

International Economics Development Economics

Academic year 2019-2020

Impact Evaluation

EI041 - Autumn - 6 ECTS

Wednesday 14h15 - 16h00

Course Description

Statistically assessing the causal impact of development policies has now become an extremely large industry. While there will always be five available identification strategies, doing things right in a policy-relevant manner is neither obvious, nor easy. The literature on impact evaluation is a subset of econometrics, sometimes with a vocabulary of its own. As such, econometric methods that you have learnt will figure prominently in what follows. There is no textbook for this course. The course will be arranged around a selection of readings that touch on various technical aspects of impact evaluation, as well as applications. Roughly one third of class time will be devoted to showing how to do the stuff mentioned in the readings using R, and you will be provided with the code shortly after the class, so that you can experiment on your own. An important component of this course is embedded in the problem sets, which mostly revolve around doing geeky stuff in R.

There will be 3 problem sets, a midterm and a final.

> PROFESSOR

[Jean-Louis Arcand](#)

[Office hours](#)

> ASSISTANT

Carolina Lemos Rêgo

[Office hours](#)

Syllabus

Statistically assessing the causal impact of development policies has now become an extremely large industry. While there will always be five available identification strategies, doing things right in a policy-relevant manner is neither obvious, nor easy.¹ And the tendency to evaluate stuff simply because it is possible to do so rigorously –instead of trying to evaluate what deserves to be evaluated from the development perspective, has sometimes led to what I call the “tail wagging the dog” syndrome. There is also the minor issue of scientific progress and actually accumulating a useful body of knowledge.

The literature on impact evaluation is a subset of econometrics, sometimes with a vocabulary of its own. As such, econometric methods that you have learnt will figure prominently in what follows. There is no textbook for this course. However, there are several great surveys of impact evaluation methods, by masters in the field. Two of my favorites, which adopt diametrically opposite philosophical stances, are:²

- Imbens, G. W. and Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47(1):5–86.
- Heckman, J. J. and Vytlačil, E. J. (2007). Econometric evaluation of social programs, part 1: Causal models, structural models, and econometric policy evaluation. In Heckman, J. J. and Leamer, E. E., editors, *Handbook of Econometrics*, volume 6B, chapter 70, pages 4779–4874. North Holland, Amsterdam.

A very nice survey of more recent approaches, including synthetic cohorts and machine learning methods is given by:

¹In case you’re wondering, the five identification strategies are: (i) wishing the problem away by assuming selection on observables and applying either OLS or some type of matching estimator (I will have little if anything to say about this in what follows), (ii) randomization, (iii) instrumental variables, (iv) regression discontinuity design and (v) some sort of covariance transformation that you can implement because you are lucky enough to have panel data. Convex combinations of these are, of course, very common as well.

²The first views the world through a-theoretical Rubin lenses, the second embraces economic theory.

- Athey, S. and Imbens, G. W. (2017b). The state of applied econometrics: Causality and policy evaluation. *Journal of Economic Perspectives*, 31(2):3–32.

For my own (cynical) view on all that follows, see Arcand, J.-L. (2014). The (lack of) impact of impact: Why impact evaluations seldom lead to evidence-based policymaking. *Revue d'Economie du Développement*, 289-311. But mere cynicism is not going to stop me from enjoying teaching this class.

There is a plethora of websites dedicated to impact evaluation-related topics that you should consult on a regular basis. A non-random sample of these includes:

- <https://blogs.worldbank.org/impactevaluations/>
- <http://www.3ieimpact.org/>
- <https://www.povertyactionlab.org/>
- <http://www.poverty-action.org/>
- <http://www.worldbank.org/en/research/dime>

+ my favorite statistics blog: <http://andrewgelman.com/>

The course will be arranged around a selection of readings that touch on various technical aspects of impact evaluation, as well as applications. For you to be able to follow what is going on, I expect you to have done the readings before class.

Roughly one third of class time will be devoted to showing how to do the stuff mentioned in the readings using R, and I will be using an .RMD file for the course, that I will make available and update on a regular basis during the semester.

An important component of this course is embedded in the problem sets and the “exams” (which are basically glorified problem sets), which mostly revolve around doing geeky stuff in R. There will be 2 problem sets, plus a take-home midterm and a take-home final.

1 Why go to all the trouble? Potential outcomes

- Ravallion, M. (2009). Evaluation in the practice of development. *World Bank Research Observer*, 24(1):29–53.
- Heckman, J. J. and Vytlačil, E. J. (2005). Structural equations, treatment effects and econometric policy evaluation. *Econometrica*, 73(3):669–738.

2 Randomization

2.1 Understanding the basics

- Heckman, J. J. (1996). Randomization as an instrumental variable. *Review of Economics and Statistics*, 78(2):336–341.

2.2 The *randomista* debate

- Deaton, A. (2010). Instruments, randomization, and learning about development. *Journal of Economic Literature*, 48(2):424–455.
- Imbens, G. W. (2010). Better LATE than nothing: Some comments on Deaton (2009) and Heckman and Urzua (2009). *Journal of Economic Literature*, 48(2):399–423.
- Barrett, C. B. and Carter, M. R. (2010). The power and pitfalls of experiments in development economics: Some non-random reflections. *Applied Economic Perspectives and Policy*, 32(4):515–548.
- Ravallion, M. (2012). Fighting poverty one experiment at a time: A review of Abhijit Banerjee and Esther Duflo’s Poor economics: A radical rethinking of the way to fight global poverty. *Journal of Economic Literature*, 50(1):103–114.
- Deaton, A. and Cartwright, N. (2018). Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine*, 210:2–21.
- Ravallion, M. (2018). Should the randomistas (continue to) rule? CGD Working Paper No. 492.

2.3 Bias and randomization inference

- Eble, A., Boone, P., and Elbourne, D. (2017). On minimizing the risk of bias in randomized controlled trials in economics. *World Bank Economic Review*, 31(3):687–707.
- Young, A. (2018). Channeling fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results. *The Quarterly Journal of Economics*, 134(2):557–598.
- Athey, S. and Imbens, G. W. (2017a). The econometrics of randomized experiments. In *Handbook of Field Experiments*, volume 1, chapter 3, pages 73–140. Elsevier.

2.4 Placebo effects and (surprise) *homo oeconomicus* is alive and kicking

- Malani, A. (2006). Identifying placebo effects with data from clinical trials. *Journal of Political Economy*, 114(2):236–256.
- Bulte, E., Beekman, G., Di Falco, S., Hella, J., and Lei, P. (2014). Behavioral responses and the impact of new agricultural technologies: Evidence from a double-blind field experiment in Tanzania. *American Journal of Agricultural Economics*, 96(3):813–830.
- Chassang, S., Padró I Miquel, G., and Snowberg, E. (2012). Selective trials: A principal-agent approach to randomized controlled experiments. *American Economic Review*, 102(4):1279–1309.
- Wing, C. and Clark, M. H. (2016). What can we learn from a doubly randomized preference trial? –An instrumental variables perspective. *Journal of Policy Analysis and Management*, 36(2):418–437.

3 Setting up an evaluation

3.1 Statistical power, survey data and just doing it...

- Bloom, H. S. (1995). Minimum detectable effects: A simple way to report the statistical power of experimental designs. *Evaluation Review*, 19(5):547–556.
- Optimal Design software: <http://hlmssoft.net/od/>
- Deaton, A. (1997). *The Analysis of Household Surveys. A Microeconomic Approach to Development Policy*. Johns Hopkins University Press, Baltimore, MD.
- List, J. A. (2011). Why economists should conduct field experiments and 14 tips for pulling one off. *Journal of Economic Perspectives*, 25(3):3–16.
- Ranganathan, A. (2018). The artisan and his audience: Identification with work and price-setting in a handicraft cluster in southern India. *Administrative Science Quarterly*, 63(3):637–667.

3.2 Stuff to keep in mind: Survey bias and Hawthorne effects

- Zwane, A. P., Zinman, J., Van Dusen, E., Pariente, W., Null, C., Miguel, E., Kremer, M., Karlan, D. S., Hornbeck, R., Giné, X., Duflo, E., Devoto, F., Crepon, B., and Banerjee, A. (2011). Being surveyed can change later behavior and related parameter estimates. *Proceedings of the National Academy of Sciences of the United States of America*, 108(5):1821–1826.

- Levitt, S. D. and List, J. A. (2011). Was there really a Hawthorne effect at the Hawthorne plant? An analysis of the original illumination experiments. *American Economic Journal: Applied Economics*, 3(1):224–238.

4 IV and RDD

4.1 The basics of IV

- Hausman, J. A. (1983). Specification and estimation of simultaneous equation models. In Griliches, Z. and Intriligator, M., editors, *Handbook of Econometrics*, volume 1, pages 391–448, The Netherlands. North Holland.

4.2 Finite sample bias, common mistakes, the GMM black box, and bootstrap inference

- Hahn, J. and Hausman, J. A. (2002). Notes on bias in estimators for simultaneous equation models. *Economics Letters*, 75(2):237–241.
- Conley, T. G., Hansen, C., and Rossi, P. E. (2012). Plausibly exogenous. *Review of Economics and Statistics*, 94(1):260–272
- Bazzi, S. and Clemens, M. A. (2013). Blunt instruments: Avoiding common pitfalls in identifying the causes of economic growth. *American Economic Journal: Macroeconomics*, 5(2):152–186.
- Young, A. (2019). Consistency without inference: Instrumental variables in practical applications. London School of Economics.

4.3 Regression discontinuity design

- Imbens, G. W. and Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2):615–635.
- Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, 48(2):281–355.

5 Panel data structures

5.1 Reviewing the basics

- Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica*, 46(1):69–85.

- Hausman, J. A. and Taylor, W. (1981). Panel data and unobservable individual effects. *Econometrica*, 49(6):1377–1398.
- Lancaster, T. (2000). The incidental parameter problem since 1948. *Journal of Econometrics*, 95(2):391–413.
- Semykina, A. and Wooldridge, J. M. (2010). Estimating panel data models in the presence of endogeneity and selection. *Journal of Econometrics*, 157(2):375–380.

5.2 Inference and clustering

- Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics*, 119(1):249–275.
- Cameron, C. A., Gelbach, J. B., and Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered standard errors. *Review of Economics and Statistics*, 90(3):414–427.

5.3 Measurement error

- Hausman, J. A. (2001). Mismeasured variables in econometric analysis: Problems from the right and problems from the left. *Journal of Economic Perspectives*, 15(4):57–67.
- Griliches, Z. and Hausman, J. A. (1986). Errors in variables in panel data. *Journal of Econometrics*, 31(1):93–118.
- Hausman, J. A., Abrevaya, J., and Scott-Morton, F. M. (1998). Misclassification of the dependent variable in a discrete-response setting. *Journal of Econometrics*, 87(2):239–269.
- Lewbel, A. (2012). Using heteroskedasticity to identify and estimate mismeasured and endogenous regressor models. *Journal of Business and Economic Statistics*, 30(1):67–80.
- Dagenais, M. G. and Dagenais, D. L. (1997). Higher moment estimators for linear regression models with errors in the variables. *Journal of Econometrics*, 76(1-2):193–221.

5.4 Pseudo-panel data and synthetic control

- Verbeek, M. and Nijman, T. (1993). Minimum MSE estimation of a regression model with fixed effects from a series of cross-sections. *Journal of Econometrics*, 59(1-2):125–136.

- Abadie, A., Diamond, A., and Hainmueller, J. (2015). Comparative politics and the synthetic control. *American Journal of Political Science*, 59(2):495–510.

6 External validity, replicability and lack of statistical significance

6.1 External validity

- Rosenzweig, M. R. and Udry, C. (2019). External validity in a stochastic world: Evidence from low-income countries. *The Review of Economic Studies*, forthcoming.

6.2 Replicability and scientific progress

- Roodman, D. and Morduch, J. (2014). The impact of microcredit on the poor in Bangladesh: Revisiting the evidence. *Journal of Development Studies*, 50(4):583–604.
- The “Worm Wars”: <https://blogs.worldbank.org/impactevaluations/worm-wars-anthology>.
- Simmons, J. P., Nelson, L. D., and Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22(11):1359–1366.
- Yong, E. (2012). Replication studies: Bad copy. *Nature*, 485:298–300.
- Brodeur, A., Cook, N., and Heyes, A. (2018). Methods matter: P-hacking and causal inference in economics. IZA DP No. 11796
- Nosek, B. A., Spies, J. R., and Motyl, M. (2013). Scientific utopia: II. Restructuring incentives and practices to promote truth over publishability. *Perspectives on Psychological Science*, 7:615–631.
- Smaldino, P. E. and McElreath, R. (2016). The natural selection of bad science. *Royal Society Open Science*, 3:160384.

6.3 Saying something intelligent when things aren’t statistically significant

- Andrews, D. W. K. (1989). Power in econometric applications. *Econometrica*, 57(5):1059–1090.

7 Heterogeneity in various flavors

7.1 Quantile regressions and random coefficient models

- Koenker, R. and Hallock, K. (2001). Quantile regression. *Journal of Economic Perspectives*, 15(4):143–156.
- Beck, N. and Katz, J. N. (2007). Random coefficient models for time-series-cross-section data: Monte carlo experiments. *Political Analysis*, 15(2):182–195.

7.2 Essential heterogeneity

- Heckman, J. J. and Vytlačil, E. J. (1999). Local instrumental variables and latent variable models for identifying and bounding treatment effects. *Proceedings of the National Academy of Sciences of the United States of America*, 96(8):4730–4734.
- Heckman, J. J., Urzua, S., and Vytlačil, E. (2006). Understanding instrumental variables in models with essential heterogeneity. *Review of Economics and Statistics*, 88(3):389–432.
- Heckman, J. J. and Navarro-Lozano, S. (2004). Using matching, instrumental variables and control functions to estimate economic choice models. *Review of Economics and Statistics*, 86(1):30–57.
- Ravallion, M. (2015). On the implications of essential heterogeneity for estimating causal impacts using social experiments. *Journal of Econometric Methodology*, 4(1):145–151.