Quantitative Methods in Health Systems Research I

HAD 5744
Fall 2025

Instructor: Alex Hoagland, Ph.D.
Assistant Professor, IHPME
155 College Street, Suite 440
alexander.hoagland@utoronto.ca

Class time and location: Mondays, 11am-1pm. **MS 4279** (1 King's College Cir, Toronto, M5S 3K3)

Office hours: Book appointments at calendly.com/Hoagland-office-hours/
- Tuesdays, 9-11am, or by appointment.
- All appointments are by Zoom unless arranged otherwise in advance.
- Please note that I typically respond to emails and other class communication during normal "business hours," and not on evenings, weekends, or holidays.

Course Description: Introduces quantitative methods frequently used in health systems research, including social epidemiology, applied health economics, health policy, and others. In many applications, researchers want to understand a process by which data and outcomes are generated; however, many data generating processes (DGPs) are possible given observed data. This course deals with how to determine *which* DGPs, and hence which "story", has generated your data. The course uses applications of statistical tests and procedures in the context of distinguishing between models and explores the applications of a range of frameworks to the types of questions addressed by social scientists and health services researchers. It is assumed that students have basic (graduate) training in statistics; if you feel uncomfortable with completing Assignment 0, I recommend you take HAD5772 prior to taking this course.

Evaluation Criteria
- Assignments: 4 group assignments. Each assignment is worth 10% of the final grade, including the presentation/video components. (40% total)
- 1 Final Paper (Phd or Msc) or 2 Referee Reports (Msc only): worth 60% of the final grade.

Assignments are to be completed using R.[1] Assignments and data sets will be posted on the course website. Due dates are listed on the schedule below. The assignments **must be submitted via Quercus.** Assignments can be done in groups of no more than 5 students, but each person must submit their own knitted output file with code.
- Late assignments will be penalized at the rate of 10 percentage points per day.
- Any requested extension longer than 3 days must be made with a formal accommodation from Accessibility Services, at least a week before the deadline.
- 25% of assignment marks (10% of overall grade) come from recorded video presentations

---

[1] I strongly encourage you to use R unless you feel very comfortable with Stata, Python, or another software. We will only discuss R coding principles and examples in class; in previous years, students have attempted to use Stata or Python and then felt they could not engage in group assignments as easily, given the lack of Stata supports. We can't cover all coding languages in only 4 months!

of assignment solutions. **These must be submitted individually**. For each assignment, you will be asked to record videos *of no more than 3 minutes* explaining how you arrived at the answer to a given problem. You should explain both the code you used as well as how you interpreted your results. If applicable, you may also discuss how you would extend your findings in an actual research setting. These videos should be submitted on Quercus in a MP4, AVI, WMV, or MOV format. Students will be randomly selected (with replacement) at the start of each class following a homework assignment to present their answers to the rest of the class. You must be present at the start of the class to receive credit for this component of the assignment if called on.

In addition, each student must complete **a major evaluation or major paper**. **PhD students are expected to do a paper; MSc students have the option to either write 2 referee reports over the course of the semester, or to be a part of a paper team.**

Major Paper**:** The paper provides students with an opportunity to undertake an investigation of a research question of their own choosing on a self-contained topic within their field. Writing will also help with communication skills and familiarization with the structure of journal manuscripts. **Students are expected to pair up in groups of no more than five (and no less than two) to produce a paper.**

To find data for the project, I recommend using publicly available data: a primer on where to look can be found here. For example, there are public use files for the Canadian Community Health Survey (CCHS). Other health related surveys can be downloaded at CHASS (https://datacentre.chass.utoronto.ca/) We will discuss these options more in class.

Students will hand in an outline of their major paper during the term. The proposal/outline must include a brief introduction and literature review, description of data, variables, and methods employed. The final paper should include these as well as results, discussion and conclusion. The paper should be presented in the form of a journal manuscript (following the outline and style of a journal in their field). It must also include a contribution statement listing each author's unique role in producing the final project, agreed upon by all project team members. The final paper is due December 19.

Referee Reports: If you are an MSc student and are not doing a paper proposal, you must instead submit **2 referee reports** over the course of the semester. A referee report critically evaluates a recent HSR paper for its statistical (internal) validity and causal inference techniques, as well as providing suggestions for improvement and future work. For each course topic, there is a list of papers indicated on the reading list (there are approximately three choices per topic). **You must select papers from two different topics.** Your referee report should include: (1) a brief summary of the paper; (2) major concerns you have, including any potential flaws or drawbacks you see in the modeling choices; and (3) minor concerns you have, including ideas for extensions and future research. Your referee report should be no more than 3 pages long. **Reports are due one week after the topic has been covered in class.** Late reports will be discounted by 10 percentage points per day late.

Grading scale: Courses taken for graduate credit are assigned a letter grade according to the School of Graduate Studies usage as follows. While course grades may be collectively "curved,"

no individual grades will be rounded. Please do not ask me to round your grade, as this introduces inequity to other students, and does not come off well for a graduate student in a required methods course of a graduate degree.

Letter Grade	Grade Meaning	Numerical Marks (%)
A+	Excellent	90%-100%
A		85%-89.9%
A-		80%-84.9%
B+	Good	77%-79.9%
B		73%-76.9%
B-		70%-72.9%
FZ	Inadequate	0-69%

Attendance policy: This course is designed for full-time, in-person attendance. I reserve the right to impose a grade penalty for too many unexcused absences. Lectures will not be recorded and virtual attendance is not permitted per IHPME and University policy. Students may create audio recordings of the lectures for their personal use. Recordings are intended to permit lecture content review so as to enhance understanding of the topics presented. Audio-recordings are not substitutes for attending class. Students should note that since audio recordings are to be permitted, their voice may be recorded by others during the class. Please speak to me if this is a concern for you.

Course website: This course has a GitHub repository that contains all relevant materials; you can access the repo at https://github.com/alex-hoagland/HAD5744. Each lecture has a folder containing slides, code, and example papers. The assignments folder contains the relevant questions and data sets. Materials will be updated and/or added throughout the semester.

Software: We will make use of R and RStudio. You do not need a license for R or RStudio and it can be downloaded for free at https://www.r-project.org/ and https://rstudio.com/. Learning R or any program for the first time can be frustrating. Sometimes you can get really bogged down from not knowing a simple command or syntax. This is normal, and part of the learning process. The best way to learn is by relying on the built-in help functions, online short tutorials, class participation or by taking a short course.

In addition to our Assignment 0, which was circulated before the beginning of the term, I recommend these additional resources for getting started with R:
- These slides: https://nickch-k.github.io/EconometricsSlides/Week_01/Week_01_Slides_2_Starting_R.html#1.

Required texts: Both books have a free e-book version or are relatively cheap buys on Amazon. I *highly* recommend owning print copies if you plan to pursue applied research – they are both very handy. **Course readings are expected to be completed prior to class.**
- "The Effect," Nick Huntington-Klein (NHK).
- "Causal Inference: The Mixtape," Scott Cunningham (SC).

Additional (non-required) reading:
- "Principles of Econometrics, 5th edition", R. Carter Hill, William E. Griffiths and Guay C. Lim (HGL).

- A useful online resource: https://tinyurl.com/bdzxbxce
- Example papers for each method are included in the GitHub repo.

Statement on Generative AI in Course Content and Materials: Students may use artificial intelligence tools, including generative AI and GitHub Copilot, in this course as learning aids or to help produce assignments. However, students are ultimately accountable for the work they submit. For each submitted assignment, please include a paragraph explaining (a) how these tools were used and (b) how you as the student integrated these materials into your assignment (e.g., how were the inputs checked for completeness or debugged?).

> *Note*: GitHub Copilot is free for students and easy to integrate into R Studio – I highly recommend you take advantage of it! Our Assignment "0" goes over details on installation of Copilot into RStudio.

Course Schedule

We are likely to move slower this term with the course modernization. I have left a bit of wiggle room in the schedule accordingly.

Session #	Date	Lecture / Readings *All case studies should be read before class*	Assignments Due
1	Sep. 8	**Understanding the Determinants of Health Outcomes: DAGs and Potential Outcomes** • SC Introduction, Chapters 2-3 (p. 96-148) • Imbens, G. W. (2020). "Potential outcome and directed acyclic graph approaches to causality: Relevance for empirical practice in economics". *Journal of Economic Literature*, *58*(4), 1129-1179.	
2	Sep. 15	**Coding/Probability Review, Univariate OLS** • SC, Chapter 1 (p. 16-58) • HK, Chapter 4	
3	Sep. 22	**Regression I: Setup & Inference** • SC, Chapter 1 (pg. 63-95) • HK, Chapter 12-13 (pg. 175-200) *Case Study:* Freedman, S., Goodman-Bacon, A., & Hammarlund, N. (2021). "Observational studies of the effect of Medicaid on health: controls are not enough." *Journal of Labor Economics, 39*(S2), S619-S650.	
4	Sep. 29	**Regression II: Additional Topics** • HK, Chapter 13 (pg. 201-266; can omit pages 226-232) *Case Study:* Michaelsen, M. M., & Tolan, S. (2012). Children at Risk: The Effect of Crop Loss on Child Health in Rural Mexico. *Ruhr Economic Paper*, (376).	• **Assignment 1 due**
5	Oct. 6	**Data Cleaning and Research Designs** • "Everything is a Remix" (YouTube video, ~30 minutes)	
N/A	Oct. 13	**No class, Fall Reading Week**	
6	Oct. 20	**Matching Methods** • SC, Chapter 4 (p.175-240) • HK, Chapter 14 *Case Study:* Shahidi, F. V., Muntaner, C., Shankardass, K., Quiñonez, C., & Siddiqi, A. (2019). The effect of unemployment benefits on health: a propensity score analysis. *Social Science & Medicine*, *226*, 198-206.	**Assignment 2 due**
N/A	Oct. 27	**Final date to drop Fall session courses without academic penalty**	
7	Oct. 27	**Instrumental Variables Regression** • HK, Chapter 19 (also chapter 9 if you're interested in front-door approaches) • SC, Chapter 7 (p. 315-385) *Case Study:* Eliason, P., Heebsh, B., League, R., McDevitt, R., & Roberts, J. (2022). The effect of bundled payments on provider behavior and patient outcomes: evidence from the dialysis industry. *R&R, American Economic Revie*	**Research Question + Teams due**

8	Nov. 3	**Limited Dependent Variables I: Binary and Count Data** • HK, pages 226-232 *Case Study:* Bruhin, A., Goette, L., Haenni, S., & Jiang, L. (2020). Spillovers of prosocial motivation: Evidence from an intervention study on blood donors. *Journal of health economics*, *70*, 102244.	**Assignment 3 due**
9	Nov. 10	**Limited Dependent Variables II: Regression Specifications and Identification** *Case Studies (2!):* - Buchmueller, T. C., Jacobson, M., & Wold, C. (2006). How far to the hospital?: The effect of hospital closures on access to care. *Journal of health economics*, *25*(4), 740-761. - Kreider, B., Manski, R. J., Moeller, J., & Pepper, J. (2015). The effect of dental insurance on the use of dental care for older adults: a partial identification analysis. *Health economics*, *24*(7), 840-858. **Optional reading:** Hoagland (2024). "Innovations and Inequities in Access to Medical Services" Working paper.	**Paper Outline due**
10	Nov. 17	**Differences-in-Differences** • HK, Chapter 18 • SC, Chapter 8-9 (p. 406-510) - *Case Study:* Kessler, J. B., & Roth, A. E. (2014). *Don't take 'no' for an answer: An experiment with actual organ donor registrations* (No. w20378). National Bureau of Economic Research.	
11	Nov. 24	**Catch up Day as needed**	
12	Dec. 1	**Class Presentations** • Please review each other's slides prior to class	
13	Dec. 8 (if needed)	**Class Presentations** • Please review each other's slides prior to class	**Assignment 4 due**
		--	
N/A	Dec. 19		**Final Paper Due**

If time permits (it likely won't), we will cover:
1. **Synthetic Control Designs**
   - SC, Chapter 10 (p. 511-539)
   - HK, Section 21.2.1 (pages 555-559)
2. **Machine Learning for Quantitative HSR**
   - Athey, S., & Imbens, G. W. (2019). "Machine learning methods that economists should know about." *Annual Review of Economics, 11*(1), 685-725.