ADVANCED CAUSAL INFERENCE DESIGNS AND APPLICATIONS

APSTA-GE 2018
Course Syllabus – Spring 2023

Professor
Joseph Cimpian
Kimball Hall, Room 203
Email: joseph.cimpian@nyu.edu
Office hours: By appointment

Class Meeting
Thursday, 2:00pm – 3:40pm
Kimball Hall, Room 211

Course Description

Social scientists and education researchers have become increasingly interested in the causal effects of specific policies or practices. For example, an education policy researcher may want to know if policy A produces higher graduation rates than policy B does for low-income high-school students. And if there is a difference in graduation rates between the two groups of students under the different policies, what is the magnitude of the effect?

Experiments (e.g., randomized control trials) are often referred to as the "gold standard" for determining the effect of a policy or practice (relative to some other policy or practice) on a population of interest. However, experimental designs can in some cases be time-consuming, costly, unethical, or otherwise impractical. Given these considerations and the wealth of already existing observational data, researchers have crafted careful approximations to randomized control trials that utilize the already existing data to learn more about social science phenomena. This class of research designs is referred to as "quasi-experimental designs." Although the counterfactual model framework these designs are based on is not a specific statistical technique, so much as it is a precise logic, a number of statistical techniques have come to be associated with quasi-experimental designs, which you will have already been exposed to in prior courses and (perhaps) your own research.

This course focuses on advanced topics and applications in two classes of designs: (1) **regression discontinuity designs** and (2) **matching estimators**. For each of these design classes, we will go in depth with the theory, assumptions and plausibility checks, applications, recent advances, combinations of these designs with other techniques (e.g., matching with difference-in-differences to create a synthetic control design), and complications of the designs. Throughout the course, you will be applying your knowledge of these designs to weekly memos and in-class discussions, as well as to an **original analysis** (with a large dataset you have identified and have access to) that will culminate in a journal-quality manuscript and final in-class presentation at the end of the semester.

Learning Objectives

By the end of the course, students should be able to:

- Identify when a specific statistical technique will be beneficial for assessing causality;

- Explain the assumptions required in order for specific combinations of quasi-experimental models to yield causal estimates;
- Evaluate the plausibility of the assumptions required for causality in a variety of applications;
- Through class discussions, hone critical analysis and discussion skills;
- Through a semester-long class project, design and execute an original statistical data analysis, explaining the logic of the approach taken, its strengths and weaknesses, and the conclusions and next steps of the research;
- Demonstrate analytical thinking, reading, writing and speaking skills.

Course Prerequisites

Graduate applied statistics course covering multiple regression (equivalent to APSTA-GE 2003: Intermediate Quantitative Methods: The General Linear Model) and graduate course in causal inference (equivalent to APSTA-GE 2012 Causal Inference).

Course Requirements

1. Attendance: Each class meeting centers a group discussion of weekly readings. Therefore, it is critical that you attend every class and arrive on time. If you must be absent or late, please let me know beforehand.

2. Weekly readings, discussion, and participation: I expect you to complete ALL of the readings for each class. Each class meeting will center around discussions on the assigned journal article readings due that week, and these discussions will count toward your participation grade.

3. Leading discussion: You will take turns leading the weekly discussions. Each of you will lead twice (once during the RD weeks, and once again during the matching weeks); more often if necessary. When leading discussion, you will introduce the papers and give a summary of the main points. You will look for and discuss overarching themes, and then present several questions to the class to spur discussion.

4. Weekly memos: Each week, you will submit a SHORT (2-3 pages maximum, double-spaced) critique of the assigned reading. Rather than summarizing the work, the critique should identify AND evaluate the most important claim or claims in the work. What did the author set out to do? What method was used, and what critical assumptions were required? Who can the results generalize to? How does the article contribute to knowledge (in your view, not the author's)? Does it succeed? What are the shortcomings? What are your lingering questions? Upload your memo to the NYU Classes course page (under the "Forum" tab, and in the folder for the week) by 12pm on the Wednesday before the reading is due.

5. Course project, final paper, and presentation

This final project will be conducted independently (i.e., not a group project). You will decide on a topic you would like to study for this course. The project must concern a question that can be

addressed through quasi-experimental designs discussed in this course (i.e., must be studying the effect of a cause on a specific outcome for a specific population using primarily a RDD or matching), and it cannot be a pure experimental design (i.e., randomized control trial).

The project will culminate in a research "journal-ready" manuscript that must include all of the following:

1) A brief critical review of the literature in a specific field, focusing particularly on the literature's strengths and weaknesses regarding causal inference, and identifying a research gap that your project will fill. This review should identify a specific causal question and then review the literature (in whatever fields are relevant) that attempts to answer this question. Depth, detail, and critical reading are valued here over breadth and vague generalities. Note that the purpose of this review is *not* a summary of the findings in the literature (e.g., what a typical literature review might look like), but rather a critical look at the analyses used to arrive at the conclusions in the literature.
2) A thoughtfully designed research study that uses one or more of the methods discussed in the course. Be sure to carefully state the following: the research question, the POI, the POS, the POCI, and the dataset being analyzed. Relevant equations should be included, and they must be explained in plain English, so that a lay audience can understand the analyses being conducted.
3) A data analysis that uses one or more of the methods discussed in the course to analyze extant data to make valid causal inference. Include a description of the rationale for the methods of analysis used, and a discussion of the assumptions made. Interpret the results, and clearly state the limitations of the study—there are sure to be limitations, so it's better you point them out before others (including me) point them out.

[As a guideline, bullets 1-3 immediately above should total approximately 8,000-12,000 words. Bullets 4-8 should not be included in that word count.]

4) An abstract of between 120-200 words.
5) A reference list in APA format (and use APA format throughout your paper).
6) Provide sensitivity analyses and robustness checks as appropriate for your questions, data, and design.
7) Your paper should include tables and figures that support your analyses and conclusions. [Again, these tables and figures do not count toward the word count.]
8) Supplemental files should contain all code (with clearly labeled annotations, so I know what you are doing throughout the code) and any other relevant data or information (e.g., questionnaires, detailed variable description, supplemental analyses) that would not go in the main body of a research article.

This project will take place throughout the semester (not just the end), with the following four assignments:

- Part 1 (due by start of class week 3)

  Submit a 2-page summary of your intended final paper. The summary should include: (a) the topic of your assignment; (b) the research question of interest; (c) a brief

summary of the most relevant, existing literature pertaining to your research question, and (d) the data source you will be using for your analysis.

In addition, you will give a brief (approximately 2 minutes) "elevator talk" to the class that week and get some initial feedback from the group.

- Part 2 (due by start of class week 8)

  Submit a 5-page summary of your project thus far. This progress report should be further along than your initial submission on your intended project. You should have a much more directed focus to your project (e.g., you should have your dataset cleaned and be in the analysis stages). List any major findings up to that point, as well as any major obstacles. Be sure to include an updated version of your conceptual model, along with any equations in the analysis. Also, include all Stata (or R) code used for your study.

  In addition, you will give a longer (approximately 10-15 minutes) update on your progress to the class that week and get some new feedback from the group. Use this as an opportunity to get feedback on things you are stuck on.

- Part 3 (in class week 14)

  In-class presentations on your final paper.

- Part 4 (due by 5pm during day of last class meeting)

  Final papers due by 5pm.

Grading

Your grade for this course will be determined as follows:

- Class participation: 10%
- Leading discussion: 10%
- Weekly memos: 20%
- Parts 1 and 2 of final project: 20%
- Final paper and presentation (parts 3 and 4): 40%

Course readings

- Readings will be articles, distributed on the NYU Classes website.
- Supplemental text: Murnane, R. J., & Willett, J. B. (2011). *Methods matter: Improving causal inference in educational and social science research*. New York: Oxford University Press.

Statistical software

Stata or R is required for this course, and it is assumed you have familiarity with whichever program you choose in advance of the course. You may use whichever you wish for your final project.

Other class information

1. NYU Classes: All required readings for this course will be made available via NYU Classes. Enrollment in the course should automatically give you access to the class site. Check frequently for new materials and announcements.

2. Absences: Please see me immediately if you have any conflicts with the schedule or if you anticipate being absent due to religious observances.

3. Class etiquette: Please make an effort to be on time (I will do the same) and please turn off your cell phone and other digital distractions while in class.

4. Academic integrity: NYU Steinhardt policies on academic integrity will be *strictly enforced* in this class. You can read the school's official statement on academic integrity (http://steinhardt.nyu.edu/policies/academic_integrity). All submitted work must be that of the individual student.

5. Withdrawal: If you wish to withdraw from the course, please do so formally with the University Registrar. If you withdraw without authorization, you are at risk for receiving a failing grade for the course. If withdrawing, please try to do so before the posted Registrar's last day for graduate and undergraduate students to withdraw without receiving a "W" on their transcripts.

6. Accommodations: Any student requiring an accommodation due to a chronic psychological, visual, auditory, mobility and/or learning disability should register with and consult with the Moses Center for Students with Disabilities at 212-998-4980, 726 Broadway, 2nd floor (www.nyu.edu/csd). Of course, I am happy to provide any and all accommodations recommended by the Moses Center.

COURSE OUTLINE

Week 1	Course Introduction; Causal inference refresher; Populations; Regression discontinuity design theory and fundamentals
Week 2	Regression discontinuity design theory and fundamentals (continued)
Week 3	Functional form considerations; "Non-parametric" bandwidth selection **Assignment**: Give a 2-minute "elevator talk" of your proposed final paper, in addition to your Part 1 hand-in
Week 4	"Fuzzy" regression discontinuities (and review of instrumental variables); Discrete-valued rating scores
Week 5	Assumption plausibility checks, density checks, donut estimators; Regression kink designs
Week 6	Multiple rating score regression discontinuities
Week 7	Comparative/difference-in-regression discontinuities; RDs in time
Week 8	**Assignment**: Longer check-in with updates on final papers, in addition to your Part 2 hand-in
Week 9	Matching estimator theory review and review of propensity score matching in practice; Propensity score methods with survey-weighted data
Week 10	Machine learning and matching estimators
Week 11	Robustness checks/sensitivity analyses
Week 12	Combining matching with other techniques (e.g., matching with difference-in-differences to create a synthetic control design)
Week 13	**Assignment**: Presentations
Week 14	**Assignment**: Final Paper Due

READING LIST
Readings are due and discussed the week listed

Key
G = General/Practical; T = Technical; A = Applied

Week 1

No readings due.

Week 2

(G) Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, *142*(2), 615-635.

(G) Bloom, H. S. (2012). Modern regression discontinuity analysis. *Journal of Research on Educational Effectiveness*, *5*(1), 43-82.

(A) Martorell, P., & McFarlin Jr, I. (2011). Help or hindrance? The effects of college remediation on academic and labor market outcomes. *The Review of Economics and Statistics*, *93*(2), 436-454.

(A) Thistlethwaite, D. L., & Campbell, D. T. (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational psychology*, *51*(6), 309-317.

Week 3

(T) Gelman, A., & Imbens, G. (2019). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, *37*(3), 447-456.

(G) De la Cuesta, B., & Imai, K. (2016). Misunderstandings about the regression discontinuity design in the study of close elections. *Annual Review of Political Science*, *19*, 375-396.

(T/A) Cattaneo, M. D., Titiunik, R., & Vazquez-Bare, G. (2017). Comparing inference approaches for RD designs: A reexamination of the effect of head start on child mortality. *Journal of Policy Analysis and Management*, *36*(3), 643-681.

Week 4

(T) Feir, D., Lemieux, T., & Marmer, V. (2016). Weak identification in fuzzy regression discontinuity designs. *Journal of Business & Economic Statistics*, *34*(2), 185-196.

(A) Winters, M. A., & Cowen, J. M. (2012). Grading New York: Accountability and student proficiency in America's largest school district. *Educational Evaluation and Policy Analysis*, *34*(3), 313-327.

(T) Kolesár, M., & Rothe, C. (2018). Inference in regression discontinuity designs with a discrete running variable. *American Economic Review*, *108*(8), 2277-2304.

Week 5

(T) Cattaneo, M. D., Titiunik, R., & Vazquez-Bare, G. (2019). The Regression Discontinuity Design. *arXiv preprint arXiv:1906.04242*.

(T/A) Barreca, A. I., Lindo, J. M., & Waddell, G. R. (2016). Heaping-induced bias in regression-discontinuity designs. *Economic Inquiry*, *54*(1), 268-293.

(G/T) Card, D., Lee, D. S., Pei, Z., & Weber, A. (2016). *Regression kink design: Theory and practice* (No. w22781). National Bureau of Economic Research.

(A) Dee, T. S., & Penner, E. K. (2017). The causal effects of cultural relevance: Evidence from an ethnic studies curriculum. *American Educational Research Journal*, *54*(1), 127-166.

Week 6

(G/T) Reardon, S. F., & Robinson, J. P. (2012). Regression discontinuity designs with multiple rating-score variables. *Journal of Research on Educational Effectiveness*, *5*(1), 83-104.

(A) Robinson, J. P. (2011). Evaluating criteria for English learner reclassification: A causal-effects approach using a binding-score regression discontinuity design with instrumental variables. *Educational Evaluation and Policy Analysis*, *33*(3), 267-292.

(A) Robinson-Cimpian, J. P., & Thompson, K. D. (2016). The effects of changing test-based policies for reclassifying English learners. *Journal of Policy Analysis and Management*, *35*(2), 279-305.

Week 7

(G) Hausman, C., & Rapson, D. S. (2018). Regression discontinuity in time: Considerations for empirical applications. *Annual Review of Resource Economics*, *10*, 533-552.

(G/A) Tang, Y., Cook, T. D., Kisbu-Sakarya, Y., Hock, H., & Chiang, H. (2017). The comparative regression discontinuity (CRD) design: An overview and demonstration of its performance relative to basic RD and the randomized experiment. In *Regression discontinuity designs: Theory and applications* (pp. 237-279). Emerald Publishing Limited.

(A) Anderson, M. L. (2014). Subways, strikes, and slowdowns: The impacts of public transit on traffic congestion. *American Economic Review*, *104*(9), 2763-96.

Week 8

No readings due.

Week 9

(G) Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, *22*(1), 31-72.

(G) DuGoff, E. H., Schuler, M., & Stuart, E. A. (2014). Generalizing observational study results: Applying propensity score methods to complex surveys. *Health services research*, *49*(1), 284-303.

(A) Reardon, S. F., Cheadle, J. E., & Robinson, J. P. (2009). The effect of Catholic schooling on math and reading development in kindergarten through fifth grade. *Journal of Research on Educational Effectiveness*, *2*(1), 45-87.

Week 10

(G/A) Lee, B. K., Lessler, J., & Stuart, E. A. (2010). Improving propensity score weighting using machine learning. *Statistics in Medicine*, *29*(3), 337-346.

(G/A) Linden, A., & Yarnold, P. R. (2016). Using machine learning to assess covariate balance in matching studies. *Journal of Evaluation in Clinical Practice*, *22*(6), 848-854.

(G/A) Hill, J., Weiss, C., & Zhai, F. (2011). Challenges with propensity score strategies in a high-dimensional setting and a potential alternative. *Multivariate Behavioral Research*, *46*(3), 477-513.

(G/A) McCaffrey, D. F., Ridgeway, G., & Morral, A. R. (2004). Propensity score estimation with boosted regression for evaluating causal effects in observational studies. *Psychological Methods*, *9*(4), 403.

(G/A) McCaffrey, D. F., Griffin, B. A., Almirall, D., Slaughter, M. E., Ramchand, R., & Burgette, L. F. (2013). A tutorial on propensity score estimation for multiple treatments using generalized boosted models. *Statistics in Medicine*, *32*(19), 3388-3414.

Week 11

(G) Lee, B. K., Lessler, J., & Stuart, E. A. (2011). Weight trimming and propensity score weighting. *PloS one*, *6*(3), e18174.

(A) TBD

(A) TBD

(A) TBD

Week 12

(G/A) Abadie, A., Diamond, A., & Hainmueller, J. (2010). 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.

(A) TBD

(A) TBD

(A) TBD

Week 13

No readings due.

Week 14

No readings due.