ECO375 Tutorial 3
Wooldridge: Chapter 4 and Appendix E

Matt Tudball

University of Toronto St. George

September 29, 2017

ECO375 Tutorial 3

Welcome back!

Today's coverage:

- Chapter 4, #2 (in slides)
- Chapter 4, #C2 (in slides)
- Chapter 4, #C8 (in slides)
- Appendix E, #3 (on the board)

Consider an equation to explain salaries of CEOs in terms of annual firm sales, return on equity (*roe*, in percentage form) and return on firm's stock (*ros*, in percentage form):

$$\log(salary) = \beta_0 + \beta_1 \log(sales) + \beta_2 roe + \beta_3 ros + u$$

i) In terms of the model parameters, state the null hypothesis that, after controlling for *sales* and *roe*, *ros* has no effect on CEO salary. State the alternative that better stock market performance increases CEO's salary.

Chapter 4, #2

Consider an equation to explain salaries of CEOs in terms of annual firm sales, return on equity (*roe*, in percentage form) and return on firm's stock (*ros*, in percentage form):

$$\log(salary) = \beta_0 + \beta_1 \log(sales) + \beta_2 roe + \beta_3 ros + u$$

i) In terms of the model parameters, state the null hypothesis that, after controlling for *sales* and *roe*, *ros* has no effect on CEO salary. State the alternative that better stock market performance increases CEO's salary.

H_0: $\beta_3 = \mathbf{0}$
H_1: $\beta_3 > \mathbf{0}$

ii) Using the data in CEOSAL1.dta, the following equation was obtained by OLS:

$$\widehat{\log(salary)} = \underset{(0.32)}{4.32} + \underset{(0.035)}{0.280} \log(sales) + \underset{(0.0041)}{0.0174}\, roe + \underset{(0.00054)}{0.00024}\, ros$$

$$n = 209, \; R^2 = 0.283$$

By what percentage is *salary* predicted to increase if *ros* increases by 50 points? Does *ros* have a practically large effect on *salary*?

ii) Using the data in CEOSAL1.dta, the following equation was obtained by OLS:

$$\widehat{\log(salary)} = \underset{(0.32)}{4.32} + \underset{(0.035)}{0.280} \log(sales) + \underset{(0.0041)}{0.0174}\ roe + \underset{(0.00054)}{0.00024}\ ros$$
$$n = 209,\ R^2 = 0.283$$

By what percentage is *salary* predicted to increase if *ros* increases by 50 points? Does *ros* have a practically large effect on *salary*?

The proportionate effect on $\widehat{\log(salary)}$ is $0.00024 \cdot 50 = 0.012$. Since this is a log-linear regression we multiply this by 100 in order to obtain the percentage change, which is 1.2%. Therefore, a 50-point ceteris paribus increase in *ros* is predicted to increase salary by only 1.2%. Practically speaking, this is a very small effect for such a large change in *ros*.

iii) Test the null hypothesis that *ros* has no effect on salary against the alternative that *ros* has a positive effect. Carry out the test at the 10% significance level.

Chapter 4, #2

iii) Test the null hypothesis that *ros* has no effect on salary against the alternative that *ros* has a positive effect. Carry out the test at the 10% significance level.

We first need to obtain the 10% critical value for a one-tailed test in Table G.2. Since n is relatively large we will use the approximation at df = ∞. This comes out to 1.282. We now need to calculate the t-statistic for this test, which is $t = (\hat{\beta}_3 - 0)/(se(\hat{\beta}_3)) = 0.00024/0.00054 = 0.44 < 1.282$. Thus, we fail to reject the null hypothesis.

iii) Would you include *ros* in a final mode explaining CEO compensation in terms of firm performance? Explain.

Chapter 4, #2

iii) Would you include *ros* in a final mode explaining CEO compensation in terms of firm performance? Explain.

Based on this sample, the estimated *ros* coefficient appears to be different from zero only because of sampling variation. However, if *ros* is very correlated with the other explanatory variables then omitting it from the regression may still bias the estimates (recall equation 3.46). We can see that the other explanatory variables are still very significant even with *ros* included in the regression, which provides evidence against this. It is ultimately a judgement call.

Chapter 4, #C2

Use the data in LAWSCH85.dta for this exercise.

i) Using the same model as Chapter 3 #4, state and test the null hypothesis that the rank of law schools has no ceteris paribus effect on median starting salary.

$$\log(salary) = \beta_0 + \beta_1 LSAT + \beta_2 GPA + \beta_3 \log(libvol) + \beta_4 \log(cost) + \beta_5 rank + u$$

We estimate that $\hat{\beta}_5$ = -0.0033 and $se(\hat{\beta}_5)$ = 0.00035, so t = -0.0033/0.00035 = -9.429. Since this is a two-sided test at 95% significance and df $\approx$ 120 we can use a critical value of 1.98. Since $|-9.429| > 1.98$ we reject the null hypothesis that $\beta_5 = 0$.

ii) Are features of the incoming class of students - namely, LSAT and GPA - individually or jointly significant for explaining *salary*? (Be sure to account for missing data on *LSAT* and *GPA*).

The t-statistic on *LSAT* is 1.17 so it is insignificant at the 95% level. The t-statistic on *GPA* is 2.75 so it is significant at the 95% level. To test joint significance, we note that R^2_{ur} from the unrestricted regression is 0.8417 and R^2_r from the restricted regression which omits *LSAT* and *GPA* is 0.8174 (with missing values from *GPA* and *LSAT* dropped). The F-statistic is $F = \frac{(R^2_{ur} - R^2_r)/q}{(1-R^2_{ur})/df_{ur}} = \frac{(0.8417-0.8174)/2}{(1-0.8417)/130} \approx 9.95$. **So *LSAT* and *GPA* are jointly very significant.**

iii) Test whether the size of the entering class (*clsize*) or the size of the faculty (*faculty*) needs to be added to this equation; carry out a single test. (Be careful to account for missing data on *clsize* and *faculty*).

We follow a very similar to the above procedure. We run the unrestricted regression first including *clsize* and *faculty*. We can save ourselves some time, however, by typing the command

```
test clsize faculty.
```

after the unrestricted regression. Stata will run the F-test for us, returning an F-statistic of 0.95 and a p-value of 0.3902 at the default 95% significance level. This indicates that *clsize* and *faculty* are not jointly significant.

iv) What factors might influence the rank of the law school that are not included in the salary regression?.

Chapter 4, #C2

iv) What factors might influence the rank of the law school that are not included in the salary regression?.

Many possible answers here. I thought of student-faculty ratio, faculty publication record, academic reputation among other academics.

The dataset 401KSUBS.dta contains information on net financial wealth (*nettfa*), age of the survey respondent (*age*), annual family income (*inc*), family size (*fsize*) and participation in certain pension plans for people in the United States. The wealth and income variables are both recorded in thousands of dollars. For this question, use only the data for single-person households (so $fsize = 1$)

i) How many single-person households are there in the dataset?

I used the command `keep if fsize == 1`.
There are 2,017 observations left.

ii) Use OLS to estimate the model:

$$nettfa = \beta_0 + \beta_1 inc + \beta_2 age + u$$

and report the results in the usual format. Be sure to use only the single-person households in the sample. Interpret the slope coefficients. Are there any surprises in the slope estimates?

We interpret $\hat{\beta}_1$ as a \$1 thousand dollar increase in income corresponds to a \$799 increase in net financial wealth. We interpret $\hat{\beta}_2$ as a 1 year increase in age corresponds to a \$842 increase in net financial wealth. Nothing too surprising here.

iii) Does the intercept from the regression in part (ii) have any interesting meaning? Explain.

$\hat{\beta}_0$ = -43.04. This is an individual's net financial wealth when their income is $0 and their age is 0, so this is the net financial wealth of newborn babies which we are unlikely to be interested in...

Chapter 4, #C8

iv) Find the p-value for the test $\mathrm{H}_0 : \beta_2 = 1$ against $\mathrm{H}_1 : \beta_2 < 1$. Do you reject H_0 at the 1% significance level?

The t-statistic is (0.843-1)/0.092 $\approx$ -1.71. Against the one-sided alternative $\mathrm{H}_1 : \beta_2 < 1$ the p-value is $\mathrm{P}(T < -1.71) \approx 0.044$ which we can find from Table G.1. Therefore we can reject the null hypothesis at the 5% significance level but not at the 1% significance level.

v) If you do a simple regression of *nettfa* on *inc*, is the estimated coefficient on *inc* much different from the estimate in part (ii)? Why or why not?

We get an estimate of $\hat{\beta}_1 = 0.821$ which is not very different from the estimate of 0.799 in the last regression. Since this is an omitted variable question we need to know the correlated between age and income, which we find to be only 0.039. This explains why the coefficient does not change much.