Homework #5

Economics 113

Introduction to Econometrics

Professor Spearot

Due Friday, March 13th

Problem 1

Using the Wage2.dta dataset from the website, we wish to compare a few different wage regressions. Use a 5% significance level for all regressions, and conduct hypothesis tests where necessary. Suppose that we start with the following model, hereafter referred to as the basic model:

$$\log(wage) = \beta_0 + \beta_1 educ + \beta_2 exper + \beta_3 tenure + u$$

a. Is the basic model preferred to the following model? Why or why not?

$$\log(wage) = \beta_0 + \beta_1 educ + +\beta_2 exper + \beta_3 tenure + \beta_4 sibs + \beta_5 brthord + u$$

b. Now compare the basic model with the following model:

$$wage = \beta_0 + \beta_1 educ + +\beta_2 exper + \beta_3 tenure + \beta_4 sibs + \beta_5 brthord + u$$

Which is preferred?

c. Does the basic model tell us anything about the log(wage)? Why or why not?

d. In the basic model, do the returns to education depend on the level of experience? Choose a specification to test this hypothesis, regress it, and test the hypothesis at the 95% level.

Problem 2

Using the Bwght.dta dataset from the website, we wish to predict the probability of smoking by expectant mothers. That is, we wish to test the following:

$$Smoke = \beta_0 + \beta_1 meduc + \beta_2 feduc + \beta_3 cigprice + u$$

where *Smoke* equals 1 if the mother smokes, and zero otherwise.

a. Generate the variable, Smoke.

b. Do cigarette prices affect the probability of smoking? Test this hypothesis, and interpret the parameter of interest.

c. What is a potential problem in running this type of regression? Do we have this problem in this particular case?