Causal Inference

University of Bayreuth - Spring Term 2024

Course Convenor: Ken Stiller (kenneth.stiller@uni-bayreuth.de)

This course introduces causal inference in the social sciences. The objective is to learn how to make and evaluate causal claims. The course will combine a theoretical component and an applied component. The course is intended as an overview of the different methods ranging from experimental designs to quasi-experimental designs. Students will develop the skills to critique methods used in recent academic work and to begin to apply these methods in their own research.
The course will start by introducing causality in social sciences, the potential outcomes framework, experiments, selection on observables/matching, panel data and fixed effects models, difference in differences, instrumental variables, and regression discontinuity design. The topics covered are extensive and I expect this course to be an introduction for students to then explore the different topics/applications on their own. Students on this course are required to be familiar with probability and statistics (OLS, hypothesis testing, logistic regression) but I otherwise assume no knowledge of causal inference. The course is divided in a Lecture component every week, and then a lab component per topic. Please bring your own laptop for the labs. If you do not have a personal laptop, please get in touch with me. Make sure R and R Studio is installed, more on this below.

Books & Readings

The main textbook is Angrist and Pischke (2014), "Mastering Metrics: The Path from Cause to Effect". This is the key introductory text that presents the essential tools of econometric research and demonstrates how to untangle cause and effect in human affairs. Most recently in 2021, one of its authors - Professor Joshua D. Angrist - along with Professor Guido Imbens (whose work we will be dealing with, too), won the Nobel Prize in Economics "for their methodological contributions to the analysis of causal relationships". You are expected to buy or borrow this book and master its contents. If you want to read something more advanced on the same topics, see Angrist and Pischke's 2009 book "Mostly Harmless Econometrics" or the recently released "Causal Inference: The Mixtape" by Scott Cunningham, which is available online. For those interested in the background and the review of statistical concepts we'll be using, Wooldridge' (2016) 'Introductory Econometrics: A Modern Approach' is an excellent companion.
I have listed additional readings for each week to allow students to explore individual topics further as described in other textbooks and articles that implement the discussed research design on interesting research questions. All papers can be found online. A more detailed description of the contents of the course can be found below.

Assessment

This section is relevant for students taking the course for credit. Your final grade will be based on the following:

• Participation during term, during lectures and labs (25% of final mark).

• After term, you will complete a take-home exam (75% of final mark): Distributed by Friday of the week of our last session, due noon on 1 August 2024.

All assignments should be submitted to me via email in PDF format. In the labs you will learn to produce PDF reports using R markdown, which allows you to produce clearly formatted documents with math notations (using Latex), R code, and R output. Please make sure to keep within the word limit.

Collaboration on Assignments

Collaboration on the take-home exam is not allowed, as is discussion of the questions or your responses on online channels. The use of artificial intelligence and natural language processing tools is strictly forbidden. Students will be penalized for violating these rules.

Course Website

Slides, course materials, data, and lab scripts we will be working with can be found on the course website (https://bayreuth-politics.github.io/CI24/).

Lab Sessions

Students will learn to apply the methods we discuss in our weekly lectures. The assignments for this course will be largely based on the implementation of these methods with data as taught in the labs. The software for instruction will be R.

Drop-In Sessions

Instead of office hours, a weekly drop-in session will take place (most likely on Wednesday, tbc). Feel free to pop in to ask any questions on or discuss the lecture materials, lab scripts or assignments. You are also welcome to discuss your dissertation or other projects related to causal inference techniques and data analysis.

About R

Below you can find some useful information about R :

• To download R, go to: https://cran.rstudio.com/

• Many people like R-Studio as a way of managing your work in R. Like R, the basic version of R-Studio is free.

• You can download it here.

• I suggest that you take a look at these websites, where you will find a number of tutorials:
– Try R: http://tryr.codeschool.com/levels/1/challenges/1
– swirl: http://swirlstats.com
– Jared Knowles R bootcamp: https://www.jaredknowles.com/r-bootcamp/

Syllabus

Week 1: Introduction

Brief overview of the course. No readings required.

Week 2 & 3: Causality, Potential Outcomes Framework & Experiments

Discussion of what is causality in the social sciences, introduction to the potential outcomes framework and the ideal experiment. The second lecture will focus on randomization and how to conduct randomized experiments.

Readings:
• Angrist, Joshua D., and Jörn-Steffen Pischke. 2014. Mastering 'Metrics: The Path from Cause to Effect. Princeton University Press: Chapter 1
• Angrist, Joshua and Jörn-Steffen Pischke. 2009. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton: Princeton University Press: Chapter 2
• Cunningham, Scott. (2021). Causal Inference: The Mixtape: Chapter 4

Additional Readings:
• for an overview of causality (philosophically and in the context of the social sciences) see Hidalgo, Daniel and Sekhon, Jashjeet (2011), click here for a copy.
• Gerber and Green (2012), Field Experiments, Ch. 1 & 2.
• Gerber, Alan S., Donald P. Green and Christopher W. Larimer. 2008. "Social Pressure and Voter Turnout: Evidence from a Large Scale Field Experiment." American Political Science Review 102(1): 1-48.

• Hainmueller, Jens, and Michael J. Hiscox. (2010). "Attitudes toward highly skilled and low-skilled immigration: Evidence from a survey experiment." American Political Science Review, 61-84.

Lab Papers:
• Week 2: Lupu, N. (2013). Party brands and partisanship: Theory with evidence from a survey experiment in Argentina. American Journal of Political Science, 57(1), 49-64.
• Week 3: Hyde, S. D. (2007). The observer effect in international politics: Evidence from a natural experiment. World politics, 60(1), 37-63.

Week 4: Selection on Observables, Multiple Regression and Matching

Experiments/randomization is not always practical. This week we discuss designs that assume that selection into the treatment groups is based on observables and how we can use multiple regression to overcome endogeneity. We discuss regression and matching techniques.

Readings:
• Angrist, Joshua D., and Jörn-Steffen Pischke. 2014. Mastering 'Metrics: The Path from Cause to Effect. Princeton University Press: Chapter 2
• Angrist, Joshua and Jörn-Steffen Pischke. 2009. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton: Princeton University Press: Chapter 3. Sections 3.1-3.3. For Matching see 3.3.1-3.3.3

Additional Readings:
• Cunningham, Scott. (2021). Causal Inference: The Mixtape: Chapter 5
• Washington, Ebonya L. (2008). "Female socialization: how daughters affect their legislator fathers." American Economic Review 98.1: 311-32.
• Dale, Stacey B. and Alan B. Krueger. (2002). "Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables," Quarterly Journal of Economics, vol. 117, no. 4, 14911527.

Lab Paper:
• Washington, E. L. (2008). Female socialization: how daughters affect their legislator fathers. American Economic Review, 98(1), 311-32.

Week 5: Panel Data and Fixed Effects Models

This week will discuss what research designs we can use when we move from cross-sectional to panel data. These empirical strategies use data with a time or cohort/spatial dimension to control for unobserved, but time-specific or cohort-specific omitted variables.

Readings:
• Angrist, Joshua and Jörn-Steffen Pischke. 2009. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton: Princeton University Press: Chapter 5
• Cunningham, Scott. (2021). Causal Inference: The Mixtape: Chapter 8

Additional Readings:
• For a more formal rigorous treatment: Wooldridge (2010) Econometric Analysis of Cross Section and Panel Data
• Egorov, Georgy, Sergei Guriev, and Konstantin Sonin. (2009). "Why resource-poor dictators allow freer media: A theory and evidence from panel data." American Political Science Review 103.4: 645-668.

Lab Paper:
• Gomez, B. T., Hansford, T. G., & Krause, G. A. (2007). The Republicans should pray for rain: Weather, turnout, and voting in US presidential elections. The Journal of Politics, 69(3), 649-663.

Week 6: Differences-In-Differences

Confounders cannot always be observed. However, many public policies are adopted at a certain point in time for a selected group which can be used to retrieve causal estimates. When there is data across time for both treatment and control groups, before and after treatment, a difference-in-difference approach can be implemented. This strategy addresses time-invariant confounders.

Readings:
• Angrist, Joshua D., and Jörn-Steffen Pischke. 2014. Mastering 'Metrics: The Path from Cause to Effect. Princeton University Press: Chapter 5
• Angrist, Joshua and Jörn-Steffen Pischke. 2009. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton: Princeton University Press: Chapter 5 - 5.1-5.3
• Cunningham, Scott. (2021). Causal Inference: The Mixtape: Chapter 9

Additional Readings:
• Card, David and Alan B. Krueger. (1994). "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania." American Economic Review 84: 772-793.
• de Chaisemartin, Clement and d'Haultfoeuille, Xavier. (2021). "Two-Way Fixed Effects and Differences-in-Differences with Heterogeneous Treatment Effects: A Survey". Available at SSRN
• Goodman-Bacon, Andrew. (2021). "Difference-in-differences with variation in treatment timing." Journal of Econometrics, Forthcoming
• Foos, Florian and Daniel Bischof. (2021). Tabloid Media Campaigns and Public Opinion: Quasi-Experimental Evidence on Euroscepticism in England. American Political Science Review, 119.doi:10.1017/S000305542100085X.

Lab Paper:
• Dinas, E., Matakos, K., Xefteris, D., & Hangartner, D. (2019). Waking up the golden dawn: does exposure to the refugee crisis increase support for extreme-right parties?. Political analysis, 27(2), 244-254.

Week 7: Instrumental Variables and 2SLS Estimation

If we discover a variable or phenomenon that is exogenously determined yet explains the variation in our explanatory variable of interest, we can use the method of instrumental variables (IV) to

uncover the causal estimate of the explanatory variable on the outcome. A valid instrument induces changes in the explanatory variable but has no independent effect on the dependent variable. We will discuss the assumptions needed for a valid IV and appropriate model specifications.

Readings:
• Angrist, Joshua D., and Jörn-Steffen Pischke. 2014. Mastering 'Metrics: The Path from Cause to Effect. Princeton University Press: Chapter 3
• Angrist, Joshua and Jörn-Steffen Pischke. 2009. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton: Princeton University Press: Chapter 4
• Cunningham, Scott. (2021). Causal Inference: The Mixtape: Chapter 7

Additional Readings:
• Sovey, Allison and Donald Green. (2010) "Instrumental Variables Estimation in Political Science: A Reader's Guide." American Journal of Political Science, 55(1): 188-200.
• Colantone, Italo, and Piero Stanig. (2018) "Global competition and Brexit." American Political Science Review, 112, no. 2: 201-218.

Lab Paper:
• Dinas, E. (2014). Does choice bring loyalty? Electoral participation and the development of party identification. American Journal of Political Science, 58(2), 449-465.

Week 8: Regression Discontinuity Designs

When selection into treatment group changes at an arbitrary cut-off for a continuous variable (i.e. a pass for a test score), a regression discontinuity design (RDD) can be implemented.

Readings:
• Angrist, Joshua D., and Jörn-Steffen Pischke. 2014. Mastering 'Metrics: The Path from Cause to Effect. Princeton University Press: Chapter 4
• Angrist, Joshua and Jörn-Steffen Pischke. 2009. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton: Princeton University Press: Chapter 6

Additional Readings:
• Lee, David S., and Lemiuex, Thomas. "Regression Discontinuity Designs in Economics." Journal of Economic Literature 2010:48: 281-355.
• Cattaneo, Titiunik and Vazquez-Bare (2020): The Regression Discontinuity Design. Handbook of Research Methods in Political Science and International Relations, Sage Publications, Ch. 44, pp. 835-857.
• Dahlgaard, Jens Olav. (2018) "Trickle-up political socialization: The causal effect on turnout of parenting a newly enfranchised voter." American Political Science Review 112, no. 3: 698-705.
• Cattaneo, Idrobo, and Titiunik. A Practical Introduction to Regression Discontinuity Designs: Foundations. Available at Cattaneo's website.
• De La Cuesta, Brandon, and Kosuke Imai. "Misunderstandings About the Regression Discontinuity Design in the Study of Close Elections." Annual Review of Political Science 19.1: 375-96

Lab Paper:

• Meyersson, E. (2014). Islamic Rule and the Empowerment of the Poor and Pious. Econometrica, 82(1), 229-269.

Week 9: Fuzzy Regression Discontinuity Design

Building upon what we discussed in the two previous weeks, we will discuss more advanced topics around instrumental variables and regression discontinuity design. Specifically, we will discuss the fuzzy regression discontinuity design where the cut-off point increases the probability of receiving treatment. Moreover, we will revisit recent advances in applications of instrumental variable design.

Readings:
• Angrist, Joshua D., and Jörn-Steffen Pischke. 2014. Mastering 'Metrics: The Path from Cause to Effect. Princeton University Press: Chapter 4 (Fuzzy RD)
• Angrist, Joshua and Jörn-Steffen Pischke. 2009. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton: Princeton University Press: Section 6.2.

Additional Readings:
• Eggers, Andrew C, and Jens Hainmueller (2009) "MPs for Sale? Returns to Office in Postwar British Politics" American Political Science Review 103.4 : 513-33
• Pons, Vincent, and Clemence Tricaud. (2018) "Expressive voting and its cost: Evidence from runoffs with two or three candidates." Econometrica 86, no. 5: 1621-1649.
• Eggers, Andrew C., Anthony Fowler, Jens Hainmueller, Andrew B. Hall, and James M. Snyder Jr. (2015). "On the validity of the regression discontinuity design for estimating electoral effects: New evidence from over 40,000 close races." American Journal of Political Science 59, no. 1: 259-274.
• Lal, Apoorva, Mackenzie William Lockhart, Yiqing Xu, and Ziwen Zu. (2021) "How Much Should We Trust Instrumental Variable Estimates in Political Science? Practical Advice based on Over 60 Replicated Studies." Working paper

Lab Paper:
• Dinas, E., Riera, P., & Roussias, N. (2015). Staying in the first league: Parliamentary representation and the electoral success of small parties. Political Science Research and Methods, 3(2), 187-204.