# POLI 666: Causal Inference with Observational Data

**Winter Semester 2022**

**Wednesday, 2.05-3.25pm and Friday, 2:05-3:25-pm**

**Leacock 517**

**Instructors:** Professor Leo Baccini
**Office:** Leacock 432
**Office hours:** Wed & Fri., 12-1pm

**TA:**
**Office:**
**Office hours:**

## Course description

This course covers empirical strategies for applied research questions. The focus of the course is on statistical methods used for causal inference in the social sciences. Using the potential outcomes framework of causality, we discuss designs and methods for data from observational studies. While the emphasis of the course is on quantitative methods, the first class will survey issues and techniques related to qualitative analysis and mixed methods. The core of the course includes panel data analysis, instrumental variables, matching, differences-in-differences, regression discontinuity designs, and advance panel techniques for both continuous and dichotomous outcomes. Examples are drawn from different social sciences. The goal of the course is to show that a carefully thought research design goes often as far as fancy statistical methods when it comes to assess causality.

## Prerequisite

Students should have taken POLI 618 (or equivalent). A basic knowledge of R or Stata is necessary.

## Software

We will use R in classes and in lab sessions. However, we will provide all the replication files also in Stata and the instructor would be happy to answer Stata-related questions. Moreover, to complete assignments and/or to implement the analysis of the final project, students are welcome to use either R or Stata.

## Course Material

The main course texts are:

1. Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics*. Princeton University Press. [MHE]

2. Angrist, Joshua D., and Jörn-Steffen Pischke. 2014. *Mastering'metrics: The path from cause to effect*. Princeton University Press [MM].

3. Bailey, Michael A. 2016. *Real Stats*. Oxford University Press.

All texts have similar content and address the same techniques. MHE is the most technical and the heaviest on math. Bailey is the most accessible and is purely applied. MM is a compromise between the two: more accessible than MHE, but less intuitive than Bailey. If you have a good background in econometrics, I would suggest relying on MHE, which would prepare you to take more advanced courses in the future. In my lectures I try to keep the math at a minimum and to do well in the course it is sufficient that you learn the applied component of each technique.

Other useful text books are:

- Cattaneo, M. D., & Escanciano, J. C. (Eds.). 2017. *Regression Discontinuity Designs: Theory and Applications* (Vol. 38). Emerald Group Publishing [RDD].
- Cleves, Mario. *An introduction to survival analysis using Stata*. Stata Press, 2008. [DURATION MODELS]
- Gelman, Andrew, and Jennifer Hill. 2006. *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press. [MULTILEVEL ANALYSIS]
- Rosenbaum, Paul. 2009. Design of Observational Studies. Heidelberg: Springer. [MATCHING]
- Wooldridge, Jeffrey M. 2010. *Econometric analysis of cross section and panel data*. MIT University Press. [PANEL DATA]

All the articles covered in classes and in assignments are freely available from the academic journals in which they are published through the McGill library. All the unpublished working papers will be made available to students before class.

**Course requirements**

**This course is delivered in-person. I will be lecturing in the assigned building/room during the scheduled time for the course.**

This course is delivered in-person. I will be lecturing in the assigned building/room during the scheduled time for the course. A tentative structure of the course entails lecturing on a specific methodology on Wednesday and discussing the material using the same methodology on Friday. Moreover, there will be 1.5-hour lab sessions, which will cover the programming component of each methodology using R, every Monday, starting in Week 3.

1. Participation 10%: participation is assessed on comments provided in classes and in lab sessions. Each week I will assign a specific reading to one or two students for discussion.

2. Homework assignments 40%: there are eight weekly assignments, which are mostly short replications of published articles or working papers. I will make available dataset for replications. Note that it is not enough to simply replicate the estimates in the assignments, it is also necessary to explain the rationale of each empirical analysis and to carefully interpret the results. While I favor working together on the assignments, each student must submit her/his own assignment, which cannot be exactly the same for two or more students. Assignments are generally due a week after they are assigned. Assignments are due on Tuesday of the following week in which the specific methodology has been covered in class. Assignments should be uploaded on Mycourses (no physical copy accepted).

3. Research project presentation 10%: students are supposed to present their final research paper in weeks 11 and 12. While the research project does not have to be completed, presentations should focus on the research design and on the data analysis. Presentations should be made in a professional academic format. Each presentation will be followed by the comments from a discussant and then by comments from the other seminar participants. Presentations should be uploaded on Mycourses.

- Research paper 40%: the research project should have the following components: 1) a clear theory proposing the causal effect of an explanatory variable(s) on an outcome variable; 2) data collection, i.e. replications of existing studies are not sufficient. While a research paper can build on the empirical analysis of an article already published, students are required to extend the dataset of existing studies, e.g. merging the dataset from two different articles; 3) data analysis using at least one method covered in the course; 4) clear and through discussion of both findings and limitations of the paper. Students may use a chapter of their master or Ph.D. thesis as a research paper. Students are supposed to submit title and detailed abstract of the research paper by the end of the reading week. The research paper is due on April 29 2022 and should be uploaded on Mycourses (no physical copy accepted).

**Grade Distribution**

- Participation: 10%
- Assignments: 40%
- Presentation: 10%
- Research paper: 40%

| Assignment | Weight | Due Date |
|---|---|---|
| Participation | 10% | Throughout the term |
| Assignments | 40% | Weakly |
| Presentation | 10% | Weeks 11 and 12 |
| Research paper | 40% | April 29, 2022 |

**Policies**

In order to calculate your final grade, I sum all of the points you have received in the class and convert the points into letter grades using the following scale:
85-100 = A
80-84 = A-
75-79 = B+
70-74 = B
65-69 = B-
60-64 = C+
55-59 = C
50-54 = D
Below 50 = F

I round up scores of 0.5 and higher and round down scores of less than 0.5. No exceptions are made to the above system of converting points into letter grades.

Students who wish to contest a grade for an assignment must do so in writing (by email, sent to the T.A.) providing the reasoning behind their challenge to the grade received, within two weeks of the day on which the assignments are returned. If they are not satisfied by the TA's answer, they can write me to re-assess the grade. In that case, I re-evaluate the paper, but also reserve the right to raise or lower the grade.

Students who need to miss a class due to a religious holiday should notify me at least fourteen days prior to the holiday. If you must miss an examination in order to observe a religious holy day, you are given an opportunity to complete the missed work within a reasonable time after the absence.

Late work will be penalized by 5 points per day.

**The Fine Print**

McGill University values academic integrity. Therefore, all students must understand the meaning and consequences of cheating, plagiarism and other academic offences under the Code of Student Conduct and Disciplinary Procedures (see www.mcgill.ca/students/srr/honest/ for more info).

In accord with McGill University's Charter of Students' Rights, students in this course have the right to submit in English or in French any written work that is to be graded.

Instructors who adopt the use of text-matching software to verify the originality of students' written course work must register for use of the software with Educational Technologies and must inform their students before the drop/add deadline, in writing, of the use of text-matching software in a course.

**Course Evaluations**

End-of-course evaluations are one of the ways that McGill works towards maintaining and improving the quality of courses and the student's learning experience. You will be notified by e-mail when the evaluations are available on Mercury, the online course evaluation system. Please note that a minimum number of responses must be received for results to be available to students.

**Final note**

The content of the course could face small revisions in light of the composition of the seminar participants. Similarly, the content of weekly assignments could be modified due to data availability.

*McGill University is on the traditional territory of the Anishinaabeg and Haudenosaunee Nations and a place which has long served as a site of meeting and exchange amongst various Indigenous nations.*

## Tentative Schedule

WEEK 1. INTRODUCTION

January 5: CLASS CANCELED

January 7: Course description and logistics

WEEK 2: INTRODUCTION AND MIXED METHODS

January 12: Theory:

MHE, Ch. 1 and 2

Lieberman, Evan S. (2005) "Nested analysis as a mixed-method strategy for comparative research." *American Political Science Review* 99(3): 435-452.

Rohlfing, Ingo (2007) "What You See and What You Get: Pitfalls and Principles of Nested Analysis in Comparative Research." *Comparative Political Studies* 20(10).

January 14: Applications (data description):

Bush, Sarah Sunn. (2011) "International politics and the spread of quotas for women in legislatures." *International Organization* 65(1): 103-137.

Howard, Marc Morjé, and Philip G. Roessler. (2006) "Liberalizing electoral outcomes in competitive authoritarian regimes." *American Journal of Political Science* 50(2): 365-381.

Murali, Kanta. (2011) "Economic Liberalization, Electoral Coalitions and Private Investment in India." Mimeo.

*Assignment 1: selecting cases on-the-line and off-the-line.*

WEEK 3: FIXED EFFECTS AND RANDOM EFFECTS

January19: Theory:

Bailey, Ch. 8 (pages 247-68)

MHE, Ch. 3.1 and 3.2 (pages 27-68) and 5 (pages 222-227 & pages 243-247)

Gelman & Hill, Ch. 11 and 12

January 21: Applications (gravity model & hierarchical trade data):

Baier, S. L., and Bergstrand, J. H. (2007). "Do free trade agreements actually increase members' international trade?" *Journal of international Economics*, 71(1), 72-95.

Gowa, Joanne, and Raymond Hicks. (2013). "Politics, Institutions, and Trade: Lessons of the Interwar Era." *International Organization* 67(3): 439-467.

Kim, In Song and Weihuang Wong. 2015. "Intra-industry Trade and Trade Liberalization: Evidence from Preferential Tariffs Data". Working paper.

*Assignment 2: gravity model using Desta data.*

## WEEK 4: DIFFERENCES-IN-DIFFERENCES

January 26: Theory:

Bailey, Ch. 8 (pages 268-286)

MHE, Ch. 5 (pages 227-243)

Card, D., & Krueger, A. B. (2000). "Minimum wages and employment: a case study of the fast-food industry in New Jersey and Pennsylvania." *American Economic Review*, 1397-1420.

Autor, D. H. (2003). "Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing." *Journal of labor economics*, *21*(1), 1-42.

January 28: Applications (IR/IPE and voting behavior):

Bechtel, M. M., and Hainmueller, J. (2011) "How Lasting Is Voter Gratitude? An Analysis of the Short-and Long-Term Electoral Returns to Beneficial Policy." *American Journal of Political Science*, *55*(4), 852-868.

Bechtel, M. M., Hangartner, D., and Schmid, L. (2015). "Does Compulsory Voting Increase Support for Leftist Policy?" *American Journal of Political Science*.

"Berger, D., Easterly, W., Nunn, N., and Satyanath, S. (2013). "Commercial imperialism? Political influence and trade during the Cold War." *American Economic Review*, 103(2).

Earle, J. S., & Gehlbach, S. (2015). "The Productivity Consequences of Political Turnover: Firm-Level Evidence from Ukraine's Orange Revolution." *American Journal of Political Science*.

*Assignment 3: replication of Bechtel and Hainmueller, J. (2011). Replication file available at https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/25582*

## WEEK 5: INSTRUMENTAL VARIABLES I

February 2: Theory:

Bailey, Ch. 9 (pages 287-300)

MHE, Ch. 4.1-4.4 (pages 113-172)

February 4: Applications:

Acemoglu, D., Johnson, S., & Robinson, J. A. (2000). "The colonial origins of comparative development: An empirical investigation." *American Economic Review*.

Baccini, Leonardo, Maria Laura Sudulich, and Matthew Wall. (2015). "Internet effects in times of political crisis: online newsgathering and attitudes towards the European Union." Forthcoming in *Public Opinion Quarterly*.

Miguel, E., Satyanath, S., & Sergenti, E. (2004). "Economic shocks and civil conflict: An instrumental variables approach." *Journal of political Economy*, 112(4), 725-753.

*Assignment 4: replication of AJR paper (due in week 5).*

## WEEK 6: INSTRUMENTAL VARIABLES II

February 9: Theory:

Bailey, Ch. 9, pages 300-324

MHE Ch. 4.5-4.6 (pages 173-218)

February 11: Applications:

Cederman, Lars-Erik, et al. "Territorial Autonomy in the Shadow of Conflict: Too Little, Too Late?." *American Political Science Review* 109.02 (2015): 354-370.

Dietrich, S., & Wright, J. (2015). "Foreign Aid Allocation Tactics and Democratic Change in Africa." *The Journal of Politics*, 77(1), 216-234.

Kern, H. L., & Hainmueller, J. (2009). "Opium for the masses: How foreign media can stabilize authoritarian regimes." *Political Analysis*.

Pinto and Zhu. (2015) "Fortune or Evil? The Effect of Inward Foreign Direct Investment on Corruption." Forthcoming in *International Studies Quarterly*.

## WEEK 7: REGRESSION DISCONTINUITY

February 16: Theory:

Bailey, Ch. 11

MHE Ch. 6

February 18: Applications:

Bertoli, Andrew. (2015) "Nationalism and Interstate Conflict: A Regression Discontinuity Analysis." Manuscript, University of California, Berkeley.

Dal Bó, E., Dal Bó, P., & Snyder, J. (2009). "Political dynasties." *The Review of Economic Studies*, *76*(1), 115-142.

Eggers, A. C., and Hainmueller, J. (2009). "MPs for sale? Returns to office in postwar British politics." *American Political Science Review*, *103*(04), 513-533.

Keele, Luke J., and Rocio Titiunik. "Geographic boundaries as regression discontinuities." *Political Analysis* (2014): mpu014.

Szakonyi, D. (2018). Businesspeople in Elected Office: Identifying Private Benefits from Firm-Level Returns. *American Political Science Review*, *112*(2), 322-338.

Software: https://sites.google.com/a/umich.edu/titiunik/software.

*Assignment 5: replication of* Szakonyi *'s paper using regression discontinuity. Replication files (R and Stata) available at: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/DO43V5*

## WEEK 8: MATCHING TECHNIQUES AND SYNTHETIC CONTROL METHODS

February 23: Theory:

MHE Ch. 3.3 (pages 68-91)

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

Hainmueller, J. (2011). "Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies." *Political Analysis*.

Iacus, S. M., King, G., and Porro, G. (2011). "Causal inference without balance checking: Coarsened exact matching." *Political analysis*.

February 25: Applications:

Abadie, A. and J. Gardeazabal (2003), "The Economic Costs of Conflict: A Case Study of the Basque Country," *American Economic Review*, 93(1): 113-132.

Billmeier, A., & Nannicini, T. (2013). "Assessing economic liberalization episodes: A synthetic control approach." *Review of Economics and Statistics*, 95(3), 983-1001.

Hill, D. W. (2010). "Estimating the effects of human rights treaties on state behavior." *The Journal of Politics*, *72*(04), 1161-1174.

Lupu, Y. (2015). "Legislative Veto Players and the Effects of International Human Rights Agreements." *American Journal of Political Science*.

Spilker, G. and Böhmelt, T. (2013). "The impact of preferential trade agreements on governmental repression revisited." *The Review of International Organizations*, 8(3): 343-361.

*Assignment 6: matching/balancing and gravity model using Desta data.*

## WEEK 9: ADVANCED PANEL DATA

March 9: Theory:

Bailey, Ch. 13 & 15 [the basics]

De Boef, Suzanna, and Luke Keele. "Taking time seriously." *American Journal of Political Science* 52(1) (2008): 184-200.

Grant, T., & Lebo, M. J. (2016). Error correction methods with political time series. *Political Analysis*, *24*(1), 3-30.

March 11: Applications:

Büthe, Tim, and Helen V. Milner. "Foreign direct investment and institutional diversity in trade agreements: Credibility, commitment, and economic flows in the developing world, 1971–2007." *World Politics* 66(1) (2014): 88-122.

Doyle, David. "Remittances and social spending." *American Political Science Review* 109(4) (2015): 785-802.

Gawande, Kishore, Devesh Kapur, and Shanker Satyanath. "Renewable Natural Resource Shocks and Conflict Intensity: Findings from India's Ongoing Maoist Insurgency." *Journal of Conflict Resolution* 61(1) (2017): 140-172.

*Assignment 7: replication of Doyle (2015).*

## WEEK 10: SURVIVAL ANALYSIS

March 16: Theory:

Box-Steffensmeier, Janet M., and Bradford S. Jones. 1997. "Time is of the Essence: Event History Models in Political Science." *American Journal of Political Science*, v. 41, no.4 (October): 336-83.

Box-Steffensmeier, Janet M., and Christopher J.W. Zorn. 2001. "Duration Models and Proportional Hazards in Political Science."*American Journal of Political Science.* 45(October): 951-67.

Box-Steffensmeier, Janet M., Suzanna De Boef, Kyle A. Joyce. 2007. "Event Dependence and Heterogeneity in Duration Models: The Conditional Frailty Model." *Political Analysis*.

March 18: Applications:

Elkins, Zachary, Andrew T. Guzman, and Beth A. Simmons. "Competing for capital: The diffusion of bilateral investment treaties, 1960–2000."*International organization* 60.04 (2006): 811-846.

Epstein, David L., et al. "Democratic transitions." *American journal of political science* 50.3 (2006): 551-569.

Gates, Scott, et al. "Institutional inconsistency and political instability: Polity duration, 1800–2000." *American Journal of Political Science* 50.4 (2006): 893-908.

Golub, Jonathan. "Survival analysis and European Union decision-making." *European Union Politics* 8.2 (2007): 155-179.

*Assignment 8: replication of Elkins et al (2006).*

WEEK 11: PRESENTATIONS OF THE FINAL PROJECT

March 23: Presentations

March 25: Presentations

WEEK 12: PRESENTATIONS OF THE FINAL PROJECT

March 30: Presentations

April 1: Presentations

WEEK 13: NONSTANDARD (ERROR) ISSUES & WRAP-UP

- April 6: *Quantile regressions*: MHE Ch. 7; *Nonstandard error issues*: MHE Ch. 8; *Mediation*: Imai, K., Keele, L., Tingley, D., & Yamamoto, T. (2011). "Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies." *American Political Science Review*, 105(04): 765-789.

- April 8: *Wrap-up article*: Keele, L. (2015). The statistics of causal inference: A view from political methodology. *Political Analysis*, *23*(3), 313-335.