

Problem Set 1  
Econometrics 410  
Prof. Taber  
Due: Thurs Feb. 3

**Problem 1.** Look for some magazine, web page, or newspaper article that discusses an empirical study for which causation/correlation is an issue. Explain what the true causal effect of interest is. That is explain what question we would ideally like to answer. Explain other possibilities that may account for the correlation that is found.

**Problem 2.** Please do the following exercise using data from stata. Turn in the stata log file. This is a case for which if you work with a classmate on the stata work, you may turn in copies of the same stata log file.

- a) Load the data set CPS78\_85.dta
- b) Summarize the data
- c) This data set has “log wage” rather than wage. Construct the wage using “gen wage=exp(lwage)”
- d) Produce an estimate of the conditional expectation of wages conditional on year
- e) produce an estimate of the conditional expectations of wages across all years by gender and marital status. That is the expected age for married men, married women, unmarried men, and unmarried women

**Problem 3.** This is another stata exercise. The directions for problem 2 hold.

- a) load the data set wine
- b) list the country name, the amount of wine consumed per capita, and the per capita death rate
- c) Produce a scatter plot of alcohol and deaths
- d) What is the correlation coefficient and the covariance between the two? (hint: command correlate)
- f) Run a regression of deaths on alcohol and plot the regression line within the scatter plot
- g) Using the summary statistics from the data and the formula for the regression coefficients, derive the regression coefficients and check that they are very close to the estimates from stata.
- h) Do you think this is causal? Give an alternative story for the results. That is, is this likely Case 1, 2 or 3. Explain why.