Applied Econometrics II
Dept of Economics, Carnegie Mellon University
73-360, Fall 2000

Solution #1

See the program hwk1.sas for the method to select the data. All references to the output below refer to hwk1.lst. Both are available on the course web site.

1. **First, please calculate the average salary for people with 8,10,12,14,16 years of schooling respectively. Do you see any pattern?**

   The results are on page 2 of the output. They are summarized in the following table:

Years Schooling	Avg Salary
8	N/A
10	$9,648
12	15,437
14	16,632
16	27,229

   It looks like people with more schooling earn more.

2. **Now, let's consider the regression model:**

   $Y_i = \beta_1 + \beta_2 yrschool + u_i$

   Based on your answer to question 1, do you think β_2 is likely to be positive or negative?

   I think β_2 is likely to be positive. When β_2 is positive, that means that higher values of yrschool go with higher values of salary, and that is what I found in question 1.

3. **Please run the regression in question 2. Interpret β_2 and relate your results back to the prior questions.**

   Please see the output, page 3. β_2 is the amount by which salary will increase when years of schooling increases by 1 — it is the "value" of a year of schooling.

   Looking at page 3, the best (by G-M Thm) estimate of β_2 is positive. We expected the true value of β_2 to be positive, so the fact that our estimate of it is positive is reassuring.

4. **Run the following regression:**

   $Y_i = \beta_1 + \beta_2 yrschool + \beta_3 age + u_i$

   Interpret β_2. Contrast your results here with those in question 3 — contrast both the interpretation of β_2 and the meaning of the value of $\widehat{\beta}_2$.

   Please see the output, page 4. In the previous question, β_2 was the amount by which salary increases when yrschool increases by 1. In the current question, β_2 is the amount by which salary increases when yrschool increases by 1 and age does not change. The difference in these two estimates is that, in the previous question, when we are comparing people with different levels of schooling, they might *also* differ in age. In this question, when we are comparing people with different levels of schooling, we are holding constant age — so the differences here cannot be due to differences in age, since that is held constant (at least, as long as the model here is true).

   As far as the estimators go, in both questions $\widehat{\beta}_{2,OLS}$ is the best linear unbiased estimator of β_2 for its model.

   As far as the estimates go, in this model (as opposed to the model of the previous question) it looks like our best estimate of the effect of schooling on salary is higher. Since our estimate is also that the effect of age on salary is positive, it seems likely that the more schooled women in our sample are the younger women, and that this age difference was hiding some of the relationship between schooling and salary above.