

Instructor Prof. Vivek Bhattacharya
vivek.bhattacharya@northwestern.edu
Office Hours: Wednesdays 10:30 am–12:00 pm in KGH 3379
and by appointment
(Please shoot me an email if you plan to stop by office hours.)

TA Zhen Xie
zhenxie@u.northwestern.edu
Office Hours: Fridays 3:30-5:00 pm in KGH 3496

Course Description: The primary purpose of this course is to teach you some simple tools to take a dataset and learn something about the patterns hidden in it. In the first part of the course, we will go over tools to estimate statistical correlations. These correlations are informative, and it's important to know how to get them right. But, such statistical correlations are often not the object of economic interest. Academic economists, businesspeople, and policy makers alike often care about the *causal* effect a policy instrument and an outcome, which need not coincide with the correlation between the policy and the outcome. We'll discuss why we care about causal effects, what sort of questions can be answered with a precise causal estimate rather than simply a correlation, and why it can be misleading to confuse correlation with causation.

This takes us to the second part of the course—and indeed the meat of it. We will develop formal models that can be used to estimate causal effects. While we will discuss the statistical properties of these models and develop estimators, the goal is not so much to learn how to plug-and-chug but rather to learn how to critically assess (i) what conceptual quantity a model is designed to estimate, and (ii) whether a model is applicable for a particular question. As such, the theory will be interspersed with applications. We'll try to devote a few classes at the end of the course either to critiquing empirical analysis in depth or to student presentations about research topics using the techniques in this class.

Through the problem sets, the recitations, and the final project, you will develop some familiarity with the “practical” parts of econometrics as well. You'll learn how to implement these estimators in statistical software and have a chance to critique them in the context of a recent study of an important policy issue.

Prerequisites: While 381-1 is the formal prerequisite, you should be fine if you've taken introductory statistics, probability, linear regression, and linear algebra.

Textbooks and Software: I will distribute lecture notes. If you'd like supplemental material, I'd recommend the following textbooks. We won't use these textbooks formally in class (i.e., I won't assign reading from them). So if you are thinking about purchasing them, feel free to buy whichever edition is cheapest. I would strongly recommend acquiring a copy of *Mostly Harmless Econometrics* if it fits with your budget; the others are certainly optional, and I wouldn't recommend buying them if you don't have them already.

- Angrist and Pischke (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. A lot of the material in my lecture notes will come from this book, although the book is

geared towards incoming graduate students rather than advanced undergraduates. If you are interested in purchasing one book for this course, I'd recommend this one. It's cheap for a textbook, fun to read, and it will also provide you with more material to digest if you decide to keep going with economics, econometrics, or any sort of policy analysis.

- Stock and Watson (2011). *Introduction to Econometrics*. This is a fairly comprehensive overview of the material we cover in class. It is maybe a more “basic” undergraduate textbook, though, and the focus is slightly different. You'll be fine in this class even if you don't have it.
- Angrist and Pischke (2014). *Mastering 'Metrics: The Path from Cause to Effect*. This book is geared towards undergraduates and covers much of the material from *Mostly Harmless Econometrics* that we will cover in class. It's pitched in a somewhat basic (and unfortunately silly) way, but it could be good supplemental material.
- Cunningham (2020). *Causal Inference: The Mixtape*. This book also covers much of the material we'll cover, and you might want to look at it if you want a different treatment from the lecture notes. It's also free online at <http://mixtape.scunning.com>.

Problem sets have been designed to use Stata. Stata is unfortunately commercial software, but you can find it through NUWorkspace at <http://nuworkspace.northwestern.edu/>. If you do not have an account, I recommend you request one immediately.

If you are interested in learning a broader-use, open-source language, you are welcome to complete the problem sets in R or Python. You might get more limited help from me with technical issues.

Grading: Problem Sets: 15%, Midterm 1: 35%, Midterm 2: 35%, Final Project: 15%.

- *Problem Sets:* There will be 5 problem sets. I'll drop your lowest score, so you can feel free to miss one assignment without penalty if things get busy during the semester.

Problem sets will be submitted on Crowdmark. I will not give extensions on problem sets, nor will I accept late problem sets, other than in truly exceptional situations (with a Dean's note).

You can work in groups of up to four (although I'd recommend no more than three) on each problem set. Groups can turn in a single copy of their solutions through Crowdmark. Make sure you have an active role in solving the problems. It is not in your best interest to free-ride off others: you'll have to do the exams yourself!

A randomly selected subset (and possibly all) of the problems will be graded carefully for accuracy. The remainder of the problem set will be graded on completion. Show all your work. For computational problems, submit the (well-documented) code and the output, but also incorporate your results into your writeup in a clean, readable fashion.

Be neat. If Zhen and I can't read your work, we won't grade it. I'd recommend typing your problem sets, especially if you have messy handwriting.

- *Exams:* The midterms will in-class on **Monday, April 24** and **Wednesday, May 24**. If you anticipate having legitimate conflicts with the midterm dates, please let me know by **April 3**.
- *Final Project:* A goal of the class is to prepare you to think critically about analyses done in various settings—academic papers, policy briefs, popular press, etc. The final project will give you an opportunity to conduct such an analysis. Approximately three weeks before the end of the class, I will distribute a piece of empirical analysis about a current policy issue. You will write a 3–5 page (single-spaced) critique of this paper. What question does it answer? Do you

find the answer convincing? Why or why not? What other analyses could the authors conduct to convince you of the results? What do you think their empirical strategies are missing? How would you do the analysis differently?

Most likely, I will also offer an alternate option of submitting a research proposal using materials from this class. Depending on interest in this option, we may devote time for student presentations of their research proposals at the end of the class.

You will have to work by yourself on this project. You can research anything online (as long as it is cited). The project should be emailed to me by **Tuesday, May 30, at 11:59 pm**. I will accept late work with no penalty until Friday, June 4, at 11:59 pm. I will not accept late work after that unless you have a good reason. More details will be distributed about the final project closer to the end of the quarter.

Academic Integrity Statement: Students in this course are required to comply with the policies found in the booklet, “Academic Integrity at Northwestern University: A Basic Guide”. All papers submitted for credit in this course must be submitted electronically unless otherwise instructed by the professor. Your written work may be tested for plagiarized content. For details regarding academic integrity at Northwestern or to download the guide, visit: <https://www.northwestern.edu/provost/policies-procedures/academic-integrity/index.html>.

AccessibleNU Notice: Northwestern University is committed to providing the most accessible learning environment as possible for students with disabilities. Should you anticipate or experience disability-related barriers in the academic setting, please contact AccessibleNU to move forward with the university’s established accommodation process (e: accessiblenu@northwestern.edu; p: 847-467-5530). If you already have established accommodations with AccessibleNU, please let me know as soon as possible, preferably within the first two weeks of the term, so we can work together to implement your disability accommodations. Disability information, including academic accommodations, is confidential under the Family Educational Rights and Privacy Act.

Support for Wellness and Mental Health: Northwestern University is committed to supporting the wellness of our students. Student Affairs has multiple resources to support student wellness and mental health. If you are feeling distressed or overwhelmed, please reach out for help. Students can access confidential resources through the Counseling and Psychological Services (CAPS), Religious and Spiritual Life (RSL) and the Center for Awareness, Response and Education (CARE). Additional information on all of the resources mentioned above can be found here:

<https://www.northwestern.edu/counseling/>
<https://www.northwestern.edu/religious-life/>
<https://www.northwestern.edu/care/>

Tentative Schedule

March 28 (Tuesday)	Introduction and Discussion of Empirical Work
March 29	Causality and Potential Outcomes
March 31 (Friday)	Ordinary Least Squares
April 3	Ordinary Least Squares
April 5 (Zoom)	Limited Dependent Variables
April 10	Limited Dependent Variables and MLE
April 12	Nonparametric Regression
April 17	Nonparametric Regression
April 19	Review for Exam 1 + Panel Data and Fixed Effects
April 24	Midterm 1

April 26	Instrumental Variables
May 1 (Maybe Zoom)	Instrumental Variables
May 3	No Class
May 8	Instrumental Variables and Heterogeneity
May 10	Differences-in-Differences
May 15	DID and Regression Discontinuity
May 17	Regression Discontinuity + Review for Exam 2
May 22	Discussion of Empirical Topics / Student Presentations
May 24	Midterm 2
May 29	Memorial Day (No Class)
May 30 (Tuesday)	Final Project Due at 11:59 pm

Problem sets will tentatively be due on April 7 (Friday), April 14 (Friday), April 20 (Thursday), May 15 (Monday), and May 21 (Sunday). The problem set deadlines are chosen so that you have time to review the solutions before the exam. I will try to stick to the proposed schedule of topics as closely as possible, but the schedule may vary a bit from what's on the syllabus. Recitations will cover additional practice problems as well as (potentially) discussions of empirical papers.