Applied Economics (040178)
Yan Xu

Homework 3

You should work in groups (up to four students) and submit only once on Moodle. Please put the name and student number of all students on top of the first page. **The deadline is Jan 16, 2023.** Late submissions will not be accepted. This homework is worth 12.5% of your overall grade. Please hand in an answer sheet (in PDF format) and your R script. For questions regarding the results of running the code, you can use the snipping tool and answer questions with screenshots or captured pictures. Please report your output from the code on your answer sheet.

1. Go to the website `https://www.rug.nl/ggdc/productivity/pwt/` and download the PWT 10.0 dataset (either excel or state format). In the data file, you will find information on income, output, input, and productivity for 183 countries. Import the dataset into R. For this exercise, we are mainly interested in the following variables for the **year 2019**:

   - *rgdpna*: Real GDP at constant prices of 2017 (in millions of USD of 2017)
   - *rnna*: Capital stock at constant prices of 2017 (in millions of USD of 2017)
   - $rtfpna$: Total factor productivity at constant prices of 2017
   - *pop*: Population (in millions)
   - *emp*: Employed population (in millions)
   - *avh*: Average hours worked per year by active people
   - *hc*: Human capital index, based on years of schooling and returns to education

   (a) Estimate a regression by OLS where the dependent variable is the real GDP (*rgdpna*) and the explanatory variable is the capital stock (*rnna*). Based on your intuition and knowledge of economics, which relevant variable(s) are we omitting in this equation?

   (b) Estimate a second regression by OLS where the dependent variable is the real GDP (*rgdpna*) and the explanatory variables are the capital stock (*rnna*), the population (*pop*) and the total factor productivity ($rtfpna$). Interpret the coefficients and test the hypothesis that the coefficients of *pop* and *rnna* are jointly zero. Which are your conclusions?

   (c) Create a graphical scatter plot of real GDP (*rgdpna*) and the stock of capital (*rnna*) and another scatter plot of real GDP (*rgdpna*) and population (*pop*). Are both relationships linear?

   (d) Create the logarithm of real GDP (*rgdpna*), the stock of capital (*rnna*), the population (*pop*) and the total factor productivity ($rtfpna$) and estimate by OLS a regression where the dependent variable is the logarithm of real GDP (*rgdpna*), and the explanatory variables are the logarithm of the stock of capital (*rnna*), the logarithm of the population (*pop*) and the logarithm of the total factor productivity ($rtfpna$). Interpret the estimated coefficients.

(e) Using the results of the previous question, test the hypothesis of constant returns to scale (that is, that the sum of the two estimated coefficients is equal to 1). Does the data verify this hypothesis? Rewrite the previous equations to test the same hypothesis but now based on a **single** estimated coefficient.

(f) Estimate by OLS a regression where the dependent variable is the logarithm of real GDP ($rgdpna$) and the explanatory variables are the logarithm of the stock of capital ($rnna$) and the logarithm of the population (pop). Then, estimate another regression where you use the logarithm of the employed population (emp) instead of the logarithm of the population (pop). Does the estimated coefficient change much?

2. In empirical studies, the regression results for different models will often be reported as in the following table. Below the coefficients for each independent variable is its standard error. I estimated three models. The dependent variables are the log of CEO salary for 177 companies. The independent variable *lsales* is the log of sales, *lmktval* is the log of the market value of the firm, *profmarg* is profit as a percentage of sales, *ceoten* is years as CEO with the current company, and *comten* is total years with the company.

   (a) Explain the coefficient on *lmktval* and answer whether the company's market value significantly affects the CEO salary.

   (b) Interpret the coefficients on *ceoten* and *comten*. Are they statistically significant individually?

   (c) Are the coefficients on *ceoten* and *comten* jointly significant from zero? (Hint: use R^2)

   (d) What do you make of the fact that longer tenure with the company, holding the other factors fixed, is associated with a lower salary?

Table 1: Regression on factors for CEO salaries

	Dependent variable: log(salary)		
	(1)	(2)	(3)
lsales	0.224	0.158	0.188
	(0.027)	(0.040)	(0.040)
lmktval		0.112	0.100
		(0.050)	(0.049)
profmarg		−0.002	−0.002
		(0.002)	(0.002)
ceoten			0.017
			(0.006)
comten			−0.009
			(0.003)
Constant	4.961	4.621	4.572
	(0.200)	(0.254)	(0.253)
Observations	177	177	177
R^2	0.281	0.303	0.353
Adjusted R^2	0.277	0.291	0.334
Residual Std. Error	0.515 (df = 175)	0.510 (df = 173)	0.495 (df = 171)
F Statistic	68.345 (df = 1; 175)	25.128 (df = 3; 173)	18.622 (df = 5; 171)