Econometric Course of Professor Sul @ UTD: first session

Meisam Hejazinia

1/14/2013

Try to attend the tutor sessions, on friday 9-10.

Dreive limiting distribution, and mathematical relations. Econometrics self study would be impossible if you do not learn to how to read in this class.

You need to study yourself to avoid forgetting that is quite an epidemy.

We will talk about theoritical econometric, which would be pure mathematic. You need to understand which test is used for what situation.

In econometric II you will learn about the applied econometrics.

Only 45 minutes from 9:00 am you will have the first test. Test format would be the same. You need to memorize until page 33 for test one, some is covered and some of them not, and you need to memorize very notations.

Econometricians have their own commonalities, and language. You need to memorize this language, and these are alphabets that you need to memorize.

Study hard for test.

Memorize all the matrix algebra formulas.

"Normal" distribution:

About 95% of statistics when size of sample increases will be the normal distribution.

If two samples is took x_1, x_2 and x_3, x_4 is took and we take the average, or you increase the size for example take three $\frac{x_1, x_2, x_3}{3}$ this is called statisitc, if the size of sample is increased, the statistic would be in the form of normal distribution.

In many cases the distribuion is not normal. Income distribution is in the form of free scale, log normal distribution, as an example.

If you take integral, it would be cumulative form.

if $x \sim N(\mu, \sigma^2)$ then if you add a number to it, then only the mean will change so we will have $x + 100 \sim N(\mu + 100, \sigma^2)$.

Do not ask average since one could be 10 million and change everything. Centeral tendency is median and you need to ask that, rather than average.

For central tendancy median is better than average.

Things that already happened is non-random, but to find evidance as a researcher it will become random.

If you multiply a number to a random variable when it is random then we will have $2x \sim N(2\mu, 4\sigma^2)$

Check wikipedia for studying the distriutions.

$Beta(\alpha, \beta)$ distribution has lots of shapes.

You may not think about total size, but think about share that you got, or portion of it, and it would be in the form of Beta distribtuion.

Beta distribution can not create the periodical form, but truncate normal can.

In normal distribution mean and variance are independant. In other distribution than normal when you change mean, variance will change.

If distribution is normal estimate of mean and variance would be independent.

Weibull distribution is used for growth models.

Logistic distribution OM and Marketing students should know.

Cauchy distribution does not have mean. It could be in the form of undefined fraction of number divide by zero.

First moment is mean, and second moment is variance.

Variance $V(X) = E((X - E(X))^2]$ variantion toward the mean.

The moment generating function helps to define distribution function based on Tailor extension.

When you have cumulant function you would be able to use Tailor extension to create the main function.

Large Sample Distribution Theory

Notation is very important in econometric, and you must use the same notation.

x_n is totally different from x.

$\rightarrow^p$ is also important notation.

x_i is cross sectional variable

x_t is time series variable

x_{it} is a panel variable

Cross sectional variable and time series ar totally different and each could lead to different result.

x_n stands for $i = 1, \ldots, n$
x_T stands for $t = 1, , T$
$x_n T$

$x_n = \frac{1}{n} \sum_{i=1}^{N} x_i$
For a dice the average would be $E(x) = 1.\frac{1}{6} + 2.\frac{1}{6} + 3.\frac{1}{6} + \ldots + 6.\frac{1}{6} = 3.5$

If you through a dice and you take the average it is not 3.5.

If we through a dice four times the average would be x_4.

If you increase the sample size x_{100} would become closre to 3.5.

When n is increases for throughing this dice x_n mean the sample mean would converge to 3.5.

How fast this estimate is converaged dependes on the method you use for estimation.

For this we will put $x_n \rightarrow^p 3.5$ mean it converges probability to 3.5, and it may not actually be equal to it.

First one shows that distant will go to zero, but second one which says $prob(lim_{n\rightarrow\infty} x_n = c) = 1$ is stronger saying that $^{a.s.}c$ mean almost sure it will go to c with probability of one.

Part three and four will tak about distribution of sample mean.

When you increase the mean the variance will go to zero, and eventually the probability will be equal to one.

$\frac{1}{n}\sum^n (x_i - c)^2 \to 0$ as $n \to \infty$.

Consistent estimator

There are many different definitions.

$plim_{n\to\infty} x_n = c$ is called consistent estimator.

Example: $x_n = c + \epsilon/n$
$E(X_n) = c + E\frac{\epsilon}{n}$
if E $\epsilon \neq 0$
$E\epsilon = 0.1$
$E(X_n) = c + \frac{0.1}{n} \neq c$
$x_n \to^p c$
$plim x_n = c$

$EX_n = c$ and $plim x_n = c \Rightarrow plim x_n = c$

$Ex_n = c$ x_n is unbiased estimator (method of estimation)

estimate: number you got.

i.i.d: identical independent distributed.

No matter how many courses you take, you may usually forgot these econometrics, if you don't use them. It took professor one year after 11 courses to review and recall econometrics for his paper.

Median can be writtne in the form of weighted mean $\sum w_i x_i$

$\frac{1}{n}\sum x_i \to^p \mu$

to calculate variance $E[\frac{1}{n}\sum_{i=1}^n (x_i - c)]^2 = E[\frac{1}{n^2}(x_1 + \ldots + x_n - c - \ldots c)^2] = \frac{1}{n^2}E((\sum x_i)^2 + c^2.n - 2\sum x_i c)$

$[\frac{1}{n}\sum(x_i - c)]^2 = [\frac{1}{n}(\sum x_i - \sum c)]^2 = \frac{1}{n}^2(\sum x_i - \sum c)^2 = \frac{1}{n^2}[(\sum x_i)^2 - 2(\sum x_i)(\sum c) + (\sum c)^2]$

$x_i \sim i.i.d(0, \sigma^2)$

if $c = 0$

$E(\frac{1}{n}\sum x_i)^2 = E\frac{1}{n^2}(\sum x_i)^2 = \frac{1}{n^2}E(x_1 + x_2 + \ldots + x_n)^2$

$x_i \sim d(0, \sigma^2)$

$E(\frac{1}{n}\sum x_i)^2 = E\frac{1}{n^2}(x_1 + \ldots + x_n)^2 = \frac{1}{n^2}E[x_1^2 + \ldots + x_n^2 + x_1(x_2 + \ldots + x_n) + x_2(x_1 + x_3 + +x_n) + \ldots + x_n(x_1 + x_2 + \ldots + x_{n-1})]$

$E(x_1^2) = E(X_2^2) = \sigma^2$ identical

$E(x_1 x_2) =?$ if not independent

For couple of weeks we assume the correlation would not exist mean $E(x_1 x_2) = 0$

Therefore we got:

$= \frac{1}{n^2}n.\sigma^2 = \frac{1}{n}\sigma^2$

$E(\frac{1}{n}\sum x_i)^2 = E(x_n - c)^2$

$x_n \to^p 0$

$x_i \sim i.i.d(\mu, \sigma^2)$

$\frac{1}{n}\sum x_i = x_n \to^p \mu$

$lim_{n\to\infty E[\frac{1}{n}\sum(x_i-\mu)]^2=0}$

$x_i \sim d(mu, \sigma^2)$

$x_i - \mu \sim d(0, \sigma^2)$

$z_i \sim d(0, \sigma^2)$

$x_i \sim i.i.d.(\mu_i, \sigma^2)$

identical assumption could be released, yet independance assumption is still required.

$plim\frac{x_n}{y_n} = \frac{plim x_n}{plim y_n}$

$plim\frac{\frac{1}{n}\sum x_i}{\frac{1}{n}\sum y_i}$

Sample mean and any statisitcs if sample size goes to infinity will converge to normal distribution.

Sample mean is central tendency, and if that central tendency is estimated will comply normal distribution.

Mean median and Mode are central tendency, yet when distribution is not symmetric, median is better.

In econometrics notation is the key. x_n and x have totally different meaning.

$x_i \sim i.i.d(0, \sigma^2)$

$\frac{1}{n} \sum_i$

$E \frac{1}{n} \sum x_i = \frac{1}{n} E(x_1 + \ldots + x_n) = \frac{1}{n} = 0$

$E(\frac{1}{n} \sum x_i)^2 = \frac{1}{n} \sigma^2$

$\frac{1}{n} \sum x_i \to N(0, \frac{1}{n} \sigma^2)$

$\sqrt{n} x_n \to^d N(0, \sigma^2)$

$\sqrt{n} \frac{1}{n} \sigma x_i \to^d N(0, (\sqrt{n}) \frac{1}{n} \sigma^2)$

$\sqrt{n} x_n \to^d N(0, \sigma^2)$

$\frac{\sqrt{n} x_n}{\sigma^2} \to^d N(0,1) :$ t-statistic

$x_i \sim i.i.d.(\mu, \sigma^2)$

$\frac{x_i - \mu}{z_i} \sim i.i.d.(0, \sigma^2)$

$x_i \sim i.i.d.(\mu, \sigma^2)$

$y_i = x_i + c$

$x_i t = \epsilon_i + \theta_t$

$\epsilon_i \sim i.i.d.(0, \sigma^2)$

$\theta_t \sim i.i.d.(0, \sigma_\theta^2)$

$x_{nt} = \frac{1}{n} \sum x_{it}$

$\sqrt{n}(x_{nt} - \theta_t) \to^d N(0, \sigma^2)$

Econometrics 28th Jan ①

- percision – Not missing any character

chapter 1 : Single Variable

Types: Ⓐ Cross sectional : lab / Field Experiment $\rightarrow x_i,\ i=1,\dots,N$

Ⓑ time series : histories $\rightarrow x_t,\ t=1,\dots,T$

$\Rightarrow$ Panel data $\rightarrow x_{it}$

most sophisticated: $x_{it} = \beta_{it} \cdot \theta_t$

$N \times T$ info $\quad$ $N \times T$ info $\quad$ T

$\boxed{\text{Econ meaning}}$

$\boxed{\text{time Varying Common factor representation}}$

θ_t: Common behavior

β_{it}: Economic distance

- in the class teacher is Common factor that shows Correlation

$E(x_{it}\, x_{jt}) = \beta_{it}\beta_{jt} E(\theta_t^2)$

$\rightarrow$ DNA is diffirent Very small – Education is Common factor

---

Complexity of panel makes us:

$x_{it} = \beta_i \theta_t + \varepsilon_{it} = \underbrace{(\beta_i + \frac{\varepsilon_{it}}{\theta_t})}_{\beta_{it}} \theta_t$: $\boxed{\text{approximate Common factor representation}}$

β_i: factor loading

ε_{it}: ideosyncratic

---

even this one is Complex, so:

$x_{it} = \alpha_i + \theta_t + \varepsilon_{it}$ for any fixed t

Two way Compound model

$x_t = 1 = -2+1+2$

we assume independence here for symplification

You observe: $x_{it} = \alpha_i + \theta_0 + \varepsilon_{i0}$

you Can Calculate $\begin{cases} \text{mean, median} \\ \text{Top 10\%, low 10\%} \\ \text{distribution} \end{cases}$

$x_{it} = \alpha_i + \theta_t + \theta_{it}$

$\wedge$: hat = estimate

$\frac{1}{N}\sum^{N} x_{it} = \hat{\mu}_x$

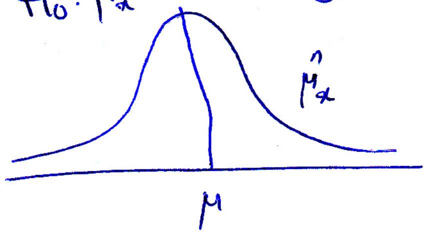
Estimator $\quad$ Estimate

Evaluate $\hat{\mu}_x$: How?

②

$H_0: \hat{\mu}_x = 0$ null hypothesis

dist. is unknown



Approximation of the unknown distribution

$\Downarrow$

Asymptotic distribution $\quad N \rightarrow \infty$

= limiting distribution

Lindberg-Levy Central limit theorem (L-L CLT)

$\frac{1}{N}\sum^{N} x_{it} = \hat{\mu}_x$

$\sqrt{N}(\hat{\mu}_x - \mu_x) \xrightarrow{d} N(0, \sigma_x^2)$

θ_t is here

$\frac{1}{N}\sum_{i=1}^{N} x_{it} = \frac{1}{N}\sum \alpha_i + \theta_t$

$\frac{1}{N}\sum^{N} x_{it} = \underbrace{\frac{1}{N}\sum^{N}\alpha_i}_{\text{I}} + \underbrace{\theta_t}_{\text{II}} + \underbrace{\frac{1}{N}\sum^{N}\varepsilon_{it}}_{\text{III}}$

I : $\frac{1}{N}\sum \alpha_i$

assume α : $\alpha_i \sim iid(\alpha, \sigma_\alpha^2)$

$\frac{1}{N}\sum \alpha_i$

① $E(\frac{1}{N}\sum \alpha_i)$: $= \alpha$

② $E(\frac{1}{N}\sum \alpha_i - \alpha)^2$: $\frac{1}{N} E[\frac{1}{N}\sum(\alpha_i - \alpha)^2]$

③ L-L CLT

assumption: $E(\alpha_i - \alpha)(\alpha_j - \alpha) = 0$ independence

in time series serial Correlation Exists and so they would not be independent

② $\Rightarrow$ $= \frac{1}{N}\sigma_\alpha^2$

③ L.L: $\hat{\alpha} = \frac{1}{N}\sum \alpha_i$

wrong Expression

$(\hat{\alpha} - \alpha) \xrightarrow{d} N(0, \frac{\sigma_\alpha^2}{N})$

Since as $N \rightarrow \infty$ $\frac{\sigma_\alpha^2}{N}$ goes to Zero

Correct one is:

$\sqrt{N}(\hat{\alpha} - \alpha) \xrightarrow{d} N(0, \sigma_\alpha^2)$ Correct

③ II: is constant so CLT can not be used

III: ① $E(\frac{1}{N}\sum^{N}\epsilon_{it})$ ② $E(\frac{1}{N}\sum^{N}\epsilon_{it}-0)^2$

③ L.L CLT:

$$\sqrt{N}\left(\frac{1}{N}\sum\epsilon_{it}\right)\xrightarrow{d} N(0,\sigma_\epsilon^2)$$

$$\Rightarrow \frac{1}{\sqrt{N}}\sum\epsilon_{it}\xrightarrow{d} N(0,\sigma_\epsilon^2)$$

Combining: I + II + III:

$E(\hat{\mu}_x)=\alpha+\theta_t=\mu_x$

over "N"

$$\sqrt{N}(\hat{\mu}_x-\mu_x)\xrightarrow{d} N(0,\sigma_\alpha^2+\sigma_\epsilon^2)$$

1990: $\hat{\mu}_{x,1990}$ $N=10$ mil 10

$\hat{\mu}_{x,2010}$ $N=10$ mil 11

as long as θ_t is different $\hat{\mu}_{x,t}$ would be different

→ in this case you need to compare relatively and fix it, and not say anything about statistics, Econometrics

- you will say compared to 1990 the mean increased by one

$$\hat{\sigma}_\epsilon^2=\frac{1}{N}\sum_{i=1}^{N}\left(\epsilon_{it}-\frac{1}{N}\sum\epsilon_{it}\right)^2$$

①②③

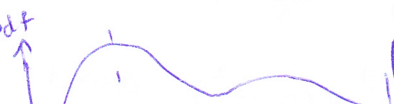
① $E[\hat{\sigma}_\epsilon^2]$

② $E\left[\frac{1}{N}\sum\left(\epsilon_{it}-\frac{1}{N}\sum\epsilon_{it}\right)^2-\sigma_\epsilon^2\right]^2$

③ L.L:

$$\sqrt{N}(\hat{\sigma}_\epsilon^2-\sigma_\epsilon^2)\xrightarrow{d} N(0,\boxed{\ })$$

pdf



② Calculate this

$\boxed{\text{Associate Variable}}$

Dummies

④

- it is called «mixed distribution» from two different population

→ Each paper different story → Fiction novel

$x_i=\alpha_i+\epsilon_{it}$ "ignor θ_t"

$x_{it}=\alpha_i+\lambda_i\theta_t+\epsilon_{it}$

e.g. IQ

e.g. Effort attitude

→ Not observable since you have only one cross section so assume it a number

λ_i & θ_t: not correlated

$$\Rightarrow x_{it}=\underbrace{\alpha_i+\lambda_i^*}_{\alpha_i^*}+\epsilon_{it}$$

⇒ same form

So for groups you will have:

$$x_i\begin{cases}\alpha_1+\epsilon_{it} & \text{if } i\in G_1\\ \alpha_2+\epsilon_{it} & \text{if } i\in G_2\\ \vdots & \vdots\end{cases}$$

- for example male and female different income:

$$\begin{cases}x_i=\alpha_1+\epsilon_i & \text{if } i\in F\\ x_i=\alpha_2+\epsilon_i & \text{if } i\in M\end{cases}$$

$$\hat{\alpha}_1\neq\hat{\alpha}_2\quad\begin{cases}\frac{1}{N_F}\sum^{N_F}x_i=\hat{\alpha}_1\\ \frac{1}{N_M}\sum^{N_M}x_i=\hat{\alpha}_2\end{cases}$$

①, ②, ③ → F.M

$$\sqrt{N_F}(\hat{\alpha}_1-\alpha_1)\xrightarrow{d} N(0,\sigma_1^2)$$

$$\sqrt{N_M}(\hat{\alpha}_2-\alpha_2)\xrightarrow{d} N(0,\sigma_2^2)$$

$\hat{\alpha}=\hat{\alpha}_1-\hat{\alpha}_2$

$E(\hat{\alpha})=\alpha$

$\hat{\alpha}-\alpha=(\hat{\alpha}_1-\hat{\alpha}_2-\alpha_1+\alpha_2)$

$$\sqrt{n}(\hat{\alpha}-\alpha)\xrightarrow{d} N(0,\underbrace{\sigma_1^2+\sigma_2^2}_{A})$$

Z or t: $\dfrac{\sqrt{N}(\hat{\alpha}-\alpha)}{\sqrt{A}}\xrightarrow{d} N(0,1)$

$\boxed{\text{Econometrics}}$ $\boxed{\text{28th Jan}}$ (5)

$$\hat{\alpha} = (Z'Z)^{-1} Z'X$$

$$Z = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}_{n\times 1} \qquad Z'Z = [1\ 1 \ldots 1]\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \sum_{i=1}^{N} 1 = N$$

$$Z'X = [1\ 1 \ldots 1]\begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} = \sum x_i$$

$$\hat{\alpha} = \frac{1}{N}\sum x_i$$

$x_i = \alpha_1 Z_i + \varepsilon_i$ if $i = F$

$x_i = \alpha_2 Z_i + \varepsilon_i$ if $i = M$

$x_i = \alpha + \beta w_i + \varepsilon_i$ (5) $\quad w_i = \begin{cases} 0 & \text{if } i \in F \\ 1 & \text{o.w.} \end{cases}$ (6)

$E(x_i) = \alpha_1$ if $i \in F$

$E(x_i) = \alpha_2$ if $i \in M$

$E(x_i)$ in (5) $= \alpha + \beta E(w_i) + \overset{0}{E(\varepsilon_i)} =$

$$\begin{cases} \alpha + \beta & \text{if } E(w_i) = 1 \\ \alpha & \text{if } E(w_i) = 0 \end{cases}$$

$\alpha + \beta = \alpha_1 \qquad \beta = \alpha_1 - \alpha_2$

$\alpha = \alpha_2$

given (6)

$$x_i = \overset{\alpha}{\alpha_2} + \overset{\beta}{(\alpha_1 - \alpha_2)} w_i + \varepsilon_i$$

$$X = Q\gamma + \varepsilon$$

$$Q = \underset{N\times 2}{\begin{bmatrix} 1 & 1 \\ \vdots & 0 \\ \vdots & 0 \\ \vdots & 0 \\ \vdots & 1 \\ \vdots & \vdots \\ 1 & 1 \end{bmatrix}} \qquad \gamma = \underset{2\times 1}{\begin{bmatrix} \alpha_1 \\ \alpha_1 - \alpha_2 \end{bmatrix}}$$

w_i

$$\hat{\gamma} = \begin{bmatrix} \hat{\alpha}_1 \\ \widehat{\alpha_1 - \alpha_2} \end{bmatrix} = (Q'Q)^{-1} Q'X$$

Skilled		nonskilled	
M	F	M	F

(6)

$$x_i \begin{cases} \alpha_1 & \text{if } i \in F\ i \in S \\ \alpha_2 & \text{if } i \in F\ i \in NS \\ \alpha_3 & \text{if } i \in M\ i \in S \\ \alpha_4 & \text{if } i \in M\ i \in NS \end{cases}$$

- You have four here, you need one more dummy in the previous regression equation therefore
- Find that variable as $\boxed{\text{Home work}}$
- if you have three level of variation You would need 8 variables including Constant

Q: what null hypothesis:

$\boxed{(\alpha_1 = \alpha_3 \ \& \ \alpha_2 = \alpha_4)}$

H_0: There is NO Gender wage Difference

H_0: There is NO skilled wage differen

H_0^A: $\alpha_1 = \alpha_3$ Not reject

H_0^B: $\alpha_2 = \alpha_4$ reject

- Matching one to one dummy helps you to test Various hypothesis
- why Asian earn more mony for example

Econometric | Feb 04 | ①

- Finish of the Dummy Variable Regression
- Threatment Effect

- Dummy Variable : Cross sectional Regression

- You try to Calc. the mean
- Median: will talk about

Identification

y: Gender — S skill — N — race

Hierarchical | struct $G^{(\alpha)}$ M, F | $S^{(\beta)}$ S, U | $R^{(\gamma)}$ W-NW | Baseline

$E(y_i) = a$ 0 0 0

$E(y_i) = a$ 0 0 γ

$E(y_i) = a$ $\boxed{0 \quad 1 \quad 0}$

$a + \beta$

$y_i = a + a_1 G_i + a_2 S_i + a_3 R_i + \underbrace{a_4 G \cdot S_i + a_5 G R_i + a_6 S R_i}_{\text{CROSS products}}$

Gender Dummy, Skill Dummy

$+ a_7 G_i S_i R_i + \underbrace{\varepsilon_i}$

Random Error → must be zero

$E[y_i] = a = b_1$ mean for $(G, S, R) = (0, 0, 0)$

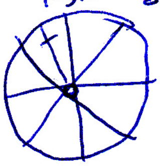
There are 8 unknowns

y_i — G: M — S — R: b_1, b_2; R: b_3, b_4; F — S — R: b_5, b_6; R: b_7, b_8

Gender Dummy, skill Dummy, Race Dummy

Once you find a's and b's it will become streight forward to check hypothesis of $b_1 \overset{?}{=} b_5$

MSW : $y_i = b_1 + \varepsilon_i$



$E\ y_i = a + a_1 = b_5$ if $G_i = 1, S_i = 0, R_i = 0$

Regional Dummy | $G \left[\begin{smallmatrix}0 \\ 1\end{smallmatrix}\right.$

- Female Salary difference of Region

Reason: Home Production

②

- Nobody pays for home production

Texas home production ~~cost~~ is higher than California, so they are not in labor market

- You write fiction model and want to test

- Theory said he is criminal but in stat you check exactly opposite: check he is not criminal

Alternative hypoth: (objective) A is a bad, A > B

Null :	$A = B$	$A \le B$
Alt :	$A \neq B$	$A > B$

- For 48 Dummy var → hierarchical approach

$$D = \begin{array}{c|ccc} & A & B & C \ldots \\ y_1 & 1 & 0 & 0 \\ y_2 & 0 & 1 & 0 \\ \vdots & 0 & \vdots & \\ y_{10} & 1 & 0 & \\ \vdots & \vdots & \vdots & \\ y_N & 0 & & \end{array}$$

$D_{1i} = \begin{cases} 1 & \text{if } i \in A \\ 0 & \text{if otherwise} \end{cases}$

$D_{2i} = \begin{cases} 1 & \text{if } i \in B \\ 0 & \text{other} \end{cases}$

$\vdots$

$D_{47} = \begin{cases} 1 \\ 0 \end{cases}$

Vector var, Vector

$y_i = a + D_i \gamma + a_2 S' + (D_i S_i) \phi$

$\gamma = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_{47} \end{bmatrix}$

47 Cross Dummies

hierarchy Theoretically possible, yet interpretation would be harder

3

Dummy Variable : Cross sectional Regression

$T_i = \begin{cases} \text{Controlled Case} & : 0 \\ \text{Treated} & : 1 \end{cases}$

test score – [Control / Treatment

You designed dummy Endogeneous

$y_i = \alpha + \beta T_i + \varepsilon_i$

$H_0 : \beta = 0$

→ attitude – Education

$\varepsilon_i = \gamma A_i + u_i$

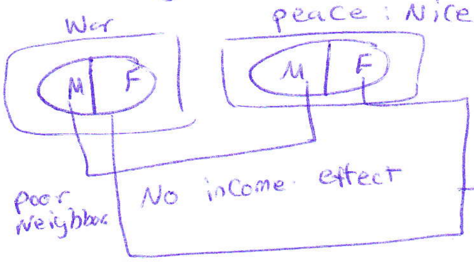
Correlation

- Dummy Should be exogenous : ① Gender ② Race

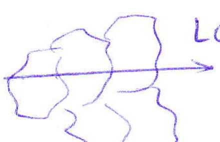
  You should not be able to Change

- exogenous : Given
- Skill Dummy is not Exogenous it is indogenous
- spanish use English income higher
- Medicine [Control : No medicine / treated : medicine
- most lab Experiment / not filed okay

if experiment in field – Condition must be exogenous



- in Lab Experiment most Variables are exogenous
- all Social Economic Variables are Endegenous



Location is Longitude not latitude is Endogenous

4

Endogeniety Could be environment

political   economical

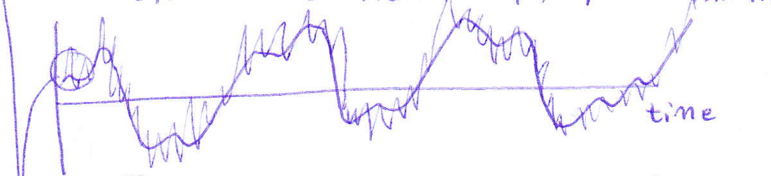
- but we define with polar opposit mean things that are not Exogenous

Single variable : time series

- James hamilton book for more info

y_t

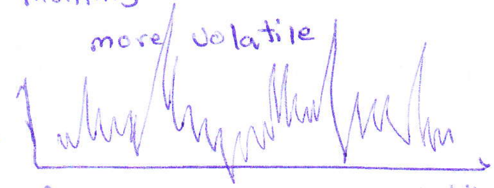
Stock market index, GDP, GPA



Frequency: Daily, monthly, annual, milisecond

→ monthly with fluctuation

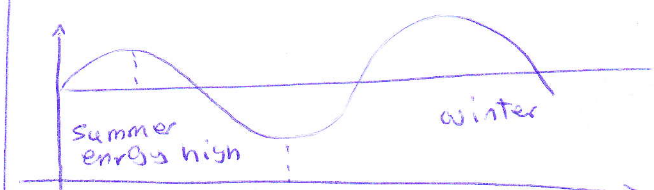
more volatile



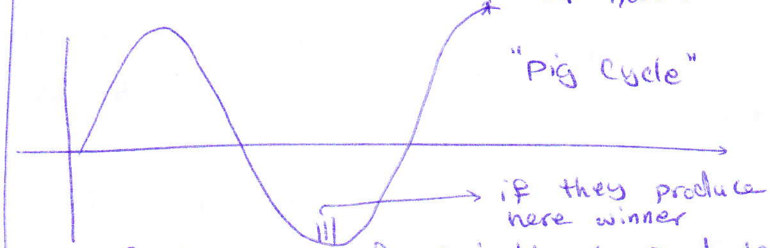
hourly – Extremely volatile

high frequence : Stock Price

(*) as frequency decreases volatility & Varience decreases



Seasonal



Cycle : price of agricultral Products

milk price, home price, looser

... could be 1,000 years : Global

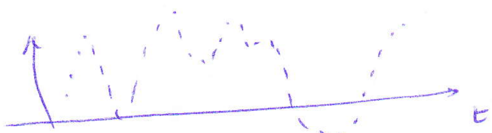
Econometrics | Feb 04 | 5

- we approximate plot assigned func.
  look for cycles

- mean we don't Care since it fluctuates
- mean is time dependent
  what we Care:

  1. variance

     distribution we don't Care, on special Cases only [special time series]



  2. Correlation: between times

Auto regressive Model: AR(1)

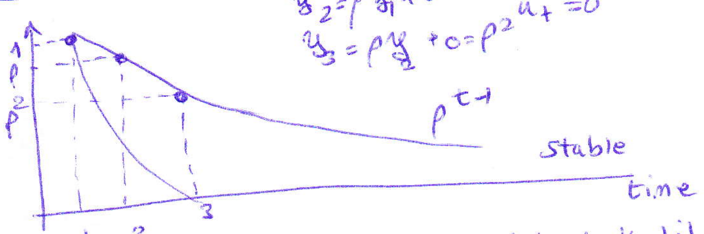
$y_t = \rho y_{t-1} + u_t$ → Shock

ρy_{t-1} ↓ lag dependent variable

Economic interpretation

$y_2 = \rho y_1 + 0 = \rho$

$y_3 = \rho y_2 + 0 = \rho^2 \quad u_t = 0$



- lab experiment if repeated will look like this

ρ: stands for decay function

$t \to \infty \longrightarrow$ will go to Zero

$E(y_t) = \rho E(y_{t-1}) + \underbrace{E(u_t)}_{0}$

$E(y_t) = 0$

6

Bubble period - Cycle

if $\rho = 1$, y_t is "Random walk"
"non-stationary"

- what ever shock yr syst will become stable

$y_t = \rho y_{t-1} + u_t = \rho(\rho y_{t-2} + u_{t-1}) + u_t =$

$\rho^2 y_{t-2} + \rho u_{t-1} + u_t = \rho^2(\rho y_{t-2} + u_{t-2}) + \rho u_{t-1} + u_t$

$= \rho^3 y_{t-3} + \rho^2 u_{t-2} + \rho u_{t-1} + u_t$

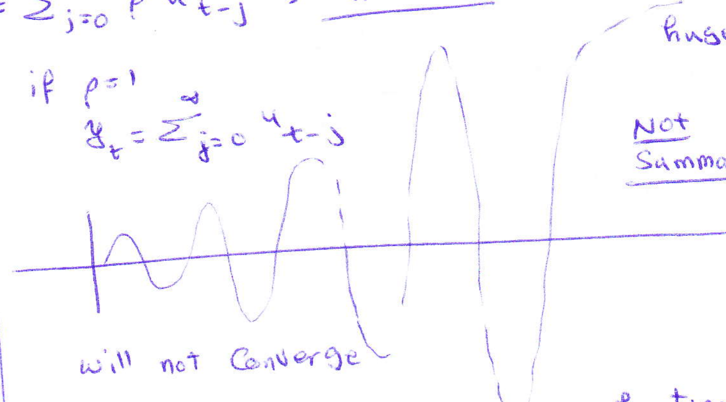
$y_t = (\rho^\infty y_{t-\infty} + \rho^{\infty-1} u_{t-\infty+} \dots + \rho u_{t-1} + u_t) + \rho^{\infty+1} y_{t-\infty+}$

$= \sum_{j=0}^{\infty} \rho^j u_{t-j}$ = summable

if $\rho = 1$

$y_t = \sum_{j=0}^{\infty} u_{t-j}$

Not Summable



If you want to forecast, become fortune teller
for non stationary Case it will be impossible

$$\boxed{\begin{aligned} & E[\tfrac{1}{T}\sum u_t] = 0 \\ & \tfrac{1}{T}\sum u_t \neq 0 \\ & \tfrac{1}{T}\sum u_t \xrightarrow{p} 0 \\ & \text{if } E\ \tfrac{1}{T}\sum u_t = 0 \\ & E(\tfrac{1}{T}\sum u_t)^2 = (\tfrac{1}{T}\sigma_u^2) \times T = \sigma_u^2 \end{aligned}}$$

Auto Correlation

$u_t = iid(0, \sigma_u^2)$

$E\, y_t^2 = E[u_t^2 \dots$

$\dfrac{E\, y_t y_{t-1}}{E\, y_t^2}$

There shld be 3 dots

$E\, y_t^2 = \dfrac{\sigma_u^2}{1-\rho^2}$

$y_t^2 = (\sum \rho^i u_{t-i})^2 = (u_t + \rho u_{t-1} + \rho^2 u_{t-2} + \dots)^2$

$$E\,y_t^2 = E[u_t^2 + \rho^2 u_t^2 + \dots] = \sigma_u^2(1+\rho^2+\rho^4+\dots)$$ (7)

$$= \sigma_u^2\left(\frac{1}{1-\rho^2}\right)$$

if $\rho=1$, $y_t = \sum_{s=0}^{t} u_s = (\overset{0}{\overset{\shortparallel}{u_0}} + u_1 + u_2 + \dots + u_t)$

$$E[y_t^2] = \sigma_y^2(t+\dots+1) = t\sigma_u^2 \text{ or } (t+1)\sigma_u^2$$

Variance increase by t

$\boxed{\text{Single variable : time series}}$

$$y_t = \rho y_{t-1} + \overset{\text{purely random shock}}{u_t}$$

$$\hat\rho = (X'X)^{-1}X'y = \frac{\sum y_{t-1}y_t}{\sum y_{t-1}^2} = \frac{\sum y_{t-1}(y_{t-1}\rho + u_t)}{\sum y_{t-1}^2}$$

$$= \rho + \underbrace{\frac{\sum y_{t-1}u_t}{\sum y_{t-1}^2}}_{\text{Noise}}$$

$$\hat\rho - \rho = \frac{\sum y_{t-1}u_t}{\sum y_{t-1}^2}$$

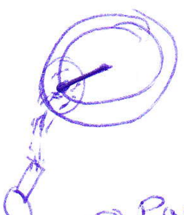
$E[\hat\rho - \rho] = ?$

if Dummy $\rightarrow$ deterministic Can kick out

unbias

$$E\,\frac{\sum y_{t-1}u_t}{\sum y_{t-1}^2} \neq \frac{E\sum y_{t-1}u_t}{E\sum y_{t-1}^2} \qquad E\left(\frac{A}{B}\right) \neq \frac{EA}{EB}$$

$$\underset{T\to\infty}{\text{plim}}\ \frac{\frac{1}{T}\sum y_{t-1}u_t}{\frac{1}{T}\sum y_{t-1}^2} = \frac{\text{plim}\ \frac{1}{T}\sum y_{t-1}u_t}{\text{plim}\ \frac{1}{T}\sum y_{t-1}^2}$$



Consistent

$$E[\sum y_{t-1}u_t] = E(y_1u_2 + y_2u_3 + \dots + y_{t-1}u_t)$$

⊛ Past even is not supposed to Correlate with Current shock

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

Assymmetric (8)

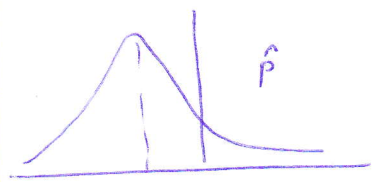
$$E(\hat\rho) = \rho - \frac{2\rho}{T}$$

$$E(\hat\rho - \rho) = -\frac{2\rho}{T} \qquad \text{This is bias}$$

⊛ $T\to\infty$ then bias would be zero

but T small we will have bias

⊛ $\hat\rho$ is Consistent estimator (plim), yet biased (but these two are different)



Extend using $y_2 = u_2 + \rho u_1 + \rho^2 u_0$

$$E(\hat\rho - \rho) = E\left(\frac{\sum y_{t-1}u_t}{\sum y_t^2}\right) = E\left(\frac{A}{B}\right) = \frac{E(A)}{E(B)} - \frac{cE\,B^2}{E\,B\,Cov(A,B)} = 0 - \frac{2\rho}{T}$$

this bias is okay, and it is small, so it is okay

$$y_t = a + \rho y_{t-1} + u_t \quad (3)$$ when a comes in pic. everything diff

$\hat\rho$

when more than one var not summation but matrix notation

$$\frac{1}{T}\sum y_t = a + \rho\frac{1}{T}\sum y_{t-1} + \frac{1}{T}\sum u_t \quad (4)$$

$$(2)-(4) = \left(y_t - \frac{1}{T}\sum y_t\right) = \rho\left(y_{t-1} - \frac{1}{T}\sum y_{t-1}\right) + \left(u_t - \frac{1}{T}\sum u_t\right) \quad (5)$$

$$\Rightarrow \hat\rho - \rho = \frac{\sum \tilde y_{t-1}u_t}{\sum \tilde y_{t-1}}$$

$$\text{let } \tilde y_t = \frac{\sum \hat y_t - u_t}{\sum \hat y_{t-1}^2}$$

$\boxed{\text{Econometrics}}$ $\boxed{\text{Feb 04}}$ (9)

$\tilde{y}_{t-1} = y_{t-1} - \frac{1}{T}\sum y_{t-1}$

$\tilde{u}_t = u_t - \frac{1}{T}\sum u_t$

$\sum \tilde{y}_{t-1}\tilde{u}_t = \sum y_{t-1}u_t - \frac{1}{T}(\sum y_{t-1})(\sum u_t)$

Proof: $\sum (y_{t-1} - \frac{1}{T}\sum y_{t-1})(u_t - \frac{1}{T}\sum u_t) =$

$\sum (y_{t-1}u_t - (\frac{1}{T}\sum y_{t-1})u_t - y_{t-1}(\frac{1}{T}\sum u_t)$

$+ (\frac{1}{T}\sum y_{t-1})(\frac{1}{T}\sum u_t)) = \sum y_{t-1}u_t - (\frac{1}{T}\sum y_{t-1})$

$(\sum u_t) - \cancel{(\sum y_{t-1})(\frac{1}{T}\sum u_t) + T\cdot(\frac{1}{T}\sum y_{t-1})(\frac{1}{T}\sum u_t)}$

$E\sum \tilde{y}_{t-1}\tilde{u}_t = \sum y_{t-1}u_t - E[(\frac{1}{T}\sum y_{t-1})(\sum u_t)]$

Correlation

$E[(\frac{1}{T}\sum y_{t-1})(\sum u_t)] = \frac{1}{T}E[y_1 + y_2 + y_3 + \dots + y_{t-2} + y_{t-1})$

$+ (u_2 + u_3 + \dots + u_{T-1} + u_T)$

$u_t \sim iid$

$\boxed{y_2 = u_2 + \rho u_1 + \rho^2 u_0 + \dots}$

$\frac{E[\frac{1}{T}(\sum \tilde{y}_{t-1})(\sum \tilde{u}_t)]}{E(\sum \tilde{y}^2_{t-1})} = \frac{1+\rho}{T}$ $\boxed{\text{Nickell Bias}}$

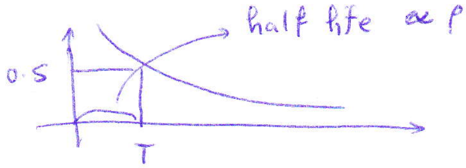
$E(\hat\rho - \rho) = E[\frac{\sum \tilde{y}_{t-1}\tilde{u}_t}{\sum \tilde{y}^2_{t-1}}] = E[\frac{\sum y_{t-1}\sum u_t}{\sum \tilde{y}^2_{t-1}}] -$

$E[\frac{\frac{1}{T}(\sum y_{t-1})(\sum u_t)}{\sum \tilde{y}^2_{t-1}}] = -\frac{2\rho}{T} - \frac{1+\rho}{T} = -\frac{1+3\rho}{T}$

$\boxed{\text{Kendall Bias}}$ Constant Case

$-\frac{1+3\rho}{T}$ CMP $\frac{-2\rho}{T}$

much Larger - people care a lot about



⊛ people try to Estimate half life

Nickel Bais $\longrightarrow$ related to panel data

$E(\hat\rho - \rho) = \frac{-2\rho}{T}$ (10)

$\begin{cases} \rho > 1 \Rightarrow \text{Bubble} \\ \rho = 1 \Rightarrow \text{unit Root} \\ \rho < 1 \Rightarrow \text{stationary} \end{cases}$

$(\hat\rho - \rho) \xrightarrow{d} ?$ $\begin{cases} \text{Normal} & \rho < 1 \\ \text{Dickey Fuller Gauss} & \rho = 1 \\ \text{Cauchy} & \rho > 1 \end{cases}$

⊛ depending on ρ value you need to use different distribution

You don't know ρ, So Conclusion of 30 years: YOU ARE SKREWED

$\boxed{\text{Econometrics Intes}}$ $\boxed{\text{Feb 8}}$

$d_i = \begin{cases} 1 & \text{if male} \\ 0 & \text{if female} \end{cases}$

$s_i = \begin{cases} 1 & \text{if skilled} \\ 0 & \text{if unskilled} \end{cases}$

unskill female	female skilled
I	II
III	IV
male unskilled	

① $E(y_i) = \begin{cases} b_1 & \text{if } i \in \text{female unskilled } (i \in I) \\ b_2 & \text{if } i \in \text{female \& skilled } (i \in II) \\ b_3 & \text{if } i \in \text{male \& unskilled } (i \in III) \\ b_4 & \text{if } i \in \text{male \& skilled } (i \in IV) \end{cases}$

$\begin{cases} y_i = b_1 + \varepsilon_i & \text{if } i \in I \\ E(\varepsilon_i) = 0 \\ y_i = b_2 + \varepsilon_i & \text{if } i \in II \\ y_i = b_3 + \varepsilon_i & \text{if } i \in III \\ y_i = b_4 + \varepsilon_i & \text{if } i \in IV \end{cases}$

test H_0^1: $b_2 = b_4$ not to reject

H_0^2: $\boxed{b \neq b_3 \; H_a}$ $\underline{b_1 = b_3 \text{ reject}}$

$y_2 = \alpha_0 + \alpha_1 d_i + \alpha_2 s_i + \alpha_3 d_i s_i + \varepsilon_i$

② $E(y_i) = \begin{cases} \alpha_0 & \text{if } i \in I, \; d_i = 0 \quad s_i = 0 \\ \alpha_0 + \alpha_2 & \text{if } i \in II \quad d_i = 0 \quad s_i = 1 \\ \alpha_0 + \alpha_1 & \text{if } i \in III \quad d_i = 1 \quad s_i = 0 \\ \alpha_0 + \alpha_1 + \alpha_2 + \alpha_3 & \text{if } i \in IV \quad d_i = 1 \quad s_i = 1 \end{cases}$

①②:

$\Rightarrow \begin{cases} \alpha_0 = b_1 \\ b_2 = \alpha_0 + \alpha_2 \\ b_3 = \alpha_0 + \alpha_1 \\ b_4 = \alpha_0 + \alpha_1 + \alpha_2 + \alpha_3 \end{cases} \Rightarrow$

$\boxed{\text{Economic interpretation}}$

$\begin{cases} \alpha_0 = b_1 & \text{Expected Earning among...} \\ \alpha_2 = b_2 - \alpha_0 = b_2 - b_1 & \text{difference b/w male... fem...} \\ \alpha_1 = b_3 - \alpha_0 = b_3 - b_1 \\ \alpha_3 = b_4 - \alpha_0 - \alpha_1 - \alpha_2 \\ \quad = b_4 - b_3 - b_2 + b_1 \end{cases}$

test $\begin{cases} H_0^1: \cancel{\alpha_0} + \cancel{\alpha_2} = \cancel{\alpha_0} + \alpha_1 + \cancel{\alpha_2} + \alpha_3 \Rightarrow \alpha_1 = -\alpha_3 \\ H_0^2: \alpha_0 \neq \alpha_0 + \alpha_1 \Rightarrow \alpha_1 \neq 0 \Rightarrow \alpha_1 = 0 \text{ reject} \end{cases}$

Econometrics | 11-Feb | (1)

Time series

order of magnitude

$1 = O(1)$

$0 = O(1)$

$T = O(T)$

$\hookrightarrow \frac{T}{T} = 1$

$\frac{1}{T} = O(1)$ → Constant

$\frac{1}{T} = O(\frac{1}{T})$ → more precise

$\frac{\frac{1}{T}}{\frac{1}{T}} = O(1)$

$\frac{1+3\rho}{T} = O(\frac{1}{T})$

$\Delta_T = 1 + \frac{1}{T} + \frac{1}{T^2} = O(1) + O(\frac{1}{T}) + O(\frac{1}{T^2}) = O(1)$

---

small o:

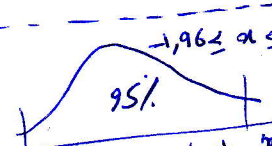
$\frac{1}{T} = o(1)$

$1 = o(T^{\alpha})$ → we don't express in this case

$\Delta_T = O(1) + o(1)$

Order in probability

$x \sim iid\ N(0,1)$



Small probability that x will go to zero, but x is finite usually

since x is Constant:

$x = O_p(1)$

Example: * $x_1 \sim N(0,\sigma^2)$

$x_1 = O_p(1)$ $\sigma^2 < \infty$

* $x_2 \sim N(0,\sigma^2 T)$

$\frac{x_2}{\sqrt{T}} \sim N(0,\sigma^2)$

$x_2 = O_p(1)$, $x_1 = O_p(\sqrt{T})$

* $x_3 \sim N(0, \frac{\sigma^2}{T}) \Rightarrow x_3 = O_p(\frac{1}{\sqrt{T}})$

* $x_4 = 1 + N(0,1) = O(1) + O_p(1) = O_p(1)$

means it is random variable, in contrast to Constant

$x_5 = 1 + N(0, \frac{1}{T}) = O(1) + O_p(\frac{1}{\sqrt{T}}) = O_p(1)$ (2)

$x_6 = N(2,T) = 2 + N(0,T) = O(1) + O_p(\sqrt{T}) = O_p(\sqrt{T})$

Rules

① large one dominate

② still random variable if added to Random Variable

From now on we will use big O since it is usefu

$x_t \sim iid\ N(0,\sigma^2)$

$\frac{1}{T}\sum x_t = O_p(\ ?\ )$

$\frac{1}{T}\sum x_t = \bar{x}_T$, let say

① $E(\bar{x}_T) = 0$

② $E(\frac{1}{T}\sum x_t - 0) = \sigma^2/T$

③ Liderberg Levy: CLT:

$(\bar{x}_t - 0) \xrightarrow{d} N(0, \frac{\sigma^2}{T})$ → Not Econometrics so after you write next line Erase this

$\sqrt{T}(\bar{x}_t - 0) \xrightarrow{d} N(0,\sigma^2)$

↳ Convergence rate

$\boxed{\sqrt{T}\, x_t = O_p(1),\ \bar{x}_t = O_p(\frac{1}{\sqrt{T}})}$

$x_t \sim iid\ N(1,\sigma^2)$

$\frac{1}{T}\sum x_t = O_p(\ ?\ )$

$\sqrt{T}(\bar{x}_T - 1) \xrightarrow{d} N(0,\sigma^2)$

$\sqrt{T}\,\bar{x}_T - \sqrt{T} = O_p(1)$

$\bar{x}_T = 1 + O_p(\frac{1}{\sqrt{T}}) = O_p(1)$

---

memorize the new notes for test 2

AR(1)

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

$V(y_t) = \frac{\sigma^2}{1-\rho^2}$

$E[\frac{1}{T}\sum y_t^2]$

$y_t = \sum_{j=0}^{\infty} \rho^j u_{t-j}$

when you learn languages study everyday to not forget & this is foreign language (Econometrics)

- Review every day
- do not underestimate this language (1 out of 10)
- your tenure depends on it

$y_t = \rho y_{t-1} + u_t = \rho(\rho y_{t-2} + u_{t-1}) + u_t$

$E(\frac{1}{T}\sum y_t^2) = E[\frac{1}{T}[\sum^t(\sum_{j=0} \rho^j u_{t-j})^2]]$

$E(y_t y_{t-1}) = \rho E(y_{t-1})^2 = \rho \frac{\sigma^2}{1-\rho^2}$

$E(\rho y_{t-1} + u_t) y_{t-1}$

$E(y_t y_{t-2}) = \frac{\sigma^2}{1-\rho^2}\rho^2$

$E(y_t y_{t-3}) = \rho^3 \frac{\sigma^2}{1-\rho^2}$

we studied this since cross products are dependent

$\frac{1}{T}\sum y_t \xrightarrow{d} ?$

① $E(\frac{1}{T}\sum y_t) = 0$

② $E(\frac{1}{T}\sum y_t - 0)^2 =$

$E(\frac{1}{T^2}(\sum y_t)^2) = \sum y_t^2 + 2\sum y_t y_{t-1} + \dots$

$E(\sum y_t)^2 = E(y_1^2 + y_2^2 + \dots + y_T^2) \neq$

$E[y_1(y_2 + \dots + y_T)] + E[y_2(y_1 + y_3 + \dots + y_T)] + \dots$

$= T\sigma_y^2 + \sigma_y^2(\rho + \rho^2 + \rho^3 + \dots) + \sigma_y^2(\rho + \rho + \rho^2 + \dots)$

$+ \sigma_y^2(\rho^2 + \rho + \rho + \rho^2 + \dots) =$

$\frac{1}{T^2} E(\sum y_t)^2 = \frac{1}{T}\frac{\sigma^2}{(1-\rho)^2}$ → long run Variance

$\frac{1}{T}\sum y_t \xrightarrow{d} N(0, \frac{1}{T}\frac{\sigma^2}{(1-\rho)^2})$

$\frac{1}{T}\sqrt{T}$

$\frac{1}{\sqrt{T}}\sum y_t \xrightarrow{d} N(0, \frac{\sigma^2}{(1-\rho)^2})$

Now on we will talk about Economics
until now it was statistics

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

α is not mean of y_t,

α does not capture mean of y_t

$E y_t = \alpha + \rho E(y_{t+1}) + \underbrace{E(u_t)}_{=0}$

$E y_t = \frac{\alpha}{1-\rho}$

it is not dummy regression to use $y_{it} = \alpha + \beta d_i + \rho y_{it-1} + u_{it}$ then the check whether $\beta = 0$

$E x_t = \frac{\alpha_1}{1-\rho_1}$

This is wrong!

This 'i' made your life prosperous

$y_{it} = \alpha + \beta d_i + \rho_i y_{it-1} + u_{it}$

They should

Latent model

$y_t = a + x_t$

$x_t = \rho x_{t-1} + u_t$

a is constant different from α

$E(y_t) = a$

$y_t = a + x_t$

$-\rho y_{t-1} = a\rho + \rho x_{t-1}$

$y_t - \rho y_{t-1} = a(1-\rho) + \underbrace{(x_t - \rho x_{t-1})}_{u_t}$

$y_t = \underbrace{a(1-\rho)}_{\alpha} + \rho y_{t-1} + u_t$

$E y_t = \frac{\alpha}{1-\rho} = \frac{a(1-\rho)}{1-\rho} = a$

- physics they don't have u_t: Error term
They do not allow mistakes

while in social science we allow Errors
approximate, asymptote, Econometrics

$y_t = \alpha + \rho y_{t-1} + u_t$ level accounting

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

$\boxed{\Delta y_t = \alpha + (\rho - 1) y_{t-1} + u_t}$: growth rate

Growth accounting

- it is insane, this is not growth, it is still ~~latent~~ latent, and since it is coming from latent, it is still latent

Non - Stationary | **Econometrics** | **11- Feb** | 5

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

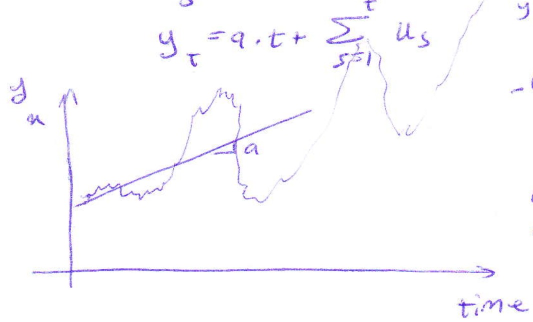
Let $\rho = 1$, $y_1 = a + u_1 = 0$ let say

$y_1 = a$

$y_2 = a + 1 \cdot a + u_2 = 2a + u_2 + u_1$

$y_3 = 3a + u_3 + u_2 + u_1$

$y_T = a \cdot t + \sum_{s=1}^{t} u_s$



- Carl marx growth rate in Capitalism about GDP. more fluctuation

---

Fore Casting

$y_t = \alpha + \beta x_{t-1} + u_t$

1 ——— T/2 ——— T

$\hat{\alpha}, \hat{\beta}$; $T/2 = P$

$\hat{y}_{P+1} = \hat{\alpha} + \hat{\beta} x_P + \hat{u}_{P+1}$ → Removed since $E_T[u_{T+1}] = 0$

if you know you are god so we remove for definition

$\hat{y}_{P+1|P} = \hat{\alpha} + \hat{\beta} x_P$ ①

$y_t = a + \rho y_{t-1} + u_t$ ② → if you dont know ρ you have to run 2nd one

1980's forecasting was popular

$y_t = a + \rho y_{t-1} + u_t$ AR(1)

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

$y_t = a + \sum_{j=1}^{p} \rho_j y_{t-j} + u_t$: AR(P)

p different fore Casts

$\hat{y}_{T+1|T} = \hat{\alpha} + \sum_{j=1}^{p} \hat{\rho}_j y_{T-j+1}$

Random walk model

$\rho = 1$, $P = 1$

$a \neq 0$ drift

6

$\hat{y}_{T+1|T} = \hat{a} + y_T$; $\hat{a} = 0$

Time Series | **Fore Cast Error**

$\varepsilon_{T+1}^{*} = y_{T^*+1} - \hat{y}_{T^*+1|T^*}$

$\frac{1}{T/2} \sum_{i=1}^{T/2} |\varepsilon_{T^*+i}|$ Mean prediction absolute error (MPAE) ✓ → This one is better

$\frac{2}{T} \sum_{i=1}^{T/2} \varepsilon_{T+i}^2$ Mean prediction Square Error (MPSE)

we want now to do second horizon

using t information and fore Cast

fortune telling to plan

$\hat{y}_{T+2|T}$

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

① $y_t = a + \rho(a + \rho y_{t-2} + u_{t-1}) + u_t = a(1+\rho) + \rho^2 y_{t-2} + \rho u_{t-1} + u_t$ multiple step

② $y_t = a_2 + \rho^2 y_{t-2} + u_{t,t-2}$ ← direct

First, from ①, $\hat{y}_{T+2|T} = \hat{a} + \hat{\rho}\, \hat{y}_{T+1|T}$

From ②, $\hat{y}_{T+2|T} = \hat{a}_2 + \hat{\rho}^2 y_T$

iterative

direct fore Cast is better if regression model is misspesified

if not,

Direct = Iterative

True: MA(1)

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

Regression: $y_t = a + \rho y_{t-1} + u_t$

Long horizon forecast

$y_{t+k} = \alpha_k + \beta_k x_t + u_{t+k,t}$

k th horizon forecast,

No evidance $\beta_1 \neq 0$

as $k \to M < \infty$

Stronger evidence $\beta_k \neq 0$ 7

first horizon

$$y_t = \alpha + \beta x_{t-1} + u_t \quad (3)$$

second horizon

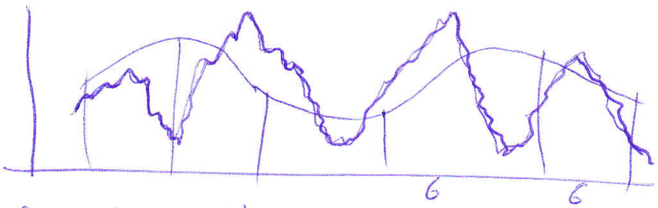
$$y_{t+1} = \alpha_2 + \beta_2 x_{t-1} + ?$$

$$x_t = \rho x_{t-1} + \varepsilon_t$$

from (3)

$$y_{t+1} = \alpha + \beta x_t + u_{t+1} = \alpha + \beta\rho x_{t-1} + \underbrace{\beta\varepsilon_t + u_{t+1}}$$

$$= \alpha + \beta_2 x'_{t-1} + u_{t+1,t}$$



for forecasting, first they try to find big cycles

- use cos - sin function split cycles:

frequency analysis

$$\underbrace{y_t = y_t^P}_{\text{permanent}} + \underbrace{y_t^C}_{\text{cycle}} \rightarrow \text{cyclical cycle}$$

- Definition of business cycle? $\boxed{Hw}$

HP

BP filter

Band past

Phillips Coordinating filter

$$\underset{\text{non stationary}}{y_t} = \underset{\text{non stationary}}{y_t^P} + \underset{\text{non stationary}}{y_t}$$

No proof

Fundamental - base line forecast

- forecast are wrong, since there is huge shock

- GDP: two years to find out true Gross Domestic Product, usually for last year is forecast & correct it

- they do in two year horizon 8

as time goes you get true information and you correct forecast

"Common factor" or

"Combined forecasts"

AR(p) iterative) 2p

p Direct

"p"

$$\sum_{i=1}^{p} w_i Y_{t+1,i|t} = \frac{1}{p} \sum \hat{y}_{t+1,i|t}$$

it is usually best forecast -

take average forecast

right now currently the best forecast is this one

Method of detecting bubble on stock - realestate

Macro switching - good state - bad state

vector autoregressive project -

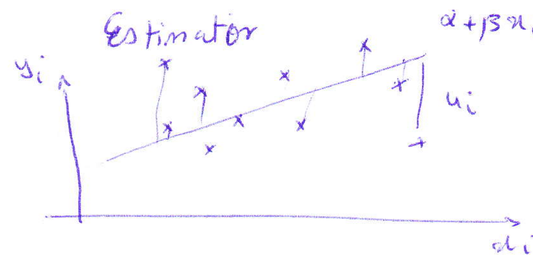
$\boxed{\text{Two Variable}}$ Econometrics

asymptotic properties

$$y_i = \alpha + \beta x_i + u_i$$

Ordinary Least Squares (OLS)

Estimator



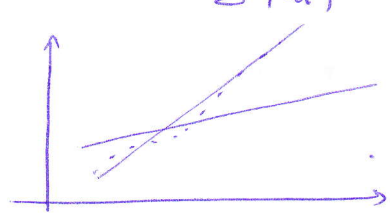
change slope, and try to minimize distance

change slope

$$\sum u_i^2$$

$$\arg\min_{\alpha,\beta} \sum u_i^2$$

$\sum |u_i|$ Least absolute Error (LAE)



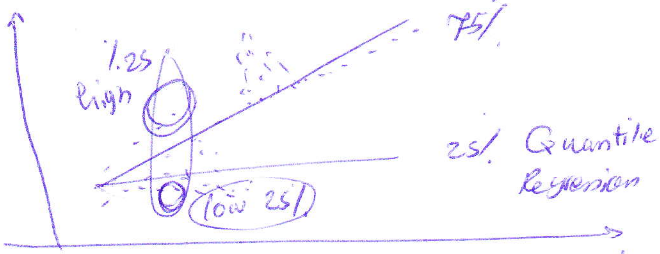
- minimize median of u_i Could be third Estimator to discard outlier

median $(u_i) \longrightarrow$ 0.5 probability

$\alpha, \beta \quad u_1, u_2, \dots, u_n$

then median estimator look like: $\underbrace{\sum w_i u_i^2}$

Weighted Least square Estimator (WLSE)



- ideas Come simply and you don't have to think about mathematics but use intuition

4th Regression Could be

Quantile Regression

- high income & low income different Regression

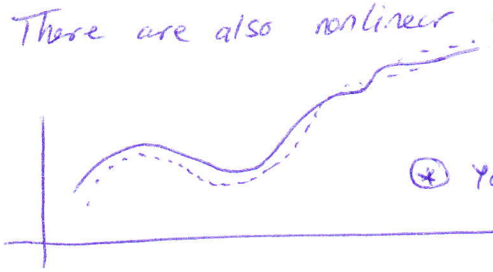
Split to two

minimize %25 Quantile

- two different Coefficient by running Regression

$\longrightarrow$ flexible models when you know more Econometrics

- There are also nonlinear estimators

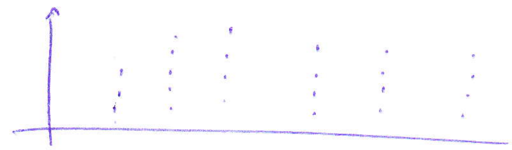


$y_i = f(x_i) + u_i$

* You try to minimize this one

- when you don't know functional form

Several points



Econometrics → tutor

Bonus Question

$y_{it} = \mu_i + \varepsilon_{it}, \quad \varepsilon_{it} \sim iid(0, \sigma_\varepsilon^2 N^{-\alpha})$

and $\mu_i \sim iid(0, \sigma_\mu^2 N^{-\beta})$

(a) $\alpha > 1, \beta = 0, \mu_i = iid(0, \sigma_\varepsilon^2)$

$\bar{y}_{TN} = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} y_{it} = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T}$

$(\mu_i + \varepsilon_{it}) = \underbrace{\frac{1}{NT} \sum_i \sum_t \mu_i}_{\frac{1}{N}\sum \mu_i} + \frac{1}{NT} \sum_i \sum_t \varepsilon_{it}$

$= \frac{1}{N} \sum_i \mu_i + \frac{1}{NT} \sum_i \sum_t \varepsilon_{it}$

$E(\bar{y}_{TN}] = E[\frac{1}{N} \sum_i \mu_i + \frac{1}{NT} \sum_i \sum_t \varepsilon_{it}]$

$= \frac{1}{N} E[\mu_1 + \mu_2 + \mu_3 + \cdots + \mu_N] + \frac{1}{NT} E(\varepsilon_{11} + \varepsilon_{12} + \cdots$

$+ \varepsilon_{1T} + \varepsilon_{21} + \varepsilon_{22} + \cdots + \varepsilon_{2T} + \cdots + \varepsilon_{N1} + \varepsilon_{N2} + \cdots$

$+ \varepsilon_{NT}) = 0 + 0 = 0$

$V(\bar{y}_{TN}) = E[\bar{y}_{TN} - 0]^2 = E[\frac{1}{N} \sum_i \mu_i + \frac{1}{NT}$

$\sum_i \sum_t \varepsilon_{it}]^2 = \frac{1}{N^2} E[\sum_i \mu_i] + \frac{1}{N^2 T^2} E[\sum_i \sum_t \varepsilon_{it}]^2$

$+ \underbrace{E[\text{Com product})}_{0}$ ε_{it} and μ_{it} indep

$\beta = 0$

$\frac{1}{N^2}(N \sigma \mu^2) + \frac{1}{N^2 T^2} NT \sigma_{\varepsilon N}^2 =$

$\frac{1}{N} \sigma_\mu^2 + \frac{1}{NT} \sigma_\varepsilon^2 = \frac{1}{N} \sigma_\mu + \frac{1}{NT} \sigma_\varepsilon^2 N^{-\alpha}$

$= \frac{1}{N} \sigma_\mu^2 + \frac{\sigma_\varepsilon^2}{N^{1+\alpha} T}$

$\bar{y}_{TN} \xrightarrow{d} N(0, \frac{1}{N} \sigma_\mu^2 + \frac{1}{N^{(1+\alpha)} T} \sigma_\varepsilon^2)$

$\sqrt{N}\, \bar{y}_{TN} \xrightarrow{d} N(0, \sigma_\mu^2 + \frac{1}{\underbrace{N^\alpha T}_{\text{on } N, T \to \infty}} \sigma_\varepsilon^2)$

0

(a) $\Rightarrow \sqrt{N}\, \bar{y}_{TN} \xrightarrow{d} N(0, \sigma_\mu^2) \quad \alpha > 1$

b) $\alpha > 0 \quad \beta > 0$

$V(\bar{y}_{TN}) = \frac{1}{N^{1+\beta}} \sigma_\mu^2 + \frac{1}{N^{1+\alpha} T} \sigma_\varepsilon^2$

$\Rightarrow \bar{y}_{TN} \xrightarrow{d} N(0, \frac{1}{N^{1+\beta}} \sigma_\mu^2 + \frac{1}{N^{1+\alpha} T} \sigma_\varepsilon^2)$

$\Rightarrow N^{\beta/2} \sqrt{N}\, \bar{y}_{TN} \xrightarrow{d} N(0, \sigma_\mu^2 + \frac{\sigma_\varepsilon^2}{N^{\alpha-\beta} T})$

① if $\alpha > \beta \quad \frac{1}{N^{\alpha-\beta} T} \to 0$

② if $\alpha = \beta \quad \frac{1}{N^{\alpha-\beta} T} \to 0$

③ if $\alpha < \beta \quad N^{\frac{1+\beta}{2}} \bar{y}_{TN} \to N(0, \sigma_\mu^2 + \frac{N^{\beta-\alpha}}{T} \sigma_u^2$

① assume $\frac{N^{\beta-\alpha}}{T} \to 0$, if $N, T \to \infty$ jointly

② assume $\frac{N^{\beta-\alpha}}{T} \to c$, if $N, T \to \infty$ jointly

③ assume $\frac{N^{\beta-\alpha}}{T} \to \infty$, if $N, T \to \infty$ jointly

① $N^{(1+\beta)/2} \bar{y}_{TN} \xrightarrow{d} N(0, \sigma_\mu^2)$

② $N^{\frac{(1+\beta)}{2}} \bar{y}_{TN} \xrightarrow{d} N(0, \sigma_\mu^2 + c\, \sigma_\varepsilon^2)$

③ $N^\alpha \sqrt{N}\, y_{TN} \xrightarrow{d} N(0, \frac{1}{N^{\beta-\alpha}} \sigma_\mu^2 + \frac{1}{T} \sigma_\varepsilon^2$

$\sqrt{T}\, N^{\alpha/2} \sqrt{N}\, \bar{y}_{TN} \xrightarrow{d} N(0, \frac{T}{N^{\beta-\alpha}} \sigma_\mu^2 + \sigma_\varepsilon^2$

Econometrics Intro

Feb 8

$$d_i = \begin{cases} 1 & \text{if male} \\ 0 & \text{if female} \end{cases}$$

$$s_i = \begin{cases} 1 & \text{if skilled} \\ 0 & \text{if unskilled} \end{cases}$$

unskill female	female skilled
I	II
III	IV

male unskilled

① $$E(y_i) = \begin{cases} b_1 & \text{if } i \in \text{female unskilled } (i \in I) \\ b_2 & \text{if } i \in \text{female \& skilled } (i \in II) \\ b_3 & \text{if } i \in \text{male \& unskilled } (i \in III) \\ b_4 & \text{if } i \in \text{male \& skilled } (i \in IV) \end{cases}$$

$$\begin{cases} y_i = b_1 + \varepsilon_i & \text{if } i \in I \\ E(\varepsilon_i) = 0 & \\ y_i = b_2 + \varepsilon_i & \text{if } i \in II \\ y_i = b_3 + \varepsilon_i & \text{if } i \in III \\ y_i = b_4 + \varepsilon_i & \text{if } i \in IV \end{cases}$$

test H_0^1: $b_2 = b_4$ not to reject

H_0^2: $\boxed{b_1 \neq b_3}$ H_a $\quad$ $b_1 = b_3$ reject

$$y_2 = \alpha_0 + \alpha_1 d_i + \alpha_2 s_i + \alpha_3 d_i s_i + \varepsilon_i$$

② $$E(y_i) = \begin{cases} \alpha_0 & \text{if } i \in I,\ d_i = 0\ \ s_i = 0 \\ \alpha_0 + \alpha_2 & \text{if } i \in II\ \ d_i = 0\ \ s_i = 1 \\ \alpha_0 + \alpha_1 & \text{if } i \in III\ \ d_i = 1\ \ s_i = 0 \\ \alpha_0 + \alpha_1 + \alpha_2 + \alpha_3 & \text{if } i \in IV\ \ d_i = 1\ \ s_i = 1 \end{cases}$$

①②:

$$\Rightarrow \begin{cases} \alpha_0 = b_1 \\ b_2 = \alpha_0 + \alpha_2 \\ b_3 = \alpha_0 + \alpha_1 \\ b_4 = \alpha_0 + \alpha_1 + \alpha_2 + \alpha_3 \end{cases} \Rightarrow$$

Economic interpretation

$$\begin{cases} \alpha_0 = b_1 \\ \alpha_2 = b_2 - \alpha_0 = b_2 - b_1 \\ \alpha_1 = b_3 - \alpha_0 = b_3 - b_1 \\ \alpha_3 = b_4 - \alpha_0 - \alpha_1 - \alpha_2 \\ \quad = b_4 - b_3 - b_2 + b_1 \end{cases}$$

Expected Earning among...

difference b/w male... fem...

test $$\begin{cases} H_0^1: \cancel{\alpha_0} + \cancel{\alpha_2} = \cancel{\alpha_0} + \alpha_1 + \cancel{\alpha_2} + \alpha_3 \Rightarrow \alpha_1 = -\alpha_3 \\ H_0^2: \alpha_0 \neq \alpha_0 + \alpha_1 \Rightarrow \alpha_1 \neq 0 \Rightarrow \alpha_1 = 0 \text{ reject} \end{cases}$$

Econometrics Tutor Feb 15 ①

Calculate Rules

① $O_p(n^\alpha) \pm O_p(n^\beta) = \begin{cases} O_p(n^\alpha) & \text{if } \alpha \geq \beta \\ O_p(n^\beta) & \text{if } \alpha < \beta \end{cases}$

$O(n^\alpha) \pm O(n^\beta) = \begin{cases} O(n^\alpha) & \text{if } \alpha \geq \beta \\ O(n^\beta) & \text{if } \alpha < \beta \end{cases}$

② $\begin{cases} O_p(n^\alpha) O_p(n^\beta) = O_p(n^{\alpha+\beta}) \\ O_p(n^\alpha) / O_p(n^\beta) = O_p(n^{\alpha-\beta}) \end{cases}$

$\begin{cases} O(n^\alpha) O(n^\beta) = O(n^{\alpha+\beta}) \\ O(n^\alpha)/O(n^\beta) = O(n^{\alpha-\beta}) \end{cases}$

③ $c \cdot O_p(n^\alpha) = O_p(n^\alpha)$ $c=3$ $3\,O_p(n^\alpha) = O_p(n^\alpha) + O_p(n^\alpha) + O_p(n^\alpha) = O_p(n^\alpha)$

c is Constant

$c\,O(n^\alpha) = O(n^\alpha)$

④ c is Constant c (LEMMA)

$c = O_p(1) = O(1)$

---

① $x_n \sim N(0, \sigma^2/n)$, $x_n = O_p(\quad)$

$\boxed{x_n \sim N(0, \sigma^2),\ x_n = O_p(1)}$ → principle

$\sqrt{n}\, x_n \sim N(0, \sigma^2) \Rightarrow \sqrt{n}\, x_n = O_p(1)$

$n \to \infty$ limiting dist $\Rightarrow x_n = \frac{1}{\sqrt{n}} O_p(1) = O_p\left(\frac{1}{\sqrt{n}}\right)$

---

② if $x_n \sim N(1, \sigma^2/n)$ $x_n = O_p(\quad)$

$x_n - 1 \sim N(0, \sigma^2/n)$

$\sqrt{n}(x_n - 1) \sim N(0, \sigma^2)$

$\sqrt{n}(x_n - 1) = O_p(1)$

$x_n - 1 = \frac{1}{\sqrt{n}} O_p(1) = O_p\left(\frac{1}{\sqrt{n}}\right)$

$x_n = O_p\left(\frac{1}{\sqrt{n}}\right) + 1 = \underbrace{O_p\left(\frac{1}{\sqrt{n}}\right)}_{n^{-1/2}} + \underbrace{O_p(1)}_{n^0} = O_p(1)$

→ random variable includes Constant

③ $\sqrt{n}(\bar{x}_n - \bar{\mu}_n) \xrightarrow{d} N(0, \sigma^2)$ then $\bar{x}_n - \bar{\mu}_n = O_p(n^{-1/2})$, $\bar{x}_n = O_p(\quad)$ ②

$\sqrt{n}(\bar{x}_n - \bar{\mu}_n) = O_p(1)$

$\bar{x}_n - \bar{\mu}_n = \frac{1}{\sqrt{n}} O_p(1) = O_p\left(\frac{1}{\sqrt{n}}\right)$

$\bar{x}_n = O_p\left(\frac{1}{\sqrt{n}}\right) + \underbrace{\bar{\mu}_n}_{\text{Constant}} = O_p\left(\frac{1}{\sqrt{n}}\right) + O_p(1) = O_p(1)$

(with $\frac{1}{\sqrt{n}} \to n^{-1/2}$, $O_p(1) \to n^0$)

---

⑥ $y_t = \alpha + u_t$ $u_t \sim \text{iid}(0,1)$

true Data generating process

$y_t = \alpha x_t + e_t$ Regression $x_t = 1$

$\hat{\alpha} = \frac{\sum x_t y_t}{\sum x_t^2}$

Generally $y_t = b x_t + u_t$ (Error term) as $\min_b L = \sum_{t=1}^T u_t^2$

$\hat{b} \quad = \sum_{t=1}^T (y_t - b x_t)^2$

$\frac{\partial L}{\partial b} = \sum_{t=1}^T 2(y_t - b x_t)(-x_t)$

$\sum_{t=1}^T (y_t - b x_t)(-x_t) = 0$

$\sum_{t=1}^T (-y_t x_t + b x_t^2) = 0$

$\sum_{t=1}^T (-y_t x_t) + \sum_{t=1}^T b x_t^2 = 0$

$b \sum_{t=1}^T x_t^2 = \sum_{t=1}^T y_t x_t \Rightarrow \hat{b} = \frac{\sum_{t=1}^T x_t y_t}{\sum_{t=1}^T x_t^2}$

* Since there are many b and we want one

$\boxed{\begin{aligned} & x'(y - bx) = 0 \Rightarrow b(x'x) = x'y \\ & \Rightarrow b = (x'x)^{-1} x'y \end{aligned}}$ my own note

$\hat{b} = \frac{\sum_{t=1}^T x_t (b x_t + u_t)}{\sum_{t=1}^T x_t^2} = \frac{b \sum_{t=1}^T x_t^2 \sum_{t=1}^T x_t u_t}{\sum_{t=1}^T x_t^2}$

$= b + \frac{\sum_{t=1}^T x_t u_t}{\sum_{t=1}^T x_t^2} = b + \frac{\sum_{t=1}^T x_t u_t}{\sum_{t=1}^T x_t^2}$

$y = b x + u$

$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = b \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ $\hat{b} = (x'x)^{-1} x'y = b + (x'x)^{-1} x'u$

$[x_1\ x_2 \dots x_T] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_T \end{bmatrix} = \sum x_t^2$

③

$$x'Y = [x_1 \; x_2 \; \cdots \; x_T] \begin{bmatrix} Y_1 \\ \vdots \\ \vdots \\ Y_T \end{bmatrix} = \sum x_t Y_t$$

$$L = u^2 = (y_t - bx_t)^2$$

$$\frac{\partial L}{\partial b} \Rightarrow \begin{cases} \dfrac{\partial Ax}{\partial x} = A' \\ \dfrac{\partial Ax}{\partial x'} = A \\ \dfrac{\partial (x'Ax)}{\partial x} = 2Ax \\ \dfrac{\partial (x'Ax)}{\partial A} = xx' \end{cases}$$

---

6. $\overset{DGP}{y_t = a + u_t}$ $\quad u_t \sim iid(0,1)$

$y_T = \alpha x_t + e_t$

$$\hat{\alpha} = \frac{\sum x_t y_t}{\sum x_t^2} = \frac{\sum y_t}{\sum 1} = \frac{\sum (a+u_t)}{T} = \frac{\sum a + \sum u_t}{T}$$

$$= \frac{Ta + \sum u_t}{T} = a + \frac{\sum u_t}{T}$$

$$\hat{a} - a = \frac{\sum u_t}{T} \qquad u_T \sim iid(0,1)$$

$$\frac{1}{T}\sum u_t \sim iid(0, \tfrac{1}{T})$$

$$E(\tfrac{1}{T}\sum u_t) = \tfrac{1}{T} E(u_1 + u_2 + \cdots + u_T) = 0$$

$$V(\tfrac{1}{T}\sum u_t) = E[\tfrac{1}{T}\sum u_t - 0]^2 = \frac{1}{T^2} \underbrace{E(\sum u_t)^2}_{T \times 1}$$

$$= \frac{1}{T^2} \times T = \frac{1}{T} \qquad \frac{1}{T}\sum u_t \xrightarrow{d} N(0, \tfrac{1}{T})$$

$$\Rightarrow \hat{\alpha} - a \xrightarrow{d} N(0, \tfrac{1}{T})$$

$$\sqrt{T}(\hat{\alpha} - a) \xrightarrow{d} N(0,1)$$

---

make your own assumption for Questions

---

7. $y_t = at + u_t \qquad u_t \sim iid(0,1)$

regression $\quad y_t = at + e_t$

$$\hat{a} = \frac{\sum t y_t}{\sum t^2} = \frac{\sum t(at + u_t)}{\sum t^2} = \frac{a \sum t^2 + \sum t u_t}{\sum t^2}$$

$$= a + \frac{\sum t u_t}{\sum t^2} \Rightarrow \hat{a} = a + \left(\frac{\sum t u_t}{\sum t^2}\right)$$

④

$$\sum t^2 = 1^2 + 2^2 + \cdots + T^2 = \frac{t(t+1)(2t+1)}{6}$$

$$= \frac{T^3}{3} + \frac{T^2}{2} + \frac{T}{6} = \frac{T^3}{3} + O(T^2)$$

$$\underbrace{O(T^2) + O(T)}_{O(T^2)}$$

$$\sum t = 1 + 2 + \cdots + T = \frac{T(T+1)}{2} = \frac{T^2}{2} + \frac{T}{2} = \frac{T^2}{2} + O(T)$$

$$\sum t^3 = \frac{1}{4}T^4 + O(T^3)$$

memrize it

$$\sum (t^4) = \frac{1}{5}T^5 + O(T^4)$$

$\vdots$

Generall $\quad \boxed{\sum t^n = \frac{1}{n+1} O(T^{n+1}) + O(T^n)}$

$\Rightarrow$ back to Question

$T \to$ infinity

$$\hat{\alpha} = \frac{\sum t u_t}{T^3/3}$$

$$E(\sum t u_t) = E(u_1 + 2u_2 + 3u_3 + \cdots + Tu_T) = 0$$

$$V(\sum t u_t) = E(\sum t u_t)^2 = E(\sum t^2 u_t^2 + \cdots)$$

$$= E[u_1^2 + 2^2 u_2^2 + \cdots + T^2 u_T^2]$$

when $T \to \infty$

plug-in formula and Calculate final answer

Two Variables

Econometrics | Feb 18 (1)

$y_i = \beta x_i + u_i$

y_i ↓ dependent, Regression

x_i → Explanatory, independent, regressor

$\hat\beta = (X'X)^{-1}X'y$

$X = [x_1, \dots, x_n]'$

$X'X = \sum_{i=1}^n x_i^2$

$X'y = \sum_{i=1}^n x_i y_i$

$\Rightarrow \hat\beta = \frac{\sum x_i y_i}{\sum x_i^2} = \frac{\sum x_i(\beta x_i + u_i)}{\sum x_i^2}$

$= \beta \frac{\sum x_i^2}{\sum x_i^2} + \frac{\sum x_i u_i}{\sum x_i^2} = \beta + \frac{\sum x_i u_i}{\sum x_i^2}$

Assumption: (I) $E x_i u_j = 0$ for all i & j

$\Longleftrightarrow$ x is exogenous variable

(II) $u_i \sim iid(0, \sigma_u^2)$

Driving Limiting Dist

(I) $\hat\beta - \beta = \frac{\frac{1}{n}\sum x_i u_i}{\frac{1}{n}\sum x_i^2}$ (II) Let $\xi_i = x_i \cdot u_i \Rightarrow \frac{1}{n}\sum \xi_i = \frac{1}{n}\sum x_i u_i$

(1) $E \frac{1}{n}\sum \xi_i = 0$

(2) $E(\frac{1}{n}\sum \xi_i)^2 = \frac{1}{n} E[\frac{1}{n}\sum (x_1u_1 + x_2u_2 + \dots + x_nu_n)]^2$ (*)

$= \frac{1}{n} E[\frac{1}{n}[x_1^2u_1^2 + x_2^2u_2^2 + \dots + x_n^2u_n^2 + x_1u_1x_2u_2 + \dots]]$

on the other hand $E\, x_1u_1x_2u_2 = E(x_1x_2)E(u_1u_2) = E(x_1x_2)\cdot 0 = 0$,

Note that x_i could be dependent.

(*) $= \frac{1}{n}\sigma_x^2\sigma_u^2$

since $E x_i^2 = \sigma_x^2$ (assumption)

(3) L.L CLT: $\frac{1}{n}\sum \xi_i \xrightarrow{d} N(0, \frac{\sigma_x^2\sigma_u^2}{n})$ $\boxed{\text{Erase}}$

$\Rightarrow \frac{1}{\sqrt n}\sum \xi_i \xrightarrow{d} N(0, \sigma_x^2\sigma_u^2)$

$C_n \xrightarrow{p} c$

$\Rightarrow C_n \cdot \frac{1}{\sqrt n}\sum \xi_i \xrightarrow{d} N(0, c^2\sigma_x^2\sigma_u^2)$

(4) $\text{plim} \frac{1}{n}\sum x_i^2 = \sigma_x^2$

(III) $\sqrt n(\hat\beta - \beta) = \frac{\frac{1}{\sqrt n}\sum x_i u_i}{\frac{1}{n}\sum x_i^2} \xrightarrow{d} N(0, \frac{\sigma_x^2\sigma_u^2}{\sigma_x^2\sigma_x^2})$

$= N(0, \frac{\sigma_u^2}{\sigma_x^2})$

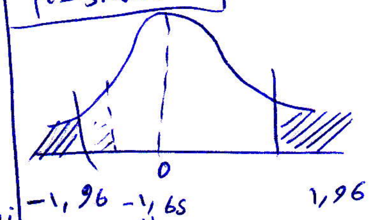
memorize step I, II, III and (1), (2), (3), (4) as well.

all you need is excercising

why we did this? (2)

$H_0: \hat\beta = 0$ $H_a: \hat\beta \neq 0$ two-side

$\boxed{\text{t-statistics}}$ t-ratio: $\frac{\sqrt n(\hat\beta - \hat\beta)}{\sqrt{\sigma_u^2/\sigma_x^2}} \xrightarrow{d} N(0,1)$



two side test

one side test $\hat\beta > 0$, $\hat\beta < 0$ one side

$\hat\sigma_u^2 = \frac{1}{n}\sum \hat u_i^2$

$(X'X)^{-1}$

$t\text{-stat} = \frac{\hat\beta}{\sqrt{\hat\sigma_u^2 (X'X)^{-1}}}$

$(X'X)^{-1} = \frac{1}{\sum x_i^2}$ Conventional t-statistics

always assumptions are not binding, so Conventional t-stat is wrong, especially for iid assumption.

Two variable Cont.

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

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

$x_t = \phi x_{t-1} + e_t$

Question: $t_{\hat\beta} = \frac{\hat\beta}{\sqrt{\hat\sigma_u^2/\sum x_t^2}} \xrightarrow{d} N(0,1)$?

- if this does not hold, you have to find mean and variance (could be $N(1,1)$)

$E[x_t, u_t] = 0\ \forall\ t$ & s exogeneous assumption, independence

(1) $\hat\beta - \beta = \frac{\frac{1}{T}\sum x_t u_t}{\sum x_t^2}$

(2) let $\xi_t = x_t u_t$

$\frac{1}{T}\sum \xi_t = \frac{1}{T}\sum x_t u_t$

(1) $\frac{1}{T}\sum \xi_t = 0$

(2) $E(\frac{1}{T}\sum \xi_t)^2 = \frac{1}{T} E \frac{1}{T}\sum \xi_t^2$

$+ E\frac{1}{T^2}\sum \xi_t \xi_{t+1} + \dots$

$\xi_t = x_t u_t$, $\xi_{t-1} = x_{t-1} \cdot u_{t-1}$

$E \xi_t \xi_{t-1} = E[x_t x_{t-1} u_t u_{t-1}] = E(x_t x_{t-1}) \cdot E(u_t u_{t-1})$

$\frac{1}{\sqrt T}\sum \xi_t \xrightarrow{d} N(0, ?)$

(iii) $\sqrt{T}(\hat\beta-\beta) = \dfrac{\frac{1}{\sqrt{T}}\sum x_t u_t}{\frac{1}{T}\sum x_t^2} \xrightarrow{d} N(0, ??)$

as <u>homework</u> drive it by yourself to memorize it

- don't just memorize result since small change in assumption can change everything

☐ final answer should be: $\boxed{\sigma_x^{-2}\,\Omega_\xi^2\,\sigma_x^{-2}}$

$\sqrt{T}(\hat\beta-\beta) \xrightarrow{d} N(0, \cancel{\Omega_\xi^2 / \Sigma x_t})$

$\Omega_\xi^2 = E\,\frac{1}{T}(\sum_t x_t u_t)^2 = \dfrac{\sigma_\varepsilon^2}{(1-\rho)^2}$

long run variance

if $\Omega_\xi = \sigma_u^2$ then:

$t_{\hat\beta} = \dfrac{\hat\beta}{\sqrt{\hat\sigma_u^2/\Sigma x_t^2}} \xrightarrow{d} N(0,1)$ is okay

long run variance is usually larger

$\Omega_\xi^2 \gg \sigma_u^2$

$\Omega_\xi^2 = ?$ $\qquad u_t = \rho u_{t-1} + \xi_t$

$E\,\frac{1}{T}(\Sigma u_t)^2 = \dfrac{\sigma_\varepsilon^2}{(1-\rho)^2} = \Omega_u^2$ long run variance

$E\,\frac{1}{T}\Sigma u_t^2 = \dfrac{\sigma_\varepsilon^2}{1-\rho^2}$ Contemporary variance

e.g.

$\rho = 0.9 \Rightarrow \rho^2 = 0.81 \Rightarrow 1-\rho^2 = 0.2$

$(1-\rho)^2 = 0.01$

$\Rightarrow$ 20 times larger

$\Omega_\xi^2 = E\,\frac{1}{T}(\Sigma u_t x_t)^2 = \bigcirc$

true t-stat

$t_{\hat\beta} = \dfrac{\hat\beta}{\sqrt{\Omega_\varepsilon^2/\Sigma x_u^2}}$

$<<<$

$t_{\hat\beta} = \dfrac{\hat\beta}{\sqrt{\sigma_u^2/\Sigma x_t^2}}$ Conventional → Stata gives you this

$t_\beta = -6.8$

true stat could be -0.68

→ on your thesis you used wrong variance

how to estimate Ω_ξ^2 ?

You need to use HAC estimator

Hetrodescedasticity Auto-Correlation Consistent

Newey (MIT) – West (Wisconsin) estimator (Econometrica 1984)

Andrews estimator

pre-whitening estimator

looks very complicated, so not cited, and others put in simple way, and get citation

try to prove yourself to never forget

$\hat\xi_t = \hat u_t x_t$

Run $\hat\xi_t = \psi\hat\xi_{t-1} + m_t$ prewhitening

$\frac{1}{T}\Sigma \hat m_t^2 = \hat\sigma_m^2$

$\hat\Omega_\xi^2 = \dfrac{\hat\sigma_m^2}{(1-\hat\psi)^2}$ ← re-coloring

$\xi_t = \psi\xi_{t-1} + m_t, \quad \psi = \rho^2$

$m_t = \xi_t e_t + \phi\varepsilon_t x_{t-1} + \rho u_{t-1} + e_t$

↓

is not serially correlate

We did serial correlation, what about hetrodescedicity?

$y_t = \beta x_i + u_i$

$u_i \sim ind(0, \sigma_i^2)$

Conv $t_{\hat\beta} = \dfrac{\hat\beta}{\sqrt{\hat\sigma_u^2/\Sigma x_i^2}}$

↓ ?

$\hat\beta - \beta = \dfrac{\frac{1}{n}\Sigma x_i u_i}{\frac{1}{n}\Sigma x_i^2}$

(2) $E(\frac{1}{n}\Sigma x_i u_i)^2 = \frac{1}{n}E(\frac{1}{n}[x_1^2u_1^2 + x_2^2u_2^2 + \dots + x_n^2u_n^2]) + \underbrace{E\ \text{Cross}}_{=0}$

why? independent

$= \frac{1}{n}\cdot\frac{1}{n}\Sigma E(x_i^2)\sigma_i^2$

if $E x_i^2 = \sigma_x^2 \ \forall i$ then $\frac{1}{n}\Sigma E(x_i^2)\sigma_i^2 = \frac{1}{n}\sigma_x^2\Sigma\sigma_i^2$

$\frac{1}{\sqrt{n}}\Sigma x_i u_i \xrightarrow{d} N(0, \sigma_x^2\sigma^2)$

where $\sigma^2 = \lim_{n\to\infty}\frac{1}{n}\Sigma\sigma_i^2$

Econometrics Feb 18

$\hat\sigma^2 = \frac{1}{n}\sum \hat u_i^2$

if $E x_i^2 \neq \sigma_x^2$,

$\frac{1}{n}\sum E(x_i^2)\sigma_i^2 = \frac{1}{n}\sum x_i^2 \hat u_i^2$

x_i is "given"

Sandwich

$\sqrt{n}(\hat\beta - \beta) \xrightarrow{d} N\left(0, \left(\frac{1}{n}\sum x_i^2\right)^{-1}\left(\frac{1}{n}\sum x_i^2 \hat u_i^2\right)\cdot\left(\frac{1}{n}\sum x_i^2\right)^{-1}\right)$

White Correlation

(*) when we have heterodescedasticity, should we use white correlation or conventional t-stat?

(*) your stat will not change much if you use conventional t-stat. (only second digit)

(*) long-run variance will lead to smaller t-stat in absolute value, so you should not use conventional in serial correlation

(*) Source of heterodescedasticity is σ_i^2, $E x_i^2 \pm \sigma_x^2$

(*) Lesson learned: Do not ignore serial correlation

by wed. next monday test hint

one hour exam, and exam would be hard - if know everything in 15' - There would be Bonus questions as well

in time series heterodescerasdicity would be serious.

- in finance second or minute data would make hetero. problem
- in macro Econ also it becomes important

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

$y_t = \alpha + \beta x_{t-1} + u_t$

$x_t = a + \rho x_{t-1} + \varepsilon_t$

$u_t = \delta \varepsilon_t + \epsilon_t$

$\rho_T = 1 - \frac{c}{T}$

$T(\hat\beta - \beta) \xrightarrow{d} \dots$ Complicated

Back to main issue: (Pure math & stat)

$y_i = \alpha + \beta x_i + u_i$

$Y = Z\gamma + u$

$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{bmatrix}_{N\times 1} \quad Z = \begin{bmatrix} 1 & x_1 \\ \vdots & x_2 \\ \vdots & \vdots \\ 1 & x_T \end{bmatrix}_{N\times 2} \quad \gamma = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$

$\hat\gamma = (Z'Z)^{-1}Z'y = (Z'Z)^{-1}Z'(Z\gamma + u)$

$= \gamma + \underbrace{(Z'Z)^{-1}}_{2\times 2}Z'u$

$Z'Z = \begin{bmatrix} \sum_{i=1}^N 1 & \sum_{i=1}^N x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix}$

$Z'u = \begin{bmatrix} \sum u_i \\ \sum u_i x_i \end{bmatrix}$

$\hat\gamma - \gamma = \left(\frac{Z'Z}{n}\right)^{-1}\left(\frac{Z'u}{n}\right)$

$\frac{Z'u}{n} = \begin{bmatrix} \frac{1}{n}\sum u_i \\ \frac{1}{n}\sum u_i x_i \end{bmatrix}$

1. $E\frac{Z'u}{n} = 0$
2. $E\left(\frac{Z'u}{n}\frac{u'Z}{n}\right) = E\begin{bmatrix} \sum u_i \\ \sum u_i x_i \end{bmatrix}\begin{bmatrix} \sum u_i & \sum x_i u_i \end{bmatrix}$

$= E\begin{bmatrix} (\sum u_i)^2 & (\sum u_i)(\sum u_i x_i) \\ (\sum u_i)(\sum u_i x_i) & (\sum u_i x_i)^2 \end{bmatrix} = \begin{bmatrix} n\sigma_u^2 & 0 \\ 0 & n\sigma_u^2\sigma_x^2 \end{bmatrix}$

3. L-L

$\frac{Z'u}{n} \xrightarrow{d} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{\sigma_u^2}{n} & 0 \\ 0 & \frac{\sigma_u^2\sigma_x^2}{n} \end{pmatrix}\right)$

$\Rightarrow \frac{Z'u}{\sqrt{n}} \xrightarrow{d} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} \sigma_u^2 & 0 \\ 0 & \sigma_u^2\sigma_x^2 \end{bmatrix}\right)$

$\operatorname{plim}\left(\frac{Z'Z}{n}\right)^{-1} = \left(\operatorname{plim}\frac{Z'Z}{n}\right)^{-1}$

$\operatorname{plim}\frac{Z'Z}{n} = \begin{bmatrix} 1 & \mu_x \\ \mu_x & \sigma_x^2 + \mu_x^2 \end{bmatrix}$

$\operatorname{plim}\frac{1}{n}\sum x_i = \mu_x$

⑦

$$\begin{bmatrix} 1 & \mu_x \\ \mu_x & \sigma_x^2 + \mu_x^2 \end{bmatrix}^{-1} = \frac{1}{\sigma_x^2} \begin{bmatrix} & ? & \end{bmatrix}$$

$= Q_2^{-1}$, let say

- "maple" ⟶ gives you inverse
- "MuPAD"
- SWP : scientific work place → derives automatically
  Calculates limiting distribution

LATEX

~~Beamer~~ ~~p~~

(*) Beamer presentation → fancy presentation

software

in finance → LATEX, stiff competition

$$\sqrt{n}(\hat\beta - \beta) = \left(\frac{Z'Z}{n}\right)^{-1}\left(\frac{Z'u}{\sqrt{n}}\right) \longrightarrow N\left(\begin{pmatrix}0\\0\end{pmatrix}, Q_2^{-1}\Sigma Q_2^{-1}\right)$$

no traspose on second part since it is Symmetric

Constant made it easy

Economics Cont.

① $y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + u_i$

② $y_i = \alpha + \beta_1 x_{1i} + e_i$

if $E(x_{1i} x_{2i}) \neq 0$ typically

① is true DGP (Data generating process) then $E(x_{1i} \cdot e_i) \neq 0$

true model = ↵ inconsistent estimate

$\hat\beta_1$ from ② $\not\xrightarrow{p} \beta_1$

always this happens.

Reason: intrinsic

y_i: growth rate x_i: political policy stability

in political science is reverse

⟹ interchanging

do not interpret x_i as exogeneous, it is always indogeneous

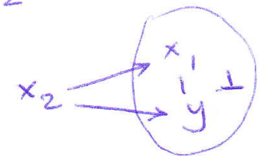
⑧

all regressors are indogeneous.

→ all variable you are using is endogeneous

run y_i on x_{1i}

x_2 → true determinant variable



when $\beta_1 = 0$
$\beta_2 \neq 0$

missing variable case.

empirical study.

- Nobody read tables in dissertation
- people see picture - important figure that explains everything.

(*) if you have multiple variables not 3d but 2D

How?

$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + u$

idempotent matrix

① run $y_i = a + b_1 x_{1i} + y^*_{1i}$

$x_{2i} = a_2 + b_2 x_{2i} + x^*_{2i}$

$\hat{y}^*_{1i} = \alpha + \beta_2 \hat{x}^*_{2i} + u_i^*$ ③

in ③ $\hat\alpha, \hat\beta_2 = \hat\alpha, \hat\beta_2$ in ①: $y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + u_i$

$\hat b_1 = (x_1' x_1)^{-1} x_1' y$

$y^*_1 = y_1 - x_1 \hat b_1 = y - x_1(x_1'x_1)^{-1}x_1'y$

$= [I - x_1(x_1'x_1)^{-1}x_1']y$

M_{x_1}

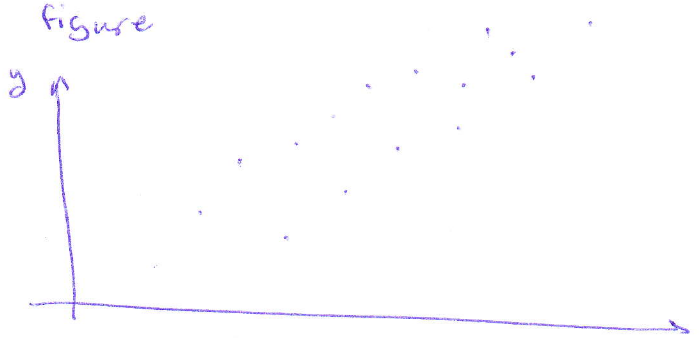
$(I - x_1(x_1'x_1)^{-1}x_1')x_1 = 0$

From ① $y = \alpha + \beta_1 x_1 + \beta_2 x_2 + u$

$M_{x_1} y = M_{x_1}\alpha + \beta \overset{0}{M_{x_1} x_1} + \beta_2 M_{x_1} x_2 + M_{x_1} u$

⟹ $y^* = \alpha + \beta_2 x_2^* + u^*$

as a result you can use the following figure



$x^* = (x_2, x_1, \ldots)$

(*) you can simply do this for multiple Variable and get x_1^*. means having Controled by other Variables.

- Control Variables

- business cares about real world and not fiction novel economic theory

$\underset{\theta}{\arg\max}\ u(x, y, \theta)$

F.O.C $\frac{\partial u}{\partial \theta} = \frac{\text{linear}}{y_i = f(x_i)}$

$y_i = \alpha + \beta x_i + u_i$

$y_i = f(x_i) = f(x_0) + \frac{\partial f}{\partial x}(x_i - x_0) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(x - x_0)^2 + \ldots = \alpha + \beta x_i + u_i$

1. approximation Error
2. Expectation Error
3. misspecification Error

- make your model simple, otherwise it takes forever; physicians use this, and gets 100 year

- cetris paribus — others the same

partial Equilibrium

$x_i \to$ endogenous

$\alpha + \beta x_i + Z\gamma + u_i$

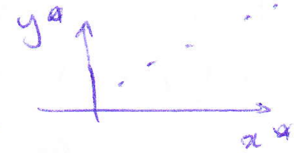
↳ Control variable

lots of variable

You don't care some are zero, or significant, since it is redundant, you don't care, and you don't criticise other people

Z

$z_1, z_2, \ldots z_k$



then you need to plot

(*) takes two weeks easily write your dissertation - empirical papers

(*) each variable one paper, since there is no theory, writing paper would be piece of cake; you run regression and write fiction novel; lucky, and good writer leads publish in Good journal

Instrument Variable $\boxed{IV}$

previous one had Control variable, but many people do not have it.

$y = \beta X + u \qquad E\,xu \neq 0$

in stat there is no instrument variable, yet in Econometrics we have

$E(ZX) \neq 0$

$E(Z \cdot u) = 0$

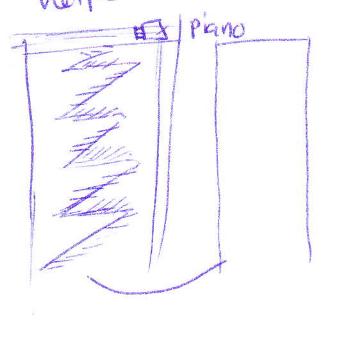
Z : IV

$\hat{\beta}_{IV} = (Z'X)^{-1} Z'y = (Z'X)^{-1} Z'(X\beta + u) = (Z'X)^{-1}(Z'X)\beta + (Z'X)(Z'u)$

$\hat{\beta}_{iv} = \beta + (Z'X)^{-1}(Z'u)$

$\text{plim}(\hat{\beta}_{IV} - \beta) = \text{plim}\left(\frac{Z'X}{n}\right)^{-1}\left(\frac{Z'u}{n}\right) = 0$

they try to find one nice instrument variable that helps to solve this problem.



- location of river (geographic, agricultural) for instrumental Variable
- instrumental Variable wrong
- Journal of applied Econometrics

- empirical micro Econ tries to find instrument var → Geographic, yet again everything in the world is Endogeneous, even location, and

why people ~~be~~come Texas

- You need to be really carefule

$\boxed{\text{Econometrics}}$ $\boxed{\text{Feb 25}}$ ①

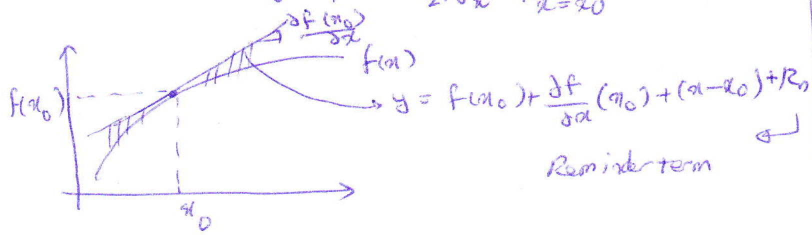
Delta method (stochastic Taylor Expansion)

Taylor Expansion

$y = f(x)$

Evaluate $f(x)$ at point x_0

$f(x_0) = f(x) + \frac{\partial f}{\partial x}\Big|_{x=x_0}(x-x_0) + \frac{\partial^2 f}{2\cdot\partial x^2}\Big|_{x=x_0}(x-x_0)^2 + \dots$



$x_0 = \mu_x$ in stoch. we replace with mean
$= E(x)$

$\sqrt{n}(x_n - \mu_x) \xrightarrow{d} N(0, \sigma_x^2)$

such as $x_n = \frac{1}{n}\sum x_i$

$x_n - \mu_x = O_p\left(\frac{1}{\sqrt{n}}\right)$

$(x_n - \mu_x)^2 = O_p\left(\frac{1}{n}\right)$

$\boxed{\text{usage}}$ $y_i = \alpha + \beta x_i + \gamma z_i + u_i$

$\delta = \beta/\gamma$, you want to estimate $\hat{\delta}$

$\hat{\delta} = \hat{\beta}/\hat{\gamma}$?

$$\left.\begin{matrix}\sqrt{n}(\hat{\beta}-\beta) \\ \sqrt{n}(\hat{\gamma}-\gamma)\end{matrix}\right] \xrightarrow{d} N\left(\begin{bmatrix}0\\0\end{bmatrix}, \Sigma\right)$$

$$\Sigma = \begin{bmatrix}\sigma_\beta^2 & \sigma_{\beta\gamma} \\ \sigma_{\beta\gamma} & \sigma_\gamma^2\end{bmatrix}$$

$\hat{\delta} = \frac{\hat{\beta}}{\hat{\gamma}} = \frac{\beta}{\gamma} + \frac{1}{\gamma}(\hat{\beta}-\beta) \dots$

derivative with respect to mean

$\hat{\delta} = \frac{\hat{\beta}}{\hat{\gamma}} = \beta/\gamma + \frac{1}{\gamma}(\hat{\beta}-\beta) - \beta/\gamma^2(\hat{\gamma}-\gamma) + 0 \times (\beta-\hat{\beta})^2$

$+ \frac{1}{2}\times\frac{1}{2}\beta \frac{1}{\gamma^3}(\hat{\gamma}-\gamma)^2$ ⊛

$\hat{\delta}/\partial\hat{\beta} = \frac{1}{\hat{\gamma}}$ $\quad \frac{\partial\hat{\delta}}{\partial\hat{\beta}}(\beta, \gamma) = \frac{1}{\gamma}$ $\quad \frac{\partial\hat{\delta}}{\partial\hat{\gamma}} = \frac{-\hat{\beta}}{\gamma^2}$

since $(\hat{\beta}-\beta) = O_p(1/\sqrt{n})$

$(\hat{\beta}-\beta)^2 = O_p\left(\frac{1}{n}\right)$, $(\hat{\gamma}-\gamma)^2 = O_p\left(\frac{1}{n}\right)$

term ⊛ is not important

②

$\hat{\delta} = \delta + \frac{1}{\gamma}(\hat{\beta}-\beta) - \beta/\gamma^2(\hat{\gamma}-\gamma) + O_p\left(\frac{1}{n}\right)$

where $\delta = \beta/\gamma$

$\sqrt{n}(\hat{\delta}-\delta) = \frac{1}{\gamma}\sqrt{n}(\hat{\beta}-\beta) - \beta/\gamma^2\sqrt{n}(\hat{\gamma}-\gamma) + O_p\left(\frac{1}{\sqrt{n}}\right)$

- you run the first reg - then you add this term, and get t-stat and publish paper $140 dinner

- Once your write $O_p\left(\frac{1}{\sqrt{n}}\right)$ you will not recieve econometrics comment on your paper - regarding Econometrics

$$\underset{1\times 2}{\begin{bmatrix}\frac{1}{\gamma} & -\beta/\gamma^2\end{bmatrix}} \underset{2\times 1}{\begin{bmatrix}\sqrt{n}(\hat{\beta}-\beta) \\ \sqrt{n}(\hat{\gamma}-\gamma)\end{bmatrix}} \xrightarrow{d} N\left(\begin{bmatrix}0\\0\end{bmatrix}, \underset{1\times2}{R}\,\underset{2\times2}{\Sigma}\,\underset{2\times}{R'}\right)$$

$\downarrow$
R

you don't know γ and β, you just know $\hat{\beta}$, $\hat{\gamma}$, $\hat{\Sigma}$.

$\hat{\beta} = \beta + O_p\left(\frac{1}{\sqrt{n}}\right)$ Econometrics is approximation so you can put this usin O_p

$\frac{1}{\hat{\beta}} = \frac{1}{\beta} + ? \rightarrow O_p\left(\frac{1}{\sqrt{n}}\right)$

use $\boxed{\text{prove it yourself}}$

$1/\hat{\beta} = \frac{1}{\hat{\beta}-\beta+\beta} = \frac{\alpha\ 1}{\hat{\beta}-\beta} + \frac{1}{\beta}$

$\downarrow$
$O_p(1/\sqrt{n})$

Ref $\hat{\Sigma} = \Sigma + O_p(1/\sqrt{n})$

$\hat{\Sigma}^{-1} = \Sigma^{-1} + O_p(1/\sqrt{n})$

$y_i = a + \beta x_i + u_i$

$\sqrt{n}(\hat{\beta}-\beta) = \frac{\sum x_i u_i}{\sum \tilde{x}_i^2} \cdot \frac{\frac{1}{n}\sum x_i u_i}{\sum \tilde{x}_i^2}$

Good term $\xrightarrow{d} N(0, \sigma^2)$

nuizance term

$\downarrow$

$\frac{1}{n}\sum x_i \sum u_i$

$\| $

$\frac{1}{\sqrt{n}}\sum x_i \frac{1}{\sqrt{n}}\sum u_i$

$\downarrow$ $O_p(1)$ $\quad \downarrow$ $O_p(1)$

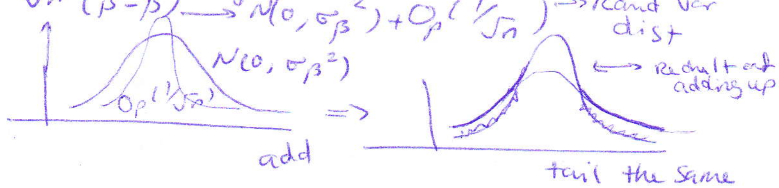
$x_i = a + x_i^\circ$, $x_i^\circ \sim d(0, \sigma_x^2)$

$\tilde{x}_i = x_i - \frac{1}{n}\sum x_i = x_i^\circ - \frac{1}{n}\sum x_i^2$

Since $x_n = O_p\left(\frac{1}{\sqrt{n}}\right)$

$\sqrt{n}\,\frac{1}{n}\sum x_i = \frac{1}{\sqrt{n}}\sum x_i = O_p(1)$

$\sqrt{n}(\hat{\beta}-\beta) \xrightarrow{d} N(0, \sigma_\beta^2) + O_p(1/\sqrt{n}) \rightarrow$ Rand var dist



③

You need to get rid of the $O_p(\frac{1}{\sqrt{n}})$

the name ~~of~~ method is: "Bootstrapping":

- you must program it yourself

$\boxed{\text{GLS : Generalized Least Square}}$

- Better due to smaller variance but in practice not efficient (Smaller variance)

- on the panel data we will discuss about it

We will now discuss NLS

$\boxed{\text{NonLiner Least square}}$ $\boxed{\text{NLS}}$

$y_t = f(x_t) + u_t = g(x_t, u_t)$ u_t inside func Complicated

$y_t = f(x_0) + \frac{\partial f}{\partial x}(x - x_0) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(x - x_0)^2 + R_n$

- most people first order, and many people (JPE, JER), say you must use the second term as well, and for this you need to use

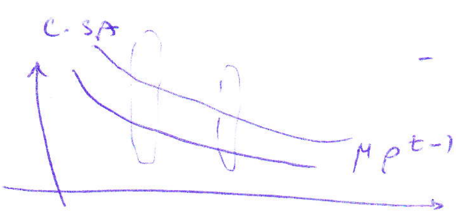
- non linear → initial var

- $y_t = \mu \rho^{t-1} + \varepsilon_t$ → dicate rate

$\frac{1}{n} \sum y_{it}$

each indiv @ t. (public good Game)

Cross sectiontional Average



- if Confidence interval used, comparison would be difficult

Round

μ: unknown

ρ: unknown

- example of non liner estimator you need Delta method, to derive limiting dist.

$\underset{\mu, \beta}{\arg\min} \sum_{t=1}^{T} \varepsilon_t^2$

$\varepsilon_t^2 = (y_t - \mu\rho^{t-1})$

put μ and increase

0.01 1

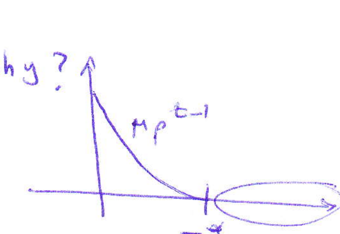
ρ 0.01 . . . 0.999 find value of μ, ρ

$\hat{\mu}_{NLS}$ $\hat{\rho}_{NLS}$ - Computer does Calculation

④

the variation b/w zero & one is for this example

$\hat{\rho}_{NLS}$ → inconsistent why?



$T \to \infty$ $y_t \to \varepsilon_t$

t is large

$y_t = \mu + \alpha\rho^{t-1} + \varepsilon_t$

have $\rho = 1$ known

limiting dist $\hat{\alpha} - \alpha = \dfrac{\sum \rho^{t-1}\varepsilon_t \to O(1)}{\sum \rho^{2t-2} \to O(1)}$

$(\sum \rho^{t-1}\varepsilon_t)^2 = O(1)$

$\Rightarrow \hat{\alpha} - \alpha \to O_p(1)$

- the variance as $n \to \infty$ does not Converge and does not shrink (assumption of Lindberg levy)

$\Rightarrow \alpha$ is not identifiable

$y = f(\alpha, x) + u_t$ Consistency Condition

$\frac{\partial f}{\partial \alpha} \Rightarrow \frac{1}{T}\sum\left[\frac{\partial f}{\partial \alpha}\right]^2\Big|_{\alpha=\alpha_0} \to M$

as $T \to \infty$

if $\alpha \neq \alpha_0$ where α_0 is true value

then $\frac{1}{T}\sum\left[\frac{\partial f}{\partial \alpha}\right]^2\Big|_{\alpha \neq \alpha_0} \to \infty$

f Could have been $f(\alpha) = \ln(x + 1/\alpha - \alpha x^2)$

$e^{-\alpha x}$

$\Rightarrow$ then $\sqrt{n}(\hat{\alpha}_{NLS} - \alpha) \xrightarrow{d} N(0, \sigma_\alpha^2)$

σ_α^2

$y = \frac{\partial f}{\partial \alpha}(\alpha - \alpha_0) + \underbrace{R_n}_{e_t}$ tailor Expansion

$\frac{\partial f}{\partial \alpha}$ $\sigma_e^2 \left(\frac{\partial f}{\partial \alpha}\right)^2 = \sigma_\alpha^2$ ☆ σ_e^2 ?

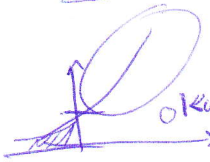
$E\left(\frac{1}{T}\sum e_t^2\right) = \sigma_e^2$

- limiting dist is Simple, and most of Cases you don't know

- whether it is Consistent or not should be analyzed by you, rather than Stat pckg

$y_t = \alpha + \beta e^{-\gamma x_i} + u_t$ → Not identifiable

$$y_t = \alpha + \beta e^{-\gamma x_i} + u_t \quad \gamma \to 0$$



okay

$e^{-2x_i^2}$

(*) non linear Reg: You need to be worried about identification

$\sum e^{-2x_t} \to O(1)$ and not to infinity

non linear function → is Complicated

Recommendation: don't use it

insted use tailor Expansion

why not crude

$$y_t = \alpha + \beta x_t + \gamma x_t^2 + u_t$$

$\beta \to \frac{\partial y}{\partial x}$ $\quad \gamma \to \frac{1}{2}\frac{\partial^2 y}{\partial x^2}$

Econometrics | March 04 | ①

- from today panel data & first difference IV
- makeup - takehome - Consult classmate, but don't copy
- Study yourself - step by step derivation
  tommorow afternoon - 20 pages of makeup

- panel data

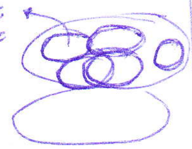
- treatment Effect
- basic stuff, and rest you pick up by yourself

~~pages~~ always you can write like this - seasonal, these days is trend

y_{it} ← y_{it}
household/Nation state — time

how multi nation

R_i race across people
T_{it}
P_{it}



Example: testscore in school districts

→ Control Variable

$T_i = \alpha + \beta P_i + Z_i \gamma + u_i$ ⇒ cross sec reg

$\beta > 0$ ↳ parent's income

$T_{it} = \alpha_i + \beta P_{it} + Z_{it}\Gamma + u_{it}$ ⇒ panel Reg (5 years)

$\beta < 0$

Pool panel Regression

↳ what does it mean?

when time series $\beta \to \beta_i \to$ (Group mean Estimator)

* if you pool, you will get general β, showing general trend

- as GDP increase test score drops

i individual Specific (Idiosyncratic

individual Specific | t

$y_i^o = \alpha + \beta x_i^o + u_i^o$

$y_t^o = \alpha + \beta_i x_t^o + u_t^o$

$y_{it}^o = \alpha + \beta x_{it}^o + u_{it}^o$ — for each t, for each i ↳ TSR

$y_{it} = \cancel{x} \; g_{it}^y + y_{it}^o$

Common — idiosyncratic

$x_{it} = g_{it}^x + x_{it}^o \longrightarrow \beta^G$

$g_{it}^y = \alpha + \beta g_{it}^x + u_{it}$

$DS_{it} = \alpha_i + \beta R_{it-1} + u_{it}$

↳ stock price — ↳ individual return

For each CSR → β^i

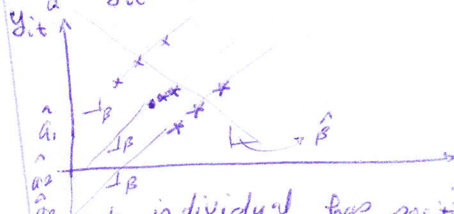
For each TSR → β^t

②

- Currently these are strivers
  probably in 20 Yrs these will be used
- 50 years Gap

Basic panel

$\hat{\alpha}$ y_{it} x_{it} $t = 1, 2, 3$



each individual has positive relationship, but overall it has negative relationship

① $y_{it} = \alpha + \beta x_{it} + u_{it} \Rightarrow \beta < 0$

$$\begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{21} \\ y_{22} \\ y_{23} \\ y_{31} \\ y_{32} \\ y_{33} \end{bmatrix} = \alpha \begin{bmatrix} 1 \\ 1 \\ \vdots \\ \vdots \\ 1 \end{bmatrix} + \beta \begin{bmatrix} x_{11} \\ x_{12} \\ \vdots \\ \vdots \\ x_{33} \end{bmatrix} + \begin{bmatrix} u_{11} \\ u_{12} \\ \vdots \\ \vdots \\ u_{33} \end{bmatrix}$$

$\hat{\beta} = ?$

② $\frac{1}{N}\frac{1}{T}\sum_{i=1}^{N}\sum_{t=1}^{T} y_{it} = \alpha + \beta \frac{1}{N}\frac{1}{T}\sum_{i=1}^{N}\sum_{t=1}^{T} x_{it} + \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} u_{it}$

①-② $\Rightarrow \underbrace{(y_{it} - \frac{1}{N}\frac{1}{T}\sum^N\sum^T y_{it})}_{\tilde{\tilde{y}}_{it}} = \beta \underbrace{(x_{it} - \frac{1}{N}\frac{1}{T}\sum^N\sum^T x_{it})}_{\tilde{\tilde{x}}_{it}} + \underbrace{u_{it} - \frac{1}{NT}\sum^N\sum^T u_{it}}_{\tilde{\tilde{u}}_{it}}$

$\hat{\beta} = \frac{\sum^{NT} \tilde{\tilde{y}}_{it}\tilde{\tilde{x}}_{it}}{\sum^{NT} \tilde{\tilde{x}}_{it}}$ Pooled ordinary Least Square POLS Estimator

- This is wrong estimator, since you need to get for each person, but if you put one parameter you will get this amount
- if you put different intercept, but same slope you will have positive relation
- for time series regression you get positive relationship
  Cross sectional is negative

Cross sectional or time series?

* Cross sectional
  think about economics
- Cross section: general difference discription
- Time series: each individual nuisance relationship and Control

→ in cross section $x_i \uparrow \Rightarrow y_i \uparrow$ by one does not have meaning, since exogen can not change

③

- $\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3$ was individual fixed effect

how calc analyt? - Dummy Constant
- keep slope coeff.

$y_{it} = a_i + \beta x_{it} + u_{it}$ ②

- Fixed Effect Regression
- Least square Dummy variable Regression (LSDV)
- within group regression (WG)

$\frac{1}{T}\sum y_{it} = a_i + \beta \frac{1}{T}\sum x_{it} + \frac{1}{T}\sum u_{it}$ ②′

② - ②′: $(y_{it} - \frac{1}{T}\sum y_{it}) = \beta(x_{it} - \frac{1}{T}\sum x_{it}) + (u_{it} - \frac{1}{T}\sum u_{it})$

$$\hat{\beta}_{WG} = \hat{\beta}_{FE} = \hat{\beta}_{LSDV} = \frac{\sum^N\sum^T (y_{it} - \frac{1}{T}\sum y_{it})(x_{it} - \frac{1}{T}\sum x_{it})}{\sum^N\sum^T (x_{it} - \frac{1}{T}\sum x_{it})^2}$$

* FE Command in Stata gives you this, but you have to sacrifice lots of degree of freedom

N: total number of individual

$N \times T$: total # of samples → POLS

$N \times (T-1)$: " → FE (fixed Effect)

marriage → 2 times don't have lots of data

Birth weights → (+) health rate; (−) mortality health

bias: due to cross sectional, but they want to have Causal relationship

→ solving problem for Birth weighs and mortality health you need to do Twin Study.

Control everything, they found Zero Correlation

Random effects

assume they exist, like moving piano

$y_{it} = a_i + \beta x_{it} + u_{it}$

$= a + \beta x_{it} + \underbrace{(a_i - a) + u_{it}}_{e_{it}}$

a is mean of a_i

$= a + \beta x_{it} + e_{it}$ serially correlate, "GLS"

OLS was for pooled ordinary OLS

→ Random effect & pooled OLS almost identical

④

-- There is test that tells us whether GLS is right or not (you can cook up your result using some tests: art of lying) Robustness

Time of asymptotics

Asymptotic Theory.

$u_{it} \sim iid(0, \sigma^2)$

$i = 1, \dots, N$

$t = 1, \dots, T$

$\frac{1}{T}\frac{1}{N}\sum^N\sum^T u_{it} = u_{NT}$, let say

① $E \frac{1}{TN}\sum^{NT} u_{it} = 0$

② $E(\frac{1}{TN}\sum^{NT} u_{it})^2 = \frac{\sigma^2}{NT}$

③ L-L $(U_{NT} - 0) \xrightarrow{d} N(0, \frac{\sigma^2}{NT})$ ← wrong Expression

$\Rightarrow \sqrt{NT}(U_{NT} - 0) \xrightarrow{d} (0, \sigma^2)$

pooled OLS:

$\hat{\beta}_{POLS} = \beta + \frac{\sum^{NT}\tilde{x}_{it}\tilde{u}_{it}}{\sum^{NT}\tilde{x}_{it}^2}$

$\frac{1}{NT}\sum^{NT}\tilde{x}_{it}\tilde{u}_{it}$ = Good + nuisance $= \frac{1}{NT}\sum^{NT} x^+_{it} u^+_{it}$

$-(\frac{1}{NT})^2(\sum^{NT} x^+_{it})(\sum^{NT} u^+_{it})$ $\frac{1}{NT}\sum u^+_{it} \to O_p(\frac{1}{\sqrt{NT}})$

$\to \frac{1}{NT}\sum x^+_{it}$ $O_p(1/\sqrt{N})$ $O_p(\frac{1}{N\sqrt{T}})\tilde{x}_{it}$

on test 2 you had to find this - look at lecture nots

$x_{it} = \alpha^x_i + x^o_{it}$, $x^o_{it} \sim d(0, \sigma_x^2)$

$x_{it} - \frac{1}{N}\sum^N \frac{1}{T}\sum x_{it}$ (α^x_N)

$\frac{1}{NT}\sum\sum x_{it} = \frac{1}{N}\sum^N \alpha^x_i + \frac{1}{NT}\sum^N\sum^T x^o_{it}$

$x_{it} - \frac{1}{NT}\sum^N\sum^T x_{it} = (\alpha^x_i - \alpha^x_N) + (x^o_{it} - \frac{1}{NT}\sum x^o_{it})$

$\tilde{x}_{it}$ x^+_{it}

$= \underbrace{(\alpha^x_i - \alpha^x_N) + x^o_{it}}_{} - \frac{1}{NT}\sum^{NT} x^o_{it}$

$\tilde{x}_{it} = x^+_{it} - \frac{1}{NT}\sum^{NT} x^o_{it}$

$\tilde{u}_{it} = u^+_{it} - \frac{1}{NT}\sum^{NT} u_{it}$

$\hat{u}_{it}$

$x_{it} \sim iid(\alpha, \sigma_i^2)$

$\frac{1}{NT}\sum^N\sum^T x_{it} = O_p(1)$ since $\alpha \neq 0$

if $\alpha = 0$, $\frac{1}{NT}\sum^N\sum^T x_{it} = O_p(\frac{1}{\sqrt{NT}})$

$\xrightarrow{d} N(0, \frac{\sigma_x^2}{NT})$

Econometrics March 04 ⑤

$\sqrt{NT}\left(\frac{1}{NT}\sum^N\sum^T x_{it}\right) \xrightarrow{d} N(0, \sigma_x^2)$
$= O_p(1)$

$\rightarrow O_p\left(\frac{1}{\sqrt{NT}}\right)$

$\frac{1}{NT}\sum\sum x_{it}^+ = \frac{1}{NT}\sum^N\sum^T(\alpha_i^x - \alpha_N^x) + \frac{1}{NT}\sum^N\sum^T x_{it}^o$

$= O_p\left(\frac{1}{\sqrt{N}}\right) + O_p\left(\frac{1}{\sqrt{NT}}\right) = O_p\left(\frac{1}{\sqrt{N}}\right)$

$\xi_{it} = x_{it}^+ \cdot u_{it}$ let say

$E\,\xi_{it} = 0 \qquad V(\xi_{it}) = \sigma_x^2\sigma_u^2$

$\frac{1}{NT}\sum^{NT} x_{it}^+ u_{it}$

$\frac{1}{\sqrt{NT}}\sum^{NT} x_{it}^+ u_{it} \xrightarrow{d} N(0, \sigma_x^2\sigma_u^2)$

$\operatorname{plim}\frac{1}{NT}\sum^{NT}\tilde{x}_{it} = \sigma_x^2$ ⊛

$\sqrt{NT}(\hat{\beta}_{POLS} - \beta) = \frac{\frac{1}{\sqrt{NT}}\sum^{NT} x_{it}^+ u_{it}}{\frac{1}{\sqrt{NT}}\sum^{NT}\tilde{x}_{it}^2}$ (I)

$- \frac{\frac{1}{\sqrt{NT}}\left(\sum^{NT} x_{it}^+\right)\frac{1}{NT}\sum u_{it}}{\frac{1}{NT}\sum^{NT}\tilde{x}_{it}^2}$ (II)

$\mathrm{I} \Rightarrow \xrightarrow{d} N\left(0, \frac{\cancel{\sigma_x^2}\sigma_u^2}{\cancel{\sigma_x^2}\sigma_x^2}\right)$

$\mathrm{II} \Rightarrow = \frac{\sqrt{NT}\left(\frac{1}{NT}\sum^{NT} x_{it}^+\right)\left(\frac{1}{NT}\sum u_{it}\right)}{\frac{1}{NT}\sum^{NT}\tilde{x}_i^2} = \frac{\sqrt{NT}\,O_p\left(\frac{1}{N\sqrt{T}}\right)}{O_p(1)}$

$= O_p\left(\frac{1}{\sqrt{N}}\right)$ since mean is constant from ⊛

$\sqrt{NT}(\hat{\beta}_{POLS} - \beta) = O_p(1) + O_p\left(\frac{1}{\sqrt{N}}\right) \xrightarrow{d} N\left(0, \frac{\sigma_u^2}{\sigma_x^2}\right)$

as $N, T \to \infty$

$x_{it} = \alpha_i^+ + x_{it}^o, \quad x_{it}^o \sim d(0, \sigma_x^2)$

$\hat{\beta}_{FE} - \beta = \frac{\frac{1}{NT}\sum^N\sum^T\tilde{x}_{it}\tilde{u}_{it}}{\frac{1}{NT}\sum^N\sum^T\tilde{x}_{it}^2}$

$\tilde{x}_{it} = x_{it} - \frac{1}{T}\sum^T x_{it}$

$x_{it} = \alpha_i^x + x_{it}^o$

$\frac{1}{T}\sum x_{it} = \alpha_i^x + \frac{1}{T}\sum x_{it}^o$

$\tilde{x}_{it} = x_{it}^o - \frac{1}{T}\sum x_{it}^o = \tilde{x}_{it}^o$

⑥

$\hat{\beta}_{FE} - \beta = \frac{\frac{1}{NT}\sum x_{it}^o u_{it}}{\frac{1}{NT}\sum(x_{it}^o)^2} - \frac{\left(\frac{1}{NT}\sum^{NT} x_{it}^o\right)\left(\frac{1}{NT}\sum^{NT} u_{it}\right)}{\frac{1}{NT}\sum^{NT}(\tilde{x}_{it}^o)^2}$

$\Downarrow$ I $\qquad O_p\left(\frac{1}{NT}\right)$

$\frac{1}{NT}\sum x_{it}^o = O_p\left(\frac{1}{\sqrt{NT}}\right)$

$\sqrt{NT}(\hat{\beta}_{FE} - \beta) = O_p(1) + O_p\left(\frac{1}{\sqrt{NT}}\right) \xrightarrow{d} N\left(0, \frac{\sigma_u^2}{\sigma_x^2}\right)$

Shrink of variance compared to POLS

- Study lecture notes, and study by yourself to solve the takehome exam

makeup

March 18 | Econometrics | ①

POLS, Random Effect

① $y_{it} = a_i + \beta x_{it} + u_{it}$

② $x_{it} = \alpha_i + x^{o}_{it}$ $x^{o}_{it} \sim iid(0, \sigma^2)$

Appendix I

Consistency : as $n \to \infty$ bias $= 0$

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

$E(\hat\rho - \rho) = \frac{-2\rho}{T}$ → means it is biased but

$\underset{T\to\infty}{plim}(\hat\rho - \rho) = 0$ ≡ { $\lim_{T\to\infty} E(\hat\rho-\rho)$ & $E(\hat\rho-\rho)^2 = c$ (constant)

Biased but Consistent

$N >> T$, $T >> N$ Could have diff result

$Y_i = \beta x_i + u_i$ $u_i \sim iid\ N(0, \sigma^2)$

$\frac{\sqrt{N}(\hat\beta - \beta)}{V(\hat\beta)} \sim t_n$ → Exact dist. - Critical liner theory

but many cases we do not have ① Normal dist of u_i ② iid Assump

so we will have $\frac{\sqrt{n}(\hat\beta-\beta)}{\cancel{\ }} \xrightarrow{d} N(0, (X'X)^{-1}(X'\Omega X)(X'X)^{-1})$

$u_i \sim d(0, \Omega)$

$\Rightarrow \frac{\sqrt{N}(\hat\beta-\beta)}{V(\hat\beta)} \xrightarrow{d} N(0,1)$ (*)

$N > 25$ → t and normal dist are the same

(*) $\frac{\sqrt{N}(\hat\beta-\beta)}{V(\hat\beta)} \simeq \underbrace{N(0,1)}_{\text{Good term}} + \underbrace{O_p\left(\frac{1}{\sqrt{n}}\right)}_{\text{bad term}}$ ★ if n small this approx will fail

---

main part (back). Cont

i) FE ⟹ Running ① $\hat\beta_{FE} \xrightarrow{p} \beta$

③ $y_{it} = a + \beta x_{it} + e_{it}$ → Correlate if a_i, α_i Correl

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

if $E(a_i - a)(\alpha_i - \alpha) \neq 0$ then $E\, x_{it} e_{it} \neq 0$ (*)

permanent bias (persistent)

lots of test show whether it exists or not

Stata gives ya $\hat\sigma_a^2$

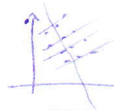
- researcher may say $\hat\sigma_a^2$ is large, so running this Reg does not make sense; since if $\frac{Var = 0}{\sigma_u^2}$ does not make condition ★ to satisfy; so this sentence may not make sense.
- a_i, α_i is Common factor for examply (You & Yr classmate Correlate), since Teacher is Common factor

②

a_i, α_i → firm size or firm profit

↳ time series mean, invariant part of size of firm

- The other firms size could affect your profit



Easy Takeaway Do not run 3, just run 1

④ $y_{it} = a_i + \theta_t + \beta x_{it} + u_{it}$

a_i: individual fixed effects (FE); θ_t: time fixed (common) effects (CTE)

	①	②	③	④
indep				
ed				
⋮				
FE	Yes	Yes	No	
CTE	NO	Yes	NO	
	I	II	POLS ③	

I, II: previously in Econ but now not in om MKT Not

③: the only case avail in econ

④: finance run this

- Gender study run ③
- why not 1 and 2 and just 3?

since you do not have time varying independent Variable (time invariant)

for example: ① parent income (P_i) ② IQ (IQ_i) ③ Gender (G_i)

$y_{it} = a_i + \beta x_i + u_{it}$ (4*)

$- \frac{\Sigma}{T} y_{it} = a_i + \beta x_{it} + \frac{\Sigma}{T} u_{it}$

$\tilde y_{it} = \tilde u_{it}$ → not identifiable, can't run

So they run: fifth Equation

$y_{it} = a + \beta x_{it} + \rho y_{it-1} + \varepsilon_{it}$ ⑤

since $\underbrace{y_{it}}_{\text{panel}} = \underbrace{a + \beta x_{it} + u_{it}}_{\text{Cross sec}}$ is silly ⟹ ⑤ ✓✓ is better

Question: why not run (4*) by taking time series mean? (4**)

$y_{iT} = a + \beta x_i + u_i^*$ this will become Cross sectional Regression

since people run usually run panel Reg today

Q Why (4**) better? since you are interested to estimate β and not ρ

$y_{it} = a + \beta x_i + (a_i - a) + u_{it}$

u_{it}^*

⚠ do not Fell pressure to use panel reg. if you have panel sometimes you need to run Cross Sections

(!) one Case you should not run mean of time series

18 months power usage

x_i

when data is monthly Cyclic (seasonality), the mean would be different. If annual you should not worry

on that case You need to run

Common time effect

$y_{it} = a + \theta_t + \beta x_i + u_{it}$ (6)

[TxN] CTE [TxN] [TxN] Time dummies (17 dummies for 18 month)

Data preparation in stata

$$\underset{T\times N}{\begin{bmatrix} y_{1t} & \cdots & y_{N1} \\ \vdots & \ddots & \vdots \\ y_{iT} & \cdots & y_{NT}\end{bmatrix}} \qquad \underset{T\times N}{\begin{bmatrix} x_{11} & \cdots & x_{N1} \\ \vdots & \ddots & \vdots \\ x_{1T} & \cdots & x_{NT}\end{bmatrix}} \qquad \theta_t = \begin{array}{c|c} & 2 \;\; 3 \;\; 4 \;\cdots\; 18 \\ \hline i=1 & \begin{bmatrix} 1 & 0 & \cdots \\ 0 & 1 & \\ \vdots & & \ddots \end{bmatrix}_T \\ i=2 & \begin{bmatrix} 1 & 0 & \cdots \\ \vdots & & \ddots \end{bmatrix} \\ \vdots & \vdots \\ i=N & \begin{bmatrix} 1 & 0 & \cdots \\ \vdots & & \ddots \end{bmatrix}_T \end{array}$$

$$\begin{bmatrix} y_{11}\\ y_{21}\\ \vdots \\ y_{1T} \\ y_{21} \\ \vdots \\ y_{2T} \\ \vdots \\ y_{TN}\end{bmatrix} = \begin{bmatrix}1\\ \vdots\\ \vdots \\ 1\end{bmatrix} + \begin{bmatrix} x_1 \\ \vdots \\ x_1 \\ x_2 \\ \vdots \\ x_2 \\ \vdots \\ x_N \\ \vdots \\ x_N\end{bmatrix} + \overbrace{\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\\ & \vdots & \end{bmatrix}}^{17}$$

same

(7) $y_{it} = \underbrace{a + \beta x_i}_{a_i} + \gamma Z_{it} + \theta_t + u_{it}$ most Complicated Panel Regression

Hausman Tailor 1987 → IV

it is easy, but IV we don't know where it has come from

y_{it} stock price, x_i firm characteristic, Z_{it} ...

they use a_i and run $\boxed{y_{it} = a_i + \gamma Z_{it} + \theta_t + u_{it}}$

alternative way:

(7-1) $y_{it} = a_i + \theta_t + \gamma Z_{it} + u_{it}$

$\hat\gamma_{FE}$

(7-2) $y_{it} - \hat\gamma Z_{it} = a + \beta x_i + \theta_t + u_{it}^*$ (θ_T)

$\frac{1}{T}\sum(y_{it} - \hat\gamma Z_{it}) = a^* + \beta x_i + u_{it}^*$ ← Take time series mean and run this Reg

$\hat\beta_{OLS} \xrightarrow{p} \beta$ as $NT \to \infty$

$\hat\gamma_{FE} \xrightarrow{p} \gamma$ as $N, T \to \infty$

Equation (8) more popular

(8) $y_{it} = \alpha_i + \rho y_{it-1} + \beta x_{it} + \theta_t + e_{it}$

Dynamic panel Regression

they are running:

(1) $y_{it} = \alpha_i + \beta x_{it} + e_{it}$

(2) $y_{it} = \alpha_i + \beta x_{it} + \rho y_{it-1} + e_{it}$

⚠ but if (1) is true then (2) is wrong → persistance

They report both to show their result is robust

$y_{it} = \alpha_i + \beta x_{it} + e_{it}$ (**)

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

$-\ \rho y_{it-1} = \alpha_i \rho + \beta\rho x_{it-1} + \rho e_{it-1}$

$y_{it} = \alpha_i(1-\rho) + \rho y_{it-1} + \beta(x_{it} - \rho x_{it-1}) + u_{it}$

(2)' $y_{it-1} = \alpha_i^* + \rho y_{t-1} + \beta x_t + \gamma x_{t-1} + u_t$

So if (1) is true you should run (2)' and not (2), since these two are not the same.

putting dummy variable (paper of difference in difference)

$y_{it} = \alpha_i + \gamma d_i + \beta x_{it} + e_{it}$ according above (**) Calc

$\Rightarrow y_{it} = \alpha_i(1-\rho_i) + \gamma(1-\rho_i)d_i + \rho_i y_{it-1} + \beta(x_{it} - \rho_i x_{it-1}) + u_i$

FE Regression

(0) $y_{it} = \alpha_i + \beta x_{it} + u_{it}$

(1) $y_{it} = \alpha_i + \beta_i x_{it} + u_{it}$ (*)

$\hat\beta_i \in [-2, 0]$

98% $\hat\beta_i < 0$

2% $\hat\beta_i \approx 0$

y_{it} Exchange rate

x_{it} interest rate

(0) $\hat\beta_{FE} \approx 0$

$$\hat\beta = \frac{\sum^{NT} \tilde x_{it}\tilde y_{it}}{\sum^{NT} \tilde x_{it}^2} = \frac{(\sum^{N} \tilde x_{1t}\tilde y_{1t} + \sum^{N} \tilde x_{2t}\tilde y_{2t} \cdots}{\sum \tilde x_{1t}^2 + \cdots + \sum \tilde x_{Nt}^2} \approx \hat\beta_N$$

Run (*) First and check Consistency

Group mean $\sum^{N} \frac{\hat\beta_i}{N} = -2$ more reliable if

This will give you central tendency

$\hat\beta_{\mu}$ from (i) $= 0$

⊛ if t is large enough to run time series reg, then ($t=2,3$: No, $t=5$ maybe, greater than 5)

Just cross sec.

You run for β_i

⊛ if t is large enough, just run time series

➡ The reason you run panel reg. is that you do not have time series ($t > 100$)

$\boxed{\text{POLS}}$

$$y_{it} = a_i + \beta x_i + u_{it}$$

$$y_{it} = a + \beta x_{it} + \underbrace{(a_i - a) + u_{it}}_{= e_{it}}$$

$$\hat\beta_{pols} \xrightarrow{p} \beta$$

iff $E(x_{it} e_{js}) = 0$ for all i, t, j, s

$\Leftrightarrow$ strongly exogenous

$$\hat\beta_{pols} = \beta + \frac{\sum^{NT}(x_{it} - \frac{\sum}{NT} x_{it})(e_{it} - \frac{\sum}{NT} e_{it})}{\sum^{NT}(x_{it} - \frac{\sum}{NT} x_{it})^2}$$

$\boxed{\text{Typical t-value}}$

$$t_{\hat\beta_{pols}} = \frac{\hat\beta_{pols} - \beta}{\sqrt{\hat\sigma_e^2 (X'X)^{-1}}}$$

$\therefore$ e_{it} is serially correlated in most of cases

$$E(e_{it}e_{is}) = E[(a_i - a) + u_{it}][(a_i - a) + u_{is}] = \sigma_a^2 + \underbrace{E u_{it} u_{is}}_{0}$$

$$= \begin{cases} \sigma_a^2 & \text{if } t \neq s \\ \sigma_a^2 + \sigma_u^2 & \text{if } t = s \end{cases}$$

$$E(\hat\beta - \beta)^2 \Rightarrow E\left[\sum^{NT}(x_{it} - \sum_{NT}^{NT} x_{it})(e_{it} - \sum_{NT}^{NT} e_{it})\right]^2$$

if variance are identical over t

$$= N \cdot T\, \sigma_x^2 \sigma_e^2 + E \text{ cross products}$$

$$E(x_{it} x_{is} e_{it} e_{is}) = E(x_{it} x_{is}) E(e_{it} e_{is}) = \gamma^{t-s} \sigma_x^2 \cdot \sigma_a^2 \text{ if } t \neq s$$

here $E(x_{it} x_{is}) = \gamma^{t-s} \sigma_x^2$ follow AR(1)

$$E \text{ Cross Products for each } i = \sigma_a^2 \sigma_x^2 (\overbrace{\gamma + \gamma^2 + \dots}^{T}) = \frac{\sigma_a^2 \sigma_x^2 (1 - \gamma^t)}{1 - \gamma}$$

$$= N \cdot T \cdot \sigma_x^2 \underset{\substack{\downarrow \\ \sigma_a^2 + \sigma_u^2}}{\sigma_e^2} + 2 N \sigma_a^2 \sigma_x^2 \cdot \left(\frac{1 - \gamma^T}{1 - \gamma}\right)$$

$$V(\hat\beta) = \frac{N \cdot T \sigma_x^2 (\sigma_a^2 + \sigma_u^2) + 2 N \sigma_a^2 \sigma_x^2 \frac{1 - \gamma^T}{1 - \gamma}}{(N \cdot T\, \cancel{\sigma_x^2})(N \cdot T)\sigma_x^2}$$

$$= \frac{\sigma_a^2 + \sigma_u^2 + (\sigma_a^2 / T)\left(\frac{1 - \gamma^T}{1 - \gamma}\right)}{N \cdot T \sigma_x^2}$$

$$\sqrt{NT}(\hat\beta - \beta) \xrightarrow{d} N\left(0, \frac{\sigma_a^2 + \sigma_u^2 + \frac{\sigma_a^2}{T}\left(\frac{1 - \gamma^T}{1 - \gamma}\right)}{\sigma_x^2}\right)$$

$\hat\sigma_e^2 (X'X)^{-1}$

as $N \to \infty$

more serial correlation, so second part could not be ignored.

True

$$t_{\hat\beta} = \frac{\hat\beta_{OLS} - \beta}{\sqrt{\frac{\frac{1}{NT}\left(\sum^{NT} \hat e_{it}^2 + \sum^{NT} e_{it} e_{is}\right)}{\sum^{NT} \tilde x_{it}^2}}}$$

vs.

Typical

$$= \frac{\hat\beta_{OLS} - \beta}{\sqrt{\frac{\frac{\sum^{NT} \hat e_{it}^2}{NT}}{\sum^{NT} \tilde x_{it}^2}}}$$

Remedy:

① Run Fixed Effect first

For small t you don't know serial correlation, and pray God u_{it} not serially correlated

---

$$Y = \beta X + e$$

$$\hat\beta = (X'X)^{-1} X'Y = \beta + (X'X)^{-1} X'e$$

$$E(\hat\beta - \beta)(\hat\beta - \beta)' = E[(X'X)^{-1} X'ee'X(X'X)^{-1}]$$

X is given

$$E[(\hat\beta - \beta)(\hat\beta - \beta)' | X] = (X'X)^{-1} X' E[ee'|X] X (X'X)^{-1}$$

$$= (X'X)^{-1} X' \Omega X (X'X)^{-1}$$

$\Omega = E(ee'|X)$ Cov matrix

if u_{it} ~ iid $\rightarrow \Omega_{11}$

$$\Omega = \begin{bmatrix} \Omega_{11} & \Omega_{12} & \cdots \\ T \times T & & \vdots \\ \cdot & T \times T & \cdots \\ \cdot & \cdots & \cdot \end{bmatrix}$$

$(NT) \times (NT)$

$$\Omega_{11} = \begin{bmatrix} \sigma_a^2+\sigma_u^2 & \sigma_a^2 & \cdots & \\ \sigma_a^2 & \ddots & & \vdots \\ \vdots & & \ddots & \\ & \cdots & & \sigma_a^2+\sigma_u^2 \end{bmatrix}$$

$\Omega_{12} = 0$ since they are independent

so for Ω the diagonal should be non Zero and other terms should be Zero

if every thing unknown :

$$\Omega_{11} = \begin{bmatrix} \sigma^2 & \cdots & \sigma_{T\times 1} \\ \sigma_{12} & \ddots & \vdots \\ \vdots & & \sigma^2 \\ \sigma_{1T} & \cdots & \end{bmatrix}$$

You will have $T\times(T-1)$ unknown element in this matrix

Rob (Robust) panel Regression

You will estimate it by your $N\times T$ data you have

Econometrics 25 March (1)

Generalized Least Square (GLS) Estimator

$y = X\beta + u$

let $E(uu') = \underset{(n\times T)\times(T\times n)}{\Omega}$ then $\hat\beta = (X'X)^{-1}X'y = \beta + (X'X)^{-1}X'u$

$E(\hat\beta - \beta | x) = E[(X'X)^{-1}X'u|x] = (X'X)^{-1}X'E(u) = 0$ Sandwitch

$E[(\hat\beta-\beta)(\hat\beta-\beta)'|X] = (X'X)^{-1}X'E(uu')X(X'X)^{-1} = \underbrace{(X'X)^{-1}X'\Omega X(X'X)^{-1}}$ POLS (I)

let $\Omega_\beta = (\hat{X'X})^{-1}\hat{X'\Omega X}(\hat{X'X})^{-1}, n$ then $\sqrt{n}(\hat\beta_{POLS} - \beta) \xrightarrow{d} N(0, \Omega_\beta)$

Now Consider $PP' = \Omega$ premultiplying P eigen value decomposition, choleskey decompose

$P'y = P'X\beta + P'u$ or $y^* = X^*\beta + u^*$

$\hat\beta_{GLS} = (X^{*\prime}X^*)^{-1}X^{*\prime}y^* = (X'PP'X)^{-1}X'PP'y = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y$

$\hat\beta_{GLS} = \beta + (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}u$

$E[(\hat\beta_{GLS} - \beta)(\hat\beta_{GLS} - \beta)'|X] = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}E(uu')\Omega^{-1}X$

$(X'\Omega^{-1}X)^{-1} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}\Omega\Omega^{-1}X(X'\Omega^{-1}X)^{-1} =$

$(X'\Omega^{-1}X)^{-1}X'\Omega X(X'\Omega^{-1}X)^{-1} = (X'\Omega^{-1}X)^{-1} = \Omega_{\hat\beta_{GLS}}^{-1}$ let say

Hence we have

$\sqrt{n}(\hat\beta_{GLS} - \beta) \xrightarrow{d} N(0, \Omega_{\hat\beta_{GLS}})$ Less Variance than Least Square Estimator (I)

Example. $y_{it} = \alpha_i + \beta x_{it} + u_{it}$

but you run $y_{it} = \alpha + \beta x_{it} + e_{it}$, $e_{it} = u_{it} + (\alpha_i - \alpha)$

then $E(e_{it}e_{is}) = \begin{cases} \sigma_\alpha^2 & \text{if } t \neq s \\ \sigma_\alpha^2 + \sigma_u^2 & \text{if } t = s \end{cases}$

$\Omega = \begin{bmatrix} \sigma_\alpha^2+\sigma_u^2 & & \sigma_\alpha^2 \\ & \ddots & \\ \sigma_\alpha^2 & & \sigma_\alpha^2+\sigma_u^2 \end{bmatrix}$ $PP' = \Omega^{-1}$

here u is not iid

PO(LS) Appendix I

$E(\frac{A}{B}|B) = \frac{E(A)}{B}$

* in practice we assume Ω is identity matrix due to iid assumption

$(X'X)^{-1}X'\Omega X(X'X)^{-1} >> \sigma_u^2(X'X)^{-1}$

if iid assumption not had t-stat would be large

(GLS) Appendix II

$\Omega_\beta > \Omega_{\hat\beta_{GLS}}$ efficient estimator

- if you replace $\hat\Omega_{\hat\beta_{GLS}}$ with $\Omega_{\beta_{GLS}}$ then it may not become efficient

---

Appendix III (2)

You don't know whether e_{it} and x_{it} are not correlated, don't use random effect

generally tests are weak, so never use random effect assuming no correlation

---

$y_{(n\times T)\times 1} = \begin{bmatrix} y_{11} \\ \vdots \\ y_{1T} \\ y_{21} \\ \vdots \\ y_{2T} \\ y_{N1} \\ \vdots \\ y_{NT} \end{bmatrix} = \begin{bmatrix} 1 & X_{11} \\ \vdots & \vdots \\ 1 & X_{NT} \end{bmatrix} = X_{(n\times T)\times 2}$

$\gamma = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$ $y = X\gamma + e$

$\underset{(n\times T)\times(n\times T)}{E\, ee'} = \begin{bmatrix} \Omega & & & 0 \\ & \Omega & & \\ & & \ddots & \\ 0 & & & \Omega \end{bmatrix}$

$E(e_{it}e_{js}) = 0$ Cross sectional independence

to estimate α_i fixed effect estimator

(1) $y_{it} = \alpha_i + \beta x_{it} + u_{it}$ $\hat\alpha_{i\,FE}$, $\hat\beta_{FE}$

(2) $\hat\alpha_i \Rightarrow \hat\sigma_\alpha^2 = \frac{1}{n} \sum_{i=1}^N (\hat\alpha_i^2 - \frac{1}{n}\sum\alpha_i)^2$

(3) $\hat\sigma_u = \frac{1}{T(n)} \sum^N \sum^T \hat u_{it}^2$

asymptotically $n-1 \approx n$

then get chol decomposition

$\Rightarrow \underset{\text{estimate}}{\hat\Omega^{-1}} \longrightarrow \hat P' \longrightarrow \hat\beta_{GLS}$

$\hat\beta_{FE} \neq \hat\beta_{POLS}$

Assumptions

(1) $u_{it} \sim iid$ no serial correlation

(2) u_{it} & x_{it} are not correlated

if serial correlation $\hat\Omega^{-1} \longrightarrow$ would not be consistent

$\hat\Omega \not\xrightarrow{P} \Omega$ $\hat\beta_{GLS} \not\xrightarrow{P} \beta$

$FE \longrightarrow GLS_{FE}$ $y_{it} = \alpha_i + \beta x_{it} + u_{it}$

$u_{it} = \rho u_{it-1} + \varepsilon_{it}$ $\sigma_u^2 = \sigma_e^2/(1-\rho^2)$

homogeneous mean same across i $\rho_i = \rho$

$\underset{T\times T}{E(uu')} = \begin{bmatrix} \sigma_u^2 & \rho\sigma_u^2 & \rho^2\sigma_u^2 & \cdots \\ \vdots & \ddots & & \vdots \\ & \cdots & & \sigma_u^2 \end{bmatrix}$

$$E\, u_i u_i' = \frac{\sigma_\varepsilon^2}{1-\rho^2}\begin{bmatrix} 1 & \rho & \cdots & \rho^{T-1} \\ \rho & 1 & & \vdots \\ \vdots & & \ddots & \vdots \\ \rho^{T-1} & \cdots & \cdots & 1 \end{bmatrix}_{T\times T} = \Omega_{ii} = \Omega^\circ$$

$$\Omega = \begin{bmatrix} \Omega^\circ & & 0 \\ & \ddots & \\ 0 & & \Omega^\circ \end{bmatrix}_{(n\times T)\times(n\times T)}$$

① Run $y_{it} = \alpha_i + \beta x_{it} + u_{it}$

then get $\hat{u}_{it} = y_{it} - \hat{\alpha}_i - \hat{\beta} x_{it}$

→ as long as $\underline{x_{it}}$ independent from $\underline{u_{it}}$ estimator is Consistent

$\hat{u}_{it} = \rho \hat{u}_{it-1} + \varepsilon_{it}^*$

② get $\hat{\rho}$

③ $\hat{\Omega}$ from ②

④ GLS, $\hat{P}\hat{P}' = \hat{\Omega}^{-1}$

$\hat{P} y = \hat{P} X \gamma + \hat{P} u$

Appendix IV

- This does not exist in stata, so people usually do not use it
- You need to be good programmer and not just run stata, if you want to be empirical

Appendix V

$$V(\hat{\beta}_{FE}) = \underline{(X'\hat{\Omega}X)^{-1}X'\hat{\Omega}X(X'\hat{\Omega}X)^{-1}} \cdot (n\times T)$$

2×2

$$\frac{\sqrt{n\cdot T}\,(\hat{\beta}_{FE} - \beta)}{\sqrt{V(\hat{\beta}_{FE})}} = t_\beta \xrightarrow{d} N(0,1)$$

Panel Robust [Covariance Matrix]

stata Jumps from ③ to here

clustering-thining ⚠ [assumes AR(1)]

$$V(\hat{\beta}_{FE\,GLS}) = (X'\hat{\Omega}^{-1}X)^{-1}$$

- retire & write 100 papers that are rubbish if you just know stata & not fit it to your problem, and write your program if needed

* what if $u_{it} = \rho u_{it-1} + \varepsilon_{it}$?

Ω is not known and you dont know it is ρ^2, AR(1), AR(3), …

or diagonal matrix is different?

$$\Omega_{T\times T} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1T} \\ & \ddots & & \vdots \\ & & \ddots & \vdots \\ & & & \sigma_T^2 \end{bmatrix}$$

total number of unknown

$\frac{T(T-1)}{2} + T = O(T^{-1})$

$\hat{\sigma}_{ts} = \frac{1}{n}\sum_{i=1}^{N} \hat{u}_{it}\hat{u}_{is}$ → from fixed effect

as long as cross sectionaly independent you can estimate that

$\hat{\Omega}_\beta = (X'\hat{\Omega}X)^{-1}X'\hat{\Omega}X(X'\hat{\Omega}X)^{-1}(n\wedge T)$

Panel Robust Covariance Matrix Estimator

✓ → This does not assume anything

$\underline{n\times T}$ data you have, so it: ~~n×T~~

- if ① $n > T$ Can identify

② $n \simeq T$ # of obs $= n^2 = T^2$

\# of unknown $\frac{T^2 - T}{2} + T$

but the degree of freedom shrink

③ $y_{it} = \alpha_i + \beta x_{it} + u_{it}$

$T > N \Rightarrow$ forget it

alternative: bootstrapping

you can account for cross sectional dependence in this case

Dynamic Panel Regression

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

βx_{it-1} → terrism research, account for Nicole bias

$E\, y_{it-1} u_{it} = 0$

$$E(\hat{\rho}_{FE} - \hat{\rho}) = -\frac{1+\rho}{T} + O(T^{-2})$$

This means $\hat{\beta}$ will also become inconsistent

$y_{it} = a_i + y_{it}^\circ$

$y_{it}^\circ = \rho y_{it-1}^\circ + u_{it}$

$\rho y_{it-1} = a_i \rho + \rho y_{it-1}^\circ$

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

$\alpha_i = (1-\rho)\cdot a_i$

$y_{it} - \frac{\sum y_{it}}{T} = \tilde{y}_{it} \Rightarrow \tilde{y}_{it} = \rho \tilde{y}_{it-1} + \tilde{u}_{it}$

Econometrics 25 march 5

$$\hat{\rho}_{FE} = \rho + \frac{\sum^{NT} \tilde{y}_{it-1}\tilde{u}_{it}}{\sum^{NT} \tilde{y}^2_{it-1}} \quad \begin{matrix} \Rightarrow \Delta = \frac{-(1+\rho)}{T} \\ \Rightarrow \frac{\sigma_u^2}{1-\rho^2} \end{matrix}$$

$$\operatorname*{plim}_{N\to\infty} \sum_N^{N,T} (y_{it-1} - \sum_T y_{it-1})(u_{it} - \sum_T u_{it})$$

$$= \operatorname*{plim}_{N\to\infty} \left[\sum_N^{N,T} y_{it-1} y_{it} + \sum_N^{N} (\sum_T y_{it-1})(\sum^T u_{it}) \right] = ?$$

memorize, don't solve — Test 3 — 100% — $0 = ?$

$$\sum^T y_{it-1} + \sum^T u_{it} =$$

$$(y_{i1} + y_{i2} + \dots + y_{it-1})(u_{i2} + u_{i3} + \dots + u_{iT-1} + u_{iT})$$

- always downward bias

$\rho = 0.9$ → True value

$\hat{\rho}_{FE} = 0.9 - \frac{1.9}{T}$ $T = 3 \Rightarrow \frac{1.9}{3} = 0.6$

$= 0.3$ → estimated value

$T = 100,\ N = 10{,}000 \Rightarrow$ $\boxed{\text{Bias okay for large } N}$

as long as T is large it is okay

$\hat{\rho}_{FE} = 0.9 - \frac{1.9}{100} = 0.9 - 0.019 = 0.881$

Okay? No! $\boxed{\text{limiting dist skrews up}}$

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

$$\hat{\rho}_{FE} = \frac{\sum^{NT} y_{it-1} u_{it}}{\sum^{NT} \tilde{y}^2_{it-1}} - \frac{\frac{1}{T}\sum^N (\sum^T y_{it-1})(\sum^T u_{it})}{\sum^{N,T} \tilde{y}^2_{it-1}} = \mathrm{I} + \mathrm{II} \quad (1)$$

$$\sqrt{NT}\,\mathrm{I}.\quad \frac{\frac{1}{\sqrt{NT}} \sum^{N,T} y_{it-1} u_{it}}{\sum_{NT}^{N,T} \tilde{y}^2_{it-1}} \xrightarrow{d} N(0, 1-\rho^2)$$

$$\mathrm{II}.\quad E(\mathrm{II}) = -\frac{1+\rho}{T} + O(T^{-2})$$

$$(1):\ \mathrm{I} + O(T^{-1})$$

$$\Rightarrow \sqrt{NT}(\hat{\rho}_{FE} - \rho) = \sqrt{NT}\cdot \mathrm{I} + O_p\left(\frac{\sqrt{NT}}{T}\right) \xrightarrow{d} N(0, 1-\rho^2)$$

$$+ O_p\left(\sqrt{\frac{N}{T}}\right)$$

(1) $\frac{N}{T} \to 0$ as $N, T \to \infty$ jointly

$T \gg N$

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

(2) $\frac{N}{T} \to \infty$ $\sqrt{NT}(\hat{\rho}_{FE} - \rho) \to$ No limiting distribution

t-stat will reject everything

(3) $N = T$

$N(0, 1-\rho^2) + O_p(1)$

Good news: There's solution
Bad news: You have to solve in test 3

it is not consistent ⇒ find instrument that does not correlate with error term 6

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

y_{it-2}

indep — dependent

Level — First Difference

$$y_{it} - y_{it-1} = \rho(y_{it-1} - y_{it-2}) + (u_{it} - u_{it-1})$$

Correlate

y_{it-2} → not correlated

Could be used for instrument

(1) Anderson & Hsiao
(2) Brundell & Bond

$$\beta_{IV} = (Z'X)^{-1} Z'Y$$

$$\hat{\rho}_{IV} = \frac{\sum^{N,T} (y_{it-2}(y_{it} - y_{it-1}))}{\sum^{N,T} (y_{it-2}(y_{it-1} - y_{it-2}))}$$

another instrument variable

$[y_{it-2} - y_{it-3}]$ level

$T = 2,\ T = 3$ 60%

Test 3 Question — Find limiting dist.

$T = 3$

$$\hat{\rho}_{IV_2} = \frac{\sum^{N,T} (y_{it-2} - y_{it-3})(y_{it} - y_{it-1})}{\sum^{N,T} (y_{it-2} - y_{it-3})(y_{it-1} - y_{it-2})}$$

$$= \frac{\sum^N (y_{i2} - y_{i1})(y_{i4} - y_{it})}{\sum^N (y_{i2} - y_{i1})(y_{i3} - y_{i2})} = \rho + \frac{\sum (y_{i2} - y_{i1})(u_{it} - u_{it-1})}{0}$$

- Cross sectional regression

$\boxed{\text{Test 3: (1) Derivation (2) True - False - explain — extremely hard}}$

Econometrics | Summary | serialize

POLS, Random Effect

FE ✓ ① $y_{it} = a_i + \beta x_{it} + u_{it}$

② $x_{it} = \alpha_i + x^o_{it}$ → time var

$x^o_{it} \sim iid(0, \sigma^2_e)$

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

$E(\hat{\rho} - \rho) = \frac{-2\rho}{T}$ $\lim_{T\to\infty}$

$\text{plim}_{T\to\infty}(\hat{\rho} - \rho) = 0$: $E(\hat{\rho} - \rho)$

$E(\hat{\rho} - \rho)^2 = c$ → Consistent

$\frac{\sqrt{N}(\hat{\beta} - \beta)}{V(\beta)} \sim t_n$

$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{d} N(0, (X'X)^{-1}(X'\Omega X)(X'X)^{-1})$

$\frac{\sqrt{N}(\hat{\beta} - \beta)}{V(\beta)} \xrightarrow{d} N(0,1) + O_p(\frac{1}{\sqrt{N}})$

Good term, nuissance term

④ $y_{it} = a_i + \theta_t + \beta x_{it} + u_{it}$

④** $y_{iT} = a + \beta x_i + [u_i^{**}]$, $(a_i - a) + u_{iT}$

④* $y_{it} = a_i + \beta x_i + u_{it}$

⑤ $y_{it} = a + \beta x_i + \rho y_{it-1} + u_{it}$

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) $y_{it} - \hat{\gamma} z_{it} = a + \beta x_{it} + \theta_t + u^*_{it}$ $\hat{\beta}_{OLS} \xrightarrow{p} \beta$ $N,T\to\infty$

$\sum_T (y_{it} - \hat{\gamma} z_{it}) = a + \beta x_i + u^*_{it}$

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 $y_{it} = a_i + \beta_i x_{it}\ u_{it}$ Ess [run first] $\sum_N \beta_i = \beta$

$E(e_{it} e_{is}) = \begin{cases} \sigma_a^2 & \text{if } t \neq s \\ \sigma_a^2 + \sigma_u^2 & \text{if } t = s \end{cases}$

POLS corr → $t_{\hat{\beta}_{POLS}} = \frac{\hat{\beta}_{POLS} - \beta}{\sqrt{\hat{\sigma}_e^2 (X'X)^{-1}}}$

ⓐ $E(\hat{\beta} - \beta)^2 = E\left[\sum^{NT}(x_{it} - \sum_{NT} x_{it})(e_{it} - \sum_{NT} e_{it})\right]^2$

$E(x_{it} x_{is} e_{it} e_{is})_{t \neq s} = \gamma^{t-s} \sigma_\alpha^2 \sigma_e^2$ since $x \sim AR(1)$: $E(x_{it} x_{is}) = \gamma^{t-s} \sigma_\alpha^2$

ⓐ $= \frac{\sigma_u^2 + \sigma_a^2 + (\frac{\sigma_a^2}{T})(\frac{1-\gamma^T}{1-\gamma})}{N \cdot T \cdot \sigma_\alpha^2} \rightarrow \hat{\sigma}_e^2 (X'X)^{-1}$

Remedy → run FE

Correct $t_{\hat{\beta}} = \frac{\hat{\beta}_{OLS} - \beta}{\sqrt{\frac{(\sum^{NT} \hat{e}_{it}^2 + \sum^{NT} e_{it} e_{is})}{\sum^{NT} \tilde{x}_{it}^2}}}$ typically

$\hat{\beta}_{FE} \neq \hat{\beta}_{POLS}$

$y = \beta X + e$ $E(\hat{\beta} - \beta)(\hat{\beta} - \beta)' = E((X'X)^{-1} X' \underbrace{\Omega}_{E(ee'|X)} X (X'X)^{-1})$

POLS

$\Omega = \begin{bmatrix} \Omega_{11}\ (T\times T) & & \dots \\ & \ddots & \\ T\times T & \dots & \end{bmatrix}_{(NT)\times(NT)}$

$\Omega_{11} = \begin{bmatrix} \sigma_a^2 + \sigma_u^2 & \sigma_a^2 & \dots \\ \sigma_a^2 & \ddots & \\ \vdots & & \sigma_a^2 + \sigma_u^2 \end{bmatrix}$

→ Condition = replicate matrix N times

iff diff

indep zero

Non zero diagonal

$T \times (T-1)$ 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

$\Omega_\beta = (X'X)^{-1} X' \Omega X (X'X)^{-1} \cdot n$ $\sqrt{N}(\hat{\beta}_{POLS} - \beta) \xrightarrow{d} N(0, \Omega_\beta)$ sandwitch $PP' = \Omega^{-1}$

GLS $E[(\hat{\beta}_{GLS} - \beta)(\hat{\beta}_{GLS} - \beta)'|X] = \boxed{(X'\Omega^{-1}X)^{-1} = \Omega_{\hat{\beta}_{GLS}} \cdot n^{-1}}$

$P'y = P'X\beta + P'u \Rightarrow \sqrt{N}(\hat{\beta}_{GLS} - \beta) \xrightarrow{d} N(0, \Omega_{\hat{\beta}_{GLS}})$ $\frac{E(A)}{B}$

$V(\hat{\beta}_{GLS}) < V(\hat{\beta}_{POLS}) \Rightarrow \sigma_u^2 (X'X)^{-1} \neq V(\hat{\beta}_{GLS})$ 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} ⊗ Corel

① $y_{it} = \alpha_i + \beta x_{it} + u_{it} \rightarrow \hat{\alpha}_{iFE}, \hat{\beta}_{FE}$

② $\hat{\alpha}_i \rightarrow \hat{\sigma}_\alpha = \sum_N^N (\hat{\alpha}_i^2 - \sum_N \hat{\alpha}_i)^2$ ⇒ ③ cholese decomposition

$\hat{\sigma}_u = \sum_{NT}^{NT} \hat{u}_{it}^2$ $\hat{\Omega}^{-1}$ →(estimate) $\hat{P}'$ → $\hat{\beta}_{GLS}$

if serial corr $\hat{\Omega}^{-1}$ → ⊗ 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

$E(uu')_{T\times T} = \begin{bmatrix} \sigma_u^2 & \rho\sigma_u^2 & \rho^2\sigma_u^2 & \rho^{T-1}\sigma_u^2 \\ \vdots & \ddots & & \vdots \\ \rho^{T-1}\sigma_u^2 & & & \sigma_u^2 \end{bmatrix} = \Omega_o = \Omega_{ii} \Rightarrow \Omega = \begin{bmatrix} \Omega_o & & 0 \\ & \ddots & \\ 0 & & \Omega_o \end{bmatrix}$

① Run $y_{it} = \alpha_i + \beta x_{it} + u_{it} \Rightarrow \hat{u}_{it} = y_{it} - \hat{\alpha}_i - \hat{\beta} x_{it}$

② get $\hat{\rho}$ $\hat{u}_{it} = \rho \hat{u}_{it-1} + \varepsilon^*_{it}$ → ③ $\hat{\Omega}$ from ② → ④ GLS

(Stata)

$\begin{cases} V(\hat{\beta}_{FE\,GLS}) = (X'\hat{\Omega}^{-1}X)^{-1} \\ \hat{P}\hat{P}' = \hat{\Omega}^{-1} \\ \hat{P}'y = \hat{P}'X\gamma + \hat{P}'u \end{cases}$

$V(\hat{\beta}_{FE}) = (X'\hat{\Omega}X)^{-1} X'\hat{\Omega}X(X'\hat{\Omega}X)^{-1} \cdot NT$ ← 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), ...

$\Omega_{T\times T} = \begin{bmatrix} \sigma_1^2 & \dots & \sigma_{1T} \\ & \ddots & \vdots \\ " & & \sigma_T^2 \end{bmatrix}$ unknown $\frac{T(T-1)}{2} + T = O(T^{-1})$

$\hat{\sigma}_{ts} = \sum_n \hat{u}_{it} \hat{u}_{is}$ $\hat{\Omega}_\beta = (X'\hat{\Omega}X)^{-1} X'\hat{\Omega}X(X'\hat{\Omega}X)^{-1}(NT)$

Panel Robust Covar matrix Estimator

identity iff $n \geq T$ ←

Dynamic panel Reg $y_{it} = \alpha_i + \rho y_{it-1} + u_{it} \Rightarrow y_{it} - \frac{\sum y_{it}}{T} = \tilde{y}_{it}$

$\hat{\rho}_{FE} - \rho = \frac{[\sum^{NT} \tilde{y}_{it-1} \tilde{u}_{it}]_A}{[\sum^{NT} \tilde{y}^2_{it-1}]_B}$ $B = \frac{\sigma_u^2}{1-\rho^2}$

$\text{plim}_{N\to\infty} A = \text{plim}_{N\to\infty} \frac{\sum^{N,T}}{N} (y_{it-1} - \sum_T y_{it-1})(u_{it} - \frac{\sum u_{it}}{T})$ Nickell B

$= \text{plim}_{N\to\infty} [\underbrace{\sum_N^{N,T} y_{it-1} u_{it}}_{C} - \underbrace{\sum_N^N (\sum_T y_{it-1})(\sum^T u_{it})}_{D}]$, $E(\frac{D}{B}) = \frac{-1+\rho}{T} + O(\cdot)$

$E(D_B)$ $\frac{1}{B}$

$\sqrt{NT}(\hat{\rho}_{FE} - \rho) = \sqrt{NT}\frac{C}{B} + O_p(\frac{\sqrt{NT}}{T}) \xrightarrow{d} N(0, 1-\rho^2)$

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

② $\frac{N}{T} \to \infty$: $\sqrt{NT}(\hat{\rho}_{FE} - \rho) \rightarrow$ No lim dist

③ $N = T$: $N(0, 1-\rho^2) + O_p(1)$

Solution: IV ① y_{it-2} ② level $y_{it-2} - y_{it-3}$

$y_{it} - y_{it-1} = \rho(y_{it-1} - y_{it-2}) + (u_{it} - u_{it-1})$ $\boxed{\hat{\beta}_{IV} = (Z'X)^{-1} Z'Y}$

Econometrics April 1st ①

Panel Regression

y_{it} x_{it} Z_i W_t

Stock return | industry size firm | interest rate, bond rate, treasury bond rate

$$y_{it} = \alpha_i + X_{it}\beta + \theta_t + u_{it} \quad (1)$$

ignore Z_i, W_t

indiv dummy, time dummy

parameter estimate of interest is β: $H_0: \beta = 0$

if your run $y_{it} = \alpha + X_{it}\beta + \theta_t + e_{it}$ (1)*

$\hat{\beta}_{pols} \not\to \beta$ if $E\, X_{it} e_{it} \neq 0$

means pols will become inconsistent

From (1) serial correlate u_{it}, x_{it}

$\hat{\beta}_{FE} \xrightarrow{p} \beta$ as long as $E\, x_{it} u_{js} = 0$ for all $i, j, t, s \Rightarrow$ x_{it} is strongly/strictly Exogenous

Consistency Condition

Strong assumption

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

$y_{it} = \alpha_i + \theta_t + \beta(a_i + \rho x_{it-1} + \varepsilon_{it}) + u_{it}$ (1)

$= \alpha_i + \theta_t + \rho\beta x_{it-1} + \beta\varepsilon_{it} + u_{it} + \beta a_i$

$= \alpha_i^* + \theta_t + \beta x_{it-1} + u_{it}^*$ (1)''

$\boxed{E\, x_{it-1} u_{it}^* = 0}$ weakly exogeneous — past var is not correlated with current shock

$\operatorname{plim}_{n\to\infty} \sum_N \left(\sum_T x_{it-1}\right)\left(\sum_T u_{it}^*\right) \neq 0$

$$\begin{bmatrix} x_{it} \\ u_{it} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_{it-1} \\ u_{it-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{it} \\ e_{it} \end{bmatrix}$$

x_{it} is exogeneous if $a_{12} \neq 0$ or $a_{21} \neq 0$

Weakly Exogenous $\Rightarrow a_{12} = a_{21} = 0$

But $E(\varepsilon_{it} e_{it}) \neq 0$

Strongly Exogenous $E(\varepsilon_{it} e_{it}) = 0$, $a_{12} = a_{21} = 0$

$y_{it} = \alpha_i + \theta_t + \beta_i x_{it-1} + u_{it}^*$

$\Delta y_{it} = \Delta\theta_t + \beta_i \Delta x_{it-1} + \Delta u_{it}^*$

$\left.\begin{matrix}\Delta x_{it-2} \\ x_{it-2}\end{matrix}\right\}$ IV

②

$\hat{\beta}_{FDIV} \xrightarrow{p} \beta$

First Difference IV

Problem of IV

$x_{it} = \rho_t x_{it-1} + \varepsilon_{it}$ if $\rho = 1$

$\rho_t = 1 - c/T \to$ load to unity

$\Delta x_{it-1} = x_{it-1} - x_{it-2} = \varepsilon_{it-1}$

$E(\varepsilon_{it-1}\varepsilon_{it-2}) = 0$

$(Z'X)^{-1}$ would become singular

$(Z'X)^{-1} \Rightarrow$ Does not Exist

$\simeq 1/0$

$V(\hat{\beta}_{FDIV}) \to \infty$ extremely volatile

$E(x_{it-2}\Delta x_{it-1}) = E(x_{it-1}\varepsilon_{it-1}) = 0$

Weak IV problem: so if $\rho \to 1$ IV beomes Invalid $\Rightarrow$ if $\rho = .9$ Do not use IV

---

so what can you do? many papers on this topic study yourself

- if $t \to \infty$ bias goes away and Consistent

$\Rightarrow$ long time series data & not run panel regression (in finance)

---

what they do in microEconomics is very strong $\to$ IV

S_{it} as instrumental variables

Geographic variable, temprature, rain falls

$\hat{\beta}_{IV} = (S'X)^{-1}(S'Y)$

- all IV papers turn out to be wrong, they report $\hat{\beta}_{IV} \neq 0$ & others test and it was wrong

alternative

$y_{it} = \alpha_i + X_{it}\beta + \underbrace{\lambda_i' F_t + u_{it}^o}_{u_{it}}$

Correlate

$x_{it} = \phi_i' G_t + X_{it}^o$

$E\, X_{it}u_{it} \neq 0$ b/c $E(G_{it}F_t) \neq 0$ but $E\, X_{it}^o u_{it}^o = 0$

like moving piano

this is called Factor Augmented panel Regression

Pesaran (2006) CCE Commonly Correlated Estimator

$y_{it} - \alpha_i - x_{it}\beta - u_{it}^o = \lambda_i' F_t$

$\sum_{NT} y_{it} - \sum_N \alpha_i - \left(\sum_N x_{it}\right)\beta - \underbrace{\sum_N u_{it}^o}_{O_p\left(\frac{1}{\sqrt{N}}\right)} = \bar{\lambda}' F_t$ proxy random var

Cross sec mean

$\bar{\lambda}' F_t = f\left(\sum_N y_{it}, \sum_N x_{it}\right)$

run $y_{it} = \alpha_i^* + x_{it}\beta + f_{yi}\left(\sum_N y_{it}\right) + f_{xi}\left(\sum_N x_{it}\right) + u_{it}^+$

$y_{it} = \alpha_i + \beta x_{it} + \gamma_i \underbrace{w_t} + u_t^*$ ⊛ alternative method ③

macro variable

GDP, unemployment rate

- macro shocks affect personal level / & we disconnect that by this

problem: Does not take all var's out

$y_{it} = \alpha_i + \beta x_{it} + \gamma_i w_t + u_{t+}^* \quad \Phi \bar{y} + \varphi \bar{x}$

add previous section to this ⊛

$y_{it} = \alpha_i + \beta x_{it} + \quad \underline{\lambda_i \hat{F}_t} + u_{it}^*$

identifiable

this study third method

④

Econometrics 8-April ①

may13 Final - may 23 Qualifier
└ Comprehensive

- In top Journal → program it yourself - understand Equation
replicate results - Program it

- today and next session review due to not good scores

① $y_{it} = \alpha_i + \theta_t + \beta x_{it} + u_{it}$

$E(x_{it} u_{js}) \neq 0$ $\quad x_{it}$ endogeneous

Sol1 : use IV - rain force - geographic

Sol2. not good idea

Stock Price/return

$y_{it} = \alpha_i + \theta_t + \beta^* \underbrace{x_{it-1}}_{\text{lag}} + u_{it}^*$ → dividend ratio → var we want to explain growth with

Growth rate GDP

why? $x_{it} = \rho\, \cancel{x}_{it-1} + e_{it}$

$\beta^* = \rho\beta$

x_{it} due to e_{it} correl u_{it}:

$x_{it} = \rho x_{it-1} + e_{it}$ ⓘ

$u_{it}^* = \delta e_{it} + \varepsilon_{it}$ ⓘⓘ

ignore θ_t:

$$\hat\beta^*_{FE} - \beta^* = \frac{\sum^N \sum^T \tilde{x}_{it-1} \tilde{u}_{it}^*}{\sum^N \sum^T \tilde{x}^2_{it-1}}$$

where $\tilde{x}_{it-1} = x_{it-1} - \frac{\sum^T x_{it-1}}{T}$

+ You have to focus on sign of 'β', NEG, POS, or Zero and not value, since value interpretation is wrong essense is sign and not point (magnitude)

⊛ ρ is not interest, parameter of interest is β.

↳ ρ is usually close to 1 - dividend ratio is close to one

$$\sum^N \sum^T \tilde{x}_{it-1} \tilde{u}^*_{it} = \underbrace{\sum^N \sum^T x_{it-1} u^*_{it}}_{\text{Good term}} - \underbrace{\frac{\sum^N}{T} (\sum^T x_{it-1})(\sum^T u^*_{it})}_{\text{Bad term}}$$

Good term does not have bias

$- E \frac{1}{T} (\sum x_{it-1})(\sum e_{it})$

- most of Panel data is T=3 or T=4 and size of N is huge e.g. 5,000.

☒ For final Exam Practice derieving for T=3,4,5

$$\sum^T u_{it} = \delta \sum^T e_{it} + \sum^T \varepsilon_{it} = \frac{\sum^N}{T} (\sum^T x_{it-1})(\sum^T u^*_{it}) =$$

$$-\frac{\delta}{T} \sum^N (\sum x_{it-1})(\sum^T e_{it}) + O_p(1)$$

{ $N > T$ → Nikel Bias $\sqrt{NT}(\hat\beta^*_{FE} - \beta^*) \xrightarrow{d} \pm\infty$ 
{ $N \cong T$ → Bias Correction $\hat\beta^*_{FE} - \beta^* = -\delta \frac{(1+\rho)}{T}$ (steps...)
{ $N < T$

②

①ρ → $x_{it} = \alpha_i + \rho x_{it-1} + e_{it}$ (Steps...)

$\hat\rho = \rho - \frac{1+\rho}{T} + O_p(T^{-2})$

$E\hat\rho = \rho - \frac{1+\rho}{T} + O(T^{-2})$

Fixed ← T → know

$\hat\rho = \rho T - (1+\rho) = -1 - \rho + \rho T = -1 + \rho(T-1)$

$\frac{\hat\rho T + 1}{T-1} = \rho$

② $E \hat\delta = \delta$

where $\hat u^*_{it} = \delta \hat e_{it}$ + error

$\hat u^*_{it} = \tilde{y}_{it} - \hat\beta^+_{FE} x_{it} = u^*_{it} + \Delta_1$

$\hat e_{it} = \tilde{x}_{it} - \hat\rho_{FE} \tilde{x}_{it-1} = e_{it} + \Delta_2$

$E(\hat\delta) = \delta$

end of sol 2: time lagged

Sol 3 use Covariate variable.

⊛ Oil price, GDP, ... Covariate with stock return
variables used in previous litrature

⊛ in your paper argue I found another meaningful variable that effects → x_{it}

$y_{it} = \alpha_i + \theta_t + \rho x_{it} \underbrace{+ Z_{it}\gamma + e_{it}}_{u_{it}}$

⊛ include variables of papers related to topic of your research and cite all those papers

⊛ multicollinearity says correlation does not exist - not determinant _ not significant

Sol 3 Cont Combine solution 2 & 3

$y_{it} = \alpha_i + \theta_t + \beta^* x_{it-1} + Z_{it-1}\gamma^* + \epsilon^*_{it}$

one of the control var becomes key var

Nichel bias could still be problem due to:

$E(e_{it}\epsilon^*_{it}) \neq 0$

② $y_{it} = \alpha_i + \theta_t + \beta x_{it} + u_{it}$

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

don't know if e=0 or not

why u_{it} is error (unknown)

y_{it} is serially correlated

x_{it} is (not) weakly serially correlated

Ex. y_{it} earning, wage, growing variable, including Price

$y_{it} = \log$ GDP ← per capita real

i: country

$y_{it} = \alpha_i + \theta_t + \beta x_{it} + u_{it}$

$\boxed{y_{it} = \alpha_i + \theta_t + \rho y_{it-1} + \beta x_{it} + u_{it}}$ ← inconsistent with Previous

$\rho y_{it-1} = \alpha_i \rho + \theta_t \rho + \beta\rho x_{it-1} + \rho u_{it-1}$

$y_{it} = \alpha_i(1-\rho) + \theta_t(1-\rho) + \beta x_{it} - \beta\rho x_{it-1} + \varepsilon_{it}$

- people follow big people who doesn't know Econonometrics and go wrong and run (W)
- or $y_{it} = \rho y_{i0} + \beta x_{it} + u$ → Convergence growth
  South african Country (HIV) QJE wrong and people extended it.

→ even under assumption of Exogeneous x_{it} mean $E(x_{it} u_{it}) \neq 0$ (W) has biased β and ρ

$E(\hat\beta_{FE} - \beta) = -E(\hat\rho_{FE} - \rho) \cdot$ → Bias

even sign is affected by direction of bias
sign($\hat\beta_{FE}$) = affected by Nickell Bias

$\rho \approx 1$ usually Nickel bias would be large

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

$T=3,\ \rho=1 \quad E(\hat\rho) = 0.25$

$E(\hat\rho - 1) = -.75$ which is huge

in Econ 3 suggest solution and program it

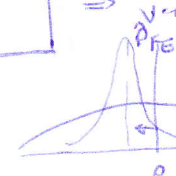
Solution $\Delta y_{it} = \Delta\theta_t + \rho\Delta y_{it-1} + \beta\Delta x_{it} + \Delta u_{it}$

IV -: $y_{it-2} \quad \Delta y_{it-2}$
=: $x_{it-2} \quad \Delta x_{it}$

Brundel & Bond Estimator ⇒ two IV

$\hat\rho_{FE} > \hat\rho_{FDIV}$. But

$\rho \approx 1 \Rightarrow (Z'X)^{-1}$ → Singular
$\Delta y_{it-1} = \varepsilon_{it}$
⇒ variance very large



[+] helps to criticize empirical papers

if ρ not close to 1 You could

$\Delta y_{it} = \alpha_i + \theta_t + \beta\Delta x_{it} + u_{it}$

Converting this to the ⟨1⟩, but people may criticise You on interpretation

⚠ Do not believe empirical research they are generally wrong. e.g. Terrorism and Economic growth

- if you Control for Nickel bias β would become insignificant for $\rho \approx 1$

⟨3⟩ $y_{it} = \alpha + \beta x_{it} + u_{it}$

$x_{it} = \begin{cases} 0 \\ 1 \end{cases}$ time variant dummies

or $x_i = \begin{cases} 0 \\ 1 \end{cases}$

Dummies Cross sec

y_{it} - Birth rate, min wage, unemployment wage

x_i Dummies: treatment, take medicine, Course

y_{it} Could be serially Correlate

e.g. policy arizona/georgia hispanic school immigr document
legalization of marijuana

state income serially Correlate

someone does: $y_{it} = \alpha_i + \theta_t + \rho y_{it-1} + \beta x_{it} + u_{it}$ (W)

WRONG! $\alpha_i = \alpha(1-\rho)$
You are Caring about mean
You don't know mean or policy

x_{it} → pretty much Exogenous

$$\boxed{V(\hat\beta_{FE} - \beta) = (X'X)^{-1} X' \Omega_u X (X'X)^{-1}}$$

⇒ panel Robust Covariance matrix (PRCM)
or clustering

u_{it} here serially Correlate
x_{it} Dummy variable ⇒ PRCM

$$t_{\hat\beta} = \frac{\hat\beta_{FE}}{\sqrt{\hat\sigma_u^2 / \sum^{NT} \tilde{x}_{it}^2}}$$

interpret $y_{it} = \alpha_i + \beta x_{it} + u_{it}$
$u_{it} = \rho u_{it-1} + \varepsilon_{it}$

interest: mean change

⚠ Do NOT do $y_{it} = \alpha_i(1-\rho) + \beta x_{it} + \beta\rho x_{it-1} + \varepsilon_{it}$ if x_{it} is binary

Ω → if know ⇒ gls

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

$$\underset{(N\times T)\times(N\times T)}{\Omega_u} = \begin{bmatrix} A & & 0 \\ & A & \\ & & \ddots & \\ 0 & & & A \end{bmatrix} \qquad A = \sigma_u^2 \begin{bmatrix} 1 & \cdots & \rho^{T-1} \\ \rho & 1 & \rho^{T-2} \\ \vdots & \ddots & \\ \rho^{T-1} & \rho^{T-2} \cdots & 1 \end{bmatrix}$$

$A_{T\times T}$

Estimate ρ: How?

Econometrics   8-April   5   6

$\hat{u}_{it} = \rho \hat{u}_{it-1} + \varepsilon_{it}$

$\hat{u}_{it} = y_{it} - \hat{\alpha}_i - \hat{\beta} x_{it}$

- x_{it} Exogeneous $\Rightarrow$ Consistent estimate of β

always

DID: difference in difference   AER

Stata assumes this and provides Clustering result

- Reveal what truth is
- she programs this herself

$$A = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1T} \\ \vdots & \ddots & & \\ \vdots & & \ddots & \\ \sigma_{1T} & & & \sigma_T^2 \end{bmatrix}$$

how estimate

$\hat{\sigma}_{1T} = \sum^{N} \frac{\hat{u}_{it}\hat{u}_{it}}{N}$

$N \gg T$

$N \approx T$ Can not identify   Forget it

$N > T^2$

total unknown

- The last one is most Conservative one

among ① panel robust Covariance matrix

② AR(1)

③ no assumptions on AR

Assumptions were:

① y_{it} persistent u_{it} serially Correlate

② x_{it} Exogeneous

③ fixed effect

Suppose you do not have

$y_{it} = \alpha + \beta x_{it} + u_{it}$ $E(x_{it} u_{is}) = 0$ for all t, s

restriction will affect this

but $y_{it} = \alpha + \beta x_{it} + e_{it}$

$\hat{\beta}_{pols} \overset{P}{\nrightarrow} \beta$ $e_{it} = (\alpha_i - \alpha) + u_{it}$

inConsistent

---

150   last Exam and final on same day

Econometrics | Tutor | 12 April

①

Q6. DGP: $\begin{cases} y_{it} = a_i + x_{it} \\ x_{it} = \rho x_{it-1} + u_{it} = \sum_{j=0}^{\infty} \rho^j u_{it-j} \\ u_{it} \sim iid(0,1) \end{cases}$

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

Generally Vector form $y = x\beta + u$ $\hat{\beta}$

OLS $x'y = x'x\beta + x'u \Rightarrow (x'x)^{-1}x'y = (x'x)^{-1}(x'x)\beta + (x'x)^{-1}x'u$

OLS $\hat{\beta} = \beta + (x'x)^{-1}x'u$ x, u must be uncorrelated

what if x, u be Correlated? Z (Instrument) $= \beta + \frac{\sum x_{it}u_{it}}{\sum x_{it}^2}$

$Z'Y = Z'x\beta + Z'u$

$(Z'x)^{-1}Z'Y = (Z'x)^{-1}Z'x\beta + (Z'x)^{-1}Z'u$

$\hat{\beta} = \beta + (Z'x)^{-1}Z'u \Rightarrow \hat{\beta} - \beta = (Z'x)^{-1}Z'u = \frac{\sum^{N,T} z_{it}u_{it}}{\sum^{N,T} z_{it}x_{it}}$

- Z and u should be uncorrelated
Generally

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

$y_{it-1} = \alpha_i + \rho y_{it-2}$

$\underbrace{y_{it} - y_{it-1}}_{\Delta y_{it}} = \rho \underbrace{(y_{it-1} - y_{it-2})}_{\Delta y_{it-1}} + \underbrace{(u_{it} - u_{it-1})}_{\Delta u_{it}}$

$y_{it-1} - y_{it-2} = (a_i + x_{it-1}) - (a_i + x_{it-2}) = x_{it-1} - x_{it-2}$

$= \sum_{j=0}^{\infty} \rho^j u_{it-1-j} - \sum_{j=0}^{\infty} \rho^j u_{it-2-j} = (u_{it-1} + \rho u_{it-2} + \rho^2 u_{it-3} \dots)$

$+ (u_{it-2} + \rho u_{it-3} + \rho^2 u_{it-4} + \dots)$ → Correl ⇒ need IV

$\Rightarrow \hat{\rho} = \rho + \frac{\sum\sum \Delta y_{it}\, IV}{\sum\sum \Delta y_{it-1}\, IV}$

T = 3 $\quad y_{i1}, y_{i2}, y_{i3}$

$y_{i3} = a_i + \rho y_{i2} + u_{i3}$

$y_{i2} = a_i + \rho y_{i1} + u_{i2}$

$y_{i3} - y_{i2} = \rho(y_{i2} - y_{i1}) + (u_{i3} - u_{i2})$

$(y_{i2} - y_{i1})$ & $(u_{i3} - u_{i2})$ are Correlated

$\hat{\rho} = \rho + \frac{\sum_i \Delta u_{it}\, IV}{\sum_i \Delta y_{it} \cdot IV} = \rho + \frac{\sum_i (u_{i3} - u_{i2}) \cdot IV}{\sum_i (y_{i2} - y_{i1}) \cdot IV}$

To find instrumental var uncorrelated with $\Delta u_{it} = u_{i3} - u_{i2}$

IV? $= y_{i1}$

$y_{i1} = a_i + x_{i1} \Rightarrow y_{i1}, \Delta u_{it}$ are independent

$\hat{\rho} = \rho + \frac{\sum_i (u_{i3} - u_{i2}) y_{i1} \times \frac{1}{N}}{\sum_i (y_{i2} - y_{i1}) y_{i1} \times \frac{1}{N}}$

②

$\hat{\rho} - \rho =$

$\boxed{\text{Denominator}}$ $\text{plim} \frac{1}{N}\sum_i (y_{i2} - y_{i1}) y_{i1} = E(\frac{1}{N}\sum_i (y_{i2} - y_{i1}) y_{i1})$

$= E[\frac{1}{N}\sum_i (\cancel{a_i} + x_{i2} - \cancel{a_i} - x_{i1})(a_i + x_{i1})]$

$= E[\frac{\sum_i (x_{i2} - x_{i1})(a_i + x_{i1})}{N}] = \frac{\sum_i}{N} E(x_{i2}a_i + x_{i2}x_{i1} - x_{i1}a_i - x_{i1}^2)$

$\boxed{\begin{aligned} &x_{it} = \sum_{j=0}^{\infty} \rho^j u_{it-j} \quad E(x_{it}) = 0 \\ &V(x_{it}) = E(x_{it})^2 = \frac{\sigma_u^2}{1-\rho^2} \\ &E(x_{it} x_{is}) = \rho^{t-s} \sigma_x^2 \end{aligned}}$

$= \frac{1}{N}\sum_i (0 + \rho\sigma_x^2 - 0 - \sigma_x^2) = (\rho - 1)\sigma_x^2 = -(1-\rho)\sigma_x^2$

$\boxed{\text{Numerator}}$

$E(\frac{\sum}{N}(u_{i3} - u_{i2}) y_{i1}) = E[\frac{\sum_i}{N}(u_{i3} - u_{i2})(a_i + x_{i1})]$

$= E[\frac{\sum_i}{N}(u_{i3}a_i + u_{i3}x_{i1} - u_{i2}a_i - u_{i2}x_{i1})] = 0$

$V(\frac{\sum_i}{N}(u_{i3} - u_{i2}) y_{i1}) = E[\frac{\sum_i}{N}(u_{i3} - u_{i2}) y_{i1}]^2 =$ independ

$E[\frac{1}{N}\sum_i (u_{i3} - u_{i2})(a_i + x_{i1})]^2 = \frac{E[\sum_i (u_{i3} - u_{i2})^2 (a_i + x_{i1})^2]}{N^2}$

$= \frac{1}{N^2}(2\sigma_u^2)(\sigma_a^2 + \sigma_x^2)$

$\boxed{\begin{aligned} &E(u_{i3} - u_{i2})^2 = 2\sigma_u^2 = E(u_{i3}^2 - 2u_{i3}u_{i2} + u_{i2}^2) \\ &E(a_i + x_{i1})^2 = E(a_i^2 + 2a_i x_{i1} + x_{i1}^2) = \sigma_a^2 + \sigma_x^2 \\ &\text{denote } E(a_i^2) = \sigma_a^2 \end{aligned}}$

Numerator $\xrightarrow{d} N(0, \frac{1}{N}(2\sigma_u^2)(\sigma_a^2 + \sigma_x^2))$

$\frac{\text{Numerator}}{\text{Denom}} = \hat{\rho} - \rho \longrightarrow N(0, \frac{(2\sigma_u^2)(\sigma_a^2 + \sigma_x^2)}{N(1-\rho)^2(\sigma_x^2)})$

$\sqrt{N}(\hat{\rho} - \rho) \xrightarrow{d} N(0, \underbrace{\frac{(2\sigma_u^2)(\sigma_a^2 + \sigma_x^2)}{(1-\rho)^2 \sigma_x^2}}_{A})$

$A = \frac{2\sigma_u^2(\sigma_a^2 + \frac{1}{1-\rho^2}\sigma_u^2)}{(1-\rho)^2 \frac{1}{(1-\rho)^2}\sigma_u^4} = \frac{2(\sigma_a^2 + \frac{1}{1-\rho^2})}{\frac{(1-\rho)^2}{(1-\rho)^2}}$

$= \frac{2[(1-\rho^2)^2\sigma_a^2 + (1-\rho^2)]}{(1-\rho)^2}$

divide $(1-\rho)$: $A = \frac{2[(1+\rho)^2(1-\rho)\frac{\sigma_a^2}{\sigma_u^2} + 1+\rho]}{1-\rho}$

$\boxed{(1-\rho^2)^2 = [(1+\rho)(1-\rho)]^2 = (1+\rho)^2(1-\rho)^2}$

$= \frac{2[(1+\rho)^2(1-\rho)\sigma_a^2]}{1-\rho} + \frac{2(1+\rho)}{1-\rho} = 2(1+\rho)^2\sigma_a^2 + \frac{2(1+\rho)}{1-\rho}$

③

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

when $\sigma_a^2 \to 0$ $\qquad N\left(0, \frac{2(1+\rho)}{1-\rho}\right)$

$$\sqrt{N}(\hat{\rho}-\rho) \to N(0, \sigma^2)$$

$\sigma_a^2 \to \infty \Rightarrow$ IV fails

④

Econometrics | Apr 11 | ①

$y_{it} = \alpha_i + \beta x_{it} + u_{it}$

$x_{it} = \begin{cases} 1 \\ 0 \end{cases}$

Difference-in-Difference

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

$\Rightarrow$ Panel Robust Covariance Estimator

$\hat{\beta}_{FE} \xrightarrow{p} \beta$ as long as $E(x_{is} u_{it}) = 0 \quad \forall\ i, t, j, s$

$\sqrt{NT}(\hat{\beta}_{FE} - \beta) \xrightarrow{d} N(0, \Omega_\beta)$

$\Omega_\beta = \left(\frac{X'X}{NT}\right)^{-1}\left(\frac{X'\Omega X}{NT}\right)\left(\frac{X'X}{NT}\right)^{-1} / NT = O(1) \rightarrow$ general form

orders NT — devide to make $O(1)$

time series Correlation

$\hat{\Omega}_{ts} = \frac{1}{N}\sum_{i=1}^{N} \hat{u}_{it}\hat{u}_{is}$ is general formula since does not need u_{it} to foll AR(1)

if $u_{it} \sim AR(1)$ $\quad \Omega_{ts} = |\rho^{t-s}|\sigma_u^2$

$\hat{\rho} \Leftarrow \hat{u}_{it} = \rho\hat{u}_{it-1} + \text{error}$

Q: $E(\hat{\rho}) = \rho$?

$\hat{\alpha}_i \rightarrow$ since it is already mean of previous term

$\hat{u}_{it} = (y_{it} - \hat{\beta}x_{it}) - \sum_T (y_{it} - \hat{\beta}x_{it}) = \tilde{u}_{it} + (\beta - \hat{\beta})\tilde{x}_{it}$

$\hookrightarrow$ NO $\quad E\left(\sum_T u_{it-1}\right)\left(\sum \text{error}\right) \neq 0$

$\Rightarrow \rho\hat{u}_{it-1} = \rho\tilde{u}_{it-1} + \rho(\beta - \hat{\beta})\tilde{x}_{it-1} \Rightarrow \tilde{u}_{it} = \rho\tilde{u}_{it-1} + \underbrace{\tilde{\varepsilon}_{it} + \rho(\beta - \hat{\beta})\tilde{x}_{it-1}}_{\text{error}}$

stata gives you this

SURE — seemingly unrelated regression estimator

$y_{it} = \alpha_i + \beta_i x_{it} + u_{it}$

Cross Section Correlation

Example: $\quad y_{1t} = \alpha_1 + \beta_1 x_{1t} + u_{1t}$

$N=2 \quad y_{2t} = \alpha_2 + \beta_2 x_{2t} + u_{2t}$

$E\, u_{1t}u_{2t} \neq 0$

"account for cross sectional dependance" Lie!

improve $\hat{\beta}$ statistical inference $\rightarrow$ shrink variance, efficient

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

$Y \qquad X\beta \qquad U$

insure $x_{1t}x_{2t}$ should not be cross sec correlated

$\Omega_\beta = \underset{2\times2}{(X'X)^{-1}}(X'\underset{2T\times2T\times2}{\Omega X})\underset{2\times2}{(X'X)^{-1}} \rightarrow$

$\sqrt{NT} \rightarrow$ reason: Variance is converging as $N \rightarrow \infty$, $T \rightarrow \infty$

rate of shrinkage — law of large number

②

$\Omega_{ij} \Rightarrow \hat{\Omega}_{12} = \frac{\sum_{t=1}^{T}\hat{u}_{1t}\hat{u}_{2t}}{T}$

rather than ts which was for the sou — the rest are identical

---

- Economic theory is uniform and limiting distribution exactly identical

---

- identification problem still there $N \times N$ so if $T > N^2$ only identifiable

$\Rightarrow$ use GLS to improve the statistical inference

Litreture: Selection of study based on topic of study

- if in "Dif-in-Dif" we have cross sectional dependance, # of unknowns would be greater so the true variance due to Cov we will have larger and larger t-stat, as N increase (variance would be smaller and smaller compa to real as N increases)

- all variables will become significant. you should think about direction - serial correlation is positiv

- Shock response is at the same direction (cross section)

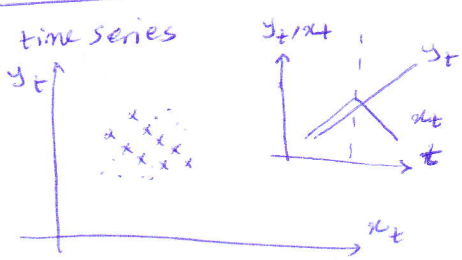
- when cancel (line vote) then cross product

$E(x_1 + \dots + x_n)^2 = E(x_1^2 + \dots + x_n^2) + 2E(\text{Cross Prod})$

- if Cross Section Corr in Dif-in-Diff we will have big problem

---

Breaks



split period - one positive one negative

$\begin{cases} y_t: \text{\# police} \\ x_t: \text{crime rate} \end{cases}$ $\quad \begin{cases} \text{government Consumption} \\ \text{Japenice Production} \end{cases}$

$\begin{cases} \text{unemployment rate} \\ \text{inflation} \end{cases}$

everytime it has Break

need particular specification of regression (3)

- Markov switching: need model
  timing – probability at timing
  markov chain – transition probability

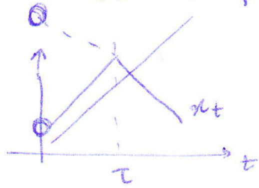
$y_t=\alpha+\beta x_t+u_t$

$y_t=\alpha_1+\beta_1 x_t+u_t \quad (t<\tau)$

(1) $y_t=\alpha_2+\beta_2 x_t+u_t \quad (t\ge\tau)$ ⟹ Good attempt But wrong [use of SURE is wrong, SURE Requires correlation]

Econometric modeling — structural equations based on basics

Dummy? $d_t=\begin{cases}0 & \text{if } t<\tau\\ 1 & \text{if } t\ge\tau\end{cases}$



$y_t=a_1+a_2d_t+b_1x_t+b_2x_td_t+u_t$

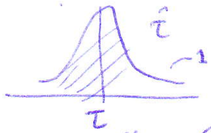
$H_0:\beta_1=\beta_2 \Rightarrow H_0: b_2=0$

to find τ $\quad t=1920\ldots2013 \quad \tau=$ from 2001 to 2000

$y_t=a_1+a_2d_t(\tau)+b_1x_t+b_2x_td_t(\tau)+u_t$

- if you cannot reject → means there would be no structural breaks
- institutional approach or reasoning to find the Break when you can not identify it

$Pr[\hat{\tau}=t]$



if u_t would be serially correlated, what you will do?

$t_{\hat{b}_2}$

$y=X\beta+u$

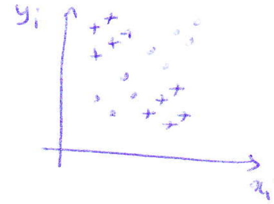
$\sqrt{T}(\hat{\beta}-\beta)\xrightarrow{d}N(0,\Omega_\beta)$

$\Omega_\beta=(X'X)^{-1}\underbrace{(X'\Omega X)}_{4\times4}(X'X)^{-1}/T$ figure out yourself

Hint: Long run variance

HAC Consistent Estimator to estimate $X'\Omega X$

(4)



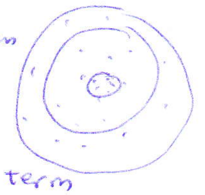
Same as before but what Dummy? important empirical question

$d_i=\begin{cases}0 & \text{if ie}\\ 1 & \text{if ie}\end{cases}$

- asian: if english at home income higher
- Construct different dummies and try to explain why structured breaks

$y_i=a_1+a_2d_i+b_1x_i+b_2x_id_i+u_i$

- The time series → you have ordering HAC
- Can we do ordering cross sectional? Grade, ID, ...
- Classify individual based on location (Spatial analysis)



  - geographic → social interaction
    Spatial analysis with Error term and do analysis

Spatial auto Correlation R

$y_i=\rho y_{\underline{i}}+x_i\underline{\beta}+u_i$

spatial coeffi

⟹ $\boxed{\text{Breaks} = \text{Dummies}}$

Testing → Lecture notes missing

$y=X\beta+u$

$H_0:\beta=0$

$H_A:\beta\neq0$ Complementary set of H_0

- null hypothesis set must be bounded; alternative we don't care — we test on boundaries

⟹ $H_0:\beta\ge0$ ✓

$H_A:\beta<0$ one sided

$H_0:\beta>0$ ⟹ Cannot test since not bounded

$H_A:\beta\le0$

$t=\frac{\hat{\beta}-*}{\sqrt{V(\hat{\beta})}} \rightarrow t=\frac{\hat{\beta}-0}{\sqrt{V(\hat{\beta})}}$

to test $\beta\ge0$

size

power → more powerful is better

type 1 Error:

Size: rejection rate of null when null is true (type 1 Error) → larger t-test has relation with this: larger t-test as N increases)

due to lower variance relative to what it should be

Econometrics April 15 (5)

- Significance 5% } → ~~two~~ ~~one~~ One side
  t-stat 1.65
  1.96 → two side

$N \to \infty$, 100% reject null
⇓ $100 - 5 = 95\%$ $\boxed{\text{size distortion}}$
Shrink the variance ⇒ larger t-stat ↑

⇒ for power of the test you need more data

power: rejection rate of the null when null is false. ⇒ max power = 1

increase sample size → higher convergence rate

- in practice do not use gls
  if want to publish paper use gls
  in experiment people use non parametric

- theoretically gls var is lower than OLS
  practically gls var is greater than ols
  when estimate Ω → total # unknown increase Exponentially (sacrifice deg of freedom)

  ⇒ Convergence rate lower than OLS ⇒ more size distortion

- Econometric theory → { $T \to \infty$, $N \to \infty$ } total # of unknown does not matter

→ gls more efficient than pols
  ↳ more likely reject → due to more size distortion and not power of the test

⇒ size distortion means:
  true 2.5 but you use 1.96

⚠ using gls is way of cheating

$\boxed{\text{Delta method}}$

$\frac{\hat{\alpha}}{\hat{\beta}} = 2$ $\quad y = X_1\alpha + X_2\beta + u$ restriction = 1 order 1

↳ t-dist → Normal
F - X^2

t-statistics → $\frac{\hat{\beta}}{\sqrt{V(\hat{\beta})}} \xrightarrow{d} N(0,1)$

$y = X\beta + u$
- sample small → only way to know $u \sim N(0, \sigma_u^2)$
  follow t-dist
- Exact sample theory

(6)

$\alpha = 2$ & $\beta = 2$ two restrictions

$\sqrt{N}\begin{bmatrix} \hat{\alpha} - \alpha \\ \hat{\beta} - \beta \end{bmatrix} \xrightarrow{d} N(0, (2 \times 2))$

Restriction

$R\underset{2\times 1}{\begin{bmatrix} \hat{\alpha} - \alpha \\ \hat{\beta} - \beta \end{bmatrix}'}\underset{2\times 1}{\begin{bmatrix} \quad \end{bmatrix}}'R \longrightarrow X^2_{R-1}$

$\boxed{\text{end of Review}}$

$\boxed{\text{New stuff}}$

GMM not thaught litrature huge
Limited rand var litrature huge
treatment effect panel logit
propensity score

$\boxed{\text{Econ 3}}$ MLE → programming

OLS you can work on paper
MLE cannot solve by paper, just program by

in this term Not MLE, but $\boxed{\text{GMM}}$
minimum distance estimator

- Standard estimation technique

Econ theory: $\underset{x_i}{\arg\max}\ u(x_i)$

First order Condition $\boxed{\frac{\partial u}{\partial x} \Rightarrow x_i^* \Leftarrow \frac{\partial u(x_i^*)}{\partial x} = 0}$

in binary choice use utility, but gmm uses FOC

$\arg\max E\, u(x_i)$

⇒ $\boxed{E\left(\frac{\partial u(x_i^*)}{\partial x}\right) = 0 \quad \text{moment Condition}}$

if you put x_i^* in your data FOC = 0

$\frac{\partial u}{\partial x}(x_i^*) = f(x_i^*, \theta)$ - heavy formation parameter

→ Gmm: estimate $\hat{\theta}$ using x_i^*
- Risk aversion
- fixed effect
- GMMIV

Example) $\arg\max\ u(x_i, \theta) = \frac{x_i^{1-\theta}}{1-\theta}$

$u(x_i, y_i, \theta) = (x_i^{\theta} + y_i^{\theta})^{1/\theta}$

You can have weighting $((1-\alpha)x_i^{\theta} + \alpha y_i^{\theta})^{\theta}$

$\frac{\partial u}{\partial x_i} = 0 \quad x_i^{-\theta} = 0 \Rightarrow E(x_i^{-\theta}) = 0$

$\Rightarrow \frac{1}{n}\sum x_i^{\theta} \xrightarrow{p} 0$

$\alpha + \beta x_i + \gamma x_i^2 + u_i$

$\beta + 2\gamma x_i = 0 \quad x_i = -\beta/2\gamma$

$\frac{1}{n}\sum x_i = -\beta/2\gamma \Rightarrow \frac{\sum x_i}{n} + \frac{\beta}{2\gamma} = 0$

here to identify separately you get second moment

Econometrics April 15 (1)

$y_{it} = \alpha_i + \beta x_{it} + u_{it}$

$x_{it} = \begin{cases} 1 \\ 0 \end{cases}$

Difference-in-Difference

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

$\Rightarrow$ Panel Robust Covariance Estimator

$\hat\beta_{FE} \xrightarrow{p} \beta$ as long as $E(x_{is} u_{it}) = 0 \quad \forall\, i, t, i, s$

$\sqrt{NT}(\hat\beta_{FE} - \beta) \xrightarrow{d} N(0, \Omega_\beta)$

$\Omega_\beta = \left(\frac{X'X}{NT}\right)^{-1}\left(\frac{X'\Omega X}{NT}\right)\left(\frac{X'X}{NT}\right)^{-1}/NT = O(1)$ → general form

orders NT — devide to make $O(1)$

time series Correlation

$\hat\Omega_{ts} = \frac{1}{N}\sum_{i=1}^{N} \hat u_{it}\hat u_{is}$ is general formula since does not need u_{it} to foll AR(1)

if $u_{it} \sim AR(1)$ $\quad \Omega_{ts} = |\rho^{t-s}|\,\sigma_u^2$

$\hat\rho \Leftarrow \hat u_{it} = \rho \hat u_{it-1} + \text{error}$

Q: $E(\hat\rho) = \rho$? $\quad \hat\alpha_i$ → since it is already mean of previous term

$\hat u_{it} = (y_{it} - \hat\beta x_{it}) - \frac{\sum}{T}(y_{it} - \hat\beta x_{it}) = \tilde u_{it} + (\beta - \hat\beta)\tilde x_{it}$

$\boxed{\text{No } E\left(\frac{\sum}{T} u_{it-1}\right)\left(\sum \text{error}\right) \neq 0}$ $\Rightarrow \rho \hat u_{it-1} = \rho \tilde u_{it-1} + \rho(\beta - \hat\beta)\tilde x_{it-1} \Rightarrow \tilde u_{it} = \rho \tilde u_{it-1} + \underbrace{\left(\tilde\varepsilon_{it} + \rho(\beta - \hat\beta)\tilde x_{it-1}\right)}_{\text{error}}$

stata gives you this

SURE seemingly unrelated regression estimator

$y_{it} = \alpha_i + \beta_i x_{it} + u_{it}$ $\quad$ Cross Section Correlation

Example: $y_{1t} = \alpha_1 + \beta_1 x_{1t} + u_{1t}$

$N=2$ $\quad y_{2t} = \alpha_2 + \beta_2 x_{2t} + u_{2t}$

$E\, u_{1t} u_{2t} \neq 0$ $\quad$ Lie!

"account for Cross sectional dependance"

improve $\hat\beta$ statistical inference → shrink variance, efficient

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

$Y \qquad X\beta \qquad U$

insure $x_{1t} x_{2t}$ should not be cross sec Correlated

$\Omega_\beta = \underset{2\times 2}{(X'X)^{-1}}\,(X'\underset{2T\times 2T}{\Omega} X)\,\underset{2\times 2}{(X'X)^{-1}}$ →

$\sqrt{NT}$ → reason: variance is converging as $N \to \infty$, $T \to \infty$

rate of shrinkage $\quad$ law of large number

(2)

$\Omega_{ij} \Rightarrow \hat\Omega_{12} = \frac{\sum_{t=1}^{T} \hat u_{1t}\hat u_{2t}}{T}$

rather than ts which was for the Sol

the rest are identical

- Economic theory is uniform and limiting distribution exactly identical
- identification problem still there $N \times N$ so if $T > N^2$ only identifiable

$\Rightarrow$ use GLS to improve the statistical inference

Litrature: selection of study based on topic of study

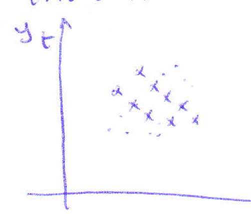
- if in "Dif-in-Dif" we have cross sectional dependance, # of unknowns would be greater so the true variance due to cov we will have larger and larger t-stat, as N increas (variance would be smaller and smaller compa to real as N increases)
- all variables will become significant. You shoul think about direction - serial correlation is positi
- Shock response is at the same direction (cross section)
- when Cancel line vote) then cross product

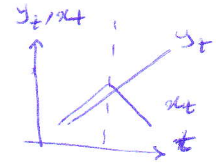
$E(x_1 + \dots + x_n)^2 = E(x_1^2 + \dots + x_n^2) + 2E(\text{Cross Prod})$

- if Cross Section Corr in Dif-in-Diff we will have big problem

Breaks

time series





split period - one positive one negative

y_t: # police
x_t: crime rate

government Consumption
Japanese Production

unemployment rate
inflation

everytime it has Break

need particular specification of regression

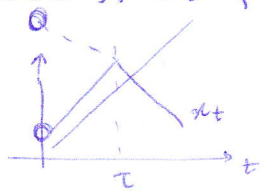
- Markov switching: need model
  timing — probability at timing
  markov chain — transition probability

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

① $y_t = \alpha_1 + \beta_1 x_t + u_t \; (t < \tau)$
$y_t = \alpha_2 + \beta_2 x_t + u_t \; (t \geq \tau)$ $\Rightarrow$ Good attempt But wrong [use of SURE is wrong, SURE Requires correlation]

Econometric modeling — structural equations based on basics

Dummy? $d_t = \begin{cases} 0 & \text{if } t < \tau \\ 1 & \text{if } t \geq \tau \end{cases}$



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

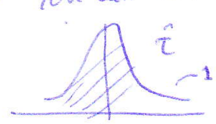
$H_0: \beta_1 = \beta_2 \Rightarrow H_0: b_2 = 0$

to find τ — $t = 1920 \ldots 2013$ $\tau =$ from 2001 to 2000

$y_t = a_1 + a_2 d_t(\tau) + b_1 x_t + b_2 x_t d_t(\tau) + u_t$

- if you cannot reject → means there would be no structural breaks
- institutional approach or reasoning to find the Break when you can not identify it

$Pr[\hat{\tau} = t]$



if u_t would be serially correlated, what you will do?

$t_{\hat{b}_2}$

$y = X\beta + u$

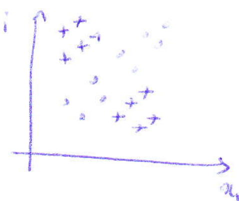
$\sqrt{T}(\hat{\beta} - \beta) \xrightarrow{d} N(0, \Omega_\beta)$

$\Omega_\beta = (X'X)^{-1}(X'\Omega X)(X'X)^{-1}/T$ — figure out yourself

4×4

Hint: Long run variance

HAC — Consistent Estimator to estimate $X'\Omega X$



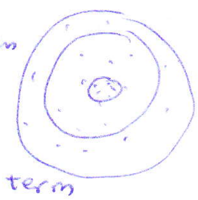
Same as before but what Dummy? — important empirical strategy

$d_i = \begin{cases} 0 & \text{if ie} \\ 1 & \text{if ie} \end{cases}$

- asian: if english at home income higher
- Construct different dummies and try to explain why structured breaks

$y_i = a_1 + a_2 d_i + b_1 x_i + b_2 x_i d_i + u_i$

- The time series → you have ordering HAC
- Can we do ordering cross sectional? Grade, ID, ...
- Classify individual based on location (Spatial analysis)



- geographic → social interaction
  Spatial analysis with Error term and do analysis

Spatial auto Correlation

$y_i = \rho y_{-i} + x_i \underline{\beta} + u_i$

spatial Coeffi

$\Rightarrow$ Breaks = Dummies

Testing → Lecture notes missing

$y = X\beta + u$

$H_0: \beta = 0$

$H_A: \beta \neq 0$ — Complementary set of H_0

- null hypothesis set must be bounded; alternative we don't care — we test on boundaries

$\Rightarrow H_0: \beta \geq 0$ ✓
$H_A: \beta < 0$ — one sided

$H_0: \beta > 0$
$H_A: \beta \leq 0$ $\Rightarrow$ Cannot test since not bounded

$t = \dfrac{\hat{\beta} - ⊛}{\sqrt{V(\hat{\beta})}} \rightarrow t = \dfrac{\hat{\beta} - 0}{\sqrt{V(\hat{\beta})}}$ to test $\beta \geq 0$

size
power → more powerful is better
type 1 Error:

Size: rejection rate of null when null is true (type 1 Error) → larger t-test (has relation with this: larger t-test as N increases)

due to lower variance relative to what it should be

Econometrics April 15 (5)

- Significance 5% } → ~~two~~ ~~one~~ One side
  t-stat 1.65
  1.96 → two side

$N \to \infty$, 100% reject null
⇓ $100 - 5 = 95\%$ size distortion
Shrink the variance ⇒ larger t-stat ↑

⇒ for power of the test you need more data
power: rejection rate of the null when null is false. ⇒ max power = 1
increase sample size → higher convergence rate

- in practice do not use gls
  if want to publish paper use gls
  in experiment people use non parametric
- theoretically gls var is lower than OLS
  practically gls var is greater than ols
  when estimate Ω → total # unknown increase Exponentially (sacrifice deg of freedom)
  ⇒ Convergence rate lower than OLS ⇒ more size distortion
- Econometric theory → $T \to \infty$, $N \to \infty$ total # of unknown does not matter
→ gls more efficient than pols
  ↳ more likely reject → due to more size distortion and not power of the test
⇒ size distortion means:
  true 2.5 but you use 1.96
⚠ using gls is way of cheating

Delta method

$\frac{\hat{\alpha}}{\hat{\beta}} = 2$ $y = X_1\alpha + X_2\beta + u$ restriction = 1, order 1

↳ t-dist → Normal
F - X^2

t-statistics → $\frac{\hat{\beta}}{\sqrt{V(\hat{\beta})}} \xrightarrow{d} N(0,1)$

$y = X\beta + u$
- sample small → only way to know $u \sim N(0, \sigma_u^2)$ Follow t-dist
- Exact sample theory

(6)

$\alpha = 2$ & $\beta = 2$ two restrictions

$$\sqrt{N}\begin{bmatrix}\hat{\alpha} - \alpha \\ \hat{\beta} - \beta\end{bmatrix} \xrightarrow{d} N(0, (2\times 2))$$

Restriction

$$R\begin{bmatrix}\hat{\alpha} - \alpha \\ \hat{\beta} - \beta\end{bmatrix}\underset{2\times 1}{}\ [\ \ ]'_{2\times 1} R \longrightarrow X^2_{R-1}$$

end of Review

New stuff

GMM not thaught litrature huge
Limited rand var litrature huge
treatment effect panel logit, propensity score

Econ 3 MLE → programming

OLS you can work on paper
MLE cannot solve by paper, just by program

in this term Not MLE, but GMM
minimum distance estimator

- Standard estimation technique
Econ theory: $\arg\max_{x_i} u(x_i)$
First order Condition $\boxed{\frac{\partial u}{\partial x} \Rightarrow x_i^* \Leftarrow \frac{\partial u(x_i^*)}{\partial x} = 0}$
in binary choice use utility, but gmm uses FOC
$\arg\max E\, u(x_i)$
⇒ $\boxed{E\left(\frac{\partial u(x_i)}{\partial x}\right) = 0 \quad \text{moment Condition}}$
if you put x_i^* in your data FOC = 0

$\frac{\partial u}{\partial x}(x_i^*) = f(x_i^*, \theta)$ - heavy formation parameter

→ Gmm: estimate $\hat{\theta}$ using x_i^*
- Risk aversion
- fixed effect
- GMMIV

Example) $\arg\max u(x_i, \theta) = \frac{x_i^{1-\theta}}{1-\theta}$

$u(x_i, y_i, \theta) = (x_i^\theta + y_i^\theta)^{1/\theta}$

You can have weighting $((1-\alpha)x_i^\theta + \alpha y_i^\theta)^{\theta}$

$\frac{\partial u}{\partial x_i} = 0$ $x_i^{-\theta} = 0 \Rightarrow E(x_i^{-\theta}) = 0$

$\Rightarrow \frac{1}{n}\sum x_i^\theta \xrightarrow{P} 0$

$\alpha + \beta x_i + \gamma x_i^2 + u_i$

$\beta + 2\gamma x_i = 0$ $x_i = -\frac{\beta}{2\gamma}$

$\frac{1}{n}\sum x_i = -\frac{\beta}{2\gamma} \Rightarrow \frac{\sum x_i}{n} + \frac{\beta}{2\gamma} = 0$

here to identify separately you get second moment

Condition → FOC
Constraint
multiple moment evaluation

19 April | Econometrics tutor

①

(part ~~II~~)

① $y_{it} = \alpha_i + \rho y_{it-1} + \beta x_{it-1} + e_{it}$

α_i ↳ Fix effect

mean over T $\frac{\sum_t y_{it}}{T} = \alpha_i + \rho \frac{\sum_t}{T} y_{it-1} + \frac{\beta}{T}\sum x_{it-1} + \frac{\sum e_{it}}{T}$ ②

① − ② ⇒ $\underbrace{y_{it} - \frac{\sum_t}{T} y_{it}}_{\tilde{y}_{it}} = \rho \underbrace{(y_{it-1} - \frac{\sum_t y_{it-1}}{T})}_{\tilde{y}_{it-1}} + \beta \underbrace{(x_{it-1} - \frac{\sum_t x_{it-1}}{T})}_{\tilde{x}_{it-1}}$

$+ \underbrace{(e_{it} - \frac{\sum_t}{T} e_{it})}_{\tilde{e}_{it}}$

$\tilde{y}_{it} = \rho \tilde{y}_{it-1} + \beta \tilde{x}_{it-1} + \tilde{e}_{it}$

show $\underset{N\to\infty}{\text{plim}} (\hat{\beta}_{FE} - \beta) = -\underset{N\to\infty}{\text{plim}} (\hat{\rho}_{FE} - \rho) \underset{N\to\infty}{\text{plim}} \frac{\sum \tilde{x}_{it-1}\tilde{y}_{it-1}}{\sum \tilde{x}^2_{it-1}}$

$\boxed{\hat{\rho}_{FE}\tilde{y}_{it-1}}$ trick

$\tilde{y}_{it} - \hat{\rho}_{FE}\tilde{y}_{it-1} = (\rho - \hat{\rho}_{FE})\tilde{y}_{it-1} + \beta \tilde{x}_{it-1} + \hat{e}_{it}$

$= \beta \tilde{x}_{it-1} + \underbrace{(\rho - \hat{\rho}_{FE})\tilde{y}_{it-1} + \tilde{e}_{it}}_{\text{error term}}$

$\text{plim}\,\hat{\beta}_{FE} - \beta = \frac{\sum^{NT} \tilde{x}_{it-1}[(\rho - \hat{\rho}_{FE})\tilde{y}_{it-1} + \tilde{e}_{it}]}{\sum^{NT}\tilde{x}^2_{it-1}}$

$= \text{plim}\frac{(\rho - \hat{\rho}_{FE})\sum_{NT}\tilde{x}_{it-1}\tilde{y}_{it-1}}{\sum^{NT}\tilde{x}^2_{it-1}} + \underbrace{\text{plim}\frac{\sum^{NT}\tilde{x}_{it-1}\tilde{e}_{it}}{\sum^{NT}\tilde{x}^2_{it-1}}}_{=0}$

$E(x_{it}e_{js}) = 0$

x_{it}, e_{js} are independent

$\tilde{x}_{it-1}, \tilde{e}_{is}$ are also independent

$\text{plim}(\hat{\beta}_{FE} - \beta) = \text{plim}(\rho - \hat{\rho}_{FE})\frac{\sum^{NT}\tilde{x}_{it-1}\tilde{y}_{it-1}}{\sum \tilde{x}^2_{it-1}}$

$= -\text{plim}(\hat{\rho}_{FE} - \rho)\frac{\sum^{NT} x_{it-1}\tilde{y}_{it-1}}{\sum \tilde{x}^2_{it-1}}$

$\boxed{\text{part I}}$ Data Generating process DGP

$y_{it} = y_{it-1} + e_{it}$

$e_{it} \sim iid(0, \sigma^2)$

regression $y_{it} = \alpha + \rho y_{it-1} + e_{it}$ ①

pols → mean $\frac{1}{NT}\sum y_{it} = \alpha + \rho\frac{\sum y_{it-1}}{NT} + \frac{\sum}{NT} e_{it}$ ②

① − ② ⇒ $y_{it} - \frac{\sum}{NT} y_{it} = \rho \boxed{(y_{it-1} - \frac{\sum}{NT} y_{it-1})}$

$+ \boxed{(e_{it} - \frac{\sum e_{it}}{NT})}$ ↳ Correlated ⇒ need IV

⇒ $\hat{\rho}_{POLS} - \rho = \frac{\sum (e_{it} - \frac{\sum e_{it}}{NT})(y_{it-1} - \frac{\sum}{NT} y_{it-1})}{\sum (y_{it-1} - \frac{\sum y_{it-1}}{NT})^2}$

②

IV can ensure consistency

$\text{plim}(\hat{\beta} - \beta) = \frac{\sum xu}{\sum x^2} = 0$ $y = x\beta + u$

Consistent when x, u are independent

x, u are correlated IV

$\text{plim}(\hat{\beta} - \beta) = \frac{\sum zu}{\sum xu} = 0$

Q 1: $N \to \infty$

Fixed T

$y_{it} = y_{it-1} + e_{it} = y_{it-2} + e_{it-1} + e_{it}$

$= y_{it-3} + e_{it-2} + e_{it-1} + e_{it} = e_{it} + e_{it-1} + e_{it-2} \cdots$

$= \sum_{j=0}^{\infty} e_{it-j}$

$y_{it} = \sum_{j=0}^{\infty} e_{it-j}$

$y_{it-1} = \sum_{j=0}^{\infty} e_{it-1-j}$

$e_{it} = O_p(1)$ since Normal

for fixed T $y_{it-1} = O_p(1)$

$\hat{\rho}_{POLS} - \rho = \frac{\frac{1}{N}\sum_i\sum_t (y_{it-1} - \frac{\sum_i\sum_t y_{it-1}}{NT})(e_{it} - \frac{\sum e_{it}}{NT})}{\frac{1}{N}\sum_i\sum_t (\underbrace{y_{it-1}}_{O_p(1)} - \underbrace{\frac{\sum_i\sum_t y_{it-1}}{NT}}_{O_p(\frac{1}{\sqrt{NT}})})^2}$

$\underbrace{\qquad\qquad}_{O_p(1)}$

Denominator $= \frac{1}{N}\sum_i \underbrace{(\sum_t O_p(1))}_{O_p(1)\text{ due to fixed }t} = \frac{\sum}{N} O_p(1) = O_p(1)$ Constant

Numerator: $\sum_i\sum_t (y_{it} - \frac{\sum_i\sum_t}{NT} y_{it-1})(e_{it} - \frac{\sum\sum}{NT} e_{it})$

$= \underbrace{\frac{1}{N}\sum_i\sum_t y_{it-1}e_{it}}_{=0} - \frac{1}{N}\frac{1}{NT}\sum_i\sum_t y_{it-1}\sum e_{it} = \frac{-\frac{1}{N}O_p(1)}{O_p(1)}$

$\underbrace{\frac{1}{\sqrt{NT}}\sum\sum_t y_{it}}_{O_p(1)}$ $\underbrace{\frac{1}{\sqrt{NT}}\sum_i\sum_t e_{it}}_{O_p(1)}$

So as $N \to \infty$ but fix T ⇒ Consistent

trick → Fixed T

$\boxed{\text{Q2}}$ fixed effect reg

$\frac{\sum}{T} y_{it} = \alpha_i + \rho\frac{\sum}{T} y_{it-1} + \frac{\sum}{T} e_{it}$

$y_{it} - \frac{\sum_t}{T} y_{it} = \rho(y_{it-1} - \frac{\sum}{T} y_{it-1}) + (e_{it} - \frac{\sum}{T} e_{it})$

③

$$\hat{\rho}_{FE} - \hat{\rho} = \frac{\frac{1}{N}\Sigma_i \Sigma_t \left(y_{it-1} - \frac{\Sigma}{T} y_{it-1} \right)\left(e_{it} - \frac{\Sigma e_{it}}{T}\right)}{\frac{1}{N}\Sigma_i\Sigma_t\left(y_{it-1} - \frac{\Sigma_t}{T} y_{it-1}\right)}$$

$O_p(1)$ from Q1

Numerator:

$$\frac{1}{N}\Sigma_i \Sigma_t\left(y_{it-1} - \frac{\Sigma}{T} y_{it-1}\right)\left(e_{it} - \frac{\Sigma}{T} e_{it}\right)$$

$$= \frac{\Sigma_i \Sigma_t}{N} \underline{y_{it-1}} \underline{e_{it}} - \frac{\Sigma_i}{N}\left(\underbrace{\frac{\Sigma}{T} y_{it-1}}_{O_p(\frac{1}{\sqrt{T}})}\right)\left(\Sigma e_{it}\right) \rightarrow O_p(\sqrt{T})$$

plim = 0 ← independent

$e_{it} \rightarrow O_p(1)$

$y_{it} = O_p(1)$

T = fixed

Second term

$$-\frac{1}{N}\Sigma_i O_p(1) = O_p(1)$$

since it is not mean zero

⇒ Final result: $\frac{O_p(1)}{O_p(1)} = O_p(1)$: not consistent

$$\underset{N \to \infty}{\text{plim}} (\hat{\rho}_{FE} - \rho) \neq \infty$$

④

Econometrics → 22 April ①

- Delta method → tailor expansion method

$$\begin{cases} y_i = \alpha + \beta x_i + \gamma z_i + u_i \\ H_0: \frac{\alpha}{\beta} = 1 \end{cases}$$

Harder test 4

Traps treasure hunt

for $\beta = 0$ & $\gamma = 1$ → Chi sq F-test

- understand Exactly what is going on

part I: Definition — test 4, final, qualifier

part II: Dummies — all from previous final Exams

part III: True false — Explain mathematic & formula

part IV: Derivation: basic not in test 4

- test 4 only harder ones

test 4

R.D. test 3 _ makeup_ tutorial, today's material

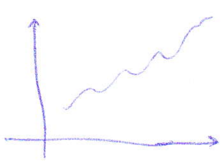
---

professor said → below average

those A: did not practice and forget and exponentially decrease memory

dissertation: logic, idea, how write, equations

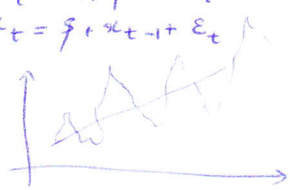
derivation: mathematical skills



deterministic trend

$x_t = \alpha + \beta t + u_t$

$x_t = \delta + x_{t-1} + \varepsilon_t$



Cointegration:

integration of: $x_t \sim I(1,t)$ non stationary; $x_t \sim (\mu, \sigma_x^2)$ W. stationary

$$\begin{cases} \underbrace{x_t}_{I(1)} = \theta_t + \underbrace{\varepsilon_t}_{I(0)} \\ y_t = \alpha\theta + \eta_t \end{cases}$$

$\alpha x_t - y_t = \alpha\varepsilon_t - \eta_t$ → stationary as a result of nonstationary & weakly stationary integration

---

Granger Causality

x_t y_t z_t

- behavior may be caused by each of them
- whether money supply causes GDP

money supply does not Granger Cause GDP but the Causality was reverse ②

how test? lagged variable

Vector autoregressive Process (VAR)

VAR1: if $\neq 0$ Granger Cause — lagged variable Caused current variable

$$\begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix}$$

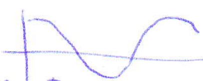
$\varepsilon_{it} \sim (0, \Sigma)$

Σ → Cholesky or spectral decomposition

$$\begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ 0 & b_{22} & b_{23} \\ 0 & 0 & b_{33} \end{pmatrix} \begin{pmatrix} u_{1t} \\ u_{2t} \\ u_{3t} \end{pmatrix}$$

→ independent

- impulse response analysis
- give shock and see what happens



- check what happen to GDP, Participation rate, ...

popular in practice

---

① Wisconsin Bruce Hansen → lecture notes of Econometrics (Condensed)

② Green book

- review his notes

G1: observe $y_{it} = \mu_{y_i} + y^{\circ}_{it}$, $x_{it} = \mu_{x_i} + x^{\circ}_{it}$ → time varying

$\mu_{y_i} = a + b\mu_{x_i} + \epsilon_i$

$y^{\circ}_{it} = \beta x^{\circ}_{it} + u_{it}$

Fix effect: $y_{it} = \alpha_i + \gamma x_{it} + e_{it}$

$\hat{\gamma}_{FE} = \beta$ — b: cross sectional variation

within transformation: $\tilde{y}_{it} = y_{it} - \frac{\sum y_{it}}{T}$

$y_{it} = \mu_{y_i} + y^{\circ}_{it}$

$\frac{\sum y_{it}}{T} = \mu_{y^{\circ}_{it}} + \frac{\sum y^{\circ}_{it}}{T}$

$\tilde{y}_{it} = \tilde{y}^{\circ}_{it}$

means in fix effect you just care about time

---

to calculate b you just take time series average

$\frac{\sum y_{it}}{T} = \mu_{y_i} + \frac{\sum y^{\circ}_{it}}{T} \to O_p(\frac{1}{\sqrt{T}})$

$\frac{\sum y_{it}}{T} = a^* + b^*(\frac{\sum x_{it}}{T}) + \text{error}$

$b^* \xrightarrow{p} b$ as $T \to \infty$, $N \to \infty$ ✓ much more important

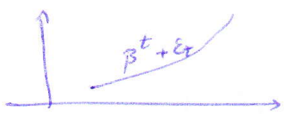
$y_{it} = a_i + x_{it}$, $x_{it} = \rho x_{it-1} + u_{it}$

$\rho = 1$

$y_{it} = \alpha y_{it-1} + e_{it}$

$\hat{\alpha} = ?$

- if series is integrated you should take first difference
- take log because GDP has exponential growth and ln makes it linear



then take first difference since they do not want to check whether series is stationary or DT

$y_{it} = a_i + x_{it}$

$- y_{it-1} = a_i + x_{it-1}$

$\Delta y_{it} = \Delta x_{it} = u_{it}$ → makes it stationary

---

$y_{it} = a_i + \rho y_{it-1} + e_{it}$ use this

$H_0 : \rho = 1$ needs I

$(\hat{\rho}_{FE} - 1) \xrightarrow{d}$

$$\hat{\rho} = \frac{\sum^{NT} y_{it} y_{it-1}}{\sum^{NT} y^2_{it-1}} = 1 + \frac{\sum^{NT} y_{it-1} u_{it}}{\sum y^2_{it-1}} =$$

1. $\text{plim}_{N \to \infty} \frac{1}{N} \sum^N \sum^T y^2_{it-1}$
2. $E\left(\frac{1}{N} \sum^N \sum^T y_{it-1} u_{it}\right)^2$

Cramer wald $\frac{(2)}{(1)}$

$y_{it} = y_{it-1} + u_{it} = y_{it-2} + u_{it-1} + u_{it} = u_{i1} + u_{i2} + u_{i3} + \dots + u_{it}$

not to infty ←

non stationary: since if past to infty total infty

if $y_{i0} \neq 0$ then Cauchy dist

here $y_{i0} = 0$

- means you made mistake in the past and you don't forget it

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

$\sum^T y^2_{it-1} = u_{i1}^2 + (u_{i1} + u_{i2})^2 + (u_{i1} + u_{i2} + u_{i3})^2 + \dots$

$= (T-1) u_{i1}^2 + (T-2) u_{i2}^2 + \dots + u^2_{it-1}$ + Cross prod

$E \frac{\sum^N}{N}$: $\text{plim} = \lim_{N \to \infty} E\left(\frac{\sum^N}{N}(\ )\right) = \sigma_u^2 [1 + \dots + T-1] = \frac{\sigma_u^2}{2} T^2 + O(T)$

$\sum t = \frac{T^2}{2} + O(T)$

$\sum t^2 = \frac{T^3}{3} + O(T^{-2})$ } → integral $\int u^x dx$

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

$E\, y^2_{it-1} u^2_{it} = \sigma_u^2 E\, y^2_{it-1} = (t-1) \sigma_u^2 \sigma_u^2$

$E\left(\sum^T y^2_{it-1} u^2_{it}\right) = (\sigma_u^4) \sum_{t=1}^{T-1} t = \sigma_u^4 \frac{T^2}{2}$

the rest you can do by yourself

- There is no nickel bias since $\frac{\sum^{NT} y_{it-1} u_{it}}{\sum y^2_{it-1}}$ has only good term and there is no bad term

---

$\sum \tilde{y}_{it-1} \tilde{u}_{it} = \sum y_{it-1} u_{it} - \frac{1}{T}\left(\sum y_{it-1}\right)\left(\sum^T u_{it}\right)$

- when there is no auto correlation then second term will be zero or $\rho = 0$
- Bad term is src of Nickel bias

- True - false questions need (1) understanding (2) derivation - both

---

Q3: check lecture note

$(X'X)(X'\Omega X)(X'X)^{-1} \cdot NT$ → This is true variance

- General formuler for panel
- for time series get rid of N

$\Omega = E(uu')$, $(N \times T) \times (N \times T)$

define u: $u = \begin{bmatrix} u_{11} \\ u_{12} \\ \vdots \\ u_{1T} \\ u_{21} \\ \vdots \\ u_{NT} \end{bmatrix}$

if no serial Correlation, no Cross sectional: $\Omega = \sigma_u^2 I$

5

$V(\hat\beta_{FE}) = (\tilde X'\tilde X)^{-1}(\tilde X'\Omega\tilde X)(\tilde X'\tilde X)^{-1}\cdot NT$ if No Ser Corr / Cross Sec

$V(\hat\beta_{FE}) = (\tilde X'\tilde X)^{-1}\sigma_u^2 I \cancel{(\tilde X'\tilde X)}\cancel{(\tilde X'\tilde X)^{-1}}\cdot(NT)$

→ Formula given in Q4

$= V(\hat\beta_{wg}) = \cancel{(\tilde X'\tilde X)^{-1}\tilde X' I \tilde X(\tilde X)}$

$$\frac{\sum^{NT}_{NT} \hat u_{it}^2}{\left(\sum^{NT}_{NT}(x_{it} - \sum^{T}_{T} x_{it})^2\right)^*}$$

$$\begin{bmatrix} [A] & & \\ & [A] & 0 \\ 0 & \ddots & \\ & & [A] \end{bmatrix} \quad A = \sigma_u^2 \lambda \begin{bmatrix} 1 & \rho & \rho^2 \dots \\ \rho & 1 \dots & \\ & \ddots & \\ & & 1 \end{bmatrix}$$

$\hat u_{it} = \hat\rho \hat u_{it} + \text{error}$ ⟶ Biased Nickel

$T << N$

panel Cov matrix

Bad term $\boxed{\text{Part I}}\ \boxed{\text{POLS}}$

as $N,T \to \infty$ $O(\frac{1}{\sqrt{NT}})$ will go zero

so we will not have Nickel bias

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

$$\hat\rho_{POLS} = \rho + \frac{\sum^{NT}(y_{it-1} - \sum^{NT} y_{it-1})(u_{it} - \frac{\sum^{NT} u_{it}}{NT})}{\sum^{NT}(y_{it} - \frac{\sum^{NT} y_{it-1}}{NT})^2}$$

denom is okay

$$A = \sum^{NT}(y_{it-1} - \frac{\sum^{N,T}}{NT} y_{it-1})(u_{it} - \frac{\sum^{N,T} u_{it}}{NT})$$

$$= \sum^{NT} y_{it-1}u_{it} - \frac{1}{NT}(\sum^{NT} y_{it-1})(\sum^{NT} u_{it})$$

$\text{plim}_{N\to\infty} A = 0$ $\quad$ $a_i = a$ ⟶ we fix it

---

$\boxed{\text{Q2}}\ \boxed{\text{Part 1}}\ \boxed{\text{POLS}}$

$a - a_i$ ⟶ will go into fixed effect

$e_{it} = (a_i - a)(1-\rho) u_{it}$

$y_{it} = a_i + x_{it-1}$ NO

$e_{it} + (a_i - a)(1-\rho) + u_{it}$

$\text{plim}_{N\to\infty} \sum^{NT} y_{it-1} e_{it} \neq 0 = \sigma_a^2(1-\rho)$

even good term is not Consistent

6

$\boxed{\text{Part II}}$ Dynamic Panel Regression with Fixed Effect

$\boxed{\text{Q4}}$ you need to take first difference and only then use IV

β_{FDIV}

$$\Delta y_{it} = \rho\Delta y_{it-1} + \Delta u_{it}$$

$$\underbrace{\underline{y_{it-1} - y_{it-1}} \quad \underline{u_{it} - u_{it-1}}}_{\text{Correlate}}$$

two $\begin{cases} y_{it-2} \\ \Delta y_{it-2} = y_{it-2} - y_{it-3} \end{cases}$

Singular _ variance goes to infinity

$\begin{cases} y_{it} - y_{it-2} = u_{it-1} \\ y_{it-2} - y_{it-1} = u_{it-2} \end{cases}$ ⟶ independent

long distance ~~is~~ good idea, but still in Consistent

- Strong Exogeneity : is reason that $E(u_{it}u_{is}) = 0$

For Part III bonus

$\begin{cases} \Omega^{-1} = PP' \quad \text{decomposition} \\ P'y = P'X\beta + P'u \\ y^* = X^*\beta + u^* \end{cases}$

$\hat\beta_{GLS} = (X^{*\prime}X^*)^{-1}(X^{*\prime}y^*)\, y_{it}$

$\boxed{\text{Test 4}}$

1. True/False
2. derivation, makeups tutorial, Sample test 3, lecture notes

whole session we will have Exam

Questions not hard compared to Test 3

sim

Final Cumulative

- everything
- Definition
- Dummy
- ordering
- ...

May 13 8:00

location

Qualifier Exam: Same as final