Supplement to "QML Estimation of the Matrix Exponential Spatial Specification Panel Data Model with Fixed Effects and Heteroskedasticity"

February 17, 2019

Yuanqing Zhang[a,b], Shuhui Feng[c], Fei Jin[d,1]

[a]School of Business, Shanghai University of International Business and Economics, Shanghai, 201620, China

[b]School of Economics, Peking University, Beijing, 100871, China

[c]Institute of Economics, Shanghai Academy of Social Sciences, Shanghai, 200020, China

[d]School of Economics, Fudan University, Shanghai, 200433, China

1 Regularity assumptions

We define $\bar{\Gamma}_{nT}(\eta) = \min_\beta E[\Gamma_{nT}(\zeta)]$ with $\eta = [\alpha, \tau]'$ and $\tilde{\mathbf{X}}_{nT} = [\tilde{X}'_{n1}, \ldots, \tilde{X}'_{nT}]'$. Then,

$$\begin{aligned}\bar{\Gamma}_{nT}(\eta) =& (\tilde{\mathbf{X}}_{nT}\beta_0)'(I_T \otimes W_n' e^{\tau_0 M_n'}) H_{nT}(\tau)(I_T \otimes e^{\tau_0 M_n} W_n)\tilde{\mathbf{X}}_{nT}\beta_0 \\ &+ tr\big[(J_T \otimes e^{-\tau_0 M_n'} e^{(\alpha-\alpha_0)W_n'} e^{\tau M_n'} e^{\tau M_n} e^{(\alpha-\alpha_0)W_n} e^{-\tau_0 M_n})\Sigma_{nT}\big].\end{aligned} \tag{A.1}$$

where $H_{nT}(\tau) = I_{nT} - (I_T \otimes e^{\tau M_n})\tilde{\mathbf{X}}_{nT}[\tilde{\mathbf{X}}'_{nT}(I_T \otimes e^{\tau M_n'} e^{\tau M_n})\tilde{\mathbf{X}}_{nT}]^{-1}\tilde{\mathbf{X}}'_{nT}(I_T \otimes e^{\tau M_n'})$ is a projection matrix. In addition to Assumptions 1–2 in the main text, we make the following regularity assumptions for the analysis on asymptotic properties of the QMLE.

Assumption A.1. *The v_{it}'s, $i = 1, 2, \cdots, n$ and $t = 1, 2, \cdots, T$, are independent $(0, \sigma_{it}^2)$, and the moments $E|v_{it}|^{4+\eta} \le k < \infty$ for some $\eta > 0$ are uniformly bounded for all i and t.*

Assumption A.2. *There exists a constant $\delta > 0$ such that $|\alpha| \le \delta$ and $|\tau| \le \delta$, and the true value $[\alpha_0, \tau_0]'$ is in the interior of the parameter space $[-\delta, \delta] \times [-\delta, \delta]$.*

Assumption A.3. *n is large and T can be finite or large.*

[1]Corresponding author (E-mail: jin.fei@live.com).

Assumption A.4. *W_n and M_n are uniformly bounded in row and column sums in absolute value (for short, UB).*

Assumption A.5. *The elements of X_{nT} are nonstochastic and bounded, uniformly in n and t. In addition, under the setting in Assumption A.3, $\lim \frac{1}{nT}\sum_{t=1}^{T}(e^{\tau M_n}\tilde{X}_{nt})'e^{\tau M_n}\tilde{X}_{nt}$ exists and is nonsingular for any $\tau \in [-\delta, \delta]$, and the sequence of the smallest eigenvalues of $e^{\tau M_n'}e^{\tau M_n}$ is bounded away from 0 uniformly in $\tau \in [-\delta, \delta]$.*

Assumption A.6. *The limit of $\frac{1}{nT}[\bar{\Gamma}_{nT}(\alpha, \tau) - (1 - \frac{1}{T}) \cdot tr(\Sigma_{nT})]$ is positive for any $\eta \neq \eta_0$.*

Assumption A.1 provides the essential property of the disturbances for the heteroskedastic case. Assumption A.2 maintains that the parameter spaces of α and τ are compact, as usual for extremum estimators, although the inverse of $e^{\alpha W_n}$ or $e^{\tau M_n}$ exists for any α or τ. In practice, the parameter spaces of α and τ may not need to be restricted, but as pointed out by Debarsy et al. (2015), in analysis, in order to ensure that $e^{\alpha W_n}$ and $e^{\tau M_n}$ be bounded in both row and column sum norms, α and τ should be bounded. In the rest of this paper, all limits are taken under Assumption A.3. The UB condition for spatial weights matrices in Assumption A.4 limits spatial correlation to a manageable degree. In Assumption A.5, elements of explanatory variables are assumed to be bounded constants for simplicity. The assumption also rules out multicollinearity among the explanatory variables. In addition, since the sequence of the smallest eigenvalues of $e^{\tau M_n'}e^{\tau M_n}$ is bounded away from 0 uniformly in $\tau \in [-\delta, \delta]$,[2] elements of $[\frac{1}{nT}\sum_{t=1}^{T}(e^{\tau M_n}\tilde{X}_{nt})'e^{\tau M_n}\tilde{X}_{nt}]^{-1}$ are bounded uniformly in τ. Assumption A.6 is a sufficient identification condition that ensures η_0 can be uniquely identified.[3]

2 Lemmas

Lemma 1. *Suppose that $\{A_n = [a_{n,ij}]\}$ and $\{B_n = [b_{n,ij}]\}$ are sequences of $n \times n$ matrices, and v_{it}'s in $V_{nt} = (v_{1t}, \cdots, v_{nt})'$ are independently distributed with mean zero (but may not be i.i.d.). Denote $\tilde{V}_{nt} = V_{nt} - \frac{1}{T}\sum_{k=1}^{T} V_{nk}$ and $B_n^s = B_n + B_n'$. Then,*

(1) $E\Big(\sum_{t=1}^{T}\tilde{V}_{nt}'A_n\tilde{V}_{nt}\Big) = tr\big[\Sigma_{nT}(J_T \otimes A_n)\big]$, *and*

(2) $E\Big[\Big(\sum_{t=1}^{T}\tilde{V}_{nt}'A_n\tilde{V}_{nt}\Big)\Big(\sum_{t=1}^{T}\tilde{V}_{nt}'B_n\tilde{V}_{nt}\Big)\Big]$
$= (\frac{T-1}{T})^2\sum_{t=1}^{T}\sum_{i=1}^{n} a_{n,ii}b_{n,ii}[E(v_{it}^4) - 3\sigma_{it}^4] + tr\big[\Sigma_{nT}(J_T \otimes A_n)\big]tr\big[\Sigma_{nT}(J_T \otimes B_n)\big] + tr\big[\Sigma_{nT}(J_T \otimes A_n)\Sigma_{nT}(J_T \otimes B_n^s)\big]$, *where $\Sigma_{nT} = diag(\Sigma_{nT,1}, \cdots, \Sigma_{nT,T})$ is a block diagonal matrix and each block $\Sigma_{nT,t} = diag(\sigma_{1t}^2, \cdots, \sigma_{nt}^2)$ with $E(v_{it}^2) = \sigma_{it}^2$.*

If A_n and B_n have zero diagonal elements, then

(3) $E\Big(\sum_{t=1}^{T}\tilde{V}_{nt}'A_n\tilde{V}_{nt}\Big) = 0$, *and*

(4) $E\Big[\Big(\sum_{t=1}^{T}\tilde{V}_{nt}'A_n\tilde{V}_{nt}\Big)\Big(\sum_{t=1}^{T}\tilde{V}_{nt}'B_n\tilde{V}_{nt}\Big)\Big] = tr\big[\Sigma_{nT}(J_T \otimes A_n)\Sigma_{nT}(J_T \otimes B_n^s)\big]$.

[2] This condition is similar to one in Debarsy et al. (2015), so it is standard.

[3] We discuss some low level conditions in the proof of Theorem 1.

Proof. Denote $\mathbf{V}_{nT} = (V'_{n1}, V'_{n2}, \cdots, V'_{nT})'$. With $J_T = I_T - \frac{1}{T} l_T l'_T$, we have $\sum_{t=1}^T \tilde{V}'_{nt} A_n \tilde{V}_{nt} = \mathbf{V}'_{nT}(J_T \otimes A_n)\mathbf{V}_{nT}$. Because v_{it}'s are mutually independent, $E\left(\sum_{t=1}^T \tilde{V}'_{nt} A_n \tilde{V}_{nt}\right) = E[\mathbf{V}'_{nT}(J_T \otimes A_n)\mathbf{V}_{nT}] = \sum_{t=1}^T \sum_{i=1}^n (1 - \frac{1}{T}) a_{n,ii} E(v_{it}^2) = tr\left[\Sigma_{nT}(J_T \otimes A_n)\right]$, then (1) holds. The (2) follows from Lemma A.2 in the supplement to Debarsy et al. (2015). When A_n and B_n have zero diagonal elements, we have $a_{ii} = 0$ and $b_{ii} = 0$, and both $J_T \otimes A_n$ and $J_T \otimes B_n$ have zero diagonal elements, then (3) and (4) hold. □

Lemma 2. *Suppose that $n \times n$ matrices A_n are UB, the elements of the $n \times k$ matrices C_{nt} are uniformly bounded, and v_{it}'s in $V_{nt} = (v_{1t}, \cdots, v_{nt})'$ are independent random variables with mean zero and variances σ_{it}^2. The $E(v_{it}^4)$ is bounded uniformly over i and t. Denote $\tilde{V}_{nt} = V_{nt} - \frac{1}{T}\sum_{m=1}^T V_{nm}$, $\tilde{C}_{nt} = C_{nt} - \frac{1}{T}\sum_{m=1}^T C_{nm}$. Then for large n and finite or large T,*

$$\frac{1}{nT}\sum_{t=1}^T \tilde{C}'_{nt} A_n \tilde{V}_{nt} = O_P\left(\frac{1}{\sqrt{nT}}\right), \tag{A.2}$$

$$\frac{1}{nT}\sum_{t=1}^T \tilde{V}'_{nt} A_n \tilde{V}_{nt} - \frac{1}{nT} E \sum_{t=1}^T \tilde{V}'_{nt} A_n \tilde{V}_{nt} = O_P\left(\frac{1}{\sqrt{nT}}\right), \tag{A.3}$$

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

Proof. Denote $\mathbf{C}_{nT} = (C'_{n1}, \cdots, C'_{nT})'$, $\mathbf{V}_{nT} = (V'_{n1}, V'_{n2}, \cdots, V'_{nT})'$. For the result (A.2), $J_T \otimes A_n$ is UB as A_n is UB. Then $\frac{1}{\sqrt{nT}}\sum_{t=1}^T C'_{nt} A_n V_{nt} = \frac{1}{\sqrt{nT}} \mathbf{C}'_{nT}(J_T \otimes A_n)\mathbf{V}_{nt}$ and Var$(\frac{1}{\sqrt{nT}}\sum_{t=1}^T C'_{nt} A_n V_{nt}) = \frac{1}{nT}\mathbf{C}'_{nT}(J_T \otimes A_n)\Sigma_{nT}(J_T \otimes A'_n)\mathbf{C}_{nT} = O(1)$. It follows that $(1/\sqrt{nT})\sum_{t=1}^T C'_{nt} A_n V_{nt} = O_P(1)$ by Chebyshev's inequality. For (A.3), when n is large but T is fixed, it is Lemma A.3 in Debarsy et al. (2015); when both n and T are large, it is a heteroskedastic version of the third term in Lemma 15 of Yu et. al. (2008). □

Lemma 3. *Suppose that $\{A_n = [a_{n,ij}]\}$ is a sequence of $n \times n$ symmetric UB matrices, $\{b_{nt} = [b_{nt,i}]\}$ is a sequence of constant column vectors such that $\sup_{n,T} \frac{1}{nT}\sum_{t=1}^T \sum_{i=1}^n |b_{nt,i}|^{2+\eta_1} < \infty$ for some $\eta_1 > 0$, and $v_{nt,i}$'s in $V_{nt} = (v_{nt,1}, \cdots, v_{nt,n})'$ are independent random variables with mean zero, variance $\sigma_{nt,i}^2$ and $\sup_{i,t,n} E(|v_{it}|^{4+\eta_2}) < \infty$ for some $\eta_2 > 0$. Denote $\Sigma_{nT} = diag(\Sigma_{nT,1}, \cdots, \Sigma_{nT,T})$, where $\Sigma_{nT,t} = diag(\sigma_{1t}^2, \cdots, \sigma_{nt}^2)$, and $\sigma_{C_{nT}}^2 = var(C_{nT})$, where $C_{nT} = \sum_{t=1}^T \left[b'_{nt} V_{nt} + V'_{nt} A_n V_{nt}\right] - tr\left[(I_T \otimes A_n)\Sigma_{nT}\right]$. Assume that $\frac{1}{nT}\sigma_{C_{nT}}^2$ is bounded away from zero. If n is large and T is finite or large, then $\frac{C_{nT}}{\sigma_{C_{nT}}^2} \xrightarrow{d} N(0,1)$.*

Proof. When n is large and T is fixed, this lemma is Lemma A.4 in Debarsy et al. (2015), which is essentially a central limit theorem originated in Kelejian and Prucha (2010). When both n and T are large, this lemma is a heteroskedastic version of that in Yu et. al. (2008), which can be proved similarly as that in Yu et. al. (2008). □

Lemma 4. *Let $\mathcal{A}_{nT}(\alpha, \tau) = \frac{1}{n(T-1)}\sum_{t=1}^T (\widetilde{X}_{nt}, A_n \widetilde{X}_{nt}\beta_0)' e^{\tau M'_n} e^{\tau M_n} (\widetilde{X}_{nt}, A_n \widetilde{X}_{nt}\beta_0)$, and $\mathcal{A}_{nT}(\alpha, \tau)$ be a 2×2 block matrix consisting of the following matrices: $\mathcal{A}_{11,nT}(\alpha, \tau) = \frac{1}{n(T-1)}\sum_{t=1}^T (e^{\tau M_n}\tilde{X}_{nt})'(e^{\tau M_n}\tilde{X}_{nt})$, $\mathcal{A}_{12,nT}(\alpha, \tau) = \frac{1}{n(T-1)}\sum_{t=1}^T \tilde{X}'_{nt} e^{\tau M'_n} e^{\tau M_n} A_n \tilde{X}_{nt}\beta_0$, $\mathcal{A}_{21,nT}(\alpha, \tau) = \mathcal{A}'_{12,nT}(\alpha, \tau)$, and*

$$\mathcal{A}_{22,nT}(\alpha, \tau) = \frac{1}{n(T-1)}\sum_{t=1}^T (A_n \tilde{X}_{nt}\beta_0)' e^{\tau M'_n} e^{\tau M_n} (A_n \tilde{X}_{nt}\beta_0),$$

where A_n is an $n \times n$ UB matrix. Suppose that $n \times n$ matrices M_n and W_n are UB, α and τ are in the parameter space $[-\delta, \delta]$ for some $\delta > 0$, the smallest eigenvalue of $e^{\tau M_n'} e^{\tau M_n}$ is bounded away from zero uniformly in τ, elements of the $n \times k$ matrix X_{nt} are uniformly bounded, and the limit of $\frac{1}{n(T-1)} \sum_{t=1}^T \tilde{X}_{nt}' e^{\tau M_n'} e^{\tau M_n} \tilde{X}_{nt}$ exists and is nonsingular for any $\tau \in [-\delta, \delta]$, where $\tilde{X}_{nt} = X_{nt} - \frac{1}{T} \sum_{k=1}^T X_{nk}$ and the limit is taken as n tends to infinite and T is fixed or tends to infinite. Then, $e^{\alpha W_n}$ is UB uniformly in $\alpha \in [-\delta, \delta]$, $e^{\tau M_n}$ is UB uniformly in $\tau \in [-\delta, \delta]$, $\mathcal{A}_{0,nT}(\alpha, \tau) = \mathcal{A}_{22,nT}(\alpha, \tau) - \mathcal{A}_{21,nT}(\alpha, \tau) \mathcal{A}_{11,nT}^{-1}(\alpha, \tau) \mathcal{A}_{12,nT}(\alpha, \tau)$ is bounded uniformly in $\alpha, \tau \in [-\delta, \delta]$.

Proof. If M_n and W_n are UB, that $e^{\alpha W_n}$ is UB uniformly in $\alpha \in [-\delta, \delta]$ and $e^{\tau M_n}$ is UB uniformly in $\tau \in [-\delta, \delta]$ can be found in the proof of Lemma A.6 of the supplement to Debarsy et al. (2015). As the smallest eigenvalue of $e^{\tau M_n'} e^{\tau M_n}$ is bounded away from zero uniformly in τ, there exists a constant $k > 0$ such that the smallest eigenvalue of $e^{\tau M_n'} e^{\tau M_n}$ is greater than or equal to k. Then $(\frac{1}{n(T-1)} \sum_{t=1}^T k \tilde{X}_{nt}' \tilde{X}_{nt})^{-1} - \mathcal{M}_{11,nT}^{-1}(\alpha, \tau)$ is positive semi-definite, as in the proof of Lemma A.6 in Debarsy et al. (2015). It follows that the elements of $\mathcal{A}_{11,nT}(\alpha, \tau)$ and $\mathcal{A}_{11,nT}^{-1}(\alpha, \tau)$ are bounded uniformly in $\tau \in [-\delta, \delta]$. Let $B_n = e^{\tau M_n'} e^{\tau M_n} A_n$. From the assumption of the lemma and the above proof, B_n is UB uniformly in $\alpha, \tau \in [-\delta, \delta]$. Since the elements of X_{nt} are uniformly bounded, so are the elements of $\tilde{X}_{nt}$. Then the elements of $\mathcal{A}_{22,nT}(\alpha, \tau)$, $\mathcal{A}_{12,nT}(\alpha, \tau)$ and $\mathcal{A}_{21,nT}(\alpha, \tau)$ are bounded uniformly in $\alpha, \tau \in [-\delta, \delta]$. It follows that $\mathcal{A}_{0,nT}(\alpha, \tau)$ is bounded uniformly in $(\alpha, \tau) \in [-\delta, \delta] \times [-\delta, \delta]$. □

Lemma 5. *Suppose that $n \times n$ matrices W_n, M_n and A_n are UB, the elements of the n-dimensional column vector $b_{nt} = [b_{nt,i}]$ are uniformly bounded, v_{it}'s in $V_{nt} = (v_{1t}, \cdots, v_{nt})'$ are independent random variables with mean zero and variances σ_{it}^2, the sequence $\{\sup_{t,i} E(v_{it}^4)\}$ is bounded, and the parameter space of $(\alpha, \tau)'$ is $\Phi = [-\delta, \delta] \times [-\delta, \delta]$ for some $\delta > 0$. Denote $\tilde{V}_{nt} = V_{nt} - \frac{1}{T} \sum_{k=1}^T V_{nk}$. Then $\frac{1}{n(T-1)} \sum_{t=1}^T b_{nt}' e^{\alpha W_n'} e^{\tau M_n'} e^{\tau M_n} e^{\alpha W_n} A_n \tilde{V}_{nt} = o_p(1)$ uniformly on Φ,*

$$\frac{1}{n(T-1)} \sum_{t=1}^T \tilde{V}_{nt} A_n' e^{\alpha W_n'} e^{\tau M_n'} e^{\tau M_n} e^{\alpha W_n} A_n \tilde{V}_{nt} - \frac{1}{n(T-1)} tr\big[(J_T \otimes A_n' e^{\alpha W_n'} e^{\tau M_n'} e^{\tau M_n} e^{\alpha W_n} A_n) \Sigma_{nT}\big] = o_p(1)$$

uniformly on Φ, and $\frac{1}{n(T-1)} tr\big[(J_T \otimes A_n' e^{\alpha W_n'} e^{\tau M_n'} e^{\tau M_n} e^{\alpha W_n} A_n) \Sigma_{nT}\big] = O(1)$ uniformly on Φ, where $\Sigma_{nT} = diag(\Sigma_{nT,1}, \ldots, \Sigma_{nT,T})$ is a block diagonal matrix with each block $\Sigma_{nT,t} = diag(\sigma_{1t}^2, \cdots, \sigma_{nt}^2)$.

Proof. By Lemma 2, $\frac{1}{n(T-1)} \sum_{t=1}^T b_{nt}' e^{\alpha W_n'} e^{\tau M_n'} e^{\tau M_n} e^{\alpha W_n} A_n \tilde{V}_{nt} = o_p(1)$,

$$\frac{1}{n(T-1)} tr\big[(J_T \otimes A_n' e^{\alpha W_n'} e^{\tau M_n'} e^{\tau M_n} e^{\alpha W_n} A_n) \Sigma_{nT}\big] = O(1)$$

and $\frac{1}{n(T-1)} \sum_{t=1}^T \tilde{V}_{nt} A_n' e^{\alpha W_n'} e^{\tau M_n'} e^{\tau M_n} e^{\alpha W_n} A_n \tilde{V}_{nt} - \frac{1}{n(T-1)} tr\big[(J_T \otimes A_n' e^{\alpha W_n'} e^{\tau M_n'} e^{\tau M_n} e^{\alpha W_n} A_n) \Sigma_{nT}\big] = o_p(1)$ for any $(\alpha, \tau) \in \Phi$. Denote $A(\alpha, \tau) = A_n' e^{\alpha W_n'} e^{\tau M_n'} e^{\tau M_n} e^{\alpha W_n} A_n$ with the (i, j)th element $a_{ij}(\alpha, \tau)$. The proof of the stochastic equicontinuity of the above three sequences, which is based on the mean value theorem, is similar to the proof of lemma A.7 in the supplement to Debarsy et al. (2015), thus it is omitted. Then the results in the lemma follow by Theorems 21.9 and 21.10 on p. 337–340 of Davidson (1994). □

3 Proofs of theorems

3.1 Proof of Theorem 1

Substituting $\hat{\beta}_{nT}(\eta) = \left(\sum_{t=1}^{T}(e^{\tau M_n}\tilde{X}_{nt})'e^{\tau M_n}\tilde{X}_{nt}\right)^{-1}\sum_{t=1}^{T}(e^{\tau M_n}\tilde{X}_{nt})'e^{\tau M_n}e^{\alpha W_n}\tilde{Y}_{nt}$ into $\Gamma_{nT}(\zeta)$, we obtain

$$\Gamma_{nT}(\eta) = \tilde{\mathbf{Y}}'_{nT}(I_T \otimes e^{\alpha W'_n}e^{\tau M'_n})H_{nT}(\eta)(I_T \otimes e^{\tau M_n}e^{\alpha W_n})\tilde{\mathbf{Y}}_{nT},$$

where $\tilde{\mathbf{Y}}_{nT} = [\tilde{Y}'_{n1}, \cdots, \tilde{Y}'_{nT}]'$. We shall prove that $\frac{1}{nT}[\Gamma_{nT}(\eta) - \bar{\Gamma}_{nT}(\eta)]$ converges in probability to zero uniformly on Φ and the identification uniqueness condition holds, where $\bar{\Gamma}_{nT}(\eta)$ is in (A.1).

We first show the uniform convergence that $\sup_{\eta\in\Phi}|\frac{1}{nT}[\Gamma_{nT}(\eta) - \bar{\Gamma}_{nT}(\eta)]| = o_p(1)$. Denote $\mathcal{M}_{nT}(\eta) = \frac{1}{n(T-1)}\sum_{t=1}^{T}(\tilde{X}_{nt},\, e^{(\alpha-\alpha_0)W_n}\tilde{X}_{nt}\beta_0)'e^{\tau M'_n}e^{\tau M_n}(\tilde{X}_{nt},\, e^{(\alpha-\alpha_0)W_n}\tilde{X}_{nt}\beta_0)$. $\mathcal{M}_{nT}(\eta)$ is a block matrix consisting of the following matrices: $\mathcal{M}_{11,nT}(\eta) = \frac{1}{n(T-1)}\sum_{t=1}^{T}(e^{\tau M_n}\tilde{X}_{nt})'(e^{\tau M_n}\tilde{X}_{nt})$, $\mathcal{M}_{21,nT}(\eta) = \mathcal{M}'_{12,nT}(\eta)$, $\mathcal{M}_{12,nT}(\eta) = \frac{1}{n(T-1)}\sum_{t=1}^{T}\tilde{X}'_{nt}e^{\tau M'_n}e^{\tau M_n}e^{(\alpha-\alpha_0)W_n}\tilde{X}_{nt}\beta_0$, and

$$\mathcal{M}_{22,nT}(\eta) = \frac{1}{n(T-1)}\sum_{t=1}^{T}(e^{(\alpha-\alpha_0)W_n}\tilde{X}_{nt}\beta_0)'e^{\tau M'_n}e^{\tau M_n}(e^{(\alpha-\alpha_0)W_n}\tilde{X}_{nt}\beta_0).$$

Let

$$\tilde{\mathcal{L}}_{\mathcal{X},nT}(\eta) = e^{\tau M_n}(e^{(\alpha-\alpha_0)W_n}\tilde{X}_{nt}\beta_0 - \tilde{X}_{nt}\mathcal{M}^{-1}_{11,nT}(\eta)\mathcal{M}_{12,nT}(\eta)),$$

and

$$\mathcal{V}_{\mathcal{X},nT}(\eta) = -\frac{1}{n(T-1)}\sum_{t=1}^{T}\tilde{X}'_{nt}e^{\tau M'_n}e^{\tau M_n}e^{(\alpha-\alpha_0)W_n}e^{-\tau_0 M_n}\tilde{V}_{nt}.$$

Then

$$\begin{aligned}\frac{1}{nT}[\Gamma_{nT}(\eta) - \bar{\Gamma}_{nT}(\eta)] = &\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}'_{nt}e^{-\tau_0 M'_n}e^{(\alpha-\alpha_0)W'_n}e^{\tau M'_n}e^{\tau M_n}e^{(\alpha-\alpha_0)W_n}e^{-\tau_0 M_n}\tilde{V}_{nt}\\ &-\frac{1}{nT}\cdot tr\big[(J_T \otimes e^{(\alpha-\alpha_0)W'_n}e^{(\tau-\tau_0)M'_n}e^{(\tau-\tau_0)M_n}e^{(\alpha-\alpha_0)W_n})\Sigma_{nT}\big]\\ &+\frac{2}{nT}\sum_{t=1}^{T}\tilde{\mathcal{L}}'_{\mathcal{X},nT}e^{\tau M_n}e^{\alpha W_n}e^{-\alpha_0 W_n}e^{-\tau_0 M_n}\tilde{V}_{nt}\\ &-\frac{T-1}{T}\mathcal{V}'_{\mathcal{X},nT}(\eta)\mathcal{M}^{-1}_{11,nT}(\eta)\mathcal{V}_{\mathcal{X},nT}(\eta).\end{aligned}$$

where

$$\frac{1}{nT}\sum_{t=1}^{T}\tilde{V}'_{nt}e^{-\tau_0 M'_n}e^{(\alpha-\alpha_0)W'_n}e^{\tau M'_n}e^{\tau M_n}e^{(\alpha-\alpha_0)W_n}e^{-\tau_0 M_n}\tilde{V}_{nt}$$

$$-\frac{1}{nT}\cdot tr\big[J_T\otimes e^{(\alpha-\alpha_0)W_n'}e^{(\tau-\tau_0)M_n'}e^{(\tau-\tau_0)M_n}e^{(\alpha-\alpha_0)W_n})\Sigma_{nT}\big]=o_p(1)$$

and the last two terms on the r.h.s. are $o_p(1)$ by Lemma 2. Thus, by Lemma 5, $\frac{1}{nT}[\Gamma_{nT}(\eta)-\bar{\Gamma}_{nT}(\eta)]=o_p(1)$ uniformly on Φ.

Second, we prove that $\frac{1}{nT}\bar{\Gamma}_{nT}(\eta)$ is uniformly equicontinuous on Φ. Similar to the proof of equicontinuity in the proof of Proposition 1 of the supplement to Debarsy et al. (2015), using the mean value theorem and Lemma 4, there exists a constant ω such that $\frac{1}{nT}|\bar{\Gamma}_{nT}(\eta_1)-\bar{\Gamma}_{nT}(\eta_2)|\leq\omega(\mid\alpha_1-\alpha_2\mid+\mid\tau_1-\tau_2\mid)$ for any $\eta_1,\eta_2\in\Phi$. Thus $\frac{1}{nT}\bar{\Gamma}_{nT}(\eta)$ is uniformly equicontinuous.

Third, we discuss the identification uniqueness condition. Let $\bar{\Gamma}_{nT}(\eta)=\bar{\Gamma}_{1n,T}(\eta)+\bar{\Gamma}_{2n,T}(\eta)$, where $\bar{\Gamma}_{1n,T}(\eta)$ and $\bar{\Gamma}_{2n,T}(\eta)$ are the first and second terms on the r.h.s. of (A.1) respectively. Using the commutativity property of W_n and M_n, the first and second order derivatives of $\bar{\Gamma}_{2n,T}(\eta)$ are

$$\frac{\partial\bar{\Gamma}_{2n,T}(\eta)}{\partial\eta}=\begin{pmatrix} tr\big[(J_T\otimes e^{(\alpha-\alpha_0)W_n'}e^{(\tau-\tau_0)M_n'}(W_n+W_n')e^{(\tau-\tau_0)M_n}e^{(\alpha-\alpha_0)W_n})\Sigma_{nT} \\ tr\big[(J_T\otimes e^{(\alpha-\alpha_0)W_n'}e^{(\tau-\tau_0)M_n'}(M_n+M_n')e^{(\tau-\tau_0)M_n}e^{(\alpha-\alpha_0)W_n})\Sigma_{nT}\big]\end{pmatrix}, \tag{A.4}$$

and

$$\frac{\partial^2\bar{\Gamma}_{2n,T}(\eta)}{\partial\eta\partial\eta'}=\begin{pmatrix}\overline{\gamma}_{11}(\eta) & * \\ \overline{\gamma}_{21}(\eta) & \overline{\gamma}_{22}(\eta)\end{pmatrix}, \tag{A.5}$$

where

$$\overline{\gamma}_{11}(\eta)=tr\Big[\Sigma_{nT}^{\frac{1}{2}}\big(J_T\otimes e^{(\alpha-\alpha_0)W_n'}e^{(\tau-\tau_0)M_n'}(W_n^2+W_n^{'2}+2W_n'W_n)e^{(\tau-\tau_0)M_n}e^{(\alpha-\alpha_0)W_n}\big)\Sigma_{nT}^{\frac{1}{2}}\Big],$$
$$\overline{\gamma}_{21}(\eta)=tr\Big[\Sigma_{nT}^{\frac{1}{2}}\big(J_T\otimes e^{(\alpha-\alpha_0)W_n'}e^{(\tau-\tau_0)M_n'}C_ne^{(\tau-\tau_0)M_n}e^{(\alpha-\alpha_0)W_n}\big)\Sigma_{nT}^{\frac{1}{2}}\Big],$$
$$\overline{\gamma}_{22}(\eta)=tr\Big[\Sigma_{nT}^{\frac{1}{2}}\big(J_T\otimes e^{(\alpha-\alpha_0)W_n'}e^{(\tau-\tau_0)M_n'}(M_n^2+M_n^{'2}+2M_n'M_n)e^{(\tau-\tau_0)M_n}e^{(\alpha-\alpha_0)W_n}\big)\Sigma_{nT}^{\frac{1}{2}}\Big],$$

where $C_n=M_n'(W_n+W_n')+(W_n+W_n')M_n$. By the Cauchy-Schwarz inequality, $\bar{\Gamma}_{1n,T}(\eta)\geq 0$ and is equal to zero when η is equal to η_0 under Assumption 1. Furthermore, under Assumption 1, the first order derivative of $\bar{\Gamma}_{2n,T}(\eta)$ at the true value η_0 is **0**. Therefore, η_0 is a stationary point of $\bar{\Gamma}_{2n,T}(\eta)$ as well as $\bar{\Gamma}_{nT}(\eta)$. If $W_n'W_n=W_nW_n'$, $M_n'M_n=M_nM_n'$ and $M_n'W_n=W_nM_n'$, then $W_n^2+W_n^{'2}+2W_n'W_n=(W_n+W_n')^2$, $M_n'(W_n+W_n')+(W_n+W_n')M_n=(W_n+W_n')(M_n+M_n')$ and $M_n^2+M_n^{'2}+2M_n'M_n=(M_n+M_n')^2$. Thus, by the Cauchy-Schwarz inequality, $(\overline{\gamma}_{12}(\eta))^2\leq\overline{\gamma}_{11}(\eta)\cdot\overline{\gamma}_{22}(\eta)$ under the conditions $W_n'W_n=W_nW_n'$, $M_n'M_n=M_nM_n'$ and $M_n'W_n=W_nM_n'$. Let $A_n^s=A_n+A_n'$ for any square matrix A_n, $J_T\otimes e^{(\alpha-\alpha_0)W_n'}e^{(\tau-\tau_0)M_n'}W_n^s=[w_{ij,e}^s]$, and $J_T\otimes e^{(\alpha-\alpha_0)W_n'}e^{(\tau-\tau_0)M_n'}M_n^s=[m_{ij,e}^s]$ for $i,j=1,\cdots,n$. If there is no constant k such that $w_{ij,e}^s=km_{ij,e}^s$ for all i,j, $\big(\overline{\gamma}_{12}(\eta)\big)^2<\overline{\gamma}_{11}(\eta)\cdot\overline{\gamma}_{22}(\eta)$. In this case, $\frac{\partial^2\bar{\Gamma}_{2n,T}(\eta)}{\partial\eta\partial\eta'}$ is positive definite and $\bar{\Gamma}_{2n,T}(\eta)$ is a strictly convex function. Thus, η_0 is the global minimizer of $\bar{\Gamma}_{2n,T}(\eta)$ as well as $\bar{\Gamma}_{nT}(\eta)$. It follows that $\frac{1}{nT}\bar{\Gamma}_{nT}(\eta)$ can have a unique minimal value at η_0 in the limit. If $w_{ij,e}^s=km_{ij,e}^s$ for a non-zero constant k, $\big(\overline{\gamma}_{12}(\eta)\big)^2=\overline{\gamma}_{11}(\eta)\cdot\overline{\gamma}_{22}(\eta)$. Then η_0 might be or might not be a global minimizer of $\bar{\Gamma}_{nT}(\eta)$ in some cases. Consider the case with $W_n=M_n$, which implies that $w_{ij,e}^s=m_{ij,e}^s$, then $\overline{\gamma}_{12}(\eta_0)=\overline{\gamma}_{11}(\eta_0)=\overline{\gamma}_{22}(\eta_0)=tr\big[\big(J_T\otimes(W_n^2+W_n^{'2}+2W_n'W_n)\big)\Sigma_{nT}\big]$. In this case, $tr\big[\big(J_T\otimes(W_n^2+W_n^{'2}+2W_n'W_n)\big)\Sigma_{nT}\big]>0$ if the elements of W_n are non-negative. It follows that

$\frac{\partial^2 \bar{\Gamma}_{2n,T}(\eta_0)}{\partial\eta\partial\eta'}$ is positive semi-definite. Then η_0 might be only a local minimizer of $\bar{\Gamma}_{2n,T}(\eta)$. Assumption A.6 is a sufficient condition that ensures the identification uniquess of the true parameter vector.

Combining the uniform convergence and identification uniqueness condition in Assumption A.6, the consistency of $\hat{\alpha}_{nT}$ and $\hat{\tau}_{nT}$ follows. For given η, minimizing $\Gamma_{nT}(\zeta)$ yields

$$\hat{\beta}_{nT}(\eta) = \Big(\sum_{t=1}^{T}(e^{\tau M_n}\tilde{X}_{nt})' e^{\tau M_n}\tilde{X}_{nt}\Big)^{-1}\sum_{t=1}^{T}(e^{\tau M_n}\tilde{X}_{nt})' e^{\tau M_n} e^{\alpha W_n}\tilde{Y}_{nt}.$$

Then we can substitute the estimators $\hat{\alpha}_{nT}$ and $\hat{\tau}_{nT}$ into the above equation to derive a consistent estimator $\hat{\beta}_{nT}$.

3.2 Proof of Theorem 2

The asymptotic distribution of $\hat{\zeta}_{nT}$ is derived from applying the mean value theorem to the first-order condition $\frac{\partial \Gamma_{nT}(\hat{\zeta}_{nT})}{\partial \zeta} = 0$ at the true value ζ_0, which yields

$$\sqrt{nT}(\hat{\zeta}_{nT} - \zeta_0) = -\left(\frac{1}{nT}\frac{\partial^2 \Gamma_{nT}(\bar{\zeta})}{\partial\zeta\partial\zeta'}\right)^{-1}\left(\frac{1}{\sqrt{nT}}\frac{\partial \Gamma_{nT}(\zeta_0)}{\partial\zeta}\right), \tag{A.6}$$

where $\bar{\zeta}_{nT}$ is between $\hat{\zeta}_{nT}$ and ζ_0. We need to show that (1) $\frac{1}{nT}\frac{\partial^2 \Gamma_{nT}(\bar{\zeta}_{nT})}{\partial\zeta\partial\zeta'} = \Sigma_{\zeta_0,nT} + o_p(1)$ and (2) the limit of $\Sigma_{\zeta_0,nT}$ is nonsingular.

Proof of (1): The second-order derivatives of $\Gamma_{nT}(\zeta)$ are

$$\begin{aligned}
\frac{\partial^2 \Gamma_{nT}(\zeta)}{\partial\beta\partial\beta'} &= 2\sum_{t=1}^{T}(e^{\tau M_n}\tilde{X}_{nt})' e^{\tau M_n}\tilde{X}_{nt},\\
\frac{\partial^2 \Gamma_{nT}(\zeta)}{\partial\alpha\partial\beta'} &= -2\sum_{t=1}^{T}(e^{\tau M_n}W_n e^{\alpha W_n}\tilde{Y}_{nt})' e^{\tau M_n}\tilde{X}_{nt},\\
\frac{\partial^2 \Gamma_{nT}(\zeta)}{\partial\tau\partial\beta'} &= -2\sum_{t=1}^{T}\tilde{V}'_{nt}(\zeta)(M'_n + M_n)e^{\tau M_n}\tilde{X}_{nt},\\
\frac{\partial^2 \Gamma_{nT}(\zeta)}{\partial\alpha\partial\alpha} &= 2\sum_{t=1}^{T}(e^{\tau M_n}W_n e^{\alpha W_n}\tilde{Y}_{nt})'(e^{\tau M_n}W_n e^{\alpha W_n}\tilde{Y}_{nt}) + \sum_{t=1}^{T}(e^{\tau M_n}W_n W_n e^{\alpha W_n}\tilde{Y}_{nt})'\tilde{V}_{nt}(\zeta),\\
\frac{\partial^2 \Gamma_{nT}(\zeta)}{\partial\tau\partial\alpha} &= 2\sum_{t=1}^{T}(e^{\tau M_n}W_n e^{\alpha W_n}\tilde{Y}_{nt})'(M'_n + M_n)\tilde{V}_{nt}(\zeta),\\
\frac{\partial^2 \Gamma_{nT}(\zeta)}{\partial\tau\partial\tau} &= 2\sum_{t=1}^{T}(M_n\tilde{V}_{nt}(\zeta))'(M'_n + M_n)\tilde{V}_{nt}(\zeta).
\end{aligned}$$

By Lemma 2 and the reduced form of $\tilde{Y}_{nt}$, $\frac{1}{nT}\sum_{t=1}^{T}\tilde{Y}'_{nt}A_n\tilde{Y}_{nt} = O_p(1)$ and $\frac{1}{nT}\sum_{t=1}^{T}\tilde{X}'_{nt}A_n\tilde{Y}_{nt} = O_p(1)$, where A_n is an $n\times n$ UB matrix. Then we can use the method in the proof of Proposition 2 in the supplement of Debarsy et al. (2015) (page 11). First, we write $e^{\bar{\alpha}W_n} = (e^{\bar{\alpha}W_n} - e^{\alpha_0 W_n}) + e^{\alpha_0 W_n}$, $e^{\bar{\tau}M_n} = (e^{\bar{\tau}M_n} - e^{\tau_0 M_n}) + e^{\tau_0 M_n}$ and $\bar{\beta} = (\bar{\beta} - \beta_0) + \beta_0$. By Lemma A.8 in the supplement of Debarsy et al. (2015), $\| e^{\bar{\alpha}W_n} - e^{\alpha_0 W_n} \|_\infty = o_p(1)$ and $\| e^{\bar{\tau}M_n} - e^{\tau_0 M_n} \|_\infty = o_p(1)$. Then from the expanded forms of each term for $\frac{1}{nT}\frac{\partial^2 \Gamma_{nT}(\bar{\zeta})}{\partial\zeta\partial\zeta'}$ and the sub-multiplicability of the row sum matrix norm, $\frac{1}{nT}\frac{\partial^2\Gamma_{nT}(\bar{\zeta})}{\partial\zeta\partial\zeta'} = \frac{1}{nT}\frac{\partial^2\Gamma_{nT}(\zeta_0)}{\partial\zeta\partial\zeta'} + o_p(1)$. The detailed expression of each entry of the difference $\frac{1}{nT}\frac{\partial^2\Gamma_{nT}(\zeta_0)}{\partial\zeta\partial\zeta'} - \Sigma_{\zeta_0,nT}$ is straightforward from (9) and the second-order derivatives of $\Gamma_{nT}(\zeta)$, and each element of the above matrix difference is a linear-quadratic form of $\tilde{V}_{nT}$. Thus, by using Lemma 2, we have $\frac{1}{nT}\frac{\partial^2\Gamma_{nT}(\zeta_0)}{\partial\zeta\partial\zeta'} - \mathcal{H}_{\zeta_0,nT} = o_p(1)$.

Proof of (2): We need to prove that $\lim \mathcal{H}_{\zeta_0,nT}c = 0$ implies $c = 0$, where $c = (c_1', c_2, c_3)'$, c_1 is a $k \times 1$ vector, and c_2 and c_3 are scalars. Denote $\mathcal{N}_{11,nT} = \sum_{t=1}^T (e^{\tau_0 M_n}\tilde{X}_{nt})'(e^{\tau_0 M_n}\tilde{X}_{nt})$,

$$\mathcal{N}_{12,nT} = -\sum_{t=1}^T (e^{\tau_0 M_n}\tilde{X}_{nt})'(e^{\tau_0 M_n}W_n\tilde{X}_{nt}\beta_0),$$

$\mathcal{N}_{21,nT} = \mathcal{N}_{12,nT}'$ and $\mathcal{N}_{22,nT} = \sum_{t=1}^T (e^{\tau_0 M_n}W_n\tilde{X}_{nt}\beta_0)'(e^{\tau_0 M_n}W_n\tilde{X}_{nt}\beta_0)$. Under Assumption A.5, $\lim \frac{1}{nT}\mathcal{N}_{11,nT}$ is nonsingular. Then we have $c_1 = \lim[\frac{1}{nT}\mathcal{N}_{11,nT}]^{-1}\frac{1}{nT}\mathcal{N}_{12,nT}c_2$. If $\lim \frac{1}{nT}tr\big[(J_T \otimes M_n^s M_n)\Sigma_{nT}\big] \neq 0$, $c_3 = -c_2 \lim_{n\to\infty} tr\big[(J_T \otimes M_n^s W_n)\Sigma_{nT}\big]/tr\big[(J_T \otimes M_n^s M_n)\Sigma_{nT}\big]$. Substituting the expressions of c_1 and c_3 into the second row block of $\lim \mathcal{H}_{\zeta_0,nT}c = 0$, we have

$$\lim \frac{1}{nT}\big((\tilde{\mathbf{X}}_{nT}\beta_0)'(I_T \otimes W_n' e^{\tau_0 M_n'})H_{nT}(\tau_0)(I_T \otimes e^{\tau_0 M_n}W_n)\tilde{\mathbf{X}}_{nT}\beta_0 + \mathcal{B}\big)c_2 = 0,$$

where $\mathcal{B} = tr[(J_T \otimes W_n^s W_n)\Sigma_{nT}] - \frac{tr^2[(J_T\otimes M_n^s W_n)\Sigma_{nT}]}{tr[(J_T\otimes M_n^s M_n)\Sigma_{nT}]}$. Thus, Assumption 2 implies that the limit of $\mathcal{H}_{\zeta_0,nT}$ is nonsingular.

Combining $\frac{1}{nT}\frac{\partial^2\Gamma_{nT}(\bar{\zeta})}{\partial\zeta\partial\zeta'} = \mathcal{H}_{\zeta_0,nT} + o_p(1)$ and the nonsingularity of the limit of $\mathcal{H}_{\zeta_0,nT}$, we have $\sqrt{nT}(\hat{\zeta}_{nT} - \zeta_0) = \mathcal{H}_{\zeta_0,nT}^{-1}\left(\frac{1}{\sqrt{nT}}\frac{\partial\Gamma_{nT}(\zeta_0)}{\partial\zeta}\right) + o_p(1)$. Each element of $\frac{1}{\sqrt{nT}}\frac{\partial\Gamma_{nT}(\zeta_0)}{\partial\zeta}$ is a linear-quadratic form of $\tilde{V}_{nT}$ with zero mean, and the variance of $\frac{1}{\sqrt{nT}}\frac{\partial\Gamma_{nT}(\zeta_0)}{\partial\zeta}$ is $\Delta_{\zeta_0,nT}$ in (9) of the main text. Thus, by Lemma 3, $\sqrt{nT}(\hat{\zeta}_{nT} - \zeta_0) \xrightarrow{d} N\big(0, \lim(\mathcal{H}_{\zeta_0,nT}^{-1}\Delta_{\zeta_0,nT}\mathcal{H}_{\zeta_0,nT}^{-1})\big)$.

3.3 Proof of Theorem 3

(I) Proof of $\hat{\mathcal{H}}_{\zeta_0,nT} - \mathcal{H}_{\zeta_0,nT} = o_p(1)$ and $\hat{\Delta}_{\zeta_0,nT} - \Delta_{\zeta_0,nT} = o_p(1)$ under the condition that both n and T are large: It is sufficient to prove that (i) $\frac{1}{nT}tr\big[\hat{\Sigma}_{nT}(J_T \otimes W_n^s)\hat{\Sigma}_{nT}(J_T \otimes M_n^s)\big] - \frac{1}{nT}tr\big[\Sigma_{nT}(J_T \otimes W_n^s)\Sigma_{nT}(J_T \otimes M_n^s)\big] = o_p(1)$, (ii) $\frac{1}{nT}tr\big[(J_T \otimes W_n^s M_n)\hat{\Sigma}_{nT})\big] - \frac{1}{nT}tr\big[(J_T \otimes W_n^s M_n)\Sigma_{nT}\big] = o_p(1)$,

$$\text{(iii) } \frac{1}{nT}\sum_{t=1}^T r_{nt}' e^{\hat{\tau}_{nT}M_n'}e^{\hat{\tau}_{nT}M_n}s_{nt} - \frac{1}{nT}\sum_{t=1}^T r_{nt}' e^{\tau_0 M_n'}e^{\tau_0 M_n}s_{nt} = o_p(1),$$

and (iv) $\frac{1}{nT}R_{nT}'(J_T \otimes e^{\hat{\tau}_{nT}M_n'})\hat{\Sigma}_{nT}(J_T \otimes e^{\hat{\tau}_{nT}M_n})S_{nT} - \frac{1}{nT}R_{nT}'(J_T \otimes e^{\tau_0 M_n'})\Sigma_{nT}(J_T \otimes e^{\tau_0 M_n})S_{nT} = o_p(1)$, where $\{r_{nt} = [r_{nt,i}]\}$ and $\{s_{nt} = [s_{nt,i}]\}$ are n-dimensional column vectors with uniformly bounded elements, $R_{nT} = (r_{n1}', \cdots, r_{nT}')'$, $S_{nT} = (s_{n1}', \cdots, s_{nT}')'$, and $\Sigma_{nT} = diag(\Sigma_{nT,1}, \cdots, \Sigma_{nT,T})$ is a block diagonal matrix with each block $\Sigma_{nT,t} = diag(\sigma_{1t}^2, \cdots, \sigma_{nt}^2)$, where $E(v_{it}^2) = \sigma_{it}^2$. Suppose that we would like to prove that $\frac{1}{nT}tr[(J_T \otimes A_n)\hat{\Sigma}_{nT}] - \frac{1}{nT}tr[(J_T \otimes A_n)\Sigma_{nT}] = o_p(1)$, where $A_n = [a_{ij}]$ is an $n \times n$ UB matrix. Note that

$$\frac{1}{nT}tr[(J_T \otimes A_n)\Sigma_{nT}] = \frac{T-1}{nT^2}\sum_{i=1}^n a_{ii}\sum_{t=1}^T \sigma_{it}^2. \tag{A.7}$$

In the above equation, suppose that we replace σ_{it}^2 with $(v_{it} - \frac{1}{T}\sum_{t=1}^T v_{it})^2$. Then we have

$$\frac{T-1}{nT^2} E \sum_{i=1}^n a_{ii} \sum_{t=1}^T \Big(v_{it} - \frac{1}{T}\sum_{t=1}^T v_{it} \Big)^2 = \frac{(T-1)^2}{nT^3} \sum_{i=1}^n a_{ii} \sum_{t=1}^T \sigma_{it}^2. \tag{A.8}$$

This expectation is not equal to $\frac{1}{nT} tr[(J_T \otimes A_n)\Sigma_{nT}]$. However, we can replace σ_{it}^2 with $\frac{T}{T-1}(v_{it} - \frac{1}{T}\sum_{t=1}^T v_{it})^2$ instead to derive a term with an expected value equal to $\frac{1}{nT} tr[(J_T \otimes A_n)\Sigma_{nT}]$. Denote $W_n^s = [w_{ij}^s]$ and $M_n^s = [m_{ij}^s]$. Note that

$$\begin{aligned}
&\frac{1}{nT} tr\big[\Sigma_{nT}(J_T \otimes W_n^s)\Sigma_{nT}(J_T \otimes M_n^s)\big] \\
&= \frac{1}{nT^3} tr\Big((\sum_{t=1}^T \Sigma_{nT,t}) W_n^s (\sum_{t=1}^T \Sigma_{nT,t}) M_n^s \Big) + \frac{T-2}{T} \Big(\sum_{t=1}^T tr(\Sigma_{nT,t} W_n^s \Sigma_{nT,t} M_n^s) \Big) \\
&= \frac{1}{nT^3} \sum_{t=1}^T \sum_{p=1}^n \sum_{q=1}^n \sum_{j=1}^T \sigma_{pt}^2 \sigma_{qj}^2 w_{pq}^s m_{qp}^s + \frac{T-2}{nT^2} \sum_{t=1}^T \sum_{p=1}^n \sum_{q=1}^n \sigma_{pt}^2 \sigma_{qt}^2 w_{pq}^s m_{qp}^s.
\end{aligned} \tag{A.9}$$

Suppose that we use $\tilde{v}_{pt}^2 = \frac{T}{T-1}(v_{pt} - \frac{1}{T}\sum_{k=1}^T v_{pk})^2$ and $\tilde{v}_{qj}^2 = \frac{T}{T-1}(v_{qj} - \frac{1}{T}\sum_{k=1}^T v_{qk})^2$ to replace σ_{pt}^2 and σ_{qj}^2 respectively in the above equation. Note that if $p = q$, $w_{pq}^s m_{qp}^s = 0$. If $p \neq q$,

$$\begin{aligned}
&(\frac{T}{T-1})^2 E\Big[\Big(v_{pt} - \frac{1}{T}\sum_{k=1}^T v_{pk} \Big)^2 \Big(v_{qj} - \frac{1}{T}\sum_{k=1}^T v_{qk} \Big)^2 \Big] \\
&= (\frac{T}{T-1})^2 \Big[(1 - \frac{2}{T})^2 \sigma_{pt}^2 \sigma_{qj}^2 + (\frac{1}{T^2} - \frac{2}{T^3}) \sigma_{pt}^2 \sum_{k=1}^T \sigma_{qk}^2 + (\frac{1}{T^2} - \frac{2}{T^3}) \sigma_{qj}^2 \sum_{k=1}^T \sigma_{pk}^2 \\
&\quad + \frac{1}{T^4} (\sum_{k=1}^T \sigma_{pk}^2) \Big(\sum_{k=1}^T \sigma_{qk}^2 \Big) \Big].
\end{aligned}$$

Then

$$\begin{aligned}
&\frac{T}{n(T-1)^2} E\Big(\frac{1}{T^2} \sum_{t=1}^T \sum_{p=1}^n \sum_{q=1}^n \sum_{j=1}^T \tilde{v}_{pt}^2 \tilde{v}_{qj}^2 w_{pq}^s m_{qp}^s + \frac{T-2}{T} \sum_{t=1}^T \sum_{p=1}^n \sum_{q=1}^n \tilde{v}_{pt}^2 \tilde{v}_{qt}^2 w_{pq}^s m_{qp}^s \Big) \\
&= \frac{T}{n(T-1)^2} \Big[\big[\frac{3}{T^2}(1 - \frac{2}{T})^2 + \frac{3}{T^3}(1 - \frac{2}{T}) + \frac{1}{T^4} \big] \sum_{t=1}^T \sum_{p=1}^n \sum_{q=1}^n \sum_{j=1}^T \sigma_{pt}^2 \sigma_{qj}^2 w_{pq}^s m_{qp}^s \\
&\quad + (1 - \frac{2}{T})^3 \sum_{t=1}^T \sum_{p=1}^n \sum_{q=1}^n \sigma_{pt}^2 \sigma_{qt}^2 w_{pq}^s m_{qp}^s \Big].
\end{aligned} \tag{A.10}$$

Note that the third line of (A.9) is not equal to the r.h.s. of (A.10) when T is fixed. Let $\sigma_{it}^2 \leqslant c$, where c is an non-negative constant. Then as W_n and M_n are UB,

$$\begin{aligned}
&\Big| \sum_{t=1}^T \sum_{p=1}^n \sum_{q=1}^n \sum_{j=1}^T \sigma_{pt}^2 \sigma_{qj}^2 w_{pq}^s m_{qp}^s \Big| \leq 2T \cdot c^2 \sum_{p=1}^n \Big(\sum_{q=1}^n |w_{pq}^s m_{qp}^s| \Big) \leq nT \cdot k, \\
&\Big| \sum_{t=1}^T \sum_{p=1}^n \sum_{q=1}^n \sigma_{pt}^2 \sigma_{qt}^2 w_{pq}^s m_{qp}^s \Big| \leq T \cdot c^2 \sum_{p=1}^n \Big(\sum_{q=1}^n |w_{pq}^s m_{qp}^s| \Big) \leq \frac{nT}{2} \cdot k
\end{aligned}$$

for some non-negative constant k. We see that as T tends to infinity, the first terms of the r.h.s of the third line of (A.9) and the r.h.s of (A.10) are $o(1)$, and the second terms of the r.h.s of the third line of (A.9) and the r.h.s of (A.10) are $O(1)$. Thus, (A.9) and (A.10) are dominated by their second terms. As T tends to infinity, the two dominant terms are asymptotically equal. It follows that we can replace σ_{it}^2 with $\frac{T}{T-1}(v_{it} - \frac{1}{T}\sum_{t=1}^T v_{it})^2$ to estimate $\frac{1}{nT}tr\big[\Sigma_{nT}(J_T \otimes W_n^s)\Sigma_{nT}(J_T \otimes M_n^s)\big]$ when T tends to infinite. In practice, we do not observe v_{it}, but we have the QML residuals $\hat{\tilde{v}}_{it} = \hat{v}_{it} - \frac{1}{T}\sum_{t=1}^T \hat{v}_{it}$. So we may let $\hat{\Sigma}_{nT} = diag(\hat{\Sigma}_{nT,1}, \cdots, \hat{\Sigma}_{nT,T})$, where $\hat{\Sigma}_{nT,t} = \frac{T}{T-1}(\hat{\tilde{v}}_{1t}^2, \cdots, \hat{\tilde{v}}_{nt}^2)$.

Proof of (i): Using $\hat{\Sigma}_{nT}$ to replace $\Sigma_{nT,t}$, we have

$$\begin{aligned}
&\frac{1}{nT}tr\big[\hat{\Sigma}_{nT}(J_T \otimes W_n^s)\hat{\Sigma}_{nT}(J_T \otimes M_n^s)\big] \\
&= \frac{1}{nT}\Big[\frac{1}{T^2}tr\big((\sum_{t=1}^T \hat{\Sigma}_{nT,t})W_n^s(\sum_{t=1}^T \hat{\Sigma}_{nT,t})M_n^s\big) + \frac{T-2}{T}\big(\sum_{t=1}^T tr(\hat{\Sigma}_{nT,t}W_n^s\hat{\Sigma}_{nT,t}M_n^s)\big)\Big].
\end{aligned}$$

Then we can rewrite (i) as

$$\frac{1}{nT}tr\big[\hat{\Sigma}_{nT}(J_T \otimes W_n^s)\hat{\Sigma}_{nT}(J_T \otimes M_n^s)\big] - \frac{1}{nT}tr\big[\Sigma_{nT}(J_T \otimes W_n^s)\Sigma_{nT}(J_T \otimes M_n^s)\big] = A_1 + A_2 = o_p(1),$$

where

$$\begin{aligned}
A_1 &= \frac{1}{nT}\Big[\frac{1}{T^2}\big[tr\big((\sum_{t=1}^T \hat{\Sigma}_{nT,t})W_n^s(\sum_{t=1}^T \hat{\Sigma}_{nT,t})M_n^s\big) - tr\big((\sum_{t=1}^T \Sigma_{nT,t})W_n^s(\sum_{t=1}^T \Sigma_{nT,t})M_n^s\big)\big]\Big], \\
A_2 &= \frac{1}{nT}\Big[\frac{T-2}{T}\big[(\sum_{t=1}^T tr(\hat{\Sigma}_{nT,t}W_n^s\hat{\Sigma}_{nT,t}M_n^s)) - (\sum_{t=1}^T tr(\Sigma_{nT,t}W_n^s\Sigma_{nT,t}M_n^s))\big]\Big].
\end{aligned}$$

We shall show that $A_1 = o_p(1)$ and $A_2 = o_p(1)$. The proof of $A_1 = o_p(1)$: Let $P_n = [p_{ij}]$ be an $n \times n$ symmetric matrix, where $p_{ij} = w_{ij}^s m_{ji}^s$. Note that $p_{ii} = 0$ and P_n is UB under Assumption A.4. To show that $A_1 = o_p(1)$, we may show that (a) $\frac{1}{nT^3}\sum_{i=1}^n\sum_{j=1}^n\big((\frac{T}{T-1}\sum_{t=1}^T \tilde{v}_{it}^2)(\frac{T}{T-1}\sum_{t=1}^T \tilde{v}_{jt}^2) - (\sum_{t=1}^T \sigma_{it}^2)(\sum_{t=1}^T \sigma_{jt}^2)\big)p_{ij} = o_p(1)$ and (b) $\frac{1}{nT^3}\sum_{i=1}^n\sum_{j=1}^n\big((\frac{T}{T-1}\sum_{t=1}^T \hat{\tilde{v}}_{it}^2)(\frac{T}{T-1}\sum_{t=1}^T \hat{\tilde{v}}_{jt}^2) - (\frac{T}{T-1}\sum_{t=1}^T \tilde{v}_{it}^2)(\frac{T}{T-1}\sum_{t=1}^T \tilde{v}_{jt}^2)\big)p_{ij} = o_p(1)$. We first show that (a) holds: Let $\{\tilde{f}_n = [\tilde{f}_{ni}]\}$ be an n-dimensional column vector with $\tilde{f}_{ni} = \frac{T}{T-1}\sum_{t=1}^T \tilde{v}_{it}^2$, and $\{f_n = [f_{ni}]\}$ be an n-dimensional column vector with $f_{ni} = \sum_{t=1}^T \sigma_{it}^2$. As

$$\tilde{f}_{ni}\tilde{f}_{nj} - f_{ni}f_{nj} = (\tilde{f}_{ni} - f_{ni})(\tilde{f}_{nj} - f_{nj}) + f_{ni}(\tilde{f}_{nj} - f_{n,j}) + f_{nj}(\tilde{f}_{ni} - f_{ni}),$$

we have $\frac{1}{nT^3}\sum_{i=1}^n\sum_{j=1}^n\big((\frac{T}{T-1}\sum_{t=1}^T \tilde{v}_{it}^2)(\frac{T}{T-1}\sum_{t=1}^T \tilde{v}_{jt}^2) - (\sum_{t=1}^T \sigma_{it}^2)(\sum_{t=1}^T \sigma_{jt}^2)\big)p_{ij} = B_{1,1} + B_{1,2} + B_{1,3}$, where $B_{1,1} = \frac{1}{nT^3}\sum_{i=1}^n\sum_{j=1}^n(\tilde{f}_{ni} - f_{ni})(\tilde{f}_{nj} - f_{n,j})p_{ij}$, $B_{1,2} = \frac{1}{nT^3}\sum_{i=1}^n\sum_{j=1}^n(\tilde{f}_{n,j} - f_{nj})f_{ni}p_{ij}$, and $B_{1,3} = \frac{1}{nT^3}\sum_{i=1}^n\sum_{j=1}^n(\tilde{f}_{ni} - f_{ni})f_{nj}p_{ij}$. Let $F_n = (F_{n1}, \cdots, F_{nn})'$ where $F_{ni} = \tilde{f}_{ni} - f_{ni}$. Denote $f_n = (f_{n1}, \cdots, f_{nn})'$. Then $B_{1,1} = \frac{1}{nT^3}F_n'P_nF_n$, $B_{1,2} = \frac{1}{nT^3}f_n'P_nF_n$, and $B_{1,3} = \frac{1}{nT^3}F_n'P_nf_n$. Denote $\Lambda_n = E(F_nF_n')$. As v_{it}'s are independent across i and t, F_{ni}'s are mutually independent and Λ_n is a diagonal matrix. Then $E(F_n'P_nF_n) = tr(P_n\Lambda_n) = 0$ as $p_{ii} = 0$. Note that under Assumption A.1, $E|\tilde{v}_{it}^4|$'s exist and are uniformly bounded for all i and t. Then similar to the proof of (i) in the proof of Proposition 2 of Lin and Lee (2010), $B_{1,1}=o_p(1)$, $B_{1,2}=o_p(1)$, and $B_{1,3}=o_p(1)$. Hence, (a) holds. We next show that (b) holds: Let $\hat{\tilde{f}}_n=[\hat{\tilde{f}}_{ni}]$ be an n-dimensional column vector with $\hat{\tilde{f}}_{ni}=\frac{T}{T-1}\sum_{t=1}^T\hat{\tilde{v}}_{it}^2$. Then $\hat{\tilde{f}}_{ni}\hat{\tilde{f}}_{nj}-\tilde{f}_{ni}\tilde{f}_{nj}=(\hat{\tilde{f}}_{ni}-\tilde{f}_{ni})\tilde{f}_{nj}+\tilde{f}_{ni}(\hat{\tilde{f}}_{nj}-\tilde{f}_{nj})+(\hat{\tilde{f}}_{ni}-\tilde{f}_{ni})(\hat{\tilde{f}}_{nj}-\tilde{f}_{nj})$. Note that

$$\frac{1}{nT^3}\sum_{i=1}^n\sum_{j=1}^n\big((\frac{T}{T-1}\sum_{t=1}^T\hat{\tilde{v}}_{it}^2)(\frac{T}{T-1}\sum_{t=1}^T\hat{\tilde{v}}_{jt}^2)-(\frac{T}{T-1}\sum_{t=1}^T\tilde{v}_{it}^2)(\frac{T}{T-1}\sum_{t=1}^T\tilde{v}_{jt}^2)\big)p_{ij}=C_1+C_2+C_3,$$

where $C_1=\frac{1}{nT^3}\sum_{i=1}^n\sum_{j=1}^n(\hat{\tilde{f}}_{ni}-\tilde{f}_{ni})\tilde{f}_{nj}p_{ij}$, $C_2=\frac{1}{nT^3}\sum_{i=1}^n\sum_{j=1}^n(\hat{\tilde{f}}_{nj}-\tilde{f}_{nj})\tilde{f}_{ni}p_{ij}$, and

$$C_3=\frac{1}{nT^3}\sum_{i=1}^n\sum_{j=1}^n(\hat{\tilde{f}}_{ni}-\tilde{f}_{ni})(\hat{\tilde{f}}_{nj}-\tilde{f}_{nj})p_{ij}.$$

We shall show that $C_i=o_p(1)$ for $i=1,2,3$. Since the proofs for different C_i's are similar, we just detail the proof for the most complicated term C_3. From the model, we have

$$\begin{aligned}\hat{\tilde{V}}_{nt}&=e^{\hat{\tau}_{nT}M_n}(e^{\hat{\alpha}_{nT}W_n}\tilde{Y}_{nt}-\tilde{X}_{nt}\hat{\beta}_{nT})\\&=[e^{\hat{\tau}_{nT}M_n}e^{(\hat{\alpha}_{nT}-\alpha_0)W_n}e^{-\tau_0M_n}-I_n]\tilde{V}_{nt}+e^{\hat{\tau}_{nT}M_n}(e^{(\hat{\alpha}_{nT}-\alpha_0)W_n}-I_n)\tilde{X}_{nt}\beta_0\\&\quad+e^{\hat{\tau}_{nT}M_n}\tilde{X}_{nt}(\beta_0-\hat{\beta}_{nT})+\tilde{V}_{nt}.\end{aligned}$$

Then in scalar form, $\hat{\tilde{v}}_{it}=a_{it}+b_{it}+c_{it}+\tilde{v}_{it}$, where $a_{it}=e_i[e^{\hat{\tau}_{nT}M_n}e^{(\hat{\alpha}_{nT}-\alpha_0)W_n}e^{-\tau_0M_n}-I_n]\tilde{V}_{nt}$, $b_{it}=e_i[e^{\hat{\tau}_{nT}M_n}(e^{(\hat{\alpha}_{nT}-\alpha_0)W_n}-I_n)]\tilde{X}_{nt}\beta_0$, and $c_{it}=e_ie^{\hat{\tau}_{nT}M_n}\tilde{X}_{nt}(\beta_0-\hat{\beta}_{nT})$, where e_i is the ith row of the $n\times n$ identity matrix I_n. Thus,

$$\begin{aligned}C_3&=\frac{1}{nT^3}\sum_{i=1}^n\sum_{j=1}^n\big(\frac{T}{T-1}\sum_{t=1}^T(a_{it}+b_{it}+c_{it}+\tilde{v}_{it})^2-\tilde{f}_{ni}\big)\big(\frac{T}{T-1}\sum_{t=1}^T(a_{jt}+b_{jt}+c_{jt}+\tilde{v}_{jt})^2-\tilde{f}_{nj}\big)p_{ij}\\&=\frac{1}{nT^3}\sum_{i=1}^n\sum_{j=1}^n\big(\frac{T}{T-1}\sum_{t=1}^T(a_{it}^2+b_{it}^2+c_{it}^2+2a_{it}b_{it}+2a_{it}c_{it}+2a_{it}\tilde{v}_{it}+2b_{it}c_{it}+2b_{it}\tilde{v}_{it}+2c_{it}\tilde{v}_{it})\big)\\&\times\big(\frac{T}{T-1}\sum_{t=1}^T(a_{jt}^2+b_{jt}^2+c_{jt}^2+2a_{jt}b_{jt}+2a_{jt}c_{jt}+2a_{jt}\tilde{v}_{jt}+2b_{jt}c_{jt}+2b_{jt}\tilde{v}_{jt}+2c_{jt}\tilde{v}_{jt})\big)p_{ij}.\end{aligned}\tag{A.11}$$

We now show that $\frac{1}{nT^3}\sum_{i=1}^n\sum_{j=1}^n(\frac{T}{T-1}\sum_{t=1}^Ta_{it}^2)(\frac{T}{T-1}\sum_{t=1}^Ta_{jt}^2)p_{ij}=o_p(1)$. By the proof of Proposition 5 of the supplement to Debarsy et al. (2015) (pages 12–13), $a_{it}^2\le5(t_{it,1}+t_{it,2}+t_{it,3}+t_{it,4}+t_{it,5})$, where $t_{it,1}=(e_iM_n\tilde{V}_{nt})^2(\hat{\tau}_{nT}-\tau_0)^2$, $t_{it,2}=(e_iW_n\tilde{V}_{nt})^2(\hat{\alpha}_{nT}-\alpha_0)^2$,

$$t_{it,3}=\frac{1}{4}\big(e_iM_n^2e^{(\tilde{\tau}-\tau_0)M_n}e^{(\tilde{\alpha}-\alpha_0)W_n}\tilde{V}_{nt}\big)^2(\hat{\tau}_{nT}-\tau_0)^4,$$

$t_{it,4}=\frac{1}{4}\big(e_iW_n^2e^{(\tilde{\tau}-\tau_0)M_n}e^{(\tilde{\alpha}-\alpha_0)W_n}\tilde{V}_{nt}\big)^2(\hat{\alpha}_{nT}-\alpha_0)^4$, and $t_{it,5}=\big(e_iW_nM_ne^{(\tilde{\tau}-\tau_0)M_n}e^{(\tilde{\alpha}-\alpha_0)W_n}\tilde{V}_{nt}\big)^2(\hat{\alpha}_{nT}-\alpha_0)^2(\hat{\tau}_{nT}-\tau_0)^2$, where $\tilde{\alpha}$ is between $\hat{\alpha}_{nT}$ and α_0, and $\tilde{\tau}$ is between $\hat{\tau}_{nT}$ and τ_0. We need to show that

$$\frac{1}{nT^3}\sum_{i=1}^n\sum_{j=1}^n(\frac{T}{T-1}\sum_{t=1}^Tt_{it,k})(\frac{T}{T-1}\sum_{t=1}^Tt_{jt,l})p_{ij}=o_p(1)\tag{A.12}$$

for $k, l = 1, \ldots, 5$. For $k = 1$ and $l = 1$,

$$\begin{aligned}
&\frac{1}{nT^3}\sum_{i=1}^{n}\sum_{j=1}^{n}(\frac{T}{T-1}\sum_{t=1}^{T}t_{it,1})(\frac{T}{T-1}\sum_{t=1}^{T}t_{jt,1})p_{ij} \\
&= \frac{1}{nT^3}\sum_{i=1}^{n}\sum_{j=1}^{n}\Big(\frac{T}{T-1}\big((e_iM_n\tilde{V}_{n1})^2+\cdots+(e_iM_n\tilde{V}_{nT})^2\big)\Big) \\
&\quad \times\Big(\frac{T}{T-1}\big((e_jM_n\tilde{V}_{n1})^2+\cdots+(e_jM_n\tilde{V}_{nT})^2\big)\Big)(\hat{\tau}_{nT}-\tau_0)^4p_{ij}.
\end{aligned}$$

Note that

$$\begin{aligned}
&\sum_{i=1}^{n}\sum_{j=1}^{n}(e_iM_n\tilde{V}_{nt_1})^2(e_jM_n\tilde{V}_{nt_2})^2p_{ij} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k_1=1}^{n}\sum_{k_2=1}^{n}\sum_{l_1=1}^{n}\sum_{l_2=1}^{n}m_{ik_1}m_{ik_2}m_{il_1}m_{il_2}\tilde{v}_{k_1t_1}\tilde{v}_{k_2t_1}\tilde{v}_{l_1t_2}\tilde{v}_{l_2t_2}p_{ij},
\end{aligned}$$

where $t_1, t_2 = 1, \ldots, T$. By the Cauchy-Schwarz inequality,

$$E|\tilde{v}_{k_1t_1}\tilde{v}_{k_2t_1}\tilde{v}_{l_1t_2}\tilde{v}_{l_2t_2}| \le E^{\frac{1}{2}}(\tilde{v}_{k_1t_1}^2\tilde{v}_{k_2t_1}^2)E^{\frac{1}{2}}(\tilde{v}_{l_1t_2}^2\tilde{v}_{l_2t_2}^2) \le E^{\frac{1}{4}}(\tilde{v}_{k_1t_1}^4)E^{\frac{1}{4}}(\tilde{v}_{k_2t_1}^4)E^{\frac{1}{4}}(\tilde{v}_{l_1t_2}^4)E^{\frac{1}{4}}(\tilde{v}_{l_2t_2}^4) \le c$$

for some constant c. Thus, as P_n and M_n are UB,

$$\begin{aligned}
&E|\sum_{i=1}^{n}\sum_{j=1}^{n}(e_iM_n\tilde{V}_{nt_1})^2(e_jM_n\tilde{V}_{nt_2})^2p_{ij}| \\
&\le c\sum_{i=1}^{n}(\sum_{j=1}^{n}|p_{ij}|)(\sum_{k_1=1}^{n}|m_{ik_1}|)(\sum_{k_2=1}^{n}|m_{ik_2}|)(\sum_{l_1=1}^{n}|m_{il_1}|)(\sum_{l_2=1}^{n}|m_{il_2}|) \le nk
\end{aligned}$$

for some constant k. It follows that

$$\frac{1}{nT^3}\sum_{i=1}^{n}\sum_{j=1}^{n}(\frac{T}{T-1}\sum_{t=1}^{T}t_{it,1})(\frac{T}{T-1}\sum_{t=1}^{T}t_{jt,1})p_{ij} = o_p(1)$$

by Markov's inequality. Similarly, the terms in the expression of (A.12) involving $t_{it,1}$ or $t_{it,2}$ are $o_p(1)$. From the proof of Proposition 5 of the supplement to Debarsy et al. (2015), $t_{it,3} \le c\tilde{V}_{nt}'\tilde{V}_{nt}(\hat{\tau}_{nt}-\tau_0)^4$ for some constant c. Then

$$\frac{1}{nT^3}\sum_{i=1}^{n}\sum_{j=1}^{n}(\frac{T}{T-1}\sum_{t=1}^{T}t_{it,3})^2|p_{ij}| \le \frac{c^2(\hat{\tau}_{nT}-\tau_0)^8}{nT(T-1)^2}\sum_{i=1}^{n}\sum_{j=1}^{n}(\tilde{V}_{n1}'\tilde{V}_{n1}+\cdots+\tilde{V}_{nT}'\tilde{V}_{nT})^2|p_{ij}|.$$

By the Cauchy-Schwarz inequality,

$$E(|\tilde{V}_{nm}'\tilde{V}_{nm}\tilde{V}_{nq}'\tilde{V}_{nq}|) \le E^{\frac{1}{2}}[(\tilde{V}_{nm}'\tilde{V}_{nm})^2]E^{\frac{1}{2}}[(\tilde{V}_{nq}'\tilde{V}_{nq})^2] \le n^2\xi$$

for some constant ξ, where m, $q = 1, \cdots, T$. Then by Markov's inequality, $\frac{1}{n^3T^2}\sum_{i=1}^{n}\sum_{j=1}^{n}(\tilde{V}_{n1}'\tilde{V}_{n1}+\cdots+\tilde{V}_{nT}'\tilde{V}_{nT})^2|p_{ij}| = O_p(1)$. As $\sqrt{nT}(\hat{\tau}_{nT}-\tau_0) = O_p(1)$, $\frac{1}{nT^3}\sum_{i=1}^{n}\sum_{j=1}^{n}(\frac{T}{T-1}\sum_{t=1}^{T}t_{it,3})^2p_{ij} = o_p(1)$. Similarly, the terms in the expression of (A.12) involving two of $t_{k,3}$, $t_{k,4}$ and $t_{k,5}$ are $o_p(1)$. For

$$\frac{1}{nT^3}\sum_{i=1}^{n}\sum_{j=1}^{n}(\frac{T}{T-1}\sum_{t=1}^{T}t_{it,1})(\frac{T}{T-1}\sum_{t=1}^{T}t_{jt,3})p_{ij},$$

$$\frac{1}{nT^3}\sum_{i=1}^{n}\sum_{j=1}^{n}(\frac{T}{T-1}\sum_{t=1}^{T}t_{it,1})(\frac{T}{T-1}\sum_{t=1}^{T}t_{jt,3})|p_{ij}|$$
$$\leq \frac{1}{nT^3}\sum_{i=1}^{n}\sum_{j=1}^{n}\Big(\frac{T}{T-1}\big((e_iM_n\tilde{V}_{n1})^2+\cdots+(e_iM_n\tilde{V}_{nT})^2\big)\Big)\Big(\frac{T}{T-1}(\tilde{V}_{n1}'\tilde{V}_{n1}+\cdots+\tilde{V}_{nt}'\tilde{V}_{nt})\Big)(\hat{\tau}_{nt}-\tau_0)^6|p_{ij}|.$$

Then with an argument similar to that for the case with $k=1$ and $l=1$, we have

$$\frac{1}{nT^3}\sum_{i=1}^{n}\sum_{j=1}^{n}(\frac{T}{T-1}\sum_{t=1}^{T}t_{it,1})(\frac{T}{T-1}\sum_{t=1}^{T}t_{jt,3})p_{ij}=o_p(1).$$

Thus, the terms in the expression of (A.12) involving $t_{it,5}$ and one of $t_{it,1}$ and $t_{it,2}$, or $t_{it,5}$ and one of $t_{it,3}$ and $t_{it,4}$ are $o_p(1)$. Hence, we have $\frac{1}{nT^3}\sum_{i=1}^{n}\sum_{j=1}^{n}(\frac{T}{T-1}\sum_{t=1}^{T}t_{it,k})(\frac{T}{T-1}\sum_{t=1}^{T}t_{jt,l})p_{ij}=o_p(1)$. As shown in the proof of Proposition 5 in the supplement to Debarsy et al. (2015) (page 12), terms in the expression of C_3 in (A.11) involving $|b_{it}|$ and $|c_{it}|$ are $o_p(1)$. Then by Markov's inequality, the terms in the expression of C_3 in (A.11) involving b_{it} or c_{it} are $o_p(1)$, and so are the terms involving a_{it}. Thus, C_3 is $o_p(1)$. Similarly, $C_1=o_p(1)$ and $C_2=o_p(1)$. Hence, (b) holds under Assumption A.2. It follows that $A_1=o_p(1)$.

The proof of $A_2=o_p(1)$: We shall show that (a) $\frac{T-2}{nT^2}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n}\big((\frac{T}{T-1}\tilde{v}_{it}^2)(\frac{T}{T-1}\tilde{v}_{jt}^2)-\sigma_{it}^2\sigma_{jt}^2\big)p_{ij}=o_p(1)$ and (b) $\frac{T-2}{nT^2}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n}\big((\frac{T}{T-1}\hat{\tilde{v}}_{it}^2)(\frac{T}{T-1}\hat{\tilde{v}}_{jt}^2)-(\frac{T}{T-1}\tilde{v}_{it}^2)(\frac{T}{T-1}\tilde{v}_{jt}^2)\big)p_{ij}=o_p(1)$. Note that

$$\frac{T-2}{nT^2}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n}((\frac{T}{T-1}\tilde{v}_{it}^2)(\frac{T}{T-1}\tilde{v}_{jt}^2)-\sigma_{it}^2\sigma_{jt}^2)p_{ij}=B_{2,1}+B_{2,2}+B_{2,3}+B_{2,4},$$

where

$$B_{2,1}=\frac{T-2}{n(T-1)^2}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n}[\tilde{v}_{it}^2-E(\tilde{v}_{it}^2)][\tilde{v}_{jt}^2-E(\tilde{v}_{jt}^2)]p_{ij},$$
$$B_{2,2}=\frac{T-2}{n(T-1)^2}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n}E(\tilde{v}_{it}^2)[\tilde{v}_{jt}^2-E(\tilde{v}_{jt}^2)]p_{ij},$$
$$B_{2,3}=\frac{T-2}{n(T-1)^2}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n}[\tilde{v}_{it}^2-E(\tilde{v}_{it}^2)]E(\tilde{v}_{jt}^2)p_{ij},$$

and $B_{2,4}=\frac{T-2}{nT^2}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n}\big(\frac{T^2}{(T-1)^2}E(\tilde{v}_{it}^2)E(\tilde{v}_{jt}^2)-\sigma_{it}^2\sigma_{jt}^2\big)p_{ij}$. For $B_{2,1}$, denote $\varrho_{nT}=[\chi_{n1}',\cdots,\chi_{nT}']'$, where $\chi_{nt}=[\chi_{nt,1},\cdots,\chi_{nt,n}]'$ with $\chi_{nt,i}=\tilde{v}_{it}^2-E(\tilde{v}_{it}^2)$. Since $E\chi_{nt,i}=0$, and if $i=j$, $p_{ii}=0$, $E\big(\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n}\chi_{nt,i}\chi_{nt,j}p_{ij}\big)=0$. As

$$\varrho_{nT}'(I_T\otimes P_n)\varrho_{nT}\varrho_{nT}'(I_T\otimes P_n)\varrho_{nT}=\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{T}\sum_{r=1}^{n}\sum_{s=1}^{n}\chi_{nt,i}\chi_{nt,j}\chi_{nk,r}\chi_{nk,s}p_{ij}p_{rs},$$

the mutual independence of $\chi_{nt,i}$'s over i, the correlation of $\chi_{nt,i}$'s over t and $p_{ii}=0$ imply that $E(\varrho_{nT}'(I_T\otimes P_n)\varrho_{nT}\varrho_{nT}'(I_T\otimes P_n)\varrho_{nT})$ only if $(i=r\neq j=s)$ or $(i=s\neq j=r)$. Note that as $E|\tilde{v}_{it}^4|$'s exist and are bounded uniformly in i and t, $E\big(|\chi_{nt,i}\chi_{nt,j}\chi_{nk,j}\chi_{nk,i}|\big)=E(|\chi_{nt,i}\chi_{nk,i}|)E(|\chi_{nt,j}\chi_{nk,j}|)$ are uniformly bounded for $i \neq j$. It follows that

$$
\begin{aligned}
&\text{Var}(\varrho'_{nT}(I_T \otimes P_n)\varrho_{nT}) = E[(\varrho'_{nT}(I_T \otimes P_n)\varrho_{nT})^2] \\
&= E\big(\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{T}\sum_{r=1}^{n}\sum_{s=1}^{n}\chi_{nt,i}\chi_{nt,j}\chi_{nk,r}\chi_{nk,s}p_{ij}p_{rs}\big) \\
&\leq \sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{T}\sum_{r=1}^{n}\sum_{s=1}^{n}E\big(|\chi_{nt,i}\chi_{nt,j}\chi_{nk,r}\chi_{nk,s}|\cdot|p_{ij}|\cdot|p_{rs}|\big) \\
&= \sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{k=1}^{T}\sum_{j\neq i}^{n}E\big(|\chi_{nt,i}\chi_{nt,j}\chi_{nk,i}\chi_{nk,j}|\cdot|p_{ij}|^2\big)+\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{k=1}^{T}\sum_{j\neq i}^{n}E\big(|\chi_{nt,i}\chi_{nt,j}\chi_{nk,j}\chi_{nk,i}|\cdot|p_{ij}|^2\big) \\
&= O(nT^2)
\end{aligned}
$$

since P_n is UB. Hence, the variance of $B_{2,1}$ is $o(1)$ as T tends to infinity. By the generalized Chebyshev inequality, $B_{2,1} = o_p(1)$. Similarly, $B_{2,2} = o_p(1)$ and $B_{2,3} = o_p(1)$ hold. As $|\frac{T}{T-1}E(\tilde{v}_{it}^2) - \sigma_{it}^2| = \frac{1}{T-1}\sigma_{it}^2 + \frac{1}{T(T-1)}\sum_{k=1}^{T}\sigma_{ik}^2 \leq \frac{c}{T-1}$ for some constant c, $B_{2,4} = o_p(1)$ as both n and T tend to infinity. The proof of (b) is omitted as it is similar to the proof of (b) in the proof of $A_1 = o_p(1)$.

The proof of (ii) is similar to the proof of (i).

Proof of (iii): As $||e^{(\hat{\tau}_{nT}-\tau_0)M_n}||_\infty = o_p(1)$, it is similar to the proof of Proposition 5(iii) of the supplement to Debarsy et al. (2015).

Proof of (iv): We can rewrite (iv) as

$$
\begin{aligned}
&\frac{1}{nT}R'_{nT}(J_T \otimes e^{\hat{\tau}_{nT}M'_n})\hat{\Sigma}_{nT}(J_T \otimes e^{\hat{\tau}_{nT}M_n})S_{nT} - \frac{1}{nT}R'_{nT}(J_T \otimes e^{\tau_0 M'_n})\Sigma_{nT}(J_T \otimes e^{\tau_0 M_n})S_{nT} \\
&= \frac{1}{nT}\Big[R'_{nT}(J_T \otimes e^{\tau_0 M'_n})(\hat{\Sigma}_{nT} - \Sigma_{nT})(J_T \otimes e^{\tau_0 M_n})S_{nT} \\
&\quad + R'_{nT}(J_T \otimes e^{\tau_0 M'_n})(\hat{\Sigma}_{nT} - \Sigma_{nT})\big(J_T \otimes (e^{\hat{\tau}_{nT}M_n} - e^{\tau_0 M_n})\big)S_{nT} \\
&\quad + R'_{nT}\big(J_T \otimes (e^{\hat{\tau}_{nT}M'_n} - e^{\tau_0 M'_n})\big)(\hat{\Sigma}_{nT} - \Sigma_{nT})(J_T \otimes e^{\hat{\tau}_{nT}M_n})S_{nT} \\
&\quad + R'_{nT}\big(J_T \otimes (e^{\hat{\tau}_{nT}M'_n} - e^{\tau_0 M'_n})\big)\Sigma_{nT}(J_T \otimes e^{\hat{\tau}M_n})S_{nT} \\
&\quad + R'_{nT}(J_T \otimes e^{\tau_0 M'_n})\Sigma_{nT}\big(J_T \otimes (e^{\hat{\tau}_{nT}M_n} - e^{\tau_0 M_n})\big)S_{nT}\Big].
\end{aligned}
\tag{A.13}
$$

For the first term on the r.h.s. of (A.13), it can be proved to be $o_p(1)$ as the term in (ii). For the second term, note that by the sub-multiplicative property of the row sum matrix norm,

$$
\begin{aligned}
&|\frac{1}{nT}R'_{nT}(J_T \otimes e^{\tau_0 M'_n})(\hat{\Sigma}_{nT} - \Sigma_{nT})\big(J_T \otimes (e^{\hat{\tau}_{nT}M_n} - e^{\tau_0 M_n})\big)S_{nT}| \\
&\leq \frac{c}{nT}||e^{\hat{\tau}_{nT}M_n} - e^{\tau_0 M_n}||_\infty||\hat{\Sigma}_{nT} - \Sigma_{nT}||_\infty \\
&\leq \frac{c}{nT}||e^{\hat{\tau}_{nT}M_n} - e^{\tau_0 M_n}||_\infty\sum_{t=1}^{T}\sum_{i=1}^{n}[(|\frac{T}{T-1}\hat{\tilde{v}}_{it}^2 - \frac{T}{T-1}\tilde{v}_{it}^2|) + (|\frac{T}{T-1}\tilde{v}_{it}^2 - \sigma_{it}^2|)]
\end{aligned}
$$

for some constant c. We note that $E\big|\frac{T}{T-1}\tilde{v}_{it}^2 - \sigma_{it}^2\big| \leq E\big(\big|\frac{T}{T-1}\tilde{v}_{it}^2\big| + \big|\sigma_{it}^2\big|\big) \leq k$ for some constant k. Then by

Markov's inequality, $\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left|\tilde{v}_{it}^{2}-\sigma_{it}^{2}\right|=O_p(1)$. The equation of

$$\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}|\hat{\tilde{v}}_{it}^{2}-\tilde{v}_{it}^{2}|=\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}|a_{it}^{2}+b_{it}^{2}+c_{it}^{2}+2a_{it}b_{it}+2a_{it}c_{it}$$
$$+2a_{it}\tilde{v}_{it}+2b_{it}c_{it}+2b_{it}\tilde{v}_{it}+2c_{it}\tilde{v}_{it}|=o_p(1)$$

can be shown as in the proof of (i)(b). By Lemma A.8 in the supplement to Debarsy et al. (2015), the second term on the r.h.s. of (A.13) is $o_p(1)$. Similarly, the third term is $o_p(1)$. By using the sub-multiplicative property of the row sum matrix norm, the last two terms are $o_p(1)$. Hence, (iv) holds.

In summary, Theorem 3 holds when both n and T are large since (i)–(iv) hold.

(II) Proof of $\hat{\mathcal{H}}_{\zeta_0,nT}-\mathcal{H}_{\zeta_0,nT}=o_p(1)$ and $\hat{\Delta}_{\zeta_0,nT}-\Delta_{\zeta_0,nT}=o_p(1)$ under the conditions that n is large, T is finite and $\sigma_{it}^{2}=\sigma_{i}^{2}$: Note that $E(\tilde{v}_{it}^{2})=\frac{T-1}{T}\sigma_{i}^{2}$. Suppose that we replace σ_{it}^{2} with $\frac{T}{T-1}(v_{it}-\frac{1}{T}\sum_{t=1}^{T}v_{it})^{2}$. Corresponding to (A.7) and (A.8), we have

$$tr[(J_T\otimes A_n)\Sigma_{nT}]=(T-1)\sum_{i=1}^{n}a_{ii}\sigma_{i}^{2}.$$

and

$$(T-1)E\sum_{i=1}^{n}a_{ii}\Big(\frac{T}{T-1}(v_{it}-\frac{1}{T}\sum_{t=1}^{T}v_{it})^{2}\Big)=(T-1)\sum_{i=1}^{n}a_{ii}\sigma_{i}^{2};$$

and corresponding to (A.9) and (A.10), we have

$$tr\big[\Sigma_{nT}(J_T\otimes W_n^{s})\Sigma_{nT}(J_T\otimes M_n^{s})\big]=\frac{T-1}{T}\sum_{t=1}^{T}\sum_{p=1}^{n}\sum_{q=1}^{n}\sigma_{p}^{2}\sigma_{q}^{2}w_{pq}^{s}m_{qp}^{s}.$$

and

$$\frac{T^{2}}{(T-1)^{2}}E\Big(\frac{1}{T^{2}}\sum_{t=1}^{T}\sum_{p=1}^{n}\sum_{q=1}^{n}\sum_{j=1}^{T}\tilde{v}_{pt}^{2}\tilde{v}_{qj}^{2}w_{pq}^{s}m_{qp}^{s}+\frac{T-2}{T}\sum_{t=1}^{T}\sum_{p=1}^{n}\sum_{q=1}^{n}\tilde{v}_{pt}^{2}\tilde{v}_{qt}^{2}w_{pq}^{s}m_{qp}^{s}\Big)$$
$$=\frac{T-1}{T}\sum_{t=1}^{T}\sum_{p=1}^{n}\sum_{q=1}^{n}\sigma_{p}^{2}\sigma_{q}^{2}w_{pq}^{s}m_{qp}^{s}.$$

Thus, we can replace σ_{it}^{2} with $\frac{T}{T-1}(v_{it}-\frac{1}{T}\sum_{t=1}^{T}v_{it})^{2}$ instead to derive terms with expected value equal to $tr[(J_T\otimes A_n)\Sigma_{nT}]$ and $tr\big[\Sigma_{nT}(J_T\otimes W_n^{s})\Sigma_{nT}(J_T\otimes M_n^{s})\big]$ respectively.

When v_{it} are set as $v_{it}\sim(0,\sigma_{i}^{2})$, Theorem 3 holds since (i), (ii), (iii) and (iv) of (I) hold. It is $E(\tilde{v}_{it}^{2})=\frac{T-1}{T}\sigma_{i}^{2}$ that ensures that as long as n tends to infinity, the four equations hold regardless of whether T is fixed or tends to infinity. In this case, the proofs of the four equations are similar to those proofs in (I), thus they are omitted.

References

[1] Davidson J. Stochastic Limit Theory [M]. Oxford University Press, (1994).

[2] Debarsy N, Jin F, Lee L F. Large sample properties of the matrix exponential spatial specification with an application to FDI [J]. *Journal of Econometrics*, 188 (1): 1–21 (2015).

[3] Kelejian H H, Prucha I R. Specification and Estimation of Spatial Autoregressive Models with Autoregressive and Heteroskedastic Disturbances [J]. *Journal of Econometrics*, 157 (1): 53–67 (2010).

[4] Lin X, Lee L.F. GMM estimation of spatial autoregressive models with unknown heteroskedasticity [J]. *Journal of Econometrics*, 157(1): 34–52 (2010).

[5] Lee L F, Yu J. Estimation of spatial autoregressive panel data models with fixed effects [J]. *Journal of Econometrics*, 154 (2): 165–185 (2010).

[6] Yu J, de Jong R., Lee L.F. Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and T are large [J]. *Journal of Econometrics*, 146(1): 118–134 (2008).