1) Omitted Variables

Vitor Kamada

July 2018

OLS Omitted Variables Bias

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k + \gamma q + v$$

$$q = \delta_0 + \delta_1 x_1 + ... + \delta_k x_k + r$$

$$y = (\beta_0 + \gamma\delta_0) + (\beta_1 + \gamma\delta_1)x_1 + (\beta_2 + \gamma\delta_2)x_2 + ... + (\beta_k + \gamma\delta_k)x_k + v + \gamma r$$

$$plim\hat{\beta}_k = \beta_k + \gamma[\frac{Cov(x_k, q)}{Var(x_k)}]$$

OVB Example

$$log(wage) = \beta_0 + \beta_1 exper + \beta_2 exper^2 + \beta_3 educ + \gamma abil + v$$

$$abil = \delta_0 + \delta_3 educ + r$$

$$\hat{\beta}_3 = \beta_3 + \gamma\delta_3$$

Proxy Variable

$$E(y|x,q,z) = E(y|x,q) \tag{1}$$

$$L(q|1,x_1,...,x_k,z) = L(q|1,z) \tag{2}$$

$$q = \theta_0 + \theta_1 z + r$$

$$Cov(x_j, r) = 0, \quad j = 1,2,...,K$$

Perfect Proxy

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k + \gamma q + v$$

$$q = \theta_0 + \theta_1 z + r$$

$$y = (\beta_0 + \gamma\theta_0) + \beta_1 x_1 + ... + \beta_k x_k + \gamma\theta_1 z + (\gamma r + v)$$

Imperfect Proxy

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k + \gamma q + v$$

$$q = \theta_0 + \rho_1 x_1 + ... + \rho_k x_k + \theta_1 z + r$$

$$plim\hat{\beta}_j = \beta_j + \gamma\rho_j$$

If $\theta_1 \neq 0$,

$$Var(\gamma r + v) < Var(\gamma q + v)$$

Blackburn and Neumark (1992)

$$\log(\hat{wage}) = \underset{(0.11)}{5.40} + \underset{(.003)}{.014\, exper} + \underset{(.002)}{.012\, tenure} + \underset{(.039)}{.199\, married}$$

$$- \underset{(.026)}{.091\, south} + \underset{(.027)}{.184\, urban} - \underset{(.038)}{.188\, black} + \underset{(.006)}{.065\, educ}$$

$$N = 935, \qquad R^2 = .253$$

$$\log(\hat{wage}) = \underset{(0.13)}{5.18} + \underset{(.003)}{.014\, exper} + \underset{(.002)}{.011\, tenure} + \underset{(.039)}{.200\, married}$$

$$- \underset{(.026)}{.080\, south} + \underset{(.027)}{.182\, urban} - \underset{(.039)}{.143\, black} + \underset{(.007)}{.054\, educ}$$

$$+ \underset{(.0010)}{.0036\, IQ}$$

$$N = 935, \qquad R^2 = .263$$

Effects of Job Training Grants on Worker Productivity

$$\underset{}{\log(\hat{scrap})} = \underset{(.240)}{.409} + \underset{(.406)}{.057}\, grant$$

$$N = 54, \qquad R^2 = .0004$$

$$\underset{}{\log(\hat{scrap})} = \underset{(.089)}{.021} - \underset{(.147)}{.254}\, grant + \underset{(.044)}{.831}\, \log(scrap_{-1})$$

$$N = 54, \qquad R^2 = .873$$

Holzer, Block, Cheatham, and Knott (1993)