

Syllabus - Quantitative Methods III: Explanation and Causation

FLS6441: Métodos Quantitativos III: Explicação e Causação

October 2019

First Semester 2020, DCP, FFLCH, USP

Thursdays, 14:00 - 17:30

Jonathan Phillips

1 Objectives

This course is the third course in a three course sequence, following Quantitative Methods I (Introductory Statistics) and Quantitative Methods II (Multivariate Regression). This is an applied course in how to *explain* the outcomes – public policies, democracy, peace, governance, corruption, inequality – we study in political science. It provides students with the skills, tools and confidence to implement a wide range of analyses commonly used in leading political science studies. By the end of this course you will be able to:

1. Describe research questions in terms of counterfactual observations;
2. Critique the assumptions, accuracy and generalizability of *any* quantitative research design;
3. Replicate the core analysis of papers published in major political science journals;
4. Apply (in R or Stata) empirical analyses using experimental, instrumental variables, difference-in-differences, regression discontinuity, matching, comparative case study and process tracing methods.
5. Design compelling research projects that effectively answer important political science questions;

Lots of these words are probably unfamiliar and intimidating. Don't worry - the aim of this course is to overcome these barriers and give students the intuition, experience and confidence to regularly use these tools. You will get lots of hands-on practice in how to replicate major studies using their original datasets, so the techniques will make much more sense and you will understand what they can and cannot help us explain.

More practically, this course is excellent preparation for students who want to understand and critique papers, develop their own research designs for dissertations, or be able to engage more in sandwich programs or at conferences.

2 Justification

All good political science research projects are causal - they do not just describe the world but seek to *explain* it: Do women govern better than men? Can corrupt politicians influence local media to get re-elected? How do authoritarian regimes survive during economic crises?

But explanation is hard because societies are extremely complex and lots of processes are happening all at the same time. We simply do not - cannot - know what the alternative outcome would have been if Angela Merkel had lost at the last election or if Hilary Clinton had won.

Even if we throw huge amounts of data and lots of control variables into a regression we would not be able to explain these outcomes convincingly. A range of problems prevent us saying that X causes Y - an omitted variable, some reverse causation, or a selection bias. Regression on its own cannot solve any of these problems. Instead, we need an alternative - smarter - strategy, which carefully selects the data we are comparing and identifies the most likely *counterfactual* observations.

For example, instead of comparing all women and men in politics we could focus on India where some government positions are randomly-allocated by gender; instead of asking all voters what they think of

corrupt politicians we could ask only those who live in places which are on the border of receiving local radio stations those politicians control; instead of comparing all authoritarian regimes we could focus on a paired comparison between two neighbouring countries which share similar histories, economies and institutions.

The essential toolkit for empirical research in political science therefore includes being able to describe and identify appropriate counterfactuals, choose appropriate research designs to maximize causal inference, apply these research designs to real data, and provide compelling *explanations* for the patterns in the data.

3 Prerequisites

Students are required to have completed FLS 6183, Quantitative Methods II, or equivalent.

4 Content

Part I focuses on reviewing the core tools and ideas for understanding the variation in our data.

1. Review: What regression does and does not do. (05/03/20)
2. A Framework for Explanation (12/03/20)

Part II practices implementing key causal methodologies. Each week we discuss the methodology, read a high-profile paper that uses the methodology, and replicate the analysis ourselves using the raw data in R or Stata.

3. Field Experiments (19/03/20)
4. Survey and Lab Experiments (26/03/20)
5. Natural Experiments (02/04/20)
6. Instrumental Variables (16/04/20)
7. Discontinuities (23/04/20)
8. Difference-in-Differences (30/04/20)
9. Controlling for Confounding (07/05/20)
10. Matching (14/05/20)
11. Comparative Cases and Process Tracing (21/05/20)
12. Generalizability, Reproducibility and Mechanisms (28/05/20)

5 Evaluation

1. **Replication Tasks (40%)** - Each week I will provide you with a pre-prepared dataset from a well-known political science paper that uses the methodology we are studying that week.
 - Rather than replicate the full paper, I will provide specific questions for you to answer to structure and guide you through the replication.
 - You will be graded on the understanding you show in your responses to these questions and in implementing the core methodology, and not on how precisely you replicate the final result. I.e. it doesn't matter at all if your standard errors are slightly different.
 - I will provide clear instructions and examples of how to apply the methodologies in class so this will be accessible for all students.
 - Students are required to submit at least 7 out of 9 replication tasks (your grade will be based on your 7 best submissions).
2. **Short Research Paper (40%)** - To apply your new skills you will pick a research question of interest to you (possibly your thesis/dissertation) and suggest a research design that would be effective at answering that question.
 - This should be a question for which data is available and you can provide a first attempt at applying the methodology - the results themselves don't matter.
 - Papers should be no more than 15 pages, and can be in English or Portuguese.
 - The deadline for submission will be 24th July 2020.

3. **Participation (20%)** - Presence and constructive contribution in class.

6 Bibliography

The list below is a general list of reference texts that we will use repeatedly throughout the course. For a detailed breakdown of readings for each class, please see the course website.

- Joshua D Angrist and Steve Pischke. Mastering 'Metrics. 2015.
- Guido W. Imbens and Donald B Rubin. Causal Inference for Statistics, Social, and Biomedical Sciences. 2015.
- Thad Dunning. Natural Experiments in the Social Sciences: A Design-Based Approach. Yale University Press, 2012.
- Joshua D Angrist and Steve Pischke. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press, 2009.
- Stephen L Morgan and Christopher Winship. Counterfactuals and Causal Inference: Methods and Principles for Social Research. Cambridge University Press, 2007.