Advanced Counterfactual Methods for Causal Inference with Panel Data

Department of Social and Political Sciences
European University Institute

Semester: Spring 2026
Dates: 27–30 April

Abstract

This 20-hour workshop provides an exploration of modern methods for causal inference in panel data settings. Starting from the potential outcomes framework and the canonical difference-in-differences (DiD) model, the course revisits core identification assumptions and discusses how they are commonly violated in applied research. We then survey recent methodological innovations that address these challenges, including staggered adoption designs, generalized DiD estimators, synthetic control methods, and hybrid approaches. Participants will gain both theoretical understanding and practical skills through hands-on applications in R using real datasets. Ample time will be devoted to developing participants' own empirical projects, and we encourage collaborative work among participants.

Organizers

Given by: Roberto Valli, Sebastian Ramirez-Ruiz

Course Description

The increased availability of time-series cross-sectional data offers new opportunities to make causally valid inferences about important relationships of interest. Methodologists have proposed relatively simple designs that allow researchers to make unbiased comparisons under transparent assumptions. Yet real-world settings often deviate from these idealized designs in ways that require us to adapt our empirical strategies to more complex comparisons that go beyond *before-vs-after* and *treated-vs-control*.

This workshop provides an overview of modern methods for causal inference in panel data settings that address these complexities. Starting from the potential outcomes framework and the canonical difference-in-differences (DiD) model, the course revisits core identification assumptions and discusses how they are commonly violated in applied research. Ample space is dedicated to understanding the type of variation and data structures that are commonly encountered in practice and their econometric implications. We then survey recent methodological innovations that address these challenges, including staggered adoption designs, generalized DiD estimators, synthetic control methods, and hybrid approaches.

The workshop has three main objectives. First, you will gain theoretical understanding of standard designs for causal inference with cross-sectional time-series data, and of common issues related to their implementation. Some space will be dedicated to coding examples and hands-on applications in R using real datasets. Yet the main goal is to develop practical problem-solving skills when confronted with challenging empirical settings.

Second, the modern literature on DiD estimators highlights the importance of transparent, explicit, and honest research design. You will have the opportunity to reflect on the role of identification in social science research, and the importance of clearly specifying causal quantities of interest, identifying variation, and identification assumptions.

Third, we aim to help you to progress in your own research. To this goal, you will receive conceptual and practical tools for statistical analysis, obtain peer and instructor feedback on your empirical projects, and dedicate time to fine tune your own projects. Moreover, the content of the class will be fine-tuned around your specific interests and the needs of your projects.

Learning Objectives

By the end of this workshop, participants will be able to:

1. Explain the potential outcomes framework and the assumptions required for causal identification with observational data
2. Derive the canonical difference-in-differences estimator and state its identifying assumptions
3. Diagnose when two-way fixed effects estimation produces biased estimates under staggered adoption and heterogeneous treatment effects
4. Compare modern DiD estimators (e.g., Callaway & Sant'Anna, Sun & Abraham, Borusyak et al.) and select the appropriate one for a given data structure
5. Implement synthetic control and generalized synthetic control methods, and assess pre-treatment fit
6. Evaluate the credibility of an empirical strategy by assessing its identification assumptions, conducting robustness checks, and identifying threats to validity
7. Design a research strategy for a causal question using panel data, declaring the estimand, the source of variation, and the chosen estimator

Course Evaluation

Participants are expected to:

- Develop and present an empirical research idea
- Attend all sessions (both morning and afternoon)
- Actively engage in lab exercises and discussions
- Provide feedback to peers on their projects
- Submit a 5-page research design

Preparation

Before the workshop, every participant will prepare **a single-page document** based on a research question that they want to answer. It can be one that they are currently studying and are facing challenges with, or one they are planning to work on and for which they need a research design. The document will contain:

1. The research question in plain English
   E.g.: What is the effect of unemployment on turnout?

2. A paragraph describing the comparisons that would be needed to answer the RQ
   E.g.: I need to compare individuals who have been fired and those who are still employed...

3. A paragraph describing the potential identification problems
   E.g.: Workers who are fired might be less productive and have different skillsets...

4. A paragraph describing the ideal data that you would need for your study
   E.g.: Monthly individual-level employment data and voter registration records

5. A paragraph describing the ideal variation that you would need for your study
   E.g.: An industry-specific shock increasing unemployment before the election

Deadline for the document is **20 April** (Monday before the course starts). During the workshop, there will be opportunities to work through the challenges of your project and to evaluate different empirical strategies to address them.

Final assignment

After the workshop, participants will hand in a **research design** that develops the initial research question. The final deliverable will be no more than 5 pages long without references.
The research design will include:

- An introduction, outlining the research question, state of the literature, empirical strategy (results, if available), and contribution of the project
- A literature review, describing succinctly the previous findings and empirical strategies in the literature, and why the question has not been answered yet
- A description of the case / setting and the data chosen for the analysis. It should explain the empirical and theoretical leverage it offers, the data-collection strategy, and what tradeoffs exist compared to alternative approaches
- A description of the empirical strategy that declares the causal quantity of interest, the variation used for estimation, and the chosen estimator together with its identification assumptions and the planned robustness tests

Any empirical results, if available, can be included in the Appendix. Deadline for the document is **Friday 29 May**.

Reading Requirements

Readings marked with an asterisk (*) are compulsory. Additional readings provide alternative perspectives, empirical applications, and extensions of specific topics. Please reach out to the instructors for additional sources and readings!

Required Software

This course uses R for all computational work. Participants should have R and RStudio or an equivalent code editor installed on their computers prior to the course. Participants should also be comfortable running linear regressions and manipulating data frames in R.

Course Schedule

The course will take place every day from 9:30 to 12:30 and from 14:00 to 17:00.

Day 1 Morning – Potential outcomes framework, causal identification, and difference-in-differences

Session overview. This session introduces the potential outcomes framework as a foundation for thinking about causality. It will cover how counterfactuals are defined and why randomized experiments set the gold standard for causal inference. You will learn core assumptions needed for causal identification: unconfoundedness, positivity, and stable unit treatment value assumptions (SUTVA). The session then moves to the canonical DiD design with two periods and two groups, explaining the parallel trends assumption, and deriving the main estimation equation with OLS. You will learn when DiD can and cannot be used, and how violations of parallel trends compromise causal inference.

Key topics

- Potential outcomes framework and causal inference assumptions
- Identification problems with observational data
- The difference-in-differences design and its assumptions
- The two-way fixed effects (TWFE) estimator
- Testing and validating parallel trends
- Event studies

Readings

- *Card, D., & Krueger, A. B. (1994). Minimum wages and employment: A case study of the fast-food industry in New Jersey and Pennsylvania. *American Economic Review*, *84*(4), 772–793
- *Chapters 1,2 & 5 of Angrist, J. D., & Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press
- Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, *66*(5), 688–701
- Angrist, J. D., & Pischke, J.-S. (1999). Empirical strategies in labor economics. *Handbook of Labor Economics*, *3*, 1277–1366
- Abadie, A. (2005). Semiparametric difference-in-differences estimators. *Review of Economic Studies*, *72*(1), 1–19

Day 1 Afternoon – Staggered adoption designs and the problem of treatment heterogeneity

Session overview. Many real-world policies are adopted at different times across units, creating staggered adoption designs. This session examines why the canonical TWFE estimator performs poorly under staggered treatment timing when treatment effects are heterogeneous across units and time periods. You will learn about recent criticisms of weighted TWFE estimators, and understand the sources of bias from "forbidden comparisons" and negative weighting.

Lab component. Illustration of bias under TWFE, decomposition of estimates and inspection of cohor weights in `R`.

Key topics

- Staggered treatment adoption and implicit comparisons in TWFEs

- Treatment effect heterogeneity and the limitations of TWFE
- Negative weights and negative weighting problems
- Implications for applied research

Readings

- *Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254–277
- *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. https://doi.org/10.1016/j.jeconom.2023.03.008
- de Chaisemartin, C., & D'Haultfœuille, X. (2022). Two-Way Fixed Effects and Differences-in-Differences with Heterogeneous Treatment Effects: A Survey. *The Econometrics Journal*, *26*(3), C1–C30. https://doi.org/10.1093/ectj/utac017
- Athey, S., Bayati, M., Imbens, G., & Qu, Z. (2019, January). *Ensemble Methods for Causal Effects in Panel Data Settings* (Working Paper No. 25675). https://doi.org/10.3386/w25675

Day 2 Morning – Modern DiD estimators: Addressing limitations of TWFE under staggered adoption

Session overview. Not long after econometricians highlighted the limitations of TWFE estimates of ATT, the field has been flooded by different estimators that address the problems. The different solutions share an explicit separation of identification and aggregation. This session surveys recent innovations in DiD estimation in settings with absorbing treatments, staggered adoption and treatment effect heterogeneity. You will learn about alternative estimators that improve upon TWFE, including clean-comparison estimators, doubly robust approaches, and imputation-based methods. The focus is on understanding how, when and why these estimators work, and how to implement them in practice.

Lab component. Exercises in R to inspect identifying variation in real datasets, and to estimate ATTs and event studies with different estimators.

Key topics

- Different solutions to the staggered adoption problem
- Partial overlap and treatment timing
- Doubly robust estimation for DiD and imputation methods
- Overview of software options
- Assumption checking

Readings

- *Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230
- *de Chaisemartin, C., & D'Haultfœuille, X. (2020). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review*, *110*(9), 2964–2996. https://doi.org/10.1257/aer.20181169
- *Liu, L., Wang, Y., & Xu, Y. (2024). A Practical Guide to Counterfactual Estimators for Causal Inference with Time-Series Cross-Sectional Data. *American Journal of Political Science*, *68*(1), 160–176. https://doi.org/https://doi.org/10.1111/ajps.12723
- *Sant'Anna, P. H., & Zhao, J. (2020). Doubly robust difference-in-differences estimators. *Journal of Econometrics*, *219*(1), 101–122
- Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, *225*(2), 175–199. https://doi.org/10.1016/j.jeconom.2020.09.006

- Kennedy, E. H. (2023). Semiparametric doubly robust estimation of treatment effects. *Statistical Science*, *38*(2), 186–206
- Borusyak, K., Jaravel, X., & Spiess, J. (2024). Revisiting event study designs: Robust and efficient estimation. *Review of Economic Studies*, *91*(6), 3253–3286

Day 2 Afternoon – Synthetic control methods

Session overview. Synthetic control methods construct a weighted combination of control units to match the pre-treatment trajectory of treated units. Whereas DiD typically frame identification as a problem of comparison, synthetic controls frame it as a prediction problem. This session explains how synthetic control estimators differ from standard DiD, their identification assumptions, and when they are preferable to other designs. You will learn both the classic approach and recent extensions for multiple treated units and staggered adoption, with applications and recommendations on best practices.

Lab component. Implementation of synthetic control methods and extensions in R.

Key topics

- Causal identification as a prediction problem
- The synthetic control estimator
- Inference and confidence intervals
- Pre-treatment fit and model validation
- Extensions to multiple treated units, staggered designs and synthetic DiD

Readings

- *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
- *Abadie, A. (2021). Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature*, *59*(2), 391–425. https://doi.org/10.1257/jel.20191450
- * Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., & Wager, S. (2021). Synthetic difference-in-differences. *American Economic Review*, *111*(12), 4088–4118. https://doi.org/10.1257/aer.20190159
- Xu, Y. (2017). Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models. *Political Analysis*, *25*(1), 57–76. https://doi.org/10.1017/pan.2016.2
- Athey, S., & Imbens, G. W. (2019). The state of applied econometrics: Causality and policy evaluation. *Journal of Economic Perspectives*, *33*(2), 3–32

Day 3 Morning – Extensions and hybrid approaches: Choose 2!

Session overview. Real-world data and research questions require comparisons that deviate the canonical DiD design or violate its assumptions. This session explores extensions and hybrid approaches that generalize the canonical DiD framework to address diverse empirical settings. Participants will choose two topics to learn about among: DiDs with continuous treatments, non-absorbing and repeated treatments, regression discontinuity designs applied to panel data (difference in discontinuities), causal mediation analysis within DiD frameworks, and more. These advanced methods extend the toolbox for researchers facing non-binary or complex treatment patterns, and allow investigation of mechanisms underlying treatment effects.

Lab component. Implementation of methods in R.

Potential topics

- Continuous treatments and dose-response analysis
- Non-absorbing and repeated treatments in DiD
- Causal mediation and moderation in DiD settings
- Spatial DiDs
- Factorial DiDs
- Spillovers and compositional change
- Hybrid estimators combining DiD with other designs
  e.g., difference-in-discontinuities, matching and DiD, instrumented and fuzzy DiD, . . .

Readings (illustrative)

- Athey, S., Bayati, M., Imbens, G., & Qu, Z. (2019, January). *Ensemble Methods for Causal Effects in Panel Data Settings* (Working Paper No. 25675). https://doi.org/10.3386/w25675
- Grembi, V., Nannicini, T., & Troiano, U. (2016). Do Fiscal Rules Matter? *American Economic Journal: Applied Economics*, *8*(3), 1–30. https://doi.org/10.1257/app.20150076
- Imai, K., Kim, I. S., & Wang, E. H. (2023). Matching methods for causal inference: A review and look forward. *Political Analysis*, *31*(2), 155–182
- Hazlett, C., & Xu, Y. (2026). Trajectory balancing: A general reweighting approach to causal inference with time-series cross-sectional data
- Xu, Y., Zhao, A., & Ding, P. (2026). Factorial Difference-in-Differences [Forthcoming]. *Journal of the American Statistical Association*. https://doi.org/10.1080/01621459.2026.2628343
- Valli, R. (2026, March). Difference-in-Differences Estimates under Selective Migration. https://osf.io/preprints/socarxiv/s7pw3_v1

Day 3 Afternoon – Student project brainstorming

Session overview. Research design is a dialogue between real-world concerns, theoretical priors, and empirical constraints. This session aims to promote out-of-the-box and rigorous research design, and to improve the participants' projects through peer feedback in small groups of peers. Each participant will have approximately 5 minutes to present their research question, data, proposed identification strategy, and the methodological challenge they are addressing. Peers will provide verbal and written feedback to each other. The session concludes summarizing the challenges, solutions, and open questions that emerged during the group work. This session allows peers and instructors to provide feedback on project design and helps identify common themes across participants' work.

Activities

- Participant presentations in groups
- Peer feedback and discussion
- Instructor guidance on identification strategies

Day 4 Morning – Selecting the right method and package for your research

Session overview. Following the project presentations, this session is dedicated to helping you select the most appropriate methodological approach and software tools for your specific research. Based on research design characteristics (treatment timing, unit heterogeneity, sample size, available data), you will learn how to evaluate and select among the methods covered in the workshop. The session covers practical considerations including robustness checks, sensitivity analysis, and implementation in R, with focus on decision-making frameworks and common pitfalls.

Key topics

- Decision framework for method selection

- Matching research design to appropriate estimators
- Available R packages and their capabilities
- Robustness checks and sensitivity analysis
- Communicating results to audiences
- Common pitfalls and how to avoid them

Readings

- *Notes and decision tree provided by the instructors.

Day 4 Afternoon – Assignment feedback, big picture concerns, and course wrap-up

Session overview. The first half of the final session offers another opportunity for participants to discuss methodological choices related to their research questions. The instructor will discuss general advice and feedback from the pre-workshop memos, and discuss the expectations for the final assignment. The second part of the session addresses big-picture questions about the role of causal inference in applied social science research. From ongoing trends and the credibility revolution, to best practices in research design and scientific communication. The session concludes with a synthesis of key course takeaways and a discussion of future directions for research and learning.

Activities

- Collective feedback and reflection on project brainstorming
- Trends in applied social science research and best practices
- "Big picture:" How much should we be concerned with bias in applied social science research?
- Feedback on course and discussion
- Course synthesis and key takeaways

Readings

- *Torreblanca, C., Dinneen, W., Grossman, G., & Xu, Y. (2026). The Credibility Revolution in Political Science. https://arxiv.org/abs/2601.11542
- *Chiu, A., Lan, X., LIU, Z., & Xu, Y. (2026). Causal Panel Analysis under Parallel Trends: Lessons from a Large Reanalysis Study. *American Political Science Review*, *120*(1), 245–266. https://doi.org/10.1017/S0003055425000243
- Goldsmith-Pinkham, P. S. (2026). *Tracking the Credibility Revolution across Fields* (Working Paper No. w35051). National Bureau of Economic Research. https://doi.org/10.3386/w35051