ECOM30003/ECOM90003: Applied Microeconometric Modelling
Tutorial 6

Please read Chapter 15 of Wooldridge before attempting the following.

1. Consider a simple model to estimate the effect of personal computer (PC) ownership on college grade point average for graduating seniors at a large public university.

$$GPA = \beta_0 + \beta_1 PC + u \tag{1}$$

where PC is a binary variable indicating PC ownership.

   (a) Why might PC ownership be correlated with u?

   (b) Explain why PC is likely to be related to parents' annual income. Does this mean parental income is a good IV for PC? why or why not?

   (c) Suppose that, four years ago, the university gave grants to buy computers to roughly one-half of the incoming students, and the students who received grants were randomly chosen. Carefully explain if you would use this information to construct an instrumental variable for PC

2. We know that

$$\begin{aligned}
plim\hat{\beta}_{1,IV} =& \beta_1 + \frac{corr(z,u)}{corr(z,x)}.\frac{\sigma_u}{\sigma_x} \\
\rightarrow \text{asymptotic bias } \hat{\beta}_{1,IV} =& \frac{corr(z,u)}{corr(z,x)}.\frac{\sigma_u}{\sigma_x} \\
plim\hat{\beta}_{1,OLS} =& \beta_1 + corr(x,u).\frac{\sigma_u}{\sigma_x} \\
\rightarrow \text{asymptotic bias } \hat{\beta}_{1,OLS} =& corr(x,u).\frac{\sigma_u}{\sigma_x}
\end{aligned}$$

Assume $\sigma_u = \sigma_x$ so that the population variation in the error term is the same as it is in x. Suppose that the instrumental variable z is slightly correlated with u, $Corr(z,u) = 0.1$. Suppose also that z and x have a somewhat stronger correlation $Corr(z,x) = 0.2$

   (a) What is the asymptotic bias in the IV estimator?

   (b) How much correlation would have to exist between x and u before OLS has more asymptotic bias than 2SLS?

3. Use the data in WAGE2.dta for this exercise.

(a) In the lecture, when we use number of siblings, *sibs* as an instrument for *educ*, the IV estimate of the return to education is 0.122. To convince yourself that using *sibs* as an IV for *educ* is not the same as just plugging *sibs* in for *educ* and running an OLS regression. Run the regression of $log(wage)$ on *sibs* and explain your findings.

(b) The variable *brthord* is birth order. Explain why *educ* and *brthord* might be negatively correlated. Regress *educ* on *brthord* to determine whether there is a statistically significant negative correlation.

(c) Use *brthord* as an IV for *educ* in equation

$$log(wage) = \beta_0 + \beta_1 educ + u \tag{2}$$

Report and interpret the results.

(d) Now suppose that we include the number of siblings as an explanatory variable in the wage equation. This controls for family background to some extent:

$$log(wage) = \beta_0 + \beta_1 educ + \beta_2 sibs + u \tag{3}$$

Suppose that we want to use *brthord* as an IV for *educ* assuming that *sibs* is exogenous. The reduced form for *educ* is

$$educ = \pi_0 + \pi_1 sibs + \pi_2 brthord + v \tag{4}$$

State and where possible test the identification assumptions.

(e) Estimate the equation from part (d) using *brthord* as an IV for *educ* and (*sibs* its own IV). Comment on the standard errors for $\hat{\beta}_{educ}$ and $\hat{\beta}_{sibs}$.

(f) Using the fitted value from part (4) $\widehat{educ}$ compute the correlation between $\widehat{educ}$ and *sibs*. Use this result to explain your findings from part (e)