Spatial Dynamic Panel Data Models with Interactive Fixed Effects
Supplementary Materials

Wei Shi

Institute for Economic and Social Research, Jinan University, Guangzhou, Guangdong 510632, China.

shiweiemail@gmail.com

Lung-fei Lee

Department of Economics, The Ohio State University, Columbus, Ohio 43210, United States.

lee.1777@osu.edu

This version: 12/12/2016.

S1. Some General Perturbation Theory

The perturbed operator for our problem is

$$\begin{aligned}T(\xi) =& T^{(0)} + \sum_{k_1=0}^{K+2}\xi_{k_1}T_{k_1}^{(1)} + \sum_{k_1=0}^{K+2}\sum_{k_2=0}^{K+2}\xi_{k_1}\xi_{k_2}T_{k_1k_2}^{(2)} \\ &+ \sum_{k_1=0}^{K+2}\sum_{k_2=0}^{K+2}\sum_{k_3=0}^{K+2}\xi_{k_1}\xi_{k_2}\xi_{k_3}T_{k_1k_2k_3}^{(3)} + \sum_{k_1=0}^{K+2}\sum_{k_2=0}^{K+2}\sum_{k_3=0}^{K+2}\sum_{k_4=0}^{K+2}\xi_{k_1}\xi_{k_2}\xi_{k_3}\xi_{k_4}T_{k_1k_2k_3k_4}^{(4)},\end{aligned}$$

where the unperturbed operator is $T(0) = T^{(0)} = \tilde{\Gamma}_n F_T' F_T \tilde{\Gamma}_n'$, and the other operators $T_{k_1}^{(1)}, \cdots, T_{k_1k_2k_3k_4}^{(4)}$ are defined as those in Eq. (C.2). The objective function (Eq. (5)) involves the sum of $n-r_0$ smallest eigenvalues of $\frac{1}{nT}T(\xi)$. The problem here is to relate the sum of the $n-r_0$ smallest eigenvalues of $T(\xi)$ to $T^{(0)}, \cdots, T^{(4)}$'s and ξ. When $\xi = 0$, $T(0) = T^{(0)}$ of which the sum of the $n-r_0$ smallest eigenvalues is 0. The general perturbation theory is applicable to any finite order of terms or even infinite series for $T(\xi)$ in ξ.

The eigenvalues of $T(\xi)$ satisfy the equation $\det(T(\xi) - \lambda I_n) = 0$. $T(\xi)$ is holomorphic for $\xi \in C$, the complex field. The roots can be viewed as branches of analytic functions of ξ with only algebraic singularities (Kato (1995), p.64). Supposing that there are R such distinct functions. $R \leq n$ and $T(\xi)$ is called permanently degenerate if $R < n$. The eigenvalues are denoted as functions $\lambda_1(\xi), \cdots, \lambda_R(\xi)$, which are holomorphic in ξ in a simply connected subdomain of C that contains no exceptional point, and they are continuous at exceptional points. An exceptional point is where some of the otherwise distinct eigenvalue functions have the same value. When $\xi = 0$, $T(\xi) = T^{(0)}$ has at most $r_0 + 1$ distinct eigenvalues, and the zero eigenvalue has multiplicity of $n - r_0$. If $T^{(0)}$ has r_0 positive eigenvalues and they are sufficiently separated from 0, then for ξ small enough, there exists a circle in the complex plane such that it contains just the $n - r_0$ eigenvalues of $T(\xi)$ that are 0 when $\varepsilon = 0$, but no other, and the set of these eigenvalues is denoted as the $\lambda-$group. They all originate from the zero eigenvalue of $T^{(0)}$ in the sense that $\lambda_s(0) = 0$ for $\lambda_s(\cdot) \in \lambda-$ group and they are close to 0 for small ξ. Kato (1995) provides a general formula for $L_{nT}(\theta) = \frac{1}{nT}\sum_{\lambda_s(\cdot)\in\lambda-\text{group}}\lambda_s(\xi)$.

Firstly consider the case where there is no perturbation, $\xi_k = 0$ for $k = 0, \cdots, K+2$. As $T^{(0)}$ is symmetric, and by Theorem 6.38 of Kato (1995), $T^{(0)}$ is diagonalizable and its eigenprojections are orthogonal projections. $T^{(0)}$ can be represented as $T^{(0)} = \sum_h \lambda_h P_h$, with $\sum_h P_h = I_n$, $P_h P_k = \delta_{hk} P_h$, $P_h T^{(0)} = T^{(0)} P_h = \lambda_h P_h$ and $\dim P_h = m_h$, where λ_h 's are distinct eigenvalues, $\delta_{hk} = 1$ if and only if $h = k$, and m_h is the algebraic (and geometric) multiplicity of λ_h. Of particular interest is λ_0 which is the zero eigenvalue. There are several useful properties. The eigenprojection associated with $\lambda_0 = 0$ is denoted as P_0 with $\dim P_0 = n - r_0$. Because $M_{\tilde{\Gamma}_n} = I_n - \tilde{\Gamma}_n(\tilde{\Gamma}_n'\tilde{\Gamma}_n)^{-1}\tilde{\Gamma}_n'$ satisfies the conditions above for eigenprojections, $P_0 = M_{\tilde{\Gamma}_n}$. The resolvent of $T^{(0)}$ is $R(\zeta) = (T^{(0)} - \zeta I)^{-1}$. There is a close relation between eigenprojection and resolvent. The singular points of $R(\zeta)$ are the eigenvalues of $T^{(0)}$. The resolvent has the Laurent series (Kato (1995), p.38, 5.13) about an eigenvalue λ_h,

$$R(\zeta) = -(\zeta - \lambda_h)^{-1}P_h + \sum_{n=0}^{\infty}(\zeta - \lambda_h)^n S_h^{n+1}, \tag{S.1}$$

where the eigennilpotent of the eigenvalue λ_h of $T^{(0)}$ is 0, because $T^{(0)}$ here is diagonalizable; and $S_h = -\sum_{k\neq h}(\lambda_h - \lambda_k)^{-1}P_k$ (Kato (1995), p.40, 5.32). The operator P_h is related to the resolvent as $P_h = -\frac{1}{2\pi i}\int_{\Xi_h} R(\zeta)d\zeta$, where Ξ_h is a positively-oriented small circle enclosing λ_h but excluding other eigenvalues of $T^{(0)}$. The operator $S_h(\zeta) = \sum_{n=0}^{\infty}(\zeta - \lambda_h)^n S_h^{n+1}$ is called the reduced resolvent of $T^{(0)}$ with respect to the eigenvalue λ_h. $S_h(\lambda_h) = S_h$. The reduced resolvent about λ_0 is $S_0(\lambda_0) = S_0$, which satisfies the properties $S_0 = S_0(\lambda_0) = \sum_{k\neq 0}\frac{1}{\lambda_k}P_k$ and $T^{(0)}S_0 = S_0 T^{(0)} = I - P_0 = \tilde{\Gamma}_n(\tilde{\Gamma}_n'\tilde{\Gamma}_n)^{-1}\tilde{\Gamma}_n'$. As a result, $T^{(0)}S_0T^{(0)} = (I - P_0)T^{(0)} = T^{(0)}$, $S_0T^{(0)}S_0 = S_0(I - P_0) = \sum_{k\neq 0}\frac{1}{\lambda_k}P_k(I - P_0) = \sum_{k\neq 0}\frac{1}{\lambda_k}P_k = S_0$, $(T^{(0)}S_0)' = T^{(0)}S_0$, and $(S_0T^{(0)})' = S_0T^{(0)}$. These conditions are the characterizations of the Moore-Penrose pseudo inverse of $T^{(0)}$. S_0 is unique since the Moore-Penrose pseudo inverse is unique. Therefore, $S_0 = S_0(\lambda_0) = \tilde{\Gamma}_n(\tilde{\Gamma}_n'\tilde{\Gamma}_n)^{-1}(F_T'F_T)^{-1}(\tilde{\Gamma}_n'\tilde{\Gamma}_n)^{-1}\tilde{\Gamma}_n'$.

The resolvent of $T(\xi)$ is $R(\zeta,\xi) = (T(\xi) - \zeta I)^{-1}$. Expanding $R(\zeta,\xi)$ as a power series in ξ with coefficients depending on ζ,

$$R(\zeta,\xi)=(T(\xi)-\zeta I)^{-1}=R(\zeta)\left(I+(T(\xi)-T^{(0)})R(\zeta)\right)^{-1}=R(\zeta)\sum_{p=0}^{\infty}(-1)^p\left((T(\xi)-T^{(0)})R(\zeta)\right)^p \tag{S.2}$$

$$=R(\zeta)+\sum_{g=1}^{\infty}\sum_{k_1=0}^{K+2}\cdots\sum_{k_g=0}^{K+2}\xi_{k_1}\xi_{k_2}\cdots\xi_{k_g}R^{(g)}_{k_1\cdots k_g}(\zeta), \tag{S.3}$$

with $R^{(g)}_{k_1\cdots k_g}(\zeta)=\sum_{p=\lceil\frac{g}{4}\rceil}^{g}(-1)^p\sum_{\substack{v_1+\cdots+v_p=g\\ v_j=1,\cdots,4}}R(\zeta)T^{(v_1)}_{k_1\cdots}R(\zeta)\cdots R(\zeta)T^{(v_p)}_{\cdots k_g}R(\zeta)$. Eq. (S.2) is a Neumann series which is absolutely convergent if $\left\|(T(\xi)-T^{(0)})R(\zeta)\right\|_2<1$ which can be true for small enough ξ, and ζ that is in the compact set not including any of the eigenvalues of $T^{(0)}$. Because $\|R(\zeta)\|_2=\frac{1}{\min_k|\zeta-\lambda_k|}$ (Kato (1995), p.60), if ζ is from $\Xi_0=\left\{\zeta|\,|\zeta-\lambda_0|=\frac{nT}{2}d^2_{\min}(\tilde{\Gamma}_n,F_T)\right\}$ which is a positively oriented circle in the complex plane enclosing $\lambda_0=0$ but no other positive eigenvalues of $T^{(0)}$, then $\|R(\zeta)\|_2\leq\frac{1}{\frac{nT}{2}d^2_{\min}(\tilde{\Gamma}_n,F_T)}$ for $\zeta\in\Xi_0$. Therefore a sufficient condition for the series expansion in Eq. (S.2) to converge is $\left\|T(\xi)-T^{(0)}\right\|_2<\frac{nT}{2}d^2_{\min}(\tilde{\Gamma}_n,F_T)$. Substituting in b_{nT} from Eq. (C.3), $\left\|T(\xi)-T^{(0)}\right\|_2\leq nTd^2_{\max}(\tilde{\Gamma}_n,F_T)\left(b_{nT}+b^2_{nT}+b^3_{nT}+b^4_{nT}\right)<4b_{nT}nTd^2_{\max}(\tilde{\Gamma}_n,F_T)$, because $b_{nT}<1$ by the assumption in Lemma 7. The condition becomes $b_{nT}<\frac{d^2_{\min}(\tilde{\Gamma}_n,F_T)}{8d^2_{\max}(\tilde{\Gamma}_n,F_T)}$. Notice that the assumption on b_{nT} in Lemma 7 is stricter than this, which is needed to derive a bound on the remainder term in the truncation of the series expansion.

Let $P(\xi)=-\frac{1}{2\pi i}\int_{\Xi_0}R(\zeta,\xi)d\zeta$. By (5.25) of Kato (1995) p.40, $P(\xi)$ equals to the sum of the eigenprojections of $T(\xi)$ corresponding to the eigenvalues that are inside Ξ_0. Substituting Eq. (S.3) into $P(\xi)$,

$$P(\xi)=P_0+\sum_{g=1}^{\infty}\sum_{k_1=0}^{K+2}\cdots\sum_{k_g=0}^{K+2}\xi_{k_1}\xi_{k_2}\cdots\xi_{k_g}P^{(g)}_{k_1\cdots k_g}, \tag{S.4}$$

where $P^{(g)}_{k_1\cdots k_g}=-\frac{1}{2\pi i}\int_{\Xi_0}R^{(g)}_{k_1\cdots k_g}(\zeta)d\zeta$. Since Eq. (S.3) is convergent for ξ small enough and $\zeta\in\Xi_0$, $P(\xi)$ is a projection depending continuously on ξ. Lemma 4.10 of Kato (1995) p.34 says that $\dim P(\xi)$ is constant, where $\dim P(\xi)=\dim M(\xi)$ and $M(\xi)=\{P(\xi)u,\,u\in\mathbb{R}^n\}$, which implies that $\dim P(\xi)=\dim P_0=n-r_0$. $P(\xi)$ is the total projection of eigenvalues in the $\lambda-$group, i.e., $P(\xi)=\sum_{\lambda_s(\cdot)\in\lambda-\text{group}}P_s(\xi)$, which means that $P(\xi)$ is the projection of the $n-r_0$ smallest eigenvalues of $\frac{1}{nT}\left(T(\xi)-T^{(0)}\right)$ and we can write $P(\xi)=\hat{B}\hat{B}'$, where columns of the $n\times(n-r_0)$ matrix $\hat{B}$ are the eigenvectors corresponding to the $n-r_0$ smallest eigenvalues of $\frac{1}{nT}\left(T(\xi)-T^{(0)}\right)$. Note that $T(\xi)P(\xi)=\sum_{\lambda_s(\cdot)\in\lambda-\text{group}}\lambda_s(\xi)P_s(\xi)$, and $\text{tr}\left(T(\xi)P(\xi)\right)=\sum_{\lambda_s(\cdot)\in\lambda-\text{group}}m_s\lambda_s(\xi)$, where m_s is the multiplicity of $\lambda_s(\xi)$.

$$T(\xi)P(\xi)=-\frac{1}{2\pi i}\int_{\Xi_0}T(\xi)R(\zeta,\xi)d\zeta=-\frac{1}{2\pi i}\int_{\Xi_0}(1+\zeta R(\zeta,\xi))\,d\zeta=-\frac{1}{2\pi i}\int_{\Xi_0}\zeta R(\zeta,\xi)d\zeta$$

$$= -\frac{1}{2\pi i}\int_{\Xi_0}\zeta\left(R(\zeta)+\sum_{g=1}^{\infty}\sum_{k_1=0}^{K+2}\cdots\sum_{k_g=0}^{K+2}\xi_{k_1}\xi_{k_2}\cdots\xi_{k_g}R^{(g)}_{k_1\cdots k_g}(\zeta)\right)d\zeta \qquad \text{from (S.3).} \tag{S.5}$$

From Eq. (S.1), the Laurent series of $R(\zeta)$ about $\lambda_0 = 0$ is $R(\zeta) = \zeta^{-1}S^{(0)} + \sum_{n=0}^{\infty}\zeta^n S^{(n+1)} = \sum_{n=0}^{\infty}\zeta^{n-1}S^{(n)}$, where for the ease of notation, $S^{(0)} = -P_0$, $S^{(n)} = S_0^n$ for $n \geq 1$, as in Lemma 7. Substituting it into the integrand in Eq. (S.5), and using the fact that only the term with ζ^{-1} contributes to the integral, $T(\xi)P(\xi) = \sum_{g=1}^{\infty}\sum_{k_1=0}^{K+2}\sum_{k_2=0}^{K+2}\cdots\sum_{k_g=0}^{K+2}\xi_{k_1}\xi_{k_2}\cdots\xi_{k_g}\tilde{T}^{(g)}_{k_1\cdots k_g}$, where

$$\tilde{T}^{(g)}_{k_1\cdots k_g} = -\sum_{p=\lceil \frac{g}{4}\rceil}^{g}(-1)^p \sum_{\substack{v_1+v_2+\cdots+v_p=g\\ m_1+\cdots+m_{p+1}=p-1\\ v_j=1,\cdots,4,\, m_j=0,1,\cdots}} S^{(m_1)}T^{(v_1)}_{k_1\cdots}S^{(m_2)}\cdots S^{(m_p)}T^{(v_p)}_{\cdots k_g}S^{(m_{p+1})}.$$

$L_{nT}(\xi) = \sum_{g=1}^{\infty}\sum_{k_1=0}^{K+2}\sum_{k_2=0}^{K+2}\cdots\sum_{k_g=0}^{K+2}\xi_{k_1}\xi_{k_2}\cdots\xi_{k_g}\frac{1}{nT}\mathrm{tr}\left(\tilde{T}^{(g)}_{k_1\cdots k_g}\right)$, because the sum of the $n-r_0$ perturbed eigenvalues is $\mathrm{tr}\left(T(\xi)P(\xi)\right)$. The following lemma provides a bound on remainder terms in the series expansion.

Lemma 12. *Under Assumption SF,* $\|S_0\|_2 \leq \frac{1}{nTd^2_{\min}(\tilde{\Gamma}_n,F_T)}$, *and for* $g \geq 5$,

$$\frac{1}{nT}\sum_{k_1,\cdots,k_g=0}^{K+2}\left|\xi_{k_1}\cdots\xi_{k_g}tr\left(\tilde{T}^{(g)}_{k_1k_2\cdots k_g}\right)\right| \leq \frac{16r_0d^2_{\max}(\tilde{\Gamma}_n,F_T)d^2_{\min}(\tilde{\Gamma}_n,F_T)}{16d^2_{\max}(\tilde{\Gamma}_n,F_T)-d^2_{\min}(\tilde{\Gamma}_n,F_T)}\left(\frac{16d^2_{\max}(\tilde{\Gamma}_n,F_T)b_{nT}}{d^2_{\min}(\tilde{\Gamma}_n,F_T)}\right)^g.$$

Proof. Firstly we show that $\|S_0\|_2 \leq \frac{1}{nTd^2_{\min}(\tilde{\Gamma}_n,F_T)}$.

$$\begin{aligned}\|S_0\|_2^2 =& \mu_1\left(\left(F_T'F_T\right)^{-1}\left(\tilde{\Gamma}_n'\tilde{\Gamma}_n\right)^{-1}\left(F_T'F_T\right)^{-1}\left(\tilde{\Gamma}_n'\tilde{\Gamma}_n\right)^{-1}\right) \leq \left[\mu_1\left(\left(F_T'F_T\right)^{-1}\left(\tilde{\Gamma}_n'\tilde{\Gamma}_n\right)^{-1}\right)\right]^2 \\ =& \frac{1}{\left[\mu_{r_0}\left(\tilde{\Gamma}_n'\tilde{\Gamma}_nF_T'F_T\right)\right]^2} = \frac{1}{\left[\mu_{r_0}\left(\tilde{\Gamma}_nF_T'F_T\tilde{\Gamma}_n'\right)\right]^2} = \left(\frac{1}{nTd^2_{\min}(\tilde{\Gamma}_n,F_T)}\right)^2.\end{aligned} \tag{S.6}$$

Under Assumption SF, $\left(\tilde{\Gamma}_n'\tilde{\Gamma}_n\right)^{-\frac{1}{2}}\left(F_T'F_T\right)^{-1}\left(\tilde{\Gamma}_n'\tilde{\Gamma}_n\right)^{-\frac{1}{2}}$ is positive definite, and the inequality in Eq. (S.6) follows from Lemma 2(4). Because $v_j = 1,\cdots,4$ in $T^{(v_j)}_{k_1\cdots k_{v_j}}$ and $v_1+\cdots+v_p = g$, for $g \geq 5$, p is at least 2. From $m_1+\cdots+m_{p+1} = p-1$, $\tilde{T}^{(g)}_{k_1\cdots k_g}$ has at least one $S^{(m)}$ with $m \geq 1$ if $g \geq 5$. The rank of $S^{(m)}$, $m \geq 1$, is at most r_0 and $\left\|S^{(m)}\right\|_2 \leq \|S_0\|_2^m \leq \left(\frac{1}{nTd^2_{\min}(\tilde{\Gamma}_n,F_T)}\right)^m$. For $g \geq 5$,

$$\begin{aligned}&\frac{1}{nT}\sum_{k_1,\cdots,k_g=0}^{K+2}\left|\xi_{k_1}\cdots\xi_{k_g}\mathrm{tr}\left(\tilde{T}^{(g)}_{k_1k_2\cdots k_g}\right)\right| \\ \leq& r_0\frac{1}{nT}\sum_{k_1,\cdots,k_g=0}^{K+2}\sum_{p=\lceil\frac{g}{4}\rceil}^{g}\|S_0\|_2^{p-1}\sum_{\substack{m_1+\cdots+m_{p+1}=p-1\\ m_j=0,1,\cdots}}\left(\sum_{\substack{v_1+\cdots+v_p=g\\ v_j=1,\cdots,4}}|\xi_{k_1\cdots}|\left\|T^{(v_1)}_{k_1\cdots}\right\|_2\cdots|\xi_{\cdots k_g}|\left\|T^{(v_p)}_{\cdots k_g}\right\|_2\right)\end{aligned}$$

$$\leq r_0 \frac{1}{nT} \sum_{k_1,\cdots,k_g=0}^{K+2} \sum_{p=1}^{g} \|S_0\|_2^{p-1} 4^p \left(\sum_{\substack{v_1+\cdots+v_p=g \\ v_j=1,\cdots,4}} |\xi_{k_1\cdots}| \left\| T_{k_1\cdots}^{(v_1)} \right\|_2 \cdots |\xi_{\cdots k_g}| \left\| T_{\cdots k_g}^{(v_p)} \right\|_2 \right) \tag{S.7}$$

$$\leq r_0 d_{\min}^2(\tilde{\Gamma}_n, F_T) \sum_{p=1}^{g} \left(\frac{4}{d_{\min}^2(\tilde{\Gamma}_n, F_T)} \right)^p \left(\sum_{\substack{v_1+\cdots+v_p=g \\ v_j=1,\cdots,4}} b_{nT}^g d_{\max}^{2p}(\tilde{\Gamma}_n, F_T) \right)$$

$$\leq r_0 d_{\min}^2(\tilde{\Gamma}_n, F_T) b_{nT}^g \sum_{p=1}^{g} \left(\frac{16 d_{\max}^2(\tilde{\Gamma}_n, F_T)}{d_{\min}^2(\tilde{\Gamma}_n, F_T)} \right)^p = r_0 d_{\min}^2(\tilde{\Gamma}_n, F_T) b_{nT}^g \frac{16 d_{\max}^2(\tilde{\Gamma}_n, F_T)}{d_{\min}^2(\tilde{\Gamma}_n, F_T)} \frac{\left(\frac{16 d_{\max}^2(\tilde{\Gamma}_n, F_T)}{d_{\min}^2(\tilde{\Gamma}_n, F_T)} \right)^g - 1}{\frac{16 d_{\max}^2(\tilde{\Gamma}_n, F_T)}{d_{\min}^2(\tilde{\Gamma}_n, F_T)} - 1} \tag{S.8}$$

$$< \frac{16 r_0 d_{\max}^2(\tilde{\Gamma}_n, F_T) d_{\min}^2(\tilde{\Gamma}_n, F_T)}{16 d_{\max}^2(\tilde{\Gamma}_n, F_T) - d_{\min}^2(\tilde{\Gamma}_n, F_T)} \left(\frac{16 d_{\max}^2(\tilde{\Gamma}_n, F_T) b_{nT}}{d_{\min}^2(\tilde{\Gamma}_n, F_T)} \right)^g.$$

In Eq, (S.7), we use the inequality,[14] $\sum_{\substack{m_1+\cdots+m_{p+1}=p-1 \\ m_j=0,1,\cdots}} 1 \leq \sum_{\substack{m_1+\cdots+m_{p+1}=p \\ m_j=0,1,\cdots}} 1 = \frac{(2p)!}{(p!)^2} \leq 4^p$. Eq. (S.8) is because $\sum_{\substack{v_1+\cdots+v_p=g \\ v_j=1,\cdots,4}} 1 \leq 4^p$. □

S2. Proof of Lemmas

Proof of Lemma 3.

Firstly for $k = 1, \cdots, k$, because $\mathbb{E}\|X_k\|_2^2 \leq \mathbb{E}\|X_k\|_F^2 = \sum_{i=1}^n \sum_{t=1}^T \mathbb{E} X_{k,it}^2 = O(nT)$ from Assumption R3, $\|X_k\|_2 = O_P(\sqrt{nT})$ follows from the generalized Chebyshev inequality. By recursive substitution, $Y_{-1} = \sum_{h=0}^{\infty} A_n^h S_n^{-1} \left(\sum_{k=1}^{K-2} X_{k,-h} \beta_k + \Gamma_n F'_{-h} + U_{-h} \right)$ where $X_{k,-h} = (X_{n1-h,k}, \cdots, X_{nT-h,k})$, $F_{-h} = (f_{1-h}, \cdots, f_{T-h})'$ and $U_{-h} = (U_{n1-h}, \cdots, U_{nT-h})$. It follows that

$$\|Y_{-1}\|_2 \leq \|Y_{-1}\|_F \leq \sum_{h=0}^{\infty} \|A_n\|_2^h \left\| S_n^{-1} \right\|_2 \left(\sum_{k=1}^{K-2} \|X_{k,-h}\|_F |\beta_k| + \left\| \Gamma_n F'_{-h} \right\|_F + \|U_{-h}\|_F \right).$$

Because by Assumption R, $\|A_n\|_2 < 1$, $\mathbb{E}\left\|\Gamma_n F'_{-h}\right\|_F \leq \left(\mathbb{E}\left\|\Gamma_n F'_{-h}\right\|_F^2 \right)^{\frac{1}{2}} = O\left(\sqrt{nT}\right)$ and $\mathbb{E}\|U_{-h}\|_F^2 \leq \left\|R_n^{-1}\right\|_2^2 \mathbb{E}\|\varepsilon_{-h}\|_F^2 = O(nT)$, $\|Y_{-1}\|_2 = O_P\left(\sqrt{nT}\right)$ follows from the Markov inequality. $\|W_n Y_{-1}\|_2 = O_P(\sqrt{nT})$ because $\|W_n Y_{-1}\|_2 \leq \|W_n\|_1^{\frac{1}{2}} \|W_n\|_\infty^{\frac{1}{2}} \|Y_{-1}\|_2 = O_P(\sqrt{nT})$. For $k = K+1$, $Z_{K+1} = \sum_{k=1}^K G_n Z_k \delta_{0k}$. Using these results,

$$\|Z_{K+1}\|_2 = \left\| \sum_{k=1}^K G_n Z_k \delta_{0k} \right\|_2 \leq \sum_{k=1}^K \|G_n\|_2 \|Z_k\|_2 |\delta_{0k}| \leq \sum_{k=1}^K \|G_n\|_\infty^{\frac{1}{2}} \|G_n\|_1^{\frac{1}{2}} \|Z_k\|_2 |\delta_{0k}| = O_P(\sqrt{nT}).$$

---

[14](1) $\sum_{\substack{m_1+\cdots+m_{p+1}=p \\ m_j=0,1,\cdots}} 1 = \frac{(2p)!}{(p!)^2}$. This can be seen by counting the number of combinations of $\dagger, \cdots, \dagger, x, \cdots, x$. For each p, there are p †'s and p x's. A combination of $\dagger, \cdots, \dagger, x, \cdots, x$ corresponds to a set of $m_1, \cdots, m_{p+1}$. The number of $\dagger' s$ between the i-th x (or the beginning) and the $i+1$-th x (or the end) is the value of m_{i+1}. For example, let $p = 2$. Then $\dagger\dagger xx$ corresponds to $(m_1, m_2, m_3) = (2, 0, 0)$ and $\dagger x \dagger x$ corresponds to $(m_1, m_2, m_3) = (1, 1, 0)$. There are $(2p)!$ such arrangements. Because the ordering of † and x does not matter, $\sum_{\substack{m_1+\cdots+m_{p+1}=p \\ m_j=0,1,\cdots}} 1 = \frac{(2p)!}{(p!)^2}$.

(2) $\frac{(2p)!}{(p!)^2} \leq 4^p$. This can be checked by taking log on both sides and using the trapezoidal rule.

In conclusion, $\|Z_k\|_2 = O_P(\sqrt{nT})$ and $\|Z_k\|_F = O_P(\sqrt{nT})$ for $k = 1,2,\cdots,K+1$.

Under Assumptions E and R, we have, for $k = 1,\cdots,K,K+1$, $\mathrm{tr}\left((A_nZ_k)'\varepsilon\right) = O_P\left(\sqrt{nT}\right)$ where A_n is a $n\times n$ nonstochastic UB matrix. This is so because

$$\begin{aligned}\mathbb{E}\mathrm{tr}\left((A_nZ_k)'\varepsilon\right)^2 =&\mathbb{E}\sum_{i,i'=1}^{n}\sum_{t,t'=1}^{T}[A_nZ_k]_{it}[A_nZ_k]_{i't'}\varepsilon_{it}\varepsilon_{i't'} = \sigma_0^2\sum_{i=1}^{n}\sum_{t=1}^{T}\mathbb{E}[A_nZ_k]_{it}^2 = \sigma_0^2\mathbb{E}\|A_nZ_k\|_F^2\\ \le&\sigma_0^2\|A_n\|_2^2\mathbb{E}\|Z_k\|_F^2 \le \sigma_0^2\mathbb{E}\|Z_k\|_F^2 = O(nT),\end{aligned}$$

where the second equality is due to $\mathbb{E}(\varepsilon_{it}|Z_{k,nt}) = 0$, and $\mathbb{E}(\varepsilon_{it}\varepsilon_{i't}|Z_{k,nt}) = 0$ for $i\neq i'$. Some of the inequalities on matrix norms in Lemma 2 are used. □

Proof of Lemma 4.

Proof. For (1), because of Assumption SF,

$$\left\|P_{\tilde\Gamma_n,F_T}\right\|_2 \le \frac{1}{nT}\|\tilde\Gamma_n\|_2\left\|\left(\frac{1}{n}\tilde\Gamma_n'\tilde\Gamma_n\right)^{-1}\right\|_2\left\|\left(\frac{1}{T}F_T'F_T\right)^{-1}\right\|_2\|F_T'\|_2 = O_P\left(\frac{1}{\sqrt{nT}}\right).$$

The second part of (1) is shown in Lemma S.8.1(d) in Moon and Weidner (2015a). For (2), the explicit expressions of $Z_k-\bar Z_k$ for various k are in Eq.(6). $Z_k-\bar Z_k = 0$ for $k = 3,\cdots,K$. Consider $k = 1$,

$$Z_1-\bar Z_1 = \sum_{h=0}^{\infty}A_n^hS_n^{-1}R_n^{-1}\tilde\varepsilon_h = \sum_{h=0}^{T}A_n^hS_n^{-1}R_n^{-1}\tilde\varepsilon_h + r_{nT},$$

where $r_{nT} = \sum_{h=T+1}^{\infty}A_n^hS_n^{-1}R_n^{-1}\tilde\varepsilon_h$. Let $\varepsilon_{n,2T}$ denote the $n\times 2T$ matrix $(\varepsilon_{n,1-T},\cdots,\varepsilon_{n,0},\varepsilon_{n,1},\cdots,\varepsilon_{n,T})$. Assumption E shows that $\|\varepsilon_{n,2T}\|_2 = O_P\left(\max\left(\sqrt{n},\sqrt{2T}\right)\right)$. Because $\tilde\varepsilon_h$ is a submatrix of $\varepsilon_{n,2T}$, $\|\tilde\varepsilon_h\|_2 \le \|\varepsilon_{n,2T}\|_2 = O_P\left(\max\left(\sqrt{n},\sqrt{2T}\right)\right)$. Therefore $\left\|\sum_{h=0}^{T}A_n^hS_n^{-1}R_n^{-1}\tilde\varepsilon_h\right\|_2 \le \left\|S_n^{-1}\right\|_2\left\|R_n^{-1}\right\|_2\|\varepsilon_{n,2T}\|_2\sum_{h=0}^{T}\|A_n\|_2^h = O_P\left(\max\left(\sqrt{n},\sqrt{2T}\right)\right)$, as $\sum_{h=0}^{T}\|A_n\|_2^h$ is bounded by Assumption R5. It remains to show that the remainder term is $o_P(\sqrt{n})$. As $[r_{nT}]_{it}^2 = \sum_{h=T+1}^{\infty}\sum_{h'=T+1}^{\infty}\sum_{j=1}^{n}\sum_{j'=1}^{n}\left[A_n^hS_n^{-1}R_n^{-1}\right]_{ij}\left[A_n^{h'}S_n^{-1}R_n^{-1}\right]_{ij'}\varepsilon_{j,t-h}\varepsilon_{j',t-h'}$,

$$\begin{aligned}\mathbb{E}\sum_{i=1}^{n}\sum_{t=1}^{T}[r_{nT}]_{it}^2 &= \sum_{i=1}^{n}\sum_{t=1}^{T}\sum_{h=T+1}^{\infty}\sum_{h'=T+1}^{\infty}\sum_{j=1}^{n}\sum_{j'=1}^{n}\left[A_n^hS_n^{-1}R_n^{-1}\right]_{ij}\left[A_n^{h'}S_n^{-1}R_n^{-1}\right]_{ij'}\mathbb{E}\left(\varepsilon_{j,t-h}\varepsilon_{j',t-h'}\right)\\ &= \sigma_0^2T\sum_{h=T+1}^{\infty}\sum_{i=1}^{n}\sum_{j=1}^{n}\left[A_n^hS_n^{-1}R_n^{-1}\right]_{ij}\left[A_n^hS_n^{-1}R_n^{-1}\right]_{ij} = \sigma_0^2T\sum_{h=T+1}^{\infty}\mathrm{tr}\left(A_n^hS_n^{-1}R_n^{-1}R_n^{-1\prime}S_n^{-1\prime}A_n^{h\prime}\right)\\ &\le \sigma_0^2nT\left\|S_n^{-1}\right\|_2^2\left\|R_n^{-1}\right\|_2^2\sum_{h=T+1}^{\infty}\|A_n\|_2^{2h} \le \sigma_0^2n\left\|S_n^{-1}\right\|_2^2\left\|R_n^{-1}\right\|_2^2\sum_{h=T+1}^{\infty}h\|A_n\|_2^{2h}\\ &= o(n),\end{aligned}$$

where $\sum_{h=T+1}^{\infty}h\|A_n\|_2^{2h} = o(1)$ because the series $\sum_{h=1}^{\infty}h\|A_n\|_2^{2h}$ converges by Assumption R5. Therefore $\|r_{nT}\|_2 \le \|r_{nT}\|_F = o_P(\sqrt{n})$. $\|Z_1-\bar Z_1\|_2 = O_P\left(\max\left(\sqrt{n},\sqrt{2T}\right)\right)$ follows. The cases for $k = 2$ and $k =$

$K+1$ are similar.

For the second part of (2),

$$\left\|Z_k\varepsilon' P_{\tilde{\Gamma}_n}\right\|_2 \leq \frac{1}{n}\left(\left\|\bar{Z}_k\varepsilon'\tilde{\Gamma}_n\right\|_2 + \left\|\left(Z_k-\bar{Z}_k\right)\varepsilon'\tilde{\Gamma}_n\right\|_2\right)\left\|\left(\frac{1}{n}\tilde{\Gamma}'_n\tilde{\Gamma}_n\right)^{-1}\right\|_2 \left\|\tilde{\Gamma}_n\right\|_2.$$

Because $\left\|\bar{Z}_k\varepsilon'\tilde{\Gamma}_n\right\|_2^2 \leq \left\|\bar{Z}_k\varepsilon'\tilde{\Gamma}_n\right\|_F^2 = \operatorname{tr}\left(\bar{Z}_k\varepsilon'\tilde{\Gamma}_n\tilde{\Gamma}'_n\varepsilon\bar{Z}'_k\right) = \operatorname{vec}(\varepsilon)'\left(\bar{Z}'_k\bar{Z}_k \otimes \tilde{\Gamma}_n\tilde{\Gamma}'_n\right)\operatorname{vec}(\varepsilon)$, which implies that $\mathbb{E}\left\|\bar{Z}_k\varepsilon'\tilde{\Gamma}_n\right\|_2^2 \leq \sigma_0^2\mathbb{E}\left\|\bar{Z}_k\right\|_F^2\left\|\tilde{\Gamma}_n\right\|_F^2 = O(n^2T)$, therefore $\left\|\bar{Z}_k\varepsilon'\tilde{\Gamma}_n\right\|_2 = O_P(n\sqrt{T})$. As $\left\|\tilde{\Gamma}_n\right\|_2 = O_P(\sqrt{n})$, $\left\|\left(Z_k-\bar{Z}_k\right)\varepsilon'\tilde{\Gamma}_n\right\|_2 \leq \left\|Z_k-\bar{Z}_k\right\|_2\left\|\varepsilon\right\|_2\left\|\tilde{\Gamma}_n\right\|_2 = O_P\left(\max\left(n\sqrt{n}, T\sqrt{n}\right)\right)$. we have $\left\|Z_k\varepsilon' P_{\tilde{\Gamma}_n}\right\|_2 = O_P\left(\max\left(n,T\right)\right)$. □

Proof of Lemma 5.

Proof. The statistics under consideration for $k=3,\cdots,K$ are identically zero because $\bar{Z}_k = Z_k$ for those k's. It remains to consider $k=1,2$, and $K+1$. Starting from $k=1$.

$$\begin{aligned}
\frac{1}{\sqrt{nT}}\operatorname{tr}\left(Q_n\left(Z_1-\bar{Z}_1\right)P_{F_T}\varepsilon'\right) &= \frac{1}{\sqrt{nT}}\operatorname{tr}\left(\sum_{h=1}^{\infty} Q_n A_n^{h-1}S_n^{-1}\tilde{\varepsilon}_h P_{F_T}\varepsilon'\right) \\
&= \frac{1}{\sqrt{nT}}\sum_{q,q'=1}^{r}\left[\left(\frac{1}{T}F'_TF_T\right)^{-1}\right]_{qq'}\frac{1}{T}\underbrace{\sum_{i,j=1}^{n}\sum_{t,t'=1}^{T}\sum_{h=1}^{\infty}\left[Q_nA_n^{h-1}S_n^{-1}\right]_{ij}F_{T,tq}F_{T,t'q'}\varepsilon_{it'}\varepsilon_{j,t-h}}_{M_{qq'}},
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}\left(M_{qq'}|\mathscr{C}_{nT}\right) =& \sigma_0^2\sum_{i=1}^{n}\sum_{t=2}^{T}\sum_{h=1}^{t-1}\left[Q_nA_n^{h-1}S_n^{-1}\right]_{ii}F_{T,tq}F_{T,t-h,q'}, \\
\left(\mathbb{E}\left(M_{qq'}|\mathscr{C}_{nT}\right)\right)^2 =& \sigma_0^4\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{t=2}^{T}\sum_{h=1}^{t-1}\sum_{t'=2}^{T}\sum_{h'=1}^{t'-1}\left[Q_nA_n^{h-1}S_n^{-1}\right]_{ii}\left[Q_nA_n^{h'-1}S_n^{-1}\right]_{jj}F_{T,tq}F_{T,t-h,q'}F_{T,t'q}F_{T,t'-h',q'}, \\
\mathbb{E}\left(M_{qq'}^2|\mathscr{C}_{nT}\right) =& \sum_{i,j,i',j'=1}^{n}\sum_{t,t',s,s'=1}^{T}\sum_{h,h'=1}^{\infty}\left[Q_nA_n^{h-1}S_n^{-1}\right]_{ij}\left[Q_nA_n^{h'-1}S_n^{-1}\right]_{i'j'}F_{T,tq}F_{T,t'q'}F_{T,sq}F_{T,s'q'}\mathbb{E}\left(\varepsilon_{it'}\varepsilon_{j,t-h}\varepsilon_{i's'}\varepsilon_{j',s-h'}\right) \\
=& \mu^{(4)}\sum_{i=1}^{n}\sum_{t=2}^{T}\sum_{h,h'=1}^{t-1}\left[Q_nA_n^{h-1}S_n^{-1}\right]_{ii}\left[Q_nA_n^{h'-1}S_n^{-1}\right]_{ii}F_{T,t+h,q}F_{T,tq'}F_{T,t+h',q}F_{T,tq'} \\
&+ \sigma_0^4\sum_{i=1}^{n}\sum_{t,t'=1}^{T}\sum_{h=1,h\neq t-t'}^{\infty}\left[Q_nA_n^{h-1}S_n^{-1}\right]_{ii}^2F_{T,tq}^2F_{T,t'q'}^2 \\
&+ 2\sigma_0^4\sum_{i=1}^{n}\sum_{t,t'=2}^{T}\sum_{h=1}^{t-1}\sum_{h'=1,h'\neq t'-t+h}^{t'-1}\left[Q_nA_n^{h-1}S_n^{-1}\right]_{ii}\left[Q_nA_n^{h'-1}S_n^{-1}\right]_{ii}F_{T,tq}F_{T,t-h,q'}F_{T,t'q}F_{T,t'-h',q'} \\
&+ \sigma_0^4\sum_{i,j=1,i\neq j}^{n}\sum_{t,t'=2}^{T}\sum_{h=1}^{t-1}\sum_{h'=1}^{t'-1}\left[Q_nA_n^{h-1}S_n^{-1}\right]_{ii}\left[Q_nA_n^{h'-1}S_n^{-1}\right]_{jj}F_{T,tq}F_{T,t-h,q'}F_{T,t'q}F_{T,t'-h',q'} \\
&+ \sigma_0^4\sum_{i,j=1,i\neq j}^{n}\sum_{t,t'=1}^{T}\sum_{h,h'=1,1\leq t-h+h'\leq T}^{\infty}\left[Q_nA_n^{h-1}S_n^{-1}\right]_{ij}\left[Q_nA_n^{h'-1}S_n^{-1}\right]_{ij}F_{T,tq}F_{T,t'q'}F_{T,t-h+h',q}F_{T,t'q'} \\
&+ \sigma_0^4\sum_{i,j=1,i\neq j}^{n}\sum_{t,t'=2}^{T}\sum_{h=1}^{t-1}\sum_{h'=1}^{t'-1}\left[Q_nA_n^{h-1}S_n^{-1}\right]_{ij}\left[Q_nA_n^{h'-1}S_n^{-1}\right]_{ji}F_{T,tq}F_{T,t'-h',q'}F_{T,t'q}F_{T,t-h,q'},
\end{aligned}$$

$$
\begin{aligned}
\operatorname{var}\left(M_{qq'}|\mathscr{C}_{nT}\right) =& \left(\mu^{(4)}-\sigma_0^4\right)\sum_{i=1}^{n}\sum_{t=2}^{T}\sum_{h,h'=1}^{t-1}\left[Q_nA_n^{h-1}S_n^{-1}\right]_{ii}\left[Q_nA_n^{h'-1}S_n^{-1}\right]_{ii}F_{T,t+h,q}F_{T,tq'}F_{T,t+h',q}F_{T,tq'} \\
&+\sigma_0^4\sum_{i=1}^{n}\sum_{t,t'=1}^{T}\sum_{h=1,h\neq t-t'}^{\infty}\left[Q_nA_n^{h-1}S_n^{-1}\right]_{ii}^2F_{T,tq}^2F_{T,t'q'}^2 \\
&+\sigma_0^4\sum_{i=1}^{n}\sum_{t,t'=2}^{T}\sum_{h=1}^{t-1}\sum_{h'=1,h'\neq t'-t+h}^{t'-1}\left[Q_nA_n^{h-1}S_n^{-1}\right]_{ii}\left[Q_nA_n^{h'-1}S_n^{-1}\right]_{ii}F_{T,tq}F_{T,t-h,q'}F_{T,t'q}F_{T,t'-h',q'} \\
&+\sigma_0^4\sum_{i,j=1,i\neq j}^{n}\sum_{t,t'=1}^{T}\sum_{h,h'=1,1\leq t-h+h'\leq T}^{\infty}\left[Q_nA_n^{h-1}S_n^{-1}\right]_{ij}\left[Q_nA_n^{h'-1}S_n^{-1}\right]_{ij}F_{T,tq}F_{T,t'q'}F_{T,t-h+h',q}F_{T,t'q'} \\
&+\sigma_0^4\sum_{i,j=1,i\neq j}^{n}\sum_{t,t'=2}^{T}\sum_{h=1}^{t-1}\sum_{h'=1}^{t'-1}\left[Q_nA_n^{h-1}S_n^{-1}\right]_{ij}\left[Q_nA_n^{h'-1}S_n^{-1}\right]_{ji}F_{T,tq}F_{T,t'-h',q'}F_{T,t'q}F_{T,t-h,q'} \\
\leq& C_1T^2\sum_{i=1}^{n}\sum_{h,h'=1}^{\infty}\operatorname{abs}\left[Q_nA_n^{h-1}S_n^{-1}\right]_{ii}\operatorname{abs}\left[Q_nA_n^{h'-1}S_n^{-1}\right]_{ii}+C_2T^2\sum_{h,h'=1}^{\infty}\left|\operatorname{tr}\left(Q_nA_n^{h-1}S_n^{-1}S_n^{-1\prime}A_n^{h'-1\prime}Q_n'\right)\right| \\
&+C_3T^2\sum_{h,h'=1}^{\infty}\left|\operatorname{tr}\left(Q_nA_n^{h-1}S_n^{-1}Q_nA_n^{h'-1}S_n^{-1}\right)\right| \\
=& O(nT^2),
\end{aligned}
$$

where C_1, C_2 and C_3 are some finite constants. By the generalized Chebyshev's (with second moment) inequality, $M_{qq'}-\mathbb{E}\left(M_{qq'}|\mathscr{C}_{nT}\right)=O_P\left(T\sqrt{n}\right)$. Therefore $\frac{1}{\sqrt{nT}}\left[\operatorname{tr}\left(Q_n\left(Z_1-\bar{Z}_1\right)P_{F_T}\varepsilon'\right)-\mathbb{E}\left(\operatorname{tr}\left(Q_n\left(Z_1-\bar{Z}_1\right)P_{F_T}\varepsilon'\right)|\mathscr{C}_{nT}\right)\right]=O_P(\frac{1}{\sqrt{T}})$. The cases for $k=2$ and $K+1$ are similar. □

Proof of Lemma 6.

We shall prove (1), and (2) and (3) can be proved using the same arguments. As $\operatorname{tr}\left(P_{\tilde{\Gamma}_n}\varepsilon\varepsilon'\right)=\operatorname{vec}(\varepsilon)'\left(I_T\otimes P_{\tilde{\Gamma}_n}\right)\operatorname{vec}(\varepsilon)$,

$$
\begin{aligned}
\operatorname{var}\left(\operatorname{tr}\left(P_{\tilde{\Gamma}_n}\varepsilon\varepsilon'\right)|\mathscr{C}_{nT}\right) =& 2\sigma_0^4\operatorname{tr}\left(I_T\otimes P_{\tilde{\Gamma}_n}\right)+\left(\mu^{(4)}-3\sigma_0^4\right)T\sum_{i=1}^{n}\left[P_{\tilde{\Gamma}_n}\right]_{ii}^2 \\
\leq& 2\sigma_0^4\operatorname{tr}\left(I_T\otimes P_{\tilde{\Gamma}_n}\right)+\left(\mu^{(4)}-3\sigma_0^4\right)T\left\|P_{\tilde{\Gamma}_n}\right\|_F^2 \\
\leq& 2\sigma_0^4\operatorname{tr}\left(I_T\otimes P_{\tilde{\Gamma}_n}\right)+\left(\mu^{(4)}-3\sigma_0^4\right)Tr_0\left\|P_{\tilde{\Gamma}_n}\right\|_2^2=O(T).
\end{aligned}
$$

Therefore (1) follows by Chebyshev's inequality. □

Proof of Lemma 8.

From Lemma 7 in Appendix C and assuming that n and T are proportional,

$$
\begin{aligned}
L_{nT}(\xi) =& \frac{1}{nT}\sum_{g=1}^{4}\sum_{k_1=0}^{K+2}\cdots\sum_{k_g=0}^{K+2}\xi_{k_1}\cdots\xi_{k_g}\frac{1}{nT}\operatorname{tr}\left(\tilde{T}_{k_1\cdots k_g}^{(g)}\right)+O_P\left(\|\theta-\theta_0\|_1^5\right)+O_P\left(\|\theta-\theta_0\|_1^4|\xi_0|\right) \\
&+O_P\left(\|\theta-\theta_0\|_1^3|\xi_0|^2\right)+O_P\left(\|\theta-\theta_0\|_1^2|\xi_0|^3\right)+O_P\left(\|\theta-\theta_0\|_1|\xi_0|^4\right)+O_P\left(|\xi_0|^5\right), \qquad \text{(S.9)}
\end{aligned}
$$

where the last equality is because $|\xi_0|=O_P\left(n^{-\frac{1}{2}}\right)$ and θ is in a small neighborhood of θ_0.

Starting with $g=1$,

$$\sum_{k_1=0}^{K+2}\xi_{k_1}\frac{1}{nT}\operatorname{tr}\left(\tilde{T}_{k_1}^{(1)}\right)=\sum_{k_1=0}^{K+2}\xi_{k_1}\frac{1}{nT}\operatorname{tr}\left(S^{(0)}T_{k_1}^{(1)}S^{(0)}\right)=\sum_{k_1=0}^{K+2}\xi_{k_1}\frac{1}{nT}\operatorname{tr}\left(M_{\tilde{\Gamma}_n}\left(V_{k_1}F_T\tilde{\Gamma}_n'+\tilde{\Gamma}_nF_T'V_{k_1}'\right)M_{\tilde{\Gamma}_n}\right)=0,$$

because $M_{\tilde{\Gamma}_n}\tilde{\Gamma}_n=0$ and $\tilde{\Gamma}_n'M_{\tilde{\Gamma}_n}=0$. For $g=2$, notice that $S^{(0)}T_k^{(1)}S^{(0)}=0$,

$$\begin{aligned}
&\sum_{k_1=0}^{K+2}\sum_{k_2=0}^{K+2}\xi_{k_1}\xi_{k_2}\frac{1}{nT}\operatorname{tr}\left(\tilde{T}_{k_1k_2}^{(2)}\right)\\
=&\sum_{k_1=0}^{K+2}\sum_{k_2=0}^{K+2}\xi_{k_1}\xi_{k_2}\frac{1}{nT}\operatorname{tr}\left(S^{(0)}T_{k_1k_2}^{(2)}S^{(0)}\right)-\sum_{k_1=0}^{K+2}\sum_{k_2=0}^{K+2}\xi_{k_1}\xi_{k_2}\frac{1}{nT}\operatorname{tr}\left(S^{(1)}T_{k_1}^{(1)}S^{(0)}T_{k_2}^{(1)}S^{(0)}\right)\\
&-\sum_{k_1=0}^{K+2}\sum_{k_2=0}^{K+2}\xi_{k_1}\xi_{k_2}\frac{1}{nT}\operatorname{tr}\left(S^{(0)}T_{k_1}^{(1)}S^{(1)}T_{k_2}^{(1)}S^{(0)}\right)-\sum_{k_1=0}^{K+2}\sum_{k_2=0}^{K+2}\xi_{k_1}\xi_{k_2}\frac{1}{nT}\operatorname{tr}\left(S^{(0)}T_{k_1}^{(1)}S^{(0)}T_{k_2}^{(1)}S^{(1)}\right)\\
=&\frac{1}{nT}\operatorname{tr}\left(M_{\tilde{\Gamma}_n}\varepsilon M_{F_T}\varepsilon'\right)+2\sum_{k=1}^{K+1}\left(\delta_{0k}-\delta_k\right)\frac{1}{nT}\operatorname{tr}\left(M_{\tilde{\Gamma}_n}R_nZ_kM_{F_T}\varepsilon'\right)\\
&+\sum_{k_1=1}^{K+1}\sum_{k_2=1}^{K+1}\left(\delta_{0k_1}-\delta_{k_1}\right)\left(\delta_{0k_2}-\delta_{k_2}\right)\frac{1}{nT}\operatorname{tr}\left(M_{\tilde{\Gamma}_n}R_nZ_{k_1}M_{F_T}Z_{k_2}'R_n'\right),
\end{aligned}\tag{S.10}$$

where for the ease of notation, δ_{K+1} denotes λ.

For $g=3$,

$$\begin{aligned}
\tilde{T}_{k_1k_2k_3}^{(3)}=&S^{(0)}T_{k_1k_2k_3}^{(3)}S^{(0)}-S^{(0)}T_{k_1k_2}^{(2)}S^{(1)}T_{k_3}^{(1)}S^{(0)}-S^{(0)}T_{k_1k_2}^{(2)}S^{(0)}T_{k_3}^{(1)}S^{(1)}-S^{(1)}T_{k_1}^{(1)}S^{(0)}T_{k_2k_3}^{(2)}S^{(0)}-S^{(0)}T_{k_1}^{(1)}S^{(1)}T_{k_2k_3}^{(2)}S^{(0)}\\
&+S^{(1)}T_{k_1}^{(1)}S^{(0)}T_{k_2}^{(1)}S^{(1)}T_{k_3}^{(1)}S^{(0)}+S^{(0)}T_{k_1}^{(1)}S^{(1)}T_{k_2}^{(1)}S^{(1)}T_{k_3}^{(1)}S^{(0)}+S^{(0)}T_{k_1}^{(1)}S^{(1)}T_{k_2}^{(1)}S^{(0)}T_{k_3}^{(1)}S^{(1)},
\end{aligned}$$

by using $S^{(0)}T_k^{(1)}S^{(0)}=0$. After simplification, we have

$$\begin{aligned}
&\sum_{k_1,k_2,k_3=0}^{K+2}\xi_{k_1}\xi_{k_2}\xi_{k_3}\frac{1}{nT}\operatorname{tr}\left(\tilde{T}_{k_1k_2k_3}^{(3)}\right)\\
=&2\sum_{k_1,k_2,k_3=0}^{K+2}\xi_{k_1}\xi_{k_2}\xi_{k_3}\frac{1}{nT}\operatorname{tr}\left(M_{\tilde{\Gamma}_n}V_{k_1k_2}M_{F_T}V_{k_3}'\right)-2\sum_{k_1,k_2,k_3=0}^{K+2}\xi_{k_1}\xi_{k_2}\xi_{k_3}\frac{1}{nT}\operatorname{tr}\left(M_{\tilde{\Gamma}_n}V_{k_1}M_{F_T}V_{k_2}'P_{\tilde{\Gamma}_n,F_T}V_{k_3}'\right),
\end{aligned}\tag{S.11}$$

where $P_{\tilde{\Gamma}_n,F_T}=\tilde{\Gamma}_n\left(\tilde{\Gamma}_n'\tilde{\Gamma}_n\right)^{-1}\left(F_T'F_T\right)^{-1}F_T'$. The first term of Eq. (S.11) is

$$\begin{aligned}
&\sum_{k_1,k_2,k_3=0}^{K+2}\xi_{k_1}\xi_{k_2}\xi_{k_3}\frac{1}{nT}\operatorname{tr}\left(M_{\tilde{\Gamma}_n}V_{k_1k_2}M_{F_T}V_{k_3}'\right)\\
=&\left(\lambda_0-\lambda\right)\frac{1}{nT}\operatorname{tr}\left(M_{\tilde{\Gamma}_n}R_nG_nR_n^{-1}\varepsilon M_{F_T}\varepsilon'\right)+\left(\alpha_0-\alpha\right)\frac{1}{nT}\operatorname{tr}\left(M_{\tilde{\Gamma}_n}\tilde{G}_n\varepsilon M_{F_T}\varepsilon'\right)+\left(\alpha_0-\alpha\right)\left(\lambda_0-\lambda\right)\frac{\sigma_0^2}{n}\operatorname{tr}\left(G_n\tilde{G}_n\right)\\
&+O_P\left(\|\theta-\theta_0\|_1^3\right)+O_P\left(\|\theta-\theta_0\|_1^2n^{-\frac{1}{2}}\right),
\end{aligned}\tag{S.12}$$

where in Eq. (S.12), many terms can be put into the remainder $O_P\left(\|\theta-\theta_0\|_1^3\right)+O_P\left(\|\theta-\theta_0\|_1^2n^{-\frac{1}{2}}\right)$. For example, $\sum_{k=1}^{K}(\lambda_0-\lambda)(\delta_{0k}-\delta_k)\frac{1}{nT}\operatorname{tr}\left(M_{\tilde{\Gamma}_n}R_nG_nR_n^{-1}\varepsilon M_{F_T}Z_k'R_n'\right)=O_P\left(\|\theta-\theta_0\|_1^2 n^{-\frac{1}{2}}\right)$. The second term of Eq. (S.11) is

$$\begin{aligned}
&\sum_{k_1,k_2,k_3=0}^{K+2}\xi_{k_1}\xi_{k_2}\xi_{k_3}\frac{1}{nT}\operatorname{tr}\left(M_{\tilde{\Gamma}_n}V_{k_1}M_{F_T}V_{k_2}'P_{\tilde{\Gamma}_n,F_T}V_{k_3}'\right)\\
=&\frac{1}{nT}\operatorname{tr}\left(M_{\tilde{\Gamma}_n}\varepsilon M_{F_T}\varepsilon' P_{\tilde{\Gamma}_n,F_T}\varepsilon'\right)+(\lambda_0-\lambda)\frac{1}{nT}\operatorname{tr}\left(M_{\tilde{\Gamma}_n}\varepsilon M_{F_T}\varepsilon' P_{\tilde{\Gamma}_n}R_n^{-1\prime}G_n'R_n'\right)+(\alpha_0-\alpha)\frac{1}{nT}\operatorname{tr}\left(M_{\tilde{\Gamma}_n}\varepsilon M_{F_T}\varepsilon' P_{\tilde{\Gamma}_n}\tilde{G}_n\right)\\
&+\sum_{k=1}^{K+1}(\delta_{0k}-\delta_k)\left(\frac{1}{nT}\operatorname{tr}\left(M_{\tilde{\Gamma}_n}\varepsilon M_{F_T}\varepsilon' P_{\tilde{\Gamma}_n,F_T}Z_k'R_n'\right)+\frac{1}{nT}\operatorname{tr}\left(M_{\tilde{\Gamma}_n}\varepsilon M_{F_T}Z_k'R_n'P_{\tilde{\Gamma}_n,F_T}\varepsilon'\right)+\frac{1}{nT}\operatorname{tr}\left(M_{\tilde{\Gamma}_n}R_nZ_kM_{F_T}\varepsilon' R_n'P_{\tilde{\Gamma}_n,F_T}\varepsilon'\right)\right)\\
&+O_P\left(\|\theta-\theta_0\|_1^3\right)+O_P\left(\|\theta-\theta_0\|_1^2 n^{-\frac{1}{2}}\right).
\end{aligned}$$

For $g=4$, we have

$$\begin{aligned}
\operatorname{tr}\left(\tilde{T}^{(4)}_{k_1k_2k_3k_4}\right)=&\operatorname{tr}\left(S^{(0)}T^{(4)}_{k_1k_2k_3k_4}S^{(0)}\right)-\operatorname{tr}\left(S^{(0)}T^{(3)}_{k_1k_2k_3}S^{(1)}T^{(1)}_{k_4}S^{(0)}\right)-\operatorname{tr}\left(S^{(0)}T^{(2)}_{k_1k_2}S^{(1)}T^{(2)}_{k_3k_4}S^{(0)}\right)-\operatorname{tr}\left(S^{(0)}T^{(1)}_{k_1}S^{(1)}T^{(3)}_{k_2k_3k_4}S^{(0)}\right)\\
&+\operatorname{tr}\left(S^{(0)}T^{(2)}_{k_1k_2}S^{(0)}T^{(1)}_{k_3}S^{(2)}T^{(1)}_{k_4}S^{(0)}\right)+\operatorname{tr}\left(S^{(0)}T^{(2)}_{k_1k_2}S^{(1)}T^{(1)}_{k_3}S^{(1)}T^{(1)}_{k_4}S^{(0)}\right)+\operatorname{tr}\left(S^{(1)}T^{(1)}_{k_1}S^{(0)}T^{(2)}_{k_2k_3}S^{(0)}T^{(1)}_{k_4}S^{(1)}\right)\\
&+\operatorname{tr}\left(S^{(0)}T^{(1)}_{k_1}S^{(1)}T^{(2)}_{k_2k_3}S^{(1)}T^{(1)}_{k_4}S^{(0)}\right)+\operatorname{tr}\left(S^{(0)}T^{(1)}_{k_1}S^{(2)}T^{(1)}_{k_2}S^{(0)}T^{(2)}_{k_3k_4}S^{(0)}\right)+\operatorname{tr}\left(S^{(0)}T^{(1)}_{k_1}S^{(1)}T^{(1)}_{k_2}S^{(1)}T^{(2)}_{k_3k_4}S^{(0)}\right)\\
&-\operatorname{tr}\left(S^{(0)}T^{(1)}_{k_1}S^{(2)}T^{(1)}_{k_2}S^{(0)}T^{(1)}_{k_3}S^{(1)}T^{(1)}_{k_4}S^{(0)}\right)-\operatorname{tr}\left(S^{(0)}T^{(1)}_{k_1}S^{(1)}T^{(1)}_{k_2}S^{(0)}T^{(1)}_{k_3}S^{(2)}T^{(1)}_{k_4}S^{(0)}\right)\\
&-\operatorname{tr}\left(S^{(1)}T^{(1)}_{k_1}S^{(0)}T^{(1)}_{k_2}S^{(1)}T^{(1)}_{k_3}S^{(0)}T^{(1)}_{k_4}S^{(1)}\right)-\operatorname{tr}\left(S^{(0)}T^{(1)}_{k_1}S^{(1)}T^{(1)}_{k_2}S^{(1)}T^{(1)}_{k_3}S^{(1)}T^{(1)}_{k_4}S^{(0)}\right).
\end{aligned}$$

Substituting in the definitions of $T^{(1)}_{k_1}$, $T^{(2)}_{k_1k_2}$, $T^{(3)}_{k_1k_2k_3}$, $T^{(4)}_{k_1k_2k_3k_4}$, $S^{(0)}$, $S^{(1)}$ and $S^{(2)}$,

$$\begin{aligned}
&\sum_{k_1,k_2,k_3,k_4=0}^{K+2}\xi_{k_1}\xi_{k_2}\xi_{k_3}\xi_{k_4}\frac{1}{nT}\operatorname{tr}\left(\tilde{T}^{(4)}_{k_1k_2k_3k_4}\right)\\
=&(\lambda_0-\lambda)^2\frac{1}{nT}\operatorname{tr}\left(R_nG_nR_n^{-1}\varepsilon\varepsilon' R_n^{-1\prime}G_n'R_n'\right)+2(\lambda_0-\lambda)(\alpha_0-\alpha)\frac{1}{nT}\operatorname{tr}\left(R_nG_nR_n^{-1}\varepsilon\varepsilon'\tilde{G}_n'\right)+(\alpha_0-\alpha)^2\frac{1}{nT}\operatorname{tr}\left(\tilde{G}_n\varepsilon\varepsilon'\tilde{G}_n'\right)\\
&+\mathscr{J}_{nT}+O_P\left(\|\theta-\theta_0\|_1^3\right)+O_P\left(\|\theta-\theta_0\|_1^2 n^{-\frac{1}{2}}\right)+O_P\left(\|\theta-\theta_0\|_1 n^{-\frac{3}{2}}\right) \qquad \text{(S.13)}\\
=&(\lambda_0-\lambda)^2\frac{\sigma_0^2}{n}\operatorname{tr}\left(R_n^{-1\prime}G_n'R_n'R_nG_nR_n^{-1}\right)+2(\lambda_0-\lambda)(\alpha_0-\alpha)\frac{\sigma_0^2}{n}\operatorname{tr}\left(\tilde{G}_n'R_nG_nR_n^{-1}\right)+(\alpha_0-\alpha)^2\frac{\sigma_0^2}{n}\operatorname{tr}\left(\tilde{G}_n'\tilde{G}_n\right)\\
&+\mathscr{J}_{nT}+O_P\left(\|\theta-\theta_0\|_1^3\right)+O_P\left(\|\theta-\theta_0\|_1^2 n^{-\frac{1}{2}}\right)+O_P\left(\|\theta-\theta_0\|_1 n^{-\frac{3}{2}}\right), \qquad \text{(S.14)}
\end{aligned}$$

recognizing that $\mathscr{J}_{nT}$ is exactly the 4-th order term in (C.6) of $L_{nT}(\theta_0)$. In Eq. (S.13), some terms are $O_P\left(\|\theta-\theta_0\|_1 n^{-\frac{3}{2}}\right)$. One example is

$$\begin{aligned}
&(\lambda_0-\lambda)\frac{1}{nT}\operatorname{tr}\left(M_{\tilde{\Gamma}_n}\left(R_nG_nR_n^{-1}\varepsilon\varepsilon'+\varepsilon\varepsilon' R_n^{-1\prime}G_n'R_n'\right)P_{\tilde{\Gamma}_n,F_T}\varepsilon'\right)\\
\leq&|\lambda_0-\lambda|\frac{2r_0}{nT}\left\|M_{\tilde{\Gamma}_n}\right\|_2\|R_n\|_2\|G_n\|_2\left\|R_n^{-1}\right\|_2\|\varepsilon\|_2^3\left\|P_{\tilde{\Gamma}_n,F_T}\right\|_2=O_P\left(|\lambda_0-\lambda|n^{-\frac{3}{2}}\right),
\end{aligned}$$

because $\|\varepsilon\|_2^3 = O_P\left(\max(n,T)^{\frac{3}{2}}\right)$ and $\left\|P_{\tilde{\Gamma}_n,F_T}\right\|_2 = O_P\left(\frac{1}{\sqrt{nT}}\right)$. From Eqs. (S.10), (S.11) and (S.14),

$$
\begin{aligned}
L_{nT}(\theta) = & \left[L_{nT}(\theta_0) - O_P\left(\left(\frac{\|\varepsilon\|_2}{\sqrt{nT}}\right)^5\right)\right] - \frac{2}{\sqrt{nT}}(\theta-\theta_0)'\left(C^{(1)}+C^{(2)}+C^{(3)}\right) + (\theta-\theta_0)'C(\theta-\theta_0) \\
& + O_P\left(\|\theta-\theta_0\|_1^3\right) + O_P\left(\|\theta-\theta_0\|_1^2 (nT)^{-\frac{1}{4}}\right) + O_P\left(\|\theta-\theta_0\|_1 (nT)^{-\frac{3}{4}}\right) \\
= & L_{nT}(\theta_0) - \frac{2}{\sqrt{nT}}(\theta-\theta_0)'\left(C^{(1)}+C^{(2)}+C^{(3)}\right) + (\theta-\theta_0)'C(\theta-\theta_0) + L_{nT}^{\text{rem}}(\theta),
\end{aligned}
$$

where $L_{nT}^{\text{rem}}(\theta)$ collect terms as $O_P\left(\|\theta-\theta_0\|_1^3\right) + O_P\left(\|\theta-\theta_0\|_1^2 (nT)^{-\frac{1}{4}}\right) + O_P\left(\|\theta-\theta_0\|_1 (nT)^{-\frac{3}{4}}\right) + O_P\left((nT)^{-\frac{5}{4}}\right)$. This proves Lemma 8. □

Proof of Lemma 9.

Consider the elements of the $(K+2)\times(K+2)$ matrix C in Eq. (C.10). For $k_1, k_2 = 1,\cdots,K+1$, $C_{1,k_1k_2} = \frac{1}{nT}\text{tr}(M_{\tilde{\Gamma}_n}R_nZ_{k_1}M_{F_T}Z'_{k_2}R'_n) = O_P(1)$, because $\left|\frac{1}{nT}\text{tr}\left(M_{\tilde{\Gamma}_n}R_nZ_{k_1}M_{F_T}Z'_{k_2}R'_n\right)\right| \le \frac{1}{nT}\left\|M_{\tilde{\Gamma}_n}R_nZ_{k_1}\right\|_F\left\|M_{F_T}Z'_{k_2}R'_n\right\|_F = O_P(1)$. $C_2 = O_P(1)$ because W_n, S_n^{-1}, $\tilde{W}_n$ and R_n^{-1} are UB. Therefore $C = O_P(1)$.

Now consider $C_k^{(1)}$.

$$
\begin{aligned}
C_k^{(1)} &= \frac{1}{\sqrt{nT}}\text{tr}\left(M_{\tilde{\Gamma}_n}R_nZ_kM_{F_T}\varepsilon'\right) = \frac{1}{\sqrt{nT}}\text{tr}\left(\mathbb{Z}_k\varepsilon'\right) + \frac{1}{\sqrt{nT}}\text{tr}\left(\left(M_{\tilde{\Gamma}_n}R_nZ_kM_{F_T} - \mathbb{Z}_k\right)\varepsilon'\right) \\
&= \frac{1}{\sqrt{nT}}\text{tr}\left(\mathbb{Z}_k\varepsilon'\right) - \frac{1}{\sqrt{nT}}\text{tr}\left(M_{\tilde{\Gamma}_n}R_n\left(Z_k-\bar{Z}_k\right)P_{F_T}\varepsilon'\right) = O_P(1). \qquad \text{(S.15)}
\end{aligned}
$$

Next we want to show that

$$
C_k^{(2)} = -\frac{1}{\sqrt{nT}}\text{tr}\left(M_{\tilde{\Gamma}_n}\varepsilon M_{F_T}\varepsilon' P_{\tilde{\Gamma}_n,F_T}Z'_kR'_n\right) - \frac{1}{\sqrt{nT}}\text{tr}\left(M_{\tilde{\Gamma}_n}\varepsilon M_{F_T}Z'_kR'_nP_{\tilde{\Gamma}_n,F_T}\varepsilon'\right) - \frac{1}{\sqrt{nT}}\text{tr}\left(M_{\tilde{\Gamma}_n}R_nZ_kM_{F_T}\varepsilon' P_{\tilde{\Gamma}_n,F_T}\varepsilon'\right) \qquad \text{(S.16)}
$$

is $o_P(1)$. Starting with the first term in Eq. (S.16),

$$
\begin{aligned}
& \frac{1}{\sqrt{nT}}\text{tr}\left(\varepsilon M_{F_T}\varepsilon' P_{\tilde{\Gamma}_n,F_T}Z'_kR'_nM_{\tilde{\Gamma}_n}\right) \\
= & \frac{1}{\sqrt{nT}}\text{tr}\left(\varepsilon\varepsilon' P_{\tilde{\Gamma}_n,F_T}\bar{Z}'_kR'_nM_{\tilde{\Gamma}_n}\right) + \frac{1}{\sqrt{nT}}\text{tr}\left(\varepsilon\varepsilon' P_{\tilde{\Gamma}_n,F_T}\left(Z_k-\bar{Z}_k\right)'R'_nM_{\tilde{\Gamma}_n}\right) - \frac{1}{\sqrt{nT}}\text{tr}\left(\varepsilon P_{F_T}\varepsilon' P_{\tilde{\Gamma}_n,F_T}Z'_kR'_nM_{\tilde{\Gamma}_n}\right) \\
= & \frac{1}{\sqrt{nT}}\text{vec}(\varepsilon)'\mathscr{M}\text{vec}(\varepsilon) + \frac{1}{\sqrt{nT}}\text{tr}\left(\varepsilon\varepsilon' P_{\tilde{\Gamma}_n,F_T}\left(Z_k-\bar{Z}_k\right)'R'_nM_{\tilde{\Gamma}_n}\right) - \frac{1}{\sqrt{nT}}\text{tr}\left(\varepsilon P_{F_T}\varepsilon' P_{\tilde{\Gamma}_n,F_T}Z'_kR'_nM_{\tilde{\Gamma}_n}\right),
\end{aligned}
$$

where $\mathscr{M} = \frac{1}{2}\left(I_T\otimes(P_{\tilde{\Gamma}_n,F_T}\bar{Z}'_kR'_nM_{\tilde{\Gamma}_n} + M_{\tilde{\Gamma}_n}R_n\bar{Z}_kP'_{\tilde{\Gamma}_n,F_T})\right)$. Notice that, because $M_{\tilde{\Gamma}_n}P_{\tilde{\Gamma}_n,F_T} = 0$,

$$
\mathbb{E}\text{tr}\left(\mathscr{M}^2\right) \le \frac{T}{2}\mathbb{E}\left|\text{tr}\left(P_{\tilde{\Gamma}_n,F_T}\bar{Z}'_kR'_nM_{\tilde{\Gamma}_n}R_n\bar{Z}_kP'_{\tilde{\Gamma}_n,F_T}\right)\right| \le \frac{Tr_0}{2}\left\|P_{\tilde{\Gamma}_n,F_T}\right\|_2^2\mathbb{E}\left\|\bar{Z}_k\right\|_2^2\left\|R_n\right\|_2^2\left\|M_{\tilde{\Gamma}_n}\right\|_2^2 = O(T),
$$

where $\left\|P_{\tilde{\Gamma}_n,F_T}\right\|_2 = \left(\frac{1}{\sqrt{nT}}\right)$ is from Lemma 4. We have $\mathbb{E}\mathrm{vec}(\varepsilon)'\mathscr{M}\mathrm{vec}(\varepsilon) = \sigma_0^2\mathbb{E}\mathrm{tr}(\mathscr{M}) = 0$, and

$$\mathbb{E}\left(\mathrm{vec}(\varepsilon)'\mathscr{M}\mathrm{vec}(\varepsilon)\right)^2 = \left(\mu^{(4)} - 3\sigma_0^4\right)\sum_{i=1}^{nT}\mathbb{E}\left[\mathscr{M}\right]_{ii}^2 + 2\sigma_0^4\mathbb{E}\mathrm{tr}\left(\mathscr{M}^2\right) = O(T). \tag{S.17}$$

Therefore $\frac{1}{\sqrt{nT}}\mathrm{vec}(\varepsilon)'\mathscr{M}\mathrm{vec}(\varepsilon) = O_P\left(\frac{1}{\sqrt{n}}\right)$ by Markov's inequality. Furthermore,

$$\frac{1}{\sqrt{nT}}\mathrm{tr}\left(\varepsilon\varepsilon' P_{\tilde{\Gamma}_n,F_T}\left(Z_k - \bar{Z}_k\right)' R_n' M_{\tilde{\Gamma}_n}\right) \leq \frac{r_0}{\sqrt{nT}}\|\varepsilon\|_2^2\left\|P_{\tilde{\Gamma}_n,F_T}\right\|_2\|Z_k - \bar{Z}_k\|_2\|R_n\|_2\left\|M_{\tilde{\Gamma}_n}\right\|_2 = O_P\left(\frac{1}{\sqrt{T}}\right), \tag{S.18}$$

because $\|Z_k - \bar{Z}_k\|_2 = O_P\left(\sqrt{n}\right)$ from Lemma 4. The last term, $\frac{1}{\sqrt{nT}}\mathrm{tr}\left(\varepsilon P_{F_T}\varepsilon' P_{\tilde{\Gamma}_n,F_T}Z_k'R_n'M_{\tilde{\Gamma}_n}\right)$ is $o_P(1)$, because

$$\frac{1}{\sqrt{nT}}\mathrm{tr}\left(\varepsilon P_{F_T}\varepsilon' P_{\tilde{\Gamma}_n,F_T}Z_k'R_n'M_{\tilde{\Gamma}_n}\right) \leq \frac{r_0}{\sqrt{nT}}\|\varepsilon\|_2\left\|P_{F_T}\varepsilon' P_{\tilde{\Gamma}_n}\right\|_2\left\|P_{\tilde{\Gamma}_n,F_T}\right\|_2\|Z_k\|_2\|R_n\|_2\left\|M_{\tilde{\Gamma}_n}\right\|_2 = O_P\left(\frac{1}{\sqrt{n}}\right), \tag{S.19}$$

where $\left\|P_{F_T}\varepsilon' P_{\tilde{\Gamma}_n}\right\|_2 = O_P(1)$, $\left\|P_{\tilde{\Gamma}_n,F_T}\right\|_2 = O_P\left(\frac{1}{\sqrt{nT}}\right)$ from Lemma 4, and $\|Z_k\|_2 = O_P\left(\sqrt{nT}\right)$ from Lemma 3. Thus, we have shown that the first term in Eq. (S.16), $\frac{1}{\sqrt{nT}}\mathrm{tr}\left(M_{\tilde{\Gamma}_n}\varepsilon M_{F_T}\varepsilon' P_{\tilde{\Gamma}_n,F_T}Z_k'R_n'\right)$ is $o_P(1)$. The next term in Eq. (S.16) is

$$\begin{aligned}
&\frac{1}{\sqrt{nT}}\mathrm{tr}\left(\varepsilon' M_{\tilde{\Gamma}_n}\varepsilon M_{F_T}Z_k'R_n'P_{\tilde{\Gamma}_n,F_T}\right) \\
=&\frac{1}{\sqrt{nT}}\mathrm{tr}\left(\varepsilon'\varepsilon M_{F_T}\bar{Z}_k'R_n'P_{\tilde{\Gamma}_n,F_T}\right) + \frac{1}{\sqrt{nT}}\mathrm{tr}\left(\varepsilon'\varepsilon M_{F_T}\left(Z_k - \bar{Z}_k\right)' R_n'P_{\tilde{\Gamma}_n,F_T}\right) - \frac{1}{\sqrt{nT}}\mathrm{tr}\left(\varepsilon' P_{\tilde{\Gamma}_n}\varepsilon M_{F_T}Z_k'R_n'P_{\tilde{\Gamma}_n,F_T}\right).
\end{aligned}$$

Using similar arguments from Eqs. (S.17)-(S.19), it can be shown that $\frac{1}{\sqrt{nT}}\mathrm{tr}\left(M_{\tilde{\Gamma}_n}\varepsilon M_{F_T}Z_k'R_n'P_{\tilde{\Gamma}_n,F_T}\varepsilon'\right) = o_P(1)$. The last term in Eq. (S.16) is

$$\begin{aligned}
&\frac{1}{\sqrt{nT}}\mathrm{tr}\left(M_{\tilde{\Gamma}_n}R_nZ_kM_{F_T}\varepsilon' P_{\tilde{\Gamma}_n,F_T}\varepsilon'\right) \\
\leq&\frac{r_0}{\sqrt{nT}}\left\|M_{\tilde{\Gamma}_n}\right\|_2\|R_n\|_2\left\|Z_k\varepsilon' P_{\tilde{\Gamma}_n}\right\|_2\left\|P_{\tilde{\Gamma}_n,F_T}\right\|_2\|\varepsilon\|_2 + \frac{r_0}{\sqrt{nT}}\left\|M_{\tilde{\Gamma}_n}\right\|_2\|R_n\|_2\|Z_k\|_2\left\|P_{F_T}\varepsilon' P_{\tilde{\Gamma}_n}\right\|_2\left\|P_{\tilde{\Gamma}_n,F_T}\right\|_2\|\varepsilon\|_2 = o_P\left(\frac{1}{\sqrt{T}}\right),
\end{aligned}$$

because $\left\|Z_k\varepsilon' P_{\tilde{\Gamma}_n}\right\|_2 = O_P(\sqrt{nT})$ and $\left\|P_{F_T}\varepsilon' P_{\tilde{\Gamma}_n}\right\|_2 = O_P(1)$ from Lemma 4. Therefore $C_k^{(2)} = o_P(1)$.

Now consider,

$$C_{K+1}^{(3)} - \sqrt{\frac{T}{n}}\mathrm{tr}\left(G_n\right)\sigma_0^2 = \frac{1}{\sqrt{nT}}\left(\mathrm{tr}\left(M_{\tilde{\Gamma}_n}R_nG_nR_n^{-1}M_{\tilde{\Gamma}_n}\varepsilon\varepsilon'\right) - \mathrm{tr}\left(M_{\tilde{\Gamma}_n}R_nG_nR_n^{-1}M_{\tilde{\Gamma}_n}\varepsilon P_{F_T}\varepsilon'\right)\right) - \sqrt{\frac{T}{n}}\mathrm{tr}\left(G_n\right)\sigma_0^2.$$

Because

$$\frac{1}{nT}\operatorname{tr}\left(M_{\tilde{\Gamma}_n}R_nG_nR_n^{-1}M_{\tilde{\Gamma}_n}\varepsilon\varepsilon'\right)=\frac{\sigma_0^2}{n}\operatorname{tr}\left(M_{\tilde{\Gamma}_n}R_nG_nR_n^{-1}\right)+O_P(\frac{1}{\sqrt{nT}})$$
$$=\frac{\sigma_0^2}{n}\operatorname{tr}\left(R_nG_nR_n^{-1}\right)+O_P\left(\frac{1}{n}\right)+O_P(\frac{1}{\sqrt{nT}})=\frac{\sigma_0^2}{n}\operatorname{tr}\left(G_n\right)+O_P(\frac{1}{\sqrt{nT}}),$$

and $\frac{1}{\sqrt{nT}}\left|\operatorname{tr}\left(R_nG_nR_n^{-1}\varepsilon P_{F_T}\varepsilon'\right)\right|\leq\frac{r_0}{\sqrt{nT}}\left\|R_nG_nR_n^{-1}\right\|_2\left\|P_{F_T}\right\|_2\left\|\varepsilon\right\|_2^2=O_P(1)$, we have $C_{K+1}^{(3)}-\sqrt{\frac{T}{n}}\operatorname{tr}\left(G_n\right)\sigma_0^2=O_P(1)$. Similarly, $C_{K+2}^{(3)}-\sqrt{\frac{T}{n}}\operatorname{tr}\left(\tilde{G}_n\right)\sigma_0^2=O_P(1)$. □

Proof of Lemma 10.

In the following, we show that $\frac{1}{nT\sigma_0^2}\operatorname{tr}\left(M_{\tilde{\Gamma}_n}R_nZ_1M_{F_T}Z_1'R_n'\right)=\frac{1}{nT\sigma_0^2}\operatorname{tr}\left(M_{\tilde{\Gamma}_n}R_n\bar{Z}_1M_{F_T}\bar{Z}_1'R_n'\right)+\frac{1}{n}\operatorname{tr}\left(R_n\mathbf{P}_nR_n'\right)+o_P(1)$, where $\mathbf{P}_n=\sum_{h=1}^{\infty}A_n^{h-1}S_n^{-1}R_n^{-1}R_n^{-1\prime}S_n^{-1\prime}A_n^{h-1\prime}$, and the rest of the lemma can be proved analogously.

$$\frac{1}{nT\sigma_0^2}\operatorname{tr}\left(M_{\tilde{\Gamma}_n}R_nZ_1M_{F_T}Z_1'R_n'\right)$$
$$=\frac{1}{nT\sigma_0^2}\operatorname{tr}\left(M_{\tilde{\Gamma}_n}R_n\bar{Z}_1M_{F_T}\bar{Z}_1'R_n'\right)+\frac{1}{nT\sigma_0^2}\operatorname{tr}\left(M_{\tilde{\Gamma}_n}R_n\left(Z_1-\bar{Z}_1\right)M_{F_T}\left(Z_1-\bar{Z}_1\right)'R_n'\right)+o_P(1). \tag{S.20}$$

In Eq. (S.20), the remainder term is $o_P(1)$, because from Eq. (6), $Z_1-\bar{Z}_1=\sum_{h=1}^{\infty}A_n^{h-1}S_n^{-1}R_n^{-1}\tilde{\varepsilon}_h$,

$$\frac{1}{nT\sigma_0^2}\left|\operatorname{tr}\left(\left(Z_1-\bar{Z}_1\right)M_{F_T}\bar{Z}_1'R_n'M_{\tilde{\Gamma}_n}R_n\right)\right|$$
$$\leq\frac{1}{nT\sigma_0^2}\left|\operatorname{tr}\left(\left(Z_1-\bar{Z}_1\right)\bar{Z}_1'R_n'R_n\right)\right|+\frac{1}{nT\sigma_0^2}\left|\operatorname{tr}\left(\left(Z_1-\bar{Z}_1\right)P_{F_T}\bar{Z}_1'R_n'R_n\right)\right|$$
$$+\frac{1}{nT\sigma_0^2}\left|\operatorname{tr}\left(\left(Z_1-\bar{Z}_1\right)\bar{Z}_1'R_n'P_{\tilde{\Gamma}_n}R_n\right)\right|+\frac{1}{nT\sigma_0^2}\left|\operatorname{tr}\left(\left(Z_1-\bar{Z}_1\right)P_{F_T}\bar{Z}_1'R_n'P_{\tilde{\Gamma}_n}R_n\right)\right|$$
$$\leq\frac{1}{nT\sigma_0^2}\left|\operatorname{tr}\left(\left(Z_1-\bar{Z}_1\right)\bar{Z}_1'R_n'R_n\right)\right|+\frac{3r_0}{nT\sigma_0^2}\left\|Z_1-\bar{Z}_1\right\|_2\left\|\bar{Z}_1\right\|_2\left\|R_n\right\|_2^2 \tag{S.21}$$
$$=\frac{1}{nT\sigma_0^2}\operatorname{tr}\left(\sum_{h=1}^{\infty}\bar{Z}_1'R_n'R_nS_n^{-1}R_n^{-1}Q_h\tilde{\varepsilon}_h\right)+o_P(1) \tag{S.22}$$

where $Q_h=R_nS_nA_n^{h-1}S_n^{-1}R_n^{-1}$ and $\left\|Z_1-\bar{Z}_1\right\|_2=O_P\left(\sqrt{n}\right)$ in (6) from Lemma 4(2). We have

$$\operatorname{tr}\left(\sum_{h=1}^{\infty}\bar{Z}_1'R_n'R_nS_n^{-1}R_n^{-1}Q_h\tilde{\varepsilon}_h\right)=\sum_{i=1}^{n}\sum_{t=1}^{T}\left[\bar{Z}_1'R_n'R_nS_n^{-1}R_n^{-1}\right]_{it}\sum_{h=1}^{\infty}\sum_{j=1}^{n}Q_{h,ij}\varepsilon_{j,t-h},$$

and hence

$$\mathbb{E}\operatorname{tr}\left(\sum_{h=1}^{\infty}\bar{Z}_1'R_n'R_nS_n^{-1}R_n^{-1}Q_h\tilde{\varepsilon}_h\right)^2$$
$$\leq C\mathbb{E}\sum_{i,i'=1}^{n}\sum_{t,t'=1}^{T}\sum_{h,h'=1}^{\infty}\sum_{j,j'=1}^{n}Q_{h,ij}Q_{h',i'j'}\varepsilon_{j,t-h}\varepsilon_{j',t'-h'}=CT\sigma_0^2\sum_{i,i',j=1}^{n}\sum_{h,h'=1}^{\infty}Q_{h,ij}Q_{h',i'j}$$

$$\leq CT\sigma_0^2 \sum_{j=1}^{n}\sum_{h=1}^{\infty}\sum_{i=1}^{n} Q_{h,ij} \sum_{h'=1}^{\infty}\sum_{i'=1}^{n} Q_{h',i'j} \leq CnT\sigma_0^2 \left(\left\| \sum_{h=1}^{\infty} Q_h \right\|_1 \right)^2 = O(nT),$$

where C is some constant and in the last equation, $\|\sum_{h=1}^{\infty} Q_h\|_1$ is uniformly bounded by Assumption R5. Therefore $\frac{1}{nT\sigma_0^2}\operatorname{tr}\left(\sum_{h=1}^{\infty} \bar{Z}_1' R_n' R_n S_n^{-1} R_n^{-1} Q_h \tilde{\varepsilon}_h\right) = O_P\left(\frac{1}{\sqrt{nT}}\right)$. Continuing from Eq. (S.20),

$$\frac{1}{nT\sigma_0^2}\operatorname{tr}\left(M_{\tilde{\Gamma}_n} R_n (Z_1 - \bar{Z}_1) M_{F_T} (Z_1 - \bar{Z}_1)' R_n'\right)$$

$$= \frac{1}{nT\sigma_0^2}\operatorname{tr}\left(R_n (Z_1 - \bar{Z}_1)(Z_1 - \bar{Z}_1)' R_n'\right) + o_P(1) \tag{S.23}$$

$$= \frac{1}{nT\sigma_0^2}\sum_{t=1}^{T} \mathbb{U}_{nt}'\mathbb{U}_{nt} + o_P(1) = \frac{1}{nT\sigma_0^2}\mathbb{E}\sum_{t=1}^{T} \mathbb{U}_{nt}'\mathbb{U}_{nt} + o_P(1) = \frac{1}{n}\operatorname{tr}\left(R_n \mathbf{P}_n R_n'\right) + o_P(1). \tag{S.24}$$

where $\mathbb{U}_{nt} = R_n \sum_{h=1}^{\infty} A_{0n}^{h-1} S_n^{-1} R_n^{-1} \varepsilon_{n,t-h}$, and Eq. (S.24) is due to Lemma 7 in Yu et al. (2008). The remainder terms in Eq. (S.23) are $o_P(1)$, because

$$\left| \frac{1}{nT\sigma_0^2}\operatorname{tr}\left(P_{\tilde{\Gamma}_n} R_n (Z_1 - \bar{Z}_1) P_{F_T} (Z_1 - \bar{Z}_1)' R_n'\right) \right| \leq \frac{r_0}{nT\sigma_0^2} \|P_{\tilde{\Gamma}_n}\|_2^2 \|R_n\|_2^2 \|Z_1 - \bar{Z}_1\|_2^2 \|P_{F_T}\|_2^2 = O_P\left(\frac{1}{T}\right),$$

where $\|Z_1 - \bar{Z}_1\|_2 = O_P(\sqrt{n})$ by Lemma 4 and the other terms are $O_P\left(\frac{1}{T}\right)$ analogously. In conclusion,

$$\frac{1}{nT\sigma_0^2}\operatorname{tr}\left(M_{\tilde{\Gamma}_n} R_n Z_1 M_{F_T} Z_1' R_n'\right) = \frac{1}{nT\sigma_0^2}\operatorname{tr}\left(M_{\tilde{\Gamma}_n} R_n \bar{Z}_1 M_{F_T} \bar{Z}_1' R_n'\right) + \frac{1}{n}\operatorname{tr}\left(R_n \mathbf{P}_n R_n'\right) + o_P(1).$$

□

Proof of Lemma 11.

Because $P(\xi) = M_{\hat{\Gamma}_n}$, the properties of $M_{\hat{\Gamma}_n}$ can be derived from examining the series expansion of $P(\xi)$. Substituting the Laurent series $R(\zeta) = \sum_{n=0}^{\infty} \zeta^{n-1} S^{(n)}$ into the integrand in Eq. (S.4), we have $M_{\hat{\Gamma}_n} = M_{\tilde{\Gamma}_n} + \sum_{g=1}^{\infty} M_{\tilde{\Gamma}_n,k_1\cdots k_g}^{(g)}$ and

$$M_{\tilde{\Gamma}_n,k_1\cdots k_g}^{(g)} = -\sum_{p=\lceil \frac{g}{4} \rceil}^{g} (-1)^p \sum_{\substack{v_1+v_2+\cdots+v_p=g \\ m_1+\cdots+m_{p+1}=p \\ v_j=1,\cdots,4,\, m_j=0,1,\cdots}} \sum_{k_1=0}^{K+2}\cdots\sum_{k_g=0}^{K+2} \xi_{k_1}\xi_{k_2}\cdots\xi_{k_g} S^{(m_1)} T_{k_1\cdots}^{(v_1)} S^{(m_2)} \cdots S^{(m_p)} T_{\cdots k_g}^{(v_p)} S^{(m_{p+1})}.$$

Because $\left\|S^{(0)}\right\|_2 = 1$, $\left\|S^{(m)}\right\|_2 \leq \left(nT d_{\min}^2(\tilde{\Gamma}_n, F_T)\right)^{-m}$ for $m \geq 1$ and use the definition of b_{nT} in Eq. (C.3),

$$\left\|M_{\tilde{\Gamma}_n,k_1\cdots k_g}^{(g)}\right\|_2 \leq \sum_{p=1}^{g} \sum_{\substack{v_1+v_2+\cdots+v_p=g \\ m_1+\cdots+m_{p+1}=p \\ v_j=1,\cdots,4,\, m_j=0,1,\cdots}} b_{nT}^g \left(\frac{d_{\max}^2(\tilde{\Gamma}_n, F_T)}{d_{\min}^2(\tilde{\Gamma}_n, F_T)}\right)^p$$

$$= b_{nT}^g \sum_{p=1}^{g} \left(\frac{d_{\max}^2(\tilde{\Gamma}_n, F_T)}{d_{\min}^2(\tilde{\Gamma}_n, F_T)}\right)^p \sum_{\substack{m_1+\cdots+m_{p+1}=p \\ m_j=0,1,\cdots}} 1 \sum_{\substack{v_1+\cdots+v_p=g \\ v_j=1,\cdots,4}} 1$$

$$
\begin{aligned}
&\leq b_{nT}^{g}\sum_{p=1}^{g}\left(\frac{16d_{\max}^{2}(\tilde{\Gamma}_n,F_T)}{d_{\min}^{2}(\tilde{\Gamma}_n,F_T)}\right)^{p}\\
&\leq \frac{16d_{\max}^{2}(\tilde{\Gamma}_n,F_T)d_{\min}^{2}(\tilde{\Gamma}_n,F_T)}{d_{\min}^{2}(\tilde{\Gamma}_n,F_T)\left(16d_{\max}^{2}(\tilde{\Gamma}_n,F_T)-d_{\min}^{2}(\tilde{\Gamma}_n,F_T)\right)}\left(\frac{16d_{\max}^{2}(\tilde{\Gamma}_n,F_T)b_{nT}}{d_{\min}^{2}(\tilde{\Gamma}_n,F_T)}\right)^{g}\\
&=O_P(b_{nT}^{g}).
\end{aligned}
$$

Under the assumptions of Theorem 3, $\hat{\theta}=\theta_0+O_P\left(\frac{1}{\sqrt{nT}}\right)$. Under Assumption E, $|\xi_0|=O_P\left(\frac{1}{\sqrt{n}}\right)$. Therefore $b_{nT}=O_P\left(\frac{1}{\sqrt{n}}\right)$ and we have $\left\|M_{\hat{\tilde{\Gamma}}_n}-M_{\tilde{\Gamma}_n}\right\|_2=O_P(b_{nT})=O_P(\frac{1}{\sqrt{n}})$. Interchanging the role of n and T, we also have $\left\|M_{\hat{F}_T}-M_{F_T}\right\|_2=O_P(b_{nT})=O_P(\frac{1}{\sqrt{T}})$. □

S3. Asymptotic Distributions of the QML Estimator in Special Cases

In this appendix, we present explicit expressions of the asymptotic variances for two special cases of the general model, which may be of interest for empirical researchers.

A. No Spatial Correlation in Disturbances

The model is $Y_{nt}=\lambda_0 W_nY_{nt}+\gamma_0 Y_{n,t-1}+\rho_0 W_nY_{n,t-1}+X_{nt}\beta_0+\Gamma_{n0}f_{t0}+\varepsilon_{nt}$ with i.i.d. ε_{it} with zero mean and finite variance σ_0^2. This is a special case with $\alpha_0=0$. Denote $\theta=(\delta',\lambda)$. The concentrated likelihood function is $Q_{nT}(\theta,\Gamma_n,F_T)=\frac{1}{n}\log|S_n(\lambda)|-\frac{1}{2}\log\left(L_{nT}(\theta)\right)$, where

$$
L_{nT}(\theta)=\frac{1}{nT}\sum_{i=r+1}^{n}\mu_i\left(\left(S_n(\lambda)Y-\sum_{k=1}^{K}Z_k\delta_k\right)\left(S_n(\lambda)Y-\sum_{k=1}^{K}Z_k\delta_k\right)'\right).
$$

In this context, Assumptions NC1 and NC2 will be simplified to the following.

Assumption NC1-A. There exists a positive constant b, such that $\min_{\eta\in B_{K+1}}\sum_{i=2r+1}^{n}\mu_i\left(\frac{1}{nT}(\eta\cdot Z)(\eta\cdot Z)'\right)\geq b>0$ wpa 1 as n, $T\to\infty$, where B_{K+1} is the unit ball of the $(K+1)-$dimensional Euclidean space; η is a $(K+1)\times 1$ nonzero vector with $\|\eta\|_2=\sqrt{\eta'\eta}=1$; $\eta\cdot Z\equiv\sum_{k=1}^{K+1}\eta_kZ_k$ is a convex linear combination of those $n\times T$ matrices Z_k's.

Assumption NC2-A.

1. Suppose that $G_nZ_{nt}\delta_0=Z_{nt}C$ for a constant vector C. There exists a positive constant b, such that $\min_{\eta\in B_K}\sum_{i=2r+1}^{n}\mu_i\left(\frac{1}{nT}(\eta\cdot Z)(\eta\cdot Z)'\right)\geq b>0$ wpa 1 as n, $T\to\infty$, where B_K is the unit ball of the $K-$dimensional Euclidean space; η is a $K\times 1$ nonzero vector with $\|\eta\|_2=\sqrt{\eta'\eta}=1$; $\eta\cdot Z\equiv\sum_{k=1}^{K}\eta_kZ_k$ is a convex linear combination of those $n\times T$ matrices Z_k's.
2. For any $\lambda\in\Theta_\lambda$ if $\lambda\neq\lambda_0$, $\liminf_{n,T\to\infty}\left(\frac{1}{n}\mathrm{tr}\left(S_n^{-1\prime}S_n(\lambda)'S_n(\lambda)S_n^{-1}\right)-\left|S_n^{-1\prime}S_n(\lambda)'S_n(\lambda)S_n^{-1}\right|^{\frac{1}{n}}\right)>0$.

To derive the asymptotic distribution of $\hat{\theta}$, we expand $L_{nT}(\theta)$ in terms of $\theta-\theta_0$ using Lemma 7 with $S_n(\lambda)Y-\sum_{k=1}^{K}Z_k\delta_k=\Gamma_nF_T'+\sum_{k=0}^{K+1}\xi_kV_k+\sum_{k_1=0,k_2=0}^{K+1}\xi_{k_1}\xi_{k_2}V_{k_1k_2}$, where $\xi_0=\frac{\|\varepsilon\|_2}{\sqrt{nT}}$, $\xi_k=\delta_{0k}-\delta_k$ for $k=1,\cdots,K$, $\xi_{K+1}=\lambda_0-\lambda$, $V_0=\frac{\sqrt{nT}\varepsilon}{\|\varepsilon\|_2}$, $V_k=Z_k$ for $k=1,\cdots,K$, $V_{K+1}=Z_{K+1}+G_n\Gamma_nF_T'$, $V_{0,K+1}=G_n\frac{\sqrt{nT}\varepsilon}{\|\varepsilon\|_2}$, and $V_{k_1k_2}=\mathbf{0}$ otherwise. $L_{nT}(\theta)$ is

$$L_{nT}(\xi)=\sum_{i=r_0+1}^{n}\mu_i\left(\frac{1}{nT}T^{(0)}+\frac{1}{nT}\sum_{k_1=0}^{K+1}\xi_{k_1}T_{k_1}^{(1)}+\frac{1}{nT}\sum_{k_1=0}^{K+1}\sum_{k_2=0}^{K+1}\xi_{k_1}\xi_{k_2}T_{k_1k_2}^{(2)}\right.$$
$$\left.+\frac{1}{nT}\sum_{k_1=0}^{K+1}\sum_{k_2=0}^{K+1}\sum_{k_3=0}^{K+1}\xi_{k_1}\xi_{k_2}\xi_{k_3}T_{k_1k_2k_3}^{(3)}+\frac{1}{nT}\sum_{k_1=0}^{K+1}\sum_{k_2=0}^{K+1}\sum_{k_3=0}^{K+1}\sum_{k_4=0}^{K+1}\xi_{k_1}\xi_{k_2}\xi_{k_3}\xi_{k_4}T_{k_1k_2k_3k_4}^{(4)}\right), \quad \text{(S.25)}$$

where the operators are $T^{(0)}=\tilde{\Gamma}_nF_T'F_T\tilde{\Gamma}_n'$, $T_{k_1}^{(1)}=V_{k_1}F_T\tilde{\Gamma}_n'+\tilde{\Gamma}_nF_T'V_{k_1}'$, $T_{k_1k_2}^{(2)}=V_{k_1k_2}F_T\tilde{\Gamma}_n'+\tilde{\Gamma}_nF_T'V_{k_1k_2}'+V_{k_1}V_{k_2}'$, $T_{k_1k_2k_3}^{(3)}=V_{k_1k_2}V_{k_3}'+V_{k_3}V_{k_1k_2}'$, and $T_{k_1k_2k_3k_4}^{(4)}=V_{k_1k_2}V_{k_3k_4}'$, with $k_j=0,\cdots,K+1$, and $j=1,2,3,4$. The difference between the series expansions of Eq. (S.25) and Eq. (C.2) is that the perturbations due to $\alpha_0-\alpha$ are not present and there is no R_n in V_k for $k=1,\cdots,K+1$. The series expansion in Lemma 8 can be adjusted accordingly. We have the following limiting distribution of the QML estimator $\hat{\theta}$ of this model.

Theorem S.1. *Let $\theta=(\delta',\lambda)'$. Assume that $\frac{T}{n}\to\kappa^2>0$; $D=plim_{n,T\to\infty}D_{nT}$ is positive definite; $\Sigma=p\lim_{n,T\to\infty}\Sigma_{nT}$; $\varphi=p\lim_{n,T\to\infty}\varphi_{nT}$; and suppose that Assumptions NC1-A (or NC2-A), E, R and SF hold, then $\sqrt{nT}\left(\hat{\theta}-\theta_0\right)-\left(\sigma_0^2D\right)^{-1}\varphi\xrightarrow{d}N\left(0,D^{-1}\left(D+\Sigma\right)D^{-1}\right)$, where*

$$\varphi_{nT}=\begin{pmatrix}-\frac{\sigma_0^2}{\sqrt{nT}}\sum_{h=1}^{T-1}tr\left(J_0P_{F_T}J_h'\right)tr\left(A_n^{h-1}S_n^{-1}\right)\\ -\frac{\sigma_0^2}{\sqrt{nT}}\sum_{h=1}^{T-1}tr\left(J_0P_{F_T}J_h'\right)tr\left(W_nA_n^{h-1}S_n^{-1}\right)\\ 0\\ \vdots\\ 0\\ -\frac{\sigma_0^2}{\sqrt{nT}}\sum_{h=1}^{T-1}tr\left(J_0P_{F_T}J_h'\right)tr\left(\left(\gamma G_n+\rho G_nW_n\right)A_n^{h-1}S_n^{-1}\right)+\sqrt{\frac{T}{n}}\sigma_0^2\left(\frac{r_0}{n}tr\left(G_n\right)-tr\left(P_{\tilde{\Gamma}_n}G_n\right)\right)\end{pmatrix},$$

$$D_{nT}=\frac{1}{\sigma_0^2}\begin{pmatrix}\frac{1}{nT}tr\left(M_{\Gamma_n}Z_1M_{F_T}Z_1'\right) & \cdots & \frac{1}{nT}tr\left(M_{\Gamma_n}Z_1M_{F_T}Z_{K+1}'\right)\\ \vdots & & \vdots\\ \frac{1}{nT}tr\left(M_{\Gamma_n}Z_1M_{F_T}Z_{K+1}'\right) & \cdots & \frac{1}{nT}tr\left(M_{\Gamma_n}Z_{K+1}M_{F_T}Z_{K+1}'\right)\end{pmatrix}+\begin{pmatrix}0 & \cdots & 0 & 0\\ \vdots & & \vdots & \vdots\\ 0 & \cdots & 0 & 0\\ 0 & \cdots & 0 & \psi_{K+1,K+1}\end{pmatrix},$$

with $\psi_{K+1,K+1}=\frac{1}{n}tr\left(G_nG_n'\right)+\frac{1}{n}tr(G_n^2)-2\left(\frac{1}{n}tr\left(G_n\right)\right)^2$ and

$$\Sigma_{nT}=\frac{\mu^{(3)}}{\sigma_0^4}\begin{pmatrix} & & & \Sigma_{nT,A}^{1,K+1}\\ & \mathbf{0}_{K\times K} & & \vdots\\ & & & \Sigma_{nT,A}^{K,K+1}\\ \Sigma_{nT,A}^{1,K+1} & \cdots & \Sigma_{nT,A}^{K,K+1} & 2\Sigma_{nT,A}^{K+1,K+1}\end{pmatrix}+\frac{\mu^{(4)}-3\sigma_0^4}{\sigma_0^4}\begin{pmatrix} & & & 0\\ & \mathbf{0}_{K\times K} & & \vdots\\ & & & 0\\ 0 & \cdots & 0 & \Sigma_{nT,B}^{K+1,K+1}\end{pmatrix},$$

where $\Sigma_{nT,A}^{k_1,K+1}=\frac{1}{nT}\sum_{i=1}^{nT}[vec\left(M_{\Gamma_n}\bar{Z}_{k_1}M_{F_T}\right)]_iH_{K+1,ii}$ for $k_1=1,\cdots,K+1$; $\Sigma_{nT,B}^{K+1,K+1}=\frac{1}{nT}\sum_{i=1}^{nT}\left(H_{K+1,ii}\right)^2$;

and $H_{K+1} = I_T \otimes \left(M_{\Gamma_n} G_n - M_{\Gamma_n} G_n P_{\Gamma_n} - \frac{1}{n} tr(G_n) M_{\Gamma_n}\right)$.

B. No Contemporaneous Spatial Interaction

If there is no contemporaneous spatial interaction, $\lambda_0 = 0$. The model is simplified to $Y_{nt} = \gamma_0 Y_{n,t-1} + \rho_0 W_n Y_{n,t-1} + X_{nt}\beta_0 + \Gamma_{n0} f_{t0} + U_{nt}$ with $U_{nt} = \alpha_0 \tilde{W}_n U_{nt} + \varepsilon_{nt}$. The concentrated likelihood function is $Q_{nT}(\theta, \tilde{\Gamma}_n, F_T) = \frac{1}{n}\log|R_n(\alpha)| - \frac{1}{2}\log(L_{nT}(\theta))$ where $L_{nT}(\theta) = \frac{1}{nT}\sum_{i=r+1}^{n}\mu_i\left(R_n(\alpha)\left(Y - \sum_{k=1}^{K} Z_k\delta_k\right)\left(Y - \sum_{k=1}^{K} Z_k\delta_k\right)' R_n(\alpha)'\right)$ and $\theta = (\delta', \alpha)$. In this case, Assumption NC can be simplified.

Assumption NC-B.

1. There exists a positive constant b, such that $\min_{\eta\in B_K, \alpha\in\Theta_\alpha}\sum_{i=2r+1}^{n}\mu_i\left(\frac{1}{nT}R_n(\alpha)(\eta\cdot Z)(\eta\cdot Z)' R_n(\alpha)'\right) \geq b > 0$ wpa 1 as n, $T\to\infty$, where B_K is the unit ball of the $K-$dimensional Euclidean space; η is a $K\times 1$ nonzero vector with $\|\eta\|_2 = \sqrt{\eta'\eta} = 1$; $\eta\cdot Z \equiv \sum_{k=1}^{K}\eta_k Z_k$ is a convex linear combination of those $n\times T$ matrices Z_k's.
2. For any $\alpha\in\Theta_\alpha$, $\alpha\neq\alpha_0$, $\liminf_{n,T\to\infty}\left(\frac{1}{n}\text{tr}\left(R_n^{-1\prime} R_n(\alpha)' R_n(\alpha) R_n^{-1}\right) - \left|R_n^{-1\prime} R_n(\alpha)' R_n(\alpha) R_n^{-1}\right|^{\frac{1}{n}}\right) > 0$.

To derive the asymptotic distribution of $\hat{\theta}$, we expand $L_{nT}(\theta)$ in terms of $\theta-\theta_0$ using Lemma 7 with $R_n(\alpha)\left(Y - \sum_{k=1}^{K} Z_k\delta_k\right) = \tilde{\Gamma}_n F_T' + \sum_{k=0}^{K+1}\xi_k V_k + \sum_{k_1=0,k_2=0}^{K+1}\xi_{k_1}\xi_{k_2}V_{k_1k_2}$, where $\xi_0 = \frac{\|\varepsilon\|_2}{\sqrt{nT}}$, $\xi_k = \delta_{0k} - \delta_k$ for $k = 1,\cdots,K$, $\xi_{K+1} = \alpha_0 - \alpha$, $V_0 = \frac{\sqrt{nT}\varepsilon}{\|\varepsilon\|_2}$, $V_k = R_n Z_k$ for $k = 1,\cdots,K$, and $V_{K+1} = \tilde{W}_n\Gamma_n F_T'$; and the $n\times T$ matrices indexed by k_1, k_2 are $V_{0,K+1} = \tilde{G}_n\frac{\sqrt{nT}\varepsilon}{\|\varepsilon\|_2}$, $V_{k_1,K+1} = \tilde{W}_n Z_{k_1}$ for $k_1 = 1,\cdots,K$ and $V_{k_1k_2} = \mathbf{0}$ otherwise. $L_{nT}(\theta)$ is

$$
\begin{aligned}
L_{nT}(\xi) = \sum_{i=r+1}^{n}\mu_i\Bigg(&\frac{1}{nT}T^{(0)} + \frac{1}{nT}\sum_{k_1=0}^{K+1}\xi_{k_1}T_{k_1}^{(1)} + \frac{1}{nT}\sum_{k_1=0}^{K+1}\sum_{k_2=0}^{K+1}\xi_{k_1}\xi_{k_2}T_{k_1k_2}^{(2)} \\
&+\frac{1}{nT}\sum_{k_1=0}^{K+1}\sum_{k_2=0}^{K+1}\sum_{k_3=0}^{K+1}\xi_{k_1}\xi_{k_2}\xi_{k_3}T_{k_1k_2k_3}^{(3)} + \frac{1}{nT}\sum_{k_1=0}^{K+1}\sum_{k_2=0}^{K+1}\sum_{k_3=0}^{K+1}\sum_{k_4=0}^{K+1}\xi_{k_1}\xi_{k_2}\xi_{k_3}\xi_{k_4}T_{k_1k_2k_3k_4}^{(4)}\Bigg),
\end{aligned}
\tag{S.26}
$$

where the operators are $T^{(0)} = \tilde{\Gamma}_n F_T' F_T \tilde{\Gamma}_n'$, $T_{k_1}^{(1)} = V_{k_1}F_T\tilde{\Gamma}_n' + \tilde{\Gamma}_n F_T' V_{k_1}'$, $T_{k_1k_2}^{(2)} = V_{k_1k_2}F_T\tilde{\Gamma}_n' + \tilde{\Gamma}_n F_T' V_{k_1k_2}' + V_{k_1}V_{k_2}'$, $T_{k_1k_2k_3}^{(3)} = V_{k_1k_2}V_{k_3}' + V_{k_3}V_{k_1k_2}'$, and $T_{k_1k_2k_3k_4}^{(4)} = V_{k_1k_2}V_{k_3k_4}'$, with $k_j = 0,\cdots,K+1$, and $j = 1,2,3,4$. The difference between the series expansions of Eq. (S.26) and Eq. (C.2) is that the perturbations due to $\lambda_0 - \lambda$ are not present. The series expansion in Lemma 8 can be adjusted accordingly. We have the following result.

Theorem S.2. *Let* $\theta = (\delta', \alpha)'$. $D = plim_{n,T\to\infty}D_{nT}$ *is positive definite;* $\Sigma = p\lim_{n,T\to\infty}\Sigma_{nT}$*;* $\varphi = p\lim_{n,T\to\infty}\varphi_{nT}$*; and suppose that Assumptions NC-B, E, R and SF hold, then* $\sqrt{nT}\left(\hat{\theta}-\theta_0\right) - \left(\sigma_0^2 D\right)^{-1}\varphi \xrightarrow{d} N\left(0, D^{-1}(D+\Sigma)D^{-1}\right)$, *where*

$$\varphi_{nT}=\begin{pmatrix} -\frac{\sigma_0^2}{\sqrt{nT}}\sum_{h=1}^{T-1}tr\left(J_0P_{F_T}J_h'\right)tr\left(R_n\dot{A}_n^{h-1}R_n^{-1}\right) \\ -\frac{\sigma_0^2}{\sqrt{nT}}\sum_{h=1}^{T-1}tr\left(J_0P_{F_T}J_h'\right)tr\left(R_nW_n\dot{A}_n^{h-1}R_n^{-1}\right) \\ 0 \\ \vdots \\ 0 \\ \sqrt{\frac{T}{n}}\sigma_0^2\left(\frac{r_0}{n}tr\left(\tilde{G}_n\right)-tr\left(P_{\tilde{\Gamma}_n}\tilde{G}_n\right)\right) \end{pmatrix},$$

$\dot{A}_n=\gamma_0I_n+\rho_0W_n,$

$$D_{nT}=\frac{1}{\sigma_0^2}\begin{pmatrix} \frac{1}{nT}tr\left(M_{\tilde{\Gamma}_n}R_nZ_1M_{F_T}Z_1'R_n'\right) & \cdots & \frac{1}{nT}tr\left(M_{\tilde{\Gamma}_n}R_nZ_1M_{F_T}Z_K'R_n'\right) & 0 \\ \vdots & & \vdots & \vdots \\ \frac{1}{nT}tr\left(M_{\tilde{\Gamma}_n}R_nZ_1M_{F_T}Z_K'R_n'\right) & \cdots & \frac{1}{nT}tr\left(M_{\tilde{\Gamma}_n}R_nZ_KM_{F_T}Z_K'R_n'\right) & 0 \\ 0 & \cdots & 0 & 0 \end{pmatrix}+\begin{pmatrix} 0 & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & \psi_{K+1,K+1} \end{pmatrix},$$

with $\psi_{K+1,K+1}=\frac{1}{n}tr\left(\tilde{G}_n\tilde{G}_n'\right)+\frac{1}{n}tr(\tilde{G}_n^2)-2\left(\frac{1}{n}tr\left(\tilde{G}_n\right)\right)^2$ *and*

$$\Sigma_{nT}=\frac{\mu^{(3)}}{\sigma_0^4}\begin{pmatrix} & & & \Sigma_{nT,A}^{1,K+1} \\ & \mathbf{0}_{K\times K} & & \vdots \\ & & & \Sigma_{nT,A}^{K,K+1} \\ \Sigma_{nT,A}^{1,K+1} & \cdots & \Sigma_{nT,A}^{K,K+1} & 0 \end{pmatrix}+\frac{\mu^{(4)}-3\sigma_0^4}{\sigma_0^4}\begin{pmatrix} & & & 0 \\ & \mathbf{0}_{K\times K} & & \vdots \\ & & & 0 \\ 0 & \cdots & 0 & \Sigma_{nT,B}^{K+1,K+1} \end{pmatrix},$$

where $\Sigma_{nT,A}^{k_1,K+1}=\frac{1}{nT}\sum_{i=1}^{nT}[vec\left(M_{\tilde{\Gamma}_n}R_n\bar{Z}_{k_1}M_{F_T}\right)]_iH_{K+1,ii}$ *for* $k_1=1,\cdots,K$; $\Sigma_{nT,B}^{K+1,K+1}=\frac{1}{nT}\sum_{i=1}^{nT}\left(H_{K+1,ii}\right)^2$; *and* $H_{K+1}=I_T\otimes\left(M_{\tilde{\Gamma}_n}\tilde{G}_n-M_{\tilde{\Gamma}_n}\tilde{W}_nP_{\tilde{\Gamma}_n}-\frac{1}{n}tr\left(\tilde{G}_n\right)M_{\tilde{\Gamma}_n}\right).$

S4. Additional Monte Carlo Results

A. Additional Monte Carlo Simulations

Tables S.1 and S.2 report the Monte Carlo simulations for the QML estimators and the bias corrected QML estimators under two alternative DGPs. When there is no spatial effect in disturbances (θ_0^c), biases are smaller and average biases of the original QML estimators and the bias-corrected estimators are similar. However, the bias-corrected estimators have better CPs, especially for α. The case with no contemporaneous spatial interaction (θ_0^d) is similar to the baseline case. The QML estimators have noticeable biases for α which are reduced by the bias correction.

B. Estimation under Misspecification

Tables S.3 and S.4 provide the Monte Carlo results for the misspecified models where either the interactive effects or the spatial effects are present but ignored. When interactive effects are ignored, the estimates

Table S.1: Performance of the QML and Bias Corrected Estimators, θ_0^c

n	T			$\hat{\lambda}$	$\hat{\gamma}$	$\hat{\rho}$	$\hat{\alpha}$	$\hat{\beta}_1$	$\hat{\beta}_2$
25	25	$\hat{\theta}_{nT}$	Bias	-0.00071	-0.00134	0.00115	-0.00194	0.00175	-0.00101
25	25	$\hat{\theta}_{nT}$	CP	0.912	0.910	0.906	0.887	0.931	0.915
25	25	$\hat{\theta}^c_{nT}$	Bias	-0.00071	-0.00124	0.00107	-0.00226	0.00180	-0.00097
25	25	$\hat{\theta}^c_{nT}$	CP	0.913	0.918	0.905	0.907	0.934	0.915
25	49	$\hat{\theta}_{nT}$	Bias	-0.00013	-0.00062	0.00010	-0.00140	-0.00024	-0.00055
25	49	$\hat{\theta}_{nT}$	CP	0.924	0.934	0.939	0.868	0.922	0.917
25	49	$\hat{\theta}^c_{nT}$	Bias	-0.00012	-0.00062	0.00009	-0.00093	-0.00022	-0.00055
25	49	$\hat{\theta}^c_{nT}$	CP	0.925	0.937	0.939	0.916	0.925	0.917
25	81	$\hat{\theta}_{nT}$	Bias	0.00024	-0.00037	0.00008	-0.00107	0.00058	-0.00099
25	81	$\hat{\theta}_{nT}$	CP	0.928	0.927	0.930	0.874	0.934	0.939
25	81	$\hat{\theta}^c_{nT}$	Bias	0.00031	-0.00039	0.00003	-0.00109	0.00057	-0.00099
25	81	$\hat{\theta}^c_{nT}$	CP	0.931	0.929	0.932	0.930	0.937	0.939
49	25	$\hat{\theta}_{nT}$	Bias	-0.00095	-0.00126	0.00159	-0.00041	-0.00004	-0.00055
49	25	$\hat{\theta}_{nT}$	CP	0.919	0.927	0.933	0.914	0.919	0.942
49	25	$\hat{\theta}^c_{nT}$	Bias	-0.00096	-0.00114	0.00149	-0.00021	-0.00001	-0.00052
49	25	$\hat{\theta}^c_{nT}$	CP	0.921	0.930	0.938	0.923	0.917	0.942
49	49	$\hat{\theta}_{nT}$	Bias	-0.00003	-0.00016	0.00011	0.00235	-0.00053	0.00012
49	49	$\hat{\theta}_{nT}$	CP	0.941	0.936	0.934	0.918	0.945	0.932
49	49	$\hat{\theta}^c_{nT}$	Bias	-0.00000	-0.00017	0.00009	0.00235	-0.0005	0.00012
49	49	$\hat{\theta}^c_{nT}$	CP	0.941	0.944	0.936	0.934	0.944	0.933
49	81	$\hat{\theta}_{nT}$	Bias	-0.00055	0.00004	0.00025	0.00004	0.00073	0.00085
49	81	$\hat{\theta}_{nT}$	CP	0.941	0.938	0.940	0.919	0.945	0.933
49	81	$\hat{\theta}^c_{nT}$	Bias	-0.00054	0.00003	0.00026	-0.00013	0.00073	0.00085
49	81	$\hat{\theta}^c_{nT}$	CP	0.944	0.939	0.941	0.949	0.945	0.933
81	25	$\hat{\theta}_{nT}$	Bias	-0.00059	0.00020	0.00004	-0.00023	0.00039	-0.00071
81	25	$\hat{\theta}_{nT}$	CP	0.933	0.938	0.930	0.909	0.921	0.935
81	25	$\hat{\theta}^c_{nT}$	Bias	-0.00060	0.00022	0.00003	-0.00016	0.00044	-0.00069
81	25	$\hat{\theta}^c_{nT}$	CP	0.934	0.945	0.935	0.914	0.925	0.936
81	49	$\hat{\theta}_{nT}$	Bias	0.00002	0.00029	-0.00060	-0.00122	0.00021	0.00062
81	49	$\hat{\theta}_{nT}$	CP	0.940	0.939	0.944	0.932	0.933	0.934
81	49	$\hat{\theta}^c_{nT}$	Bias	0.00002	0.00027	-0.00058	-0.00128	0.00023	0.00062
81	49	$\hat{\theta}^c_{nT}$	CP	0.940	0.944	0.947	0.936	0.933	0.934
81	81	$\hat{\theta}_{nT}$	Bias	-0.00037	-0.00014	0.00034	0.00048	0.00032	-0.00004
81	81	$\hat{\theta}_{nT}$	CP	0.926	0.944	0.920	0.933	0.942	0.950
81	81	$\hat{\theta}^c_{nT}$	Bias	-0.00037	-0.00013	0.00033	0.00049	0.00032	-0.00004
81	81	$\hat{\theta}^c_{nT}$	CP	0.927	0.946	0.919	0.939	0.942	0.950

The true parameters are $\theta_0^c = (0.3, 0.3, 0.3, 0, 1, 1)$ with $\theta = (\lambda, \gamma, \rho, \alpha, \beta_1, \beta_2)$. $\vartheta = 1$. $\hat{\theta}_{nT}$ is the QML estimator and $\hat{\theta}^c_{nT}$ is the bias corrected estimator from Theorem 4.

Table S.2: Performance of the QML and Bias Corrected Estimators, θ_0^d

n	T			$\hat{\lambda}$	$\hat{\gamma}$	$\hat{\rho}$	$\hat{\alpha}$	$\hat{\beta}_1$	$\hat{\beta}_2$
25	25	$\hat{\theta}_{nT}$	Bias	0.00114	-0.00172	-0.00084	0.02021	0.00156	-0.00127
			CP	0.912	0.922	0.918	0.863	0.934	0.921
		$\hat{\theta}_{nT}^c$	Bias	0.00049	-0.00153	-0.00061	0.00176	0.00155	-0.00136
			CP	0.909	0.927	0.916	0.904	0.932	0.922
	49	$\hat{\theta}_{nT}$	Bias	0.00097	-0.00119	-0.00142	0.02070	-0.00026	-0.00027
			CP	0.934	0.924	0.931	0.874	0.914	0.921
		$\hat{\theta}_{nT}^c$	Bias	0.00044	-0.00111	-0.00123	0.00238	-0.00030	-0.00036
			CP	0.931	0.927	0.935	0.928	0.914	0.921
	81	$\hat{\theta}_{nT}$	Bias	0.00195	-0.00048	-0.00045	0.02102	0.00067	-0.00068
			CP	0.935	0.928	0.911	0.834	0.933	0.930
		$\hat{\theta}_{nT}^c$	Bias	0.00132	-0.00039	-0.00020	0.00200	0.00063	-0.00078
			CP	0.934	0.928	0.911	0.934	0.933	0.931
49	25	$\hat{\theta}_{nT}$	Bias	-0.00008	-0.00141	0.00014	0.00974	-0.00040	-0.00082
			CP	0.917	0.927	0.948	0.917	0.923	0.939
		$\hat{\theta}_{nT}^c$	Bias	-0.00035	-0.00126	0.00020	0.00024	-0.00038	-0.00085
			CP	0.915	0.930	0.948	0.927	0.923	0.938
	49	$\hat{\theta}_{nT}$	Bias	0.00069	-0.00030	-0.00049	0.01269	-0.00054	0.00026
			CP	0.944	0.936	0.927	0.895	0.946	0.935
		$\hat{\theta}_{nT}^c$	Bias	0.00041	-0.00026	-0.00039	0.00282	-0.00055	0.00021
			CP	0.945	0.942	0.929	0.937	0.949	0.935
	81	$\hat{\theta}_{nT}$	Bias	-0.00005	-0.00015	-0.00029	0.01075	0.00062	0.00089
			CP	0.946	0.936	0.950	0.900	0.942	0.939
		$\hat{\theta}_{nT}^c$	Bias	-0.00032	-0.00012	-0.00018	0.00059	0.00061	0.00084
			CP	0.944	0.937	0.950	0.949	0.942	0.939
81	25	$\hat{\theta}_{nT}$	Bias	-0.00025	-0.00006	-0.00088	0.00576	0.00033	-0.00077
			CP	0.938	0.925	0.941	0.905	0.923	0.943
		$\hat{\theta}_{nT}^c$	Bias	-0.00036	-0.00003	-0.00087	-0.00016	0.00037	-0.00077
			CP	0.939	0.940	0.943	0.923	0.925	0.943
	49	$\hat{\theta}_{nT}$	Bias	0.00038	0.00006	-0.00075	0.00496	0.00008	0.00063
			CP	0.940	0.937	0.945	0.924	0.932	0.927
		$\hat{\theta}_{nT}^c$	Bias	0.00025	0.00006	-0.00070	-0.00120	0.00009	0.00061
			CP	0.941	0.945	0.945	0.941	0.932	0.927
	81	$\hat{\theta}_{nT}$	Bias	-0.00014	-0.00024	0.00011	0.00687	0.00017	-0.00005
			CP	0.926	0.944	0.930	0.912	0.947	0.955
		$\hat{\theta}_{nT}^c$	Bias	-0.00029	-0.00020	0.00017	0.00072	0.00016	-0.00008
			CP	0.927	0.947	0.926	0.937	0.947	0.955

The true parameters are $\theta_0^d = (0, 0.3, 0.3, 0.3, 1, 1)$ with $\theta = (\lambda, \gamma, \rho, \alpha, \beta_1, \beta_2)$. $\vartheta = 1$. $\hat{\theta}_{nT}$ is the QML estimator and $\hat{\theta}_{nT}^c$ is the bias corrected estimator from Theorem 4.

for λ and β_1 have large biases, which is expected because $X_{nt,1}$ and the spatially generated regressor correlate with the interactive effects. When spatial effects are ignored, the estimates for other spatial features (ρ and α) have large biases. The biases persist in large samples.

The two experiments in the previous section show that the estimators from our joint modeling of all the spatial features and interactive effects have satisfactory finite sample properties even if some features are not present in the true DGP. Given that large biases likely occur when these features are incorrectly omitted, it therefore may be preferable in empirical applications that researchers not impose a priori assumptions on which features are present.

C. Estimation with Redundant Factors

Tables S.5 and S.6 extend the simulations in Table 3 by allowing more redundant factors. The coverage probability deteriorates with more redundant factors in small samples, indicating the importance to include a correct number of factors. Note that the biases and CP improve in large samples (e.g., $n = 81$ and $T = 81$), and this is consistent with the results of Moon and Weidner (2015b) that the limiting distribution is invariant to the number of over-specified factors. Therefore for valid inference in small sample, it is important that a correct number of factors is chosen. The estimators are less sensitive to redundant factors in large samples.

Table S.3: Estimation under Misspecification: Ignore Interactive Effects

θ_0	n	T		$\hat{\lambda}$	$\hat{\gamma}$	$\hat{\rho}$	$\hat{\alpha}$	$\hat{\beta}_1$	$\hat{\beta}_2$
θ_0^a	25	25	Bias	-0.03452	-0.00831	0.02095	-0.15025	0.20509	-0.00033
			RMSE	0.07176	0.04795	0.06317	0.18686	0.22366	0.07058
	25	49	Bias	-0.03589	-0.00913	0.02439	-0.14125	0.20510	0.00062
			RMSE	0.06441	0.03567	0.05431	0.17147	0.21661	0.05054
	25	81	Bias	-0.03607	-0.00828	0.02488	-0.14681	0.20319	-0.00189
			RMSE	0.05898	0.02895	0.04647	0.17475	0.21087	0.03798
	49	25	Bias	-0.03630	-0.00794	0.02385	-0.14074	0.20293	-0.00034
			RMSE	0.06080	0.04566	0.05519	0.16397	0.21481	0.04711
	49	49	Bias	-0.03717	-0.00838	0.02584	-0.14068	0.20492	-0.00081
			RMSE	0.05479	0.03237	0.04611	0.15864	0.21135	0.03455
	49	81	Bias	-0.03957	-0.00828	0.02803	-0.14417	0.20605	0.00184
			RMSE	0.05320	0.02622	0.04279	0.16175	0.21066	0.02690
	81	25	Bias	-0.03525	-0.00887	0.02396	-0.14474	0.20560	-0.00272
			RMSE	0.05303	0.04167	0.04946	0.16059	0.21329	0.03922
	81	49	Bias	-0.03937	-0.00960	0.02865	-0.14412	0.20858	0.00024
			RMSE	0.05123	0.03171	0.04449	0.15638	0.21332	0.0272
	81	81	Bias	-0.04006	-0.00648	0.02702	-0.14037	0.20658	-0.00083
			RMSE	0.04974	0.02527	0.03940	0.15144	0.20989	0.02112
θ_0^b	25	25	Bias	-0.03708	-0.01462	-0.00633	-0.13768	0.18761	-0.00809
			RMSE	0.07963	0.04810	0.04321	0.17305	0.20680	0.07224
	25	49	Bias	-0.03620	-0.01556	-0.00628	-0.13295	0.18750	-0.00681
			RMSE	0.07119	0.03670	0.03365	0.16145	0.19913	0.05339
	25	81	Bias	-0.04249	-0.01327	-0.00176	-0.13240	0.18548	-0.01084
			RMSE	0.07079	0.02984	0.02738	0.15771	0.19338	0.04140
	49	25	Bias	-0.04032	-0.01333	-0.00359	-0.13060	0.18568	-0.00832
			RMSE	0.06698	0.04566	0.03308	0.15256	0.19751	0.04861
	49	49	Bias	-0.04341	-0.01360	-0.00266	-0.12870	0.18745	-0.00956
			RMSE	0.06526	0.03324	0.02549	0.14635	0.19412	0.03735
	49	81	Bias	-0.04481	-0.01372	-0.00148	-0.13376	0.18804	-0.00706
			RMSE	0.06236	0.02758	0.02135	0.15044	0.19264	0.02935
	81	25	Bias	-0.03905	-0.01472	-0.00614	-0.13634	0.18920	-0.00999
			RMSE	0.05898	0.04208	0.02803	0.15075	0.19669	0.04124
	81	49	Bias	-0.04396	-0.01509	-0.00258	-0.13541	0.19110	-0.00800
			RMSE	0.05850	0.03274	0.02055	0.14647	0.19569	0.02893
	81	81	Bias	-0.04435	-0.01188	-0.00126	-0.13243	0.18910	-0.00920
			RMSE	0.05656	0.02624	0.01781	0.14281	0.19236	0.02349

$\theta_0^a = (0.3, 0.3, 0.3, 0.3, 1, 1)$, $\theta_0^b = (-0.3, 0.3, 0.3, 0.3, 1, 1)$, and $\theta = (\lambda, \gamma, \rho, \alpha, \beta_1, \beta_2)$. $\vartheta = 1$. The DGP is the same as in Table 1 with both spatial interaction and interactive effects. The estimation uses the objective function of an SAR model with spatial error.

Table S.4: Estimation under Misspecification: Ignore Spatial Effects

θ_0	n	T		$\hat{\gamma}$	$\hat{\rho}$	$\hat{\alpha}$	$\hat{\beta}_1$	$\hat{\beta}_2$
θ_0^a	25	25	Bias	0.07626	0.17745	0.29915	-0.06485	-0.07442
			RMSE	0.07915	0.17943	0.30306	0.07471	0.08670
	25	49	Bias	0.07804	0.17980	0.30189	-0.06704	-0.07484
			RMSE	0.07938	0.18078	0.30384	0.07225	0.08099
	25	81	Bias	0.07910	0.18079	0.30447	-0.06774	-0.07604
			RMSE	0.08006	0.18151	0.30575	0.07108	0.07959
	49	25	Bias	0.07361	0.18294	0.2954	-0.06216	-0.07032
			RMSE	0.07504	0.18388	0.29754	0.06744	0.07656
	49	49	Bias	0.07593	0.18417	0.30016	-0.06275	-0.07061
			RMSE	0.07660	0.18467	0.30118	0.06529	0.07381
	49	81	Bias	0.07623	0.18519	0.30063	-0.06150	-0.06988
			RMSE	0.07668	0.18554	0.30132	0.06333	0.07179
	81	25	Bias	0.07407	0.18559	0.29531	-0.05922	-0.06850
			RMSE	0.07492	0.18614	0.29665	0.06254	0.07226
	81	49	Bias	0.07460	0.18629	0.29739	-0.05944	-0.06743
			RMSE	0.07502	0.18663	0.29807	0.06123	0.06954
	81	81	Bias	0.07411	0.18702	0.29972	-0.05948	-0.06845
			RMSE	0.07439	0.18726	0.30016	0.06057	0.06956
θ_0^b	25	25	Bias	-0.03563	-0.10417	-0.29197	-0.01949	0.02364
			RMSE	0.04216	0.10968	0.30396	0.04814	0.05599
	25	49	Bias	-0.03609	-0.10817	-0.29384	-0.02452	0.02478
			RMSE	0.03955	0.11071	0.30075	0.04114	0.04356
	25	81	Bias	-0.03515	-0.10742	-0.29229	-0.02363	0.02444
			RMSE	0.03733	0.10926	0.29721	0.03581	0.03647
	49	25	Bias	-0.03459	-0.10696	-0.31113	-0.02164	0.02253
			RMSE	0.03822	0.10989	0.31593	0.03878	0.04058
	49	49	Bias	-0.03397	-0.10964	-0.30857	-0.02486	0.02415
			RMSE	0.03565	0.11106	0.31095	0.03353	0.03420
	49	81	Bias	-0.03365	-0.11154	-0.31353	-0.02637	0.02349
			RMSE	0.03473	0.11249	0.31507	0.03283	0.03010
	81	25	Bias	-0.03270	-0.11188	-0.31923	-0.02301	0.02134
			RMSE	0.03500	0.11358	0.32197	0.03319	0.03393
	81	49	Bias	-0.03280	-0.11230	-0.32016	-0.02551	0.02230
			RMSE	0.03387	0.11324	0.32152	0.03082	0.02949
	81	81	Bias	-0.03295	-0.11198	-0.31978	-0.02657	0.02195
			RMSE	0.03361	0.11262	0.32065	0.03061	0.02593

$\theta_0^a = (0.3, 0.3, 0.3, 0.3, 1, 1)$, $\theta_0^b = (-0.3, 0.3, 0.3, 0.3, 1, 1)$, and $\theta = (\lambda, \gamma, \rho, \alpha, \beta_1, \beta_2)$. $\vartheta = 1$. The DGP is the same as in Table 1 with both spatial interaction and interactive effects. The estimation here uses otherwise the same objective function but ignores the spatial interaction $\lambda W_n Y_{nt}$.

Table S.5: Performance of the Bias Corrected Estimator When the Number of Factors is Overspecified by 2

θ_0	n	T		$\hat{\lambda}$	$\hat{\gamma}$	$\hat{\rho}$	$\hat{\alpha}$	$\hat{\beta}_1$	$\hat{\beta}_2$
θ_0^a	25	25	Bias	0.00089	-0.00252	-0.00003	0.00435	0.00237	-0.00114
			CP	0.825	0.819	0.813	0.705	0.855	0.840
	25	49	Bias	0.00053	-0.00079	-0.00093	0.00810	0.00021	0.00029
			CP	0.865	0.860	0.865	0.738	0.859	0.864
	25	81	Bias	0.00046	-0.0005	-0.00021	0.01040	0.00070	-0.00057
			CP	0.867	0.885	0.879	0.729	0.894	0.889
	49	25	Bias	-0.00001	-0.00121	0.00037	-0.00091	-0.00067	-0.00179
			CP	0.878	0.884	0.860	0.810	0.858	0.877
	49	49	Bias	0.00017	-0.00060	0.00010	0.00584	-0.00071	0.00026
			CP	0.905	0.901	0.908	0.883	0.903	0.898
	49	81	Bias	-0.00007	-0.00006	-0.00025	0.00342	0.00079	0.00062
			CP	0.920	0.901	0.920	0.879	0.919	0.921
	81	25	Bias	0.00008	-0.00016	-0.00061	-0.00127	0.00014	-0.00110
			CP	0.904	0.874	0.888	0.842	0.885	0.896
	81	49	Bias	0.00021	0.00027	-0.00085	-0.00111	0.00041	0.00044
			CP	0.910	0.907	0.914	0.920	0.900	0.906
	81	81	Bias	-0.00043	-0.00013	0.00033	0.00101	0.00025	-0.00001
			CP	0.914	0.923	0.905	0.900	0.923	0.935
θ_0^b	25	25	Bias	0.00532	-0.00345	-0.00302	0.00559	0.00251	-0.00041
			CP	0.813	0.823	0.821	0.708	0.845	0.849
	25	49	Bias	0.00279	-0.00178	-0.00279	0.00790	0.00054	0.00088
			CP	0.857	0.864	0.853	0.730	0.842	0.859
	25	81	Bias	0.00320	-0.00062	-0.00074	0.00766	0.00118	0.00010
			CP	0.858	0.876	0.872	0.714	0.879	0.892
	49	25	Bias	0.00210	-0.00140	-0.00069	-0.00008	-0.00042	-0.00131
			CP	0.879	0.874	0.886	0.832	0.864	0.878
	49	49	Bias	0.00117	-0.00063	-0.00046	0.00616	-0.00061	0.00052
			CP	0.911	0.907	0.902	0.880	0.911	0.899
	49	81	Bias	0.00031	-0.00027	-0.00061	0.00372	0.00077	0.00067
			CP	0.925	0.899	0.921	0.885	0.920	0.920
	81	25	Bias	0.00064	-0.00056	-0.00137	-0.00027	0.00012	-0.00096
			CP	0.890	0.878	0.892	0.859	0.887	0.885
	81	49	Bias	0.00056	0.00019	-0.00052	-0.00073	0.00036	0.00056
			CP	0.904	0.904	0.913	0.919	0.903	0.900
	81	81	Bias	-0.00018	-0.00015	0.00024	0.00117	0.00015	-0.00003
			CP	0.907	0.912	0.909	0.906	0.922	0.933

$\theta_0^a = (0.3, 0.3, 0.3, 0.3, 1, 1)$, $\theta_0^b = (-0.3, 0.3, 0.3, 0.3, 1, 1)$, and $\theta = (\lambda, \gamma, \rho, \alpha, \beta_1, \beta_2)$. $\vartheta = 1$. The DGP is the same as described in the text. The true number of factors is 2 and the estimation assumes 4 factors. $\hat{\theta}$ is the bias corrected QML estimator assuming 4 factors.

Table S.6: Performance of the Bias Corrected Estimator When the Number of Factors is Overspecified by 3

θ_0	n	T		$\hat{\lambda}$	$\hat{\gamma}$	$\hat{\rho}$	$\hat{\alpha}$	$\hat{\beta}_1$	$\hat{\beta}_2$
θ_0^a	25	25	Bias	-0.00156	-0.00234	0.00004	-0.00545	0.00093	-0.00288
			CP	0.763	0.748	0.764	0.563	0.787	0.793
	25	49	Bias	-0.00027	-0.00078	-0.00062	0.00418	-0.00037	0.00024
			CP	0.826	0.827	0.826	0.565	0.814	0.844
	25	81	Bias	-0.00035	-0.00053	0.00034	0.00008	0.00079	-0.00069
			CP	0.827	0.854	0.844	0.454	0.853	0.869
	49	25	Bias	-0.00022	-0.00114	0.00041	-0.00106	-0.00095	-0.00192
			CP	0.840	0.840	0.829	0.764	0.813	0.846
	49	49	Bias	0.00038	-0.00059	-0.00017	0.00601	-0.00070	0.00007
			CP	0.886	0.877	0.883	0.835	0.880	0.887
	49	81	Bias	-0.00002	-0.00003	-0.00027	0.00512	0.00070	0.00052
			CP	0.903	0.886	0.905	0.837	0.908	0.898
	81	25	Bias	0.00015	-0.00031	-0.00068	-0.00125	-0.00019	-0.00129
			CP	0.868	0.806	0.854	0.818	0.838	0.862
	81	49	Bias	0.00018	0.00027	-0.00091	-0.00081	0.00046	0.00052
			CP	0.903	0.891	0.912	0.895	0.881	0.888
	81	81	Bias	-0.00054	-0.00004	0.00032	0.00165	0.00027	-0.00000
			CP	0.904	0.911	0.899	0.882	0.914	0.925
θ_0^b	25	25	Bias	0.00743	-0.00313	-0.00398	-0.00422	0.00216	-0.00050
			CP	0.746	0.753	0.762	0.537	0.759	0.786
	25	49	Bias	0.00471	-0.00180	-0.00366	-0.00690	0.00028	0.00111
			CP	0.812	0.824	0.829	0.538	0.807	0.841
	25	81	Bias	0.00567	-0.00083	-0.00137	-0.01715	0.00205	0.00068
			CP	0.824	0.853	0.844	0.397	0.842	0.868
	49	25	Bias	0.00240	-0.00142	-0.00081	-0.00014	-0.00072	-0.00140
			CP	0.833	0.822	0.841	0.767	0.819	0.841
	49	49	Bias	0.00160	-0.00070	-0.00082	0.00660	-0.00047	0.00045
			CP	0.896	0.877	0.875	0.829	0.885	0.889
	49	81	Bias	0.00052	-0.00019	-0.00019	0.00539	0.00073	0.00064
			CP	0.900	0.892	0.902	0.845	0.909	0.911
	81	25	Bias	0.00075	-0.00080	-0.00161	-0.00018	-0.00018	-0.00112
			CP	0.868	0.824	0.853	0.822	0.854	0.865
	81	49	Bias	0.00047	0.00015	-0.00058	-0.00037	0.00036	0.00061
			CP	0.892	0.892	0.895	0.894	0.888	0.887
	81	81	Bias	-0.00032	-0.00008	0.00023	0.00183	0.00012	-0.00007
			CP	0.893	0.903	0.896	0.882	0.914	0.921

$\theta_0^a = (0.3, 0.3, 0.3, 0.3, 1, 1)$, $\theta_0^b = (-0.3, 0.3, 0.3, 0.3, 1, 1)$, and $\theta = (\lambda, \gamma, \rho, \alpha, \beta_1, \beta_2)$. $\vartheta = 1$. The DGP is the same as described in the text. The true number of factors is 2 and the estimation assumes 5 factors. $\hat{\theta}$ is the bias corrected QML estimator assuming 5 factors.