POLI 666: Causal Inference

Professor Elissa Berwick

Winter 2021

Instructor: **Elissa Berwick**
elissa.berwick@mcgill.ca

Fixed lectures: WF 8:35-9:55 am
Office hours: by appointment

TA: Olivier Bergeron-Boutin
olivier.bergeron-boutin@mail.mcgill.ca

Lab Hours:
TBD

---

1 Overview and Goals

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, we discuss designs and methods for predominantly for observational data but also draw examples from experimental data. 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.

Course Structure

The course will meet for short, fixed lectures of approximately 50 minutes twice a week, on Wednesday and Fridays. All students who can attend are encouraged to do so, but these sessions will also be recorded for students who are unable to attend. Optional review sessions and open office hours with the TA will be held on Mondays.

This course is divided into five units: introduction to causal inference, randomized experiments, selection on observables, cross-sectional designs, and longitudinal designs. Each unit except the first will include one to two class sessions entirely devoted to discussing papers that apply the methods taught in that unit. Students must pick one of the assigned papers to read, and come to class prepared to discuss in detail how it uses the methods we have studied. **Attending the discussion sessions is required**. If due to unforeseen

circumstances a student cannot attend a discussion section, they must submit a one page discussion of one of the assigned papers via GitHub.

Who this course is for?

- You have taken POLI618 or equivalent and have a solid knowledge of R or are willing to learn R quickly.
- You are willing to spend time considerable outside of the classroom to learn the course materials, as data analysis is a skill learned by doing.

2 Course Materials

Required Textbooks

Given people's various backgrounds, we will have four main textbooks in addition to the assigned articles. The Bailey book is a great book and very applied. Everyone should read it. For those of you seriously interested in pursuing quantitative analysis, you should then read the Angrist and Pischke *MHE* book on top of the Bailey book. Angrist and Pischke's *MM* provides a simpler version of *MHE*. The Morgan and Winship book is a more technical, DAG-centric approach to causal inference. It often helps to read this material from several different angles. The books are all available for purchase through the McGill bookstore. Electronic availability indicated separately for each text.

Bailey, Michael A. (2015). *Real Stats: Using Econometrics for Political Science and Public Policy*. 1st. New York: Oxford University Press. *Also available for purchase as an e-book via https://redshelf.com/book/356389/real-stats-356389-9780190262198-michael-a-bailey*

Angrist, Joshua D. and Jörn-Steffen Pischke (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press. *Also available for free as an e-book via McGill libraries*

Morgan, Stephen L. and Christopher Winship (2015). *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Analytical methods for social research. New York, NY: Cambridge University Press. *Also available for free as an e-book via McGill libraries*

Angrist, Joshua D. and Jörn-Steffen Pischke (2014). *Mastering 'Metrics: The Path from Cause to Effect*. Google-Books-ID: s2eYDwAAQBAJ. Princeton University Press. *Also available for purchase as an e-book via various online retailers including Princeton University Press and Amazon*

Required Software

In this course, we will be using the statistical computing environment R, a FREE open source language used by data scientists and statisticians across the world. R consists of

a base environment for data manipulation, calculation and graphical display as well as numerous user-made packages that bundle together more specialized functions.

We will also be using a FREE integrated development environment (IDE) for R called RStudio that makes learning and exploring R easier. While the learning curve in R is steeper than in more expensive programs (such as Stata and SPSS), there is much more you can do with it!

There are many free online tutorials for downloading and installing R and RStudio. The RStudio team also makes great "cheatsheets" for using their interface (see here) as well as other R packages.

3 Course Requirements

There are 1000 points available in the class. Therefore, for each 10% of the grade, 100 points are allotted.

- **50%**. *Problem sets.* Either completed alone or as part of a rotating, randomly assigned group. **Seven total assignments**.
- **40%** *Final project.* The final project consists of a proposal, progress report, and final paper. The project will be based on quantitative data analysis — likely a replication exercise or part of your dissertation or MA Thesis. One portion of this grade (5% of total grade for the course) will come from the on-time submission of a **project proposal by March 10th**. Another portion of this grade (10% of total grade for the course) will come from the submission of a **progress report by April 12th** and participation in subsequent class discussions of the different projects on April 14th and April 16th. The remaining 35% of the grade will be based on a short final paper. **The final paper is due on April 21st**.
- **10%** *Participation in discussion sessions.* Based on participation in the paper discussions that will occur one to two times per unit. May be replaced by submission of a one page response paper if necessary.

Problem sets

Problem sets must be submitted via *GitHub*. To complete your homework, you must use an R variant, the GitHub compatible `.Rmd` (otherwise known as `rmarkdown`) files. Do not submit your homework using Microsoft Word or any other document editor. It will not be graded.

There will be approximately 7 problem sets. I reserve the right to lower or raise the number of problem sets if needed. In principle, one problem set will be assigned for each topic covered.

As in 618, there will be two tracks for completing problem sets. In the first track, I will divide students into randomly assigned small groups to work together. Group members must still submit problem sets individually, with their own write-ups and explanations of results. However, all group members will be expected to work together on coding, and for individuals in this track, 25% of the problem set grade will come from an average of the grades received by the rest of the group. (However, if a member of the group turns in a problem set late this will not count against other group members). Group assignments will be switched around between each problem set so that students get the opportunity to work with different people. The number of people assigned to each group will depend on how many students are on the track.

On the second track, students may opt out of group work, and submit assignments entirely on their own, with 100% of the grade coming from the problem set. Students are welcome to ask questions of classmates, but the assumption is that they are primarily working on their own. This is an opt-out rather than an opt-in system, since the process of explaining concepts to classmates is often the best way to internalize them yourself, and seeing other ways of approaching a problem can be very rewarding. Students who start on one track are always welcome to switch to the other track for subsequent problem sets, just not mid-problem set.

Final paper

All students will submit a final paper of SHORT article length (10-20 pages). This will be done via *Github*. This can be *either* a replication paper with an extension of the original paper *or* a new paper. An original paper must contain 1) a clear theory proposing a relationship between explanatory variable(s) on an outcome variable; 2) use of methods covered in this course 3) a clear 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. Replication papers must use one of the methods covered in the course, and cannot use papers that are discussed in class or analyzed in a problem set.

The final version of the paper is due by Wednesday, **April 21st**. Papers should focus on the research methods and do not have to have extensive citations or literature reviews. If you are a TA, you may be granted an extension based on department guidelines. If you need an extension to finish, please talk to me.

Project proposal

All students are required to submit a one page project proposal with a write up of the data set they are going to use and the research question they are going to ask by Wednesday, **March 10th**. The proposals should be a one page write up in `rmarkdown` explaining the data set/s which you are going to use and the question you will ask. You should also highlight your outcome variable. Submission will be via *Github*. For full credit, it is only

necessary to submit a proposal on time.

Progress report

Students will submit a progress report by **April 12th**. The progress report will contain preliminary results and discuss any concerns about how to proceed with the project. Reports will be posted to an open repository on *GitHub*, where other students are expected to read them and come to class on April 14th and April 16th prepared to discuss each others' work and how to improve it before the final paper is due. If students are unable to attend these discussions, their classmates will be asked to submit comments electronically via *GitHub*.

4 Class Policies

Extraordinary Circumstances Statement

In the event of extraordinary circumstances beyond the University's control, the content and/or evaluation scheme in this course is subject to change.

Re-Grading

Students who wish to contest a grade for an assignment or exam must do so in writing (by email, sent to me) providing the reasoning behind their challenge to the grade received within two weeks of the day on which the assignments are returned. I will re-evaluate the paper, but also reserve the right to **raise or lower the grade**. Please also see (http://www.mcgill.ca/politicalscience/files/politicalscience/assessment_and_re-read_policy_final.pdf).

Make-Up Work

If you are unable to complete a homework assignment for documented emergency medical or family reasons, an alternative submission date will be arranged. The alternative arrangement is only open to those who can provide a valid medical/family reason. If you cannot provide a valid reason for failing to submit an assignment on time, you will receive zero points for the submission.

Copyright of Lectures

All slides, video recordings, lecture notes, etc. remain the instructor's intellectual property. As such, you may use these only for your own learning (and research, with proper referencing/citation) ends. You are not permitted to disseminate or share these materials; doing so may violate the instructor's intellectual property rights and could be cause for disciplinary action.

I remind everyone of their responsibility in ensuring that this video and associated material are not reproduced or placed in the public domain. This means that each of you can use

it for your educational (and research) purposes, but you cannot allow others to use it, by putting it up on the Internet or by giving it or selling it to others who may also copy it and make it available. Please refer to McGill's Guidelines for Instructors and Students on Remote Teaching and Learning for further information.

Recording

By enrolling in a remote course, you accept that fixed sessions will be recorded. You must consent to being recorded if you are attending a lecture or participating in a component of a course that is being recorded. You will be notified through a "pop-up" box in Zoom if a lecture or portion of a class is being recorded. If you are not comfortable being in a class that is recorded, you may decide to not take part by logging off Zoom. Students who log off will be able to later watch the video recording in MyCourses.

For pedagogical reasons and for the enrichment of the experience of all students, attendance may be monitored and/or active participation may be expected or required during fixed (synchronous) class time. As such, you may be asked to turn on your camera and audio. If you do not have the necessary resources (e.g., adequate Internet bandwidth or equipment) to do so, inform your instructor at the beginning of term so that appropriate accommodations can be made.

In addition to the recording of your image and voice, your name (or preferred name) may be displayed on screen, and your instructor may call your name during the lecture. As such, this personal information will be disclosed to classmates, whether during the lecture or in viewing the recording. By remaining in classes that are being recorded, you accept that personal information of this kind may be disclosed to others, whether during the lecture or in viewing the recording.

Recordings will be deleted at the end of the semester to protect students' privacy.

Netiquette

The University recognizes the importance of maintaining teaching spaces that are respectful and inclusive for all involved. To this end, offensive, violent, or harmful language arising in contexts such as the following may be cause for disciplinary action:

1. Username (use only your legal or preferred name)
2. Visual backgrounds
3. Chat boxes

To maintain a clear and uninterrupted learning space for all, you should keep your microphone muted throughout your class, unless invited by the instructor to speak. You should follow instructors' directions about the use of the chat function on remote learning

platforms, and should NOT use the chat function for private conversations during class time.

Academic Integrity

Course Policy on Computer Code

As discussed in the problems set section, verbatim copying other people's computer code constitutes plagiarism. Moreover, data programming is learned through trial and error. *Please do not under any circumstances copy another students code.* You may of course collaborate with colleagues, but please write your own code! If you are found to have plagiarized, you may be referred to the appropriate Dean. The instructors reserve the right to use software to compare the code that has been written by different students.

McGill Policy

"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 information).

Language of Submission

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.

Conformément à la Charte des droits de l'étudiant de l'Université McGill, chaque étudiant a le droit de soumettre en français ou en anglais tout travail écrit devant être noté (sauf dans le cas des cours dont l'un des objets est la maîtrise d'une langue).

Disabilities Policy

As the instructor of this course I endeavor to provide an inclusive learning environment. However, if you experience barriers to learning in this course, do not hesitate to discuss them with me and the Office for Students with Disabilities, 514-398-6009.

End of 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. Please note that a minimum number of responses must be received for results to be available to students.

5 Class Schedule

Unit I: Introduction to causal inference

1. *(January 8th)* Introduction
2. *(January 13th)* Potential outcomes framework
    - READING:
        - MHE Ch. 1-2.1
        - MW 2-3
        - Sekhon, Jasjeet S. (2004). "Quality Meets Quantity: Case Studies, Conditional Probability, and Counterfactuals". In: Perspectives on Politics 2.2, pp. 281–293. DOI: 10.1017/S1537592704040150.

Unit II: The Randomized Ideal

1. *(January 15th & 20th)* Identification and estimation
    - READING:
        - Bailey Ch. 1 (Review)
        - MHE Ch. 2
        - Duflo, Esther, Rachel Glennerster, and Michael Kremer. 2006. "Using Randomization in Development Economics: A Toolkit." Handbook of Development Economics.
2. *(January 22nd & 27th)* Inference
    - READING:
        - MHE Ch 8.1
        - Fisher, Ronald Aylmer. 1966 [1935]. The Design of Experiments. Edinburgh; London: Oliver and Boyd. Part II.
        - Efron, Bradley, and R. J. Tibshirani. 1993. An Introduction to the Bootstrap. New York: Chapman and Hall/CRC. Chapters 2 and 6
3. *(January 29th)* Applying randomized experiments: **DISCUSSION**
    - PICK ONE:
        - Beamen, Lori et al. (2009). "Powerful Women: Does Exposure Reduce Bias?" In: *Quarterly Journal of Economics*, p. 1497 - 1540.
        - Björkman, Martina, and Jakob Svensson (2009). "Power to the people: evidence from a randomized field experiment on community-based monitoring in Uganda." In *The Quarterly Journal of Economics* 124, no. 2: 735-769.

  - Gerber, Alan S., Donald P. Green and Christopher W. Larimer (2008). "Social Pressure and Voter Turnout: Evidence from a Large Scale Field Experiment." In *American Political Science Review* 102(1): 1-48

Unit III: Selection on observables

1. *(February 3rd & 5th)* Matching and weighting
   - READING:
     - MHE 4-5
     - Iacus, Stefano M, Gary King, and Giuseppe Porro (2012). "Causal Inference without Balance Checking: Coarsened Exact Matching". In: *Political Analysis* 20.1, pp. 1–24.
2. *(February 10th)* Revisiting regression
   - READING:
     - MHE Ch. 3
     - MW 6-7
3. *(February 12th)* Applying matching: **DISCUSSION**
   - PICK ONE:
     - Blattman, Christopher and Jeannie Annan (2010). "The Consequences of Child Soldiering". In: *Review of Economics and Statistics* 92.4, pp. 882–898.
     - Gilligan, Michael J., and Ernest J. Sergenti (2008). "Do UN interventions cause peace? Using matching to improve causal inference." In: *Quarterly Journal of Political Science* 3, no. 2: 89-122
     - Lyall, Jason. "Does indiscriminate violence incite insurgent attacks? Evidence from Chechnya." Journal of Conflict Resolution 53, no. 3 (2009): 331-362.

Unit IV: Cross-section Designs

1. *(February 17th, 19th, & 24th)* Instrumental variables
   - READING:
     - MHE: Chapter 4
     - MW: Chapter 9
     - Bailey, Ch. 9.1-9.5,10.2,10.3
2. *(February 26th)* Applying instrumental variables: **DISCUSSION**

- PICK ONE:
  - Acemoglu, Daron, Simon Johnson, and James A. Robinson (2001). "The Colonial Origins of Comparative Development: An Empirical Investigation". In: *The American Economic Review* 91.5, pp. 1369–1401
  - Miguel, Edward, Shanker Satyanath, and Ernest Sergenti (2004). "Economic Shocks and Civil Conflict: An Instrumental Variables Approach". In: *Journal of Political Economy* 112.4, pp. 725–753.
  - Clingingsmith, David, Asim Ijaz Khwaja, and Michael Kremer (2009). "Estimating the Impact of the Hajj: Religion and Tolerance in Islam's Global Gathering". In: *Quarterly Journal of Economics*

3. *(March 10th & 12th)* Regression discontinuity designs

   - **Final project proposals (data check-in) due March 10th**
   - READING:
     - MHE: Ch. 6
     - Bailey, Ch. 1
     - McCrary, Justin (2008). "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test". In: *Journal of Econometrics* 142.2, pp. 698–714
     - Cattaneo, Matias D, Nicolás Idrobo, and Rocío Titiunik (2020). *A Practical Introduction to Regression Discontinuity Designs: Foundations*. Place of publication not identified: Cambridge University Press. https://cattaneo.princeton.edu/books/Cattaneo-Idrobo-Titiunik_2019_CUP-Vol1.pdf

4. *(March 17th)* Applying regression discontinuity designs: **DISCUSSION**

   - PICK ONE:
     - Cepaluni, Gabriel and F. Daniel Hidalgo (2016). "Compulsory Voting Can Increase Political Inequality: Evidence from Brazil". In: *Political Analysis* 24.2. Publisher: Cambridge University Press, pp. 273–280
     - Hinnerich, Björn Tyrefors, and Per Pettersson-Lidbom (2014). "Democracy, redistribution, and political participation: Evidence from Sweden 1919–1938." *Econometrica* 82, no. 3: 961-993.
     - Dunning, Thad, and Janhavi Nilekani (2013). "Ethnic quotas and political mobilization: caste, parties, and distribution in Indian village councils." *American Political Science Review* 107, no. 1: 35-56.

Unit V: Longitudinal Designs

1. *(March 19th & 24th)* Differences-in-differences

   - READING:

  - Bailey 8.5
  - MHE Ch.h (pp. 227-243)
  - Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan (2004). "How Much Should We Trust Differences-In-Differences Estimates?" In: *The Quarterly Journal of Economics* 119.1, pp. 249–275.

2. *(March 26th)* Applying differences-in-differences: **DISCUSSION**

   - PICK ONE:
     - Bechtel, Michael M. and Jens Hainmueller (2011). "How Lasting Is Voter Gratitude? An Analysis of the Short- and Long-Term Electoral Returns to Beneficial Policy". In: *American Journal of Political Science* 55.4, pp. 852–868.
     - Sances, Michael W. (2016). "The Distributional Impact of Greater Responsiveness: Evidence from New York Towns". In: *The Journal of Politics* 78.1
     - Earle, John S and Scott Gehlbach (2015). "The Productivity Consequences of Political Turnover: Firm-Level Evidence from Ukraine's Orange Revolution". In: *American Journal of Political Science* 59.3, p. 17.

3. *(March 31st & April 7th)* Fixed and random effects

   - READING:
     - MHE 5 (pages 222-227 and pages 243-247)

4. *(April 9th)* Applying fixed effects: **DISCUSSION**

   - PICK ONE:
     - Hainmueller, Jens and Dominik Hangartner (2019). "Does Direct Democracy Hurt Immigrant Minorities? Evidence from Naturalization Decisions in Switzerland". In: *American Journal of Political Science* 63.3, pp. 530–547
     - Scheve, Kenneth and David Stasavage (2012). "Democracy, War, and Wealth: Lessons from Two Centuries of Inheritance Taxation". In: *American Political Science Review* 106.01, pp. 81–102.
     - La Ferrara, Eliana, Alberto Chong, and Suzanne Duryea (2012). "Soap Operas and Fertility: Evidence from Brazil". In: *American Economic Journal: Applied Economics* 4.4, pp. 1–31.

5. *(April 14th & 16th)* Discussions of progress reports.

   - **Progress reports due April 12th**
   - **Final paper due April 21st**