

Assignment 4: Due December 8

PLEASE BE SURE TO READ THE DOCUMENT ENTITLED “GENERAL ASSIGNMENT GUIDELINES” BEFORE YOU BEGIN THIS ASSIGNMENT. ALL REFERENCES ARE TO THE 7th EDITION OF STUDENMUND. UNLESS SPECIFIED OTHERWISE, USE A 5% SIGNIFICANCE LEVEL FOR ALL TESTS. ASSIGNMENTS SHOULD BE SUBMITTED THROUGH BRIGHTSPACE EITHER ON OR BEFORE THE DUE DATE.

IF IN DOUBT, PROVIDE MORE DETAIL IN YOUR ANSWERS, RATHER THAN LESS.

NOTE: If a statistical table does not give an entry for the appropriate number(s) of degrees of freedom, then use the closest number(s) of degrees of freedom available.

A. Consider the following variation on the wages model from Assignment 3

$$\begin{aligned}
 WAGES_i = & \beta_0 + \beta_1 SCHOOL_i + \beta_2 EXPER_i + \beta_3 SCHOOLEXPER_i + \beta_4 MALE_i \\
 & + \beta_5 ETHBLACK_i + \beta_6 ETHHISP_i + \epsilon_i
 \end{aligned}$$

$i = 1, 2, \dots, 400$

where the variables are defined as follows:

WAGES = current hourly wages in \$
 SCHOOL = years of schooling
 EXPER = years of out-of-school work experience
 SCHOOLEXPER = interaction between SCHOOL and EXPER
 MALE = gender of respondent (1 if male, 0 otherwise)
 ETHBLACK = African-American (1 if African-American, 0 otherwise)
 ETHHISP = Hispanic (1 if Hispanic, 0 otherwise)

together with the U.S. dataset Wages.dta, which is available in Brightspace.

- a) Using an appropriate set of STATA commands, estimate this model, and then copy and paste the output into your assignment. Be sure to include ALL of your STATA commands in this output.
- b) Using STATA, compute the correlation matrix for the explanatory variables, and then cut and paste this correlation matrix into your assignment.
- c) Using STATA to run the relevant regressions, calculate the VIFs for this model using the steps outlined on p. 234, together with the formula in equation (8.16).
- d) Using Klein’s Rule of Thumb (discussed in footnote 6 on p. 235), do any of the R^2 values from the auxiliary regressions run in part c) suggest the presence of multicollinearity? Explain.

e) Check your answers to part c), by using the STATA vif post-estimation command. Be sure to copy and paste the relevant output into your assignment.

f) Review your answers to parts a), b), c), d), and e), and then draw an overall conclusion as to whether there is any evidence of a multicollinearity issue in the model. Explain.

g) Can you suggest a remedy for the multicollinearity issue, if any, in this model? Explain.

B. [85 marks] Consider the time-series Cobb-Douglas production function regression model from Assignment 3:

$$\ln Y_t = \beta_0 + \beta_1 \ln K_t + \beta_2 \ln L_t + \epsilon_t \quad t = 1, 2, \dots, 20$$

where Y = Output (millions of dollars), K = Capital (millions of dollars), and L = Labour (thousands of workers)

together with the NewCobbDouglas.xlsx dataset, which is available in Brightspace.

a) Using an appropriate set of STATA commands, estimate this double-log regression model, and then copy and paste the output into your assignment. Be sure to include ALL of your STATA commands in this output. NOTE: At the beginning of your STATA program, be sure to tell STATA that this is a time-series regression and that “YEAR” is the time variable. (See p. 9-1 in Chapter 9 of Using STATA: A Practical Guide.)

b) Using STATA, plot the residuals from your estimated regression equation in a line graph against year. (To produce a line graph in STATA, you can just replace “scatter” with “line” in the graph command which I discussed in class.) Cut and paste your line graph into your assignment. Does your graph exhibit any signs of serial correlation? Explain.

c) Conduct a Durbin-Watson test for positive, first-order serial correlation. Be sure to state the null and alternative hypotheses, the appropriate lower and upper critical values from Table B-4, the calculated value of the Durbin-Watson statistic, and your conclusion.

d) Using STATA, together with the steps outlined on p. 290, conduct an LM test for first-order serial correlation. Be sure to state the null and alternative hypotheses, the appropriate critical value from Table B-6, the calculated value of the LM statistic, and your conclusion.

e) Repeat part d), but this time conduct an LM test for second-order serial correlation.

f) Check the calculation of your LM statistics in parts d) and e) by using the single step STATA command, which is discussed in Chapter 9 of Using STATA: A Practical Guide. Be sure to cut and paste the output from this STATA command into your assignment.

- g) Based on your various results above, do you believe that there is a problem with serial correlation in this model? Explain.
- h) Re-estimate the model using the STATA prais command, and then copy and paste the output into your assignment.
- i) Are these GLS coefficient estimates and t-statistics the same as the OLS estimates and t-statistics from part a) above? Why or why not? Explain.
- j) Using STATA, re-estimate the original model from part a) with the appropriate Newey-West standard errors and then copy and paste the output into your assignment.
- k) Are the Newey-West standard errors different from the original OLS standard errors? Why or why not? Explain. Which standard errors do you believe are more appropriate in this context? Explain.