Omitted Variable Bias

True: $y_i = \beta_1 + \beta_2 x_{1i} + \beta_3 X_{2i} + \varepsilon_i$

Est : $y_i' = \beta_1 + \beta_2 X_{1i} + u_i$

$u_i = \beta_3 X_{2i} + \varepsilon_i$

$E(u_i) \neq 0$

$E(\varepsilon_i) = 0$ — Det.

$E(\varepsilon_i | X) = 0$ — Stoch

X - exogenous if $cov(x, \varepsilon) = 0$

X - endogenous if $cov(x, \varepsilon) \neq 0$

X, Y - independent

$\Downarrow$

$E(X|Y) = E(X)$ - non-predictable

$\Downarrow$

$Cov(X,Y) = 0$ - uncorrelated

True: $y_i = \beta_1 + \beta_2 x_{1i} + \beta_3 X_{2i} + \varepsilon_i$

Est : $\hat{y}_i' = \hat{\beta}_1 + \hat{\beta}_2 X_{1i} + u_i$

$$\hat{\beta}_2 = \frac{\hat{Cov}(x_1, y)}{\hat{Var}(x_1)} = \frac{\hat{Cov}(x_1, \beta_1 + \beta_2 X_1 + \beta_3 X_2 + \varepsilon)}{\hat{Var}(x_1)} =$$

$$= \frac{\beta_2 \hat{Var}(x_1) + \beta_3 \hat{Cov}(x_1, x_2) + \hat{Cov}(x_1, \varepsilon)^{\neq 0}}{\hat{Var}(x_1)}$$

$$= \beta_2 + \beta_3 \frac{\hat{Cov}(X_1, X_2)}{\hat{Var}(X_1)} + \frac{\hat{Cov}(X_1, \varepsilon)}{\hat{Var}(X_1)}$$

$$E(\hat{\beta}_2) = \beta_2 + \beta_3 E\left(\frac{\hat{Cov}(X_1, X_2)}{\hat{Var}(X_1)} \right) + E\left(\frac{\hat{Cov}(X_1, \varepsilon)}{\hat{Var}(X_1)} \right)$$

$$\hat{Var}(X_1) = E\left(\frac{1}{n-1} \Sigma (X_i - \bar{X})^2 \right)$$

$$= \beta_2 + \beta_3 \cdot \frac{\hat{Cov}(X_1, X_2)}{\hat{Var}(X_1)} + \frac{E(\hat{Cov}(X_1, \varepsilon))}{\hat{Var}(X_1)} = \beta_2 + \beta_3 \frac{\hat{Cov}(X_1, X_2)}{\hat{Var}(X_1)}$$

$$\circledast \quad E\left(\frac{1}{n-1} \Sigma (X_i - \bar{X})(\varepsilon_i - \bar{\varepsilon}) \right) =$$

$$\frac{1}{n-1} \Sigma (X_i - \bar{X}) \Big(\underbrace{E \varepsilon_i}_{=0} - \underbrace{E(\bar{\varepsilon})}_{=0} \Big) = 0$$

$$E(\hat{\beta}_2) = \beta_2 + \underbrace{\beta_3 \frac{\hat{Cov}(X_1, X_2)}{\hat{Var}(X_1)}}_{\text{bias}}$$

$\nearrow$ $\beta_3 = 0$ $\quad X_2 \not\to y$

$\searrow$ $\hat{Cov}(X_1, X_2) = 0$

(*) Stochastic Regressor case

$\bar{X} \xrightarrow{P} E(X)$

$\hat{Var}(X) \xrightarrow{P} Var(X)$

$u_i = \beta_3 \cdot X_{2i} + \varepsilon_i$ $\quad Cov(X_{1i}, u_i) = Cov(X_{1i}, \beta_3 X_{2i} + \varepsilon_i)$

$\rightarrow$ True: $y_i = \beta_1 + \beta_2 X_{1i} + \beta_3 X_{2i} + \varepsilon_i$ $\quad$ TGM:

Est: $\hat{y}_i' = \hat{\beta}_1 + \hat{\beta}_2 X_{1i} + \underline{u_i}$

① Deterministic (t)

$y_i = \beta_1 + \beta_2 \cdot t + \varepsilon_i$

② Stochastic (π_t)

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

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

$$\operatorname*{plim}_{n\to\infty} \hat{\beta}_2 = \operatorname*{plim}_{n\to\infty} \frac{\hat{cov}(X_1, y)}{\hat{Var}(X_1)} =$$

$$= \frac{Cov(X_1, y)}{Var(X_1)} =$$

$$= \frac{Cov(X_{1i}, \beta_1 + \beta_2 \cdot X_{1i} + \beta_3 X_{2i} + \varepsilon_i)}{Var(X_{1i})} =$$

$$= \frac{\beta_2 \cdot Var(X_{1i}) + \beta_3 Cov(X_{1i}, X_{2i}) + \overbrace{Cov(X_{1i}, \varepsilon_i)}^{=0}}{Var(X_{1i})} =$$

$$= \beta_2 + \underbrace{\frac{\overset{<0}{\beta_3} \cdot \overset{>0}{cov(X_{1i}, X_{2i})}}{Var(X_{1i})}}_{\text{bias}} \Rightarrow \hat{\beta}_2 - \text{inconsistent}$$

bias $\nearrow$ $\beta_3 = 0$

$\searrow$ $cov(X_1, X_2)$

$C \mid Y, U$ $\qquad C \mid Y$

$\hat{\beta}_y$

$\beta_3 < 0$

$cov(U, Y) < 0$

$\Rightarrow$ bias > 0

True: $y_i = \beta_1 + \beta_2 X_{1i} + u_i$

Est: $\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_{1i} + \hat{\beta}_3 X_{2i} + \varepsilon_i$

$E(\varepsilon_i) = 0$

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

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

TGM: $\hat{\beta}$

$$Var(\hat{\beta}_i) = \frac{\sigma_\varepsilon^2}{TSS_i(1 - R_i^2)}$$

• $Var(\hat{\beta}_2) = \dfrac{\sigma_\varepsilon^2}{TSS(1 - \hat{\rho}^2_{X_1, X_2})}$

R_i^2 : R^2 $X_i | X_{-i}$

• t - test

$t \sim t_{n-k}$

TSS : $\Sigma(X_i - \bar{X})^2$

df ↓

$$\bullet\ \hat{\beta}_2 = \frac{\widehat{Cov}(X_1, Y) \cdot \widehat{Var}(X_2) - \widehat{Cov}(X_2, Y) \cdot \widehat{Cov}(X_1, X_2)}{\widehat{Var}(X_1)\widehat{Var}(X_2) - \widehat{Cov}(X_1, X_2)}$$

Q2

$y_i = \gamma_1 + \gamma_2 \cdot x_{1i} + \gamma_3 X_{2i} + \varepsilon_i \Rightarrow \hat{y}$

RESET (Ramsey) Test

$\hat{y}^2 = (\hat{\gamma}_1 + \hat{\gamma}_2 x_{1i} + \hat{\gamma}_3 x_{2i})^2$

$y_i = \hat{\gamma}_1 + \hat{\gamma}_2 \cdot x_{1i} + \hat{\gamma}_3 X_{2i} + \hat{\gamma}_4 \hat{y}_i^2 + \varepsilon_i$

$H_0 : \gamma_4 = 0$

$t \sim t_{n-4}$