Microeconometrics

Dmitry Arkhangelsky
darkhangel@cemfi.es

Fall 2024

Logistics

Schedule

Lectures: Tue 9:30-11:00, 11:30 - 13:00;
Exercises: Mon 9:30-11:00;
Office hours: By appointment
TA: Wisse Rutgers

Home assignments

Home assignments will include theoretical, computational, and empirical exercises. I strongly encourage you to use R for the coding problems. Home assignments should be done in groups with a minimal size of two and a maximal size of three students.

Grading

Grades will be based on group home assignments (10%), presentations (10%), and a final exam (80%). The final exam will consist of two equally weighted parts: a replication project and a take-home exam.

Course Objectives

The primary goal of this course is to guarantee that students can (a) understand, (b) implement, and (c) discuss possible improvements of any empirical paper in economics/finance (from the econometric perspective). We will develop the necessary theory and will illustrate it with empirical exercises. A secondary goal is to introduce modern statistical techniques (ML and high-dimensional statistics) and show how they can be used to improve standard estimation procedures.

Prerequisites

Satisfactory performance in the first-year statistical and econometric courses (including Applied Microeconometrics, if applicable).

Plan

1. Prediction: Empirical Risk Minimization and Regularization;
2. Basic Toolkit: Experiments, Weak Instruments, Unconfoundedness, RDD, Two-Way Models;
3. Beyond MHE: Selection Models, Marginal Treatment Effects;
4. General view: GMM and Related Problems;
5. Topics: Bipartite Networks, Quantile Methods, Shift-Share Designs;
6. Uncertainty Quantification: Design-Based vs. Model-Based Inference;

Literature

Textbooks

Lecture notes will be uploaded weakly on the course website. There is no textbook for the course, but the references below can be useful:

- Econometrics textbook by Bruce Hansen (old version available online);
- "Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction", by Guido W. Imbens and Donald B. Rubin;
- Manuel Arellano's lecture notes (available on his website);
- Stefan Wager's lecture notes for STATS 361 (available on his website)
- "Mostly Harmless Econometrics" by Joshua D. Angrist & Jorn-Steffen Pischke;
- "Elements of Statistical Learning" by Hastie, Tibshirani, Friedman (available online).
- "Computer Age Statistical Inference" by Hastie and Efron (available online);

Papers

Lecture notes will be partially based on the papers below (the list will be updated throughout the course). I do not expect you to read all or any of them (some are very technical), but I will let you know the most useful ones.

References

A. Abadie, A. Diamond, and J. Hainmueller. 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, 2010.

A. Abadie, S. Athey, G. W. Imbens, and J. M. Wooldridge. Finite population causal standard errors. Technical report, National Bureau of Economic Research, 2014.

A. Abadie, S. Athey, G. W. Imbens, and J. Wooldridge. When should you adjust standard errors for clustering? Technical report, National Bureau of Economic Research, 2017.

R. Adao, M. Kolesár, and E. Morales. Shift-share designs: Theory and inference. Technical report, National Bureau of Economic Research, 2018.

S. Agrawal and N. Goyal. Analysis of thompson sampling for the multi-armed bandit problem. In *Conference on learning theory*, pages 39–1. JMLR Workshop and Conference Proceedings, 2012.

I. Andrews, J. H. Stock, and L. Sun. Weak instruments in instrumental variables regression: Theory and practice. *Annual Review of Economics*, 11:727–753, 2019.

J. Angrist and M. Kolesár. One instrument to rule them all: The bias and coverage of just-id iv. Technical report, National Bureau of Economic Research, 2021.

M. Arellano and S. Bond. Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *The review of economic studies*, 58(2):277–297, 1991.

M. Arellano and S. Bonhomme. Nonlinear panel data estimation via quantile regressions, 2016.

M. Arellano and S. Bonhomme. Quantile selection models with an application to understanding changes in wage inequality. *Econometrica*, 85(1):1–28, 2017.

D. Arkhangelsky and G. Imbens. The role of the propensity score in fixed effect models. Technical report, National Bureau of Economic Research, 2018.

D. Arkhangelsky, S. Athey, D. A. Hirshberg, G. W. Imbens, and S. Wager. Synthetic difference in differences. Technical report, National Bureau of Economic Research, 2019.

T. B. Armstrong and M. Kolesár. Optimal inference in a class of regression models. *Econometrica*, 86(2):655–683, 2018.

O. C. Ashenfelter and D. Card. Using the longitudinal structure of earnings to estimate the effect of training programs, 1984.

S. Athey and G. W. Imbens. Identification and inference in nonlinear difference-in-differences models. *Econometrica*, 74(2):431–497, 2006.

S. Athey, G. W. Imbens, and S. Wager. Approximate residual balancing: debiased inference of average treatment effects in high dimensions. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 2016.

S. Athey, M. Bayati, N. Doudchenko, G. Imbens, and K. Khosravi. Matrix completion methods for causal panel data models. *arXiv preprint arXiv:1710.10251*, 2017.

J. Bai. Panel data models with interactive fixed effects. *Econometrica*, 77(4):1229–1279, 2009.

A. Belloni, V. Chernozhukov, and C. Hansen. Inference on treatment effects after selection among high-dimensional controls. *The Review of Economic Studies*, 81(2):608–650, 2014.

M. Bertrand, E. Duflo, and S. Mullainathan. How much should we trust differences-in-differences estimates? *The Quarterly journal of economics*, 119(1):249–275, 2004.

S. Bonhomme and E. Manresa. Grouped patterns of heterogeneity in panel data. *Econometrica*, 83 (3):1147–1184, 2015.

S. Bonhomme, T. Lamadon, and E. Manresa. Discretizing unobserved heterogeneity. *University of Chicago, Becker Friedman Institute for Economics Working Paper*, (2019-16), 2017.

S. Bonhomme, T. Lamadon, and E. Manresa. A distributional framework for matched employer employee data. *Econometrica*, 87(3):699–739, 2019.

S. Bonhomme, K. Holzheu, T. Lamadon, E. Manresa, M. Mogstad, and B. Setzler. How much should we trust estimates of firm effects and worker sorting? Technical report, National Bureau of Economic Research, 2020.

K. Borusyak, P. Hull, and X. Jaravel. Quasi-experimental shift-share research designs. Technical report, National Bureau of Economic Research, 2018.

K. Borusyak, X. Jaravel, and J. Spiess. Revisiting event study designs: Robust and efficient estimation. *arXiv preprint arXiv:2108.12419*, 2021.

G. Chamberlain. Asymptotic efficiency in estimation with conditional moment restrictions. *Journal of Econometrics*, 34(3):305–334, 1987.

G. Chamberlain. Efficiency bounds for semiparametric regression. *Econometrica: Journal of the Econometric Society*, pages 567–596, 1992.

V. Chernozhukov, D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, et al. Double machine learning for treatment and causal parameters. *arXiv preprint arXiv:1608.00060*, 2016.

V. Chernozhukov, W. Newey, and J. Robins. Double/de-biased machine learning using regularized riesz representers. *arXiv preprint arXiv:1802.08667*, 2018.

M. Cytrynbaum. Designing representative and balanced experiments by local randomization. *arXiv preprint arXiv:2111.08157*, 2021.

M. Das, W. K. Newey, and F. Vella. Nonparametric estimation of sample selection models. *The Review of Economic Studies*, 70(1):33–58, 2003.

C. De Chaisemartin and X. d'Haultfoeuille. Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–96, 2020.

D. Ghanem, P. H. Sant'Anna, and K. Wüthrich. Selection and parallel trends. *arXiv preprint arXiv:2203.09001*, 2022.

A. Goodman-Bacon. Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277, 2021.

B. Graham and A. de Paula. *The Econometric Analysis of Network Data*. Academic Press, 2019.

L. P. Hansen. Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, pages 1029–1054, 1982.

D. A. Hirshberg and S. Wager. Balancing out regression error: efficient treatment effect estimation without smooth propensities. *arXiv preprint arXiv:1712.00038*, 2017.

G. Imbens and S. Wager. Optimized regression discontinuity designs. *arXiv preprint arXiv:1705.01677*, 2017.

M. Kasy and A. Sautmann. Adaptive treatment assignment in experiments for policy choice. *Econometrica*, 89(1):113–132, 2021.

D. S. Lee, J. McCrary, M. J. Moreira, and J. R. Porter. Valid t-ratio inference for iv. Technical report, National Bureau of Economic Research, 2021.

C. F. Manski. Identification of endogenous social effects: The reflection problem. *The review of economic studies*, 60(3):531–542, 1993.

R. Meager. Aggregating distributional treatment effects: A bayesian hierarchical analysis of the microcredit literature. *American Economic Review*, 112(6):1818–47, 2022.

H. R. Moon and M. Weidner. Linear regression for panel with unknown number of factors as interactive fixed effects. *Econometrica*, 83(4):1543–1579, 2015.

D. B. Rubin. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, 66(5):688, 1974.

P. H. Sant'Anna and J. Zhao. Doubly robust difference-in-differences estimators. *Journal of Econometrics*, 219(1):101–122, 2020.

S. M. Schennach. Recent advances in the measurement error literature. *Annual Review of Economics*, 8:341–377, 2016.

L. Sun and S. Abraham. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199, 2021.

S. Wager, W. Du, J. Taylor, and R. J. Tibshirani. High-dimensional regression adjustments in randomized experiments. *Proceedings of the National Academy of Sciences*, 113(45):12673–12678, 2016.

J. M. Wooldridge. Two-way fixed effects, the two-way mundlak regression, and difference-in-differences estimators. *Available at SSRN 3906345*, 2021.