Spatial-temporal Model with Heterogeneous Random Effects

Xingdong Feng, Wenyu Li and Qianqian Zhu

School of Statistics and Management

Shanghai University of Finance and Economics

Supplementary Material

This supplementary material includes a residual-based bootstrapping procedure for the QMLE together with its asymptotic validity, and provides additional simulation results on the finite-sample performance of the WCQE and WQAE. It also includes the detailed forms of the matrices Σ_N, $\Sigma_N^\star$, $\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)$ in Section 3 and the vector $\boldsymbol{Y}_{-1}$, as well as technical details for Theorems 1–6 and Propositions 1–2. Moreover, Lemmas 1–10 are introduced to establish Theorems 1–6.

Throughout the appendix, $\otimes$ denotes the Kronecker product of two matrices, $|\cdot|$ denotes the absolute value of a scalar/vector or the determinant of a matrix, $\|\cdot\|$ denotes the Euclidean norm of a vector, and $tr(\cdot)$ denotes the trace of a matrix. Denote N and T as the total numbers of spatial units and time periods, respectively. Define $I(\cdot)$ as the indicator function. For a positive integer m, denote I_m as an $m \times m$ identity matrix, $\mathbf{0}_m$ as an $m \times 1$ vector of zeros and $\boldsymbol{\iota}_m$ as an $m \times 1$ vector of ones, and let $J_m \equiv \boldsymbol{\iota}_m \boldsymbol{\iota}_m'$. For an $n \times m$ matrix A, denote $A = [a_{ij}]$, where a_{ij} is the ith row and jth column element of A. The operator E denotes the expectation with respect to the probability measure, $\mathbb{E}_n$ denotes the expectation with respect to the empirical measure, and $\mathbb{G}_n = \sqrt{n}(\mathbb{E}_n - E)$. In addition, $\ell^\infty(\mathcal{T})$ denotes the space of all uniformly bounded functions on $\mathcal{T}$. Moreover, $\rightarrow_p$ denotes convergence in probability, $\rightarrow_d$ denotes convergence in distribution, and $\rightsquigarrow$ denotes weak convergence.

S1 Bootstrapping Procedure

To circumvent difficulties of calculating the covariance matrix of $\widehat{\boldsymbol{\zeta}}$ in practice, we propose a residual-based bootstrapping procedure to approximate matrix Γ in (3.9); see also Su and Yang (2015). Denote the score function $\partial \ln \mathcal{L}_{NT}(\boldsymbol{\zeta})/\partial \boldsymbol{\zeta}$ by $\boldsymbol{s}(\boldsymbol{Y}_{-1}, \boldsymbol{V}_{NT}, \boldsymbol{\zeta})$. Let $\bar{A}_N(\boldsymbol{\beta}^\star) = A_N(\boldsymbol{\beta}^\star) + I_N$ and $\bar{A}_{0N} = \bar{A}_N(\boldsymbol{\beta}_0^\star)$. The bootstrapping procedure for estimating Γ is summarized as follows:

(1) Based on the QMLE $\widehat{\boldsymbol{\zeta}}$, compute the residuals $\widehat{\boldsymbol{v}}_t = \boldsymbol{y}_t - \widehat{\alpha}\boldsymbol{y}_{t-1} - \widehat{\lambda} W_N \boldsymbol{y}_t - Z_t\widehat{\boldsymbol{\gamma}} - X\widehat{\boldsymbol{\psi}}$ and the transformed residuals $\widehat{\boldsymbol{r}}_t = \bar{A}_N^{-1/2}(\widehat{\boldsymbol{\beta}}^\star)\widehat{\boldsymbol{v}}_t$ for $t = 1, \ldots, T$. For each t, center $\widehat{\boldsymbol{r}}_t$ by its mean. (Notice that for $\boldsymbol{r}_t = \bar{A}_{0N}^{-1/2}\boldsymbol{v}_t$ with $\boldsymbol{v}_t = \boldsymbol{y}_t - \alpha_0\boldsymbol{y}_{t-1} - \lambda_0 W_N \boldsymbol{y}_t - Z_t\boldsymbol{\gamma}_0 - X\boldsymbol{\psi}_0$, $E(\boldsymbol{r}_t) = 0$ and $\mathrm{cov}(\boldsymbol{r}_t) = \sigma_{\varepsilon 0}^2 I_N$.)

(2) Make N random draws from the rows of $(\widehat{\boldsymbol{r}}_1, \ldots, \widehat{\boldsymbol{r}}_T)$ to give T matched bootstrap samples, $\{\widehat{\boldsymbol{r}}_1^b, \ldots, \widehat{\boldsymbol{r}}_T^b\}$, of the transformed residuals.

(3) Conditional on $\boldsymbol{y}_0$, X, $Z_1, \ldots, Z_T$, and the QMLE $\widehat{\boldsymbol{\zeta}}$, generate the bootstrap data according to

$$\widehat{\boldsymbol{y}}_1^b = \left(I_N - \widehat{\lambda} W_N\right)^{-1}\left(\widehat{\alpha}\boldsymbol{y}_0 + Z_1\widehat{\boldsymbol{\gamma}} + X\widehat{\boldsymbol{\psi}} + \widehat{\boldsymbol{v}}_1^b\right),$$
$$\widehat{\boldsymbol{y}}_t^b = \left(I_N - \widehat{\lambda} W_N\right)^{-1}\left(\widehat{\alpha}\widehat{\boldsymbol{y}}_{t-1}^b + Z_t\widehat{\boldsymbol{\gamma}} + X\widehat{\boldsymbol{\psi}} + \widehat{\boldsymbol{v}}_t^b\right),\ t = 2, \ldots, T,$$

where $\widehat{\boldsymbol{v}}_t^b = \bar{A}_N^{1/2}(\widehat{\boldsymbol{\beta}}^\star)\widehat{\boldsymbol{r}}_t^b$ for $t = 1, \ldots, T$. Then the bootstrapped values of $\boldsymbol{V}_{NT}$, $\boldsymbol{Y}$ and $\boldsymbol{Y}_{-1}$ are given by $\widehat{\boldsymbol{V}}_{NT}^b = \mathrm{vec}(\widehat{\boldsymbol{v}}_1^b, \ldots, \widehat{\boldsymbol{v}}_T^b)$, $\widehat{\boldsymbol{Y}}^b = \mathrm{vec}(\widehat{\boldsymbol{y}}_1^b, \ldots, \widehat{\boldsymbol{y}}_T^b)$ and $\widehat{\boldsymbol{Y}}_{-1}^b = \mathrm{vec}(\boldsymbol{y}_0, \ldots, \widehat{\boldsymbol{y}}_{T-1}^b)$, respectively, where the operator $\mathrm{vec}(\cdot)$ vectorizes a matrix to a vector by columns.

(4) Compute the bootstrapped score function by $\boldsymbol{s}^b = \boldsymbol{s}\left(\widehat{\boldsymbol{Y}}_{-1}^b, \widehat{\boldsymbol{V}}_{NT}^b, \widehat{\boldsymbol{\zeta}}\right)$.

(5) Repeat steps (2)–(4) B times, and the bootstrap estimator of Γ is defined as

$$\widehat{\Gamma}_B = \frac{1}{NT}\left[\frac{1}{B}\sum_{b=1}^{B}(\boldsymbol{s}^b - \bar{\boldsymbol{s}}_B)(\boldsymbol{s}^b - \bar{\boldsymbol{s}}_B)'\right],$$

where $\bar{\boldsymbol{s}}_B = \frac{1}{B}\sum_{b=1}^{B}\boldsymbol{s}^b$.

Proposition 2. *Suppose the conditions of Theorem 2 hold, then as $N \to \infty$,*

$$\frac{1}{NT}E^*\left[\left(\boldsymbol{s}^b - E^*\left(\boldsymbol{s}^b\right)\right)\left(\boldsymbol{s}^b - E^*\left(\boldsymbol{s}^b\right)\right)'\right] - \Gamma = o_p(1),$$

where the expectation operator E^* corresponds to the bootstrap probability space.

Proposition 2 implies $\widehat{\Gamma}_B - \Gamma = o_p(1)$ as $N \to \infty$ and $B \to \infty$, which establishes the validity of the proposed residual-based bootstrapping procedure.

S2 Additional simulation results

In this section, we provide additional results for Tables 2 and 3 in simulation study.

From Table 2 in the manuscript, we can see that the ESDs of WCQE are obviously smaller than the corresponding ASDs at $\tau = 0.5$ due to the same sign restriction in optimization. Here, we provide the results of WCQE at $\tau = 0.5$ for larger N in Table S.1. It can be found that the ESDs will move closer to their corresponding ASDs for larger N.

Table S.2 reports the results of WQAE $\widehat{\beta}_1(\widehat{\boldsymbol{\pi}}_{opt,K})$ for settings $(N, T) = (1500, 300)$ and $(3000, 600)$. Comparing with the results of small samples in Table 3 of the manuscript, we can see that the ESDs of $\widehat{\beta}_1(\widehat{\boldsymbol{\pi}}_{opt,K})$ will move closer to their corresponding ASDs for larger N.

S3 Structures of Matrices and Vectors

S3.1 Σ_N and $\Sigma_N^\star$

This subsection includes detailed structure of matrices Σ_N and $\Sigma_N^\star$ in Section 3. For simplicity, hereinafter the first derivative of $\ln \mathcal{L}_{NT}(\boldsymbol{\zeta}_0)$ is denoted as $\dot{\boldsymbol{d}}_\zeta = (\dot{\boldsymbol{d}}_\phi', \dot{d}_{\sigma_\varepsilon^2}, \dot{d}_\lambda, \dot{\boldsymbol{d}}_{\beta^\star}')' \triangleq \left(\frac{\partial \ln \mathcal{L}_{NT}(\boldsymbol{\zeta}_0)}{\partial \boldsymbol{\phi}'}, \frac{\partial \ln \mathcal{L}_{NT}(\boldsymbol{\zeta}_0)}{\partial \sigma_\varepsilon^2}, \frac{\partial \ln \mathcal{L}_{NT}(\boldsymbol{\zeta}_0)}{\partial \lambda}, \frac{\partial \ln \mathcal{L}_{NT}(\boldsymbol{\zeta}_0)}{\partial \boldsymbol{\beta}^{\star\prime}}\right)'$, and its second derivatives are denoted as $\ddot{d}_{\phi\phi'} \triangleq \frac{\partial^2 \ln \mathcal{L}_{NT}(\boldsymbol{\zeta}_0)}{\partial \boldsymbol{\phi} \partial \boldsymbol{\phi}'}$, $\ddot{\boldsymbol{d}}_{\sigma_\varepsilon^2\phi'} \triangleq \frac{\partial^2 \ln \mathcal{L}_{NT}(\boldsymbol{\zeta}_0)}{\partial \sigma_\varepsilon^2 \partial \boldsymbol{\phi}'}$, $\ddot{d}_{\sigma_\varepsilon^2\sigma_\varepsilon^2} \triangleq \frac{\partial^2 \ln \mathcal{L}_{NT}(\boldsymbol{\zeta}_0)}{\partial \sigma_\varepsilon^2 \partial \sigma_\varepsilon^2}$, and so on.

Then Σ_N and $\Sigma_N^\star$ can be rewritten as follows,

$$\Sigma_N = -E\left(\frac{1}{NT}\frac{\partial^2 \ln \mathcal{L}_{NT}(\boldsymbol{\zeta}_0)}{\partial \boldsymbol{\zeta} \partial \boldsymbol{\zeta}'}\right) = -\frac{1}{NT}E\begin{pmatrix} \ddot{d}_{\phi\phi'} & * & * & * \\ \ddot{\boldsymbol{d}}_{\sigma_\varepsilon^2\phi'} & \ddot{d}_{\sigma_\varepsilon^2\sigma_\varepsilon^2} & * & * \\ \ddot{\boldsymbol{d}}_{\lambda\phi'} & \ddot{d}_{\lambda\sigma_\varepsilon^2} & \ddot{d}_{\lambda\lambda} & * \\ \ddot{d}_{\beta^\star\phi'} & \ddot{\boldsymbol{d}}_{\beta^\star\sigma_\varepsilon^2} & \ddot{\boldsymbol{d}}_{\beta^\star\lambda} & \ddot{d}_{\beta^\star\beta^{\star\prime}} \end{pmatrix},$$

$$\begin{aligned}\Sigma_N^\star &= E\left(\frac{1}{\sqrt{NT}}\frac{\partial \ln \mathcal{L}_{NT}(\boldsymbol{\zeta}_0)}{\partial \boldsymbol{\zeta}}\frac{1}{\sqrt{NT}}\frac{\partial \ln \mathcal{L}_{NT}(\boldsymbol{\zeta}_0)}{\partial \boldsymbol{\zeta}'}\right) - \Sigma_N \\ &= \frac{1}{NT}E\begin{pmatrix} \dot{\boldsymbol{d}}_\phi \dot{\boldsymbol{d}}_\phi' & * & * & * \\ \dot{d}_{\sigma_\varepsilon^2}\dot{\boldsymbol{d}}_\phi' & \dot{d}_{\sigma_\varepsilon^2}\dot{d}_{\sigma_\varepsilon^2} & * & * \\ \dot{d}_\lambda \dot{\boldsymbol{d}}_\phi' & \dot{d}_\lambda \dot{d}_{\sigma_\varepsilon^2} & \dot{d}_\lambda \dot{d}_\lambda & * \\ \dot{\boldsymbol{d}}_{\beta^\star}\dot{\boldsymbol{d}}_\phi' & \dot{\boldsymbol{d}}_{\beta^\star}\dot{d}_{\sigma_\varepsilon^2} & \dot{\boldsymbol{d}}_{\beta^\star}\dot{d}_\lambda & \dot{\boldsymbol{d}}_{\beta^\star}\dot{\boldsymbol{d}}_{\beta^\star}' \end{pmatrix} - \Sigma_N,\end{aligned}$$

where

$$\dot{\boldsymbol{d}}_\zeta = \begin{pmatrix} \dot{\boldsymbol{d}}_\phi \\ \dot{d}_{\sigma_\varepsilon^2} \\ \dot{d}_\lambda \\ \dot{\boldsymbol{d}}_{\beta^\star} \end{pmatrix} = \begin{pmatrix} (\sigma_{\varepsilon 0}^2)^{-1}\widetilde{Z}'\Omega_{0NT}^{-1}\boldsymbol{V}_{NT} \\ (2\sigma_{\varepsilon 0}^4)^{-1}\boldsymbol{V}_{NT}'\Omega_{0NT}^{-1}\boldsymbol{V}_{NT} - (2\sigma_{\varepsilon 0}^2)^{-1}NT \\ -tr\left[S_{0NT}^{-1}(I_T \otimes W_N)\right] + (\sigma_{\varepsilon 0}^2)^{-1}\boldsymbol{Y}'(I_T \otimes W_N)'\Omega_{0NT}^{-1}\boldsymbol{V}_{NT} \\ (\dot{d}_{\beta_0^\star}, \ldots, \dot{d}_{\beta_p^\star})' \end{pmatrix},$$

$$\ddot{d}_{\phi\phi'} = -\frac{1}{\sigma_{\varepsilon 0}^2}\widetilde{Z}'\Omega_{0NT}^{-1}\widetilde{Z}, \quad \ddot{\boldsymbol{d}}_{\sigma_\varepsilon^2\phi'} = -\frac{1}{\sigma_{\varepsilon 0}^4}\boldsymbol{V}_{NT}'\Omega_{0NT}^{-1}\widetilde{Z},$$

$$\ddot{d}_{\sigma_\varepsilon^2\sigma_\varepsilon^2} = -\frac{1}{\sigma_{\varepsilon 0}^6}\boldsymbol{V}'_{NT}\Omega_{0NT}^{-1}\boldsymbol{V}_{NT} + \frac{NT}{2\sigma_{\varepsilon 0}^4}, \quad \ddot{\boldsymbol{d}}_{\lambda\phi'} = -\frac{1}{\sigma_{\varepsilon 0}^2}\boldsymbol{Y}'(I_T \otimes W_N)'\Omega_{0NT}^{-1}\widetilde{Z},$$

$$\ddot{d}_{\lambda\sigma_\varepsilon^2} = -\frac{1}{\sigma_{\varepsilon 0}^4}\boldsymbol{Y}'(I_T \otimes W_N)'\Omega_{0NT}^{-1}\boldsymbol{V}_{NT},$$

$$\ddot{d}_{\lambda\lambda} = -tr\left[\left(S_{0NT}^{-1}(I_T \otimes W_N)\right)^2\right] - \frac{1}{\sigma_{\varepsilon 0}^2}\boldsymbol{Y}'(I_T \otimes W_N)'\Omega_{0NT}^{-1}(I_T \otimes W_N)\boldsymbol{Y},$$

$$\ddot{d}_{\beta^\star\phi'} = \begin{pmatrix} \ddot{d}_{\beta_0^\star\phi'} \\ \vdots \\ \ddot{d}_{\beta_p^\star\phi'} \end{pmatrix}, \quad \ddot{\boldsymbol{d}}_{\beta^\star\sigma_\varepsilon^2} = \begin{pmatrix} \ddot{\boldsymbol{d}}_{\beta_0^\star\sigma_\varepsilon^2} \\ \vdots \\ \ddot{\boldsymbol{d}}_{\beta_p^\star\sigma_\varepsilon^2} \end{pmatrix}, \quad \ddot{\boldsymbol{d}}_{\beta^\star\lambda} = \begin{pmatrix} \ddot{\boldsymbol{d}}_{\beta_0^\star\lambda} \\ \vdots \\ \ddot{\boldsymbol{d}}_{\beta_p^\star\lambda} \end{pmatrix},$$

$$\ddot{d}_{\beta^\star\beta^{\star\prime}} = \begin{pmatrix} \ddot{d}_{\beta_0^\star\beta_0^\star} & \cdots & \ddot{d}_{\beta_0^\star\beta_p^\star} \\ \vdots & & \\ \ddot{d}_{\beta_p^\star\beta_0^\star} & \cdots & \ddot{d}_{\beta_p^\star\beta_p^\star} \end{pmatrix},$$

with

$$\dot{d}_{\beta_k^\star} = -\frac{1}{2}tr\left[\Omega_{0NT}^{-1}\frac{\partial\Omega_{NT}(\boldsymbol{\beta}_0^\star)}{\partial\beta_k^\star}\right] + \frac{1}{2\sigma_{\varepsilon 0}^2}\boldsymbol{V}'_{NT}\Omega_{0NT}^{-1}\frac{\partial\Omega_{NT}(\boldsymbol{\beta}_0^\star)}{\partial\beta_k^\star}\Omega_{0NT}^{-1}\boldsymbol{V}_{NT},$$

$$\ddot{d}_{\beta_\ell^\star\phi'} = -\frac{1}{\sigma_{\varepsilon 0}^2}\boldsymbol{V}'_{NT}\Omega_{0NT}^{-1}\frac{\partial\Omega_{NT}(\boldsymbol{\beta}_0^\star)}{\partial\beta_\ell^\star}\Omega_{0NT}^{-1}\tilde{Z},$$

$$\ddot{\boldsymbol{d}}_{\beta_\ell^\star\sigma_\varepsilon^2} = -\frac{1}{2\sigma_{\varepsilon 0}^4}\boldsymbol{V}'_{NT}\Omega_{0NT}^{-1}\frac{\partial\Omega_{NT}(\boldsymbol{\beta}_0^\star)}{\partial\beta_\ell^\star}\Omega_{0NT}^{-1}\boldsymbol{V}_{NT},$$

$$\ddot{\boldsymbol{d}}_{\beta_\ell^\star\lambda} = -\frac{1}{\sigma_{\varepsilon 0}^2}\boldsymbol{Y}'(I_T \otimes W_N)'\Omega_{0NT}^{-1}\frac{\partial\Omega_{NT}(\boldsymbol{\beta}_0^\star)}{\partial\beta_\ell^\star}\Omega_{0NT}^{-1}\boldsymbol{V}_{NT},$$

$$\begin{aligned}\ddot{d}_{\beta_\ell^\star\beta_k^\star} = &-\frac{1}{2}tr\left[-\Omega_{0NT}^{-1}\frac{\partial\Omega_{NT}(\boldsymbol{\beta}_0^\star)}{\partial\beta_\ell^\star}\Omega_{0NT}^{-1}\frac{\partial\Omega_{NT}(\boldsymbol{\beta}_0^\star)}{\partial\beta_k^\star} + \Omega_{0NT}^{-1}\frac{\partial^2\Omega_{NT}(\boldsymbol{\beta}_0^\star)}{\partial\beta_\ell^\star\partial\beta_k^\star}\right] \\ &+\frac{1}{2\sigma_{\varepsilon 0}^2}\left[-2\boldsymbol{V}'_{NT}\Omega_{0NT}^{-1}\frac{\partial\Omega_{NT}(\boldsymbol{\beta}_0^\star)}{\partial\beta_\ell^\star}\Omega_{0NT}^{-1}\frac{\partial\Omega_{NT}(\boldsymbol{\beta}_0^\star)}{\partial\beta_k^\star}\Omega_{0NT}^{-1}\boldsymbol{V}_{NT}\right. \\ &\left.+\boldsymbol{V}'_{NT}\Omega_{0NT}^{-1}\frac{\partial^2\Omega_{NT}(\boldsymbol{\beta}_0^\star)}{\partial\beta_\ell^\star\partial\beta_k^\star}\Omega_{0NT}^{-1}\boldsymbol{V}_{NT}\right],\end{aligned}$$

and

$$\frac{\partial\Omega_{NT}(\boldsymbol{\beta}_0^\star)}{\partial\beta_k^\star} = J_T \otimes \mathrm{diag}\left\{2|x_{k1}|\boldsymbol{x}'_{a1}\boldsymbol{\beta}^\star, \ldots, 2|x_{kN}|\boldsymbol{x}'_{aN}\boldsymbol{\beta}^\star\right\},$$

$$\frac{\partial^2 \Omega_{NT}(\boldsymbol{\beta}_0^\star)}{\partial \beta_\ell^\star \partial \beta_k^\star} = J_T \otimes \operatorname{diag}\left\{2|x_{\ell 1}||x_{k1}|, \ldots, 2|x_{\ell N}||x_{kN}|\right\}.$$

S3.2 $\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)$

Recall that $\Omega_{NT}(\boldsymbol{\beta}^\star) = J_T \otimes A_N(\boldsymbol{\beta}^\star) + I_{NT}$ with $A_N(\boldsymbol{\beta}^\star) = \operatorname{diag}\{(\boldsymbol{x}_{a1}'\boldsymbol{\beta}^\star)^2, \ldots, (\boldsymbol{x}_{aN}'\boldsymbol{\beta}^\star)^2\}$. By Lemma 2.1 of Magnus (1982), we have that

$$\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star) = \frac{1}{T} J_T \otimes (I_N + T A_N(\boldsymbol{\beta}^\star))^{-1} + \left(I_T - \frac{1}{T} J_T\right) \otimes I_N. \tag{S3.1}$$

S3.3 $\boldsymbol{Y}_{-1}$

By model (2.1), for $t = 1, \ldots, T$, we have

$$\begin{aligned}
\boldsymbol{y}_t = {} & \alpha_0^t \left(B_{0N}^{-1}\right)^t \boldsymbol{y}_0 + \left(\alpha_0^{t-1} \left(B_{0N}^{-1}\right)^t Z_1 + \alpha_0^{t-2} \left(B_{0N}^{-1}\right)^{t-1} Z_2 + \cdots + B_{0N}^{-1} Z_t\right) \boldsymbol{\gamma}_0 \\
& + \left(\alpha_0^{t-1} \left(B_{0N}^{-1}\right)^t + \alpha_0^{t-2} \left(B_{0N}^{-1}\right)^{t-1} + \cdots + B_{0N}^{-1}\right) \boldsymbol{\theta} \\
& + \left(\alpha_0^{t-1} \left(B_{0N}^{-1}\right)^t \boldsymbol{\varepsilon}_1 + \alpha_0^{t-2} \left(B_{0N}^{-1}\right)^{t-1} \boldsymbol{\varepsilon}_2 + \cdots + B_{0N}^{-1} \boldsymbol{\varepsilon}_t\right),
\end{aligned} \tag{S3.2}$$

and then we have

$$\boldsymbol{Y}_{-1} = (\boldsymbol{y}_0', \ldots, \boldsymbol{y}_{T-1}')' = M_1 \boldsymbol{y}_0 + M_2 Z \boldsymbol{\gamma}_0 + M_3 \boldsymbol{\theta} + M_2 \boldsymbol{\varepsilon}, \tag{S3.3}$$

where

$$M_1 = \begin{pmatrix} I_N \\ \alpha_0 B_{0N}^{-1} \\ \vdots \\ \alpha_0^{T-2} \left(B_{0N}^{-1}\right)^{T-2} \\ \alpha_0^{T-1} \left(B_{0N}^{-1}\right)^{T-1} \end{pmatrix}_{NT \times N,} \qquad M_3 = \begin{pmatrix} 0 \\ B_{0N}^{-1} \\ \vdots \\ \sum_{t=1}^{T-2} \alpha_0^{t-1} \left(B_{0N}^{-1}\right)^t \\ \sum_{t=1}^{T-1} \alpha_0^{t-1} \left(B_{0N}^{-1}\right)^t \end{pmatrix}_{NT \times N,}$$

and

$$M_2 = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ B_{0N}^{-1} & 0 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ \alpha_0^{T-3}\left(B_{0N}^{-1}\right)^{T-2} & \alpha_0^{T-4}\left(B_{0N}^{-1}\right)^{T-3} & \dots & 0 & 0 \\ \alpha_0^{T-2}\left(B_{0N}^{-1}\right)^{T-1} & \alpha_0^{T-3}\left(B_{0N}^{-1}\right)^{T-2} & \dots & B_{0N}^{-1} & 0 \end{pmatrix}_{NT\times NT.}$$

S4 Lemmas

This section includes some lemmas to show Theorems 1–6, and their proofs are relegated at the end of this section. Particularly, Lemmas 1–7 are general results that are useful for all theorems. Throughout this section, n and t are positive integers and nt denotes $n \times t$.

Lemma 1. *Suppose that elements $h_{n,ij}(x)$ of the sequence of $n \times n$ matrices $\{H_n(x)\}$, are uniformly bounded with respect to all i, j, n, and x in X; and $\{G_n(x)\}$ is a sequence of conformable $n \times n$ matrices.*

(i) If $\{G_n(x)\}$ are uniformly bounded in column sums, uniformly on X, then elements $c_{n,ij}(x)$ of $\{H_n(x)G_n(x)\}$ are uniformly bounded with respect to all i, j and n on X.

(ii) If $\{G_n(x)\}$ are uniformly bounded in row sums, uniformly on X, then elements $c_{n,ij}(x)$ of $\{G_n(x)H_n(x)\}$ are uniformly bounded with respect to all i, j and n on X.

For both cases (i) and (ii), it holds that $tr\,[H_n(x)G_n(x)] = tr\,[G_n(x)H_n(x)] = O(n)$, uniformly on X.

Lemma 2. *For $x \in X$, let $\{H_n(x)\}$ and $\{G_n(x)\}$ be sequences of conformable $n\times n$ matrices.*

(i) If $\{H_n(x)\}$ and $\{G_n(x)\}$ are both uniformly bounded in row sums, uniformly on X, then it holds that $\{H_n(x)G_n(x)\}$ are also uniformly bounded in row sums, uniformly on X.

(ii) *If* $\{H_n(x)\}$ *and* $\{G_n(x)\}$ *are both uniformly bounded in column sums, uniformly on* X*, then it holds that* $\{H_n(x)G_n(x)\}$ *are also uniformly bounded in column sums, uniformly on* X*.*

(iii) *If* $\{H_n(x)\}$ *and* $\{G_n(x)\}$ *are both uniformly bounded in both row and column sums, uniformly on* X*, then it holds that* $\{H_n(x)G_n(x)\}$ *are also uniformly bounded in both row and column sums, uniformly on* X*.*

Lemma 3. *Suppose that elements of the sequences of vectors* $\boldsymbol{p}_n = (p_{n,1}, \ldots, p_{n,n})'$ *and* $\boldsymbol{q}_n = (q_{n,1}, \ldots, q_{n,n})'$ *are uniformly bounded. For* $x \in X$*,* $\{H_n(x)\}$ *is a sequence of conformable* $n \times n$ *matrices.*

(i) *If* $\{H_n(x)\}$ *are uniformly bounded in either row or column sums, uniformly in* x *on* X*, then it holds that* $|\boldsymbol{q}_n' H_n(x) \boldsymbol{p}_n| = O(n)$ *uniformly on* X*.*

(ii) *If* $\{H_n(x)\}$ *are uniformly bounded in row sums, uniformly on* X*, and* $\{G_n\}$ *are uniformly bounded in row sums, then it holds that* $\left|\boldsymbol{g}_{i,n} H_n(x) \boldsymbol{p}_n\right| = O(1)$ *uniformly in all* i*, uniformly on* X*, where* $\boldsymbol{g}_{i,n}$ *is the* i*th row of* G_n*.*

Lemma 4. *Suppose that elements of random vectors* $\boldsymbol{u}_n = (u_{n,1}, \ldots, u_{n,n})'$ *and* $\overline{\boldsymbol{u}}_{nt} = (\overline{u}_{nt,1}, \ldots, \overline{u}_{nt,nt})'$ *are i.i.d. with zero mean and finite variance* σ^2*, and elements of the sequence of vector* $\boldsymbol{p}_n = (p_{n,1}, \ldots, p_{n,n})'$ *are uniformly bounded for all* n*. For* $x \in X$*,* $\{H_n(x)\}$ *and* $\left\{\overline{H}_{nt,n}(x)\right\}$ *are sequences of conformable* $n \times n$ *and* $nt \times n$ *matrices, respectively.*

(i) *If* $\{H_n(x)\}$ *are uniformly bounded in row sums, uniformly in* x *on* X*, then it holds that* $\left|\frac{1}{n}\boldsymbol{u}_n' H_n(x) \boldsymbol{p}_n\right| = O_p(1)$ *uniformly on* X*.*

(ii) *If* $\left\{\overline{H}_{nt,n}(x)\right\}$ *are uniformly bounded in row sums, uniformly on* X*, then it holds that* $\left|\frac{1}{nt}\overline{\boldsymbol{u}}_{nt}' \overline{H}_{nt,n}(x) \boldsymbol{p}_n\right| = O_p(1)$ *uniformly on* X*.*

(iii) If $\{H_n(x)\}$ *are uniformly bounded in both row and column sums, uniformly on* X*,* $\left\{\frac{\partial}{\partial x}H_n(x)\right\}$ *are uniformly bounded in row sums, uniformly on* X*, and* X *is a compact set, then it holds that* $E\left(\boldsymbol{u}_n' H_n(x)\boldsymbol{p}_n\right) = 0$*,* $\operatorname{var}\left(\boldsymbol{u}_n' H_n(x)\boldsymbol{p}_n\right) = O(n)$*, and* $\left|\frac{1}{n}\boldsymbol{u}_n' H_n(x)\boldsymbol{p}_n\right| = o_p(1)$ *uniformly on* X*.*

(iv) If $\left\{\overline{H}_{nt,n}(x)\right\}$ *and* $\left\{\frac{\partial}{\partial x}\overline{H}_{nt,n}(x)\right\}$ *are both uniformly bounded in row sums, uniformly on* X*, the column sums of* $\left\{\overline{H}_{nt,n}(x)\right\}$ *are* $O(t)$ *uniformly on* X*, and* X *is a compact set, then it holds that* $E\left(\overline{\boldsymbol{u}}_{nt}'\overline{H}_{nt,n}(x)\boldsymbol{p}_n\right) = 0$*,* $\operatorname{var}\left(\frac{1}{t}\overline{\boldsymbol{u}}_{nt}'\overline{H}_{nt,n}(x)\boldsymbol{p}_n\right) = O(n)$*, and* $\left|\frac{1}{nt}\overline{\boldsymbol{u}}_{nt}'\overline{H}_{nt,n}(x)\boldsymbol{p}_n\right| = o_p(1)$ *uniformly on* X*.*

If elements of $\boldsymbol{u}_n$ *and* $\overline{\boldsymbol{u}}_{nt}$ *are not i.i.d., but* $E(u_{n,i})$ *and* $E(\overline{u}_{nt,i})$ *are uniformly bounded, (i) and (ii) still hold. If elements of* $\boldsymbol{u}_n$ *and* $\overline{\boldsymbol{u}}_{nt}$ *are not i.i.d., but the covariance matrices of* $\boldsymbol{u}_n$ *and* $\overline{\boldsymbol{u}}_{nt}$ *are uniformly bounded in row (column) sums, (iii) and (iv) still hold.*

Lemma 5. *Suppose that elements of random vector* $\boldsymbol{u}_n = (u_{n,1}, \ldots, u_{n,n})'$ *are i.i.d. with finite variance* σ^2*. For* $x \in X$*,* $\{H_n(x)\}$ *is a sequence of conformable* $n \times n$ *matrices.*

(i) If elements of $\{H_n(x)\}$ *are uniformly bounded, uniformly in* x *on* X*, then it holds that* $\left|\boldsymbol{u}_n' H_n(x)\boldsymbol{u}_n\right| = O_p(n)$*,* $E\left(\boldsymbol{u}_n' H_n(x)\boldsymbol{u}_n\right) = O(n)$*, and* $\left|\frac{1}{n}\boldsymbol{u}_n' H_n(x)\boldsymbol{u}_n - \frac{1}{n}E\left(\boldsymbol{u}_n' H_n(x)\boldsymbol{u}_n\right)\right| = O_p(1)$ *uniformly on* X*.*

(ii) If $E\left(u_{n,i}^4\right) < \infty$*,* $\{H_n(x)\}$ *are uniformly bounded in either row or column sums, uniformly on* X*, elements of* $\left\{\frac{\partial}{\partial x}H_n(x)\right\}$ *are uniformly bounded, uniformly on* X*, and* X *is a compact set, then it holds that* $E\left(\boldsymbol{u}_n' H_n(x)\boldsymbol{u}_n\right) = O(n)$*,* $\operatorname{var}\left(\boldsymbol{u}_n' H_n(x)\boldsymbol{u}_n\right) = O(n)$*,* $\left|\boldsymbol{u}_n' H_n(x)\boldsymbol{u}_n\right| = O_p(n)$*, and* $\left|\frac{1}{n}\boldsymbol{u}_n' H_n(x)\boldsymbol{u}_n - \frac{1}{n}E\left(\boldsymbol{u}_n' H_n(x)\boldsymbol{u}_n\right)\right| = o_p(1)$ *uniformly on* X*.*

If elements of $\boldsymbol{u}_n$ *are not i.i.d., but the covariance matrix of* $|\boldsymbol{u}_n|$ *is uniformly bounded in*

row (column) sums, (i) still holds, furthermore, if for any $1 \leq i, j \leq N$*, we also have*

$$\sum_{s=1}^{N}\sum_{t=1}^{N}\left[E\left(\boldsymbol{u}_i\boldsymbol{u}_j\boldsymbol{u}_s\boldsymbol{u}_t\right) - E\left(\boldsymbol{u}_i\boldsymbol{u}_j\right)E\left(\boldsymbol{u}_s\boldsymbol{u}_t\right)\right] = O(1), \tag{S4.4}$$

(ii) still holds, where equation (S4.4) *is to ensure that* $\operatorname{var}\left(\boldsymbol{u}_n' H_n(x)\boldsymbol{u}_n\right) = O(n)$.

Lemma 6. *Suppose that elements of random vectors* $\overline{\boldsymbol{u}}_{nt} = (\overline{u}_{nt,1}, \ldots, \overline{u}_{nt,nt})'$, $\boldsymbol{u}_n = (u_{n,1}, \ldots, u_{n,n})'$ *and* $\boldsymbol{v}_n = (v_{n,1}, \ldots, v_{n,n})'$ *are i.i.d. with finite variance, respectively, and both* $\overline{u}_{nt,i}$ *and* $u_{n,i}$ *are independent with* $v_{n,j}$*. For* $x \in X$*,* $\{H_n(x)\}$ *and* $\left\{\overline{H}_{nt,n}(x)\right\}$ *are sequences of conformable* $n \times n$ *and* $nt \times n$ *matrices, respectively.*

(i) If elements of $\{H_n(x)\}$ *are uniformly bounded, uniformly in* x *on* X*, then it holds that* $\left|\frac{1}{n}\boldsymbol{u}_n' H_n(x)\boldsymbol{v}_n\right| = O_p(1)$ *uniformly on* X*.*

(ii) If elements of $\left\{\overline{H}_{nt,n}(x)\right\}$ *are uniformly bounded, uniformly in* x *on* X*, then it holds that* $\left|\frac{1}{nt}\overline{\boldsymbol{u}}_{nt}'\overline{H}_{nt,n}(x)\boldsymbol{v}_n\right| = O_p(1)$ *uniformly on* X*.*

(iii) If $E\left(u_{n,i}^4\right) < \infty$ *and* $E\left(v_{n,i}^4\right) < \infty$*,* $\{H_n(x)\}$ *are uniformly bounded in either row or column sums, uniformly on* X*, elements of* $\left\{\frac{\partial}{\partial x}H_n(x)\right\}$ *are uniformly bounded, uniformly on* X*, and* X *is a compact set, then it holds that* $E\left(\boldsymbol{u}_n' H_n(x)\boldsymbol{v}_n\right) = 0$*,* $\operatorname{var}\left(\boldsymbol{u}_n' H_n(x)\boldsymbol{v}_n\right) = O(n)$*, and* $\left|\frac{1}{n}\boldsymbol{u}_n' H_n(x)\boldsymbol{v}_n\right| = o_p(1)$ *uniformly on* X*.*

(iv) If $E\left(\overline{u}_{nt,i}^4\right) < \infty$ *and* $E\left(v_{n,i}^4\right) < \infty$*, elements of* $\left\{\overline{H}_{nt,n}(x)\right\}$ *and* $\left\{\frac{\partial}{\partial x}\overline{H}_{nt,n}(x)\right\}$ *are uniformly bounded, uniformly on* X*,* X *is a compact set, and either (a)* $\left\{\overline{H}_{nt,n}(x)\right\}$ *are uniformly bounded in row sums, uniformly on* X*; or (b)* $\left\{\frac{1}{t}\overline{H}_{nt,n}(x)\right\}$ *are uniformly bounded in column sums, uniformly on* X*, then it holds that* $E\left(\overline{\boldsymbol{u}}_{nt}'\overline{H}_{nt,n}(x)\boldsymbol{v}_n\right) = 0$*,* $\operatorname{var}\left(\overline{\boldsymbol{u}}_{nt}'\overline{H}_{nt,n}(x)\boldsymbol{v}_n\right) = O(nt)$*, and* $\left|\frac{1}{nt}\overline{\boldsymbol{u}}_{nt}'\overline{H}_{nt,n}(x)\boldsymbol{v}_n\right| = o_p(1)$ *uniformly on* X*.*

If elements in $\overline{\boldsymbol{u}}_{nt}$*,* $\boldsymbol{u}_n$ *and* $\boldsymbol{v}_n$ *are not i.i.d., (i) and (ii) still hold. Furthermore, denote* $\Sigma_{\overline{u}}$*,* Σ_u *and* Σ_v *as covariance matrices of* $\overline{\boldsymbol{u}}_{nt}$*,* $\boldsymbol{u}_n$ *and* $\boldsymbol{v}_n$*, respectively. If (c)* Σ_u *is uniformly*

bounded in row (column) sums when $\{H_n(x)\}$ *are uniformly bounded in row sums; or (d)* Σ_v *is uniformly bounded in row (column) sums when* $\{H_n(x)\}$ *are uniformly bounded in column sums, (iii) still holds. If (e)* $\Sigma_{\overline{u}}$ *is uniformly bounded in row (column) sums when (a) holds; or (f)* Σ_v *is uniformly bounded in row (column) sums when (b) holds, (iv) still holds.*

Lemma 7. *Suppose that* $\{H_{1,n}\}$ *and* $\{H_{2,nt}\}$ *are sequences of* $n \times n$ *and* $nt \times nt$ *symmetric matrices, respectively,* $\{G_{1,nt,n}\}$, $\{G_{2,nt}\}$ *and* $\{K_{nt,n}\}$ *are* $nt \times n$, $nt \times nt$ *and* $nt \times n$ *matrices, respectively, all of the sequences of aforementioned matrices are uniformly bounded in both row and column sums. Let* $\{\boldsymbol{c}_{1,n}\}$ *and* $\{\boldsymbol{c}_{2,nt}\}$ *be sequences of constant vectors with uniformly bounded elements. Denote* $\boldsymbol{u}_n = (u_1, \ldots, u_n)'$, $\boldsymbol{\varepsilon}_{nt} = (\varepsilon_1, \ldots, \varepsilon_{nt})'$ *and* $\boldsymbol{b}_{nt} = (b_1, \ldots, b_{nt})'$, *for* $1 \le i \le n$, $1 \le j \le nt$, *elements* u_i*'s,* ε_j*'s,* b_j*'s are mutually independent with zero mean, moreover,* $\sup_{1\le i\le n} E|u_i|^{4+\epsilon} < \infty$, $\sup_{1\le j\le nt} E|\varepsilon_j|^{4+\epsilon} < \infty$ *and* $\sup_{1\le j\le nt} E|b_j|^{4+\epsilon} < \infty$ *hold for some* $\epsilon > 0$. *Let* $\mu_Q = E\left(Q_{n(t+1)}\right)$ *and* $\sigma_Q^2 = \mathrm{var}\left(Q_{n(t+1)}\right)$, *where*

$$\begin{aligned} Q_{n(t+1)} = {} & \boldsymbol{c}_{1,n}'\boldsymbol{u}_n + \boldsymbol{b}_{nt}'G_{1,nt,n}\boldsymbol{u}_n + \boldsymbol{u}_n'H_{1,n}\boldsymbol{u}_n \\ & + \boldsymbol{c}_{2,nt}'\boldsymbol{\varepsilon}_{nt} + \boldsymbol{b}_{nt}'G_{2,nt}\boldsymbol{\varepsilon}_{nt} + \boldsymbol{\varepsilon}_{nt}'H_{2,nt}\boldsymbol{\varepsilon}_{nt} + \boldsymbol{\varepsilon}_{nt}'K_{nt,n}\boldsymbol{u}_n\,. \end{aligned} \tag{S4.5}$$

Suppose that t *is a nondecreasing function of* n, *then as* $n \to \infty$, *we have*

$$\frac{Q_{n(t+1)} - \mu_Q}{\sigma_Q} \to_d N(0,1). \tag{S4.6}$$

Notice that $\boldsymbol{u}_n'H_{1,n}\boldsymbol{u}_n = \boldsymbol{u}_n'\frac{H_{1,n}+H_{1,n}'}{2}\boldsymbol{u}_n$ *and* $\boldsymbol{\varepsilon}_{nt}'H_{2,nt}\boldsymbol{\varepsilon}_{nt} = \boldsymbol{\varepsilon}_{nt}'\frac{H_{2,nt}+H_{2,nt}'}{2}\boldsymbol{\varepsilon}_{nt}$, *hence the symmetry assumption on* $\{H_{1,n}\}$ *and* $\{H_{2,nt}\}$ *can be relaxed. Moreover, if elements in* $\boldsymbol{b}_{nt}$ *are not mutually independent, but the covariance matrix of* $\boldsymbol{b}_{nt}$ *is uniformly bounded in row (column) sums, then* (S4.6) *still holds.*

Under Assumptions 3.1, 3.4–3.6, Lemmas 8 and 9 hold for our proposed model, and these two lemmas are mainly used in the proof of Theorems 1–2.

Lemma 8. $E\left(\widetilde{Z}'\Omega_{0NT}^{-1}\boldsymbol{V}_{NT}\right)=0.$

Lemma 9. *Denote $\Delta(\boldsymbol{\beta}^\star)$ as the parameter space of $\boldsymbol{\beta}^\star$. For $\Omega_{NT}(\boldsymbol{\beta}^\star)$, the following results hold:*

(i) $\{T^{-1}\Omega_{NT}(\boldsymbol{\beta}^\star)\}$ are uniformly bounded in both row and column sums, uniformly in $\Delta(\boldsymbol{\beta}^\star)$, and thus $\{T^{-1}\Omega_{0NT}\}$ are uniformly bounded in both row and column sums.

(ii) $\left\{T\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)\right\}$ are uniformly bounded in both row and column sums, uniformly in $\Delta(\boldsymbol{\beta}^\star)$, and thus $\left\{T\Omega_{0NT}^{-1}\right\}$ are uniformly bounded in both row and column sums.

(iii) $\left\{T^{-1}\frac{\partial}{\partial\beta_\ell^\star}\Omega_{NT}(\boldsymbol{\beta}^\star)\right\}$ are uniformly bounded in both row and column sums, uniformly in $\Delta(\boldsymbol{\beta}^\star)$, for all $\ell=0,1,\ldots,p$.

Under Assumptions 3.3–3.9, Lemma 10 holds for our proposed model, and this lemma is mainly used in the proof of Theorem 5.

Lemma 10. *Denote $\Psi_\tau(u)=\tau-I(u<0)$ and $D_0=\lim\limits_{N\to\infty}\frac{1}{N}\sum\limits_{i=1}^{N}\frac{\boldsymbol{x}_{ai}\boldsymbol{x}_{ai}'}{(\boldsymbol{x}_{ai}'\boldsymbol{\beta}_0)^2}$.*

(i) $\mathbb{G}_N\left(\frac{\boldsymbol{x}_{ai}}{\boldsymbol{x}_{ai}'\boldsymbol{\beta}_0}\Psi_\bullet(\eta_i-Q_\eta(\cdot))\right)\rightsquigarrow\mathbb{G}(\cdot)$ in $\ell^\infty(\mathcal{T})$, where $\mathbb{G}$ is a Gaussian process with zero mean and covariance function $E\left[\mathbb{G}(\tau)\mathbb{G}'(\tau')\right]=(\min\{\tau,\tau'\}-\tau\tau')D_0$, and D_0 is assumed to be well-defined.

(ii) If $\sup\limits_{\tau\in\mathcal{T}}\|\widehat{\boldsymbol{\varphi}}(\tau)-\boldsymbol{\varphi}_0(\tau)\|=o_p(1)$ and $\max\limits_{1\le i\le N}|\widehat{r}_i|=o_p(1)$, then it holds that
$$\sup_{\tau\in\mathcal{T}}\left\|\mathbb{G}_N\left(\frac{\boldsymbol{x}_{ai}}{\boldsymbol{x}_{ai}'\boldsymbol{\beta}_0}\Psi_\tau(\vartheta_i-\boldsymbol{x}_{ai}'\widehat{\boldsymbol{\varphi}}(\tau)+\widehat{r}_i)\right)-\mathbb{G}_N\left(\frac{\boldsymbol{x}_{ai}}{\boldsymbol{x}_{ai}'\boldsymbol{\beta}_0}\Psi_\tau(\vartheta_i-\boldsymbol{x}_{ai}'\boldsymbol{\varphi}_0(\tau))\right)\right\|=o_p(1).$$

Proof of Lemma 1. This lemma is an extension of Lemma A.8 of Lee (2004).

We first show (i). There exist constants h_0 and a_0 such that $\sup\limits_x|h_{n,ij}(x)|\le h_0$ and $\sup\limits_x\sum\limits_{k=1}^{n}|g_{n,kj}(x)|\le a_0$ for all i, j and n, where $G_n(x)=[g_{n,ij}(x)]$. It follows that $\sup\limits_x|c_{n,ij}(x)|=\sup\limits_x|\sum\limits_{k=1}^{n}h_{n,ik}g_{n,kj}(x)|\le h_0a_0$ for all i, j and n, that is elements of $\{H_n(x)G_n(x)\}$ are uniformly bounded, uniformly on X.

Similarly, we can show (ii) holds, and $tr\left[H_n(x)G_n(x)\right] = tr\left[G_n(x)H_n(x)\right] = O(n)$ uniformly on X holds. □

Proof of Lemma 2. This lemma is an extension of Lemma B.1 of Su and Yang (2015).

We first show that (i) holds. Let $H_n(x) = [h_{n,ij}(x)]$ and $G_n(x) = [g_{n,ij}(x)]$, there exist constants a_1 and b_1, such that $\sup_x \sum_{j=1}^n |h_{n,ij}(x)| \leq a_1$ and $\sup_x \sum_{j=1}^n |g_{n,ij}(x)| \leq b_1$ for all i and n. Then, let $H_n(x)G_n(x) = [c_{n,ij}(x)]$, for all i and n it holds that

$$\sup_x \sum_{j=1}^n |c_{n,ij}(x)| = \sup_x \sum_{j=1}^n \left|\sum_{k=1}^n h_{n,ik}(x)g_{n,kj}(x)\right| \leq \sup_x \sum_{k=1}^n |h_{n,ik}(x)| \sum_{j=1}^n |g_{n,kj}(x)| \leq a_1 b_1.$$

As a result, $\{H_n(x)G_n(x)\}$ are uniformly bounded in row sums, uniformly on X.

Similarly, we can prove (ii), and then (iii) follows with (i) and (ii). □

Proof of Lemma 3. This lemma is an extension of Lemma A.6 of Lee (2004).

There exists constants c_1 and c_2 such that $|p_{n,i}| \leq c_1$ and $|q_{n,i}| \leq c_2$ for all i and n.

We first show (i). Let $H_n(x) = [h_{n,ij}(x)]$. There exists a constant a_1 such that $\sup_x \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n |h_{n,ij}(x)| \leq a_1$. It then follows that

$$\sup_x |\boldsymbol{q}_n' H_n(x) \boldsymbol{p}_n| = \sup_x \left|\sum_{i=1}^n \sum_{j=1}^n q_{n,i} h_{n,ij}(x) p_{n,j}\right| \leq n a_1 c_1 c_2.$$

Hence, we obtain that $|\boldsymbol{q}_n' H_n(x) \boldsymbol{p}_n| = O(n)$ uniformly on X.

Next we show (ii). Let $G_n = [g_{n,ij}]$. There exist constants b_1 and b_2 such that $\sup_x \sum_{j=1}^n |h_{n,ij}(x)| \leq b_1$ and $\sum_{j=1}^n |g_{n,ij}| \leq b_2$ for all i and n. Then $\sup_x \sum_{j=1}^n |h_{n,ij}(x) p_{n,j}| \leq b_1 c_1$ for all i and n, so it follows that

$$\sup_x |\boldsymbol{g}_{i,n} H_n(x) \boldsymbol{p}_n| = \sup_x \left|\sum_{k=1}^n \sum_{j=1}^n g_{n,ik} h_{n,kj}(x) p_{n,j}\right| \leq \sum_{k=1}^n |g_{n,ik}| \sup_x \sum_{j=1}^n |h_{n,kj}(x) p_{n,j}| \leq b_2 b_1 c_1,$$

for all i. Thus $\left|\boldsymbol{g}_{i,n} H_n(x) \boldsymbol{p}_n\right| = O(1)$ holds uniformly in all i, uniformly on X. □

Proof of Lemma 4. There exist constants c_1 and c_2 such that $|p_{n,i}| \leq c_1$ and $|\overline{p}_{nt,i}| \leq c_2$ for all i, n and t. Let $H_n(x) = [h_{n,ij}(x)]$ and $\overline{H}_{n,nt}(x) = \left[\overline{h}_{n,nt,ij}(x)\right]$.

To show (i), it suffices to show that $\sup_n E\left(\sup_x \left|\frac{1}{n}\boldsymbol{u}_n' H_n(x)\boldsymbol{p}_n\right|\right) < \infty$ by page 339 of Davidson (1994). There exists a constant a_1 such that $\sup_x \sum_{j=1}^n |h_{n,ij}(x)| \leq a_1$ for all i and n. Then it holds that

$$\begin{aligned}
\sup_n E\left(\sup_x \left|\frac{1}{n}\boldsymbol{u}_n' H_n(x)\boldsymbol{p}_n\right|\right) &= \sup_n E\left(\sup_x \left|\frac{1}{n}\sum_{i=1}^n\sum_{j=1}^n u_{n,i}h_{n,ij}(x)p_{n,j}\right|\right) \\
&\leq \sup_n E\left(\frac{1}{n}\sum_{i=1}^n |u_{n,i}| \sup_x \sum_{j=1}^n |h_{n,ij}(x)||p_{n,j}|\right) \\
&\leq \sup_n E\left(\frac{a_1c_1}{n}\sum_{i=1}^n |u_{n,i}|\right) = a_1c_1E|u_{n,j}| < \infty.
\end{aligned}$$

Therefore $\left|\frac{1}{n}\boldsymbol{u}_n' H_n(x)\boldsymbol{p}_n\right| = O_p(1)$ uniformly on X. Similarly, we can show (ii) holds.

Next we show (iii). It is obvious that $E\left(\boldsymbol{u}_n' H_n(x)\boldsymbol{p}_n\right) = 0$. Notice that $\{H_n'(x)H_n(x)\}$ are uniformly bounded in both row and column sums, uniformly on X by Lemma 2(iii), then by Lemma 3(i), we have that

$$\operatorname{var}\left(\boldsymbol{u}_n' H_n(x)\boldsymbol{p}_n\right) = E\left(\boldsymbol{u}_n' H_n(x)\boldsymbol{p}_n\right)^2 = \sigma^2\boldsymbol{p}_n' H_n'(x)H_n(x)\boldsymbol{p}_n = O(n),$$

which holds uniformly on X. It then follows that $\left|\frac{1}{n}\boldsymbol{u}_n' H_n(x)\boldsymbol{p}_n\right| = o_p(1)$ for all $x \in X$ by the Chebyshev's inequality, which verifies the pointwise convergence. Furthermore, for x_1, $x_2 \in X$, by the mean value theorem, it holds that

$$\begin{aligned}
\left|\frac{1}{n}\boldsymbol{u}_n' H_n(x_1)\boldsymbol{p}_n - \frac{1}{n}\boldsymbol{u}_n' H_n(x_2)\boldsymbol{p}_n\right| &= \left|\frac{1}{n}\boldsymbol{u}_n' \frac{\partial}{\partial x}H_n(\bar{x})\boldsymbol{p}_n\right| |x_1 - x_2| \\
&\leq \sup_x \left|\frac{1}{n}\boldsymbol{u}_n' \frac{\partial}{\partial x}H_n(x)\boldsymbol{p}_n\right| |x_1 - x_2|,
\end{aligned}$$

where $\bar{x}$ lies between x_1 and x_2. This together with $\sup_x \left|\frac{1}{n}\boldsymbol{u}_n' \frac{\partial}{\partial x}H_n(x)\boldsymbol{p}_n\right| = O_p(1)$ by (i), implies that $\frac{1}{n}\boldsymbol{u}_n' H_n(x)\boldsymbol{p}_n$ is stochastically equicontinuous by Theorem 21.10 of Davidson (1994). The result that $\left|\frac{1}{n}\boldsymbol{u}_n' H_n(x)\boldsymbol{p}_n\right| = o_p(1)$ holds uniformly on X, by Lemma 2.8 of Newey and McFadden (1994).

Finally, we show that (iv) holds. Notice that the sequence $\left\{\frac{1}{t}\overline{H}'_{nt,n}(x)\overline{H}_{nt,n}(x)\right\}$ are uniformly bounded in row sums, uniformly on X by Lemma 2(i), it then follows that $\text{var}\left(\frac{1}{t}\overline{\boldsymbol{u}}'_{nt}\overline{H}_{nt,n}(x)\boldsymbol{p}_n\right) = O(n)$ holds uniformly on X by Lemma 3(i) and $\sup_x \left|\frac{1}{nt}\overline{\boldsymbol{u}}'_{nt}\frac{\partial}{\partial x}\overline{H}_{nt,n}(x)\boldsymbol{p}_n\right| = O_p(1)$ by (ii). As a result, we can prove (iv) with the similar argument as that of (iii). □

Proof of Lemma 5. We first show (i). Let $H_n(x) = [h_{n,ij}(x)]$. There exists a constant c_1 such that $\sup_x |h_{n,ij}(x)| \leq c_1$ for all i, j and n. On one hand, it can be shown that

$$\begin{aligned}\sup_n E\left(\sup_x \left|\frac{1}{n}\boldsymbol{u}'_n H_n(x)\boldsymbol{u}_n\right|\right) &= \sup_n E\left(\sup_x \left|\frac{1}{n}\sum_{i=1}^n\sum_{j=1}^n u_{n,i}h_{n,ij}(x)u_{n,j}\right|\right)\\ &\leq \sup_n \frac{c_1}{n}\sum_{i=1}^n\sum_{j=1}^n E|u_{n,i}||u_{n,j}|\\ &= \sup_n \frac{c_1}{n}\sum_{i=1}^n Eu_{n,i}^2 = c_1\sigma^2 < \infty,\end{aligned}$$

so we have $\sup_x \left|\frac{1}{n}\boldsymbol{u}'_n H_n(x)\boldsymbol{u}_n\right| = O_p(1)$ by the argument of page 339 in Davidson (1994), that is $|\boldsymbol{u}'_n H_n(x)\boldsymbol{u}_n| = O_p(n)$ uniformly on X. On the other hand, we have

$$\sup_x \left|E\left(\frac{1}{n}\boldsymbol{u}'_n H_n(x)\boldsymbol{u}_n\right)\right| = \sup_x \left|E\left(\frac{1}{n}\sum_{i=1}^n\sum_{j=1}^n u_{n,i}h_{n,ij}(x)u_{n,j}\right)\right| \leq c_1\sigma^2 < \infty,$$

which implies $E\left(\boldsymbol{u}'_n H_n(x)\boldsymbol{u}_n\right) = O(n)$ uniformly on X. As a result, we have that $\left|\frac{1}{n}\boldsymbol{u}'_n H_n(x)\boldsymbol{u}_n - \frac{1}{n}E\left(\boldsymbol{u}'_n H_n(x)\boldsymbol{u}_n\right)\right| = O_p(1)$, uniformly on X.

We next show (ii) holds. $\left|\frac{1}{n}\boldsymbol{u}'_n H_n(x)\boldsymbol{u}_n - \frac{1}{n}E\left(\boldsymbol{u}'_n H_n(x)\boldsymbol{u}_n\right)\right| = o_p(1)$ holds for all $x \in X$ by Lemma A.12 of Lee (2004). For x_1, $x_2 \in X$, by the mean value theorem, it holds that

$$\begin{aligned}&\left|\left[\frac{1}{n}\boldsymbol{u}'_n H_n(x_1)\boldsymbol{u}_n - \frac{1}{n}E\left(\boldsymbol{u}'_n H_n(x_1)\boldsymbol{u}_n\right)\right] - \left[\frac{1}{n}\boldsymbol{u}'_n H_n(x_2)\boldsymbol{u}_n - \frac{1}{n}E\left(\boldsymbol{u}'_n H_n(x_2)\boldsymbol{u}_n\right)\right]\right|\\ &\leq \sup_x \left|\frac{1}{n}\boldsymbol{u}'_n\frac{\partial}{\partial x}H_n(x)\boldsymbol{u}_n - \frac{1}{n}E\left(\boldsymbol{u}'_n\frac{\partial}{\partial x}H_n(x)\boldsymbol{u}_n\right)\right| |x_1 - x_2|,\end{aligned}$$

where $\bar{x}$ lies between x_1 and x_2, and $\sup_x \left|\frac{1}{n}\boldsymbol{u}'_n\frac{\partial}{\partial x}H_n(x)\boldsymbol{u}_n - \frac{1}{n}E\left(\boldsymbol{u}'_n\frac{\partial}{\partial x}H_n(x)\boldsymbol{u}_n\right)\right| = O_p(1)$ by (i), then $\frac{1}{n}\boldsymbol{u}'_n H_n(x)\boldsymbol{u}_n - \frac{1}{n}E\left(\boldsymbol{u}'_n H_n(x)\boldsymbol{u}_n\right)$ is stochastically equicontinuous by Theorem

21.10 of Davidson (1994). This together with $\left|\frac{1}{n}\boldsymbol{u}_n' H_n(x)\boldsymbol{u}_n - \frac{1}{n}E\left(\boldsymbol{u}_n' H_n(x)\boldsymbol{u}_n\right)\right| = o_p(1)$ for all $x \in X$, implies that $\left|\frac{1}{n}\boldsymbol{u}_n' H_n(x)\boldsymbol{u}_n - \frac{1}{n}E\left(\boldsymbol{u}_n' H_n(x)\boldsymbol{u}_n\right)\right| = o_p(1)$ holds uniformly on X by Lemma 2.8 of Newey and McFadden (1994). □

Proof of Lemma 6. (i) and (ii) follows from that $\sup_n E\left(\sup_x \left|\frac{1}{n}\boldsymbol{u}_n' H_n(x)\boldsymbol{v}_n\right|\right) = 0$ and that $\sup_n E\left(\sup_x \left|\frac{1}{nt}\overline{\boldsymbol{u}}_{nt}' \overline{H}_{nt,n}(x)\boldsymbol{v}_n\right|\right) = 0$.

We next show (iii) holds. Let $\sigma_u^2 = \mathrm{var}(u_i)$ and $\sigma_v^2 = \mathrm{var}(v_i)$. $E\left(\boldsymbol{u}_n' H_n(x)\boldsymbol{v}_n\right) = 0$, and it can be shown that

$$\mathrm{var}\left(\boldsymbol{u}_n' H_n(x)\boldsymbol{v}_n\right) = \sigma_u^2\sigma_v^2 tr\left[H_n(x)H_n'(x)\right] = O(n),$$

by Lemma 1. Then $\left|\frac{1}{n}\boldsymbol{u}_n' H_n(x)\boldsymbol{v}_n\right| = o_p(1)$ holds for all $x \in X$ by Chebyshev's inequality. This together with the stochastical equicontinuity of $\frac{1}{n}\boldsymbol{u}_n' H_n(x)\boldsymbol{v}_n$ by similar argument of the proof of Lemma 5 (ii), implies that $\left|\frac{1}{n}\boldsymbol{u}_n' H_n(x)\boldsymbol{v}_n\right| = o_p(1)$ holds uniformly on X by Lemma 2.8 of Newey and McFadden (1994).

The proof of (iv) is analogous to that of (iii). □

Proof of Lemma 7. Let $\boldsymbol{v}_{n(t+1)} = (\boldsymbol{u}_n', \boldsymbol{\varepsilon}_{nt}')' = \{v_i, 1 \le i \le n(t+1)\}$, $\boldsymbol{c}_{n(t+1)} = (\boldsymbol{c}_{1,n}', \boldsymbol{c}_{2,nt}')' = \{c_i, 1 \le i \le n(t+1)\}$, $\sigma_{v_i}^2 = \mathrm{var}(v_i)$, and

$$G_{nt,n(t+1)} = (G_{1,nt,n}, G_{2,nt}) = [g_{ij}, 1 \le i \le nt, 1 \le j \le n(t+1)],$$

$$H_{n(t+1)} = \begin{pmatrix} H_{1,n} & \frac{1}{2}K_{nt,n}' \\ \frac{1}{2}K_{nt,n} & H_{2,nt} \end{pmatrix} = [h_{ij}, 1 \le i \le n(t+1), 1 \le j \le n(t+1)].$$

Note that $E(\boldsymbol{v}_{n(t+1)}) = \boldsymbol{0}$ and v_i's are mutually independent, so it holds that

$$\begin{aligned} Q_{n(t+1)} - \mu_Q = {} & \boldsymbol{c}_{n(t+1)}'\boldsymbol{v}_{n(t+1)} + \boldsymbol{b}_{nt}G_{nt,n(t+1)}\boldsymbol{v}_{n(t+1)} \\ & + \left[\boldsymbol{v}_{n(t+1)}' H_{n(t+1)}\boldsymbol{v}_{n(t+1)} - E\left(\boldsymbol{v}_{n(t+1)}' H_{n(t+1)}\boldsymbol{v}_{n(t+1)}\right)\right] \end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{n(t+1)} c_i v_i + \sum_{i=1}^{n(t+1)} \sum_{j=1}^{nt} g_{ji} b_j v_i + \sum_{i=1}^{n(t+1)} \sum_{j=1}^{n(t+1)} h_{ij} v_i v_j - \sum_{i=1}^{n(t+1)} h_{ii} \sigma_{v_i}^2 \\
&= \sum_{i=1}^{n(t+1)} \left[h_{ii} \left(v_i^2 - \sigma_{v_i}^2 \right) + 2 v_i \sum_{j=1}^{i-1} h_{ij} v_j + v_i \sum_{j=1}^{nt} g_{ji} b_j + c_i v_i \right] \\
&\triangleq \sum_{i=1}^{n(t+1)} Y_{i,n(t+1)}^{\star}.
\end{aligned}
$$

Define the σ-field $\mathcal{F}_{i,n(t+1)}$ generated by $\{\boldsymbol{b}_{nt}, v_1, \ldots, v_i\}$ for $1 \le i \le n(t+1)$. For each i, $\mathcal{F}_{i-1,n(t+1)} \subset \mathcal{F}_{i,n(t+1)}$, $Y_{i,n(t+1)}^{\star}$ is known given $\mathcal{F}_{i,n(t+1)}$ and $E\left(Y_{i,n(t+1)}^{\star} | \mathcal{F}_{i-1,n(t+1)}\right) = 0$, so $\{Y_{i,n(t+1)}^{\star}, \mathcal{F}_{i,n(t+1)}, 1 \le i \le n(t+1), n, t \ge 1\}$ forms a martingale difference array. Moreover, since $Y_{i,n(t+1)}^{\star}$'s are martingale differences, it holds that

$$
\sigma_Q^2 = E\left(\sum_{i=1}^{n(t+1)} Y_{i,n(t+1)}^{\star} \right)^2 = \sum_{i=1}^{n(t+1)} E\left(Y_{i,n(t+1)}^{\star 2} \right).
$$

Let $X_{i,n(t+1)}^{\star} = \sigma_Q^{-1} Y_{i,n(t+1)}^{\star}$. We have that $\{X_{i,n(t+1)}^{\star}, \mathcal{F}_{i,n(t+1)}, 1 \le i \le n(t+1), n, t \ge 1\}$ also form a martingale difference array.

To prove that $\frac{Q_{n(t+1)} - \mu_Q}{\sigma_Q} = \sum_{i=1}^{n(t+1)} X_{i,n(t+1)}^{\star} \to_d N(0,1)$, by Theorem A.1 of Kelejian and Prucha (2001), it suffices to show

a) $\sum_{i=1}^{n(t+1)} E\left[E\left(\left| X_{i,n(t+1)}^{\star} \right|^{2+\delta} \middle| \mathcal{F}_{i-1,n(t+1)} \right) \right] \to 0$ for some $\delta > 0$;

b) $\sum_{i=1}^{n(t+1)} E\left(X_{i,n(t+1)}^{\star 2} | \mathcal{F}_{i-1,n(t+1)} \right) \to_p 1$.

We first show a). Take $0 < \delta \le \epsilon/2$, for any arbitrary $1 \le i, j \le n(t+1)$, $1 \le \ell \le nt$ and $s \le 2 + \delta$, there exist constants k_e and k_p such that

$$
\sigma_{v_i}^2 \le k_e,\ E|v_i|^s \le k_e,\ E|v_i|^s E|v_j|^s \le k_e,\ E|v_i|^s E|b_\ell|^s \le k_e,\ E\left| v_i^2 - \sigma_{v_i}^2 \right|^s \le k_e, \text{ and}
$$

$$
\sum_{i=1}^{n(t+1)} |h_{ij}| \le k_p,\ \sum_{j=1}^{n(t+1)} |h_{ij}| \le k_p,\ \sum_{i=1}^{nt} |g_{ij}| \le k_p,\ \sum_{j=1}^{n(t+1)} |g_{ij}| \le k_p,\ |c_i| < k_p.
$$

Let $q = 2+\delta$ and $\frac{1}{q}+\frac{1}{p}=1$. By Minkowski inequality and Hölder inequality, we obtain that

$$\begin{aligned}
\left|Y^{\star}_{i,n(t+1)}\right|^q &= \left|h_{ii}\left(v_i^2-\sigma^2_{v_i}\right)+2v_i\sum_{j=1}^{i-1}h_{ij}v_j+c_iv_i+v_i\sum_{j=1}^{nt}g_{ji}b_j\right|^q \\
&\le 2^q\left|h_{ii}\left(v_i^2-\sigma^2_{v_i}\right)+2v_i\sum_{j=1}^{i-1}h_{ij}v_j\right|^q+2^{2q}\left|c_iv_i\right|^q+2^{2q}\left|v_i\sum_{j=1}^{nt}g_{ji}b_j\right|^q \\
&\le 2^q\left(|h_{ii}|^{\frac{1}{p}}|h_{ii}|^{\frac{1}{q}}\left|v_i^2-\sigma^2_{v_i}\right|+\sum_{j=1}^{i-1}|h_{ij}|^{\frac{1}{p}}|h_{ij}|^{\frac{1}{q}}|2v_iv_j|\right)^q+2^{2q}|c_i|^q|v_i|^q \\
&\quad +2^{2q}|v_i|^q\left(\sum_{j=1}^{nt}|g_{ji}|^{\frac{1}{p}}|g_{ji}|^{\frac{1}{q}}|b_j|\right)^q \\
&\le 2^q\left[\left(\sum_{j=1}^{i}|h_{ij}|\right)^{\frac{1}{p}}\left(|h_{ii}|\left|v_i^2-\sigma^2_{v_i}\right|^q+\sum_{j=1}^{i-1}|h_{ij}||2v_iv_j|^q\right)^{\frac{1}{q}}\right]^q \\
&\quad +2^{2q}|c_i|^q|v_i|^q+2^{2q}|v_i|^q\left[\left(\sum_{j=1}^{nt}|g_{ji}|\right)^{\frac{1}{p}}\left(\sum_{j=1}^{nt}|g_{ji}||b_j|^q\right)^{\frac{1}{q}}\right]^q \\
&= 2^q\left(\sum_{j=1}^{i}|h_{ij}|\right)^{\frac{q}{p}}\left(|h_{ii}|\left|v_i^2-\sigma^2_{v_i}\right|^q+2^q\sum_{j=1}^{i-1}|h_{ij}||v_i|^q|v_j|^q\right) \\
&\quad +2^{2q}|c_i|^q|v_i|^q+2^{2q}|v_i|^q\left(\sum_{j=1}^{nt}|g_{ji}|\right)^{\frac{q}{p}}\sum_{j=1}^{nt}|g_{ji}||b_j|^q.
\end{aligned}$$

Then it follows that

$$\begin{aligned}
&\sum_{i=1}^{n(t+1)}E\left[E\left(\left|Y^{\star}_{i,n(t+1)}\right|^{2+\delta}\middle|\mathcal{F}_{i-1,n(t+1)}\right)\right] \\
&\le 2^q\sum_{i=1}^{n(t+1)}\left(\sum_{j=1}^{i}|h_{ij}|\right)^{\frac{q}{p}}\left(|h_{ii}|E\left|v_i^2-\sigma^2_{v_i}\right|^q+2^q\sum_{j=1}^{i-1}|h_{ij}|E|v_i|^qE|v_j|^q\right) \\
&\quad +2^{2q}\sum_{i=1}^{n(t+1)}|c_i|^qE|v_i|^q+2^{2q}\sum_{i=1}^{n(t+1)}E|v_i|^q\left(\sum_{j=1}^{nt}|g_{ji}|\right)^{\frac{q}{p}}\sum_{j=1}^{nt}|g_{ji}|E|b_j|^q \\
&\le 2^{2q}\left(n(t+1)k_ek_p^{\frac{q}{p}+1}+n(t+1)k_ek_p^q+n(t+1)k_ek_p^{\frac{q}{p}+1}\right)=cn(t+1),
\end{aligned}\tag{S4.7}$$

where c is a constant. Moreover, by algebra calculation, we obtain that

$$\sigma_Q^2=\sum_{i=1}^{n(t+1)}\left[h_{ii}^2(\mu^{(4)}_{v_i}-\sigma^4_{v_i})+4\sum_{j=1}^{i-1}h_{ij}^2\sigma^2_{v_i}\sigma^2_{v_j}+c_i^2\sigma^2_{v_i}+2c_ih_{ii}\mu^{(3)}_{v_i}+\sum_{j=1}^{nt}g_{ji}^2\sigma^2_{b_j}\sigma^2_{v_i}\right]$$

$$= O(n(t+1)), \tag{S4.8}$$

where $\mu_{v_i}^{(3)} = E(v_i^3)$, $\mu_{v_i}^{(4)} = E(v_i^4)$ and $\sigma_{b_j}^2 = \mathrm{var}(b_j)$. Then by (S4.7) and (S4.8), it holds that

$$\begin{aligned}
&\sum_{i=1}^{n(t+1)} E\left[E\left(\left|X_{i,n(t+1)}^{\star}\right|^{2+\delta}\Big|\mathcal{F}_{i-1,n(t+1)}\right)\right] \\
=&\frac{1}{\sigma_Q^{2+\delta}}\sum_{i=1}^{n(t+1)} E\left[E\left(\left|Y_{i,n(t+1)}^{\star}\right|^{2+\delta}\Big|\mathcal{F}_{i-1,n(t+1)}\right)\right] \le \left(\frac{n(t+1)}{\sigma_Q^2}\right)^{1+\frac{\delta}{2}} \frac{c}{(n(t+1))^{\frac{\delta}{2}}} \to 0,
\end{aligned}$$

as $n \to \infty$. Hence a) holds.

It remains to show that b) holds. It can be verified that

$$\begin{aligned}
&\sum_{i=1}^{n(t+1)} E\left(X_{i,n(t+1)}^{\star 2}|\mathcal{F}_{i-1,n(t+1)}\right) - 1 \\
=& \frac{1}{\sigma_Q^2}\sum_{i=1}^{n(t+1)} E\left(Y_{i,n(t+1)}^{\star 2}|\mathcal{F}_{i-1,n(t+1)}\right) - \frac{1}{\sigma_Q^2}\sum_{i=1}^{n(t+1)} E\left(Y_{i,n(t+1)}^{\star 2}\right) \\
=& \frac{1}{\sigma_Q^2}\sum_{i=1}^{n(t+1)} \left[E\left(Y_{i,n(t+1)}^{\star 2}|\mathcal{F}_{i-1,n(t+1)}\right) - E\left(Y_{i,n(t+1)}^{\star 2}\right)\right] \\
=& \frac{n(t+1)}{\sigma_Q^2}\frac{1}{n(t+1)}\sum_{i=1}^{n(t+1)}\Bigg\{8\sum_{j=1}^{i-1}\sum_{k=1}^{j-1} h_{ij}h_{ik}\sigma_{v_i}^2 v_j v_k + 4\sum_{j=1}^{i-1} h_{ij}^2\sigma_{v_i}^2(v_j^2 - \sigma_{v_j}^2) \\
&+2\sum_{j=1}^{nt}\sum_{k=1}^{j-1} g_{ji}h_{ki}\sigma_{v_i}^2 b_j b_k + \sum_{j=1}^{nt} g_{ji}^2\sigma_{v_i}^2(b_j^2 - \sigma_{b_j}^2) + 4\sum_{j=1}^{i-1}\sum_{k=1}^{nt} h_{ij}g_{ki}\sigma_{v_i}^2 v_j b_k \\
&+ 4\sum_{j=1}^{i-1}\left(h_{ii}\mu_{v_i}^{(3)} + c_i\sigma_{v_i}^2\right) h_{ij}v_j + 2\sum_{j=1}^{nt}\left(h_{ii}\mu_{v_i}^{(3)} + c_i\sigma_{v_i}^2\right) g_{ji}b_j\Bigg\}.
\end{aligned}$$

Note that $\frac{n(t+1)}{\sigma_Q^2} = O(1)$. It suffices to show that the rest of the last equality is $o_p(1)$. We next show that b1) $s_1 \triangleq \frac{1}{n(t+1)}\sum_{i=1}^{n(t+1)}\sum_{j=1}^{i-1}\sum_{k=1}^{j-1} h_{ij}h_{ik}\sigma_{v_i}^2 v_j v_k = o_p(1)$ and b2) $s_2 \triangleq \frac{1}{n(t+1)}\sum_{i=1}^{n(t+1)}\sum_{j=1}^{i-1} h_{ij}^2\sigma_{v_i}^2(v_j^2 - \sigma_{v_j}^2) = o_p(1)$, and the results on other terms can be established similarly.

For b1), obviously $E(s_1) = 0$. Moreover, we have

$$\begin{aligned}
E\left(s_1^2\right) &= \frac{1}{(n(t+1))^2}\sum_{i=1}^{n(t+1)}\sum_{j=1}^{i-1}\sum_{k=1}^{j-1}\sum_{r=j+1}^{n(t+1)} h_{ij}h_{ik}h_{rj}h_{rk}\sigma_{v_i}^2\sigma_{v_r}^2\sigma_{v_j}^2\sigma_{v_k}^2 \\
&\le \frac{k_e^4}{(n(t+1))^2}\sum_{i=1}^{n(t+1)}\sum_{j=1}^{n(t+1)}\sum_{k=1}^{n(t+1)}\sum_{r=1}^{n(t+1)} |h_{ij}||h_{ik}||h_{rj}||h_{rk}|
\end{aligned}$$

$$= \frac{k_e^4}{(n(t+1))^2} \sum_{i=1}^{n(t+1)} \sum_{j=1}^{n(t+1)} |h_{ij}| \sum_{k=1}^{n(t+1)} |h_{ik}| \sum_{r=1}^{n(t+1)} |h_{rj}||h_{rk}|$$
$$\leq \frac{k_e^4}{(n(t+1))^2} \left(n(t+1)k_p^4\right) = \frac{k_e^4 k_p^4}{n(t+1)} \to 0,$$

as $n \to \infty$. By Chebyshev inequality, we have $s_1 = o_p(1)$.

For b2), it holds that

$$s_2 = \frac{1}{n(t+1)} \sum_{i=1}^{n(t+1)} \sum_{j=1}^{i-1} h_{ij}^2 \sigma_{v_i}^2 (v_j^2 - \sigma_{v_j}^2)$$
$$= \frac{1}{n(t+1)} \sum_{j=1}^{n(t+1)-1} \sum_{i=j+1}^{n(t+1)} h_{ij}^2 \sigma_{v_i}^2 (v_j^2 - \sigma_{v_j}^2)$$
$$= \frac{1}{n(t+1)} \sum_{i=1}^{n(t+1)-1} \sum_{j=i+1}^{n(t+1)} h_{ji}^2 \sigma_{v_j}^2 (v_i^2 - \sigma_{v_i}^2) \triangleq \sum_{i=1}^{n(t+1)-1} \chi_{i,n(t+1)} (v_i^2 - \sigma_{v_i}^2),$$

where $\chi_{i,n(t+1)} = \frac{1}{n(t+1)} \sum_{j=i+1}^{n(t+1)} h_{ji}^2 \sigma_{v_j}^2$. Note that $\chi_{i,n(t+1)}(v_i^2 - \sigma_{v_i}^2)$'s are mutually independent with zero mean. To verify $s_2 \to_p 0$ which is implied by $s_2 \to 0$ in L_1 space, by Theorem 19.7 of Davidson (1994), it suffices to show that $\{|v_i^2 - \sigma_{v_i}^2|\}$ is uniformly integrable, $\limsup_{n\to\infty} \sum_{i=1}^{n(t+1)-1} \chi_{i,n(t+1)} < \infty$ and $\lim_{n\to\infty} \sum_{i=1}^{n(t+1)-1} \chi_{i,n(t+1)}^2 = 0$. Recall that $E\left|v_i^2 - \sigma_{v_i}^2\right|^{1+\delta} \leq K_e$. Thus $\{|v_i^2 - \sigma_{v_i}^2|\}$ is uniformly integrable. Furthermore,

$$\limsup_{n\to\infty} \sum_{i=1}^{n(t+1)-1} \chi_{i,n(t+1)} \leq \limsup_{n\to\infty} \frac{1}{n(t+1)} \sum_{i=1}^{n(t+1)} \sum_{j=1}^{n(t+1)} h_{ji}^2 \sigma_{v_j}^2 \leq k_e k_p^2 < \infty, \quad \text{and}$$

$$\lim_{n\to\infty} \sum_{i=1}^{n(t+1)-1} \chi_{i,n(t+1)}^2 \leq \lim_{n\to\infty} \frac{1}{(n(t+1))^2} \sum_{i=1}^{n(t+1)} \sum_{j=1}^{n(t+1)} \sum_{k=1}^{n(t+1)} h_{ji}^2 h_{ki}^2 \sigma_{v_j}^2 \sigma_{v_k}^2 \leq \lim_{n\to\infty} \frac{k_e^2 k_p^4}{n(t+1)} = 0.$$

Then $s_2 \to 0$ in L_1 space, and hence $s_2 \to_p 0$. We complete this lemma by a) and b). □

Proof of Lemma 8. Recall that $\widetilde{Z} = (\boldsymbol{Y}_{-1}, Z, (\boldsymbol{\iota}_T \otimes I_N)X)$. Note that $E\left(Z'\Omega_{0NT}^{-1}\boldsymbol{V}_{NT}\right) = 0$ and $E\left(X'(\boldsymbol{\iota}_T \otimes I_N)'\Omega_{0NT}^{-1}\boldsymbol{V}_{NT}\right) = 0$, so it suffices to to show $E\left(\boldsymbol{Y}'_{-1}\Omega_{0NT}^{-1}\boldsymbol{V}_{NT}\right) = 0$ to verify that $E\left(\widetilde{Z}'\Omega_{0NT}^{-1}\boldsymbol{V}_{NT}\right) = 0$. By (S3.1) and (S3.3), it can be shown that

$$E\left(\boldsymbol{Y}'_{-1}\Omega_{0NT}^{-1}\boldsymbol{V}_{NT}\right) = tr\left[E(\boldsymbol{V}_{NT}\boldsymbol{Y}'_{-1}\Omega_{0NT}^{-1})\right] = \sigma_{\varepsilon 0}^2 tr\left[\Omega_{0NT}^{-1}\left((\boldsymbol{\iota}_T \otimes I_N)A_{0N}M_3' + M_2'\right)\right],$$

where $A_{0N} = A_N(\boldsymbol{\beta}_0^\star)$, and

$$
\begin{aligned}
& tr\left[\Omega_{0NT}^{-1}\left((\boldsymbol{\iota}_T \otimes I_N)A_{0N}M_3' + M_2'\right)\right] \\
&= tr\left[\left(\frac{1}{T}J_T \otimes (I_N + TA_{0N})^{-1} + \left(I_T - \frac{1}{T}J_T\right) \otimes I_N\right)\left((\boldsymbol{\iota}_T \otimes I_N)A_{0N}M_3' + M_2'\right)\right] \\
&= tr\left[\left(\frac{1}{T}J_T \otimes (I_N + TA_{0N})^{-1} + \left(I_T - \frac{1}{T}J_T\right) \otimes I_N\right)(\boldsymbol{\iota}_T \otimes I_N)A_{0N}M_3'\right] \\
&\quad + tr\left[\left(\frac{1}{T}J_T \otimes (I_N + TA_{0N})^{-1} + \left(I_T - \frac{1}{T}J_T\right) \otimes I_N\right)M_2'\right] \\
&= tr\left[\left(\boldsymbol{\iota}_T \otimes \left((I_N + TA_{0N})^{-1}A_{0N}\right)\right)M_3'\right] + tr\left[\left(\frac{J_T}{T} \otimes \left((I_N + TA_{0N})^{-1} - I_N\right)\right)M_2'\right] = 0.
\end{aligned}
$$

Then the proof of this lemma is accomplished. □

Proof of Lemma 9. We first show (i). Recall that $\Omega_{NT}(\boldsymbol{\beta}^\star) = J_T \otimes A_N(\boldsymbol{\beta}^\star) + I_{NT}$ with $A_N(\boldsymbol{\beta}^\star) = \operatorname{diag}\{(\boldsymbol{x}_{a1}'\boldsymbol{\beta}^\star)^2, \ldots, (\boldsymbol{x}_{aN}'\boldsymbol{\beta}^\star)^2\} \triangleq \operatorname{diag}\{a_1(\boldsymbol{\beta}^\star), \ldots, a_N(\boldsymbol{\beta}^\star)\}$. For $T^{-1}\Omega_{NT}(\boldsymbol{\beta}^\star)$, it suffices to show that the row sums corresponding to the first N rows are uniformly bounded, uniformly on $\Delta(\boldsymbol{\beta}^\star)$. Since $\Delta(\boldsymbol{\beta}^\star)$ is a compact set and x_i is bounded by Assumption 3.6(i), there exists a positive constant a_0 such that $\sup_{\boldsymbol{\beta}^\star} a_i(\boldsymbol{\beta}^\star) \le a_0 < \infty$ for $1 \le i \le N$. The ith row sum is $a_i(\boldsymbol{\beta}^\star) + T^{-1}$, then it follows that

$$
0 < \sup_{\boldsymbol{\beta}^\star}\left(a_i(\boldsymbol{\beta}^\star) + T^{-1}\right) \le a_0 + 1 < \infty.
$$

Therefore, $\{T^{-1}\Omega_{NT}(\boldsymbol{\beta}^\star)\}$ are uniformly bounded in both row and column sums, uniformly in $\Delta(\boldsymbol{\beta}^\star)$.

Next we show (ii). Recall that $\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star) = \frac{1}{T}J_T \otimes (I_N + TA_N(\boldsymbol{\beta}^\star))^{-1} + \left(I_T - \frac{1}{T}J_T\right) \otimes I_N$, then it follows that

$$
T\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star) = J_T \otimes \operatorname{diag}\left\{\frac{1}{1 + Ta_1(\boldsymbol{\beta}^\star)} - 1, \ldots, \frac{1}{1 + Ta_N(\boldsymbol{\beta}^\star)} - 1\right\} + TI_{NT}.
$$

For $T\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)$, we only need to show that the row sums corresponding to the first N rows are uniformly bounded, uniformly on $\Delta(\boldsymbol{\beta}^\star)$. For $1 \le i \le N$, note that $\inf_{\boldsymbol{\beta}^\star} a_i(\boldsymbol{\beta}^\star) =$

$\inf_{\boldsymbol{\beta}^\star}(\boldsymbol{x}'_{ai}\boldsymbol{\beta}^\star)^2 \geq \inf_{\boldsymbol{\beta}^\star}\beta_0^{\star 2} = \inf_{\boldsymbol{\beta}^\star}\beta_0^2\frac{\sigma_\eta^2}{\sigma_\varepsilon^2} > 0$. Then the ith row sum is $T\left(\frac{1}{1+Ta_i(\boldsymbol{\beta}^\star)}-1\right)+T = \left(1-\frac{1}{1+Ta_i(\boldsymbol{\beta}^\star)}\right)\frac{1}{a_i(\boldsymbol{\beta}^\star)}$, so it holds that

$$0 < \sup_{\boldsymbol{\beta}^\star}\left(1-\frac{1}{1+Ta_i(\boldsymbol{\beta}^\star)}\right)\frac{1}{a_i(\boldsymbol{\beta}^\star)} \leq \sup_{\boldsymbol{\beta}^\star}\frac{1}{a_i(\boldsymbol{\beta}^\star)} = \frac{1}{\inf_{\boldsymbol{\beta}^\star}a_i(\boldsymbol{\beta}^\star)} < \infty.$$

Hence, (ii) follows.

The proof of (iii) is analogous to that of (i). The proof of this lemma is complete. □

Proof of Lemma 10. This lemma directly follows from Lemma A.1 of Canay (2011) and Lemma B.2 of Chernozhukov and Hansen (2006), so its proof is omitted here. □

S5 Technical proofs

S5.1 Proof of Theorem 1

To establish the consistency of $\widehat{\boldsymbol{\delta}}$ and hence that of $\widehat{\boldsymbol{\phi}}$ and $\widehat{\sigma}_\varepsilon^2$, by Theorem 3.4 of White (1996), it suffices to show that:

(A1) the identification uniqueness condition holds, that is $\limsup_{N\to\infty}\max_{\boldsymbol{\delta}\in N_\epsilon^c(\boldsymbol{\delta}_0)}\frac{1}{NT}[\ln\ell_{NT}^\star(\boldsymbol{\delta}) - \ln\ell_{NT}^\star(\boldsymbol{\delta_0})] < 0$ for any $\epsilon > 0$, where $N_\epsilon^c(\boldsymbol{\delta}_0)$ is the complement of an open neighborhood of $\boldsymbol{\delta}_0$ on Δ of radius ϵ;

(A2) $\frac{1}{NT}[\ln\ell_{NT}(\boldsymbol{\delta}) - \ln\ell_{NT}^\star(\boldsymbol{\delta_0})] \to_p 0$, uniformly in $\boldsymbol{\delta}$ on Δ;

where $\ln\ell_{NT}(\boldsymbol{\delta}) = \max_{\boldsymbol{\phi},\sigma_\varepsilon^2}\ln\mathcal{L}_{NT}(\boldsymbol{\zeta})$ and $\ln\ell_{NT}^\star(\boldsymbol{\delta}) = \max_{\boldsymbol{\phi},\sigma_\varepsilon^2}E[\ln\mathcal{L}_{NT}(\boldsymbol{\zeta})]$ can be rewritten for our model in the following forms

$$\ln\ell_{NT}(\boldsymbol{\delta}) = -\frac{NT}{2}(\ln(2\pi)+1) - \frac{NT}{2}\ln\widehat{\sigma}_\varepsilon^2(\boldsymbol{\delta}) - \frac{1}{2}\ln|\Omega_{NT}(\boldsymbol{\beta}^\star)| + \ln|S_{NT}(\lambda)|,$$
$$\ln\ell_{NT}^\star(\boldsymbol{\delta}) = -\frac{NT}{2}(\ln(2\pi)+1) - \frac{NT}{2}\ln\widetilde{\sigma}_\varepsilon^2(\boldsymbol{\delta}) - \frac{1}{2}\ln|\Omega_{NT}(\boldsymbol{\beta}^\star)| + \ln|S_{NT}(\lambda)|.$$

First, we show (A1) by contradiction, which needs $\frac{1}{NT}\left[\ln \ell^{\star}_{NT}(\boldsymbol{\delta}) - \ln \ell^{\star}_{NT}(\boldsymbol{\delta_0})\right] \leq 0$ for all $\boldsymbol{\delta}$ on Δ, and $\frac{1}{NT}\ln \ell^{\star}_{NT}(\boldsymbol{\delta})$ is uniformly equicontinuous on Δ.

(Show that $\frac{1}{NT}\left[\ln \ell^{\star}_{NT}(\boldsymbol{\delta}) - \ln \ell^{\star}_{NT}(\boldsymbol{\delta_0})\right] \leq 0$ for all $\boldsymbol{\delta}$ on Δ) By Jensen's inequality, it can be verified that

$$0 = \ln 1 = \ln E\frac{\mathcal{L}_{NT}(\boldsymbol{\zeta})}{\mathcal{L}_{NT}(\boldsymbol{\zeta}_0)} \geq E\ln\frac{\mathcal{L}_{NT}(\boldsymbol{\zeta})}{\mathcal{L}_{NT}(\boldsymbol{\zeta}_0)} = E\ln\mathcal{L}_{NT}(\boldsymbol{\zeta}) - E\ln\mathcal{L}_{NT}(\boldsymbol{\zeta}_0),$$

so it follows that

$$\ln \ell^{\star}_{NT}(\boldsymbol{\delta}) = \max_{\boldsymbol{\phi},\sigma^2_{\varepsilon}} E\ln\mathcal{L}_{NT}(\boldsymbol{\zeta}) \leq \max_{\boldsymbol{\phi},\sigma^2_{\varepsilon}} E\ln\mathcal{L}_{NT}(\boldsymbol{\zeta}_0) = \ln \ell^{\star}_{NT}(\boldsymbol{\delta_0}),$$

and thus $\frac{1}{NT}\left[\ln \ell^{\star}_{NT}(\boldsymbol{\delta}) - \ln \ell^{\star}_{NT}(\boldsymbol{\delta_0})\right] \leq 0$ holds for all $\boldsymbol{\delta}$ on Δ.

(Show that $\frac{1}{NT}\ln \ell^{\star}_{NT}(\boldsymbol{\delta})$ is uniformly equicontinuous on Δ) It suffices to prove that $\frac{1}{NT}\ln|\Omega_{NT}(\boldsymbol{\beta}^{\star})|$, $\frac{1}{NT}\ln|S_{NT}(\lambda)|$ and $\ln\widetilde{\sigma}^2_{\varepsilon}(\boldsymbol{\delta})$ are all uniformly equicontinuous on Δ. Recall that Δ is the parameter space of $\boldsymbol{\delta} = (\lambda, \boldsymbol{\beta}^{\star\prime})'$, and denote $\Delta(\boldsymbol{\beta}^{\star})$ and $\Delta(\lambda)$ as the parameter space of $\boldsymbol{\beta}^{\star}$ and λ, respectively.

For $\frac{1}{NT}\ln|\Omega_{NT}(\boldsymbol{\beta}^{\star})|$, let $\boldsymbol{\beta}^{\star}_1 = (\beta^{\star}_{10}, \ldots, \beta^{\star}_{1p})'$ and $\boldsymbol{\beta}^{\star}_2 = (\beta^{\star}_{20}, \ldots, \beta^{\star}_{2p})'$ be in $\Delta(\boldsymbol{\beta}^{\star})$, by the mean value theorem, it holds that

$$\begin{aligned}\left|\frac{1}{NT}\ln|\Omega_{NT}(\boldsymbol{\beta}^{\star}_1)| - \frac{1}{NT}\ln|\Omega_{NT}(\boldsymbol{\beta}^{\star}_2)|\right| &= \left|\sum_{\ell=0}^{p}\frac{1}{NT}\frac{\partial}{\partial\beta^{\star}_{\ell}}\ln\left|\Omega_{NT}(\bar{\boldsymbol{\beta}}^{\star})\right|(\beta^{\star}_{1\ell} - \beta^{\star}_{2\ell})\right| \\ &\leq \sum_{\ell=0}^{p}\left|\frac{1}{NT}\frac{\partial}{\partial\beta^{\star}_{\ell}}\ln\left|\Omega_{NT}(\bar{\boldsymbol{\beta}}^{\star})\right|\right||\beta^{\star}_{1\ell} - \beta^{\star}_{2\ell}|,\end{aligned}$$

where $\bar{\boldsymbol{\beta}}^{\star}$ lies between $\boldsymbol{\beta}^{\star}_1$ and $\boldsymbol{\beta}^{\star}_2$. Moreover, by Lemmas 1 and 9, we have that

$$\frac{\partial}{\partial\beta^{\star}_{\ell}}\ln|\Omega_{NT}(\boldsymbol{\beta}^{\star})| = tr\left[T\Omega^{-1}_{NT}(\boldsymbol{\beta}^{\star})\frac{\partial}{\partial\beta^{\star}_{\ell}}T^{-1}\Omega_{NT}(\boldsymbol{\beta}^{\star})\right] = O(NT)$$

holds uniformly in $\boldsymbol{\beta}^{\star} \in \Delta(\boldsymbol{\beta}^{\star})$. Therefore, we can obtain that $\frac{1}{NT}\ln|\Omega_{NT}(\boldsymbol{\beta}^{\star})|$ is uniformly equicontinuous on Δ by its definition. Similarly, we can show the uniform equicontinuity of $\frac{1}{NT}\ln|S_{NT}(\lambda)|$.

For $\ln \widetilde{\sigma}_{\varepsilon}^2(\boldsymbol{\delta})$, we show the uniform equicontinuity of $\ln \widetilde{\sigma}_{\varepsilon}^2(\boldsymbol{\delta})$ by showing that 1) $\widetilde{\sigma}_{\varepsilon}^2(\boldsymbol{\delta})$ is uniformly equicontinuous on Δ, and 2) $\widetilde{\sigma}_{\varepsilon}^2(\boldsymbol{\delta})$ is uniformly bounded away from zero on Δ. We first consider 1). Notice that $\boldsymbol{Y} = S_{0NT}^{-1}\left(\widetilde{Z}\boldsymbol{\phi}_0 + \boldsymbol{V}_{NT}\right)$, and (3.8) implies that

$$\widetilde{\sigma}_{\varepsilon}^2(\boldsymbol{\delta}) = \frac{1}{NT} E\left(S_{NT}(\lambda)\boldsymbol{Y} - \widetilde{Z}\widetilde{\boldsymbol{\phi}}(\boldsymbol{\delta})\right)' \Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)\left(S_{NT}(\lambda)\boldsymbol{Y} - \widetilde{Z}\widetilde{\boldsymbol{\phi}}(\boldsymbol{\delta})\right) = \sum_{i=1}^{3} \widetilde{\sigma}_{\varepsilon i}^2(\boldsymbol{\delta}),$$

where

$$\begin{aligned}
\widetilde{\sigma}_{\varepsilon 1}^2(\boldsymbol{\delta}) &= \frac{1}{NT} E\left(\boldsymbol{V}_{NT}' \left(S_{0NT}^{-1}\right)' S_{NT}'(\lambda)\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star) S_{NT}(\lambda) S_{0NT}^{-1}\boldsymbol{V}_{NT}\right),\\
\widetilde{\sigma}_{\varepsilon 2}^2(\boldsymbol{\delta}) &= \frac{2}{NT} E\left[\left(S_{NT}(\lambda)S_{0NT}^{-1}\widetilde{Z}\boldsymbol{\phi}_0 - \widetilde{Z}\widetilde{\boldsymbol{\phi}}(\boldsymbol{\delta})\right)' \Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)\left(S_{NT}(\lambda)S_{0NT}^{-1}\boldsymbol{V}_{NT}\right)\right], \text{ and}\\
\widetilde{\sigma}_{\varepsilon 3}^2(\boldsymbol{\delta}) &= \frac{1}{NT} E\left[\left(S_{NT}(\lambda)S_{0NT}^{-1}\widetilde{Z}\boldsymbol{\phi}_0 - \widetilde{Z}\widetilde{\boldsymbol{\phi}}(\boldsymbol{\delta})\right)' \Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)\left(S_{NT}(\lambda)S_{0NT}^{-1}\widetilde{Z}\boldsymbol{\phi}_0 - \widetilde{Z}\widetilde{\boldsymbol{\phi}}(\boldsymbol{\delta})\right)\right].
\end{aligned}$$

Then the uniform equicontinuity of $\widetilde{\sigma}_{\varepsilon}^2(\boldsymbol{\delta})$ follows from that of $\widetilde{\sigma}_{\varepsilon 1}^2(\boldsymbol{\delta})$, $\widetilde{\sigma}_{\varepsilon 2}^2(\boldsymbol{\delta})$ and $\widetilde{\sigma}_{\varepsilon 3}^2(\boldsymbol{\delta})$. Similar to the proof of that of $\frac{1}{NT}\ln \ell_{NT}^\star(\boldsymbol{\delta})$, we can verify the uniform equicontinuity of $\widetilde{\sigma}_{\varepsilon i}^2(\boldsymbol{\delta})$ for $i = 1, 2, 3$. Hence, 1) holds. Next we prove 2) by contradiction. We have shown that $\frac{1}{NT}\left[\ln \ell_{NT}^\star(\boldsymbol{\delta}) - \ln \ell_{NT}^\star(\boldsymbol{\delta_0})\right] \le 0$ for all $\boldsymbol{\delta}$ on Δ, and

$$\begin{aligned}
\frac{1}{NT}\left[\ln \ell_{NT}^\star(\boldsymbol{\delta}) - \ln \ell_{NT}^\star(\boldsymbol{\delta_0})\right] = &-\frac{1}{2}\ln \widetilde{\sigma}_{\varepsilon}^2(\boldsymbol{\delta}) + \frac{1}{2}\ln \sigma_{\varepsilon 0}^2 - \frac{1}{2NT}\left[\ln|\Omega_{NT}(\boldsymbol{\beta}^\star)| - \ln|\Omega_{0NT}|\right]\\
&+\frac{1}{NT}\left[\ln|S_{NT}(\lambda)| - \ln|S_{0NT}|\right].
\end{aligned}$$

Then by the mean value theorem, it follows that

$$\begin{aligned}
&-\frac{1}{2}\ln \widetilde{\sigma}_{\varepsilon}^2(\boldsymbol{\delta})\\
\le& -\frac{1}{2}\ln \sigma_{\varepsilon 0}^2 + \frac{1}{2NT}\left[\ln|\Omega_{NT}(\boldsymbol{\beta}^\star)| - \ln|\Omega_{0NT}|\right] - \frac{1}{NT}\left[\ln|S_{NT}(\lambda)| - \ln|S_{0NT}|\right]\\
=& -\frac{1}{2}\ln \sigma_{\varepsilon 0}^2 + \frac{1}{2NT}\sum_{\ell=0}^{p}\frac{\partial}{\partial \beta_\ell^\star}\ln\left|\Omega_{NT}(\bar{\boldsymbol{\beta}}^\star)\right|(\beta_\ell^\star - \beta_{\ell 0}^\star) - \frac{1}{NT}\frac{\partial}{\partial \lambda}\ln\left|S_{NT}(\bar{\lambda})\right|(\lambda - \lambda_0),
\end{aligned}$$

where $\bar{\boldsymbol{\delta}} = (\bar{\lambda}, \bar{\boldsymbol{\beta}}^{\star\prime})'$ lies between $\boldsymbol{\delta} = (\lambda, \boldsymbol{\beta}^{\star\prime})'$ and $\boldsymbol{\delta}_0 = (\lambda_0, \boldsymbol{\beta}_0^{\star\prime})'$. This together with compactness of Δ, $\frac{\partial}{\partial \beta_\ell^\star}\ln\left|\Omega_{NT}(\bar{\boldsymbol{\beta}}^\star)\right| = O(NT)$ and $\frac{\partial}{\partial \lambda}\ln\left|S_{NT}(\bar{\lambda})\right| = O(NT)$ which are shown previously, implies that $-\frac{1}{2}\ln \widetilde{\sigma}_{\varepsilon}^2(\boldsymbol{\delta})$ is bounded above. Suppose that 2) does not hold, then

there exists a sequence $\{\boldsymbol{\delta}_N\}$ in Δ such that $\lim_{N\to\infty} \widetilde{\sigma}_\varepsilon^2(\boldsymbol{\delta}_N) = 0$, which is in contradiction with that $-\frac{1}{2}\ln\widetilde{\sigma}_\varepsilon^2(\boldsymbol{\delta})$ is bounded above. Hence 2) holds by contradiction. As a result, the uniform equicontinuity of $\ln\widetilde{\sigma}_\varepsilon^2(\boldsymbol{\delta})$ follows from 1) and 2).

(Show (A1) by contradiction) Suppose that the identification uniqueness condition does not hold, then there exists a constant $\epsilon_1 > 0$ and a sequence $\{\boldsymbol{\delta}_N\}$ in $N_{\epsilon_1}^c(\boldsymbol{\delta}_0)$ such that $\limsup_{N\to\infty} \frac{1}{NT}\left[\ln \ell_{NT}^\star(\boldsymbol{\delta}_N) - \ln \ell_{NT}^\star(\boldsymbol{\delta_0})\right] = 0$. Furthermore, because of the compactness of $N_{\epsilon_1}^c(\boldsymbol{\delta}_0)$, there exists a convergent subsequence $\{\boldsymbol{\delta}_{N_k}\}$ of $\{\boldsymbol{\delta}_N\}$ such that $\lim_{N_k\to\infty} \boldsymbol{\delta}_{N_k} = \boldsymbol{\delta}_1 \in N_{\epsilon_1}^c(\boldsymbol{\delta}_0)$, obviously $\boldsymbol{\delta}_1 \neq \boldsymbol{\delta}_0$. It follows that $\lim_{N_k\to\infty} \frac{1}{N_kT}\left[\ln \ell_{N_kT}^\star(\boldsymbol{\delta}_{N_k}) - \ln \ell_{N_kT}^\star(\boldsymbol{\delta_0})\right] = 0$, and then by the uniform equicontinuity of $\frac{1}{NT}\ln \ell_{NT}^\star(\boldsymbol{\delta})$, we have that

$$\lim_{N_k\to\infty} \frac{1}{N_kT}\left[\ln \ell_{N_kT}^\star(\boldsymbol{\delta}_1) - \ln \ell_{N_kT}^\star(\boldsymbol{\delta_0})\right] = 0. \tag{S5.9}$$

However, by Assumption 3.7, it holds that, for any $\boldsymbol{\delta} \neq \boldsymbol{\delta}_0$,

$$\begin{aligned}
&\lim_{N\to\infty} \frac{1}{NT}\left[\ln \ell_{NT}^\star(\boldsymbol{\delta}) - \ln \ell_{NT}^\star(\boldsymbol{\delta_0})\right] \\
&= \lim_{N\to\infty} \frac{1}{2NT}\left[\ln\left|\widetilde{\sigma}_\varepsilon^{-2}(\boldsymbol{\delta})\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)S_{NT}(\lambda)\right| - \ln\left|\sigma_{\varepsilon 0}^{-2}\Omega_{0NT}^{-1}S_{0NT}\right|\right] \neq 0,
\end{aligned}$$

which is in contradiction with (S5.9). Hence, the identification uniqueness condition holds.

It remains to show (A2). Since $\frac{1}{NT}\left[\ln \ell_{NT}(\boldsymbol{\delta}) - \ln \ell_{NT}^\star(\boldsymbol{\delta_0})\right] = -\frac{1}{2}\left[\ln\widehat{\sigma}_\varepsilon^2(\boldsymbol{\delta}) - \ln\widetilde{\sigma}_\varepsilon^2(\boldsymbol{\delta})\right]$, it suffices to show that $\widehat{\sigma}_\varepsilon^2(\boldsymbol{\delta}) - \widetilde{\sigma}_\varepsilon^2(\boldsymbol{\delta}) = o_p(1)$ uniformly in $\boldsymbol{\delta} \in \Delta$, and $\widetilde{\sigma}_\varepsilon^2(\boldsymbol{\delta})$ is uniformly bounded away from zero on Δ. The latter has been verified in (A1), hence we are left to show $\widehat{\sigma}_\varepsilon^2(\boldsymbol{\delta}) - \widetilde{\sigma}_\varepsilon^2(\boldsymbol{\delta}) = o_p(1)$ uniformly in $\boldsymbol{\delta} \in \Delta$. Denote

$$\begin{aligned}
&\overline{\Omega}(\boldsymbol{\delta}) = \Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)S_{NT}(\lambda)S_{0NT}^{-1},\ \widetilde{\Omega}(\boldsymbol{\delta}) = \left(S_{0NT}^{-1}\right)' S_{NT}'(\lambda)\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)S_{NT}(\lambda)S_{0NT}^{-1}, \\
&K_0(\boldsymbol{\delta}) = \frac{1}{NT}\widetilde{Z}'\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)\widetilde{Z},\ K_{10}(\boldsymbol{\delta}) = \frac{1}{NT}\widetilde{Z}'\overline{\Omega}(\boldsymbol{\delta})\widetilde{Z},\ K_{11}(\boldsymbol{\delta}) = \frac{1}{NT}\widetilde{Z}'\overline{\Omega}(\boldsymbol{\delta})\boldsymbol{V}_{NT}, \\
&K_{20}(\boldsymbol{\delta}) = \frac{1}{NT}\widetilde{Z}'\widetilde{\Omega}(\boldsymbol{\delta})\widetilde{Z},\ K_{21}(\boldsymbol{\delta}) = \frac{1}{NT}\widetilde{Z}'\widetilde{\Omega}(\boldsymbol{\delta})\boldsymbol{V}_{NT},\ K_{22}(\boldsymbol{\delta}) = \frac{1}{NT}\boldsymbol{V}_{NT}'\widetilde{\Omega}(\boldsymbol{\delta})\boldsymbol{V}_{NT}.
\end{aligned}$$

Then by (3.3), (3.7) and $\boldsymbol{Y} = S_{0NT}^{-1}\left(\widetilde{Z}\boldsymbol{\phi}_0 + \boldsymbol{V}_{NT}\right)$, it follows that

$$\begin{aligned}
\widehat{\boldsymbol{V}}_{NT}(\boldsymbol{\delta}) &= S_{NT}(\lambda)\boldsymbol{Y} - \widetilde{Z}\widehat{\boldsymbol{\phi}}(\boldsymbol{\delta}) \\
&= S_{NT}(\lambda)S_{0NT}^{-1}\left(\widetilde{Z}\boldsymbol{\phi}_0 + \boldsymbol{V}_{NT}\right) \\
&\quad -\widetilde{Z}\left(\widetilde{Z}'\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)\widetilde{Z}\right)^{-1}\widetilde{Z}'\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)S_{NT}(\lambda)S_{0NT}^{-1}\left(\widetilde{Z}\boldsymbol{\phi}_0 + \boldsymbol{V}_{NT}\right) \\
&= \left[S_{NT}(\lambda)S_{0NT}^{-1}\widetilde{Z}\boldsymbol{\phi}_0 - \widetilde{Z}K_0^{-1}(\boldsymbol{\delta})K_{10}(\boldsymbol{\delta})\boldsymbol{\phi}_0\right] \\
&\quad + \left[S_{NT}(\lambda)S_{0NT}^{-1}\boldsymbol{V}_{NT} - \widetilde{Z}K_0^{-1}(\boldsymbol{\delta})K_{11}(\boldsymbol{\delta})\right] \\
&\triangleq Q_1(\boldsymbol{\delta}) + Q_2(\boldsymbol{\delta}), \\
\widetilde{\boldsymbol{V}}_{NT}(\boldsymbol{\delta}) &= S_{NT}(\lambda)\boldsymbol{Y} - \widetilde{Z}\widetilde{\boldsymbol{\phi}}(\boldsymbol{\delta}) \\
&= S_{NT}(\lambda)S_{0NT}^{-1}\left(\widetilde{Z}\boldsymbol{\phi}_0 + \boldsymbol{V}_{NT}\right) \\
&\quad -\widetilde{Z}\left[E\left(\widetilde{Z}'\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)\widetilde{Z}\right)\right]^{-1}E\left[\widetilde{Z}'\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)S_{NT}(\lambda)S_{0NT}^{-1}\left(\widetilde{Z}\boldsymbol{\phi}_0 + \boldsymbol{V}_{NT}\right)\right] \\
&= \left\{S_{NT}(\lambda)S_{0NT}^{-1}\widetilde{Z}\boldsymbol{\phi}_0 - \widetilde{Z}\left[E\left(K_0(\boldsymbol{\delta})\right)\right]^{-1}E\left(K_{10}(\boldsymbol{\delta})\right)\boldsymbol{\phi}_0\right\} \\
&\quad + \left\{S_{NT}(\lambda)S_{0NT}^{-1}\boldsymbol{V}_{NT} - \widetilde{Z}\left[E\left(K_0(\boldsymbol{\delta})\right)\right]^{-1}E\left(K_{11}(\boldsymbol{\delta})\right)\right\} \\
&\triangleq P_1(\boldsymbol{\delta}) + P_2(\boldsymbol{\delta}).
\end{aligned}$$

These together with (3.4) and (3.8), imply that

$$\begin{aligned}
\widehat{\sigma}_\varepsilon^2(\boldsymbol{\delta}) - \widetilde{\sigma}_\varepsilon^2(\boldsymbol{\delta}) &= \frac{1}{NT}\widehat{\boldsymbol{V}}'_{NT}(\boldsymbol{\delta})\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)\widehat{\boldsymbol{V}}_{NT}(\boldsymbol{\delta}) - \frac{1}{NT}E\left(\widetilde{\boldsymbol{V}}'_{NT}(\boldsymbol{\delta})\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)\widetilde{\boldsymbol{V}}_{NT}(\boldsymbol{\delta})\right) \\
&= \frac{1}{NT}\left[Q_1'(\boldsymbol{\delta})\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)Q_1(\boldsymbol{\delta}) - E\left(P_1'(\boldsymbol{\delta})\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)P_1(\boldsymbol{\delta})\right)\right] \\
&\quad + \frac{1}{NT}\left[Q_2'(\boldsymbol{\delta})\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)Q_2(\boldsymbol{\delta}) - E\left(P_2'(\boldsymbol{\delta})\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)P_2(\boldsymbol{\delta})\right)\right] \\
&\quad + \frac{2}{NT}\left[Q_1'(\boldsymbol{\delta})\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)Q_2(\boldsymbol{\delta}) - E\left(P_1'(\boldsymbol{\delta})\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)P_2(\boldsymbol{\delta})\right)\right] \\
&\triangleq \Pi_1(\boldsymbol{\delta}) + \Pi_2(\boldsymbol{\delta}) + 2\Pi_3(\boldsymbol{\delta}).
\end{aligned}$$

Therefore, to show $\widehat{\sigma}_\varepsilon^2(\boldsymbol{\delta}) - \widetilde{\sigma}_\varepsilon^2(\boldsymbol{\delta}) = o_p(1)$ uniformly in $\boldsymbol{\delta} \in \Delta$, it suffices to show that $\Pi_1(\boldsymbol{\delta})$,

$\Pi_2(\boldsymbol{\delta})$ and $\Pi_3(\boldsymbol{\delta})$ are all $o_p(1)$ uniformly in $\boldsymbol{\delta} \in \Delta$. Next we rewrite $\Pi_i(\boldsymbol{\delta})$ for $i = 1, 2, 3$ below

$$
\begin{aligned}
\Pi_1(\boldsymbol{\delta}) &= \frac{1}{NT}\left[Q_1'(\boldsymbol{\delta})\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)Q_1(\boldsymbol{\delta}) - E\left(P_1'(\boldsymbol{\delta})\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)P_1(\boldsymbol{\delta})\right)\right] \\
&= \boldsymbol{\phi}_0'\Big\{ \left[K_{20}(\boldsymbol{\delta}) - E\left(K_{20}(\boldsymbol{\delta})\right)\right] \\
&\quad - \left[K_{10}'(\boldsymbol{\delta})K_0^{-1}(\boldsymbol{\delta})K_{10}(\boldsymbol{\delta}) - \left[E\left(K_{10}(\boldsymbol{\delta})\right)\right]'\left[E\left(K_0(\boldsymbol{\delta})\right)\right]^{-1}E\left(K_{10}(\boldsymbol{\delta})\right)\right]\Big\}\boldsymbol{\phi}_0 \\
&= \boldsymbol{\phi}_0'\Big\{ \left[K_{20}(\boldsymbol{\delta}) - E\left(K_{20}(\boldsymbol{\delta})\right)\right] \\
&\quad - \left[K_{10}(\boldsymbol{\delta}) - E\left(K_{10}(\boldsymbol{\delta})\right)\right]' K_0^{-1}(\boldsymbol{\delta})K_{10}(\boldsymbol{\delta}) \\
&\quad + \left[E\left(K_{10}(\boldsymbol{\delta})\right)\right]' K_0^{-1}(\boldsymbol{\delta})\left[K_0(\boldsymbol{\delta}) - E\left(K_0(\boldsymbol{\delta})\right)\right]\left[E\left(K_0(\boldsymbol{\delta})\right)\right]^{-1}K_{10}(\boldsymbol{\delta}) \\
&\quad - \left[E\left(K_{10}(\boldsymbol{\delta})\right)\right]'\left[E\left(K_0(\boldsymbol{\delta})\right)\right]^{-1}\left[K_{10}(\boldsymbol{\delta}) - E\left(K_{10}(\boldsymbol{\delta})\right)\right]\Big\}\boldsymbol{\phi}_0 \\
&\triangleq \boldsymbol{\phi}_0'\left\{\Pi_{11}(\boldsymbol{\delta}) - \Pi_{12}(\boldsymbol{\delta}) + \Pi_{13}(\boldsymbol{\delta}) - \Pi_{14}(\boldsymbol{\delta})\right\}\boldsymbol{\phi}_0, \\
\Pi_2(\boldsymbol{\delta}) &= \frac{1}{NT}\left[Q_2'(\boldsymbol{\delta})\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)Q_2(\boldsymbol{\delta}) - E\left(P_2'(\boldsymbol{\delta})\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)P_2(\boldsymbol{\delta})\right)\right] \\
&= \left[K_{22}(\boldsymbol{\delta}) - E\left(K_{22}(\boldsymbol{\delta})\right)\right] \\
&\quad - \left[K_{11}(\boldsymbol{\delta}) - E\left(K_{11}(\boldsymbol{\delta})\right)\right]' K_0^{-1}(\boldsymbol{\delta})K_{11}(\boldsymbol{\delta}) \\
&\quad + \left[E\left(K_{11}(\boldsymbol{\delta})\right)\right]' K_0^{-1}(\boldsymbol{\delta})\left[K_0(\boldsymbol{\delta}) - E\left(K_0(\boldsymbol{\delta})\right)\right]\left[E\left(K_0(\boldsymbol{\delta})\right)\right]^{-1}K_{11}(\boldsymbol{\delta}) \\
&\quad - \left[E\left(K_{11}(\boldsymbol{\delta})\right)\right]'\left[E\left(K_0(\boldsymbol{\delta})\right)\right]^{-1}\left[K_{11}(\boldsymbol{\delta}) - E\left(K_{11}(\boldsymbol{\delta})\right)\right] \\
&\triangleq \Pi_{21}(\boldsymbol{\delta}) - \Pi_{22}(\boldsymbol{\delta}) + \Pi_{23}(\boldsymbol{\delta}) - \Pi_{24}(\boldsymbol{\delta}), \\
\Pi_3(\boldsymbol{\delta}) &= \frac{1}{NT}\left[Q_1'(\boldsymbol{\delta})\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)Q_2(\boldsymbol{\delta}) - E\left(P_1'(\boldsymbol{\delta})\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)P_2(\boldsymbol{\delta})\right)\right] \\
&= \boldsymbol{\phi}_0'\Big\{ \left[K_{21}(\boldsymbol{\delta}) - E\left(K_{21}(\boldsymbol{\delta})\right)\right] \\
&\quad - \left[K_{10}(\boldsymbol{\delta}) - E\left(K_{10}(\boldsymbol{\delta})\right)\right]' K_0^{-1}(\boldsymbol{\delta})K_{11}(\boldsymbol{\delta}) \\
&\quad + \left[E\left(K_{10}(\boldsymbol{\delta})\right)\right]' K_0^{-1}(\boldsymbol{\delta})\left[K_0(\boldsymbol{\delta}) - E\left(K_0(\boldsymbol{\delta})\right)\right]\left[E\left(K_0(\boldsymbol{\delta})\right)\right]^{-1}K_{11}(\boldsymbol{\delta}) \\
&\quad - \left[E\left(K_{10}(\boldsymbol{\delta})\right)\right]'\left[E\left(K_0(\boldsymbol{\delta})\right)\right]^{-1}\left[K_{11}(\boldsymbol{\delta}) - E\left(K_{11}(\boldsymbol{\delta})\right)\right]\Big\} \\
&\triangleq \boldsymbol{\phi}_0'\left\{\Pi_{31}(\boldsymbol{\delta}) - \Pi_{32}(\boldsymbol{\delta}) + \Pi_{33}(\boldsymbol{\delta}) - \Pi_{34}(\boldsymbol{\delta})\right\}.
\end{aligned}
$$

(Show that $\Pi_1(\boldsymbol{\delta}) = o_p(1)$ uniformly in $\boldsymbol{\delta} \in \Delta$) It suffices to show that all elements of $\Pi_{11}(\boldsymbol{\delta})$, $\Pi_{12}(\boldsymbol{\delta})$, $\Pi_{13}(\boldsymbol{\delta})$ and $\Pi_{14}(\boldsymbol{\delta})$ are $o_p(1)$ uniformly in $\boldsymbol{\delta} \in \Delta$, respectively. We establish the proof by verifying the following parts i)–iv).

i) We first show that $K_{20}(\boldsymbol{\delta}) - E\left(K_{20}(\boldsymbol{\delta})\right)$, $K_{10}(\boldsymbol{\delta}) - E\left(K_{10}(\boldsymbol{\delta})\right)$ and $K_0(\boldsymbol{\delta}) - E\left(K_0(\boldsymbol{\delta})\right)$ are all $o_p(1)$ uniformly on Δ. For $K_{20}(\boldsymbol{\delta}) - E\left(K_{20}(\boldsymbol{\delta})\right)$, we have that

$$K_{20}(\boldsymbol{\delta}) - E\left(K_{20}(\boldsymbol{\delta})\right) = \begin{pmatrix} e_{11} & * & * \\ \boldsymbol{e}_{21} & e_{22} & * \\ \boldsymbol{e}_{31} & e_{32} & e_{33} \end{pmatrix}, \tag{S5.10}$$

where

$$\begin{aligned}
e_{11} &= \frac{1}{NT}\left[\boldsymbol{Y}'_{-1}\widetilde{\Omega}(\boldsymbol{\delta})\boldsymbol{Y}_{-1} - E\left(\boldsymbol{Y}'_{-1}\widetilde{\Omega}(\boldsymbol{\delta})\boldsymbol{Y}_{-1}\right)\right], \\
\boldsymbol{e}_{21} &= \frac{1}{NT}\left[Z'\widetilde{\Omega}(\boldsymbol{\delta})\boldsymbol{Y}_{-1} - E\left(Z'\widetilde{\Omega}(\boldsymbol{\delta})\boldsymbol{Y}_{-1}\right)\right], \\
e_{22} &= \frac{1}{NT}\left[Z'\widetilde{\Omega}(\boldsymbol{\delta})Z - E\left(Z'\widetilde{\Omega}(\boldsymbol{\delta})Z\right)\right], \\
\boldsymbol{e}_{31} &= \frac{1}{NT}\left[X'(\boldsymbol{\iota}_{NT}\otimes I_N)'\widetilde{\Omega}(\boldsymbol{\delta})\boldsymbol{Y}_{-1} - E\left(X'(\boldsymbol{\iota}_{NT}\otimes I_N)'\widetilde{\Omega}(\boldsymbol{\delta})\boldsymbol{Y}_{-1}\right)\right], \\
e_{32} &= \frac{1}{NT}\left[X'(\boldsymbol{\iota}_{NT}\otimes I_N)'\widetilde{\Omega}(\boldsymbol{\delta})Z - E\left(X'(\boldsymbol{\iota}_{NT}\otimes I_N)'\widetilde{\Omega}(\boldsymbol{\delta})Z\right)\right], \\
e_{33} &= \frac{1}{NT}\left[X'(\boldsymbol{\iota}_{NT}\otimes I_N)'\widetilde{\Omega}(\boldsymbol{\delta})(\boldsymbol{\iota}_{NT}\otimes I_N)X - E\left(X'(\boldsymbol{\iota}_{NT}\otimes I_N)'\widetilde{\Omega}(\boldsymbol{\delta})(\boldsymbol{\iota}_{NT}\otimes I_N)X\right)\right].
\end{aligned}$$

Here we only show the element $e_{11} = o_p(1)$ uniformly on Δ, and the others can be verified similarly. By (S3.3), we have $\boldsymbol{Y}_{-1} = M_1\boldsymbol{y}_0 + M_2Z\boldsymbol{\gamma}_0 + M_3(X\boldsymbol{\psi}_0 + \boldsymbol{\vartheta}) + M_2\boldsymbol{\varepsilon}$, then it follows that $e_{11} \triangleq 2(S_1 + S_2)$, where

$$\begin{aligned}
S_1 = &\frac{1}{NT}\boldsymbol{\gamma}_0'Z'M_2'\widetilde{\Omega}(\boldsymbol{\delta})M_1\boldsymbol{y}_0 + \frac{1}{NT}\boldsymbol{\varepsilon}'M_2'\widetilde{\Omega}(\boldsymbol{\delta})M_1\boldsymbol{y}_0 + \frac{1}{NT}\boldsymbol{\vartheta}'M_3'\widetilde{\Omega}(\boldsymbol{\delta})M_1\boldsymbol{y}_0 \\
&+\frac{1}{NT}\boldsymbol{\gamma}_0'Z'M_2'\widetilde{\Omega}(\boldsymbol{\delta})M_3X\boldsymbol{\psi}_0 + \frac{1}{NT}\boldsymbol{\varepsilon}'M_2'\widetilde{\Omega}(\boldsymbol{\delta})M_3X\boldsymbol{\psi}_0 \\
&+\frac{1}{NT}\boldsymbol{\vartheta}'M_3'\widetilde{\Omega}(\boldsymbol{\delta})M_3X\boldsymbol{\psi}_0, \\
S_2 = &\frac{1}{2NT}\left[\boldsymbol{\gamma}_0'Z'M_2'\widetilde{\Omega}(\boldsymbol{\delta})M_2Z\boldsymbol{\gamma}_0 - E\left(\boldsymbol{\gamma}_0'Z'M_2'\widetilde{\Omega}(\boldsymbol{\delta})M_2Z\boldsymbol{\gamma}_0\right)\right]
\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{2NT}\left[\boldsymbol{\varepsilon}' M_2' \widetilde{\Omega}(\boldsymbol{\delta}) M_2 \boldsymbol{\varepsilon} - E\left(\boldsymbol{\varepsilon}' M_2' \widetilde{\Omega}(\boldsymbol{\delta}) M_2 \boldsymbol{\varepsilon}\right)\right] \\
&+\frac{1}{2NT}\left[\boldsymbol{\vartheta}' M_3' \widetilde{\Omega}(\boldsymbol{\delta}) M_3 \boldsymbol{\vartheta} - E\left(\boldsymbol{\vartheta}' M_3' \widetilde{\Omega}(\boldsymbol{\delta}) M_3 \boldsymbol{\vartheta}\right)\right] \\
&+\frac{1}{NT}\boldsymbol{\gamma}_0' Z' M_2' \widetilde{\Omega}(\boldsymbol{\delta}) M_2 \boldsymbol{\varepsilon} + \frac{1}{NT}\boldsymbol{\gamma}_0' Z' M_2' \widetilde{\Omega}(\boldsymbol{\delta}) M_3 \boldsymbol{\vartheta} + \frac{1}{NT}\boldsymbol{\varepsilon}' M_2' \widetilde{\Omega}(\boldsymbol{\delta}) M_3 \boldsymbol{\vartheta}.
\end{aligned}
$$

Note that under Assumption 3.5, M_1, M_2 and $\frac{1}{T}M_3$ are all uniformly bounded in both row and column sums, and $T\widetilde{\Omega}(\boldsymbol{\delta}) = \left(S_{0NT}^{-1}\right)' S_{NT}'(\lambda) T \Omega_{NT}^{-1}(\boldsymbol{\beta}^\star) S_{NT}(\lambda) S_{0NT}^{-1}$ and $T\frac{\partial}{\partial \boldsymbol{\delta}}\widetilde{\Omega}(\boldsymbol{\delta})$ are both uniformly bounded in both row and column sums, uniformly on Δ by Lemmas 2 and 9. Then we have that $M_2'\widetilde{\Omega}(\boldsymbol{\delta})M_1$, $M_2'\widetilde{\Omega}(\boldsymbol{\delta})M_2$, $M_3'\widetilde{\Omega}(\boldsymbol{\delta})M_1 = \frac{1}{T}M_3'T\widetilde{\Omega}(\boldsymbol{\delta})M_1$, $M_2'\widetilde{\Omega}(\boldsymbol{\delta})M_3 = M_2'T\widetilde{\Omega}(\boldsymbol{\delta})\frac{1}{T}M_3$ and $\frac{1}{T}M_3'\widetilde{\Omega}(\boldsymbol{\delta})M_3 = \frac{1}{T}M_3'T\widetilde{\Omega}(\boldsymbol{\delta})\frac{1}{T}M_3$ as well as their derivatives with respect to $\boldsymbol{\delta}$ are uniformly bounded in both row and column sums uniformly on Δ by Lemma 2(iii). Moreover, elements of previous matrices are all uniformly bounded on Δ by Lemma 1. It follows that $S_1 = o_p(1)$ holds uniformly on Δ by Lemma 4, and $S_2 = o_p(1)$ holds uniformly on Δ by Lemma 5 and Lemma 6. As a result, $K_{20}(\boldsymbol{\delta}) - E\left(K_{20}(\boldsymbol{\delta})\right) = o_p(1)$ holds uniformly on Δ. Similarly, we can show that $K_{10}(\boldsymbol{\delta}) - E\left(K_{10}(\boldsymbol{\delta})\right) = o_p(1)$ and $K_0(\boldsymbol{\delta}) - E\left(K_0(\boldsymbol{\delta})\right) = o_p(1)$ hold uniformly on Δ.

ii) Next we prove that $K_0^{-1}(\boldsymbol{\delta}) = O_p(1)$. All elements of $K_0^{-1}(\boldsymbol{\delta})$ are $O_p(1)$ holds uniformly on Δ if and only if $\left\|K_0^{-1}(\boldsymbol{\delta})\right\|_F^2 = O_p(1)$ uniformly on Δ. Denote $d_0 = 2+p+q$, by Bernstein (2009), we have

$$
\begin{aligned}
\sup_{\boldsymbol{\beta}^\star}\left\|K_0^{-1}(\boldsymbol{\delta})\right\|_F^2 &= \sup_{\boldsymbol{\beta}^\star} tr\left[\left(\frac{1}{NT}\widetilde{Z}'\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)\widetilde{Z}\right)^{-1}\left(\frac{1}{NT}\widetilde{Z}'\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)\widetilde{Z}\right)^{-1}\right] \\
&\leq d_0 \sup_{\boldsymbol{\beta}^\star}\lambda_{\max}\left[\left(\frac{1}{NT}\widetilde{Z}'\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)\widetilde{Z}\right)^{-1}\left(\frac{1}{NT}\widetilde{Z}'\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)\widetilde{Z}\right)^{-1}\right] \\
&= d_0 \sup_{\boldsymbol{\beta}^\star}\lambda_{\max}^2\left[\left(\frac{1}{NT}\widetilde{Z}'\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)\widetilde{Z}\right)^{-1}\right] \\
&= d_0 \sup_{\boldsymbol{\beta}^\star}\lambda_{\min}^{-2}\left(\frac{1}{NT}\widetilde{Z}'\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)\widetilde{Z}\right)
\end{aligned}
$$

$$= d_0 \sup_{\boldsymbol{\beta}^\star} \left[\lambda_{\min}^2 \left(\frac{1}{NT} \widetilde{Z}' \Omega_{NT}^{-1}(\boldsymbol{\beta}^\star) \widetilde{Z} \right) \right]^{-1},$$

where $\|\cdot\|_F$ denotes the Frobenius norm of a matrix, $\lambda_{\min}(\cdot)$ and $\lambda_{\max}(\cdot)$ are the smallest and largest eigenvalues of a matrix, respectively. It follows $\sup_N E \sup_{\boldsymbol{\beta}^\star} \left\| K_0^{-1}(\boldsymbol{\delta}) \right\|_F^2 \leq d_0 \sup_N E \sup_{\boldsymbol{\beta}^\star} \left[\lambda_{\min}^2 \left(\frac{1}{NT} \widetilde{Z}' \Omega_{NT}^{-1}(\boldsymbol{\beta}^\star) \widetilde{Z} \right) \right]^{-1} < \infty$ by Assumption 3.6(iv), then $\left\| K_0^{-1}(\boldsymbol{\delta}) \right\|_F^2 = O_p(1)$ holds uniformly on Δ by page 339 in Davidson (1994), and equivalently, all elements of $K_0^{-1}(\boldsymbol{\delta})$ are $O_p(1)$ uniformly on Δ.

iii) Below we show that all elements of $E\left(K_{10}(\boldsymbol{\delta})\right)$ and $\left[E\left(K_0(\boldsymbol{\delta})\right)\right]^{-1}$ are uniformly bounded on Δ. We only need to show that each element of $E\left(K_{10}(\boldsymbol{\delta})\right)$ and $E\left(K_0(\boldsymbol{\delta})\right)$ is uniformly bounded on Δ. As for (S5.10), we next show the element $\frac{1}{NT} E\left(\boldsymbol{Y}'_{-1} \Omega_{NT}^{-1}(\boldsymbol{\beta}^\star) \boldsymbol{Y}_{-1} \right)$ in $E\left(K_0(\boldsymbol{\delta})\right)$ is $O(1)$ uniformly on Δ. Note that elements of M_1, M_2 and M_3 are uniformly bounded, and M_1, M_2 and $\frac{1}{T} M_3$ are all uniformly bounded in both row and column sums. Then for $M_2' \Omega_{NT}^{-1}(\boldsymbol{\beta}^\star) M_2$ and $M_3' \Omega_{NT}^{-1}(\boldsymbol{\beta}^\star) M_3 = (M_3)' T \Omega_{NT}^{-1}(\boldsymbol{\beta}^\star) \frac{1}{T} M_3$, their elements are all uniformly bounded on Δ by Lemma 1, so by Lemma 5 (i), we obtain that

$$\begin{aligned} \frac{1}{NT} E\left(\boldsymbol{Y}'_{-1} \Omega_{NT}^{-1}(\boldsymbol{\beta}^\star) \boldsymbol{Y}_{-1} \right) &= \frac{1}{NT} E\left(\boldsymbol{\vartheta}' M_3' \Omega_{NT}^{-1}(\boldsymbol{\beta}^\star) M_3 \boldsymbol{\vartheta} \right) + \frac{1}{NT} E\left(\boldsymbol{\varepsilon}' M_2' \Omega_{NT}^{-1}(\boldsymbol{\beta}^\star) M_2 \boldsymbol{\varepsilon} \right) \\ &\quad + \frac{1}{NT} E\left(\boldsymbol{\gamma}_0' Z' M_2' \Omega_{NT}^{-1}(\boldsymbol{\beta}^\star) M_2 Z \boldsymbol{\gamma}_0 \right) = O(1) \end{aligned}$$

holds uniformly on Δ.

iv) Analogous to the proofs of i) and iii), we can show that $K_{11}(\boldsymbol{\delta}) - E\left(K_{11}(\boldsymbol{\delta})\right) = o_p(1)$ and $E\left(K_{11}(\boldsymbol{\delta})\right) = O(1)$ hold uniformly on Δ. It follows that $K_{11}(\boldsymbol{\delta}) = O_p(1)$ holds uniformly on Δ. As a result, with i)–iv), we can obtain that $\Pi_1(\boldsymbol{\delta}) = o_p(1)$ holds uniformly in $\boldsymbol{\delta} \in \Delta$.

With the similar argument, we can show that $\Pi_2(\boldsymbol{\delta})$ and $\Pi_3(\boldsymbol{\delta})$ are also $o_p(1)$ uniformly in $\boldsymbol{\delta} \in \Delta$. This completes the proof of (A2).

The consistency of $\widehat{\boldsymbol{\delta}}$ and hence that of $\widehat{\boldsymbol{\phi}}$ and $\widehat{\sigma}_\varepsilon^2$ follow from (A1) and (A2). We accomplish the proof of this theorem.

S5.2 Proof of Theorem 2

By Taylor's expansion, we have

$$0=\frac{1}{\sqrt{NT}}\frac{\partial\ln\mathcal{L}_{NT}(\widehat{\boldsymbol{\zeta}})}{\partial\boldsymbol{\zeta}}=\frac{1}{\sqrt{NT}}\frac{\partial\ln\mathcal{L}_{NT}(\boldsymbol{\zeta}_0)}{\partial\boldsymbol{\zeta}}+\frac{1}{NT}\frac{\partial^2\ln\mathcal{L}_{NT}(\bar{\boldsymbol{\zeta}})}{\partial\boldsymbol{\zeta}\partial\boldsymbol{\zeta}'}\sqrt{NT}(\widehat{\boldsymbol{\zeta}}-\boldsymbol{\zeta}_0),$$

where elements of $\bar{\boldsymbol{\zeta}}=(\bar{\boldsymbol{\phi}}',\bar{\sigma}_\varepsilon^2,\bar{\lambda},\bar{\boldsymbol{\beta}}^{\star\prime})'$ lie in the segment joining the corresponding elements of $\widehat{\boldsymbol{\zeta}}$ and $\boldsymbol{\zeta}_0$. It follows that

$$\sqrt{NT}(\widehat{\boldsymbol{\zeta}}-\boldsymbol{\zeta}_0)=-\left(\frac{1}{NT}\frac{\partial^2\ln\mathcal{L}_{NT}(\bar{\boldsymbol{\zeta}})}{\partial\boldsymbol{\zeta}\partial\boldsymbol{\zeta}'}\right)^{-1}\frac{1}{\sqrt{NT}}\frac{\partial\ln\mathcal{L}_{NT}(\boldsymbol{\zeta}_0)}{\partial\boldsymbol{\zeta}}.$$

Then, to verify Theorem 2, it suffices to show that

(B1) $\frac{1}{\sqrt{NT}}\frac{\partial\ln\mathcal{L}_{NT}(\boldsymbol{\zeta}_0)}{\partial\boldsymbol{\zeta}}\to_d N\left(0,\Sigma+\Sigma^\star\right)$;

(B2) $\frac{1}{NT}\frac{\partial^2\ln\mathcal{L}_{NT}(\bar{\boldsymbol{\zeta}})}{\partial\boldsymbol{\zeta}\partial\boldsymbol{\zeta}'}-\frac{1}{NT}\frac{\partial^2\ln\mathcal{L}_{NT}(\boldsymbol{\zeta}_0)}{\partial\boldsymbol{\zeta}\partial\boldsymbol{\zeta}'}=o_p(1)$;

(B3) $\frac{1}{NT}\frac{\partial^2\ln\mathcal{L}_{NT}(\boldsymbol{\zeta}_0)}{\partial\boldsymbol{\zeta}\partial\boldsymbol{\zeta}'}-E\left(\frac{1}{NT}\frac{\partial^2\ln\mathcal{L}_{NT}(\boldsymbol{\zeta}_0)}{\partial\boldsymbol{\zeta}\partial\boldsymbol{\zeta}'}\right)=o_p(1)$.

It can be verified that all elements of $\frac{1}{\sqrt{NT}}\frac{\partial\ln\mathcal{L}_{NT}(\boldsymbol{\zeta}_0)}{\partial\boldsymbol{\zeta}}$ can be rewritten in form of (S4.5). For example, $\dot{d}_{\sigma_\varepsilon^2}$ can be decomposed as follows

$$\dot{d}_{\sigma_\varepsilon^2}+\frac{NT}{2\sigma_{\varepsilon0}^2}=\left(2\sigma_{\varepsilon0}^4\right)^{-1}\left[\boldsymbol{\vartheta}'(\boldsymbol{\iota}_T\otimes I_N)'\Omega_{0NT}^{-1}(\boldsymbol{\iota}_T\otimes I_N)\boldsymbol{\vartheta}+2\boldsymbol{\vartheta}'(\boldsymbol{\iota}_T\otimes I_N)'\Omega_{0NT}^{-1}\boldsymbol{\varepsilon}+\boldsymbol{\varepsilon}'\Omega_{0NT}^{-1}\boldsymbol{\varepsilon}\right].$$

Note that Ω_{0NT}^{-1} and $\Omega_{0NT}^{-1}(\boldsymbol{\iota}_T\otimes I_N)=T\Omega_{0NT}^{-1}\frac{1}{T}(\boldsymbol{\iota}_T\otimes I_N)$ are both uniformly bounded in both row and column sums, and $\frac{1}{\sqrt{NT}}(\boldsymbol{\iota}_T\otimes I_N)'\Omega_{0NT}^{-1}(\boldsymbol{\iota}_T\otimes I_N)=\sqrt{\frac{T}{N}}\frac{1}{T}(\boldsymbol{\iota}_T\otimes I_N)'T\Omega_{0NT}^{-1}\frac{1}{T}(\boldsymbol{\iota}_T\otimes I_N)$ is also uniformly bounded in both row and column sums. Then (B1) holds by Lemma 7. Moreover, with the similar argument of the proof of Theorem 1, we can show that all elements of $\frac{1}{NT}\frac{\partial^2\ln\mathcal{L}_{NT}(\boldsymbol{\zeta}_0)}{\partial\boldsymbol{\zeta}\partial\boldsymbol{\zeta}'}-E\left(\frac{1}{NT}\frac{\partial^2\ln\mathcal{L}_{NT}(\boldsymbol{\zeta}_0)}{\partial\boldsymbol{\zeta}\partial\boldsymbol{\zeta}'}\right)$ are $o_p(1)$ and thus (B3) holds. Then we are left to prove the result of (B2).

(Show (B2)) Recall that $\boldsymbol{\zeta}=(\boldsymbol{\phi}',\sigma_\varepsilon^2,\lambda,\boldsymbol{\beta}^{\star\prime})'$. We first show that $\frac{1}{NT}\frac{\partial^2\ln\mathcal{L}_{NT}(\bar{\boldsymbol{\zeta}})}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}'}-\frac{1}{NT}\frac{\partial^2\ln\mathcal{L}_{NT}(\boldsymbol{\zeta}_0)}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}'}=o_p(1)$. Note that $\frac{1}{NT}\frac{\partial^2\ln\mathcal{L}_{NT}(\boldsymbol{\zeta})}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}'}=-\frac{1}{\sigma_\varepsilon^2}\frac{1}{NT}\widetilde{Z}'\Omega_{NT}^{-1}(\boldsymbol{\beta}^\star)\widetilde{Z}$, by the mean value theorem, it can

be verified that

$$
\begin{aligned}
&\frac{1}{NT}\frac{\partial^2 \ln \mathcal{L}_{NT}(\bar{\boldsymbol{\zeta}})}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}'} - \frac{1}{NT}\frac{\partial^2 \ln \mathcal{L}_{NT}(\boldsymbol{\zeta}_0)}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}'} \\
=& -\frac{1}{\bar{\sigma}_\varepsilon^2}\frac{1}{NT}\widetilde{Z}'\Omega_{NT}^{-1}(\bar{\boldsymbol{\beta}}^\star)\widetilde{Z} + \frac{1}{\sigma_{\varepsilon 0}^2}\frac{1}{NT}\widetilde{Z}'\Omega_{0NT}^{-1}\widetilde{Z} \\
=& \left(\frac{1}{\sigma_{\varepsilon 0}^2} - \frac{1}{\bar{\sigma}_\varepsilon^2}\right)\frac{1}{NT}\widetilde{Z}'\Omega_{NT}^{-1}(\bar{\boldsymbol{\beta}}^\star)\widetilde{Z} + \frac{1}{\sigma_{\varepsilon 0}^2}\frac{1}{NT}\widetilde{Z}'\Omega_{NT}^{-1}(\bar{\boldsymbol{\beta}}^\star)\left(\Omega_{NT}(\bar{\boldsymbol{\beta}}^\star) - \Omega_{0NT}\right)\Omega_{0NT}^{-1}\widetilde{Z} \\
=& \left(\frac{1}{\sigma_{\varepsilon 0}^2} - \frac{1}{\bar{\sigma}_\varepsilon^2}\right)\frac{1}{NT}\widetilde{Z}'\Omega_{NT}^{-1}(\bar{\boldsymbol{\beta}}^\star)\widetilde{Z} \\
&+\frac{1}{\sigma_{\varepsilon 0}^2}\frac{1}{NT}\widetilde{Z}'\Omega_{NT}^{-1}(\bar{\boldsymbol{\beta}}^\star)\left[\sum_{\ell=0}^{p}\frac{\partial}{\partial\beta_\ell^\star}\Omega_{NT}(\underline{\boldsymbol{\beta}}^\star)\left(\bar{\beta}_\ell^\star - \beta_{\ell 0}^\star\right)\right]\Omega_{0NT}^{-1}\widetilde{Z} \\
=& \frac{\bar{\sigma}_\varepsilon^2 - \sigma_{\varepsilon 0}^2}{\bar{\sigma}_\varepsilon^2\sigma_{\varepsilon 0}^2}\frac{1}{NT}\widetilde{Z}'\Omega_{NT}^{-1}(\bar{\boldsymbol{\beta}}^\star)\widetilde{Z} \\
&+\frac{1}{\sigma_{\varepsilon 0}^2}\sum_{\ell=0}^{p}\left(\bar{\beta}_\ell^\star - \beta_{\ell 0}^\star\right)\frac{1}{NT}\widetilde{Z}'\Omega_{NT}^{-1}(\bar{\boldsymbol{\beta}}^\star)\frac{\partial}{\partial\beta_\ell^\star}\Omega_{NT}(\underline{\boldsymbol{\beta}}^\star)\Omega_{0NT}^{-1}\widetilde{Z},
\end{aligned}
$$

where $\underline{\boldsymbol{\beta}}^\star$ lies between $\bar{\boldsymbol{\beta}}^\star$ and $\boldsymbol{\beta}_0^\star$. By Theorem 1, $\bar{\sigma}_\varepsilon^2 \to_p \sigma_{\varepsilon 0}^2$ and $\bar{\boldsymbol{\beta}}^\star \to_p \boldsymbol{\beta}_0^\star$. Moreover, by Lemmas 3–6, it can be verified that $\frac{1}{NT}\widetilde{Z}'\Omega_{NT}^{-1}(\bar{\boldsymbol{\beta}}^\star)\frac{\partial}{\partial\beta_\ell^\star}\Omega_{NT}(\underline{\boldsymbol{\beta}}^\star)\Omega_{0NT}^{-1}\widetilde{Z} = O_p(1)$ and $\frac{1}{NT}\widetilde{Z}'\Omega_{NT}^{-1}(\bar{\boldsymbol{\beta}}^\star)\widetilde{Z} = O_p(1)$. Hence, we have $\frac{1}{NT}\frac{\partial^2 \ln \mathcal{L}_{NT}(\bar{\boldsymbol{\zeta}})}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}'} - \frac{1}{NT}\frac{\partial^2 \ln \mathcal{L}_{NT}(\boldsymbol{\zeta}_0)}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}'} = o_p(1)$. Furthermore, we can show that other terms of $\frac{1}{NT}\frac{\partial^2 \ln \mathcal{L}_{NT}(\bar{\boldsymbol{\zeta}})}{\partial\boldsymbol{\zeta}\partial\boldsymbol{\zeta}'} - \frac{1}{NT}\frac{\partial^2 \ln \mathcal{L}_{NT}(\boldsymbol{\zeta}_0)}{\partial\boldsymbol{\zeta}\partial\boldsymbol{\zeta}'}$ converges to zero in probability similarly, and then (B2) holds.

The proof of this theorem is then accomplished.

S5.3 Proof of Theorem 3

Recall that $\widehat{\boldsymbol{\vartheta}} = \frac{1}{T}\sum_{t=1}^{T}\left[B_N(\widehat{\lambda})\boldsymbol{y}_t - \widehat{\alpha}\boldsymbol{y}_{t-1} - Z_t\widehat{\boldsymbol{\gamma}} - X\widehat{\boldsymbol{\psi}}\right]$, so it holds that

$$
\widehat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta} = \frac{1}{T}\sum_{t=1}^{T}\left[-(\widehat{\lambda} - \lambda_0)W_N\boldsymbol{y}_t - (\widehat{\alpha} - \alpha_0)\boldsymbol{y}_{t-1} - Z_t(\widehat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0) - X(\widehat{\boldsymbol{\psi}} - \boldsymbol{\psi}_0) + \boldsymbol{\varepsilon}_t\right]. \quad \text{(S5.11)}
$$

We first show Theorem 3(i), that is $\widehat{\vartheta}_i - \vartheta_i = o_p(1)$ for $1 \le i \le N$. By (S5.11), we have

$$
\widehat{\vartheta}_i - \vartheta_i = -(\widehat{\lambda} - \lambda_0)\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{w}_{i,N}\boldsymbol{y}_t - (\widehat{\alpha} - \alpha_0)\frac{1}{T}\sum_{t=1}^{T}y_{i,t-1} - \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{z}_{it}'(\widehat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0)
$$

$$-\boldsymbol{x}_i'(\widehat{\boldsymbol{\psi}}-\boldsymbol{\psi}_0)+\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it}, \tag{S5.12}$$

where $\boldsymbol{w}_{i,N}$ is the ith row of W_N. Note that $\widehat{\boldsymbol{\zeta}}-\boldsymbol{\zeta}_0=o_p(1)$ by Theorem 1, $\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{z}_{it}=o_p(1)$, $\frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it}=o_p(1)$ and $x_i=O(1)$ by Assumption 3.6. To prove that $\widehat{\vartheta}_i-\vartheta_i=o_p(1)$, it remains to show that $\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{w}_{i,N}\boldsymbol{y}_t=O_p(1)$ and $\frac{1}{T}\sum_{t=1}^{T}y_{i,t-1}=O_p(1)$. Since the row sums of W_N are uniformly bounded, we only need to show $\frac{1}{T}\sum_{t=1}^{T}y_{it}=O_p(1)$. By (S3.2), we have

$$\boldsymbol{y}_t=\alpha_0^t(B_{0N}^{-1})^t\boldsymbol{y}_0+\sum_{j=1}^{t}\left[\alpha_0^{j-1}(B_{0N}^{-1})^jZ_{t-j+1}\right]\boldsymbol{\gamma}_0+\sum_{j=1}^{t}\left[\alpha_0^{j-1}(B_{0N}^{-1})^j\right]\boldsymbol{\theta}+\sum_{j=1}^{t}\left[\alpha_0^{j-1}(B_{0N}^{-1})^j\boldsymbol{\varepsilon}_{t-j+1}\right].$$

By Assumption 3.5, we have $E\left(\sum_{t=1}^{T}y_{it}\right)=O(T)$ and $\mathrm{var}\left(\sum_{t=1}^{T}y_{it}\right)=O(T)$, then it follows that $\frac{1}{T}\sum_{t=1}^{T}(y_{it}-Ey_{it})=o_p(1)$ by Chebyshev inequality and thus $\frac{1}{T}\sum_{t=1}^{T}y_{it}=O_p(1)$. Therefore, $\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{w}_{i,N}\boldsymbol{y}_t=O_p(1)$ and $\frac{1}{T}\sum_{t=1}^{T}y_{i,t-1}=O_p(1)$ holds. This completes the proof of Theorem 3(i).

Next we consider Theorem 3(ii). By (S5.12), it follows that

$$\begin{aligned}
&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\widehat{\vartheta}_i-\vartheta_i)\\
&=\frac{1}{\sqrt{N}T}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[-(\widehat{\lambda}-\lambda_0)\boldsymbol{w}_{i,N}\boldsymbol{y}_t-(\widehat{\alpha}-\alpha_0)y_{i,t-1}-\boldsymbol{z}_{it}'(\widehat{\boldsymbol{\gamma}}-\boldsymbol{\gamma}_0)-\boldsymbol{x}_i'(\widehat{\boldsymbol{\psi}}-\boldsymbol{\psi}_0)+\varepsilon_{it}\right]\\
&=-\sqrt{NT}(\widehat{\lambda}-\lambda_0)\frac{1}{NT^{3/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\boldsymbol{w}_{i,N}\boldsymbol{y}_t-\sqrt{NT}(\widehat{\alpha}-\alpha_0)\frac{1}{NT^{3/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}y_{i,t-1}\\
&\quad-\frac{1}{NT^{3/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\boldsymbol{z}_{it}'\sqrt{NT}(\widehat{\boldsymbol{\gamma}}-\boldsymbol{\gamma}_0)-\frac{1}{NT^{1/2}}\sum_{i=1}^{N}\boldsymbol{x}_i'\sqrt{NT}(\widehat{\boldsymbol{\psi}}-\boldsymbol{\psi}_0)+\frac{1}{N^{1/2}T}\sum_{i=1}^{N}\sum_{t=1}^{T}\varepsilon_{it}.
\end{aligned}$$

Notice that $\sqrt{NT}(\widehat{\boldsymbol{\zeta}}-\boldsymbol{\zeta}_0)=O_p(1)$ by Theorem 2, $\frac{1}{NT^{3/2}}\sum_{i=1}^{N}\sum_{t=1}^{T}\boldsymbol{z}_{it}=o_p(1)$, $\frac{1}{NT^{1/2}}\sum_{i=1}^{N}\boldsymbol{x}_i=o_p(1)$ and $\frac{1}{N^{1/2}T}\sum_{i=1}^{N}\sum_{t=1}^{T}\varepsilon_{it}=o_p(1)$ by Assumption 3.6. Moreover, by Assumption 3.5, it can be verified that

$$E\left(\sum_{i=1}^{N}\sum_{t=1}^{T}y_{it}\right)=O(NT)\quad\text{and}\quad\mathrm{var}\left(\sum_{i=1}^{N}\sum_{t=1}^{T}y_{it}\right)=O(N^2T),$$

then it follows that $\frac{1}{NT^{3/2}}\sum_{i=1}^{N}\sum_{t=1}^{T} y_{i,t} = o_p(1)$ by Chebyshev inequality. Since the row sums of W_N are uniformly bounded, we obtain that $\frac{1}{NT^{3/2}}\sum_{i=1}^{N}\sum_{t=1}^{T} \boldsymbol{w}_{i,N}\boldsymbol{y}_t = o_p(1)$ and $\frac{1}{NT^{3/2}}\sum_{i=1}^{N}\sum_{t=1}^{T} y_{i,t-1} = o_p(1)$ holds. Thus $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\widehat{\vartheta}_i - \vartheta_i) = o_p(1)$, that is Theorem 3$(ii)$ holds.

The proof of Theorem 3 is accomplished.

S5.4 Proof of Theorem 4

Define

$$Q_N(\tau,\boldsymbol{\varphi}) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\boldsymbol{x}'_{ai}\widetilde{\boldsymbol{\beta}}_c}\rho_\tau(\widehat{\vartheta}_i - \boldsymbol{x}'_{ai}\boldsymbol{\varphi}) \text{ and } Q(\tau,\boldsymbol{\varphi}) = \frac{1}{N}\sum_{i=1}^{N}E\frac{1}{\boldsymbol{x}'_{ai}\boldsymbol{\beta}_0}\rho_\tau(\vartheta_i - \boldsymbol{x}'_{ai}\boldsymbol{\varphi}).$$

We first show that $Q_N(\tau,\boldsymbol{\varphi}) - Q(\tau,\boldsymbol{\varphi}) = o_p(1)$ holds uniformly on $\mathcal{T}\times\Phi$. To achieve this, it suffices to show that, as $N\to\infty$,

(C1) $\sup_{(\tau,\boldsymbol{\varphi})}\left|\frac{1}{N}\sum_{i=1}^{N}E\frac{1}{\boldsymbol{x}'_{ai}\boldsymbol{\beta}_0}\rho_\tau(\widehat{\vartheta}_i - \boldsymbol{x}'_{ai}\boldsymbol{\varphi}) - \frac{1}{N}\sum_{i=1}^{N}E\frac{1}{\boldsymbol{x}'_{ai}\boldsymbol{\beta}_0}\rho_\tau(\vartheta_i - \boldsymbol{x}'_{ai}\boldsymbol{\varphi})\right| \to 0;$

(C2) $\sup_{(\tau,\boldsymbol{\varphi})}\left|\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\boldsymbol{x}'_{ai}\boldsymbol{\beta}_0}\rho_\tau(\widehat{\vartheta}_i - \boldsymbol{x}'_{ai}\boldsymbol{\varphi}) - \frac{1}{N}\sum_{i=1}^{N}E\frac{1}{\boldsymbol{x}'_{ai}\boldsymbol{\beta}_0}\rho_\tau(\widehat{\vartheta}_i - \boldsymbol{x}'_{ai}\boldsymbol{\varphi})\right| = o_p(1);$

(C3) $\sup_{(\tau,\boldsymbol{\varphi})}\left|\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\boldsymbol{x}'_{ai}\widetilde{\boldsymbol{\beta}}_c}\rho_\tau(\widehat{\vartheta}_i - \boldsymbol{x}'_{ai}\boldsymbol{\varphi}) - \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\boldsymbol{x}'_{ai}\boldsymbol{\beta}_0}\rho_\tau(\widehat{\vartheta}_i - \boldsymbol{x}'_{ai}\boldsymbol{\varphi})\right| = o_p(1).$

(Show (C1)) By Theorem 3, we obtain that $\widehat{\vartheta}_i - \vartheta_i = o_p(1)$ and hence $\widehat{\vartheta}_i \to_d \vartheta_i$. Then by Lemma 2.2 of Van der Vaart (2000), it holds that

$$\left|E\frac{1}{\boldsymbol{x}'_{ai}\boldsymbol{\beta}_0}\rho_\tau(\widehat{\vartheta}_i - \boldsymbol{x}'_{ai}\boldsymbol{\varphi}) - E\frac{1}{\boldsymbol{x}'_{ai}\boldsymbol{\beta}_0}\rho_\tau(\vartheta_i - \boldsymbol{x}'_{ai}\boldsymbol{\varphi})\right| \to 0$$

as $N\to\infty$, because $\frac{1}{\boldsymbol{x}'_{a}\boldsymbol{\beta}_0}\rho_\tau(\vartheta - \boldsymbol{x}'_{a}\boldsymbol{\varphi})$ is a bounded (which is due to the compactness of Φ and $\mathscr{B}$) and continuous function. Furthermore, since the set $\mathcal{T}\times\Phi$ is compact and $E\frac{1}{\boldsymbol{x}'_{ai}\boldsymbol{\beta}_0}\rho_\tau(\widehat{\vartheta}_i - \boldsymbol{x}'_{ai}\boldsymbol{\varphi}) - E\frac{1}{\boldsymbol{x}'_{ai}\boldsymbol{\beta}_0}\rho_\tau(\vartheta_i - \boldsymbol{x}'_{ai}\boldsymbol{\varphi})$ is continuous on $\mathcal{T}\times\Phi$, we have that

$$\sup_{(\tau,\boldsymbol{\varphi})}\left|E\frac{1}{\boldsymbol{x}'_{ai}\boldsymbol{\beta}_0}\rho_\tau(\widehat{\vartheta}_i - \boldsymbol{x}'_{ai}\boldsymbol{\varphi}) - E\frac{1}{\boldsymbol{x}'_{ai}\boldsymbol{\beta}_0}\rho_\tau(\vartheta_i - \boldsymbol{x}'_{ai}\boldsymbol{\varphi})\right| \to 0.$$

It follows that (C1) holds.

(Show (C2)) Let $\mathcal{H} = \left\{(\tau, \boldsymbol{\varphi}) \mapsto \frac{1}{\boldsymbol{x}_a'\boldsymbol{\beta}_0}\rho_\tau(\vartheta - \boldsymbol{x}_a'\boldsymbol{\varphi})\right\}$. Since $\mathcal{T} \times \Phi$ is compact and $\frac{1}{\boldsymbol{x}_a'\boldsymbol{\beta}_0}\rho_\tau(\vartheta - \boldsymbol{x}_a'\boldsymbol{\varphi})$ is Lipschitz over $\mathcal{T} \times \Phi$, $\mathcal{H}$ is Donsker by Theorem 2.10.6 of Van Der Vaart and Wellner (1996). Following Chernozhukov and Hansen (2006) and Canay (2011), donskerness implies a uniform law of large number, that is

$$\sup_{(\tau,\boldsymbol{\varphi})} \left| \frac{1}{N}\sum_{i=1}^{N}\left[\frac{1}{\boldsymbol{x}_{ai}'\boldsymbol{\beta}_0}\rho_\tau(\vartheta_i - \boldsymbol{x}_{ai}'\boldsymbol{\varphi}) - E\frac{1}{\boldsymbol{x}_{ai}'\boldsymbol{\beta}_0}\rho_\tau(\vartheta_i - \boldsymbol{x}_{ai}'\boldsymbol{\varphi})\right]\right| = o_p(1),$$

which gives

$$\sup_{(\tau,\boldsymbol{\varphi})} \left| \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\boldsymbol{x}_{ai}'\boldsymbol{\beta}_0}\rho_\tau(\widehat{\vartheta}_i - \boldsymbol{x}_{ai}'\boldsymbol{\varphi}) - \frac{1}{N}\sum_{i=1}^{N}E\frac{1}{\boldsymbol{x}_{ai}'\boldsymbol{\beta}_0}\rho_\tau(\widehat{\vartheta}_i - \boldsymbol{x}_{ai}'\boldsymbol{\varphi})\right| = o_p(1),$$

then (C2) holds; see also Van der Vaart (2000).

(Show (C3)) Note that $\widetilde{\boldsymbol{\beta}}_c - \boldsymbol{\beta}_0 = o_p(1)$, so we have that

$$\begin{aligned}
&\left| \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\boldsymbol{x}_{ai}'\widetilde{\boldsymbol{\beta}}_c}\rho_\tau(\widehat{\vartheta}_i - \boldsymbol{x}_{ai}'\boldsymbol{\varphi}) - \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\boldsymbol{x}_{ai}'\boldsymbol{\beta}_0}\rho_\tau(\widehat{\vartheta}_i - \boldsymbol{x}_{ai}'\boldsymbol{\varphi})\right| \\
= &\left| \frac{1}{N}\sum_{i=1}^{N}\frac{\boldsymbol{x}_{ai}'(\boldsymbol{\beta}_0 - \widetilde{\boldsymbol{\beta}}_c)}{(\boldsymbol{x}_{ai}'\widetilde{\boldsymbol{\beta}}_c)(\boldsymbol{x}_{ai}'\boldsymbol{\beta}_0)}\rho_\tau(\widehat{\vartheta}_i - \boldsymbol{x}_{ai}'\boldsymbol{\varphi})\right| = o_p(1).
\end{aligned}$$

Furthermore, $\mathcal{T} \times \Phi$ is compact and $\frac{1}{N}\sum_{i=1}^{N}\frac{1}{\boldsymbol{x}_{ai}'\widetilde{\boldsymbol{\beta}}_c}\rho_\tau(\widehat{\vartheta}_i - \boldsymbol{x}_{ai}'\boldsymbol{\varphi}) - \frac{1}{N}\sum_{i=1}^{N}\frac{1}{\boldsymbol{x}_{ai}'\boldsymbol{\beta}_0}\rho_\tau(\widehat{\vartheta}_i - \boldsymbol{x}_{ai}'\boldsymbol{\varphi})$ is stochastically equicontinuous on $\mathcal{T} \times \Phi$ (Chernozhukov and Hansen, 2006), so (C3) holds.

By (C1)–(C3), we obtain that $Q_N(\tau, \boldsymbol{\varphi}) - Q(\tau, \boldsymbol{\varphi}) = o_p(1)$ holds uniformly on $\mathcal{T} \times \Phi$. This together with $\boldsymbol{\varphi}_0(\tau) = \underset{\boldsymbol{\varphi}\in\Phi}{\operatorname{argmin}}\, Q(\tau, \boldsymbol{\varphi})$, implies that $\sup_{\tau\in\mathcal{T}} \|\widehat{\boldsymbol{\varphi}}(\tau) - \boldsymbol{\varphi}_0(\tau)\| = o_p(1)$ by Lemma B.1 of Chernozhukov and Hansen (2006). As a result, we complete the proof of Theorem 4.

S5.5 Proof of Theorem 5

Following Canay (2011) and Chernozhukov and Hansen (2006), by the computational properties of standard quantile regression, it holds that

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{1}{\boldsymbol{x}_{ai}'\widetilde{\boldsymbol{\beta}}_c}\Psi_{\bullet}(\widehat{\vartheta}_i-\boldsymbol{x}_{ai}'\widehat{\boldsymbol{\varphi}}(\cdot))\boldsymbol{x}_{ai}=O_p\left(1/\sqrt{N}\right)\quad\text{in}\quad\ell^{\infty}(\mathcal{T}). \tag{S5.13}$$

Denote

$$F_1\triangleq\mathbb{G}_N\left(\frac{1}{\boldsymbol{x}_{ai}'\boldsymbol{\beta}_0}\Psi_{\bullet}(\vartheta_i-\boldsymbol{x}_{ai}'\widehat{\boldsymbol{\varphi}}(\cdot)+\widehat{r}_i)\boldsymbol{x}_{ai}\right),$$
$$F_2\triangleq\frac{1}{\sqrt{N}}\sum_{i=1}^{N}E\frac{1}{\boldsymbol{x}_{ai}'\boldsymbol{\beta}_0}\Psi_{\bullet}(\vartheta_i-\boldsymbol{x}_{ai}'\widehat{\boldsymbol{\varphi}}(\cdot)+\widehat{r}_i)\boldsymbol{x}_{ai},$$

where $\widehat{r}_i\equiv\widehat{\vartheta}_i-\vartheta_i=o_p(1)$ for $1\le i\le N$ by Theorem 3(i). By (S5.13) and $\widetilde{\boldsymbol{\beta}}_c-\boldsymbol{\beta}_0=o_p(1)$, we obtain that

$$\begin{aligned}F_1+F_2&=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{1}{\boldsymbol{x}_{ai}'\boldsymbol{\beta}_0}\Psi_{\bullet}(\vartheta_i-\boldsymbol{x}_{ai}'\widehat{\boldsymbol{\varphi}}(\cdot)+\widehat{r}_i)\boldsymbol{x}_{ai}\\&=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{1}{\boldsymbol{x}_{ai}'\widetilde{\boldsymbol{\beta}}_c}\Psi_{\bullet}(\widehat{\vartheta}_i-\boldsymbol{x}_{ai}'\widehat{\boldsymbol{\varphi}}(\cdot))\boldsymbol{x}_{ai}+o_p(1)\\&=o_p(1)\quad\text{in}\quad\ell^{\infty}(\mathcal{T}).\end{aligned} \tag{S5.14}$$

(For F_1) Note that $\sup_{\tau\in\mathcal{T}}\|\widehat{\boldsymbol{\varphi}}(\tau)-\boldsymbol{\varphi}_0(\tau)\|=o_p(1)$ by Theorem 4, so by Lemma 10(ii),

$$\begin{aligned}F_1&=\mathbb{G}_N\left(\frac{1}{\boldsymbol{x}_{ai}'\boldsymbol{\beta}_0}\Psi_{\bullet}(\vartheta_i-\boldsymbol{x}_{ai}'\widehat{\boldsymbol{\varphi}}(\cdot)+\widehat{r}_i)\boldsymbol{x}_{ai}\right)\\&=\mathbb{G}_N\left(\frac{1}{\boldsymbol{x}_{ai}'\boldsymbol{\beta}_0}\Psi_{\bullet}(\vartheta_i-\boldsymbol{x}_{ai}'\boldsymbol{\varphi}_0(\cdot))\boldsymbol{x}_{ai}\right)+o_p(1)\\&=\mathbb{G}_N\left(\frac{1}{\boldsymbol{x}_{ai}'\boldsymbol{\beta}_0}\Psi_{\bullet}(\eta_i-Q_{\eta}(\cdot))\boldsymbol{x}_{ai}\right)+o_p(1)\quad\text{in}\quad\ell^{\infty}(\mathcal{T}).\end{aligned} \tag{S5.15}$$

(For F_2) Following Canay (2011) and Chernozhukov and Hansen (2006), we have

$$F_2=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}E\frac{1}{\boldsymbol{x}_{ai}'\boldsymbol{\beta}_0}\Psi_{\bullet}(\vartheta_i-\boldsymbol{x}_{ai}'\widehat{\boldsymbol{\varphi}}(\cdot)+\widehat{r}_i)\boldsymbol{x}_{ai}$$

$$= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \Bigg[E \frac{1}{\boldsymbol{x}_{ai}' \boldsymbol{\beta}_0} \Psi_{\bullet}(\vartheta_i - \boldsymbol{x}_{ai}' \boldsymbol{\varphi}_0(\cdot) + 0) \boldsymbol{x}_{ai}$$

$$+ J_{1i}(\cdot, \boldsymbol{\varphi}^\star(\cdot), r_i^\star)(\widehat{\boldsymbol{\varphi}}(\cdot) - \boldsymbol{\varphi}_0(\cdot)) + \boldsymbol{j}_{2i}(\cdot, \boldsymbol{\varphi}^\star(\cdot), r_i^\star)(\widehat{r}_i - 0) \Bigg],$$

where $(\boldsymbol{\varphi}^\star(\tau), r_1^\star, \ldots, r_N^\star)$ lies between $(\boldsymbol{\varphi}_0(\tau), 0, \ldots, 0)$ and $(\widehat{\boldsymbol{\varphi}}(\tau), \widehat{r}_1, \ldots, \widehat{r}_N)$, and

$$J_{1i}(\tau, \boldsymbol{\varphi}, r_i) = \frac{\partial}{\partial \boldsymbol{\varphi}'} E \frac{1}{\boldsymbol{x}_{ai}' \boldsymbol{\beta}_0} \Psi_\tau(\vartheta_i - \boldsymbol{x}_{ai}' \boldsymbol{\varphi} + r_i) \boldsymbol{x}_{ai},$$

$$\boldsymbol{j}_{2i}(\tau, \boldsymbol{\varphi}, r_i) = \frac{\partial}{\partial r_i} E \frac{1}{\boldsymbol{x}_{ai}' \boldsymbol{\beta}_0} \Psi_\tau(\vartheta_i - \boldsymbol{x}_{ai}' \boldsymbol{\varphi} + r_i) \boldsymbol{x}_{ai}.$$

By Theorems 3(i) and 4, $\boldsymbol{\varphi}^\star(\cdot) - \boldsymbol{\varphi}_0(\cdot) = o_p(1)$ in $\ell^\infty(\mathcal{T})$, and $r_i^\star = o_p(1)$ for $1 \le i \le N$. Then by Assumption 3.9, we have that

$$F_2 = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \Bigg[E \frac{1}{\boldsymbol{x}_{ai}' \boldsymbol{\beta}_0} \Psi_{\bullet}(\vartheta_i - \boldsymbol{x}_{ai}' \boldsymbol{\varphi}_0(\cdot)) \boldsymbol{x}_{ai}$$

$$+ J_{1i}(\cdot, \boldsymbol{\varphi}_0(\cdot), 0)(\widehat{\boldsymbol{\varphi}}(\cdot) - \boldsymbol{\varphi}_0(\cdot)) + \boldsymbol{j}_{2i}(\cdot, \boldsymbol{\varphi}_0(\cdot), 0) \widehat{r}_i \Bigg] + o_p(1)$$

$$= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \Bigg[E \frac{1}{\boldsymbol{x}_{ai}' \boldsymbol{\beta}_0} \Psi_{\bullet}(\eta_i - Q_\eta(\cdot)) \boldsymbol{x}_{ai}$$

$$+ J_{1i}(\cdot, \boldsymbol{\varphi}_0(\cdot), 0)(\widehat{\boldsymbol{\varphi}}(\cdot) - \boldsymbol{\varphi}_0(\cdot)) + \boldsymbol{j}_{2i}(\cdot, \boldsymbol{\varphi}_0(\cdot), 0) \widehat{r}_i \Bigg] + o_p(1),$$

in $\ell^\infty(\mathcal{T})$. Furthermore, we have

$$J_{1i}(\tau, \boldsymbol{\varphi}_0(\tau), 0) = -\frac{\boldsymbol{x}_{ai}' \boldsymbol{x}_{ai}}{(\boldsymbol{x}_{ai}' \boldsymbol{\beta}_0)^2} f_\eta(Q_\eta(\tau)) \quad \text{and} \quad \boldsymbol{j}_{2i}(\tau, \boldsymbol{\varphi}_0(\tau), 0) = \frac{\boldsymbol{x}_{ai}}{(\boldsymbol{x}_{ai}' \boldsymbol{\beta}_0)^2} f_\eta(Q_\eta(\tau)).$$

These together with $E\Psi_\tau(\eta_i - Q_\eta(\tau)) = 0$ and $\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{\boldsymbol{x}_{ai}}{(\boldsymbol{x}_{ai}' \boldsymbol{\beta}_0)^2} (\widehat{\vartheta}_i - \vartheta_i) = o_p(1)$ by Theorem 3(ii), imply that

$$F_2 = -D_{1N}(\cdot) \sqrt{N} (\widehat{\boldsymbol{\varphi}}(\cdot) - \boldsymbol{\varphi}_0(\cdot)) + \Xi(\cdot) \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{\boldsymbol{x}_{ai}}{(\boldsymbol{x}_{ai}' \boldsymbol{\beta}_0)^2} (\widehat{\vartheta}_i - \vartheta_i) + o_p(1)$$

$$= -D_{1N}(\cdot) \sqrt{N} (\widehat{\boldsymbol{\varphi}}(\cdot) - \boldsymbol{\varphi}_0(\cdot)) + o_p(1) \quad \text{in} \quad \ell^\infty(\mathcal{T}). \tag{S5.16}$$

By (S5.14)–(S5.16), it follows that

$$\mathbb{G}_N\left(\frac{\boldsymbol{x}_{ai}}{\boldsymbol{x}'_{ai}\boldsymbol{\beta}_0}\Psi_{\bullet}(\eta_i - Q_\eta(\cdot))\right) - D_{1N}(\cdot)\sqrt{N}(\widehat{\boldsymbol{\varphi}}(\cdot) - \boldsymbol{\varphi}_0(\cdot)) = o_p(1) \quad \text{in} \quad \ell^\infty(\mathcal{T}). \tag{S5.17}$$

This together with Lemma 10, implies that

$$\sqrt{N}\left(\widehat{\boldsymbol{\varphi}}(\cdot) - \boldsymbol{\varphi}_0(\cdot)\right) = D_{1N}^{-1}(\cdot)\mathbb{G}_N\left(\frac{\boldsymbol{x}_{ai}}{\boldsymbol{x}'_{ai}\boldsymbol{\beta}_0}\Psi_{\bullet}\left(\eta_i - Q_\eta(\cdot)\right)\right) + o_p(1) \rightsquigarrow \mathbb{G}(\cdot) \quad \text{in} \quad \ell^\infty(\mathcal{T}),$$

where $\mathbb{G}(\cdot)$ is a Gaussian process with zero mean and covariance kernel $\Xi^{-1}(\tau)S(\tau,\tau')\Xi^{-1}(\tau')D_0^{-1}$. This completes the proof of Theorem 5.

S5.6 Proof of Theorem 6

Recall that $\widehat{\boldsymbol{\beta}}(\boldsymbol{\pi}_K) = \sum_{k=1}^{K}\pi_k\widehat{\boldsymbol{\varphi}}(\tau_k)$ and $\sum_{k=1}^{K}\pi_k Q_\eta(\tau_k) = 1$, then by Theorem 5, we have

$$\begin{aligned}
\sqrt{N}\left(\widehat{\boldsymbol{\beta}}(\boldsymbol{\pi}_K) - \boldsymbol{\beta}_0\right) &= \sqrt{N}\left(\sum_{k=1}^{K}\pi_k\widehat{\boldsymbol{\varphi}}(\tau_k) - \sum_{k=1}^{K}\pi_k Q_\eta(\tau_k)\boldsymbol{\beta}_0\right) \\
&= \sqrt{N}\left(\sum_{k=1}^{K}\pi_k\widehat{\boldsymbol{\varphi}}(\tau_k) - \sum_{k=1}^{K}\pi_k\boldsymbol{\varphi}_0(\tau_k)\right) \\
&= \sum_{k=1}^{K}\pi_k\sqrt{N}(\widehat{\boldsymbol{\varphi}}(\tau_k) - \boldsymbol{\varphi}_0(\tau_k)) \to_d N\left(0, \Upsilon(\boldsymbol{\pi}_K)\right),
\end{aligned}$$

where the covariance matrix is in form of

$$\Upsilon(\boldsymbol{\pi}_K) = \sum_{k_1=1}^{K}\sum_{k_2=1}^{K}\pi_{k_1}\pi_{k_2}\Xi^{-1}(\tau_{k_1})S(\tau_{k_1},\tau_{k_2})\Xi^{-1}(\tau_{k_2})D_0^{-1} = \boldsymbol{\pi}'_K H\boldsymbol{\pi}_K D_0^{-1}.$$

As for Theorem 5 of Zhao and Xiao (2014), the optimal weight follows from Lagrange multiplier method, that is

$$\boldsymbol{\pi}_{opt,K} = \underset{\boldsymbol{\pi},\,\text{s.t.}\,\boldsymbol{\pi}'\boldsymbol{q}=1}{\operatorname{argmin}}\ \Upsilon(\boldsymbol{\pi}_K) = \frac{H^{-1}\boldsymbol{q}}{\boldsymbol{q}'H^{-1}\boldsymbol{q}}.$$

It follows that the asymptotic covariance of the optimal WQAE $\widehat{\boldsymbol{\beta}}(\boldsymbol{\pi}_{opt,K})$ is $\Upsilon(\boldsymbol{\pi}_{opt,K}) = (\boldsymbol{q}'H^{-1}\boldsymbol{q})^{-1}D_0^{-1}$. Hence, the proof of this theorem is accomplished.

S5.7 Proof of Proposition 1

Let $\{\boldsymbol{\xi}_t\}$ be a vector sequence of random variables with

$$\boldsymbol{\xi}_t = \sum_{m=0}^{\infty} \alpha^m (B_N^{-1})^{m+1} (\boldsymbol{\theta} + Z_{t-m}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}_{t-m}), \tag{S5.18}$$

where $B_N = B_N(\lambda)$. Recall that $\{\varepsilon_{it}\}$ are *i.i.d.* across i and t, we have $E(\boldsymbol{\varepsilon}_t) = \mathbf{0}$, $\text{cov}(\boldsymbol{\varepsilon}_t) = \sigma_\varepsilon^2 I_N$ and $\text{cov}(\boldsymbol{\varepsilon}_t, \boldsymbol{\varepsilon}_{t-h}) = 0$ for $h \neq 0$. Moreover, since $\{\boldsymbol{z}_{it}\}$ is covariance stationary by assumption, we have $E(\boldsymbol{z}_{it}) = \boldsymbol{\mu}_z$ for a constant vector $\boldsymbol{\mu}_z$ and $\text{cov}(\boldsymbol{z}_{it}, \boldsymbol{z}_{i,t-h})$ is independent of t. Without loss of generality, we assume $\boldsymbol{\mu}_z = \mathbf{0}$. It follows that $E(Z_t\boldsymbol{\gamma}) = \mathbf{0}$ and $\text{cov}(Z_t\boldsymbol{\gamma}, Z_{t-h}\boldsymbol{\gamma}) = E(Z_t\boldsymbol{\gamma}\boldsymbol{\gamma}' Z_{t-h}')$ is independent of t. Since $\left|\alpha B_N^{-1}\right| < 1$ by assumption, $\boldsymbol{\theta} = \boldsymbol{\vartheta} + X\boldsymbol{\psi}$ and $E(\boldsymbol{\vartheta}) = \mathbf{0}$, it holds that

$$\begin{aligned} E\left(B_N\boldsymbol{\xi}_t\right) &= E\left[\sum_{m=0}^{\infty} \left(\alpha B_N^{-1}\right)^m \left(\boldsymbol{\vartheta} + X\boldsymbol{\psi} + Z_{t-m}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}_{t-m}\right)\right] \\ &= \sum_{m=0}^{\infty} \left(\alpha B_N^{-1}\right)^m X\boldsymbol{\psi} = \left(I_N - \alpha B_N^{-1}\right)^{-1} X\boldsymbol{\psi}, \end{aligned}$$

and by independence of $\boldsymbol{\vartheta}$, $\{\boldsymbol{\varepsilon}_t\}$ and $\{\boldsymbol{z}_{it}\}$,

$$\begin{aligned} &\text{cov}\left(B_N\boldsymbol{\xi}_t, B_N\boldsymbol{\xi}_{t-h}\right) = E\left[B_N\boldsymbol{\xi}_t - E\left(B_N\boldsymbol{\xi}_t\right)\right]\left[B_N\boldsymbol{\xi}_{t-h} - E\left(B_N\boldsymbol{\xi}_{t-h}\right)\right]' \\ =& E\left\{\left[\sum_{m=0}^{\infty}\left(\alpha B_N^{-1}\right)^m\left(\boldsymbol{\vartheta} + Z_{t-m}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}_{t-m}\right)\right]\left[\sum_{k=0}^{\infty}\left(\alpha B_N^{-1}\right)^k\left(\boldsymbol{\vartheta} + Z_{t-h-k}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}_{t-h-k}\right)\right]'\right\} \\ =& \sum_{m,k=0}^{\infty}\left(\alpha B_N^{-1}\right)^m E\left(\boldsymbol{\vartheta}\boldsymbol{\vartheta}'\right)\left[\left(\alpha B_N^{-1}\right)^k\right]' + \sum_{m,k=0}^{\infty}\left(\alpha B_N^{-1}\right)^m E\left(\boldsymbol{\varepsilon}_{t-m}\boldsymbol{\varepsilon}_{t-h-k}'\right)\left[\left(\alpha B_N^{-1}\right)^k\right]' \\ &+ \sum_{m,k=0}^{\infty}\left(\alpha B_N^{-1}\right)^m E\left(Z_{t-m}\boldsymbol{\gamma}\boldsymbol{\gamma}' Z_{t-h-k}'\right)\left[\left(\alpha B_N^{-1}\right)^k\right]', \end{aligned}$$

where $E(B_N\boldsymbol{\xi}_t)$ is a vector of constants and $\text{cov}(B_N\boldsymbol{\xi}_t, B_N\boldsymbol{\xi}_{t-h})$ is independent of t. Thus $\{\boldsymbol{\xi}_t\}$ is covariance stationary. Note that models (2.1) and (2.2) can be rewritten as

$$B_N\boldsymbol{y}_t = \boldsymbol{\theta} + \alpha\boldsymbol{y}_{t-1} + Z_t\boldsymbol{\gamma} + \boldsymbol{\varepsilon}_t. \tag{S5.19}$$

It can be shown that $\{\boldsymbol{\xi}_t\}$ satisfies the recursive equation

$$B_N\boldsymbol{\xi}_t = \boldsymbol{\theta} + \alpha\boldsymbol{\xi}_{t-1} + Z_t\boldsymbol{\gamma} + \boldsymbol{\varepsilon}_t.$$

Hence we prove the existence of a covariance stationary solution to models (2.1) and (2.2) by setting $\boldsymbol{y}_t = \boldsymbol{\xi}_t$.

Next suppose that $\{\boldsymbol{y}_t\}$ is a covariance stationary solution to models (2.1) and (2.2). Then, for any integer $h > 0$, by successively substituting $\boldsymbol{y}_{t-1}$'s in (S5.19) h times, we have

$$B_N \boldsymbol{y}_t = \sum_{m=0}^{h} \left(\alpha B_N^{-1}\right)^m \left(\boldsymbol{\theta} + Z_{t-m}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}_{t-m}\right) + R_{t,h},$$

where $R_{t,h} = (\alpha B_N^{-1})^h \alpha \boldsymbol{y}_{t-h-1} \to 0$ almost surely as $h \to \infty$, since $\left|\alpha B_N^{-1}\right| < 1$ by assumption. Thus $B_N \boldsymbol{y}_t = B_N \boldsymbol{\xi}_t$ and then $\boldsymbol{y}_t = \boldsymbol{\xi}_t$ almost surely. This completes the proof of Proposition 1.

S5.8 Proof of Proposition 2

Recall that $\boldsymbol{Y} = S_0^{-1}(\widetilde{Z}\boldsymbol{\phi}_0 + \boldsymbol{V}_{NT})$ and $\widetilde{Z} = (\boldsymbol{Y}_{-1}, Z, (\boldsymbol{\iota}_T \otimes I_N)X)$. Since $M_3 = M_2(\boldsymbol{\iota}_T \otimes I_N)$, $\boldsymbol{Y}_{-1}$ can be rewritten as $M_1\boldsymbol{y}_0 + M_2 Z\boldsymbol{\gamma}_0 + M_3 X\boldsymbol{\psi}_0 + M_2\boldsymbol{V}_{NT}$ by (S3.3). Then elements of the score vector $\boldsymbol{s}(\boldsymbol{Y}_{-1}, \boldsymbol{V}_{NT}, \boldsymbol{\zeta}_0) = \partial \ln \mathcal{L}_{NT}(\boldsymbol{\zeta}_0)/\partial\boldsymbol{\zeta}$ can be expressed in terms of $\boldsymbol{\zeta}_0$ and $\boldsymbol{V}_{NT}$ as follows:

$$\boldsymbol{s}(\boldsymbol{Y}_{-1}, \boldsymbol{V}_{NT}, \boldsymbol{\zeta}_0) = (\dot{\boldsymbol{d}}_\phi', \dot{d}_{\sigma_\varepsilon^2}, \dot{d}_\lambda, \dot{\boldsymbol{d}}_{\beta^\star}')' \text{ with } \dot{\boldsymbol{d}}_\phi = (\dot{d}_\alpha, \dot{\boldsymbol{d}}_\gamma', \dot{\boldsymbol{d}}_\psi')',\ \dot{\boldsymbol{d}}_{\beta^\star} = (\dot{d}_{\beta_0^\star}, \ldots, \dot{d}_{\beta_p^\star})',$$

where,

$$\begin{aligned}
\dot{d}_\alpha &= \left(\sigma_{\varepsilon 0}^2\right)^{-1} \left[(M_1\boldsymbol{y}_0 + M_2 Z\boldsymbol{\gamma}_0 + M_3 X\boldsymbol{\psi}_0)' \Omega_{0NT}^{-1}\boldsymbol{V}_{NT} + \boldsymbol{V}_{NT}' M_2' \Omega_{0NT}^{-1}\boldsymbol{V}_{NT}\right],\\
\dot{\boldsymbol{d}}_\gamma &= \left(\sigma_{\varepsilon 0}^2\right)^{-1} Z'\Omega_{0NT}^{-1}\boldsymbol{V}_{NT},\ \dot{\boldsymbol{d}}_\psi = \left(\sigma_{\varepsilon 0}^2\right)^{-1} X'(\boldsymbol{\iota}_T \otimes I_N)'\Omega_{0NT}^{-1}\boldsymbol{V}_{NT},\\
\dot{d}_{\sigma_\varepsilon^2} &= \left(2\sigma_{\varepsilon 0}^4\right)^{-1} \boldsymbol{V}_{NT}'\Omega_{0NT}^{-1}\boldsymbol{V}_{NT} - \left(2\sigma_{\varepsilon 0}^2\right)^{-1} NT,\\
\dot{d}_\lambda &= \left(\sigma_{\varepsilon 0}^2\right)^{-1} \big\{[\alpha_0 M_1\boldsymbol{y}_0 + (\alpha_0 M_2 + I_N)Z\boldsymbol{\gamma}_0 + (\alpha_0 M_3 + \boldsymbol{\iota}_T \otimes I_N)X\boldsymbol{\psi}_0]'\\
&\quad (S_0^{-1})'(I_T \otimes W_N)'\Omega_{0NT}^{-1}\boldsymbol{V}_{NT} + \boldsymbol{V}_{NT}'(\alpha_0 M_2 + I_N)'(S_0^{-1})'(I_T \otimes W_N)'\Omega_{0NT}^{-1}\boldsymbol{V}_{NT}\big\}\\
&\quad -tr\left[S_{0NT}^{-1}(I_T \otimes W_N)\right],
\end{aligned}$$

$$\dot{d}_{\beta_\ell^\star} = (2\sigma_{\varepsilon 0}^2)^{-1}\boldsymbol{V}'_{NT}\Omega_{0NT}^{-1}\frac{\partial\Omega_{NT}(\boldsymbol{\beta}_0^\star)}{\partial\beta_\ell^\star}\Omega_{0NT}^{-1}\boldsymbol{V}_{NT} - \frac{1}{2}tr\left[\Omega_{0NT}^{-1}\frac{\partial\Omega_{NT}(\boldsymbol{\beta}_0^\star)}{\partial\beta_\ell^\star}\right],\ \ell = 0, 1, \ldots, p.$$

Let $\boldsymbol{r}_{NT} = (\boldsymbol{r}'_1, \ldots, \boldsymbol{r}'_T)'$ with $\boldsymbol{r}_t = \bar{A}_{0N}^{-1/2}\boldsymbol{v}_t$, then $\boldsymbol{r}_{NT} = (I_T \otimes \bar{A}_{0N}^{-1/2})\boldsymbol{V}_{NT}$. It follows that elements of $\boldsymbol{s}(\boldsymbol{Y}_{-1}, \boldsymbol{V}_{NT}, \boldsymbol{\zeta}_0)$ can be rewritten in terms of $\boldsymbol{\zeta}_0$ and $\boldsymbol{r}_{NT}$, which contains three types of terms: constant C, linear $\boldsymbol{g}'\boldsymbol{r}_{NT}$ and quadratic $\boldsymbol{r}'_{NT}Q\boldsymbol{r}_{NT}$. Hence the result in Proposition 2 follows if the following conditions hold:

(a) $\frac{1}{NT}\left[\text{var}^*(\widehat{\boldsymbol{g}}'\widehat{\boldsymbol{r}}^b_{NT}) - \text{var}(\boldsymbol{g}'\boldsymbol{r}_{NT})\right] = o_p(1)$;

(b) $\frac{1}{NT}\left[\text{var}^*(\boldsymbol{r}^{b'}_{NT}\widehat{Q}\boldsymbol{r}^b_{NT}) - \text{var}(\boldsymbol{r}'_{NT}Q\boldsymbol{r}_{NT})\right] = o_p(1)$;

(c) $\frac{1}{NT}\left[\text{cov}^*(\widehat{\boldsymbol{g}}'\widehat{\boldsymbol{r}}^b_{NT}, \boldsymbol{r}^{b'}_{NT}\widehat{Q}\boldsymbol{r}^b_{NT}) - \text{cov}(\boldsymbol{g}'\boldsymbol{r}_{NT}, \boldsymbol{r}'_{NT}Q\boldsymbol{r}_{NT})\right] = o_p(1)$,

where the operators var* and cov* denote variance and covariance corresponding to the bootstrap probability space, $\widehat{\boldsymbol{g}}$ and $\widehat{Q}$ are estimates of $\boldsymbol{g}$ and Q by plugging in QMLE $\widehat{\boldsymbol{\zeta}}$, respectively, and $\boldsymbol{r}^b_{NT} = (\boldsymbol{r}^{b'}_1, \ldots, \boldsymbol{r}^{b'}_T)'$. Note that matrices S_0^{-1}, $I_T \otimes W_N$, $I_T \otimes \bar{A}_{0N}^{-1/2}$, M_1, M_2, $T\Omega_{0NT}^{-1}$ and $\frac{1}{T}\frac{\partial}{\partial\beta_p^\star}\Omega_{NT}(\boldsymbol{\beta}_0^\star)$ are uniformly bounded in both row and column sums, and $\boldsymbol{\iota}_T \otimes I_N$ and M_3 are uniformly bounded in row sums. By Lemmas 1 and 2 together with the compactness of the parameter space, it can be shown that $\boldsymbol{g}_a = T\boldsymbol{g} = O_p(1)$, and $Q_a = TQ$ is uniformly bounded in both row and column sums. Therefore, to show (a)–(c), it suffices to show that

(i) $\frac{1}{NT^3}\left[\text{var}^*(\widehat{\boldsymbol{g}}'_a\widehat{\boldsymbol{r}}^b_{NT}) - \text{var}(\boldsymbol{g}'_a\boldsymbol{r}_{NT})\right] = o_p(1)$;

(ii) $\frac{1}{NT^3}\left[\text{var}^*(\boldsymbol{r}^{b'}_{NT}\widehat{Q}_a\boldsymbol{r}^b_{NT}) - \text{var}(\boldsymbol{r}'_{NT}Q_a\boldsymbol{r}_{NT})\right] = o_p(1)$;

(iii) $\frac{1}{NT^3}\left[\text{cov}^*(\widehat{\boldsymbol{g}}'_a\widehat{\boldsymbol{r}}^b_{NT}, \boldsymbol{r}^{b'}_{NT}\widehat{Q}_a\boldsymbol{r}^b_{NT}) - \text{cov}(\boldsymbol{g}'_a\boldsymbol{r}_{NT}, \boldsymbol{r}'_{NT}Q_a\boldsymbol{r}_{NT})\right] = o_p(1)$,

where $\widehat{\boldsymbol{g}}_a = T\widehat{\boldsymbol{g}}$ and $\widehat{Q}_a = T\widehat{Q}$. It can be verified that (i)–(iii) hold by the proof of Proposition 5.1 in Su and Yang (2015). As a result, the proof of this proposition is accomplished.

References

Bernstein, D. S. (2009). *Matrix mathematics: Theory, facts, and formulas.* Princeton University Press.

Canay, I. A. (2011). A simple approach to quantile regression for panel data. *The Econometrics Journal 14*, 368–386.

Chernozhukov, V. and C. Hansen (2006). Instrumental quantile regression inference for structural and treatment effect models. *Journal of Econometrics 132*, 491–525.

Davidson, J. (1994). *Stochastic limit theory: An introduction for econometricians.* Oxford University Press.

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.

Lee, L.-F. (2004). Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica 72*, 1899–1925.

Magnus, J. R. (1982). Multivariate error components analysis of linear and nonlinear regression models by maximum likelihood. *Journal of Econometrics 19*, 239–285.

Newey, K. and D. McFadden (1994). Large sample estimation and hypothesis. *Handbook of Econometrics 4*, 2112–2245.

Su, L. and Z. Yang (2015). QML estimation of dynamic panel data models with spatial errors. *Journal of Econometrics 185*, 230–258.

Van der Vaart, A. W. (2000). *Asymptotic statistics.* Cambridge University Press.

Van Der Vaart, A. W. and J. A. Wellner (1996). *Weak convergence and empirical processes.* Springer.

White, H. (1996). *Estimation, inference and specification analysis.* Cambridge University Press.

Zhao, Z. and Z. Xiao (2014). Efficient regressions via optimally combining quantile information. *Econometric Theory 30*, 1272–1314.

Table S.1: Biases, ESDs and ASDs of the WCQE $\widehat{\boldsymbol{\varphi}}(\tau)$ at $\tau = 0.5$, when the innovations $\{\eta_i\}$ follow the standard normal or uniform distribution.

				$N(0,1)$			$U(-\sqrt{3},\sqrt{3})$		
		N	T	Bias	ESD	ASD	Bias	ESD	ASD
$\tau = 0.5$	φ_0	20	20	0.0164	0.2520	0.8681	0.0037	0.2654	1.0016
		100	50	0.0038	0.1092	0.3529	0.0035	0.1334	0.4541
		300	200	0.0045	0.0589	0.1851	0.0041	0.0750	0.2427
		1500	300	0.0003	0.0272	0.0792	0.0010	0.0321	0.1068
		3000	600	0.0001	0.0196	0.0557	0.0003	0.0243	0.0756
	φ_1	20	20	0.0143	0.3411	1.6275	0.0253	0.4375	1.8757
		100	50	0.0104	0.1718	0.6852	0.0028	0.2384	0.8802
		300	200	0.0052	0.1125	0.3844	0.0015	0.1558	0.5042
		1500	300	0.0030	0.0505	0.1680	0.0014	0.0751	0.2268
		3000	600	0.0004	0.0353	0.1172	0.0021	0.0501	0.1593

Table S.2: Biases, ESDs and ASDs of the WQAE $\widehat{\boldsymbol{\beta}}(\widehat{\boldsymbol{\pi}}_{opt,K})$, when the innovations $\{\eta_i\}$ follow the standard normal or uniform distribution.

			$N(0,1)$			$U(-\sqrt{3},\sqrt{3})$		
	N	T	Bias	ESD	ASD	Bias	ESD	ASD
$\widehat{\beta}_1$	20	20	0.5116	1.2505	1.1530	0.3418	1.1946	1.0260
	100	50	0.1165	0.8976	0.4603	0.0669	0.5035	0.3477
	300	200	0.0348	0.3578	0.2455	0.0606	0.2372	0.1730
	1500	300	0.0195	0.1578	0.1095	0.0151	0.0951	0.0698
	3000	600	0.0136	0.1111	0.0771	0.0110	0.0655	0.0477