# Econometrics-I

05.02.2021

## Lecture 7

STATISTICAL PROPERTIES OF THE OLS estimator:

population model: $y = \beta_0 + \beta_1 x + u$

"random" sample is drawn

Size $= N$, $\{x_i, y_i\}_{i=1}^{N}$

error.

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

; $i$: person/state/country/firm

Estimate $\beta_0$ and $\beta_1$ using OLS strategy.

$$\min_{\beta_0, \beta_1} \sum_{i=1}^{n} u_i^2 \longrightarrow \hat{\beta}_0(y_i; x_i),\ \hat{\beta}_1(y_i; x_i)$$

residual

$$\hat{u}_i = y_i - \hat{y}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i$$

$$\hat{\beta}_{1,ols} = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{N}(x_i - \bar{x})^2}$$

$$= \frac{\sum_{i=1}^{N}(x_i - \bar{x})\,y_i}{\sum_{i=1}^{N}(x_i - \bar{x})^2} - \cancel{\frac{\bar{y}\sum_{i=1}^{N}(x_i - \bar{x})}{\sum_{i=1}^{N}(x_i - \bar{x})^2}}^{\,0}$$

$$\bar{y}\sum_{i=1}^{N}(x_i - \bar{x}) = 0$$

$$\beta_{1,ols} = \frac{\sum_{i=1}^{N}(x_i - \bar{x})\,y_i}{\underbrace{\sum_i (x_i - \bar{x})^2}} \Rightarrow \hat{\beta}_{1,ols} \text{ is itself a r.v.!!}$$

$$\underline{y_i} = \underbrace{\beta_0 + \beta_1 x_i} + \underset{\text{random variable}}{u_i} \qquad E(u_i|x) = 0$$

$\downarrow$

itself is a random variable

$$\hat{\beta}_{1,OLS} = \frac{\sum_{i=1}^{N}(x_i-\bar{x})y_i}{\sum_{i=1}^{N}(x_i-\bar{x})^2}$$

$$y_i = \beta_0 + \beta_1 x_i + u_i \rightarrow \text{population relation}$$

$$\hat{\beta}_{1\,OLS} = \frac{\sum_{i=1}^{N}(x_i-\bar{x})(\beta_0+\beta_1 x_i + u_i)}{\sum_{i=1}^{N}(x_i-\bar{x})^2}$$

$$= \beta_0 \frac{\cancelto{0}{\overbrace{\sum_{i=1}^{N}(x_i-\bar{x})}^{\text{sum of deviations from mean}}}}{\sum_{i=1}^{N}(x_i-\bar{x})^2} + \beta_1 \frac{\sum_{i=1}^{N}(x_i-\bar{x})x_i}{\sum_{i=1}^{N}(x_i-\bar{x})^2} + \frac{\sum_{i=1}^{N}(x_i-\bar{x})u_i}{\sum_{i=1}^{N}(x_i-\bar{x})^2}$$

$$\hat{\beta}_{1,OLS} = \underbrace{\beta_1 \frac{\sum_{i=1}^{N}(x_i-\bar{x})x_i}{\sum_{i=1}^{N}(x_i-\bar{x})^2}}_{T2} + \underbrace{\frac{\sum_{i=1}^{N}(x_i-\bar{x})u_i}{\sum_{i=1}^{N}(x_i-\bar{x})^2}}_{T3}$$

$$T2 = \beta_1 \frac{\sum_{i=1}^{N}(x_i-\bar{x})x_i}{\sum_{i=1}^{N}(x_i-\bar{x})(x_i-\bar{x})} = \beta_1 \frac{\sum_{i=1}^{N}(x_i-\bar{x})x_i}{\sum_{i=1}^{N}(x_i-\bar{x})x_i - \bar{x}\cancelto{0}{\sum_{i=1}^{N}(x_i-\bar{x})}} = \beta_1$$

$$\hat{\beta}_{1,OLS} = \beta_1 + \underbrace{\frac{\sum_{i=1}^{N}(x_i-\bar{x})u_i}{\sum_{i=1}^{N}(x_i-\bar{x})^2}}_{TB}$$

$TB$ = bias if causality is not attained

$$TB = \underbrace{\frac{\sum_{i=1}^{N} x_i u_i}{\sum_{i=1}^{N}(x_i-\bar{x})^2}}_{=0} - \underbrace{\frac{\bar{x}\sum_{i=1}^{N} u_i}{\sum_{i=1}^{N}(x_i-\bar{x})^2}}_{=0}$$

by using the assumption $E(u|x) = 0$     by using $E(u) = 0$

$$\boxed{E(\hat{\beta}_{1,OLS}) = \beta_1}$$ only if $E(u) = 0$ and $E(u|x) = 0$

Ceteris paribus experiment and causality.

Unbiasedness of the OLS estimator.

Bottomline: If $E(u|x)=0$ fails, the OLS is biased.

$$\underline{E(\hat{\beta}_{1,OLS}) = \beta_1} \quad \checkmark$$

$$V(\hat{\beta}_{1,OLS}) = ?$$

random variable

$$\Rightarrow V(\hat{\beta}_{1,OLS}|x) = V\left(\frac{\sum_{i=1}^{N}(x_i-\bar{x})\,y_i}{\sum_{i=1}^{N}(x_i-\bar{x})^2}\right)$$

$$= \frac{1}{\left[\sum_{i=1}^{N}(x_i-\bar{x})^2\right]^2} V\left(\sum_{i=1}^{N}(x_i-\bar{x})y_i\right)$$

$$V(\hat{\beta}_{1,OLS}|x) = \frac{1}{\left[\sum_{i=1}^{N}(x_i-\bar{x})^2\right]^2} \sum_{i=1}^{N}(x_i-\bar{x})^2\, \underline{V(y_i|x_i)}$$

$$y_i = \beta_0 + \beta_1 x_i + \underbrace{u_i}_{\text{random error.}}$$

we add an assumption that errors are "homoskedastic".

A3 $\quad V(u|x) = \sigma^2$

$$V(u|x) = E_x[u - E_x(u)]^2$$

$$= E_x(u^2) - [E_x(u)]^2$$

$$= E(u^2|x) - \underbrace{[E(u|x)]^2}_{=0}$$

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

Interpretation : Expectation of $u^2$ is independent of $x$.

Remember: • A3 is based on the second moment of the error term

• It is different from $E(u|x)=0$

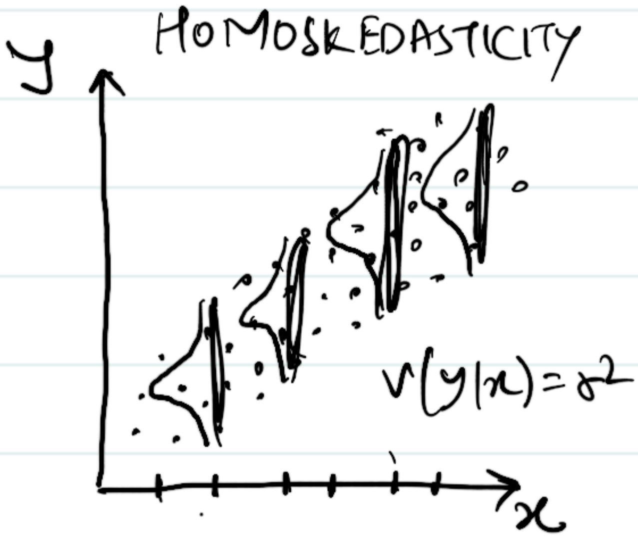
• Homoskedasticity (i.e., A3) is not needed for unbiasedness of the OLS estimator.

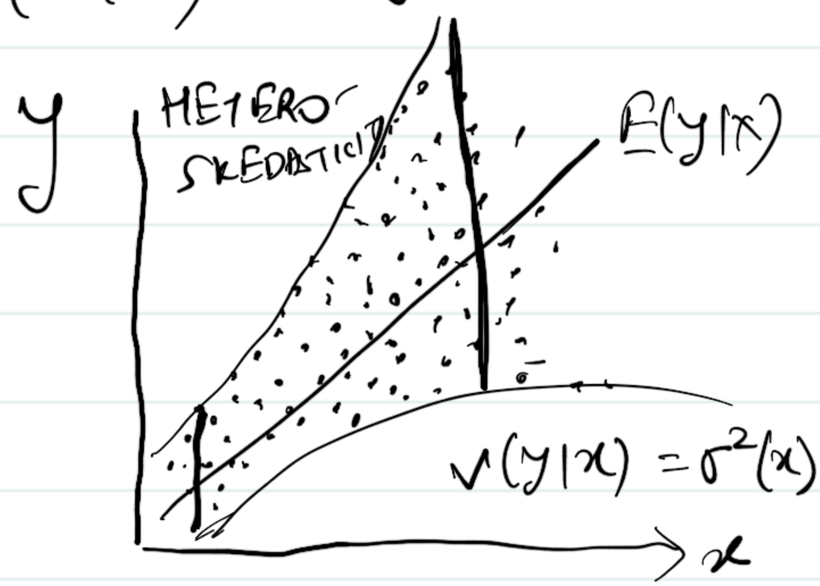
Population model:

$$y = \underline{\beta_0} + \underline{\beta_1 \; x} + u$$

$$E(y|x) = \beta_0 + \beta_1 x$$

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



Therefore,

$$V(\hat{\beta}_{1,ols} \mid X) = \frac{1}{\left[\sum_{i=1}^{N}(x_i-\bar{x})^2\right]^2} \sum_{i=1}^{N}(x_i-\bar{x})^2 \cdot \sigma^2$$

$$= \sigma^2 \frac{\sum_{i=1}^{N}(x_i-\bar{x})^2}{\left[\sum_{i=1}^{N}(x_i-\bar{x})^2\right]^2}$$

$$\boxed{V(\hat{\beta}_{1,ols}) = \frac{\sigma^2}{\sum_{i=1}^{N}(x_i-\bar{x})^2}}$$

$$V(\hat{\beta}_{0,ols}) = \frac{\sigma^2 n^{-1} \sum_{i=1}^{N} x_i^2}{\sum_{i=1}^{N}(x_i-\bar{x})^2}$$ * Home assignment

Aside (Redone for clarity)

$$\hat{\beta}_{1,OLS} = \frac{\sum_{i=1}^{N} (x_i - \bar{x}) y_i}{\sum_{i=1}^{N} (x_i - \bar{x})^2}$$

$$V(\hat{\beta}_{1,OLS}|X) = V\left( \frac{\sum_{i=1}^{N} (x_i - \bar{x}) y_i}{\sum_{i=1}^{N} (x_i - \bar{x})^2} \right)$$

$$= \frac{1}{\left[\sum_{i=1}^{N} (x_i - \bar{x})^2\right]^2} V\left( \sum_{i=1}^{N} (x_i - \bar{x}) y_i \right)$$

$$= \frac{1}{[\cdot]^2} \sum_{i=1}^{N} \left\{ V\left[ (x_i - \bar{x}) y_i \right] \right\}$$

$$= \frac{1}{[\cdot]^2} \sum_{i=1}^{N} (x_i - \bar{x})^2 \underbrace{V(y_i|X_i)}_{= \sigma^2 \text{ (By A3)}}$$

$$= \frac{\sigma^2}{\sum_{i=1}^{N} (x_i - \bar{x})^2}$$

By now: $\hat{\beta}_{1,OLS}$ ✓, $\hat{\beta}_{0,OLS}$ ✓ $\rightarrow$ functions of data.

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

✓ $E(\hat{\beta}_{1,OLS}) = \beta_1$ $\quad$ ($\because$ A1, A2)

✓ $E(\hat{\beta}_{0,OLS}) = \beta_0$

✓ $V(\hat{\beta}_{1,OLS}) = \dfrac{\sigma^2}{\sum_{i=1}^{N}(x_i - \bar{x})^2}$, $\quad V(\hat{\beta}_{0,OLS}) =$ HOME ASSIGN-MENT

What is $\sigma^2 = ?$

Need $\hat{\sigma}^2_{OLS}$ from the data: s.t. $E(\hat{\sigma}^2) = \sigma^2$.

- $\sigma^2$ is a second-order moment of $u_i$; $\boxed{V(u_i|x_i) = \sigma^2}$
- So, the sample estimate can be written as:

✓ $$\hat{\sigma}^2 \overset{?}{=} \frac{\sum_{i=1}^{N} \hat{u}_i^2}{N - \underbrace{2}_{??}}$$ $\rightarrow$ Unbiased sample estimate of $V(u|x) = \sigma^2$?

Aside

we say an unbiased sample estimator of mean of a r.v. $Z$ is

$$V(Z) = \frac{\sum_{i=1}^{N} (Z_i - \bar{Z})^2}{N-1}$$

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{N} (\hat{u}_i - \overset{=0}{\bar{\hat{u}}})^2}{\underbrace{N-2}} = \frac{\sum_{i=1}^{N} \hat{u}_i^2}{N-2}$$

$$\left.\begin{aligned} \because \sum_{i=1}^{N} \hat{u}_i &= 0 \\ \sum_{i=1}^{N} \hat{u}_i x_i &= 0 \end{aligned}\right\}$$ we know these two conditions will always be true

SAMPLE PROP. OLS estimator

$$\boxed{E(\hat{\sigma}^2) = \sigma^2}$$

* → will Do the proof in next lecture.