MRM 2024 3rd Term

Prof. Valentina Raponi
Prof. e-mail: vraponi@iese.edu
Office: E-306
Phone: 504021

Assistant: Mónica Trujillo-Bencomo
Assistant e-mail: mtrujillo-bencomo@iese.edu
Office: E-314
Phone: 506551

# ECONOMETRICS II

## Introduction

*This course is based on Econometric I. The first part of the course focuses on causal inference and introduces core statistical topics and techniques necessary for a causal interpretation of regression results. Topics include Differences-in-Differences, Synthetic Controls, and Regression Discontinuity Designs.*

*The second part of the course introduces modern Statistical Learning tools for modeling and understanding complex datasets. We will cover methods like Lasso and Sparse Regression, Discriminant Analysis, Regression Trees, and Unsupervised Learning techniques such as Principal Component Analysis and Clustering Methods.*

*To facilitate the understanding of all the methods considered in the course, as well as to motivate their practical usage, each session will walk you through a realistic application of the methods using R tutorials.*

*This course is intended for people who are already conversant with probability, statistics and basic econometrics, as well as the basics of R programming.*

## Objectives

*This course will enable you to:*

- *Understand the definition of causality and why causality matters: why correlation is not causation, what causation is, and how one can demonstrate it in practice*
- *Apply statistical procedures and implement empirical strategies to arrive at the estimate of the causal effect using non-experimental data*
- *Have a good understanding of the fundamental issues and challenges of machine learning: data, model selection, and model complexity, etc.*
- *Learn core concepts of supervised and unsupervised learning*

## Learning Outcomes

*After successful completion of this course, you will be able to:*

- *Formulate sufficiently well-defined causal questions for comparative effectiveness research*
- *Appreciation of the relationship between statistical estimation and causal identification*
- *Understand how to evaluate models generated from data*
- *Design and implement various machine learning algorithms in a range of real-world applications.*

## Content

*The first part of the course introduces causal inference and the potential outcome framework. The focus will be on the identifying assumptions and the estimation of causal effects using techniques such as Difference-in-Differences, Regression Discontinuity Designs and Synthetic Control.*

*In the second part of the course , basic elements of Machine Learning will be provided, covering methods like and Sparse Regressions, Classification, Regression Trees, and Unsupervised Learning techniques such as Principal Component Analysis and Clustering Methods.*

## Methodology

***Lecture Notes**: Lecture slides will be provided in each class.*

***R codes**: The R codes used in class will be also available after each lecture.*

***Main References**: The main references for the course are:*

*Part I*

*- Angrist, J.D., and Pischke, J-S., 2008," Mostly Harmless Econometrics: An Empiricist's Companion".*
*- W. Greene, "Econometric Analysis"*

*Part II*
*- James, G., Witten, D., Hastie, T., Tibshirani, R. "An Introduction to Statistical Learning" Springer.*

***Required readings**: Other relevant papers/readings will be posted during each class.*

***Software Requirements**: We will use RStudio throughout the course. Both R and RStudio are free and open source. To set them up on your own computer, first download and install R from http://cran.r-project.org/.*
*Then, download and install RStudio from http://rstudio.org/download/desktop.*

***Optional Texts:***
*For students who want a deeper theoretical grounding in the material covered as well as the basics of R programming, additional books are listed below:*

- *"The Elements of Statistical Learning" by Hastie, Tibshirani, and Friedman.*
- *"R for Data Science", by Wickham & Grolemund*

## Evaluation

*Grades for this course will be based on individual assignments (40%) and a final exam (60%).*

***Individual Assignments***
*I will assign several problem-sets over the course. Each problem set will be due a deadline specified in class. Problem set solutions should be submitted by email and include both the pdf report and the R code to replicate your results. You may discuss problem set questions and how to solve them with your colleagues, but your code and write-up must be your own.*

***Final Exam***
*General exam instructions will be given during the course*

## Contact

*You will find assignment questions, readings and other course material on the course web page at https://campus.iese.edu. Occasionally, I will also send e-mails to the class regarding logistical matters, such as information about when an assignment is due. Please check your e-mail before coming to class.*

*The best way to get in touch with me is by e-mail: vraponi@iese.edu. You can drop by my office (E-306) if you have questions about the course or would like to talk about a research idea. I greatly value your feedback on any aspects of this course. Please feel free to contact me in person or by e-mail with suggestions.*

# Course Outline

*The topics scheduled below are approximate. I will post any updated version of this outline on the course web page.*

***PART I - Causality***

| *Session* | *Session Topics* |
|---|---|
| *1-2* | ***Review of Panel Data.*** |
| *3-4* | ***Treatment evaluation and causality (I).***<br>*Selection Bias and the Difference-in-Differences estimator* |
| *5-6* | ***Treatment evaluation and causality (II).***<br>*Synthetic Control* |
| *7-8* | ***Treatment evaluation and causality (III).***<br>*Regression Discontinuity Design* |

***PART II - Statistical Learning***

| | |
|---|---|
| *9-10* | ***Introduction to supervised and unsupervised learning.***<br>*Test error, training data, MSE, Cross-validation and prediction* |
| *11-12* | ***Supervised Learning for quantitative responses (I).***<br>*Ridge, Lasso and Elastic-Net Regressions* |
| *13-14* | ***Supervised Learning for qualitative responses (II).***<br>*Logistic regression and Discriminant Analysis* |
| *15-16* | ***Tree based methods.***<br>*Decision trees, bagging, random forest and boosting* |
| *17-18* | ***Unsupervised Learning.***<br>*Principal Component Analysis and Clustering Methods* |
| *19-20* | ***Exam.*** |

# Bibliography


- *Angrist, J.D., and Pischke, J-S., 2008,"Mostly Harmless Econometrics: An Em-piricist's Companion".*
- *James, G., Witten, D., Hastie, T., Tibshirani, R. "An Introduction to Statistical Learning" Springer.*
- *"The Elements of Statistical Learning" by Hastie, Tibshirani, and Friedman.*
- *"Econometric Analysis", by W. Greene.*
- *"R for Data Science",by Wickham & Grolemund.*


# Professor's Biography


**Prof. Valentina Raponi**
Assistant Professor of Financial Management

Ph.D. in Finance, Imperial College Business School
Ph.D. in Methodological Statistics, University of Rome La Sapienza
M.Sc in Econometrics and Mathematical Economics, London School of Economics and Political Science
M.Sc in Statistics and Economics, University of Rome La Sapienza
B.Sc in Statistics and Economics, University of Rome La Sapienza
E-mail: vraponi@iese.edu


**Biography**

Valentina Raponi is Assistant Professor at the Financial Management Department of IESE. She holds a Ph.D. in Finance from Imperial College Business School, a PhD in Methodological Statistics from the University of Rome La Sapienza, an MSc in Econometrics and Mathematical Economics from the London School of Economics and Political Science, an MSc in Statistics and Economics from the University of Rome La Sapienza and a BSc in Statistics and Economics from the University of Rome La Sapienza.

Her research interests include Financial Econometrics, Asset Pricing and Portfolio Choice, Econometric Theory and Machine Learning. Her dissertation has been published in the Review of Financial Studies.

Before joining IESE, Valentina has also worked as intern at the European Central Bank, the Bank of Italy and the Italian Ministry of Economy and Finance.

**Areas of interest**
* Financial Econometrics
* Empirical asset pricing
* Asset pricing anomalies
* Portfolio Choice
* Econometric Theory
* Machine Learning