Quantitative Methods in Health Systems Research I

HAD 5744
Winter 2026

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

Class time and location: Mondays, 11am-1pm. **HS 740** (155 College St, 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.
- 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 submitted as a group of no more than 5 students.

- 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 come from recorded video presentations 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

[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!

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.

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.

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 text: "The Effect," Nick Huntington-Klein (NHK). There is a free e-book version or it's a relatively cheap buy on Amazon. I *highly* recommend this book and the SC book below if you plan to pursue applied research – they are both very handy. **Course readings are expected to be completed prior to class.**

Additional (non-required) reading:

- "Causal Inference: The Mixtape," Scott Cunningham (SC) (also free e-version – highly recommend, although a bit mathier! (Optional) readings are listed on syllabus)
- 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! Assignment "0" goes over installation details.

Course Schedule

Session #	Date	Lecture / Readings *All readings + case studies should be read before class*	Assignments Due
1	Jan. 5	**Understanding the Determinants of Health Outcomes: DAGs and Potential Outcomes** • 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. *Optional:* SC Introduction, Chapters 2-3 (p. 96-148)	
2	Jan. 12	**Regression I: Univariate Ordinary Least Squares** • HK, Chapter 4 • Khullar, D., & Jena, A. B. (2021). "Natural experiments" in health care research. *JAMA Health Forum* *Case Studies:* - Worsham, C., Bray, C., & Jena, A. (2024). Halloween, ADHD, and Subjectivity in Medical Diagnosis (No. w33232). *Optional:* SC, Chapter 1 (p. 16-58)	
3	Jan. 19	**Regression II: Multivariate Regression Setup & Inference** • HK, Chapter 12-13 (pg. 175-200) *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). *Optional:* SC, Chapter 1 (pg. 63-95)	
4	Jan. 26	**Regression III: Regressions in Practice** • HK, Chapter 13 (pg. 201-266; can omit pages 226-232) *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.	**Assignment 1 due**
5	Feb. 2	**Data Cleaning and Research Designs** • "Everything is a Remix" (YouTube video, parts 1 through 3 only, ~15 minutes)	
6	Feb. 9	**Matching Methods** • 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. *Optional:* SC, Chapter 4 (p.175-240)	**Assignment 2 due**
N/A	Feb. 16	**No class, Reading Week**	
7	Feb. 23	**Instrumental Variables Regression** • HK, Chapters 9 and 19 *Case Study:* Wong, K., Campitelli, M. A., Stukel, T. A., & Kwong, J. C. (2012). Estimating influenza vaccine effectiveness in community-dwelling	**Research Question + Teams due**

		elderly patients using the instrumental variable analysis method. *Archives of internal medicine,* 172(6), 484-491. *Optional:* - SC, Chapter 7 (p. 315-385) - Eliason, P., Heebsh, B., League, R., McDevitt, R., & Roberts, J. (2026). The effect of bundled payments on provider behavior and patient outcomes: evidence from the dialysis industry. *American Economic Review*	
N/A	Feb. 27	**Final date to drop course without academic penalty**	
8	Mar. 2	**Differences-in-Differences** • HK, Chapter 18 *Case Study:* Dranove, D., Gaynor, M., & Geddes, E. (2025). *Expecting Harm? The Impact of Rural Hospital Acquisitions on Maternal Health Care* (No. w34159). National Bureau of Economic Research. *Optional:* SC, Chapter 8-9 (p. 406-510)	**Assignment 3 due**
9	Mar. 9	**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.	**Paper Outline due**
10	Mar. 16	**Limited Dependent Variables II: Regression Specifications and Combining Identification Strategies** *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. - Jacob, B. J., Sutradhar, R., Moineddin, R., Baxter, N. N., & Urbach, D. R. (2013). Methodological approaches to population based research of screening procedures in the presence of selection bias and exposure measurement error: colonoscopy and colorectal cancer outcomes in Ontario. *BMC medical research methodology*, *13*(1), 59.	
11	Mar. 23	**1. Synthetic Control Designs** • HK, Section 21.2.1 (pages 555-559) • *Case Study:* Geloso, V., & Pavlik, J. B. (2021). The Cuban revolution and infant mortality: A synthetic control approach. *Explorations in Economic History*, *80*, 101376. **2. Machine Learning for Quantitative HSR** • *Case Study/Reading:* Athey, S., & Imbens, G. W. (2019). "Machine learning methods that economists should know about." *Annual Review of Economics, 11*(1), 685-725. *Optional:* SC, Chapter 10 (p. 511-539)	**Assignment 4 due**
12	Mar. 30	**Class Presentations** • Please review each other's slides prior to class	
13	Apr. 6 (if needed)	**Class Presentations** • Please review each other's slides prior to class	
N/A	Apr. 17		**Final Paper Due**