ECO203: Introductory Econometrics

Assignment IV

To be submitted over email before 17:00 on 5th May, 2023. Email should be sent to the instructor. The teaching assistant should be copied on the email.

The objective of this assignment is to ensure that you learn how to handle data using Stata. You are expected to write a do file that contains all the commands that you are using to answer the questions. Feel free to write comments in the dofile to explain what you are trying to achieve with a certain set of commands. You are required to submit the do-file along with a word/pdf file that contains answers to the questions.

Download the dataset titled *Card1995.dta* (all Stata data files have a *.dta* extension) and save it in a folder you can easily access. You can create a folder titled *ECO203_Stata* in your *D* drive and save the dataset in the folder. Open a new Stata dofile . Save the dofile as *ECO203_AssignmentIV.do* in the folder. This might be the same folder in which you have saved the dataset or a subfolder within that folder. Now start writing commands in the dofile. Keep executing the commands in the dofile to ensure that there are no errors in the command.

The dataset contains information on 52 variables for 3613 observations. The pdf document *Card1995_description* defines the variables in the dataset.

Answer the following questions:

1. Consider the following model. <u>Assume</u> that CLRM assumptions are true.

$$log(wage_i) = \beta_0 + \beta_1 education_i + \beta_2 experience_i + \beta_3 black_i + \beta_4 south_i + \beta_5 urban_i + \varepsilon_i$$

   Using the $Card1995$ data estimate the following model and on the basis of the estimation results answer the questions that follow:

   (a) Construction a publication style table in MSWord to report the regression results. The table should contain information on the number of observations used in estimation, OLS estimates, standard errors, R^2, and adjusted R^2.

   (b) What is the percentage increase in wages due to an additional year of education? Explain.

   (c) Does this regression specification allow the marginal returns to education to be different for Blacks and Non- Blacks? Explain.

   (d) 'Everything else being the same, there is no Black–Non-Black wage gap'
   - Is the above sentence true in light of the statement estimation results?

   (e) Is the variable *urban* statistically significant?

   (f) Test: $H_0 : \beta_3 <= -0.1$ against $H_A : \beta_3 > -0.1$. Can the null hypothesis be rejected at a 5% level of significance? What is the p-value?

   (g) Test: $H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$ against $H_A : (\sim H_0)$. Can the null hypothesis be rejected at a 5% level of significance?

2. Consider the following model. <u>Assume</u> that CLRM assumptions are true.

$$log(wage_i) = \beta_0 + \beta_1 education_i + \beta_2 experience_i + \beta_3 \frac{experience^2}{100} + \beta_4 black_i + \beta_5 south_i + \beta_6 urban_i + \varepsilon_i$$

Using the $Card1995$ data estimate the following model and on the basis of the estimation results answer the questions that follow:

(a) What is the percentage increase in wage due to an additional year of experience? Does it depend on the initial level of experience?

(b) What proportion of the total variation in log(wage) is explained by the covariates of this model?

(c) What is the value of adjusted R^2?