METHODS IN HEALTH SERVICES RESEARCH

HPM 583 (Spring 2017)

Instructor: Xi Chen, Assistant Professor of Health Policy and Economics
203-737-3556; xi.chen@yale.edu

Instructor Office Hours: LEPH 301, Thursday 12:00-1:00pm and by appointment

Teaching Assistants:
Diana Gonzalez Garcia (diana.gonzalez@yale.edu)
Office Hours: Tuesdays 12:00-1:00pm, Computer Lab @ 47 College St

Maya Mahin (maya.mahin@yale.edu)
Office Hours: Mondays 9:00-10:00am, Computer Lab @ 47 College St

Lecture Day/Time: Mondays, 10:00-11:50am, Winslow Auditorium

Lab Sessions Day/Time:1 Wednesday, 5:00-6:20pm
Thursday, 1:00-2:20pm

Course Description:
This is a course about the concepts, tools and approaches used to answer questions empirically. Students will learn a set of statistical tools and research designs that are useful in conducting good empirical research on health policy and management topics. We will consider four types of questions. Description: what is the value (or distribution) of outcome variable Y? Association: does Y vary by explanatory variable X? Evaluation: what is the causal effect of X on Y? Forecasting: what will Y be in the future or under some alternate scenario?

The course will emphasize the importance of research design for the identification of causal effects and the limitations in the applicability of commonly used techniques. We will illustrate five most well-received econometric methods, i.e., random assignment, regression, instrumental variables, regression discontinuity designs, and differences-in-differences through well-crafted real-world examples. For example, randomized experiments help answer whether health insurance makes us healthier, regression discontinuity designs reveal whether early life health interventions benefit academic performance, instrumental variables help examine whether money could relieve depression, differences-in-differences help test whether legal access to alcohol affects death rates, and so on.

1 Each week two lab sessions will be held in the Computer Lab @ 47 College St. Please click the online spreadsheet to sign up for a lab session and stick to it for ALL remaining lab assignments – no switching between sessions without prior approval of the TFs. The maximum number of students per lab session is 25 students.
https://docs.google.com/spreadsheets/d/17XhfnyGLIKMLgc2R_o9B9-yyafIDus9dwysOVo1MV1U/edit#gid=0
The pedagogical philosophy behind the class is colorfully summarized by Ed Leamer (“Lets take the Con out of Econometrics”, *American Economic Review* 1992):

This rhetoric is understandably tiring. Methodology, like sex, is better demonstrated than discussed, though often better anticipated than experienced.

Accordingly, while we will learn the statistical properties of a variety of common estimators, the formal requirements of the class will consist primarily of problem sets and a group research project requiring students to analyze real data in the hope that doing applied work will help you to learn the theory behind it.

**Prerequisites:**

Students should be familiar with basic matrix algebra and basic statistics/econometrics. The formal requirements for the class is one semester of biostatistics or permission of the instructor.

**Readings:**

You can get away without purchasing any books for this class, but you will learn a lot more if you work through Angrist and Pischke’s introductory level or more advanced level textbooks as we cover individual topics. The two textbooks present the essential tools of quantitative research and demonstrate why learning quantitative methods is exciting and useful. These textbooks will also be excellent references if you see yourself doing any data analysis at all in the future. Allison’s textbook covers additional introductory topics. If you’re going to continue writing SAS programs, I highly recommend the SAS book. The articles assigned are either overviews of methods or good examples of methods in action. All the books will be on reserve at the medical library and the articles are posted to classes*v2.


- articles (available online through classes*v2)
Assignments:

Lab assignments. There will be 6 lab homework assignments involving SAS/STATA programming and application of the material taught. No labs will be accepted for credit after the due date. The lab answers will be posted in a timely fashion. Labs will each be graded on a 0 to 4 point basis, and all labs should be turned in via the Classes V2 website.

In-class quizzes. The first quiz will be given on March 6 and the second on April 17. The quizzes will cover material presented in lecture and lab. No SAS/STATA programming issues will be covered in the two quizzes.

Group research project. In randomly-assigned groups, students will conduct a multi-step research project that consists of: (1) developing a health-related (broadly considered) research question, (2) cleaning a repeated cross-sectional survey data collected in the last two years or an external dataset (with the permission of the instructor) and creating analytic variables, (4) analyzing the data (using SAS/STATA), (5) presenting the project in class, and (6) writing up a research report that describes the question, analysis, results, and strengths and limitations of the approach used. Part of the lab help sessions on 4/12, 4/13, 4/19, 4/20 will be devoted to discussions of this assignment. Write-ups are due by 5pm on Friday, 5/5.

Requirements and Grading:

- Lab assignments (4x6) 24%
- In-class quizzes (20x2) 40%
- Group research project 30%
- Class participation 6%

Honor Code: All students are expected to comply with the Yale School of Public Health Code of Academic and Professional Integrity.

Laptop Use Policy: Laptops (and other computing devices) may be used only for class-related purposes, and laptop users are encouraged to sit near the back of the room to minimize class disruption.
LECTURE AND LAB SCHEDULE

FRI 1/20: Course introduction and overview
FRI 1/20: Lab instruction: SAS/STATA fundamentals Part I
* HW #1 assigned on Friday 1/20, due by Monday 1/30

MON 1/23: Survey design, Introduction to regression analysis
MON 1/23: Lab instruction: SAS/STATA fundamentals Part II
WED&THU help sessions for HW #1 (basic data analysis I) and HW #2 (basic data analysis II)
* HW #2 assigned on Monday 1/23, due by Monday 2/6

MON 1/30: Linear regression
MON 1/30: Lab instruction: Linear regression (SAS/STATA)
WED&THU help sessions for HW #2 (basic data analysis II) and HW #3 (linear regression)
* HW #3 assigned on Monday 1/30 due by Monday 2/13

MON 2/6: Logistic regression
MON 2/6: Lab instruction: Logistic regression (SAS/STATA)
WED&THU help sessions for HW #3 (linear regression) and HW #4 (logistic regression)
* HW #4 assigned on Monday 2/6 due by Monday 2/20

MON 2/13: Causality and evaluation
MON 2/13: Group project help session (I)
WED&THU help sessions for HW #4 (logistic regression)

MON 2/20: Model building and selection
MON 2/20: Lab instruction: Model building and selection (SAS/STATA)
WED&THU help sessions for HW #5 (model building and selection)
* HW #5 assigned on Monday 2/20 due by Friday 3/3

MON 2/27: Model-based prediction and policy simulation
MON 2/27: Quiz #1 review
WED&THU help sessions for HW #5 (model building and selection)

MON 3/6: QUIZ #1 (covers survey design, regressions, model building and selection)
NO LAB SESSION
NO WED&THU HELP SESSIONS

SPRING BREAK (3/11 – 3/26)

MON 3/27: Differences-in-differences
MON 3/27: Lab instruction: Differences-in-differences (SAS/STATA)
WED&THU help sessions for HW #6 (DID and IV)
* HW #6 assigned on Monday 3/27 due by Monday 4/10
MON 4/3: Instrumental variables
MON 4/3: Lab instruction: Instrument Variables (SAS/STATA)
WED&THU help sessions for HW #6 (DID and IV)

MON 4/10: Regression Discontinuity Designs
MON 4/10: Quiz #2 review
WED&THU help sessions for group project (I)

MON 4/17: QUIZ #2 (causality and evaluation, DID, IV, RDD)
WED&THU help sessions for group project (II)

MON 4/24: Student presentations of group research project
WED&THU: TBA
*Final group research project write-up is due by 5pm on Friday 5/5
COURSE READINGS BY TOPIC (subject to change)

I have attempted to provide a number of readings under each topic that might aide your understanding. I neither require nor expect you to get through all of this. Where I list many readings, I note the more important readings with a “*”.

1. Course introduction and overview (1/20)


2. Survey design (1/23)


3. Linear regression (1/23, 1/30)


4. **Logistic regression (2/6)**


5. **Causality and evaluation (2/13)**


6. **Model building and selection (2/20)**

7. Model-based prediction and policy simulation (2/27)


SPRING BREAK


Dowd B, Town R. Does x really cause y? Changes in Health Care Financing and Organization brief, 2002. (pp. 1-10)


9. Instrumental Variables (4/3)


10. **Regression Discontinuity Designs (4/10)**


11. **Further readings**