

**Note:**

- We assume that the standard OLS assumptions hold.
- If not stated otherwise, use  $\alpha = 5\%$ .
- Whenever asked to perform a test, please indicate which test you use, your null hypothesis, the test-statistic, the critical value, the behavior under the null, the decision rule and the decision.
- For any type of test, you only need to do this once, then you can directly conclude while specifying how you have got to your conclusion.
- If you need additional assumptions to draw your conclusion, state them explicitly.
- Make your code readable and understandable. (Spaces, comments, ...)

EXERCISE 1. Please write a well structured and commented code for the following exercises. Your code should be able to run without any error once the working directory is changed. You can get 1 point for a code running without errors, 1 point for the readability and commenting of your code. 1 [2 points]

EXERCISE 2. Environmental effects of openness [10 points]

**Objective:** In their 2005 paper, “Is Trade Good or Bad for the Environment: Sorting out the Causality”, Jeffrey Frankel and Andrew Rose examine the environmental effects of openness to trade in a statistical cross-section of countries in 1995. More specifically, the overall goal of the paper is to empirically address the question, “Do countries which are more open to international trade incur more (or less) environmental damage?” (Frankel, J. A. and A. K. Rose (2005) “Is Trade Good or Bad for the Environment? Sorting Out the Causality.” The Review of Economics and Statistics 87(1), 85-91.) The description of the variables is in Table 1. Take into account that even though the sample size of the dataset is small, you can apply asymptotic results. The data is in the file “*Trade.txt*”.)

1. [2 points] Estimate the following model using OLS. Which coefficients are significant at 5% level? Interpret the coefficient on *inc* and *polity*.

$$\log \text{sulfdm} = \beta_0 + \beta_1 \text{inc}_i + \beta_2 \text{pwtopen}_i + \beta_3 \text{polity}_i + \beta_4 \text{lareapc}_i + \epsilon_i$$

2. [2 points] Add an interaction term of the logarithm of income and OECD country to the model, together with the OECD dummy variable, and estimate it via OLS. What is the average marginal effect of income on concentration of sulphur? Test if there is a significant difference between the marginal effect of income for OECD countries vs non OECD country.
3. [3 points] We want to test for the presence of heteroscedasticity. We suppose that potential heteroscedasticity might be coming from *inc*, *pwtopen*, *polity* and *lareapc* and their squared terms. Test for the presence of heteroscedasticity. Do not forget to write down your null and alternative hypothesis.
4. [3 points] Reestimate the model of question (1) according to the information you have about the conditional variance of the error term. Write down the necessary steps and report the results.

---

<sup>1</sup>You may need the following libraries in R: lmtest, sandwich, MASS, car, gap.

- (a) Assume you have no information on  $Var(u_i|X)$  .
- (b) Assume that the heteroscedasticity is of the form ( $\gamma$  not known):

$$Var(u_i|X) = e^{\gamma_0 + \gamma_1 inc_i + \gamma_2 pwtopen_i + \gamma_3 polity_i + \gamma_4 lareapc_i}$$

- (c) Assume the heteroscedasticity is of the form:  $Var(\epsilon_i|x) = \sigma^2 inc_i$  ( $\sigma^2$  is known).

### EXERCISE 3. Gender differences in wages [8 points]

**Objective:** We would like to study the effects of education on wages and how this effect differs by gender. We have information from the “Current Population Survey” (CPS), which provides data on labor force characteristics of the population, including the level of employment and earnings. The data is in the file *cps\_85.txt*

- [2 points] Estimate a regression of *logwage* on *education*, *experience*, *married* and *south*.
- [2 points] Are the previous estimated coefficients significant at 1%? Report the Confidence Intervals at 1% level. Interpret the coefficient on *married*.
- [4 points] We want to check if the model is the same for men and women. Do a stability test. State the null and the alternative hypotheses, the test statistic, its behavior under the null hypothesis, the rejection rule, and conclude.<sup>2</sup>

Table 1: Description of Variables

sulfdm	mean $SO^2$ concentration in micrograms per cubic meter (from 1990)
inc	logarithm of real per capita DGP
pwtopen	(imports+exports)/DGP*100
polity	index of democratic (+10) vs autocratic (-10) institutions
lareapc	logarithm of land area per capita
oecd	dummy variable =1 for oecd countries
country	country name (41 countries)

Table 2: Description of Variables for Exercise 2

wage	hourly wage in dollars
education	Number of years of education
experience	Number of years of potential work experience
age	age in years
south	1 if region is South, 0 for other regions
female	1 if female, 0 if male
married	1 if married
manufacture	1 if works in manufacture, 0 if works in other sectors

<sup>2</sup>In R, you can use `chowtest` command using library `gap`