Econometric Methods in Impact Evaluation
GHP228, Spring 2022, Fridays 8:00am-11:15am
Harvard T. H. Chan School of Public Health, Kresge 201

Instructor's Information

Jessica Cohen
Bruce A. Beal, Robert L. Beal and Alexander S. Beal Associate Professor of Global Health
Department of Global Health and Population
Email: cohenj@hsph.harvard.edu
Office Hours: Tuesdays 3pm-4:30pm (But must schedule appointment)
Office Address: HSPH Building 1, Room 1209 or Zoom
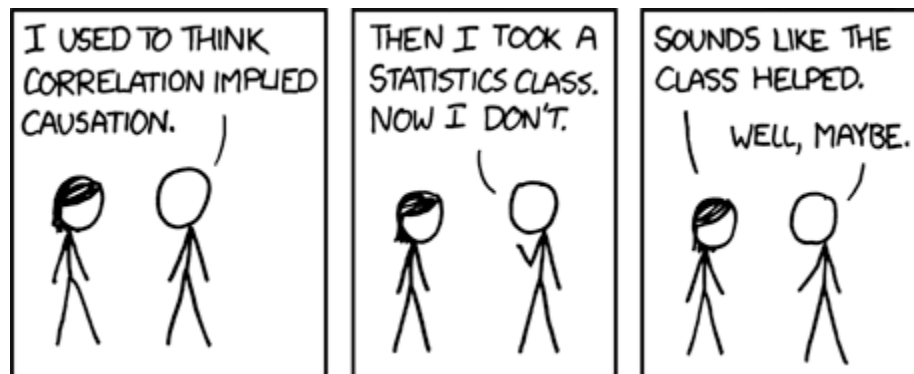
Teaching Assistant's Information

Daniel Arias
PhD Candidate
Department of Global Health and Population
Email: darias@g.harvard.edu
Office Hours: Wednesdays 2-3pm and by appointment
Office Hour location: Zoom (link on canvas)

Source: https://xkcd.com/552

Course Objectives

The objective of this course is to provide students with a set of theoretical, econometric, and reasoning skills to assess causality and impact. The course will introduce students to a variety of econometric techniques in impact evaluation and a set of reasoning skills intended to help them become both a consumer and producer of applied empirical research. Students will learn to both critically analyze evaluation research and gauge how convincing the research is in identifying a causal impact. They will use these skills to develop an evaluation plan for a topic of their own, with the aim of stimulating ideas for dissertation research.

Examples from the readings explore the causal effect of policies, laws, governmental and non-governmental programs, and "natural experiments" on health, education, poverty, and other outcomes. We will, for the most part, approach impact evaluation from an economics perspective and will discuss differences and similarities between how economists establish causality and how causality is established in the medical and public health fields. We will go beyond estimating causal effects to analyze the channels through which the causal impact was likely achieved, discussing the ways that economists think about measuring causal pathways.

This is a methods class that relies heavily on familiarity with econometrics and (at least a basic level of) microeconomics. These are pre-requisites for the course without exception. The course is intended for doctoral students who are finishing their course work or for master's students planning to pursue a doctoral degree, with the aim of helping them begin the transition into independent research.

At the end of the course the student will be able to:

- Understand and apply a variety of econometric methods for estimating impact, including randomized controlled trials and quasi-experimental designs ("natural experiments") including regression discontinuity designs, difference-in-differences, synthetic control, and interrupted time series.

- Critically analyze impact evaluation research in economics and public health and gauge the validity of causal estimates

- Understand evaluation design, including methods for designing randomized-controlled trials and how to spot a valid quasi-experimental design

- Learn how to develop meaningful hypotheses that are amenable to evaluation and test them using econometric techniques

What This Course is Not

I'd like for this course to cover all aspects of impact evaluation, but there is limited class time and you have limited time for assignments and so the focus is on reading, writing and critical thinking. *What the course focuses on is academic research in impact evaluation—how to pick it apart, how to know when a piece of research has credibly established causality, understand the benefits and limits to different approaches to evaluation, and get you thinking about your own research.*

Here is what it does *not* do:

- It does not cover the practical aspects of program implementation or evaluation in much detail, though I do try to share some of my own experiences doing evaluations. It's not a monitoring and evaluation or measurement class.

- It does not cover all of the statistical properties of the econometric estimators used in impact evaluation. We cover the basics, and the most commonly used procedures and fixes, but students wanting to apply these methods (well) will want to dig deeper.

- It does not build up your skills in statistical analysis software or build experience with programming.

Who Can Take This Course?

The aim of this course is to prepare doctoral students in the Population Health Sciences PhD Program and in the Health Policy PhD Program for the dissertation phase of their research and thus they will be given priority in enrollment. The course is also open to other graduate students in these programs and in other Harvard Chan departments, as well as doctoral students from other departments and schools, *conditional on having completed the pre-requisites and the course having enough space.*

Enrollment in the course will be capped at 12 students. Once space has been offered to those students for whom the class is required, slots will be offered based on students' level of preparedness for the course and opportunities to take it in the future. I also prioritize having a mix of backgrounds and perspectives in the class when possible.

Pre-Requisites

Econometrics and microeconomics are required for this course. While students can get by with just these two subjects, some previous experience with regression analysis and applied economic research will be advantageous. Students seeing applied regression analysis for the first time in this course may struggle with the reading.

Outcome Measures

Throughout this course, students will be asked to complete a range of different assignments. These include:

- **Weekly reading reviews:** Each week—except for those in which students have in-class presentations—we will be discussing an empirical paper that explores a particular econometric method in detail. To prepare for this discussion, students will be asked to write a reading review in advance of class. Details are provided in the "Reading and Discussion Guide" section of this syllabus. While reviews will only be graded for timely completion, the class participation grade will be lowered if reviews are consistently late or of poor quality (see "Criteria for Course Participation Evaluation" below)**.** Please upload the reading reviews to the designated space on Canvas no later than 8 AM on the day of class.

- **Randomized Controlled Trial (RCT) Proposal:** Following the completion of our module on RCT's, you will be asked to prepare a short paper proposing an RCT to evaluate the impact of an NGO or community-based program. Students will present their proposals in class. Details on this assignment will be provided separately.

- **Final Paper (submitted in four parts):** Throughout the course, students will develop a final paper based on a paper topic from the news, media, policy debate, or popular non-fiction. The topic can be from any field (economics, public health, sociology, criminology, anthropology, history...) but should be motivated by a news item, a policy question, a general curiosity, etc. rather than an academic text or paper. Health-related topics (interpreted broadly) are encouraged for PHS and Health Policy doctoral students, but other students are welcome to explore non-health outcomes. Students will develop a few questions on this topic that are amenable to impact evaluation techniques and then thoroughly describe how they would ideally analyze these questions (conceptual framework, data sets, sample, econometric specifications, etc.) and how they might go about analyzing it in practice.

  Students will write the paper in sections, submitted as three separate, shorter assignments throughout the course. Students will receive feedback as their work progresses and will be able to revise their work before submitting the final paper at the end of the term. All students are required to meet with the instructor and the teaching assistant to discuss these assignments and get feedback on their progress on the final paper overall.

Late Assignments: One third of a letter grade (e.g., from B+ to B) will be deducted for each day a written assignment is late (including the day it is due, if it is turned in past the appropriate time) unless permission for the late submission has been granted.

Assignment due dates

All assignments—except for reading reviews, which are due before class at 8AM—are due at midnight Eastern Time on the following days:

- March 3: Slides for RCT Proposal Presentation
- March 3: RCT Proposal (*Final version*)
- April 1: Part 1 (Introduction Section) of Final Paper
- April 15: Parts 2 and 3 of Final Paper (Conceptual Framework and Literature Review); *revisions to Part 1 are optional*
- April 28: Part 4 (Methods Section) of Final Paper (*revisions to Parts 1, 2 and 3 are optional*)
- April 28: Slides for Final Paper Presentations
- May 12: Final Paper, with all section completed

Written assignment should be submitted as Word Documents, while slides may be submitted either as .pdf or .pptx files. Submissions should be made on Canvas.

Grading Criteria

Grades will be given according to the following criteria:
- RCT Proposal: 20%
- Parts 1, 2-3, and 4 of Final Paper: 45% (15% each)
- Final Paper: 25%
- Preparation, reading reviews, class participation, and presentations: 10%

Participation

While the technical aspects of the readings will be presented in lecture format, the course will be heavily focused on discussion of the readings and will rely on student contributions to discussion. Discussion will be based largely on the questions raised in the "Reading and Discussion Guide" below but will frequently skip around and occasionally pursue somewhat off-topic ideas and critiques.

Students will be graded (10% of final grade) for the extent to which they meaningfully contribute to the critiques and ideas discussed in class and for the submission of reading reviews. **I expect each student to contribute to class discussion *at least* once or twice per session.** Criteria for evaluating participation is below in "Criteria for Course Participation Evaluation".

Absences

Our course relies heavily on participation and only meets once a week so absences should be avoided whenever possible. If you anticipate being absent more than once over the course of the semester, discuss this with the instructor at the beginning of the semester.

COVID-19 Contingency Plans

In the event that an exposure or infection means you are unable to attend in-person, please alert the teaching team that you will need accommodation—we will work with you to provide support and flexibility. If you feel well enough to participate remotely, we can arrange hybrid instruction as needed. In the event that I am not able to teach a session in person, we will meet remotely. If I am unable to teach at all, we will reschedule our session or rearrange the order of

sessions. In case of any scheduling changes, a message will be sent out on Canvas to alert students beforehand.

Texts and Reading Materials

The only **required** text for the class is:

> Angrist, Joshua D, and Jörn-Steffen Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press, 2009. (Noted as "MHE" in reading list below.)

Students are also **strongly recommended** to be familiar with:

> Wooldridge, Jeffrey M. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, Mass: MIT Press, 2010.

> Huntington-Klein, Nick. *The Effect*. 1st edition. Boca Raton: Routledge, 2021.

These are also two standard references:

> Handbook of Development Economics. Volumes 1-5.

> Handbook of Econometrics. Volumes 1-6.

Each class will draw on several technical and applied readings as specified below. Students are responsible for reading the required materials (marked with *) and are invited to read the optional readings for a broader understanding of each topic.

Session 1: Course Intro, Objectives, Expectations; Overview of Impact Evaluation
January 28, 2022

Topics Covered:

Counterfactuals and the fundamental problem of causal inference; selection/omitted variable bias/confounders; types of program evaluation; what is impact evaluation; case study: randomized trials relative to other methods to remove bias; potential outcomes framework ; types of randomization; types of impact evaluation; internal vs. external validity
How to read an economics paper.

Session 2: RCTs Part 1
February 4, 2022

Topics covered:

Regression and use of regression in potential outcomes framework; Average treatment effects; Treatment on the treated; Good controls vs. bad controls; Specifications and sub-group analysis; Stratification; Types of randomization; Clustering; Spillovers and choosing the level of randomization

Required Readings:

*Duflo, Glennerster, and Kremer. 2007. "Using Randomization in Development Economics Research: A Toolkit," Sections 2 & 3. *Handbook of Development Economics*, 4.

URL: https://search-proquest-com.ezp-prod1.hul.harvard.edu/docview/1698804581?accountid=11311&rfr_id=info%3Axri%2Fsid%3Aprimo

*Cohen, Jessica and Pascaline Dupas. 2010. "Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment." *Quarterly Journal of Economics.* 125(1), pp. 1–45.
[You can ignore the section on cost-effectiveness.]
URL: http://web.stanford.edu/~pdupas/CohenDupas.pdf

Optional Readings:

Shadish, Cook, and Campbell. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference,* Chapter 1, pp. 1–18. URL:
https://moodle2.units.it/pluginfile.php/132646/mod_resource/content/1/Estratto_ShadishCookCampbellExperimental2002.pdf

The New Yorker on the Poverty Action Lab RCT Approach:
https://www.newyorker.com/magazine/2010/05/17/the-poverty-lab

Development Impact Blog on Pre-Analysis Plans:
https://blogs.worldbank.org/impactevaluations/a-pre-analysis-plan-checklist

Development Impact Blog on Balance Tests in RCTs:
https://blogs.worldbank.org/impactevaluations/should-we-require-balance-t-tests-baseline-observables-randomized-experiments

Session 3: RCTs Part II
February 11, 2022

Topics covered:

Partial compliance in RCTs; Encouragement designs and "Intention to Treat (ITT)"; Use of instrumental variables in RCT analysis

Required Readings:

* MHE, Chapters 1, 2 & 3.2. URL:
https://www.researchgate.net/publication/51992844_Mostly_Harmless_Econometrics_An_Empiricist's_Companion

* Hanna, Rema, Esther Duflo and Michael Greenstone. 2016. "Up in Smoke: The Influence of Household Behavior on the Long-Run Impact of Improved Cookstoves." *American Economic Journal: Economic Policy*. 8(1), pp. 80-114.
URL: https://www.aeaweb.org/articles?id=10.1257/pol.20140008

Optional Readings:

Banerjee, Abhijit Vinayak, Esther Duflo, and Michael Kremer. 2016. "The Influence of Randomized Controlled Trials on Development Economics Research and on Development Policy." Published in Basu, Kaushik, David Rosenblatt, and Claudia Sepúlveda, eds. *The State*

of Economics, the State of the World. The MIT Press, 2020. URL:
https://economics.mit.edu/files/16485

Lant Pritchett's overview of RCT critiques: https://lantpritchett.org/rct/

Bryan, Kevin. "What Randomisation Can and Cannot Do: The 2019 Nobel Prize." *VoxEU.Org*
(blog), October 29, 2019. https://voxeu.org/article/what-randomisation-can-and-cannot-do-2019-
nobel-prize. Good summary by Kevin Bryan on Voxeu on pros and cons of RCTs.

Deaton, Angus. "Randomization in the Tropics Revisited: A Theme and Eleven Variations."
Working Paper. Working Paper Series. National Bureau of Economic Research, July 2020.
https://doi.org/10.3386/w27600.

Session 4: RCTs Part III
February 18, 2022

Topics covered:

Local Average Treatment Effects (LATE); External validity; Attrition; Drawbacks of RCTs;
Mechanism Experiments

Required Readings:

*Duflo, Esther, Rachel Glennerster, and Michael Kremer. 2007. "Using Randomization in
Development Economics Research: A Toolkit," Sections 4 - 6 & 8. *Handbook of Development
Economics*, 4. URL: http://www.sciencedirect.com.ezp-
prod1.hul.harvard.edu/science/article/pii/S1573447107040612

*Finkelstein, et al. 2012. "The Oregon Health Insurance Experiment: Evidence from the First
Year" *Quarterly Journal of Economics*, Vol. 127, Issue 3, pp. 1057-1106. URL: http://ezp-
prod1.hul.harvard.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=heh&A
N=79311104&site=ehost-live&scope=site

Optional Reading:

Overview of "Mechanism Experiments: Ludwig, Kling and Mullainathan. ""Mechanism
Experiments and Policy Evaluations". *Journal of Economic Perspectives*, 25 (3), pp: 17-38.
URL: https://www.aeaweb.org/articles?id=10.1257/jep.25.3.17

Continuation of the debate regarding the value of RCTs: Imbens, Guido. 2010. "Better LATE
Than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009)." *Journal of
Economic Literature* 48(2), pp. 399-423. URL: http://pubs.aeaweb.org.ezp-
prod1.hul.harvard.edu/doi/pdfplus/10.1257/jel.48.2.399

Some perspectives on the ethics of clinical trials and public policy RCTs:
Nardini, Cecilia. "The Ethics of Clinical Trials." *Ecancermedicalscience* 8 (January 16,
2014): 387. URL: https://doi.org/10.3332/ecancer.2014.387.

MacKay, Douglas. "The Ethics of Public Policy RCTs: The Principle of Policy
Equipoise." *Bioethics* 32, no. 1 (2018): 59–67. URL: https://onlinelibrary-wiley-com.ezp-
prod1.hul.harvard.edu/doi/full/10.1111/bioe.12403.

Session 5: Difference-in-Differences Part I
February 25, 2022

Topics covered:

Introduction to fixed effects; introduction to difference-in-differences; regression specifications for D-D; D-D as IV

Required Readings:

*MHE Sections 5.1-5.3 (inclusive)

*Duflo, Esther. 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment," *American Economic Review*, 91(4), pp. 795–813. URL: http://www.jstor.org.ezp-prod1.hul.harvard.edu/stable/2677813

Optional Readings:

Ryan, et al. 2015. "Why we Should not be Indifferent to Specification Choices for Difference-in-Differences", Health Services Research.50(4): 1211-1235.
https://doi.org/10.1111/1475-6773.12270

Daw and Hatfield 2018. "Matching and Regression to the Mean in Difference-in-Differences Analysis" Health Services Research 53(6): 4111-4117.
Link = https://onlinelibrary.wiley.com/doi/abs/10.1111/1475-6773.12993

Rokicki S, et al. 2018. "Inference with Difference-in-Differences with a Small Number of Groups: A Review, Simulation Study, and Empirical Application Using Share Data." Medical Care, 56(1): 97-105.
Link = https://www.jstor.org/stable/23015923?seq=1

Session 6: RCT Proposal Presentations
March 4, 2022

Students will present their RCT proposals in groups

Session 7: Difference-in-Differences Part II: Two Way Fixed Effects and Event Study
March 11, 2022

Topics covered:

Extensions of difference in differences: event study models, triple differences, multi-period D-D

Required Readings:

*Gruber J, and Kleiner SA. "Do Strikes Kill? Evidence from New York State." *American Economic Journal: Econ Policy*. 2012;4(1):127-157. doi:10.1257/pol.4.1.127. URL: https://www-aeaweb-org.ezp-prod1.hul.harvard.edu/articles?id=10.1257/pol.4.1.127

*Jakiela, Pamela. 2019. What Are We Estimating When We Estimate Difference-in-Differences? World Bank Blogs. URL: http://blogs.worldbank.org/impactevaluations/what-are-we-estimating-when-we-estimate-difference-differences

* McKenzie, David. 2022. A New Synthesis and Key Lessons from the Recent Difference-in-Differences Literature. World Bank Blogs. URL:

https://blogs.worldbank.org/impactevaluations/new-synthesis-and-key-lessons-recent-difference-differences-literature?cid=SHR_BlogSiteShare_EN_EXT

Optional Readings:

Anttila-Hughes, Jesse K., Lia CH Fernald, Paul J. Gertler, Patrick Krause, and Bruce Wydick. Mortality from Nestlé's marketing of infant formula in low and middle-income countries. No. w24452. National Bureau of Economic Research, 2018.
URL: https://drive.google.com/file/d/1ZajT15VuV9bqVIKoXNrwmeG_56DiMl2D/view

Brooks, Nina, Eran Bendavid, and Grant Miller. "USA Aid Policy and Induced Abortion in Sub-Saharan Africa: An Analysis of the Mexico City Policy." *The Lancet Global Health* 7, no. 8 (August 1, 2019): e1046–53. https://doi.org/10.1016/S2214-109X(19)30267-0.

Baker, Andrew. 2019. "Difference-in-Differences Methodology."
https://andrewcbaker.netlify.com/2019/09/25/difference-in-differences-methodology/.

Baker, Andrew. 2019. "So You've Been Told to Do My Difference-in-Difference Thing: A Guide."
https://cdn.vanderbilt.edu/vu-my/wp-content/uploads/sites/2318/2019/10/09023516/so_youve_been_told_dd_10_9_2019.pdf

Kondylis, Florence and John Loeser. 2019. Econometrics Sandbox: Event Study Design & Co. World Bank Blogs. https://blogs.worldbank.org/impactevaluations/econometrics-sandbox-event-study-designs-co

Session 8: Interrupted Time Series
March 25, 2022

Topics covered:

Introduction to ITS; regression specification choices; ITS with comparator

Required Readings:

*Jayachandran, S, Lleras-Muney A, Smith K. "Modern Medicine and the Twentieth Century Decline in Mortality: Evidence on the Impact of Sulfa Drugs". *American Economic Journal: Applied Economics* 2(April 2010): 118-146.
http://faculty.wcas.northwestern.edu/~sjv340/sulfa.pdf

Optional Readings:

Bernal, James Lopez, Steven Cummins, and Antonio Gasparrini. "Interrupted Time Series Regression for the Evaluation of Public Health Interventions: A Tutorial." *International Journal of Epidemiology* 46, no. 1 (February 1, 2017): 348–55. https://doi.org/10.1093/ije/dyw098.

Humphreys, David K., Antonio Gasparrini, and Douglas J. Wiebe. "Evaluating the Impact of Florida's 'Stand Your Ground' Self-Defense Law on Homicide and Suicide by Firearm: An Interrupted Time Series Study." *JAMA Internal Medicine* 177, no. 1 (January 1, 2017): 44–50. https://doi.org/10.1001/jamainternmed.2016.6811.

Silaba, Micah et al. "Effect of 10-valent pneumococcal conjugate vaccine on the incidence of radiologically-confirmed pneumonia and clinically-defined pneumonia in Kenyan children: an interrupted time series analysis". *Lancet Global Health*. 7(3): PE337-E346.
https://www.thelancet.com/journals/langlo/article/PIIS2214-109X(18)30491-1/fulltext

Mehta, Shivan et al. 2015. "ACA-Mandated Elimination of Cost Sharing for Preventive Screening Has Had Limited Early Impact." *The American Journal of Managed Care,* 21(7), pp. 511-517. https://www.ajmc.com/journals/issue/2015/2015-vol21-n7/aca-mandated-elimination-of-cost-sharing-for-preventive-screening-has-had-limited-early-impact

Lu, Christine et al. 2014. "Changes in antidepressant use by young people and suicidal behavior after FDA warnings and media coverage: quasi-experimental study." *BMJ*, 348:g3596. https://www.bmj.com/content/348/bmj.g3596

Session 9: Synthetic Control Method (SCM)
April 1, 2022

Topics Covered:

Introduction to SCM; requirements and assumptions of SCM; inference in SCM; robustness and specification checks

Required Readings:

*Cunningham, 2019 "Causal Inference, The Mixtape – Chapter on Synthetic Controls. https://mixtape.scunning.com

* Barber, Andrew, and Jeremy West. "Conditional cash lotteries increase COVID-19 vaccination rates. Download Conditional cash lotteries increase COVID-19 vaccination rates." *Journal of Health Economics* 81 (2022): 102578.

Optional Readings:

Rokicki Slawa, 2021. "Impact of Family Law Reform on Adolescent Reproductive Health in Ethiopia: A Quasi-Experimental Study".*World Development*, Vol 144. URL: https://www.sciencedirect.com/science/article/abs/pii/S0305750X21000966

Alpert, Abby et al. 2018. "Supply-Side Drug Policy in the Presence of Substitutes: Evidence from the Introduction of Abuse-Deterrent Opioids," *American Economic Journal: Economic Policy*, 10(4), pp. 1-35. https://www.aeaweb.org/articles?id=10.1257/pol.20170082

Abadie, Alberto et al. 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), pp. 493-505. https://amstat.tandfonline.com/doi/abs/10.1198/jasa.2009.ap08746#.XBkXWFVKhaQ

Session 10: Regression Discontinuity
April 8, 2022

Topics Covered:

RD design assumptions and implementation; fuzzy vs. sharp RD; key robustness checks for RD; RD as an IV strategy

Required Readings:

* MHE Chapter 6

* Geiger, Caroline K., Mark A. Clapp, and Jessica L. Cohen. "Association of Prenatal Care Services, Maternal Morbidity, and Perinatal Mortality With the Advanced Maternal Age Cutoff of 35 Years." *JAMA Health Forum* 2, no. 12 (December 3, 2021): e214044. https://doi.org/10.1001/jamahealthforum.2021.4044.

Optional Readings:

Almond, et al. 2010. "Estimating Marginal Returns to Medical Care: Evidence from At-Risk Newborns." *Quarterly Journal of Economics*, 125(2), pp. 591–631. URL: https://www-jstor-org.ezp-prod1.hul.harvard.edu/stable/27867491

Alaouf, Mattan, Sarah Miller, and Laura R. Wherry. "What Difference Does a Diagnosis Make? Evidence from the Marginal Patients." No. w26363. National Bureau of Economic Research, 2019. https://www.nber.org/papers/w26363

Bor, Jacob, Matthew P. Fox, Sydney Rosen, Atheendar Venkataramani, Frank Tanser, Deenan Pillay, and Till Baernighausen. "Treatment eligibility and retention in clinical HIV care: a regression discontinuity study in south Africa." PLoSmedicine 14, no. 11 (2017). https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1002463

Angrist, Joshua and Victor Lavy. 1999. "Using Maimonides Rule to Estimate the Effect of Class Size on Scholastic Achievement." *Quarterly Journal of Economics*, 114(2), pp. 533–75. URL: http://www.jstor.org.ezp-prod1.hul.harvard.edu/stable/pdfplus/2587016.pdf?acceptTC=true

Dykstra, Sarah, Amanda Glassman, Charles Kenny, and Justin Sandefur. "Regression Discontinuity Analysis of Gavi's Impact on Vaccination Rates." *Journal of Development Economics* 140 (September 1, 2019): 12–25. https://doi.org/10.1016/j.jdeveco.2019.04.005.

Chapter 20 – Regression Discontinuity in Huntington-Klein, Nick. *The Effect*. 1st edition. Boca Raton: Routledge, 2021.

Session 11: Non-randomized Instrumental Variables
April 15, 2022
Conditions for valid instruments; Reduced form/First Stage; Exclusion restrictions; Weak instruments;

*Angrist/Pischke, MHE, Sections 4.1, 4.4.1-4.4.2

*Miguel, Edward & Shanker Satyanath & Ernest Sergenti, 2004. "Economic Shocks and Civil Conflict: An Instrumental Variables Approach," *Journal of Political Economy*, 112(4), pp. 725–53. URL: http://www.jstor.org.ezp-prod1.hul.harvard.edu/stable/3555136

De Neve, Jan-Walter, Günther Fink, S. V. Subramanian, Sikhulile Moyo, and Jacob Bor. "Length of secondary schooling and risk of HIV infection in Botswana: evidence from a natural experiment." The Lancet Global Health 3, no. 8 (2015): e470-e477. Link: https://www.thelancet.com/journals/langlo/article/PIIS2214-109X(15)00087-X/fulltext

Angrist, Joshua and Alan Krueger. 1991. "Does Compulsory Schooling Attendance Affect Schooling and Earnings?" *Quarterly Journal of Economics*, 106(91), pp. 976–1014. URL: http://www.jstor.org.ezp-prod1.hul.harvard.edu/stable/2937954

Wooldridge, Chapter 15 (p.510-529 in the most recent version; pp.484-503 in the older version)

Anderson, Michael and David Matsa. 2011. "Are Restaurants Supersizing America?" *American Economic Journal: Applied Economics* 3, pp. 152-188.
http://pubs.aeaweb.org.ezp-prod1.hul.harvard.edu/doi/pdfplus/10.1257/app.3.1.152

Garabedian, Laura et al. 2014. "Potential Bias of Instrumental Variable Analyses for Observational Comparative Effectiveness Research." *Annals of Internal Medicine*, 161(2), pp. 131-139. http://annals.org/aim/article-abstract/1887030/potential-bias-instrumental-variable-analyses-observational-comparative-effectiveness-research?doi=10.7326%2fM13-1887

Session 12: Considerations in Survey Design, Data Collection and Power Calculations
April 22, 2022

Topics Covered:

Considerations in sample size selection and power calculations; measurement error in surveys: recall bias, telescoping, surveys can influence outcomes, social desirability bias, list randomization and other approaches to measuring sensitive topics

Required Readings:

*Zwane, Zinman, Dusen, et al., 2011. "Being surveyed can change later behavior and related parameter estimates." *Proceedings of the National Academy of Sciences* 108(5): 1821-1826. URL: http://www.pnas.org.ezp-prod1.hul.harvard.edu/content/108/5/1821.full.pdf+html

* (Introduction) Ambler, Herskowitz, Maredia. 2021. "Are We Done Yet? Response Fatigue and Rural Livelihoods". Journal of Development Economics. Vol. 153.
URL: https://www.sciencedirect.com/sciencNotes e/article/pii/S0304387821001073

* (Introduction) Kerwin and Reynoso. 2021. "You Know What I Know: Interviewer Knowledge Effects in Subjective Expectation Elicitation". Demography. 58(1): 1-29. URL: https://read.dukeupress.edu/demography/article/58/1/1/167833/You-Know-What-I-Know-Interviewer-Knowledge-Effects

* (Introduction) Das, Hammer, Sánchez-Paramo. 2012. "The Impact of Recall Periods on Reported Morbidity and Health Seeking Behavior." Journal of Development Economics. 98(1), pp. 76-88. URL: https://www.sciencedirect.com/science/article/abs/pii/S0304387811000708

Optional Readings:

Celhay, Meyer, Mittag. 2022. "What Leads to Measurement Errors? Evidence from Reports of Program Participation in Three Surveys". NBER Working Paper # 29652
URL:
https://www.nber.org/papers/w29652?utm_campaign=ntwh&utm_medium=email&utm_source=ntwg13

Hickey, Graeme L, Stuart W Grant, Joel Dunning, and Matthias Siepe. "Statistical Primer: Sample Size and Power Calculations—Why, When and How?†." *European Journal of Cardio-Thoracic Surgery* 54, no. 1 (July 1, 2018): 4–9. https://doi.org/10.1093/ejcts/ezy169.

Sessions 13, 14, and 14.5: Student Presentations of Research Proposal
April 29, May 6, and May 10 (2:00 – 3:30 PM)

Session 15 Held for Spillover Topics and Input on Summer Impact Evaluation Coursework

May 13, 2022

Criteria for Course Participation Evaluation

	Exemplary (90%-100%)	Proficient (80%-90%)	Developing (70%-80%)	Unacceptable (<70%)
Frequency of Participation in Class	Student is always able to answer discussion questions when called on and initiates contributions several times in each class session.	Student is mostly able to answer discussion questions when called on and initiates contributions at least once in each class.	Student is able to answer discussion questions when called on half of the time and initiates contributions in half of the class sessions.	Student mostly is unable to answer discussion questions when called on and rarely initiates contributions in class sessions.
Listening/Attentiveness	Student listens attentively both during regular sessions and when others present materials and perspectives, and *regularly offers comments that build on the class discussion and others' remarks* (i.e. the student hears what others say and contributes to the dialogue)	Student is mostly attentive during class, including other students' presentations, and offers comments that build on others' remarks.	Student is sometimes inattentive and rarely makes comments based on others' contributions.	Does not pay attention in class or listen to others' presentations and thoughts. .
Quality of Comments	Responses to discussion questions and student-initiated contributions always indicate a careful readings of the assignments and are always insightful and constructive; uses appropriate terminology.	Responses to discussion questions and student-initiated contributions mostly indicate a careful readings of the assignments and are mostly insightful and constructive; mostly uses appropriate terminology.	Comments are sometimes constructive and informed, with occasional signs of insight. Student does not use appropriate terminology and struggles with concepts.	Comments do not reflect careful reading and are not constructive. Student does not use appropriate terminology. Comments are not relevant to discussion.
Reading Reviews	All reading reviews completed.	All except one reading review completed.	All except two reading reviews completed.	More than two missing reading reviews.

Reading and Discussion Guide: Quantitative Methods for Impact Evaluation (GHP 228)
Work in Progress; Suggestions Welcome

This guide is intended to help guide you through the causal arguments and presentation in each paper and to help facilitate in-class discussion. Since the papers we will read will take different empirical approaches and have variation in focus, some of the questions will be more appropriate for certain readings than others and your answers can reflect this. Feel free to make your responses as broad or detailed as you want—the main goal is to help prepare you for class discussion and ensure you have read the paper carefully.

Reading Reviews

For each of the required readings (but not for the methods/didactic texts such as MHE), you should prepare a 1-2 page reading review that addresses these questions and submit them to Canvas by 8:00a on the day of class. Your discussion of the paper does not need to address every question in this guide and you can feel free to focus more on certain questions than others and to only give overviews/broad descriptions. However, the more detailed your notes are on these questions, the more prepared you will be for in-class discussion. You do not need to write up notes about the tables and figures from tables/figures guide in your reading review. **You will not receive grades on these reviews and will only be marked for whether or not you turned them in. However, they will be briefly reviewed each week and if the review is consistently too thin and brief, we will get in touch with you (especially if your in-class participation reflects this).** No reading reviews will be due the weeks that there are in class presentations. Your feedback on the questions in this guide is very welcome. This is a work in progress.

In-Class Discussion

The in-class discussion of these papers is intended to help you understand in great detail how the authors investigate a causal question and how they use econometrics and rhetoric to convince the reader that the effect they have identified is indeed real. It is as much to help you be a top-notch producer of applied empirical research as it is to make you a clever reader of others' research. You should feel free to ask any question about the papers, no matter how minute or seemingly off-topic. Questions about why the authors present the data in such and such a way or perform a certain robustness check are particularly encouraged as are questions about the validity of the authors' arguments. While this isn't a program evaluation or M&E class, I have done a lot of field work and will try to answer your questions about how these interventions work in practice (e.g., how do you do the randomization in the field?) to the extent that I can and that there is class time.

The order in which we discuss the issues in this guide will vary somewhat by paper and by the flow of class discussion, so you should expect a lot of skipping around these questions, as well as some discussion of questions and comments that are somewhat off-topic if they seem useful. You should be prepared to be called on to discuss any of the questions below and to discuss any of the tables and figures in the paper (see Tables/Figures discussion guide below).

Discussion Questions

Background/Significance:
1) What is the motivation for this paper? What is the focus of inquiry (i.e. in a general way, what is the broad question of interest)? Does this question have policy relevance?

Overview:
2) What is the main causal question being asked in this paper?
   a. This paper estimates the impact of _____(X) on _____(Y).
   b. How are X and Y measured?
3) What is the basic empirical challenge that the paper faces in tackling this causal question?
   a. Let's say you know the association between X and Y from a large observational retrospective dataset. What are the sources of bias here (OVB/confounding, selection, etc.)?

Identification Strategy:
4) What is the identification strategy? (Overview)
   a. What is the general class of identification strategy (RCT, diff-in-diff, IV, etc.)?
   b. How does the paper propose to obtain an unbiased (or relatively unbiased) estimate of impacts? Describe either the intuition or "thought experiment"
   c. If the paper is an RCT, give an overview of the intervention. If it evaluates a policy, provide an overview of the policy.
5) What is the identification Strategy? (Technical)
   a. What are the main regressions for this strategy?
   b. What do each of the key regressors represent, and how are their coefficients interpreted?
   c. Who are the treatment and control group?
   d. Who are the compliers?
6) What is the exclusion restriction?
   a. What would cause the exclusion restriction to fail?

Data:
7) How is the dataset constructed?
   a. What kind of dataset is this (observational, experimental, etc.)?
   b. At what level(s) are the data measured (individual, household, village, etc.)?
   c. At what level(s) are the data grouped (household, village, school, etc.)?
   d. Are there issues with measurement error? Attrition?
   e. Do all variables actually capture the intended concept? E.g., does data on income adequately capture household consumption?

Findings:
8) What are the main findings?
   a. Interpret the magnitude of the coefficients.
   b. Does the magnitude seem reasonable? Does it seem like a meaningful effect size? (e.g., is it so small that it would never make a difference?)
   c. What do you think of how the main findings are presented? (e.g., are they only in tables and should be presented graphically? Are the tables hard to interpret?)
   d. If relevant (e.g., for IV or D-D): do you think the identification strategy is presented convincingly?

e. How do the findings compare to previous research (if relevant)? Are the differences between these findings and previous results what you would have expected?

Theory/Mechanisms:
9) Does the paper discuss its theory of change or proposed mechanism? That is, what is the proposed pathway from intervention to outcome?
    a. Is the proposed mechanism plausible?
10) Is the paper able to assess the theory or mechanism? For example, does it collect data on intermediate outcomes or behaviors? Or is it a black box where T goes in, Y goes out, and we don't really know why?
11) If the intervention is found to have a significant impact, what evidence is provided for why the program was successful? If the program was not successful, what evidence is provided for why the program was not successful?

Threats to internal validity:
12) What robustness checks (e.g., falsification tests, placebo tests) are used to assess threats to internal validity? Do you find them convincing?
    a. What specification checks are done?
13) What potential sources of bias remain? How realistic are these and how large of a source of bias might they be (e.g. would they change the impact estimates a lot or a little? Would they change the sign of the coefficient/direction of the impact?)?
    a. Can you think of any additional placebo/falsification tests that could have been done?

External validity:
14) Think about the program being evaluated, the method of intervention (e.g. NGO-run, government-run, etc.) and the population that is being considered.
    a. How representative are these of the sort of interventions that might address a similar problem and the sort of populations that might be affected by this intervention?
    b. Does the paper discuss the generalizability of the results? Do you find it convincing?

Improvements to the Study:
15) What would you have done differently?
    a. If you had the same or similar dataset, is there an alternative identification strategy? A better/more realistic intervention to analyze?
    b. Could you answer the same question with a different kind of data? Example: use better data on final or intermediate outcome. See whether results generalize in a different population.
    c. Would you have discussed limitations or advantages of the design differently?

Guide to table/charts and figures (NOTE: YOU DON'T NEED TO SUBMIT NOTES ON THESE IN YOUR READING REVIEW)
1) What results does the table/figure show?
    a. If this is the result of a regression equation, what is the regression equation (including controls, fixed effects, etc.)?
    b. How is the table/chart sub-divided?
2) Why is this table/chart in the paper?
    a. Are these the primary/secondary results? Robustness/falsification tests? Data description?
    b. What does this table/chart tell us about the identification strategy in the paper?
    c. Why are results sub-divided in this way?

3) What are the main take-away points from this table/chart?
   a. What are the most important variables here?
   b. Do the estimated values of the main variables change within the table/chart for different specifications?
   c. Can you interpret these differences?
4) Does this table/chart agree with the authors' interpretation of their results?
   a. Are there any results shown which seem to undermine the argument of the authors?
   b. Is there anything missing from this table/chart which you would have liked to see?
   c. Do the authors address these issues?
5) Is the design of the table/chart appropriate?
   a. If this table/chart stood alone, would you be able to understand it?
   b. Would you have organized it differently?