- Inference allowing for general heteroskedasticity.

$E(u_i^2|x_i) = \sigma^2$ does not need to hold.

We allow $E(u_i^2|x_i) = \sigma^2(x_i)$

Asymptotic variance-covariance matrix of OLS : $E\{x_i x_i'\}^{-1} E\{u_i^2 x_i x_i'\} E\{x_i x_i'\}^{-1}$

$$E\{u_i^2 x_i x_i'\} \xleftarrow{a.s.} \frac{1}{N}\sum_{i=1}^N (y_i - x_i'\hat{\beta})^2 x_i x_i' , \quad E\{x_i x_i'\} \xleftarrow{a.s.} \frac{1}{N}\sum_{i=1}^N x_i x_i'$$

Let $\hat{\Gamma} \to \Gamma$ where $\Gamma\Gamma' = E\{u_i^2 x_i x_i'\}$ $\quad \left(\hat{\Gamma}\hat{\Gamma}' = \frac{1}{N}\sum_{i=1}^N \hat{u}_i^2 x_i x_i' \xrightarrow{a.s.} \Gamma\Gamma' = E\{u_i^2 x_i x_i'\}\right)$

Define $\bar{\Gamma} = \left(\frac{1}{N}\sum_{i=1}^N x_i x_i'\right)^{-1}\hat{\Gamma}$

- $\bar{\Gamma}\bar{\Gamma}' = \left(\frac{1}{N}\sum_{i=1}^N x_i x_i'\right)^{-1}\hat{\Gamma}\hat{\Gamma}'\left(\frac{1}{N}\sum_{i=1}^N x_i x_i'\right)^{-1\prime} \xrightarrow{a.s.} E(x_i x_i')^{-1} E\{u_i^2 x_i x_i'\} E(x_i x_i')^{-1}$

$$\left(\frac{1}{N}\sum_{i=1}^N x_i x_i'\right)^{-1} \xrightarrow{a.s.} E(x_i x_i')^{-1}, \quad \hat{\Gamma}\hat{\Gamma}' \to \Gamma\Gamma', \quad \left(\frac{1}{N}\sum_{i=1}^N x_i x_i'\right)^{-1\prime} \xrightarrow{a.s} E(x_i x_i')^{-1}$$

- $\bar{\Gamma}^{-1} = \hat{\Gamma}^{-1}\left(\frac{1}{N}\sum_{i=1}^N x_i x_i'\right)$

$\sqrt{N}(\hat{\beta} - \beta) \xrightarrow{d} N(0, E\{x_i x_i'\}^{-1} E\{u_i^2 x_i x_i'\} E\{x_i x_i'\}^{-1\prime})$ (By CLT)

$\downarrow$ Multiply $\bar{\Gamma}^{-1}$

$\bar{\Gamma}^{-1}\sqrt{N}(\hat{\beta} - \beta) \xrightarrow{d} \Gamma^{-1}E\{x_i x_i'\} \cdot N(0, E\{x_i x_i'\}^{-1} E\{u_i^2 x_i x_i'\} E\{x_i x_i'\}^{-1})$

$$= N(0, \underbrace{\Gamma^{-1}E\{u_i^2 x_i x_i'\}(\Gamma^{-1})'}_{= I_K})$$

$$\boxed{\text{Note} \quad \text{Let } Z \sim N(0, \underset{r\times r}{V}). \quad \text{Then, } Z'V^{-1}Z \sim \chi^2_{(r)}}$$

By Note,

$\sqrt{N}(\hat{\beta} - \beta)'(\bar{\Gamma}^{-1})'\bar{\Gamma}^{-1}\sqrt{N}(\hat{\beta} - \beta) \xrightarrow{d} \chi^2_{(K)}$

i.e. $\bar{\Gamma}\bar{\Gamma}' = \left(\frac{1}{N}\sum_{i=1}^N x_i x_i'\right)^{-1}\hat{\Gamma}\hat{\Gamma}'\left(\frac{1}{N}\sum_{i=1}^N x_i x_i'\right)^{-1\prime}$ & $\hat{\Gamma}\hat{\Gamma}' = \frac{1}{N}\sum_{i=1}^N \hat{u}_i^2 x_i x_i'$

$\sqrt{N}(\hat{\beta} - \beta)'\left(\frac{1}{N}\sum_{i=1}^N x_i x_i'\right)^{-1}\left(\frac{1}{N}\sum_{i=1}^N \hat{u}_i^2 x_i x_i'\right)\left(\frac{1}{N}\sum_{i=1}^N x_i x_i'\right)\sqrt{N}(\hat{\beta} - \beta) \xrightarrow{d} \chi^2_{(K)}$

→ If we use $\underset{r\times k}{C'}$ for Multiple case, C' is Multiplied to T

→ Considering Wald test,

Under Homoskedasticity,

$E(u_i^2|x_i) = \sigma^2$

$$(C\hat{\beta}-A)'(C'(X'X)^{-1}C)^{-1}(C\hat{\beta}-A)/r \sim F_{(r,N-k)}$$

$$\sqrt{N}\,C'(\hat{\beta}-\beta) \xrightarrow{d} (0, C'E(u_i^2 x_i x_i')C)$$

Under Heteroskedasticity,

$E(u_i^2|x_i) = \sigma^2(x_i)$

$$(C\hat{\beta}-A)'(C'(X'X)^{-1}X'\Omega X(X'X)^{-1}C)^{-1}(C\hat{\beta}-A)/r \sim F(r, N-k)$$

Since

$$\sqrt{N}\,C'(\hat{\beta}-\beta) \xrightarrow{d} (0, C'E(x_i x_i')^{-1}E(u_i^2 x_i x_i')E\{x_i x_i'\}^{-1}C)$$

→ In addition, SSE (restricted, unrestricted) should be changed by Heteroskedasticity.

※ Under Heteroskedasticity,

① Normality assumption can be used. (∵ By CLT, asymotically Normal. Thus, Normal assumption is valid.)

② Ω should be applied to show the Heteroskedasticity

③ SSE (restricted, unrestricted) should be changed for Heteroskedasticiry.

- Collapse of the assumption $E(u_i|x_i)=0$

$y_i = x_i'\beta + u_i$, $E(u_i|x_i)=0 \Leftrightarrow E(y_i|x_i) = x_i'\beta$: correct specification.

☆☆☆ When do we have $E(u_i|x_i) \neq 0$?

① Misspecification (functional)

② Endogeneity of regression (A regressor is correlated with u_i)

$E(u_i|x_i) \neq 0 \Rightarrow E(u_i x_i) \neq 0$.

③ Sample Selection problem (Sample is selected in a way related to dependent variable).

④ Measurement error among regressors.

⑤ Lagged dependent variable among regressors and serial correlation in the residual.

① Misspecification

→ $R'\beta = c^*$ is true but if $R'\beta = c$ is used, estimators may be biased.

② Endogeneity

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

data generation : $y_i = g(x_i) + v_i$, $E(v_i|x_i) = 0$

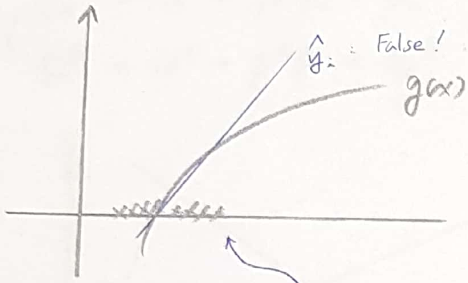
↓ In this case,

$$y_i = \alpha + \beta x_i + \underbrace{g(x_i) - (\alpha + \beta x_i) + v_i}_{= u_i}$$

$$E(u_i|x_i) = \underbrace{g(x_i) - (\alpha + \beta x_i)}_{\neq 0} + \underbrace{E(v_i|x_i)}_{0}$$

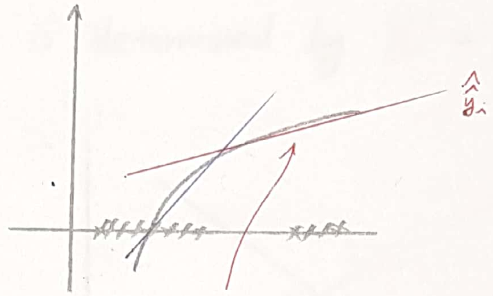
However, we can interpret OLS estimation to be estimating α^* & β^* which solves the following minimization problem :

$$(\alpha^*, \beta^*) = \arg\min_{\alpha,\beta} E\{[g(x_i) - (a+bx_i)]^2\}$$



→ $\hat{y}_i$ is estimated by here

→



→ $\hat{\hat{y}}_i$ is estimated by there

→ Comparing two equations, F-distributions of them are totally different.

→ If we are going to compare two or more groups, we need to make sure the distribution of regressors are comparable.

* Two types of Endogeneity
  - Including variables which we don't need to
  - Not Including variables which we should have.

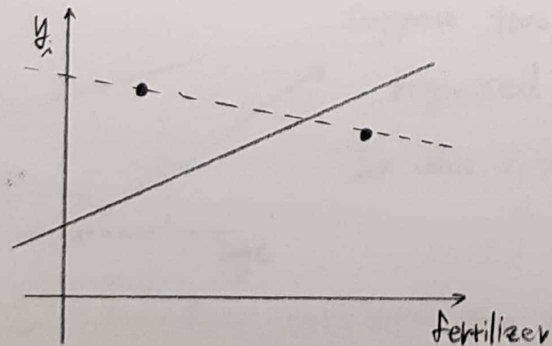
ⓐ production function.

$$\underset{\text{output}}{y_i} = \alpha + \beta \cdot labor_i + \gamma \cdot capital_i + \delta \cdot fertilizer_i + u_i$$

↑ unobserved shock (ex) land quality, weather ...)

⇓

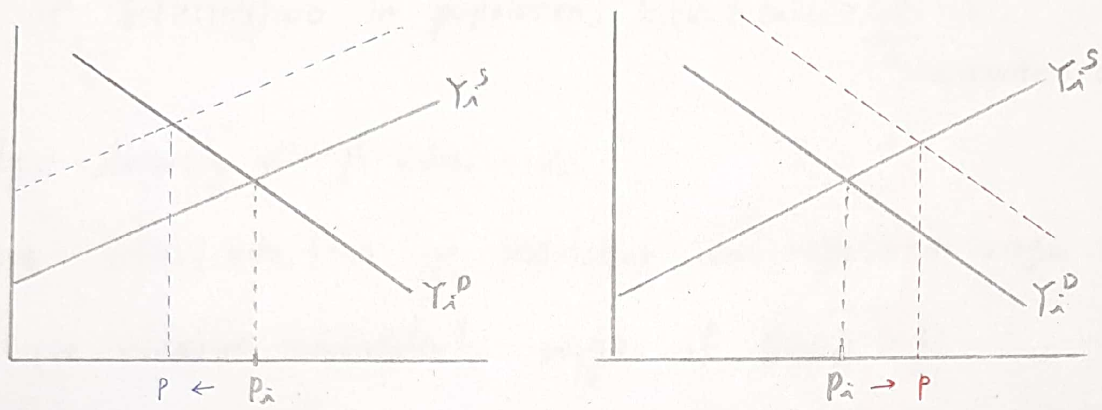
$E(fertilizer_i \cdot u_i) \neq 0$



→ there exists endogeneity, so estimators may be biased.

ⓑ Simultaneity (ex) : supply & demand)

$$\begin{cases} y_i^S = \alpha^S + \beta^S P_i + u_i^S \\ y_i^D = \alpha^D + \beta^D P_i + u_i^D \end{cases}, \quad P_i \text{ is determined by } y_i^S = y_i^D$$



→ Equilibrium model implies correlation of $\begin{pmatrix} u_i^S \;\&\; P_i \\ u_i^D \;\&\; P_i \end{pmatrix}$

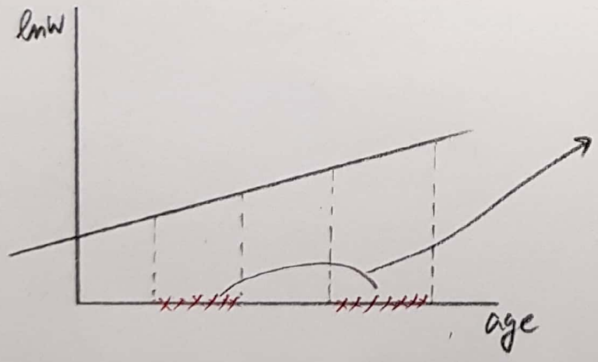
→ Therefore, if we want to use OLS, we should choose the more stable one.

D is stable, then S shifts → Thus, estimate D

S 〃 , then D 〃 → Thus, estimate S

③ Sample selection problem.

$E(u_i | x_i) \neq 0$ does not always arise if sample selection is on x_i's.



Suppose two different age groups are regressed by OLS.

In this case, No problem.

※ Sample selection on dependent variable give rise to $E(u_i | x_i) \neq 0$

→ A leading example is the women's labor supply problem

$\ln w_i = \alpha + \beta \cdot edu_i + \gamma \cdot age_i + \cdots + u_i$

wage is observed for workers only

Even if $E(u_i | x_i) = 0$ in population, $E(u_i | \ln w_i > \ln R_i) \neq 0$.

reservation wage

Simply, $\ln w_i = \alpha + \beta \cdot edu_i + u_i$

Suppose $E(u_i | edu_i) = 0$ in population and reservation wage $= R$.

But we observe individual's wage if $\ln w_i > R$

reservation wage

Then, $E\{\ln w_i | \ln w_i > R, edu_i\}$

$$= E\{\alpha + \beta \cdot edu_i + u_i | \ln w_i > R, edu_i\}$$

$$= \alpha + \beta \cdot edu_i + E\{u_i | \ln w_i > R, edu_i\}$$

$$= \alpha + \beta \cdot edu_i + \underline{E\{u_i | u_i > R - \alpha - \beta edu_i, edu_i\}}$$

Generally, this is a function of edu_i

→ Therefore, we can handle sample selection problems by MLE or semi-parametric/non-parametric methods.