

Epi 268 Econometric Methods for Causal Inference

Instructor: Justin White, PhD
 E-mail: Justin.White@ucsf.edu
 Phone: (415) 276-8045
 Office hours: E-mail to set up meeting

Class time: Fridays, 1:30 – 4:00pm (2¼ hours, plus 15-min break)

Class location: Mission Hall, Room 1407

Units: 2 or 3

Course description

Epidemiologists and clinical researchers are increasingly seeking to estimate the causal effects of health-related policies, programs, and interventions. Economists have long had similar interests, and have developed and refined methods to estimate causal relationships. Examples include difference-in-differences, instrumental variables, and regression discontinuity. This course introduces a set of econometric tools and research designs in the context of health-related questions. The course topics are especially useful for evaluating natural experiments—situations in which comparable groups of people are exposed or not exposed to conditions determined by “nature” (not by a researcher), as occurs with a government policy or a disease outbreak.

Prerequisites: Biostat 200, Biostat 208, Epi 203, or equivalent experience. Experience with Stata.

Objectives

- Identify opportunities for rigorous evaluation of natural experiments
- Enumerate the (oft unstated) assumptions involved in interpreting research that aims to draw causal inferences
- Implement the research techniques covered in class

Requirements

Grades will be based on:

Item	Description	% grade, 2 units	% grade, 3 units
Weekly quizzes	6 quizzes	40%	20%
Problem sets	2 problem sets	50%	25%
Discussion leader	1 journal article	10%	5%
Empirical paper	1 paper	-	40%
Presentation	1 presentation	-	10%

Weekly Quizzes

Quizzes will be completed online, typically due seven days after each class. Late quizzes will not be graded except in extraordinary circumstances. The main purpose of the quizzes is to ensure that you have completed and understood the content of the assigned readings and in-class lecture for that week. You may *not* work on the quizzes with others.

Problem Sets

There are two problem sets, each due at the beginning of class on the due date. Late assignments will not be graded except in extraordinary circumstances. Assignments may be handwritten or typed, but must be legible. You can work with others, but each student must turn in their own assignment, and must acknowledge intellectual contributions of anyone who has assisted them.

Discussion Leader

Each student will be responsible for facilitating class discussion about one assigned reading. The student will be graded on his/her ability to stimulate discussion about the strengths and limitations of the paper, with focus on the paper's strategy for identifying causal effects, analytic framework, and presentation of results. The discussion leader should schedule a meeting with the instructor prior to class, in order to review the article and key points for discussion.

Original Empirical Paper (3-unit option only)

The paper will be based on a topic of your choice, using data of your choice, and should apply a technique covered in this class. You are encouraged to discuss your paper with others, but as always you must acknowledge any intellectual contributions made by others. See the Box course folder for more detailed guidelines. Extra class sessions will be offered to help students in the 3-unit option to develop an empirical paper. A one-page proposal for the paper is due by email on October 27. The paper is due by email on December 14.

Readings

Each week you will be assigned one or more textbook chapters or a summary article and an empirical paper that applies the method under study. You are encouraged to complete these readings *prior* to class on the day they are listed.

Required Textbook

- Joshua Angrist and Jörn-Steffen Pischke (2015). Mastering 'Metrics: The Path from Cause to Effect. Princeton University Press.

Additional Textbooks

- Guido W. Imbens and Donald B. Rubin (2015). Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction. Cambridge University Press. (The most thorough and current statement of potential outcomes models for causal inference.)
- Joshua Angrist and Jörn-Steffen Pischke (2009). Mostly Harmless Econometrics. Princeton University Press. (The more mathematically advanced companion to Mastering 'Metrics.)
- Jeffrey M. Wooldridge (2013). Introductory Econometrics: A Modern Approach, 5th Edition. South-Western College Publishers. (A thorough, introductory treatment of a broad range of econometric applications.)
- Peter Kennedy (2008). A Guide to Econometrics, 8th Edition. Wiley-Blackwell. (Intuitive feel for econometric concepts, alongside more technical discussion.)
- A. Colin Cameron and Pravin Trivedi (2009). Microeconometrics Using Stata. Cambridge University Press. (A useful, though aging, primer on using Stata for econometrics; the book also has parallel content to their econometrics textbook.)

- Partha Deb, Edward C. Norton, Willard G. Manning (2017). Health Econometrics Using Stata. Stata Press. (Overview of a variety of econometric modeling approaches in Stata, including many not covered in this course, e.g., GLM, count models, and a mass at zero.)

Guidelines for Requesting Accommodations and Leave

If you wish to be evaluated for classroom accommodations due to a disability, medical condition, or pregnancy-related medical condition, please contact the Student Disability Services Office (sds.ucsf.edu) or HR Disability Management (hr.ucsf.edu) to review the process for requesting accommodations. If you have been approved for accommodations, please request a letter be sent to me detailing the accommodations. For parental- or pregnancy-related accommodations and leave, please contact your program director to develop a plan with respect to coursework.

Course Schedule

Sept 29: Causal Inference and Regression

Topics: course introduction, RCTs as gold standard, potential outcomes framework, conditional expectation function, regression notation, regression mechanics, interpreting coefficients and effect sizes

Required reading: MM, Intro, Ch. 1, 2

Further reading: Holland (1986); Imbens, Rubin (2015), Ch. 1-3; Rubin (1986)

Oct 6: Selection on Observables Designs and Stata Programming

Topics: regression adjustment, omitted variable bias, nonparametric regression, introduction to Stata, Monte Carlo simulation (bias analysis)

Required reading: Freedman (1991); Baum (2011); Baum (2007)

Further reading: Arnold et al. (2011); Kennedy (2008), Ch. 2

Oct 13: Difference-in-Differences I

Topics: introduction to panel data, difference-in-differences analysis, DiD assumptions, clustering standard errors in DiD

Required reading: MM, Ch. 5; Rajaram et al. (2014)

Further reading: Bertrand, Duflo, Mullainathan (2004); Wherry, Miller (2016); Ryan, Burgess, Dimick (2015)

Oct 20: Difference-in-Differences II

Topics: DiD analyses in Stata, triple differences estimator, case studies with synthetic controls

Required reading: Joyce (2004); Rudolf et al. (2015)

Further reading: Abadie, Diamond, Hainmueller (2010); Basu et al. (2016); Gruber (1994)

Oct 27: **Fixed Effects**

Empirical paper proposal due

Topics: dummy variable version, first difference estimator, deviations from means estimator, FE vs. random effects, specification tests

Required reading: Allison (2009), Ch. 2; Cylus, Glymour, Avendano (2014)

Further reading: Colen, Ramey (2014); Currie, Thomas (1995); Frankenberg (1995); Wooldridge (2013), Ch. 13, 14

Nov 3: **Instrumental Variables**

Problem Set #1 due

Topics: IV intuition, IV assumptions, two-stage least squares, local average treatment effects, IV specification testing

Required reading: MM, Ch. 3; Hearst, Newman, Hulley (1986); Angrist (1990)

Further reading: Angrist, Imbens, Rubin (1996); Baiocchi, Cheng, Small (2014); Glymour, Tchechen Tchechen, Robins (2012); Harris, Remler (1998); Imbens (2010); McClellan, McNeil, Newhouse (1994); Pizer (2015)

Nov 17: **Regression Discontinuity Designs I**

Topics: RD graphical intuition, forcing variable, sharp RD, sharp RD estimation

Required reading: MM, Ch. 4.1; Bor, Moscoe, Bärnighausen (2015); Hansen (2015)

Further reading: Green et al. (2009);

Dec 1: **Regression Discontinuity Designs II**

Topics: Fuzzy RD, fuzzy RD estimation, RD specification testing, RD in Stata

Required reading: MM, Ch. 4.2; Lee, Lemieux (2010); Hamad, Cohen, Rehkopf (2016)

Further reading: Card, Dobkin, Maestas (2009); Imbens, Lemieux (2009)

Dec 8: **Student Presentations**

Problem Set #2 due

Dec 14: *Empirical paper due*

References

- Abadie A, Diamond A, Hainmueller J (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association* 105(490), 493–505.
- Allison, PD (2009). Fixed Effects Regression Models. SAGE Publications.
- Angrist JD (1990). Lifetime earnings and the Vietnam era draft lottery: Evidence from social security administrative records. *American Economic Review* 80(3): 313-336.
- Angrist JD, Imbens GW, Rubin DB (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association* 91(434), 444–455.
- Angrist J, Pischke JS (2009). Mostly Harmless Econometrics. Princeton University Press.
- Arnold BF, Hogan DR, Colford JM, Hubbard AE (2011). Simulation methods to estimate design power: An overview for applied research. *BMC Medical Research Methodology* 11(94): 1-10.
- Baiocchi M, Cheng J, Small DS. (2014). Instrumental variable methods for causal inference. *Statistics in Medicine* 33: 2297-2340.
- Basu S, Rehkopf DH, Siddiqi A, Glymour MM, Kawachi I. (2016). Health behaviors, mental health, and health care utilization among single mothers after welfare reforms in the 1990s. *American Journal of Epidemiology* 183(6): 531-538.
- Baum CF (2007). Monte Carlo Simulation in Stata. Presentation. July 2007.
- Baum CF (2011). Introduction to Stata. Presentation. August 2011.
- Bertrand M, Duflo E, Mullainathan S (2004). How much should we trust difference-in-differences estimates? *Quarterly Journal of Economics* 119(1): 249-275.
- Bor J, Moscoe E, Bärnighausen T (2015). Three approaches to causal inference in regression discontinuity designs. *Epidemiology* 26(2): e28-30.
- Card D, Dobkin C, Maestas N (2009). Does Medicare save lives? *Quarterly Journal of Economics* 124(2): 597-636.
- Colen CG, Ramey DM. Is breast truly best? Estimating the effects of breastfeeding on long-term child health and wellbeing in the United States using sibling comparisons. *Social Science & Medicine* 109 (2014): 55-65.
- Currie J, Thomas D (1995). Medical care for children: public insurance, private insurance, and racial differences in utilization. *Journal of Human Resources* 30(1): 135-162.
- Cylus J, Glymour MM, Avendano M (2014). Do generous unemployment benefit programs reduce suicide rates? A state fixed-effect analysis covering 1968-2008. *American Journal of Epidemiology* 180(1):45-52.
- Frankenberg E (1995). The effects of access to health care on infant mortality in Indonesia. *Health Transition Review* 5: 143-163.

- Freedman DA (1991). Statistical models and shoe leather. *Sociological Methodology* 21, 291–313.
- Glymour MM, Tchetgen Tchetgen EJ, Robins JM (2012). Credible Mendelian randomization studies: approaches for evaluating the instrumental variable assumptions. *American Journal of Epidemiology* 175(4): 332–9.
- Green DP, Leong TY, Kern HL, Gerber AS, Larimer CW (2009). Testing the accuracy of regression discontinuity analysis using experimental benchmarks. *Political Analysis* 17(4), 400–417.
- Gruber J (1994). The incidence of mandated maternity benefits. *American Economic Review* 84(3): 622–641.
- Hamad R, Cohen AK, Rehkopf DH (2016). Changing national guidelines is not enough: The impact of 1990 IOM recommendations on gestational weight gain among US women. *International Journal of Obesity*: 40(10), 1529–1534.
- Hansen B. (2015). Punishment and deterrence: Evidence from drunk driving. *American Economic Review* 105(4): 1581–1617.
- Harris KM, Remler DK (1998). Who is the marginal patient? Understanding instrumental variables estimates of treatment effects. *Health Services Research* 33(5): 1337–1360.
- Hearst N, Newman TB, Hulley SB (1986). Delayed effects of the military draft on mortality. *New England Journal of Medicine* 314(10): 620–624.
- Holland PW (1986). Statistics and causal inference. *Journal of the American Statistical Association* 81(396), 945–960.
- Imbens GW (2010). Better LATE than nothing: Some comments on Deaton (2009) and Heckman and Urzua (2009). *Journal of Economic Literature* 48: 399–423.
- Imbens GW, Lemieux T (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics* 142(2), 615–635.
- Imbens GW, Rubin DB (2015). Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction. Cambridge University Press.
- Joyce T (2004). Did legalized abortion lower crime? *Journal of Human Resources* 39(1): 1–28.
- Kennedy P (2008). A Guide to Econometrics, 8th Edition. Wiley-Blackwell.
- Lee DS, Lemieux T (2010). Regression discontinuity designs in economics. *Journal of Economic Literature* 48: 281–355.
- McClellan M, McNeil BJ, Newhouse JP (1994). Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables. *JAMA* 272: 859–866.
- Pizer SD (2015). Falsification testing of instrumental variables methods for comparative effectiveness research. *Health Services Research*. 51(2): 790–811.

Rajaram R, Chung JW, Jones AT, Cohen ME, Dahlke AR, Ko CY, Tarpley JL, Lewis FR, Hoyt DB, Bilimoria KY (2014). Association of the 2011 ACGME resident duty hour reform with general surgery patient outcomes and with resident examination performance. *JAMA* 312(22): 2374-2384.

Rubin DB (1986). Comment: Which ifs have causal answers. *Journal of the American Statistical Association* 81(396): 961–962.

Rudolph KE, Stuart EA, Vernick JS, Webster DW (2015). Association between Connecticut's permit-to-purchase handgun law and homicides. *American Journal of Public Health*, 105(8): e49-e54.

Ryan AM, Burgess Jr., JF, Dimick JB (2015). Why we should not be indifferent to specification choices for difference-in-differences. *Health Services Research* 50(4): 1211-1235.

Wherry LR, Miller S (2016). Early coverage, access, utilization, and health effects associated with the Affordable Care Act Medicaid expansions: A quasi-experimental study. *Annals of Internal Medicine* 164(12): 795-803.

Wooldridge JM (2013). *Introductory Econometrics: A Modern Approach*, 5th Edition. South-Western College Publishers.