Published in final edited form as:

Reg Sci Urban Econ. 2015 March ; 51: 37–46. doi:10.1016/j.regsciurbeco.2015.01.003.

Estimation of Partially Specified Dynamic Spatial Panel Data Models with Fixed-Effects

Yuanqing Zhang^{1,2,3}

¹School of Finance and Business, Shanghai Normal University, No.100 Guilin Rd. Shanghai, 200234, P. R. China

²Key Laboratory of Mathematical Economics (SUFE), Ministry of Education, Shanghai 200433, P. R. China

³School of Statistics and Management, Shanghai University of Finance and Economics, Shanghai, 200433, P. R. China

Abstract

This paper studies estimation of a partially specified spatial panel data linear regression with fixed-effects. Under the assumption of strictly exogenous regressors and strictly exogenous spatial weighting matrix, the unknown parameter is estimated by applying the instrumental variable estimation. Under some sufficient conditions, the proposed estimator for the finite dimensional parameter is shown to be root-N consistent and asymptotically normally distributed; The proposed estimator for the unknown function is shown to be consistent and asymptotically distributed as well, though at a rate slower than root-N. Consistent estimators for the asymptotic variance-covariance matrices of both estimators are provided. The results can be generalized to several spatial weighting matrix and spatial matrix which vary with time. The simulation results suggest that the proposed approach has some practical value.

Keywords

Spatial; Panel Data; Partially Linear; Dynamic

1 Introduction

Dynamic panel data can not only capture dynamics of economic activities but also enable researchers to control unobservable heterogeneity across units. So, the spatial econometrics literature has exhibited a growing interest in the specification and estimation of dynamic regression equations based on spatial panels in the last decade. Both Kukenova and Monteiro (2009) consider a dynamic spatial panel data model and find that the system GMM estimator substantially reduces the bias in the parameter estimate of the WY_t variable, and that the

Correspondence to: Yuanqing Zhang.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

system GMM estimator outperforms the Arrelano and Bond difference GMM estimator. Yu and Lee (2008) study the quasi maximum likelihood (QML) estimation for spatial dynamic panel data models, Korniotis (2010) constructs a bias-corrected LSDV estimator for a dynamic panel data model in the fixed effects setting. Lee and Yu (2010c) propose an optimal GMM estimator based on linear moment conditions, which are standard, and quadratic moment conditions and prove that this GMM estimator is consistent, also when T is small relative to N . These spatial panel data models have many applications, for instance, transportation research (Parent O, LeSage JP (2011)), growth convergence of countries and regions (Baltagi et al., 2007; Ertur and Koch, 2007), regional markets (Keller and Shiue, 2007), labor economics (Foote, 2007), economic growth (Elhorst et al. (2010)), to name a few.

All of these studies specify a parametric regression function. Parametric regression models provide a parsimonious description of the relationship between the explained variables and the explanatory variables. However, a badly misspecified one may have large model approximation error leading to large modeling bias in practice. So an intermediate strategy is to apply a semiparametric form, among which partially linear models are widely used. Example studies include Cuzick (1992), Severini and Wong (1992), Liang, et al (1999), among others. More methods and references can be found in the monograph by Härdle, Liang and Gao (2000). On the application side, researchers have started addressing the importance of semiparametric modeling in spatial econometrics recently. for example, Basile and Gress(2005) propose a semiparametric spatial auto-covariance specification of the growth model for the European economy and find that nonlinearities are important in regional growth in Europe even when the spatial dependence is controlled. As a result, assuming a common linear relationship between economic growth and inputs is misleading.

To follow this line of literature, we wish to specify the following partially specified dynamic spatial panel regression (hereafter PSDSPR):

$$y_{it} = \gamma_0 y_{i,t-1} + \lambda_0 \sum_{j=1}^N w_{ij} y_{jt} + x'_{it} \beta_0 + g_0(z_{it}) + c_i + \varepsilon_{it}, i=1, \dots, N, t=2, \dots, T, \quad (1)$$

where y_{it} denotes the dependent variable of individual i in period t , $x_{it} = (x_{it1}, x_{it2}, \dots, x_{itq})'$ and z_{it} denote explanatory variables, c_i denotes the unobserved and time invariant individual effect, w_{ij} denotes the spatial weight between individual i and j , ε_{it} denotes random noise, and $\delta_0 = (\gamma_0, \lambda_0, \beta_0)'$ denotes the unknown true parameter value. where $g_0(\cdot)$ is an unknown function.

The model (1) is a generalization of many usual parametric, semi-parametric, panel data, spatial models. When $\gamma_0 \equiv 0$ and $T \equiv 1$, the model is the partially linear spatial autoregressive model study by Su and Jin (2010), Su (2012). When $\gamma_0 \equiv 0$ and $\lambda_0 = 0$, the model is the partially linear panel data model study by Li, Q. and T. Stengos, (1996), the book edited by Ai, C. and Li, Qi. (2008). When $\gamma_0 \equiv 0$, $g_0 \equiv 0$ and $T = 1$, the model is the spatial autoregressive model study by Kelejian and Prucha (1998, 1999). When $\gamma_0 \equiv 0$ and

$g_0 \equiv 0$, the model is the spatial panel data model study which that Various versions of theirs have been studied extensively in a series papers, see the reviews in Anselin (2001), the works of Elhorst (2001, 2003), as well as the recent papers by Baltagi et al (2003, 2006). , Kapoor et al., and Lee et al (2010)., among others. When $\lambda_0 \equiv 0$ and $\gamma_0 \equiv 0$, the model is the dynamic panel data model, see the Ahn, SC., P. Schmidt.(1995), for details. When $\gamma_0 \equiv 0$, the model is the partially linear spatial panel data model study by Ai and Zhang (2014) . When $g_0 \equiv 0$, the model is the dynamic spatial panel data model study by Kukenova and Monteiro (2009), Korniotis (2010) , Lee and Yu (2010c) , J. Paul Elhorst (2012). When $\lambda_0 \equiv 0$, the model is the partially linear dynamic panel data model study by Qi Li, Thomas J. Kniesner (2002), Byeong U. Park, Robin C. Sickles, Léopold Simar(2007) . Our estimation method in the paper could be applied successfully to the above special cases.

To the best of our knowledge, this is the first work in which the estimating problem of modeling partially specified dynamic spatial panel data is investigated. The main objective of this paper is to propose an approach to estimate δ_0 and $g_0(\cdot)$ consistently under the fixed-effects assumption which has the advantage of robustness in that the individual effects are allowed to correlate with included regressors in the model , to establish the asymptotic properties for the proposed estimators and to report on a simulation study.

The paper is organized as follows: Section 2 proposes an estimation of (1); Section 3 establishes the asymptotic properties of the proposed estimators; Section 4 presents some consistent covariance matrices; Section 5 reports some Monte Carlo simulation results. Concluding remarks are in Section 6, and all technical proofs are relegated to the appendices.

2 Estimation

Throughout the paper, we shall consider the model (1) only for the case of small and fixed T but large N . The case of large T and large N will be dealt with in a separate study. There are three issues with estimation of the model (1). First is the potential correlation between the unobserved individual effects c_i and the other explanatory variables $\{(x_{it}, z_{it}), t = 2, \dots, T\}$

(i.e., the fixed-effects). Second is the correlation between the spatial term $\sum_{j=1}^N w_{ij} y_{jt}$ and the error term ε_{it} (i.e., endogeneity). Third is the infinite dimensionality of the unknown parameter $g_0(z_{it})$. We tackle the infinite dimensionality problem by applying the sieve method (see Ai and Chen (2003)). Specifically, let $p^K(z_{it}) = (p_1(z_{it}), \dots, p_K(z_{it}))'$ denote a sequence of known basis functions that can approximate any measurable function arbitrarily well in the functional space endowed by a norm $\|\cdot\|_s$. Examples of sieves include polynomials, Fourier series, splines, wavelets etc. For each $K = K(N)$, there exists some constant $\pi_0 = \pi_{0,K} = \pi_{0,K(N)}$, such that $g_0(z) = p^{K(N)}(z) \pi_{0,K(N)} + v_0(z)$ with $v_0(z)$ as the approximation error satisfying $\|v_0(z)\|_s \rightarrow 0$ as $K \rightarrow \infty$. For simplicity, we denote $g_0(z) = p^K(z)' \pi_0 + v_0(z)$. Denote $v_{0it} = v_0(z_{it})$. We rewrite the model as

$$y_{it} = \gamma_0 y_{i,t-1} + \lambda_0 \sum_{j=1}^N w_{ij} y_{jt} + x'_{it} \beta_0 + p^K(z_{it})' \pi_0 + c_i + \varepsilon_{it} + v_{0it}, i=1, \dots, N, t=2, 3, \dots, T.$$

Let Δ denote the first difference operator and denote $p_{it} = p^K(z_{it}) - p^K(z_{it-1})$. We tackle the fixed-effects issue by applying the first difference

$$\Delta y_{it} = \gamma_0 \Delta y_{i,t-1} + \lambda_0 \sum_{j=1}^N w_{ij} \Delta y_{jt} + \Delta x'_{it} \beta_0 + \Delta p'_{it} \pi_0 + \Delta \varepsilon_{it} + \Delta v_{0it}, i=1, \dots, N, t=3, 4, \dots, T.$$

To simplify notation and to aid exposition, we stack the data in two ways and use whichever is convenient. One way is to stack data first by individual then by time. Specifically, define:

$$\Delta X_t = (\Delta x_{1t}, \Delta x_{2t}, \dots, \Delta x_{Nt})' \text{ and } \Delta X = (\Delta X'_3, \Delta X'_4, \dots, \Delta X'_T)'$$

Define $Y_t, Y, P_t, P, \varepsilon_t, \varepsilon, V_{0t}, V_0$ analogously using $y_{it}, p_{it}, \varepsilon_{it}, v_{0it}$ respectively. Let $W = (w_{ij})$ denote the weighting matrix and let $D_w = \text{diag}\{W, \dots, W\}$ denote the block diagonal matrix. With $\delta_0 = (\gamma_0, \lambda_0, \beta'_0)'$, we write

$$\Delta Y_t = \gamma_0 \Delta Y_{t-1} + \lambda_0 W \Delta Y_t + \Delta X_t \beta_0 + \Delta P_t \pi_0 + \Delta \varepsilon_t + \Delta V_{0t}$$

or

$$\begin{aligned} \Delta Y &= \gamma_0 \Delta Y_{-1} + \lambda_0 D_w \Delta Y + \Delta X \beta_0 + \Delta P \pi_0 + \Delta \varepsilon + \Delta V_0 \\ &= \Delta B \delta_0 + \Delta P \pi_0 + \Delta \varepsilon + \Delta V_0, \end{aligned}$$

where

$$\Delta B = (\Delta B'_1, \Delta B'_2, \dots, \Delta B'_T)' \text{ and } \Delta B_t = (\Delta Y_{t-1}, W \Delta Y_t, \Delta X_t).$$

The second way of stacking data is to stack data first by time then by individual. We use a bar over the variable to distinguish this stacking from the first one. Specifically, define

$$\Delta \bar{X}_i = (\Delta x_{i2}, \Delta x_{i3}, \dots, \Delta x_{iT})' \text{ and } \Delta \bar{X} = (\Delta \bar{X}'_1, \Delta \bar{X}'_2, \dots, \Delta \bar{X}'_N)'$$

Define $\bar{Y}_i, \bar{Y}, \bar{P}_i, \bar{P}, \bar{\varepsilon}_i, \bar{\varepsilon}, \bar{V}_{0i}, \bar{V}_0$ analogously.

Let $S = P(P'P)^{-1}P'$ denote the projection matrix onto the space spanned by P . Define $\bar{S}$ using $\bar{P}$. Partialing out the sieve approximation, we obtain

$$(I-S)\Delta Y = (I-S)\Delta B \delta_0 + (I-S)\Delta \varepsilon + (I-S)\Delta V_0.$$

Finally, we tackle the endogeneity issue by applying the 2SLS. Let h_i denote a column vector of instrumental variables

$$h_i = \begin{pmatrix} [h_{i1}] & & & 0 \\ & [h_{i2}] & & \\ & & \ddots & \\ 0 & & & [h_{i,T-2}] \end{pmatrix}$$

$h_{i1} = (y_{i1}, \sum_{j=1}^N w_{ij} x_{j1}, x'_{i1}, x'_{i2}); h_{i2} = (y_{i1}, y_{i2}, \sum_{j=1}^N w_{ij} x_{j1}, \sum_{j=1}^N w_{ij} x_{j2}, x'_{i1}, x'_{i2}, x'_{i3}); h_{i,T-2} = (y_{i1}, \dots, y_{i,T-2}, \sum_{j=1}^N w_{ij} x_{j1}, \dots, \sum_{j=1}^N w_{ij} x_{j,T-2}, x'_{i1}, \dots, x'_{i,T-1})$. By computing we could obtain that the dimension of h_i is $(T-1) \times \frac{(T-2)(2qT+T-1)}{2}$ where q denotes the dimension of x_{it} .

Define $H = (H'_1, \dots, H'_N)'$ and $M = H(H'H)^{-1}H'$. Define $\bar{H}, \bar{H}, \bar{M}$ analogously. The 2SLS estimators for the unknown parameters $\delta_0, \pi_0, g_0(z)$ are respectively given by

$$\begin{aligned} \hat{\delta} &= \{\Delta B' (I-S)M(I-S)\Delta B\}^{-1} \Delta B' (I-S)M(I-S)\Delta Y \\ &= \{\Delta \bar{B}' (I-\bar{S})\bar{M}(I-\bar{S})\Delta \bar{B}\}^{-1} \Delta \bar{B}' (I-\bar{S})\bar{M}(I-\bar{S})\Delta \bar{Y}; \quad (2) \end{aligned}$$

$$\begin{aligned} \hat{\pi} &= \{\Delta P' \Delta P\}^{-1} \Delta P' \{\Delta Y - \Delta B \hat{\delta}\} \\ &= \{\Delta \bar{P}' \Delta \bar{P}\}^{-1} \Delta \bar{P}' \{\Delta \bar{Y} - \Delta \bar{B} \hat{\delta}\}; \quad (3) \end{aligned}$$

$$\hat{g}(z) = p^K(z)' \hat{\pi}. \quad (4)$$

3 Asymptotic Properties

We now present sufficient conditions and establish the asymptotic properties of the proposed estimators. We begin with the assumptions that are normally imposed for the fixed-effects panel regression, for the spatial regression, and for the series regression. Define $\bar{X}_i = (x_{i1}, x_{i1}, \dots, x_{iT})$ and define $\bar{\varepsilon}_i, \bar{Z}_i$ similarly. Denote

$$\sum_i = E(\Delta \bar{\varepsilon}_i \Delta \bar{\varepsilon}_i' | \bar{X}_i, \bar{Z}_i, \bar{H}_i, W).$$

The following assumptions will be maintained throughout the paper.

Assumption 1

(i) The elements of x_{it} and z_{it} are nonstochastic and bounded, uniformly in i and t ; (ii) for all i , $E[\tilde{\varepsilon}_i' \tilde{X}_i \tilde{Z}_i \tilde{H}_i' W] = 0$; (iii) there exist some constants $\lambda_{\min}, \lambda_{\max} > 0$ such that $\lambda_{\max} I - \Sigma_i \lambda_{\min} I$ holds in positive semidefinite matrix sense uniformly over i ; (iv) for all (i, t) , $E(\tilde{\varepsilon}_{it})^4 | X_i, Z_i, H_i, W$ is uniformly bounded in absolute value.

Assumption 2

(i) The diagonal element of the spatial weighting matrix W is zero.; (ii) the matrix $I - \lambda_0 W$ is nonsingular with $|\lambda_0| < 1$, where I is an $N \times N$ identity matrix; (iii) the row and column sums of the matrices $W, (I - \lambda_0 W)^{-1}$ are bounded uniformly in absolute value; and (iv) with $C = W(I - \lambda_0 W)^{-1}$, there exists a constant λ_c such that $\lambda_c I - CC'$ is positive semidefinite for all N .

Assumption 1 is similar to those commonly assumed in the panel fixed-effects regression literature. It requires the explanatory variables (x, z), the instruments h and the spatial weighting matrix W to be strictly exogenous. While maintaining the cross sectional independence, arbitrary time series dependence is permitted. Moreover, The instruments are allowed to have cross sectional dependence. Clearly these conditions are too restrictive, ruling out predetermined and endogenous regressors as well as the predetermined or endogenous spatial weighting matrix. Extension of the proposed estimation to the case of predetermined or endogenous explanatory variables x is straightforward. However, extension to the case of predetermined/endogenous spatial weighting matrix is nontrivial. Assumption 1(iv) is needed for establishing the asymptotic distribution of the nonparametric estimator $\hat{g}(z)$. See Newey (1997).

Assumption 2 imposes restrictions on the spatial weighting matrix. These restrictions are commonly imposed in the spatial regression literature. The next set of conditions are imposed on the sieve approximations. Since the constant in the fixed-effects setting is not identified, we must impose some normalization on $g_0(\cdot)$ such as $g_0(z_0) = 0$ at some point z_0 so that $g_0(\cdot)$ can be identified. The approximating series $P^K(z)$ shall be constructed to satisfy this normalization. In addition, we impose the following conditions:

Assumption 3

For every K ; (i) the smallest eigenvalue of

$$E \left[(T-1)^{-1} \Delta \bar{P}_i' \Delta \bar{P}_i \right]$$

is bounded away from zero uniformly in K and (ii) there is a sequence of constants $\zeta_0(K)$ satisfying $\sup_{z \in \mathcal{S}} \|P^K(z)\| \leq \zeta_0(K)$, and $K = K(N)$ such that $\zeta_0(K)^2 K/N \rightarrow 0$ as $N \rightarrow \infty$.

Assumption 4

For any function $g(\cdot)$ satisfying the normalization of $g_0(\cdot)$, (i) there exist some $\alpha = \alpha(N)(> 0)$, $\pi = \pi(K)$, $\sup_{z \in \mathcal{Z}} |g(z) - P^K(z)' \pi| = O(K^{-\alpha})$ as $K \rightarrow \infty$ uniformly over all measurable function $g(\cdot)$; (ii) $\sqrt{N}K^{-\alpha} \rightarrow 0$ and $N \rightarrow \infty$. (iii) $g_0(z_{it})$ is uniformly bounded in absolute value.

Assumption 3 and 4 are familiar in the series regression literature (e.g., Newey (1997)) except that Assumption 3(i) is a generalization to the panel data setting. This condition allows $\{z_{it}, t = 1, 2, \dots, T\}$ to be time series dependent and nonidentically distributed. Nevertheless, approximating series of this sort can always be constructed to satisfy this condition through normalization. Without loss of generality, we will assume

$E [T^{-1} \bar{P}_i' \bar{P}_i] = I$. These conditions restrict the rate of K , K must increase with N . However, these conditions do not yield a unique choice of K , neither too fast nor too slow. In practice, we use the Leave-one-out cross-validation to select the series term K , that is, select $\hat{K}$ to minimize

$$CV_K = \sum_{i=1}^N \sum_{t=2}^T [y_{it} - b_{it} \hat{\delta}_{-i} - \hat{g}_{-i}(z_{it})] \quad (5)$$

where

$$\begin{aligned} \hat{\delta}_{-i} &= \{\Delta \bar{B}_{-i}' (I - \bar{S}) \bar{M} (I - \bar{S}) \Delta \bar{B}_{-i}\}^{-1} \Delta \bar{B}_{-i}' (I - \bar{S}) \bar{M} (I - \bar{S}) \Delta \bar{Y}_{-i}; \\ \hat{\pi}_{-i} &= \{\Delta \bar{P}_{-i}' \Delta \bar{P}_{-i}\}^{-1} \Delta \bar{P}_{-i}' \{\Delta \bar{Y}_{-i} - \Delta \bar{B}_{-i} \hat{\delta}_{-i}\}; \\ \hat{g}_{-i}(z_{it}) &= p^K(z_{it})' \hat{\pi}_{-i}. \end{aligned}$$

obtained by removing (x_i, z_i, y_i) .

Assumption 5

$\Gamma_N' \Gamma_N / [N(T-1)]$ converges in probability to a positive definite matrix Γ .

Define $D_c = \text{diag}\{C, \dots, C\}$,

$$\Delta B^* = (D_c(\Delta G_0 + \Delta X \beta_0), \Delta X).$$

Define B^* by restacking B^* first by time then by individual. Define

$$\begin{aligned}\Gamma'_N &= \Delta \overline{B}^{*'} (I - \overline{S}) \overline{M} (I - \overline{S}); \\ \Lambda_N &= \text{diag}\{\sum_1, \dots, \sum_N\}; \\ \Omega_N &= (\Gamma'_N \Gamma_N)^{-1} \Gamma'_N \Lambda_N \Gamma_N (\Gamma'_N \Gamma_N)^{-1}.\end{aligned}$$

Theorem 1

Suppose Assumption 1–5 are satisfied, then we can establish

$$\Omega_N^{-1/2} (\hat{\delta} - \delta_0) \rightarrow N(0, I) \text{ in distribution.}$$

For any given value z_0 , denote

$$\sigma_N^2(z_0) = p^K(z_0)' \left\{ \Delta \overline{P}' \Delta \overline{P} \right\}^{-1} \Delta \overline{P}' \Lambda_N \Delta \overline{P} \left\{ \Delta \overline{P}' \Delta \overline{P} \right\}^{-1} p^K(z_0).$$

Theorem 2

Suppose Assumption 1–5 are satisfied, then we can establish: (i)

$$\sup_z |\hat{g}(z) - g_0(z)| = O_p(\zeta_0(K) [\sqrt{K}/\sqrt{N} + K^{-\alpha}]); \text{ (ii) } E[\hat{g}(z) - g_0(z)]^2 = O_p([K/N + K^{-2\alpha}]);$$

$$\text{(iii) } \frac{\hat{g}(z) - g_0(z)}{\sigma_N(z_0)} \rightarrow N(0, 1) \text{ in distribution.}$$

The above two theorems allow us to conduct statistical inference on both δ_0 and $g_0(z)$, provided some consistent covariances are provided. Notice that the single spatial parameter is a too restrictive specification in some applications where we may wish to divide the cross sectional units into several groups. For each i , let $N_{i1}, N_{i2}, \dots, N_{iJ}$ denote a partition of $\{1, 2, \dots, N\}$ so that N_{i1} is the closed to i and N_{iJ} is the farther from i by some distance measure. The proposed estimation can be easily extended to the following more general model

Define

$$\begin{aligned}b_{it} &= \left(y_{i,t-1}, \sum_{j \in N_{i1}} w_{ij} y_{jt}, \sum_{j \in N_{i2}} w_{ij} y_{jt}, \dots, \sum_{j \in N_{i,J}} w_{ij} y_{jt}, x'_{it} \right)'; \\ \Delta B_t &= (\Delta b_{1t}, \Delta b_{2t}, \dots, \Delta b_{Nt})'; \\ \delta_0 &= (\lambda_{10}, \lambda_{20}, \dots, \lambda_{J0}, \beta'_0)'. \end{aligned}$$

With this modification, the estimators are still given by (2)–(4). With appropriate modification to B^* , Theorem 1–2 still hold true.

In some applications, the spatial weighting matrix may vary with time; in these applications we define

$$\Delta B_t = (\Delta Y_{t-1}, \Delta(W_t Y_t), \Delta X_t).$$

With this modification, the estimators are still given by (2)–(4). With appropriate modification to ΔB_t^* and Ω_N and σ_N^2 , Theorem 1–2 still hold true.

4 Covariance

We begin with the covariance matrix for the δ_0 estimator. The asymptotic derivation in the appendix suggests the following consistent estimator for Γ_N :

$$\hat{\Gamma}_N' = \Delta \bar{B}' (I - \bar{S}) \bar{M} (I - \bar{S}).$$

We estimate the residuals by $\hat{\varepsilon}_i = \bar{Y}_i - \bar{B}_i \bar{\delta} - \bar{P}_i \hat{\pi}$ and estimate Σ_i by $\hat{\Sigma}_i = \Delta \hat{\varepsilon}_i \Delta \hat{\varepsilon}_i'$. Denote

$$\begin{aligned} \hat{\Lambda}_N &= \text{diag}\{\hat{\Sigma}_1, \dots, \hat{\Sigma}_N\}; \\ \hat{\Omega}_N &= \left(\hat{\Gamma}_N' \hat{\Gamma}_N \right)^{-1} \hat{\Gamma}_N' \hat{\Lambda}_N \hat{\Gamma}_N \left(\hat{\Gamma}_N' \hat{\Gamma}_N \right)^{-1}. \end{aligned}$$

The following theorem establishes the consistency of Ω_N .

Theorem 3

Suppose Assumption 1–5 are satisfied, then we can establish

$$N(\hat{\Omega}_N - \Omega_N) = o_p(1).$$

Next, we estimate the variance $\sigma_N^2(z_0)$ by

$$\hat{\sigma}_N^2(z_0) = p^K(z_0)' \left\{ \Delta \bar{P}' \Delta \bar{P} \right\}^{-1} \Delta \bar{P}' \hat{\Lambda}_N \Delta \bar{P} \left\{ \Delta \bar{P}' \Delta \bar{P} \right\}^{-1} p^K(z_0).$$

The following theorem proves the consistency

Theorem 4

Suppose Assumption 1–5 are satisfied, then we can establish

$$\frac{N(\hat{\sigma}_N^2(z_0) - \sigma_N^2(z_0))}{\|p^K(z_0)\|^2} = o_p(1).$$

The above theorems imply

$$\begin{aligned}\hat{\Omega}_N^{-1/2} \left(\hat{\delta} - \delta_0 \right) &\rightarrow N(0, I) \text{ in distribution;} \\ \hat{\sigma}_N(z_0)^{-1} \left(\hat{g}(z_0) - g(z_0) \right) &\rightarrow N(0, 1) \text{ in distribution.}\end{aligned}$$

Therefore restrictions on both the finite dimensional parameters and the infinite dimensional parameter can be tested statistically by applying a Wald test.

5 Monte Carlo Simulation

In this section we report the results of a small scale Monte Carlo study in which we evaluate the finite sample properties of the proposed. The response variable is given by

$$y_{it} = \gamma_0 y_{i,t-1} + \lambda_0 \sum_{j=1}^N w_{ij} y_{jt} + x'_{it} \beta_0 + g_0(z_{it}) + c_i + \varepsilon_{it}, i=1, \dots, N, t=2, \dots, T. \quad (6)$$

The explanatory variables, the fixed-effects, and the error terms are generated as follows:

$x_{i1} = \eta_{i1} / \sqrt{1-0.7^2} + f_i / (1-0.7)$, and $x_{it} = 0.7x_{it-1} + f_i + \eta_{it}$, with η_{it} drawn independently from the uniform distribution over $[-2, 2]$; f_i are drawn independently from $N(0, 0.25)$, z_{it} are drawn independently from the uniform distribution over $[0, 1]$; $g_0(z) = e^{3z^2} \sin(3\pi z)$; $c_i \equiv f_i$ and ε_{it} are drawn independently from $N(0, 0.09)$. we take the spatial weights matrix as the spatial scenario in Case (1991) with R number of units, m members in each unit, and with each neighbor of a member in a unit given equal weight, i.e., $W = I_R \otimes B_m$, where

$B_m = (l_m l'_m - I_m) / (m-1)$, where l_m denotes m-dimensional column vector with all elements of ones, I denotes identity matrix, $m = 10$. The approximating series is the polynomial series $p^K(z) = (z, z^2, \dots, z^K)'$, we use equation(5) to select the optimal K .

Calculations were done in language R. For each parameter value, we generate a sample of size $N (=200,500)$ and compute the estimators (2)–(4). This step is repeated for 1000 times. Let BIAS denote the average of the estimation bias of the parameter estimates, let RMSE denote the root mean error, let SEE denote the sampling standard error of the estimates, and let ESE denote the estimated the asymptotic standard error. The integrated squared error (ISE) of the nonparametric estimator $\hat{g}$ is calculated as

$$ISE(\hat{g}) = E \left[|\hat{g}(z) - g_0(z)|^2 \right] = \int_z \{ \hat{g}(z) - g(z) \}^2 dz.$$

Table summarize the results for λ and β estimates.

It is evident from these tables that the proposed approach works quite well. The finite dimensional parameter estimate is approximately unbiased, The RMSE, ESE and SEE are close. The figure shows that the nonparametric estimator is quite close to the truth.

Acknowledgments

Work is supported by the National Natural Science Foundation of China (Grant No.71371118, 71471117), National Social Science Fund of China (Grant No.14BJY012), Program for Changjiang Scholars and Innovative Research Team in University (PCSIRTIRT13077), State Key Program of National Natural Science of China (Grant No. 71331006), NSF grants DMS-0905777 and DMS-1208978, NIH NIAID grant 2R37 AI054165-10.

References

1. Ai C, Chen X. Efficient Estimation of Models with Conditional Moment Restrictions Containing Unknown Functions. *Econometrica*. 2003; 71:1795–1843.
2. Ai, C., Li, Q. Semi-parametric and on-parametric Methods in Panel Data Models. In: Matyas, L., Sevestre, P., editors. *The Econometrics of Panel Data*. Springer-Verlag; Berlin: 2008. p. 625-660.
3. Ai C, YQ. Estimation of A Partly Specified Spatial Panel Data Model. *Forthcoming in Econometric Reviews*. 2014
4. Ahn SC, Schmidt P. Efficient Estimation of Models for Dynamic Panel Data. *Journal of Econometrics*. 1995; 68:5–27.
5. Anselin, L. Spatial Econometrics. In: Baltagi, B., editor. *A Companion to Theoretical Econometrics*. Blackwell; Oxford: 2001. p. 310-330.
6. Baltagi B, Song SH, Koh W. Testing panel data regression models with spatial error correlation. *Journal of Econometrics*. 2003; 117:123–150.
7. Baltagi B, Li D. Prediction in the panel data model with spatial correlation: the case of liquor. *Spatial Economic Analysis*. 2006; 1:175–185.
8. Baltagi, Badi, Pfaffermayr, Eggerand M. A Generalized Spatial panel Data Model with Random Effects[W]. *Syracuse University*; 2007.
9. Basile R, Gress B. Semi-parametric spatial auto-covariance models of regional growth behavior in Europe. *Region et Developpement*. 2005; 21:97–118.
10. Case AC. Spatial patterns in household demand. *Econometrica*. 1991; 59:953–965.
11. Cuzick J. Semiparametric additive regression. *J Royal Statist Soc, B*. 1992; 54:831–843.
12. Elhorst J, Paul. *Dynamic Models in Space and Time*. *Geographical Analysis*. 2001; 33:119–140.
13. Elhorst, JP. Matlab software for spatial panels. Paper presented at 4th world conference of the spatial econometric association; Chicago. 2010.
14. Elhorst J, Paul. Specification and Estimation of Spatial Panel Data Models. *Journal of Regional Science Review*. 2003; 26:223–243.
15. Ertur C, Koch W. Growth, technological interdependence and spatial externalities: theory and evidence. *Journal of Applied Econometrics*. 2007; 22:1033–1062.
16. Foote, CL. Space and time in macroeconomic panel data: young workers and state-level unemployment revisited. *Federal Reserve Bank of Boston*; 2007. Working Paper No. 07-10
17. Härdle, W., Liang, H., Gao, J. *Partially Linear Models*. New York: Springer; 2000.
18. Kapoor M, Kelejian H, Prucha I. Panel data models with spatially correlated error components. *Journal of Econometrics*. 2007; 140:97–130.
19. Kelejian HH, Prucha IR. A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbance. *Journal of Real Estate Finance and Economics*. 1998; 17:99–121.
20. Kelejian HH, Prucha IR. A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review*. 1999; 40:509–533.
21. Keller W, Shiue CH. The origin of spatial interaction. *Journal of Econometrics*. 2007; 140:304–332.
22. Korniotis GM. Estimating panel models with internal and external habit formation. *J Bus Econ Stat*. 2010; 28:145–158.
23. Kukenova, M., Monteiro, JA. Spatial dynamic panel model and system GMM: a Monte Carlo investigation. 2009. <http://ideas.repec.org/p/pramprapa/11569.html>

24. Lee LF, Yu J. Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics*. 2010a; 154:165–185.
25. Lee LF, Yu J. A spatial dynamic panel data with both time and individual fixed effects. *Econometric Theory*. 2010b; 26:564–597.
26. Lee, LF., Yu, J. Efficient GMM estimation of spatial dynamic panel data models with fixed effects. 2010c. <http://www.economics.smu.edu.sg/events/Paper/LungfeiLee.pdf>
27. Li Q, Stengos T. Semiparametric estimation of partially linear panel data models. *Journal of Econometrics*. 1996; 71:389–397.
28. Liang, Hua, Härdle, W., Carroll, RJ. Estimation in a semiparametric partially linear errors-in-variables model. *Ann Statist*. 1999; 27:1519–1535.
29. Su, Liangjun. Semiparametric GMM estimation of spatial autoregressive models. *Journal of Econometrics*. 2012; 167:543–560.
30. Newey WK. Convergence rates and asymptotic normality for series estimators. *Journal of Econometrics*. 1997; 79:147–168.
31. Parent O, LeSage JP. A space-time filter for panel data models containing random effects. *Comput Stat Data Anal*. 2011; 55:475–490.
32. Park, Byeong U., Sickles, Robin C., Simar, Leopold. Semiparametric efficient estimation of dynamic panel data models. *Journal of Econometrics*. 2007; 136:281–301.
33. Li, Qi, Kniesner, Thomas J. Nonlinearity in dynamic adjustment: Semiparametric estimation of panel labor supply. *Empirical Economics*. 2002; 27:131–148.
34. Severini TA, Wong WH. Generalized profile likelihood and conditional parametric models. *Ann Statist*. 1992; 20:1768–1802.
35. Su L, Jin S. Profile quasi-maximum likelihood estimation of spatial autoregressive models. *Journal of Econometrics*. 2010; 157:18–33.
36. Yu J, de Jong R, Lee LF. Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and T are large. *Journal of Econometrics*. 2008; 146:118–134.

6 Appendix. Technicalities

Before we prove the main results, we first establish some useful convergence results. By Assumption 3(i), we assume without loss of generality that:

$$E \left\{ \sum_{t=2}^T \frac{(p_{kK}(z_{it}) - p_{kK}(z_{it-1}))(p_{jK}(z_{it}) - p_{jK}(z_{it-1}))}{T-1} \right\} = I_{jk},$$

where I_{jk} denotes the (j, k) th element of the identity matrix. Let $\|\cdot\|$ denote the Euclidean

norm. With $Q = \frac{\Delta P' \Delta P}{N(T-1)}$, employing the same line of proof as in Newey (1997, pp. 161–162) and applying Assumption 1(i) and 3, we obtain

$$\begin{aligned}
& E \left[\|Q - I\|^2 \right] \\
&= \sum_{k=1}^K \sum_{j=1}^K E \left[\left\{ \sum_{i=1}^N \sum_{t=2}^T \frac{(p_{kK}(z_{it}) - p_{kK}(z_{it-1})) (p_{jK}(z_{it}) - p_{jK}(z_{it-1}))}{N(T-1)} - I_{jk} \right\}^2 \right] \\
&= \sum_{k=1}^K \sum_{j=1}^K \frac{1}{N^2} \sum_{i=1}^N E \left[\left\{ \sum_{t=2}^T \frac{(p_{kK}(z_{it}) - p_{kK}(z_{it-1})) (p_{jK}(z_{it}) - p_{jK}(z_{it-1}))}{(T-1)} - I_{jk} \right\}^2 \right] \\
&\leq \sum_{k=1}^K \sum_{j=1}^K \frac{1}{N^2} \sum_{i=1}^N E \left[\left\{ \sum_{t=2}^T \frac{(p_{kK}(z_{it}) - p_{kK}(z_{it-1})) (p_{jK}(z_{it}) - p_{jK}(z_{it-1}))}{T-1} \right\}^2 \right] \\
&\leq \sum_{k=1}^K \sum_{j=1}^K \frac{1}{N(T-1)} \sum_{t=2}^T E \left[(p_{kK}(z_{it}) - p_{kK}(z_{it-1}))^2 (p_{jK}(z_{it}) - p_{jK}(z_{it-1}))^2 \right] \\
&= E \left[\frac{2\zeta_0(K)^2}{N(T-1)} \sum_{t=2}^T E \left\{ \sum_{j=1}^K (p_{jK}(z_{it}) - p_{jK}(z_{it-1}))^2 \right\} \right] \leq \frac{2\zeta_0(K)^2 K}{N(T-1)} \rightarrow 0.
\end{aligned}$$

Hence

$$\|Q - I\| = O_p(\zeta_0(K) \sqrt{K} / \sqrt{N}) = o_p(1). \quad (7)$$

Let 1_N denote the indicator function for the smallest eigenvalue of Q being greater than $1/2$. Then $\lim_{N \rightarrow \infty} P(1_N = 1) = 1$. By Assumption 1, we obtain

$$\begin{aligned}
& E \left[1_N \left\| \frac{Q^{-1/2} \Delta P' \Delta \varepsilon}{N(T-1)} \right\|^2 \right] = E \left[\frac{1_N \Delta \varepsilon' \Delta P (\Delta P' \Delta P)^{-1} \Delta P' \Delta \varepsilon}{N(T-1)} \right] \\
&= E \left[\frac{1_N \text{tr}\{(\Delta P' \Delta P)^{-1/2} \Delta P' E[\Delta \varepsilon \Delta \varepsilon' | Z] \Delta P (\Delta P' \Delta P)^{-1/2}\}}{N(T-1)} \right] \\
&\leq \lambda_{\max} E \left[\frac{1_N \text{tr}\{I\}}{N(T-1)} \right] = O \left(\frac{K}{N(T-1)} \right).
\end{aligned}$$

Hence,

$$1_N \|Q^{-1/2} \Delta P' \Delta \varepsilon / N\| = O_p(\sqrt{K} / \sqrt{N}) \text{ and} \quad (8)$$

$$1_N \|Q^{-1} \Delta P' \Delta \varepsilon / N\| = O_p(\sqrt{K} / \sqrt{N}) \quad (9)$$

by applying (7). Employing the same techniques, we obtain

$$1_N \|Q^{-1} \Delta P' \Delta V_0 / N\| = O_p(K^{-\alpha}). \quad (10)$$

Proof of Theorem 1

Substituting for Y , we obtain

$$\begin{aligned}\hat{\delta} &= \{\Delta B' (I-S)M(I-S)\Delta B\}^{-1} \Delta B' (I-S)M(I-S)\Delta Y \\ &= \delta_0 + \{\Delta B' (I-S)M(I-S)\Delta B\}^{-1} \Delta B' (I-S)M(I-S)\Delta \varepsilon + \{\Delta B' (I-S)M(I-S)\Delta B\}^{-1} \Delta B' (I-S)M(I-S)\Delta V_0.\end{aligned}$$

Consider the term $\Delta B' (I-S)M(I-S)\Delta \varepsilon$ first. Assumption 1 implies $E[\varepsilon \varepsilon' | Z, H] < \lambda_{\max} I$ and

$$\begin{aligned}& E \left[\frac{\Delta \varepsilon' (I-S)M(I-S)\Delta \varepsilon}{N} \right] \\ &= \text{trace} \left\{ E \left[\frac{(H' H)^{-1/2} H' (I-S)\Delta \varepsilon \Delta \varepsilon' (I-S)H (H' H)^{-1/2}}{N} \right] \right\} \\ &\leq \lambda_{\max} \text{trace} \left\{ E \left[\frac{(H' H)^{-1/2} H' (I-S)H (H' H)^{-1/2}}{N} \right] \right\} \\ &\leq \lambda_{\max} \text{trace} \left\{ E \left[\frac{(H' H)^{-1/2} H' H (H' H)^{-1/2}}{N} \right] \right\} = O(N^{-1}).\end{aligned}$$

Hence

$$\frac{\Delta \varepsilon' (I-S)M(I-S)\Delta \varepsilon}{N} = O_p(N^{-1}). \quad (11)$$

Employing the same line of proof and applying Assumption 2 we obtain

$$\begin{aligned}& E \left[\frac{\Delta \varepsilon' D_c' (I-S)M(I-S)D_c \Delta \varepsilon}{N} \right] \\ &= \text{trace} \left\{ E \left[\frac{(H' H)^{-1/2} H' D_c (I-S)\Delta \varepsilon \Delta \varepsilon' (I-S)D_c' H (H' H)^{-1/2}}{N} \right] \right\} \\ &\leq \lambda_{\max} \text{trace} \left\{ E \left[\frac{(H' H)^{-1/2} H' D_c (I-S)D_c' H (H' H)^{-1/2}}{N} \right] \right\} \\ &\leq \lambda_{\max} \text{trace} \left\{ E \left[\frac{(H' H)^{-1/2} H' D_c D_c' H (H' H)^{-1/2}}{N} \right] \right\} \\ &\leq \lambda_{\max} \lambda_c \text{trace} \left\{ E \left[\frac{(H' H)^{-1/2} H' H (H' H)^{-1/2}}{N} \right] \right\} = O(N^{-1}).\end{aligned}$$

Hence

$$\frac{\Delta \varepsilon' D_c' (I-S)M(I-S)D_c \Delta \varepsilon}{N} = O_p(N^{-1}). \quad (12)$$

Applying the triangular inequality and applying (11) and (12), we obtain

$$\frac{\Delta \varepsilon' D_c' (I-S) M (I-S) \Delta \varepsilon}{N} = O_p(N^{-1}). \quad (13)$$

Recall that

$$\Delta B^* = (D_c(\Delta G_0 + \Delta X \beta_0), \Delta X).$$

Then

$$\begin{aligned} & E \left[\left\| \frac{\Delta B^* (I-S) M (I-S) \Delta \varepsilon}{N(T-1)} \right\|^2 \right] \\ &= \text{trace} \left\{ E \left[\frac{\Delta B^* (I-S) M (I-S) \Delta \varepsilon \Delta \varepsilon' (I-S) M (I-S) \Delta B^*}{N^2(T-1)^2} \right] \right\} \\ &\leq \lambda_{\max} \text{trace} \left\{ E \left[\frac{\Delta B^* (I-S) M (I-S) M (I-S) \Delta B^*}{N^2(T-1)^2} \right] \right\} \\ &\leq \lambda_{\max} \text{trace} \left\{ E \left[\frac{\Delta B^* (I-S) M (I-S) \Delta B^*}{N^2(T-1)^2} \right] \right\} = O_p(N^{-1}) \end{aligned}$$

by Assumption 5. Hence

$$\frac{\Delta B^* (I-S) M (I-S) \Delta \varepsilon}{N} = O_p(N^{-1/2}). \quad (14)$$

Let e denote the column vector with the first element equal to one and all other elements equal to zero. Then we can write $B = B^* + D_c e e'$ and

$$\begin{aligned} \frac{\Delta B' (I-S) M (I-S) \Delta \varepsilon}{N} &= \frac{\Delta B^* (I-S) M (I-S) \Delta \varepsilon}{N} + \frac{e \Delta \varepsilon' D_c' (I-S) M (I-S) \Delta \varepsilon}{N} \\ &= O_p(N^{-1/2}) + O_p(N^{-1}) = O_p(N^{-1/2}) \end{aligned}$$

by applying (13) and (14).

Next, we write

$$\begin{aligned}
& \Delta B' (I-S)M(I-S)\Delta B \\
&= \Delta B'^* (I-S)M(I-S)\Delta B^* + e\Delta \varepsilon' D_c' (I-S)M(I-S)D_c \Delta \varepsilon e' \\
&+ \Delta B'^* (I-S)M(I-S)D_c \Delta \varepsilon e' + e\Delta \varepsilon' D_c' (I-S)M(I-S)\Delta B^* \\
&= \Delta B'^* (I-S)M(I-S)\Delta B^* + O_p(N^{1/2})
\end{aligned}$$

by applying (13) and (14). Hence,

$$\frac{\Delta B' (I-S)M(I-S)\Delta B}{N} = \frac{\Delta B'^* (I-S)M(I-S)\Delta B^*}{N} + O_p(N^{-1/2}). \quad (15)$$

Notice that for any constant vector c satisfying $c'c = 1$

$$\left| \frac{c' \Delta B' (I-S)M(I-S)\Delta V_0}{N} \right| \leq \sqrt{\frac{c' \Delta B' (I-S)M(I-S)\Delta B c}{N}} \sqrt{\frac{\Delta V_0' (I-S)M(I-S)\Delta V_0}{N}} \quad (16)$$

$$\leq O_p(1) \left\| \frac{\Delta V_0' \Delta V_0}{N} \right\| = O_p(K^{-\alpha}) \quad (17)$$

by applying (10). Combining (14)–(17), we obtain

$$\begin{aligned}
\hat{\delta} &= \delta_0 + \left(\Delta B'^* (I-S)M(I-S)\Delta B^* \right)^{-1} \Delta B'^* (I-S)M(I-S)\Delta \varepsilon + O_p(K^{-\alpha} + N^{-1}) \\
&= \delta_0 + \left(\Delta \bar{B}'^* (I-\bar{S})\bar{M}(I-\bar{S})\Delta \bar{B}^* \right)^{-1} \Delta \bar{B}'^* (I-\bar{S})\bar{M}(I-\bar{S})\Delta \bar{\varepsilon} + O_p(K^{-\alpha} + N^{-1}).
\end{aligned}$$

Recall that

$$\begin{aligned}
\Gamma_N' &= \Delta \bar{B}'^* (I-\bar{S})\bar{M}(I-\bar{S}); \\
\Lambda_N &= \text{diag}\{\sum_1, \dots, \sum_N\}; \\
\Omega_N &= (\Gamma_N' \Gamma_N)^{-1} \Gamma_N' \Lambda_N \Gamma_N (\Gamma_N' \Gamma_N)^{-1}.
\end{aligned}$$

Then

$$\hat{\delta} - \delta_0 = (\Gamma'_N \Gamma_N)^{-1} \Gamma'_N \Delta \varepsilon + O_p(K^{-\alpha} + N^{-1}). \quad (18)$$

By Assumption 1 and 5, $\mathbf{\Lambda}\mathbf{\Omega}_N$ converges to a positive definite matrix and $\sqrt{N}K^{-\alpha} \rightarrow 0$. The theorem is now proven by application of a central limit theorem.

Proof of Theorem 2

We write

$$\begin{aligned} \hat{\pi} &= \{\Delta P' \Delta P\}^{-1} \Delta P' \{\Delta Y - \Delta B \hat{\delta}\} \\ &= \{\Delta P' \Delta P\}^{-1} \Delta P' \{\Delta Y - \Delta B \delta_0\} + \{\Delta P' \Delta P\}^{-1} \Delta P' \Delta B (\hat{\delta} - \delta_0). \end{aligned}$$

Consider the second term. Notice that the largest eigenvalue of $\frac{\Delta P' \Delta P}{N(T-1)}$ converges to 1 in probability. We have

$$\begin{aligned} & \|(\Delta P' \Delta P)^{-1} \Delta P' \Delta B (\hat{\delta} - \delta_0)\|^2 \\ &= (\hat{\delta} - \delta_0)' \frac{\Delta B' \Delta P}{N(T-1)} \left(\frac{\Delta P' \Delta P}{N(T-1)} \right)^{-2} \frac{\Delta P' \Delta B}{N(T-1)} (\hat{\delta} - \delta_0) \\ &= O_p\left(\frac{1}{N(T-1)}\right) (\hat{\delta} - \delta_0)' \Delta B' S \Delta B (\hat{\delta} - \delta_0) \\ &= O_p\left(\frac{1}{N(T-1)}\right) \text{ by Assumption 5.} \end{aligned}$$

Thus

$$\begin{aligned} \hat{\pi} &= \{\Delta P' \Delta P\}^{-1} \Delta P' \{\Delta Y - \Delta B \delta_0\} + O_p(N^{-1/2}) \\ &= \pi_0 + \{\Delta P' \Delta P\}^{-1} \Delta P' \Delta \varepsilon + \{\Delta P' \Delta P\}^{-1} \Delta P' \Delta V_0 + O_p(N^{-1/2}) \\ &= \pi_0 + O_p(\sqrt{K/N} + K^{-\alpha}) \end{aligned}$$

by applying (10). Notice that

$$\begin{aligned} 1_N |\hat{g}(z) - g_0(z)| &\leq |p^K(z)' (\hat{\pi} - \pi_0)| + |p^K(z)' \pi_0 - g_0(z)| \\ &\leq \zeta_0(K) 1_N \|\hat{\pi} - \pi_0\| + O(K^{-\alpha}) = O_p(\zeta_0(K) [\sqrt{K}/\sqrt{N} + K^{-\alpha}]). \end{aligned}$$

We obtain

$$\hat{g}(z) = g_0(z) + O_p(\zeta_0(K) [\sqrt{K}/\sqrt{N} + K^{-\alpha}])$$

uniformly over z and

$$\sup_{z \in Z} |\hat{g}(z) - g_0(z)| = O_p(\zeta_0(K) [\sqrt{K}/\sqrt{N} + K^{-\alpha}]).$$

Furthermore, applying (7) and Assumption 3(ii), we obtain

$$\begin{aligned} 1_N E [(\hat{g}(z) - g_0(z))^2] &= 1_N E [p^K(z)' (\hat{\pi} - \pi_0) + (p^K(z)' \pi_0 - g_0(z))]^2 \\ &\leq 2 * 1_N \|\hat{\pi} - \pi_0\|^2 + 2 * 1_N E [p^K(z)' \pi_0 - g_0(z)]^2 \\ &= O_p(K/N + K^{-2\alpha}) + O(K^{-2\alpha}) \\ &= O_p(K/N + K^{-2\alpha}). \end{aligned}$$

Notice that

$$p^K(z_0)' (\Delta P' \Delta P)^{-1} \Delta P' \Delta B (\hat{\delta} - \delta_0) = O_p(\zeta_0(K) N^{-1/2}).$$

By substitution and at a fixed z_0 , we have

$$\begin{aligned} &\hat{g}(z_0) - g_0(z_0) \\ &= p^K(z_0)' \hat{\pi} - [p^K(z_0)' \pi_0 + O(K^{-\alpha})] \\ &= p^K(z_0)' [\pi_0 + \{\Delta P' \Delta P\}^{-1} \Delta P' \Delta \varepsilon + \{\Delta P' \Delta P\}^{-1} \Delta P' \Delta V_0 + O_p(N^{-1/2})] - [p^K(z_0)' \pi_0 + O(K^{-\alpha})] \\ &= p^K(z_0)' \{\Delta P' \Delta P\}^{-1} \Delta P' \Delta \varepsilon + O_p(\zeta_0(K) [N^{-1/2} + K^{-\alpha}]) \\ &= p^K(z_0)' \{\Delta \bar{P}' \Delta \bar{P}\}^{-1} \Delta \bar{P}' \Delta \bar{\varepsilon} + O_p(\zeta_0(K) [N^{-1/2} + K^{-\alpha}]). \end{aligned}$$

Recall that

$$\begin{aligned} \sigma_N^2(z_0) &= E \left[\left(p^K(z_0)' \{\Delta \bar{P}' \Delta \bar{P}\}^{-1} \Delta \bar{P}' \Delta \bar{\varepsilon} \right)^2 | Z \right] \\ &= p^K(z_0)' \{\Delta \bar{P}' \Delta \bar{P}\}^{-1} \Delta \bar{P}' \Lambda_N \Delta \bar{P} \{\Delta \bar{P}' \Delta \bar{P}\}^{-1} p^K(z_0). \end{aligned}$$

By Assumption 1,

$$\begin{aligned} \sigma_N^2(z_0) &\geq \lambda_{\min} p^K(z_0)' \{\Delta \bar{P}' \Delta \bar{P}\}^{-1} p^K(z_0) \\ &= O_p(N^{-1}) \|p^{(K)}(z_0)\|^2. \end{aligned}$$

Similar, we can also show

$$\sigma_N^2(z_0) \leq O_p(N^{-1}) \|p^K(z_0)\|^2.$$

Applying the Lindbergh-Feller central limit theorem,

$$\frac{p^K(z_0)' \left\{ \Delta \bar{P}' \Delta \bar{P} \right\}^{-1} \Delta \bar{P}' \Delta \bar{\varepsilon}}{\sigma_N(z_0)} \rightarrow N(0, 1) \text{ in distribution.}$$

Theorem 2(iv) now follows from Assumption 4.

Proof of Theorem 3

Write

$$N \hat{\Omega}_N = \left(\frac{\hat{\Gamma}_N' \hat{\Gamma}_N}{N} \right)^{-1} \frac{\hat{\Gamma}_N' \hat{\Lambda}_N \hat{\Gamma}_N}{N} \left(\frac{\hat{\Gamma}_N' \hat{\Gamma}_N}{N} \right)^{-1}.$$

We complete the proof by showing

$$\begin{aligned} \frac{\hat{\Gamma}_N' \hat{\Gamma}_N}{N} - \frac{\Gamma_N' \Gamma_N}{N} &= o_p(1); \\ \frac{\hat{\Gamma}_N' \hat{\Lambda}_N \hat{\Gamma}_N}{N} - \frac{\Gamma_N' \Lambda_N \Gamma_N}{N} &= o_p(1). \end{aligned}$$

Notice that by substituting for $B = B^* + D_c \varepsilon \varepsilon'$ we obtain

$$\hat{\Gamma}_N' - \Gamma_N' = e \Delta \varepsilon' D_c' (I - S) M (I - S).$$

Applying (12) yields

$$\begin{aligned} & \frac{(\hat{\Gamma}_N - \Gamma_N)' (\hat{\Gamma}_N - \Gamma_N)}{N} \\ &= \frac{e \Delta \varepsilon' D_c' (I - S) M (I - S) M (I - S) D_c \Delta \varepsilon e'}{N} \\ &\leq \frac{e \Delta \varepsilon' D_c' (I - S) M (I - S) D_c \Delta \varepsilon e'}{N} = O_p(N^{-1}). \end{aligned}$$

$$\begin{aligned}
& E \left\{ \left\| \frac{(\hat{\Gamma}_N - \Gamma_N)' \Gamma_N}{N} \right\|_E^2 \right\} \\
&= \text{trace} \left\{ E \left[\frac{\Gamma_N' (I-S) M (I-S) D_c \Delta \varepsilon \varepsilon' e \Delta \varepsilon' D_c' (I-S) M (I-S) \Gamma_N}{N^2} \right] \right\} \\
&\leq \text{trace} \left\{ E \left[\frac{\Gamma_N' (I-S) M (I-S) D_c \Delta \varepsilon \varepsilon' D_c' (I-S) M (I-S) \Gamma_N}{N^2} \right] \right\} \\
&\leq \lambda_{\max} \text{trace} \left\{ E \left[\frac{\Gamma_N' (I-S) M (I-S) D_c D_c' (I-S) M (I-S) \Gamma_N}{N^2} \right] \right\} \\
&\leq \lambda_{\max} \lambda_c * \text{trace} \left\{ E \left[\frac{\Gamma_N' (I-S) M (I-S) M (I-S) \Gamma_N}{N^2} \right] \right\} \\
&= O(N^{-1}) \text{ by Assumption 5.}
\end{aligned}$$

Hence

$$\frac{(\hat{\Gamma}_N - \Gamma_N)' \Gamma_N}{N} = O_p(N^{-1/2}).$$

This implies

$$\begin{aligned}
& \frac{\hat{\Gamma}_N' \hat{\Gamma}_N}{N} - \frac{\hat{\Gamma}_N' \Gamma_N}{N} = \frac{(\hat{\Gamma}_N - \Gamma_N)' (\hat{\Gamma}_N - \Gamma_N)}{N} + \frac{\Gamma_N' (\hat{\Gamma}_N - \Gamma_N)}{N} + \frac{(\hat{\Gamma}_N - \Gamma_N)' \Gamma_N}{N} \\
&= O_p(N^{-1}) + O_p(N^{-1/2}) + O_p(N^{-1/2}) = O_p(N^{-1/2}) = o_p(1).
\end{aligned}$$

Write

$$\Delta \hat{\varepsilon}_i \Delta \hat{\varepsilon}_i' - \Delta \bar{\varepsilon}_i \Delta \bar{\varepsilon}_i' = (\Delta \hat{\varepsilon}_i - \Delta \bar{\varepsilon}_i) (\Delta \hat{\varepsilon}_i - \Delta \bar{\varepsilon}_i)' + \Delta \bar{\varepsilon}_i (\Delta \hat{\varepsilon}_i - \Delta \bar{\varepsilon}_i)' + (\Delta \hat{\varepsilon}_i - \Delta \bar{\varepsilon}_i) \Delta \bar{\varepsilon}_i'$$

Notice that

$$\Delta \hat{\varepsilon}_i - \Delta \bar{\varepsilon}_i = \Delta \bar{B}_i (\delta_0 - \hat{\delta}) + \Delta \bar{P}_i (\pi_0 - \hat{\pi}) + \Delta \bar{V}_{i0} \quad (19)$$

where $\bar{V}_{i0} = (v_{i02}, \dots, v_{i0T})'$. Then

$$\begin{aligned}
& (\Delta \hat{\varepsilon}_i - \Delta \bar{\varepsilon}_i) (\Delta \hat{\varepsilon}_i - \Delta \bar{\varepsilon}_i)' \leq 2 \left(\Delta \bar{B}_i (\hat{\delta} - \delta_0) (\hat{\delta} - \delta_0)' \Delta \bar{B}_i' + \Delta \bar{P}_i (\pi_0 - \hat{\pi}) (\pi_0 - \hat{\pi})' \Delta \bar{P}_i' + \Delta \bar{V}_{i0} \Delta \bar{V}_{i0}' \right); \\
& \Delta \bar{\varepsilon}_i (\Delta \hat{\varepsilon}_i - \Delta \bar{\varepsilon}_i)' = \Delta \bar{\varepsilon}_i (\delta_0 - \hat{\delta})' \Delta \bar{B}_i' + \Delta \bar{\varepsilon}_i (\pi_0 - \hat{\pi})' \Delta \bar{P}_i' + \Delta \bar{\varepsilon}_i \Delta \bar{V}_{i0}'.
\end{aligned}$$

Let $\hat{H}_i$ denote the fitted values from regressing $\bar{H}_i$ on $\bar{P}_i$. Then

$$\begin{aligned}
& \frac{1}{N} \sum_{i=1}^N (\bar{H}_i - \hat{H}_i)' (\Delta \hat{\varepsilon}_i \Delta \hat{\varepsilon}_i' - \Delta \bar{\varepsilon}_i \Delta \bar{\varepsilon}_i') (\bar{H}_i - \hat{H}_i) \\
&= \frac{1}{N} \sum_{i=1}^N (\bar{H}_i - \hat{H}_i)' (\Delta \hat{\varepsilon}_i - \Delta \bar{\varepsilon}_i) (\Delta \hat{\varepsilon}_i - \Delta \bar{\varepsilon}_i)' (\bar{H}_i - \hat{H}_i) \\
&\quad + \frac{1}{N} \sum_{i=1}^N (\bar{H}_i - \hat{H}_i)' \Delta \bar{\varepsilon}_i (\Delta \hat{\varepsilon}_i - \Delta \bar{\varepsilon}_i)' (\bar{H}_i - \hat{H}_i) \\
&\quad + \frac{1}{N} \sum_{i=1}^N (\bar{H}_i - \hat{H}_i)' (\Delta \hat{\varepsilon}_i - \Delta \bar{\varepsilon}_i) \Delta \bar{\varepsilon}_i' (\bar{H}_i - \hat{H}_i).
\end{aligned}$$

Consider the first term. Notice that applying the triangular inequality we obtain

$$(\Delta \hat{\varepsilon}_i - \Delta \bar{\varepsilon}_i) (\Delta \hat{\varepsilon}_i - \Delta \bar{\varepsilon}_i)' \leq 2\Delta \bar{B}_i (\hat{\delta} - \delta_0) (\hat{\delta} - \delta_0)' \Delta \bar{B}_i' + 2\Delta \bar{P}_i (\pi_0 - \hat{\pi}) (\pi_0 - \hat{\pi})' \Delta \bar{P}_i' + 2\Delta \bar{V}_{i0} \Delta \bar{V}_{i0}'.$$

Hence

$$\begin{aligned}
& \frac{1}{N} \sum_{i=1}^N (\bar{H}_i - \hat{H}_i)' (\Delta \hat{\varepsilon}_i - \Delta \bar{\varepsilon}_i) (\Delta \hat{\varepsilon}_i - \Delta \bar{\varepsilon}_i)' (\bar{H}_i - \hat{H}_i) \\
&\leq \frac{2}{N} \sum_{i=1}^N (\bar{H}_i - \hat{H}_i)' \Delta \bar{B}_i (\hat{\delta} - \delta_0) (\hat{\delta} - \delta_0)' \Delta \bar{B}_i' (\bar{H}_i - \hat{H}_i) \\
&\quad + \frac{2}{N} \sum_{i=1}^N (\bar{H}_i - \hat{H}_i)' \Delta \bar{P}_i (\pi_0 - \hat{\pi}) (\pi_0 - \hat{\pi})' \Delta \bar{P}_i' (\bar{H}_i - \hat{H}_i) \\
&\quad + \frac{2}{N} \sum_{i=1}^N (\bar{H}_i - \hat{H}_i)' \Delta \bar{V}_{i0} \Delta \bar{V}_{i0}' (\bar{H}_i - \hat{H}_i).
\end{aligned}$$

Since the largest eigenvalue of $(\hat{\delta} - \delta_0) (\hat{\delta} - \delta_0)'$ is $O_p(N^{-1})$, we obtain

$$\begin{aligned}
& \frac{1}{N} \sum_{i=1}^N (\bar{H}_i - \hat{H}_i)' \Delta \bar{B}_i (\hat{\delta} - \delta_0) (\hat{\delta} - \delta_0)' \Delta \bar{B}_i' (\bar{H}_i - \hat{H}_i) \\
&\leq \frac{1}{N} \sum_{i=1}^N (\bar{H}_i - \hat{H}_i)' \Delta \bar{B}_i \Delta \bar{B}_i' (\bar{H}_i - \hat{H}_i) \times O_p(N^{-1}) \\
&= O_p(1) \times O_p(N^{-1}) = o_p(1) \text{ by Assumption 2.}
\end{aligned}$$

Since the largest eigenvalue of $(\pi_0 - \hat{\pi}) (\pi_0 - \hat{\pi})'$ is $O_p(N^{-1}K + K^{-2\alpha})$ and

$$\Delta \bar{P}_i \Delta \bar{P}_i' \leq \|\Delta \bar{P}_i\|^2 I \leq \varsigma_0(K)^2 I, \text{ we obtain}$$

$$\begin{aligned}
& \frac{1}{N} \sum_{i=1}^N (\bar{H}_i - \hat{H}_i)' \Delta \bar{P}_i (\pi_0 - \hat{\pi}) (\pi_0 - \hat{\pi})' \Delta \bar{P}_i' (\bar{H}_i - \hat{H}_i) \\
&\leq \frac{1}{N} \sum_{i=1}^N (\bar{H}_i - \hat{H}_i)' \Delta \bar{P}_i \Delta \bar{P}_i' (\bar{H}_i - \hat{H}_i) \times O_p(N^{-1}K + K^{-2\alpha}) \\
&= \frac{1}{N} \sum_{i=1}^N (\bar{H}_i - \hat{H}_i)' (\bar{H}_i - \hat{H}_i) \times O_p(\varsigma_0(K)^2 N^{-1}K + \varsigma_0(K)^2 K^{-2\alpha}) \\
&= o_p(1) \text{ by Assumption 3 and Assumption 4(ii).}
\end{aligned}$$

Since the largest eigenvalue of $\Delta \bar{V}_{i0} \Delta \bar{V}_{i0}'$ is $O(K^{-2\alpha})$ by Assumption 4, we obtain

$$\begin{aligned} & \frac{1}{N} \sum_{i=1}^N (\bar{H}_i - \hat{H}_i)' \Delta \bar{V}_{i0} \Delta \bar{V}_{i0}' (\bar{H}_i - \hat{H}_i) \\ & \leq \frac{1}{N} \sum_{i=1}^N (\bar{H}_i - \hat{H}_i)' (\bar{H}_i - \hat{H}_i) \times O(K^{-2\alpha}) \\ & = o_p(1) \text{ by Assumption 4.} \end{aligned}$$

By substitution, we have

$$\begin{aligned} & \frac{1}{N} \sum_{i=1}^N (\bar{H}_i - \hat{H}_i)' \Delta \bar{\varepsilon}_i (\Delta \hat{\varepsilon}_i - \Delta \bar{\varepsilon}_i)' (\bar{H}_i - \hat{H}_i) \\ & = \frac{1}{N} \sum_{i=1}^N (\bar{H}_i - \hat{H}_i)' \Delta \bar{\varepsilon}_i (\delta_0 - \hat{\delta})' \Delta \bar{B}_i' (\bar{H}_i - \hat{H}_i) \\ & + \frac{1}{N} \sum_{i=1}^N (\bar{H}_i - \hat{H}_i)' \Delta \bar{\varepsilon}_i (\pi_0 - \hat{\pi})' \Delta \bar{P}_i' (\bar{H}_i - \hat{H}_i) \\ & + \frac{1}{N} \sum_{i=1}^N (\bar{H}_i - \hat{H}_i)' \Delta \bar{\varepsilon}_i \Delta \bar{V}_{i0}' (\bar{H}_i - \hat{H}_i). \end{aligned}$$

Let I_s denote the s th column of the identity matrix I . Applying the Cauchy inequality, the (s, m) th element is bounded by

$$\begin{aligned} & \left| \frac{1}{N} \sum_{i=1}^N I_s' (\bar{H}_i - \hat{H}_i)' \Delta \bar{\varepsilon}_i (\delta_0 - \hat{\delta})' \Delta \bar{B}_i' (\bar{H}_i - \hat{H}_i) I_m \right| \\ & \leq \sqrt{\frac{1}{N} \sum_{i=1}^N I_s' (\bar{H}_i - \hat{H}_i)' \Delta \bar{\varepsilon}_i \Delta \bar{\varepsilon}_i' (\bar{H}_i - \hat{H}_i) I_s} \times \\ & \sqrt{\frac{1}{N} \sum_{i=1}^N I_m' (\bar{H}_i - \hat{H}_i)' \Delta \bar{B}_i (\delta_0 - \hat{\delta}) (\delta_0 - \hat{\delta})' \Delta \bar{B}_i' (\bar{H}_i - \hat{H}_i) I_m} \\ & = O_p(1) \times O_p(N^{-1/2}) = o_p(1). \end{aligned}$$

Employing the same line of proof, we have

$$\begin{aligned} & \left| \frac{1}{N} \sum_{i=1}^N I_s' (\bar{H}_i - \hat{H}_i)' \Delta \bar{\varepsilon}_i (\pi_0 - \hat{\pi})' \Delta \bar{P}_i' (\bar{H}_i - \hat{H}_i) I_m \right| \\ & \leq \sqrt{\frac{1}{N} \sum_{i=1}^N I_s' (\bar{H}_i - \hat{H}_i)' \Delta \bar{\varepsilon}_i \Delta \bar{\varepsilon}_i' (\bar{H}_i - \hat{H}_i) I_s} \times \\ & \sqrt{\frac{1}{N} \sum_{i=1}^N I_m' (\bar{H}_i - \hat{H}_i)' \Delta \bar{P}_i (\pi_0 - \hat{\pi}) (\pi_0 - \hat{\pi})' \Delta \bar{P}_i' (\bar{H}_i - \hat{H}_i) I_m} \\ & = O_p \left(\zeta_0(K) \left[\sqrt{\frac{K}{N}} + K^{-\alpha} \right] \right); \end{aligned}$$

$$\begin{aligned}
& \left| \frac{1}{N} \sum_{i=1}^N I'_s (\bar{H}_i - \hat{H}_i)' \Delta \bar{\varepsilon}_i \Delta \bar{V}'_{i0} (\bar{H}_i - \hat{H}_i) I_m \right| \\
& \leq \sqrt{\frac{1}{N} \sum_{i=1}^N I'_s (\bar{H}_i - \hat{H}_i)' \Delta \bar{\varepsilon}_i \Delta \bar{\varepsilon}'_i (\bar{H}_i - \hat{H}_i) I_s} \times \\
& \sqrt{\frac{1}{N} \sum_{i=1}^N I'_m (\bar{H}_i - \hat{H}_i)' \Delta \bar{V}_{i0} \Delta \bar{V}'_{i0} (\bar{H}_i - \hat{H}_i) I_m} \\
& = O_p(1) \times O_p(K^{-\alpha}).
\end{aligned}$$

Combining all of the above results, we obtain

$$\frac{\hat{\Gamma}'_N (\hat{\Lambda}_N - \Lambda_N) \hat{\Gamma}_N}{N} = O_p \left(\zeta_0(K) \left[\sqrt{\frac{K}{N}} + K^{-\alpha} \right] \right) = o_p(1).$$

Recall that $\bar{B} = \bar{B}^* + \bar{D}_c \bar{e} \bar{e}'$. Applying Assumption 1, equation (12), and equation (15), we obtain

$$\begin{aligned}
& \frac{\hat{\Gamma}'_N \Lambda_N \hat{\Gamma}_N}{N} - \frac{\Gamma'_N \Lambda_N \Gamma_N}{N} \\
& = \frac{\left[\Delta \bar{B}' (I - \bar{S}) \bar{H} (\bar{H}' \bar{H})^{-1} \sum_{i=1}^N (\bar{H}_i - \hat{H}_i)' (\Delta \bar{\varepsilon}_i \Delta \bar{\varepsilon}'_i) * (\bar{H}_i - \hat{H}_i) (\bar{H}' \bar{H})^{-1} \bar{H}' (I - \bar{S}) \Delta \bar{D}_c \Delta \bar{\varepsilon} e' \right]}{N} \\
& + \frac{\left[e \Delta \bar{\varepsilon}' \Delta \bar{D}'_c (I - \bar{S}) \bar{H} (\bar{H}' \bar{H})^{-1} \sum_{i=1}^N (\bar{H}_i - \hat{H}_i)' (\Delta \bar{\varepsilon}_i \Delta \bar{\varepsilon}'_i) * (\bar{H}_i - \hat{H}_i) (\bar{H}' \bar{H})^{-1} \bar{H}' (I - \bar{S}) \Delta \bar{B} \right]}{N} \\
& + \frac{\left[e \Delta \bar{\varepsilon}' \Delta \bar{D}'_c (I - \bar{S}) \bar{H} (\bar{H}' \bar{H})^{-1} \sum_{i=1}^N (\bar{H}_i - \hat{H}_i)' (\Delta \bar{\varepsilon}_i \Delta \bar{\varepsilon}'_i) * (\bar{H}_i - \hat{H}_i) (\bar{H}' \bar{H})^{-1} \bar{H}' (I - \bar{S}) \Delta \bar{D}_c \Delta \bar{\varepsilon} e' \right]}{N} \\
& \leq \frac{\lambda_{\max} \times \Delta \bar{B}' (I - \bar{S}) \bar{M} (I - \bar{S}) \bar{D}_c \Delta \bar{\varepsilon} e'}{N} \\
& + \frac{\lambda_{\max} \times e \Delta \bar{\varepsilon}' \bar{D}'_c (I - \bar{S}) \bar{M} (I - \bar{S}) \Delta \bar{B}}{N} \\
& + \frac{\lambda_{\max} \times e \Delta \bar{\varepsilon}' \bar{D}'_c (I - \bar{S}) \bar{M} (I - \bar{S}) \bar{D}_c \Delta \bar{\varepsilon} e'}{N} \\
& = O_p(N^{-1/2}) + O_p(N^{-1/2}) + O_p(N^{-1}) \\
& = O_p(N^{-1/2}) = o_p(1).
\end{aligned}$$

This implies

$$\frac{\hat{\Gamma}'_N \hat{\Lambda}_N \hat{\Gamma}_N}{N} - \frac{\Gamma'_N \Lambda_N \Gamma_N}{N} = \frac{\hat{\Gamma}'_N (\hat{\Lambda}_N - \Lambda_N) \hat{\Gamma}_N}{N} + \frac{\hat{\Gamma}'_N \Lambda_N \hat{\Gamma}_N}{N} - \frac{\Gamma'_N \Lambda_N \Gamma_N}{N} = o_p(1)$$

hence the theorem.

Proof of Theorem 4

Write

$$\Delta \hat{\varepsilon}_i \Delta \hat{\varepsilon}_i' - \Delta \bar{\varepsilon}_i \Delta \bar{\varepsilon}_i' = (\Delta \hat{\varepsilon}_i - \Delta \bar{\varepsilon}_i)(\Delta \hat{\varepsilon}_i - \Delta \bar{\varepsilon}_i)' + \Delta \bar{\varepsilon}_i(\Delta \hat{\varepsilon}_i - \Delta \bar{\varepsilon}_i)' + (\Delta \hat{\varepsilon}_i - \Delta \bar{\varepsilon}_i)\Delta \bar{\varepsilon}_i'.$$

By substitution, we have

$$\begin{aligned} & \frac{\Delta \bar{P}' \Lambda_N \Delta \bar{P}}{N} - \frac{\Delta \bar{P}' \Lambda_N \Delta \bar{P}}{N} \\ &= \frac{\sum_{i=1}^N \Delta \bar{P}_i (\Delta \hat{\varepsilon}_i - \Delta \bar{\varepsilon}_i)(\Delta \hat{\varepsilon}_i - \Delta \bar{\varepsilon}_i)' \Delta \bar{P}_i}{N} \\ & \quad + \frac{\sum_{i=1}^N \Delta \bar{P}_i' \Delta \bar{\varepsilon}_i (\Delta \hat{\varepsilon}_i - \Delta \bar{\varepsilon}_i)' \Delta \bar{P}_i}{N} \\ & \quad + \frac{\sum_{i=1}^N \Delta \bar{P}_i' (\Delta \hat{\varepsilon}_i - \Delta \bar{\varepsilon}_i)' \Delta \bar{\varepsilon}_i' \Delta \bar{P}_i}{N}. \end{aligned}$$

By equation (19) we obtain

$$\begin{aligned} & \frac{\sum_{i=1}^N \Delta \bar{P}_i' (\Delta \hat{\varepsilon}_i - \Delta \bar{\varepsilon}_i)(\Delta \hat{\varepsilon}_i - \Delta \bar{\varepsilon}_i)' \Delta \bar{P}_i}{N} \\ & \leq \frac{2}{N} \sum_{i=1}^N \Delta \bar{P}_i' \Delta \bar{B}_i' (\delta_0 - \hat{\delta})(\delta_0 - \hat{\delta})' \Delta \bar{B}_i' \Delta \bar{P}_i \\ & \quad + \frac{2}{N} \sum_{i=1}^N \Delta \bar{P}_i' \Delta \bar{P}_i (\pi_0 - \hat{\pi})(\pi_0 - \hat{\pi})' \Delta \bar{P}_i' \Delta \bar{P}_i \\ & \quad + \frac{2}{N} \sum_{i=1}^N \Delta \bar{P}_i' \bar{V}_{i0} \bar{V}_{i0}' \Delta \bar{P}_i. \end{aligned}$$

Since the largest eigenvalue of $(\delta_0 - \hat{\delta})(\delta_0 - \hat{\delta})'$ is $O_p(N^{-1})$, we obtain

$$\begin{aligned} & \frac{1}{N} \sum_{i=1}^N \Delta \bar{P}_i' \Delta \bar{B}_i' (\delta_0 - \hat{\delta})(\delta_0 - \hat{\delta})' \Delta \bar{B}_i' \Delta \bar{P}_i \\ & \leq \frac{1}{N} \sum_{i=1}^N \Delta \bar{P}_i' \Delta \bar{B}_i' \Delta \bar{B}_i' \Delta \bar{P}_i \times O_p(N^{-1}) \\ & = O_p(1) \times O_p(N^{-1}) = o_p(1). \end{aligned}$$

Since the largest eigenvalue of $(\pi_0 - \hat{\pi})(\pi_0 - \hat{\pi})'$ is $O_p(N^{-1}K + K^{-\alpha})$, we obtain

$$\begin{aligned} & \frac{1}{N} \sum_{i=1}^N \Delta \bar{P}_i' \Delta \bar{P}_i (\pi_0 - \hat{\pi})(\pi_0 - \hat{\pi})' \Delta \bar{P}_i' \Delta \bar{P}_i \\ & \leq \frac{1}{N} \sum_{i=1}^N \Delta \bar{P}_i' \Delta \bar{P}_i \Delta \bar{P}_i' \Delta \bar{P}_i \times O_p(N^{-1}K + K^{-\alpha}) \\ & = O_p(1) \times O_p(N^{-1}K + K^{-\alpha}) = o_p(1). \end{aligned}$$

Since the largest eigenvalue of $\bar{V}_{i0} \bar{V}_{i0}'$ is $O(K^{-2\alpha})$, we obtain

$$\begin{aligned}
& \frac{1}{N} \sum_{i=1}^N \Delta \bar{P}_i' \bar{V}_{i0} \bar{V}_{i0}' \Delta \bar{P}_i \\
& \leq \frac{1}{N} \sum_{i=1}^N \Delta \bar{P}_i' \Delta \bar{P}_i \times O(K^{-2\alpha}) \\
& = O_p(1) \times O(K^{-2\alpha}) = o_p(1).
\end{aligned}$$

By equation (19) we obtain

$$\begin{aligned}
& \frac{\sum_{i=1}^N \Delta \bar{P}_i' \Delta \bar{\varepsilon}_i (\Delta \hat{\varepsilon}_i - \Delta \bar{\varepsilon}_i)' \Delta \bar{P}_i}{N} \\
& = \frac{1}{N} \sum_{i=1}^N \Delta \bar{P}_i' \Delta \bar{\varepsilon}_i (\delta_0 - \hat{\delta})' \Delta \bar{B}_i' \Delta \bar{P}_i \\
& + \frac{1}{N} \sum_{i=1}^N \Delta \bar{P}_i' \Delta \bar{\varepsilon}_i (\pi_0 - \hat{\pi})' \Delta \bar{P}_i' \Delta \bar{P}_i \\
& + \frac{1}{N} \sum_{i=1}^N \Delta \bar{P}_i' \Delta \bar{\varepsilon}_i \bar{V}_{i0}' \Delta \bar{P}_i.
\end{aligned}$$

Notice that

$$\begin{aligned}
& \left| \frac{1}{N} \sum_{i=1}^N I_s' \Delta \bar{P}_i' \Delta \bar{\varepsilon}_i' (\delta_0 - \hat{\delta})' \Delta \bar{B}_i' \Delta \bar{P}_i I_m \right| \\
& \leq \sqrt{\frac{1}{N} \sum_{i=1}^N I_s' \Delta \bar{P}_i' \Delta \bar{\varepsilon}_i \Delta \bar{\varepsilon}_i' \Delta \bar{P}_i I_s} \sqrt{\frac{1}{N} \sum_{i=1}^N I_m' \Delta \bar{P}_i' \Delta \bar{B}_i (\delta_0 - \hat{\delta}) (\delta_0 - \hat{\delta})' \Delta \bar{B}_i' \Delta \bar{P}_i I_m} \\
& = O_p(1) \times O_p(N^{-1/2}) = o_p(1).
\end{aligned}$$

Employing the same line of proof, we obtain

$$\begin{aligned}
& \left| \frac{1}{N} \sum_{i=1}^N I_s' \Delta \bar{P}_i' \Delta \bar{\varepsilon}_i (\pi_0 - \hat{\pi})' \Delta \bar{P}_i' \Delta \bar{P}_i I_m \right| \\
& \leq \sqrt{\frac{1}{N} \sum_{i=1}^N I_s' \Delta \bar{P}_i' \Delta \bar{\varepsilon}_i \Delta \bar{\varepsilon}_i' \Delta \bar{P}_i I_s} \sqrt{\frac{1}{N} \sum_{i=1}^N I_m' \Delta \bar{P}_i' \Delta \bar{P}_i (\pi_0 - \hat{\pi}) (\pi_0 - \hat{\pi})' \Delta \bar{P}_i' \Delta \bar{P}_i I_m} \\
& = O_p(1) \times O_p(\sqrt{K}/\sqrt{N} + K^{-\alpha}) = o_p(1).
\end{aligned}$$

and

$$\begin{aligned}
& \left| \frac{1}{N} \sum_{i=1}^N I_s' \Delta \bar{P}_i' \Delta \bar{\varepsilon}_i \bar{V}_{i0}' \Delta \bar{P}_i I_m \right| \\
& \leq \sqrt{\frac{1}{N} \sum_{i=1}^N I_s' \Delta \bar{P}_i' \Delta \bar{\varepsilon}_i \Delta \bar{\varepsilon}_i' \Delta \bar{P}_i I_s} \sqrt{\frac{1}{N} \sum_{i=1}^N I_m' \Delta \bar{P}_i' \bar{V}_{i0} \bar{V}_{i0}' \Delta \bar{P}_i I_m} \\
& = O_p(1) \times O_p(K^{-\alpha}) = o_p(1).
\end{aligned}$$

Combining all results, we obtain

$$\frac{\Delta \bar{P}' \hat{\Lambda}_N \Delta \bar{P} - \Delta \bar{P}' \Lambda_N \Delta \bar{P}}{N} = O_p(\sqrt{K/N} + K^{-\alpha}) = o_p(1).$$

Now we can show

$$\begin{aligned} &= p^K(z_0)' \left\{ \frac{\Delta \bar{P}' \Delta \bar{P}}{N} \right\}^{-1} \left\{ \frac{\Delta \bar{P}' \hat{\Lambda}_N \Delta \bar{P} - \Delta \bar{P}' \Lambda_N \Delta \bar{P}}{N} \right\} \left\{ \frac{\Delta \bar{P}' \Delta \bar{P}}{N} \right\}^{-1} p^K(z_0) \\ &= O_p(1) O_p(\sqrt{K/N} + K^{-\alpha}) O_p(1) O_p(\zeta_0(K)^2) \\ &= O_p(\zeta_0(K)^2 [\sqrt{K/N} + K^{-\alpha}]). \end{aligned}$$

Thus,

$$\frac{N(\hat{\sigma}_N^2(z_0) - \sigma_N^2(z_0))}{\|p^K(z_0)\|^2} = O_p(\sqrt{K/N} + K^{-\alpha}) = o_p(1).$$

Applying Theorem 3, we obtain

$$N(\hat{\Omega}_N - \Omega_N) = o_p(1)$$

and

$$\hat{\Omega}_N = \Omega_N + o_p(1/N).$$

Thus,

$$\hat{\Omega}_N^{-\frac{1}{2}} = \Omega_N^{-\frac{1}{2}} + o_p\left(\frac{1}{\sqrt{N}}\right).$$

$$\begin{aligned} &\hat{\Omega}_N^{-1/2} (\hat{\delta} - \delta_0) \\ &= \Omega_N^{-\frac{1}{2}} (\hat{\delta} - \delta_0) + o_p\left(\frac{1}{\sqrt{N}}\right) (\hat{\delta} - \delta_0) \\ &= \Omega_N^{-\frac{1}{2}} (\hat{\delta} - \delta_0) + o_p(1) \\ &\rightarrow N(0, I) \text{ in distribution.} \end{aligned}$$

Applying theorem 4, we have

$$\frac{N(\hat{\sigma}_N^2(z_0) - \sigma_N^2(z_0))}{\|p^K(z_0)\|^2} = O_p\left(\sqrt{K/N} + K^{-\alpha}\right).$$

Thus

$$\hat{\sigma}_N^2(z_0) = \sigma_N^2(z_0) + O_p\left(\frac{\zeta_0(K)^2}{N} [\sqrt{K/N} + K^{-\alpha}]\right)$$

and

$$\begin{aligned}\hat{\sigma}_N(z_0)^{-1}(\hat{g}(z_0) - g(z_0)) &= \sigma_N(z_0)^{-1}(\hat{g}(z_0) - g(z_0)) + O_p\left(\frac{\zeta_0(K)^2}{\sqrt{N}} [\sqrt{K}/\sqrt{N} + K^{-\alpha}]^{\frac{3}{2}}\right) \\ &= \sigma_N(z_0)^{-1}(\hat{g}(z_0) - g(z_0)) + o_p(1) \\ &\rightarrow N(0, 1) \text{ in distribution.}\end{aligned}$$

Figure 1.

The plots of the estimates of $g_0(z) = e^{3z^2} \sin(3\pi z)$ under model (6) with $N=200$ and $T=7$. The black dashed line is the true function. The dotted red line is the average estimate of 1000 simulations. The grey lines are the estimated $g(z)$ in 20 simulations.

Table 1

Finite Sample Performance

$\gamma_0 = -0.5, \lambda_0 = 0, \beta_0 = 0, \gamma_0 = 0, \lambda_0 = 0, \beta_0 = -2$					
	$N = 200$	$N = 500$	$N = 200$	$N = 500$	$N = 500$
γ_0	BIAS	-0.0002	-0.0002	-0.0004	-0.0002
	RMSE	0.0043	0.0026	0.0052	0.0034
	SEE	0.0043	0.0026	0.0052	0.0033
	ESE	0.0044	0.0028	0.0057	0.0037
λ_0	BIAS	0.0056	0.0023	0.0113	0.0091
	RMSE	0.0214	0.0190	0.0201	0.0193
	SEE	0.0206	0.0188	0.0166	0.0170
	ESE	0.0235	0.0182	0.0206	0.0209
β_0	BIAS	0.0069	-0.0011	0.0119	0.0044
	RMSE	0.0467	0.0342	0.0475	0.0357
	SEE	0.0462	0.0342	0.0459	0.0355
	ESE	0.0488	0.0317	0.0504	0.0328
$\hat{g}(z)$	ISE	0.0415	0.0240	0.0415	0.0241

Table 2

Finite Sample Performance

$\gamma_0 = -0.6, \lambda_0 = 0.7, \beta_0 = 0.5, \gamma_0 = -0.5, \lambda_0 = 0.5, \beta_0 = 0.5$					
	$N = 200$	$N = 500$	$N = 200$	$N = 500$	$N = 500$
γ_0	BIAS	0.0003	-4.39e-05	-0.0030	-0.0011
	RMSE	0.0044	0.0029	0.0092	0.0060
	SEE	0.0044	0.0029	0.0087	0.0059
	ESE	0.0046	0.0030	0.0091	0.0063
λ_0	BIAS	0.0001	-1.68e-05	0.0236	0.0167
	RMSE	0.0022	0.0015	0.0569	0.0524
	SEE	0.0021	0.0015	0.0518	0.0496
	ESE	0.0023	0.0015	0.0505	0.0458
β_0	BIAS	0.0014	-0.0014	-0.0205	-0.0126
	RMSE	0.0482	0.0345	0.0700	0.0518
	SEE	0.0482	0.0345	0.0669	0.0502
	ESE	0.0486	0.0318	0.0705	0.0506
$\hat{g}(z)$	ISE	0.0411	0.0238	0.0413	0.0241