

ECON 211B
Winter 2017

Econometrics
Homework 1

The primary objective of this course is to learn how to use various econometric methods to get causal estimates. One of the secondary objectives of this course is to learn how to present and describe your analysis clearly. To this end please type up the homework in LaTeX. If you do not already have it set up on your computer you can download it free online. I like TeXworks but there are many other good LaTeX editors (a popular alternative is LyX which has a more user friendly interface). An important part of communicating clearly is creating readable tables and figures so please make sure all tables and figures have descriptive titles, column headings and detailed footnotes. They should look like the tables and figures you see in the journal articles you are reading as part of the course. Your goal is to create a table or figure clear enough that a reader should be able to interpret it correctly without having to refer to any other source. If you are using Stata the `outreg2` procedure (ado file) will automatically output tables in LaTeX which you can then enrich with footnotes, titles and column headings. Most other packages also have a simple way of outputting regressions to LaTeX. Please also keep a program file that has all your code with comments included. Good coding and documentation practices are a critical part of doing empirical work. Please submit your code file along with your home work.

1. X and Y are discrete random variables whose expectations exist. Present a proof for the law of iterated expectation $E[Y] = E[E[Y|X]]$.
2. Create a figure with hourly wage plotted against educational attainment for men in the US between the ages of 30 and 40 in January of 2014. There are numerous sites you can download the Current Population Survey from. One source among many is <http://ceprdata.org/cps-uniform-data-extracts/cps-outgoing-rotation-group/cps-org-data/>
3. Estimate the CEF using OLS and plot the fitted values in the figure created in part 2.
4. Why is the CEF linear in this case and show that OLS will generate a consistent estimate of the CEF?
5. We wish to estimate the causal effect on earnings of college attendance relative to only completing high school. Please use the framework of the Rubin Causal Model to assess if your comparison from question 4 gives you a causal estimate. What is D_i what is $E[Y_i(0)|D_i = 1]$ etc.