- When do we violate $E(u_i|x_i)=0$ assumption?

④ Measurement Error among regressor.

$$y_i = \alpha + \beta x_i^* + u_i \quad (x_i^* : \text{Not observed})$$

→ Until now, when we think of data, those are considered as "observed".

Thus, we can refer the data as "fixed ones".

If X can be controlled, X : fixed variable.

If X cannot be controlled, X : stochastic variable.

→ Now, let's think of a case: x_i^* cannot be observed. (Assume $E(u_i|x_i^*)=0$)

$\Downarrow$ implies

$E(x_i^* u_i)=0$

Instead, we can see $x_i = x_i^* + v_i$ under $E(u_i|x_i^*)=0$

$\Downarrow$

- x_i^*, v_i are independent. ($x_i^* \perp v_i$ and $E(v_i)=0$)

$\Rightarrow E(x_i^* v_i)=0$

- v_i, u_i are also independent. ($v_i \perp u_i$ and $E(u_i)=0$)

$\Rightarrow E(u_i v_i)=0$

Classical Measurement error model.

(That is, assume $\boxed{\begin{array}{l} ① \ E(x_i^* u_i)=0 \\ ② \ E(x_i^* v_i)=0 \\ ③ \ E(u_i v_i)=0 \end{array}}$

- used model: $y_i = \alpha + \beta x_i + \varepsilon_i$ (Not true model)

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

$$= \alpha + \beta(x_i - v_i) + u_i = \alpha + \beta x_i + (u_i - \beta v_i)$$

Check $E(x_i \varepsilon_i)=0$: $E(x_i\varepsilon_i) = E[(x_i^* + v_i)(u_i - \beta v_i)]$

$$= E[(x_i^* u_i + v_i u_i - \beta x_i^* v_i - \beta v_i^2)]=0$$

$\downarrow$ $E(x_i^* u_i)=0$ and $E(x_i^* v_i)=0$

$$= E(v_i u_i) - \beta E(v_i^2) \underset{\uparrow}{=} -\beta E(v_i^2) = \boxed{-\beta\sigma_v^2 \neq 0}$$

By assumption, v_i, u_i are independent.

Note Several assumption. (Not Lecture Note)

Strong assump. ↑

	$E(\hat{\beta})=\beta$ (unbiased)	$\hat{\beta} \xrightarrow{p} \beta$ (consistent)
x_i^* : constant (Always x_i^*, u_i independant)	O	O
x_i^* : stochastic & x_i^*, u_i are independent ($x_i \perp u_i$ and $E(u_i)=0$)	O	O
x_i^* : Stochastic & $E(u_i \mid x_i)=0$	O	O
x_i^* : Stochastic & $COV(x_i, u_i)=0$, $E(u_i)=0$	Not always	O

① x_i : stochastic & x_i, u_i are independent.

- $E(\hat{\beta})-\beta = E\left[(\Sigma x_i^2)^{-1}\Sigma x_i u_i\right] = E\left(\frac{\Sigma x_i}{\Sigma x_i^2} u_i\right) = E[\Sigma w_i u_i]$ where $w_i = \frac{x_i}{\Sigma x_i^2}$

$= \Sigma E(w_i u_i) = \Sigma E(w_i)E(u_i) = 0$. ($\because$ x_i, u_i independant $\Rightarrow E(u_i)=0$)

Therefore, $E(\hat{\beta})=\beta$ (unbiased)

- $\hat{\beta} = (\Sigma x_i^2)^{-1}\Sigma x_i y_i = \beta + \frac{\Sigma x_i u_i}{\Sigma x_i^2} = \beta + \frac{\frac{1}{N}\Sigma x_i u_i}{\frac{1}{N}\Sigma x_i^2}$ $\quad \frac{1}{N}\Sigma x_i u_i \xrightarrow{p} E(x_i u_i)$ by LLN, $\quad \frac{1}{N}\Sigma x_i^2 \xrightarrow{p} E(x_i^2)>0$ by LLN

By x_i^*, u_i independant, $E(x_i u_i) = E(x_i)E(u_i) = E(x_i)\cdot 0 = 0$

Therefore, $\hat{\beta} \xrightarrow{p} \beta$ (consistent).

② x_i : Stochastic & $E(u_i \mid x_i)=0$

- $E(\hat{\beta}) = \beta + E\left(\frac{\Sigma x_i u_i}{\Sigma x_i^2}\right) = \beta + E(\Sigma w_i u_i)$ (where $w_i = \frac{x_i}{\Sigma x_i^2}$)

$= \beta + E(E(\Sigma w_i u_i \mid x_i))$ by law of Iterative Expectation

$= \beta + E(\Sigma w_i \underbrace{E(u_i \mid x_i)}_{=0}) = \beta$

- $\hat{\beta} = \beta + \frac{\frac{1}{N}\Sigma x_i u_i}{\frac{1}{N}\Sigma x_i^2}$ $\quad \frac{1}{N}\Sigma x_i u_i \xrightarrow{p} E(x_i u_i)$, $\quad \frac{1}{N}\Sigma x_i^2 \xrightarrow{p} E(x_i^2)>0$ by LLN. and then

$E(x_i u_i) = E(E(x_i u_i \mid x_i)) = E(x_i \cdot E(u_i \mid x_i)) = 0$ Thus, $\hat{\beta} \xrightarrow{p} \beta$.

③ x_i : stochastic & $cov(x_i, u_i) = 0$, $E(u_i) = 0$

$$\hat{\beta} = \beta + \frac{\frac{1}{N}\Sigma x_i u_i \xrightarrow{P} E(x_i u_i)}{\frac{1}{N}\Sigma x_i^2 \xrightarrow{P} E(x_i^2)} \quad \text{by LLN.}$$

We need $E(x_i u_i)$ for consistency.

$\because cov(x_i, u_i) = 0$

$$cov(x_i, u_i) = E(x_i - E(x_i))(u_i - E(u_i)) = E[x_i u_i - E(x_i)u_i] = E(x_i u_i) \overset{\downarrow}{=} 0$$

Therefore, $E(x_i u_i) = 0$. Thus, $\hat{\beta} \xrightarrow{P} \beta$

※ We usually use assumption for "x_i is Stochastic" : $E(u_i | x_i) = 0$.

(For unbiasedness & consistency)

Under classical measurement error assumption for regressor, $E(x_i^* v_i) = 0$, $E(x_i^* u_i) = 0$, $E(u_i v_i) = 0$

$y_i = \beta \cdot x_i^* + u_i$, $E(u_i | x_i^*) = 0$

↑ we cannot see this

$$= \beta \cdot (x_i - v_i) + u_i = \beta \cdot x_i + \underbrace{(u_i - \beta v_i)}_{:= \varepsilon_i}$$

↑ we can see this.

(Assume $\sigma_x^2 = E(x_i^{*2})$, $\sigma_v^2 = E(v_i^2)$)

$$\hat{\beta} = \beta + \frac{\frac{1}{N}\Sigma x_i \varepsilon_i}{\frac{1}{N}\Sigma x_i^2}$$

① $\frac{1}{N}\Sigma x_i^2 = \frac{1}{N}\Sigma (x_i^* + v_i)^2 = \frac{1}{N}\Sigma x_i^{*2} + \frac{2}{N}\Sigma x_i^* v_i + \frac{1}{N}\Sigma v_i^2 \xrightarrow{P} \sigma_x^2 + \sigma_v^2$

($\frac{2}{N}\Sigma x_i^* v_i \xrightarrow{P} E(x_i^* v_i) = 0$ by LLN)

② $\frac{1}{N}\Sigma x_i \varepsilon_i = \frac{1}{N}\Sigma (x_i^* + v_i)(u_i - \beta v_i) = \frac{1}{N}\Sigma x_i^* u_i + \frac{1}{N}\Sigma u_i v_i + \frac{1}{N}\Sigma x_i^* (-\beta v_i) - \beta \cdot \frac{1}{N}\Sigma v_i^2$

$\xrightarrow{P} E(x_i^* u_i) = 0$, $E(u_i v_i) = 0$, $-\beta E(x_i^* v_i) = 0$, $-\beta E(v_i^2)$

$$\xrightarrow{P} -\beta \cdot \sigma_v^2$$

$$\hat{\beta} = \beta + \frac{\frac{1}{N}\Sigma x_i \varepsilon_i}{\frac{1}{N}\Sigma x_i^2} \xrightarrow{P} \beta + \frac{-\beta \sigma_v^2}{\sigma_x^2 + \sigma_v^2} = \left(1 - \frac{\sigma_v^2}{\sigma_x^2 + \sigma_v^2}\right)\beta < \beta$$

=> If ∃ measurement error, $\hat{\beta}$ converge to less number than β.

(Not consistent)

c-f) Measurement Error on the dependent variable

→ If the measurement error is on the dependent variable, the same argument for the previous case does not hold.

$$\underline{\underline{y_i^* = \alpha + \beta x_i + u_i}}$$ (y_i^* : Not observed) (Assume $E(u_i|x_i)=0$)

$y_i = y_i^* + v_i$: y_i observed.

$$\downarrow$$

$$y_i = y_i^* + v_i = \alpha + \beta x_i + \boxed{u_i + v_i}$$

Logically / Usually x_i, v_i are not correlated. maybe.

$$E(x_i \cdot (u_i + v_i)) = E(x_i u_i) + E(x_i v_i) = E(x_i v_i) = 0.$$

$$\hookrightarrow E(x_i u_i) = E(E(x_i u_i | x_i)) = E(x_i E(u_i | x_i)) = 0.$$

※ Therefore, we could assume that measurement error on y_i^* and x_i are Not Correlated without falling into logical inconsistency.

c-f 2) Dummy variable case.

→ If x_i^* is discrete (so that x_i is also discrete) then the classical measurement error does not hold.

ex) $x_i^* = \begin{cases} 0 \\ 1 \end{cases}$ $\quad x_i = \begin{cases} 0 \\ 1 \end{cases}$

$$x_i = x_i^* + v_i \Rightarrow \left.\begin{array}{ll} x_i^* = 1 & v_i = 0 \text{ or } -1 \\ x_i^* = 0 & v_i = 0 \text{ or } 1 \end{array}\right\} \text{ Not independent.}$$

Negatively correlated. : This shows dummy variable case break classical assumption for measurement error.

(May be, $E(x_i^* v_i) \leq 0$).

$\Downarrow$

This case is a little different from the previous dummy.

: Comment on proxy

$$y_i = \alpha + \beta x_i + u_i, \quad E(u_i|x_i)=0$$

When we use proxy, we can make Z_i as follows ;

$$Z_i \begin{cases} = 1 & \text{if } f(x_i) \geq 0 \\ = 0 & \text{if } f(x_i) < 0 \end{cases}$$

$$y_i = \underline{\theta_0 + \theta_1 Z_i} + \underline{\varepsilon_i}$$

$$= \underline{E(y_i|Z_i)} + \underline{y_i - E(y_i|Z_i)}$$

→ In this case, we don't want to get β so that there is No measurement error for the proxy.

- When do we violate $E(u_i|x_i)=0$ assumption?

⑤ Lagged Dependent Variable among regressors
& serial correlation in the residual term

$$\begin{cases} y_t = \alpha + \beta \cdot y_{t-1} + u_t \\ \text{and } u_t = \rho u_{t-1} + v_t \end{cases}$$

When we see y_{t-1},

$$y_{t-1} = \alpha + \beta y_{t-2} + u_{t-1}$$

Thus, y_t and y_{t-1} are generally correlated.

※ Solution for each case of the violation of $E(u_i|x_i)=0$ assumption.

① Measurement Error
② Simultaneity
③ Lagged Dependent Variable & Serial Correlation.

⇒ · Instrumental Variable
· Panel Data analysis.

④ Sample selection ⇒ MLE, Control function Approach, Semi-parametric Analysis.

⑤ Functual Misspecification ⇒ Non-parametric or semi-parametric Analysis.

⇒ As I showed, we should be careful to take a regression when $E(u_i|x_i) \neq 0$, which implies it is possible to violate consistency.

- What happens to the OLS Estimator when $E(u_i|x_i) \neq 0$?

$$\hat{\beta} = \beta + (X'X)^{-1} X'u = \beta + \left(\frac{1}{N} X'X\right)^{-1} \frac{1}{N} X'u \xrightarrow{p} \beta + E(x_i x_i') \cdot E(x_i) \cdot E(u_i|x_i)$$

$$\frac{1}{N} X'X \xrightarrow{p} E(x_i x_i') \qquad \frac{1}{N} X'u \xrightarrow{p} E(x_i u_i) = E(E(x_i u_i|x_i)) = E(x_i E(u_i|x_i)) = E(x_i) \cdot E(u_i|x_i)$$

→ Thus, when $E(u_i|x_i) \neq 0$, inconsistency is possible.

- Effect of function misspecification

① Including a regressor which is not needed among the regressors.

$$E(y|x) = \beta_1 + \beta_2 x_2 + \beta_3 x_3 \quad \& \quad \beta_3 = 0.$$

→ Even though $\beta_3 = 0$ is true, we cannot know this.

Thus, we usually put x_2, x_3 for check the real structure for data.

→ In this case, even if we eliminate $\beta_3 x_3$, the above model still holds so that conditional unbiasedness and consistency for OLS estimator hold.

ex) $u_i = y_i - E(y_i|x_i)$

$= y_i - [\beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i}]$

$y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + u_i$, $E(u_i|x_i) = 0$ ⇒ Still classic.

but the conditional variance is bigger.

ex2) $Var(\hat{\beta}_2|X) = \dfrac{\sigma^2}{\sum_{i=1}^{N} \hat{v}_{2i}^2}$, $\hat{v}_{2i}$: the residual of auxiliary regression.

② Omitting a regressor which should be included?

$E(y|x) = \beta_1 + \beta_2 x_2 + \beta_3 x_3, \quad \beta_3 \neq 0.$ (True model)

But regress y on $1, x_2$: $\underline{E(y|x) = \beta_1 + \beta_2 x_2}$ [Short regression]

$\Downarrow$

$y_i = \beta_1 + \beta_2 x_{2i} + \varepsilon_i$

$y_i = \beta_1 + \beta_2 x_{2i} + \boxed{\beta_3 x_{3i} + u_i}$

This reg. problem is, in this case, u_i and x_{3i} correlated.

If x_{2i}, x_{3i} are correlated,

then x_{2i}, ε_i are correlated, which implies $\hat{\beta}_2$ is inconsistent.

("omitted variable bias" causes Direction of inconsistency problem.)

- How do we assess the direction of the omitted variable problem?

⇒ use auxiliary regression analysis.

$y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + u_i$ (True model).

$y_i = \hat{\beta}_1 + \hat{\beta}_2 x_{2i} + \hat{\beta}_3 x_{3i} + \hat{u}_i$ (Ideal estimation)

Auxiliary regression x_{3i} on 1 and x_{2i}

: $x_{3i} = \hat{\pi}_0 + \hat{\pi}_1 x_{2i} + \hat{v}_i$

$$
\begin{aligned}
y_i &= \hat{\beta}_1 + \hat{\beta}_2 x_{2i} + \hat{\beta}_3 x_{3i} + \hat{u}_i \\
&= \hat{\beta}_1 + \hat{\beta}_2 x_{2i} + \hat{\beta}_3 [\hat{\pi}_0 + \hat{\pi}_1 x_{2i} + \hat{v}_i] + \hat{u}_i \\
&= \hat{\beta}_1 + \hat{\beta}_3 \hat{\pi}_0 + \underline{(\hat{\beta}_2 + \hat{\beta}_3 \hat{\pi}_1)} x_{2i} + \hat{u}_i + \hat{\beta}_3 \hat{v}_i
\end{aligned}
$$

$\xrightarrow{\not p} \beta_2$: inconsistent if $\hat{\beta}_3 > 0$, $\hat{\pi}_1 > 0$.

(Continuing ...)