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

Jihai Yu, Robert de Jong, Lung-fei Lee†
Department of Economics
The Ohio State University

August 25, 2007

Abstract

Yu, de Jong and Lee (2006) established 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. This paper covers a nonstationary case where there are unit roots in the data generating process. When not all the roots in the DGP are unit, the estimators' rates of convergence will be the same as the stationary case, and the estimators can be asymptotically normal. The presence of the nonstationary components however will make the estimators' asymptotic variance matrix singular. Consequently, a linear combination of the spatial and dynamic effects can converge at a higher rate. We also propose a bias correction for our estimator. 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, Quasi-maximum likelihood estimation, Bias correction, Unit root, Nonstationarity

*We would like to thank participants of the Econometrics Seminar at The Ohio State University (March 2007) and the Third Symposium on Econometric Theory and Applications at HKUST (April 2007) for helpful comments.

†Lee acknowledges financial support for his research from NSF under Grant No. SES-0519204.

1 Introduction

This paper investigates the properties of maximum likelihood (ML) estimators and quasi-maximum likelihood (QML) estimators for spatial dynamic panel data models with individual fixed effects when both the number of individuals n and the number of time periods T are large for a nonstationary case.

In Yu, de Jong and Lee (2006), the consistency and asymptotic distribution of the QML estimators are established for the stationary case. Also, a bias correction procedure for the estimators is proposed. It is shown that as long as T grows faster than $n^{1/3}$, the correction will asymptotically eliminate the bias and will yield a centered confidence interval. When there are unit roots in the process so that the assumption of absolute summability in Yu, de Jong and Lee (2006) does not hold, the analysis of the properties of estimators for the stationary case, which is crucially based on the absolute summability condition, will not be valid. In this paper, we will show that when the spatial weights matrix is row normalized from a symmetric matrix, we can still obtain the consistency and asymptotic normality of the ML and QML estimators with the same rate of convergence as in the stationary case. The difference is that the variance matrix is different from the stationary case, and it is singular in the limit. Also, for this nonstationary case, there is a linear combination of common parameters that will have a higher rate of convergence.

The nonstationary case we consider is relevant in empirical applications. In Yu (2006), a spatial dynamic panel data model is applied to study the growth convergence of 48 contiguous states. In the estimation result, the spatial effects are significant and the sum of estimators of spatial and dynamic effects equals nearly to the 1. This implies that there may be nonstationary components in the DGP (see discussion in Section 2.1 for details), which motivates deriving asymptotic theory for the estimators under nonstationarity. Also, in Tao's (2006) study on the education spending of local school districts using spatial dynamic panel model, we have significant spatial effects and the sum of estimators of spatial and dynamic effects equals nearly to 1.

There is growing research interest in nonstationary panels in recent years. For independent panels, we have Maddala and Wu (1999), Levin, Lin and Chu (2002), Im, Pesaran and Shin (2003), etc. For cross-sectionally correlated panels, we have Pesaran (2003), Phillips and Sul (2003), Moon and Perron (2004), etc, where the cross sectional dependence is specified by common factors. This paper covers a case of nonstationary panel data where the cross sectional dependence is specified by spatial correlation among units directly. There are already extensive empirical applications for nonstationary panel data[1]. We expect that our model can shed light on existing nonstationary panel data models and empirical applications.

This paper is organized as follows. In Section 2, the model is introduced. We then explain our method of estimation, which is a concentrated QML estimation. Several lemmas on matrix algebra and a central limit theorem are stated. Section 3 derives the consistency and asymptotic distribution of the spatial effect parameter. Using the results of Section 3, we establish the asymptotic distribution of the common parameters

[1] The applications include purchasing power parity, growth and convergence, money demand, monetary exchange rate model, inflation-rate convergence, interest rate, health care expenditure, hysteresis in unemployment, etc. See Choi (2004) for more details.

in Section 4. Also, a bias correction procedure is proposed and simulation results are reported. Section 5 concludes the paper. Some useful lemmas and proofs are collected in the Appendix.

2 The Model and The 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 a known $n \times n$ spatial weights matrix which is nonstochastic 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 individual 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 $k_x + 4$. W_n is usually row normalized from a symmetric matrix such that its ith row is $[c_{n,i1}, c_{n,i2}, \cdots, c_{n,in}]/\sum_{j=1}^{n} c_{n,ij}$, where $c_{n,ij}$ represents a function of the spatial distance of different units in some space. As a normalization, $c_{n,ii} = 0$. It is a common practice in empirical work that W_n is row normalized, which ensures that all the weights are between 0 and 1 and weighting operations can be interpreted as an average of the neighboring values. Also, a weights matrix row normalized from a symmetric matrix has real eigenvalues, with all its eigenvalues less than or equal to one in absolute value and its largest eigenvalue always 1 (see Ord (1975)). Such a spatial weights matrix is also diagonalizable (see Proposition B.1 in Appendix B).

Define $S_n(\lambda) = I_n - \lambda W_n$ and denote $S_n \equiv S_n(\lambda_0) = I_n - \lambda_0 W_n$. 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

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

A nonstationary case occurs if some eigenvalues d_{ni} of A_n are equal to 1, i.e., $d_{ni} = \frac{\gamma_0 + \rho_0 \varpi_{ni}}{1 - \lambda_0 \varpi_{ni}} = 1$ for some i where ϖ_{ni} is an eigenvalue of W_n. For the nonstationary case, we can decompose Y_{nt} into a stationary part and a nonstationary part. To do that, we can first diagonalize[2] A_n as $A_n = R_n D_n R_n^{-1}$ where R_n is the eigenvectors of A_n and $D_n = diag(d_{n1}, d_{n2}, \cdots, d_{nn})$ where d_{ni}'s are eigenvalues of A_n. When $d_{n,\max} = 1$ and $d_{n,\min} > -1$ where $d_{n,\max}$ and $d_{n,\min}$ are respectively the largest and smallest eigenvalues of A_n, without loss of generality, suppose that $d_{ni} = 1$ for $i = 1, 2, \cdots, m_n$ and $|d_{ni}| < 1$ for $m_n + 1 \leq i \leq n$ where m_n is the number of unit roots. Let $B_n = R_n \tilde{D}_n R_n^{-1}$ with $\tilde{D}_n = Diag(0, \cdots, 0, d_{n,m_n+1}, \cdots, d_{nn})$ so that $D_n = J_n + \tilde{D}_n$ where $J_n = Diag\{\mathbf{1}'_{m_n}, 0, \cdots, 0\}$ with $\mathbf{1}_{m_n}$ being an $m_n \times 1$ vector of ones. As J_n is idempotent

[2] See Proposition B.2 in Appendix B for diagonalizability of A_n.

and $J_n \cdot \tilde{D}_n = \mathbf{0}$, $A_n^h = M_n + B_n^h$ for any $h = 1, 2, \cdots$ where $M_n = R_n J_n R_n^{-1}$. Then (see Proposition B.5 in Appendix B), for $t \geq 0$, we can decompose Y_{nt} into sum of a stationary part and a nonstationary part:

$$Y_{nt} = Y_{nt}^u + Y_{nt}^s, \tag{2.3}$$

where

$$Y_{nt}^u = M_n \left(Y_{n,-1} + \mathbf{c}_{n0} \frac{t}{(1-\lambda_0)} + \frac{\sum_{h=0}^{t-1} X_{nh}\beta_0}{(1-\lambda_0)} + \frac{\sum_{h=0}^{t-1} V_{nh}}{(1-\lambda_0)} \right), \tag{2.4}$$

$$Y_{nt}^s = \sum_{h=0}^{\infty} B_n^h S_n^{-1} \mathbf{c}_{n0} + \sum_{h=0}^{\infty} B_n^h S_n^{-1} X_{n,t-h}\beta_0 + \sum_{h=0}^{\infty} B_n^h S_n^{-1} V_{n,t-h}. \tag{2.5}$$

Compared to the stationary case, the model has a time trend attachment $M_n \mathbf{c}_{n0} \frac{t}{(1-\lambda_0)} + M_n \frac{\sum_{h=0}^{t-1} X_{nh}\beta_0}{(1-\lambda_0)}$, a random walk attachment $M_n \frac{\sum_{h=0}^{t-1} V_{nh}}{(1-\lambda_0)}$ and a nonstationary initial value component $M_n Y_{n,-1}$.

Using (2.3), (2.4) and (2.5), we have[3]

$$\tilde{Y}_{nt} = \tilde{Y}_{nt}^u + \tilde{Y}_{nt}^s, \quad t = 0, 1, \cdots, T, \tag{2.6}$$

where

$$\tilde{Y}_{nt}^u = \frac{1}{(1-\lambda_0)} M_n \left(\mathbf{c}_{n0}\tilde{t} + \tilde{\mathbb{X}}_{nt}\beta_0 + \tilde{\xi}_{nt} \right), \; \tilde{Y}_{nt}^s = \tilde{\mathcal{X}}_{nt}^s \beta_0 + \tilde{U}_{nt}^s, \tag{2.7}$$

with $\tilde{t} = t - \frac{T+1}{2}$, $\mathbb{X}_{nt} = \sum_{h=0}^{t-1} X_{nh}$, $\xi_{nt} = \sum_{h=0}^{t-1} V_{nh}$, $\mathcal{X}_{nt}^s = \sum_{h=0}^{\infty} B_n^h S_n^{-1} X_{n,t-h}$ and $U_{nt}^s = \sum_{h=0}^{\infty} B_n^h S_n^{-1} V_{n,t-h}$.

To analyze the model, the following assumptions are needed.

Assumption 1. W_n is a nonstochastic spatial weights matrix, row normalized from a symmetric weights matrix.

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. n is a nondecreasing function of T.

Assumption 4. The elements of X_{nt} and $\mathbf{c}_{n0}$ are nonstochastic and bounded, uniformly in n and t, and $\lim_{T\to\infty} \frac{1}{nT} \sum_{t=1}^{T} \tilde{X}_{nt}' \tilde{X}_{nt}$ exists and is nonsingular. Also, $\lim_{T\to\infty} \frac{1}{nT^3} \sum_{t=1}^{T} (\mathbf{c}_{n0}\tilde{t} + \tilde{\mathbb{X}}_{nt}\beta_0)' M_n' M_n (\mathbf{c}_{n0}\tilde{t} + \tilde{\mathbb{X}}_{nt}\beta_0) \neq 0$.

Assumption 5. $S_n(\lambda)$ is invertible for all $\lambda \in \Lambda$. Furthermore, Λ is compact[4] and the true parameter λ_0 with $|\lambda_0| < 1$ is in the interior of Λ.

Assumption 6. $\rho_0 + \gamma_0 + \lambda_0 = 1$ with $\gamma_0 \neq 1$. Also, $d_{n,\max} = 1$ and $d_{n,\min} > -1$, where $d_{n,\max}$ and $d_{n,\min}$ are the largest and smallest eigenvalues of A_n.

Assumption 7. The row and column sums of W_n and $S_n^{-1}(\lambda)$ are bounded uniformly[5] in n, also uniformly in $\lambda \in \Lambda$ for $S_n^{-1}(\lambda)$.

---

[3] For notational purpose, we define for any $n \times 1$ vector at period t, Υ_{nt}, we have $\tilde{\Upsilon}_{nt} = \Upsilon_{nt} - \bar{\Upsilon}_{nT}$ and $\tilde{\tilde{\Upsilon}}_{n,t-1} = \Upsilon_{n,t-1} - \bar{\Upsilon}_{nT,-1}$ for $t = 1, 2, \cdots, T$ where $\bar{\Upsilon}_{nT} = \frac{1}{T} \sum_{t=1}^{T} \Upsilon_{nt}$ and $\bar{\Upsilon}_{nT,-1} = \frac{1}{T} \sum_{t=1}^{T} \Upsilon_{n,t-1}$.

[4] Note that in the literature, Λ is typically assumed to be a compact subset of $(-1, 1)$.

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

Assumption 8. The row and column sums of $\sum_{h=1}^{\infty} abs(B_n^h)$ are bounded uniformly in n, where $[abs(B_n)]_{ij} = |B_{n,ij}|$.

Assumptions 1 and 2 provide essential features of the weights matrix and disturbances of the model. Assumption 3 allows two cases: (i) $n \to \infty$ as $T \to \infty$; (ii) n is fixed as $T \to \infty$. For case (i), we say that $n, T \to \infty$ simultaneously. When exogenous variables X_{nt} are included in the model, it is convenient to assume that the exogenous regressors are uniformly bounded, as is done in Assumption 4. Also, we make the assumption that either $\mathbf{c}_{n0}$ or X_{nt} is relevant in the model. A simple consequence of Assumption 5 is that, for the system (2.1), Y_{nt} can be solved in terms of $\mathbf{c}_{n0}$, X_{nt} and V_{nt}. Assumption 6 specifies that some roots of A_n are equal to 1, while the other roots are less than 1 in absolute value. The first part of this assumption rules out explicitly the pure unit root time series case without spatial interaction; more generally, it rules out the case where $\gamma_0 = 1$ and $\rho_0 + \lambda_0 = 0$. A sufficient condition for Assumption 6 is $\rho_0 < 1$ with $|\gamma_0| < 1$ and $|\lambda_0| < 1$ under $\rho_0 + \gamma_0 + \lambda_0 = 1$ (see Proposition B.3 in Appendix B). Assumption 7 is originated by Kelejian and Prucha (1998, 2001). The uniform boundedness of W_n and $S_n^{-1}(\lambda)$ is a condition to limit the spatial correlation to a manageable degree. Assumption 8 is the absolute summability condition and the row and column sum boundedness condition, which will play an important role to derive asymptotic properties of QML estimators. This assumption is essential for the model because it limits the dependence between time series and between cross sectional units for the stationary component Y_{nt}^s in the process. In order to justify the absolute summability of B_n in (2.5) and Assumption 8, a sufficient condition is $\|B_n\| < 1$ for any matrix norm (see Horn and Johnson (1985), Corollary 5.6.16) that satisfies $\|B_n\| = \|abs\,(B_n)\|$. When $\|B_n\| < 1$, $\sum_{h=0}^{\infty} B_n^h$ exists and can be defined as $(I_n - B_n)^{-1}$ (see Appendix B.1 for an example where A_n has some eigenvalues equalling to one but others strictly less than one in absolute value).

2.2 Concentrated Likelihood Function

Denote $Z_{nt} = (Y_{n,t-1},\ W_n Y_{n,t-1},\ X_{nt})$ and $\theta = (\delta', \lambda, \sigma^2)'$ where $\delta = (\gamma, \rho, \beta')'$. The log likelihood function of (2.1) is

$$\ln L_{n,T}(\theta, \mathbf{c}_n) = -\frac{nT}{2} \ln 2\pi - \frac{nT}{2} \ln \sigma^2 + T \ln |S_n(\lambda)| - \frac{1}{2\sigma^2} \sum_{t=1}^{T} V_{nt}'(\zeta) V_{nt}(\zeta), \tag{2.8}$$

where $V_{nt}(\zeta) = S_n(\lambda) Y_{nt} - Z_{nt}\delta - \mathbf{c}_n$ and $\zeta = (\delta', \lambda, \mathbf{c}_n')'$. The QML estimators $\hat{\theta}_{nT}$ and $\hat{\mathbf{c}}_{nT}$ are the extreme estimators derived from the maximization of (2.8). 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 use the concentrating approach. We will concentrate $\mathbf{c}_n$ and δ out and focus asymptotic analysis on the estimator of λ_0 via the concentrated likelihood function[6]. For the concentrated likelihood function, the dimension of parameter space does not change as n and/or T increase.

[6] The reason to concentrate out δ is to avoid technical complication in the consistency proof and deriving the asymptotic distribution jointly for the common parameters. See footnote 15 for details.

From (2.8), using the first order conditions, we can get the concentrated estimators given λ:

$$\begin{aligned}
\hat{\delta}_{nT}(\lambda) &= [\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}'_{nt}\tilde{Z}_{nt}]^{-1}[\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}'_{nt}S_n(\lambda)\tilde{Y}_{nt}],\ \hat{\mathbf{c}}_{nT}(\lambda)=\frac{1}{T}\sum_{t=1}^{T}(S_n(\lambda)Y_{nt}-Z_{nt}\hat{\delta}_{nT}(\lambda)),\\
\hat{\sigma}^2_{nT}(\lambda) &= \frac{1}{nT}\sum_{t=1}^{T}(S_n(\lambda)\tilde{Y}_{nt}-\tilde{Z}_{nt}\hat{\delta}_{nT}(\lambda))'(S_n(\lambda)\tilde{Y}_{nt}-\tilde{Z}_{nt}\hat{\delta}_{nT}(\lambda)), \qquad (2.9)
\end{aligned}$$

and the concentrated likelihood is

$$\ln L_{n,T}(\lambda) = -\frac{nT}{2}(\ln 2\pi + 1) - \frac{nT}{2}\ln\hat{\sigma}^2_{nT}(\lambda) + T\ln|S_n(\lambda)|. \qquad (2.10)$$

The QML estimator $\hat{\lambda}_{nT}$ maximizes the concentrated likelihood function (2.10), and the QML estimators of δ_0, σ_0^2 and $\mathbf{c}_{n0}$ are $\hat{\delta}_{nT}(\hat{\lambda}_{nT})$, $\hat{\sigma}^2_{nT}(\hat{\lambda}_{nT})$ and $\hat{\mathbf{c}}_n(\hat{\lambda}_{nT})$.

Also, the reduced form of (2.1) can be represented as

$$\begin{aligned}
Y_{nt} &= S_n^{-1}\left(Z_{nt}\delta_0 + \mathbf{c}_{n0} + V_{nt}\right) \qquad (2.11)\\
&= Z_{nt}\delta_0 + \lambda_0 G_n Z_{nt}\delta_0 + S_n^{-1}(\mathbf{c}_{n0}+V_{nt}),\ t=0,1,...,T,
\end{aligned}$$

because $I_n + \lambda_0 G_n = S_n^{-1}$ where $G_n = W_n S_n^{-1}$. Denote $\mathcal{H}_{nT} = \frac{1}{nT}\sum_{t=1}^{T}(\tilde{Z}_{nt}, G_n\tilde{Z}_{nt}\delta_0)'(\tilde{Z}_{nt}, G_n\tilde{Z}_{nt}\delta_0) = \begin{pmatrix} \mathcal{H}_{1,nT} & \mathcal{H}_{2,nT} \\ \mathcal{H}'_{2,nT} & \mathcal{H}_{3,nT} \end{pmatrix}$ where $\mathcal{H}_{1,nT} = \frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}'_{nt}\tilde{Z}_{nt}$, $\mathcal{H}_{2,nT} = \frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}'_{nt}G_n\tilde{Z}_{nt}\delta_0$ and $\mathcal{H}_{3,nT} = \frac{1}{nT}\sum_{t=1}^{T}\delta_0'\tilde{Z}'_{nt}G'_nG_n\tilde{Z}_{nt}\delta_0$. Hence, $\mathcal{H}_{nT}$ is the covariance matrix of the explanatory variables of the reduced form (2.11) after taking difference from time average, which is crucial for our asymptotic analysis of QML estimators because $\hat{\sigma}^2_{nT}(\lambda)$ in (2.10) involves $\mathcal{H}_{i,nT}$ terms for $i=1,2,3$. To study $\mathcal{H}_{nT}$, it is desirable to decompose $\tilde{Z}_{nt}$ into a stationary part and a nonstationary part such that $\tilde{Z}_{nt} = \tilde{Z}^u_{nt} + \tilde{Z}^s_{nt}$ where

$$\tilde{Z}^u_{nt} = (\tilde{\tilde{Y}}^u_{n,t-1}, W_n\tilde{\tilde{Y}}^u_{n,t-1}, \mathbf{0}_{n\times kx}),\ \tilde{Z}^s_{nt} = (\tilde{\tilde{Y}}^s_{n,t-1}, W_n\tilde{\tilde{Y}}^s_{n,t-1}, \tilde{X}_{nt}). \qquad (2.12)$$

As (see Proposition B.4 in Appendix B) $\tilde{\tilde{Y}}^u_{n,t-1} = W_n\tilde{\tilde{Y}}^u_{n,t-1} = G_n\tilde{Z}^u_{nt}\delta_0$, we have $\tilde{Z}^u_{nt} = \tilde{\tilde{Y}}^u_{n,t-1}\cdot c'$ and $(\tilde{Z}^u_{nt}, G_n\tilde{Z}^u_{nt}\delta_0) = \tilde{\tilde{Y}}^u_{n,t-1}\cdot c^{*\prime}$ where $c = (1,1,\mathbf{0}_{1\times k_x})'$ and $c^* = (c',1)'$. Hence, denoting $\mathcal{H}^s_{nT} = \frac{1}{nT}\sum_{t=1}^{T}(\tilde{Z}^s_{nt}, G_n\tilde{Z}^s_{nt}\delta_0)'(\tilde{Z}^s_{nt}, G_n\tilde{Z}^s_{nt}\delta_0)$, we can express $\mathcal{H}_{nT}$ in terms of vectors such that

$$\mathcal{H}_{nT} \equiv \omega_{nT}\left(T^2\cdot c^*c^{*\prime} + T\cdot d_{nT}\cdot c^{*\prime} + T\cdot c^*\cdot d'_{nT} + \mathcal{H}^s_{nT}/\omega_{nT}\right), \qquad (2.13)$$

where $\omega_{nT} = \frac{1}{nT^3}\sum_{t=1}^{T}\tilde{\tilde{Y}}^{u\prime}_{n,t-1}\tilde{\tilde{Y}}^u_{n,t-1}$, $d_{nT} = \frac{1}{\omega_{nT}}(\frac{1}{nT^2}\sum_{t=1}^{T}(\tilde{Z}^s_{nt}, G_n\tilde{Z}^s_{nt}\delta_0)'\tilde{\tilde{Y}}^u_{n,t-1})'$. Similarly, we can express $\mathcal{H}_{i,nT}$ in terms of vectors such that

$$\begin{aligned}
\mathcal{H}_{1,nT} &= \omega_{nT}\left(T^2\cdot cc' + T\cdot d_{1,nT}\cdot c' + T\cdot c\cdot d'_{1,nT} + \mathcal{H}^s_{1,nT}/\omega_{nT}\right), \qquad (2.14a)\\
\mathcal{H}_{2,nT} &= \omega_{nT}\left(T^2\cdot c + T\cdot d_{1,nT} + T\cdot d_{2,nT}\cdot c + \mathcal{H}^s_{2,nT}/\omega_{nT}\right), \qquad (2.14b)\\
\mathcal{H}_{3,nT} &= \omega_{nT}\left(T^2 + 2T\cdot d_{2,nT} + \mathcal{H}^s_{3,nT}/\omega_{nT}\right), \qquad (2.14c)
\end{aligned}$$

where $d_{1,nT} = \frac{1}{\omega_{nT}}\frac{1}{nT^2}\sum_{t=1}^{T}\tilde{Z}^{s\prime}_{nt}\tilde{\tilde{Y}}^u_{n,t-1}$, $d_{2,nT} = \frac{1}{\omega_{nT}}\frac{1}{nT^2}\sum_{t=1}^{T}(G_n\tilde{Z}^s_{nt}\delta_0)'\tilde{\tilde{Y}}^u_{n,t-1}$. We notice that elements of $\mathcal{H}_{nT}$ are of the order $O(T^2)$ and $T^{-2}\mathcal{H}_{nT}$ is singular in the limit. However, because of the pattern of the nonstationary component, $\mathcal{H}^{-1}_{nT}$ exists and $\mathcal{H}^{-1}_{nT}\cdot c^*$ has a lower order of $O(T^{-1})$ from Proposition 2.1 below.

2.3 Two Technical Propositions

To study the asymptotic behavior of $\mathcal{H}_{nT}^{-1}$, we need the following proposition about matrix algebra.

Proposition 2.1 *Let $K_T = T^2 c_T c_T' + T(c_T d_T' + d_T c_T') + A_T$, where c_T, d_T are m-dimensional column random vectors, $plim_{T\to\infty} c_T \neq \mathbf{0}$ and is nonstochastic, A_T is positive definite for large enough T with probability one, $plim_{T\to\infty} A_T$ exists and is an $m\times m$ positive definite matrix. Denote $\Delta_T = 1 - \left(d_T' A_T^{-1} d_T - \frac{(d_T' A_T^{-1} c_T)^2}{c_T' A_T^{-1} c_T}\right)$. Under the assumption that $plim_{T\to\infty}\Delta_T \neq 0$, the sequence $\{K_T\}$ has the following properties:*

(a) the elements of K_T^{-1} are $O_p(1)$;

(b) the elements of $K_T^{-1} c_T$ are $O_p(T^{-1})$;

(c) $T^2 c_T' K_T^{-1} c_T = 1 + O_p(T^{-1})$.

Proof. *See the proof for Proposition B.13 in Appendix B.3.* ■

In our application, we can apply K_T to $\mathcal{H}_{nT}$ in (2.13) and $\mathcal{H}_{1,nT}$ in (2.14). To apply Proposition 2.1, we need an additional assumption.

Assumption 9. $\mathcal{H}_{nT}^s$ is nonsingular for large enough T with probability one, $\text{plim}_{T\to\infty}\mathcal{H}_{nT}^s$ exists and is nonsingular.

As $\mathcal{H}_{nT}^s = \frac{1}{nT}\sum_{t=1}^{T}(\tilde{Z}_{nt}^s, G_n\tilde{Z}_{nt}^s\delta_0)'(\tilde{Z}_{nt}^s, G_n\tilde{Z}_{nt}^s\delta_0)$ is always positive semidefinite, with Assumption 9, $\mathcal{H}_{nT}^s$ is positive definite for large enough T and $\text{plim}_{T\to\infty}\mathcal{H}_{nT}^s$ will also be positive definite.

In this paper, we need a central limit theorem for linear and quadratic forms of V_{nt}. Denote $Q_{nT} = Q_{nT}^s + Q_{nT}^u$ where

$$Q_{nT}^s = \sum_{t=1}^{T}\left(\mathbb{U}_{n,t-1}' V_{nt} + D_{nt}' V_{nt} + V_{nt}'\mathcal{B}_n V_{nt} - \sigma_0^2 tr\mathcal{B}_n\right), \tag{2.15a}$$

$$Q_{nT}^u = \frac{k_T}{T}\sum_{t=1}^{T}\left(M_n\left(\mathbf{c}_{n0}\tilde{t}_{-1} + \tilde{\mathbb{X}}_{n,t-1}\beta_0 + \xi_{n,t-1}\right)\right)'\cdot V_{nt}. \tag{2.15b}$$

Here, $\mathbb{U}_{nt} = \sum_{h=1}^{\infty} P_{nt,h}V_{n,t+1-h}$ where $\{P_{nt,h}\}_{h=1}^{\infty}$ is a sequence of $n\times n$ nonstochastic square matrices, D_{nt} is $n\times 1$ vector, which is nonstochastic and bounded, uniformly in n and t, $\mathcal{B}_n$ is a nonstochastic $n\times n$ symmetric matrix[7] and its row and column sums are bounded uniformly in n and k_T is $O(1)$. Denote the mean and variance of Q_{nT} as $\mu_{Q_{nT}}$ and $\sigma_{Q_{nT}}^2$ respectively with $\mu_{Q_{nT}} = 0$, we have the following proposition.

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. The row and column sums of $\sum_{h=1}^{\infty} abs(P_{nt,h})$ are bounded uniformly in n and t.

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.

Proposition 2.2 *Assume that row and column sums of $\mathcal{B}_n$ are bounded uniformly in n and assume the sequence $\frac{1}{nT}\sigma_{Q_{nT}}^2$ is bounded away from zero. Then under Assumptions A1, A2, A3 and A4, $\frac{Q_{nT}}{\sigma_{Q_{nT}}} \xrightarrow{d} N(0,1)$.*

Proof. *See Appendix B.4.* ■

[7] The assumption that $\mathcal{B}_n$ is symmetric is maintained w.l.o.g. since $V_{nt}'\mathcal{B}_n V_{nt} = V_{nt}'[(\mathcal{B}_n + \mathcal{B}_n')/2]V_{nt}$.

3 Consistency and Asymptotic Distribution of $\hat{\lambda}_{nT}$

We have the Taylor expansion $\sqrt{nT}(\lambda-\lambda_0) = \left(-\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\bar{\lambda})}{\partial\lambda^2}\right)^{-1}\left(\frac{1}{\sqrt{nT}}\frac{\partial \ln L_{n,T}(\lambda_0)}{\partial\lambda}\right)$ where $\bar{\lambda}$ lies between λ and λ_0. From concentrated likelihood function (2.10):

$$
\begin{aligned}
\frac{1}{nT}\frac{\partial \ln L_{n,T}(\lambda)}{\partial\lambda} &= -\frac{1}{2\hat{\sigma}^2_{nT}(\lambda)}\frac{\partial\hat{\sigma}^2_{nT}(\lambda)}{\partial\lambda} - \frac{1}{n}trG_n(\lambda), && (3.16a)\\
\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\lambda)}{\partial\lambda^2} &= -\frac{1}{2\hat{\sigma}^4_{n,T}(\lambda)}\left[\frac{\partial^2\hat{\sigma}^2_{nT}(\lambda)}{\partial\lambda^2}\hat{\sigma}^2_{nT}(\lambda) - (\frac{\partial\hat{\sigma}^2_{nT}(\lambda)}{\partial\lambda})^2\right] - \frac{1}{n}tr(G_n^2(\lambda)). && (3.16b)
\end{aligned}
$$

The $\hat{\sigma}^2_{nT}(\lambda)$, $\frac{\partial\hat{\sigma}^2_{nT}(\lambda)}{\partial\lambda}$ and $\frac{\partial^2\hat{\sigma}^2_{nT}(\lambda)}{\partial\lambda^2}$ have the explicit forms (see (B.49), (B.50) and (B.51)) implied by (2.9). Using Proposition B.14, we have (derived in Appendix B.5)

$$
\begin{aligned}
\hat{\sigma}^2_{nT}(\lambda) &= \sigma_0^2 + |\lambda-\lambda_0|\cdot O_p(1) + O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right), && (3.17a)\\
\frac{\partial\hat{\sigma}^2_{nT}(\lambda)}{\partial\lambda} &= -\frac{2}{n}\sigma_0^2 trG_n + |\lambda-\lambda_0|\cdot O_p(1) + O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right), && (3.17b)\\
\frac{\partial^2\hat{\sigma}^2_{nT}(\lambda)}{\partial\lambda^2} &= 2(\mathcal{H}_{3,nT} - \mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT}) + 2\sigma_0^2\frac{1}{n}trG'_nG_n + O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right), && (3.17c)\\
\sqrt{nT}\frac{\partial\hat{\sigma}^2_{nT}(\lambda_0)}{\partial\lambda} &= -\frac{2}{\sqrt{nT}}\sum_{t=1}^T \tilde{V}'_{nt}G'_n\tilde{V}_{nt} - \frac{2}{\sqrt{nT}}\sum_{t=1}^T(\delta'_0\tilde{Z}'_{nt}G'_n - \mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\tilde{Z}'_{nt})\tilde{V}_{nt} && (3.17d)\\
&\quad + O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T},\sqrt{\frac{n}{T^3}}\right)\right),
\end{aligned}
$$

where the $O_p(1)$, $O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right)$ and $O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T},\sqrt{\frac{n}{T^3}}\right)\right)$ are uniform in λ. (3.16) through (3.17) will be used to derive the consistency and asymptotic distribution of the spatial effect parameter λ.

3.1 Consistency of $\hat{\lambda}_{nT}$

For the log likelihood function (2.10) divided by the sample size nT, we have corresponding $Q_{n,T}(\lambda) = \max_{\delta,\mathbf{c}_n,\sigma^2} E\frac{1}{nT}\ln L_{n,T}(\theta)$ and the optimal solution to the problem is (equation: concentrated estimators expect)

$$
\begin{aligned}
\delta^*_{nT}(\lambda) &= [E\frac{1}{nT}\sum_{t=1}^T\tilde{Z}'_{nt}\tilde{Z}_{nt}]^{-1}[E\frac{1}{nT}\sum_{t=1}^T\tilde{Z}'_{nt}S_n(\lambda)\tilde{Y}_{nt}],\ \mathbf{c}^*_{nT}(\lambda) = E\frac{1}{T}\sum_{t=1}^T(S_n(\lambda)Y_{nt} - Z_{nt}\delta^*_{nT}(\lambda)),\\
\sigma^{*2}_{nT}(\lambda) &= E\frac{1}{nT}\sum_{t=1}^T(S_n(\lambda)\tilde{Y}_{nt} - \tilde{Z}_{nt}\delta^*_{nT}(\lambda))'(S_n(\lambda)\tilde{Y}_{nt} - \tilde{Z}_{nt}\delta^*_{nT}(\lambda)). && (3.18)
\end{aligned}
$$

Hence,

$$
Q_{n,T}(\lambda) = -\frac{1}{2}(\ln 2\pi + 1) - \frac{1}{2}\ln\sigma^{*2}_{nT}(\lambda) + \frac{1}{n}\ln|S_n(\lambda)|. \qquad (3.19)
$$

Claim 3.1 *Under Assumptions 1-9, $\frac{1}{nT}\ln L_{n,T}(\lambda) - Q_{n,T}(\lambda) \xrightarrow{p} 0$ uniformly in λ in any compact parameter space Λ and $Q_{n,T}(\lambda)$ is uniformly equicontinuous for $\lambda\in\Lambda$.*

Proof. *See Appendix C.1.* ■

From (3.19), we have

$$\frac{\partial^2 Q_{n,T}(\lambda)}{\partial\lambda^2} = -\frac{1}{2\sigma^{*4}_{n,T}(\lambda)}\left[\frac{\partial^2\sigma^{*2}_{nT}(\lambda)}{\partial\lambda^2}\sigma^{*2}_{nT}(\lambda) - (\frac{\partial\sigma^{*2}_{nT}(\lambda)}{\partial\lambda})^2\right] - \frac{1}{n}tr(G_n^2(\lambda)). \tag{3.20}$$

Using (B.60) about $\sigma^{*2}_{nT}(\lambda)$, $\frac{\partial\sigma^{*2}_{nT}(\lambda)}{\partial\lambda}$ and $\frac{\partial^2\sigma^{*2}_{nT}(\lambda)}{\partial\lambda^2}$, we have

$$\begin{aligned}\partial^2 Q_{n,T}(\lambda_0)/\partial\lambda^2 &= -\frac{1}{\sigma_0^2}(E\mathcal{H}_{3,nT} - E\mathcal{H}'_{2,nT}(E\mathcal{H}_{1,nT})^{-1}E\mathcal{H}_{2,nT}) \\ &\quad -\frac{1}{n}\left(trG'_nG_n + trG_n^2 - \frac{2(trG_n)^2}{n}\right) + O\left(\frac{1}{T}\right),\end{aligned} \tag{3.21}$$

and its limit will be negative if $\lim_{T\to\infty}\left(\mathcal{H}_{3,nT} - \mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT}\right) \neq 0$ or $\lim_{n\to\infty}\frac{1}{n}tr(\mathcal{C}_n+\mathcal{C}'_n)(\mathcal{C}_n+\mathcal{C}'_n)' \neq 0$ where $\mathcal{C}_n = G_n - \frac{trG_n}{n}I_n$ (see Appendix C.2). Claim 3.1 is the uniform convergence condition, combined with identification, we can get the consistency of QML estimators.

Theorem 3.2 *Under Assumptions 1-9, λ_0 is globally identified and $\hat{\lambda}_{nT}$ is consistent.*

Proof. *See Appendix C.3.* ■

3.2 Distribution of $\hat{\lambda}_{nT}$

Plugging (3.17) into $\frac{\partial \ln L_{n,T}(\lambda)}{\partial\lambda}$ in (3.16a), we have

$$\begin{aligned}&\frac{1}{\sqrt{nT}}\frac{\partial \ln L_{n,T}(\lambda_0)}{\partial\lambda} \\ =\ &\frac{1}{\sigma^2_{nT}(\lambda_0)}\left(\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}\tilde{V}'_{nt}(G'_n - \frac{1}{n}trG_n\cdot I_n)\tilde{V}_{nt} + \frac{1}{\sqrt{nT}}\sum_{t=1}^{T}(\delta'_0\tilde{Z}'_{nt}G'_n - \mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\tilde{Z}'_{nt})\tilde{V}_{nt}\right) \\ &+O_p\left(\max\left(\frac{1}{\sqrt{nT}}, \frac{1}{T}, \sqrt{\frac{n}{T^3}}\right)\right).\end{aligned} \tag{3.22}$$

As $\tilde{Z}_{nt}$ has stationary and nonstationary parts (see (2.12)), we can decompose $\frac{1}{\sqrt{nT}}\frac{\partial \ln L_{n,T}(\lambda_0)}{\partial\lambda}$ into two parts accordingly such that $\frac{1}{\sqrt{nT}}\frac{\partial \ln L_{n,T}(\lambda_0)}{\partial\lambda} = \frac{1}{\sqrt{nT}}\frac{\partial \ln L^s_{n,T}(\lambda_0)}{\partial\lambda} + \frac{1}{\sqrt{nT}}\frac{\partial \ln L^u_{n,T}(\lambda_0)}{\partial\lambda} + O_p\left(\max\left(\frac{1}{\sqrt{nT}}, \frac{1}{T}, \sqrt{\frac{n}{T^3}}\right)\right)$ where $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^s_{n,T}(\lambda_0)}{\partial\lambda}$ is the stationary part and $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^u_{n,T}(\lambda_0)}{\partial\lambda}$ is the nonstationary part as defined via (C.5)-(C.9). For $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^s_{n,T}(\lambda_0)}{\partial\lambda}$, it has two parts $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^s_{n,T}(\lambda_0)}{\partial\lambda} = \frac{1}{\sqrt{nT}}\frac{\partial \ln L^{s*}_{n,T}(\lambda_0)}{\partial\lambda} - \Delta_{\lambda_0,nT}$ (defined in (C.5) and (C.6) respectively) where $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^{s*}_{n,T}(\lambda_0)}{\partial\lambda}$ has zero mean and $\Delta_{\lambda_0,nT}$ has nonzero mean because the latter involves $\bar{V}_{nT}$. For $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^u_{n,T}(\lambda_0)}{\partial\lambda}$, it also has two parts $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^u_{n,T}(\lambda_0)}{\partial\lambda} = \frac{1}{\sqrt{nT}}\frac{\partial \ln L^{u*}_{n,T}(\lambda_0)}{\partial\lambda} - \blacktriangle_{\lambda_0,nT}$ (defined in (C.8) and (C.9) respectively) where $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^{u*}_{n,T}(\lambda_0)}{\partial\lambda}$ has zero mean and $\blacktriangle_{\lambda_0,nT}$ has nonzero mean. To study the asymptotic behavior of $\frac{1}{\sqrt{nT}}\frac{\partial \ln L_{n,T}(\lambda_0)}{\partial\lambda}$, we will first study $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^{s*}_{n,T}(\lambda_0)}{\partial\lambda} + \frac{1}{\sqrt{nT}}\frac{\partial \ln L^{u*}_{n,T}(\lambda_0)}{\partial\lambda}$ (using Proposition 2.2) , then $\Delta_{\lambda_0,nT} + \blacktriangle_{\lambda_0,nT}$ (using Lemma B.11).

Theorem 3.3 *Under Assumptions 1-9*[8]*,*

$$\frac{1}{\sqrt{nT}}\frac{\partial \ln L_{n,T}(\lambda_0)}{\partial\lambda} + \sqrt{\frac{n}{T}}(a^s_{\lambda_0,nT} + \frac{m_n}{n}\cdot a^u_{\lambda_0,nT}) + O_p\left(\max\left(\sqrt{\frac{n}{T^3}}, \frac{1}{\sqrt{T}}\right)\right) \xrightarrow{p} N(0, \Sigma_{\lambda_0} + \Omega_{\lambda_0}). \tag{3.23}$$

[8] Only parts of Assumptions 5 and 7 are required. Specifically, S_n is invertible; and the row and column sums of W_n and S_n^{-1} are uniformly bounded in n.

where

$$\Sigma_{\lambda_0} = \frac{1}{\sigma_0^2}\lim_{T\to\infty}(\mathcal{H}_{3,nT}-\mathcal{H}_{2,nT}'\mathcal{H}_{1,nT}^{-1}\mathcal{H}_{2,nT})+\lim_{n\to\infty}\frac{1}{n}(trG_n'G_n+trG_n^2-\frac{2(trG_n)^2}{n}), \tag{3.24}$$

$$\Omega_{\lambda_0} = \frac{\mu_4-3\sigma_0^4}{\sigma_0^4}\lim_{n\to\infty}\sum_{i=1}^n G_{n,ii}^2, \tag{3.25}$$

$$a_{\lambda_0,nT}^s = \frac{1}{n}tr\left(G_n\gamma_0-(\mathcal{H}_{1,nT}^{-1}\mathcal{H}_{2,nT})_1 I_n\right)\left(\sum\nolimits_{h=0}^{\infty}B_n^h\right)S_n^{-1} + \frac{1}{n}tr\left(G_n\rho_0-(\mathcal{H}_{1,nT}^{-1}\mathcal{H}_{2,nT})_2 I_n\right)\left(\sum\nolimits_{h=0}^{\infty}W_nB_n^h\right)S_n^{-1}, \tag{3.26}$$

$$a_{\lambda_0,nT}^u = T\cdot(1-c'\mathcal{H}_{1,nT}^{-1}\mathcal{H}_{2,nT})\cdot\frac{1}{2(1-\lambda_0)}. \tag{3.27}$$

Proof. *See Appendix C.4.* ■

Also, we have the following claims.

Claim 3.4 *Under Assumptions 1-9,* $\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\lambda)}{\partial\lambda^2}-\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\lambda_0)}{\partial\lambda^2}=|\lambda-\lambda_0|\cdot O(1)+O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right)$.

Proof. *See Appendix C.5.* ■

Claim 3.5 *Under Assumptions 1-9,* $\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\lambda_0)}{\partial\lambda^2}-\frac{\partial^2 Q_{n,T}(\lambda_0)}{\partial\lambda^2}=O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right)$.

Proof. *See Appendix C.6.* ■

Using Theorem 3.3, Claim 3.4 and Claim 3.5, we have the following theorem:

Theorem 3.6 *Under Assumptions 1-9,*

$$\sqrt{nT}(\hat{\lambda}_{nT}-\lambda_0)+\sqrt{\frac{n}{T}}b_{\lambda_0,nT}+O_p\left(\max\left(\sqrt{\frac{n}{T^3}},\frac{1}{\sqrt{T}}\right)\right)\xrightarrow{d}N(0,\Sigma_{\lambda_0}^{-1}+\Sigma_{\lambda_0}^{-2}\Omega_{\lambda_0}), \tag{3.28}$$

where

$$b_{\lambda_0,nT}=\Sigma_{\lambda_0}^{-1}\left(a_{\lambda_0,nT}^s+\frac{m_n}{n}a_{\lambda_0,nT}^u\right). \tag{3.29}$$

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

$$\sqrt{nT}(\hat{\lambda}_{nT}-\lambda_0)\xrightarrow{d}N(0,\Sigma_{\lambda_0}^{-1}+\Sigma_{\lambda_0}^{-2}\Omega_{\lambda_0}). \tag{3.30}$$

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

$$\sqrt{nT}(\hat{\lambda}_{nT}-\lambda_0)+\sqrt{k}b_{\lambda_0,nT}\xrightarrow{d}N(0,\Sigma_{\lambda_0}^{-1}+\Sigma_{\lambda_0}^{-2}\Omega_{\lambda_0}). \tag{3.31}$$

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

$$T(\hat{\lambda}_{nT}-\lambda_0)+b_{\lambda_0,nT}\xrightarrow{p}0. \tag{3.32}$$

Additionally ,if v_{it} *is normal, (3.28) becomes*

$$\sqrt{nT}(\hat{\lambda}_{nT}-\lambda_0)+\sqrt{\frac{n}{T}}b_{\lambda_0,nT}+O_p\left(\max\left(\sqrt{\frac{n}{T^3}},\frac{1}{\sqrt{T}}\right)\right)\xrightarrow{d}N(0,\Sigma_{\lambda_0}^{-1}). \tag{3.33}$$

Proof. *See Appendix C.7.* ■

4 Distribution of QML Estimator $\hat{\theta}_{nT}$ and Bias Corrected $\hat{\theta}^1_{nT}$

4.1 QML Estimator $\hat{\theta}_{nT}$

After we get the distribution of $\hat{\lambda}_{nT}$, the distribution of $\hat{\delta}_{n,T} = \hat{\delta}_{n,T}(\hat{\lambda}_{nT})$, $\hat{\sigma}^2_{nT} = \hat{\sigma}^2_{nT}(\hat{\lambda}_{nT})$ and $\hat{\mathbf{c}}_{n,T} = \hat{\mathbf{c}}_{n,T}(\hat{\lambda}_{nT})$ can be derived from (2.9). As is derived in Appendix C.8,

$$\sqrt{nT}\left(\hat{\theta}_{nT} - \theta_0\right) = \Sigma^{-1}_{\theta_0,nT} \cdot \frac{1}{\sqrt{nT}} \frac{\partial \ln L_{nT}(\theta_0)}{\partial \theta} + O_p\left(\max\left(\frac{1}{\sqrt{nT}}, \frac{1}{T}, \sqrt{\frac{n}{T^3}}\right)\right), \tag{4.1}$$

where

$$\frac{1}{\sqrt{nT}} \frac{\partial \ln L_{nT}(\theta_0)}{\partial \theta} = \begin{pmatrix} \frac{1}{\sigma_0^2} \frac{1}{\sqrt{nT}} \sum_{t=1}^{T} \tilde{Z}'_{nt} \tilde{V}_{nt} \\ \frac{1}{\sigma_0^2} \frac{1}{\sqrt{nT}} \sum_{t=1}^{T} (\tilde{V}'_{nt} G_n \tilde{V}_{nt} - \sigma_0^2 tr G_n) + \frac{1}{\sigma_0^2} \frac{1}{\sqrt{nT}} \sum_{t=1}^{T} (G_n \tilde{Z}_{nt} \delta_0)' \tilde{V}_{nt} \\ \frac{1}{2\sigma_0^4} \frac{1}{\sqrt{nT}} \sum_{t=1}^{T} \left(\tilde{V}'_{nt} \tilde{V}_{nt} - n\sigma_0^2\right) \end{pmatrix},$$

$$\Sigma_{\theta_0,nT} = \frac{1}{\sigma_0^2} \begin{pmatrix} E\mathcal{H}_{nT} & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix} + \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{1}{n}\left[tr(G'_n G_n) + tr(G_n^2)\right] & \frac{1}{\sigma_0^2 n} tr(G_n) \\ \mathbf{0} & \frac{1}{\sigma_0^2 n} tr(G_n) & \frac{1}{2\sigma_0^4} \end{pmatrix}.$$

Using the central limit theorem for martingale difference arrays (see Proposition 2.2), we have the joint distribution of the common parameters in the following theorem. Denote

$$\Omega_{\theta_0,n} = \frac{\mu_4 - 3\sigma_0^4}{\sigma_0^4} \cdot \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{1}{n}\sum_{i=1}^{n} G^2_{n,ii} & \frac{1}{2\sigma_0^2 n} tr G_n \\ \mathbf{0} & \frac{1}{2\sigma_0^2 n} tr G_n & \frac{1}{4\sigma_0^4} \end{pmatrix}, \tag{4.2}$$

$$b_{\theta_0,nT} \equiv \Sigma^{-1}_{\theta_0,nT} \cdot a_{n,\theta_0}, \tag{4.3}$$

where $a_{\theta_0,nT} = a^s_{\theta_0,n} + \frac{m_n}{n} a^u_{\theta_0,T}$ with

$$a^s_{\theta_0,n} = \begin{pmatrix} \frac{1}{n} tr\left(\left(\sum_{h=0}^{\infty} B_n^h\right) S_n^{-1}\right) \\ \frac{1}{n} tr\left(W_n \left(\sum_{h=0}^{\infty} B_n^h\right) S_n^{-1}\right) \\ \mathbf{0} \\ \frac{1}{n}\gamma_0 tr(G_n \left(\sum_{h=0}^{\infty} B_n^h\right) S_n^{-1}) + \frac{1}{n}\rho_0 tr(G_n W_n \left(\sum_{h=0}^{\infty} B_n^h\right) S_n^{-1}) + \frac{1}{n} tr G_n \\ \frac{1}{2\sigma_0^2} \end{pmatrix}, \tag{4.4}$$

$$a^u_{\theta_0,T} = T \cdot \frac{1}{2(1-\lambda_0)} \cdot (c^{*\prime}, 0)'. \tag{4.5}$$

Theorem 4.1 *Under Assumptions 1-9,*

$$\sqrt{nT}(\hat{\theta}_{nT} - \theta_0) + \sqrt{\frac{n}{T}} b_{\theta_0,nT} + O_p\left(\max\left(\frac{1}{\sqrt{T}}, \sqrt{\frac{n}{T^3}}\right)\right) \xrightarrow{d} N(0, \lim_{T\to\infty} \Sigma^{-1}_{\theta_0,nT} + \lim_{T\to\infty} \Sigma^{-1}_{\theta_0,nT} \Omega_{\theta_0,n} \Sigma^{-1}_{\theta_0,nT}). \tag{4.6}$$

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

$$\sqrt{nT}(\hat{\theta}_{nT} - \theta_0) \xrightarrow{d} N(0, \lim_{T\to\infty} \Sigma_{\theta_0,nT}^{-1} + \lim_{T\to\infty} \Sigma_{\theta_0,nT}^{-1}\Omega_{\theta_0,n}\Sigma_{\theta_0,nT}^{-1}). \tag{4.7}$$

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

$$\sqrt{nT}(\hat{\theta}_{nT} - \theta_0) + \sqrt{k}b_{\theta_0,nT} \xrightarrow{d} N(0, \lim_{T\to\infty} \Sigma_{\theta_0,nT}^{-1} + \lim_{T\to\infty} \Sigma_{\theta_0,nT}^{-1}\Omega_{\theta_0,n}\Sigma_{\theta_0,nT}^{-1}). \tag{4.8}$$

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

$$T(\hat{\theta}_{nT} - \theta_0) + b_{\theta_0,nT} \xrightarrow{p} 0. \tag{4.9}$$

Additionally ,if v_{it} *is normal, (4.6) becomes*

$$\sqrt{nT}(\hat{\theta}_{nT} - \theta_0) + \sqrt{\frac{n}{T}}b_{\theta_0,nT} + O_p\left(\max\left(\frac{1}{\sqrt{T}}, \sqrt{\frac{n}{T^3}}\right)\right) \xrightarrow{d} N(0, \lim_{T\to\infty} \Sigma_{\theta_0,nT}^{-1}). \tag{4.10}$$

Proof. *See Appendix C.9.* ■

Hence, $\hat{\theta}_{nT}$ has the bias of the order $O(T^{-1})$. Also, the asymptotic variance matrix of $\sqrt{nT}\hat{\theta}_{nT}$ is singular because $\Sigma_{\theta_0,nT}^{-1} \cdot (c^{*\prime}, 0) = O(T^{-1})$. This implies that we have a different rate of convergence of $(c^{*\prime}, 0) \cdot (\hat{\theta}_{nT} - \theta_0) = \hat{\lambda}_{nT} + \hat{\gamma}_{nT} + \hat{\rho}_{nT} - 1$ using $\mathcal{H}_{nT} \cdot c^* = O(T^{-1})$ in Proposition 2.1.

Theorem 4.2 *Under Assumptions 1-9,*

$$\begin{aligned} &\sqrt{nT^3}(c^{*\prime}, 0)(\hat{\theta}_{nT} - \theta_0) + \sqrt{\frac{n}{T}}\left(T(c^{*\prime}, 0)b_{\theta_0,nT}\right) + O_p\left(\max\left(\frac{1}{\sqrt{T}}, \sqrt{\frac{n}{T^3}}\right)\right) \\ &\xrightarrow{d} N\left(0, \lim_{T\to\infty} \omega_{nT}^{-1} + \lim_{T\to\infty} T^2(c^{*\prime}, 0)(\lim_{T\to\infty} \Sigma_{\theta_0,nT}^{-1}\Omega_{\theta_0,nT}\Sigma_{\theta_0,nT}^{-1})(c^{*\prime}, 0)'\right) \end{aligned} \tag{4.11}$$

Proof. *See Appendix C.10.* ■

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

Theorem 4.3 *Under Assumptions 1-9, if* $\left(Y_{n,-1}/T\right)_i - E\left(Y_{n,-1}/T\right)_i = o_p(1)$ *and* $E\left(Y_{n,-1}/T\right)_i = O(1)$ *uniformly in* n *and* i*, then, for* $i = 1, 2, \cdots, n$, $\sqrt{T}\left(\hat{c}_{i,nT} - c_{i,0}\right) \xrightarrow{d} N(0, \Phi_{n,c_i})$ *where* Φ_{n,c_i} *is in (C.40). When* n *also goes to infinity,* $\sqrt{T}\left(\hat{c}_{i,nT} - c_{i,0}\right) \xrightarrow{d} N(0, \sigma_0^2)$.

Proof. *See Appendix C.11.* ■

4.2 Bias Corrected Estimators $\hat{\theta}^1_{nT}$

From (4.6), the QML estimator has the bias $-\frac{1}{T}b_{\theta_0,nT}$ where $b_{\theta_0,nT} \equiv \Sigma_{\theta_0,nT}^{-1} \cdot \left(a^s_{\theta_0,n} + \frac{m_n}{n}a^u_{\theta_0,T}\right)$ 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 (4.9). 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

$$\hat{\theta}^1_{nT} = \hat{\theta}_{nT} - \frac{\hat{B}_{nT}}{T}. \tag{4.12}$$

From Theorem 4.1, $B_{nT} = -\Sigma_{\theta_0,nT}^{-1} \cdot \left(a_{\theta_0,n}^s + \frac{m_n}{n} a_{\theta_0,T}^u \right)$ and a natural candidate for $\hat{B}_{nT}$ is $\left[-\Sigma_{\theta,nT}^{-1} \cdot a_{\theta,nT} \right] \Big|_{\theta=\hat{\theta}_{nT}}$. As $\Sigma_{\hat{\theta}_{nT},nT}^{-1}$ involves $E\mathcal{H}_{nT}(\hat{\theta}_{nT})$ (see (B.47)) which is hard to evaluate, our alternative estimate is

$$\hat{B}_{nT} = \left[-\ddot{\Sigma}_{\theta,nT}^{-1} \cdot a_{\theta,nT} \right] \Big|_{\theta=\hat{\theta}_{nT}}, \tag{4.13}$$

and $\ddot{\Sigma}_{\theta,nT}^{-1}$ is defined in (B.36) where $E\mathcal{H}_{nT}(\hat{\theta}_{nT})$ in $\Sigma_{\hat{\theta}_{nT},nT}$ is replaced with $\mathcal{H}_{nT}(\hat{\theta}_{nT})$. We show that when $n/T^3 \to 0$, $\hat{\theta}_{nT}^1$ is $\sqrt{nT}$ consistent and asymptotically centered normal even when $n/T \to \infty$.

To show our result for the bias corrected estimators, we need the following additional assumption.

Assumption 10. Either row sum or column sum of $\sum_{h=0}^{\infty} B_n^h(\theta)$ and $\sum_{h=1}^{\infty} hB_n^{h-1}(\theta)$ are bounded uniformly in n and in a neighborhood of θ_0.

Assumption 10 can be verified through the following lemma.

Lemma 4.4 *If $\sup_n\{\|B_n(\theta_0))\|_\infty\} < 1$ (resp: $\sup_n\{\|B_n(\theta_0))\|_1\} < 1$), then the row sum (resp: column sum) of $\sum_{h=0}^{\infty} B_n^h(\theta)$ and $\sum_{h=1}^{\infty} hB_n^{h-1}(\theta)$ are bounded uniformly in n and in a neighborhood of θ_0.*

Proof. *This is Lemma 3.9 in Yu, de Jong and Lee (2006).* ■

Our result for the bias corrected estimator is as follows.

Theorem 4.5 *Under Assumptions 1-10, if $\frac{n}{T^3} \to 0$,*

$$\sqrt{nT}(\hat{\theta}_{nT}^1 - \theta_0) \xrightarrow{d} N(0, \lim_{T\to\infty} \Sigma_{\theta_0,nT}^{-1} + \lim_{T\to\infty} \Sigma_{\theta_0,nT}^{-1} \Omega_{\theta_0,n} \Sigma_{\theta_0,nT}^{-1}). \tag{4.14}$$

Additionally,

$$\sqrt{nT^3}(c^{*\prime}, 0)(\hat{\theta}_{nT}^1 - \theta_0) \xrightarrow{d} N\left(0, \lim_{T\to\infty} \omega_{nT}^{-1} + \lim_{T\to\infty} T^2 (c^{*\prime}, 0) \left(\Sigma_{\theta_0,nT}^{-1} \Omega_{\theta_0,n} \Sigma_{\theta_0,nT}^{-1} \right) (c^{*\prime}, 0)' \right). \tag{4.15}$$

Proof. *See Appendix C.12.* ■

4.3 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 (2.1) using $\theta_0^a = (0.4, 0.2, 1, 0.4, 1)'$ and $\theta_0^b = (0.6, -0.4, 1, 0.8, 1)'$ 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 normal distributions[9] and the spatial weights matrix we use is a rook matrix. We use $T = 10$, 50 and $n = 49$, 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 $\hat{\theta}_{nT}^1$ and evaluate the bias $\hat{\theta}_{nT}^1 - \theta_0$. We do this for 1000 times to see if the bias is reduced on average by using the analytical bias correction procedure,

[9] 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. Also, in our example, the second largest eigenvalue of W_n is 0.94107. If we count it as a unit root, the bias corrected estimator does not change much.

Table 1: Performance of QMLs and Their Bias Corrected Estimators: 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	–0.0758	0.0187	–0.0135	-0.0107	-0.1211
				-0.0021	0.0161	0.0015	-0.0042	-0.0346
(2)	10	49	θ_0^b	-0.0939	0.0785	-0.0180	-0.0087	-0.1234
				-0.0050	0.0124	0.0026	-0.0063	-0.0374
(3)	10	196	θ_0^a	-0.0749	0.0160	-0.0135	-0.0108	-0.1147
				-0.0019	0.0163	0.0015	-0.0039	-0.0276
(4)	10	196	θ_0^b	-0.0919	0.0745	-0.0184	0.0071	-0.1179
				-0.0042	0.0119	0.0020	-0.0046	-0.0312
(5)	50	49	θ_0^a	-0.0139	0.0081	-0.0009	-0.0018	-0.0219
				0.0004	0.0024	-0.0000	-0.0030	-0.0020
(6)	50	49	θ_0^b	-0.0170	0.0172	-0.0003	-0.0029	-0.0204
				-0.0002	0.0031	0.0008	-0.0030	-0.0008
(7)	50	196	θ_0^a	-0.0142	0.0087	-0.0005	-0.0019	-0.0208
				0.0002	0.0040	0.0004	-0.0031	-0.0010
(8)	50	196	θ_0^b	-0.0172	0.0166	-0.0003	-0.0019	-0.0202
				-0.0004	0.0028	0.0008	-0.0023	-0.0003

Note: $\theta_0^a = (0.4, 0.2, 1, 0.4, 1)$ and $\theta_0^b = (0.6, -0.4, 1, 0.8, 1)$.

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 biases and Table 2 is for the standard errors of estimators.

We see that both estimators have some biases, but the bias corrected estimators reduce those biases. 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.

For different cases of n and T, we see that for each given n, when T is larger, the biases of two sets of estimators will be smaller and the variances 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 variances will be smaller. This is consistent with our theoretical prediction.

Table 2: Performance of QMLs and Their Bias Corrected Estimators: 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	0.0320	0.0534	0.0454	0.0426	0.0568
				0.0336	0.0572	0.0476	0.0428	0.0625
(2)	10	49	θ_0^b	0.0312	0.0415	0.0460	0.0237	0.0582
				0.0327	0.0441	0.0482	0.0237	0.0639
(3)	10	196	θ_0^a	0.0160	0.0276	0.0227	0.0221	0.0286
				0.0168	0.0296	0.0238	0.0222	0.0315
(4)	10	196	θ_0^b	0.0156	0.0214	0.0230	0.0126	0.0292
				0.0163	0.0228	0.0241	0.0126	0.0321
(5)	50	49	θ_0^a	0.0136	0.0219	0.0203	0.0184	0.0283
				0.0137	0.0222	0.0205	0.0185	0.0289
(6)	50	49	θ_0^b	0.0124	0.0165	0.0206	0.0102	0.0290
				0.0125	0.0167	0.0208	0.0103	0.0296
(7)	50	196	θ_0^a	0.0068	0.0113	0.0102	0.0095	0.0142
				0.0069	0.0114	0.0103	0.0096	0.0144
(8)	50	196	θ_0^b	0.0062	0.0085	0.0103	0.0054	0.0145
				0.0062	0.0086	0.0104	0.0055	0.0148

5 Conclusion

In this paper, we derived the properties of QML estimators of a nonstationary spatial dynamic panel data with fixed effects when both n and T are large. For the distribution of the common parameters, when 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. Compared to Yu, de Jong and Lee (2006), the estimators' rate of convergence will be the same, but the asymptotic variance matrix will be driven by the nonstationary component and it is singular. Also, the sum of the spatial effect coefficients and dynamic effect coefficient will have a higher rate of convergence. We also propose a bias correction for our estimators. We show that as long as T grows faster than $n^{1/3}$, the correction will eliminate the bias of order $O(T^{-1})$ and yield a centered confidence interval.

Appendices

A Notations

The following list summarizes some frequently used notations in the text:

$S_n(\lambda) = I_n - \lambda W_n$ for any possible λ.

$S_n = I_n - \lambda_0 W_n$.

$G_n = W_n S_n^{-1}$. $A_n = S_n^{-1}(\gamma_0 I_n + \rho_0 W_n)$.

$A_n = R_n D_n R_n^{-1}$ where R_n is the eigenvectors and D_n is the diagonal matrix of eigenvalues.

$J_n = Diag\{\mathbf{1}'_{m_n}, 0, \cdots, 0\}$ where $\mathbf{1}_{m_n}$ is an $m_n \times 1$ vector of ones.

$M_n = R_n J_n R_n^{-1}$.

$c = (1, 1, \mathbf{0}_{1\times k_x})'$ and $c^* = (c', 1)'$.

$Z_{nt} = (Y_{n,t-1},\ W_n Y_{n,t-1},\ X_{nt})$.

$\theta = (\delta', \lambda, \sigma^2)'$ where $\delta = (\gamma, \rho, \beta')'$.

$\ln L_{n,T}(\theta, \mathbf{c}_n)$ is the log-likelihood of θ and $\mathbf{c}_n$.

$\ln L_{n,T}(\lambda)$ is the concentrated log-likelihood of λ.

$Q_{n,T}(\lambda) = \max_{\delta, c_n, \sigma^2} E \frac{1}{nT} \ln L_{n,T}(\theta, \mathbf{c}_n)$.

$\tilde{\Upsilon}_{nt} = \Upsilon_{nt} - \bar{\Upsilon}_{nT}$ and $\tilde{\tilde{\Upsilon}}_{n,t-1} = \Upsilon_{n,t-1} - \bar{\Upsilon}_{nT,-1}$ where $\bar{\Upsilon}_{nT} = \frac{1}{T}\sum_{t=1}^{T}\Upsilon_{nt}$ and $\bar{\Upsilon}_{nT,-1} = \frac{1}{T}\sum_{t=1}^{T}\Upsilon_{n,t-1}$.

$\omega_{nT} = \frac{1}{nT^3}\sum_{t=1}^{T}\tilde{\tilde{Y}}^{u\prime}_{n,t-1}\tilde{\tilde{Y}}^{u}_{n,t-1}$ (see (2.7) for $\tilde{\tilde{Y}}^{u}_{n,t-1}$).

B Algebra for the Nonstationary Case

B.1 An Example to Justify the Assumptions

Consider the group case with equal weights for peers, i.e., W_n is a block diagonal matrix with its jth block being $W_{jn} = \frac{1}{n_j - 1}[l_{n_j} l'_{n_j} - I_{n_j}]$, $j = 1, \cdots, R$, where R is the total number of groups.

The eigenvalues are roots of the characteristic polynomial

$$|W_{jn} - \lambda I_{n_j}| = |\frac{1}{n_j - 1} l_{n_j} l'_{n_j} - (\lambda + \frac{1}{n_j - 1}) I_{n_j}| = (-1)^{n_j}(\lambda + \frac{1}{n_j - 1})^{n_j - 1}(\lambda - 1),$$

by using the property of a determinant that $|A + \alpha b d'| = |A|(1 + \alpha d' A^{-1} b)$ (Proposition 31 in Dhrymes (1978)). Hence the eigenvalues of W_{jn} are a single root with the unit, and $(n_j - 1)$ multiple roots with the value $(-\frac{1}{n_j - 1})$ for the jth group. As W_n is a block diagonal matrix, its determinant is the product of the determinants of the diagonal block matrices. It follows that there are R-multiple roots of the unit, and $(n_j - 1)$-multiple roots of the value $(-\frac{1}{n_j - 1})$ for each $j = 1, \cdots, R$.

As the total number of unit eigenvalues of W_n is R, the corresponding orthonormal matrix of eigenvectors of W_n is $R_n = (R_{n,R}, R_{n,n-R})$, where

$$R_{n,R} = \begin{pmatrix} \frac{l_{n_1}}{\sqrt{n_1}} & 0 & \cdots & 0 \\ 0 & \frac{l_{n_2}}{\sqrt{n_2}} & \cdots & 0 \\ \vdots & \cdots & \ddots & \cdots \\ 0 & 0 & \cdots & \frac{l_{n_R}}{\sqrt{n_R}} \end{pmatrix}.$$

As $J_n = \begin{pmatrix} I_R & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$, we have

$$R_n J_n R_n' = (R_{n,R}, R_{n,n-R}) J_n (R_{n,R}, R_{n,n-R})' = R_{n,R} R_{n,R}' = \begin{pmatrix} \frac{l_{n_1} l_{n_1}'}{n_1} & \cdots & 0 \\ 0 & \ddots & 0 \\ 0 & \cdots & \frac{l_{n_R} l_{n_R}'}{n_R} \end{pmatrix},$$

which is uniformly bounded in both row and column sums.

The matrix $A_n = (I_n - \lambda_0 W_n)^{-1}(\gamma_0 I_n + \rho_0 W_n)$ for this group setting is a diagonal block matrix. Because $B_n = A_n - R_n J_n R_n$ as defined, B_n is also a block diagonal matrix. Consider the first diagonal block of A_n which is $A_{n1} = (I_{n_1} - \lambda_0 W_{1n})^{-1}(\gamma_0 I_{n_1} + \rho_0 W_{1n})$. Note that

$$\begin{aligned} (I_{n_1} - \lambda_0 W_{1n})^{-1} &= \frac{n_1 - 1}{n_1 - 1 + \lambda_0}(I_{n_1} - \frac{\lambda_0}{n_1 - 1 + \lambda_0} l_{n_1} l_{n_1}')^{-1} \\ &= \frac{n_1 - 1}{n_1 - 1 + \lambda_0}(I_{n_1} + \frac{\lambda_0}{(n_1 - 1)(1 - \lambda_0)} l_{n_1} l_{n_1}'). \end{aligned}$$

As $\gamma_0 + \rho_0 = 1 - \lambda_0$, it follows that

$$\begin{aligned} A_{n1} &= (I_{n_1} - \lambda_0 W_{1n})^{-1}(\gamma_0 I_{n_1} + \rho_0 W_{1n}) \\ &= \frac{n_1 - 1}{n_1 - 1 + \lambda_0}\{(\gamma_0 - \frac{\rho_0}{n_1 - 1}) I_{n_1} + \frac{\lambda_0 + \rho_0}{n_1 - 1} l_{n_1} l_{n_1}'). \end{aligned}$$

Hence,

$$B_{n1} = A_{n1} - \frac{l_{n_1} l_{n_1}'}{n_1} = (\frac{n_1 \gamma_0 - 1 + \lambda_0}{n_1 - 1 + \lambda_0})(I_{n_1} - \frac{1}{n_1} l_{n_1} l_{n_1}').$$

Because $(I_{n_1} - \frac{1}{n_1} l_{n_1} l_{n_1}')$ is an idempotent matrix, it follows that for any positive integer h,

$$B_{n1}^h = (\frac{n_1 \gamma_0 - 1 + \lambda_0}{n_1 - 1 + \lambda_0})^h (I_{n_1} - \frac{1}{n_1} l_{n_1} l_{n_1}').$$

The $(I_{n_1} - \frac{1}{n_1} l_{n_1} l_{n_1}')$ is uniformly bounded in both row and column sums, so $\sum_{h=0}^{\infty} abs(B_{n1}^h)$ will be uniformly bounded in both row and column sum if $|(\frac{n_1 \gamma_0 - 1 + \lambda_0}{n_1 - 1 + \lambda_0})| < 1$. The corresponding B_n will be so if $\max_{j=1,\cdots,R} |(\frac{n_j \gamma_0 - 1 + \lambda_0}{n_j - 1 + \lambda_0})| < 1$. A sufficient condition for this to occur is that $|\lambda_0| < 1$, $\gamma_0 < 1$ and $\rho_0 < 1$. This is so as follows. Define the function $f(x) = \frac{x\gamma_0 - 1 + \lambda_0}{x - 1 + \lambda_0}$. The derivative of $f(x)$ is $\frac{df(x)}{dx} = \frac{(1-\lambda_0)(1-\gamma_0)}{(x-1+\lambda_0)^2}$ which will be positive if $1 > \lambda_0$ and $1 > \gamma_0$. As the upper bound of $f(x)$ will be γ_0 and its lower bound is $f(2)=\frac{2\gamma_0-1+\lambda_0}{1+\lambda_0}=\frac{1-\lambda_0-2\rho_0}{1+\lambda_0}>-1$ because $1+\lambda_0>0$, $1>\rho_0$ and $x\geq 2$. Under this situation, we can justify the Assumption 8 in the text for this example.

The same consideration will also justify that the smallest eigenvalue of A_n is less than one in absolute value in Assumption 6[10]. Because A_n is a block diagonal matrix, it is sufficient to consider the eigenvalues of each block A_{nj}. An eigenvalue λ of A_{nj} for some j will also be the eigenvalue of A_n. This is so because if λ is an eigenvalue of A_{nj} with eigenvector x_{nj} such that $A_{nj}x_{nj}=\lambda x_{nj}$, then $A_n x_n=\lambda x_n$ where $x_n=(0,\cdots,0,x_{nj}',0,\cdots,0)'$. Consider the eigenvalue $(-\frac{1}{n_1-1})$ of W_{n1} (and the remaining eigenvalue is one) and the corresponding eigenvalue x_1. As

$$A_{n1}x_1=(I_{n_1}-\lambda_0 W_{n1})^{-1}(\gamma_0 I_{n_1}+\rho_0 W_{n_1})x_1=(\frac{n_1\gamma_0-1+\lambda_0}{n_1-1+\lambda_0})x_1,$$

thus the corresponding eigenvalue of A_n is $(\frac{n_1\gamma_0-1+\lambda_0}{n_1-1+\lambda_0})$, which lies in $(-1,\gamma_0)$ with $\gamma_0<1$, as previously shown.

B.2 Some Basic Lemmas

Proposition B.1 *Suppose that W_n is a weights matrix row normalized from a symmetric matrix C_n, i.e., $W_n=\Lambda_n^{-1}C_n$, where Λ_n is a diagonal matrix with its diagonal elements formed by the row sums of C_n. Then, the eigenvalues of W_n are all real and W_n is diagonalizable.*

Proposition B.2 *Suppose that $A_n=(I_n-\lambda_0 W_n)^{-1}(\gamma_0 I_n+\rho_0 W_n)$, where W_n is the row normalized weights matrix in Proposition B.1. Then, A_n is diagonalizable with all real eigenvalues. If W_n is diagonalizable as $W_n=R_n D_n^* R_n^{-1}$, then A_n can be diagonalizable as $A_n=R_n D_n R_n^{-1}$, with its eigenvalue matrix $D_n=(I_n-\lambda_0 D_n^*)^{-1}(\gamma_0 I_n+\rho_0 D_n^*)$.*

Proposition B.3 *Denote $d_{n,i}$'s the eigenvalues of A_n. Under Assumption 1 for W_n, $|\lambda_0|<1$ and $\rho_0+\gamma_0+\lambda_0=1$, (1) if $\rho_0+\gamma_0\lambda_0>0$ and $\frac{\gamma_0-\rho_0}{1+\lambda_0}>-1$, we have $d_{n,\max}=1$ and $d_{n,\min}>-1$; (2) when $\rho_0+\gamma_0+\lambda_0=1$, "$\rho_0<1$, $|\gamma_0|<1$ and $|\lambda_0|<1$" implies "$\rho_0+\gamma_0\lambda_0>0$ and $\frac{\gamma_0-\rho_0}{1+\lambda_0}>-1$".*

Proposition B.4 *(1) Suppose that $|\lambda_0|<1$ and $\gamma_0\neq 1$, then the unit eigenvalues of W_n correspond to unit eigenvalues of A_n via the relation $\frac{\gamma_0+\rho_0\varpi_{ni}}{1-\lambda_0\varpi_{ni}}$, if and only if $\rho_0+\gamma_0+\lambda_0=1$.*

(2) $A_n R_n J_n R_n^{-1}=R_n J_n R_n^{-1}A_n=R_n J_n R_n^{-1}$.

(3) Assuming that the unit eigenvalues of W_n correspond to unit eigenvalues of A_n, then,

(3i) $W_n R_n J_n R_n^{-1}=R_n J_n R_n^{-1}$;

(3ii) $S_n^{-1}R_n J_n R_n^{-1}=R_n J_n R_n^{-1}S_n^{-1}=\frac{1}{1-\lambda_0}R_n J_n R_n^{-1}$.

[10] See also sufficient conditions on parameters in Proposition B.3, which guarantee Assumption 6.

Proposition B.5 *Under Assumptions 5 and 6, for Y_{nt} in (2.1), $Y_{nt} = Y_{nt}^u + Y_{nt}^s$ where*

$$Y_{nt}^u = (R_n J_n R_n^{-1}) \left(Y_{n,-1} + \mathbf{c}_{n0} \frac{t}{(1-\lambda_0)} + \frac{\sum_{h=0}^{t-1} X_{nh}\beta_0}{(1-\lambda_0)} + \frac{\sum_{h=0}^{t-1} V_{nh}}{(1-\lambda_0)} \right), \quad \text{(B.1)}$$

$$Y_{nt}^s = \sum_{h=0}^{\infty} B_n^h S_n^{-1} \mathbf{c}_{n0} + \sum_{h=0}^{\infty} B_n^h S_n^{-1} X_{n,t-h}\beta_0 + \sum_{h=0}^{\infty} B_n^h S_n^{-1} V_{n,t-h}. \quad \text{(B.2)}$$

Furthermore,

$$\tilde{\tilde{Y}}_{n,t-1}^u = \frac{R_n J_n R_n^{-1}}{(1-\lambda_0)} \left(\mathbf{c}_{n0}[(t-1) - (\frac{T-1}{2})] + \sum_{h=0}^{t-2} (X_{nh}\beta_0 + V_{nh}) - \frac{1}{T} \sum_{h=0}^{T-2} (T-1-h)(X_{nh}\beta_0 + V_{nh}) \right),$$

$$\tilde{\tilde{Y}}_{n,t-1}^s = S_n^{-1} \sum_{h=0}^{\infty} B_n^h [(X_{n,t-1-h} - \frac{1}{T} \sum_{t=0}^{T-1} X_{n,t-h})\beta_0 + (V_{n,t-1-h} - \frac{1}{T} \sum_{t=0}^{T-1} V_{n,t-h})].$$

Proposition B.6 *Under Assumptions 5 and 6, for the nonstationary part $Y_{n,t-1}^u$ of $Y_{n,t-1}$,*

$$W_n Y_{n,t-1}^u = Y_{n,t-1}^u, \; G_n Y_{n,t-1}^u = \frac{1}{1-\lambda_0} Y_{n,t-1}^u. \quad \text{(B.3)}$$

Also, for nonstationary part Z_{nt}^u of Z_{nt}, denote $c = (1, 1, \mathbf{0}_{1\times k_x})'$, we have

$$Z_{nt}^u = (Y_{n,t-1}^u, W_n Y_{n,t-1}^u, \mathbf{0}_{1\times k_x}) = Y_{n,t-1}^u \cdot c', \; G_n Z_{nt}^u \delta_0 = Y_{n,t-1}^u. \quad \text{(B.4)}$$

Denote $\xi_{nt} = \sum_{h=0}^{t-1} V_{nh}$, $\mathbb{X}_{nt} = \sum_{h=0}^{t-1} X_{nh}$, $\mathbb{U}_{nt} = \sum_{h=1}^{\infty} P_{nt,h} V_{n,t+1-h}$ and $\mathbb{W}_{nt} = \sum_{h=1}^{\infty} Q_{nt,h} V_{n,t+1-h}$ where $P_{nt,h}$ and $Q_{nt,h}$ are $n \times n$ nonstochastic matrices and the row and column sums of $\sum_{h=1}^{\infty} abs(P_{nt,h})$ and $\sum_{h=1}^{\infty} abs(Q_{nt,h})$ are bounded uniformly in n and t. Also, $n \times 1$ vector D_{nt} are nonstochastic and bounded, uniformly in n and t. We note that $\bar{\xi}_{nT} = \frac{1}{T}\sum_{t=1}^{T} \xi_{nt} = \sum_{h=1}^{T} \frac{h}{T} V_{n,T-h}$. Also, as $\bar{\mathbb{U}}_{nT} = \left(\sum_{t=1}^{T} \mathbb{U}_{nt}\right)/T$, we have $\bar{\mathbb{U}}_{nT} = \sum_{h=1}^{\infty} \bar{P}_{nT,h} V_{n,T+1-h}$, where

$$\bar{P}_{nT,h} = \begin{cases} \frac{1}{T}(P_{nT,1} + P_{nT,2} + \cdots + P_{nT,h}) = \frac{1}{T}\sum_{g=1}^{h} P_{n,T-h+g,g} & \text{for } h \le T \\ \frac{1}{T}\sum_{g=h-T+1}^{h} P_{n,T-h+g,g} & \text{for } h > T. \end{cases} \quad \text{(B.5)}$$

Lemma B.7 *Under Assumption 2, for $t \ge s$,*

$$E(\mathbb{U}_{nt}\mathbb{W}_{ns}') = \sigma_0^2 \left(\sum_{h=1}^{\infty} P_{nt,t-s+h} Q_{ns,h}' \right), \; E(\mathbb{U}_{nt}'\mathbb{W}_{ns}) = \sigma_0^2 tr \left(\sum_{h=1}^{\infty} P_{nt,t-s+h}' Q_{ns,h} \right), \quad \text{(B.6)}$$

$$Cov(\mathbb{U}_{nt}'\mathbb{W}_{nt}, \mathbb{U}_{ns}'\mathbb{W}_{ns}) = (\mu_4 - 3\sigma_0^4) \sum_{h=1}^{\infty}\sum_{i=1}^{n} (P_{nt,t-s+h}' Q_{nt,t-s+h})_{ii} (P_{ns,h}' Q_{ns,h})_{ii} +$$
$$\sigma_0^4 tr \left[\left(\sum_{h=1}^{\infty} P_{ns,h} P_{nt,t-s+h}' \right) \left(\sum_{h=1}^{\infty} Q_{nt,t-s+h} Q_{ns,h}' \right) + \left(\sum_{h=1}^{\infty} Q_{ns,h} P_{nt,t-s+h}' \right) \left(\sum_{h=1}^{\infty} Q_{nt,t-s+h} P_{ns,h}' \right) \right]$$

Lemma B.8 [11] *Denote $\mathcal{B}_n$ an $n \times n$ nonstochastic matrix which is row sum and column sum bounded uniformly in n. Under Assumptions 1-8, for ξ_{nt}, $\mathbf{c}_{n0}$, $\mathbb{X}_{nt}$, $\mathbb{U}_{nt}$, D_{nt} and their cross products, we have*

$$\frac{1}{nT} \sum_{t=1}^{T} (\mathbf{c}_{n0}\tilde{t} + \tilde{\mathbb{X}}_{nt}\beta_0)' \mathcal{B}_n (\mathbf{c}_{n0}\tilde{t} + \tilde{\mathbb{X}}_{nt}\beta_0) = O\left(T^2\right); \quad \text{(B.7)}$$

[11] For $M_n = R_n J_n R_n^{-1} = A_n - B_n$, as $A_n = S_n^{-1}(\gamma_0 I_n + \rho_0 W_n)$ is row sum and column sum bounded, and B_n is also row sum and column sum bounded implied by Assumption 8, M_n is also row sum and column sum bounded. Hence, we can replace $\mathcal{B}_n$ with M_n or $M_n' M_n$ to apply following lemmas.

$$\frac{1}{nT}\sum\nolimits_{t=1}^{T}(\mathbf{c}_{n0}\tilde{t}+\tilde{\mathbb{X}}_{nt}\beta_0)'\mathcal{B}_n\tilde{\xi}_{nt}=O_p\left(\sqrt{\frac{T^3}{n}}\right) \text{ with zero mean;} \tag{B.8}$$

$$\frac{1}{nT}\sum\nolimits_{t=1}^{T}\tilde{\xi}_{nt}'\mathcal{B}_n\tilde{\xi}_{nt}-E(\frac{1}{nT}\sum\nolimits_{t=1}^{T}\tilde{\xi}_{nt}'\mathcal{B}_n\tilde{\xi}_{nt})=O_p\left(\frac{T}{\sqrt{n}}\right) \tag{B.9}$$

where $E(\frac{1}{nT}\sum_{t=1}^{T}\tilde{\xi}_{nt}'\mathcal{B}_n\tilde{\xi}_{nt})=O(T)$*;*

$$\frac{1}{nT}\sum\nolimits_{t=1}^{T}(\mathbf{c}_{n0}\tilde{t}+\tilde{\mathbb{X}}_{nt}\beta_0)'\mathcal{B}_nD_{nt}=O\left(T\right); \tag{B.10}$$

$$\frac{1}{nT}\sum\nolimits_{t=1}^{T}(\mathbf{c}_{n0}\tilde{t}+\tilde{\mathbb{X}}_{nt}\beta_0)'\tilde{\mathbb{U}}_{nt}=O_p\left(\sqrt{\frac{T}{n}}\right) \text{ with mean zero;} \tag{B.11}$$

$$\frac{1}{nT}\sum\nolimits_{t=1}^{T}\tilde{\xi}_{nt}'\mathcal{B}_nD_{nt}=O_p\left(\sqrt{\frac{T}{n}}\right) \text{ with mean zero;} \tag{B.12}$$

$$\frac{1}{nT}\sum\nolimits_{t=1}^{T}\tilde{\xi}_{nt}'\tilde{\mathbb{U}}_{nt}-E(\frac{1}{nT}\sum\nolimits_{t=1}^{T}\tilde{\xi}_{nt}'\tilde{\mathbb{U}}_{nt})=O_p\left(\sqrt{\frac{T}{n}}\right), \tag{B.13}$$

where $E(\frac{1}{nT}\sum_{t=1}^{T}\tilde{\xi}_{nt}'\tilde{\mathbb{U}}_{nt})=O(1)$ *and*

$$\frac{1}{nT}\sum\nolimits_{t=1}^{T}\tilde{\mathbb{U}}_{nt}'\tilde{\mathbb{W}}_{nt}-E(\frac{1}{nT}\sum\nolimits_{t=1}^{T}\tilde{\mathbb{U}}_{nt}'\tilde{\mathbb{W}}_{nt})=O\left(\frac{1}{\sqrt{nT}}\right), \tag{B.14}$$

where $E(\frac{1}{nT}\sum_{t=1}^{T}\tilde{\mathbb{U}}_{nt}'\tilde{\mathbb{W}}_{nt})=O(1)$*.*

Lemma B.9 *Denote* $\mathcal{B}_n$ *an* $n\times n$ *nonstochastic matrix which is row sum and column sum bounded uniformly in* n*. Under Assumptions 1-8, for* $\mathbf{c}_{n0}$*,* $\mathbb{X}_{nt}$*,* ξ_{nt}*,* V_{nt}*,* $\mathbb{U}_{nt}$ *and their cross products,*

$$\frac{1}{nT}\sum\nolimits_{t=1}^{T}(\mathbf{c}_{n0}\tilde{t}_{-1}+\tilde{\mathbb{X}}_{n,t-1}\beta_0)'\mathcal{B}_nV_{nt}=O_p\left(\sqrt{\frac{T}{n}}\right); \tag{B.15}$$

$$\frac{1}{nT}\sum\nolimits_{t=1}^{T}\xi_{n,t-1}'\mathcal{B}_nV_{nt}=O_p\left(\frac{1}{\sqrt{n}}\right); \tag{B.16}$$

$$\frac{1}{n}\bar{\xi}_{n,T-1}'\mathcal{B}_n\bar{V}_{nT}-E\frac{1}{n}\bar{\xi}_{n,T-1}'\mathcal{B}_n\bar{V}_{nT}=O_p\left(\frac{1}{\sqrt{n}}\right), \tag{B.17}$$

where $E\frac{1}{n}\bar{\xi}_{n,T-1}'\mathcal{B}_n\bar{V}_{nT}=O(1)$ *and for* $\mathcal{B}_n=M_n'$*,* $E\frac{1}{n}(M_n\bar{\xi}_{n,T-1})'\bar{V}_{nT}=\sigma_0^2\frac{(T-1)(T-2)m_n}{2T^2n}=O(\frac{m_n}{n})$*;*

$$\frac{1}{n}\bar{\mathbb{U}}_{n,T-1}'\bar{V}_{nT}-E\frac{1}{n}\bar{\mathbb{U}}_{n,T-1}'\bar{V}_{nT}=O_p\left(\frac{1}{\sqrt{n}}\right), \tag{B.18}$$

where $E\frac{1}{n}\bar{\mathbb{U}}_{n,T-1}'\bar{V}_{nT}=O(\frac{1}{T})$*;*

$$\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}_{nt}'\mathcal{B}_n\tilde{V}_{nt}=(1-\frac{1}{T})\sigma_0^2\frac{1}{n}tr(\mathcal{B}_n)+O_p\left(\frac{1}{\sqrt{nT}}\right). \tag{B.19}$$

Lemma B.10 *Denote $\mathcal{B}_n$ an $n \times n$ nonstochastic matrix which is row sum and column sum bounded uniformly in n. Under Assumptions 1-8,*

$$\frac{1}{nT}\sum_{t=1}^{T}\tilde{\tilde{Y}}^{u\prime}_{n,t-1}\mathcal{B}_n\tilde{\tilde{Y}}^{u}_{n,t-1} - E\frac{1}{nT}\sum_{t=1}^{T}\tilde{\tilde{Y}}^{u\prime}_{n,t-1}\mathcal{B}_n\tilde{\tilde{Y}}^{u}_{n,t-1} = O_p\left(T\cdot\sqrt{\frac{T}{n}}\right), \tag{B.20}$$

$$\frac{1}{nT}\sum_{t=1}^{T}\tilde{\tilde{Y}}^{u\prime}_{n,t-1}\mathcal{B}_n\tilde{\tilde{Y}}^{s}_{n,t-1} - E\frac{1}{nT}\sum_{t=1}^{T}\tilde{\tilde{Y}}^{u\prime}_{n,t-1}\mathcal{B}_n\tilde{\tilde{Y}}^{s}_{n,t-1} = O_p\left(\sqrt{\frac{T}{n}}\right), \tag{B.21}$$

and

$$\frac{1}{nT}\sum_{t=1}^{T}\tilde{\tilde{Y}}^{s\prime}_{n,t-1}\mathcal{B}_n\tilde{\tilde{Y}}^{s}_{n,t-1} - E\frac{1}{nT}\sum_{t=1}^{T}\tilde{\tilde{Y}}^{s\prime}_{n,t-1}\mathcal{B}_n\tilde{\tilde{Y}}^{s}_{n,t-1} = O_p\left(\frac{1}{\sqrt{nT}}\right), \tag{B.22}$$

where $E\frac{1}{nT}\sum_{t=1}^{T}\tilde{\tilde{Y}}^{u\prime}_{n,t-1}\mathcal{B}_n\tilde{\tilde{Y}}^{u}_{n,t-1} = O(T^2)$, $E\frac{1}{nT}\sum_{t=1}^{T}\tilde{\tilde{Y}}^{u}_{n,t-1}\mathcal{B}_n\tilde{\tilde{Y}}^{s}_{n,t-1} = O(T)$ *and* $E\frac{1}{nT}\sum_{t=1}^{T}\tilde{\tilde{Y}}^{s}_{n,t-1}\mathcal{B}_n\tilde{\tilde{Y}}^{s}_{n,t-1} = O(1)$.

Lemma B.11 *Under Assumptions 1-8 and $\mathcal{B}_n$ is an $n \times n$ nonstochastic matrix which is row sum and column sum bounded uniformly in n,*

$$\frac{1}{nT}\sum_{t=1}^{T}\tilde{\tilde{Y}}^{u\prime}_{n,t-1}\mathcal{B}_nV_{nt} - E\frac{1}{nT}\sum_{t=1}^{T}\tilde{\tilde{Y}}^{u\prime}_{n,t-1}\mathcal{B}_nV_{nt} = O_p\left(\sqrt{\frac{T}{n}}\right), \tag{B.23}$$

$$\frac{1}{nT}\sum_{t=1}^{T}\tilde{\tilde{Y}}^{s\prime}_{n,t-1}\mathcal{B}_nV_{nt} - E\frac{1}{nT}\sum_{t=1}^{T}\tilde{\tilde{Y}}^{s\prime}_{n,t-1}\mathcal{B}_nV_{nt} = O_p\left(\frac{1}{\sqrt{nT}}\right), \tag{B.24}$$

where $E\frac{1}{nT}\sum_{t=1}^{T}\tilde{\tilde{Y}}^{u\prime}_{n,t-1}\mathcal{B}_nV_{nt} = O(1)$ *and* $E\frac{1}{nT}\sum_{t=1}^{T}\tilde{\tilde{Y}}^{s\prime}_{n,t-1}\mathcal{B}_nV_{nt} = O(\frac{1}{T})$. *For the special case with* $\mathcal{B}_n = I_n$, *we have* $E\frac{1}{nT}\sum_{t=1}^{T}\tilde{\tilde{Y}}^{u\prime}_{n,t-1}V_{nt} = \sigma_0^2\frac{(T-1)(T-2)m_n}{2T^2n}\frac{1}{1-\lambda_0} = O(\frac{m_n}{n})$ *where m_n is the number of unit roots.*

Proposition B.12 *Consider the $m \times m$ square matrix $H_T = I_m + T(g_Td_T' + h_Tb_T')$, where g_T, h_T, b_T, and d_T are all m-dimensional column vectors. Then, under the assumption $\Delta_T \neq 0$,*

$$H_T^{-1} = I_m - \frac{T}{\Delta_T}B_T, \tag{B.25}$$

where

$$\Delta_T = 1 + T(b_T'h_T + d_T'g_T) - T^2Det\left((b_T, d_T)'(g_T, h_T)\right), \tag{B.26}$$

and

$$B_T = (h_Tb_T' + g_Td_T') - T[(d_T'h_T)g_Tb_T' + (b_T'g_T)h_Td_T' - (d_T'g_T)h_Tb_T' - (b_T'h_T)g_Td_T']. \tag{B.27}$$

Proposition B.13 *Consider the $m \times m$ stochastic matrix K_T,*

$$K_T = T^2c_Tc_T' + T(b_Td_T' + d_T'b_T) + A_T,$$

where c_T, b_T and d_T are m-dimensional column random vectors with c_T proportional to b_T such that $c_T = \omega_T \cdot b_T$, where ω_T is a nonzero random variable with probability one. Suppose that, as $T \to \infty$, c_T, b_T, d_T, and ω_T converge in probability, respectively, to finite limits c, b, d and ω where c and ω are nonstochastic and

nonzero. Assume that A_T is positive definite for large enough T with probability one and its limit A exists and is also a positive definite matrix. Then, under the condition that $\Delta = 1 - \frac{1}{\omega^2}\left[d'A^{-1}d - \frac{(d'A^{-1}b)^2}{b'A^{-1}b}\right] \neq 0$,
(a) the limit of K_T^{-1} is a finite matrix L_k, where

$$L_k = (A^{-1} - \frac{1}{b'A^{-1}b}A^{-1}bb'A^{-1}) + \frac{1}{\omega^2\Delta}A^{-1}(d - \frac{d'A^{-1}b}{b'A^{-1}b}b)(d - \frac{d'A^{-1}b}{b'A^{-1}b}b)'A^{-1};$$

(b) $K_T^{-1}\cdot c_T = O_p(T^{-1})$;
(c) $c_T'K_T^{-1}c_T = O_p(T^{-2})$ and $T^2c_T'K_T^{-1}c_T = 1 + O_p(T^{-1})$.

Proposition B.14 *Assuming plim$_{T\to\infty}\mathcal{H}_{nT}^s$ is positive definite, we have*

$$\mathcal{H}_{1,nT}^{-1}c = O_p(T^{-1}), \tag{B.28}$$

$$c'\mathcal{H}_{1,nT}^{-1}c = O_p(T^{-2}), \tag{B.29}$$

$$\mathcal{H}_{1,nT}^{-1}\mathcal{H}_{2,nT} = O_p(1), \tag{B.30}$$

$$1 - c'\mathcal{H}_{1,nT}^{-1}\mathcal{H}_{2,nT} = O_p(T^{-1}), \tag{B.31}$$

$$\left(\mathcal{H}_{3,nT} - \mathcal{H}_{2,nT}'\mathcal{H}_{1,nT}^{-1}\mathcal{H}_{2,nT}\right)^{-1} \text{ exists and is } O_p(1), \tag{B.32}$$

$$\text{plim}_{T\to\infty}\left(\mathcal{H}_{3,nT} - \mathcal{H}_{2,nT}'\mathcal{H}_{1,nT}^{-1}\mathcal{H}_{2,nT}\right)^{-1} \neq 0, \tag{B.33}$$

$$\left(\mathcal{H}_{3,nT} - \mathcal{H}_{2,nT}'\mathcal{H}_{1,nT}^{-1}\mathcal{H}_{2,nT}\right) - \left(E\mathcal{H}_{3,nT} - E\mathcal{H}_{2,nT}'(E\mathcal{H}_{1,nT})^{-1}E\mathcal{H}_{2,nT}\right) \xrightarrow{p} 0, \tag{B.34}$$

$$\text{plim}_{T\to\infty}\left(\mathcal{H}_{3,nT} - \mathcal{H}_{2,nT}'\mathcal{H}_{1,nT}^{-1}\mathcal{H}_{2,nT}\right) = \lim_{T\to\infty}\left(E\mathcal{H}_{3,nT} - E\mathcal{H}_{2,nT}'(E\mathcal{H}_{1,nT})^{-1}E\mathcal{H}_{2,nT}\right). \tag{B.35}$$

Proposition B.15 *For QML estimator $\hat\theta_{nT}$ in Theorem 4.1, define*

$$\ddot\Sigma_{\hat\theta_{nT},nT} = \frac{1}{\hat\sigma_{nT}^2}\begin{pmatrix} \mathcal{H}_{nT}(\hat\theta_{nT}) & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix} + \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{1}{n}\left[tr(G_n'(\hat\lambda_{nT})G_n(\hat\lambda_{nT})) + tr(G_n^2(\hat\lambda_{nT}))\right] & \frac{1}{\hat\sigma_{nT}^2 n}tr(G_n(\hat\lambda_{nT})) \\ \mathbf{0} & \frac{1}{\hat\sigma_{nT}^2 n}tr(G_n(\hat\lambda_{nT})) & \frac{1}{2\hat\sigma_{nT}^4} \end{pmatrix}, \tag{B.36}$$

where $\mathcal{H}_{nT}(\hat\theta_{nT})$ is $\mathcal{H}_{nT}(\theta)$ (see (B.47)) evaluated at $\hat\theta_{nT}$, then,

$$\ddot\Sigma_{\hat\theta_{nT},nT}^{-1} - \Sigma_{\theta_0,nT}^{-1} = O_p\left(\max\left(\frac{1}{\sqrt{nT}}, \frac{1}{T}\right)\right), \tag{B.37}$$

$$T\cdot\left[\ddot\Sigma_{\hat\theta_{nT},nT}^{-1} - \Sigma_{\theta_0,nT}^{-1}\right](c^{*\prime}, 0)' = O_p\left(\max\left(\frac{1}{\sqrt{nT}}, \frac{1}{T}\right)\right). \tag{B.38}$$

B.3 Proofs for Basic Lemmas

Proof for Proposition B.1: The first part is known in Ord (1975). To show that W_n is diagonalizable, note that as $W_n = \Lambda_n^{-1} C_n$, it implies that $\Lambda_n^{\frac{1}{2}} W_n \Lambda_n^{-\frac{1}{2}} = \Lambda_n^{-\frac{1}{2}} C_n \Lambda_n^{-\frac{1}{2}}$, which is a symmetric matrix. Let D_n^* be the eigenvalue matrix of $\Lambda_n^{-\frac{1}{2}} C_n \Lambda_n^{-\frac{1}{2}}$, which is real; and let R_n^* be the corresponding orthonormal matrix such that $\Lambda_n^{-\frac{1}{2}} C_n \Lambda_n^{-\frac{1}{2}} = R_n^* D_n^* R_n^{*'}$. Hence, $W_n = \Lambda_n^{-\frac{1}{2}} (R_n^* D_n^* R_n^{*'}) \Lambda_n^{\frac{1}{2}} = R_n D_n^* R_n^{-1}$ where $R_n = \Lambda_n^{-\frac{1}{2}} R_n^*$ is an eigenvector matrix of W_n, and D_n^* is the eigenvalue matrix for W_n. ■

Proof for Proposition B.2: Because $W_n = R_n D_n^* R_n^{-1}$ from Proposition B.1, it follows that

$$\begin{aligned} A_n &= (I_n - \lambda_0 W_n)^{-1} (\gamma_0 I_n + \rho_0 W_n) \\ &= (I_n - \lambda_0 R_n D_n^* R_n^{-1})^{-1} (\gamma_0 I_n + \rho_0 R_n D_n^* R_n^{-1}) \\ &= R_n (I_n - \lambda_0 D_n^*)^{-1} R_n^{-1} \cdot R_n (\gamma_0 I_n + \rho_0 D_n^*) R_n^{-1} \\ &= R_n (I_n - \lambda_0 D_n^*)^{-1} (\gamma_0 I_n + \rho_0 D_n^*) R_n^{-1}. \end{aligned}$$

Note that $D_n = (I_n - \lambda_0 D_n^*)^{-1} (\gamma_0 I_n + \rho_0 D_n^*)$ is a diagonal real matrix because D_n^* is diagonal and real, and $(I_n - \lambda_0 D_n^*)$ is invertible because $(I_n - \lambda_0 W_n)$ is assumed to be invertible to begin with. ■

Proof for Proposition B.3: For (1): The eigenvalue of A_n has the formula $\frac{\gamma_0 + \rho_0 \varpi_{ni}}{1 - \lambda_0 \varpi_{ni}}$ where ϖ_{ni} is an eigenvalue of W_n with $|\varpi_{ni}| \leq 1$ for all i and $\varpi_{ni} = 1$ for some i (see Ord (1975)). Because $\frac{\partial \left(\frac{\gamma_0 + \rho_0 \varpi_{ni}}{1 - \lambda_0 \varpi_{ni}} \right)}{\partial \varpi_{ni}} = \frac{\rho_0 + \gamma_0 \lambda_0}{(1 - \lambda_0 \varpi_{ni})^2}$ and $|\varpi_{ni}| \leq 1$ for all i, $\rho_0 + \gamma_0 \lambda_0 > 0$ will imply that $\frac{\gamma_0 + \rho_0 \varpi_{ni}}{1 - \lambda_0 \varpi_{ni}}$ is an increasing function of ϖ_{ni}. As $\varpi_{n,\max} = 1$, the maximum value of $\frac{\gamma_0 + \rho_0 \varpi_{ni}}{1 - \lambda_0 \varpi_{ni}}$ will be achieved at $\varpi_{ni} = 1$; additionally, as $\varpi_{n,\min} \geq -1$, $\frac{\gamma_0 - \rho_0}{1 + \lambda_0} > -1$ will assure that the minimum value of $\frac{\gamma_0 + \rho_0 \varpi_{ni}}{1 - \lambda_0 \varpi_{ni}}$ will be greater than -1. For (2): If $\rho_0 + \gamma_0 + \lambda_0 = 1$, then $\rho_0 + \gamma_0 \lambda_0 > 0$ is equivalent to $(1 - \gamma_0)(1 - \lambda_0) > 0$; also, $\frac{\gamma_0 - \rho_0}{1 + \lambda_0} > -1$ if $\gamma_0 + \lambda_0 > 0$ and $\lambda_0 > -1$. Under $\rho_0 + \gamma_0 + \lambda_0 = 1$, $\gamma_0 + \lambda_0 > 0$ is equivalent to $\rho_0 < 1$. The conclusion in (2) follows. ■

Proof for Proposition B.4: (1) An eigenvalue d_{ni} of A_n has the form $d_{ni} = \frac{\gamma_0 + \rho_0 \varpi_{ni}}{1 - \lambda_0 \varpi_{ni}}$ for some eigenvalue ϖ_{ni} of W_n. Thus, $d_{ni} = 1$ is equivalent to $\gamma_0 + \rho_0 \varpi_{ni} = 1 - \lambda_0 \varpi_{ni}$ when $|\lambda_0| < 1$. It is apparent that $\varpi_{ni} = 1$ is equivalent to $d_{ni} = 1$ when $\rho_0 + \gamma_0 + \lambda_0 = 1$ and $\gamma_0 \neq 1$. That $\rho_0 + \gamma_0 + \lambda_0 = 1$ is a necessary condition is trivial. (2) Because $A_n = R_n D_n R_n^{-1}$ and $D_n J_n = (J_n + \tilde{D}_n) J_n = J_n$, we have $A_n R_n J_n R_n^{-1} = R_n D_n J_n R_n^{-1} = R_n J_n R_n^{-1}$. Note that because J_n and D_n are diagonal matrices, $R_n J_n R_n^{-1} A_n = R_n J_n D_n R_n^{-1} = R_n D_n J_n R_n^{-1} = A_n R_n J_n R_n^{-1}$. (3) From Proposition B.1, $W_n = R_n D_n^* R_n^{-1}$. Hence, $W_n R_n J_n R_n^{-1} = R_n D_n^* J_n R_n^{-1} = R_n J_n R_n^{-1}$ as $D_n^* J_n = J_n$ when the unit eigenvalues of W_n correspond to unit eigenvalues of A_n. As $S_n^{-1} = R_n (I_n - \lambda_0 D_n^*)^{-1} R_n^{-1}$, we have $S_n^{-1} R_n J_n = R_n (I_n - \lambda_0 D_n^*)^{-1} J_n = \frac{1}{1 - \lambda_0} R_n J_n$ because $(I_n - \lambda_0 D_n^*)^{-1} J_n = \frac{1}{1 - \lambda_0} J_n$. It follows that $S_n^{-1} R_n J_n R_n^{-1} = \frac{1}{1 - \lambda_0} R_n J_n R_n^{-1}$. Furthermore, $R_n J_n R_n^{-1} S_n^{-1} = R_n J_n (I_n - \lambda_0 D_n^*)^{-1} R_n^{-1} = R_n (I_n - \lambda_0 D_n^*)^{-1} J_n R_n^{-1} = S_n^{-1} R_n J_n R_n^{-1}$. ■

Proof for Proposition B.5: Suppose that the number of unit roots of A_n is m_n, then $D_n = J_n + \tilde{D}_n$ where $J_n = Diag\{\mathbf{1}'_{m_n}, 0, \cdots, 0\}$ and $\tilde{D}_n = Diag\{0, \cdots, 0, d_{n,m_n+1}, \cdots, d_{nn}\}$ with $|d_{nj}| < 1$ for all $j =$

$m_n+1,\cdots,n$. As J_n is idempotent and $J_n\cdot\tilde{D}_n=\mathbf{0}$, we have $A_n^h=R_nJ_nR_n^{-1}+B_n^h$ where $B_n^h=R_n\tilde{D}_n^hR_n^{-1}$ for any $h=1,2,3,\cdots$.

Because $Y_{nt}=A_nY_{n,t-1}+S_n^{-1}(X_{nt}\beta_0+\mathbf{c}_{n0}+V_{nt})$, we can decompose Y_{nt} as $Y_{nt}=Y_{nt}^u+Y_{nt}^s$ where $Y_{nt}^u=R_nJ_nR_n^{-1}Y_{n,t-1}$ and $Y_{nt}^s=B_nY_{n,t-1}+S_n^{-1}(X_{nt}\beta_0+\mathbf{c}_{n0}+V_{nt})$. By using $B_nA_n=B_n^2$ and $B_nS_n^{-1}=S_n^{-1}B_n$, Y_{nt}^s can be written as an infinite sum of the past by recursive induction for any integer t:

$$\begin{aligned}
Y_{nt}^s &= B_nY_{n,t-1}+S_n^{-1}(X_{nt}\beta_0+\mathbf{c}_{n0}+V_{nt})\\
&= B_n[A_nY_{n,t-2}+S_n^{-1}(X_{n,t-1}\beta_0+\mathbf{c}_{n0}+V_{n,t-1})]+S_n^{-1}(X_{nt}\beta_0+\mathbf{c}_{n0}+V_{nt})\\
&= S_n^{-1}(\sum_{h=0}^{\infty}B_n^h)\mathbf{c}_{n0}+S_n^{-1}\sum_{h=0}^{\infty}B_n^h(X_{n,t-h}\beta_0+V_{n,t-h}).
\end{aligned}$$

For Y_{nt}^u, there are two versions which will be useful. By using $R_nJ_nR_n^{-1}A_n=R_nJ_nR_n^{-1}$ and $R_nJ_nR_n^{-1}S_n^{-1}=S_n^{-1}R_nJ_nR_n^{-1}$,

$$\begin{aligned}
Y_{nt}^u &= R_nJ_nR_n^{-1}Y_{n,t-1}\\
&= R_nJ_nR_n^{-1}[A_nY_{n,t-2}+S_n^{-1}(X_{n,t-1}\beta_0+\mathbf{c}_{n0}+V_{n,t-1})]\\
&= R_nJ_nR_n^{-1}Y_{n,t-2}+S_n^{-1}R_nJ_nR_n^{-1}(X_{n,t-1}\beta_0+\mathbf{c}_{n0}+V_{n,t-1})\\
&= R_nJ_nR_n^{-1}Y_{n,0}+(t-1)S_n^{-1}R_nJ_nR_n^{-1}\mathbf{c}_{n0}+S_n^{-1}R_nJ_nR_n^{-1}\sum_{h=1}^{t-1}(X_{nh}\beta_0+V_{nh}),
\end{aligned}$$

for $t=1,2,\cdots$, where $\sum_{h=1}^{0}$ is a zero as a convention. Another version is to expand Y_{nt}^u to $Y_{n,-1}$ as

$$Y_{nt}^u=R_nJ_nR_n^{-1}Y_{n,-1}+tS_n^{-1}R_nJ_nR_n^{-1}\mathbf{c}_{n0}+S_n^{-1}R_nJ_nR_n^{-1}\sum_{h=0}^{t-1}(X_{nh}\beta_0+V_{nh}), \tag{B.39}$$

for $t=0,1,2,\cdots$.

Using $\frac{1}{T}\sum_{t=1}^{T}(t-1)=\frac{1}{T}\sum_{t=1}^{T-1}t=\frac{T-1}{2}$, $\frac{1}{T}\sum_{t=1}^{T-1}\sum_{h=0}^{t-1}z_h=\frac{1}{T}\sum_{t=1}^{T-1}(T-t)z_{t-1}$ and $\frac{1}{T}\sum_{t=2}^{T}\sum_{h=1}^{t-1}z_h=\frac{1}{T}\sum_{t=1}^{T-1}(T-t)z_t$, it follows that

$$\begin{aligned}
\bar{Y}_{nT}^u &= \frac{1}{T}\sum_{t=1}^{T}Y_{nt}^u\\
&= R_nJ_nR_n^{-1}Y_{n0}+S_n^{-1}R_nJ_nR_n^{-1}\mathbf{c}_{n0}\frac{1}{T}\sum_{t=1}^{T}(t-1)+S_n^{-1}R_nJ_nR_n^{-1}\frac{1}{T}\sum_{t=2}^{T}\sum_{h=1}^{t-1}(X_{nh}\beta_0+V_{nh})\\
&= R_nJ_nR_n^{-1}Y_{n0}+S_n^{-1}R_nJ_nR_n^{-1}\mathbf{c}_{n0}(\frac{T-1}{2})+S_n^{-1}R_nJ_nR_n^{-1}\frac{1}{T}\sum_{t=1}^{T-1}(T-t)(X_{nt}\beta_0+V_{nt}),
\end{aligned}$$

and

$$\begin{aligned}
\bar{Y}_{nT,-1}^u &= \frac{1}{T}\sum_{t=0}^{T-1}Y_{nt}^u\\
&= R_nJ_nR_n^{-1}Y_{n,-1}+S_n^{-1}R_nJ_nR_n^{-1}\mathbf{c}_{n0}\frac{1}{T}\sum_{t=1}^{T-1}t+S_n^{-1}R_nJ_nR_n^{-1}\frac{1}{T}\sum_{t=1}^{T-1}\sum_{h=0}^{t-1}(X_{nh}\beta_0+V_{nh})\\
&= R_nJ_nR_n^{-1}Y_{n,-1}+S_n^{-1}R_nJ_nR_n^{-1}\mathbf{c}_{n0}(\frac{T-1}{2})+S_n^{-1}R_nJ_nR_n^{-1}\frac{1}{T}\sum_{t=0}^{T-2}(T-1-t)(X_{nt}\beta_0+V_{nt}).
\end{aligned}$$

Hence,

$$\begin{aligned}\tilde{\tilde{Y}}^u_{n,t-1} &= Y^u_{n,t-1} - \bar{Y}^u_{nT,-1} \\ &= S_n^{-1} R_n J_n R_n^{-1} \{ \mathbf{c}_{n0}[(t-1) - (\frac{T-1}{2})] + \sum_{h=0}^{t-2}(X_{nh}\beta_0 + V_{nh}) - \frac{1}{T}\sum_{h=0}^{T-2}(T-1-h)(X_{nh}\beta_0 + V_{nh})\} \\ &= \frac{R_n J_n R_n^{-1}}{(1-\lambda_0)} \{ \mathbf{c}_{n0}[(t-1) - (\frac{T-1}{2})] + \sum_{h=0}^{t-2}(X_{nh}\beta_0 + V_{nh}) - \frac{1}{T}\sum_{h=0}^{T-2}(T-1-h)(X_{nh}\beta_0 + V_{nh})\}\end{aligned}$$

because $S_n^{-1} R_n J_n R_n = \frac{1}{1-\lambda_0} R_n J_n R_n^{-1}$ from Proposition B.4 (3ii). For the stationary component,

$$\bar{Y}^s_{nT,-1} = \frac{1}{T}\sum_{t=0}^{T-1} Y^s_{nt} = S_n^{-1}(\sum_{h=0}^{\infty} B_n^h)\mathbf{c}_{n0} + S_n^{-1}\sum_{h=0}^{\infty} B_n^h \frac{1}{T}\sum_{t=0}^{T-1}(X_{n,t-h}\beta_0 + V_{n,t-h}),$$

and

$$\tilde{\tilde{Y}}^s_{n,t-1} = Y^s_{n,t-1} - \bar{Y}^s_{nT,-1} = S_n^{-1}\sum_{h=0}^{\infty} B_n^h[(X_{n,t-1-h} - \frac{1}{T}\sum_{t=0}^{T-1} X_{n,t-h})\beta_0 + (V_{n,t-1-h} - \frac{1}{T}\sum_{t=0}^{T-1} V_{n,t-h})]. \blacksquare$$

Proof for Proposition B.6: We use the result of Proposition B.4 to prove the result here. Conditions there are satisfied under Assumptions 5 and 6. That $W_n Y^u_{n,t-1} = Y^u_{n,t-1}$ follows from (B.1) of Proposition B.5 using $W_n R_n J_n R_n^{-1} = R_n J_n R_n^{-1}$ from Proposition B.4. For $G_n Y^u_{n,t-1} = \frac{1}{1-\lambda_0} Y^u_{n,t-1}$, this is so because (1) $S_n^{-1} Y^u_{n,t-1} = \frac{1}{1-\lambda_0} Y^u_{n,t-1}$ using $S_n^{-1} R_n J_n R_n^{-1} = \frac{1}{1-\lambda_0} R_n J_n R_n^{-1}$ and (2) $G_n = W_n S_n^{-1} = S_n^{-1} W_n$. Also, as $Z^u_{nt} = (Y^u_{n,t-1}, W_n Y^u_{n,t-1}, \mathbf{0}_{1\times k_x}) = Y^u_{n,t-1}(1, 1, \mathbf{0}_{1\times k_x})'$, we have $G_n Z^u_{nt}\delta_0 = Y^u_{n,t-1}$. This follows because $G_n Z^u_{nt}\delta_0 = G_n Y^u_{n,t-1}(\gamma_0 + \rho_0)$ and $\gamma_0 + \rho_0 = 1 - \lambda_0$. $\blacksquare$

Proof for Lemma B.7: See Lemma A.2 and A.4 in Yu, de Jong and Lee (2006). $\blacksquare$

Proof for Lemma B.8:

Equation (B.7): Let $\rho(\mathcal{B}_n)$ be its spectral radius (the largest eigenvalue in absolute value) and $||\cdot||$ be a matrix norm. It is known from matrix theory that $\rho(\mathcal{B}_n) \leq ||\mathcal{B}_n||$ (see Horn and Johnson (1985)). Taking $\|\cdot\|$ to be either $\|\cdot\|_\infty$ or $\|\cdot\|_1$, it follows that $\{\|\mathcal{B}_n\|\}$ is bounded because $\mathcal{B}_n$ is row sum and column sum bounded. With the above settings,

$$\begin{aligned}&\left|\frac{1}{nT}\sum_{t=1}^{T}(\mathbf{c}_{n0}\tilde{t} + \tilde{\mathbb{X}}_{nt}\beta_0)'\mathcal{B}_n(\mathbf{c}_{n0}\tilde{t} + \tilde{\mathbb{X}}_{nt}\beta_0)\right| \\ \leq\ & \rho(\mathcal{B}_n)\cdot\left|\frac{1}{nT}\sum_{t=1}^{T}(\mathbf{c}_{n0}\tilde{t} + \tilde{\mathbb{X}}_{nt}\beta_0)'(\mathbf{c}_{n0}\tilde{t} + \tilde{\mathbb{X}}_{nt}\beta_0)\right| \\ \leq\ & ||\mathcal{B}_n||\cdot\left|\frac{1}{nT}\sum_{t=1}^{T}(\mathbf{c}_{n0}\tilde{t} + \tilde{\mathbb{X}}_{nt}\beta_0)'(\mathbf{c}_{n0}\tilde{t} + \tilde{\mathbb{X}}_{nt}\beta_0)\right| \\ =\ & ||\mathcal{B}_n||\cdot\frac{1}{nT}\sum_{t=1}^{T}(\mathbf{c}'_{n0}\mathbf{c}_{n0}\tilde{t}^2 + 2\mathbf{c}'_{n0}\tilde{\mathbb{X}}_{nt}\beta_0\tilde{t} + (\tilde{\mathbb{X}}_{nt}\beta_0)'(\tilde{\mathbb{X}}_{nt}\beta_0)).\end{aligned}$$

Because $\frac{1}{T}\sum_{t=1}^{T} t^2 = \frac{1}{6}(T+1)(2T+1) = O(T^2)$, $\left|\tilde{\mathbb{X}}_{nt}\beta_0\right| = \left|\sum_{h=0}^{t-1}\tilde{X}_{nh}\beta_0\right| \leq t\cdot\sup_{n,t}\left|\tilde{X}_{nt}\beta_0\right|$, $\tilde{X}_{nt}$ is bounded uniformly in all n and t, and elements of $\mathbf{c}_{n0}$ are also uniformly bounded, we have the result that $\left|\frac{1}{nT}\sum_{t=1}^{T}(\mathbf{c}_{n0}\tilde{t}+\tilde{\mathbb{X}}_{nt}\beta_0)'\mathcal{B}_n(\mathbf{c}_{n0}\tilde{t}+\tilde{\mathbb{X}}_{nt}\beta_0)\right| = O(T^2)$.

Equation (B.8): As $E\xi_{nt}\xi_{ns}' = \sigma_0^2\min\{t,s\}I_n$, we have

$$
\begin{aligned}
& Var(\frac{1}{nT}\sum_{t=1}^{T}(\mathbf{c}_{n0}\tilde{t}+\tilde{\mathbb{X}}_{nt}\beta_0)'\mathcal{B}_n\tilde{\xi}_{nt}) = Var(\frac{1}{nT}\sum_{t=1}^{T}(\mathbf{c}_{n0}\tilde{t}+\tilde{\mathbb{X}}_{nt}\beta_0)'\mathcal{B}_n\xi_{nt}) \\
= & \frac{1}{n^2T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}(\mathbf{c}_{n0}\tilde{t}+\tilde{\mathbb{X}}_{nt}\beta_0)'\mathcal{B}_nE\xi_{nt}\xi_{ns}'\mathcal{B}_n'(\mathbf{c}_{n0}\tilde{s}+\tilde{\mathbb{X}}_{ns}\beta_0) \\
= & \frac{\sigma_0^2}{n^2T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\min\{s,t\}(\mathbf{c}_{n0}\tilde{t}+\tilde{\mathbb{X}}_{nt}\beta_0)'\mathcal{B}_n^2(\mathbf{c}_{n0}\tilde{s}+\tilde{\mathbb{X}}_{ns}\beta_0) \\
\leq & \frac{\sigma_0^2}{n^2T^2}\left(\sum_{t=1}^{T}t(\mathbf{c}_{n0}\tilde{t}+\tilde{\mathbb{X}}_{nt}\beta_0)\right)\mathcal{B}_n^2\left(\sum_{s=1}^{T}\left(\mathbf{c}_{n0}\tilde{s}+\tilde{\mathbb{X}}_{ns}\beta_0\right)\right) \\
= & \frac{\sigma_0^2T^3}{n}\frac{1}{n}\left(\frac{1}{T^3}\sum_{t=1}^{T}t(\mathbf{c}_{n0}\tilde{t}+\tilde{\mathbb{X}}_{nt}\beta_0)\right)\mathcal{B}_n^2\left(\frac{1}{T^2}\sum_{s=1}^{T}\left(\mathbf{c}_{n0}\tilde{s}+\tilde{\mathbb{X}}_{ns}\beta_0\right)\right) = O(\frac{T^3}{n})
\end{aligned}
$$

by the uniform boundedness elements of $\mathbf{c}_{n0}$ and X_{nt}, and the uniform boundedness of $\mathcal{B}_n^2$ in row and column sums. The result follows.

Equation (B.9): We have $\frac{1}{nT}\sum_{t=1}^{T}\tilde{\xi}_{nt}'\mathcal{B}_n\tilde{\xi}_{nt} = \frac{1}{nT}\sum_{t=1}^{T}(\xi_{nt}'\mathcal{B}_n\xi_{nt}) - \frac{1}{n}\bar{\xi}_{nT}'\mathcal{B}_n\bar{\xi}_{nT}$.

For the first part, $E(\frac{1}{nT}\sum_{t=1}^{T}(\xi_{nt}'\mathcal{B}_n\xi_{nt})) = \sigma_0^2 tr(\mathcal{B}_n)\left(\frac{1}{nT}\sum_{t=1}^{T}t\right) = O(T)$ and $Var(\frac{1}{nT}\sum_{t=1}^{T}\xi_{nt}'\mathcal{B}_n\xi_{nt}) = \frac{1}{n^2T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}Cov(\xi_{nt}'\mathcal{B}_n\xi_{nt},\xi_{ns}'\mathcal{B}_n\xi_{ns})$. Using Lemma B.7 for covariance between $\mathbb{U}_{nt}'\mathbb{W}_{nt}$ and $\mathbb{U}_{nt}'\mathbb{W}_{nt}$ (in our case here, $\mathbb{U}_{nt} = \sum_{h=1}^{\infty}P_{nt,h}V_{n,t+1-h}$ and $\mathbb{W}_{nt} = \sum_{h=1}^{\infty}Q_{nt,h}V_{n,t+1-h}$, where $P_{nt,h} = I_n$ and $Q_{nt,h} = \mathcal{B}_n$ for $h \leq t$, and $P_{nt,h} = Q_{nt,h} = 0$ for $h > t$.), we have $Var(\frac{1}{nT}\sum_{t=1}^{T}\xi_{nt}'\mathcal{B}_n\xi_{nt}) = \frac{T^2}{n}$.

For the second part, $E(\frac{1}{n}\bar{\xi}_{nT}'\mathcal{B}_n\bar{\xi}_{nT}) = E(\frac{1}{n}\ddot{\mathbb{U}}_{nT}'\ddot{\mathbb{W}}_{nT})$ where $\ddot{\mathbb{U}}_{nT} = \sum_{h=1}^{\infty}\ddot{P}_{nT,h}V_{n,T-h}$ and $\ddot{\mathbb{W}}_{nT} = \sum_{h=1}^{\infty}\ddot{Q}_{nT,h}V_{n,T-h}$ with

$$
\ddot{P}_{nT,h} = \begin{cases} I_n\frac{h}{T} & \text{for } h \leq T \\ 0 & \text{for } h > T \end{cases} \quad \text{and } \ddot{Q}_{nT,h} = \begin{cases} \mathcal{B}_n\frac{h}{T} & \text{for } h \leq T \\ 0 & \text{for } h > T \end{cases}. \tag{B.40}
$$

Then, using Lemma B.7, $E(\frac{1}{n}\bar{\xi}_{nT}'\mathcal{B}_n\bar{\xi}_{nT}) = O(T)$ and $Var(\frac{1}{n}\bar{\xi}_{nT}'\mathcal{B}_n\bar{\xi}_{nT}) = \frac{1}{n^2}Cov(\ddot{\mathbb{U}}_{nT}'\ddot{\mathbb{W}}_{nT},\ddot{\mathbb{U}}_{nT}'\ddot{\mathbb{W}}_{nT}) = O(\frac{T^2}{n})$ because $\sum_{h=1}^{\infty}\ddot{P}_{nT,h}\ddot{P}_{nT,h}' = \sum_{h=1}^{T}(\frac{h}{T})^2I_n$, $\sum_{h=1}^{\infty}\ddot{Q}_{nT,h}'\ddot{P}_{nT,h} = \sum_{h=1}^{T}(\frac{h}{T})^2\mathcal{B}_n'$, $\sum_{h=1}^{\infty}\ddot{Q}_{nT,h}\ddot{Q}_{nT,h}' = \sum_{h=1}^{T}(\frac{h}{T})^2\mathcal{B}_n\mathcal{B}_n'$ and $\sum_{h=1}^{T}h^2 = O(T^3)$.

Equation (B.10): Because of the uniform boundedness of $\mathbf{c}_{n0}$, $\tilde{D}_{nt}$ and $\mathcal{B}_n$, there exist finite constants c_1

and c_2 such that

$$\begin{aligned}
& \left|\frac{1}{nT}\sum_{t=1}^{T}(\mathbf{c}_{n0}\tilde{t}+\tilde{\mathbb{X}}_{nt}\beta_0)'\mathcal{B}_n\tilde{D}_{nt}\right| \\
\leq & \frac{1}{T}\sum_{t=1}^{T}\left|\frac{\mathbf{c}_{n0}'\mathcal{B}_n\tilde{D}_{nt}}{n}\right|\cdot\left|\tilde{t}\right|+\frac{1}{T}\sum_{t=1}^{T}\left|\frac{(\tilde{\mathbb{X}}_{nt}\beta_0)'\mathcal{B}_n\tilde{D}_{nt}}{n}\right|\leq\frac{c_1}{T}\sum_{t=1}^{T}|\tilde{t}|+\frac{c_2}{T}\sum_{t=1}^{T}|\tilde{t}|=O(T).
\end{aligned}$$

Equation (B.11): $\frac{1}{nT}\sum_{t=1}^{T}(\mathbf{c}_{n0}\tilde{t}+\tilde{\mathbb{X}}_{nt}\beta_0)'\tilde{\mathbb{U}}_{nt}=\frac{1}{n}\sum_{t=1}^{T}(\mathbf{c}_{n0}\frac{\tilde{t}}{T}+\frac{1}{T}\tilde{\mathbb{X}}_{nt}\beta_0)'\tilde{\mathbb{U}}_{nt}=T[\frac{1}{nT}\sum_{t=1}^{T}(\mathbf{c}_{n0}\frac{\tilde{t}}{T}+\frac{1}{T}\tilde{\mathbb{X}}_{nt}\beta_0)'\tilde{\mathbb{U}}_{nt}]$. As $\frac{\tilde{t}}{T}$ and $\frac{1}{T}\tilde{\mathbb{X}}_{nt}\beta_0$ are bounded, using Theorem A.8 in Yu, de Jong and Lee (2006), $\frac{1}{nT}\sum_{t=1}^{T}(\mathbf{c}_{n0}\frac{\tilde{t}}{T}+\frac{1}{T}\tilde{\mathbb{X}}_{nt}\beta_0)'\tilde{\mathbb{U}}_{nt}=O_p(\frac{1}{\sqrt{nT}})$. Hence, $\frac{1}{nT}\sum_{t=1}^{T}(\mathbf{c}_{n0}\tilde{t}+\tilde{\mathbb{X}}_{nt}\beta_0)'\tilde{\mathbb{U}}_{nt}=O_p(\sqrt{\frac{T}{n}})$.

Equation (B.12): As $E\xi_{nt}\xi_{ns}'=\sigma_0^2\min\{t,s\}I_n$, we have

$$\begin{aligned}
& Var(\frac{1}{nT}\sum_{t=1}^{T}\tilde{\xi}_{nt}'\mathcal{B}_n\tilde{D}_{nt}) \\
= & Var(\frac{1}{nT}\sum_{t=1}^{T}(\xi_{nt}'\mathcal{B}_n\tilde{D}_{nt})=\frac{1}{n^2T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}(\tilde{D}_{nt}'\mathcal{B}_n'(E(\xi_{nt}\xi_{ns}'))\mathcal{B}_n\tilde{D}_{ns}) \\
= & \frac{\sigma_0^2}{n^2T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\min\{s,t\}(\tilde{D}_{nt}'\mathcal{B}_n'\mathcal{B}_n\tilde{D}_{ns})=O(\frac{T}{n}).
\end{aligned}$$

Equation (B.13): We have $\frac{1}{nT}\sum_{t=1}^{T}\tilde{\xi}_{nt}'\tilde{\mathbb{U}}_{nt}=\frac{1}{nT}\sum_{t=1}^{T}\tilde{\xi}_{nt}'\tilde{\mathbb{U}}_{nt}-\frac{1}{n}\bar{\xi}_{nT}'\bar{\mathbb{U}}_{nT}$.

For the first part, using Lemma B.7,

$$\begin{aligned}
& E(\frac{1}{nT}\sum_{t=1}^{T}\xi_{nt}'\mathbb{U}_{nt})=\frac{1}{nT}\sum_{t=1}^{T}\left(\sigma_0^2tr(\sum_{h=1}^{t}P_{nt,h})\right) \\
= & \sigma_0^2\frac{1}{nT}\left(tr(\sum_{t=1}^{T}\sum_{h=1}^{t}P_{nt,h})\right)=O(1),
\end{aligned}$$

because $\sum_{h=1}^{\infty}abs(P_{nt,h})$ is row sum and column sum bounded. Also,

$$Var(\frac{1}{nT}\sum_{t=1}^{T}\xi_{nt}'\mathbb{U}_{nt})=\frac{1}{n^2T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}Cov(\xi_{nt}'\mathbb{U}_{nt},\xi_{ns}'\mathbb{U}_{ns})=O(\frac{T}{n}).$$

This is so as follows. As $\xi_{nt}=\sum_{h=1}^{\infty}Q_{nt,h}V_{n,t-h}$ where $Q_{nt,h}=\begin{cases} I_n & \text{for } h=1,2,\cdots,t \\ 0 & \text{for } h\geq t+1\end{cases}$, we have $Var(\frac{1}{nT}\sum_{t=1}^{T}\xi_{nt}'\mathbb{U}_{nt})=O(\frac{T}{n})$ using Lemma B.7 because the leading factor $\sum_{h=1}^{\infty}Q_{ns,h}Q_{ns,h}'=\sum_{h=1}^{s}I_n=s\cdot I_n$ and $\sum_{h=1}^{T}\sum_{s=1}^{t}s=O(T^3)$.

For the second part, $E(\frac{1}{n}\bar{\xi}_{nT}'\bar{\mathbb{U}}_{nT})=\frac{\sigma_0^2}{nT}tr\left(\sum_{h=1}^{T}h\cdot\bar{P}_{nT,h}\right)$ where $\bar{P}_{nT,h}$ is specified in (B.5). So, $E(\frac{1}{n}\bar{\xi}_{nT}'\bar{\mathbb{U}}_{nT})=O(1)$. Also, $Var(\frac{1}{n}\bar{\xi}_{nT}'\bar{\mathbb{U}}_{nT})=\frac{1}{n^2}Cov(\ddot{\mathbb{W}}_{nT}'\bar{\mathbb{U}}_{nT},\ddot{\mathbb{W}}_{nT}'\bar{\mathbb{U}}_{nT})$ where $\bar{\mathbb{U}}_{nT}=\sum_{h=1}^{\infty}\bar{P}_{nT,h}V_{n,t+1-h}$ and $\ddot{\mathbb{W}}_{nT}=\sum_{h=1}^{\infty}\ddot{P}_{nT,h}V_{n,t+1-h}$ with $\bar{P}_{nT,h}$ specified in (B.5) and $\ddot{P}_{nT,h}$ specified in (B.40). Then, using Lemma B.7, we have $Var(\frac{1}{n}\bar{\xi}_{nT}'\bar{\mathbb{U}}_{nT})=O(\frac{T}{n})$.

Equation (B.14): This is Theorem A.7 in Yu, de Jong and Lee (2006).

Proof for Lemma B.9

Equation (B.15): Denote $\rho(\mathcal{B}_n\mathcal{B}_n'\mathcal{B}_n')$ be the spectral radius of $\mathcal{B}_n\mathcal{B}_n'$. Then,

$$\begin{aligned}
& Var(\frac{1}{nT}\sum\nolimits_{t=1}^{T}(\mathbf{c}_{n0}\tilde{t}_{-1}+\tilde{\mathbb{X}}_{n,t-1}\beta_0)'\mathcal{B}_nV_{nt}) \\
= \quad & \frac{\sigma_0^2}{n^2T^2}\sum\nolimits_{t=1}^{T}(\mathbf{c}_{n0}\tilde{t}_{-1}+\tilde{\mathbb{X}}_{n,t-1}\beta_0)'\mathcal{B}_n\mathcal{B}_n'(\mathbf{c}_{n0}\tilde{t}_{-1}+\tilde{\mathbb{X}}_{n,t-1}\beta_0) \\
\leq \quad & \frac{\sigma_0^2}{n^2T^2}\cdot\rho(\mathcal{B}_n\mathcal{B}_n')\cdot\left|\sum\nolimits_{t=1}^{T}(\mathbf{c}_{n0}\tilde{t}_{-1}+\tilde{\mathbb{X}}_{n,t-1}\beta_0)'(\mathbf{c}_{n0}\tilde{t}_{-1}+\tilde{\mathbb{X}}_{n,t-1}\beta_0)\right| \\
\leq \quad & \frac{\sigma_0^2}{n^2T^2}\cdot\|\mathcal{B}_n\mathcal{B}_n'\|_\infty\cdot\left|\sum\nolimits_{t=1}^{T}(\mathbf{c}_{n0}\tilde{t}_{-1}+\tilde{\mathbb{X}}_{n,t-1}\beta_0)'(\mathbf{c}_{n0}\tilde{t}_{-1}+\tilde{\mathbb{X}}_{n,t-1}\beta_0)\right| = O(\frac{T}{n}),
\end{aligned}$$

because $\sum_{t=1}^{T}t^2=O(T^3)$ in the leading term.

Equation (B.16):

$$\begin{aligned}
& Var\left(\frac{1}{nT}\sum\nolimits_{t=1}^{T}\xi_{n,t-1}'\mathcal{B}_nV_{nt}\right) \\
= \quad & \frac{\sigma_0^2}{n^2T^2}\sum\nolimits_{t=1}^{T}E(\xi_{n,t-1}'\mathcal{B}_n\mathcal{B}_n'\xi_{n,t-1}) = \frac{\sigma_0^2}{n^2T^2}\sum\nolimits_{t=1}^{T}tr\left(E(\mathcal{B}_n\mathcal{B}_n'\xi_{n,t-1}\xi_{n,t-1}')\right) \\
= \quad & \frac{\sigma_0^4}{n^2T^2}\cdot tr(\mathcal{B}_n\mathcal{B}_n')\cdot\sum\nolimits_{t=1}^{T}(t-1) = O(\frac{1}{n}).
\end{aligned}$$

Then, $\frac{1}{nT}\sum_{t=1}^{T}\xi_{n,t-1}'\mathcal{B}_nV_{nt}=O_p\left(\frac{1}{\sqrt{n}}\right)$ with mean zero.

Equation (B.17): As $\bar{\xi}_{n,T-1}=\frac{1}{T}\sum_{t=0}^{T-1}\xi_{nt}=\frac{1}{T}\sum_{t=0}^{T-1}\sum_{h=0}^{t-1}V_{nh}=\frac{1}{T}\sum_{t=1}^{T-1}(T-t)V_{n,t-1}$, we have

$$E(\bar{\xi}_{n,T-1}'\mathcal{B}_n\bar{V}_{nT})=\frac{1}{T^2}E[(\sum\nolimits_{t=1}^{T-1}(T-t)V_{n,t-1})'\mathcal{B}_n\sum\nolimits_{t=1}^{T}V_{nt}]=\sigma_0^2\frac{(T-1)(T-2)}{2T^2}tr(\mathcal{B}_n). \tag{B.41}$$

For the special case where $\mathcal{B}_n=M_n$, $E(\bar{\xi}_{n,T-1}'M_n\bar{V}_{nT})=\sigma_0^2\frac{(T-1)(T-2)}{2T^2}tr(M_n)=\sigma_0^2\frac{(T-1)(T-2)}{2T^2}m_n$ because $tr(M_n)=tr(R_nJ_nR_n^{-1})=tr(J_n)=m_n$. Also,

$$\begin{aligned}
& Var(\bar{\xi}_{n,T-1}'\mathcal{B}_n\bar{V}_{nT}) \\
= \quad & Var(\frac{1}{T^2}(\sum\nolimits_{t=1}^{T-1}(T-t)V_{n,t-1})'\mathcal{B}_n\sum\nolimits_{t=1}^{T}V_{nt})=\frac{1}{T^4}Var(\sum\nolimits_{t=1}^{T}V_{nt}'\mathcal{B}_n'(\sum\nolimits_{t=1}^{T-1}(T-t)V_{n,t-1})) \\
= \quad & \frac{1}{T^4}Var(\mathbb{U}_{nT,-1}'\mathbb{W}_{nT,-1}),
\end{aligned}$$

where $\mathbb{U}_{nT,-1}=\sum_{h=1}^{\infty}P_{nt,h}V_{n,t+1-h}$ with $P_{nt,h}=\begin{cases} I_n & \text{for } h\leq T \\ 0 & \text{for } h>T\end{cases}$ and $\mathbb{W}_{nT,-1}=\sum_{h=1}^{\infty}Q_{nt,h}V_{n,t+1-h}$

with $Q_{nt,h} = \begin{cases} \mathcal{B}_n' \cdot (h-1) & \text{for } h \le T \\ 0 & \text{for } h > T \end{cases}$. From Lemma B.7,

$$
\begin{aligned}
& Var(\bar{\xi}_{n,T-1}' \mathcal{B}_n \bar{V}_{nT}) \\
= \; & \frac{1}{T^4}\sigma_0^4 tr\left[I_n \cdot T \times \mathcal{B}_n'\mathcal{B}_n \cdot \sum\nolimits_{h=1}^T (h-1)^2 + \sum\nolimits_{h=1}^T (h-1) \cdot \sum\nolimits_{h=1}^T (h-1)\mathcal{B}_n'\mathcal{B}_n'\right] \\
& + \frac{1}{T^4}(\mu_4 - 3\sigma_0^4) \sum\nolimits_{h=1}^T (h-1)^2 \cdot \sum\nolimits_{i=1}^n (\mathcal{B}_n)_{ii}(\mathcal{B}_n)_{ii} \\
= \; & \frac{1}{T^4}\sigma_0^4 \left[tr(\mathcal{B}_n'\mathcal{B}_n) \cdot T \cdot \sum\nolimits_{h=1}^T (h-1)^2 + (tr(\mathcal{B}_n'\mathcal{B}_n') \cdot \frac{(T-1)^2 T^2}{4}\right] \\
& + \sum\nolimits_{i=1}^n (\mathcal{B}_n)_{ii}(\mathcal{B}_n)_{ii} \cdot \frac{1}{T^4}(\mu_4 - 3\sigma_0^4) \sum\nolimits_{h=1}^T (h-1)^2 = O(n).
\end{aligned}
$$

So, $E(\bar{\xi}_{n,T-1}' \mathcal{B}_n \bar{V}_{nT}) = O(n)$ and $Var(\bar{\xi}_{n,T-1}' \mathcal{B}_n \bar{V}_{nT}) = O(n)$.

Equation (B.18): This is implied by Theorem A.11 in Yu, de Jong and Lee (2006). ■

Proof for Lemma B.10: Using Lemma B.8 and that Y_{nt} has those components, we have the result.

For (B.20), we use (B.7), (B.8) and (B.9) in Lemma B.8. For (B.21), we use (B.10), (B.11), (B.12) and (B.13) in Lemma B.8. For (B.22), it is implied by Lemma B.1 in Yu, de Jong and Lee (2006). ■

Proof for Lemma B.11: Using Lemma B.9 and that Y_{nt} has those components, we have the result.

For (B.23), we use (B.15), (B.16) and (B.17) in Lemma B.9. For (B.24), it is in Lemma B.1 in Yu, de Jong and Lee (2006). ■

Proof for Proposition B.12: The form of the inverse of H_T can be checked by direct multiplication of H_T with the right hand side matrix expression (of H_T^{-1}), which will result in an identity matrix. The explicit expression of H_T^{-1} is complicated. But it can be derived by the following motivations. Define $Q_T = I_m + T g_T d_T'$ and $R_T = I_m + T Q_T^{-1} h_T b_T'$. It follows, by construction, that $H_T = Q_T R_T$. If both R_T and Q_T are invertible, then H_T must be invertible. By the familiar pattern of Q_T, its inverse will have the form (see Dhrymes (1978)) $Q_T^{-1} = I_m - \frac{T}{1+T d_T' g_T} g_T d_T'$, and also, the inverse of R_T has the form $R_T^{-1} = I_m - \frac{T}{1+T b_T' Q_T^{-1} h_T} Q_T^{-1} h_T b_T'$. The final expression of $H_T^{-1} = R_T^{-1} Q_T^{-1}$ can be derived by exploring the explicit expressions of Q_T^{-1}, R_T^{-1} and their multiplication. ■

Proof for Proposition B.13: The following proof is for the case that K_T is nonrandom. After we get the result, it can be extended to the case that K_T is random as long as A_T is nonsingular with probability 1.

Using the notations in the Proposition B.12, $K_T = P_T H_T$ and $K_T^{-1} = H_T^{-1} P_T^{-1}$, where H_T^{-1} is in (B.25) and

$$P_T^{-1} = (T^2 c_T c_T' + A_T)^{-1} = A_T^{-1} - \frac{T^2}{1 + T^2 c_T' A_T^{-1} c_T} A_T^{-1} c_T c_T' A_T^{-1}. \tag{B.42}$$

Furthermore, denote $h_T = P_T^{-1} d_T$ and $g_T = P_T^{-1} b_T$, where h_T and g_T are in Proposition B.12. As c_T is proportional to b_T and the explicit inverse formula of P_T involves c_T, $g_T = P_T^{-1} b_T = \frac{1}{1+T^2 c_T' A_T^{-1} c_T} A_T^{-1} b_T$.

This implies the following scalar values: $b'_T g_T = \frac{b'_T A_T^{-1} b_T}{1+T^2 c'_T A_T^{-1} c_T}$, $d'_T g_T = b'_T h_T = \frac{b'_T A_T^{-1} d_T}{1+T^2 c'_T A_T^{-1} c_T}$, and $d'_T h_T = d'_T A_T^{-1} d_T - \frac{T^2 (c'_T A_T^{-1} d_T)^2}{1+T^2 c'_T A_T^{-1} c_T}$. In terms of orders of magnitude, we have $d'_T h_T = O(1)$, $b'_T g_T = O(\frac{1}{T^2})$, $d'_T g_T = O(\frac{1}{T^2})$, and $b'_T h_T = O(\frac{1}{T^2})$.

With these, one can evaluate Δ_T in H_T^{-1} and its limit. The two terms of Δ_T are $T(b'_T h_T + d'_T g_T) = 2\frac{b'_T A_T^{-1} d_T}{T(T^{-2} + c'_T A_T^{-1} c_T)} = O(\frac{1}{T})$, and

$$
\begin{aligned}
& T^2 (b'_T g_T \cdot d'_T h_T - b'_T h_T \cdot d'_T g_T) \\
= \quad & (\frac{b'_T A_T^{-1} b_T}{T^{-2} + c'_T A_T^{-1} c_T})(d'_T A_T^{-1} d_T - \frac{(c'_T A_T^{-1} d_T)^2}{T^{-2} + c'_T A_T^{-1} c_T}) - (\frac{T b'_T A_T^{-1} d_T}{1 + T^2 c'_T A_T^{-1} c_T})^2 \\
\longrightarrow \quad & \frac{1}{\omega^2}(d' A^{-1} d - \frac{(c' A^{-1} d)^2}{c' A^{-1} c}).
\end{aligned}
$$

Hence,

$$
\Delta = \lim_{T\to\infty} \Delta_T = 1 - \frac{1}{\omega^2}\left(d' A^{-1} d - \frac{(d' A^{-1} b)^2}{b' A^{-1} b} \right).
$$

As $K_T^{-1} = H_T^{-1} P_T^{-1} = (I_m - \frac{T B_T}{\Delta_T}) P_T^{-1}$, it remains to consider the limiting behavior of $T B_T \cdot P_T^{-1}$ where B_T is in (B.27) and P_T^{-1} is in (B.42). As $b'_T P_T^{-1} = \frac{1}{1+T^2 c'_T A_T^{-1} c_T} b'_T A_T^{-1}$ and $d'_T P_T^{-1} = (d_T - \frac{T^2 c'_T A_T^{-1} d_T}{1+T^2 c'_T A_T^{-1} c_T} c_T)' A_T^{-1}$, these imply the following matrices

$$
g_T b'_T P_T^{-1} = P_T^{-1} b_T b'_T P_T^{-1} = \frac{1}{(1 + T^2 c'_T A_T^{-1} c_T)^2} A_T^{-1} b_T b'_T A_T^{-1},
$$

$$
h_T d'_T P_T^{-1} = P_T^{-1} d_T d'_T P_T^{-1} = A_T^{-1}(d_T - \frac{T^2 c'_T A_T^{-1} d_T}{1 + T^2 c'_T A_T^{-1} c_T} c_T)(d_T - \frac{T^2 c'_T A_T^{-1} d_T}{1 + T^2 c'_T A_T^{-1} c_T} c_T)' A_T^{-1},
$$

and

$$
g_T d'_T P_T^{-1} = (h_T b'_T P_T^{-1})' = P_T^{-1} b_T d'_T P_T^{-1} = \frac{1}{(1 + T^2 c'_T A_T^{-1} c_T)} A_T^{-1} b_T (d_T - \frac{T^2 c_T A_T^{-1} d_T}{1 + T^2 c'_T A_T^{-1} c_T} c_T)' A_T^{-1}.
$$

In terms of orders of magnitude, $h_T d'_T P_T^{-1} = O(1)$, $h_T b'_T P_T^{-1} = O(\frac{1}{T^2})$, $g_T d'_T P_T^{-1} = O(\frac{1}{T^2})$ and $g_T b'_T P_T^{-1} = O(\frac{1}{T^4})$. Therefore, for $T B_T \cdot P_T^{-1}$ where B_T is in (B.27) and P_T^{-1} is in (B.42), we have $T(h_T b'_T + g_T d'_T) P_T^{-1} = O(\frac{1}{T})$ and

$$
\begin{aligned}
& T^2 [(d'_T h_T) g_T b'_T + (b'_T g_T) h_T d'_T - (d'_T g_T) h_T b'_T - (b'_T h_T) g_T d'_T] P_T^{-1} \\
= \quad & T^2 (b'_T g_T) h_T d'_T P_T^{-1} + O(\frac{1}{T^2}) \\
= \quad & \frac{T^2 b'_T A_T^{-1} b_T}{1 + T^2 c'_T A_T^{-1} c_T} A_T^{-1}(d_T - \frac{T^2 d'_T A_T^{-1} c_T}{1 + T^2 c'_T A_T^{-1} c_T} c_T)(d_T - \frac{T^2 d'_T A_T^{-1} c_T}{1 + T^2 c'_T A_T^{-1} c_T} c_T)' A_T^{-1} + O(\frac{1}{T^2}) \\
\longrightarrow \quad & \frac{1}{\omega^2} A^{-1}(d - \frac{d' A^{-1} b}{b' A^{-1} b} b)(d - \frac{d' A^{-1} b}{b' A^{-1} b} b)' A^{-1},
\end{aligned}
$$

because $c = \omega b$, which is the limit of $(-T B_T P_T^{-1})$.

Thus, $K_T^{-1} = (I_m - \frac{T B_T}{\Delta_T}) P_T^{-1} = P_T^{-1} - \frac{1}{\Delta_T} T B_T P_T^{-1}$ converges to L_k, where

$$
L_k = (A^{-1} - \frac{1}{b' A^{-1} b} A^{-1} b b' A^{-1}) + \frac{1}{\omega^2 \Delta} A^{-1}(d - \frac{d' A^{-1} b}{b' A^{-1} b} b)(d - \frac{d' A^{-1} b}{b' A^{-1} b} b)' A^{-1}.
$$

For (b) and (c), we have $K_T^{-1} = (P_T H_T)^{-1}$ where $P_T^{-1} = A_T^{-1} - \frac{T^2}{1+T^2 c_T' A_T^{-1} c_T} A_T^{-1} c_T c_T' A_T^{-1}$ and $H_T^{-1} = I_m - \frac{T}{\Delta_T} B_T$ with $\Delta_T = 1 + T(b_T' h_T + d_T' g_T) - T^2 |(b_T, d_T)'(g_T, h_T)|$ and $B_T = (h_T b_T' + g_T d_T') - T[(d_T' h_T) g_T b_T' + (b_T' g_T) h_T d_T' - (d_T' g_T) h_T b_T' - (b_T' h_T) g_T d_T']$. Hence,

$$
\begin{aligned}
K_T^{-1} &= P_T^{-1} - \frac{T}{\Delta_T} B_T P_T^{-1} \qquad \text{(B.43)} \\
&= P_T^{-1} - \frac{T}{\Delta_T} \left((h_T b_T' + g_T d_T') - T[(d_T' h_T) g_T b_T' + (b_T' g_T) h_T d_T' - (d_T' g_T) h_T b_T' - (b_T' h_T) g_T d_T'] \right) P_T^{-1} \\
&= P_T^{-1} - \frac{T}{\Delta_T} (h_T b_T' + g_T d_T') P_T^{-1} + \frac{T^2}{\Delta_T} [(d_T' h_T) g_T b_T' + (b_T' g_T) h_T d_T' - (d_T' g_T) h_T b_T' - (b_T' h_T) g_T d_T'] P_T^{-1} \\
&= P_T^{-1} - \frac{T}{\Delta_T} (P_T^{-1} d_T b_T' P_T^{-1} + P_T^{-1} b_T d_T' P_T^{-1}) \\
&\quad + \frac{T^2}{\Delta_T} [(d_T' P_T^{-1} d_T) P_T^{-1} b_T b_T' P_T^{-1} + (b_T' P_T^{-1} b_T) P_T^{-1} d_T d_T' P_T^{-1}] \\
&\quad - \frac{T^2}{\Delta_T} [(d_T' P_T^{-1} b_T) P_T^{-1} d_T b_T' P_T^{-1} + (b_T' P_T^{-1} d_T) P_T^{-1} b_T d_T' P_T^{-1}].
\end{aligned}
$$

As

$$
\begin{aligned}
c_T' P_T^{-1} c_T &= \frac{c_T' A_T^{-1} c_T}{1 + T^2 c_T' A_T^{-1} c_T} = O(T^{-2}),\ c_T' P_T^{-1} d_T = \frac{c_T' A_T^{-1} d_T}{1 + T^2 c_T' A_T^{-1} c_T} = O(T^{-2}), \qquad \text{(B.44a)} \\
d_T' P_T^{-1} d_T &= d_T' A_T^{-1} d_T - \frac{T^2 (c_T' A_T^{-1} d_T)^2}{1 + T^2 c_T' A_T^{-1} c_T} = O(1) \qquad \text{(B.44b)}
\end{aligned}
$$

and b_T is proportional to c_T, we have

$$
\begin{aligned}
c_T' K_T^{-1} c_T &= c_T' P_T^{-1} c_T - \frac{2T}{\Delta_T} (c_T' P_T^{-1} d_T b_T' P_T^{-1} c_T) \\
&\quad + \frac{T^2}{\Delta_T} [(d_T' P_T^{-1} d_T) c_T' P_T^{-1} b_T b_T' P_T^{-1} c_T + (b_T' P_T^{-1} b_T) c_T' P_T^{-1} d_T d_T' P_T^{-1} c_T] \\
&\quad + \frac{2T^2}{\Delta_T} [-(d_T' P_T^{-1} b_T) c_T' P_T^{-1} d_T b_T' P_T^{-1} c_T].
\end{aligned}
$$

Using (B.44), we have $c_T' K_T^{-1} c_T = O(T^{-2})$ and similarly, $K_T^{-1} c_T = O(T^{-1})$. Also, we have that $T^2 c_T' K_T^{-1} c_T = T^2 c_T' P_T^{-1} c_T + O(T^{-1})$. This is so because $T^2 c_T' P_T^{-1} c_T = \frac{T^2 c_T' A_T^{-1} c_T}{1 + T^2 c_T' A_T^{-1} c_T}$ and $1 - \frac{T^2 c_T' A_T^{-1} c_T}{1 + T^2 c_T' A_T^{-1} c_T} = O(T^{-2})$.

When this Proposition is applied to Proposition 2.1 and B.14, we have $b_T = c_T$. Also, it can be extended to the case where d_T and A_T are stochastic. ■

Proof for Proposition B.14: We are going to use Proposition 2.1 to prove in Proposition B.14. First, we need to show that $\Delta_T = 1 - \left(d_T' A_T^{-1} d_T - \frac{(d_T' A_T^{-1} c_T)^2}{c_T' A_T^{-1} c_T} \right)$ has the property that $\Delta \equiv \text{plim}_{T\to\infty} \Delta_T \neq 0$.

In our paper, $\Delta = 1 - \left[d' A^{-1} d - \frac{(d' A^{-1} c)^2}{c' A^{-1} c} \right] = 1 - d' A^{-1} d \left[1 - \frac{(d' A^{-1} c)^2}{c' A^{-1} c d' A^{-1} d} \right]$. Using Cauchy inequality, $\frac{(d' A^{-1} c)^2}{c' A^{-1} c d' A^{-1} d} \leq 1$; using positive definiteness of A, $\frac{(d' A^{-1} c)^2}{c' A^{-1} c d' A^{-1} d} \geq 0$. Hence, $0 \leq 1 - \frac{(d' A^{-1} c)^2}{c' A^{-1} c d' A^{-1} d} \leq 1$. Also, $d' A^{-1} d < 1$ in our application because it is equivalent to $\lim_{T\to\infty} (d_{nT})' (\mathcal{H}_{nT}^s / \omega_{nT})^{-1} (d_{nT}) < 1$ where $d_{nT} = \frac{1}{\omega_{nT}} \left(\frac{1}{nT^2} \sum_{t=1}^{T} (\tilde{Z}_{nt}^s, G_n \tilde{Z}_{nt}^s \delta_0)' \tilde{Y}_{n,t-1}^u \right)$, $\mathcal{H}_{nT}^s = \frac{1}{nT} \sum_{t=1}^{T} (\tilde{Z}_{nt}^s, G_n \tilde{Z}_{nt}^s \delta_0)' (\tilde{Z}_{nt}^s, G_n \tilde{Z}_{nt}^s \delta_0)$ and $\omega_{nT} =$

$\frac{1}{nT^3}\sum_{t=1}^{T}\tilde{Y}^{u\prime}_{n,t-1}\tilde{Y}^{u}_{n,t-1}$. This is so, as we have

$$\left(\frac{1}{nT^2}\sum_{t=1}^{T}\tilde{Y}^{u\prime}_{n,t-1}(\tilde{Z}^s_{nt},G_n\tilde{Z}^s_{nt}\delta_0)\right)\left(\frac{1}{nT^3}\sum_{t=1}^{T}\tilde{Y}^{u\prime}_{n,t-1}\tilde{Y}^{u}_{n,t-1}\right)^{-1}\times$$
$$\left(\frac{1}{nT}\sum_{t=1}^{T}(\tilde{Z}^s_{nt},G_n\tilde{Z}^s_{nt}\delta_0)'(\tilde{Z}^s_{nt},G_n\tilde{Z}^s_{nt}\delta_0)\right)^{-1}\left(\frac{1}{nT^2}\sum_{t=1}^{T}\tilde{Y}^{u\prime}_{n,t-1}(\tilde{Z}^s_{nt},G_n\tilde{Z}^s_{nt}\delta_0)\right)'<1$$

and $d'A^{-1}d<1$ because of the generalized Schwartz inequality and that, for large enough T, $(\tilde{Y}^u_{n,t-1})_i$ is not a linear function of $(\tilde{Z}^s_{nt},G_n\tilde{Z}^s_{nt}\delta_0)_i$ with some positive probability. Hence, combined with $0\leq 1-\frac{(d'A^{-1}c)^2}{c'A^{-1}cd'A^{-1}d}\leq 1$ and $d'A^{-1}d<1$, $\Delta>0$. ■

Proof for (B.28) and (B.29): This is implied by (b) and (c) in Proposition 2.1 when K_T there is taken to be $\mathcal{H}_{1,nT}$ here. ■

Proof for (B.30): To prove it, we take $\mathcal{H}_{1,nT}/\omega_{nT}$ to be K_T. As $K_T^{-1}=O(1)$ and $K_T^{-1}c_T=O_p(T^{-1})$, we need to show $K_T^{-1}(T^2c_T+Td_T)=O_p(1)$.

From (B.43), we have

$$\begin{aligned}K_T^{-1}=&P_T^{-1}-\frac{1}{\Delta_T}\{TP_T^{-1}(d_Tb_T'+b_Td_T')P_T^{-1}-T^2[(d_T'P_T^{-1}d_T)P_T^{-1}b_Tb_T'P_T^{-1}\\&+(b_T'P_T^{-1}b_T)P_T^{-1}d_Td_T'P_T^{-1}-(b_T'P_T^{-1}d_T)P_T^{-1}(d_Tb_T'+b_Td_T')P_T^{-1}]\},\end{aligned}$$

where $\Delta_T=1+T(b_T'P_T^{-1}d_T+d_T'P_T^{-1}b_T)-T^2[(b_T'P_T^{-1}b_T)(d_T'P_T^{-1}d_T)-(b_T'P_T^{-1}d_T)(d_T'P_T^{-1}b_T)]$. It follows that

$$K_T^{-1}d_T=\frac{1}{\Delta_T}\{P_T^{-1}d_T+T(b_T'P_T^{-1}d_T)P_T^{-1}d_T-T(d_T'P_T^{-1}d_T)P_T^{-1}b_T\},$$

and

$$K_T^{-1}b_T=\frac{1}{\Delta_T}\{P_T^{-1}b_T+T(b_T'P_T^{-1}d_T)P_T^{-1}b_T-T(b_T'P_T^{-1}b_T)P_T^{-1}d_T\}.\tag{B.45}$$

As $b_T=c_T$ in our case, after arrangement of terms,

$$K_T^{-1}(T^2c_T+Td_T)=\frac{1}{\Delta_T}\{T^2(1+Tc_T'P_T^{-1}d_T-d_T'P_T^{-1}d_T)P_T^{-1}c_T+T(1+Tc_T'P_T^{-1}d_T-T^2c_T'P_T^{-1}c_T)P_T^{-1}d_T\}.$$

The first part on the right hand side is of order $O_p(1)$ because $P_T^{-1}c_T=O_p(\frac{1}{T^2})$ and $c_T'P_T^{-1}d_T=O_p(\frac{1}{T^2})$. It is of interest to see that for the second half, because

$$\begin{aligned}&T+T^2c_T'P_T^{-1}d_T-T^3c_T'P_T^{-1}c_T\\=\;&T[1+\frac{T}{1+T^2c_T'A_T^{-1}c_T}c_T'A_T^{-1}d_T-\frac{T^2}{1+T^2c_T'A_T^{-1}c_T}c_T'A_T^{-1}c_T]=\frac{T(1+Tc_T'A_T^{-1}d_T)}{1+T^2c_T'A_T^{-1}c_T}=O_p(1),\end{aligned}$$

so $K_T^{-1}(T^2c_T+Td_T)=O_p(1)$. ■

Proof for (B.31): As $\mathcal{H}_{1,nT}$ and $\mathcal{H}_{nT}$ have the form specified in (2.14), we need to prove that for $K_T=T^2c_Tc_T'+T(c_Td_T'+d_Tc_T')+A_T$, we have $1-c_T'K_T^{-1}(T^2c_T+Td_T+Td_{2,nT}c_T+\mathcal{H}^s_{2,nT}/\omega_{nT})=O_p(T^{-1})$

where $d_{2,nT}$ and $\mathcal{H}^s_{2,nT}/\omega_{nT}$ are defined in (2.14). As $c'_T K_T^{-1} = O_p(T^{-1})$ and $c'_T K_T^{-1} c_T = O_p(T^{-2})$, we need to show $1 - c'_T K_T^{-1}(T^2 c_T + T d_T) = O_p(T^{-1})$.

From (B.43),

$$\begin{aligned} T^2 c'_T K_T^{-1} c_T &= T^2 c'_T P_T^{-1} c_T - \frac{T^3}{\Delta_T}(c'_T P_T^{-1} d_T b'_T P_T^{-1} c_T + c'_T P_T^{-1} b_T d'_T P_T^{-1} c_T) \\ &+ \frac{T^4}{\Delta_T}[(d'_T P_T^{-1} d_T) c'_T P_T^{-1} b_T b'_T P_T^{-1} c_T + (b'_T P_T^{-1} b_T) c'_T P_T^{-1} d_T d'_T P_T^{-1} c_T] \\ &+ \frac{T^4}{\Delta_T}[-(d'_T P_T^{-1} b_T) c'_T P_T^{-1} d_T b'_T P_T^{-1} c_T - (b'_T P_T^{-1} d_T) c'_T P_T^{-1} b_T d'_T P_T^{-1} c_T] \\ &= \frac{T^2 c'_T A_T^{-1} c_T}{1 + T^2 c'_T A_T^{-1} c_T} + \frac{1}{\Delta_T}(T^4 c'_T P_T^{-1} b_T b'_T P_T^{-1} c_T)(d'_T P_T^{-1} d_T) + O_p(T^{-1}) \end{aligned}$$

by using (B.44) because $b_T = c_T$ for our case, and, similarly,

$$T c'_T K_T^{-1} d_T = -\frac{1}{\Delta_T}(T^2 c'_T P_T^{-1} b_T)(d'_T P_T^{-1} d_T) + O_p(T^{-1}).$$

Hence,

$$\begin{aligned} & T^2 c'_T K_T^{-1} c_T + T c'_T K_T^{-1} d_T \\ = & \frac{T^2 c'_T A_T^{-1} c_T}{1 + T^2 c'_T A_T^{-1} c_T} + \frac{1}{\Delta_T}(T^4 c'_T P_T^{-1} b_T b'_T P_T^{-1} c_T)(d'_T P_T^{-1} d_T) - \frac{1}{\Delta_T}(T^2 c'_T P_T^{-1} b_T)(d'_T P_T^{-1} d_T) + O_p(T^{-1}) \\ = & \frac{T^2 c'_T A_T^{-1} c_T}{1 + T^2 c'_T A_T^{-1} c_T} + \frac{d'_T P_T^{-1} d_T}{\Delta_T}(T^2 c'_T P_T^{-1} b_T)((T^2 c'_T P_T^{-1} b_T) - 1) + O_p(T^{-1}). \end{aligned}$$

As $b_T = c_T$, $(T^2 c'_T P_T^{-1} b_T) - 1 = -\frac{1}{1 + T^2 c'_T A_T^{-1} c_T} = O_p(T^{-2})$, using the fact that Δ_T and $d'_T P_T^{-1} d_T$ are $O_p(1)$, we have $1 - c'_T K_T^{-1}(T^2 c_T + T d_T) = O_p(T^{-1})$. ■

Proof for (B.32): $\begin{pmatrix} \mathcal{H}_{1,nT} & \mathcal{H}_{2,nT} \\ \mathcal{H}'_{2,nT} & \mathcal{H}_{3,nT} \end{pmatrix} = \omega_{nT}\left(T^2 \cdot c^* c^{*\prime} + T \cdot d_{nT} \cdot c^{*\prime} + T \cdot c^* \cdot d'_{nT} + \mathcal{H}^s_{nT}/\omega_{nT}\right)$ where $c^* = (c', 1)'$ and $d_{nT} = (d'_{1,nT}, d_{2,nT})'$. As we have already established that $\Delta > 0$, by using Proposition B.13, inverse of $\begin{pmatrix} \mathcal{H}_{1,nT} & \mathcal{H}_{2,nT} \\ \mathcal{H}'_{2,nT} & \mathcal{H}_{3,nT} \end{pmatrix}$ exists and is $O_p(1)$. Using the formula of inverting a partitioned matrix, we can get $\left(\mathcal{H}_{3,nT} - \mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT}\right)^{-1}$ exists and is $O_p(1)$. ■

Proof for (B.33): To prove $\text{plim}_{T\to\infty}\left(\mathcal{H}_{3,nT} - \mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT}\right)^{-1} \neq 0$, we will make use of the matrix algebra result of Proposition B.13 we have developed. Denoting $\mathcal{H}_{nT} = \begin{pmatrix} \mathcal{H}_{1,nT} & \mathcal{H}_{2,nT} \\ \mathcal{H}'_{2,nT} & \mathcal{H}_{3,nT} \end{pmatrix}$ and $L_{\mathcal{H}} = \text{plim}_{T\to\infty}\mathcal{H}^{-1}_{nT}$, we are going to prove that $e' L_{\mathcal{H}} e \neq 0$ where e is a unit vector such that $e = (0, \cdots, 0, 1)'$. Here, $\mathcal{H}_{nT}$ takes the form of $\mathcal{H}_{nT} = \omega_{nT} \cdot \left(T^2 \cdot c^* c^{*\prime} + T \cdot d_{nT} \cdot c^{*\prime} + T \cdot c^* \cdot d'_{nT} + \mathcal{H}^s_{nT}/\omega_{nT}\right)$. From Proposition B.13, for $K_T = \left(T^2 \cdot c_T c'_T + T \cdot d_{nT} \cdot c'_T + T \cdot c_T \cdot d'_{nT} + A_T\right)$, the limit of K_T^{-1} is

$$L_k = (A^{-1} - \frac{1}{c' A^{-1} c} A^{-1} c c' A^{-1}) + \frac{1}{\Delta} A^{-1}(d - \frac{d' A^{-1} c}{c' A^{-1} c} c)(d - \frac{d' A^{-1} c}{c' A^{-1} c} c)' A^{-1}$$

where $\Delta = 1 - \left[d'A^{-1}d - \frac{(d'A^{-1}c)^2}{c'A^{-1}c}\right]$. Then,

$$e'L_k e = (e'A^{-1}e - \frac{(e'A^{-1}c)^2}{c'A^{-1}c}) + \frac{1}{\Delta}\left(e'A^{-1}d - \frac{d'A^{-1}c}{c'A^{-1}c}e'A^{-1}c\right)^2.$$

The Cauchy inequality guarantees that $e'A^{-1}e - \frac{(e'A^{-1}c)^2}{c'A^{-1}c} > 0$ as e and c are not proportional. The second part of $e'L_k e$ will be nonnegative if $\Delta > 0$ where $\Delta = 1 - \left[d'A^{-1}d - \frac{(d'A^{-1}c)^2}{c'A^{-1}c}\right] = 1 - d'A^{-1}d\left[1 - \frac{(d'A^{-1}c)^2}{c'A^{-1}cd'A^{-1}d}\right]$, which is proved in the beginning of the proof for Proposition B.14. ■

Proof for (B.34): From (B.32) and (B.33), $\mathcal{H}_{3,nT} - \mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT}$ is $O_p(1)$ and is a function of ω_{nT} and d_{nT}; more explicitly, ω_{nT} and $\omega_{nT}d_{nT}$ in (2.14). As we have $\omega_{nT} - E\omega_{nT} \xrightarrow{p} 0$ and $\omega_{nT}d_{nT} - E\omega_{nT}d_{nT} \xrightarrow{p} 0$ (by using Lemma B.10) where ω_{nT} and d_{nT} are $O_p(1)$, we have the result. ■

Proof for (B.35): (B.33) states that $\text{plim}_{T\to\infty}\left(\mathcal{H}_{3,nT} - \mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT}\right)^{-1} \neq 0$. As $\mathcal{H}_{3,nT} - \mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT}$ is a scalar, we have $\text{plim}_{T\to\infty}\left(\mathcal{H}_{3,nT} - \mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT}\right)$ also exists.

Similarly, $\text{plim}_{T\to\infty}\left(E\mathcal{H}_{3,nT} - E\mathcal{H}'_{2,nT}(E\mathcal{H}_{1,nT})^{-1}E\mathcal{H}_{2,nT}\right)$ also exists. Using (B.34), we have the result. ■

Proof for Proposition B.15: From (B.45), using $b_T = c_T$ for our case, we have

$$TK_T^{-1}c_T = \frac{1}{\Delta_T}\{TP_T^{-1}c_T + (T^2c'_TP_T^{-1}d_T)P_T^{-1}c_T - (T^2c'_TP_T^{-1}c_T)P_T^{-1}d_T\}.$$

As $\Delta_T = O_p(1)$ and bounded away from zero, $c'_TP_T^{-1}d_T = O_p(T^{-2})$ and $P_T^{-1}c_T = O_p(T^{-2})$, we have $TK_T^{-1}c_T = -\frac{1}{\Delta_T}(T^2c'_TP_T^{-1}c_T)P_T^{-1}d_T + O_p(T^{-1})$. Also, as $P_T^{-1} = A_T^{-1} - \frac{T^2}{1+T^2c'_TA_T^{-1}c_T}A_T^{-1}c_Tc'_TA_T^{-1}$ and $T^2c'_TP_T^{-1}c_T = \frac{T^2c'_TA_T^{-1}c_T}{1+T^2c'_TA_T^{-1}c_T}$, we have

$$TK_T^{-1}c_T = -\frac{1}{\Delta_T}\frac{T^2c'_TA_T^{-1}c_T}{1+T^2c'_TA_T^{-1}c_T}(A_T^{-1}d_T - \frac{T^2}{1+T^2c'_TA_T^{-1}c_T}A_T^{-1}c_Tc'_TA_T^{-1}d_T) + O_p(T^{-1}). \quad \text{(B.46)}$$

Also, for K_T^{-1}, using (B.43), it is $O_p(1)$ and is just a function of A_T^{-1}, c_T and d_T. We are going to apply the above results to $\Sigma_{\theta,nT}$ where

$$\Sigma_{\theta,nT} = \frac{1}{\sigma^2}\begin{pmatrix} E\mathcal{H}_{nT}(\theta) & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix} + \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{1}{n}\left[tr(G'_n(\lambda)G_n(\lambda)) + tr(G_n^2(\lambda))\right] & \frac{1}{\sigma^2 n}tr(G_n(\lambda)) \\ \mathbf{0} & \frac{1}{\sigma^2 n}tr(G_n(\lambda)) & \frac{1}{2\sigma^4} \end{pmatrix},$$

and

$$\mathcal{H}_{nT}(\theta) = \frac{1}{nT}\sum_{t=1}^{T}(\tilde{Z}_{nt}, G_n(\lambda)\tilde{Z}_{nt}\delta)'(\tilde{Z}_{nt}, G_n(\lambda)\tilde{Z}_{nt}\delta). \quad \text{(B.47)}$$

As $\tilde{\tilde{Y}}^u_{n,t-1} = W_n\tilde{\tilde{Y}}^u_{n,t-1} = G_n\tilde{Z}^u_{nt}\delta_0$ (see Proposition B.4 in Appendix B), we have $\tilde{Z}^u_{nt} = \tilde{\tilde{Y}}^u_{n,t-1}\cdot c'$ where $c = (1, 1, \mathbf{0}_{1\times k_x})'$ and $G_n(\lambda)\tilde{Z}^u_{nt}\delta = \frac{\gamma+\rho}{1-\lambda}\tilde{\tilde{Y}}^u_{n,t-1}$. Hence, we have $(\tilde{Z}^u_{nt}, G_n(\lambda)\tilde{Z}^u_{nt}\delta) = \tilde{\tilde{Y}}^u_{n,t-1}\cdot(c', \frac{\gamma+\rho}{1-\lambda})'$.

Therefore, denote $c(\theta)=(c',\frac{\gamma+\rho}{1-\lambda},0)'$, we can write $\Sigma_{\theta,nT}$ as

$$\Sigma_{\theta,nT}=(E\omega_{nT})\left(T^2\cdot c(\theta)c'(\theta)+T\cdot(Ed_{nT}(\theta))\cdot c'(\theta)+T\cdot c'(\theta)\cdot(Ed_{nT}(\theta))'+\Sigma^s_{\theta,nT}/E\omega_{nT}\right),$$

where $c(\theta)=(1,1,\mathbf{0}_{1\times k_x},\frac{\gamma+\rho}{1-\lambda},0)'$, $d_{nT}(\theta)=\frac{1}{E\omega_{nT}}(\frac{1}{nT^2}\sum_{t=1}^{T}(\tilde{Z}^s_{nt},G_n(\lambda)\tilde{Z}^s_{nt}\delta,0)'\tilde{\tilde{Y}}^u_{n,t-1})'$, $\Sigma^s_{\theta,nT}=\frac{1}{\sigma^2}\begin{pmatrix}E\mathcal{H}^s_{nT}(\theta) & \mathbf{0}\\ \mathbf{0} & 0\end{pmatrix}+\begin{pmatrix}\mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \frac{1}{n}\left[tr(G'_n(\lambda)G_n(\lambda))+tr(G^2_n(\lambda))\right] & \frac{1}{\sigma^2 n}tr(G_n(\lambda))\\ \mathbf{0} & \frac{1}{\sigma^2 n}tr(G_n(\lambda)) & \frac{1}{2\sigma^4}\end{pmatrix}$ and $\mathcal{H}^s_{nT}(\theta)=\frac{1}{nT}\sum_{t=1}^{T}(\tilde{Z}^s_{nt},G_n(\lambda)\tilde{Z}^s_{nt}\delta)'(\tilde{Z}^s_{nt},G_n(\lambda)\tilde{Z}^s_{nt}\delta)$.

For $c(\theta)$ evaluated at θ_0 and $\hat{\theta}_{nT}$, we have $c(\theta_0)=(c',\frac{\gamma_0+\rho_0}{1-\lambda_0},0)'=(c',1,0)'$ (because $\gamma_0+\rho_0+\lambda_0=1$ under Assumption 6) and $c(\hat{\theta}_{nT})=(c',\frac{\hat{\gamma}_{nT}+\hat{\rho}_{nT}}{1-\hat{\lambda}_{nT}},0)'$. Also, $\frac{\hat{\gamma}_{nT}+\hat{\rho}_{nT}}{1-\hat{\lambda}_{nT}}=1+O_p\left(\max\left(\frac{1}{\sqrt{nT^3}},\frac{1}{T^2}\right)\right)$. This is so as follows. From Theorem 4.2, $\hat{\gamma}_{nT}+\hat{\rho}_{nT}+\hat{\lambda}_{nT}-1=O_p\left(\max\left(\frac{1}{\sqrt{nT^3}},\frac{1}{T^2}\right)\right)$. As $\hat{\lambda}_{nT}-1\neq 0$ for large enough T with probability close to one[12], $\frac{\hat{\gamma}_{nT}+\hat{\rho}_{nT}}{1-\hat{\lambda}_{nT}}=1+\frac{1}{1-\hat{\lambda}_{nT}}O_p\left(\max\left(\frac{1}{\sqrt{nT^3}},\frac{1}{T^2}\right)\right)$. Hence, $\frac{\hat{\gamma}_{nT}+\hat{\rho}_{nT}}{1-\hat{\lambda}_{nT}}=1+O_p\left(\max\left(\frac{1}{\sqrt{nT^3}},\frac{1}{T^2}\right)\right)$.

Hence, for $\Sigma_{\theta_0,nT}$, we have

$$\Sigma_{\theta_0,nT}=(E\omega_{nT})\left(T^2\cdot c(\theta_0)c'(\theta_0)+T\cdot Ed_{nT}(\theta_0)\cdot c'(\theta_0)+T\cdot c'(\theta_0)\cdot(Ed_{nT}(\theta_0))'+\Sigma^s_{\theta_0,nT}/E\omega_{nT}\right),$$

where $c(\theta_0)=(c^{*\prime},0)'$, $d_{nT}(\theta_0)=\frac{1}{E\omega_{nT}}(\frac{1}{nT^2}\sum_{t=1}^{T}(\tilde{Z}^s_{nt},G_n\tilde{Z}^s_{nt}\delta_0,0)'\tilde{\tilde{Y}}^u_{n,t-1})'$ and $\Sigma^s_{\theta_0,nT}=\frac{1}{\sigma_0^2}\begin{pmatrix}E\mathcal{H}^s_{nT} & \mathbf{0}\\ \mathbf{0} & 0\end{pmatrix}+\begin{pmatrix}\mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \frac{1}{n}\left[tr(G'_nG_n)+tr(G^2_n)\right] & \frac{1}{\sigma_0^2 n}tr(G_n)\\ \mathbf{0} & \frac{1}{\sigma_0^2 n}tr(G_n) & \frac{1}{2\sigma_0^4}\end{pmatrix}$. For $\ddot{\Sigma}_{\hat{\theta}_{nT},nT}$ (see (B.36)), we have

$$\ddot{\Sigma}_{\hat{\theta}_{nT},nT}=\omega_{nT}\left(T^2\cdot c(\hat{\theta}_{nT})c'(\hat{\theta}_{nT})+T\cdot d_{nT}(\hat{\theta}_{nT})\cdot c'(\hat{\theta}_{nT})+T\cdot c'(\hat{\theta}_{nT})\cdot d'_{nT}(\hat{\theta}_{nT})+\ddot{\Sigma}^s_{\hat{\theta}_{nT},nT}/\omega_{nT}\right),\tag{B.48}$$

where $c(\hat{\theta}_{nT})=(1,1,\mathbf{0},\frac{\hat{\gamma}_{nT}+\hat{\rho}_{nT}}{1-\hat{\lambda}_{nT}},0)'$, $d_{nT}(\hat{\theta}_{nT})=\frac{1}{\omega_{nT}}(\frac{1}{nT^2}\sum_{t=1}^{T}(\tilde{Z}^s_{nt},G_n(\hat{\lambda}_{nT})\tilde{Z}^s_{nt}\hat{\delta}_{nT},0)'\tilde{\tilde{Y}}^u_{n,t-1})'$ and $\ddot{\Sigma}^s_{\hat{\theta}_{nT},nT}=\frac{1}{\hat{\sigma}^2_{nT}}\begin{pmatrix}\mathcal{H}^s_{nT}(\hat{\theta}_{nT}) & \mathbf{0}\\ \mathbf{0} & 0\end{pmatrix}+\begin{pmatrix}\mathbf{0} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \frac{1}{n}\left[tr(G'_n(\hat{\lambda}_{nT})G_n(\hat{\lambda}_{nT}))+tr(G^2_n(\hat{\lambda}_{nT}))\right] & \frac{1}{\hat{\sigma}^2_{nT}n}tr(G_n(\hat{\lambda}_{nT}))\\ \mathbf{0} & \frac{1}{\hat{\sigma}^2_{nT}n}tr(G_n(\hat{\lambda}_{nT})) & \frac{1}{2\hat{\sigma}^4_{nT}}\end{pmatrix}$.

As $\frac{\hat{\gamma}_{nT}+\hat{\rho}_{nT}}{1-\hat{\lambda}_{nT}}=1+O_p\left(\max\left(\frac{1}{\sqrt{nT^3}},\frac{1}{T^2}\right)\right)$, we have $c(\hat{\theta}_{nT})-c(\theta_0)=(0,0,\mathbf{0},O_p(\max(\frac{1}{\sqrt{nT^3}},\frac{1}{T^2})),0)'$. For $d_{nT}(\hat{\theta}_{nT})-Ed_{nT}(\theta_0)=[d_{nT}(\hat{\theta}_{nT})-d_{nT}(\theta_0)]+[d_{nT}(\theta_0)-Ed_{nT}(\theta_0)]$, $d_{nT}(\hat{\theta}_{nT})-d_{nT}(\theta_0)$ is $O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right)$ as $\left\|\hat{\theta}_{nT}-\theta_0\right\|=O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right)$ from (4.6); also, $d_{nT}(\theta_0)-Ed_{nT}(\theta_0)=O_p\left(\frac{1}{\sqrt{nT}}\right)$ using (B.21) in Lemma B.10. Hence, $d_{nT}(\hat{\theta}_{nT})-Ed_{nT}(\theta_0)=O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right)$. From Equation (C.9) and (C.10) in Yu,

[12]From (3.28), $\hat{\lambda}_{nT}-\lambda_0=O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right)$, this implies that $1-\hat{\lambda}_{nT}=1-\lambda_0+O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right)$. As $\lambda_0\neq 1$ under Assumption 5 (If $\lambda_0=1$, $S_n(\lambda_0)=I_n-W_n$ would not be invertible because W_n is row normalized under Assumption 1.), $1-\hat{\lambda}_{nT}\neq 0$ for large enough T with probability close to one.

de Jong and Lee (2006), $(\ddot{\Sigma}^s_{\hat{\theta}_{nT},nT})^{-1} - (\Sigma^s_{\theta_0,nT})^{-1} = O_p\left(\max\left(\frac{1}{\sqrt{nT}}, \frac{1}{T}\right)\right)$. Also, from (B.8) and (B.9) in Lemma B.8, $\omega_{nT} - E\omega_{nT} = O_p\left(\max\left(\frac{1}{\sqrt{nT}}, \frac{1}{T}\right)\right)$.

To prove (B.37): From (B.43), $\ddot{\Sigma}^{-1}_{\hat{\theta}_{nT},nT}$ is $O_p(1)$ and is a simple function of $\omega_{nT}, c(\hat{\theta}_{nT})$, $d_{nT}(\hat{\theta}_{nT})$, $(\ddot{\Sigma}^s_{\hat{\theta}_{nT},nT})^{-1}$. Similarly, $\Sigma^{-1}_{\theta_0,nT}$ is $O(1)$ and is a simple function of $E\omega_{nT}, c(\theta_0)$, $Ed_{nT}(\theta_0)$, $(\ddot{\Sigma}^s_{\theta_0,nT})^{-1}$. As $\omega_{nT} - E\omega_{nT}$, $c(\hat{\theta}_{nT}) - c(\theta_0)$, $d_{nT}(\hat{\theta}_{nT}) - Ed_{nT}(\theta_0)$ and $(\ddot{\Sigma}^s_{\hat{\theta}_{nT},nT})^{-1} - (\Sigma^s_{\theta_0,nT})^{-1}$ are all having at most the order $O_p\left(\max\left(\frac{1}{\sqrt{nT}}, \frac{1}{T}\right)\right)$, it implies that elements of $\ddot{\Sigma}^{-1}_{\hat{\theta}_{nT},nT} - \Sigma^{-1}_{\theta_0,nT}$ will be of the order $O_p\left(\max\left(\frac{1}{\sqrt{nT}}, \frac{1}{T}\right)\right)$.

To prove (B.38): We are going to show first that $T\cdot\left[\ddot{\Sigma}^{-1}_{\hat{\theta}_{nT},nT}c(\hat{\theta}_{nT}) - \Sigma^{-1}_{\theta_0,nT}c(\theta_0)\right] = O_p\left(\max\left(\frac{1}{\sqrt{nT}}, \frac{1}{T}\right)\right)$. From (B.46), $T\cdot\ddot{\Sigma}^{-1}_{\hat{\theta}_{nT},nT}\cdot c(\hat{\theta}_{nT})$ is a simple function of ω_{nT}, $c(\hat{\theta}_{nT})$, $d_{nT}(\hat{\theta}_{nT})$, $(\ddot{\Sigma}^s_{\hat{\theta}_{nT},nT})^{-1}$ and $T\cdot\Sigma^{-1}_{\theta_0,nT}\cdot c(\theta_0)$ is a simple function of $E\omega_{nT}$, $c(\theta_0)$, $Ed_{nT}(\theta_0)$, $(\Sigma^s_{\theta_0,nT})^{-1}$. As $\omega_{nT} - E\omega_{nT}$, $c(\hat{\theta}_{nT}) - c(\theta_0)$, $d_{nT}(\hat{\theta}_{nT}) - Ed_{nT}(\theta_0)$ and $(\ddot{\Sigma}^s_{\hat{\theta}_{nT},nT})^{-1} - (\Sigma^s_{\theta_0,nT})^{-1}$ are all having at most the order $O_p\left(\max\left(\frac{1}{\sqrt{nT}}, \frac{1}{T}\right)\right)$, it implies that elements of $T\cdot\left[\ddot{\Sigma}^{-1}_{\hat{\theta}_{nT},nT}c(\hat{\theta}_{nT}) - \Sigma^{-1}_{\theta_0,nT}c(\theta_0)\right]$ will be of the order $O_p\left(\max\left(\frac{1}{\sqrt{nT}}, \frac{1}{T}\right)\right)$. As $T\cdot\left[\ddot{\Sigma}^{-1}_{\hat{\theta}_{nT},nT} - \Sigma^{-1}_{\theta_0,nT}\right]\cdot c(\theta_0) = T\cdot\ddot{\Sigma}^{-1}_{\hat{\theta}_{nT},nT}(c(\theta_0) - c(\hat{\theta}_{nT})) + T\cdot\ddot{\Sigma}^{-1}_{\hat{\theta}_{nT},nT}c(\hat{\theta}_{nT}) - \Sigma^{-1}_{\theta_0,nT}c(\theta_0)$, (B.38) follows because $T\cdot(c(\hat{\theta}_{nT}) - c(\theta_0)) = O_p\left(\max\left(\frac{1}{\sqrt{nT}}, \frac{1}{T}\right)\right)$ and $\ddot{\Sigma}^{-1}_{\hat{\theta}_{nT},nT}$ is $O_p(1)$. ■

B.4 Proof for Proposition 2.2

We have already proved the central limit theorem for statistics like Q^s_{nT} for stationary case (see Theorem 2.4 in Yu, de Jong and Lee (2006)). As $Q_{nT} = Q^s_{nT} + Q^u_{nT}$ specified in (2.15), we can prove that Q_{nT} still behaves like Q^s_{nT}. Rewrite Q_{nT} as

$$\begin{aligned} Q_{nT} &= \sum_{t=1}^{T}\left(\mathbb{U}_{n,t-1} + \frac{k_T}{T}(R_nJ_nR_n^{-1})\xi_{n,t-1}\right)'\cdot V_{nt} \\ &+ \sum_{t=1}^{T}\left(D_{nt} + \frac{k_T}{T}(R_nJ_nR_n^{-1})\left(\mathbf{c}_{n0}\tilde{t}_{-1} + \tilde{\mathbb{X}}_{n,t-1}\beta_0\right)\right)'\cdot V_{nt} \\ &+ \sum_{t=1}^{T}\left(V'_{nt}\mathcal{B}_nV_{nt} - \sigma_0^2 tr\mathcal{B}_n\right), \end{aligned}$$

then Q_{nT} has just another form of Q^s_{nT} so that the central limit theorem in Yu, de Jong and Lee (2006) is applicable. To confirm this, we need to show that

(1) For $\mathbb{W}_{nt} = \mathbb{U}_{nt} + \frac{k_T}{T}(R_nJ_nR_n^{-1})\xi_{n,t} = \sum_{h=1}^{\infty}Q_{nt,h}V_{n,t+1-h}$, $\sum_{h=1}^{\infty}abs(Q_{nt,h})$ is row sum and column sum bounded uniformly in n and t;

(2) For $\mathbb{D}_{nt} = D_{nt} + \frac{k_T}{T}(R_nJ_nR_n^{-1})\left(\mathbf{c}_{n0}\tilde{t}_{-1} + \tilde{\mathbb{X}}_{n,t-1}\beta_0\right)$, elements of $\mathbb{D}_{nt}$ is bounded uniformly in n and t.

For (1), as $\xi_{nt} = \sum_{h=0}^{t-1}V_{nh}$, we have $Q_{nt,h} = \begin{cases} P_{nt,h} + \frac{k_T}{T}(R_nJ_nR_n^{-1}) & \text{for } h \le t-1 \\ P_{nt,h} & \text{for } h > t-1 \end{cases}$. Hence,

$$\sum_{h=1}^{\infty}abs(Q_{nt,h}) = \sum_{h=1}^{t-1}abs(P_{nt,h} + \frac{k_T}{T}(R_nJ_nR_n^{-1})) + \sum_{h=t}^{\infty}abs(P_{nt,h}).$$

This implies that $\|\sum_{h=1}^{\infty} abs(Q_{nt,h})\| \leq \|\sum_{h=1}^{\infty} abs(P_{nt,h})\| + \left\|\sum_{h=1}^{t-1} abs(\frac{k_T}{T}(R_n J_n R_n^{-1}))\right\|$, where $\|\cdot\|$ represents either the row or column sum norm. Here, $\sum_{h=1}^{\infty} abs(P_{nt,h})$ is row sum and column sum bounded uniformly in n and t. Also, $\left\|\sum_{h=1}^{t-1} abs(\frac{k_T}{T}(R_n J_n R_n^{-1}))\right\|$ is row sum and column sum bounded uniformly in n and t because $R_n J_n R_n^{-1}$ is row sum and column sum bounded and K_T is $O(1)$.

For (2), $\mathbb{D}_{nt} = D_{nt} + \frac{k_T}{T}(R_n J_n R_n^{-1})\left(\mathbf{c}_{n0}\tilde{t}_{-1} + \tilde{\mathbb{X}}_{n,t-1}\beta_0\right)$ where $\tilde{\mathbb{X}}_{n,t} = \sum_{h=0}^{t-1} \tilde{X}_{nt}$. As elements of $\mathbf{c}_{n0}$, X_{nt} and D_{nt} are uniformly bounded in n and t and $R_n J_n R_n^{-1}$ is row sum and column sum bounded, using $\frac{t}{T} \leq 1$ for $t = 1, 2, ..., T$, elements of $\mathbb{D}_{nt}$ are uniformly bounded. ■

B.5 Proof about $\hat{\sigma}^2_{nT}(\lambda)$ and $\hat{\sigma}^2_{nT}(\lambda_0)$

Using (2.9), we have

$$\begin{aligned}
\hat{\sigma}^2_{nT}(\lambda) &= (\lambda_0-\lambda)^2(\mathcal{H}_{3,nT} - \mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT}) + \frac{1}{nT}\sum_{t=1}^{T}\tilde{V}'_{nt}S'^{-1}_n S'_n(\lambda)S_n(\lambda)S_n^{-1}\tilde{V}_{nt} \\
&\quad +2(\lambda_0-\lambda)\frac{1}{nT}\sum_{t=1}^{T}(\delta'_0\tilde{Z}'_{nt}G'_n - \mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\tilde{Z}'_{nt})S_n(\lambda)S_n^{-1}\tilde{V}_{nt} \\
&\quad -\left(\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}'_{nt}S_n(\lambda)S_n^{-1}\tilde{V}_{nt}\right)'\mathcal{H}^{-1}_{1,nT}\left(\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}'_{nt}S_n(\lambda)S_n^{-1}\tilde{V}_{nt}\right),
\end{aligned} \tag{B.49}$$

$$\begin{aligned}
\frac{\partial\hat{\sigma}^2_{nT}(\lambda)}{\partial\lambda} &= 2(\lambda-\lambda_0)(\mathcal{H}_{3,nT} - \mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT}) - \frac{2}{nT}\sum_{t=1}^{T}\tilde{V}'_{nt}G'_nS_n(\lambda)S_n^{-1}\tilde{V}_{nt} \\
&\quad -2(\lambda_0-\lambda)\frac{1}{nT}\sum_{t=1}^{T}(\delta'_0\tilde{Z}'_{nt}G'_n - \mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\tilde{Z}'_{nt})G_n\tilde{V}_{nt} \\
&\quad -\frac{2}{nT}\sum_{t=1}^{T}(\delta'_0\tilde{Z}'_{nt}G'_n - \mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\tilde{Z}'_{nt})S_n(\lambda)S_n^{-1}\tilde{V}_{nt} \\
&\quad +2\left(\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}'_{nt}S_n(\lambda)S_n^{-1}\tilde{V}_{nt}\right)'\mathcal{H}^{-1}_{1,nT}\left(\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}'_{nt}G_n\tilde{V}_{nt}\right),
\end{aligned} \tag{B.50}$$

$$\begin{aligned}
\frac{\partial^2\hat{\sigma}^2_{nT}(\lambda)}{\partial\lambda^2} &= 2(\mathcal{H}_{3,nT} - \mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT}) + \frac{2}{nT}\sum_{t=1}^{T}\tilde{V}'_{nt}G'_nG_n\tilde{V}_{nt} \\
&\quad +\frac{4}{nT}\sum_{t=1}^{T}(\delta'_0\tilde{Z}'_{nt}G'_n - \mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\tilde{Z}'_{nt})G_n\tilde{V}_{nt} \\
&\quad -2\left(\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}'_{nt}G_n\tilde{V}_{nt}\right)'\mathcal{H}^{-1}_{1,nT}\left(\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}'_{nt}G_n\tilde{V}_{nt}\right).
\end{aligned} \tag{B.51}$$

To study $\hat{\sigma}^2_{nT}(\lambda)$, $\frac{\partial\hat{\sigma}^2_{nT}(\lambda)}{\partial\lambda}$ and $\frac{\partial^2\hat{\sigma}^2_{nT}(\lambda)}{\partial\lambda^2}$, we need the following (B.52).

$$\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}^{u\prime}_{nt}B_n\tilde{V}_{nt} = \left(\frac{1}{nT}\sum_{t=1}^{T}\tilde{Y}^{u\prime}_{n,t-1}B_n\tilde{V}_{nt}\right)\cdot c, \tag{B.52a}$$

$$G_n\tilde{Z}^u_{nt}\delta_0 - \tilde{Z}^u_{nt}\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT} = \tilde{Y}^{u\prime}_{n,t-1}\cdot(1 - c'\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT}). \tag{B.52b}$$

For $\hat{\sigma}^2_{nT}(\lambda)$ in (B.49), using (B.32) in Proposition B.14, $(\lambda_0-\lambda)^2(\mathcal{H}_{3,nT}-\mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT})$ is $|\lambda-\lambda_0|\cdot O_p(1)$. Using (B.19) and $S_n(\lambda)S_n^{-1}=I_n-(\lambda-\lambda_0)G_n$, $\frac{1}{nT}\sum_{t=1}^T \tilde{V}'_{nt}S_n'^{-1}S'_n(\lambda)S_n(\lambda)S_n^{-1}\tilde{V}_{nt}=\sigma_0^2+|\lambda-\lambda_0|\cdot O_p(1)+O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right)$. For $2(\lambda_0-\lambda)\frac{1}{nT}\sum_{t=1}^T(\delta_0'\tilde{Z}'_{nt}G'_n-\mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\tilde{Z}'_{nt})S_n(\lambda)S_n^{-1}\tilde{V}_{nt}$, it can be rewritten as a sum of stationary component and nonstationary component. For the stationary component $2(\lambda_0-\lambda)\frac{1}{nT}\sum_{t=1}^T(\delta_0'\tilde{Z}^{s\prime}_{nt}G'_n-\mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\tilde{Z}^{s\prime}_{nt})S_n(\lambda)S_n^{-1}\tilde{V}_{nt}$, using (B.24), it is $|\lambda-\lambda_0|\cdot O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right)$. For the nonstationary component $2(\lambda_0-\lambda)\frac{1}{nT}\sum_{t=1}^T(\delta_0'\tilde{Z}^{u\prime}_{nt}G'_n-\mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\tilde{Z}^{u\prime}_{nt})S_n(\lambda)S_n^{-1}\tilde{V}_{nt}$, it is equal to

$$2(\lambda_0-\lambda)(1-c'\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT})\frac{1}{nT}\sum_{t=1}^T\tilde{Y}^{u\prime}_{n,t-1}S_n(\lambda)S_n^{-1}\tilde{V}_{nt}$$

using (B.52). Then, using (B.23) and $(1-c'\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT})=O_p(T^{-1})$ from (B.31), $2(\lambda_0-\lambda)\frac{1}{nT}\sum_{t=1}^T(\delta_0'\tilde{Z}^{u\prime}_{nt}G'_n-\mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\tilde{Z}^{u\prime}_{nt})S_n(\lambda)S_n^{-1}\tilde{V}_{nt}$ is $|\lambda-\lambda_0|\cdot O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right)$. For the last term of $\hat{\sigma}^2_{nT}(\lambda)$ in (B.49), we have $\frac{1}{nT}\sum_{t=1}^T\tilde{Z}'_{nt}S_n(\lambda)S_n^{-1}\tilde{V}_{nt}=\frac{1}{nT}\sum_{t=1}^T\tilde{Z}^{u\prime}_{nt}S_n(\lambda)S_n^{-1}\tilde{V}_{nt}+\frac{1}{nT}\sum_{t=1}^T\tilde{Z}^{s\prime}_{nt}S_n(\lambda)S_n^{-1}\tilde{V}_{nt}$. As $\frac{1}{nT}\sum_{t=1}^T\tilde{Z}^{s\prime}_{nt}S_n(\lambda)S_n^{-1}\tilde{V}_{nt}=O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right)$ from (B.24) and $\frac{1}{nT}\sum_{t=1}^T\tilde{Z}^{u\prime}_{nt}S_n(\lambda)S_n^{-1}\tilde{V}_{nt}=c\cdot O_p\left(\max\left(1,\sqrt{\frac{T}{n}}\right)\right)$ from (B.52) and (B.23), using $\mathcal{H}^{-1}_{1,nT}$ is $O_p(1)$, $\mathcal{H}^{-1}_{1,nT}c$ is $O_p(T^{-1})$ and $c'\mathcal{H}^{-1}_{1,nT}c$ is $O_p(T^{-2})$ from Proposition B.14, we have

$$\left(\frac{1}{nT}\sum_{t=1}^T\tilde{Z}'_{nt}S_n(\lambda)S_n^{-1}\tilde{V}_{nt}\right)'\mathcal{H}^{-1}_{1,nT}\left(\frac{1}{nT}\sum_{t=1}^T\tilde{Z}'_{nt}S_n(\lambda)S_n^{-1}\tilde{V}_{nt}\right)=O_p\left(\max\left(\frac{1}{nT},\frac{1}{\sqrt{nT^3}},\frac{1}{T^2}\right)\right). \quad \text{(B.53)}$$

Hence, $\hat{\sigma}^2_{nT}(\lambda)=\sigma_0^2+|\lambda-\lambda_0|\cdot O_p(1)+O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right)$. Also,

$$\hat{\sigma}^2_{nT}(\lambda)=(\lambda-\lambda_0)^2(\mathcal{H}_{3,nT}-\mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT})+\sigma_0^2\frac{1}{n}tr(S_n'^{-1}S'_n(\lambda)S_n(\lambda)S_n^{-1})+O_p\left(\max\left(\frac{1}{T},\frac{1}{\sqrt{nT}}\right)\right). \quad \text{(B.54)}$$

Similarly, we can get the behavior of $\frac{\partial\hat{\sigma}^2_{nT}(\lambda)}{\partial\lambda}$ and $\frac{\partial^2\hat{\sigma}^2_{nT}(\lambda)}{\partial\lambda^2}$ by using Lemma B.11, (B.52) and Proposition B.14. The results are summarized as follows.

$$\hat{\sigma}^2_{nT}(\lambda) = \sigma_0^2+|\lambda-\lambda_0|\cdot O_p(1)+O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right), \quad \text{(B.55a)}$$

$$\frac{\partial\hat{\sigma}^2_{nT}(\lambda)}{\partial\lambda} = -\sigma_0^2\frac{2}{n}trG_n+|\lambda-\lambda_0|\cdot O_p(1)+O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right), \quad \text{(B.55b)}$$

$$\frac{\partial^2\hat{\sigma}^2_{nT}(\lambda)}{\partial\lambda^2} = 2(\mathcal{H}_{3,nT}-\mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT})+2\sigma_0^2\frac{1}{n}trG'_nG_n+O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right). \quad \text{(B.55c)}$$

Furthermore, at $\lambda=\lambda_0$, we have, from (B.50),

$$\begin{aligned}\sqrt{nT}\frac{\partial\hat{\sigma}^2_{nT}(\lambda_0)}{\partial\lambda} &= -\frac{2}{\sqrt{nT}}\sum_{t=1}^T\tilde{V}'_{nt}G'_n\tilde{V}_{nt}-\frac{2}{\sqrt{nT}}\sum_{t=1}^T(\delta_0'\tilde{Z}'_{nt}G'_n-\mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\tilde{Z}'_{nt})\tilde{V}_{nt}\\ &\quad +2\left(\frac{1}{\sqrt{nT}}\sum_{t=1}^T\tilde{Z}'_{nt}\tilde{V}_{nt}\right)'\mathcal{H}^{-1}_{1,nT}\left(\frac{1}{nT}\sum_{t=1}^T\tilde{Z}'_{nt}G_n\tilde{V}_{nt}\right).\end{aligned}$$

Using (B.53), we have

$$\sqrt{nT}\frac{\partial\hat{\sigma}^2_{nT}(\lambda_0)}{\partial\lambda} = -\frac{2}{\sqrt{nT}}\sum_{t=1}^{T}\tilde{V}'_{nt}G'_n\tilde{V}_{nt} - \frac{2}{\sqrt{nT}}\sum_{t=1}^{T}(\delta'_0\tilde{Z}'_{nt}G'_n - \mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\tilde{Z}'_{nt})\tilde{V}_{nt} + O_p\left(\max\left(\frac{1}{\sqrt{nT}}, \frac{1}{T}, \sqrt{\frac{n}{T^3}}\right)\right). \blacksquare \quad \text{(B.56)}$$

B.6 Proof about $\hat{\sigma}^{*2}_{nT}(\lambda)$ and $\hat{\sigma}^{*2}_{nT}(\lambda_0)$

From (3.18), we have

$$\begin{aligned}\sigma^{*2}_{nT}(\lambda) &= (\lambda_0-\lambda)^2(E\mathcal{H}_{3,nT} - E\mathcal{H}'_{2,nT}(E\mathcal{H}_{1,nT})^{-1}E\mathcal{H}_{2,nT}) + \frac{1}{nT}E\sum_{t=1}^{T}\tilde{V}'_{nt}S'^{-1}_nS'_n(\lambda)S_n(\lambda)S^{-1}_n\tilde{V}_{nt}\\ &+2(\lambda_0-\lambda)\frac{1}{nT}E\sum_{t=1}^{T}(\delta'_0\tilde{Z}'_{nt}G'_n - E\mathcal{H}'_{2,nT}(E\mathcal{H}_{1,nT})^{-1}\tilde{Z}'_{nt})S_n(\lambda)S^{-1}_n\tilde{V}_{nt}\\ &-\left(\frac{1}{nT}E\sum_{t=1}^{T}\tilde{Z}'_{nt}S_n(\lambda)S^{-1}_n\tilde{V}_{nt}\right)'(E\mathcal{H}_{1,nT})^{-1}\left(E\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}'_{nt}S_n(\lambda)S^{-1}_n\tilde{V}_{nt}\right), \end{aligned} \quad \text{(B.57)}$$

$$\begin{aligned}\frac{\partial\sigma^{*2}_{nT}(\lambda)}{\partial\lambda} &= 2(\lambda-\lambda_0)(E\mathcal{H}_{3,nT} - E\mathcal{H}'_{2,nT}(E\mathcal{H}_{1,nT})^{-1}E\mathcal{H}_{2,nT}) - \frac{2}{nT}E\sum_{t=1}^{T}\tilde{V}'_{nt}G'_nS_n(\lambda)S^{-1}_n\tilde{V}_{nt}\\ &-2(\lambda_0-\lambda)\frac{1}{nT}E\sum_{t=1}^{T}(\delta'_0\tilde{Z}'_{nt}G'_n - E\mathcal{H}'_{2,nT}(E\mathcal{H}_{1,nT})^{-1}\tilde{Z}'_{nt})G_n\tilde{V}_{nt}\\ &-\frac{2}{nT}E\sum_{t=1}^{T}(\delta'_0\tilde{Z}'_{nt}G'_n - E\mathcal{H}'_{2,nT}(E\mathcal{H}_{1,nT})^{-1}\tilde{Z}'_{nt})S_n(\lambda)S^{-1}_n\tilde{V}_{nt}\\ &+2\left(E\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}'_{nt}S_n(\lambda)S^{-1}_n\tilde{V}_{nt}\right)'(E\mathcal{H}_{1,nT})^{-1}\left(E\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}'_{nt}G_n\tilde{V}_{nt}\right), \end{aligned} \quad \text{(B.58)}$$

$$\begin{aligned}\frac{\partial^2\sigma^{*2}_{nT}(\lambda)}{\partial\lambda^2} &= 2(E\mathcal{H}_{3,nT} - E\mathcal{H}'_{2,nT}(E\mathcal{H}_{1,nT})^{-1}E\mathcal{H}_{2,nT}) + \frac{2}{nT}E\sum_{t=1}^{T}\tilde{V}'_{nt}G'_nG_n\tilde{V}_{nt}\\ &+\frac{4}{nT}E\sum_{t=1}^{T}(\delta'_0\tilde{Z}'_{nt}G'_n - E\mathcal{H}'_{2,nT}(E\mathcal{H}_{1,nT})^{-1}\tilde{Z}'_{nt})G_n\tilde{V}_{nt}\\ &-2\left(\frac{1}{nT}E\sum_{t=1}^{T}\tilde{Z}'_{nt}G_n\tilde{V}_{nt}\right)'(E\mathcal{H}_{1,nT})^{-1}\left(\frac{1}{nT}E\sum_{t=1}^{T}\tilde{Z}'_{nt}G_n\tilde{V}_{nt}\right). \end{aligned} \quad \text{(B.59)}$$

Using Lemma B.11, (B.52) and Proposition B.14, similarly as we derived (B.55), we have

$$\sigma^{*2}_{nT}(\lambda) = \sigma^2_0 + |\lambda-\lambda_0|\cdot O_p(1) + O\left(\frac{1}{T}\right), \quad \text{(B.60a)}$$

$$\frac{\partial\sigma^{*2}_{nT}(\lambda)}{\partial\lambda} = -\sigma^2_0\frac{2}{n}trG_n + |\lambda-\lambda_0|\cdot O_p(1) + O\left(\frac{1}{T}\right), \quad \text{(B.60b)}$$

$$\frac{\partial^2\sigma^{*2}_{nT}(\lambda)}{\partial\lambda^2} = 2(E\mathcal{H}_{3,nT} - E\mathcal{H}'_{2,nT}(E\mathcal{H}_{1,nT})^{-1}E\mathcal{H}_{2,nT}) + 2\sigma^2_0\frac{1}{n}trG'_nG_n + \left(\frac{1}{T}\right). \quad \text{(B.60c)}$$

Furthermore,

$$\sigma_{nT}^{*2}(\lambda) = (\lambda-\lambda_0)^2(E\mathcal{H}_{3,nT} - E\mathcal{H}_{2,nT}'(E\mathcal{H}_{1,nT})^{-1}E\mathcal{H}_{2,nT}) + \sigma_0^2\frac{1}{n}tr(S_n'^{-1}S_n'(\lambda)S_n(\lambda)S_n^{-1}) + O\left(\frac{1}{T}\right). \blacksquare \tag{B.61}$$

C Proof for Theorems

C.1 Proof of Claim 3.1

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

As $\ln L_{n,T}(\lambda) = -\frac{nT}{2}(\ln 2\pi+1) - \frac{nT}{2}\ln\hat{\sigma}_{nT}^2(\lambda) + T\ln|S_n(\lambda)|$ and $Q_{n,T}(\lambda) = -\frac{1}{2}(\ln 2\pi+1) - \frac{1}{2}\ln\sigma_{nT}^{*2}(\lambda) + \frac{1}{n}\ln|S_n(\lambda)|$ ((2.10) and (3.19)), $\frac{1}{nT}\ln L_{n,T}(\lambda) - Q_{n,T}(\lambda) = \frac{1}{2}\ln\sigma_{nT}^{*2}(\lambda) - \frac{1}{2}\ln\hat{\sigma}_{nT}^2(\lambda)$. By the mean value theorem, $\frac{1}{nT}\ln L_{n,T}(\lambda) - Q_{n,T}(\lambda) = -\frac{1}{2}\frac{1}{\tilde{\sigma}_{n,T}^2(\lambda)}(\hat{\sigma}_{nT}^2(\lambda) - \sigma_{nT}^{*2}(\lambda))$ where $\tilde{\sigma}_{nT}^2(\lambda)$ lies between $\hat{\sigma}_{nT}^2(\lambda)$ and $\sigma_{nT}^{*2}(\lambda)$. We need to show that (1) $\hat{\sigma}_{nT}^2(\lambda) - \sigma_{nT}^{*2}(\lambda) \xrightarrow{p} 0$ uniformly in λ and (2) $\tilde{\sigma}_{nT}^2(\lambda)$ is bounded away from zero uniformly in Λ with probability one.

To prove (1): We have $\hat{\sigma}_{nT}^2(\lambda)$ and $\sigma_{nT}^{*2}(\lambda)$ in (B.54) and (B.61). Using (B.34) in Proposition B.14, $\hat{\sigma}_{nT}^2(\lambda) - \sigma_{nT}^{*2}(\lambda) \xrightarrow{p} 0$ uniformly in λ.

To prove (2): As $\tilde{\sigma}_{nT}^2(\lambda)$ lies between $\hat{\sigma}_{nT}^2(\lambda)$ and $\sigma_{nT}^{*2}(\lambda)$, we have $\frac{1}{\tilde{\sigma}_{nT}^2(\lambda)} \le \max\{\frac{1}{\hat{\sigma}_{nT}^2(\lambda)}, \frac{1}{\sigma_{nT}^{*2}(\lambda)}\}$. Denote $\sigma_{nT}^2(\lambda) = \sigma_0^2\frac{1}{n}tr(S_n'^{-1}S_n'(\lambda)S_n(\lambda)S_n^{-1})$, then $\sigma_{nT}^2(\lambda)$ is uniformly bounded away from zero[13]. As $\mathcal{H}_{3,nT} - \mathcal{H}_{2,nT}'\mathcal{H}_{1,nT}^{-1}\mathcal{H}_{2,nT}$ is nonnegative[14], $\hat{\sigma}_{nT}^2(\lambda)$ and $\sigma_{nT}^{*2}(\lambda)$ are uniformly bounded away from zero. So, $\frac{1}{\tilde{\sigma}_{nT}^2(\lambda)}$ is uniformly bounded.

Combining (1) $\hat{\sigma}_{nT}^2(\lambda) - \sigma_{nT}^{*2}(\lambda) \xrightarrow{p} 0$ uniformly in λ and (2) $\frac{1}{\tilde{\sigma}_{nT}^2(\lambda)}$ is $O_p(1)$ uniformly in λ, we have $\frac{1}{nT}\ln L_{n,T}(\lambda) - Q_{n,T}(\lambda) \xrightarrow{p} 0$ uniformly in λ.

To prove $Q_{n,T}(\lambda)$ *is uniformly equicontinuous in* λ *in any compact parameter space* Λ:

To prove this property, from the expression of $Q_{n,T}(\lambda)$, the followings are sufficient: (1) $\frac{1}{n}\ln|S_n(\lambda)|$ is uniformly equicontinuous; (2) $(\lambda-\lambda_0)^2(E\mathcal{H}_{3,nT} - E\mathcal{H}_{2,nT}'(E\mathcal{H}_{1,nT})^{-1}E\mathcal{H}_{2,nT})$ is uniformly equicontinuous; (4) $\sigma_n^2(\lambda)$ is uniformly equicontinuous.

For (1), $\frac{1}{n}\ln|S_n(\lambda_2)| - \frac{1}{n}\ln|S_n(\lambda_1)| = \frac{1}{n}tr\left(W_nS_n^{-1}\left(\bar{\lambda}\right)\right)(\lambda_2-\lambda_1)$ 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 $\lambda \in \Lambda$, $\frac{1}{n}tr\left(W_nS_n^{-1}\left(\bar{\lambda}\right)\right)$ is bounded, we have $\frac{1}{n}\ln|S_n(\lambda)|$ is uniformly equicontinuous. For (2), because λ is bounded and because $E\mathcal{H}_{3,nT} - E\mathcal{H}_{2,nT}'(E\mathcal{H}_{1,nT})^{-1}E\mathcal{H}_{2,nT}$ is $O(1)$ according to Proposition B.14, the result follows. For (3), $\sigma_n^2(\lambda_2) - \sigma_n^2(\lambda_1) = \frac{\sigma_0^2}{n}tr(S_n'^{-1}S_n'(\lambda_2)S_n(\lambda_2)S_n^{-1}) - \frac{\sigma_0^2}{n}tr(S_n'^{-1}S_n'(\lambda_1)S_n(\lambda_1)S_n^{-1})$. Using $S_n(\lambda)S_n^{-1} = I_n - (\lambda-\lambda_0)G_n$, $\sigma_n^2(\lambda_2) - \sigma_n^2(\lambda_1) = \sigma_0^2\left[(\lambda_2-\lambda_1)(\lambda_2+\lambda_1-2\lambda_0)\frac{trG_n'G_n}{n} - (\lambda_2-\lambda_1)\frac{tr\left(G_n'+G_n\right)}{n}\right]$. As elements of $G_n'G_n$ and G_n are uniformly bounded, $\sigma_n^2(\lambda)$ is uniformly equicontinuous. ■

[13]See the supplement to Lee (2004), Page 8 for the proof of consistency, available in http://economics.sbs.ohio-state.edu/lee/.

[14]Here, $\mathcal{H}_{3,nT} - \mathcal{H}_{2,nT}'\mathcal{H}_{1,nT}^{-1}\mathcal{H}_{2,nT} \ge 0$ because of the Cauchy inequality.

C.2 Proof of Nonsingularity of Information Matrix (Scalar)

From (3.21), $\partial^2 Q_{n,T}(\lambda_0)/\partial\lambda^2 = -\frac{1}{\sigma_0^2}(E\mathcal{H}_{3,nT} - E\mathcal{H}'_{2,nT}(E\mathcal{H}_{1,nT})^{-1}E\mathcal{H}_{2,nT}) - \frac{1}{n}\left(trG'_nG_n + trG_n^2 - \frac{2(trG_n)^2}{n}\right) + O_p\left(T^{-1}\right)$. Then, using (B.35) in Proposition B.14, $\lim_{T\to\infty}\frac{\partial^2 Q_{n,T}(\lambda_0)}{\partial\lambda^2} = -\frac{1}{\sigma_0^2}\text{plim}_{T\to\infty}(\mathcal{H}_{3,nT} - \mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT}) - \lim_{n\to\infty}\frac{1}{n}[trG'_nG_n + trG_n^2 - \frac{2(trG_n)^2}{n}]$. If $\text{plim}_{T\to\infty}(\mathcal{H}_{3,nT} - \mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT}) \neq 0$ or $\lim_{n\to\infty}\frac{1}{n}(trG'_nG_n + trG_n^2 - \frac{2(trG_n)^2}{n}) \neq 0$, $\lim_{T\to\infty} -\frac{\partial^2 Q_{n,T}(\lambda_0)}{\partial\lambda^2}$ is positive. Here, $(\mathcal{H}_{3,nT} - \mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT}) > 0$ for large enough T because of the Cauchy inequality; also, denote $\mathcal{C}_n = G_n - \frac{trG_n}{n}I_n$, then, $\frac{1}{n}\{trG'_nG_n + trG_n^2 - \frac{2(trG_n)^2}{n}\} = \frac{1}{n}tr(\mathcal{C}_n + \mathcal{C}'_n)(\mathcal{C}_n + \mathcal{C}'_n)' \geq 0$. ■

C.3 Proof of Theorem 3.2

We have

$$Q_{n,T}(\lambda) = -\frac{1}{2}(\ln 2\pi + 1) - \frac{1}{2}\ln\sigma^{*2}_{nT}(\lambda) + \frac{1}{n}\ln|S_n(\lambda)| \tag{C.1}$$

where $\sigma^{*2}_{nT}(\lambda) = (\lambda_0 - \lambda)^2(E\mathcal{H}_{3,nT} - E\mathcal{H}'_{2,nT}(E\mathcal{H}_{1,nT})^{-1}E\mathcal{H}_{2,nT}) + \frac{1}{n}\sigma_0^2 tr(S_n'^{-1}S_n(\lambda)S_n(\lambda)S_n^{-1}) + O(\frac{1}{T})$ and the $O(\frac{1}{T})$ is uniformly in λ. At $\lambda = \lambda_0$, $Q_{n,T}(\lambda_0) = -\frac{1}{2}(\ln 2\pi + 1) - \frac{1}{2}\ln\sigma^{*2}_{nT}(\lambda_0) + \frac{1}{n}\ln|S_n(\lambda_0)|$. We are going to prove that $\lim_{T\to\infty} Q_{n,T}(\lambda) < \lim_{T\to\infty} Q_{n,T}(\lambda_0)$ for any $\lambda \neq \lambda_0$.

$$\begin{aligned} Q_{n,T}(\lambda) - Q_{n,T}(\lambda_0) &= -\frac{1}{2}[\ln\sigma^{*2}_{nT}(\lambda) - \ln\sigma^{*2}_{nT}(\lambda_0)] + \frac{1}{n}\ln|S_n(\lambda)| - \frac{1}{n}\ln|S_n(\lambda_0)| \\ &= T_{1,nT} - T_{2,nT} + O(\frac{1}{T}) \end{aligned}$$

where

$$\begin{aligned} T_{1,nT} &= -\frac{1}{2}[\ln\{\frac{1}{n}\sigma_0^2 tr(S_n'^{-1}S_n(\lambda)S_n(\lambda)S_n^{-1})\} - \ln\sigma^{*2}_{nT}(\lambda_0)] + \frac{1}{n}\ln|S_n(\lambda)| - \frac{1}{n}\ln|S_n(\lambda_0)| \\ T_{2,nT} &= \ln\left(1 + \frac{(\lambda_0 - \lambda)^2(E\mathcal{H}_{3,nT} - E\mathcal{H}'_{2,nT}(E\mathcal{H}_{1,nT})^{-1}E\mathcal{H}_{2,nT})}{\sigma_0^2 tr(S_n'^{-1}S_n(\lambda)S_n(\lambda)S_n^{-1})/n}\right). \end{aligned}$$

Consider the pure spatial dynamic panel process $Y_{nt} = \lambda_0 W_n Y_{nt} + \mathbf{c}_{n0} + V_{nt}$, the concentrated log likelihood function of this process is

$$\ln L_{p,n,T}(\theta) = -\frac{nT}{2}\ln 2\pi - \frac{nT}{2}\ln\sigma^2 + T\ln|S_n(\lambda)| - \frac{1}{2\sigma^2}\sum_{t=1}^{T}(S_n(\lambda)Y_{nt} - \mathbf{c}_{n0})'(S_n(\lambda)Y_{nt} - \mathbf{c}_{n0}), \tag{C.2}$$

and the concentrated likelihood is

$$\ln L_{p,n,T}(\lambda) = -\frac{nT}{2}(\ln 2\pi + 1) - \frac{nT}{2}\ln\hat{\sigma}^2_{p,nT}(\lambda) + T\ln|S_n(\lambda)|, \tag{C.3}$$

where $\hat{c}_{p,nT}(\lambda) = \frac{1}{T}\sum\limits_{t=1}^{T} S_n(\lambda)Y_{nt}$ and $\hat{\sigma}^2_{p,nT}(\lambda) = \frac{1}{nT}\sum\limits_{t=1}^{T}(S_n(\lambda)\tilde{Y}_{nt})'S_n(\lambda)\tilde{Y}_{nt}$. Then, $E\ln L_{p,n,T}(\lambda) - E\ln L_{p,n,T}(\lambda_0)$ would be equal to $T_{1,nT}$. By information inequality, $E\ln L_{p,n,T}(\lambda) - E\ln L_{p,n,T}(\lambda_0) \leq 0$. Thus, $T_{1,nT} \leq 0$ for any λ. Also, $\lim_{T\to\infty} T_{2,nT} > 0$ as long as $\lim_{T\to\infty}(E\mathcal{H}_{3,nT} - E\mathcal{H}'_{2,nT}(E\mathcal{H}_{1,nT})^{-1}E\mathcal{H}_{2,nT}) \neq 0$. Under Assumption 9, $\lim_{T\to\infty}(E\mathcal{H}_{3,nT} - E\mathcal{H}'_{2,nT}(E\mathcal{H}_{1,nT})^{-1}E\mathcal{H}_{2,nT}) \neq 0$ from Proposition B.14. This proves the global identification. The consistency then follows from the global identification, uniform convergence and uniform equicontinuity in Claim 3.1. ■

C.4 Proof of Theorem 3.3

As $\tilde{Z}_{nt}$ has stationary and nonstationary parts (see (2.12)), we can decompose $\frac{1}{\sqrt{nT}}\frac{\partial \ln L_{n,T}(\lambda_0)}{\partial\lambda}$ from (3.22) into two parts accordingly:

$$\frac{1}{\sqrt{nT}}\frac{\partial \ln L_{n,T}(\lambda_0)}{\partial\lambda} = \frac{1}{\sqrt{nT}}\frac{\partial \ln L^s_{n,T}(\lambda_0)}{\partial\lambda} + \frac{1}{\sqrt{nT}}\frac{\partial \ln L^u_{n,T}(\lambda_0)}{\partial\lambda} + O_p\left(\max\left(\frac{1}{\sqrt{nT}}, \frac{1}{T}, \sqrt{\frac{n}{T^3}}\right)\right), \quad \text{(C.4)}$$

where $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^s_{n,T}(\lambda_0)}{\partial\lambda}$ is the stationary part and $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^u_{n,T}(\lambda_0)}{\partial\lambda}$ is the nonstationary part as defined via (C.5)-(C.9). The $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^s_{n,T}(\lambda_0)}{\partial\lambda}$ has two parts, namely, $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^s_{n,T}(\lambda_0)}{\partial\lambda} = \frac{1}{\sqrt{nT}}\frac{\partial \ln L^{s*}_{n,T}(\lambda_0)}{\partial\lambda} - \Delta_{\lambda,nT}$ where

$$\frac{1}{\sqrt{nT}}\frac{\partial \ln L^{s*}_{n,T}(\lambda_0)}{\partial\lambda} = \frac{1}{\hat{\sigma}^2_{nT}(\lambda_0)}\left(\frac{1}{\sqrt{nT}}\sum_{t=1}^T V'_{nt}(G'_n - \frac{1}{n}trG_n\cdot I_n)V_{nt} + \frac{1}{\sqrt{nT}}\sum_{t=1}^T(\delta'_0\ddot{Z}^{s*\prime}_{nt}G'_n - \mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\ddot{Z}^{s*\prime}_{nt})V_{nt}\right) \quad \text{(C.5)}$$

and

$$\Delta_{\lambda_0,nT} = \frac{1}{\hat{\sigma}^2_{n,T}(\lambda_0)}\left(\sqrt{\frac{T}{n}}\bar{V}'_{nT}(G'_n - \frac{1}{n}trG_n\cdot I_n)\bar{V}_{nT}\right) \quad \text{(C.6)}$$
$$+\frac{1}{\hat{\sigma}^2_{n,T}(\lambda_0)}\left(\sqrt{\frac{T}{n}}(\delta'_0(\bar{U}^s_{nT,-1}, W_n\bar{U}^s_{nT,-1}, \mathbf{0}_{n\times k_x})'G'_n - \mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}(\bar{U}^s_{nT,-1}, W_n\bar{U}^s_{nT,-1}, \mathbf{0}_{n\times k_x})')\bar{V}_{nT}\right),$$

Here, $\ddot{Z}^{s*}_{nt}$ is the component of $\tilde{Z}^s_{nt}$, which is uncorrelated with V_{nt} such that

$$\tilde{Z}^s_{nt} = \ddot{Z}^{s*}_{nt} - (\bar{U}^s_{nT,-1},\ W_n\bar{U}^s_{nT,-1},\ \mathbf{0}_{n\times k_x}) \quad \text{(C.7)}$$

and $\ddot{Z}^{s*}_{nt} = ((\tilde{\tilde{\mathcal{X}}}^s_{n,t-1}\beta_0 + U^s_{n,t-1}),\ (W_n\tilde{\tilde{\mathcal{X}}}^s_{n,t-1}\beta_0 + W_nU^s_{n,t-1}),\ \tilde{X}_{nt})$ with $\tilde{\tilde{\mathcal{X}}}^s_{n,t-1} = \mathcal{X}^s_{n,t-1} - \bar{\mathcal{X}}^s_{nT,-1}$. For $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^u_{n,T}(\lambda_0)}{\partial\lambda}$, it also has two parts $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^u_{n,T}(\lambda_0)}{\partial\lambda} = \frac{1}{\sqrt{nT}}\frac{\partial \ln L^{u*}_{n,T}(\lambda_0)}{\partial\lambda} - \blacktriangle_{\lambda_0,nT}$ where

$$\frac{1}{\sqrt{nT}}\frac{\partial \ln L^{u*}_{n,T}(\lambda_0)}{\partial\lambda} = \frac{1}{\hat{\sigma}^2_{n,T}(\lambda_0)}\left\{(1 - c'\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT})\cdot\frac{1}{\sqrt{nT}}\sum_{t=1}^T \ddot{Y}^{u*\prime}_{n,t-1}V_{nt}\right\} \quad \text{(C.8)}$$

with $\ddot{Y}^{u*}_{n,t-1} = \frac{1}{(1-\lambda_0)}M_n\left(\mathbf{c}_{n0}\tilde{t}_{-1} + \tilde{\tilde{\mathbb{X}}}_{n,t-1}\beta_0 + \xi_{n,t-1}\right)$, $\tilde{t}_{-1} = (t-1) - \frac{T-1}{2}$ and

$$\blacktriangle_{\lambda_0,nT} = \frac{T(1 - c'\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT})}{\hat{\sigma}^2_{n,T}(\lambda_0)(1-\lambda_0)}\left\{\sqrt{\frac{n}{T}}\frac{1}{n}(M_n\bar{\xi}_{n,T-1})'\cdot\bar{V}_{nT}\right\}. \quad \text{(C.9)}$$

For $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\lambda_0)}{\partial\lambda} = \frac{1}{\sqrt{nT}}\frac{\partial \ln L^{s*}_{n,T}(\lambda_0)}{\partial\lambda} + \frac{1}{\sqrt{nT}}\frac{\partial \ln L^{u*}_{n,T}(\lambda_0)}{\partial\lambda}$ from (C.5) and (C.8), denote

$$\ddot{Z}^*_{nt} = \ddot{Z}^{s*}_{nt} + \ddot{Z}^{u*}_{nt}, \quad \text{(C.10)}$$

we have

$$\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\lambda_0)}{\partial\lambda} = \frac{1}{\hat{\sigma}^2_{nT}(\lambda_0)}\left(\frac{1}{\sqrt{nT}}\sum_{t=1}^T V'_{nt}(G'_n - \frac{1}{n}trG_n\cdot I_n)V_{nt} + \frac{1}{\sqrt{nT}}\sum_{t=1}^T(\delta'_0\ddot{Z}^{*\prime}_{nt}G'_n - \mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\ddot{Z}^{*\prime}_{nt})V_{nt}\right).$$

Proposition 2.2 implies that it will be asymptotically normally distributed because $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^{s*}_{n,T}(\lambda_0)}{\partial\lambda}$ from (C.5) and $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^{u*}_{n,T}(\lambda_0)}{\partial\lambda}$ from (C.8) are counterparts of Q^s_{nT} and Q^u_{nT} in Proposition 2.2. To calculate the limit variance, using uncorrelatedness of $\ddot{Z}^*_{nt}$ and V_{nt}, we have

$$\begin{aligned} & Cov\left(\frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}(V'_{nt}(G'_n - \sigma_0^2\frac{1}{n}trG_n)V_{nt}), \frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}(\delta_0\ddot{Z}^{*\prime}_{nt}G'_n - \mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\ddot{Z}^{*\prime}_{nt})V_{nt}\right) \\ = & \frac{\mu_3}{\sigma_0^4}\frac{1}{nT}\sum_{i=1}^{n}(G_{n,ii} - \frac{1}{n}trG_n)\left(E\sum_{t=1}^{T}(G_n\ddot{Z}^*_{nt}\delta_0 - \ddot{Z}^*_{nt}\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT})\right)_i = 0 \end{aligned}$$

because $E\sum\limits_{t=1}^{T}G_n\ddot{Z}^*_{nt}\delta_0 = 0$ and $E\sum\limits_{t=1}^{T}\ddot{Z}^*_{nt} = \mathbf{0}$. Hence,

$$\frac{1}{\sqrt{nT}}\frac{\partial \ln L^{s*}_{n,T}(\lambda_0)}{\partial\lambda} + \frac{1}{\sqrt{nT}}\frac{\partial \ln L^{u*}_{n,T}(\lambda_0)}{\partial\lambda} \xrightarrow{d} N(0, \Sigma_{\lambda_0} + \Omega_{\lambda_0}) \tag{C.11}$$

where

$$\Sigma_{\lambda_0} = \frac{1}{\sigma_0^2}\lim_{T\to\infty}(\mathcal{H}_{3,nT} - \mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT}) + \lim_{n\to\infty}\frac{1}{n}(trG'_nG_n + trG_n^2 - \frac{2(trG_n)^2}{n}), \tag{C.12}$$

$$\Omega_{\lambda_0} = \frac{\mu_4 - 3\sigma_0^4}{\sigma_0^4}\lim_{n\to\infty}\sum_{i=1}^{n}G^2_{n,ii}. \tag{C.13}$$

For $\Delta_{\lambda_0,nT}$, using the results in Yu, de Jong and Lee (2006) (Theorem A.11, page 19), we have $\Delta_{\lambda_0,nT}\cdot\frac{\hat{\sigma}^2_{nT}(\lambda_0)}{\sigma_0^2} = \sqrt{\frac{n}{T}}a^s_{\lambda_0,nT} + O_p\left(\max\left(\sqrt{\frac{n}{T^3}}, \frac{1}{\sqrt{T}}\right)\right)$ where

$$\begin{aligned} a^s_{\lambda_0,nT} = & \frac{1}{n}tr\left(G_n\gamma_0 - (\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT})_1 I_n\right)\left(\sum\nolimits_{h=0}^{\infty}B_n^h\right)S_n^{-1} \\ & + \frac{1}{n}tr\left(G_n\rho_0 - (\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT})_2 I_n\right)\left(\sum\nolimits_{h=0}^{\infty}W_nB_n^h\right)S_n^{-1} \end{aligned}$$

is $O(1)$. As $\hat{\sigma}^2_{nT}(\lambda_0) = \sigma_0^2 + O_p\left(\max\left(\frac{1}{\sqrt{nT}}, \frac{1}{T}\right)\right)$ from (3.17) so that $\frac{\sigma_0^2}{\hat{\sigma}^2_{nT}(\lambda_0)} = 1 + O_p\left(\max\left(\frac{1}{\sqrt{nT}}, \frac{1}{T}\right)\right)$, $\Delta_{\lambda_0,nT} = \sqrt{\frac{n}{T}}a^s_{\lambda_0,nT} + O_p\left(\max\left(\sqrt{\frac{n}{T^3}}, \frac{1}{\sqrt{T}}\right)\right)$. Also, using (B.17) and (C.9), we have $\blacktriangle_{\lambda_0,nT}\cdot\frac{\hat{\sigma}^2_{nT}(\lambda_0)}{\sigma_0^2} = \sqrt{\frac{n}{T}}\cdot\frac{m_n}{n}\cdot a^u_{\lambda_0,nT} + O_p\left(\max\left(\sqrt{\frac{n}{T^3}}, \frac{1}{\sqrt{T}}\right)\right)$ where $a^u_{\lambda_0,nT} = T\cdot(1 - c'\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT})\frac{1}{2(1-\lambda_0)}$. As $\frac{\sigma_0^2}{\hat{\sigma}^2_{nT}(\lambda_0)} = 1 + O_p\left(\max\left(\frac{1}{\sqrt{nT}}, \frac{1}{T}\right)\right)$, $\blacktriangle_{\lambda_0,nT} = \sqrt{\frac{n}{T}}\cdot\frac{m_n}{n}\cdot a^u_{\lambda_0,nT} + O_p\left(\max\left(\sqrt{\frac{n}{T^3}}, \frac{1}{\sqrt{T}}\right)\right)$. Hence,

$$\Delta_{\lambda_0,nT} + \blacktriangle_{\lambda_0,nT} = \sqrt{\frac{n}{T}}\cdot(a^s_{\lambda_0,nT} + \frac{m_n}{n}\cdot a^u_{\lambda_0,nT}) + O_p\left(\max\left(\sqrt{\frac{n}{T^3}}, \frac{1}{\sqrt{T}}\right)\right). \blacksquare \tag{C.14}$$

C.5 Proof of Claim 3.4

First, by the mean value theorem, $tr(G_n^2(\lambda)) = tr(G_n^2) + 2tr(G_n^3(\bar{\lambda}))(\lambda - \lambda_0)$ where $\bar{\lambda}$ lies between λ and λ_0. So, $\frac{1}{n}tr(G_n^2(\lambda)) = \frac{1}{n}tr(G_n^2) + |\lambda - \lambda_0|\cdot O(1)$ as $\frac{1}{n}tr(G_n^3(\bar{\lambda}))$ is uniformly bounded (see Lee(2001), Lemma A.8 on page 22). Second, using (3.16), we can express $\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\lambda)}{\partial\lambda^2}$ in terms of $\hat{\sigma}^2_{nT}(\lambda)$, $\frac{\partial\hat{\sigma}^2_{nT}(\lambda)}{\partial\lambda}$ and $\frac{\partial^2\hat{\sigma}^2_{nT}(\lambda)}{\partial\lambda^2}$. Then, using (3.17), we have the result that $\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\lambda)}{\partial\lambda^2} - \frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\lambda_0)}{\partial\lambda^2} = |\lambda - \lambda_0|\cdot O_p(1) + O_p\left(\max\left(\frac{1}{\sqrt{nT}}, \frac{1}{T}\right)\right)$. $\blacksquare$

C.6 Proof of Claim 3.5

Using (3.16), we can express $\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\lambda_0)}{\partial\lambda^2}$ in terms of $\hat{\sigma}^2_{nT}(\lambda_0)$, $\frac{\partial\hat{\sigma}^2_{nT}(\lambda_0)}{\partial\lambda}$ and $\frac{\partial^2\hat{\sigma}^2_{nT}(\lambda_0)}{\partial\lambda^2}$ where, using (3.17),

$$\begin{aligned}
\hat{\sigma}^2_{nT}(\lambda_0) &= \sigma_0^2 + O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right),\ \frac{\partial\hat{\sigma}^2_{nT}(\lambda_0)}{\partial\lambda} = -2\frac{1}{n}\sigma_0^2 trG_n + O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right) \\
\frac{\partial^2\hat{\sigma}^2_{nT}(\lambda_0)}{\partial\lambda^2} &= 2(\mathcal{H}_{3,nT} - \mathcal{H}'_{2,nT}\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT}) + 2\sigma_0^2\frac{1}{n}trG'_nG_n + O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right).
\end{aligned}$$

Similarly, we can express $\frac{\partial^2 Q_{n,T}(\lambda_0)}{\partial\lambda^2}$ in terms of $\sigma^{*2}_{nT}(\lambda_0)$, $\frac{\partial\sigma^{*2}_{nT}(\lambda_0)}{\partial\lambda}$ and $\frac{\partial^2\sigma^{*2}_{nT}(\lambda_0)}{\partial\lambda^2}$ via (3.20) where, using (B.60),

$$\begin{aligned}
\sigma^{*2}_{nT}(\lambda_0) &= \sigma_0^2 + O\left(\frac{1}{T}\right),\ \frac{\partial\sigma^{*2}_{nT}(\lambda_0)}{\partial\lambda} = -2\sigma_0^2\frac{1}{n}trG_n + O\left(\frac{1}{T}\right), \\
\frac{\partial^2\sigma^{*2}_{nT}(\lambda_0)}{\partial\lambda^2} &= 2(E\mathcal{H}_{3,nT} - E\mathcal{H}'_{2,nT}(E\mathcal{H}_{1,nT})^{-1}E\mathcal{H}_{2,nT}) + 2\sigma_0^2\frac{1}{n}trG'_nG_n + \left(\frac{1}{T}\right).
\end{aligned}$$

Hence, we have the result $\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\lambda_0)}{\partial\lambda^2} - \frac{\partial^2 Q_{n,T}(\lambda_0)}{\partial\lambda^2} = O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right)$. ■

C.7 Proof of Theorem 3.6

(3.28) follows from the Taylor expansion $(\hat{\lambda}_{nT} - \lambda_0) = (-\frac{\partial^2 \ln L_{n,T}(\bar{\lambda}_{nT})}{\partial\lambda^2})^{-1}\frac{\partial \ln L_{n,T}(\lambda_0)}{\partial\lambda}$ where $\bar{\lambda}_{nT}$ lies between λ_0 and $\hat{\lambda}_{nT}$. Note that because

$$-\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\bar{\lambda}_{nT})}{\partial\lambda^2} = \left(-\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\bar{\lambda}_{nT})}{\partial\lambda^2} - \left(-\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\lambda_0)}{\partial\lambda^2}\right)\right) + \left(-\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\lambda_0)}{\partial\lambda^2} - \Sigma_{\lambda_0,nT}\right) + \Sigma_{\lambda_0,nT}$$

where $\Sigma_{\lambda_0,nT} \equiv -\frac{\partial^2 Q_{n,T}(\lambda_0)}{\partial\lambda^2}$, we have $-\frac{\partial^2 \ln L_{n,T}(\bar{\lambda}_{nT})}{\partial\lambda^2} = \Sigma_{\lambda_0,nT} + \left|\hat{\lambda}_{nT} - \lambda_0\right| \cdot O_p(1) + O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right)$ according to Claim C.5 and C.6. Because $\left|\hat{\lambda}_{nT} - \lambda_0\right| = o_p(1)$ as $\hat{\lambda}_{nT}$ is consistent and $\Sigma_{\lambda_0,nT}$ is positive in the limit from Appendix C.2, we have $-\frac{\partial^2 \ln L_{n,T}(\bar{\lambda}_{nT})}{\partial\lambda^2}$ is invertible for large T and $\left(-\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\bar{\lambda}_{nT})}{\partial\lambda\partial\lambda'}\right)^{-1}$ is $O_p(1)$.

According to the Taylor expansion, $\sqrt{nT}(\hat{\lambda}_{nT}-\lambda_0) = \left(-\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\bar{\lambda}_{nT})}{\partial\lambda^2}\right)^{-1}\cdot\left(\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\lambda_0)}{\partial\lambda} - \Delta_{\lambda_0,nT} - \blacktriangle_{\lambda_0,nT}\right)$ where $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\lambda_0)}{\partial\lambda} \xrightarrow{d} N(0, \Sigma_{\lambda_0} + \Omega_{\lambda_0})$ from (C.11) and $\Delta_{\lambda_0,nT} + \blacktriangle_{\lambda_0,nT} = \sqrt{\frac{n}{T}} \cdot (a^s_{\lambda_0,nT} + \frac{m_n}{n} \cdot a^u_{\lambda_0,nT}) + O_p\left(\max\left(\sqrt{\frac{n}{T^3}}, \frac{1}{\sqrt{T}}\right)\right)$ with $a^s_{\lambda_0,nT} + \frac{m_n}{n}\cdot a^u_{\lambda_0,nT} = O(1)$ from (C.14). Then, $\sqrt{nT}(\hat{\lambda}_{nT} - \lambda_0) = O_p(1)\cdot\left(O_p(1) + O\left(\sqrt{\frac{n}{T}}\right)\right)$, which implies that $\hat{\lambda}_{nT} - \lambda_0 = O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right)$. Hence,

$$\begin{aligned}
\sqrt{nT}(\hat{\lambda}_{nT} - \lambda_0) &= \left(-\frac{1}{nT}\frac{\partial^2 \ln L_{n,T}(\bar{\lambda}_{nT})}{\partial\lambda^2}\right)^{-1}\cdot\left(\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\lambda_0)}{\partial\lambda} - \Delta_{\lambda_0,nT} - \blacktriangle_{\lambda_0,nT}\right) \qquad \text{(C.17)} \\
&= \left(\Sigma_{\lambda_0,nT} + O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right)\right)^{-1}\cdot\left(\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\lambda_0)}{\partial\lambda} - \Delta_{\lambda_0,nT} - \blacktriangle_{\lambda_0,nT}\right)
\end{aligned}$$

using Claim C.6. Using the fact that

$$\left(\Sigma_{\lambda_0,nT} + O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right)\right)^{-1} = \Sigma^{-1}_{\lambda_0,nT} + O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right) \qquad \text{(C.18)}$$

given that $\Sigma_{\lambda_0,nT}$ is positive in the limit, we have

$$\begin{aligned}\sqrt{nT}(\hat{\lambda}_{nT}-\lambda_0) &= \left(\Sigma^{-1}_{\lambda_0,nT}+O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right)\right)\cdot\left(\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\lambda_0)}{\partial\lambda}-\Delta_{\lambda_0,nT}-\blacktriangle_{\lambda_0,nT}\right)\\ &= \Sigma^{-1}_{\lambda_0,nT}\cdot\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\lambda_0)}{\partial\lambda}+O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right)\cdot\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\lambda_0)}{\partial\lambda}\\ &\quad -\Sigma^{-1}_{\lambda_0,nT}\cdot(\Delta_{\lambda_0,nT}+\blacktriangle_{\lambda_0,nT})-O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right)\cdot(\Delta_{\lambda_0,nT}+\blacktriangle_{\lambda_0,nT}),\end{aligned}$$

which implies that

$$\begin{aligned}&\sqrt{nT}(\hat{\lambda}_{nT}-\lambda_0)+\Sigma^{-1}_{\lambda_0,nT}\cdot(\Delta_{\lambda_0,nT}+\blacktriangle_{\lambda_0,nT})+O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right)\cdot(\Delta_{\lambda_0,nT}+\blacktriangle_{\lambda_0,nT})\\ =\ &(\Sigma^{-1}_{\lambda_0,nT}+o_p(1))\cdot\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\lambda_0)}{\partial\lambda}. \qquad (C.19)\end{aligned}$$

As $\Sigma_{\lambda_0}=\lim_{T\to\infty}\Sigma_{\lambda_0,nT}$ exists, then using Theorem 3.3 and that $\Delta_{\lambda_0,nT}+\blacktriangle_{\lambda_0,nT}=\sqrt{\frac{n}{T}}\cdot(a^s_{\lambda_0,nT}+\frac{m_n}{n}\cdot a^u_{\lambda_0,nT})+O_p\left(\max\left(\sqrt{\frac{n}{T^3}},\frac{1}{\sqrt{T}}\right)\right)$, we have $\sqrt{nT}(\hat{\lambda}_{nT}-\lambda_0)+\sqrt{\frac{n}{T}}b_{\lambda_0,nT}+O_p\left(\max\left(\sqrt{\frac{n}{T^3}},\frac{1}{\sqrt{T}}\right)\right)\xrightarrow{d} N(0,\Sigma^{-1}_{\lambda_0}+\Sigma^{-2}_{\lambda_0}\Omega_{\lambda_0})$. The results in (3.30)-(3.32) are immediate consequences of (3.28). ■

C.8 Proof for (4.1)[15]

From concentrated estimators ((2.9)), $\hat{\delta}_{nT}(\lambda)=\delta_0-(\lambda-\lambda_0)\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT}+\mathcal{H}^{-1}_{1,nT}\left[\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}'_{nt}S_n(\lambda)S_n^{-1}\tilde{V}_{nt}\right]$. Using $S_n(\lambda)S_n^{-1}=I_n-(\lambda-\lambda_0)G_n$,

$$\begin{aligned}\sqrt{nT}\left(\hat{\delta}_{n,T}(\hat{\lambda}_{nT})-\delta_0\right) &= -\sqrt{nT}(\hat{\lambda}_{nT}-\lambda_0)\left(\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT}+\mathcal{H}^{-1}_{1,nT}\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}'_{nt}G_n\tilde{V}_{nt}\right) \qquad (C.20)\\ &\quad +\mathcal{H}^{-1}_{1,nT}\left(\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}\tilde{Z}'_{nt}\tilde{V}_{nt}\right)\\ &= -\sqrt{nT}(\hat{\lambda}_{nT}-\lambda_0)\mathcal{H}^{-1}_{1,nT}\mathcal{H}_{2,nT}+\mathcal{H}^{-1}_{1,nT}\left(\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}\tilde{Z}'_{nt}\tilde{V}_{nt}\right)+R_{\hat{\delta}_{nT}}\end{aligned}$$

where $R_{\hat{\delta}_{nT}}=-\sqrt{nT}(\hat{\lambda}_{nT}-\lambda_0)\mathcal{H}^{-1}_{1,nT}\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}'_{nt}G_n\tilde{V}_{nt}$.

For the term $\mathcal{H}^{-1}_{1,nT}\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}'_{nt}G_n\tilde{V}_{nt}=\mathcal{H}^{-1}_{1,nT}\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}^{s\prime}_{nt}G_n\tilde{V}_{nt}+\mathcal{H}^{-1}_{1,nT}\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}^{u\prime}_{nt}G_n\tilde{V}_{nt}$, we have $\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}^{s\prime}_{nt}G_n\tilde{V}_{nt}=O_p\left(\max\left(\frac{1}{T},\frac{1}{\sqrt{nT}}\right)\right)$ from Theorem A.7 in Yu, de Jong and Lee (2006) and $\mathcal{H}^{-1}_{1,nT}\frac{1}{nT}\sum_{t=1}^{T}\tilde{Z}^{u\prime}_{nt}G_n\tilde{V}_{nt}=\mathcal{H}^{-1}_{1,nT}c\cdot\frac{1}{nT}\sum_{t=1}^{T}\tilde{\tilde{Y}}^{u\prime}_{n,t-1}G_n\tilde{V}_{nt}=O_p\left(\max\left(\frac{1}{T},\frac{1}{\sqrt{nT}}\right)\right)$ because $\frac{1}{nT}\sum_{t=1}^{T}\tilde{\tilde{Y}}^{u\prime}_{n,t-1}G_n\tilde{V}_{nt}=O_p\left(\max\left(1,\sqrt{\frac{T}{n}}\right)\right)$ from Lemma B.11 and $\mathcal{H}^{-1}_{1,nT}\cdot c=O_p(T^{-1})$. Then, because $\hat{\lambda}_{nT}-\lambda_0=O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right)$ from

[15] Note that the derivation of (4.1) is built up from the estimates of various components of $\hat{\theta}_{nT}$ in (2.9). The reason is that the conventional mean value theorem can not be directly applied to the $\frac{\partial^2 \ln L_{nT}(\theta)}{\partial\theta\partial\theta'}$ at θ_0 for analysis due to technical complication.

(3.28), $\mathcal{H}_{1,nT}^{-1} \cdot c = O_p(T^{-1})$ and $c'\mathcal{H}_{1,nT}^{-1} \cdot c = O_p(T^{-2})$, we have $R_{\hat{\delta}_{nT}} = O_p\left(\max\left(\frac{1}{T}, \frac{1}{\sqrt{nT}}, \sqrt{\frac{n}{T^3}}\right)\right)$ and $c' \cdot R_{\hat{\delta}_{nT}} = O_p\left(\max\left(\frac{1}{T^2}, \frac{1}{\sqrt{nT^3}}, \sqrt{\frac{n}{T^5}}\right)\right)$.

Also, from (B.49) and $\hat{\lambda}_{nT} - \lambda_0 = O_p\left(\max\left(\frac{1}{\sqrt{nT}}, \frac{1}{T}\right)\right)$,

$$\sqrt{nT}\left(\hat{\sigma}_{n,T}^2(\hat{\lambda}_{nT}) - \sigma_0^2\right) = \frac{1}{\sqrt{nT}}\left(\sum_{t=1}^{T} \tilde{V}_{nt}'\tilde{V}_{nt} - n\sigma_0^2\right) - 2\sigma_0^2 \frac{1}{n} trG_n \left(\sqrt{nT}(\hat{\lambda}_{nT} - \lambda_0)\right) + R_{\hat{\sigma}_{nT}^2} \quad \text{(C.21a)}$$

$$R_{\hat{\sigma}_{nT}^2} = O_p\left(\max\left(\frac{1}{T}, \frac{1}{\sqrt{nT}}, \sqrt{\frac{n}{T^3}}\right)\right). \quad \text{(C.21b)}$$

From the Taylor expansion, $\sqrt{nT}\left(\hat{\lambda}_{nT} - \lambda_0\right) = \left(-\frac{\partial^2 \ln L_{n,T}(\bar{\lambda}_{nT})}{\partial \lambda^2}\right)^{-1} \frac{1}{\sqrt{nT}} \frac{\partial \ln L_{n,T}(\lambda_0)}{\partial \lambda}$ where $\bar{\lambda}_{nT}$ lies between λ_0 and $\hat{\lambda}_{nT}$. From Claim 3.4, Claim 3.5 and (C.18), $\left(-\frac{\partial^2 \ln L_{n,T}(\bar{\lambda}_{nT})}{\partial \lambda^2}\right)^{-1} = \Sigma_{\lambda_0}^{-1} + O_p\left(\max\left(\frac{1}{\sqrt{nT}}, \frac{1}{T}\right)\right)$. Using Theorem 3.3,

$$\sqrt{nT}\left(\hat{\lambda}_{nT} - \lambda_0\right) = \Sigma_{\lambda_0}^{-1} \frac{1}{\sqrt{nT}} \frac{\partial \ln L_{n,T}(\lambda_0)}{\partial \lambda} + R_{\hat{\lambda}_{nT}} \text{ and } R_{\hat{\lambda}_{nT}} = O_p\left(\max\left(\frac{1}{T}, \frac{1}{\sqrt{nT}}, \sqrt{\frac{n}{T^3}}\right)\right). \quad \text{(C.22a)}$$

Hence, we have

$$\begin{aligned}
&\sqrt{nT}\begin{pmatrix} I_{k_x+2} & \mathcal{H}_{1,nT}^{-1}\mathcal{H}_{2,nT} & \mathbf{0} \\ \mathbf{0} & 1 & 0 \\ \mathbf{0} & 2\frac{trG_n}{n}\sigma_0^2 & 1 \end{pmatrix}\begin{pmatrix} \hat{\delta}_{n,T}(\hat{\lambda}_{nT}) - \delta_0 \\ \hat{\lambda}_{nT} - \lambda_0 \\ \hat{\sigma}_{nT}^2(\hat{\lambda}_{nT}) - \sigma_0^2 \end{pmatrix} \\
&= \begin{pmatrix} \mathcal{H}_{1,nT}^{-1}\left(\frac{1}{\sqrt{nT}}\sum_{t=1}^{T} \tilde{Z}_{nt}'\tilde{V}_{nt}\right) \\ \Sigma_{\lambda_0}^{-1}\left(\frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\sum_{t=1}^{T} \tilde{V}_{nt}'(G_n' - \frac{1}{n}trG_n)\tilde{V}_{nt} + \frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}(\delta_0'\tilde{Z}_{nt}'G_n' - \mathcal{H}_{2,nT}'\mathcal{H}_{1,nT}^{-1}\tilde{Z}_{nt}')\tilde{V}_{nt}\right) \\ \frac{1}{\sqrt{nT}}\sum_{t=1}^{T}\left(\tilde{V}_{nt}'\tilde{V}_{nt} - n\sigma_0^2\right) \end{pmatrix} \\
&+(R_{\hat{\delta}_{nT}}', R_{\hat{\lambda}_{nT}}, R_{\hat{\sigma}_{nT}^2})' \\
&= \begin{pmatrix} \mathcal{H}_{1,nT}^{-1}\sigma_0^2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \Sigma_{\lambda_0}^{-1} & 0 \\ \mathbf{0} & 0 & 2\sigma_0^4 \end{pmatrix} \times \begin{pmatrix} \left(\frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}\tilde{Z}_{nt}'\tilde{V}_{nt}\right) \\ \left(\begin{array}{c} \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}{\sigma_0^2}\frac{1}{n}trG_n\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}(\tilde{V}_{nt}'\tilde{V}_{nt} - n\sigma_0^2) \\ +\frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}(\delta_0'\tilde{Z}_{nt}'G_n' - \mathcal{H}_{2,nT}'\mathcal{H}_{1,nT}^{-1}\tilde{Z}_{nt}')\tilde{V}_{nt} \end{array}\right) \\ \frac{1}{2\sigma_0^4}\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}\left(\tilde{V}_{nt}'\tilde{V}_{nt} - n\sigma_0^2\right) \end{pmatrix} \\
&+(R_{\hat{\delta}_{nT}}', R_{\hat{\lambda}_{nT}}, R_{\hat{\sigma}_{nT}^2})' \\
&= \begin{pmatrix} \mathcal{H}_{1,nT}^{-1}\sigma_0^2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \Sigma_{\lambda_0}^{-1} & 0 \\ \mathbf{0} & 0 & 2\sigma_0^4 \end{pmatrix} \times \begin{pmatrix} I_{k_x+2} & \mathbf{0} & \mathbf{0} \\ -\mathcal{H}_{2,nT}'\mathcal{H}_{1,nT}^{-1} & 1 & -\frac{2}{n}\sigma_0^2 trG_n \\ \mathbf{0} & 0 & 1 \end{pmatrix} \\
&\times \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}(\tilde{V}_{nt}'G_n'\tilde{V}_{nt} - \sigma_0^2 trG_n) + \frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}(\delta_0'\tilde{Z}_{nt}'G_n')\tilde{V}_{nt} \\ \frac{1}{2\sigma_0^4}\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}\left(\tilde{V}_{nt}'\tilde{V}_{nt} - n\sigma_0^2\right) \end{pmatrix} + (R_{\hat{\delta}_{nT}}', R_{\hat{\lambda}_{nT}}, R_{\hat{\sigma}_{nT}^2})'.
\end{aligned}$$

Hence,

$$\sqrt{nT}\begin{pmatrix} \hat{\delta}_{n,T}(\hat{\lambda}_{nT})-\delta_0 \\ \hat{\lambda}_{nT}-\lambda_0 \\ \hat{\sigma}^2_{nT}(\hat{\lambda}_{nT})-\sigma_0^2 \end{pmatrix} = C_{1,nT}\times\begin{pmatrix} \frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\sum_{t=1}^T \tilde{Z}'_{nt}\tilde{V}_{nt} \\ \frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\sum_{t=1}^T(\tilde{V}'_{nt}G'_n\tilde{V}_{nt}-\sigma_0^2 trG_n)+\frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\sum_{t=1}^T(\delta_0'\tilde{Z}'_{nt}G'_n)\tilde{V}_{nt} \\ \frac{1}{2\sigma_0^4}\frac{1}{\sqrt{nT}}\sum_{t=1}^T\left(\tilde{V}'_{nt}\tilde{V}_{nt}-n\sigma_0^2\right) \end{pmatrix}$$

$$+\begin{pmatrix} I_{k_x+2} & \mathcal{H}_{1,nT}^{-1}\mathcal{H}_{2,nT} & \mathbf{0} \\ \mathbf{0} & 1 & 0 \\ \mathbf{0} & \frac{2}{n}\sigma_0^2 trG_n & 1 \end{pmatrix}^{-1}\cdot(R'_{\hat{\delta}_{nT}},R_{\hat{\lambda}_{nT}},R_{\hat{\sigma}^2_{nT}})'$$

where

$$C_{1,nT}=\begin{pmatrix} I_{k_x+2} & \mathcal{H}_{1,nT}^{-1}\mathcal{H}_{2,nT} & \mathbf{0} \\ \mathbf{0} & 1 & 0 \\ \mathbf{0} & \frac{2}{n}\sigma_0^2 trG_n & 1 \end{pmatrix}^{-1}\begin{pmatrix} \mathcal{H}_{1,nT}^{-1}\sigma_0^2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \Sigma_{\lambda_0}^{-1} & 0 \\ \mathbf{0} & 0 & 2\sigma_0^4 \end{pmatrix}\times\begin{pmatrix} I_{k_x+2} & \mathbf{0} & \mathbf{0} \\ -\mathcal{H}'_{2,nT}\mathcal{H}_{1,nT}^{-1} & 1 & -\frac{2}{n}\sigma_0^2 trG_n \\ \mathbf{0} & 0 & 1 \end{pmatrix}$$

$$=\begin{pmatrix} I_{k_x+2} & -\mathcal{H}_{1,nT}^{-1}\mathcal{H}_{2,nT} & \mathbf{0} \\ \mathbf{0} & 1 & 0 \\ \mathbf{0} & -\frac{2}{n}\sigma_0^2 trG_n & 1 \end{pmatrix}\times\begin{pmatrix} \mathcal{H}_{1,nT}^{-1}\sigma_0^2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \Sigma_{\lambda_0}^{-1} & 0 \\ \mathbf{0} & 0 & 2\sigma_0^4 \end{pmatrix}\times\begin{pmatrix} I_{k_x+2} & \mathbf{0} & \mathbf{0} \\ -\mathcal{H}'_{2,nT}\mathcal{H}_{1,nT}^{-1} & 1 & -\frac{2}{n}\sigma_0^2 trG_n \\ \mathbf{0} & 0 & 1 \end{pmatrix}$$

$$=\Sigma_{\theta_0,nT}^{-1}.$$

We note that, from the log likelihood in (2.8), by concentrating out $\mathbf{c}_n$ in terms of $\theta=(\delta',\lambda,\sigma^2)'$, the concentrated likelihood of θ is

$$\ln L_{n,T}(\theta)=-\frac{nT}{2}\ln 2\pi-\frac{nT}{2}\ln\sigma^2+T\ln|S_n(\lambda)|-\frac{1}{2\sigma^2}\sum_{t=1}^T\tilde{V}'_{nt}(\theta)\tilde{V}_{nt}(\theta) \tag{C.23}$$

where $\tilde{V}_{nt}(\theta)=S_n(\lambda)\tilde{Y}_{nt}-\tilde{Z}_{nt}\delta$. It follows that the first derivative of $\ln L_{n,T}(\theta)$ with θ evaluated at θ_0 is

$$\frac{1}{\sqrt{nT}}\frac{\partial\ln L_{nT}(\theta_0)}{\partial\theta}\equiv\begin{pmatrix} \frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\sum_{t=1}^T \tilde{Z}'_{nt}\tilde{V}_{nt} \\ \frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\sum_{t=1}^T(\tilde{V}'_{nt}G'_n\tilde{V}_{nt}-\sigma_0^2 trG_n)+\frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\sum_{t=1}^T(\delta_0'\tilde{Z}'_{nt}G'_n)\tilde{V}_{nt} \\ \frac{1}{2\sigma_0^4}\frac{1}{\sqrt{nT}}\sum_{t=1}^T\left(\tilde{V}'_{nt}\tilde{V}_{nt}-n\sigma_0^2\right) \end{pmatrix}. \tag{C.24}$$

Hence,

$$\sqrt{nT}\begin{pmatrix} \hat{\delta}_{n,T}(\hat{\lambda}_{nT})-\delta_0 \\ \hat{\lambda}_{nT}-\lambda_0 \\ \hat{\sigma}^2_{nT}(\hat{\lambda}_{nT})-\sigma_0^2 \end{pmatrix}=\Sigma_{\theta_0,nT}^{-1}\cdot\frac{1}{\sqrt{nT}}\frac{\partial\ln L_{nT}(\theta_0)}{\partial\theta}+\begin{pmatrix} I_{k_x+2} & \mathcal{H}_{1,nT}^{-1}\mathcal{H}_{2,nT} & \mathbf{0} \\ \mathbf{0} & 1 & 0 \\ \mathbf{0} & \frac{2}{n}\sigma_0^2 trG_n & 1 \end{pmatrix}^{-1}\cdot(R'_{\hat{\delta}_{nT}},R_{\hat{\lambda}_{nT}},R_{\hat{\sigma}^2_{nT}})'. \tag{C.25}$$

As $\mathcal{H}_{1,nT}^{-1}\mathcal{H}_{2,nT}=O_p(1)$ from Proposition B.14 and elements of $R'_{\hat{\delta}_{nT}},R_{\hat{\lambda}_{nT}},R_{\hat{\sigma}^2_{nT}}$ are $O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T},\sqrt{\frac{n}{T^3}}\right)\right)$, we have $\sqrt{nT}\left(\hat{\theta}_{nT}-\theta_0\right)=\Sigma_{\theta_0,nT}^{-1}\cdot\frac{1}{\sqrt{nT}}\frac{\partial\ln L_{nT}(\theta_0)}{\partial\theta}+O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T},\sqrt{\frac{n}{T^3}}\right)\right)$.

C.9 Proof for Theorem 4.1

As $\tilde{Z}_{nt}$ has stationary and nonstationary parts ((2.12)), we can decompose $\frac{1}{\sqrt{nT}}\frac{\partial \ln L_{nT}(\theta_0)}{\partial\theta}$ into two parts:

$$\frac{1}{\sqrt{nT}}\frac{\partial \ln L_{n,T}(\theta_0)}{\partial\theta} = \frac{1}{\sqrt{nT}}\frac{\partial \ln L^s_{n,T}(\theta_0)}{\partial\theta} + \frac{1}{\sqrt{nT}}\frac{\partial \ln L^u_{n,T}(\theta_0)}{\partial\theta} \tag{C.26}$$

where $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^s_{n,T}(\theta_0)}{\partial\theta}$ is the stationary part and $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^u_{n,T}(\theta_0)}{\partial\theta}$ is the nonstationary part as follows. For $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^s_{n,T}(\theta_0)}{\partial\theta}$, it has two parts $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^s_{n,T}(\theta_0)}{\partial\theta} = \frac{1}{\sqrt{nT}}\frac{\partial \ln L^{s*}_{n,T}(\theta_0)}{\partial\theta} - \Delta_{\theta_0,nT}$ where

$$\frac{1}{\sqrt{nT}}\frac{\partial \ln L^{s*}_{nT}(\theta_0)}{\partial\theta} = \begin{pmatrix} \frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\sum_{t=1}^T \ddot{Z}^{s*\prime}_{nt}V_{nt} \\ \frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\sum_{t=1}^T (V'_{nt}G'_nV_{nt} - \sigma_0^2 trG_n) + \frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\sum_{t=1}^T(\delta_0'\ddot{Z}^{s*\prime}_{nt}G'_n)V_{nt} \\ \frac{1}{2\sigma_0^4}\frac{1}{\sqrt{nT}}\sum_{t=1}^T \left(V'_{nt}V_{nt} - n\sigma_0^2\right) \end{pmatrix} \tag{C.27}$$

and

$$\Delta_{\theta_0,nT} = \begin{pmatrix} \frac{1}{\sigma_0^2}\sqrt{\frac{T}{n}}(\bar{U}^s_{nT,-1}, W_n\bar{U}^s_{nT,-1}, \mathbf{0}_{n\times k_x})'\bar{V}_{nT} \\ \frac{1}{\sigma_0^2}\sqrt{\frac{T}{n}}\bar{V}'_{nT}G'_n\bar{V}_{nT} + \frac{1}{\sigma_0^2}\sqrt{\frac{T}{n}}(\delta_0'(\bar{U}^s_{nT,-1}, W_n\bar{U}^s_{nT,-1}, \mathbf{0}_{n\times k_x})'G'_n)\bar{V}_{nT} \\ \frac{1}{2\sigma_0^4}\sqrt{\frac{T}{n}}\bar{V}'_{nT}\bar{V}_{nT} \end{pmatrix}. \tag{C.28}$$

For $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^u_{n,T}(\theta_0)}{\partial\theta}$, it also has two parts $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^u_{n,T}(\theta_0)}{\partial\theta} = \frac{1}{\sqrt{nT}}\frac{\partial \ln L^{u*}_{n,T}(\theta_0)}{\partial\theta} - \blacktriangle_{\theta_0,nT}$ where

$$\frac{1}{\sqrt{nT}}\frac{\partial \ln L^{u*}_{n,T}(\theta_0)}{\partial\theta} = \frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\sum_{t=1}^T \ddot{Y}^{u*\prime}_{n,t-1}V_{nt}\cdot(c^{*\prime},0)' \tag{C.29}$$

with $\ddot{Y}^{u*}_{n,t-1} = \frac{1}{(1-\lambda_0)}(R_nJ_nR_n^{-1})\left(\mathbf{c}_{n0}\tilde{t}_{-1} + \approx{\mathbb{X}}_{n,t-1}\beta_0 + \xi_{n,t-1}\right)$, $\tilde{t}_{-1} = (t-1) - \frac{T-1}{2}$ and

$$\blacktriangle_{\theta_0,nT} = \frac{1}{\sigma_0^2}\left\{\frac{1}{(1-\lambda_0)}\sqrt{\frac{T}{n}}(M_n\bar{\xi}_{n,T-1})'\cdot\bar{V}_{nT}\right\}\cdot(c^{*\prime},0)'. \tag{C.30}$$

Let $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\theta_0)}{\partial\theta} = \frac{1}{\sqrt{nT}}\frac{\partial \ln L^{s*}_{n,T}(\theta_0)}{\partial\theta} + \frac{1}{\sqrt{nT}}\frac{\partial \ln L^{u*}_{n,T}(\theta_0)}{\partial\theta}$. For $\Sigma^{-1}_{\theta_0,nT}\cdot\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\theta_0)}{\partial\theta}$, because $\Sigma^{-1}_{\theta_0,nT}\cdot(c^{*\prime},0)' = O(T^{-1})$ according to Proposition 2.1, it has the form of the CLT in Proposition 2.2 and is normally. For its variance, we can write $E(\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\theta_0)}{\partial\theta}\cdot\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{n,T}(\theta_0)}{\partial\theta'}) =$

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

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

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

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

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

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

As $E\sum_{t=1}^{T}\ddot{Z}_{nt}^*=\mathbf{0}$ and $E\sum_{t=1}^{T}G_n\ddot{Z}_{nt}^*\delta_0=\mathbf{0}$, the second matrix equals

$$\Omega_{\theta_0,n}=\frac{\mu_4-3\sigma_0^4}{\sigma_0^4}\begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{1}{n}\sum_{i=1}^{n}G_{n,ii}^2 & \frac{1}{2\sigma_0^2 n}trG_n \\ \mathbf{0} & \frac{1}{2\sigma_0^2 n}trG_n & \frac{1}{4\sigma_0^4} \end{pmatrix}.$$

When V_{nt} are normally distributed, $\Omega_{\theta_0,n}=0$ because $\mu_4-3\sigma_0^4=0$ for a normal distribution. For the first matrix, premultiplying and postmultiplying it with $\Sigma_{\theta_0,nT}^{-1}$ will yield $\Sigma_{\theta_0,nT}^{-1}+O\left(\frac{1}{T}\right)$. To see this, denote $\Delta_{z,nT}=(\bar{U}_{n,T-1}^s,W_n\bar{U}_{n,T-1}^s,\mathbf{0}_{n\times k_x})$ and $\blacktriangle_{z,nT}=\left(\frac{1}{1-\lambda_0}M_n\bar{\xi}_{n,T-1}\right)\cdot c'$, then, $\tilde{Z}_{nt}=\ddot{Z}_{nt}^*-\Delta_{z,nT}-\blacktriangle_{z,nT}$. This implies that $E(\frac{1}{\sqrt{nT}}\frac{\partial \ln L_{n,T}^*(\theta_0)}{\partial\theta}\cdot\frac{1}{\sqrt{nT}}\frac{\partial \ln L_{n,T}^*(\theta_0)}{\partial\theta'})=\Sigma_{\theta_0,nT}+\Omega_{\theta_0,n}+\Xi_{\theta_0,nT}$ where

$$\Xi_{\theta_0,nT}$$

$$=\begin{pmatrix} \frac{1}{\sigma_0^2 nT}E\sum_{t=1}^{T}\tilde{Z}_{nt}'(\blacktriangle_{z,nT}+\Delta_{z,nT}) & \frac{1}{\sigma_0^2 nT}E\sum_{t=1}^{T}\tilde{Z}_{nt}'G_n(\blacktriangle_{z,nT}+\Delta_{z,nT})\delta_0 & \mathbf{0} \\ \frac{1}{\sigma_0^2 nT}E\sum_{t=1}^{T}(G_n\tilde{Z}_{nt}\delta_0)'(\blacktriangle_{z,nT}+\Delta_{z,nT}) & \frac{2}{\sigma_0^2 nT}E\sum_{t=1}^{T}(G_n\tilde{Z}_{nt}\delta_0)'G_n(\blacktriangle_{z,nT}+\Delta_{z,nT})\delta_0 & 0 \\ \mathbf{0} & 0 & 0 \end{pmatrix}$$

$$+\begin{pmatrix} \frac{1}{\sigma_0^2 nT}E\sum_{t=1}^{T}(\blacktriangle_{z,nT}+\Delta_{z,nT})'\tilde{Z}_{nt} & \frac{1}{\sigma_0^2 nT}E\sum_{t=1}^{T}(\blacktriangle_{z,nT}+\Delta_{z,nT})'G_n\tilde{Z}_{nt}\delta_0 & \mathbf{0} \\ \frac{1}{\sigma_0^2 nT}E\sum_{t=1}^{T}(G_n(\blacktriangle_{z,nT}+\Delta_{z,nT})\delta_0)'\tilde{Z}_{nt} & 0 & 0 \\ \mathbf{0} & 0 & 0 \end{pmatrix}$$

$$+\begin{pmatrix} \frac{1}{\sigma_0^2 nT}E\sum_{t=1}^{T}(\blacktriangle_{z,nT}+\Delta_{z,nT})'(\blacktriangle_{z,nT}+\Delta_{z,nT}) & \frac{1}{\sigma_0^2 nT}E\sum_{t=1}^{T}(\blacktriangle_{z,nT}+\Delta_{z,nT})'G_n(\blacktriangle_{z,nT}+\Delta_{z,nT})\delta_0 & \mathbf{0} \\ \frac{1}{\sigma_0^2 nT}E\sum_{t=1}^{T}(G_n(\blacktriangle_{z,nT}+\Delta_{z,nT})\delta_0)'(\blacktriangle_{z,nT}+\Delta_{z,nT}) & \frac{1}{\sigma_0^2 nT}E\sum_{t=1}^{T}(G_n(\blacktriangle_{z,nT}+\Delta_{z,nT})\delta_0)'G_n(\blacktriangle_{z,nT}+\Delta_{z,nT})\delta_0 & 0 \\ \mathbf{0} & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{\sigma_0^2 nT} E \sum_{t=1}^{T} (\blacktriangle_{z,nT} + \Delta_{z,nT})'(\blacktriangle_{z,nT} + \Delta_{z,nT}) & \frac{1}{\sigma_0^2 nT} E \sum_{t=1}^{T} (\blacktriangle_{z,nT} + \Delta_{z,nT})' G_n (\blacktriangle_{z,nT} + \Delta_{z,nT})\delta_0 & \mathbf{0} \\ \frac{1}{\sigma_0^2 nT} E \sum_{t=1}^{T} (G_n(\blacktriangle_{z,nT} + \Delta_{z,nT})\delta_0)'(\blacktriangle_{z,nT} + \Delta_{z,nT}) & \frac{1}{\sigma_0^2 nT} E \sum_{t=1}^{T} (G_n(\blacktriangle_{z,nT} + \Delta_{z,nT})\delta_0)' G_n (\blacktriangle_{z,nT} + \Delta_{z,nT})\delta_0 & 0 \\ \mathbf{0} & 0 & 0 \end{pmatrix}$$

because the expectations in the first two matrices are all zero.

We have

$$\begin{aligned} & E \frac{1}{\sigma_0^2 nT} E \sum_{t=1}^{T} (\blacktriangle_{z,nT} + \Delta_{z,nT})'(\blacktriangle_{z,nT} + \Delta_{z,nT}) \\ = \;& c \cdot \left(\frac{1}{(1-\lambda_0)^2} \frac{1}{\sigma_0^2 n} E \bar{\xi}'_{n,T-1} M_n' M_n \bar{\xi}_{n,T-1} \right) \cdot c' + \left(\frac{1}{\sigma_0^2 n} E \Delta'_{z,nT} \Delta_{z,nT} \right). \\ & + c \cdot \left(-\frac{1}{(1-\lambda_0)} \frac{1}{\sigma_0^2 n} E \bar{\xi}'_{n,T-1} M_n' \Delta_{z,nT} \right) + \left(-\frac{1}{1-\lambda_0} \frac{1}{\sigma_0^2 n} E \Delta'_{z,nT} M_n \bar{\xi}_{n,T-1} \right) c'. \end{aligned}$$

Similarly we can expand $\frac{1}{\sigma_0^2 nT} E \sum_{t=1}^{T} (\blacktriangle_{z,nT} + \Delta_{z,nT})' G_n (\blacktriangle_{z,nT} + \Delta_{z,nT})\delta_0$ and $\frac{1}{\sigma_0^2 nT} E \sum_{t=1}^{T} (G_n(\blacktriangle_{z,nT} + \Delta_{z,nT})\delta_0)' G_n (\blacktriangle_{z,nT} + \Delta_{z,nT})\delta_0$. Using the orders of relevant terms from Lemma B.8 and Lemma B.10, we have $\frac{1}{n} E \Delta'_{z,nT} \mathcal{B}_n \bar{\xi}_{n,T-1} = O(1)$, $\frac{1}{n} E \Delta'_{z,nT} \mathcal{B}_n \Delta_{z,nT} = O(T^{-1})$ and $\frac{1}{n} E \bar{\xi}'_{n,T-1} \mathcal{B}_n \bar{\xi}_{n,T-1} = O(T)$ where $\mathcal{B}_n$ is a row sum and column sum bounded matrix. Using $\Sigma^{-1}_{\theta_0,nT} \cdot (c^{*\prime}, 0)' = O(T^{-1})$ and $(c^{*\prime}, 0) \cdot \Sigma^{-1}_{\theta_0,nT} \cdot (c^{*\prime}, 0)' = O(T^{-2})$ (see Proposition 2.1) and the above, it follows that $\Sigma^{-1}_{\theta_0,nT} \cdot \Xi_{\theta_0,nT} \cdot \Sigma^{-1}_{\theta_0,nT} = O(T^{-1})$. Hence,

$$\Sigma^{-1}_{\theta_0,nT} E\Big(\frac{1}{\sqrt{nT}} \frac{\partial \ln L^*_{n,T}(\theta_0)}{\partial \theta} \cdot \frac{1}{\sqrt{nT}} \frac{\partial \ln L^*_{n,T}(\theta_0)}{\partial \theta'}\Big) \Sigma^{-1}_{\theta_0,nT} = \Sigma^{-1}_{\theta_0,nT} + \Sigma^{-1}_{\theta_0,nT} \Omega_{\theta_0,n} \Sigma^{-1}_{\theta_0,nT} + O\left(T^{-1}\right). \quad \text{(C.31)}$$

Therefore,

$$\Sigma^{-1}_{\theta_0,nT} \left(\frac{1}{\sqrt{nT}} \frac{\partial \ln L^*_{n,T}(\theta_0)}{\partial \theta'} \right) \xrightarrow{p} N(0, \lim_{T\to\infty} \Sigma^{-1}_{\theta_0,nT} + \lim_{T\to\infty} \Sigma^{-1}_{\theta_0,nT} \Omega_{\theta_0,n} \Sigma^{-1}_{\theta_0,nT}) \quad \text{(C.32)}$$

Using the results in Yu, de Jong and Lee (2006) (Theorem A.11, page 19), we have $\Delta_{\theta_0,nT} = \sqrt{\frac{n}{T}} a^s_{\theta_0,n} + O_p\left(\max\left(\sqrt{\frac{n}{T^3}}, \frac{1}{\sqrt{T}}\right)\right)$. Using (B.17) and (C.30), we have $\blacktriangle_{\theta_0,nT} = \sqrt{\frac{n}{T}} \cdot \frac{m_n}{n} \cdot a^u_{\theta_0,T} + T \cdot (c^{*\prime}, 0) \cdot O_p\left(\max\left(\sqrt{\frac{n}{T^3}}, \frac{1}{\sqrt{T}}\right)\right)$ where

$$\begin{aligned} a^s_{\theta_0,n} &= \begin{pmatrix} \frac{1}{n} tr\left(\left(\sum_{h=0}^{\infty} B_n^h\right) S_n^{-1}\right) \\ \frac{1}{n} tr\left(W_n \left(\sum_{h=0}^{\infty} B_n^h\right) S_n^{-1}\right) \\ \mathbf{0} \\ \frac{1}{n}\gamma_0 tr(G_n \left(\sum_{h=0}^{\infty} B_n^h\right) S_n^{-1}) + \frac{1}{n}\rho_0 tr(G_n W_n \left(\sum_{h=0}^{\infty} B_n^h\right) S_n^{-1}) + \frac{1}{n} tr G_n \\ \frac{1}{2\sigma_0^2} \end{pmatrix} \quad \text{(C.33)} \\ a^u_{\theta_0,T} &= T \cdot \frac{1}{2(1-\lambda_0)} \cdot (c^{*\prime}, 0)'. \end{aligned}$$

Hence,

$$\Delta_{\theta_0,nT} + \blacktriangle_{\theta_0,nT} = \sqrt{\frac{n}{T}} \cdot (a^s_{\theta_0,n} + a^u_{\theta_0,T} \frac{m_n}{n}) + T \cdot (c^{*\prime}, 0) \cdot O_p\left(\max\left(\sqrt{\frac{n}{T^3}}, \frac{1}{\sqrt{T}}\right)\right) + O_p\left(\max\left(\sqrt{\frac{n}{T^3}}, \frac{1}{\sqrt{T}}\right)\right). \quad \text{(C.34)}$$

Using $\Sigma^{-1}_{\theta_0,nT} \cdot (c^{*\prime}, 0)' = O(T^{-1})$, we have $\Sigma^{-1}_{\theta_0,nT} \left(\Delta_{\theta_0,nT} + \blacktriangle_{\theta_0,nT}\right) = \sqrt{\frac{n}{T}} \cdot b_{\theta_0,nT} + O_p\left(\max\left(\sqrt{\frac{n}{T^3}}, \frac{1}{\sqrt{T}}\right)\right)$.

Hence, combining (4.1), (C.32) and the equation above, we have the result. ■

C.10 Proof for Theorem 4.2

For the remainder term in (C.25), using (C.20), (C.22) and $1-c'\mathcal{H}_{1,nT}^{-1}\mathcal{H}_{2,nT}=O(T^{-1})$ from Proposition B.14, we have

$$(c^{*\prime},0)'\begin{pmatrix} I_{k_x+2} & \mathcal{H}_{1,nT}^{-1}\mathcal{H}_{2,nT} & \mathbf{0} \\ \mathbf{0} & 1 & 0 \\ \mathbf{0} & \frac{2}{n}\sigma_0^2 trG_n & 1 \end{pmatrix}^{-1}\begin{pmatrix} R_{\hat{\delta}_{nT}} \\ R_{\hat{\lambda}_{nT}} \\ R_{\hat{\sigma}^2_{nT}} \end{pmatrix}=(c^{*\prime},0)'\begin{pmatrix} I_{k_x+2} & -\mathcal{H}_{1,nT}^{-1}\mathcal{H}_{2,nT} & \mathbf{0} \\ \mathbf{0} & 1 & 0 \\ \mathbf{0} & -\frac{2}{n}\sigma_0^2 trG_n & 1 \end{pmatrix}\begin{pmatrix} R_{\hat{\delta}_{nT}} \\ R_{\hat{\lambda}_{nT}} \\ R_{\hat{\sigma}^2_{nT}} \end{pmatrix}$$
$$=c'R_{\hat{\delta}_{nT}}+(1-c'\mathcal{H}_{1,nT}^{-1}\mathcal{H}_{2,nT})R_{\hat{\lambda}_{nT}}=O_p\left(\max\left(\frac{1}{T^2},\frac{1}{\sqrt{nT^3}},\sqrt{\frac{n}{T^5}}\right)\right).$$

Hence,

$$\begin{aligned}
&\sqrt{nT^3}(c^{*\prime},0)(\hat{\theta}_{nT}-\theta_0) \\
=\ & T(c^{*\prime},0)\Sigma_{\theta_0,nT}^{-1}\cdot\frac{1}{\sqrt{nT}}\frac{\partial\ln L_{nT}(\theta_0)}{\partial\theta}+O_p\left(\max\left(\frac{1}{T},\frac{1}{\sqrt{nT}},\sqrt{\frac{n}{T^3}}\right)\right) \\
=\ & T(c^{*\prime},0)\Sigma_{\theta_0,nT}^{-1}\cdot\frac{1}{\sqrt{nT}}\frac{\partial\ln L^{s*}_{nT}(\theta_0)}{\partial\theta}+T(c^{*\prime},0)\Sigma_{\theta_0,nT}^{-1}\cdot\frac{1}{\sqrt{nT}}\frac{\partial\ln L^{u*}_{nT}(\theta_0)}{\partial\theta} \\
& -T(c^{*\prime},0)\Sigma_{\theta_0,nT}^{-1}(\Delta_{\theta_0,nT}+\blacktriangle_{\theta_0,nT})+O_p\left(\max\left(\frac{1}{T},\frac{1}{\sqrt{nT}},\sqrt{\frac{n}{T^3}}\right)\right)
\end{aligned}\tag{C.35}$$

where $\frac{1}{\sqrt{nT}}\frac{\partial\ln L^{s*}_{nT}(\theta_0)}{\partial\theta}$ is in (C.27), $\frac{1}{\sqrt{nT}}\frac{\partial\ln L^{u*}_{nT}(\theta_0)}{\partial\theta}$ is in (C.29), $\Delta_{\theta_0,nT}$ is in (C.28) and $\blacktriangle_{\theta_0,nT}$ is in (C.30). We shall investigate the orders of those terms.

For stationary terms, from Yu, de Jong and Lee (2006) (Claim 3.4, page 10), $\frac{1}{\sqrt{nT}}\frac{\partial\ln L^{s*}_{nT}(\theta_0)}{\partial\theta}$ has the typical $O_p(1)$ and $\Delta_{\theta_0,nT}-E(\Delta_{\theta_0,nT})=O_p(\frac{1}{\sqrt{T}})$ where $E(\Delta_{\theta_0,nT})=O(\sqrt{\frac{n}{T}})$. For the nonstationary term $T(c^{*\prime},0)\Sigma_{\theta_0,nT}^{-1}\blacktriangle_{\theta_0,nT}=T(c^{*\prime},0)\Sigma_{\theta_0,nT}^{-1}(c^{*\prime},0)\left\{\frac{1}{\sigma_0^2(1-\lambda_0)}\sqrt{\frac{T}{n}}(M_n\bar{\xi}_{n,T-1})'\cdot\bar{V}_{nT}\right\}$, we have $T(c^{*\prime},0)\Sigma_{\theta_0,nT}^{-1}\blacktriangle_{\theta_0,nT}-E\left(T(c^{*\prime},0)\Sigma_{\theta_0,nT}^{-1}\blacktriangle_{\theta_0,nT}\right)=O_p(\frac{1}{\sqrt{T}})$ where $E\left(T(c^{*\prime},0)\Sigma_{\theta_0,nT}^{-1}\blacktriangle_{\theta_0,nT}\right)=O(\sqrt{\frac{n}{T}})$ by using (B.17) in Lemma B.9 and $(c^{*\prime},0)\Sigma_{\theta_0,nT}^{-1}(c^{*\prime},0)'=O(T^{-2})$. For nonstationary term $T(c^{*\prime},0)\Sigma_{\theta_0,nT}^{-1}\cdot\frac{1}{\sqrt{nT}}\frac{\partial\ln L^{u*}_{nT}(\theta_0)}{\partial\theta}$, we have

$$\frac{1}{\sqrt{nT}}\frac{\partial\ln L^{u*}_{nT}(\theta_0)}{\partial\theta}=\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}\frac{1}{\sigma_0^2(1-\lambda_0)}M_n\left(\mathbf{c}_{n0}\tilde{t}_{-1}+\tilde{\tilde{\mathbb{X}}}_{n,t-1}\beta_0+\xi_{n,t-1}\right)'V_{nt}\cdot(c^{*\prime},0)'.$$

Hence,

$$\begin{aligned}
&\sqrt{nT^3}(c^{*\prime},0)(\hat{\theta}_{nT}-\theta_0) \\
=\ & T(c^{*\prime},0)\Sigma_{\theta_0,nT}^{-1}\cdot\frac{1}{\sqrt{nT}}\frac{\partial\ln L^{s*}_{nT}(\theta_0)}{\partial\theta} \\
& +T^2(c^{*\prime},0)\Sigma_{\theta_0,nT}^{-1}(c^{*\prime},0)'\cdot\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}\frac{1}{\sigma_0^2(1-\lambda_0)}M_n\frac{1}{T}\left(\mathbf{c}_{n0}\tilde{t}_{-1}+\tilde{\tilde{\mathbb{X}}}_{n,t-1}\beta_0+\xi_{n,t-1}\right)'V_{nt} \\
& -T(c^{*\prime},0)\Sigma_{\theta_0,nT}^{-1}E(\Delta_{\theta_0,nT}+\blacktriangle_{\theta_0,nT})+O_p(\frac{1}{\sqrt{T}})+O_p\left(\max\left(\frac{1}{T},\frac{1}{\sqrt{nT}},\sqrt{\frac{n}{T^3}}\right)\right)
\end{aligned}$$

where $T(c^{*\prime},0)\Sigma_{\theta_0,nT}^{-1}E(\Delta_{\theta_0,nT}+\blacktriangle_{\theta_0,nT})=O(\sqrt{\frac{n}{T}})$ represents the asymptotic bias term and the first two terms will be asymptotically jointly normally distributed. As $(c^{*\prime},0)(\hat{\theta}_{nT}-\theta_0)=(\hat{\lambda}_{nT}+\hat{\gamma}_{nT}+\hat{\rho}_{nT}-1)$, the rate of convergence of $\hat{\lambda}_{nT}+\hat{\gamma}_{nT}+\hat{\rho}_{nT}$ to the unit is of higher order $O(\frac{1}{\sqrt{nT^3}})$ as long as $\frac{n}{T^3}\to 0$.

So, we have

$$\sqrt{nT^3}(c^{*\prime},0)(\hat{\theta}_{nT}-\theta_0)+\sqrt{\frac{n}{T}}T(c^{*\prime},0)b_{\theta_0,nT}+O_p\left(\max\left(\frac{1}{\sqrt{T}},\sqrt{\frac{n}{T^3}}\right)\right)$$
$$\xrightarrow{d} N\left(0,\lim_{T\to\infty}T^2(c^{*\prime},0)(\Sigma^{-1}_{\theta_0,nT}+\lim_{T\to\infty}\Sigma^{-1}_{\theta_0,nT}\Omega_{\theta_0}\Sigma^{-1}_{\theta_0,nT})(c^{*\prime},0)'\right). \quad \text{(C.36)}$$

Also, using Proposition 2.1, $\lim_{T\to\infty}T^2(c^{*\prime},0)\Sigma^{-1}_{\theta_0,nT}(c^{*\prime},0)'=\lim_{T\to\infty}\omega^{-1}_{nT}$, where ω_{nT} is defined in (2.13).

C.11 Proof for Theorem 4.3

From the first order condition that $\frac{\partial \ln L_{n,T}(\theta,\mathbf{c}_n)}{\partial \mathbf{c}_n}=\frac{1}{\sigma^2}\sum_{t=1}^{T}V_{nt}(\zeta)=0$, 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,

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

As $\frac{1}{T}\sum_{t=1}^{T}\left[Z^u_{nt}\left(\delta-\delta_0\right)+(\lambda-\lambda_0)\left(G_nZ^u_{nt}\delta_0\right)\right]=\left(\frac{1}{T}\sum_{t=1}^{T}Y^u_{n,t-1}\right)(\gamma+\rho+\lambda-1)$, we have
$\frac{1}{T}\sum_{t=1}^{T}\left[Z^u_{nt}\left(\hat{\delta}_{nT}-\delta_0\right)+\left(\hat{\lambda}_{nT}-\lambda_0\right)\left(G_nZ^u_{nt}\delta_0\right)\right]=\left(\frac{1}{T^2}\sum_{t=1}^{T}Y^u_{n,t-1}\right)T\cdot(\hat{\gamma}_{nT}+\hat{\rho}_{nT}+\hat{\lambda}_{nT}-1)$.

From Theorem 4.2, $T\cdot(\hat{\gamma}_{nT}+\hat{\rho}_{nT}+\hat{\lambda}_{nT}-1)=O_p\left(\max\left(\frac{1}{T},\frac{1}{\sqrt{nT}}\right)\right)$. From (2.4), elements of $\frac{1}{T^2}\sum_{t=1}^{T}Y^u_{n,t-1}$ are $O_p(1)$ if elements of $\frac{Y_{n,-1}}{T}$ are $O_p(1)$. Then, for each fixed effect, we have

$$\begin{aligned}\hat{c}_{i,nT}(\hat{\theta}_{nT})-c_{i,0} &= -\frac{1}{T}\sum_{t=1}^{T}((G_n\mathbf{c}_{n0}+G_nZ^s_{nt}\delta_0)_i\,,(Z^s_{nt})_i)\times\begin{pmatrix}\hat{\lambda}_{nT}-\lambda_0\\ \hat{\delta}_{nT}-\delta_0\end{pmatrix}+\frac{1}{T}\sum_{t=1}^{T}\left\{\left(I_n-(\hat{\lambda}_{nT}-\lambda_0)G_n\right)V_{nt}\right\}_i\\
&\quad +O_p\left(\max\left(\frac{1}{T},\frac{1}{\sqrt{nT}}\right)\right),\end{aligned} \quad \text{(C.37)}$$

where $(Z^s_{nt})_i$ is the ith row of Z^s_{nt} and $(G_n\mathbf{c}_{n0}+G_nZ^s_{nt}\delta_0)_i$ is the ith element of $(G_n\mathbf{c}_{n0}+G_nZ^s_{nt}\delta_0)$. As elements of $\frac{1}{T}\sum_{t=1}^{T}((G_n\mathbf{c}_{n0}+G_nZ^s_{nt}\delta_0)_i,\ (Z^s_{nt})_i)$ are $O_p(1)$ uniformly in n and i implied by Lemma B.4 of Yu, de Jong and Lee (2006) and $\hat{\theta}_{nT}-\theta_0=O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right)$ by Theorem 3.6, the dominant term of $\sqrt{T}(\hat{c}_{i,nT}(\hat{\theta}_{nT})-c_{i,0})$ would be $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}v_{it}+O_p\left(\frac{1}{\sqrt{n}}\right)$ when $T\to\infty$ where the $O_p\left(\frac{1}{\sqrt{n}}\right)$ term is the

$\left(\frac{1}{T^2}\sum_{t=1}^{T} Y^u_{n,t-1}\right)$ term multiplied by the distribution part of $T\cdot(\hat{\gamma}_{nT}+\hat{\rho}_{nT}+\hat{\lambda}_{nT}-1)$, which is $T(c^{*\prime},0)\Sigma^{-1}_{\theta_0,nT}\cdot \frac{1}{n\sqrt{T}}\frac{\partial \ln L^*_{nT}(\theta_0)}{\partial\theta}$ (see (C.35)). So, for each fixed effect,

$$\sqrt{T}\left(\hat{c}_{i,nT}(\hat{\theta}_{nT})-c_{i,0}\right)=\frac{1}{\sqrt{T}}\sum_{t=1}^{T} v_{it}+\frac{1}{\sqrt{n}}\left(\frac{1}{T^2}\sum_{t=1}^{T} Y^u_{n,t-1}\right)_i\left([T(c^{*\prime},0)\Sigma^{-1}_{\theta_0,nT}]\times\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{nT}(\theta_0)}{\partial\theta}\right)+O_p\left(\frac{1}{\sqrt{T}}\right) \tag{C.38}$$

where $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{nT}(\theta_0)}{\partial\theta}=\frac{1}{\sqrt{nT}}\frac{\partial \ln L^{s*}_{nT}(\theta_0)}{\partial\theta}+\frac{1}{\sqrt{nT}}\frac{\partial \ln L^{u*}_{nT}(\theta_0)}{\partial\theta}$ (defined in (C.27) and (C.29)) and $[T(c^{*\prime},0)\Sigma^{-1}_{\theta_0,nT}]\times \frac{1}{\sqrt{nT}}\frac{\partial \ln L^*(\theta_0)}{\partial\theta}$ is normally distributed asymptotically with the variance specified in (4.11). Hence, $\sqrt{T}\left(\hat{c}_{i,nT}(\hat{\theta}_{nT})-c_{i,0}\right)$ is a linear and quadratic form of V_{nt} and it will be normally distributed asymptotically using the central limit theorem by Proposition 2.2. We need to calculate its variance.

Under the assumption that $(Y_{n,-1}/T)_i-E(Y_{n,-1}/T)_i=o_p(1)$ and $E(Y_{n,-1}/T)_i=O(1)$ uniformly in n and i, we have $\frac{1}{T^2}\sum_{t=1}^{T}(Y^u_{n,t-1})_i=E\frac{1}{T^2}\sum_{t=1}^{T}(Y^u_{n,t-1})_i+o_p(1)$ where $E\frac{1}{T^2}\sum_{t=1}^{T}(Y^u_{n,t-1})_i$ is $O(1)$. Hence,

$$\sqrt{T}\left(\hat{c}_{i,nT}(\hat{\theta}_{nT})-c_{i,0}\right)=\frac{1}{\sqrt{T}}\sum_{t=1}^{T} v_{it}+\frac{1}{\sqrt{n}}\left(E\frac{1}{T^2}\sum_{t=1}^{T} Y^u_{n,t-1}\right)_i\left([T(c^{*\prime},0)\Sigma^{-1}_{\theta_0,nT}]\times\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{nT}(\theta_0)}{\partial\theta}\right)+o_p(1). \tag{C.39}$$

As $\frac{1}{\sqrt{nT}}\frac{\partial \ln L^*_{nT}(\theta_0)}{\partial\theta}=\begin{pmatrix}\frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}\ddot{Z}^{*\prime}_{nt}V_{nt}\\ \frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}(V'_{nt}G'_nV_{nt}-\sigma_0^2 trG_n)+\frac{1}{\sigma_0^2}\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}(\delta_0'\ddot{Z}^{*\prime}_{nt}G'_n)V_{nt}\\ \frac{1}{2\sigma_0^4}\frac{1}{\sqrt{nT}}\sum_{t=1}^{T}\left(V'_{nt}V_{nt}-n\sigma_0^2\right)\end{pmatrix}$ where $\ddot{Z}^*_{nt}$ is defined in (C.10), the asymptotic variance of $\sqrt{T}\left(\hat{c}_{i,nT}(\hat{\theta}_{nT})-c_{i,0}\right)$ would be Φ_{n,c_i} where

$$\begin{aligned}\Phi_{n,c_i} \;=\;& \sigma_0^2+\frac{2}{n}\left(E\frac{1}{T^2}\sum_{t=1}^{T} Y^u_{n,t-1}\right)_i\left(\mu_3\left([T(c^{*\prime},0)\Sigma^{-1}_{\theta_0,nT}]\cdot[0,G_{ii},1]'\right)\right)\\ &+\frac{2}{n}\left(E\frac{1}{T^2}\sum_{t=1}^{T} Y^u_{n,t-1}\right)_i\left(\sigma_0^2\left([T(c^{*\prime},0)\Sigma^{-1}_{\theta_0,nT}]\cdot[(\sum_{t=1}^{T}E\ddot{Z}^*_{nt},)_i,(\sum_{t=1}^{T}EG_n\ddot{Z}^*_{nt}\delta_0,)_i,0]'\right)\right)\\ &+\frac{1}{n}\left(\left(E\frac{1}{T^2}\sum_{t=1}^{T} Y^u_{n,t-1}\right)_i^2\left(\lim_{T\to\infty}\omega^{-1}_{nT}+\lim_{T\to\infty}T^2(c^{*\prime},0)(\lim_{T\to\infty}\Sigma^{-1}_{\theta_0,nT}\Omega_{\theta_0,nT}\Sigma^{-1}_{\theta_0,nT})(c^{*\prime},0)'\right)\right).\end{aligned} \tag{C.40}$$

When $n\to\infty$, we have $\Phi_{n,c_i}\to\sigma_0^2$. ■

C.12 Proof for Theorem 4.5

Theorem 4.1 states that $\sqrt{nT}(\hat{\theta}_{nT}-\theta_0)+\sqrt{\frac{n}{T}}b_{\theta_0,nT}+O_p\left(\max\left(\frac{1}{T},\sqrt{\frac{n}{T^3}}\right)\right)\xrightarrow{d}N(0,\lim_{T\to\infty}\Sigma^{-1}_{\theta_0,nT}+\lim_{T\to\infty}\Sigma^{-1}_{\theta_0,nT}\Omega_{\theta_0,n}\Sigma^{-1}_{\theta_0,nT})$. As the bias corrected estimator $\hat{\theta}^1_{nT}=\hat{\theta}_{nT}+\frac{1}{T}b_{\hat{\theta}_{nT},nT}$, we have $\sqrt{nT}(\hat{\theta}^1_{nT}-\theta_0)\xrightarrow{d}N(0,\lim_{T\to\infty}\Sigma^{-1}_{\theta_0,nT}+\lim_{T\to\infty}\Sigma^{-1}_{\theta_0,nT}\Omega_{\theta_0,n}\Sigma^{-1}_{\theta_0,nT})$ if $\sqrt{\frac{n}{T}}\left(b_{\hat{\theta}_{nT},nT}-b_{\theta_0,nT}\right)\xrightarrow{p}0$ and $\frac{n}{T^3}\to 0$. So, given $\frac{n}{T^3}\to 0$, we are going to prove that $\sqrt{\frac{n}{T}}\left(b_{\hat{\theta}_{nT},nT}-b_{\theta_0,nT}\right)\xrightarrow{p}0$ where $b_{\theta_0,nT}=\Sigma^{-1}_{\theta_0,nT}\cdot\left(a^s_{\theta_0,n}+a^u_{\theta_0,T}\frac{m_n}{n}\right)$ and $b_{\hat{\theta}_{nT},nT}=\ddot{\Sigma}^{-1}_{\hat{\theta}_{nT},nT}\cdot\left(a^s_{\hat{\theta}_{nT},n}+a^u_{\hat{\theta}_{nT},T}\frac{m_n}{n}\right)$. As $\Sigma^{-1}_{\theta_0,nT}=\ddot{\Sigma}^{-1}_{\hat{\theta}_{nT},nT}+O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right)$

and $T\cdot\left[\ddot{\Sigma}^{-1}_{\hat{\theta}_{nT},nT}-\Sigma^{-1}_{\theta_0,nT}\right]\cdot(c^{*\prime},0)'=O_p\left(\max\left(\frac{1}{\sqrt{nT}},\frac{1}{T}\right)\right)$ from Proposition B.15, $\sqrt{\frac{n}{T}}\left(b_{\hat{\theta}_{nT},nT}-b_{\theta_0,nT}\right)\xrightarrow{p}0$ is reduced to

$$\sqrt{\frac{n}{T}}\left(\Sigma^{-1}_{\theta_0,nT}a^u_{\hat{\theta}_{nT},n}-\Sigma^{-1}_{\theta_0,nT}a^u_{\theta_0,T}\right)\xrightarrow{p}0 \tag{C.41}$$

and

$$\sqrt{\frac{n}{T}}\left(a^s_{\hat{\theta}_{nT},n}-a^s_{\theta_0,n}\right)\xrightarrow{p}0. \tag{C.42}$$

For (C.41), as $a^u_{\theta_0,nT}=T\frac{1}{2(1-\lambda_0)}(c^{*\prime},0)'$ with $\Sigma^{-1}_{\theta_0,nT}\cdot(c^{*\prime},0)'=O(T^{-1})$, $\sqrt{\frac{n}{T}}\left(\Sigma^{-1}_{\theta_0,nT}a^u_{\hat{\theta}_{nT},n}-\Sigma^{-1}_{\theta_0,nT}a^u_{\theta_0,n}\right)=\sqrt{\frac{n}{T}}\left(T\cdot\Sigma^{-1}_{\theta_0,nT}(c^{*\prime},0)'\right)\cdot\left(\frac{1}{2(1-\hat{\lambda}_{nT})}-\frac{1}{2(1-\lambda_0)}\right)=\sqrt{\frac{n}{T}}\left(T\cdot\Sigma^{-1}_{\theta_0,nT}(c^{*\prime},0)'\right)\cdot\left(\frac{\hat{\lambda}_{nT}-\lambda_0}{2(1-\hat{\lambda}_{nT})(1-\lambda_0)}\right)$. As $\hat{\lambda}_{nT}-\lambda_0=O_p\left(\max\left(\frac{1}{T},\frac{1}{\sqrt{nT}}\right)\right)$, we have $\sqrt{\frac{n}{T}}\left(\Sigma^{-1}_{\theta_0,nT}a^u_{\hat{\theta}_{nT},n}-\Sigma^{-1}_{\theta_0,nT}a^u_{\theta_0,n}\right)\xrightarrow{p}0$ if $\frac{n}{T^3}\to0$.

For (C.42), as $\hat{\theta}_{nT}-\theta_0=O_p\left(\max\left(\frac{1}{T},\frac{1}{\sqrt{nT}}\right)\right)$ and $a^s_n(\theta_0)$ is $O(1)$ where $a^s_n(\theta_0)\equiv a^s_{\theta_0,n}$, according to the Taylor expansion of $a^s_n(\hat{\theta}_{nT})$ around $a^s_n(\theta_0)$, to prove (C.42) is reduced to proving that elements of $\frac{\partial a^s_n(\bar{\theta}_{nT})}{\partial\theta'}$ are $O(1)$ where $\bar{\theta}_{nT}$ lies between $\hat{\theta}_{nT}$ and θ_0 and

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

From Proposition B.2, for $A_n(\theta)=(I_n-\lambda W_n)^{-1}(\gamma I_n+\rho W_n)$ where W_n is diagonalizable as $W_n=R_nD^*_nR^{-1}_n$, we have that $A_n(\theta)$ is diagonalizable as $A_n(\theta)=R_nD_n(\theta)R^{-1}_n$, with its eigenvalue matrix $D_n(\theta)=(I_n-\lambda D^*_n)^{-1}(\gamma I_n+\rho D^*_n)$. As $B_n(\theta)=R_n\tilde{D}_n(\theta)R^{-1}_n$ with $\tilde{D}_n(\theta)=Diag(0,\cdots,0,d_{n,m_n+1},\cdots,d_{nn})$ so that $D_n(\theta)=J_n+\tilde{D}_n(\theta)$ where $J_n=Diag\{1'_{m_n},0,\cdots,0\}$, we have

$$B_n(\theta)=R_n\left(D_n(\theta)-J_n\right)R^{-1}_n=R_n(I_n-\lambda D^*_n)^{-1}(\gamma I_n+\rho D^*_n)R^{-1}_n-R_nJ_nR^{-1}_n.$$

With $B_n(\theta)$ as a function explicitly in θ, $\frac{\partial B_n(\theta)}{\partial\theta'}$ can be easily evaluated. Because $\frac{\partial B^h_n(\theta)}{\partial\theta'}=hB^{h-1}_n(\theta)\frac{\partial B_n(\theta)}{\partial\theta'}$ for $h\geq1$ (see footnote 9 in Yu, de Jong and Lee (2006)), we have $\sum_{h=1}^{\infty}\frac{\partial B^h_n(\theta)}{\partial\theta'}=\sum_{h=1}^{\infty}hB^{h-1}_n(\theta)\frac{\partial B_n(\theta)}{\partial\theta'}$. As (1) $\sum_{h=0}^{\infty}B^h_n(\theta)$ and $\sum_{h=1}^{\infty}hB^{h-1}_n(\theta)$ are uniformly bounded in either row sum or column sum, uniformly in a neighborhood of θ_0, (2) $S^{-1}_n(\lambda)$ 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 $\frac{\partial a^s_n(\theta)}{\partial\theta'}$ will be uniformly bounded in n in a neighborhood of θ_0. As $\bar{\theta}_{nT}$ converges in probability to θ_0, we conclude that elements of $\frac{\partial a^s_n(\bar{\theta}_{nT})}{\partial\theta'}$ are $O_p(1)$.

For (4.15), we can start from (4.11). Similarly, we can prove $\sqrt{\frac{n}{T}}T(c^{*\prime},0)\left(b_{\hat{\theta}^1_{nT},nT}-b_{\theta_0,nT}\right)\xrightarrow{p}0$. Hence, $\sqrt{nT^3}(c^{*\prime},0)(\hat{\theta}^1_{nT}-\theta_0)\xrightarrow{d}N\left(0,\lim_{T\to\infty}T^2(c^{*\prime},0)\left(\Sigma^{-1}_{\theta_0,nT}+\Sigma^{-1}_{\theta_0,nT}\Omega_{\theta_0,n}\Sigma^{-1}_{\theta_0,nT}\right)(c^{*\prime},0)'\right)$ under $\frac{n}{T^3}\to0$, where $\lim_{T\to\infty}T^2(c^{*\prime},0)\Sigma^{-1}_{\theta_0,nT}(c^{*\prime},0)'=\lim_{T\to\infty}\omega^{-1}_{nT}$ using Proposition 2.1. ■

References

[1] Dhrymes, P (1978), *Mathematics for Econometrics*, Springer-Verlag.

[2] Choi, I. (2004), Nonstationary Panels, *Palgrave Handbooks of Econometrics*, Vol. 1, forthcoming (Jul., 2004).

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

[4] Im, K.S., M.H. Pesaran, S. Shin (2003), Testing for Unit Roots in Heterogeneous Panels, *Journal of Econometrics*, 115, 53-74

[5] 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.

[6] 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.

[7] Lee, L.F. (2001), Asymptotic Distributions of Quasi-Maximum Likelihood Estimators for Spatial Econometric Models I: Spatial Autoregressive Process, Working Paper, The Ohio State University.

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

[9] Levin, A., C-F. Lin and C-S.J. Chu (2002), Unit Root tests in Panel Data: Asymptotic and Finite Sample Properties, *Journal of Econometrics,* 108, 1-24.

[10] Maddala, G.S. and S. Wu (1999), A Comparative Study of Unit Root Tests With Panel Data and A New Simple Test, *Oxford Bulletin of Economics and Statistics*, 61, 631-652.

[11] Moon, H.R. and B. Perron (2004), Testing for A Unit Root in Panels With Dynamic Factors, *Journal of Econometrics*, 122, 81-126.

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

[13] Pesaran, M.H. (2003), A Simple Panel Unit Root Test in the Presence of Cross Section Dependence, *Mimeo*, Trinity College, Cambridge.

[14] Phillips, P.C.B. and D. Sul (2003), Dynamic Panel Estimation and Homogeneity Testing Under Cross Section Dependence, *Econometrics Journal*, 6, 217-259.

[15] Tao, J. (2006), Analyzing Local School Expenditure in A Dynamic Game, Working Paper, Shanghai University of Finance and Economics.

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

[17] Yu, J. (2006), Convergence: A Spatial Dynamic Panel Data Approach, Working Paper, The Ohio State University.

[18] Yu, J., R. de Jong and L-F. Lee (2006), Quasi-Maximum Likelihood Estimators For Spatial Dynamic Panel Data With Fixed Effects When Both n and T Are Large, Working Paper, The Ohio State University.