Course Description
This course is a Ph.D. seminar focused on analyzing, designing, and conducting empirical causal inference using quantitative methods. Inquiries into cause and effect lie at the heart of almost all social and behavioral research. For example: What are the effects of media exposure on children’s educational attainment? Which kinds of outreach and incentives promote participation in voluntary organizations? Can social pressure increase political engagement? Answering these questions requires methods of analysis that identify the effects of particular causes.

Learning Objectives and Course Design
By the end of the quarter, students should be able to: (1) discuss and compare multiple research methods of causal inference; (2) assess causal identification strategies and claims in empirical research; (3) design research to address causal theories and questions; (4) manipulate and analyze empirical data sets to identify, estimate, and interpret causal effects; (5) identify and address threats to the validity of findings; and (6) apply methods of causal inference to empirical questions and data beyond those covered in course readings and assignments.

The course consists of two “units.” The first unit focuses on the fundamentals of causal inference. Then, in the second unit, we will survey a variety of specific methodological approaches, such as field experiments, natural experiments, regression discontinuity designs, difference in differences, matching, and instrumental variable estimation. For each of these approaches, I recommend you read at least one text providing an instructional overview of the method as well as at least one empirical application of the method.

Every week, I will assign problem sets and in-class activities aimed at helping you cultivate practical experience applying the methods in question. In general, the problem sets will lag a week behind the readings so that you can get a feel for each approach before you apply it.

During the course, you will also develop two larger-scale projects. The first of these is a class exercise in which you will work (perhaps in a small team) over several weeks to plan, design, execute,
analyze, and report the results of a field experiment. The second is a planning document for an empirical research project that you plan to execute (or perhaps have already begun executing). The first assignment is intended to help you cultivate experience with the craft of designing, conducting, and writing up an empirical study to address a causal question. We will workshop these group projects throughout the quarter. The second assignment is intended to provide an opportunity to apply your skills and experiences from the class to the design of an independent project that directly advances your research interests. I especially encourage you to use this second assignment to pursue something bold and creative.

Prerequisites

There are no formal prerequisites for this course. However, you should be familiar with (i.e. have taken at least one graduate-level class on) social scientific research methods and should be comfortable with (i.e. have taken at least one graduate-level class on) quantitative analysis techniques including multivariate regression. Some experience conducting empirical research will be useful, but is not strictly necessary. Knowledge of some statistical software and/or programming will also prove very helpful, but is not required. If you have questions or concerns, please email me.

Requirements

The day-to-day course requirements consist in performing standard graduate seminar activities in a standard timely fashion: attending and participating in seminar meetings, as well as completing all assignments. I believe it is crucial for all participants of the seminar to do all of these things. If you believe you require some sort of exemption or exception from any of these activities, please negotiate with me before the end of the first week of the quarter.

Course Website and Materials

Most materials for the class (including a “live” version of the course schedule, readings, problem sets, datasets, and solution code) will be made available through the course website.

http://aaronshaw.org/teaching/2015/causal

Once the quarter begins, I will not issue revised PDF versions of the syllabus, so check the website! I may distribute some readings and materials via Canvas.

Assignments

There are four kinds of assignments for this class: readings, problem sets, your field experiment, and your research planning document. Along the way, I will also ask you to submit incremental portions of the two larger assignments (the field experiment and the planning document).

Any assignments you turn in must be submitted to me via Canvas no less than one day (24 hours) before the class meeting for that week. I strongly prefer PDFs and plain-text formats because they don’t require special software. For problem sets, please include your (well-commented) code as well as the final output or answers. I will evaluate problem sets at random throughout the quarter.
Mostly, you are on your own recognizance as far as checking your solutions and following up when you do not understand any content.

Problem sets will encourage conceptual understanding, provide experience applying techniques, elicit analytical interpretation, and (to much a lesser extent) facilitate computational skill-building. I strongly encourage you to complete the problem sets in a small group (ideally no more than 2-3 people). If you work in a group, I also recommend that you (1) try to complete the work on your own first; (2) submit your work individually and acknowledge your collaborators by name.

Readings

During the first few weeks of the course we will use the following as a textbook:


This book provides a clear, practically-oriented introduction to the potential outcomes framework for causal inference focused on field experiments. I recommend that you get a copy.

For all subsequent topics I will assign at least one instructional reading along with at least one empirical reading. I assign several instructional readings about natural and quasi-experimental methods from:


These authors emphasize conceptual understanding and examples over formal math. If their approach works for you (and it may not!), get a copy of this book too. If Murnane & Willett does not work, I still recommend that you acquire a “textbook” treatment that covers all of the methods we talk about in a consistent way. All of the following are widely used:


Depending on your training and aspirations, you may prefer these other approaches and you should feel free to substitute corresponding selections for instructional readings throughout the latter half of the quarter.
For each topic, I have also listed additional, “optional” readings that you may find useful. If you have trouble locating any of the readings, please let me know.

Last, if you want to revisit some of the statistical foundations of the material covered in this course (e.g. Regression, Probability, Linear Algebra), I recommend the following textbooks:


**Statistical Computing**

You will need to use statistical computing software in this class. You are welcome to use any statistical software you like (Stata, SPSS, Python, etc.); however, I will use (heavily commented) R for all in-class examples and solutions to the problem sets.

I do all of my statistical computing and analysis in R (http://r-project.org), which is free software available for use on any operating system and extremely powerful for just about any kind of statistical application. If you have not yet invested heavily in another statistical computing language/environment (and sometimes even if you have) I highly recommend learning R and using it to complete assignments for this class. I'll say more during the first class session.

**Major Projects**

You are required to complete two major projects in this course: (1) a field experiment and (2) a final project consisting of a detailed research planning document for another project. The field experiment provides an opportunity to design, plan, execute, and write-up a research project estimating a causal effect. The final project provides an opportunity to apply the course material to a research design that will lead to a major research product (e.g. journal/conference submission, qualifying paper, dissertation chapter, replication study, research proposal). I will provide additional details about my requirements and expectations for each project early in the quarter.

Field experiments are due via Canvas by May 20, 2015 at 5pm, CDT.

Final projects are due via Canvas by June 9, 2015 at 5pm, CDT.

**Evaluation**

Your final grade for the course will be based on my evaluation on your engagement with the material in class (25%) as well as the successful and timely completion of the field experiment project (25%), final project (25%), and all other assignments throughout the quarter (25%).
Course Policies

Confidentiality of Peers’ Work
You will be receiving, reading and commenting on classmates’ writing. These writing assignments are for class use only. You may not share them with anybody outside of class without explicit written permission from the document’s author pertaining to the specific piece.

Confidentiality of In-Class Discussions
It is essential to the success of this class that all participants feel comfortable sharing questions, ideas, fears, reservations, apprehensions, and confusions about works-in-progress, writing, the research process, and related experiences during discussions. Therefore, you may not create any audio or video recordings during class time nor share verbatim comments with those not in class nor are you allowed to share using other methods – e.g., social media – comments linked to people’s identities unless you get a person’s permission.

Academic Integrity
You are responsible for reading and abiding by the Northwestern University Principles Regarding Academic Integrity as well as the applicable school-specific academic integrity policies. The bottom line: make sure to document all of your work and acknowledge the ideas and work of others. When in doubt, err on the side of giving more credit to the original source rather than less. The sanctions for violations of these principles are severe. Feel free to ask me (the instructor) for clarification about related matters.

Deadlines, Absences, etc.
If something causes you to miss a deadline or a class, please contact me. If you request, and obtain, an incomplete for the course and/or an extension on the final project (note: I strongly discourage this!), please allow at least 1 month (4 weeks) after you submit your completed work for me to submit a grade. Keep this in mind if you will need the grade in order to receive your fellowship/diploma/visa/etc. by a particular date.

Students with Disabilities
Any student requesting accommodations related to a disability or other condition is required to register with AccessibleNU (accessiblenu@northwestern.edu; 847-467-5530) and provide professors with an accommodation notification from AccessibleNU, preferably within the first two weeks of class. All information will remain confidential.

Sexual Harassment
All participants in this class are bound by the Northwestern University sexual harassment policies. Please note that the core of the policy states, "no member of the Northwestern community may sexually harass any other member of the community." Review the policy and speak to me (or some other appropriate person) if you have any questions or concerns.
Course Outline
As described above, all readings and assigned tasks are due prior to the class alongside which they are listed. Items listed in the “Additional Readings” subsection for each week are not required. I may adjust the list of readings or the schedule as needed throughout the quarter, so please consult the online version of the schedule for the most up-to-date information.

April 1 – Introductions & Causality
Readings
Gerber and Green – Chapter 1.

Tasks
Describe three research questions addressing some causal relationship relevant to your research interests. Write the research questions down for the first class meeting and be prepared to discuss at least one of them.

Make sure you have a satisfactory setup for statistical computing. If you are working with R, Download & install it. You will also want to pick and install a development environment for R: I recommend Rstudio unless you use Emacs already, in which case Emacs Speaks Statistics (ESS) may be more appealing.

Learn (or review) statistical computing basics: R users: if you’re completely new, take the Try R course. Also, this Getting Started with R page from York University contains a variety of tutorials, tips, and instructional resources. You may also want to check out the startR resources page that I created for this course.

Optional: Attend an “Intro to R” session that I will lead on Friday, April 3, 2:00-4:00pm in Frances Searle 2-378. We’ll cover basic data input and manipulation; some operators and mathematical functions; as well as writing loops and simple functions of your own.

Additional Readings

Murnane and Willett – Chapters 1-2, 13.
Morgan and Winship – Chapters 1-2.
Watch Ned Hall and L.A. Paul discuss causation on Philosophy TV.
April 8 – Causal Inference & Experimentation

Readings

Gerber and Green – Chapter 2; Appendix B.

Tasks

Complete Problem Set 1.

Complete (if you have not done so already) the required CITI training for human subjects research through the Northwestern Institutional Review Board (IRB).

Additional Readings


April 15 – Sampling Distributions, Inference, & Hypothesis Testing

Readings

Gerber and Green – Chapter 3.


Tasks

Complete Problem Set 2.

Complete Field Experiment Assignment 1.

Additional Readings


April 22 – Covariates in Experimental Design & Analysis

**Readings**

Gerber and Green – Chapter 4 (& skim Chapter 9).


**Tasks**

Complete Problem Set 3.

**Additional Readings**


April 29 – Validity: Planning, Executing, & Archiving Research Effectively

Readings


Gelman, A. and Loken, E. (2013). The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time. [Available via Gelman's website]

Tasks

Complete Problem Set 4.

Complete Mid-quarter course assessment (distributed in class).

Additional Readings


Gerber and Green. Chapter 11.

May 6 – Natural Experiments

Readings

Murnane and Willett – Chapter 8.

**Tasks**

Complete Problem Set 5.

Complete Field Experiment Assignment 2 (Submit Planning Document for your field experiment).

**Additional Readings**


**May 13 – Regression Discontinuity & Difference-in-Differences**

**Readings**

Murnane and Willett – Chapter 9.


**Tasks**

Complete Problem Set 5.
**Additional Readings**

Morgan and Winship – Chapter 9.


**May 20 – No class meeting**

No class meeting today.

Submit all materials for your field experiment (planning document; dataset and reproducible analysis code; research report).

**May 27 – Instrumental Variables**

*Readings*

Murnane and Willett – Chapters 10-11.


**Additional Readings**

Morgan and Winship – Chapter 7.


Tasks

Complete peer evaluation of field experiments.

Work on final projects.

Download & read through example IVE analysis.

June 3 – Lightning talks & Assessment

In-class ≤ 5 minute presentations of your final projects and discussion. Collective reflection on the design of the course and other related questions of profound significance. Much rejoicing.

Final Projects Due June 9

Final projects due no later than 5pm Central Time, June 9, 2015. Please submit your project via Canvas.

References


Gelman, A. and Loken, E. (2013). The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking"and the research hypothesis was posited ahead of time.


