L(\delta_0)}{\partial \rho_2 \partial \xi_\epsilon^2} &= \frac{-1}{T(\xi_{\epsilon 0}^2 + T\xi_{b0}^2)^2} - \frac{\rho_{20}(\xi_{\epsilon 0}^2 + T\xi_{b0}^2)}{\xi_\epsilon^4} + o(1), \\
\frac{1}{NT} \frac{\partial^2 \log L(\delta_0)}{\partial \rho_2 \partial \xi_b^2} &= -\frac{1}{\xi_{\epsilon 0}^2 + T\xi_{b0}^2}, \quad \frac{1}{NT} \frac{\partial^2 \log L(\delta_0)}{\partial \gamma \partial \xi_\epsilon^2} = 0, \quad \frac{1}{NT} \frac{\partial^2 \log L(\delta_0)}{\partial \gamma \partial \xi_b^2} = 0, \\
\frac{1}{NT} \frac{\partial^2 \log L(\delta_0)}{\partial \gamma \partial \gamma'} &= -\frac{Z'_1 C'_0 \Gamma_0^{\frac{1}{2}} H \Gamma_0^{\frac{1}{2}} C_0 Z_1 - Z'(\tilde{C}_0 A_1 - \rho_{20} A_2) H(\tilde{C}_0 A_1 - \rho_{20} A_2) Z}{NT(\xi_{\epsilon 0}^2 + T\xi_{b0}^2)} \\
&\quad - \frac{Z'_1 C'_0 \Gamma_0^{\frac{1}{2}} (I - H) \Gamma_0^{\frac{1}{2}} C_0 Z_1 + Z'(\tilde{C}_0 A_1 - \rho_{20} A_2) (I - H) (\tilde{C}_0 A_1 - \rho_{20} A_2) Z}{NT\xi_{\epsilon 0}^2}, \\
\frac{1}{NT} \frac{\partial^2 \log L(\delta_0)}{\partial \xi_\epsilon^2 \partial \xi_\epsilon^2} &= \frac{1}{2\xi_{\epsilon 0}^4}, \quad \frac{1}{NT} \frac{\partial^2 \log L(\delta_0)}{\partial \xi_\epsilon^2 \partial \xi_b^2} = \frac{1}{2(\xi_{\epsilon 0}^2 + T\xi_{b0}^2)^2} + o(1), \\
\frac{1}{NT} \frac{\partial^2 \log L(\delta_0)}{\partial \xi_b^2 \partial \xi_b^2} &= -\frac{T^2}{(\xi_{\epsilon 0}^2 + T\xi_{b0}^2)^2} + o(1),
\end{aligned}$$

So, according to Assumption 2 (iv) and Fact 2 in Lee [38], it is important to know

$$\begin{aligned}
&\frac{1}{NT} \frac{\partial^2 \log L_1(\delta_0)}{\partial \rho_1^2} - \frac{1}{NT} \frac{\partial^2 \log L_1(\tilde{\delta})}{\partial \rho_1^2} \\
&= \frac{11}{NT} \{ \rho_{20}^2 \text{tr}[(C_0^{-1}W)^2] - \rho_{20}^2 \text{tr}[(C^{-1}W)^2] + \rho_{20}^2 \text{tr}[(C^{-1}W)^2] - \rho_2^2 \text{tr}[(C^{-1}W)^2] \} \\
&\quad + \frac{\{ \text{tr}[(C_0^{-1}W)^2] - \text{tr}[(C^{-1}(\tilde{\rho}_1)W)^2] \}}{NT} + \frac{\{ \rho_{20}^2 \xi_{b0}^2 \text{tr}[(C_0^{-1}W)^2] - \rho_2^2 \xi_{b0}^2 \text{tr}[(C_0^{-1}W)^2] \}}{N} \\
&\quad + \frac{\{ \rho_2^2 \xi_{b0}^2 \text{tr}[(C_0^{-1}W)^2] - \rho_2^2 \xi_b^2 \text{tr}[(C_0^{-1}W)^2] \}}{N} + \frac{\{ \rho_2^2 \xi_b^2 \text{tr}[(C_0^{-1}W)^2] - \rho_2^2 \xi_b^2 \text{tr}[(C^{-1}W)^2] \}}{N} \\
&= op(1).
\end{aligned}$$

Then, we have

$$\frac{1}{NT} \frac{\partial^2 \log L(\tilde{\delta})}{\partial \delta \partial \delta'} - \frac{1}{NT} \frac{\partial^2 \log L(\delta_0)}{\partial \delta \partial \delta'} = op(1).$$

Now, to prove (A6), from Assumption 1 and Theorem 1 in Kelejian and Prucha [25], the conclusion that $\frac{1}{\sqrt{NT}} \frac{\partial \log L(\delta_0)}{\partial \delta}$ is an asymptotically normal distribution with mean 0 is obtained. Therefore, the variance is calculated as follows:

$$\begin{aligned} \text{Var}\left(\frac{1}{\sqrt{NT}} \frac{\partial \log L(\delta_0)}{\partial \delta}\right) &= E\left(\frac{1}{\sqrt{NT}} \frac{\partial \log L(\delta_0)}{\partial \delta} \cdot \frac{1}{\sqrt{NT}} \frac{\partial \log L(\delta_0)}{\partial \delta'}\right) \\ &= -E\left(\frac{1}{NT} \frac{\partial^2 \log L(\delta_0)}{\partial \delta \partial \delta'}\right) + \Omega_{\delta_0}, \end{aligned} \quad (\text{A7})$$

where

$$\begin{aligned} -\frac{1}{NT} \frac{\partial^2 \log L(\delta_0)}{\partial \delta \partial \delta'} &= \begin{pmatrix} L_{11} & L_{12} & L_{13} & L_{14} & L_{15} \\ * & L_{22} & L_{23} & L_{24} & L_{25} \\ * & * & L_{33} & L_{34} & L_{35} \\ * & * & * & L_{44} & L_{45} \\ * & * & * & * & L_{55} \end{pmatrix} + op(1) \equiv \Sigma_{\delta_0} + op(1), \\ \Omega_{\delta_0} &= \begin{pmatrix} \Pi_{11} & \Pi_{12} & \Pi_{13} & \Pi_{14} & \Pi_{15} \\ * & \Pi_{22} & \Pi_{23} & \Pi_{24} & \Pi_{25} \\ * & * & \Pi_{33} & \Pi_{34} & \Pi_{35} \\ * & * & * & \Pi_{44} & \Pi_{45} \\ * & * & * & * & \Pi_{55} \end{pmatrix}, \\ L_{11} &= \left(\frac{1}{NT} + \frac{11\rho_{20}^2}{NT} + \frac{2\rho_{20}^2\bar{\zeta}_{b0}^2}{N}\right)tr[(C_0^{-1}W)^2], \quad L_{12} = -\frac{\rho_{20}^2 + 2\rho_{20}}{NT}tr(C_0^{-1}W), \quad L_{13} = 0, \\ L_{14} &= \frac{tr(\Gamma_0 C_0 W)}{NT(\bar{\zeta}_{\epsilon 0}^2 + T\bar{\zeta}_{b0}^2)} + \frac{\rho_{20}^2}{NT}tr(C_0^{-1}W), \quad L_{15} = \frac{tr(\Gamma_0 C_0 W)}{N(\bar{\zeta}_{\epsilon 0}^2 + T\bar{\zeta}_{b0}^2)} + \frac{\rho_{20}^2 tr(C_0^{-1}W)}{N(\bar{\zeta}_{\epsilon 0}^2 + T\bar{\zeta}_{b0}^2)} + \frac{tr(A_4 \bar{W})}{N(\bar{\zeta}_{\epsilon 0}^2 + T\bar{\zeta}_{b0}^2)}, \\ L_{22} &= \frac{(T-1)\bar{\zeta}_{b0}^2}{\bar{\zeta}_{\epsilon 0}^2}, \quad L_{23} = 0, \quad L_{24} = \frac{1}{T(\bar{\zeta}_{\epsilon 0}^2 + T\bar{\zeta}_{b0}^2)^2} + \frac{\rho_{20}(\bar{\zeta}_{\epsilon 0}^2 + T\bar{\zeta}_{b0}^2)}{\bar{\zeta}_{\epsilon}^4}, \quad L_{25} = \frac{1}{\bar{\zeta}_{\epsilon 0}^2 + T\bar{\zeta}_{b0}^2}, \\ L_{33} &= \frac{Z_1' C_0' \Gamma_0^{\frac{1}{2}} H \Gamma_0^{\frac{1}{2}} C_0 Z_1 + Z'(\tilde{C}_0 A_1 - \rho_{20} A_2) H(\tilde{C}_0 A_1 - \rho_{20} A_2) Z}{NT(\bar{\zeta}_{\epsilon 0}^2 + T\bar{\zeta}_{b0}^2)} \\ &\quad + \frac{Z_1' C_0' \Gamma_0^{\frac{1}{2}} (I - H) \Gamma_0^{\frac{1}{2}} C_0 Z_1 + Z'(\tilde{C}_0 A_1 - \rho_{20} A_2)(I - H)(\tilde{C}_0 A_1 - \rho_{20} A_2) Z}{NT\bar{\zeta}_{\epsilon 0}^2}, \\ L_{34} &= 0, \quad L_{35} = 0, \quad L_{44} = -\frac{1}{2\bar{\zeta}_{\epsilon 0}^4}, \quad L_{45} = -\frac{1}{2(\bar{\zeta}_{\epsilon 0}^2 + T\bar{\zeta}_{b0}^2)^2}, \quad L_{55} = \frac{T^2}{(\bar{\zeta}_{\epsilon 0}^2 + T\bar{\zeta}_{b0}^2)^2}, \\ \Pi_{11} &= \frac{(\mu_4 - 3\bar{\zeta}_{\epsilon 0}^4) \sum_{i=1}^{NT} (A_{60ii})^2}{NT(\bar{\zeta}_{\epsilon 0}^2 + T\bar{\zeta}_{b0}^2)^2} + \frac{(\mu_4 - 3\sigma_0^4) \sum_{i=1}^{NT} (A_{70ii})^2}{NT\bar{\zeta}_{\epsilon 0}^4}, \quad \Pi_{13} = \frac{\mu_3 Z' A_{120} \text{diag} A_{70}}{NT(\bar{\zeta}_{\epsilon 0}^2 + T\bar{\zeta}_{b0}^2)^2}, \\ \Pi_{12} &= \frac{(\mu_4 - 3\bar{\zeta}_{\epsilon 0}^4) \sum_{i=1}^{NT} (A_{60ii} A_{90ii})}{NT\bar{\zeta}_{\epsilon 0}^2 (\bar{\zeta}_{\epsilon 0}^2 + T\bar{\zeta}_{b0}^2)} + \frac{(\mu_4 - 3\bar{\zeta}_{\epsilon 0}^4) \sum_{i=1}^{NT} (A_{90ii}^2) + \mu_3 Z' A_{130} \text{diag} A_{70}}{NT\bar{\zeta}_{\epsilon 0}^4}, \\ &\quad + \frac{(\mu_4 - 3\bar{\zeta}_{\epsilon 0}^4) \sum_{i=1}^{NT} (A_{60ii} A_{80ii})}{NT(\bar{\zeta}_{\epsilon 0}^2 + T\bar{\zeta}_{b0}^2)^2}, \quad \Pi_{14} = \frac{\mu_4 \sum_{i=1}^{NT} (A_{60ii}^2) - 3\bar{\zeta}_{\epsilon 0}^4 \sum_{i=1}^{NT} (A_{60ii}^2)}{NT(\bar{\zeta}_{\epsilon 0}^2 + T\bar{\zeta}_{b0}^2)^3} \\ &\quad + \frac{(\mu_4 - 3\bar{\zeta}_{\epsilon 0}^4) \sum_{i=1}^{NT} (A_{70ii} A_{100ii})}{NT\bar{\zeta}_{\epsilon 0}^2 (\bar{\zeta}_{\epsilon 0}^2 + T\bar{\zeta}_{b0}^2)^2} + \frac{(\mu_4 - 3\bar{\zeta}_{\epsilon 0}^4) \sum_{i=1}^{NT} (A_{70ii} A_{110ii})}{NT\bar{\zeta}_{\epsilon 0}^8}, \\ \Pi_{15} &= \frac{(\mu_4 - 3\bar{\zeta}_{\epsilon 0}^4) \sum_{i=1}^{NT} (A_{60ii} A_{100ii})}{2N(\bar{\zeta}_{\epsilon 0}^2 + T\bar{\zeta}_{b0}^2)^4} + \frac{(\mu_4 - 3\bar{\zeta}_{\epsilon 0}^4) \sum_{i=1}^{NT} (A_{70ii} A_{100ii})}{2N\bar{\zeta}_{\epsilon 0}^2 (\bar{\zeta}_{\epsilon 0}^2 + T\bar{\zeta}_{b0}^2)}, \\ \Pi_{22} &= \frac{(\mu_4 - 3\bar{\zeta}_{\epsilon 0}^4) \sum_{i=1}^{NT} (A_{80ii}^2)}{NT(\bar{\zeta}_{\epsilon 0}^2 + T\bar{\zeta}_{b0}^2)^2} + \frac{(\mu_4 - 3\bar{\zeta}_{\epsilon 0}^4) \sum_{i=1}^{NT} (A_{80ii} A_{90ii})}{NT\bar{\zeta}_{\epsilon 0}^2 (\bar{\zeta}_{\epsilon 0}^2 + T\bar{\zeta}_{b0}^2)} + \frac{(\mu_4 - 3\bar{\zeta}_{\epsilon 0}^4) \sum_{i=1}^{NT} (A_{90ii}^2)}{NT\bar{\zeta}_{\epsilon 0}^4}, \\ \Pi_{33} &= 0, \quad \Pi_{25} = \frac{(\mu_4 - 3\bar{\zeta}_{\epsilon 0}^4) \sum_{i=1}^{NT} (A_{80ii} A_{110ii})}{2N(\bar{\zeta}_{\epsilon 0}^2 + T\bar{\zeta}_{b0}^2)^3} + \frac{(\mu_4 - 3\bar{\zeta}_{\epsilon 0}^4) \sum_{i=1}^{NT} (A_{70ii} A_{110ii})}{2\bar{\zeta}_{\epsilon 0}^2 (\bar{\zeta}_{\epsilon 0}^2 + T\bar{\zeta}_{b0}^2)^2}, \end{aligned}$$

$$\begin{aligned}\Pi_{24} &= \frac{(\mu_4 - 3\zeta_{\epsilon 0}^4) \sum_{i=1}^{NT} (A_{80ii} A_{100ii})}{2NT(\zeta_{\epsilon 0}^2 + T\zeta_{b0}^2)^3} + \frac{(\mu_4 - 3\zeta_{\epsilon 0}^4) \sum_{i=1}^{NT} (A_{80ii} A_{110ii})}{NT\zeta_{\epsilon 0}^4(\zeta_{\epsilon 0}^2 + T\zeta_{b0}^2)} + \frac{(\mu_4 - 3\zeta_{\epsilon 0}^4) \sum_{i=1}^{NT} (A_{90ii} A_{110ii})}{NT\zeta_{\epsilon 0}^6}, \\ \Pi_{34} &= \frac{\mu_3 Z' A_{120} \text{diag} A_{100}}{2NT(\zeta_{\epsilon 0}^2 + T\zeta_{b0}^2)^3} + \frac{\mu_3 Z' A_{120} \text{diag} A_{110}}{NT\zeta_{\epsilon 0}^4(\zeta_{\epsilon 0}^2 + T\zeta_{b0}^2)}, \quad \Pi_{35} = \frac{\mu_3 Z' A_{120} \text{diag} A_{100}}{2N(\zeta_{\epsilon 0}^2 + T\zeta_{b0}^2)^3} + \frac{\mu_3 Z' A_{130} \text{diag} A_{100}}{2N\zeta_{\epsilon 0}^2(\zeta_{\epsilon 0}^2 + T\zeta_{b0}^2)^2}, \\ \Pi_{45} &= \frac{(\mu_4 - 3\zeta_{\epsilon 0}^4) \sum_{i=1}^{NT} (A_{100ii}^2)}{2N(\zeta_{\epsilon 0}^2 + T\zeta_{b0}^2)^4} + \frac{(\mu_4 - 3\zeta_{\epsilon 0}^4) \sum_{i=1}^{NT} (A_{100ii} A_{110ii})}{2N\zeta_{\epsilon 0}^4(\zeta_{\epsilon 0}^2 + T\zeta_{b0}^2)^2}, \quad \Pi_{55} = \frac{T(\mu_4 - 3\zeta_{\epsilon 0}^4) \sum_{i=1}^{NT} (A_{100ii}^2)}{2N(\zeta_{\epsilon 0}^2 + T\zeta_{b0}^2)^4}, \\ \Pi_{44} &= \frac{(\mu_4 - 3\zeta_{\epsilon 0}^4) \sum_{i=1}^{NT} (A_{100ii}^2)}{2NT(\zeta_{\epsilon 0}^2 + T\zeta_{b0}^2)^4} + \frac{(\mu_4 - 3\zeta_{\epsilon 0}^4) \sum_{i=1}^{NT} (A_{100ii} A_{110ii})}{2NT\zeta_{\epsilon 0}^4(\zeta_{\epsilon 0}^2 + T\zeta_{b0}^2)^2} + \frac{(\mu_4 - 3\zeta_{\epsilon 0}^4) \sum_{i=1}^{NT} (A_{110ii}^2)}{4NT\zeta_{\epsilon 0}^8}, \\ \Pi_{23} &= \frac{\mu_3 Z' A_{120} \text{diag} A_{80}}{NT(\zeta_{\epsilon 0}^2 + T\zeta_{b0}^2)^2} + \frac{\mu_3 Z' A_{130} \text{diag} A_{80}}{NT\zeta_{\epsilon 0}^2(\zeta_{\epsilon 0}^2 + T\zeta_{b0}^2)}, \quad A_1 = (\mathbf{0}_N, I_{N(T-1) \times N(T-1)}), \\ A_2 &= (I_{N(T-1) \times N(T-1)}, \mathbf{0}_N)', \quad A_5 = (I_N, \mathbf{0}_{N \times N(T-1)}), \quad A_6 = A_4' \tilde{W} H A_1, \quad A_7 = A_4' \tilde{W} (I - H) A_1, \\ A_8 &= A_2' A_3' H A_1, \quad A_9 = A_2' A_3' (I - H) A_1, \quad A_{10} = A_1' H A_1, \quad A_{11} = A_1' (I - H) A_1, \quad A_{12} = (\tilde{C} A_1 - \rho_2 A_2) H A_1,\end{aligned}$$

$$A_3 = \begin{pmatrix} C^{-1} \Gamma^{-\frac{1}{2}} & 0 & 0 & \cdots & 0 \\ \rho_2 C^{-2} \Gamma^{-\frac{1}{2}} & C^{-1} & 0 & \cdots & 0 \\ \rho_2^2 C^{-3} \Gamma^{-\frac{1}{2}} & \rho_2 C^{-2} \Gamma^{-\frac{1}{2}} & C^{-1} & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \rho_2^{T-2} C^{-T+1} \Gamma^{-\frac{1}{2}} & \rho_2^{T-3} C^{-T+2} \Gamma^{-\frac{1}{2}} & \rho_2^{T-4} C^{-T+3} \Gamma^{-\frac{1}{2}} & \cdots & C^{-1} \end{pmatrix}_{N(T-1) \times N(T-1)},$$

$$A_4 = \begin{pmatrix} \rho_2 C^{-2} \Gamma^{-\frac{1}{2}} & C^{-1} & 0 & \cdots & 0 \\ \rho_2^2 C^{-3} \Gamma^{-\frac{1}{2}} & \rho_2 C^{-2} \Gamma^{-\frac{1}{2}} & C^{-1} & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \rho_2^{T-1} C^{-T} \Gamma^{-\frac{1}{2}} & \rho_2^{T-2} C^{-T+1} \Gamma^{-\frac{1}{2}} & \rho_2^{T-3} C^{-T+2} \Gamma^{-\frac{1}{2}} & \cdots & C^{-1} \end{pmatrix}_{N(T-1) \times NT},$$

μ_3 and μ_4 are the third and fourth moments, and $A_{i0} (i = 1, \dots, 13)$ are the true value of $A_i (i = 1, \dots, 13)$, $\tilde{C}_0$ is the true value of $\tilde{C}$, respectively.

Therefore, according to Assumption 6 and the forms of Σ_{δ_0} and Ω_{δ_0} , we obtain

$$\sqrt{NT}(\hat{\delta} - \delta_0) \xrightarrow{D} \mathcal{N}(0, \lim_{N \rightarrow \infty} (\Sigma_{\delta_0}^{-1} + \Sigma_{\delta_0}^{-1} \Omega_{\delta_0} \Sigma_{\delta_0}^{-1})).$$

□

Proof of Theorem 3. Recall $\hat{\phi} = \iota(\iota)(Y - Z\hat{\gamma})$ and

$$Y = \begin{pmatrix} F_{10} + Z_1 \gamma_0 + C_0^{-1} \Gamma_0^{-\frac{1}{2}} \eta_{10} \\ F_{20} + Z_2 \gamma_0 + \rho_{20} C_0^{-2} \Gamma_0^{-\frac{1}{2}} \eta_{10} + C_0^{-1} \eta_{20} \\ F_{30} + Z_3 \gamma_0 + \rho_{20}^2 C_0^{-3} \Gamma_0^{-\frac{1}{2}} \eta_{10} + \rho_{20} C_0^{-2} \eta_{20} + C_0^{-1} \eta_{30} \\ \vdots \\ F_{T0} + Z_T \gamma_0 + \rho_{20}^{T-1} C_0^{-T} \Gamma_0^{-\frac{1}{2}} \eta_{10} + \rho_{20}^{T-2} C_0^{-T+1} \eta_{20} + \dots + C_0^{-1} \eta_{T0} \end{pmatrix},$$

then

$$\begin{aligned}\hat{\phi} &= \iota(\iota)(Y - Z\hat{\gamma}) \\ &= \iota(\iota)[f_0(\iota) + \tilde{\eta} - Z(\hat{\gamma} - \gamma_0)],\end{aligned}$$

where $\tilde{\eta} = \hat{G}(\eta_{10}, \eta_{20}, \dots, \eta_{T0})'$ and

$$\hat{G} = \begin{pmatrix} C_0^{-1}\Gamma_0^{-\frac{1}{2}} & 0 & \dots & 0 \\ \rho_{20}C_0^{-2}\Gamma_0^{-\frac{1}{2}} & C_0^{-1} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ \rho_{20}^{T-1}C_0^{-T}\Gamma_0^{-\frac{1}{2}} & \rho_{20}^{T-2}C_0^{-T+1} & \dots & C_0^{-1} \end{pmatrix}.$$

Let $\phi = \begin{pmatrix} f(\iota) \\ S\dot{f}(\iota) \end{pmatrix}$, by the second order Taylor expression, we know that $f(l_{it}) = \tilde{l}'_{it}(\iota)\phi + \frac{1}{2}(l_{it} - \iota)' \ddot{f}(\iota)(l_{it} - \iota) + op(\|S\|^2)$. It is easy to obtain

$$\begin{aligned} \iota(\iota)f(l_{it}) &= [\tilde{l}'(\iota)K_S(\iota)\tilde{l}(\iota)]^{-1}\tilde{l}'(\iota)K_S(\iota)\tilde{l}(\iota)\phi + \frac{1}{2}\iota(l_{it})(l_{it} - \iota)' \ddot{f}(\iota)(l_{it} - \iota) \\ &= \phi + \frac{1}{2}\iota(l_{it})(l_{it} - \iota)' \ddot{f}(\iota)(l_{it} - \iota). \end{aligned} \quad (A8)$$

Therefore,

$$\hat{\phi} - \phi = \frac{1}{2} \sum_{i=1}^N \sum_{t=1}^T \iota(l_{it})(l_{it} - \iota)' \ddot{f}(\iota)(l_{it} - \iota) + \iota(\iota)\tilde{\eta} - \iota(\iota)Z(\hat{\gamma} - \gamma)$$

and

$$\sqrt{N|S|}(\hat{\phi} - \phi) = B_{11} + B_{12} - B_{13},$$

where $B_{11} = \frac{\sqrt{N|S|}}{2} \sum_{i=1}^N \sum_{t=1}^T \iota(l_{it})(l_{it} - \iota)' \ddot{f}(\iota)(l_{it} - \iota)$, $B_{12} = \sqrt{N|S|}\iota(\iota)\tilde{\eta}$ and $B_{13} = \sqrt{N|S|}\iota(\iota)Z(\hat{\gamma} - \gamma)$.

For B_{11} , write

$$\begin{aligned} & \frac{\sqrt{N|S|}}{2} \sum_{i=1}^N \sum_{t=1}^T \iota(l_{it})(l_{it} - \iota)' \ddot{f}(\iota)(l_{it} - \iota) \\ &= \frac{\sqrt{N|S|}}{2} Q^{-1} \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \tilde{l}'_{it}(\iota)K_S(l_{it} - \iota)(l_{it} - \iota)' \ddot{f}(\iota)(l_{it} - \iota) + op(1) \\ &= Q^{-1} \frac{\sqrt{N|S|}}{2} \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T [1(l_{it} - \iota)'S^{-1}]' K_S(l_{it} - \iota)(l_{it} - \iota)' \ddot{f}(\iota)(l_{it} - \iota) + op(1) \\ &= Q^{-1} \frac{\sqrt{N|S|}}{2} \frac{1}{NT} \left(\frac{\sum_{i=1}^N \sum_{t=1}^T K_S(l_{it} - \iota)(l_{it} - \iota)' \ddot{f}(\iota)(l_{it} - \iota)}{\sum_{i=1}^N \sum_{t=1}^T (l_{it} - \iota)' S^{-1} K_S(l_{it} - \iota)(l_{it} - \iota)' \ddot{f}(\iota)(l_{it} - \iota)} \right) + op(1) \\ &= Q^{-1} \frac{\sqrt{N|S|}}{2} O(|S|^{-3}s^{-3}) + op(1) \\ &= o(1). \end{aligned}$$

For B_{12} , it is easy to know that $E[B_{12}] = 0$ and $Var(B_{12}) = E[B_{12}^2] = O(1)$.

For B_{13} , by Lemma 2 and Facts 1, 2 (Lee [38]) and Theorem 1, we can show that $\|\iota(\iota)\| = Op(N^{1/2}|S|^{-1})$ and $\hat{\gamma} - \gamma_0 = Op(\frac{1}{NT})$, then $B_{13} = op(1)$ can be obtained. Following Slutsky's theorem and central limit theorem, we obtain

$$\sqrt{N|S|}(\hat{\phi} - \phi) \xrightarrow{D} \mathcal{N}(0, \sigma^2(\phi)),$$

where $\sigma^2(\phi) = \hat{G}'\iota'(\iota)\sigma_{\eta}^2\iota(\iota)\hat{G}$. $\square$

Proof of Theorem 4. It can be seen that

$$\begin{aligned}
 RSS(H_1) &= \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \hat{f}(\iota_{it}) - z'_{it}\hat{\gamma}]^2 \\
 &= \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \hat{f}(\iota_{it}) - z'_{it}\gamma]^2 \\
 &\quad + \left\{ \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \hat{f}(\iota_{it}) - z'_{it}\hat{\gamma}]^2 - \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \hat{f}(\iota_{it}) - z'_{it}\gamma]^2 \right\} \\
 &= RSS^*(H_1) + T_1,
 \end{aligned}$$

where $RSS^*(H_1) = \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \hat{f}(\iota_{it}) - z'_{it}\gamma]^2$ and $T_1 = \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \hat{f}(\iota_{it}) - z'_{it}\hat{\gamma}]^2 - \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \hat{f}(\iota_{it}) - z'_{it}\gamma]^2$. For T_1 , it can be obtained by Theorems 1–3 and the calculation that

$$\begin{aligned}
 T_1 &= \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \hat{f}(\iota_{it}) - z'_{it}\hat{\gamma}]^2 - \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \hat{f}(\iota_{it}) - z'_{it}\gamma]^2 \\
 &= \sum_{i=1}^N \sum_{t=1}^T [z'_{it}\gamma + f(\iota_{it}) + v_{it} - \hat{f}(\iota_{it}) - z'_{it}\hat{\gamma}]^2 \\
 &\quad - \sum_{i=1}^N \sum_{t=1}^T [z'_{it}\gamma + f(\iota_{it}) + v_{it} - \hat{f}(\iota_{it}) - z'_{it}\gamma]^2 \\
 &= \sum_{i=1}^N \sum_{t=1}^T \left[Op\left(\frac{1}{\sqrt{NT}}\right) + O(|S|^2) + |S|^2 Op\left(\frac{1}{\sqrt{N}|S|}\right) + v \right]^2 \\
 &\quad - \sum_{i=1}^N \sum_{t=1}^T \left[O(|S|^2) + |S|^2 Op\left(\frac{1}{\sqrt{N}|S|}\right) + v \right]^2 \\
 &= O(1).
 \end{aligned}$$

Similarly, we have that

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

where $RSS^*(H_0) = \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \tilde{f}(\iota_{it}) - z'_{it}\gamma]^2$, and $T_0 = \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \tilde{f}(\iota_{it}) - z'_{it}\tilde{\gamma}]^2 - \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \tilde{f}(\iota_{it}) - z'_{it}\gamma]^2$. For T_0 , it can be obtained by Theorems 1–3 and the calculation that

$$\begin{aligned} T_0 &= \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \tilde{f}(\iota_{it}) - z'_{it}\tilde{\gamma}]^2 - \sum_{i=1}^N \sum_{t=1}^T [y_{it} - \tilde{f}(\iota_{it}) - z'_{it}\gamma]^2 \\ &= \sum_{i=1}^N \sum_{t=1}^T [z'_{it}\gamma + f(\iota_{it}) + v_{it} - \tilde{f}(\iota_{it}) - z'_{it}\tilde{\gamma}]^2 \\ &\quad - \sum_{i=1}^N \sum_{t=1}^T [z'_{it}\gamma + f(\iota_{it}) + v_{it} - \tilde{f}(\iota_{it}) - z'_{it}\gamma]^2 \\ &= \sum_{i=1}^N \sum_{t=1}^T [Op(\frac{1}{\sqrt{NT}}) + \tilde{f}(\iota_{it}) - z'_{it}\tilde{\gamma}]^2 \\ &\quad - \sum_{i=1}^N \sum_{t=1}^T [\tilde{f}(\iota_{it}) - z'_{it}\tilde{\gamma}]^2 \\ &= Op(1). \end{aligned}$$

Furthermore, we obtain that

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

According to the Remark 3.4 in Fan et al. [36], under H_0 , we have

$$r_k F_{NT} \xrightarrow{D} \chi^2_{r_k c_k |\Delta|/|S|},$$

where $r_k = \frac{K(0) - \frac{1}{2} \int K^2(\iota) d\iota}{\int (K(\iota) - \frac{1}{2} K * K(\iota))^2 d\iota}$, $c_k = K(0) - \frac{1}{2} \int K^2(\iota) d\iota$, Δ is the support set of ι , $K(\iota) = \text{diag}(S^{-1}k(S^{-1}(\iota_{11} - \iota)), \dots, S^{-1}k(S^{-1}(\iota_{NT} - \iota)))$, and $K * K$ is the convolution of K . $\square$

References

- Cheng, S.; Chen, J. GMM estimation of partially linear additive spatial autoregressive model. *Comput. Stat. Data Anal.* **2023**, *182*, 107712. [\[CrossRef\]](#)
- Chamberlain, G. Multivariate regression models for panel data. *J. Econ.* **1982**, *18*, 5–46. [\[CrossRef\]](#)
- Chamberlain, G. Panel data. *Handb. Econom.* **1984**, *2*, 1247–1318.
- Hsiao, C. *Analysis of Panel Data*; Cambridge University Press: Cambridge, UK, 2022; Volume 64.
- Baltagi, B.H. *Econometric Analysis of Panel Data*; Springer: Berlin/Heidelberg, Germany, 2008.
- Arellano, M. *Panel Data Econometrics*; Oxford University Press: Oxford, UK, 2003.
- Wooldridge, J.M. *Econometric Analysis of Cross Section and Panel Data*; MIT Press: Cambridge, MA, USA, 2010.
- Thapa, S.; Lomholt, M.A.; Krog, J.; Cherstvy, A.G.; Metzler, R. Bayesian analysis of single-particle tracking data using the nested-sampling algorithm: Maximum-likelihood model selection applied to stochastic-diffusivity data. *Phys. Chem. Chem. Phys.* **2018**, *20*, 29018–29037. [\[CrossRef\]](#) [\[PubMed\]](#)
- Rehfeldt, F.; Weiss, M. The random walker's toolbox for analyzing single-particle tracking data. *Soft Matter* **2023**, *19*, 5206–5222. [\[CrossRef\]](#) [\[PubMed\]](#)
- Elhorst, J.P. *Spatial Econometrics: From Cross-Sectional Data to Spatial Panels*; Springer: Berlin/Heidelberg, Germany, 2014.
- Brueckner, J.K. Strategic interaction among governments: An overview of empirical studies. *Int. Reg. Sci. Rev.* **2003**, *26*, 175–188. [\[CrossRef\]](#)
- LeSage, J.P. An introduction to spatial econometrics. *Rev. Econ. Ind.* **2008**, *123*, 19–44.
- Elhorst, J.P. Applied spatial econometrics: Raising the bar. *Spat. Econ. Anal.* **2010**, *5*, 9–28. [\[CrossRef\]](#)
- Allers, M.; Elhorst, J. Tax mimicking and yardstick competition among local governments in The Netherlands. *Int. Tax Public Financ.* **2005**, *12*, 493–513. [\[CrossRef\]](#)
- Elhorst, J.P. Specification and estimation of spatial panel data models. *Int. Reg. Sci. Rev.* **2003**, *26*, 244–268. [\[CrossRef\]](#)
- Baltagi, B.H.; Li, D. Prediction in the panel data model with spatial correlation. In *Advances in Spatial Econometrics: Methodology, Tools and Applications*; Springer: Berlin/Heidelberg, Germany, 2004; pp. 283–295.
- Pesaran, M.H. General diagnostic tests for cross-sectional dependence in panels. *Empir. Econ.* **2021**, *60*, 13–50. [\[CrossRef\]](#)
- Kapoor, M.; Kelejian, H.H.; Prucha, I.R. Panel data models with spatially correlated error components. *J. Econ.* **2007**, *140*, 97–130. [\[CrossRef\]](#)

19. Lee, L.; Yu, J. Estimation of spatial autoregressive panel data models with fixed effects. *J. Econ.* **2010**, *154*, 165–185. [\[CrossRef\]](#)
20. Mutl, J.; Pfaffermayr, M. The Hausman test in a Cliff and Ord panel model. *Econ. J.* **2011**, *14*, 48–76. [\[CrossRef\]](#)
21. Baltagi, B.H.; Egger, P.; Pfaffermayr, M. A generalized spatial panel data model with random effects. *Econ. Rev.* **2013**, *32*, 650–685. [\[CrossRef\]](#)
22. Anselin, L. The Scope of Spatial Econometrics. In *Spatial Econometrics: Methods and Models*; Springer: Dordrecht, The Netherlands, 1988; pp. 7–15.
23. Baltagi, B.H.; Song, S.H.; Koh, W. Testing panel data regression models with spatial error correlation. *J. Econ.* **2003**, *117*, 123–150. [\[CrossRef\]](#)
24. Bordignon, M.; Cerniglia, F.; Revelli, F. In search of yardstick competition: A spatial analysis of Italian municipality property tax setting. *J. Urban Econ.* **2003**, *54*, 199–217. [\[CrossRef\]](#)
25. Kelejian, H.H.; Prucha, I.R. On the asymptotic distribution of the Moran I test statistic with applications. *J. Econ.* **2001**, *104*, 219–257. [\[CrossRef\]](#)
26. Cohen, J.P.; Paul, C.J. Public infrastructure investment, interstate spatial spillovers, and manufacturing costs. *Rev. Econ. Stat.* **2004**, *86*, 551–560. [\[CrossRef\]](#)
27. Lee, L.; Yu, J. Spatial panels: random components versus fixed effects. *Int. Econ. Rev.* **2012**, *53*, 1369–1412. [\[CrossRef\]](#)
28. Parent, O.; LeSage, J.P. A space–time filter for panel data models containing random effects. *Comput. Stat. Data Anal.* **2011**, *55*, 475–490. [\[CrossRef\]](#)
29. Elhorst, J.P. Serial and spatial error correlation. *Econ. Lett.* **2008**, *100*, 422–424. [\[CrossRef\]](#)
30. Lee, L.; Yu, J. Estimation of fixed effects panel regression models with separable and nonseparable space–time filters. *J. Econ.* **2015**, *184*, 174–192. [\[CrossRef\]](#)
31. 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 Theory Appl.* **2020**, *9*, 2150005. [\[CrossRef\]](#)
32. 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. [\[CrossRef\]](#)
33. 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. [\[CrossRef\]](#)
34. Bai, Y.; Hu, J.; You, J. Panel data partially linear varying-coefficient model with errors correlated in space and time. *Stat. Sin.* **2015**, *35*, 275–294.
35. Cai, Z. Trending time-varying coefficient time series models with serially correlated errors. *J. Econ.* **2007**, *136*, 163–188. [\[CrossRef\]](#)
36. Fan, J.; Zhang, C.; Zhang, J. Generalized likelihood ratio statistics and Wilks phenomenon. *Ann. Stat.* **2001**, *29*, 153–193. [\[CrossRef\]](#)
37. Su, L.; Jin, S. Profile quasi-maximum likelihood estimation of partially linear spatial autoregressive models. *J. Econ.* **2010**, *157*, 18–33. [\[CrossRef\]](#)
38. Lee, L. Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica* **2004**, *72*, 1899–1925. [\[CrossRef\]](#)
39. Hamilton, S.A.; Truong, Y.K. Local linear estimation in partly linear models. *J. Multivar. Anal.* **1997**, *60*, 1–19. [\[CrossRef\]](#)
40. Su, L.; Ullah, A. Nonparametric and semiparametric panel econometric models: Estimation and testing. In *Handbook of Empirical Economics and Finance*; CRC: Boca Raton, FL, USA, 2011; pp. 455–497.
41. Su, L.; Ullah, A. Profile likelihood estimation of partially linear panel data models with fixed effects. *Econ. Lett.* **2006**, *92*, 75–81. [\[CrossRef\]](#)
42. Su, L. Semiparametric GMM estimation of spatial autoregressive models. *J. Econ.* **2012**, *167*, 543–560. [\[CrossRef\]](#)
43. Mack, Y.; Silverman, B.W. Weak and strong uniform consistency of kernel regression estimates. *Z. Wahrscheinlichkeitstheorie Verwandte Geb.* **1982**, *61*, 405–415. [\[CrossRef\]](#)
44. Feng, Q.; Horrace, W.C. Alternative technical efficiency measures: Skew, bias, and scale. *J. Appl. Econ.* **2010**, *27*, 253–268. [\[CrossRef\]](#)
45. Druska, V.; Horrace, W.C. Generalized moments estimation for spatial panel data: Indonesian rice farming. *Am. J. Agric. Econ.* **2004**, *86*, 185–198. [\[CrossRef\]](#)
46. Cheng, S.; Chen, J. Estimation of partially linear single-index spatial autoregressive model. *Stat. Pap.* **2021**, *62*, 495–531. [\[CrossRef\]](#)
47. Qu, A.; Lindsay, B.G.; Li, B. Improving generalized estimating equations using quadratic inference functions. *Biometrika* **2000**, *4*, 823–836. [\[CrossRef\]](#)
48. White, H. *Estimation, Inference and Specification Analysis*; Cambridge University Press: Cambridge, UK, 1996.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.