POLSCI 699
Statistical Methods in Political Research II

Winter 2024

Yuki Shiraito Adam Rauh

Department of Political Science
University of Michigan

This course is the first graduate-level course in applied statistical methods for political scientists. We begin by studying randomized experiments and the difference-in-means estimator of a treatment effect based on the potential outcomes framework for statistical causal inference. Students will then learn applications of linear regression for causal and predictive inference, including factorial experiments, linear interaction models, regression discontinuity designs, difference-in-differences, and encouragement designs. The course will conclude with an introduction to maximum likelihood estimation for nonlinear limited dependent variable models.

1 Contact Information

	Yuki Shiraito (Faculty Instructor)	Adam Rauh (GSI)
Office:	4259 ISR-Thompson	
Phone:	734–615–5165	
Email:	shiraito@umich.edu	amrauh@umich.edu

2 Logistics

	Day and Time	**Location**
Lectures:	Monday and Wednesday, 4:00–5:20pm	7603 Haven Hall
Sections:	Wednesday, 5:30–6:20pm	7603 Haven Hall
Yuki's office hours:	Stop by anytime	4259 ISR-Thompson
Adam's office hours:	Monday, noon–2:00pm	6670 Haven Hall

3 Questions and Announcements

In addition to sections and office hours, please use the Piazza Discussion Board when asking questions about lectures, problem sets, and other course materials. This allows all students to benefit from the 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. Engagement in Piazza discussions is part of course requirements (see Section 5). All course announcements will also be made through Piazza. Canvas will still be used for hosting all class materials.

All students are required to sign up for Piazza. Because the POLSCI 699 Piazza site has been synced with the Canvas course roster, you do not need to enroll the course site yourself. Once you create your account in Piazza, the Piazza course page can also be accessed by logging in from `https://piazza.com` or its mobile apps.

4 Prerequisites

The following materials are the prerequisites of this course.

1. Foundations of mathematical statistics covered in POLSCI 599 or an equivalent course (e.g., STATS 412). In particular, students are assumed to be familiar with the topics covered in Chapters 1–5, 6.1–3, 7.5–6, 8.1–5, 8.7–8 of:

   DeGroot, Morris H. and Mark J. Schervish. 2012. *Probability and Statistics*. 4th ed. Boston: Pearson Education.

   For probability theory, an alternative reference is Chapters 1–5, 7, 8.1–2, 9–10 of:

   Blitzstein, Joseph K. and Jessica Hwang. 2019. *Introduction to Probability*. 2nd ed. Chapman and Hall/CRC.

2. Proficiency in **R** at the level of:

   Imai, Kosuke. 2017. *Quantitative Social Science: An Introduction*. Princeton: Princeton University Press.

3. Concurrently enrolling in POLSCI 514 is strongly recommended.

5 Course Requirements

The final grades are based on the following items:

- **Participation** (10%): The level of engagement in lectures, sections, and Piazza discussions.
- **Problem sets** (40%): There will be four problem sets throughout the semester. Each problem set will equally contribute to the final grade and contain both analytical and data analysis questions. The following instructions will apply to all problem sets:
  - *Collaboration policy*: Collaboration on the problem sets is encouraged, because it provides opportunities for learning from each other. However, to facilitate individual learning, you should try all problem set questions before working together with other students, and you are required to submit your own answers for each problem set. Copying any part of someone else's code or answers is considered as academic misconduct.
  - *Submission policy*: Submit your answers individually through Canvas. All answers must be typed via LaTeX and your **R** source code should be included as part of the LaTeX document. Please ensure your code adheres to the Google's **R** Style Guide rules available at `https://google.github.io/styleguide/Rguide.xml`. No late submission will be accepted, unless the instructor's permission is given in advance.
- **Take-home midterm exam** (15%): The take-home open-book midterm exam will be given on **March 8**. You have one week to complete the exam. The exam consists of both analytical and data analysis questions. Except that *no collaboration is allowed*, the submission policy for the problem sets also applies to the exam.
- **Final project** (35%): The final project must be reanalysis of a published study in the field of your interest. After the reanalysis, you can choose to improve the original analysis either methodologically or substantively (or ideally both). This extension is optional, but you are encouraged to do so. During the entire process, the instructor and the GSI are available to help you with both substantive and technical questions.

  Be aware of the following key deadlines. Late submission will be penalized.

- **March 4 (Project identification and data acquisition):** By this date, you should identify your project and acquire the data to be analyzed (e.g., by downloading the data from a replication archive or requesting the original authors to share the data). Upload one paragraph (up to 150 words) description of your project and the data set to Canvas.
- **March 25 (Descriptive analysis):** By this date, you should finish your descriptive analysis. Turn in a brief summary of your descriptive analysis of the data (up to three pages including figures and tables). This memo should not simply list tables of descriptive statistics and graphs of variables. Rather, you should describe the aspect of the data upon which you aim to improve the original analysis. Look into the details of the data, and if there are any discrepancies or irregularities in the data, you should find them at this stage.
- **April 15 (Reanalysis and proposed extensions):** By this date, you should finish your reanalysis of the data (e.g., reproducing the original results) and optionally come up with your proposed extensions of the analysis. Turn in a brief summary of your replication and proposed extensions of the original analysis (up to five pages including tables and figures). Meet with the instructor to get feedback for your memo by **April 19**.
- **May 2 (Final write-up):** You must electronically submit a final write-up to Canvas. *Note that the write-up should not be a paper.* Instead, it should consist of a title, abstract, introduction, and tables and figures with informative captions. In the introduction, you should concisely summarize the motivation, research question, methods, contributions, and empirical findings of the project. The write-up must be no longer than 10 pages. The project will be graded based on its overall quality. Specifically, we will look at the effectiveness of presentation and writing, as well as the novelty of substantive and methodological contributions if applicable.

- **Incomplete Policy**: No incompletes will be given.

6 Other Course Policies

- **COVID-19:** The course follows the University of Michigan's COVID-19 Policies & Guidelines at `https://healthresponse.umich.edu/policies-guidance/`.

- **Student Sexual Misconduct Policy:** Title IX prohibits sex discrimination to include sexual misconduct: harassment, domestic and dating violence, sexual assault, and stalking. If you or someone you know has been harassed or assaulted, you can receive confidential support and academic advocacy at the Sexual Assault Prevention and Awareness Center (SAPAC). SAPAC can be contacted on their 24-hour crisis line, 734–936–3333 and online at sapac.umich.edu. Alleged violations can be reported non-confidentially to the Office for Institutional Equity (OIE) at institutional.equity@umich.edu. Reports to law enforcement can be made to University of Michigan Police Department at 734–763–3434.[1]

- **Accommodations for Students with Disabilities:** If you think you need an accommodation for a disability, please let me know at your earliest convenience. Some aspects of this course, the assignments, the in-class activities, and the way the course is usually taught may be modified to facilitate your participation and progress. As soon as you make me aware of your needs, we can work with the Services for Students with Disabilities (SSD)

[1] This statement is taken from: `https://sapac.umich.edu/article/faculty-resources-sample-syllabus-language`.

office to help us determine appropriate academic accommodations. SSD (734–763–3000; http://ssd.umich.edu) typically recommends accommodations through a Verified Individualized Services and Accommodations (VISA) form. Any information you provide is private and confidential and will be treated as such.[2]

- **Religious-Academic Conflicts:** While the university does not observe religious holidays, it is the policy of the University of Michigan to make every reasonable effort to allow members of the university community to observe their religious holidays without academic penalty. Absensce from classes or examinations for religious reasons does not relieve students from responsibility for any part of the course work required during the period ob absence. Students who expect to miss classes as a consequence of their religious observance shall be provided with a reasonable alternative opportunity to make-up missed academic work. It is the obligation of students to provide faculty with reasonable notice of the dates on which they will be absent. When the absence coincides with an exam or other assignment due date, the options to make up that missed work may be limited and will be determined by the instructor within the boundaries of the respective class.[3]

- **Academic Misconduct:** The University of Michigan community functions best when its members treat one another with honesty, fairness, respect, and trust. The college promotes the assumption of personal responsibility and integrity, and prohibits all forms of academic dishonesty and misconduct. All cases of academic misconduct will be referred to the Office of the Assistant Dean for Undergraduate Education. Being found responsible for academic misconduct will usually result in a grade sanction, in addition to any sanction from the college. For more information, including examples of behaviors that are considered academic misconduct and potential sanctions, please see https://lsa.umich.edu/lsa/academics/academic-integrity.html.[4]

- **Student Mental Health and Wellbeing:** The University of Michigan is committed to advancing the mental health and wellbeing of its students. If you or someone you know is feeling overwhelmed, depressed, and/or in need of support, services are available. For help, contact *Counseling and Psychological Services (CAPS)* at (734) 764-8312 and https://caps.umich.edu/ during and after hours, on weekends and holidays, or through its counselors physically located in schools on both North and Central Campus. You may also consult *University Health Service (UHS)* at (734) 764-8320 and https://www.uhs.umich.edu/mentalhealthsvcs, or for alcohol or drug concerns, see https://www.uhs.umich.edu/aodresources. For a listing of other mental health resources available on and off campus, visit: http://umich.edu/~health.[5]

- **Use of Generative AI:** In principle you may submit AI-generated code, or code that is based on or derived from AI-generated code, as long as this use is properly documented in the comments: you need to include the prompt and the significant parts of the response. AI tools may help you avoid syntax errors, but there is no guarantee that the generated code is correct. It is your responsibility to identify errors in program logic through comprehensive, documented testing. Moreover, generated code, even if syntactically correct, may have significant scope for improvement, in particular regarding separation of concerns and avoiding

[2]This statement is taken from: https://ssd.umich.edu/article/syllabus-statement.
[3]This statement is taken from: *Handbook for Faculty and Instructional Staff 2018*, p. 17.
[4]This statement is taken from: *Handbook for Faculty and Instructional Staff 2018*, p. 16.
[5]This statement is taken from: *Handbook for Faculty and Instructional Staff 2018*, p. 16.

repetitions. The submission itself should meet our standards of attribution and validation.[6]

7 Textbooks

There is no single required textbook for this course, but lectures are based on relevant chapters and sections of the following books:

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

Cattaneo, Matias D., Nicolás Idrobo and Rocio Titiunik. 2019. *A Practical Introduction to Regression Discontinuity Designs*. Elements in Quantitative and Computational Methods for the Social Sciences New York: Cambridge University Press.

Freedman, David A. 2009. *Statistical Models: Theory and Practice*. 2nd ed. New York: Cambridge University Press.

Hayashi, Fumio. 2000. *Econometrics*. Princeton University Press.

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

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. 2nd ed. The MIT Press.

Also, the following books are useful for supplementary readings on some of the topics covered in the class:

Hansen, Bruce E. 2022. *Econometrics*. Princeton: Princeton University Press.

McCullagh, P. and J.A. Nelder. 1989. *Generalized Linear Models*. 2nd ed. Chapman & Hall/CRC.

8 Course Outline

- **Week 0 (Jan. 10)**
  - *On-demand video lecture, no in-person meeting*
  - Course logistics

Randomized Experiments

- **Week 1 (Jan. 17)**
  - Potential outcomes framework for causal inference
  - Fundamental problem of causal inference
  - Causal estimands
  - Supplementary reading: Imbens and Rubin (2015, Ch. 1)
- **Week 2 (Jan. 22 and 24)**

[6] Boris Steipe (2023) "Syllabus Resources." The Sentient Syllabus Project `http://sentientsyllabus.org`.

  - Causal identification through randomized experiment
  - Estimation and inference for the difference-in-means estimator
  - Supplementary readings: Imbens and Rubin (2015, Chs. 3, 4, and 6); Angrist and Pischke (2009, Ch. 2)

- **Week 3-4 (Jan. 29 and 31, and Feb. 5)**
  - Simple linear regression and ordinary least squares (OLS)
  - Heterogeneous effects
  - Cluster randomized experiments
  - Supplementary readings: Imbens and Rubin (2015, 7.1–4); Angrist and Pischke (2009, 3.1.1–3);

    Samii, Cyrus and Peter M. Aronow. 2012. "On Equivalencies between Design-based and Regression-based Variance Estimators for Randomized Experiments." *Statistics & Probability Letters* 82(2):365–370.

∗ **Problem Set 1 (Jan. 31–Feb. 9)**

- **Week 4-5 (Feb. 7, 12, and 14)**
  - Factorial experiments
  - Combination and marginal effects
  - Multiple hypothesis testing
  - Interaction effects
  - Linear interaction models
  - Supplementary readings: Angrist and Pischke (2009, 3.1.4);

    Hainmueller, Jens, Daniel J. Hopkins and Teppei Yamamoto. 2014. "Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices via Stated Preference Experiments." *Political Analysis* 22(1):1–30.

    Liu, Guoer and Yuki Shiraito. 2023. "Multiple Hypothesis Testing in Conjoint Analysis." *Political Analysis* 31(3):380–395.

  - Advanced reading:

    Benjamini, Yoav and Yosef Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society: Series B (Methodological)* 57(1):289–300.

    Ferreira, J. A. and A. H. Zwinderman. 2006. "On the Benjamini–Hochberg method." *The Annals of Statistics* 34(4):1827–1849.

∗ **Problem Set 2 (Feb. 14–23)**

Linear Regression

- **Week 6 (Feb. 19 and 21)**
  - Linear regression for predictive inference
  - Geometry of linear regression
  - Frisch-Waugh-Lovell theorem
  - Supplementary readings: Angrist and Pischke (2009, 3.1.1–2); Hansen (2022, 3.10–18)
- **No class meetings on Feb. 26 and 28**

∗ **Final Project: One-paragraph project description is due Mar. 4**

- **Week 7-8 (Mar. 4, 6, and 11)**
  - Causal estimands and linear regression with covariates
  - Statistical inference for the OLS estimator
  - Prediction diagnostics
  - Cross validation
  - Supplementary readings: Angrist and Pischke (2009, 3.2–5); Imbens and Rubin (2015, 7.5–11); Hayashi (2000, 1.1–3, 1.6, 2.1, 2.3–5); Hansen (2022, 3.18–21, Chs. 4 and 6–7)

∗ **Midterm Exam (Mar. 8–15)**

- **No lecture on Mar. 13**

Non-experimental Research Designs for Causal Identification

- **Week 9 (Mar. 18 and 20)**
  - Regression discontinuity design (RDD)
  - Identification for RDD
  - Local linear regression for RDD
  - Optimal bandwidth selection
  - RDD with covariates
  - Supplementary readings: Cattaneo, Idrobo and Titiunik (2019);

    Cattaneo, Matias D and Rocio Titiunik. 2022. "Regression Discontinuity Designs." *Annual Review of Economics* 14:821–851.

    de la Cuesta, Brandon and Kosuke Imai. 2016. "Misunderstandings about the Regression Discontinuity Design in the Study of Close Elections." *Annual Review of Political Science* 19:375–396.

    Gelman, Andrew and Guido Imbens. 2019. "Why High-order Polynomials Should Not Be Used in Regression Discontinuity Designs." *Journal of Business & Economic Statistics* 37(3):447–456.
  - Advanced readings:

Hahn, Jinyong, Petra Todd and Wilbert Van der Klaauw. 2001. "Identification and Estimation of Treatment Effects with a Regression-discontinuity Design." *Econometrica* 69(1):201–209.

Imbens, Guido and Karthik Kalyanaraman. 2012. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *The Review of economic studies* 79(3):933–959.

∗ **Final Project: Three-page descriptive analysis is due Mar. 25**

- **Week 10 (Mar. 25 and 27)**
  - Difference-in-differences design
  - Parallel trend assumption
  - Two-way fixed effects regression
  - Supplementary readings: Angrist and Pischke (2009, Ch. 5);

    Imai, Kosuke and In Song Kim. 2021. "On the Use of Two-way Fixed Effects Regression Models for Causal Inference with Panel Data." *Political Analysis* 29(3):405–415.

∗ **Problem Set 3 (Mar. 27–Apr. 5)**

- **Week 11 (Apr. 1 and 3)**
  - Encouragement designs and instrumental variables
  - Identification of the complier average treatment effect
  - Wald estimator and two-stage least squares estimator
  - Fuzzy RDD
  - Supplementary readings: Imbens and Rubin (2015, Chs. 23–4); Angrist and Pischke (2009, Ch. 4);

Introduction to Discrete Choice Models

- **Week 12 (Apr. 8 and 10)**
  - Discrete choice models
  - Maximimum likelihood estimation and inference

∗ **Final Project: Five-page reanalysis is due Apr. 15**

- **Week 13 (Apr. 15 and 17)**
  - Cramer-Rao lower bound
  - Newton-Raphson algorithm
  - Calculating quantities of interest

∗ **Problem Set 4 (Apr. 17–26)**

- **Week 14 (Apr. 22)**
  - Delta method
  - Bootstrap
  - Quasi-Bayesian Monte Carlo approximation

∗ **Final Project: Ten-page final write-up is due May. 2**