

Syllabus ECON42720

Causal Inference and Policy Evaluation

Schedule

This course is offered every spring trimester. In spring 2025, our weekly sessions are on *Wednesdays from 9-11am* in D201 Newman. There is a two-week midterm break around St. Patrick's Day. The course has no scheduled lab sessions. Instead, several sessions will be run in flipped classroom mode, that is, students watch lecture videos at home and we discuss applications during the in-person sessions.

Instructor

Prof. Benjamin Elsner, PhD
University College Dublin
School of Economics Office
Newman G206
Email: benjamin.elsner@ucd.ie

I am available for questions before, during, and after the sessions on Wednesday. I am also happy to chat to you 1:1; please email me.

Module Description and Learning Outcomes

This course introduces students to the most commonly used techniques in causal inference and policy evaluation. These techniques have a wide range of applications in the social sciences. Upon completion of this module, students will achieve the following learning outcomes:

1. Develop identification strategies to uncover causal effects for a given research question and in a given context.
2. Applying state-of-the-art causal inference techniques to policy-relevant questions using real-world data and statistical software.
3. Critically evaluate the assumptions and limitations of various causal inference methods.

4. Apply visualisation techniques and effectively communicate and interpret research findings.
5. Set up an effective workflow with version control and markdown.

We will cover the following topics:

1. Foundations of Causality: DAGs ([Mixtape](#))([Effect](#))
2. Econometrics recap ([Mixtape](#))([Effect](#))
3. Randomised Experiments and the Potential Outcomes Framework ([Mixtape](#))
4. Matching and inverse probability weighting ([Mixtape](#))([Effect](#))
5. Instrumental variables ([Mixtape](#))([Effect](#))
6. Regression Discontinuity ([Mixtape](#))([Effect](#))
7. Difference-in-differences ([Mixtape](#))([Effect](#))
8. Synthetic control ([Mixtape](#))([Effect](#))

Prerequisites

Students should have a firm understanding of basic statistical concepts (descriptive statistics, hypothesis testing), regression analysis, and programming with R or a comparable statistical programming language. There are no strict prerequisites regarding courses: it is possible to take this course without having taken MSc-level econometrics before, as long as students have a good idea of the aforementioned concepts/techniques.

Assessment

The assessment will consist of a final exam worth 60% of the grade and two assignments (20% each). Students will get one letter grade in the end.

Final exam

The final exam will take place in the RDS or on campus during the exam period. It's a two-hour closed book exam. Students cannot bring notes. Usually, there is a study week between the last teaching week and the exam period. The exam period lasts two weeks. The exam timetable is set by the UCD exam office. We can accommodate special arrangements (for example for students with disabilities), but we cannot accommodate requests to sit an exam on a different date or online.

Assignments

We will have two empirical assignments that require students to analyse a dataset, visualise content, and write up their results. The first assignment will be given around week 4; the second assignment will be given around week 9 (assuming a 14-week term).

Assessment Policies

- **AI Policy** I am a firm believer in AI as a helpful tool for our work, and I encourage students to use it. I see two major uses for the assignments: helping with coding (debugging, suggesting more efficient solutions), and language editing. If you use AI for your assignment, please add an AI statement at the beginning and explain very briefly what you used it for. You will not lose points if you use AI; I am just curious what students used it for. Also, be careful when using AI: it can produce nonsense results, and for those you will lose a lot of points.
- **Late submission policy:** assignments have to be submitted by the deadline. UCD's late submission policy does not apply. If you know that you will not be able to submit an assignment on time, please let me know as soon as possible.
- **Plagiarism policy:** UCD's [Student Plagiarism Policy](#) applies. I reserve the right to use a plagiarism checker.

Grade Scale

For all assessments, I will apply the **alternative linear grade scale**; the pass mark is 40%:

| Grades | Lower % | Upper % |
|--------|---------|---------|
| A+ | 95 | 100 |
| A | 90 | <95 |
| A- | 85 | <90 |
| B+ | 80 | <85 |
| B | 75 | <80 |
| B- | 70 | <75 |
| C+ | 65 | <70 |
| C | 60 | <65 |
| C- | 55 | <60 |
| D+ | 50 | <55 |
| D | 45 | <50 |
| D- | 40 | <45 |
| E+ | 35 | <40 |
| E | 30 | <35 |
| E- | 25 | <30 |

| Grades | Lower % | Upper % |
|--------|----------------------------|---------|
| F+ | 20 | <25 |
| F | 15 | <20 |
| F- | 10 | <15 |
| G+ | 5 | <10 |
| G | 0.02 | <5 |
| G- | 0.01 | <0.02 |
| NM | 0 | <0.01 |
| ABS | No work submitted / absent | |

Software

TL/DR; you will need:

1. [R](#) and [R Studio](#); please download and install these at the start of the course.
2. A word processor: MS Word is ok but I strongly recommend [Quarto](#).
3. Github: please make an account and download [Github Desktop](#)

Statistical Software

In this course, we will use R. What this means is that all the examples are given in R. Using R comes with lots of advantages: it is a very flexible object-oriented programming language with a huge comparative advantage in data analytics. It is free, which is great, and it has excellent AI support through ChatGPT and Github Copilot.

To work with R, you need to install R itself ([link here](#)) and a user interface. As a user interface, I recommend RStudio ([link here](#)), which is very user friendly and has been optimised for use with R. RStudio Desktop is free and is sufficient for all our applications. An alternative is Visual Studio Code ([link here](#)), which is also great.

Can you use other statistical software packages for the assignments? It depends. I strongly discourage the use of commercial software (STATA, EViews, SPSS). There is simply no need for that. I would approve Python and Julia, but no commercial software or more general languages such as Java or C++. If you use anything else than R, please seek approval before. Also, I cannot provide advice on languages other than R.

Word processing

For writing your assignments, you can use any word processing software. However, I strongly recommend that you learn to use markdown software. Markdown allows for an integrated workflow of data cleaning, exploration, visualisation and write-up. It can handle inputs from

many programming languages (R, Python, SQL, Julia, etc) and markdown languages (R mark-down, html, LaTeX, etc) and create nice-looking outputs in pdf, html, doc, or ppt format. So you can create presentation slides, websites, academic papers, CVs, and many more outputs with markdown. For example, this website was made with a markdown software, as were the lecture slides and the tutorial sheets. If you are interested in a career in data science or data analytics, you will have to learn to use markdown.

Sounds daunting? The basics are actually quite easy to learn. I highly recommend [Quarto](#), which can be used from within RStudio or VSCode.

Version Control

Version control is a crucial tool for anyone writing software. And, yes, when you do empirical work you do write software. Version controls allows efficient management of project files, collaboration with teammates, and tracking of changes over time. With version control, you can maintain different versions of your work, experiment, and easily revert if needed. It promotes transparency by recording every modification, which is valuable when working on complex projects. Additionally, it enables seamless collaboration and merging of changes.

In your assignments you will be asked to show proof that you use version control. The main version control system is [Github](#), which offers free accounts. I particularly recommend [Github Desktop](#), a very user-friendly interface for Github. I will provide a brief introduction to Github.

Textbook(s)

The main textbook for this course is

- Cunningham, Scott. *Causal Inference: The Mixtape*. Yale University Press, 2021. <https://doi.org/10.2307/j.ctv1c29t27>.

I recommend purchasing this book, although it is also available in a free html version at <https://mixtape.scunning.com/>

Other excellent books are

- Huntington-Klein, N. (2021). *The Effect: An Introduction to Research Design and Causality* (1st ed.). Chapman and Hall/CRC. <https://doi.org/10.1201/9781003226055>; see <https://theeffectbook.net/> for a free html version.
- Angrist, J. and J.-S. Pischke (2014). *Mastering 'Metrics*. Princeton University Press. Additional material here: <https://www.masteringmetrics.com/> This book provides a non-technical introduction to the canonical research designs in causal inference.

- Angrist, J. and J.-S. Pischke (2009). *Mostly Harmless Econometrics*. Princeton University Press. Additional material here: <https://www.mostlyharmlesseconometrics.com/>. MHE has become the standard textbook for causal inference in economics. The presentation of the material is more technical than in other books listed here.
- Huber, M. (2023). *Causal Analysis: Impact Evaluation and Causal Machine Learning with Applications in R*. MIT Press. Free E-book version available [here](#). This book covers the standard causal inference techniques covered in this course as well as newer techniques based on machine learning. It is more technical than the other books listed here but well worth reading.
- Békés, G. and G. Kézdi (2021). *Data Analysis for Business, Economics, and Policy*. Cambridge University Press. Additional material here: <https://gabors-data-analysis.com/>. This is an excellent introductory textbook for data analysis in general. It covers data exploration, regression analysis, prediction with basic machine learning techniques, and causal analysis. The chapter on causal analysis is not as detailed as in books dedicated to this topic. However, if you want to have a book that introduces you to lots of important empirical techniques in a fairly non-technical fashion, this book is for you.

Tentative Schedule

The table below provides a tentative schedule. Some topics may not require two weeks, and we will use some of the sessions for labs or tutorials. Details will be announced on Brightspace closer to the time.

| Weeks | Topic | Chapter Mixtape | Chapter Effect |
|-------|---------------------------|-----------------|----------------|
| 1 | Foundations of Causality | 3 | 1-9 |
| 2 | Regression re-cap | 2 | 13 |
| 3 | Potential Outcomes | 4 | N.A. |
| 4/5 | Matching and IPW | 5 | 14 |
| 6/7 | Instrumental Variables | 7 | 19 |
| 8/9 | Midterm break | | |
| 10/11 | Regression Discontinuity | 6 | 20 |
| 12/13 | Difference-in-differences | 9 | 18 |
| 14 | Synthetic Control | 10 | 18 |