



## ECON 1313: BASIC ECONOMETRICS

**SEMESTER 3, 2022**

Assessment #2 **Empirical Project Individual Report (40%)**

**Due: on Monday 12 December 2022, at 9:00 AM**

### Submission Instructions:

Your submission needs:

- To be submitted in TII through Canvas
- Word limit: **3000 words** (excluding references and any Appendix). The font used should be Times New Roman 12pt. Include page numbers in either the footer or the header). Please note that an **Appendix** contains information that is **not essential** to your report but supports your analysis and discussion. If you put important information or discussion in an Appendix, it will not be marked.
- You need to submit a Word file **NOT** a PDF file and avoid a high similarity rate in TurnItIn .
- Late work: Deduction of 10% of total marks of the assessment for each day being late.
- There is a penalty of: 5% of the assessed mark for each 100 words over the limit. It is compulsory to specify the word count on the cover page. A deduction of 10% of total marks of the assessment will be made for papers without a word count or which provide an incorrect word count on the cover page.
- The front page of the report should state: (i) Your name and student number; (ii) The lecturer's name; (iii) Your class number; (iv) The word count.
- Be sure to provide a reference list of the resources you have used for your report, and spell check it before submission. Please note that RMIT treats plagiarism very strictly , so be sure to avoid plagiarizing your report or enabling plagiarism of another student's report.

### Required Documents

- You should **submit both Word report and an R script file showing the commands used for your assignment**. (Note: the R script file is only for reference and will not be marked, hence make sure to include your estimation output in the Word report)

In this assessment, you will prepare a report as a part of consulting research project. You have been approached by a client to prepare a report on the determinants of health care expenditure in a cross-section of **one (1)** of the groupings of developing and transitional countries<sup>1</sup> for **one (1)** of the years between 2009 and-2018.

You will estimate multiple regressions of health care expenditures on the following variables from the file named ass2\_data.csv:<sup>2</sup>

- Current health expenditure per capita (Current US\$) (PCHE)
- Birth rate, crude (per 1000 people) (CBR)
- Urbanization (urban population, % of the total population) (Urban)
- Share of the population that is under 15 years of age or above 65 years of age as a percentage of the population (DepRatio)
- GDP per capita (current US\$) (GDPPC)
- Net official development assistance received (current US\$) (NETODA)

You have been assigned to complete this project and submit a report on its findings. **Your report should include the five sections detailed below.**

## **PART 1: Overview and data collection**

### **1. Overview of topic**

Find at least one recent research article from reliable resources about the determinants of health care expenditures in developing and transitional countries. Briefly summarize and discuss what variables are identified in the article that influences health care expenditures and their expected effects.

### **2. Data Importation and Descriptive Statistics**

- a) Import the data for your assigned year and group of countries into R. Prepare a table of descriptive statistics the variables you will analyse Identify if there are any missing observations and decide what you will do about these.
- b) Consider use of the natural logarithm of PCHE, CBR, GDPPC, and NETODA in a model and

---

<sup>1</sup> Countries are divided into groups (A, B or C). Lecturers will assign country groups to students.

<sup>2</sup> This data has been downloaded from the World Bank's open data website ( <https://data.worldbank.org/>) from which you can find further information on the definition of these variables.

briefly explain why you should take the natural logarithm of these variables in a model.

## **PART 2: Descriptive statistics and initial estimation**

1. Briefly discuss the descriptive statistics for ALL variables.
2. Write a population multiple linear regression function of  $\log(PCHE)$  on  $\log(CBR)$ , *Urban*, *DepRatio*,  $\log(GDP \text{ per capita})$ , and *NETODA* (**Model 1**)
3. Estimate **Model 1** using Ordinary Least Squares (OLS) and present your estimated results in a table.

## **PART 3: Interpretation**

1. Write out your results for **Model 1** in standard regression format. Interpret the Goodness-of-Fit -  $R^2$  and comment on the adjusted  $R^2$  from the regression. .
2. Discuss the hypothesis and results of the F-test for **Model 1**, using one of these significance levels: 1%, 5%, and 10%.
3. State the hypothesis and discuss the results of t-tests for each variable in **Model 1**. Interpret the coefficients of all significant variables, using one of these significance levels: 1%, 5%, and 10%.
4. Based on what you have learned the research article in PART 1 or economic theory, do the estimated slope coefficient of **Model 1** support what you had expected? Explain your reasoning briefly.
5. Use a relevant method to check for potential multicollinearity in Model 1. Discuss the results of this method. Discuss the remedies you could apply to treat multicollinearity if it exists.

## **PART 4: Further Estimation**

1. Calculate Median of *DepRatio*. Generate the dummy variable **High DepPop (HDP)** ( $HDP = 1$ , if countries have *Dependent Populations* higher than the median, and 0 = otherwise). Re-estimate Model 1 by replacing *DepRatio* by the dummy variable *HDP*. Write out your sample multiple linear regression of **Model 2**.
2. State the hypothesis and discuss the result of the t-test on the variable *HDP ONLY* in **Model2**. Interpret the coefficient of this variable if it is statistically significant.
3. Generate an **interaction term** between variable *Per Capita GDP* and variable *HDP*. Re-estimate Model 2 adding this interaction term to the model (**Model 3**). Write out your sample

multiple linear regression of **Model 3** in standard regression format.

4. State the hypothesis and discuss the results of the t-test on the **interaction term ONLY** in **Model 3**? Interpret the coefficient of the interaction variable if it is statistically significant.
5. Your colleague suggests re-working the model by using a different functional form or different variables to estimate PCHE (**Model 4**). Explain what the reasons are for using a different functional forms or different variables. Write out your estimating equation for **Model 4**
6. Suggest a final, preferred model (from 1 to 4) you will use to estimate PCHE and interpret all the coefficients in your preferred model.

### **Part 5: Conclusion**

1. Summarize the main findings in your report (from Parts 1 to 4) using non-technical language.
2. Based on your main findings, propose policy recommendations for improving health expenditures in your assigned group of countries.
3. Discuss the limitations of your report and suggest what you can do to improve the estimations in your models.

**- End of the Assessment -**