ECON 626: Empirical Microeconomics

("Mostly Painless Econometrics")

Department of Economics
University of Maryland, College Park
Fall 2019
Revised November 1, 2019

1 Course Information

1.1 Instructors:

Dr. Pamela Jakiela
Center for Global Development
Email: pjakiela@cgdev.org

Dr. Owen Ozier
Development Research Group
The World Bank
Email: oozier@worldbank.org

1.2 Teaching Assistant:

Ming Fang
Economics Department
University of Maryland
Office: 3115A Tydings Hall
Email: mfang@umd.edu

1.3 Lectures:

Time: Friday from 1:30 PM to 4:30 PM
Location: 2108 Tydings

1.4 Course Website:

http://economics.ozier.com/econ626

1.5 Office Hours

Professor Jakiela and Professor Ozier will hold joint office hours on Fridays from 10:30 AM to 12:30 PM in 3115F Tydings. TA Ming Fang will hold office hours on Wednesdays from 10:30 AM to 12:30 PM in 3115A Tydings. We encourage you to come by during office hours to introduce yourself or discuss any topic related to the course material or economics more broadly. If you cannot attend these office hours, please email us to set up an appointment.

2 About the Course

2.1 Course Description

This course provides an overview of modern microeconometric methods. Topics include linear and nonlinear models, causal inference (instrumental variables, difference-in-differences, regression discontinuity), and techniques for correct statistical inference (clustering, etc).

2.2 Course Objectives

1. To provide a foundation in linear models in microeconomics, emphasizing both the practical implementation of these models and the application of these models to the question of causal inference.

2. To introduce students to more advanced techniques that are used both to assess data needs before empirical work and to carry out robustness checks in the completion of such work.

3. To familiarize students with modern statistical and econometric software in order to use these models and techniques.

4. To equip students with the basic econometric tools needed to implement non-linear models in modern software.

2.3 Prerequisites

This course is intended for PhD students in the economics department who have completed their department's first-year courses.

2.4 Readings & Lecture Notes

Many readings are academic articles and working papers which are available online (either through JSTOR or via the authors' websites). The following econometric references are not required, but will prove useful (both in this class and in life):

- Angrist, Joshua, and Jörn-Steffen Pischke. (2008) *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.
- Cameron, Colin, and Pravin Travedi. (2005) *Microeconometrics: Methods and Applications*. Cambridge University Press.
- Cunningham, Scott. (2018) *Causal Inference: The Mixtape. Available for free on Cunningham's website.*
- Gerber, Alan S. and Donald P. Green. (2012) *Field Experiments*. W. W. Norton & Company.
- Imbens, Guido W. and Donald B. Rubin. (2015) *Causal Inference for Statistics, Social, and Biomedical Sciences*. Cambridge University Press.
- Train, Kenneth. (2009) *Discrete Choice Methods with Simulation*. Cambridge University Press. *Available for free on Train's website.*
- Wooldridge, Jeffrey. (2010) *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

2.5 Assignments & Grading

- **Problem sets**: you will complete six problem sets, each worth 10 points. Each problem set will involve applied work in Stata and occasionally R or MATLAB. You are encouraged to work in groups, but each student must turn in his or her own write-up of the solutions. Assignments will typically be handed out or announced during lecture. Supporting documents will be made available online.

PS	Handed Out	Due Date
1	September 6	September 12
2	September 13	September 19
3	October 4	October 10
4	October 25	October 31
5	November 8	November 14
6	November 22	December 5

- **Research notes**: you will complete four short writing assignments of a page or two, each worth 5 points. The purpose of these assignments is to keep you active in generating research ideas, and to provide feedback that can improve your written articulation of those ideas.

Dual submission guidelines: Sometimes other courses assign similar research notes; if you would like to submit a single research note (or two works that overlap substantially in content) to both this class and another class, that is completely fine with us—provided two conditions are met: (1) you clearly write on the assignment what two courses it is being submitted to, and who the faculty are; and (2) the faculty teaching the other course must also be informed of, and okay with, with this arrangement.

The exact format of the note can vary. You might imagine it as part of project proposal to a dissertation committee or to a small grant funding facility. Your notes should be brief: the first three notes could each be one page, but really should not exceed two pages each. The essentials of a research note (for empirical ideas, which most of yours should be in this class, as we will have the best ability to give feedback on empirical ideas) would include:

1. In a sentence, what is the research question?
2. Why is it an important question?
3. Cite at least one high-profile paper to which the current idea relates.
4. How does this idea represent an advance over what has been done before?
5. What data would you need? How would you get the data?
6. What is your identification strategy - what variation in the data is essential for estimating the parameter of interest?
7. Can you write down the equation you would estimate?
8. What other details do you know already that show your strategy to be feasible? Can you calculate a statistic or produce a diagram using public data that strengthen the argument that this is feasible, interesting, important, and so on?

The last of the four research notes will either be an expansion of one of your previous three, or a new idea. We will discuss these possibilities in class.

Note	**Due Date**
1	October 4
2	October 25
3	November 8
4	November 22

- **Class participation**: your participation in class over the course of the semester (attendance, asking questions, having done the readings) will be worth 17 points in total. To receive full credit, you must attend class regularly and participate in class discussions in a manner that demonstrates familiarity with the readings and course material.

- **Midterm and Final exams**: you will take two exams during this course, each worth 20 points. The midterm (October 18; see the notable dates below) is intended to cover the material in the first half of the course, roughly through regression discontinuity. The final (December 6) will focus on the second half of the course, but since the second half depends on the first, there will be some overlap. Exams will be closed-book.

The total number of points above is 137; the grade is a score out of 137.

2.6 Tentative Schedule

The schedule below is approximate, and will be updated as we progress through the semester.

1. Linear Models: Empirical Approaches to Causal Inference (6 classes)
   (a) Aug. 30: Rubin Causal Model, Experimental Ideal, Probability
   (b) Sep. 6: Regression, Data-Generating Processes
   (c) Sep. 13: Difference-in-Differences
   (d) Sep. 20: Instrumental Variables
   (e) Oct. 4: Regression Discontinuity
   (f) Oct. 11: Selection on Observables
2. Linear Models: Advanced Variations (4 classes)
   (a) Oct. 25: Statistical Power, Clustering, Experimental Design
   (b) Nov. 1: Randomization Inference, Wild Cluster Bootstrapping, Permutation Tests
   (c) Nov. 8: Multiple Hypothesis Testing, Attrition and Bounds
   (d) Nov. 15: Causal Mediation
3. Nov. 22: Maximum Likelihood Estimation, Non-linear Models (1 class)

Notable dates. The Midterm exam will take place in-class on October 18; the Final exam will take place in class on December 6. There will be no class on September 27 because of a very important football game. There will be no class on November 29 because of the Thanksgiving holiday.

3 Readings by week/topic

1. August 30: Rubin Causal Model, Experimental Ideal

   Required readings:

Fisher (1971[1935]): *The Design of Experiments*, "II. The Principles of Experimentation, Illustrated by a Psycho-Physical Experiment," pp. 11–26.

Recommended readings:

Angrist and Pishke (2009): *Mostly Harmless Econometrics*, chapters 1 and 2

Gerber and Green (2012): *Field Experiments*, chapters 1 and 2

Imbens and Ruben (2016): *Causal Inference*, chapters 1 and 2

Jamison (2019): "The Entry of Randomized Assignment into the Social Sciences," *Journal of Causal Inference*, forthcoming

Lilienfeld (1982): "The Fielding H. Garrison Lecture: Ceteris Paribus: The Evolution of the Clinical Trial," *Bulletin of the History of Medicine* 56(1): 1–18.

2. Sep. 6: Regression, Data-Generating Processes

Required readings:

Angrist and Pishke (2009): *Mostly Harmless Econometrics*, chapter 3

Recommended readings:

Cameron and Trivedi (2005): *Microeconomics*, chapter 4

3. Sep. 13: Difference-in-Differences

Required readings:

Angrist and Pishke (2009): *Mostly Harmless Econometrics*, chapter 5

Goodman-Bacon (2019): "Difference-in-Differences with Variation in Treatment Timing," working paper

Recommended readings:

Duflo (2001): "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment," *American Economic Review*, 91 (4): 795–813)

Bertrand, Duflo, and Mullainathan (2004): "How Much Should We Trust Differences-In-Differences Estimates?" *Quarterly Journal of Economics*, 119 (1): 249–275

Bleakley (2010): "Malaria Eradication in the Americas: A Retrospective Analysis of Childhood Exposure," *American Economic Journal: Applied*, 2: 1–45

4. Sep. 20: Instrumental Variables

   Required readings:

   Angrist and Pishke (2009): *Mostly Harmless Econometrics*, chapter 4

   Recommended readings:

   Imbens and Angrist (1994): "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62(2): 467–475

   Young (2019): "Consistency without Inference: Instrumental Variables in Practical Application," working paper, available from Author's website

5. Oct. 4: Regression Discontinuity

   Required readings:

   Angrist and Pishke (2009): *Mostly Harmless Econometrics*, chapter 6

   Recommended readings:

   Hahn, Todd, and Van Der Klaauw (2001): "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, 69(1): 201–209

   Lee and Lemieux (2010): "Regression Discontinuity Designs in Economics," *Journal of Economic Literature* 48: 281–355

   Keele and Titiunik (2015): "Geographic Boundaries as Regression Discontinuities," *Political Analysis* 23: 127–155.

6. Oct. 11: Conditional Independence, Selection on Observables

   Required readings:

   Altonji, Elder, and Taber (2005): "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools," *Journal of Political Economy*, 113 (1): 151–184

   James, Witten, Hastie, and Tibshirani (2013): *An Introduction to Statistical Learning*, "Linear Model Selection and Regularization," pp. 203–243. (A pdf of the book is available online: `http://faculty.marshall.usc.edu/gareth-james/ISL/ISLR%20Seventh%20Printing.pdf`.)

   Recommended readings:

   Cunningham (2018): *Causal Inference: The Mixtape*, "Directed Acyclical Graphs," pp. 67–80

Morgan and Winship (2014): *Counterfactuals and Causal Inference*

Mullainathan and Spiess (2017): "Machine Learning: An Applied Econometric Approach," *Journal of Economic Perspectives*, 31(2): 87–106

Oster (2019): "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business & Economic Statistics* 37:2 187–204. Also available from Author's website

7. Oct. 25: Statistical Power, Clustering, Experimental Design

Required readings:

Bruhn and McKenzie (2009): "In Pursuit of Balance: Randomization in Practice in Development Field Experiments," *AEJ: Applied*, 1(4): 200–232

McKenzie (2012): "Beyond baseline and follow-up: The case for more T in experiments," *Journal of Development Economics*, 2: 210–221

Duflo, Glennerster, and Kremer (2007): "Using Randomization in Development Economics Research: A Toolkit," *Handbook of Development Economics*, Volume 4, 2007, Chapter 61, pages 3895–3962, available from Elsevier or MIT/CEPR

Recommended readings:

Imbens (2010): "Better LATE Than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009)," *JEL*, 48: 399–423

Gerber and Green (2012): *Field Experiments*, chapters 3 and 4

8. Nov. 1: Randomization Inference, Wild Cluster Bootstrap, Permutation Tests

Required readings:

Young (2019): "Channelling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results," *Quarterly Journal of Economics*, 134(2): 557-598

Recommended readings:

Cameron and Miller (2015): "A Practitioner's Guide to Cluster-Robust Inference," *Journal of Human Resources*, 50(2): 317–372

Cameron, Gelbach, and Miller (2008): "Bootstrap-based improvements for inference with clustered errors," *REStat*, 90(3): 414–427

9. Nov. 8: Multiple Hypothesis Testing

   Required readings:

   Anderson (2008): "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Re-evaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association*, 103(84): 1481–1495

10. Nov. 15: Attrition and Lee Bounds

    Required readings:

    Lee (2009). "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects," *Review of Economics Studies*, 76: 1071—1102

    Recommended readings:

    Gerber and Green (2012): *Field Experiments*, chapter 7

11. Nov. 22: Maximum Likelihood Estimation of Non-linear Models

    Recommended readings:

    Train (2003): *Discrete Choice Methods with Simulation*, chapters 1–6