PSC 85509:
Applied Quantitative Research II

Instructor: Thomas Leavitt

Official Course Description

This course is the third in the Political Science program's quantitative methods sequence, following PSC 85507 (Quantitative Analysis I: Regression Analysis). The course is intended for students who want to develop the skills to use quantitative methods in their own research and to think systematically about empirical data in political science. The course begins with a review of the foundations of statistical modeling, estimation, and inference, including ordinary least squares (OLS) and maximum likelihood, and then examines how to apply these tools in common empirical settings. Topics include methods for panel, multilevel (hierarchical), event history (survival), and spatial data, as well as the use of these methods in causal inference designs, such as regression discontinuity, difference-in-differences, and instrumental variables.

Students will work with real datasets, replicate and extend published empirical studies, and produce an original empirical research paper applying these methods to a substantive question in political science. The course uses both `R` and `Stata` for statistical analysis. By the end of the semester, students will be able to identify appropriate quantitative methods for different types of data, implement these methods using statistical software, and apply them to conduct and evaluate research in political science.

Course Organization

The course assumes the material of PSC 85507. From that course you should bring descriptive statistics and the normal distribution; bivariate and multiple regression fit by least squares; hypothesis tests, p-values, and confidence intervals; dummy variables, interactions, polynomials, and logarithms as ways of changing a regression's functional form; the omitted-variable-bias formula; heteroskedasticity, autocorrelation, multicollinearity, and influential observations; logit and probit, their coefficients converted to predicted probabilities, odds ratios, and average marginal effects; and `Stata`. This course assumes no prior experience with `R`.

The course has two parts. Part I (Weeks 1–5) builds the framework, and three of its five weeks — Weeks 2, 4, and 5 — are also anchored in a published application. Part II (Weeks 6–13) applies the framework to causal questions and to dependent data, and anchors every week in a published application. Week 14 adds final-paper presentations.

The course gives you a framework that outlasts any particular method. New methods keep arriving: The omitted-variable-bias sensitivity analysis taught in Week 8, the difference-in-differences estimators surveyed in Week 11, and double machine learning (Chernozhukov et al., 2018; Kennedy, 2024), which this course does not cover, all emerged within the last decade. With the fundamentals in hand, you can teach yourself whatever comes next.

Part I (Weeks 1–5): The framework

Week 1: Estimands, estimators, and the plug-in principle

We begin not with a method but a question: *What is the quantity you want to know?* That quantity is the *estimand*; a recipe turning data into a guess of the estimand is an estimator, and the guess the estimator returns on a particular sample is an estimate. An estimand is a feature of the population distribution P from which each observation is drawn, written $\psi(P)$.

We never observe P, only a sample of n observations, from which we form the *empirical distribution* $\hat{P}$ putting probability $1/n$ on each observation. The *plug-in principle* says: To estimate $\psi(P)$, compute the same function on $\hat{P}$. For a population mean, $\psi(\hat{P})$ is the familiar sample mean, and the gap between $\psi(\hat{P})$ and $\psi(P)$ is what Weeks 3 and 4 quantify. Nearly every estimator in this course is a plug-in estimator, and the first question we ask of any procedure is: *Which population quantity is this procedure the plug-in for?*

Week 2: The CEF, the BLP, and OLS as the BLP's plug-in

Week 2 takes the most familiar tool in applied work — OLS, the recipe that picks the coefficients of a linear function of predictors to minimize the average squared residual in your sample — and reframes it around two features of the population distribution, each a $\psi(P)$. The first, the *conditional expectation function* (CEF), is the average of the outcome Y at each value of the predictors X. The second, the *best linear predictor* (BLP), is the linear function of X whose predictions of Y have the smallest average squared error under P. When the CEF is linear, it coincides with the BLP; otherwise the CEF and the BLP are genuinely different quantities.

What does OLS estimate? *Under random sampling, OLS converges to the best linear predictor of Y given X in the population* — whether or not the CEF is linear, whether or not the error (the deviation of Y from the BLP) is normal, and whether or not that error has constant variance. We therefore treat OLS not as a model but as the plug-in estimator of the BLP. The BLP minimizes average squared error over linear functions of X under P; running that same minimization under $\hat{P}$ replaces those averages with sample averages and yields precisely least squares. Which BLP OLS converges to depends on the predictors you include: Each set of predictors defines its own BLP, so dropping one predictor generally shifts the coefficients on the predictors that remain, and the omitted-variable-bias formula computes that shift exactly, without any model of the population.

Week 3: Consistency — the gap between estimator and estimand, and how fast that gap shrinks

Weeks 3–4 come first because every later method rests on what they develop: the foundations that make statistical inference possible without assuming population models. Those foundations are *asymptotic* — exact as n grows indefinitely, approximate at the n we actually have — so every

later guarantee inherits that approximation. Week 3 establishes *consistency*. The WEAK LAW OF LARGE NUMBERS sends the sample mean $\bar{X}$ to the population mean μ in probability, meaning that for any margin, however small, the probability that $\bar{X}$ falls within that margin of μ tends to one as n grows indefinitely. The CONTINUOUS MAPPING THEOREM carries that convergence through continuous functions. Nearly every estimand in the course is a continuous function of population means, and the plug-in estimator of that estimand is the same function of the corresponding sample means, so each plug-in estimator converges to $\psi(P)$. The *rate of convergence* says how fast: The typical magnitude of the estimator's error — the gap between $\psi(\hat{P})$ and $\psi(P)$ — shrinks in proportion to $1/\sqrt{n}$, so a sample four times as large halves that magnitude.

Week 4: Agnostic inference — the distribution of the estimator's error, with no model of the population

To enable calculations of how unusual an estimate $\psi(\hat{P})$ is given a hypothetical value of $\psi(P)$, we need a probability distribution for the error $\psi(\hat{P}) - \psi(P)$. In the absence of a model of the population, Week 4 draws on the CENTRAL LIMIT THEOREM and associated theory to supply the distribution of the error. The central limit theorem makes that error, once rescaled by $\sqrt{n}$, approximately normal in large samples, and SLUTSKY'S THEOREM licenses the plug-in estimation of that distribution's standard deviation, i.e., the *standard error*. The DELTA METHOD carries that approximation to differentiable functions of population means, and the *bootstrap* reaches the error distribution by a second route, repeatedly drawing samples of size n from $\hat{P}$ itself — that is, resampling the observed data with replacement — rather than appealing to the normal approximation.

The template of Weeks 3–4 recurs for every method in the course: *Choose the estimand first, confirm the recipe tracks it, and get the distribution of the estimator's error from the normal approximation with an estimator of the standard error or from the bootstrap*. The few estimators that depart from the template do so instructively, and Part II flags the departures that matter. Nothing that follows assumes the population is normal or that the fitted equation matches the CEF. Nor do we assume the errors have constant variance, which the classical formula for the standard error requires, so heteroskedasticity-robust ("sandwich") estimators of the standard error are our default, with cluster-robust versions for independently drawn groups of dependent members.

Week 5: Parametric models as plug-ins — what maximum likelihood targets when the model is false

A model is *parametric* when a fixed, finite list of parameters describes the outcome's distribution at each value of the predictors. Logistic regression, probit regression, and their generalized-linear relatives are all parametric, each pairing a distribution for the outcome with a linear function of the predictors. *Maximum likelihood* fits them by taking the parameter values that maximize the average *log-likelihood* of the sample, a measure of how probable the model at those parameter values makes the data we observed. Under conditions we state — most importantly that the population problem has a single best answer — this maximization on our sample converges to the parameter values that solve the population version of the same problem, even if the model defining the likelihood is false. Just as the BLP is the linear function of X with the smallest average squared error in the population, the parameter values that solve the population problem are those with the largest average log-likelihood in the population. The same convergence holds

for the quantities we compute from fitted parameter values, each with a population target of its own. A difference in predicted probabilities, for instance, has a target determined by the two values you contrast and by where you hold the remaining predictors. Week 4's delta method then extends a sandwich standard error to the difference itself, so inference remains valid in large samples even when the model is false.

Part II (Weeks 6–13): Causal questions and dependent data

So far we have described populations and predicted outcomes. *Causal* inference adds a step: Specify a causal estimand — a treatment effect, how outcomes would differ if treatment were different. For example, suppose A equals one for the treated and zero for the untreated — the two *treatment arms* — and Y is the outcome. The population target $\psi(P)$ might then contrast the distribution of outcomes among the treated with the distribution among the untreated, a comparison of two different parts of one population as it actually is. The analogous causal estimand contrasts the distribution of outcomes the whole population would have were everyone treated with the distribution that same population would have were no one treated.

The causal estimand is a feature of the distributions the population would have under each treatment, not of P, the distribution the population actually has. Even if you knew P exactly, the causal estimand would remain undetermined. Identification tells you when $\psi(P)$ coincides with the causal estimand, so that your statistical tools warrant causal conclusions. Of every method in Part II we therefore ask not only *what feature the recipe targets in the distribution P that you actually sample from*, but also *under what assumptions that feature is a causal estimand*.

Weeks 6–8: Identification from observed covariates — causal estimands, conditional ignorability, and sensitivity analysis

Week 6 opens with the potential outcomes framework, which rigorously defines causal effects, and with the identifying assumption we lean on most, *conditional ignorability*. Fix a covariate profile — a single set of values of the predictors — at which both treatments occur, so that P supplies a distribution of outcomes among the treated and a distribution among the untreated (*positivity*). Conditional ignorability implies that at that profile, the distribution of outcomes among the untreated is also the distribution the treated would have had were they untreated, and vice versa. Causal identification follows: The distribution of outcomes among the treated at each profile, averaged over profiles in proportion to their shares of the population, is the distribution of outcomes the whole population would have were everyone treated, and likewise for the untreated. The causal estimand is typically the difference between the means of those two distributions.

Week 7 turns from identification to estimation. Two routes reach that difference by plug-in. One models the outcome, predicting each unit's outcome under each treatment and averaging the predictions. The other models the treatment — the *propensity score*, the probability that a unit with a given covariate profile receives treatment — and weights each unit by the inverse of the estimated probability of the treatment that unit received. The outcome model is correct only when its predictions converge to the CEF, which for OLS means the BLP equals the CEF; the treatment model is correct only when its predicted probabilities converge to the propensity score. Combining the two yields a doubly robust estimator, consistent when either model is correct.

Conditional ignorability itself is untestable. When you estimate the treatment effect by regression, Week 8's sensitivity analysis runs Week 2's omitted-variable-bias formula backward, asking how strong unmeasured confounding would have to be, in its association with the treatment and with the outcome, to overturn your conclusion.

Weeks 9–11: Identification from instruments, thresholds, and timing — exclusion restrictions, continuity at the cutoff, and parallel trends

Weeks 9–11 turn to identifying assumptions that do not rest on conditional ignorability. These assumptions attach instead to an instrument that moves treatment without affecting the outcome directly (*the exclusion restriction*, Week 9), an eligibility threshold that renders units just above and just below a cutoff comparable (*continuity at the cutoff*, Week 10), and the timing of a policy's adoption, with never-adopting units tracing the path the adopters would otherwise have followed (*parallel trends*, Week 11). The two questions of Part II recur concretely for each design's workhorse recipe — two-stage least squares, the local polynomial fit, and the two-way fixed-effects regression. Of each we ask which $\psi(P)$ it converges to, and under what assumptions that $\psi(P)$ is the effect the design advertises. Some of these recipes depart from the template of Weeks 3–4, in what they target and in the distribution of their error — the two-way fixed-effects regression when units adopt treatment at different times (*staggered adoption*, Week 11) and the local polynomial fit (Week 10).

Weeks 12–13: Dependent and incomplete samples — partial pooling, censoring, and spatial dependence

Weeks 12–13 confront the premise the course has so far relied on quietly: that observations reach us as independent, fully observed draws from P. Week 12 covers multilevel models, which pool information across groups (*partial pooling*) — voters within districts, students within schools — whose members are not independent draws, and event-history models, which follow units toward events that may not occur before the observation window closes (*censoring*). Week 13 covers spatial models, which let one unit's outcome depend on the outcomes of neighboring units (*spatial dependence*). Each of these topics departs from the template of Weeks 3–4, and each departure tells us which premise failed and what estimation and inference must do instead.

Course Learning Goals

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

1. Articulate what an estimator targets in a population, and reason carefully about whether the estimator reaches that target under assumptions you are willing to defend.
2. Distinguish the *statistical* question from the *causal* question in any analysis, your own or a published study's, and evaluate the identifying assumptions that let a statistical quantity answer the causal question.
3. Implement the core methods of quantitative political science research in `R` or `Stata` (course code is provided in both): OLS with sandwich and cluster-robust inference, maximum-likelihood and other plug-in estimators, propensity-score weighting, doubly robust estima-

tion, sensitivity analysis for unmeasured confounding, instrumental variables, regression discontinuity, difference-in-differences, multilevel models, event history analysis, and spatial models.

4. Read and critically evaluate published quantitative political science research, including identifying threats to the validity of its conclusions and proposing principled extensions.

5. Replicate and extend a published empirical study, communicating your analysis through reproducible code, well-formatted tables and figures, and clear written interpretation.

6. Most importantly, learn a new method on your own — from a methods paper, a package vignette, or a replication archive — by asking what the method targets, under what assumptions, and how to quantify the error of its estimator.

Course Assignments

Problem sets (4)

Each student will complete four problem sets, each due at the start of class in the week marked in the schedule below. The four sets follow the course's arc: Problem set 1 covers the plug-in framework, OLS, and consistency (Weeks 1–3); problem set 2, agnostic inference and maximum likelihood (Weeks 4–5); problem set 3, causal identification, weighting, and sensitivity analysis (Weeks 6–8); and problem set 4, instrumental variables, regression discontinuity, and difference-in-differences (Weeks 9–11).

Problem sets will combine analytical questions and replication exercises in `R` and/or `Stata`. Problem sets 3 and 4 add an "Application of Methods" component in which student groups select a published political science article that uses that problem set's methods and evaluate the article's use of those methods. Each group should select a different article; groups coordinate their choices in a shared spreadsheet, first come, first served. Articles covered in lecture or listed in the slide references are not eligible for the Application of Methods component.

On any problem set's due date, I may ask you to complete a short check of understanding tied to the problem set just submitted: a brief handwritten answer, no laptop, to one of its questions, or a short oral account of your own submitted code.

Topic-week replication and co-teaching (Weeks 7–13)

Beginning with Week 7 (the first week after the causal-inference foundations of Week 6), each topic week features a **designated student** who will replicate a published application of that week's method and co-teach the application portion of the session with me. The role is substantial: You walk the class through the data, the analysis, the identification strategy, the diagnostics, and the substantive interpretation, and then lead the discussion. Your job is to connect the week's conceptual material to a worked empirical example.

There are seven topic weeks (Weeks 7–13), and each student claims one week, first come, first served, at the start of the term. I will lead the application in the one week that goes unclaimed. Choose a topic aligned with your interests — ideally one connected to your final-paper topic, so that the replication contributes directly to the final paper. Each week's lead application is listed in the readings below, but you may propose an alternative application (subject to my approval),

provided it is published, uses the appropriate method on the appropriate data structure, and has available replication data.

You will submit a written replication report — your full code, in the language of your choice, plus a 3–5 page memo describing the design, analysis, and any extensions or critiques — one day before your presentation. The presentation itself should run roughly 20–25 minutes, leaving the rest of the session for instructor-led content and full-class discussion.

Final research paper

Each student will complete an original empirical research paper applying the course's methods to a substantive question of their choosing. The paper should be a polished, professional document suitable as a writing sample, a draft for journal submission, or a policy report, depending on each student's career objectives.

I strongly encourage students who completed a research paper or replication project in PSC 85507 (or another course) to continue developing that work in PSC 85509, ideally toward a publishable piece. The methods you will learn this term will let you take a paper you wrote with the toolkit of PSC 85507 and deepen the paper's identification strategy, extend the paper to a different data structure, or reframe the paper as a more credible empirical study. The project milestones described below are intended to give you time to develop the paper substantively rather than start from scratch.

Project milestones

- **Research proposal** (due Week 6): a 2–3 page document outlining the research question, motivating literature, proposed methodology, and intended data source(s).
- **Draft of final paper** (due at the end of Week 13): a complete first draft of the paper, shared with your discussant on submission.
- **In-class presentation** (Week 14): a 15-minute talk on the paper, with slides. Each student also serves as the discussant for one classmate's paper, having read the classmate's draft, and delivers 5 minutes of remarks, also with slides. Both slide decks are due by the end of the day before the presentation.
- **Final paper** (due the last day of finals week): the polished final version, incorporating feedback from the draft and presentation.

Grades

Final grades weight the course assessments as follows.

- **28%**: four problem sets (**7%** each).
- **20%**: the topic-week replication and presentation.
- **5%**: the research paper proposal.
- **5%**: the draft of the final paper.

- **30%**: the final paper.
- **12%**: the in-class final-paper presentation, attendance, and participation (including service as a discussant for classmates' presentations).

Course Readings

This course is ZTC ("Zero Textbook Cost"). I will post all required readings on the course website, and the reading list draws primarily from the following books, all freely available online or through the library. I will also upload the dataset for each week's application to the course website, so no assignment requires you to track down data.

1. Hansen, Bruce E. (2022). *Probability and Statistics for Economists*. Princeton, NJ: Princeton University Press. (Cited below as "Hansen Vol I.")
2. Hansen, Bruce E. (2022). *Econometrics*. Princeton, NJ: Princeton University Press. (Cited below as "Hansen Vol II.")
3. Aronow, Peter M., and Benjamin T. Miller. (2019). *Foundations of Agnostic Statistics*. New York, NY: Cambridge University Press.
4. Gelman, Andrew, Jennifer Hill, and Aki Vehtari. (2021). *Regression and Other Stories*. New York, NY: Cambridge University Press.
5. Hernán, Miguel A., and James M. Robins. (2020). *Causal Inference: What If*. Boca Raton, FL: Chapman & Hall/CRC. Available online at https://miguelhernan.org/whatifbook. (Page references follow the online edition dated November 21, 2025.)
6. Cattaneo, Matias D., Nicolás Idrobo, and Rocío Titiunik. (2020). *A Practical Introduction to Regression Discontinuity Designs: Foundations*. New York, NY: Cambridge University Press. (*Elements in Quantitative and Computational Methods for the Social Sciences*.)
7. Gelman, Andrew, and Jennifer Hill. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York, NY: Cambridge University Press.
8. Roback, Paul, and Julie Legler. (2021). *Beyond Multiple Linear Regression: Applied Generalized Linear Models and Multilevel Models in R*. Boca Raton, FL: Chapman & Hall/CRC. Available online at https://bookdown.org/roback/bookdown-BeyondMLR/.
9. Box-Steffensmeier, Janet M., and Bradford S. Jones. (2004). *Event History Modeling: A Guide for Social Scientists*. New York, NY: Cambridge University Press.
10. Ward, Michael D., and Kristian Skrede Gleditsch. (2007). *An Introduction to Spatial Regression Models in the Social Sciences*. Unpublished manuscript, June 15, 2007. (Week 13's page references follow this manuscript.)

In addition to texts focused on statistical and causal inference methods, virtually every week's reading list includes one or more articles applying those methods to substantive empirical questions. We will draw applications from American politics, comparative politics, and international relations.

Course Standards

Electronics

To the extent possible, **please bring a personal laptop to each class** for class-related purposes, mostly in-class exercises using statistical software. However, please do not use your cell phones or other electronic devices during class. If a cell phone or another device is important for your participation in the course, please let me know; I will adjust the policy for the entire class rather than single anyone out.

Attendance, Participation, and Collaboration

Attendance and active participation are extremely important to learning from this course. I strongly encourage you to collaborate with other students on problem sets, and the "Application of Methods" components are explicitly group-based. However, all written submissions should be in your own words.

If you need to miss class, please do **not** contact me ahead of time. A missed class affects the participation component of your grade the same way whether or not you contact me beforehand. Please contact me only if you have to miss a class in which you are scheduled to present.

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. If you turn in the final paper late without my prior consent, I will deduct one-third of a letter grade per day. Because I typically review problem sets in class, I will not accept late problem sets for credit.

Use of Artificial Intelligence (AI) Tools

Because this course is fundamentally about quantitative reasoning and statistical computing, I see little sense in forbidding the very tools that are reshaping how such work is done. You are welcome — encouraged, even — to use large language models (ChatGPT, Claude, Copilot, and the like) to debug your `R` or `Stata` code, decipher cryptic error messages, brainstorm questions worth asking of the data, and draft the routine syntax you would otherwise type out by rote.

These tools have become genuinely capable: A single well-posed prompt can already take you a surprising distance. Part of what I hope you take from this course, though, is a feel for how much further you can go once you pair prompting with real code, sensible data structures, and a little discipline in your workflow. Time and again we will see that a scripted pipeline — one that can sweep through thousands of records, merge them with other sources, and keep a versioned log of every step — buys you a degree of scale, auditability, and reproducibility that ad hoc prompting cannot.

That latitude comes with one non-negotiable condition: You own everything you submit.

- **The output is yours.** "The model hallucinated" and "that is just what the chatbot gave me" are never valid defenses. If your code quietly returns the wrong answer, or your paper cites a study that was never written, the error is yours and no one else's. Treat these tools as you

would an eager assistant who is frequently confident and occasionally completely wrong: Verify everything they produce before your name goes on it.

- **You must be able to explain it.** You may use AI to help write code, but you are answerable for every line you hand in. If I ask you in class or in office hours to walk me through a particular function or loop in your work and you cannot say how the function or loop behaves or why it is there, that portion of the assignment may receive no credit. More important still, I want you to ask at every step what could be wrong with a given result — what sanity check, cross-tabulation, or quick plot would tell you whether you are truly on the right track.
- **Think before you prompt.** These tools are excellent at carrying out a task you have already framed and poor at supplying the insight that framed it. Do not surrender your judgment to them; use them to move faster on questions that are already your own, not to do your thinking for you.
- **Disclosure.** Any assignment on which you relied on AI tools must include a brief disclosure statement. As journals ask of their authors, it should (i) name the specific tool(s) and the version or model you used (for example, the product name together with its model or version identifier), (ii) state what you used each one for — which tasks, and which parts of the work — and (iii) confirm that you reviewed and verified the output and take full responsibility for it. As in those venues, an AI tool is never a co-author and never answerable for the work; that responsibility is yours alone.

The course's academic-integrity standards apply to AI-assisted work exactly as they do to everything else, and I treat undisclosed use of AI as a breach of them. These tools also routinely misattribute or invent their sources, so catching and correcting such errors before you submit is part of the same responsibility.

Academic Integrity

I will not tolerate violations of academic integrity under any circumstances. Please familiarize yourself with the CUNY Graduate Center Academic Integrity Policy. If you are having difficulty with an assignment, do not attempt to present another author's work as your own; please come talk to me instead.

Software and Computing

This course uses both `R` and `Stata`, and the bilingualism is my job, not yours: **All code I provide — in lectures, problem sets, and replication exercises — will be available in both languages**, so that the course accommodates every student's choice. **There is no bilingual requirement for any student work.** Submit every assignment — problem sets, the replication, the final paper — in whichever language you prefer. PSC 85507 is taught in `Stata`; the goal of this course is for you to leave equally comfortable in `R`.

Why R is worth the investment

If you have not yet used R seriously, the start of this course will involve some adjustment. The investment is genuinely worth it for several reasons.

- **Frontier methods are R-first.** New causal-inference and machine-learning methods typically appear in R years before they reach Stata, if they reach it at all. The did, DoubleML, HonestDiD, and sensemakr packages all reached Stata as later reimplementations of the R originals; causal forests (grf) took years to acquire a Stata counterpart; and at the time of this writing, kernel balancing (kbal) has no Stata implementation at all. Working at the methodological frontier requires R.
- **Free and open source.** R runs on any platform, requires no license, and is fully accessible from anywhere — no institutional affiliation required. That accessibility matters for collaborators, for replication archives, and for your work after graduate school.
- **The standard in political science methodology.** Replication archives for *Political Analysis*, the *American Journal of Political Science*, the *Journal of Politics*, and similar venues now overwhelmingly ship R code. Reading and understanding recent methods papers in political science requires literacy in R.
- **Reproducible research workflows.** R Markdown and Quarto integrate code, output, citations, and prose into a single document. You can write a polished, fully reproducible paper or report entirely within this workflow.
- **Visualization.** The ggplot2 package is more flexible and publication-friendly than Stata's graphics. Political science journals increasingly publish figures produced with ggplot2.
- **Job market.** Data-science, government-analyst, and industry positions overwhelmingly expect R or Python (often both). Stata alone narrows your opportunities outside academic econometrics. Adding R to your toolkit is the single biggest investment in your software skills that you can make in this course.

Installing the software

To download R, please go to https://cloud.r-project.org and follow the installation instructions. After installing R, please download and install RStudio from https://posit.co/download/rstudio-desktop/. CUNY provides access to Stata, which you can download for your personal computer; Stata is also installed on lab computers at the Graduate Center.

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

Course Schedule and Readings

The CUNY Graduate Center's academic calendar is available via the Office of the Registrar here.

Within a week, anything **due** comes first, marked ⋆ and set in **blue**; anything distributed that week follows, marked ○.

Part I (Weeks 1–5): The framework

Week 1: Estimands, estimators, and the plug-in principle

Required readings

Theory

- Aronow and Miller (2019) §§2.1.1–2.1.3 (pp. 45–59), §3.1 (pp. 91–96), §3.3 (pp. 116–120)

Recommended readings

- Hansen Vol I §§6.4–6.11 (pp. 130–137)
- Wickham et al. (2023): available here

Week 2: The CEF, the BLP, and OLS as the BLP's plug-in

- Problem set 1 distributed

Required readings

Theory

- Aronow and Miller (2019) §§2.2.3–2.2.5 (pp. 67–84), §4.1 (pp. 143–151) — read §§4.1.1–4.1.2 (pp. 144–145) before §4.1.3 (p. 147)

Theory (omitted variable bias)

- Hansen Vol II §2.24 (pp. 45–46)

Application

- Mincer (1974), used throughout Hansen Vol II Chs. 2–4

Recommended readings

- Hansen Vol II §§2.3–2.18 (pp. 16–40)

Week 3: Consistency — the gap between estimator and estimand, and how fast that gap shrinks

Required readings

Theory

- Aronow and Miller (2019) §3.2.1 (pp. 96–102, convergence in probability, the continuous mapping theorem, the weak law of large numbers), §3.2.2 (pp. 102–105, bias, consistency, mean squared error), §3.2.3 (pp. 105–108, the plug-in sample variance, the unbiased sample variance)
- Aronow and Miller (2019) §3.3.1 (pp. 120–121, the plug-in regularity conditions)

Recommended readings

- Hansen Vol I §7.3 (p. 149), §7.5 (p. 152), §7.10 (p. 155)

Week 4: Agnostic inference — the distribution of the estimator's error, with no model of the population

⋆ **Problem set 1 due**

Required readings

Theory (asymptotic normality)

- Aronow and Miller (2019) §3.2.4 (pp. 108–111, convergence in distribution, the central limit theorem, Slutsky's theorem), §3.2.5 (pp. 111–114, asymptotic normality)
- Hansen Vol I §8.9 (p. 172, the delta method)

Theory (sandwich)

- Aronow and Miller (2019) §3.2.6 (pp. 114–120), §3.4 (pp. 124–135), §3.5 (pp. 135–141), §4.2 (pp. 151–156)
- Hansen Vol II §§4.15–4.16 (pp. 111–114) and §§4.22–4.23 (pp. 123–128, cluster-robust)

Theory (bootstrap)

- Aronow and Miller (2019) §3.4.3 (the bootstrap, within §3.4 above)
- Hansen Vol II §10.29 (p. 299) and §10.30 (p. 301, wild and cluster bootstrap)

Application

- Gilens and Page (2014)

Recommended readings

- Hansen Vol I §8.11 (p. 173, the asymptotic distribution of a plug-in estimator), §8.2 (p. 165), §8.6 (p. 170)
- Hansen Vol II §4.24 (pp. 128–130, at what level to cluster), §7.3 (p. 164, asymptotic normality of OLS)
- Cameron and Miller (2015)
- Cameron et al. (2008)
- Efron and Tibshirani (1993, Chs. 1–7)
- Gelman et al. (2021, §§4.4–4.5, pp. 57–62, Type S and Type M error, the significance filter, forking paths)

Week 5: Parametric models as plug-ins — what maximum likelihood targets when the model is false

○ Problem set 2 distributed

Required readings

Theory

- Aronow and Miller (2019) §5.1 (pp. 178–185), §5.2 (pp. 185–203), especially §5.2.3 (p. 194, maximum likelihood under misspecification), §5.2.4 (p. 197, the maximum-likelihood plug-in estimator), §5.2.7 (p. 202)
- Gelman et al. (2021, §14.4, pp. 249–252, average predictive comparisons)

Application

- Fearon and Laitin (2003)

Recommended readings

- Hansen Vol II Ch. 22 (pp. 769–777, M-estimation as a general class)

Part II (Weeks 6–13): Causal questions and dependent data

Week 6: Causal estimands, identification, and conditional ignorability

⋆ **Research proposal due**

Required readings

Theory

- Aronow and Miller (2019) §7.1 (pp. 236–256), especially §7.1.2 (p. 238, missing data), §7.1.5 (p. 247, ignorability), §7.1.6 (p. 250, the propensity score), §7.1.7 (p. 252, post-treatment variables)
- Hernán and Robins (2020, Chs. 1–3, pp. 3–46)

Application

- Hainmueller and Hangartner (2013)

Recommended readings

- Imbens and Rubin (2015, Chs. 1–3)
- Gelman et al. (2021, Ch. 18, pp. 339–362; identification at p. 351, the stable unit treatment value assumption (SUTVA) at pp. 352–354)

Week 7: Estimation after identification — outcome models, propensity scores, and doubly robust estimation

⋆ **Problem set 2 due**

○ Problem set 3 distributed

Required readings

Theory (outcome models)

- Aronow and Miller (2019) §7.2.1 (p. 256), §7.2.2 (p. 258), §7.2.3 (p. 261, maximum-likelihood plug-in estimation)

Theory (weighting)

- Aronow and Miller (2019) §7.2.5 (p. 264), §7.2.6 (pp. 264–267), §7.2.7 (pp. 267–270, doubly robust estimators), §7.2.8 (pp. 270–271), §7.3 (pp. 271–276, overlap and positivity)

Application

- Blattman (2009)

Recommended readings

- Gelman et al. (2021, Ch. 20, pp. 399–413, the propensity-score workflow)

Week 8: Sensitivity analysis — how strong unmeasured confounding would have to be

Required readings

Theory

- Cinelli and Hazlett (2020, §3.1, pp. 43–45, omitted variable bias; §4.3, pp. 50–52, the robustness value; §4.4, pp. 52–54, benchmarking)
- Hansen Vol II §2.24 (pp. 45–46), revisited from Week 2

Application

- Hazlett (2020)

Recommended readings

- Gelman et al. (2021, Ch. 20, pp. 383–394, omitted variable bias; the algebra at p. 385)
- Hosman et al. (2010): the earlier omitted-variable-bias approach to sensitivity analysis that Cinelli and Hazlett (2020) §6.1.2 compares itself against

Week 9: Instrumental variables — the effect for compliers

⋆ **Problem set 3 due**

Required readings

Theory

- Angrist et al. (1996)
- Hansen Vol II §§12.1–12.12 (pp. 337–352), §12.34 (p. 388, the local average treatment effect), §12.36 (p. 392), §12.38 (p. 399, weak instruments)

Application

- Acemoglu et al. (2001)

Recommended readings

- Hernán and Robins (2020, Ch. 16, pp. 209–226)
- Sovey and Green (2011)
- Gelman et al. (2021, §21.1, pp. 421–432, the complier average causal effect)

Week 10: Regression discontinuity — the effect at the cutoff

- Problem set 4 distributed

Required readings

Theory

- Cattaneo et al. (2020, Ch. 2, pp. 8–19; Ch. 3, pp. 20–38; Ch. 4, pp. 39–87; Ch. 5, pp. 88–108)

Application

- Lee (2008)

Recommended readings

- Lee and Lemieux (2010)
- Hansen Vol II Ch. 21 (pp. 755–767)
- Caughey and Sekhon (2011)
- Eggers et al. (2015)
- Gelman et al. (2021, §21.3, pp. 432–440, fuzzy regression discontinuity as an IV problem)

Week 11: Difference-in-differences on panel data — the effect for the treated

Required readings

Theory (panel structure)

- Hansen Vol II §§17.1–17.13 (pp. 609–625), §§17.24–17.25 (p. 637)
- Beck and Katz (1995)

Theory (difference-in-differences)

- Hansen Vol II Ch. 18 (pp. 664–676)
- Roth et al. (2023)

Application

- Paglayan (2019)

Recommended readings

- Goodman-Bacon (2021)
- Callaway and Sant'Anna (2021)
- Ward (2020)
- Gelman et al. (2021, §21.4, pp. 442–445, two-group difference-in-differences and parallel trends; no staggered case)

Week 12: Grouped and censored data — multilevel and event history models

⋆ **Problem set 4 due**

Required readings

Theory (multilevel)

- Gelman and Hill (2007, Ch. 12, pp. 251–278)
- Bell et al. (2019)

Theory (event history)

- Hernán and Robins (2020, Ch. 17, pp. 227–240, causal survival analysis)
- Box-Steffensmeier and Jones (2004, Ch. 2, pp. 7–20, censoring and the hazard and survivor functions)

Application

- Barnes and Burchard (2013)

Recommended readings

- Gelman and Hill (2007, Ch. 11, pp. 237–247, fixed versus random effects; Ch. 13, pp. 279–283, varying slopes)
- Hansen Vol I Ch. 15 (pp. 303–312, shrinkage)
- Hansen Vol II §§17.5–17.6 (pp. 613–616, random effects)
- Hansen Vol II Ch. 27 (pp. 859–862, censoring and selection)
- Roback and Legler (2021, Chs. 8–9): a free alternative, available here
- Box-Steffensmeier and Jones (2004, Ch. 4, pp. 47–68, the Cox proportional hazards model and the risk set)
- Steenbergen and Jones (2002)
- Box-Steffensmeier and Zorn (2002)

Week 13: Spatial data — dependence among neighboring outcomes

⋆ **Draft of final paper due at the end of the week**

Required readings

Theory

- Ward and Gleditsch (2007, Ch. 1 §§3–7, pp. 7–28; Ch. 2, pp. 29–54; Ch. 3, pp. 55–64, especially §5, p. 58, spatially lagged y versus spatial errors)

Application

- Franzese and Hays (2008), reanalyzing Swank and Steinmo (2002)

Recommended readings

- Franzese and Hays (2007)
- Pebesma and Bivand (2023): the `sf` and `spdep` packages, available here

Week 14: Final paper presentations

- ⋆ **Final paper due the last day of finals week**

  A 15-minute talk on your final paper, plus 5 minutes of discussant remarks on one classmate's paper; slides for both are due by the end of the day before, so the six presentations run back-to-back without setup time. The full session is devoted to the presentations.

CUNY Graduate Center 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, please access it here.

Student Disability Services

The CUNY Graduate Center's Office of Student Disability Services provides accommodations to students with disabilities to promote equal access to the GC's programs and services. Please request accommodations as soon as possible by contacting Student Disability Services directly. More information is available here.

Writing Support

The Graduate Center's Writing Center offers free writing support for graduate students. More information is available here.

References

Acemoglu, D., S. Johnson, and J. A. Robinson (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review 91*(5), 1369–1401.

Angrist, J. D., G. W. Imbens, and D. B. Rubin (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association 91*(434), 444–455.

Aronow, P. M. and B. T. Miller (2019). *Foundations of Agnostic Statistics*. New York, NY: Cambridge University Press.

Barnes, T. D. and S. M. Burchard (2013). "Engendering" politics: The impact of descriptive representation on women's political engagement in Sub-Saharan Africa. *Comparative Political Studies 46*(7), 767–790.

Beck, N. and J. N. Katz (1995). What to do (and not to do) with time-series cross-section data. *American Political Science Review 89*(3), 634–647.

Bell, A., M. Fairbrother, and K. Jones (2019). Fixed and random effects models: Making an informed choice. *Quality & Quantity 53*(2), 1051–1074.

Blattman, C. (2009). From violence to voting: War and political participation in Uganda. *American Political Science Review 103*(2), 231–247.

Box-Steffensmeier, J. M. and B. S. Jones (2004). *Event History Modeling: A Guide for Social Scientists*. New York, NY: Cambridge University Press.

Box-Steffensmeier, J. M. and C. Zorn (2002). Duration models for repeated events. *Journal of Politics 64*(4), 1069–1094.

Callaway, B. and P. H. C. Sant'Anna (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics 225*(2), 200–230.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics 90*(3), 414–427.

Cameron, A. C. and D. L. Miller (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources 50*(2), 317–372.

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.

Caughey, D. and J. S. Sekhon (2011). Elections and the regression discontinuity design: Lessons from close U.S. House races, 1942–2008. *Political Analysis 19*(4), 385–408.

Chernozhukov, V., D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, W. Newey, and J. Robins (2018). Double/debiased machine learning for treatment and structural parameters. *Econometrics Journal 21*(1), C1–C68.

Cinelli, C. and C. Hazlett (2020). Making sense of sensitivity: Extending omitted variable bias. *Journal of the Royal Statistical Society Series B: Statistical Methodology 82*(1), 39–67.

Efron, B. and R. J. Tibshirani (1993). *An Introduction to the Bootstrap*. Boca Raton, FL: Chapman & Hall/CRC.

Eggers, A. C., A. Fowler, J. Hainmueller, A. B. Hall, and J. M. Snyder (2015). On the validity of the regression discontinuity design for estimating electoral effects: New evidence from over 40,000 close races. *American Journal of Political Science 59*(1), 259–274.

Fearon, J. D. and D. D. Laitin (2003). Ethnicity, insurgency, and civil war. *American Political Science Review 97*(1), 75–90.

Franzese, R. J. and J. C. Hays (2007). Spatial econometric models of cross-sectional interdependence in political science panel and time-series-cross-section data. *Political Analysis 15*(2), 140–164.

Franzese, Jr., R. J. and J. C. Hays (2008). Interdependence in comparative politics: Substance, theory, empirics, substance. *Comparative Political Studies 41*(4–5), 742–780.

Gelman, A. and J. Hill (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York, NY: Cambridge University Press.

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

Gilens, M. and B. I. Page (2014). Testing theories of American politics: Elites, interest groups, and average citizens. *Perspectives on Politics 12*(3), 564–581.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics 225*(2), 254–277.

Hainmueller, J. and D. Hangartner (2013). Who gets a Swiss passport? A natural experiment in immigrant discrimination. *American Political Science Review 107*(1), 159–187.

Hansen, B. E. (2022a). *Econometrics*. Princeton, NJ: Princeton University Press.

Hansen, B. E. (2022b). *Probability and Statistics for Economists*. Princeton, NJ: Princeton University Press.

Hazlett, C. (2020). Angry or weary? How violence impacts attitudes toward peace among Darfurian refugees. *Journal of Conflict Resolution 64*(5), 844–870.

Hernán, M. A. and J. M. Robins (2020). *Causal Inference: What If*. Boca Raton, FL: Chapman & Hall/CRC. Online edition dated November 21, 2025, available at https://miguelhernan.org/whatifbook; page ranges refer to that edition.

Hosman, C. A., B. B. Hansen, and P. W. Holland (2010). The sensitivity of linear regression coefficients' confidence limits to the omission of a confounder. *Annals of Applied Statistics 4*(2), 849–870.

Imbens, G. W. and D. B. Rubin (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. New York, NY: Cambridge University Press.

Kennedy, E. H. (2024). Semiparametric doubly robust targeted double machine learning: A review. In E. B. Laber, B. Chakraborty, E. E. M. Moodie, T. Cai, and M. van der Laan (Eds.), *Handbook of Statistical Methods for Precision Medicine*, Chapman & Hall/CRC Handbooks of Modern Statistical Methods, Chapter 10, pp. 207–236. Boca Raton, FL: Chapman & Hall/CRC.

Lee, D. S. (2008). Randomized experiments from non-random selection in U.S. House elections. *Journal of Econometrics 142*(2), 675–697.

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

Mincer, J. (1974). *Schooling, Experience, and Earnings*. Number 2 in Human Behavior and Social Institutions. New York, NY: National Bureau of Economic Research. Distributed by Columbia University Press.

Paglayan, A. S. (2019). Public-sector unions and the size of government. *American Journal of Political Science 63*(1), 21–36.

Pebesma, E. and R. Bivand (2023). *Spatial Data Science: With Applications in R*. Boca Raton, FL: Chapman & Hall/CRC.

Roback, P. and J. Legler (2021). *Beyond Multiple Linear Regression: Applied Generalized Linear Models and Multilevel Models in R*. Boca Raton, FL: Chapman & Hall/CRC.

Roth, J., P. H. C. Sant'Anna, A. Bilinski, and J. Poe (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics 235*(2), 2218–2244.

Sovey, A. J. and D. P. Green (2011). Instrumental variables estimation in political science: A readers' guide. *American Journal of Political Science 55*(1), 188–200.

Steenbergen, M. R. and B. S. Jones (2002). Modeling multilevel data structures. *American Journal of Political Science 46*(1), 218–237.

Swank, D. and S. Steinmo (2002). The new political economy of taxation in advanced capitalist democracies. *American Journal of Political Science 46*(3), 642–655.

Ward, G. (2020). Happiness and voting: Evidence from four decades of elections in Europe. *American Journal of Political Science 64*(3), 504–518.

Ward, M. D. and K. S. Gleditsch (2007). An introduction to spatial regression models in the social sciences. Manuscript dated June 15, 2007; page ranges refer to this version.

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.