**Course Modality**
This course will be run as a flipped class. Lecture content is available online along with the slides and Stata examples. Students will be required to view the lectures, respond to discussion questions, and complete Stata exercises online **before** the live session. In-person class time will be spent focusing on those questions and issues that students flagged as unclear or benefitting from more discussion. After class, you will have a lab with our TA and have another chance to over any questions.

**Course Description and Objectives**
This course is an introduction to the empirical toolkit widely used to estimate causal effects in wide range of settings -- from program evaluation, medical trials, public health interventions, and labor training programs to management HR interventions to policies implemented at the local, state, and national level. More generally, the methods we will discuss are part of the advanced toolkit of any data analyst -- they are the next step beyond multiple regression, although as we will see regression remains a bedrock tool that we will use often.

Another way to think about this course is that it primarily about understanding how to use variation in the data to estimate policy and causal effects, and having understood that how to seek out or create such variation to evaluate policies of interest.

We begin by discussing the strengths and limitations of multiple regression analysis and the relationship between regression and causal modeling. We then develop a sequence of extensions and alternatives, including: randomized trials, regression discontinuity, matching methods, difference-in-differences and panel data, and instrumental variables.
Having learned the methods, we will also learn how to implement them on large-scale data from real policy settings. This will entail expanding your comfort and expertise with advanced use of Stata, and learning how to conceptualize, clean, and handle large data sets in an abstract manner -- essential when dealing with data sets that are sufficiently large that you cannot eyeball and hand code data.

After taking this course, you will:

1. Understand the concept of causality and how it applies to policy research;
2. Understand the limitation of using direct comparisons of sample data to estimate policy effects (also known as sample selection bias);
3. Understand the importance of variation in policy exposure, in defining causal effects, in interpreting them, and in seeking out or even creating variation to evaluate policies of interest to you;
4. Understand (even better) the multiple regression model;
5. Understand at an advanced level the toolkit of methods we discuss
   • randomized controlled trials - both analyzing and designing
   • instrumental variable methods - what is valid instrument, how to find one, how to test for one, and how to estimate and interpret
   • regression discontinuity methods - how to look for RD’s, how to analyze them, and how to test the key assumptions
   • matching methods - understand the key assumptions and when they are plausible, understand their use as a data descriptive tool, and implement both direct and propensity score matching;
   • difference in differences estimation - understand why this is the most widely applied method, learn to test its assumptions, and estimate effects
6. Develop your advanced Stata skills - both for the sake of understanding a widely used package and also to get a bit of the coder's mindset;
7. Learn how to obtain a large data set, look at the codebook, clean and organize the data, and estimate the results;
8. Learn how to conceptualize, propose, and implement a strategy to estimate a policy or treatment effect of interest;
9. Become a proficient producer of empirical research; and
10. Become an advanced, critical, and thoughtful consumer of empirical research.

Grading
The course will be evaluated through five problem sets (12 points each), a pop quiz or two (up to one bonus point each), a semester-long replication exercise (30%), and class participation (completing online asynchronous material, attending and participating in live sessions) (10%). All problem sets will make use of Stata, so please ensure you are familiar with how to access this program at NYU.

Late Policy
Assignments are due prior to 6.45 pm on the dates indicated below on NYU classes. Late submission of problem sets will lead to a two-point reduction for missing the deadline, another one-point reduction for a one-week delay, and zero thereafter. Late submission of the replication exercise will lead to a three-point reduction for missing the deadline, two additional points for a one-week delay, and zero thereafter. Any waivers due to emergencies must be discussed with Student Affairs, who will then communicate with all relevant faculty.

Course Structure
The class includes lectures, readings, and independent computer lab work. You are strongly encouraged to relate the general material of the course to your specific policy interests throughout the course. Class attendance is critical as interaction within the classroom is an essential aspect of this course and the learning process associated with it.

Expectations
Reading e-mail: I will communicate to you through your NYU e-mail. I’m assuming you read it at least once a day.

Preparation before class: come prepared for each class having read the required material carefully, viewed the online lecture content, and completed the discussion questions and Stata exercises. Given the nature of the material, you may find your first pass at a required article challenging; the first time through focus on the concepts and intuition (often found in the introduction).
Classroom Norms and netiquette: You are expected to join each class on time. We will often use computers for Stata exercises in class. Please do not multitask. Please keep other windows, including social media and e-mail, closed, and remember that class participation is part of your grade.

Academic Integrity
Academic integrity is a vital component of Wagner and NYU. All students enrolled in this class are required to read and abide by Wagner's Academic Code. All Wagner students have already read and signed the Wagner Academic Oath. Plagiarism of any form will not be tolerated and students in this class are expected to report violations to me. If any student in this class is unsure about what is expected of you and how to abide by the academic code, you should consult me.

Henry and Lucy Moses Center for Students with Disabilities at NYU
Academic accommodations are available for students with disabilities. Please visit the Moses Center for Students with Disabilities (CSD) website and click the "Get Started" button. You can also call or email CSD (212-998-4980 or mosescsd@nyu.edu) for information. Students who are requesting academic accommodations are strongly advised to reach out to the Moses Center as early as possible in the semester for assistance.

Technology Support
You have 24/7 support via NYU’s IT services. Explore the NYU servicelink knowledgebase for troubleshooting and student guides for all NYU-supported tools (NYU Classes, Zoom, etc). Contact askIT@nyu.edu or 1-212-998-3333 (24/7) for technology assistance, or contact Zoom’s 24/7 technical support (includes a chat function), or review Zoom’s support resources. Your peers are another source of support, so you could ask a friend or classmate for help or tips.

If you do not have the appropriate hardware technology nor financial resources to purchase the technology, consider applying for the NYU Emergency Relief Grant.

NYU’s Calendar Policy of Religious Holidays
NYU’s Calendar Policy of Religious Holidays states that members of any religious group may, without penalty, absent themselves when required in compliance with their religious obligations. Please notify me in advance of religious holidays that might coincide with exams to schedule mutually acceptable alternatives.
Books
There is no textbook for this course. The assigned readings (mainly articles) and lecture notes together will provide a self-contained treatment.

An introductory level book that covers similar material is:


This book blends an introduction to econometrics with the basic tools of causal inference we discuss in this course.

Two more technical references are:


These books are more technically advanced than the material our course, so are useful for students who wish to get a flavor of more advanced material or have an advanced reference at hand.

Schedule
8 September, Lecture 1: Preliminaries and review
15 September, Lecture 2: Introduction to causality
22 September, Lecture 3: Randomized controlled trials: basics
29 September, Lecture 4: Randomized controlled trials: advanced
6 October, Lecture 5: Introduction to Stata and discussion of PS 1
13 October, Lecture 6: Instrumental variables, a (re-) introduction, and problem set 1 due + replication data
20 October, Lecture 7: Instrumental variables: advanced
27 October, Lecture 8: Class discussion and tutorial for replication exercise, PS 2 discussion
3 November, Lecture 9: Regression discontinuity, introduction, and problem set 2 due
10 November, Lecture 10: Regression discontinuity, advanced
17 November, Lecture 11: Matching, basics, and problem set 3 due
24 November, Lecture 12: Replication exercise
1 December, Lecture 13: Matching, advanced
8 December, Lecture 14: Difference-in-differences, **problem set 4 due**
and PS 5 discussion
15 December: No class: **Problem set 5 and replication assignments due.**

*Note:* These due dates are tentative; they will be postponed if we have not covered the necessary material.
Readings
The asterisk (*) denotes required, the pound sign (#) denotes an advanced reading where you should only skim the non-technical portions. IR denotes chapters from Imbens- Rubin. MM denotes chapters from Angrist-Pischke’s Mastering Metrics. Articles (not books) will available on our Brightspace site.

Topic 1: Causality and the Treatment Effect

IR, chapter 1.


Topic 2: Introduction to Randomized MM, Chapter 1.


Marianne Bertrand, Dean Karlan, Sendhil Mullainathan, Eldar Shafir, and Jonathan Zinman, “What’s Psychology Worth?”


Cox, The Planning of Experiments, selected chapters (currently unavailable online; check the library if you are interested in this topic).

Topic 3: Randomized Experiments, Advanced Issues IR, Chapters 4-7
and 9-11.


**Topic 4: Instrumental Variables, A Re-Introduction**


MM, Chapter 3.

Pitt, Mark, Mark Rosenzweig, and Nazmul Hassan, “Identifying the Hidden Costs of a Public Health Success: Arsenic Well Water Contamination and Productivity in Bangladesh”

**Topic 5: Instrumental Variables, Advanced Topics**


Imbens, Guido, and Jeffrey Wooldridge (2009), “Instrumental Variables with Treatment Effect Heterogeneity: Local Average Treatment Effects,” manuscript.

AR, Chapters 23-25.

**Topic 6: Regression Discontinuity, Introduction**


Topic 7: Regression Discontinuity, Advanced Topics


Topic 8: Matching, Introduction
* Stuart, Elizabeth (2009), “Matching Methods for Causal Inference: A Review and a Look Forward,” manuscript, Department of Biostatistics, Johns Hopkins University


IR, Chapter 18.

Topic 9: Matching, Advanced Topics


# Matias Busso, John DiNardo, and Justin McCrary, "New Evidence on the Finite Sample Properties of Propensity Score Reweighting and Matching Estimators," manuscript.

IR, Chapters 12-17.

**Topic 10: Difference-in-Differences**


Meyer, Bruce, “Natural and Quasi Experiments in Economics,” December 1994


MM, Chapter 5.