ECON 7910
Econometrics I
Philip Shaw
Problem Set 7
Due Date: December 14, 2023

Problem 1

Consider the following system of equations:

$$growth_i = \gamma_1 cored_i + \gamma_2 corjud_i + z_1\delta_1 + u_{i1} \tag{1}$$

$$cored_i = \gamma_3 growth_i + \gamma_4 corjud_i + z_2\delta_2 + u_{i2} \tag{2}$$

$$corjud_i = \gamma_5 growth_i + \gamma_6 cored_i + z_3\delta_3 + u_{i3} \tag{3}$$

where *growth* is the average growth rate for a country over the period 1980-2010, *cored* is the proportion of households that reporting paying a bribe at educational institutions, and *corjud* is the proportion of households that reporting paying a bribe at legal institutions. z_1, z_2, and z_3 contain control variables for each equation.

a. Assume that $E(z'u_g) = 0$ for $g = 1, 2, 3$ where $z = [z_1, z_2, z_3]$. Also assume that $E(cored'u_g) \neq 0$ for $g = 1, 3$, $E(corjud'u_g) \neq 0$ for $g = 1, 2$, and $E(growth'u_g) \neq 0$ for $g = 2, 3$. Under what conditions can you identify each equation? Describe the approach to estimating each equation using 2SLS.

b. Ignoring any potential endogeneity issues above, estimate Equation (1) via OLS.* Interpret how each type of corruption impacts expected average economic growth.

c. Now assume we observe an potential IV *edext* for *cored*. The variable *edext* measures the percentage of housholds that believe corruption in education is an extreme problem. In your opinion is this IV valid? Explain carefully.

*Use data file *corrupt.csv*.

d. Assuming $\gamma_2 = 0$, estimate Equation (1) via 2SLS using *edext* as an IV for *cored*. Is the IV sufficiently strong? How do you know?

e. Interpret the marginal effect of educational corruption on average economic growth. Is this marginal effect statistically different from zero?

f. Now test whether *cored* is endogenous in Equation (1). What do you conclude?

g. Suppose you think that the proposed IV above is not a valid IV. How else can you identify Equation (1) under the assumption that *cored* is still endogenous in the structural equation? Be specific.

h. Now implement the approach of Lewbel (2010) to estimate the model via 2SLS using generated IVs $(z - \bar{z})\hat{\epsilon}_2$. Are the instruments sufficiently strong?