2017 STATA ECONOMETRICS WINTER SCHOOL
Porto, 2017

IV Exercise

We can use the data from Card (1995ab) to estimate the impact of education on wages, where nearness to a college is posited as a source of exogenous variation in educational attainment. For this exercise use card.dta, available at http://fmwww.bc.edu/ec-p/data/wooldridge/card.dta.

[in stata type: use http://fmwww.bc.edu/ec-p/data/wooldridge/card.dta]

1. 2SLS by hand

Regress educ against exper nearc4. Generate predicted value – call this educ_hat. Now regress lwage on educ_hat and exper. Compare the results with the ones from command ivregress.

Note: The third two-stage version of the IV strategy, which applies in the case of one endogenous variable and one excluded instrument, is sometimes called the Wald estimator. Try the following commands:

reg educ nearc4 exper
loc p=_b[nearc4]
reg lwage nearc4 exper
loc g=_b[nearc4]
di 'g'/'p'

2 Example

Try the following set of commands:

loc x "exper* smsa* south mar black reg662-reg669"
reg lw educ 'x'
ivreg2 lw 'x'(educ=nearc2 nearc4), first endog(educ)

to find the return to an additional year of education is about 7% using ordinary regression or 16% using IV methods. The Sargan statistic fails to reject that excluded instruments are valid, the test of endogeneity is marginally significant (giving different results at the 95% and 90% levels), and the Anderson-Rubin and Stock-Wright tests of identification strongly reject that the model is underidentified.

Discuss the results found for the weak identification tests. Discuss the differences found on the return to education when you used OLS or IV.