**HUDM 5133**
**Causal Inference for Program Evaluation**
**Spring 2025**
Teachers College, Columbia University

---

**Instructor:** Youmi Suk, Ph.D
**Email**: ysuk@tc.columbia.edu
**Office:** GD 552
**Office Hours**: Tue. 3:00-4:00pm, in-person. Online sessions available upon request.

**Co-Instructor:** Bryan Keller, Ph.D
**Email**: keller4@tc.columbia.edu
**Office:** GD 456
**Office Hours**: In person or online sessions available upon request.

**Lecture:** Tue. (In-Person) 12:55-2:55pm, ZB 406

**Course Website:** Canvas

**Course Overview and Learning Outcomes**:
HUDM 5133 will prepare students to start research in causal inference. This course will provide an introduction to the theoretical and practical aspects of causal inference methods, along with real-world applications in education. Topics include: the Neyman-Rubin potential outcomes framework, Pearl's directed acyclic graphical models, single-world intervention graphs, non-equivalent control group designs using both traditional methods (e.g., matching, weighting, regression) and machine learning methods, optimal treatment regimes, regression discontinuity designs, and sensitivity analysis.

At the end of the course, students will be able to:

1. Understand key concepts in causal inference, such as confounding, counterfactuals, and missing data.
2. Identify and estimate causal estimands in both randomization trials and quasi-experimental designs.
3. Evaluate the strengths and weakness of different quasi-experimental designs.
4. Apply causal inference methods to real-world examples and accurately interpret the results.

**Software:**
The primary software for this course is R. While prior experience with R is not required, experience using any statistical software will be helpful. Although you are welcome to use any software for analyses, R is recommended because our examples in lectures and notes will be written in R.

**Software sources:**
Download the free software from the links below.

- The Comprehensive R Archive Network (CRAN) at http://cran.us.rproject.org has links to downloads, manuals, and searchable mailing list archives, all related to R.
- The RStudio IDE for R may be downloaded from posit at https://posit.co/products/open-source/rstudio/.
- If you run R on MacOS, install XQuartz for full graphical functionality. Get it here: https://www.xquartz.org/.

**Supporting Texts:**

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

Morgan, S. L. & Winship, C. (2015). *Counterfactuals and Causal Inference: Methods and Principles for Social Research* (2nd Ed.). New York, NY: Cambridge.

Pearl, J. (2009). *Causality: Models, Reasoning, and Inference* (2nd ed.). New York, NY: Cambridge.

**Other Useful Texts:**

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Belmont, CA: Wadsworth.

VanderWeele, T. J. (2015). *Explanation in Causal Inference*. New York, NY: Oxford.

## Course Schedule

| | |
|---|---|
| W1<br>(Jan 21) | Syllabus, introduction to the course, statistics review<br><br>- Casella & Berger, p. 3 (theorem 1.1.4), p. 11 (theorem 1.2.11.a), 1.3 (conditional probability), 2.2 (expectation), 4.1 (joint/marginal), 4.2 (conditional distributions) |
| W2<br>(Jan 28) | The Neyman-Rubin potential outcomes notation<br><br>- Keller, B. & Branson, Z. (2024). Defining, identifying, and estimating causal effects with the potential outcomes framework: a review for education research. *Asia Pacific Educ. Rev.* 25, 575–594. https://osf.io/preprints/psyarxiv/58qmp<br>- Morgan and Winship, chapters 1-2<br>Optional:<br>- Rubin, D. B. (2007). The design versus the analysis of observational studies for causal effects: Parallels with the design of randomized trials. *Statistics in Medicine*, 26, 20–36. https://doi.org/10.1002/sim.2739<br>- Imbens & Rubin, 1.1-1.8 |
| W3<br>(Feb 4) | Introduction to causal graphs<br><br>Problem set 1 assigned.<br><br>- Pearl, chapter 1<br>- Morgan & Winship, chapters 3-3<br>Optional:<br>- Rohrer, J. M. (2018). Thinking clearly about correlations and causation: Graphical causal models for observational data. *Advances in Methods and Practices in Psychological Science*, 1, 27-42. https://doi.org/10.1177/2515245917745629<br>- Steiner, P. M., Kim, Y., Hall, C. E. & Su, D. (2015). Graphical models for quasiexperimental designs. *Sociological Methods & Research, 46*, 155-188. https://doi.org/10.1177/0049124115582272 |
| W4<br>(Feb 11) | Single-world intervention graphs (SWIGs)<br><br>Problem set 1 due.<br><br>- Richardson, T. S., & Robins, J. M. (2013). Single world intervention graphs: a primer. In *Second UAI workshop on causal structure learning, Bellevue, Washington*.<br>- Bezuidenhout, D., Forthal, S., Rudolph, K., & Lamb, M. R. (2024) Single world intervention graphs (SWIGs): A practical guide, *American Journal of Epidemiology.* https://doi.org/10.1093/aje/kwae353<br>Optional:<br>Ocampo, A. & Bather, J. R. (2023) Single-world intervention graphs for defining, identifying, and communicating estimands in clinical trials, *Statistics in Medicine. 42*(21), 3892-3902. https://doi.org/10.1002/sim.9833 |
| W5<br>(Feb 18) | Design-based identification and estimation<br><br>- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55. https://doi.org/10.1093/biomet/70.1.41<br>- Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science*, 25(1), 1–21. https://doi.org/10.1214/09-STS313 |

| | |
|---|---|
| | Optional:<br>- Imbens & Rubin, chapter 12, 17.3, 18.3<br>- Keller & Branson, 3.2 and 4.2<br>- Keller, B. & Tipton, E. (2016). Propensity score analysis in R: A software review. Journal of Educational and Behavioral Statistics, 41: 326-348. https://doi.org/10.3102/1076998616631744 |
| W6<br>(Feb 25) | Outcome and treatment-based identification and estimation<br><br>Problem set 2 assigned.<br><br>- Schafer, J. L., & Kang, J. (2008). Average causal effects from nonrandomized studies: A practical guide and simulated example. *Psychological Methods, 13*(4), 279–313. https://doi.org/10.1037/a0014268<br>- Keller and Branson, 3.2 and 4.1<br>Optional:<br>- Morgan & Winship, chapter 6 |
| W7<br>(Mar 4) | Doubly-robust identification and estimation<br><br>Problem set 2 due.<br><br>- Kang, J. D. Y., & Schafer, J. L. (2007). Demystifying double robustness: A comparison of alternative strategies for estimating a population mean from incomplete data. *Statistical Science, 22*(4), 523–539. https://doi.org/10.1214/07-STS227<br>- Glynn, A. N., & Quinn, K. M. (2010). An introduction to the augmented inverse propensity weighted estimator. *Political Analysis, 18*(1), 36–56. https://doi.org/10.1093/pan/mpp036<br>- Keller and Branson, 3.2 and 4.1<br>Optional:<br>- Morgan & Winship, chapter 7.3 |
| W8<br>(Mar 11) | In-class mid-term examination |
| W9<br>(Mar 18) | *Spring break* |
| W10<br>(Mar 25) | Machine learning for causal inference<br><br>Problem set 3 assigned.<br><br>- Hill, J., Perrett, G., & Dorie, V. (2023). Machine Learning for Causal Inference. In *Handbook of Matching and Weighting Adjustments for Causal Inference* (pp. 415-444). Chapman and Hall/CRC. http://dx.doi.org/10.1201/9781003102670-20<br>- Athey, S., & Wager, S. (2019). Estimating treatment effects with causal forests: An application. *Observational Studies, 5*(2), 37–51. https://doi.org/10.1353/obs.2019.0001<br>Optional:<br>- Shah, V., Kreif, N., & Jones, A. M. (2021). Machine learning for causal inference: estimating heterogeneous treatment effects. In *Handbook of Research Methods and Applications in Empirical Microeconomics*. https://doi.org/10.4337/9781788976480.00025 |
| W11<br>(Apr 1) | Optimal treatment regimes |

| | |
|---|---|
| | Problem set 3 due.<br><br>- Tsiatis, A. A., Davidian, M., Holloway, S. T., & Laber, E. B. (2019). *Single Decision Treatment Regimes: Fundamentals. Dynamic Treatment Regimes*, 51–67. (Chapter 3.1- 3.4) https://doi.org/10.1201/9780429192692-3<br>- Suk, Y., & Park, C. (2023). Designing optimal, data-driven policies from multisite randomized trials. *Psychometrika. 88*, 1171–1196. Sections 1 & 2 https://doi.org/10.1007/s11336-023-09937-2 |
| W12 (Apr 8) | Regression discontinuity designs: sharp RD designs<br><br>Problem set 4 assigned.<br><br>- Lee, D., & Lemieux, T. (2014). Regression discontinuity designs in social sciences. In Best, H., & Wolf, C. (eds.). *The SAGE handbook of regression analysis and causal inference*. SAGE Publications Ltd. https://doi.org/10.4135/9781446288146.n14<br>Optional:<br>- Shadish, Cook, & Campbell, Chapter 7 |
| W13 (Apr 15) | Regression discontinuity designs: fuzzy RD designs and other variations<br><br>Problem set 4 due.<br><br>- Suk, Y. (2024). Regression discontinuity designs in education: a practitioner's guide, *Asian Pacific Education Review, 25,* 629-645. https://doi.org/10.1007/s12564-024-09956-3 |
| W14 (Apr 22) | Sensitivity analysis<br><br>- Rosenbaum, P. R. (2005). Sensitivity analysis in observational studies. *Encyclopedia of statistics in behavioral science*, *4*, 1809-1814.<br>- VanderWeele, T. J., & Arah, O. A. (2011). Bias formulas for sensitivity analysis of unmeasured confounding for general outcomes, treatments, and confounders. *Epidemiology, 22*(1), 42–52. https://doi.org/10.1097/ede.0b013e3181f74493 |
| W15 (Apr 29) | Presentation for problem sets |
| W16 (May 6) | Presentation for problem sets |

**Grading and assignments:**

Grading will be based on weekly reading responses (10%), problem sets (40%), a midterm examination (40%), and presentations (10%).

1. **Reading responses.** Each week with assigned readings will include a discussion thread on Canvas. You are expected to post at least one question or discussion point based on the assigned readings and respond to at least one classmate's post. Posts do not need to be long, nor is it necessary to cover all the assigned readings for the week, although you're welcome to do so if you have something to share for each.
2. **Problem sets.** The problem sets will be due one week after they are assigned. These will include some combination of theoretical work, coding work, and write-ups to answer

questions. The results should be written in a word processing software. This could be in R Markdown, Microsoft Word, or LaTeX, for example. If you use R Markdown, you may include embedded code snippets. If you choose other formats, you should upload a separate script file containing your commented and replicable code.

3. **Midterm examination.** The midterm examination will take place in class on March 11 during the regular class period. It will cover material from the first seven weeks of the course. Additional details will be provided in class.
4. **Presentations for problem sets.** Each student will present their analysis methods and results for a problem set of their choice. Specific presentation formats will be announced in class.
5. **Attendance.** Regular in-person attendance is required for this class. Missing more than three classes will have a negative impact on your final grade.

**Late work:**

Reading responses are due on class days by the beginning of class (12:55 pm). Reading responses received after the start of class will receive a maximum of 50% credit. Due dates for problem sets will be posted on Canvas. Problem sets submitted after the due date and time will be discounted by 25% per day late. No make-ups will be permitted for the exam except in the case of a documented emergency.

**AI Policy:**

Large language models (LLMs) such as ChatGPT can be useful to help you get started on an early version of a coding problem or to help you find errors or bugs in your code. For written work and/or mathematical arguments, LLMs can be helpful for proof reading or finding errors. You are permitted to use LLMs for these purposes. However, be aware that LLMs often generate incorrect or incomplete answers, so always critically evaluate their output.

Your work and ideas should primarily reflect your own understanding. Do not copy and paste code or text generated by an LLM (or any other source) directly into your work. Doing so constitutes plagiarism and may result in a failing grade for the assignment, as well as potential disciplinary action. For further details, please refer to the TC Student Conduct Code (here), specifically part 2 on Academic Integrity.

The Provost and Dean of the College in conjunction with the Faculty has adopted the following statements to be included on all Teachers College syllabi.

1. **Accommodations** – The College will make reasonable accommodations for persons with documented disabilities. Students are encouraged to contact the Office of Access and Services for Individuals with Disabilities (OASID) for information about registration. You can reach OASID by email at oasid@tc.columbia.edu, stop by 163 Thorndike Hall or call 212-678-3689. Services are available only to students who have registered and submit appropriate documentation. As your instructor, I am happy to discuss specific needs with you as well. Please report any access related concerns about instructional material to OASID and to me as your instructor.
2. **Incomplete Grades** – For the full text of the Incomplete Grade policy please refer to http://www.tc.columbia.edu/policylibrary/Incomplete Grades
3. **Student Responsibility for Monitoring TC email account** – Students are expected to monitor their TC email accounts. For the full text of the Student Responsibility for Monitoring TC email account please refer to

http://www.tc.columbia.edu/policylibrary/Student Responsibility for Monitoring TC Email Account

4. **Religious Observance** – For the full text of the Religious Observance policy, please refer to http://www.tc.columbia.edu/policylibrary/provost/religious-observance/
5. **Sexual Harassment and Violence Reporting** – Teachers College is committed to maintaining a safe environment for students. Because of this commitment and because of federal and state regulations, we must advise you that if you tell any of your instructors about sexual harassment or gender-based misconduct involving a member of the campus community, your instructor is required to report this information to the Title IX Coordinator, Janice Robinson. She will treat this information as private, but will need to follow up with you and possibly look into the matter. The Ombuds officer for Gender-Based Misconduct is a confidential resource available for students, staff and faculty. "Gender-based misconduct" includes sexual assault, stalking, sexual harassment, dating violence, domestic violence, sexual exploitation, and gender-based harassment. For more information, see http://sexualrespect.columbia.edu/gender-based-misconduct-policy-students.

**Emergency Plan:**
TC is prepared for a wide range of emergencies. After declaring an emergency situation, the President/Provost will provide the community with critical information on procedures and available assistance. If travel to campus is not feasible, instructors will facilitate academic continuity through Canvas and other technologies, if possible.

1. It is the student's responsibility to ensure that they are set to receive email notifications from TC and communications from their instructor at their TC email address.
2. Within the first two sessions for the course, students are expected to review and be prepared to follow the instructions stated in the emergency plan.
3. The plan may consist of downloading or obtaining all available readings for the course or the instructor may provide other instructions.