Part IIA, Paper 3, Supervision 5

A. IV

We are interested in the determinants of crime.

1. Download the dataset CRIME4.dta

2. See what the variables mean by typing "des"

3. Summarize the data using the command "summ"

4. Run an OLS regression with the command "reg crmrte polpc west central urban"

6. What do you infer from the row corresponding to "west"?

7. Interpret the positive and highly significant coefficient on polpc.

8. We intend to use "taxpc" as an IV for polpc. In terms of a simultaneous equations model, explain which coefficients should be zero or non-zero, in order for "taxpc" to be a valid IV for polpc. Be careful about what equation is being identified.

9. Run a 2SLS using "taxpc" as an IV. Compare with the OLS output.

10. If taxpc were a valid instrument, what does the 2SLS estimate tell us about the effect of police on crime? What does it tell us about the effect of crime on police employment?

11. Can you suggest a reason why taxpc may not be a valid instrument?

12. If taxpc is indeed a valid IV, how would you test the endogeniety of polpc in the OLS regression of 4. Perform this test and write down your conclusion.

B. Simultaneous Equations

We are interested in estimating how women's hours of work respond to the wages they get. Consider the simultaneous equations relating wages earned and hours worked for women in the labour force:

Hours=b0+b1*wage+b2*kids+u (1)

Wage=c0+c1*hours+c2*kids+v (2)

A. Where do you think these equations come from—i.e., what are the economic rationale for these equations?

B. Would an OLS of Hours on wage and kids give us consistent estimates of the causal effect of wage on hours of work? Please justify your answer.

C. Solve for hours and wage in terms of kids, u and v by solving the equations (1) and (2). Now consider equation (1) and check if any of the Gauss-Markov assumptions is necessarily violated for this equation.

D. We want to use "kids" as an IV to get rid of the simultaneity problem. How would you do that? Please state carefully what assumptions you need to make and what equation is being identified.