ISS5096: Experiments and Causal Inference

Jaewon Yoo

Assistant Professor
Institute of Service Science
National Tsing Hua University

Spring 2026

Course Duration: Feb 2026 – Jul 2026	*Classroom:* TSMC Bldg. R406
Class Time: Thursday, 14:20 pm- 17:20 pm	*TA:* TBA

COURSE DESCRIPTION:
This course introduces experimental methods for causal inference that are widely used in a broad array of domains such as marketing and information systems. The focus of the course is on delivering a breadth of substantive topics and methodological considerations that emerge in utilizing the identification-oriented methods. Throughout the course, we will discuss topics that are related to methods such as randomized controlled trials (RCT), difference-in-differences (DiD), matching methods such as propensity score matching (PSM) and coarsened exact matching (CEM), and more advanced topics such as regression discontinuity designs (RDD), double-debiased machine learning (DML), synthetic control methods (SCM), and its extensions such as synthetic difference-in-differences (synth. DiD). To foster learning experiences, students will review relevant research papers on each topic and be asked to actively engage in presentations and discussions about the nature of causation and alternative means of inferring causal relationships.

Students will also carry out a collaborative group project for which they will design an experiment and associated plan of analysis in an attempt to draw business insights. Leading tech companies such as Netflix, Microsoft, and Amazon are growingly paying attention to the importance of business experimentation and causal inference to enhance their decision making, and are hiring employees that are equipped with such analytical tools/mindsets (e.g., watch a recent video highlighting this by causal inference science team at Netflix). In line with this growing trend, this course also caters to those that are interested in joining the industry after graduation with one condition: a willingness to work hard on possibly unfamiliar materials.

COURSE GOALS:
- Learn how to determine which methods and results best support empirical inference questions.
- Be familiar with causal inference methods that are widely used for business analytics.
- Understand the trade-offs in the design, analysis, and reporting of field/quasi/natural experiment methods.

PREREQUISITES:
- **Math:** undergraduate-level probability and statistics; some experiences with regression analysis/econometrics will be helpful.
- **Programming:** Knowledge of statistical programming (e.g., R and Python)

TEXTBOOK:
Students can choose to purchase either the hard copy or the e-book for both textbooks 1 and 2, but please be prepared to have the textbooks ready before the beginning of the semester. If you choose to order

online, be aware that it might take time for the textbooks to be delivered. So, if you plan to take the course, please make sure to place your order early enough so the books arrive in time.

1. **Mostly Harmless Econometrics: an empiricist's companion** by J. D. Angrist and Jorn-Steffen Pischke, Princeton University Press, 2008.
2. **Causal Inference for Statistics, Social, and Biomedical Sciences** by G. Imbens and D. Rubin, Cambridge University Press, 2015.
3. **The Effect: An Introduction to Research Design and Causality** by N. Huntington-Klein, CRC Press, 2021. (Opensource PDF version provided by the author **here**.)

The following books are optional but may prove useful for additional coverage of some of the course topics.

- J. D. Angrist, and Jorn-Steffen Pischke. 2015. *Mastering 'metrics: The path from cause to effect.* Princeton University Press.
- Cunningham, Scott. *Causal Inference: The Mixtape.* 2021. Yale University Press.
- Hernán, Miguel A., and James M. Robins. 2020. *Causal inference: What If.* CRC Press. Taylor and Francis Group. (Opensource PDF version provided by the authors **here**.)
- Morgan, Stephen L., and Christopher Winship. 2015. *Counterfactuals and Causal Inference, Second Edition.* Cambridge University Press.
- Pearl, Judea. 2000. *Causality: Models, Reasoning, and Inference.* New York: Cambridge University Press.
- Blackwell, M., *A User's Guide to Statistical Inference and Regression.* (Opensource PDF version provided by the authors **here**.)
- Rosenbaum, Paul R. 2002. *Observational Studies, 2nd edition*. Springer-Verlag.
- Rosenbaum, Paul R. 2009. *Design of Observational Studies*. Springer Series in Statistics.
- Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data, Second Edition.* MIT Press.
- Rubin, Donald. 2006. *Matched Sampling for Causal Effects.* Cambridge University Press.
- Manski, Charles F. 1995. *Identification Problems in the Social Sciences.* Cambridge: Harvard University Press.
- Chernozhukov, Victor, Christian Hansen, Nathan Kallus, Martin Spindler, and Vasilis Syrgkanis. 2025. *Applied Causal Inference Powered by ML and AI.* (Unpublished. Opensource PDF version provided the authors **here**.)

COURSE EVALUATION:

Grading Policy

The Effect assignments	25%
Problem sets	20%
Final research project	35%
One-page summaries & paper presentations (peer-reviewed)	10%
Attitude/Participation	5%

Attendance	5%
Total	100%

Course Requirements

1. **AI Use Policy*:** Students are permitted to utilize AI tools or Large Language Models (LLMs; e.g., ChatGPT, Claude 2, Bard) responsibly as aids for brainstorming and generating initial drafts for assignments. <u>However, it is imperative that the final submission predominantly reflects the student's understanding and personal input.</u> To maintain academic integrity, students who choose to use AI tools must adhere to the following guidelines:
   a. **Thorough Explanation:** Students must provide a detailed explanation of how the AI tool was used in the completion of the assignment.
   b. **Original AI Responses:** Alongside their submission, students are required to submit the original responses generated by the AI tool.
   c. **Critical Evaluation and Personal Input:** Students should critically evaluate the AI-generated content and clearly indicate the portions of the work that have been modified or expanded upon with their own insights and understanding.
   d. **Proper Attribution:** Proper attribution should be given to the AI-generated content used, clearly indicating the sections that are AI-generated.

   ** Remember, <u>attempting to cheat the system by heavily relying on AI-generated content without substantial personal input is ultimately the student's loss</u>, as it undermines the learning process and personal growth.

2. **Academic Honesty and Plagiarism*:** All work submitted for academic evaluation must be the student's own. The penalty for violation of academic integrity will result in a zero for that assignment for the first time. Subsequent violation(s) will result in a failing grade for the course. Plagiarism will also not be tolerated.

   Academic dishonesty comprises of, but is not limited to, the following:
   - Cheating: Copying from other students' quizzes and assignments or allowing others to copy from one's own.
   - Plagiarism: Using other people's original work without giving appropriate credit or acknowledgment to the authors or sources.
   - Self-plagiarism: Submitting a piece of work in more than one course without the explicit permission of the instructors involved.
   - Misrepresentation of authorship: Submitting work as one's own, which has been prepared by or purchased from others.

   Students will be asked to upload their submission materials to Turnitin.com, an online plagiarism checker, to ensure academic integrity. Read more about online submission on http://learning.site.nthu.edu.tw/p/412-1319-7120.php?Lang=en.

3. **Attendance:** All students are expected to attend every class. Please bring your own hard copy of the course materials, which will be distributed by the instructor before class. If you have any urgent reason to miss a class, you are still responsible for the materials covered during the class and are expected to complete the required work. Attendance will be taken on a regular basis and will count

towards your participation score (5%). Class missing will cause about 1% loss of final grade. Students who miss a class should inform me or the TA prior to the class via email or phone call.

4. **Attitude/Participation:** In class, the most important thing for the students is to stay active and engaged about the topic being discussed. Positive contributions to class discussions will increase your score towards attitude. When we discuss a topic in class, effective discussions are only possible if everyone is well prepared. Please, be prepared to open and engage in discussions with your opinions and thoughts.

5. **Problem Sets and '*The Effect*' Assignments:** Methods are tools, and it isn't very instructive to read a lot about hammers or watch someone else wield a hammer -- you need to get your hands on a hammer or two. In this course, you will have 1) problem sets, which will be a mix of conceptual questions, analytic problems, computer simulations, and data analysis that closely resemble what we discuss in lectures, and 2) assignments from our third textbook, *The Effect*. I encourage you to work in groups on the homework, but you always need to write your own solutions including your computer code. Also, it is hugely beneficial to attempt the problems sets on your own before working in groups. Assignments are to be submitted on the due date. *Late submissions will be penalized 1 percentage point of the assignment's weight per day* (e.g., for an assignment worth 7 % of the course grade, turning it in 3 days late caps the attainable score at 4%).

6. **Journal Article Review**
    a. **One-Page Summaries:** Before class starts, every student should submit 1-page summary of the assigned papers which will be discussed in class. A guideline will be distributed for students to use when summarizing the papers.
    b. **Paper Introduction Presentations (Peer-Reviewed):** Throughout the semester, students will be reading academic publications/papers that utilize different causal inference methods. Students will be asked to present the papers they reviewed each week (reading list will be distributed) by taking turns. The contents to be necessarily included are the value of the topic and motivational story. Other audience should also read the paper prior to the class for discussion. The presentations will be peer-reviewed.

7. **Research Project:** In lieu of a final exam, this course requires students to write a short paper applying or extending the causal inference methods we learn in this class. It should be no longer than 20 double-spaced pages and focus on the research design, data, methodology, results, and analysis.
    a. **Progress Report & Project Presentation (Peer-Reviewed):** Students will generate a potential research idea based on what we will discuss and suggest how to collect data and what causal inference methods/models can be used to carry out the research. The final project presentation will be peer-reviewed.

Here is a brief timetable for the projects:

b. **By Week 2:** Find a collaborator or obtain permission from the instructor to work on an individual project.
c. **By Week 5:** Submit a short (half-page) description of your proposed project and a feasible plan for carrying out the research.
d. **By Week 10:** Submit a brief (no longer than 5 page) page memo of your main results, including tables, figures, and brief analysis. For methodological projects, this should include a description of the method and any analytical/simulation results. You will be

required to give feedback on another group's project, which will be counted toward the overall grade based on attentiveness and usefulness of the feedback provided.

e. **By Week 16:** Submit your final version of the project.

COURSE SCHEDULE:

Week	Topics	Reading
1	**Introduction and Potential Outcomes:** - Neyman-Rubin causal model - Fundamental problem of causal inference (FPOCI) - Causal estimands - Post-treatment bias under the truncation by death problem Group forming & Ice breaking activities.	- Imbens & Rubin, Ch. 1 - Angrist & Pischke, Ch.1 - Holland, P. W. 1986. "Statistics and Causal Inference". Journal of the American Statistical Association, Vol. 81, No. 396: 945-960.
2	**Randomization Inference:** - Randomized experiments - Fisher's approach to inference, permutation tests - Sharp null, randomization distribution Find a collaborator for the group project.	- Imbens & Rubin, Ch. 5 (Skim Ch. 4 for some definitions.) - Rosenbaum, Paul R. 2002. *Observational Studies*. Springer-Verlag. 2nd edition. Ch. 2
3	**Inference for the Average Treatment Effect:** - Neyman's approach to inference for the ATE - Finite-sample vs superpopulation inference	- Imbens & Rubin, Ch. 6, 9 (Skip 9.6–9.7), and 10 (Skip 10.6–10.7) - Angrist & Pischke: Ch. 2
4	**Linear Regression and Randomized Experiments:** - Simple linear regression in experiments - Covariate adjustment in experiments with regression Students' Paper Introduction Presentations: 1.1. and 1.2.	- Imbens & Rubin, Ch. 7, 9 (9.6–9.7), and 10 (10.6–10.7) - Winston Lin. "Agnostic notes on regression adjustments to experimental data: Reexamining Freedman's critique." Ann. Appl. Stat. 7 (1) 295 - 318, March 2013. - Freedman, David. 2008. "On regression adjustments to experimental data."
5	**Observational Studies I:** - Selection on observables - Regression for observational data Students' Paper Introduction Presentations: 1.3. and 1.4. Submit a half-page description of your proposed project & plan.	- Angrist & Pischke, Ch. 3

6	**Observational Studies II:** - DAGs and the back-door criterion - Sensitivity analysis - Partial identification Construct a DAG for your own research project!	- Imbens & Rubin, Ch. 21 and 22 - Morgan & Winship, Ch. 4
7	In-Class Group Project Meetings (30 minutes per group): The purpose of the meetings is to help each group materialize their ideas/topics for the group project.	
8	**Instrumental Variables I:** - Noncompliance in randomized experiments - Instrumental variables in observational studies Students' Paper Introduction Presentations: 2.1. and 2.2.	- Imbens & Rubin, Ch. 23 and 24 - Angrist & Pischke, Ch. 4
9	**Instrumental Variables II:** - Two-stage least squares - Review of IV applications Students' Paper Introduction Presentations: 2.3. and 2.4.	
10	**Panel Data, Fixed Effects, and Difference-in-Differences:** - Fixed effects and first differences - Difference-in-differences * Submit a progress report. Students' Paper Introduction Presentations: 3.1. and 3.2.	- Angrist & Pischke, Ch. 5 - Imai, Kosuke and In Song Kim. (2019). When Should We Use Unit Fixed Effects Regression Models for Causal Inference with Longitudinal Data? American Journal of Political Science, Vol. 63, No. 2 (April), pp. 467-490.
11	**Difference-in-Differences Cont.:** - Applications of difference-in-differences Students' Paper Introduction Presentations: 3.3. and 3.4.	
12	**Matching and Weighting Estimators:** - Propensity scores, matching, and weighting Students' Paper Introduction Presentations: 4.1. and 4.2.	- Imbens & Rubin, Ch.13, 15, and 18 - Stuart, Elizabeth. 2010. Matching Methods for Causal Inference: A Review and a Look Forward. Stat. Sci. Vol. 25, No. 1: 1–21
13	**Regression Discontinuity Designs:** - Sharp RD designs, identification - Estimation and bandwidth selection Students' Paper Introduction Presentations: 4.3. and 4.4.	- Angrist & Pischke, Ch. 6 - Imbens, Guido W., and Thomas Lemieux. 2008. Regression Discontinuity Designs: A Guide to Practice. J. Econom. 142: 615-35.

14	**Regression Discontinuity Designs Cont.:** - Fuzzy RD designs. - Applications of RDDs. Students' Paper Introduction Presentations: 5.1. and 5.2.	- Cattaneo, M., Idrobo, N., & Titiunik, R. (2020). A Practical Introduction to Regression Discontinuity Designs: Foundations (Elements in Quantitative and Computational Methods for the Social Sciences). Cambridge: Cambridge University Press. doi:10.1017/9781108684606 / Preprint Version
15	**Causal Mechanisms and Effect Heterogeneity** - Causal heterogeneity and effect modification. - Controlled direct effects, natural direct and indirect effects. - Causal mediation analysis. - Individual treatment rules. * Submit a final project report.	- Imai, K., Keele, L., & Yamamoto, T. (2010). Identification, inference and sensitivity analysis for causal mediation effects. Statistical Science, 25(1):51–71. - 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. - Acharya, A., Blackwell, M., & Sen, M. (2016). Explaining causal findings without bias: Detecting and assessing direct effects. American Political Science Review, 110(3), 512-529.
16	**Final Group Project Presentations**	

READING ASSIGNMENTS:

1. Reading Materials for Field Experiments:

1) *Chapter 2. The Experimental Ideal* in Joshua D. Angrist & Jörn-Steffen Pischke, 2009. "Mostly Harmless Econometrics: An Empiricist's Companion," Economics Books, Princeton University Press, edition 1, number 8769.
2) Elisa Montaguti, Scott A. Neslin, Sara Valentini (2016) Can Marketing Campaigns Induce Multichannel Buying and More Profitable Customers? A Field Experiment. Marketing Science 35(2):201-217.
3) Navdeep S. Sahni, Dan Zou, Pradeep K. Chintagunta (2017) Do Targeted Discount Offers Serve as Advertising? Evidence from 70 Field Experiments. Management Science 63(8):2688-2705.

4) Bapna, Ravi, Jui Ramaprasad, Galit Shmueli, and Akhmed Umyarov. "One-way mirrors in online dating: A randomized field experiment." Management Science 62, no. 11 (2016): 3100-3122.
5) *Chapter 1. Experiments and Generalized Causal Inference* in Cook, Thomas D., Donald Thomas Campbell, and William Shadish. Experimental and quasi-experimental designs for generalized causal inference. Boston, MA: Houghton Mifflin, 2002.

2. Reading Materials for Instrumental Variables (IVs):

1) *Chapter 4. Instrumental Variables in Action: Sometimes You Get What You Need* in Joshua D. Angrist & Jörn-Steffen Pischke, 2009. "Mostly Harmless Econometrics: An Empiricist's Companion," Economics Books, Princeton University Press, edition 1, number 8769.
2) Angrist, Imbens, G. W., & Rubin, D. B. (1996). Identification of Causal Effects Using Instrumental Variables. Journal of the American Statistical Association, 91(434), 444–455.
3) Florian Zettelmeyer, Fiona Scott Morton, & Jorge Silva-Risso. (2006). How the Internet Lowers Prices: Evidence from Matched Survey and Automobile Transaction Data. Journal of Marketing Research, 43(2), 168–181.
4) Dewan S, Ramaprasad J (2012) Research Note—Music Blogging, Online Sampling, and the Long Tail. Inf. Syst. Res. 23(3-part-2):1056–1067.
5) Barron, Kyle, Edward Kung, and Davide Proserpio. "The effect of home-sharing on house prices and rents: Evidence from Airbnb." Marketing Science 40.1 (2021): 23-47.

3. Reading Materials for Difference in Differences (DiD):

1) *Chapter 5. Parallel Worlds: Fixed Effects, Difference-in-Differences, and Panel Data* in Joshua D. Angrist & Jörn-Steffen Pischke, 2009. "Mostly Harmless Econometrics: An Empiricist's Companion," Economics Books, Princeton University Press, edition 1, number 8769.
2) Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. "How much should we trust differences-in-differences estimates?" The Quarterly journal of economics 119.1 (2004): 249-275.
3) Dranove, David, et al. "Is more information better? The effects of “report cards” on health care providers." Journal of political Economy 111.3 (2003): 555-588.
4) Goldfarb, Avi, and Catherine E. Tucker. "Conducting research with quasi-experiments: A guide for marketers." Rotman School of Management Working Paper 2420920 (2014).
5) Foerderer, Jens, Nele Lueker, and Armin Heinzl. "And the Winner Is…? The Desirable and Undesirable Effects of Platform Awards." Information Systems Research 32, no. 4 (2021): 1155-1172.

4. Reading Materials for Matching Methods:

1) Gordon, Brett R., et al. "A comparison of approaches to advertising measurement: Evidence from big field experiments at Facebook." Marketing Science 38.2 (2019): 193-225.

2) Xu, Kaiquan, Jason Chan, Anindya Ghose, and Sang Pil Han. "Battle of the channels: The impact of tablets on digital commerce." Management Science 63, no. 5 (2017): 1469-1492.
3) Adamopoulos, Panagiotis, Vilma Todri, and Anindya Ghose. "Demand effects of the Internet-of-things sales channel: Evidence from automating the purchase process." Information Systems Research 32.1 (2020): 238-267.
4) Kim, Jun Hyung, et al. "Home-tutoring services assisted with technology: Investigating the role of artificial intelligence using a randomized field experiment." Journal of Marketing Research (2021): 00222437211050351.
5) Son, Yoonseock, Wonseok Oh, Sang Pil Han, and Sungho Park. "When loyalty goes mobile: Effects of mobile loyalty apps on purchase, redemption, and competition." Information Systems Research 31, no. 3 (2020): 835-847.

5. Reading Materials for Regression Discontinuity Design (RDD):

1) *Chapter 6. Getting a Little Jumpy: Regression Discontinuity Designs* in Joshua D. Angrist & Jörn-Steffen Pischke, 2009. "Mostly Harmless Econometrics: An Empiricist's Companion," Economics Books, Princeton University Press, edition 1, number 8769.
2) Caroline Flammer (2015) Does Corporate Social Responsibility Lead to Superior Financial Performance? A Regression Discontinuity Approach. Management Science 61(11):2549-2568.
3) Wesley Hartmann, Harikesh S. Nair, Sridhar Narayanan, (2011) Identifying Causal Marketing Mix Effects Using a Regression Discontinuity Design. Marketing Science 30(6):1079-1097.
4) Jo, Wooyong, et al. "Protecting consumers from themselves: Assessing consequences of usage restriction laws on online game usage and spending." Marketing Science 39.1 (2020): 117-133.
5) Flammer, Caroline, and Pratima Bansal. "Does a long-term orientation create value? Evidence from a regression discontinuity." Strategic Management Journal 38, no. 9 (2017): 1827-1847.

ADDITIONAL READING MATERIALS:

- Holland, Paul W. "Statistics and causal inference." Journal of the American statistical Association 81, no. 396 (1986): 945-960.
- Lin, Winston. "Agnostic notes on regression adjustments to experimental data: Reexamining Freedman’s critique." The Annals of Applied Statistics 7, no. 1 (2013): 295-318.
- Freedman, David A. "On regression adjustments to experimental data." Advances in Applied Mathematics 40, no. 2 (2008): 180-193.
- Stuart, Elizabeth A. "Matching methods for causal inference: A review and a look forward." Statistical science: a review journal of the Institute of Mathematical Statistics 25, no. 1 (2010): 1-21.
- Imai, Kosuke, and In Song Kim. "When should we use unit fixed effects regression models for causal inference with longitudinal data?" American Journal of Political Science 63, no. 2 (2019): 467-490.

- Baker, Andrew C., David F. Larcker, and Charles CY Wang. "How much should we trust staggered difference-in-differences estimates?" Journal of Financial Economics 144, no. 2 (2022): 370-395.
- Arkhangelsky, Dmitry, Susan Athey, David A. Hirshberg, Guido W. Imbens and Stefan Wager. 2021. "Synthetic Difference-in-Differences." American Economic Review, 111 (12): 4088–4118.
- Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. Journal of Econometrics, 225(2):254–277.
- Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. Journal of Econometrics, 225(2):200–230.
- Victor Chernozhukov, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, James Robins, Double/debiased machine learning for treatment and structural parameters, The Econometrics Journal, Volume 21, Issue 1, 1 February 2018, Pages C1–C68, https://doi.org/10.1111/ectj.12097
- Angrist, Joshua D., and Jörn-Steffen Pischke. 2017. "Undergraduate Econometrics Instruction: Through Our Classes, Darkly." Journal of Economic Perspectives 31 (2): 125–44.
- Angrist, Joshua D., and Jörn-Steffen Pischke. 2010. "The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics." Journal of Economic Perspectives 24 (2): 3–30.