Objectives:
Once the research question is established, there are three core steps to effective empirical work:
   (1) Establish what relationships are in the data
   (2) Interpret those relationships in light of your research objectives
   (3) Communicate those relationships as clearly, completely, and convincingly as possible

The main focus of the course will be on methods for establishing causal relationships in field data. This means we will discuss how to establish what relationships exist in the data, when you can interpret these relationships as causal, and how you can convince your audience of your results (without overselling).

Because methods aren’t too useful without interesting questions to answer, we will also spend time developing our “taste” for what constitutes a quality empirical research paper. The ultimate goal is for you to leave prepared to undertake your own empirical research.

We will also think carefully about the interaction between empirical design and theory, especially the importance of careful theoretical thinking for empirical research.

Preparation and Prerequisites: This course is designed to complement a graduate sequence in econometrics, but it should be accessible to students with basic knowledge of statistics and probability. We will focus on intuition and understanding how statistical models relate to the underlying data (and theory). Still, there will be technical material in readings, discussions, and assignments.

Grading:
Class participation: 10%
Class Discussant 1: 10%
Class Discussant 2: 10%
Research paper proposal: 5% (due Jan 31)
Replication exercise 1: 10%
Replication exercise 2: 10%
Research paper class presentation: 20% (April 25)
Research paper extended abstract: 25% (due Apr 30)

Class participation: Attendance is mandatory. Class discussion is a critical component of the seminar. You will be expected to complete the readings assigned before class and come to class with comments and questions.
Class discussant: Discuss assigned class readings market with *. No written report required. I will circulate a sign-up sheet during the first day of class.

Research paper (individual or group): Individually or in a group of no more than two, select a research paper topic of your choice. We'll work on it throughout the duration of the course to exemplify the concepts learned. At the end of the course, you will present your project in front of the class. You will also submit a written extended abstract of maximum 10 double-spaced pages of text plus tables, figures and references. It is possible that some projects require lengthy data collection. In that case, you should at least know what the data are like and how you can obtain the data so that the paper can be as specific as possible about the implementation.

Class Structure: Most sessions we will spend time on: 1) Listening to the class discussants’ report, 2) Lecture plus reviewing the assigned readings (all of them) and 3) Discussing your projects in light of session learnings. Two sessions are dedicated to replicating empirical papers. Read the paper to be replicated in advance. Note: All readings are required. The * marks readings assigned to class discussants, but everybody must reading everything.

Books:
Required:

Other books, not required for this class, but good references for anyone doing empirical work:

<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>January 10 Getting Started</td>
</tr>
<tr>
<td>2</td>
<td>January 17 Causal effects: Why do we care and why so hard to establish? (I)</td>
</tr>
<tr>
<td>3</td>
<td>January 24 Project proposal</td>
</tr>
<tr>
<td>4</td>
<td>January 31 Causal effects: Why do we care and why so hard to establish? (II)</td>
</tr>
<tr>
<td>5</td>
<td>February 7 Difference-in-Differences (I)</td>
</tr>
<tr>
<td>6</td>
<td>February 14 Difference-in-Differences (II)</td>
</tr>
<tr>
<td>7</td>
<td>February 21 Replication exercise</td>
</tr>
<tr>
<td>8</td>
<td>February 28 Instrumental variables</td>
</tr>
<tr>
<td>9</td>
<td>March 7 Matching</td>
</tr>
<tr>
<td></td>
<td>March Break (March 11-18)</td>
</tr>
<tr>
<td>10</td>
<td>March 21 Regression Discontinuity</td>
</tr>
<tr>
<td>11</td>
<td>March 28 Quantile regression; Event studies</td>
</tr>
<tr>
<td>12</td>
<td>April 4 Replication exercise</td>
</tr>
<tr>
<td>13</td>
<td>April 11 Continuous, discrete and binary dependent variables</td>
</tr>
<tr>
<td>14</td>
<td>April 18 Testing</td>
</tr>
<tr>
<td>15</td>
<td>April 25 Project Presentations</td>
</tr>
</tbody>
</table>
Week 1 (January 10): Getting started

Course goals

Syllabus

Getting started:
1) Data structure: Panel, Cross section; Unit of analysis
2) How to read (and write) an empirical paper (e.g., research question, data structure/unit of analysis, estimating equation(s); main effect, contingency, mechanism; robustness, falsification tests)

Week 2 (January 17): Causal effects: Why do we care and why so hard to establish? (I)

Theory: What is identification?
1) *MM Chapter 1
2) MHE Chapters 1 and 2

In practice

Week 3 (January 24): Develop project proposals

Week 4 (January 31): Causal effects: Why do we care and why so hard to establish? (II)

Theory: Regression fundamentals
1) *MM Chapter 2
2) MHE Chapter 3 (Only 3.2 (all subsections) and 3.4.3)

In practice
Field experiments

Describing interesting data

Week 5 (February 7): Difference-in-differences (I)

Theory: Difference-in-differences fundamentals
1) *MM Chapter 5 (Only 5.1 (all subsections))

In practice

Week 6 (February 14): Difference-in-differences (II)

Theory: Difference-in-differences, fixed effects, lagged variables and controls
1) *MHE Chapter 5

**In practice**

**Week 7 (February 21): Paper replication exercise**
*Class exercise; data will be provided in class. Bring your laptop.*
Empirical paper TBD

**Week 8 (February 28): Instrumental variables**
**Theory:** *IV design (including 2SLS), Language of treatment effects*
1) *MM Chapter 3*
   (Pay closer attention to sections 2.1, 2.2, 3.1, 5.1, 5.3, 5.4, 5.5, 6.3, 6.4, and 6.5)

**In practice**

**Week 9 (March 7): Matching**
**Theory:**
1) MHE 3.3

**In practice**
   This paper is a rebuttal of Jaffe, A., M. Trajtenberg and R. Henderson (1993), “Geographic Knowledge Spillovers as Evidenced by Patent Citations” Quarterly Journal of Economics, 108(3):577-98. It may be useful to skim this original paper for a better understanding of the original study.

(March 11-18: March Break)

**Week 10 (March 21): Regression Discontinuity**
**Theory:** *Sharp RD, Fuzzy RD*
1) *MHE Chapter 6*
2) MM Chapter 4

**In practice**
2) *Fehder, Dan. 2015. “Startup Accelerators and Ecosystems: Complements or Substitutes?” (job market paper)*
Week 11 (March 28): Quantile regression, Event studies

Theory

Event studies

Quantile regression
2) *MHE Chapter 7

In practice

---

Week 12 (April 4): Paper replication exercise

Class exercise; data will be provided in class. Bring your laptop.

Empirical paper TBD

---

Week 13 (April 11): Continuous, discrete and binary dependent variables

Theory

Continuous and discrete dependent variables (OLS and Poisson)
1) *MHE Chapter 3 (Only 3.1 (all subsections))

Binary dependent variables (Probit, Logit, Tobit)
4) *MHE Chapter 3 (Only 3.4.2) and Chapter 4 (Only 4.6.3)

Interaction terms
(Interpreting interaction terms in linear models)

In practice

Count data

Binary Dependent Variables

**Week 14 (April 18): Testing**

**Theory**

*Economic vs. statistical significance*


*Random effects and Hausman tests*


**In practice**

*Economic vs. statistical significance*


*Random effects and Hausman tests*


*Fit, R-squared, and explanation vs. prediction*


**Week 15 (April 25): Project presentation**