

Problem Set 1

Econometrics 20203, Part I

Due by midnight of 17 March 2025

To start, unzip the file “assignment1_class20_data.zip” in “Assignment I” folder and extract your group dataset: “assignment_data_group_.dta”, where instead of * you will find the index that corresponds to your group ID.*

In this problem set, you are asked to work with a simulated cross-section of firms. Each firm belongs to one of four sectors, as indicated by the variable sector. You have to estimate a linear regression model using Stata.

n = natural logarithm of employment

w = natural logarithm of relative wage

k = natural logarithm of physical capital

ys = sectoral output

sector = codes for 4 sectors indicating the firm sectoral membership. Each firm belongs to one and only one sector.

You have to carry out the following tasks

1. Describe the data
2. Estimate an employment equation through an OLS regression with dependent variable n and regressors w and k .
3. Test the significance of coefficients on w and k and comment.
4. By using ys and/or sectoral indicators test the null of homogeneity across sectors
5. If you find evidence of sectoral heterogeneity, find a convenient way to accommodate it in your regression model.
6. Test the null of conditional homoskedasticity and comment.
7. If you find evidence of heteroskedasticity, find a way to address the issue and rerun regression analysis with the correction.

Once done, send a zip file to (your emails) with

- (i) your simulated initial dataset (“assignment_data_group_*.dta”)
- (ii) the dofile performing the requested tasks
- (iii) a pdf file with the results and the comments (no more than 2 pages).