Lemma 1 :- (CLT) Let $\{X_{nt}\}$ be a sequence such that $EX_{nt}=0$ for all n,t and

(i) α coefficients are of size $\frac{-p}{p-2}$, $p>2$

(ii) $\sup_t E|X_{nt}|^p < \Delta$ for all n

(iii) $w_n := Var\left(\frac{1}{\sqrt{n}}\sum_{t=1}^n X_{nt}\right) > \delta > 0$ for all n sufficiently large

Then

$$\frac{1}{\sqrt{n}}\sum_{t=1}^n \frac{X_{nt}}{w_n^{1/2}} \xrightarrow{d} N(0,1).$$

(*) The books show CLT for random variables. We can show the same for random vectors.

Definition .- Let $\{M_n\}$ be a sequence of $k\times k$ matrices. Let e_n be the smallest eigenvalue of M_n. Then M_n is said to be uniformly positive definite if for all n sufficiently large $e_n > \delta > 0$ uniformly in n.

Proposition 6 .- Let $\{X_{nt}\}$ be an α-mixing sequence of random vectors such that $EX_{nt}=0$ for all n,t and for some $p>2$ and $\Delta>0$,

(*) Recall $X_{nt} = \begin{pmatrix} X_{nt1} \\ \vdots \\ X_{ntk}\end{pmatrix}_{k\times 1}$

(i) α is of size $\frac{-p}{p-2}$

(ii) $E|X_{nt}|^p \leq \Delta$ for all t,n

(iii) $\Omega_n = Var\left(\frac{1}{\sqrt{n}}\sum_{t=1}^n X_{nt}\right)$ is uniformly positive definite.

Then

$$\Omega_n^{-1/2}\frac{1}{\sqrt{n}}\sum_{t=1}^n X_{nt} \xrightarrow{d} N(0,I_k)$$

Proof: Let $\lambda\in\mathbb{R}^k$ such that $\|\lambda\|=1$. Then by the Cramér-Wold device we want to show that $\lambda'\Omega_n^{-1/2}\frac{1}{\sqrt{n}}\sum_{t=1}^n X_{nt} \xrightarrow{d} N(0,1)$ using Lemma 1.

We need to check for the conditions of Lemma 1

- $E(\lambda'\Omega_n^{-1/2}X_{nt}) = \lambda'\Omega_n^{-1/2}E(X_{nt}) = 0$ for all n,t. new random variable (linear comb)

(i) • Define the process $\{Y_{nt}\} := \{\lambda'\Omega_n^{-1/2}X_{nt}\} = g(X_{nt1},\dots,X_{ntk})$ where g is measurable. Then by Proposition 1 $\{Y_{nt}\}$ is α-mixing of the same size $\frac{-p}{p-2}$.

(ii). $\displaystyle \sup_t E\Big|\underbrace{\lambda'}_{1\times k}\underbrace{\Omega_n^{-1/2}}_{\substack{k\times k\\ \text{depends on } n}}\underbrace{X_{nt}}_{k\times 1}\Big|^p = \sup_t E\Big|\sum_{j=1}^{k} c_{jn} X_{ntj}\Big|^p$

$$= \sup_t \Big(E\Big|\sum_{j=1}^{k} c_{jn} X_{ntj}\Big|^p\Big)^{1/p\cdot p}$$

$$\underset{\text{Minkowski's Inequality}}{\le} \sup_t \Big\{\sum_{j=1}^{k}|c_{jn}|\left(E|X_{ntj}|^p\right)^{1/p}\Big\}^p$$

$$\le \Delta\Big(\sum_{j=1}^{k}\underbrace{|c_{jn}|}_{L_1\text{ norm}}\Big)^p$$

$$\underset{\text{Norm inequality}}{\le} \Delta\Big(\sum_{j=1}^{k}|c_{jn}|^2\Big)^{p/2}$$

$$= \Delta\left(\lambda'\Omega_n^{-1}\lambda\right)^{p/2}$$

$C_n\Lambda_n^{-1}C_n'$ where $C_nC_n' = I_k$, $\|\lambda\| = 1$

$$\underset{\text{spectral decomposition}}{=} \Delta\big(\underbrace{\lambda' C_n}_{d_n'}\Lambda_n^{-1}\underbrace{C_n'\lambda}_{d_n}\big)^{p/2}$$

where $d_n'd_n = 1$ by construction

$$= \Delta\left(d_n'\Lambda_n^{-1}d_n\right)^{p/2}$$

largest eigenv of Ω_n^{-1} is the smallest eigenv of Ω

$$\le \Delta\Big(\underbrace{e_n^{-1}}_{<\delta}\underbrace{\sum_{t=1}^{n}d_{nt}^2}_{\|d_n\| = d_n'd_n = 1}\Big)^{p/2}$$

e_n^{-1} is the largest eigenvalue of Ω_n^{-1}

$$< \Delta\,\delta^{-p/2}$$

$$< \infty.$$

- $\displaystyle \mathrm{Var}\Big(\lambda'\Omega_n^{-1/2}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}X_{nt}\Big) = \lambda'\Omega_n^{-1/2}\,\mathrm{Var}\Big(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}X_{nt}\Big)\,\Omega_n^{-1/2}\lambda$

$$= \lambda'\Omega_n^{-1/2}\Omega_n\Omega_n^{-1/2}\lambda = 1 > 0 \quad \text{no matter } n, t$$

Therefore, by Lemma 1 the desired result holds and the proof is complete. ∎

Linear Regression with weakly dependent data

Consider the usual regression model

$$y_t = x_t'\beta + u_t$$

Consistency

Provided

(a) $\{(x_t', u_t)\}$ is α-mixing of any size

(b) $\lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} x_t u_t = 0$

(c) $E|x_{tj}|^{2+\eta} < \Delta$ for all t and $j=1,\dots,k$ and some $\eta>0$

(d) $E|u_t|^{2+\eta} < \Delta$ for some $\eta > 0$

(e) $M_n = \frac{1}{n}\sum_{t=1}^{n} E x_t x_t'$ is uniformly positive definite over n.

Then $\hat\beta_n - \beta = o_p(1)$

⊛ $\{x_t u_t\}$ and $\{x_t x_t'\}$ are also α-mixing of the same size, by Proposition 1.

$$\|x_t u_t\|^{1+\eta} \le (\|x_t\|^{2+\eta})^{1/2} (\|u_t\|^{2+\eta})^{1/2}$$

Proof:

We start from the moment condition

$$0 = \lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} E x_t u_t = \lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} E x_t (y_t - x_t'\beta)$$

$$\Rightarrow \quad \beta = \left(\lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} E x_t x_t'\right)^{-1} \left(\lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} E x_t y_t\right)$$

Sample analogue: drop the "E"!

$$\hat\beta_n = \left(\frac{1}{n}\sum_{t=1}^{n} x_t x_t'\right)^{-1} \left(\frac{1}{n}\sum_{t=1}^{n} x_t y_t\right)$$

$$= \beta + \left(\frac{1}{n}\sum_{t=1}^{n} x_t x_t'\right)^{-1} \frac{1}{n}\sum_{t=1}^{n} x_t u_t$$

$$= \beta + \Big(\underbrace{\frac{1}{n}\sum_{t=1}^{n} x_t x_t' - M_n}_{o_p(1)} + M_n\Big)^{-1} \frac{1}{n}\sum_{t=1}^{n} x_t u_t$$

$E\|x_t x_t'\|^{1+\eta} \le (E\|x_t\|^{2+2\eta} E\|x_t\|^{2+2\eta})^{1/2}$ by Cauchy-Schwarz

WLLN $\{x_t x_t'\}$

$$= \beta + \Big(o_p(1)\underbrace{M_n^{-1}}_{O(1)} + M_n M_n^{-1}\Big)^{-1} M_n^{-1} \frac{1}{n}\sum_{t=1}^{n} x_t u_t$$

$$= \beta + \left(o_p(1)\, O(1) + I_K\right)^{-1} O(1)\ \frac{1}{n}\sum_{t=1}^{n} x_t u_t$$

$$= \beta + [I_K + o_p(1)]\ O(1)\ \frac{1}{n}\sum_{t=1}^{n} X_t u_t$$

$$= \beta + [I_K + o_p(1)]\ O(1) \left[\frac{1}{n}\sum_{t=1}^{n} X_t u_t - \frac{1}{n}\sum_{t=1}^{n} EX_t u_t + \frac{1}{n}\sum_{t=1}^{n} EX_t u_t \right]$$

$E\|X_t u_t\|^{1+\eta} \le (E\|X_t\|^{2+2\eta} E|u_t|^{2(2\eta)})^{1/2}$ by Cauchy-Schwarz

$$= \beta + [I_K + o_p(1)]\ O(1) \left[\underbrace{o_p(1)}_{\text{WLLN, } \{X_t u_t\}} + o(1) \right]$$

$$= \beta + I_K\ O(1)\ o_p(1) + o_p(1)\ O(1)\ o_p(1)$$

$$= \beta + o_p(1). \quad \blacksquare$$

Asymptotic Normality

(*) Reminder: conditions for CLT in Proposition 6
(i) α-mixing size $-p/p-2$
(ii) $E\|X_{nt}\|^p \le \Delta$ for all t, n
(iii) $\Omega_n = Var\left(\frac{1}{\sqrt{n}}\sum X_{nt}\right)$ is uniformly p.d. and $p > 2$.

Provided

(a) $\{(X_t', u_t)\}$ is α-mixing of size $-p/p-2$

(b) $\frac{1}{\sqrt{n}}\sum_{t=1}^{n} EX_t u_t = o(1)$

(c) $E|X_{tj}|^{2p} \le \Delta$ for all t and $j = 1, \dots, K$

(d) $E|u_t|^{2p} \le \Delta$ for all t

(e) $M_n = \frac{1}{n}\sum_{t=1}^{n} EX_t X_t'$ is uniformly positive definite over n

(f) $\Omega_n = Var\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n} X_t u_t\right)$ is uniformly positive definite over n

Then

$$V_n^{-1/2}\sqrt{n}(\hat{\beta}_n - \beta) \xrightarrow{d} N(0, I_K)$$

where $V_n = M_n^{-1}\Omega_n M_n^{-1}$.

proof:

From the previous proposition we get

$$\sqrt{n}(\hat{\beta}_n - \beta) = (I_K + o_p(1))\ M_n^{-1}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\{X_t u_t - EX_t u_t\} + o_p(1)$$

$$= M_n^{-1}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\{X_t u_t - EX_t u_t\} + o_p(1)\ O(1)\ \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\{X_t u_t - EX_t u_t\} + o_p(1)$$

$$= \frac{1}{\sqrt{n}}\sum_{t=1}^{n} \underbrace{M_n^{-1}\{X_t u_t - EX_t u_t\}}_{\xi_{nt}} + \underbrace{o_p(1)\ \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\{X_t u_t - EX_t u_t\}}_{o_p(1)\ O(1)\ O_p(1) = o_p(1)} + o_p(1)$$

$$= \underbrace{\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \xi_{n,t}}_{A} + \underbrace{o_p(1) \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \{X_t u_t - E X_t u_t\}}_{B} + o_p(1)$$

- We will deal with B first. We're interested in the process $\{X_t u_t - E X_t u_t\}$, so we check for the conditions

(i) $\{X_t u_t - E X_t u_t\}$ is a measurable function of $\{X_t', u_t\}$ so by Proposition 1 this is α-mixing of size $-p/p-2$.

(ii) $$E\| X_t u_t - E X_t u_t \|^p \underset{\text{Minkowski's Inequality}}{\leq} \{(E\|X_t u_t\|^p)^{1/p} + (E\|X_t u_t\|^p)^{1/p}\}^p$$

$$\underset{\text{Cauchy-Schwarz}}{\leq} 2^p (E\|X_t\|^{2p})^{1/2} (E\|u_t\|^{2p})^{1/2}$$

$$< \infty$$

(iii) $$\mathrm{Var}\left(\frac{1}{\sqrt{n}} \sum_{t=1}^{n} [X_t u_t - E X_t u_t]\right) = \mathrm{Var}\left(\frac{1}{\sqrt{n}} \sum_{t=1}^{n} X_t u_t\right) = \Omega_n$$ which is uniformly p.d. by assumption.

Then, by the CLT $\Omega_n^{-1/2} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \{X_t u_t - E X_t u_t\} = O_p(1)$.

We write B as:

$$o_p(1) \underbrace{\Omega_n^{-1/2} \Omega_n}_{\Omega_n^{1/2}} \Omega_n^{-1/2} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \{X_t u_t - E X_t u_t\} = o_p(1)\, \Omega_n^{-1/2} \Omega_n\, O_p(1)$$

$$= o_p(1)\, O(1)\, O_p(1)$$

$$= o_p(1).$$

- Now we can deal with A. We're interested in the array $\{M_n^{-1}(X_t u_t - E X_t u_t)\}$, so we will check for the conditions

(i) $\{M_n^{-1}(X_t u_t - E X_t u_t)\}$ is a measurable function of $\{X_t', u_t\}$ so by Proposition 1 it's α-mixing of size $-p/p-2$.

$E X_t X_t'$ non random

(ii) $$E\| M_n^{-1}(X_t u_t - E X_t u_t)\|^p \leq \|M_n^{-1}\| \; E\|X_t u_t - E X_t u_t\|^p$$

$$\leq O(1) \quad O(1)$$

$$< \infty$$

(iii) $$\mathrm{Var}\left(\frac{1}{\sqrt{n}} \sum_{t=1}^{n} M_n^{-1}(X_t u_t - E X_t u_t)\right) = M_n^{-1} \Omega_n M_n^{-1}$$ must be uniformly positive definite.

This requires that for arbitrary x s.t. $\|x\| = 1$

$$x' M_n^{-1} \Omega_n M_n^{-1} x > 0$$

$$
\begin{aligned}
x' M_n^{-1} \Omega_n M_n^{-1} x &= \underbrace{x' M_n^{-1} C_n}_{y_n'} \Delta_n \underbrace{C_n M_n^{-1} x}_{y_n} \\
&= y_n' \Delta_n y_n \\
&= \sum_{i=1}^{k} e_{ni} \, y_{ni}^2 \\
&\geq \underline{e_n} \, \| y_n \|^2 \\
&\geq \delta \, x' M_n^{-1} \underbrace{C_n C_n'}_{= I_k} M_n^{-1} x \\
&= \delta \, x' M_n^{-2} x \\
&= \delta \, x' D_n P_n^{-2} D_n x \\
&= \delta \sum_{i=1}^{k} d_{ni}^{-2} \, w_{ni}^2 \\
&\geq \frac{\delta}{\bar{d}_n^2} \, \| x' D_n \| \\
&= \frac{\delta}{\bar{d}_n^2} \\
&> 0
\end{aligned}
$$

because

$$
\begin{aligned}
\| M_n \| &\leq \frac{1}{n} \sum_{t=1}^{n} E \| X_t X_t' \| \\
&\leq \frac{1}{n} \sum_{t=1}^{n} E \| X_t \|^2 \\
&\leq \Delta < \infty
\end{aligned}
$$

Then, by the CLT $\quad V_n^{-1/2} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} M_n^{-1} (X_t u_t - E X_t u_t) \} \xrightarrow{d} N(0, I)$.

Putting it all together yields:

$$
V_n^{-1/2} \sqrt{n} (\hat{\beta}_n - \beta) = V_n^{-1/2} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} M_n^{-1} (X_t u_t - E X_t u_t) + V_n^{-1/2} o_p(1)
$$

bounded b.c. $Var(\frac{1}{\sqrt{n}} \sum \xi_{nt})$ is unif. p.d.

$$
\begin{aligned}
V_n^{-1/2} \sqrt{n} (\hat{\beta}_n - \beta) &= V_n^{-1/2} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} M_n^{-1} (X_t u_t - E X_t u_t) \} + O(1) \, o_p(1) \\
&= V_n^{-1/2} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} M_n^{-1} (X_t u_t - E X_t u_t) \} + o_p(1) \\
&\xrightarrow{d} N(0, I_k). \quad \blacksquare
\end{aligned}
$$

Estimation of Asymptotic Variance Matrix

Recall that $V_n = M_n^{-1} \Omega_n M_n^{-1}$ and $M_n = \frac{1}{n} \sum_{t=1}^{n} E X_t X_t'$. Then we can estimate M_n using

$$\hat{M_n} = \frac{1}{n} \sum^{n} X_t X_t' \quad \text{and hope that} \quad \hat{M_n} - M_n = o_p(1).$$

To estimate Ω_n we need $\Omega_n(h) = \frac{1}{n} \sum_{t=h+1}^{n} \underbrace{E[X_t u_t (X_{t-h} u_{t-h})']}_{\text{because they are centered at } 0}$.

Now, our initial estimator could be

$$\tilde{\Omega_n} = \tilde{\Omega_n}(0) + \sum_{h=1}^{n-1} (\tilde{\Omega_n}(h) + \tilde{\Omega_n}(h)')$$

$$\text{where} \quad \tilde{\Omega_n}(h) = \frac{1}{n} \sum_{t=h+1}^{n} [X_t u_t (X_{t-h} u_{t-h})']$$

- Problem: we need to ensure that $(\tilde{\Omega_n}(h) + \tilde{\Omega_n}(h)')$ grow slower than n. A solution would be to allow for autocovariances to grow slower than n.

- New problem: when we truncate we can get non positive definite matrix, so we need to put weights in the sum.

The (infeasible) HAC estimator of variance is

(*) Infeasible: use true u_t
Feasible: use $\hat{u}_t := Y_t - X_t'\hat{\beta}$

$$\tilde{\Omega_n} = \tilde{\Omega_n}(0) + \sum_{j=1}^{m_n} w(j, m_n) (\tilde{\Omega_n}(j) + \tilde{\Omega_n}(j)')$$

Proposition HAC 1.- Suppose that for some $p > 2$ and $\Delta, \delta, C > 0$

(a) $\{(X_t', u_t)\}$ is α-mixing of size $-p/p-2$

(b) $E X_t u_t = 0$ for all t

(c) $E|X_{tj}|^{4p+\delta} \leq \Delta$ for all t and all $j = 1, \dots, K$

(d) $E|u_t|^{4p+\delta} \leq \Delta$ for all t

(e) $|w(j,m)| \leq C$ for all j and m

(g) $\lim_{m \to \infty} w(j,m) = 1$ for all j

(h) $m_n = o(n^{1/4})$.

Then

$$\tilde{\Omega_n} - \Omega_n = o_p(1).$$

<u>proof:</u> By the Cramer-Wold device it suffices to show that

$$c'(\hat{\Omega}_n - \Omega_n)c = o_p(1) \quad \text{for all } c \in \mathbb{R}^k.$$

Now, define $h_t = c' X_t u_t$ and notice that by Proposition 1 it is α-mixing of size $-p/p-2$. Then we write

$$c'(\hat{\Omega}_n - \Omega_n)c = \underbrace{\frac{1}{n}\sum_{t=1}^{n}(h_t - Eh_t)}_{R_{n,0} = o_p(1) \text{ by LLN}} + \underbrace{2\sum_{j=1}^{m_n} w(j,m_n)\frac{1}{n}\sum(h_t h_{t-j} - Eh_t h_{t-j})}_{R_{n,1} := \text{regular estimation error of covariances}} + \underbrace{2\sum_{j=1}^{m_n}(w(j,m_n)-1)\frac{1}{n}\sum_{t=h+1}^{n} Eh_t h_{t-j}}_{R_{n,2} := \text{bias due to using weights}}$$

$$\underbrace{-\,2\sum_{j=m_n+1}^{n-1}\frac{1}{n}\sum_{t=h+1}^{n} Eh_t h_{t-j}}_{R_{n,3} := \text{bias due to truncation of autocovariances}}$$

- $R_{n,2}$: We want to use the covariance inequalities, so we need to show that $\sup_t E|h_t|^p < \infty$. To see this

$$E|h_t|^p \leq \|c\| \; E\|X_t u_t\|^p$$

$$\leq \|c\| \left(E\|X_t\|^{2p}\right)^{1/2} \left(E\|u_t\|^{2p}\right)^{1/2}$$

$$< \infty$$

Then

$$|R_{n,2}| \leq \sum_{j=1}^{m_n} |w(j,m_n) - 1| \frac{1}{n}\sum_{t=h+1}^{n} Eh_t h_{t-j}$$

$$\underset{\text{Proposition 4}}{\leq} \sum_{j=1}^{m_n} |w(j,m_n) - 1| \frac{1}{n}\sum_{t=h+1}^{n} K\,\alpha(j)^{1-2/p}$$

$$= \sum_{j=1}^{m_n} |w(j,m_n) - 1| \frac{1}{n}\sum_{t=h+1}^{n} K\, j^{\left(-\frac{p}{p-2}-\epsilon\right)\left(\frac{p-2}{p}\right)}$$

$$= \sum_{j=1}^{m_n} |w(j,m_n) - 1| \frac{1}{n}\sum_{t=h+1}^{n} K\, j^{-1-\eta} \qquad \text{for some } \eta > 0$$

$$\leq \sum_{j=1}^{m_n} |w(j,m_n) - 1| \frac{1}{n} K\, j^{-1-\eta}\,(n-(h+1)+1)$$

$$\leq K \sum_{j=1}^{\infty} |w(j,m_n) - 1|\, j^{-1-\eta}$$

$$\lim_{n\to\infty} |R_{n,2}| \leq \lim_{n\to\infty} K \sum_{j=1}^{\infty} |w(j,m_n) - 1|\, j^{-1-\eta}$$

Dominated Convergence Theorem

$$= K \sum_{j=1}^{\infty} \left| \lim_{n\to\infty} w(j,m_n) - 1 \right| j^{-1-\eta}$$

$K(C+1)\, j^{-1-\eta}$ can be the dominating function and it's integrable/summable.

$$= 0 .$$

- $R_{n,3}$: We will use the same idea as in the previous part.

$$|R_{n,3}| \le \sum_{j=m_n+1}^{n-1} \frac{1}{n} \sum_{t=1}^{n} |E h_t h_{t-j}|$$

$$\underset{\text{Using bounds computed in } R_{n,2}}{\le} K \sum_{j=m_n+1}^{n-1} j^{-1-\eta}$$

$$\le -\frac{1}{\eta} n^{-\eta} K + \frac{K}{\eta} m_n^{-\eta}$$

$$\lim_{n\to\infty} |R_{n,3}| \le \lim_{n\to\infty} -\frac{K}{\eta} n^{-\eta} + \lim_{n\to\infty} \frac{K}{\eta} m_n^{-\eta}$$

$$= 0 .$$

- $R_{n,1}$: before we deal with this I will write this term again to see why this can be difficult to check.

$$R_{1,n} := \sum_{j=1}^{m_n} w(j,m_n) \frac{1}{n} \sum_{t=j+1}^{n} \underbrace{(h_t h_{t-j} - E h_t h_{t-j})}$$

Call this process Z_{jt}. Moreover, notice that $Z_{jt} = g(h_t, h_{t-j})$ so by Proposition 1 it is α-mixing and $\alpha_{Z_j}(\ell) \le \underbrace{\alpha_h(\ell - j)}$ for all $\ell = j+1, j+2, \dots$

This number must be positive!

You will see why this is important later. Mark this as (*).

We want to show that the object is $o_P(1)$, so we write

$$P\left(\left| \underbrace{\sum_{j=1}^{m_n} w(j,m_n)}_{\le C} \frac{1}{n} \sum_{t=j+1}^{n} Z_{jt} \right| > \varepsilon \right) \le P\left(\sum_{j=1}^{m_n} \left| \sum_{t=j+1}^{n} Z_{jt} \right| > \frac{\varepsilon \cdot n}{C} \right)$$

$$\le P\left(\left| \sum_{t=1+1}^{n} Z_{1t} \right| > \frac{\varepsilon . n}{C \cdot m_n} \right) + \dots + P\left(\left| \sum_{t=m_n+1}^{n} Z_{m_n t} \right| > \frac{\varepsilon \cdot n}{C \cdot m_n} \right)$$

$$= \sum_{j=1}^{m_n} P\left(\left| \sum_{t=j+1}^{n} Z_{jt} \right| > \frac{\varepsilon \cdot n}{C \cdot m_n} \right)$$

$$\underset{\text{Markov's Inequality}}{\le} \sum_{j=1}^{m_n} \frac{C^2 m_n^2}{\varepsilon^2 n^2} E \left| \sum_{t=j+1}^{n} Z_{jt} \right|^2$$

Claim 1.- If $E|\sum^{n} z_{jt}|^2 \leq K \cdot n \cdot (j+2)$ then $R_{1,n} = o_p(1)$.

Using this claim we get

$$\leq \sum_{j=1}^{m_n} \frac{c^2 m_n^2}{\varepsilon^2 n^2} K \cdot n \cdot (j+2)$$

$$= K \frac{c^2 m_n^2}{n \varepsilon^2} \sum_{j=1}^{m_n} (j+2)$$

$\frac{1}{m_n} \sum_{j=1}^{m_n} (j+2) = \frac{\text{first} + \text{last}}{2}$

$$= K \frac{c^2 m_n^2}{n \varepsilon^2} \left[\frac{(m_n+2)+3}{2} \right] m_n$$

$$\leq K \frac{m_n^4}{n}$$

$$= K \frac{1}{n} o(n)$$

$$= o(1).$$

To finish the proof we only need to show that the assumption for Claim 1 is true. write

$$E|\sum^{n} z_{jt}|^2 = \underbrace{\sum_{t=j+1}^{n} E z_{jt}^2}_{\text{Variance formula}} + 2 \sum_{\ell=1}^{n-j-1} \sum_{t=\ell+j+1}^{n} E(z_{jt} z_{jt-\ell})$$

$$\leq \sum_{t=j+1}^{n} \sup_t E|h_t|^4 + 2 \sum_{\ell=1}^{n-j-1} \sum_{t=\ell+j+1}^{n} E(z_{jt} z_{jt-\ell})$$

$$\underset{\text{Cauchy Schwarz}}{\leq} \sum_{t=j+1}^{n} \left(\sup_t E|h_t|^8 \cdot \sup_t E|c' x_t|^8 \right)^{1/2} + 2 \sum_{\ell=1}^{n-j-1} \sum_{t=\ell+j+1}^{n} E(z_{jt} z_{jt-\ell})$$

$$\leq K \cdot n + 2 \sum_{\ell=1}^{n-j-1} \sum_{t=\ell+j+1}^{n} E(z_{jt} z_{jt-\ell})$$

split the sum

$$= K \cdot n + 2 \underbrace{\sum_{\ell=1}^{j} \sum_{t=\ell+j+1}^{n} |E z_{jt} z_{jt-\ell}|}_{\substack{\text{we cannot use mixing coefficient properties here} \\ \text{but we can use Cauchy Schwarz, recall (*)}}} + 2 \underbrace{\sum_{\ell=j+1}^{n-j-1} \sum_{t=j+1+\ell}^{n} |E z_{jt} z_{jt-\ell}|}_{\substack{\text{we can use mixing coefficients properties here because} \\ \ell \geq j+1 \text{, recall (*)}}}$$

Cauchy Schwarz + Proposition 4

$$\leq K \cdot n + 2 \sum_{\ell=1}^{j} \sum_{t=\ell+j+1}^{n} \underbrace{\left(E|z_{jt}|^2 \, E|z_{jt-\ell}|^2 \right)^{1/2}}_{\text{we just showed these are bounded}} + 2 \sum_{\ell=1}^{n-j-1} (n-\ell-j) \, \ell^{-1-\eta}$$

$$\leq Kn + K \cdot n \cdot j + 2n \sum_{\ell=1}^{n-j-\ell} \ell^{-1-\eta}$$

By summability

$$\leq K \cdot n + K \cdot n \cdot j + K \cdot n = K \cdot n (j+2). \blacksquare$$

Proposition HAC 2. - Suppose that for some $p > 2$ and $\Delta, \delta, C > 0$

(a) $\{(X_t', u_t)\}$ is α-mixing of size $-p/p-2$

(b) $E X_t u_t = 0$ for all t

(c) $E|X_{tj}|^{4p+\delta} \le \Delta$ for all t and all $j = 1, \dots, K$

(d) $E|u_t|^{4p+\delta} \le \Delta$ for all t

(e) $|w(j,m)| \le C$ for all j and m

(g) $\lim_{m\to\infty} w(j,m) = 1$ for all j

(h) $m_n = o(n^{1/4})$.

Then

$$\hat{\Omega}_n - \Omega_n = o_p(1).$$

where $\hat{\Omega}_n = \hat{\Omega}_n(0) + \sum_{j=1}^{m_n} w(j, m_n)\,(\hat{\Omega}_n(j) + \hat{\Omega}_n(j)')$ is the feasible estimator.

$$\hat{\Omega}_n(j) = \frac{1}{n}\sum_{t=j+1}^{n} (X_t \hat{u}_t)(X_{t-j}\hat{u}_{t-j})'$$

proof:

$$\hat{\Omega}_n - \Omega_n = \underbrace{\hat{\Omega}_n - \tilde{\Omega}_n}_{\text{we just need to show that this is } o_p(1).} + \underbrace{\tilde{\Omega}_n - \Omega_n}_{o_p(1) \text{ by Proposition HAC 1}}$$

Recall that $\hat{u}_t = u_t - X_t'(\hat{\beta}_n - \beta)$. We now write

$$\hat{\Omega}_n - \tilde{\Omega}_n = \left(\frac{1}{n}\sum_{t=1}^{n} \hat{u}_t^2 X_t X_t' + \sum_{j=1}^{m_n} w(j,m_n)\frac{1}{n}\sum \hat{u}_t \hat{u}_{t-j}\,(X_t X_{t-j}' + X_{t-j}X_t')\right)$$

$$-$$

$$\left(\frac{1}{n}\sum_{t=1}^{n} u_t^2 X_t X_t' + \sum_{j=1}^{m_n} w(j,m_n)\frac{1}{n}\sum u_t u_{t-j}\,(X_t X_{t-j}' + X_{t-j}X_t')\right)$$

$$= \underbrace{-2\,\frac{1}{n}\sum_{t=1}^{n}\left((\hat{\beta}_n-\beta)' X_t u_t\right) X_t X_t'}_{B_{n,1}} + \underbrace{\frac{1}{n}\sum_{t=1}^{n}\left((\hat{\beta}_n-\beta)' X_t\right)^2 X_t X_t'}_{B_{n,2}}$$

$$- \underbrace{\sum_{j=1}^{m_n} w(j,m_n)\frac{1}{n}\sum_{t=j+1}^{n}\left((\hat{\beta}_n-\beta)' X_t u_{t-j}\right)(X_t X_{t-j}' + X_{t-j}X_t')}_{B_{n,3}}$$

$$- \underbrace{\sum_{j=1}^{m_n} w(j,m_n)\frac{1}{n}\sum_{t=j+1}^{n}\left((\hat{\beta}_n-\beta)' X_{t-j} u_t\right)(X_t X_{t-j}' + X_{t-j}X_t')}_{B_{n,4}}$$

$$+ \underbrace{\sum_{j=1}^{m_n} w(j,m_n) \frac{1}{n} \sum_{t=j+1}^{n} \left((\hat{\beta}_n - \beta)' x_t \right) \left((\hat{\beta}_n - \beta)' x_{t-j} \right) \left(x_t x_{t-j}' + x_{t-j} x_t' \right)}_{B_{n,5}}$$

Then by the cramer wold device we will work with $c' (\hat{\Omega}_n - \tilde{\Omega}_n) c$ for $c \in \mathbb{R}^k$. We will work with each term separately.

- $B_{n,1}$:

$$|c' B_{n,1} c| = \left| \frac{1}{n} \sum^{n} \left((\hat{\beta}_n - \beta)' x_t u_t \right) c' x_t x_t' c \right|$$

$$\leq \|\hat{\beta}_n - \beta\| \frac{1}{n} \sum_{t=1}^{n} \|x_t\|^3 \|c\|^2 |u_t|$$

$$\underset{\substack{\text{using} \\ \|c\| = O(1)}}{=} o_p(1)\, O(1) \left[\frac{1}{n} \sum_{t=1}^{n} \|x_t\|^3 |u_t| - \frac{1}{n} \sum_{t=1}^{n} E\|x_t\|^3 |u_t| + \frac{1}{n} \sum_{t=1}^{n} E\|x_t\|^3 |u_t| \right]$$

$$= o_p(1)\, O(1) \left[\frac{1}{n} \sum_{t=1}^{n} E\|x_t\|^3 |u_t| + o_p(1) \right]$$

$$\underset{\substack{\text{Hölder inequality} \\ p=4}}{\leq} o_p(1)\, O(1) \left[\frac{1}{n} \sum_{t=1}^{n} \left(E|u_t|^4 \right)^{1/4} \left(E\|x_t\|^4 \right)^{3/4} + o_p(1) \right]$$

$$= o_p(1)\, O(1) \left[O(1) + o_p(1) \right]$$

$$= o_p(1).$$

- $B_{n,2}$:

$$|c' B_{n,2} c| = \left| \frac{1}{n} \sum_{t=1}^{n} \left((\hat{\beta}_n - \beta)' x_t \right)^2 c' x_t x_t' c \right|$$

$$\leq \|c\|^2 \|\hat{\beta}_n - \beta\|^2 \left[\frac{1}{n} \sum_{t=1}^{n} \|x_t\|^4 - \frac{1}{n} \sum_{t=1}^{n} E\|x_t\|^4 + \frac{1}{n} \sum_{t=1}^{n} E\|x_t\|^4 \right]$$

$$= O(1)\, o_p(1) \left[o_p(1) + O(1) \right]$$

$$= o_p(1).$$

- $B_{n,3}$:

$$|c' B_{n,3} c| \leq 2 \|c\|^2 \|\hat{\beta}_n - \beta\| \sum_{j=1}^{m_n} w(j,m_n) \frac{1}{n} \sum_{t=j+1}^{n} \left(|u_{t-j}| \cdot \|x_t\|^2 \cdot \|x_{t-j}\| \right)$$

$$= o_p(1)\sum_{j=1}^{m_n}\omega(j,m_n)\,\frac{1}{n}\sum_{t=j+}^{n}\Big\{\underbrace{|u_{t-j}|\cdot\|x_t\|^2\|x_{t-j}\| - E|u_{t-j}|\cdot\|x_t\|^2\|x_{t-j}\|}_{\text{Define as } Z_{j,t} \text{ and notice that it's } \alpha\text{-mixing of size } \frac{-p}{p-2}} + E|u_{t-j}|\,\|x_t\|^2\|x_{t-j}\|\Big\}$$

$R_{n,1}$ in Proposition HAC 1 + Cauchy Schwarz

$$\le O_p(1)\sum_{j=1}^{m_n}\omega(j,m_n)\,\frac{1}{n}\sum_{t=j+}^{n}\Big\{o_p(1) + \left(E|u_{t-j}|^2\|x_{t-j}\|^2\,E\|x_t\|^4\right)^{1/2}\Big\}$$

Cauchy Schwarz

$$\le O_p(1)\sum_{j=1}^{m_n}\omega(j,m_n)\,\frac{1}{n}\sum_{t=j+}^{n}\Big\{o_p(1) + \left(\left(E|u_{t-j}|^4\,E\|x_{t-j}\|^4\right)^{1/2}E\|x_t\|^4\right)^{1/2}\Big\}$$

$$\le O_p(1)\quad O(m_n)$$

we need to use $\|\hat\beta_n - \beta\|$

$$\le \|V_n\|^{1/2}\,\|V_n^{-1/2}\sqrt{n}(\hat\beta_n-\beta)\|\quad\frac{O(m_n)}{\sqrt{n}}$$

$$\le O(1)\quad O_p(1)\quad\frac{o(n^{1/4})}{n^{1/2}}$$

$$= K\quad O_p(1)\quad o(n^{-1/4})$$

$$= o_p(1).$$

- $B_{n,4}$:

$$|c'B_{n,4}c| \le 2\|c\|^2\,\|\hat\beta_n-\beta\|\sum_{j=1}^{m_n}\omega(j,m_n)\,\frac{1}{n}\sum_{t=j+1}^{n}|u_{t-j}|\,\|x_t\|\,\|x_{t-j}\|^2$$

$$= o_p(1)\sum_{j=1}^{m_n}\omega(j,m_n)\,\frac{1}{n}\sum_{t=j+}^{n}\Big\{\underbrace{|u_{t-j}|\cdot\|x_t\|\cdot\|x_{t-j}\|^2 - E|u_{t-j}|\cdot\|x_t\|\,\|x_{t-j}\|^2}_{\text{Define as } Z_{j,t} \text{ and notice that it's } \alpha\text{-mixing of size } \frac{-p}{p-2}} + E|u_{t-j}|\,\|x_t\|\,\|x_{t-j}\|^2\Big\}$$

$$= o_p(1).$$

By same steps as $B_{n,3}$

- $B_{n,5}$:

$$|c'B_{n,5}c| \le 2\|c\|^2\,\|\hat\beta_n-\beta\|^2\sum_{j=1}^{m_n}\omega(j,m_n)\,\frac{1}{n}\sum_{t=j+1}^{n}\|x_t\|^2\,\|x_{t-j}\|^2$$

$$= o_p(1)\sum_{j=1}^{m_n}\omega(j,m_n)\,\frac{1}{n}\sum_{t=j+}^{n}\Big\{\underbrace{\|x_t\|\cdot\|x_{t-j}\|^2 - E\,\|x_t\|^2\,\|x_{t-j}\|^2}_{\text{Define as } Z_{j,t} \text{ and notice that it's } \alpha\text{-mixing of size } \frac{-p}{p-2}} + E\,\|x_t\|^2\,\|x_{t-j}\|^2\Big\}$$

$$= o_p(1).$$

By same steps as $B_{n,3}$

■

Block / Cluster Dependence (Hansen, JoE 2020)

Consider the sample mean $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ as an estimator of $\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}X_i$

Define cluster sums

$$\tilde{X}_g = \sum_{j=1}^{n} X_{gj}$$

mutually independent under clustered sampling for $g \neq g'$.

We may rewrite the sample mean as

$$\bar{X}_n = \frac{1}{n}\sum_{g=1}^{G} \tilde{X}_g$$

Assumption 1.- As $n \to \infty$

$$\max_{g \leq G} \frac{n_g}{n} \to 0$$

(i.e. n_g is asymptotically negligible so implicitly $G \to \infty$)

$$= \left(\max_{g \leq G} \frac{n_g^2}{n^2} \right)^{1/2}$$

$$\leq \left(\sum_{g=1}^{G} \frac{n_g^2}{n^2} \right)^{1/2}$$

Theorem 1.- If A1 holds and

$$\lim_{M \to \infty} \sup_i \left(\underbrace{\mathbb{E}\|X_i\| \mathbb{1}(\|X_i\| > M)}_{} \right) = 0$$

$$\leq (\mathbb{E}\|X_i\|^p)^{1/p} \left(P(\|X_i\| > M)\right)^{1/q} \quad , \frac{1}{p}+\frac{1}{q}=1, \; p>1$$

$$\leq (\mathbb{E}\|X_i\|^p)^{1/p} \left(\frac{\mathbb{E}\|X_i\|^p}{M^p} \right)^{1/q}$$

so $\sup_i \mathbb{E}\|X_i\|^p < \infty$ is sufficient, $p > 1$.

Then, as $n \to \infty$

$$\|\bar{X}_n - \mathbb{E}\bar{X}_n\| \xrightarrow{p} 0.$$

Lemma 1.- For random vectors X_i, set $\tilde{X}_m = \sum_{i=1}^{m} X_i$. For $r \geq 1$ if

$$\lim_{B \to \infty} \sup_i \mathbb{E}\left(\|X_i\|^r \mathbb{1}\{\|X_i\| > B\} \right) = 0$$

then

$$\lim_{B \to \infty} \sup_m \mathbb{E}\left(\|m^{-1}\tilde{X}_m\|^r \mathbb{1}\{\|m^{-1}\tilde{X}_m\| > B\} \right) = 0$$

<u>Proof of Lemma:</u>

$$\lim_{B\to\infty} \sup_i E\left(\|X_i\|^r \mathbb{1}\{\|X_i\| > B\} \right) = 0 \iff \sup_i E\|X_i\|^r \le C, \quad r > 1.$$

By Cr inequality

$$\left\| \frac{1}{m} \tilde{X}_m \right\|^r = \frac{1}{m^r} \left\| \sum_{i=1}^m x_i \right\|^r \le \frac{1}{m} \sum_{i=1}^m \|X_i\|^r$$

Then

$$E\| m^{-1} \tilde{X}_m \|^r \le \frac{1}{m} \sum_{i=1}^m E\|X_i\|^r \le C.$$

Write

$$E\left(\|m^{-1} \tilde{X}_m\|^r \mathbb{1}\{\| m^{-1} \tilde{X}_m \| > B\} \right)$$

$$\le \frac{1}{m} \sum_{i=1}^m E\left(\|X_i\|^r \mathbb{1}\{\| m^{-1} \tilde{X}_m \|\} > B\} \right)$$

$$= \frac{1}{m} \sum_{i=1}^m E\left(\|X_i\|^r \mathbb{1}\{\| m^{-1} \tilde{X}_m \|\} > B\} \mathbb{1}\{\|X_i\| > \sqrt{B}\} \right)$$

$$+$$

$$\frac{1}{m} \sum_{i=1}^m E\left(\|X_i\|^r \mathbb{1}\{\| m^{-1} \tilde{X}_m \|\} > B\} \mathbb{1}\{\|X_i\| \le \sqrt{B}\} \right)$$

$$\le \frac{1}{m} \sum_{i=1}^m E\left(\|X_i\|^r \mathbb{1}\{\|X_i\| > \sqrt{B}\} \right) + B^{r/2} E \mathbb{1}\{\| m^{-1} \tilde{X}_m \| > B\}$$

$$\le \underbrace{\frac{1}{m} \sum_{i=1}^m E\left(\|X_i\|^r \mathbb{1}\{\|X_i\| > \sqrt{B}\} \right)}_{\text{want this } < \varepsilon/2} + \underbrace{B^{r/2} \frac{E\| m^{-1} \tilde{X}_m \|^r}{B^r}}_{\text{want this } < \varepsilon/2}$$

$\Downarrow$

$B^{r/2} \ge 2C/\varepsilon$ does the work

$$\le \frac{\varepsilon}{2} + \frac{\varepsilon}{2}. \qquad \blacksquare$$

by assumption we can find B large enough

<u>proof of Theorem 1:</u> Without loss of generality assume $E x_i = 0$.

By Lemma 1 under $r=1$ we can pick B such that

$$\sup_g E \| (n_g^{-1} \tilde{X}_g \mathbb{1}\{\|n_g^{-1} \tilde{X}_g\| > B\}) - E(n_g^{-1} \tilde{X}_g \mathbb{1}\{\|n_g^{-1} \tilde{X}_g\| > B\}) \| \le \varepsilon \quad (*)$$

Notice that

$$P(\|\bar{X}_n - E\bar{X}_n\| > \varepsilon) \le \frac{2E\|\bar{X}_n\|}{\varepsilon} \quad \text{by Markov}$$

Then, write

$$\begin{aligned}
E\|\bar{X}_n\| &= E\left\| \frac{1}{n} \sum_{g=1}^{G} \tilde{X}_g \right\| \\
&\le E\left\| \frac{1}{n} \sum_{g=1}^{G} \left[\tilde{X}_g \mathbb{1}(\|n_g^{-1} \tilde{X}_g\| \le B) - E(\tilde{X}_g \mathbb{1}(\|n_g^{-1} \tilde{X}_g\| \le B)) \right] \right\| \\
&\quad + \frac{1}{n} \sum_{g=1}^{G} E\left\| \left[\tilde{X}_g \mathbb{1}(\|n_g^{-1} \tilde{X}_g\| > B) - E(\tilde{X}_g \mathbb{1}(\|n_g^{-1} \tilde{X}_g\| > B)) \right] \right\| \\
&\overset{\text{by } (*)}{\le} E\left\| \frac{1}{n} \sum_{g=1}^{G} \left[\tilde{X}_g \mathbb{1}(\|n_g^{-1} \tilde{X}_g\| \le B) - E(\tilde{X}_g \mathbb{1}(\|n_g^{-1} \tilde{X}_g\| \le B)) \right] \right\| \\
&\quad + \frac{1}{n} \sum_{g=1}^{G} \varepsilon n_g \\
&\overset{\text{norm inequality}}{\le} \left(E\left\| \frac{1}{n} \sum_{g=1}^{G} \left[\tilde{X}_g \mathbb{1}(\|n_g^{-1} \tilde{X}_g\| \le B) - E(\tilde{X}_g \mathbb{1}(\|n_g^{-1} \tilde{X}_g\| \le B)) \right] \right\|^2 \right)^{1/2} \\
&\quad + \varepsilon \\
&\overset{\text{by cluster uncorrelatedness}}{=} \left(\frac{1}{n^2} \sum_{g=1}^{G} E \left\| \tilde{X}_g \mathbb{1}\{\|n_g^{-1} \tilde{X}_g\| \le B\} - E(\tilde{X}_g \mathbb{1}\{\|n_g^{-1} \tilde{X}_g\| \le B\}) \right\|^2 \right)^{1/2} + \varepsilon \\
&\le \left(\frac{1}{n^2} \sum_{g=1}^{G} (2 B n_g)^2 \right)^{1/2} + \varepsilon \\
&= \left(4B^2 \sum_{g=1}^{G} \frac{n_g^2}{n^2} \right)^{1/2} + \varepsilon \le o(1) + \varepsilon \quad \text{by (A1)}.
\end{aligned}$$

CLT

$$\Omega_n := \mathrm{Var}(\sqrt{n}\,\bar{X}_n) = E\left(n(\bar{X}_n - E\bar{X}_n)(\bar{X}_n - E\bar{X}_n)' \right)$$
$$= \frac{1}{n} \sum_{g=1}^{G} E\left((X_g^{\sim} - EX_g^{\sim})(X_g^{\sim} - EX_g^{\sim})' \right)$$

Assumption 2.- For some $2 \le r < \infty$

- $\frac{1}{n} \left(\sum_{g=1}^{G} n_g^r \right)^{1/2} \le C < \infty$
- $\max_{g \le G} \frac{n_g^2}{n} \to 0$ as $n \to \infty$.

Theorem 2.- If for some $2 \le r < \infty$ (A2) holds and

(i) $\lim_{M \to \infty} \sup_i \left(E\|X_i\|^r \mathbb{1}\{\|X_i\| > M\} \right) = 0$

(ii) $\lambda_n = \lambda_{min}(\Omega_n) \ge \lambda > 0$.

Then as $n \to \infty$

$$\Omega_n^{-1/2} \sqrt{n}(\bar{X}_n - E\bar{X}_n) \xrightarrow{d} N(0, I_p)$$

proof: WLOG assume $EX_i = 0$. Note that

$$\Omega_n^{-1/2} \sqrt{n}\,\bar{X}_n = \Omega_n^{-1/2} \sum_{g=1}^{G} \frac{1}{\sqrt{n}} \tilde{X}_g$$

inid random vectors

We can apply Lindeberg-Feller multivariate CLT if Lindeberg's condition holds

$$\frac{1}{n\lambda_n} \sum_{g=1}^{G} E\left(\|\tilde{X}_g\|^2 \mathbb{1}\{\|\tilde{X}_g\|^2 \ge n\lambda_n \varepsilon\} \right) \to 0 \text{ as } n \to \infty.$$

Fix $\varepsilon > 0$ and $\delta > 0$ and pick B large by Lemma 1 such that

$$\sup_g E\left(\|n_g^{-1} \tilde{X}_g\|^r \mathbb{1}(\|n_g^{-1}\tilde{X}_g\| > B) \right) \le \frac{\delta \varepsilon^{r/2-1}}{C^{r/2}}$$

By A2 $\quad \max_{g \le G} \frac{1}{\lambda_n} \frac{n_g^2}{n} \le \frac{1}{\lambda} \max_{g \le G} \frac{n_g^2}{n} = o(1)$

Then we pick n large enough so that $\quad \max_{g \le G} \frac{n_g^2}{n\lambda_n} \le \frac{\varepsilon}{B^2}$

Thus

$$\frac{1}{n\lambda_n} \sum_{g=1}^{G} E\left(\|\tilde{X}_g\|^2 \mathbb{1}\{ \|\tilde{X}_g\|^2 \geq n\lambda_n \varepsilon \} \right)$$

$$= \frac{1}{n\lambda_n} \sum_{g=1}^{G} E\left(\frac{\|\tilde{X}_g\|^r}{\|\tilde{X}_g\|^{r-2}} \mathbb{1}\{ \|\tilde{X}_g\| \geq (n\lambda_n \varepsilon)^{1/2} \} \right)$$

$$\leq \frac{1}{n\lambda_n} \cdot \frac{1}{(n\lambda_n\varepsilon)^{\frac{r-2}{2}}} \sum_{g=1}^{G} E\left(\|\tilde{X}_g\|^r \mathbb{1}\{ \|\tilde{X}_g\| \geq (n\lambda_n \varepsilon)^{1/2} \} \right)$$

$$\leq \frac{1}{\varepsilon^{r/2-1} (n\lambda_n)^{r/2}} \sum_{g=1}^{G} n_g^r E\left(\|n_g^{-1} \tilde{X}_g\|^r \mathbb{1}\{ \|n_g^{-1} \tilde{X}_g\| \geq B \} \right)$$

$$\leq \frac{\delta}{C^{r/2}} \sum_{g=1}^{G} \frac{n_g^r}{(n\lambda_n)^{r/2}}$$

$\leq \delta$ and the proof is complete because δ can be arbitrarily small. ∎

$$\underset{n_g\times 1}{y_g} = \underset{n_g\times k}{x_g}' \underset{k\times 1}{\beta} + \underset{n_g\times 1}{u_g}$$

Now, consider a case where X_i is mean zero. Then

$$\Omega_n = \frac{1}{n} \sum_{g=1}^{G} E(\tilde{X}_g \tilde{X}_g') \qquad G \to \infty$$

and we can estimate

$$\tilde{\Omega}_n = \frac{1}{n} \sum_{g=1}^{G} \tilde{x}_g \tilde{x}_g'$$

⊛ Heteroskedasticity

$$\frac{1}{n} \sum_{i=1}^{n} E\left((X_i u_i)(X_i u_i)' \right)$$

Cluster

$$\frac{1}{n} \sum_{g=1}^{G} E\left((\underset{k\times n_g}{x_g} \underset{n_g\times 1}{u_g})(x_g u_g)' \right)$$

Still white approach!

Theorem 3.- Under Theorem 2 if $EX_i = 0$. Then as $n \to \infty$

$$\Omega_n^{-1/2} \tilde{\Omega}_n \Omega_n^{-1/2} \xrightarrow{p} I_p$$

$$\sum_{g=1}^{G} (\Omega_n^{-1/2} \tilde{X}_g)(\Omega_n^{-1/2} \tilde{X}_g)'$$

and

$$\tilde{\Omega}_n^{-1/2} \sqrt{n}\, \bar{X}_n \xrightarrow{d} N(0, I_p).$$

proof: Fix $\delta > 0$ and set $\varepsilon = \delta^2/4p$. Define $\tilde{X}_g^* = \Omega_n^{-1/2} \tilde{X}_g$ and

$$\tilde{Y}_g = \tilde{X}_g \mathbb{1}\{ \|\tilde{X}_g^*\|^2 \leq n\varepsilon \}$$

$$\tilde{\Omega}_n^* = \frac{1}{n} \sum_{g=1}^{G} \tilde{X}_g^* \tilde{X}_g^{*\prime}$$

$$= \frac{1}{n} \sum_{g=1}^{G} \tilde{Y}_g \tilde{Y}_g' + \frac{1}{n} \sum_{g=1}^{G} \tilde{X}_g \tilde{X}_g^{*\prime} \mathbb{1}\{ \|\tilde{X}_g^*\|^2 > n\varepsilon \}$$

Then $P\left(\| \tilde{\Omega}_n^* - I_p \| > \varepsilon \right) \leq \frac{1}{\varepsilon} \underbrace{E \| \tilde{\Omega}_n^* - I_p \|}_{\text{Bound this}}$

$$E \| \tilde{\Omega}_n^* - I_p \| \underset{\text{triangle ineq}}{\leq} \frac{1}{n} E \left\| \sum_{g=1}^{G} \left(\tilde{Y}_g \tilde{Y}_g' - E \tilde{Y}_g \tilde{Y}_g' \right) \right\| + \underbrace{\frac{2}{n} \sum_{g=1}^{G} E\left(\| \tilde{X}_g \|^2 \mathbb{1}\{ \|\tilde{X}_g\|^2 > n\varepsilon \} \right)}_{\text{From Theorem 2 we can bound an object like this for n large}}$$

$$\leq \frac{1}{n} E \left\| \sum_{g=1}^{G} \left(\tilde{Y}_g \tilde{Y}_g' - E \tilde{Y}_g \tilde{Y}_g' \right) \right\| + 2\delta$$

norm ineq

$$\leq \frac{1}{n} \left(E \left\| \sum_{g=1}^{G} \left(\tilde{Y}_g \tilde{Y}_g' - E \tilde{Y}_g \tilde{Y}_g' \right) \right\|^2 \right)^{1/2} + 2\delta$$

$$= \frac{1}{n} \left(\sum_{g=1}^{G} E \| \tilde{Y}_g \tilde{Y}_g' - E (\tilde{Y}_g \tilde{Y}_g') \|^2 \right)^{1/2} + 2\delta$$

$$\leq \frac{2}{n} \left(\sum_{g=1}^{G} E \| \tilde{Y}_g \tilde{Y}_g' \|^2 \right)^{1/2} + 2\delta$$

use $\| \tilde{Y}_g \tilde{Y}_g' \| \leq n\varepsilon$ $\Rightarrow$ multiply $\| \tilde{Y}_g \tilde{Y}_g' \|^2 \leq n\varepsilon \| \tilde{X}_g^* \|^2$

$\| \tilde{Y}_g \tilde{Y}_g' \| \leq \| \tilde{X}_g^* \|^2$

$$\leq 2 \varepsilon^{1/2} \left(\frac{1}{n} \sum_{g=1}^{G} E \| \tilde{X}_g^* \|^2 \right)^{1/2} + 2\delta$$

by cluster indep

$$= 2\varepsilon^{1/2} \left(\frac{1}{n} E \left\| \sum_{g=1}^{G} \tilde{X}_g^* \right\|^2 \right)^{1/2} + 2\delta$$

$$= 2\varepsilon^{1/2} \left(n \operatorname{Var}(\bar{X}_n^*) \right)^{1/2} + 2\delta$$

$$= 2\varepsilon^{1/2} \left(\operatorname{tr}(I_p) \right)^{1/2} + 2\delta$$

$$= \delta + 2\delta .$$

by choice of δ

The second result is

$$\tilde{\Omega}_n^{-1/2} \sqrt{n} \, \bar{X}_n = \tilde{\Omega}_n^{-1/2} \Omega_n^{1/2} \Omega_n^{-1/2} \sqrt{n} \, \bar{X}_n$$

$$= \Omega_n^{-1/4} \underbrace{\Omega_n^{1/4} \tilde{\Omega}_n^{-1/2} \Omega_n^{1/4}}_{} \Omega_n^{1/4} \Omega_n^{-1/2} \sqrt{n} \, \bar{X}_n$$

$$= \Omega_n^{-1/4} \tilde{\Omega}_n^{*\,-1/2} \Omega_n^{1/4} \quad \Omega_n^{-1/2} \sqrt{n} \, \bar{X}_n$$

$$= \left[I_p + o_p(1) \right] \Omega_n^{-1/2} \sqrt{n}\, \bar{X}_n$$

by the previous result and continuous mapping

$$= \Omega_n^{-1/2} \sqrt{n}\, \bar{X}_n + o_p(1)\, \Omega_n^{-1/2} \sqrt{n}\, \bar{X}_n$$

$$= \Omega_n^{-1/2} \sqrt{n}\, \bar{X}_n + o_p(1)\, O_p(1)$$

$$\xrightarrow{d} N(0, I_p). \quad \blacksquare$$

(*) Other option : • wild bootstrap. (Cameron)

Obtain draws (y_g^*, x_g) such that

$$y_g^* = x_g \hat{\beta} + \tilde{u}_g$$

where

$$\tilde{u}_g^* := \begin{cases} \frac{1-\sqrt{5}}{2}\, \hat{u}_g & , \text{prob } \frac{1+\sqrt{5}}{2\sqrt{5}} \\ \left(1 - \frac{1-\sqrt{5}}{2}\right) \hat{u}_g & , \text{prob } 1 - \frac{1+\sqrt{5}}{2\sqrt{5}} \end{cases}$$

Basically, we multiply $\hat{u}_g$ by a random weight using a distribution resembling white noise. The most common (shown above) comes from Mammen 1993.

• Parametric bootstrap

Draw $\tilde{u}_g$ from some parametric family (e.g. Normal $(0, 1)$).

We compute our statistic (e.g. clustered s.e) across all simulations to obtain critical values.