Condition to use panel with missing time series on Stationary series: $N > \frac{T_s}{T}$ $T_s \rightarrow$ time series max non missing; $T \rightarrow$ ~~time series~~ panel's max non miss

$\boxed{\text{non stationary Cont.}}$ k: fraction of sample have to sacrifice to use additional N cross. sec.

$$\sqrt{N} > \frac{T_s}{T} = \frac{T_s}{(1-k)T_s} = \frac{1}{1-k} \Rightarrow N > \left(\frac{1}{1-k}\right)^2 \equiv \text{same power Condition}$$

$\boxed{\text{Decision}}$

1. if interest: Correl among level Variables $\Rightarrow$ panel with more $T \equiv$ largest NT^2 rather than NT
2. interest: Correl among (quasi) diff. variables (e.g growth rate) $\Rightarrow$ Panel with higher total obs $\equiv NT$ largest

$\boxed{\text{Cover Matrix}}$

$$\Omega_A = \frac{1}{NT} E\left(\Sigma^{NT} x_{it} u_{it}\right)^2 = \frac{\Sigma x_{it}^2 u_{it}^2}{NT} + E(\text{Cross product})$$

$E(u_{it} u_{is}) \neq 0$ in serial correl. $\Rightarrow E(\text{cross}) \neq 0$

White $\Rightarrow$ $\boxed{\text{Hetroskedasticity Consistent Estimator}}$

$$\hat{\Omega}_A^2 = \frac{\Sigma^N X_i' \hat{u}_i \hat{u}_i' X_i}{N}$$

$\hat{u}_i = (\hat{u}_{i1}, \dots, \hat{u}_{iT})'$ $\quad \tilde{X}_i = (\tilde{x}_{i1}, \dots, \tilde{x}_{iT})'$

$$\Rightarrow V(\hat{b}) = \left(\Sigma_{i=1}^N \tilde{X}_i' \tilde{X}\right)^{-1} \left(\Sigma_{i=1}^N \tilde{X}_i' \hat{u}_i \hat{u}_i' \tilde{X}_i\right) \left(\Sigma_{i=1}^N \tilde{X}_i' \tilde{X}\right)^{-1}$$

$$\text{t-stat: } t_{\hat{b}} = \frac{\hat{b}}{\sqrt{\left(\Sigma_{i=1}^N \tilde{X}_i' \tilde{X}\right)^{-1} \left(\Sigma_{i=1}^N \tilde{X}_i' \hat{u}_i \hat{u}_i' \tilde{X}_i\right) \left(\Sigma_{i=1}^N \tilde{X}_i' \tilde{X}\right)^{-1}}}$$

if u_{it}, x_{it} iid $\Rightarrow$ simplify $V(\hat{b}) = \hat{\sigma}_u^2 \left(\Sigma_{i=1}^N \tilde{X}_i' \tilde{X}\right)^{-1}$

$\boxed{\text{Caution}}$

1. when heteroskedasticity or autocorrel t-ratio much larger than it's true value
2. T large enough t-ratio ~~N deg~~ fine $\equiv$ Normal dist; if T small $\Rightarrow$ t-ratio $\equiv$ t-dist with $N-1$ deg. Fredom under homoskedasticity

$\boxed{\text{Trend Regression}}$ $y_t = t + \varepsilon_t \quad \varepsilon_t \sim$ iid

For $y_t = b_2 \sqrt{t} + e_t \rightarrow$ write ~~et~~ $e_t = f(t, \varepsilon_t, \sqrt{t})$

$$\hat{\sigma}^2 = \frac{\Sigma_{i=1}^n (x_i - \mu)^2}{n} = \frac{n-1}{n}\sigma^2 \Rightarrow \begin{cases} \text{① Bias exists} & n-1 \approx n \\ \text{② But Consistent} & n \to \infty \end{cases}$$

Bias Correction: $\frac{n}{n-1}\hat{\sigma}^2 = \sigma^2 \equiv s^2$ Standard dev

$$\begin{cases} y = X_1\beta_1 + e & \rightarrow R^2 \\ y = X_1\beta_1 + X_2\beta_2 + u & \rightarrow \bar{R}^2 \end{cases} \quad R^2 < \bar{R}^2$$

$e_i \sim$ iid $N(0, \sigma^2)$ $\quad s^2 = \frac{\hat{u}'\hat{u}}{n-1}$ $\quad \hat{\beta}_1$ & $\hat{s}_2 \sim$ independent

$$\frac{\hat{\beta}_1 - \beta}{\sqrt{s^2 (X_1' X_1)^{-1}}} \sim t_1$$

$$\begin{cases} y_i^* = \mu_y + y_i \\ x_i^* = \mu_x + x_i \\ y_i = \beta x_i + e_i \end{cases} \Rightarrow y_i^* = \underbrace{\mu_y - \beta\mu_x}_{\alpha} + \beta\underbrace{(\mu_x + x_i)}_{x_i^*} + u_i \Rightarrow y_i^* = \alpha + \beta x_i^* + u_i$$

$$\Rightarrow \tilde{y}_i^* = \beta \tilde{x}_i^* + u_i^*$$

$$\begin{aligned} E\,\tilde{u}_i^2 &= \frac{n-1}{n}\sigma_u^2 \\ E(\tilde{u}_i \tilde{u}_{i+1}) &= -\frac{\sigma_u^2}{n} = O(n^{-1}) \\ E(\Sigma \tilde{x}_i u_i)^2 &= \sigma_u^2 \Sigma \tilde{x}_i^2 \\ E(\hat{\beta} - \beta)^2 &= \frac{\sigma_u^2}{\Sigma \tilde{x}_i^2} \\ E\,\tilde{x}_i^2 &= \frac{n-1}{n}\sigma_x^2 \end{aligned} \Rightarrow \sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} N\left(0, \frac{\sigma_u^2}{\frac{\Sigma}{n}\tilde{x}_i^2}\right) \text{ or } \frac{\hat{\beta} - \beta}{\sqrt{\frac{\sigma_u^2}{\Sigma \tilde{x}_i^2}}} \xrightarrow{d} N(0,1)$$

$x_i \sim$ stochastic $\Rightarrow$ instead of Expectation $\rightarrow$ prob. limit (Core Consistency)

$$\Rightarrow \text{plim}_{n\to\infty}(\hat{\beta}_n - \beta) = \frac{\text{plim}_{n\to\infty} \frac{1}{n}\Sigma \tilde{x}_i \tilde{u}_i}{\text{plim}_{n\to\infty} \frac{\Sigma \tilde{x}_i^2}{n}}$$

Cross term

$$E(\Sigma \tilde{x}_i \tilde{u}_i)^2 = n\left(\frac{n-1}{n}\sigma_x^2\right)\left(\frac{n-1}{n}\sigma_u^2\right) - n(n-1)\left(\frac{\sigma_x^2}{n}\right)\left(\frac{\sigma_u^2}{n}\right)$$

$$= n\sigma_x^2\sigma_u^2\left[1 - \frac{3n+2}{n^2}\right] = n\sigma_x^2\sigma_u^2\left[1 - \frac{3}{n}\right] = n\sigma_x^2\sigma_u^2 + O(n^{-1})$$

$$\text{plim} \frac{\Sigma}{n}\tilde{x}_i^2 = \sigma_x^2 - \frac{\sigma_x^2}{n} = \sigma_x^2 + O(n^{-1})$$

$$\frac{\Sigma}{n}\tilde{x}_i \tilde{u}_i \xrightarrow{d} N\left(0, \frac{\sigma_x^2 \sigma_u^2}{n}\right) \Rightarrow \frac{\frac{1}{\sqrt{n}}\Sigma \tilde{x}_i \tilde{u}_i}{\frac{1}{n}\Sigma \tilde{x}_i^2} \xrightarrow{d} N\left(0, \frac{\sigma_u^2}{\sigma_x^2}\right)$$

$E(\hat{\beta} - \beta | X)$, $E(\hat{\beta} - \beta | X)^2$ $\equiv$ treat x_i as non-stochastic term

$$\sqrt{n}(\hat{\beta} - \beta | X) \xrightarrow{d} N(0, \sigma_u^2 Q_x^{-1}) \qquad Q_x = \lim_{n\to\infty}\left(\frac{X'X}{n}\right)^{-1}$$

$\boxed{\text{Dummy var Reg}}$

$y_i = \alpha + \beta S_i + u_i$ $\quad S_i = \begin{cases} 0 & i = \text{female} \\ 1 & i = \text{male} \end{cases}$ $\quad \# n_1 = \frac{n}{2}$, $\# n_2 = \frac{n}{2}$

$$\hat{\beta} = \beta + \frac{\Sigma \tilde{S}_i \tilde{u}_i}{\Sigma \tilde{S}_i^2} \qquad \Sigma \tilde{S}_i^2 = \Sigma S_i^2 - \frac{1}{n}(\Sigma S_i)^2 = \frac{n}{4}$$

$$\text{plim} \frac{\Sigma}{n} \tilde{S}_i \tilde{u}_i = 0 \qquad \frac{1}{\sqrt{n}}\Sigma \tilde{S}_i \tilde{u}_i = \frac{\Sigma_{i=1}^n u_i^*}{\sqrt{n}} \quad \frac{\sigma_u^2}{4n}$$

$$u_i^* = \begin{cases} \frac{\tilde{u}_i}{2} & i = \text{male} \\ -\frac{\tilde{u}_i}{2} & i = \text{female} \end{cases} \Rightarrow E\left(\frac{\Sigma^n}{\sqrt{n}} u_i^*\right)^2 = \frac{\sigma_u^2}{4} + O(n^{-1})$$

$$\Rightarrow \frac{\Sigma \tilde{S}_i \tilde{u}_i}{\sqrt{n}} \xrightarrow{d} N\left(0, \frac{\sigma_u^2}{4}\right) \Rightarrow \sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} N(0, 4\sigma_u^2)$$

$$\Rightarrow \sqrt{n}\frac{\hat{\beta} - \beta}{\sqrt{4\sigma_u^2}} \xrightarrow{d} N(0,1) \qquad \text{Asym Var}(\hat{\beta}) = \frac{4\sigma_u^2}{n}$$

$\boxed{\text{Two Dependent Dummies}}$

	unskilled	skilled
Female	n_{11}	n_{12}
male	n_{21}	n_{22}

$\begin{cases} n_{11} + n_{12} = n_{21} + n_{22} \\ n_{11} = 2n_{21} \end{cases}$

$y_i = \alpha + \beta S_i + \gamma U_i + \varepsilon_i$ $\quad U_i \begin{cases} 0 & i = \text{non skilled} \\ 1 & i = \text{skilled} \end{cases}$

$y_i = \alpha + \beta S_i + u_i \Rightarrow u_i = \gamma U_i + \varepsilon_i$

$\frac{\Sigma S_i U_i}{n} \neq 0$ $\quad \begin{cases} n_{11} = 1/3 = n_{22} \\ n_{12} = \frac{1}{6} = n_{21} \end{cases}$ $\quad \boxed{\text{hierarchical gender difference}}$

Female unskilled: $S_i = U_i = 0$ $\quad E(y_i) = \alpha$

Female skilled: $S_i = 0, U_i = 1$ $\quad E(y_i) = \alpha + \gamma$

~~Female~~ male & unskilled: $S_i = 1, U_i = 0$ $\quad E(y_i) = \alpha + \beta$

Male & skilled: $S_i = 1, U_i = 1$ $\quad E(y_i) = \alpha + \beta + \gamma$

$\boxed{\text{Omitted Variable}}$ $E(S_i u_i) \neq 0 \Rightarrow$ OLS estimator $\hat{\beta}$ inconsistent

$\boxed{\text{Cross Dummies}}$ $y_i = \alpha + \beta S_i + \gamma U_i + \delta S_i U_i + e_i$

Econometrics summary note (1)

$w=(x,z)' \Rightarrow w \sim N(0,\Sigma)$, $\Sigma = \begin{bmatrix} \sigma_x^2 & \sigma_{xz} \\ \sigma_{xz} & \sigma_z^2 \end{bmatrix}$ $\equiv$ not orthogonal = Correlation, Synergy

$w \sim N(\mu,\Sigma) \Rightarrow w'\Sigma^{-1}w$ non Central χ^2_J param $\mu'\Sigma^{-1}\mu/2$

$\hookrightarrow (x_1,\dots,x_J)'$

Normal MGF $= e^{\mu t + \frac{\sigma^2 t^2}{2}}$

Normal CF $= e^{i\mu t - \frac{\sigma^2 t^2}{2}}$

Gamma func $\Gamma(p) = \int_0^\infty t^{p-1} e^{-t} dt = (p-1)!$

Log normal $= \frac{1}{\sqrt{2\pi}\,\sigma x} e^{-1/2\left(\frac{\ln x - \mu}{\sigma}\right)^2}$

$E(x) = e^{\mu + \frac{\sigma^2}{2}} \Rightarrow \mu = \ln E(x) - \frac{1}{2}\ln\left(1+\frac{V(x)}{E(x)^2}\right)$

$V(x) = e^{2\mu+\sigma^2}(e^{\sigma^2}-1) \Rightarrow \sigma^2 = \ln\left(1+\frac{V(x)}{E(x)^2}\right)$

$\text{mode}(x) = e^{\mu-\sigma^2}$ Median $(x) = e^{\mu}$

$x \sim N(\mu,\sigma^2) \Rightarrow e^x \sim LN(\mu,\sigma^2)$

$x \sim LN(\mu,\sigma^2) \Rightarrow \ln(x) \sim N(\mu,\sigma^2)$

$x \sim LN(\mu,\sigma^2) \Rightarrow y = x + c$ (Shift) $\sim LN$ $E(y) = E(x)+c$, $V(y) = V(x+c) = V(x)$

$x \sim LN(\mu,\sigma^2) \Rightarrow ax = y \sim LN(\ln a + \mu, \sigma^2)$

$x \sim LN(\mu,\sigma^2) \Rightarrow y = 1/x \sim LN(-\mu,\sigma^2)$

$x \sim LN(\mu_1,\sigma_1^2)$, $y \sim LN(\mu_2,\sigma_2^2)$, indep $\Rightarrow xy \sim LN(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$

Gamma Dist $f(x) = \frac{\lambda^p x^{p-1}}{\Gamma(p)} e^{-\lambda x}$ $\lambda, x, p > 0$

$p=1 \Rightarrow$ exponential $\sim \lambda e^{-\lambda x}$

$p = n/2$, $\lambda = 1/2 \Rightarrow \sim \chi^2$

$p \in N^+ \Rightarrow$ Erlang

Beta $0<x<c$

$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \left(\frac{x}{c}\right)^{\alpha-1}\left(1-\frac{x}{c}\right)^{\beta-1}\frac{1}{c}$

1. $\alpha=\beta \Rightarrow$ symmetry
2. $\alpha=\beta=1 \sim \text{Unif}(0,1)$
3. $\alpha, \beta < 1 \Rightarrow$ U shape
4. $\alpha=1, \beta>2 \Rightarrow$ strictly Convex
5. $\alpha=1, \beta=2$: Straight line
6. $\alpha=1$, $1<\beta<2$ strictly concave

- mean $\frac{c\alpha}{\alpha+\beta}$
- Var $\frac{c^2\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$

Logistic Dist $f(x) = \frac{e^{-\frac{(x-\mu)}{s}}}{s\left(1+e^{-\frac{(x-\mu)}{s}}\right)^2}$ $s>0$

$F(x) = \frac{1}{1+e^{-(x-\mu)/s}}$

Weibull Distribution

$f(x) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k}$

$k=1 \Rightarrow$ weibull $\equiv$ exponential

Cauchy Distr $f(x) = \frac{1}{\pi}\left[\frac{\gamma}{(x-x_0)^2+\gamma^2}\right]$ No mean and variance

x_0 : location

γ : scale

$\Rightarrow$ t-dist with $v=1 \equiv f(x) = \frac{1}{\pi(1+x^2)}$ Standard Cauchy dist

---

Survival Fun $S(x) = 1-F(x) = P[X \geq x]$

Hazard Func $h(x) = \frac{f(x)}{S(x)} = \frac{f(x)}{1-F(x)}$ exponent. Dist: $h(t)=\lambda$

moment Gen func (mdf) $M_X(t) = E(e^{tx})$ $t \in R$ $\nexists$ Cauchy

Characteristic Function (CF) $\phi(t) = E(e^{itx})$ always Exists

Cauchy: $\phi(t) = e^{x_0 it - \gamma|t|}$

Cumulant Gent. Func. $g(t) = \log[E(e^{tx})]$

$k_1 = \mu = g'(0)$

$k_2 = \sigma^2 = g''(0)$

$k_n = g^{(n)}(0)$

Joint dist for x, y

$\text{Prob}(a \leq x \leq b, c \leq y \leq d) = \begin{cases} \sum_a^b \sum_c^d f(x,y) \\ \int_a^b \int_c^d f(x,y)\,dx\,dy \end{cases}$

$\begin{pmatrix} x \\ y \end{pmatrix} \sim N\left(\begin{bmatrix}\mu_x \\ \mu_y\end{bmatrix}, \begin{bmatrix}\sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2\end{bmatrix}\right)$

bivar normal $f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} e^{\frac{-1/2\,\varepsilon_x^2+\varepsilon_y^2-2\varepsilon_x\varepsilon_y\rho}{1-\rho^2}}$

$\varepsilon_x = \frac{x-\mu_x}{\sigma_x}$ $\varepsilon_y = \frac{y-\mu_y}{\sigma_y}$ $\rho = \frac{\sigma_{xy}}{\sigma_x\sigma_y}$

---

marginal (2)

$P(x=x_0) = \sum_y P(x=x_0|y=y_0)\cdot P(y=y_0) = f_x(x) = \begin{cases} \sum_y f(x,y) \\ \int_y f(x,y)\,dy \end{cases}$

$f(x,y) = f_x(x)\cdot f_y(y)$ $F(x,y) = F_x(x)\cdot F_y(y)$

$f_x(x) = N(\mu_x,\sigma_x^2)$ $f_y(y) = N(\mu_y,\sigma_y^2)$

$E(x) = \begin{cases} \sum_x f_x(x) = \sum_x x \sum_y f(x,y) \\ \int_x x f_x(x)\,dx = \int_x \int_y x f(x,y)\,dy\,dx \end{cases}$

$\text{Cov}(x,y) = E[(x-\mu_x)(y-\mu_y)]$ $\text{Corr}(x,y) = \frac{\sigma_{xy}}{\sigma_x\sigma_y}$

$V(x+y) = V(x)+V(y)+2\text{Cov}(x,y)$

$f(y|x) = \frac{f(x,y)}{f_x(x)}$ $f(y|x) = N(\alpha+\beta x, \sigma_y^2(1-\rho^2))$

$\alpha = \mu_y - \beta\mu_x$ $\beta = \frac{\sigma_{xy}}{\sigma_x^2}$

Regresion $E(y|x) = \begin{cases} \sum_y y f(y|x) \\ \int_y y f(y|x)\,dy \end{cases}$

Conditional mean

$y = E(y|x) + (y - E(y|x)) = E(y|x) + \varepsilon$

e.g. $y = a+bx+\varepsilon \Rightarrow E(y|x) = a+bx$ deterministic part

$V(y|x) = E((y-E(y|x))^2|x) = E(y^2|x) - E^2(y|x)$

Multivar normal $x = (x_1,\dots,x_n)'$

$f(x) = (2\pi)^{-n/2}|\Sigma|^{1/2} e^{1/2(x-\mu)'\Sigma^{-1}(x-\mu)}$

1. $x \sim N(\mu,\Sigma) \Rightarrow Ax+b \sim N(A\mu+b, A\Sigma A')$
2. $x \sim N(\mu,\Sigma) \Rightarrow (x-\mu)'\Sigma^{-1}(x-\mu) \sim \chi^2_n$
3. $x \sim N(\mu,\Sigma) \Rightarrow \Sigma^{-1/2}(x-\mu) \sim N(0, I_n)$

---

mean $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$

Standard Error (deviation): $S_x = \left[\frac{\sum_{i=1}^n (x_i-\bar{x})^2}{n-1}\right]^{1/2}$

Covariance: $S_{xy} = \frac{\sum_{i=1}^n (x_i-\bar{x})(y_i-\bar{y})}{n-1}$

Correlation: $r_{xy} = \frac{S_{xy}}{S_x S_y}$

$y_i = bx_i + u_i$

$u_i \sim \text{iid } N(\mu,\sigma^2)$

$\hat{b}$: estimate computed for pop. mean μ

σ : Standard dev vs. S_x : Standard Error

u_i : error $\hat{u}_i$: residuals

Consistent estimator : -- Consistently estimates...

Asymptotic = approximated (how works as $n \to \infty$)

---

1. unbiased $E(\hat{\theta}-\theta) = 0$ $\forall n$ $\begin{cases} \text{- param } \theta \\ \text{- mean of sample } \hat{\theta} \text{ dist} \end{cases}$
2. Efficient : $V(\hat{\theta}_1) < V(\hat{\theta}_2)$ $\rightarrow$ effect on t-value $\Rightarrow$ / acceptance of true null
3. (MSE) mean square Error: $MSE(\hat{\theta}) = E(\hat{\theta}-\theta)^2 = E[(\hat{\theta}-E\hat{\theta}+E\hat{\theta}-\theta)^2] = V(\hat{\theta}) + E(\hat{\theta}-\theta)^2$
4. Likelihood : $u_i = y_i - bx_i$
   $f(u_1,\dots,u_n|b) = f(u_1|b)f(u_2|b)\dots f(u_n|b) = \prod_{i=1}^n f(u_i|b)$
   $= L(b|x_1,\dots,x_n)$
5. Cramer Rao Lower Bound : (1) Regularity Condition (2) Variance of unbiased Estimator of θ at least as large as $[I(\theta)]^{-1} = \left(-E\left[\frac{\partial^2 \ln L(\theta)}{\partial\theta^2}\right]\right)^{-1} = \left(E\left(\frac{\partial \ln L(\theta)}{\partial\theta}\right)^2\right)^{-1}$
   $\hookrightarrow$ information number for sample

Trick - seperate the calculation of ΣT and ΣN to decrease Complexity

Key Sometimes take a time to get big picture and only then write

Econometrics summary note

- Large sample dist. theory

① Convergence in prob. : n: sample size

$\lim_{n\to\infty} \text{prob}(|x_n - c| < \varepsilon) = 0 \quad \forall \varepsilon > 0 \Rightarrow x_n \xrightarrow{p} c, \ \text{plim}_{n\to\infty} x_n = c$

② almost sure convergence : $p(\lim_{n\to\infty} x_n = c) = 1$

$x_n \xrightarrow{a.s.} c$

③ Convergence in r-th mean : $\lim_{n\to\infty} E(|x_n - c|^r) = 0$

$x_n \xrightarrow{L^r} c$

④ Consistent Estimator : iff $\text{plim}_{n\to\infty} \hat{\theta}_n = \theta$

⑤ Kinchine's weak law of large number : $E(x_i) = \mu$ → random sample

$\text{plim}_{n\to\infty} \bar{x}_n = \text{plim}_{n\to\infty} \frac{\sum_{i=1}^n x_i}{n} = \mu$

⑥ Chebychev's weak law of large number

(1) $E(x_i) = \mu_i < \infty$ (2) $V(x_i) = \sigma_i^2 < \infty$ (3) $\frac{\bar{\sigma}_n^2}{n} = \frac{\sum_{i=1}^n \sigma_i^2}{n^2} \xrightarrow[n\to\infty]{} 0$

$\Rightarrow \text{plim}_{n\to\infty} (\bar{x}_n - \bar{\mu}_n) = 0 \quad \bar{\mu}_n = \frac{\sum_{i=1}^n \mu_i}{n}$

⑦ Kolmogorov Strong LLN : $x_i \sim$ iid sequence

(1) $E(x_i) = \mu_i < \infty$ (2) $V(x_i) = \sigma_i^2 < \infty$ (3) $\sum_{s=1}^{\infty} \frac{\sigma_s^2}{s^2} < \infty$

$\Rightarrow \bar{x}_n - \bar{\mu}_n \xrightarrow[n\to\infty]{a.s} 0$

⑦ Corollary : (1) $E(x_i) = \mu < \infty$ (2) $E|x_i| < \infty \Rightarrow \bar{x}_n - \mu \xrightarrow{a.s} 0$

⑧ Markov Strong LLN : $x_i \sim$ iid seq rand var

(1) $E(x_i) = \mu_i < \infty$ (2) $\exists \delta > 0, \ \sum_{i=1}^{\infty} \frac{E[|x_i - \mu_i|^{1+\delta}]}{i^{1+\delta}} < \infty$

$\Rightarrow \bar{x}_n - \bar{\mu}_n \xrightarrow{a.s} 0$

Probability limit

① $\begin{cases} \text{plim}\ x_n = b \\ \text{plim}\ y_n = c \end{cases} \Rightarrow \begin{array}{l} \text{plim}(x_n + y_n) = b + c \\ \text{plim}\ x_n y_n = bc \\ \text{plim}\ x_n / y_n = b/c \ \text{if } c \neq 0 \end{array}$

② W_n : matrix rand var $\Rightarrow \text{plim}\ W_n = \Omega \Rightarrow \text{plim}\ W_n^{-1} = \Omega^{-1}$

(3) X_n, Y_n random matrices $\begin{array}{l} \text{plim}\ X_n = B \\ \text{plim}\ Y_n = C \end{array} \Rightarrow \text{plim}\ X_n Y_n = BC$

Convergence in Dist

① $\lim_{n\to\infty} |F(x_n) - F(x)| = 0 \quad \forall$ cont. point of $F(x) \Rightarrow$ Convergence in Distribution

$\Rightarrow F(x)$: Limiting Dist. of $(x_n) \Rightarrow x_n \xrightarrow{d} x$

② Cramer-Wold Device : $\begin{cases} x_n \xrightarrow{d} x \\ c \in \mathbb{R} \end{cases} \Rightarrow c' x_n \xrightarrow{d} c' x$

③ Lindberg Levy CLT (Central limit theorem) : $x_1, \dots, x_n$ random sample from prob. dist. (1) finite mean $\mu < \infty$

(2) finite var $\sigma^2 < \infty \Rightarrow$ sample mean $\bar{x}_n = \frac{\sum_{i=1}^n x_i}{n}$:

$\Rightarrow \sqrt{n}(\bar{x}_n - \mu) \xrightarrow{d} N(0, \sigma^2)$

④ Lindberg Feller CLT : $x_1, \dots, x_n$ rand. sample from prob dist. (1) finite mean $\mu_i < \infty$ (2) finite var $\sigma_i^2 < \infty$

(3) $\bar{\sigma}_n = \frac{\sum_{i=1}^n \sigma_i^2}{n}$ (4) $\bar{\mu}_n = \frac{\sum_{i=1}^n \mu_i}{n}$ (5) $\lim_{n\to\infty} \frac{\max(\sigma_i)}{n\bar{\sigma}} = 0$

(6) $\bar{\sigma}_n^2 : \lim_{n\to\infty} \bar{\sigma}_n^2 = \bar{\sigma}^2 < \infty \Rightarrow$ sample mean $\bar{x}_n = \sum_{i=1}^n x_i$:

$\Rightarrow \sqrt{n}(\bar{x}_n - \bar{\mu}_n) \xrightarrow{d} N(0, \bar{\sigma}^2)$, or $\frac{\sqrt{n}(\bar{x}_n - \bar{\mu}_n)}{\bar{\sigma}} \xrightarrow{d} N(0,1)$

⑤ Liapnov CLT : $\{x_i\}$ seq of iid r.v (1) finite mean $\mu_i < \infty$

(2) finite positive var $\sigma_i^2 < \infty$ (3) $E(|x_i - \mu_i|^{2+\delta}) < \infty \ \exists \delta > 0$

(4) $\bar{\sigma}_n < \infty \ \forall n$ (5) $\bar{\sigma}_n > 0 \ \forall n \Rightarrow \frac{\sqrt{n}(\bar{x}_n - \bar{\mu}_n)}{\bar{\sigma}_n} \xrightarrow{d} N(0,1)$

⑥ multivar Lindberg ~~Feller~~ CLT : $\sqrt{n}(\bar{x}_n - \bar{\mu}_n) \xrightarrow{d} N(0, Q)$

(1) $V(x_i) = Q_i$ (2) $\lim \bar{Q}_n = Q$

⑦ Asymptotic Cov Matrix : (1) $\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N(0, V)$

$\Rightarrow$ Asympt. Cov matrix $V(\hat{\theta}_n) = \frac{1}{n} V$

Order of a sequence

① Sequence c_n of order n^δ : $O(n^\delta)$ iff $\text{plim}_{n\to\infty} \frac{c_n}{n^\delta} = c < \infty$

② seq c_n of order less than n^δ iff $\text{plim}_{n\to\infty} \frac{c_n}{n^\delta} = 0 \quad o(n^\delta)$

③ seq var $x_n = O_p(g(n))$ iff $\exists N_\varepsilon : \varepsilon > 0, \ \forall n > N_\varepsilon$

$\text{prob}\left(\left|\frac{x_n}{g(n)}\right| < c\right) > 1 - \varepsilon$

Ⅰ $x_n \sim N(0, \sigma^2) \Rightarrow x_n = O_p(1)$

Ⅱ $O_p(n^a) O_p(n^b) = O_p(n^{a+b})$

Ⅲ if $\sqrt{n}(\bar{x}_n - \bar{\mu}_n) \xrightarrow{d} N(0, \bar{\sigma}^2) \Rightarrow (\bar{x}_n - \bar{\mu}_n) = O_p(n^{-1/2})$

$\bar{x}_n = O_p(1) \rightarrow$ due to mean

④ $x_n = o_p(g_n)$: $\frac{x_n}{g_n} \xrightarrow{p} 0$

(a) if $\sqrt{n}\bar{x}_n \xrightarrow{d} N(0, \bar{\sigma}^2) \Rightarrow \bar{x}_n = O_p(n^{-1/2}), \ \bar{x}_n = o_p(1)$

(b) $o_p(n^a) o_p(n^b) = o_p(n^{a+b})$

① Linear : $y = Xb + u < \infty$

② ① $X \neq$ stochastic ② Finite ③ $n \times k$

③ $X'X$ nonsingular $\forall n \geq k$

④ $E(u) = 0$

⑤ $u \sim N(0, \sigma^2 I)$

If X = stochastic

④ Exogeneity of indep var : $E(u|X) = 0$

⑤ ① Homoscedasticity & no autoc

② Normal dist

① unbiasedness $E(\hat{\beta}) = \beta$ Given (1-4)

① $\exists \hat{\beta} \ \forall n \geq k$, unique (Given 1-3)

② Normality (Given 1-5) $\hat{\beta} \sim N(\beta, \sigma^2 (X'X)^{-1})$

③ Gauss Markov Theorem : OLS : mini Var linear unbiased Estimator

Dummy Variable test : $y_i = \alpha_0 + \alpha_1 d_i + \alpha_2 s_i + \alpha_3 d_i s_i + \varepsilon_i$

skilled : $\alpha_1 = -\alpha_3$ unskilled : $\alpha_1 = 0$

matrix form $x_i = \overset{\alpha+}{\beta} w_i + \varepsilon_i \quad w_i \begin{cases} 0 & \text{if } i \in F \\ 1 & \text{o.w} \end{cases}$

$\begin{cases} \alpha_1 = \alpha + \beta & \text{if } E(w_i) = 1 \\ \alpha_2 = \alpha & \text{if } E(w_i) = 0 \end{cases} \Rightarrow x_i = \alpha_2 + (\alpha_1 - \alpha_2) w_i + \varepsilon_i$

$X = QY + \varepsilon$

$Q = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ \vdots & \vdots \end{bmatrix} \quad Y = \begin{bmatrix} \alpha_1 \\ \alpha_1 - \alpha_2 \end{bmatrix} \quad \hat{Y} = \begin{bmatrix} \hat{\alpha}_1 \\ \hat{\alpha}_1 - \hat{\alpha}_2 \end{bmatrix} = (Q'Q)^{-1} Q' X$

- in selection k^* AIC : overestimates asymptotically
BIC & Hannan-Quinn Consistent : $\lim_{T\to\infty} \phi(T) = \infty$

- General to specific (GS) : ① accumulate prob of rejection
: overestimation probability > significance level ⇒
Pötscher 1983 : Significance level α_T - $\frac{\log q_T}{\sqrt{T}} \to 0$, if $q_T = 1 - \Phi(C_T)$ ← normal

$x_t = \rho^2 x_{t-1} + \varepsilon_t, \ \varepsilon_t = \rho u_{t-1/2} + u_t$, ARMA(1,1)

$z_t = \rho z_{t-1} + \varepsilon_t, \ \varepsilon_t = (\phi - \rho) \sum_{j=0}^{\infty} \phi^j e_{t-j-1} + u_t + e_t$ ARMA(1,∞)

$E(\bar{y}_T - \mu)^2 = \frac{1}{T^2}[T\gamma_0 + 2(T-1)\gamma_1 + \cdots + 2\gamma_{T-1}] \leq \frac{1}{T}[\gamma_0 + 2\gamma_1 + \cdots + \gamma_{T-1}]$

$\lim_{T\to\infty} T \cdot E(\bar{y}_T - \mu)^2 = \sum_{-\infty}^{\infty} \gamma_j$

$y_t = u_t \quad u_t = \rho u_{t-1} + \varepsilon_t \quad \varepsilon_t \sim \text{iid}(0, \sigma^2)$

$E\left(\frac{\sum u_t}{T}\right)^2 = \frac{1}{T} \frac{\sigma^2}{1-\rho^2}\left[1 + 2\frac{\rho}{1-\rho}\right] + O(T^{-2}) = \frac{1}{T} \frac{\sigma^2}{(1-\rho)^2} + O(T^{-2})$

Econometrics summary note

CLT for MDS (martingale Diff. Seq)

(1) $\{y_t\}_{t=1}^{\infty}$: mds (2) $\bar{y}_T = \frac{\sum_{t=1}^{T} y_t}{T}$ (3) $E(y_t^2)=\sigma_t^2>0$

(4) $\frac{\sum_{t=1}^{T}\sigma_t^2}{T} \to \sigma^2>0$ (5) $E|y_t|^r<\infty$ $\Rightarrow \sqrt{T}\,\bar{y}_T \xrightarrow{d} N(0,\sigma^2)$

CLT for stationary stochastic process

(1) $y_t=\mu+\sum_{j=0}^{\infty}\psi_j\varepsilon_{t-j}$ (2) $\varepsilon_t\sim iid$ (3) $E(\varepsilon_t^2)<\infty$

(4) $\sum_{j=0}^{\infty}|\psi_j|<\infty \Rightarrow \sqrt{T}(\bar{y}_T-\mu)\xrightarrow{d} N(0,\sum_{j=-\infty}^{\infty}\gamma_j)$

e.g ① $y_t=a+\sum_{j=0}^{\infty}\rho^j e_{t-j} \Rightarrow \sqrt{T}(\bar{y}_T-a)\xrightarrow{d} N(0,\frac{\sigma_e^2}{(1-\rho)^2})$

when $y_t=a+u_t$ $u_t=\rho u_{t-1}+e_t$ $e_t\sim iid(0,\sigma_e^2)$

② $y_t=a(1-\rho)+\rho y_{t-1}+\varepsilon_t$ $\varepsilon_t\sim iid(0,\sigma^2)$ Denom $=\frac{\sigma^2}{1-\rho^2}$

$\Rightarrow \sqrt{T}(\hat\rho-\rho)\xrightarrow{d} N(0,1-\rho^2)$

Bias using simple Taylor Expansion

$E(\frac{A}{B})=\frac{a}{b}-\frac{1}{b^2}Cov(A,B)+\frac{a}{b^3}Var(B)+O(T^{-2})$

$E(A)=a$ $E(B)=b$ $E(A-a)=0$ → set theory

$E(\frac{A}{B})=\frac{E(A)}{E(B)}\left(1-\frac{Cov(A,B)}{E(A)E(B)}+\frac{Var(B)}{E(B)^2}\right)+O(T^{-2})$

model: $y_t=a+\rho y_{t-1}+e_t$ $e_t\sim iid\ N(0,1)$ (*)

Key in time series: use $Y_1,\dots,Y_n$ Covariance & not normal ols

Trick: module (I) $E(x_t x_{t+k} x_{t+k+l} x_{t+k+l+m})=\frac{\rho^{k+m}(1+2\rho^{2l})}{(1-\rho^2)^2}$

$\frac{1}{T^2}E(\sum x_t^2)^2=\frac{1}{T^2}\left[\frac{3T}{(1-\rho^2)^2}+2\sum_{t=1}^{T-1}(T-i)\frac{1+2\rho^{2t}}{(1-\rho^2)^2}\right]$

use module (I): $E(y_{t-1}y_t y_{t+l}^2)=\frac{\rho(1+2\rho^{2l})}{(1-\rho^2)^2}$

$\Rightarrow Cov(\frac{\sum y_t y_{t-1}}{T},\frac{\sum y_t^2}{T})=\frac{2\rho}{(1-\rho^2)^2}$

$\Rightarrow$ for model (*): $E(\hat\rho-\rho)=-\frac{1+3\rho}{T}+O(T^{-2})$

non Constant Case: $x_t=\rho x_{t-1}+e_t \Rightarrow E(\hat\rho-\rho)=-\frac{2\rho}{T}+O(T^{-2})$

For trend Case: $y_t=a+bt+\rho y_{t-1}+e_t \Rightarrow E(\hat\rho-\rho)=\frac{-2(1+2\rho)}{T}+O(T^{-2})$

main key for any problem: Break down problem and start from the leaves to solve it

Edgeworth series

- approx prob dist. in terms of cumulants (Hermit polynomial)

$e^{tX}=1+tX+\frac{t^2X^2}{2!}+\frac{t^3X^3}{3!}+\dots+\frac{t^nX^n}{n!}=\lim_{n\to\infty}(1+\frac{tx}{n})^n$ → pop estimate

$M_X(t)=E(e^{tX})=1+tm_1+\frac{t^2m_2}{2!}+\frac{t^3m_3}{3!}+\dots$ (MGF)

Taylor: $f(x)=f(a)+\frac{f'(a)}{1!}(x-a)+\frac{f''(a)}{2!}(x-a)^2+\dots=\sum_{n=0}^{\infty}\frac{f^{(n)}(a)}{n!}(x-a)^n$

↳ point estimate

non constant case Philips 1977: $\hat\rho-\rho=\frac{\sum y_{t-1}u_t}{\sum\tilde{y}_{t-1}^2}=\frac{\sum y_{t-1}y_t-\rho\sum y_{t-1}^2}{\sum y_{t-1}^2}$

$\Rightarrow$ Express in moments vector (m): $\sqrt{T}(\hat\rho-\rho)=\sqrt{T}e(m)$

Taylor Expansion: $\sqrt{T}e(m)=\sqrt{T}\{e_r m_r+\frac{1}{2}e_{rs}m_r m_s+\frac{1}{6}e_{rst}m_r m_s m_t+O_p(\frac{1}{T^2})\}$: $e_r=\frac{\partial e(0)}{\partial m_r}$

$Pr\left(\frac{\sqrt{T}(\hat\rho-\rho)}{\sqrt{1-\rho^2}}\le\omega\right)=\Phi(\omega)+\frac{\phi(\omega)}{\sqrt{T}}\left(\frac{\rho}{\sqrt{1-\rho^2}}\right)(\omega^2+1)$

$\omega=\frac{x}{\sqrt{1-\rho^2}}$

Constant Case (Tanaka, 1983): $Pr\left(\frac{\sqrt{T}(\hat\rho-\rho)}{\sqrt{1-\rho^2}}\le\omega\right)=\Phi(\omega)+\frac{\phi(\omega)}{\sqrt{T}}\left(\frac{\rho+\rho\omega^2}{\sqrt{1-\rho^2}}\right)$

Unit root process

Integrated of order one $I(1)$ if $m=1$ root of the charac. polynomial $m^p-m^{p-1}a_1-m^{p-2}a_2-\dots-a_p=0$ for model $y_t=a_1y_{t-1}+a_2y_{t-2}+\dots+a_py_{t-p}+\varepsilon_t$

Charact. polynomial: $P_A=\det(tI-A)$ encodes eigenvalue and determinant of matrix (with its trace)

problem of unit root: not stationary. e.g $y_t=a_1y_{t-1}+\varepsilon_t$

$a_1=1$, $m-a_1=0\Rightarrow m=1$ ↳ moment of the stoch. process depend on t: $y_t=y_0+\sum_{j=1}^{t}\varepsilon_j$ $Var(y_t)=\sum_{j=1}^{t}\sigma^2=t\sigma^2$

Unit root

$y_t=a+u_t$ $u_t=u_{t-1}+\varepsilon_t$

$y_T=y_{t-1}+\varepsilon_t$ $y_t=\varepsilon_1+\dots+\varepsilon_t$ $\varepsilon_t\sim iid\ N(0,1)$

$y_t\sim N(0,t)$ $y_t-y_{t-1}=\varepsilon_t\sim N(0,1)\to$ difference solves

$y_t-y_s\sim N(0,t-s)$

Standard Brownian Motion: (1) $W(\cdot)$ Cont-time stoch process (2) date $t\in[0,1]$ (3) scalar $W(t)$: ① $W(0)=0$ ② $[W(s)-W(t)]\sim N(0,s-t)$ ③ $W(1)\sim N(0,1)$

Transition from Discrete to Continuous

→ Solve problem of not invariant variance

$\frac{1}{\sqrt{T}}\frac{\sum_{t=1}^{T}y_t}{T}=\frac{\sum_{t=1}^{T}\frac{y_t}{T}}{\sqrt{T}}=\frac{1}{\sqrt{T}}\left(\frac{\varepsilon_1}{T}+\frac{\varepsilon_1+\varepsilon_2}{T}+\dots+\frac{1}{T}\sum_{t=1}^{T}\varepsilon_T\right)\xrightarrow{d}\int_0^1 W(r)dr$

$\Rightarrow$ ① $\frac{\sum_{t=1}^{T}y_t^2}{T^2}\xrightarrow{d}\int_0^1 W^2dr$ ② $\frac{\sum_{t=1}^{T}ty_{t-1}}{T^{5/2}}\xrightarrow{d}\int_0^1 rWdr$

if $\varepsilon_t\sim N(0,\sigma^2)\Rightarrow T^{-3/2}\sum y_t\xrightarrow{d}\sigma^2\int_0^1 Wdr$

Take away: divide by $\frac{1}{\sqrt{T}}$ make continuous and integrate

Example: $y_t=y_{t-1}+e_t$ $e_t\sim iid(0,\sigma^2)$ $y_0=O_p(1)$

① $T\to\infty\Rightarrow\frac{\sum e_t}{\sqrt{T}}\xrightarrow{d}\sigma W(1)$ ☒ $=N(0,\sigma^2)$

② $\frac{\sum y_{t-1}e_t}{\sqrt{T}}\to\infty$, $\frac{1}{\sqrt{t}}y_t=\frac{1}{\sqrt{t}}\sum_{s=1}^{t}e_s\xrightarrow{d}N(0,\sigma^2)$

③ $\frac{1}{\sigma^2t}y_t^2=\frac{1}{\sigma^2t}(\sum_{s=1}^{t}e_s)^2\xrightarrow{d}N(0,1)^2=X_1^2$ (chisq) large t @

④ $\frac{\sum y_{t-1}e_t}{T}\xrightarrow{d}\frac{\sigma^2}{2}[W(1)^2-1]$ ▲

Proof ▲: $y_{t-1}e_t=y_te_t-e_t^2=y_t^2-y_ty_{t-1}-e_t^2$

$y_ty_{t-1}=y_{t-1}^2+y_{t-1}e_t \Rightarrow y_{t-1}e_t=\frac{1}{2}(y_t^2-y_{t-1}^2-e_t^2)$

$\Rightarrow\frac{\sum y_{t-1}e_t}{T}=\frac{1}{2}(y_T^2-y_0^2)-\frac{\sum e_t^2}{2T}\xrightarrow{d}\frac{1}{2}\sigma^2X^2-\frac{\sigma^2}{2}=$ @

$\frac{1}{2}\sigma^2[W(1)^2-1]$

⑤ $\frac{\sum^T y_{t-1}}{T^{3/2}}=\frac{\sum^T e_t}{\sqrt{T}}-\frac{\sum^T te_t}{T^{3/2}}\xrightarrow{d}\sigma W(1)-\sigma\int_0^1 W(r)dr$

⑥ $y_t=\rho y_{t-1}+e_t\Rightarrow\hat\rho=\frac{\sum y_{t-1}e_t}{\sum y_{t-1}^2}$ $\xrightarrow{d}\frac{1}{2}\sigma^2[W(1)-1]$ / $\xrightarrow{d}\sigma^2\int_0^1 W^2dr$

$\Rightarrow T(\hat\rho-1)\xrightarrow{d}(\int W^2dr)^{-1}(\frac{1}{2}[W(1)^2-1])=(\int W^2dr)^{-1}\int WdW$

t-ratio: $t_\rho=\frac{\hat\rho-1}{\sqrt{\hat\sigma_T}}=\frac{\hat\rho-1}{\sqrt{(\sigma_T^2\sum y_{t-1}^2)}}$ $\sigma_T^2=\frac{\sum(y-\hat\rho y_{t-1})^2}{T}\xrightarrow{p}\sigma^2$

$t_\rho=\frac{\hat\rho-1}{\sqrt{(\sigma_T^2\sum y_{t-1}^2)}}\xrightarrow{d}(\int W^2dr)^{-1/2}(\frac{1}{2}[W(1)^2-1])$ $\ne\{^{1.96}_{1.65}$

upper and lower 2.5 (5.0)%

(-2.23, 1.62) (-1.95, 1.28)

Econometrics summary note

$y_t = a + \rho y_{t-1} + e_t$ Unit root Process I : Constant

$\rho = 1 \quad a = 0$

$$\begin{bmatrix} \hat{a}-0 \\ \hat{\rho}-1 \end{bmatrix} = \begin{bmatrix} T & \Sigma y_{t-1} \\ \Sigma y_{t-1} & \Sigma y_{t-1}^2 \end{bmatrix}^{-1} \begin{bmatrix} \Sigma e_t \\ \Sigma y_{t-1} e_t \end{bmatrix}$$

$$\begin{bmatrix} \sqrt{T} & 0 \\ 0 & T \end{bmatrix} \begin{bmatrix} \hat{a}-0 \\ \hat{\rho}-1 \end{bmatrix} = \begin{bmatrix} 1 & \frac{\Sigma y_{t-1}}{T^{3/2}} \\ \frac{\Sigma y_{t-1}}{T^{3/2}} & \frac{\Sigma y_{t-1}^2}{T^2} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\Sigma e_t}{\sqrt{T}} \\ \frac{\Sigma y_{t-1} e_t}{T} \end{bmatrix} \xrightarrow{d}$$

$$\begin{bmatrix} 1 & \sigma \int w(r)dr \\ \sigma \int w(r)dr & \sigma^2 \int w(r)^2 dr \end{bmatrix}^{-1} \begin{bmatrix} \sigma w(1) \\ \frac{\sigma^2}{2}[w(1)^2 - 1] \end{bmatrix}$$

$$T(\hat{\rho}-1) \xrightarrow{d} \frac{\frac{1}{2}[w(1)^2-1] - w(1)\int w(r)}{\{\int w^2 dr - (\int w dr)^2\}} = \frac{\int \tilde{w} dw}{\int \tilde{w}^2 dw}$$

t-ratio statistics: $t_\rho \xrightarrow{d} \dfrac{\frac{1}{2}[w(1)^2-1] - w(1)\int w dr}{\{\int w^2 dr - (\int w dr)^2\}^{1/2}}$

Unit Root Test $\quad y_t = a + \rho y_{t-1} + \Sigma \phi_j \Delta y_{t-j} + e_t$

$\dfrac{\Sigma}{T} \Delta y_{t-j} = O_p\left(\frac{1}{\sqrt{T}}\right) \xrightarrow[T\to\infty]{} 0$

Rejection of null $\neq$ stationary @

Economic mean of nonstationary

① No steady state. ≡ ① No static mean or average exists

② never converge to mean (random around mean)

② No Equilibrium, since no steady state ① Cannot forecast or predict future value without cons other nonstationary variable

③ Fast Convergence rate

@ example: $x_t = \rho x_{t-1} + \varepsilon_t \quad \varepsilon_t = \sqrt{t}\, e_t \quad e_t \sim iid(0, \sigma^2)$

at $\rho = 0$ non stationary but not unit root (from $\rho = 0$)

$\hat{\rho} - \rho = \dfrac{\Sigma x_{t-1}\varepsilon_t}{\Sigma x_{t-1}^2} \qquad E(\Sigma x_{t-1}\varepsilon_t)^2 = E(\Sigma \varepsilon_{t-1}\varepsilon_t)^2 = E(\Sigma t e_{t-1}^2 e_t^2)$

$= (\sigma^2)^2 \frac{T^2}{2} + O(T), \quad \dfrac{\Sigma x_{t-1}\varepsilon_t}{T} \xrightarrow{d} N(0, \frac{\sigma^4}{2})$

$\Sigma x_{t-1}^2 = \Sigma \varepsilon_{t-1}^2 = \Sigma t e_{t-1}^2 \xrightarrow{p} \frac{\sigma^2 T^2}{2} + O(T)$

$\Rightarrow T(\hat{\rho}-\rho) = \dfrac{\frac{\Sigma}{T} x_{t-1}\varepsilon_t}{\frac{\Sigma x_{t-1}^2}{T^2}} \xrightarrow{d} N(0,2) = \sqrt{2}\, W(1)$

Convergence rate

limiting Dist. not function of Brownian motion

Unit root testing Considering Finite Sample Bias — Better than ADF

recursive mean adjustment: $\bar{y}_t = \dfrac{\Sigma_{s=1}^{t-1} y_s}{t-1}$

$y_t - \bar{y}_t = a + \rho(y_t - \bar{y}_t) + (\rho - 1)\bar{y}_t + e_t$

$a = 0, \rho = 1 \Rightarrow y_t - \bar{y}_t = \rho(y_t - \bar{y}_t) + e_t$

$\Rightarrow T(\hat{\rho}_{RD} - 1) \xrightarrow{d} \left(\int w dr - r^{-1}\int_0^r w(r) dr\right)^{-1}\left(\int w dw - r^{-1}\int_0^r w dw\right)$

Weak stationary & local to Unity

$y_t = \rho_n y_{t-1} + u_t \quad t = 1, \dots, n \quad \sqrt{n}(\hat{\rho}_n - \rho_n) \xrightarrow{d} N(0, 1-\rho_n^2)$

$\rho_n = 1 - \dfrac{c}{n^\alpha} \quad 0 < \alpha < 1, \ c > 0 \quad \dfrac{\sqrt{n}}{\sqrt{1-\rho_n^2}}(\hat{\rho}_n - \rho_n) \xrightarrow{d} N(0,1)$

$1 - \rho_n^2 = \dfrac{c^2}{n^{2\alpha}} + \dfrac{2c}{n^\alpha} = O(n^{-\alpha}) + O(n^{-2\alpha})$

$\sqrt{1-\rho_n^2} = \sqrt{\frac{2c}{n^\alpha}\left(1 + \frac{c}{2n^\alpha}\right)} \Rightarrow \dfrac{\sqrt{n}}{\sqrt{1-\rho_n^2}} = n^{1/2} n^{\alpha/2} \dfrac{1}{\sqrt{2c}} + O(n^{\frac{\alpha+1}{2}})$

$\Rightarrow n^{\frac{\alpha+1}{2}}(\hat{\rho}_n - \rho_n) \xrightarrow{d} N(0, 2c) \qquad \alpha = 1 \Rightarrow$ local to unity

DGP: $y_t = \rho_n y_{t-1} + u_t \quad \rho_n = \exp\left(\frac{c}{T}\right) \approx 1 + \frac{c}{T}$

$J(r) = \int_0^r e^{(r-s)c} dw(s)$: Gaussian Process

↳ for fixed $r > 0$: $J(r) \sim N\left(0, \frac{e^{2rc} - 1}{2c}\right)$ Ornstein-Uhlenbeck process

$J(r) = w(r) + c\int_0^r e^{(r-s)c} w(s) ds$

$n(\hat{\rho}_n - \rho_n) \xrightarrow{d} \left[\int J dw + \frac{1}{2}\left(1 - \frac{\sigma_u^2}{\sigma^2}\right)\right]\left[\int J^2 dr\right]^{-1}$

σ: long run variance of u_t

$n\rho_n = n + c \quad n(\hat{\rho}_n - 1 - c) \xrightarrow{d} \left[\int J dw + \frac{1}{2}\left(1 - \frac{\sigma_u^2}{\sigma^2}\right)\right]\left[\int J^2 dr\right]^{-1}$

$\sigma_u^2 = \sigma^2 \Rightarrow n(\hat{\rho}_n - 1) \xrightarrow{d} c + \left[\int J dw\right]\left[\int J^2 dr\right]^{-1} \quad c < 0$

Explosive series $\quad \rho_n = 1 + \dfrac{c}{n^\alpha} \quad c > 0$

$n \to \infty \Rightarrow \rho_n \to 1 \qquad$ Fixed n: $\rho_n > 1$

$\rho > 1, y_0 = 0 \Rightarrow \dfrac{\rho^n}{\rho^2 - 1}(\hat{\rho} - \rho) \xrightarrow[n\to\infty]{d} C \to$ Cauchy

C: Cauchy $\qquad \rho_n^2 - 1 = \dfrac{2c}{n^\alpha} + \dfrac{c^2}{n^{2\alpha}}$

$\Rightarrow \dfrac{\rho_n^n}{\rho_n^2 - 1} = \dfrac{\rho_n^n}{2cn^{-\alpha}} = \rho_n^n n^\alpha / 2c$

$\Rightarrow (\rho_n^n n^\alpha / 2c)(\hat{\rho}_n - \rho_n) \xrightarrow{d} C \qquad$ MGF → PDF

Size of test: rejection rate of null when null is true → significance level

Power of test: rejection rate of null when alternative is true

making wrong decision at the 5% level

5% → 1.95 permit to make wrong decision at the 5% level

smaller size → more Conservative ⇒ ↓ Power of test

Size distortion: over-reject the null = oversize distortion

undersize distortion acceptable = MK less mistake

AR(1): size distortion $\xrightarrow{T\to\infty}$ goes away

$y_t = a + \rho y_{t-1} + u_t \quad t_{\hat{\rho}} = \dfrac{\hat{\rho}}{\sqrt{V(\hat{\rho})}} \quad \underline{\rho < 1}$ (I)

Panel data: ① size distortion exists as $\begin{cases} \text{Fixed } T \\ N \to \infty \end{cases}$

↳ multiplication of size distortion of (I) to N.

② if Hetroskedasticity for $\begin{cases} \text{Fixed } N \\ T \to \infty \end{cases}$ ⇒ more size distortion

---

LSDV: ① robust & ② Consistent even if $Cov(x_{it}, a_i) \neq 0$

GLS: (random effect Estimator) → efficient & Consistent

only if $Cov(x_{it}, a_i) = 0$

two metods for $Cov(x_{it}, a_i) = 0$:

① pooled OLS Reg residuals (Breusch & pagan 1980s) BP's LM test

② different b/w LSDV and GLS: Hausman Spec test

Pooled OLS Reg residuals — BP's LM test

$H_0: a_i = a \ \forall i \quad : E(\hat{e}_{it}\hat{e}_{is}) = 0 \ \forall i \to E(\Sigma_{t=1}^T \hat{e}_{it})^2 = E(\Sigma_{t=1}^T \hat{e}_{it}^2)$

$H_A: a_i \neq a \ \forall i \quad : E(\hat{e}_{it}\hat{e}_{is}) \neq 0 \ \forall i \Rightarrow E(\Sigma^T \hat{e}_{it})^2 > E(\Sigma_{t=1}^T \hat{e}_{it}^2)$

$\hat{e}_{it} = y_{it} - \hat{a} - \hat{b}_{pols} x_{it}$ ← test stat

$LM = \dfrac{NT}{2(T-1)}\left[\dfrac{\Sigma_{i=1}^N (\Sigma_{t=1}^T \hat{e}_{it})^2}{\Sigma_{i=1}^N \Sigma_{t=1}^T \hat{e}_{it}^2} - 1\right]^2 \xrightarrow{d} \chi_1^2$

Hausman's Specification test

$H_0: a_i = a \ \forall i \quad plim_{N,T\to\infty} \hat{b}_{LSDV} = plim_{N,T\to\infty} \hat{b}_{GLS}$

$H_A: plim_{N,T\to\infty} (\hat{b}_{GLS} - \hat{b}_{LSDV}) \neq 0$

test stat → $(\hat{b}_{GLS} - \hat{b}_{LSDV})'\left[Var(\hat{b}_{GLS} - \hat{b}_{LSDV})\right]^{-1}$

$\times (\hat{b}_{GLS} - \hat{b}_{LSDV})' \xrightarrow{d} \chi_k^2 \quad : \dfrac{(\hat{b}_{GLS} - \hat{b}_{LSDV})^2}{Var(\hat{b}_{GLS} - \hat{b}_{LSDV})} \xrightarrow{d} \chi_1^2$

Econometrics summary note (9)

$Var(\hat{b}_{GLS} - \hat{b}_{LSDV}) = \hat{\sigma}_u^2 \left[\sum^N \sum^T \tilde{x}_{it}^2\right]^{-1} - \left[\sum^N \sum^N \sum^T \hat{w}_{ij}\, \tilde{x}_{it} x_{jt}\right]^{-1}$

$[\hat{w}_{ij}] = \hat{\Omega}^{-1}$

Dynamic Panel Regression

$y_{it} = y_i^* + (y_{i0} - y_i^*) e^{-\beta t}$ (e.g. logarithm of income, wage)

$y_i^* \longrightarrow$ steady state outcome

$\Rightarrow y_{it} = y_i^*(1-e^{-\beta}) + e^{-\beta} y_{it-1} \Rightarrow \rho = e^{-\beta},\ a_i = y_i^*$

$\Rightarrow y_{it} = a_i(1-\rho) + \rho y_{it-1} + u_{it} \quad t = 1, \dots, T$ (time dependent growth model)

$u_{it} \sim$ exogenous iid (1) measurement Error (2) transitory Shock

General Regression

Serial Correlation b/c regression is not balanced

example $y_{it} = a_i + bx_{it} + u_{it}$

① $\begin{cases} y_{it} \sim \text{linear trend} \\ b \neq 0 \\ x_{it} \sim \text{linear trend or not} \\ x_{it} \sim \text{linear trend} \\ u_{it} \sim \text{linear trend} \end{cases} \Rightarrow$ include trend in Reg.

$\Rightarrow \begin{cases} y_{it} = a_{iy} + c_i t + y_{it}^o \\ x_{it} = a_{ix} + d_i t + x_{it}^o \end{cases}$

Cant write $c_i t = a + b(d_i t) + e_{it}$

reason { ① dependent $\rightarrow$ not stochastic
② if stoch. : $\zeta_{it} = c_i t + \varepsilon_{it}$
$\zeta_{it} = a + b(d_i t) + e_{it}$ Component

as long as $c_i \neq bd_i \Rightarrow$ Error term has linear trend

✱ interest should be: $y_{it}^o \overset{?}{\sim} x_{it}^o$

eliminate trend term by linear trend in Reg.

interest in analyzing growth rate $\Rightarrow$ First difference

$\Delta y_{it} = a_i + b\Delta x_{it} + \text{error}_{it}$

② $\begin{cases} y_{it} \sim \text{serially correlated} \\ x_{it} \sim \text{not serially correlated} \\ \Rightarrow u_{it} \sim \text{serially correlated} \end{cases}$

$\Rightarrow$ run dynamic panel Regression

$y_{it} = a_i + \rho y_{it-1} + bx_{it} + \gamma x_{it-1} + \varepsilon_{it}$

due to $u_{it} = \rho u_{it-1} + \varepsilon_{it}$ AR(1)

③ $\begin{cases} y_{it} \sim \text{not serially correlated} \\ x_{it} \sim \text{very persistent } (\rho \text{ near to unity}) \\ b \neq 0 \end{cases}$

$\Rightarrow$ unbalanced Regression

to balance out : $Cov(u_{it}, x_{it}) = -a$ (NEG)

eg $y_{it} = a_i + bx_{it} + u_{it}$ ① serially Correlate ② $Cov(u_{it}, x_{it}) < 0$

y_{it}: Stock return or depriciation rate $\equiv$ white noise

x_{it}: interest rate differential UIP or dividend ratio (stock return)

Modeling Dynamic Panel Regression

M_1: $y_{it} = a_i + \beta x_{it} + u_{it}$ $\quad u_{it} = \rho u_{it-1} + \varepsilon_{it}$

$u_{it} \sim y_{it}$ Correl in levels

$\begin{cases} z_{it} = a_i + u_{it} \\ z_{it} = y_{it} - \beta x_{it} \\ u_{it} = \rho u_{it-1} + \varepsilon_{it} \end{cases}$

(10)

M_2: $y_{it} = a_i + \rho y_{it-1} + \beta x_{it} + \varepsilon_{it}$

$\Rightarrow \begin{cases} y_{it} = a_i + u_{it} \\ u_{it} = \rho u_{it-1} + e_{it} \\ e_{it} = \beta x_{it} + \varepsilon_{it} \end{cases}$ $\quad x_{it}$ Correl with quasi differenced $x_{it}^{\circ} \sim y_{it}$

$\Rightarrow$ if M_2 true M_1 is not misspecified, since

M_1: $y_{it} = a_i + \rho y_{it-1} + \beta x_{it} + \gamma x_{it-1} + \varepsilon_{it}$ M_3

simply $\gamma = 0$

But if M_1 true $\Rightarrow$ M_2 is misspecified

$\Rightarrow$ inconsistent estimates $\hat{\beta}$ and $\hat{\rho}$

M_1 or M_3 nests M_2

Economic interpretation

M_2 follow white noisy process (no serial Correl)

M_1 no white noise restriction

Nickel Bias

showed left $T \to \infty \ = 0$ LSDV

in LSDV two kinds of Bias: $\tilde{y}_t = a + \rho \tilde{y}_{t-1} + \tilde{u}_t$ transfo

① Correlation b/w $\tilde{y}_{t-1}$ and $\tilde{u}_t$ (regressor & Error within) $\xrightarrow{N \to \infty} 0$ in panel

② assymetric distribution of $\hat{\rho}$

$N \to \infty$ part ① remains permanent : $-\frac{1+\rho}{T} + O(T^{-2})$

$\rho = 1$ $\quad y_t = \sum_{s=1}^t u_s \quad E\left(\sum_{t=2}^T y_{it-1}^2\right) = \frac{\sigma_u^2}{2} T$

$\Rightarrow E(\hat{\rho}_{LSDV} - 1) = -\frac{3}{T} + O(T^{-2}) < -\frac{2}{T} + O(T^{-2})$

Inconsistency in POLS

$y_{it} = a + \rho y_{it-1} + e_{it} \quad e_{it} = a_i - a + u_{it}$

$E([a_i - a] + y_{it-1}^o)([a_i - a](1-\rho) + u_{it}) = \sigma_a^2 (1-\rho)$

$E(\sum^N \sum^T y_{it-1})^2 = \sigma_a^2 + \sigma_y^2 \quad \eta = \frac{\sigma_a^2}{\sigma_u^2}$

Asymptotic Dist of LSDV

$\hat{\rho}_{LSDV} - \rho \xrightarrow{d} N(0, 1-\rho^2)$

$\frac{N}{T} \to c,\ N, T \to \infty$: $\sqrt{NT}(\hat{\rho}_{LSDV} - \rho) \xrightarrow{d} -(1+\rho)c + N(0, 1-\rho^2)$

$\frac{N}{T} \to \infty,\ N, T \to \infty$: $\sqrt{NT}(\hat{\rho}_{LSDV} - \rho) \xrightarrow{p} \infty$

$\frac{N}{T} \to 0,\ N, T \to \infty$: $\sqrt{NT}(\hat{\rho}_{LSDV} - \rho) \xrightarrow{d} N(0, 1-\rho^2)$

Panel Dummy

$\begin{cases} y_{it} : \text{Nominal wage} \\ S_i : \text{treatment variable / Dummy} \end{cases}$

$\begin{cases} y_{it} = \alpha + \beta S_i + u_{it} \\ u_{it} = \rho u_{it-1} + \varepsilon_{it} \\ \varepsilon_{it} \sim iid\ N(0, \sigma^2) \end{cases}$ ① Transform $\frac{\Sigma}{N}$ ② $S_i = \begin{cases} 0 & i = 1, \dots, \frac{N}{2} \\ 1 & i = \frac{N}{2}+1, \dots, N \end{cases}$

③ $E\left[\sum^N \tilde{S}_i \left(\sum^T \tilde{u}_{it}\right)\right]^2 = \frac{NT}{4} \frac{\sigma^2}{(1-\rho)^2} \quad E\left(\sum^T \tilde{u}_{it}\right)^2 = \frac{\sigma^2}{(1-\rho)^2}$

Good term $G_{NT} = \frac{\sum^{NT} y_{it-1} u_{it}}{\sum^{NT} \tilde{y}_{it-1}}$ $\quad$ Bad term $N_{NT} = \frac{-\frac{1}{T} \sum^N (\sum^T y_{it-1})(\sum^T u_{it})}{\sum^{NT} \tilde{y}_{,T-1}^2}$

$\sqrt{NT}\, G_{NT} \xrightarrow{d} N(0, V^2)$

$N_{NT} = \frac{-\frac{1}{T} \frac{\sum^N O_p(1) O_p(1)}{N}}{O_p(1)} = \frac{1}{\sqrt{N}} O_p(1) \times \frac{1}{T} = O_p\left(\frac{1}{\sqrt{N}T}\right)$

$\Rightarrow \sqrt{NT}(\hat{\rho}_{LSDV} - \rho) = O_p(1) + O_p\left(\frac{1}{\sqrt{T}}\right)$ $\quad T \to \infty$ ignore second term

Econometrics Summary note

$\begin{cases} x_t & AR(1) \\ u_t & AR(1) \end{cases}$ $\begin{cases} y_t = bx_t + u_t \\ u_t = \rho u_{t-1} + \varepsilon_t \quad \varepsilon_t \sim iid(0,\sigma^2) \end{cases}$

$\Rightarrow E(x_s x_t u_t u_s) = \rho^{2(t-s)} \sigma_x^2 \sigma_u^2$

$\Rightarrow E(\sum x_t u_t)^2 = T\sigma_x^2\sigma_u^2 + \frac{2\sigma_x^2\sigma_u^2\rho^2}{1-\rho^2} T + O(1) =$

$T\sigma_x^2\sigma_u^2 \left(\frac{1+\rho^2}{1-\rho^2}\right) + O(1)$

$\Rightarrow \sqrt{T}(\hat{b}-b) \xrightarrow{d} N(0, \omega_b^2)$, $\omega_b = \frac{\sigma_x^2\sigma_u^2\left(\frac{1+\rho^2}{1-\rho^2}\right)}{\sigma_x^2\sigma_x^2} = \left(\frac{1+\rho^2}{1-\rho^2}\right)\frac{\sigma_u^2}{\sigma_x^2} > \frac{\sigma_u^2}{\sigma_x^2}$

effect on Variance

$\Rightarrow \sqrt{T}(\hat{b}-b) \xrightarrow{d} N(0, \sigma_u^2 (x'x)^{-1})$ does not work here

---

Strongly Stationary time series (ts) $\{Z_t\}_{t=-\infty}^{t=\infty}$

if join prob dist. of any set of k obs $\{Z_t, \dots, Z_{t+k}\} \equiv$ Same regardless of origin

Weakly Stationary $\{Z_t\}$ (i) $E(Z_t) < \infty$ (ii) $Cov(Z_t, Z_{t-k}) < \infty$

(iii) $Cov(Z_t, Z_{t-k}) = \gamma_k$ (not time varying)

Ergodicity (i) strongly stationary (ii)

(ii) $\lim_{k\to\infty} |E[f(Z_t, Z_{t+1}, \dots, Z_{t+a}) g(Z_{t+k}, Z_{t+k+1}, \dots, Z_{t+k+b})]|$

$= |E f(Z_t, Z_{t+1}, \dots, Z_{t+a})| \, |E g(Z_{t+k}, Z_{t+k+1}, \dots, Z_{t+k+b})|$

another def (i) $\sum_{t=1}^T y_t \xrightarrow{P} E(y_t)$

(ii) $\sum_{j=0}^{\infty} |\gamma_j| < \infty$

{Greek word} → work path ≡ Same behavior $\begin{cases} \text{AVG (time) as} \\ \text{AVG (Syst. state)} \end{cases}$

e.g $\lim_{k\to\infty} |E(Z_t Z_{t+k})| = \lim_{k\to\infty} |\rho^k \sigma_z^2| = 0 \neq E|Z_t| \, |E(Z_{t+k})|$

for $Z_t = \rho Z_{t-1} + u_t \quad u_t \sim iid(0,1)$

Ergodic Theorem (i) Z_t: strongly stationary
(ii) Ergodic
(iii) $E|Z_t| = c < \infty$ (finite cte)

$\Rightarrow \bar{Z}_T = \frac{\sum Z_t}{T} \xrightarrow{a.s} \mu = E(Z_t)$

martingale Sequence $E(Z_t | Z_{t-1}, Z_{t-2}, \dots) = Z_{t-1}$

eg $Z_t = Z_{t-1} + u_t$, $E(Z_t | Z_{t-1}, Z_{t-2}, \dots) = Z_{t-1}$

Martingale difference sequence $E(Z_t | Z_{t-1}, Z_{t-2}, \dots) = 0$

White Noise process (i) stationary (ii) but not auto corr process

Long run Variance $\begin{cases} u_t = \rho u_{t-1} + e_t \Rightarrow \\ e_t \sim \text{white noise}(0, \sigma^2), \sigma^2 < \infty \\ \mu_T = \frac{\sum^T u_t}{T} \end{cases}$

$\Rightarrow \mu_T = \frac{\sum^T u_t}{\sqrt{T}} \xrightarrow{d} N\left(0, \frac{\sigma^2}{(1-\rho)^2}\right)$

Estimation of Long Run Variance **HAC Estimation**

Estimation of Long Run Variance

$u_t \sim AR(T)$ or $ARMA(p,q)$

Problem estimate $\hat{\sigma}^2$ not possible $\begin{cases} \text{Homosc.: } \frac{T(T-1)}{2} + 1 \text{ params} \\ \text{Hetrosc.: } \frac{T(T-1)}{2} + T \text{ params} \end{cases}$

Solution: impose regularity:
(i) Ergodic & stationary
(ii) $E(u_t u_{t-k}) \simeq 0$ for large k

e.g $E\left(\frac{\sum_t^{T-1}\sum_{s=t}^{T} u_t u_s}{T}\right) = E\left(\frac{\sum_t^{T-1}\sum_{s=t}^{t+k} u_t u_s}{T}\right) \longrightarrow$ discard $s > t+k$

Newey and West Estimator

$\omega^2 = \omega_0^2 + \sum_{j=1}^{\infty}(\omega_j^2 + \omega_{-j}^2)$ $\quad \omega_j^2 = E\, u_t u_{t-j}$

$= \omega_0^2 + \sum_{j=1}^{k}(\hat{\omega}_j^2 + \hat{\omega}_{-j}^2)$ truncate

Andrews 1991 $\hat{\omega}^2 = \hat{\omega}_0^2 + \sum_{j=1}^{k} w_j(\hat{\omega}_j^2 + \hat{\omega}_{-j}^2)$

optimal weight

Newey and West 1992 $w_j = 1 - \frac{j}{k+1}$ $\quad k = int(T^{1/3})$

Berlett Kernel weight $\Rightarrow$ Consistent

Pre whitening HAC Estimator
1. parametric
2. Andrews & Monahan's

$e_t \sim \begin{cases} \text{① Stationary} \\ \text{② Ergodic} \end{cases}$ $\begin{cases} u_t = \rho u_{t-1} + e_t \\ E\left(\frac{\sum u_t}{\sqrt{T}}\right)^2 = \frac{w_e^2}{(1-\rho)^2} \longrightarrow \text{long run var of } e_t \end{cases}$

$\hat{\omega}_u^2 = \frac{\hat{\omega}_e^2}{(1-\hat{\rho})^2} \longrightarrow$ estimate first & replace

---

Matrix Form $y_t = X_t' b + u$

$\sqrt{T}(\hat{b}-b) \xrightarrow{d} N(0, V_b)$

$V_b = \left(\frac{\sum X_t X_t'}{T}\right)^{-1} \left(\frac{\sum_s^T \sum_t^T E(u_s X_s)(u_t X_t')}{T}\right) \times \left(\frac{\sum X_t X_t'}{T}\right)^{-1}$

$\zeta_t = u_t \cdot X_t = (u_t x_{1t}, u_t x_{2t}, \dots, u_t x_{kt})$

$\Omega^2 = \Omega_0 + \Omega_j + \Omega_{-j}$

$\hat{\Omega}^2 = \hat{\Omega}_0 + \sum_{j=1}^{k} w_j(\hat{\Omega}_j + \hat{\Omega}_{-j})$

$\Rightarrow \hat{V}_{\hat{b}} = \left(\frac{\sum X_t X_t'}{T}\right)^{-1} \hat{\Omega}^2 \left(\frac{\sum X_t X_t'}{T}\right)^{-1}$

Alternative approach AR(p)

① $y_t = x_t' b + u_t$ ② $u_t = \sum_{j=1}^{p} \rho_j u_{t-j} + e_t$

$\rho_i y_{t-i} = \rho_i X_{t-i}' b + \rho_i u_{t-i} \quad 1 \le i \le p$ ⊛

$\Rightarrow$ subtract ① − ⊛ : $y_t = X_t' b - \sum_{j=1}^{p} \rho_j X_{t-j}' b + \sum_{j=1}^{p} \rho_j y_{t-j} + e_t$

$= Z_t \gamma + e$

$Z_t = (X_t, X_{t-1}, \dots, X_{t-p}, y_{t-1}, \dots, y_{t-p})$

$\Rightarrow \sqrt{T}(\hat{\gamma} - \gamma) \xrightarrow{d} N(0, \sigma_e^2 Q_z^{-1}) \quad Q_z = \text{plim}_{T\to\infty} \frac{Z'Z}{T}$

Econometrics summary note

$AR(1): u_t = \rho u_{t-1} + e_t$

$E(uu') = \Omega_{TT} = \frac{\sigma_e^2}{1-\rho^2}\begin{bmatrix} 1 & \rho & \cdots & \rho^{T-1} \\ \rho & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \rho \\ \rho^{T-1} & \cdots & \rho & 1 \end{bmatrix}$ $\quad \Omega = C\Lambda C'$, $C'C = I$

$\Omega^{-1} = C\Lambda^{-1}C' = P'P \Rightarrow Py = Pxb + Pu \Rightarrow y^* = X^*b + u^*$

$\hat{b}_{gls} = (X^{*\prime}X^*)^{-1}X^{*\prime}y^*$ or $\hat{b}_{gls} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y$

$E u^* u^{*\prime} = P\Omega P' = I \Rightarrow \sqrt{n}(\hat{b}_{gls} - b) \xrightarrow{d} N(0, (\frac{X'\Omega^{-1}X}{n})^{-1})$

Feasible GLS $(\hat\Omega)$: $\hat{b}_{fgls} = (X'\hat\Omega X)^{-1}X'\hat\Omega^{-1}y$

Heteroscedasticity i or t ① $E u_i^2 = \sigma_i^2 \neq \sigma_j^2 = E u_j^2$
② $E[u_i u_j] = 0$

$\Rightarrow X'E(uu')X = X'\Omega X \neq X'X$

$[x_1, \ldots, x_n]\begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ \vdots & \sigma_2^2 & \cdots & \vdots \\ \vdots & & \ddots & \sigma_2^2 \\ 0 & \cdots & 0 & \sigma_n^2 \end{bmatrix}\begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix} = \sum_{i=1}^n \sigma_i^2 X_i X_i$

$\Rightarrow \sqrt{n}(\hat b - b) \xrightarrow{d} N(0, V_b)$, $V_b = (X'X)^{-1}(\sum_{i=1}^n \sigma_i^2 X_i X_i)(X'X)^{-1}$

White heteroskedasticity ① $\hat\sigma_i^2 = \hat u_i^2$ for σ_i^2
② is Consistent

$\begin{cases} y_i = a * x_i + u_i \\ u_i = \beta x_i + e_i \end{cases}$ $\Rightarrow \hat a = a + (X'X)^{-1}X'u = a + \beta + (X'X)^{-1}X'e$

$E(u_i e_j) = 0 \ \forall i,j$ $\quad E(\hat a - a | x) = \beta \neq 0 \Rightarrow$ endogeneous

---

① test of Exogeneity : Hausman

② Lagged var bias (lagged dependent):

$y_t = a + y_t^o$, $y_t^o = \rho y_{t-1}^o + u_t$ $\Rightarrow y_t = a(1-\rho) + \rho y_{t-1} + u_t \Rightarrow \tilde y_t = \rho \tilde y_{t-1} + \tilde u_t$

$E(\tilde y_{t-1}\tilde u_t) \neq 0$ but $O_p(T^{-1}) \to 0$, $t \to \infty$

③ Measurement Error: ①

True model: $y_i = \alpha x_i + u_i$

observe (measurement error): $x_i^* = x_i + e_i$

run: $y_i = \alpha x_i^* + u_i$

Behavior of true phenomenon: $y_i = \alpha(x_i + e_i) \underbrace{-\alpha e_i + u_i}_{v_i} = \alpha x_i^* + v_i$

$E(v_i x_i^*) = E(u_i - \alpha e_i)(x_i + e_i) \neq 0$

Solution to endogeneity

① Control variable: $y_i = \alpha x_i + w_i'\gamma + v_i$
$w_i = (w_{1i}, \ldots, w_{ki})' \rightarrow$ proxy for u_i

Problem Don't know how many

② instrument variable z_i: ① $E(x_i z_i) \neq 0$
② $E(z_i u_i) = 0$

$\Rightarrow \hat\alpha_{IV} = (Z'X)^{-1}Z'Y = \alpha + (Z'X)^{-1}Z'u$

[1] $plim(\hat\alpha_{IV} - \alpha) = 0 \Rightarrow$ Consistent

[2] Asympt. Variance: $E(\hat\alpha_{IV} - \alpha)(\hat\alpha_{IV} - \alpha)' = (Z'X)^{-1}Z'\Omega_u Z(Z'X)^{-1}$

Measurement Error ① $\Rightarrow$ $\begin{cases} y_i = \alpha x_i^* + v_i \\ x_i^* = x_i + e_i \\ v_i = -\alpha e_i + u_i \end{cases}$

Right IV $\begin{cases} z_i = \beta x_i + m_i \\ E(m_i e_i) = 0 \\ E(m_i u_i) = 0 \end{cases}$

Pooling

why?	why not
① Economic theory hold all indiv ② more data (cross sec or time series obs) ⇒ ① more eff ② more powerful	① account for indiv. heterogeneity $y_{it} = a_i + bx_{it} + u_{it}$ ② fixed or random ③ a_i observable: (1) gender (2) edu (3) age

Random effect

$\begin{cases} y_{it} = a + bx_{it} + e_{it} \\ e_{it} = a_i - a + u_{it} = \mu_i + u_{it} \end{cases}$

Assump. (A1) $E(\mu_i x_{it}) = 0 \ \forall i$
(A2) $E(\mu_i u_{it}) = 0 \ \forall i$

(A1), (A2) $\Rightarrow$ Consistent, but not Efficient

$\Rightarrow$ reg err $\neq$ iid

$E(e_{i1}e_{i2}) = \sigma_\mu^2$
$E(e_{i1}^2) = \sigma_\mu^2 + \sigma_u^2$ $\Rightarrow$ GLS ① Efficient ② Consistent

Proc. GLS:

① Run. $y_{it} = a + bx_{it} + e_{it} \Rightarrow \hat e_{it}, \hat b_{pols}, \hat a_{pols}$

② Construct:

ⓐ $\hat\sigma_e^2 = \frac{\sum^{NT}(y_{it} - \hat a_{pols} - \hat b_{pols}x_{it})^2}{NT} = \frac{\sum^{NT}\hat e_{it}^2}{NT}$

ⓑ $\hat\mu_i = \frac{\sum^T \hat e_{it}}{T}$ $\quad \hat u_{it} = \hat e_{it} - \hat\mu_i$

ⓒ $\hat\sigma_\mu^2 = \frac{\sum^N(\hat\mu_i - \frac{\sum\hat\mu_i}{N})^2}{N}$

ⓓ $\hat\sigma_u^2 = \frac{\sum^{NT}(\hat u_{it} - \frac{\sum^{NT}\hat u_{it}}{NT})^2}{NT}$

if T small ⇓ use T-1 for T

③ Construct sample Cov matrix:

$\hat\Omega = \begin{bmatrix} \hat\sigma_\mu^2 + \hat\sigma_u^2 & \hat\sigma_\mu^2 & \cdots & \hat\sigma_\mu^2 \\ \hat\sigma_\mu^2 & \hat\sigma_\mu^2 + \hat\sigma_u^2 & \cdots & \vdots \\ \hat\sigma_\mu^2 & \hat\sigma_\mu^2 & \cdots & \hat\sigma_\mu^2 + \hat\sigma_u^2 \end{bmatrix}$

④ Feasible GLS:

$\hat b_{fgls} = (\sum_{i=1}^N X_i'\hat\Omega^{-1}X_i)^{-1}(\sum_{i=1}^N X_i'\hat\Omega^{-1}Y_i)$

$X_i = [x_{1i}, \ldots, x_{Ti}]'$

$Y_i = [y_{1i}, \ldots, y_{Ti}]'$

Src of inconsistency $\rightarrow$ (A1) if not hold
$\equiv$ indiv error ~ Regressor $\neq 0 \equiv$ Corr
$\Rightarrow \hat b_{OLS}, \hat b_{FGLS}$ are inconsistent

including observed indiv effect

even if (A1) not hold if μ_i observed $\Rightarrow$ as regressor

$y_{it} = a + \gamma_1\mu_{1i} + \gamma_2\mu_{2i} + \cdots + bx_{it} + u_{it}$

Econometrics summary note

(15)

Fixed Effects: micro panel

Eyeball approach

① draw graph

② explain relation y_{it}/x_{it} (Single Explanatory Var)

① unique across i
② fixed effects (+ & +)
③ fixed effects (+ but −)
④ heterogeneity (+ but −)
⑤ demean (within transform) $\tilde{y}_{it}=b\tilde{x}_{it}+\tilde{u}_{it}$

more than two explan. variable

$y_{it}=a_i+bx_{it}+cZ_{it}+u_{it}$

Ⓧ plot $\tilde{y}_{it}/\tilde{x}_{it}, \tilde{Z}_{it}$

reason: $\equiv \tilde{y}_{it}=b\tilde{x}_{it}+\tilde{e}_{it}$

$\tilde{e}_{it}=c\tilde{Z}_{it}+\tilde{u}_{it}$

Solution — same $E(\tilde{x}_{it}\tilde{Z}_{it})\neq 0$ $\hat{b}$ inconsistent

① $\tilde{y}_{it}=a_1\tilde{Z}_{it}+\tilde{y}^{+}_{it}$

$\tilde{x}_{it}=a_2\tilde{Z}_{it}+\tilde{x}^{+}_{it}$

$\Rightarrow$ get resid. $\tilde{y}^{+}_{it}, \tilde{x}^{+}_{it} \Rightarrow$ plot

② $\tilde{y}_{it}=b_1\tilde{x}_{it}+\tilde{y}^{*}_{it}$

$\tilde{Z}_{it}=b_2\tilde{x}_{it}+\tilde{Z}^{*}_{it}$

$\Rightarrow$ plot $\tilde{y}^{*}_{it}$ on $\tilde{Z}^{*}_{it}$

mathematically $\equiv$ projection approach

$I-Z(Z'Z)^{-1}Z'=M_Z$ or M_X

Common Time Effect

$y_{it}=a_i+\lambda_t+bx_{it}+u_{it}$ $\hat{b}=?$

Proc

① demeaning over t ($\sum_T$ transform)
$\equiv$ eliminate fixed effect

② Cross sec mean ($\sum_N$ transform)
on $\tilde{y}_{it}=\tilde{\lambda}_t+b\tilde{x}_{it}+\tilde{u}_{it}$

$\Rightarrow \tilde{y}_{it}-\sum^{N}_{N}\tilde{y}_{it}=b(\tilde{x}_{it}-\sum^{N}_{N}\tilde{x}_{it})+(\tilde{u}_{it}-\sum^{N}_{N}\tilde{u}_{it})$

③ evaluate within transfo.:

$y^{\dagger}_{it}=y_{it}-\frac{\sum^{T}}{T}y_{it}-\frac{\sum^{N}}{N}y_{it}+\frac{\sum^{T}\sum^{N}}{NT}y_{it}$

(FE) Fixed effect estimator $\equiv$ Least Square Dummy Variable (LSDV)
$\equiv$ within Group (WG) estimator

DGP
$$\begin{cases} y_{it}=\mu_{y,i}+y^{o}_{it} \\ x_{it}=\mu_{x,i}+x^{o}_{it} \end{cases}$$

$$\begin{cases} \mu_{y,i}=a+b\mu_{x,i}+e_i \\ y^{o}_{it}=\alpha_i+\beta x^{o}_{it}+u_{it} \end{cases}$$

① You run: $y_{i1}=c_1+\gamma_1 x_{i1}+\varepsilon_{i1}$ (OLS) $\Rightarrow \operatorname{plim}_{N\to\infty}\hat{\gamma}_1\neq b$ $\equiv$ inconsistent

Cross sec Reg for t=1

② run cross sectional regression with ts AVG

$\bar{y}_i=c+\gamma\bar{x}_i+\bar{\varepsilon}_i$ $\bar{y}_i=\frac{\sum^{T}y_{it}}{T}$ $\bar{x}_i=\frac{\sum^{T}x_{it}}{T}$

(16)

DGP
$$\begin{cases} y_{it}=a_i+y^{o}_{it} \\ y^{o}_{it}=\rho y^{o}_{it-1}+u_{it} \\ u_{it}\sim iid(0,\sigma^2) \end{cases}$$

run pols: $y_{it}=a+\rho y_{it-1}+e_{it}$

$\Rightarrow \rho<1$ $\rho_{POLS}\rightarrow$ inconsistent

Dynamic panel Regression

DID: $y_{it}=a_i+\lambda_t+bx_{it}+u_{it}$ (I) Difference in Difference

$u_{it}=\rho u_{it-1}+\varepsilon_{it}$

① $x_{it}\sim$ Exogen $\Rightarrow$ ① b_{LSDV} Consist.
② stat inf. (t-value) $\rightarrow$ issue (Conditioned)

② $T>>0$ large
more efficient estimator by running Dynamic Panel Regression

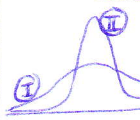
$y_{it}=(1-\rho)a_i+\rho\lambda_t+\rho\lambda_{t-1}+\rho y_{it-1}+bx_{it}-b\rho x_{it-1}+\varepsilon_{it}$ (II)

or $\equiv$ $y_{it}=\alpha_i+\theta_t+\rho y_{it-1}+bx_{it}+\gamma x_{it-1}+\varepsilon_{it}$

Dynamic

within transfo $\rightarrow y^{\dagger}_{it}=\rho y^{\dagger}_{it-1}+bx^{\dagger}_{it}+\gamma x^{\dagger}_{it-1}+\varepsilon^{\dagger}_{it}$

Dynamic



Consistency ① ρ_{LSDV} inconsistent
② γ_{LSDV} inconsistent
Param of interest ⇐ ③ β_{LSDV} Consistent
↳ t-stat be Careful

Benchmark Model (*)
① Strongly exogen
② single Regressor
③ Fixed effect

model $y_{it}=a_i+bx_{it}+u_{it}$

Assump. ① $E(x_{it}u_{js})=0$ $\forall i,j,s,t$

t-stat to test hypoth: $t^{o}_{b}=\frac{\hat{b}}{\sqrt{var(\hat{b})}}$ ← Limiting dist

---

$y_s=bx_s+u_s$ s = i or t

$\hat{b}-b=(\frac{1}{\sqrt{S}})\frac{A_S}{B_S}$ $A_S\xrightarrow{d}N(0,\Omega^2_A)$ $B_S\xrightarrow{P}Q_B$ as $s\to\infty$

$\Rightarrow$ Cramer's theorem: $\sqrt{S}(\hat{b}-b)\xrightarrow{d}N(0,Q_B^{-1}\Omega^2_A Q_B^{-1})$

$\Rightarrow \frac{\sqrt{S}(\hat{b}-b)}{\sqrt{Q_B^{-1}\Omega^2_A Q_B^{-1}}}\xrightarrow{d}N(0,1)\Rightarrow\frac{\sqrt{S}\hat{b}}{\sqrt{Q_B^{-1}\Omega^2_A Q_B^{-1}}}\xrightarrow{d}N(\frac{\sqrt{S}b}{\sqrt{Q_B^{-1}\Omega^2_A Q_B^{-1}}},1)$

power of test larger when { How of reject null when Alt.=True
① true value of |b| larger
② Var b is getting smaller
③ number of obs, S, getting larger $\Rightarrow$ Only Control

For panel $\sqrt{NT}(\hat{b}_{panel}-b)\xrightarrow{d}N(0,Q_B^{-1}\Omega^2_A Q_B^{-1})$

I(1) $\equiv$ non stationary $\rightarrow$ need large cross sec $\Rightarrow(\int B_x du)(\int B_x du)^{2}$

↳ Convergence rate is T not $\sqrt{T}$ $\Rightarrow$ min Condition

For ur panel: $\sqrt{N}>\frac{T_S}{T}$ $T_S\rightarrow$ Just time Seris, $T\rightarrow$ panel with missing data

Econometrics summary note

Linear non linear Restriction

$$y = Xb + u = x_1 b_1 + x_2 b_2 + u$$

$$\sqrt{n}(\hat{b} - b) \xrightarrow{d} N(0, \Sigma_b)$$

$$\sqrt{n}\begin{pmatrix} \hat{b}_1 - b_1 \\ \hat{b}_2 - b_2 \end{pmatrix} \xrightarrow{d} N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{11}^2 & \sigma_{12} \\ \sigma_{21} & \sigma_{22}^2 \end{bmatrix} \right)$$

Linear Restriction $\alpha_0 \hat{b}_1 + \alpha_1 \hat{b}_2 = \alpha_2$

or $f(\hat{b}_1, \hat{b}_2) = \alpha_0 \hat{b}_1 + \alpha_1 \hat{b}_2 + \alpha_2 = \hat{\gamma}$

Taylor Expansion $f(\hat{b}_1, \hat{b}_2) = f(b_1, b_2) + \frac{\partial f(b_1,b_2)}{\partial b_1}(\hat{b}_1 - b_1)$

$$+ \frac{\partial f(b_1,b_2)}{\partial b_2}(\hat{b}_2 - b_2) + \frac{1}{2}\frac{\partial^2 f(b_1,b_2)}{\partial b_1^2}(\hat{b}_1 - b_1)^2 + \frac{1}{2}\frac{\partial^2 f(b_1,b_2)}{\partial b_2^2}(\hat{b}_2 - b_2)^2$$

$$+ \frac{1}{2}\frac{\partial^2 f(b_1,b_2)}{\partial b_1 \partial b_2}(\hat{b}_1 - b_1)(\hat{b}_2 - b_2) + \cdots$$

$$\left\{ \begin{array}{l} \frac{\partial f(b_1,b_2)}{\partial b_1} = \alpha_0 \\ \frac{\partial f(b_1,b_2)}{\partial b_2} = \alpha_1 \\ \frac{\partial^2 f(b_1,b_2)}{\partial b_1^2} = \frac{\partial^2 f(b_1,b_2)}{\partial b_2^2} = \frac{\partial^2 f(b_1,b_2)}{\partial b_1 \partial b_2} = 0 \equiv \text{cross} = 0 \end{array} \right.$$

$$\Rightarrow \hat{\gamma} = \alpha_0 b_1 + \alpha_1 b_2 + \alpha_2 + \alpha_0(\hat{b}_1 - b_1) + \alpha_1(\hat{b}_2 - b_2) = \alpha_0 b_1 + \alpha_1 b_2 + \alpha_2 + (\alpha_0, \alpha_1)\begin{bmatrix} \hat{b}_1 - b_1 \\ \hat{b}_2 - b_2 \end{bmatrix}$$

$$\Rightarrow \sqrt{n}(\hat{\gamma} - \gamma) = [\alpha_0 \ \alpha_1]\begin{bmatrix} \sqrt{n}(\hat{b}_1 - b_1) \\ \sqrt{n}(\hat{b}_2 - b_2) \end{bmatrix} \xrightarrow{d} N\left(0, [\alpha_0 \ \alpha_1]\begin{bmatrix} \sigma_{11}^2 & \sigma_{12} \\ \sigma_{12} & \sigma_{22}^2 \end{bmatrix}\begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix}\right)$$

$$\Rightarrow \frac{\sqrt{n}(\hat{\gamma} - \gamma)}{\sqrt{\sigma_\gamma^2}} \xrightarrow{d} N(0,1)$$ or alternatively

$R = [\alpha_0 \ \alpha_1]$ $q = \alpha_2$ $W = (Rb - q)'[\sigma^2 R (X'X)^{-1} R']^{-1}(Rb - q)$

$W \xrightarrow{d} \chi_1^2$

$\sigma^2 (X'X)^{-1} = \Sigma_b$

Non Linear Restriction second & cross $\neq 0$

$$f(\hat{b}_1, \hat{b}_2) - f(b_1, b_2) = \frac{\partial f(b_1,b_2)}{\partial b_1}(\hat{b}_1 - b_1) + \frac{\partial f(b_1,b_2)}{\partial b_2}(\hat{b}_2 - b_1) + R_n$$

$$\hat{\gamma} - \gamma = \frac{\partial f(b_1,b_2)}{\partial b_1}(\hat{b}_1 - b_1) + \frac{\partial f(b_1,b_2)}{\partial b_2}(\hat{b}_2 - b_2) + R_n$$

$$\left\{ \begin{array}{l} \sqrt{n}(\hat{b}_1 - b_1) = O_p(1) \\ \hat{b}_1 - b_1 = O_p(\frac{1}{\sqrt{n}}) \\ (\hat{b}_1 - b_1)(\hat{b}_2 - b_2) = O_p(n^{-1}) \\ (\hat{b}_1 - b_1)^2 = O_p(n^{-1}) \end{array} \right. \Rightarrow R_n = O_p(n^{-1}) \neq 0$$

$$\sqrt{n}(\hat{\gamma} - \gamma) = \frac{\partial f(b_1,b_2)}{\partial b_1}\sqrt{n}(\hat{b}_1 - b_1) + \frac{\partial f(b_1,b_2)}{\partial b_2}\sqrt{n}(\hat{b}_2 - b_2) + O_p(\frac{1}{\sqrt{n}})$$

teilor streight

$$\Rightarrow \sqrt{n}(\hat{\gamma} - \gamma) = \left[\frac{\partial f(b_1,b_2)}{\partial b_1} \ \ \frac{\partial f(b_1,b_2)}{\partial b_2}\right]\begin{bmatrix} \sqrt{n}(\hat{b}_1 - b_1) \\ \sqrt{n}(\hat{b}_2 - b_2) \end{bmatrix} + O_p(\frac{1}{\sqrt{n}})$$

$$\xrightarrow{d} N\left(0, \left[\frac{\partial f(b_1,b_2)}{\partial b_1} \ \ \frac{\partial f(b_1,b_2)}{\partial b_2}\right]\begin{bmatrix} \sigma_{11}^2 & \sigma_{12} \\ \sigma_{12} & \sigma_{22}^2 \end{bmatrix}\begin{bmatrix} \frac{\partial f(b_1,b_2)}{\partial b_1} \\ \frac{\partial f(b_1,b_2)}{\partial b_2} \end{bmatrix}\right)$$

e.g. $\frac{\hat{b}_1}{\hat{b}_2} = 0 \Rightarrow \frac{\partial f(b_1,b_2)}{\partial b_1} = \frac{1}{b_1}$

$\frac{\partial f(b_1,b_2)}{\partial b_2} = -\frac{b_1}{b_2^2}$

Sandwich is matrix form of square, since as you multiply something to a value, variance is multiplied to square

procedure order: ① demean & remove common term

e.g when $e_{it} = a + u_{it}$

leads to removal of a

② seperate to good term & Bad term

$$\Rightarrow \sum^{NT} \tilde{e}_{it} \tilde{x}_{it} = \sum e_{it} x_{it} - \sum e_{it} \sum x_{it}$$

Method of moment (MM)

moment Cond. : $E(\xi_t - \mu) = 0$

ξ_t → Random var

μ → unknown mean (param of interest)

Criteria: $\underset{\mu}{\arg\min}\, V_T = \underset{\mu}{\arg\min}\, \frac{2}{T}\sum^T (\xi_t - \mu)^2$ obj func : mean variance

Sol: $\frac{\partial V_T}{\partial \mu} = \frac{-2}{T}\sum^T (\xi_t - \mu) = 0 \Rightarrow \frac{\sum^T}{T}\xi_t = \mu$

example:

$E(\xi_t - \mu) = 0$

$E[(\xi_t - \mu)^2 - \gamma_0] = 0$

$E[(\xi_t - \mu)(\xi_{t-1} - \mu) - \gamma_1] = 0$

$E[(\xi_t - \mu)(\xi_{t-2} - \mu) - \gamma_2] = 0$

unknown: $\mu, \gamma_0, \gamma_1, \gamma_2$

Solve numerically: $\frac{\sum^T}{T}\xi_t,\ \frac{\sum^T}{T}\xi_t^2,\ \frac{\sum}{T}\xi_t\xi_{t-1},\ \frac{\sum^T}{T}\xi_t\xi_{t-2}$

further Restrictions: $\xi_t \to AR(1) \Rightarrow \begin{cases}\gamma_1 = \rho\gamma_0 \\ \gamma_2 = \rho\gamma_0\end{cases}$

⇒ reduce unknown → (γ_0, ρ, μ)

$\psi_T = \left(\frac{\sum^T}{T}\xi_t,\ \frac{\sum^T}{T}\xi_t^2,\ \frac{\sum^T}{T}\xi_t\xi_{t-1},\ \frac{\sum^T}{T}\xi_t\xi_{t-2}\right)'$

$\Rightarrow \begin{cases} E\,\frac{\sum^T}{T}\xi_t^2 = \gamma_0 - \mu^2 \\ E\,\frac{\sum^T}{T}\xi_t\xi_{t-1} = \rho\gamma_0 - \mu^2 \end{cases}$

estimation: $\underset{\mu,\rho,\gamma_0}{\arg\min}\, [\psi_T - \psi(\theta)]'[\psi_T - \psi(\theta)]$

θ → param of interest (true param: μ, ρ, γ_0)

MM is min distance estimator

problem: optimizes second moment rather than both first and second ⇒ need weight ⇒ generalized method of moments (GMM)

Asymptotic properties of MM

Taylor expansion: $\psi_T = \psi(\theta) + \frac{\partial \psi_T(\theta)}{\partial \theta'}(\hat\theta - \theta) + O_p\left(\frac{1}{T}\right)$

$\Rightarrow \sqrt{T}(\hat\theta - \theta) = \sqrt{T}[\psi_T - \psi(\theta)]\, G(\theta)^{-1} + O_p\left(\frac{1}{T}\right)$

$G_T(\theta) = \frac{\partial \psi_T(\theta)}{\partial \theta'}$

$\sqrt{T}(\psi_T - \psi(\theta)) \xrightarrow{d} N(0, \Phi)$

$\sqrt{T}(\hat\theta - \theta) \xrightarrow{d} N(0, G(\theta)^{-1}\Phi\, G(\theta)'^{-1})$

$G_T(\theta) \xrightarrow{p} G(\theta)$

GMM: infeasible generalized version of method of moments

$\underset{\mu,\rho,\gamma_0}{\arg\min}\, [\psi_T - \psi(\theta)]'\,\Phi^{-1}[\psi_T - \psi(\theta)]$

Φ^{-1} → true unknown weighting matrix

feasible: $\underset{\mu,\rho,\gamma_0}{\arg\min}\, [\psi_T - \psi(\theta)]'\, W_T\, [\psi_T - \psi(\theta)] = \underset{\mu,\rho,\gamma_0}{\arg\min}\, G_T'(\theta) W_T G_T(\theta)$

W_T: Consistent estimators of Φ^{-1}

$V_T = [\psi_T - \psi(\theta)]'\, W_T\, [\psi_T - \psi(\theta)]$

GMM: $\frac{\partial V_T(\hat\theta_{GMM})}{\partial \hat\theta_{GMM}} = 2\,G_T(\hat\theta_{GMM})'\, W_T\, [\psi_T - \psi(\hat\theta_{GMM})] = 0$

$\psi(\hat\theta_{GMM}) = \psi_T(\theta) + G_T(\theta)(\hat\theta_{GMM} - \theta) + O_p\left(\frac{1}{T}\right)$

$\Rightarrow G_T(\hat\theta_{GMM})'\, W_T\, [\psi_T - \psi(\hat\theta_{GMM})] = G_T(\hat\theta_{GMM})'\, W_T\, [\psi_T - \psi(\hat\theta_{GMM})]$

$+ G_T(\hat\theta_{GMM})'\, W_T\, G_T(\theta)(\hat\theta_{GMM} - \theta) = 0$

$\Rightarrow (\hat\theta_{GMM} - \theta) = -\{G_T(\hat\theta_{GMM})'\, W_T\, G_T(\theta)\}^{-1} G_T(\hat\theta_{GMM})'$

$\cdot [\psi_T - \psi_T(\hat\theta_{GMM})]$

$\sqrt{T}(\hat\theta_{GMM} - \theta) \xrightarrow{d} N(0, V)$

$V = \frac{1}{T}\{G'WG\}^{-1} G'W\Phi WG\,\{G'WG\}^{-1}$

$\to W = \Phi^{-1} \Rightarrow V = \frac{1}{T}\{G'\Phi^{-1}G\}^{-1} G'\Phi^{-1}G\,\{G'\Phi^{-1}G\}^{-1}$

$= \frac{1}{T}\{G'\Phi^{-1}G\}^{-1}$

Pols ①

model: (two variable)

$$y_{it} = \alpha + \beta x_{it} + \gamma w_i + u_{it}$$

$u_{it} \sim iid(0, \sigma^2)$, $x_{it} \sim iid(0, \sigma_x^2)$, $w_i \sim iid(0, \sigma_w^2)$

$E(x_{it} u_{is}) = 0 \quad \forall i,j,t,s$

$E(w_i u_{jt}) = 0$

$\boxed{\text{if } \alpha_i = \alpha}$

$$\Rightarrow E\left(\frac{Z'uu'Z}{N^2T^2}\right) = \frac{1}{NT}\begin{bmatrix} \sigma^2 & 0 & 0 \\ 0 & \sigma_x^2\sigma^2 & 0 \\ 0 & 0 & \sigma_w^2\sigma^2 \end{bmatrix} = \frac{\Omega}{NT}$$

$$\frac{Z'Z}{NT} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sigma_x^2 & 0 \\ 0 & 0 & \sigma_w^2 \end{bmatrix} = Q$$

$$\sqrt{NT}\begin{bmatrix} (\hat\alpha - \alpha) \\ (\hat\beta - \beta) \\ (\hat\gamma - \gamma) \end{bmatrix} \xrightarrow{d} N(0, Q^{-1}\Omega Q^{-1})$$

$\boxed{\text{if } \alpha_i \neq \alpha}$ $\boxed{\text{But } E(\alpha_i x_{it}) = 0} \to \boxed{\text{critical}}$

trick: First Calculate $Z'u$ and then its square result is $Z'uu'Z$

$\alpha_i = \alpha + \epsilon_i$ → Constant with respect to T

$$\Rightarrow a_{11} = \left(\frac{\Sigma^{NT} e_{it}}{NT}\right)^2 = \boxed{\frac{\sigma_\epsilon^2}{N}} + \frac{\sigma^2}{NT}$$

$$a_{22} = E\left(\frac{\Sigma^{NT} x_{it}^2 u_{it}}{NT}\right)^2 = \frac{\sigma_x^2\sigma^2}{NT}$$

$$a_{33} = E\left(\frac{\Sigma^{NT} w_i u_{it}}{NT}\right)^2 = \frac{\sigma_w^2\sigma_\epsilon^2}{N} + \frac{\sigma_w^2\sigma^2}{NT}$$

$$\Rightarrow E(Z'uu'Z) = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} \frac{\sigma_\epsilon^2}{N} + \frac{\sigma^2}{NT} & 0 & 0 \\ 0 & \frac{\sigma_x^2\sigma^2}{NT} & 0 \\ 0 & 0 & \frac{\sigma_w^2\sigma_\epsilon^2}{N} + \frac{\sigma_w^2\sigma^2}{NT} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{N} & 0 & 0 \\ 0 & \frac{1}{NT} & 0 \\ 0 & 0 & \frac{1}{N} \end{bmatrix}\begin{bmatrix} \sigma_\epsilon^2 + O(T^{-1}) & 0 & 0 \\ 0 & \sigma_x^2\sigma^2 & 0 \\ 0 & 0 & \sigma_w^2\sigma_\epsilon^2 + O(T^{-1}) \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} \sqrt{N} & 0 & 0 \\ 0 & \sqrt{NT} & 0 \\ 0 & 0 & \sqrt{N} \end{bmatrix}\begin{bmatrix} \hat\alpha - \alpha \\ \hat\beta - \beta \\ \hat\gamma - \gamma \end{bmatrix} \xrightarrow{d} N(0, Q^{-1}\Omega Q^{-1})$$

→ due to difference $\sqrt{NT}$ & $\sqrt{N}$

$\boxed{\text{model Reality}}$ $y_{it} = y_{it-1} + e_{it}$ $\quad e_{it} \sim iid(0, \sigma^2)$

$y_{i1} = e_{i1}$

$\boxed{\text{run:}}$ $y_{it} = \alpha + \rho y_{it-1} + e_{it}$ $\quad AR(1)$

Pols is Consistent as $N \to \infty$ $\quad \boxed{\text{Trick: Fixed T}}$

$\boxed{\text{run:}}$ $y_{it} = \alpha_i + \rho y_{it-1} + e_{it}$

$O(\frac{1}{T})$ For bad term $\Rightarrow N \to \infty$ $\quad \boxed{\text{not Consisten since Fixed T } O_{p1}}$

$\boxed{\text{Reality model}}$ $y_{it} = \alpha_i + \rho y_{it-1} + \beta x_{it-1} + e_{it}$

$E(x_{it} e_{is}) = 0 \quad \forall i,t,j,s$

$$\Rightarrow \underset{N\to\infty}{plim}(\hat\beta_{FE} - \beta) = -\underset{N\to\infty}{plim}(\hat\rho_{FE} - \rho)\,\underset{N\to\infty}{plim}\frac{\Sigma^{NT}\tilde x_{it-1}\tilde y_{it-1}}{\Sigma^{NT}\tilde x_{it-1}}$$

$\boxed{\text{trick: deduct } \hat\rho_{FE}\tilde y_{it-1} \text{ from both side}}$

②

$$\tilde y_{it} - \hat\rho_{FE}\tilde y_{it-1} = (\rho - \hat\rho_{FE})\tilde y_{it-1} + \beta\tilde x_{it-1} + \tilde e_{it}$$

$$\Rightarrow = \beta x_{it-1} + \underbrace{(\rho - \hat\rho_{FE})\tilde y_{it-1} + \tilde e_{it}}_{\text{error term}}$$ → since bias uncorrelated with variable → indep

$$\Rightarrow plim\,\hat\beta_{FE} - \beta = \frac{\Sigma^{NT}\tilde x_{it-1}[(\rho - \hat\rho_{FE})\tilde y_{it-1} + \tilde e_{it}]}{\Sigma^{NT}}$$

$$= \frac{plim(\rho - \hat\rho_{FE})\,\Sigma^{NT}\tilde x_{it-1}\tilde y_{it-1}}{\Sigma^{NT}\tilde x_{it-1}^2}$$

$$\Rightarrow plim(\hat\beta_{FE} - \beta) = plim(\rho - \hat\rho_{FE})\frac{\Sigma^{NT}\tilde x_{it-1}\tilde y_{it-1}}{\Sigma\tilde x_{it-1}^2}$$

$$= -plim(\hat\rho_{FE} - \rho)\frac{\Sigma^{NT}\tilde x_{it-1}\tilde y_{it-1}}{\Sigma\tilde x_{it-1}^2}$$

---

① what you run is created from sum up of real model: e.g real $y_{it} = \alpha_i + u_{it}$ $\quad u_{it} \sim iid(0,1)$

You run $y_{it} = \alpha_i + \rho y_{it-1} + e_{it}$

$\alpha_i = (1-\rho)\alpha_i \quad e_{it} = u_{it} - \rho u_{it-1}$

when $\rho = 0 \Rightarrow e_{it} = u_{it}$

---

② use intuition

---

③ when $\rho = 1$ $\quad y_{it} = u_{i1} + u_{i2} + \cdots + u_{it}$

$\boxed{\text{Key:}}$ Build intuition and make sense of equations when deriving them

math: ① Form structure
② math intuition: picture
High level ← ③ Econometrics intuition: Data model
↓ missing in Computer
④ math technique intuition of Derivative
⑤ hierarchical Category order of mind (move only by translation)

Redundant but needed due to human Error

Quality work only result of Polish repitition and remove Faults

Economic theory of human behavior Explaining

Short term: only time of study ← ordered, organized world

long term: intuition, IQ, Experi
= internalization Chance
→ result of curiosity

---

matrix derivative $\frac{\partial Ax}{\partial x} = A$ $\quad [T\times T][T\times 1] = [T\times 1]$, $[T\times 1]$, $[T\times T]$

each line derivative of element (Ax) w.r.t. each of x vector element like transpose of x

---

(iii)

$x \sim N(0,1) \Rightarrow x^2 \sim \chi_1^2$

$x_1, \dots, x_n \sim \chi_1^2 \Rightarrow \Sigma x_i \sim \chi_n^2$

$x_1, \dots, x_n \sim N(0, \sigma^2) \Rightarrow \Sigma\frac{x_i^2}{\sigma^2} \sim \chi_n^2$

$x_1 \sim \chi_n^2, x_2 \sim \chi_m^2 \Rightarrow x_1 + x_2 \sim \chi_{n+m}^2$

$\Rightarrow \frac{x_1/n}{x_2/m} \sim F(n,m)$

$Z \sim N(0,1)$, $x \sim \chi_n^2$ $\Rightarrow t_n = \frac{Z}{\sqrt{x/n}} \xrightarrow[n\to\infty]{} N(0,1)$ (Student)

$x \sim t_n \Rightarrow x^2 \sim F(1,n)$

Gama(α,β) $\quad \frac{x^{\alpha-1}e^{-x/\beta}}{\beta^\alpha\Gamma(\alpha)}$

Chisq $\quad \alpha = n/2 \quad \beta = 2$

$$f(x,k) = \frac{1}{2^{k/2}\Gamma(k/2)}x^{\frac{k-2}{2}}e^{-x/2}$$

Student t (t_n):

$$f(x) = \frac{\Gamma(\frac{\upsilon+1}{2})}{\sqrt{\upsilon\pi}\,\Gamma(\upsilon/2)}\left(1 + \frac{x^2}{\upsilon}\right)^{-\frac{\upsilon+1}{2}}$$

Final test

Part II : panel Regression with Fixed Effect

model : $y_{it} = a_i + x_{it}$ fixed effect ⇒ withing group

$x_{it} = \rho x_{it-1} + u_{it}$ $u_{it} \sim iid(0,1)$

$T=3$, $\rho=1$ Nickel Bias?

$y_{it} = (1-\rho) a_i + \rho y_{it-1} + u_{it}$

$\Rightarrow \hat\rho - \rho = \dfrac{\sum \tilde y_{it-1} \tilde u_{it}}{\sum \tilde y_{it-1}^2}$ (Nom) (Denom)

First trik: ① $\tilde u_{it}$ NOT u_{it-1}

$t=3$: $\tilde y_{it2} = y_{i2} - \dfrac{y_{i2}+y_{i1}}{2} = \dfrac{y_{i2}-y_{i1}}{2} = \dfrac{x_{i2}-x_{i1}}{2}$

$t=2$: $\tilde y_{i1} = y_{i1} - \dfrac{y_{i2}+y_{i1}}{2} = \dfrac{y_{i1}-y_{i2}}{2} = \dfrac{x_{i1}-x_{i2}}{2}$

$t=3$: $\tilde u_{i3} = u_{i3} - \dfrac{u_{i2}+u_{i3}}{2} = \dfrac{u_{i3}-u_{i2}}{2}$

$t=2$: $\tilde u_{i2} = u_{i2} - \dfrac{u_{i2}+u_{i3}}{2} = \dfrac{u_{i2}-u_{i3}}{2}$

Nom : $\left(\dfrac{x_{i2}-x_{i1}}{2}\right)\left(\dfrac{u_{i3}-u_{i2}}{2}\right) + \left(\dfrac{x_{i1}-x_{i2}}{2}\right)\left(\dfrac{u_{i2}-u_{i3}}{2}\right)$

$= -\dfrac{\sigma_u^2}{2}$

Denom: $\left(\dfrac{x_{i2}-x_{i1}}{2}\right)^2 + \left(\dfrac{x_{i1}-x_{i2}}{2}\right)^2 = \dfrac{2(2\sigma_x^2 - 2x_{i2}x_{i1})}{4}$

$= (1-\rho)\sigma_x^2 = \dfrac{1-\rho}{1-\rho^2}\sigma_u^2 = \dfrac{\sigma_u^2}{1+\rho}$

$\Rightarrow \hat\rho - \rho = \dfrac{\text{Nom}}{\text{Denom}} = -\left(\dfrac{1+\rho}{2}\right)$

Second trick: For denomenator do not decompose to y and calculate only based on x

$\rho = 1 \Rightarrow$ This is key ⇒ $\boxed{x_{it} = u_{i1} + u_{i2} + \dots + u_{it}}$, π^0

in this Case write every thing in terms of u

$\tilde y_{i3} = \dfrac{2u_{i3}+u_{i2}}{3}$ ⇒ Denom $= 4/3\,\sigma_u^2$

$\tilde y_{i2} = \dfrac{u_{i2}-u_{i3}}{3}$ Nom $= \left(\dfrac{2u_{i3}+u_{i2}}{3}\right)\left(\dfrac{2u_{i4}-u_{i3}-u_{i2}}{3}\right)$

$y_{i1} = \dfrac{-2u_{i2}-u_{i3}}{3}$ $+ \left(\dfrac{u_{i2}-u_{i3}}{3}\right)\left(\dfrac{2u_{i3}-u_{i2}-u_{i4}}{3}\right) + \left(\dfrac{-u_{i3}-2u_{i2}}{3}\right)\left(\dfrac{2u_{i2}-u_{i4}-u_{i3}}{3}\right) =$

$(-1/3 - 1/3 - 1/3)\sigma_u^2 = -\sigma_u^2$

$\Rightarrow \dfrac{\text{Nom}}{\text{Denom}} = \dfrac{-\sigma_u^2}{4/3\,\sigma_u^2} = -\dfrac{3}{4} = 0.75$

FDIV ① Calculate mean

② Calculate variance of numerator

③ Calculate plim Denom

④ Calculate $Var\left(\dfrac{\text{Nom}}{\text{Denom}}\right)$

trick ① Don't Forget first difference of Error

② Don't Forget (plim Denom)^2 square

$y_{it} = a_i(1-\rho) + \rho y_{i,t-1} + u_{it}$

$Z = y_{iT-2}$

$\hat\beta_{FDIV} = (Z'X)^{-1}(Z'Y) = \dfrac{\sum_N y_{it-2}\Delta u_{it}}{\sum_N y_{it-2}\Delta y_{it-1}}$

Nom: $E(x_{i1}+a_i)(u_{i3}-u_{i2}) = 0$

$Var(\text{nom}) = 2(\sigma_x^2 + \sigma_a^2)\sigma_u^2 = \dfrac{2\sigma_u^4}{1-\rho^2} + 2\sigma_u^2\sigma_a^2$

Denom: $y_{i2} = y_{i2} - y_{i1} = x_{i2} - x_{i1}$

$x_{i1}(x_{i2} - x_{i1}) = (\rho - 1)\sigma_x^2$

$\Rightarrow Var\left(\dfrac{\text{Nom}}{\text{Denom}}\right) = \dfrac{2(\sigma_x^2+\sigma_a^2)\sigma_u^2}{(\rho-1)^2\dfrac{(\sigma_u^4)}{(1-\rho^2)^2}} = \dfrac{2(1+\rho)}{(1-\rho)} + \dfrac{\sigma_a^2}{\sigma_u^2}(1+\rho)^2$

Q4 DGP: $y_{it} = a_i + \beta x_{it} + u_{it}$ β_{LSDV} : least square Dummy variable = WG, FE

Error Serially Correlated $\begin{cases} u_{it} = \rho u_{it-1} + \varepsilon_{it} \\ x_{it} = \rho x_{it-1} + v_{it} \end{cases}$

independent/ regressor serially correlated

$\Rightarrow E(uu') = \begin{bmatrix} [A] & 0 & \dots & 0 \\ 0 & [A] & \dots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \dots & [A] \end{bmatrix} = \Sigma$ $[A] = \dfrac{\sigma_u^2}{1-\rho^2}\begin{bmatrix} 1 & \rho & \dots & \rho^{T-1} \\ \rho & 1 & \dots & \rho^{T-2} \\ \vdots & & \ddots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \dots & 1 \end{bmatrix}$

$\Rightarrow (\hat\beta_{FE} - \beta) \xrightarrow{d} N(0, (\tilde X'\tilde X)^{-1}\tilde X'\Sigma\tilde X(\tilde X'\tilde X)^{-1})$

$\Rightarrow \sqrt{NT}(\hat\beta_{FE} - \beta) \xrightarrow{d} N\left(0, \left(\dfrac{\tilde X'\tilde X}{NT}\right)^{-1}\dfrac{\tilde X'\Sigma\tilde X}{NT}\left(\dfrac{\tilde X'\tilde X}{NT}\right)^{-1}\right)$

$\tilde X = [\tilde X_1, \dots, \tilde x_N]'$ $\tilde X_1 = [\tilde x_{11} \dots \tilde x_{1T}]$ $\tilde x_{it} = x_{it} - \dfrac{\sum^T x_{it}}{T}$

$\hat\sigma_{st} = \dfrac{\sum^N \hat u_{is}\hat u_{it}}{N}$ $\hat\beta_{FE} = \beta + \dfrac{\sum^N\sum^T\left(x_{it} - \frac{\sum^T x_{it}}{T}\right)\left(u_{it} - \frac{\sum^T u_{it}}{T}\right)}{\sum^N\sum^T\left(x_{it} - \frac{\sum^T x_{it}}{T}\right)}$

$\hat a_i = \dfrac{\sum^T}{T}(y_{it} - \hat\beta_{FE}x_{it})$

$\Sigma^{-1} = PP'$ $\beta_{FE,GLS} = (X^{*\prime}X^*)^{-1}X^{*\prime}y^* = (\tilde X'\Sigma^{-1}\tilde X)^{-1}\tilde X'\Sigma^{-1}\tilde y$

$y^* = P'\tilde y$ $= \beta + (\tilde X'\Sigma^{-1}\tilde X)^{-1}\tilde X'\Sigma^{-1}\tilde u$

$X^* = P'\tilde X$

$\Rightarrow \sqrt{NT}(\hat\beta_{FE,GLS} - \beta) \xrightarrow{d} N\left(0, \left(\dfrac{\tilde X'\Sigma^{-1}\tilde X}{NT}\right)^{-1}\right)$

reason of inconsistency of $V(\hat\beta_{WG}) = \dfrac{\frac{\sum^{NT}\hat u_{it}^2}{NT}}{\frac{\sum^{NT}(\tilde x)^2}{NT}}$

assum: $\sum_{k=1}^T \hat u_{is}\hat u_{is+k} = 0$

but $\sum_{k=1}^T \hat u_{is}\hat u_{is+k} = \sum(\tilde u_{is} + (\hat\beta-\beta)\tilde x_{it})(\tilde u_{is+k} + (\hat\beta-\beta)\tilde x_{is+k})$

$= \rho^k\sigma_u^2 + \rho^k(\hat\beta-\beta)\sigma_u^2$

$\Rightarrow \sum_{k=1}^T \hat u_{is}\hat u_{is+k} = \dfrac{\rho - \rho^{T+1}}{1-\rho}(\sigma_u^2 + (\hat\beta-\beta)^2\sigma_x^2) \neq 0$

Test 3 | Econometrics | Post III | 1

Bonus Question

DGP: $y_{it} = \alpha_i + \rho y_{it-1} + u_{it}$

Run { AR(2)) $y_{it} = \alpha_i + \rho_1 y_{it-1} + \rho_2 y_{it-2} + u_{it}$

$\rho_2 = 0$

- withing group estimator : $\sum_T$ transformation

$$\begin{bmatrix} \tilde{y}_{t-1} & \cdots \\ \tilde{y}_{t-2} & \cdots \end{bmatrix}\begin{bmatrix} \tilde{y}_{t-1} & \tilde{y}_{t-2} \\ \vdots & \vdots \end{bmatrix} = \begin{bmatrix} \sum \tilde{y}_{t-1}^2 & \sum \tilde{y}_{t-2}\tilde{y}_{t-1} \\ \sum \tilde{y}_{t-1}\tilde{y}_{t-2} & \sum \tilde{y}_{t-2} \end{bmatrix}^{-1} = \left(\frac{X'X}{N}\right)^{-1}$$

model $$\underbrace{\begin{bmatrix} \tilde{y}_t & \tilde{y}_{t-1} \\ \vdots & \vdots \end{bmatrix}}_{} = \underbrace{\begin{bmatrix} \tilde{y}_{t-1} & \tilde{y}_{t-2} \\ \vdots & \vdots \end{bmatrix}}_{X}\underbrace{\begin{bmatrix} \rho_1 & 1 \\ \rho_2 & 0 \end{bmatrix}}_{\gamma} + [\tilde{u}_{it} \quad 0]$$

$$(X'_{-}\times U) = \begin{bmatrix} \tilde{y}_{t-1} & \cdots \\ y_{t-2} & \cdots \end{bmatrix}\begin{bmatrix} \tilde{u}_{it} & 0 \\ \vdots & \vdots \\ & 0 \end{bmatrix} = \begin{bmatrix} \sum \tilde{y}_{t-1}\tilde{u}_{it} & 0 \\ \sum \tilde{y}_{t-2}\tilde{u}_{it} & 0 \end{bmatrix}$$

$$\Rightarrow \left(\frac{X'X}{N}\right)^{-1} = \begin{bmatrix} \sigma_y^2 & \rho_1\sigma_y^2 \\ \rho_1\sigma_y^2 & \sigma_y^2 \end{bmatrix}^{-1} = \frac{\sigma_y^2}{1-\rho_1^2}\begin{bmatrix} 1 & -\rho_1 \\ -\rho_1 & 1 \end{bmatrix}$$

$$\hat{\gamma} = \left(\frac{X'X}{T}\right)^{-1}\left(\frac{X'u}{T}\right) = \frac{\sigma_y^2}{1-\rho_1^2}\begin{bmatrix} 1 & -\rho_1 \\ -\rho_1 & 1 \end{bmatrix}\begin{bmatrix} \frac{\sum}{T}\tilde{y}_{t-1}\tilde{u}_{it} & 0 \\ \frac{\sum}{T}\tilde{y}_{t-2}\tilde{u}_{it} & 0 \end{bmatrix}$$

$$= \frac{\sigma_y^2}{1-\rho_1^2}\begin{bmatrix} \frac{\sum \tilde{y}_{t-1}\tilde{u}_{it}}{T} - \rho_1\frac{\sum \tilde{y}_{t-2}\tilde{u}_{it}}{T} & 0 \\ -\rho_1\frac{\sum \tilde{y}_{t-1}\tilde{u}_t}{T} + \frac{\sum \tilde{y}_{t-2}\tilde{u}_t}{T} & 0 \end{bmatrix}$$

asymptotic inconsistency $\rightarrow$ stem from bad term

$$= \frac{\sigma_y^{-2}}{1-\rho_1^2}\begin{bmatrix} \frac{\sum}{T}\tilde{y}_{it-1}\frac{\sum u_{it}}{T} - \rho_1\frac{\sum y_{it-2}}{T}\frac{\sum u_{it}}{T} & 0 \\ -\rho_1\frac{\sum y_{it-1}}{T}\frac{\sum u_{it}}{T} + \frac{\sum y_{t-2}}{T}\frac{\sum u_t}{T} & 0 \end{bmatrix}$$

from square of two side of $\tilde{y}_{it} = \rho_1\tilde{y}_{it-1} + \rho_2\tilde{y}_{it-2} + \tilde{u}_{it}$

$$\Rightarrow \sigma_y^2 = \frac{(1-\rho_2)\sigma_u^2}{[(1-\rho_1)^2(1-\rho_2) - \rho_2^2]}$$

---

if $y_t = \rho y_{t-1} + u_t \Rightarrow u_t = y_t - \rho y_{t-1}$

bad term

$$\frac{\sum y_t}{T}\frac{\sum u_t}{T} = \frac{\sum y_t}{T}\frac{\sum(y_t - \rho y_{t-1})}{T} = (1-\rho)(\sum y_t)^2$$

Long run variance

$$= \frac{\sigma_u^2}{1-\rho} \text{ when } (\sum y_t)^2 = \frac{\sigma_u^2}{1-\rho^2}(1 + (T-1)y_t y_{t-1} + (T-2)y_t y_{t-2} + \cdots) = \sigma_u^2/(1-\rho^2)\left(1 + \frac{1}{1-\rho} + O\left(\frac{1}{T}\right)\right) = \frac{1}{(1-\rho)^2}$$

$y_t = \rho_1 y_{t-1} + \rho_2 y_{t-2} + u_t$

$\gamma_1 = E(y_{t-1}y_t) = \frac{\rho_1\sigma_y^2}{1-\rho_2}$ $\quad \gamma_2 = \left[\rho_2 + \frac{\rho_1^2}{1-\rho_2}\right]\sigma_y^2$

$\gamma_3 = \frac{(\rho_1\rho_2 + \rho_1(\rho_1+\rho_2))\sigma_y^2}{1-\rho_2}$

$\gamma_4 = \left[(\rho_1^2 + \rho_2 + \rho_2^2) + \frac{\rho_1^4 + 2\rho_1^2\rho_2}{1-\rho^2}\right]\sigma_y^2$

---

trick You do not need Cover's since true model was AR(1) $y_{it} = \alpha_i + \rho y_{it-1} + u_{it}$ and what you ran was not true or is true but assumption $\boxed{\rho_2 = 0}$

$\Rightarrow \gamma_1 = y_t y_{t-1} = (\rho_1 y_{t-1} + \rho_2 y_{t-2} + u_t)y_{t-1} = \rho_1$

$\gamma_2 = y_t y_{t-2} = (\rho_1 y_{t-1})(y_{t-2}) = \rho_1^2$

$\gamma_3 = y_t y_{t-3} = (\rho_1^2 y_{t-2})y_{t-3}) = \rho_1^3 \ldots \gamma_t = \rho_1^4$

$\Rightarrow$ Form of AR(1) when $\rho_2 = 0$

$\Rightarrow \sum y_t \sum u_t = \sum y_t \sum(y_t - \rho_1 y_{t-1} + \rho_2 y_{t-2}) =$

Bad term $(1-\rho_1)(\sum y_t)^2$ when $\rho_2 = 0$

$\Rightarrow$ Bad term= Nomerator $\frac{\sigma_u^2}{(1-\rho_1)^2}(1-\rho_1) = \frac{\sigma_u^2}{1-\rho_1}$ $\Rightarrow$ total bad $= -\frac{(1+\rho)}{T}$

$$\text{Bias (Bad term)} = \frac{1}{1-\rho^2}\begin{bmatrix} \frac{\sum}{T}y_{t-1}\frac{\sum u_t}{T} - \rho_1\frac{\sum y_{t-1}}{T}\frac{\sum u_t}{T} \\ -\rho_1\frac{\sum}{T}y_{t-1}\frac{\sum u_t}{T} + \frac{\sum y_{t-2}}{T}\frac{\sum u_t}{T} \end{bmatrix} \times \frac{1}{\sigma_y^2}$$

$$= \frac{1}{1+\rho}\begin{bmatrix} \frac{\sum y_t}{T}\frac{\sum u_t}{T} \\ -\frac{\sum y_t}{T}\frac{\sum u_t}{T} \end{bmatrix} \times \frac{1}{\sigma_y^2} = \frac{1}{1+\rho \times \frac{\sigma_u^2}{1-\rho^2}}\begin{bmatrix} \frac{\sigma_u^2}{1-\rho_1} \\ \frac{-\sigma_u^2}{1-\rho_1} \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Variance : matrix form : $(X'X)^{-1}(X'\Omega X)(X'X)^{-1} \times NT$

$$\Rightarrow \Omega = \begin{bmatrix} \Omega_{11} & \cdots & \\ & \ddots & \\ & \cdots & \Omega_{NN} \end{bmatrix} \quad \underbrace{\Omega_{ii}}_{T\times T} = \begin{bmatrix} \hat{\sigma}_{11} & \cdots & \\ & \ddots & \\ & \cdots & \hat{\sigma}_{TT} \end{bmatrix} \quad \hat{\sigma}_{st} = \frac{\sum \hat{u}_{is}\hat{u}_{it}}{N}$$

after correction

mean:

$$\sqrt{NT}\begin{bmatrix} \hat{\rho}_1 - \rho - 1 \\ \hat{\rho}_2 + 1 \end{bmatrix} \xrightarrow{d} (X'X)^{-1}(X'\Omega X)(X'X)^{-1}$$

---

trick

- There is theory = reality but you don't observe
- There is your assumption and empirical that you ran
- The respond to your model is according to reality which is true theory

Previous Final Exam

$$\begin{cases} y_{it} = \mu_{y,i} + y_{it}^{o} & \text{① } y_{it}^{o} \sim iid(0, \sigma_y^2) \\ x_{it} = \mu_{x,i} + x_{it}^{o} & \text{② } x_{it}^{o} \sim iid(0, \sigma_x^2) \end{cases}$$

only observable → Fixed Effect; time varying

$$\mu_{y,i} = a + b\mu_{x,i} + \varepsilon_i \quad ③ \quad \varepsilon_i \sim iid(0, \sigma_\varepsilon^2)$$ Cross section Variation

unexplained of mean

$$y_{it}^{o} = \beta x_{it}^{o} + u_{it} \quad ④ \quad u_{it} \sim iid(0, \sigma_\varepsilon^2)$$

→ Behavioral difference from mean of each Individual (parameter of interest in fixed effect) → time

①,④: $y_{it} = \alpha_i + \gamma x_{it} + e_{it}$, $\alpha_i \to \mu_{y,i}$

① How estimat b? time series AVG

$$\frac{\sum y_{it}}{T} = a + b\left(\frac{\sum x_{it}}{T}\right) + Error$$

FE

$$\tilde{y}_{it} = y_{it} - \frac{\sum y_{it}}{T}$$
$$y_{it} = \mu_{y_i} + y_{it}^{o}$$
$$\frac{\sum y_{it}}{T} = \mu_{y_i} + \frac{\sum y_{it}^{o}}{T}$$
$$\tilde{y}_{it} = \tilde{y}_{it}^{o}$$

① $\operatorname{plim}_{N\to\infty} \hat{\gamma} = ?$

$$\begin{cases} \bar{y}_i = c + \gamma \bar{x}_i + \varepsilon_i \\ \bar{y}_i = \frac{\sum y_{it}}{T} \\ \bar{x}_i = \frac{\sum x_{it}}{T} \end{cases}$$ → within transformation WG

$$\tilde{y}_i = \bar{x}_i - \frac{\sum \bar{x}_i}{N} = \mu_{x,i} + \underbrace{\frac{\sum^T x_{it}^o}{T}}_{O(\frac{1}{\sqrt{T}})} - \underbrace{\frac{\sum^N \mu_{xi}}{N}}_{O(\frac{1}{\sqrt{N}})} - \underbrace{\frac{\sum^T\sum^N x_{it}^o}{NT}}_{O(\frac{1}{\sqrt{NT}})}$$

$= \mu_{xi}$

Good term portion

$$\operatorname{plim}_{N\to\infty} \hat{\gamma} = \gamma + \frac{\frac{1}{N}\sum \mu_{xi}\varepsilon_i}{\frac{1}{N}\sum \mu_{xi}^2} = \gamma$$ since μ_{xi}, ε_i uncorr $Cov(\mu_{xi}, \varepsilon_i) = 0$

② $\sqrt{N}(\hat{\gamma} - \gamma) \longrightarrow ?$

① Nom $= E\left(\frac{1}{N}\sum \mu_{xi}\varepsilon_i\right) = 0$

② $V\left(\frac{1}{N}\sum \mu_{xi}\varepsilon_i\right)^2 = \frac{1}{N}\sigma_{\mu_x}^2\sigma_\varepsilon^2 + \underbrace{Cross}_{=0}$

③ Denom: $\frac{1}{N}\sum \mu_{xi}^2 = \sigma_{\mu_x}^2$

④ $\sqrt{N}(\hat{\gamma} - \gamma) \xrightarrow{d} N\left(0, \frac{\sigma_\varepsilon^2}{\sigma_{\mu_x}^2}\right)$ $\quad \sigma_{\mu_x}^2 = \left(\frac{\sigma_x}{\sqrt{T}}\right)^2 + \frac{\sigma_x^2}{T}$ Sample time

$\Rightarrow \sqrt{N}(\hat{\gamma} - \gamma) \xrightarrow{d} N\left(0, \frac{T\sigma_\varepsilon^2}{\sigma_x^2}\right)$

(*) Large N netralizes $\frac{-\rho}{T}$ bias of $y_t = \rho y_{t-1} + u_t$ when $\frac{1}{NT}\sum$ transformation of POLS, since root of problem is Bad term, but still Problem in T-stat if correlation missed

① (*) Cramer-Wold if $x_n \xrightarrow{d} x$ $\quad c \in R$ ②

$c'x_n \to c'x$

used in $\left[\frac{V(\;)}{\operatorname{plim}\ldots}\right]$ to calculate limiting dist

Q2 $y_{it} = \alpha_i + x_{it}$

DGP $x_{it} = \rho x_{it-1} + u_{it} \sim iid(0,1)$

① $\rho=1$ run $y_{it} = \alpha y_{it-1} + e_{it} \Rightarrow$ OLS Consistent {Fixed T, large N}

$y_{it} = \rho y_{it-1} + \underbrace{(1-\rho)\alpha_i + u_{it}}_{e_{it}}$ makes this term zero

$$\hat{\alpha} = \frac{\frac{\sum}{N} y_{it-1}e_{it}}{\frac{\sum}{N} y_{it-1}^2} = \frac{\frac{\sum}{N}[((1-\rho)\alpha_i + u_{it})y_{it-1}]}{\frac{\sum}{N} y_{it-1}^2} + \alpha$$

as $N \to \infty$ $\quad T \ll N$ $\quad = 0 \Rightarrow$ Consisten

Trick → assumption of $\boxed{\rho = 1}$ makes Fix effect zero

Q3 assume $\boxed{\rho < 1}$ you run $\boxed{y_{it} = \alpha_i + \rho y_{it-1} + u_{it}}$

⇒ inconsistent withing group (Fixed Effect) Estimator

$$E(\hat{\rho}_{wg} - \rho) = -\frac{1+\rho}{T}$$

$$\sqrt{NT}(\hat{\rho}_{wg} - \rho) \begin{cases} \infty & \text{when } \frac{N}{T} \to \infty \\ \text{Correction} & N \simeq T \\ N(0, 1-\rho^2) & N > T \end{cases}$$

$N > T$ $$Var\left(\frac{\sum y_{it-1}u_{it}}{\sum y_{it-1}^2}\right) = \frac{\frac{\sigma_u^4}{1-\rho^2}}{\frac{\sigma_u^4}{(1-\rho^2)^2}} = (1-\rho^2)$$

$N \simeq T$ $\quad \sqrt{NT}(\hat{\rho}_{WG} - \rho + 1 + \rho) \xrightarrow{d} N(0, 1-\rho^2)$

$\Rightarrow \sqrt{NT}(\hat{\rho}_{wg} + 1) \xrightarrow{d} N(0, 1-\rho^2)$

Q2 P2 $\boxed{\rho=1}$

$y_{it} = u_{i1} + u_{i2} + \dots + u_{it}$ (*)

$\Rightarrow \sum y_{it-1}^2 = (T-1)u_{i1}^2 + (T-2)u_{i2}^2 + \dots + u_{it-1}^2 + $ Cross

$= \frac{\sigma_u^2 T^2}{2} + O(T) \to \boxed{\text{Denom}}$

$$\sum t = \frac{T^2}{2}$$
$$\sum t^2 = \frac{T^3}{3}$$
$$\sum\left(t - \frac{\sum t}{T}\right)^2 = \left(\frac{1}{3} - \frac{1}{4}\right)T^2 = \frac{T^2}{12}$$ Good term, Bad term

Nomin (*) $\sum y_{it-1}^2 u_{it}^2 = \sum(T-1)\sigma_u^4$

$\Rightarrow$ Denom $= \frac{T^2}{2}\sigma_u^4$

$$\Rightarrow \frac{\text{Nom}}{(\text{Denom})^2} = \frac{\frac{T^2}{2}\sigma_u^4}{\left(\frac{T^2}{2}\sigma_u^2\right)^2} = \frac{2}{T^2}$$

$$\Rightarrow \sqrt{NT}(\hat{\alpha}_{pols} - \alpha) \xrightarrow{d} N\left(0, \frac{2}{T}\right)$$

Trick even when Consistent pols, auto corr still Causes problem in Variance

Previous final Exam

Q4 model:

$y_{it} = a_i + \beta x_{it} + u_{it}$ Fixed effect on dependent

$x_{it} = \rho x_{it-1} + v_{it}$ $v_{it} \sim iid(0,1)$ } auto corr error & dependent

$u_{it} = \rho u_{it-1} + m_{it}$ $m_{it} \sim iid(0,1)$

$\rho < 1$ assumption

Conventional t-stat: $V(\hat{\beta}_{wg}) = \left(\frac{\sum^N \sum^T \hat{u}_{it}^2}{NT}\right) \Big/ \sum^N \sum^T \tilde{x}_{it}^2$ (I)

conv

problem: $Cov\left(\sum^N \sum^T u_{it} u_{is}\right)$ not in model t-stat

show in other way

true variance $(\tilde{x}'\tilde{x})^{-1}(\tilde{x}'\Omega \tilde{x})(\tilde{x}'\tilde{x})^{-1} NT$ → takes into account covariances

$\Omega = E(uu')_{(NT)(NT)}$ $u = \begin{bmatrix} u_{11} \\ \vdots \\ u_{1T} \\ u_{21} \\ \vdots \\ u_{2T} \\ \vdots \\ u_{NT} \end{bmatrix}$

if no serial / cross sectional correlation: $\Omega = \sigma_u^2 I$

→ (I) which is wrong: $V(\hat{\beta}_{FE}) = (\tilde{x}'\tilde{x})^{-1} \sigma_u I (\tilde{x}'\tilde{x}) \cdot (\tilde{x}'\tilde{x})^{-1} (NT)$

when auto corr: $\frac{\sigma_u^2}{1-\rho^2}\begin{bmatrix} 1 & \rho & \rho^2 & \cdots \\ \rho & 1 & \cdots & \\ & & \ddots & \\ & & & 1 \end{bmatrix}$

$\hat{u}_{it} = \rho \hat{u}_{it} + error$ $T \ll N$

Pooled OLS DGP $\begin{cases} y_{it} = a_i + x_{it} \\ x_{it} = \rho x_{it-1} + u_{it} \sim iid(0,1) \end{cases}$ Fixed effect

Q1 $a_i = a$ → assumption explicit that there is no fix effect

run: $y_{it} = \alpha + \rho y_{it-1} + u_{it}$

- no distortion in good term ⇒ $E(\sum \overbrace{y_{it-1} u_{it}}^{indep}) = 0$
- Bad term $O(\frac{1}{\sqrt{N}})$ so $N \to \infty$ zero!
- $\frac{\frac{\sum y_{it-1}}{NT} \frac{\sum u_{it}}{NT}}{\sum \bar{y}_{it}^2} = \frac{\frac{1}{(1-\rho) NT}}{\frac{1}{1-\rho^2}} = -\frac{1+\rho}{NT} \Rightarrow \underset{N \to \infty}{plim} = 0$

⇒ Consistent $\forall N, T$

Q2 $a_i \neq a$ $a_i \sim iid(a, \sigma_a^2)$

Pols ≠ Consistent

reason good term $\frac{\sum}{NT} y_{it-1} e_{it} = \frac{\sum}{NT} y_{it}(a_i - a + u_{it})$

$\neq 0$

$\underset{N \to \infty}{plim} \sum^{NT} y_{it-1} e_{it} \neq 0 = \sigma_u^2 (1-\rho) \to \boxed{(a_i - a)(1-\rho)}$

---

trick of Dummy var → ① Coding ② interaction ③ Formalization of hypothesis

Kendall bias $-\frac{2\rho}{T} - \frac{1+\rho}{T} = -\left(\frac{1+3\rho}{T}\right)$ { first from good term, second bad term

Part III Bonus

$E(x_{it} u_{js}) = 0 \ \forall i, j, s, t$ → strong Exogeneity

GLS: $\begin{cases} \Omega^{-1} = PP' \\ P'y = P'x\beta + P'u \\ y^* = x^*\beta \\ \beta_{GLS} = (x^{*\prime} x^*)^{-1}(x^{*\prime} y^*) \end{cases}$

---

in Breaks → Dummy

two effects: ① unexplained intercept ② coefficient (rate) differen $\begin{cases} a_2 = 0 \\ b_2 = 0 \end{cases}$

addition part adds nothing

Model: $y_t = a_1 + a_2 d_t + b_1 x_t + b_2 x_t d_t + u_t$

$\Omega_\beta = \underbrace{(x'x)^{-1}(x'\Omega x)(x'x)^{-1}}_{4 \times 4} / T$ → long run variance HAC

t-stat { 5% one sided 1.65; 5% two sided 1.96

Power: rejection rate of null when actually false < 1

Size: variance and t-stat validation

SURE → stack both → reason → Ω - correlation error

$\begin{bmatrix} y_{1t} \\ \vdots \\ y_{2t} \\ \vdots \\ y_{2t} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \vdots & 0 \\ 1 & \vdots \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} + \begin{bmatrix} x_{11} & 0 \\ \vdots & 0 \\ x_{1T} & 0 \\ 0 & x_{21} \\ 0 & x_{2T} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} u_{11} \\ \vdots \\ u_{1T} \\ u_{21} \\ \vdots \\ u_{2T} \end{bmatrix}$

stack into effect → $\Omega_\beta = \underset{2\times 2}{(x'x)^{-1}} \underset{2\times T \ T\times T \ T\times 2}{(x'\Omega x)} \underset{2 \times 2}{(x'x)^{-1}}$

writing size to avoid mistake

$\hat{\Omega}_{12} = \frac{\sum_{t=1}^T \hat{u}_{1t} \hat{u}_{2t}}{T}$ → mean over time in contrast to cross sec mean we now on auto corr

$\sqrt{NT}$: variance converge as $N \to \infty$, $T \to \infty$

multivariate asymptotic

Delta method $g(u)$ Regular

$\sqrt{N}(\hat{\mu} - \mu) \xrightarrow{d} \xi \underset{N \to \infty}{\Rightarrow} \sqrt{N}(g(\hat{\mu}) - g(\mu)) \xrightarrow{d} G'\xi$

$\begin{cases} G(u) = \frac{\partial}{\partial u} g(u)' \\ G = G(\mu) \\ \xi \sim N(0, V) \end{cases} \underset{N \to \infty}{\Rightarrow} \sqrt{N}(g(\hat{\mu}) - g(\mu)) \xrightarrow{d} N(0, G'VG)$ Sandwich of covariance

tailor ⇒ $\boxed{g(x_n) = g(\theta) + g'(\tilde{\theta})(x_n - \theta)}$

tailor Expansion $E\left(\frac{A}{B}\right) = \frac{E(A)}{E(B)}\left(1 - \frac{Cov(A,B)}{E(A)E(B)} + \frac{Var(B)}{E(B^2)}\right)$ $+ O(T^{-2})$

FDW T=3

DGP: $\begin{cases} y_{it} = a_i + x_{it} \\ x_{it} = \rho x_{it-1} + u_{it} \\ u_{it} \sim iid(0,1) \end{cases}$ $E(x_{it} x_{is}) = \rho^{t-s} \sigma_x^2$

Denominator $= -(1-\rho)\sigma_x^2$

Numerator

- write based on x_{it} ⇒ use auto corr

Num $\xrightarrow{d} N\left(0, \frac{1}{N}(2\sigma_u^2)(\sigma_u^2 + \sigma_x^2)\right)$

⇒ $\hat{\rho} - \rho \xrightarrow{d} N\left(0, \frac{(2\sigma_u^2)(\sigma_u^2 + \sigma_x^2)}{N(1-\rho)^2(\sigma_x^2)}\right)$

Test 3 | Question solution ①

model: $y_{it} = a_i + \beta x_{it} + u_{it}$ real: fixed effect

$x_{it} = b_i + \underbrace{x_{it}^o}_{\text{deviation of individual during time}}$

$\begin{cases} a_i \sim iid(0,1) \\ b_i \sim iid(0,1) \\ E(a_i b_i) = \beta \sigma_x^2 \neq 0 \quad \text{(I)} \\ E(a_i x_{it}^o) = 0 \end{cases}$ person's change not depend on first point (or its mean) moment

---

it is inconsistent if $y_{it} = a + \beta x_{it} + e_{it}$ (POLS) → artificial constraint

→ since $\underset{N\to\infty,\ \forall T}{\text{plim}}\ \hat\beta - \beta = \frac{\sum^{N,T}_{NT} a_i b_i}{\sum \tilde{x}_{it}^2} = \frac{\beta\sigma_x^2}{\sigma_x^2} = \beta \neq 0$ (*)

Good term problem $\Rightarrow \boxed{\underset{N\to\infty}{\text{plim}}\ \hat\beta - 2\beta = 0}$

⇒ even true variance will not help in t-stat without removing bias

true variance: $(x'x)(x'\Omega x)(x'x)^{-1} \cdot NT$

$\Omega = \hat u \hat u' \Rightarrow = \frac{\hat\sigma_e^2}{\sum^{NT}(x_{it} - \frac{\sum^{NT} x_{it}}{NT})^2}$ (Var)

where $\begin{cases} \hat\sigma_e^2 = \sum^{NT} \hat e_{it}^2 \\ \hat e_{it} = y_{it} - \hat a - \hat\beta x_{it} \quad ① \end{cases}$

module ③ $e_{it} = (a_i - a) + u_{it}$

$\hat e_{it} = \tilde e_{it} + (\beta - \hat\beta) x_{it}$ ②

(Var of Nominator) $\overset{①}{=} E(\hat e_{it}^2) \overset{②}{=} E(\tilde e_{it} + (\beta - \hat\beta)\tilde x_{it})^2$

$\overset{③}{=} E(\underbrace{\tilde a_i + u_{it}}_{\text{right based on real theory}} + (\beta - \hat\beta)(\hat b_i + \tilde x_{it}^o))^2 = \sigma_a^2 + \sigma_u^2$

based on ④ mean of each indiv → since part correlated with error

(*) $+ \beta^2 \sigma_x^2 + 2\beta E(\tilde a_i \tilde b_i) = \sigma_a^2 + \sigma_u^2 + \beta^2\sigma_x^2 + 2\beta^2\sigma_x^2$ (I)

$\sigma_b^2 + \sigma_{x^o}^2$ ④ (*)

$= \sigma_a^2 + \sigma_u^2 + 3\beta^2 \sigma_x^2$

$\Rightarrow$ (Var) $= \frac{\sigma_a^2 + \sigma_u^2 + 3\beta^2\sigma_x^2}{\sigma_x^2} = \frac{\sigma_a^2 + \sigma_u^2}{\sigma_x^2} + 3\beta^2$ ⑤

$\Rightarrow$ if $t_\beta = \frac{\hat\beta - \beta}{\sqrt{\frac{\hat\sigma_e^2}{\sum^{NT}\tilde x_{it}^2}}}$ it's invalid since ⑥

decompose to: $2 \cdot \frac{\sqrt{NT}(\hat\beta - 2\beta)}{\sqrt{\frac{\sigma_u^2}{\sigma_x^2}}} \xrightarrow{d} N(0,1)$ ← Var

$t_\beta = \frac{\sqrt{NT}(\hat\beta - \beta)}{\sqrt{\frac{\sigma_a^2 + \sigma_u^2}{\sigma_x^2} + 3\beta^2}}$ ⑤ $= \left(\underbrace{\frac{\sqrt{NT}(\hat\beta - 2\beta)}{\sqrt{\sigma_u^2/\sigma_x^2}}}_{\text{from ⑥ has zero mean}} + \frac{\sqrt{NT}\beta}{\sqrt{\sigma_u^2/\sigma_x^2}}\right) \times$

$\frac{\sqrt{\frac{\sigma_u^2}{\sigma_x^2}}}{\sqrt{\frac{\sigma_a^2 + \sigma_u^2}{\sigma_x^2} + 3\beta^2}}$ ~~$\sqrt{NT}$~~

$\propto \sqrt{NT} \Rightarrow \begin{cases} N \to \infty \\ t_\beta \to \infty \end{cases}$

---

trick → when partially correlate break down first and then sum up to get σ_u^2

---

t-test

① Calculate mean $\hat\beta$

② Calculate estimated residuals $\hat u_{it}$

③ Calculate estimated variance $\frac{\sum u_{it} u_{is}}{NT}$ = Std Err

④ do the t-test on $\frac{\hat\beta - \beta}{\text{Std Err}}$ 1.96

---

when cross section: assume Heteroscedasticity

$\hat\sigma_{st} = \frac{\sum \hat u_{is} \hat u_{st}}{N}$ → Covariance off error terms across time (mean of cross section)

assumption of stationary

identifiable only if $N >> T$

→ but heteroscedasticity effect only second digit

not big problem ⇒ use conventional t-stat

- main problem: auto corr (serial correl)

long run var $(\sum u_{it})^2$ → lead smaller variance (not taking into account covariance)

⇒ smaller var ⇒ bigger t-stat ⇒ you reject H_0 while you should have accepted it

---

if Heteroscedasticity: may want to use

① HAC estimator (Heteroscedasticity auto-corr consistent)

② Newey-West estimator (Econometrica 1984)

③ Andrew estimator

④ pre whitening estimator

Conventional t-stat (stata) $t_{\hat\beta} = \frac{\hat\beta}{\sqrt{\sigma_u^2 / \sum x_t^2}}$

which is bigger than real when serial correl

real t-stat

$t_{\hat\beta} = \frac{\hat\beta}{\sqrt{\Omega_\varepsilon^2 / \sum x_t^2}}$

Ω_ε^2 → matrix sums up not only variance but also covariances

$\Omega_i = E(e_i e_i') = E\begin{bmatrix} e_{i1}^2 & e_{i1}e_{i2} & \cdots & e_{i1}e_{iT} \\ e_{i2}e_{i1} & e_{i2}^2 & \cdots & e_{i2}e_{iT} \\ \vdots & \vdots & & \vdots \\ e_{iT}e_{i1} & e_{iT}e_{i2} & \cdots & e_{iT}^2 \end{bmatrix}$

$b_{gls} = (X'\Omega^{-1}X)^{-1} X'\Omega^{-1} Y$ generalized least square

$V(b_{gls}) = (X'\Omega^{-1}X)^{-1} \times NT$

Panel Robust Regression

$(x'x)^{-1} x'\Omega x (x'x)^{-1}$

Estimate based on residuals $T(T-1)$ unknown

Robust since captures both variance and covariance

$\Omega = E(ee'|X)$

$\Omega = \begin{bmatrix} \Omega_{11} & \cdots & \Omega_{N1} \\ \vdots & \ddots & \vdots \\ \Omega_{1N} & \cdots & \Omega_{NN} \end{bmatrix}$ $\Omega_{ij} = [T \times T]$

$\Omega = [(NT) \times (NT)]$ if POLS $\begin{bmatrix} \sigma_a^2 + \sigma_u^2 & \sigma_a^2 & \cdots \\ \vdots & & \\ \sigma_a^2 & \sigma_a^2 & \sigma_a^2 + \sigma_u^2 \end{bmatrix}$

common a_i inside error

Econometrics summary

LSDV & POLS

$x_{it} = a_i + \theta_t + x_{it}^{\circ}$

time and indiv fixed Effect

$\begin{cases} a_i \sim iid(a, \sigma_a^2) \\ \theta_t \sim iid(\theta, \sigma_\theta^2) \\ x_{it}^{\circ} \sim iid(0, \sigma^2) \end{cases}$

POLS $\tilde{x}_{it}^{+} = x_{it} - \frac{\sum^N \sum^T x_{it}}{NT}$

Fixed Eff (indiv) $\tilde{x}_{it} = x_{it} - \frac{\sum^T x_{it}}{T}$

time fixed effect $\tilde{x}_{it}^{++} = x_{it} - \frac{\sum^N x_{it}}{N}$

POLS: Good term: $\frac{\frac{1}{\sqrt{NT}} \sum^{NT} (a_i^{\circ} + \theta_t^{\circ} + x_{it}^{\circ}) u_{it}}{\frac{1}{NT} \sum^{NT} (\tilde{x}_{it}^{+})^2} \xrightarrow{d} N(0, \frac{\sigma_u^2}{\sigma_a^2 + \sigma_\theta^2 + \sigma^2})$

(*) other terms removed for their lower order of probability

$\begin{cases} a_i^{\circ} \sim iid(0, \sigma_a^2) \\ \theta_t^{\circ} \sim iid(0, \sigma_\theta^2) \\ x_{it}^{\circ} \sim iid(0, \sigma^2) \end{cases}$

model $\boxed{\begin{array}{l} y_{it} = a + \beta x_{it} + u_{it} \\ u_{it} \sim iid(0, \sigma_u^2) \end{array}}$

Fixed Effect β_{FE} : $y_{it} = a_i + \beta x_{it} + u_{it} \sim iid(0, \sigma_u^2)$

$\sqrt{NT}(\hat{\beta}_{FE} - \beta) \xrightarrow{d} N(0, \frac{\sigma_u^2}{\sigma_\theta^2 + \sigma^2})$

(*) same comment as above

Fixed time & indiv

$\boxed{\begin{array}{l} y_{it} = a_i + \eta_t + \beta x_{it} + u_{it} \\ u_{it} \sim iid(0, \sigma_u^2) \end{array}}$

$\boxed{\sqrt{NT}(\hat{\beta}_{FE} - \hat{\beta}) \xrightarrow{d} N(0, \frac{\sigma_u^2}{\sigma^2})}$

(*) need to use $\left[\cdot \frac{\sum}{T} - \frac{\sum}{N} + \frac{\sum\sum}{NT} \right]$ transformation

---

take away
- mean on time will remove cross section
- mean on cross section will remove time

- when indiv ~~Fixed~~ Effect we are interested about time

---

trick ① First calculate module $\tilde{u}_{it}, \tilde{x}_{it}, \dots$ manually, & remove element with order lower than $O_p(1)$, mean $O_p(\frac{1}{\sqrt{N}}), O_p(\frac{1}{\sqrt{T}}), O_p(\frac{1}{\sqrt{NT}})$

② for Bad term calculate order (if $\frac{\sum\sum \theta_t}{NT}$, convert to $\frac{\sum \theta_t}{T}$ so $O_p(\frac{1}{\sqrt{T}})$), sometimes $O(\frac{1}{\sqrt{CN}})$ $\begin{cases} N > T \\ N < T \end{cases}$

⚠ important Before breaking to good term and bad term common part should be removed: $\frac{\sum^T}{T}$ within transform

when $x_{it} = a_t + x_{it}^{\circ} + \theta_t$

$= (\theta_t + x_{it}^{\circ} - \frac{\sum^T}{T}(\theta_t + x_{it}^{\circ}))(u_{it} - \frac{\sum^T}{T} u_{it})$

$= (\theta_t + x_{it}^{\circ}) u_{it} - \frac{\sum(\theta_t + x_{it}^{\circ})}{T} \frac{\sum}{T} u_{it}$

math definition | Previous Final Exam

① Ergodicity = Ergodic for mean | Cov. stationary process

math $\left[\begin{array}{l}\frac{\Sigma^T y_t}{T} \xrightarrow{P} E(y_t)=\mu \\ \Sigma_{j=0}^{\infty} |\gamma_j| < \infty\end{array}\right.$ Condition ← $\left\{\begin{array}{l}\text{all lag cov} = \gamma_j \\ \text{mean} = \mu\end{array}\right.$

{dynamic system same behavior averaged over time over all system state} thermodynamic

---

② martingle $\overset{Dif}{=}$ sequence : martingle difference sequence (MDS) : Expectation w.r.t past is zero

Zero mean with respect to past

if (mds) then means not serially correlated

math $\left\{\begin{array}{l}E(y_t)=0 \\ E(y_t|\Omega_{t-1})=0\end{array}\right.$ $\forall t$ | or $E(e_t|\mathcal{F}_{t-1})=0$, $\mathcal{F}_{t-1}=\{y_{t-1}, y_{t-2}, \dots\}$, $y_t = \rho_1 y_{t-1} + \rho_2 y_{t-2} + \dots + e_t$

---

martingle : model of fair game

sequence knowledge of past never helps future winning

- not correlation of info
- If I know past I expect to know nothing about future

math $E(|Y_n|)<\infty$, $E(Y_{n+1}|X_1, \dots, X_n)=Y_n$ } Discrete | e.g random walk: $Z_t = Z_{t-1} + u_t$

Continuous $\left\{\begin{array}{l}E(|Y_t|)<\infty \\ E(Y_t|\{X_\tau, \tau \le s\}) = Y_s \quad \forall s \le t\end{array}\right.$

You will expect what will happen only based on your info set | Green | $E(Z_t|Z_{t-1}, Z_{t-2}, \dots)=Z_{t-1}$

Error Correction model

$I(1)$ $y_t, Z_t \rightarrow$ cointegrated

$y_t = \mu + y_{t-1} + \varepsilon_t$ ← $I(1) \rightarrow$ integrated of order 1 because taking first difference produces stationary process

$$\Delta y_t = X_t\beta + \gamma(\Delta Z_t) + \lambda(y_{t-1} - \theta Z_{t-1}) + \varepsilon_t$$

$\Delta y_t = y_t - y_{t-1}$

$x_t, y_t - \theta_t Z_t \sim I(0)$ stationary

- describes the variation in y_t around its long run trend

ΔZ_t: " Z_t " "

Panel Robust Cov. Matrix clustering $u_{it} \sim$ Exogenous

③ $V(\hat\beta_{FE} - \beta) = (X'X)^{-1} X' \Omega_u X (X'X)^{-1}$ | LSDV | Least square Dummy Variable

PRCM sandwich

deviation from cluster means → cluster of individuals that have fixed effect

---

White heteroscedasticity Consistent Estimator

Est. Asym. $Var[b] = \frac{1}{n}\left(\frac{1}{n}X'X\right)^{-1}\left(\frac{\Sigma_{i=1}^n e_i^2 x_i x_i'}{n}\right)\left(\frac{1}{n}X'X\right)^{-1}$

$= n(X'X)^{-1} S_0 (X'X)^{-1}$

$s^2 = \frac{\Sigma_i \hat{u}_i^2}{n} \longrightarrow$ Correction when $E(uu') = \sigma^2 I_n$ does not hold

---

Cointegration

$y_t, x_t \rightarrow I(1)$ $\exists \beta$: $\varepsilon_t = y_t - \beta x_t$ $\varepsilon_t = I(0)$ means stationary

mean fixed mean

---

non stationary = $\left\{\begin{array}{l}\text{random walk} \\ \text{sum of random} \\ \text{not fix mean} \\ \text{Variance larger \& larger}\end{array}\right.$ e.g GDP, Stock price

$x_t \sim N(0, t)$

Random walk = is kind of martingle sequence

$Z_t = Z_{t-1} + u_t$ $Cov[u_t, u_s]=0$ $t \ne s$

---

martingle sequence $E(Z_t|Z_{t-1}, Z_{t-2}, \dots) = Z_{t-1}$

---

martingle Difference sequence (MDS) $E(Z_t|Z_{t-1}, Z_{t-2}, \dots) = 0$

MDS CLT Z_t $\left\{\begin{array}{l}\text{Stationary} \\ \text{Ergodic} \\ \text{MDS}\end{array}\right.$ $\bar{Z}_t = \frac{\Sigma^T Z_t}{T}$

$E(Z_t Z_t') = \Sigma$ finite positive definite $\Rightarrow \sqrt{T}\,\bar{Z}_T \xrightarrow{d} N[0, \Sigma]$

Unit root e.g $\left\{\begin{array}{l}Z_t = \mu + Z_{t-1} + \varepsilon_t \\ Z_t = \mu + \beta t + \varepsilon_t \\ Z_t = Z_{t-1} + \varepsilon_t\end{array}\right.$ $\left[\begin{array}{l}DGP: (1-L)Z_t = \alpha + v_t \\ \alpha = \mu, \beta, 0 \text{ respectively} \\ v_t : \text{Stationary process}\end{array}\right.$

White noise non auto correlated $u_t \sim iid$

Weak stationary $\{Z_t\}_{t=-\infty}^{\infty}$ $\left\{\begin{array}{l}\forall t \quad E(Z_t) = \mu < \infty \text{ finite constant} \\ E(Z_t - \mu)(Z_{t-j} - \mu) = \gamma_j\end{array}\right.$

= Cover stationary

Cov independent of location ← but distance

Strong stationary $\{Z_t\}_{t=-\infty}^{\infty}$ $\left\{\begin{array}{l}u_t \sim N(0, \sigma_u^2) \text{ iid} \\ \varepsilon_t = \rho\varepsilon_{t-1} + u_t\end{array}\right.$ → strongly stationary

joint prob. dist. $\forall$ set (k)

$[Z_t, Z_{t+1}, \dots, Z_{t+k-1}]$ same regardless of origin t

---

Granger Causality

absent when $f(x_t|x_{t-1}, y_{t-1}) = f(x_t|x_{t-1})$

use lagged variable $\equiv$ lag y_{t-1} does not add info

feedback b/w changes (lag) to explain movement

Granger Causality test VAR : Vector of auto regressive process

$$\begin{bmatrix} x_t \\ y_t \\ Z_t \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \\ Z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix}$$

lagged var cause current variable $\varepsilon_{it} \sim (0, \Sigma)$

$$\begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ 0 & b_{22} & b_{23} \\ 0 & 0 & b_{33} \end{pmatrix} \begin{pmatrix} u_{1t} \\ u_{2t} \\ u_{3t} \end{pmatrix}$$

Cholesky spectral decompos

impulse response analysis → Give shock what happen?

Simple definitions | math + intuition

General to Specific process (GS): *sequential test*

Step 1: run $AR(k_{max})$ check if the last coefficient is significantly different from zero

Step 2: if not, let $k_{max} = k_{max} - 1$, repeat step 1 until the last coeff is significant

problem: overestimation _ accumulation of prob of reject

solution: mitigated, asymptotically when level of test depend on sample size:

Critical value c_T: $\begin{cases} (i)\ c_T \to \infty \\ (ii)\ c_T/\sqrt{T} \to 0 \end{cases}$ $T \to \infty$

$\alpha_T = 1 - \Phi(c_T)$

α_T: significance level; Φ: standard normal cdf

$(i),(ii) \equiv \begin{cases} \alpha_T \to 0 \\ \dfrac{-\log \alpha_T}{\sqrt{T}} \to 0 \end{cases}$

(LLN) law of large number for a cov stationary process

Sample mean → population mean: $\bar{y}_T = \dfrac{\sum y_t}{T}$ $\quad E(\bar{y}_T) = \mu$ $\quad T \to \infty$

Variance of estimate/mean: $E(\bar{y}_T - \mu)^2 = E\left(\dfrac{1}{T^2}\left\{\sum(y_t - \mu)\right\}^2\right)$

$= \dfrac{1}{T^2}\left[\underbrace{T\gamma_0}_{\text{variances}} + \underbrace{2(T-1)\gamma_1}_{\text{first lag}<T} + \cdots + \underbrace{2\gamma_{T-1}}_{t-1 \text{ lag} < T}\right] \leq \dfrac{1}{T}\sum_{t=0}^{T-1} \gamma_t$

$\Rightarrow \lim_{T\to\infty} T \cdot E(\bar{y}_T - \mu)^2 = \sum_{-\infty}^{\infty} \gamma_j$

Variance of estimate/mean: $y_t = u_t$

Autocorrelation: $u_t = \rho u_{t-1} + \varepsilon_t$ $\quad \varepsilon_t \sim iid(0, \sigma^2)$

$E\left(\dfrac{\sum u_t}{T}\right)^2 = \dfrac{1}{T}\dfrac{\sigma^2}{(1-\rho)^2} + O(T^{-2})$ → select combination module

simply write based on ε_t

$E\left(\dfrac{\sum u_t}{T}\right)^2 = \dfrac{1}{T^2}\left(\gamma_0 + 2(T-1)\gamma_1 + \cdots + 2\gamma_{n-1}\right) =$

$\dfrac{1}{T}\left(\dfrac{\sigma^2}{1-\rho^2}\right)\left(1 + \dfrac{2\rho(T-1)}{T} + \dfrac{2\rho^2(T-2)}{T} + \cdots + \dfrac{2(\rho)^{T-1}}{T}\right)$

$= \dfrac{1}{T}\left(\dfrac{\sigma^2}{1-\rho^2}\right)\left(1 + \underbrace{\dfrac{2\rho}{(1-\rho)}}_{\text{Sumarize series}} + \dfrac{X}{T}\right) = \dfrac{\sigma^2}{(1-\rho)^2} + O\left(\dfrac{1}{T^{-2}}\right)$

$\Rightarrow \lim_{T\to\infty} E\left(\dfrac{\sum u_t}{T}\right)^2 = \dfrac{\sigma^2}{(1-\rho)^2}$ → **long run variance**

$E\left(\dfrac{\sum u_t^2}{T}\right) = \dfrac{\sigma_\varepsilon^2}{1-\rho^2}$ **Contemporary variance**

Simple definitions | math ⊕ intuition

① model: time invariant mean (mean ✗ time) not depend

$\boxed{y_t = \mu + \varepsilon_t}$

② AutoCovariance : time lag covariance → shock covariance

③ $\gamma_{jt} = E(y_t - \mu)(y_{t-j} - \mu) = E(\varepsilon_t \varepsilon_{t-j})$

④ Stationary { neither mean μ, autoCov γ } depend on data t

= Cov station

= weakly stationary ; mat: $\begin{cases} E(y_t) = \mu \ \forall t \\ E(y_t - \mu)(y_{t-j} - \mu) = \gamma_j \end{cases}$

$\forall t, j$ = any lag

⑤ ergodic for second moment if: (lag Cov probability γ_j)

$\frac{1}{T-j} \sum_{t=j+1}^{T} (y_t - \mu)(y_{t-j} - \mu) \xrightarrow{p} \gamma_j \quad \forall j$

⑥ white noise : series ε_t if (① uncorrelated ③ ② zero mean, constant Var ④ Combination of Noises)

① $E(\varepsilon_t) = 0$ ② $E(\varepsilon_t^2) = \sigma^2$ ③ $E(\varepsilon_t \varepsilon_s) = 0 \ \forall t, s$

Moving Average ① 1st order (MA1) $\boxed{y_t = \mu + \varepsilon_t + \theta \varepsilon_{t-1}}$

lagged shock to time invarient

$\varepsilon_t \sim iid(0, \sigma_\varepsilon^2)$

variance: $E(y_t - \mu)^2 = (1 + \theta^2)\sigma_\varepsilon^2 = \gamma_0$ lag variance increase variance

$E(y_t - \mu)(y_{t-1} - \mu) = \gamma_1 = \theta \sigma_\varepsilon^2$ lag affects first cover

$E(y_t - \mu)(y_{t-2} - \mu) = \gamma_2 = 0$ lag does not afect cov any more

---

MA2 two lag shock effect $\boxed{y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}}$

Variance ~~First Cov~~ lag : $\gamma_0 = (1 + \theta_1^2 + \theta_2^2)\sigma_\varepsilon^2$

First lag Cov : $\gamma_1 = (\theta_1 + \theta_2 \theta_1)\sigma_\varepsilon^2$

Second lag Cov: $\gamma_2 = \theta_2 \sigma_\varepsilon^2$

next Cov : $\gamma_3 = \gamma_4 = \dots = 0$

MA(∞) $\boxed{y_t = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}}$

stationary if: $\sum_{j=0}^{\infty} \psi_j^2 < \infty$ square sum effect is finite

---

Autoregressive process shocks are auto correlated

time invariant mean but auto corr shocks $\begin{cases} y_t = a + u_t \\ u_t = \rho u_{t-1} + \varepsilon_t \end{cases}$ → $\boxed{y_t = a(1-\rho) + \rho y_{t-1} + \varepsilon_t}$

un correlated shock

finite long run variance $\sum_{j=0}^{\infty} \rho^{2j} = \frac{1}{1-\rho^2} < \infty$

Covariance $\begin{cases} \text{1st lag Cov} : \gamma_0 = \frac{1}{1-\rho^2}\sigma_\varepsilon^2 \\ \text{2nd lag cov} : \gamma_1 = \frac{1}{1-\rho^2}\rho\sigma_\varepsilon^2 \\ t \text{ lag Cov} : \gamma_t = \rho \gamma_{t-1} \end{cases}$

AR(p) = $y_t = a(1-\rho) + \rho_1 y_{t-1} + \dots + \rho_p y_{t-p} + \varepsilon_t$

lag effect $\hookrightarrow \sum_{j=1}^{p} \rho_j$

Cover of t : $\gamma_t = \rho_1 \gamma_{t-1} + \dots + \rho_p \gamma_{t-p}$ Yule walker Equation

augmented form AR(2) ; change to ρ form AR(1)+X

$\boxed{y_t = a(1-\rho) + \rho y_{t-1} - \rho_2 \Delta y_{t-1} + \varepsilon_t}$ $\boxed{\rho = \rho_1 + \rho_2}$

unit root testing : first different + first lag

$\boxed{\Delta y_t = a(1-\rho) + (1-\rho) y_{t-1} - \rho_2 \Delta y_{t-1} + \varepsilon_t}$ white Noise

---

ARMA 2 Auto corr $\begin{cases} y_{it} = \rho y_{it-1} + u_t \\ x_t = \phi x_{it-1} + e_t \end{cases}$ → white noise ①

add two Auto Corr : $\rho(y_{t-1} + x_{t-1}) + \varepsilon_t = z_t = x_{it} + y_{it}$

$\varepsilon_t = (\phi - \rho) x_{t-1} + u_t + e_t$

↳ could be written in e_t ①

$\Rightarrow \boxed{\varepsilon_t = (\phi - \rho) \sum_{j=0}^{\infty} \phi^j e_{t-j-1} + u_t + e_t}$

$\Rightarrow z_t = \rho z_{t-1} + \varepsilon_t \Rightarrow$ ARMA(1,∞) = sum two autocorr

MA = lag sum of shocks, Noises

---

only even event observed $y_s = \rho y_{s-1} + u_s$

$y_s = \rho^2 y_{s-2} + \rho u_{s-1} + u_s$

only observable

$\Rightarrow x_t = y_s$ $t = 1, \dots T$, $s = 2, \dots S$ $\Rightarrow x_t = \rho^2 x_{t-1} + \varepsilon_t$

↳ $= \rho u_{\frac{t-1}{2}} + u_t$

ARMA(1,1) ← lag of shock sum

---

model selection Criteria function : to evaluate

$C_n = -2 \frac{\ln L(k)}{T} + k \frac{\phi(T)}{T}$ Goal : Find true lag length k^*

helps minimize lag length ; ↳ deterministic function

$\Rightarrow \arg\min_k C_T(k)$ $\boxed{\text{plim}_{T\to\infty} \ln L(k^*) > \text{plim}_{T\to\infty} \ln L(k) \ \forall k}$

AIC $\phi(T) = 2$

BIC (shwartz): $\phi(T) = \ln T$

Hann-Quinn : $\phi(T) = 2\ln(\ln T)$

→ deterministic Function for each criteria

$k < k^* \ \forall \phi(T)$

① $\lim_{T\to\infty} \Pr\left[\frac{\ln(k^*)}{T} - \frac{\ln L(k)}{T} \le \frac{1}{2}(k^* - k)\frac{\phi(T)}{T}\right] = 0$ (mean)

② $k > k^*$: $2[\ln L(k) - \ln(k^*)] \xrightarrow{D} \chi^2_{k-k^*}$ likelihood ratio test

AIC $T(C_T(k^*) - C_T(k)) = 2[\ln L(k) - \ln L(k^*)] - 2(k - k^*)$

$\longrightarrow \chi^2_{k-k^*} - 2(k - k^*) \Rightarrow \lim_{T\to\infty} \Pr[C_T(k^*) \ge C_T(k)]$

$= \Pr[\chi^2_{k-k^*} \ge 2(k - k^*)] > 0$

⇒ AIC asymptotically overestimate the lag length

two other criteria $\lim_{T\to\infty} \phi(T) = \infty$

$T(C_T(k^*) - C_T(k)) = 2[\ln L(k) - \ln L(k^*)] - 2(k - k^*)\phi(T)$

$\xrightarrow{D} \chi^2_{k-k^*} - 2(k - k^*)\phi(T)$

$\Rightarrow \lim_{T\to\infty} \Pr[C_T(k^*) \ge C_T(k)] = \lim_{T\to\infty} \Pr[\chi^2_{k-k^*} \ge 2(k - k^*)\phi(T)]$

= 0 = probability that some value > ∞ is zero

⇒ BIC & Hann Quinn's criteria Consistently estimate true lag length

Econometrics

Correl → Covariance↑ → cross↑ ≠ 0 → ~~denom t-ratio~~

⇒ denom-t-ratio < real denom ⇒ t-ratio Calc > t-ratio real
(not Consider corr)

t-ratio real (no reject H_0)

t-ratio Calc (Reject H_0)

8 April Sumer

model Name: Dynamic panel Regression

$y_{it} = \alpha_i + \rho y_{it-1} + u_{it}$ Src of AR $\boxed{\sum^{N}_{N}\left(\sum^{T}_{T}\right)\left(\sum^{T}\right)}$

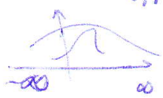
$E(y_{it-1} u_{it}) = 0$

$E(\hat{\rho}_{FE} - \hat{\rho}) = -\frac{1+\rho}{T} + O(T^{-2})$ → Result of Bad term
we know ←

$\cancel{N}$ $\sqrt{NT}(\quad) \rightarrow \frac{\rho}{} -\frac{\sqrt{N}(1+\rho)}{\sqrt{T}}$ * Nickel bias of AutoCorr multiplies # of observations

$N \ll T$: $\frac{N}{T} \to 0$ as $N, T \to \infty$ $\sqrt{NT}(\hat{\rho}_{FE} - \rho) \xrightarrow{d} N(0, 1-\rho^2)$

$N > T$: $\frac{N}{T} \to \infty$: Nickel Bias



$N = T$ → $\sqrt{NT}(\hat{\rho}_{FE} - \rho) \rightarrow N(0, 1-\rho^2) + O(1)$ *

$E(x_{it} u_{is}) \neq 0$ endogenous

Root of model: $y_{it} = \alpha_i + \theta_t + \beta x_{it} + u_{it}$
① $x_{it} = \alpha_i + \rho x_{it-1} + e_{it}$
→ independent AutoCorr

$\frac{1}{1-\rho^2}$, $\frac{}{(1-\rho^2)^2}$

* $\hat{\beta}_{FE} - \beta^* = \frac{-S(1+\rho)}{T}$: $E(\hat{\rho}) = \rho - \frac{1+\rho}{T} + O_p(T^{-2})$

$E(\hat{S}) = S$ ⇒ $\boxed{\frac{T\hat{\rho}+1}{T-1} = \rho}$ $\hat{u}^*_{it} = S\hat{e}_{it} + \text{error}$

② second Case $E(x_{it} u_{is}) \neq 0$
→ is due to AR

* independent Auto Corr: means Endogeneity,
Person changes : $x_{it} = \rho x_{it-1} + e_{it}$

⇒ $y_{it} = \alpha_i + \theta_t + \beta^* x_{it-1} + u^*_{it}$ $\begin{cases} x_{it} = \rho x_{it-1} + e_{it} \\ u^*_{it} = Se_{it} + \varepsilon_{it} \end{cases}$
→ Bad term $\boxed{\beta^* = \rho\beta}$

* Lag creates Nickel bias for $N \simeq T$, and for $N > T$
we need First Difference
same time & cross section ← ; only two Period, large cross section ←

* Bad term: what goes around Comes around

---

steps to get β for $y_{it} = \alpha_i + \theta_t + \beta x_{it} + u_{it}$
when independent x_{it} auto corr $N \simeq T$

① Correct ρ: $\frac{T\hat{\rho}+1}{T-1} = \rho$

② get S: $\hat{u}^*_{it} = S\hat{e}_{it} + \text{error}$

③ $\hat{\beta}^*_{FE} - \beta^* = -\frac{S(1+\rho)}{T}$ → Correct bias

---

two ways to remove fixed effect:

$\begin{cases} ① \ \frac{\Sigma}{T}, \frac{\Sigma}{NT}, \frac{\Sigma}{N} \text{ transformation } (N \simeq T) \rightarrow \text{take into account Correction} \\ ② \text{ First difference } (N \gg T) \end{cases}$

larger t-stat than reality → due to lower variance than real (due to not taking into account Covars (Cross's))

$t_\beta = \frac{\hat{\beta} - \beta}{\sqrt{V/N}} \leftarrow N(0, V)$ → result of 3rd step

$T = 3$ β_{FDIV} → (1) mean
(2) Variance to get limiting dist.

* true variance takes into account all Covariences

$y_{it} = \alpha_i + \theta_t + \rho y_{it-1} + \beta x_{it} + u_{it}$ $\boxed{\text{Wrong}}$

$\boxed{\text{reason}}$ when direct relation $y_{it} \sim x_{it}$ then when y_{it-1} exists x_{it-1} will exist → $\boxed{\text{auto corre}}$

$\boxed{\text{math}}$ true model: ① $y_{it} = \alpha_i + \theta_t + \beta x_{it} + u_{it}$
② $u_{it} = \rho u_{it-1} + \varepsilon_{it}$ $- \{\rho y_{it-1} = \alpha_i\rho + \theta_t\rho + \beta\rho x_{it-1} + \rho u_{it-1}$

$\ldots = + \ldots \quad e_{it}$

$\boxed{\text{Conclusion}}$
when y_{it} auto corr, it means u_{it} is auto Correlated — remove auto Correlation

---

- when you take out Z_{it} from X_{it} as Covar's of u_{it} the correlated part with Error is now out of Error and independent Error will hold:

$y_{it} = \alpha_i + \theta_t + \rho x_{it} + \boxed{Z_{it}\gamma + \epsilon_{it}}$ → u_{it}

$\hat{u}^*_{it} = \tilde{y}_{it} - \hat{\beta}_{FE}\tilde{x}_{it} = u_{it} + (\beta - \hat{\beta}_{FE})\tilde{x}_{it}$

$\hat{e}_{it} = \tilde{x}_{it} + \hat{\rho}_{FE}\tilde{x}_{it} = e_{it} + (\rho - \hat{\rho}_{FE})\tilde{x}_{it}$

⇒ $\hat{u}^*_{it} = S\hat{e}_{it} + \text{error}$

$\text{error} = u_{it} - Se_{it} + (\beta - \hat{\beta}_{FE})\tilde{x}_{it} - S(\rho - \hat{\rho}_{FE})\tilde{x}_{it}$

Econometrics (1)

DID: $\begin{cases} y_{it} = \alpha_i + \beta x_{it} + u_{it} \\ u_{it} = \rho u_{it-1} + \varepsilon_{it} \end{cases}$ $x_{it} \begin{cases} 1 \\ 0 \end{cases}$

Estimate of $\hat{\rho}$: from $\hat{u}_{it} = \rho \hat{u}_{it-1} + \zeta_{it}$

$E(\hat{\beta}) \neq \rho$ since $\rho = \frac{\sum \hat{u}_{it-1} \zeta_{it}}{\sum \hat{u}_{it-1}^2}$

$\zeta_{it} = ?$ $\hat{u}_{it} = y_{it} - \hat{\beta} x_{it} - \hat{\alpha}_i \Rightarrow \hat{u}_{it} = (\alpha_i - \hat{\alpha}_i) + (\beta - \hat{\beta}) x_{it} + u_{it}$

$\tilde{y}_{it} = \beta \tilde{x}_{it} + \tilde{u}_{it}$ =

$\hat{u}_{it} = (\alpha_i - \hat{\alpha}_i) + (\beta - \hat{\beta}) x_{it} + u_{it}$

$\hat{u}_{it} = (\beta - \hat{\beta}) \tilde{x}_{it} + \tilde{u}_{it}$ ⟶ transformation to remove α_i

$-\rho \hat{u}_{it-1} = \rho(\beta - \hat{\beta}) \tilde{x}_{it-1} + \rho \tilde{u}_{it-1}$

$\hat{u}_{it} - \rho \hat{u}_{it-1} = (\beta - \hat{\beta})(\tilde{x}_{it} - \rho \tilde{x}_{it-1}) + \underbrace{\tilde{u}_{it} - \tilde{u}_{it-1}}_{\tilde{\varepsilon}_{it}}$

$\Rightarrow \begin{cases} \zeta_{it} = (\beta - \hat{\beta})(x_{it} - \tilde{x}_{it-1}) + \tilde{\varepsilon}_{it} \\ \hat{u}_{it-1} = (\beta - \hat{\beta}) \tilde{x}_{it-1} + \tilde{u}_{it-1} \end{cases}$

Common Factor ⟵ ⟹ ρ estimate is biased

meaning & words

Assumption (1) y_{it} persistent u_{it} serially Correlated

(2) x_{it} exogeneous

(3) fixed effect

x_{it} ⟶ Exogeneous (A1) Dummy DID treatment

(A2) u_{it} serially Correlate

model $y_{it} = \alpha + \beta x_{it} + u_{it}$

(1) panel Robust Covariance matrix (PRCM) clustering

$V(\hat{\beta}_{FE} - \beta) = (X'X)^{-1} X' \Omega_u X (X'X)^{-1} \times NT$ True Variance

$t_{\hat{\beta}} = \frac{\hat{\beta}_{FE}}{\sqrt{\hat{\sigma}_u^2 / \sum^{NT} \tilde{x}_{it}^2}}$ Auto Correlat

Time series inconsistency

Bad term ⟶ Src of Nickel Bias

(1) Fix time effect wrong ⟶ care about mean (β)

$\frac{1}{T}\sum^N \sum^T x_{it+} \sum^T u_{it}^*$

⚠ x_{it-1} when binary ⟶ non sense

(2) $u_{it} = \rho u_{it-1} + \varepsilon_{it}$ AR(1) Assump

$\underset{NT \times NT}{\Omega_u} = \begin{bmatrix} A & & 0 \\ & A & \\ 0 & & A \end{bmatrix}$ $A = \sigma_u^2 \begin{bmatrix} 1 & \cdots & \rho^{T-1} \\ \rho & 1 \cdots & \rho^{T-2} \\ \rho^{T-1} & \rho^{T-2} & 1 \end{bmatrix}$

(3) no assumpt: $A = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1T} \\ \vdots & & \ddots & \vdots \\ \sigma_{1T} & \cdots & & \sigma_T^2 \end{bmatrix}$ $T \times T$

$\sigma_{st} = \sum_N \hat{u}_{is} \hat{u}_{it}$

cross section sum

more Robust

Src $\frac{1}{T}\sum^T$

if no fixed Effect: $\alpha_i = \alpha$

Then ⟶ pols $O(\frac{1}{\sqrt{NT}})$

so Zero

Cross section inconsistency ⟶ due to Restriction pols

(2)

$y_{it} = \underbrace{\alpha}_{\text{Restriction}} + \beta x_{it} + u_{it}$ $E(x_{it} u_{is}) = 0$ $\forall t, s$ strong exog

but $y_{it} = \alpha + \beta x_{it} + e_{it}$

$e_{it} = (\alpha_i - \alpha) + u_{it}$

α_i, x_{it} ⟶ Generally Common factor of response and Covariate

$E(x_{it} u_{js}) \neq 0$ x_{it} = endogeneous

Sol 1 IV $\hat{\beta}_{IV} = (S'X)^{-1}(S'Y)$

Sol 2 lag - but bad term still problem ⟹ Correction

time lag

$E(\hat{\beta}) = \rho - \frac{1+\rho}{T} + O(T^{-2}) \Rightarrow \rho = \frac{\hat{\rho}T + 1}{T - 1}$ $0 = E(x_{it}^0 u_{it})$

Sol 3 use Covariate ⟶ remove endogeneity $\underbrace{u_{it}}_{\lambda_i' F_t + u_{it}^0}$

FAPR Factor Augmented panel Regression

(2) model: $y_{it} = \alpha_i + \theta_t + \beta x_{it} + u_{it}$

$u_{it} = \rho u_{it-1} + \varepsilon_{it}$

⚠ only addition of lag of y not enough, but lag of x_{it}, θ_t

$y_{it} = \rho y_{it-1} + \alpha_i(1-\rho) + (\theta_t - \rho\theta_{t-1}) + \beta x_{it} - \beta\rho x_{it-1} + \varepsilon_{it}$

removing auto Correlation from error term

⟶ Fix effect ⟹ Nickel bias ⟹ Solution first difference & $\hat{\beta}_{FDIV}$

⚠ $\rho \approx 1$ $(Z'X)^{-1}$ Singular $\Delta y_{it} = \varepsilon_{it}$ ⟶ unCorrelated

= Very large Variance

(*) $E\, x_{it} u_{is} = 0$ i, j, t, s

Strongly/strictly Exogeneous = Consistent

inConsistent when $E\, X_{it} e_{it} \neq 0$

Fixed effect, Bad term (Nickel)

Constraint pols, when fixed effect exists

Past var not Correl Current Shock

(*) weakly Exogeneous: $E(x_{it-1} u_{it}^+) = 0$

in [2x2] off diagonal element = 0, but $E(\varepsilon_{it} e_{it}) \neq 0$

$\begin{bmatrix} x_{it} \\ u_{it} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_{it-1} \\ u_{it-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{it} \\ e_{it} \end{bmatrix}$

FAPR (I) $\lambda' F_t = f(\sum_N y_{it}, \sum x_{it})$ Function of means

altrntv macro Variables?

(II) $y_{it} = \alpha_i + \beta x_{it} + \gamma_i w_t + u_t^+$

altrntv 2 Combine (I) and (II)

test 4 Econometrics

$y_{it} = y_{it-1} + e_{it}$ $\quad e_{it} \sim iid(0, \sigma^2)$

$y_{i1} = e_{i1}$

AR (1)

$y_{it} = \alpha + \rho y_{it-1} + e_{it}$

① $\tilde{y}_{it} = \rho \tilde{y}_{it-1} + \tilde{e}_{it}$ → Bad term $\neq 0$

⇒ Not Consistent

② $y_{it} = \alpha_i + \rho y_{it-1} + e_{it}$

$\hat{\rho}_{FE}$ → Consistent? $\quad$ No not Consistent

$N \to \infty$ $\quad$ Bad term $\neq 0$

$\tilde{y}^{*}_{it} = \rho \tilde{y}^{*}_{it-1} + \tilde{e}^{*}_{it}$

II $y_{it} = \alpha_i + \rho y_{it-1} + \beta x_{it-1} + e_{it}$

$E\, x_{it} e_{js} = 0 \quad \forall\, i, t, j, s$

$\operatorname{plim}_{N\to\infty} (\hat{\beta}_{FE} - \beta) = -\operatorname{plim}_{N\to\infty} (\hat{\rho}_{FE} - \rho) \cdot \operatorname{plim}_{N\to\infty} \frac{\sum^{NT} \tilde{x}_{it-1} \tilde{y}_{it-1}}{\sum \tilde{x}^2_{it-1}}$

$\hat{\rho}_{FE} - \rho \to \underline{\rho_{F_{DEV}}}$ Consistent

$\operatorname{plim}_{N\to\infty} (\hat{\beta}_{FE} - \beta) = (x'x)^{-1} x'y = \beta + \frac{\rho \sum^{NT} \tilde{x}_{it-1} \tilde{y}_{it-1}}{\sum^{NT} \tilde{x}^2_{it-1}}$

$\left(\frac{x'x}{\sigma^2} + B_0\right)^{-1} * \left(x'y + B_0^{-1}\beta_0\right)$ $\quad + \frac{\sum \tilde{x}_{it-1} \tilde{e}_{it}}{\sum^{NT} \tilde{x}^2_{it-1}}$ ⇒

$\left(\frac{x'x}{\sigma^2} + B_0\right)^{-1} \left(x'y + B_0^{-1}\beta\right)$ $\quad 0 <$

$\hat{\beta} - \beta = \operatorname{plim}_{N\to\infty} (\hat{\rho}_{FE} - \rho) \cdot \operatorname{plim} \frac{\sum \tilde{x}_{it-1} \tilde{y}_{it-1}}{\sum \tilde{x}^2_{it-1}}$

$y_{it} = \alpha_i + \beta x_{it} + u_{it}$ $\quad \hat{\beta}_{FE} \xrightarrow{p} \beta \quad E(x_{it} u_{it}) = 0 \quad \forall i, t, j, s$

$x_{it} = \{^1_0$ $\quad$ Diff-in-Diff

$u_{it} = \rho u_{it-1} + e_{it} \quad \sqrt{NT}(\hat{\beta}_{FE} - \beta) \xrightarrow{d} N(0, \Omega_\beta)$

$\Omega_\beta = \left(\frac{xx}{NT}\right)^{-1} \left(\frac{x'\Omega x}{NT}\right) \left(\frac{x'x}{NT}\right)^{-1} / NT = O(1)$

$\hat{\Omega}_\beta = \sum_N^N \hat{u}_{it} \hat{u}_{is}$ $\quad u_{it} \sim AR(1) \quad \Omega_{ts} = |\rho^{t-s}| \sigma_u^2$

$\hat{\rho}$ ⇐ $\hat{u}_{it} = \rho \hat{u}_{it-1} + \text{error}$ $\quad E(\hat{\rho}) = \rho$?

$\hat{u}_{it} = (y_{it} - \hat{\beta} x_{it}) - \underbrace{\sum_T (y_{it} - \hat{\beta} x_{it})}_{\bar{x}_{it}} = \tilde{u}_{it} + (\beta - \hat{\beta}) \tilde{x}_{it}$

$\hat{u}_{it} = (y_{it} - \hat{\beta} x_{it}) - \underbrace{\sum_T (y_{it} - \hat{\beta} x_{it})}_{\bar{x}_{it}} = \tilde{u}_{it} + (\beta - \hat{\beta}) \tilde{x}_{it}$

⇒ $\hat{\rho} u_{it-1} = \rho \hat{u}_{iT-1} + \rho(\beta - \hat{\beta}) x_{it-1}$ ⇒ $\tilde{u}_{it} = \rho \tilde{u}_{it-1}$ Error

$E\left(\sum_T u_{it-1}\right)\left(\sum^T \text{error}\right) \neq 0$ $\quad$ → $\boxed{\tilde{\varepsilon}_{it} + \rho(\beta - \hat{\beta}) \tilde{x}_{it-1}}$

SURE $\quad y_{it} = \alpha_i + \beta_i x_{it} + u_{it}$

$\begin{cases} y_{1t} = \alpha_1 + \beta_1 x_{1t} + u_{1t} \\ y_{2t} = \alpha_2 + \beta_2 x_{2t} + u_{2t} \end{cases}$ $\quad E(u_{1t} u_{2t}) \neq 0$

$\begin{bmatrix} y_{1t} \\ y_{2t} \\ \vdots \\ y_{it} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \\ 0 & 1 \\ \vdots & \vdots \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} + \begin{bmatrix} x_{11} & 0 \\ \vdots & \vdots \\ x_{1T} & 0 \\ 0 & x_{21} \\ \vdots & \vdots \\ 0 & x_{2T} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} u_{11} \\ \vdots \\ u_{1T} \\ u_{21} \\ \vdots \\ u_{2T} \end{bmatrix}$

$Y = \alpha + X\beta + U$ $\quad \Omega_\beta = (x'x)^{-1} (x' \Omega_\beta x)(x'x)^{-1}$

$2T \times T \times 2$

$\sqrt{NT}$ → Converge us $\quad x_{1t} x_{2t}$ → Crosssec Correl

Ω_{ij} ⇒ $\hat{\Omega}_{12} = \sum_T^T \hat{u}_{1t} \hat{u}_{2t}$

$T > N^2$ identifiable $\quad$ GLS stat inference

literature selecting of study based on topic of study

Larger t-stat $\quad N \to \infty$

Econometrics | test3 summary

model (Pols)

$y_{it} = a_i + \beta x_{it} + u_{it}$ $\quad u_{it} \sim iid(0,1)$

$x_{it} = b_i + x^o_{it}$ $\quad x^o_{it} \sim iid(0,1)$

$E(a_i b_i) = \beta\sigma_{x}^2 \neq 0$ $\quad b_i \sim iid(0,1)$

$E(a_i x^o_{it}) = 0$

$y_{it} = \alpha + \beta x_{it} + e_{it}$ $\quad \text{Plim}_{N\to\infty} \hat\beta = \beta + \frac{\sigma_x^2}{\sigma_x^2}\beta = 2\beta$

$\hat e_{it} = \tilde y_{it} - \hat\beta \tilde x_{it} = \tilde e_{it} + (\hat\beta - \beta)\tilde x_{it}$ $\quad e_{it} = a_i - \bar a + u_{it}$

$\sqrt{NT}(\hat\beta - 2\beta) \xrightarrow{d} N(0, \sigma_u^2 \sigma_x^{-2})$ $\quad \tilde e_{it} = \tilde a_i + \tilde u_{it}$

T2 $E(\hat\beta - \beta)^2 = \beta^2$ $\quad E\hat e_{it}^2 = E(\tilde e_{it} + (\hat\beta-\beta)\tilde x_{it})^2 = \underbrace{\sigma_a^2 + \sigma_u^2}_{\sigma_e^2} + 3\beta^2\sigma_x^2$

invalid $t_\beta = \dfrac{\hat\beta - \beta}{\sqrt{\sigma_e^2 / \sum^{NT}(x_{it} - \frac{\sum^{NT} x_{it}}{NT})^2}} = \left(\dfrac{\sqrt{NT}(\hat\beta - 2\beta)}{\sqrt{\sigma_u^2\sigma_x^{-2}}} + \dfrac{\sqrt{NT}\beta}{\sqrt{\sigma_u^2\sigma_x^{-2}}}\right)\sqrt{\dfrac{\sigma_u^2}{\sigma_a^2 + \sigma_u^2 + 3\beta^2\sigma_x^2}}$

$N \to \infty \quad t_\beta \to \infty$

valid t: $Z = \dfrac{\sqrt{NT}(\hat\beta - 2\beta)}{\sqrt{\sigma_u^2\sigma_x^{-2}}} \xrightarrow{d} N(0,1)$

model Dynamic panel Reg with Fixed Effect

$y_{it} = a_i + x_{it}$ $\quad u_{it} \sim iid(0,1)$ $\quad$ key → everything in u_{it}

$x_{it} = \rho x_{it-1} + u_{it}$

$\xrightarrow{\text{run}} y_{it} = a_i + \rho y_{it-1} + u_{it}$

T=3: $\dfrac{E(A)}{E(B)} = \dfrac{-\frac{1}{2}\sigma_u^2}{\sigma_x^2(1-\rho)} = \dfrac{-\frac{1}{2}\sigma_u^2}{\frac{\sigma_u^2}{1-\rho^2}(1-\rho)} = \dfrac{-(1+\rho)}{2}$

T=4, $\rho = 1$: ① first calc $\bar x = \dfrac{x_{i1} + x_{i2} + x_{i3}}{3} = \dfrac{3u_{i1} + 2u_{i2} + u_{i3}}{3}$

Denom: $\frac{4}{3}\sigma_u^2$ $\quad$ Nom: $-\sigma_u^2 \Rightarrow E(\hat\rho - 1) = -\frac{3}{4} = -.75$

T=3 FDIV y_{i3}, y_{i2} only $\Rightarrow y_{i3} - y_{i2} = \rho(y_{i2} - y_{i1}) + (u_{i3} - u_{i2})$

indep of $(u_{i3} - u_{i2})$? y_{i1} : $\hat\beta - \beta = (Z'X)^{-1}Z'u$

$\hat\rho = \rho + \dfrac{\sum_i (u_{i3} - u_{i2})(a_i + x_{i1})}{\sum_i (y_{i2} - y_{i1}) y_{i1}}$ $\quad \to \text{var} = (\sigma_a^2 + \sigma_x^2)\, 2\sigma_u^2$ $\quad \to \text{var} = (1-\rho)^2\sigma_x^2$

$\Rightarrow \sqrt{N}(\hat\rho - \rho) \xrightarrow{d} N\left(0, 2(1+\rho)^2\dfrac{\sigma_a^2}{\sigma_u^2} + \dfrac{2(1+\rho)}{(1-\rho)}\right)$

- when no explicit assumption of independence ⇒ inconsistent

e.g. $y_{it} = a_i + \beta x_{it} + u_{it}$ $\quad x_{it}\ \{^1_0$ DID

$\hat u_{it} = \rho\hat u_{it-1} + \tilde\varepsilon_{it} + \rho(\beta - \hat\beta)x_{it}$ $\quad E\left(\sum_T u_{it-1}\right)\left(\sum^T err\right) \neq 0$

$\Omega_\beta = \left(\dfrac{X'X}{NT}\right)^{-1}\left(\dfrac{X'\Omega X}{NT}\right)\left(\dfrac{X'X}{NT}\right)^{-1} \cancel{X}NT = O(1)$

SURE: seemingly unrelated regression estimator

$$\begin{bmatrix} y_{1t} \\ \vdots \\ y_{2t} \\ \vdots \\ y_{2T} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \vdots & \vdots \\ 1 & 0 \\ 0 & 1 \\ \vdots & \vdots \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} + \begin{bmatrix} x_{11} & 0 \\ \vdots & \vdots \\ x_{1T} & 0 \\ 0 & x_{21} \\ \vdots & \vdots \\ 0 & x_{2T} \end{bmatrix}\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} u_{11} \\ \vdots \\ u_{1T} \\ u_{21} \\ \vdots \\ u_{2T} \end{bmatrix}$$

→ not account for cross sec dependence

- cross sec. dependence → our t-stat cmp with real larger and larger → our variance cmp real smaller & smaller

$E(x_1 + \cdots + x_n)^2 = E(x_1^2 + \cdots + x_n^2) + 2(\text{cross product})$

Breaks $d_t = \begin{cases} 0 & t < \tau \\ 1 & t \geq \tau \end{cases}$ $\quad y_t = a_1 + a_2 d_t + b_1 x_t + b_2 x_t d_t + u_t$

Cannot reject = no structural breaks

Exam note $y_{it}=\alpha_i+\beta x_{it}+\gamma w_i+u_{it}$ Case 1 $\alpha_i=\alpha$

$\frac{Z'Z}{NT}=\begin{bmatrix}1&0&0\\0&\sigma_x^2&0\\0&0&\sigma_w^2\end{bmatrix}$ $\begin{cases}u_{it}\sim iid(0,\sigma^2)\\x_{it}\sim iid(0,\sigma_x^2)\\w_{it}\sim iid(0,\sigma_w^2)\end{cases}$ $\frac{Z'u}{NT}=\begin{bmatrix}\sum^{NT}u_{it}/NT\\\sum^{NT}x_{it}u_{it}/NT\\\sum^{NT}w_iu_{it}/NT\end{bmatrix}$

$E\left(\frac{Z'\Omega Z}{(NT)^2}\right)=\frac{1}{NT}\begin{bmatrix}\sigma^2&0&0\\0&\sigma_x^2\sigma^2&0\\0&0&\sigma_w^2\sigma^2\end{bmatrix}$ $\sqrt{NT}\begin{bmatrix}(\hat\alpha-\alpha)\\(\hat\beta-\beta)\\(\hat\gamma-\gamma)\end{bmatrix}\xrightarrow{d}N(0,Q^{-1}\Omega Q^{-1})$

Case 2 $\alpha_i\neq\alpha$ $\alpha_i=\alpha+\varepsilon_i$ module $\left(\sum^{NT}_{NT}e_{it}\right)^2=\frac{\sigma_\varepsilon^2}{N}+\frac{\sigma^2}{N^2T^2}$

but $E\alpha_i x_{it}=0$ $e_{it}=\varepsilon_i+u_{it}$

$E\frac{Z'uu'Z}{N^2T^2}=\begin{bmatrix}\frac{\sigma_\varepsilon^2}{N}+\frac{\sigma^2}{NT}&0&0\\0&\frac{\sigma_x^2\sigma^2}{NT}&0\\0&0&\frac{\sigma_w^2\sigma_\varepsilon^2}{N}+\frac{\sigma_w^2\sigma^2}{NT}\end{bmatrix}$

$\sqrt{NT}\begin{bmatrix}(\hat\alpha-\alpha)\\(\hat\beta-\beta)\\(\hat\gamma-\gamma)\end{bmatrix}\xrightarrow{d}N(0,Q^{-1}\Omega Q^{-1})$

not consistent

DGP $y_{it}=\alpha_i+\beta x_{it}+u_{it}$ $Euu'=\begin{bmatrix}[A]&0&\cdots&0\\0&[A]&&\vdots\\0&\cdots&0&[A]\end{bmatrix}=\Sigma=PP'$

LSDV & GLS $u_t=\rho u_{it-1}+\varepsilon_t$

$x_{it}=\rho x_{it-1}+v_{it}$ $[A]=\frac{\sigma_u^2}{1-\rho^2}\begin{bmatrix}1&\rho&\cdots&\rho^{T-1}\\\rho&1&&\vdots\\\vdots&&\ddots&\\\rho^{T-1}&\rho^{T-2}&\cdots&1\end{bmatrix}$

$(\hat\beta_{FE}-\beta)\xrightarrow{d}N(0,(\tilde x'\tilde x)^{-1}\tilde x'\Sigma\tilde x(\tilde x'\tilde x)^{-1})$

$\sqrt{NT}(\hat\beta_{FE}-\beta)\xrightarrow{d}N(0,(\frac{\tilde X'\tilde X}{NT})^{-1}\tilde X'\Sigma\tilde X(\tilde X'\tilde X)^{-1})$

$\sqrt{NT}(\hat\beta_{FE/GLS}-\beta)\xrightarrow{d}N(0,(\frac{\tilde X'\Sigma^{-1}\tilde X}{NT})^{-1})$

DGP $\begin{cases}y_{it}=\alpha_i+x_{it}\\x_{it}=\rho x_{it-1}+u_{it}\\u_{it}\sim iid(0,1)\end{cases}$ $\xrightarrow{\rho=1}$ $y_{it}=y_{it-1}+e_{it}$

$\hat\alpha_{pols}\to 1$ $N\to\infty$

$E\left(\sum^{NT}_{NT}y_{it-1}e_{it}\right)^2=\sum^{NT}_{NT}E(e_{i1}+\dots+e_{it-1})^2e_{it}^2=\sigma_e^4\frac{T}{2N}+O(\frac{1}{N})$

$E\left(\sum^{NT}_{NT}y_{it-1}^2\right)=\sigma_e^2\frac{T}{2}+O(1)\Rightarrow\sqrt{NT}(\hat\alpha_{pols}-1)\xrightarrow{d}N(0,2)$

same: $y_{it}=\rho y_{it-1}+\alpha_i(1-\rho)+e_{it}$ main $y_{it-1}=\alpha_i+x_{it-1}$

inconsistent if $\rho>1$ since $E(y_{it-1}u_{it})=(1-\rho)(\mu_u^2+\mu^2)$

same

$\hat\beta_{FE}=\frac{\sum\tilde y_{it}\tilde e_{it}}{\sum\tilde y_{it}^2}=\frac{\sum^{NT}_{NT}u_{it}y_{it}}{\sum^{NT}_{NT}\tilde y_{it}^2}$ AN/AD $-\frac{\sum^{NT}_{NT}y_{it}\sum^{T}_{T}u_{it}}{\sum^{NT}_{NT}\tilde y_{it}^2}$ BN/BD

A B

$BN_1^2=\left(\frac{\sum^{NT}y_i}{N^2T^2}\right)^2=\frac{N\sigma_u^2}{N(1-\rho)^2}$ $BN_1^2=\frac{(\sum u_{it})^2}{T}=\frac{\sigma_u^2}{T}\Rightarrow BN=\frac{\sigma_u^4}{(1-\rho)^2\times T}$

e_{i1}
$\rho e_{i1}+e_{i2}$
$\rho e_{i1}+\cdots$
$\rho^{T-1}+\rho^{T-2}$
$\frac{1-\rho^T}{1-\rho}$

$AN=\frac{\sigma_u^4}{1-\rho^2}\frac{1}{NT}$ $AD=\frac{\sigma_u^4}{(1-\rho^2)^2}=BD$

$\Rightarrow\sqrt{NT}(\hat\beta_{FE}-\beta)\xrightarrow{d}N(0,(1-\rho^2))-(1+\rho)\times\frac{N}{T}$

$y_{it}-y_{it-1}=\rho(y_{it-1}-y_{it-2})+e_{it}-e_{it-1}$

IV: (1) $(y_{it-2}-y_{it-3})$ (2) y_{it-2}

$t=3$ use ☒ and ⊛ $E(DN)=-\sigma_u^2(1-\rho)^2$ for IV(1)

$\sqrt{N}(\hat\beta_{FDIV1}-\rho)\xrightarrow{d}N(0,\frac{4(1-\rho^2)}{(1-\rho)^3})$ $E(N_c)=\frac{4\sigma_u^4}{N(1-\rho^2)}(1-\rho)$

$T=4$

$\hat\rho_{FDIV2}=\frac{\sum^{N,T}(y_{it}-y_{it-1})y_{it-2}}{\sum^{N,T}(y_{it-1}-y_{it-2})y_{it-2}}=$

$t=3,4$ $\frac{\sum^N(y_{i4}-y_{i3})y_{i2}+(y_{i3}-y_{i2})y_{i1}}{\sum^N(y_{i3}-y_{i2})y_{i2}+(y_{i2}-y_{i1})y_{i1}}$

(1) key exam look ahead modifier → [illegible]

(2) if constant (eg $w_i|t$) $\to\frac{\Sigma}{NT}\to\frac{\Sigma}{N}$ (fixing)

(3) know answer - just try to make it (intuition)

(4) for y_{it-1} use main relation $y_{it-1}=\alpha_i+x_{it-1}$

(5) trade off big denom for small denom → eg remove $\frac{1}{NT}$ when $\frac{1}{N}$ exists

(6) ⚠ power two on variance when $\frac{1}{\sum x_i^2}$

(7) match کنی نه fix کتاب پیچیده رو mathematics

$\alpha=\frac{\sum x_i}{N}$

(I) $y_{it}=\alpha_i+\beta x_{it-1}+u_{it}$ $\}$ $u_{it}=\gamma e_{it}+\varepsilon_{it}$ $\}$ $\varepsilon_{it}\sim N(0,\sigma^2)$

(II) $x_{it}=b_i+\rho x_{it-1}+e_{it}$ $E(a_ib_i)=0$ $a_i\sim iid(0,\sigma_a^2)$

write (I) → lead to auto reg form of (II)

2nd level of intelligence → (1) Review, (2) memorize summarize & change problem to look like problem you have seen

$\frac{\Sigma\Sigma}{N\,T}$ $\frac{-1+\rho}{T}$ $+\frac{1}{T^2}\Sigma$ $O(T^{-2})$

Fixed $T=\frac{\sum^{N,T}}{T}$

$x_{i0}=0$

put zero for uncorrelated no serial correl

β_{POLS} fixed T $N\to\infty$ unbias due to $\frac{1}{NT}$

β_{LSDV} fixed T $N\to\infty$ bias due to $\frac{1}{T}\to\frac{-1+\rho}{T}$

$\underset{N\to\infty}{plim}\,\gamma\frac{\frac{1}{N}\sum^{N,T}(x_{it-1}-\frac{\sum^T}{T}x_{it-1})(e_{it}-\frac{\sum^T}{T}e_{it})}{\frac{1}{N}\sum^{N,T}(x_{it-1}-\frac{\sum^T}{T}x_{it-1})^2}$

create x_{it} based on e_{it}

$y_{it}=a_i+\beta x_{it-1}\,u_{it}$ … (I), (II)

plim is for bias calc at infty & nothing to do with variance

$\frac{\Sigma}{N}\left[\frac{\Sigma^T}{T}(0-\frac{\Sigma 0}{T})(0-\frac{\Sigma 0}{T})\right]=\frac{\Sigma}{N}\left[\frac{\Sigma 00}{T}-\frac{\Sigma 0}{T}\frac{\Sigma 0}{T}\right]$

just by fixing N ☺

auto corr $x_{it}=\sum\rho^je_{it-j}$

when $y_{it}=\alpha_i+\beta x_{it-1}+u_{it}$ $\Rightarrow\Sigma 0\Sigma 0\to$ has $E\frac{(a_ib_i)}{1-\rho}$ (e_{it}) pols NT will not conv

$x_{it}=b_i+\rho x_{it-1}+e_{it}$

Panel stambaugh bias (PSB) since constant on $T\to\sum\rho^jb_i$

⚠ Nom Denom → power 2 make sure not miss

$\boxed{\text{Econometrics Summary}}$ ①

model : $y_{it} = \alpha_i + x_{it}\beta + \theta_t + u_{it}$
$x_{it} = a_i + \rho x_{it-1} + \varepsilon_{it}$
$\Rightarrow y_{it} = \alpha_i^* + \theta_t + \beta x_{it-1} + u_{it}^*$

Assump: ① y_{it} Pesist ② u_{it} Serial Corr ③ x_{it} Exogen ④ fixed effect

$y_{it} = \alpha + \beta x_{it} + u_{it}$ $E(\alpha_i u_{is}) = 0$ → Restriction affect

$y_{it} = \alpha + \beta x_{it} + e_{it}$ $e_{it} = (\alpha_i - \alpha) + u_{it}$

$\hat\beta_{pols} \not\xrightarrow{p} \beta$ inconsistent

unlike PSB that fix effect of x and $E(a_i b_i) = 0$ does not allow this

---

Consistency only if $\boxed{E(u_{is} x_{it})\ \forall i, j, t, s = 0}$

$$\begin{bmatrix} x_{it} \\ u_{it} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_{it-1} \\ u_{it-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{it} \\ \epsilon_{it} \end{bmatrix}$$

$a_{12} = a_{21} = 0$ $\boxed{\text{WE) Weakly Exogen}}$

WE & $E(\varepsilon_{it}\epsilon_{it}) = 0$ $\boxed{\text{Strongly Exogen}}$

IV $\{ \Delta x_{it-2},\ x_{it-2}$ · FDIV : First difference IV

$\boxed{\hat\rho_{FE} > \hat\rho_{FIDV}}$ ($\rho \approx 1$) → $\boxed{(Z'X)^{-1} \simeq \frac{1}{0}\ \text{Singular}}$ ← problem (ρ)

(ρ) $\rho \to 1$: IV = invalid

(Sol 1) strong IV $\hat\beta_{IV} = (S'X)^{-1}(S'Y)$ geographic, temperature

(Sol 2) Factor augmented panel Reg (Pesaran 2006)

$y_{it} = \alpha_i + x_{it}\beta + \lambda_i F_t + u_{it}^0$ (corr) $\Rightarrow \frac{\sum y_{it}}{NT} - \frac{\sum \alpha_i}{NT} - \left(\frac{\sum x_{it}}{NT}\right)\beta$

$x_{it} = \phi_i G_t + x_{it}^0$ $- \frac{\sum u_{it}^0}{N} = \bar\lambda' F_t$

$\Rightarrow$ run: $y_{it} = \alpha_i^* + x_{it}\beta + F_{y_i}\left(\frac{\sum y_{it}}{N}\right) + F_{x_i}\left(\frac{\sum x_{it}}{N}\right) + u_{it}^+$

(Sol 3) Just use long time series

(Sol 4) $y_{it} = \alpha_i + \beta x_{it} + \gamma_i w_t + u_t^+$ → Macro Var : GDP

$\boxed{\text{Sol 5}}$ $y_{it} = \alpha_i + \beta x_{it} + \gamma_i w_t + u_t^+ + \Phi \bar y + \Psi_x \bar x$ → identifiable

$\boxed{\text{Sol 6}}$ $y_{it} = \alpha_i + \beta x_{it} + \lambda_i \hat F_t + u_{it}^+$

$\rho \approx 1$ Sign is important

---

① $y_{it} = \alpha_i + \theta_t + \beta x_{it} + u_{it}$ $\boxed{\text{Sol ①}}$ IV - geographic, rain fall

Sol ② $y_{it} = \alpha_i + \theta_t + \beta^* x_{it-1} + u_{it}^*$ $x_{it} = \rho x_{it-1} + e_{it}$ ⓘ

$\{ N > T \to$ Nikel bias $\sqrt{NT}(\hat\beta_{FE}^* - \beta^*) \to d \ne 0$ (Second Bad term) due to

$N \simeq T \to$ Bias Correction $\hat\beta_{FE}^* - \beta^* = \frac{-\delta(1+\rho)}{T}$

$N < T$

$u_{it}^* = \delta e_{it} + \tilde\varepsilon_{it}$ (error) ⓘⓘ $\hat\rho = \rho - \frac{1+\rho}{T} + O_p(T^{-2}) \Rightarrow \frac{\hat\rho T + 1}{T-1} = \rho$

$\hat u_{it}^* = \tilde y_{it} - \hat\beta_{FE}^* \tilde x_{it} = u_{it}^* + \Delta_1$ $\hat e_{it} = \tilde x_{it} - \hat\rho_{FE} \tilde x_{it-1} = e_{it} + \Delta 2$

$E(\hat\delta) = \delta \longrightarrow \boxed{\text{time lagged}}$ $\hat u_{it} \to$ created by $\hat\beta$: $x_{it} \frac{\sum \tilde x_{it} u_{it}}{\sum \tilde x_{it}^2}$

$\boxed{\text{Sol 3}}$ Covariate of literature $y_{it} = \alpha_i + \theta_t + \beta x_{it} + z_{it}\gamma + \varepsilon_{it}$ → u_{it}

Nickel still exists if $E(e_{it}\varepsilon_{it}) \ne 0$

② $u_{it} = \rho u_{it-1} + \varepsilon_{it}$ ⓘⓘⓘ

$y_{it} = \alpha_i(1-\rho) + \theta_t(1-\rho) + \beta x_{it} - \beta\rho x_{it-1} + \rho y_{it-1} + \varepsilon_{it}$ (IV) Brundell & Bond Estimate

$\hookrightarrow y_{it} = \alpha_i + \theta_t + \rho y_{it-1} + \beta x_{it} + u_{it}$ $\{ y_{it-2}\ \ \Delta y_{it-2}$

inconsist $\hookrightarrow \Delta y_{it} = \Delta\theta_t + \rho\Delta y_{it-1} + \beta\Delta x_{it} + \Delta u_{it}$ $x_{it-2}\ \ \Delta x_{it-2}$

③ $y_{it} = \alpha + \beta x_{it} + u_{it}$ → Dummy (put $\{x_{it-1}, y_{it-1}$ Wrong) (PRCM)

$V(\hat\beta_{FE} - \beta) = (X'X)^{-1} X' \Omega_u X (X'X)^{-1}$ panel Robust Covariance Matrix or clustering

→ interest: mean change

$t_{\hat\beta} = \frac{\hat\beta_{FE}}{\sqrt{\hat\sigma_u^2 / \sum^{NT} \tilde x_{it}^2}}$ $u_{it} = \rho u_{it-1} + \varepsilon_{it}$ $A = \sigma_u^2 \begin{bmatrix} \rho^1 & \cdots & \rho^{T-1} \\ & \ddots & \rho^{T-2} \\ \rho^{T-1} & \rho^{T-2} & \end{bmatrix}$

$\Omega_u = \begin{bmatrix} A & 0 \\ 0 & \ddots & A \end{bmatrix}$ $\hat u_{it} = \rho \hat u_{it-1} + \varepsilon_{it}$

x_{it} : Exogen (DID) $\hat u_{it} = y_{it} - \hat\alpha_i - \beta x_{it}$

$A = \begin{bmatrix} \sigma_{i1}^2 & \sigma_{12} & \cdots & \sigma_{1T} \\ \vdots & & \ddots & \\ \sigma_{1T} & & & \sigma_T^2 \end{bmatrix}$ → Conservative among $\hat\sigma_{iT} = \frac{\sum^N \hat u_{it}\hat u_{iT}}{N}$ $N >> T$ identifiable

{ ① PRCM no assumption ② AR(1) " ③ No assumption

Econometrics   Summery

① Dummy matrix

② exogneous mean difference

$\sqrt{N_F}(\hat{\alpha}_1-\alpha_1) \xrightarrow{d} N(0,\sigma_1^2)$ Z i-t $\frac{\sqrt{N}(\hat{\alpha}-\alpha)}{\sqrt{A}} \xrightarrow{d} N(0,1)$

$\sqrt{N_m}(\hat{\alpha}_2-\alpha_2) \xrightarrow{d} N(0,\sigma_2^2)$ $x_i = \underbrace{\alpha_2}_{\alpha} + \underbrace{(\alpha_1-\alpha_2)}_{\beta}\omega_i + \varepsilon_i$

$\sqrt{N}(\hat{\alpha}-\alpha) \xrightarrow{d} N(0, \underbrace{\sigma_1^2+\sigma_2^2}_{A})$ $X = Q\cdot\gamma+\varepsilon$ $\gamma = \begin{bmatrix}\alpha_1 \\ \alpha_1-\alpha_2\end{bmatrix}$ $[N\times 2]\cdot[2\times 1]$

Lindberg levy CLT
① —
② —
③ —

③ interactions $y_i = a + D_i\gamma + a_2 S_i + (D_i S_i)\varphi$

- endogienety - longitude, lattitude, location

- timeseries → ① Core variance & mean (fulct.)
② correl b/w ~~ts~~'s

AR(1) Auto regressive

Ⅰ $y_t = \rho y_{t-1} + u_t$ shock
decay rate ← lag

techic fixing $t=1, t=2 \ldots$

$\rho = 1$: Random walk non stationary

T2 $y_t = \sum_{j=0}^{\infty} \rho^j u_{t-i}$

T3 components $E(\frac{\sum u_t}{T})^2 = \sigma_u^2$ → $\rho \neq 1$

$t\sigma_u^2 : \rho = 1$ ← $E(y_t^2) = \frac{\sigma_u^2}{1-\rho^2}$ from $y_t^2 = (\sum \rho^i u_{t-i})^2$

$E(\frac{A}{B}) = \frac{E(A)}{E(B)}\left(1 - \frac{E(B^2)}{E(A)\cdot Cov(A,B)} + \ldots\right)$

$A = \sum y_{t-1}u_t$ $B = \sum y_{t-1}^2$ $\boxed{E(\hat{\rho}-\rho) = \frac{-2\rho}{T}}$ Bias

Cauchy ↑ $\rho>1$ bubble, $\rho=1$ unit root, $\rho<1$ stationary, Normal

Ⅱ $y_t = a + \rho y_{t-1} + u_t$ $\tilde{y}_t = \rho\tilde{y}_{t-1} + \tilde{u}_t$

→ Dicky Gauss Fuller

$$\frac{E(\sum_T \tilde{y}\sum\tilde{u}_{t-1})}{E(\sum \tilde{y}_{t-1}^2)} = \frac{1+\rho}{T}$$ Nickell Bias $\frac{\sigma_u^2/(1-\rho)}{\sigma_u^2/(1-\rho)^2}$ Ⓐ

$$E\left[\frac{\frac{\sum}{T} y_{t-1}u_t}{\sum y_{t-1}^2}\right] = -\frac{1+3\rho}{T}$$ Sum Ⓐ Ⓑ Kendall Bias

AR1 Corr

$E(y_t y_{t-i}) = \frac{\rho^i \sigma^2}{1-\rho^2} \Rightarrow \boxed{E\left(\frac{\sum y_T}{T}\right)^2 = \frac{\sigma^2}{T(1-\rho)^2}}$ long run variance

$x_i \sim N(0,\sigma^2) \Rightarrow x_i \sim O_p(1)$, $\frac{1}{T}\sum x_t = O_p(\frac{1}{\sqrt{T}})$, $\sum x_t = O_p(\sqrt{T})$

$\sqrt{T}(\frac{1}{T}\sum x_t - 1) = O_p(1) \Rightarrow \frac{\sum}{T} x_t = 1 + O_p(\frac{1}{\sqrt{T}}) = O_p(1)$

Ⓧ $E(y_t) = \alpha + \rho E(y_{t-1}) + \overset{=0}{E(u_t)} \Rightarrow E(y_t) = \frac{\alpha}{1-\rho}$ ~~wrong~~ Ⓐ

Latent model

$y_t = \alpha + \rho y_{t-1} + u_t \Leftarrow \begin{cases} y_t = a + x_t \\ x_t = \rho x_{t-1} + u_t \end{cases} \Rightarrow E(y_t) = a$

① add dummy streigh ⇒ wrong

$\Delta y_t = \alpha + (\rho-1)y_{t-1} + u_t$ level, NOT growth

Forecast ② $y_t = \alpha + \beta x_{t-1} + u_t \Rightarrow$ ① $\hat{y}_{p+1|p} = \hat{\alpha} + \hat{\beta}x_p$

if not know P ⇒ ② $y_t = a + \rho y_{t-1} + u_t$

MSPE: Mean Predition Square Error

$\frac{1}{T}\sum^{T/2} \varepsilon_{T+i}^2$, $\varepsilon^*_{T+1} = y^*_{T+1} - \hat{y}^*_{T+1|T}$ $\hat{y}_{T+1|T} = \hat{\alpha} + \sum^{P} \hat{\rho}_j y_{T-j+1}$

Two Horizon forecast ③ $\hat{y}_{T+2|T} = \hat{a}_2 + \hat{\rho}^2 y_T$

② LAE → leas absolute Error $\Sigma|u_i|$

③ WLSE → weighted least square: median Estimater $\Sigma\omega_i u_i$

④ Quantile Regression → minimize 1/5 Quantile

⑤ non liner Estimater → $y_i = f(x_i) + u_i$

t-ratio: $\frac{\sqrt{n}(\hat{\beta}-\beta)}{\sqrt{\sigma_u^2/\sigma_x^2}} \xrightarrow{d} N(0,1)$ t-stat: $\frac{\hat{\beta}}{\sqrt{\sigma_u^2(x'x)^{-1}}}$ → $\frac{1}{\Sigma x_i^2}$

$y_i = \beta x_i + u_i$

$y_t = \beta x_t + u_t$, $u_t = \rho u_{t-1} + \varepsilon_t$, $s.e.: \sigma_\varepsilon^2$

$\sqrt{T}(\hat{\beta}-\beta) \xrightarrow{d} N(0, \sigma_x^{-2}\Omega_\varepsilon^2\sigma_x^{-2})$ $\frac{\sigma_\varepsilon^2}{(1-\rho)^2}$ long run variance

$E\left[\frac{\sum u_t^2}{T}\right] = \frac{\sigma_\varepsilon^2}{1-\rho^2}$ Contemporary variance

HAC: Heterodescedasticity Auto Corr Estimator ① Neway west, ② andrews, ③ prewhitening "

$\sqrt{n}(\hat{\beta}-\beta) \xrightarrow{d} N(0, (\sum_n x_i^2)^{-1}(\sum_n x_i^2 u_i^2)(\sum_n x_i^2)^{-1})$

$y_i = \alpha + \beta x_i + u_i$ $y = Z\gamma + u$, $\hat{\gamma} = (Z'Z)^{-1}Z'y = \gamma + [Z'Z]^{-1}Z'u$

$\underset{N\times 1}{y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}}$ $\underset{N\times 2}{Z = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_T \end{bmatrix}}$ $\gamma = \begin{bmatrix}\alpha \\ \beta\end{bmatrix}$ $plim(\frac{Z'Z}{n})^{-1} = (plim\frac{Z'Z}{n})^{-1}$

$plim\frac{Z'Z}{n} = \begin{bmatrix} 1 & M_x \\ M_x & \sigma_x^2 + M_x^2 \end{bmatrix} = Q_Z^{-1}$ $\hat{\gamma}-\gamma = (\frac{Z'Z}{n})^{-1}(\frac{Z'u}{n})$

$\sqrt{N}(\hat{\beta}-\beta) \xrightarrow{d} N(\begin{pmatrix}0\\0\end{pmatrix}, Q_Z^{-1}\Sigma Q_Z^{-1})$ $\frac{Z'u}{\sqrt{n}} \xrightarrow{d} N(\begin{pmatrix}0\\0\end{pmatrix}, \begin{bmatrix}\sigma_u^2 & 0 \\ 0 & \sigma_u^2\sigma_x^2\end{bmatrix})$

$M_{x1} = (I - x_1(x_1'x_1)^{-1}x_1')$ u_i Error: ① approximation Error ② Expectation Error ③ misspecification Error

Instrument Variable (IV) ① $E(Zx) \neq 0$ $E(Zu) = 0$

② $\hat{\beta}_{IV} = (Z'x)^{-1}Z'y$ $plim(\hat{\beta}_{IV}-\beta) = 0$

Random Effect $y_{it} = a + bx_{it} + e_{it}$ $e_{it} = a_i - a + u_{it} = \mu_i + u_{it}$

(A1). $E(\mu_i x_{it}) = 0 \; \forall i$ (A2). $E(\mu_i u_{it}) = 0 \; \forall i$ → OLS Consistent but not efficient

Econometrics | Summery | serialize

POLS, Random Effect

FE ① $y_{it} = a_i + \beta x_{it} + u_{it}$

② $x_{it} = \alpha_i + x^{o}_{it}$ → time ver

$x^{o}_{it} \sim iid(0, \sigma^2)$

POLS ③ $y_{it} = a + \beta x_{it} + e_{it}$

$E(a_i - a)(\alpha_i - \alpha) \neq 0$

α_i, a_i → Common factor persistance

Bias

$E(\hat\rho - \rho) = \frac{-2\rho}{T}$ $\lim_{T\to\infty}$

$\text{Plim}_{T\to\infty}(\hat\beta - \rho) = 0 : E(\hat\rho - \rho)$

$E(\hat\rho - \rho)^2 = c$ → Consistent

$\frac{\sqrt{N}(\hat\beta - \beta)}{V(\hat\beta)} \sim t_n$

$\sqrt{n}(\hat\beta - \beta) \xrightarrow{d} N(0, (X'X)^{-1}(X'\Omega X)(X'X)^{-1})$

$\frac{\sqrt{N}(\hat\beta - \beta)}{V(\beta)} \approx^{d} N(0,1) + O_p(\frac{1}{\sqrt{N}})$

Goodterm | nuissance term

④ $y_{it} = a_i + \theta_t + \beta x_{it} + u_{it}$ → $(a_i - a) + u_{iT}$ → ④-1 $y_{it} = a_i + \beta x_i + u_{it}$

u^{**}_{it} $y_{iT} = a + \beta x_i + [u^{*}_i]$

⑤ $y_{it} = a + \beta x_{it} + \rho y_{it-1} + u_{it}$

seasonal: ⑥ $y_{it} = a + \theta_t + \beta x_i + u_{it}$

$[T\times N]$ $[T\times N]$ $[T\times N]$

⑦ $y_{it} = \underbrace{\alpha + \beta x_i}_{a_i} + \gamma z_{it} + \theta_t + u_{it}$

7-1 $y_{it} = a_i + \theta_t + \gamma z_{it} + u_{it} \Rightarrow \hat\gamma_{FE} \xrightarrow{p} \gamma$ $N,T\to\infty$

7-2 $y_{it} - \hat\gamma z_{it} = a + \beta x_{it} + \theta_t + u^{*}_{it}$ $\hat\beta_{OLS} \xrightarrow{p} \beta$ $N,T\to\infty$

$\sum_T (y_{it} - \hat\gamma z_{it}) = \overset{\sigma_t}{a} + \beta x_i + u^{*}_{it}$

⑧ $y_{it} = \alpha_i + \rho y_{it-1} + \beta x_{it} + \theta_t + e_{it}$

8-1 $y_{it} = \alpha_i + \beta x_{it} + e_{it}$

8-2 $y_{it-1} = \underset{\alpha_i(1-\rho)}{\alpha^{*}_i} + \rho y_{it-1} + \beta x_{it} + \underset{-\beta\rho}{\gamma x_{it-1}} + u_{it}$

⑨ FE Reg.

iff $E(x_{it} e_{it}) = 0$

⓪ $y_{it} = \alpha_i + \beta x_{it} + u_{it}$

① $y_{it} = \alpha_i + \beta_i x_{it}\ u_{it}$ Ess [run first] $\sum^{N}_{N} \beta_i = \beta$

POLS Corr → $t_{\hat\beta_{POLS}} = \frac{\hat\beta_{POLS} - \beta}{\sqrt{\hat\sigma_e^2 (X'X)^{-1}}}$

$E(e_{it} e_{is}) = \begin{cases} \sigma_a^2 & \text{if } t \neq s \\ \sigma_a^2 + \sigma_u^2 & \text{if } t = s \end{cases}$

ⓐ $E(\hat\beta - \beta)^2 = E\left[\sum^{NT}\left(x_{it} - \sum_{NT}^{} x_{it}\right)\left(e_{it} - \sum_{NT}^{NT} e_{it}\right)\right]^2$

$E(x_i x_i e_{it} e_{is})_{t\neq s} = \gamma^{t-s} \sigma_\alpha^2 \sigma_e^2$ since $x \sim AR(1): E(x_{it} x_{is}) = \gamma^{t-s} \sigma_\alpha^2$

ⓐ $= \left(\sigma_u^2 + \sigma_a^2 + \left(\frac{\sigma_a^2}{T}\right)\left(\frac{1-\gamma^T}{1-\gamma}\right)\right)$ $N \cdot T \cdot \sigma_\alpha^2$ $\Big] \to \hat\sigma_e^2 (X'X)^{-1}$

① ②

Correct

Remedy → run FE

$t_{\hat\beta} = \frac{\hat\beta_{OLS} - \beta}{\sqrt{\frac{(\sum^{NT} \hat e_{it}^2 + \sum^{NT} e_{it} e_{is})}{\sum^{NT} \hat x_{it}^2}}}$

typical

$\hat\beta_{FE} \neq \hat\beta_{POLS}$

$y = \beta X + e$ $E(\hat\beta - \beta)(\hat\beta - \beta)' = E((X'X)^{-1} X' \Omega X (X'X)^{-1})$

$E(ee'|x)$ → Condition = replicate matrix N times

POLS

$\Omega = \begin{bmatrix} \Omega_{11} & & \\ & T\times T & \\ & & T\times T \end{bmatrix}_{(NT)\times(NT)}$ $\Omega_{11} = \begin{bmatrix} \sigma_a^2 + \sigma_u^2 & \sigma_a^2 & \cdots \\ \sigma_a^2 & \ddots & \\ \vdots & & \sigma_a^2 + \sigma_u^2 \end{bmatrix}$

indep zero → Non zero diagonal

iff diff

$T\times(T-1)$ unknown

Rob (Robust panel Regression)

$E[(\hat\beta - \beta)|x] = 0$ $E[(\hat\beta - \beta)(\hat\beta - \beta)'|x] = (X'X)^{-1} X' \Omega X (X'X)^{-1}$

Decom value Eign | $PP' = \Omega^{-1}$

$\Omega_\beta = (X'X)^{-1} X' \Omega X (X'X)^{-1} \cdot n$ $\sqrt{N}(\hat\beta_{POLS} - \beta) \xrightarrow{d} N(0, \Omega_\beta)$ sandwitch

GLS $E[(\hat\beta_{GLS} - \beta)(\hat\beta_{GLS} - \beta)|X] = (X'\Omega^{-1}X)^{-1} = \Omega_{\hat\beta_{GLS}} \cdot n^{-1}$

$P'y = P'X\beta + P'u \Rightarrow \sqrt{N}(\hat\beta_{GLS} - \beta) \xrightarrow{d} N(0, \Omega_{\hat\beta_{GLS}})$ $\frac{E(A)}{B}$

$V(\hat\beta_{GLS}) < V(\hat\beta_{POLS}) \Rightarrow \sigma_u^2 (X'X)^{-1} = V(\hat\beta_{GLS})$ since $E(\frac{A}{B}|B)$

iid assump not hold ⇒ t-test would be large → Efficient

$y_{(NT\times 1)} = \begin{bmatrix} y_{11} \\ y_{1T} \\ y_{21} \\ y_{N1} \\ y_{NT} \end{bmatrix}$ $X_{(NT)\times 2} = \begin{bmatrix} 1 & x_{11} \\ 1 & x_{12} \\ \vdots & \vdots \\ 1 & x_{NT} \end{bmatrix}$ $\gamma = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$ $y = X\gamma + e$

$E(e_{it} e_{js}) = 0$ Cross sec. independ

proc: estimate α_i fixed effect estimator

Assump
① $u_{it} \sim iid$ ⊗ serial corre
② u_{it} & x_{it} ⊗ Correl

① $y_{it} = \alpha_i + \beta x_{it} + u_{it} \to \hat\alpha_{iFE}, \hat\beta_{FE}$

② $\hat\alpha_i \to \hat\sigma_\alpha = \sum^{N}_{N}(\hat\alpha_i^2 - \sum_N \alpha_i)^2$ ⇒ ③ Cholesc decomposition

$\hat\sigma_u = \sum^{NT}_{NT} \hat u_{it}^2$ $\hat\Omega^{-1} \xrightarrow{\text{estimate}} \hat P' \to \hat\beta_{GLS}$

if serial corr $\hat\Omega^{-1} \to$ ⊗ Consistent $\hat\Omega \not\xrightarrow{p} \Omega$ $\hat\beta_{GLS} \not\xrightarrow{p} \beta$

if serial corr | FE → GLS$_{FE}$

$\begin{cases} y_{it} = \alpha_i + \beta x_{it} + u_{it} \\ u_{it} = \rho u_{it-1} + \varepsilon_{it} \end{cases}$ $\sigma_u^2 = \frac{\sigma_\varepsilon^2}{(1-\rho^2)}$

$\rho_i = \rho$ homogen across i mean

$E(uu')_{T\times T} = \begin{bmatrix} \sigma_u^2 & \rho\sigma_u^2 & \rho^2\sigma_u^2 \cdots \rho^{T-1}\sigma_u^2 \\ & \ddots & \vdots \\ \rho^{T-1}\sigma_u^2 & \cdots & \sigma_u^2 \end{bmatrix} = \Omega_0 = \Omega_{ii} \Rightarrow \Omega = \begin{bmatrix} \Omega_0 & & 0 \\ & \ddots & \\ 0 & & \Omega_0 \end{bmatrix}$

① Run $y_{it} = \alpha_i + \beta x_{it} + u_{it} \Rightarrow \hat u_{it} = y_{it} - \hat\alpha_i - \hat\beta x_{it}$

$\hat u_{it} = \rho \hat u_{it-1} + \varepsilon^{*}_{it}$

② get $\hat\rho$ → ③ $\hat\Omega$ from ② → ④ GLS

Study

$\begin{cases} V(\hat\beta_{FE\,GLS}) = (X'\hat\Omega^{-1}X)^{-1} \\ \hat P \hat P' = \hat\Omega^{-1} \\ \hat P' y = \hat P' X\gamma + \hat P' u \end{cases}$

$V(\hat\beta_{FE}) = (X'\hat\Omega X)^{-1} X' \hat\Omega X (X'\hat\Omega X)^{-1} \cdot NT$ ← Panel Robust (Covariance matrix)

$\frac{\sqrt{N\cdot T}(\hat\beta_{FE} - \beta)}{\sqrt{V(\hat\beta_{FE})}} = t_{\hat\beta} \xrightarrow{d} N(0,1)$ ⚠ Assumption: AR(1)

as long as x_{it} indep. from u_{it} estimator is Consistent

if not know? AR(1), AR(3), ...

$\Omega_{T\times T} = \begin{bmatrix} \sigma_1^2 & \cdots & \sigma_{1T} \\ & \ddots & \vdots \\ " & & \sigma_T^2 \end{bmatrix}$ unknown $\frac{T(T-1)}{2} + T = O(T^{-1})$

$\hat\sigma_{ts} = \sum_n \hat u_{it} \hat u_{is}$ $\hat\Omega_\beta = (X'\hat\Omega X)^{-1} X' \hat\Omega X (X'\hat\Omega X)^{-1} (NT)$

identity iff $n \geq T$ ← Panel Robust Cover matrix Estimator

Dynamic panel Reg $y_{it} = \alpha_i + \rho y_{it-1} + u_{it} \Rightarrow y_{it} - \frac{\sum y_{it}}{T} = \tilde y_{it}$

$\hat\rho_{FE} - \rho = \frac{[\sum^{NT} \tilde y_{it-1} \tilde u_{it}]_A}{[\sum^{NT} y^2_{it-1}]_B}$ $B = \frac{\sigma_u^2}{1-\rho^2}$

$\text{Plim}_{N\to\infty} A = \text{Plim}_{N\to\infty} \frac{\sum^{N,T}}{N}\left(y_{it-1} - \sum_T y_{it-1}\right)\left(u_{it} - \frac{\sum u_{it}}{T}\right)$ Nickell B

$= \text{plim}_{N\to\infty}\Big[\underbrace{\sum^{N,T}_{N} y_{it-1} u_{it}}_{} - \underbrace{\sum^{N}_{N}\left(\sum_T y_{it-1}\right)\left(\sum^{T} u_{it}\right)}_{C}\Big]$, $E\left(\frac{D}{B}\right) = \frac{-1+\rho}{T} + O(\cdot)$

$E(D_\beta)$ D $\frac{1}{B}$

$\sqrt{NT}(\hat\rho_{FE} - \rho) = \sqrt{NT}\frac{C}{B} + O_p\left(\frac{\sqrt{NT}}{T}\right) \xrightarrow{d} N(0, 1-\rho^2)$

$N,T\to\infty$
① $\frac{N}{T} \to 0$: $\sqrt{NT}(\hat\rho_{FE} - \rho) \xrightarrow{d} N(0, 1-\rho^2)$
② $\frac{N}{T} \to \infty$: $\sqrt{NT}(\hat\rho_{FE} - \rho) \to$ No lim dist → $\frac{N}{T}$
③ $N = T$: $N(0, 1-\rho^2) + O_p(1) = -(1+\rho)c + N(0, 1-\rho^2)$

Solution: IV ① y_{it-2} ② level $y_{it-2} - y_{it-3}$

$y_{it} - y_{it-1} = \rho(y_{it-1} - y_{it-2}) + (u_{it} - u_{it-1})$ $\hat\beta_{IV} = (Z'X)^{-1} Z' Y$

$y_t = bx_t + u_t \quad , u_t = \rho u_{t-1} + \varepsilon_t$

$E(x_t x_s u_t u_s) = \rho^{t-s} E(x_t^2)\, \rho^{t-s} E(u_t^2) = \rho^{2(t-s)} \sigma_x^2 \sigma_u^2$

GLS AR(1) $u_t = \rho u_{t-1} + e_t$

$$E\, uu' = \begin{bmatrix} \frac{1}{1-\rho^2} & \frac{\rho}{1-\rho^2} & \cdots & \frac{\rho^{T-1}}{1-\rho^2} \\ \frac{\rho}{1-\rho^2} & \frac{1}{1-\rho^2} & \cdots & \frac{\rho}{1-\rho^2} \\ \frac{\rho^{T-1}}{1-\rho^2} & \cdots & \frac{\rho}{1-\rho^2} & \frac{1}{1-\rho^2} \end{bmatrix} = \frac{\sigma_e^2}{1-\rho^2} \begin{bmatrix} 1 & \rho & \cdots & \rho^{T-1} \\ \rho & \ddots & & \vdots \\ \vdots & & \ddots & \rho \\ \rho^{T-1} & \cdots & \rho & 1 \end{bmatrix}$$

$\sqrt{n}(\hat b_{gls} - b) \xrightarrow{d} N\left(0, \left(\frac{X'\Omega^{-1}X}{n}\right)^{-1}\right) \quad \hat b_{fgls} = (X'\hat\Omega^{-1}X)^{-1} X'\hat\Omega^{-1} y$

Heteroskedasticity $E(u_i^2) = \sigma_i^2 \neq \sigma_j^2 = E(u_j^2) \quad E(u_i u_j) = 0$

$$\underset{T\times T}{E(uu') = \Omega} \begin{bmatrix} \sigma_1^2 & & & 0 \\ & \sigma_2^2 & & \\ & & \ddots & \\ 0 & & 0 & \sigma_n^2 \end{bmatrix}$$

$X'E(uu')X = \sum^n \sigma_i^2 x_i' x_i$

$V_b = (X'X)^{-1}(X'\Omega X)^{-1}(X'X)^{-1}$

$= (X'X)^{-1}\left(\sum^n \sigma_i^2 x_i' x_i\right)(X'X)^{-1}$

$\hat\sigma_i^2 = \hat u_i^2 \rightarrow$ white heteroskedasticity

– test of Exogeneity → Hausman test $\boxed{IV}$

$plim(\hat\alpha_{IV} - \alpha) = plim\left(\frac{Z'X}{n}\right)^{-1} plim\left(\frac{Z'u}{n}\right) = Q_{zx} \cdot 0 = 0$

$E[(\hat\alpha_{IV} - \alpha)(\hat\alpha_{IV} - \alpha)'] = E[(Z'X)^{-1} Z'u\, u'Z (Z'u)^{-1}]$

$= (Z'X)^{-1} Z' \Omega_u Z (Z'X)^{-1}$

– Random Effects

assumptions

(A1) $E(\mu_i x_{it}) = 0 \;\forall i$

(A2) $E(\mu_i u_{it}) = 0 \;\forall i$

→ indiv char. ⊥̸ regressor {inconsistency}

(I) $y_{it} = a + bx_{it} + e_{it}$

(II) $e_{it} = a_i - a + u_{it} = \mu_i + u_{it}$

$E(e_{it} e_{it}) = \sigma_\mu^2 + \sigma_u^2$ (IV)

① $\hat e_{it}$, $\hat b_{pols}$, $\hat a_{pols}$ from (I)

② $\hat\sigma_e^2$ using ① & (I)

$\hat\mu_i = \frac{\sum}{T} \hat e_{it} \rightarrow \hat u_{it} = \hat e_{it} - \hat\mu_{it}$ (III)

③ from (III) → $\hat\sigma_\mu^2$, $\hat\sigma_u^2$ from (IV) ⇒ Construct Ω

$\rightarrow \hat b_{fgls} \left(\sum x_i' \hat\Omega^{-1} x_i\right)^{-1} \left(\sum x_i' \hat\Omega^{-1} y_i\right)$

$[x_{1i}, \ldots, x_{Ti}]' \quad [y_{1i}, \ldots, y_{Ti}]$

Fixed Effects

Approaches:

① use M_x

② use diff steps
- $-\frac{\sum x_{it}}{N}$ remove fixed indiv effect
- $-\frac{\sum x_{it}}{T}$ remove indiv time effect

(LSDV)

Fixed effect estimator = 'Least square Dummies Variable'

Consistency Condition $\quad \underset{N\to\infty}{plim}\, \hat r_1 = b \rightarrow$ from DGP

show residual is not zero $O(1)$

Dynamic panel Reg

$y_{it}^{+} = \rho y_{it-1}^{+} + b x_{it}^{+} + \gamma x_{it}^{+} + \varepsilon_{it}^{+}$

$\left\{ \begin{array}{l} y_{it} = a_i + d_t + bx_{it} + u_{it} \\ u_{it} = \rho u_{it-1} + \varepsilon_{it} \end{array} \right.$

LSDV → ρ & γ → inconsistent
↳ β → Consistent

$\hat b - b = \frac{1}{\sqrt{S}} \frac{A_S}{B_S} \Rightarrow \sqrt{S}(\hat b - b) \xrightarrow{d} N(0, Q_B^{-1} \Omega_A^2 Q_B^{-1})$ – Cramer theory

$\frac{\sqrt{S}(\hat b - b)}{\sqrt{Q_B^{-1}\Omega_A^2 Q_B^{-1}}} \xrightarrow{d} N(0,1)$

Nickel Bias $E\left[\sum_{t=2}^{T} \tilde y_{t-1} \tilde u_t\right] = 0 - \frac{1}{T}\left(\sum_{t=2}^{T} y_{t-1}\right)\left(\sum_{t=2}^{T} u_t\right) \neq 0$

$\rightarrow E(y_1 u_1) = \sigma_u^2$

$\rho\sigma_u^2 = E(y_2 u_1) = E(B$

$E\left[\frac{A_T}{B_T}\right] = \frac{EA_T}{EB_T}[1 - E(C_T)]$

$E(C_T) = \frac{Cov(A_T B_T)}{E(A_T)E(B_T)} + \frac{Var(B_T)}{[E(B_T)]^2}$

LSDV → inconsistent

$y_{it} = a_i + \rho y_{it-1} + u_{it}$

$\sqrt{NT}(\hat\rho_{LSDV} - \rho) = -\frac{(1+\rho)}{T}\sqrt{NT} + \frac{\frac{1}{\sqrt{NT}}\sum^{N,T} y_{it-1} u_{it}}{\frac{1}{NT}\sum^{N,T} \tilde y_{it-1}}$

in $\frac{\sum \tilde x_{it} \tilde u_{it}}{\sum \tilde x_{it}}$

nuisance term ↓ affect tail

→ (I) $-\frac{\sum}{N}$ removes common term from Both nominator & Denom → $O\left(\frac{1}{\sqrt{NT}}\right)$

→ $-\frac{\sum}{NT}$ → if fix effect exists shrink reduces to $\frac{\sum}{T}$ & $\frac{\sum}{N}$ by order ② $O\left(\frac{1}{\sqrt{N}}\right)$

→ $-\frac{\sum}{T}$ same as (I) but Careful Nickel bias

NLS

tailor series: $f(x_0) = f(x_0) + \frac{\partial f}{\partial x}\Big|_{x=x_0}(x - x_0) + \frac{\partial^2 f}{2\cdot\partial x^2}\Big|_{x=x_0}(x - x_0)^2 \cdots$

model $y_i = \alpha + \beta x_i + \gamma z_i + u_i$

want to estimate $\delta = \frac{\hat\beta}{\hat\gamma} = ?$ second term

$\left\{ \begin{array}{l} \hat\delta = \frac{\hat\beta}{\hat\gamma} = \frac{\beta}{\gamma} + \frac{1}{\gamma}(\hat\beta - \beta) - \frac{\beta}{\gamma^2}(\hat\gamma - \gamma) \\ + 1/4\, \beta/\gamma^3 (\hat\gamma - \gamma)^2 \end{array} \right.$

P 34 | Econ lec note ~~- all own~~ Solution & Summary.

① $y = Xb + u$ $<\infty$

② X : non stoc. $n \times k$

③ $X'X$ non singular $\forall\, n \geq K$

④ $E(u) = 0$

⑤ $u \sim N(0, \sigma^2 I)$

X : Stoch

① Exogen of indep var $E(U|X) = 0$

② Homosced. & no auto cor

③ $\sim$ Normal $(0, \sigma_x^2)$

(1-1) (1-2) (1-3) $\Rightarrow \exists \hat{\beta}\ \forall\, n \geq k$

Ⅰ unique

Ⅱ unbiasedness $E(\hat{\beta}) = \beta$

Ⅲ Normality (1-1)–(1-5)

$\hat{\beta} \sim N(\beta, \sigma^2 (X'X)^{-1})$

Ⅳ Gauss Markov theorem

OLS ① min var

② liner unbiased

intercept : $\lim_{n\to\infty} \frac{\sigma_u^2}{\sum \tilde{x}_i^2} = 0$ if

$\sum \sqrt{T} = \int \sqrt{T}\, dt$

Quadratic loss: $\sum_{i=1}^{n} u_i^2$

$\frac{1}{n} \sum \tilde{x}_i^2 = c$: $E(\hat{\beta} - \beta)^2 = \frac{\sigma_u^2}{\sum \tilde{x}_i^2}$ $\Longrightarrow \frac{\hat{\beta} - \beta}{\sqrt{\frac{\sigma_u^2}{\sum \tilde{x}_i^2}}} \xrightarrow{d} N(0,1)$

$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} N\left(0, \frac{\sigma_u^2}{\frac{1}{n}\sum \tilde{x}_i^2}\right)$

$E\, \tilde{u}_i u_{i+1} = -\frac{1}{n} \sigma_u^2$

$\tilde{x} = \frac{1}{n} \sum_{i=1}^{n} \tilde{x}_i = \bar{x}$

$E(\tilde{u}_i^2) = \frac{n-1}{n} \sigma_u^2$

x_i : stoch : $\text{plim}_{n\to\infty} (\hat{\beta}_n - \beta) = \frac{\text{plim}_{n\to\infty} \frac{1}{n} \sum \tilde{x}_i \tilde{u}_i}{\text{plim}_{n\to\infty} \frac{1}{n} \sum \tilde{x}_i^2} = 0$

$E(\hat{\beta} - \beta | x)$

$E(\hat{\beta} - \beta | x)$

$\sqrt{n}(\hat{\beta} - \beta | x) \xrightarrow{d} N(0, \sigma_u^2 Q_x^{-1})$

$Q_x = \lim_{n\to\infty} \left(\frac{X'X}{n}\right)^{-1}$

$E(\tilde{x}_i^2) = \frac{n-1}{n} \sigma_x^2$

$E\, \tilde{x}_i \tilde{x}_{i+1} = -\frac{1}{n} \sigma_x^2$

$\text{plim}_{n\to\infty} \frac{1}{n} \sum \tilde{x}_i = \sigma_x^2 + O(n^{-1})$

~~plim~~ $E(\sum \tilde{x}_i u_i)^2 = n \sigma_x^2 \sigma_u^2 + O(n^{-1})$

Cramer Wald: $\frac{\frac{1}{\sqrt{n}} \sum \tilde{x}_i u_i}{\frac{1}{n} \sum \tilde{x}_i^2} \xrightarrow{d} N\left(0, \frac{\sigma_u^2}{\sigma_x^2}\right)$

$\sum \tilde{S}_i^2 = \sum S_i^2 - \frac{1}{n} (\sum S_i)^2$

Pooled ordinary Least Square (POLS)

$\hat{\beta} = \frac{\sum^{NT} \tilde{y}_{it} \tilde{x}_{it}}{\sum^{NT} \tilde{x}_{it}}$ $N \times T$ (df)

$y = Xb + u$

$\sqrt{n}(\hat{\beta} - b) \xrightarrow{d} N(0, \Sigma_b)$

$\sqrt{n} \begin{pmatrix} \hat{b}_1 - b_1 \\ \hat{b}_2 - b_2 \end{pmatrix} \xrightarrow{d} N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{pmatrix} \sigma_{11}^2 & \sigma_{12} \\ \sigma_{12} & \sigma_{22}^2 \end{pmatrix} \right)$

$f(\hat{b}_1, \hat{b}_2) = f(b_1, b_2) + \frac{\partial f(b_1, b_2)}{\partial b_1} (\hat{b}_1 - b_1) + \frac{\partial f(b_1, b_2)}{\partial b_2} (\hat{b}_2 - b_2)$

$+ \frac{1}{2} \frac{\partial^2 f(b_1, b_2)}{\partial b_1^2} (\hat{b}_1 - b_1)^2 + \frac{1}{2} \frac{\partial^2 f(b_1, b_2)}{\partial b_2^2} (\hat{b}_2 - b_2)^2 + \frac{1}{2} \frac{\partial f(b_1, b_2)}{\partial b_1 b_2}$

$(\hat{b}_1 - b_1)(\hat{b}_2 - b_2) + \ldots$

$\sqrt{n} (\hat{r} - r) = [\alpha_0 \quad \alpha_1] \begin{bmatrix} \sqrt{n}(\hat{b}_1 - b_1) \\ \sqrt{n}(\hat{b}_2 - b_2) \end{bmatrix}$

$\longrightarrow N\left(0, [\alpha_0 \quad \alpha_1] \begin{bmatrix} \sigma_{11}^2 & \sigma_{12} \\ \sigma_{21} & \sigma_{22}^2 \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix} \right)$

$= N(0, \sigma_r^2)$

$W = (Rb - q)' [\sigma^2 R (X'X)^{-1} R']^{-1} (Rb - q) \xrightarrow{d} \chi_1^2$

$\sigma^2 (X'X)^{-1} = \Sigma_b$

$\frac{\sqrt{n}(\hat{r} - r)}{\sqrt{\sigma_r^2}} \xrightarrow{d} N(0,1)$

$f(\hat{b}_1, \hat{b}_2) - f(b_1, b_2) = \frac{\partial f(b_1, b_2)}{\partial b_1} (\hat{b}_1 - b_1) + \frac{\partial f(b_1, b_2)}{\partial b_2} (\hat{b}_2 - b_2) + R_N$

$\Rightarrow \hat{r} - r = \frac{\partial f(b_1, b_2)}{\partial b_1} \underbrace{(\hat{b}_1 - b_1)}_{O_p(n^{-1/2})} + \frac{\partial f(b_1, b_2)}{\partial b_2} (\hat{b}_2 - b_2) + R_N \rightarrow O_p(n^{-1})$

$\hookrightarrow (\hat{b}_1 - b_1)(\hat{b}_2 - b_2) = O_p(n^{-1})$

$(\hat{b}_1 - b_1)^2 = O_p(n^{-1})$

$\Rightarrow \sqrt{N} (\hat{r} - r) = \frac{\partial f(b_1, b_2)}{\partial b_1} \sqrt{n}(\hat{b}_1 - b_1) + \frac{\partial f(b_1, b_2)}{\partial b_2} \sqrt{n}(\hat{b}_2 - b_2) + O_p\left(\frac{1}{\sqrt{n}}\right)$

$= \left[\frac{\partial f(b_1, b_2)}{\partial b_1} \quad \frac{\partial f(b_1, b_2)}{\partial b_2} \right] \begin{bmatrix} \sqrt{n}(\hat{b}_1 - b_1) \\ \sqrt{n}(\hat{b}_2 - b_2) \end{bmatrix} + O_p\left(\frac{1}{\sqrt{n}}\right)$

$\xrightarrow{d} N\left(0, \left[\frac{\partial f(b_1, b_2)}{\partial b_1} \quad \frac{\partial f(b_1, b_2)}{\partial b_2} \right] \begin{bmatrix} \sigma_{11}^2 & \sigma_{12} \\ \sigma_{21} & \sigma_{22}^2 \end{bmatrix} \begin{bmatrix} \frac{\partial f(b_1, b_2)}{\partial b_1} \\ \frac{\partial f(b_1, b_2)}{\partial b_2} \end{bmatrix} \right)$

- Least Sq Dummy Reg (LSDV)
- Withins Group (WG)
- Fixed Effect (FE)

$\hat{\beta}_{WG} = \hat{\beta}_{FE} = \hat{\beta}_{LSDV}$

$= \frac{\sum^N \sum^T (y_{it} - \frac{1}{T} \sum y_{it}) \times (x_{it} - \frac{1}{T} \sum x_{it})}{\sum^N \sum^T (x_{it} - \frac{1}{T} \sum x_{it})^2}$

$df = N(T-1)$

Random Effect

$y_{it} = a_i + \beta x_{it} + u_{it} = a + \beta x_i + \underbrace{a_i - a + u_{it}}_{e_{it}}$

Serially Correlated

GLS ⚠ PRCM

POLS