Omitted Variable Bias Calculation

Full Model: TS ~ B0 + B1(CS) + B2(SES) + e1

Naïve Model: TS ~ b0 + b1(CS) + e2

Policy variable: Classroom Size

Omitted variable: Socio-Economic Status

URL <- "https://raw.githubusercontent.com/DS4PS/cpp-523-fall-2019/master/labs/class-size-seed-1234.csv"

dat <- read.csv( URL )

m.full <- lm( test ~ csize + ses, data=dat )

summary( m.full ) # Model 4

m.naive <- lm( test ~ csize, data=dat )

summary( m.naive) # Model 1

m.auxiliary <- lm( ses ~ csize, data=dat )

summary( m.auxiliary )

# lm(formula = ses ~ csize, data = dat)

# Coefficients:

# ----------------------------------

# Estimate Std. Error t value Pr(>|t|)

# (Intercept) 4.469458 0.009033 494.8 <2e-16 \*\*\*

# csize -0.094876 0.000326 -291.0 <2e-16 \*\*\*

# ----------------------------------

# b1 = B1 + bias

# b1 - B1 = bias

b1 <- -4.22

B1 <- -2.67

b1 - B1

# bias = a1\*B2

a1 <- -0.0949

B2 <- 16.34

a1\*B2



