Econometrics II: Instrumental variables

March 29th, 2017

1. Are the OLS regressions estimated in the following examples appropriate? Why? Provide examples of instruments where needed.
   a. A researcher is trying to find out if putting criminals in jail reduces crime. He uses data on crime rate (*crimes per 100,000*) as dependent and incarceration rate (*prisoners per 100,000*), economic conditions and demographics as explanatory variables.
   b. A paper is trying to estimate the impact of employment during school on college student academic performance. The estimation is based on the regression of GPA (grade point average) on hours worked and some other control variables (gender, age, race, year in school, health status).
   c. We want to measure the impact of education on wages. Thus, we consider a regression of log-earnings on years of schooling, work experience and gender.

2. (Stock & Watson) For this exercise use *fertility.dta*. The data set contains information on married women aged 21-35 with two or more children. We are interested in the relationship between fertility and labor supply. (Provide relevant Stata output where appropriate.)
   a. Produce summary statistics and explain what it shows.
   b. Regress *weeksm1* on the *morekids*. Explain the results.
   c. Is the previous regression appropriate for estimating the casual effect of fertility (*morekids*) on labor supply (*weeksm1*)?
   d. Explain why *samesex* is a valid instrument for the IV regression of *weeksm1* on the *morekids*. Is it a weak instrument?
   e. Estimate the regression of *weeksm1* on *morekids* using *samesex* as an instrument (use `ivregress 2sls`). Compare the results with OLS estimates.
   f. Now include variables *agem1, black, hispan, and othrace*. Do the results change? Why?
   g. Estimate the same regression using `regress` and `predict` instead of `ivregress`. Compare the results.

3. (Wooldridge) Assume that $\sigma_x = \sigma_u$, 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 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?

   (Hint: Probability limit of IV estimator is: $plim\hat{\beta} = \beta + \frac{Corr(z,u)}{Corr(z,x)}\frac{\sigma_u}{\sigma_x}$ and of OLS estimator: $plim\hat{\beta} = \beta + Corr(x,u)\frac{\sigma_u}{\sigma_x}$.)