Econometrics - 20.01.26

I Cross section - aim to arrive at → cause-effect relationship ⇒ proxy variable; High correlation → consumption, income, savings → collect data; cannot be spurious correlation → observational studies → with unobserved factors in model; Experimental studies & with unobs u - cause & effect might not entirely established

1. Bivariate - CLRM
2. Multivariate - MLRM
3. Inference, Hypothesis Testing,
4. Relaxing assumptions - heteroskedasticity, Autocorrelation
6. Endogeneity - Experimental studies → DiD, propensity score matching, RCTs.
5. Simultaneous eq^n system & Instrumental Variable
7. Limited Dependent variable Model - Logit, Probit, Tobit

II. Time Series - stationarity, AR(1), ARMA, ARIMA, ARCH, GARCH.

III. Panel Data

References:

1. Introductory Econometrics - Wooldridge (7e).
2. Econometric Analysis of cross section & panel data - Wooldridge.

Causal inference:

3. Mastering metrics - Angrist & Pischke
4. Mostly Harmless econometrics - Angrist
   a) Causal inference - the mixtape - Scott Cunningham.

5. Time Series Analysis - Hamilton
6. Econometrics - Baltagi

ssc install bcuse

bcuse wage1

save

20 January 2026.

Recap Ordinary Least Square - Method of Estimation.

- Population distribution (parameter θ: unknown).

↓

Sample (statistic).

↓

Sample mean ($\bar{X}$)

↓

Draw an inference about unknown population parameter θ using sample statistic $\bar{X}$.

If $E(\bar{X}) = \theta$: unbiased.

$\underset{n\to\infty}{p\lim} E(\bar{X}) \to \theta$: Consistent

$V(\bar{X}) < V(\tilde{X})$: Minimum variance unbiased estimator (MVUE).

↓

Best

Apply these for bivariate OLS model parameters

→ $\hat{\beta}_1$ and β_0.

Assumptions of OLS:

1. Linear in parameter
2. Random sampling: $\{(x_i, y_i): i = 1, 2, \ldots, n\}$
3. $E(u|x_i) = 0$
4. $V(u|x) = \sigma^2 < \infty$
5. $Cov(x, u) = 0$.
6. There exists some variation in x_i ∀ $i = 1, 2 \ldots, n$

Holds even without perfect multicollinearity and ~~[illegible]~~ autocorrelation.

From the assumption $E(u|x) = 0$ → get to population regression function.

Model: $y_i = \beta_0 + \beta_1 x_i + u_i$ —(1)

$E(y_i|x_i) = \beta_0 + \beta_1 x_i + 0 = \beta_0 + \beta_1 X_i$ ~~$\beta_0 + \beta_1 x$~~ → Population Regression Function (PRF).

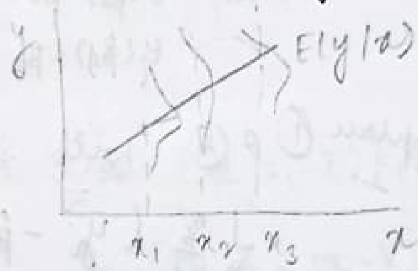
Note: y → random → ∵ u → random

x — given → exogenous → constant → not random.

Note: 1. $E(y_i|x = x_i)$ is a linear function of x

i) ⇒ one unit change in x, changes the conditional expected value of y (mean), given x, by β_1 units.

ii) For any given value of x_i, the distribution of y_i is centred about — $E(y_i|x = x_i)$



Minimise error — How?

1. Method of moments
2. Fitted values

OLS: Let, $\{(x_i, y_i) : i = 1, 2, \ldots, n\}$

$y_i = \beta_0 + \beta_1 x_i + u_i$ → stochastic factors other than x_i affecting y_i

↓ endogenous variable

x_i = obs. factor - education.

Objective: $\hat{\beta}_0, \hat{\beta}_1$ — Two methods - 1. Method of moments
2. Fitted values - differential calculus

• Let the popl^n parameter $= \theta$
• and $\theta = f(\mu)$: μ = population mean.

So, sample counterpart is $\bar{y}$

if $E(\bar{y}) = \mu$
$\text{plim}_{n\to\infty} E(\bar{Y}) \to \mu$ } replace μ with $\bar{Y}$.

Replace μ with $\bar{Y}$ → $f(\bar{y})$

Now, $\text{plim}_{n\to\infty} E(f(y)) = \theta$ and if μ is a linear function of $f(\mu)$ then, $E(f(\bar{Y})) = \theta$ → method of moments.

Apply this to eqn (1) - $y_i = \beta_0 + \beta_1 x_i + u_i$

use, $E(u_i | x = x_i) = 0$

or, $E(y_i - \beta_0 - \beta_1 x_i) = 0$ — (2)

Also, $\operatorname{cov}(u_i, x_i) = 0$

$\Rightarrow E(u_i x_i) = 0$

$\Rightarrow E(x_i(y_i - \beta_0 - \beta_1 x_i)) = 0$ — (3)

If sample projection on popl^n is correct, then unbiasedness & consistency holds.

Then $E(\hat{\beta}_0) = \beta_0$ and are consistent.
$E(\hat{\beta}_1) = \beta_1$

Replace ① & ② with sample counter part.

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0 \quad \text{— (4)}$$

$$\frac{1}{n}\sum_{i=1}^{n}(x_i(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)) = 0 \quad \text{— (5)}$$

from (4)

$$\bar{y} - \hat{\beta}_0 - \hat{\beta}_1 \bar{x} = 0.$$

$$\Rightarrow \bar{y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{x} \quad \text{— (6)} \qquad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

from (5) $\frac{1}{n}\sum_{i=1}^{n}(x_i y_i - \hat{\beta}_0 x_i - \hat{\beta}_1 x_i^2) = 0.$

$$\Rightarrow \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \hat{\beta}_0 \bar{x} - \hat{\beta}_1 \frac{1}{n}\sum_{i=1}^{n} x_i^2 = 0.$$

from (6)

$$\Rightarrow \frac{1}{n}\sum_{i=1}^{n} x_i y_i = (\bar{y} - \hat{\beta}_1 \bar{x})\bar{x} + \hat{\beta}_1 \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

$$\Rightarrow \frac{1}{n}\sum_{i=1}^{n} x_i y_i = \bar{x}\bar{y} - \hat{\beta}_1 (\bar{x})^2 + \hat{\beta}_1 \frac{1}{n}\sum_{i=1}^{n} x_i^2$$

$$\Rightarrow \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\bar{y} = \hat{\beta}_1 \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \hat{\beta}_1 (\bar{x})^2$$

$$\Rightarrow \quad Cov(x,y) = \hat{\beta}_1 \left(\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2 \right)$$

$$\Rightarrow \frac{S_{xy}}{S_{xx}} = \frac{Cov(x,y)}{Var(x)} = \hat{\beta}_1 = \frac{\frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\cancel{var}\ \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

Unbiasedness of $\hat{\beta}_1$ for β_1

check?

$$\hat{\beta}_1 = \frac{\frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2} = \frac{\frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})y_i}{\frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

$$\hat{\beta}_1 = \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})(\beta_0 + \beta_1 x_i + u_i) = \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})\beta_0 + \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})\beta_1 x_i + \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})u_i$$

$$= \frac{\beta_0}{n}\sum_{i=1}^{n} (x_i - \bar{x}) + \frac{\beta_1}{n}\sum_{i=1}^{n} (x_i - \bar{x})x_i + \frac{1}{n}\sum (x_i - \bar{x})u_i$$

$$= 0 + \frac{\beta_1}{n}\sum_{i=1}^{n} (x_i)^2 - \frac{\beta_1}{n}\sum_{i=1}^{n} \bar{x}x_i + \frac{1}{n}\sum (x_i - \bar{x})u_i$$

$$= 0 + \frac{\beta_1}{n}\sum_{i=1}^{n} (x_i)^2 - \frac{\beta_1 \bar{x}}{n}\sum_{i=1}^{n} x_i + \frac{1}{n}\sum (x_i - \bar{x})u_i$$

$$= \frac{\beta_1}{n}\sum_{i=1}^{n} (x_i)^2 - \frac{\beta_1}{n}(\bar{x})^2 + \frac{1}{n}\sum (x_i - \bar{x})u_i$$

$$\frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})(\beta_0 + \beta_1 x_i + u_i) = \frac{\beta_1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2 + \frac{1}{n}\sum (x_i - \bar{x})u_i$$

On dividing by $\frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2$

$$\frac{\frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})(\beta_0 + \beta_1 x_i + u_i)}{\frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2} = \beta_1 + \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})u_i$$

$$\hat{\beta}_1 = \beta_1 + \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})u_i$$

On taking the expectation

$$E(\hat{\beta}_1) = \beta_1 + \frac{1}{n}\sum_{i=1}^{n} E\left((x_i-\bar{x})u_i\right)$$

$$= \beta_1 + \frac{1}{n}\sum_{i=1}^{n}\sum_{i=1}^{n} E\left((x_i-\bar{x})(u_i-\bar{u})\right)$$

$$E(\hat{\beta}_1) = \beta_1 + 0$$

$$E(\hat{\beta}_1) = \beta_1$$

$\therefore$ unbiased.

Sign of $\hat{\beta}_1 = \dfrac{\overset{(+)\text{ or }(-)}{Cov(x,y)}}{\underset{(+)}{V(x)}}$. If $V(x)=0$, then $\hat{\beta}_1$ is undetermined

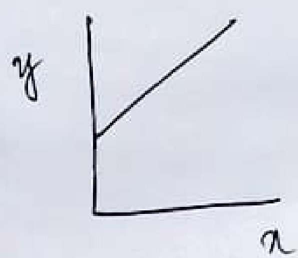
Method (2) - continious method (Differential calculus)

Population Model: $y_i = \beta_0 + \beta_1 x_i + \underbrace{u_i}_{\text{error}}$ — (1)

Estimated: $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ — (2) residual

$\Big\}$ not equal

$$Q = \sum_{i=1}^{n}\hat{u}_i = \sum_{i=1}^{n}(y_i-\hat{y}_i)^2$$



$$Q = \sum_{i=1}^{n}(y_i-\hat{\beta}_0-\hat{\beta}_1 x)^2$$

FOC: $\dfrac{\partial Q}{\partial \hat{\beta}_0} = 0$ or $\dfrac{\partial Q}{\partial \hat{\beta}_1} = -2\sum_{i=1}^{n}(y_i-\hat{\beta}_0-\hat{\beta}_1 x) = 0$

$$\Rightarrow \sum_{i=1}^{n} y_i - n\hat{\beta}_0 - \hat{\beta}_1\sum_{i=1}^{n} x = 0$$

$$\Rightarrow \boxed{\bar{y} = \hat{\beta}_0 + \hat{\beta}_1\bar{x}} \quad \text{— (3)}$$

Do the SOC as well, will get true for both.

Property of sample could be potential estimator.

It can be shown as,

$$\frac{\partial Q}{\partial \hat{\beta}_0} = \sum_i (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0$$

or, $\sum_i (y_i - \hat{y}_i) = 0$

$\sum_i u_i = 0$.

$\bar{u} = 0$

Average of residuals are zero, that does not mean $\bar{u}$ is a constant.

consider, the second normal equation,

$$\frac{\partial Q}{\partial \hat{\beta}_1} = \sum_{i=1}^{n} x_i (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) \text{ --- (4)}$$

from (3) & (4), we get, $\hat{\beta}_1 = \frac{\text{cov}(x,y)}{V(x)}$.

$$\hat{\beta}_1 = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

$$= \frac{\sum_{i=1}^{n} (x_i - \bar{x}) y_i}{\sum (x_i - \bar{x})^2}$$

$$= \beta_1 + \frac{\sum_{i=1}^{n} (x_i - \bar{x}) u_i}{\sum_i (x_i - \bar{x})^2} \text{ --- (6)}$$

from (6) $E(\hat{\beta}_1) = \beta$ $[\because E(u_i|x) = 0]$

$$V(\hat{\beta}_1) = \frac{\sigma^2}{SST_x}$$

from (3) $E(\hat{\beta}_0) = \beta_0$

$$V(\hat{\beta}_0) = \frac{\sigma^2 n^{-1} \sum_{i=1}^{n} x_i^2}{SST_x}$$

Both β_0 and β_1 have expectations and variation so both of them have a distribution.
we assumed it is normal.

where $SST_x = \sum (x_i - \bar{x})^2$

Algebraic properties of OLS:

1. $\sum_{i=1}^{n} \hat{u}_i = 0$ or $\bar{u} = 0$

2. $\sum x_i \hat{u}_i = 0$

3. $(\bar{x}, \bar{y})$ lie on the OLS regression line.

4. $\hat{u}_i = y_i - \hat{y}_i$

$SST_y = \sum_{i=1}^{n} (y_i - \bar{y})^2$

$SSE = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$

$SSR = \sum_{i=1}^{n} \hat{u}_i^2$

or $(y_i - \bar{y}) = (\underbrace{\hat{y}_i - \bar{y}}_{a} + \underbrace{\hat{u}_i}_{b})^2$

or $\sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum (\hat{y} - \bar{y})^2 + 2(\hat{y} - \bar{y})(\hat{u}_i) + (\hat{u}_i)^2$

$\Rightarrow \sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} (\hat{y} - \bar{y})^2 + \underbrace{2\sum_{i=1}^{n} (\hat{y} - \bar{y})(\hat{u}_i)}_{0} + \sum_{i=1}^{n} (\hat{u}_i)^2$

$SST_y = \underbrace{SSE_y}_{\text{explained term}} + SSR_y + 0$

then write in terms of x.

$\hat{y} - \hat{\beta}_0 - \hat{\beta}_1 x_i - u_i$

$\beta_0 - \beta_1 x_i - u_i - \hat{\beta}_0 - \hat{\beta}_1 x_i - u_i$

How much of the variation is explained.

GOODNESS OF FIT

$R^2 = \frac{SSE}{SST}$

$R^2 = [0, 1]$

Both the extreme are not good, means x is redundant

$R^2 \uparrow \Rightarrow$ Better fitted model

$R^2 \downarrow \Rightarrow$ bad fit

$R^2 = 1$ - if $SSE = SST$

$\Rightarrow$ Perfect fit but may not be real

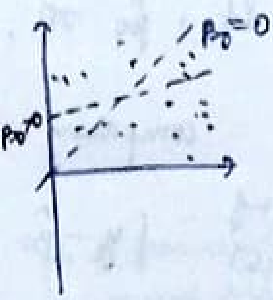
$\Rightarrow$ you are dealing with the population itself

$R^2 = 0$ if $SSE = 0$.

Regression through origin - special case

If $\beta_0 = 0$

$$y_i = \tilde{\beta}_1 x_i + u_i \quad \text{—(1)}$$



Fitted $\hat{y}_i = \hat{\beta}_1 x_i$ —(2)

Run OLS : $\sum_{i=1}^{n} \hat{u}_i^2 = \sum_{i=1}^{n} (y_i - \hat{\tilde{\beta}}_1 x_i)^2$ —(3)

$\beta_0 < 0$

Here compared to normal cases, you have only one foc.

$$\text{FOC : } \frac{\partial \sum_{i=1}^{n} \hat{u}_i^2}{\partial \tilde{\beta}_1} = \sum_{i=1}^{n} x_i (y_i - \hat{\tilde{\beta}}_1) = 0 \quad \text{—(4)}$$

$$\Rightarrow \sum_{i=1}^{n} x_i y_i - \cancel{\tilde{\beta}_1 x_i = 0} \; \tilde{\beta}_1 \sum_{i=1}^{n} x_i = 0$$

$$\Rightarrow \quad \hat{\tilde{\beta}}_1 = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2} \quad \text{—(5)}$$

from (5) - $\hat{\tilde{\beta}}_1 = \dfrac{\sum_{i=1}^{n} x_i (\hat{\tilde{\beta}}_1 x_i + u_i)}{\sum_{i=1}^{n} x_i^2}$

$$\hat{\tilde{\beta}}_1 = \tilde{\beta}_1 + \frac{\sum_i x_i u_i}{\sum_i x_i^2}$$

$E(\hat{\tilde{\beta}}) = \hat{\tilde{\beta}}_1 \longrightarrow$ unbiased.

Running Goodness of fit,

Case 1 : $R^2|_{\beta_0 \neq 0} = \dfrac{SSE}{SST} = 1 - \dfrac{SSR}{SST} = 1 - \left.\dfrac{\sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)}{SST}\right\} a$

Case 2 : $R^2|_{\beta_0 = 0} = 1 - \left.\dfrac{\sum_{i=1}^{n} (y_i - \hat{\tilde{\beta}}_1 x_i)^2}{SST}\right\} b$ showing the possiblity of negative R^2

Imposing the same conditions,

If, $\hat{\beta}_0 \neq 0$, $\tilde{\hat{\beta}}_1 = \hat{\beta}_1$

comparing a and b

Deviating by SST

$$\frac{(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2}{SST} < \frac{(y_i - \hat{\beta}_1 x_i)^2}{SST}$$

subtracting from 1

$$1 - \frac{(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2}{SST} > 1 - \frac{(y_i - \hat{\beta}_1 x_i)^2}{SST}$$

$$R^2 |_{\beta_0 \neq 0} > \tilde{R}^2 |_{\beta_0 = 0}$$

To address this, we use,

$$\tilde{R}^2 |_{\beta=0} = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{\beta}_1 x_i)^2}{\sum_{i=1}^{n} y_i^2}$$

29. Jan. 26.

CAUSAL INFERENCE

Ch-2 - wooldridge end of chapter

(1) $E(u|x) = 0$

(2) $E(xu|x) = 0$ } important assumption.

$\hookrightarrow cov(x,u) = 0$

$\hookrightarrow \rho_{xy} = 0$

OLS - cause-effect

$x \perp u$ i.e variations in x is not influenced in u or vice versa.

Eg: $wage = \beta_0 + \beta_1 e + u$ - here $e \perp u$ i.e $cov(e,u) = 0$

Means education is independent of the error term which is not exactly true, i.e, the are factors in u that might influence wage. These are called ommitted variables.

Omitted variable Bias -

from the example - experience, parents education, sectors, gender, social markup.

These many variables are omitted from the model then all of this is included in u, so there is a high correlation between wage and u.

so, unable to identify e as a sole factor causing an identification problem. And x variable - education is no more exogenous variable so $E(u|x) \neq 0$ and $E(xu|x) \neq 0$.

Example: $w = \beta_0 + \beta_1 edu + u$
└ continuous

making education discrete -

$edu = \begin{cases} \text{No schooling (illiterate)} = 0 \\ \text{Schooling (literate)} = 1 \end{cases}$

0 → Reference group, group you are going to measure
1 → interest group, main group

creating a binary variable -
loss of information

edu	bin_edu
	0
0	
10	1
15	1
20	1
10	1
0	0
1	1
0	0

β_1 was the average
now what is it

$w = \beta_0 + \beta_1 \overset{0\ 1}{\overset{\vee}{e}} + u \qquad e = \begin{cases} 0 \\ 1 \end{cases}$

$E(w|x=1) = \beta_0 + \beta_1 X_1 + E(u|x=1)$
$= \beta_0 + \beta_1$

$E(w|x=0) = \beta_0 + \beta_1(0) + E(u|x=0)$
$= \beta_0$

$E(w|x=1) - E(w|x=0) = \beta_1$

∴ β_1 is the difference b/w illiterate & literate.

Causal inference is usually taken to understand policy cause-effect or policy choice. It is borrowed from medical science to understand the effectiveness of medicine.
Now introducing treatment variable T.

$w = \beta_0 + \beta_1 T + u$ → everything exactly same

$T = \begin{cases} 0 \text{, not eligible group - who did not recieve T} \\ 1 \text{ eligible, who recieved T.} \end{cases}$

Problems of comparing 1s and 0s

→ more 1s and less 0s will end up in selective bias

- becomes like comparing apples & oranges.

$$\underbrace{E(w|x=1)}_{\text{motivated}} - E(w|x=0) = 0$$

↳ How do we solve this? - Randomisation

(1)	(0)
Treatment Group	Control group
C-19 vaccine	placebo

T|C

p - psychological disorder

$$E(w|T=1) - E(w|T=0)$$

$$= \beta_0 + \beta_1 + p_1 - \beta_0 - p_1$$

$= \beta_1$ - this β_1 is pure causal effect because there is no omitted variable

Formally written as:

$$w = \beta_0 + \beta_1 T + u \qquad v = u + p$$

$$\tilde{w} = \beta_0 + \beta_1 T + v$$

$$= \beta_0 + \beta_1 T + p + u$$

Before finding the causal effect - make it comparable.

$$= E(\tilde{w}|T=1) - E(\tilde{w}|T=0)$$

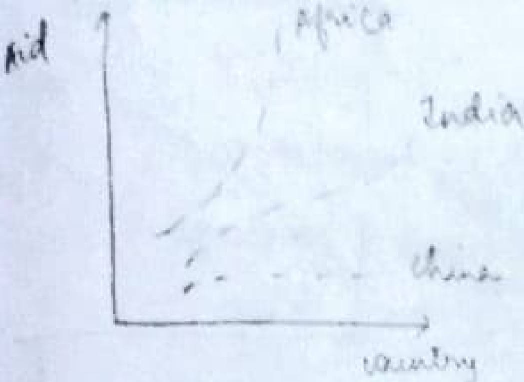
$$= \beta_0 + \beta_1 T + p - \beta_1 - p$$

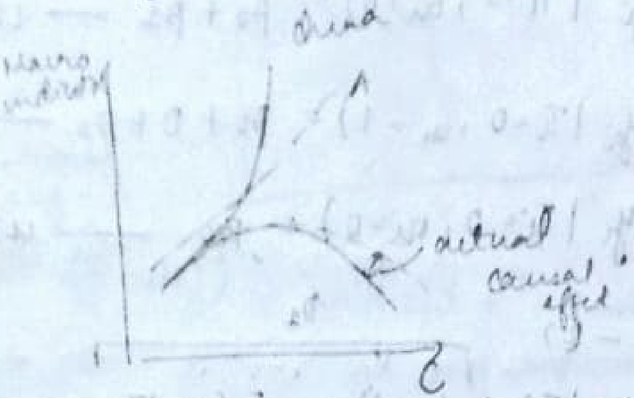
$= \beta_1$ → cause and effect relationship

→ A T E
Average Treatment effect

$\boxed{T=0 : \text{control counter factor}}$

ex In medical field placebo is easy but in social science how will you give placebo in such a scenario.
→ counter factual – a parallel hypothetical world.



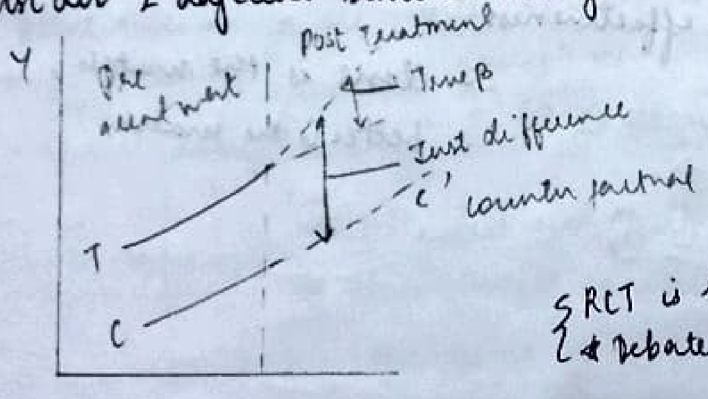


Methods to valid counterfactual group:

1. Randomized Control Trial (RCT) → field experiment
2. Difference in difference (DID) →
3. Propensity score matching (PSM)
4. Regression discontinuity design (RDD)

} Quasi – natural experiment

1. RCT
Consider 2 adjacent similar villages



To compare – you have to make counterfactual then compare

{RCT is said to be "golden-rule"}
{& Debate b/w Banerjee & Deaton on RCT}

2. DID – time is introduced:

$$y_i = \beta_0 + \beta_1 T_t + \beta_2 G_i + \underbrace{\gamma T_t \times G_i}_{\text{interaction term}} + u_i$$

pure causal effect

$T_t = \begin{cases} 0 \\ 1 \end{cases}$ pre and post group, $G_i \begin{cases} 0 \\ 1 \end{cases}$ } Have to take 2 difference one for time effect and other for time variate & time invariate

$$E(y_i \mid T_i=1, G_i=1) = \beta_0 + \beta_1 + \beta_2 + \gamma$$

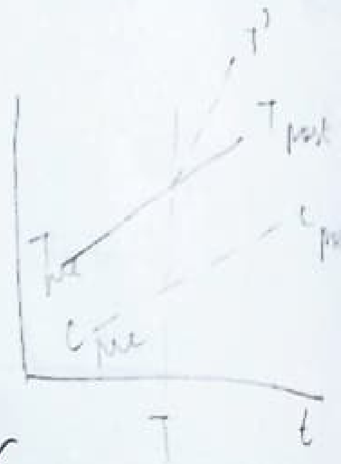
$$E(y_i \mid T_i=1, G_i=0) = \beta_0 + \beta_1 \quad \text{— (2)}$$

Time effect is captured

$$E(y_i \mid T_i=0, G_i=1) = \beta_0 + 0 + \beta_2 \quad \text{— (3)}$$

$$E(y_i \mid T_i=0, G_i=0) = \beta_0 \quad \text{— (4)}$$

2nd time effect is captured



(1) − (2)

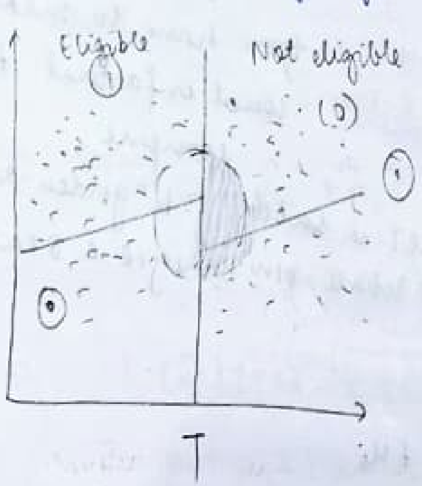
$$\overbrace{E(y_i \mid T_i=1, G_i=1) - E(y_i \mid T_i=1, G_i=0)}^{D_1} = \beta_2 + \gamma$$

(3) − (4)

$$\overbrace{E(y_i \mid T_i=0, G_i=1) - E(y_i \mid T_i=0, G_i=0)}^{D_2} = \beta_2$$

$$\{1 - 2\} - \{3 - 4\} = D_1 - D_2 = \gamma$$

3. RDD - consider policy effectiveness.

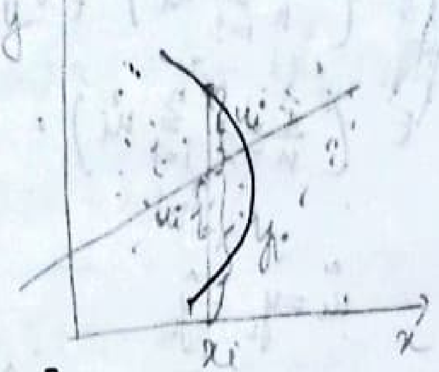


Closer is the match, better is the match.

Error Variance: Obj - estimate variance of error.

3 February · 2026.

u: Population errors.



Homoskedasticity:

$V(u|x) = \sigma^2$

$E(u|x) = 0 \rightarrow$ conditional avg = 0.

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

$= \sigma^2 - 0 = \sigma^2$

without σ^2 - cannot fit my regression - so to do this need to estimate σ^2.

$E(u^2) = \sigma^2$

$E\left(\frac{1}{n}\sum_{i=1}^{n} u_i^2\right) = \frac{1}{n}\sum E(u_i^2) = \frac{n\sigma^2}{n} = \underline{\underline{\sigma^2}}$

Homoskedastic $\Rightarrow$ constant variance

How do we estimate $\sigma^2 \rightarrow$ from sample $\rightarrow$ sample counterpart - $\hat{\sigma}^2$

How $\rightarrow$ using $\hat{u}_i$.

Substitute $\frac{1}{n}\sum_i u_i^2$ with sample counterpart $\rightarrow \frac{1}{n}\sum_{i=1}^{n} \hat{u}_i^2$

Then, Is $E\left(\frac{1}{n}\sum_{i=1}^{n} \hat{u}_i^2\right) = \hat{\sigma}^2$?? is this equivalent sample variance

- no, it is not unbiased
- is this sample variance?
- apply the logic of $S^2 = \frac{1}{n}\sum(X_i - \bar{X})^2$ is still biased.

but $s^2 = \frac{1}{n-1}\sum(X_i - \bar{X})^2$ is unbiased. — STATS. — here mean is adjusted for - the mean is restricted

we need an unbiased estimator

adjust by taking restriction of mean

$E\left(\frac{SSR}{n}\right) \rightarrow$ biased, so the unbiased estimator — 2 parameter restrictions

$\hat{\beta}_0, \hat{\beta}_1 \rightarrow$ 2 restrictions

$E\left(\frac{1}{n}\sum_{i=1}^{n}\hat{u}_i^2\right) \neq \sigma^2$: Biased

$E\left(\frac{1}{n-2}\sum_{i=1}^{n}\hat{u}_i^2\right) = \sigma^2$: Unbiased.

$\hat{u}_i = y_i - \hat{y}_i$ $\qquad y_i = \beta_0 + \beta_1 x_i + u_i$

$= \beta_0 + \beta_1 x_i + u_i - \hat{\beta}_0 - \hat{\beta}_1 x_i$ $\qquad \hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$

$\hat{u}_i = u_i - (\hat{\beta}_0 - \beta_0) - (\hat{\beta}_1 - \beta_1) x_i$

On taking the expectation –

$E(\hat{u}_i) = E(u_i) - E(\hat{\beta}_0) + E(\beta_0) - E[(\hat{\beta}_1 - \beta_1) x_i]$

$= 0 - \beta_0 + \beta_0 - 0$

$E(\hat{u}_i) = 0 \neq u_i$ $\quad$ So $\hat{u}_i$ is not an unbiased estimator of u_i.

Also, we have seen before, that – $\frac{1}{n}\sum_{i=1}^{n}\hat{u}_i^2$ is a biased estimator, because it does not consider two restrictions on the parameters – β_0 and β_1. and correspondingly,

$\frac{\partial \hat{u}_i^2}{\partial \hat{\beta}_1} = 0$ or $\sum_{i=1}^{n}\hat{u}_i = 0$ and $\frac{\partial \sum_{i=1}^{n}}{\partial \hat{\beta}_1} = 0$ or $\sum_{i=1}^{n} x_i \hat{u}_i = 0$

How do we show the adjustment algebraically –

From (2) $\hat{u}_i = u_i - (\hat{\beta}_0 - \beta_0) - (\hat{\beta}_1 - \beta_1) x_i$ —— (2)

Sample average – $\qquad \sum(\hat{\beta}_1 - \beta_1) x_i$

$\frac{1}{n}\sum_{i=1}^{n}\hat{u}_i = \frac{1}{n}\sum_{i=1}^{n} u_i - \frac{1}{\cancel{n}}\cancel{n}(\hat{\beta}_0 - \beta_0) - (\hat{\beta}_1 - \beta_1)\frac{1}{n}\sum_{i=1}^{n} x_i$ $\qquad \sum x_i$

$\bar{\hat{u}}_i = \bar{u} - (\hat{\beta}_0 - \beta_0) - (\hat{\beta}_1 - \beta_1)\bar{x}$ —— (3)

$2\cdot(nx_i)$

$\sum_{i=1}^{n} c = nc$ $\qquad \sum_{i=1}^{n} c\, x_i = c\left(\sum_{i=1}^{n} x_i\right)$

$2x$ $\quad \sum 2x$

(2) − (3)

$$\hat{u}_i - \bar{\hat{u}}_i = u_i - \cancel{(\hat{\beta}_0 - \beta_0)} - (\hat{\beta}_1 - \beta_1)x_i - \bar{u} - \cancel{(\hat{\beta}_0 - \beta_0)} - (\hat{\beta}_1 - \beta_1)\bar{x}$$

$$\hat{u}_i - \underset{\overset{\parallel}{0}}{\bar{\hat{u}}_i} = (u_i - \bar{u}) - (\hat{\beta}_1 - \beta_1)(x_i - \bar{x})$$

$$\hat{u}_i = (u_i - \bar{u}) - (\hat{\beta}_1 - \beta_1)(x_i - \bar{x})$$

On squaring both sides −

$$\hat{u}_i^2 = \left[(u_i - \bar{u}) - (\hat{\beta}_1 - \beta_1)(x_i - \bar{x})\right]^2$$

$$= (u_i - \bar{u})^2 + \left[(\hat{\beta}_1 - \beta_1)(x_i - \bar{x})\right]^2 - 2(u_i - \bar{u})(\hat{\beta}_1 - \beta_1)(x_i - \bar{x})$$

On summation on both sides −

$$\sum \hat{u}_i^2 = \sum_{i=1}^{n} (u_i - \bar{u})^2 + \sum_i \left[(\hat{\beta}_1 - \beta_1)(x_i - \bar{x})\right]^2 - 2\sum_i (u_i - \bar{u})(\hat{\beta}_1 - \beta_1)(x_i - \bar{x})$$

$$= (n-1)\sigma^2 + (\hat{\beta}_1 - \beta_1)^2 \underbrace{\sum_i (x_i - \bar{x})^2}_{\text{SST of } x} - 2(\hat{\beta}_1 - \beta_1)\underbrace{\sum_i (x_i - \bar{x})(u_i - \bar{u})}_{\text{cov}(x,y)}$$

~~$(n-1)\sigma^2 + (\hat{\beta}_1 - \beta_1)^2 \sum_i (x_i - \bar{x})^2$~~

On ~~taking expectation~~ =

$$E\left(\sum_{i=1}^{n} \hat{u}_i^2\right) = (n-1)\sigma^2 + E\underbrace{(\hat{\beta}_1 - \beta_1)^2}_{V(\hat{\beta}_1)} \cdot SST_x \quad -2E$$

$$= (n-1)\sigma^2 + \frac{\sigma^2}{SST_x} \cdot SST_x$$

but recall,

$$\hat{\beta}_1 = \beta_1 + \frac{\sum (x_i - \bar{x})(u_i - \bar{u})}{\sum (x_i - \bar{x})^2}$$

$$\Rightarrow \sum (x_i - \bar{x})^2 (\hat{\beta}_1 - \beta_1) = \sum (x_i - \bar{x})(u_i - \bar{u})$$

$$(n-1)\sigma^2 + (\hat{\beta}_1 - \beta_1)^2 \sum_i (x_i - \bar{x})^2 - 2(\hat{\beta}_1 - \beta_1)\sum_i (x_i - \bar{x})(x_i - \bar{u})$$

$$E\left(\sum_i \hat{u}_i^2\right) = (n-1)\sigma^2 + \frac{\sigma^2}{SST_x} \cdot SST_x - 2E\left[\underbrace{(\hat{\beta}_1 - \beta_1)^2}_{V(\hat{\beta}_1)}\right] \sum_i (x_i - \bar{x})^2$$

$$= (n-1)\sigma^2 + \sigma^2 - 2\frac{\sigma^2}{SST_x} \cdot SST_x$$

$$= (n-1)\sigma^2 + \sigma^2 - 2\sigma^2$$

$$= (n-2)\sigma^2$$

$$E\left(\frac{1}{n-2}\sum_{i=1}^{n} u_i^2\right) = \sigma^2$$

$$E\left(\frac{SSR}{n-2}\right) = \sigma^2$$

Variance of $\hat{\beta}_1$

$$V(\hat{\beta}_1) = \frac{\sigma^2}{SST_x}$$

Observe: $\hat{\sigma}^2 \rightarrow \frac{1}{n-2} \sum_{i=1}^{n} \hat{u}_i^2$

$$E[V(\hat{\beta}_1)] = E\left[\frac{\frac{1}{n-2}\sum_{i=1}^{n} u_i^2}{SST_x}\right] = \frac{\sigma^2}{SST_x}$$

$\therefore$ It is unbiased estimator.

Variance of $\hat{\beta}_0$: $= \frac{\sigma^2}{n} \frac{\sum_i x_i^2}{\sum (x_i - \bar{x})^2}$

$$E(V(\hat{\beta}_0)) = \frac{\sigma^2}{n}$$

standard deviation:

$$V(\hat{\beta}_1) = \frac{\sigma^2}{SST_x}$$

$$S.D(\hat{\beta}_1) = \frac{\sigma}{\sqrt{SST_x}} = \frac{\sigma}{(\sum_i (x_i - \bar{x})^2)^{1/2}}$$

Replace with standard error to find the estimator -

$$S.E(\hat{\beta}_1) = \frac{\hat{\sigma}^2}{[\sum (x_i - \bar{x})^2]^{1/2}}$$

Multiple Linear Regression Model

3. Feb. 2020.

Example: $w = f(\text{educ}, \text{experience})$ confounding variable, causal literature

$$w_i = \beta_0 + \beta_1 \, edu_i + \beta_2 \, exp_i + u_i$$

edu_i: main interest variable

exp_i: control variable, regressor, co-variate, independent variable.

∂ first var - interest variable

- these models - partial equilibrium models - $\dfrac{\partial w_i}{\partial edu_i} = \beta_1$ - ceteris paribus (other things constant).
  - also called level-level models.

- Log level → $\log w_i = \beta_0 + \beta_1 edu_i + \beta_2 exp_i + u_i$

$$\frac{1}{w}\frac{\partial w_i}{\partial edu_i} \qquad \frac{1}{w}\partial w = \beta_1 \partial edu_i$$

$$\frac{\dfrac{\partial w}{\partial edu_i}}{w} = \beta_1$$

- ~~log~~ level log model - $\log w_i = \beta_0 + \beta_1 \log edu_i + \beta_2 \log exp_i + u_i$

$$\Delta w = \beta_1 \Delta edu_i$$

The model: ~~$\hat{y}_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + u_i$~~ $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + u_i$

$$\tilde{Q} = \sum_{i=1}^{n} \hat{u}_i^2 = \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2})^2 = 0$$

FOC: $\dfrac{\partial \tilde{Q}}{\partial \hat{\beta}_0} = 2 \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2})(-1) = 0.$

$$\Rightarrow \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2}) = 0$$

$$\Rightarrow \sum_i y_i = n\hat{\beta}_0 + \hat{\beta}_1 \sum_i x_{i1} + \hat{\beta}_2 \sum_i x_{i2}$$

$$\Rightarrow \frac{1}{n}\sum_i y_i = \hat{\beta}_0 + \hat{\beta}_1 \frac{1}{n}\sum_i x_{i1} + \hat{\beta}_2 \frac{1}{n}\sum_i x_{i2}$$

$$\Rightarrow \boxed{\bar{Y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{x}_1 + \hat{\beta}_2 \bar{x}_2} \quad \text{---(1)}$$

$$\frac{\partial \tilde{R}}{\partial \hat{\beta}_1} = 2\sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2})(-x_{i1}) = 0$$

$$\Rightarrow 2\sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2})(x_{i1}) = 0$$

$$\Rightarrow \cancel{\sum_{i=1}^{n} y_i x_i - \hat{\beta}_0 x_{i1}} \quad \sum$$

$$\Rightarrow \sum_{i=1}^{n} \left(y_i x_{i1} - \hat{\beta}_0 x_{i1} - \hat{\beta}_1 x_{i1}^2 - \hat{\beta}_2 x_{i2}(x_{i1})\right) = 0.$$

Substituting $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}_1 - \hat{\beta}_2 \bar{x}_2$

$$\Rightarrow \sum y_i x_{i1} = \hat{\beta}_0 \sum x_{i1} + \hat{\beta}_1 \sum x_{i1}^2 + \hat{\beta}_2 \sum x_{i2} x_{i1}$$

$$\Rightarrow \sum y_i x_{i1} = (\bar{y} - \hat{\beta}_1 \bar{x}_1 - \hat{\beta}_2 \bar{x}_2) \sum x_{i1} + \hat{\beta}_1 \sum (x_{i1})^2 + \hat{\beta}_2 \sum x_{i2} x_{i1}$$

$$\Rightarrow \sum y_i x_{i1} = (\bar{y} - \hat{\beta}_1 \bar{x}_1 - \hat{\beta}_2 \bar{x}_2) n\bar{x}_1 + \hat{\beta}_1 \sum (x_{i1})^2 + \hat{\beta}_2 \sum x_{i2} x_{i1}$$

$$\Rightarrow \sum y_i x_{i1} = n\bar{y}\bar{x}_1 - n\hat{\beta}_1 \bar{x}_1^2 - n\hat{\beta}_2 \bar{x}_1 \bar{x}_2 + \hat{\beta}_1 \sum (x_{i1})^2 + \hat{\beta}_2 \sum x_{i2} x_{i1}$$

$$\Rightarrow \sum y_i x_{i1} - n\bar{y}\bar{x}_1 = \hat{\beta}_1 \left(\sum x_{i1}^2 - n\bar{x}_1^2\right) + \hat{\beta}_2 \left(\sum x_{i2} x_{i1} - n\bar{x}_1 \bar{x}_2\right)$$

$\div n$

$$\Rightarrow S_{1y} = \hat{\beta}_1 S_{11} + \hat{\beta}_2 S_{12} \quad \text{——} \quad (2)$$

where

$$S_{1y} = \sum_i (y_i - \bar{y})(x_{i1} - \bar{x}_1) \qquad S_{11} = \sum_i (x_{i1} - \bar{x})^2$$

$$= \sum_i x_{i1} y_i - n\bar{x}_1 \bar{y} \qquad = \sum x_{i1}^2 - n\bar{x}_1^2$$

$$S_{12} = \sum_i (x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2)$$

$$S_{12} = \sum x_{i1} x_{2i} - n\bar{x}_1 \bar{x}_2$$

Similarly, $\frac{\partial \tilde{Q}}{\partial \beta_2} = 2\sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2})(-x_{i2}) = 0$

$$\Rightarrow \sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2})(x_{i2}) = 0.$$

$$\Rightarrow \sum_{i=1}^{n} y_i x_{i2} - \sum \hat{\beta}_0 x_{i2} - \sum \hat{\beta}_1 \sum_{i=1}^{n} x_{i1} x_{i2} - \beta_2 \sum_{i=1}^{n} x_{i2}^2 = 0.$$

substituting $\hat{\beta}_0$, we get -

$$\Rightarrow \sum_{i=1}^{n} y_i x_{i2} - n\bar{x}_2 \bar{y} = \hat{\beta}_1 \left(\sum_i x_{i1} x_{i2} - n\bar{x}_1 \bar{x}_2\right) + \hat{\beta}_2 \left(\sum x_{2i}^2 - n\bar{x}_2^2\right)$$

$$\Rightarrow S_{2y} = \hat{\beta}_1 S_{21} + \hat{\beta}_2 S_{22} \quad \text{--- (3)}$$

where $S_{2y} = \sum_i (y_i - \bar{y})(x_{i2} - \bar{x}_2) = \sum_i x_{2i} y_i - n\bar{x}_2 \bar{y}$

$S_{22} = \sum_i (x_{i2} - \bar{x}_2)^2.$

So we have 2 equations & 2 unknowns $\Rightarrow$ unique sol^n exists.
So we have 2 equations by using successive elimination one can solve for these 2 eq^n

or $S_{1y} = \hat{\beta}_1 S_{11} + \hat{\beta}_2 S_{12}$

$S_{2y} = \hat{\beta}_1 S_{21} + \hat{\beta}_2 S_{22}$

Using matrix notation,

$$\begin{bmatrix} S_{1y} \\ S_{2y} \end{bmatrix}_{2\times1} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}_{2\times2} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix}_{2\times1}$$

$$S_{xy} = \underbrace{S_{xx}}_{\text{non singular matrix}} \hat{\beta}$$

$\Rightarrow S_{xx}^{-1}$ exists

$$\hat{\beta} = S_{xx}^{-1} S_{xy}.$$

For 3 equation,

$$S_{xx} = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix}$$

$$S_{xy} = \begin{bmatrix} S_{1y} \\ S_{2y} \\ S_{3y} \end{bmatrix}$$

$$\hat{\beta}_1 = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix}$$

Again $\hat{\beta} = S_{xx}^{-1} S_{xy}$

In matrix notation.

$$\bar{X} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{bmatrix}$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}' \bar{X}$$

$$RSS = S_{yy} - \hat{\beta}_1' S_{xy} \; ; \; S_{yy} = \text{scalar}$$

$$R^2_{y \cdot x_1 x_2 x_3} = \frac{\hat{\beta}' S_{xy}}{S_{yy}}$$

Let's take a simple numerical example:

$$5\hat{\beta}_1 + 15\hat{\beta}_2 + 25\hat{\beta}_3 = 20$$

$$15\hat{\beta}_1 + 55\hat{\beta}_2 + 81\hat{\beta}_3 = 76$$

$$25\hat{\beta}_1 + 81\hat{\beta}_2 + 129\hat{\beta}_3 = 109$$

$$\begin{bmatrix} 5 & 15 & 25 \\ 15 & 55 & 81 \\ 25 & 81 & 129 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} = \begin{bmatrix} 20 \\ 76 \\ 109 \end{bmatrix}$$

$R_2 - 3R_1 \rightarrow R_2$

$$\begin{bmatrix} 5 & 15 & 25 \\ 0 & 10 & 6 \\ 25 & 81 & 129 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ 109 \end{bmatrix} = \begin{bmatrix} 20 \\ 16 \\ 109 \end{bmatrix}$$

$R_3 - 5R_1 \rightarrow R_3$

$$\begin{bmatrix} 5 & 15 & 25 \\ 0 & 10 & 6 \\ 0 & 6 & 4 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} = \begin{bmatrix} 20 \\ 16 \\ 109 \end{bmatrix}$$

$R_3 - \frac{6}{10} R_2 \rightarrow R_3$

$$\begin{bmatrix} 5 & 15 & 25 \\ 0 & 10 & 6 \\ 0 & 0 & 0.4 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix} = \begin{bmatrix} 20 \\ 16 \\ -0.6 \end{bmatrix}$$

$$\hat{\beta}_3 = -0.6/0.4 = -1.5$$

$$10\hat{\beta}_2 + 6\hat{\beta}_3 = 16$$

$$\Rightarrow 10\hat{\beta}_2 - 0.6(6) = 16$$

$$\Rightarrow \hat{\beta}_2 = \frac{16+9}{10} = 2.5$$

$$5(\hat{\beta}_1) + 15(2.5) + 25(-1.5) = 20$$

$$\hat{\beta}_1 = 4$$

Generalised k-variable model - MLRM

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \cdots + \beta_k x_{ik} + u_i$$

$$y_i = \beta_0 + \sum_{j=1}^{k} \beta_j x_{ij} + u_i \quad , \forall i, i = 1(1)n \qquad \begin{array}{l} i = 1(1)n \\ j = 1(1)k \end{array}$$

$\forall i, i = 1(1)n$

$i = 1 - \quad y_1 = \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \beta_3 x_{13} + \cdots + \beta_k x_{1k} + u_1$

$i = 2 : \quad y_2 = \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + \beta_3 x_{23} + \cdots + \beta_k x_{2k} + u_2$

$\vdots$

$i = n : \quad y_n = \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \cdots + \beta_k x_{nk} + u_n$

In matrix form:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}_{n\times 1} = \begin{pmatrix} 1 & x_{11} & x_{12} & x_{13} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & x_{23} & \cdots & x_{2k} \\ \vdots & \vdots & & \ddots & & \vdots \\ 1 & x_{n1} & x_{n2} & x_{n3} & \cdots & x_{nk} \end{pmatrix}_{n\times(k+1)} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix}_{(k+1)\times 1} + \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}_{n\times 1}$$

$$\underset{\text{vector}}{y_{i\,n\times 1}} = (1 \;\; x_{i1} \;\; x_{i2} \;\cdots\; x_{ik}) \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{pmatrix}_{(k+1)\times 1} + \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}_{n\times 1} \text{, vector}$$

Matrix notation: $Y_i = X_i \beta + U_i$

$$Y_i = n \times 1$$
$$X_i = (n \times (k+1))$$
$$\hat{\beta} : ((k+1) \times 1)$$
$$X_i' : ((k+1) \times n)$$

$$SSR = \sum_{i=1}^{n} (Y_i - X_i \hat{\beta})^2$$

$$\frac{\partial(SSR)}{\partial \hat{\beta}} = 0$$

$$= \sum_{i=1}^{n} x_i' (y_i - x_i \hat{\beta}) = 0$$

$$x_i' = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{21} & \cdots & x_{n1} \\ \vdots & x_{22} & \ddots & \vdots \\ x_{1k} & x_{2k} & \cdots & x_{nk} \end{pmatrix}$$

$$\frac{\partial(SSR)}{\partial \hat{\beta}} = \sum_{i=1}^{n} x_i' (y_i - x_i \hat{\beta}) = 0$$

$$= \sum_{i=1}^{n} \left(x_i' y_i - x_i' x_i \hat{\beta} \right) = 0$$

$$\cancel{= x' y - x'}$$

$$= \cancel{x'x}\; x'y - x'x\hat{\beta} = 0$$

$$x'y = \underbrace{x'x}_{n \times n}\hat{\beta}$$

$n \times n$ – non singular matrix

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

$$SSR = \sum_{i=1}^{n} u_i^2 = \hat{u}_1^2 + \hat{u}_2^2 + \cdots + \hat{u}_n^2$$

we get, $SSR = \underbrace{(\hat{u}_1 \; \hat{u}_2 \; \cdots \; \hat{u}_n)}_{1 \times n} \underbrace{\begin{pmatrix} \hat{u}_1 \\ \hat{u}_2 \\ \vdots \\ \hat{u}_n \end{pmatrix}}_{n \times 1} = u'u = (y - x\hat{\beta})'(y - x\hat{\beta})$

$$= (y - x\hat{\beta})'(y - x\hat{\beta})$$

$$= y'(y - x\hat{\beta}) - (x\hat{\beta})'(y - x\hat{\beta}) = y'y - \hat{\beta}'x'y - y'x\hat{\beta} + \hat{\beta}'x'x\hat{\beta}$$

pre multiply for multiplication

$$\cancel{= y'y - y'x\hat{\beta} - (x\hat{\beta})'y}$$

$$SSR = y'y - 2\hat{\beta}'x'y + \hat{\beta}'x'x\hat{\beta}$$

$$SSR = Y'Y - 2\hat{\beta}'X'Y + \hat{\beta}'X'X\hat{\beta}$$

$\hat{\beta}'\hat{\beta} = \hat{\beta}^2$

$$\frac{\partial SSR}{\partial \hat{\beta}} = 0 \Rightarrow -2X'Y + 2\hat{\beta}X'X = 0$$

$[x \ y \ z]\begin{bmatrix} x \\ y \\ z \end{bmatrix} = x^2 + y^2 + z^2$

$$\Rightarrow \cancel{2}\hat{\beta}X'X = \cancel{2}X'Y$$

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

Wooldridge

6 Feb 2026

Recall, $y_i = (1 \quad x_{i1} \quad x_{i2} \quad \cdots \quad x_{ik}) \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} + u_i$

$n \times (k+1)$ $\quad$ $n \times 1$

$n \times 1$

$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$

$$\frac{\partial SSR}{\partial \hat{\beta}} \equiv 0$$

$$\hat{\beta} = \underset{k \times n}{(X'X)^{-1}} \underbrace{\overset{(k+1)\times n}{X'} \overset{n \times 1}{Y}}_{(k+1)\times 1}$$

$\hat{\beta}$ - Unbiasedness Property

How do we show unbiasedness with matrix notation?

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

1. Unbiased - $E(\hat{\beta}) = \beta$ [$\because E(u|x) = 0$].

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

$$= (X'X)^{-1} X' (X\beta + u)$$

$$= (X'X)^{-1} X'X\beta + (X'X)^{-1} X'u$$

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

$$E(\hat{\beta}) = \beta + 0$$

2. Homoskedasticity - $V(u|x) = \sigma^2 I_n$ ⇒ all diagonal elements are 0, other elements - co variance

3. No serial auto correlation - $Cov(u_t, u_s | x) = 0 \; \forall \; t \neq s$.

→ errors uncorrelated;

~~4 Variance covariance matrix:~~

(a) Conditional variance of $\hat{\beta}$:

$$V(\hat{\beta}|x) = V(\underset{\sim}{(X'X)^{-1}} X'Y|X)$$

exogenous/ non stochastic → canceled-out.

$$= V((X'X)^{-1} X'(X\beta + u)|X)$$

$$= V((X'X)^{-1}X'X\beta + (X'X^{-1})X'u|X)$$

$$= V(\beta) + V[(X'X)^{-1}X'u|X]$$

$$\cancel{= V(\beta) + (X'X)^{-1}X'[Var(u|X)X(XX)^{-1}]}$$

$$= V(\beta) + (X'X)^{-1}X'[V(u|x)].X(X'X)^{-1}$$

$$= 0 + (X'X)^{-1}X'X \; (X'X)^{-1}\sigma^2$$

$$= \sigma^2 (X'X)^{-1} \quad \left(\rightarrow \frac{\sigma^2}{S_{xx}}\right)$$

This is a more general case for a k-variable

Try to show if bivariate results hold for multivariate case.

Decomposition: SST = SSE + SSR.

$$\hat{u} = Y - \hat{Y} \quad \text{(substitute results)}$$

$$\Rightarrow Y = \hat{u} + \hat{Y} \qquad SSR = \sum_{i=1}^{n} u_i^2$$

$$\cancel{= \hat{u} + \hat{\beta}_0 x'} \qquad = u_1^2 + u_2^2 + \dots + u_n^2$$

$$Y = \hat{u} + X\hat{\beta} \qquad = (\hat{u}_1, \hat{u}_2 \dots \hat{u}_n)\begin{pmatrix} \hat{u}_1 \\ \hat{u}_2 \\ \vdots \\ \hat{u}_n \end{pmatrix}$$

$$Y' = (X\hat{\beta} + \hat{u})'$$

$$Y'Y \overset{-n\bar{Y}}{=} (X\hat{\beta} + \hat{u})'(\hat{u} + X\hat{\beta}) \quad +n\bar{Y} \text{ adjusted}$$

$$= (X'\hat{\beta}'_{y^T} + \hat{u}'_{n\times 1})(\hat{u}_{n\times 1} + X\hat{\beta}_{n\times 1}) = (Y^T + \hat{u}^T)(\hat{u} + Y)$$

$$Y'Y = \hat{\beta}'X'X\hat{\beta} + \cancel{Y'\hat{u}} + \cancel{\hat{u}'Y} + \hat{u}'\hat{u}$$

$\Sigma x u = 0$

True model
$Y = X\beta + u$

$$= \hat{\beta}'X'X\hat{\beta} + \hat{u}'\hat{u}$$

ESS + SSR

Total variation in Y.

$y_1 = \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + u_1$

$y_2 = \beta_0 + \beta_1 x_{21} + \beta_2 x_{22} + u_2$

We could have adjusted in $Y'Y$

$$\bar{y} = \hat{\beta_0} + \hat{\beta_1}\bar{x}_{i1} + \hat{\beta_2}\bar{x}_{i2}$$

adj in $Y'Y$

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{12} \\ 1 & x_{21} & x_{22} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

$$\underset{1\times 3}{(\hat{\beta}_0\ \hat{\beta}_1\ \hat{\beta}_2)} \underset{3\times 2}{\begin{pmatrix} 1 & 1 \\ x_{11} & x_{21} \\ x_{12} & x_{22} \end{pmatrix}} \underset{2\times 3}{\begin{pmatrix} 1 & x_{11} & x_{12} \\ 1 & x_{21} & x_{22} \end{pmatrix}} \underset{3\times 1}{\begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix}}$$

Partialling Out Effect → Frisch-Waugh Theorem

We take $K=2$ for simplicity:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + u_i \quad : \text{MLRM}$$

$$w_i = \beta_0 + \beta_1\, \text{education}_i + \beta_2\, \text{experience}_i + u_i$$

main obj — establish this relationship

↳ main interest variable

this may confound my result.

- Model assumes $x_{i1} \perp x_{i2}$
- Reality - expect - some correlation b/w x_{i1} & x_{i2}.
  ⇒ Some correlation is acceptable.
- Correl (edu, exp) $\neq 0$.

Objective: establish cause effect rel^n b/w education & wage.

→ So, take education in such a way that the correl. b/w ~~tot~~ edu & exp is ~~..~~ ~~...~~ excluded → this means -

we hope to "net-out" (or) "partial-out" or "tease out" the correlation b/w main interest variable & confounding factor.

Objective: To capture the effect of x_{i1}, netting out / partialling out the effect of x_{i2} or incorporate that part of x_{i1} which is not correlated with x_{i2}.

Two stages involved?-

Stage 1: Run OLS - regress x_{i1} on x_{i2}

→ all the effects captured by coefficients → which can be adjusted for.

$$x_{i1} = \alpha + \beta x_{i2} + r_{i1} \quad \text{where } r_{i1} \to \text{error}.$$

Estimated: $\hat{x}_{i1} = \hat{\alpha} + \hat{\beta} x_{i2}$

Therefore, $\hat{r}_{i1} = x_{i1} - \hat{x}_{i1}$

$$\hat{r}_{i1} = x_{i1} - \hat{\alpha} - \hat{\beta} x_{i2}$$

Stage 2: Recall original regression—

$$y_i = \beta_0 + \hat{\beta}_1 \hat{r}_{i1} + u_i$$

↳ gives the cause effect.

$\hat{r}_{i1}$ - captures that part of x_{i1} which is uncorrelated with x_{i2}.

Observation: 1. True / Fundamental model ⇒ MLRM.

2. Stage two model - is of SLRM type

↳ captures sole effect of x_{i1}. or net effect.

$$\hat{\beta}_1 = \frac{\mathrm{Cov}(X,Y)}{V(X)} \quad \text{now?}$$

$$\bar{y}_i = \beta_0 + \hat{\beta}_1 \bar{\hat{r}}_{i1} + \bar{u}_i$$

STATA

- Test FW Theo

reg wage (educ exper tenure)

reg educ

13·Feb·2026

~~COMPARING~~ ~~SLRM and MLRM~~

FW Theorem → convert an MLRM to SLRM.

$$\hat{\beta}_1 = \frac{cov(\hat{r}_{i1}, y_i)}{V(\hat{r}_i)} = \frac{\sum(\hat{r}_{i1} - \bar{\hat{r}}_{i1})(y_i - \bar{y})}{\sum(\hat{r}_i)^2}$$

(sum of residual $= 0$)

$$= \frac{\sum(\hat{r}_i - \bar{r}_{i1})(y_i - \bar{y})}{\sum(\hat{r}_i - \bar{\hat{r}}_i)^2} = \frac{\sum(\hat{r}_{i1} - 0)(y_i - \bar{y})}{\sum(\hat{r}_i)^2}$$

$$= \frac{\sum(\hat{r}_{i1} y_i - \hat{r}_{i1}\bar{y} - \bar{r}_{i1} y_i + \bar{r}_{i1}\bar{y})}{\sum(\hat{r}_i - \bar{\hat{r}}_i)^2} = \frac{cov(\hat{r}_i, y_i)}{V(\hat{r}_1)}$$

$$= \frac{\sum \hat{r}_{i1} y_i - \cancel{\bar{y} \sum r_{i1}} - \bar{\hat{r}}_{i1} \sum y_i + \cancel{n\bar{r}_{i1}\bar{y}}}{(\hat{r}_i - \bar{\hat{r}}_i)^2}$$

$= \bar{u}_i = 0 \quad \bar{r}_{i1} \sum y_i + n\bar{r}_{i1}$

$\bar{y} = \beta_0 - \hat{\beta}_1 \bar{\hat{r}} = 0$

$$= \frac{\sum \hat{r}_i y_i - 0}{\sum(\hat{r}_i^2)} = \frac{cov(\hat{r}_{i1}, y_i)}{V(\hat{r}_i)}$$

COMPARING SLRM and MLRM

Model: $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + u_i$ —— (1)

Estimate: $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2}$ —— (2)

Estimated SLRM: $\hat{\tilde{y}}_i = \hat{\tilde{\beta}}_0 + \hat{\tilde{\beta}}_1 x_{i1}$ —— (3)

Population model: SLRM: $\tilde{y}_i = \tilde{\beta}_0 + \tilde{\beta}_1 x_{i1} + \varepsilon_i$ —— (4)

From (1) - Run OLS x_{i2} on x_{i1}

$x_{i2} = \alpha + \delta x_{i1} + r_{i1}$ —— (5)

comparing (5) and (1) - substitute (5) in (1)

$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2(\alpha + \delta x_{i1} + r_{i1}) + u_i$ —— (6)

$= \beta_0 + \beta_1 x_{i1} + \beta_2 \alpha + \beta_2 \delta x_{i1} + \beta_2 r_{i1} + u_i$

$= \beta_0 + \beta_2 \alpha + \beta_1 x_{i1} + \beta_2 \delta x_{i1} + \beta_2 r_{i1} + u_i$

$= \underbrace{(\beta_0 + \beta_2 \alpha)}_{\text{intercept}} + (\overbrace{\beta_1 + \beta_2 \delta}^{\text{slope}}) x_{i1} + (\beta_2 r_{i1} + u_i)$ follows Normal. —— (7)

Another SLRM Model. residual.

$\beta' + \beta_2' x_{i1} + v_i'$ constant relationship.

On running OLS for (7) - putting $\hat{r}_{i1}$

$\hat{y}_i = (\hat{\beta}_0 + \hat{\beta}_2 \hat{\alpha}) + (\hat{\beta}_1 + \hat{\beta}_2 \hat{\delta}) x_{i1}$ —— (8)

comparing eq^n (8) with (3) → typically SLRM.

$\hat{\tilde{\beta}}_1 = \hat{\beta}_1 + \hat{\beta}_2 \hat{\delta}$.

$\hat{\tilde{\beta}}_1 = \hat{\beta}_1$ when $\hat{\beta}_2 = 0$ or $\hat{\delta} = 0$.

If $\hat{\beta}_2 = 0$, then x_{i2} and y_i have no relationship

1. when $\hat{\beta}_2 = 0$, $\tilde{\beta}_1 = \hat{\beta}_1$, this implies x_{i2} has no relationship/ effect with y_i.

2. when $\hat{\delta} = 0$, $\tilde{\beta}_1 = \hat{\beta}_1$, this implies $\text{cov}(x_{i1}, x_{i2}) = 0$ or correlation of x_{i1} & $x_{i2} = 0$.

→ Does this mean you need to have intuition quality

CASE 1: $\hat{\beta}_2 = 0$ → may endup giving you an insight into the problem of — OVERSPECIFICATION → false illusion of R^2 being a good fit

$E(\hat{\beta}_2) - \beta_2 = 0$.

$\frac{SSE \uparrow}{SST}$? can be spurious. Eg: Child birth & rainfall.

From (1) → Population Regression fn.

$E(y_i | X = x_1, x_2) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2}$ $\because [E(u_i|X) = 0]$

If PRF if $\beta_2 = 0$, then $E(\hat{\beta}_2) = \beta_2 = 0$.

x_{i2} is called a redundant variable.

~~Other~~ Other possiblity — $\beta_2 \neq 0$ (population) $\because$ sampling bias, non representative samples etc
but $\hat{\beta}_2 = 0$ (estimated).

This is a trivial case, by chance the relationship is null. We cannot say x_{i2} is a redundant variable. cannot say the problem is overspecification.

OMITTED VARIABLE BIAS

Model: $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + u_i$

Example-1: $W_i = \alpha + \delta_1 \, educ_i + \delta_2 \, ability + \varepsilon_i$

unobserved →
- add as covariate
- derive your covariate

Example 2: $C_i = \phi + \gamma_1 y_i + \gamma_2$ psychological factor$_i$ $+ \theta_i$

(unobserved.

Owner's age, between, or high/day wage

(3) $(\text{Crime rate})_i = \psi + \lambda_1 \, \text{unemployment}_i + \lambda_2 \begin{pmatrix}\text{police}\\ \text{station}\end{pmatrix}_i + \lambda_3 \, \text{Temperature} + \varepsilon$

confounding.

Date: 17 Feb 2026

Recap: Overspecification → Redundant variables.

$$\beta_3 = 0$$

$$E(\hat{\beta}_3) = 0$$

CASE 2: Non-Redundant, but omitted variables:

Did he wake up on the angel/wrong side

Model: $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + u_i \rightarrow$ MLRM, $k=2$ — (i)

Model: (CLRM Type): $y_i = \beta_0 + \beta_1 x_{i1} + \varepsilon_i$ — (ii)

Estimated: $\tilde{y}_i = \tilde{\beta}_0 + \tilde{\beta}_1 x_{i1}$ — (iii)

Est

Estimated: $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2}$ — (iv)

I-stage regression of x_{i2} and x_{i1}.

$$x_{i2} = \alpha + \delta x_{i1} + \gamma_{i1} \text{ — (v)}$$

Substituting (v) in (i)

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2(\alpha + \delta x_{i1} + \gamma_{i1}) + u_i$$

$$= \underbrace{(\beta_0 + \beta_2 \alpha)} + \underbrace{(\beta_1 + \beta_2 \delta)} x_{i1} + \underbrace{(u_i + \beta_2 \gamma_{i1})}_{\varepsilon_i} \text{ — (vi)}$$

error has additive property.

~~Recap~~:- From (vi) and ~~&~~ (iii)

$$\hat{y}_i = (\hat{\beta}_0 + \hat{\beta}_2 \alpha) + (\hat{\beta}_1 + \hat{\beta}_2 \tilde{\delta}) x_{i1} \text{ --- (vii)}$$

compare (iii) - and (vii)

$$\tilde{\beta}_1 = \hat{\beta}_1 + \hat{\beta}_2 \delta$$

$$E(\tilde{\beta}_1) = E(\hat{\beta}_1) + E(\hat{\beta}_2 \delta)$$

$$\boxed{E(\tilde{\beta}_1) = \beta_1 + \beta_2 \delta}$$

$$\therefore \text{Bias} = E(\tilde{\beta}_1) - \beta.$$

$$= \beta_2 \delta$$

unbiased:

$E(\tilde{\beta}_1) = \beta_1$. iff $\beta_2 \delta = 0 \Rightarrow \beta_2 = 0$ or $\tilde{\delta} = 0$.

If $\beta_2 = 0$, $x_{i2} \Rightarrow$ redundant.

If biased $\Rightarrow E(\tilde{\beta}_1) \neq \beta_1$.

Bias = positive or negative?

$\rightarrow$ ~~dep~~ If biased $\Rightarrow$ Bias $= E(\tilde{\beta}_1) = \beta_1 + \beta_2 \delta$

If $E(\tilde{\beta}_1) \neq \beta$, then $\beta_2 > 0, \tilde{\delta} > 0$ — (+)

$\beta_2 < 0, \tilde{\delta} > 0$ — (-)

$\beta_2 > 0, \tilde{\delta} < 0$ — (-)

$\beta_2 < 0, \tilde{\delta} < 0 \rightarrow$ (+)

From (v) $\rightarrow$ what is $\tilde{\delta}$?

~~Substituting $\bar{x}_1$ viii to~~

Lets take the avg.

$$\bar{x}_2 = \alpha + \delta \bar{x}_1 + \cancel{\ldots}^{0} + \bar{r}_{i1}$$

$$\cancel{\bar{x}_2 = \alpha + \delta \bar{x}_1 \text{ --- (viii)}}$$

Subtracting VIII from V.

$$(x_{i2}-\bar{x}_2) = \delta(x_{i1}-\bar{x}_1) + (r_{i1}-\bar{r}_1)$$

Multiplying both sides with $(x_{i1}-\bar{x})$

$$(x_{i2}-\bar{x}_2)(x_{i1}-\bar{x}) = \delta(x_{i1}-\bar{x}_1)(x_{i1}-\bar{x}_1) + (r_{i1}-\bar{r}_1)(x_{i1}-\bar{x}_1)$$

Summing both sides, dividing by n.

$$\sum_{i=1}^{n}(x_{i2}-\bar{x}_2)(x_{i1}-\bar{x}) = \delta\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2 + \sum_{i=1}^{n}(r_{i1}-\bar{r}_1)(x_{i1}-\bar{x}_1)$$

$$\frac{\sum_{i=1}^{n}(x_{i2}-\bar{x}_2)(x_{i2}-\bar{x})}{\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2} = \delta + \sum_{i=1}^{n}\frac{(r_{i1}-\bar{r}_1)(x_{i1}-\bar{x}_1)}{\sum_{i=1}^{n}(x_{i1}-\bar{x}_1)^2}$$

$$\frac{CoV(x_1,x_2)}{V(x_1)} = \delta + \frac{CoV(\overset{\text{error}}{r},x)}{V(x_1)} \quad = 0.$$

By OLS assumption

can be +/-

$$\frac{CoV(x_1,x_2)}{V(x_1)} = \delta \quad \text{depends on correlation of } x_1 \,\&\, x_2.$$

> 0

	$\rho_{x_1x_2} > 0$	$\rho_{x_1x_2} < 0$
$\beta_2 > 0$	(+)	(−)
$\beta_2 < 0$	(−)	(+)

Standardisation

Learning outcome = $\beta_0 + \beta_1$ years of edu $+ \beta_2 x_2 + \cdots + \beta_k x_k + u_i$

1 SD $\Delta \rightarrow$ what happens to learning outcome in years of edu

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + u_i$$

compute $\bar{y} = \beta_0 + \beta_1 \bar{x}_1 + \beta_2 \bar{x}_2 + \bar{u}$

$$(y_i - \bar{y}) = \cancel{\beta_0 - \beta_0} \; \beta_1(x_{i1} - \bar{x}_1) + \beta_2(x_{i2} - \bar{x}_2) + u_i - \bar{u}$$

$$(y_i - \bar{y}) = \beta_1(x_{i1} - \bar{x}_1) + \beta_2(x_{i2} - \bar{x}_2) + u_i - \bar{u}$$

Let σ_y: standard deviation of y.

σ_{x_j}: s.d of x_i $\forall j = 1(1)2$.

On dividing with σ_y

$$\frac{(y_i - \bar{y})}{\sigma_y} = \left(\frac{\sigma_{x1}}{\sigma_y}\beta_1\right)\left(\frac{x_{i1} - \bar{x}_1}{\sigma_{x1}}\right) + \left(\frac{\sigma_{x2}}{\sigma_y}\beta_2\right)\left(\frac{x_{i2} - \bar{x}_2}{\sigma_{x2}}\right) + \left(\frac{u_i - \bar{u}}{\sigma_y}\right)$$

$$z_y = b_1 z_{x1} + b_2 z_{x2} + z_u$$

Estimated eq^n: $\hat{z}_y = \hat{b}_1 z_{x1} + \hat{b}_2 z_{x2}$, where $b_1 = \frac{\sigma_{x1}}{\sigma_y}\beta_1$

$b_2 = \frac{\sigma_{x2}}{\sigma_y}\beta_2$

1 S.D $\uparrow$ in x_1, $y \uparrow$ by b_1.

$\frac{\partial z_y}{\partial z_x} = b_1$ $\quad \uparrow 1 \, \partial z_x \Rightarrow \uparrow z_y$ by b_1.

Scaling : $y_i = \beta_0 + \beta_1 x_i + u_i$

If y_i is child birth weight

$(b\_wt) = \beta_0 + \beta_1 (\text{smoking habit}) + u_i$

b_wt ↓ unit of measurement in grams ↓ convert this to kg

smoking habit ↓ $\hookrightarrow$ no. of ciggarette / day (c-pd).

$$\frac{\hat{b\_wt}}{1000} = \frac{\hat{\beta}_0}{1000} + \frac{\hat{\beta}_1}{1000} \text{c-pd} + \frac{u_i}{1000} \;?$$

↳ not possible as they are all estimates from sample, not from population.

→ this coefficients would scale down.

Rescaling dependent variable post estimation

→ convert cigarette to packets → 1 pack = 12 ciggarette.

$$\frac{\hat{b\text{-}wt}}{12} = \frac{\hat{\beta}_0}{12} + \hat{\beta}_1 \frac{\text{c-pd}}{12}$$

or

$$\hat{b\text{-}wt} = \hat{\beta}_0 + 12\hat{\beta}_1 \frac{\text{c-pd}}{12}$$

↳ remains same.

Gauss Markove Theorem $\rightarrow$ BLUE (Best Linear Unbiased Estimator).

SLRM MLRM

Ch-3 Appendix

Unbiasedness: $E(\hat{\beta}_1) = \beta_1 \longrightarrow E(\hat{\beta}_j) = \hat{\beta}_j \}\ j = 1(1)k.$

Linear : $\hat{\beta}_1 = \dfrac{\sum(x_i - \bar{x})\, y_i}{(x_i - \bar{x})^2} = \sum_{i=1}^{n} w_i y_i$

for simplicity, let, $i = 1(1)2$.

$$\hat{\beta}_1 = \frac{\sum_i (x_1 - \bar{x})}{\sum_i (x_1 - \bar{x})^2} y_1 + \frac{\sum_i (x_2 - \bar{x})}{\sum_i (x_2 - \bar{x})^2} y_2$$

$$\hat{\beta}_1 = w_1 y_1 + w_2 y_2, \text{ where } w_i = \frac{x_i - \bar{x}}{\sum_i (x_i - \bar{x})^2}$$

$$\hat{\beta}_1 = \sum_{i=1}^{n} w_i y_i$$

Similarly for MLRM: $\hat{\beta}_j = \sum_{i=1}^{n} w_{ij}\, y_i \ ;\ j = 1(1)k$

Best: Minimum variance.

To Prove: $V(\hat{\beta}_{1OLS}) \leq V(\tilde{\beta}_1)_{OLS}$

Aside**

$V(c_1u_1 + c_2u_2)$

$= c_1^2\sigma^2 + c_2^2\sigma^2 + 2\,cov(u_1,u_2)$

$= \sigma^2(c_1^2 + c_2^2)$

Scalar. $\tilde{\beta}_1 = \sum_i c_i y_i$ (not $\hat{\beta}_1$)

$\boxed{\text{(Dealing w/ unbiased estimator}}$

$$\tilde{\beta}_1 = \sum_i c_i(\beta_0 + \beta_1 x_i + u_i)$$

$$\tilde{\beta}_1 = \beta_0 \sum_{i=1}^{n} c_i + \beta_1 \sum_{i=1}^{n} x_i c_i + \underbrace{\sum_{i=1}^{n} c_i u_i}_{=0 \text{ by assumption}} \quad \text{— (1)}$$

Now, $\underbrace{E(\tilde{\beta}_1) = \beta_1}_{\text{Has to be unbiased}}$ requires $\sum_{i=1}^{n} c_i = 0$ (2) and $\sum_{i=1}^{n} x_i c_i = 1$ — (3)

$E(u_i|x) = 0$

$V(u|x) = \sigma^2$

$cov(u_i, u_j) = 0$.

$$V(\tilde{\beta}_1) = V\left(\sum_i c_i y_i\right) = V\left(\sum c_i(\beta_0 + \beta_1 x_i + u_i)\right)$$

$$= V\left(\beta_0 \sum c_i\right) + V\left(\beta \sum c_i x_i\right) + V\left(\sum c_i u_i\right)$$

Aside (**)

$$= 0 + 0 + V\left(\sum c_i u_i\right) = \sigma^2 \sum c_i^2$$

Objective: minimum variance

given $\sigma^2 = $ const $\Rightarrow$ min $\sum_i c_i^2$ subject to $\sum_i c_i = 0$ & $\sum_i c_i x_i = 1$

min $\sum_i c_i^2 \rightarrow V(\hat{\beta}_{1\,OLS})$.

Using lagrange multiplier

$$\mathcal{L} = \sum c_i^2 + \lambda\left(\sum c_i\right) + \mu\left(\sum_i c_i x_i - 1\right)$$

$$\frac{\partial \mathcal{L}}{\partial c_i} = 2c_i + \lambda + \mu x_i = 0.$$

$$\Rightarrow 2c_i = -\lambda - \mu x_i$$

$$c_i = -\frac{\lambda}{2} - \frac{\mu}{2} x_i$$

Let $c_i = A + Bx_i$, where $A = -\frac{\lambda}{2}$, $B = -\frac{\mu}{2}$

Given $\sum c_i = 0$

$$\sum c_i = nA + B\sum x_i$$

$$\Rightarrow 0 = nA + B\sum x_i$$

$$\Rightarrow -B\sum x_i = nA$$

$$\Rightarrow \boxed{-B\bar{x} = A.}$$

$$\sum c_i x_i = \sum (A + Bx_i)x_i$$

$$1 = \sum Ax_i + B\sum x_i^2$$

$$1 = \sum(-B\bar{x})x_i + B\sum x_i^2$$

$$1 = B\sum x_i^2 - B\sum(\bar{x}x_i)$$

$$\Rightarrow 1 = B\sum x_i^2 - B\bar{x}\sum x_i$$

$$\Rightarrow 1 = B\sum x_i^2 - B\bar{x}(n\bar{x})$$

$\Rightarrow 1 = B\sum x_i^2 - nB\bar{x}^2$

$\Rightarrow \cancel{1 = B\sum_{i=1}^{n}(\sum x_i^2 - \bar{x})^2}$

$\Rightarrow 1 = B(\sum x_i^2 - n\bar{x}^2)$

$\therefore \quad B = \dfrac{1}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$

$c_i = A + Bx_i$

$c_i = -B\bar{x} + Bx_i$

$c_i = B(x_i - \bar{x})$

$c_i = \dfrac{(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$

$c_i^2 = \dfrac{(x_i - \bar{x})^2}{(\sum(x_i - \bar{x})^2)^2}$

$\sum c_i^2 = \dfrac{\cancel{\sum(x_i - \bar{x})^2}}{(\sum(x_i - \bar{x})^2)^{\cancel{2}}}$

$\sum c_i^2 = \dfrac{1}{\sum(x_i - \bar{x})^2}$

On Multiplying with σ^2 on both sides

$$\sigma^2 \sum c_i^2 = \frac{\sigma^2}{\sum_i (x_i - \bar{x})^2} = \frac{\sigma^2}{SST_x} = V(\hat{\beta}_{1\,OLS})$$

$$\therefore \sigma^2 \sum c_i^2 = \frac{\sigma^2}{SST_x} = V(\hat{\beta}_{1\,OLS})$$

Scalar notation using → Cauchy Schwartz inequality

$$\left(\sum u_i v_i\right)^2 \leq \left(\sum_i u_i^2\right)\left(\sum_i v_i^2\right)$$

General proof: Using Matrix notation:

$\underset{k \times n}{\tilde{\beta}_1} = \underset{n \times k+1}{CY}$ where $Y = \underset{n \times k+1}{X}\underset{k+1 \times 1}{\beta} + u$ — $n \times 1$.

$\tilde{\beta}_1 = C(X\beta + u)$

$\tilde{\beta}_1 = CX\beta + \underbrace{Cu}_{0}$

$E(\tilde{\beta}_1) = \beta$, this requires $CX = I$.

$C = (X'X)^{-1}X' + D$

Aside:

To satisfy - $\cancel{CX=0}$

Let, $C_0 = (X'X)^{-1}X'$

$C_0 X = (X'X)^{-1}X'X = I$

$C_0X - C\cancel{X} = I - I = 0$

Define $D = C - (X'X)^{-1}X'$

and $DX = CX - (X'X)^{-1}X'X = 0$

$V(\tilde{\beta}) = V(CY)$

$= V(C(X\beta + u))$

$= V[((X'x)^{-1}X' + D)(X\beta + u)]$

$= V[(X'X)^{-1}X'X\beta + (X'X)^{-1}X'u + DX\beta + Du]$

$= V[\beta + (X'X)^{-1}X'u + 0 + Du]$

$= V[\beta + (X'X)^{-1}X'u + Du]$

$= V[(X'X)^{-1}X'u + Du] = \sigma^2[V((X'X)^{-1}X' + D)]$

$\cancel{= \sigma^2[(X'X)^{-1}X' + D]}$

$$= \sigma^2 V\left((X'X)^{-1}X' + D\right)$$ → nothing but square

$$= \sigma^2 \left[\left((XX')^{-1}X' + D\right)\left(X(X'X)^{-1} + D'\right) \right]$$

$$= \sigma^2 \left[(X'X)^{-1}X'X(X'X)^{-1} + DD' \right]$$ [∵ all DX terms are 0]

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

$$= V(\hat{\beta}) + \sigma^2 DD'$$ → positive semi definite matrix

$$V(\tilde{\beta}) = V(\hat{\beta}) + \sigma^2 DD'$$

$$V(\tilde{\beta}) \geq V(\hat{\beta}_{OLS})$$