Week 4 Discussion Section Questions

Ryan Martin

January 30, 2018

Table of Contents

Question 1

Does the return to education differ by race and gender? For this exercise, use the file cps4.dat (This is a large file with 4,838 observations. If you are using the student version of Stata software, you can use the smaller file cps4_small.dat. If you are using R, the size shouldn't be a problem. R runs into problems around the 10 million entries point, whereupon you may need to do some fancier technique and use some packages, but can still get the job done without paying.) In this exercise you will extract subsamples of observations consisting of (i) all males, (ii) all females (iii) all whites, (iv) all blacks, (v) white males, (vii) black males and (vii) black females.

4.15 ii

a For each sample partition, obtain the summary statistics of *WAGE*

b A variable's *coefficient of variation* is 100 times the ratio of its sample standard deviation to its sample mean. That is, for a variable y it is

$$CV(y) := 100 \times \frac{s_y}{\bar{y}}$$

where s_y is y's standard deviation and $\bar{y}$ is y's average. What is the coefficient of variation for *WAGE* within each sample partition?

c For each sample partition, estimate hte log-linear model

$$\log(WAGE) = \beta_1 + \beta_2 EDUC + e$$

What is the approximate percentage return to another year of education for each group?

d Does the model fit the data equally well for each sample partition?

e For each sample partition, test the null hypothesis that the rate of return to education is 10% against the alternative that it is not, using a two-tail test at the 5% level of significance.