POLI 666: Causal Inference with Observational Data

Winter Semester 2019
Tuesday, 4.05-5.25pm and Thursday, 4:05-5:25-pm

Instructors: Leo Baccini & Costin Ciobanu
Office: Leacock 432 (Leo) & McTavish 33-2 (Costin)
Office hours: Tue., 2-3pm (Leo) & Monday 9:30-10:30am (Costin)

TA: Ms Aleksandra Conevska
Office: LEA 424
Office hours: Friday, 1:30-3pm

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 Stata, R or similar statistical software 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 both instructors and TA 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 two main course texts are:

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 textbooks are:

- Cleves, Mario. *An introduction to survival analysis using Stata*. Stata Press, 2008. [DURATION MODELS]

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**

1. **Participation 10%**: participation is assessed on comments provided in classes and in lab sessions.

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.

- **Research project presentation 10%**: students are supposed to present their final research paper in weeks 10, 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.

- **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 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/](http://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.

Late work will be penalized by one letter grade (e.g. from B to B-) per day. If you drop-off late work outside of class, you must have it time-stamped and signed by a secretary in the political science department.

**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 email 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.
Tentative Schedule

1. **INTRODUCTION AND MIXED METHODS**

   **Theory:**

   MHE, Ch. 1 and 2


   **Applications (data description):**


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

2. **FIXED EFFECTS AND RANDOM EFFECTS**

   **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

   **Applications (gravity model & hierarchical trade data):**


*Assignment 2: gravity model using Desta data.*

3. **Differences-in-differences**

**Theory:**

Bailey, Ch. 8 (pages 268-286)

MHE, Ch. 5 (pages 227-243)


**Applications (IR/IPE and voting behavior):**


4. **Instrumental Variables I**

**Theory:**

Bailey, Ch. 9 (pages 287-300)

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

**Applications:**


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

5. **Instrumental Variables II**

**Theory:**

Bailey, Ch. 9, pages 300-324

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

**Applications:**


6. **Regression Discontinuity**

**Theory:**

Bailey, Ch. 11

MHE Ch. 6

**Applications:**


Software: [https://sites.google.com/a/umich.edu/titiunik/software](https://sites.google.com/a/umich.edu/titiunik/software).


7. **Matching Techniques and Synthetic Control Methods**

**Theory:**

MHE Ch. 3.3 (pages 68-91)


Applications:


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

8. **Advanced Panel Data**

Theory:

Bailey, Ch. 13 & 15 [the basics]


Applications:


9. **Survival Analysis**

**Theory:**


**Applications:**


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

10. **Presentations of the Final Project**

11. **Presentations of the Final Project**

12. **Presentations of the Final Project**

13. **Nonstandard (Error) Issues & Wrap-up**

   - *Quantile regressions*: MHE Ch. 7
   - *Nonstandard error issues*: MHE Ch. 8