Syllabus

[cmcd.economia@fgv.br]

Course: ***Microeconometrics 1***

Professors:

Teaching Assistant: TBD

2026 THIRD QUARTER

PROGRAM

This course is the first part of a sequence of two courses that presents core topics in Microconometrics. It covers topics about Causal Inference and Randomized Experiments, and Selection on Observables. Its main goal is to prepare the applied researcher to use econometric tools appropriately.

TOPICS

1) Causal Inference and Random Experiments
   a. Causality and Potential Outcomes
   b. The Spectrum of Experimentation in Economics
   c. Sample Size, Design and Power Calculation
   d. Inference: Clustering, Bootstrap, Randomization Inference, Multiple Testing
   e. Attrition and Sample Selection
   f. Externalities and Peer Effects
   g. Internal v. External Validity

2) Observational Data: Selection on Observables
   a. Conditional Independence Assumption
   b. Regression, Inverse Probability Weighting (IPW) and Matching Estimators
   c. Estimation and Inference
   d. Specification Tests and Sensitivity Analysis
   e. Multi-valued Discrete and Continuous Treatments

GRADING

Presentation: 30%

- The length of the presentation will depend on the number of enrolled students.
- Whether the presentation will be individual or in pairs will depend on the number of enrolled students.
- We will post another document with more information on what we expect from these presentations. Read it carefully. Here, we highlight that the goal of the presentation is to offer a thoughtful critique of the article, discussing its weak points and suggesting possible improvements.
- To help you prepare your presentation, we recommend two readings:
    - Chapter 3 in *Doing Economics What You Should Have Learned in Grad School—But Didn't* by Marc F. Bellemare (https://mitpress.mit.edu/books/doing-economics)
    - Chapter 3 in *A Guide for the Young Economist* by William Thomson (https://www.amazon.com.br/Guide-Young-Economist-William-Thomson/dp/026251589X)

Referee Reports: 50% [Deadline: TBD]

- You will write two referee reports in this course. One of them must use methods discussed by and the other must use methods discussed by
- Each referee report must have at most 1000 words.
- To help you prepare your referee report, we recommend two readings:
    - Chapter 6.3 in *Doing Economics What You Should Have Learned in Grad School—But Didn't* by Marc F. Bellemare (https://mitpress.mit.edu/books/doing-economics)
    - Chapter 4 in *A Guide for the Young Economist* by William Thomson (https://www.amazon.com.br/Guide-Young-Economist-William-Thomson/dp/026251589X)
- There is no need to write a “Letter to the Editor”.

Problem Sets: 20%. The problem sets may cover:

- Some theoretical questions from the lectures. It may include particular examples or lengthier proofs.
- Coding exercises (for example, Monte Carlo simulations or specific empirical applications).
- Analyzing and criticizing papers.
- The student may submit each problem set one week after the official deadline. In case of a late submission, the student's score will be penalized in 20 p.p. in this problem set. For example, suppose the student scored 250 points out of 300 points in a late submission. In this case, the student's score in this problem set will be $S = 100 \times \frac{250}{300} - 20 = 63.3$. This policy is set as a "no-questions asked" excuse to help students who, for some reason, are unable to complete submit their problem sets on time.

To pass the course, the student must achieve a final score greater than or equal to 60.

(TENTATIVE) SCHEDULE AND BIBLIOGRAPHY

Lecture 1) TBD**:**

a. Introduction to Experiments in Economics & Different Kinds of Randomization

- AP, Chapters 2 and 3
- Toolkit, Session 1 to 3
- GT, Chapter 6

Lecture 2) TBD**:**

a. Sample Size, Design and Power Calculation

- Toolkit, Session 4
- GT, Chapter 9

Lecture 3) TBD**:**

a. Inference: Clustering, Bootstrap, Randomization Inference, Multiple Testing

- Toolkit, Chapter 7
- TBA

Lecture 4) TBD**:**

a. Sampling-based vs design-based uncertainty

- Abadie, A., Athey, S., Imbens, G. W., and Wooldridge, J. M. (2020). Sampling-based versus design-based uncertainty in regression analysis. Econometrica, 88(1):265–296.
- Abadie, A., Athey, S., Imbens, G. W., and Wooldridge, J. M. (2022). When Should You Adjust Standard Errors for Clustering?. The Quarterly Journal of Economics

Lecture 5) TBD**:**

a. Attrition and Sample Selection

- Lee, D. (2009) Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects. *The Review of Economic Studies*, 76 (3)
- GT, Chapter 7

Lecture 6) TBD**:**

a. Externalities and Peer Effects

b. Internal v. External Validity

- Toolkit, Session 8

Lecture 7) TBD TBD**:**

a. Conditional Independence Assumption

- IW, Sessions 3, 5.1, 5.4, 5.10

Lecture 8) TBD**:**

a. Regression (Distribution, Quantile, Mean), IPW and Matching

- IW, Sessions 5.3-5.8
- Mixtape, Chapter 5
- Koenker, R. & Hallock, K. (2001). Quantile Regression: An Introduction
- Koenker, R (2019). Quantile Regression in R: A Vignette

Lecture 9) TBD**:**

a. Estimation and Inference under CIA

- IW, Sessions 5.2, 5.9, 5.12

Lecture 10) TBD**:**

a. Specification Tests and Sensitivity Analysis

- IW, Session 5.11, 5.13, 6.2
- Pre-testing Issues
- Sant'Anna, P. & Song, X. (2019). Specification Tests for the Propensity Score. *Journal of Econometrics*, 210 (2)
- Cinelli, C. & Hazlett, C. (2020). Making Sense of Sensitivity: Extending Omitted Variable Bias. *Journal of the Royal Statistical Society – Series B*, 82 (1)

Lecture 11) TBD**:**

a. Multi-value Discrete and Continuous Treatments

- Calonico, S. & Cattaneo, M. & Farrell, M. (2019). nprobust: Nonparametric Kernel-Based Estimation and Robust Bias-Corrected Inference. *Journal of Statistical Software* 91 (8).
- Goldsmith-Pinkham, P. & Hull, P. & Kolesár, M. (2022). Contamination Bias in Linear Regressions. Working Paper (https://about.peterhull.net/wp)

Lecture 12) TBD **and** TBD **:** Students' Presentation

CONTACT INFORMATION

Instructor:
Email:
Office:

Instructor:
Email:
Office:
Office Hours: TBD

- Office hours are a pre-defined period of time that is entirely dedicated to students. During this period, feel free to drop by my office for any reason. You can ask me any questions related to the course, to academia, to your career, to Economics or to your life as a graduate student.

TA: TBD
Contact: TBD

ACCEPTABLE USE POLICY

You are free to use any published materials (e.g., a textbook), in preparing your assignments or for learning the material more generally. Similarly, you are free to use online resources such as stackoverflow questions or R tutorials. You are also strongly encouraged to work with others in your class. This is particularly helpful for learning to code. Each person must turn in their own assignment.

We do not tolerate cheating and plagiarism. Cheating or plagiarism will result in a 0 on the assignment and will be reported to the department. You are welcome to work together in groups, but you are required to submit your own write-up and your own code.

Please take precautions to avoid putting us in a situation where we are forced to decide if two documents are “too similar”. As future researchers, consultants, bankers, entrepreneurs etc, learning to do honest work in a timely manner is more important than getting everything correct.

If you are uncertain, please add proper citation. For example, if you relied heavily on a group-member's code for one part of an assignment, then you should make a footnote highlighting this fact. As long as proper credit is clearly given, it does not constitute cheating.

TEXTBOOKS AND MAIN SURVEY PAPERS

This course covers a large variety of topics. Consequently, we use many different references, including textbooks, survey papers and specific papers. Here, we list the references that are repeatedly used through this course.

- **Toolkit:** Kremer, M., Glennerster, R., & Duflo, E. (2007). Using Randomization in Development Economics Research: A Toolkit.
- **AP:** Angrist, J. D. & Pischke, J. (2009). Mostly Harmless Econometrics. Princeton University Press.

- **IW:** Imbens, G. & Wooldridge, J. (2009). Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature*, 47 (1)
- **Mixtape:** Cunningham, S. (2020). Causal Inference. The Mixtape, 1. (https://mixtape.scunning.com/)
- **GT:** Glewwe and Todd (2022). Impact Evaluation in International Development: Theory, Methods, and Practice. (http://hdl.handle.net/10986/37152)

There are many textbooks that cover topics related to the ones covered in this course. A standard textbook about cross-sectional and panel data is Wooldridge's *Econometric Analysis of Cross Section and Panel Data*. Another standard option that covers cross-section, panel data and time series topics is *Econometrics* by Hayashi.

I don't think that there is an optimal textbook. Some books work best for some students and other books work best for other students. Sometimes, you will need to read more than one reference to really understand a topic. Therefore, feel free to check any references that you find useful.

SOFTWARE

Much of the course work in this course will involve analysis of data using *R*, an open-source implementation of the object-oriented programming language *S*. It is widely used by applied statisticians and its libraries implement a wide variety of statistical and graphical techniques with applications to a range of disciplines, such as the agricultural and biological sciences, genetics, neuroscience and economics.

R can be downloaded from https://cran.r-project.org. The program documentation is excellent. There are also many excellent and free *R* references available online, for example, *Econometrics in R* (https://cran.r-project.org/doc/contrib/Farnsworth-EconometricsInR.pdf) by G. Farnsworth and *Applied Econometrics with R* (https://link.springer.com/book/10.1007/978-0-387-77318-6) by Christian Kleiber and Achim Zeileis. If your time permits and you want to dig deeper, there are also more programming-oriented references such as *An Introduction to R* (https://cran.r-project.org/doc/manuals/R-intro.pdf) by W. N. Venables, D. M. Smith and the *R* Core Team, and *Hands-on Programming with R* (https://rstudio-education.github.io/hopr/) by Garrett Grolemund. However, I recommend learning by trial and error, as it is the most time efficient approach and sufficient for the type of coding problems that we will consider.

If you have never used *R* (and have never used another programming language), it might be helpful to check Chapter 1-12 in this free book (https://bookdown.org/ndphillips/YaRrr/) by

Nathaniel D. Phillips. There are two more online tutorials that are very informative and quick: *Introduction to R* (https://hhsievertsen.github.io/applied_econ_with_r/) by Hans H. Sievertsen. Moreover, Kyle F. Butts, Nick C. Huntington-Klein and Grant McDermott developed an online tutorial focusing on data cleaning and regression analysis: https://stata2r.github.io/.

Although introductory, two of my favorite references combining *R* and Econometrics is *Introduction to Econometrics with R* (https://www.econometrics-with-r.org/) by Christoph Hanck, Martin Arnold, Alexander Gerber, and Martin Schmelzer, and *Using R for Introductory Econometrics* (http://urfie.net/read/index.html) by Florian Heiss. The theoretical depth of these books is surely not sufficient for a graduate course in Econometrics, but it covers most of the topics in this course and it brings many coding examples that may be helpful to you.

A deeper reference is *Causal Inference: The Mixtape* (https://mixtape.scunning.com/) by Scott Cunningham. It covers many recent econometric techniques and brings coding examples in *R* and *Stata*. However, it focuses exclusively on cross-section and panel data topics.

A similar resource is vikjam's and Shusuke-Hori's GitHub repository (https://github.com/vikjam/mostly-harmless-replication) that replicates the book *Mostly Harmless Econometrics* by Joshua Angrist and Jörn-Steffen Pischke. In this repository, you can find R, Stata, Python and Julia codes.

Moreover, there is a new book by Martin Huber that focus on policy evaluation with causal machine learning: https://drive.switch.ch/index.php/s/tNhKQmkGB48bjfz. It includes many examples in R.

If you are interested in *Python*, you may want to check the following introductory books:

- *Causal Inference for the Brave and True* (https://matheusfacure.github.io/python-causality-handbook/landing-page.html) by Matheus Facure Alves.
- *The Effect* (https://www.theeffectbook.net/) by Nick Huntington-Klein.

CLASSROOMS POLICIES

This class is committed to an inclusive learning environment. All students, teaching staff, and the professor are expected to treat each other with respect and dignity at all times. All community members should enjoy an environment free of any form of harassment, sexual misconduct, discrimination, or intimate partner violence.

Mental Health issues are an extremely common problem among graduate students (https://scholar.harvard.edu/bolotnyy/publications/graduate-student-mental-health-lessons-american-economics-departments).

To help any students, FGV offers the Pró-Saúde Program: https://eesp.fgv.br/pro-saude-fgv. Feel free to contact them at any time.

Moreover, there are many affordable options outside FGV that may be helpful to you too, such as

- Clínica Psicológica Ana Maria Poppovic – PUC-SP: https://www.pucsp.br/clinica/index.html.
- Psicodrama Público no Centro Cultura São Paulo: http://centrocultural.sp.gov.br/2020/03/05/subjetividades-e-espaco-publico/ e https://spcultura.prefeitura.sp.gov.br/evento/26868/.

Furthermore, there are many mindfulness and meditation apps that you may find useful:

- Ten Percent Happier: https://www.tenpercent.com/
- Unwinding Anxiety: https://unwindinganxiety.com/
- Meditopia: https://meditopia.com/pt/ (em Português)