61

$$Y = X\beta + \epsilon$$

Nonzero Mean of Disturbances

$$E(\epsilon) = \psi \neq 0 \qquad \text{where} \quad \psi = \begin{bmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_T \end{bmatrix}$$

– other ass. hol

Here $E(\epsilon_t) = \psi_t \neq 0$,
and is not necessarily constant $\forall\, t$.

Thus the noise, ϵ, represents a part of Y not explained adequately by our model.

Thm: $\hat{\beta} \sim N\left[\beta + (X'X)^{-1}X'\psi,\ \sigma^2 (X'X)^{-1}\right]$

Proof: let $\epsilon = \psi + v$ where $\underbrace{v \sim N(0, \sigma^2 I_T)}_{v \text{ is "nice"}}$

then $v = \epsilon - \psi$.

$$\begin{aligned}
\hat{\beta} &= (X'X)^{-1}X'Y \\
&= (X'X)^{-1}X'(X\beta + \epsilon) \\
&= (X'X)^{-1}X'(X\beta + \psi + v) \\
&= \beta + (X'X)^{-1}X'\psi + (X'X)^{-1}X'v
\end{aligned}$$

3 requirements for proof: — ϵ_t is normal

(i) $\hat{\beta}$ is Normal — l.c. of v's

(ii) $E(\hat{\beta}) = \beta + (X'X)^{-1}X'\psi$

(iii)
$$\begin{aligned}
Cov(\hat{\beta}) &= E[\hat{\beta} - E(\hat{\beta})][\hat{\beta} - E(\hat{\beta})]' \\
&= E[(X'X)^{-1}X'v][v'X(X'X)^{-1}] \\
&= (X'X)^{-1}X'E(vv')X(X'X)^{-1} \\
&= (X'X)^{-1}X'\sigma^2 I X(X'X)^{-1} \\
&= \sigma^2 (X'X)^{-1}
\end{aligned}$$

QED

<u>Note:</u> $\hat{\beta}$ is biased (unless $X'\psi = 0$).

Furthermore, The bias, $(X'X)^{-1}X'\psi$, is just the "regression" of ψ on X.

<u>Thm:</u> $$\text{plim}\ \hat{\beta} = \beta + \text{plim}\left(\frac{X'X}{T}\right)^{-1} \text{plim}\left(\frac{X'\psi}{T}\right)$$

Proof:
$$\begin{aligned}
\text{plim}\ \hat{\beta} &= \text{plim}\left[\beta + (X'X)^{-1}X'\psi + (X'X)^{-1}X'v\right] \\
&= \beta + \text{plim}\left(\frac{X'X}{T}\right)^{-1} \text{plim}\left(\frac{X'\psi}{T}\right) + \underbrace{\text{plim}\left(\frac{X'X}{T}\right)^{-1}}_{\text{finite}} \underbrace{\text{plim}\left(\frac{X'v}{T}\right)}_{0} \\
&= \beta + \text{plim}\left(\frac{X'X}{T}\right)^{-1} \text{plim}\left(\frac{X'\psi}{T}\right)
\end{aligned}$$

QED

<u>Note:</u> $\hat{\beta}$ is inconsistent (unless $\text{plim}\left(\frac{X'\psi}{T}\right) = 0$).

[X "uncorrelated with" ψ]

$$E\left(\frac{X'v}{T}\right) = 0$$

$$\text{Var}\left(\frac{X'v}{T}\right) \to 0$$

Thm: $E(s^2) = \sigma^2 + \frac{1}{T-K}\psi' M \psi$;
$\text{plim}\, s^2 = \sigma^2 + \text{plim}\, \frac{1}{T}\psi' M \psi$.

Proof: $s^2 = \frac{SSE}{T-K} = \frac{1}{T-K} e'e = \frac{1}{T-K}\epsilon' M \epsilon$

$$= \frac{1}{T-K}(\psi+\upsilon)' M (\psi+\upsilon)$$
$$= \frac{1}{T-K}\left[\psi' M \psi + \upsilon' M \upsilon + 2\psi' M \upsilon\right]$$

Thus $E(s^2) = \frac{1}{T-K}E(\upsilon' M \upsilon) + \frac{1}{T-K}E(\psi' M \psi) + \underbrace{\frac{2}{T-K}E(\psi' M \upsilon)}_{0 \;(\psi \text{ \& } \upsilon \text{ uncorrel.})}$

$$= \underbrace{\frac{1}{T-K}(T-K)\sigma^2}_{\text{since } \upsilon \text{ is nice (proved before)}} + \frac{1}{T-K}\psi' M \psi$$

$$= \sigma^2 + \frac{1}{T-K}\psi' M \psi$$

$$\text{plim}\, s^2 = \text{plim}\, \frac{1}{T-K}\left[\upsilon' M \upsilon + \psi' M \psi + 2\psi' M \upsilon\right]$$

$$= \text{plim}\, \frac{1}{T}\upsilon' M \upsilon + \text{plim}\, \frac{1}{T}\psi' M \psi + \underbrace{\text{plim}\, \frac{2}{T}\psi' M \upsilon}_{0^*}$$

$$= \underbrace{\sigma^2}_{\text{since } \upsilon \text{ is nice (proved before)}} + \text{plim}\, \frac{1}{T}\psi' M \psi$$

$$\parallel \; \text{plim}\, \frac{1}{T-K}\psi' M \psi$$

QED

* $\text{plim}\, \frac{\psi' M \upsilon}{T} = 0$, since $E\left(\frac{\psi' M \upsilon}{T}\right) = 0$ from above, and
$\text{Cov}\left(\frac{\psi' M \upsilon}{T}\right) = E\left(\frac{\psi' M \upsilon}{T}\right)\left(\frac{\upsilon' M \psi}{T}\right)$
$= \frac{1}{T}\psi' M E(\upsilon\upsilon') M \psi = \frac{\sigma^2}{T}\psi' M \psi \to 0$

[Because $\left(\frac{\psi' M \psi}{T}\right) \to$ finite #]

Note: s^2 is biased & inconsistent.

$$E(s^2) \geq \sigma^2 \quad \text{since} \quad \frac{1}{T-K}\,\psi' M \psi \quad \text{is psd.}$$
$$\text{plim}\, s^2 \geq \sigma^2 \quad \text{since} \quad \text{plim}\, \frac{\psi' M \psi}{T} \quad \text{is psd.}$$

Thus s^2 is biased upward.

Summary: ① $\hat{\beta} = \beta + (X'X)^{-1}X'\psi + (X'X)^{-1}X'v$
is biased unless $X'\psi = 0$.

② $s^2 = \frac{SSE}{T-K}$
is biased unless $\psi' M \psi = 0$.

* It is impossible to have $X'\psi = 0$ and $\psi' M \psi = 0$

∴ At least one (& usually both) will be biased.

*(Proof)
Assuming $\psi \neq 0$;

(i) If $X'\psi = 0$, then $\psi' M \psi = \psi'\psi - \psi' X (X'X)^{-1} X'\psi$
$= \psi'\psi$

i.e. If $\hat{\beta}$ is unbiased, s^2 is biased.

(ii) If $\psi' M \psi = 0$, then $\psi'\psi = \psi' X (X'X)^{-1} X'\psi$
and $X'\psi \neq 0$

i.e. If s^2 is unbiased, $\hat{\beta}$ is biased.

Again, in general, both will be biased.

This (nonzero mean of ϵ) is serious problem. We have generally biased estimates of β, and no confidence in the t-statistics (since s^2 is biased).

Consider a special case of this problem.

Suppose $\psi = \begin{bmatrix} \psi \\ \psi \\ \vdots \\ \psi \end{bmatrix}$; or $E(\epsilon_t) = \psi \ \forall t$.

[i.e. constant mean of ϵ.]

Here, $$y_t = \alpha + \beta_1 X_{t1} + \cdots + \beta_K X_{tK} + \epsilon_t$$

$$= (\alpha + \psi) + \beta_1 X_{t1} + \cdots + \beta_K X_{tK} + (\epsilon_t - \psi)$$

(this is now well-behaved) v_t!

If $\exists$ a constant term in the regression, this special case is no problem! The error's constant effect simply gets washed into the constant term.

The estimate of α is now biased, but the other estimates are nice.

Intuitive explanation:

$$E(\hat{\beta}) = \beta + (X'X)^{-1}X'\psi$$
$$E(s^2) = \sigma^2 + \frac{1}{T-K}\psi' M \psi$$

It turns out that in general, some (but not all) of the ψ gets absorbed into the estimates of β, $\hat{\beta}$, and $\therefore$ biases them.

The rest of the ψ gets absorbed into the estimate of the variance, s^2, and $\therefore$ biases that.

This special case is one in which ψ happens to get totally absorbed into <u>one</u> regressor — the constant.

Hence it doesn't bias the other estimates of $\hat{\beta}$ or s^2 in this case.

Specification Error

Consider the model, $Y = X\beta + \epsilon$,
satisfying ideal conditions.

Partition X & β :

$$Y = (X_1 \quad X_2)\begin{bmatrix}\beta_1 \\ \beta_2\end{bmatrix} + \epsilon \qquad \begin{bmatrix} H \text{ columns in } X_1 \\ K-H \quad " \quad " \quad X_2 \end{bmatrix}$$

$$= X_1\beta_1 + X_2\beta_2 + \epsilon$$

Suppose we set up the model;

$$Y = X_1\beta_1 + \epsilon^* \qquad \text{where} \quad \epsilon^* = X_2\beta_2 + \epsilon$$
$$E(\epsilon^*) = X_2\beta_2 \quad (= \psi)$$

This is <u>Specification Error</u>
by the <u>Omission of Relevant Variables</u>.

How serious is this problem?
- depends on how significant the omitted variables are.
- We must always omit variables,
  but must include everything <u>important</u> if ϵ is to be nice.

Case (i)

Regressing Y on X_1:

$$\hat\beta_1 = (X_1'X_1)^{-1}X_1'Y$$
$$= (X_1'X_1)^{-1}X_1'(X_1\beta_1 + X_2\beta_2 + \epsilon)$$
$$= \beta_1 + (X_1'X_1)^{-1}X_1'X_2\beta_2 + (X_1'X_1)^{-1}X_1'\epsilon$$

Already shown this in gen. case

$$E(\hat\beta_1) = \beta_1 + \underbrace{(X_1'X_1)^{-1}X_1'X_2\beta_2}_{}$$

$\Psi = X_2\beta_2$

Bias $= f$(omitted variables)

$=$ "the regression of $X_2\beta_2$ on X_1"

Hence, the extent of bias depends on the extent to which $X_2\beta_2$ & X_1 are <u>correlated</u>.

$\Longrightarrow$ Bias $= 0$ if $\beta_2 = 0$

or if $X_1'X_2 = 0$.

or if $X_2\beta_2 = X_1$

Consider the estimate of σ^2;

$$s^2 = \frac{1}{T-H}SSE^* = \frac{1}{T-H}e^{*\prime}e^* = \frac{1}{T-H}\epsilon^{*\prime}M_1\epsilon^*$$

$$= \frac{1}{T-H}(X_2\beta_2+\epsilon)'M_1(X_2\beta_2+\epsilon)$$
$$= \frac{1}{T-H}[\beta_2'X_2'M_1X_2\beta_2 + 2\beta_2'X_2'M_1\epsilon + \epsilon'M_1\epsilon]$$
$$= \frac{1}{T-H}\epsilon'M_1\epsilon + \frac{1}{T-H}[\beta_2'X_2'M_1X_2\beta_2 + 2\beta_2'X_2'M_1\epsilon]$$

$$E(s^2) = \sigma^2 + \frac{1}{T-H}[\beta_2'X_2'M_1X_2\beta_2] \quad \leftarrow \text{Biased}$$

DIGRESSION on Partitioned Inverse Rule

Suppose that a square matrix, Z, can be partitioned as follows.

$$Z = \begin{bmatrix} E & F \\ G & H \end{bmatrix} \quad \text{Then} \quad Z^{-1} = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \quad \text{with} \quad \begin{bmatrix} E & F \\ G & H \end{bmatrix}\begin{bmatrix} a & c \\ b & d \end{bmatrix} = I.$$

Given $E, F, G,$ & H, find $a, b, c,$ and d.

The following 4 statements hold.

① $Ea + Fb = I$

② $Ec + Fd = 0 \implies$ from ②: $Ec = -Fd$

③ $Ga + Hb = 0$ $\qquad c = -E^{-1}Fd$ (if E nonsingular)

④ $Gc + Hd = I$

Substitute into ④: $G(-E^{-1}Fd) + Hd = I$

$$\underbrace{(H - GE^{-1}F)}_{D} d = I$$

$$Dd = I$$

$$\boxed{\begin{aligned} d &= D^{-1}I = D^{-1} \\ c &= -E^{-1}FD^{-1} \end{aligned}}$$

?

From ①: $Ea = I - Fb$

$$a = E^{-1}(I - Fb)$$

Substitute into ③: $GE^{-1}(I - Fb) + Hb = 0$

$$GE^{-1} - GE^{-1}Fb + Hb = 0$$

$$\underbrace{(H - GE^{-1}F)}_{D} b = -GE^{-1}$$

$$Db = -GE^{-1}$$

$$\boxed{\begin{aligned} b &= -D^{-1}GE^{-1} \\ a &= E^{-1}(I + FD^{-1}GE^{-1}) \end{aligned}}$$

$$\therefore \quad Z^{-1} = \begin{bmatrix} E^{-1} + E^{-1}FD^{-1}GE^{-1} & -E^{-1}FD^{-1} \\ -D^{-1}GE^{-1} & D^{-1} \end{bmatrix}$$

[* put $\mathbb{D}^{-1}$ in top left corner.]

Alternatively, from (3): $\quad Hb = -Ga$

$$b = -H^{-1}Ga \qquad \text{(if } H \text{ nonsingular)}$$

Substitute into (1):

$$Ea + F(-H^{-1}Ga) = I$$
$$\underbrace{(E - FH^{-1}G)}_{\mathbb{D}} a = I$$
$$\mathbb{D} a = I$$
$$\boxed{\begin{aligned} a &= \mathbb{D}^{-1} \\ b &= -H^{-1}G\mathbb{D}^{-1} \end{aligned}}$$

From (4):

$$Hd = I - Gc$$
$$d = H^{-1}(I - Gc)$$

Substitute into (2):

$$Ec + FH^{-1}(I - Gc) = 0$$
$$Ec + FH^{-1} - FH^{-1}Gc = 0$$
$$\underbrace{(E - FH^{-1}G)}_{\mathbb{D}} c = -FH^{-1}$$
$$\mathbb{D}c = -FH^{-1}$$
$$c = -\mathbb{D}^{-1}FH^{-1}$$
$$d = H^{-1}(I + G\mathbb{D}^{-1}FH^{-1})$$

And Z^{-1} can also be written:

$$Z^{-1} = \begin{bmatrix} \mathbb{D}^{-1} & -\mathbb{D}^{-1}FH^{-1} \\ -H^{-1}G\mathbb{D}^{-1} & H^{-1} + H^{-1}G\mathbb{D}^{-1}FH^{-1} \end{bmatrix}$$

Thus $$Z^{-1} = \begin{bmatrix} E^{-1} + E^{-1}FD^{-1}GE^{-1} & -E^{-1}FD^{-1} \\ -D^{-1}GE^{-1} & D^{-1} \end{bmatrix} = \begin{bmatrix} \mathbb{D}^{-1} & -\mathbb{D}^{-1}FH^{-1} \\ -H^{-1}G\mathbb{D}^{-1} & H^{-1} + H^{-1}G\mathbb{D}^{-1}FH^{-1} \end{bmatrix}$$

where $D = H - GE^{-1}F$ and $\mathbb{D} = E - FH^{-1}G$

Why the Digression?

Consider the Covariance Matrix ;

$$Cov(\hat{\beta}_1) = \sigma^2 (X_1'X_1)^{-1}$$

(This is too small, as shown below.)

The true Covariance Matrix of the true OLS estimates are as follows.

(Case (ii): Regress Y on X to get $\hat{\beta}$.

Let $\hat{\beta}_1$ = top H estimates in $\hat{\beta}$. These are unbiased.

$$Cov(\hat{\beta}) = \sigma^2 (X'X)^{-1} = \sigma^2 \begin{bmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{bmatrix}^{-1}$$

$$= \sigma^2 \begin{bmatrix} (X_1'M_2X_1)^{-1} & \text{something} \\ \text{something} & \text{something} \end{bmatrix}$$

using Partitioned Inverse Rule with $\mathbb{D} = (X_1'M_2X_1)$ in top left.

Then $Cov(\hat{\beta}_1)$ = top left corner

$$= \sigma^2 (X_1'M_2X_1)^{-1}$$

We need to show that $(X_1'M_2X_1)^{-1} - (X_1'X_1)^{-1}$ is psd.

This is true iff $(X_1'X_1) - (X_1'M_2X_1)$ is psd.

— imposing Rule that $A-B$ is psd iff $B^{-1} - A^{-1}$ is psd.

$$\Longrightarrow (X_1'X_1) - X_1'[I_T - X_2(X_2'X_2)^{-1}X_2']X_1$$

$$= X_1'\underbrace{X_2(X_2'X_2)^{-1}X_2'}_{\text{psd}}X_1 \checkmark$$

$\rightarrow X_2(X_2'X_2)^{-1}X_2'$ is idempotent

INSERT A

Could show this in another way.

$$Cov(\hat{\beta}) = \sigma^2 \begin{bmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{bmatrix}^{-1}$$

$$= \sigma^2 \begin{bmatrix} (X_1'X_1)^{-1} + (X_1'X_1)^{-1}X_1'X_2(X_2'M_1X_2)^{-1}X_2'X_1(X_1'X_1)^{-1} & \text{something} \\ \text{something} & \text{something} \end{bmatrix}$$

using Partitioned Inverse Rule
with $D = (X_2'M_1X_2)$ in bottom right.

Then the Case (ii) $Cov(\hat{\beta}_1)$ can also be expressed as,

$$\sigma^2[(X_1'X_1)^{-1} + (X_1'X_1)^{-1}X_1'X_2(X_2'M_1X_2)^{-1}X_2'X_1(X_1'X_1)^{-1}]$$

Thus [Case (ii) $Cov(\hat{\beta}_1)$ − Case (i) $Cov(\hat{\beta}_1)$] also =

$$(X_1'X_1)^{-1}X_1'X_2(X_2'M_1X_2)^{-1}X_2'X_1(X_1'X_1)^{-1}$$

— also psd ✓

Implications :

Omitting relevant variables results in —

(i) Biased estimates, $\hat{\beta}_1$.

(ii) Biased estimate of σ^2, s^2 .

(iii) Variance of estimates in $\hat{\beta}_1$ reduced .

Note: These implications already discussed under the topic of restrictions.

$\Longrightarrow$ Omitting variables is one way to impose restrictions !

→ If true, no bias introduced, but variance reduced (gai... effi...

→ If untrue, Bias introduced, Variance reduced.

(iv) t-statistics and F-statistics are messed up. [s^2 is biased]
∴ Testing problems.

In general, this is a serious problem.

— Don't omit relevant variables !

Prelim Question 1

What is wrong with the following "solution" to Multicollinearity?

- Drop variable that is uncorrelated with the other variables, so that the Bias, $(X_1'X_1)^{-1}X_1'X_2\beta_2$, will not amount to much.

Answer — Dropping a variable (uncorrelated one) does nothing about Multicollinearity

Example of Misspecification of Functional Form

Demand for Money;

$$M_t = \alpha + \beta Y_t + \epsilon_t \qquad \text{(true form)}$$

Run log form;

$$\log M_t = \delta + \gamma \log Y_t + v_t$$

$$\Rightarrow v_t = \log M_t - \delta - \gamma \log Y_t$$

$$= \underbrace{\log(\alpha + \beta Y_t + \epsilon_t) - \delta - \gamma \log Y_t}_{\text{(this is a mess; doesn't satisfy ideal cond.)}}$$

Point: anything causing nonzero mean of disturbance results in these problems.

Some people use log-linear form because they want to get "elasticities" as coefficients.

This is bad reasoning.
<u>Should</u> use the functional form that is <u>appropriate</u> –
(i) by a priori theory;
(ii) by best fit of data.

So, what should determine Specification?

(i) A priori Theory;
(ii) Look at the data;
(iii) Try various statistical tests.