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

Solution #2

1. **You think that wages are determined by age, schooling, and hours worked. Do not run any regression to answer this question!**

   (a) **Please write down a model appropriate to analyze these data in light of these ideas.**

   $$wage = \beta_1 + \beta_2 age + \beta_3 yrschool + \beta_4 hours + u$$

   (b) **Tell me what each parameter means and what sign (±) you expect for it.**

Parameter	Interpretation	Expect Sign
β_1	intercept: average wage for a newborn with no schooling working 0 hours	?
β_2	Amount by which wages increase when person ages by 1 year, without changing their years of schooling or hours.	+
β_3	Amount by which wages increase when person obtains one more year of schooling without changing age or hours.	+
β_4	Amount by which wages increase when person increases their hours by 1 without aging or increasing their years of schooling	+

   (c) **How would you test the theory that hours do not affect wages if you knew the parameters?**

   I would look at β_4. If $\beta_4 = 0$ the theory is true; if not, it is false.

   (d) **How would you test the theory that a one year increase in age gives the same increase to wages that a 4 year increase in hours does?**

   I would look at β_2 and $4\beta_4$. If $\beta_2 = 4\beta_4$, then the theory is true. If not, then the theory is false.

2. **Please estimate the model you constructed in Question 1.**

   (a) **Using your best estimates, does it look like you were right about the signs of the variables?**

   My best estimates come from running an OLS regression (by the G-M Thm). Those estimates appear on page 2 of the output. It looks like I was right about all three I guessed at. Maybe I should have guessed negative for β_1 since it seems likely that a newborn would generate negative value for a potential employer.

   (b) **Please test to see if the each of coefficients in your model is different from zero.**

   I'll choose the 1% significance level. I can just look at the P-value column of the table on page 2. Since all the P-values are less than 0.01, except for β_2, I can reject each of the hypotheses that the β are zero at the 1% level, except for β_2. So, I am at least 99% confident that each of hours and years of schooling affect wages, but I am not 99% sure that age does.

   (c) **Please make confidence intervals for each parameter.**

   I'll make 95% confidence intervals for each parameter. Looking at the t-table at the 5% level with $N - K = 364 - 4 = 360$ degrees of freedom, we get about 1.96. All results come from page 2 of the output. So, the 95% confidence intervals are:

Parameter	Estimate	$\pm$	C.I.
β_1	-25.99	15.93	(10.06,41.92)
β_2	0.12	0.18	(-0.06, 0.30)
β_3	0.35	0.22	(0.13, 0.57)
β_4	1.72	0.98	(0.74, 2.70)

   (d) **In light of the results from Questions 2a, 2b, and 2c what can you say about whether or not you were right in your answer to Question 1b.**

   Well, since the confidence intervals for β_3 and β_4 do not overlap 0 and the estimates are positive, I am at least 95% sure I was right about those. As for β_2, it looks like I was right, since the estimate is positive, but I can't rule out (at 95% confidence) the possibility that $\beta_2 < 0$.

   (e) **Please test the theory described in 1d.**

   The answer here could be a t-test (in which case you would need to use the /covb option and calculate the variance of $\hat{\beta}_{2,OLS} - 4\hat{\beta}_{4,OLS}$) or an F-test. The F-test is easier, so I'll do that. The UR model is on page 2 and the R model is on page 3. We are testing $H_0 : \beta_2 = 4\beta_4$.

$$
\begin{aligned}
F - stat &= \frac{(SSE_R - SSE_{UR})/q}{SSE_{UR}/(N-K)} \\
&= \frac{(98786 - 98749)/1}{98749/360} \\
&= 0.13
\end{aligned}
$$

Let's choose the 5% significance level and look at the $F_{1,360}$ table. We get a critical value of about 3.89.

So, we accept the H_0 and conclude that there is not enough information in these data to be 95% sure that $\beta_2 \neq 4\beta_4$.