Draft Course Syllabus

Econometrics for Managers:

Correlation & Causation in a Big Data World

Fall 2018

Faculty:

Professor Joseph Doyle
E62-516 jjdoyle@mit.edu

Teaching Assistant:
Mert Demirer mdemirer@mit.edu
Course Goals

This course enables students to transform data into knowledge to make better business decisions. Whether students anticipate producing or consuming data analytics, the course provides the tools to identify problems and a framework to solve them. That framework is econometrics: modern empirical methods used by economists.

How will you learn these tools? We will introduce the topics and then use in-class labs so students are learning by doing. To complement the in-class learning, we will also have short problem sets and group projects. At the end of the course, students will be able to evaluate the quality of evidence supported by data, and implement a whole toolkit aimed at providing credible answers to difficult questions.

Specifically, we will develop the foundations of econometrics and regression analysis. We offer some theory, but mostly practical examples including: forecasting using high-frequency financial data, developing a trading model, running experiments to test the effects of pricing, product quality, and marketing decisions, and discovering “natural” experiments that allow analysts to explore Big Data to go beyond correlations. One workhorse model we will introduce is instrumental-variable estimation.

Big Data is a hot topic these days, and we will explore the potential opportunities and pitfalls that arise with the flood of data currently being generated.

Course Complements & Pre-Requisites

There are no pre-requisites for the course, as we will define terms as we go. The course is complementary to other data-analysis courses at the school, such as Analytics Lab, Statistical Consulting, Analytics Edge and Data Mining; we believe the more you see these concepts the better you will be at using them. What distinguishes this course is that we view these statistical concepts through the framework provided by economics and econometrics: a way of modeling the empirical problem to evaluate correlations and “get to causality”.

Readings

We will rely on articles posted on Stellar, plus book chapters in the Course Reader available at study.net, including textbook treatments of the main topics. For those interested in a textbook, you will see from the readings that we are partial to Stock and Watson’s Introduction to Econometrics, Angrist & Pischke’s Mastering ‘Metrics and Wooldridge’s Introductory Econometrics. Another useful text is Greene’s Econometric Analysis. We will have these textbooks on reserve at Dewey Library.
Recitations/Office Hours
We will hold weekly recitations to aid you in learning how to use R and answer questions about the course and the projects. These are optional but we have found them to be appreciated by students. We will also offer office hours; we are aiming for Fridays depending on class demand.

Evaluation:
Grading will be determined as follows:

<table>
<thead>
<tr>
<th>Component</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class Participation</td>
<td>10%</td>
</tr>
<tr>
<td>Problem Sets</td>
<td>20%</td>
</tr>
<tr>
<td>Take-home quiz</td>
<td>20%</td>
</tr>
<tr>
<td>Team Projects</td>
<td></td>
</tr>
<tr>
<td>Data Exploration Project</td>
<td>10%</td>
</tr>
<tr>
<td>034 Hedge Fund Project</td>
<td>20%</td>
</tr>
<tr>
<td>034 Consulting Project</td>
<td>20%</td>
</tr>
</tbody>
</table>

Class participation will reward students for comments that help the class, including discussion of experiences related to the material.

Problem sets are not designed to be difficult; rather they will enable you to become more competent with statistical programming for the projects and internalize the lessons from the course. The take-home quiz will be somewhat longer in format, will be completed individually, and will test your understanding of the lessons from the first half of the course.

For the team projects, students may join groups of up to five people. The first project is a data exploration exercise using cutting-edge data used to understand the extent to which management matters for firm performance.

The second project will use Billion Prices Project data to predict official measures of inflation and build a trading model using these predictions (the 034 Hedge Fund!).

The third project asks you to investigate whether core teams (study groups) matter for performance at Sloan and beyond. We will provide de-identified data including admissions criteria (e.g. GMAT scores), performance at Sloan (e.g. GPA), and exit survey results (including salary and industry). This is a unique dataset that you can analyze: you are the experts on how such study groups function. The answers can help shape core-team design (034 Consulting!).
The coursework for 15.0341, the undergraduate course number that “meets with” 15.034, is the same. We recognize that undergraduate students have less work experience compared to graduate students and their course grades will be graded separately relative to other students in 15.0341.

We welcome suggestions to substitute the above projects with a question and data of your own choosing. The data would need to be accessible to the class and we would determine if it is appropriate for a semester-long course.

**Software**

As part of the course you will learn to use R, which is a straightforward but powerful statistical software package. This is a valuable tool to go well beyond Excel in your analysis of data. This will be the basis for in-class labs as well as problem sets and projects.

R is free software and we will post how to download it to stellar.
Course Outline (Readings to be added)

**Sept 6  Thurs**  Course Introduction:  Big Data & Getting to Causality

**Sept 10  Mon**  Recitation

**Sept 11  Tues**  Big Data Introduction

Read:  The Billion Prices Project, Cavallo and Rigobon JEP

The Economist:  The return of the machinery question

Due:  Problem Set 1

**Sept 13  Thurs**  Big Data Tools 1:  Regression Modeling and Overfitting

Read:  The Parable of Google Flu

The Economist:  From not working to neural networking

Watch:  Machine Learning 1 Video on Stellar (Minute 31-end)

**Sept 18  Tues**  Big Data Tools 2:  CART & Random Forest

**Sept 20  Thurs**  Big Data Tools 3:  LASSO Estimation

**Sept 24  Mon**  Recitation

**Sept 25  Tues**  Causality 1:  What does a multivariate regression actually do?

Read:  Wooldridge p69-81

Stock & Watson p123-126:  Wooldridge p80-84

Due:  Problem Set 2

**Sept 27  Thurs**  Human Resources Application:  Measuring Discrimination & Returns to Diversity
|------|--------|-----|------|--------|-----|------|---------|-----|------|--------|-----|------|---------|-----|------|---------|-----|------|--------|-----|------|---------|-----|------|---------|-----|
Oct 29 Mon No Recitation

Oct 30 Tues Introduction to Freakonomics: Natural Experiments
Read: *Freakonomics* (excerpt)
Review: Returns to Physician Human Capital
Due: Take-home Quiz

Nov 1 Thurs Natural Experiments 2: Regression Discontinuity Designs & Event Studies

Nov 5 Mon Recitation

Nov 6 Tues Natural Experiments 3: Panel Data Methods
Read: Wooldridge p484-491
Due: Group Project 2 (Part 1)

Nov 8 Thurs Natural Experiments 4: Difference-in-Differences
Read: Mostly Harmless Econometrics: Ch1-2
Review: Vertical Relationships and Competition in Retail Gasoline Markets

Nov 9 Fri Due: Group Project 2 (Update)

Nov 12 Mon No Recitation: Holiday

Nov 13 Tues Group Project Presentations

Nov 15 Thurs Introduction to Instrumental Variable Estimation
Read: Wooldridge p512-521; 528-530
Review: Evidence of a Shift in the Short-Run Price Elasticity of Gasoline Demand
Due: Problem Set 3
<table>
<thead>
<tr>
<th>Date</th>
<th>Day</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nov 19</td>
<td>Mon</td>
<td>No Recitation</td>
</tr>
<tr>
<td>Nov 20</td>
<td>Tues</td>
<td>No Class</td>
</tr>
<tr>
<td>Nov 22</td>
<td>Thurs</td>
<td>No Class: Holiday</td>
</tr>
<tr>
<td>Nov 26</td>
<td>Mon</td>
<td>Recitation</td>
</tr>
<tr>
<td>Nov 27</td>
<td>Tues</td>
<td>Instrumental Variables Estimation 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Read: Instrumental Variables and the Search for Identification</td>
</tr>
<tr>
<td>Nov 29</td>
<td>Thurs</td>
<td>Application: Measuring the Effects of Walmart</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Review: Walmart and Local Economic Development</td>
</tr>
<tr>
<td>Dec 3</td>
<td>Mon</td>
<td>Recitation</td>
</tr>
<tr>
<td>Dec 4</td>
<td>Tues</td>
<td>Probit/Logit &amp; Selection Models</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Read: Stock &amp; Watson: Chapter 11, p383-396, 400-408</td>
</tr>
<tr>
<td>Dec 5</td>
<td>Wed</td>
<td>Due: Write up for Group Project 3</td>
</tr>
<tr>
<td>Dec 6</td>
<td>Thurs</td>
<td>Group Project 3: Team Presentations</td>
</tr>
<tr>
<td>Dec 10</td>
<td>Mon</td>
<td>No Recitation</td>
</tr>
<tr>
<td>Dec 11</td>
<td>Tues</td>
<td>Lessons: Getting to Causality vs. Getting to Forecasts</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Due: Problem Set 4</td>
</tr>
</tbody>
</table>
STUDENT SUPPORT SERVICES

If you are dealing with a personal or medical issue that is impacting your ability to attend class, complete work, or take an exam, please discuss this with Student Support Services (S3). The deans in S3 will verify your situation, and then discuss with you how to address the missed work. Students will not be excused from coursework without verification from Student Support Services. You may consult with Student Support Services in 5-104 or at 617-253-4861.

STUDENT DISABILITY SERVICES

MIT is committed to the principle of equal access. Students who need disability accommodations are encouraged to speak with Kathleen Monagle, Associate Dean, prior to or early in the semester so that accommodation requests can be evaluated and addressed in a timely fashion. Even if you are not planning to use accommodations, it is recommended that you meet with SDS staff to familiarize yourself with the services and resources of the office. You may also consult with Student Disability Services in 5-104 or at 617-253-1674. If you have already been approved for accommodations, please contact your professor early in the semester so that we can work together to get your accommodation logistics in place.