



UNIVERSITÀ  
DEGLI STUDI DI MILANO-BICOCCA

## SYLLABUS DEL CORSO

### Economics For Data Science M

2021-2-F8204B020

---

#### Learning objectives

At the end of the course, you will learn how to:

- Build an economic model to identify and estimate a causal effect.
- Compare and assess alternative applied econometrics strategies to estimate a causal effect.
- Use big data and machine learning for causal inference.
- Understand the advantages and value added of using big data for applied research in social sciences.

At the end of the course, you will be able to:

- Be familiar with the most important approaches to program evaluation to address a variety of policy-relevant research questions.
- Effectively use big data to address important research questions in Economics.

#### Contents

This course introduces the emerging field that merges Economics and Data Science to answer policy relevant research questions. The main goal of the course is to discuss the most important machine learning and applied econometrics techniques to estimate a causal effect, and the advantages of using big data to answer relevant research questions in several applications.

We will discuss four main topics:

- 1) Model's Validity and Causal Inference.
- 2) Machine Learning and Causal Inference.
- 3) Structural Models, Experiments and Machine Learning.
- 4) Empirical Applications Using Big Data.

## **Detailed program**

Topic 1: Model's Validity and Causal Inference.

- Internal and external validity
- Big data: new frontiers for economic analysis

Topic 2: Machine Learning and Causal Inference.

- Program evaluation and the "missing" counterfactuals
- Causal inference and machine learning

Topic 3: Structural Models, Experiments and Machine Learning

- Developing, estimating and using structural models
- Comparing experiments, structural models and machine learning methods for causal inference

Topic 4: Empirical Applications Using Big Data

- Students' presentations: present your own work, one of the papers from a list of suggested papers that will be provided, or a paper of your choice that uses machine learning methods, possibly replicating the results of the paper you choose to present.

## **Prerequisites**

Principles of applied econometrics and statistical quantitative methods for data analysis.

## Teaching methods

Lectures and students' projects (individual or group).

## Assessment methods

Presentations and projects that apply the economic models and methods to the data.

## Textbooks and Reading Materials

Textbooks: there is no given recommended textbook for this course. Below a list of some Econometrics textbooks for reference. All listed textbooks are available as e-books with the exception of Wooldridge (2020), which is available at the University Library (both Central Site and Science Site).

Advanced:

- W. H. Greene. *Econometric Analysis*, 5<sup>th</sup> Edition, Prentice Hall International, 2003.

Simpler/less math:

- J. Wooldridge. *Introductory Econometrics: A Modern Approach*, 7<sup>th</sup> Edition, Cengage Learning, 2020. (for IV and 2 stage least squares)
- Stock and Watson, *Introduction to Econometrics*, 3<sup>rd</sup> Edition. (Basic statistics and regression analysis; companion website with datasets and files to replicate empirical results: [https://wps.pearsoned.com/aw\\_stock\\_ie\\_3/178/45691/11696965.cw/index.html](https://wps.pearsoned.com/aw_stock_ie_3/178/45691/11696965.cw/index.html))
- Angrist and Pischke, *Mostly Harmless Econometrics*, Princeton and Oxford University Press, 2009. (Excellent for concept of causality, experiments, diff-in-diff, and RD)

Journal articles and book chapters: the discussion of each the four topics that we will discuss will make reference to the following main articles.

Topic 1: Model's Validity and Causal Inference.

- Angrist and Pischke, *Mostly Harmless Econometrics*, Princeton and Oxford University Press, 2009, Chapter 1 pages 3-8.
- Athey, Susan, 2017. "Beyond prediction: Using big data for policy problems", *Science*, 355, 483-485.

- Einav L., and J. Levin. 2014. "Economics in the Age of Big Data", *Science*, Vol 346, Issue 6210: 1243089.
- Kleinberg, John, Jens Ludwig, Sendhil Mullainathan, and Ziad Obermeyer. 2015. "Prediction Policy Problems" *American Economic Review*, 105(5): 491–95.
- Kleinberg, John, Jens Ludwig, and Sendhil Mullainathan. 2016. "A Guide to Solving Social Problems with Machine Learning", *Harvard Business Review*.
- Shmueli, G. 2010. "To Explain or to Predict?", *Statistical Science*, Vol. 25, No. 3, 289-310.
- Varian, H. 2014. "Big Data: New Tricks for Econometrics", *Journal of Economic Perspectives* 28, 3-28

Topic 2: Machine Learning and Causal Inference.

- Athey S. and G. Imbens. 2015. "Machine Learning Methods in Economics and Econometrics", *American Economic Review: Papers and Proceedings*, 105(5): 476-480.
- Varian, H. (2016) "Causal Inference in Economics and Marketing", *PNAS*, Vol. 113, No. 27, pages 7310-7315.
- Angrist and Pischke, *Mostly Harmless Econometrics*, Princeton and Oxford University Press, 2009, Chapter 2 pages 11-24, Chapter 6 pages 251-267.

Topic 3: Structural Models, Experiments and Machine Learning.

- Attanasio O., Costas Meghir and Ana Santiago. 2011. "Education choices in Mexico: using a structural model and a randomised experiment to evaluate Progres", *Review of Economic Studies*.
- Todd, Petra, and Kenneth, I. Wolpin. 2006. "Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioural Model of Child Schooling and Fertility", *American Economic Review*, Vol. 96, Issue 5, pages 1384-1417.

Topic 4: Empirical Applications Using Big Data.

References: a list of papers will be provided.

## **Semester**

Second semester.

## **Teaching language**

English.

---