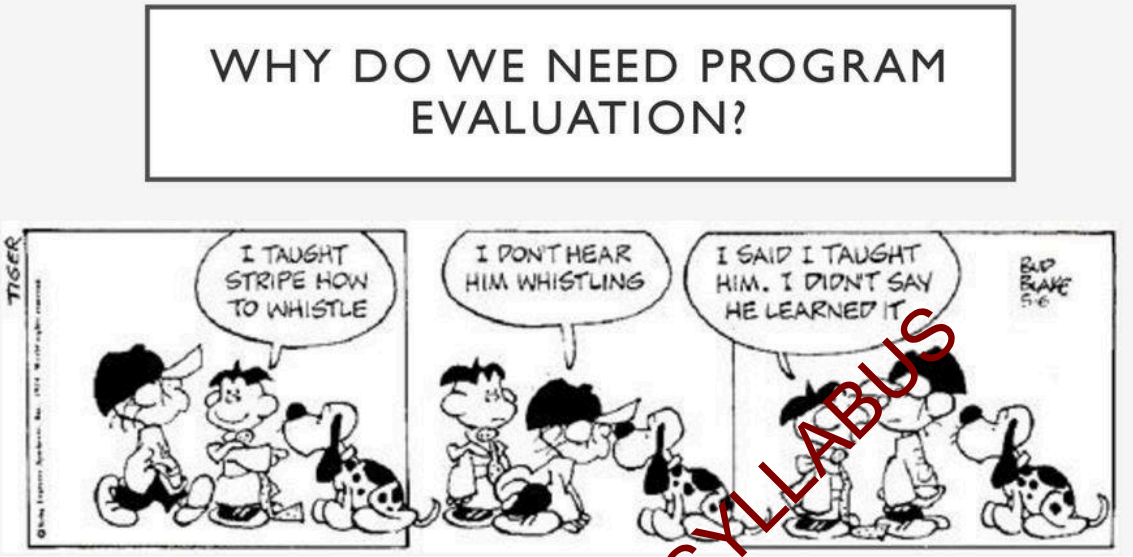
PBPL 26831: Quantitative Evaluation for Public Policy
Section 1 - Summer Session



Class Meetings	Lectures: MTWRF morning Lab Sessions: MTWRF afternoon Location: Keller Center Room 1002 (first floor of Keller Center)
Instructor	Professor Sheng-Hao Lo (shenghaolo@uchicago.edu) Harris School of Public Policy, The University of Chicago
Teaching Assistants	Kunjian Li (kunjianli@uchicago.edu) Isabella Nascimento (isabellapestana@uchicago.edu)

Course Overview

This course is designed for enterprising high school students with an interest in or curiosity about quantitative public policy analysis. Much of the class is dedicated to understanding the effects of policies and addressing empirical, policy-relevant questions using observational data. In doing so, the course provides an introduction to critical, quantitative thinking. Students will be introduced to the essential toolkit of policy analysis, which includes sampling, hypothesis testing, regression, experiments, regression discontinuity, difference-in-differences and matching. Additionally, students will learn to use statistical software to organize and analyze data. More importantly, they will learn the principles of critical thinking that are essential for careful and credible policy analysis.

Course Objectives:
Students can expect to:

- Acquire a toolkit for policy analysis.
- Understand how to consume quantitative evidence responsibly, critically, and skeptically.
- Understand what to evaluate and why we should care about impacts of policies.
- Learn a number of techniques, including foundational statistical programming skills, designed to address that problem and inform the effects of different interventions.
- Understand how to do basic quantitative analysis in their own work.
- Recognize the limits of the tools covered, specifically, when they can and cannot be used to produce meaningful information.

To achieve these goals, we will help you develop rigorous statistical thinking and quantitative reasoning skills, enabling you to make appropriate inferences from data. Together, we will build a powerful toolkit for conducting quantitative analysis, which will serve as a foundation for more advanced tools and techniques you will encounter in Statistics and related courses in the College.

Course Structure:
We will be meeting in person from 9:00 AM to 3:00 PM every weekday for the entire 3-week Summer Immersion Program. Our full-day class will typically be structured as follows:

- **Lectures (9:00 AM-11:30 AM)**: Each morning, we will have two 70-minute lectures with a 10-min break in between. These lectures, along with accompanying slides, will serve as the primary conceptual resources for completing assignments and studying for assessments. Lecture slides will be posted on Canvas in advance of class to allow students to take notes directly on them. Please note that not all information will be pre-written on the slides, so it is important for students to actively take notes during class. These sessions will also include demonstrations of statistical concepts through practice problems, occasional R tutorials for coding concepts, and ample time for questions.
- **Lab sessions (1:00 PM-3:00 PM)**: In the afternoons, our teaching assistants will lead lab sessions that include short lectures on coding in R, providing instructions and practice problems to help you acquire the coding and practical knowledge necessary for completing homework assignments. These sessions also offer another opportunity to work on statistical exercises and ask questions about the material.

Textbooks and Programming in R:
Required:

- Lecture slides and assigned readings (available on Canvas).
- *Critical Thinking in a Data-Driven World*, by Bueno de Mesquita, Ethan and Anthony Fowler. 2019.
- *OpenIntro Statistics (4th edition)*, by Diez, Cetinkaya-Rundel and Barr. 2019.

These are all available on our Canvas site. You can also buy the textbook for the second one.

There is no assigned textbook for coding in R. However, the material of this course is classic and is covered in several excellent textbooks. Feel free to use the following three books.

- James, Witten, Hastie and Tibshirani. *An Introduction to Statistical Learning, with Applications in R.*
- Kosuke Imai. *Quantitative Social Science: An Introduction.*
- Wickham and Grolemund. *R for Data Science.*

This course will require you to follow lectures and complete assignments using the statistical software **R** (and its companion software, **RStudio**). From the outset, we will be using R, so please ensure it is fully installed and functioning on your computer before the course begins. Fortunately, R, RStudio, and all the R libraries we will use in the course are available free of charge.

No prior experience with statistical software or coding is expected. You will need to use R to complete assignments. To facilitate learning, TAs will frequently demonstrate data analysis in R during the sessions, providing hands-on practice. It is beneficial for many students to work in R simultaneously. We will provide the relevant R files on Canvas for you to access.

Supplementary Readings:

- *The Lady Tasting Tea*, by David Salsburg.
- *Naked Statistics: Stripping the Dread from Data*, by Charles Wheelan.

While these reading assignments are not per se required, they are strongly recommended, especially for any student who seeks for more readings in the course. These resources can be very useful for any topics for which you think you need additional practice. I will also provide a suggested reading list by topic for the supplemental textbooks on Canvas.

Communications:

Communication from instructor to students will happen primarily through the posting of materials on Canvas including postings to Announcements and the discussion board. Please note that you are responsible for reading all Canvas Announcements related to the course. To ensure receipt, you may wish to confirm that you have email notifications enabled for Canvas Announcements.

As there are many students in this course, emailing your instructor directly is an ineffective way to have either a logistical or a pedagogical issue resolved. Therefore, we suggest and request that communication from students take the following forms:

- Questions regarding scheduling and other course logistics should be directed to the TA: Kunjian Li (kunjianli@uchicago.edu).
- Questions regarding course material may be posted to the course discussion board, a forum that is monitored by the teaching assistants and instructor. Please note that, while we strive to respond expeditiously to student questions posted on the discussion board,

you should not expect to always receive prompt replies, especially if your question is posted on the weekend or after normal business hours. So, please do not expect to receive a quick response at 9:00 PM on a Friday evening.

Grading and Class Deliverables:
The composition of your overall course grade is as follows:

Lab Assignments and Problem Sets	55%
Two in-class Quizzes	15%
Final Group Presentation	25%
Participation	5%

- **Lab Assignments and Problem Sets**: There are daily assignments throughout the course. In the afternoon sessions, you will work on lab assignments with your group members. These lab assignments will be graded based on completion rather than accuracy. For problem sets, you will collaborate with your group to discuss and solve the problems. Problem sets will be graded based on correctness.

  For all lab assignments and problem sets, please submit only one submission per group rather than individual submissions. We will be using Gradescope to manage assignments and grading. You can find the Gradescope shortcut on the left side of your Canvas menu. You must submit a PDF version of your assignment via that link, and attach any relevant R files in PDF format when needed. We will not consider submissions if they are not uploaded to Gradescope.

  All assignments must be submitted by 11:59 PM Central Time on the specified due date. Late assignments will incur a 25% grade reduction if submitted within 9 hours past the due date and time. Assignments submitted more than 9 hours late will not be accepted. Please note that delayed submissions hinder the timely posting of suggested solutions, which negatively impacts your classmates.

  Your work should reflect your individual knowledge and effort. This means that your write-up for any assignment must be entirely in your own words. If your write-up contains identical language to that of other groups, it will be considered a violation of the collaboration policy. Occasionally, certain parts of an assignment may require individual work with no collaboration allowed. These instances will be clearly indicated in the assignment instructions, and you must complete these sections on your own. If you have any questions about what is permissible, proper attribution, or any other academic honesty issues, please ask the teaching staff!

All assignments must be formatted professionally, as if you were presenting them to a boss, client, or colleague. If you have any confusion about the instructions or materials, feel free to ask specific and clarifying questions about the assignment.

- **In-class Open-book Quizzes**: Instead of midterm exams, you will have short quizzes on 6/23 and 6/30, respectively. The purpose of these quizzes is to provide students with feedback on their performance in a timed situation. They are short quizzes intended to gauge your understanding of the course material. Each quiz accounts for an average of 7.5% of your overall grade. They are designed to be diagnostic rather than punitive, particularly if you are encountering difficulties in grasping the material. Out of the two quizzes, the one in which you achieve the higher score contributes 10% to your overall grade, while the one with your lower score contributes 5%. This weighting scheme aims to reduce the impact of a single poor performance and to recognize your best assessment.

- **Final Group Presentation**: On our final day of class, you (along with your group members) will deliver a brief (15-minute) PowerPoint presentation on your selected topic/paper in public policy. Firstly, you will introduce the topic/issue of your policy interest. Following this, you are expected to summarize, elucidate, and critique how the paper you studied applies to the policy evaluation techniques covered in class. This includes discussing how data analysis could contribute to resolving the issue, identifying complications and complexities to consider, suggesting the type of analysis to conduct (e.g., experiments, regression discontinuity, diff-in-diff, matching, etc.), and proposing recommendations to enhance the study's persuasiveness. Additionally, you are required to listen to other groups' presentations and provide feedback after their presentations.

- **Participation**: Attendance is mandatory unless unforeseen circumstances prevent it. Full credits will be automatically awarded if attendance is taken in both morning and afternoon sessions. To ensure accurate credit allocation and assist the instructional team in identifying you, each student must bring their own name tent to every class.

Grading Scale:

Below is the grading scale that will be utilized for assignments, assessments, and final grades in this course. Please note that A is the highest attainable letter grade at UChicago, and there is no A+ in the grading system employed here:

- ≥93: A | 90-92.99: A-
- 86-89.99: B+ | 83-85.99: B | 80-82.99: B-
- 76-79.99: C+ | 73-75.99: C | 70-72.99: C-
- 66-69.99: D+ | 63-65.99: D | 60-62.99: D-
- <60: F

Tentative Course Outline (subject to change):

- **Part 1: Introduction, Probability and Statistics**
    - **Introduction**: We will go over the syllabus for this course and explore the significance of quantitative methods in public policy. Additionally, we will ensure that R/RStudio is installed and set up on everyone's laptops.
    - **Probability Theory and Random Variables**: Before delving into quantitative policy analysis, it's crucial to grasp the fundamentals of probability and random variables as they form the foundation of statistics. We will cover classical probability as well as delve into more advanced concepts like conditional probability.
    - **Summary Statistics & Exploratory Data Analysis**: We will delve into the fundamentals of descriptive statistics, covering concepts such as measuring center and spread. Additionally, we will conduct basic data analysis using a case study in R/RStudio.

- **Part 2: Descriptive Inference and Causal Inference**
    - **Statistical Inference, Sampling Distribution & Central Limit Theorem**: We will embark on a formal study of inferential statistics, which serves as the cornerstone of social science research. The concept revolves around researchers extracting representative samples of data and drawing inferences about a population based on the characteristics of that sample. Additionally, we will introduce the Central Limit Theorem, which serves as a bridge from selecting a sample to utilizing it to draw inferences about the population.
    - **Linear Regression for Description and Prediction**: Prediction holds broad utility across various fields of study. For instance, can a school district predict which students will drop out to tailor counseling interventions? Can Amazon forecast which product a customer is likely to purchase based on their past transactions? To address such inquiries, we will introduce the OLS regression model, widely employed for predicting and describing the linear relationship between two (or more) variables.
    - **Hypothesis Testing**: The primary objective of statistics is to test hypotheses. For instance, you may conduct an experiment and determine that a particular drug effectively treats headaches. However, if the experiment cannot be replicated, your findings will lack credibility. We will dedicate two days to this crucial topic.
    - **From Correlation to Causation**: Does correlation imply causation? It's essential to underscore the fundamental distinction between these two concepts before arriving at a conclusive recommendation for policy. We will explore evaluation issues, including before-after and simple cross-sectional estimators, as well as treatment parameters of interest and their estimators.

- **<u>Part 3: Toolkit for Program Evaluation & Final Group Presentation</u>**
    - **Experiments**: As we shift our focus towards causal analysis, we will begin introducing additional toolkits for program evaluation, including experiments (also known as randomized controlled trials). Specifically, our aim is to empirically determine the extent to which a change in one variable causes a response in another.
    - **Regression Discontinuity Designs**: We transition to our exploration of quasi-experimental research designs, recognizing that experiments can be costly and not always feasible. Regression discontinuity designs offer valuable insights into estimating the causal effects of various interventions based on eligibility criteria that occur at specific thresholds. This specification enables the identification of treatment effects in non-experimental settings by comparing individuals just above and below a cutoff point.
    - **Difference-in-differences**: We will proceed with our examination of quasi-experimental research designs. While our previous designs have necessitated some level of randomization, many real-world scenarios lack valid randomization. However, with panel data, it becomes feasible to derive a valid estimator—the difference-in-differences estimator—to gauge the treatment effect.
    - **Matching**: This marks the final quasi-experimental research design we will explore in this course. The objective of matching is to mitigate bias in the estimated treatment effect within observational-data studies. This is achieved by identifying, for each treated unit, one or more non-treated units with comparable observable characteristics, thereby balancing out the covariates. Through matching treated units to similar non-treated units, we facilitate a comparison of outcomes between treated and non-treated units, enabling the estimation of the treatment effect.
    - **Final Group Presentation**: On our final day of class, you (along with your group members) will deliver a brief (15-minute) PowerPoint presentation on your selected topic/paper in public policy. See more details from the "Final Group Presentation" description above.

Course Policies:

- General
    - **Regular attendance is necessary to succeed in this course.** However, if you get sick, are caring for a sick relative, have a tough situation, or anything else that becomes an obstacle to your coursework, please inform me and the course assistant as soon as you are able. We will work together to develop appropriate accommodations.
    - The class webpage is available through the Canvas site. I will use it to post announcements, assignments, course materials and grades. Communication will

happen primarily through the posting of materials on Canvas, including postings to Announcements and sending emails. Please note that you are responsible for checking your UChicago email account and Canvas site regularly.

- Learning in Various Modes
  - Public health and/or personal health circumstances vary across individual members of the University community and may change abruptly with limited notice. Students, TAs, and instructors may need to participate remotely for a short time or, in some limited instances, for the entire quarter. To guide expectations and plans, please note the following:
    - **If you are experiencing COVID-19 symptoms or are required to isolate, do not come to class! As soon as possible, contact me or TAs by email if you cannot attend class for this reason.**
    - Students are expected to abide by the University's COVID-19 health requirements AND its specific Protocol for Addressing Confirmed or Suspected COVID-19 Exposures. Note that the Protocol, which addresses self-monitoring, testing, and isolating requirements, represents evolving guidance and is subject to change.

- Recording, Deletion and Copyright
  - All course materials (including, but not limited to, class lectures and discussions, meetings, handouts, exams and web materials) and the intellectual content of the course itself are protected by United States Federal Copyright Law. Students are permitted to make notes solely for their own private educational use. Students and all other persons are expressly forbidden from recording lectures or discussions and from distributing or selling lecture notes and all other course materials without the prior written permission of the instructor. Because the instructor owns the copyright to the classroom presentations and all course materials, any notes taken during those presentations and subsequently sold or distributed to others would constitute an unauthorized derivative work and expose the person or persons involved to individual copyright infringement actions by the instructor.

- Academic Integrity
  - All University of Chicago students are expected to uphold the highest standards of academic integrity and honesty. It is contrary to justice, academic integrity, and to the spirit of intellectual inquiry to submit another's statements or ideas as one's own work. This means that students shall not represent another's work as their own, use impermissible materials during quizzes, or otherwise gain an unfair academic advantage. To do so is plagiarism or cheating, offenses punishable under the University's disciplinary system. Because these offenses undercut the

distinctive moral and intellectual character of the University, we take them very seriously.

- Use of Generative Artificial Intelligence
    - Generative AI tools, such as ChatGPT, are incredibly powerful and potentially transformative. Students are permitted to use these tools for assignments in the course as they would any other internet resource. With that in mind, students should remember that:
        - Generative AI tools, like other internet resources, are sometimes inaccurate. Large language models can generate remarkably coherent and correct-sounding text that is profoundly inaccurate. There is some evidence that these models are especially prone to producing inaccurate responses to prompts in the quantitative domain. As a result, compared to traditional internet resources (the sources of which are known), these tools may confuse and mislead rather than provide guidance towards the solution. Students are always *fully responsible for the accuracy of their written work for the course*.
        - Just as it is not acceptable to copy and paste text into an assignment from a traditional internet resource, it is not acceptable to copy and paste text from generative AI tools. As the collaboration policy states, "*Your* write-up and code should only reflect *your* understanding of the material. As such, these should be written in *your own words*." These principles and this policy are applicable to the generative AI context as well. To comply with the university policies on academic honesty, you must always provide proper documentation and citation of the generative AI tools that you used. When you are not sure about appropriate usage, please ask for clarification.

- Americans with Disabilities Act
    - Students with disabilities needing academic accommodation should contact UChicago's Student Disability Services (SDS). Please see their webpage for contact information (https://disabilities.uchicago.edu).

- Mental Health Services
    - Students differ in how much they know about mental health services. Your use of UChicago's Student Health and Counseling Services (SHCS) is free, confidential and not linked to your academic file. There are no gains from suffering in silence, so please do not hesitate to make use of the services provided by SHCS if you need them. Please see SHCS' mental health webpage for services and contact information (https://wellness.uchicago.edu/mental-health/). And if you are having

serious mental, physical, or other problems, immediately contact the urgent medical care line at (773) 834-WELL.

- Diversity and Inclusion
    - UChicago welcomes, values, and respects students from a wide range of backgrounds and experiences, and we believe that rigorous inquiry and effective public policy problem-solving requires the expression and understanding of diverse viewpoints, experiences, and traditions.
    - UChicago is committed to diversity and rigorous inquiry that arises from multiple perspectives, and encourages thought-provoking discourse that involves not only speaking freely about all issues but also listening carefully and respectfully to the views of others. I concur with this commitment and view the diversity that students bring to my class as a valuable resource and a benefit to learning. I expect to maintain a productive learning environment based on open communication, mutual respect, and non-discrimination, and strive to present materials in a way that is respectful of diverse student backgrounds. As there can always be a gap between intent and execution, suggestions for promoting a positive and open environment are welcomed. Please feel free to correct me on your preferred name and gender pronouns if necessary.