

Spring 2020
2pm – 4:50pm Wednesdays HOH 506

Professor: Florenta Teodoridis
Office: HOH 520
Office Hours: By appointment
E-mail: florenta.teodoridis@marshall.usc.edu

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: 15%
- Replication report (individual or groups of max two students): 20% (**due any time after February 19 and before April 22**)
- Research paper proposal (individual): 15% (**due February 19**)
- Research paper class presentation (individual): 20% (**April 22 and 29**)
- Research paper extended abstract (individual): 30% (**due May 6**)

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. All readings under the "In Practice" headings are particularly relevant for class participation. They are meant to provoke you to think how you might go about implementing the theoretical empirical approaches in research projects.

Replication Report: You will replicate the main results of an existing empirical paper and discuss (or critique) the robustness of the results. These should be written as referee reports that focus only on the empirical implementation. You can work on this report individuals or in a group of no more than two students.

While there is no page limit, I expect the report to be 1-4 pages of text, 2-5 tables and/or figures, and a Stata .log output file. One to the tables would likely replicate results from the existing paper and one or two more would perhaps show results with the data that are not shown in the paper. If you can communicate the core ideas in less space and less tables, no problem. If you need more space, that's fine too. **The key is to show that you could replicate the main result(s) and that you tried some other specifications to check robustness, and that these other specifications are informed by what we do in class!**

In selecting your papers to replicate and review, check papers in top journals that post their data. There are increasingly more papers published along with their datasets. You can also ask your advisor or a faculty member to share one of their working or published papers and their data. I am happy to share any of mine and certainly don't mind the critique!

Please see me before starting any replication effort. I want to confirm that (a) replication will be feasible and (b) that there is room for robustness tests that can be performed using the techniques we learn in the class.

Your replication report is due any time after February 19 and before April 22. The reason for this is twofold. First, on February 19 we will conduct a replication exercise in class, and I would like for you to go through that experience before engaging in your own replication exercise. Second, I don't want to restrict your exercise to any particular techniques, hence I give you the option to select the time when you think you have learned all the tools you need to write your report. For example, some of you might chose to replicate a difference-in-differences paper with robustness that uses alternative specifications to the same technique. Since we learn this technique early in the class, you could, in principle, turn in your report right after February 19. Conversely, if you would like to focus on projects that include probability models, you will need to wait until we discuss these techniques in class towards the end of the course.

Research paper: Individually, 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.

Books:

Required:

- 1) Angrist, Joshua D., Pischke, Jörn-Steffen. 2009. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press: Princeton NJ.
- 2) Angrist, Joshua D., Pischke, Jörn-Steffen. 2014. Mastering 'Metrics: The Path from Cause to Effect. Princeton University Press: Princeton NJ.

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

- 3) Wooldridge, J. 2010. Econometric Analysis of Cross Section and Panel Data (Second Edition). MIT Press.
- 4) Greene, William .2008. Econometric Analysis 6th edition. Prentice Hall.
- 5) Imbens, W. Guido, Rubin, B. Donald. 2015. Causal Inference for Statistics, Social and Biomedical Sciences: An Introduction. Cambridge University Press

	Date	Topic
1	January 15	Getting Started
2	January 22	Causal effects: Why do we care and why so hard to establish? (I)
3	January 29	Causal effects: Why do we care and why so hard to establish? (II)
4	February 5	Difference-in-Differences (I)
5	February 12	Difference-in-Differences (II)
6	February 19	Replication exercise
7	February 26	Instrumental variables
8	March 4	Matching
9	March 11	Regression Discontinuity
March Break (March 15-22)		
10	March 25	Quantile regression; Event studies
11	April 1	Continuous, discrete and binary dependent variables
12	April 8	Testing
13	April 15	Introduction to machine learning techniques in the context of causal inference
14	April 22	Project Presentations
15	April 29	Project Presentations and Wrap-up

Week 1 (January 15): 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 22): 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
- 3) Heckman, J. 2000. Causal Parameters and Policy Analysis in Economics: A Twentieth Century Retrospective. *Quarterly Journal of Economics*, 115: 45-97.

In practice

- 1) Azoulay, P., Graff Zivin, J, Wang, J. 2010. Superstar Extinction. *Quarterly Journal of Economics*, 125(2): 549-589

Week 3 (January 29): 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

- 1) Bertrand, M., Mullainathan, S. 2004. Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *American Economic Review*, 94(4): 991–1013.

Describing interesting data

- 2) Jones, B. F., Wuchty, S, Uzzi, B. 2008. Multi-University Research Teams: Shifting Impact, Geography, and Stratification in Science. *Science*, 322: 1259-1262.

Week 4 (February 5): Difference-in-differences (I)

Theory: *Difference-in-differences fundamentals*

- 1) MM Chapter 5 (Only 5.1 (all subsections))
- 2) Bertrand, M., Duflo, E. Mullainathan, S. 2004. How Much Should We Trust Differences-in-Differences Estimates?. *Quarterly Journal of Economics*, 119: 249-76.
(Only the introduction is required)
- 3) Donald, S., Lang, K. 2007. Inference with Difference in Differences and Other Panel Data. *Review of Economics and Statistics*, 2: 221-233.

In practice

- 1) Agrawal, A., Goldfarb, A. and Teodoridis, F. 2016. Understanding the Changing Structure of Scientific Inquiry. *American Economic Journal: Applied Economics*, 8(1), 100-128.

Week 5 (February 12): Difference-in-differences (II)

Theory: *Difference-in-differences, fixed effects, lagged variables and controls*

- 1) MHE Chapter 5

In practice

- 1) Stern, S. 2004. Do Scientists Pay to Be Scientists?. *Management Science* 50(6), 835-853.
- 2) Furman, J., Stern, S. 2011. Climbing Atop the Shoulders of Giants: The Impact of Institutions on Cumulative Research. *American Economic Review* 101(5): 1933-1963

Week 6 (February 19): Paper replication exercise

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

Paper TBD

Week 7 (February 26): Instrumental variables

Theory: *IV design (including 2SLS), Language of treatment effects*

- 1) MM Chapter 3
- 2) Imbens, Guido W., and Jeffrey M. Wooldridge. 2009. "Recent Developments in the Econometrics of Program Evaluation." *Journal of Economic Literature*, 47(1): 5–86.
(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

- 1) Angrist, Joshua D. 1990. "Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records." *American Economic Review* 80(3), 313-336.

Week 8 (March 4): Matching

Theory:

- 1) *MHE 3.3
- 2) Todd, P. (2006), "Matching Estimators"
- 3) Iacus, King, and Porro (2011). "Causal Inference without Balance Checking: Coarsened Exact Matching" *Political Analysis*. (Skip section 5)

In practice

- 1) Thompson, P., and M. Fox-Kean (2005): "Patent citations and the geography of knowledge spillovers: a reassessment". *American Economic Review*, 95(1): 450-460.

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.

- 2) Levine, D. I., & Toffel, M. W. (2010). Quality management and job quality: How the ISO 9001 standard for quality management systems affects employees and employers. *Management Science*, 56(6), 978-996. (including the online appendix)
- 3) *Azoulay, P., Stuart, T., & Wang, Y. (2013). Matthew: Effect or fable? *Management Science*, 60(1), 92-109.

Week 09 (March 11): Regression Discontinuity

Theory: Sharp RD, Fuzzy RD

- 1) MHE Chapter 6
- 2) MM Chapter 4

In practice

- 1) Busse, Meghan, Jorge Silva-Risso, and Florian Zettelmeyer. 2006. "\$1,000 Cash Back: The Pass-Through of Auto Manufacturer Promotions." *American Economic Review*, 96(4): 1253–1270.
- 2) Fehder, Dan. 2015. "Startup Accelerators and Ecosystems: Complements or Substitutes?" (job market paper)

(March 15-22: March Break)

Week 10 (March 25): Quantile regression, Event studies

Theory

Event studies

- 1) MacKinlay, A. C. (1997) "Event Studies in Economics and Finance" *Journal of Economic Literature*, 35, 13-39.

Quantile regression

- 2) MHE Chapter 7

In practice

- 1) Oxley, Joanne, Rachelle C. Sampson and Brian Silverman. 2009. "Arms Race or Détente? How Inter-firm Alliance Announcements Change the Stock Market Valuation of Rivals." *Management Science* 55(8): 1321-37
- 2) Goldberg, Pinelopi. 1996. "Dealer Price Discrimination in New Car Purchases: Evidence from the Consumer Expenditure Survey," *Journal of Political Economy*.

Week 11 (April 01): Continuous, discrete and binary dependent variables

Theory

Continuous and discrete dependent variables (OLS and Poisson)

- 1) MHE Chapter 3 (Only 3.1 (all subsections))
- 2) Santos Silva, J. M. C. and Silvana Tenreyro. 2006. "The Log of Gravity." *Review of Economics and Statistics* 88(4): 641-658.
- 3) Hausman, Jerry, Hall, Bronwyn H & Griliches, Zvi, 1984. "Econometric Models for Count Data with an Application to the Patents-R&D Relationship," *Econometrica*, 52(4): 909-38.

Binary dependent variables (Probit, Logit, Tobit)

- 4) MHE Chapter 3 (Only 3.4.2) and Chapter 4 (Only 4.6.3)
- 5) Angrist, Joshua D. 2001. "Estimation of Limited Dependent Variable Models with Dummy Endogenous Regressors: Simple Strategies for Empirical Practice" *Journal of Business and Economic Statistics* 19(1), 2-16.

Interaction terms

(Interpreting interaction terms in linear models)

- 6) Brambor, Thomas, WR Clark, M Golder. 2006. "Understanding Interaction Models: Improving Empirical Analysis" *Political Analysis* 14:63-82
(Calculating interaction terms in nonlinear models)
- 7) Ai, Chunrong, and Edward C. Norton. 2003. "Interaction Terms in Logit and Probit." *Economics Letters* 80, 123-129.

- 8) Zelter BA. 2009. Using Simulation to Interpret Results From Logit, Probit, and Other Nonlinear Models. *Strategic Management Journal* 30(12): 1335-1348

In practice

Count data

- 1) Azoulay, P., Graff Zivin, J, Wang, J. 2010. Superstar Extinction. *Quarterly Journal of Economics*, 125(2): 549-589

Binary Dependent Variables

- 2) Agrawal, Ajay, and Avi Goldfarb. 2008. "Restructuring Research: Communication Costs and the Democratization of University Innovation." *American Economic Review* 98(4), 1578-1590.
- 3) Forbes, Silke Januszewski, and Mara Lederman. 2009. "Adaptation and Vertical Integration in the Airline Industry." *American Economic Review*, 99(5): 1831-49.

Week 12 (April 08): Testing

Theory

Economic vs. statistical significance

- 1) McCloskey D, Ziliak S T. (1996). "The Standard Error of Regressions." *Journal of Economic Literature* 34(1): 97-114.

Random effects and Hausman tests

- 2) Hausman, Jerry A., and William E. Taylor. 1981. Panel Data and Unobservable Individual Effects. *Econometrica* 49(6), 1377-1398.

In practice

Economic vs. statistical significance

- 1) Milyo, J., and J. Waldfogel. 1999. "The Effect of Price Advertising on Prices: Evidence in the Wake of 44 Liquormart," *American Economic Review* 89: 1081-96
- 2) Kuhn, P., Skuterud, M. (2004), "Internet job search and unemployment durations", *American Economic Review*, 94(1), 218-32.
- 3) Kyle, Margaret, and Anita McGahan. 2012. Investments in Pharmaceuticals Before and After TRIPS. *Review of Economics and Statistics* 94(4):1157-1172

Random effects and Hausman tests

- 4) Evans, William N., and Ioannis N. Kessides. 1993. "Localized Market Power in the U.S. Airline Industry." *Review of Economics and Statistics* 75(1), 66-75.

Fit, R-squared, and explanation vs. prediction

- 5) Athey, Susan and Scott Stern. 2002. "The Impact of Information Technology on Emergency Health Care Outcomes," *RAND Journal of Economics*, 33(3): 399-432.
- 6) Donohue, John J. III, and Stephen D. Levitt. 2001. "The Impact of Legalized Abortion on Crime." *Quarterly Journal of Economics* 116(2): 379-420.

Week 13 (April 15): Introduction to machine learning techniques in the context of causal inference

Theory

- 1) Choudhury, Prithwiraj, Ryan Allen, and Michael G. Endres. "Implementation, Evaluation, and Interpretation of Machine Learning Methods for Pattern Detection", working paper

In practice

- 1) Furman, J., Teodoridis, F. 2019. "Automation, Research Technology and Researchers' Trajectories: Evidence from Computer Science and Electrical Engineering", *Organization Science*, forthcoming
- 2) Choudhury, Prithwiraj, Dan Wang, Natalie A. Carlson, and Tarun Khanna. "Machine Learning Approaches to Facial and Text Analysis: Discovering CEO Oral Communication Styles." *Strategic Management Journal*, forthcoming

Week 14 & 15 (April 22 and 29): Project presentations