BEM 106: Homework #3 Description

Goal: The goal of this homework is to help you understand omitted variable bias in a controlled setting. That is, this homework uses artificially generated data from a known conditional expectation function and asks you to estimate omitted variable bias.

Part 1

The (artificially created) model of the world is

$$y_1 = 1.3x_1 + 4.5x_2 + error$$

Question 1: What is the correlation between x_1 and x_2?

Question 2: You're thinking about estimating the following model: $y_1 = b_0 + b_1 * x_1 + error$

That is, you are omitting x_2 from your estimation. Before estimating, do you think your estimate of b_1 will be biased? Why or why not? If yes, will your estimate for b_1 be an overestimate or an underestimate?

Question 3: Estimate the model: $y_1 = b_0 + b_1 * x_1 + error$ using OLS with robust standard errors. Report the estimated b_0 and b_1. Was your hypothesis in Question 2 correct? Why or why not?

Part 2

The (artificially created) model of the world is now:

$$y_2 = 0.2x_1 + 0.1x_2 + 13x_3 + error$$

Question 4: What is the correlation between x_2 and x_3?

Question 5: You're thinking about estimating the model: $y_2 = b_0 + b_1 * x_1 + b_2 * x_2 + error$. That is, you are omitting x_3. Before estimating, do you think b_2 will be biased? Why or why not? If yes, will your estimate for b_2 be an overestimate or an underestimate?

Question 6: Estimate the model: $y_2 = b_0 + b_1 * x_1 + b_2 * x_2 + error$. Report the estimated b's. Was your hypothesis in Question 5 correct? Why or why not?

Part 3

The (artificially created) model of the world is now:

$$y_3 = 0.2x_1 + 0.1x_2 - 13x_3 + error$$

Question 7: What is the correlation between x_2 and x_3?

Question 8: You're thinking about estimating the model: $y_3 = b_0 + b_1*x_1 + b_2*x_2 + error$
That is, you are omitting x_3. Before estimating, do you think b_2 will be biased? Why or why not? If yes, will your estimate for b_2 be an overestimate or an underestimate?

Question 9: Estimate the model: $y_3 = b_0 + b_1*x_1 + b_2*x_2 + error$. Report the estimated b's.. Was your hypothesis in Question 8 correct? Why or why not?