

Prof. Robert Porter
KGGH 3337
r-porter@northwestern.edu

Economics 381-2
Econometrics
Spring 2022

This is the second quarter of the econometrics sequence. You should be familiar with the material covered in Economics 381-1. The attached list of topics describes the material to be covered in this course.

The primary purpose of this course is to teach you some tools to analyze a data set. 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 statistical correlations are often not the object of economic interest. Academic economists, businesspeople and policy makers alike often care about the causal effect of a policy instrument on 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 causal estimate rather than a correlation, and why it can be misleading to confuse correlation with causation.

This takes us to the second part of the course. We will develop formal models that can be used to estimate causal effects. We will discuss the statistical properties of these models and develop estimators, but the goal is not to learn how to plug-and-chug. The goal is instead to learn how to assess critically (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 devote a few classes at the end of the course to critiquing some empirical analyses in depth and to student presentations about research projects using the techniques discussed in this class.

Through the problem sets, the TA sessions, 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 to critique them in the context of a recent study of an important policy issue.

There will be several problem sets, most of which will involve some analysis of data. This is an applications course, and the best way to learn the material is to employ it in problem solving. Problem sets must be submitted by 3 pm on their due dates, via Crowdmark. Late problem sets will not be accepted. The problem sets are designed to be done in groups of two people, who can turn in their answers as a group. Discussing problem sets with other groups is allowed and encouraged, but problem sets are meant to be reflective of each group's work. Copying another group's assignment, permitting another group to copy from your assignment, or jointly writing up answers across groups will result in no credit for the entire assignment for all involved

students. Furthermore, dividing up problem set questions across students is discouraged. Problem sets are intended to help you study the material and to learn important skills, and skipping a problem entirely will be detrimental to your learning.

In each problem set, some questions will guide you through the process of developing an original research question. At the end of the quarter, we will ask you to put everything together and make a 15 minute presentation concerning a research proposal. This is also to be done in your problem set group of two. We strongly suggest taking the problem set questions regarding the research proposal seriously during the quarter, as doing so will ensure that you will not have a difficult time preparing your presentation. Research proposal grades are a combination of the presentation itself and the questions you ask during your classmates' presentations. To that end, there is an expectation that you will attend class on days when you are not presenting. However, attendance is not mandatory. This is meant to allow for some flexibility for students with early internships while still incentivizing participation. The presentation dates will be the class and TA session times during reading week (primarily on June 1, and June 3 if necessary). I will randomly assign students to presentation times. If you have an early internship and cannot present on one of the two allotted dates, please let me know by May 16. No action is needed if you have no conflicts. More details concerning the research proposal are covered in the "Research Proposal" document on Canvas.

In addition, there will be two midterm exams. The midterms will be during class times, on Wednesday, April 27 and Wednesday, May 25. Both exams will be open book and open notes. The approximate weighting scheme for the final grade will be:

Problem Sets	20%
Midterm Exams	30% each
Research Proposal	20%

Any student who misses the midterms without a good reason will receive no credit for the exams. Medical reasons require written confirmation from the University health services or your doctor. Non-medical reasons require a written excuse from some official of the University (e.g. an official from the Dean's Office.) A makeup for the midterms will be permitted only if there is a good reason for missing the exam. Suspicion of academic dishonesty will be immediately referred to the Dean's office.

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 notice from AccessibleNU, preferably within the first two weeks of class. All information will remain confidential.

The Teaching Assistant for the course is David Stillerman, davidstillerman2021@u.northwestern.edu. David will hold discussion sections on Friday at the scheduled class time. The discussion sessions will cover problem sets, additional practice problems, as well as discussions of empirical papers.

The plans for how the course will be run are described below. However, we may modify various aspects, depending on campus restrictions, as well as feedback we receive from you and our own evaluations. We always welcome feedback on how to improve the course, and your feedback is especially welcome this year.

All lectures and discussion sections will be conducted in person, and attendance is strongly recommended. Current university guidelines stipulate that students must wear masks in class, but instructors may remain unmasked while teaching as long as they do not have close contact with others. Office hours will be also conducted in person. If you would like a meeting outside office hours, either in person or via Zoom, this can be arranged by appointment with either myself or David.

I will post copies of lecture notes before the topics are covered in class. You may find it useful to review the material in advance. As far as exams are concerned, you are only responsible for the material covered in class. If you'd like supplemental material, I'd recommend the following textbooks. We won't use these books formally in class (i.e., I won't assign reading from them). If you are thinking about purchasing them, feel free to buy whichever edition is cheapest. I recommend acquiring a copy of *Mostly Harmless Econometrics* if it fits with your budget. The others are 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 the 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 (2019), *Introduction to Econometrics*. This is a comprehensive overview of the material we cover in class. It is a more "basic" undergraduate textbook, however, and the focus is slightly different. Most of you should have this textbook, since it was required in 381-1, but you'll be fine in this class 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 simplistic (and silly) way, but it could be a good supplement.

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> .

Course Outline

0. Introduction and Logistics
1. Causality and Potential Outcomes
2. Ordinary Least Squares
3. Limited Dependent Variables
4. Nonparametric Regression
5. Instrumental Variables
6. Differences-in-Differences
7. Regression Discontinuity
8. Discussion of Empirical Topics

If we have extra time, we will also discuss matching estimators or quantile regressions.