Online Supplement to
"Identifying Latent Grouped Patterns in Cointegrated Panels"

Wenxin Huang[a], Sainan Jin[b], and Liangjun Su[b]
[a] Antai College of Economics and Management, Shanghai Jiao Tong University
[b] School of Economics, Singapore Management University

This supplement is composed of four parts. Section B contains the proofs of the technical lemmas in the above paper. Section C studies the determination of the number of groups. Section D provides some details on the practical implementation of the C-Lasso procedure. Sections E and F contain some additional simulation and application results, respectively.

B Proofs of the Technical Lemmas

Proof of Lemma A.1. (i) By Park and Phillips (1988, 1989), we can readily show that

$$\begin{aligned}
\frac{1}{T^2}\tilde{x}_{1,i}'\tilde{x}_{1,i} &= \frac{1}{T^2}\sum_{t=1}^{T} x_{1,it}x_{1,it}' - \frac{1}{\sqrt{T}}\bar{x}_{1,i}\frac{1}{\sqrt{T}}\bar{x}_{1,i}' \\
&\Rightarrow \int_0^1 B_{1,i}B_{1,i}' - \int_0^1 B_{1,i}\int_0^1 B_{1,i} = \int_0^1 \tilde{B}_{1,i}\tilde{B}_{1,i}', \\
\frac{1}{T}\tilde{x}_{1,i}'\tilde{x}_{2,i}\mathbb{S}_2' &= \frac{1}{T}\sum_{t=1}^{T} x_{1,it}x_{2,it}'\mathbb{S}_2' - \frac{1}{\sqrt{T}}\bar{x}_{1,i}\sqrt{T}\bar{x}_{2,i}'\mathbb{S}_2' \\
&\Rightarrow \left(\int_0^1 B_{1,i}dB_{2,i}' + \Delta_{12,i}\right)\mathbb{S}_2' - \int_0^1 B_{1,i}B_{2,i}(1)\,\mathbb{S}_2' = \left(\int_0^1 \tilde{B}_{1,i}dB_{2,i}' + \Delta_{12,i}\right)\mathbb{S}_2', \\
\frac{1}{T}\mathbb{S}_2\tilde{x}_{2,i}'\tilde{x}_{2,i}\mathbb{S}_2' &= \frac{1}{T}\sum_{t=1}^{T}\mathbb{S}_2 x_{2,it}x_{2,it}'\mathbb{S}_2' - \mathbb{S}_2\bar{x}_{2,i}\bar{x}_{2,i}'\mathbb{S}_2' = \mathbb{S}_2\Sigma_{22,i}\mathbb{S}_2' + O_P\left(T^{-1/2}\right).
\end{aligned}$$

It follows that $\mathbb{S}D_T\hat{Q}_{i,\tilde{x}\tilde{x}}D_T\mathbb{S}' = \mathbb{S}\begin{pmatrix} \frac{1}{T^2}\sum_{t=1}^T \tilde{x}_{1,it}\tilde{x}_{1,it}' & \frac{1}{T^{3/2}}\sum_{t=1}^T \tilde{x}_{1,it}\tilde{x}_{2,it}' \\ \frac{1}{T^{3/2}}\sum_{t=1}^T \tilde{x}_{2,it}\tilde{x}_{1,it}' & \frac{1}{T}\sum_{t=1}^T \tilde{x}_{2,it}\tilde{x}_{2,it}' \end{pmatrix}\mathbb{S}' \Rightarrow \mathbb{S}\begin{pmatrix} \int_0^1 \tilde{B}_{1,i}\tilde{B}_{1,i}' & 0 \\ 0 & \Sigma_{22,i}\end{pmatrix}\mathbb{S}'.$

(ii) and (iii). By Park and Phillips (1988, 1989),

$$\begin{aligned}
\frac{1}{T}\tilde{x}_{1,i}'\tilde{u}_i &= \frac{1}{T}\sum_{t=1}^{T} x_{1,it}u_{it} - \frac{1}{\sqrt{T}}\bar{x}_{1,i}\sqrt{T}\bar{u}_i \\
&\Rightarrow \left(\int_0^1 B_{1,i}dB_{0,i} + \Delta_{10,i}\right) - \int_0^1 B_{1,i}B_{0,i}(1) = \int_0^1 \tilde{B}_{1,i}dB_{0,i}' + \Delta_{10,i},
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{S}_2\left(\frac{1}{\sqrt{T}}\tilde{x}_{2,i}'\tilde{u}_i - \sqrt{T}\Sigma_{20,i}\right) &= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathbb{S}_2\left(x_{2,it}u_{it} - \Sigma_{20,i}\right) - \frac{1}{\sqrt{T}}\sqrt{T}\mathbb{S}_2\bar{x}_{2,i}\sqrt{T}\bar{u}_i \\
&= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathbb{S}_2\left(x_{2,it}u_{it} - \Sigma_{20,i}\right) - O_P\left(T^{-1/2}\right) \\
&\Rightarrow N(0, \mathbb{S}_2 V_{20,i}^0\mathbb{S}_2'),
\end{aligned}$$

where we allow that $\Sigma_{20,i}=E\left(x_{2,it}u_{it}\right)$ to be nonzero and $V_{20,i}^{0}$ denotes the long-run covariance of $x_{2,it}u_{it}-\Sigma_{20,i}$. It follows that

$$\begin{aligned}
T\hat{Q}_{i,\tilde{x}_1\tilde{u}^*} &= \frac{1}{T}\tilde{x}_{1,i}'\tilde{u}_i^* = \frac{1}{T}\tilde{x}_{1,i}'\tilde{u}_i - \frac{1}{T}\tilde{x}_{1,i}'\tilde{x}_{2,i}\Sigma_{22,i}^{-1}\Sigma_{20,i} \\
&\Rightarrow \int_0^1 \tilde{B}_{1,i}dB_{0,i}+\Delta_{10,i}-\left(\int_0^1 \tilde{B}_{1,i}dB_{2,i}'+\Delta_{12,i}\right)\Sigma_{22,i}^{-1}\Sigma_{20,i}, \text{ and} \\
T^{3/2}\mathbb{S}_2\hat{Q}_{i,\tilde{x}_2\tilde{u}^*} &= \frac{1}{\sqrt{T}}\mathbb{S}_2\tilde{x}_{2,i}'\tilde{u}_i^* = \frac{1}{\sqrt{T}}\mathbb{S}_2\left(\tilde{x}_{2,i}'\tilde{u}_i-\tilde{x}_{2,i}'\tilde{x}_{2,i}\Sigma_{22,i}^{-1}\Sigma_{20,i}\right) \\
&= \mathbb{S}_2\left(\frac{1}{\sqrt{T}}x_{2,i}'u_i-\sqrt{T}\Sigma_{20,i}\right)-\sqrt{T}\mathbb{S}_2\left(\frac{1}{T}x_{2,i}'x_{2,i}-\Sigma_{22,i}\right)\Sigma_{22,i}^{-1}\Sigma_{20,i}+O_P\left(T^{-1/2}\right) \\
&= \mathbb{S}_2 J_{1,i}\left(T^{-1/2}\sum_{t=1}^{T}\left(w_{it}w_{it}'-\Sigma_i\right)\right)J_{2,i}'+o_P(1) \\
&\Rightarrow \mathbb{S}_2\left(J_{1,i}\otimes J_{2,i}\right)N\left(0,V_i^0\right),
\end{aligned}$$

where $J_{1,i}=\left(\mathbf{0}_{p_2\times 1},\mathbf{0}_{p_2\times p_1},I_{p_2}\right)$ and $J_{2,i}=\left(1,\mathbf{0}_{1\times p_1},-\Sigma_{20,i}'\Sigma_{22,i}^{-1}\right)$. Combining the above results yields the results in (ii)-(iii).

(iv) and (v). Our conditions ensure that $\left\|\frac{1}{T}\tilde{x}_{2,i}'\tilde{x}_{2,i}-\Sigma_{22,i}\right\|=O_P\left(p_2T^{-1/2}\right)=o_P(1)$ and $\lambda_{\min}(\frac{1}{T}\tilde{x}_{2,i}'\tilde{x}_{2,i}) \geq \lambda_{\min}\left(\Sigma_{22,i}\right)-\left\|\frac{1}{T}\tilde{x}_{2,i}'\tilde{x}_{2,i}-\Sigma_{22,i}\right\| \geq \underline{c}_{22}/2$ with probability approaching 1 (w.p.a.1). [See the proof of Lemma A.2(iv) which is related to the former claim.] By (i)-(iii),

$$\begin{aligned}
\frac{1}{T^2}\tilde{x}_{1,i}'M_{2,i}\tilde{x}_{1,i} &= \frac{1}{T^2}\tilde{x}_{1,i}'\tilde{x}_{1,i}-\frac{1}{T}\left(\frac{1}{T}\tilde{x}_{1,i}'\tilde{x}_{2,i}\right)\left(\frac{1}{T}\tilde{x}_{2,i}'\tilde{x}_{2,i}\right)^{-1}\left(\frac{1}{T}\tilde{x}_{2,i}'\tilde{x}_{1,i}\right)\Rightarrow\int_0^1\tilde{B}_{1,i}\tilde{B}_{1,i}', \\
\frac{1}{T}\mathbb{S}_2\tilde{x}_{2,i}'M_{1,i}\tilde{x}_{2,i}\mathbb{S}_2' &= \frac{1}{T}\mathbb{S}_2\tilde{x}_{2,i}'\tilde{x}_{2,i}\mathbb{S}_2'-\frac{1}{T}\left(\frac{1}{T}\mathbb{S}_2\tilde{x}_{2,i}'\tilde{x}_{1,i}\right)\left(\frac{1}{T^2}\tilde{x}_{1,i}'\tilde{x}_{1,i}\right)^{-1}\left(\frac{1}{T}\tilde{x}_{1,i}'\tilde{x}_{2,i}\mathbb{S}_2'\right) \\
&= \mathbb{S}_2\Sigma_{22,i}\mathbb{S}_2'+O_P\left(T^{-1}\right), \\
\frac{1}{T}\tilde{x}_{1,i}'M_{2,i}\tilde{u}_i &= \frac{1}{T}\tilde{x}_{1,i}'\tilde{u}_i-\left(\frac{1}{T}\tilde{x}_{1,i}'\tilde{x}_{2,i}\right)\left(\frac{1}{T}\tilde{x}_{2,i}'\tilde{x}_{2,i}\right)^{-1}\left(\frac{1}{T}\tilde{x}_{2,i}'\tilde{u}_i\right) \\
&\Rightarrow \left(\int_0^1\tilde{B}_{1,i}dB_{0,i}'+\Delta_{10,i}\right)-\left(\int_0^1\tilde{B}_{1,i}dB_{2,i}'+\Delta_{12,i}\right)\Sigma_{22,i}^{-1}\Sigma_{20,i},
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{S}_2\left(\frac{1}{\sqrt{T}}\tilde{x}_{2,i}'M_{1,i}\tilde{u}_i-\sqrt{T}\Sigma_{20,i}\right) &= \mathbb{S}_2\left(\frac{1}{\sqrt{T}}\tilde{x}_{2,i}'\tilde{u}_i-\sqrt{T}\Sigma_{20,i}\right) \\
&\quad -\frac{1}{\sqrt{T}}\left(\frac{1}{T}\mathbb{S}_2\tilde{x}_{2,i}'\tilde{x}_{1,i}\right)\left(\frac{1}{T^2}\tilde{x}_{1,i}'\tilde{x}_{1,i}\right)^{-1}\left(\frac{1}{T}\tilde{x}_{1,i}'\tilde{u}_i\right) \\
&= \mathbb{S}_2\left(\frac{1}{\sqrt{T}}x_{2,i}'u_i-\sqrt{T}\Sigma_{20,i}\right)-O_P\left(T^{-1/2}\right)\Rightarrow N(0,\mathbb{S}_2V_{20,i}^0\mathbb{S}_2').
\end{aligned}$$

Consequently, we have

$$\begin{aligned}
T\left(\tilde{\beta}_{1,i}-\beta_{1,i}^0\right) &= \left(\frac{1}{T^2}\tilde{x}_{1,i}'M_{2,i}\tilde{x}_{1,i}\right)^{-1}\frac{1}{T}\tilde{x}_{1,i}'M_{2,i}\tilde{u}_i \\
&\Rightarrow \left(\int_0^1\tilde{B}_{1,i}\tilde{B}_{1,i}'\right)^{-1}\left[\left(\int_0^1\tilde{B}_{1,i}dB_{0,i}'+\Delta_{10,i}\right)-\left(\int_0^1\tilde{B}_{1,i}dB_{2,i}'+\Delta_{12,i}\right)\Sigma_{22,i}^{-1}\Sigma_{20,i}\right],
\end{aligned}$$

and

$$\begin{aligned}
\sqrt{T}\mathbb{S}_2\left(\tilde{\beta}_{2,i}-\beta_{2,i}^*\right) &= \mathbb{S}_2\left(\frac{1}{T}\tilde{x}_{2,i}'M_{1,i}\tilde{x}_{2,i}\right)^{-1}\frac{1}{\sqrt{T}}\tilde{x}_{2,i}'M_{1,i}\tilde{u}_i-\sqrt{T}\Sigma_{22,i}^{-1}\Sigma_{20,i} \\
&= \mathbb{S}_2\left(\frac{1}{T}\tilde{x}_{2,i}'M_{1,i}\tilde{x}_{2,i}\right)^{-1}\left(\frac{1}{\sqrt{T}}\tilde{x}_{2,i}'M_{1,i}\tilde{u}_i-\sqrt{T}\Sigma_{20,i}\right) \\
&\quad -\mathbb{S}_2\left(\frac{1}{T}\tilde{x}_{2,i}'M_{1,i}\tilde{x}_{2,i}\right)^{-1}\sqrt{T}\left(\frac{1}{T}\tilde{x}_{2,i}'M_{1,i}\tilde{x}_{2,i}-\Sigma_{22,i}\right)\Sigma_{22,i}^{-1}\Sigma_{20,i} \\
&= \mathbb{S}_2\Sigma_{22,i}^{-1}\left(\frac{1}{\sqrt{T}}x_{2,i}'u_i-\sqrt{T}\Sigma_{20,i}\right)-\mathbb{S}_2\Sigma_{22,i}^{-1}\sqrt{T}\left(\frac{1}{T}x_{2,i}'x_{2,i}-\Sigma_{22,i}\right)\Sigma_{22,i}^{-1}\Sigma_{20,i}+o_P(1) \\
&= \mathbb{S}_2\Sigma_{22,i}^{-1}J_{1,i}\left(T^{-1/2}\sum_{t=1}^{T}(w_{it}w_{it}'-\Sigma_i)\right)J_{2,i}'+o_P(1) \\
&\Rightarrow \mathbb{S}_2\left(\Sigma_{22,i}^{-1}J_{1,i}\otimes J_{2,i}\right)N\left(0,V_i^0\right)\equiv N\left(0,\mathbb{S}_2V_i\mathbb{S}_2'\right),
\end{aligned}$$

where $\beta_{2,i}^*=\beta_{2,i}^0+\Sigma_{22,i}^{-1}\Sigma_{20,i}$, $J_{1,i}=\left(\mathbf{0}_{p_2\times 1},\mathbf{0}_{p_2\times p_1},\Sigma_{22,i}^{-1}\right)$, $J_{2,i}=\left(1,\mathbf{0}_{1\times p_1},-\Sigma_{20,i}'\Sigma_{22,i}^{-1}\right)$, and $V_i=(J_{1.i}\otimes J_{2.i})V_i^0\left(J_{1,i}'\otimes J_{2,i}'\right)$. ■

Proof of Lemma A.2. (i) Noting that $\frac{1}{T^2}\tilde{x}_{1,i}'\tilde{u}_i=\frac{1}{T^2}x_{1,i}'u_i-\frac{1}{T}\bar{x}_{1,i}\bar{u}_i$, it suffices to prove (i) by showing that (i1) $P(\max_{1\le i\le N}\frac{1}{T^2}\left\|x_{1,i}'u_i\right\|\ge ca_{1NT})=o\left(N^{-1}\right)$ and (i2) $P(\max_{1\le i\le N}\frac{1}{T}\left\|\bar{x}_{1,i}\bar{u}_i\right\|\ge ca_{1NT})$ $=o\left(N^{-1}\right)$ for any fixed constant $c>0$. Recall that $\varepsilon_{it}=(u_{it},\varepsilon_{1,it}',\varepsilon_{2,it}')'$. Let S_0 and S_1 be $1\times(1+p_1+p_2)$ and $p_1\times(1+p_1+p_2)$ selection matrices such that $S_0\varepsilon_{it}=u_{it}$ and $S_1\varepsilon_{it}=\varepsilon_{1,it}$. Noting that $x_{1,it}=\sum_{s=1}^{t-1}\varepsilon_{1,is}+\varepsilon_{1,it}$ and $\varepsilon_{it}=\psi_{i0}e_{it}+\varepsilon_{i,t-1}^{\dagger}$ where $\varepsilon_{i,t-1}^{\dagger}=\sum_{j=1}^{\infty}\psi_{ij}e_{i,t-j}$, we have

$$\begin{aligned}
\frac{1}{T^2}x_{1,i}'u_i &= \frac{1}{T^2}\sum_{t=1}^{T}x_{1,it}u_{it}=\frac{1}{T^2}\sum_{t=1}^{T}\left(\sum_{s=1}^{t-1}\varepsilon_{1,is}+\varepsilon_{1,it}\right)u_{it} \\
&= S_1\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{t-1}\varepsilon_{is}\varepsilon_{it}'S_0'+S_1\frac{1}{T^2}\sum_{t=1}^{T}\varepsilon_{it}\varepsilon_{it}'S_0' \\
&= S_1\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{t-1}\varepsilon_{is}e_{it}'\psi_{i0}'S_0'+S_1\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{t-1}\varepsilon_{is}\varepsilon_{i,t-1}^{\dagger\prime}S_0'+S_1\frac{1}{T^2}\sum_{t=1}^{T}\varepsilon_{it}\varepsilon_{it}'S_0' \\
&\equiv b_{1i}+b_{2i}+b_{3i},\ \text{say.} \qquad \text{(B.1)}
\end{aligned}$$

We prove (i1) by showing that $N\cdot P\left(\max_{1\le i\le N}|b_{li}|\ge ca_{1NT}\right)=o(1)$ for any fixed constant $c>0$ and $l=1,2,3$. For notational simplicity, we assume that $p_1=1$.

We first study b_{1i}. Let $z_{it}=S_1\sum_{s=1}^{t-1}\varepsilon_{is}e_{it}'\psi_{i0}'S_0'$ and $\mathcal{F}_{i,t}=\sigma\left(e_{it},e_{i,t-1},...\right)$, the sigma-field generated by the series $\{e_{it}\}$. Then $b_{1i}=\frac{1}{T^2}\sum_{t=1}^{T}z_{it}$. Noting that $E\left(z_{it}|\mathcal{F}_{i,t-1}\right)=0$ by construction, we want to apply the exponential inequality for martingales (see, e.g., Freedman (1975, Proposition 2.1)). Let $c_{1NT}=N^{2/q}T^{1/2}$. We make the following decomposition:

$$b_{1i}=\frac{1}{T^2}\sum_{t=1}^{T}z_{1it}+\frac{1}{T^2}\sum_{t=1}^{T}z_{2it}-\frac{1}{T^2}\sum_{t=1}^{T}E\left[z_{2it}|\mathcal{F}_{i,t-1}\right]\equiv b_{1i,1}+b_{1i,2}-b_{1i,3},$$

where $z_{1it}=z_{it}\mathbf{1}_{it}-E\left[z_{it}\mathbf{1}_{it}|\mathcal{F}_{i,t-1}\right]$, $z_{2it}=z_{it}\bar{\mathbf{1}}_{it}$, $\mathbf{1}_{it}=\mathbf{1}\{|z_{it}|\le c_{1NT}\}$, and $\bar{\mathbf{1}}_{it}=1-\mathbf{1}_{it}$. It suffices to show that $N\cdot P\left(\max_{1\le i\le N}|b_{1i,l}|\ge ca_{1NT}\right)=o(1)$ for $l=1,2,3$.

Let $V_{iT}=\sum_{t=1}^{T}E\left[z_{1i,t}^2|\mathcal{F}_{i,t-1}\right]$, $v_{NT}=N^{2/q}T^2$, and $\mathbf{1}_{it}=\mathbf{1}\{|z_{it}|\le c_{1NT}\}$. Then by the Hölder's and

Jensen's inequalities and the law of iterated expectations,

$$\begin{aligned} E\left(V_{iT}^{q}\right) &= E\left(\sum_{t=1}^{T} E\left[z_{1it}^{2}|\mathcal{F}_{i,t-1}\right]\right)^{q} \leq T^{q-1}\sum_{t=1}^{T} E\left[\left\{E\left[z_{1it}^{2}|\mathcal{F}_{i,t-1}\right]\right\}^{q}\right] \\ &\leq T^{q-1}\sum_{t=1}^{T} E\left|z_{1it}\right|^{2q} \leq 2^{2q}T^{q-1}\sum_{t=1}^{T} E\left|z_{it}\right|^{2q} \leq CT^{q-1}\sum_{t=1}^{T} t^{q} \leq CT^{2q}, \end{aligned}$$

where the fourth inequality follows because

$$E\left|z_{it}\right|^{2q} = E\left|S_1\sum_{s=1}^{t-1}\varepsilon_{is}e_{it}'\psi_{i0}'S_0'\right|^{2q} \leq CE\left\|\sum_{s=1}^{t-1}\varepsilon_{is}\right\|^{2q} E\left\|e_{it}\right\|^{2q} \leq Ct^{q}.$$

Here we use the independence between e_{it} and $\sum_{s=1}^{t-1}\varepsilon_{is}$ and the fact that $E\left\|\sum_{s=1}^{t-1}\varepsilon_{is}\right\|^{2q} \leq Ct^{q}$. [Recall that we allow the constant C to vary across places.] It follows that $N^2\max_{1\leq i\leq N} P\left(V_{iT} > v_{NT}\right) \leq N^2\max_{1\leq i\leq N} v_{NT}^{-q}E\left(V_{iT}^{q}\mathbf{1}\left\{V_{iT} > v_{NT}\right\}\right) = o\left(N^2T^{2q}v_{NT}^{-q}\right) = o\left(1\right)$. By Proposition 2.1 in Freedman (1975), we have

$$\begin{aligned} N\cdot P\left(\max_{1\leq i\leq N}\left|b_{1i,1}\right| \geq ca_{1NT}\right) &\leq N^2\max_{1\leq i\leq N} P\left(\left|\frac{1}{T^2}\sum_{t=1}^{T} z_{1it}\right| \geq ca_{1NT}\right) \\ &\leq N^2\max_{1\leq i\leq N} P\left(\left|\sum_{t=1}^{T} z_{1it}\right| \geq CT^2a_{1NT}, V_{iT} \leq v_{NT}\right) \\ &\quad + N^2\max_{1\leq i\leq N} P\left(V_{iT} > v_{NT}\right) \\ &= N^2\cdot\exp\left(-\frac{c^2T^4a_{1NT}^2}{2v_{NT}+4cT^2a_{1NT}c_{1NT}}\right) + o\left(N^2T^{2q}v_{NT}^{-q}\right) \\ &= o\left(1\right) + o\left(1\right) = o\left(1\right) \end{aligned}$$

as $T^4a_{1NT}^2/v_{NT} = \frac{T^4a_{1NT}^2}{N^{2/q}T^2} = T^2N^{-2/q}a_{1NT}^2 \geq \left(\log N\right)^{1+\epsilon}$ and $T^2a_{1NT}/c_{1NT} \geq \left(\log N\right)^{1+\epsilon}$ for some $\epsilon > 0$ by Assumption A.3(iii).

$$\begin{aligned} N\cdot P\left(\max_{1\leq i\leq N}\left|b_{1i,2}\right| \geq ca_{1NT}\right) &\leq N\cdot P\left(\max_{1\leq i\leq N}\max_{1\leq t\leq T}\left|z_{it}\right| \geq c_{1NT}\right) \\ &\leq \frac{N^2T}{c_{1NT}^{2q}}\cdot\max_{1\leq i\leq N}\max_{1\leq t\leq T} E\left[\left|z_{it}\right|^{2q}\mathbf{1}\left\{\left|z_{it}\right| \geq c_{1NT}\right\}\right] \\ &= o\left(N^2T^{q+1}c_{1NT}^{-2q}\right) = o\left(N^{-2}T\right) = o\left(1\right). \end{aligned}$$

Similarly, we can show that $N\cdot P\left(\max_{1\leq i\leq N}\left|b_{1i,3}\right| \geq ca_{1NT}\right) = o\left(1\right)$. Consequently, we have $N\cdot P(\max_{1\leq i\leq N}\left|b_{1i}\right| \geq ca_{1NT}) = o\left(1\right)$.

To study b_{2i}, we apply the Beveridge-Nelson (BN) decomposition (see, e.g., Lemma 2.1 in Phillips and Solo (1992)) to obtain

$$\varepsilon_{it} = \psi_i\left(1\right)e_{it} + \breve{e}_{i,t-1} - \breve{e}_{it}$$

where $\psi_i\left(1\right) = \sum_{j=0}^{\infty}\psi_{ij}$, $\breve{e}_{it} = \breve{\psi}_i\left(L\right)e_{it} = \sum_{j=0}^{\infty}\breve{\psi}_{ij}e_{i,t-j}$, and $\breve{\psi}_{ij} = \sum_{k=j+1}^{\infty}\psi_{ik}$. Then $\sum_{s=1}^{t-1}\varepsilon_{is} = \psi_i\left(1\right)\sum_{s=1}^{t-1}e_{is} + \breve{e}_{i0} - \breve{e}_{i,t-1}$ and $\varepsilon_{i,t-1}^{\dagger} = \varepsilon_{it} - \psi_{i0}e_{it} = \psi_i^{\dagger}\left(1\right)e_{it} + \breve{e}_{i,t-1} - \breve{e}_{it}$, where $\psi_i^{\dagger}\left(1\right) = \sum_{j=1}^{\infty}\psi_{ij}$. It follows that

$$
\begin{aligned}
b_{2i} &= S_1 \frac{1}{T^2} \sum_{t=1}^{T} \left(\sum_{s=1}^{t-1} \varepsilon_{is} \right) \varepsilon_{i,t-1}^{\dagger\prime} S_0' \\
&= S_1 \frac{1}{T^2} \sum_{t=1}^{T} \left(\psi_i(1) \sum_{s=1}^{t-1} e_{is} + \breve{e}_{i,0} - \breve{e}_{i,t-1} \right) \left(\psi_i^{\dagger}(1) e_{it} + \breve{e}_{i,t-1} - \breve{e}_{it} \right)' S_0' \\
&= S_1 \frac{1}{T^2} \sum_{t=1}^{T} \psi_i(1) \sum_{s=1}^{t-1} e_{is} e_{it}' \psi_i^{\dagger}(1)' S_0' + S_1 \frac{1}{T^2} \sum_{t=1}^{T} \psi_i(1) \sum_{s=1}^{t-1} e_{is} \left(\breve{e}_{i,t-1} - \breve{e}_{it} \right)' S_0' \\
&\quad + S_1 \frac{1}{T^2} \sum_{t=1}^{T} \left(\breve{e}_{i0} - \breve{e}_{i,t-1} \right) e_{it}' \psi_i^{\dagger}(1)' S_0' + S_1 \frac{1}{T^2} \sum_{t=1}^{T} \left(\breve{e}_{i0} - \breve{e}_{i,t-1} \right) \left(\breve{e}_{i,t-1} - \breve{e}_{it} \right)' S_0' \\
&\equiv b_{2i,1} + b_{2i,2} + b_{2i,3} + b_{2i,4}.
\end{aligned}
$$

Noting that $b_{2i,1} = \frac{1}{T^2} \sum_{t=1}^{T} z_{it}^{\dagger}$ and $b_{2i,3} = \frac{1}{T^2} \sum_{t=1}^{T} z_{it}^{\ddagger}$ where $z_{it}^{\dagger} = S_1 \psi_i(1) \sum_{s=1}^{t-1} e_{is} e_{it}' \psi_i^{\dagger}(1)' S_0'$ and $z_{it}^{\ddagger} = S_1 \left(\breve{e}_{i,0} - \breve{e}_{i,t-1} \right) e_{it}' \psi_i^{\dagger}(1)' S_0'$ satisfy $E\left\{ z_{it}^{\dagger} | \mathcal{F}_{i,t-1} \right\} = 0$ and $E\left\{ z_{it}^{\ddagger} | \mathcal{F}_{i,t-1} \right\} = 0$, we can also follow the analysis of b_{1i} and show that $N \cdot P\left(\max_{1 \le i \le N} |b_{2i,l}| \ge c a_{1NT} \right) = o(1)$ for $l = 1, 3$. It remains to show that $N \cdot P\left(\max_{1 \le i \le N} |b_{2i,l}| \ge c a_{1NT} \right) = o(1)$ for $l = 2, 4$. For $b_{2i,2}$, we have

$$
\begin{aligned}
b_{2i,2} &= S_1 \frac{1}{T^2} \sum_{s=1}^{T-1} \psi_i(1) e_{is} \sum_{t=s+1}^{T} \left(\breve{e}_{i,t-1} - \breve{e}_{it} \right)' S_0' \\
&= S_1 \frac{1}{T^2} \sum_{s=1}^{T-1} \psi_i(1) e_{is} \left(\breve{e}_{is} - \breve{e}_{iT} \right)' S_0' \\
&= S_1 \frac{1}{T^2} \sum_{s=1}^{T-1} \psi_i(1) e_{is} \breve{e}_{is}' S_0' - S_1 \frac{1}{T^2} \sum_{s=1}^{T-1} \psi_i(1) e_{is} \breve{e}_{iT}' S_0' \equiv b_{2i,2a} - b_{2i,2b}.
\end{aligned}
$$

For $b_{2i,2b}$, we have

$$
\begin{aligned}
N \cdot P\left(\max_{1 \le i \le N} |b_{2i,2b}| \ge c a_{1NT} \right) &\le N^2 \max_{1 \le i \le N} P\left(|b_{2i,2b}| \ge c a_{1NT} \right) \\
&\le N^2 T^{-2q} (c a_{1NT})^{-q} \max_{1 \le i \le N} E \left\| S_1 \sum_{s=1}^{T-1} \psi_i(1) e_{is} \breve{e}_{iT}' S_0' \right\|^q \\
&\le C N^2 T^{-2q} (c a_{1NT})^{-q} \max_{1 \le i \le N} \left\{ E \left\| \sum_{s=1}^{T-1} S_1 \psi_i(1) e_{is} \right\|^{2q} E \left\| S_0 \breve{e}_{iT} \right\|^{2q} \right\}^{1/2} \\
&\le C N^2 T^{-2q} (c a_{1NT})^{-q} O(T^{q/2}) = O(N^2 T^{-3q/2} a_{1NT}^{-q}) = o(1).
\end{aligned}
$$

For $b_{2i,2a}$, we can make further decomposition

$$
\begin{aligned}
b_{2i,2a} &= S_1 \psi_i(1) \frac{1}{T^2} \sum_{s=1}^{T-1} \Omega_i \breve{\psi}_{i0}' S_0' + S_1 \psi_i(1) \frac{1}{T^2} \sum_{s=1}^{T-1} \left(e_{is} e_{is}' - \Omega_i \right) \breve{\psi}_{i0}' S_0' \\
&\quad + S_1 \psi_i(1) \frac{1}{T^2} \sum_{s=1}^{T-1} e_{is} \left(\breve{e}_{is} - \breve{\psi}_{i0} e_{is} \right)' S_0' \equiv b_{2i,2a}(1) + b_{2i,2a}(2) + b_{2i,2a}(3).
\end{aligned}
$$

Apparently, $\max_{1\le i\le N}|b_{2i,2a}(1)| \le \frac{1}{T}\max_{1\le i\le N}\left\{\|S_1\psi_i(1)\|\left\|S_0\breve{\psi}_{i0}\right\|\right\}\max_{1\le i\le N}\|\Omega_i\|_{\mathrm{sp}} \le \frac{C}{T}$. Noting that $E\left(e_{is}e_{is}' - \Omega_i|\mathcal{F}_{i,s-1}\right) = 0$, we can apply the Markov and Burkholder's inequalities to obtain

$$\begin{aligned}
N\cdot P\left(\max_{1\le i\le N}|b_{2i,2a}(2)| \ge ca_{1NT}\right) &\le N^2\max_{1\le i\le N}P\left(|b_{2i,2a}(2)| \ge ca_{1NT}\right)\\
&\le N^2T^{-2q}(ca_{1NT})^{-q}\max_{1\le i\le N}E\left\|S_1\psi_i(1)\sum_{s=1}^{T-1}\left(e_{is}e_{is}'-\Omega_i\right)\breve{\psi}_{i0}'S_0'\right\|^q\\
&\le CN^2T^{-2q}(ca_{1NT})^{-q}T^{q/2} = O(N^2T^{-3/2q}a_{1NT}^{-q}) = o(1).
\end{aligned}$$

Similarly, noting that $E\left[e_{is}(\breve{e}_{is}-\breve{\psi}_{i0}e_{is})|\mathcal{F}_{i,s-1}\right] = 0$, we can apply the Markov and Burkholder's inequalities (e.g., Hall and Heyde, 1980, p.23) to obtain $N\cdot P\left(\max_{1\le i\le N}|b_{2i,2a}(3)| \ge ca_{1NT}\right) = o(1)$. Consequently, we have $N\cdot P(\max_{1\le i\le N}|b_{2i,2}| \ge ca_{1NT}) = o(1)$.

For $b_{2i,4}$, we make the following decomposition

$$b_{2i,4} = S_1\frac{1}{T^2}\breve{e}_{i0}\left(\breve{e}_{i0}-\breve{e}_{iT}\right)'S_0' - S_1\frac{1}{T^2}\sum_{t=1}^{T}\breve{e}_{i,t-1}\left(\breve{e}_{i,t-1}-\breve{e}_{it}\right)'S_0' \equiv b_{2i,4a} - b_{2i,4b}.$$

By the Markov inequality

$$\begin{aligned}
N\cdot P\left(\max_{1\le i\le N}|b_{2i,4a}| \ge ca_{1NT}\right) &\le N^2\max_{1\le i\le N}P\left(|b_{2i,4a}| \ge ca_{1NT}\right)\\
&\le N^2T^{-2q}(ca_{1NT})^{-q}E\left\|S_1\breve{e}_{i0}\left(\breve{e}_{i0}-\breve{e}_{iT}\right)S_0'\right\|\\
&= O(N^2T^{-2q}a_{1NT}^{-q}) = o(1).
\end{aligned}$$

As in the analysis of $b_{2i,2a}$, we can show that $N\cdot P\left(\max_{1\le i\le N}|b_{2i,4b}| \ge ca_{1NT}\right) = o(1)$. Then $N\cdot P(\max_{1\le i\le N}|b_{2i,4}| \ge ca_{1NT}) = o(1)$. In sum, we have $N\cdot P\left(\max_{1\le i\le N}|b_{2i}| \ge ca_{1NT}\right) = o(1)$.

For b_{3i}, we make the following decomposition

$$\begin{aligned}
b_{3i} &= S_1\frac{1}{T^2}\sum_{t=1}^{T}\sum_{j=1}^{\infty}\sum_{l=1}^{\infty}\psi_{ij}e_{i,t-j}e_{i,t-l}'\psi_{il}'S_0'\\
&= S_1\frac{1}{T^2}\sum_{t=1}^{T}\sum_{j=1}^{\infty}\psi_{ij}\Omega_i\psi_{ij}'S_0' + S_1\frac{1}{T^2}\sum_{t=1}^{T}\sum_{j=1}^{\infty}\psi_{ij}\left[e_{i,t-j}e_{i,t-j}'-\Omega_i\right]\psi_{ij}'S_0'\\
&\quad + S_1\frac{1}{T^2}\sum_{t=1}^{T}\sum_{j=1}^{\infty}\sum_{l=j+1}^{\infty}\psi_{ij}e_{i,t-j}e_{i,t-l}'\psi_{il}'S_0' + S_1\frac{1}{T^2}\sum_{t=1}^{T}\sum_{l=1}^{\infty}\sum_{j=l+1}^{\infty}\psi_{ij}e_{i,t-j}e_{i,t-l}'\psi_{il}'S_0'\\
&\equiv b_{3i,1}+b_{3i,2}+b_{3i,3}+b_{3i,4}.
\end{aligned}$$

It is easy to show that $\max_{1\le i\le N}|b_{3i,1}| \le C/T$. For $b_{3i,l}$ with $l = 2,3,4$, by tedious calculations we can show that $E\left(b_{3,l}\right)^4 \le CT^{-6}$. With this, we can apply the Markov inequality to show that $N\cdot P(\max_{1\le i\le N}|b_{3i,l}| \ge ca_{1NT}) \le N^2(ca_{1NT})^{-4}\max_{1\le i\le N}E\left|b_{3i,l}\right|^4 = O\left(N^2T^{-6}a_{1NT}^{-4}\right) = o(1)$ for $l = 2,3,4$. Then $N\cdot P(\max_{1\le i\le N}|b_{3i}| \ge ca_{1NT}) = o(1)$ for each fixed constant $c > 0$.

Consequently, we have shown (i1).

We now show (i2) $P\left(\max_{1\le i\le N}\frac{1}{T}\|\bar{x}_{1,i}\bar{u}_i\| \ge ca_{1NT}\right) = o\left(N^{-1}\right)$ for any fixed constant $c > 0$. Noting that by Lemma S1.2 in Su, Shi, and Phillips (2016b, hereafter SSPb),

$$P\left(\max_{1\leq i\leq N}|\bar{u}_i|\geq cT^{-1/2}(\log T)^3\right)=o\left(N^{-1}\right).$$

Using $\varepsilon_{it}=\psi_i(1)\,e_{it}+\breve{e}_{i,t-1}-\breve{e}_{it}$,

$$\begin{aligned}\frac{1}{T}\bar{x}_{1,i} &= \frac{1}{T^2}\sum_{t=1}^{T}x_{1,it}=\frac{1}{T^2}\sum_{t=1}^{T}S_1\sum_{s=1}^{t}\varepsilon_{is}\\ &= \frac{1}{T^2}S_1\psi_i(1)\sum_{t=1}^{T}\sum_{s=1}^{t}e_{is}+\frac{1}{T}S_1\breve{e}_{i0}-\frac{1}{T^2}S_1\sum_{t=1}^{T}\breve{e}_{it}\\ &\equiv c_{1,i}+c_{2,i}+c_{3,i},\ \text{say.}\end{aligned}$$

As in the analysis of b_{1i}, we can show that $N\cdot P\left(\max_{1\leq i\leq N}|c_{1,i}|\geq ca_{1NT}\right)=o(1)$. By the Markov inequality, we can show that

$$\begin{aligned}N\cdot P\left(\max_{1\leq i\leq N}|c_{2,i}|\geq ca_{1NT}T^{1/2}(\log T)^{-3}\right) &\leq N^2\max_{1\leq i\leq N}P\left(\|S_1\breve{e}_{i0}\|\geq a_{1NT}T^{3/2}(\log T)^{-3}\right)\\ &= N^2 o\left(a_{1NT}^{-2q}T^{-3q}(\log T)^{6q}\right)=o(1).\end{aligned}$$

For $c_{3,i}$, we use the fact that $E\left\|\sum_{t=1}^{T}S_1\breve{e}_{it}\right\|^{2q}\leq CT^q$ and the Markov inequality to obtain

$$\begin{aligned}N\cdot P\left(\max_{1\leq i\leq N}|c_{3,i}|\geq ca_{1NT}T^{1/2}(\log T)^{-3}\right) &\leq N^2\max_{1\leq i\leq N}P\left(\frac{1}{T^2}\left\|\sum_{t=1}^{T}S_1\breve{e}_{it}\right\|\geq ca_{1NT}T^{1/2}(\log T)^{-3}\right)\\ &\leq \frac{N^2T^{-4q}}{\left(ca_{1NT}T^{1/2}(\log T)^{-3}\right)^{2q}}O(T^q)\\ &= O\left(N^2a_{1NT}^{-2q}T^{-4q}(\log T)^{6q}\right)=o(1).\end{aligned}$$

Consequently, $P(\max_{1\leq i\leq N}\frac{1}{T}\|\bar{x}_{1,i}\|\geq ca_{1NT}T^{1/2}(\log T)^{-3})=o\left(N^{-1}\right)$. Let $z_{it}=\sum_{s=1}^{t}S_1\psi_i(1)\,e_{is}$. Then $c_{1,i}=\frac{1}{T^2}\sum_{t=1}^{T}z_{it}$ where $E(z_{it}|\mathcal{F}_{i,t-1})=z_{i,t-1}$. We can readily follow the analysis of b_{2i} and show that $P\left(\max_{1\leq i\leq N}|c_{1,i}|\geq ca_{1NT}\right)=o\left(N^{-1}\right)$. It follows that $N\cdot P(\max_{1\leq i\leq N}|\bar{x}_{1,i}|\geq ca_{1NT}T^{1/2}(\log T)^{-3})=o(1)$. Then

$$\begin{aligned}& P\left(\max_{1\leq i\leq N}\frac{1}{T}\|\bar{x}_{1,i}\bar{u}_i\|\geq ca_{1NT}\right)\\ \leq\ & P\left(\max_{1\leq i\leq N}\frac{1}{T}\|\bar{x}_{1,i}\bar{u}_i\|\geq ca_{1NT},\ \max_{1\leq i\leq N}|\bar{u}_i|\leq T^{-1/2}(\log T)^3\right)+P\left(\max_{1\leq i\leq N}|\bar{u}_i|\geq T^{-1/2}(\log T)^3\right)\\ \leq\ & P\left(\max_{1\leq i\leq N}\frac{1}{T}\|\bar{x}_{1,i}\|\geq ca_{1NT}T^{1/2}(\log T)^{-3}\right)+o\left(N^{-1}\right)\\ =\ & o\left(N^{-1}\right)+o\left(N^{-1}\right)=o\left(N^{-1}\right).\end{aligned}$$

(ii) Noting that $\frac{1}{T}\tilde{x}_{2,i}'\tilde{u}_i=\frac{1}{T}x_{2,i}'u_i-\bar{x}_{2,i}\bar{u}_i$, it suffices to prove (ii) by showing that (i1) $P(\max_{1\leq i\leq N}\|\frac{1}{T}x_{2,i}'u_i-\Sigma_{20,i}\|\geq cp_2^{1/2}a_{2NT})=o\left(N^{-1}\right)$ and (i2) $P(\max_{1\leq i\leq N}\frac{1}{T}\|\bar{x}_{2,i}\bar{u}_i\|\geq cp_2^{1/2}a_{2NT})=o\left(N^{-1}\right)$ for any fixed constant $c>0$, where $a_{2NT}=T^{-1/2}(\log T)^3$. (i1) follows directly from a modification of the proof of Lemma S.1.2 in SSPb. Noting that both $x_{2,it}$ and u_{it} have zero mean, we can follow SSPb and show that $P\left(\max_{1\leq i\leq N}\|\bar{x}_{2,i}\|\geq cp_2^{1/2}a_{2NT}\right)=o\left(N^{-1}\right)$ and $P\left(\max_{1\leq i\leq N}|\bar{u}_i|\geq ca_{2NT}\right)=o\left(N^{-1}\right)$, implying that $P\left(\max_{1\leq i\leq N}\|\bar{x}_{2,i}\bar{u}_i\|\geq ca_{2NT}^2\right)=o\left(N^{-1}\right)$. Consequently, we have $P(\max_{1\leq i\leq N}||\frac{1}{T}\tilde{x}_{2,i}'\tilde{u}_i-\Sigma_{20,i}||\geq cp_2^{1/2}a_{2NT})=o\left(N^{-1}\right)$.

(iii) Noting that $E\left(x_{2,it}\right)=0$, the proof is analogous to that of (i) and thus omitted.

(iv) Note that $\frac{1}{T}\tilde{x}_{2,i}'\tilde{x}_{2,i}-\Sigma_{22,i}=\frac{1}{T}\sum_{t=1}^T\left(x_{2,it}x_{2,it}'-\Sigma_{22,i}\right)-\bar{x}_{2,i}\bar{x}_{2,i}'$. Using Lemma S1.2 in SSPb, we can readily show that $P(\max_{1\leq i\leq N}||\frac{1}{T}\sum_{t=1}^T\left(x_{2,it}x_{2,it}'-\Sigma_{22,i}\right)||\geq cp_2a_{2NT})=o\left(N^{-1}\right)$ and $P(\max_{1\leq i\leq N}\|\bar{x}_{2,i}\|\geq cp_2^{1/2}a_{2NT})=o\left(N^{-1}\right)$ for any $c>0$. Thus (iv) follows.

(v) Note that $\hat{Q}_{i,\tilde{x}_1\tilde{u}^*}=\frac{1}{T^2}\tilde{x}_{1,i}'\tilde{u}_i-\frac{1}{T^2}\tilde{x}_{1,i}'\tilde{x}_{2,i}\Sigma_{22,i}^{-1}\Sigma_{20,i}$. The condition in Assumption A.2(ii)-(iii) ensures that $\tilde{x}_{2,it}'\Sigma_{22,i}^{-1}\Sigma_{20,i}$ behaves like $\tilde{u}_{it}$ despite the possible divergence of p_2. As a result, part (i) also holds when $\frac{1}{T^2}\tilde{x}_{1,i}'\tilde{u}_i$ is replaced by $\frac{1}{T^2}\tilde{x}_{1,i}'\tilde{x}_{2,i}\Sigma_{22,i}^{-1}\Sigma_{20,i}$. Then

$$\begin{aligned}P\left(\max_{1\leq i\leq N}\left\|\hat{Q}_{i,\tilde{x}_1\tilde{u}^*}\right\|\geq ca_{1NT}\right) &\leq P\left(\max_{1\leq i\leq N}\left\|\frac{1}{T^2}\tilde{x}_{i,1}'\tilde{u}_i\right\|\geq ca_{1NT}/2\right)\\ &\quad+P\left(\max_{1\leq i\leq N}\left\|\frac{1}{T^2}\tilde{x}_{i,1}'\tilde{x}_{i,2}\Sigma_{22,i}^{-1}\Sigma_{20,i}\right\|\geq ca_{1NT}/2\right)\\ &= o\left(N^{-1}\right)+o\left(N^{-1}\right)=o\left(N^{-1}\right).\end{aligned}$$

(vi) Note that $T\hat{Q}_{i,\tilde{x}_2\tilde{u}^*}=\frac{1}{T}\tilde{x}_{2,i}'\left(\tilde{u}_i-\tilde{x}_{2,i}'\Sigma_{22,i}^{-1}\Sigma_{20,i}\right)=\frac{1}{T}x_{2,i}'\left(u_i-x_{2,i}'\Sigma_{22,i}^{-1}\Sigma_{20,i}\right)-\bar{x}_{2,i}'\left(\bar{u}_i-\bar{x}_{2,i}'\Sigma_{22,i}^{-1}\Sigma_{20,i}\right)$. Since $E\left(u_{it}-x_{2,it}'\Sigma_{22,i}^{-1}\Sigma_{20,i}\right)=0$, we can use Lemma S1.2 in SSPb and show that $P(\max_{1\leq i\leq N}||\frac{1}{T}x_{2,i}'(u_i-x_{2,i}'\Sigma_{22,i}^{-1}\Sigma_{20,i})||\geq cp_2^{1/2}a_{2NT}/2)=o\left(N^{-1}\right)$ for any fixed $c>0$. In addition, $P(\max_{1\leq i\leq N}||\bar{x}_{2,i}||\geq cp_2^{1/2}a_{2NT})=o\left(N^{-1}\right)$ and $P(\max_{1\leq i\leq N}||\bar{u}_i||\geq ca_{2NT})=o\left(N^{-1}\right)$, from which we can readily show that $P(\max_{1\leq i\leq N}||\frac{1}{T}\bar{x}_{2,i}'(\bar{u}_i-\bar{x}_{2,i}'\Sigma_{22,i}^{-1}\Sigma_{20,i})||\geq cp_2^{1/2}a_{2NT}/2)=o\left(N^{-1}\right)$. Then (vi) follows. ■

Proof of Lemma A.3. (i) Let $v\in\mathbb{R}^{p_1}$ be an arbitrary vector such that $\|v\|=1$. Let $d_T=\sqrt{2T\log\log T}$. By arguments used in the proof of Lemma 2.1 of Corradi (1999), we can verify the conditions in Theorem 2 of Eberlein (1986) and obtain

$$\begin{aligned}\frac{1}{d_{[Tr]}}\Omega_{11,i}^{-1/2}\tilde{x}_{1,i,[Tr]} &= \frac{1}{d_{[Tr]}}\Omega_{11,i}^{-1/2}\left[x_{1,i,[Tr]}-\bar{x}_{1,i}\right]=\frac{1}{d_{[Tr]}}\Omega_{11,i}^{-1/2}\left[\sum_{s=1}^{[Tr]}\varepsilon_{1,is}-\frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{t}\varepsilon_{1,is}\right]\\ &= \frac{1}{d_{[Tr]}}\tilde{B}_i\left(Tr\right)+o_{a.s.}\left(1\right)\end{aligned}$$

for each $r\in[0,1]$. This result can be strengthened to hold uniformly in $r\in[0,1]$ with $d_{[Tr]}$ replaced by d_T. Let $t=[Tr]$ and $\Omega_{11,i}^{-1/2}\sum_{s=1}^{[Tr]}\varepsilon_{1,is}=S_{i,[Tr]}=S_{i,T}\left(r\right)$. Define

$$\eta_{i,T}\left(r\right)=\left([Tr]+1-Tr\right)S_{i,T}\left(r\right)+\left(Tr-[Tr]\right)S_{i,T+1}\left(r\right).$$

Then $\sup_{r\in[0,1]}\frac{1}{d_T}\left\|S_{i,T}\left(r\right)-B_i\left(Tr\right)\right\|=o_{a.s.}\left(1\right)$, $\sup_{r\in[0,1]}\frac{1}{d_T}||S_{i,T}\left(r\right)-\eta_{i,T}\left(r\right)||=o_{a.s.}\left(1\right)$ as $T\to\infty$, and the set of norm limit points of $\left\{d_T^{-1}S_{i,T}\right\}$ and $\left\{d_T^{-1}\eta_{i,T}\right\}$ coincides with the set of norm limit points of $\left\{d_T^{-1}B_{i,T}\right\}$ with probability one. By Theorem 1 in Strassen (1964), the latter is relatively norm compact with the set of limit points coinciding a.s. with $\mathcal{K}$, where

$$\mathcal{K}=\left\{f:[0,1]\rightarrow\mathbb{R}^{p_1}, f(0)=0,\ f \text{ is absolutely continuous}, \ \int_0^1 \left\|\dot{f}(r)\right\|^2 dr\leq 1\right\}.$$

Here $\dot{f}(r)=\partial f(r)/\nabla r$. First, observe that

$$P\left(\limsup_{T\rightarrow\infty}\max_{\|v\|=1} v'\Omega_{11,i}^{-1/2}\frac{1}{Td_T^2}\sum_{t=1}^{T}\tilde{x}_{1,it}\tilde{x}_{1,it}'\Omega_{11,i}^{-1/2}v=\limsup_{T\rightarrow\infty}\max_{\|v\|=1}\frac{1}{d_T^2}v'\int_0^1\tilde{\eta}_{i,T}(r)\tilde{\eta}_{i,T}(r)'drv\right)=1,$$

where $\tilde{\eta}_{i,T}(r)=\eta_{i,T}(r)-\int_0^1\eta_{i,T}(r)dr$.

Now, let ϕ_v denote the continuous map from the space of p_1-dimensional continuous functions on $[0,1]$, closed with respect to the sup norm, to the Euclidean space such that $\phi_v(f)=v'\int_0^1\tilde{f}(r)\tilde{f}(r)'drv$ where $\tilde{f}(r)=f(r)-\int_0^1 f(r)dr$. By the Corollary of Theorem 3 in Strassen (1964), with probability one $\left\{\phi_v\left(d_T^{-1}\eta_{i,T}\right)\right\}$ is relatively norm compact with the set of norm limit points coinciding almost surely with $\phi_v(\mathbb{K})$. This implies that

$$P\left(\limsup_{T\rightarrow\infty}\max_{\|v\|=1}\frac{1}{d_T^2}v'\int_0^1\tilde{\eta}_{i,T}(r)\tilde{\eta}_{i,T}(r)'drv=\sup_{f\in\mathcal{K}}\max_{\|v\|=1}\phi_v(f)\right).$$

By the definition of ϕ_v and K,

$$\begin{aligned}
\sup_{f\in\mathcal{K}}\max_{\|v\|=1}\phi_v(f) &= \sup_{f\in\mathcal{K}}\max_{\|v\|=1} v'\int_0^1\tilde{f}(r)\tilde{f}(r)'drv\\
&\leq \sup_{f\in\mathcal{K}}\max_{\|v\|=1} v'\left(\int_0^1 f(r)f(r)'dr-\int_0^1 f(r)dr\int_0^1 f(r)'dr\right)v\\
&\leq \sup_{f\in\mathcal{K}}\max_{\|v\|=1}\int_0^1\left(\int_0^r v'\dot{f}(s)ds\right)^2 dr\leq\sup_{f\in\mathcal{K}}\max_{\|v\|=1}\int_0^1\left(\int_0^r 1^2ds\right)\int_0^r\left(v'\dot{f}(s)\right)^2ds\\
&= \sup_{f\in\mathcal{K}}\max_{\|v\|=1}\int_0^1 r\left(v'\int_0^r\dot{f}(s)\dot{f}(s)'dsv\right)dr\leq\int_0^1 rdr=\frac{1}{2},
\end{aligned}$$

where the second inequality follows from the Hölder's inequality, and the third follows from the fact that

$$\max_{\|v\|=1}v'\int_0^r\dot{f}(s)\dot{f}(s)'dsv\leq\lambda_{\max}\left(\int_0^r\dot{f}(s)\dot{f}(s)'ds\right)\leq\mathrm{tr}\left(\int_0^r\dot{f}(s)\dot{f}(s)'ds\right)\leq\int_0^1\left\|\dot{f}(s)\right\|^2ds=1,$$

for any $r\in[0,1]$ and $f\in\mathcal{K}$. It follows that

$$\limsup_{T\rightarrow\infty}\lambda_{\max}\left(\Omega_{11,i}^{-1/2}\frac{1}{Td_T^2}\sum_{t=1}^{T}\tilde{x}_{1,it}\tilde{x}_{1,it}'\Omega_{11,i}^{-1/2}\right)\leq\frac{1}{2}+c \text{ a.s. for any } c>0,$$

and

$$\begin{aligned}
\limsup_{T\to\infty}\lambda_{\max}\left(\frac{1}{Td_T^2}\sum_{t=1}^T\tilde{x}_{1,it}\tilde{x}_{1,it}'\right) &= \limsup_{T\to\infty}\lambda_{\max}\left(\Omega_{11,i}^{-1/2}\frac{1}{Td_T^2}\sum_{t=1}^T\tilde{x}_{1,it}\tilde{x}_{1,it}'\Omega_{11,i}^{-1/2}\Omega_{11,i}\right)\\
&\leq \limsup_{T\to\infty}\lambda_{\max}\left(\Omega_{11,i}^{-1/2}\frac{1}{Td_T^2}\sum_{t=1}^T\tilde{x}_{1,it}\tilde{x}_{1,it}'\Omega_{11,i}^{-1/2}\right)\max_{1\leq i\leq N}\lambda_{\max}\left(\Omega_{11,i}\right)\\
&\leq \left(\frac{1}{2}+c\right)\bar{c}_{\Omega_{11}} \text{ a.s.,}
\end{aligned}$$

where recall that $\bar{c}_{\Omega_{11}}$ denotes the upper bound for $\lambda_{\max}\left(\Omega_{11,i}\right)$.

(ii) Let v be a $p_2\times 1$ vector such that $\|v\|=1$. By Lemma A.2(iv), with probability $1-o\left(N^{-1}\right)$ we have

$$\begin{aligned}
\min_{1\leq i\leq N}\inf_{\|v\|=1}v'T\hat{Q}_{i,\tilde{x}_2\tilde{x}_2}v &= \min_{1\leq i\leq N}\inf_{\|v\|=1}\left(v'\Sigma_{22,i}v+v'\left(T\hat{Q}_{i,\tilde{x}_2\tilde{x}_2}-\Sigma_{22,i}\right)v\right)\\
&\geq \min_{1\leq i\leq N}\inf_{\|v\|=1}v'\Sigma_{22,i}v_1-\max_{1\leq i\leq N}\left\|T\hat{Q}_{i,\tilde{x}_2\tilde{x}_2}-\Sigma_{22,i}\right\|\\
&\geq \min_{1\leq i\leq N}\lambda_{\min}(\Sigma_{22,i})-o\left(1\right)\geq \underline{c}_{22}/2.
\end{aligned}$$

(iii) Note that $D_T\hat{Q}_{i,\tilde{x}\tilde{x}}D_T=\begin{pmatrix}\frac{1}{T^2}\sum_{t=1}^T\tilde{x}_{1,it}\tilde{x}_{1,it}' & \frac{1}{T^{3/2}}\sum_{t=1}^T\tilde{x}_{1,it}\tilde{x}_{2,it}'\\ \frac{1}{T^{3/2}}\sum_{t=1}^T\tilde{x}_{2,it}\tilde{x}_{1,it}' & \frac{1}{T}\sum_{t=1}^T\tilde{x}_{2,it}\tilde{x}_{2,it}'\end{pmatrix}=\begin{pmatrix}\hat{Q}_{i,x_1x_1} & \sqrt{T}\hat{Q}_{i,x_1x_2}\\ \sqrt{T}\hat{Q}_{i,x_1x_2}' & T\hat{Q}_{i,x_2x_2}\end{pmatrix}$.
Let $v=(v_1',v_2')'$ be a $(p_1+p_2)\times 1$ vector such that $\|v\|=1$. Then by Lemmas A.2(iii)-(iv) and Assumptions A.2(i), A.2(iii), and A.3(iv), with probability $1-o\left(N^{-1}\right)$

$$\begin{aligned}
\min_{1\leq i\leq N}\inf_{\|v\|=1}v'D_T\hat{Q}_{i,\tilde{x}\tilde{x}}D_Tv &= \min_{1\leq i\leq N}\inf_{\|v\|=1}\left(v_1'\hat{Q}_{i,\tilde{x}_1\tilde{x}_1}v_1+v_2'\hat{Q}_{i,\tilde{x}_2\tilde{x}_2}v_2+2v_1'\hat{Q}_{i,\tilde{x}_1\tilde{x}_2}v_2\right)\\
&\geq \min_{1\leq i\leq N}\inf_{\|v\|=1}\left(v_1'\hat{Q}_{i,\tilde{x}_1\tilde{x}_1}v_1+v_2'T\hat{Q}_{i,\tilde{x}_2\tilde{x}_2}v_2\right)-2\max_{1\leq i\leq N}\left\|\sqrt{T}\hat{Q}_{i,\tilde{x}_1\tilde{x}_2}\right\|\\
&\geq \min_{1\leq i\leq N}\left[\min\lambda_{\min}(\hat{Q}_{i,\tilde{x}_1\tilde{x}_1}),\ \lambda_{\min}(T\hat{Q}_{i,\tilde{x}_2\tilde{x}_2})\right]-2\max_{1\leq i\leq N}\left\|\sqrt{T}\hat{Q}_{i,\tilde{x}_1\tilde{x}_2}\right\|\\
&\geq \underline{c}_{11}/(2b_T).
\end{aligned}$$

Then (iii) follows. ■

Proof of Lemma A.4. (i) Noting that $\frac{1}{T^2}\tilde{x}_{1,i}'M_{2,i}\tilde{x}_{1,i}-\frac{1}{T^2}\tilde{x}_{1,i}'\tilde{x}_{1,i}=-T(\frac{1}{T^2}\tilde{x}_{1,i}'\tilde{x}_{2,i})(\frac{1}{T}\tilde{x}_{2,i}'\tilde{x}_{2,i})^{-1}$ $\times(\frac{1}{T^2}\tilde{x}_{2,i}'\tilde{x}_{1,i})$, it suffices to show that

$$\max_{1\leq i\leq N}\left\|T(\frac{1}{T^2}\tilde{x}_{1,i}'\tilde{x}_{2,i})(\frac{1}{T}\tilde{x}_{2,i}'\tilde{x}_{2,i})^{-1}(\frac{1}{T^2}\tilde{x}_{2,i}'\tilde{x}_{1,i})\right\|$$

is $o(1)$ with probability $1-o\left(N^{-1}\right)$. This follows because by Lemma A.2(iii)-(iv) and Assumptions A.2(iii) and A.3(iv), with probability $1-o\left(N^{-1}\right)$ we have

$$\begin{aligned}
\max_{1\leq i\leq N}\left\|T(\frac{1}{T^2}\tilde{x}_{1,i}'\tilde{x}_{2,i})(\frac{1}{T}\tilde{x}_{2,i}'\tilde{x}_{2,i})^{-1}(\frac{1}{T^2}\tilde{x}_{2,i}'\tilde{x}_{1,i})\right\| &\leq T\max_{1\leq i\leq N}\left\|\frac{1}{T^2}\tilde{x}_{1,i}'\tilde{x}_{2,i}\right\|^2\left[\min_{1\leq i\leq N}\lambda_{\min}\left(\frac{1}{T}\tilde{x}_{2,i}'\tilde{x}_{2,i}\right)\right]^{-1}\\
&= To\left(p_2a_{1NT}^2\right)O\left(1\right)=o\left(b_T^{-1}\right).
\end{aligned}$$

(ii) Noting that $\frac{1}{T}\tilde{x}_{2,i}'M_{1,i}\tilde{x}_{2,i} = \frac{1}{T}\tilde{x}_{2,i}'\tilde{x}_{2,i} - Tb_T(\frac{1}{T^2}\tilde{x}_{2,i}'\tilde{x}_{1,i})(\frac{b_T}{T^2}\tilde{x}_{1,i}'\tilde{x}_{1,i})^{-1}(\frac{1}{T^2}\tilde{x}_{1,i}'\tilde{x}_{2,i})$, the result follows from Lemmas A.2(iii)-(iv) and Assumption A.2(i) and the fact that $Tb_T(\sqrt{p_2}a_{1NT})^2 = O(p_2 a_{2NT})$. The detailed arguments are analogous to those used in the proof of (iii) below.

(iii) Note that $\frac{1}{T}\tilde{x}_{1,i}'M_{2,i}\tilde{u}_i^* = \frac{1}{T}\tilde{x}_{1,i}'\tilde{u}_i^* - \frac{1}{T}\tilde{x}_{1,i}'\tilde{x}_{2,i}(\frac{1}{T}\tilde{x}_{2,i}'\tilde{x}_{2,i})^{-1}\frac{1}{T}\tilde{x}_{2,i}'\tilde{u}_i^*$. By Lemma A.2(v), $P(\max_{1\leq i\leq N} ||\frac{1}{T^2}\tilde{x}_{1,i}'\tilde{u}^*|| > ca_{1NT}/2) = o\left(N^{-1}\right)$. Define the following two events:

$$E_{1NT} = \left\{\min_{1\leq i\leq N}\lambda_{\min}\left(\frac{1}{T}\tilde{x}_{2,i}'\tilde{x}_{2,i}\right) \geq \underline{c}_{22}/2\right\} \text{ and } E_{2NT} = \left\{\max_{1\leq i\leq N}\left\|\frac{1}{T^2}\tilde{x}_{1,i}'\tilde{x}_{2,i}\right\|_{\text{sp}} \leq p_2^{1/2}a_{1NT}\right\}.$$

By Lemma A.2(iii)-(iv), $P\left(E_{lNT}\right) = 1 - o\left(N^{-1}\right)$ for $l = 1, 2$. Denote the complement of $E_{\ell NT}$ as $E_{\ell NT}^c$ for $\ell = 1, 2$. Then, in view of the fact that $\|AB\| \leq \|A\|_{\text{sp}}\|B\|$ and $\|A\|_{\text{sp}} \leq \|A\|$ for any two conformable matrices A and B, we have

$$\begin{aligned}
& P\left(\max_{1\leq i\leq N}\left\|\frac{1}{T^2}\tilde{x}_{1,i}'\tilde{x}_{2,i}\left(\frac{1}{T}\tilde{x}_{2,i}'\tilde{x}_{2,i}\right)^{-1}\left(\frac{1}{T}\tilde{x}_{2,i}'\tilde{u}_i^*\right)\right\| > ca_{1NT}/2,\right) \\
\leq\ & P\left(\max_{1\leq i\leq N}\left\|\frac{1}{T^2}\tilde{x}_{1,i}'\tilde{x}_{2,i}\right\|_{\text{sp}}\left\|\frac{1}{T}\tilde{x}_{2,i}'\tilde{u}_i^*\right\|\left[\min_{1\leq i\leq N}\lambda_{\min}\left(\frac{1}{T}\tilde{x}_{2,i}'\tilde{x}_{2,i}\right)\right]^{-1} > ca_{1NT}/2, E_{1NT}\cap E_{2NT}\right) \\
& + P\left(E_{1NT}^c \cup E_{2NT}^c\right) \\
\leq\ & P\left(\max_{1\leq i\leq N}\left\|\frac{1}{T}\tilde{x}_{2,i}'\tilde{u}_i^*\right\| > c\cdot\underline{c}_{22}p_2^{-1/2}/4\right) + o\left(N^{-1}\right) \\
=\ & o\left(N^{-1}\right) + o\left(N^{-1}\right) = o\left(N^{-1}\right),
\end{aligned}$$

where the first equality follows by Lemma A.2(vi) and the fact that $p_2^{1/2}a_{2NT} = o(p_2^{-1/2})$. Consequently, the result in (iii) follows.

(iv) Note that $\frac{1}{T}\tilde{x}_{2,i}'M_{1,i}\tilde{u}_i^* = \frac{1}{T}\tilde{x}_{2,i}'\tilde{u}_i^* - T(\frac{1}{T^2}\tilde{x}_{2,i}'\tilde{x}_{1,i})(\frac{1}{T^2}\tilde{x}_{1,i}'\tilde{x}_{1,i})^{-1}(\frac{1}{T^2}\tilde{x}_{1,i}'\tilde{u}_i^*) \equiv I_{1i} - I_{2i}$, say. By Lemma A.2(vi), $P(\max_{1\leq i\leq N}\|I_{1i}\| \geq cp_2^{1/2}a_{2NT}/2) = o\left(N^{-1}\right)$. Noting that

$$\|I_{2i}\| \leq b_T\left\|\frac{1}{T^2}\tilde{x}_{2,i}'\tilde{x}_{1,i}\right\|_{\text{sp}}\left[\lambda_{\min}\left(\frac{b_T}{T^2}\tilde{x}_{1,i}'\tilde{x}_{1,i}\right)\right]^{-1}\left\|\frac{1}{T}\tilde{x}_{1,i}'\tilde{u}_i^*\right\|$$

and $b_T a_{1NT} = o(1)$, we can readily apply Lemmas A.2(iii), A.2(v), and A.3(i) to show that $P(\max_{1\leq i\leq N} \|I_{2i}\| \geq cp_2^{1/2}a_{2NT}/2) = o\left(N^{-1}\right)$. ■

Proof of Lemma A.5. (i) Noting that $\tilde{\beta}_{1,i} - \beta_{1,i}^0 = \left(\frac{b_T}{T^2}\tilde{x}_{1,i}'M_{2,i}\tilde{x}_{1,i}\right)^{-1}\frac{b_T}{T^2}\tilde{x}_{1,i}'M_{2,i}\tilde{u}_i^*$, the result follows from Lemmas A.4(i) and (iii), and Assumption A.2(i).

(ii) Noting that

$$\begin{aligned}
\tilde{\beta}_{2,i} - \beta_{2,i}^* &= \left(\frac{1}{T}\tilde{x}_{2,i}'M_{1,i}\tilde{x}_{2,i}\right)^{-1}\frac{1}{T}\tilde{x}_{2,i}'M_{1,i}\tilde{u}_i^* \\
&= \left[\left(\frac{1}{T}\tilde{x}_{2,i}'M_{1,i}\tilde{x}_{2,i}\right)^{-1} - \Sigma_{22,i}^{-1}\right]\frac{1}{T}\tilde{x}_{2,i}'M_{1,i}\tilde{u}_i^* + \Sigma_{22,i}^{-1}\frac{1}{T}\tilde{x}_{2,i}'M_{1,i}\tilde{u}_i^*,
\end{aligned}$$

the result follows from Lemma A.4(ii) and (iv) and Assumption A.2(iii).

(iii) Let $\beta_i^* = (\beta_{1,i}^{0\prime}, \beta_{2,i}^{*\prime})'$, which is $\beta_i^0 = (\beta_{1,i}^{0\prime}, \beta_{2,i}^{0\prime})'$ if $\Sigma_{20,i} = 0$. Noting that $\tilde{y}_{it} = \tilde{x}_{1,it}'\beta_{1,i}^0 + \tilde{x}_{2,it}'\beta_{2,i}^0 +$

$\tilde{u}_{it} = \tilde{x}_{it}'\beta_i^* + \tilde{u}_{it}^*$ with $\tilde{u}_{it}^* = \tilde{u}_{it} - \tilde{x}_{2,it}'\Sigma_{22,i}^{-1}\Sigma_{20,i}$, we have

$$\begin{aligned}
\tilde{\sigma}_i^2 &= \frac{1}{T}\sum_{t=1}^{T}[\tilde{y}_{it} - \tilde{\beta}_i'\tilde{x}_{it}]^2 = \frac{1}{T}\sum_{t=1}^{T}[\tilde{u}_{it}^* + \tilde{x}_{it}'(\beta_i^* - \tilde{\beta}_i)]^2 \\
&= \frac{1}{T}\sum_{t=1}^{T}(\tilde{u}_{it}^*)^2 + (\beta_i^* - \tilde{\beta}_i)'\frac{1}{T}\sum_{t=1}^{T}\tilde{x}_{it}\tilde{x}_{it}'(\beta_i^* - \tilde{\beta}_i) + 2(\beta_i^* - \tilde{\beta}_i)'\frac{1}{T}\sum_{t=1}^{T}\tilde{x}_{it}\tilde{u}_{it}^* \\
&\equiv D_{1i} + D_{2i} + 2D_{3i}, \text{ say.}
\end{aligned}$$

We prove (i) by showing that (i1) $P\left(\max_{1\le i\le N}\left|D_{1i} - \Sigma_{0.2,i}^*\right| > \epsilon\right) = o\left(N^{-1}\right)$, (i2) $P\left(\max_{1\le i\le N}|D_{2i}| > \epsilon\right) = o\left(N^{-1}\right)$, and (i3) $P\left(\max_{1\le i\le N}|D_{3i}| > \epsilon\right) = o\left(N^{-1}\right)$ for any $\epsilon > 0$. Noting that

$$\begin{aligned}
D_{1i} - \Sigma_{0.2,i}^* &= \frac{1}{T}\sum_{t=1}^{T}\left(\tilde{u}_{it} - \tilde{x}_{2,it}'\Sigma_{22,i}^{-1}\Sigma_{20,i}\right)^2 - \left(\Sigma_{00,i} - \Sigma_{02,i}\Sigma_{22,i}^{-1}\Sigma_{20,i}\right) \\
&= \frac{1}{T}\sum_{t=1}^{T}\left[u_{it}^2 - E\left(u_{it}^2\right)\right] - \overline{u}_i^2 + \Sigma_{02,i}\Sigma_{22,i}^{-1}\left(\frac{1}{T}\sum_{t=1}^{T}\tilde{x}_{2,it}\tilde{x}_{2,it}' - \Sigma_{22,i}\right)\Sigma_{22,i}^{-1}\Sigma_{20,i} \\
&\quad -2\left(\frac{1}{T}\sum_{t=1}^{T}\tilde{u}_{it}\tilde{x}_{2,it}' - \Sigma_{02,i}\right)\Sigma_{22,i}^{-1}\Sigma_{20,i} \\
&\equiv D_{1i,1} + D_{1i,2} + D_{1i,3} + D_{1i,4}.
\end{aligned}$$

By a simple application of Lemma S1.2 in SSPb, we can show that $P\left(\max_{1\le i\le N}|D_{1i,\ell}| > \epsilon/4\right) = o\left(N^{-1}\right)$ for $\ell = 1, 2$. By Lemma A.2(iv) and Assumption A.2(iv), $P\left(\max_{1\le i\le N}|D_{1i,3}| > \epsilon/4\right) = o\left(N^{-1}\right)$. By Lemma A.2(ii) and Assumption A.2(iv), $P\left(\max_{1\le i\le N}|D_{1i,4}| > \epsilon/4\right) = o\left(N^{-1}\right)$. It follows that

$$P\left(\max_{1\le i\le N}\left|D_{1i} - \Sigma_{0.2,i}^*\right| > \epsilon\right) = o(N^{-1}).$$

For D_{2i}, we have by the Cauchy-Schwarz inequality

$$\begin{aligned}
D_{2i} &\le 2(\beta_{1,i}^0 - \tilde{\beta}_{1,i})'\frac{1}{T}\sum_{t=1}^{T}\tilde{x}_{1,it}\tilde{x}_{1,it}'(\beta_{1,i}^0 - \tilde{\beta}_{1,i}) + 2(\beta_{2,i}^* - \tilde{\beta}_{2,i})'\frac{1}{T}\sum_{t=1}^{T}\tilde{x}_{2,it}\tilde{x}_{2,it}'(\beta_{2,i}^* - \tilde{\beta}_{2,i}) \\
&\equiv 2D_{2i,1} + 2D_{2i,2}.
\end{aligned}$$

With probability $1 - o\left(N^{-1}\right)$, $D_{2i,1}$ is bounded above by

$$2T\log\log T\max_{1\le i\le N}\left\|\beta_{1,i}^0 - \tilde{\beta}_{1,i}\right\|^2\max_{1\le i\le N}\left\|\frac{1}{2T^2\log\log T}\sum_{t=1}^{T}\tilde{x}_{1,it}\tilde{x}_{1,it}'\right\|_{\text{sp}} = o\left(T\log\log Tb_T^2a_{1NT}^2\right) = o\left(1\right),$$

by Lemma A.3(i) and part (i) and Assumption A.3(iii). And $D_{2i,2}$ bounded above by

$$\max_{1\le i\le N}\left\|\beta_{2,i}^* - \tilde{\beta}_{2,i}\right\|^2\max_{1\le i\le N}\left\|\frac{1}{T}\sum_{t=1}^{T}\tilde{x}_{2,it}\tilde{x}_{2,it}'\right\|_{\text{sp}} = o(p_2a_{2NT}^2) = o\left(1\right),$$

by Lemma A.2(iv), Assumption A.2(iii), and part (ii). It follows that $P\left(\max_{1\le i\le N}|D_{2i}| > \epsilon\right) = o\left(N^{-1}\right)$.

Similarly, with probability $1-o\left(N^{-1}\right)$,

$$\begin{aligned}
|D_{3i}| &\leq \left|(\beta_{1,i}^{0}-\tilde{\beta}_{1,i})'\frac{1}{T}\sum_{t=1}^{T}\tilde{x}_{1,it}\tilde{u}_{it}^{*}\right|+\left|(\beta_{2,i}^{*}-\tilde{\beta}_{2,i})'\frac{1}{T}\sum_{t=1}^{T}\tilde{x}_{2,it}\tilde{u}_{it}^{*}\right| \\
&\leq T\left\|\beta_{1,i}^{0}-\tilde{\beta}_{1,i}\right\|\left\|\frac{1}{T^{2}}\sum_{t=1}^{T}\tilde{x}_{1,it}\tilde{u}_{it}^{*}\right\|+\left\|\beta_{2,i}^{*}-\tilde{\beta}_{2,i}\right\|\left\|\frac{1}{T}\sum_{t=1}^{T}\tilde{x}_{2,it}\tilde{u}_{it}^{*}\right\| \\
&= To\left(b_{T}a_{1NT}\right)o\left(a_{1NT}\right)+o(p_{2}^{1/2}a_{2NT})o(p_{2}^{1/2}a_{2NT})=o\left(1\right),
\end{aligned}$$

by Lemma A.2(v)-(iv), parts (i)-(ii), and Assumption A.3(iii). It follows that $P\left(\max_{1\leq i\leq N}|D_{2i}|>\epsilon\right)=o\left(N^{-1}\right)$. ■

Proof of Lemma A.6. (i) Noting that $\frac{1}{T^{2}}\tilde{x}_{1,i}'\tilde{u}_{i}^{*}=\frac{1}{T^{2}}\tilde{x}_{1,i}'\tilde{u}_{i}-\frac{1}{T^{2}}\tilde{x}_{1,i}'\tilde{x}_{2,i}\Sigma_{22,i}^{-1}\Sigma_{20,i}$, it suffices to show $\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{1}{T^{2}}\tilde{x}_{1,i}'\tilde{u}_{i}\right\|^{2}=O_{P}\left(T^{-2}\right)$ and $\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{1}{T^{2}}\tilde{x}_{1,i}'\tilde{x}_{2,i}\Sigma_{22,i}^{-1}\Sigma_{20,i}\right\|^{2}=O_{P}\left(T^{-2}\right)$. We only show the former one as the proof of the latter claim is similar under the side condition $\left\|\Sigma_{22,i}^{-1}\Sigma_{20,i}\right\|\leq C<\infty$, which is ensured by Assumption A.2(ii)-(iii). By equation (B.1) and the Cauchy-Schwarz inequality

$$\begin{aligned}
\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{1}{T^{2}}\tilde{x}_{1,i}'\tilde{u}_{i}\right\|^{2} &\leq \frac{2}{N}\sum_{i=1}^{N}\left\|\frac{1}{T^{2}}x_{1,i}'u_{i}\right\|^{2}+\frac{2}{NT^{2}}\sum_{i=1}^{N}\left\|\bar{x}_{1,i}\bar{u}_{i}\right\|^{2} \\
&\leq \frac{6}{N}\sum_{i=1}^{N}\left(\|b_{1i}\|^{2}+\|b_{2i}\|^{2}+\|b_{3i}\|^{2}\right)+\frac{2}{NT^{2}}\sum_{i=1}^{N}\left\|\bar{x}_{1,i}\bar{u}_{i}\right\|^{2} \\
&\equiv 6d_{1}+6d_{2}+6d_{3}+2d_{4},\ \text{say.}
\end{aligned}$$

For d_1, we have

$$\begin{aligned}
E(d_{1}) &= \frac{1}{N}\sum_{i=1}^{N}E\left(\left\|S_{1}\frac{1}{T^{2}}\sum_{t=1}^{T}\sum_{s=1}^{t-1}\varepsilon_{is}e_{it}'\psi_{i0}'S_{0}'\right\|^{2}\right) \\
&= \frac{1}{NT^{4}}\sum_{i=1}^{N}E\left(\left\|\sum_{t=1}^{T}z_{it}\right\|^{2}\right)=\frac{1}{NT^{4}}\sum_{i=1}^{N}\sum_{t=1}^{T}E\left(z_{it}^{2}\right) \\
&\leq \frac{C}{NT^{4}}\sum_{i=1}^{N}\sum_{t=1}^{T}t=O\left(T^{-2}\right),
\end{aligned}$$

where $z_{it}=S_{1}\sum_{s=1}^{t-1}\varepsilon_{is}e_{it}'\psi_{i0}'S_{0}'$ satisfies $E\left(z_{it}|\mathcal{F}_{i,t-1}\right)=0$ and the first inequality follows because we can show that $E\left(z_{it}^{2}\right)\leq Ct$. For d_2, we can follow the analysis of $b_{2,i}$ in the proof of Lemma A.2 and show that $E(d_{2})\leq\frac{4}{N}\sum_{i=1}^{N}E[\|b_{2i,1}\|^{2}+\|b_{2i,2}\|^{2}+\|b_{2i,3}\|^{2}+\|b_{2i,4}\|^{2}]=O(T^{-2})$. In addition,

$$\begin{aligned}
E(d_{3}) &= \frac{1}{N}\sum_{i=1}^{N}E\left\|S_{1}\frac{1}{T^{2}}\sum_{t=1}^{T}\varepsilon_{it}\varepsilon_{it}'S_{0}'\right\|^{2}\leq\frac{C}{NT^{2}}\sum_{i=1}^{N}E\left\{\frac{1}{T}\sum_{t=1}^{T}\|\varepsilon_{it}\|^{2}\right\}^{2} \\
&\leq \frac{C}{NT^{2}}\sum_{i=1}^{N}\frac{1}{T}\sum_{t=1}^{T}E\left\|\varepsilon_{it}\right\|^{4}=O\left(T^{-2}\right),\ \text{and} \\
E(d_{4}) &= \frac{1}{NT^{2}}\sum_{i=1}^{N}E\left\|\bar{x}_{1,i}\bar{u}_{i}\right\|^{2}\leq\frac{1}{NT^{2}}\sum_{i=1}^{N}\left\{E(\|\bar{x}_{1,i}\|^{4})E\left(\bar{u}_{i}^{4}\right)\right\}^{1/2}=O\left(T^{-2}\right),
\end{aligned}$$

where the last equality follows from the fact that $E(\bar{u}_i^4) \leq CT^{-2}$ and $E(\|\bar{x}_{1,i}\|^4) \leq CT^2$. Consequently, $\frac{1}{N}\sum_{i=1}^N E\left\|\frac{1}{T^2}\tilde{x}_{1,i}'\bar{u}_i\right\|^2 = O\left(T^{-2}\right)$ and $\frac{1}{N}\sum_{i=1}^N \left\|\frac{1}{T^2}\tilde{x}_{1,i}'\bar{u}_i\right\|^2 = O_P\left(T^{-2}\right)$ by the Markov inequality.

(ii) Noting that

$$\begin{aligned}
\frac{1}{T^{3/2}}\tilde{x}_{2,i}'\tilde{u}_i^* &= \frac{1}{T^{3/2}}\tilde{x}_{2,i}'\tilde{u}_i - \frac{1}{T^{3/2}}\tilde{x}_{2,i}'\tilde{x}_{2,i}\Sigma_{22,i}^{-1}\Sigma_{20,i} \\
&= \frac{1}{T^{3/2}}\left(\tilde{x}_{2,i}'\tilde{u}_i - \Sigma_{20,i}\right) - \frac{1}{T^{3/2}}\left(\tilde{x}_{2,i}'\tilde{x}_{2,i} - \Sigma_{22,i}\right)\Sigma_{22,i}^{-1}\Sigma_{20,i} \\
&= \frac{1}{T^{3/2}}\left(x_{2,i}'u_i - \Sigma_{20,i}\right) - \frac{1}{T^{3/2}}\left(x_{2,i}'x_{2,i} - \Sigma_{22,i}\right)\Sigma_{22,i}^{-1}\Sigma_{20,i} - \frac{1}{T^{1/2}}\bar{x}_{2,i}'\bar{u}_i \\
&\quad + \frac{1}{T^{1/2}}\bar{x}_{2,i}'\bar{x}_{2,i}\Sigma_{22,i}^{-1}\Sigma_{20,i} \equiv d_{1i} + d_{2i} + d_{3i} + d_{4i}, \text{ say,}
\end{aligned}$$

it suffices to show $\frac{1}{N}\sum_{i=1}^N \|d_{\ell i}\|^2 = O_P\left(p_2 T^{-2}\right)$ for $\ell = 1,2,3,4$. We can prove these by the Markov inequality. Then (ii) follows.

(iii) By the BN decomposition $x_{1,it} = S_1[\psi_i(1)\sum_{s=1}^t e_{it} + \breve{e}_{i,0} - \breve{e}_{i,t}]$ and the Cauchy-Schwarz inequality,

$$\begin{aligned}
\frac{1}{N}\sum_{i=1}^N E\left\|\hat{Q}_{1i}\right\|^2 &\leq \frac{1}{N}\sum_{i=1}^N E\left\|\frac{1}{T^2}\sum_{t=1}^T x_{1,it}x_{1,it}'\right\|^2 \\
&\leq \frac{1}{N}\sum_{i=1}^N E\left\|\frac{1}{T^2}\sum_{t=1}^T S_1\psi_i(1)\sum_{s=1}^t e_{it}\sum_{r=1}^t e_{ir}'\psi_i(1)' S_1'\right\|^2 \\
&\quad + \frac{1}{N}\sum_{i=1}^N E\left\|\frac{1}{T^2}\sum_{t=1}^T S_1[(\breve{e}_{i,0} - \breve{e}_{i,t})(\breve{e}_{i,0} - \breve{e}_{i,t})' S_1'\right\|^2.
\end{aligned}$$

By straightforward moment calculation, we can bound the first term in the last expression by $O(1)$ and the second term by $O\left(T^{-2}\right)$. Then $\frac{1}{N}\sum_{i=1}^N \left\|\hat{Q}_{1i}\right\|^2 = O_P(1)$ by the Markov inequality.

(iv) The proof is analogous to that of (i) and thus omitted.

(v) Noting that $\tilde{x}_{1,i}'M_{2,i}\tilde{u}_i^* = \tilde{x}_{1,i}'\tilde{u}_i^* - \tilde{x}_{1,i}'\tilde{x}_{2,i}(\tilde{x}_{2,i}'\tilde{x}_{2,i})^{-1}\tilde{x}_{2,i}'\tilde{u}_i^*$, we have by Lemmas A.2(ii), A.2(iv), part (iv), and Assumption A.3(iii)

$$\begin{aligned}
\frac{1}{N}\sum_{i=1}^N \left\|\frac{1}{T^2}\tilde{x}_{1,i}'M_{2,i}\tilde{u}_i^*\right\|^2 &\leq \frac{2}{N}\sum_{i=1}^N \left\|\frac{1}{T^2}\tilde{x}_{1,i}'\tilde{u}_i^*\right\|^2 + \frac{2}{NT^2}\sum_{i=1}^N \left\|\frac{1}{T^2}\tilde{x}_{1,i}'\tilde{x}_{2,i}\left(\frac{1}{T}\tilde{x}_{2,i}'\tilde{x}_{2,i}\right)^{-1}\frac{1}{T}\tilde{x}_{2,i}'\tilde{u}_i^*\right\|^2 \\
&\leq O_P\left(T^{-2}\right) + \max_i \left\|\left(\frac{1}{T}\tilde{x}_{2,i}'\tilde{x}_{2,i}\right)^{-1}\right\|_{\text{sp}} \max_i \left\|\frac{1}{T}\tilde{x}_{2,i}'\tilde{u}_i^*\right\|^2 \frac{2}{N}\sum_{i=1}^N \left\|\frac{1}{T^2}\tilde{x}_{1,i}'\tilde{x}_{2,i}\right\|^2 \\
&\leq O_P\left(T^{-2}\right) + O_P(1)\, o_P(p_2^{1/2}a_{2NT})O_P\left(p_2T^{-2}\right) = O_P\left(T^{-2}\right). \blacksquare
\end{aligned}$$

Proof of Lemma A.7. (i) Note that $Q_{k,NT} = \frac{1}{N_kT^2}\sum_{i\in G_k^0}\tilde{x}_{1,i}'\tilde{x}_{1,i} - \frac{1}{N_kT^2}\sum_{i\in G_k^0}\left(\tilde{x}_{1,i}'\tilde{x}_{2,i}\right)\left(\tilde{x}_{2,i}'\tilde{x}_{2,i}\right)^{-1}$ $\times\left(\tilde{x}_{2,i}'\tilde{x}_{1,i}\right) \equiv Q_{1k,NT} - Q_{2k,NT}$. By Lemmas A.2(ii)-(iii), $\|Q_{2k,NT}\| \leq T\max_{i\in G_k^0}\left\|\frac{1}{T^2}\tilde{x}_{1,i}'\tilde{x}_{2,i}\right\|^2 \left\|(\frac{1}{T}\tilde{x}_{2,i}'\tilde{x}_{2,i})^{-1}\right\|_{\text{sp}}$

$= To_P\left(p_2 a_{1NT}^2\right) = o_P(1)$. By the arguments used in Phillips and Moon (1999, Section 4), we can show that

$$
\begin{aligned}
Q_{1k,NT} &= \frac{1}{N_k}\sum_{i\in G_k^0} E\left(\int_0^1 \tilde{B}_{1,i}\tilde{B}_{1,i}'\right) + o_P(1) = \frac{1}{N_k}\sum_{i\in G_k^0} S_1\psi_i(1)\, E\left(\int_0^1 \tilde{W}_{1,i}\tilde{W}_{1,i}'\right)\psi_i(1)' S_1' + o_P(1)\\
&= \frac{1}{6N_k}\sum_{i\in G_k^0} S_1\psi_i(1)\psi_i(1)' S_1' + o_P(1),
\end{aligned}
$$

where we use the fact that $E(\int_0^1 \tilde{W}_{1,i}\tilde{W}_{1,i}') = E(\int_0^1 W_{1,i}W_{1,i}') - E(\int_0^1 W_{1,i}\int_0^1 W_{1,i}') = \frac{1}{2}I_{p_1} - \frac{1}{3}I_{p_1} = \frac{1}{6}I_{p_1}$.Thus (i) follows.

(ii) Let $E_{it} \equiv \sum_{s=1}^t e_{is}$, $e_{it}^* \equiv e_{it} - \frac{1}{T}\sum_{s=t}^T e_{is}$, and $p = 1 + p_1 + p_2$. Note that $\tilde{e}_{it} = e_{it} - \bar{e}_i = e_{it}^* - \frac{1}{T}E_{i,t-1}$. We apply the arguments as used in the proof of Theorem 16 in Phillips and Moon (1999, PM hereafter) and derive the limiting distribution of $V_{1k,NT}$ below.[1]

First, we apply the BN decompositions. Noting that $\varepsilon_{it} = \psi_i(1)e_{it} + \breve{e}_{i,t-1} - \breve{e}_{i,t}$, we have $x_{1,it} = S_1\left[\psi_i(1)E_{it} + \breve{e}_{i,0} - \breve{e}_{i,t}\right]$ and

$$
u_{it} - x_{2,it}'\Sigma_{22,i}^{-1}\Sigma_{20,i} = \left[\psi_i(1)e_{it} + \breve{e}_{i,t-1} - \breve{e}_{i,t}\right]'\left[S_0' - S_2'\Sigma_{22,i}^{-1}\Sigma_{20,i}\right] = \left[e_{it}'\psi_i(1)' + \breve{e}_{i,t-1}' - \breve{e}_{i,t}'\right]s_i,
$$

where $s_i = S_0' - S_2'\Sigma_{22,i}^{-1}\Sigma_{20,i}$ is a $p\times 1$ vector. It follows that we can write the demeaned versions of $x_{1,it}$ and $u_{it} - x_{2,it}'\Sigma_{22,i}^{-1}\Sigma_{20,i}$ as

$$
\tilde{x}_{1,it} = S_1\left[\psi_i(1)\tilde{E}_{it} + \tilde{\breve{e}}_{i,0} - \tilde{\breve{e}}_{i,t}\right] \text{ and } \tilde{u}_{it} - \tilde{x}_{2,it}'\Sigma_{22,i}^{-1}\Sigma_{20,i} = \left[\tilde{e}_{it}'\psi_i(1)' + \tilde{\breve{e}}_{i,t-1}' - \tilde{\breve{e}}_{i,t}'\right]s_i,
$$

where $\tilde{E}_{it} = E_{it} - \frac{1}{T}\sum_{s=1}^T E_{is}$ and $\tilde{\breve{e}}_{i,t} = \breve{e}_{i,t} - \frac{1}{T}\sum_{s=1}^T \tilde{\breve{e}}_{i,s}$. Let $S_{i,t}^{\varepsilon} = \sum_{s=1}^t \varepsilon_{is}$ and $\tilde{S}_{i,t}^{\varepsilon} = S_{i,t}^{\varepsilon} - \frac{1}{T}\sum_{s=1}^T S_{i,t}^{\varepsilon}$. As in PM (p.1105), we can obtaining the following decomposition,

$$
\begin{aligned}
\bar{V}_{k,NT} - \mathbb{B}_{k,NT} =& \frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0}\frac{1}{T}\sum_{t=1}^T \tilde{x}_{1,it}\left(\tilde{u}_{it} - \tilde{x}_{2,it}'\Sigma_{22,i}^{-1}\Sigma_{20,i}\right) - \mathbb{B}_{k,NT}\\
=& \frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0}\left\{\frac{1}{T}\sum_{t=1}^T S_1\psi_i(1)\left[\tilde{E}_{it}\tilde{e}_{it}' - \left(1 - \frac{T+1}{2T}\right)I_{p+1}\right]\psi_i(1)' s_i\right.\\
&+ \frac{1}{T}\sum_{t=1}^{T-1} S_1\left(\tilde{\varepsilon}_{i,t+1}\tilde{\breve{\varepsilon}}_{it}' + \sum_{s=0}^{\infty}\psi_{i,s+1}\breve{\psi}_{i,s}'\right)s_i - \frac{1}{T}\sum_{s=0}^{\infty} S_1\psi_{i,s+1}\breve{\psi}_{i,s}' s_i\\
&- \frac{1}{T}\sum_{t=1}^T S_1\psi_i(1)\left[\tilde{\breve{\varepsilon}}_{it}\tilde{e}_{it}'\psi_i(1)' - \breve{\psi}_{i,0}\psi_i(1)'\right]s_i + \frac{1}{T}\sum_{t=1}^T S_1\tilde{\breve{\varepsilon}}_{i0}\tilde{e}_{it}'\psi_i(1)' s_i\\
&\left. - \frac{1}{T}S_1\tilde{S}_{i,T}^{\varepsilon}\tilde{\breve{\varepsilon}}_{iT}' s_i + \frac{1}{T}S_1\tilde{S}_{i,1}^{\varepsilon}\tilde{\breve{\varepsilon}}_{i0} s_i\right\}\\
=& \frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0}\left\{Q_{i,T} + R_{1i,T} + R_{2i,T} + R_{3i,T} + R_{4i,T} + R_{5i,T} + R_{6i,T}\right\}, \text{ say.}
\end{aligned}
$$

Note that our terms $Q_{i,T}$ and $R_{\ell i,T}$ $(\ell = 1, 2, ..., 6)$ parallel the corresponding term in PM. There are three main differences: (1) all variables involved here are time-demeaned versions of those in PM; (2) we need to center $\tilde{E}_{it}\tilde{e}_{it}'$ around its expectation $(1 - \frac{T+1}{2T})I_{p+1}$ while PM centers the non-demeaned version of $\tilde{E}_{it}\tilde{e}_{it}'$ around its expectation I_{p+1}, and the difference between the two centering terms, namely, $-\frac{T+1}{2T}I_{p+1}$, enters the bias term $\mathbb{B}_{2k,NT}$ and reflects the contribution of time-demeaning of random variables in the regression; (3) the sign $R_{2i,T}$ is negative rather than positive. One can verify that $\sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\psi_{i,s+r}\psi_{i,s}' = \psi_i(1)\psi_i(1)' + \sum_{s=0}^{\infty}\psi_{i,s+1}\breve{\psi}_{i,s}' - \breve{\psi}_{i,0}\psi_i(1)'$.

Second, we study the asymptotic distribution of $N_k^{-1/2}\sum_{i\in G_k^0} Q_{i,T}$. Noting that $\frac{1}{T}\sum_{t=1}^{T}\tilde{E}_{it}\tilde{e}_{it}' = \frac{1}{T}\sum_{t=1}^{T}E_{it}\tilde{e}_{it}'$, and $\tilde{e}_{it} = e_{it} - \frac{1}{T}E_{iT}$, we have

$$
\begin{aligned}
\frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0} Q_{i,T} &= \frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0}\frac{1}{T}\sum_{t=1}^{T} S_1\psi_i(1)\left[\tilde{E}_{it}\tilde{e}_{it}' - \left(1-\frac{T+1}{2T}\right)I_p\right]\psi_i(1)' s_i \\
&= \frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0}\frac{1}{T}\sum_{t=1}^{T} S_1\psi_i(1)E_{it-1}e_{it}'\psi_i(1)' s_i - \frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0}\frac{1}{T^2}\sum_{t=1}^{T} S_1\psi_i(1)\left(E_{it}E_{iT}' - tI_p\right)\psi_i(1)' s_i \\
&+ \frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0}\frac{1}{T}\sum_{t=1}^{T} S_1\psi_i(1)\left[e_{it}e_{it}' - I_p\right]\psi_i(1)' s_i \\
&= \frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0}\left\{Q_{1i,T} - Q_{2i,T} + Q_{3i,T}\right\}, \text{ say.}
\end{aligned}
$$

By direct moment calculations, we can readily show that

$$
\left\|E\left(\frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0} Q_{3i,T}\right)\right\| \le \frac{1}{T}\frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0}\left\|S_1\psi_i(1)\psi_i(1)' s_i\right\| = O\left(\sqrt{N_k p_2}/T\right) = o(1)
$$

and $\left\|\mathrm{Var}\left(\frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0} Q_{3i,T}\right)\right\| = \left\|\frac{1}{N_k}\sum_{i\in G_k^0}\mathrm{Var}\left(Q_{3i,T}\right)\right\| \le \frac{1}{N_k}\sum_{i\in G_k^0} E\left\|Q_{3i,T}\right\|^2 = O\left(p_2 T^{-2}\right) = o(1)$. Then $\frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0} Q_{3i,T} = o_P(1)$ by the Chebyshev inequality. It follows that

$$
\mathrm{Var}\left(\frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0} Q_{i,T}\right) = \frac{1}{N_k}\sum_{i\in G_k^0}\left\{\mathrm{Var}\left(Q_{1i,T}\right) + \mathrm{Var}\left(Q_{2i,T}\right) - \mathrm{Cov}\left(Q_{1i,T}, Q_{2i,T}\right) - \mathrm{Cov}\left(Q_{1i,T}, Q_{2i,T}\right)'\right\}
$$

Then we study the asymptotic variance by terms. For $Q_{1i,T}$,

$$
\begin{aligned}
\mathrm{Var}\left(Q_{1i,T}\right) &= \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T} E\left[S_1\psi_i(1)E_{i,t-1}e_{it}'\psi_i(1)' s_i \; s_i'\psi_i(1)e_{is}E_{i,s-1}'\psi_i(1)' S_1'\right] \\
&= \frac{1}{T^2}\sum_{t=1}^{T} E\left\{S_1\psi_i(1)E_{i,t-1}e_{it}'\psi_i(1)' s_i \; s_i'\psi_i(1)e_{it}E_{i,t-1}'\psi_i(1)' S_1'\right\} \\
&= \frac{1}{T^2}\sum_{t=1}^{T}\left(s_i'\psi_i(1)\otimes S_1\psi_i(1)\right)E\left\{\mathrm{vec}\left(E_{i,t-1}e_{it}'\right) \quad \left[\mathrm{vec}\left(E_{i,t-1}e_{it}'\right)\right]' \right\}\left(s_i'\psi_i(1)\otimes S_1\psi_i(1)\right)' \\
&= \frac{1}{T^2}\sum_{t=1}^{T}\left(s_i'\psi_i(1)\otimes S_1\psi_i(1)\right)E\left(e_{it}e_{it}'\otimes E_{i,t-1}E_{i,t-1}'\right)\left(s_i'\psi_i(1)\otimes S_1\psi_i(1)\right)' \\
&= \frac{1}{T^2}\sum_{t=1}^{T}(t-1)\left(s_i'\psi_i(1)\otimes S_1\psi_i(1)\right)\left(s_i'\psi_i(1)\otimes S_1\psi_i(1)\right)'
\end{aligned}
$$

$$= \frac{1}{2} s_i' \psi_i (1) \psi_i (1)' s_i \otimes S_1 \psi_i (1) \psi_i (1)' S_1' + O\left(T^{-1}\right) = \frac{1}{2} s_i' \Omega_i s_i S_1 \Omega_i S_1' + O\left(T^{-1}\right),$$

where the second equality follows from the fact that $\{E_{i,t-1}e_{it}', \mathcal{F}_{i,t}\}$ is a martingale difference sequence (m.d.s.), the third equality holds because $\text{vec}(A_1 A_2 A_3) = (A_3' \otimes A_1)\text{vec}(A_2)$ with $A_1 = S_1 \psi_i (1)$, $A_2 = E_{i,t-1}e_{it}'$, and $A_3 = \psi_i (1)' s_i$, the fourth equality follows from the fact that $\text{vec}(a_1 a_2') = a_2 \otimes a_1$ and $(a_2 \otimes a_1)(a_2 \otimes a_1)' = a_2 a_2' \otimes a_1 a_1'$, the fifth equality holds because $E\left(e_{it}e_{it}' \otimes E_{i,t-1}E_{i,t-1}'\right) = E\left(e_{it}e_{it}'\right) \otimes E\left(E_{i,t-1}E_{i,t-1}'\right) = (t-1) I_{(1+p)^2}$. Similarly, we have for $Q_{2i,T}$,

$$\begin{aligned}
& \text{Var}\left(Q_{2i,T}\right) \\
= & \frac{1}{T^2} \sum_{t=1}^{T} \sum_{s=1}^{T} E\left[S_1 \psi_i (1) \left(E_{i,t}E_{i,T}' - t\right) \psi_i (1)' s_i \; s_i' \psi_i (1) \left(E_{i,s}E_{i,T}' - s\right) \psi_i (1)' S_1'\right] \\
= & \left(s_i' \psi_i (1) \otimes S_1 \psi_i (1)\right) \frac{1}{T^4} \sum_{t=1}^{T} \sum_{s=1}^{T} E[\text{vec}\left(E_{i,t}E_{i,T}'\right) \text{vec}\left(E_{i,s}E_{i,T}'\right)' - t\text{vec}\left(I_{1+p}\right) \text{vec}\left(E_{i,s}E_{i,T}'\right)' \\
& -s\text{vec}\left(E_{i,t}E_{i,T}'\right) \text{vec}\left(I_{1+p}\right)' + ts\text{vec}\left(I_{1+p}\right) \text{vec}\left(I_{1+p}\right)'] \left(s_i' \psi_i (1) \otimes S_1 \psi_i (1)\right)' \\
= & [s_i' \psi_i (1) \otimes S_1 \psi_i (1)] \frac{1}{T^4} \sum_{t=1}^{T} \sum_{s=1}^{T} E[\text{vec}\left(E_{i,t}E_{i,T}'\right) \text{vec}\left(E_{i,s}E_{i,T}'\right)' - ts\text{vec}\left(I_{1+p}\right) \text{vec}\left(I_{1+p}\right)'][s_i' \psi_i (1) \otimes S_1 \psi_i (1)]' \\
= & [s_i' \psi_i (1) \otimes S_1 \psi_i (1)] \frac{1}{T^4} \sum_{t=1}^{T} \sum_{s=1}^{T} E[\left(E_{i,T}E_{i,T}' \otimes E_{i,t}E_{i,s}'\right) - ts\text{vec}\left(I_{1+p}\right) \text{vec}\left(I_{1+p}\right)'][s_i' \psi_i (1) \otimes S_1 \psi_i (1)]'
\end{aligned}$$

and

$$\begin{aligned}
& \frac{1}{T^4} \sum_{t=1}^{T} \sum_{s=1}^{T} E\left(E_{i,T}E_{i,T}' \otimes E_{i,t}E_{i,s}'\right) = \frac{1}{T^4} \sum_{t=1}^{T} \sum_{s=1}^{T} \left[T (t \wedge s) I_{(p+1)^2} + ts \left(K_{1+p} + \text{vec}(I_{p+1})\text{vec}(I_{p+1})'\right)\right] \\
= & \left\{\frac{1}{3} I_{(1+p)^2} + \frac{1}{4}\left(K_{p+1} + \text{vec}(I_{p+1})\text{vec}(I_{p+1})'\right)\right\} + O\left(T^{-1}\right).
\end{aligned}$$

where K_{1+p} is the $(p+1)^2 \times (p+1)^2$ commutation matrix such that $K_{1+p}\text{vec}(A)$ =vec(A') for any $(p+1) \times (p+1)$ matrix A. It follows that

$$\text{Var}\left(Q_{2i,T}\right) = \left(s_i' \psi_i (1) \otimes S_1 \psi_i (1)\right) \left(\frac{1}{3} I_{(1+p)^2} + \frac{1}{4} K_{1+p}\right) \left(s_i' \psi_i (1) \otimes S_1 \psi_i (1)\right)' + O\left(T^{-1}\right).$$

For Cov$(Q_{1i,T}, Q_{2i,T})$,

$$\begin{aligned}
\text{Cov}\left(Q_{1i,T}, Q_{2i,T}\right) &= \frac{1}{T^3} \sum_{t=1}^{T} \sum_{s=1}^{T} E\left[S_1 \psi_i (1) E_{i,t-1}e_{it}' \psi_i (1)' s_i \; s_i' \psi_i (1) E_{i,s}E_{i,T}' \psi_i (1)' S_1'\right] \\
&= [s_i' \psi_i (1) \otimes S_1 \psi_i (1)] \frac{1}{T^3} \sum_{t=1}^{T} \sum_{s=1}^{T} E\left\{\text{vec}\left(E_{i,t-1}e_{it}'\right) \text{vec}\left(E_{i,s}E_{i,T}'\right)'\right\} [s_i' \psi_i (1) \otimes S_1 \psi_i (1)]' \\
&= [s_i' \psi_i (1) \otimes S_1 \psi_i (1)] \frac{1}{T^3} \sum_{t=1}^{T} \sum_{s=1}^{T} E\left(e_{it}E_{i,T}' \otimes E_{i,t-1}E_{i,s}'\right) [s_i' \psi_i (1) \otimes S_1 \psi_i (1)]' \\
&= [s_i' \psi_i (1) \otimes S_1 \psi_i (1)] \left\{\frac{1}{3} I_{(1+p)^2} + \frac{1}{6} K_{1+p}\right\} [s_i' \psi_i (1) \otimes S_1 \psi_i (1)]' + O\left(T^{-1}\right),
\end{aligned}$$

where we use the fact that

$$
\begin{aligned}
& \frac{1}{T^3}\sum_{t=1}^{T}\sum_{s=1}^{T}E\left(e_{it}E_{i,T}'\otimes E_{i,t-1}E_{i,s}'\right) \\
= & \frac{1}{T^3}\sum_{t=1}^{T}\sum_{s=1}^{t-1}E\left(e_{it}E_{i,T}'\otimes E_{i,t-1}E_{i,s}'\right)+\frac{1}{T^3}\sum_{t=1}^{T}\sum_{s=t}^{T}E\left(e_{it}E_{i,T}'\otimes E_{i,t-1}E_{i,s}'\right) \\
= & \frac{1}{T^3}\sum_{t=1}^{T}\sum_{s=1}^{t-1}E\left(e_{it}E_{i,T}'\otimes E_{i,t-1}E_{i,s}'\right) \\
& +\frac{1}{T^3}\sum_{t=1}^{T}\sum_{s=t}^{T}\left\{E\left(e_{it}E_{i,T}'\otimes E_{i,t-1}E_{i,t-1}'\right)+E[e_{it}E_{i,T}'\otimes E_{i,t-1}\left(E_{i,s}-E_{i,t-1}\right)']\right\} \\
= & \frac{1}{T^3}\sum_{t=1}^{T}\sum_{s=1}^{t-1}sI_{(1+p)^2}+\frac{1}{T^3}\sum_{t=1}^{T}\sum_{s=t}^{T}(t-1)\left(I_{(1+p)^2}+K_{1+p}\right) \\
= & \frac{1}{3}I_{(1+p)^2}+\frac{1}{6}K_{1+p}+O\left(T^{-1}\right).
\end{aligned}
$$

Thus we have

$$
\begin{aligned}
\operatorname{Var}\left(\frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0}Q_{i,T}\right) &= \frac{1}{N_k}\sum_{i\in G_k^0}\left\{\operatorname{Var}\left(Q_{1i,T}\right)+\operatorname{Var}\left(Q_{2i,T}\right)-2\operatorname{Cov}\left(Q_{1i,T},Q_{2i,T}\right)\right\} \\
&=\frac{1}{N_k}\sum_{i\in G_k^0}\left\{\frac{1}{6}s_i'\Omega_i s_i S_1\Omega_i S_1'-\frac{1}{12}[s_i'\psi_i\left(1\right)\otimes S_1\psi_i\left(1\right)]K_{1+p}[s_i'\psi_i\left(1\right)\otimes S_1\psi_i\left(1\right)]'\right\} \\
&\quad +O\left(T^{-1}\right) \\
&=\frac{1}{N_k}\sum_{i\in G_k^0}\left\{\frac{1}{6}s_i'\Omega_i s_i S_1\Omega_i S_1'-\frac{1}{12}\left(s_i'\Omega_i S_1'\otimes S_1\Omega_i s_i\right)K_{p_1,1}\right\}+O\left(T^{-1}\right),
\end{aligned}
$$

where $K_{p_1,1}$ is the $p_1\times p_1$ commutation matrix. It follows that $\operatorname{Var}(N_K^{-1/2}\sum_{i\in G_k^0}Q_{i,T})\to\lim_{N_k\to\infty}\frac{1}{N_k}\sum_{i\in G_k^0}$ $[\frac{1}{6}s_i'\Omega_i s_i S_1\Omega_i S_1'-\frac{1}{12}\left(s_i'\Omega_i S_1'\otimes S_1\Omega_i s_i\right)K_{p_1,1}]\equiv V_{(k)}$. This limit contributes to the asymptotic variance of our estimator. In addition, we can verify that $\sum_{i=1}^{N}E\left\|N_K^{-1/2}Q_{i,T}\right\|^4=O\left(N_k^{-1}\right)$, which verifies the Lyapunov condition for the central limit theorem for independent but non-identically distributed (i.n.i.d.) observations. Consequently, we have shown that $Q_{i,T}\Rightarrow N\left(0,\mathbb{V}_{(k)}\right)$. Third, we study $R_{1i,T}$:

$$
\begin{aligned}
\frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0}R_{1i,T} &= \frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0}\frac{1}{T}\sum_{t=1}^{T-1}S_1\left(\tilde{\varepsilon}_{i,t+1}\tilde{\breve{\varepsilon}}_{it}'-\sum_{s=0}^{\infty}\psi_{i,s+1}\breve{\psi}_{i,s}'\right)s_i \\
&= \frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0}\left\{\frac{1}{T}\sum_{t=1}^{T-1}S_1\left(\varepsilon_{i,t+1}\breve{\varepsilon}_{it}'-\sum_{s=0}^{\infty}\psi_{i,s+1}\breve{\psi}_{i,s}'\right)s_i-\frac{1}{T}\sum_{t=1}^{T-1}S_1\frac{1}{T}\sum_{s=1}^{T}\varepsilon_{is}\breve{\varepsilon}_{it}'s_i\right. \\
&\quad \left.-\frac{1}{T}\sum_{t=1}^{T-1}S_1\varepsilon_{i,t+1}\frac{1}{T}\sum_{r=1}^{T}\breve{\varepsilon}_{ir}'s_i+\frac{T-1}{T}S_1\left(\frac{1}{T}\sum_{s=1}^{T}\varepsilon_{is}\frac{1}{T}\sum_{r=1}^{T}\breve{\varepsilon}_{ir}'\right)s_i\right\} \\
&\equiv \frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0}\left\{R_{1i,T1}-R_{1i,T2}-R_{1i,T3}+R_{1i,T4}\right\}.
\end{aligned}
$$

Following PM, we can show that $E\left\|N_k^{-1/2}\sum_{i\in G_k^0}R_{1i,T1}\right\|^2=O\left(p_2T^{-1}\right)$. For $R_{1i,T2}$, we apply the Cauchy-Schwarz and Markov inequalities

$$\begin{aligned}\left\|\frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0}R_{1i,T2}\right\| &\le \sqrt{N_k}\left\{\frac{1}{N_k}\sum_{i\in G_k^0}\left\|S_1\frac{1}{T}\sum_{s=1}^{T}\varepsilon_{is}\right\|^2\right\}^{1/2}\left\{\frac{1}{N_k}\sum_{i\in G_k^0}\left\|\frac{1}{T}\sum_{r=1}^{T-1}\breve{\varepsilon}_{ir}'s_i\right\|^2\right\}^{1/2}\\ &= \sqrt{N_k}O_P(T^{-1/2})O_P(p_2^{1/2}T^{-1/2})=o_P(1),\end{aligned}$$

where we use the fact $\frac{1}{N_k}\sum_{i\in G_k^0}E\left\|S_1\frac{1}{T}\sum_{s=1}^{T}\varepsilon_{is}\right\|^2\le CT^{-1}\frac{1}{N_k}\sum_{i\in G_k^0}\mathrm{tr}(S_1\Omega_iS_1')\le CT^{-1}\mathrm{tr}(S_1S_1')=O\left(T^{-1}\right)$. Similarly, we can show that $\frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0}R_{1i,T\ell}=o_P(1)$ for $\ell=3,4$. Thus we have $\frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0}R_{1i,T}=o_P(1)$.

Fourth,

$$\begin{aligned}\frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0}R_{3i,T} &= \frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0}\left\{\frac{1}{T}\sum_{t=1}^{T}S_1\psi_i(1)\left[\breve{\varepsilon}_{it}e_{it}'\psi_i(1)'-\breve{\psi}_{i,0}\psi_i(1)'\right]s_i-\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}S_1\psi_i(1)\breve{\varepsilon}_{it}e_{is}'\psi_i(1)'\right\}\\ &\equiv \frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0}\left\{R_{3i,T1}-R_{3i,T2}\right\}.\end{aligned}$$

It is easy to show that $E\left\|N_k^{-1/2}\sum_{i\in G_k^0}R_{3i,T1}\right\|^2=O\left(p_2T^{-1}\right)$, implying that $N_k^{-1/2}\sum_{i\in G_k^0}R_{3i,T1}=o_P(1)$. As in the analysis of $R_{1i,T2}$, we can show that $N_k^{-1/2}\sum_{i\in G_k^0}R_{3i,T2}=O_P(\sqrt{N_kp_2}/T)=o_P(1)$. Thus $N_k^{-1/2}\sum_{i\in G_k^0}R_{3i,T}=o_P(1)$.

Fifth,

$$\begin{aligned}\frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0}R_{4i,T} &= \frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0}\left\{\frac{1}{T}\sum_{t=1}^{T}S_1\breve{\varepsilon}_{i0}e_{it}'\psi_i(1)'s_i-\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}S_1\breve{\varepsilon}_{it}e_{is}'\psi_i(1)'s_i\right\}\\ &\equiv \frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0}\left\{R_{4i,T1}-R_{4i,T2}\right\}.\end{aligned}$$

Noting that $R_{4i,T2}=R_{3i,T2}$, $N_k^{-1/2}\sum_{i\in G_k^0}R_{4i,T2}=o_P(1)$. For $R_{4i,T1}$, in view of the fact $\breve{\varepsilon}_{i0}=\sum_{s=0}^{\infty}\tilde{\psi}_{i,s}e_{i,-s}$ and $\{e_{it},t\ge1\}$ are mutually independent, we can readily show that $E(R_{4i,T1})=0$ and

$$\begin{aligned}E\left\|\frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0}R_{4i,T1}\right\|^2 &= \frac{1}{N_k}\sum_{i\in G_k^0}E\left\|\frac{1}{T}\sum_{t=1}^{T}S_1\breve{\varepsilon}_{i0}e_{it}'\psi_i(1)'s_i\right\|^2\\ &= \frac{1}{T^2N_k}\sum_{i\in G_k^0}\sum_{t=1}^{T}\sum_{s=1}^{T}E\left(s_i'\psi_i(1)e_{is}\breve{\varepsilon}_{i0}'S_1'S_1\breve{\varepsilon}_{i0}e_{it}'\psi_i(1)'s_i\right)\\ &= \frac{1}{T^2N_k}\sum_{i\in G_k^0}\sum_{t=1}^{T}\mathrm{tr}\left[E\left(\breve{\varepsilon}_{i0}'S_1'S_1\breve{\varepsilon}_{i0}\right)E\left(e_{it}'\psi_i(1)'s_is_i'\psi_i(1)e_{it}\right)\right]\\ &= \frac{1}{TN_k}\sum_{i\in G_k^0}E\left(\breve{\varepsilon}_{i0}'S_1'S_1\breve{\varepsilon}_{i0}\right)s_i'\psi_i(1)\psi_i(1)'s_i=O\left(p_2^2T^{-1}\right),\end{aligned}$$

where we use the fact that

$$
\begin{aligned}
s_i'\psi_i(1)\psi_i(1)'s_i &= s_i'\Omega_i s_i \le \lambda_{\max}(\Omega_i)s_i's_i \le 2\lambda_{\max}(\Omega_i)\left(S_0S_0' + \Sigma_{20,i}'\Sigma_{22,i}^{-1}S_2S_2'\Sigma_{22,i}^{-1}\Sigma_{20,i}\right) \\
&\le 2\lambda_{\max}(\Omega_i)\left[1+\lambda_{\max}(S_2S_2')\left[\lambda_{\min}(\Sigma_{22,i})\right]^{-2}\Sigma_{20,i}'\Sigma_{20,i}\right] \le Cp_2
\end{aligned}
$$

and $E\left(\breve{\varepsilon}_{i0}'S_1'S_1\breve{\varepsilon}_{i0}\right) \le E\left(\breve{\varepsilon}_{i0}'\breve{\varepsilon}_{i0}\right) \le Cp_2$. Then $N_k^{-1/2}\sum_{i\in G_k^0} R_{4i,T1} = O_P\left(p_2T^{-1/2}\right)$ and $N_k^{-1/2}\sum_{i\in G_k^0} R_{4i,T}$ $= o_P(1)$.

Sixth, we can show that

$$
\begin{aligned}
N_k^{-1/2}\sum_{i\in G_k^0} R_{5i,T} &= N_k^{-1/2}\sum_{i\in G_k^0}\frac{1}{T}S_1\left\{S_{i,T}^{\varepsilon}\breve{\varepsilon}_{iT}' - \frac{1}{T}\sum_{t=1}^{T}S_{i,t}^{\varepsilon}\breve{\varepsilon}_{iT}' - \frac{1}{T}S_{i,T}^{\varepsilon}\sum_{t=1}^{T}\breve{\varepsilon}_{it}' + \frac{1}{T^2}\sum_{t=1}^{T}S_{i,t}^{\varepsilon}\sum_{t=1}^{T}\breve{\varepsilon}_{it}'\right\}s_i \\
&= N_k^{-1/2}\sum_{i\in G_k^0}\frac{1}{T}S_1S_{i,T}^{\varepsilon}\breve{\varepsilon}_{iT}'s_i + o_P(1) = N_k^{-1/2}\sum_{i\in G_k^0}\bar{R}_{5i,T} + o_P(1), \text{ say.}
\end{aligned}
$$

Note that

$$
\begin{aligned}
\left\|E\left(N_k^{-1/2}\sum_{i\in G_k^0}\bar{R}_{5i,T}\right)\right\| &= \left\|N_k^{-1/2}\sum_{i\in G_k^0}\frac{1}{T}S_1\psi_i(1)\sum_{t=1}^{T}\sum_{r=1}^{\infty}E\left(e_{it}e_{i,T-r}'\right)\breve{\psi}_{i,r}'s_i\right\| \\
&= \left\|N_k^{-1/2}\sum_{i\in G_k^0}\frac{1}{T}S_1\psi_i(1)\sum_{t=1}^{T}\breve{\psi}_{i,T-t}'s_i\right\| \\
&\le N_k^{-1/2}\sum_{i\in G_k^0}\frac{1}{T}\left\|S_1\psi_i(1)\right\|_{\text{sp}}\sum_{r=1}^{\infty}\left\|\breve{\psi}_{i,r}'s_i\right\|_{\text{sp}} = O\left(\sqrt{N_k}p_2/T\right) = o(1).
\end{aligned}
$$

Similarly, we can verify that $\left\|\text{Var}\left(N_k^{-1/2}\sum_{i\in G_k^0}\bar{R}_{5i,T}\right)\right\| \le N_k^{-1}\sum_{i\in G_k^0}\left\|\text{Var}\left(\bar{R}_{5i,T}\right)\right\| = O\left(p_2^2/T\right) = o(1)$. It follows that $N_k^{-1/2}\sum_{i\in G_k^0} R_{5i,T} = o_P(1)$.

Last, it is trivial to show $\left\|N_k^{-1/2}\sum_{i\in G_k^0} R_{2i,T}\right\| = O\left(\sqrt{N_k}p_2/T\right) = o(1)$ and $\left\|N_k^{-1/2}\sum_{i\in G_k^0} R_{6i,T}\right\| =$ $O_P\left(\sqrt{N_k}p_2/T\right) = o_P(1)$.

In sum, we have shown $V_{1k,NT} - \mathbb{B}_{1k,NT} \Rightarrow N\left(0,\mathbb{V}_{(k)}\right)$. This completes the proof of (ii).

(iii) For $V_{2k,NT}$, we have

$$
\begin{aligned}
V_{2k,NT} &= \frac{1}{\sqrt{N_k}T}\sum_{i\in G_k^0}\tilde{x}_{1,i}'\tilde{x}_{2,i}\Sigma_{22,i}^{-1}\left(\Sigma_{20,i} - \frac{1}{T}\tilde{x}_{2,i}'\tilde{u}_i\right) \\
&= \frac{1}{\sqrt{N_k}T}\sum_{i\in G_k^0}x_{1,i}'x_{2,i}\Sigma_{22,i}^{-1}\left(\Sigma_{20,i} - \frac{1}{T}x_{2,i}'u_i\right) + \frac{1}{\sqrt{N_k}T}\sum_{i\in G_k^0}x_{1,i}'x_{2,i}\Sigma_{22,i}^{-1}\bar{x}_{2,i}\bar{u}_i \\
&\quad + \frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0}\bar{x}_{1,i}\bar{x}_{2,i}'\Sigma_{22,i}^{-1}\left(\frac{1}{T}x_{2,i}'u_i - \Sigma_{20,i}\right) - \frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0}\bar{x}_{1,i}\bar{x}_{2,i}'\Sigma_{22,i}^{-1}\bar{x}_{2,i}\bar{u}_i \\
&\equiv V_{2k,NTa} + V_{2k,NTb} + V_{2k,NTc} + V_{2k,NTd}, \text{ say.}
\end{aligned}
$$

Noting that $\frac{1}{N_kT}\sum_{i\in G_k^0}\|\bar{x}_{1,i}\|^2 = O_P(1)$, $\frac{1}{N_k}\sum_{i\in G_k^0}\|\bar{x}_{2,i}\|^2 = O_P\left(p_2T^{-1}\right)$, and $\frac{1}{N_kT^2}\sum_{i\in G_k^0}\left\|x_{1,i}'x_{2,i}\right\|^2 =$ $O_P(p_2)$ by direct moment calculation and Markov inequality and $\max_{1\le i\le N}\|\bar{x}_{2,i}\| = o_P(p_2^{1/2}a_{2NT})$ and

$\max_{1\le i\le N}|\bar{u}_i| = o_P(a_{2NT})$ by a simple application of Lemma S.1.2 in Su, Shi and Phillips (2016b, SSPb hereafter), we have

$$\begin{aligned}
\|V_{2k,NTb}\| &\le \sqrt{N_k}\left\{\frac{1}{N_kT^2}\sum_{i\in G_k^0}\left\|x_{1,i}'x_{2,i}\right\|^2\right\}^{1/2}\max_{i\in G_k^0}\|\bar{x}_{2,i}\|\max_{i\in G_k^0}\|\bar{u}_i\|\left\|\Sigma_{22,i}^{-1}\right\|_{\text{sp}} \\
&= \sqrt{N_k}O_P(p_2^{1/2})o_P(p_2^{1/2}a_{2NT})o_P(a_{2NT}) = o_P(1), \\
\|V_{2k,NTc}\| &\le \sqrt{N_k}\left\{\frac{1}{N_kT}\sum_{i\in G_k^0}\|\bar{x}_{1,i}\|^2\right\}^{1/2}\left\{\frac{1}{N_k}\sum_{i\in G_k^0}\|\bar{x}_{2,i}\|^2\right\}^{1/2}\max_{i\in G_k^0}\left\|\frac{1}{T}x_{2,i}'u_i-\Sigma_{20,i}\right\|\left\|\Sigma_{22,i}^{-1}\right\|_{\text{sp}} \\
&= \sqrt{N_k}O_P(1)O_P(p_2^{1/2}T^{-1/2})o_P(p_2^{1/2}a_{2NT}) = o_P(1),
\end{aligned}$$

and

$$\begin{aligned}
\|V_{2k,NTd}\| &\le \sqrt{N_kT}\left\{\frac{1}{N_kT}\sum_{i\in G_k^0}\|\bar{x}_{1,i}\|^2\right\}^{1/2}\max_{i\in G_k^0}\|\bar{x}_{2,i}\|^2\max_{i\in G_k^0}\|\bar{u}_i\|\left\|\Sigma_{22,i}^{-1}\right\|_{\text{sp}} \\
&= \sqrt{N_kT}o_P(p_2a_{2NT}^2)o_P(a_{2NT}) = o_P(1).
\end{aligned}$$

For $V_{2k,NTa}$, it is easy to see that

$$\begin{aligned}
\|V_{2k,NTa}\| &\le \sqrt{N_k}\left\{\frac{1}{N_kT^2}\sum_{i\in G_k^0}\left\|x_{1,i}'x_{2,i}\right\|^2\right\}^{1/2}\left\{\frac{1}{N_k}\sum_{i\in G_k^0}\left\|\Sigma_{20,i}-\frac{1}{T}x_{2,i}'u_i\right\|^2\right\}^{1/2}\max_{1\le i\le N}\left\|\Sigma_{22,i}^{-1}\right\|_{\text{sp}} \\
&= \sqrt{N_k}O_P\left(p_2^{1/2}\right)O_P\left(T^{-1/2}p_2^{1/2}\right)O(1) = O_P\left(p_2\sqrt{N_k/T}\right)
\end{aligned}$$

which is $o_P(1)$ if we assume that $p_2^2N_k/T = o(1)$. But this is a very strong assumption that we try to avoid. To do this, we can employ the BN decomposition and write $x_{1,it} = S_1\left[\psi_i(1)\sum_{s=1}^t e_{is} + \breve{e}_{i,0} - \breve{e}_{i,t}\right]$ and $x_{2,it} = S_2\left[\psi_i(1)e_{it} + \breve{e}_{i,t-1} - \breve{e}_{i,t}\right]$. Let $B_{iT} = \Sigma_{22,i}^{-1}\left(\Sigma_{20,i} - \frac{1}{T}x_{2,i}'u_i\right)$. As in the analysis of $V_{1k,NT}$, we can show that

$$\begin{aligned}
V_{2k,NTa} &= \frac{1}{\sqrt{N_k}T}\sum_{i\in G_k^0}x_{1,i}'x_{2,i}B_{iT} \\
&= \frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0}\frac{1}{T}\sum_{t=1}^T S_1\psi_i(1)E_{it}e_{it}'\psi_i(1)'S_2'B_{iT} + o_P(1) \\
&= \frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0}\left\{S_1\psi_i(1)\psi_i(1)'S_2'B_{iT} + \frac{1}{T}\sum_{t=1}^T S_1\psi_i(1)(e_{it}e_{it}'-I_p)\psi_i(1)'S_2'B_{iT}\right. \\
&\quad \left.\frac{1}{T}\sum_{t=1}^T S_1\psi_i(1)E_{i,t-1}e_{it}'\psi_i(1)'S_2'B_{iT}\right\} + o_P(1) \\
&\equiv \frac{1}{\sqrt{N_k}}\sum_{i\in G_k^0}\{R_{2iT,1}+R_{2iT,2}+R_{2iT,3}\} + o_P(1).
\end{aligned}$$

Noting that $E(R_{2iT,1}) = 0$ and

$$\begin{aligned}
\left\| \mathrm{Var}\left(\frac{1}{\sqrt{N_k}} \sum_{i \in G_k^0} R_{2iT,1} \right) \right\| &\leq \frac{1}{N_k} \sum_{i \in G_k^0} \left\| \mathrm{Var}\left(R_{2iT,1} \right) \right\| = \frac{1}{N_k} \sum_{i \in G_k^0} \left\| S_1 \Omega_i S_2' E\left(B_{iT} B_{iT}' \right) S_2 \Omega_i S_1' \right\| \\
&\leq \frac{1}{N_k} \sum_{i \in G_k^0} \left\| E\left(B_{iT} B_{iT}' \right) \right\|_{\mathrm{sp}} \left\| S_1 \Omega_i S_2' S_2 \Omega_i S_1' \right\| \\
&\leq \frac{C}{N_k} \sum_{i \in G_k^0} \left\| E\left(B_{iT} B_{iT}' \right) \right\| = O\left(p_2 T^{-1} \right) = o(1),
\end{aligned}$$

we have $\frac{1}{\sqrt{N_k}} \sum_{i \in G_k^0} R_{2iT,1} = o_P(1)$. Next,

$$\begin{aligned}
\left\| \frac{1}{\sqrt{N_k}} \sum_{i \in G_k^0} R_{2iT,2} \right\| &\leq \sqrt{N_k} \left\{ \frac{1}{N_k} \sum_{i \in G_k^0} \left\| \frac{1}{T} \sum_{t=1}^{T} S_1 \psi_i(1) \left(e_{it} e_{it}' - I_p \right) \psi_i(1)' S_2' \right\|^2 \right\}^{1/2} \left\{ \frac{1}{N_k} \sum_{i \in G_k^0} \left\| B_{iT} \right\|^2 \right\}^{1/2} \\
&= \sqrt{N_k} O_P(p_2^{1/2} T^{-1/2}) O_P(p_2^{1/2} T^{-1/2}) = o_P(1).
\end{aligned}$$

Let $z_{it} = S_1 \psi_i(1) E_{i,t-1} e_{it}' \psi_i(1)' S_2'$. Then $\frac{1}{\sqrt{N_k}} \sum_{i \in G_k^0} R_{2iT,3} = \frac{1}{\sqrt{N_k}} \sum_{i \in G_k^0} \frac{1}{T} \sum_{t=1}^{T} z_{it} B_{iT}$. Noting that $E(z_{it} | \mathcal{F}_{i,t-1}) = 0$, by the Burkholder's and Hölder's inequalities, for any $r \geq 2$,

$$E \left\| \sum_{t=1}^{T} z_{it} \right\|^r \leq C E \left\{ \sum_{t=1}^{T} \left\| z_{it} \right\|^2 \right\}^{r/2} \leq C_1 \left\{ \sum_{t=1}^{T} E\left(\left\| z_{it} \right\|^r \right) \right\}^{r/2} \leq C_2 p_2^{r/2} \sum_{t=1}^{T} t^{r/2} \leq C_2 p_2^{r/2} T^{r/2+1}.$$

where C_1 and C_2 are constants that depend on r. Then by the Hölder's inequality

$$\begin{aligned}
E \left\| \frac{1}{\sqrt{N_k}} \sum_{i \in G_k^0} R_{2iT,3} \right\| &\leq \frac{1}{\sqrt{N_k}} \sum_{i \in G_k^0} \frac{1}{T} E \left\| \sum_{t=1}^{T} z_{it} B_{iT} \right\| \\
&\leq \frac{1}{\sqrt{N_k}} \sum_{i \in G_k^0} \frac{1}{T} \left\{ E \left\| \sum_{t=1}^{T} z_{it} \right\|^r \right\}^{1/r} \left\{ E \left\| B_{iT} \right\|^q \right\}^{1/q} \\
&\leq \frac{C}{\sqrt{N_k}} \sum_{i \in G_k^0} \frac{1}{T} \left\{ p_2^{r/2} T^{r/2+1} \right\}^{1/r} p_2^{1/2} T^{-1/2} \\
&\leq C \sqrt{N_k} p_2 T^{-1+1/r} = o(1).
\end{aligned}$$

where $\frac{1}{q} + \frac{1}{r} = 1$.

Consequently, we have shown that $V_{2k,NT} = o_P(1)$.

(iv) Following the analysis of $V_{4k,NT}$ below, we can readily show that

$$\begin{aligned} V_{3k,NT} &= \frac{1}{\sqrt{N_k}T}\sum_{i\in G_k^0}\tilde{x}_{1,i}'\tilde{x}_{2,i}\Sigma_{22,i}^{-1}\left(\frac{1}{T}\tilde{x}_{2,i}'\tilde{x}_{2,i}-\Sigma_{22,i}\right)\left(\frac{1}{T}\tilde{x}_{2,i}'\tilde{x}_{2,i}\right)^{-1}\Sigma_{20,i} \\ &= \frac{1}{\sqrt{N_k}T}\sum_{i\in G_k^0}x_{1,i}'x_{2,i}\Sigma_{22,i}^{-1}\left(\frac{1}{T}x_{2,i}'x_{2,i}-\Sigma_{22,i}\right)\Sigma_{22,i}^{-1}\Sigma_{20,i}+o_P(1) \\ &\equiv V_{3k,NTa}+o_P(1). \end{aligned}$$

Following the analysis of $V_{2k,NT}$, we can show that $\|V_{3k,NTa}\| = o_P(1)$ by resorting to the BN decomposition, moment calculations, and Chebyshev inequality.

(v) For $V_{4k,NT}$, by the Cauchy-Schwarz inequality and Lemmas A.2(iii)-(iv) and A.6(iv),

$$\begin{aligned} \|V_{4k,NT}\| &\le \sqrt{N_k}\left\{\frac{1}{N_kT^2}\sum_{i\in G_k^0}\left\|\tilde{x}_{1,i}'\tilde{x}_{2,i}\right\|^2\right\}^{1/2}\max_{i\in G_k^0}\left\|\left(\frac{1}{T}\tilde{x}_{2,i}'\tilde{x}_{2,i}\right)^{-1}-\Sigma_{22,i}^{-1}\right\|_{\mathrm{sp}}\max_{i\in G_k^0}\left\|\frac{1}{T}\tilde{x}_{2,i}'\tilde{u}_i-\Sigma_{20,i}\right\| \\ &= \sqrt{N_k}O_P(p_2^{1/2})O_P(p_2a_{2NT})O_P(p_2^{1/2}a_{2NT}) = o_P(1). \end{aligned}$$

(vi) This follows from (i)-(v). ■

Proof of Lemma A.8. Let $\mathcal{V}_{k,NT}^a = \frac{1}{\sqrt{N_k}T}\sum_{i\in G_k^0}\tilde{x}_{1,i}'M_{2,i}\tilde{v}_i^a$. We make the following decomposition

$$\mathcal{V}_{k,NT}^a = \frac{1}{\sqrt{N_k}T}\sum_{i\in G_k^0}\tilde{x}_{1,i}'\tilde{v}_i^a - \frac{1}{\sqrt{N_k}T}\sum_{i\in G_k^0}\tilde{x}_{1,i}'\tilde{x}_{2,i}\left(\tilde{x}_{2,i}'\tilde{x}_{2,i}\right)^{-1}\tilde{x}_{2,i}'\tilde{v}_i^a \equiv \mathcal{V}_{1k,NT}^a - \mathcal{V}_{2k,NT}^a.$$

Noting that $v_{it}^a = \sum_{|j|\ge\bar{p}_2}\gamma_{i,j}'\varepsilon_{1,t+j}$, we have

$$\max_{i,t}E\left[(v_{it}^a)^2\right] = \max_{i,t}\sum_{|j|\ge\bar{p}_2}\sum_{|l|\ge\bar{p}_2}\gamma_{i,j}'E\left(\varepsilon_{1,t+j}\varepsilon_{1,t+j}'\right)\gamma_{i,l} \le C\max_i\left(\sum_{|j|\ge\bar{p}_2}\|\gamma_{i,j}\|\right)^2 \le CT^{-2a},$$

and

$$\max_i E\left[(\bar{v}_i^a)^2\right] = \max_{i,t}E\left(\frac{1}{T_{\bar{p}_2}}\sum_{t=\bar{p}_2+1}^{T-\bar{p}_2}v_{it}^a\right)^2 \le \frac{1}{T_{\bar{p}_2}}\sum_{t=\bar{p}_2+1}^{T-\bar{p}_2}\max_{i,t}E\left(v_{it}^a\right)^2 \le CT^{-2a}.$$

Then $\max_{i,t}E\left[(\tilde{v}_{it}^a)^2\right] \le 2\max_{i,t}E\left[(v_{it}^a)^2\right] + 2\max_i E\left[(\bar{v}_i^a)^2\right] \le 4CT^{-2a}$. Analogously, we can show that

$$\max_{i,t}E\left[(\tilde{v}_{it}^a)^4\right] = C\max_i\left(\sum_{|j|\ge\bar{p}_2}\|\gamma_{i,j}\|\right)^4 \le CT^{-4a}.$$

It follows that

$$
\begin{aligned}
E\left\|V_{1k,NT}^a\right\| &= E\left\|\frac{1}{\sqrt{N_k}T_{\bar{p}_2}}\sum_{i\in G_k^0}\sum_{t=\bar{p}_2+1}^{T-\bar{p}_2}x_{1,it}\tilde{v}_{it}^a\right\| \le \frac{1}{\sqrt{N_k}T_{\bar{p}_2}}\sum_{i\in G_k^0}\sum_{t=\bar{p}_2+1}^{T-\bar{p}_2}E\left\|x_{1,it}\tilde{v}_{it}^a\right\| \\
&\le \frac{1}{\sqrt{N_k}T_{\bar{p}_2}}\sum_{i\in G_k^0}\sum_{t=\bar{p}_2+1}^{T-\bar{p}_2}\left\{E\left\|x_{1,it}\right\|^2\right\}^{1/2}\left\{E\left\|\tilde{v}_{it}^a\right\|^2\right\}^{1/2} \\
&\le C\sqrt{N_k}T_{p_2}^{-a-1}\sum_{t=\bar{p}_2+1}^{T-\bar{p}_2}t^{1/2} = O\left(N_k^{1/2}T^{-a+\frac{1}{2}}\right) = o(1),
\end{aligned}
$$

where we use the fact that $\max_{1\le i\le N}E\left\|x_{1,it}\right\|^2 \le Ct$. Then $V_{1k,NT}^a = o_P(1)$ by the Markov inequality.

Next, noting that $\left\|\frac{1}{T}\tilde{x}_{2,i}'\tilde{x}_{2,i} - \Sigma_{22,i}\right\|_{\text{sp}} = o_P(1)$ by Lemma A.2(iv) and

$$
\begin{aligned}
\max_i E\left\|x_{2,i}'\tilde{v}_i^a\right\|^2 &\le \max_i\sum_{t=1}^T\sum_{s=1}^T E\left(x_{2,is}'x_{2,it}\tilde{v}_{it}^a\tilde{v}_{is}^{a2}\right) \\
&\le C\max_i\sum_{t=1}^T\sum_{s=1}^T\left\{E\left\|x_{2,it}\right\|^4\right\}^{1/2}\left\{E\left(\tilde{v}_{it}^a\right)^4\right\}^{1/2} \\
&\le C\sum_{t=1}^T\sum_{s=1}^T p_2T^{-2a} = Cp_2T^{-2a+2},
\end{aligned}
$$

we have

$$
\begin{aligned}
\left\|\mathcal{V}_{k,NT}^a\right\| &\le \frac{1+o_P(1)}{\sqrt{N_k}T^2}\sum_{i\in G_k^0}\left\|\tilde{x}_{1,i}'\tilde{x}_{2,i}\right\|\left\|\Sigma_{22,i}^{-1}\right\|_{\text{sp}}\left\|\tilde{x}_{2,i}'\tilde{v}_i^a\right\| \\
&\le \underline{c}_{22}^{-1}\frac{1+o_P(1)}{\sqrt{N_k}T^2}\sum_{i\in G_k^0}\left\|\tilde{x}_{1,i}'\tilde{x}_{2,i}\right\|\left\|x_{2,i}'\tilde{v}_i^a\right\| \\
&\le \underline{c}_{22}^{-1}\left(1+o_P(1)\right)\sqrt{N_k}\left\{\frac{1}{N_kT^2}\sum_{i\in G_k^0}\left\|\tilde{x}_{1,i}'\tilde{x}_{2,i}\right\|^2\right\}^{1/2}\left\{\frac{1}{N_kT^2}\sum_{i\in G_k^0}\left\|x_{2,i}'\tilde{v}_i^a\right\|^2\right\}^{1/2} \\
&= \sqrt{N_k}O_P\left(p_2^{1/2}\right)O_P\left(p_2^{1/2}T^{-a}\right) = O_P\left(p_2N_k^{1/2}T^{-a}\right) = o_P(1).
\end{aligned}
$$

In sum, we have shown that $\mathcal{V}_{k,NT}^a = o_P(1)$. ■

C Determination of the Number of Groups

In this section, we now propose a BIC-type information criterion to choose K, the number of groups. We now use K_0 to denote the true number of groups and K a generic number of groups. We assume that the true number of groups is bounded from above by a finite integer $K_{\max}$ and $1 \le K_0 \le K_{\max}$. By minimizing the objective function in (2.7), we obtain the C-Lasso estimators $\{\hat{\alpha}_k(K,\lambda), \hat{\beta}_{1,i}(K,\lambda), \hat{\beta}_{2,i}(K,\lambda)\}$ of $\{\alpha_k, \beta_{1,i}, \beta_{2,i},\}$ where we make the dependence of these estimators on (K,λ) explicit. We classify individual

i into group $\hat{G}_k(K,\lambda)$ if and only if $\hat{\beta}_{1,i}(K,\lambda) = \hat{\alpha}_k(K,\lambda)$, i.e.,

$$\hat{G}_k(K,\lambda) = \{i = \{1,2,...,N\} : \hat{\beta}_{1,i}(K,\lambda) = \hat{\alpha}_k(K,\lambda)\} \text{ for } k = 1,...K. \tag{C.1}$$

Let $\hat{G}(K,\lambda) = \{\hat{G}_1(K,\lambda), \cdots, \hat{G}_K(K,\lambda)\}$. We can define the post-Lasso estimators of α_k as

$$\hat{\alpha}^{\text{post}}_{\hat{G}_k(K,\lambda)} = \left(\sum_{i\in\hat{G}_k(K,\lambda)} \tilde{x}'_{1,i}M_{2,i}\tilde{x}_{1,i}\right)^+ \sum_{i\in\hat{G}_k(K,\lambda)} \tilde{x}'_{1,i}M_{2,i}\tilde{y}_i.$$

$\hat{\beta}^{\text{post}}_{2,i}(\hat{G}_k(K,\lambda))$ is defined as before but now we also make its dependence on $\hat{G}_k(K,\lambda)$ explicit. Let $\hat{\sigma}^2_{\hat{G}_k(K,\lambda)} = \frac{1}{NT}\sum_{k=1}^K \sum_{i\in\hat{G}_k(K,\lambda)} \sum_{t=1}^T [\hat{u}_{it}(k)]^2$, where $\hat{u}_{it}(k) = \tilde{y}_{it} - \tilde{x}'_{1,it}\hat{\alpha}^{\text{post}}_{\hat{G}_k(K,\lambda)} - \tilde{x}'_{2,it}\hat{\beta}^{\text{post}}_{2,i}(\hat{G}_k(K,\lambda))$ for $i \in \hat{G}_k(K,\lambda)$. We choose $\hat{K} = \hat{K}(\lambda)$ to minimize the following information criterion:

$$IC(K,\lambda) = \ln[\hat{\sigma}^2_{\hat{G}_k(K,\lambda)}] + \rho_{NT}p_1K, \tag{C.2}$$

where ρ_{NT} is a tuning parameter.

Let $G^{(K)} = (G_{K,1},...,G_{K,K})$ be any K-partition of the set of $\{1,2,...,N\}$ and $\mathcal{G}_K$ is a collection of such partitions. Let $\hat{\sigma}^2_{G^{(K)}} = \frac{1}{NT}\sum_{k=1}^K \sum_{i\in G_{K,k}} \sum_{t=1}^T [\tilde{y}_{it} - \tilde{x}'_{1,it}\hat{\alpha}_{G_{K,k}} - \tilde{x}'_{2,it}\hat{\beta}_{2,i}(G_{K,k})]^2$, where $\hat{\alpha}_{G_{K,k}} = (\sum_{i\in G_{K,k}} \tilde{x}'_{1,i}M_{2,i}\tilde{x}_{1,i})^+ \sum_{i\in G_{K,k}} \tilde{x}'_{1,i}M_{2,i}\tilde{y}_i$ and $\hat{\beta}_{2,i}(G_{K,k}) = \left(\tilde{x}'_{2,i}\tilde{x}_{2,i}\right)^{-1} \tilde{x}'_{2,i}\left(\tilde{y}_i - \tilde{x}_{1,i}\hat{\alpha}_{G_{K,k}}\right)$ for any $i \in G_{K,k}$. Define

$$\nu_{NT} = \begin{cases} N^{1/2}T^{1/2} & \text{in Case 1 where } x_{2,it} \text{ is absent in (2.1) and there is no endogeneity in } x_{1,it}, \\ T^{1/2} & \text{in Case 2 where } x_{2,it} \text{ is absent in (2.1) and there is endogeneity in } x_{1,it}, \\ p_2^{-1/2}T^{1/2} & \text{in Case 3 where } x_{2,it} \text{ is present in (2.1).} \end{cases} \tag{C.3}$$

Let $\sigma^2_{0,NT} = \frac{1}{NT}\sum_{i=1}^N \sum_{i=1}^T \tilde{u}^2_{it}$ in Cases 1-2 and $= \frac{1}{NT}\sum_{i=1}^N \sum_{i=1}^T (\tilde{u}^*_{it})^2$ in the Case 3. We can show that $\hat{\sigma}^2_{\hat{G}_k(K_0,\lambda)} - \sigma^2_{0,NT} = O_P(N^{-1}T^{-1})$, $O_P(T^{-1})$, and $O_P(p_2T^{-1})$ corresponding to the above three cases, respectively.

We add the following assumption.

Assumption A.5 (i) As $(N,T) \to \infty$, $\min_{1\le K<K_0} \inf_{G^{(K)}\in\mathcal{G}_K} \hat{\sigma}^2_{G^{(K)}} \to \underline{\sigma}^2 > \sigma^2_0$, where $\sigma^2_0 = \text{plim}_{(N,T)\to\infty}$ $\sigma^2_{0,NT}$.

(ii) As $(N,T) \to \infty$, $\rho_{NT} \to 0$ and $\rho_{NT}\nu^2_{NT} \to \infty$ where ν_{NT} is as defined in (C.3).

Assumption A.5(i) guarantees that all under-grouped models yield asymptotic mean square errors that are larger than σ^2_0, which can be obtained from the true model. Assumption A.5(ii) imposes the usual type of conditions for the consistency of model selection: the penalty coefficient ρ_{NT} cannot shrink to zero either too fast or too slowly.

The following theorem suggests that in large samples we can determine the correct number of groups by minimizing the information criterion defined in (C.2).

Theorem C.1 *Suppose that Assumptions A.1, A.3 and A.5 hold. Suppose that there exists a constant* $\underline{c}_{00}$ *such that* $\min_{1\le i\le N} \Sigma_{00,i} \ge \underline{c}_{00} > 0$. *Then* $P(\hat{K} = K_0) \to 1$ *as* $(N,T) \to \infty$.

Theorem C.1 indicates that w.p.a.1 the use of IC(K,λ) in (C.2) determines the correct number of groups. A natural question is how to choose the tuning parameter ρ_{NT} empirically.

In simulations and applications, we recommend the use of DPLS estimation so that Case 3 applies. We will choose $p_2 = [T^{1/4}]$ and set $\rho_{NT} = \frac{1}{3}(NT)^{-1/3}$. Note that this rate converges to zero much slower than the usual $(NT)^{-1}\ln(NT)$-rate that works in Case 1. One can verify that the conditions in Lemma A.5(ii) are satisfied in this case when N and T diverge to infinity at roughly the same rate. Our simulations suggest that the choice of ρ_{NT} has little effect on the results.

Proof of Theorem C.1. Let $\mathcal{K} = \{1, 2, ..., K_{\max}\}$ where $K_{\max} \geq K_0$. We divide K into three subsets K_0, K_- and $K_+ : \mathcal{K}_0 = \{K_0\}$, $\mathcal{K}_- = \{K \in \mathcal{K} : K < K_0\}$, and $\mathcal{K}_+ = \{K \in \mathcal{K} : K > K_0\}$. First, using arguments as used in the proof of Lemma A.5(iii) we can show that

$$\hat{\sigma}^2_{\hat{G}(K_0,\lambda)} = \frac{1}{NT}\sum_{k=1}^{K_0}\sum_{i\in\hat{G}_k(K_0,\lambda)}\sum_{t=1}^{T}\left[\tilde{y}_{it} - \tilde{x}'_{1,it}\hat{\alpha}_k^{D,\text{post}}(K_0,\lambda) - \tilde{x}'_{2,it}\hat{\beta}_{2,i}^{D,\text{ post}}\left(\hat{G}_k(K_0,\lambda)\right)\right]^2 = \sigma_0^2 + o_P(1).$$

It follows that $IC(K_0,\lambda) = \ln[\hat{\sigma}^2_{\hat{G}(K_0,\lambda)}] + \rho_{NT}p_1K_0 = \ln[\hat{\sigma}^2_{\hat{G}(K_0,\lambda)}] + o(1) \xrightarrow{P} \ln(\sigma_0^2)$. We consider the cases of under- and over-fitted models separately.

Case 1: Under-fitted model ($K \in \mathcal{K}_-$). Noting that

$$\begin{aligned}
\hat{\sigma}^2_{\hat{G}(K,\lambda)} =& \frac{1}{NT}\sum_{k=1}^{K}\sum_{i\in\ \hat{G}_k(K,\lambda)}\sum_{t=1}^{T}\left[\tilde{y}_{it} - \tilde{x}'_{1,it}\hat{\alpha}_k^{D,\text{post}}(K,\lambda) - \tilde{x}'_{2,it}\hat{\beta}_{2,i}^{D,\text{ post}}\left(\hat{G}_k(K,\lambda)\right)\right]^2 \\
&\geq \min_{1\leq K<K_0}\inf_{G^{(K)}\in\mathcal{G}_K}\frac{1}{NT}\sum_{k=1}^{K}\sum_{i\in G_{K,k}}\sum_{t=1}^{T}\left[\tilde{y}_{it} - \tilde{x}'_{1,it}\hat{\alpha}^D_{\hat{G}_{K,k}} - \tilde{x}'_{2,it}\hat{\beta}^D_{2,i}(\hat{G}_{K,k})\right]^2 = \min_{1\leq K<K_0}\inf_{G^{(K)}\in\mathcal{G}_K}\hat{\sigma}^2_{G^{(K)}}.
\end{aligned}$$

By Assumption 4.2, we demonstrate

$$\min_{1\leq K<K_0} IC(K,\lambda) \geq \min_{1\leq K<K_0}\inf_{G^{(K)}\in G_K}\ln(\hat{\sigma}^2_{G^{(K)}}) + \rho_{NT}pK \xrightarrow{P} \ln(\underline{\sigma}^2) > \ln(\sigma_0^2).$$

It follows that $P(\min_{K\in\mathcal{K}_-} IC(K,\lambda) > IC(K_0,\lambda)) \to 1$.

Case 2: Over-fitted model ($K \in \mathcal{K}_+$).

$$\begin{aligned}
&P\left(\min_{K\in\mathcal{K}_+} IC(K,\lambda) > IC(K_0,\lambda)\right) \\
&= P\left(\min_{K\in\mathcal{K}_+}\nu_{NT}^2\ln(\hat{\sigma}^2_{\hat{G}(K,\lambda)}/\hat{\sigma}^2_{\hat{G}(K_0,\lambda)}) + \nu_{NT}^2\rho_{NT}(K-K_0) > 0\right) \\
&= P\left(\min_{K\in\mathcal{K}_+}\nu_{NT}^2(\hat{\sigma}^2_{\hat{G}(K,\lambda)} - \hat{\sigma}^2_{\hat{G}(K_0,\lambda)})/\hat{\sigma}^2_{\hat{G}(K_0,\lambda)} + \nu_{NT}^2\rho_{NT}(K-K_0) + o_P(1) > 0\right) \\
&\to 1 \text{ as } (N,T)\to\infty,
\end{aligned}$$

where $\min_{K\in K^+}\nu_{NT}^2(\hat{\sigma}^2_{\hat{G}(K,\lambda)} - \hat{\sigma}^2_{\hat{G}(K_0,\lambda)}) = O_p(1)$ by Lemma C.2 below and $\nu_{NT}^2\rho_{NT} \to \infty$ by Assumption A.5. ■

Lemma C.2 *Let* $\bar{\sigma}^2_{0,NT} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{u}_{it}^2$. *Let the conditions in Theorem C.1 hold. Then* $\max_{K_0<K\leq K_{\max}}$ $|\hat{\sigma}^2_{\hat{G}(K,\lambda)} - \bar{\sigma}^2_{0,NT}| = O_p(\nu_{NT}^{-2})$.

Proof of Lemma C.2. When $K > K_0$, following the proof of Theorem 4.1, we can show that $\|\hat{\beta}_{1,i} - \beta^0_{1,i}\| = O_P(T^{-1}+\lambda)$, $\|\hat{\beta}_{2,i} - \beta^*_{2,i}\| = O_P(p_2^{1/2}(T^{-1/2}+\lambda))$, and $\frac{1}{N}\sum_{i=1}^{N}\prod_{k=1}^{K}\|\beta^0_{1,i} - \hat{\alpha}_k\| = O_P(b_TT^{-1})$.

Noting that $\beta^0_{1,i},\ i=1,...,N$, only take K_0 distinct values, the latter implies that the collection $\mathcal{C} \equiv \{\hat{\alpha}_1,...,\hat{\alpha}_K\}$ contains at least K_0 distinct vectors, say, $\hat{\alpha}_1,...,\hat{\alpha}_{K_0}$, such that $\hat{\alpha}_k - \alpha^0_k = O_P(b_T T^{-1})$ for $k=1,...,K_0$. For notational simplicity, we rename the other vectors in the above collection as $\hat{\alpha}_{K_0+1},...,\hat{\alpha}_K$. By the pointwise convergence of $\hat{\beta}^D_{1,i}$, $\hat{\alpha}_{K_0+1},...,\hat{\alpha}_K$ must converge in probability to one of the true values in $\{\alpha^0_1,...,\alpha^0_{K_0}\}$.

We classify $i \in \hat{G}_k(K,\lambda)$ if $||\hat{\beta}_{1,i}-\hat{\alpha}_k||=0$ for $k=1,...,K$, and $i \in \hat{G}_0(K,\lambda)$ otherwise. Using arguments like those used in the proof of Theorem 4.3 and that of Lemma S1.14 in SSPb, we can show that

$$\sum_{i\in G^0_k} P(\hat{E}_{kNT,i}) = o(1) \text{ and } \sum_{i\in \hat{G}_k(K,\lambda)} P(\hat{F}_{kNT,i}) = o(1) \text{ for } k=1,...,K_0. \tag{C.4}$$

This implies that

$$\sum_{i=1}^{N} P(i \in \hat{G}_0(K,\lambda) \cup \hat{G}_{K_0+1}(K,\lambda) \cup ... \cup \hat{G}_K(K,\lambda)) = o(1). \tag{C.5}$$

That is, the 'redundant' last $K-K_0$ groups containing empty elements asymptotically. Using the fact that $\mathbf{1}\{i\in\hat{G}_k\} = \mathbf{1}\{i\in G^0_k\} + \mathbf{1}\{i\in\hat{G}_k\backslash G^0_k\} - \mathbf{1}\{i\in G^0_k\backslash\hat{G}_k\}$, we have

$$\hat{\sigma}^2_{\hat{G}(K,\lambda)} = \frac{1}{NT}\sum_{k=1}^{K}\sum_{i\in\hat{G}_k(K,\lambda)}\sum_{t=1}^{T}[\hat{u}_{it}(k)]^2 = D_{1NT} + D_{2NT} - D_{3NT} + D_{4NT},$$

where $D_{1NT} = \frac{1}{NT}\sum_{k=1}^{K_0}\sum_{i\in G^0_k}\sum_{t=1}^{T}[\hat{u}_{it}(k)]^2$, $D_{2NT} = \frac{1}{NT}\sum_{k=1}^{K_0}\sum_{i\in\hat{G}_k(K,\lambda)\backslash G^0_k}\sum_{t=1}^{T}[\hat{u}_{it}(k)]^2$, $D_{3NT} = \frac{1}{NT}\sum_{k=1}^{K_0}\sum_{i\in G^0_k\backslash\hat{G}_k(K,\lambda)}\sum_{t=1}^{T}[\hat{u}_{it}(k)]^2$, and $D_{4NT} = \frac{1}{NT}\sum_{k=K_0+1}^{K}\sum_{i\in\hat{G}_k}\sum_{t=1}^{T}[\hat{u}_{it}(k)]^2$. By (C.4)-(C.5), we can readily show that $D_{\ell NT} = o_P\left((NT)^{-1}\right)$ for $\ell=2,3,4$. For D_{1NT}, we discuss several cases: (1) When $x_{2,it}$ is absent in the cointegrating regression and there is no endogeneity in $x_{1,it}$, we can apply the fact $\hat{\alpha}^{\text{post}}_k$, $k=1,...,K_0$, converge to their true values at $N^{-1/2}T^{-1}$ and show that $D_{1NT} - \bar{\sigma}^2_{0NT} = O_P\left(N^{-1}T^{-1}\right)$; (2) When $x_{2,it}$ is absent in the cointegrating regression and there is endogeneity in $x_{1,it}$, we can apply the fact $\hat{\alpha}^{\text{post}}_k$, $k=1,...,K_0$, converge to their true values at rate T^{-1} and show that $D_{1NT} - \bar{\sigma}^2_{0NT} = O_P\left(T^{-1}\right)$; (3) When both $x_{1,it}$ and $x_{2,it}$ are present, we observe that $\hat{\beta}_{2,i}$ converge to their (pseudo) true values at rate $p_2^{1/2}T^{-1/2}$ and show that $D_{1NT} - \sigma^2_0 = O_P\left(pT^{-1}\right)$. As a result, we have $\hat{\sigma}^2_{\hat{G}(K,\lambda)} = \sigma^2_0 + O_P(\nu^{-2}_{NT})$ where $\nu_{NT} = N^{1/2}T^{1/2}$, $T^{1/2}$, and $p_2^{-1/2}T^{1/2}$ in the above three cases, respectively. This completes the proof of the lemma. ■

D Practical Implementation of the C-Lasso Procedure

In this section, we provide more details on the practical implementation of the C-Lasso procedure in the followings steps.

1. **Initial estimates based on the heterogenous nonstationary panels.** Obtain the initial estimates $\tilde{\beta}_{1,i}$ and $\tilde{\beta}_{2,i}$ from the LS time-series regression of $\tilde{y}_{it}$ on $\left(\tilde{x}'_{1,it},\tilde{x}'_{2,it}\right)$. Let $Q_{NT}(\boldsymbol{\beta_1},\boldsymbol{\beta}_2) = \frac{1}{NT^2}\sum_{i=1}^{N}\left\|\tilde{y}_i - \tilde{x}_{1,i}\beta_{1,i} - \tilde{x}_{2,i}\beta_{2,i}\right\|^2$, $\tilde{\sigma}^2_i = \frac{1}{T}\sum_{t=1}^{T}(\tilde{y}_{it} - \tilde{\beta}'_i\tilde{x}_{it})^2$, and $\hat{Q}_{1i} = \frac{1}{T^2}\sum_{t=1}^{T}\tilde{x}_{1,it}\tilde{x}'_{1,it}$.

2. **Determining the number (K) of groups along with the tuning parameter λ.** Let

$$\Lambda \equiv \left\{\lambda = c_j T^{-3/4},\ c_j = c_0\gamma^j \text{ for } j=0,...,J\right\} \text{ for some } c_0>0 \text{ and } \gamma>1.$$

Given any $K \in \{1, 2, ..., K_{\max}\}$ and $\lambda \in \Lambda$, compute $\mathrm{IC}(K, \lambda)$ and $\mathrm{IC}(\hat{K}(\lambda), \lambda)$ where $\hat{K}(\lambda) = \arg\min_{1 \le K \le K_{\max}} \mathrm{IC}(K, \lambda)$. Choose $\hat{\lambda} \in \Lambda$ such that $\mathrm{IC}(\hat{K}(\lambda), \lambda)$ is minimized. The estimated number of groups is then given by

$$\hat{K} = \min_{\lambda \in \Lambda} \hat{K}(\lambda).$$

Note that the above procedure fine-tunes the tuning parameter λ for the determination of the number of groups and is recommended by Su, Shi, and Phillips (2016a, SSPa hereafter). We find in simulations $c_0 = 0.025$, $\gamma = 2$, and $J = 3$ work fairly well for all DGPs under our investigation. If $\hat{K} = 1$, stop here and estimate a homogenous nonstationary panel as usual. Otherwise, move to the next step.

3. **C-Lasso estimation.** Given $\hat{\lambda}$ and $\hat{K} > 1$, solve the PLS problem

$$Q_{NT,\lambda}^{K}(\boldsymbol{\beta_1}, \boldsymbol{\beta}_2, \boldsymbol{\alpha}) = Q_{NT}(\boldsymbol{\beta_1}, \boldsymbol{\beta}_2) + \frac{\hat{\lambda}}{N} \sum_{i=1}^{N} (\tilde{\sigma}_i)^{2-\hat{K}} \prod_{k=1}^{\hat{K}} \left\| \hat{Q}_{1i}(\beta_{1,i} - \alpha_k) \right\|.$$

Obtain the C-Lasso estimates $\{\hat{\alpha}_k\}$ for the group-specific parameters and $\{\hat{G}_k, k = 1, ..., \hat{K}\}$ for the estimated group membership.

4. **Post-Lasso estimator with bias correction:** Given the estimated groups, $\{\hat{G}_k, k = 1, ..., \hat{K}\}$, we can obtain the post-Lasso estimators of α_k and $\beta_{2,i}$ as

$$\begin{aligned} \hat{\alpha}_k^{\text{post}} &= \left(\sum_{i \in \hat{G}_k} \tilde{x}_{1,i}' M_{2,i} \tilde{x}_{1,i} \right)^{-1} \sum_{i \in \hat{G}_k} \tilde{x}_{1,i}' M_{2,i} \tilde{y}_i \quad \text{for } k = 1, ..., \hat{K}, \\ \hat{\beta}_{2,i}^{\text{post}} &= \left(\tilde{x}_{2,i}' \tilde{x}_{2,i} \right)^{-1} \tilde{x}_{2,i}' (\tilde{y}_i - \tilde{x}_{1,i} \hat{\alpha}_k^{\text{post}}) \text{ for } i \in \hat{G}_k, \end{aligned}$$

where to remove the bias we apply the dynamic OLS method in the post-Lasso estimation by including the lags and leads of $\Delta x_{1,it}$ into $x_{2,it}$ as in Section 4.4. If $x_{2,it}$ only contains the lags and leads of $\Delta x_{1,it}$ but no other stationary regressors, we compute the standard errors for the elements of $\hat{\alpha}_k^{\text{post}}$ as the square roots of the diagonal elements of $\frac{1}{\hat{N}_k T^2} \hat{\mathbb{Q}}_{(k)}^{-1} \hat{\mathbb{V}}_{(k)}^{\dagger} \hat{\mathbb{Q}}_{(k)}^{-1}$ where

$$\hat{\mathbb{Q}}_{(k)} = \frac{1}{\hat{N}_k T^2} \sum_{i \in \hat{G}_k} \tilde{x}_{1,i}' M_{2,i} \tilde{x}_{1,i} \text{ and } \hat{\mathbb{V}}_{(k)}^{\dagger} \equiv \frac{1}{\hat{N}_k} \sum_{i \in \hat{G}_k} \frac{1}{6} \hat{\Omega}_{00,i}^{\dagger} \hat{\Omega}_{11,i} \text{ for } k = 1, ..., \hat{K},$$

and $\hat{\Omega}_{00,i}^{\dagger}$ and $\hat{\Omega}_{11,i}$ are as defined in Section 4.4. If $x_{2,it}$ also contains other stationary covariates, then we can compute the standard errors for the elements of $\hat{\alpha}_k^{\text{post}}$ as the square roots of the diagonal elements of $\frac{1}{\hat{N}_k T^2} \hat{\mathbb{Q}}_{(k)}^{-1} \hat{\mathbb{V}}_{(k)} \hat{\mathbb{Q}}_{(k)}^{-1}$ where

$$\hat{\mathbb{V}}_{(k)} = \frac{1}{\hat{N}_k} \sum_{i \in \hat{G}_k} \left[\frac{1}{6} \hat{s}_i' \hat{\Omega}_i \hat{s}_i S_1 \hat{\Omega}_i S_1' - \frac{1}{12} \left(\hat{s}_i' \hat{\Omega}_i S_1' \otimes S_1 \hat{\Omega}_i \hat{s}_i \right) K_{p_1,1} \right],$$

$\hat{s}_i = S_0' - S_2' \hat{\Sigma}_{22,i}^{-1} \hat{\Sigma}_{20,i}$, $\hat{\Omega}_i$ denotes the HAC estimator of the long-run variance-covariance in Ω_i, and $\hat{\Sigma}_{22,i}$ and $\hat{\Sigma}_{20,i}$ denote the plug-in estimators of the short-run variance covariance submatrices $\Sigma_{22,i}$ and $\Sigma_{20,i}$ of Σ_i.

Table A.1: Frequency for selecting K=1, 2,..., 6 groups

	N	T	1	2	3	4	5	6
DGP1	50	40	0	0	0.992	0.008	0	0
	50	80	0	0	1	0	0	0
	100	40	0	0	1	0	0	0
	100	80	0	0	1	0	0	0
DGP2	50	40	0	0	0.966	0.034	0	0
	50	80	0	0	0.998	0.002	0	0
	100	40	0	0	0.982	0.018	0	0
	100	80	0	0	1	0	0	0
DGP3	50	40	0	0	0.988	0.012	0	0
	50	80	0	0	1	0	0	0
	100	40	0	0	1	0	0	0
	100	80	0	0	1	0	0	0
DGP4	50	40	0	0.976	0.024	0	0	0
	50	80	0	1	0	0	0	0
	100	40	0	0.956	0.044	0	0	0
	100	80	0	1	0	0	0	0
DGP5	50	40	0	0	0.990	0.010	0	0
	50	80	0	0	1	0	0	0
	100	40	0	0	0.986	0.014	0	0
	100	80	0	0	1	0	0	0

E Additional Simulation Results

In this appendix, we assess the performance of the information criterion (IC) proposed in Section C. We set $\rho_{NT} = \frac{1}{3}(NT)^{-1/3}$ and $\lambda = c_\lambda T^{-3/4}$ where $c_\lambda = 0.025$, 0.05, 0.1, or 0.2. We find that the results are not sensitive to the choice of c_λ and will only report the simulation results for the case where $c_\lambda = 0.1$ to save space. Table A.1 displays the empirical probability with which a particular group number from 1 to 6 is selected according to IC based on 500 replications for each DGP. Note that the true number of groups is 3 for DGPs 1, 2, 3, and 5 and 2 for DGP 4. When $T = 40$, the probabilities of correct choices are higher than 95 % in all cases and they reach the unity when $T = 80$. The simulation results show that our information criterion works fairly well.

F Additional application results

In this section, we report some additional results for the empirical application

F.1 Information criterion for the quarterly data

Table A.2 reports the information criterion (IC) for the quarterly data with different tuning parameter values: $\lambda = c_\lambda \times T^{-3/4}$ where $c_\lambda = 0.025$, 0.05, 0.1, and 0.2. Following the majority rule, we decide to select $K = 2$ groups for the period 1975.Q1-1998.Q4 and $K = 3$ groups for the period 1999.Q1-2014.Q2. Note that the IC is minimized at $c_\lambda = 0.1$ and 0.05 for the first and second subsamples, respectively. For this reason, we choose $c_\lambda = 0.1$ and 0.05 for these two subsamples, in the paper.

Table A.2: The information criterion for different numbers of groups (quarterly data)

	From 1975.Q1-1998.Q4				From 1999.Q1-2014.Q2			
K/c_λ	0.025	0.05	0.10	0.20	0.025	0.05	0.10	0.20
1	-0.7503	-0.7503	-0.7503	-0.7503	-0.2074	-0.2074	-0.2074	-0.2074
2	-1.1262	**-1.1262**	**-1.1262**	**-1.0716**	-0.4719	-0.4730	-0.4902	**-0.4836**
3	**-1.1622**	-0.7961	-1.0956	-0.7135	**-0.5230**	**-0.5319**	**-0.5268**	-0.4418
4	-0.7719	-0.7507	-0.7507	-1.0596	-0.5037	-0.4994	-0.4958	-0.3815
5	-0.7233	-0.7203	-0.6750	-0.6750	-0.4789	-0.4749	-0.3499	-0.2093
6	-0.6946	-0.6405	-0.6005	-0.6844	-0.4454	-0.4358	-0.3566	-0.1720

Table A.3: The information criterion for different numbers of groups (monthly data)

	From 1975-1998				From 1999-2014			
$K\backslash c_\lambda$	0.025	0.05	0.10	0.20	0.025	0.05	0.10	0.20
1	-0.8042	-0.8042	-0.8042	-0.8042	-0.1953	-0.1953	-0.1953	-0.1953
2	-1.0907	**-1.0907**	**-1.0907**	**-1.0140**	-0.4686	-0.4753	-0.4837	**-0.4748**
3	**-1.1460**	-0.8480	-0.8404	-0.8365	**-0.5312**	**-0.5311**	**-0.5230**	-0.3940
4	-1.0966	-1.0897	-0.8292	-0.9159	-0.5161	-0.5132	-0.5086	-0.3139
5	-0.9044	-1.0646	-0.9047	-0.7949	-0.5032	-0.4987	-0.3630	-0.2711
6	-0.8782	-1.0379	-0.7875	-0.7678	-0.4768	-0.4753	-0.3016	-0.2466

F.2 Results for the monthly data

In this section we provide the application results for the monthly data.

Table A.3 reports the information criterion (IC) for the monthly data with different tuning parameter values: $\lambda = c_\lambda \times T^{-3/4}$ where $c_\lambda = 0.025$, 0.05, 0.1, and 0.2. As is evident from Table A.3, for the monthly data our information criterion tends to choose 2 groups for the first subsample and 3 groups for the second subsample, too. We set $c_\lambda = 0.05$ to report the estimation results in Table A.4 and classification results in Table A.5.

Comparing the estimation results in Table 4 for the quarterly data with those in Table A.4 for the monthly data, we find that the estimates for either group in either subsample period of the monthly data are reasonably close to the corresponding estimates based on the quarterly data. This suggests the robustness of our results. The countries in bold in Table A.5 suggest good coincidences of the classification results based on the monthly and quarterly datasets.

References

Corradi, V. (1999) Deciding between I(0) and I(1) via FLIL-based bounds. *Econometric Theory* 15(5), 643-663.

Eberlein, E. (1986) On strong invariance principles under dependence assumptions. *Annals of Probability* 14, 260-270.

Freedman, D.A. (1975) On tail probabilities for martingales. *Annals of Probability* 3, 100-118.

Hall, P. & C.C. Heyde (1980) *Martingale Limit Theory and Its Application.* Academic Press, New York.

Park, J.Y. & P.C.B. Phillips (1988) Statistical inference in regressions with integrated processes: Part 1. *Econometric Theory* 4(3), 468-497.

Park, J.Y. & P.C.B. Phillips (1989) Statistical inference in regressions with integrated processes: Part 2. *Econometric Theory* 5(1), 95-131.

Table A.4: Estimation results for the monthly data

	Panel A: From 1975.M1-1998.M12						
	Pool	Group 1			Group 2		
	DOLS	C-Lasso	post-Lasso		C-Lasso	post-Lasso	
β_i	0.7553	0.8350	0.8350		-0.7175	-0.7175	
	(0.0116)	(0.0097)	(0.0097)		(0.0554)	(0.0554)	
	Panel B: From 1999.M1-2014.M7						
	Pool	Group 1		Group 2		Group 3	
	DOLS	C-Lasso	post-Lasso	C-Lasso	post-Lasso	C-Lasso	post-Lasso
β_i	0.3886	0.8197	0.8197	-0.5097	-0.5096	0.1577	0.1577
	(0.0069)	(0.0086)	(0.0086)	(0.0154)	(0.0154)	(0.0179)	(0.0179)

Phillips, P.C.B. & H.R. Moon (1999) Linear regression limit theory for nonstationary panel data. *Econometrica* 67, 1057-1111.

Phillips, P.C.B. & V. Solo (1992). Asymptotics for linear processes. *Annals of Statistics* 20, 971-1001.

Pollard, D. (1984) *Convergence of stochastic processes.* Springer-Verlag, New York.

Strassen, V. (1964) An invariance principle for the law of the iterated logarithm. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 3, 211-226.

Su, L., Z. Shi, & P.C.B. Phillips (2016a) Identifying latent structures in panel data. *Econometrica* 84(6), 2215-2264.

Su, L., Z. Shi, & P.C.B. Phillips (2016b) Supplement to 'Identifying latent structures in panel data', *Econometrica Supplemental Material* 84, http://dx.doi.org/10.3982/ECTA12560.

Table A.5: Classification results for the monthly data

Panel A: From 1975.M1-1998.M12				
Group 1 ($N_1 = 53$)				
Algeria	**Austria**	**Bahrain**	**Belgium**	**Bolivia**
Botswana	**Canada**	**Colombia**	**Costa Rica**	**Cyprus**
Denmark	**Egypt**	**Finland**	**France**	**Ghana**
Greece	**Honduras**	**Hungary**	**India**	**Indonesia**
Israel	**Italy**	**Ivory Coast**	**Jamaica**	**Japan**
Jordan	**Kenya**	**South Korea**	**Luxembourg**	**Malta**
Mauritius	**Mexico**	**Morocco**	**Nepal**	**Netherlands**
Nigeria	**Norway**	**Pakistan**	**Paraguay**	**Philippines**
Portugal	**Singapore**	**South Africa**	**Spain**	**Sri Lanka**
Sudan	**Sweden**	**Switzerland**	**Thailand**	**Trinidad and Tobago**
Turkey	**Uruguay**	**Venezuela**		
Group 2 ($N_2 = 3$)				
Ecuador	**Kuwait**	**Myanmar**		
Panel B: From 1999.M1-2014.M7				
Group 1 ($N_1 = 53$)				
Angola	**Argentina**	**Austria**	**Bangladesh**	**Belgium**
Botswana	Cambodia	**Canada**	**Costa Rica**	**Denmark**
Dominican	Egypt	**Europe**	**Finland**	**France**
Germany	**Ghana**	**Honduras**	**Iceland**	India
Iran	**Italy**	**Jamaica**	**Japan**	**Jordan**
Luxembourg	**Malawi**	Mauritius	**Mexico**	**Mongolia**
Morocco	**Mozambique**	Nepal	**Netherlands**	**Nigeria**
Norway	**Pakistan**	**Romania**	**Saudi Arabia**	**Sri Lanka**
Sudan	**Sweden**	**Switzerland**	**Tanzania**	**Trinidad and Tobago**
Tunisia	**Turkey**	**Uganda**	**United Kingdom**	**Ukraine**
Uruguay	**Venezuela**	**Viet Nam**		
Group 2 ($N_2 = 20$)				
Albania	**Armenia**	**Brazil**	**Bulgaria**	**Colombia**
Congo	**Croatia**	**Georgia**	**Hungary**	**Ireland**
Ivory Coast	**Kuwait**	**Latvia**	Lithuania	**Macau**
Moldova	**Peru**	**Philippines**	**Spain**	**Thailand**
Group 3 ($N_3 = 21$)				
Algeria	Bolivia	**Czech Republic**	**Guatemala**	**Hong Kong**
Indonesia	**Israel**	**Kazakhstan**	**Kenya**	**South Korea**
Kyrgyzstan	**Laos**	**Macedonia**	**Malaysia**	**Myanmar**
Paraguay	**Poland**	**Portugal**	**Russia**	**Singapore**
South Africa				

Note: Countries in bold denote coincidences of the classification results based on the monthly and quarterly datasets.