# ECON4008 Applied Microeconometrics

### *GROUP PROJECT*:
### Attempt either Project A or Project B

Submission deadline:  **May 15, 2025  at 3pm.**

The word limit for this project is 3000. This limit does not include appendices where unedited regression outputs and your Stata log  file should be reported.

**PROJECT A**

The aim of this project is to estimate the impact of **boardroom gender diversity** on **gender pay gap**  propensity score-based estimators.

The dataset **econ4008_fame.*dta*** contains data on the treatment variables and potential confounders (ultimately it is your decision on which variables to use as confounders ,and also on the type of data cleaning and transformation to undertake prior to the analysis).

This database was constructed from the FAME (Financial Analysis Made Easy) database which can be accessed via **NUsearch,** although it is not necessary for you to access this database unless you want to add more variables to your analysis. Please note that in the provided dataset, the company identifier is called **CompanyNumber**, and we have yearly data for the calendar years 2017-2023. It is up to you to decide which years to use for your analysis.

The treatment variable should be  defined as some  measure of  female presence in the boardroom which you can construct from the variable **NF_dir** which  gives the total number of female directors in the company. It is always a good idea to  explore more than one definition of the treatment variable.

You should collect your own outcome variable(s) form the UK government website https://gender-pay-gap.service.gov.uk/viewing/download. As of the writing of these notes, data (CSV files) were available from  2017-18  to 2014-25 fiscal years. For the

purpose of this project , take 2017-18 fiscal year as 2017; 2018-17 fiscal year as 2018, etc.

Each CSV file identifies companies by their **CompanyNumber.** Hence, this variable. can be used to merge the gender pay gap data with the econ4008_fame.dta. You should of course clean the former and convert them to Stata format first. Naturally, you should not expect to have a perfect match between the two datasets (because they are provided by two different entities).

You will find out that there are variously defined gender pay gap variables (please read the documentation at https://gender-pay-gap.service.gov.uk/, in particular the "Making your calculations" section ).The "standard" pay gap variable is perhaps the mean (average) gender pay gap in hourly pay as a percentage of men's pay (**DiffMeanHourlyPercent)**, but it is always a good idea to explore more than one outcome variable.

Your report should be structured as follows:

**Section 1: [weight 10%]:**

Provide a brief and selective literature review of maximum five studies looking at the relationship between boardroom gender diversity and company performance (ideally focusing on gender pay-gap. Your review should concentrate on the econometric approaches and identification strategies of the papers.

**Section 2: [weight 20%]:**

Provide a discussion of your databases merging process. Explain clearly any data cleaning and transformation exercises you may have carried out. Show the most important features of the data, describe any interesting patterns or correlations, and provide some summary statistics/graphs of the main variables of interest.

**Section 3: [weight 40%]:**

Explain very briefly your econometric approach to evaluate the casual effects of boardroom gender diversity on the outcome variable(s) of your choice (you can assume that your readers know the basic principles of propensity score-based estimation). You are encouraged to estimate more than one model and probe the sensitivity of your findings to alternative model specifications. Write a report on your

main findings, indicating which of the estimators, if any, would you prefer most in the context of this exercise, and why?

**Section 4: [weight 30%]:**

Try to answer the question as to whether your conclusions from Section 3 change if you re-estimate the casual effects of boardroom gender diversity differently. You are encouraged to experiment with different treatment effect estimators as well as alternative outcome variables (even if the extent of data availability may differ from one gender pay gap measure to another).

**Section 5: [weight 10%]:**

This is a summary and conclusion section where you should give an overall evaluation of your work including possible shortcomings.

**Appendix**: Your Stata log file.

# Project B:

This project involves an extended replication exercise based on a published article satisfying the following criteria:

(i) It is published in an economics journal ( e.g. *American Economic Review* or *American Economic Journal: Applied Economics*) from 2010 onwards.

(ii) It uses differences-in-differences to identify some causal effects.

(iii) There is time variation in the treatment under study (i.e. we have staggered treatment adoption), but the paper used a version of the standard two-way fixed effects model.

(iv) The data used by the paper is made publicly available by the author(s) in the relevant journal website.

(v) According to Google Scholar, the paper is cited at least 20 times.

The aim of the project is to investigate whether the main conclusions of the paper can be replicated if one or more of the new difference-in-differences estimators in the presence of staggered treatments are used instead of the standard two-way fixed effects model.

The report of your replication effort should be structured as follows:

**Section 1 [weight 15%]**

A non-technical summary of the paper you seek to replicate. Try to explain in your own words the context of the article and identify its main research question(s) and why it is interesting to answer them. Also describe the data used by the study, and its main conclusions.

**Section 2 [weight 70%]:**

Re-estimate the main models using one or more of the "staggered" difference-in-differences estimators. Also investigate the sensitivity of these results to alternative approaches such as (i) the inclusion or exclusion of certain variables; (ii) estimations on different subsamples; (iii) or even a different modelling approach if you believe staggered difference-in-differences is not the best method to use in the present context. Write a report of your main findings.

**Section 3 [weight 15%]:**

A brief conclusion explaining how close you have come to replicating the results. If you have not been close enough at replicating the results, try to explain as to why this might be the case.

**Appendix**: Your Stata log file.

## Criteria for Marking

Good marks will be awarded for projects that have the following qualities:

- ❑ Accurate results.
- ❑ Evidence of critical thought and analysis.
- ❑ Evidence of reading beyond the lecture material and main textbooks.
- ❑ Logical structure.
- ❑ Good standard of sentence construction, grammar, and presentation.

## Further guidelines:

a. A group can consist of a <u>maximum of 3 students</u>.

b. Students are responsible for forming their own group, and for this purpose a group sign-up list is set up on Moodle.

c. Each member of a group will receive the same mark for the group project unless representations are received in writing from *all* group members requesting an unequal mark allocation. Whilst any case for an unequal allocation of marks may be submitted with a group project, no case for an unequal allocation of marks from any group member can be considered once a project has been marked.

d. Students who do not join a group of appropriate size and who do not inform the module convenor of the same by March **15, 2025,** will receive a mark of zero for the coursework unless they submit an individual coursework

e. Students who report that they have been unable to join a group of appropriate size, or who leave a group and are unable or unwilling to join another will have to undertake a substantial piece of individual coursework in lieu of the group project.

Please email me if you are going to submit an individual coursework stating (i) the reason why you can't be part of group, and (ii) which of the two projects you want to work on.

Professor Sourafel Girma
[sourafel.girma@nottingham.ac.uk](mailto:sourafel.girma@nottingham.ac.uk)