Fall 2015 | Thu., 2-4pm | Classroom: CGIS K354

Gov 2002: Causal Inference

Matthew Blackwell
Office: CGIS K305
Office Hours: just drop by
mblackwell@gov.harvard.edu
http://www.mattblackwell.org

TF: Stephen Pettigrew
Office Hours: Wednesdays 10am–12pm (CGIS-K Cafe)
pettigrew@fas.harvard.edu

General Information

This course will cover a growing field in political science and the social sciences more generally: causal inference. That is, we will primarily be concerned with how and when we can make causal claims from empirical research. This is a difficult task and, in general, there are no formulas or computer programs that can give your research a causal interpretation. We always require assumptions. This course will focus on providing you the analytic and quantitative framework to assess and implement these assumptions. Perhaps at the most basic level, this course will sharpen your senses toward the issues of causal reasoning in the social sciences.

The structure of the class will mix lectures, discussions, and computer work. You’ll learn these technique by actually implementing them. Having said this, this will not be a computing class. After this class you should understand and be able to apply the standard set of causal inference tools in the social sciences. These include randomized experiments, matching, weighting, fixed effects, differences-in-differences, instrumental variables, regression discontinuity designs, marginal structural models, and sensitivity analyses.

Qualified undergraduates or students from other departments are welcome to take this course.
Who should and should not take this class?

Students in the class should plan to work hard at understanding some difficult material. You'll need to have some proficiency with R. You should have taken GOV 2000 and GOV 2001 or the equivalent. Of course there are always exceptions to this, so feel free to chat with me about your background to see if the fit is right.

Course Details

Reading

There are readings for each topic and they mostly cover the theory of the method along with some applications. Obviously, read the required readings and any others that pique your curiosity. In addition, though, engage with the readings: take notes, write down your impressions or confusions, talk with your classmates. All of your classes should be pushing your research forward and you will be more creative the more you actively read.

Homeworks

Methods are tools and it isn’t very instructive to read a lot about hammers or watch someone else wield a hammer. You need to get your hands on a hammer or two. Thus, in this course, you will have homeworks on a (roughly) biweekly basis. They will be a mix of analytic problems, computer simulations, and data analysis. These homeworks should be typed and well-formatted. We will grade on a (+, √, −) basis. I encourage you to work in groups on the homework, but you always need to write your own solutions including your computer code. Also, it is hugely beneficial to attempt the problems sets on your own before working in groups.

Student Project

In lieu of a final exam, this course requires students to write a short empirical paper on a research topic of their choice. This paper should apply at least some of the methods in the course to an empirical problem. It should be 5-15 pages and focus on the research design, data, methodology, results, and analysis. Literature reviews or background material should be omitted or included in an appendix. Co-authored projects are strongly encouraged: working with collaborators will be the cornerstone of your career from now on. It’s crucial to get to know this process sooner rather than later. Of course, with a co-authored project, we will have higher expectations on the paper and presentation.

Here is a brief timetable for the projects:

- **October 30th**: Email me and Stephen a short (half-page) description of your pro-
posed project. You should meet with me and Stephen in office hours to discuss your project, especially if you are interested in applying methods that we cover later in the course (which are great!).

- **November 19th:** Email me and Stephen a 2-3 page description of your progress, including tables, figures, preliminary results, and analysis.

- **December 10th:** Extra class meeting where each student will give a short, 5-10 minute presentation on their project with a brief summary of the data, methods, and results.

- **December 11th:** Email me and Stephen your final version of the project.

**Books**

The following texts are **required** for this course:


- Hernán, Miguel A. and James M. Robins. 2012. *Causal Inference*. Forthcoming, Cambridge University Press. (Note that this book is still being written and you can find draft PDFs on the linked page above.)

**Optional Books**

- **The following books are optional but may prove useful for additional coverage of some of the course topics.**

- **Reference Book for Panel Methods**

- **Causal Inference**
Computing

We’ll use R in this class, which you can download for free at http://www.r-project.org. R is open source and available on all major platforms (including Solaris, so no excuses). You can find a virtually endless set of resources for R on the internet, including this Getting Started With R page. You may also be interested in using RStudio, an editor and development environment for R. If you are completely new to R, you should complete this online short course, Try R.

Grading

- biweekly homework assignments (50% of final grade)
- student project (40% of final grade)
- participation and presentation (10% of final grade).

Preliminary Schedule

The following is a preliminary schedule of course topics. We may adjust the schedule due to time or interest. Required readings are marked with a (⋆). Note that there is no class on March 14th due to Spring break.

§1 Introduction

- Overview, Course Requirements, Course Outline
§2  Review of Statistical Concepts Useful for Causal Inference
   · Random Variables, Measures of Location and Dispersion
   · Inference and Properties of Estimators
   · Conditional mean function

§3  The Potential Outcome Model
   · Counterfactual Responses and the Fundamental Identification Problem
   · Graphical Causal Models
   · Estimands and Assignment Mechanisms

Readings
   · Hernán and Robins, Chapter 1, 6. (⋆)
   · Angrist and Pischke: Chapter 1. (⋆)

§4  Randomized Experiments
   · Identification of Causal Effects under Randomization
   · Implementation, Estimation, Diagnostics, Blocking
   · Effect Modification

Readings: Theory of Experiments
   · Angrist and Pischke: Chapter 2. (⋆)
   · Hernán and Robins, Chapter 2.1-2.3, 4.1-4.3. (⋆)
   · Imbens and Rubin, Chapter 4.

Readings: Application of Experiments


Readings: Application of Natural Experiments


Hyde, Susan D. 2007. The Observer Effect in International Politics: Evidence from a Natural Experiment. World Politics 60(1): 37-63. (⋆)


Readings: Experiments Review Articles


· List, John A., and Steven Levitt. 2006. *What Do Laboratory Experiments Tell Us About the Real World?* University of Chicago and NBER.


Readings: Useful Methodological Guides for Experiments


· Harvard University Committee on the Use of Human Subjects http://cuhs.harvard.edu
§5 \textbf{Randomization Inference}

- Fisher’s approach to inference, permutation tests, Lady Tasting Tea.
- Sharp null of no effect, randomization distribution

\textit{Readings: Fisher’s randomization inference}

- Bower, Jake and Panagopoulos, Costas. 2011. \textit{Fisher’s randomization mode of inference, then and now}. (⋆)


- Imbens and Rubin, Chapter 5.


§6 \textbf{Causal Effects under Selection on Observables}

\textbf{Selection on Observables}

- Identification under Selection on Observables, Back-door criterion
- Subclassification

\textit{Readings}

- Hernán and Robins, Chapter 7. (⋆)


**Matching Methods**

- Covariate Matching, Balance Checks, Properties of Matching Estimators

**Readings: Matching Theory**


- Rubin. 2006. *Chapters 3 to 5*.


**Readings: Matching Applications**


### Propensity Score Methods

- Identification, Propensity Score Estimation, Matching on the Propensity Score, Weighting on the Propensity Score, Reweighting methods

*Readings: Propensity Score Methods Theory*

- Morgan and Winship: Chapter 3.

*Readings: Propensity Score Methods Applications*


### Regression

- Non-parametric Regression, Identification with Regression

*Readings*

- Angrist and Pischke: Chapter 3. (⋆)
- Morgan and Winship: Chapter 5. (⋆)
- Chapter in Winship and Morgan on Matching vs Regression.


**Conclusion: Selection on Observables**

- Can Non-Experimental Method Recover Causal Effects?

**Readings: Comparison of Experimental and Non-experimental Methods**


**§7 Causal Effects under Selection on Time-Invariant Characteristics**

**Panel Data Methods**

- Fixed Effects and Random Effects Estimation
Readings: Panel Methods Theory

- Angrist and Pischke: Chapter 5.1 (⋆)
- Angrist and Pischke: Chapter 8 (⋆)

Readings: Panel Methods Applications


Difference-in-Differences Estimators

- Identification, Estimation, Falsification tests

Readings: DID Theory

- Angrist and Pischke: Chapter 5.2-5.4 (⋆)

Readings: DID Applications


§8 Causal Effects under Selection on Time-variant Characteristics

Instrumental Variables

- Identification: Using Exogenous Variation in Treatment Intake Given by Instruments
- Imperfect Compliance in Randomized Studies
- Wald Estimator, Local Average Treatment Effects, 2SLS

Readings: Instrument Variable Theory

- Angrist and Pischke: Chapter 4 (⋆)
- Morgan and Winship: Chapter 7

Readings: Instrument Variable Critique

Readings: Instrument Variable Applications


- Angrist and Krueger. 2001 Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments


The Regression Discontinuity Design

- Sharp and Fuzzy Designs, Identification, Estimation, Falsification Checks

Readings: RDD Theory


- Angrist and Pischke: Chapter 6 (⋆)


Readings: RDD Applications


§9 Causal Mediation

· Interactions

· Direct, Indirect Effects, Mechanisms/Mediation

*Readings*

· Hernán and Robins, Chapter 5.1-5.3. (⋆)


§10 Causal Inference with Time-varying Treatments

· Time-varying confounders, sequential ignorability, marginal structural models.

*Readings*


## §11 Odds, Ends, and Practicals

### Sensitivity Analysis

- Nonparametric Bounds
- Formal sensitivity tests

**Readings**

- Morgan and Winship: Chapter 6 (⋆)

### Synthetic Control Methods

**Readings**


Quantile Regression

Readings

- Angrist and Pischke: Chapter 7 (⋆)

Distributional Effects in Difference-in-Differences

Readings