ECON 629 – Applied Econometric Methods for Empirical Research and Policy Evaluation

Instructor: Fernando Rios-Avila
Office: Room 307, Blithewood
Office hours: By appointment
Email: friosavi@levy.org
Website: https://riosavila.github.io/.

Course Description

This course is designed to introduce students to econometric methods used in empirical research, with emphasis on the analysis and identification of causal effects and Policy Evaluation. It aims to provide an overview and general guidelines on practical implementation, discussing the intuition behind the empirical methods.

The course is divided in two sections. The first section focuses on a review of advanced estimation methods for the analysis of data under the linearity assumption: ordinary least squares, parametric and semiparametric model estimation, quantile regression, and RIF regressions. We also review the use of maximum likelihood estimation (MLE), and Nonlinear Least Squares (NLS) as methodologies that allows the estimation of both linear and nonlinear models.

In the second part of the course, we re-introduce the problem of identification of causal effects, the role of experimental designs, and the idea of potential outcome models, discussing the main limitations on the identification of causal effects. We follow by discussing various methods focused on the identification of causal effects including, fixed effects / Panel data models, matching and reweighting, instrumental variables, Regression Discontinuity, Differences in Differences and extensions, and Synthetic controls. In all cases, we discuss limitations and interpretation of treatment effects.

By the end of the course students should have a good understanding of the estimation of causal effects, as well as the main methods used to address this problem, and critically evaluate and interpret the output of such analyses.

Course outline

This course syllabus provides a general plan for the course; deviations may be necessary:

Part I: Advanced Estimation Methods

1. Introduction: General Overview of the course
2. Linear Regression:

- Estimation, Assumptions, Inference
- Deviation from Main Assumptions
- Alternative Approaches: Robust-Standard errors, Cluster Standard errors, Bootstrap Standard error.
- Model Selection, and Shrinkage Methods

    - MHE – Chapter 3
    - CI – Chapter 2
    - MMA - Chapter 11
    - ISL – Chapter 3, Chapter 6.1 – 6.2
    - MacKinnon, et al (2023) "Cluster-Robust Inference: A Guide to Empirical Practice" Journal of Econometrics

3. Parametric, non-parametric and Semi-parametric Regressions:

- Difference between Parametric and Non-Parametric models

- Distributions, densities and histograms

- Advantages and disadvantages of non-parametric models (bias vs variance)

- Estimation: Kernel vs Splines; Semi - Parametric vs non-parametric

    - MMA Chapter 9
    - ISL Chapter 7
    - Henderson and Parmeter 2016 Teaching Nonparametric Econometrics to Undergraduates
    - Racine, 2008 Nonparametric Econometrics: A Primer
    - Rios-Avila, 2020 Smooth Varying-coefficient Models in Stata
    - Verardi and Debarsy 2012 Robinson's square root of N consistent semiparametric regression estimator in Stata

4. Going Beyond the Mean: [conditional] Quantile Regression

- Difference between Linear Regression and CQuantile Regression

- Estimation and Statistical Inference

- Quantile Regression: Modeling Heteroskedasticity
- Quantile Regression: As a semiparametric model with unobserved Heterogeneity

  – MHE Chapter 7
  – Koenker and Hallok (2001) Quantile Regression
  – Machado and Santos Silva (2019) Quantiles via moments
  – Wenz (2018) What does QR does and Doesn't do
  – Castro et al (2019) Smoothed GMM for quantile models

5. Going Beyond the Mean II: RIF – regressions

- Influence Functions and Recentered Influence functions
- Usage with distributional Statistics
- RIF-Regressions: Unconditional Quantile Regressions, Partially Conditional Quantile Regressions

  – Rios-Avila and Maroto (2022) Moving Beyond Linear Regression: Implementing and Interpreting Quantile Regression Models With Fixed Effects
  – Borah and Basu (2013) Highlighting Differences Between Conditional and Unconditional Quantile Regression Approaches Through An Application To Assess Medication Adherence
  – Firpo Fortin and Lemieux (2009) Unconditional Quantile Regressions
  – Rios-Avila (2020) Recentered influence functions (RIFs) in Stata: RIF regression and RIF decomposition

6. Maximum Likelihood Estimation

- Linear vs Non-Linear models
- Non-Linear Least Squares
- MLE for Logit/probit/poisson
- MLE for heckman/tobit
- MLE for other NL models

  – MMA Chapter 4
  – EA Chapter 7, 14 (Greene)
  – Rios-Avila (2018) Standard-error correction in two-stage optimization models: A quasi–maximum likelihood estimation approach
  – Kapteyn and Ypma (2007) Measurement Error and Misclassification: A comparison of Survey and Administrative Data.
  – Ransom (1987) An Empirical Model of Discrete and Continuous Choice in Family Labor Supply

Part II: Causal Inference

7. Potential Outcomes Causal Model

- Assumptions, Implications and Imputations
- Randomized Control Trials
- Regression in Experimental settings

    – MHE Chapter 1
    – CI Chapter 4
    – IEP Chapter 3 – 4
    – MM Chapter 1-2
    – TE Chapter 10 - 11

8. Panel Data and Fixed Effects

- Random Effects vs Fixed effects

- Correlated Random Effects Model

- 2 or more FE: Frisch-Waugh-Lovell Theorem

    – TE Chapter 16
    – MHE Chapter 5
    – CI Chapter 8
    – Rios-Avila (2015) Feasible fitting of linear models with N fixed effects
    – Correia (2016) A Feasible Estimator for Linear Models with Multi-Way Fixed Effects
    – Wooldridge (2019) Correlated random effects models with unbalanced panels
    – Machado & Santos Silva (2019) Quantiles via moments

9. Instrumental Variables - IV

- Theory, assumptions and Limitations

- Randomization and Natural Experiments

- Heterogeneity and Local Average Treatment Effects

- Leading Examples and Limitations

    – CI Chapter 7
    – MHE Chapter 4
    – MM Chapter 3
    – TE Chapter 19

10. Matching and Reweighting

- Stratification, Exact Matching, Approximate Matching

- Curse of Dimensionality

- Weighting and Re-weighting Analysis

  – CI Chapter 5
  – TE Chapter 14
  – IEP Chapter 8
  – Austin (2011) An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies
  – King and Nielsen (2019) Why Propensity Scores Should Not Be Used for Matching
  – Caliendo et al (2008) Some Practical Guidance for the Implementation Of Propensity Score Matching

11. Regression Discontinuity

- Identification, Assumptions and Limitations

- Leveraging threshold: Estimation – Parametric and non-parametric

- Sharp vs Fuzzy RD

  – CI Chapter 6
  – MHE Chapter 6
  – TE Chapter 20
  – MM Chapter 4
  – Cattaneo & Titiunik 2022 Regression Discontinuity Designs
  – Lee & Lemieux 2010 Regression Discontinuity Designs in Economics

12. Differences in Differences

- Assumptions and Identification of Canonical 2x2 DID

- Allowing for Covariates

- Generalized DID, TWFE, and Limitations

- Event Studies

- Timing and cohort Heterogeneity

  – CI Chapter 9
  – MHE Chapter 5
  – TE Chapter 18
  – MM Chapter 5
  – Roth et al (2022) What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature

– Chaisemartin & D'Haultfeuille 2022 Two-Way Fixed Effects and Differences-in-Differences with Heterogeneous Treatment Effects: A Survey

13. Synthetic Control

- Identification and Estimation
- Internal vs external Validity
- Exact Statistics and Extensions

  – CI Chapter 10
  – TE Chapter 21
  – Abadie (2020), "Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects," JEL

Suggested Readings

MHE - Mostly Harmless Econometrics: An Empiricist's Companion
by Joshua D. Angrist and Jorn-Steffen Pischke

CI - Causal Inference: The Mixtape
By Scott Cunningham

TE - The Effect: An Introduction to Research Design and Causality
By Nick Huntington-Klein

MM - Mastering Metrics: The path from Cause to Effect
by Joshua D. Angrist and Jorn-Steffen Pischke

MMA - Microeconometrics: Methods and Applications
By Colin Cameron and Pravin Trivedi (Selected Chapters)

EA - Econometric Analysis
By William Greene (Selected Chapters)

ISL - An Introduction to Statistical Learning By Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani (Selected Chapters)

Further Readings

Potential Outcomes Causal Model

Blattman, Christopher, Nathan Fiala, and Sebastian Martinez. 2014. "Generating Skilled Self-Employment in Developing Countries: Experimental Evidence from Uganda." Quarterly Journal of Economics 129 (2): 697–752. doi: 10.1093/qje/qjt057.

Schultz, Paul. 2004. "School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program." Journal of Development Economics 74 (1): 199–250

Manning, Willard G., et al. "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment." The American Economic Review, vol. 77, no. 3, 1987, pp. 251–77. JSTOR, http://www.jstor.org/stable/1804094. Accessed 18 Jan. 2023.

Panel Data and Fixed Effects

Cornwell, C. And Rupert, P. (1997), Unobservable Individual Effects, Marriage And The Earnings Of Young Men. Economic Inquiry, 35: 285-294. Https://Doi.Org/10.1111/J.1465-7295.1997.Tb01910.X

Freeman, R. B. (1984). Longitudinal Analyses of the Effects of Trade Unions. Journal of Labor Economics, 2(1), 1–26. http://www.jstor.org/stable/2535015Instrumental Variables - IV

Ashenfelter, O., & Rouse, C. (1998). Income, Schooling, and Ability: Evidence from a New Sample of Identical Twins. The Quarterly Journal of Economics, 113(1), 253–284. http://www.jstor.org/stable/2586991

Sandra E. Black, Paul J. Devereux, Kjell G. Salvanes, From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes, The Quarterly Journal of Economics, Volume 122, Issue 1, February 2007, Pages 409–439, https://doi.org/10.1162/qjec.122.1.409

Instrumental Variables - IV

Cygan-Rehm, K., & Wunder, C. (2018). Do working hours affect health? Evidence from statutory workweek regulations in Germany. Labour Economics, 53, 162-171. https://doi.org/10.1016/j.labeco.2018.05.003

Bulman, George, Robert Fairlie, Sarena Goodman, and Adam Isen. 2021. "Parental Resources and College Attendance: Evidence from Lottery Wins." American Economic Review, 111 (4): 1201-40.

Tsai, A. C., & Venkataramani, A. S. (2015). The causal effect of education on HIV stigma in Uganda: Evidence from a natural experiment. Social Science & Medicine, 142, 37-46. https://doi.org/10.1016/j.socscimed.2015.08.009

Koppel, Stephen, Bergin, Tiffany, Ropac, René, Randolph, Imani, & Joseph, Hannah. (2022). Examining the causal effect of pretrial detention on case outcomes: a judge fixed effect instrumental variable approach. Journal of Experimental Criminology.

Matching and Reweighting

Jyotsna Jalan & Martin Ravallion (2003) Estimating the Benefit Incidence of an Antipoverty Program by Propensity-Score Matching, Journal of Business & Economic Statistics, 21:1, 19-30, DOI: 10.1198/073500102288618720 Arceneaux, K., Gerber, A., & Green, D. (2006). Comparing experimental and matching methods using a largescale voter mobilization experiment. Political Analysis, 14, 1-26.

Ren Mu & Dominique van de Walle (2011) Rural Roads and Local Market Development in Vietnam, The Journal of Development Studies, 47:5, 709-734, DOI: 10.1080/00220381003599436

Regression Discontinuity

Card, D., Dobkin, C., & Maestas, N. (2008). The Impact of Nearly Universal Insurance Coverage on Health Care Utilization: Evidence from Medicare. The American Economic Review, 98(5), 2242–2258. http://www.jstor.org/stable/29730170

Carpenter, Christopher, and Carlos Dobkin. 2009. "The Effect of Alcohol Consumption on Mortality: Regression Discontinuity Evidence from the Minimum Drinking Age." American Economic Journal: Applied Economics, 1 (1): 164-82.

Bleemer, Zachary, and Aashish Mehta. 2022. "Will Studying Economics Make You Rich? A Regression Discontinuity Analysis of the Returns to College Major." American Economic Journal: Applied Economics, 14 (2): 1-22.

Differences in Differences

Nollenberger, Natalia and Rodriguez-Planas, Núria, (2015), Full-time universal childcare in a context of low maternal employment: Quasi-experimental evidence from Spain, Labour Economics, 36, issue C, p. 124-136

Gousse, Marion and Leturcq, Marion, (2022), More or Less unmarried. The impact of legal settings of cohabitation on labour market outcomes, European Economic Review 149 (2022)

Hotchkiss, J. L., Moore, R. E., & Rios-Avila, F. (2014). Reevaluation of the Employment Impact of the 1996 Summer Olympic Games. Southern Economic Journal, 140619071514006. doi:10.4284/0038-4038-2013.063

Synthetic Control

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. Journal of the American Statistical Association, 105(490), 493–505. doi:10.1198/jasa.2009.ap08746

Peri, G., & Yasenov, V. (2018). The Labor Market Effects of a Refugee Wave: Synthetic Control Method Meets the Mariel Boatlift. Journal of Human Resources, 0217_8561R1.

Viana, J. H., Barbosa, A. V., & Sampaio, B. (2018). Does the World Cup get the economic ball rolling? Evidence from a synthetic control approach. EconomiA, 19(3), 330-349. https://doi.org/10.1016/j.econ.2018.05.001

Course website

We will be using `Github` for all materials, presentations and announcements. The course website is located at https://riosavila.github.io/.

Attendance:

Class attendance is highly recommended, but not compulsory. Material for homeworks will come from class lectures.

Course Software

There are several statistical packages for analyzing data. In this course, I will be using `Stata` to cover all materials in class. The Institute will be providing you with licenses for Stata/BE for length of the course. However, you may also use other-software of your liking. Two of the books used for the course, Causal Effect and The Effect, provides codes for most of the material we cover here using `Stata`, `Python` and `R`. `Julia` is also a viable option.

`Stata` offers many free short webinars and video tutorials that may be useful if you never used Stata before, or even if you have some experience with it. See the resources for links to the webinars offered by `Stata`.

There are also a series of Cheat sheets to remember basics of Data processing, analysis, and visualization here stata-cheat-sheets

Additional Information:

All students are responsible for knowing Bard's Policy on Academic Honesty as publish in Bard College Student Handbook.

Grading

- Presentations 20%: Students will have to do a series of presentations during the second half of the semester based on the suggested material (further reading) or other papers the students may be interested in.

The main requirement. The paper should implement any of the methodologies we will be covering in class. The presentation should emphasize the Research question, assumptions used, methodology, and results. If possible, provide comments and suggestions to extend the paper.

- Homework 40%: Homeworks will be provided for you to practice and implement the different methodologies discussed in class. They can be carried out individually or in groups. This will include making a brief description of the results.

- Paper Project 40: Write a term paper that can be of two types:

  - **Paper Replication**: You can choose to write a replication paper on a methodological paper, or applied empirical paper.

In either case, the replication will have to extend the analysis of the original paper to a different setup (empirical paper), different software (if replication paper), or other extensions to the original analysis/methodology.

  - **Original Research:** A 20-25 pages paper where students answer a research question of their choice, using any of the methodologies presented in class. Standard structure of the paper applies.

10% of the grade will be based on a presentation of the paper.