27:202:641:01
QUANTITATIVE METHODS FOR CAUSAL INFERENCE
Fall 2023 Syllabus

Professor:	Robert (Bob) Apel, Ph.D.
Office:	Center for Law and Justice, Room 579H
E-mail:	robert.apel@rutgers.edu

Class Location:	Center for Law and Justice, Room 572
Class Time:	Tuesday, 4:00-6:40
Office Hours:	By appointment

Teaching Assistant:	Raven Lewis
E-mail:	ral221@scj.rutgers.edu

COURSE DESCRIPTION

This course provides a survey of a large class of methods designed for causal inference from observational or non-experimental data. Evaluation of the impact of some "treatment" from observational data is complicated by the problem of confounding or selection bias—the technical expression is endogenous treatment assignment. This describes a situation where units that are unusually prone to experience treatment might also be unusually likely to have a higher- or lower-than-average mean outcome, for reasons having nothing to do with treatment. Thus, even if there is no causal effect of treatment on the outcome, there could appear to be one because of selection bias. A classic example is the question of the impact of educational attainment on earnings, where people who attain higher education, on average, tend to have more innate academic ability or a more advantaged family background, among other sources of confounding. This means people who attain higher education would probably have higher earnings even if they did not attain higher education, at least in part. The task of the analyst is to isolate the portion of the association between education and earnings that plausibly reflects the causal effect of the former on the latter.

The theory and best practices of the major quasi-experimental designs and their associated statistical models will be reviewed in this course. As the name implies, the classic experimental design is the inspiration for most quasi-experimental designs. The latter are generally divided into two classes assuming different underlying selection mechanisms: (1) selection on observables and (2) selection on unobservables. A theme to be repeatedly emphasized is that statistical models are not a substitute for, and in fact are often inferior to, thoughtful *a priori* decisions concerning research design—a clearly defined and novel treatment, a comparison group that could have actually been treated, and an assignment mechanism that renders treatment as good as randomly assigned. A well-specified statistical model can then be used after these design choices to eliminate remaining sources of confounding. This framework is best captured by Donald Rubin's observation that "for objective causal inference, design trumps analysis." But analysis still matters greatly.

Each topic will be presented in a workshop format designed to introduce students to a specific type of quasi-experimental design associated statistical models, along with research application. Class discussion will be based on research reported in peer-review journal articles, and data analysis will provide concrete, step-by-step application of key methods. While applications will draw heavily from examples in criminology and criminal justice, the course material will cover broader applications from education, psychology, sociology, political science, economics, and public policy.

All statistical analyses in the course will be performed using R and RStudio. Prior versions of the course have been taught using Stata, so students wishing to have Stata materials will be able to obtain them from the instructor.

Course Objectives

- Describe the potential outcomes framework and the relevance of counterfactual thinking for making causal statements. Use the framework to clearly articulate the counterfactual for a given causal question.
- Identify key features of the classic experimental design, and classify different counterfactual designs in non-experimental settings by the assignment mechanism (e.g., selection on observables, selection on unobservables).
- Articulate the different statistical models for the analysis of observational data, including their underlying assumptions, data requirements, implementation, interpretation, and sensitivity analyses.
- Acquire familiarity with the use of R and RStudio for a wide range of research applications involving the use of quasi-experimental methods.

Course Prerequisite

The goal of this course is to make students knowledgeable consumers (and perhaps users) of the econometric methods that will be covered, under the assumption that one or more of the methods will be applicable to students' own research projects (e.g., comprehensive exam, dissertation, other empirical paper). The material will include readings as well as exercises in data analysis. Because the readings will often be of a technical nature, students are expected to have successfully completed the doctoral-level courses in introductory and intermediate statistics and research methods, and to be comfortable with the use of some statistical program (R, Stata). For doctoral students in the School of Criminal Justice, the prerequisites include passing grades in Introductory Statistics (27:202:542), Intermediate Statistics (27:202:543), and Research Methods (27:202:640). For doctoral students from other programs, this includes full coverage of linear regression analysis as well as maximum likelihood estimation (e.g., logistic regression analysis).

COURSE MATERIALS

In this course, we will use R and RStudio. Students are strongly encouraged to bring a laptop with them to class each week, because time will be devoted each week to data analysis. Before the start of the course, they should download and install the latest versions of R (https://cran.r-project.org) and RStudio Desktop (https://www.rstudio.com) compatible with their operating system. Because these are open-source programs, there will be no need to purchase any software licenses. Students who already have R and RStudio loaded on their laptop are advised to update them before the start of the course. Although Stata will not be used, notes on Stata will be available for all of the designs and statistical models in the course, for students who wish to have them.

Powerpoint slides covering weekly material will be made available by the instructor, along with the data files and R scripts used for in-class exercises. Students should print up the slides before each class and use them for notetaking purposes, and have the data files and R scripts loaded onto their laptops and ready for analysis.

COURSE GRADING

Course grading will be based on the following criteria, described in more detail below:

Class Preparation	40%
Do It Yourself	60%
	100%

The grading scale that will be used for the final semester grades is as follows:

A	90.0% or higher
B	80.0% to 89.9%
C	70.0% or 79.9%
F	69.9% or lower

Class Preparation (40%)

Students are expected to attend each class meeting and to have read and to be conversant with (to the extent possible) all of the required reading material. Some of this material will be of a technical nature, so the goal of the class meetings will be to help students understand what they have read (both conceptually and mathematically), and to work through empirical applications of key concepts. The weekly class meetings will also be an opportunity for students to bring their own questions or data challenges to the attention of the instructor and their classmates.

Do It Yourself (60%)

Each week, a short assignment requires students to expand on work performed during class, using the same dataset and design. These tasks might entail including an alternative set of variables, or estimating an alternative model, for example. Students will write up a short memo with relevant commands and output from the assignment, along with a brief narrative description of the findings. These are expected to be no more than 1-2 pages in length.

COURSE POLICIES

Class Announcements

As needed, e-mail will be utilized to post course announcements (e.g., class cancellation due to inclement weather) as well as to occasionally provide links to items that are relevant for the topics covered in this course (e.g., newspaper articles, journal articles).

Classroom Climate

Disruptive behavior in the classroom cheats other students of the opportunity to learn. Examples include arriving late to class, leaving and re-entering the classroom during the seminar, talking excessively, using cell phones, eating, reading outside material, and persisting in speaking without being recognized. The instructor reserves the right to ask disruptive students to leave the classroom.

Academic Integrity

The instructor will uphold Rutgers University policies concerning ethical behavior and academic integrity, and students are expected to familiarize themselves with these policies. The relevant principles, policies, and disciplinary procedures can be accessed from the university's website at http://academicintegrity.rutgers.edu.

ACCOMMODATION AND SUPPORT STATEMENT

Rutgers University Newark (RU-N) is committed to the creation of an inclusive and safe learning environment for all students and the University as a whole. RU-N has identified the following resources to further the mission of access and support:

For Individuals Experiencing Disability: The Office of Disability Services (ODS) works with students with medical, physical, and/or mental conditions who encounter disabling barriers in order to determine reasonable and appropriate accommodations for access. Students who have completed the process with ODS and have approved accommodations are provided a Letter of Accommodation (LOA) specific to each course. To initiate accommodations for their course students must both provide the LOA to and have a conversation with the course instructor about the accommodations. This should occur as early in the semester as possible. More information can be found at the RU-N ODS website (ods.newark.rutgers.edu). Contact ODS at (973) 353-5375 or via email at ods@newark.rutgers.edu.

For Individuals Who Are Pregnant: The Office of Title IX and ADA Compliance is available to assist with any concerns or potential accommodations related to pregnancy. Students may contact the Office of Title IX and ADA Compliance at (973) 353-1906 or via email at TitleIX@newark.rutgers.edu.

For Short-Term Absence Verification: The Office of the Dean of Students can provide assistance for absences related to religious observance, emergency or unavoidable conflict (illness, personal or family emergency, etc.). Students should refer to University Policy 10.2.7 for information about expectations and responsibilities. The Office of the Dean of Students can be contacted by calling (973) 353-5063 or emailing deanofstudents@newark.rutgers.edu.

For Individuals with Temporary Conditions/Injuries: The Office of the Dean of Students can assist students who are experiencing a temporary condition or injury (broken or sprained limbs, concussions, or recovery from surgery). Students experiencing a temporary condition or injury should submit a request using the following link: https://temporaryconditions.rutgers.edu.

For Gender or Sex-Based Discrimination or Harassment: The Office of Title IX and ADA Compliance can assist students who are experiencing any form of gender or sex-based discrimination or harassment, including sexual assault, sexual harassment, relationship violence, or stalking. Students can report an incident to the Office of Title IX and ADA Compliance by calling (973) 353-1906 or emailing TitleIX@newark.rutgers.edu. Incidents may also be reported by using the following link: tinyurl.com/RUNReportingForm. For more information, students should refer to the University's Title IX Policy and Grievance Procedures located at https://uec.rutgers.edu/wp-content/uploads/60-1-33-current-1.pdf.

For Support Related to Interpersonal Violence: The Office for Violence Prevention and Victim Assistance (VPVA) can provide any student with confidential support. The office does not have a reporting obligation to Title IX. Students can contact the office by calling (973) 353-1918 or emailing run.vpva@rutgers.edu. There is also a confidential text-based helpline available to students; students can text (973) 339-0734 for support. Students do not need to be a victim/survivor of violence; any student can receive services, information and support.

For Crisis and Concerns: The Campus Awareness Response and Education (CARE) Team works with students in crisis to develop a plan of support plan and address personal situations that might impact their academic performance. Connect with the CARE Team by using the following link: tinyurl.com/RUNCARE or emailing careteam@rutgers.edu.

For Stress, Worry, or Concerns about Well-Being: The Counseling Center has confidential therapists available to support students. Students should reach out to the Counseling Center to schedule an appointment: counseling@newark.rutgers.edu or (973) 353-5805.

Additional support is available to any RU-N student through Uwill services:

- Umatch: Teletherapy with flexible scheduling, starting with a free account.
- Uhelp: Crisis support at 833-646-1526 (available 24/7/365).
- Urise: Wellness-based video collection with a free account.

Access Uwill@RUN at https://my.rutgers.edu using your netid. Services are confidential and free.

For emergencies, call 911 or contact Rutgers University Police Department (RUPD) by calling (973) 353-5111.

COURSE SCHEDULE

Students are not expected to read all of the course readings each week, but they are asked to read some of them and to be minimally conversant with the causal inference method used. This schedule is subject to change depending on time demands, and odds are that it will indeed change.

Class Date	Causal Design Topic	Readings
Tue, Sep 5	**Course Introduction** **Counterfactual Inference**	Neyman (1923); Rubin (1974, 1990); Holland (1986)
Tue, Sep 12	**Classic Experimental Design** *Bonus Methods*: Randomization Inference, Regression Adjustment, Lasso Regression	Heckman and Smith (1995); Sampson (2010); Athey and Imbens (2017)
Tue, Sep 19	(cont'd)	(cont'd)
Tue, Sep 26	**Natural Experimental Design** *Bonus Method*: Bootstrap Inference	Freedman (1991); Meyer (1995); Dunning (2008)
Tue, Oct 3	**Regression Discontinuity Design** *Bonus Methods*: Fractional Outcomes, Kernel Smoothing	Thistlethwaite and Campbell (1960); Imbens and Lemieux (2008); Lee and Lemieux (2010)
Tue, Oct 10	(cont'd)	(cont'd)
Tue, Oct 17	**Difference in Differences Design** *Bonus Methods*: Event Study, Staggered Difference in Differences	Ashenfelter (1978); Ashenfelter and Card (1985); Heckman and Smith (1999); Roth et al. (2023)
Tue, Oct 24	(cont'd)	(cont'd)
Tue, Oct 31	**Instrumental Variables Design** *Bonus Methods*: Endogenous Sample Selection, Marginal Treatment Effect	Imbens and Angrist (1994); Angrist et al. (1996); Angrist and Krueger (2001)
Tue, Nov 7	(cont'd)	(cont'd)
Tue, Nov 14	Catch-Up Day	NA
Tue, Nov 21	NO CLASS – THURSDAY SCHEDULE	NA
Tue, Nov 28	**Propensity Score Design** *Bonus Methods*: Mahalanobis Distance Matching, Coarsened Exact Matching, Inverse Probability Weighting, Doubly Robust Estimation, Entropy Balancing	Rosenbaum and Rubin (1983, 1984, 1985); Hirano et al. (2003); Imbens (2004); Smith and Todd (2005)
Tue, Dec 5	(cont'd)	(cont'd)
Tue, Dec 12	**Synthetic Control Design**	Abadie et al. (2015); Abadie (2021)

COURSE READINGS

Abadie, A. (2021). Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature* 59: 391-425.

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

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association* 91: 444-55.

Angrist, J. D., & Krueger, A. B. (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments. *Journal of Economic Perspectives* 15: 69-85.

Ashenfelter, O. (1978). Estimating the effect of training programs on earnings. *Review of Economics and Statistics* 60: 47-57.

Ashenfelter, O., & Card, D. (1985). Using the longitudinal structure of earnings to estimate the effect of training programs. *Review of Economics and Statistics* 67: 648-660.

Athey, S., & Imbens, G.W. (2017). The econometrics of randomized experiments. In Abhijit Vinayak Banerjee and Esther Duflo (Eds.), *Handbook of Economic Field Experiments, Vol. 1* (pp. 73-140). North-Holland: Elsevier.

Dunning, T. (2008). Improving causal inference: Strengths and limitations of natural experiments. *Political Research Quarterly* 61: 282-293.

Freedman, D. A. (1991). Statistical models and shoe leather. *Sociological Methodology* 21: 291-313.

Heckman, J. J., & Smith, J. A. (1995). Assessing the case for social experiments. *Journal of Economic Perspectives* 9: 85-110.

Heckman, J. J., & Smith, J. A. (1999). The pre-programme earnings dip and the determinants of participation in a social programme: Implications for simple programme evaluation strategies. *Economic Journal* 109: 313-348.

Hirano, K., Imbens, G. W., & Ridder G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica* 71: 1161-1189.

Holland, P. W. (1986). Statistics and causal inference (with comments and rejoinder). *Journal of the American Statistical Association* 81: 945-970.

Imbens, G. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and Statistics* 86: 4-30.

Imbens, G. M., & Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica* 62: 467-475.

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

Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature* 48: 281-355.

Meyer, B. D. (1995). Natural and quasi-experiments in economics. *Journal of Business and Economic Statistics* 13: 151-161.

Neyman, J. S. (1923 [1990]). On the application of probability theory to agricultural experiments (translated from Polish by D. M. Dabrowska and T. P. Speed). *Statistical Science* 5: 465-472.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika* 70: 41-55.

Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association* 79: 516-524.

Rosenbaum, P. R., & Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity. *American Statistician* 39: 33-38.

Roth, J., Sant'Anna, P. H. C., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics* 235: 2218-2244.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* 66: 688-701.

Rubin, D. B. (1990). Comment: Neyman (1923) and causal inference in experiments and observational studies. *Statistical Science* 5: 472-480.

Sampson, R. J. (2010). Gold standard myths: Observations on the experimental turn in quantitative criminology. *Journal of Quantitative Criminology* 26: 485-500.

Smith, J. A., & Todd, P. E. (2005). Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of Econometrics* 125: 305-353.

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