Introduction to Econometrics
Session 9 – A Final Recap

December 2025

1 Problem

The purpose of this problem is to summarize as much as possible everything that has been covered since the beginning of the semester. In order to get as close as possible to a real-world situation, the names of the variables to be used in the tables, as well as the functions and packages to rely on, are no longer specified: it is up to you to think carefully about how to use the documentation and to choose the appropriate tools.

1. Download the OECD dataset *"Full-time and part-time employment based on OECD-harmonized definition"*, without filtering, in `.csv` format.
2. Download the OECD dataset *"Gender wage gap"*, without filtering, in `.csv` format.
3. Create two tables in R from the downloaded files.
4. Restrict the table related to working time to observations covering the total population.
5. For each sex, each country, and each available year, compute the share of part-time work in salaried employment.
6. For each country and each available year, compute the ratio between the share of part-time work among women and the share of part-time work among men.
7. Restrict the wage gap table to average wage gaps.
8. Merge the two resulting tables by country and year of observation.
9. How many different countries are represented? How many years are available?
10. Plot the gender wage gap as a function of the ratio, using a different color for each year.

11. Restricting the sample to the period starting in 1980, estimate the regression of the average gender wage gap on the ratio of part-time employment shares, controlling for the year of observation. How should the estimated coefficient be interpreted?

12. Compute the 95% confidence interval for this coefficient, using the appropriate variance–covariance matrix.

13. Plot the gender wage gap as a function of the ratio, using a different color for each country.

14. Still restricting to the same period, estimate the regression of the average gender wage gap on the ratio of part-time employment shares, including country fixed effects. How should the estimated coefficient be interpreted?

15. How does this coefficient compare to the previous one?

16. Perform a Student t-test at the 5% level for this coefficient.

17. What criticism can be made of this way of estimating uncertainty?