Econometrics-I

2.2.2021

Lecture 6

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

$$\hat{\beta}_{0,OLS} = \bar{y} - \hat{\beta}_{1,OLS}\,\bar{x}$$

How?

$$\underbrace{\min_{\beta_0, \beta_1} \sum_{i=1}^{N} u_i^2}$$

Properties of the OLS estimators on any sample of data.

$$\hat{y}_i = \hat{\beta}_{0,OLS} + \hat{\beta}_{1,OLS}\, x_i \quad ; \quad y_i = \hat{y}_i + \hat{u}_i$$

$$\underbrace{\hat{u}_i}_{\text{(residuals)}} = y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i) = \underbrace{y_i}_{\substack{\text{actuals} \\ \text{(truth)}}} - \underbrace{\hat{y}_i}_{\text{prediction}}$$

- due to the OLS estimators based on the given sample.
- not the same as the error term in the original model.

All OLS estimators: $\hat{\beta}_0$, $\hat{\beta}_1$, $\hat{u}_i$

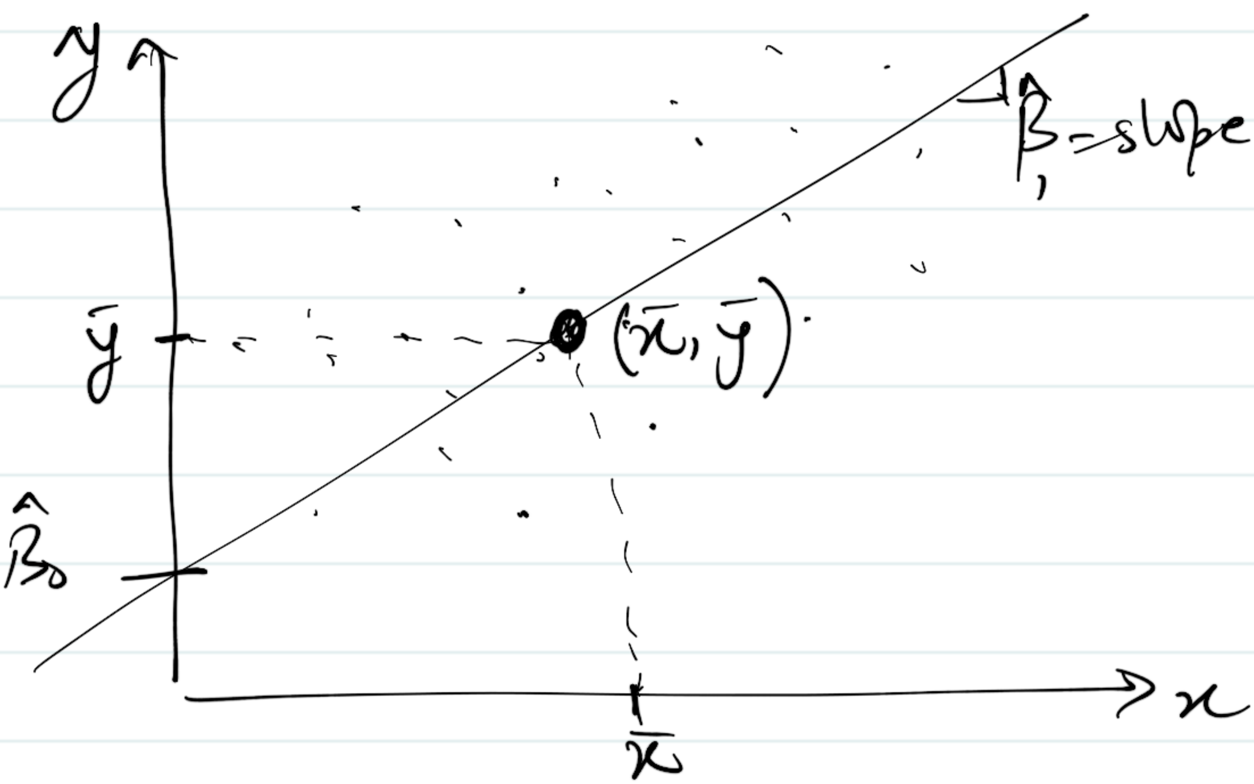
(i) $\sum_{i=1}^{N} \hat{u}_i = 0$ $\Rightarrow \sum_{i=1}^{N} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0$

FOCs, optimization

(ii) $\sum_{i=1}^{N} \hat{u}_i x_i = 0$ $\Rightarrow \sum_{i=1}^{N} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) x_i = 0$

(iii) The point $(\bar{x}, \bar{y})$ is always on the $\underbrace{\text{OLS line.}}$

$$\bar{y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{x}$$



In essense,

$$\underbrace{y_i}_{\substack{\text{actual}\\ \text{variation}\\ \text{in data.}}} = \underbrace{\hat{y}_i}_{\substack{\text{predicted}\\ \text{variation}\\ \text{in data}}} + \underbrace{\hat{u}_i}_{\substack{\text{residual variation}\\ \text{in data.}}}$$

$$\boxed{\text{Goodness of fit}}$$

Define:

$$SST = \sum_{i=1}^{N} (y_i - \bar{y})^2 \Rightarrow \text{Sum of squares - total}$$

$$SSE = \sum_{i=1}^{N} (\hat{y}_i - \bar{y})^2 \Rightarrow \text{sum of squares - Explained}$$

$$SSR = \sum_{i=1}^{N} (\hat{u}_i - \bar{\hat{u}})^2 \Rightarrow \text{Sum of squares - Residual}$$

Interpretation: SST → total variation in data
SSE → explained variation in data
SSR → residual variation in data

$$SST = SSE + SSR$$

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

Proof :

Start w/ LHS:

$$\overbrace{\sum_{i=1}^{N}(y_i - \bar{y})^2}^{SST} = \sum_{i=1}^{N}\left(y_i - \bar{y} + \hat{y}_i - \hat{y}_i\right)^2$$

$$= \sum_{i=1}^{N}\Big[\underbrace{(y_i - \hat{y}_i)}_{=\hat{u}_i} + (\hat{y}_i - \bar{y})\Big]^2$$

$$= \sum_{i=1}^{N}\left[\hat{u}_i^2 + (\hat{y}_i - \bar{y})^2 + 2\hat{u}_i(\hat{y}_i - \bar{y})\right]$$

$$= \underbrace{\sum_{i=1}^{N}\hat{u}_i^2}_{SSR} + \underbrace{\sum_{i=1}^{N}(\hat{y}_i - \bar{y})^2}_{SSE} + \underbrace{2\sum_{i=1}^{N}\hat{u}_i(\hat{y}_i - \bar{y})}_{=0}$$

(How and why?)

$$\sum_{i=1}^{N} \hat{u}_i(\hat{y}_i - \bar{y}) = \underbrace{\sum_{i=1}^{N} \hat{u}_i \hat{y}_i}_{} - \bar{y}\underbrace{\sum_{i=1}^{N} \hat{u}_i}_{=0}$$

$\sum_{i=1}^{N} \hat{u}_i = 0$: 1st prop of OLS estimator

$$= \sum_{i=1}^{N} \hat{u}_i(\hat{\beta}_0 + \hat{\beta}_1 x_i)$$

$$= \hat{\beta}_0 \underbrace{\sum_{i=1}^{N} \hat{u}_i}_{=0} + \hat{\beta}_1 \underbrace{\sum_{i=1}^{N} \hat{u}_i x_i}_{=0}$$

1st prop of OLS estimator | 2nd prop of OLS estimator.

Hence:

$$\underbrace{\sum_{i=1}^{N}(y_i - \bar{y})^2}_{\text{Total variation in data}} = \underbrace{\sum_{i=1}^{N}(\hat{y}_i - \bar{y})^2}_{\text{explained}} + \underbrace{\sum_{i=1}^{N} \hat{u}_i^2}_{\text{residual}}$$

Goodness of fit: $\dfrac{SSE}{SST} \times 100$

% of sample variation in y_i explained by x_i via SLRM.

$$R^2 = \frac{SSE}{SST} = 1 - \frac{SSR}{SST}$$

$R^2 \in (0,1)$

$100 \cdot R^2 \rightarrow$ % of sample variation explained by the SLRM.

Note: Low R^2 values are not uncommon in social sciences

In cross-sectional regression, low R^2 does not mean that the regression is useless.

$E(u|x) = 0 \rightarrow$ key assumption for causality (simulating the ceteris paribus experiment)

The function form of the simple linear regression model and interpretation of β_1

$$\rightarrow y_i = \underbrace{\beta_0 + \beta_1 x_i} + u_i$$

$$\frac{\partial y_i}{\partial x_i} = \beta_1$$ for all x_i values

Model	Y	X	Interpretation (β_1)
Level-level	y_i	x_i	$\Delta y = \beta_1 \Delta x$
log-level	$\log(y_i)$	x_i	$\% \Delta y = \beta_1 \Delta x$
level-log	y_i	$\log(x_i)$	$\Delta y = \beta_1 (\% \Delta x)$
log-log	$\log(y_i)$	$\log(x_i)$	$\% \Delta y = \beta_1 (\% \Delta x)$

$$\log(y_i) = \beta_0 + \beta_1 x_i + u_i \Rightarrow \beta_1 = \frac{\partial \log y_i}{\partial x_i} = \frac{\frac{1}{y_i}\partial y_i}{\partial x_i}$$

$$\beta_1 = \frac{1}{y_i}\frac{dy_i}{dx_i}$$ = % change in y upon 1 unit change in x.

$$\Rightarrow \boxed{\beta_1 y_i} = dy_i/dx_i = f(y_i)$$

log - log model

$$\log(y_i) = \beta_0 + \beta_1 \log(x_i) + u_i$$

why is this a SLRM ?

Linear in parameters!

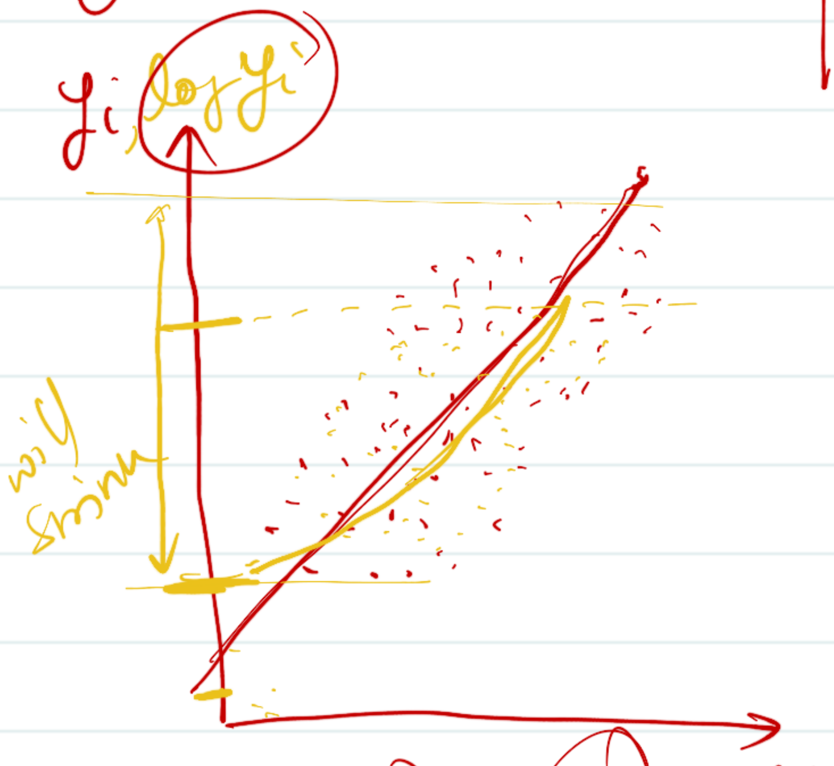
$$\beta_1 = \frac{\partial \log(y_i)}{\partial \log(x_i)} = \frac{\frac{1}{y_i}\partial y_i}{\frac{1}{x_i}\partial x_i} = \frac{\%\text{ change in } y}{\%\text{ change in } x}$$

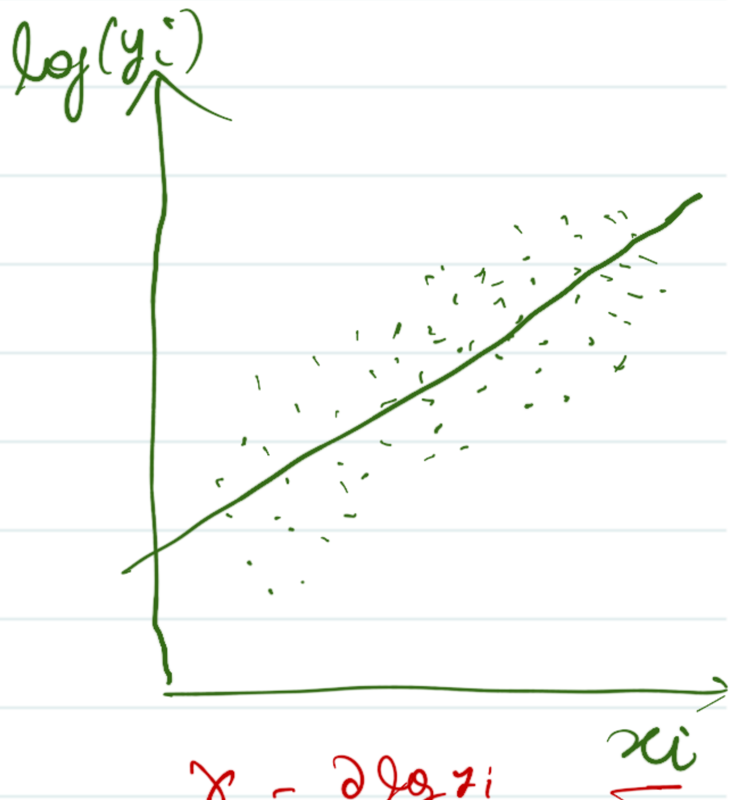
$\Rightarrow$ elasticity of y w.r.t. change in x.

$$\frac{\partial y_i}{\partial x_i} = \beta_1 \left[\frac{y_i}{x_i}\right]$$

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

$\log(y_i) = \gamma_0 + \gamma_1 x_i + u_i$



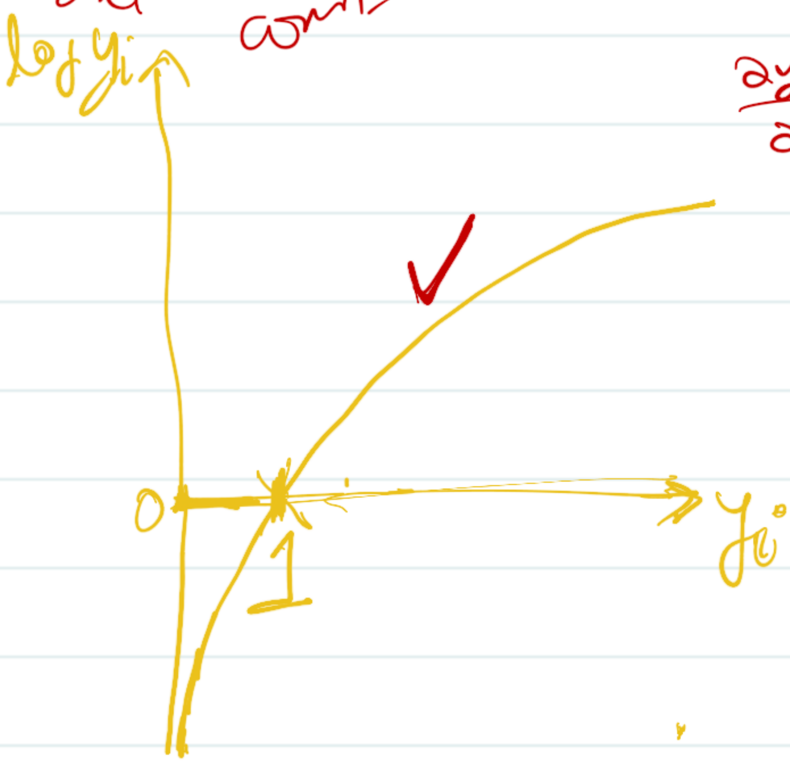


$\frac{\partial y_i}{\partial x_i}$ = β_1 const

$\gamma_1 = \frac{\partial \log y_i}{\partial x_i}$

$\frac{\partial y_i}{\partial x_i} = \gamma_1 y_i$ not a const.

varies by y_i



$y = f(x) + \varepsilon$