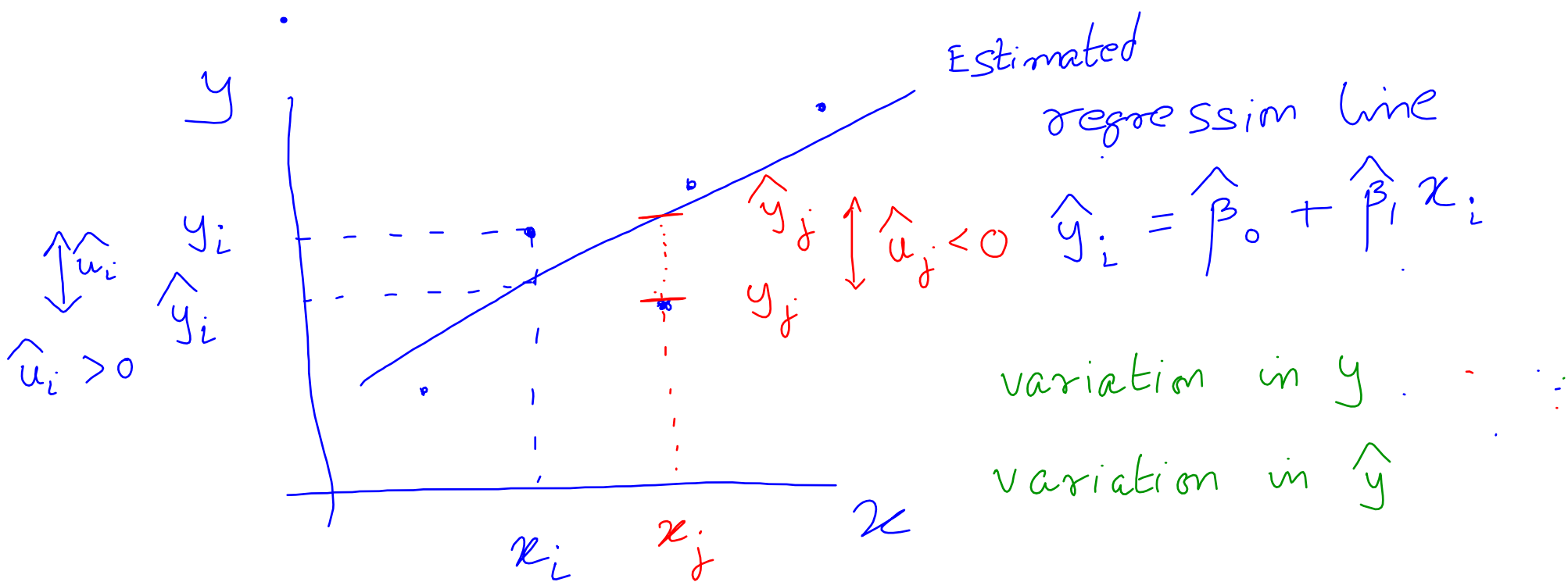




For i th obs. y_i : dep. var. value
 $\hat{y}_i$: fitted value

residual $\leftarrow \hat{u}_i = y_i - \hat{y}_i$

$\hat{\beta}_0$ and $\hat{\beta}_1$ also minimize the
sum of squared residuals (SSR)

$$\rightarrow \sum_{i=1}^n \hat{u}_i^2$$

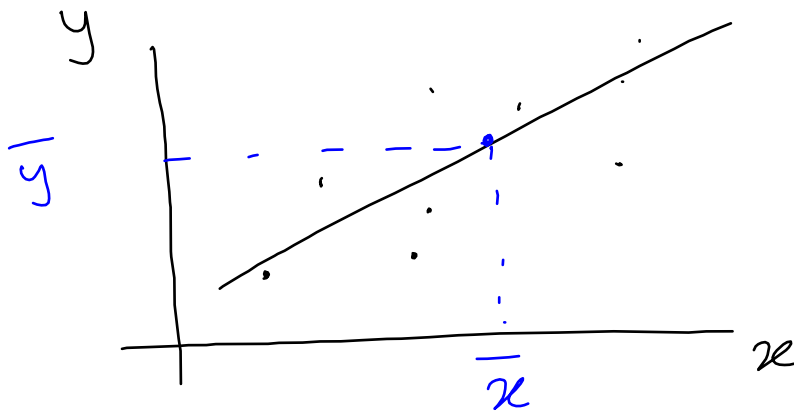
Method $\rightarrow$ also called ordinary
least squares (OLS)

$$\hat{\beta}_1 = \frac{\text{how } x \text{ \& } y \text{ covary}}{\text{how } x \text{ varies}}$$

$$= \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

$\bar{x}, \bar{y}$:

sample means



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

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

Properties of OLS

Sum/avg. of OLS residuals = 0

$$\sum_{i=1}^n \hat{u}_i = 0$$

Sample correl. b/w $\hat{y}$ & $\hat{u} = 0$

Sample correl. b/w x & $\hat{u} =$

$$\sum_{i=1}^n x_i \hat{u}_i = 0$$

$$\sum_{i=1}^n \hat{y}_i \hat{u}_i = 0$$

$(\bar{x}, \bar{y}) \rightarrow$ on the OLS reg. line

$$\text{Total variation} = \underset{\substack{\text{by} \\ \text{model}}}{\text{explained}} + \underset{\text{residual}}{\text{not explained}}$$

For each obs.

$$y_i = \hat{y}_i + \hat{u}_i$$

Diagram illustrating the decomposition of total variation:

- y_i (Total variation) is the sum of $\hat{y}_i$ (explained variation) and $\hat{u}_i$ (residual variation).
- Total variation (sum of squares) SST is represented by the formula: $\sum_{i=1}^n (y_i - \bar{y})^2$.
- Explained variation SSE is represented by the formula: $\sum_{i=1}^n (\hat{y}_i - \bar{y})^2$.
- Residual variation SSR is represented by the formula: $\sum_{i=1}^n \hat{u}_i^2$.

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

$$0 \leq R^2 \leq 1$$

R^2 : also square of correl. b/w y and $\hat{y}$

High R^2 : not the ultimate objective in causal analysis