Quasi-Maximum Likelihood Estimators For Spatial Dynamic Panel Data With Fixed Effects When Both n and T Are Large*

Jihai Yu, Robert de Jong, Lung-fei Lee

Department of Economics

The Ohio State University

November, 2006

Abstract

This paper investigates the asymptotic properties of quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both the number of individuals n and the number of time periods T are large. We consider the case where T is asymptotically large relative to n, the case where T is asymptotically proportional to n, and the case where n is asymptotically large relative to T. In the case where T is asymptotically large relative to n, the estimators are $\sqrt{nT}$ consistent and asymptotically normal, with the limit distribution centered around 0. When n is asymptotically proportional to T, the estimators are $\sqrt{nT}$ consistent and asymptotically normal, but the limit distribution is not centered around 0; and when n is large relative to T, the estimators are consistent with rate T, and have a degenerate limit distribution. The estimators of the fixed effects are $\sqrt{T}$ consistent and asymptotically normal. We also propose a bias correction for our estimators. We show that when T grows faster than $n^{1/3}$, the correction will asymptotically eliminate the bias and yield a centered confidence interval.

JEL classification: C13; C23

Keywords: Spatial autoregression, Dynamic panels, Fixed effects, Maximum likelihood estimation, Quasi-Maximum likelihood estimation, Bias correction

*We would like to thank participants of the Econometrics Seminar at The Ohio State University, the International Workshop on Spatial Econometrics and Statistics 2006 (Rome, Italy) and the Far Eastern Meeting of Econometric Society 2006 (Beijing, China) for helpful comments. Lee acknowledges financial support from NSF under Grant No. SES-0519204 and research assistantship support from the Department of Economics in The Ohio State University.

1 Introduction

Spatial econometrics deals with the spatial interactions of economic units in cross-sectional and/or panel data, and has received much attention recently. The spatial autoregressive (SAR) model by Cliff and Ord (1973) has received the most attention. It extends autocorrelation in time series to spatial dimensions. Early development in estimation and testing is summarized in Paelinck and Klaassen (1979), Doreian (1980), Anselin (1988,1992), Haining (1990), Kelejian and Robinson (1993), Cressie (1993), Anselin and Florax (1995), Anselin and Rey (1997), and Anselin and Bera (1998).

For the dynamic panel version of the SAR model, when the time dimension T is fixed, we are likely to encounter the "incidental parameters" problem discussed in Neyman and Scott (1948). This is because the introduction of fixed effects increases the number of parameters to be estimated, where the increase is the number of cross sectional units n. Furthermore, the dynamic nature of the model will give rise to the "initial conditions" problem (Hsiao (1986), section 4.2). Because of this "initial conditions" problem, the maximum likelihood (ML) estimator is inconsistent if T is finite for dynamic panel data.

When T goes to infinity, the estimators of dynamic panel data can be consistent. We can not only estimate the fixed effects[1], but also avoid the "initial conditions" problem as well as the "incidental parameters" problem. Recently, there is a growing literature on the estimation of dynamic panel data models when both n and T are large (see Phillips and Moon (1999), Hahn and Kuersteiner (2002), Hahn and Newey (2004), etc). To the best knowledge of the authors, there is little work done on testing and estimation of spatial correlation in panel data when both n and T are large, with the exception of Baltagi, Song and Koh (2003) and Kapoor, Kelejian and Prucha (2004). This paper investigates the properties of quasi-maximum likelihood (QML) estimator for spatial dynamic panel data with fixed effects when both n and T are large.

When $T/n \to 0$ or T and n are proportional asymptotically, there are asymptotic biases in the estimators of dynamic panel models. There is currently active ongoing research on bias correction procedures in order to improve the estimation, see e.g. Hahn and Kuersteiner (2002), Andrews and Guggenberger (2003), Hahn and Newey (2004), Bun and Carree (2005). For the spatial dynamic panel data model, we will face the same issue. In order to improve the estimation for our model, we investigate a possible bias correction procedure for our estimators. Therefore, our paper extends the current literature on bias correction for dynamic panel models to the spatial setting.

This paper is organized as follows. In Section 2, we introduce the model and explain our estimation

[1] The consistent estimation of fixed effects is of great interest if there are time invariant explanatory variables present in the model. In that situation, the effect of those time invariant variables can be recovered from a regression of estimated fixed effects on the time invariant explanatory variables (see Hausman and Taylor (1981)).

method, which is concentrated quasi-maximum likelihood estimation. Also, the law of large numbers and central limit theorem for our setting are developed. Section 3 establishes the consistency and asymptotic distribution of ML estimator and QML estimator. We also propose an analytical bias correction for our estimators. We show that when T grows faster than $n^{1/3}$, this correction will eliminate the bias and yield a centered confidence interval. Section 4 concludes the paper. Some useful lemmas and proofs are collected in the Appendix.

2 The Model and Concentrated Likelihood Function

2.1 The Model

The model considered in this paper is

$$Y_{nt} = \lambda_0 W_n Y_{nt} + \gamma_0 Y_{n,t-1} + \rho_0 W_n Y_{n,t-1} + X_{nt}\beta_0 + \mathbf{c}_{n0} + V_{nt}, \quad t = 1, 2, ..., T, \tag{2.1}$$

where $Y_{nt} = (y_{1t}, y_{2t}, ..., y_{nt})'$ and $V_{nt} = (v_{1t}, v_{2t}, ..., v_{nt})'$ are $n \times 1$ column vectors and v_{it} is *i.i.d.* across i and t with zero mean and variance σ_0^2, W_n is $n \times n$ spatial weights matrix which is predetermined and generates the spatial dependence between cross sectional units y_{it}, X_{nt} is an $n \times k_x$ matrix of nonstochastic regressors, and $\mathbf{c}_{n0}$ is $n \times 1$ column vector of fixed effects. Therefore, the total number of parameters in this model is equal to the number of individuals n plus the dimension of the common parameters $(\gamma, \rho, \beta', \lambda, \sigma^2)'$, which is $k_x + 4$.

Define $S_n \equiv S_n(\lambda_0) = I_n - \lambda_0 W_n$. Then, presuming S_n is invertible and denoting $A_n = S_n^{-1}(\gamma_0 I_n + \rho_0 W_n)$, (2.1) can be rewritten as

$$Y_{nt} = A_n Y_{n,t-1} + S_n^{-1} X_{nt}\beta_0 + S_n^{-1}\mathbf{c}_{n0} + S_n^{-1} V_{nt}. \tag{2.2}$$

Assuming that the infinite sums are well-defined, by continuous substitution of (2.2),

$$Y_{nt} = \sum_{h=0}^{\infty} A_n^h S_n^{-1}(\mathbf{c}_{n0} + X_{n,t-h}\beta_0 + V_{n,t-h}) = \mu_n + \mathcal{X}_{nt}\beta_0 + U_{nt} \tag{2.3}$$

where $\mu_n \equiv \sum_{h=0}^{\infty} A_n^h S_n^{-1}\mathbf{c}_{n0}$, $\mathcal{X}_{nt} \equiv \sum_{h=0}^{\infty} A_n^h S_n^{-1} X_{n,t-h}$, and $U_{nt} \equiv \sum_{h=0}^{\infty} A_n^h S_n^{-1} V_{n,t-h}$.

2.2 Concentrated Likelihood Function

Denote $\theta = (\delta', \lambda, \sigma^2)'$ where $\delta = (\gamma, \rho, \beta')'$. The likelihood function of (2.1) is

$$\ln L_{n,T}(\theta, \mathbf{c}_n) = -\frac{nT}{2}\ln 2\pi - \frac{nT}{2}\ln \sigma^2 + T\ln|S_n(\lambda)| - \frac{1}{2\sigma^2}\sum_{t=1}^{T} V_{nt}'(\zeta)V_{nt}(\zeta) \tag{2.4}$$

where $V_{nt}(\zeta) = S_n(\lambda)Y_{nt} - \gamma Y_{n,t-1} - \rho W_n Y_{n,t-1} - X_{nt}\beta - \mathbf{c}_n$ and $\zeta = (\delta', \lambda, \mathbf{c}_n')'$. Thus, $V_{nt} = V_{nt}(\zeta_0)$.

The QML estimators $\hat{\theta}_{nT}$ and $\hat{\mathbf{c}}_{nT}$ are the extreme estimators derived from the maximization of (2.4). When the V_{nt}'s are normally distributed, $\hat{\theta}_{nT}$ and $\hat{\mathbf{c}}_{nT}$ are the ML estimators; when the V_{nt}'s are not normally distributed, $\hat{\theta}_{nT}$ and $\hat{\mathbf{c}}_{nT}$ are QML estimators. As the number of parameters goes to infinity when n goes to infinity, it's convenient to concentrate $\mathbf{c}_n$ out and focus asymptotic analysis on the estimator of θ_0 via the concentrated likelihood function. For the concentrated likelihood function, the dimension of parameter space does not change as n and/or T increase.

For notational purpose, we define $\tilde{Y}_{nt} = Y_{nt} - \bar{Y}_{nT}$ and $\tilde{Y}_{n,t-1} = Y_{n,t-1} - \bar{Y}_{nT,-1}$ for $t = 1, 2, \cdots, T$ where $\bar{Y}_{nT} = \frac{1}{T}\sum_{t=1}^{T} Y_{nt}$ and $\bar{Y}_{nT,-1} = \frac{1}{T}\sum_{t=1}^{T} Y_{n,t-1}$. Similarly, we define $\tilde{X}_{nt} = X_{nt} - \bar{X}_{nT}$ and $\tilde{V}_{nt} = V_{nt} - \bar{V}_{nT}$.

Denote $Z_{nt} = (Y_{n,t-1}, W_n Y_{n,t-1}, X_{nt})$, then from (2.4), using the first order condition that $\frac{\partial \ln L_{n,T}(\theta, \mathbf{c}_n)}{\partial \mathbf{c}_n} = \frac{1}{\sigma^2}\sum_{t=1}^{T} V_{nt}(\zeta)$, the concentrated estimators of $\mathbf{c}_{n0}$ given θ are

$$\hat{\mathbf{c}}_{nT}(\theta) = \frac{1}{T}\sum_{t=1}^{T}(S_n(\lambda)Y_{nt} - Z_{nt}\delta)$$

and the concentrated likelihood is

$$\ln L_{n,T}(\theta) = -\frac{nT}{2}\ln 2\pi - \frac{nT}{2}\ln \sigma^2 + T\ln|S_n(\lambda)| - \frac{1}{2\sigma^2}\sum_{t=1}^{T} \tilde{V}_{nt}'(\zeta)\tilde{V}_{nt}(\zeta) \tag{2.5}$$

where $\tilde{V}_{nt}(\zeta) = S_n(\lambda)\tilde{Y}_{nt} - \tilde{Z}_{nt}\delta$ and $\tilde{Z}_{nt} = (Y_{n,t-1} - \bar{Y}_{nT,-1},\ W_n Y_{n,t-1} - W_n\bar{Y}_{nT,-1},\ X_{nt} - \bar{X}_{nT})$. The QML estimator $\hat{\theta}_{nT}$ maximizes the concentrated likelihood function (2.5), and the QML estimator of $\mathbf{c}_{n0}$ is $\hat{\mathbf{c}}_{nT}(\hat{\theta}_{nT})$.

For the concentrated likelihood (2.5), the first and second order derivatives of the likelihood function can be derived; see Equation (C.3) and (C.4) in Appendix C.

2.3 The Law of Large Numbers and Central Limit Theorem for Our Setting

Let D_{nt} be an $n \times 1$ vector with nonstochastic elements for all n and t and denote

$$\mathbb{U}_{nt} = \sum_{h=1}^{\infty} P_{nh}V_{n,t+1-h} \tag{2.6}$$

$$\mathbb{W}_{nt} = \sum_{h=1}^{\infty} Q_{nh}V_{n,t+1-h} \tag{2.7}$$

where $\{P_{nh}\}_{h=1}^{\infty}$and $\{Q_{nh}\}_{h=1}^{\infty}$ are sequences of $n \times n$ nonstochastic square matrices. Below, we state the law of large numbers and central limit theorem useful to derive the asymptotic properties of our estimators.

Assumption A1. The disturbances $\{v_{it}\}$, $i = 1, 2, ..., n$ and $t = 1, 2, ..., T$, are $i.i.d$ across i and t with zero mean, variance σ_0^2 and $E\,|v_{it}|^{4+\eta} < \infty$ for some $\eta > 0$.

Assumption A2. $P_{nh} = B_{1n}P_n^h$ and $Q_{nh} = B_{2n}Q_n^h$. Furthermore, the row and column sums of B_{1n}, B_{2n}, $\sum\limits_{h=1}^{\infty} abs(P_n^h)$ and $\sum\limits_{h=1}^{\infty} abs(Q_n^h)$ are bounded uniformly in n.[2]

Assumption A3. The elements of $n \times 1$ vector D_{nt} are nonstochastic and bounded, uniformly in n and t.

Assumption A4. n is a nondecreasing function of T.

Theorem 2.1 *Under Assumption A1, A2 and A4,*

$$\frac{1}{nT}\sum_{t=1}^{T}\mathbb{U}'_{nt}\mathbb{W}_{nt} - E\left(\frac{1}{nT}\sum_{t=1}^{T}\mathbb{U}'_{nt}\mathbb{W}_{nt}\right) = O_p\left(\frac{1}{\sqrt{nT}}\right),$$

$$\frac{1}{n}\bar{\mathbb{U}}'_{nT}\bar{\mathbb{W}}_{nT} - E\left(\frac{1}{n}\bar{\mathbb{U}}'_{nT}\bar{\mathbb{W}}_{nT}\right) = O_p\left(\frac{1}{\sqrt{nT^2}}\right),$$

and

$$\frac{1}{nT}\sum_{t=1}^{T}\tilde{\mathbb{U}}'_{nt}\tilde{\mathbb{W}}_{nt} - E\left(\frac{1}{nT}\sum_{t=1}^{T}\tilde{\mathbb{U}}'_{nt}\tilde{\mathbb{W}}_{nt}\right) = O_p\left(\frac{1}{\sqrt{nT}}\right)$$

where $E(\frac{1}{nT}\sum\limits_{t=1}^{T}\mathbb{U}'_{nt}\mathbb{W}_{nt})$ *is* $O(1)$, $E(\frac{1}{n}\bar{\mathbb{U}}'_{nT}\bar{\mathbb{W}}_{nT})$ *is* $O\left(\frac{1}{T}\right)$.

Proof. *This is Theorem A.7 in Appendix A.* ■

Theorem 2.2 *Under Assumption A1, A2, A3 and A4,* $\frac{1}{nT}\sum\limits_{t=1}^{T}\tilde{D}'_{nt}\tilde{\mathbb{U}}_{nt} = O_p\left(\frac{1}{\sqrt{nT}}\right)$.

Proof. *This is Lemma A.8 in Appendix A.* ■

We also need a different type of law of large numbers stated below.

Theorem 2.3 *Under Assumption A1, A2 and A4*

$$\sqrt{\frac{T}{n}}\bar{\mathbb{U}}'_{nT,-1}\bar{V}_{nT} - E\left(\sqrt{\frac{T}{n}}\bar{\mathbb{U}}'_{nT,-1}\bar{V}_{nT}\right) = O_p\left(\frac{1}{\sqrt{T}}\right) \tag{2.8}$$

where $E(\sqrt{\frac{T}{n}}\bar{\mathbb{U}}'_{nT,-1}\bar{V}_{nT}) = \sqrt{\frac{n}{T}}\frac{\sigma_0^2}{n}tr\left(\sum\limits_{h=1}^{\infty}P_{nh}\right) + O\left(\sqrt{\frac{n}{T^3}}\right)$ *and*

[2] We say the row and column sums of a (sequence of $n \times n$) matrix P_n are uniformly bounded in n if $\sup_{1\le i,j\le n, n\ge 1}\sum_{j=1}^{n}|p_{ij,n}| < \infty$ and $\sup_{1\le i,j\le n,n\ge 1}\sum_{i=1}^{n}|p_{ij,n}| < \infty$.

$$\sqrt{\frac{T}{n}}\bar{V}_{nT}'B_n\bar{V}_{nT} - E(\sqrt{\frac{T}{n}}\bar{V}_{nT}'B_n\bar{V}_{nT}) = O_p\left(\frac{1}{\sqrt{T}}\right) \tag{2.9}$$

where $E(\sqrt{\frac{T}{n}}\bar{V}_{nT}'B_n\bar{V}_{nT}) = \sqrt{\frac{n}{T}}\frac{\sigma_0^2}{n}tr(B_n)$.

Proof. *(2.8) is in Theorem A.11 and (2.9) is in Theorem A.9.* ∎

For the central limit theorem, consider the following form:

$$Q_{nT} = \sum_{t=1}^{T}\left(\mathbb{U}_{n,t-1}'V_{nt} + D_{nt}'V_{nt} + V_{nt}'B_nV_{nt} - \sigma_0^2 trB_n\right)$$

where B_n is a nonstochastic $n \times n$ symmetric matrix[3] and its row and column sums are bounded uniformly in n. Denote the mean and variance of Q_{nT} as $\mu_{Q_{nT}}$ and $\sigma_{Q_{nT}}^2$ respectively. We have $\mu_{Q_{nT}} = 0$ and

$$\sigma_{Q_{nT}}^2 = T\sigma_0^4 tr\left(\sum_{h=1}^{\infty}P_{nh}'P_{nh}\right) + \sigma_0^2\sum_{t=1}^{T}D_{nt}'D_{nt} + T\left(\left(\mu_4 - 3\sigma_0^4\right)\sum_{i=1}^{n}b_{n,ii}^2 + 2\sigma_0^4 tr(B_n^2)\right) + 2\mu_3\sum_{t=1}^{T}\sum_{i=1}^{n}d_{nti}b_{n,ii},$$

where $\mu_s = Ev_{it}^s$ for $s = 3, 4$, $b_{n,ii}$'s are diagonal elements of B_n and d_{nti} is the ith element of D_{nt}.

Theorem 2.4 *Assume that row and column sums of* B_n *are bounded uniformly in* n *and assume the sequence* $\frac{1}{nT}\sigma_{Q_{nT}}^2$ *is bounded away from zero. Then under Assumption A1, A2, A3, A4,*

$$\frac{Q_{nT}}{\sigma_{Q_{nT}}} \xrightarrow{d} N(0,1)$$

when $T \to \infty$.

Proof. *This is Theorem A.13 in Appendix A.* ∎

3 Quasi Maximum Likelihood Estimation

For our analysis of the asymptotic properties of estimators, we need the following assumptions:

Assumption 1. W_n is a constant spatial weights matrix and its diagonal elements satisfy $w_{n,ii} = 0$ for $i = 1, 2, \cdots, n$.

Assumption 2. The disturbances $\{v_{it}\}$, $i = 1, 2, ..., n$ and $t = 1, 2, ..., T$, are $i.i.d$ across i and t with zero mean, variance σ_0^2 and $E\left|v_{it}\right|^{4+\eta} < \infty$ for some $\eta > 0$.

Assumption 3. $S_n(\lambda)$ is invertible for all $\lambda \in \Lambda$. Furthermore, Λ is compact and the true parameter λ_0 is in the interior of Λ.

[3] The assumption that B_n is symmetric is maintained w.l.o.g. since $V_{nt}'B_nV_{nt} = V_{nt}'[(B_n + B_n')/2]V_{nt}$.

Assumption 4. The elements of X_{nt} are nonstochastic and bounded, uniformly in n and t.

Assumption 5. The row and column sums of W_n and $S_n^{-1}(\lambda)$ are bounded uniformly in n, also uniformly in $\lambda \in \Lambda$ for $S_n^{-1}(\lambda)$.

Assumption 6. The row and column sums of $\sum_{h=1}^{\infty} abs(A_n^h)$ are bounded uniformly in n, where $[abs(A_n)]_{ij} = |A_{n,ij}|$.

Assumption 7. n is a nondecreasing function of T.

Assumptions 1 is a standard normalization assumption in spatial econometrics, and Assumption 2 provides regularity assumptions for v_{it}. Assumption 3 guarantees that Equation (2.2) is valid. When exogenous variables X_{nt} are included in the model, it is convenient to assume that the exogenous regressors are uniformly bounded as in Assumption 4. Assumption 5 is also made by Kelejian and Prucha (1998, 2001). In many empirical applications, the rows of W_n sum to 1, which ensures that all the weights are between 0 and 1. The uniform boundedness of W_n and $S_n^{-1}(\lambda)$ is a condition that limits the spatial correlation to a manageable degree. Assumption 6 is the absolute summability condition and row/column sum boundedness condition, which will play an important role to derive asymptotic properties of QML estimators. This assumption is essential for the paper because it limits the dependence between time series and between cross sectional units. In order to justify the absolute summability of A_n in Equation (2.3) and Assumption 6, a sufficient condition is $\|A_n\| < 1$ for any matrix norm (see Horn and Johnson (1985), Corollary 5.6.16) that satisfies $\|A_n\| = \|abs\,(A_n)\|$. When $\|A_n\| < 1$, $\sum_{h=0}^{\infty} A_n^h$ exists and can be defined as $(I_n - A_n)^{-1}$. Assumption 7 allows two cases: (i) $n \to \infty$ as $T \to \infty$; (ii) n is fixed as $T \to \infty$. Because (ii) is similar to a vector autoregressive (VAR) model, our main interest is in (i). If Assumption 7 holds, then we say that $n, T \to \infty$ simultaneously.

3.1 Consistency of the Concentrated Estimator $\hat{\theta}_{nT}$

For the log likelihood function (2.5) divided by the sample size nT, we have corresponding $Q_{n,T}(\theta) = E \max_{\mathbf{c}_n} \frac{1}{nT} \ln L_{n,T}(\theta, \mathbf{c}_n)$. Hence,

$$Q_{n,T}(\theta) = \frac{1}{nT} E \ln L_{n,T}(\theta) = -\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln \sigma^2 + \frac{1}{n}\ln |S_n(\lambda)| - \frac{1}{2\sigma^2} E \frac{1}{nT}\sum_{t=1}^{T} \tilde{V}'_{nt}(\zeta)\tilde{V}_{nt}(\zeta). \tag{3.1}$$

To get the consistency proof, we need the following uniform convergence result.

Claim 3.1 *Let Θ be any compact parameter space. Then under Assumption 1-7, $\frac{1}{nT}\ln L_{n,T}(\theta) - Q_{n,T}(\theta) \xrightarrow{p} 0$ uniformly in $\theta \in \Theta$ and $Q_{n,T}(\theta)$ is uniformly equicontinuous for $\theta \in \Theta$.*

Proof. *See Appendix D.1.* ■

For identification, if the information matrix $-E\left(\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\theta_0)}{\partial\theta\partial\theta'}\right)$ is nonsingular and $-E\left(\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\theta)}{\partial\theta\partial\theta'}\right)$ has full rank for any θ in some neighborhood $N(\theta_0)$ of θ_0, the parameters are locally identified (see Rothenberg (1971)). To get the information matrix, denote $G_n \equiv W_n S_n^{-1}$, then from Appendix C, the information matrix, which will be denoted by $\Sigma_{\theta_0,nT}$, is Equation (C.6) in Appendix C. Denote $\mathcal{H}_{nT} = \frac{1}{nT}\sum_{t=1}^{T}(\tilde{Z}_{nt}, G_n\tilde{Z}_{nt}\delta_0)'(\tilde{Z}_{nt}, G_n\tilde{Z}_{nt}\delta_0)$, we can see that from Equation (C.6),

$$\Sigma_{\theta_0,nT} = \begin{pmatrix} E\mathcal{H}_{nT} & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{n}\left[tr(G_n'G_n) + tr(G_n^2)\right] & \frac{1}{\sigma_0^2 n}tr(G_n) \\ 0 & \frac{1}{\sigma_0^2 n}tr(G_n) & \frac{1}{2\sigma_0^4} \end{pmatrix} + O\left(\frac{1}{T}\right), \tag{3.2}$$

which is nonsingular if $E\mathcal{H}_{nT}$ is nonsingular or $\frac{1}{n}(trG_n'G_n + trG_n^2 - \frac{2(trG_n)^2}{n})$ is nonzero[4]. Also, its rank does not change in a small neighborhood of θ_0 (see Equation (C.10)).

When $\lim_{T\to\infty} E\mathcal{H}_{nT}$ is nonsingular, we can get the global identification of the parameters.

Theorem 3.2 *Under Assumption 1-7, if* $\lim_{T\to\infty} E\mathcal{H}_{nT}$ *is nonsingular,* θ_0 *is globally identified and* $\hat{\theta}_{nT} \xrightarrow{p} \theta_0$.

Proof. *See Appendix D.3.* ■

When $\lim_{T\to\infty} E\mathcal{H}_{nT}$ is singular, global identification can still be obtained from the following theorem. Denote $\sigma_n^2(\lambda) = \frac{\sigma_0^2}{n}tr(S_n'^{-1}S_n'(\lambda)S_n(\lambda)S_n^{-1})$.

Theorem 3.3 *Under Assumption 1-7,* θ_0 *is globally identified if* $\lim_{n\to\infty}\left(\frac{1}{n}\ln\left|\sigma_0^2 S_n^{-1\prime}S_n^{-1}\right| - \frac{1}{n}\ln\left|\sigma_n^2(\lambda)S_n^{-1}(\lambda)'S_n^{-1}(\lambda)\right|\right) \neq 0$ *for* $\lambda \neq \lambda_0$.[5]*, and if this condition holds,* $\hat{\theta}_{nT} \xrightarrow{p} \theta_0$.

Proof. See Appendix D.4. ■

3.2 Distribution of QML Estimator

The asymptotic distribution of the QML estimator $\hat{\theta}_{nT}$ can be derived from the Taylor expansion of $\frac{\partial \ln L_{n,T}(\hat{\theta})}{\partial\theta}$ around θ_0. At θ_0, from Equation (C.2) and (C.3), the first order derivative of the concentrated likelihood function at θ_0 is in (C.5) of Appendix C, which involves both linear and quadratic functions of $\tilde{V}_{nt}$.

[4] See Appendix D.2 for proof.

[5] When n is finite, the condition is $\left(\frac{1}{n}\ln\left|\sigma_0^2 S_n'^{-1}S_n^{-1}\right| - \frac{1}{n}\ln\left|\sigma_n^2(\lambda)S_n'^{-1}(\lambda)S_n^{-1}(\lambda)\right|\right) \neq 0$ for $\lambda \neq \lambda_0$.

From Equation (2.3),

$$\tilde{Z}_{nt} = \tilde{Z}^*_{nt} - (\bar{U}_{nT,-1},\ W_n\bar{U}_{nT,-1},\ \mathbf{0}) \tag{3.3}$$

where $\tilde{Z}^*_{nt} = ((\tilde{\tilde{\mathcal{X}}}_{n,t-1} + U_{n,t-1}),\ (W_n\tilde{\tilde{\mathcal{X}}}_{n,t-1} + W_nU_{n,t-1}),\ \tilde{X}_{nt})$ with $\tilde{\tilde{\mathcal{X}}}_{n,t-1} = \mathcal{X}_{n,t-1} - \bar{\mathcal{X}}_{nT,-1}$. Hence, $\tilde{Z}_{nt}$ has two components: one is $\tilde{Z}^*_{nt}$, which is uncorrelated with V_{nt}; the other is $-(\bar{U}_{nT,-1},\ W_n\bar{U}_{nT,-1},\ \mathbf{0})$, which is correlated with V_{nt} when $t \leq T-1$.

Correspondingly, $\frac{1}{\sqrt{nT}}\frac{\partial \ln L_{n,T}(\theta_0)}{\partial\theta} = \frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\theta_0)}{\partial\theta} - \Delta_{nT}$ where

$$\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\theta_0)}{\partial\theta} = \begin{pmatrix} \frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}\tilde{Z}^{*\prime}_{nt}V_{nt} \\ \frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}(G_n\tilde{Z}^*_{nt}\delta_0)'V_{nt} + \frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}(V'_{nt}G'_nV_{nt} - \sigma_0^2 trG_n) \\ \frac{1}{2\sigma_0^4}\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}(V'_{nt}V_{nt} - n\sigma_0^2) \end{pmatrix} \tag{3.4}$$

and

$$\Delta_{nT} = \begin{pmatrix} \frac{1}{\sigma_0^2}\sqrt{\frac{T}{n}}(\bar{U}_{nT,-1},\ W_n\bar{U}_{nT,-1},\ \mathbf{0})'\bar{V}_{nT} \\ \frac{1}{\sigma_0^2}\sqrt{\frac{T}{n}}(G_n(\bar{U}_{nT,-1},\ W_n\bar{U}_{nT,-1},\ \mathbf{0})\delta_0)'\bar{V}_{nT} + \frac{1}{\sigma_0^2}\sqrt{\frac{T}{n}}\bar{V}'_{nT}G'_n\bar{V}_{nT} \\ \frac{1}{2\sigma_0^4}\sqrt{\frac{T}{n}}\bar{V}'_{nT}\bar{V}_{nT} \end{pmatrix}. \tag{3.5}$$

As is derived in Appendix C.3, the variance matrix of $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\theta_0)}{\partial\theta}$ is equal to

$$E\left(\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\theta_0)}{\partial\theta}\cdot\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\theta_0)}{\partial\theta'}\right) = \Sigma_{\theta_0,nT} + \Omega_{\theta_0,nT} + O\left(\frac{1}{T}\right) \tag{3.6}$$

and $\Omega_{\theta_0,nT} = \frac{\mu_4 - 3\sigma_0^4}{\sigma_0^4}\begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{n}\sum_{i=1}^{n}G^2_{n,ii} & \frac{1}{2\sigma_0^2 n}trG_n \\ 0 & \frac{1}{2\sigma_0^2 n}trG_n & \frac{1}{4\sigma_0^4} \end{pmatrix}$ is a symmetric matrix with μ_4 being the fourth moment of v_{it}, where $G_{n,ii}$ is the (i,i) entry of G_n. When V_{nt} are normally distributed, $\Omega_{\theta_0,nT} = 0$ because $\mu_4 - 3\sigma_0^4 = 0$ for a normal distribution. Denote $\Sigma_{\theta_0} = \lim_{T\to\infty}\Sigma_{\theta_0,nT}$ and $\Omega_{\theta_0} = \lim_{T\to\infty}\Omega_{\theta_0,nT}$, then,

$$\lim_{T\to\infty} E\left(\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\theta_0)}{\partial\theta}\cdot\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\theta_0)}{\partial\theta'}\right) = \Sigma_{\theta_0} + \Omega_{\theta_0}. \tag{3.7}$$

The asymptotic distribution of $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\theta_0)}{\partial\theta}$ can be derived from the central limit theorem for martingale difference arrays (Theorem 2.4). For the term Δ_{nT}, from Theorem 2.3, $\Delta_{nT} = \sqrt{\frac{n}{T}}b_{2n} + O(\sqrt{\frac{n}{T^3}}) +$

$O_p(\frac{1}{\sqrt{T}})$ where

$$b_{2n} = \begin{pmatrix} b_{2n}^{\delta} \\ b_{2n}^{\lambda} \\ b_{2n}^{\sigma^2} \end{pmatrix} = \begin{pmatrix} \frac{1}{n}tr\left(\left(\sum_{h=0}^{\infty} A_n^h\right) S_n^{-1}\right) \\ \frac{1}{n}tr\left(W_n\left(\sum_{h=0}^{\infty} A_n^h\right) S_n^{-1}\right) \\ \mathbf{0} \\ \frac{1}{n}\gamma_0 tr(G_n\left(\sum_{h=0}^{\infty} A_n^h\right) S_n^{-1}) + \frac{1}{n}\rho_0 tr(G_n W_n\left(\sum_{h=0}^{\infty} A_n^h\right) S_n^{-1}) + \frac{1}{n}trG_n \\ \frac{1}{2\sigma_0^2} \end{pmatrix} \tag{3.8}$$

is $O(1)$.[6]

Assumption 8. $\lim_{T\to\infty} E\mathcal{H}_{nT}$ is nonsingular or $\lim_{n\to\infty} \frac{1}{n}(trG_n'G_n + trG_n^2 - \frac{2(trG_n)^2}{n}) \neq 0$.

Assumption 8 is a condition for the nonsingularity of the information matrix Σ_{θ_0}. When $\lim_{T\to\infty} E\mathcal{H}_{nT}$ is singular, as long as we have $\lim_{n\to\infty} \frac{1}{n}(trG_n'G_n + trG_n^2 - \frac{2(trG_n)^2}{n}) \neq 0$, the information matrix Σ_{θ_0} is still nonsingular (see Appendix D.2).

Claim 3.4 *Under Assumption 1-8,*

$$\frac{1}{\sqrt{nT}}\frac{\partial \ln L_{n,T}(\theta_0)}{\partial\theta} + \Delta_{nT} \xrightarrow{d} N(0, \Sigma_{\theta_0} + \Omega_{\theta_0}), \tag{3.9}$$

where Δ_{nT} *is in (3.5) and* $\Delta_{nT} = \sqrt{\frac{n}{T}}b_{2n} + O_p\left(\max\left(\sqrt{\frac{n}{T^3}}, \sqrt{\frac{1}{T}}\right)\right)$ *with* b_{2n} *in Equation (3.8). When* $\{v_{it}\}$, $i = 1, 2, ..., n$ *and* $t = 1, 2, ..., T$, *are normal,* $\frac{1}{\sqrt{nT}}\frac{\partial \ln L_{n,T}(\theta_0)}{\partial\theta} + \Delta_{nT} \xrightarrow{d} N(0, \Sigma_{\theta_0})$.

Proof. *See Appendix D.5.* ■

Claim 3.5 *Under Assumption 1-8,* $\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\theta)}{\partial\theta\partial\theta'} - \frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\theta_0)}{\partial\theta\partial\theta'} = \|\theta - \theta_0\| \cdot O_p(1)$.

Proof. *See Appendix D.6.* ■

Claim 3.6 *Under Assumption 1-8,* $\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\theta_0)}{\partial\theta\partial\theta'} - \frac{\partial^2 Q_{n,T}(\theta_0)}{\partial\theta\partial\theta'} = O_p\left(\frac{1}{\sqrt{nT}}\right)$.

Proof. *See Appendix D.7.* ■

Using Claim 3.4, Claim 3.5 and Claim 3.6, we have the following theorem for the distribution of $\hat{\theta}_{nT}$.

Theorem 3.7 *Under Assumption 1-8,*

$$\begin{aligned} & \sqrt{nT}\left(\hat{\theta}_{nT} - \left(\theta_0 - \frac{b_{\theta_0,nT}}{T}\right)\right) + O_p\left(\max\left(\sqrt{\frac{n}{T^3}}, \sqrt{\frac{1}{T}}\right)\right) \\ = \; & \sqrt{nT}(\hat{\theta}_{nT} - \theta_0) + \sqrt{\frac{n}{T}}b_{\theta_0,nT} + O_p\left(\max\left(\sqrt{\frac{n}{T^3}}, \sqrt{\frac{1}{T}}\right)\right) \xrightarrow{d} N(0, \Sigma_{\theta_0}^{-1}(\Sigma_{\theta_0} + \Omega_{\theta_0})\Sigma_{\theta_0}^{-1}) \end{aligned} \tag{3.10}$$

[6] When $\gamma_0 = \rho_0 = 0$, $b_{2n} = (b_{2,n}^{\delta\prime}, b_{2n}^{\lambda}, b_{2n}^{\sigma^2})' = (\frac{1}{n}trS_n^{-1}, \frac{1}{n}trG_n, \mathbf{0}, \frac{1}{n}trG_n, \frac{1}{2\sigma_0^2})'$.

where $b_{\theta_0,nT} = \Sigma_{\theta_0,nT}^{-1} b_{2n}$ *is* $O(1)$.

When $\frac{n}{T} \to 0$,

$$\sqrt{nT}(\hat{\theta}_{nT} - \theta_0) \xrightarrow{d} N(0, \Sigma_{\theta_0}^{-1}(\Sigma_{\theta_0} + \Omega_{\theta_0})\Sigma_{\theta_0}^{-1}). \tag{3.11}$$

When $\frac{n}{T} \to k < \infty$,

$$\sqrt{nT}(\hat{\theta}_{nT} - \theta_0) + \sqrt{k} b_{\theta_0,nT} \xrightarrow{d} N(0, \Sigma_{\theta_0}^{-1}(\Sigma_{\theta_0} + \Omega_{\theta_0})\Sigma_{\theta_0}^{-1}). \tag{3.12}$$

When $\frac{n}{T} \to \infty$,

$$T(\hat{\theta}_{nT} - \theta_0) + b_{\theta_0,nT} \xrightarrow{p} 0. \tag{3.13}$$

Additionally, if $\{v_{it}\}$, $i = 1, 2, ..., n$ *and* $t = 1, 2, ..., T$, *are normal, (3.10) becomes*

$$\sqrt{nT}(\hat{\theta}_{nT} - \theta_0) + \sqrt{\frac{n}{T}} b_{\theta_0,nT} + O_p\left(\max\left(\sqrt{\frac{n}{T^3}}, \sqrt{\frac{1}{T}}\right)\right) \xrightarrow{d} N(0, \Sigma_{\theta_0}^{-1}). \tag{3.14}$$

Proof. See Appendix D.8. ∎

Hence, $\hat{\theta}_{nT}$ is consistent but is biased with a bias of the order $O(T^{-1})$. For the distribution of $\hat{\theta}_{nT}$, when T is relatively large, the QML estimators are $\sqrt{nT}$ consistent and asymptotically properly centered normal; when n is asymptotically proportional to T, the estimators are $\sqrt{nT}$ consistent and asymptotically normal, but the limit distribution does not center around the truth; when n is relatively large, the estimators are consistent with rate T and have a degenerate distribution.

The estimators of fixed effects are $\sqrt{T}$ consistent and asymptotically centered normal, as shown below.

Theorem 3.8 *Assume that the elements of* $\mathbf{c}_{n0}$ *are bounded. Then under Assumption 1-8, for* $i = 1, 2, \cdots, n$,

$$\sqrt{T}\,(\hat{c}_{i,nT} - c_{i,0}) \xrightarrow{d} N(0, \sigma_0^2) \tag{3.15}$$

and they are asymptotically independent with each other.

Proof. *See Appendix D.9.* ∎

3.3 Bias Reduction

From Equation (3.10), the QML estimator $\hat{\theta}_{nT}$ has the bias $-\frac{1}{T} b_{\theta_0,nT}$ and the confidence interval is not centered when $\frac{n}{T} \to k$ where $0 < k < \infty$. Furthermore, when T is small relative to n in the sense that $\frac{n}{T} \to \infty$, the presence of $b_{\theta_0,nT}$ causes $\hat{\theta}_{nT}$ to have the slower T^{-1} rate of convergence in (3.13). An analytical bias reduction procedure is to correct the bias $B_{nT} = -b_{\theta_0,nT}$ by constructing an estimator $\hat{B}_{nT}$ and defining the bias corrected estimator as

$$\hat{\theta}^1_{nT} = \hat{\theta}_{nT} - \frac{\hat{B}_{nT}}{T}. \tag{3.16}$$

From Theorem 3.7, $B_{nT} = -\Sigma^{-1}_{\theta_0,nT} b_{2n}$ where b_{2n} is defined in (3.8), and we choose

$$\hat{B}_{nT} = \left[\left(E \left(\frac{1}{nT} \frac{\partial^2 \ln L_{n,T}(\theta)}{\partial\theta\partial\theta'} \right) \right)^{-1} b_{2n}(\theta) \right] \Bigg|_{\theta=\hat{\theta}_{nT}}. \tag{3.17}$$

We show that when $T/n^{1/3} \to \infty$, $\hat{\theta}^1_{nT}$ is $\sqrt{nT}$ consistent and asymptotically centered normal even when $n/T \to \infty$.

For showing our result for the bias corrected estimator, we need the following additional assumption.

Assumption 9. $\sum_{h=0}^{\infty} A_n^h(\theta)$ and $\sum_{h=1}^{\infty} h A_n^{h-1}(\theta)$ are uniformly bounded in either row sum or column sums, uniformly in a neighborhood of θ_0.

Assumption 9 can be verified through the following lemma.

Lemma 3.9 *If $\|A_n(\theta_0))\|_\infty < 1$ (resp: $\|A_n(\theta_0))\|_1 < 1$), then the row sum (resp: columns sum) of $\sum_{h=0}^{\infty} A_n^h(\theta)$ and $\sum_{h=1}^{\infty} h A_n^{h-1}(\theta)$ are bounded uniformly in n and in a neighborhood of θ_0.*

Proof. *See Appendix D.10.* ■

Our result for the bias corrected estimator is as follows.

Theorem 3.10 *If $T/n^{1/3} \to \infty$, under Assumption 1-9,*

$$\sqrt{nT}(\hat{\theta}^1_{nT} - \theta_0) \xrightarrow{d} N(0, \Sigma^{-1}_{\theta_0}(\Sigma_{\theta_0} + \Omega_{\theta_0})\Sigma^{-1}_{\theta_0}). \tag{3.18}$$

Proof. *See Appendix D.11* ■

Hence, if T grows faster than $n^{1/3}$, the analytical bias correction will give us estimators that are asymptotically normal and centered around θ_0. For the case $\frac{n}{T} \to k$, $\hat{\theta}^1_{nT}$ has removed the asymptotic bias $b_{\theta_0,nT}$. Note that $\frac{n}{T} \to k$ implies $T/n^{1/3} \to \infty$. For the case $\frac{n}{T} \to \infty$, as long as $T/n^{1/3} \to \infty$, the rate of convergence of $\hat{\theta}^1_{nT}$ is $\sqrt{nT}$, which is also an improvement upon the T rate of convergence of $\hat{\theta}_{nT}$. Thus, $\hat{\theta}^1_{nT}$ might have better performance in economic applications, especially when n is much larger than T.

3.4 Monte Carlo Results

We conduct a small Monte Carlo experiment to evaluate the performance of our ML estimators and the bias corrected estimators. We generate samples from Equation (2.1) using $\theta_0^a = (0.2, 0.2, 1, 0.2, 1)'$ and $\theta_0^b = (0.3, 0.3, 1, 0.3, 1)'$ where $\theta_0 = (\gamma_0, \rho_0, \beta_0', \lambda_0, \sigma_0^2)'$, and X_{nt}, $\mathbf{c}_{n0}$ and V_{nt} are generated from independent

Table 1: Performance of Bias Corrected Estimators: the Biases

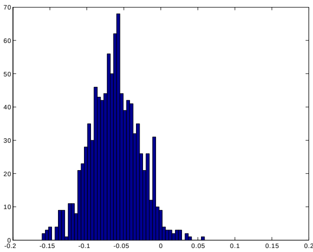
Case				Bias of $\hat{\theta}_{nT}$ (1st line) and $\hat{\theta}^1_{nT}$ (2nd line)				
	T	n	θ_0	γ	ρ	β	λ	σ^2
(1)	10	49	θ_0^a	-0.0628	-0.0031	-0.0077	-0.0024	-0.1168
				-0.0049	-0.0030	-0.0010	0.0166	-0.0488
(2)	10	49	θ_0^b	-0.0701	-0.0080	-0.0111	-0.0105	-0.1193
				-0.0067	-0.0050	-0.0019	-0.0262	-0.0555
(3)	10	196	θ_0^a	-0.0625	-0.0036	-0.0076	-0.0024	-0.1105
				-0.0050	-0.0036	-0.0009	0.0175	-0.0418
(4)	10	196	θ_0^b	-0.0691	-0.0067	-0.0109	-0.0091	-0.1129
				-0.0065	-0.0073	-0.0021	0.0281	-0.0481
(5)	50	49	θ_0^a	-0.0121	-0.0018	-0.0008	0.0005	-0.0220
				-0.0005	-0.0029	-0.0007	0.0052	-0.0038
(6)	50	49	θ_0^b	-0.0132	-0.0024	-0.0009	-0.0006	-0.0221
				-0.0010	-0.0055	-0.0011	0.0071	-0.0047
(7)	50	196	θ_0^a	-0.0122	-0.0002	-0.0004	0.0012	-0.0211
				-0.0005	-0.0014	-0.0004	0.0062	-0.0028
(8)	50	196	θ_0^b	-0.0133	-0.0008	-0.0005	0.0004	-0.0212
				-0.0011	-0.0042	-0.0007	0.0086	-0.0038

Note: $\theta_0^a = (0.2, 0.2, 1, 0.2, 1)$ and $\theta_0^b = (0.3, 0.3, 1, 0.3, 1)$.

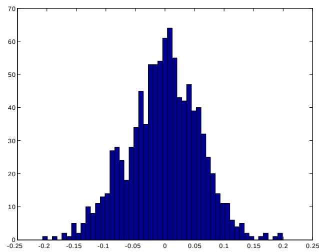
Table 2: Performance of Bias Corrected Estimators: the Standard Errors

Case				S.E. of $\hat{\theta}_{nT}$ (1st line) and $\hat{\theta}^1_{nT}$ (2nd line)				
	T	n	θ_0	γ	ρ	β	λ	σ^2
(1)	10	49	θ_0^a	0.0322	0.0591	0.0452	0.0477	0.0566
				0.0334	0.0617	0.0469	0.0478	0.0610
(2)	10	49	θ_0^b	0.0322	0.0570	0.0453	0.0457	0.0567
				0.0333	0.0599	0.0469	0.0451	0.0609
(3)	10	196	θ_0^a	0.0161	0.0304	0.0226	0.0246	0.0285
				0.0167	0.0317	0.0234	0.0247	0.0307
(4)	10	196	θ_0^b	0.0160	0.0292	0.0226	0.0236	0.0285
				0.0166	0.0307	0.0234	0.0233	0.0307
(5)	50	49	θ_0^a	0.0141	0.0260	0.0202	0.0213	0.0280
				0.0143	0.0263	0.0204	0.0213	0.0286
(6)	50	49	θ_0^b	0.0139	0.0243	0.0203	0.0201	0.0281
				0.0140	0.0246	0.0205	0.0200	0.0287
(7)	50	196	θ_0^a	0.0071	0.0134	0.0101	0.0110	0.0140
				0.0071	0.0136	0.0102	0.0110	0.0143
(8)	50	196	θ_0^b	0.0070	0.0125	0.0101	0.0103	0.0141
				0.0070	0.0127	0.0102	0.0103	0.0143

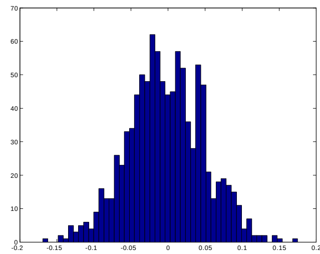
Table 3: Empirical Densities of Biases when $n = 49$ and $T = 10$.



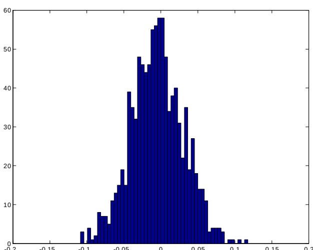
$\hat{\gamma}_{nT} - \gamma_0$



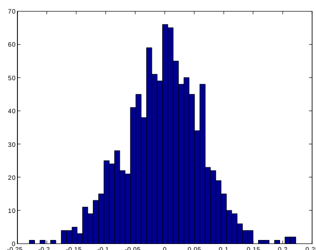
$\hat{\rho}_{nT} - \rho_0$



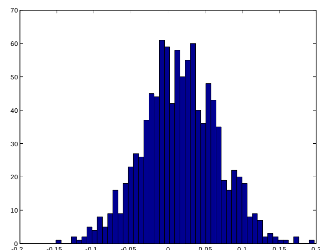
$\hat{\lambda}_{nT} - \lambda_0$



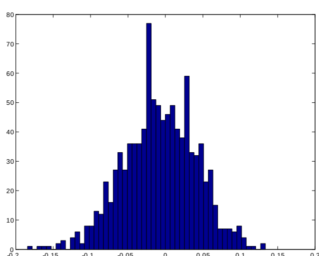
$\hat{\gamma}^1_{nT} - \gamma_0$



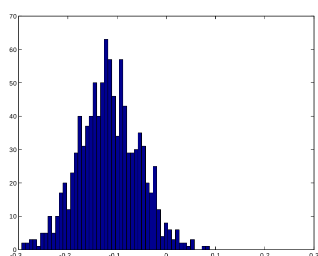
$\hat{\rho}^1_{nT} - \rho_0$



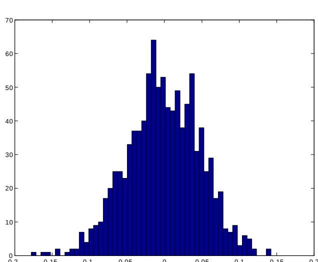
$\hat{\lambda}^1_{nT} - \lambda_0$



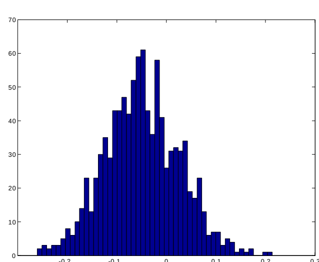
$\hat{\beta}_{nT} - \beta_0$



$\hat{\sigma}^2_{nT} - \sigma^2_0$



$\hat{\beta}^1_{nT} - \beta_0$



$(\hat{\sigma}^2_{nT})^1 - \sigma^2_0$

normal distributions[7] and the spatial weights matrix we use is a rook matrix. We use $T = 10$ and $T = 50$, and $n = 49$ and $n = 196$. For each set of generated sample observations, we calculate the ML estimator $\hat{\theta}_{nT}$ and evaluate the bias $\hat{\theta}_{nT} - \theta_0$; we then construct the bias corrected estimator $\hat{\theta}^1_{nT}$ and evaluate the bias $\hat{\theta}^1_{nT} - \theta_0$. We do this for 1000 times to see if the bias is reduced on average by using the analytical bias correction procedure, i.e., to compare $\frac{1}{1000}\sum_{i=1}^{1000}(\hat{\theta}_{nT} - \theta_0)_i$ with $\frac{1}{1000}\sum_{i=1}^{1000}(\hat{\theta}^1_{nT} - \theta_0)_i$. With two different values of θ_0 for each n and T, finite sample properties of both estimators are summarized in Table 1 and Table 2, where Table 1 is for the magnitude of biases and Table 2 is for standard errors of estimators.

We can see that both estimators have some biases, but the bias corrected estimators reduce those biases which are relatively larger. This is consistent with our asymptotic analysis, because the bias corrected estimators will eliminate the bias of order $O(T^{-1})$. Also, the bias reduction is achieved while there is no significant increase in the variance of the estimators, as can be seen from Table 2. Table 3 gives us empirical densities of $\hat{\theta}_{nT} - \theta_0$ and $\hat{\theta}^1_{nT} - \theta_0$ when $n = 49$ and $T = 10$.

For different cases of n and T, we can see that for each given n, when T is larger, the biases of two sets of estimators will be smaller and the variance will be smaller; for each given T, when n is larger, the biases of two sets of estimators will be nearly the same, but the variance will be smaller. This is consistent with our theoretical prediction, because the bias is of the order $O(T^{-1})$ order and the variance of the estimators is of the order $O(\frac{1}{nT})$. Also, for different values of θ_0, the biases become larger when θ_0 is larger, and the variances nearly do not change.

4 Conclusion

In this paper, we derived the properties of QML estimators of spatial dynamic panel data with fixed effects when both n and T are large. The estimators of the fixed effects are $\sqrt{T}$ consistent and asymptotically normally distributed. For the distribution of the common parameters, where T is asymptotically large relative to n, the estimators are $\sqrt{nT}$ consistent and asymptotically normal, with the limit distribution centered around 0; when n is asymptotically proportional to T, the estimators are $\sqrt{nT}$ consistent and asymptotically normal, but the limit distribution is not centered around 0; and when n is large relative to T, the estimators are consistent with rate T, and have a degenerate limit distribution. We also propose a bias correction for our estimators. We show that when T grows faster than $n^{1/3}$, the correction will eliminate the bias of order $O(T^{-1})$ and yield a centered confidence interval. The contribution of this paper is that it

[7] We generated the spatial panel data with $20 + T$ periods and then take the last T periods as our sample. And the initial value is generated as $N(0, I_n)$ in the simulation. We have also generated the data with a much longer history $1000 + T$ and the results are similar.

establishes the asymptotic properties of QML estimators and bias-corrected estimators of the spatial dynamic panel model when both n and T are large.

Appendices

A Some Basic Lemmas and Proofs

A.1 Basic Lemmas

Let $V_{nt} = (v_{1t}, v_{2t}, \cdots, v_{nt})'$ be $n \times 1$ column vector. We assume that $\{v_{it}\}$, $i = 1, 2, \cdots, n$ and $t = 1, 2, \cdots, T$, are *i.i.d.* across i and t with zero mean, variance σ_0^2 and $E\,|v_{it}|^{4+\eta} < \infty$ for some $\eta > 0$. Also, let D_{nt} be $n \times 1$ vector of uniformly bounded constants for all n and t. Denote

$$\mathbb{U}_{nt} = \sum_{h=1}^{\infty} P_{nh} V_{n,t+1-h} \tag{A.1}$$

and

$$\mathbb{W}_{nt} = \sum_{h=1}^{\infty} Q_{nh} V_{n,t+1-h} \tag{A.2}$$

where $\{P_{nh}\}_{h=1}^{\infty}$ and $\{Q_{nh}\}_{h=1}^{\infty}$ are sequences of $n \times n$ nonstochastic square matrices.

Denote $\tilde{\mathbb{U}}_{nt} = \mathbb{U}_{nt} - \bar{\mathbb{U}}_{nT}$ where $\bar{\mathbb{U}}_{nT} = \left(\sum_{t=1}^{T} \mathbb{U}_{nt}\right)/T$, and $\tilde{\tilde{\mathbb{U}}}_{n,t-1} = \mathbb{U}_{n,t-1} - \bar{\mathbb{U}}_{nT,-1}$ where $\bar{\mathbb{U}}_{nT,-1} = \left(\sum_{t=0}^{T-1} \mathbb{U}_{nt}\right)/T$. Also $\tilde{\mathbb{W}}_{nt}$, $\tilde{\tilde{\mathbb{W}}}_{n,t-1}$ and and $\tilde{V}_{nt}$ are defined similarly.

Lemma A.1 *With* $\mathbb{U}_{nt}$ *in (A.1) and* $\mathbb{W}_{nt}$ *in (A.2),* $\bar{\mathbb{U}}_{nT} = \sum_{h=1}^{\infty} \ddot{P}_{nh} V_{n,T+1-h}$ *and* $\bar{\mathbb{W}}_{nT} = \sum_{h=1}^{T} \ddot{Q}_{nh} V_{n,T+1-h}$ *where*

$$\ddot{P}_{nh} = \begin{cases} \frac{1}{T}(P_{n1} + P_{n2} + \cdots + P_{nh}) = \frac{1}{T}\sum_{g=1}^{h} P_{ng} & \textit{for } h \le T \\ \frac{1}{T}\sum_{g=1}^{T} P_{n,h-T+g} & \textit{for } h > T \end{cases} \tag{A.3}$$

and $\ddot{Q}_{nh}$ *has the same pattern. Furthermore,* $\sum_{h=1}^{\infty} \ddot{P}_{nh} = \sum_{h=1}^{\infty} P_{nh}$*, and* $\sum_{h=1}^{\infty} \ddot{Q}_{nh} = \sum_{h=1}^{\infty} Q_{nh}$*.*

Lemma A.2 *Under Assumption A1, for* $t \ge s$*,*

$$E(\mathbb{U}_{nt}\mathbb{W}'_{ns}) = \sigma_0^2 \left(\sum_{h=1}^{\infty} P_{n,t-s+h} Q'_{nh} \right) \tag{A.4}$$

and

$$E(\mathbb{U}'_{nt}\mathbb{W}_{ns}) = \sigma_0^2 tr\left(\sum_{h=1}^{\infty} P'_{n,t-s+h}Q_{nh}\right). \tag{A.5}$$

Lemma A.3 *Under Assumption A1,*

$$E(V'_{nt}B_{1n}V_{ns})(V'_{ng}B_{2n}V_{nh})$$
$$= \begin{cases} (\mu_4 - 3\sigma_0^4)\sum_{i=1}^n B_{1,ii}B_{2,ii} + \sigma_0^4(trB_{1n}\times trB_{2n} + trB_{1n}B_{2n} + trB_{1n}B'_{2n}) & for\ t=s=g=h \\ \sigma_0^4 trB_{1n}\times trB_{2n} & for\ t=s\neq g=h \\ \sigma_0^4 tr(B_{1n}B'_{2n}) & for\ t=g\neq s=h \\ \sigma_0^4 tr(B_{1n}B_{2n}) & for\ t=h\neq s=g \\ 0 & otherwise \end{cases}.$$

Lemma A.4 *Under Assumption A1, for $t \geq s$,*

$$\begin{aligned} & Cov(\mathbb{U}'_{nt}\mathbb{W}_{nt}, \mathbb{U}'_{ns}\mathbb{W}_{ns}) \\ = \; & \sigma_0^4 tr\left[\left(\sum_{h=1}^{\infty} P_{nh}P'_{n,t-s+h}\right)\left(\sum_{h=1}^{\infty} Q_{n,t-s+h}Q'_{nh}\right) + \left(\sum_{h=1}^{\infty} Q_{nh}P'_{n,t-s+h}\right)\left(\sum_{h=1}^{\infty} Q_{n,t-s+h}P'_{nh}\right)\right] \\ & +(\mu_4 - 3\sigma_0^4)\sum_{h=1}^{\infty}\sum_{i=1}^{n}(P'_{n,t-s+h}Q_{n,t-s+h})_{ii}(P'_{nh}Q_{nh})_{ii} \end{aligned} \tag{A.6}$$

Lemma A.5 *Suppose B_n, C_{nh} and D_{nh} are $n \times n$ square matrices with all elements being non-negative, and B_n, $\sum_{h=1}^{\infty} C_{nh}$ and $\sum_{h=1}^{\infty} D_{nh}$ are uniformly bounded in both row and column sums. Then, the product $\sum_{h=1}^{\infty} C_{nh}B_nD_{nh}$ is uniformly bounded in both row and column sums.*

Lemma A.6 *Under Assumption A1, A2 and A4,*

$$Var(\sum_{t=1}^{T}\mathbb{U}'_{nt}\mathbb{W}_{nt}) = \sum_{t=1}^{T}\sum_{s=1}^{T} Cov(\mathbb{U}'_{nt}\mathbb{W}_{nt}, \mathbb{U}'_{ns}\mathbb{W}_{ns}) = O(nT).$$

Theorem A.7 *Under Assumption A1, A2 and A4,*

$$\frac{1}{nT}\sum_{t=1}^{T}\mathbb{U}'_{nt}\mathbb{W}_{nt} - E(\frac{1}{nT}\sum_{t=1}^{T}\mathbb{U}'_{nt}\mathbb{W}_{nt}) = O_p\left(\frac{1}{\sqrt{nT}}\right), \tag{A.7}$$

$$\frac{1}{n}\bar{\mathbb{U}}'_{nT}\bar{\mathbb{W}}_{nT} - E(\frac{1}{n}\bar{\mathbb{U}}'_{nT}\bar{\mathbb{W}}_{nT}) = O_p\left(\frac{1}{\sqrt{nT^2}}\right) \tag{A.8}$$

and

$$\frac{1}{nT}\sum_{t=1}^{T}\tilde{\mathbb{U}}'_{nt}\tilde{\mathbb{W}}_{nt} - E(\frac{1}{nT}\sum_{t=1}^{T}\tilde{\mathbb{U}}'_{nt}\tilde{\mathbb{W}}_{nt}) = O_p\left(\frac{1}{\sqrt{nT}}\right) \tag{A.9}$$

where $E(\frac{1}{nT}\sum_{t=1}^{T}\mathbb{U}'_{nt}\mathbb{W}_{nt}) = \frac{\sigma_0^2}{n}tr\left(\sum_{h=1}^{\infty}P'_{nh}Q_{nh}\right) = O(1)$ *and* $E(\frac{1}{n}\bar{\mathbb{U}}'_{nT}\bar{\mathbb{W}}_{nT}) = \frac{\sigma_0^2}{n}tr\left(\sum_{h=1}^{\infty}\ddot{P}'_{nh}\ddot{Q}_{nh}\right) = O(\frac{1}{T})$ *where* $\ddot{P}_{nh}$ *and* $\ddot{Q}_{nh}$ *are defined in (A.3) in Lemma A.1.*

Theorem A.8 *Under Assumption A1, A2, A3 and A4,*

$$\frac{1}{nT}\sum_{t=1}^{T}\tilde{D}'_{nt}\tilde{\mathbb{U}}_{nt} = \frac{1}{nT}\sum_{t=1}^{T}\tilde{D}'_{nt}\mathbb{U}_{nt} = O_p\left(\frac{1}{\sqrt{nT}}\right), \tag{A.10}$$

and

$$\frac{1}{nT}\sum_{t=1}^{T}\bar{D}'_{nt}\bar{\mathbb{U}}_{nt} = O_p\left(\frac{1}{\sqrt{nT}}\right). \tag{A.11}$$

Lemma A.9 *Under Assumption A1 and A4, for an* $n\times n$ *matrix* B_n*, uniformly bounded in row and column sums,*

$$\frac{1}{nT}\sum_{t=1}^{T}V'_{nt}B_nV_{nt} - E(\frac{1}{nT}\sum_{t=1}^{T}V'_{nt}B_nV_{nt}) = O_p(\frac{1}{\sqrt{nT}}), \tag{A.12}$$

$$\frac{1}{n}\bar{V}'_{nT}B_n\bar{V}_{nT} - E(\frac{1}{n}\bar{V}'_{nT}B_n\bar{V}_{nT}) = O_p(\frac{1}{\sqrt{nT^2}}) \tag{A.13}$$

and

$$\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}'_{nt}B_n\tilde{V}_{nt} - E(\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}'_{nt}B_n\tilde{V}_{nt}) = O_p(\frac{1}{\sqrt{nT}}), \tag{A.14}$$

where $E(\frac{1}{nT}\sum_{t=1}^{T}V'_{nt}B_nV_{nt}) = \frac{1}{n}\sigma_0^2tr(B_n) = O(1)$ *and* $E(\frac{1}{n}\bar{V}'_{nT}B_n\bar{V}_{nT}) = \frac{1}{nT}\sigma_0^2tr(B_n) = O\left(\frac{1}{T}\right)$.

Lemma A.10 *Under Assumption A1,*

$$E([(\mathbb{U}_{n,t-1})_i]^4) = (\mu_4 - 3\sigma_0^4)\sum_{h=1}^{\infty}\sum_{j=1}^{n}[(P_{nh})_{ij}]^4 + 3\sigma_0^4[\sum_{h=1}^{\infty}(P_{nh}P'_{nh})_{ii}]^2.$$

Theorem A.11 *Under Assumption A1, A2 and A4,*

$$\sqrt{\frac{T}{n}}\left(\bar{\mathbb{U}}'_{nT,-1}\bar{V}_{nT} - E\left(\bar{\mathbb{U}}'_{nT,-1}\bar{V}_{nT}\right)\right) = O_p\left(\frac{1}{\sqrt{T}}\right) \tag{A.15}$$

where $\sqrt{\frac{T}{n}}E\left(\bar{\mathbb{U}}'_{nT,-1}\bar{V}_{nT}\right) = \sqrt{\frac{n}{T}}\frac{1}{n}\sigma_0^2tr\left(\sum_{h=1}^{\infty}P_{nh}\right) + O\left(\sqrt{\frac{n}{T^3}}\right)$ *when* $T\to\infty$.

Lemma A.12 *Let* B_n^- *denote the lower diagonal matrix constructed from symmetric* B_n *by deleting the diagonal and the upper triangle entries. Under Assumptions A1, A2 and if* B_n *is uniformly bounded in row and column sums, and* K_n *is an n-dimensional nonstochastic vector with all its elements uniformly bounded,*

(a) $\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}(\sum_{j=1}^{i-1}b_{nij}v_{jt})^2-\frac{1}{2}\sigma_0^2[tr(B_n^2)-vec_D'(B_n)vec_D(B_n)]$
$=\frac{1}{nT}\sum_{t=1}^{T}[V_{nt}'B_n^{-'}B_n^{-}V_{nt}-\sigma_0^2tr(B_n^{-'}B_n^{-})]=O_p\left(\frac{1}{\sqrt{nT}}\right).$
(b) $\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}k_{ni}(\sum_{j=1}^{i-1}b_{nij}v_{jt})=\frac{1}{nT}\sum_{t=1}^{T}K_n'B_n^{-}V_{nt}=O_p\left(\frac{1}{\sqrt{nT}}\right).$
(c) $\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}u_{n,t-1,i}(\sum_{j=1}^{i-1}b_{nij}v_{jt})=\frac{1}{nT}\sum_{t=1}^{T}\mathbb{U}_{n,t-1}'B_n^{-}V_{nt}=O_p\left(\frac{1}{\sqrt{nT}}\right).$
(d) $\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}k_{ni}u_{n,t-1,i}=\frac{1}{nT}\sum_{t=1}^{T}K_n'\mathbb{U}_{n,t-1}=O_p\left(\frac{1}{\sqrt{nT}}\right)$
where $vec_D(B_n)$ *is the n-dimensional column vector formed by the the diagonal elements of* B_n.

For the theorem that follows, we will consider the following form:

$$Q_{nT}=\sum_{t=1}^{T}\left(\mathbb{U}_{n,t-1}'V_{nt}+D_{nt}'V_{nt}+V_{nt}'B_nV_{nt}-\sigma_0^2trB_n\right)=\sum_{t=1}^{T}\sum_{i=1}^{n}z_{nt,i},$$

where B_n is a $n\times n$ symmetric matrix which is uniformly bounded in both row and column sums, and $z_{nt,i}=(u_{i,t-1}+d_{nti})v_{it}+b_{n,ii}(v_{it}^2-\sigma_0^2)+2(\sum_{j=1}^{i-1}b_{n,ij}v_{jt})v_{it}$. Then, for the mean and variance of Q_{nT}, $\mu_{Q_{nT}}=0$ and

$$\sigma_{Q_{nT}}^2=T\sigma_0^4tr\left(\sum_{h=1}^{\infty}P_{nh}'P_{nh}\right)+\sigma_0^2\sum_{t=1}^{T}D_{nt}'D_{nt}+T\left((\mu_4-3\sigma_0^4)\sum_{i=1}^{n}b_{n,ii}^2+2\sigma_0^4tr(B_n^2)\right)+2\mu_3\sum_{t=1}^{T}\sum_{i=1}^{n}d_{nti}b_{n,ii},$$

where $\mu_s=Ev_{it}^s$ for $s=3,4$, $b_{n,ii}$'s are diagonal elements of B_n and d_{nti} is the ith element of D_{nt}.

Theorem A.13 *Under Assumption A1, A2, A3, A4 and that row and column sums of* B_n *are bounded uniformly in n, if the sequence* $\frac{1}{nT}\sigma_{Q_{nT}}^2$ *is bounded away from zero, then,*

$$\frac{Q_{nT}}{\sigma_{Q_{nT}}}\xrightarrow{d}N(0,1). \tag{A.16}$$

A.2 Proofs of the Basic Lemmas

Proof for Lemma A.1

As $\mathbb{U}_{nt}=\sum_{h=1}^{\infty}P_{nh}V_{n,t+1-h}$ so that

$$\begin{aligned}\mathbb{U}_{n1}&=P_{n1}V_{n1}+P_{n2}V_{n0}+P_{n3}V_{n,-1}+\cdots,\\ \mathbb{U}_{n2}&=P_{n1}V_{n2}+P_{n2}V_{n1}+P_{n3}V_{n,0}+P_{n4}V_{n,-1}+\cdots,\\ &\vdots\\ \mathbb{U}_{nT}&=P_{n1}V_{nT}\cdots+P_{nT}V_{n1}+P_{n,T+1}V_{n,0}+P_{n,T+2}V_{n,-1}+\cdots,\end{aligned}$$

we have $\sum_{t=1}^{T}\mathbb{U}_{nt}=P_{n1}V_{nT}+(P_{n1}+P_{n2})V_{n,T-1}+(P_{n1}+P_{n2}+P_{n3})V_{n,T-2}+\cdots+(\sum_{g=1}^{T}P_{ng})V_{n1}$
$+(P_{n2}+P_{n3}+\cdots+P_{n,T+1})V_{n0}+(P_{n3}+P_{n4}+\cdots+P_{n,T+2})V_{n,-1}+\cdots.$

As $\bar{\mathbb{U}}_{nT} = \left(\sum_{t=1}^{T} \mathbb{U}_{nt}\right)/T$, we have $\bar{\mathbb{U}}_{nT} = \sum_{h=1}^{\infty} \ddot{P}_{nh} V_{n,T+1-h}$, where

$$\ddot{P}_{nh} = \begin{cases} \frac{1}{T}(P_{n1} + P_{n2} + \cdots + P_{nh}) = \frac{1}{T}\sum_{g=1}^{h} P_{ng} & \text{for } h \leq T \\ \frac{1}{T}\sum_{g=1}^{T} P_{n,h-T+g} & \text{for } h > T. \end{cases} \tag{A.17}$$

Similarly, $\ddot{Q}_{nh}$ has the same pattern. Furthermore,

$$\begin{aligned} \sum_{h=1}^{\infty} \ddot{P}_{nh} &= \sum_{h=1}^{T}(\frac{1}{T}\sum_{g=1}^{h} P_{ng}) + \sum_{h=T+1}^{\infty}(\frac{1}{T}\sum_{g=1}^{T} P_{n,h-T+g}) = \frac{1}{T}\sum_{h=1}^{T}\sum_{g=1}^{h} P_{ng} + \frac{1}{T}\sum_{h=T+1}^{\infty}(\sum_{g=1}^{T} P_{n,h-T+g}) \\ &= \frac{1}{T}\sum_{h=1}^{T}\sum_{g=1}^{h} P_{ng} + \frac{1}{T}\sum_{h=1}^{T}(\sum_{g=1}^{\infty} P_{n,g+h}) = \frac{1}{T}\sum_{h=1}^{T}\left(\sum_{g=1}^{h} P_{ng} + \sum_{g=h+1}^{\infty} P_{ng}\right) = \sum_{h=1}^{\infty} P_{nh}. \ \blacksquare \end{aligned}$$

Proof for Lemma A.2

First, we have the result that for any nonstochastic $n \times n$ matrix B_n, $EV_{nt}B_nV'_{ns} = \sigma_0^2 B_n$ if $t = s$ and $EV_{nt}B_nV'_{ns} = \mathbf{0}$ otherwise.

For $t \geq s$, defining summations over empty index sets as 0, $\mathbb{U}_{nt} = \sum_{h=1}^{\infty} P_{nh}V_{n,t+1-h} = \sum_{h=1}^{t-s} P_{nh}V_{n,t+1-h} + \sum_{g=1}^{\infty} P_{n,t-s+g}V_{n,s+1-g}$. Hence,

$$\begin{aligned} E(\mathbb{U}_{nt}\mathbb{W}'_{ns}) &= E\left(\sum_{h=1}^{t-s} P_{nh}V_{n,t+1-h} + \sum_{g=1}^{\infty} P_{n,t-s+g}V_{n,s+1-g}\right)\left(\sum_{g=1}^{\infty} Q_{ng}V_{n,s+1-g}\right)' \\ &= E\left(\sum_{g=1}^{\infty} P_{n,t-s+g}V_{n,s+1-g}\right)\left(\sum_{g=1}^{\infty} Q_{ng}V_{n,s+1-g}\right)' = \sigma_0^2\left(\sum_{g=1}^{\infty} P_{n,t-s+g}Q'_{ng}\right). \end{aligned}$$

Also, $E(\mathbb{U}'_{nt}\mathbb{W}_{ns}) = trE(\mathbb{U}_{nt}\mathbb{W}'_{ns}) = \sigma_0^2 tr\left(\sum_{g=1}^{\infty} P_{n,t-s+g}Q'_{ng}\right) = \sigma_0^2 tr\left(\sum_{h=1}^{\infty} P'_{n,t-s+h}Q_{nh}\right)$. $\blacksquare$

Proof for Lemma A.3

For (i) $t = s = g = h$, the result for $E(V'_{nt}B_{1n}V_{nt})(V'_{nt}B_{2n}V_{nt})$ can be found in Lee (2001). For (ii) $t = s \neq g = h$, the result follows directly from the independence of $V'_{nt}B_{1n}V_{nt}$ and $V'_{ng}B_{2n}V_{ng}$. For (iii) $t = g \neq s = h$, note that

$$\begin{aligned} (V'_{nt}B_{1n}V_{ns}) \cdot (V'_{nt}B_{2n}V_{ns}) &= \left(\sum_{j=1}^{n}\sum_{i=1}^{n} b_{1,ij}v_{it}v_{js}\right) \cdot \left(\sum_{l=1}^{n}\sum_{k=1}^{n} b_{2,kl}v_{kt}v_{ls}\right) \\ &= \sum_{l=1}^{n}\sum_{k=1}^{n}\sum_{j=1}^{n}\sum_{i=1}^{n} b_{1,ij}b_{2,kl}[v_{it}v_{kt}][v_{js}v_{ls}]. \end{aligned}$$

As $E(v_{it}v_{kt}) = 0$ whenever $i \neq k$, and $E(v_{it}v_{kt}) = \sigma_0^2$ when $i = k$ for any t, it follows that, for $t \neq s$, $E(V_{nt}'B_{1n}V_{ns})(V_{nt}'B_{2n}V_{ns}) = \sigma_0^4 \sum_{i=1}^n \sum_{j=1}^n b_{1,ij}b_{2,ij} = \sigma_0^4 tr(B_{1n}B_{2n}')$. For (iv) $t = h \neq s = g$, we use $E(V_{nt}'B_{1n}V_{ns}) \cdot (V_{ns}'B_{2n}V_{nt}) = E(V_{nt}'B_{1n}V_{ns}) \cdot (V_{nt}'B_{2n}'V_{ns}) = \sigma_0^4 tr(B_{1n}B_{2n})$. ■

Proof for Lemma A.4

Using Lemma A.3, we have for $t \geq s$,

$$\begin{aligned} E(\mathbb{U}_{nt}'\mathbb{W}_{nt} \times \mathbb{U}_{ns}'\mathbb{W}_{ns}) &= E\left[\begin{array}{c} (\sum_{h=1}^{t-s} P_{nh}V_{n,t+1-h} + \sum_{g=1}^{\infty} P_{n,t-s+g}V_{n,s+1-g})' \\ \times(\sum_{h=1}^{t-s} Q_{nh}V_{n,t+1-h} + \sum_{g=1}^{\infty} Q_{n,t-s+g}V_{n,s+1-g}) \\ \times(\sum_{g=1}^{\infty} P_{ng}V_{n,s+1-g})'(\sum_{g=1}^{\infty} Q_{ng}V_{n,s+1-g}) \end{array}\right] \\ &= E_1 + E_2 \end{aligned}$$

where

$$\begin{aligned} E_1 &= E(\sum_{h=1}^{t-s} P_{nh}V_{n,t+1-h})'(\sum_{h=1}^{t-s} Q_{nh}V_{n,t+1-h}) \times (\sum_{g=1}^{\infty} P_{ng}V_{n,s+1-g})'(\sum_{g=1}^{\infty} Q_{ng}V_{n,s+1-g}) \\ &= \sigma_0^4 tr\left(\sum_{h=1}^{t-s} P_{nh}'Q_{nh}\right) \times tr\left(\sum_{g=1}^{\infty} P_{ng}'Q_{ng}\right) \end{aligned}$$

and

$$\begin{aligned} E_2 &= E(\sum_{g=1}^{\infty} P_{n,t-s+g}V_{n,s+1-g})'(\sum_{g=1}^{\infty} Q_{n,t-s+g}V_{n,s+1-g}) \times (\sum_{g=1}^{\infty} P_{ng}V_{n,s+1-g})'(\sum_{g=1}^{\infty} Q_{ng}V_{n,s+1-g}) \\ &= E\left[\begin{array}{c} (\sum_{g=1}^{\infty}(P_{n,t-s+g}V_{n,s+1-g})'Q_{n,t-s+g}V_{n,s+1-g} + \sum_{g=1}^{\infty}\sum_{h\neq g}^{\infty}(P_{n,t-s+h}V_{n,s+1-h})'Q_{n,t-s+g}V_{n,s+1-g}) \\ \times(\sum_{g=1}^{\infty}(P_{ng}V_{n,s+1-g})'Q_{ng}V_{n,s+1-g} + \sum_{g=1}^{\infty}\sum_{h\neq g}^{\infty}(P_{nh}V_{n,s+1-h})'Q_{ng}V_{n,s+1-g}) \end{array}\right] \\ &= E\left((\sum_{g=1}^{\infty}(P_{n,t-s+g}V_{n,s+1-g})'Q_{n,t-s+g}V_{n,s+1-g}) \times (\sum_{g=1}^{\infty}(P_{ng}V_{n,s+1-g})'Q_{ng}V_{n,s+1-g})\right) \\ &\quad + E\left((\sum_{g=1}^{\infty}\sum_{h\neq g}^{\infty}(P_{n,t-s+h}V_{n,s+1-h})'Q_{n,t-s+g}V_{n,s+1-g}) \times (\sum_{g=1}^{\infty}\sum_{h\neq g}^{\infty}(P_{nh}V_{n,s+1-h})'Q_{ng}V_{n,s+1-g})\right) \\ &= E_{21} + E_{22}. \end{aligned}$$

Here,

$$\begin{aligned} E_{21} = E&\left(\sum_{g=1}^{\infty}(P_{n,t-s+g}V_{n,s+1-g})'Q_{n,t-s+g}V_{n,s+1-g}(P_{ng}V_{n,s+1-g})'Q_{ng}V_{n,s+1-g}\right) \\ +E&\left(\sum_{g=1}^{\infty}\sum_{h\neq g}^{\infty}(P_{n,t-s+h}V_{n,s+1-h})'Q_{n,t-s+h}V_{n,s+1-h}(P_{ng}V_{n,s+1-g})'Q_{ng}V_{n,s+1-g}\right) \end{aligned}$$

$$= E_{21}^a + E_{21}^b$$

where

$$E_{21}^a = E\left(\sum_{g=1}^{\infty}(P_{n,t-s+g}V_{n,s+1-g})'Q_{n,t-s+g}V_{n,s+1-g}(P_{ng}V_{n,s+1-g})'Q_{ng}V_{n,s+1-g}\right)$$
$$= (\mu_4 - 3\sigma_0^4)\sum_{g=1}^{\infty}\sum_{i=1}^{n}(P'_{n,t-s+g}Q_{n,t-s+g})_{ii}(P'_{ng}Q_{ng})_{ii}$$
$$+\sigma_0^4\sum_{g=1}^{\infty} tr(P'_{n,t-s+g}Q_{n,t-s+g}) \times tr(P'_{ng}Q_{ng}) + \sigma_0^4\sum_{g=1}^{\infty} tr(P'_{n,t-s+g}Q_{n,t-s+g}P'_{ng}Q_{ng})$$
$$+\sigma_0^4\sum_{g=1}^{\infty} tr(P'_{n,t-s+g}Q_{n,t-s+g}Q'_{ng}P_{ng})$$

and

$$E_{21}^b = E\left(\sum_{g=1}^{\infty}\sum_{h\neq g}^{\infty}(P_{n,t-s+h}V_{n,s+1-h})'Q_{n,t-s+h}V_{n,s+1-h}(P_{ng}V_{n,s+1-g})'Q_{ng}V_{n,s+1-g}\right)$$
$$= \sigma_0^4\left(\sum_{g=1}^{\infty} tr(P'_{n,t-s+g}Q_{n,t-s+g})\right)\sum_{g=1}^{\infty} tr(P'_{ng}Q_{ng}) - \sigma_0^4\sum_{g=1}^{\infty} tr(P'_{n,t-s+g}Q_{n,t-s+g}) \times tr(P'_{ng}Q_{ng}),$$

hence,

$$E_{21} = (\mu_4 - 3\sigma_0^4)\sum_{g=1}^{\infty}\sum_{i=1}^{n}(P'_{n,t-s+g}Q_{n,t-s+g})_{ii}(P'_{ng}Q_{ng})_{ii}$$
$$+\sigma_0^4 tr\left(\sum_{g=1}^{\infty}(P'_{n,t-s+g}Q_{n,t-s+g})(P'_{ng}Q_{ng})\right) + \sigma_0^4 tr\left(\sum_{g=1}^{\infty}(P'_{n,t-s+g}Q_{n,t-s+g}Q'_{ng}P_{ng})\right)$$
$$+\sigma_0^4 tr\left(\sum_{g=1}^{\infty}(P'_{n,t-s+g}Q_{n,t-s+g})\right) tr\left(\sum_{g=1}^{\infty} P'_{ng}Q_{ng}\right).$$

Also,

$$E_{22} = E\left[(\sum_{g=1}^{\infty}\sum_{h\neq g}^{\infty}(P_{n,t-s+h}V_{n,s+1-h})'Q_{n,t-s+g}V_{n,s+1-g}) \times (\sum_{g=1}^{\infty}\sum_{h\neq g}^{\infty}(P_{nh}V_{n,s+1-h})'Q_{ng}V_{n,s+1-g})\right]$$
$$= E_{22}^a + E_{22}^b$$

where

$$E_{22}^a = E\left(\sum_{g=1}^{\infty}\sum_{h\neq g}^{\infty}[(P_{n,t-s+h}V_{n,s+1-h})'Q_{n,t-s+g}V_{n,s+1-g} \times (P_{nh}V_{n,s+1-h})'Q_{ng}V_{n,s+1-g}]\right)$$
$$= \sigma_0^4\sum_{g=1}^{\infty}\sum_{h=1}^{\infty} tr\left(P'_{n,t-s+h}Q_{n,t-s+g}Q'_{ng}P_{nh}\right) - \sigma_0^4 tr\left(\sum_{h=1}^{\infty} P'_{n,t-s+h}Q_{n,t-s+h}Q'_{nh}P_{nh}\right)$$
$$= \sigma_0^4\sum_{g=1}^{\infty}\sum_{h=1}^{\infty} tr\left(P_{nh}P'_{n,t-s+h}Q_{n,t-s+g}Q'_{ng}\right) - \sigma_0^4 tr\left(\sum_{h=1}^{\infty} P_{nh}P'_{n,t-s+h}Q_{n,t-s+h}Q'_{nh}\right)$$

and

$$E_{22}^b = E\left(\sum_{g=1}^{\infty}\sum_{h\neq g}^{\infty}[(P_{n,t-s+h}V_{n,s+1-h})'Q_{n,t-s+g}V_{n,s+1-g} \times (P_{ng}V_{n,s+1-g})'Q_{nh}V_{n,s+1-h}]\right)$$
$$= \sigma_0^4\sum_{g=1}^{\infty}\sum_{h=1}^{\infty} tr\left(P'_{n,t-s+h}Q_{n,t-s+g}P'_{ng}Q_{nh}\right) - \sigma_0^4 tr\left(\sum_{h=1}^{\infty} P'_{n,t-s+h}Q_{n,t-s+h}P'_{nh}Q_{nh}\right)$$
$$= \sigma_0^4\sum_{g=1}^{\infty}\sum_{h=1}^{\infty} tr\left(Q_{nh}P'_{n,t-s+h}Q_{n,t-s+g}P'_{ng}\right) - \sigma_0^4 tr\left(\sum_{h=1}^{\infty} Q_{nh}P'_{n,t-s+h}Q_{n,t-s+h}P'_{nh}\right),$$

hence,

$$E_{22} = \sigma_0^4 tr\left[\left(\sum_{h=1}^{\infty} P_{nh}P'_{n,t-s+h}\right)(\sum_{h=1}^{\infty} Q_{n,t-s+h}Q'_{nh}) + \left(\sum_{h=1}^{\infty} Q_{nh}P'_{n,t-s+h}\right)(\sum_{h=1}^{\infty} Q_{n,t-s+h}P'_{nh})\right]$$
$$-\sigma_0^4 tr\left(\sum_{h=1}^{\infty} P_{nh}P'_{n,t-s+h}Q_{n,t-s+h}Q'_{nh}\right) - \sigma_0^4 tr\left(\sum_{h=1}^{\infty} Q_{nh}P'_{n,t-s+h}Q_{n,t-s+h}P'_{nh}\right).$$

Therefore,

$$\begin{aligned} E_2 &= E_{21} + E_{22} = (\mu_4 - 3\sigma_0^4) \sum_{g=1}^{\infty} \sum_{i=1}^{n} (P'_{n,t-s+g} Q_{n,t-s+g})_{ii} (P'_{ng} Q_{ng})_{ii} \\ &\quad + \sigma_0^4 tr \left(\sum_{g=1}^{\infty} P'_{n,t-s+g} Q_{n,t-s+g} \right) tr \left(\sum_{g=1}^{\infty} P'_{ng} Q_{ng} \right) \\ &\quad + \sigma_0^4 tr \left[\left(\sum_{h=1}^{\infty} P_{nh} P'_{n,t-s+h} \right) \left(\sum_{h=1}^{\infty} Q_{n,t-s+h} Q'_{nh} \right) + \left(\sum_{h=1}^{\infty} Q_{nh} P'_{n,t-s+h} \right) \left(\sum_{h=1}^{\infty} Q_{n,t-s+h} P'_{nh} \right) \right]. \end{aligned}$$

As $Cov(\mathbb{U}'_{nt}\mathbb{W}_{nt}, \mathbb{U}'_{ns}\mathbb{W}_{ns}) = E(\mathbb{U}'_{nt}\mathbb{W}_{nt} \times \mathbb{U}'_{ns}\mathbb{W}_{ns}) - E\mathbb{U}'_{nt}\mathbb{W}_{nt} \times E\mathbb{U}'_{ns}\mathbb{W}_{ns}$, using Lemma A.2,

$$\begin{aligned} & Cov(\mathbb{U}'_{nt}\mathbb{W}_{nt}, \mathbb{U}'_{ns}\mathbb{W}_{ns}) \\ = {} & E_1 + E_2 - \sigma_0^4 tr \left(\sum_{h=1}^{\infty} P'_{nh} Q_{nh} \right) tr \left(\sum_{h=1}^{\infty} P'_{nh} Q_{nh} \right) \\ = {} & \sigma_0^4 tr \left[\left(\sum_{h=1}^{\infty} P_{nh} P'_{n,t-s+h} \right) \left(\sum_{h=1}^{\infty} Q_{n,t-s+h} Q'_{nh} \right) + \left(\sum_{h=1}^{\infty} Q_{nh} P'_{n,t-s+h} \right) \left(\sum_{h=1}^{\infty} Q_{n,t-s+h} P'_{nh} \right) \right] \\ & + (\mu_4 - 3\sigma_0^4) \sum_{h=1}^{\infty} \sum_{i=1}^{n} (P'_{n,t-s+h} Q_{n,t-s+h})_{ii} (P'_{nh} Q_{nh})_{ii}. \ \blacksquare \end{aligned}$$

Proof for Lemma A.5

As B_n, $\sum_{h=1}^{\infty} C_{nh}$ and $\sum_{h=1}^{\infty} D_{nh}$ are uniformly bounded in row sum and column sums, their product $(\sum_{h=1}^{\infty} C_{nh})B_n(\sum_{h=1}^{\infty} D_{nh})$ is also a uniformly bounded in row and column sums. Also, for every (i,j) entry of a matrix, $(\sum_{h=1}^{\infty} C_{nh} B_n D_{nh})_{ij} \leq ((\sum_{h=1}^{\infty} C_{nh})B_n(\sum_{h=1}^{\infty} D_{nh}))_{ij}$. Hence, $\sum_{h=1}^{\infty} C_{nh} B_n D_{nh}$ is uniformly bounded in row sum and columns. $\blacksquare$

Proof for Lemma A.6

From Lemma A.4, we have

$$\begin{aligned} Var(\sum_{t=1}^{T} \mathbb{U}'_{nt}\mathbb{W}_{nt}) &= \sum_{t=1}^{T} \sum_{s=1}^{T} Cov(\mathbb{U}'_{nt}\mathbb{W}_{nt}, \mathbb{U}'_{ns}\mathbb{W}_{ns}) \\ &= \left\{ \begin{array}{c} \sigma_0^4 \sum_{t=1}^{T} \sum_{s=1}^{T} tr \left[\left(\sum_{h=1}^{\infty} P_{nh} P'_{n,|t-s|+h} \right) \left(\sum_{h=1}^{\infty} Q_{n,|t-s|+h} Q'_{nh} \right) \right] \\ + \sigma_0^4 \sum_{t=1}^{T} \sum_{s=1}^{T} tr \left[\left(\sum_{h=1}^{\infty} Q_{nh} P'_{n,|t-s|+h} \right) \left(\sum_{h=1}^{\infty} Q_{n,|t-s|+h} P'_{nh} \right) \right] \\ + (\mu_4 - 3\sigma_0^4) \sum_{t=1}^{T} \sum_{s=1}^{T} \sum_{h=1}^{\infty} \sum_{i=1}^{n} (P'_{n,|t-s|+h} Q_{n,|t-s|+h})_{ii} (P'_{nh} Q_{nh})_{ii} \end{array} \right\}. \end{aligned}$$

Denoting $\mathcal{A}_{n,PP} = \sum_{h=1}^{\infty} P_n^h P_n^{h\prime}$ and $\mathcal{A}_{n,QQ} = \sum_{h=1}^{\infty} Q_n^h Q_n^{h\prime}$, both of which are uniformly bounded in both row and column sums by Lemma A.5, we have

$$\begin{aligned}
& \sigma_0^4 \sum_{t=1}^{T}\sum_{s=1}^{T} tr\left(\sum_{h=1}^{\infty} P_{nh}P'_{n,|t-s|+h}\right)\left(\sum_{h=1}^{\infty} Q_{n,|t-s|+h}Q'_{nh}\right) \\
= & \sigma_0^4 \sum_{t=1}^{T}\sum_{s=1}^{T} tr\left[B_{1n}\left(\sum_{h=1}^{\infty} P_n^h P_n^{\prime h}\right) P_n^{\prime|t-s|} B'_{1n} B_{2n} Q_n^{|t-s|}\left(\sum_{h=1}^{\infty} Q_n^h Q_n^{\prime h}\right) B'_{2n}\right] \\
= & \sigma_0^4 \sum_{t=1}^{T}\sum_{s=1}^{T} tr\left[B_{1n}\mathcal{A}_{n,PP} P_n^{\prime|t-s|} B'_{1n} B_{2n} Q_n^{|t-s|}\mathcal{A}_{n,QQ} B'_{2n}\right] \\
= & \sigma_0^4 tr\left[\left(\sum_{t=1}^{T}\sum_{s=1}^{T} P_n^{\prime|t-s|} B'_{1n} B_{2n} Q_n^{|t-s|}\right)\mathcal{A}_{n,QQ} B'_{2n} B_{1n}\mathcal{A}_{n,PP}\right].
\end{aligned}$$

Consider $\sum_{t=1}^{T}\sum_{s=1}^{T} P_n^{\prime|t-s|}\mathcal{B}_n Q_n^{|t-s|}\mathcal{C}_n$, $\mathcal{B}_n$ and $\mathcal{C}_n$ are uniformly bounded in row and column sums, then

$$\sum_{t=1}^{T}\sum_{s=1}^{T} P_n^{\prime|t-s|}\mathcal{B}_n Q_n^{|t-s|}\mathcal{C}_n = T\mathcal{B}_n\mathcal{C}_n + 2(T-1)P'_n\mathcal{B}_n Q_n\mathcal{C}_n + 2(T-2)P_n^{\prime 2}\mathcal{B}_n Q_n^2\mathcal{C}_n + \cdots + 2P_n^{\prime T-1}\mathcal{B}_n Q_n^{T-1}\mathcal{C}_n.$$

Hence

$$\left|tr\left(\sum_{t=1}^{T}\sum_{s=1}^{T} P_n^{\prime|t-s|}\mathcal{B}_n Q_n^{|t-s|}\mathcal{C}_n\right)\right| \le 2T\cdot tr\left[abs(\mathcal{B}_n)abs(\mathcal{C}_n) + \sum_{h=1}^{\infty}(abs(P'_n))^h abs(\mathcal{B}_n)(abs(Q_n))^h abs(\mathcal{C}_n)\right], \tag{A.18}$$

which is of order $O(nT)$ by the uniform boundedness of the matrix within the trace operator. Hence,

$$\sigma_0^4 \sum_{t=1}^{T}\sum_{s=1}^{T} tr\left(\sum_{h=1}^{\infty} P_{nh}P'_{n,|t-s|+h}\right)\left(\sum_{h=1}^{\infty} Q_{n,|t-s|+h}Q'_{nh}\right) = O(nT)$$

and similarly, $\sigma_0^4 \sum_{t=1}^{T}\sum_{s=1}^{T}\left(\sum_{h=1}^{\infty} Q_{nh}P'_{n,|t-s|+h}\right)\left(\sum_{h=1}^{\infty} Q_{n,|t-s|+h}P'_{nh}\right) = O(nT)$.

Now consider the term $\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{h=1}^{\infty}\sum_{i=1}^{n}(P'_{n,|t-s|+h}Q_{n,|t-s|+h})_{ii}(P'_{nh}Q_{nh})_{ii}$. As $\sum_{h=1}^{\infty} abs(P'_{nh}Q_{nh})$ is uniformly bounded in both row and column sums by Lemma A.5, there exists a finite constant c such that $\sup_n \sup_{i=1,\cdots,n}\sum_{h=1}^{\infty}|(P'_{nh}Q_{nh})_{ii}| \le c$. Hence

$$\left|\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{h=1}^{\infty}\sum_{i=1}^{n}(P'_{n,|t-s|+h}Q_{n,|t-s|+h})_{ii}(P'_{nh}Q_{nh})_{ii}\right|$$
$$\leq c\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{h=1}^{\infty}\sum_{i=1}^{n}\left|(P'_{n,|t-s|+h}Q_{n,|t-s|+h})_{ii}\right| \leq c\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{h=1}^{\infty} tr\left(abs(P'_{n,|t-s|+h})abs(Q_{n,|t-s|+h})\right)$$
$$\leq c\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{h=1}^{\infty} tr\left(abs(P_n'^{|t-s|})\left(abs(P_n'^{h}B'_{1n})abs(B_{2n}Q_n^h)\right)abs(Q_n^{|t-s|})\right)$$
$$\leq c\sum_{t=1}^{T}\sum_{s=1}^{T} tr\left(abs(P_n'^{|t-s|})\left(\sum_{h=1}^{\infty}\left(abs(P_n'^{h})abs(B'_{1n})abs(B_{2n})abs(Q_n^h)\right)\right)abs(Q_n^{|t-s|})\right),$$

which is of order $O(nT)$ from (A.18) because $\sum_{h=1}^{\infty}\left(abs(P_n'^{h})abs(B'_{1n})abs(B_{2n})abs(Q_n^h)\right)$ is row and column sum bounded by Lemma A.5. Hence, $Var(\sum_{t=1}^{T}\mathbb{U}'_{nt}\mathbb{W}_{nt}) = O(nT)$. ■

Proof for Equation (A.7) in Theorem A.7

First, from Lemma A.2, $E\frac{1}{nT}\sum_{t=1}^{T}\mathbb{U}'_{nt}\mathbb{W}_{nt} = \frac{1}{n}\sigma_0^2 tr\,(\sum_{h=1}^{\infty}P'_{nh}Q_{nh})$. Second, from Lemma A.6, $Var(\sum_{t=1}^{T}\mathbb{U}'_{nt}\mathbb{W}_{nt}) = O(nT)$. Using Chebyshev's inequality, the result follows. ■

Proof for Equation (A.8) in Theorem A.7

From Lemma A.1, $\bar{\mathbb{U}}_{nT} = \sum_{h=1}^{\infty}\ddot{P}_{nh}V_{n,T+1-h}$ and $\bar{\mathbb{W}}_{nT} = \sum_{h=1}^{T}\ddot{Q}_{nh}V_{n,T+1-h}$ where

$$\ddot{P}_{nh} = \begin{cases} \frac{1}{T}(P_{n1}+P_{n2}+\cdots+P_{nh}) = \frac{1}{T}\sum_{g=1}^{h}P_{ng} & \text{for } h\leq T \\ \frac{1}{T}\sum_{g=1}^{T}P_{n,h-T+g} & \text{for } h>T \end{cases} \tag{A.19}$$

and $\ddot{Q}_{nh}$ has the same pattern.

First, using Lemma A.2, $E\bar{\mathbb{U}}'_{nT}\bar{\mathbb{W}}_{nT} = \sigma_0^2 tr\left(\sum_{h=1}^{\infty}\ddot{P}_{nh}\ddot{Q}'_{nh}\right)$. Second, using Lemma A.4,

$$\begin{aligned} Var(\bar{\mathbb{U}}'_{nT}\bar{\mathbb{W}}_{nT}) &= Cov(\bar{\mathbb{U}}'_{nT}\bar{\mathbb{W}}_{nT}, \bar{\mathbb{U}}'_{nT}\bar{\mathbb{W}}_{nT}) \\ &= \sigma_0^4 tr\left[\left(\sum_{h=1}^{\infty}\ddot{P}_{nh}\ddot{P}'_{nh}\right)\left(\sum_{h=1}^{\infty}\ddot{Q}_{nh}\ddot{Q}'_{nh}\right)+\left(\sum_{h=1}^{\infty}\ddot{Q}_{nh}\ddot{P}'_{nh}\right)\left(\sum_{h=1}^{\infty}\ddot{Q}_{nh}\ddot{P}'_{nh}\right)\right] \\ &\quad +(\mu_4-3\sigma_0^4)\sum_{h=1}^{\infty}\sum_{i=1}^{n}(\ddot{P}'_{nh}\ddot{Q}_{nh})_{ii}(\ddot{P}'_{nh}\ddot{Q}_{nh})_{ii}. \end{aligned}$$

In order to get the orders of $E\bar{\mathbb{U}}'_{nT}\bar{\mathbb{W}}_{nT}$ and $Var(\bar{\mathbb{U}}'_{nT}\bar{\mathbb{W}}_{nT})$, we consider the order of the elements in $(\sum_{h=1}^{\infty}\ddot{Q}_{nh}\ddot{P}'_{nh})$; the analysis of the elements of $\sum_{h=1}^{\infty}\ddot{P}_{nh}\ddot{P}'_{nh}$ and $\sum_{h=1}^{\infty}\ddot{Q}_{nh}\ddot{Q}'_{nh}$ are analogous.

As $\sum_{h=1}^{\infty}\ddot{Q}_{nh}\ddot{P}'_{nh} = \sum_{h=1}^{T}\ddot{Q}_{nh}\ddot{P}'_{nh} + \sum_{h=T+1}^{\infty}\ddot{Q}_{nh}\ddot{P}'_{nh}$, from (A.19), we have for $p=1$ (the column sum norm) and $p=\infty$ (the row sum norm),

$$\left\|\sum_{h=1}^{T}\ddot{Q}_{nh}\ddot{P}'_{nh}\right\|_p = \left\|\sum_{h=1}^{T}(\frac{1}{T}\sum_{g=1}^{h}Q_{ng})(\frac{1}{T}\sum_{g=1}^{h}P'_{ng})\right\|_p \leq \frac{1}{T}\left\|\sum_{g=1}^{\infty}abs(Q_{ng})\right\|_p \cdot \left\|\sum_{g=1}^{\infty}abs(P'_{ng})\right\|_p = O\left(\frac{1}{T}\right),$$

and

$$\begin{aligned}\left\|\sum_{h=T+1}^{\infty}\ddot{Q}_{nh}\ddot{P}'_{nh}\right\|_p &= \left\|\sum_{h=T+1}^{\infty}(\frac{1}{T}\sum_{g=1}^{T}Q_{n,h-T+g})(\frac{1}{T}\sum_{g=1}^{T}P_{n,h-T+g})'\right\|_p \\ &= \left\|\sum_{h=T+1}^{\infty}B_{2n}(\frac{1}{T}\sum_{g=1}^{T}Q_n^g)Q_n^{h-T}P_n'^{h-T}(\frac{1}{T}\sum_{g=1}^{T}P_n'^g)B'_{1n}\right\|_p \\ &\leq \frac{1}{T^2}\|B_{2n}\|\cdot\left\|\sum_{g=1}^{T}abs(Q_n^g)\right\|_p\cdot\left\|\sum_{h=1}^{\infty}abs(Q_n^hP_n'^h)\right\|_p\cdot\left\|\sum_{g=1}^{T}abs(P_n'^g)\right\|_p\cdot\|B'_{1n}\|_p = O\left(\frac{1}{T^2}\right),\end{aligned}$$

because $\sum_{g=1}^{\infty}abs(P_n'^g)$, $\sum_{g=1}^{\infty}abs(Q_n^g)$, $\sum_{h=1}^{\infty}abs(Q_n^hP_n'^h)$, B_{1n} and B_{2n} are uniformly bounded in both row and column sum norms. This implies, in particular, that all elements of $\sum_{h=1}^{\infty}\ddot{Q}_{nh}\ddot{P}'_{nh}$ are of the order $O(\frac{1}{T})$.

Furthermore, as

$$\begin{aligned}(\sum_{h=1}^{\infty}\ddot{Q}_{nh}\ddot{P}'_{nh})(\sum_{h=1}^{\infty}\ddot{Q}_{nh}\ddot{P}'_{nh}) &= (\sum_{h=1}^{T}\ddot{Q}_{nh}\ddot{P}'_{nh})(\sum_{h=1}^{T}\ddot{Q}_{nh}\ddot{P}'_{nh})+(\sum_{h=1}^{T}\ddot{Q}_{nh}\ddot{P}'_{nh})(\sum_{h=T+1}^{\infty}\ddot{Q}_{nh}\ddot{P}'_{nh}) \\ &\quad +(\sum_{h=T+1}^{\infty}\ddot{Q}_{nh}\ddot{P}'_{nh})(\sum_{h=1}^{T}\ddot{Q}_{nh}\ddot{P}'_{nh})+(\sum_{h=T+1}^{\infty}\ddot{Q}_{nh}\ddot{P}'_{nh})(\sum_{h=T+1}^{\infty}\ddot{Q}_{nh}\ddot{P}'_{nh}),\end{aligned}$$

it follows that $\left\|(\sum_{h=1}^{\infty}\ddot{Q}_{nh}\ddot{P}'_{nh})(\sum_{h=1}^{\infty}\ddot{Q}_{nh}\ddot{P}'_{nh})\right\|_p \leq O\left(\frac{1}{T^2}\right)+O\left(\frac{1}{T^3}\right)+O\left(\frac{1}{T^4}\right)=O\left(\frac{1}{T^2}\right)$. This implies, in particular, that all elements of $(\sum_{h=1}^{\infty}\ddot{Q}_{nh}\ddot{P}'_{nh})^2$ are of the order $O(\frac{1}{T^2})$.

Similarly, we can get the order of the elements of $\sum_{h=1}^{\infty}\ddot{P}_{nh}\ddot{P}'_{nh}$, $\sum_{h=1}^{\infty}\ddot{Q}_{nh}\ddot{Q}'_{nh}$ and their products respectively, implying that

$$E(\bar{\mathbb{U}}'_{nT}\bar{\mathbb{W}}_{nT}) = \sigma^2 tr(\sum_{h=1}^{\infty}\ddot{P}_{nh}\ddot{Q}'_{nh}) = O\left(\frac{n}{T}\right),$$

$$tr\left[\left(\sum_{h=1}^{\infty}\ddot{P}_{nh}\ddot{P}'_{nh}\right)\left(\sum_{h=1}^{\infty}\ddot{Q}_{nh}\ddot{Q}'_{nh}\right)+\left(\sum_{h=1}^{\infty}\ddot{Q}_{nh}\ddot{P}'_{nh}\right)\left(\sum_{h=1}^{\infty}\ddot{Q}_{nh}\ddot{P}'_{nh}\right)\right] = O\left(\frac{n}{T^2}\right),$$

and

$$\left|\sum_{h=1}^{\infty}\sum_{i=1}^{n}(\ddot{P}'_{nh}\ddot{Q}_{nh})_{ii}(\ddot{P}'_{nh}\ddot{Q}_{nh})_{ii}\right| \leq \left(\sup_{i,n}\left|(\ddot{P}'_{nh}\ddot{Q}_{nh})_{ii}\right|\right)\cdot\left(\sum_{i=1}^{n}\left|\sum_{h=1}^{\infty}(\ddot{P}'_{nh}\ddot{Q}_{nh})_{ii}\right|\right) = O\left(\frac{1}{T}\right)\cdot O\left(\frac{n}{T}\right) = O\left(\frac{n}{T^2}\right).$$

To sum up, $E\frac{1}{n}\bar{\mathbb{U}}'_{nT}\bar{\mathbb{W}}_{nT} = O\left(\frac{1}{T}\right)$ and $Var(\frac{1}{n}\bar{\mathbb{U}}'_{nT}\bar{\mathbb{W}}_{nT}) = O\left(\frac{1}{nT^2}\right)$. Hence, $\frac{1}{n}\bar{\mathbb{U}}'_{nT}\bar{\mathbb{W}}_{nT} - E\frac{1}{n}\bar{\mathbb{U}}'_{nT}\bar{\mathbb{W}}_{nT} = O_p\left(\frac{1}{\sqrt{nT^2}}\right)$. ■

Proof for Equation (A.9) in Theorem A.7

$\frac{1}{nT}\sum_{t=1}^{T}\tilde{\mathbb{U}}'_{nt}\tilde{\mathbb{W}}_{nt} = \frac{1}{nT}\sum_{t=1}^{T}\mathbb{U}'_{nt}\mathbb{W}_{nt} - \frac{1}{n}\bar{\mathbb{U}}'_{nT}\bar{\mathbb{W}}_{nT}$. Using (A.7) and (A.8), the result follows. ■

Proof for Theorem A.8

To prove (A.10):

First, as $\tilde{D}'_{nt}$ is nonstochastic, $E\frac{1}{nT}\sum_{t=1}^{T}\tilde{D}'_{nt}\tilde{\mathbb{U}}_{nt} = E\frac{1}{nT}\sum_{t=1}^{T}\tilde{D}'_{nt}\mathbb{U}_{nt} = 0$. Second, as $\tilde{D}'_{nt}$ is the deviation from the sample mean of D'_{nt}, $\frac{1}{nT}\sum_{t=1}^{T}\tilde{D}'_{nt}\tilde{\mathbb{U}}_{nt} = \frac{1}{nT}\sum_{t=1}^{T}\tilde{D}'_{nt}\mathbb{U}_{nt}$. Hence,

$$Var\left(\frac{1}{nT}\sum_{t=1}^{T}\tilde{D}'_{nt}\mathbb{U}_{nt}\right) = \frac{1}{n^2T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}Cov(\tilde{D}'_{nt}\mathbb{U}_{nt}, \tilde{D}'_{ns}\mathbb{U}_{ns}) = \frac{1}{n^2T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\tilde{D}'_{nt}E(\mathbb{U}_{nt}\mathbb{U}'_{ns})\tilde{D}_{ns}$$

where $E(\mathbb{U}_{nt}\mathbb{U}'_{ns}) = \sigma_0^2 B_{1n}(\sum_{h=1}^{\infty}P_n^{t-s+h}P_n'^h)B'_{1n}$ for $t \geq s$ by Lemma A.2.

As elements of D_{nt} are bounded, denote the maximum element of $\tilde{D}_{nt}$ for all n and t as $\tilde{d}_{\max}$. Furthermore, let M be a constant such that for any row and column sum uniformly bounded matrix, its row and column sums are smaller than M. Then, for $t \geq s$,

$$\begin{aligned}\left|Cov(\tilde{D}'_{nt}\mathbb{U}_{nt}, \tilde{D}'_{ns}\mathbb{U}_{ns})\right| &\leq \sigma_0^2\tilde{x}_{\max}^2 M^2\sum_{i=1}^{n}\sum_{j=1}^{n}abs(\sum_{h=1}^{\infty}P_n^{t-s+h}P_n'^h)_{ij} \\ &\leq \sigma_0^2\tilde{x}_{\max}^2 M^2\sum_{i=1}^{n}\sum_{j=1}^{n}\left(abs(P_n^{t-s})\times abs(\sum_{h=1}^{\infty}P_n^hP_n'^h)\right)_{ij}.\end{aligned}$$

Hence, $\sum_{t=1}^{T}\sum_{s=1}^{T}\left|Cov(\tilde{D}'_{nt}\mathbb{U}_{nt}, \tilde{D}'_{ns}\mathbb{U}_{ns})\right| \leq 2\sigma_0^2 T\tilde{x}_{\max}^2 M^2\sum_{i=1}^{n}\sum_{j=1}^{n}\left((\sum_{h=1}^{\infty}abs(P_n^h))\times abs(\sum_{h=1}^{\infty}P_n^hP_n'^h)\right)_{ij}$.

As $\sum_{h=1}^{\infty}abs(P_n^h)$ is uniformly bounded in row and column sums, $abs(\sum_{h=1}^{\infty}P_n^hP_n'^h)$ is also uniformly bounded in row and column sums. Hence, $\sum_{t=1}^{T}\sum_{s=1}^{T}\left|Cov(\tilde{D}'_{nt}\mathbb{U}_{nt}, \tilde{D}'_{ns}\mathbb{U}_{ns})\right| \leq 2\sigma_0^2 nT\tilde{x}_{\max}^2 M^3$, which implies that elements of $Var\left(\frac{1}{nT}\sum_{t=1}^{T}\tilde{D}'_{nt}\mathbb{U}_{nt}\right)$ are $O(\frac{1}{nT})$. Using Chebyshev's inequality, $\frac{1}{nT}\sum_{t=1}^{T}\tilde{D}'_{nt}\tilde{\mathbb{U}}_{nt} = O\left(\frac{1}{\sqrt{nT}}\right)$.

To prove (A.11):

By Lemma A.1, $Var(\frac{1}{n}\bar{D}'_{nT}\bar{\mathbb{U}}_{nT}) = \frac{1}{n^2}\bar{D}'_{nT}E(\bar{\mathbb{U}}_{nT}\bar{\mathbb{U}}'_{nT})\bar{D}_{nT} = \frac{\sigma_0^2}{n^2}\bar{D}'_{nT}(\sum_{h=1}^{\infty}\ddot{P}_n^h\ddot{P}_n'^h)\bar{D}_{nT}$. Hence, $Var(\frac{1}{n}\bar{D}'_{nT}\bar{\mathbb{U}}_{nT}) \leq \frac{\sigma_0^2}{n^2}\left\|\bar{D}'_{nT}\right\|_\infty \cdot \left\|(\sum_{h=1}^{\infty}\ddot{P}_n^h\ddot{P}_n'^h)\right\|_\infty \cdot \left\|\bar{D}_{nT}\right\|_\infty = O\left(\frac{1}{nT}\right)$ because $\left\|\bar{D}'_{nT}\right\|_\infty = \left\|\bar{D}_{nT}\right\|_1 = O(n)$, $\left\|\bar{D}_{nT}\right\|_\infty = O(1)$, and $\left\|(\sum_{h=1}^{\infty}\ddot{P}_n^h\ddot{P}_n'^h)\right\|_\infty = O\left(\frac{1}{T}\right)$ in the proof of (A.8) in Theorem A.7. ■

Proof for Lemma A.9

Using the formula for the variance of quadratic forms of disturbance (see Lemma A.3), we can get the result for (A.12). For (A.13), as $Var((\bar{V}_{nT})_i) = O\left(\frac{1}{T}\right)$ and $E((\bar{V}_{nT})_i^4) = O\left(\frac{1}{T^2}\right)$, the variance of $\frac{1}{n}\bar{V}'_{nT}B_n\bar{V}_{nT}$ has the order $O\left(\frac{1}{nT^2}\right)$. (A.14) follows because $\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}'_{nt}B_n\tilde{V}_{nt} = \frac{1}{nT}\sum_{t=1}^{T}V'_{nt}B_nV_{nt} - \frac{1}{n}\bar{V}'_{nT}B_n\bar{V}_{nT}$. ■

Proof of Lemma A.10

Let e_{ni} be the ith unit vector of R^n. Then $(\mathbb{U}_{n,t-1})_i = e'_{ni}\sum_{h=1}^{\infty}P_{nh}V_{n,t-h} = \sum_{h=1}^{\infty}V'_{n,t-h}P'_{nh}e_{ni}$. Hence,

$$\begin{aligned} E([(\mathbb{U}_{n,t-1})_i]^4) &= \sum_{h=1}^{\infty}\sum_{k=1}^{\infty}\sum_{l=1}^{\infty}\sum_{s=1}^{\infty}E[V'_{n,t-h}(P'_{nh}e_{ni}e'_{ni}P_{nk})V_{n,t-k}\cdot V'_{n,t-l}(P'_{nl}e_{ni}e'_{ni}P_{ns})V_{n,t-s}] \\ &= \sum_{h=1}^{\infty}E[V'_{n,t-h}(P'_{nh}e_{ni}e'_{ni}P_{nh})V_{n,t-h}\cdot V'_{n,t-h}(P'_{nh}e_{ni}e'_{ni}P_{nh})V_{n,t-h}] \\ &\quad +\sum_{h=1}^{\infty}\sum_{k\neq h}^{\infty}E[V'_{n,t-h}(P'_{nh}e_{ni}e'_{ni}P_{nh})V_{n,t-h}\cdot V'_{n,t-k}(P'_{nk}e_{ni}e'_{ni}P_{nk})V_{n,t-k}] \\ &\quad +2\sum_{h=1}^{\infty}\sum_{k\neq h}^{\infty}E([V'_{n,t-h}(P'_{nh}e_{ni}e'_{ni}P_{nk})V_{n,t-k}\cdot V'_{n,t-h}(P'_{nh}e_{ni}e'_{ni}P_{nk})V_{n,t-k}]). \end{aligned}$$

By using the formula for $E[(V'_{nt}B_{1n}V_{ns})(V'_{ng}B_{2n}V_{nh})$ in Lemma A.3, one has

$$\begin{aligned} E([(\mathbb{U}_{n,t-1})_i]^4) &= (\mu_4-3\sigma_0^4)\sum_{h=1}^{\infty}\sum_{j=1}^{n}[(P'_{nh}e_{ni}e'_{ni}P_{nh})_{jj}]^2+\sigma_0^4\sum_{h=1}^{\infty}\{tr^2(P'_{nh}e_{ni}e'_{ni}P_{nh})+2tr[(P'_{nh}e_{ni}e'_{ni}P_{nh})^2]\} \\ &\quad +\sigma_0^4\sum_{h=1}^{\infty}\sum_{k\neq h}^{\infty}tr(P'_{nh}e_{ni}e'_{ni}P_{nh})tr(P'_{nk}e_{ni}e'_{ni}P_{nk})+2\sigma_0^4\sum_{h=1}^{\infty}\sum_{k\neq h}^{\infty}tr(P'_{nh}e_{ni}e'_{ni}P_{nk}\cdot P'_{nk}e_{ni}e'_{ni}P_{nh}) \\ &= (\mu_4-3\sigma_0^4)\sum_{h=1}^{\infty}\sum_{j=1}^{n}[(P'_{nh}e_{ni}e'_{ni}P_{nh})_{jj}]^2 \\ &\quad +\sigma_0^4\sum_{h=1}^{\infty}\sum_{k=1}^{\infty}tr(P'_{nh}e_{ni}e'_{ni}P_{nh})tr(P'_{nk}e_{ni}e'_{ni}P_{nk})+2\sigma_0^4\sum_{h=1}^{\infty}\sum_{k=1}^{\infty}tr(P'_{nh}e_{ni}e'_{ni}P_{nk}\cdot P'_{nk}e_{ni}e'_{ni}P_{nh}) \\ &= (\mu_4-3\sigma_0^4)\sum_{h=1}^{\infty}\sum_{j=1}^{n}(e'_{ni}P_{nh}e_{nj})^4+3\sigma_0^4(\sum_{h=1}^{\infty}e'_{ni}P_{nh}P'_{nh}e_{ni})^2. \ ■ \end{aligned}$$

Proof for Theorem A.11

Let $\ddot{\mathbb{U}}_{nT}=\bar{\mathbb{U}}_{nT,-1}$ and $\ddot{\mathbb{W}}_{nT}=\bar{V}_{nT}$. As $\mathbb{U}_{nt}=\sum_{h=1}^{\infty}P_{nh}V_{n,t+1-h}$, we have $\ddot{\mathbb{U}}_{nT}=\sum_{h=1}^{\infty}\ddot{P}_{nh}V_{n,T+1-h}$ and $\ddot{\mathbb{W}}_{nT}=\sum_{h=1}^{\infty}\ddot{Q}_{nh}V_{n,T+1-h}$ where

$$\ddot{P}_{nh}=\begin{cases}0 & \text{for } h=1\\ \frac{1}{T}(P_{n1}+P_{n2}+\cdots+P_{n,h-1})=\frac{1}{T}\sum_{g=1}^{h-1}P_{ng} & \text{for } 2\le h\le T\\ \frac{1}{T}\sum_{g=0}^{T-1}P_{n,h-T+g} & \text{for } h>T\end{cases}\quad\text{and }\ddot{Q}_{nh}=\begin{cases}\frac{1}{T}I_n & \text{for } h\le T\\ 0 & \text{for } h>T\end{cases}.$$

First, using Lemma A.2, $E\ddot{\mathbb{U}}_{nT}'\ddot{\mathbb{W}}_{nT}=\sigma_0^2 tr\left(\sum_{h=1}^{\infty}\ddot{P}_{nh}\ddot{Q}_{nh}'\right)$. As $\ddot{Q}_{nh}$ is specified above, $E\bar{\mathbb{U}}_{nT,-1}'\bar{V}_{nT}=\frac{1}{T}\sigma_0^2 tr\left(\sum_{h=1}^{T}\ddot{P}_{nh}\right)$. Also, as $\ddot{P}_{nh}$ is specified above and $P_{nh}=B_{1n}P_n^h$,

$$\sum_{h=1}^{T}\ddot{P}_{nh}=B_{1n}\frac{1}{T}\sum_{h=2}^{T}\left(\sum_{g=1}^{h-1}P_n^g\right)=B_{1n}\sum_{g=1}^{\infty}P_n^g+R_{nT}\tag{A.20}$$

where $R_{nT}=B_{1n}\frac{1}{T}\sum_{h=2}^{T}\left(\sum_{g=1}^{h-1}P_n^g\right)-B_{1n}\sum_{g=1}^{\infty}P_n^g$. As $tr(B_{1n}\sum_{g=1}^{\infty}P_n^g)$ is $O(n)$, it remains to investigate the order of elements of the remainder R_{nT}. Because $R_{nT}=B_{1n}\frac{1}{T}\sum_{h=2}^{T}\left(\sum_{g=1}^{h-1}P_n^g\right)-B_{1n}\sum_{g=1}^{\infty}P_n^g=-B_{1n}\frac{1}{T}\sum_{h=0}^{T-1}P_n^h\sum_{l=1}^{\infty}P_n^l$, it follows that $\|R_{nT}\|_\infty\le\frac{1}{T}\|B_{1n}\|_\infty\cdot\left\|\sum_{h=0}^{T-1}P_n^h\right\|_\infty\cdot\left\|\sum_{l=1}^{\infty}P_n^l\right\|_\infty=O(\frac{1}{T})$, uniformly in n, because $\left\|\sum_{h=0}^{T-1}P_n^h\right\|_\infty\le\left\|\sum_{h=0}^{\infty}abs\left(P_n^h\right)\right\|_\infty$ and $\sup_n\left\|\sum_{h=0}^{\infty}abs\left(P_n^h\right)\right\|_\infty<\infty$. This implies that elements of R_{nT} are of order $O(\frac{1}{T})$ uniformly in n. Hence,

$$E\bar{\mathbb{U}}_{nT,-1}'\bar{V}_{nT}=\frac{1}{T}\sigma_0^2 tr\left(B_{1n}\sum_{g=1}^{\infty}P_n^g+R_{nT}\right)=\frac{1}{T}\sigma_0^2 tr\left(B_{1n}\sum_{g=1}^{\infty}P_n^g\right)+\frac{1}{T}\sigma_0^2 tr\left(R_{nT}\right)=O\left(\frac{n}{T}\right)+O\left(\frac{n}{T^2}\right).$$

Second, using Lemma A.4,

$$\begin{aligned}Var(\bar{\mathbb{U}}_{nT,-1}'\bar{V}_{nT}) &= Cov(\ddot{\mathbb{U}}_{nT}'\ddot{\mathbb{W}}_{nT},\ddot{\mathbb{U}}_{nT}'\ddot{\mathbb{W}}_{nT})\\ &= \sigma_0^4 tr\left[\left(\sum_{h=1}^{\infty}\ddot{P}_{nh}\ddot{P}_{nh}'\right)\left(\sum_{h=1}^{\infty}\ddot{Q}_{nh}\ddot{Q}_{nh}'\right)+\left(\sum_{h=1}^{\infty}\ddot{Q}_{nh}\ddot{P}_{nh}'\right)\left(\sum_{h=1}^{\infty}\ddot{Q}_{nh}\ddot{P}_{nh}'\right)\right]\\ &\quad +(\mu_4-3\sigma_0^4)\sum_{h=1}^{\infty}\sum_{i=1}^{n}(\ddot{P}_{nh}'\ddot{Q}_{nh})_{ii}(\ddot{P}_{nh}'\ddot{Q}_{nh})_{ii}.\end{aligned}$$

For the first term, because of the special form of $\ddot{Q}_{nh}$, it becomes $\sigma_0^4 tr\left[\frac{1}{T^2}\left(\sum_{h=1}^{\infty}\ddot{P}_{nh}\ddot{P}_{nh}'\right)+\frac{1}{T^2}\left(\sum_{h=1}^{T}\ddot{P}_{nh}'\right)\left(\sum_{h=1}^{T}\ddot{P}_{nh}'\right)\right]$, which is $O\left(\frac{n}{T^2}\right)$. For the second term, it is $\frac{1}{T^2}(\mu_4-3\sigma_0^4)\sum_{h=1}^{T}\sum_{i=1}^{n}\left[(\ddot{P}_{nh}')_{ii}\right]^2$; also, its order is $O(\frac{n}{T^3})$ because elements of $\sum_{h=1}^{\infty}\ddot{P}_{nh}\ddot{P}_{nh}'$ have order $O(\frac{1}{T})$ uniformly in n, as shown in the proof of (A.8) in Theorem A.7. Also, each element of $\ddot{P}_{nh}$ is of $O(\frac{1}{T})$ uniformly in n by construction. Therefore, $Var(\bar{\mathbb{U}}_{nT,-1}'\bar{V}_{nT})=O\left(\frac{n}{T^2}\right)$ and hence, $Var(\sqrt{\frac{T}{n}}\bar{\mathbb{U}}_{nT,-1}'\bar{V}_{nT})=O\left(\frac{1}{T}\right)$. Using Chebyshev's inequality, $\sqrt{\frac{T}{n}}\left(\bar{\mathbb{U}}_{nT,-1}'\bar{V}_{nT}-E\left(\bar{\mathbb{U}}_{nT,-1}'\bar{V}_{nT}\right)\right)=O_p\left(\frac{1}{\sqrt{T}}\right)$. ■

Proof of Lemma A.12

As B_n^- is the lower diagonal matrix constructed from B_n by deleting the diagonal and the upper triangle entries,

$$B_n^- = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ b_{n21} & 0 & \cdots & 0 & 0 \\ b_{n31} & b_{n32} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ b_{nn1} & b_{nn2} & \cdots & b_{nn,n-1} & 0 \end{pmatrix}.$$

As B_n is uniformly bounded in both row and column sums, B_n^- will also be uniformly bounded in both row and column sums. We note that

$$\sum_{i=1}^{n} k_{ni}(\sum_{j=1}^{i-1} b_{nij}v_{jt}) = K_n' B_n^- V_{nt},$$

$$\sum_{i=1}^{n} u_{n,t-1,i}(\sum_{j=1}^{i-1} b_{nij}v_{jt}) = U_{n,t-1}' B_n^- V_{nt},$$

and

$$\sum_{i=1}^{n} k_{ni}u_{n,t-1,i} = K_n' U_{n,t-1}.$$

Furthermore, because $(\sum_{j=1}^{i-1} b_{nij}v_{jt})^2 = V_{nt}' B_n^{-'} e_{ni} \cdot e_{ni}' B_n^- V_{nt}$ with e_{ni} being the ith unit vector of R^n and $\sum_{i=1}^{n} e_{ni}e_{ni}' = I_n$, we have $\sum_{i=1}^{n}(\sum_{j=1}^{i-1} b_{nij}v_{jt})^2 = V_{nt}' B_n^{-'} B_n^- V_{nt}$.

For (a), $\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}(\sum_{j=1}^{i-1} b_{nij}v_{jt})^2 - \frac{1}{2}\sigma_0^2[tr(B_n^2) - vec_D'(B_n)vec_D(B_n)] = \frac{1}{nT}\sum_{t=1}^{T}[V_{nt}' B_n^{-'} B_n^- V_{nt} - \sigma_0^2 tr(B_n^{-'} B_n^-)]$. It's variance is

$$\frac{1}{(nT)^2}\sum_{t=1}^{T} Var(V_{nt}' B_n^{-'} B_n^- V_{nt}) = \frac{T}{(nT)^2}\left\{(\mu_4 - 3\sigma_0^4)\sum_{i=1}^{n}[(B_n^{-'} B_n^-)_{ii}]^2 + 2\sigma_0^4 tr[(B_n^{-'} B_n^-)^2]\right\} = O\left(\frac{1}{nT}\right),$$

because $B_n^{-'} B_n^-$ is uniformly bounded in both row and column sums. Thus (a) follows by Chebyshev's inequality.

For (b), $\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n} k_{ni}(\sum_{j=1}^{i-1} b_{nij}v_{jt}) = \frac{1}{nT}\sum_{t=1}^{T} K_n' B_n^- V_{nt}$. Its variance is

$$\frac{1}{(nT)^2}\sum_{t=1}^{T} Var(K_n' B_n^- V_{nt}) = \frac{\sigma_0^2 T}{(nT)^2} K_n' B_n^- B_n^{-'} K_n = O\left(\frac{1}{nT}\right),$$

because elements of K_n are uniformly bounded and $B_n^- B_n^{-'}$ is uniformly bounded in both row and column sums. This proves (b).

For (c), $\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}u_{n,t-1,i}(\sum_{j=1}^{i-1}b_{nij}v_{jt})=\frac{1}{nT}\sum_{t=1}^{T}\mathbb{U}_{n,t-1}'B_n^{-}V_{nt}$. Its variance is

$$\frac{1}{(nT)^2}\sum_{t=1}^{T}Var(\mathbb{U}_{n,t-1}'B_n^{-}V_{nt})=\frac{\sigma_0^2T}{(nT)^2}E(\mathbb{U}_{n,t-1}'B_n^{-}B_n^{-'}\mathbb{U}_{n,t-1})$$
$$=\frac{\sigma_0^2T}{(nT)^2}tr(B_n^{-}B_n^{-'}E(\mathbb{U}_{n,t-1}\mathbb{U}_{n,t-1}'))=\frac{\sigma_0^4T}{(nT)^2}tr(B_n^{-}B_n^{-'}\sum_{h=1}^{\infty}P_{nh}P_{nh}')=O\left(\frac{1}{nT}\right),$$

because $E(\mathbb{U}_{n,t-1}\mathbb{U}_{n,t-1}')=\sigma_0^2\sum_{h=1}^{\infty}P_{nh}P_{nh}'$ (from Lemma A.2) is uniformly bounded in both row and column sums. This proves (c).

For (d), $\frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}k_{ni}u_{n,t-1,i}=\frac{1}{nT}\sum_{t=1}^{T}K_n'\mathbb{U}_{n,t-1}$. Its variance is

$$\frac{1}{(nT)^2}\sum_{t=1}^{T}K_n'E(\mathbb{U}_{n,t-1}\mathbb{U}_{n,t-1}')K_n=\frac{\sigma_0^2T}{(nT)^2}K_n'(\sum_{h=1}^{\infty}P_{nh}P_{nh}')K_n=O\left(\frac{1}{nT}\right).$$

This proves (d). ■

Proof for Theorem A.13

We are going to use the following Theorem (Gänsler and Stute (1977, p.365)) to prove our CLT.

Theorem A.1 Let $\{X_{i,n},\mathcal{F}_{i,n},1\le i\le k_n,n\ge 1\}$ be a square integrable martingale difference array. Suppose that for all $\epsilon>0$, (1) $\sum_{i=1}^{k_n}E\left[X_{i,n}^2\mathbf{I}(|X_{i,n}|>\epsilon)|\mathcal{F}_{i-1,n}\right]\xrightarrow{p}0$ and (2) $\sum_{i=1}^{k_n}E\left[X_{i,n}^2|\mathcal{F}_{i-1,n}\right]\xrightarrow{p}1$, then $\sum_{i=1}^{k_n}X_{i,n}\xrightarrow{D}N(0,1)$. And a sufficient condition for (1) is $\sum_{i=1}^{k_n}E\left[|X_{i,n}|^{2+\delta}\right]\xrightarrow{p}0$ for some $\delta>0$ (see Pötscher and Prucha (1997), p.235).

Consider the σ-field

$$\mathcal{F}_{n,t,i}=\sigma(v_{11},v_{21},\cdots,v_{n1},\cdots,v_{1,t-1},\cdots,v_{n,t-1},v_{1t},\cdots,v_{it}),\qquad\text{(A.21)}$$

then $E(z_{nt,i}|\mathcal{F}_{n,t,i-1})=0$ and $E(z_{nt,i}|\mathcal{F}_{n,t-1,n})=0$. As a convention, define $\mathcal{F}_{n,t,0}=\mathcal{F}_{n,t-1,n}$. Thus, $\{z_{nt,i},\mathcal{F}_{n,t,i},\ 1\le t\le T,\ 1\le i\le n\}$ forms a martingale difference array. To see explicitly that this is a difference array, let $j=n(t-1)+i$ for $1\le i\le n$ and $1\le t\le T$. Thus, j takes integer values from 1 to J where $J=nT$. The σ-field $\mathcal{F}_{n,t,i}$ can be renamed as $\mathcal{F}_{J,j}$ and $z_{Jj}=z_{nt,i}$. As $E(z_{J,j}|\mathcal{F}_{J,j-1})=0$ because $E(z_{nt,i}|\mathcal{F}_{n,t,i-1})=0$ and $E(z_{nt,i}|\mathcal{F}_{n,t-1,n})=0$, $\{z_{nt,i},\mathcal{F}_{n,t,i}\}=\{z_{J,j},\mathcal{F}_{J,j-1}\}$ is a martingale difference array.

Denote $z_{Jj}^*=z_{nt,i}^*=\frac{z_{nt,i}}{\sigma_{Q_{nT}}}$, where $z_{nt,i}=(u_{i,t-1}+d_{nti})v_{it}+b_{n,ii}(v_{it}^2-\sigma_0^2)+2(\sum_{j=1}^{i-1}b_{n,ij}v_{jt})v_{it}$, we will apply Theorem A.1 to $\sum_{j=1}^{nT}z_{Jj}=\sum_{t=1}^{T}\sum_{i=1}^{n}z_{nt,i}$. The sufficient conditions are (i) $\frac{1}{\sigma_{Q_{nT}}^{2+\delta}}\sum_{t=1}^{T}\sum_{i=1}^{n}E|z_{nt,i}|^{2+\delta}\to 0$ and (ii) $\frac{1}{\sigma_{Q_{nT}}^2}\sum_{t=1}^{T}\sum_{i=1}^{n}E(z_{nt,i}^2|\mathcal{F}_{nt,i-1})\xrightarrow{p}1$. To apply Theorem A.1, we have $k_n=nT$, $X_{i,n}=z_{Jj}^*=\frac{z_{nt,i}}{\sigma_{Q_{nT}}}$ in our case and $\mathcal{F}_{i,n}$ is defined in Equation (A.21).

To show (i): For any $p > 0$ and $q > 0$ such that $\frac{1}{p} + \frac{1}{q} = 1$, as

$$\begin{aligned} |z_{nt,i}| &\leq (|u_{n,t-1,i}| + |d_{nt,i}|)|v_{it}| + |b_{nii}| \cdot |v_{it}^2 - \sigma_0^2| + 2|v_{it}| \cdot \sum_{j=1}^{i-1} |b_{nij}||v_{jt}| \\ &= (|u_{n,t-1,i}| + |d_{nt,i}|)|v_{it}| + |b_{nii}|^{\frac{1}{p}}|b_{nii}|^{\frac{1}{q}} \cdot |v_{it}^2 - \sigma_0^2| + 2|v_{it}| \cdot \sum_{j=1}^{i-1} |b_{nij}|^{\frac{1}{p}}|b_{nij}|^{\frac{1}{q}}|v_{jt}|, \end{aligned}$$

the Holder inequality implies that

$$|z_{nt,i}| \leq \left[(|u_{n,t-1,i}| + |d_{nt,i}|)^p + \sum_{j=1}^{i} |b_{nij}| \right]^{\frac{1}{p}} \cdot \left[|v_{it}|^q + |b_{nii}| \cdot |v_{it}^2 - \sigma_0^2|^q + 2^q |v_{it}|^q \cdot (\sum_{j=1}^{i-1} |b_{nij}||v_{jt}|^q) \right]^{\frac{1}{q}}.$$

Hence,

$$\begin{aligned} E|z_{nt,i}|^q &\leq E\left[(|u_{n,t-1,i}| + |d_{nt,i}|)^p + \sum_{j=1}^{i} |b_{nij}| \right]^{\frac{q}{p}} \\ &\cdot \left[E|v_{it}|^q + |b_{nii}| \cdot E|v_{it}^2 - \sigma_0^2|^q + 2^q E|v_{it}|^q \cdot (\sum_{j=1}^{i-1} |b_{nij}| E|v_{jt}|^q) \right]. \end{aligned}$$

Because the fourth and more moments of v_{it} exists, by taking $q = 2 + \delta$ for some small δ, there exists a constant $c_1 > 0$ such that $E|z_{nt,i}|^q \leq c_1 E[(|u_{n,t-1,i}| + |d_{nti}|)^p + \sum_{j=1}^{i} |b_{n,ij}|]^{\frac{q}{p}}$. By the c_r-inequality and because B_n is bounded in row sum norm, there exist constants $c_2 > 0$, $c_3 > 0$ and $c_4 > 0$ such that

$$\begin{aligned} &\left[(|u_{n,t-1,i}| + |d_{nti}|)^p + \sum_{j=1}^{i} |b_{n,ij}| \right]^{\frac{q}{p}} \leq 2^{\frac{q}{p}-1} \left[(|u_{n,t-1,i}| + |d_{nti}|)^q + c_3 \right] \\ \leq\ & 2^{\frac{q}{p}-1}[2^{q-1}(|u_{n,t-1,i}|^q + |d_{nti}|^q) + c_3] \leq c_2 |u_{n,t-1,i}|^{2+\delta} + c_4, \end{aligned}$$

as $q = 2 + \delta$ implies $\frac{q}{p} = 1 + \delta$. As $E|u_{nt,i}|^4 = O(1)$ uniformly in n, t and i (from Lemma A.10), it follows that $E|z_{nt,i}|^{2+\delta} \leq c_1 c_2 E|u_{n,t-1,i}|^{2+\delta} + c_1 c_4 = O(1)$ uniformly. Because $\sigma_{QnT}^{2+\delta} = O[(nT)^{1+\frac{\delta}{2}}]$, one has $\frac{1}{\sigma_{Q_{nT}}^{2+\delta}} \sum_{t=1}^{T} \sum_{i=1}^{n} E|z_{nt,i}|^{2+\delta} = O\left(\frac{1}{(nT)^{\frac{\delta}{2}}} \right)$ which goes to zero. This proves (i).

To show (ii): Because $z_{nt,i} = (u_{n,t-1,i} + d_{nti} + 2\sum_{j=1}^{i-1} b_{nij} v_{jt}) v_{it} + b_{nii}(v_{it}^2 - \sigma_0^2)$, it implies that

$$E(z_{nt,i}^2 | \mathcal{F}_{nt,i-1}) = \sigma_0^2 (u_{n,t-1,i} + d_{nti} + 2\sum_{j=1}^{i-1} b_{nij} v_{jt})^2 + (\mu_4 - \sigma_0^4) b_{nii}^2 + 2\mu_3 b_{nii}(u_{n,t-1,i} + d_{nti} + 2\sum_{j=1}^{i-1} b_{nij} v_{jt}),$$

as $E(v_{it}(v_{it}^2-\sigma_0^2))=\mu_3$ and $E(v_{it}^2-\sigma_0^2)^2=\mu_4-\sigma_0^4$. Therefore,

$$
\begin{aligned}
&\sum_{t=1}^{T}\sum_{i=1}^{n}E(z_{nt,i}^2|\mathcal{F}_{nt,i-1})\\
=\;&\sigma_0^2\sum_{t=1}^{T}\sum_{i=1}^{n}(u_{n,t-1,i}+2\sum_{j=1}^{i-1}b_{nij}v_{jt})^2+2\sum_{t=1}^{T}\sum_{i=1}^{n}[\sigma_0^2d_{nti}+\mu_3b_{nii}](u_{n,t-1,i}+2\sum_{j=1}^{i-1}b_{nij}v_{jt})\\
&+(\mu_4-\sigma_0^4)T\sum_{i=1}^{n}b_{nii}^2+2\mu_3\sum_{t=1}^{T}\sum_{i=1}^{n}b_{nii}d_{nti}+\sigma_0^2\sum_{t=1}^{T}\sum_{i=1}^{n}d_{nti}^2.
\end{aligned}
$$

This can be compared with $\sigma_{Q_{nT}}^2$, which can be rewritten as

$$\sigma_{Q_{nT}}^2=T\sigma_0^4tr(\sum_{h=1}^{\infty}P_{nh}'P_{nh})+2\sigma_0^4T[tr(B_n^2)-\sum_{i=1}^{n}b_{nii}^2]+T(\mu_4-\sigma_0^4)\sum_{i=1}^{n}b_{nii}^2+2\mu_3\sum_{t=1}^{T}\sum_{i=1}^{n}d_{nti}b_{nii}+\sigma_0^2\sum_{t=1}^{T}D_{nt}'D_{nt}.$$

From these, we can see that (ii) follows from the results in Lemma A.12 and Theorem A.7. ■

B Lemmas for Some Statistics in the Model

Denote $Z_{nt}=(Y_{n,t-1},\ W_nY_{n,t-1},\ X_{nt})$, we are going to provide some lemmas related to $\tilde{Z}_{nt}$, $\bar{Z}_{nt}$ and $\tilde{V}_{nt}$, $\bar{V}_{nt}$.

Lemma B.1 *Under Assumption 1-7, for an $n\times n$ nonstochastic matrix B_n, uniformly bounded in row and column sums,*

$$\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}_{nt}'B_n\tilde{Z}_{nt}-E\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}_{nt}'B_n\tilde{Z}_{nt}=O_p\left(\frac{1}{\sqrt{nT}}\right) \tag{B.1}$$

where $E\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}_{nt}'B_n\tilde{Z}_{nt}$ is $O(1)$;

$$\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}_{nt}'B_n\tilde{V}_{nt}-E\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}_{nt}'B_n\tilde{V}_{nt}=O_p\left(\frac{1}{\sqrt{nT}}\right) \tag{B.2}$$

where $E\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}_{nt}'B_n\tilde{V}_{nt}$ is $O\left(\frac{1}{T}\right)$;

$$\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}_{nt}'B_n\tilde{V}_{nt}-E\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}_{nt}'B_n\tilde{V}_{nt}=O_p\left(\frac{1}{\sqrt{nT}}\right) \tag{B.3}$$

where $E\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}_{nt}'B_n\tilde{V}_{nt}$ is $O(1)$.

Lemma B.2 *Under Assumption 1-7, for an $n \times n$ nonstochastic matrix B_n, uniformly bounded in row and column sums,*

$$\frac{1}{n}\bar{Z}'_{nT}B_n\bar{Z}_{nT} - E\frac{1}{n}\bar{Z}'_{nT}B_n\bar{Z}_{nT} = O_p(\frac{1}{\sqrt{nT}}) \tag{B.4}$$

where $E\frac{1}{n}\bar{Z}'_{nT}B_n\bar{Z}_{nT}$ is $O(1)$;

$$\frac{1}{n}\bar{Z}'_{nT}B_n\bar{V}_{nT} - E\frac{1}{n}\bar{Z}'_{nT}B_n\bar{V}_{nT} = O_p\left(\frac{1}{\sqrt{nT}}\right) \tag{B.5}$$

where $E\frac{1}{n}\bar{Z}'_{nT}B_n\bar{V}_{nT}$ is $O\left(\frac{1}{T}\right)$;

$$\frac{1}{n}\bar{V}'_{nT}B_n\bar{V}_{nT} - E\frac{1}{n}\bar{V}'_{nT}B_n\bar{V}_{nT} = O_p\left(\frac{1}{\sqrt{nT^2}}\right) \tag{B.6}$$

where $E\frac{1}{n}\bar{V}'_{nT}B_n\bar{V}_{nT}$ is $O\left(\frac{1}{T}\right)$.

Also, from Equation (2.3),

$$\tilde{Z}_{nt} = \tilde{Z}^*_{nt} - (\bar{U}_{nT,-1}, \quad W_n\bar{U}_{nT,-1}, \quad \mathbf{0})$$

where $\tilde{Z}^*_{nt} = ((\approx{\mathcal{X}}_{n,t-1} + U_{n,t-1}), \ (W_n\approx{\mathcal{X}}_{n,t-1} + W_nU_{n,t-1}), \ \tilde{X}_{nt})$ with $\approx{\mathcal{X}}_{n,t-1} = \mathcal{X}_{n,t-1} - \bar{\mathcal{X}}_{nT,-1}$. Hence Z_{nt} has two components: one is $\tilde{Z}^*_{nt}$, which is uncorrelated with V_{nt}; the other is $-(\bar{U}_{nT,-1} \ W_n\bar{U}_{nT,-1} \ \mathbf{0})$, but is correlated with V_{nt} when $t \le T-1$. Following is a lemma related to $\tilde{Z}^*_{nt}$ and Z_{nt}.

Lemma B.3 *Under Assumption 1-7, for an $n \times n$ nonstochastic matrix B_n, uniformly bounded in row and column sums,*

$$E\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}'_{nt}B_n\tilde{Z}_{nt} - E\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}^{*\prime}_{nt}B_n\tilde{Z}^*_{nt} = O\left(\frac{1}{T}\right) \tag{B.7}$$

where $E\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}^{*\prime}_{nt}B_n\tilde{Z}^*_{nt}$ is $O(1)$.

Lemma B.4 *Under Assumption 1-7, if the elements of $\mathbf{c}_{n0}$ are uniformly bounded, then the elements of $\frac{1}{T}\sum_{t=1}^{T}((G_n\mathbf{c}_{n0} + G_nZ_{nt}\delta_0)_i, \ (Z_{nt})_i)$ are $O_p(1)$ uniformly in n and i, where $(G_n\mathbf{c}_{n0} + G_nZ_{nt}\delta_0)_i$ is the ith element of $(G_n\mathbf{c}_{n0} + G_nZ_{nt}\delta_0)$ and $(Z_{nt})_i$ is the ith row of Z_{nt}.*

Proof for Lemma B.1

As $Z_{nt} = (Y_{n,t-1}, \ W_nY_{n,t-1}, \ X_{nt})$ and $Y_{nt} = \mu_n + \mathcal{X}_{nt}\beta_0 + U_{nt}$ according to (2.3), elements of $\tilde{Z}_{nt}$ have two components in general: one is $\approx{\mathbb{U}}_{n,t-1}$ (or $\approx{\mathbb{W}}_{n,t-1}$) and the other is a certain nonstochastic n-dimensional vector, say D_{nt}. From (A.9) in Theorem A.7, $\frac{1}{nT}\sum_{t=1}^{T}\approx{\mathbb{U}}'_{n,t-1}\approx{\mathbb{W}}_{n,t-1} - E(\frac{1}{nT}\sum_{t=1}^{T}\approx{\mathbb{U}}'_{n,t-1}\approx{\mathbb{W}}_{n,t-1}) =$

$O_p(\frac{1}{\sqrt{nT}})$ where $E(\frac{1}{nT}\sum_{t=1}^{T}\overset{\approx}{\mathbb{U}}'_{n,t-1}\overset{\approx}{\mathbb{W}}_{n,t-1})$ is $O(1)$. From Theorem A.8, $\frac{1}{nT}\sum_{t=1}^{T}\tilde{D}'_{nt}\overset{\approx}{\mathbb{U}}_{n,t-1} = O_p(\frac{1}{\sqrt{nT}})$. Hence, $\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}'_{nt}B_n\tilde{Z}_{nt} - E\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}'_{nt}B_n\tilde{Z}_{nt} = O_p(\frac{1}{\sqrt{nT}})$ where $E\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}'_{nt}B_n\tilde{Z}_{nt}$ is $O(1)$. This proves (B.1).

For (B.2), $\frac{1}{nT}\sum_{t=1}^{T}\tilde{D}'_{nt}\tilde{V}_{nt} = \frac{1}{nT}\sum_{t=1}^{T}\tilde{D}'_{nt}V_{nt} = O_p\left(\frac{1}{\sqrt{nT}}\right)$ because $Var(\frac{1}{nT}\sum_{t=1}^{T}\tilde{D}'_{nt}V_{nt}) = \frac{\sigma_0^2}{n^2T^2}\sum_{t=1}^{T}\tilde{D}'_{nt}\tilde{D}_{nt} = O\left(\frac{1}{nT}\right)$. Also, $\frac{1}{nT}\sum_{t=1}^{T}\overset{\approx}{\mathbb{U}}'_{n,t-1}\tilde{V}_{nt} = E\left(\frac{1}{nT}\sum_{t=1}^{T}\overset{\approx}{\mathbb{U}}'_{n,t-1}\tilde{V}_{nt}\right) + O_p\left(\frac{1}{\sqrt{nT}}\right)$ where $E\left(\frac{1}{nT}\sum_{t=1}^{T}\overset{\approx}{\mathbb{U}}'_{n,t-1}\tilde{V}_{nt}\right)$ is $O\left(\frac{1}{T}\right)$ according to Theorem A.11. As elements of $\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}'_{nt}\tilde{V}_{nt}$ are just $\frac{1}{nT}\sum_{t=1}^{T}\tilde{D}'_{nt}\tilde{V}_{nt}$ and $\frac{1}{nT}\sum_{t=1}^{T}\overset{\approx}{\mathbb{U}}'_{n,t-1}\tilde{V}_{nt}$, (B.2) is proved. And (B.3) is just (A.14) in Lemma A.9. ■

Proof for Lemma B.2

Similarly to the proof of Lemma B.1, elements of $\bar{Z}_{nt}$ have two components in general: one is $\bar{\mathbb{U}}_{nT,-1}$ (or $\bar{\mathbb{W}}_{nT,-1}$) and the other is $\bar{D}_{nT}$. From (A.8) in Theorem A.7, $\frac{1}{n}\bar{\mathbb{U}}'_{nT}\bar{\mathbb{W}}_{nT} - E(\frac{1}{n}\bar{\mathbb{U}}'_{nT}\bar{\mathbb{W}}_{nT}) = O_p\left(\frac{1}{\sqrt{nT^2}}\right)$ where $E(\frac{1}{n}\bar{\mathbb{U}}'_{nT}\bar{\mathbb{W}}_{nT})$ is $O\left(\frac{1}{T}\right)$. From Theorem A.8, $\frac{1}{n}\bar{D}'_{nT}\bar{\mathbb{U}}_{nT,-1} = O_p\left(\frac{1}{\sqrt{nT}}\right)$ because elements of D_{nt} are uniformly bounded. However, $\frac{1}{n}\bar{D}'_{nT}B_n\bar{D}_{nT} = O(1)$. Hence, $\frac{1}{n}\bar{Z}'_{nT}B_n\bar{Z}_{nT} - E\frac{1}{n}\bar{Z}'_{nT}B_n\bar{Z}_{nT} = O_p\left(\frac{1}{\sqrt{nT}}\right)$ where $E\frac{1}{n}\bar{Z}'_{nT}B_n\bar{Z}_{nT}$ is $O(1)$. This proves (B.4).

For (B.5), $\frac{1}{n}\bar{D}'_{nT}\bar{V}_{nT} = O_p\left(\frac{1}{\sqrt{nT}}\right)$ because $Var(\frac{1}{n}\bar{D}'_{nT}\bar{V}_{nT}) = O\left(\frac{1}{nT}\right)$. Also, $\frac{1}{n}\bar{\mathbb{U}}'_{nT,-1}\bar{V}_{nT} = E\left(\frac{1}{n}\bar{\mathbb{U}}'_{nT,-1}\bar{V}_{nT}\right) + O_p\left(\frac{1}{\sqrt{nT^2}}\right)$ where $E\left(\frac{1}{n}\bar{\mathbb{U}}'_{nT,-1}\bar{V}_{nT}\right) = O\left(\frac{1}{T}\right)$ according to Theorem A.11. As elements of $\frac{1}{n}\bar{Z}'_{nT}\bar{V}_{nT}$ are just $\frac{1}{n}\bar{D}'_{nT}\bar{V}_{nT}$ and $\frac{1}{n}\bar{\mathbb{U}}'_{nT,-1}\bar{V}_{nT}$, (B.5) is proved. And (B.6) is just (A.13) in Lemma A.9. ■

Proof for Lemma B.3

From (3.3), $\tilde{Z}_{nt} = \tilde{Z}^*_{nt} - (\bar{U}_{nT,-1},\ W_n\bar{U}_{nT,-1},\ \mathbf{0})$ where $\tilde{Z}^*_{nt} = ((\overset{\approx}{\mathcal{X}}_{n,t-1} + U_{n,t-1}), (W_n\overset{\approx}{\mathcal{X}}_{n,t-1} + W_nU_{n,t-1}), \tilde{X}_{nt})$. Hence, elements of $E\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}'_{nt}B_n\tilde{Z}_{nt} - E\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}^{*\prime}_{nt}B_n\tilde{Z}^*_{nt}$ have the form of $E(\frac{1}{n}\bar{\mathbb{U}}'_{nT}\bar{\mathbb{W}}_{nT})$, which is $O\left(\frac{1}{T}\right)$ according to (A.8) in Theorem A.7. ■

Proof for Lemma B.4

First, $\frac{1}{T}\sum_{t=1}^{T}(G_n\mathbf{c}_{n0})_i = (G_n\mathbf{c}_{n0})_i = (W_nS_n^{-1}\mathbf{c}_{n0})_i = e'_{ni}W_nS_n^{-1}\mathbf{c}_{n0}$ where e_{ni} is the ith unit vector of R^n. If the elements of $\mathbf{c}_{n0}$ are bounded, $e_{ni}W_nS_n^{-1}\mathbf{c}_{n0}$ is bounded because W_n and S_n^{-1} are row sum bounded.

Second, $\frac{1}{T}\sum_{t=1}^{T}(G_nZ_{nt}\delta_0)_i = \frac{1}{T}\sum_{t=1}^{T}e'_{ni}G_nZ_{nt}\delta_0$. As $Z_{nt} = (Y_{n,t-1},\ W_nY_{n,t-1},\ X_{nt})$ and elements of X_{nt} are uniformly bounded, it is sufficient to show that for a nonstochastic $n\times n$ matrix B_n which is uniformly bounded in row and column sums, $\frac{1}{T}\sum_{t=1}^{T}e'_{ni}B_nY_{n,t-1}$ is stochastically bounded. From Equation (2.3) that $Y_{nt} = \sum_{h=0}^{\infty}A_n^hS_n^{-1}(\mathbf{c}_{n0} + X_{n,t-h}\beta_0) + U_{nt}$, we only need to show $e'_{ni}B_n\left(\sum_{h=0}^{\infty}A_n^hS_n^{-1}\mathbf{c}_{n0}\right)$ and $\frac{1}{T}\sum_{t=1}^{T}e'_{ni}B_nU_{n,t-1}$ are stochastically bounded. $e'_{ni}B_n\left(\sum_{h=0}^{\infty}A_n^hS_n^{-1}\mathbf{c}_{n0}\right)$ is bounded under the assumption that elements of $\mathbf{c}_{n0}$ are bounded. For $\frac{1}{T}\sum_{t=1}^{T} e'_{ni}B_nU_{n,t-1}$, it has zero mean and $Var(\frac{1}{T}\sum_{t=1}^{T} e'_{ni}B_nU_{n,t-1}) = \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T} e'_{ni}B_n\left(E\left(U_{n,t-1}U'_{n,s-1}\right)\right)B'_ne_{ni}$.

By Lemma A.2, for $t \geq s$, $E(U_{n,t-1}U'_{n,s-1}) = \sigma_0^2(\sum_{h=1}^{\infty} P_{n,t-s+h}P'_{nh})$ and, for $s \geq t$, $E(U_{n,t-1}U'_{n,s-1}) = \sigma_0^2(\sum_{h=1}^{\infty} P_{nh}P'_{n,s-t+h})$. For our case, $U_{n,t} = \sum_{h=0}^{\infty} A_n^hS_n^{-1}V_{n,t-h} = \sum_{h=1}^{\infty} A_n^{h-1}S_n^{-1}V_{n,t+1-h}$ and, hence, $P_{nh} = A_n^{h-1}S_n^{-1}$, $h = 1, 2, \cdots$. So, for $t \geq s$, $P_{n,t-s+h}P'_{nh} = A_n^{t-s}P_{nh}P'_{nh}$; and, for $s \geq t$, $P_{nh}P'_{n,s-t+h} = P_{nh}P'_{nh}A_n^{\prime(s-t)}$. It follows that

$$\sum_{s=1}^{t}(\sum_{h=1}^{\infty} P_{n,t-s+h}P'_{nh}) = (\sum_{l=0}^{t-1} A_n^l)(\sum_{h=1}^{\infty} P_{nh}P'_{nh}),$$

and

$$\sum_{s=t+1}^{T}(\sum_{h=1}^{\infty} P_{nh}P'_{n,s-t+h}) = (\sum_{h=1}^{\infty} P_{nh}P'_{nh})(\sum_{l=1}^{T-t} A_n^{\prime l}).$$

Both matrices are uniformly bounded in row and column sums. As B_n is also uniformly bounded in both row and column sums, it follows that $B_n\sum_{s=1}^{T} E(U_{n,t-1}U'_{n,s-1})B'_n$ is uniformly bounded in row and column sums, and $\left[B_n\sum_{s=1}^{T} E(U_{n,t-1}U'_{n,s-1})B'_n\right]_{ii}$ is uniformly bounded in n and i. Hence, $Var(\frac{1}{T}\sum_{t=1}^{T} e'_{ni}B_nU_{n,t-1}) = O\left(\frac{1}{T}\right)$. Therefore, $\frac{1}{T}\sum_{t=1}^{T} e'_{ni}B_nU_{n,t-1} = O_p\left(\frac{1}{\sqrt{T}}\right)$. The results in this lemma follow by taking B_n equal to G_n, G_nW_n, W_n and I_n. ■

C Concentrated QML Estimator

C.1 Reduced Form of (2.1)

As $Z_{nt} = (Y_{n,t-1}, W_nY_{n,t-1}, X_{nt})$, the reduced form of (2.1) can be represented as

$$\begin{aligned} Y_{nt} &= S_n^{-1}Z_{nt}\delta_0 + S_n^{-1}\left(\mathbf{c}_{n0} + V_{nt}\right) \\ &= Z_{nt}\delta_0 + \lambda_0G_nZ_{nt}\delta_0 + S_n^{-1}(\mathbf{c}_{n0} + V_{nt}), \ \ t = 1, 2, ..., T \end{aligned} \tag{C.1}$$

because $S_n^{-1} = I_n + \lambda_0G_n$ where $G_n \equiv W_nS_n^{-1}$. This implies that

$$\begin{aligned} \tilde{Y}_{nt} &= S_n^{-1}\tilde{Z}_{nt}\delta_0 + S_n^{-1}\tilde{V}_{nt} \\ &= \tilde{Z}_{nt}\delta_0 + \lambda_0G_n\tilde{Z}_{nt}\delta_0 + S_n^{-1}\tilde{V}_{nt}, \ \ t = 1, 2, ..., T. \end{aligned} \tag{C.2}$$

C.2 The First Order Condition and Second Order Condition

For the concentrated likelihood function (2.5), the first order derivatives are

$$\frac{1}{\sqrt{nT}}\frac{\partial \ln L_{n,T}(\theta)}{\partial\theta} = \frac{1}{\sqrt{nT}}\begin{pmatrix} \frac{\partial \ln L_{n,T}(\theta)}{\partial\delta} \\ \frac{\partial \ln L_{n,T}(\theta)}{\partial\lambda} \\ \frac{\partial \ln L_{n,T}(\theta)}{\partial\sigma^2} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sigma^2}\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}\tilde{Z}'_{nt}\tilde{V}_{nt}(\zeta) \\ \frac{1}{\sigma^2}\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}\Big((W_n\tilde{Y}_n)'\tilde{V}_{nt}(\zeta) - trG_n(\lambda)\Big) \\ \frac{1}{2\sigma^4}\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}(\tilde{V}'_{nt}(\zeta)\tilde{V}_{nt}(\zeta) - n\sigma^2) \end{pmatrix}, \tag{C.3}$$

and the second order derivatives are

$$\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\theta)}{\partial\theta\partial\theta'} = -\frac{1}{nT}\begin{pmatrix} \frac{1}{\sigma^2}\sum_{t=1}^{T}\tilde{Z}'_{nt}\tilde{Z}_{nt} & \frac{1}{\sigma^2}\sum_{t=1}^{T}\tilde{Z}'_{nt}W_n\tilde{Y}_{nt} & \frac{1}{\sigma^4}\sum_{t=1}^{T}\tilde{Z}'_{nt}\tilde{V}_{nt}(\zeta) \\ * & \frac{1}{\sigma^2}\sum_{t=1}^{T}\Big((W_n\tilde{Y}_{nt})'W_n\tilde{Y}_{nt} + tr(G_n^2(\lambda))\Big) & \frac{1}{\sigma^4}\sum_{t=1}^{T}(W_n\tilde{Y}_{nt})'\tilde{V}_{nt}(\zeta) \\ * & * & -\frac{nT}{2\sigma^4} + \frac{1}{\sigma^6}\sum_{t=1}^{T}\tilde{V}'_{nt}(\zeta)\tilde{V}_{nt}(\zeta) \end{pmatrix}. \tag{C.4}$$

Hence,

$$\frac{1}{\sqrt{nT}}\frac{\partial \ln L_{n,T}(\theta_0)}{\partial\theta} = \frac{1}{\sqrt{nT}}\begin{pmatrix} \frac{\partial \ln L_{n,T}(\theta_0)}{\partial\delta} \\ \frac{\partial \ln L_{n,T}(\theta_0)}{\partial\lambda} \\ \frac{\partial \ln L_{n,T}(\theta_0)}{\partial\sigma^2} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}\tilde{Z}'_{nt}\tilde{V}_{nt} \\ \frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}(G_n\tilde{Z}_{nt}\delta_0)'\tilde{V}_{nt} + \frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}(\tilde{V}'_{nt}G'_n\tilde{V}_{nt} - \sigma_0^2 trG_n) \\ \frac{1}{2\sigma_0^4}\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}(\tilde{V}'_{nt}\tilde{V}_{nt} - n\sigma_0^2) \end{pmatrix} \tag{C.5}$$

and the information matrix is equal to

$$\begin{aligned}
\Sigma_{\theta_0,nT} &= -E\left(\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\theta_0)}{\partial\theta\partial\theta'}\right) \\
&= \begin{pmatrix} \frac{1}{\sigma_0^2 nT}E\sum_{t=1}^{T}\tilde{Z}'_{nt}\tilde{Z}_{nt} & \frac{1}{\sigma_0^2 nT}E\sum_{t=1}^{T}\tilde{Z}'_{nt}G_n\tilde{Z}_{nt}\delta_0 & 0 \\ \frac{1}{\sigma_0^2 nT}E\sum_{t=1}^{T}(G_n\tilde{Z}_{nt}\delta_0)'\tilde{Z}_{nt} & \frac{1}{\sigma_0^2 nT}E\sum_{t=1}^{T}(G_n\tilde{Z}_{nt}\delta_0)'G_n\tilde{Z}_{nt}\delta_0 + \frac{1}{n}\left[tr(G'_nG_n) + tr(G_n^2)\right] & \frac{1}{\sigma_0^2 n}tr(G_n) \\ 0 & \frac{1}{\sigma_0^2 n}tr(G_n) & \frac{1}{2\sigma_0^4} \end{pmatrix} \\
&\quad - \begin{pmatrix} 0 & * & * \\ \frac{1}{\sigma_0^2 n}E(G_n\bar{V}_{nT})'\bar{Z}_{nT} & \frac{2}{\sigma_0^2 n}E\left[(G_n\bar{Z}_{nT}\delta_0)'G_n\bar{V}_{nT}\right] + \frac{1}{nT}tr(G'_nG_n) & * \\ \frac{1}{\sigma_0^4 n}E\left(\bar{Z}'_{nT}\bar{V}_{nT}\right)' & \frac{1}{\sigma_0^4 n}E\left[(G_n\bar{Z}_{nT}\delta_0)'\bar{V}_{nT}\right]' + \frac{1}{\sigma_0^2 nT}tr(G_n) & \frac{1}{T}\frac{1}{\sigma_0^4} \end{pmatrix}.
\end{aligned}$$

Using Lemma B.2, the second matrix of above is $O\left(\frac{1}{T}\right)$, so, $\Sigma_{\theta_0,nT} = -E\left(\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\theta_0)}{\partial\theta\partial\theta'}\right) =$

$$\begin{pmatrix} \frac{1}{\sigma_0^2 nT}E\sum_{t=1}^{T}\tilde{Z}_{nt}'\tilde{Z}_{nt} & \frac{1}{\sigma_0^2 nT}E\sum_{t=1}^{T}\tilde{Z}_{nt}'G_n\tilde{Z}_{nt}\delta_0 & 0 \\ \frac{1}{\sigma_0^2 nT}E\sum_{t=1}^{T}(G_n\tilde{Z}_{nt}\delta_0)'\tilde{Z}_{nt} & \frac{1}{\sigma_0^2 nT}E\sum_{t=1}^{T}(G_n\tilde{Z}_{nt}\delta_0)'G_n\tilde{Z}_{nt}\delta_0 + \frac{1}{n}\left[tr(G_n'G_n)+tr(G_n^2)\right] & \frac{1}{\sigma_0^2 n}tr(G_n) \\ 0 & \frac{1}{\sigma_0^2 n}tr(G_n) & \frac{1}{2\sigma_0^4} \end{pmatrix} + O\left(\frac{1}{T}\right). \tag{C.6}$$

C.3 The Variance of the Gradient

We are going to show that for $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\theta_0)}{\partial\theta}$ in (3.4), $E(\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\theta_0)}{\partial\theta}\cdot\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\theta_0)}{\partial\theta'}) = \Sigma_{\theta_0,nT} + \Omega_{\theta_0,nT} + O\left(\frac{1}{T}\right)$ where $\Omega_{\theta_0,nT} = \frac{\mu_4-3\sigma_0^4}{\sigma_0^4}\begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{n}\sum_{i=1}^{n}G_{n,ii}^2 & \frac{1}{2\sigma_0^2 n}trG_n \\ 0 & \frac{1}{2\sigma_0^2 n}trG_n & \frac{1}{4\sigma_0^4}\end{pmatrix}$. First, we can write $E(\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\theta_0)}{\partial\theta}\cdot\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\theta_0)}{\partial\theta'}) =$

$$E\frac{1}{nT}\begin{pmatrix} \frac{1}{\sigma_0^4}\left(\sum_{t=1}^{T}\tilde{Z}_{nt}^{*\prime}V_{nt}\right)\left(\sum_{t=1}^{T}\tilde{Z}_{nt}^{*\prime}V_{nt}\right)' & * & * \\ \frac{1}{\sigma_0^4}\left(\sum_{t=1}^{T}(G_n\tilde{Z}_{nt}^*\delta_0)'V_{nt} + \sum_{t=1}^{T}(V_{nt}'G_n'V_{nt}-\sigma_0^2 trG_n)\right)\left(\sum_{t=1}^{T}\tilde{Z}_{nt}^{*\prime}V_{nt}\right)' & 0 & 0 \\ \frac{1}{2\sigma_0^6}\left(\sum_{t=1}^{T}(V_{nt}'V_{nt}-n\sigma_0^2)\right)\left(\sum_{t=1}^{T}\tilde{Z}_{nt}^{*\prime}V_{nt}\right)' & 0 & 0 \end{pmatrix}$$

$$+E\frac{1}{nT}\begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{\sigma_0^4}\left(\sum_{t=1}^{T}(G_n\tilde{Z}_{nt}^*\delta_0)'V_{nt} + \sum_{t=1}^{T}(V_{nt}'G_n'V_{nt}-\sigma_0^2 trG_n)\right)^2 & * \\ 0 & \frac{1}{2\sigma_0^4}\left(\sum_{t=1}^{T}(G_n\tilde{Z}_{nt}^*\delta_0)'V_{nt} + \sum_{t=1}^{T}(V_{nt}'G_n'V_{nt}-\sigma_0^2 trG_n)\right)\left(\sum_{t=1}^{T}(V_{nt}'V_{nt}-n\sigma_0^2)\right)' & 0 \end{pmatrix}$$

$$+E\frac{1}{nT}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{1}{4\sigma_0^8}\left(\sum_{t=1}^{T}(V_{nt}'V_{nt}-n\sigma_0^2)\right)\left(\sum_{t=1}^{T}(V_{nt}'V_{nt}-n\sigma_0^2)\right)' \end{pmatrix}.$$

As $\tilde{Z}_{nt}^*$ is uncorrelated with V_{nt}, we have $E(\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\theta_0)}{\partial\theta}\cdot\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\theta_0)}{\partial\theta'})$

$$= \begin{pmatrix} \frac{1}{\sigma_0^2 nT}E\sum_{t=1}^{T}\tilde{Z}_{nt}^{*\prime}\tilde{Z}_{nt}^* & \frac{1}{\sigma_0^2 nT}E\sum_{t=1}^{T}\tilde{Z}_{nt}^{*\prime}G_n\tilde{Z}_{nt}^*\delta_0 & 0 \\ \frac{1}{\sigma_0^2 nT}E\sum_{t=1}^{T}(G_n\tilde{Z}_{nt}^*\delta_0)'\tilde{Z}_{nt}^* & \frac{1}{\sigma_0^2 nT}E\sum_{t=1}^{T}(G_n\tilde{Z}_{nt}^*\delta_0)'G_n\tilde{Z}_{nt}^*\delta_0 + \frac{1}{n}\left[tr(G_n'G_n)+tr(G_n^2)\right] & \frac{1}{\sigma_0^2 n}tr(G_n) \\ 0 & \frac{1}{\sigma_0^2 n}tr(G_n) & \frac{1}{2\sigma_0^4} \end{pmatrix}$$

$$+\begin{pmatrix} 0 & * & * \\ \frac{\mu_3}{\sigma_0^4 nT}\sum_{i=1}^{n} G_{n,ii}E(\sum_{t=1}^{T}\tilde{Z}_{nt}^*)_i & \frac{2\mu_3}{\sigma_0^4 nT}\sum_{i=1}^{n} G_{n,ii}E(\sum_{t=1}^{T} G_n\tilde{Z}_{nt}^*\delta_0)_i + \frac{\mu_4-3\sigma_0^4}{\sigma_0^4 n}\sum_{i=1}^{n} G_{n,ii}^2 & * \\ \frac{\mu_3}{2\sigma_0^6 nT} l_n' E\sum_{t=1}^{T}\tilde{Z}_{nt}^* & \frac{1}{2\sigma_0^6 nT}\mu_3 l_n' E\sum_{t=1}^{T} G_n\tilde{Z}_{nt}^*\delta_0 + \frac{\mu_4-3\sigma_0^4}{2\sigma_0^6 n} trG_n & \frac{\mu_4-3\sigma_0^4}{4\sigma_0^8} \end{pmatrix}$$

. For the first matrix, it is equal to $\Sigma_{\theta_0,nT} + O\left(\frac{1}{T}\right)$ using Lemma B.3. For the second matrix, $E\sum_{t=1}^{T}\tilde{Z}_{nt}^* = \mathbf{0}$ and $E\sum_{t=1}^{T} G_n\tilde{Z}_{nt}^*\delta_0 = \mathbf{0}$. Hence, $E(\frac{1}{\sqrt{nT}}\frac{\partial \ln L_{n,T}^*(\theta_0)}{\partial\theta}\cdot\frac{1}{\sqrt{nT}}\frac{\partial \ln L_{n,T}^*(\theta_0)}{\partial\theta'}) = \Sigma_{\theta_0,nT} + \Omega_{\theta_0,nT} + O\left(\frac{1}{T}\right)$. When V_{nt} are normally distributed, $\Omega_{\theta_0,nT} = 0$ because $\mu_4 - 3\sigma_0^4 = 0$ for a normal distribution.

C.4 About $-\frac{1}{nT}E\frac{\partial^2 \ln L_{nT}(\theta)}{\partial\theta\partial\theta'}$, $-\frac{1}{nT}\frac{\partial^2 \ln L_{nT}(\theta)}{\partial\theta\partial\theta'}$, $-\frac{1}{nT}E\frac{\partial^2 \ln L_{nT}(\theta_0)}{\partial\theta\partial\theta'}$ and $-\frac{1}{nT}\frac{\partial^2 \ln L_{nT}(\theta_0)}{\partial\theta\partial\theta'}$

Denote $\|\theta-\theta_0\|$ as the Euclidean norm of $\theta-\theta_0$, and Θ_1 as a neighborhood of θ_0, then, we have

$$-\frac{1}{nT}\frac{\partial^2 \ln L_{nT}(\theta)}{\partial\theta\partial\theta'} - (-\frac{1}{nT}\frac{\partial^2 \ln L_{nT}(\theta_0)}{\partial\theta\partial\theta'}) = \|\theta-\theta_0\|\cdot O_p(1), \tag{C.7}$$

$$(-\frac{1}{nT}\frac{\partial^2 \ln L_{nT}(\theta_0)}{\partial\theta\partial\theta'}) - \Sigma_{\theta_0,nT} = O_p\left(\frac{1}{\sqrt{nT}}\right), \tag{C.8}$$

$$\sup_{\theta\in\Theta}\left|-\frac{1}{nT}\frac{\partial^2 \ln L_{nT}(\theta)}{\partial\theta\partial\theta'} - (-\frac{1}{nT}E\frac{\partial^2 \ln L_{nT}(\theta)}{\partial\theta\partial\theta'})\right|_{ij} = O_p\left(\frac{1}{\sqrt{nT}}\right) \tag{C.9}$$

and

$$\sup_{\theta\in\Theta_1}\left|-\frac{1}{nT}E\frac{\partial^2 \ln L_{nT}(\theta)}{\partial\theta\partial\theta'} - \Sigma_{\theta_0,nT}\right|_{ij} = \sup_{\theta\in\Theta_1}\|\theta-\theta_0\|\cdot O(1) \tag{C.10}$$

for all $i,j = 1,2,\cdots,k_x+4$.

Proof for Equation (C.7)

$-\frac{1}{nT}\frac{\partial^2 \ln L_{nT}(\theta)}{\partial\theta\partial\theta'} - (-\frac{1}{nT}\frac{\partial^2 \ln L_{nT}(\theta_0)}{\partial\theta\partial\theta'}) =$

$$\begin{pmatrix} (\frac{1}{\sigma^2}-\frac{1}{\sigma_0^2})\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}_{nt}'\tilde{Z}_{nt} & * & * \\ (\frac{1}{\sigma^2}-\frac{1}{\sigma_0^2})\frac{1}{nT}\sum_{t=1}^{T}\left(W_n\tilde{Y}_{nt}\right)'\tilde{Z}_{nt} & 0 & 0 \\ \frac{1}{\sigma^4}\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}_{nt}'(\zeta)\tilde{Z}_{nt} - \frac{1}{\sigma_0^4}\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}_{nt}'\tilde{Z}_{nt} & 0 & 0 \end{pmatrix}$$

$$+\begin{pmatrix} 0 & 0 & 0 \\ 0 & (\frac{1}{\sigma^2}-\frac{1}{\sigma_0^2})\frac{1}{nT}\sum_{t=1}^{T}(W_n\tilde{Y}_{nt})'W_n\tilde{Y}_{nt} + \frac{1}{n}tr(G_n^2(\lambda)-G_n^2) & * \\ 0 & \frac{1}{\sigma^4}\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}_{nt}'(\zeta)W_n\tilde{Y}_{nt} - \frac{1}{\sigma_0^4}\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}_{nt}'W_n\tilde{Y}_{nt} & 0 \end{pmatrix}$$

$$+\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sigma^6}\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}_{nt}'(\zeta)\tilde{V}_{nt}(\zeta)-\frac{1}{2\sigma^4}-\frac{1}{\sigma_0^6}\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}_{nt}'\tilde{V}_{nt}+\frac{1}{2\sigma_0^4} \end{pmatrix}.$$

First, $\frac{1}{n}tr(G_n^2(\lambda)-G_n^2) = \frac{1}{n}tr[(G_n(\bar{\lambda}))^3](\lambda-\lambda_0)$ where $\bar{\lambda}$ lies between λ and λ_0. As $\frac{1}{n}tr[(G_n(\lambda))^3]$ is uniformly bounded by Lemma A.7 in Lee (2004), $\frac{1}{n}tr(G_n^2(\lambda)-G_n^2)$ is of the order $|\lambda-\lambda_0|\cdot O(1)$. Second, as $\tilde{V}_{nt}(\zeta)=\tilde{V}_{nt}-(\lambda-\lambda_0)W_n\tilde{Y}_{nt}-\tilde{Z}_{nt}(\delta-\delta_0)$ and $W_n\tilde{Y}_{nt}=G_n\tilde{Z}_{nt}\delta_0+G_n\tilde{V}_{nt}$, using Lemma B.1, all the entries in the above matrices are of the same order as $\|\theta-\theta_0\|$ multiplied by stochastic terms of orders not bigger than $O_p(1)$. Hence, $-\frac{1}{nT}\frac{\partial^2\ln L_{nT}(\theta)}{\partial\theta\partial\theta'}-(-\frac{1}{nT}\frac{\partial^2\ln L_{nT}(\theta_0)}{\partial\theta\partial\theta'})=\|\theta-\theta_0\|\cdot O_p(1)$. ■

Proof for Equation (C.8)

$$(-\frac{1}{nT}\frac{\partial^2\ln L_{nT}(\theta_0)}{\partial\theta\partial\theta'})-\Sigma_{\theta_0,nT}=$$

$$\begin{pmatrix} \frac{1}{\sigma_0^2}\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}_{nt}'\tilde{Z}_{nt}-\frac{1}{\sigma_0^2}E(\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}_{nt}'\tilde{Z}_{nt}) & * & * \\ \frac{1}{\sigma_0^2}\frac{1}{nT}\sum_{t=1}^{T}\left(W_n\tilde{Y}_{nt}\right)'\tilde{Z}_{nt}-\frac{1}{\sigma_0^2}E(\frac{1}{nT}\sum_{t=1}^{T}\left(W_n\tilde{Y}_{nt}\right)'\tilde{Z}_{nt}) & 0 & 0 \\ \frac{1}{\sigma_0^4}\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}_{nt}'\tilde{Z}_{nt}-\frac{1}{\sigma_0^4}E(\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}_{nt}'\tilde{Z}_{nt}) & 0 & 0 \end{pmatrix}$$

$$+\begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{\sigma_0^2}\frac{1}{nT}\sum_{t=1}^{T}(W_n\tilde{Y}_{nt})'W_n\tilde{Y}_{nt}-\frac{1}{\sigma_0^2}E(\frac{1}{nT}\sum_{t=1}^{T}(W_n\tilde{Y}_{nt})'W_n\tilde{Y}_{nt}) & * \\ 0 & \frac{1}{\sigma_0^4}\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}_{nt}'W_n\tilde{Y}_{nt}-\frac{1}{\sigma_0^4}E\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}_{nt}'W_n\tilde{Y}_{nt} & 0 \end{pmatrix}$$

$$+\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sigma_0^4}\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}_{nt}'\tilde{V}_{nt}-\frac{1}{\sigma_0^4}E\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}_{nt}'\tilde{V}_{nt} \end{pmatrix}.$$

As $\tilde{V}_{nt}(\zeta)=\tilde{V}_{nt}-(\lambda-\lambda_0)W_n\tilde{Y}_{nt}-\tilde{Z}_{nt}(\delta-\delta_0)$ and $W_n\tilde{Y}_{nt}=G_n\tilde{Z}_{nt}\delta_0+G_n\tilde{V}_{nt}$, using Lemma B.1, all the entries in above matrices are of the order $O_p\left(\frac{1}{\sqrt{nT}}\right)$. ■

Proof for Equation (C.9)

$$-\frac{1}{nT}\frac{\partial^2\ln L_{nT}(\theta)}{\partial\theta\partial\theta'}-(-\frac{1}{nT}E\frac{\partial^2\ln L_{nT}(\theta)}{\partial\theta\partial\theta'})=$$

$$\begin{pmatrix} \frac{1}{\sigma^2}\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}_{nt}'\tilde{Z}_{nt}-\frac{1}{\sigma^2}E(\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}_{nt}'\tilde{Z}_{nt}) & * & * \\ \frac{1}{\sigma^2}\frac{1}{nT}\sum_{t=1}^{T}(W_n\tilde{Y}_{nt})'\tilde{Z}_{nt}-\frac{1}{\sigma^2}E(\frac{1}{nT}\sum_{t=1}^{T}(W_n\tilde{Y}_{nt})'\tilde{Z}_{nt}) & 0 & 0 \\ \frac{1}{\sigma^4}\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}_{nt}'(\zeta)\tilde{Z}_{nt}-\frac{1}{\sigma^4}E(\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}_{nt}'(\zeta)\tilde{Z}_{nt}) & 0 & 0 \end{pmatrix}$$

$$+\begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{\sigma^2}\frac{1}{nT}\sum_{t=1}^{T}(W_n\tilde{Y}_{nt})'W_n\tilde{Y}_{nt} - \frac{1}{\sigma^2}E(\frac{1}{nT}\sum_{t=1}^{T}(W_n\tilde{Y}_{nt})'W_n\tilde{Y}_{nt}) & * \\ 0 & \frac{1}{\sigma^4}\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}'_{nt}(\zeta)W_n\tilde{Y}_{nt} - \frac{1}{\sigma^4}E\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}'_{nt}(\zeta)W_n\tilde{Y}_{nt} & 0 \end{pmatrix}$$

$$+\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sigma^6}\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}'_{nt}(\zeta)\tilde{V}_{nt}(\zeta) - \frac{1}{\sigma^6}E\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}'_{nt}(\zeta)\tilde{V}_{nt}(\zeta) \end{pmatrix}.$$

As $\tilde{V}_{nt}(\zeta) = \tilde{V}_{nt} - (\lambda-\lambda_0)W_n\tilde{Y}_{nt} - \tilde{Z}_{nt}(\delta-\delta_0)$ and $W_n\tilde{Y}_{nt} = G_n\tilde{Z}_{nt}\delta_0 + G_n\tilde{V}_{nt}$, by Lemma B.1, we have $\sup_{\theta\in\Theta}\left|-\frac{1}{nT}\frac{\partial^2 \ln L_{nT}(\theta)}{\partial\theta\partial\theta'} - \Sigma_{\theta_0,nT}\right|_{ij} = O_p\left(\frac{1}{\sqrt{nT}}\right)$ because Θ is bounded. ■

Proof for Equation (C.10)

$$-\frac{1}{nT}E\frac{\partial^2 \ln L_{nT}(\theta)}{\partial\theta\partial\theta'} - \Sigma_{\theta_0,nT} =$$

$$\begin{pmatrix} \frac{1}{\sigma^2}E(\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}'_{nt}\tilde{Z}_{nt}) - \frac{1}{\sigma_0^2}E(\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}'_{nt}\tilde{Z}_{nt}) & * & * \\ \frac{1}{\sigma^2}E(\frac{1}{nT}\sum_{t=1}^{T}(W_n\tilde{Y}_{nt})'\tilde{Z}_{nt}) - \frac{1}{\sigma_0^2}E(\frac{1}{nT}\sum_{t=1}^{T}(W_n\tilde{Y}_{nt})'\tilde{Z}_{nt}) & 0 & 0 \\ \frac{1}{\sigma^4}E(\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}'_{nt}(\zeta)\tilde{Z}'_{nt}) - \frac{1}{\sigma_0^4}E(\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}'_{nt}\tilde{Z}_{nt}) & 0 & 0 \end{pmatrix}$$

$$+\begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{\sigma^2}E(\frac{1}{nT}\sum_{t=1}^{T}(W_n\tilde{Y}_{nt})'W_n\tilde{Y}_{nt}) - \frac{1}{\sigma_0^2}E(\frac{1}{nT}\sum_{t=1}^{T}(W_n\tilde{Y}_{nt})'W_n\tilde{Y}_{nt}) & * \\ 0 & \frac{1}{\sigma^4}E\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}'_{nt}(\zeta)W_n\tilde{Y}_{nt} - \frac{1}{\sigma_0^4}E\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}'_{nt}W_n\tilde{Y}_{nt} & 0 \end{pmatrix}$$

$$+\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sigma^6}E\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}'_{nt}(\zeta)\tilde{V}_{nt}(\zeta) - \frac{1}{\sigma_0^6}E\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}'_{nt}\tilde{V}_{nt} \end{pmatrix}.$$

As $\tilde{V}_{nt}(\zeta) = \tilde{V}_{nt} - (\lambda-\lambda_0)W_n\tilde{Y}_{nt} - \tilde{Z}_{nt}(\delta-\delta_0)$ and $W_n\tilde{Y}_{nt} = G_n\tilde{Z}_{nt}\delta_0 + G_n\tilde{V}_{nt}$, and with expectations of the relevant terms are of orders no larger than $O(1)$ by Lemma B.1, we have $\sup_{\theta\in\Theta_1}\left|-\frac{1}{nT}E\frac{\partial^2 \ln L_{nT}(\theta)}{\partial\theta\partial\theta'} - \Sigma_{\theta,nT}\right|_{ij} = \sup_{\theta\in\Theta_1}\|\theta-\theta_0\|\cdot O(1)$ because Θ_1 is bounded. ■

D Proofs For Claims and Theorems in the Text

D.1 Proof of Claim 3.1

To prove $\frac{1}{nT}\ln L_{n,T}(\theta) - Q_{n,T}(\theta) \xrightarrow{p} 0$ *uniformly in* θ *in any compact parameter space* Θ:

From $\tilde{V}_{nt}(\zeta) = \tilde{Y}_{nt} - \lambda W_n\tilde{Y}_{nt} - \tilde{Z}_{nt}\delta$ and $\tilde{V}_{nt} = \tilde{V}_{nt}(\zeta_0)$,

$$\tilde{V}_{nt}(\zeta) = \tilde{V}_{nt} - (\lambda - \lambda_0) W_n \tilde{Y}_{nt} - \tilde{Z}_{nt}(\delta - \delta_0) \tag{D.1}$$

Hence,

$$\begin{aligned} \tilde{V}'_{nt}(\zeta)\tilde{V}_{nt}(\zeta) &= \tilde{V}'_{nt}\tilde{V}_{nt} + (\lambda-\lambda_0)^2 (W_n\tilde{Y}_{nt})' W_n \tilde{Y}_{nt} + (\delta-\delta_0)'\tilde{Z}'_{nt}\tilde{Z}_{nt}(\delta-\delta_0) \\ &\quad + 2(\lambda-\lambda_0)(W_n\tilde{Y}_{nt})'\tilde{Z}_{nt}(\delta-\delta_0) - 2(\lambda-\lambda_0)(W_n\tilde{Y}_{nt})'\tilde{V}_{nt} - 2(\delta-\delta_0)'\tilde{Z}'_{nt}\tilde{V}_{nt} \end{aligned} \tag{D.2}$$

where, using $W_n\tilde{Y}_{nt} = G_n\tilde{Z}_{nt}\delta_0 + G_n\tilde{V}_{nt}$ which is implied by (C.2),

$$(W_n\tilde{Y}_{nt})'W_n\tilde{Y}_{nt} = (G_n\tilde{Z}_{nt}\delta_0)'(G_n\tilde{Z}_{nt}\delta_0) + 2(G_n\tilde{Z}_{nt}\delta_0)'G_n\tilde{V}_{nt} + (G_n\tilde{V}_{nt})'G_n\tilde{V}_{nt}.$$

Using Lemma B.1,
$\frac{1}{nT}\sum_{t=1}^T \tilde{V}'_{nt}\tilde{V}_{nt} - E\frac{1}{nT}\sum_{t=1}^T \tilde{V}'_{nt}\tilde{V}_{nt} \xrightarrow{p} 0$,
$\frac{1}{nT}\sum_{t=1}^T (W_n\tilde{Y}_{nt})'W_n\tilde{Y}_{nt} - E\frac{1}{nT}\sum_{t=1}^T (W_n\tilde{Y}_{nt})'W_n\tilde{Y}_{nt} \xrightarrow{p} 0$,
$\frac{1}{nT}\sum_{t=1}^T \tilde{Z}'_{nt}\tilde{Z}_{nt} - E\frac{1}{nT}\sum_{t=1}^T \tilde{Z}'_{nt}\tilde{Z}_{nt} \xrightarrow{p} 0$,
$\frac{1}{nT}\sum_{t=1}^T (W_n\tilde{Y}_{nt})'\tilde{Z}_{nt} - E\frac{1}{nT}\sum_{t=1}^T (W_n\tilde{Y}_{nt})'\tilde{Z}_{nt} \xrightarrow{p} 0$,
$\frac{1}{nT}\sum_{t=1}^T (W_n\tilde{Y}_{nt})'\tilde{V}_{nt} - E\frac{1}{nT}\sum_{t=1}^T (W_n\tilde{Y}_{nt})'\tilde{V}_{nt} \xrightarrow{p} 0$ and
$\frac{1}{nT}\sum_{t=1}^T \tilde{Z}'_{nt}\tilde{V}_{nt} - E\frac{1}{nT}\sum_{t=1}^T \tilde{Z}'_{nt}\tilde{V}_{nt} \xrightarrow{p} 0$.
As Θ is compact so that λ, δ are bounded in Θ, we have

$$\frac{1}{nT}\sum_{t=1}^T \tilde{V}'_{nt}(\zeta)\tilde{V}_{nt}(\zeta) - \frac{1}{nT}E\sum_{t=1}^T \tilde{V}'_{nt}(\zeta)\tilde{V}_{nt}(\zeta) \xrightarrow{p} 0 \text{ uniformly in } \theta \text{ in } \Theta.$$

Also, $\frac{1}{nT}\ln L_{n,T}(\theta) = -\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln\sigma^2 + \frac{1}{n}\ln|S_n(\lambda)| - \frac{1}{2\sigma^2 nT}\sum_{t=1}^T \tilde{V}'_{nt}(\zeta)\tilde{V}_{nt}(\zeta)$ and $Q_{n,T}(\theta) = E\frac{1}{nT}\ln L_{n,T}(\theta)$, using the fact that σ^2 is bounded away from zero in Θ,

$$\frac{1}{nT}\ln L_{n,T}(\theta) - Q_{n,T}(\theta) = -\frac{1}{2\sigma^2}\left(\frac{1}{nT}\sum_{t=1}^T \tilde{V}'_{nt}(\zeta)\tilde{V}_{nt}(\zeta) - \frac{1}{nT}E\sum_{t=1}^T \tilde{V}'_{nt}(\zeta)\tilde{V}_{nt}(\zeta)\right) \xrightarrow{p} 0$$

uniformly in θ in any compact parameter space Θ.

To prove $Q_{n,T}(\theta)$ is uniformly equicontinuous in θ in any compact parameter space Θ:

We have $Q_{nT}(\theta) = E\frac{1}{nT}\ln L_{n,T}(\theta) = -\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln\sigma^2 + \frac{1}{n}\ln|S(\lambda)| - \frac{1}{2\sigma^2 nT}E\sum_{t=1}^T \tilde{V}'_{nt}(\zeta)\tilde{V}_{nt}(\zeta)$. As $\tilde{V}_{nt}(\zeta) = S_n(\lambda)\tilde{Y}_{nt} - \tilde{Z}_{nt}\delta$ and $\tilde{Y}_{nt} = S_n^{-1}\tilde{Z}_{nt}\delta_0 + S_n^{-1}\tilde{V}_{nt}$,

$$\tilde{V}_{nt}(\zeta) = S_n(\lambda)S_n^{-1}\tilde{Z}_{nt}\delta_0 - \tilde{Z}_{nt}\delta + S_n(\lambda)S_n^{-1}\tilde{V}_{nt}.$$

Hence,

$$
\begin{aligned}
E\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}'_{nt}(\zeta)\tilde{V}_{nt}(\zeta) &= \frac{1}{nT}E\sum_{t=1}^{T}(S_n(\lambda)S_n^{-1}\tilde{Z}_{nt}\delta_0-\tilde{Z}_{nt}\delta)'(S_n(\lambda)S_n^{-1}\tilde{Z}_{nt}\delta_0-\tilde{Z}_{nt}\delta) \\
&+\frac{1}{n}\frac{T-1}{T}\sigma_0^2 tr(S_n^{-1'}S'_n(\lambda)S_n(\lambda)S_n^{-1}) \\
&+\frac{2}{nT}E\sum_{t=1}^{T}(S_n(\lambda)S_n^{-1}\tilde{Z}_{nt}\delta_0-\tilde{Z}_{nt}\delta)'S_n(\lambda)S_n^{-1}\tilde{V}_{nt}.
\end{aligned} \tag{D.3}
$$

The third term $\frac{2}{nT}E\sum_{t=1}^{T}(S_n(\lambda)S_n^{-1}\tilde{Z}_{nt}\delta_0-\tilde{Z}_{nt}\delta)'S_n(\lambda)S_n^{-1}\tilde{V}_{nt}$ is $O\left(\frac{1}{T}\right)$ according to expected values in Lemma B.1 and the order $O\left(\frac{1}{T}\right)$ is uniformly in θ in Θ because it is a polynomial function in θ and Θ is a bounded set. The first term is equal to $(\delta'-\delta'_0,\lambda-\lambda_0)E\mathcal{H}_{nT}(\delta'-\delta'_0,\lambda-\lambda_0)'$ using $S_n(\lambda)S_n^{-1}=I_n-(\lambda-\lambda_0)G_n$; the second term is equal to $\frac{T-1}{T}\sigma_n^2(\lambda)$ where $\sigma_n^2(\lambda)=\frac{\sigma_0^2}{n}tr(S_n'^{-1}S'_n(\lambda)S_n(\lambda)S_n^{-1})$, which are all polynomial functions of θ. To prove $Q_{n,T}(\theta)$ is uniformly equicontinuous in θ in the compact parameter space Θ, the followings are sufficient: (1) $\ln\sigma^2$ is uniformly continuous; (2) $\frac{1}{n}\ln|S_n(\lambda)|$ is uniformly equicontinuous; (3) $(\delta'-\delta'_0,\lambda-\lambda_0)\mathcal{H}_{nT}(\delta'-\delta'_0,\lambda-\lambda_0)'$ is uniformly equicontinuous; (4) $\sigma_n^2(\lambda)$ is uniformly equicontinuous.

(1) is obvious because σ^2 is bounded away from zero in Θ. For (2), $\frac{1}{n}\ln|S_n(\lambda_2)|-\frac{1}{n}\ln|S_n(\lambda_1)|=\frac{1}{n}tr\left(W_nS_n^{-1}\left(\bar{\lambda}\right)\right)(\lambda_2-\lambda_1)$ where $\bar{\lambda}$ lies between λ_2 and λ_1. As $S_n^{-1}(\lambda)$ is uniformly bounded in row and column sums, uniformly in $\theta\in\Theta$, $\frac{1}{n}tr\left(W_nS_n^{-1}\left(\bar{\lambda}\right)\right)$ is bounded, we have $\frac{1}{n}\ln|S(\lambda)|$ is uniformly equicontinuous. For (3), because δ and λ are bounded and because $E\mathcal{H}_{nT}$ is $O(1)$ according to Lemma B.1, the result follows. For (4), $\sigma_n^2(\lambda_2)-\sigma_n^2(\lambda_1)=\frac{\sigma_0^2}{n}tr(S_n'^{-1}S'_n(\lambda_2)S_n(\lambda_2)S_n^{-1})-\frac{\sigma_0^2}{n}tr(S_n'^{-1}S'_n(\lambda_1)S_n(\lambda_1)S_n^{-1})$. Using $S_n(\lambda)S_n^{-1}=I_n-(\lambda-\lambda_0)G_n$, $\sigma_n^2(\lambda_2)-\sigma_n^2(\lambda_1)=\sigma_0^2\left[(\lambda_2-\lambda_1)(\lambda_2+\lambda_1-2\lambda_0)\frac{trG'_nG_n}{n}-(\lambda_2-\lambda_1)\frac{tr\left(G'_n+G_n\right)}{n}\right]$. As G'_nG_n and G_n are uniformly bounded in row and column sums, $\sigma_n^2(\lambda)$ is uniformly equicontinuous. ■

D.2 Proof of nonsingularity of the information matrix (page 8, footnote 3)

We can prove the result by using an argument by contradiction (similar to Lee (2004)). For $\Sigma_{\theta_0}\equiv\lim_{T\to\infty}\Sigma_{\theta_0,nT}$, where $\Sigma_{\theta_0,nT}$ is (C.6) in Appendix C, we need to prove that $\Sigma_{\theta_0}\alpha=0$ implies $\alpha=0$ where $\alpha=(\alpha'_1,\alpha_2,\alpha_3)'$ and α_2,α_3 are scalars and α_1 is $(k_x+2)\times 1$ vector. If this is true, then, columns of Σ_{θ_0} would be linear independent so that Σ_{θ_0} would be nonsingular. Denote $\mathcal{H}_\delta=\lim_{T\to\infty}\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}'_{nt}\tilde{Z}_{nt}$, $\mathcal{H}_{\delta\lambda}=\lim_{T\to\infty}\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}'_{nt}G_n\tilde{Z}_{nt}\delta_0$, $\mathcal{H}_{\lambda\delta}=\mathcal{H}'_{\delta\lambda}$ and $\mathcal{H}_\lambda=\lim_{T\to\infty}\frac{1}{nT}\sum_{t=1}^{T}(G_n\tilde{Z}_{nt}\delta_0)'G_n\tilde{Z}_{nt}\delta_0$, then

$$
\Sigma_{\theta_0}=\frac{1}{\sigma_0^2}\begin{pmatrix} E\mathcal{H}_\delta & E\mathcal{H}_{\delta\lambda} & 0 \\ E\mathcal{H}_{\lambda\delta} & E\mathcal{H}_\lambda+\lim_{n\to\infty}\frac{\sigma_0^2}{n}\left[tr(G'_nG_n)+tr(G_n^2)\right] & \lim_{n\to\infty}\frac{1}{n}tr(G_n) \\ 0 & \lim_{n\to\infty}\frac{1}{n}tr(G_n) & \frac{1}{2\sigma_0^2}\end{pmatrix}.
$$

Hence, $\Sigma_{\theta_0}\alpha=0$ implies

$$\frac{1}{\sigma_0^2} E\mathcal{H}_\delta \times \alpha_1 + \frac{1}{\sigma_0^2} E\mathcal{H}_{\delta\lambda} \times \alpha_2 = 0,$$

$$\frac{1}{\sigma_0^2} E\mathcal{H}_{\lambda\delta} \times \alpha_1 + \left(\frac{1}{\sigma_0^2} E\mathcal{H}_\lambda + \lim_{n\to\infty} \frac{1}{n} \left[tr(G_n' G_n) + tr(G_n^2) \right] \right) \times \alpha_2 + \lim_{n\to\infty} \frac{1}{\sigma_0^2 n} tr(G_n) \times \alpha_3 = 0,$$

$$\lim_{n\to\infty} \frac{1}{\sigma_0^2 n} tr(G_n) \times \alpha_2 + \frac{1}{2\sigma_0^4} \times \alpha_3 = 0.$$

From the first equation, $\alpha_1 = -(E\mathcal{H}_\delta)^{-1} E\mathcal{H}_{\delta\lambda} \times \alpha_2$; from the third equation, $\alpha_3 = -2\lim_{n\to\infty} \frac{\sigma_0^2}{n} tr(G_n) \times \alpha_2$. By eliminating α_1 and α_3, the remaining equation becomes

$$\left\{ \left(\frac{1}{\sigma_0^2} \left(E\mathcal{H}_\lambda - E\mathcal{H}_{\lambda\delta} (E\mathcal{H}_\delta)^{-1} E\mathcal{H}_{\delta\lambda} \right) \right) + \lim_{n\to\infty} \frac{1}{n} \left[tr(G_n' G_n) + tr(G_n^2) - 2\frac{tr^2(G_n)}{n} \right] \right\} \times \alpha_2 = 0.$$

Because $tr(G_n' G_n) + tr(G_n^2) - 2\frac{tr^2(G_n)}{n} = \frac{1}{2} tr\left[(C_n' + C_n)(C_n' + C_n)' \right] \geq 0$ where $C_n = G_n - \frac{trG_n}{n} I_n$, combined with the condition that $\lim_{T\to\infty} E\mathcal{H}_{nT}$ is nonsingular, we have $\alpha_2 = 0$ and hence $\alpha = 0$. ■

D.3 Proof of Theorem 3.2

As $E\sum_{t=1}^{T} \tilde{V}_{nt}' \tilde{V}_{nt} = n(T-1)\sigma_0^2$ according to Lemma A.9, at θ_0, the expected log likelihood from (3.1) implies

$$E \ln L_{n,T}(\theta_0) = -\frac{nT}{2} \ln 2\pi - \frac{nT}{2} \ln \sigma_0^2 + T \ln |S_n| - \frac{n(T-1)}{2}.$$

Denote $\sigma_n^2(\lambda) = \frac{\sigma_0^2}{n} tr(S_n^{-1\prime} S_n'(\lambda) S_n(\lambda) S_n^{-1})$. By using $S_n(\lambda) S_n^{-1} = I_n + (\lambda_0 - \lambda) G_n$ for (D.3), it follows that

$$\begin{aligned} &\frac{1}{nT} E \ln L_{n,T}(\theta) - \frac{1}{nT} E \ln L_{n,T}(\theta_0) \\ &= -\frac{1}{2}(\ln \sigma^2 - \ln \sigma_0^2) + \frac{1}{n} \ln |S_n(\lambda)| - \frac{1}{n} \ln |S_n| - \left(\frac{1}{2\sigma^2} \frac{1}{nT} \sum_{t=1}^{T} E\tilde{V}_{nt}'(\zeta) \tilde{V}_{nt}(\zeta) - \frac{T-1}{2T} \right) \\ &= T_{1,n}(\lambda, \sigma^2) - \frac{1}{2\sigma^2} T_{2,n,T}(\delta, \lambda) + o(1) \end{aligned}$$

where $T_{1,n}(\lambda, \sigma^2) = -\frac{1}{2}(\ln \sigma^2 - \ln \sigma_0^2) + \frac{1}{n} \ln |S_n(\lambda)| - \frac{1}{n} \ln |S_n| - \frac{1}{2\sigma^2}(\sigma_n^2(\lambda) - \sigma^2)$ and

$$T_{2,n,T}(\delta, \lambda) = \frac{1}{nT} \sum_{t=1}^{T} E \left\{ (\tilde{Z}_{nt}(\delta_0 - \delta) + (\lambda_0 - \lambda) G_n \tilde{Z}_{nt} \delta_0)' (\tilde{Z}_{nt}(\delta_0 - \delta) + (\lambda_0 - \lambda) G_n \tilde{Z}_{nt} \delta_0) \right\}.$$

Consider the pure spatial process $Y_{nt} = \lambda_0 W_n Y_{nt} + V_{nt}$ for a period t, the log likelihood function of this process is

$$\ln L_{p,n}(\lambda, \sigma^2) = -\frac{n}{2} \ln 2\pi - \frac{n}{2} \ln \sigma^2 + \ln |S_n(\lambda)| - \frac{1}{2\sigma^2} (S_n(\lambda) Y_{nt})' S_n(\lambda) Y_{nt}. \quad \text{(D.4)}$$

Let $E_p(\cdot)$ be the expectation operator for Y_{nt} based on this pure spatial autoregressive process. It follows that

$E_p(\frac{1}{n}\ln L_{p,n}(\lambda,\sigma^2)) - E_p(\frac{1}{n}\ln L_{p,n}(\lambda_0,\sigma_0^2))$

$= -\frac{1}{2}(\ln\sigma^2 - \ln\sigma_0^2) + \frac{1}{n}\ln|S_n(\lambda)| - \frac{1}{n}\ln|S_n(\lambda_0)| - \frac{1}{2\sigma^2}(\sigma_n^2(\lambda)-\sigma^2)$, which equals to $T_{1,n}(\lambda,\sigma^2)$. By the information inequality, $\ln L_{p,n}(\lambda,\sigma^2) - \ln L_{p,n}(\lambda_0,\sigma_0^2) \le 0$. Thus, $T_{1,n}(\lambda,\sigma^2)\le 0$ for any (λ,σ^2).

For $T_{2,n,T}(\delta,\lambda)$, it is a quadratic function of δ and λ. Under the assumed condition that $\lim_{T\to\infty} E\mathcal{H}_{nT}$ is nonsingular, $T_{2,n,T}(\delta,\lambda) > 0$ whenever $(\delta,\lambda)\neq(\delta_0,\lambda_0)$, so, (δ,λ) is globally identified. Given λ_0, σ_0^2 is the unique maximizer of $T_{1,n}(\lambda_0,\sigma^2)$. Hence, $(\delta,\lambda,\sigma^2)$ is globally identified.

Combined with uniform convergence and equicontinuity in Claim 3.1, the consistency follows. ■

D.4 Proof of Theorem 3.3

From the Proof of Theorem 3.2 (see Appendix D.3),

$$\frac{1}{nT}E\ln L_{n,T}(\theta) - \frac{1}{nT}E\ln L_{n,T}(\theta_0) = T_{1,n}(\lambda,\sigma^2) - \frac{1}{2\sigma^2}T_{2,n,T}(\delta,\lambda) + o(1).$$

When $E\begin{pmatrix}\mathcal{H}_\delta & \mathcal{H}_{\delta\lambda}\\ \mathcal{H}'_{\delta\lambda} & \mathcal{H}_\lambda\end{pmatrix}$ is singular, δ_0 and λ_0 cannot be identified from $T_{2,n,T}(\delta,\lambda)$. Global identification requires that the limit of $T_{1,n}(\lambda,\sigma^2)$ is strictly less than zero. As $T_{1,n}(\lambda,\sigma^2)\le 0$ by the information inequality, $T_{1,n}(\lambda,\sigma^2)\neq 0$ is equivalent to $\frac{1}{n}\ln\left|\sigma_0^2 S_n^{-1}S_n^{-1\prime}\right| \neq \frac{1}{n}\ln\left|\sigma_n^2(\lambda)S_n^{-1}(\lambda)S_n^{-1\prime}(\lambda)\right|$ (see Lee (2004), Proof of Theorem 4.1). After λ_0 and σ_0^2 are identified, given λ_0, δ_0 can be identified from $T_{2,n,T}(\delta,\lambda)$. Combined with uniform convergence and equicontinuity in Claim 3.1, the consistency follows. ■

D.5 Proof of Claim 3.4

Denote $Z_{nt} = (Y_{n,t-1},\ W_nY_{n,t-1},\ X_{nt})$, then from Equation (2.3),

$$\tilde{Z}_{nt} = \tilde{Z}^*_{nt} - (\bar{U}_{nT,-1},\ W_n\bar{U}_{nT,-1},\ \mathbf{0})$$

where $\tilde{Z}^*_{nt} = ((\tilde{\tilde{\mathcal{X}}}_{n,t-1} + U_{n,t-1}),\ \ (W_n\tilde{\tilde{\mathcal{X}}}_{n,t-1} + W_nU_{n,t-1}),\ \ \tilde{X}_{nt})$ with $\tilde{\tilde{\mathcal{X}}}_{n,t-1} = \mathcal{X}_{n,t-1} - \bar{\mathcal{X}}_{nT,-1}$ and $\mathcal{X}_{n,t-1}$ is . Hence, Z_{nt} has two components: one is $\tilde{Z}^*_{nt}$, which is uncorrelated with V_{nt}; the other is $-(\bar{U}_{nT,-1},\ W_n\bar{U}_{nT,-1},\ \mathbf{0})$, which is correlated with V_{nt} when $t\le T-1$. Correspondingly, $\frac{1}{\sqrt{nT}}\frac{\partial\ln L_{n,T}(\theta_0)}{\partial\theta} = \frac{1}{\sqrt{nT}}\frac{\partial\ln L^*_{n,T}(\theta_0)}{\partial\theta} - \Delta_{nT}$ where $\frac{1}{\sqrt{nT}}\frac{\partial\ln L^*_{n,T}(\theta_0)}{\partial\theta}$ is in Equation (3.4) and Δ_{nT} is in Equation (3.5).

For $\frac{1}{\sqrt{nT}}\frac{\partial\ln L^*_{n,T}(\theta_0)}{\partial\theta}$, the central limit theorem of martingale difference arrays (Theorem 2.4) can be applied. For Δ_{nT}, using Theorem 2.3, it is equal to $\sqrt{\frac{n}{T}}b_{2n} + O\left(\sqrt{\frac{n}{T^3}}\right) + O_p\left(\sqrt{\frac{1}{T}}\right)$ where b_{2n} is $O(1)$ in (3.8). ■

D.6 Proof of Claim 3.5

This is Equation (C.7).

D.7 Proof of Claim 3.6

This is Equation (C.8).

D.8 Proof of Theorem 3.7

Equation (3.10) follows from the Taylor expansion $\sqrt{nT}(\hat{\theta}_{nT}-\theta_0) = (-\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\bar{\theta}_{nT})}{\partial\theta\partial\theta'})^{-1}\frac{1}{\sqrt{nT}}\frac{\partial \ln L_{n,T}(\theta_0)}{\partial\theta}$ where $\bar{\theta}_{nT}$ lies between θ_0 and $\hat{\theta}_{nT}$.

As $-\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\bar{\theta}_{nT})}{\partial\theta\partial\theta'} = \left(-\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\bar{\theta}_{nT})}{\partial\theta\partial\theta'} - \left(-\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\theta_0)}{\partial\theta\partial\theta'}\right)\right) + \left(-\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\theta_0)}{\partial\theta\partial\theta'} - \Sigma_{\theta_0,nT}\right) + \Sigma_{\theta_0,nT}$ where the first term is $\left\|\bar{\theta}_{nT}-\theta_0\right\|\cdot O_p(1)$ (Equation (C.7)) and the second term is $O_p\left(\frac{1}{\sqrt{nT}}\right)$ (Equation (C.8)), we have $-\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\bar{\theta}_{nT})}{\partial\theta\partial\theta'} = \left\|\bar{\theta}_{nT}-\theta_0\right\|\cdot O_p(1) + O_p\left(\frac{1}{\sqrt{nT}}\right) + \Sigma_{\theta_0,nT}$. Because (1) $\left\|\bar{\theta}_{nT}-\theta_0\right\| = o_p(1)$ as $\hat{\theta}_{nT}$ is consistent and (2) $\Sigma_{\theta_0,nT}$ is the nonsingular in the limit according to Assumption 8, we have $-\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\bar{\theta}_{nT})}{\partial\theta\partial\theta'}$ is invertible for large n and T and $\left(-\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\bar{\theta}_{nT})}{\partial\theta\partial\theta'}\right)^{-1}$ is $O_p(1)$.

According to Taylor expansion, $\sqrt{nT}(\hat{\theta}_{nT}-\theta_0) = \left(-\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\bar{\theta}_{nT})}{\partial\theta\partial\theta'}\right)^{-1}\cdot\left(\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\theta_0)}{\partial\theta} - \Delta_{nT}\right)$ where $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\theta_0)}{\partial\theta} \xrightarrow{d} N(0, \Sigma_{\theta_0}+\Omega_{\theta_0})$ and $\Delta_{nT} = \sqrt{\frac{n}{T}}b_{2n} + O\left(\sqrt{\frac{n}{T^3}}\right) + O_p\left(\sqrt{\frac{1}{T}}\right)$ with $b_{2n} = O(1)$. Then, $\sqrt{nT}(\hat{\theta}_{nT}-\theta_0) = O_p(1)\cdot\left(O_p(1) + O\left(\sqrt{\frac{n}{T}}\right)\right)$, which implies that

$$\hat{\theta}_{nT}-\theta_0 = O_p\left(\max\left(\sqrt{\frac{1}{nT}},\frac{1}{T}\right)\right). \tag{D.5}$$

Hence,

$$\begin{aligned}\sqrt{nT}(\hat{\theta}_{nT}-\theta_0) &= \left(-\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\bar{\theta}_{nT})}{\partial\theta\partial\theta'}\right)^{-1}\cdot\left(\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\theta_0)}{\partial\theta} - \Delta_{nT}\right) \\ &= \left(\Sigma_{\theta_0,nT} + O_p\left(\max\left(\sqrt{\frac{1}{nT}},\frac{1}{T}\right)\right)\right)^{-1}\cdot\left(\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\theta_0)}{\partial\theta} - \Delta_{nT}\right). \quad \text{(D.6)}\end{aligned}$$

Using the fact that[8]

$$\left(\Sigma_{\theta_0,nT} + O_p\left(\max\left(\sqrt{\frac{1}{nT}},\frac{1}{T}\right)\right)\right)^{-1} = \Sigma^{-1}_{\theta_0,nT} + O_p\left(\max\left(\sqrt{\frac{1}{nT}},\frac{1}{T}\right)\right) \tag{D.7}$$

given that $\Sigma_{\theta_0,nT}$ is nonsingular, we have

[8] For two matrices C_k and D_k which are nonsingular and $C_k - D_k = O_p(T^{-\eta})$ for $\eta > 0$, we have $C_k^{-1} - D_k^{-1} = C_k^{-1}(D_k - C_k)D_k^{-1} = O_p(T^{-\eta})$.

$$\begin{aligned}\sqrt{nT}(\hat{\theta}_{nT}-\theta_0) &= \left(\Sigma_{\theta_0,nT}^{-1}+O_p\left(\max\left(\sqrt{\frac{1}{nT}},\frac{1}{T}\right)\right)\right)\cdot\left(\frac{1}{\sqrt{nT}}\frac{\partial \ln L_{n,T}^*(\theta_0)}{\partial\theta}-\Delta_{nT}\right)\\
&= \Sigma_{\theta_0,nT}^{-1}\cdot\frac{1}{\sqrt{nT}}\frac{\partial \ln L_{n,T}^*(\theta_0)}{\partial\theta}+O_p\left(\max\left(\sqrt{\frac{1}{nT}},\frac{1}{T}\right)\right)\cdot\frac{1}{\sqrt{nT}}\frac{\partial \ln L_{n,T}^*(\theta_0)}{\partial\theta}\\
&\quad -\Sigma_{\theta_0,nT}^{-1}\cdot\Delta_{nT}-O_p\left(\max\left(\sqrt{\frac{1}{nT}},\frac{1}{T}\right)\right)\cdot\Delta_{nT},\end{aligned}$$

which implies that

$$\begin{aligned}&\sqrt{nT}(\hat{\theta}_{nT}-\theta_0)+\Sigma_{\theta_0,nT}^{-1}\cdot\Delta_{nT}+O_p\left(\max\left(\sqrt{\frac{1}{nT}},\frac{1}{T}\right)\right)\Delta_{nT}\\
=\;&(\Sigma_{\theta_0,nT}^{-1}+o_p(1))\cdot\frac{1}{\sqrt{nT}}\frac{\partial \ln L_{n,T}^*(\theta_0)}{\partial\theta}.\end{aligned}\tag{D.8}$$

As $\Sigma_{\theta_0}=\lim_{T\to\infty}\Sigma_{\theta_0,nT}$ exists, then using Claim 3.4 and that $\Delta_{nT}=\sqrt{\frac{n}{T}}b_{2n}+O\left(\sqrt{\frac{n}{T^3}}\right)+O_p\left(\sqrt{\frac{1}{T}}\right)$,

$$\sqrt{nT}(\hat{\theta}_{nT}-\theta_0)+\sqrt{\frac{n}{T}}\Sigma_{\theta_0,nT}^{-1}b_{2n}+O_p\left(\max\left(\sqrt{\frac{n}{T^3}},\sqrt{\frac{1}{T}}\right)\right)\xrightarrow{d}N(0,\Sigma_{\theta_0}^{-1}(\Sigma_{\theta_0}+\Omega_{\theta_0})\Sigma_{\theta_0}^{-1}).\ \blacksquare\tag{D.9}$$

D.9 Proof of Theorem 3.8

From the first order condition that $\frac{\partial \ln L_{n,T}(\theta,\mathbf{c}_n)}{\partial \mathbf{c}_n}=\frac{1}{\sigma^2}\sum_{t=1}^{T}V_{nt}(\zeta)$, we have $\hat{\mathbf{c}}_{nT}(\theta)=\frac{1}{T}\sum_{t=1}^{T}(S_n(\lambda)Y_{nt}-Z_{nt}\delta)$. As $S_nY_{nt}=Z_{nt}\delta_0+\mathbf{c}_{n0}+V_{nt}$ and $S_n(\lambda)S_n^{-1}=I_n-(\lambda-\lambda_0)G_n$, it implies that $\hat{\mathbf{c}}_{nT}(\theta)=\frac{1}{T}\sum_{t=1}^{T}\left((I_n-(\lambda-\lambda_0)G_n)\left(Z_{nt}\delta_0+\mathbf{c}_{n0}+V_{nt}\right)-Z_{nt}\delta\right)$. Hence,

$$\begin{aligned}\hat{\mathbf{c}}_{nT}(\theta)-\mathbf{c}_{n0} &= \frac{1}{T}\sum_{t=1}^{T}\left((I_n-(\lambda-\lambda_0)G_n)\left(Z_{nt}\delta_0+\mathbf{c}_{n0}+V_{nt}\right)-Z_{nt}\delta\right)-\mathbf{c}_{n0}\\
&= -\frac{1}{T}\sum_{t=1}^{T}\left[Z_{nt}\left(\delta-\delta_0\right)+\left(\lambda-\lambda_0\right)\left(G_n\mathbf{c}_{n0}+G_nZ_{nt}\delta_0\right)-\left(I_n-(\lambda-\lambda_0)G_n\right)V_{nt}\right].\end{aligned}$$

That means, for each fixed effect,

$$\hat{c}_{i,nT}(\hat{\theta}_{nT})-c_{i,0}=-\frac{1}{T}\sum_{t=1}^{T}((G_n\mathbf{c}_{n0}+G_nZ_{nt}\delta_0)_i\,,\,(Z_{nt})_i)\times\begin{pmatrix}\hat{\lambda}_{nT}-\lambda_0\\ \hat{\delta}_{nT}-\delta_0\end{pmatrix}+\frac{1}{T}\sum_{t=1}^{T}\left\{\left(I_n-(\hat{\lambda}_{nT}-\lambda_0)G_n\right)V_{nt}\right\}_i\tag{D.10}$$

where $(Z_{nt})_i$ is the ith row of Z_{nt} and $(G_n\mathbf{c}_{n0} + G_n Z_{nt}\delta_0)_i$ is the ith element of $(G_n\mathbf{c}_{n0} + G_n Z_{nt}\delta_0)$. As elements of $\frac{1}{T}\sum_{t=1}^{T}((G_n\mathbf{c}_{n0} + G_n Z_{nt}\delta_0)_i, (Z_{nt})_i)$ are $O_p(1)$ uniformly in n and i implied by Lemma B.4 and $\hat{\theta}_{nT} = O_p\left(\max\left(\sqrt{\frac{1}{nT}}, \frac{1}{T}\right)\right)$ implied by Theorem 3.7, the dominant term of $\hat{c}_{i,nT}(\hat{\theta}_{nT}) - c_{i,0}$ would be $\frac{1}{T}\sum_{t=1}^{T} v_{it}$. So, for each fixed effect,

$$\sqrt{T}\left(\hat{c}_{i,nT}(\hat{\theta}_{nT}) - c_0\right) \xrightarrow{d} N(0, \sigma_0^2) \tag{D.11}$$

and they are independent from each other asymptotically. ■

D.10 Proof of Lemma 3.9

Step 1: $\|A_n(\theta_0))\| < 1$ for some matrix norm implies $\|A_n(\theta)\| < 1$ uniformly in a neighborhood of θ_0.

To show this,

$$\begin{aligned} A_n(\theta) - A_n(\theta_0) &= S_n^{-1}(\lambda)(\gamma I_n + \rho W_n) - S_n^{-1}(\lambda_0)(\gamma_0 I_n + \rho_0 W_n) \\ &= [S_n^{-1}(\lambda) - S_n^{-1}(\lambda_0)](\gamma I_n + \rho W_n) + S_n^{-1}(\lambda_0)[(\gamma - \gamma_0)I_n + (\rho - \rho_0)W_n] \\ &= (\lambda - \lambda_0)S_n^{-1}(\lambda)W_n S_n^{-1}(\lambda_0)(\gamma I_n + \rho W_n) + S_n^{-1}(\lambda_0)[(\gamma - \gamma_0)I_n + (\rho - \rho_0)W_n], \end{aligned}$$

because $S_n^{-1}(\lambda) - S_n^{-1}(\lambda_0) = S_n^{-1}(\lambda)[S_n(\lambda_0) - S_n(\lambda)]S_n^{-1}(\lambda_0) = (\lambda - \lambda_0)S_n^{-1}(\lambda)W_n S_n^{-1}(\lambda_0)$. According to Lemma A.3 in Lee (2004) that $S_n^{-1}(\lambda_0)$ is uniformly bounded in row and column sums norms, it implies that $S_n^{-1}(\lambda)$ will be uniformly bounded in row and column sums, uniformly in λ in a neighborhood of λ_0. From above expression, it becomes apparent that, for any $\epsilon > 0$, there exists a small neighborhood on which $||A_n(\theta)|| - ||A_n(\theta_0)|| < \epsilon$. Therefore, when $||A_n(\theta_0)|| < 1$, it implies $||A_n(\theta)|| < 1$ for θ in a neighborhood $N(\theta_0)$ of θ_0.

Step 2. Let $\delta = \sup_{\theta \in N(\theta_0)} ||A_n(\theta)||$. From Step 1, $\delta < 1$. Note that $||A_n(\theta)|| < 1$ implies that $I_n - A_n(\theta)$ is invertible and $(I_n - A_n(\theta))^{-1} = \sum_{h=0}^{\infty} A_n^h(\theta)$ (Corollary 5.6.16 in Horn and Johnson (1985)). Therefore,

$$\sup_{\theta \in N(\theta_0)} ||I_n - A_n(\theta)||^{-1} \leq \sup_{\theta \in N(\theta_0)} \sum_{h=0}^{\infty} ||A_n^h(\theta)|| \leq \sum_{h=0}^{\infty} \delta^h = \frac{1}{1-\delta}.$$

Step 3. Let $\mathbb{A}_n(\theta) = \sum_{h=1}^{\infty} h A_n^{h-1}(\theta) = I_n + 2A_n(\theta) + 3A_n^2(\theta) + \cdots$. It follows that $A_n(\theta)\mathbb{A}_n(\theta) = A_n(\theta) + 2A_n^2(\theta) + 3A_n^3(\theta) + \cdots$. Hence, $\mathbb{A}_n(\theta) - A_n(\theta)\mathbb{A}_n(\theta) = \sum_{h=0}^{\infty} A_n^h(\theta)$ and $\mathbb{A}_n(\theta) = (I_n - A_n(\theta))^{-1}\sum_{h=0}^{\infty} A_n^h(\theta) = (I_n - A_n(\theta))^{-2}$. Therefore, $\sup_{\theta \in N(\theta_0)} ||\mathbb{A}_n(\theta)|| \leq \sup_{\theta \in N(\theta_0)} ||I_n - A_n(\theta)||^{-2} = \frac{1}{(1-\delta)^2} < \infty$. ■

D.11 Proof of Theorem 3.10

Theorem 3.7 states that $\sqrt{nT}(\hat{\theta}_{nT}-\theta_0)+\sqrt{\frac{n}{T}}\Sigma^{-1}_{\theta_0,nT}b_{2n}+O_p\left(\max\left(\sqrt{\frac{n}{T^3}},\sqrt{\frac{1}{T}}\right)\right)\xrightarrow{d}N(0,\Sigma^{-1}_{\theta_0}(\Sigma_{\theta_0}+\Omega_{\theta_0})\Sigma^{-1}_{\theta_0})$. As the bias corrected estimator $\hat{\theta}^1_{nT}=\hat{\theta}_{nT}+\frac{1}{T}(-\frac{1}{nT}E\frac{\partial^2\ln L_{nT}(\hat{\theta}_{nT})}{\partial\theta\partial\theta'})^{-1}b_{2n}(\hat{\theta}_{nT})$, $\sqrt{nT}(\hat{\theta}^1_{nT}-\theta_0)\xrightarrow{d}N(0,\Sigma^{-1}_{\theta_0}(\Sigma_{\theta_0}+\Omega_{\theta_0})\Sigma^{-1}_{\theta_0})$ if $\sqrt{\frac{n}{T}}\left(\left(-\frac{1}{nT}E\frac{\partial^2\ln L_{nT}(\hat{\theta}_{nT})}{\partial\theta\partial\theta'}\right)^{-1}b_{2n}(\hat{\theta}_{nT})-\Sigma^{-1}_{\theta_0,nT}b_{2n}(\theta_0)\right)\xrightarrow{p}0$ and $\frac{n}{T^3}\to 0$. Assuming that $\frac{n}{T^3}\to 0$, we are going to prove that

$$\sqrt{\frac{n}{T}}\left(\left(-\frac{1}{nT}E\frac{\partial^2\ln L_{nT}(\hat{\theta}_{nT})}{\partial\theta\partial\theta'}\right)^{-1}b_{2n}(\hat{\theta}_{nT})-\Sigma^{-1}_{\theta_0,nT}b_{2n}(\theta_0)\right)\xrightarrow{p}0. \tag{D.12}$$

From Equation (D.7) and that $\hat{\theta}_{nT}-\theta_0=O_p\left(\max\left(\frac{1}{T},\frac{1}{\sqrt{nT}}\right)\right)$ from Equation (D.5),

$$-\frac{1}{nT}E\frac{\partial^2\ln L_{nT}(\hat{\theta}_{nT})}{\partial\theta\partial\theta'}=\Sigma^{-1}_{\theta_0,nT}+O_p\left(\max\left(\frac{1}{T},\frac{1}{\sqrt{nT}}\right)\right).$$

Hence,

$$\begin{aligned}
&\sqrt{\frac{n}{T}}\left\{\left(-\frac{1}{nT}E\frac{\partial^2\ln L_{nT}(\hat{\theta}_{nT})}{\partial\theta\partial\theta'}\right)^{-1}b_{2n}(\hat{\theta}_{nT})-\Sigma^{-1}_{\theta_0,nT}b_{2n}(\theta_0)\right\}\\
=&\sqrt{\frac{n}{T}}\left\{\left(\Sigma^{-1}_{\theta_0,nT}+O_p\left(\max\left(\frac{1}{T},\frac{1}{\sqrt{nT}}\right)\right)\right)b_{2n}(\hat{\theta}_{nT})-\Sigma^{-1}_{\theta_0,nT}b_{2n}(\theta_0)\right\}\\
=&\sqrt{\frac{n}{T}}\left\{\Sigma^{-1}_{\theta_0,nT}\left(b_{2n}(\hat{\theta}_{nT})-b_{2n}(\theta_0)\right)\right\}+\sqrt{\frac{n}{T}}b_{2n}(\hat{\theta}_{nT})\times O_p\left(\max\left(\frac{1}{T},\frac{1}{\sqrt{nT}}\right)\right).
\end{aligned}$$

As $\hat{\theta}_{nT}-\theta_0=O_p\left(\max\left(\frac{1}{T},\frac{1}{\sqrt{nT}}\right)\right)$ and $b_{2n}(\theta_0)$ is $O(1)$, according to Taylor expansion of $b_{2n}(\hat{\theta}_{nT})$ around $b_{2n}(\theta_0)$, to prove Equation (D.12) is reduced to prove that elements of $\frac{\partial b_{2n}(\bar{\theta}_{nT})}{\partial\theta'}<\infty$ where $\bar{\theta}_{nT}$ lies between $\hat{\theta}_{nT}$ and θ_0 and

$$b_{2n}(\theta)=\begin{pmatrix} b^{\delta}_{2n}(\theta)\\ b^{\lambda}_{2n}(\theta)\\ b^{\sigma^2}_{2n}(\theta)\end{pmatrix}=\begin{pmatrix}\frac{1}{n}tr\left(\left(\sum_{h=0}^{\infty}A^h_n(\theta)\right)S^{-1}_n(\lambda)\right)\\ \frac{1}{n}tr\left(W_n\left(\sum_{h=0}^{\infty}A^h_n(\theta)\right)S^{-1}_n(\lambda)\right)\\ \mathbf{0}\\ \frac{1}{n}\gamma tr(G_n(\lambda)\left(\sum_{h=0}^{\infty}A^h_n(\theta)\right)S^{-1}_n(\lambda))+\frac{1}{n}\rho tr(G_n(\lambda)W_n\left(\sum_{h=0}^{\infty}A^h_n(\theta)\right)S^{-1}_n(\lambda))+\frac{1}{n}trG_n(\lambda)\\ \frac{1}{2\sigma^2}\end{pmatrix}.$$

As $A_n(\theta)=S^{-1}_n(\lambda)(\gamma I_n+\rho W_n)$ and $G_n(\lambda)=W_nS^{-1}_n(\lambda)$, we have $\frac{\partial A_n(\theta)}{\partial\gamma}=S^{-1}_n(\lambda)$, $\frac{\partial A_n(\theta)}{\partial\rho}=S^{-1}_n(\lambda)W_n$, $\frac{\partial A_n(\theta)}{\partial\beta_i}=0$ for $i=1,2,\cdots,k_x$ and $\frac{\partial A_n(\theta)}{\partial\lambda}=S^{-1}_n(\lambda)W_nS^{-1}_n(\lambda)(\gamma I_n+\rho W_n)$. Because[9] $\frac{\partial A^h_n(\theta)}{\partial\theta'}=hA^{h-1}_n(\theta)\frac{\partial A_n(\theta)}{\partial\theta'}$ for $h\geq 1$, $\sum_{h=1}^{\infty}\frac{\partial A^h_n(\theta)}{\partial\theta'}=\sum_{h=1}^{\infty}hA^{h-1}_n(\theta)\frac{\partial A_n(\theta)}{\partial\theta'}$. As (1) $\sum_{h=0}^{\infty}A^h_n(\theta)$ and $\sum_{h=1}^{\infty}hA^{h-1}_n(\theta)$ are uniformly

[9]This can be proved by mathematical induction. Step (i) For $h=2$, $\frac{\partial A^2_n(\theta)}{\partial\lambda}=A_n(\theta)\frac{\partial A_n(\theta)}{\partial\lambda}+\frac{\partial A_n(\theta)}{\partial\lambda}A_n(\theta)$.

bounded in either row sum or column sum, uniformly in a neighborhood of θ_0, (2) $S_n^{-1}(\lambda)$ is uniformly bounded in both row and column sums, also uniformly in λ in a neighborhood of λ_0 and (3) W_n is uniformly bounded in both row and column sums, we have the result that the elements of $\frac{\partial b_{2n}(\theta)}{\partial \theta'}$ will be uniformly bounded in a neighborhood of θ_0. As $\bar{\theta}_{nT}$ converges in probability to θ_0, we conclude that elements of $\frac{\partial b_{2n}(\bar{\theta}_{nT})}{\partial \theta'}$ are $O_p(1)$. ■

Using $W_n S_n^{-1}(\lambda) = S_n^{-1}(\lambda) W_n$, $\frac{\partial A_n(\theta)}{\partial \lambda} A_n(\theta) = S_n^{-1}(\lambda) W_n S_n^{-1}(\lambda)(\gamma I_n + \rho W_n) S_n^{-1}(\lambda)(\gamma I_n + \rho W_n) = S_n^{-1}(\lambda)(\gamma I_n + \rho W_n) S_n^{-1}(\lambda) W_n S_n^{-1}(\lambda)(\gamma I_n + \rho W_n) = A_n(\theta)\frac{\partial A_n(\theta)}{\partial \lambda}$. So, $A_n(\theta)\frac{\partial A_n(\theta)}{\partial \lambda} = \frac{\partial A_n(\theta)}{\partial \lambda} A_n(\theta)$ and $\frac{\partial A_n^2(\theta)}{\partial \gamma} = 2A_n(\theta)\frac{\partial A_n(\theta)}{\partial \gamma}$. Step (ii) Soppose $\frac{\partial A_n^h(\theta)}{\partial \lambda} = hA_n^{h-1}(\theta)\frac{\partial A_n(\theta)}{\partial \lambda}$, then $\frac{\partial A_n^{h+1}(\theta)}{\partial \lambda} = \frac{\partial A_n^h(\theta)\cdot A_n(\theta)}{\partial \lambda} = hA_n^{h-1}(\theta)\frac{\partial A_n(\theta)}{\partial \lambda} A_n(\theta) + A_n^h(\theta)\frac{\partial A_n(\theta)}{\partial \lambda} = hA_n^{h-1}(\theta) A_n(\theta)\frac{\partial A_n(\theta)}{\partial \lambda} + A_n^h(\theta)\frac{\partial A_n(\theta)}{\partial \lambda} = (h+1)A_n^h(\theta)\frac{\partial A_n(\theta)}{\partial \lambda}$. Same arguments can be applied to other components of $\frac{\partial A_n^h(\theta)}{\partial \theta}$.

References

[1] Andrews, D.W.K. and P. Guggenberger (2003), A Bias-Reduced Log-Periodogram Regression Estimator for the Long-Memory Parameter, *Econometrica,* Vol.71, No.2, 675-712.

[2] Anselin, L. (1988), *Spatial Econometrics: Methods and Models*, Kluwer Academic Publishers, The Netherlands.

[3] Anselin, L. (1992), Space and Applied Econometrics, Anselin (ed.), Special Issue, *Regional Science and Urban Economics 22*.

[4] Anselin, L. and A.K. Bera (1998), Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics, A. Ullah and D.E.A. Giles (eds.), *Handbook of Applied Economics Statistics*, Marcel Dekker, New York.

[5] Anselin, L. and R. Florax (1995), *New Directions in Spatial Econometrics*, Springer-Verlag, Berlin.

[6] Anselin, L. and S. Rey (1997), Spatial Econometrics, Anselin, L. and S. Rey (eds.), Special Issue, *International Regional Science Review 20.*

[7] Baltagi, B., Song, S.H. and W. Kon (2003), Testing Panel Data Regression Models with Spatial Error Correlation, *Journal of Econometrics,* Vol.117, 123-150.

[8] Bun, M. and M. Carree (2005), Bias-Corrected Estimation in Dynamic Panel Data Models, *Journal of Business & Economic Statistics*, Vol.3, No.2, 200-210(11).

[9] Cliff, A.D. and J.K. Ord (1973), *Spatial Autocorrelation*, London: Pion Ltd.

[10] Cressie, N. (1993), *Statistics for Spatial Data*, Wiley, New York.

[11] Davidson, J. (1994), *Stochastic Limit Theory*, Oxford University Press.

[12] Doreian, P. (1980), Linear Models with Spatially Distributed Data, Spatial Disturbances, or Spatial Effects, *Sociological Methods and Research 9, 29-60.*

[13] Gänsler, P. and W. Stute (1977), *Wahrscheinlichkeitstheorie*, Spring Verlag, New York.

[14] Hamilton, J. (1994), *Times Series Analysis*, Princeton University Press.

[15] Hahn, J. and G. Kuersteiner (2002), Asymptotically Unbiased Inference for a Dynamic Panel Model with Fixed Effects When Both n and T Are Large, *Econometrica,* Vol.70, No.4, 1639-1657.

[16] Hahn, J. and W. Newey (2004), Jackknife and Analytical Bias Reduction for Nonlinear Panel Models, *Econometrica*, Vol.72, No.4, 1295-1319.

[17] Haining, R. (1990), *Spatial Data Analysis in the Social and Environmental Sciences*, Cambridge University Press.

[18] Harville, D. (1997), *Matrix Algebra From A Statistician's Perspective*, Springer, New York.

[19] Hausman, J. and W. Taylor (1981), Panel Data and Unobservable Individual Effects, *Econometrica*, Vol.49, No.6, 1377-1398.

[20] Horn, R. and C. Johnson (1985), *Matrix Algebra*, Cambridge University Press.

[21] Hsiao, C. (1986), *Analysis of Panel Data*, Cambridge University Press.

[22] Kapoor, M., Kelejian, H.H. and I.R. Prucha (2004), Panel Data Models with Spatially Correlated Error Components, forthcoming in *Journal of Econometrics.*

[23] Kelejian, H.H. and I.R. Prucha (1998), A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbance, *Journal of Real Estate Finance and Economics*, Vol. 17:1, 99-121.

[24] Kelejian H.H. and I.R. Prucha (2001), On the Asymptotic Distribution of the Moran I Test Statistic With Applications, *Journal of Econometrics*, 104, 219-257.

[25] Kelejian, H.H. and D. Robinson (1993), A Suggested Method of Estimation for Spatial Interdependent Models with Autocorrelated Errors, and An Application to A Vounty Expenditure Model, Papers in *Regional Science* 72, 297-312.

[26] Lee, L.F. (2001), GMM and 2SLS Estimation of Mixed Regressive, Spatial Autoregressive Models, forthcoming in *Journal of Econometrics.*

[27] Lee, L.F. (2002), Consistency and Efficiency of Least Square Estimation for Mixed Regressive, Spatial Autoregressive Models, *Econometric Theory*, 18, 252-277.

[28] Lee, L.F. (2003), Best Spatial Two-Stage Least Squares Estimator for a Spatial Autoregresive Model with Autoregressive Disturbances, *Econometric Reviews*, Vol.22, No.4, 307-335.

[29] Lee, L.F. (2004), Asymptotic Distributions of Quasi-Maximum Likelihood Estimators for Spatial Econometric Models, *Econometrica*, Vol. 72, No.6, 1899-1925.

[30] Neyman, J. and E. Scott (1948), Consistent Estimates Based on Partially Consistent Observations, *Econometrica,* Vol.16, No.1, 1-32.

[31] Ord, J.K. (1975), Estimation Methods for Models of Spatial Interaction, *Journal of the American Statistical Association* 70, 120-297.

[32] Phillips, P.C.B. and H.R. Moon (1999), Linear Regression Limit Theory for Nonstationary Panel Data, *Econometrica*, Vol.67, No.5, 1057-1111.

[33] Pötscher, B.M. and I.R. Prucha (1997), *Dynamic Nonlinear Econometric Models, Asymptotic Theory*, Springer, New York.

[34] Paelinck, J. and L. Klaassen (1979), *Spatial Econometrics*, Saxon House, Farnborough.

[35] Rothenberg, T.J. (1971), Identification in Parametric Models, *Econometrica*, Vol. 39, No.3, 577-591.

[36] White, H. (1994), *Estimation, Inference and Specification Analysis*, Cambridge University Press.