PA 9950 Quantitative Methods

Problem Set #1

This homework is to accompany program #1 which you should all receive via email.

Answer all questions fully, please be sure to turn in a log file with your assignment. You should write up your results as though you are providing a reply to a reviewer.

**Note: the bootstrap and error generation commands rely on a random number generator that uses a “seed” to set the first random number.**

Question for Section 1:

1. Examine the means and the medians for wages and education. Why is the median less than the mean? What does this imply about the distribution of wages?
2. Examine the average wage by education and gender table. Why is the average wage for females with 18 years of education higher than that of males with 18 years of education? Why does the average wage for men fall going from 17 to 18 years of education? (Be sure to examine the data!)
3. Interpret the coefficient estimate on education for each regression for each regression, calculate the predicted wage for a person with 1 year of education. Comment on the result. Do the same for a person with 10 years of education. Comment.

Questions for Section 2:

1. Using the STATA help command [help bootstrap] discuss what the bootstrap "summarize wage" wagebs=r(mean), saving($datadir/bs.dta, replace) command is doing.
2. Discuss the sampling variation.

Again, note that each time you run this portion of the program you will get a new sample and new set of means (Why?)

Questions for Section 3

1. Discuss how wage1, wage2 and wage3 differ from wage.
2. How does the estimate of wage1= a+b educ compare to the baseline model (wage=a+b educ)
3. What happens when we change the dependent variable to wage2? Discuss why this changes.
4. Discuss the effects of using wage3 as our dependent variable.
5. Interpret the differences in the betas for each model.
6. Discuss the effect on the R-squared.