The Hong Kong University of Science and Technology
UG Course Syllabus

Experiments and Quasi-experiments in the Social Sciences (Fall 2025)

SOSC4250/SOSC5260
3 Credits
Pre-requisites: SOSC 2400

Instructor

- **Name:** David HENDRY
- **Email:** hendry@ust.hk
- **Office Hours:** Wednesdays and Fridays, 12:00-13:00 (or by appointment)

Teaching Assistant

- **Name:** LI Jiajun
- **Email:** jiajun.li@connect.ust.hk

Course Description

This course explores the most popular class of statistical methods used for causal inference in the social sciences. Working within the potential outcomes framework, we discuss how the logic of inference for randomized experiments is the same as for non-randomized (observational) studies under certain additional assumptions. Though randomized experiments serve as the gold standard for causal inference, we note how it may sometimes be reasonable to treat non-experimental data as if it had been drawn from an experiment. Usually, this involves some knowledge about how the natural world produced the data through a quasi-random process. Research designs and methods covered include randomized experiments, matching, instrumental variables, difference- in-differences, synthetic control, and regression discontinuity designs. In turn, we discuss how all of these methods require a unique set of assumptions to allow us to make valid causal inferences. Throughout the course we will draw examples from across the social sciences to illustrate the vast range of applications of these methods. Furthermore, the course will include computing sessions during which students are taught how to implement the techniques using modern statistical software.

Intended Learning Outcomes (ILOs)

By the end of this course, students should be able to:

1. Understand the history and development of the experimental method across the social sciences.
2. Understand the history and development of quasi-experimental research designs across the social sciences.
3. Understand the history and development of the potential outcomes framework.
4. Identify and understand the major identification assumptions and data structures required for credible causal inference in modern applied social science statistics.
5. Conduct and interpret statistical analyses of data from social science research designs using experimental and quasi-experimental designs.
6. Apply their knowledge of how to conduct and interpret statistical analyses to original social science problems.

Assessment and Grading

This course will be assessed using criterion-referencing and grades will not be assigned using a curve. Detailed rubrics for each assignment are provided below, outlining the criteria used for evaluation.

Assessments:

Assessment Task	Contribution to Overall Course grade (%)	Due date
Fundamentals Problem Set	10%	September 26
Group Presentation in Experiment Workshop	10%	September 29
Computing Problem Sets	50%	5 Problem Sets: • October 3 • October 17 • October 31 • November 14 • November 28
Final Paper	20%	During exam period
Attendance	10%	

* Assessment marks for individual assessed tasks will be released within two weeks of the due date.

Mapping of Course ILOs to Assessment Tasks

Assessed Task	Mapped ILOs	Explanation
Fundamentals Problem Set	ILO1, ILO2, ILO3, ILO4	Students will complete one homework assignment consisting of questions about the potential outcomes framework. Responding to these questions will involve some basic mathematics, understanding of philosophical issues regarding causality in the potential outcomes framework, and interpretation of statistical results. Students are encouraged to use any class notes and books or supplemental materials that they find useful, and to work with other students in the class. However, each student must submit an individual assignment. Though cooperation and use of notes and books is encouraged, students must put answers into their own words and plagiarism will not be tolerated.
Group Presentation in Experiment Workshop	ILO4, ILO5	We will hold a one-day workshop during the semester focused on applied social science journal articles that use experimental methods. During the workshop, students will give group presentations focused on a summary and critique of an article of their choosing in consultation with the instructor. A detailed description of the assignment, as well as a schedule for the presentations, will be provided in class and on Canvas.
Computing Problem Sets	ILO4, ILO5, ILO6	Throughout the semester, students will complete a series of structured problem sets primarily focused on

		performing statistical analysis using R and accompanied written interpretation of statistical results. For each problem set, students will be provided with a dataset and a series of tasks to perform. Answers should be submitted in pdf format, with the computer code used to produce the results included (rendered R Markdown documents with the computer code embedded inline in the document are welcome, but not required). Examples of similar analyses will be covered during the lectures. Students are encouraged to use any class notes, readings, or supplemental materials that they find useful, and to work with other students in the class. However, each student must submit an individual assignment. Though cooperation and use of notes and books is encouraged, students must put answers into their own words and plagiarism will not be tolerated. **Note: When calculating final grades on Computing Problem Sets for SOSC4250 students, the lowest score will be dropped; for SOSC5250 students, all scores will be counted.**
Final Paper	ILO4, ILO5, ILO6	In a paper (of about 3000 words for SOSC4250 and about 5000 words for SOSC5250), students will conduct an original data analysis on a topic of their choosing and write up the results in the style of a research note journal article. Student topics should be chosen in consultation with the instructor. A detailed description of the assignment will be provided in class and on Canvas.
Attendance	ILO1, ILO2, ILO3, ILO4, ILO5, ILO6	After the Add/Drop period, attendance will count toward students' final grade. Students can miss two class sessions for any reason without penalty. Any additional absences will be penalized unless they are valid excuses backed up by documentation.

Grading Rubrics

Fundamentals Problem Set

This assignment will involve a mix of basic arithmetic calculations, interpretations of statistical results, and synthesis and application of lecture and reading material.

- The number of points assigned to each question will be given to the students.
- Where calculations are required, a correct final answer will receive full credit. Incorrect final answers will be downgraded by examining the steps that the student took to reach their final answers. Markers will be encouraged to grant partial credit where possible. It is therefore important for students to show all of their work.
- Where interpretations of statistical results are required, there is usually a correct version of precise language that can be drawn from the course material. Students are therefore encouraged to use the exact language for each situation that is presented in the lectures and readings. Use of the correct precise language will receive full credit. Imprecise but generally correct statements will be downgraded, while incorrect application of language will receive no credit. Markers will be encouraged to grant partial credit where possible.
- Where students are asked to present their own original ideas (e.g., a substantive real-world example that fits certain criteria), or opinions (e.g., choosing which among two situations is a more important assumption for some statistical technique), markers will focus on the reasoning given by the student. It is therefore very important for students to focus on using precise language to explain why their final judgments sensible or logical.
- English prose is important, but markers will be encouraged to focus more on the substance of answers than grammar, style, etc.

Computing Problem Sets

These assignments will involve providing the students with a dataset, a description of the dataset (and how it has been used in published work, if applicable), and an R Markdown template with a series of questions. The questions will involve a mix of asking students to write R code to produce statistical results, interpretation of statistical results, and synthesis and application of lecture and reading material.

- The number of points assigned to each question will be given to the students.
- Where generation of statistical results is required, a correct final answer will receive full credit. Incorrect final answers will be downgraded by examining the steps that students took to reach their final answers. Markers will be encouraged to grant partial credit where possible. It is therefore important for students to include all of their computing code.
- Where interpretations of statistical results are required, there is usually a correct version of precise language that can be drawn from the course material. Students are therefore encouraged to use the exact language for each situation that is presented in the lectures and readings. Use of the correct precise language will receive full credit. Imprecise but generally correct statements will be downgraded, while incorrect application of language will receive no credit. Markers will be encouraged to grant partial credit where possible.
- Where students are asked to present their own original ideas (e.g., a substantive real-world example that fits certain criteria), or opinions (e.g., choosing which among two situations is a more important assumption for some statistical technique), markers will focus on the reasoning given by the student. It is therefore very important for students to focus on using precise language to explain why their final judgments sensible or logical.
- English prose is important, but markers will be encouraged to focus more on the substance of answers than grammar, style, etc.

Group Presentation in Experiment Workshop

Students will be assigned to a group with other students and asked to present on the details of an published study using experimental methods.

- Marking distinctions will be based on the following criteria:

- A: Demonstrates correct descriptions of the argument and methodology of the study, as well as insightful analysis of the strengths and weaknesses, as related to course material learned through readings and lectures.
- B: Demonstrates generally correct descriptions of the argument an methodology of the study, as well as clear analysis of the strengths and weaknesses, with some connections to course material.
- C: Fails to cover one or more important aspects of the argument and methodology of the study or its strengths and weaknesses; connections to course material somewhat limited.
- D: Presentation is lacking in clear and correct description of the argument and methodology of the study; analysis of strengths and weaknesses not connected to course material.
- F: Clear and correct description of the argument and methodology of the study are lacking and no analysis of strengths and weaknesses is provided.

- Students will be asked to evaluate the work of fellow group members. This information will be incorporated into individual students marks after due diligence if problems are identified.

Final Paper

Students will be asked to construct a short version of a paper that emulates the style of published journal articles. Students will choose their own substantive topics and data sources, in consultation with the instructor. Papers can be completely original analyses or replications and extensions of published work, but must utilize one of the methodological approaches covered in the course.

- Marking distinctions will be based on the following criteria.
  - A: Student provides a clear and insightful (1) description of the substantive topic, (2) the research question or puzzle, (3) hypotheses, (4) statistical results, (5) narrative interpretation of statistical results, (6) analysis of the validity of identification assumptions, (7) conclusion that synthesizes the statistical results and argument about identification assumptions to make a judgment about causal effects.
  - B: Generally all of the criteria for an A grade are covered, but detail or insight are lacking; or one or two of the criteria are not covered, but the criteria that are covered are done so with sufficient clarity and insight.
  - C: Generally only some of the criteria for an A grade are covered and clarity and improvements could be made with respect to clarity and insight for the criteria that are covered.
  - D: Analysis is conducted but all the criteria for an A grade are either missing or lacking sufficient explanation.
  - F: Not enough information is presented to demonstrate understanding of the course material.

Final Grade Descriptors:

Grades	Short Description	Elaboration on subject grading description
A	Excellent Performance	Demonstrates a comprehensive grasp of subject matter, expertise in problem-solving, and significant creativity in thinking. Exhibits a high capacity for scholarship and collaboration, going beyond core requirements to achieve learning goals.
B	Good Performance	Shows good knowledge and understanding of the main subject matter, competence in problem-solving, and the ability to analyze and evaluate issues. Displays high motivation to learn and the ability to work effectively with others.
C	Satisfactory Performance	Possesses adequate knowledge of core subject matter, competence in dealing with familiar problems, and some capacity for analysis and critical thinking. Shows persistence and effort to achieve broadly defined learning goals.
D	Marginal Pass	Has threshold knowledge of core subject matter, potential to achieve key professional skills, and the ability to make basic

		judgments. Benefits from the course and has the potential to develop in the discipline.
F	Fail	Demonstrates insufficient understanding of the subject matter and lacks the necessary problem-solving skills. Shows limited ability to think critically or analytically and exhibits minimal effort towards achieving learning goals. Does not meet the threshold requirements for professional practice or development in the discipline.

Course AI Policy

The use of generative AI to construct original writing is not permitted. Students are particularly discouraged from over-reliance on generative AI for the generation of original ideas. Students are encouraged to use AI tools to assist with editing English prose.

Communication and Feedback

Assessment marks for individual assessed tasks will be communicated via Canvas within two weeks of submission. Feedback on problem sets will include comments on why any points were deducted, as well as a model answer. Feedback on presentations and final papers will include feedback about strengths and areas for improvement. Students who have further questions about the feedback including marks should consult the instructor within five working days after the feedback is received.

Required Texts and Materials

Required readings should be completed prior to the date they are listed on the schedule. All readings will be provided through Canvas. There is no text that is perfect for this course, and therefore there is no text that students are required to purchase. However, the following books are highly recommended, particularly if students wish to dive deeper into the topics covered in this course or pursue them in their own research.

- Angrist, Joshua D., and Jörn-Steffen Pischke. 2015. *Mastering 'Metrics: The Path from Cause to Effect*. Princeton: Princeton University Press.
- Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.
- Gerber, Alan S., and Donald P. Green. 2012. *Field Experiments: Design, Analysis, and Interpretation*. New York: W.W. Norton & Co.
- Imbens, Guido W., and Donald R. Rubin. 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences*. Cambridge: Cambridge University Press.
- Morgan, Stephen L., and Christopher Winship. 2007. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge: Cambridge University Press.
- Rosenbaum, Paul R. 2010. *Design of Observational Studies*. New York: Springer.
- Shadish, William R., Thomas D. Cook, and Donald T. Campbell. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. New York: Houghton and Mifflin.

Academic Integrity

Students are expected to adhere to the university's academic integrity policy. Students are expected to uphold HKUST's Academic Honor Code and to maintain the highest standards of academic integrity. The University has zero tolerance of academic misconduct. Please refer to Academic Integrity | HKUST – Academic Registry for the University's definition of plagiarism and ways to avoid cheating and plagiarism.