POLS GU4722: Statistical Theory and Causal Inference

Department of Political Science, Columbia University

Spring 2026

Lectures: Tuesdays and Thursdays, 4:10pm-5:25pm, 703 Hamilton Hall
Sections: Mondays, 10:10am-11:00pm, 707 Hamilton Hall

Instructor: John Marshall
Office: 705 International Affairs Building
Office hours: Wednesdays, 2:00-4:00pm; signup
Email: john.marshall@columbia.edu

Teaching assistant: Sydney White
Office: 736 International Affairs Building
Office hours: Tuesdays, 9:00am-11:00am
Email: slw2192@columbia.edu

Course overview

This is the second course in the graduate-level sequence on quantitative political methodology offered in the Department of Political Science.[1] It addresses one of the essential goals of social science: drawing causal inferences from data. The primary objective is to empower students to understand, evaluate, and implement (i) a framework and design-based identification strategies for making causal inferences from experimental and observational data and (ii) estimation strategies for causal inference designs. Topics include: randomized experiments, estimation under conditional ignorability, instrumental variables, regression discontinuity, difference-in-differences, causal mechanisms, and more complex inferences.

Enrollment

This course is intended for PhD, masters, and advanced undergraduate students interested in quantitative causal inference in the social sciences. This course will assume knowledge of mathematics, probability, and statistics, at the level of the content covered in POLS GU4700 and GU4720 (optimization, linear algebra, laws of probability, common estimators and their properties, inference and testing, etc.). As a guide, students should be familiar with the content of Simon and Blume's (1994) *Mathematics for Economists*. Students will also need to be able to use a programming language—whether Python, R, Stata, or another of their choice—to complete problem sets; the course will be demonstrated in R. Enrollment will be capped at 30 students.

[1] This course owes a debt of gratitude to Naoki Egami, who kindly shared his materials from previous iterations of the class.

Students who wish to audit the course are expected to complete all problem sets. If you wish to audit the course, please email the instructor about what the grading option you want to use (e.g. "P/F" or "R").

Course requirements

Final grades for the course will be assigned according to:

- **Problem sets** (40%). There will be six problem sets throughout the semester, released roughly every two weeks on CourseWorks with the completed assignment due a week later. Problem sets will include both analytical problems and data analysis questions. Mathematical derivations should include and explain all key steps, as you would expect from a proof in an academic paper. All code used to answer data analysis questions should be provided; any tables or figures should be appropriately formatted to resemble journal presentation styles (i.e. clear to a casual reader, use a sensible number of decimal points, use footnotes to provide additional details). Students are encouraged, but not required, to write answers in LaTeX, Rmarkdown in R, or markstat in Stata to develop their computational skills.

  Assignments will be submitted electronically via CourseWorks, and graded blind. No late submission will be accepted without prior approval from the instructor; approval will only be granted in case of a documented medical or family emergency.

  Students are encouraged to (inclusively) work in groups on problem sets, especially after tackling problems on their own first. However, each student must submit their own succinct write-up of solutions and note which students they worked with; duplicated answers will receive half credit.

  The tentative schedule for problem sets is as follows:

  - Problem Set 1: assigned on 1/31 and due on 2/6;
  - Problem Set 2: assigned on 2/13 and due on 2/20;
  - Problem Set 3: assigned on 2/28 and due on 3/6;
  - Problem Set 4: assigned on 3/20 and due on 3/27;
  - Problem Set 5: assigned on 4/3 and due on 4/10;
  - Problem Set 6: assigned on 4/17 and due on 4/24.

- **Midterm exam** (20%). The midterm exam will be inn-person on 3/12. This exam will cover the first half of the course materials. Details will be provide by the instructor later in the semester. The exam will be graded blind.

- **Final exam** (40%). The final exam will be taken in-person; subject to confirmation from the registrar, this will be the 4:10pm-7:00pm time slot on 5/12. This exam will cover materials from throughout the course. Details will be provide by the instructor later in the semester. The exam will be graded blind.

Computing

Students are welcome to use a statistical software of their choice, but will need to be able to program in that software, e.g. to construct estimators from scratch rather than use only canned commands. The course will principally be taught in R, an open-source statistical computing environment used widely in statistics and political science; you can download it for free from www.r-project.org. You can find a virtually endless set of resources for R on the internet. You may also be interested in using RStudio (https://rstudio.com), an editor and development environment for R.

If you have taken POLS GU4720 or its equivalent course where you use R for data analysis, you have the required level of mastery. If you have not used a statistical programming language before, it might be challenging for you to make the most of this course. If you have questions or concerns, please contact the instructor before committing to the class.

Class logistics

Lectures will be twice-weekly and start promptly due to the high volume of content to cover. The goal of lectures is to explain core theoretical ideas and demonstrate the relevance through prominent applied studies. Lectures will not focus on implementing methods in statistical software.

Sections will meet weekly, with a focus on reviewing key concepts, covering problem set solutions, and especially implementing methods covered in lecture in R. Attendance is mandatory to receive full credit for the class.

We will use CourseWorks to:

- Share all class announcements;
- Share class materials from lectures and sections; and
- Share and submit problem sets.

As well as sections and office hours, please use the Piazza Discussion Board at https://piazza.com to ask questions about lectures, problem sets, and other course materials. This enables all students to benefit from discussion and to help each other understand the materials. Both students and instructors are encouraged to participate in discussions and answer any questions that are posted. To join the POLS GU4722 Piazza site, click https://piazza.com/columbia/spring2026/pols4722. You will then be prompted to enter your Columbia email address to confirm your registration.

Useful textbooks

Lectures will draw from textbooks and academic papers, with references provided in the course outline below. The following textbooks will be used frequently, and are available as e-books through Columbia libraries unless stated otherwise:

AP Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

GG Gerber, Alan S., and Donald P. Green. 2012. *Field Experiments*. W.W. Norton. [Not available as e-book.]

IR Imbens, Guido W., and Donald B. Rubin. 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press.

MW Morgan, Stephen L., and Christopher Winship. 2014. *Counterfactuals and Causal Inference: Methods and Principles for Social Research, 2nd Edition*. Cambridge University Press.

In general, AP and IR provide the more technical we cover in class than GG and MW.

The following textbooks provide comprehensive treatments of econometric techniques from a less design-based perspective:

H Hansen, Bruce. 2022. *Econometrics*. Princeton University Press. [Not available as e-book, but online version from 2021 available here.]

W Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. MIT Press. [Not available as e-book.]

If you prefer hard copies, books can be ordered from Columbia's official bookstore here.

Academic integrity

All students are expected to submit work that is their own and to comply with the University's standards for academic integrity (https://www.college.columbia.edu/academics/academicintegrity). If you have any questions, please contact the instructor or the Office of Academic Affairs.

Students are welcome to use artificial intelligence (AI) tools to support their understanding and for problems sets (especially to aid coding), but are encouraged to limit use to maximize their learning and prepare for in-person exams where AI tools are prohibited. Any use of AI in written assignments must be appropriately cited. Each student is responsible for assessing the validity and applicability of any generative AI output that is submitted, bearing final responsibility for the work submitted, including any inaccuracies or plagiarism. Violations of this policy will be considered academic misconduct. Please note that different courses at Columbia may have different AI policies, and it is the student's responsibility to conform to the expectations for each course.

Students requiring academic accommodations

Students with documented disabilities or other conditions who require additional accommodations should contact the instructor as soon as possible to make arrangements. Please find the faculty statement on disability accommodation at http://www.college.columbia.edu/rightsandresponsibilities.

Course outline

The current course plan is listed below, but may change depending on how students are keeping up with the course materials. There are suggested readings for each lecture; some readings—especially textbooks—cover similar content in different ways.

1 – Potential outcomes and randomized experiments

Lecture 1, 1/20 – Introduction and potential outcomes

- GG chapters 2.1, 2.2, and 2.7.
- IR chapter 1.
- MW chapters 2.1-2.5 and 2.8-2.11.
- Torreblanca, Carolina, William Dinneen, Grossman, Guy, and Yiqing Xu. 2025. "The Credibility Revolution in Political Science." OSF working paper.

Lecture 2, 1/22 – Randomized experiments with finite sample inference

- AP chapters 2.1-2.2.
- GG chapters 2.3-2.6.
- IR chapters 3-5.
- MW chapters 2.6 and 2.7.

Lecture 3, 1/27 – Randomized experiments with asymptotic inference

- Abadie, Alberto, Susan Athey, Guido W. Imbens, and Jeffrey M. Wooldridge. 2020. "Sampling-based versus design-based uncertainty in regression analysis." *Econometrica* 88(1):265-296.
- GG chapter 3.
- IR chapter 6.
- For a more comprehensive technical treatment of conditional expectations and asymptotic theory:
  - H chapters 2 and 7.
  - W chapters 2-3.

Lecture 4, 1/29 – Randomized experiments with regression

- AP chapters 2.3 and 3.1.

- Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014. "Inference on treatment effects after selection among high-dimensional controls." *Review of Economic Studies* 81(2):608-650.
- GG chapters 4.1, 4.2, and 4.4.
- IR chapter 7 (also 9 and 10 for blocked designs).
- Lin, Winston. 2013. "Agnostic notes on regression adjustments to experimental data: Reexamining Freedman's critique." *Annals of Applied Statistics* 7(1):295-318.
- For a more comprehensive technical treatment of OLS estimation:
  - H chapters 3-6, 9-10.
  - W chapter 4.

2 – Observational studies

Lecture 5, 2/3 – Conditional ignorability and DAGs

- Cinelli, Carlos, Andrew Forney, and Judea Pearl. 2022. "A Crash Course in Good and Bad Controls." *Sociological Methods & Research* 53(3):1071-1104.
- Deaton, Angus, and Nancy Cartwright. 2018. "Understanding and Misunderstanding Randomized Controlled Trials." *Social Science and Medicine* 210:2-21.
- GG chapter 4.3.
- Hartman, Erin, and F. Daniel Hidalgo. 2018. "An equivalence approach to balance and placebo tests." *American Journal of Political Science* 62(4):1000-1013.
- IR chapters 12 and 14.
- MW chapters 3 and 4.
- For a more comprehensive technical treatment of DAGs:
  - Imbens, Guido W. 2020. "Potential Outcome and Directed Acyclic Graph Approaches to Causality: Relevance for Empirical Practice in Economics." *Journal of Economic Literature* 58(4):1129-1179.
  - Pearl, Judea. 2009. *Causality: Models, Reasoning, and Inference 2nd Edition*. Cambridge University Press. Chapter 3.

Lecture 6, 2/5 – Regression adjustment

- AP chapter 3.

- Aronow, Peter M., and Cyrus Samii. 2016. "Does regression produce representative estimates of causal effects?" *American Journal of Political Science* 60(1):250-267.
- Cinelli, Carlos, and Chad Hazlett. 2020. "Making sense of sensitivity: Extending omitted variable bias." *Journal of the Royal Statistical Society Series B: Statistical Methodology* 82(1):39-67.
- GG chapter 4.5.
- IR chapter 22.
- MW chapter 12.

Lecture 7, 2/10 – Matching

- AP chapter 3.
- IR chapters 13 and 15-18.
- MW chapter 5.
- Stuart, Elizabeth A. 2010. "Matching methods for causal inference: A review and a look forward." *Statistical Science* 25(1):1-21.

Lecture 8, 2/12 – Weighting

- Chattopadhyay, Ambarish, Christopher H. Hase, and José R. Zubizarreta. 2020. "Balancing vs modeling approaches to weighting in practice." *Statistics in Medicine* 39(24):3227-3254.
- Hainmueller, Jens. 2012. "Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies." *Political Analysis* 20(1):25-46.
- MW chapter 7.

3 – Instrumental variables

Lecture 9, 2/17 – IV design and estimation

- AP chapters 4.1-4.4.
- GG chapters 5 and 6.
- IR chapters 23-25.
- MW chapter 9.
- For a more comprehensive technical treatment of IV estimation:

- H chapter 12.
- W chapter 4.

Lecture 10, 2/19 – IV sensitivity

- Lal, Apoorva, Mackenzie Lockhart, Yiqing Xu, and Ziwen Zu. 2024. "How much should we trust instrumental variable estimates in political science? Practical advice based on 67 replicated studies." *Political Analysis* 32(4):521-540.
- Marbach, Moritz, and Dominik Hangartner. 2020. "Profiling compliers and noncompliers for instrumental-variable analysis." *Political Analysis* 28(3):435-444.

Lecture 11, 2/24 – IV extensions

- Angrist, Joshua D., and Guido W. Imbens. 1995. "Two-stage least squares estimation of average causal effects in models with variable treatment intensity." *Journal of the American Statistical Association* 90(430):431-442.
- AP chapter 4.5.
- Chyn, Eric, Brigham Frandsen, and Emily Leslie. 2025. "Examiner and Judge Designs in Economics: A Practitioner's Guide." *Journal of Economic Literature* 63(2):401-439.

Lecture 12, 2/26 – Shift-share and formula instruments

- Borusyak, Kirill, Peter Hull, and Xavier Jaravel. 2025. "A practical guide to shift-share instruments." *Journal of Economic Perspectives* 39(1):181-204.
- Borusyak, Kirill, Peter Hull, and Xavier Jaravel. 2025. "Design-based identification with formula instruments: A review." *Econometrics Journal* 28(1):83-108.

4 – Regression discontinuity

Lecture 13, 3/3 – RD design and estimation

- AP chapter 6.1.
- Cattaneo, Matias D., Nicolás Idrobo, and Rocío Titiunik. 2020. *A Practical Introduction to Regression Discontinuity Designs: Foundations.* Cambridge University Press. Chapters 2-4.
- H chapter 21.

Lecture 14, 3/5 – RD extensions

- AP chapter 6.2.
- Cattaneo, Matias D., Nicolás Idrobo, and Rocío Titiunik. 2020. *A Practical Introduction to Regression Discontinuity Designs: Extensions.* Cambridge University Press. Chapters 2-4.

- Cattaneo, Matias D., Nicolás Idrobo, and Rocío Titiunik. 2020. *A Practical Introduction to Regression Discontinuity Designs: Foundations*. Cambridge University Press. Chapter 5.
- Stommes, Drew, P.M. Aronow, and Fredrik Sävje. 2023. "On the reliability of published findings using the regression discontinuity design in political science." *Research & Politics* 10(2):1-12.

Lecture 15, 3/10 – Multiple dimensional RD

- Cattaneo, Matias D., Nicolás Idrobo, and Rocío Titiunik. 2020. *A Practical Introduction to Regression Discontinuity Designs: Extensions*. Cambridge University Press. Chapter 5.
- Eggers, Andrew C., Ronny Freier, Veronica Grembi, and Tommaso Nannicini. 2018. "Regression discontinuity designs based on population thresholds: Pitfalls and solutions." *American Journal of Political Science* 62(1):210-229.
- Luke J. Keele and Rocío Titiunik. 2015. "Geographic Boundaries as Regression Discontinuities." *Political Analysis* 23(1):127-155.

Midterm exam

Lecture 16, 3/12 – In-class exam

5 – Difference-in-differences

Lecture 17, 3/24 – DD design and estimation

- AP chapter 5.2.
- de Chaisemartin, Clément, and Xavier D'Haultfœuille. Forthcoming. *Credible Answers to Hard Questions: Differences-in-Differences for Natural Experiments*. Princeton University. Chapters 3.4 and 7. Pre-print on SSRN here.
- Egami, Naoki, and Soichiro Yamauchi. 2023. "Using multiple pretreatment periods to improve difference-in-differences and staggered adoption designs." *Political Analysis* 31(2):195-212.
- H chapter 18.

Lecture 18, 3/26 – DD extensions

- Callaway, Brantly, Andrew Goodman-Bacon, and Pedro H.C. Sant'Anna. 2025. "Difference-in-Differences with a Continuous Treatment." NBER working paper.
- de Chaisemartin, Clément, and Xavier D'Haultfœuille. Forthcoming. *Credible Answers to Hard Questions: Differences-in-Differences for Natural Experiments*. Princeton University. Chapters 2-3.3 and 4.1. Pre-print on SSRN here.

- Roth, Jonathan, Ashesh Rambachan. 2023. "A More Credible Approach to Parallel Trends." *Review of Economic Studies* 90(5):2555-2591.

Lecture 19, 3/31 – Staggered DD

- Chiu, Albert, Xingchen Lan, Ziyi Liu, and Yiqing Xu. Forthcoming. "Causal panel analysis under parallel trends: lessons from a large reanalysis study." *American Political Science Review.*

- de Chaisemartin, Clément, and Xavier D'Haultfœuille. Forthcoming. *Credible Answers to Hard Questions: Differences-in-Differences for Natural Experiments.* Princeton University. Chapters 5-6. Pre-print on SSRN here.

- Roth, Jonathan, Pedro Sant'Anna, Alyssa Bilinski, and John Poe. 2023. "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature." *Journal of Econometrics* 235(2):2218-2244.

- Wing, Coady, Seth M. Freedman, and Alex Hollingsworth. 2024. "Stacked difference-in-differences." NBER working paper.

Lecture 20, 4/2 – Matching DD

- Abadie, Alberto. 2021. "Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects." *Journal of Economic Literature* 59(2):391-425.

- Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2015. "Comparative Politics and the Synthetic Control Method." *American Journal of Political Science* 59(2):495-510.

- Arkhangelsky, Dmitry, and Guido W. Imbens. 2024. "Causal models for longitudinal and panel data: A survey." *Econometrics Journal* 27(3):1-61.

- de Chaisemartin, Clément, and Xavier D'Haultfœuille. Forthcoming. *Credible Answers to Hard Questions: Differences-in-Differences for Natural Experiments.* Princeton University. Chapter 4.2. Pre-print on SSRN here.

- Ham, Dae Woong, and Luke Miratrix. 2024. "Benefits and costs of matching prior to a difference in differences analysis when parallel trends does not hold." *Annals of Applied Statistics* 18(3):2096-2122.

- Imai, Kosuke, In Song Kim, and Erik H. Wang. 2023. "Matching methods for causal inference with time-series cross-sectional data." *American Journal of Political Science* 67(3):587-605.

6 – Causal mechanisms

Lecture 21, 4/7 – Mediation

- Acharya, Avidit, Matthew Blackwell, and Maya Sen. 2016. "Explaining causal findings without bias: Detecting and assessing direct effects." *American Political Science Review* 110(3):512-529.
- Blackwell, Matthew, Ruofan Ma, and Aleksei Opacic. 2024. "Assumption smuggling in intermediate outcome tests of causal mechanisms." arXiv working paper.
- GG chapter 10.
- Imai, Kosuke, Luke Keele, and Teppei Yamamoto. 2010. "Identification, Inference and Sensitivity Analysis for Causal Mediation Effects." *Statistical Science* 25(1):51-71.

Lecture 22, 4/9 – Moderation

- Athey, Susan, Julie Tibshirani, and Stefan Wager. 2019. "Generalized random forests." *Annals of Statistics* 47(2):1148-1178.
- Fu, Jiawei, and Tara Slough. 2025. "Heterogeneous Treatment Effects and Causal Mechanisms." arXiv working paper.
- GG chapter 9.
- Hainmueller, Jens, Jonathan Mummolo, and Yiqing Xu. 2019. "How much should we trust estimates from multiplicative interaction models? Simple tools to improve empirical practice." *Political Analysis* 27(2):163-192.

7 – More complex inference

Lecture 23, 4/14 – Multiple comparisons

- Anderson, Michael L. 2008. "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association* 103(484):1481-1495..
- Benjamini, Yoav, Abba M. Krieger, and Daniel Yekutieli. 2006. "Adaptive linear step-up procedures that control the false discovery rate." *Biometrika* 93(3):491-507.
- Viviano, Davide, Kaspar Wüthrich and Paul Niehaus. 2025. "A model of multiple hypothesis testing." arXiv working paper.

Lecture 24, 4/16 – Power calculations, clustered designs, and inference

- Abadie, Alberto, Susan Athey, Guido W. Imbens, and Jeffrey M. Wooldridge. 2023. "When should you adjust standard errors for clustering?" *Quarterly Journal of Economics* 138(1):1-35.

- AP chapter 8.
- Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. "Bootstrap-based improvements for inference with clustered errors." *Review of Economics and Statistics* 90(3):414-427.
- IR chapter 9.

Lecture 25, 4/21 – Interference and estimating spillovers

- Aronow, Peter M., and Cyrus Samii. 2017. "Estimating average causal effects under general interference, with application to a social network experiment." *Annals of Applied Statistics* 11(4):1912-1947.
- Baird, Sarah, J. Aislinn Bohren, Craig McIntosh, and Berk Özler. 2018. "Optimal design of experiments in the presence of interference." *Review of Economics and Statistics* 100(5):844-860.
- GG chapter 8.
- Viviano, Davide. 2025. "Policy targeting under network interference." *Review of Economic Studies* 92(2):1257-1292.

Lecture 26, 4/23 – Missing data and imputation

- GG chapter 7.
- Lee, David S. 2009. "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects." *Review of Economic Studies* 76(3):1071-1102.
- Zubizarreta, José R. 2015. "Stable weights that balance covariates for estimation with incomplete outcome data." *Journal of the American Statistical Association* 110(511):910-922.

Lecture 27, 4/28 – External validity

- Egami, Naoki, and Erin Hartman. 2023. "Elements of External Validity: Design, Framework, and Analysis." *American Political Science Review* 117(3):1070-1088.
- GG chapter 11.
- Humphreys, Macartan, and Alan M. Jacobs. 2023. *Integrating Inferences: Causal Models for Qualitative and Mixed-Method Research.* Cambridge University Press.
- Slough, Tara, and Scott A. Tyson. 2024. *External Validity and Evidence Accumulation.* Cambridge University Press.

8 – Review

Lecture 28, 4/30 – Review for final exam