Winter 2025 | Lecture: Tues./Thurs., 2pm - 3:20pm, Lab: Fri., 2:30pm - 3:20pm | Room: Pick Hall 506 (Lecture), Cobb Hall 101 (Lab) | Units: 100

PLSC 30600: Causal Inference

Anton Strezhnev

Office: Pick Hall 328, 3rd floor
Office Hours: Tuesdays 4pm-6pm or schedule an appointment by e-mail
astrezhnev@uchicago.edu
http://www.antonstrezhnev.com

Teaching Assistant: Cindy Wang
Office Hours: TBA
wangxinyu@uchicago.edu

Last Updated: **01/06/2025**

Course Overview

Questions of cause and effect are central to the study of political science and to the social sciences more broadly. But making inferences about causation from empirical data is a significant challenge. Critically, there is no simple, assumption-free process for learning about a causal relationship from the data alone. Causal inference requires researchers to make assumptions about the underlying data generating process in order to identify and estimate causal effects. The goal of this course is to provide students with a structured statistical framework for articulating the assumptions behind causal research designs and estimating effects using quantitative data.

The course begins by introducing the counterfactual framework of causal inference as a way of defining causal quantities of interest such as the "average treatment effect." It then proceeds to illustrate a variety of different designs for identifying and estimating these quantities. We will start with the most basic experimental designs and progress to more complex experimental and observational methods. For each approach, we will discuss the necessary assumptions that a researcher needs to make about the process that generated the data, how to assess whether these assumptions are reasonable, how to interpret the quantity being estimated and ultimately how to conduct the analysis.

This course will involve a combination of lectures, sections and problem sets. Lectures will focus on introducing the core theoretical concepts being taught in this course. Sections will emphasize application and demonstrate how to implement various causal inference techniques with real data sets. Problem sets will contain a mixture of both theoretical and applied questions and serve to reinforce key concepts and allow students to assess their progress and understanding throughout the course.

Assignments will involve analysis of data using the R programming language. This is a free and open source language for statistical computing that is used extensively for data analysis in many fields. Prior experience with the fundamentals of R programming is required.

Prerequisites

This course is the second in the political science graduate methodology sequence. Completing the introductory course prior to this sequence should prepare you for the material in this class. We will rely on some background knowledge of core concepts in probability, statistics and inference as well as experience with statistical programming in R. However, there are no strict, specific course pre-requisites as many different disciplines and departments offer introductory statistics classes that cover the relevant material. In general, you should have had some introduction to probability theory and should be familiar with concepts like the properties of random variables (especially expectation and variance), estimands and estimators, and statistical inference. Familiarity with regression modeling is a plus but not stricty required. Please contact the instructor at (astrezhnev@uchicago.edu) if you are interested in enrolling but are unsure of the requirements.

Logistics

Lectures: Tuesdays/Thursdays from 2pm - 3:20pm – Location: Pick Hall, Room 506
Sections: Friday from 2:30pm - 3:20pm – Location: Cobb Hall, Room 101

You should attend sections regularly as they comprise a significant element of the course instruction.

Disucssion Forum: We will use a private Ed discussion forum as a course discussion platform. You should already be enrolled and able to access the forum See the Canvas page for more details.

Course Materials: Lecture materials, problem sets and section code will be posted on the course GitHub page at https://github.com/UChicago-pol-methods/plsc-30600-causal-inference/. Problem set solutions will be posted after the due date on Canvas.

Readings will be posted on the Canvas website. You can find them under the "Modules" section organized by week.

Textbooks

The course will involve readings from a variety of different textbooks and published papers. The class will not require the purchase of a single, specific, text and all excerpts from textbooks are available online (either directly or through library resources). However, we do recommend considering obtaining some of these texts to use as a personal reference and they may be valuable to you in the future.

In general, I have found the following books useful. You do not need to purchase *all* of them, but it is worth being aware of them as they provide very good overviews from a variety of disciplines – from econometrics to statistics to epidemiology.

- Cunningham, Scott. *Causal inference: The Mixtape.* Yale University Press, 2021.
- Huntington-Klein, Nick. *The Effect: An Introduction to Research Design and Causality.* Chapman and Hall/CRC, 2021.
- Angrist, Joshua D., and Jorn-Steffen Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton University Press. 2009.
- Imbens, Guido W. and Donald B. Rubin. *Causal Inference for Statistics, Social, and Biomedical Sciences.* Cambridge University Press. 2010.
- Hernán, Miguel A. and James M. Robins. *Causal Inference: What If.* Chapman & Hall/CRC. 2020. (PDF available at: https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/)
- Morgan, Stephen L., and Christopher Winship. *Counterfactuals and Causal Inference.* Cambridge University Press, 2015.

Requirements

Students' final grades are based on three components:

- **Problem sets** (25% of the course grade). Students will complete a total of three problem sets throughout the quarter. Problem sets will primarily cover topics from the lecture and section for that week and the previous week.

The goal of the problem sets is to encourage exploration of the material and to provide you with a clear and credible means of assessing your understanding and progress through the course. As such, problem sets are *designed* to be challenging and we expect students to find some questions difficult.

Problem sets will be graded on a (+/✓/-) scale with a + awarded for complete and near-perfect work, a ✓ awarded for generally good work with clear effort shown but with some errors, and a - awarded for significantly incomplete work with major conceptual errors and little effort shown.

- *Collaboration policy*: We strongly encourage collaboration between students on the problem sets and highly recommend that students discuss problems with each other either in person or via Ed. However, each student is expected to submit their own write-up of the answers and any relevant code.
- *Office hours and online discussion*: Students should feel free to discuss any questions about the problem sets with the teaching staff during sections and office hours. We also strongly encourage students to post questions about both the problem sets and the assigned readings on the course Ed discussion board and respond to other students' questions. Responding to other students' questions will contribute to your participation grade.
- *Submission guidelines*: Problem sets will be distributed as `PDF` and `Rmarkdown` files (`.Rmd`). You should submit your answers and any relevant R code in the same format: including an `Rmarkdown` file (`.Rmd` extension) and a corresponding rendered `.pdf` file as your submission. `Rmarkdown` combines the text formatting syntax of Markdown markup language with the ability to embed and execute chunks of `R` code directly into a text document. This allows you to present your code, graphical output, and discussion/write-up all in the same document. We highly recommend that you edit the distributed `Rmarkdown` assignment file for each problem set directly to make organization easier.

· **Take-home midterm** (30% of the course grade). The take-home midterm will have the same format and structure as the problem sets but with one key difference. You are **not** permitted to collaborate with other students or any other individual on the exams. The teaching staff will answer any clarifying questions on the Ed discussion board, but will not answer substantive questions.

· **In-person final** (35% of the course grade). The final exam will take place in person during exams week (March 11 - March 14). The exact date and time will be announced once final exam room assignments have been made. Exams will take the form of a standard pen-and-paper timed examination involving both theory and practical analysis of sample code and results. More details will be provided towards the latter half of the course.

- **Participation** (10% of the course grade). We expect students to take an active role in learning in both lecture and section. Engagement with the teaching staff by asking and answering questions will contribute to this grade as will interaction on the Ed board.

Computing

This course will use the `R` programming language. This is a free and open source programming language that is available for nearly all computing platforms. You should download and install it from http://www.r-project.org. Unless you have strong preferences for a specific coding environment, we also highly recommend that you use the free RStudio Desktop Integrated Development Environment (IDE) which you can download from https://rstudio.com/products/rstudio/download/#download. In addition to being a great and simple to use environment for editing code, `RStudio` makes it very easy to write and compile `Rmarkdown` documents: the format in which problem sets will be distributed. In addition to base `R`, we will be frequently using data management and processing tools found in the tidyverse set of packages along with basic graphics and visualization using ggplot2.

Policy on Generative Large Language Models

The rapid growth in both the capabilities and the accessibility of generative large language models (LLMs) such as the GPT series, PaLM, LLaMa, etc… has introduced some novel challenges to the classroom. On the one hand, generative text models can be used as a tool to improve the quality of students' writing. On the other hand, they can be readily used to represent another's work as one's own – that is, to commit plagiarism. Additionally, LLMs may appear to be useful for some tasks – such as summarizing a set of texts or finding new sources on a particular topic – when in fact the outputs are arguably sub-optimal relative to conventional research methods.

My view in short: Large language models are marvels of **engineering**. You should use them for **engineering** tasks, but the task of research is not purely engineering and LLMs are much less effective for the task of doing **science**.

By "engineering," I mean the the iterative task of solving a problem by brainstorming potential solutions, implementing those solutions, and then subsequently *evaluating* the solutions with respect to some clearly defined criteria. The key components here are both the existence of a well-defined problem and the ability to assess whether the proposed solutions are effective.

Currently, the most obvious and effective use-case for large language models is in coding. I am perfectly happy for you to experiment with using LLMs in debugging code. The interactivity is great for beginning programmers who may have an idea of what they want their code to do, but are unfamiliar with the syntax of a particular language. Likewise, it's an incredibly valuable tool for experienced programmers who want to quickly generate some prototype code that is customized to their particular problem.

Why is programming an ideal use case? Programming is fundamentally an engineering task. There is a clearly defined problem that a programmer needs to solve via code and there is a straightforward way to evaluate whether a block of code works. As a result, mistakes are easy to catch – if the code throws an error, something needs to be changed. There is always a human in the loop who is capable of evaluating the output.

Outside of coding, I do not think LLM outputs are too useful, especially for generating text that is to be submitted without further refinement. In general, you should be cautious about any LLM outputs that you are not able to verify or evaluate yourself.

Irrespective of whether LLM outputs are "good" or not, it is absolutely clear that presenting LLM-generated output as one's own ideas is clearly plagiarism and will be treated as such. This does not rule out all uses of LLM-generated text, but it does rule out most. One use that I would consider acceptable is cleaning up original text that you have written to eliminate grammar mistakes or to rephrase the text to have a clearer style. We already accept the use of spellcheckers and thesauruses that are embedded in most word processors and I don't see this use case as substantively different as long as your original writing is the input. It is important, however, that you are able to evaluate the output and determine that it is conveying exactly what you want to say in exactly the way that you want to say it, just as you would when using any other writing tool.

Beyond this particular use, **submitting LLM-generated text as a substitute for your own thinking is not permitted in this class and will be considered plagiarism.** This includes prompting an LLM to compose all or part of your writing and submitting that output either verbatim or with some editing. This policy also applies to generating posts on the Ed discussion board.

In general, I do not think that presently there are too many good uses for LLMs for the particular tasks that you will be doing in this class. Although these models can be utilized for things like brainstorming, summarizing text, and search - acting as something of a personalized tutor - and the quality of the model outputs does appear to be steadily growing, I think that you will find significant value to working through the course material directly and asking questions to the teaching staff and to your colleagues in the class.

Schedule

A schedule of topics and readings is provided below. Each week will cover a single topic or group of topics. Tuesday lectures will typically be an introduction to the topic while Thursday lectures will go into greater detail and involve some applications of the method. You should make sure to review the readings prior to that week's lectures with an aim towards completing the reading assignments prior to Thursday's lecture.

Week 1: Introduction to Potential Outcomes (January 7)

- Review of random variables, estimators and inference.
- Counterfactual reasoning and the "Fundamental Problem of Causal Inference"
- The "potential outcomes" model
- Estimands and causal quantities of interest

Readings

- Chapter 1, Imbens, Guido W. and Donald B. Rubin. *Causal Inference for Statistics, Social, and Biomedical Sciences.* Cambridge University Press. 2010.
- Chapter 1, Hernán, Miguel A. and James M. Robins. *Causal Inference: What If.* Chapman & Hall/CRC. 2020.
- Lundberg, Ian, Rebecca Johnson, and Brandon M. Stewart. "What is your estimand? Defining the target quantity connects statistical evidence to theory." *American Sociological Review* 86.3 (2021): 532-565.

Problem Set 1 Assigned January 7, Due January 20

Week 2: Randomized Experiments (January 14)

- What assumptions are needed to identify average treatment effects
- Why randomized experiments satisfy these assumptions
- Estimation and randomization inference in standard experimental designs

Readings

- Sections 1-5, Athey and Imbens, "The Econometrics of Randomized Experiments," *Handbook of economic field experiments.* Vol. 1. North-Holland, 2017. 73-140.
- Chapter 2, Hernán, Miguel A. and James M. Robins. *Causal Inference: What If.* Chapman & Hall/CRC. 2020.
- Druckman, James N., et al. "The growth and development of experimental research in political science." American Political Science Review 100.4 (2006): 627-635.

Applications

- Gerber, Alan S., Donald P. Green, and Christopher W. Larimer. "Social pressure and voter turnout: Evidence from a large-scale field experiment." American political Science review 102.1 (2008): 33-48.
- Mutz, Diana C., and Byron Reeves. "The new videomalaise: Effects of televised incivility on political trust." American Political Science Review 99.1 (2005): 1-15.

Week 3: Experiments Continued (January 21)

- Stratification and using covariates in experiments
- Analysis of cluster-randomized experiments
- Problems of non-compliance

Readings

- Sections 6-12, Athey and Imbens, "The Econometrics of Randomized Experiments," *Handbook of economic field experiments.* Vol. 1. North-Holland, 2017. 73-140.
- Montgomery, Jacob M., Brendan Nyhan, and Michelle Torres. "How conditioning on posttreatment variables can ruin your experiment and what to do about it." American Journal of Political Science 62.3 (2018): 760-775.
- Aronow, P. M., Jonathon Baron, and Lauren Pinson. "A note on dropping experimental subjects who fail a manipulation check." Political Analysis 27.4 (2019): 572-589.
- Lin, Winston. "Agnostic notes on regression adjustments to experimental data: Reexamining Freedman's critique." The Annals of Applied Statistics 7.1 (2013): 295-318.
  - **Bonus**: Samii, C., and P. M. Aronow. "On equivalencies between design-based and regression-based variance estimators for randomized experiments." Statistics & Probability Letters 82.2 (2012): 365-370.

Applications

- Casey, K., Glennerster, R., & Miguel, E. (2012). Reshaping institutions: Evidence on aid impacts using a preanalysis plan. The Quarterly Journal of Economics, 127(4), 1755-1812.
- Crépon, B., Devoto, F., Duflo, E., & Parienté, W. (2015). Estimating the impact of microcredit on those who take it up: Evidence from a randomized experiment in Morocco. American Economic Journal: Applied Economics, 7(1), 123-50.

Problem Set 2 Assigned January 21, Due February 3

Week 4: Selection-on-observables (January 28)

- What to do when random assignment of treatment is not possible – common challenges of observational designs
- Assumptions behind "no unobserved confounding" designs
- Representing assumptions using graphical models
- Covariate adjustment via subclassification

Readings

- Chapter 12. Imbens and Rubin.
- Chapters 6-8. Huntington-Klein, Nick. *The Effect: An introduction to research design and causality*. Chapman and Hall/CRC, 2021.
- Chapter 3 Hernán and Robins
- Chapter 6-8 Hernán and Robins

Applications

- Washington, Ebonya L. "Female socialization: how daughters affect their legislator fathers." American Economic Review 98, no. 1 (2008): 311-32.
- Ba, Bocar A., Dean Knox, Jonathan Mummolo, and Roman Rivera. "The role of officer race and gender in police-civilian interactions in Chicago." Science 371, no. 6530 (2021): 696-702.

Week 5: Selection-on-observables Continued (February 4)

- Propensity scores and covariate adjustment via weighting
- Matching estimators
- Regression estimators and "doubly-robust" estimators

Readings

- Chapter 13. Imbens and Rubin.
- Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. Review of Economics and statistics, 86(1), 4-29.
- Aronow, Peter M., and Cyrus Samii. "Does regression produce representative estimates of causal effects?." American Journal of Political Science 60.1 (2016): 250-267.
- Glynn, Adam N., and Kevin M. Quinn. "An introduction to the augmented inverse propensity weighted estimator." Political Analysis 18.1 (2010): 36-56.
- Abadie, A., & Imbens, G. W. (2011). Bias-corrected matching estimators for average treatment effects. Journal of Business & Economic Statistics, 29(1), 1-11.

Midterm Exam Assigned February 4, Due February 10

Week 6: Instrumental Variables (February 11)

- Estimating effects under unobserved confounding using exogenous variation in treatment induced by an instrument
- Assumptions behind the instrumental variable strategy – exogeneity, relevance, "exclusion restriction"
- Estimation via the Wald Estimator and Two-Stage Least Squares
- Interpreting the IV estimand – Local Average Treatment Effect
- What makes a good instrument?

Readings

- Cunningham, Causal Inference: The Mixtape, Chapter 7 - Instrumental Variables
- Angrist, Imbens and Rubin (1996) "Identification of causal effects using instrumental variables." Journal of the American Statistical Association, 91:434, 444-455
- Sovey, Allison J., and Donald P. Green. "Instrumental variables estimation in political science: A readers' guide." American Journal of Political Science 55, no. 1 (2011): 188-200.
- Andrews, Isaiah, James H. Stock, and Liyang Sun. "Weak instruments in instrumental variables regression: Theory and practice." Annual Review of Economics 11 (2019): 727-753.

Applications

- Gerber, Alan S., and Donald P. Green. "The effects of canvassing, telephone calls, and direct mail on voter turnout: A field experiment." American political science review 94.3 (2000): 653-663.
- Dobbie, Will, Jacob Goldin, and Crystal S. Yang. "The effects of pretrial detention on conviction, future crime, and employment: Evidence from randomly assigned judges." American Economic Review 108.2 (2018): 201-40.

Week 7: Differences-in-differences (February 18)

- Weakening "selection on observables" by studying changes over time.
- Assumptions behind the "differences-in-differences" strategy – parallel trends
- Estimation and diagnostics for the identification assumptions.
- Pitfalls and challenges when units initiate treatment at different times.

Readings

- Cunningham, The Causal Inference Mixtape, Chapter 9 - Differences-in-differences
- Roth, J., Sant'Anna, P. H., Bilinski, A., & Poe, J. (2022). What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature. arXiv preprint arXiv:2201.01194.
- Imai, Kosuke, In Song Kim, and Erik H. Wang. "Matching Methods for Causal Inference with Time-Series Cross-Sectional Data." American Journal of Political Science (2021).

Applications

- Malesky, E. J., Nguyen, C. V., & Tran, A. (2014). The impact of recentralization on public services: A difference-in-differences analysis of the abolition of elected councils in Vietnam. American Political Science Review, 108(1), 144-168.

- Miller, S., Johnson, N., & Wherry, L. R. (2021). Medicaid and mortality: new evidence from linked survey and administrative data. The Quarterly Journal of Economics, 136(3), 1783-1829.

Problem Set 3 Assigned February 18, Due March 3

Week 8: Regression Discontinuity Designs (February 25)

- Estimating effects under unobserved confounding using quasi-random assignment at a cutpoint.

- Common applications: Elections, test scores

- Estimation and sensitivity to modeling assumptions.

Readings

- Chapters 1, 2 and 5. Cattaneo, Matias D., Nicolás Idrobo, and Rocío Titiunik. A practical introduction to regression discontinuity designs: Foundations. Cambridge University Press, 2019.

- Eggers, A. C., Freier, R., Grembi, V., & Nannicini, T. (2018). Regression discontinuity designs based on population thresholds: Pitfalls and solutions. American Journal of Political Science, 62(1), 210-229.

- Keele, Luke J., and Rocio Titiunik. "Geographic boundaries as regression discontinuities." Political Analysis 23.1 (2015): 127-155.

Applications

- Hidalgo, F. Daniel, and Simeon Nichter. "Voter buying: Shaping the electorate through clientelism." American Journal of Political Science 60.2 (2016): 436-455.

- Bleemer, Zachary, and Aashish Mehta. "Will studying economics make you rich? A regression discontinuity analysis of the returns to college major." American Economic Journal: Applied Economics 14.2 (2022): 1-22.

Week 9: Mediation and Sensitivity Analysis (March 4)

- How to define and identify indirect and direct effects of treatment
- How to assess the robustness of results to violations of identification assumptions.

Readings

- Blackwell, Matthew. "A selection bias approach to sensitivity analysis for causal effects." Political Analysis 22.2 (2014): 169-182.
- Cinelli, Carlos, and Chad Hazlett. "Making sense of sensitivity: Extending omitted variable bias." Journal of the Royal Statistical Society: Series B (Statistical Methodology) 82.1 (2020): 39-67.
- Imai, K., Keele, L., Tingley, D., & Yamamoto, T. (2011). Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies. American Political Science Review, 105(4), 765-789.
  - **Bonus**: Green, Donald P., Shang E. Ha, and John G. Bullock. "Enough already about "black box" experiments: Studying mediation is more difficult than most scholars suppose." The Annals of the American Academy of Political and Social Science 628.1 (2010): 200-208.
- Acharya, Avidit, Matthew Blackwell, and Maya Sen. "Explaining causal findings without bias: Detecting and assessing direct effects." American Political Science Review 110.3 (2016): 512-529.

Final Exam during Exam Week (March 11-March 14)

Assignment Schedule

- Problem Set 1: Assigned January 7, Due January 20
- Problem Set 2: Assigned January 21, Due February 3
- **Midterm Exam**: Assigned February 4, Due February 10
- Problem Set 3: Assigned February 18, Due March 3
- **Final Exam**: In-Class, TBA - Between March 11 and March 14

Acknowledgments

This course is indebted to the many wonderful and generous scholars who have developed causal inference curricula in political science departments throughout the world and who have made their course materials available to the public. In particular, I thank Matthew Blackwell, Brandon Stewart, Molly Roberts, Kosuke Imai, Teppei Yamamoto, Jens Hainmueller, Adam Glynn, Gary King, Justin Grimmer whose lecture notes and syllabi have been immensely valuable in the creation of this course. I also thank Andy Eggers, Molly Offer-Westort and Bobby Gulotty whose comments and feedback have been essential to the development of this class.

Special thanks to previous TAs of this class: Arthur Yu, Oscar Cuadros and Zikai Li.