

Course Syllabus**GRAD-C6-2001: Statistics II: Statistical Modeling and Causal Inference with R****1. General information**

Course Format	Onsite
Instructor	Asya Magazinnik
Instructor's E-mail	a.magazinnik@hertie-school.org
Faculty Assistant Lab Assistant	Amanda Slater; slater@hertie-school.org Oliver Lang
Instructor's Office Hours	By appointment, Thursdays, 13-15. Please email Amanda Slater to schedule an appointment.

Link to [Study, Examination and Admission Rules and MIA, MDS and MPP Module Handbooks](#)

For information on **course room, times and session dates**, please consult the [Course Plan](#) on *MyStudies*.

Instructor Information:

Asya Magazinnik is the Professor of Social Data Science at the Hertie School. Her research interests include electoral geography, federalism, local politics, and law enforcement. She also works on political methodology, in particular at the intersection of causal inference and formal theory. Her work has appeared in the *American Journal of Political Science*, the *Journal of Politics*, and other outlets. Previously, she was an Assistant Professor of Political Science at MIT. She earned a PhD in Political Science from Princeton University in 2020 and holds an MPP from the Harris School at the University of Chicago.

2. Course Contents and Learning ObjectivesCourse contents:

This course continues the sequence in statistical modeling. Assuming prior knowledge in simple and multiple linear regression modeling, it introduces you to a new perspective on studying causes and effects in social science. The course agenda covers various strategies to uncover causal relationships using statistical tools. We start with reflecting about causality and the ideal research design, and then we learn to use a framework to study causal effects. Then, we revisit common regression estimators of causal effects and learn about their limitations. Next, we focus on matching, instrumental variables, difference-in-differences and fixed effects estimators, regression discontinuity designs, and techniques to explore moderated relationships. All classes divide time between theory and application.

Main learning objectives:

The goals are to (1) acquaint you with some of the most common statistical methods used to demonstrate causality, (2) enable you to implement these with statistical software, and (3) prepare you for Hertie's methods electives.

Software:

We will work with R, RStudio and RMarkdown to implement and practice these techniques. It is assumed that you have some basic knowledge of these tools from Statistics I. If not, we strongly encourage you to familiarize yourself with them prior to the course so as to be able to focus on the substance. I suggest the following resources for self-guided learning:

R and RStudio:

- <https://education.rstudio.com/learn/beginner/>
- <https://posit.co/resources/videos/a-gentle-introduction-to-tidy-statistics-in-r/>
- <https://r4ds.had.co.nz/>

RMarkdown:

- <https://rmarkdown.rstudio.com/lesson-1.html>

Teaching style:

Each session will consist of an in-person lecture, which, alongside the course readings, serves to familiarize you with the session's topic. To become acquainted with R and to learn how to implement various statistical techniques in practice, you will be enrolled in a small-group support lab session taught weekly by a teaching assistant.

Prerequisites:

Statistics I, basic command of R

Diversity Statement:

In this classroom and at the Hertie School more broadly, we respect and value differences in race, gender, gender identity, ethnicity, age, physical and language abilities, culture, religion, sexual orientation, and other identities. Such differences contribute to the strength of our community, and upholding an atmosphere of inclusion and mutual respect is of the utmost importance.

This class will also contain a diversity of backgrounds and experiences with mathematics and data science. It is everyone's responsibility to promote and maintain an environment where all students feel comfortable asking questions and voicing their struggle. Mutually supporting one another will be at the heart of everyone's success in this class, and I hope that more advanced students will help their peers.

Students are encouraged to bring up any concerns relating to these policies, or any other hardships that they encounter, with me as early as possible, and I will make every effort to resolve them.

3. Grading and Assignments

Composition of Final Grade:

Assignment 1: 3 take-home problem sets	Deadline: about 10 days after publication	Submit via Moodle	45% (15% each)
Assignment 2: Midterm exam	Midterm week	On campus	25%
Assignment 3: Replication project	Deadline: December 16	Submit via Moodle	30%

Assignment Details

Assignment 1

This assignment comprises 3 take-home problem sets evenly spaced throughout the semester. Problem sets emphasize hands-on learning and application of the tools of the course to real-world data. The assignment will have to be submitted as a knitted R Markdown file containing both R code and your answers to the questions.

Assignment 2

There will be a midterm exam taking place roughly half-way through the semester (date TBA). The exam will cover the mathematical foundations and theoretical components of the course. The exam may include some pen and paper computations or proofs as well as open-ended discussion questions related to research design.

Assignment 3

As a final assignment, course participants are asked to replicate, check, and amend the analyses of an existing academic paper. The instructor will provide a small set of replication data sets, and their accompanying code and manuscript. Participants will be asked to re-run the analyses reported in the manuscript, as well as go beyond them, by exploring additional questions with the data. This assignment will draw on skills built up during the weekly assignments. The replication project will be brief (6-8 pages). However, please keep in mind that replicating and understanding code written by someone else takes considerable time.

Late submission of assignments: For each day the assignment is turned in late, the grade will be reduced by 10% (e.g. submission two days after the deadline would result in 20% grade deduction).

Attendance: Students are expected to be present and prepared for every class session. Active participation during lectures and seminar discussions is essential. If unavoidable circumstances arise which prevent attendance or preparation, the instructor should be advised by email with as much

advance notice as possible. Please note that students cannot miss more than two out of 12 course sessions. For further information please consult the [Examination Rules](#) §10.

Academic Integrity: The Hertie School is committed to the standards of good academic and ethical conduct. Any violation of these standards shall be subject to disciplinary action. Plagiarism, misuse of AI, free riding in group work, and other deceitful actions are not tolerated. See [Examination Rules](#) §16, the Hertie [Plagiarism Policy](#), and [the Hertie Guidelines for Artificial Intelligence Tools](#).

Compensation for Disadvantages: If a student furnishes evidence that he or she is not able to take an examination as required in whole or in part due to disability or permanent illness, the Examination Committee may upon written request approve learning accommodation(s). In this respect, the submission of adequate certificates may be required. See [Examination Rules](#) §14.

Extenuating circumstances: An extension can be granted due to extenuating circumstances (i.e., for reasons like illness, personal loss or hardship, or caring duties). In such cases, please contact the course instructor and Examination Office *in advance* of the assignment deadline.

4. General Readings

- **Mastering 'Metrics**

Angrist, J. D., & Pischke, J.-S. (2015). *Mastering 'Metrics: The Path from Cause to Effect*. Princeton: Princeton University Press. **Shelfmark: I.04.04_37**

- **The Book of Why**

Pearl, J., & Mackenzie, D. (2018). *The Book of Why: The New Science of Cause and Effect*. New York: Basic Books. **Shelfmark: I.04.01_5+5**

- **The Mixtape**

Cunningham, S. (2021). *Causal Inference: The Mixtape*. New Haven: Yale University Press. Online version available at: <https://mixtape.scunning.com/>

- **Introduction to Econometrics**

Stock, J. H., & Watson, M. W. (2020). *Introduction to Econometrics* (Fourth edition, global edition). London: Pearson. Online version available [here](#).

- **The Effect**

Huntington-Klein, N. (2022). *The Effect: An Introduction to Research Design and Causality*. CRC Press, Taylor & Francis Group. Online version available [here](#).

Articles listed under **Application Reading** serve as examples during the lecture and are also mandatory, but you are not required to read every page closely. It is a good skill to develop to read articles quickly to grasp the key elements: research question, research design, results, and implications.

There are more books out there that people commonly consult to learn about these topics, including:

1. Hernán, M. A., & Robins, J. A. (2020). *Causal Inference: What If*. Boca Raton, FL: CRC Press. Preprint available at: https://www.hsph.harvard.edu/miguel-hernan/wp-content/uploads/sites/1268/2023/05/hernanrobins_WhatIf_14may23.pdf
2. Gelman, A., & Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York: Cambridge University Press. [a reference book for regression modeling and, particularly, mixed-effects models, from a Bayesian perspective] **Shelfmark: I.04.04_26**
3. Enders, C. K. (2022). *Applied Missing Data Analysis*. New York: The Guilford Press. [an accessible introduction into the problems of missing data for our analyses; includes chapters on maximum likelihood estimation] **Shelfmark: I.04.02_208 (2022) or as an e-book available [here](#).**
4. Fox, J. (2016). *Applied Regression Analysis and Generalized Linear Models* (3rd ed.). London: Sage Publications. [theoretical treatment of OLS regression and GLMs, along with selected additional topics, e.g., multilevel models] **Shelfmark: I.04.04_64**
5. Freedman, D. A. (2009). *Statistical Models: Theory and Practice*. Cambridge: Cambridge University Press. **Shelfmark: I.04.02_204 or as an e-book available [here](#).**
6. Wooldridge, J. M. (2020). *Introductory Econometrics: A Modern Approach* (5th ed.). Mason, OH: Cengage Learning. [an econometric classic] **Shelfmark: I.04.04_8 (2020)**

Finally, there are tons of video and audio resources around causal inference online. Here are some that I can recommend:

1. Causal Inference Bootcamp: <https://mattmasten.github.io/bootcamp/>
2. Online Causal Inference Seminar: <https://sites.google.com/view/ocis/>, <https://www.youtube.com/channel/UCiiOj5GSES6uw21kfXnxj3A/videos>
3. Causal Inference Podcast, <https://casualinfer.libsyn.com/>

5. Course Sessions and Readings

All course readings can be accessed on the course Moodle page.

Session 1: Counterfactual Causality

Required Reading	The Book of Why (Pearl, Judea and Dana Mackenzie. (2018). <i>The Book of Why: The New Science of Cause and Effect</i> . New York: Basic Books): Introduction .
------------------	--

Session 2: The Potential Outcomes Framework and Experiments

Required Reading	Mastering Metrics (Angrist, Joshua D. and Jörn-Steffen Pischke. (2015). <i>Mastering 'Metrics: The Path from Cause to Effect</i> . Princeton: Princeton University Press): Chapter 1 .
Application Reading	1. Guess, Andrew M. et al. "How Do Social Media Feed Algorithms Affect Attitudes and Behavior in an Election Campaign?" <i>Science</i> 381(6656): 398-404. 2. Beaman, Lori, Esther Duflo, Rohini Pande, and Petia Topalova. "Female Leadership Raises Aspirations and Educational Attainment for Girls: A Policy Experiment in India." <i>Science</i> 335(6068): 582-586.

Session 3: Causal Graphs

Required Reading	The Mixtape (Cunningham, S. (2021). <i>Causal Inference: The Mixtape</i> . New Haven: Yale University Press): Chapter 3 .
Application Reading	Knox, Dean, Will Lowe, and Jonathan Mummolo. 2020. "Administrative Records Mask Racially Biased Policing." <i>American Political Science Review</i> 114(3): 619-637.

Session 4: Regression Estimators of Causal Effects I: The Standard Linear Model

Required Reading	Mastering 'Metrics: Chapter 2 .
Application Reading	Ziblatt, Daniel, Hanno Hilbig, and Daniel Bischof. 2024. "Wealth of Tongues: Why Peripheral Regions Vote for the Radical Right in Germany." <i>American Political Science Review</i> 118(3): 1480-1496.

Session 5: Regression Estimators of Causal Effects II: Categorical Data and Nonlinearities

Required Reading	Introduction to Econometrics (Stock, J. H., & Watson, M. W. (2020). <i>Introduction to Econometrics</i> (Fourth edition, global edition). London: Pearson): Chapter 8 (8.1 and 8.2 only) and 11.
Application Reading	Emeriau, Mathilde. (2024). "Victim or Threat? Shipwrecks, Terrorist Attacks, and Asylum Decisions in France." <i>American Journal of Political Science</i> 68(4): 1187-1204.

Session 6: Matching

Required Reading	The Mixtape: Chapter 5.
Application Reading	Gilligan, Michael J. and Ernest J. Sergenti. 2008. "Do UN Interventions Cause Peace? Using Matching to Improve Causal Inference." <i>Quarterly Journal of Political Science</i> 3: 89-122.

Session 7: Natural Experiments I: Instrumental Variables

Required Reading	Mastering Metrics: Chapter 3.
Application Reading	1. Miguel, Edward, Shanker Satyanath and Ernest Sergenti. 2004. "Economic Shocks and Civil Conflict: An Instrumental Variables Approach." <i>Journal of Political Economy</i> 112(4): 725-753. 2. Hangartner, Dominik, Elias Dinas, Moritz Marbach, Konstantinos Matakos, and Dimitrios Xefteris. 2019. "Does Exposure to the Refugee Crisis Make Natives More Hostile?" <i>American Political Science Review</i> 113(2): 442-455.

Session 8: Natural Experiments II: Regression Discontinuity

Required Reading	Mastering Metrics: Chapter 4.
Application Reading	1. Zonszein, Stephanie and Guy Grossman. 2025. "Voted In, Standing Out: Public Response to Immigrants' Political Accession." <i>American Journal of Political Science</i> 69(2): 718-733. 2. Dell, Melissa. 2015. "Trafficking Networks and the Mexican Drug War." <i>American Economic Review</i> 105(6): 1738-1779.

Session 9: Difference-in-Differences

Required Reading	Mastering Metrics: Chapter 5.
------------------	--------------------------------------

Application Reading	1. Card, David and Krueger, Alan. (1994). Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. <i>The American Economic Review</i>, 84(4), 772–793. 2. Selb, Peter and Simon Munzert. (2018). "Examining a Most Likely Case for Strong Campaign Effects: Hitler's Speeches and the Rise of the Nazi Party, 1927–1933." <i>American Political Science Review</i> 112(4): 1050–1066.
----------------------------	---

Session 10: Panel Data and Fixed Effects

Required Reading	The Effect (Huntington-Klein, Nick. (2022). <i>The Effect: An Introduction to Research Design and Causality</i> . CRC Press, Taylor & Francis Group): Chapter 16 .
Application Reading	Magazinnik, Asya. (2025). "An Agency Perspective on Immigration Federalism." <i>Journal of Politics</i> .

Session 11: Mediation, Moderation, and Mechanisms

Required Readings	Introduction to Econometrics: Chapter 8 (8.3-8.5). Nguyen, Trang Quynh, Ian Schmid, and Elizabeth A. Stuart. (2021). "Clarifying Causal Mediation Analysis for the Applied Researcher: Defining Effects Based on What We Want to Learn." <i>Psychological Methods</i> 26(2): 255-271.
Application Reading	Hoffmann, Roman, Raya Muttarak, Jonas Peisker, and Piero Stanig. (2022). "Climate Change Experiences Raise Environmental Concerns and Promote Green Voting." <i>Nature Climate Change</i> 12: 148–155.

Session 12: Miscellany and Wrap-Up

Required Readings	Data Visualization (Healy, Kieran. (2019). <i>Data Visualization: A Practical Introduction</i> . Princeton: Princeton University Press): Chapters 1 and 3 . Ali, Arden, Sally Haslanger, Jerome Hodges, and Lily Hu. 2021. "Encoding Race and Philosophical Perspectives."
Suggested Reading	The Visual Display of Quantitative Information (Tufte, Edward R. (1983). <i>The Visual Display of Quantitative Information</i> . Graphics Press.)