Quantitative Methods for Causal Inference (QMCI)

MBR Program (A/I course)
WS 2025/2026

Lecturer : Dr. Ambre Nicolle
Ecole Nationale Supérieure des Mines Paris
PSL Research University
(Mines Paris – PSL)

This version: 07.07.2025
Subject to minor changes

Course description

The course will provide PhD students with a comprehensive understanding of contemporary causal inference techniques. Focusing on quasi-experimental methods like Difference-in-Differences, Regression Discontinuity Design, and Synthetic Control Methods, the course will emphasize both theoretical foundations and practical applications. Participants will engage in hands-on empirical exercises relying on datasets from published papers in Economics and Management journals. The course aims to enhance students' ability to conduct robust causal analysis in their research.

Stata will be the software used for examples and solutions. However, participants can use Python or R at their convenience during the hands-on session and during the exam.

Course material

- Lecture slides and data files will be made available on Moodle prior to each session.
- Suggested solutions to empirical exercises will be provided after each session (only Stata code).
- All material will be in English.

Prerequisites

The course is open to PhD students currently enrolled in the MBR program and is credited as an A/I course.

- **Important**: Students must (i) have participated in the course Quantitative Methods *before* participating in QMCI **or** (ii) **be enrolled** in Quantitative Methods **this semester** (Winter Semester 25/26).
- Participants to the course should feel comfortable working with Stata – or be proficient in Python or R if they opt for working with alternative software.

Organization of the sessions

The course will be organized in four sessions of five hours and a half (22 hours in total).
Sessions are held **in person** at *Luk-Pool, Ludwigstr. 28 VG, 2.Stock, Raum 207*.

- Session 1: 23.01.26 (9:00 – 12:00 + 13:30 – 16:00)
- Session 2: 26.01.26 (9:00 – 12:00 + 13:30 – 16:00)
- Session 3: 27.01.26 (9:00 – 12:00 + 13:30 – 16:00)
- Session 4: 30.01.26 (9:00 – 12:00 + 13:30 – 16:00)

Attendance to all teaching sessions is mandatory. Like for most MBR courses, an absence of 2 hours overall is tolerated. If you have to leave earlier or arrive later one of the days, please write in advance to let me know.

Examination

The exam will take the form of an in-person **exam of 3 hours** which will combine:

- A set of theoretical questions
  - 50% of the grade.
  - No books, no notes, or computer will be allowed.

- An empirical "open book" exercise similar to one of the exercises solved during class
  - 50% of the grade.
  - Students will use the software of their choice to perform this exercise.
  - Code, figures and tables produced by the students will be evaluated.

Two dates for the exam are offered (course participants have to choose **one**):

- 27.02.2026 **or** 06.03.2026, 09:00 -12:00.
- Location: Luk-Pool, Ludwigstr. 28 VG, 2.Stock, Raum 207 or ISTO's Library.
- Please arrive at 08:30 on the day of the exam.

Important note: No alternative dates or examination type (e.g. homework) will be offered. If you cannot take the exam, please do not register for the course.

Course structure

The course is organized around six topics.

Each topic combines:
- a **theoretical section** where we will present the setting, the assumptions, the properties of estimators, as well as the pros and cons of approaches (~120 minutes)
- an **applied section** where examples from one or several published papers will be presented and discussed (~30 minutes).
- an **exercise** relying on a dataset, solved by the course participants (~90 min). Solutions (in Stata) will be provided at the end of the session.

Outline of the course:
1. Course presentation: outline, organization, examination
2. Introduction and overview of the methods
3. Matching methods
4. Difference in Differences – advanced tools
5. Synthetic Control
6. Regression Discontinuity Design

A detailed version of the outline is presented at the end of this document, with references to all papers and datasets.

Main references

- Cameron, A. C., & Trivedi, P. K. (2010). Microeconometrics using Stata. Stata press.
- Cameron, A. C., & Trivedi, P. K. (2022). Microeconometrics using Stata: Volume 2 Non Linear Models and Causal Inference Methods. Stata press.
- Cunningham, S. (2020). Causal Inference. The Mixtape, 1.
- Huntington-Klein, N. (2021). The effect : An introduction to research design and causality.

Additional references and datasets (*subject to minor changes*)

- Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. Journal of the American Statistical Association, 105(490), 493-505. [SCG]
- Abadie, A., Diamond, A., & Hainmueller, J. (2015). Comparative politics and the synthetic control method. American Journal of Political Science, 59(2), 495-510. [SCG, dataset available]
- Abadie, Alberto. Using synthetic controls: Feasibility, data requirements, and methodological aspects. Journal of Economic Literature 59.2 (2021): 391-425. [SCG, dataset available]
- Almond, D., Chay, K. Y., Lee, D. S. (2005). The costs of low birth weight. The Quarterly Journal of Economics, 120(3), 1031-1083. [PSM]
- Angrist, J. D., & Pischke, J. S. (2009). Mostly harmless econometrics: An empiricist's companion. Princeton university press.
- Angrist, J. D., & Pischke, J. S. (2014). Mastering'metrics: The path from cause to effect. Princeton university press.
- Athey, S., & Imbens, G. W. (2017). The state of applied econometrics: Causality and policy evaluation. Journal of Economic perspectives, 31(2), 3-32.
- Baker, A. C., Larcker, D. F., & Wang, C. C. (2022). How much should we trust staggered difference-in-differences estimates?. Journal of Financial Economics, 144(2), 370-395. [DiD]
- Butts, K. (2021). Difference-in-differences estimation with spatial spillovers. arXiv preprint arXiv:2105.03737. [DiD]
- Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. Journal of Econometrics, 225(2), 200-230. [DiD, dataset:see Cunningham blog below, Stata implementation]
- Callaway, B., Goodman-Bacon, A., & Sant'Anna, P. H. (2024). Difference-in-differences with a continuous treatment (No. w32117). National Bureau of Economic Research.
- Cattaneo, M. D. (2010). Efficient semiparametric estimation of multi-valued treatment effects under ignorability. Journal of Econometrics, 155(2), 138-154. [PSM, dataset available]
- Cheng, C., & Hoekstra, M. (2013). Does strengthening self-defense law deter crime or escalate violence?: Evidence from expansions to castle doctrine. Journal of Human Resources, 48(3), 821-854. [DiD, data available]
- Flammer, C., & Bansal, P. (2017). Does a long-term orientation create value? Evidence from a regression discontinuity. Strategic Management Journal, 38(9), 1827-1847. [RDD]
- Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. Journal of Econometrics, 225(2), 254-277. [DiD]
- Imbens, G. W., & Rubin, D. B. (2015). Causal inference in statistics, social, and biomedical sciences. Cambridge University Press.
- Kreitmeier, D. & Raschky,, P. (2023): The Unintended Consequences of Censoring Digital Technology – Evidence from Italy's ChatGPT Ban. Working paper. [DiD]
- Kretschmer, T., & Peukert, C. (2020). Video killed the radio star? Online music videos and recorded music sales. Information Systems Research, 31(3), 776-800. [DiD]
- Manacorda, M., Miguel, E., Vigorito, A. (2011). Government transfers and political support. American Economic Journal: Applied Economics, 3(3), 1-28. [RDD, dataset available]
- Olden, A., & Møen, J. (2022). The triple difference estimator. The Econometrics Journal, 25(3), 531-553. [Triple DiD]
- Pearl, J., Glymour, M., & Jewell, N. P. (2016). Causal inference in statistics: A primer. John Wiley & Sons.

- Roth, J., Sant'Anna, P. H., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature. Journal of Econometrics, 235(2), 2218-2244.
- Zuo, G. W. (2021). Wired and hired: Employment effects of subsidized broadband Internet for low-income Americans. American Economic Journal: Economic Policy, 13(3), 447-482. [Triple DiD, dataset available]

Detailed outline (*subject to minor changes*)

Topic 0

- Course presentation: outline, organization, and examination.

Topic 1

Introduction and overview of the methods

- Causal Inference : Why ?
- Causal Inference : How ?
- Key definitions
- Key assumptions

📝 *Main references*: Athey and Imbens (2017), Cameron and Trivedi (2022), Cunningham (2020), Imbens and Rubin (2015).

Topic 2

Matching methods

- Matching Methods : Why ?
- Matching Methods : How ?
- A closer look at the Methods
    - Subclassification
    - Matching & IPW
        - Propensity Score Matching
        - Inverse Probability Weighting
        - Nearest Neighbors Matching
        - Coarsened Exact Matching
        - Inverse Probability Weighting
- Paper Illustration
- Hands-on Session

📝 *Main references*: Abadie and Imbens (2006), Athey and Imbens (2017), Cunningham (2020), Dehejia and Wahba (2002), Huntington-Klein (2021).

🖫 *Datasets*: birth_weight from Cattaneo (2010), training_example from Cunningham (2020)

Topic 3a

Difference in Differences – Advanced tools (Part 1)

- Reminders on Simple DiD
    - Canonical DiD
    - TWFE DiD

- The DiD Revolution
  - What's Trending in DiD ?
  - Non Parallel Trends
  - Variation in the Treatment Timing
  - Sampling Issues
- Paper Illustration
- Hands-on Session

📝 *Main references*: Baker et al (2022), Callaway and Sant'Anna (2021), Cunningham (2020), Goodman-Bacon (2021), Roth et al (2023).

🖫 *Datasets:* castle_doctrine_law from Cheng and Hoekstra (2013), callaway from Cunningham (subset from Callaway and Sant'Anna (2021)).

Topic 3b

Difference in Differences – Advanced tools (Part 2)

- The DiD Revolution (Part II)
  - Triple Difference
  - Continuous Treatment DiD
  - Spillovers
- Matching+DiD
  - PSM-DiD
  - DiD-CEM
- Paper Illustration
- Hands-on Session

📝 *Main references* : Butt (2021), Callaway et al (2024), Cameron and Trivedi (2022), Cunningham (2020), Sant'Anna and Zhao (2020)

🖫 *Datasets*: organs from Kessler and Roth (2012), michelin from Favaron et al (2022)

Topic 4

Synthetic Control

- Synthetic Control : Why ?
- Synthetic Control : How ?
- Key Definitions
- Key Assumptions
- A closer look at the Methods
  - The Comparative Case Study
  - Synthetic Controls with Large Datasets
  - SG-DiD
- Paper Illustration
- Hands-on Session

📝 *Main references*: Abadie et al (2010), Abadie et al (2015), Abadie (2021), Cameron and Trivedi (2022), Cunnigham (2020), Zohrehvand et al (2023).

🖫 *Datasets:* texas from Cornwell and Cunningham (2016), repgermany from Abadie et al (2015), selfdriving_cars from Courthoud (2022).

Topic 5

Regression Discontinuity Design

- Regression Design Discontinuity : Why ?
- Regression Design Discontinuity : How ?
- A closer Look at the Methods
  - Sharp RDD
  - Fuzzy RDD
  - Regressions Discontinuity in Time (RDiT)
  - RDD-DiD
- Paper Illustration
- Hands-on Session

📝 *Main references*: Cameron and Trivedi (2022), Cunningham (2020), Huntington-Klein (2021), Flammer and Bansal (2017).

🖫 *Datasets:* gov_transfs from Manacorda et al (2011), mortgage from Feller (2013), diplomas from Clark and Martorell (2014).