MRU Mastering Econometrics **M. Joshway**
Spring 2020

Problem Set 3

1. Most private schools in the US are run by religious organizations. Economists and educators have long argued over the relative merits of parochial schools and traditional public schools. This motivates Evans and Schwab (1995; https://academic.oup.com/qje/article/110/4/941/1870567) to estimate the effects of attending a Catholic high school on the probability of high school graduation and college attendance. Let $COLL_i$ indicate college attendance and let CHS_i indicate Catholic high school attendance. A linear probability model for Catholic school effects on college-going is

$$COLL_i = \beta_0 + \beta_1 CHS_i + \beta_2 X_i + u_i, \quad (1)$$

where the controls, X_i, include variables like gender, race, family income, and parental education. Evans and Schwab compute 2SLS estimates of this model using a dummy for being Catholic as an instrument for CHS_i. Call the Catholic instrument Z_i.

   (a) Use the notation above to describe the first and second stage equations that generate the 2SLS estimates reported in Tables VI and VII of the Evans and Schwab paper. That is, write out the equations, explain which variables appear in each and note the roles they play.

   (b) Under what assumptions does this 2SLS procedure capture the causal effects of Catholic high school attendance on college-going?

   (c) Many American marriages mix religious affiliations (for example, a Jewish man marries a non-Jewish woman).

      i. Explain how you might exploit information on the religious affiliation of husbands and wives to come up with a new and perhaps improved instrument for a child's Catholic school attendance. What assumptions does your IV proposal require? (Hint: What do Angrist and Evans (1998) samesex instruments control for? Why is this helpful?)

      ii. Many marriages dissolve through divorce, separation, or the death of a spouse. How might information on marital dissolution be combined with info on parents' religion to produce yet another IV strategy for the effects of Catholic school attendance?

2. This question uses the Angrist & Krueger (1991) data.

   (a) Replicate MM Table 6.5, which shows OLS and IV estimates of the returns to schooling using quarter-of-birth instruments. Report your results in a table similar to Table 6.5.

   (b) Compute manual 2SLS estimates for one of these models and compare your results with estimates from Stata's "ivregress 2sls" command. How and why do these differ?

3. IV methods remove OVB and mitigate attenuation bias from measurement error. These are the two scenarios you're most likely to see IV used today. But IV was born as a means of estimating supply and demand elasticities in simultaneous equations models (SEMs) .

   (a) Learn the theory behind this by studying these notes and the Angrist, Graddy, and Imbens (2000) paper on the Fulton Fish Market. Focus on Section 5 of the paper, which uses instrumental variables derived from offshore weather conditions to construct instruments for the price of fish in a demand equation. Using SEM theory, explain why weather instruments identify demand elasticities rather than supply elasticities.

   (b) Use fish.dta to replicate the estimates of demand elasticities reported at the end of the lecture notes.[1] Explain your procedure: what is the unit of observation? What are the endogenous variables, instrumental variables, and exogenous covariates? Identify first-stage, reduced form, and 2SLS estimates in your output. How are these related? How and why do 2SLS and OLS estimates differ?

[1] For this, it will help to know that stormy3=(speed3>18)*(wave3>4.5); mixed3=(1-stormy3)*(speed3>15)*(wave3>3); stormy2=(speed2>12)*(wave2>5.5); mixed2=(1-stormy2)*(speed2>10)*(wave2>3).

(c) Use Stata's `reshape` command to change the data structure from a time series of prices and quantities for Asians and whites to panel data that stacks the ethnic groups in a format where ethnicity is identified by a dummy rather than a variable name. Use the reshaped data to compute separate demand elasticities for Asians and Whites and test whether they differ. Discuss your results in light of the fact that Asian buyers appear to pay less than whites for fish.