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.

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.

Prerequisites

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

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. (21/03/19)
2. Causation, potential outcomes, counterfactuals and the fundamental problem of causal inference. (28/03/19)

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 (04/04/19)
4. Survey and Lab Experiments (11/04/19)
5. Randomized Natural Experiments (18/04/19)
6. Instrumental Variables (25/04/19)
7. Discontinuities (02/05/19)
8. Difference-in-Differences (09/05/19)
9. Controlling for Confounding (16/05/19)
10. Matching (23/05/19)
11. Comparative Cases (30/05/19)
12. Process Tracing (06/06/19)

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, and you will replicate the core analysis of the paper in R or Stata.
   - I will provide clear instructions and examples of how to do this in class so this will be accessible for all students.
   - Note that the emphasis is not on using precisely the same specification or reproducing identical standard errors, but implementing the core methodology to understand what is happening.
   - You should submit two files each week - the analysis code and a short write-up (in English, for practice!) containing your final graphs/tables and interpreting and explaining your results (no more than 400 words written). You are encouraged to use R markdown to simplify this task, but it is not required.
   - Students are required to submit at least 8 out of 10 replication tasks (your grade will be based on your 8 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.
- The deadline for submission will be three weeks after the last class of the semester, and can be in English or Portuguese.

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

**Bibliography**

- Mandatory reading is marked in *italics*.
- The paper we will replicate is marked with a (R*), though some of these are still provisional.

**Week 1 - Concepts**

[No readings]

**Week 2 - Concepts**

*Guido W. Imbens and Donald B Rubin. Causal Inference for Statistics, Social, and Biomedical Sciences. 2015. Chs. 1 & 3*

**Week 3 - Field Experiments**


**Week 4 - Survey and Lab Experiments**

*Steven D Levitt and John A List. What Do Laboratory Experiments Tell Us About the Real World? 2006.*


**Week 5 - Natural Experiments**


**Week 6 - Instrumental Variables**


Week 7 - Discontinuities


Week 8 - Difference-in-Differences


Week 9 - Controlling for Counterfactuals


Week 10 - Matching


Week 11 - Comparative Cases


Week 12 - Process Tracing
