ECON 6511: Advanced Applied Econometrics
Homework 6
Due in class on February 28, 2018

1. (Wooldridge, Chapter 15, Problem 1) Use the data in WAGE2.dta for this exercise.

   (a) In class we considered using *sibs* as a instrument for *educ* (which would yield an IV estimate of the return to education of 0.122). To convince yourself that using *sibs* as an IV for *educ* is <u>not</u> 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 (*brthord* is one for a first-born child, two for a second-born child, and so on.) 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 the equation:

   $$\log(wage) = \beta_0 + \beta_1 educ + u.$$

   Report and interpret the results.

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

   $$\log(wage) = \beta_0 + \beta_1 educ + \beta_2 sibs + u.$$

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

   $$educ = \pi_0 + \pi_1 sibs + \pi_2 brthord + v$$

   State and test the identification assumption.

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

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

2. (Wooldridge, Chapter 15, Problem 2) The data in FERTIL2.dta include, for women in Botswana during 1988, information on number of children, years of education, age, and religious and economic status variables.

(a) Estimate the model

$$children = \beta_0 + \beta_1 educ + \beta_2 age + \beta_3 age^2 + u$$

by OLS, and interpre the estimates. In particular, holding age fixed, what is the estimated effect of another year of education on fertility? If 100 women receive another year of education, how many fewer children are they expected to have?

(b) The variable $frsthalf$ is a dummy variable equal to one if the woman was born during the first six months of the year. Assuming that $frsthalf$ is uncorrelated with the error term from part (a), show that $frsthalf$ is a reasonable IV candidate for $educ$. ($Hint$: You need to do a regression.)

(c) Estimate the model from part (a) using $frsthalf$ as an IV for $educ$. Compare the estimated effect of education with the OLS estimate from part (a).

3. (Wooldridge, Chapter 15, Problem 6) Use the data in MURDER.dta for this exercise. The variable $mrdrte$ is the muder rate, that is, the number of murders per 100,000 people. The variable $exec$ is the total number of prisoners executed for the current and prior two years; $unem$ is the state unemployment rate.

(a) How many states executed at least one prisoner in 1991, 1992, or 1993? Which state had the most executions?

(b) Using the two years 1990 and 1993, do a pooled regression of $mrdrte$ on $d93$, $exec$, and $unem$. What do you make of the coefficient on $exec$?

(c) Using the changes from 1990 to 1993 only (for a total of 51 observations), estimate the equation

$$\Delta mrdrte = \beta_0 + \beta_1 \Delta exec + \beta_2 \Delta unem + \Delta u$$

by OLS and report the results in the usual form. Now, does capital punishment appear to have a deterrent effect?

(d) The change in executions may be at least partly related to changes in the expected murder rate, so that $\Delta exec$ is correlated with Δu in part (c). It might be reasonable to assume that $\Delta exec_{-1}$ is uncorrelated with Δu. (After all, $\Delta exec_{-1}$ depends on executions that occurred three ore more years ago!) Regress $\Delta exec$ on $\Delta exec_{-1}$ to see if they are sufficiently correlated; interpret the coefficient on $\Delta exec_{-1}$.

(e) Reestimate the equation in part (c), using $\Delta exec_{-1}$ as an IV for $\Delta exec$. Assume that $\Delta unem$ is exogenous. How do your conclusions change from part (c)?

4. The purpose of this exercise is to compare the estimates and standard errors obtained by correctly using 2SLS with those obtained using inappropriate procedures. Use the data file WAGE2.dta.

   (a) Use a 2SLS routine to estimate the equation:

   $$\log(wage) = \beta_0 + \beta_1 educ + \beta_2 exper + \beta_3 tenure + \beta_4 black + u,$$

   where *sibs* is the IV for *educ*. Report the results in the usual form.

   (b) Now, manually carry out 2SLS. That is, first regress $educ_i$ on $sibs_i$, $exper_i$, $tenure_i$, and $black_i$ and obtain the fitted values, $\widehat{educ}_i$, $i = 1, \ldots, n$. Then, run the second stage regression $\log(wage_i)$ on $\widehat{educ}_i$, $exper_i$, $tenure_i$, and $black_i$, $i = 1, \ldots, n$. Verify that the $\hat{\beta}_j$ are identical to those obtained from part (a), but that the standard errors are somewhat different. The standard errors obtained from the second stage regression when manually carrying out 2SLS are generally inappropriate.