PS2030 Political Research and Analysis

Michaël Aklin

Spring 2020

E-mail: aklin@pitt.edu
Office Hours: TBC
Office: 4815 Posvar Hall

Web: michaelaklin.github.io
Class Hours: Monday 9-11:55am
Class Room: 4430 Posvar Hall

Course Description

This course follows PS2010 and PS2020. It is an introduction to the linear model (and ordinary least squares) and what we can learn from it. The course ends with an introduction to nonlinear models and maximum likelihood.

Note: this is a live document, and I'll update it when I come across topics/papers that should be added.

Here's the link to the dropbox folder that contains material for this course: https://www.dropbox.com/sh/swfqzzgbdo4pcsn/AAC_eEvErFqoNjIsBVGSWgOha?dl=0.

Course Objectives

Requirements

1. Five homework (5x10%). They need to be completed in RMarkdown or Stata Markdown.
2. A final exam (50%). In-class, open book, no internet.

Grading follows the usual scale (http://www.pitt.edu/~graduate/reggrades.html).

Key dates (note: these dates can change; all changes will be announced in class.)

- HW1: handed out on Feb 3 (every homework is due one week later by 9am)
- HW2: handed out on Feb 17
- HW3: handed out on March 2
- HW4: handed out on March 23
- HW5: handed out on April 6
- Final exam: in class, April 20.

Textbooks

Below, I link each topic to a chapter by Wooldridge (henceforth **W**):

Wooldridge, Jeffrey M. (2012). *Introductory Econometrics*. 5th. South-Western Cengage Learning.

This being said: (a) feel free to pick an earlier version if it's cheaper; (b) you should select whatever textbook you're most comfortable with. There are many alternative textbooks which have all their strengths and weaknesses. Hayashi is a deeper theoretical textbook, Wooldridge provides a good overview of panel data, Gailmard provides a very good overview that connects probability theory, statistics, and econometrics; and Greene and Cameron and Trivedi are useful references. Gelman and Hill are a great resource that bridges several areas of quantitative methods.

Cameron, A. Colin and Pravin K. Trivedi (2005). *Microeconometrics: Methods and Applications*. London: Cambridge University Press.

Gailmard, Sean (2014). *Statistical Modeling and Inference for Social Science*. New York: Cambridge University Press.

Gelman, Andrew and Jennifer Hill (2007). *Data Analysis Using Regression and Multilevel–Hierarchical Models*. Cambridge: Cambridge University Press.

Greene, William H. (2008). *Econometric Analysis*. 6th. New York: Pearson.

Hayashi, Fumio (2000). *Econometrics*. Princeton: Princeton University Press.

Wooldridge, Jeffrey M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT press.

There exist good free resources. These include:

- Scott Cinnungham's "mixtape": https://scunning.com/cunningham_mixtape.pdf.
- Bruce Hansen's textbook: https://www.ssc.wisc.edu/~bhansen/econometrics/Econometrics.pdf

The following two references are harder to find substitute for, and I would encourage you to get your hands on a copy (from now on: **AP** and **MW**):

Angrist, Joshua and Steffan J. Pischke (2008). *Mostly Harmless Econometics*. Princeton: Princeton University Press.

Morgan, Steven L. and Christopher Winship (2014). *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. 2nd Edition. Cambridge: Cambridge University Press.

Misc

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. View the complete policy at www.cfo.pitt.edu/policies/policy/02/02-03-02.html.

Class Schedule

This is not a weekly schedule. Each topic can last more (rarely less) than a week.

Introduction

Introduction to the class. Potential outcome framework. Causal identification. Treatment effects. Math, statistics, and R/Stata review.

To read: **MW**, 1.

Regression - basics

Linear model. Ordinary least squares. Bivariate regression.

To read: **W**, 2.

Regression - advanced

Multiple regression. Mechanics of regression. Frisch–Waugh–Lovell theorem. Properties of estimators (bias, efficiency, constitency)

To read: **AP**, 3. **W**, Appendix D, E.1, 3 up to to 3.3 (included), Appendix 3A.

Inference in the linear model

Sampling distribution. Standard errors (robust, clustered) and heteroskedasticity. Hypothesis testing. Statistical power. P-values. Multiple comparison. Confidence intervals. F statistic. R^2. Review of the assumptions of regression.

To read: **W**, Appendix C.5, C.6, E.2, 3.4, 4, 8 to 8.3.

Articles:

Cameron, A Colin and Douglas L Miller (2015). "A practitioner's guide to cluster-robust inference". In: *Journal of Human Resources* 50.2, pp. 317-372.

Greenland, Sander, Stephen J. Senn, Kenneth J. Rothman, John B. Carlin, Charles Poole, Steven N. Goodman, and Douglas G. Altman (2016). "Statistical Tests, P Values, Confidence Intervals, and Power: A Guide to Misinterpretations". In: *Eur J Epidemiol* 31, pp. 337-350.

Problems and Solutions: Model Specification

Specifying a functional form. Outliers. Nonlinear effects. Interaction effects. Time permitting: matching, semi-parametric regression.

To read: **W**, 6.2, 7 up to 7.4 (included).

Articles:

Brambor, Thomas, William R. Clark, and Matt Golder (2006). "Understanding Interaction Models: Improving Empirical Analyses". In: *Political Analysis* 14.1, p. 63.

Hainmueller, Jens, Jonathan Mummolo, and Yiqing Xu (2019). "How Much Should We Trust Estimates from Multiplicative Interaction Models? Simple Tools to Improve Empirical Practice". In: *Political Analysis* 27.2, pp. 163–192.

Problems and Solutions: Omitted Variable Bias and Endogeneity

Endogeneity. Omitted variable bias. Randomized controlled trial. Instrumental variables and two-stage least squares. Regression discontinuity.

To read: **W**, 15. **MW**: 4, 9. **AP**: 1, 4, 6.

Articles:

RCT: Paler, Laura (2013). "Keeping the public purse: An experiment in windfalls, taxes, and the incentives to restrain government". In: *American Political Science Review* 107.4, pp. 706-725.

Instrumental variables: Erikson, Robert S and Laura Stoker (2011). "Caught in the draft: The effects of Vietnam draft lottery status on political attitudes". In: *American Political Science Review* 105.2, pp. 221-237.

Regression discontinuity: Dell, Melissa (2010). "The Persistent Effects of Peru's Mining Mita". In: *Econometrica* 78.6, pp. 1863-1903.

Problems and Solutions: Temporal Variation

Serial correlation. Cross section and time series. Fixed effects. Differences-in-differences.

To read: **W**, 10 (optional: **W** 12, 13). **AP**, 5. **MW**, 11.

Articles:

Differences-in-differences: Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan (2004). "How Much Should We Trust Differences-In-Differences Estimates?" In: *Quarterly Journal of Economics* 119.1, pp. 249-275.

Dube, Arindrajit, Oeindrila Dube, and Omar Garcia-Ponce (2013). "Cross-Border Spillover: U.S. Gun Laws and Violence in Mexico". In: *American Political Science Review* 107.3, pp. 397-417.

Nonlinear Models

Linear probability model. Logit. Probit. Ordered logit. Maximum likelihood.

To read: **W**, 17 up to 17.1 (included)

Problems and Solutions: Missing Observations

Missing observations. Selection issues.

To read: **W**, 17. **MW**, 4.3.2.