HOMEWORK 4

PAUL SCHRIMPF

DUE: FRIDAY, NOVEMBER 4TH

UNIVERSITY OF BRITISH COLUMBIA

ECONOMICS 628: TOPICS IN ECONOMETRICS

This assignment asks you to recreate some of the analysis of Carneiro, Heckman, and Vytlacil (2010)[1]. We will be using slightly different data and slightly different estimators, but we will estimate some of the same things.

The dataset for this assignment is hw4-nlsy.csv. It is an extract from the NLSY79. It consists of white men who are not missing any of the variables that will be used in our analysis. The variables are:

(1) `yob`: year of birth
(2) `mhgc_mi`: mother's highest grade of schooling completed
(3) `numsibs`: number of siblings
(4) `urban14`: whether lived in an urban area at age 14
(5) `tuit4`: tuition at nearest four year college
(6) `locCol`: whether there is a college in the county of residence at age 14
(7) `afqt`: armed forces qualifying test score, adjusted for education at time of test and standardized
(8) `college`: whether ever attended college
(9) `avgEarn89_00`: average earnings from 1989-2000 in year 2000 dollars

This data is not exactly the same as in Carneiro, Heckman, and Vytlacil (2010). The earnings measure is different, and the local college variable also seems to be messed up. It also included a local tuition measure, which we will use as an instrument. There is some code to get you started with the first few parts of the assignment on the course web page.

Part 1. Estimate the return to college using OLS and IV. Regress log earnings on college. Include AFQT as a control, and use tuition as an instrument. Report the coefficient on college and the standard error. Divide by four to make it a return per year of college, since that's what people usually report.

Part 2. Flexibly estimate the propensity score, the probability of college as a function of x and z. As in part 1 just use AFQT as x and tuition as z. You may use either series or kernel regression. Plot you estimate of the propensity score as a function of x and z. Also plot the support of the estimated propensity score as a function of x. Vary the number of series terms or bandwidth, and report how the results change.

Part 3. Flexibly estimate the marginal treatment effect conditional on AFQT. Plot the estimated marginal treatment effect as a function of u and AFQT. Also plot the marginal treatment effect as a function of u at the average AFQT. Again, investigate the sensitivity of the results to the number of series terms or bandwidth.

[1]The 2009 version of this paper is listed on the syllabus, but there is a newer 2010 version.

Part 4. Compute uniform confidence bands for your plot of the marginal treatment effect as a function of u at the average AFQT. Draw the plot with the uniform confidence bands included.

Part 5. Specify an alternate policy and estimate the policy relevant treatment effect. The policy could be something like reduce tuition by some percent, or it could be something more artificial like increase $\pi(z, x)$ in some way. Compute the standard error of your estimate.

REFERENCES

CARNEIRO, P., J. J. HECKMAN, AND E. J. VYTLACIL (2010): "Estimating Marginal Returns to Education," Working Paper 16474, National Bureau of Economic Research.