Supplement to "Determining the number of groups in latent panel structures with an application to income and democracy"

(*Quantitative Economics*, Vol. 8, No. 3, November 2017, 729–760)

Xun Lu
Department of Economics, Hong Kong University of Science and Technology

Liangjun Su
School of Economics, Singapore Management University

This supplement is composed of two parts. Appendix A contains the proofs of the main results in the paper. Appendix B provides some technical lemmas that are used in the proofs of the main results.

Appendix A: Proofs of the main results

Proof of Theorem 3.1. Let $\hat{u}_i = (\hat{u}_{i1}, \ldots, \hat{u}_{iT})'$ and $\bar{P}_{X_i} = X_i(X_i' M_0 X_i)^{-1} X_i'$. Then, by (2.7) and the fact that $\bar{P}_{X_i} M_0 X_i = X_i$, we have

$$\hat{u}_i = M_0 u_i + M_0 X_i \big(\beta_i^0 - \hat{\beta}_i\big) \tag{A.1}$$

and

$$\begin{aligned}
&\sqrt{V_{NT}} J_{1NT}(K_0) \\
&\quad = \left(N^{-1/2} \sum_{i=1}^{N} u_i' M_0 \bar{P}_{X_i} M_0 u_i - B_{NT} \right) + N^{-1/2} \sum_{i=1}^{N} \big(\beta_i^0 - \hat{\beta}_i\big)' X_i' M_0 X_i \big(\beta_i^0 - \hat{\beta}_i\big) \\
&\qquad + 2N^{-1/2} \sum_{i=1}^{N} u_i' M_0 X_i \big(\beta_i^0 - \hat{\beta}_i\big) \\
&\quad \equiv (A_{1NT} - B_{NT}) + A_{2NT} + 2A_{3NT}, \quad \text{say.}
\end{aligned} \tag{A.2}$$

We complete the proof by showing that under $\mathbb{H}_0(K_0)$, (i) $A_{1NT} - B_{NT} \to^D N(0, V_0)$, (ii) $A_{2NT} = o_P(1)$, and (iii) $A_{3NT} = o_P(1)$, where $V_0 = \lim_{(N,T)\to\infty} V_{NT}$. We prove (i)–(iii) in Propositions A.1–A.3 below.

Xun Lu: xunlu@ust.hk
Liangjun Su: ljsu@smu.edu.sg
Acknowledgments made in the leading note in the main paper apply also to this supplement. In particular, Lu acknowledges support from the Hong Kong Research Grants Council (RGC) under Grant 699513. Su gratefully acknowledges the Singapore Ministry of Education for Academic Research Fund under Grant MOE2012-T2-2-021 and the funding support provided by the Lee Kong Chian Fund for Excellence.

DOI: 10.3982/QE517

PROPOSITION A.1. *We have* $A_{1NT} - B_{NT} \to^D N(0, V_0)$ *under* $\mathbb{H}_0(K_0)$.

PROOF. Recall that $H_i = M_0 \bar{P}_{X_i} M_0$ and that $h_{i,ts}$ denotes the (t,s)th element of H_i: $h_{i,ts} = T^{-1} \sum_{r=1}^{T} \sum_{q=1}^{T} \eta_{tr} X_{ir}' (T^{-1} X_i' M_0 X_i)^{-1} X_{iq} \eta_{qs}$, where $\eta_{tr} = \mathbf{1}_{tr} - T^{-1}$ and $\mathbf{1}_{tr} = \mathbf{1}\{t = r\}$. Let $\bar{h}_{i,ts} \equiv T^{-1} \sum_{r=1}^{T} \sum_{q=1}^{T} \eta_{tr} X_{ir}' \Omega_i^{-1} X_{iq} \eta_{qs}$. Observe that $A_{1NT} - B_{NT} = 2N^{-1/2} \sum_{i=1}^{N} \sum_{1 \le s < t \le T} u_{it} \times u_{is} \bar{h}_{i,ts} + 2N^{-1/2} \sum_{i=1}^{N} \sum_{1 \le s < t \le T} u_{it} u_{is} (h_{i,ts} - \bar{h}_{i,ts}) \equiv A_{1NT,1} + A_{1NT,2}$. It suffices to show that (i) $A_{1NT,1} \to^D N(0, V_0)$ and (ii) $A_{1NT,2} = o_P(1)$.

First, we show (i). Using $\eta_{tr} = \mathbf{1}_{tr} - T^{-1}$, we have

$$
\begin{aligned}
A_{1NT,1} &= \frac{2}{T\sqrt{N}} \sum_{i=1}^{N} \sum_{1 \le s < t \le T} \sum_{r=1}^{T} \sum_{q=1}^{T} u_{it} u_{is} \eta_{tr} X_{ir}' \Omega_i^{-1} X_{iq} \eta_{qs} \\
&= \frac{2}{T\sqrt{N}} \sum_{i=1}^{N} \sum_{1 \le s < t \le T} u_{it} u_{is} X_{it}' \Omega_i^{-1} X_{is} - \frac{2}{T^2\sqrt{N}} \sum_{i=1}^{N} \sum_{1 \le s < t \le T} \sum_{r=1}^{T} u_{it} u_{is} X_{ir}' \Omega_i^{-1} X_{is} \\
&\quad - \frac{2}{T^2\sqrt{N}} \sum_{i=1}^{N} \sum_{1 \le s < t \le T} \sum_{q=1}^{T} u_{it} u_{is} X_{it}' \Omega_i^{-1} X_{iq} \\
&\quad + \frac{2}{T^3\sqrt{N}} \sum_{i=1}^{N} \sum_{1 \le s < t \le T} \sum_{r=1}^{T} \sum_{q=1}^{T} u_{it} u_{is} X_{ir}' \Omega_i^{-1} X_{iq} \\
&= \frac{2}{T\sqrt{N}} \sum_{i=1}^{N} \sum_{1 \le s < t \le T} u_{it} u_{is} X_{it}^{\dagger\prime} \Omega_i^{-1} X_{is}^{\dagger} \\
&\quad - \frac{2}{T^2\sqrt{N}} \sum_{i=1}^{N} \sum_{1 \le s < t \le T} \sum_{r=1}^{T} u_{it} u_{is} \big[X_{ir} - E(X_{ir})\big]' \Omega_i^{-1} X_{is} \\
&\quad - \frac{2}{T^2\sqrt{N}} \sum_{i=1}^{N} \sum_{1 \le s < t \le T} \sum_{q=1}^{T} u_{it} u_{is} X_{it}' \Omega_i^{-1} \big[X_{iq} - E(X_{iq})\big] \\
&\quad + \frac{2}{T^3\sqrt{N}} \sum_{i=1}^{N} \sum_{1 \le s < t \le T} \sum_{r=1}^{T} \sum_{q=1}^{T} u_{it} u_{is} \big[X_{ir} - E(X_{ir})\big]' \Omega_i^{-1} \big[X_{iq} - E(X_{iq})\big] \\
&\quad + \frac{4}{T^3\sqrt{N}} \sum_{i=1}^{N} \sum_{1 \le s < t \le T} \sum_{r=1}^{T} \sum_{q=1}^{T} u_{it} u_{is} \big[X_{ir} - E(X_{ir})\big]' \Omega_i^{-1} E(X_{iq}) \\
&\equiv A_{1NT,11} + A_{1NT,12} + A_{1NT,13} + A_{1NT,14} + A_{1NT,15}, \quad \text{say.}
\end{aligned}
$$

By Lemma B.4(ii)–(v), $A_{1NT,12} + A_{1NT,13} + A_{1NT,14} + A_{1NT,15} = o_P(1)$. We are left to show that $A_{1NT,11} \to^D N(0, V_0)$. Observe that

$$
A_{1NT,11} = \frac{2}{T\sqrt{N}} \sum_{i=1}^{N} \sum_{1 \le s < t \le T} u_{it} u_{is} X_{it}^{\dagger\prime} \Omega_i^{-1} X_{is}^{\dagger} = \sum_{t=2}^{T} Z_{NT,t},
$$

where $Z_{NT,t} \equiv 2T^{-1}N^{-1/2}\sum_{i=1}^{N}\sum_{s=1}^{t-1} u_{it}u_{is}\bar{b}_{it}'\bar{b}_{is}$ and $\bar{b}_{it} \equiv \Omega_i^{-1/2}X_{it}^{\dagger}$. By Assumption 1(v),

$$E(Z_{NT,t} \mid \mathcal{F}_{NT,t-1}) \equiv 2T^{-1}N^{-1/2}\sum_{i=1}^{N}\sum_{s=1}^{t-1} u_{is}\bar{b}_{it}'\bar{b}_{is}E(u_{it} \mid \mathcal{F}_{NT,t-1}) = 0.$$

That is, $\{Z_{NT,t}, \mathcal{F}_{NT,t}\}$ is an m.d.s. By the martingale central limit theorem (CLT) (e.g., Pollard (1984, p. 171)), it suffices to show that

$$\mathcal{Z} \equiv \sum_{t=2}^{T} E_{\mathcal{F}_{NT,t-1}}\big[|Z_{NT,t}|^4\big] = o_P(1), \quad \text{and} \quad \sum_{t=2}^{T} Z_{NT,t}^2 - V_{NT} = o_P(1), \tag{A.3}$$

where $E_{\mathcal{F}_{NT,t-1}}$ denotes expectation conditional on $\mathcal{F}_{NT,t-1}$. Observing that $\mathcal{Z} \ge 0$, it suffices to show that $\mathcal{Z} = o_P(1)$ by showing that $E(\mathcal{Z}) = o_P(1)$ by Markov inequality. By Assumption 1(iv) and (v),

$$\begin{aligned} E(\mathcal{Z}) = {} & \frac{16}{T^4N^2} \\ & \times \sum_{t=2}^{T}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N}\sum_{l=1}^{N}\sum_{1\le r,s,q,v\le t-1} E\big(\bar{b}_{it}'\bar{b}_{is}\bar{b}_{jt}'\bar{b}_{jr}\bar{b}_{kt}'\bar{b}_{kq}\bar{b}_{lt}'\bar{b}_{lv}u_{is}u_{jr}u_{kq}u_{lv}u_{it}u_{jt}u_{kt}u_{lt}\big) \\ = {} & 48\mathcal{Z}_1 + 16\mathcal{Z}_2, \end{aligned}$$

where

$$\begin{aligned} \mathcal{Z}_1 \equiv {} & \frac{1}{T^4N^2}\sum_{t=2}^{T}\sum_{i=1}^{N}\sum_{j=1,j\ne i}^{N}\sum_{1\le r,s,q,v\le t-1} \\ & \times E\big(\bar{b}_{it}'\bar{b}_{is}\bar{b}_{it}'\bar{b}_{ir}u_{is}u_{ir}u_{it}^2\big)E\big(\bar{b}_{jt}'\bar{b}_{jq}\bar{b}_{jt}'\bar{b}_{jv}u_{jq}u_{jv}u_{jt}^2\big), \end{aligned} \tag{A.4}$$

$$\mathcal{Z}_2 \equiv \frac{1}{T^4N^2}\sum_{t=2}^{T}\sum_{i=1}^{N}\sum_{1\le r,s,q,v\le t-1} E\big(\bar{b}_{it}'\bar{b}_{is}\bar{b}_{it}'\bar{b}_{ir}\bar{b}_{it}'\bar{b}_{iq}\bar{b}_{it}'\bar{b}_{iv}u_{is}u_{ir}u_{iq}u_{iv}u_{it}^4\big). \tag{A.5}$$

For the moment we assume that $p = 1$ so that we can treat the $p \times 1$ vector $\bar{b}_{it}$ as a scalar. (The general case follows from the Slutsky lemma and the fact that $\bar{b}_{it}'\bar{b}_{is}\bar{b}_{it}'\bar{b}_{ir} = \sum_{k=1}^{p}\sum_{l=1}^{p}\bar{b}_{it,k}\bar{b}_{is,k}\bar{b}_{it,l}\bar{b}_{ir,l}$, where $\bar{b}_{it,k}$ denotes the kth element of $\bar{b}_{it}$.) To bound the summation in (A.4), we consider three cases for the time indices in $S \equiv \{r, s, q, v, t-1\}$: (a) $\#S = 5$, (b) $\#S = 4$, and (c) $\#S \le 3$. We use EZ_{1a}, EZ_{1b}, and EZ_{1c} to denote the corresponding summations when the time indices are restricted to cases (a), (b), and (c), respectively. In case (a), applying Davydov inequality (e.g., Hall and Heyde (1980, p. 278)) yields

$$\big|E\big(\bar{b}_{it}\bar{b}_{is}\bar{b}_{it}\bar{b}_{ir}u_{is}u_{ir}u_{it}^2\big)\big| \le 8C_{i,tsr}(t,s,r)\alpha\big(t-1-(s\vee r)\big)^{(1+\sigma)/(2+\sigma)}, \tag{A.6}$$

where $a \vee b \equiv \max(a, b)$ and $C_{1i,tsr} \equiv \max_{i,t,s,r}\|\bar{b}_{is}\bar{b}_{ir}u_{is}u_{ir}\|_{4+2\sigma}\|\bar{b}_{it}^2u_{it}^2\|_{4+2\sigma}$. A similar inequality holds for $E(\bar{b}_{jt}\bar{b}_{jq}\bar{b}_{jt}\bar{b}_{jv}u_{jq}u_{jv}u_{jt}^2)$. By the repeated use of Cauchy–Schwarz and

Jensen inequalities and Assumption 1(i),

$$
\begin{aligned}
|C_{1i,tsr}| &\leq \frac{1}{2}\big[\|\bar{b}_{is}\bar{b}_{ir}u_{is}u_{ir}\|_{4+2\sigma}^2 + \big\|\bar{b}_{it}^2 u_{it}^2\big\|_{4+2\sigma}^2\big] \\
&\leq \frac{1}{4}\big\{\|\bar{b}_{is}u_{is}\|_{8+4\sigma}^2 + \|\bar{b}_{ir}u_{ir}\|_{8+4\sigma}^2 + 2\big\|\bar{b}_{it}^2 u_{it}^2\big\|_{4+2\sigma}^2\big\} \leq C_1
\end{aligned}
\tag{A.7}
$$

for some $C_1 < \infty$. With this, we can readily show that under Assumption 1(iii),

$$
EZ_{1a} \leq \frac{64C_1^2}{T^4}\sum_{t=2}^{T}\bigg\{\sum_{1\leq s,r\leq t-1}\alpha\big(t-1-(s\vee r)\big)^{(1+\sigma)/(2+\sigma)}\bigg\}^2 = O\big(T^{-1}\big).
$$

In case (b), we consider two subcases: (b1) one and only one of r, s, q, v equals $t-1$; (b2) $\#\{r, s, q, v\} = 3$. We use EZ_{1b1} and EZ_{1b2} to denote the corresponding summations when the individual indices are restricted to subcases (b1) and (b2), respectively. In subcase (b1), wlog we assume that $v = t-1$ and apply

$$
\big|E\big(\bar{b}_{jt}\bar{b}_{jq}\bar{b}_{jt}\bar{b}_{j,t-1}u_{jq}u_{j,t-1}u_{jt}^2\big)\big| \leq 8C_{2j,tq}\alpha(t-1-q)^{(1+\sigma)/(2+\sigma)}
$$

for $C_{2j,tq} \equiv \|\bar{b}_{jq}u_{jq}\|_{8+4\sigma,\mathcal{D}}\|\bar{b}_{jt}^2\bar{b}_{j,t-1}u_{j,t-1}u_{jt}^2\|_{(8+4\sigma)/3,\mathcal{D}} \leq C_2$ for some $C_2 < \infty$ and (A.6) and (A.7) to obtain

$$
\begin{aligned}
EZ_{1b1} &\leq \frac{64C_1C_2}{T^3}\sum_{t=2}^{T}\bigg\{\frac{1}{T}\sum_{1\leq s,r\leq t-1}\alpha\big(t-1-(s\vee r)\big)^{(1+\sigma)/(2+\sigma)}\bigg\} \\
&\quad \times \bigg\{\sum_{1\leq q\leq t-1}\alpha(t-1-q)^{(1+\sigma)/(2+\sigma)}\bigg\} \\
&= O\big(T^{-2}\big).
\end{aligned}
$$

In subcase (b2), wlog we assume that $q = v$ and $r < s < t-1$. We consider two subsubcases: (b21) either $t-1-s > \tau_*$ or $s-r > \tau_*$; (b22) $t-1-s \leq \tau_*$ and $s-t \leq \tau_*$. In the first case, we have

$$
\big|E\big(\bar{b}_{it}\bar{b}_{is}\bar{b}_{it}\bar{b}_{ir}u_{is}u_{ir}u_{it}^2\big)\big| \leq \begin{cases} 8C_{3i,tsr}\alpha(\tau_*)^{(1+\sigma)/(2+\sigma)} & \text{if } t-1-s > \tau_*, \\ 8C_{4i,tsr}(t,s,r)\alpha(\tau_*)^{(1+\sigma)/(2+\sigma)} & \text{if } s-r > \tau_*, \end{cases}
$$

where $C_{3i,tsr} \equiv \|\bar{b}_{it}\bar{b}_{it}u_{it}^2\|_{4+2\sigma}\|\bar{b}_{is}\bar{b}_{ir}u_{is}u_{ir}\|_{4+2\sigma} \leq C_3 < \infty$ and $C_{4i,tsr} \equiv \|\bar{b}_{it}\bar{b}_{is}\bar{b}_{it}u_{is} \times u_{it}^2\|_{(8+4\sigma)/3} \times \|\bar{b}_{ir}u_{ir}\|_{8+4\sigma} \leq C_4 < \infty$. These results, in conjunction with the fact that the total number of terms in the summation in subcase (b22) is of $O(N^2T^3\tau_*^2)$ and Assumption 1(iii), imply that

$$
\begin{aligned}
EZ_{1b2} &\leq O\big[T^2\alpha(\tau_*)^{(1+\sigma)/(2+\sigma)} + T^{-4}N^{-2}N^2T^3\tau_*^2\big] \\
&= O\big(T^2\alpha(\tau_*)^{(1+\sigma)/(2+\sigma)} + T^{-1}\tau_*^2\big) = o(1).
\end{aligned}
$$

Consequently, $EZ_{1b} = o(1)$. In case (c), we have $EZ_{1c} = O(T^{-1})$ as the number of terms in the summation is $O(N^2T^3)$ and each term in absolute value has a bounded expectation. It follows that $\mathcal{Z}_1 = o_P(1)$.

To bound $\mathcal{Z}_2$, we consider two cases for the set of indices $S \equiv \{r, s, q, v, t-1\}$: (a) $\#S = 5$ and (b) all the other cases. We use EZ_{2a} and EZ_{2b} to denote the corresponding summations when the individual indices are restricted to subcases (a) and (b), respectively. In the first case, letting $c = \max(s, r, q, v)$, we have

$$\left|E\left(\bar{b}_{it}^4 \bar{b}_{is}\bar{b}_{ir}\bar{b}_{iq}\bar{b}_{iv}u_{is}u_{ir}u_{iq}u_{iv}u_{it}^4\right)\right| \le 8C_{5i,t,s,r,q,v}(t, s, r, q, v)\alpha(t-1-c)^{\sigma/(2+\sigma)},$$

where $C_{5i,t,s,r,q,v} \equiv \|\bar{b}_{is}\bar{b}_{ir}\bar{b}_{iq}\bar{b}_{iv}u_{is}u_{ir}u_{iq}u_{iv}\|_{2+\sigma}\|\bar{b}_{it}^4 u_{it}^4\|_{2+\sigma} \le C_5 < \infty$. Then $EZ_{2a} \le 8CN^{-1}\sum_{s=1}^{T}\alpha(s)^{\sigma/(2+\sigma)} = O(N^{-1})$. In case (b), we have $EZ_{2b} = O(N^{-1})$. It follows that $\mathcal{Z}_2 = O(N^{-1})$ and thus $\mathcal{Z} = o_P(1)$. Consequently the first part of (A.3) follows.

For the second part of (A.3), by Assumption 1(iv) and (v), we have

$$\begin{aligned}\sum_{t=2}^{T} E\left(Z_{NT,t}^2\right) &= 4T^{-2}N^{-1}\sum_{t=2}^{T} E\left[\sum_{i=1}^{N}\sum_{s=1}^{t-1} u_{it}u_{is}\bar{b}_{it}'\bar{b}_{is}\right]^2 \\ &= 4T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\sum_{s=1}^{t-1}\sum_{r=1}^{t-1} E\left(u_{it}^2 u_{is}u_{ir}\bar{b}_{it}'\bar{b}_{is}\bar{b}_{it}'\bar{b}_{ir}\right) = V_{NT}.\end{aligned}$$

In addition, we can show by straightforward moment calculations that $E(\sum_{t=2}^{T} Z_{NT,t}^2)^2 = V_{NT}^2 + o(1)$. Thus $\mathrm{Var}(\sum_{t=2}^{T} Z_{NT,t}^2) = o(1)$ and the second part of (A.3) follows. This completes the proof of (i).

In addition, by Lemma B.4(i), $A_{1NT,2} = o_P(1)$. □

Proposition A.2. *We have* $A_{2NT} = o_P(1)$ *under* $\mathbb{H}_0(K_0)$.

Proof. Noting that $\mathbf{1}\{i \in G_k^0\} = \mathbf{1}\{i \in \hat{G}_k\} + \mathbf{1}\{i \in G_k^0 \setminus \hat{G}_k\} - \mathbf{1}\{i \in \hat{G}_k \setminus G_k^0\}$, under $\mathbb{H}_0(K_0)$ we have

$$\begin{aligned}A_{2NT} &= N^{-1/2}\sum_{k=1}^{K_0}\sum_{i\in G_k^0}\left(\alpha_k^0 - \hat{\beta}_i\right)' X_i' M_0 X_i\left(\alpha_k^0 - \hat{\beta}_i\right) \\ &= N^{-1/2}\sum_{k=1}^{K_0}\sum_{i\in \hat{G}_k}\left(\alpha_k^0 - \hat{\alpha}_k\right)' X_i' M_0 X_i\left(\alpha_k^0 - \hat{\alpha}_k\right) \\ &\quad + N^{-1/2}\sum_{k=1}^{K_0}\sum_{i\in G_k^0\setminus \hat{G}_k}\left(\alpha_k^0 - \hat{\beta}_i\right)' X_i' M_0 X_i\left(\alpha_k^0 - \hat{\beta}_i\right) \\ &\quad - N^{-1/2}\sum_{k=1}^{K_0}\sum_{i\in \hat{G}_k\setminus G_k^0}\left(\alpha_k^0 - \hat{\alpha}_k\right)' X_i' M_0 X_i\left(\alpha_k^0 - \hat{\alpha}_k\right) \\ &\equiv A_{2NT,1} + A_{2NT,2} - A_{2NT,3}, \quad \text{say.}\end{aligned}$$

Let $\|\cdot\|_{\mathrm{sp}}$ denote the spectral norm. Note that $\|A\| \le \mathrm{rank}(A)\|A\|_{\mathrm{sp}}$ and $\|A\|_{\mathrm{sp}} \le \|A\|$ for any matrix A. By these properties, the submultiplicative property of the spectral norm, the fact that $\|M_0\|_{\mathrm{sp}} = 1$, and Assumptions 1(i) and 2(ii),

$$A_{2NT,1} \le N^{-1/2} \sum_{k=1}^{K_0} \big\|\alpha_k^0 - \hat{\alpha}_k\big\|^2 \sum_{i \in \hat{G}_k} \big\|X_i' M_0 X_i\big\| \le p N^{-1/2} \sum_{k=1}^{K_0} \big\|\alpha_k^0 - \hat{\alpha}_k\big\|^2 \sum_{i \in \hat{G}_k} \|X_i\|^2$$
$$= N^{-1/2} O_P\big((NT)^{-1} + T^{-2}\big) O_P(NT) = N^{-1/2} O_P\big(1 + NT^{-1}\big) = o_P(1).$$

By Assumption 2(iii), for any $\epsilon > 0$,

$$P(A_{2NT,2} \ge \epsilon) \le P\left(\bigcup_{k=1}^{K_0} \hat{E}_{kNT}\right) \to 0, \quad \text{and} \quad P(A_{2NT,3} \ge \epsilon) \le P\left(\bigcup_{k=1}^{K_0} \hat{F}_{kNT}\right) \to 0.$$

It follows that $A_{2NT,2} = o_P(1)$ and $A_{2NT,3} = o_P(1)$. Consequently, $A_{2NT} = o_P(1)$ under $\mathbb{H}_0(K_0)$. □

PROPOSITION A.3. *We have* $A_{3NT} = o_P(1)$ *under* $\mathbb{H}_0(K_0)$.

PROOF. As in the proof of Proposition A.2, we make the decomposition

$$\begin{aligned} A_{3NT} &= N^{-1/2} \sum_{k=1}^{K_0} \sum_{i \in G_k^0} u_i' M_0 X_i \big(\alpha_k^0 - \hat{\beta}_i\big) \\ &= N^{-1/2} \sum_{k=1}^{K_0} \sum_{i \in \hat{G}_k} u_i' M_0 X_i \big(\alpha_k^0 - \hat{\alpha}_k\big) + N^{-1/2} \sum_{k=1}^{K_0} \sum_{i \in G_k^0 \setminus \hat{G}_k} u_i' M_0 X_i \big(\alpha_k^0 - \hat{\beta}_i\big) \\ &\quad - N^{-1/2} \sum_{k=1}^{K_0} \sum_{i \in \hat{G}_k \setminus G_k^0} u_i' M_0 X_i \big(\alpha_k^0 - \hat{\alpha}_k\big) \\ &\equiv A_{3NT,1} + A_{3NT,2} - A_{3NT,3}, \quad \text{say.} \end{aligned}$$

Using the same arguments as those used in the study of $A_{2NT,2}$ and $A_{2NT,3}$, we can show that $A_{3NT,2} = o_P(1)$ and $A_{3NT,3} = o_P(1)$. Noting that $\alpha_k^0 - \hat{\alpha}_k = O_P((NT)^{-1/2} + T^{-1})$ for $k = 1, \ldots, K_0$ under $\mathbb{H}_0(K_0)$ by Assumption 2(ii), it suffices to prove that $A_{3NT,1} = o_P(1)$ by showing that

$$\bar{A}_{3NT,1k} \equiv N^{-1/2} \sum_{i \in \hat{G}_k} X_i' M_0 u_i = o_P\big(\min\big((NT)^{1/2}, T\big)\big) \quad \text{for } k = 1, \ldots, K_0.$$

By the fact that $\mathbf{1}\{i \in \hat{G}_k\} = \mathbf{1}\{i \in G_k^0\} + \mathbf{1}\{i \in \hat{G}_k \setminus G_k^0\} - \mathbf{1}\{i \in G_k^0 \setminus \hat{G}_k\}$ and the arguments used in the study of $A_{2NT,2}$ and $A_{2NT,3}$, we can show that $\bar{A}_{3NT,1k} \equiv \dot{A}_{3NT,1k} + o_P(1)$, where $\dot{A}_{3NT,1k} = N^{-1/2} \sum_{i \in G_k^0} X_i' M_0 u_i$. Using $X_i' M_0 u_i = \sum_{t=1}^{T} X_{it}(u_{it} - \bar{u}_{i\cdot})$, we can decompose $\dot{A}_{3NT,1k}$ as

$$\dot{A}_{3NT,1k} = N^{-1/2} \sum_{i \in G_k^0} \sum_{t=1}^{T} X_{it} u_{it} - N^{-1/2} T^{-1} \sum_{i \in G_k^0} \sum_{t=1}^{T} \sum_{s=1}^{T} X_{it} u_{is}$$

$$
\begin{aligned}
&= (1 - T^{-1}) N^{-1/2} \sum_{i \in G_k^0} \sum_{t=1}^{T} X_{it} u_{it} - N^{-1/2} T^{-1} \sum_{i \in G_k^0} \sum_{1 \le t < s \le T} X_{it} u_{is} \\
&\quad - N^{-1/2} T^{-1} \sum_{i \in G_k^0} \sum_{1 \le s < t \le T} X_{it} u_{is} \\
&\equiv \dot{A}_{3NT,1k1} - \dot{A}_{3NT,1k2} - \dot{A}_{3NT,1k3}, \quad \text{say.}
\end{aligned}
$$

Using Chebyshev inequality, we can readily show that $\dot{A}_{3NT,1k1} = O_P(T^{1/2})$ under Assumption 1(i), (iv), and (v). Let $\omega_p = (\omega_{p1}, \ldots, \omega_{pp})'$ be an arbitrary $p \times 1$ nonrandom vector with $\|\omega_p\| = 1$. By Assumption 1(i), (iv), and (v) and Jensen inequality,

$$
\begin{aligned}
E\big[(\omega_p' \dot{A}_{3NT,1k2})^2\big] &= N^{-1} T^{-2} \sum_{i \in G_k^0} \sum_{j \in G_k^0} \sum_{1 \le t < s \le T} \sum_{1 \le r < q \le T} E(\omega_p' X_{it} u_{is} \omega_p' X_{jr} u_{jq}) \\
&= N^{-1} T^{-2} \sum_{i \in G_k^0} \sum_{1 \le t < s \le T} \sum_{1 \le r < q \le T} E(\omega_p' X_{it} u_{is} \omega_p' X_{ir} u_{iq}) \\
&= N^{-1} T^{-2} \sum_{i \in G_k^0} \sum_{1 \le t, r < s \le T} E(\omega_p' X_{it} \omega_p' X_{ir} u_{is}^2) = O(T).
\end{aligned}
$$

Then $\dot{A}_{3NT,1k2} = O_P(T^{1/2})$ by Chebyshev inequality. Next,

$$
\begin{aligned}
E\big[(\omega_p' \dot{A}_{3NT,1k3})^2\big] &= N^{-1} T^{-2} \sum_{i \in G_k^0} \sum_{1 \le s < t \le T} \sum_{1 \le q < r \le T} E(\omega_p' X_{it} u_{is} \omega_p' X_{ir} u_{iq}) \\
&\quad + N^{-1} T^{-2} \sum_{i \in G_k^0} \sum_{j \in G_k^0, j \ne i} \sum_{1 \le s < t \le T} \sum_{1 \le q < r \le T} E(\omega_p' X_{it} u_{is}) E(\omega_p' X_{jr} u_{jq}) \\
&\equiv \mathrm{I} + \mathrm{II}, \quad \text{say.}
\end{aligned}
$$

Let $S \equiv \{t, s, q, r\}$. To bound I, we consider two cases, (a) $\#S = 4$ and (b) $\#S \le 3$, and denote the corresponding summations as I_a and I_b such that $\mathrm{I} = I_a + I_b$. Apparently, $I_b = O(T)$. For I_a, wlog we consider three subcases, (a1) $s < t < q < r$, (a2) $s < q < t < r$, and (a3) $s < q < r < t$, and denote the corresponding summations as I_{a1}, I_{a2}, and I_{a3}, respectively. (Note that $I_a = 2(I_{a1} + I_{a2} + I_{a3})$.) In subcase (a1), we apply Davydov inequality to obtain

$$
\begin{aligned}
|I_{a1}| &\le 8 N^{-1} T^{-2} \sum_{i \in G_k^0} \sum_{1 \le s < t < q < r \le T} c_{i,tsrq} \alpha(t - s)^{(1+\sigma)/(2+\sigma)} \\
&\le 8 C T \sum_{\tau=1}^{\infty} \alpha(\tau)^{(1+\sigma)/(2+\sigma)} = O(T),
\end{aligned}
$$

where $c_{i,tsrq} = \|u_{is}\|_{8+4\sigma} \|\omega_p' X_{it} \omega_p' X_{ir} u_{iq}\|_{(8+4\sigma)/3} \le C < \infty$ by Assumption 1(i) and Jensen inequality. Analogously, we can show that $I_{a2} = O(T)$ and $I_{a3} = O(T)$. It follows that $\mathrm{I} = O(T)$. For II, we apply Davydov inequality to obtain

$$|\mathrm{II}| \leq N^{-1}T^{-2}\Big\{\sum_{i\in G_k^0}\sum_{1\leq s<t\leq T}\big|E\big(\omega_p' X_{it}u_{is}\big)\big|\Big\}^2$$
$$\leq N^{-1}T^{-2}\Big\{\sum_{i\in G_k^0}\sum_{1\leq s<t\leq T} c_{i,ts}\alpha(t-s)^{(3+2\sigma)/(4+2\sigma)}\Big\}^2$$
$$= N^{-1}T^{-2}O\big(N^2T^2\big) = O(N),$$

where $c_{i,ts} = \|\omega_p' X_{it}\|_{8+4\sigma}\|\omega_p' X_{it}\|_{8+4\sigma} \leq C < \infty$ by Assumption 1(i). Consequently,

$$E\big\{\big[\omega_p' \dot{A}_{3NT,1k1}(3)\big]^2\big\} = O(N+T) \quad \text{and} \quad \dot{A}_{3NT,1k3} = O_P\big(N^{1/2}+T^{1/2}\big).$$

In sum, we have $\dot{A}_{3NT,1k} = O_P(N^{1/2}+T^{1/2})$. It follows that $\bar{A}_{3NT,1k} = o_P(\min((N\times T)^{1/2}, T))$. □

□

Proof of Theorem 3.2. By Theorem 3.1 and the Slutsky lemma, it suffices to prove the first two parts of the theorem.

Step 1: We prove (i) $\hat{B}_{NT}(K_0) = B_{NT} + o_P(1)$ under $\mathbb{H}_0(K_0)$. Let ι_t denote a $T\times 1$ vector with 1 in the tth position and 0s everywhere else. Then $h_{i,ts} = \iota_t' M_0\bar{P}_{X_i}M_0\iota_s = \sum_{r=1}^T\sum_{q=1}^T \eta_{tr}X_{ir}'(X_iM_0X_i)^{-1}X_{iq}\eta_{qs}$. Using $\hat{u}_{it}^2 - u_{it}^2 = (\hat{u}_{it}-u_{it})^2 + 2(\hat{u}_{it}-u_{it})u_{it}$, we decompose $\hat{B}_{NT} - B_{NT}$ as

$$\hat{B}_{NT}(K_0) - B_{NT} = \frac{1}{\sqrt{N}}\sum_{i=1}^N\sum_{t=1}^T(\hat{u}_{it}-u_{it})^2h_{i,tt} + \frac{2}{\sqrt{N}}\sum_{i=1}^N\sum_{t=1}^T(\hat{u}_{it}-u_{it})u_{it}h_{i,tt}$$
$$\equiv \hat{B}_{NT,1} + 2\hat{B}_{NT,2}, \quad \text{say.}$$

Noting that $\mathrm{diag}(H_i)$ is p.s.d., we have by (A.1) and Cauchy–Schwarz inequality that

$$\hat{B}_{NT,1} = N^{-1/2}\sum_{i=1}^N(\hat{u}_i-u_i)'\,\mathrm{diag}(H_i)(\hat{u}_i-u_i)$$
$$\leq 2N^{-1/2}\sum_{i=1}^N u_i'P_0\,\mathrm{diag}(H_i)P_0u_i$$
$$+ 2N^{-1/2}\sum_{i=1}^N\big(\hat{\beta}_i-\beta_i^0\big)'X_i'M_0\,\mathrm{diag}(H_i)M_0X_i\big(\hat{\beta}_i-\beta_i^0\big)$$
$$\equiv 2\hat{B}_{NT,11} + 2\hat{B}_{NT,12}, \quad \text{say.}$$

We will show that $\hat{B}_{NT,1s} = o_P(1)$ for $s=1$ and 2. By the fact that $\sum_{t=1}^T\iota_t\iota_t' = I_T$ and M_0 is idempotent, we have $\mathbf{i}_T'\,\mathrm{diag}(H_i)\mathbf{i}_T = \mathrm{tr}[\mathbf{i}_T'\,\mathrm{diag}(H_i)\mathbf{i}_T] = \sum_{t=1}^T\mathrm{tr}(\iota_t'M_0\bar{P}_{X_i}M_0\iota_t) = \mathrm{tr}(M_0\bar{P}_{X_i}M_0) = \mathrm{tr}[M_0\times X_i(X_i'M_0X_i)^{-1}X_i'M_0] = p$. This, in conjunction with Davydov inequality, implies that

$$
\begin{aligned}
E|\hat{B}_{NT,11}| &= T^{-2}N^{-1/2}\sum_{i=1}^{N} E\big[u_i'\mathbf{i}_T\big(\mathbf{i}_T' \operatorname{diag}(H_i)\mathbf{i}_T\big)\mathbf{i}_T' u_i\big] = pT^{-2}N^{-1/2}\sum_{i=1}^{N} E\big(u_i'\mathbf{i}_T\mathbf{i}_T' u_i\big) \\
&= pT^{-2}N^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T} E\big(u_{it}^2\big) + 2pT^{-2}N^{-1/2}\sum_{i=1}^{N}\sum_{1\le t<s\le T} E(u_{it}u_{is}) \\
&= O\big(N^{1/2}T^{-1}\big) + O\big(N^{1/2}T^{-1}\big) = O\big(N^{1/2}T^{-1}\big).
\end{aligned}
$$

Consequently $\hat{B}_{NT,11} = O_P(N^{1/2}T^{-1}) = o_P(1)$ by Markov inequality. Using $\mathbf{1}\{i\in G_k^0\} = \mathbf{1}\{i\in \hat{G}_k\} + \mathbf{1}\{i\in G_k^0 \setminus \hat{G}_k\} - \mathbf{1}\{i\in \hat{G}_k \setminus G_k^0\}$, following similar arguments as those used in the proof of Proposition A.2, and by Assumptions 1(i) and 2(ii), we can show that under $\mathbb{H}_0(K_0)$,

$$
\begin{aligned}
\hat{B}_{NT,12} &= N^{-1/2}\sum_{k=1}^{K_0}\sum_{i\in\hat{G}_k}\big(\hat{\alpha}_k-\alpha_k^0\big)' X_i' M_0 \operatorname{diag}(H_i) M_0 X_i\big(\hat{\alpha}_k-\alpha_k^0\big) + o_P(1) \\
&\le N^{-1/2}\sum_{k=1}^{K_0}\big\|\hat{\alpha}_k-\alpha_k^0\big\|^2 \sum_{i\in\hat{G}_k}\big\|X_i' M_0 \operatorname{diag}(H_i) M_0 X_i\big\| + o_P(1) \\
&= N^{-1/2} O_P\big((NT)^{-1}+T^{-2}\big) O_P(N) + o_P(1) = o_P(1),
\end{aligned}
$$

based on the fact that $M_0 A M_0 \le A$ for any p.s.d. matrix A, and

$$
\begin{aligned}
&\sum_{i\in\hat{G}_k}\big\|X_i' M_0 \operatorname{diag}(H_i) M_0 X_i\big\| \\
&\quad\le \sum_{i\in\hat{G}_k}\big\|X_i' \operatorname{diag}(H_i) X_i\big\| \le \sum_{i\in\hat{G}_k}\sum_{t=1}^{T}\big\|X_{it}'\iota_t' M_0 \bar{P}_{X_i} M_0 \iota_t X_{it}\big\| \\
&\quad\le \sum_{i\in\hat{G}_k}\sum_{t=1}^{T}\big\|X_{it}'\iota_t' \bar{P}_{X_i} \iota_t X_{it}\big\| = \sum_{i\in\hat{G}_k}\sum_{t=1}^{T}\big\|X_{it}'X_{it}\big(X_i' M_0 X_i\big)^{-1} X_{it}' X_{it}\big\| \\
&\quad\le \max_{1\le i\le N}\big\|\hat{\Omega}_i^{-1}\big\| T^{-1}\sum_{i\in\hat{G}_k}\sum_{t=1}^{T}\|X_{it}\|^4 = O_P(N) \quad \text{by Lemma B.3(v)}.
\end{aligned}
$$

Consequently, we have shown that $\hat{B}_{NT,1} = o_P(1)$.

For $\hat{B}_{NT,2}$, we first apply (A.1) to decompose it as

$$
\begin{aligned}
\hat{B}_{NT,2} &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\hat{u}_i-u_i)' \operatorname{diag}(H_i) u_i \\
&= \frac{1}{\sqrt{N}}\sum_{i=1}^{N} u_i' P_0 \operatorname{diag}(H_i) u_i + \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\big(\hat{\beta}_i-\beta_i^0\big)' X_i' M_0 \operatorname{diag}(H_i) u_i \equiv \hat{B}_{NT,21} + \hat{B}_{NT,22}.
\end{aligned}
$$

Observe that

$$\begin{aligned}\hat{B}_{NT,21} &= \frac{1}{T^2\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T} u_{it}\iota_s' M_0 X_i \hat{\Omega}_i^{-1} X_i' M_0 \iota_s u_{is} \\ &= \frac{1}{T^2\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T} u_{it}\iota_s' M_0 X_i \Omega_i^{-1} X_i' M_0 \iota_s u_{is} \\ &\quad + \frac{1}{T^2\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T} u_{it}\iota_s' M_0 X_i \big[\hat{\Omega}_i^{-1} - \Omega_i^{-1}\big] X_i' M_0 \iota_s u_{is} \\ &\equiv \hat{B}_{NT,211} + \hat{B}_{NT,212}, \quad \text{say.}\end{aligned}$$

Noting that $\|\iota_s' M_0 X_i\| \le \|\iota_s' X_i\| = \|X_{is}\|$, we apply Lemma B.3(v) to obtain

$$\begin{aligned}|\hat{B}_{NT,212}| &\le \max_{1\le i\le N}\big\|\hat{\Omega}_i^{-1} - \Omega_i^{-1}\big\| \frac{1}{T^2\sqrt{N}}\sum_{i=1}^{N}\left|\sum_{t=1}^{T} u_{it}\right|\sum_{s=1}^{T}\|X_{is}\|^2\|u_{is}\| \\ &= O_P(a_{NT})O_P\big(N^{1/2}T^{-1/2}\big) = o_P(1).\end{aligned}$$

For $\hat{B}_{NT,211}$, we have

$$\begin{aligned}\hat{B}_{NT,211} &= \frac{1}{T^2\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}\sum_{q=1}^{T} u_{it}\eta_{sr} X_{ir}' \Omega_i^{-1} X_{iq}\eta_{qs} u_{is} \\ &= \frac{1}{T^2\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T} u_{it} X_{is}' \Omega_i^{-1} X_{is} u_{is} - \frac{2}{T^3\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T} u_{it} X_{ir}' \Omega_i^{-1} X_{is} u_{is} \\ &\quad + \frac{1}{T^4\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}\sum_{q=1}^{T} u_{it} X_{ir}' \Omega_i^{-1} X_{iq} u_{is} \\ &\equiv \hat{B}_{NT,211a} - 2\hat{B}_{NT,211b} + \hat{B}_{NT,221c}, \quad \text{say.}\end{aligned}$$

We further decompose $\hat{B}_{NT,211a}$ as $\hat{B}_{NT,211a} = \frac{1}{T^2\sqrt{N}}\sum_{i=1}^{N}\sum_{t=1}^{T} X_{it}'\Omega_i^{-1}X_{it}u_{it}^2 + \frac{1}{T^2\sqrt{N}} \times \sum_{i=1}^{N}\sum_{1\le t<s\le T} u_{it}X_{is}'\Omega_i^{-1}X_{is}u_{is} + \frac{1}{T^2\sqrt{N}}\sum_{i=1}^{N}\sum_{1\le s<t\le T} u_{it}X_{is}'\Omega_i^{-1}X_{is}u_{is} = \hat{B}_{NT,211a}(1) + \hat{B}_{NT,211a}(2) + \hat{B}_{NT,211a}(3)$. Apparently, $\hat{B}_{NT,211a}(1) = O_P(N^{1/2}T^{-1})$ by Markov inequality. Noting that $E[\hat{B}_{NT,211a}(2)] = 0$, by Davydov inequality we can readily show that

$$\begin{aligned}E\big[\hat{B}_{NT,211a}(2)\big]^2 &= \frac{1}{T^4 N}\sum_{i=1}^{N}\sum_{1\le t<s\le T}\sum_{1\le r<q\le T} E\big[u_{it}X_{is}'\Omega_i^{-1}X_{is}u_{is}u_{ir}X_{iq}'\Omega_i^{-1}X_{iq}u_{iq}\big] \\ &= O\big(T^{-1}\big).\end{aligned}$$

It follows that $\hat{B}_{NT,211a}(2) = O_P(T^{-1/2})$. Similarly, $\hat{B}_{NT,211a}(3) = O_P(T^{-1/2})$. Then $\hat{B}_{NT,211a} = O_P(N^{1/2}T^{-1} + T^{-1/2}) = o_P(1)$. Analogously, we can show that $\hat{B}_{NT,211s} = o_P(1)$ for $s = b, c$. Then we have $\hat{B}_{NT,211} = o_P(1)$ and $\hat{B}_{NT,21} = o_P(1)$.

For $\hat{B}_{NT,22}$, using the same arguments as those used in the proof of Proposition A.2, we can show that under $\mathbb{H}_0(K_0)$,

$$\begin{aligned}\hat{B}_{NT,22} &= \frac{1}{\sqrt{N}}\sum_{k=1}^{K_0}\left(\hat{\alpha}_k-\alpha_k^0\right)'\sum_{i\in\hat{G}_k}X_i'M_0\,\mathrm{diag}(H_i)u_i+o_P(1)\\ &= \frac{1}{\sqrt{N}}\sum_{k=1}^{K_0}\left(\hat{\alpha}_k-\alpha_k^0\right)'\sum_{i\in G_k^0}X_i'M_0\,\mathrm{diag}(H_i)u_i+o_P(1)\equiv\bar{B}_{NT,22}+o_P(1).\end{aligned}$$

Let $B_k=\frac{1}{\sqrt{N}}\sum_{i\in G_k^0}X_i'M_0\,\mathrm{diag}(H_i)u_i$. Then as in the proof of Proposition A.2 and the analysis of $\hat{B}_{NT,211a}(2)$, we can show that

$$\begin{aligned}B_k &= \frac{1}{T\sqrt{N}}\sum_{i\in G_k^0}\sum_{t=1}^{T}\sum_{s=1}^{T}X_{it}\eta_{ts}\iota_sM_0X_i\hat{\Omega}_i^{-1}X_i'M_0\iota_su_{is}\\ &= \frac{1}{T\sqrt{N}}\sum_{i\in G_k^0}\sum_{t=1}^{T}\sum_{s=1}^{T}X_{it}\eta_{ts}\iota_sM_0X_i\Omega_i^{-1}X_i'M_0\iota_su_{is}+o_P(1)\\ &= \frac{1}{T\sqrt{N}}\sum_{i\in G_k^0}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}\sum_{q=1}^{T}X_{it}\eta_{ts}\eta_{sr}X_{ir}'\Omega_i^{-1}X_{iq}\eta_{qs}u_{is}+o_P(1)=O_P\left(N^{1/2}+T^{1/2}\right).\end{aligned}$$

It follows that $\hat{B}_{NT,22}=O_P((NT)^{-1/2}+T^{-1})O_P(N^{1/2}+T^{1/2})=o_P(1)$. This completes the proof of (i1).

Step 2: We prove (ii) $\hat{V}_{NT}(K_0)=V_{NT}+o_P(1)$. Observe that $\hat{V}_{NT}(K_0)-V_{NT}=4V_{NT,1}+4V_{NT,2}$, where

$$\begin{aligned}V_{NT,1} &= T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\left\{\left[\hat{u}_{it}\hat{b}_{it}'\sum_{s=1}^{t-1}\hat{b}_{is}\hat{u}_{is}\right]^2-\left[u_{it}\bar{b}_{it}'\sum_{s=1}^{t-1}\bar{b}_{is}u_{is}\right]^2\right\},\\ V_{NT,2} &= T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\left\{\left[u_{it}\bar{b}_{it}'\sum_{s=1}^{t-1}\bar{b}_{is}u_{is}\right]^2-E\left[u_{it}\bar{b}_{it}'\sum_{s=1}^{t-1}\bar{b}_{is}u_{is}\right]^2\right\}.\end{aligned}$$

Noting that $E(V_{NT,2})=0$ and $\mathrm{Var}(V_{NT,2})=o(1)$ by direct moment calculations, we have $V_{NT,2}=o_P(1)$ by Chebyshev's inequality. Thus we are left to show that $V_{NT,1}=o_P(1)$. Again, using $a^2-b^2=(a-b)^2+2(a-b)b$, we have

$$\begin{aligned}V_{NT,1} &= T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\left[\hat{u}_{it}\hat{b}_{it}'\sum_{s=1}^{t-1}\hat{b}_{is}\hat{u}_{is}-u_{it}\bar{b}_{it}'\sum_{s=1}^{t-1}\bar{b}_{is}u_{is}\right]^2\\ &\quad+2T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\left[\hat{u}_{it}\hat{b}_{it}'\sum_{s=1}^{t-1}\hat{b}_{is}\hat{u}_{is}-u_{it}\bar{b}_{it}'\sum_{s=1}^{t-1}\bar{b}_{is}u_{is}\right]u_{it}\bar{b}_{it}'\sum_{r=1}^{t-1}\bar{b}_{ir}u_{ir}\\ &\equiv V_{NT,11}+2V_{NT,12}.\end{aligned}$$

Let $\bar{V}_{NT,12} \equiv T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}[u_{it}\bar{b}_{it}'\sum_{r=1}^{t-1}\bar{b}_{ir}u_{ir}]^2$. By Cauchy–Schwarz inequality, $V_{NT,12} \leq \{V_{NT,11}\}^{1/2}\{\bar{V}_{NT,12}\}^{1/2}$. It is straightforward to show that $\bar{V}_{NT,12} = O_P(1)$ so that we can prove that $V_{NT,1} = o_P(1)$ by showing that $V_{NT,11} = o_P(1)$. Using $\hat{u}_{it}\hat{b}_{it} = (\hat{u}_{it}\hat{b}_{it} - u_{it}\bar{b}_{it}) + u_{it}\bar{b}_{it}$ and Cauchy–Schwarz inequality,

$$\begin{aligned} V_{NT,11} \leq{}& 3T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\left[(\hat{u}_{it}\hat{b}_{it} - u_{it}\bar{b}_{it})'\sum_{s=1}^{t-1}\bar{b}_{is}u_{is}\right]^2 \\ &+ 3T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\left[u_{it}\bar{b}_{it}'\sum_{s=1}^{t-1}(\hat{b}_{is}\hat{u}_{is} - \bar{b}_{is}u_{is})\right]^2 \\ &+ 3T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\left[(\hat{u}_{it}\hat{b}_{it} - u_{it}\bar{b}_{it})'\sum_{s=1}^{t-1}(\hat{b}_{is}\hat{u}_{is} - \bar{b}_{is}u_{is})\right]^2 \\ \equiv{}& 3V_{NT,111} + 3V_{NT,112} + 3V_{NT,113}. \end{aligned}$$

We complete the proof of (ii) by showing that (ii1) $V_{NT,111} = o_P(1)$, (ii2) $V_{NT,112} = o_P(1)$, and (ii3) $V_{NT,113} = o_P(1)$.

We first show (ii1) $V_{NT,111} = o_P(1)$. Using $\hat{u}_{it}\hat{b}_{it} - u_{it}\bar{b}_{it} = (\hat{u}_{it} - u_{it})\bar{b}_{it} + u_{it}(\hat{b}_{it} - \bar{b}_{it}) + (\hat{u}_{it} - u_{it})(\hat{b}_{it} - \bar{b}_{it})$ and Cauchy–Schwarz inequality, we have

$$\begin{aligned} V_{NT,111} \leq{}& 3T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\left[(\hat{u}_{it} - u_{it})\bar{b}_{it}'\sum_{s=1}^{t-1}\bar{b}_{is}u_{is}\right]^2 \\ &+ 3T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\left[u_{it}(\hat{b}_{it} - \bar{b}_{it})'\sum_{s=1}^{t-1}\bar{b}_{is}u_{is}\right]^2 \\ &+ 3T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\left[(\hat{u}_{it} - u_{it})(\hat{b}_{it} - \bar{b}_{it})'\sum_{s=1}^{t-1}\bar{b}_{is}u_{is}\right]^2 \\ \equiv{}& 3V_{NT,111a} + 3V_{NT,111b} + 3V_{NT,111c}. \end{aligned}$$

By Markov and Davydov inequalities, we can show that $T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}[\bar{b}_{it}'\sum_{s=1}^{t-1}\bar{b}_{is} \times u_{is}]^2 = O_P(1)$. By Boole inequality and Doob inequality (e.g., Hall and Heyde (1980, pp. 14–15)) for m.d.s., and then Davydov inequality, for any $\epsilon > 0$ we have

$$\begin{aligned} P\left(\max_{1\leq i\leq N}\max_{2\leq t\leq T}\left\|T^{-1/2}\sum_{s=1}^{t-1}\bar{b}_{is}u_{is}\right\| > N^{1/8}\epsilon\right) &\leq \sum_{i=1}^{N}P\left(\max_{2\leq t\leq T}\left\|T^{-1/2}\sum_{s=1}^{t-1}\bar{b}_{is}u_{is}\right\| > N^{1/8}\epsilon\right) \\ &\leq \frac{1}{NT^4\epsilon^8}\sum_{i=1}^{N}E\left\|\sum_{s=1}^{T-1}\bar{b}_{is}u_{is}\right\|^8 = O(1). \end{aligned}$$

It follows that

$$\max_{1\leq i\leq N}\max_{2\leq t\leq T}\left\|\sum_{s=1}^{t-1}\bar{b}_{is}u_{is}\right\| = O_P\big(T^{1/2}N^{1/8}\big). \tag{A.8}$$

The same conclusion follows when one replaces $\bar{b}_{is}$ by X_{is} or 1. Let $\bar{B}_i = \text{diag}(\|\bar{b}_{i1}\|^2, \ldots, \|\bar{b}_{iT}\|^2)$. By (A.1), we can readily show that

$$N^{-1}\sum_{i=1}^{N}\|\hat{u}_i - u_i\|^2 \le 2N^{-1}\sum_{i=1}^{N}\|P_0 u_i\|^2 + 2N^{-1}\sum_{i=1}^{N}\big\|M_0 X_i\big(\beta_i^0 - \hat{\beta}_i\big)\big\|^2 \\ = O_P(1) + o_P(1) = O_P(1) \tag{A.9}$$

and

$$\begin{aligned} &N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}(\hat{u}_{it} - u_{it})^2\|\bar{b}_{it}\|^2 \\ &= N^{-1}\sum_{i=1}^{N}(\hat{u}_i - u_i)'\bar{B}_i(\hat{u}_i - u_i) \\ &\le 2N^{-1}\sum_{i=1}^{N}u_i'P_0\bar{B}_iP_0u_i + 2N^{-1}\sum_{i=1}^{N}\big(\beta_i^0 - \hat{\beta}_i\big)'X_i'M_0\bar{B}_iM_0X_i\big(\beta_i^0 - \hat{\beta}_i\big) \\ &= O_P(1) + o_P(1) = O_P(1). \end{aligned} \tag{A.10}$$

By (A.8), (A.10), and Assumption 3,

$$\begin{aligned} V_{NT,111a} &= T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}(\hat{u}_{it} - u_{it})^2\left[\bar{b}_{it}'\sum_{s=1}^{t-1}\bar{b}_{is}u_{is}\right]^2 \\ &\le T^{-2}\max_{1\le i\le N}\max_{2\le t\le T}\left\|\sum_{s=1}^{t-1}\bar{b}_{is}u_{is}\right\|^2\left\{N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}(\hat{u}_{it} - u_{it})^2\|\bar{b}_{it}\|^2\right\} \\ &= T^{-2}O_P\big(TN^{1/4}\big)O_P(1) = o_P(1). \end{aligned}$$

To determine the probability order of $V_{NT,111b}$ and $V_{NT,111c}$, we use the uniform probability order of $\hat{b}_{it} - \bar{b}_{it}$. We decompose $\hat{b}_{it} - \bar{b}_{it}$ as

$$\begin{aligned} \hat{b}_{it} - \bar{b}_{it} &= \hat{\Omega}_i^{-1/2}\left[X_{it} - T^{-1}\sum_{r=1}^{T}X_{ir}\right] - \Omega_i^{-1/2}\left[X_{it} - T^{-1}\sum_{r=1}^{T}E(X_{ir})\right] \\ &= e_iX_{it} - e_iT^{-1}\sum_{r=1}^{T}X_{ir} - \Omega_i^{-1/2}T^{-1}\sum_{r=1}^{T}\big[X_{ir} - E(X_{ir})\big] \\ &\equiv b_{1it} - b_{2it} - b_{3it}, \end{aligned} \tag{A.11}$$

where $e_i \equiv \hat{\Omega}_i^{-1/2} - \Omega_i^{-1/2}$. By Lemma B.3 and the fact that $\max_{1\le i\le N}\max_{1\le t\le T}\|X_{it}\| = o_P((NT)^{1/(8+4\sigma)})$ by Boole and Markov inequalities, we have $\max_{1\le i\le N}\max_{1\le t\le T}\|b_{1it}\| =$

$o_P(a_{NT} \times (NT)^{1/(8+4\sigma)})$. Following the proof of Lemma B.3(v), we can show that

$$\max_{1\le i\le N}\left\|T^{-1}\sum_{r=1}^{T}X_{ir}\right\| \le \max_{1\le i\le N}\left\|T^{-1}\sum_{r=1}^{T}\left[X_{ir}-E(X_{ir})\right]\right\| + \max_{1\le i\le N}\left\|T^{-1}\sum_{r=1}^{T}E(X_{ir})\right\|$$
$$= O_P\big(\max\big\{(NT)^{1/(8+4\sigma)}\log(NT)/T, \big(\log(NT)/T\big)^{1/2}\big\}\big)+O(1)$$
$$= O(1). \tag{A.12}$$

It follows that $\max_{1\le i\le N}\max_{1\le t\le T}\|b_{2it}\| = O_P(a_{NT})$. Also, $\max_{1\le i\le N}\max_{1\le t\le T}\|b_{3it}\| = O_P(a_{NT})$. Thus $\max_{1\le i\le N}\max_{1\le t\le T}\|\hat{b}_{it}-\bar{b}_{it}\| = o_P(a_{NT}(NT)^{1/(8+4\sigma)})$. In addition, using (A.11) and the above bounds, we have

$$T^{-1}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}u_{it}^2\|\hat{b}_{it}-\bar{b}_{it}\|^2$$
$$\le 3T^{-1}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}u_{it}^2\|b_{1it}\|^2 + 3T^{-1}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}u_{it}^2\|b_{2it}\|^2 + 3T^{-1}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}u_{it}^2\|b_{3it}\|^2$$
$$\le 3T^{-1}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}u_{it}^2\|b_{1it}\|^2 + O_P\big(a_{NT}^2\big) + O_P\big(a_{NT}^2\big) = O_P\big(a_{NT}^2\big),$$

where the last equality follows from the fact that $T^{-1}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}u_{it}^2\|b_{1it}\|^2 \le \max_{1\le i\le N}\|e_i\|^2T^{-1}N^{-1}\times\sum_{t=2}^{T}\sum_{i=1}^{N}u_{it}^2\|X_{it}\|^2 = O_P(a_{NT}^2)$. Then by Assumption 3,

$$V_{NT,111b} = T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\left[u_{it}(\hat{b}_{it}-\bar{b}_{it})'\sum_{s=1}^{t-1}\bar{b}_{is}u_{is}\right]^2$$
$$\le T^{-1}\max_{1\le i\le N}\max_{1\le t\le T}\left[\sum_{s=1}^{t-1}\bar{b}_{is}u_{is}\right]^2\left[T^{-1}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}u_{it}^2\|\hat{b}_{it}-\bar{b}_{it}\|^2\right]$$
$$= T^{-1}O_P\big(TN^{1/4}\big)O_P\big(a_{NT}^2\big) = o_P\big(N^{1/4}a_{NT}^2\big) = o_P(1)$$

and

$$V_{NT,111c} = T^{-2}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\left[(\hat{u}_{it}-u_{it})(\hat{b}_{it}-\bar{b}_{it})'\sum_{s=1}^{t-1}\bar{b}_{is}u_{is}\right]^2$$
$$= T^{-2}\max_{1\le i\le N}\max_{1\le t\le T}\|\hat{b}_{it}-\bar{b}_{it}\|^2\max_{1\le i\le N}\max_{1\le t\le T}\left\|\sum_{s=1}^{t-1}\bar{b}_{is}u_{is}\right\|^2 N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}(\hat{u}_{it}-u_{it})^2$$
$$= T^{-2}o_P\big(a_{NT}^2(NT)^{1/(4+2\sigma)}\big)O_P\big(TN^{1/4}\big)O_P(1) = o_P\big(N^{1/4}a_{NT}^2\big) = o_P(1).$$

It follows that $V_{NT,111} = o_P(1)$.

To show (ii2) and (ii3), we find that it is convenient to bound $S_{it} \equiv T^{-1/2} \sum_{s=1}^{t-1} (\hat{b}_{is}\hat{u}_{is} - \bar{b}_{is}u_{is})$. Using $\hat{u}_{it}\hat{b}_{it} - u_{it}\bar{b}_{it} = u_{it}(\hat{b}_{it} - \bar{b}_{it}) + (\hat{u}_{it} - u_{it})\bar{b}_{it} + (\hat{u}_{it} - u_{it})(\hat{b}_{it} - \bar{b}_{it})$, we have

$$\begin{aligned} S_{it} &\equiv T^{-1/2} \sum_{s=1}^{t-1} u_{is}(\hat{b}_{is} - \bar{b}_{is}) + T^{-1/2} \sum_{s=1}^{t-1} (\hat{u}_{is} - u_{is})\bar{b}_{is} + T^{-1/2} \sum_{s=1}^{t-1} (\hat{u}_{is} - u_{is})(\hat{b}_{is} - \bar{b}_{is}) \\ &\equiv S_{1it} + S_{2it} + S_{3it}, \quad \text{say.} \end{aligned}$$

By (A.11), $S_{1it} = T^{-1/2} \sum_{s=1}^{t-1} u_{is}(b_{1is} - b_{2is} - b_{3is}) \equiv S_{1it,1} - S_{1it,2} - S_{1it,3}$, say. By the Remark after (A.8), (A.12), and Lemma B.3,

$$\begin{aligned} \max_{1\le i\le N} \max_{1\le t\le T} \|S_{1it,1}\| &\le \max_{1\le i\le N} \|e_i\| \max_{1\le i\le N} \max_{1\le t\le T} \left\| T^{-1/2} \sum_{s=1}^{t-1} X_{is}u_{is} \right\| \\ &= O_P(a_{NT}) O_P\big(N^{1/4}\big) = o_P(1), \\ \max_{1\le i\le N} \max_{1\le t\le T} \|S_{1it,2}\| &\le \max_{1\le i\le N} \|e_i\| \max_{1\le i\le N} \left\| T^{-1} \sum_{r=1}^{T} X_{ir} \right\| \max_{1\le i\le N} \max_{1\le t\le T} \left\| T^{-1/2} \sum_{s=1}^{t-1} u_{is} \right\| \\ &= O_P(a_{NT}) O_P(1) O_P\big(N^{1/8}\big) = o_P(1) \end{aligned}$$

and

$$\begin{aligned} &\max_{1\le i\le N} \max_{1\le t\le T} \|S_{1it,3}\| \\ &\quad \le \max_{1\le i\le N} \big\|\Omega_i^{-1/2}\big\| \max_{1\le i\le N} \left\| T^{-1} \sum_{r=1}^{T} \big[X_{ir} - E(X_{ir})\big] \right\| \max_{1\le i\le N} \max_{1\le t\le T} \left\| T^{-1/2} \sum_{s=1}^{t-1} u_{is} \right\| \\ &\quad = O(1) O_P(a_{NT}) O_P\big(N^{1/8}\big) = o_P(1). \end{aligned}$$

It follows that $\max_{1\le i\le N} \max_{1\le t\le T} \|S_{1it}\| = o_P(1)$. Similarly we can show $\max_{1\le i\le N} \max_{1\le t\le T} \|S_{2it}\| = o_P(1)$ and $\max_{1\le i\le N} \max_{1\le t\le T} \|S_{3it}\| = o_P(1)$. Hence $\max_{1\le i\le N} \max_{1\le t\le T} \|S_{it}\| = o_P(1)$. It follows that

$$\begin{aligned} V_{NT,112} &= T^{-2} N^{-1} \sum_{t=2}^{T} \sum_{i=1}^{N} \left[u_{it}\bar{b}_{it}' \sum_{s=1}^{t-1} (\hat{b}_{is}\hat{u}_{is} - \bar{b}_{is}u_{is}) \right]^2 \\ &\le \left\{ \max_{1\le i\le N} \max_{1\le t\le T} \|S_{it}\|^2 \right\} \left\{ T^{-1} N^{-1} \sum_{t=2}^{T} \sum_{i=1}^{N} \big\| u_{it}\bar{b}_{it}' \big\|^2 \right\} = o_P(1) O_P(1) = o_P(1) \end{aligned}$$

and

$$\begin{aligned} V_{NT,113} &= T^{-2} N^{-1} \sum_{t=2}^{T} \sum_{i=1}^{N} \left[(\hat{u}_{it}\hat{b}_{it} - u_{it}\bar{b}_{it})' \sum_{s=1}^{t-1} (\hat{b}_{is}\hat{u}_{is} - \bar{b}_{is}u_{is}) \right]^2 \\ &\le \left\{ \max_{1\le i\le N} \max_{1\le t\le T} \|S_{it}\|^2 \right\} \left\{ T^{-1} N^{-1} \sum_{t=2}^{T} \sum_{i=1}^{N} \|\hat{u}_{it}\hat{b}_{it} - u_{it}\bar{b}_{it}\|^2 \right\} = o_P(1) o_P(1) = o_P(1), \end{aligned}$$

as one can readily show that $T^{-1}N^{-1}\sum_{t=2}^{T}\sum_{i=1}^{N}\|\hat{u}_{it}\hat{b}_{it}-u_{it}\bar{b}_{it}\|^2=o_P(1)$. Thus $V_{NT,11}=o_P(1)$. This completes the proof of (ii). □

Proof of Theorem 3.3. Observe that $\sqrt{\hat{V}_{NT}(K_0)}\hat{J}_{1NT}(K_0)=A_{1NT}-\hat{B}_{NT}(K_0)+A_{2NT}+2A_{3NT}$, where A_{1NT}, A_{2NT}, and A_{3NT} are as defined in (A.2). We study the probability order of each term in the last expression.

Noting that $\|X_i'M_0u_i\|^2\leq 2\|X_i'u_i\|^2+2\|X_i'P_0u_i\|^2$, we have $N^{-1}T^{-2}\sum_{i=1}^{N}\|X_i'M_0u_i\|^2\leq 2a_1+2a_2$, where $a_1=2N^{-1}T^{-2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}u_{it}X_{it}'X_{is}u_{is}$ and $a_2=N^{-1}T^{-4}\times\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}\sum_{q=1}^{T}u_{it}X_{is}'X_{ir}u_{iq}$. By Assumption 1 and Markov inequality, we can readily show that $a_1=O_P(T^{-1})$ and $a_2=O_P(T^{-1})$. It follows that $N^{-1}T^{-2}\sum_{i=1}^{N}\|X_i'\times M_0u_i\|^2=O_P(T^{-1})$. Then by Lemma B.3(v),

$$\begin{aligned}N^{-1/2}T^{-1}A_{1NT}&=N^{-1}T^{-1}\sum_{i=1}^{N}u_i'M_0\bar{P}_{X_i}M_0u_i\\&\leq\max_{1\leq i\leq N}\lambda_{\max}(\hat{Q}_i)N^{-1}T^{-2}\sum_{i=1}^{N}u_i'M_0X_iX_i'M_0u_i\\&=O_P(1)O_P(T^{-1})=O_P(T^{-1}).\end{aligned}$$

By (A.1) and Cauchy–Schwarz inequality,

$$\begin{aligned}N^{-1/2}T^{-1}\hat{B}_{NT}&=N^{-1}T^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{u}_{it}^2h_{i,tt}=N^{-1}T^{-1}\sum_{i=1}^{N}\hat{u}_i'\operatorname{diag}(H_i)\hat{u}_i\\&\leq 2N^{-1}T^{-1}\sum_{i=1}^{N}u_i'M_0\operatorname{diag}(H_i)M_0u_i,\\&\quad+2N^{-1}T^{-1}\sum_{i=1}^{N}\big(\beta_i^0-\hat{\beta}_i\big)'X_i'M_0\operatorname{diag}(H_i)M_0X_i\big(\beta_i^0-\hat{\beta}_i\big)\\&\equiv 2b_1+2b_2,\quad\text{say.}\end{aligned}$$

By Cauchy–Schwarz inequality, $b_1\leq 2N^{-1}T^{-1}\sum_{i=1}^{N}u_i'\operatorname{diag}(H_i)u_i+2N^{-1}T^{-1}\sum_{i=1}^{N}u_i'P_0\times\operatorname{diag}(H_i)P_0u_i\equiv 2b_{1,1}+2b_{1,2}$, say. By the fact $H_i=M_0\bar{P}_{X_i}M_0\leq[\lambda_{\min}(\hat{Q}_i)]^{-1}T^{-1}M_0X_iX_i'M_0$ and Lemma B.3(v), we have

$$\begin{aligned}b_{1,1}&\leq\big[\lambda_{\min}(\hat{Q}_i)\big]^{-1}N^{-1}T^{-2}\sum_{i=1}^{N}u_i'\operatorname{diag}\big(M_0X_iX_i'M_0\big)u_i\\&=\big[\lambda_{\min}(\hat{Q}_i)\big]^{-1}N^{-1}T^{-2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}u_{it}\eta_{ts}X_{is}'X_{ir}\eta_{tr}u_{ir}=O_P\big(T^{-1}\big),\end{aligned}$$

as we can readily show that $N^{-1}T^{-2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}u_{it}\eta_{ts}X_{is}'X_{ir}\eta_{tr}u_{ir}=O_P(T^{-1})$ based on the fact that $\eta_{ts}=\mathbf{1}\{t=s\}-T^{-1}$ and Markov inequality. As in the analysis of

$\hat{B}_{NT,11}$, we can readily apply the fact that $\mathbf{i}_T' \operatorname{diag}(H_i)\mathbf{i}_T = p$ to obtain

$$\begin{aligned}
b_{1,2} &= N^{-1}T^{-3}\sum_{i=1}^{N} u_i'\mathbf{i}_T\mathbf{i}_T' \operatorname{diag}(H_i)\mathbf{i}_T\mathbf{i}_T' u_i = pN^{-1}T^{-3}\sum_{i=1}^{N} u_i'\mathbf{i}_T\mathbf{i}_T' u_i \\
&= pN^{-1}T^{-3}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T} u_{it}u_{is} = O_P\big(T^{-2}\big).
\end{aligned}$$

It follows that $b_1 = O_P(T^{-1})$. By Cauchy–Schwarz inequality, $b_2 \le 2N^{-1}T^{-1}\sum_{i=1}^{N}(\beta_i^0 - \hat{\beta}_i)' X_i' \operatorname{diag}(H_i) \times X_i(\beta_i^0 - \hat{\beta}_i) + 2N^{-1}T^{-1}\sum_{i=1}^{N}(\beta_i^0 - \hat{\beta}_i)' X_i' P_0 \operatorname{diag}(H_i) P_0 X_i(\beta_i^0 - \hat{\beta}_i) \equiv 2b_{2,1} + 2b_{2,2}$, say. Noting that $\operatorname{diag}(M_0 X_i X_i' M_0) \le \operatorname{diag}(X_i X_i')$, we have

$$\begin{aligned}
b_{2,1} &\le \big[\lambda_{\min}(\hat{Q}_i)\big]^{-1}N^{-1}T^{-2}\sum_{i=1}^{N}\big(\beta_i^0 - \hat{\beta}_i\big)' X_i' \operatorname{diag}\big(M_0 X_i X_i' M_0\big) X_i\big(\beta_i^0 - \hat{\beta}_i\big) \\
&\le \big[\lambda_{\min}(\hat{Q}_i)\big]^{-1}N^{-1}T^{-2}\sum_{i=1}^{N}\big(\beta_i^0 - \hat{\beta}_i\big)' X_i' \operatorname{diag}\big(X_i X_i'\big) X_i\big(\beta_i^0 - \hat{\beta}_i\big) \\
&= \big[\lambda_{\min}(\hat{Q}_i)\big]^{-1}N^{-1}T^{-2}\sum_{i=1}^{N}\big(\beta_i^0 - \hat{\beta}_i\big)' X_{it} X_{it}' X_{it} X_{it}'\big(\beta_i^0 - \hat{\beta}_i\big) \\
&\le T^{-1}\big[\lambda_{\min}(\hat{Q}_i)\big]^{-1}\max_{1\le i\le N} T^{-1}\sum_{t=1}^{T}\|X_{it}\|^4 N^{-1}\sum_{i=1}^{N}\big\|\beta_i^0 - \hat{\beta}_i\big\|^2 \\
&= T^{-1}O_P(1)O_P(1)O_P(1) = O_P\big(T^{-1}\big).
\end{aligned}$$

Using $\mathbf{i}_T' \operatorname{diag}(H_i)\mathbf{i}_T = p$, we have

$$\begin{aligned}
b_{2,2} &= N^{-1}T^{-3}\sum_{i=1}^{N}\big(\beta_i^0 - \hat{\beta}_i\big)' X_i'\mathbf{i}_T\mathbf{i}_T' \operatorname{diag}(H_i)\mathbf{i}_T\mathbf{i}_T' X_i\big(\beta_i^0 - \hat{\beta}_i\big) \\
&= pN^{-1}T^{-3}\sum_{i=1}^{N}\big(\beta_i^0 - \hat{\beta}_i\big) X_i'\mathbf{i}_T\mathbf{i}_T' X_i\big(\beta_i^0 - \hat{\beta}_i\big) \\
&\le T^{-1}\max_{1\le i\le N}\|\bar{X}_{i\cdot}\|^2 N^{-1}\sum_{i=1}^{N}\big\|\beta_i^0 - \hat{\beta}_i\big\|^2 = O_P\big(T^{-1}\big).
\end{aligned}$$

It follows that $b_2 = O_P(T^{-1})$ and $N^{-1/2}T^{-1}\hat{B}_{NT}(K_0) = O_P(T^{-1})$.

By Assumption 4(ii) and Lemma B.3(v), w.p.a.1,

$$\begin{aligned}
N^{-1/2}T^{-1}A_{2NT} &= N^{-1}T^{-1}\sum_{i=1}^{N}\big(\beta_i^0 - \hat{\beta}_i\big)' X_i' M_0 X_i\big(\beta_i^0 - \hat{\beta}_i\big) \\
&\ge \lambda_{\min}(\hat{Q}_i)N^{-1}\sum_{k=1}^{K_0}\sum_{i\in\hat{G}_k}\big\|\beta_i^0 - \hat{\alpha}_k\big\|^2 \ge \frac{1}{2}\lambda_{\min}(Q_i)\underline{c}_{K_0}.
\end{aligned}$$

Now we decompose $N^{-1/2}T^{-1}A_{3NT}$ as $N^{-1/2}T^{-1}A_{3NT} = N^{-1}T^{-1}\sum_{i=1}^{N} u_i' X_i(\beta_i^0 - \hat{\beta}_i) - N^{-1}T^{-2} \times \sum_{i=1}^{N} u_i' \mathbf{i}_T \mathbf{i}_T' X_i(\beta_i^0 - \hat{\beta}_i) \equiv A_{3NT,1} + A_{3NT,2}$, say. For the first term, we apply Cauchy–Schwarz inequality to obtain

$$\begin{aligned}|A_{3NT,1}| &\leq T^{-1/2}\left\{N^{-1}T^{-1}\sum_{i=1}^{N}\left\|u_i' X_i\right\|^2\right\}^{1/2} \times \left\{N^{-1}\sum_{k=1}^{K_0}\sum_{i\in\hat{G}_k}\left\|\beta_i^0 - \hat{\alpha}_k\right\|^2\right\}^{1/2}\\ &= T^{-1/2}O_P(1)O_P(1) = o_P(1),\end{aligned}$$

where we use the fact that $N^{-1}T^{-1}\sum_{i=1}^{N} E\|u_i' X_i\|^2 = O(1)$ under Assumption 1. Similarly,

$$\begin{aligned}|A_{3NT,2}| &\leq \max_{1\leq i\leq N}\left\|T^{-1}u_i'\mathbf{i}_T\right\| \max_{1\leq i\leq N}\left\|T^{-1}X_i'\mathbf{i}_T\right\| N^{-1}\sum_{i=1}^{N}\left\|\beta_i^0 - \hat{\beta}_i\right\|\\ &\leq \max_{1\leq i\leq N}\left\|T^{-1}u_i'\mathbf{i}_T\right\| \max_{1\leq i\leq N}\left\|T^{-1}X_i'\mathbf{i}_T\right\|\left\{N^{-1}\sum_{k=1}^{K_0}\sum_{i\in\hat{G}_k}\left\|\beta_i^0 - \hat{\alpha}_k\right\|^2\right\}^{1/2}\\ &= O_P(\alpha_{NT})O_P(1)O_P(1) = o_P(1).\end{aligned}$$

It follows that $N^{-1/2}T^{-1}A_{3NT} = o_P(1)$.

In sum, we have $N^{-1/2}T^{-1}\sqrt{\hat{V}_{NT}(K_0)}\hat{J}_{1NT}(K_0) \geq \frac{1}{2}\lambda_{\min}(Q_i)\underline{c}_{K_0} + o_P(1)$ w.p.a.1. In addition, we can show that $\hat{V}_{NT}(K_0)$ has a positive probability limit under $\mathbb{H}_1(K_0)$. It follows that under $\mathbb{H}_1(K_0)$, $P(\hat{J}_{1NT}(K_0) \geq c_{NT}) \to 1$ as $(N,T) \to \infty$ for any $c_{NT} = o(N^{1/2}T)$. □

Proof of Theorem 4.1. (i) Let $\ddot{y}_i = (\ddot{y}_{i1}, \ldots, \ddot{y}_{iT})'$, $\ddot{u}_i = (\ddot{u}_{i1}, \ldots, \ddot{u}_{iT})'$, and $\tilde{X}_i = (\tilde{X}_{i1}, \ldots, \tilde{X}_{iT})'$. The minimization problem in (4.4) can be rewritten as

$$\min_{\boldsymbol{\beta},\boldsymbol{\alpha}} Q_{2NT,\lambda}^{(K_0)}(\boldsymbol{\beta},\boldsymbol{\alpha}) = \frac{1}{NT}\sum_{i=1}^{N}\left\|\ddot{y}_i - \tilde{X}_i\beta_i + \frac{1}{N}\sum_{j=1}^{N}\tilde{X}_j\beta_j\right\|^2 + \frac{\lambda}{N}\sum_{i=1}^{N}\prod_{k=1}^{K_0}\|\beta_i - \alpha_k\|. \quad \text{(A.13)}$$

Let $\boldsymbol{\beta} = \boldsymbol{\beta}^0 + T^{-1/2}\mathbf{v}$, where $\mathbf{v} = (v_1, \ldots, v_N)$ is a $p \times N$ matrix. We want to show that for any given $\epsilon^* > 0$, there exists a large constant $L = L(\epsilon^*)$ such that for sufficiently large N and T, we have

$$P\left\{\inf_{N^{-1}\sum_{i=1}^{N}\|v_i\|^2 = L} Q_{2NT,\lambda}^{(K_0)}\left(\boldsymbol{\beta}^0 + T^{-1/2}\mathbf{v}, \tilde{\boldsymbol{\alpha}}\right) > Q_{2NT,\lambda}^{(K_0)}\left(\boldsymbol{\beta}^0, \boldsymbol{\alpha}^0\right)\right\} \geq 1 - \epsilon^*, \quad \text{(A.14)}$$

where $\tilde{\boldsymbol{\alpha}} \equiv \tilde{\boldsymbol{\alpha}}(\mathbf{v})$ is chosen such that $(\boldsymbol{\beta}^0 + T^{-1/2}\mathbf{v}, \tilde{\boldsymbol{\alpha}})$ minimizes $Q_{2NT,\lambda}^{(K_0)}(\boldsymbol{\beta},\boldsymbol{\alpha})$ for some given $\mathbf{v}$. This implies that w.p.a.1 there is a local minimum $\{\tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\alpha}}\}$ such that N^{-1}, where $\tilde{b}_i = \tilde{\beta}_i - \beta_i^0$.

Observe that

$$
\begin{aligned}
&T\big[Q^{(K_0)}_{2NT,\lambda}(\boldsymbol{\beta}^0+T^{-1/2}\mathbf{v},\tilde{\boldsymbol{\alpha}})-Q^{(K_0)}_{2NT,\lambda}(\boldsymbol{\beta}^0,\boldsymbol{\alpha}^0)\big]\\
&\geq \frac{1}{N}\sum_{i=1}^{N}\left\{\left\|\ddot{y}_i-\tilde{X}_i\big(\beta_i^0+T^{-1/2}v_i\big)+\frac{1}{N}\sum_{j=1}^{N}\tilde{X}_j\big(\beta_j^0+T^{-1/2}v_j\big)\right\|^2-\|\ddot{u}_i\|^2\right\}\\
&=\frac{1}{N}\sum_{i=1}^{N}\left\{\left\|\ddot{u}_i-T^{-1/2}\left(\tilde{X}_iv_i-\frac{1}{N}\sum_{j=1}^{N}\tilde{X}_jv_j\right)\right\|^2-\|\ddot{u}_i\|^2\right\}\\
&=\frac{1}{NT}\sum_{i=1}^{N}\left\|\tilde{X}_iv_i-\frac{1}{N}\sum_{j=1}^{N}\tilde{X}_jv_j\right\|^2-\frac{2}{NT^{1/2}}\sum_{i=1}^{N}\ddot{u}_i\left(\tilde{X}_iv_i-\frac{1}{N}\sum_{j=1}^{N}\tilde{X}_jv_j\right)\\
&=\frac{1}{NT}\sum_{i=1}^{N}v_i'\tilde{X}_i'\tilde{X}_iv_i-\frac{1}{N^2T}\sum_{i=1}^{N}\sum_{j=1}^{N}v_i'\tilde{X}_i'\tilde{X}_jv_j-\frac{2}{NT^{1/2}}\sum_{i=1}^{N}\ddot{u}_i'\tilde{X}_iv_i\\
&=\frac{1}{N}v'\hat{\boldsymbol{\Omega}}v-\frac{1}{N}v'\mathbf{A}v-\frac{2}{NT^{1/2}}\sum_{i=1}^{N}\ddot{u}_i'\tilde{X}_iv_i,
\end{aligned}
\tag{A.15}
$$

where $\hat{\boldsymbol{\Omega}}\equiv\operatorname{diag}(\hat{\Omega}_1,\dots,\hat{\Omega}_N)$, $\mathbf{A}$ is an $Np\times Np$ matrix with a typical $p\times p$ block submatrix $(NT)^{-1}\tilde{X}_i'\tilde{X}_j$, and we use the fact that $\sum_{i=1}^{N}\ddot{u}_i=0$. By Lemma B.3(v) and Assumption 1(ii),

$$
\frac{1}{N}v'\hat{\boldsymbol{\Omega}}v\geq\frac{1}{N}\|v\|^2\min_{1\leq i\leq N}\lambda_{\min}(\hat{\Omega}_i)\geq\frac{1}{N}\|v\|^2\underline{c}_{XX}/2\quad\text{w.p.a.1.}
\tag{A.16}
$$

Define the upper block-triangular matrix

$$
A_1=\frac{1}{NT}\begin{pmatrix}\tilde{X}_1'\tilde{X}_1 & \tilde{X}_1'\tilde{X}_2 & \cdots & \tilde{X}_1'\tilde{X}_N\\ 0 & \tilde{X}_2'\tilde{X}_2 & \cdots & \tilde{X}_2'\tilde{X}_N\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \tilde{X}_N'\tilde{X}_N\end{pmatrix}.
$$

Note that $\mathbf{A}=A_1+A_1'-A_d$, where $A_d=N^{-1}\hat{\boldsymbol{\Omega}}$. By the fact that the eigenvalues of a block upper/lower triangular matrix are the combined eigenvalues of its diagonal block matrices, Weyl inequality, Lemma B.3(v), and Assumption 1(ii), we have $\lambda_{\max}(\mathbf{A})\leq 2\lambda_{\max}(A_1)-\lambda_{\min}(A_d)\leq 2N^{-1}\max_{1\leq i\leq N}\lambda_{\max}(\hat{\Omega}_i)=O_P(N^{-1})$. It follows that

$$
\frac{1}{N}v'\mathbf{A}v\leq\frac{1}{N}\|v\|^2\lambda_{\max}(\mathbf{A})=\frac{1}{N}\|v\|^2O_P\big(N^{-1}\big).
\tag{A.17}
$$

In addition, by Cauchy–Schwarz and Markov inequalities, we can readily show that

$$
\frac{1}{NT^{1/2}}\left|\sum_{i=1}^{N}\ddot{u}_i'\tilde{X}_iv_i\right|\leq\left\{\frac{1}{N}\sum_{i=1}^{N}v_i'v_i\right\}^{1/2}\left\{\frac{1}{NT}\sum_{i=1}^{N}\ddot{u}_i'\tilde{X}_i\tilde{X}_i'\ddot{u}_i\right\}^{1/2}=N^{-1/2}\|v\|O_P(1).
\tag{A.18}
$$

Combining (A.15)–(A.18) yields

$$T\big[Q_{2NT,\lambda}^{(K_0)}\big(\boldsymbol{\beta}^0+T^{-1/2}\mathbf{v},\tilde{\boldsymbol{\alpha}}\big)-Q_{2NT,\lambda}^{(K_0)}\big(\boldsymbol{\beta}^0,\boldsymbol{\alpha}^0\big)\big]\geq\frac{1}{N}\|v\|^2\big[\underline{c}_{XX}/2-o_P(1)\big]-N^{-1/2}\|v\|O_P(1).$$

The first term dominates the second term in the last display for sufficiently large L. That is, $T[Q_{2NT,\lambda}^{(K_0)}(\boldsymbol{\beta}^0+T^{-1/2}\mathbf{v},\tilde{\boldsymbol{\alpha}})-Q_{2NT,\lambda}^{(K_0)}(\boldsymbol{\beta}^0,\boldsymbol{\alpha}^0)]>0$ for sufficiently large L. Consequently, the minimizer $\tilde{\boldsymbol{\beta}}$ must satisfy $N^{-1}\sum_{i=1}^N\|\tilde{b}_i\|^2=O_P(T^{-1})$.

(ii) Given (i), we can readily argue the pointwise convergence of $\tilde{\beta}_i$ as in the proof of (iii) below. Let $P_{NT}(\boldsymbol{\beta},\boldsymbol{\alpha})=\frac{1}{N}\sum_{i=1}^N\prod_{k=1}^{K_0}\|\beta_i-\alpha_k\|$ and $\tilde{c}_{iNT}(\boldsymbol{\alpha})=\prod_{k=1}^{K_0-1}\|\tilde{\beta}_i-\alpha_k\|+\prod_{k=1}^{K_0-2}\|\tilde{\beta}_i-\alpha_k\|\times\|\beta_i^0-\alpha_{K_0}\|+\cdots+\prod_{k=2}^{K_0}\|\beta_i^0-\alpha_k\|$. By the repeated use of Minkowski inequality (see, e.g., Su, Shi, and Phillips (2016, SSP hereafter)), we have that as $(N,T)\to\infty$,

$$\left|\prod_{k=1}^{K_0}\|\tilde{\beta}_i-\alpha_k\|-\prod_{k=1}^{K_0}\|\tilde{\beta}_i^0-\alpha_k\|\right|\leq\tilde{c}_{iNT}(\boldsymbol{\alpha})\|\beta_i-\beta_i^0\|\quad\text{and}$$
$$\tilde{c}_{iNT}(\boldsymbol{\alpha})\leq C_{K_0NT}(\boldsymbol{\alpha})\big(1+2\|\tilde{\beta}_i-\beta_i^0\|\big),$$

where $C_{K_0NT}(\boldsymbol{\alpha})=\max_{1\leq i\leq N}\max_{1\leq s\leq k\leq K_0-1}\prod_{k=1}^{s}a_{ks}\|\beta_i^0-\alpha_k\|^{K_0-1-s}=\max_{1\leq l\leq K_0}\max_{1\leq s\leq k\leq K_0-1}\prod_{k=1}^{s}a_{ks}\|\alpha_l^0-\alpha_k\|^{K_0-1-s}=O(1)$ and a_{ks}s are finite integers. It follows that as $(N,T)\to\infty$,

$$\begin{aligned}\big|P_{NT}(\tilde{\boldsymbol{\beta}},\boldsymbol{\alpha})-P_{NT}\big(\boldsymbol{\beta}^0,\boldsymbol{\alpha}\big)\big|&\leq C_{K_0NT}(\boldsymbol{\alpha})\frac{1}{N}\sum_{i=1}^N\|\tilde{b}_i\|+2C_{K_0NT}(\boldsymbol{\alpha})\frac{1}{N}\sum_{i=1}^N\|\tilde{b}_i\|^2\\&\leq C_{K_0NT}(\boldsymbol{\alpha})\left\{\frac{1}{N}\sum_{i=1}^N\|\tilde{b}_i\|^2\right\}^{1/2}+O_P\big(T^{-1}\big)=O_P\big(T^{-1/2}\big).\end{aligned}\tag{A.19}$$

By (A.19) and the fact that $P_{NT}(\boldsymbol{\beta}^0,\boldsymbol{\alpha}^0)=0$, we have

$$\begin{aligned}0\geq P_{NT}(\tilde{\boldsymbol{\beta}},\tilde{\boldsymbol{\alpha}})-P_{NT}\big(\tilde{\boldsymbol{\beta}},\boldsymbol{\alpha}^0\big)&=P_{NT}\big(\boldsymbol{\beta}^0,\tilde{\boldsymbol{\alpha}}\big)-P_{NT}\big(\boldsymbol{\beta}^0,\boldsymbol{\alpha}^0\big)+O_P\big(T^{-1/2}\big)\\&=\frac{1}{N}\sum_{i=1}^N\prod_{k=1}^{K_0}\|\beta_i^0-\tilde{\alpha}_k\|+O_P\big(T^{-1/2}\big)\\&=\frac{N_1}{N}\prod_{k=1}^{K_0}\|\tilde{\alpha}_k-\alpha_1^0\|+\cdots+\frac{N_{K_0}}{N}\prod_{k=1}^{K_0}\|\tilde{\alpha}_k-\alpha_{K_0}^0\|+O_P\big(T^{-1/2}\big).\end{aligned}\tag{A.20}$$

Then by Assumption 2(i), we have $\prod_{k=1}^{K_0}\|\tilde{\alpha}_k-\alpha_l^0\|=O_P(T^{-1/2})$ for $l=1,\ldots,K_0$. It follows that $(\tilde{\alpha}_{(1)},\ldots,\tilde{\alpha}_{(K_0)})-(\alpha_1^0,\ldots,\alpha_{K_0}^0)=O_P(T^{-1/2})$ for some suitable permutation $(\tilde{\alpha}_{(1)},\ldots,\tilde{\alpha}_{(K_0)})$ of $(\tilde{\alpha}_1,\ldots,\tilde{\alpha}_{K_0})$.

(iii) We invoke subdifferential calculus (e.g., Bertsekas (1995, Appendix B.5)). A necessary condition for $\{\tilde{\beta}_i\}$, and $\{\tilde{\alpha}_k\}$ to minimize the objective function in (A.13) is that for each $i=1,\ldots,N$ (resp. $k=1,\ldots,K_0$), $\mathbf{0}_{p\times 1}$ belongs to the subdifferential of

$Q_{2NT,\lambda}^{(K_0)}(\boldsymbol{\beta},\boldsymbol{\alpha})$ with respect to β_i (resp. α_k) evaluated at $\{\tilde{\beta}_i\}$ and $\{\tilde{\alpha}_k\}$. That is, for each $i=1,\dots,N$ and $k=1,\dots,K_0$, we have

$$\mathbf{0}_{p\times 1}=-\frac{2}{T}\frac{N-1}{N}\tilde{X}_i'\left(\ddot{y}_i-\tilde{X}_i\tilde{\beta}_i+\frac{1}{N}\sum_{j=1}^N\tilde{X}_j\tilde{\beta}_j\right)+\lambda\sum_{j=1}^{K_0}\tilde{e}_{ij}\prod_{l=1,l\neq j}^{K_0}\|\tilde{\beta}_i-\tilde{\alpha}_l\|, \tag{A.21}$$

where $\tilde{e}_{ij}=\frac{\tilde{\beta}_i-\tilde{\alpha}_j}{\|\tilde{\beta}_i-\tilde{\alpha}_j\|}$ if $\|\tilde{\beta}_i-\tilde{\alpha}_j\|\neq 0$ and $\|\tilde{e}_{ij}\|\leq 1$ if $\|\tilde{\beta}_i-\tilde{\alpha}_j\|=0$. Noting that $\ddot{y}_i=\tilde{X}_i\beta_i^0-\frac{1}{N}\sum_{j=1}^N\tilde{X}_j\beta_j^0+\ddot{u}_i$, (A.21) implies that

$$\hat{\Omega}_i\tilde{b}_i=\frac{1}{T}\tilde{X}_i'\ddot{u}_i+\frac{1}{NT}\sum_{j=1}^N\tilde{X}_i'\tilde{X}_j\tilde{b}_j-\frac{\lambda N}{2(N-1)}\sum_{j=1}^{K_0}\tilde{e}_{ij}\prod_{l=1,l\neq j}^{K_0}\|\tilde{\beta}_i-\tilde{\alpha}_l\|, \tag{A.22}$$

where $\tilde{b}_i=\tilde{\beta}_i-\beta_i^0$. $\hat{\Omega}_i$ is asymptotically nonsingular uniformly in i by Lemma B.3(v) and Assumption 1(ii). Using the arguments as used in the proof of Lemma B.3, we can readily show that

$$\begin{aligned}\max_{1\leq i\leq N}\left\|\frac{1}{T}\tilde{X}_i'\ddot{u}_i\right\| &= \max_{1\leq i\leq N}\left\|\frac{1}{T}\sum_{t=1}^T(X_{it}-\bar{X}_{i\cdot})(u_{it}-\bar{u}_{\cdot t})\right\| \\ &\leq \max_{1\leq i\leq N}\left\|\frac{1}{T}\sum_{t=1}^T X_{it}u_{it}\right\|+\max_{1\leq i\leq N}\left\|\frac{1}{T}\sum_{t=1}^T X_{it}\bar{u}_{\cdot t}\right\| \\ &\quad+\max_{1\leq i\leq N}\|\bar{X}_{i\cdot}\|\max_{1\leq i\leq N}|\bar{u}_{i\cdot}|+\max_{1\leq i\leq N}\|\bar{X}_{i\cdot}\|\bar{u} \\ &= O_P(a_{NT}),\end{aligned}$$

where $a_{NT}=\max\{(NT)^{1/(4+2\sigma)}\log(NT)/T,(\log(NT)/T)^{1/2}\}$. By Assumption 1 and Lemma B.3,

$$\begin{aligned}\frac{1}{NT}\left\|\sum_{j=1}^N\tilde{X}_i'\tilde{X}_j\tilde{b}_j\right\| &\leq \max_{1\leq i\leq N}\frac{1}{T^{1/2}}\|\tilde{X}_i\|\left\{\frac{1}{NT}\sum_{j=1}^N\|\tilde{X}_j\|^2\right\}^{1/2}\left\{\frac{1}{N}\sum_{j=1}^N\|\tilde{b}_j\|^2\right\}^{1/2} \\ &\equiv c_{1NT}=O_P\big(T^{-1/2}\big).\end{aligned}$$

Let $R_i=\frac{\lambda N}{2(N-1)}\sum_{j=1}^{K_0}\tilde{e}_{ij}\prod_{l=1,l\neq j}^{K_0}\|\tilde{\beta}_i-\tilde{\alpha}_l\|$. By the fact that $|\prod_{l=1,l\neq j}^{K_0}\|\tilde{\beta}_i-\tilde{\alpha}_l\|-\prod_{l=1,l\neq j}^{K_0}\|\beta_i^0-\tilde{\alpha}_l\||\leq C_{K_0}(\tilde{\boldsymbol{\alpha}})(1+2\|\tilde{\beta}_i-\beta_i^0\|)$ for some $C_{K_0}(\tilde{\boldsymbol{\alpha}})=O_P(1)$, we have

$$\begin{aligned}\|R_i\| &\leq \lambda\sum_{j=1}^{K_0}\prod_{l=1,l\neq j}^{K_0}\|\tilde{\beta}_i-\tilde{\alpha}_l\| \\ &\leq \lambda\sum_{j=1}^{K_0}\left(\prod_{l=1,l\neq j}^{K_0}\|\tilde{\beta}_i-\tilde{\alpha}_l\|-\prod_{l=1,l\neq j}^{K_0}\left\|\beta_i^0-\tilde{\alpha}_l\right\|\right)+\lambda\sum_{j=1}^{K_0}\prod_{l=1,l\neq j}^{K_0}\left\|\beta_i^0-\tilde{\alpha}_l\right\| \\ &\leq 2\lambda K_0C_{K_0}(\tilde{\boldsymbol{\alpha}})\|\tilde{b}_i\|+r_{NT},\end{aligned}$$

where $r_{NT} = \lambda K_0 C_{K_0}(\tilde{\boldsymbol{\alpha}}) + \lambda \max_{1\leq i\leq N} \sum_{j=1}^{K_0} \prod_{l=1,l\neq j}^{K_0} \|\beta_i^0 - \tilde{\alpha}_l\| = O(\lambda)$ as β_i^0s can only take K_0 values. It follows that

$$\begin{aligned}\|\tilde{b}_i\| &\leq \|\hat{\Omega}_i^{-1}\|_{\text{sp}}\left(\left\|\frac{1}{T}\tilde{X}_i'\ddot{u}_i\right\| + \left\|\frac{1}{NT}\sum_{j=1}^N \tilde{X}_i'\tilde{X}_j\tilde{b}_j\right\| + 2\lambda K_0 C_{K_0}(\tilde{\boldsymbol{\alpha}})\|\tilde{b}_i\| + r_{NT}\right)\\ &\leq \|\hat{\Omega}_i^{-1}\|_{\text{sp}}\left(\left\|\frac{1}{T}\tilde{X}_i'\ddot{u}_i\right\| + c_{1NT} + 2\lambda K_0 C_{K_0}(\tilde{\boldsymbol{\alpha}})\|\tilde{b}_i\| + r_{NT}\right)\end{aligned}$$

and

$$\begin{aligned}\max_{1\leq i\leq N}\|\tilde{b}_i\| &\leq \frac{c_{2NT}}{1-2c_{2NT}\lambda K_0 C_{K_0}(\tilde{\boldsymbol{\alpha}})}\left(\max_{1\leq i\leq N}\left\|\frac{1}{T}\tilde{X}_i'\ddot{u}_i\right\| + c_{1NT} + r_{NT}\right)\\ &= O_P(c_{2NT})\left\{\max_{1\leq i\leq N}\left\|\frac{1}{T}\tilde{X}_i'\ddot{u}_i\right\| + c_{1NT} + r_{NT}\right\} = O_P(a_{NT}+\lambda),\end{aligned}$$

where $\|\cdot\|_{\text{sp}}$ denotes the spectral norm and $c_{2NT} = [\min_{1\leq i\leq N}\lambda_{\min}(\hat{\Omega}_i)]^{-1} = O_P(1)$ by Lemma B.3(v). □

Proof of Theorem 4.2. (i) Fix $k \in \{1,\ldots,K_0\}$. By the consistency of $\tilde{\alpha}_k$ and $\tilde{\beta}_i$ in Theorem 4.1, we have $\tilde{\beta}_i - \tilde{\alpha}_l \to^P \alpha_k^0 - \alpha_l^0 \neq 0$ for all $i \in G_k^0$ and $l \neq k$, and $\tilde{c}_{ki} = \prod_{l=1,l\neq k}^{K_0}\|\tilde{\beta}_i - \tilde{\alpha}_l\| \to^P c_k^0 \equiv \prod_{l=1,l\neq k}^{K_0}\|\alpha_k^0 - \alpha_l^0\| > 0$ for $i \in G_k^0$. Now suppose that $\|\tilde{\beta}_i - \tilde{\alpha}_k\| \neq 0$ for some $i \in G_k^0$. Then the first order condition (with respect to β_i) for the minimization problem implies that

$$\begin{aligned}\mathbf{0}_{p\times 1} &= -\frac{2}{\sqrt{T}}\tilde{X}_i'\left(\ddot{y}_i - \tilde{X}_i\tilde{\beta}_i + \frac{1}{N}\sum_{j=1}^N \tilde{X}_j\tilde{\beta}_j\right) + \frac{N\sqrt{T}\lambda}{N-1}\sum_{j=1}^{K_0}\tilde{e}_{ij}\prod_{l=1,l\neq j}^{K_0}\|\tilde{\beta}_i - \tilde{\alpha}_l\|\\ &= -\frac{2}{\sqrt{T}}\tilde{X}_i'\left[\ddot{u}_i - \tilde{X}_i(\tilde{\beta}_i - \beta_i^0) + \frac{1}{N}\sum_{j=1}^N \tilde{X}_j(\tilde{\beta}_j - \beta_j^0)\right]\\ &\quad + \frac{N\sqrt{T}\lambda}{N-1}\sum_{j=1}^{K_0}\tilde{e}_{ij}\prod_{l=1,l\neq j}^{K_0}\|\tilde{\beta}_i - \tilde{\alpha}_l\|\\ &= -\frac{2}{\sqrt{T}}\tilde{X}_i'\ddot{u}_i + \left(\frac{2}{T}\tilde{X}_i'\tilde{X}_i + \frac{N}{N-1}\frac{\lambda\tilde{c}_{ki}}{\|\tilde{\beta}_i - \tilde{\alpha}_k\|}I_p\right)\sqrt{T}(\tilde{\beta}_i - \tilde{\alpha}_k)\\ &\quad + \frac{2}{T}\tilde{X}_i'\tilde{X}_i\sqrt{T}(\tilde{\alpha}_k - \alpha_k^0)\\ &\quad - \frac{2}{\sqrt{T}N}\sum_{j=1}^N \tilde{X}_i'\tilde{X}_j(\tilde{\beta}_j - \beta_j^0) + \frac{N\sqrt{T}\lambda}{N-1}\sum_{j=1,j\neq k}^{K_0}\tilde{e}_{ij}\prod_{l=1,l\neq j}^{K_0}\|\tilde{\beta}_i - \tilde{\alpha}_l\|\\ &\equiv \tilde{B}_{i1} + \tilde{B}_{i2} + \tilde{B}_{i3} + \tilde{B}_{i4} + \tilde{B}_{i5}, \text{ say,}\end{aligned} \tag{A.23}$$

where $\tilde{e}_{ij} = \frac{\tilde{\beta}_i - \tilde{\alpha}_j}{\|\tilde{\beta}_i - \tilde{\alpha}_j\|}$ if $\|\tilde{\beta}_i - \tilde{\alpha}_j\| \neq 0$ and $\|\tilde{e}_{ij}\| \leq 1$ if $\|\tilde{\beta}_i - \tilde{\alpha}_j\| = 0$.

Let $\chi_{NT} = T^{-1/2}(\ln T)^{3+\upsilon} + \lambda(\ln T)^{\upsilon}$. Following SSP and using Assumption 5(i), we can strengthen the result in Theorem 4.1 to obtain that

$$P\Big(\max_{1\le i\le N}\|\hat{b}_i\| \ge C\chi_{NT}\Big) = o\big(N^{-1}\big) \quad \text{for any } C>0. \tag{A.24}$$

This, in conjunction with the proof of Theorem 4.1(ii), implies that

$$\begin{aligned} &P\big(\|\tilde{\alpha}_k-\alpha_k^0\| \ge CT^{-1/2}(\ln T)^{\upsilon}\big) = o\big(N^{-1}\big) \quad \text{and} \\ &P\Big(\max_{i\in G_k^0}\big|\tilde{c}_{ki}-c_k^0\big| \ge c_k^0/2\Big) = o\big(N^{-1}\big). \end{aligned} \tag{A.25}$$

By (A.24) and (A.25), Lemma B.4(v), and Assumption 1,

$$P\Big(\max_{i\in G_k^0}\|\tilde{B}_{i3}\| > C(\ln T)^{3+\upsilon}\Big) = o\big(N^{-1}\big) \quad \text{and} \quad P\Big(\max_{i\in G_k^0}\|\tilde{B}_{i5}\| > C\sqrt{T}\kappa\chi_{NT}\Big) = o\big(N^{-1}\big).$$

Noting that $\frac{1}{\sqrt{TN}}\|\sum_{j=1}^N \tilde{X}_i'\tilde{X}_j(\tilde{\beta}_j-\beta_j^0)\| \le T^{-1/2}\|\tilde{X}_i\|\{\frac{1}{TN}\sum_{j=1}^N\|\tilde{X}_j\|^2\}^{1/2}\{\frac{T}{N}\sum_{j=1}^N\|\tilde{\beta}_j-\beta_j^0\|^2\}^{1/2} = O_P(1)$, we can readily show that $P(\max_{i\in G_k^0}\|\tilde{B}_{i4}\| > C(\ln T)^{3+\upsilon}) = o(N^{-1})$. It follows that $P(\Xi_{kNT}) = 1-o(N^{-1})$, where

$$\begin{aligned} \Xi_{kNT} \equiv{}& \Big\{\max_{i\in G_k^0}\big|\tilde{c}_{ki}-c_k^0\big| \le c_k^0/2\Big\} \cap \Big\{\max_{i\in G_k^0}\|\tilde{B}_{i3}\| \le C(\ln T)^{3+\upsilon}\Big\} \cap \Big\{\max_{i\in G_k^0}\|\tilde{B}_{i4}\| \le C(\ln T)^{3+\upsilon}\Big\} \\ &\cap \Big\{\max_{i\in G_k^0}\|\tilde{B}_{i5}\| \le C\sqrt{T}\lambda\chi_{NT}\Big\}. \end{aligned}$$

Then, conditional on Ξ_{kNT}, we have that uniformly in $i\in G_k^0$,

$$\begin{aligned} \big\|(\tilde{\beta}_i-\tilde{\alpha}_k)'(\tilde{B}_{i2}+\tilde{B}_{i3}+\tilde{B}_{i4}+\tilde{B}_{i5})\big\| &\ge \big\|(\tilde{\beta}_i-\tilde{\alpha}_k)'\tilde{B}_{i2}\big\| - \big\|(\tilde{\beta}_i-\tilde{\alpha}_k)'(\tilde{B}_{i3}+\tilde{B}_{i4}+\tilde{B}_{i5})\big\| \\ &\ge \frac{N}{N-1}\sqrt{T}\lambda\tilde{c}_{ki}\|\tilde{\beta}_i-\tilde{\alpha}_k\| \\ &\quad -C\|\tilde{\beta}_i-\tilde{\alpha}_k\|\big\{2(\ln T)^{3+\upsilon}+\sqrt{T}\lambda\chi_{NT}\big\} \\ &\ge \sqrt{T}\lambda c_k^0\|\tilde{\beta}_i-\tilde{\alpha}_k\|/4 \text{ for sufficiently large } (N,T), \end{aligned}$$

where the last inequality follows because $\tilde{c}_{ki} \ge c_k^0/2$ on Ξ_{kNT}, $\sqrt{T}\lambda \gg 3(\ln T)^{3+\upsilon} + \sqrt{T}\lambda\chi_{NT}$ under Assumption 5(ii). It follows that for all $i\in G_k^0$,

$$\begin{aligned} P(\tilde{E}_{kNT,i}) &= P\big(i\notin\tilde{G}_k \mid i\in G_k^0\big) \\ &= P(-\tilde{B}_{i1} = \tilde{B}_{i2}+\tilde{B}_{i3}+\tilde{B}_{i4}+\tilde{B}_{i5}) \\ &\le P\big(\big|(\tilde{\beta}_i-\tilde{\alpha}_k)'\tilde{B}_{i1}\big| \ge \big|(\tilde{\beta}_i-\tilde{\alpha}_k)'(\tilde{B}_{i2}+\tilde{B}_{i3}+\tilde{B}_{i4}+\tilde{B}_{i5})\big|\big) \\ &\le P\big(\|\tilde{\beta}_i-\tilde{\alpha}_k\|\|\tilde{B}_{i1}\| \ge \sqrt{T}\kappa c_k^0\|\hat{\beta}_i-\hat{\alpha}_k\|/4, \Xi_{kNT}\big) + P\big(\Xi_{kNT}^*\big) \\ &\le P\big(\|\tilde{B}_{i1}\| \ge \sqrt{T}\kappa c_k^0/4\big) + P\big(\Xi_{kNT}^*\big) = o\big(N^{-1}\big), \end{aligned}$$

where Ξ^*_{kNT} denotes the complement of Ξ_{kNT} and the convergence follows by Assumptions 1(i), 2(iv), and 3(i) (see the Remark after Assumption 3), and the fact that $P(\Xi^*_{kNT}) = o(N^{-1})$. Consequently, we can conclude that with probability $1 - o(N^{-1})$, that $\tilde{\beta}_i - \tilde{\alpha}_k$ must be in a position where $\|\beta_i - \alpha_k\|$ is not differentiable with respect to β_i for any $i \in G^0_k$. That is,

$$P\big(\|\tilde{\beta}_i - \tilde{\alpha}_k\| = 0 \mid i \in G^0_k\big) = 1 - o\big(N^{-1}\big) \quad \text{as } (N,T) \to \infty.$$

Then the rest of the proof follows SSP. □

Proof of Theorem 4.3. For the post-Lasso estimates $\hat{\alpha}_k$, we have the first order conditions

$$\frac{1}{N_k T}\sum_{i\in\hat{G}_k} \hat{\tilde{\mathbb{X}}}_i'\left(\ddot{y}_i - \tilde{X}_i\hat{\alpha}_k + \frac{1}{N}\sum_{l=1}^{K_0}\sum_{j\in\hat{G}_l}\tilde{X}_j\hat{\alpha}_l\right) = 0 \quad \text{for } k = 1,\dots,K_0,$$

where $\hat{\tilde{\mathbb{X}}}_i = \tilde{X}_i - \frac{1}{N}\sum_{j\in\hat{G}_k}\tilde{X}_j$ for $i \in \hat{G}_k$. It follows that $\mathrm{vec}(\hat{\boldsymbol{\alpha}}_{K_0}) = \hat{\mathbb{Q}}^{-1}_{NT}\hat{\mathbb{V}}_{NT}$. Let $\hat{\mathbb{Q}}_{kNT}$, $\hat{\mathbb{Q}}_{k,l}$, and $\hat{\mathbb{V}}_{kNT}$ be analogously defined as $\mathbb{Q}_{kNT}$, $\mathbb{Q}_{k,l}$, and $\mathbb{V}_{kNT}$ for $k, l = 1,\dots,K_0$. By Theorem 4.2 and using similar arguments as used in the proof of Proposition A.2, we can readily show that $\hat{\mathbb{Q}}_{kNT} = \mathbb{Q}_{kNT} + o_P((NT)^{-1/2})$, $\hat{\mathbb{Q}}_{k,l} = \mathbb{Q}_{k,l} + o_P((NT)^{-1/2})$, and $\hat{\mathbb{V}}_{kNT} = \mathbb{V}_{kNT} + o_P((NT)^{-1/2})$. It follows that

$$\mathrm{vec}(\hat{\boldsymbol{\alpha}}_{K_0}) = \mathbb{Q}^{-1}_{NT}\mathbb{V}_{NT} + o_P\big((NT)^{-1/2}\big).$$

That is, $\hat{\boldsymbol{\alpha}}$ is asymptotically equivalent to the infeasible oracle estimator $\bar{\boldsymbol{\alpha}}_{K_0}$ under $\mathbb{H}(K_0)$. Using $\ddot{y}_i = \tilde{X}_i\alpha^0_k - \frac{1}{N}\sum_{l=1}^{K_0}\sum_{j\in G^0_l}\tilde{X}_j\alpha^0_l + \ddot{u}_i$, we can readily show that

$$\mathrm{vec}\big(\hat{\boldsymbol{\alpha}}_{K_0} - \boldsymbol{\alpha}^0_{K_0}\big) = \mathbb{Q}^{-1}_{NT}\mathbb{U}^0_{NT} + o_P\big((NT)^{-1/2}\big),$$

where $\mathbb{U}^0_{NT} = (\mathbb{U}^{0\prime}_{1NT},\dots,\mathbb{U}^{0\prime}_{K_0NT})'$ and $\mathbb{U}_{kNT} = \frac{1}{N_k T}\sum_{i\in G^0_k}\tilde{\mathbb{X}}_i'\ddot{u}_i$. Noting that $\ddot{u}_i = M_0u_i - \bar{u}_{\cdot} + \bar{u}\mathbf{i}_T$ and $\tilde{\mathbb{X}}_i = \tilde{X}_i - \frac{1}{N}\sum_{j\in G^0_k}\tilde{X}_j$ for $i \in G^0_k$, we have

$$\mathbb{U}^0_{kNT} = \frac{1}{N_k T}\sum_{i\in G^0_k}\sum_{t=1}^{T}\left(\tilde{X}_{it} - \frac{1}{N}\sum_{j\in G^0_k}\tilde{X}_{jt}\right)(u_{it} - \bar{u}_{\cdot t}).$$

It is easy to show that $\mathbb{U}^0_{kNT} = O_P((NT)^{-1/2} + T^{-1})$ by moment calculations and Chebyshev inequality under Assumptions 1 and 2(i). Then $\mathrm{vec}(\hat{\boldsymbol{\alpha}}_{K_0} - \boldsymbol{\alpha}^0_{K_0}) = O_P((NT)^{-1/2} + T^{-1})$ under Assumption 6.

(ii) We make the decomposition

$$\begin{aligned}\mathbb{U}^0_{kNT} &= \frac{1}{N_k T}\sum_{i\in G^0_k}\sum_{t=1}^{T}(X_{it} - \bar{X}_{i\cdot})(u_{it} - \bar{u}_{\cdot t}) - \frac{1}{N_k T}\sum_{i\in G^0_k}\sum_{t=1}^{T}\left(\frac{1}{N}\sum_{j\in G^0_k}(X_{jt} - \bar{X}_{j\cdot})\right)(u_{it} - \bar{u}_{\cdot t})\\ &\equiv \mathbb{U}^0_{kNT,1} - \mathbb{U}^0_{kNT,2}.\end{aligned}$$

Under Assumptions 1 and 2(i), we can readily show that

$$\mathbb{U}^0_{kNT,1} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} u_{it}\left\{\frac{N}{N_k}1\{i \in G^0_k\}X_{it} - E\big(\bar{X}^{(k)}_{\cdot t} - \bar{X}^{(k)}\big)\right\} - \mathbb{B}_{kNT} + O_P\big(N^{-1}T^{-1/2}\big)$$

and

$$\mathbb{U}^0_{kNT,2} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\frac{N}{N_k}\big[1\{i \in G^0_k\} - 1\big]X_{it}u_{it}E\big(\bar{X}^{(k)}_{\cdot t} - \bar{X}^{(k)}\big) + O_P\big(N^{-1}T^{-1/2}\big),$$

where $\mathbb{B}_{kNT}$ is defined in Section 4.2. It follows that

$$\begin{aligned}\mathbb{U}^0_{kNT} &= \frac{1}{N_k T}\sum_{i \in G^0_k}\sum_{t=1}^{T}\big\{X_{it} - E\big(\bar{X}^{(k)}_{\cdot t} - \bar{X}^{(k)}\big)\big\}u_{it} - \mathbb{B}_{kNT} + O_P\big(N^{-1}T^{-1/2}\big) \\ &= \mathbb{U}_{kNT} - \mathbb{B}_{kNT} + O_P\big(N^{-1}T^{-1/2}\big).\end{aligned}$$

Then by Assumption 6(i) and (ii), $\sqrt{NT}\,\mathrm{vec}(\hat{\boldsymbol{\alpha}}_{K_0} - \boldsymbol{\alpha}^0_{K_0} + \mathbb{Q}^{-1}_{NT}\mathbb{B}_{NT}) = \mathbb{Q}^{-1}_{NT}\sqrt{NT}\mathbb{U}_{NT} + o_P(1) \to^D N(0, \mathbb{Q}^{-1}_0 \Omega_0 \mathbb{Q}^{-1}_0)$. □

Proof of Theorem 4.4. Let $\bar{u}_\cdot = (\bar{u}_{\cdot 1}, \ldots, \bar{u}_{\cdot T})'$. Noting that $\widehat{\ddot{u}}_i = \ddot{u}_i + \tilde{X}_i(\beta^0_i - \hat{\beta}_i) + \frac{1}{N}\sum_{j=1}^{N}\tilde{X}_j(\hat{\beta}_j - \beta^0_j) = M_0 u_i - \bar{u}_\cdot + \bar{u}\mathbf{i}_T + M_0 X_i(\beta^0_i - \hat{\beta}_i) + \frac{1}{N}\sum_{j=1}^{N} M_0 X_j(\hat{\beta}_j - \beta^0_j)$, $\bar{P}_{X_i} M_0 X_i = X_i$, and $M_0\mathbf{i}_T = 0$, we have

$$\begin{aligned}N^{-1/2}\,\mathrm{LM}_{2NT}(K_0) &= N^{-1/2}\sum_{i=1}^{N}\widehat{\ddot{u}}_i' M_0 \bar{P}_{X_i} M_0 \widehat{\ddot{u}}_i \\ &= N^{-1/2}\sum_{i=1}^{N}(M_0 u_i - \bar{u}_\cdot)' M_0 \bar{P}_{X_i} M_0 (M_0 u_i - \bar{u}_\cdot) \\ &\quad + N^{-1/2}\sum_{i=1}^{N}\big(\beta^0_i - \hat{\beta}_i\big)' X_i' M_0 X_i \big(\beta^0_i - \hat{\beta}_i\big) \\ &\quad + N^{-1/2}\sum_{i=1}^{N}\left(\frac{1}{N}\sum_{j=1}^{N}\big(\hat{\beta}_j - \beta^0_j\big)' X_j'\right) M_0 \bar{P}_{X_i} M_0 \left(\frac{1}{N}\sum_{j=1}^{N} X_j\big(\hat{\beta}_j - \beta^0_j\big)\right) \\ &\quad + 2N^{-1/2}\sum_{i=1}^{N}(M_0 u_i - \bar{u}_\cdot)' M_0 X_i \big(\beta^0_i - \hat{\beta}_i\big) \\ &\quad + 2N^{-1/2}\sum_{i=1}^{N}(M_0 u_i - \bar{u}_\cdot)' M_0 \bar{P}_{X_i} M_0 \left(\frac{1}{N}\sum_{j=1}^{N} X_j\big(\hat{\beta}_j - \beta^0_j\big)\right) \\ &\quad + 2N^{-1/2}\sum_{i=1}^{N}\big(\beta^0_i - \hat{\beta}_i\big)' X_i M_0 \left(\frac{1}{N}\sum_{j=1}^{N} X_j\big(\hat{\beta}_j - \beta^0_j\big)\right) \\ &\equiv D_{1NT} + D_{2NT} + D_{3NT} + 2D_{4NT} + 2D_{5NT} + 2D_{6NT}, \quad \text{say.}\end{aligned}$$

We analyze D_{lNT}, $l = 1, 2, \ldots, 6$, in turn.

(i) For D_{1NT}, we make the decomposition

$$D_{1NT} = N^{-1/2}\sum_{i=1}^{N} u_i' M_0 \bar{P}_{X_i} M_0 u_i + N^{-1/2}\sum_{i=1}^{N} \bar{u}_\cdot' M_0 \bar{P}_{X_i} M_0 \bar{u}_\cdot - 2N^{-1/2}\sum_{i=1}^{N} u_i' M_0 \bar{P}_{X_i} M_0 \bar{u}_\cdot$$
$$\equiv D_{1NT,1} + D_{1NT,2} - 2D_{1NT,3}.$$

The term $D_{1NT,1}$ is identical to A_{1NT} studied in the proof of Theorem 3.1 and thus we have $D_{1NT,1} - B_{NT} \to^D N(0, V_0)$. Using $M_0 = I_T - \frac{1}{T}\mathbf{i}_T\mathbf{i}_T'$, we can decompose $D_{1NT,2}$ as

$$D_{1NT,2} = N^{-1/2}\sum_{i=1}^{N} \bar{u}_\cdot' M_0 \bar{P}_{X_i} M_0 \bar{u}_\cdot$$
$$= N^{-1/2}\sum_{i=1}^{N} \bar{u}_\cdot' \bar{P}_{X_i} \bar{u}_\cdot + T^{-2}N^{-1/2}\sum_{i=1}^{N} \bar{u}_\cdot' \mathbf{i}_T\mathbf{i}_T' \bar{P}_{X_i} \mathbf{i}_T\mathbf{i}_T' \bar{u}_\cdot - 2T^{-1}N^{-1/2}\sum_{i=1}^{N} \bar{u}_\cdot' \bar{P}_{X_i} \mathbf{i}_T\mathbf{i}_T' \bar{u}_\cdot$$
$$\equiv D_{1NT,2a} + D_{1NT,2b} - 2D_{1NT,2c}, \quad \text{say.}$$

Note that $D_{1NT,2a} = T^{-1}N^{-1/2}\sum_{i=1}^{N} \bar{u}_\cdot' X_i \hat{\Omega}_i^{-1} X_i' \bar{u}_\cdot \le c_{2NT}\bar{D}_{1NT,2a}$, where $c_{2NT} \equiv [\min_{1\le i\le N} \lambda_{\min}(\hat{\Omega}_i)]^{-1} = O_P(1)$ and $\bar{D}_{1NT,2a} = T^{-1}N^{-1/2}\sum_{i=1}^{N} \bar{u}_\cdot' X_i X_i' \bar{u}_\cdot$. In view of the fact that $X_i'\bar{u}_\cdot = \sum_{t=1}^{T} \bar{u}_{\cdot t} X_{it} = \frac{1}{N}\sum_{j=1}^{N}\sum_{t=1}^{T} u_{jt} X_{it}$, and by straightforward moment calculations and Assumption 1(iv) and (v), we can show that

$$E|\bar{D}_{1NT,2a}| = T^{-1}N^{-1/2}\sum_{i=1}^{N} E\big(\bar{u}_\cdot' X_i X_i' \bar{u}_\cdot\big) = T^{-1}N^{-5/2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T} E\big(u_{jt}u_{ks}X_{it}'X_{is}\big)$$
$$= T^{-1}N^{-5/2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T} E\big(u_{jt}^2 X_{it}'X_{it}\big) = O\big(N^{-1/2}\big).$$

Then $\bar{D}_{1NT,2a} = O_P(N^{-1/2})$ by Markov inequality and $D_{1NT,2a} = O_P(N^{-1/2})$. For $D_{1NT,2b}$, we have

$$D_{1NT,2b} = T^{-3}N^{-1/2}\sum_{i=1}^{N} \bar{u}_\cdot' \mathbf{i}_T\mathbf{i}_T' X_i \hat{\Omega}_i^{-1} X_i' \mathbf{i}_T\mathbf{i}_T' \bar{u}_\cdot \le c_{3NT}\bar{D}_{1NT,2b},$$

where $c_{3NT} \equiv \max_{1\le i\le NT}(T^{-1}\mathbf{i}_T' X_i)^2 c_{2NT} = O_P(1)$ and $\bar{D}_{1NT,2b} = T^{-1}N^{-1/2}\sum_{i=1}^{N} \bar{u}_\cdot' \mathbf{i}_T \times \mathbf{i}_T' \bar{u}_\cdot$. Noting that $\mathbf{i}_T'\bar{u}_\cdot = \frac{1}{N}\sum_{j=1}^{N}\sum_{t=1}^{T} u_{jt}$, we have

$$E|\bar{D}_{1NT,2b}| = T^{-1}N^{-1/2}\sum_{i=1}^{N} E\big(\bar{u}_\cdot' \mathbf{i}_T\mathbf{i}_T' \bar{u}_\cdot\big) = T^{-1}N^{-5/2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T} E(u_{jt}u_{ks})$$
$$= T^{-1}N^{-5/2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T} E\big(u_{jt}^2\big) = O\big(N^{-1/2}\big).$$

Then $\bar{D}_{1NT,2b} = O_P(N^{-1/2})$ by Markov inequality and $D_{1NT,2b} = O_P(N^{-1/2})$. By the Cauchy–Schwarz inequality, $|D_{1NT,2c}| \le \{D_{1NT,2a}\}^{1/2}\{D_{1NT,2b}\}^{1/2} = O_P(N^{-1/2})$. Consequently we have shown that $D_{1NT,2} = O_P(N^{-1/2})$.

Now we study $D_{1NT,3}$. Recall that $h_{i,ts}$ denotes the (t,s) element of $M_0\bar{P}_{X_i}M_0$: $h_{i,ts} = T^{-1}\sum_{r=1}^{T}\sum_{q=1}^{T}\eta_{tr}X_{ir}'\hat{\Omega}_i^{-1}X_{iq}\eta_{qs}$, and $\bar{h}_{i,ts} \equiv T^{-1}\sum_{r=1}^{T}\sum_{q=1}^{T}\eta_{tr}X_{ir}'\Omega_i^{-1}X_{iq}\eta_{qs}$, where $\eta_{tr} = \mathbf{1}\{t=r\} - T^{-1}$. Following the proof of Lemma B.4(i), we can show that $D_{1NT,3} = N^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}u_{it}\bar{u}_{\cdot s}h_{i,ts} = \bar{D}_{1NT,3} + o_P(1)$, where $\bar{D}_{1NT,3} = N^{-1/2} \times \sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}u_{it}\bar{u}_{\cdot s}\bar{h}_{i,ts}$. Now, in view of the fact that

$$\begin{aligned}\bar{h}_{i,ts} &= T^{-1}X_{it}'\Omega_i^{-1}X_{is} - T^{-2}\sum_{q=1}^{T}X_{it}'\Omega_i^{-1}X_{iq}\\ &\quad - T^{-2}\sum_{r=1}^{T}X_{ir}'\Omega_i^{-1}X_{is} + T^{-3}\sum_{r=1}^{T}\sum_{q=1}^{T}X_{ir}'\Omega_i^{-1}X_{iq},\end{aligned} \tag{A.26}$$

we make the decomposition

$$\begin{aligned}\bar{D}_{1NT,3} &= N^{-3/2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}u_{it}u_{js}\bar{h}_{i,ts}\\ &= T^{-1}N^{-3/2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}u_{it}u_{js}X_{it}'\Omega_i^{-1}X_{is}\\ &\quad - T^{-2}N^{-3/2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}u_{it}u_{js}X_{it}'\Omega_i^{-1}X_{ir}\\ &\quad - T^{-2}N^{-3/2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}u_{it}u_{js}X_{ir}'\Omega_i^{-1}X_{is}\\ &\quad + T^{-3}N^{-3/2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}\sum_{q=1}^{T}u_{it}u_{js}X_{ir}'\Omega_i^{-1}X_{iq}\\ &\equiv \bar{D}_{1NT,3a} - \bar{D}_{1NT,3b} - \bar{D}_{1NT,3c} + \bar{D}_{1NT,3d}.\end{aligned}$$

By Assumption 1(iv) and (v), we can show that $E(\bar{D}_{1NT,3a}^2) = O(N^{-1})$, implying that $\bar{D}_{1NT,3a} = O_P(N^{-1/2})$. For $\bar{D}_{1NT,3b}$, we have

$$\begin{aligned}E(\bar{D}_{1NT,3b}) &= T^{-2}N^{-3/2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}E\big(u_{it}u_{is}X_{it}'\Omega_i^{-1}X_{ir}\big)\\ &= T^{-2}N^{-3/2}\sum_{i=1}^{N}\sum_{1\le t\ne s\ne r\le T}E\big(u_{it}u_{is}X_{it}'\Omega_i^{-1}X_{ir}\big) + O\big(N^{-1/2}\big)\\ &= d_1 + d_2 + O\big(N^{-1/2}\big),\end{aligned}$$

where $d_1 = T^{-2}N^{-3/2}\sum_{i=1}^{N}\sum_{1\le t<s<r\le T}E(u_{it}u_{is}X_{it}'\Omega_i^{-1}X_{ir})$ and $d_2 = T^{-2}N^{-3/2}\sum_{i=1}^{N}\sum_{1\le s<t<r\le T}E(u_{it}u_{is}X_{it}'\Omega_i^{-1}X_{ir})$. By Assumption 1(i), (iii), and (v) and Davydov inequality,

$$|d_1| \le CT^{-2}N^{-1/2}\sum_{1\le t<s<r\le T}\alpha(s-t)^{(1+\sigma)/(2+\sigma)} = O\big(N^{-1/2}\big).$$

Similarly, $d_2 = O(N^{-1/2})$ and $E(\bar{D}_{1NT,3b}) = O(N^{-1/2})$. Analogously, we can show that $E(\bar{D}_{1NT,3b}^2) = O(N^{-1})$. It follows that $\bar{D}_{1NT,3b} = O_P(N^{-1/2})$. By the same token, we can show that $\bar{D}_{1NT,3c} = o_P(1)$. Noting that

$$\bar{D}_{1NT,3d} = T^{-3}N^{-3/2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}\sum_{q=1}^{T}u_{it}u_{js}X_{ir}'\Omega_i^{-1}X_{iq},$$

we have

$$E\big(\bar{D}_{1NT,3d}^2\big) = T^{-6}N^{-3}\sum_{1\le i_1,i_2,i_3,i_4\le N}\ \sum_{1\le t_1,t_2,\ldots,t_8\le N}E\big(u_{i_1t_1}u_{i_2t_2}u_{i_3t_5}u_{i_4t_6}X_{i_1t_3}'\Omega_{i_1}^{-1}X_{i_1t_4}X_{i_3t_7}'\Omega_{i_3}^{-1}X_{i_3t_8}\big).$$

Clearly, the expectation in the last summation is 0 if $\#\{i_1, i_2, i_3, i_4\} \ge 3$. With this observation, we can readily apply Assumption 1(i)–(v) and Davydov inequality to show that $E(\bar{D}_{1NT,3d}^2) = O(N^{-1} + N^{-2}T) = o(1)$. Then $\bar{D}_{1NT,3d} = o_P(1)$. Consequently, we have $D_{1NT,3} = o_P(1)$.

In sum, we have shown that $D_{1NT} - B_{NT} \to^{D} N(0, V_0)$.

(ii) The term D_{2NT} is the same as A_{2NT} in the proof of Theorem 3.1 and thus is $o_P(1)$.

(iii) Noting that $M_0\bar{P}_{X_i}M_0$ is a projection matrix with maximum eigenvalue given by 1 and using Cauchy–Schwarz inequality, we have

$$\begin{aligned} D_{3NT} &\le N^{1/2}\left(\frac{1}{N}\sum_{j=1}^{N}\big(\hat{\beta}_j-\beta_j^0\big)'X_j'\right)\left(\frac{1}{N}\sum_{j=1}^{N}X_j\big(\hat{\beta}_j-\beta_j^0\big)\right) \\ &= TN^{1/2}\left\{\frac{1}{N}\sum_{j=1}^{N}\big\|\hat{\beta}_j-\beta_j^0\big\|^2\right\}\left\{\frac{1}{NT}\sum_{j=1}^{N}\|X_j\|^2\right\} \\ &= TN^{1/2}O_P\big((NT)^{-1}+T^{-2}\big)O_P(1) = o_P(1). \end{aligned}$$

(iv) For D_{4NT}, we make the decomposition

$$D_{4NT} = N^{-1/2}\sum_{i=1}^{N}u_i'M_0X_i\big(\beta_i^0-\hat{\beta}_i\big) - N^{-1/2}\sum_{i=1}^{N}\bar{u}_\cdot'M_0X_i\big(\beta_i^0-\hat{\beta}_i\big) \equiv D_{4NT,1} - D_{4NT,2}.$$

The term $D_{4NT,1}$ is the same as A_{3NT} studied in the proof of Theorem 3.1 and thus $D_{4NT,1} = o_P(1)$. For $D_{4NT,2}$, we can apply the Cauchy–Schwarz inequality to obtain

$$\begin{aligned} |D_{4NT,2}| &\le \left\{\sum_{i=1}^{N}\bar{u}_\cdot'M_0X_iX_i'M_0\bar{u}_\cdot\right\}^{1/2}\left\{N^{-1}\sum_{i=1}^{N}\big\|\beta_i^0-\hat{\beta}_i\big\|^2\right\}^{1/2} \\ &= O_P\big(T^{1/2}\big)O_P\big((NT)^{-1/2}+T^{-1}\big) = o_P(1) \end{aligned}$$

because $\sum_{i=1}^N \bar{u}'_{\cdot} M_0 X_i X_i' M_0 \bar{u}_{\cdot} = \sum_{i=1}^N \mathrm{tr}(M_0 X_i X_i' M_0 \bar{u}_{\cdot} \bar{u}'_{\cdot}) \le \sum_{i=1}^N \mathrm{tr}(X_i X_i' M_0 \bar{u}_{\cdot} \bar{u}'_{\cdot}) \le \sum_{i=1}^N \bar{u}'_{\cdot} X_i X_i' \bar{u}_{\cdot}$ and

$$\sum_{i=1}^N E\big(\bar{u}'_{\cdot} X_i X_i' \bar{u}_{\cdot}\big) = \frac{1}{N^2}\sum_{i=1}^N\sum_{j=1}^N\sum_{k=1}^N\sum_{t=1}^T\sum_{s=1}^T E\big(u_{jt}u_{ks}X_{it}'X_{is}\big) = \frac{1}{N^2}\sum_{i=1}^N\sum_{j=1}^N\sum_{t=1}^T E\big(u_{jt}^2 X_{it}'X_{it}\big)$$
$$= O(T).$$

Hence $D_{4NT} = o_P(1)$.

(v) For D_{5NT}, we make the decomposition

$$D_{5NT} = N^{-1/2}\sum_{i=1}^N u_i' M_0 \bar{P}_{X_i} M_0 \left(\frac{1}{N}\sum_{j=1}^N X_j\big(\hat{\beta}_j - \beta_j^0\big)\right)$$
$$- N^{-1/2}\sum_{i=1}^N \bar{u}'_{\cdot} M_0 \bar{P}_{X_i} M_0 \left(\frac{1}{N}\sum_{j=1}^N X_j\big(\hat{\beta}_j - \beta_j^0\big)\right)$$
$$\equiv D_{5NT,1} - D_{5NT,2}.$$

Using $M_0 = I_T - \frac{1}{T}\mathbf{i}_T\mathbf{i}_T'$, we can decompose $D_{5NT,1}$ as

$$D_{5NT,1} = N^{-3/2}\sum_{i=1}^N\sum_{j=1}^N u_i' M_0 \bar{P}_{X_i} M_0 X_j\big(\hat{\beta}_j - \beta_j^0\big)$$
$$= \sum_{k=1}^{K_0} N^{-3/2}\sum_{j\in\hat{G}_k}\sum_{i=1}^N u_i' M_0 \bar{P}_{X_i} M_0 X_j\big(\hat{\alpha}_k - \alpha_k^0\big) = \sum_{k=1}^{K_0} D_{5NT,1k}\big(\hat{\alpha}_k - \alpha_k^0\big),$$

where $D_{5NT,1k} = N^{-3/2}\sum_{j\in\hat{G}_k}\sum_{i=1}^N u_i' M_0 \bar{P}_{X_i} M_0 X_j$. Let $b_{NT} = \min((NT)^{1/2}, T)$. To show that $D_{5NT,1} = o_P(1)$, it suffices to prove that $D_{5NT,1k} = o_P(b_{NT})$ for $k = 1, \ldots, K_0$ as $\hat{\alpha}_k - \alpha_k^0 = O_P((NT)^{-1/2} + T^{-1})$. Following the proofs of Propositions A.3 and A.1, we can show that

$$D_{5NT,1k} = N^{-3/2}\sum_{j\in G_k^0}\sum_{i=1}^N u_i' M_0 \bar{P}_{X_i} M_0 X_j + o_P(1)$$
$$= N^{-3/2}\sum_{j\in G_k^0}\sum_{i=1}^N\sum_{t=1}^T\sum_{s=1}^T u_{it}h_{i,ts}X_{js} + o_P(1) = \bar{D}_{5NT,1k} + o_P(b_{NT}),$$

where $\bar{D}_{5NT,1k} = N^{-3/2}\sum_{j\in G_k^0}\sum_{i=1}^N\sum_{t=1}^T\sum_{s=1}^T u_{it}X_{js}\bar{h}_{i,ts}$. Now using (A.26) we make the decomposition

$$\bar{D}_{5NT,1k} = T^{-1}N^{-3/2}\sum_{j\in G_k^0}\sum_{i=1}^N\sum_{t=1}^T\sum_{s=1}^T u_{it}X_{js}X_{it}'\Omega_i^{-1}X_{is}$$
$$- T^{-2}N^{-3/2}\sum_{j\in G_k^0}\sum_{i=1}^N\sum_{t=1}^T\sum_{s=1}^T\sum_{r=1}^T u_{it}X_{js}X_{it}'\Omega_i^{-1}X_{ir}$$

$$- T^{-2}N^{-3/2}\sum_{j\in G_k^0}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}u_{it}X_{js}X_{ir}'\Omega_i^{-1}X_{is}$$

$$+ T^{-3}N^{-3/2}\sum_{j\in G_k^0}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}\sum_{q=1}^{T}u_{it}X_{js}X_{ir}'\Omega_i^{-1}X_{iq}$$

$$\equiv \bar{D}_{5NT,1k}(1)-\bar{D}_{5NT,1k}(2)-\bar{D}_{5NT,1k}(3)+\bar{D}_{5NT,1k}(4).$$

As in the analysis of $\bar{D}_{1NT,3}$, we can show that $\bar{D}_{5NT,1k}(l)=o_P(b_{NT})$ for $l=1,\ldots,4$. For example, we can easily show that $\bar{D}_{5NT,1k}(1)$ has the same probability order as $T^{-1}N^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}u_{it}X_{it}'\Omega_i^{-1}X_{is}c_s$ with $c_s=\frac{1}{N}\sum_{j\in G_k^0}E(X_{js})$, and the squared Frobenius norm of the last term has expectation given by

$$T^{-2}N^{-1}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}\sum_{q=1}^{T}E\big(u_{it}u_{jr}X_{it}'\Omega_i^{-1}X_{is}X_{js}'\Omega_j^{-1}X_{jr}\big)c_s'c_q$$

$$=T^{-2}N^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}\sum_{q=1}^{T}E\big(u_{it}u_{ir}X_{it}'\Omega_i^{-1}X_{is}X_{is}'\Omega_i^{-1}X_{ir}\big)c_s'c_q$$

$$+T^{-2}N^{-1}\sum_{i=1}^{N}\sum_{j=1,j\neq i}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}\sum_{q=1}^{T}E\big(u_{it}X_{it}'\Omega_i^{-1}X_{is}\big)E\big(u_{jr}X_{js}'\Omega_j^{-1}X_{jr}\big)c_s'c_q$$

$$=O(T)+O(N)$$

by the repeated use of Davydov inequality. It follows that $\bar{D}_{5NT,1k}(1)=O_P(N^{1/2}+T^{1/2})=o_P(b_{NT})$. Then we have $D_{5NT,1k}=o_P(b_{NT})$.

For $D_{5NT,2}$, we apply $M_0=I_T-\frac{1}{T}\mathbf{i}_T\mathbf{i}_T'$ to make the decomposition

$$D_{5NT,2}=N^{-1/2}\sum_{i=1}^{N}\bar{u}_\cdot' X_i\big(X_i'M_0X_i\big)^{-1}X_i'M_0\left(\frac{1}{N}\sum_{j=1}^{N}X_j\big(\hat{\beta}_j-\beta_j^0\big)\right)$$

$$-T^{-1}N^{-1/2}\sum_{i=1}^{N}\bar{u}_\cdot'\mathbf{i}_T\mathbf{i}_T'\big(X_i'M_0X_i\big)^{-1}X_i'M_0\left(\frac{1}{N}\sum_{j=1}^{N}X_j\big(\hat{\beta}_j-\beta_j^0\big)\right)$$

$$=D_{5NT,2a}-D_{5NT,2b}.$$

By Cauchy–Schwarz inequality, we have $D_{5NT,2a}\le\{D_{5NT,2a}(1)\}^{1/2}\{D_{5NT,2a}(2)\}^{1/2}$, where $D_{5NT,2a}(1)=T^{-1}\sum_{i=1}^{N}\bar{u}_\cdot' X_iX_i'\bar{u}_\cdot$ and

$$D_{5NT,2a}(2)=N^{-1}T\sum_{i=1}^{N}\left(\frac{1}{N}\sum_{j=1}^{N}X_j\big(\hat{\beta}_j-\beta_j^0\big)\right)'$$

$$\times M_0X_i\big(X_i'M_0X_i\big)^{-2}X_i'M_0\left(\frac{1}{N}\sum_{j=1}^{N}X_j\big(\hat{\beta}_j-\beta_j^0\big)\right)^{1/2}.$$

By Markov inequality, we can readily show that $D_{5NT,2a}(1) = O_P(1)$. In addition, noting that $M_0\bar{P}_{X_i}M_0$ is a projection matrix and has maximum eigenvalue 1, we have

$$D_{5NT,2a}(2) \le c_{2NT}N^{-1}\sum_{i=1}^{N}\left(\frac{1}{N}\sum_{j=1}^{N}X_j(\hat{\beta}_j-\beta_j^0)\right)' M_0\bar{P}_{X_i}M_0\left(\frac{1}{N}\sum_{j=1}^{N}X_j(\hat{\beta}_j-\beta_j^0)\right)^{1/2}$$

$$\le c_{2NT}\left\|\frac{1}{N}\sum_{j=1}^{N}(\hat{\beta}_j-\beta_j^0)'X_j'\right\|^2 \le Tc_{2NT}\frac{1}{N}\sum_{j=1}^{N}\|\hat{\beta}_j-\beta_j^0\|^2\frac{1}{NT}\sum_{j=1}^{N}\|X_j\|^2$$

$$= TO_P(1)O_P\big((NT)^{-1}+T^{-2}\big)O_P(1) = O_P\big(N^{-1}+T^{-1}\big).$$

It follows that $D_{5NT,2a} = O_P(N^{-1/2}+T^{-1/2}) = o_P(1)$. Similarly, we can show that $D_{5NT,2b} = o_P(1)$. Thus $D_{5NT,2} = o_P(1)$ and $D_{5NT} = o_P(1)$.

(vi) By the Cauchy–Schwarz inequality and the results in (ii) and (iii), $D_{6NT} \le \{D_{2NT}\}^{1/2}\{D_{3NT}\}^{1/2} = o_P(1)$.

Combining the results in (i)–(vi) completes the proof of the theorem. □

Appendix B: Some technical lemmas

Define the mth order U statistic $\mathcal{U}_T = \binom{T}{m}^{-1}\sum_{1\le t_1<\cdots<t_m\le T}\vartheta(\xi_{t_1},\ldots,\xi_{t_m})$, where ϑ is symmetric in its arguments. Let $F_t(\cdot)$ denote the distribution function of ξ_t. Let $\vartheta_{(0)} = \int\cdots\int\vartheta(v_{t_1},\ldots,v_{t_m})\prod_{s=1}^{m}dF_{t_s}(v_{t_s})$ and $\vartheta_{(c)}(v_1,\ldots,v_c) = \int\cdots\int\vartheta(v_1,\ldots,v_c,v_{t_{c+1}},\ldots,v_{t_m})\prod_{s=c+1}^{m}dF_{t_s}(v_{t_s})$ for $c=1,\ldots,m$. Let $h^{(1)}(v) = \vartheta_{(1)}(v)-\vartheta_{(0)}$ and $h^{(c)}(v_1,\ldots,v_c) = \vartheta_{(c)}(v_1,\ldots,v_c) - \sum_{j=1}^{c-1}\sum_{(c,j)}h^{(j)}(v_{t_1},\ldots,v_{t_j}) - \vartheta_{(0)}$ for $c=2,\ldots,m$, where the sum $\sum_{(c,j)}$ is taken over all subsets $1\le t_1<t_2<\cdots<t_j\le c$ of $\{1,2,\ldots,c\}$. Let $\mathcal{H}_T^{(c)} = \binom{T}{c}^{-1}\sum_{1\le t_1<\cdots<t_c\le T}h^{(c)}(\xi_{t_1},\ldots,\xi_{t_c})$. Then by Theorem 1 in Lee (1990, p. 26), we have the Hoeffding decomposition

$$\mathcal{U}_T = \vartheta_{(0)} + \sum_{c=1}^{m}\binom{m}{c}\mathcal{H}_T^{(c)}. \tag{B.1}$$

To study the second moment of $\mathcal{H}_T^{(c)}$ for $3\le c\le m$, we need the following lemma.

Lemma B.1. *Let $\{\xi_t, t\ge 1\}$ be an l-dimensional strong mixing process with mixing coefficient $\alpha(\cdot)$ and distribution function $F_t(\cdot)$. Let the integers $(t_1,\ldots,t_m)$ be such that $1\le t_1<t_2<\cdots<t_m\le T$. Suppose that $\max\{\int|\vartheta(v_1,\ldots,v_m)|^{1+\tilde{\sigma}}dF_{t_1,\ldots,t_m}(v_1,\ldots,v_m), \int|\vartheta(v_1,\ldots,v_m)|^{1+\tilde{\sigma}}dF_{t_1,\ldots,t_j}(v_1,\ldots,v_j)\,dF_{t_{j+1},\ldots,t_m}(v_{j+1},\ldots,v_m)\}\le C$ for some $\tilde{\sigma}>0$, where, for example, $F_{t_1,\ldots,t_m}(v_1,\ldots,v_m)$ denotes the distribution function of $(\xi_{t_1},\ldots,\xi_{t_m})$. Then*

$$\left|\int\vartheta(v_1,\ldots,v_m)\,dF_{t_1,\ldots,t_m}(v_1,\ldots,v_m) - \int\vartheta(v_1,\ldots,v_m)\,dF^{(1)}_{t_1,\ldots,t_j}(v_1,\ldots,v_j)\,dF_{t_{j+1},\ldots,t_m}(v_{j+1},\ldots,v_m)\right| \le 4C^{1/(1+\tilde{\sigma})}\alpha(t_{j+1}-t_j)^{\tilde{\sigma}/(1+\tilde{\sigma})}.$$

For the proof, see Lemma 2.1 in Sun and Chiang (1997).

Lemma B.2. *Let $\{\xi_t, t \geq 1\}$ be an l-dimensional strong mixing process with mixing coefficient $\alpha(\cdot)$ and distribution function $F_t(\cdot)$. Suppose that $\alpha(s) = O(s^{-3(2+\sigma)/\sigma-\epsilon})$. If there exists $\sigma > 0$ such that*

$$L_T \equiv \max\left\{\int \left|\vartheta(v_{t_1}, \ldots, v_{t_m})\right|^{2+\sigma} \prod_{s=1}^{m} dF_{t_s}(v_{t_s}), E\left|\vartheta(\xi_{t_1}, \ldots, \xi_{t_m})\right|^{2+\sigma}\right\} \leq \sum_{q=1}^{m} C_q(t_q)$$

and $T^{-1}\sum_{q=1}^{m}\sum_{t_q=1}^{T} C_q(t_q) = O(1)$, then $E[\mathcal{H}_T^{(c)}]^2 = O_P(T^{-3})$ for $3 \leq c \leq m$.

The proof is analogous to that of Lemma A.6 in Su and Chen (2013), who consider conditional strong mixing processes instead.

Lemma B.3. *Recall that $\hat{\Omega}_i \equiv X_i' M_0 X_i / T$ and $\Omega_i \equiv E(\hat{\Omega}_i)$. Let $\hat{\Omega}_{1i} \equiv X_i' X_i / T$ and $\Omega_{1i} \equiv E(\hat{\Omega}_{1i})$. Suppose Assumptions 1–3 hold. Then* (i) $\lambda_{\max}(\hat{\Omega}_{1i}) \leq \lambda_{\max}(\Omega_{1i}) + O_P(T^{-1/2})$, (ii) $\lambda_{\min}(\hat{\Omega}_{1i}) \geq \mu_{\min}(\Omega_{1i}) - O_P(T^{-1/2})$, (iii) $\max_{1\leq i\leq N}\|\hat{\Omega}_{1i} - \Omega_{1i}\| = O_P(a_{NT})$, (iv) $\max_{1\leq i\leq N}\|\hat{\Omega}_{1i}^{-1} - \Omega_{1i}^{-1}\| = O_P(a_{NT})$, *and* (v) $\max_{1\leq i\leq N}\|\hat{\Omega}_i - \Omega_i\| = O_P(a_{NT})$ *and* $\max_{1\leq i\leq N}\|\hat{\Omega}_i^{-1} - \Omega_i^{-1}\| = O_P(a_{NT})$, *where* $a_{NT} \equiv \max\{(NT)^{1/(4+2\sigma)} \times \log(NT)/T, (\log(NT)/T)^{1/2}\}$.

Proof. The results in (i) and (ii) follow from Lemma A.1(iv) and (v) in Su and Jin (2012). Su and Chen (2013, Lemma A.7) prove (iii) for the conditional strong mixing process. The result also holds for strong mixing processes with a simple application of the Bernstein-type inequality for strong mixing processes (see, e.g., Lemma 2.2 in Sun and Chiang (1997)). Result (iv) follows from (i)–(iii) and the submultiplicative property of the Frobenius norm.

Now we show (v). Using $M_0 = I_T - T^{-1}\mathbf{i}_T\mathbf{i}_T'$, we can decompose $\hat{\Omega}_i - \Omega_i$ as

$$\begin{aligned}
\hat{\Omega}_i - \Omega_i &= T^{-1}\left[X_i' M_0 X_i - E\left(X_i' M_0 X_i\right)\right] \\
&= (\hat{\Omega}_{1i} - \Omega_{1i}) - \bar{X}_{i\cdot}\bar{X}_{i\cdot}' + E\left(\bar{X}_{i\cdot}\bar{X}_{i\cdot}'\right) \\
&= (\hat{\Omega}_{1i} - \Omega_{1i}) - \left[\bar{X}_{i\cdot} - E(\bar{X}_{i\cdot})\right]\left[\bar{X}_{i\cdot} - E(\bar{X}_{i\cdot})\right]' - \left[\bar{X}_{i\cdot} - E(\bar{X}_{i\cdot})\right]E\left(\bar{X}_{i\cdot}'\right) \\
&\quad - E(\bar{X}_{i\cdot})\left[\bar{X}_{i\cdot} - E(\bar{X}_{i\cdot})\right]' + \left[E\left(\bar{X}_{i\cdot}\bar{X}_{i\cdot}'\right) - E(\bar{X}_{i\cdot})E\left(\bar{X}_{i\cdot}'\right)\right].
\end{aligned}$$

Following the proof of (iii), we can show that $\max_{1\leq i\leq N}\|\bar{X}_{i\cdot} - E(\bar{X}_{i\cdot})\| = O_P(a_{1NT})$, where $a_{1NT} \equiv \max\{(NT)^{1/(8+4\sigma)}\log(NT)/T, (\log(NT)/T)^{1/2}\} = O(a_{NT})$. $\max_{1\leq i\leq N}\|E(\bar{X}_{i\cdot})\| = O(1)$ by Assumption 1(i). Let $b_{i,kl}$ denote the (k,l)th element of $E(\bar{X}_{i\cdot}\bar{X}_{i\cdot}') - E(\bar{X}_{i\cdot})E(\bar{X}_{i\cdot}')$ for $k, l = 1, \ldots, p$. Then by triangle inequality, Davydov inequality, and Assumption 2(iii),

$$\begin{aligned}
|b_{i,kl}| &= \frac{1}{T^2}\left|\sum_{t=1}^{T}\sum_{s=1}^{T}\mathrm{cov}(X_{it,k}, X_{is,l})\right| \\
&\leq \frac{1}{T^2}\sum_{t=1}^{T}\left|\mathrm{cov}(X_{it,k}, X_{it,l})\right| + \frac{1}{T^2}\sum_{1\leq t\neq s\leq T}\left|\mathrm{cov}(X_{it,k}, X_{is,l})\right| \\
&\leq O\left(T^{-1}\right) + \frac{8c_k c_l}{T}\sum_{\tau=1}^{\infty}\alpha(\tau)^{(3+2\sigma)/(4+2\sigma)} = O\left(T^{-1}\right),
\end{aligned}$$

where $c_k \le \sup_{T,N\ge 1} \max_{1\le i\le N} \max_{1\le t\le T} \|X_{it,k}\|_{8+4\sigma}$. Then by the triangle inequality, we have $\max_{1\le i\le N} \|\hat{\Omega}_i - \Omega_i\| = O_P(a_{NT})$. Result (vi) follows from (v) and Assumption 1(ii). □

LEMMA B.4. *Let $h_{i,ts}$ and $\bar{h}_{i,ts}$ be as defined in the proof of Theorem* 3.1. *Suppose Assumptions* 1–3 *hold. Then*

(i) $D_{1NT} \equiv N^{-1/2} \sum_{i=1}^N \sum_{1\le s\neq t\le T} u_{it}u_{is}(h_{i,ts} - \bar{h}_{i,ts}) = o_P(1)$,

(ii) $D_{2NT} \equiv T^{-2}N^{-1/2} \sum_{i=1}^N \sum_{1\le s<t\le T} \sum_{r=1}^T u_{it}u_{is}[X_{ir} - E(X_{ir})]'\Omega_i^{-1}X_{is} = o_P(1)$,

(iii) $D_{3NT} \equiv T^{-2}N^{-1/2} \sum_{i=1}^N \sum_{1\le s<t\le T} \sum_{q=1}^T u_{it}u_{is}X_{it}'\Omega_i^{-1}[X_{is} - E(X_{iq})] = o_P(1)$,

(iv) $D_{4NT} \equiv T^{-3}N^{-1/2} \sum_{i=1}^N \sum_{1\le s<t\le T} \sum_{r=1}^T \sum_{q=1}^T u_{it}u_{is}[X_{ir} - E(X_{ir})]'\Omega_i^{-1}[X_{iq} - E(X_{iq})] = o_P(1)$,

(v) $D_{5NT} \equiv T^{-3}N^{-1/2} \sum_{i=1}^N \sum_{1\le s<t\le T} \sum_{r=1}^T \sum_{q=1}^T u_{it}u_{is}[X_{ir} - E(X_{ir})]'\Omega_i^{-1}E(X_{iq}) = o_P(1)$.

PROOF. The proof of (i) is analogous to that of Lemma A.8 in Su and Chen (2013) except that we replace their Lemmas A.5–A.7 by Lemmas B.1–B.3. To show (ii), letting $c_{i,ts} \equiv [X_{it} - E(X_{it})]'\Omega_i^{-1}X_{is}$, we can decompose D_{2NT} as

$$
\begin{aligned}
D_{2NT} &= \frac{1}{T^2\sqrt{N}} \sum_{i=1}^N \sum_{1\le s<t\le T} \sum_{r=1, r\neq t,s}^T u_{it}u_{is}c_{i,rs} + \frac{1}{T^2\sqrt{N}} \sum_{i=1}^N \sum_{1\le s<t\le T} u_{it}u_{is}c_{i,ts} \\
&\quad + \frac{1}{T^2\sqrt{N}} \sum_{i=1}^N \sum_{1\le s<t\le T} u_{it}u_{is}c_{i,ss} \\
&\equiv D_{2NT,1} + D_{2NT,2} + D_{2NT,3}, \quad \text{say.}
\end{aligned}
$$

Let $\xi_{it} = (u_{it}, X_{it}')'$, $\varphi_0(\xi_{it}, \xi_{is}, \xi_{ir}) = u_{it}u_{is}c_{i,rs}$, and $\varphi(\xi_{it}, \xi_{is}, \xi_{ir}) = [\varphi_0(\xi_{it}, \xi_{is}, \xi_{ir}) + \varphi_0(\xi_{it}, \xi_{ir}, \xi_{is}) + \varphi_0(\xi_{is}, \xi_{it}, \xi_{ir}) + \varphi_0(\xi_{is}, \xi_{ir}, \xi_{it}) + \varphi_0(\xi_{ir}, \xi_{it}, \xi_{is}) + \varphi_0(\xi_{ir}, \xi_{is}, \xi_{it})]/6$. Let $d_{iNT} \equiv \binom{T}{3}^{-1} \sum_{1\le r<s<t\le T} \varphi(\xi_{it}, \xi_{is}, \xi_{ir})$. Then $D_{2NT,1} = \frac{a_T}{\sqrt{N}} \sum_{i=1}^N d_{iNT}$, where $a_T = \frac{(T-1)(T-2)}{2T}$. By Assumption 1 and Lemma B.2, $E(D_{2NT,1}^2) = \frac{a_T^2}{N} \sum_{i=1}^N E(d_{iNT}^2) = a_T^2 \times O_P(T^{-3}) = O_P(T^{-1})$. It follows that $D_{2NT,1} = O_P(T^{-1/2})$. Noting that $E_{\mathcal{D}}(D_{2NT,2}) = 0$ and $E(D_{2NT,2}^2) = \frac{1}{T^4N} \sum_{i=1}^N \sum_{1\le s,r<t\le T} E(u_{it}^2u_{is}u_{ir}c_{i,ts}c_{i,tr}) = O_P(T^{-1})$, we have $D_{2NT,2} = O_P(T^{-1/2})$. Similarly, $D_{2NT,3} = O_P(T^{-1/2})$. Then (ii) follows. The proofs of (iii)–(v) are analogous to that of (ii) and thus are omitted. □

REFERENCES

Bertsekas, D. P. (1995), *Nonlinear Programming*. Athena Scientific, Belmont, MA. [20]

Hall, P. and C. C. Heyde (1980), *Martingale Limit Theory and Its Applications*. Academic Press, New York. [3, 12]

Lee, A. J. (1990), *U-Statistics: Theory and Practice*. Marcel Dekker, New York. [31]

Pollard, D. (1984), *Convergence of Stochastic Processes*. Springer-Verlag, New York. [3]

Su, L. and Q. Chen (2013), "Testing homogeneity in panel data models with interactive fixed effects." *Econometric Theory*, 29, 1079–1135. [32, 33]

Su, L. and S. Jin (2012), "Sieve estimation of panel data models with cross section dependence." *Journal of Econometrics*, 169, 34–47. [32]

Su, L., Z. Shi, and P. C. B. Phillips (2016), "Identifying latent structures in panel." *Econometrica*, 84, 2215–2264. [20]

Sun, S. and C.-Y. Chiang (1997), "Limiting behavior of the perturbed empirical distribution functions evaluated at U-statistics for strongly mixing sequences of random variables." *Journal of Applied Mathematics and Stochastic Analysis*, 10, 3–20. [31, 32]

Co-editor Frank Schorfheide handled this manuscript.

Manuscript received 24 November, 2014; final version accepted 27 January, 2017; available online 2 February, 2017.