Econometrics A
Practice Problems #2

1. Consider the following causal model of the effect of X on Y:

$$Y = \beta_0 + \beta_1 X + U \ .$$

Suppose $0 < P\{X = 1\} < 1$.

   (a) Does $E[U] = 0$? What about $E[XU]$? Explain briefly.

   (b) Let Z be an instrument for X.

      i. Define instrument exogeneity and instrument relevance.

      ii. Express β_1 in terms of features of the distribution of (Y, X, Z).

   (c) Let Z be a binary instrument for X. Show that

$$\beta_1 = \frac{E[Y|Z=1] - E[Y|Z=0]}{E[X|Z=1] - E[X|Z=0]} \ .$$

2. Let $X_1, ..., X_m$ be an i.i.d. sample from X. Let

$$W_n = \frac{1}{n^2} \sum_{i=1}^{n} X_i \ .$$

Let $\mu = E(X)$.

   (a) Suppose that $n > 1$. Show that W_n is an unbiased estimator of μ if and only if $\mu = 0$.

   (b) Show that W_n is a consistent estimator of μ if and only if $\mu = 0$.

   (c) What is $\text{Var}[W_n]$?

3. Consider the following model of the determinants of crime:

$$\log(crime) = \beta_0 + \beta_1 \log(pol) + \beta_2 \log(inc) + U \ ,$$

where

$$\begin{aligned} crime &= \text{number of crimes committed in the town per year} \\ pol &= \text{the number of police officers working in the town} \\ inc &= \text{the average family income in the town measured in dollars} \ . \end{aligned}$$

Suppose U is uncorrelated with $\log(pol)$ and $\log(inc)$ and that there is no perfect colinearity in $(1, \log(pol), \log(inc))$. Suppose further that the fourth moments of $\log(crime)$, $\log(pol)$ and $\log(inc)$ exist.

(a) Interpret U. Does $E[U] = 0$?

(b) Interpret β_1 and β_2.

(c) What signs do you expect for β_1 and β_2? Explain briefly.

(d) Suppose you have access to an i.i.d. sample of observations on *crime*, *pol*, and *inc_cents*, where

$$inc\_cents = inc \times 100 \ .$$

If you regressed $\log(crime)$ on $\log(pol)$ and $\log(inc\_cents)$, do you believe that your estimate of β_1 will be consistent? If not, do you think the limit in probability of $(\hat{\beta}_1)$ will be bigger or smaller than β_1? Explain briefly.

(e) Suppose you have access to an i.i.d. sample of observations on *crime* and *pol*, but not *inc*. If you regressed $\log(crime)$ on $\log(pol)$, omitting $\log(inc)$, do you believe that your estimate of β_1 will be consistent? If not, do you think the limit in probability of $(\hat{\beta}_1)$ will be bigger or smaller than β_1? Explain briefly.

4. A researcher runs an experiment to determine whether having smaller class sizes improves kindergarten school performance. To this end, kindergarten students are assigned at random to either a "regular" class with no aide, a "regular" class with an aide, and a "small" class. At the end of the year, students are given a standardized test. Based on this experiment, the researcher writes down the following model of the determinants of test scores:

$$testscore = \beta_0 + \beta_1 smallclass + \beta_2 regaide + U \ ,$$

where

$$\begin{aligned} testscore &= \text{test score on a standardized test} \\ smallclass &= \text{student assigned to a ``small'' class} \\ regaide &= \text{student assigned to a ``regular'' class with an aide .} \end{aligned}$$

The research has an i.i.d. sample of observations on *testscore*, *smallclass* and *regaide*. Suppose the fourth moments of these variable exist and that there is no perfect colinearity in $(1, smallclass, regaide)$. The researcher estimates the equation by OLS and find that

$$\begin{aligned} \hat{\beta}_0 &= 918.04, \quad \text{s.e.}(\hat{\beta}_0) = 1.63 \\ \hat{\beta}_1 &= 13.90, \quad \text{s.e.}(\hat{\beta}_1) = 2.45 \\ \hat{\beta}_2 &= .31, \quad \text{s.e.}(\hat{\beta}_2) = 2.27 \end{aligned}$$

The R^2 from the regression is .01.

(a) Given that the students were randomly assigned to the different types of classes, do you think it is plausible that U is uncorrelated with the regressors?

(b) Does the low R^2 imply that the model is incorrect?

(c) Interpret each coefficient.

(d) Do you think it would be a good idea to include an indicator variable for being in a "regular" class with no aide? Explain briefly.

(e) According to these estimates, is a "small" class without a teaching aide more effective or less effective than a "regular" class with a teaching aide?

(f) Suppose you wish to test the null hypothesis that a "small" class without a teaching aide is as effective as a "regular" class size with a teaching aide versus the alternative that a "small" class without a teaching aide is more effective than a "regular" class with a teaching aide.

    i. Formally state the null and alternative hypotheses.

    ii. Do you have enough information to perform your test? If not, describe what additional information you would need and how you would perform your test.