# TECHNICAL APPENDIX FOR ASYMPTOTIC PROPERTIES OF A ROBUST VARIANCE MATRIX ESTIMATOR FOR PANEL DATA WHEN $T$ IS LARGE (TO BE OMITTED FROM PUBLICATION)

CHRISTIAN HANSEN

## Appendix A. Preliminaries

Throughout the appendix, let $\|A\| = [\text{trace}(A'A)]^{1/2}$ be the Euclidean norm of a matrix $A$, and let $\sum_i = \sum_{i=1}^n$, $\sum_t = \sum_{t=1}^T$, and $\sum_h = \sum_{h=1}^k$. Repeated use will be made of the following simple results, which are stated here for convenience.

**Lemma A.1.** *For matrices $A$ and $B$,* $\text{E}\|A \otimes B\|^r \leq (\text{E}\|A\|^{2r}\text{E}\|B\|^{2r})^{1/2}$.

**Proof.** $\text{E}\|A \otimes B\|^r = \text{E}\{[\text{trace}(AA' \otimes BB')]^{r/2}\} = \text{E}\{[\text{trace}(AA')]^{r/2}[\text{trace}(BB')]^{r/2}\} = \text{E}(\|A\|^r\|B\|^r)$ $\leq (\text{E}\|A\|^{2r}\text{E}\|B\|^{2r})^{1/2}$ where the equalities follow from the definition of $\|A\|$ and properties of the Kronecker product and the inequality results from the Cauchy-Schwarz inequality. ■

**Lemma A.2.** *Suppose $\{Z_{i,T}\}$ are independent across $i$ for all $T$ with $\text{E}[Z_{i,T}] = \mu_{i,T}$ and $\text{E}|Z_{i,T}|^{1+\delta} < \Delta < \infty$ for some $\delta > 0$ and all $i$, $T$. Then $\frac{1}{n}\sum_{i=1}^n(Z_{i,T} - \mu_{i,T}) \xrightarrow{p} 0$ as $\{n,\ T\} \to \infty$ jointly.*

**Proof.** The proof follows from standard arguments, cf. Chung (2001) Chapter 5. Details are given in Appendix G. ■

**Lemma A.3.** *For $k \times 1$ vectors $Z_{i,T}$, suppose $\{Z_{i,T}\}$ are independent across $i$ for all $T$ with $\text{E}[Z_{i,T}] = 0$, $\text{E}[Z_{i,T}Z_{i,T}'] = \Omega_{i,T}$, and $\text{E}\|Z_{i,T}\|^{2+\delta} < \Delta < \infty$ for some $\delta > 0$. Assume $\Omega = \lim_{n,T}\frac{1}{n}\sum_{i=1}^n\Omega_{i,T}$ is positive definite with minimum eigenvalue $\lambda_{\min} > 0$. Then $\frac{1}{\sqrt{n}}\sum_{i=1}^n Z_{i,T} \xrightarrow{d} N(0,W)$ as $\{n,\ T\} \to \infty$ jointly.*

**Proof.** The result follows from verifying the Lindeberg condition of Theorem 2 in Phillips and Moon (1999) using an argument similar to that used in the proof of Theorem 3 in Phillips and Moon (1999). Details are given in Appendix G. ■

The final lemma simply restates Doukhan (1994) Theorem 2 with a slight change of notation. Its proof may be found in Doukhan (1994) p. 25-30.

**Lemma A.4.** *Let $\{z_t\}$ be a strong mixing sequence with $\text{E}[z_t] = 0$, $\text{E}\|z_t\|^{\tau+\epsilon} < \Delta < \infty$, and mixing coefficient $\alpha(m)$ of size $\frac{(1-c)r}{r-c}$ where $c \in 2\mathbb{N}$, $c \geq \tau$, and $r > c$. Then there is a constant $C$ depending only on $\tau$ and $\alpha(m)$ such that $\text{E}|\sum_{t=1}^T y_t|^\tau \leq CD(\tau,\epsilon,T)$ with $D(\tau,\epsilon,T)$ defined in Doukhan (1994) and satisfying $D(\tau,\epsilon,T) = O(T)$ if $\tau \leq 2$ and $D(\tau,\epsilon,T) = O(T^{\tau/2})$ if $\tau > 2$.*

Finally note

$$\widehat{W} = \frac{1}{nT}\sum_i x_i'\widehat{\epsilon}_i\widehat{\epsilon}_i' x_i$$

$$= \frac{1}{nT}\sum_i x_i'\epsilon_i\epsilon_i' x_i \tag{O.1}$$

$$- \frac{1}{nT}\sum_i x_i' x_i(\widehat{\beta}-\beta)\epsilon_i' x_i \tag{O.2}$$

$$- \frac{1}{nT}\sum_i x_i'\epsilon_i(\widehat{\beta}-\beta)' x_i' x_i \tag{O.3}$$

$$+ \frac{1}{nT}\sum_i x_i' x_i(\widehat{\beta}-\beta)(\widehat{\beta}-\beta)' x_i' x_i, \tag{O.4}$$

and

$$\begin{aligned}\widehat{V} &= \frac{1}{nT}\sum_i [(\mathrm{vec}(x_i'\widehat{\epsilon}_i\widehat{\epsilon}_i' x_i - \widehat{W}))(\mathrm{vec}(x_i'\widehat{\epsilon}_i\widehat{\epsilon}_i' x_i - \widehat{W}))'] \\ &= \frac{1}{nT}\sum_i (\mathrm{vec}(x_i'\widehat{\epsilon}_i\widehat{\epsilon}_i' x_i))(\mathrm{vec}(x_i'\widehat{\epsilon}_i\widehat{\epsilon}_i' x_i))' - \frac{1}{nT}\sum_i (\mathrm{vec}(\widehat{W}))\frac{1}{nT}\sum_i (\mathrm{vec}(\widehat{W}))'\end{aligned}$$

where

$$\frac{1}{nT}\sum_i (\mathrm{vec}(x_i'\widehat{\epsilon}_i\widehat{\epsilon}_i' x_i))(\mathrm{vec}(x_i'\widehat{\epsilon}_i\widehat{\epsilon}_i' x_i))'$$

$$= \frac{1}{nT}\sum_i (\mathrm{vec}(x_i'\epsilon_i\epsilon_i' x_i))(\mathrm{vec}(x_i'\epsilon_i\epsilon_i' x_i))' \tag{V.1}$$

$$- \frac{1}{nT}\sum_i (\mathrm{vec}(x_i'\epsilon_i\epsilon_i' x_i))(\mathrm{vec}(x_i'\epsilon_i(\widehat{\beta}-\beta)' x_i' x_i))' \tag{V.2}$$

$$- \frac{1}{nT}\sum_i (\mathrm{vec}(x_i'\epsilon_i\epsilon_i' x_i))(\mathrm{vec}(x_i' x_i(\widehat{\beta}-\beta)\epsilon_i' x_i))' \tag{V.3}$$

$$+ \frac{1}{nT}\sum_i (\mathrm{vec}(x_i'\epsilon_i\epsilon_i' x_i))(\mathrm{vec}(x_i' x_i(\widehat{\beta}-\beta)(\widehat{\beta}-\beta)' x_i' x_i))' \tag{V.4}$$

$$- \frac{1}{nT}\sum_i (\mathrm{vec}(x_i'\epsilon_i(\widehat{\beta}-\beta)' x_i' x_i))(\mathrm{vec}(x_i'\epsilon_i\epsilon_i' x_i))' \tag{V.5}$$

$$+ \frac{1}{nT}\sum_i (\mathrm{vec}(x_i'\epsilon_i(\widehat{\beta}-\beta)' x_i' x_i))(\mathrm{vec}(x_i'\epsilon_i(\widehat{\beta}-\beta)' x_i' x_i))' \tag{V.6}$$

$$+ \frac{1}{nT}\sum_i (\mathrm{vec}(x_i'\epsilon_i(\widehat{\beta}-\beta)' x_i' x_i))(\mathrm{vec}(x_i' x_i(\widehat{\beta}-\beta)\epsilon_i' x_i))' \tag{V.7}$$

$$- \frac{1}{nT}\sum_i (\mathrm{vec}(x_i'\epsilon_i(\widehat{\beta}-\beta)' x_i' x_i))(\mathrm{vec}(x_i' x_i(\widehat{\beta}-\beta)(\widehat{\beta}-\beta)' x_i' x_i))' \tag{V.8}$$

$$- \frac{1}{nT}\sum_i (\mathrm{vec}(x_i' x_i(\widehat{\beta}-\beta)\epsilon_i' x_i))(\mathrm{vec}(x_i'\epsilon_i\epsilon_i' x_i))' \tag{V.9}$$

$$+\frac{1}{nT}\sum_i(\text{vec}(x_i'x_i(\widehat{\beta}-\beta)\epsilon_i'x_i))(\text{vec}(x_i'\epsilon_i(\widehat{\beta}-\beta)'x_i'x_i))' \tag{V.10}$$

$$+\frac{1}{nT}\sum_i(\text{vec}(x_i'x_i(\widehat{\beta}-\beta)\epsilon_i'x_i))(\text{vec}(x_i'x_i(\widehat{\beta}-\beta)\epsilon_i'x_i))' \tag{V.11}$$

$$-\frac{1}{nT}\sum_i(\text{vec}(x_i'x_i(\widehat{\beta}-\beta)\epsilon_i'x_i))(\text{vec}(x_i'x_i(\widehat{\beta}-\beta)(\widehat{\beta}-\beta)'x_i'x_i))' \tag{V.12}$$

$$+\frac{1}{nT}\sum_i(\text{vec}(x_i'x_i(\widehat{\beta}-\beta)(\widehat{\beta}-\beta)'x_i'x_i))(\text{vec}(x_i'\epsilon_i\epsilon_i'x_i))' \tag{V.13}$$

$$-\frac{1}{nT}\sum_i(\text{vec}(x_i'x_i(\widehat{\beta}-\beta)(\widehat{\beta}-\beta)'x_i'x_i))(\text{vec}(x_i'\epsilon_i(\widehat{\beta}-\beta)'x_i'x_i))' \tag{V.14}$$

$$-\frac{1}{nT}\sum_i(\text{vec}(x_i'x_i(\widehat{\beta}-\beta)(\widehat{\beta}-\beta)'x_i'x_i))(\text{vec}(x_i'x_i(\widehat{\beta}-\beta)\epsilon_i'x_i))' \tag{V.15}$$

$$+\frac{1}{nT}\sum_i(\text{vec}(x_i'x_i(\widehat{\beta}-\beta)(\widehat{\beta}-\beta)'x_i'x_i))(\text{vec}(x_i'x_i(\widehat{\beta}-\beta)(\widehat{\beta}-\beta)'x_i'x_i))'. \tag{V.16}$$

Theorems 1-4 then follow by examining the properties of equations (O.1)-(O.4) and (V.1)-(V.16).

## Appendix B. Proof of Theorem 1

**(i)** $\widehat{\beta}-\beta \xrightarrow{p} 0$ and $\sqrt{nT}(\widehat{\beta}-\beta) \xrightarrow{d} Q^{-1}N(0, W = \lim_n \frac{1}{nT}\sum_{i=1}^n \text{E}[x_i'\Omega_i x_i])$ follow immediately under the conditions of Theorem 1 from the Markov LLN and the Liapounov CLT.

We also have that $\text{E}\|x_i'x_i/T\|^{2+2\delta} \overset{(a)}{\leq} \text{E}\|x_i/\sqrt{T}\|^{4+4\delta} \overset{(b)}{\leq} \frac{[\sum_t\sum_h(\text{E}|x_{ith}^2|^{2+2\delta})^{1/(2+2\delta)}]^{2+2\delta}}{T^{2+2\delta}} \overset{(c)}{<} \frac{(kT)^{2+2\delta}\Delta}{T^{2+2\delta}}$, where (a) is by the Cauchy-Schwarz inequality, (b) follows from the definition of $\|A\|$ and Minkowski's inequality, and (c) is by $\text{E}|x_{ith}|^{4+\delta} < \Delta$. Making use of $\text{E}|\epsilon_{ith}|^{4+\delta} < \Delta$, $\text{E}\|x_i'\epsilon_i/T\|^{2+2\delta} < k^{2+2\delta}\Delta$ follows similarly. Noting that $\text{E}[x_i'x_i \otimes \epsilon_i'x_i] = 0$ by Assumption 3.a, it follows that terms (O.2)-(O.4) of $\widehat{W}$ are $o_p(n^{-1/2})$ by the Markov LLN.[1] The Markov LLN also yields $\frac{1}{nT}\sum_i x_i'\epsilon_i\epsilon_i'x_i \xrightarrow{p} W$, which implies $\widehat{W} \xrightarrow{p} W$.

**(ii)** In addition, under $\text{E}|x_{ith}|^{8+\delta} < \Delta$ and $\text{E}|\epsilon_{ith}|^{8+\delta} < \Delta$, $\text{E}\|x_i'\epsilon_i/T\|^{4+4\delta} < k^{4+4\delta}\Delta$ and $\text{E}\|x_i'x_i/T\|^{4+4\delta} < k^{4+4\delta}\Delta$ follow by an argument similar to that used to show $\text{E}\|x_i'x_i/T\|^{2+2\delta} < k^{2+2\delta}\Delta$. Then, using that (O.2)-(O.4) of $\widehat{W}$ are $o_p(n^{-1/2})$, it follows that

$$\sqrt{nT}[\text{vec}(\widehat{W}-W)] = \sqrt{nT}[\text{vec}(\frac{1}{nT}\sum_i x_i'\epsilon_i\epsilon_i'x_i - W)] + o_p(1) \xrightarrow{d} N(0,V)$$

where $V = \lim_n \frac{1}{nT}\sum_{i=1}^n \text{E}[(\text{vec}(x_i'\epsilon_i\epsilon_i'x_i - W))(\text{vec}(x_i'\epsilon_i\epsilon_i'x_i - W))']$ by the Liapounov CLT.

It is also straightforward to show that $\widehat{\beta}-\beta \xrightarrow{p} 0$, $\text{E}\|x_i'\epsilon_i/T\|^{4+4\delta} < k^{4+4\delta}\Delta$, and $\text{E}\|x_i'x_i/T\|^{4+4\delta} < k^{4+4\delta}\Delta$ imply that terms (V.2)-(V.16) of $\widehat{V}$ are $o_p(1)$ and that $\text{E}\|x_i'\epsilon_i/T\|^{4+4\delta} < k^{4+4\delta}\Delta$ implies

$$\frac{1}{nT}\sum_i[(\text{vec}(x_i'\epsilon_i\epsilon_i'x_i))(\text{vec}(x_i'\epsilon_i\epsilon_i'x_i))' - \text{E}[(\text{vec}(x_i'\epsilon_i\epsilon_i'x_i))(\text{vec}(x_i'\epsilon_i\epsilon_i'x_i))']] \xrightarrow{p} 0.$$

[1]Under Assumption 3.b, (O.2)-(O.4) will be $o_p(1)$ and consistency of $\widehat{W}$ follows.

We then have $\widehat{V} - V \xrightarrow{p} 0$. ■

## Appendix C. Proof of Theorem 2

**(i)** $\mathrm{E}\|x_i'x_i/T\|^{2+2\delta} \overset{(a)}{\leq} \mathrm{E}\|x_i/\sqrt{T}\|^{4+4\delta} \overset{(b)}{\leq} \frac{[\sum_t \sum_h (\mathrm{E}|x_{ith}^2|^{2+2\delta})^{1/(2+2\delta)}]^{2+2\delta}}{T^{2+2\delta}} \overset{(c)}{<} \frac{(kT)^{2+2\delta}\Delta}{T^{2+2\delta}}$, where (a) is by the Cauchy-Schwarz inequality, (b) follows from the definition of $\|A\|$ and Minkowski's inequality, and (c) is by $\mathrm{E}|x_{ith}|^{4+\delta} < \Delta$. Making use of $\mathrm{E}|\epsilon_{ith}|^{4+\delta} < \Delta$, $\mathrm{E}\|x_i'\epsilon_i/T\|^{2+2\delta} < k^{2+2\delta}\Delta$ follows similarly. Then Lemma A.2 gives $\frac{1}{n}\sum_i x_i'x_i/T \xrightarrow{p} Q$ and $\frac{1}{n}\sum_i x_i'\epsilon/T \xrightarrow{p} 0$ as $\{n,\ T\} \to \infty$ jointly, so $\widehat{\beta} - \beta \xrightarrow{p} 0$. In addition, since $\mathrm{E}\|x_i'\epsilon_i/T\|^{2+2\delta} < k^{2+2\delta}\Delta$, we have that $\frac{1}{\sqrt{n}}\sum_i x_i'\epsilon/T \xrightarrow{d} N(0, W = \lim_{n,T} \frac{1}{nT^2}\sum_{i=1}^n \mathrm{E}[x_i'\Omega_i x_i])$ by Lemma A.3, so $\sqrt{n}(\widehat{\beta} - \beta) \xrightarrow{d} Q^{-1}N(0, W)$.

Now consider term (O.2) of $\widehat{W}/T$:

$$\mathrm{vec}(\frac{1}{nT^2}\sum_i x_i'x_i(\widehat{\beta} - \beta)\epsilon_i'x_i) = \frac{1}{\sqrt{n}}(\frac{1}{nT^2}\sum_i x_i'\epsilon_i \otimes x_i'x_i)\sqrt{n}(\widehat{\beta} - \beta) = \frac{1}{\sqrt{n}}O_p(1)O_p(1)$$

by Lemma A.2 since $\mathrm{E}\|x_i'\epsilon_i/T \otimes x_i'x_i/T\|^{1+\delta} \leq (\mathrm{E}\|x_i'\epsilon_i/T\|^{2+2\delta}\mathrm{E}\|x_i'x_i/T\|^{2+2\delta})^{1/2} < C$ by Lemma A.1 and the argument in the preceding paragraph. That (O.3) is $O_p(1/\sqrt{n})$ and (O.4) is $O_p(1/n)$ follow similarly. Finally, $\mathrm{E}\|x_i'\epsilon_i\epsilon_i'x_i/T^2\|^{1+\delta} \leq \mathrm{E}\|x_i'\epsilon_i/T\|^{2+2\delta} < C$ where the first inequality follows from the Cauchy-Schwarz inequality and the second inequality is proven above. Hence, by Lemma A.2 $\frac{1}{nT^2}\sum_i \mathrm{vec}(x_i'\epsilon_i\epsilon_i'x_i - \mathrm{E}[x_i'\epsilon_i\epsilon_i'x_i]) = o_p(1)$. It follows immediately that $\widehat{W}/T - W \xrightarrow{p} 0$.

**(ii)** Under Assumption 3.a, we also have that $\frac{1}{\sqrt{n}}(\frac{1}{nT^2}\sum_i x_i'\epsilon_i \otimes x_i'x_i)\sqrt{n}(\widehat{\beta} - \beta) = \frac{1}{\sqrt{n}}o_p(1)O_p(1) = o_p(\frac{1}{\sqrt{n}})$ since $E[x_i'\epsilon_i \otimes x_i'x_i] = 0$. Similarly, $\frac{1}{\sqrt{n}}(\frac{1}{nT^2}\sum_i x_i'x_i \otimes \epsilon_i'x_i) = o_p(\frac{1}{\sqrt{n}})$. It follows from Lemma A.3 that $\sqrt{n}(\widehat{W}/T - W) = \sqrt{n}(\frac{1}{nT^2}\sum_i \mathrm{vec}(x_i'\epsilon_i\epsilon_i'x_i - \mathrm{E}[x_i'\epsilon_i\epsilon_i'x_i])) + o_p(1) \xrightarrow{d} N(0, V)$ where

$$V = \lim_{n,T} \frac{1}{nT^4}\sum_{i=1}^n \mathrm{E}[(\mathrm{vec}(x_i'\epsilon_i\epsilon_i'x_i - W))(\mathrm{vec}(x_i'\epsilon_i\epsilon_i'x_i - W))']$$

since $\mathrm{E}\|x_i'\epsilon_i\epsilon_i'x_i/T^2\|^{2+\delta} \leq \mathrm{E}\|x_i'\epsilon_i/T\|^{4+2\delta} < C$ by an argument similar to that used to show $\mathrm{E}\|x_i'\epsilon_i/T\|^{2+2\delta} < k^{2+2\delta}\Delta$ as long as $\mathrm{E}|x_{ith}|^{8+\delta} < \Delta$ and $\mathrm{E}|\epsilon_{ith}|^{8+\delta} < \Delta$.

To show $\widehat{V}/T^3 - V \xrightarrow{p} 0$, consider

$$\mathrm{E}\|\mathrm{vec}(x_i'\epsilon_i\epsilon_i'x_i/T^2)\mathrm{vec}(x_i'\epsilon_i\epsilon_i'x_i/T^2)'\|^{1+\delta} \leq \mathrm{E}\|x_i'\epsilon_i/T\|^{4+4\delta} < C$$

where the first inequality is by repeated application of the Cauchy-Schwarz inequality and the second is by an argument similar to that used to show $\mathrm{E}\|x_i'\epsilon_i/T\|^{2+2\delta} < k^{2+2\delta}\Delta$ which holds if $\mathrm{E}|x_{ith}|^{8+\delta} < \Delta$ and $\mathrm{E}|\epsilon_{ith}|^{8+\delta} < \Delta$. It then follows by Lemma A.2 that

$$\frac{1}{nT^4}\sum_i [\mathrm{vec}(x_i'\epsilon_i\epsilon_i'x_i)\mathrm{vec}(x_i'\epsilon_i\epsilon_i'x_i)' - \mathrm{E}[\mathrm{vec}(x_i'\epsilon_i\epsilon_i'x_i)\mathrm{vec}(x_i'\epsilon_i\epsilon_i'x_i)']] \xrightarrow{p} 0.$$

Turning to (V.2), we have

$$\mathrm{vec}[(\mathrm{vec}(x_i'\epsilon_i\epsilon_i'x_i/T^2))(\mathrm{vec}(x_i'\epsilon_i(\widehat{\beta} - \beta)'x_i'x_i/T^2))'] = [(x_i'\epsilon_i/T \otimes x_i'\epsilon_i/T) \otimes (x_i'x_i/T \otimes \epsilon_i'x_i/T)]\mathrm{vec}(\widehat{\beta} - \beta).$$

$\mathrm{E}\|(x_i'\epsilon_i/T \otimes x_i'\epsilon_i/T) \otimes (x_i'x_i/T \otimes \epsilon_i'x_i/T)\|^{1+\delta} \leq [(\mathrm{E}\|x_i'\epsilon_i/T\|^{4+4\delta})^3\mathrm{E}\|x_i'x_i\|^{4+4\delta}]^{1/4} < C$ where the first inequality is by repeated application of Lemma A.1 and the second inequality is by the moment conditions. It then follows from Lemma A.2 that $\frac{1}{n}\sum_i(x_i'\epsilon_i/T \otimes x_i'\epsilon_i/T) \otimes (x_i'x_i/T \otimes \epsilon_i'x_i/T) = O_p(1)$, so (V.2) is $o_p(1)$. Using similar arguments, it is also straightforward to show that terms (V.3)-(V.16) are $o_p(1)$, and the conclusion follows. ■

## Appendix D. Proof of Theorem 3

**(i)** $\mathrm{E}\|x_i'x_i/T\|^{2+2\delta} \overset{(a)}{\leq} \mathrm{E}\|x_i/\sqrt{T}\|^{4+4\delta} \overset{(b)}{\leq} \frac{[\sum_t\sum_h(\mathrm{E}|x_{ith}^2|^{2+2\delta})^{1/(2+2\delta)}]^{2+2\delta}}{T^{2+2\delta}} \overset{(c)}{<} \frac{(kT)^{2+2\delta}\Delta}{T^{2+2\delta}}$, where (a) is by the Cauchy-Schwarz inequality, (b) by definition of $\|A\|$ and Minkowski's inequality, and (c) by $\mathrm{E}|x_{ith}|^{r+\delta} < \Delta$. Also, $\mathrm{E}\|x_i'\epsilon_i/\sqrt{T}\|^{2+2\delta} < C$ by Lemma A.4, $\mathrm{E}|\epsilon_{it}|^{r+\delta} < \Delta$, and the mixing condition that $\alpha$ is of size $-3r/(r-4)$ for $r > 4$. It follows by Lemmas A.2 and A.3 that $\sqrt{nT}(\widehat{\beta}-\beta) \xrightarrow{d} Q^{-1}N(0, W = \lim_{n,T}\frac{1}{nT}\sum_{i=1}^n \mathrm{E}[x_i'\Xi_i x_i])$ as $\{n,T\} \to \infty$.

Now consider term (O.2) of $\widehat{W}$: $\mathrm{vec}(\frac{1}{nT}\sum_i x_i'x_i(\widehat{\beta}-\beta)\epsilon_i'x_i) = (\frac{1}{(nT)^{3/2}}\sum_i x_i'\epsilon_i \otimes x_i'x_i)\sqrt{nT}(\widehat{\beta}-\beta) = \frac{1}{\sqrt{n}}o_p(1)O_p(1)$ by Lemma A.2 since $\mathrm{E}[x_i'\epsilon_i \otimes x_i'x_i] = 0$ by Assumption 3.a and $\mathrm{E}\|x_i'\epsilon_i/\sqrt{T} \otimes x_i'x_i/T\|^{1+\delta} \leq (\mathrm{E}\|x_i'\epsilon_i/\sqrt{T}\|^{2+2\delta}\mathrm{E}\|x_i'x_i/T\|^{2+2\delta})^{1/2} < C$ by Lemma A.1 and the argument in the preceding paragraph. That (O.3) is $o_p(1/\sqrt{n})$ and (O.4) is $O_p(1/n)$ follow similarly.[2] Finally, $\mathrm{E}\|x_i'\epsilon_i\epsilon_i'x_i/T\|^{1+\delta}\| \leq \mathrm{E}\|x_i'\epsilon_i/\sqrt{T}\|^{2+2\delta} < C$ by the Cauchy-Schwarz inequality and the preceding argument. It then follows that $\widehat{W} - W \xrightarrow{p} 0$ by Lemma A.2.

**(ii)** In addition, under the mixing condition that $\alpha$ is of size $-7r/(r-8)$ for $r > 8$, $\mathrm{E}|x_{ith}|^{r+\delta} < \Delta$, and $\mathrm{E}|\epsilon_{ith}|^{r+\delta} < \Delta$, $\mathrm{E}\|x_i'\epsilon_i/\sqrt{T}\|^{4+4\delta} < C$ follows by an argument similar to that used to show $\mathrm{E}\|x_i'\epsilon_i/\sqrt{T}\|^{2+2\delta} < C$. Then, since $\mathrm{E}\|x_i'\epsilon_i\epsilon_i'x_i/T\|^{2+\delta} \leq \mathrm{E}\|x_i'\epsilon_i/\sqrt{T}\|^{4+4\delta} < C$ and using that (O.2)-(O.4) of $\widehat{W}$ are $o_p(n^{-1/2})$, the second conclusion follows from Lemma A.3 since

$$\sqrt{n}[\mathrm{vec}(\widehat{W} - W)] = \sqrt{n}[\mathrm{vec}(\frac{1}{nT}\sum_i x_i'\epsilon_i\epsilon_i'x_i - W)] + o_p(1) \xrightarrow{d} N(0, V)$$

where

$$V = \lim_{n,T}\frac{1}{nT^2}\sum_{i=1}^n \mathrm{E}[(\mathrm{vec}(x_i'\epsilon_i\epsilon_i'x_i - W))(\mathrm{vec}(x_i'\epsilon_i\epsilon_i'x_i - W))']$$

.

To show $\widehat{V}/T - V \xrightarrow{p} 0$, consider

$$\mathrm{E}\|\mathrm{vec}(x_i'\epsilon_i\epsilon_i'x_i/T)\mathrm{vec}(x_i'\epsilon_i\epsilon_i'x_i/T)'\|^{1+\delta} \leq \mathrm{E}\|x_i'\epsilon_i/\sqrt{T}\|^{4+4\delta} < C$$

where the first inequality is by repeated application of the Cauchy-Schwarz inequality and the second is by an argument similar to that used above which holds if $\mathrm{E}|\epsilon_{it}|^{8+\delta} < \Delta < \infty$, $\mathrm{E}|x_{ith}|^{8+\delta} < \Delta < \infty$, and the strong mixing coefficient $\alpha$ is of size $-7r/(r-8)$ for $r > 8$. It then follows from Lemma A.2 that

$$\frac{1}{nT^2}\sum_i[\mathrm{vec}(x_i'\epsilon_i\epsilon_i'x_i)\mathrm{vec}(x_i'\epsilon_i\epsilon_i'x_i)' - \mathrm{E}[\mathrm{vec}(x_i'\epsilon_i\epsilon_i'x_i)\mathrm{vec}(x_i'\epsilon_i\epsilon_i'x_i)']] \xrightarrow{p} 0.$$

[2]Under Assumption 3.b, (O.2)-(O.4) will be $o_p(1)$ and consistency of $\widehat{W}$ follows.

Turning to (V.2), we have

$$\frac{1}{n^{3/2}}\sum_i \text{vec}[(\text{vec}(x_i'\epsilon_i\epsilon_i'x_i/T))(\text{vec}(x_i'\epsilon_i(\sqrt{nT}(\widehat{\beta}-\beta))'x_i'x_i/T^{3/2}))']$$

$$= \frac{1}{n^{3/2}}\sum_i [(x_i'\epsilon_i/\sqrt{T}\otimes x_i'\epsilon_i/\sqrt{T})\otimes(x_i'x_i/T\otimes\epsilon_i'x_i/\sqrt{T})]\text{vec}(\sqrt{nT}(\widehat{\beta}-\beta)).$$

By Lemma A.2, we have that $\frac{1}{n}\sum_i(x_i'\epsilon_i/\sqrt{T}\otimes x_i'\epsilon_i/\sqrt{T})\otimes(x_i'x_i/T\otimes\epsilon_i'x_i/\sqrt{T}) = O_p(1)$ since

$$\text{E}\|(x_i'\epsilon_i/\sqrt{T}\otimes x_i'\epsilon_i/\sqrt{T})\otimes(x_i'x_i/T\otimes\epsilon_i'x_i/\sqrt{T})\|^{1+\delta} \leq [(\text{E}\|x_i'\epsilon_i/\sqrt{T}\|^{4+4\delta})^3\text{E}\|x_i'x_i\|^{4+4\delta}]^{1/4} < C$$

where the first inequality is by repeated application of Lemma A.1 and the second inequality is by the moment and mixing conditions. It then follows that (V.2) is $O_p(1/\sqrt{n})$. It can similarly be shown that terms (V.3)-(V.16) are $O_p(1/\sqrt{n})$, and the conclusion follows. ■

## Appendix E. Proof of Theorem 4

Under the hypotheses of the theorem, $\sqrt{n}(\widehat{\beta}-\beta)\xrightarrow{d}Q^{-1}N(0,W)$, $x_i'x_i/T-Q_i\xrightarrow{p}0$, and $x_i'\epsilon_i/\sqrt{T}\xrightarrow{d}N(0,W_i)$ are immediate from a LLN and CLT for mixing sequences, cf. White (2001) Theorems 3.47 and 5.20. The conclusion then follows from the definition of $\widehat{W}$ and $\widehat{\epsilon}_i$.

## Appendix F. Proof of Corollary 4.1

Consider $t^* = \frac{\sqrt{nT}(R\widehat{\beta}-r)}{\sqrt{R\widehat{Q}^{-1}\widehat{W}\widehat{Q}^{-1}R'}}$. Under the null hypothesis, $R\beta = r$, so the numerator of $t*$ is $\sqrt{nT}R(\widehat{\beta}-\beta) = R(\frac{1}{nT}\sum_i x_i'x_i)^{-1}(\frac{1}{\sqrt{nT}}\sum_i x_i'\epsilon_i)\xrightarrow{d}RQ^{-1}\Lambda\sum_i B_i/\sqrt{n}$. From Theorem 4 and the hypotheses of the Corollary, the denominator of $t^*$ converges in distribution to $\sqrt{RQ^{-1}\frac{1}{n}\Lambda(\sum_{i=1}^n B_iB_i'-\frac{1}{n}\sum_{i=1}^n B_i\sum_{i=1}^n B_i')\Lambda Q^{-1}R'}$. It follows from the Continuous Mapping Theorem that

$$t^*\xrightarrow{d}\frac{RQ^{-1}\Lambda\sum_i B_i/\sqrt{n}}{\sqrt{\frac{1}{n}RQ^{-1}\Lambda(\sum_{i=1}^n B_iB_i'-\frac{1}{n}\sum_{i=1}^n B_i\sum_{i=1}^n B_i')\Lambda Q^{-1}R'}}.$$

Define $\delta = (RQ^{-1}\Lambda\Lambda Q^{-1}R')^{1/2}$, so

$$t^*\xrightarrow{d}U = \frac{\delta\sum_i B_{1,i}/\sqrt{n}}{\sqrt{\frac{\delta^2}{n}(\sum_{i=1}^n B_{1,i}B_{1,i}'-\frac{1}{n}\sum_{i=1}^n B_{1,i}\sum_{i=1}^n B_{1,i}')}}$$

$$= \frac{\widetilde{B}_{1,n}}{\sqrt{\frac{1}{n}(\sum_i B_{1,i}^2-\widetilde{B}_{1,n}^2)}}.$$

It is straightforward to show that $\widetilde{B}_{1,n}\sim N(0,1)$, that $\sum_i B_{1,i}^2-\widetilde{B}_{1,n}^2\sim\chi_{n-1}^2$, and that $\sum_i B_{1,i}^2-\widetilde{B}_{1,n}^2$ and $\widetilde{B}_{1,n}$ are independent, from which it follows that

$$U = \left(\frac{n}{n-1}\right)^{1/2}\frac{\widetilde{B}_{1,n}}{\sqrt{(\sum_i B_{1,i}^2-\widetilde{B}_{1,n}^2)/(n-1)}}\sim\left(\frac{n}{n-1}\right)^{1/2}t_{n-1}.$$

The result for $F^*$ is obtained through a similar argument. ■

## Appendix G. Proof of Lemma A.2 and Lemma A.3

**Proof of Lemma A.2** Instead of proving Lemma A.2 directly, we prove the following: If $\{Z_{i,T}\}$ are independent across $i$ for all $T$ with $\mathrm{E}[Z_{i,T}] = 0$ and $\mathrm{E}|Z_{i,T}|^{1+\delta} < \Delta < \infty$ for some $\delta > 0$ and all $i$, $T$, $\frac{1}{n}\sum_i Z_{i,T} \xrightarrow{p} 0$ as $\{n, T\} \to \infty$ jointly. Lemma A.2 then follows immediately from an argument similar to that used to prove White (2001) Corollary 3.9.

Define $Y_{i,T} = Z_{i,T}1(|Z_{i,T}| \leq n)$. Then $Var(\sum_i Y_{i,T}/n) = \sum_i Var(Y_{i,T}/n)$ by independence of $Z_{i,T}$ across $i$. Now $\sum_i Var(Y_{i,T}/n) \leq \mathrm{E}(Y_{i,T}/n)^2 = \sum_i \int_{|Z|\leq n}(Z^2/n^2)dF_{i,T}(Z)$ where $F_{i,T}$ is the distribution function of $Z_{i,T}$. $Z^2/n^2 \leq Z^{1+\delta}/n^{1+\delta}$ for $|Z| \leq n$ implies

$$\sum_i \int_{|Z|\leq n}(Z^2/n^2)dF_{i,T}(Z) \leq \sum_i \int_{|Z|\leq n}(Z^{1+\delta}/n^{1+\delta})dF_{i,T}(Z) \leq \sum_i \int (Z^{1+\delta}/n^{1+\delta})dF_{i,T}(Z) < \Delta/n^\delta$$

where the last inequality results from $\mathrm{E}|Z_{i,T}|^{1+\delta} < \Delta$. It follows that

$$Var(\sum_i Y_{i,T}/n) < \Delta/n^\delta \to 0 \tag{1}$$

as $\{n,\ T\} \to \infty$ jointly.

Now consider

$$\begin{aligned}
|\mathrm{E}\frac{1}{n}\sum_i Y_{i,T}| &= |\sum_i \int_{|Z|\leq n}(Z/n)dF_{i,T}(Z)| \\
&= |\sum_i \int (Z/n)dF_{i,T}(Z) - \sum_i \int_{|Z|>n}(Z/n)dF_{i,T}(Z)| \\
&= |\sum_i \int_{|Z|>n}(Z/n)dF_{i,T}(Z)| \\
&\leq \sum_i \int_{|Z|>n}(|Z|/n)dF_{i,T}(Z)
\end{aligned}$$

by the Triangle inequality and Jensen's inequality. For $|Z| > n$, $|Z|/n \leq |Z|^{1+\delta}/n^{1+\delta}$, so

$$\sum_i \int_{|Z|>n}\frac{|Z|}{n}dF_{i,T}(Z) \leq \sum_i \int_{|Z|>n}\frac{|Z|^{1+\delta}}{n^{1+\delta}}dF_{i,T}(Z) \leq \sum_i \int \frac{|Z|^{1+\delta}}{n^{1+\delta}}dF_{i,T}(Z) \leq \Delta/n^\delta \to 0, \tag{2}$$

which yields

$$|\mathrm{E}\frac{1}{n}\sum_i Y_{i,T}| \to 0. \tag{3}$$

By Chebyshev's inequality and (1),

$$\lim_{n,T} P(|\sum_i (Y_{i,T} - \mathrm{E}[Y_{i,T}])/n| \geq \epsilon) \leq \lim_{n,T} Var(\sum_i Y_{i,T}/n)/\epsilon^2 = 0,$$

so $\frac{1}{n}\sum_i Y_{i,T} - \frac{1}{n}\sum_i \mathrm{E}[Y_{i,T}] \xrightarrow{p} 0$, which implies, using (3), that $\frac{1}{n}\sum_i Y_{i,T} \xrightarrow{p} 0$.

Finally, consider

$$\begin{aligned} P(|\frac{1}{n}\sum_i Z_{i,T} - \frac{1}{n}\sum_i Y_{i,T}| \geq \epsilon) &= P(|\frac{1}{n}\sum_i (1 - 1(|Z_{i,T}| \leq n))Z_{i,T}| \geq \epsilon) \\ &\leq \mathrm{E}|\frac{1}{n}\sum_i (1 - 1(|Z_{i,T}| \leq n))Z_{i,T}|/\epsilon \end{aligned}$$

by the Markov inequality. $\mathrm{E}|\frac{1}{n}\sum_i (1 - 1(|Z_{i,T}| \leq n))Z_{i,T}| \leq \frac{1}{n}\sum_i \mathrm{E}|(1 - 1(|Z_{i,T}| \leq n))Z_{i,T}|$ by the Triangle inequality, and

$$\begin{aligned} \frac{1}{n}\sum_i \mathrm{E}|(1 - 1(|Z_{i,T}| \leq n))Z_{i,T}| &= \frac{1}{n}\sum_i [\int |Z| dF_{i,T}(Z) - \int_{|Z|\leq n} |Z| dF_{i,T}(Z) \\ &= \sum_i \int_{|Z|>n} \frac{|Z|}{n} dF_{i,T}(Z) \to 0 \end{aligned}$$

by (2). It then follows that $\frac{1}{n}\sum_i Z_{i,T} - \frac{1}{n}\sum_i Y_{i,T} \xrightarrow{p} 0$ which, with $\frac{1}{n}\sum_i Y_{i,T} \xrightarrow{p} 0$, implies $\frac{1}{n}\sum_i Z_{i,T} \xrightarrow{p} 0$. ■

**Proof of Lemma A.3** Define $\xi_{i,n,T} = \Omega_{n,T}^{-1/2} Z_{i,T}$ where $\Omega_{n,T} = \sum_i \Omega_{i,T}$. By the Cramer-Wold device, $\sum_i \xi_{i,n,T} \xrightarrow{d} N(0, I_k)$ as $\{n,\ T\} \to \infty$ jointly if $\forall\, c \in \mathbb{R}^k$ with $\|c\| = 1$, $c' \sum_i \xi_{i,n,T} \xrightarrow{d} N(0,1)$ as $\{n,\ T\} \to \infty$ jointly. Then, by $\Omega = \lim_{n,T} \frac{1}{n}\sum_{i=1}^n \Omega_{i,T} > 0$, $\frac{1}{\sqrt{n}}\sum_{i=1}^n Z_{i,T} \xrightarrow{d} N(0, \Omega)$ as $\{n,\ T\} \to \infty$ jointly.

To establish $c' \sum_i \xi_{i,n,T} \xrightarrow{d} N(0,1)$, it is sufficient to verify

$$\lim_{n,T} \sum_i \mathrm{E}[\xi_{i,n,T}^2 1(|\xi_{i,n,T}^2 > \epsilon)] = 0. \tag{4}$$

For a given $\epsilon > 0$ and $c \in \mathbb{R}^k$ with $\|c\| = 1$,

$$\begin{aligned} & c' \sum_i \mathrm{E}[\xi_{i,n,T} \xi_{i,n,T}' 1(|c' \xi_{i,n,T} \xi_{i,n,T}' c| > \epsilon)] c \\ & \qquad = c' \Omega_{n,T}^{-1/2} \sum_i \mathrm{E}[Z_{i,T} Z_{i,T}' 1(|c' \Omega_{n,T}^{-1/2} Z_{i,T} Z_{i,T}' \Omega_{n,T}^{-1/2} c| > \epsilon)] \Omega_{n,T}^{-1/2} c. \end{aligned}$$

Considering first the indicator function, we have

$$\begin{aligned} 1(|c' \Omega_{n,T}^{-1/2} Z_{i,T} Z_{i,T}' \Omega_{n,T}^{-1/2} c| > \epsilon) &\leq 1(\|c\| \|\Omega_{n,T}^{-1/2} Z_{i,T} Z_{i,T}' \Omega_{n,T}^{-1/2}\| > \epsilon) \\ &\leq 1(\lambda_{\max}(\Omega_{n,T}^{-1}) \|Z_{i,T}\|^2 > \epsilon) \\ &= 1(\|Z_{i,T}\|^2 > \epsilon \lambda_{\min}(\Omega_{n,T})) \end{aligned}$$

Then

$$
\begin{aligned}
c'\Omega_{n,T}^{-1/2}\sum_i \mathrm{E}[Z_{i,T}Z_{i,T}'1(|c'\Omega_{n,T}^{-1/2}Z_{i,T}Z_{i,T}'\Omega_{n,T}^{-1/2}c|>\epsilon)]\Omega_{n,T}^{-1/2}c \\
&\leq \|c\|^2\|\Omega_{n,T}^{-1/2}\sum_i \mathrm{E}[Z_{i,T}Z_{i,T}'1(|c'\Omega_{n,T}^{-1/2}Z_{i,T}Z_{i,T}'\Omega_{n,T}^{-1/2}c|>\epsilon)]\Omega_{n,T}^{-1/2}\| \\
&\leq \lambda_{\max}(\Omega_{n,T}^{-1})\|\sum_i \mathrm{E}[Z_{i,T}Z_{i,T}'1(\|Z_{i,T}\|^2>\epsilon\lambda_{\min}(\Omega_{n,T}))]\| \\
&\leq \frac{1}{\lambda_{\min}(\Omega_{n,T})}\sum_i \mathrm{E}[\|Z_{i,T}\|^2 1(\|Z_{i,T}\|^2>\epsilon\lambda_{\min}(\Omega_{n,T}))] \\
&\leq \frac{1}{\lambda_{\min}(\Omega_{n,T})}\sum_i \frac{\mathrm{E}\|Z_{i,T}\|^{2+\delta}}{(\epsilon\lambda_{\min}(\Omega_{n,T}))^\delta} \\
&\leq \frac{n\Delta}{\epsilon^\delta[n\lambda_{\min}(\frac{1}{n}\sum_i\Omega_{i,T})]^{1+\delta}} = \frac{\Delta}{\epsilon^\delta n^\delta \lambda_{\min}^{1+\delta}} \to 0
\end{aligned}
$$

as $\{n,\ T\}\to\infty$ jointly, and it follows that (4) $\to 0$ as $\{n,\ T\}\to\infty$ jointly.

## References


Chung, K. L. (2001): *A Course in Probability Theory.* San Diego, California: Academic Press, third edn.

Doukhan, P. (1994): *Mixing: Properties and Examples*, vol. 85 of *Lecture Notes in Statistics (Springer-Verlag).* New York: Springer-Verlag.

Phillips, P. C. B., and H. R. Moon (1999): "Linear Regression Limit Theory for Nonstationary Panel Data," *Econometrica*, 67(5), 1057–1111.

White, H. (2001): *Asymptotic Theory for Econometricians.* San Diego: Academic Press, revised edn.