PAF 9177:
Advanced Quantitative Methods

Instructor: Thomas Leavitt

[Semester Redacted]

E-mail: [Email Redacted]
Office Hours: By Appointment via [Zoom Link Redacted]

Class Hours: 6:05 - 9:00 PM
Class Room: [Classroom Redacted]

Official Course Description

Intended for students interested in advanced quantitative research methods used in policy analysis, this course focuses on causal effects, especially of programs or policies. Topics include random assignment, multiple regression, instrumental variables, and difference-in-differences estimation. Students learn these approaches and techniques through hands-on projects and exercises on contemporary policy problems using real data and statistical software.

Course Modality

This course modality is In-Person. We will meet in [Classroom Redacted], located in the Newman Vertical Campus (55 Lexington Ave, New York, NY 10010).

Course Organization

This course begins by covering many of the topics introduced in PAF 9272 (Causal Analysis and Inference), but does so with greater depth, rigor and accompanying statistical software.

The course first covers concepts and methods to draw conclusions about policy impacts from comparisons of intervention and control groups, starting first with randomized experiments, also known as an A/B tests or RCTs (randomized controlled trials). The course's emphasis is on *randomization inference* whereby research design forms the foundation for subsequent statistical analysis. The course continues with this framework in its treatment of observational studies, which depend on stronger assumptions than randomized experiments, and sensitivity analyses in which we assess how our causal conclusions would change should these assumptions in observational studies be false.

The subsequent portion of the course offers a rigorous introduction to ordinary least squares (OLS) regression. The course then turns to two important classes of research designs: regression

discontinuity designs (RDDs) and pre-post/controlled pre-post designs. These two classes of designs draw upon regression-based tools and principles to draw causal inferences.

For all portions of the course, we will engage with real-world examples in a variety of policy domains. We will re-code, analyze and interpret existing data. We will also spend a considerable amount of time learning how to communicate via tables and figures generated in `R`.

Course Learning Goals

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

1. Articulate and develop causal claims in rigorous, counterfactual terms.
2. Understand statistical concepts, including the relationship between research designs and analysis strategies, as well as threats to reliable inference and how they can be addressed at both the design and analysis stages of research.
3. Critically evaluate existing causal analyses of both qualitative and quantitative data, to draw valid and useful conclusions using the framework(s) developed in this course.
4. Conduct and interpret the results of statistical analyses in `R`, including the estimation of average effects, variance estimation, tests of causal hypotheses, matching and others.
5. Code with statistical software (specifically `R` using the `tidyverse` collection of packages) to clean data, merge datasets, transform variables and visualize data.
6. Conduct research and compile findings in reproducible research reports, with tables, figures and embedded code, in `R Markdown` or related publishing systems

Course Assignments

Each student should complete 6 problem sets, each of which is due by 5:00 PM on the date listed in the course schedule below. I will endeavor to return all graded assignments to students within one week from when students have turned in the assignments.

Each problem set will be anchored around specific real-world applications:

- "Get out the vote" (GOTV) experiment in Kansas City, Missouri (Arceneaux, 2005)
- Survey experiment about the effects of television programming on attitudes (Albertson and Lawrence, 2009)
- Matched observational design about whether United Nations (UN) peacekeeping interventions promote peace (Gilligan and Sergenti, 2008)
- RDD about the effects of class size on student achievement (Angrist and Lavy, 1999)
- Controlled pre-post design about a gun policy's effects on violent crime (Webster et al., 2014; Hasegawa et al., 2019)

Students should also complete a final data analysis project by 5:00 PM on [Due Date Redacted]. In the process of completing this final project, students should also complete a final project draft (due by the start of class on [Due Date Redacted]). The final projects will consist of an extended analysis and report pertaining to one of the aforementioned applications from the problem sets.

Each week, students should also complete the `R` assignments on the course's DataCamp site, which can be accessed [DataCamp Link Redacted]. Each DataCamp assignment automatically expires by the beginning of class (6:05 PM) on the day the assignment is due. You will not be able to complete each assignment afterward.

Grades

I will use the standard Baruch College grading scale. I reserve the right to curve the scale depending on the distribution of class scores at the end of the semester. No curve will lower any individual grades relative to what they would have been without the curve. Final course grades will weight the course assessments using the following percentages:

- **45%** of your grade will be determined by 6 problem sets (**7.5%** each).
- **35%** of your grade will be determined by the final project
- **10%** of your grade will be determined by completion of weekly DataCamp exercises in `R`
- **10%** of your grade will be determined by course attendance and participation

Course Readings

This course is ZTC ("Zero Textbook Cost"). All course readings will be available on the course's Brightspace site. The course's readings will draw primarily from the following books:

1. Gerber, Alan S. and Donald P. Green. (2012). *Field Experiments: Design, Analysis, and Interpretation*. New York, NY: W. W. Norton.
2. Green, Donald P. (2022). *Social Science Experiments: A Hands-on Introduction*. New York, NY: Cambridge University Press.
3. Rosenbaum, Paul R. (2017). *Observation and Experiment: An Introduction to Causal Inference*. Cambridge, MA: Harvard University Press.
4. Rosenbaum, Paul R. (2023). *Causal Inference*. Cambridge, MA: The MIT Press.

In addition to texts focused on statistical and causal inference methods, the readings for virtually every week include a corresponding article (or articles) applying those methods. The applications are drawn from an electic set of fields, including economics, political science, criminology, public health and others.

Course Standards

Electronics

To the extent possible, **all students should bring to each class a personal laptop** to use for class-related purposes, mostly in-class exercises using statistical software. However, please do not use your cell phones or other electronic devices in class. If there are any students for whom cell phones or other electronics are important for course participation, please let me know. I will adjust the electronics policy for the entire class accordingly without singling out anyone.

Attendance, Participation and Collaboration

It goes without saying that attendance and active participation are extremely important to learning from this course. Collaboration with other students is strongly encouraged. However, all take-home assignments should be written individually in students' own words.

If you need to miss class, please do **not** contact me ahead of time. The implications of missing class for the attendance and participation components of your grade are the same regardless of whether you contact me beforehand or not. Please only contact me if you have to miss a class in which you are scheduled to present material.

Incomplete and Late Assignments

If you think that you will be unable to complete an assignment by its due date because of extenuating circumstances, please **do** let me know ahead of time. I may change the assignment's deadline for the entire class. For the final project, if students turn it in late without prior consent from me, the project's grade will be penalized by one-third of a letter grade per day (e.g., an A- assignment that a student turns in one day late becomes a B+). For problem sets, because I typically review them in class, I will not accept late assignments for credit.

Zoom Recordings

I can post recordings of class sessions on Zoom to the course's Brightspace site.

Students who participate in this class with their camera on or use a profile image are agreeing to have their video or image recorded solely for the purpose of creating a record for students enrolled in the class to refer to, including those enrolled students who are unable to attend live. If you are unwilling to consent to have your profile or video image recorded, be sure to keep your camera off and do not use a profile image. Likewise, students who un-mute during class and participate orally are agreeing to have their voices recorded. If you are not willing to consent to have your voice recorded during class, you will need to keep your mute button activated and communicate exclusively using the "chat" feature, which allows students to type questions and comments live.

Use of Artificial Intelligence (AI) Tools

Learning the material in this course requires thoughtful completion of all exercises in the course through the completion of your own work. You may use AI tools thoughtfully and with attribu-

tion when it supports your learning. If you do use AI tools for an assignment, include at the end of the assignment a brief explanation of which AI tool you used (e.g., Grammarly, ChatGPT, etc.) to accomplish which particular tasks (e.g., coding, language and grammar assistance, etc.)

Suitable uses of AI tools include brainstorming initial ideas, language and grammar assistance, formatting documents, and coding (among others). For other uses, I caution that the content generated by AI tools is frequently (if not usually) incorrect. AI tools will likely be helpful only insofar as they are a complement to (not a replacement of) thoughtful completion of the course material.

In addition, note that content generated via AI tools may not properly attribute or cite sources. You must nevertheless make sure to adhere to the academic integrity standards above. Failure to do so will be regarded as a violation of academic integrity like any other.

Academic Integrity

I will not tolerate violations of academic integrity under any circumstances. Please familiarize yourself with CUNY's Academic Integrity Policy. Additional resources to consult are Graff and Birkenstein (2021), Lipson (2008) and Neville (2016). If you are having difficulty with a written assignment, do not attempt to present another author's work as your own; please come talk to me instead.

Instructions on how to properly cite generative AI are here.

Academic Student Support

Baruch's Marxe School offers academic student support services for both writing and quantitative skills. I highly recommend looking into the Quantitative Student Support department's tutoring services and workshops. More information about these tutoring services and workshops is available here.

Software and computing

This course will use the `R` statistical programming language. To download `R`, please go to `https://cloud.r-project.org` and then follow the relevant instructions for installation.

Once you have successfully installed `R`, please make sure to download and install `RStudio`, which is an integrated development environment for `R`. You can download and install `RStudio` at `https://posit.co/download/rstudio-desktop/`.

You should install both `R` and `RStudio` on a personal laptop that you bring to every class. If you do not have access to a personal laptop or other technology, Baruch College's Technology Loan Program may be able to help. You can find more information about this Technology Loan Program here. You may also want to consider use of Posit Cloud, which enables you to use `R` and `RStudio` directly in an internet browser without downloading them to a personal device.

A valuable resource for learning `R` is Wickham et al. (2023), which is available for free at `https://r4ds.hadley.nz/`.

In addition, to facilitate students' learning `R`, all students should complete all assignments for the course's DataCamp site [DataCamp Link Redacted].

AI tools can be especially helpful for coding with statistical software, especially `R`. For help with such coding, I recommend ChatGPT, which you can access here. Use of ChatGPT requires signing up for an account, but a free plan is available.

Course Schedule and Readings

Baruch College's academic calendar is available via the Office of the Registrar here.

Class 1: Introduction to course and `R` programming language

- ⋆ Problem set 1 distributed

Recommmended readings

Wickham et al. (2023): Available here

Class 2: Causality and research design

- ⋆ Problem set 2 distributed
- ⋆ Problem set 1 due
- ⋆ DataCamp Assignment due
  - Introduction to the `Tidyverse`: Data wrangling

Required readings

Theory: Green (2022, pp. 1–22) and Rosenbaum (2023, pp. 1–45)
Application(s): Mousa (2020)

Recommended readings

Green (2022, pp. 42-66)

Class 3: Estimation and inference of average effects in randomized experiments

- ⋆ DataCamp Assignment due
  - Introduction to the `Tidyverse`: Data visualization

Required readings

Theory: Rosenbaum (2017, pp. 16–28) and Gerber and Green (2012, pp. 51–66)
Application(s): Mousa (2020)

Class 4: Exact hypothesis tests in randomized experiments

- ⋆ DataCamp Assignment due
  - Introduction to the `Tidyverse`: Grouping and summarizing

Required readings

Theory: Rosenbaum (2017, pp. 30–52)
Application(s): Mousa (2020)

Class 5: Covariance adjustment in randomized experiments

- ⋆ DataCamp Assignment due
  - Introduction to the `Tidyverse`: Types of visualizations

Required readings

Theory: Gerber and Green (2012, pp. 71–79, and pp. 95–102)
Application(s): Mousa (2020)

Class 6: Noncompliance and loss to follow-up in randomized experiments

- ⋆ Problem set 3 distributed
- ⋆ Problem set 2 due
- ⋆ DataCamp Assignment due
  - Introduction to Data Visualization with `ggplot2`: Introduction

Required readings

Theory: Rosenbaum (2017, pp. 261–270) and Gerber and Green (2012, pp. 211-219)
Application(s): Adams and Smith (1980) and Nickerson (2008)

Class 7: Natural experiments

- ⋆ DataCamp Assignment due
  - Introduction to Data Visualization with `ggplot2`: Aesthetics

Required readings

Theory: Rosenbaum (2023, pp. 117–128)
Application(s): Clingingsmith et al. (2009)

Class 8: Observational designs and adjustment for measured covariates

⋆ DataCamp Assignment due

– Introduction to Data Visualization with `ggplot2`: Geometries

Required readings

Theory: Rosenbaum (2017, pp. 65–77)
Application(s): Silber et al. (2022)

Recommended readings

Rosenbaum (2023, pp. 47–65)

Class 9: Matching methods

⋆ Problem set 4 distributed

⋆ Problem set 3 due

⋆ DataCamp Assignment due

– Introduction to Data Visualization with `ggplot2`: Themes

Required readings

Theory: Hansen (2011)
Application(s): Cerdá et al. (2012)

Recommended readings

Rosenbaum (2017, pp. 212–233)
Rosenbaum (2023, pp. 67–84)

Class 10: Sensitivity to unmeasured covariates in observational designs

⋆ DataCamp Assignment due

– Reshaping Data with `tidyr`: Tidy Data

Required readings

Theory: Rosenbaum (2017, pp. 170–181) and Rosenbaum (2023, pp. 85–102)
Application(s): Cornfield et al. (1959)

Class 11: Ordinary Least Squares (OLS) regression

- ⋆ Problem set 5 distributed
- ⋆ Problem set 4 due
- ⋆ DataCamp Assignment due
  - Reshaping Data with `tidyr`: From Wide to Long and Back

Required readings

Theory: Freedman et al. (2007, pp. 158–217)

Recommended readings

Gelman et al. (2021, pp. 81–112 and pp. 131–152)

Class 12: Regression discontinuity designs (RDDs)

- ⋆ DataCamp Assignment due
  - Reshaping Data with `tidyr`: Expanding Data

Required readings

Theory: Lee and Lemieux (2010)
Application(s): Klašnja and Titiunik (2017)

Recommended readings

Cattaneo et al. (2020, 2024)

Class 13: Pre-post & controlled pre-post designs using regression

- ⋆ Problem set 6 distributed
- ⋆ Problem set 5 due
- ⋆ DataCamp Assignment due
  - Reshaping Data with `tidyr`: Rectangling Data

Required readings

Theory: Bloom (2003) and Lechner (2011)
Application(s): Cruz-Cano and Mead (2019) and Montalvo (2011)

Class 14: Workshop and discussion of final project drafts

- ⋆ Problem set 6 due
- ⋆ Final project due on [Due Date Redacted]
- ⋆ DataCamp Assignment due
  - Reporting with `R Markdown`

Baruch Policies and Resources

Equal Opportunity and Non-Discrimination

You can submit a report of discrimination and/or retaliation via CUNY's centralized reporting platform here. For more information about CUNY's Policy on Equal Opportunity and Non-Discrimination ("EO Policy"), please access it here. You can also access CUNY's revised Policy on Sexual Misconduct here.

Student Disability Services

Baruch College's Student Disability Services provides accommodations to students with disabilities to promote equal access to Baruch's programs and services. To request accommodations, please do so as soon as possible by directly contacting Student Disability Services here.

Campus Intervention Team (CIT)

The CIT provides a support system for students in crisis. Any member of Baruch's community (faculty, students, staff) can report a concern about a student to CIT. **All reports are confidential.** You can learn more about CIT and the services it offers here.

To report a concern related to concerning behavior, student conduct, academic integrity or student grievance, please use Baruch College's Portal for Incident Reporting available here.

Writing Support

The Schwartz Communication Institute

Writing support services for graduate students in the Marxe school are available through a partnership with the Bernard L. Schwartz Communication Institute at the Center for Teaching and Learning. Please reach out to the Assistant Director for Writing in Public and International Affairs, Melina Moore, for more information on one-to-one writing consultations. You can also visit the website for writing resources and workshops designed to support graduate students at Marxe. You can learn more here.

The Baruch Writing Center

The Writing Center offers free, professional writing support for all undergraduate and graduate students at Baruch, through one-to-one consultations, workshops, peer review groups, written feedback, online resources, and a journal of outstanding student writing. The Writing Center support faculty through classroom visits, in- class workshops, referral forms, and workshop lesson plans, and it is always available for conversations about teaching and writing. More information is available here, by calling 646-312-4012 or by emailing the center at writing.center@baruch.cuny.edu.

Academic Advisement

The Marxe school provides a range of academic advisement services to ensure the successful completion of students' degree programs. For undergraduate advisement (Bachelor of Science

in Public Affairs BSPA), you can write an e-mail to mspia.bspa@baruch.cuny.edu. For graduate advisement (Master of Public Administration, Executive MPA, Master of International Affairs, and Master of Science in Education-HEA or any of Marxe's certificate programs), you can write an e-mail to mspia.advisement@baruch.cuny.edu.

You can also schedule an appointment with an academic advisor via Baruch College's Navigate platform available here.

Career Services and Resources

For personalized career guidance, please make an appointment by logging in to Marxe Career-Connect. For general inquiries, you can write an e-mail to mspia.careerservices@baruch.cuny.edu. More information about career services is available here; a variety of career resources, including important links, documents and opportunities, is available here.

References

Adams, W. C. and D. J. Smith (1980). Effects of telephone canvassing on turnout and preferences: A field experiment. *The Public Opinion Quarterly 44*(3), 389–395.

Albertson, B. and A. Lawrence (2009). After the credits roll: The long-term effects of educational television on public knowledge and attitudes. *American Politics Research 37*(2), 275–300.

Angrist, J. D. and V. Lavy (1999). Using maimonides' rule to estimate the effect of class size on scholastic achievement. *The Quarterly Journal of Economics 114*(2), 533–575.

Arceneaux, K. (2005). Using cluster randomized field experiments to study voting behavior. *The Annals of the American Academy of Political and Social Science 601*(1), 169–179.

Bloom, H. S. (2003). Using "short" interrupted time-series analysis to measure the impacts of whole-school reforms: With applications to a study of accelerated schools. *Evaluation Review 27*(1), 3–49.

Cattaneo, M. D., N. Idrobo, and R. Titiunik (2020). *A Practical Introduction to Regression Discontinuity Designs: Foundations*. Elements in Quantitative and Computational Methods for the Social Sciences. New York, NY: Cambridge University Press.

Cattaneo, M. D., N. Idrobo, and R. Titiunik (2024). *A Practical Introduction to Regression Discontinuity Designs: Extensions*. Elements in Quantitative and Computational Methods for the Social Sciences. New York, NY: Cambridge University Press.

Cerdá, M., J. D. Morenoff, B. B. Hansen, K. J. T. Hicks, L. F. Duque, A. Restrepo, and A. V. Diez-Roux (2012). Reducing violence by transforming neighborhoods: A natural experiment in medellín, colombia. *American Journal of Epidemiology 175*(10), 1045–1053.

Clingingsmith, D., A. I. Khwaja, and M. Kremer (2009). Estimating the impact of The Hajj: Religion and tolerance in Islam's global gathering. *The Quarterly Journal of Economics 124*(3), 1133–1170.

Cornfield, J., W. Haenszel, E. C. Hammond, A. M. Lilienfeld, M. B. Shimkin, and E. L. Wynder (1959). Smoking and lung cancer: Recent evidence and a discussion of some questions. *JNCI: Journal of the National Cancer Institute 22*(1), 173–203.

Cruz-Cano, R. and E. L. Mead (2019). Causes of excess deaths in Puerto Rico after Hurricane Maria: A time-series estimation. *American Journal of Public Health 109*(7), 1050–1052.

Freedman, D., R. Pisani, and R. Purves (2007). *Statistics* (4th ed.). New York, NY: W. W. Norton & Company.

Gelman, A., J. Hill, and A. Vehtari (2021). *Regression and Other Stories*. New York, NY: Cambridge University Press.

Gerber, A. S. and D. P. Green (2012). *Field Experiments: Design, Analysis, and Interpretation*. New York, NY: W.W. Norton.

Gilligan, M. J. and E. J. Sergenti (2008). Do UN interventions cause peace? Using matching to improve causal inference. *Quarterly Journal of Political Science 3*(2), 89–122.

Graff, G. and C. Birkenstein (2021). *They Say / I Say* (5th ed.). New York, NY: W. W. Norton & Company.

Green, D. P. (2022). *Social Science Experiments: A Hands-on Introduction*. New York, NY: Cambridge University Press.

Hansen, B. B. (2011). Propensity score matching to extract latent experiments from nonexperimental data: A case study. In N. J. Dorans and S. Sinharay (Eds.), *Looking Back: Proceedings of a Conference in Honor of Paul W. Holland*, Volume 202 of *Lecture Notes in Statistics*, Chapter 9, pp. 149–181. New York, NY: Springer.

Hasegawa, R. B., D. W. Webster, and D. S. Small (2019). Evaluating Missouri's handgun purchaser law: A bracketing method for addressing concerns about history interacting with group. *Epidemiology 30*(3), 371–379.

Klašnja, M. and R. Titiunik (2017). The incumbency curse: Weak parties, term limits, and unfulfilled accountability. *The American Political Science Review 111*(1), 129–148.

Lechner, M. (2011). The estimation of causal effects by difference-in-difference methods. *Foundations and Trends in Econometrics 4*(3), 165–224.

Lee, D. S. and T. Lemieux (2010). Regression discontinuity designs in economics. *Journal of Economic Literature 48*(2), 281–355.

Lipson, C. (2008). *Doing Honest Work in College: How to Prepare Citations, Avoid Plagiarism, and Achieve Real Academic Success* (2nd ed.). Chicago, IL: University of Chicago Press.

Montalvo, J. G. (2011). Voting after the bombings: A natural experiment on the effect of terrorist attacks on democratic elections. *The Review of Economics and Statistics 93*(4), 1146–1154.

Mousa, S. (2020). Building social cohesion between Christians and Muslims through soccer in post-ISIS Iraq. *Science 369*(6505), 866–870.

Neville, C. (2016). *The Complete Guide to Referencing and Avoiding Plagiarism* (3rd ed.). Buckingham, UK: Open University Press.

Nickerson, D. (2008). Is voting contagious? Evidence from two field experiments. *The American Political Science Review 102*(1), 49–57.

Rosenbaum, P. R. (2017). *Observation and Experiment: An Introduction to Causal Inference*. Cambridge, MA: Harvard University Press.

Rosenbaum, P. R. (2023). *Causal Inference*. The MIT Press Essential Knowledge Series. Cambridge, MA: The MIT Press.

Silber, J. H., P. R. Rosenbaum, J. G. Reiter, A. S. Hill, S. Jain, D. A. Wolk, D. Small, S. Hashemi, B. A. Niknam, M. D. Neuman, L. A. Fleisher, and R. Eckenhoff (2022). Alzheimer's dementia after exposure to anesthesia and surgery in the elderly: A matched natural experiment using appendicitis. *Annals of Surgery 276*(5), e377–e385.

Webster, D., C. K. Crifasi, and J. S. Vernick (2014). Effects of the repeal of Missouri's handgun purchaser licensing law on homicides. *Journal of Urban Health: Bulletin of the New York Academy of Medicine 91*(2), 293–302.

Wickham, H., M. Çetinkaya Rundel, and G. Grolemund (2023). *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data* (2nd ed.). Sebastopol, CA: O'Reilly Media.