

## Empirical Methods II

### Contacts

Dr. Osea Giuntella

**Office:** Room 4118

**Office Hours:** Thursday 3.45pm-5.15pm or appointment by email

**Email:** osea.giuntella@pitt.edu

**Personal Website:** [www.oseagiuntella.com](http://www.oseagiuntella.com)

### Course Description

This course will cover cutting-edge methods typically used in applied micro-economic research and it is designed to help you learn how to apply the econometric techniques learned in the econometrics sequence. It is a natural prosecution of the empirical methods class. The course will cover methods for spatial analysis introducing students to the use of ArcGIS for applied microeconomic research. Furthermore, it will introduce students to scraping and data mining techniques using Python and R. The goal of this class is to provide students with the tools needed to become critical readers of empirical work and teach them techniques that they can apply to their own original research.

### Requirements and Grading

I expect you to read the papers in the reading list before class and to be prepared to discuss them. Please bring a copy (paper or electronic) of the papers we are discussing with you to class. 20% of your grade will be based on the problem sets. 20% of your grade will be based on a verbal presentation of a paper. You should briefly summarize the paper (touching on data, methods, and findings) and then point out any major issues with the approach or conclusions. Finally, 60% of your grade will be based on the replication and extension of an empirical project. The paper replication must be turned in on time.

### Problem Sets (20%)

There will be 4 problem sets. The problem sets will be data exercises that will allow you to practice the econometric techniques covered in class and familiarize with STATA and ArcGis. If you don't have it yet, you should obtain a copy of **STATA** and **ArcGIS** through mypitt.edu:

- STATA: <http://technology.pitt.edu/software/stata-for-students>
- ArcGIS: <http://technology.pitt.edu/software/arcgis-for-students>

Each problem set will count for 5% of your grade. Time management is an important part of this exercise, so there will be no extensions.

## **Presentation (20%)**

Presenting work at conferences, meetings, seminars is a key part of our profession. Each student will be asked to critically discuss one of the papers in the reading list. You should briefly summarize the content of the paper and then point out any major limitation with the contribution, identification, interpretation, or generalizability of the results. You need to select a topic (and you can express a preference for a paper) no later than EB 11. Presentation will follow the schedule of each topic (see below).

## **Empirical Project (60%)**

You are expected to replicate and extend a research paper on a topic you choose. I recommend you choose a published paper which uses publicly available data. Most top journals in Economics make data freely available on their websites. The paper should apply one of the methods covered in the course. You can extend the paper by changing the outcome of interest, the sample (years, country etc.), or improving on the identification strategy. I will also provide a list of potential papers that you may replicate and extend. You will need to have the topic approved by me no later than Feb 4. You will need to turn in an outline or sketch of the project no later than March 4. This will count for 10% of your grade. The final draft is due no later than April 19 and should be no longer than 20 pages (including text and tables). The final draft will count for 40% of your grade. The presentation of your work will count 10% of your grade and will occur during the last week of class (20 minute talk). Time management is an important part of this exercise, so there will be no extensions.

## **Outline**

### Topic 1: Statistical Learning

1. Machine Learning and Big Data in Econ Research (Jan 7)
2. Introduction to Statistical Learning (Jan 9-14)
3. Linear Models (Jan 23)
4. Classification (Jan 25)
5. Resampling methods (Feb 4)
6. Linear model selection and regularization (Feb 6)

7. Moving beyond linearity (Feb 11)
8. Tree-based methods (Feb 13)
9. Support vector machines (Feb 18)
10. Unsupervised learning (Feb 20)

Topic 2: Spatial Analysis

- Intro to Arc Gis for Spatial Analysis (Feb 8)
- Arc Gis for Spatial Analysis (Feb 15)

Topic 2: ML Applications

- Feb 25
- Feb 27
- March 4
- March 6

Topic 2: Microeconometrics and MATLAB

- Optimization problems (March 18-20)
- Discrete choice models: applications (March 25-27)
- Dynamic choice: applications (April 1- April 3)

Topic 3: Mediation analysis (April 8-10)

Topic 4: Coping with Selection Bias:

- Using selection on observables (AET, 2005; Oster, 2017)
- Imperfect IV: plausibly exogenous instruments (April 15)

Topic 5: Testing: Hypothesis Testing in Quasi-Experimental Design (April 17)

- Multiple Hypothesis Testing
- Multiple Hypothesis Testing Correction Methods
- Pre-registering Trials

## General Readings and Textbooks

There is no required textbook for this course. Below a list of textbooks than can be used as a useful reference for the course.

- An Introduction to Statistical Learning: James, Gareth, et al. An introduction to statistical learning. Vol. 112. New York: springer, 2013. (available for free online)  
:<http://www-bcf.usc.edu/~gareth/ISL/>
  - Adams, Abi, Damian Clarke, and Simon Quinn. Microeconometrics and MATLAB: An Introduction. Oxford University Press, 2015.
  - Wooldridge, Jeffrey M. Econometric Analysis of Cross Section and Panel Data. MIT Press, 2010, 2002 (Second Edition).
- (W)

## Reading List

### Machine Learning

1. Athey, Susan. "The impact of machine learning on economics." The Economics of Artificial Intelligence: An Agenda. University of Chicago Press, 2018. (covered on Jan 7-Jan 9)
2. Oster, Emily. "Diabetes and diet: purchasing behavior change in response to health information." American Economic Journal: Applied Economics 10.4 (2018): 308-48. (covered on Jan 9)
3. Mullainathan, Sendhil, and Jann Spiess. "Machine learning: an applied econometric approach." Journal of Economic Perspectives 31.2 (2017): 87-106. (covered on Jan 9-Jan 14)
4. Atalay, Enghin, et al. "New Technologies and the Labor Market." Journal of Monetary Economics (2018). (covered on Jan 14)
5. Wu, Alice H. "Gender stereotyping in academia: Evidence from economics job market rumors forum." (2017). (covered on Jan 25)
6. Nickerson, David W., and Todd Rogers. "Political campaigns and big data." Journal of Economic Perspectives 28.2 (2014): 51-74.
7. Andini, Monica, et al. "Targeting with machine learning: An application to a tax rebate program in Italy." Journal of Economic Behavior Organization 156 (2018): 86-102.

8. Beattie, Graham, Jean-William P. Lalibert, and Philip Oreopoulos. "Thrivers and divers: Using non-academic measures to predict college success and failure." *Economics of Education Review* 62 (2018): 170-182.
9. Bonica, Adam. "Inferring RollCall Scores from Campaign Contributions Using Supervised Machine Learning." *American Journal of Political Science* 62.4 (2018): 830-848.
10. Kleinberg, Jon, et al. "Human decisions and machine predictions." *The quarterly journal of economics* 133.1 (2017): 237-293.
11. Kleinberg, Jon, et al. "Algorithmic fairness." *AEA Papers and Proceedings*. Vol. 108. 2018.
12. Sansone, Dario. "Beyond Early Warning Indicators: High School Dropout and Machine Learning." (2017).
13. McKenzie, David, and Dario Sansone. *Man vs. machine in predicting successful entrepreneurs: evidence from a business plan competition in Nigeria*. The World Bank, 2017.
14. Chalfin, Aaron, et al. "Productivity and selection of human capital with machine learning." *American Economic Review* 106.5 (2016): 124-27.
15. Jean, Neal, et al. "Combining satellite imagery and machine learning to predict poverty." *Science* 353.6301 (2016): 790-794.
16. Luca, Michael, Jon Kleinberg, and Sendhil Mullainathan. "Algorithms need managers, too." *Harvard business review* 94.1 (2016): 20.
17. Kleinberg, Jon, et al. "Prediction policy problems." *American Economic Review* 105.5 (2015): 491-95.

### **Structural Methods**

1. Low, Hamish, and Costas Meghir. "The use of structural models in econometrics." *Journal of Economic Perspectives* 31.2 (2017): 33-58.
2. Keane, Michael P. "A structural perspective on the experimentalist school." *Journal of Economic Perspectives* 24.2 (2010): 47-58.
3. Adda, Jerome, Christian Dustmann, and Katrien Stevens. "The career costs of children." *Journal of Political Economy* 125.2 (2017): 293-337.
4. Chan, Tat Y., Barton H. Hamilton, and Nicholas W. Papageorge. "Health, risky behaviour and the value of medical innovation for infectious disease." *The Review of Economic Studies* 83.4 (2015): 1465-1510.

5. Papageorge, Nicholas W., and Kevin Thom. Genes, education, and labor market outcomes: evidence from the health and retirement study. No. w25114. National Bureau of Economic Research, 2018.
6. Eckstein, Zvi, Michael Keane, and Osnat Lifshitz. "Career and Family Decisions: Cohorts born 1935-1975." *Econometrica* (2018).

### Spatial Analysis

1. Kudamatsu "GIS for credible identification strategies in economics research." *CESifo Economic Studies* 64.2 (2018): 327-338.
2. Alsan, Marcella. 2015. "The Effect of the TseTse Fly on African Development." *American Economic Review*, 105(1): 382-410.
3. Donaldson, Dave, and Adam Storeygard. 2016. "The View from Above: Applications of Satellite Data in Economics." *Journal of Economic Perspectives*, 30(4): 171-198.
4. Conley, Timothy G., and Christopher R. Udry. 2010. "Learning about a New Technology: Pineapple in Ghana." *American Economic Review*, 100(1): 356-9.
5. Nunn, Nathan. 2008. "The Long-Term Effects of Africa's Slave Trades." *Quarterly Journal of Economics*, 123(1): 139-176.
6. Michalopoulos, Stelios. 2012. "The Origins of Ethnolinguistic Diversity." *The American Economic Review*, 102(4): 1508-1539.
7. Duflo, Esther, and Rohini Pande. 2007. "Dams." *Quarterly Journal of Economics*, 122(2): 601-646.
8. Dell, Melissa. 2010. "The Persistent Effects of Peru's Mining Mita." *Econometrica*, 78(6): 1863-1903.
9. Mayshar, Joram, Omer Moav, Zvika Neeman, and Luigi Pascali. 2015. *Cereals, Appropriability and Hierarchy*. CEPR Discussion Paper, no. 10742.

### Mediation Analysis

1. Heckman J, Pinto R. *Econometric Mediation Analyses: Identifying the Sources of Treatment Effects from Experimentally Estimated Production Technologies with Unmeasured and Mismeasured Inputs*. *Econom Rev*. 2014;34(1-2):6-31.
2. Keele, Luke, Dustin Tingley, and Teppei Yamamoto. "Identifying mechanisms behind policy interventions via causal mediation analysis." *Journal of Policy Analysis and Management* 34.4 (2015): 937-963.
3. Imai, Kosuke, et al. "Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies." *American Political Science Review* 105.4 (2011): 765-789.

4. Pearl, Judea. "The causal mediation formulaa guide to the assessment of pathways and mechanisms." *Prevention science* 13.4 (2012): 426-436.
5. Tingley, Dustin, et al. "Mediation: R package for causal mediation analysis." (2014).

## Other Useful Resources

- [Melissa Dell's class notes: GIS Analysis for Applied Economists](#) (Very comprehensive, good reference for definitions of ArcGIS tools, packages, etc., no exercises to replicate)
- [Masayuki Kudamatsu's class notes: ArcGIS 10 for Applied Microeconometric Research](#) (Set of 7 lectures with replication exercises and corresponding data for replication, great learning resource)
- Ogura, L. M. 2010. [Template-Based Introductory Guide to LaTeX for Economics, Grand Valley State University.](#)
- Quick Stata Guide, Liz Foster: <http://www.princeton.edu/wwac/academic-review/stata/>
- A guide to ArcGis: <https://sites.google.com/site/mkudamatsu/gis>

## **Academic Integrity**

Cheating/plagiarism will not be tolerated. Students suspected of violating the University of Pittsburgh Policy on Academic Integrity, from the February 1974 Senate Committee on Tenure and Academic Freedom reported to the Senate Council, will be required to participate in the outlined procedural process as initiated by the instructor. A minimum sanction of a zero score for the quiz or exam will be imposed.

## **Disability Resource**

If you have a disability for which you are or may be requesting an accommodation, you are encouraged to contact both your instructor and the Office of Disability Resources and Services, 140 William Pitt Union, 412-648-7890/412-624-3346 (Fax), as early as possible in the term. Disability Resources and Services will verify your disability and determine reasonable accommodations for this course. For more information, visit [www.studentaffairs.pitt.edu/drsabout](http://www.studentaffairs.pitt.edu/drsabout).