*Econometrica Supplementary Material*

# SUPPLEMENT TO "POWER ENHANCEMENT IN HIGH-DIMENSIONAL CROSS-SECTIONAL TESTS"

(*Econometrica*, Vol. 83, No. 4, July 2015, 1497–1541)

BY JIANQING FAN, YUAN LIAO, AND JIAWEI YAO

This supplement contains additional proofs of the main paper.

## APPENDIX D: AUXILIARY LEMMAS FOR THE PROOF OF PROPOSITION 4.2

DEFINE $\mathbf{e}_t = \boldsymbol{\Sigma}_u^{-1}\mathbf{u}_t = (e_{1t}, \dots, e_{Nt})'$, which is an $N$-dimensional vector with mean zero and covariance $\boldsymbol{\Sigma}_u^{-1}$, whose entries are stochastically bounded. Let $\bar{\mathbf{w}} = (E\mathbf{f}_t\mathbf{f}_t')^{-1}E\mathbf{f}_t$. Also recall that

$$a_1 = \frac{T}{\sqrt{N}}\sum_{i=1}^{N}\big(\widehat{\boldsymbol{\theta}}'\boldsymbol{\Sigma}_u^{-1}\big)_i^2(\widehat{\sigma}_{ii} - \sigma_{ii}),$$

$$a_2 = \frac{T}{\sqrt{N}}\sum_{i\neq j,(i,j)\in S_U}\big(\widehat{\boldsymbol{\theta}}'\boldsymbol{\Sigma}_u^{-1}\big)_i\big(\widehat{\boldsymbol{\theta}}'\boldsymbol{\Sigma}_u^{-1}\big)_j(\widehat{\sigma}_{ij} - \sigma_{ij}).$$

One of the key steps of proving $a_1 = o_P(1)$, $a_2 = o_P(1)$ is to establish the following two convergences:

$$\text{(D.1)}\qquad \frac{1}{T}E\left|\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{t=1}^{T}\big(u_{it}^2 - Eu_{it}^2\big)\left(\frac{1}{\sqrt{T}}\sum_{s=1}^{T}e_{is}\big(1 - \mathbf{f}_s'\bar{\mathbf{w}}\big)\right)^2\right|^2 = o(1),$$

$$\text{(D.2)}\qquad \frac{1}{T}E\left|\frac{1}{\sqrt{NT}}\sum_{i\neq j,(i,j)\in S_U}\sum_{t=1}^{T}(u_{it}u_{jt} - Eu_{it}u_{jt}) \times\left[\frac{1}{\sqrt{T}}\sum_{s=1}^{T}e_{is}\big(1 - \mathbf{f}_s'\bar{\mathbf{w}}\big)\right]\left[\frac{1}{\sqrt{T}}\sum_{k=1}^{T}e_{jk}\big(1 - \mathbf{f}_k'\bar{\mathbf{w}}\big)\right]\right|^2 = o(1),$$

where $S_U = \{(i,j) : (\boldsymbol{\Sigma}_u)_{ij} \neq 0\}$. The proofs of (D.1) and (D.2) are given later below.

LEMMA D.1: *Under $H_0$, $a_1 = o_P(1)$.*

 DOI: 10.3982/ECTA12749

PROOF: We have $a_1 = \frac{T}{\sqrt{N}} \sum_{i=1}^N (\widehat{\boldsymbol\theta}' \boldsymbol\Sigma_u^{-1})_i^2 \frac{1}{T} \sum_{t=1}^T (\widehat{u}_{it}^2 - Eu_{it}^2)$, which is

$$\begin{aligned}&\frac{T}{\sqrt{N}} \sum_{i=1}^N \big(\widehat{\boldsymbol\theta}' \boldsymbol\Sigma_u^{-1}\big)_i^2 \frac{1}{T} \sum_{t=1}^T (\widehat{u}_{it}^2 - u_{it}^2)\\ &\quad + \frac{T}{\sqrt{N}} \sum_{i=1}^N \big(\widehat{\boldsymbol\theta}' \boldsymbol\Sigma_u^{-1}\big)_i^2 \frac{1}{T} \sum_{t=1}^T \big(u_{it}^2 - Eu_{it}^2\big)\\ &= a_{11} + a_{12}.\end{aligned}$$

For $a_{12}$, note that

$$\big(\widehat{\boldsymbol\theta}' \boldsymbol\Sigma_u^{-1}\big)_i = \big(1 - \bar{\mathbf{f}}'\mathbf{w}\big)^{-1} \frac{1}{T} \sum_{s=1}^T \big(1 - \mathbf{f}_s'\mathbf{w}\big)\big(\mathbf{u}_s' \boldsymbol\Sigma_u^{-1}\big)_i = c \frac{1}{T} \sum_{s=1}^T \big(1 - \mathbf{f}_s'\mathbf{w}\big) e_{is},$$

where $c = (1 - \bar{\mathbf{f}}'\mathbf{w})^{-1} = O_P(1)$. Hence

$$a_{12} = \frac{Tc}{\sqrt{N}} \sum_{i=1}^N \left( \frac{1}{T} \sum_{s=1}^T \big(1 - \mathbf{f}_s'\mathbf{w}\big) e_{is} \right)^2 \frac{1}{T} \sum_{t=1}^T \big(u_{it}^2 - Eu_{it}^2\big).$$

By (D.1), $Ea_{12}^2 = o(1)$. On the other hand,

$$\begin{aligned}a_{11} &= \frac{T}{\sqrt{N}} \sum_{i=1}^N \big(\widehat{\boldsymbol\theta}' \boldsymbol\Sigma_u^{-1}\big)_i^2 \frac{1}{T} \sum_{t=1}^T (\widehat{u}_{it} - u_{it})^2\\ &\quad + \frac{2T}{\sqrt{N}} \sum_{i=1}^N \big(\widehat{\boldsymbol\theta}' \boldsymbol\Sigma_u^{-1}\big)_i^2 \frac{1}{T} \sum_{t=1}^T u_{it} (\widehat{u}_{it} - u_{it})\\ &= a_{111} + a_{112}.\end{aligned}$$

Note that $\max_{i \le N} \frac{1}{T} \sum_{t=1}^T (\widehat{u}_{it} - u_{it})^2 = O_P(\frac{\log N}{T})$ by Lemma 3.1 of Fan, Liao, and Mincheva (2011). Since $\|\widehat{\boldsymbol\theta}\|^2 = O_P(\frac{N \log N}{T})$, $\|\boldsymbol\Sigma_u^{-1}\|_2 = O(1)$, and $N(\log N)^3 = o(T^2)$,

$$a_{111} \le O_P\left(\frac{\log N}{T}\right) \frac{T}{\sqrt{N}} \big\|\widehat{\boldsymbol\theta}' \boldsymbol\Sigma_u^{-1}\big\|^2 = O_P\left(\frac{(\log N)^2 \sqrt{N}}{T}\right) = o_P(1).$$

To bound $a_{112}$, note that

$$\begin{aligned}&\widehat{u}_{it} - u_{it} = \widehat{\theta}_i - \theta_i + (\widehat{\mathbf{b}}_i - \mathbf{b}_i)' \mathbf{f}_t,\\ &\max_i |\widehat{\theta}_i - \theta_i| = O_P\left(\sqrt{\frac{\log N}{T}}\right) = \max_i \|\widehat{\mathbf{b}}_i - \mathbf{b}_i\|.\end{aligned}$$

Also, $\max_i |\frac{1}{T}\sum_{t=1}^T u_{it}| = O_P(\sqrt{\frac{\log N}{T}}) = \max_i \|\frac{1}{T}\sum_{t=1}^T u_{it}\mathbf{f}_t\|$. Hence

$$a_{112} = \frac{2T}{\sqrt{N}}\sum_{i=1}^N (\widehat{\boldsymbol{\theta}}'\boldsymbol{\Sigma}_u^{-1})_i^2 \frac{1}{T}\sum_{t=1}^T u_{it}(\widehat{\theta}_i - \theta_i) + \frac{2T}{\sqrt{N}}\sum_{i=1}^N (\widehat{\boldsymbol{\theta}}'\boldsymbol{\Sigma}_u^{-1})_i^2 (\widehat{\mathbf{b}}_i - \mathbf{b}_i)'\frac{1}{T}\sum_{t=1}^T \mathbf{f}_t u_{it} \leq O_P\left(\frac{\log N}{\sqrt{N}}\right)\|\widehat{\boldsymbol{\theta}}'\boldsymbol{\Sigma}_u^{-1}\|^2 = o_P(1).$$

In summary, $a_1 = a_{12} + a_{111} + a_{112} = o_P(1)$. *Q.E.D.*

LEMMA D.2: *Under $H_0$, $a_2 = o_P(1)$.*

PROOF: We have $a_2 = \frac{T}{\sqrt{N}}\sum_{i\neq j,(i,j)\in S_U}(\widehat{\boldsymbol{\theta}}'\boldsymbol{\Sigma}_u^{-1})_i(\widehat{\boldsymbol{\theta}}'\boldsymbol{\Sigma}_u^{-1})_j\frac{1}{T}\sum_{t=1}^T(\widehat{u}_{it}\widehat{u}_{jt} - Eu_{it}u_{jt})$, which is

$$\frac{T}{\sqrt{N}}\sum_{i\neq j,(i,j)\in S_U}(\widehat{\boldsymbol{\theta}}'\boldsymbol{\Sigma}_u^{-1})_i(\widehat{\boldsymbol{\theta}}'\boldsymbol{\Sigma}_u^{-1})_j \times \left(\frac{1}{T}\sum_{t=1}^T(\widehat{u}_{it}\widehat{u}_{jt} - u_{it}u_{jt}) + \frac{1}{T}\sum_{t=1}^T(u_{it}u_{jt} - Eu_{it}u_{jt})\right) = a_{21} + a_{22},$$

where

$$a_{21} = \frac{T}{\sqrt{N}}\sum_{i\neq j,(i,j)\in S_U}(\widehat{\boldsymbol{\theta}}'\boldsymbol{\Sigma}_u^{-1})_i(\widehat{\boldsymbol{\theta}}'\boldsymbol{\Sigma}_u^{-1})_j\frac{1}{T}\sum_{t=1}^T(\widehat{u}_{it}\widehat{u}_{jt} - u_{it}u_{jt}).$$

Under $H_0$, $\boldsymbol{\Sigma}_u^{-1}\widehat{\boldsymbol{\theta}} = \frac{1}{T}(1 - \bar{\mathbf{f}}'\mathbf{w})^{-1}\sum_{t=1}^T \boldsymbol{\Sigma}_u^{-1}\mathbf{u}_t(1 - \mathbf{f}_t'\mathbf{w})$, and $\mathbf{e}_t = \boldsymbol{\Sigma}_u^{-1}\mathbf{u}_t$, we have

$$a_{22} = \frac{T}{\sqrt{N}}\sum_{i\neq j,(i,j)\in S_U}(\widehat{\boldsymbol{\theta}}'\boldsymbol{\Sigma}_u^{-1})_i(\widehat{\boldsymbol{\theta}}'\boldsymbol{\Sigma}_u^{-1})_j\frac{1}{T}\sum_{t=1}^T(u_{it}u_{jt} - Eu_{it}u_{jt}) = \frac{Tc}{\sqrt{N}}\sum_{i\neq j,(i,j)\in S_U}\frac{1}{T}\sum_{s=1}^T(1 - \mathbf{f}_s'\mathbf{w})e_{is}\frac{1}{T}\sum_{k=1}^T(1 - \mathbf{f}_k'\mathbf{w})e_{jk} \times \frac{1}{T}\sum_{t=1}^T(u_{it}u_{jt} - Eu_{it}u_{jt}).$$

By (D.2), $Ea_{22}^2 = o(1)$.

On the other hand, $a_{21} = a_{211} + a_{212}$, where

$$a_{211} = \frac{T}{\sqrt{N}} \sum_{i \neq j, (i,j) \in S_U} \big(\widehat{\boldsymbol{\theta}}' \boldsymbol{\Sigma}_u^{-1}\big)_i \big(\widehat{\boldsymbol{\theta}}' \boldsymbol{\Sigma}_u^{-1}\big)_j \frac{1}{T} \sum_{t=1}^{T} (\widehat{u}_{it} - u_{it})(\widehat{u}_{jt} - u_{jt}),$$

$$a_{212} = \frac{2T}{\sqrt{N}} \sum_{i \neq j, (i,j) \in S_U} \big(\widehat{\boldsymbol{\theta}}' \boldsymbol{\Sigma}_u^{-1}\big)_i \big(\widehat{\boldsymbol{\theta}}' \boldsymbol{\Sigma}_u^{-1}\big)_j \frac{1}{T} \sum_{t=1}^{T} u_{it}(\widehat{u}_{jt} - u_{jt}).$$

By the Cauchy–Schwarz inequality, $\max_{ij} |\frac{1}{T} \sum_{t=1}^{T} (\widehat{u}_{it} - u_{it})(\widehat{u}_{jt} - u_{jt})| = O_P(\frac{\log N}{T})$. Hence

$$\begin{aligned}
|a_{211}| &\leq O_P\bigg(\frac{\log N}{\sqrt{N}}\bigg) \sum_{i \neq j, (i,j) \in S_U} \big|\big(\widehat{\boldsymbol{\theta}}' \boldsymbol{\Sigma}_u^{-1}\big)_i\big|\big|\big(\widehat{\boldsymbol{\theta}}' \boldsymbol{\Sigma}_u^{-1}\big)_j\big| \\
&\leq O_P\bigg(\frac{\log N}{\sqrt{N}}\bigg) \bigg(\sum_{i \neq j, (i,j) \in S_U} \big(\widehat{\boldsymbol{\theta}}' \boldsymbol{\Sigma}_u^{-1}\big)_i^2\bigg)^{1/2} \bigg(\sum_{i \neq j, (i,j) \in S_U} \big(\widehat{\boldsymbol{\theta}}' \boldsymbol{\Sigma}_u^{-1}\big)_j^2\bigg)^{1/2} \\
&= O_P\bigg(\frac{\log N}{\sqrt{N}}\bigg) \sum_{i=1}^{N} \big(\widehat{\boldsymbol{\theta}}' \boldsymbol{\Sigma}_u^{-1}\big)_i^2 \sum_{j : (\boldsymbol{\Sigma}_u)_{ij} \neq 0} 1 \\
&\leq O_P\bigg(\frac{\log N}{\sqrt{N}}\bigg) \big\|\widehat{\boldsymbol{\theta}}' \boldsymbol{\Sigma}_u^{-1}\big\|^2 m_N \\
&= O_P\bigg(\frac{m_N \sqrt{N} (\log N)^2}{T}\bigg) = o_P(1).
\end{aligned}$$

Similarly to the proof of term $a_{112}$ in Lemma D.1, $\max_{ij} |\frac{1}{T} \sum_{t=1}^{T} u_{it}(\widehat{u}_{jt} - u_{jt})| = O_P(\frac{\log N}{T})$,

$$\begin{aligned}
|a_{212}| &\leq O_P\bigg(\frac{\log N}{\sqrt{N}}\bigg) \sum_{i \neq j, (i,j) \in S_U} \big|\big(\widehat{\boldsymbol{\theta}}' \boldsymbol{\Sigma}_u^{-1}\big)_i\big|\big|\big(\widehat{\boldsymbol{\theta}}' \boldsymbol{\Sigma}_u^{-1}\big)_j\big| \\
&= O_P\bigg(\frac{m_N \sqrt{N} (\log N)^2}{T}\bigg) = o_P(1).
\end{aligned}$$

In summary, $a_2 = a_{22} + a_{211} + a_{212} = o_P(1)$. *Q.E.D.*

### D.1. *Proof of (D.1) and (D.2)*

For any index set $A$, we let $|A|_0$ denote its number of elements.

LEMMA D.3: *Recall that* $\mathbf{e}_t = \boldsymbol{\Sigma}_u^{-1} \mathbf{u}_t$. $e_{it}$ *and* $u_{jt}$ *are independent if* $i \neq j$.

PROOF: Because $\mathbf{u}_t$ is Gaussian, it suffices to show that $\mathrm{cov}(e_{it}, u_{jt}) = 0$ when $i \neq j$. Consider the vector $(\mathbf{u}_t', \mathbf{e}_t')' = \mathbf{A}(\mathbf{u}_t', \mathbf{u}_t')'$, where

$$\mathbf{A} = \begin{pmatrix} \mathbf{I}_N & 0 \\ 0 & \boldsymbol{\Sigma}_u^{-1} \end{pmatrix}.$$

Then $\mathrm{cov}(\mathbf{u}_t', \mathbf{e}_t') = \mathbf{A}\,\mathrm{cov}(\mathbf{u}_t', \mathbf{u}_t')\mathbf{A}$, which is

$$\begin{pmatrix} \mathbf{I}_N & 0 \\ 0 & \boldsymbol{\Sigma}_u^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{\Sigma}_u & \boldsymbol{\Sigma}_u \\ \boldsymbol{\Sigma}_u & \boldsymbol{\Sigma}_u \end{pmatrix} \begin{pmatrix} \mathbf{I}_N & 0 \\ 0 & \boldsymbol{\Sigma}_u^{-1} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Sigma}_u & \mathbf{I}_N \\ \mathbf{I}_N & \boldsymbol{\Sigma}_u^{-1} \end{pmatrix}.$$

This completes the proof. *Q.E.D.*

PROOF OF (D.1): Let $X = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T} (u_{it}^2 - E u_{it}^2)(\frac{1}{\sqrt{T}} \sum_{s=1}^{T} e_{is}(1 - \mathbf{f}_s'\mathbf{w}))^2$. The goal is to show $EX^2 = o(T)$. We show respectively $\frac{1}{T}(EX)^2 = o(1)$ and $\frac{1}{T}\mathrm{var}(X) = o(1)$. The proof of (D.1) is the same regardless of the type of sparsity in Assumption 4.2. For notational simplicity, let

$$\xi_{it} = u_{it}^2 - E u_{it}^2, \quad \zeta_{is} = e_{is}\big(1 - \mathbf{f}_s'\mathbf{w}\big).$$

Then $X = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T} \xi_{it}(\frac{1}{\sqrt{T}} \sum_{s=1}^{T} \zeta_{is})^2$. Because of the serial independence, $\xi_{it}$ is independent of $\zeta_{js}$ if $t \neq s$, for any $i, j \leq N$, which implies $\mathrm{cov}(\xi_{it}, \zeta_{is}\zeta_{ik}) = 0$ as long as either $s \neq t$ or $k \neq t$.

*Expectation*

For the expectation,

$$\begin{aligned} EX &= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T} \mathrm{cov}\left(\xi_{it}, \left(\frac{1}{\sqrt{T}} \sum_{s=1}^{T} \zeta_{is}\right)^2\right) \\ &= \frac{1}{T\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} \sum_{k=1}^{T} \mathrm{cov}(\xi_{it}, \zeta_{is}\zeta_{ik}) \\ &= \frac{1}{T\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T} \left(\mathrm{cov}\big(\xi_{it}, \zeta_{it}^2\big) + 2\sum_{k \neq t} \mathrm{cov}(\xi_{it}, \zeta_{it}\zeta_{ik})\right) \\ &= \frac{1}{T\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T} \mathrm{cov}\big(\xi_{it}, \zeta_{it}^2\big) = O\left(\sqrt{\frac{N}{T}}\right), \end{aligned}$$

where the second last equality follows since $E\xi_{it} = E\zeta_{it} = 0$ and when $k \neq t$, $\mathrm{cov}(\xi_{it}, \zeta_{it}\zeta_{ik}) = E\xi_{it}\zeta_{it}\zeta_{ik} = E\xi_{it}\zeta_{it}E\zeta_{ik} = 0$. It then follows that $\frac{1}{T}(EX)^2 = O(\frac{N}{T^2}) = o(1)$, given $N = o(T^2)$.

*Variance*

Consider the variance. We have

$$\begin{aligned}\operatorname{var}(X) &= \frac{1}{N}\sum_{i=1}^{N}\operatorname{var}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\xi_{it}\left(\frac{1}{\sqrt{T}}\sum_{s=1}^{T}\zeta_{is}\right)^2\right)\\ &\quad + \frac{1}{NT^3}\sum_{i\neq j}\sum_{t,s,k,l,v,p\leq T}\operatorname{cov}(\xi_{it}\zeta_{is}\zeta_{ik},\xi_{jl}\zeta_{jv}\zeta_{jp})\\ &= B_1 + B_2.\end{aligned}$$

$B_1$ can be bounded by the Cauchy–Schwarz inequality. Note that $E\xi_{it} = E\zeta_{js} = 0$,

$$\begin{aligned}B_1 &\leq \frac{1}{N}\sum_{i=1}^{N}E\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\xi_{it}\left(\frac{1}{\sqrt{T}}\sum_{s=1}^{T}\zeta_{is}\right)^2\right)^2\\ &\leq \frac{1}{N}\sum_{i=1}^{N}\left[E\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\xi_{it}\right)^4\right]^{1/2}\left[E\left(\frac{1}{\sqrt{T}}\sum_{s=1}^{T}\zeta_{is}\right)^8\right]^{1/2}.\end{aligned}$$

Hence $B_1 = O(1)$.

We now show $\frac{1}{T}B_2 = o(1)$. Once this is done, it implies $\frac{1}{T}\operatorname{var}(X) = o(1)$. The proof of (D.1) is then completed because $\frac{1}{T}EX^2 = \frac{1}{T}(EX)^2 + \frac{1}{T}\operatorname{var}(X) = o(1)$.

For two variables $X, Y$, write $X \perp Y$ if they are independent. Note that $E\xi_{it} = E\zeta_{is} = 0$, and when $t \neq s$, $\xi_{it} \perp \zeta_{js}$, $\xi_{it} \perp \xi_{js}$, $\zeta_{it} \perp \zeta_{js}$ for any $i, j \leq N$. Therefore, it is straightforward to verify that if the set $\{t, s, k, l, v, p\}$ contains more than three distinct elements, then $\operatorname{cov}(\xi_{it}\zeta_{is}\zeta_{ik}, \xi_{jl}\zeta_{jv}\zeta_{jp}) = 0$. Hence if we denote $\Xi$ as the set of $(t, s, k, l, v, p)$ such that $\{t, s, k, l, v, p\}$ contains no more than three distinct elements, then its cardinality satisfies: $|\Xi|_0 \leq CT^3$ for some $C > 1$, and

$$\sum_{t,s,k,l,v,p\leq T}\operatorname{cov}(\xi_{it}\zeta_{is}\zeta_{ik},\xi_{jl}\zeta_{jv}\zeta_{jp}) = \sum_{(t,s,k,l,v,p)\in\Xi}\operatorname{cov}(\xi_{it}\zeta_{is}\zeta_{ik},\xi_{jl}\zeta_{jv}\zeta_{jp}).$$

Hence

$$B_2 = \frac{1}{NT^3}\sum_{i\neq j}\sum_{(t,s,k,l,v,p)\in\Xi}\operatorname{cov}(\xi_{it}\zeta_{is}\zeta_{ik},\xi_{jl}\zeta_{jv}\zeta_{jp}).$$

Let us partition $\Xi$ into $\Xi_1 \cup \Xi_2$, where each element $(t, s, k, l, v, p)$ in $\Xi_1$ contains exactly three distinct indices, while each element in $\Xi_2$ contains less than three distinct indices. We know that

$$\frac{1}{NT^3}\sum_{i\neq j}\sum_{(t,s,k,l,v,p)\in\Xi_2}\operatorname{cov}(\xi_{it}\zeta_{is}\zeta_{ik},\xi_{jl}\zeta_{jv}\zeta_{jp}) = O\left(\frac{1}{NT^3}N^2T^2\right) = O\left(\frac{N}{T}\right),$$

which implies

$$\frac{1}{T}B_2 = \frac{1}{NT^4}\sum_{i\neq j}\sum_{(t,s,k,l,v,p)\in\Xi_1}\operatorname{cov}(\xi_{it}\zeta_{is}\zeta_{ik},\xi_{jl}\zeta_{jv}\zeta_{jp}) + O_p\left(\frac{N}{T^2}\right).$$

The first term on the right hand side can be written as $\sum_{h=1}^5 B_{2h}$. Each of these five terms is defined and analyzed separately as below:

$$\begin{aligned}B_{21} &= \frac{1}{NT^4}\sum_{i\neq j}\sum_{t=1}^{T}\sum_{s\neq t}\sum_{l\neq s,t}E\xi_{it}\xi_{jt}E\zeta_{is}^2E\zeta_{jl}^2\\ &\leq O\left(\frac{1}{NT}\right)\sum_{i\neq j}|E\xi_{it}\xi_{jt}|.\end{aligned}$$

Note that if $(\boldsymbol{\Sigma}_u)_{ij}=0$, $u_{it}$ and $u_{jt}$ are independent, and hence $E\xi_{it}\xi_{jt}=0$. This implies $\sum_{i\neq j}|E\xi_{it}\xi_{jt}| \leq O(1)\sum_{i\neq j,(i,j)\in S_U} 1 = O(N)$. Hence $B_{21}=o(1)$.

$$B_{22} = \frac{1}{NT^4}\sum_{i\neq j}\sum_{t=1}^{T}\sum_{s\neq t}\sum_{l\neq s,t}E\xi_{it}\zeta_{it}E\zeta_{is}\xi_{js}E\zeta_{jl}^2.$$

By Lemma D.3, $u_{js}$ and $e_{is}$ are independent for $i\neq j$. Also, $u_{js}$ and $\mathbf{f}_s$ are independent, which implies $\xi_{js}$ and $\zeta_{is}$ are independent. So $E\xi_{js}\zeta_{is}=0$. It follows that $B_{22}=0$.

$$\begin{aligned}B_{23} &= \frac{1}{NT^4}\sum_{i\neq j}\sum_{t=1}^{T}\sum_{s\neq t}\sum_{l\neq s,t}E\xi_{it}\zeta_{it}E\zeta_{is}\zeta_{js}E\xi_{jl}\zeta_{jl}\\ &= O\left(\frac{1}{NT}\right)\sum_{i\neq j}|E\zeta_{is}\zeta_{js}|\\ &= O\left(\frac{1}{NT}\right)\sum_{i\neq j}\left|Ee_{is}e_{js}E\left(1-\mathbf{f}_s'\mathbf{w}\right)^2\right|\\ &= O\left(\frac{1}{NT}\right)\sum_{i\neq j}|Ee_{is}e_{js}|.\end{aligned}$$

By the definition $\mathbf{e}_s = \boldsymbol{\Sigma}_u^{-1}\mathbf{u}_s$, $\mathrm{cov}(\mathbf{e}_s) = \boldsymbol{\Sigma}_u^{-1}$. Hence $Ee_{is}e_{js} = (\boldsymbol{\Sigma}_u^{-1})_{ij}$, which implies $B_{23} \leq O(\frac{N}{NT})\|\boldsymbol{\Sigma}_u^{-1}\|_1 = o(1)$.

$$B_{24} = \frac{1}{NT^4}\sum_{i\neq j}\sum_{t=1}^{T}\sum_{s\neq t}\sum_{l\neq s,t} E\xi_{it}\xi_{jt}E\zeta_{is}\zeta_{js}E\zeta_{il}\zeta_{jl} = O\left(\frac{1}{T}\right),$$

which is analyzed in the same way as $B_{21}$.

Finally, $B_{25} = \frac{1}{NT^4}\sum_{i\neq j}\sum_{t=1}^{T}\sum_{s\neq t}\sum_{l\neq s,t} E\xi_{it}\zeta_{jt}E\zeta_{is}\xi_{js}E\zeta_{il}\zeta_{jl} = 0$, because $E\zeta_{is}\xi_{js} = 0$ when $i \neq j$, following from Lemma D.3. Therefore, $\frac{1}{T}B_2 = o(1) + O(\frac{N}{T^2}) = o(1)$. *Q.E.D.*

PROOF OF (D.2): For notational simplicity, let $\xi_{ijt} = u_{it}u_{jt} - Eu_{it}u_{jt}$. Because of the serial independence and the Gaussianity, $\mathrm{cov}(\xi_{ijt}, \zeta_{ls}\zeta_{nk}) = 0$ when either $s \neq t$ or $k \neq t$, for any $i, j, l, n \leq N$. In addition, define a set

$$H = \{(i, j) \in S_U : i \neq j\}.$$

Then by the sparsity assumption, $\sum_{(i,j)\in H} 1 = D_N = O(N)$. Now let

$$\begin{aligned}
Z &= \frac{1}{\sqrt{NT}}\sum_{(i,j)\in H}\sum_{t=1}^{T}(u_{it}u_{jt} - Eu_{it}u_{jt}) \\
&\quad\times\left[\frac{1}{\sqrt{T}}\sum_{s=1}^{T} e_{is}(1 - \mathbf{f}_s'\mathbf{w})\right]\left[\frac{1}{\sqrt{T}}\sum_{k=1}^{T} e_{jk}(1 - \mathbf{f}_k'\mathbf{w})\right] \\
&= \frac{1}{\sqrt{NT}}\sum_{(i,j)\in H}\sum_{t=1}^{T}\xi_{ijt}\left[\frac{1}{\sqrt{T}}\sum_{s=1}^{T}\zeta_{is}\right]\left[\frac{1}{\sqrt{T}}\sum_{k=1}^{T}\zeta_{jk}\right] \\
&= \frac{1}{T\sqrt{NT}}\sum_{(i,j)\in H}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{k=1}^{T}\xi_{ijt}\zeta_{is}\zeta_{jk}.
\end{aligned}$$

The goal is to show $\frac{1}{T}EZ^2 = o(1)$. We respectively show $\frac{1}{T}(EZ)^2 = o(1) = \frac{1}{T}\mathrm{var}(Z)$.

*Expectation*

The proof for the expectation is the same regardless of the type of sparsity in Assumption 4.2, and is very similar to that of (D.1). In fact,

$$\begin{aligned}
EZ &= \frac{1}{T\sqrt{NT}}\sum_{(i,j)\in H}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{k=1}^{T}\mathrm{cov}(\xi_{ijt}, \zeta_{is}\zeta_{jk}) \\
&= \frac{1}{T\sqrt{NT}}\sum_{(i,j)\in H}\sum_{t=1}^{T}\mathrm{cov}(\xi_{ijt}, \zeta_{it}^2).
\end{aligned}$$

Because $\sum_{(i,j)\in H} 1 = O(N)$, $EZ = O(\sqrt{\frac{N}{T}})$. Thus $\frac{1}{T}(EZ)^2 = o(1)$.

*Variance*

For the variance, we have

$$\begin{aligned}
&\operatorname{var}(Z)\\
&= \frac{1}{T^3 N} \sum_{(i,j)\in H} \operatorname{var}\left(\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{k=1}^{T} \xi_{ijt}\zeta_{is}\zeta_{jk}\right)\\
&\quad + \frac{1}{T^3 N} \sum_{(i,j)\in H,} \sum_{(m,n)\in H,(m,n)\neq(i,j),} \sum_{t,s,k,l,v,p\leq T} \operatorname{cov}(\xi_{ijt}\zeta_{is}\zeta_{jk}, \xi_{mnl}\zeta_{mv}\zeta_{np})\\
&= A_1 + A_2.
\end{aligned}$$

By the Cauchy–Schwarz inequality and the serial independence of $\xi_{ijt}$,

$$\begin{aligned}
A_1 &\leq \frac{1}{N} \sum_{(i,j)\in H} E\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\xi_{ijt}\frac{1}{\sqrt{T}}\sum_{s=1}^{T}\zeta_{is}\frac{1}{\sqrt{T}}\sum_{k=1}^{T}\zeta_{jk}\right]^2\\
&\leq \frac{1}{N} \sum_{(i,j)\in H} \left[E\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\xi_{ijt}\right)^4\right]^{1/2}\\
&\quad \times \left[E\left(\frac{1}{\sqrt{T}}\sum_{s=1}^{T}\zeta_{is}\right)^8\right]^{1/4}\left[E\left(\frac{1}{\sqrt{T}}\sum_{k=1}^{T}\zeta_{jk}\right)^8\right]^{1/4}.
\end{aligned}$$

So $A_1 = O(1)$.

Note that $E\xi_{ijt} = E\zeta_{is} = 0$, and when $t \neq s$, $\xi_{ijt} \perp \zeta_{ms}$, $\xi_{ijt} \perp \xi_{mns}$, $\zeta_{it} \perp \zeta_{js}$ (independent) for any $i, j, m, n \leq N$. Therefore, it is straightforward to verify that if the set $\{t, s, k, l, v, p\}$ contains more than three distinct elements, then $\operatorname{cov}(\xi_{ijt}\zeta_{is}\zeta_{jk}, \xi_{mnl}\zeta_{mv}\zeta_{np}) = 0$. Hence for the same set $\Xi$ defined as before, it satisfies: $|\Xi|_0 \leq CT^3$ for some $C > 1$, and

$$\begin{aligned}
&\sum_{t,s,k,l,v,p\leq T} \operatorname{cov}(\xi_{ijt}\zeta_{is}\zeta_{jk}, \xi_{mnl}\zeta_{mv}\zeta_{np})\\
&= \sum_{(t,s,k,l,v,p)\in\Xi} \operatorname{cov}(\xi_{ijt}\zeta_{is}\zeta_{jk}, \xi_{mnl}\zeta_{mv}\zeta_{np}).
\end{aligned}$$

We proceed by studying the two cases of Assumption 4.2 separately, and show that in both cases, $\frac{1}{T}A_2 = o(1)$. Once this is done, because we have just shown $A_1 = O(1)$, then $\frac{1}{T}\operatorname{var}(Z) = o(1)$. The proof is then completed because $\frac{1}{T}EZ^2 = \frac{1}{T}(EZ)^2 + \frac{1}{T}\operatorname{var}(Z) = o(1)$.

*When* $D_N = O(\sqrt{N})$

Because $|\Xi|_0 \leq CT^3$ and $|H|_0 = D_N = O(\sqrt{N})$, and $|\operatorname{cov}(\xi_{ijt}\zeta_{is}\zeta_{jk}, \xi_{mnl}\zeta_{mv}\zeta_{np})|$ is bounded uniformly in $i, j, m, n \leq N$, we have

$$\frac{1}{T}A_2 = \frac{1}{T^4 N}\sum_{(i,j)\in H,}\sum_{(m,n)\in H,(m,n)\neq(i,j),}\sum_{t,s,k,l,v,p\in\Xi}\operatorname{cov}(\xi_{ijt}\zeta_{is}\zeta_{jk}, \xi_{mnl}\zeta_{mv}\zeta_{np})$$
$$= O\left(\frac{1}{T}\right).$$

*When* $D_n = O(N)$, *and* $m_N = O(1)$

Similarly to the proof of the first statement, for the same set $\Xi_1$ that contains exactly three distinct indices in each of its element (recall $|H|_0 = O(N)$),

$$\frac{1}{T}A_2 = \frac{1}{NT^4}\sum_{(i,j)\in H,}\sum_{(m,n)\in H,(m,n)\neq(i,j),}\sum_{t,s,k,l,v,p\in\Xi_1}\operatorname{cov}(\xi_{ijt}\zeta_{is}\zeta_{jk}, \xi_{mnl}\zeta_{mv}\zeta_{np})$$
$$+ O\left(\frac{N}{T^2}\right).$$

The first term on the right hand side can be written as $\sum_{h=1}^{5} A_{2h}$. Each of these five terms is defined and analyzed separately as below. Before that, let us introduce a useful lemma.

The following lemma is needed when $\boldsymbol{\Sigma}_u$ has bounded number of nonzero entries in each row ($m_N = O(1)$). Let $|S|_0$ denote the number of elements in a set $S$ if $S$ is countable. For any $i \leq N$, let

$$A(i) = \{j \leq N : \operatorname{cov}(u_{it}, u_{jt}) \neq 0\} = \{j \leq N : (i, j) \in S_U\}.$$

LEMMA D.4: *Suppose* $m_N = O(1)$. *For any* $i, j \leq N$, *let* $B(i, j)$ *be a set of* $k \in \{1, \dots, N\}$ *such that*:

(i) $k \notin A(i) \cup A(j)$,

(ii) *there is* $p \in A(k)$ *such that* $\operatorname{cov}(u_{it}u_{jt}, u_{kt}u_{pt}) \neq 0$.

*Then* $\max_{i,j\leq N} |B(i, j)|_0 = O(1)$.

PROOF: First we note that if $B(i, j) = \emptyset$, then $|B(i, j)|_0 = 0$. If it is not empty, for any $k \in B(i, j)$, by definition, $k \notin A(i) \cup A(j)$, which implies $\operatorname{cov}(u_{it}, u_{kt}) = \operatorname{cov}(u_{jt}, u_{kt}) = 0$. By the Gaussianity, $u_{kt}$ is independent of $(u_{it}, u_{jt})$. Hence if $p \in A(k)$ is such that $\operatorname{cov}(u_{it}u_{jt}, u_{kt}u_{pt}) \neq 0$, then $u_{pt}$ should be correlated with either $u_{it}$ or $u_{jt}$. We thus must have $p \in A(i) \cup A(j)$. In other words, there is $p \in A(i) \cup A(j)$ such that $\operatorname{cov}(u_{kt}, u_{pt}) \neq 0$, which implies $k \in A(p)$.

Hence,

$$k \in \bigcup_{p \in A(i) \cup A(j)} A(p) \equiv M(i, j),$$

and thus $B(i, j) \subset M(i, j)$. Because $m_N = O(1)$, $\max_{i \leq N} |A(i)|_0 = O(1)$, which implies $\max_{i,j} |M(i, j)|_0 = O(1)$, yielding the result. *Q.E.D.*

Now we define and bound each of $A_{2h}$. For any $(i, j) \in H = \{(i, j) : (\boldsymbol{\Sigma}_u)_{ij} \neq 0\}$, we must have $j \in A(i)$. So

$$\begin{aligned} A_{21} &= \frac{1}{NT^4} \sum_{(i,j)\in H,} \sum_{(m,n)\in H, (m,n)\neq(i,j),} \sum_{t=1}^{T} \sum_{s\neq t} \sum_{l \neq t,s} E\xi_{ijt}\xi_{mnt} E\zeta_{is}\zeta_{js} E\zeta_{ml}\zeta_{nl} \\ &\leq O\left(\frac{1}{NT}\right) \sum_{(i,j)\in H,} \sum_{(m,n)\in H, (m,n)\neq(i,j)} |E\xi_{ijt}\xi_{mnt}| \\ &\leq O\left(\frac{1}{NT}\right) \\ &\quad \times \sum_{(i,j)\in H} \left( \sum_{m \in A(i)\cup A(j)} \sum_{n \in A(m)} + \sum_{m \notin A(i)\cup A(j)} \sum_{n\in A(m)} \right) \bigl|\operatorname{cov}(u_{it}u_{jt}, u_{mt}u_{nt})\bigr|. \end{aligned}$$

The first term is $O(\frac{1}{T})$ because $|H|_0 = O(N)$ and $|A(i)|_0$ is bounded uniformly by $m_N = O(1)$. So the number of summands in $\sum_{m\in A(i)\cup A(j)} \sum_{n\in A(m)}$ is bounded. For the second term, if $m \notin A(i) \cup A(j)$, $n \in A(m)$, and $\operatorname{cov}(u_{it}u_{jt}, u_{mt}u_{nt}) \neq 0$, then $m \in B(i, j)$. Hence the second term is bounded by $O(\frac{1}{NT}) \sum_{(i,j)\in H} \sum_{m\in B(i,j)} \sum_{n\in A(m)} |\operatorname{cov}(u_{it}u_{jt}, u_{mt}u_{nt})|$, which is also $O(\frac{1}{T})$ by Lemma D.4. Hence $A_{21} = o(1)$.

Similarly, applying Lemma D.4,

$$\begin{aligned} A_{22} &= \frac{1}{NT^4} \sum_{(i,j)\in H,} \sum_{(m,n)\in H, (m,n)\neq(i,j),} \sum_{t=1}^{T} \sum_{s\neq t} \sum_{l\neq t,s} E\xi_{ijt}\xi_{mnt} E\zeta_{is}\zeta_{ms} E\zeta_{jl}\zeta_{nl} \\ &= o(1), \end{aligned}$$

which is proved in the same lines of those of $A_{21}$.

Also note three simple facts: (1) $\max_{j\leq N} |A(j)|_0 = O(1)$, (2) $(m, n) \in H$ implies $n \in A(m)$, and (3) $\xi_{mms} = \xi_{nms}$. The term $A_{23}$ is defined as

$$\begin{aligned} A_{23} &= \frac{1}{NT^4} \sum_{(i,j)\in H,} \sum_{(m,n)\in H, (m,n)\neq(i,j),} \sum_{t=1}^{T} \sum_{s\neq t} \sum_{l\neq t,s} E\xi_{ijt}\zeta_{it} E\zeta_{js}\xi_{mns} E\zeta_{ml}\zeta_{nl} \\ &\leq O\left(\frac{1}{NT}\right) \sum_{j=1}^{N} \sum_{i\in A(j)} 1 \sum_{(m,n)\in H, (m,n)\neq(i,j)} |E\zeta_{js}\xi_{mns}| \end{aligned}$$

$$\leq O\left(\frac{2}{NT}\right)\sum_{j=1}^{N}\sum_{n\in A(j)}|E\zeta_{js}\xi_{jns}|+O\left(\frac{1}{NT}\right)\sum_{j=1}^{N}\sum_{m\neq j,n\neq j}|E\zeta_{js}\xi_{mns}|$$
$$= a+b.$$

Term $a=O(\frac{1}{T})$. For $b$, note that Lemma D.3 implies that when $m,n\neq j$, $u_{ms}u_{ns}$ and $e_{js}$ are independent because of the Gaussianity. Also because $\mathbf{u}_s$ and $\mathbf{f}_s$ are independent, hence $\zeta_{js}$ and $\xi_{mms}$ are independent, which implies that $b=0$. Hence $A_{23}=o(1)$.

The same argument as of $A_{23}$ also implies

$$A_{24}=\frac{1}{NT^4}\sum_{(i,j)\in H,}\ \sum_{(m,n)\in H,(m,n)\neq(i,j),}\ \sum_{t=1}^{T}\sum_{s\neq t}\sum_{l\neq t,s}E\xi_{ijt}\zeta_{mt}E\zeta_{is}\xi_{mns}E\zeta_{il}\zeta_{nl}$$
$$=o(1).$$

Finally, because $\sum_{(i,j)\in H}1\leq\sum_{i=1}^{N}\sum_{j\in A(i)}1\leq m_N\sum_{i=1}^{N}1$, and $m_N=O(1)$, we have

$$A_{25}=\frac{1}{NT^4}\sum_{(i,j)\in H,}\ \sum_{(m,n)\in H,(m,n)\neq(i,j),}\ \sum_{t=1}^{T}\sum_{s\neq t}\sum_{l\neq t,s}E\xi_{ijt}\zeta_{it}E\zeta_{is}\zeta_{ms}E\xi_{mnl}\zeta_{nl}$$
$$\leq O\left(\frac{1}{NT}\right)\sum_{(i,j)\in H,}\ \sum_{(m,n)\in H,(m,n)\neq(i,j)}|E\xi_{ijt}\zeta_{it}E\zeta_{is}\zeta_{ms}E\xi_{mnl}\zeta_{nl}|$$
$$\leq O\left(\frac{1}{NT}\right)\sum_{i=1}^{N}\sum_{m=1}^{N}|E\zeta_{is}\zeta_{ms}|$$
$$\leq O\left(\frac{1}{NT}\right)\sum_{i=1}^{N}\sum_{m=1}^{N}\left|\left(\boldsymbol{\Sigma}_u^{-1}\right)_{im}\right|E\left(1-\mathbf{f}_s'\mathbf{w}\right)^2$$
$$\leq O\left(\frac{N}{NT}\right)\left\|\boldsymbol{\Sigma}_u^{-1}\right\|_1=o(1).$$

In summary, $\frac{1}{T}A_2=o(1)+O(\frac{N}{T^2})=o(1)$. *Q.E.D.*

## APPENDIX E: FURTHER TECHNICAL LEMMAS FOR SECTION 4

We cite a lemma that will be needed throughout the proofs.

LEMMA E.1: *Under Assumption* 4.1, *there is* $C>0$,

(i) $P(\max_{i,j\leq N}|\frac{1}{T}\sum_{t=1}^{T}u_{it}u_{jt}-Eu_{it}u_{jt}|>C\sqrt{\frac{\log N}{T}})\to 0$,

(ii) $P(\max_{i\leq K,j\leq N}|\frac{1}{T}\sum_{t=1}^{T}f_{it}u_{jt}|>C\sqrt{\frac{\log N}{T}})\to 0$,

(iii) $P(\max_{j\leq N}|\frac{1}{T}\sum_{t=1}^{T}u_{jt}|>C\sqrt{\frac{\log N}{T}})\to 0$.

PROOF: The proof follows from Lemmas A.3 and B.1 in Fan, Liao, and Mincheva (2011). *Q.E.D.*

LEMMA E.2: *When the distribution of* $(\mathbf{u}_t, \mathbf{f}_t)$ *is independent of* $\boldsymbol{\theta}$, *there is* $C > 0$,

(i) $\sup_{\boldsymbol{\theta}\in\Theta} P(\max_{j\leq N} |\widehat{\theta}_j - \theta_j| > C\sqrt{\frac{\log N}{T}}|\boldsymbol{\theta}) \to 0$,

(ii) $\sup_{\boldsymbol{\theta}\in\Theta} P(\max_{i,j\leq N} |\widehat{\sigma}_{ij} - \sigma_{ij}| > C\sqrt{\frac{\log N}{T}}|\boldsymbol{\theta}) \to 0$,

(iii) $\sup_{\boldsymbol{\theta}\in\Theta} P(\max_{i\leq N} |\widehat{\sigma}_i - \sigma_i| > C\sqrt{\frac{\log N}{T}}|\boldsymbol{\theta}) \to 0$.

PROOF: Note that $\widehat{\theta}_j - \theta_j = \frac{1}{a_{f,T}T}\sum_{t=1}^T u_{jt}(1 - \mathbf{f}_t'\mathbf{w})$. Here $a_{f,T} = 1 - \bar{\mathbf{f}}'\mathbf{w} \to^p 1 - E\mathbf{f}_t'(E\mathbf{f}_t\mathbf{f}_t')^{-1}E\mathbf{f}_t > 0$, hence $a_{f,T}$ is bounded away from zero with probability approaching 1. Thus by Lemma E.1, there is $C > 0$ independent of $\boldsymbol{\theta}$, such that

$$\sup_{\boldsymbol{\theta}\in\Theta} P\left(\max_{j\leq N} |\widehat{\theta}_j - \theta_j| > C\sqrt{\frac{\log N}{T}}\middle|\boldsymbol{\theta}\right)$$
$$= P\left(\max_j \left|\frac{1}{a_{f,T}T}\sum_{t=1}^T u_{jt}(1 - \mathbf{f}_t'\mathbf{w})\right| > C\sqrt{\frac{\log N}{T}}\right) \to 0.$$

(ii) There is $C$ independent of $\boldsymbol{\theta}$, such that the event

$$A = \left\{\max_{i,j}\left|\frac{1}{T}\sum_{t=1}^T u_{it}u_{jt} - \sigma_{ij}\right| < C\sqrt{\frac{\log N}{T}}, \frac{1}{T}\sum_{t=1}^T \|\mathbf{f}_t\|^2 < C\right\}$$

has probability approaching 1. Also, there is $C_2$ also independent of $\boldsymbol{\theta}$ such that the event $B = \{\max_i \frac{1}{T}\sum_t u_{it}^2 < C_2\}$ occurs with probability approaching 1. Then on the event $A \cap B$, by the triangular and Cauchy–Schwarz inequalities,

$$|\widehat{\sigma}_{ij} - \sigma_{ij}| \leq C\sqrt{\frac{\log N}{T}} + 2\max_i \sqrt{\frac{1}{T}\sum_t (\widehat{u}_{it} - u_{it})^2 C_2}$$
$$+ \max_i \frac{1}{T}\sum_t (u_{it} - \widehat{u}_{it})^2.$$

It can be shown that

$$\max_{i\leq N} \frac{1}{T}\sum_{t=1}^T (\widehat{u}_{it} - u_{it})^2$$
$$\leq \max_i\left(\|\widehat{\mathbf{b}}_i - \mathbf{b}_i\|^2 + (\widehat{\theta}_i - \theta_i)^2\right)\left(\frac{1}{T}\sum_{t=1}^T \|\mathbf{f}_t\|^2 + 1\right).$$

Note that $\widehat{\mathbf{b}}_i - \mathbf{b}_i$ and $\widehat{\theta}_i - \theta_i$ only depend on $(\mathbf{f}_t, \mathbf{u}_t)$ (independent of $\boldsymbol{\theta}$). By Lemma 3.1 of Fan, Liao, and Mincheva (2011), there is $C_3 > 0$ such that $\sup_{\mathbf{b},\boldsymbol{\theta}} P(\max_{i\le N} \|\widehat{\mathbf{b}}_i - \mathbf{b}_i\|^2 + (\widehat{\theta}_i - \theta_i)^2 > C_3 \frac{\log N}{T}) = o(1)$. Combining the last two displayed inequalities yields, for $C_4 = (C+1)C_3$,

$$\sup_{\boldsymbol{\theta}} P\left(\max_{i\le N} \frac{1}{T}\sum_{t=1}^{T}(\widehat{u}_{it} - u_{it})^2 > C_4 \frac{\log N}{T}\Big|\boldsymbol{\theta}\right) = o(1),$$

which yields the desired result.

(iii) Recall $\widehat{\sigma}_j^2 = \widehat{\sigma}_{jj}/a_{f,T}$, and $\sigma_j^2 = \sigma_{jj}/(1 - E\mathbf{f}_t'(E\mathbf{f}_t\mathbf{f}_t')^{-1}E\mathbf{f}_t)$. Moreover, $a_{f,T}$ is independent of $\boldsymbol{\theta}$. The result follows immediately from part (ii). *Q.E.D.*

LEMMA E.3: *For any* $\varepsilon > 0$, $\sup_{\boldsymbol{\theta}} P(\|\widehat{\boldsymbol{\Sigma}}_u^{-1} - \boldsymbol{\Sigma}_u^{-1}\| > \varepsilon|\boldsymbol{\theta}) = o(1)$.

PROOF: By Lemma E.2(ii), $\sup_{\boldsymbol{\theta}\in\Theta} P(\max_{i,j\le N} |\widehat{\sigma}_{ij} - \sigma_{ij}| > C\sqrt{\frac{\log N}{T}}|\boldsymbol{\theta}) \to 1$. By Fan, Liao, and Mincheva (2011), on the event $\max_{i,j\le N} |\widehat{\sigma}_{ij} - \sigma_{ij}| \le C\sqrt{\frac{\log N}{T}}$, there is constant $C'$ that is independent of $\boldsymbol{\theta}$, $\|\widehat{\boldsymbol{\Sigma}}_u^{-1} - \boldsymbol{\Sigma}_u^{-1}\| \le C' m_N (\frac{\log N}{T})^{1/2}$. Hence the result follows due to the sparse condition $m_N(\frac{\log N}{T})^{1/2} = o(1)$. *Q.E.D.*

## REFERENCE

FAN, J., Y. LIAO, AND M. MINCHEVA (2011): "High Dimensional Covariance Matrix Estimation in Approximate Factor Models," *The Annals of Statistics*, 39, 3320–3356. [2,13,14]

*Dept. of Operations Research and Financial Engineering, Princeton University, Princeton, NJ 08544, U.S.A., International School of Economics and Management, Capital University of Economics and Business, Beijing, China, and Lingnan College, Sun Yat-sen University, Guangzhou, China; jqfan@princeton.edu,*

*Dept. of Mathematics, University of Maryland, College Park, MD 20742, U.S.A.; yuanliao@umd.edu,*

*and*

*Dept. of Operations Research and Financial Engineering, Princeton University, Princeton, NJ 08544, U.S.A.; jiaweiy@princeton.edu.*

*Manuscript received August, 2014; final revision received February, 2015.*