Econometrics 1.3

M2 PPD-APE

Exercise 1 Consider the multiple regression model with 3 independent variables, under the classical linear model assumptions MLR.1 through MLR.6:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_2 x_3 + u$$

You would like to test the null hypothesis H_0: $\beta_1 - 3\beta_2 = 1$.

1. Let $\widehat{\beta}_1$ and $\widehat{\beta}_2$ denote the OLS estimators of β_1 and β_2. Find $Var(\widehat{\beta}_1 - 3\widehat{\beta}_2)$ in terms of the variances of $\widehat{\beta}_1$ and $\widehat{\beta}_2$ and the covariance between them. What is the standard error of $\widehat{\beta}_1 - 3\widehat{\beta}_2$?
2. Write the t statistic for testing H_0
3. Define $\theta_1 = \beta_1 - 3\beta_2$ and $\widehat{\theta}_1 = \widehat{\beta}_1 - 3\widehat{\beta}_2$. Write a regression equation involving β_0, θ_1, β_2, and β_3 that allows you to directly obtain $\widehat{\theta}_1$ and its standard error.

Exercise 2 In exercise sheet #2, we estimated the equation

$$\widehat{sleep} = \underset{(112.28)}{3638.25} - \underset{(.017)}{.148}\, totwrk - \underset{(5.88)}{11.13} educ + \underset{(1.45)}{2.20}\, age$$

with $n = 706$ and $R^2 = .113$, where we now report standard errors along with the estimates.

1. Is either educ or age individually significant at the 5% level against a two-sided alternative?
2. Dropping *educ* and *age* from the equation gives

   $$\widehat{sleep} = \underset{(38.91)}{3586.38} - \underset{(.017)}{.151}\, totwrk$$

   with $n = 706$ and $R^2 = .103$. Are *educ* and *age* jointly significant in the original equation at the 5% level? Justify your answer.
3. Does including *educ* and *age* in the model greatly affect the estimated tradeoff between sleeping and working?
4. Suppose that the sleep equation contains heteroskedasticity. What does this mean about the tests computed in questions 1 and 2?

Exercise 3 Download Jacoby, H. (1993), "Shadow wages and Peasant Family Labour Supply: an Econometric Application to the Peruvian Sierra" *The Review of Economic Studies* 60 (no.4) Oct 1993: 903-921 from the course's website. The exercise bears on section IV (Estimation of the agricultural technology) and section V.A (Labour supply estimation) of the paper and does not necessitate to read the whole paper. You can neglect all the results and discussions related to the IV (instrumental variables). Note: a shadow wage is the marginal productivity of labour on land.

1. The paper starts by estimating a Cobb-Douglas function. In this specification, explain what the estimated coefficients are in first column of Table II.

2. What is the ratio of marginal productivity of a teenager and of a man in a Cobb-Douglas function? Give its estimate based on the 1st column of Table II and taken at mean values (Table I).

3. Estimated marginal product for male labour is quite similar in the 2 following specifications: Cobb-Douglas estimated by OLS (column 1) and simplified Translog also estimated by OLS (column 4 of Table II). However, their standard errors are quite different. How do you explain that? can you confirm this explanation by looking at another statistics?

4. The economic theory shows that, individuals taking optimal decisions, they should equate their wage on the labour market and their marginal productivity of labour on the farm. Explain the regression of interest (eq. 11) and why Jacoby wishes to test the hypothesis $(a, b) = (0, 1)$.

5. Could we run a similar regression but with marginal product as the explanatory variable and wage as the dependent variable to test the same hypothesis?

6. Define the restricted and the unrestricted model in order to test this hypothesis. What are their respective R^2? (see Table IV)

7. Do the F-test and conclude.

Exercise 4 Download the data in wage2.dta from the website.

1. Consider the standard wage equation:

$$log(wage) = \beta_0 + \beta_1 educ + \beta_2 exper + \beta_3 tenure + u$$

. State the null hypothesis that another year of general workforce experience has the same effect on log(wage) as another year of tenure with the current employer.

2. Test the null hypothesis in question 1 against a two-sided alternative, at the 5% significance level. What do you conclude?

3. Now you wonder if you errors were satisfying the normality assumption. For sake of comparison, estimate the equation

$$wage = \delta_0 + \delta_1 educ + \delta_2 exper + \delta_3 tenure + v$$

Save the residuals (use the "predict *name_of_residuals_predicted*, res" command) and plot a historgram (use the "hist *name_of_the_variable_to_plot* command).

4. Repeat last question, but with $log(wage)$ as the dependent variable.

5. Would you say that the normality assumption is closer to being satisfied for the level-level model or the log-level model?

6. Conclusion for the test performed at the beginning of the exercise?