

Econometric Methods in Impact Evaluation
GHP228, Spring, Tuesdays 3:30pm-6:30pm
FXB G3

Course Site: <http://isites.harvard.edu/icb/icb.do?keyword=k104358&pageid=icb.page671067>

Instructors' Information

Jessica Cohen

Assistant Professor of Global Health
Global Health and Population
cohenj@hsph.harvard.edu

Office Hours: Thursdays 1pm-3pm (Let me know you are planning to come)

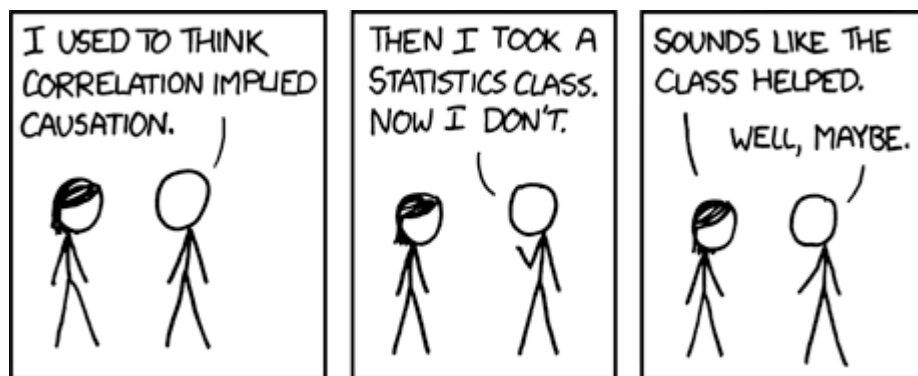
Office Address: 1209, Building 1

Teaching Assistant's Information

Noah Haber

nhaber@mail.harvard.edu

Office Hours: Thursdays, 4pm-6pm, 10th floor conference room (access via 11th floor), building 1



Source: <http://xkcd.com> (hat tip Chris Blattman & Michael Clemens)

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 critically analyze evaluation research and to 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, NGO programs and “natural experiments” on health, education, poverty, fertility 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 field. We will go beyond estimating causal effects to analyze the channels through which the causal impact was likely achieved. This will require that the students are comfortable with microeconomic theories of incentives, institutions, social networks, etc.

This is a methods class that relies heavily on familiarity with econometrics and microeconomics. These are pre-requisites for the course without exception. The course is intended for doctoral students who are finishing their course work and aims to help them 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, quasi-experimental designs (such as “natural experiments” and regression discontinuity designs), difference-in-differences 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 field trials
- 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 homework time 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 in field research. 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 health systems track of the GHP department for the dissertation phase of their research and thus they will be given priority in enrollment. Second year students in the Harvard FAS Health Policy PhD program (Evaluative Sciences and Statistics track) are also required to take the course and have priority. The course is also open to other GHP doctoral students, other GHP master's students and 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.

Pre-Requisites

Econometrics and intermediate micro-economics (i.e. graduate level economics with calculus) 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 a huge advantage.** Students seeing applied regression analysis for the first time in this course will most likely struggle with the reading.

Outcome Measures

Students will be given 4 short written assignments throughout the course of the semester and a 15 page final paper. Students will work in groups of 3-4 for the first short written assignment, and will be asked to propose a randomized trial to evaluate the impact of an NGO or government policy of interest to them. The other short written assignments will be done individually and will be geared toward helping

students develop their final paper. These will be drafts of different sections of the final paper. The final paper for the class will ask students to pick a paper topic from either the media 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 GHP and Health Policy GSAS students in the course. 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 actually go about analyzing it in practice. All students are required to meet with the instructors and the teaching assistant to discuss these short written assignments and get feedback on their progress on the final paper overall.

Weekly reading reviews on required papers (not technical texts) are required, except in weeks when papers are not assigned (such as weeks when students have in-class presentations). The purpose of these reviews is to guide you through the papers and to help prepare you for class discussion. Details are provided in the "Reading and Discussion Guide" below. These are only graded for whether or not they were submitted (on time), but they will be reviewed by the teaching staff and the class participation grade reflects their quality (see "Criteria for Course Participation Evaluation" below). If the reading reviews are consistently poor we will be in touch with you.

All written assignments and in-class presentation materials (slides) are due no later than 3pm on the day of class. One third of a letter grade (e.g. from 90 to 85) will be deducted for each day a written assignment is late.

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.

A list of student names in a randomly assigned sequence will be generated before each class and used to call on students to discuss the various questions in the guide. This encourages active learning and balanced participation. Students are also of course welcome to offer ideas and ask questions whenever they want.

Students will be graded (**15% 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. Criteria for evaluating participation is below in "Criteria for Course Participation Evaluation".

Absences

This 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.

Grading Criteria

Grades will be given according to the following criteria:

4 short written assignments: 40% (10% each)

Final paper: 30%

Class presentations: 15% (5% for presentation 1, 10% for presentation 2)

Preparation, reading reviews and class participation: 15%

Texts and Reading Materials

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

Angrist, Joshua D., and J. S. Pischke. 2008. "Mostly Harmless Econometrics: An Empiricists' Companion." Princeton, NJ: Princeton University Press. (Noted as "MHE" in reading list below)

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

Wooldridge, Jeffrey. 2002. (Or the 2nd edition from 2010) *Econometric Analysis of Cross-Section and Panel Data*. Cambridge: MIT Press. URL: http://nrs.harvard.edu/urn-3:hul.ebook:NLIB_78079

These are also two standard references:

Handbook of Development Economics. Volumes 1-5. URL: <http://www.sciencedirect.com.ezp-prod1.hul.harvard.edu/science/handbooks/15734471>

Handbook of Econometrics. Volumes 1-6. URL: <http://www.sciencedirect.com.ezp-prod1.hul.harvard.edu/science/handbooks/15734412>

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 recommended readings for a broader and deeper understanding of each topic.

I. Overview of Causal Inference (1 session); February 3, 2015

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

Shadish, Cook, and Campbell. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*, Chapter 1, pp. 1–18. URL: <http://depts.washington.edu/methods/readings/Shadish.pdf>

Banerjee, A. 2006. "Making Aid Work," *Boston Review*, July/August. URL: <http://bostonreview.net/BR31.4/banerjee.php>

The New York Times, March 9, 1993, Tuesday, Late Edition - Final, Section C; Page 1; Column 5; Science Desk, "Like a New Drug, Social Programs Are Put to the Test," By Peter Passell. URL: <http://www.nytimes.com/1993/03/09/science/like-a-new-drug-social-programs-are-put-to-the-test.html>

LaLonde, Robert J. 1986. "Evaluating the Econometric Evaluations of Training Programs Using Experimental Data," *American Economic Review* 76, 602–20. URL: <http://www.jstor.org.ezp-prod1.hul.harvard.edu/stable/1806062>

Martin Ravallion. 2001. "The Mystery of the Vanishing Benefits: An Introduction to Impact Evaluation." *World Bank Economic Review*, 15. Pp 2-29. URL: <http://wber.oxfordjournals.org.ezp-prod1.hul.harvard.edu/content/15/1/115.abstract>

Cook, Shadish, and Wong. 2008. "Three Conditions under Which Experiments and Observational Studies Produce Comparable Causal Estimates: New Findings from Within-Study Comparisons," *Journal of Policy Analysis and Management* 27(4), 724–50. URL: <http://onlinelibrary.wiley.com.ezp-prod1.hul.harvard.edu/doi/10.1002/pam.20375/abstract>

II. Randomized Trials (Part I: The Basics & RCT Implementation) (2 sessions); February 17, 2015 & February 24, 2015

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 standard errors; Spillovers and choosing the level of randomization

Session 1 Readings:

* MHE, Chapters 1, 2 & 3.2

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

URL: <http://www.sciencedirect.com.ezp-prod1.hul.harvard.edu/science/article/pii/S1573447107040612>

*Dupas, Pascaline. 2009. "Do Teenagers Respond to HIV Risk Information? Evidence from a Field Experiment in Kenya," *American Economic Journal: Applied Economics* 3(Jan), pp. 1-13.

URL: <http://pubs.aeaweb.org.ezp-prod1.hul.harvard.edu/doi/pdfplus/10.1257/app.3.1.1>

Ashraf, Nava, Field, Erica and Jean Lee. 2013. "Household Bargaining and Excess Fertility: An Experimental Study in Zambia" URL:

<https://www.dropbox.com/s/8lh855t8r5klxrv/Ashraf%202013.pdf?dl=0>

Olken, Benjamin. 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy*. 115(2). pp. 200–49. URL: <http://www.jstor.org.ezp-prod1.hul.harvard.edu/stable/10.1086/517935>

Session 2 Readings

*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. URL: <http://www.mitpressjournals.org.ezp-prod1.hul.harvard.edu/doi/abs/10.1162/qjec.2010.125.1.1>

*Bjorkman, Martina and Jakob Svensson. 2009. "Power to the People: Evidence from a Randomized Field Experiment on Community-Based Monitoring in Uganda." *Quarterly Journal of Economics* 124(2), pp. 735-769.

URL: <http://qje.oxfordjournals.org.ezp-prod1.hul.harvard.edu/content/124/2/735.short>

Dupas, Pascaline and Jonathan Robinson. 2013. "Why Don't the Poor Save More? Evidence from a Health Savings Experiment." *American Economic Review* 103(4): 1138-71.

http://www.stanford.edu/~pdupas/DupasRobinson_HealthSavings.pdf

Rodrik, Dani. 2009. "The New Development Economics: We Shall Experiment, but How Shall we Learn?" in *What Works in Development, Thinking Big and Thinking Small* (Cohen and Easterly eds.)

<http://www.hks.harvard.edu/fs/drodrik/Research%20papers/The%20New%20Development%20Economics.pdf>

III. Randomized Trials (Part II: RCTs as IV, Pros and Cons of RCTs, Ethics of RCTs) (2 sessions); March 3, 2015 & March 10, 2015

Partial compliance in RCTs; Encouragement designs and “Intention to Treat (ITT)”; Using instrumental variables in RCTs; Local Average Treatment Effects (LATE); External validity; Attrition; Drawbacks of RCTs; Mechanism Experiments; Next generation of RCTs

Session 1 readings:

*Duflo, Glennerster, and 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&AN=79311104&site=ehost-live&scope=site>

MHE, Section 4.4.3

Session 2 readings:

*Thornton, Rebecca. 2008. “The Demand for, and Impact of, Learning HIV Status.” *American Economic Review* 98(5), pp. 1829–63. URL: <http://pubs.aeaweb.org.ezp-prod1.hul.harvard.edu/doi/pdfplus/10.1257/aer.98.5.1829>

*Kling, et al. 2005. “Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment,” *Quarterly Journal of Economics*, Vol. 120, pp. 87–130. URL: <http://qje.oxfordjournals.org.ezp-prod1.hul.harvard.edu/content/120/1/87.short>

Cohen, Jessica and William Easterly, eds. 2009. *What Works in Development? Thinking Big and Thinking Small*. Introduction pp. 8–21. http://www.brookings.edu/~media/press/books/2009/whatworksindevelopment/whatworksindevelopment_chapter.pdf

Ludwig, Jens, Jeffrey R. Kling, and Sendhil Mullainathan. 2011. “Mechanism Experiments and Policy Evaluations.” *Journal of Economic Perspectives*, 25(3): 17–38. URL: <http://pubs.aeaweb.org.ezp-prod1.hul.harvard.edu/doi/pdfplus/10.1257/jep.25.3.17>

Deaton, Angus. 2010. “Instruments, Randomization, and Learning about Development.” *Journal of Economic Literature*. 48(2), pp. 424-55. URL: <http://pubs.aeaweb.org.ezp-prod1.hul.harvard.edu/doi/pdf/10.1257/jel.48.2.424>

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>

Blattman, Chris. 2008. Impact Evaluation 2.0. Lecture to DFID: http://chrisblattman.com/documents/policy/2008.ImpactEvaluation2.DFID_talk.pdf

First Written Assignment (Group RCT) Due Tuesday, March 10

III. Student Presentations and Discussion of Proposed Randomized Trial (1 session) Thursday, March 26, 2015 (Make-up session)

4pm-6pm, 10th floor conference room (access via 11th floor), building 1

IV. Instrumental Variables. (2 sessions); March 24 & March 31, 2015

Conditions for valid instruments; Reduced form/First Stage; Exclusion restrictions; Weak instruments;

Second Written Assignment Due March 31

Session 1 Readings:

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

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

*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>

*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>

Session 2 Readings:

*Angrist, Joshua and William Evans. 1998. "Children and Their Parents Labor Supply: Evidence from Exogenous Variation in Family Size." *American Economic Review*, 88, 450–77. URL: <http://www.jstor.org.ezp-prod1.hul.harvard.edu/stable/116844>

*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>

Some readings on IV from an Epidemiological Perspective:

Hernan, Miguel and James Robins. 2006. "Instruments for Causal Inference: An Epidemiologist's Dream?" *Epidemiology*, 17(4), 360-372
http://www.jstor.org.ezp-prod1.hul.harvard.edu/stable/20486236?seq=1#page_scan_tab_contents

Greenland, Sander. 2000. "An Introduction to Instrumental Variables for Epidemiologists" *International Journal of Epidemiology*, 29, 722-729.
<http://ije.oxfordjournals.org.ezp-prod1.hul.harvard.edu/content/29/4/722.full.pdf>

V. Difference-in-Differences (1 session); April 7, 2015

*MHE Sections 5.1-5.3 (inclusive)

*Duflo, Esther. 2003. "Grandmothers and Granddaughters: Old Age Pension and Intra-household Allocation in South Africa," *World Bank Economic Review*, 17(1): 1–25. URL: <http://wber.oxfordjournals.org.ezp-prod1.hul.harvard.edu/content/17/1/1.full.pdf+html>

Goldstein, et al. 2013. "The Effect of Absenteeism and Clinic Protocol on Health Outcomes: The Case of Mother-to-Child Transmission of HIV in Kenya." *American Economic Journal: Applied Economics*, 5(2), 58-85.

http://www.aeaweb.org.ezp-prod1.hul.harvard.edu/atypon.php?return_to=/doi/pdfplus/10.1257/app.5.2.58

Ward, Courtney. 2014. "Influenza Vaccination Campaigns: Is an Ounce of Prevention Worth a Pound of Cure?" *American Economic Journal: Applied Economics*, 6(1), 38-72.

<https://www.aeaweb.org/articles.php?doi=10.1257/app.6.1.38>

Ferrara, et al. 2012. "Soap Operas and Fertility: Evidence from Brazil." *American Economic Journal: Applied Economics*, 4(4), pp 1-31.

http://www.aeaweb.org.ezp-prod1.hul.harvard.edu/atypon.php?return_to=/doi/pdfplus/10.1257/app.4.4.1

Bertrand, et al. 2004. "How Much Should we Trust Difference-in-Differences Estimates?"

Quarterly Journal of Economics, 119, pp. 249–75. URL: <http://www.istor.org.ezp-prod1.hul.harvard.edu/stable/25098683>

VI. Difference-in-Difference as IV (1 session); April 14, 2015

Third Short Written Assignment Due

*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>

*Cutler, David et al. 2010. "Early-life Malaria Exposure and Adult Outcomes: Evidence from Malaria Eradication in India." *American Economic Journal: Applied Economics*, 2, 72-94.

http://dash.harvard.edu/bitstream/handle/1/5344529/malaria_nov09_final.pdf?sequence=1

Jacob, Brian. 2004. "Public Housing, Housing Vouchers and Student Achievement: Evidence from Public Housing Demolitions in Chicago," *American Economic Review*, 94(1), pp. 233–58.

URL: <http://www.jstor.org.ezp-prod1.hul.harvard.edu/stable/info/3592777>

Qian, Nancy. 2009. "Quantity-Quality and the One-Child Policy: The Only-Child Disadvantage in School Enrollment in Rural China." *BREAD Working Paper #228*.

https://ipl.econ.duke.edu/bread/system/files/bread_wpapers/228.pdf

VII. Regression Discontinuity (1 session); April 21, 2015

* MHE Chapter 6

*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:

<http://qje.oxfordjournals.org.ezp-prod1.hul.harvard.edu/content/125/2/591.full.pdf+html>

Anderson, Michael et al. 2012. "The Effect of Health Insurance Coverage on the Use of Medical Services." *American Economic Journal: Economic Policy*, 4(1), 1-27.

http://www.aeaweb.org.ezp-prod1.hul.harvard.edu/atypon.php?return_to=/doi/pdfplus/10.1257/pol.4.1.1

Lee, David and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature*, 48(2), pp. 281–355. URL: http://www.aeaweb.org.ezp-prod1.hul.harvard.edu/atypon.php?return_to=/doi/pdfplus/10.1257/jel.48.2.281

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>

VIII. Student Presentations of Proposed Evaluation; April 28 & May 5, 2015

May 5: Fourth Short Written Assignment Due

XI. Data Issues, Power Calculations, Sampling (1 session); May 12, 2015

Deaton, Angus, 1997. *The Analysis of Household Surveys. A Microeconometric Approach to Development Policy*, Johns Hopkins University Press, Baltimore. Chapter 1 and 2. URL: http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/1997/07/01/000009265_3980420172958/Rendered/PDF/multi_page.pdf

*Das, J, Hammer J and Sanchez-Paramo C. 2011. "The Impact of Recall Periods on Reported Morbidity and Health Seeking Behavior." World Bank Policy Research Working Paper #5778. http://www-wds.worldbank.org/servlet/WDSContentServer/WDSP/IB/2011/08/24/000158349_20110824082326/Rendered/PDF/WPS5778.pdf

*Zwane, Alex, Jonathan Zinman, Eric Van 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>

Martinelli, Cesar and Susan Parker. 2009. "Deception and Misreporting in a Social Program." *Journal of the European Economic Association* 7(4):886–908. URL: <http://onlinelibrary.wiley.com.ezp-prod1.hul.harvard.edu/doi/10.1162/JEEA.2009.7.4.886/pdf>
MHE Chapter 8

Delavande, et al. 2011. "Measuring Subjective Expectations in Developing Countries: A Critical Review and New Evidence." *Journal of Development Economics* 94(2), pp. 151–163 URL: <http://www.sciencedirect.com.ezp-prod1.hul.harvard.edu/science/article/pii/S0304387810000106#>

Leonard, K and M. Masatu. 2009 "Using the Hawthorne Effect to Examine the Gap Between a Doctor's Best Possible Practice and Actual Performance," *Journal of Development Economics* 92(2), pp. 226–34. URL: <http://dx.doi.org.ezp-prod1.hul.harvard.edu/10.1016/j.jdeveco.2009.11.001>

Final Papers Due: May 15, 2015 by 5pm (please submit electronically, as .doc file type)

School Requirement: Course Evaluation

Completion of the evaluation is a requirement for each course. Your grade will not be available until you submit the evaluation. In addition, registration for future terms will be blocked until you have completed evaluations for courses in prior terms.

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 more than once in each class session.	Student is mostly able to answer discussion questions when called on and initiates contributions 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 when others present materials and perspectives, as indicated by comments that build on others' remarks (i.e. the student hears what others say and contributes to the dialogue)	Student is mostly attentive when others present materials and perspectives, as indicated by comments that build on others' remarks. Occasionally, student needs encouragement or reminder from teaching staff to focus their comments.	Student is often inattentive and needs reminder to focus in class. Occasionally student makes disruptive comments while others are speaking.	Does not listen to others; regularly talks while others speak or does not pay attention while others speak; detracts from discussion.
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; can illustrate points on the board extremely well when called upon to do so.	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; can illustrate points on the board well when called upon to do so.	Comments are sometimes constructive and informed, with occasional signs of insight. Student does not use appropriate terminology and struggles to illustrate points on the board. Comments not always relevant.	Comments do not reflect careful reading and are not constructive. Student does not use appropriate terminology. Comments are not relevant to discussion.
Reading Reviews	Reading reviews always reflect careful, thoughtful, thorough reading of the assignments. Demonstrates detailed understanding of the paper and offers interesting possibilities for improvement.	Reading reviews mostly reflect careful, thorough reading of the assignments. Mostly demonstrates detailed understanding of the paper and sometimes offers interesting possibilities for improvement.	Reading reviews sometimes reflect careful reading of the assignments with occasional evidence of skimming and incomplete reading. Demonstrates basic understanding of the paper.	Reading reviews usually do not reflect careful reading of the assignments and usually show evidence of skimming and incomplete reading. Demonstrates limited understanding of the paper.

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, you should prepare a 2ish page (no less than 1.25 pages) reading review that addresses these questions and email them to the instructor and TA by 3:00p 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 reviewed each week and if the review is too thin and brief consistently 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.

Before each class, we will prepare a randomized order of names to be called on for class discussion. This is to ensure that everyone participates extensively and that we get a rounded view of opinions from everyone. This type of active participation will also help you understand the material better and be able to use it in your own work. Some discussion questions will be open for anyone to answer and some we will call on from the list. 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). We may also ask you to come to the board to explain (e.g. to write a regression equation or draw a graph that illustrates the identification strategy).

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?

Source: <http://publichealthryangosling.tumblr.com/>

Hey girl.

Sorry I
said
causation
when I
meant
correlation.

