Causal Inference and Public Policy in Europe

MA Course, Summer 2020

Basismodul Seminar Politikwissenschaft V (CGS)

Lecture videos will be uploaded every Monday from 13/04 to 25/05
Online consultations for R practice: Tuesdays 12:30 – 13:30, from April 21 to May 26.
Final projects presentation: June 30.

Instructor:
Bruno Castanho Silva, PhD
Cologne Center for Comparative Politics (CCCP)
Email: bcsilva@wiso.uni-koeln.de
Office hours: By appointment.

---

1 How This Course is Structured Online

Usually this course would be structured with weekly 3-hour meetings, in which half the session is a lecture/seminar, and the other half is practice in R. Due to the extraordinary circumstances this term, the weekly schedule will be the following:

- A few short video lectures adding up to 45-60 minutes on the topic of that week, uploaded every Monday;
- A video lecture of the R tutorial, also uploaded every Monday;

- A 1h online R consultation through Zoom, every Tuesday at 12:30 starting from April 21, for the weekly R tutorial. If this slot does not work for you, write me an email and we can talk another time.
- Discussions using the Ilias course forum (see more details below).

2 Course Description

The course focuses on statistical inference for the analysis of public policy across a wide range of contexts. The rise of 'evidence-based policy making' has brought the attention of governments, international organizations, and NGOs to rigorous methods of causal inference for impact evaluation. This course covers issues related to the design, implementation, and evaluation of policy changes. Technical aspects will focus on computational approaches and real-world challenges.

Goals

Students will learn to model cause-and-effect relationships and develop counterfactual scenarios. They will gain experience using computational methods to predict the impacts of policies, interventions, and events, while learning to avoid common pitfalls. By the end, students will be able to: (1) think through which of the methods covered in class (if any) would be best suited to solve a given decision problem and what data would be required; (2) perform appropriate analysis and interpret results; (3) connect those results to strategic decision-making; (4) critically examine statistical causal claims put forward by others; and (5) present findings and recommendations effectively for audiences of varying sophistication.

Prerequisites

Students are expected to be familiar with basic statistical methods for analysis and inference (i.e., run and interpret a linear regression). You should have taken Introduction to Quantitative Analysis with R or a similar course before starting this. Students should also have a basic familiarity with R.

Software

All students should have R installed in their computers before the first class. Students are also strongly encouraged to install RStudio. This is a more user-friendly interface

for R with integration to other packages we will use throughout the course. RStudio is available for free at https://www.rstudio.com/.

3 Course Requirements

Students will be assessed based on the following exercises (all are mandatory to pass the course):

- **Two reaction papers (10 points each).** After each session, from weeks 2–7, there is an extra reading related to that topic. During the semester, each student should pick any two sessions/topics and write a reaction paper for each. The paper should be a critical evaluation of that reading, based on the lecture and other readings, focusing on questions such as: are the assumptions met? Is the method properly used, and the right one to test that that theory? Are the results correctly interpreted? Are the conclusions valid, given the analyses? What parts of the analysis could be improved? The deadline for uploading the reaction paper is the Monday after the respective video for the topic was uploaded, so you have a week to complete the assignment. The reaction paper should be **no longer than one page**.

- **One of two take-home assignments (20 points).** Two take-home exercises will be posted during the semester, on **May 06** and **May 20** with two weeks to complete each of them – deadlines on **May 20, at 23:55 CET**, and **June 03, 23:55 CET**. Students will be given a dataset and asked to perform analyses in accordance to methods covered in the class thus far. The length of assignments should not exceed **five pages**. **Each student should do only one of the two – pick your favorite.**

- **Final course project (60 points).** For the final project, students should find data and use one of the methods discussed in class to analyze it, in order to answer a policy relevant causal question. Results should be presented as an **infographic**, targeted at a lay audience (and not academic), with an accompanying 2-page technical memo on the data and analyses performed. The infographics should be uploaded by **June 30**, and we will see about the possibility of having a live online seminar for presentations on that day.

Points are converted to final grades as follows:

Points	Grade
100–95	1,0
94.5–90	1,3
89.5–85	1,7
84.5–80	2,0
79.5–75	2,3
74.5–70	2,7
69.5–65	3,0
64.5–60	3,3
59.5–55	3,7
54.5–50	4,0
49–0	5,0

Bonus Points

It is possible to earn up to 10 bonus points in this course, meaning the maximum final grade can be 110. After each session, I will post a question on the Ilias forum for discussion, related to the weekly class. Those who want bonus points are encouraged to take part in the discussion, answering the questions and reacting to their colleagues' answers. You can get up to two bonus point each week, depending on your participation on the discussion – you will be evaluated based on the quality of your answer/arguments, and your interaction with your peers. *Students can only earn bonus points if they fulfill all the mandatory course requirements (meaning, you need to reach at least 50 points with the mandatory assignments alone.)* Bonus points can increase your grade, but never decrease it.

3.1 A Note on Professional Presentation

For the one take-home assignment, I recommend you use RMarkdown. RStudio comes with a powerful authoring format called R Markdown. R Markdown documents look like a mix of a text document and R code. They enable easy creation of data analysis reports directly from R. Rather than copying and pasting into Word, your report is created automatically. R Markdown combines the core syntax of markdown (an easy-to-write plain text format) with embedded R code chunks that are run so their output can be included in the final document. R Markdown documents are fully reproducible (they can be automatically regenerated whenever underlying R code or data changes). Markdown is simple to use as it enables the use of a syntax

like plain-text.

1. You need to install LaTeX on your machine. This is a free typesetting software which R Markdown uses. Mac users should install TeXLive, freely available at https://www.tug.org/texlive/; Windows users should install `tinytex`, which can be done directly within RStudio, with the following three commands:
    - install.packages('tinytex')
    - library(tinytex)
    - install_tinytex()
2. To use R Markdown, simply create a new R Markdown document in RStudio. This will load a sample document. Select "Knit PDF" to produce a PDF output file with the write up and the code output.
3. More information on the R Markdown syntax is available here: http://rmarkdown.rstudio.com/.

NB! Installing and running R and RMarkdown can be tricky. If you are having trouble, please contact the instructor in advance.

4 Schedule

Week 1 (Apr 14): The Potential Outcomes framework and the experimental 'gold standard'

* This week exceptionally we do not have an R consultation session on Tuesday.

Mandatory readings:

Imbens, Guido W., and Donald B. Rubin. 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences: an Introduction.* Cambridge: Cambridge University Press, Chapter 1.

Altmann, Steffen, Armin Falk, Simon Jäger, and Florian Zimmermann. 2018. "Learning about job search: A field experiment with job seekers in Germany". *Journal of Public Economics* 164: 33–49.

Week 2 (Apr 21): What to do when we can't experiment? Introduction to matching

Mandatory readings:

Ho, Daniel E., Kosuke Imai, Gary King, and Elizabeth A. Stuart. 2007. "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference", *Political Analysis* 15(3): 199–236.

Guill, Karin, Oliver Lüdtke, and Olaf Köller. 2017. "Academic tracking is related to gains in students' intelligence over four years: Evidence from a propensity score matching study" *Learning and Instruction* 47: 43–52.

Week 3 (Apr 28): How to find better matches

Mandatory readings:

Diamond, Alexis, and Jasjeet S. Sekhon. 2013. "Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies", *The Review of Economics and Statistics* 95(3): 932–945.

Wood, Reed M., and Christopher Sullivan. 2015. "Doing Harm by Doing Good? The Negative Externalities of Humanitarian Aid Provision during Civil Conflict", *The Journal of Politics* 77(3): 736–748.

Week 4 (May 05): Regression Discontinuity

Mandatory readings:

Lee, David S. 2008. "Randomized Experiments from Non-Random Selection in U.S. House Elections." *Journal of Econometrics* 142(2): 675–697.

Pinotti, Paolo. 2017. "Clicking on Heaven's Door: The Effect of Immigrant Legalization on Crime." *American Economic Review* 107(1): 138–168.

Cavaille, Charlotte, and John Marshall. 2018. "Education and Anti-Immigration Attitudes: Evidence from Compulsory Schooling Reforms across Western Europe." *American Political Science Review*, 1–10.

Week 5 (May 12): Before-after comparisons: Differences-in-Differences and Introduction to Synthetic Controls

Mandatory readings:

Angrist, Joshua B., and Jörn-Steffen Pischke. 2008. *Mostly Harmless Econometrics.* Princeton University Press: pp. 169–174.

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2015. "Comparative Politics and the Synthetic Control Method." *American Journal of Political Science* 59(2): 495–510.

Last updated: April 8, 2020

Podestà, Federico. 2017. "The impact of 'free choice': Family reforms of France and Belgium, a synthetic control analysis." *International Journal of Social Welfare*, pp. 1–13.

Week 6 (May 19): Synthetic Controls II – Multiple treated units at many points in time

Mandatory readings:

Kreif, Noémi, Richard Grieve, Dominik Hangartner, Alex James Turner, Silviya Nikolova, and Matt Sutton. 2016. "Examination of the Synthetic Control Method for Evaluating Health Policies with Multiple Treated Units." *Health Economics* 25(12): 1514–1528.

Cavallo, Eduardo, Sebastian Galiani, Ilan Noy, and Juan Pantano. 2013. "Catastrophic Natural Disasters and Economic Growth." *The Review of Economics and Statistics* 95(5): 1549–1561.

Week 7 (May 26): Instrumental Variables and Natural Experiments

Mandatory readings:

Sovey, Allison J., and Donald P. Green. 2011. "Instrumental Variables Estimation in Political Science: A Readers' Guide." *American Journal of Political Science* 55(1): 188–200

Barone, Guglielmo, and Gaia Narciso. 2015. "Organized crime and business subsidies: Where does the money go?." *Journal of Urban Economics* 86: 98–110.

Becker, Sascha O., Katrin Boeckh, Christa Hainz, and Ludger Woessmann. 2016. "The Empire is Dead, Long Live the Empire! Long-Run Persistence of Trust and Corruption in the Bureaucracy." *The Economic Journal* 126(590): 40–74.

Week 8 (June 30): Final Course Project Presentation