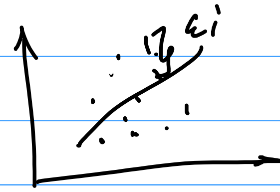
Omitted Variable Bias

$\bar{x} \xrightarrow{P} E(X_i)$

$\widehat{Cov}(x,y) \xrightarrow{P} Cov(X_i, Y_i)$

Deterministic Regressors:

True: $y_i = \beta_1 + \beta_2 \cdot X_i + \beta_3 W_i + \varepsilon_i$



Est: $y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + u_i$

$E(\varepsilon_i) = 0$, $E(\varepsilon_i^2) = \sigma^2$,

$E(\varepsilon_i, \varepsilon_j) = 0$

$\Rightarrow \hat{\beta}$ - BLUE

$u_i = \beta_3 W_i + \varepsilon_i$

$E(u_i) = \beta_3 W_i$

$$\hat{\beta}_2 = \frac{\widehat{Cov}(x,y)}{\widehat{Var}(x)} = \frac{\widehat{Cov}(x, \beta_1 + \beta_2 \cdot X + \beta_3 W + \varepsilon)}{\widehat{Var}(x)} =$$

$$= \beta_2 + \beta_3 \frac{\widehat{Cov}(x, w)}{\widehat{Var}(x)} + \frac{\widehat{Cov}(x, \varepsilon)}{\widehat{Var}(x)} \neq 0$$

$\bar{\hat{\varepsilon}} = 0$

$E(\varepsilon) = 0$

$\bar{\varepsilon} \neq 0$

$$E(\hat{\beta}_2) = \beta_2 + \beta_3 E\left(\frac{\widehat{Cov}(x,w)}{\widehat{Var}(x)}\right) + E\left(\frac{\widehat{Cov}(x,\varepsilon)}{\widehat{Var}(X)}\right)$$

$$= \beta_2 + \beta_3 \frac{\widehat{Cov}(x,w)}{\widehat{Var}(x)} + \frac{E(\widehat{Cov}(x,\varepsilon))^{=0}}{\widehat{Var}(x)}$$

$\frac{\sum \varepsilon_i}{n}$

$$E\left(\frac{1}{n-1} \sum (x_i - \bar{x})(\varepsilon_i - \bar{\varepsilon})\right) = \sum \frac{1}{n-1} (x_i - \bar{x}) (E(\varepsilon_i) - E(\bar{\varepsilon}))$$

$E(\varepsilon_i) = 0$, $E(\bar{\varepsilon}) = 0$

$$E(\hat{\beta}_2) = \beta_2 + \underbrace{\beta_3 \cdot \frac{\widehat{Cov}(x,w)}{\widehat{Var}(x)}}$$

if $\beta_3 = 0$ or $x \perp w$ $\Rightarrow$ unbiased

$\sum x_i w_i = 0$

Stochastic Regressors:

True: $y_i = \beta_1 + \beta_2 \cdot X_i + \beta_3 W_i + \varepsilon_i$

Est: $y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + u_i$

$\varepsilon \perp X \Rightarrow E(\varepsilon_i | X) = 0 \Rightarrow \operatorname{cov}(\varepsilon_i, X_i) = 0$

$$\left[\begin{array}{l} E(\varepsilon_i^2 | X) = 0 \\ E(\varepsilon_i \varepsilon_j | X) = 0 \end{array} \right]$$

$\operatorname{cov}(\varepsilon_i, X_i) = 0$

$\Rightarrow X_i$ - exogenous

$\operatorname{cov}(\varepsilon_i, X_i) \neq 0$

$\Rightarrow X_i$ - endogenous

① Deterministic (t)

$y_t = \beta_1 + \beta_2 \cdot t + \varepsilon_t$

② Stochastic (π_{t-1})

$[\pi_t = \beta_1 + \beta_2 \pi_{t-1} + \varepsilon_t]$

$\pi_{t-1} = \beta_1 + \beta_2 \pi_{t-2} + \varepsilon_{t-1}$

True: $y_i = \beta_1 + \beta_2 \cdot X_i + \beta_3 W_i + \varepsilon_i$

Est: $y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + u_i$

$$\widehat{\operatorname{cov}}(X,Y) = \sum \frac{1}{n-1}(x_i - \bar{x})(y_i - \bar{y})$$

$$\operatorname*{plim}_{n \to \infty} \hat{\beta}_2 = \operatorname*{plim}_{n \to \infty} \frac{\widehat{\operatorname{cov}}(X,Y)}{\widehat{\operatorname{var}}(X)} = \frac{\operatorname{Cov}(X_1, Y_1)}{\operatorname{Var}(X_1)} =$$

$$= \frac{\operatorname{Cov}(X_1, \beta_1 + \beta_2 X_1 + \beta_3 W_1 + \varepsilon_1)}{\operatorname{Var}(X_1)} =$$

$$= \beta_2 + \beta_3 \frac{\operatorname{Cov}(X_1, W_1)}{\operatorname{Var}(X_1)} + \frac{\overbrace{\operatorname{Cov}(X_1, \varepsilon_1)}^{=0}}{\operatorname{Var}(X_1)}$$

$\operatorname{Cov}(X_1, X_1) = \operatorname{var}(X_1)$

$\star \quad \left[E(\varepsilon|X)\right] = 0 \quad \Rightarrow \quad Cov(\varepsilon, X) = 0$

$$Cov(\varepsilon, x) = E(\varepsilon \cdot X) - \underset{\underset{E(E(\varepsilon|X))}{\parallel}}{E(\varepsilon)}E(X) = 0$$

$$E(\varepsilon \cdot X) = E(E(\varepsilon \cdot X|X)) = E(X \cdot \underset{\overset{\parallel}{0}}{E(\varepsilon|X)})$$

Deterministic Regressors:

True: $y_i = \beta_1 + \beta_2 X_i + \varepsilon_i$

Est: $y_i = \hat{\beta}_1 + \hat{\beta}_2 \cdot X_i + \hat{\beta}_3 \omega_i + \hat{\varepsilon}_i$

- $Var(\hat{\beta}_k) = \dfrac{\sigma^2_\varepsilon}{TSS_k(1-R^2_k)}$ $\quad R_i$: R^2 of $x_i|X_{-i}$

$$Var(\hat{\beta}_2) = \frac{\sigma^2_\varepsilon}{TSS_2 \cdot (1-\hat{\rho}^2_{x,y})}$$

- $df \downarrow \quad$ t-test: $t \sim t_{N-K}$