Omitted Variable Bias

November 15, 2023

Suppose the true relationship is

$$Y = a + \beta X + \gamma S + \varepsilon$$

But you estimate

$$Y = a^* + \beta^* X + \varepsilon^*$$

What is β^*?

$$\begin{aligned}
\beta^* &= \frac{Cov\left(Y, X\right)}{Var(X)} \\
&= \frac{Cov\left(a + \beta X + \gamma S + \varepsilon, X\right)}{Var(X)} \\
&= \frac{\beta Var\left(X\right) + \gamma Cov\left(S, X\right)}{Var(X)} \\
&= \beta + \gamma \frac{Cov\left(S, X\right)}{Var\left(X\right)}
\end{aligned}$$

So β^* is the true β, plus the product of the effect γ of the omitted variable times the coefficient from a regression of S on X.