

PLSC 508b

Dissertation Workshop:

Research Design & Causal Inference

Syllabus
Yale University
Spring 2010

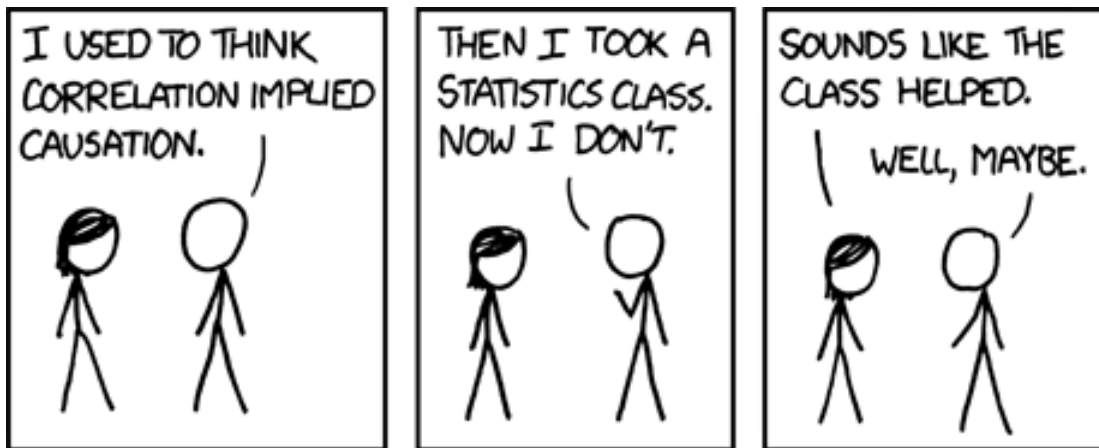
Workshop time: Tuesdays, 1:30-3:30

Workshop location: RKZ 202

Instructor: Chris Blattman, Departments of Political Science and Economics, christopher.blattman@yale.edu

Office Hours: Tuesdays 9am to 12pm, 77 Prospect St., Room 103

Instructor website: <http://chrisblattman.com/>



From <http://xkcd.com/552/>

Purpose and Nature of the Course:

Short story: Regardless of your specialty, this workshop is designed to improve your research plans, your dissertations, and your job prospects.

Using applications ranging from politics to health to criminology to economics, you'll learn to rip apart research designs and causal claims. The goal is to help students sharpen their skills as consumers and producers of applied empirical research.

Before spring break, your job will be to read applied papers and write scathing review reports. In class we'll take published papers and find their fatal flaws. We'll spot bad science, poor theory, selection problems, spurious correlations, endogenous relations, weak instruments, and measurement error.

After break, we'll do the same to your own research plans, papers, or paper replications. Students will present their own work. In class we'll break your project down and build it back up again, stronger than before.

You'll learn how to design better data collection, test for robustness and sensitivity, interpret coefficients, and fix causal concerns you never knew you had. It's a dissertation paper boot camp. By the end of class, you'll know when to hold it, when to fold it, when to walk away, and when to run.

Requirements

Research design and causal inference is first and foremost an intellectual and philosophical skill, and so solid qualitative researchers should be able to take this class, excel, and improve their own research.

You need to have taken 503 (Regression analysis) and 504 (Advanced quant) or the equivalent. Course 507 covers a similar set of topics, but is a complement to rather than a substitute for this class (and is not a prerequisite).

Credit versus audit

Auditors are welcome, but the class is capped at roughly 15 students. If the number of applicants exceeds 15, priorities are as follows:

1. Political science PhD students (2nd year or higher) who want to take the course for credit
2. Political science PhD students (3rd year or higher) who want to audit
3. PhD students from outside the department who want to take the course for credit
4. 1st year political science PhD students who want to take the course for credit
5. Undergraduate students who want to take the course for credit
6. All other categories

Second-year PhD students must take the class for credit (no auditing option).

Requirements & Grading

For credit

Your final grade has four components:

- Participation: 30%
- Weekly assignments: 0% (see below)
- In-class midterm: 40%
- Student paper, project plan, or replication: 30%

Weekly assignments (discussed below) do not have direct bearing on the final grade (except through their effects on quality of participations and the midterm).

Assuming we have space you can audit the class, but you'll be required to attend, participate, and present your own projects. If there are space constraints, auditors will be permitted under exceptional circumstances only.

Auditors

Auditors will be expected to do the following:

- Perform all readings
- Attend all student project presentations, providing verbal comments to the presenters

Participation

The meetings will be student driven. Students will be expected to provide critical analysis of the readings and student projects in class. You will be graded on the quality and (to a lesser extent) the quantity of your contributions.

Weekly Assignments

Each week you will write and submit a summary and review of the week's readings, plus any other notes you want to add. These weekly reports will not be graded. Rather, I will spend the class discussing the papers and methods and main lessons I want you to learn from the readings. During the midterm, however, you will receive a folder with all of your reviews.

The report should summarize key elements of the readings, and focus on critiques of each paper's data and empirical strategy as well as suggestions for improvement or alternative analysis. See "How to read and review an empirical paper" at the end of this syllabus for suggestions on how to approach this review.

The report can include other notes, including notes from the previous section, econometric notes, etc.

Each assignment should be 8 pages or less, in 12 point font, with 1 inch margins. These should be submitted in paper form in class.

Midterm

There will be an in-class midterm on March 23 that covers the readings and the econometric concepts and critiques discussed in class. All class discussion and all readings are testable—both the applied papers but also the material in the required books/papers on theory.

During the midterm, I will hand back to you all of the weekly assignments you submitted during the semester for use in the midterm. The more paper details, class notes, etc that you place in the weekly assignments, the better off you will be on the midterm.

You will also be allowed one 8x11 "cheat sheet" with anything you want to write on it.

Student projects

Description

Students are expected to present a short paper, project, or replication exercise.

Three other types of projects are permissible, with the instructor's permission:

- Replication of a published paper
- A plan or prospectus (i.e. a research design and plan, survey instrument development, etc.) for research to be conducted in the summer or academic year following the course
- Empirical papers and projects that are partly completed (i.e. a draft paper has been written). In these cases, students are still expected to do a significant amount of work on their data, analysis and presentation during the course. Papers that begin more polished will have higher expectations and be graded accordingly. What kind of paper or project are we talking about? Ideally you will produce a new or in-progress paper, where the data are available but analysis has just begun (and a draft paper is not yet written). A second-year “research and writing” paper is a good candidate, so long as it is not complete.

Co-authored projects are acceptable, but expectations will be higher.

Presentations

Approximately two to three students will present each class. Presentations will be no more than five minutes in length, and will be oral with the aid of printed tables and figures. The remainder of the class will be spent in a critique and defense of the presented materials. Everyone else will have read that submission prior to coming to class.

The student will submit preliminary topic ideas to the instructor by e-mail (up to one page) **by Jan 26**. We will sometimes spend the first 15 minutes of class discussing people’s paper ideas.

It should go without saying: all topics should address the issue of causal identification.

A 2 to 3 page description of progress, analysis, and preliminary results will be submitted by e-mail to the instructor **no later than March 2**. Students are encouraged to visit me in office hours before March 5 to discuss the project and progress. I will be in Africa over the spring break and not available to meet.

You will be required to email a draft paper to me and the class **on the Monday before your presentation**, so that all have a chance to read and comment on them in advance.

Materials ideally should include:

- Tables and figures
- Text of 10 to 15 pages in length (double-spaced, 12-point font, 1 inch margins) organized as follows: research question; research design; model summary (if relevant); data; empirical strategy; results; conclusions; and scope. Literature reviews, background, lengthy motivations, etc should be omitted or included as an appendix.
- Optional: a copy of the dataset and analysis code
- **Please follow the assignment guidelines.** If you have excess material, submit in appendices. If you have a full paper (e.g. your research and writing paper) do not submit that full paper unless it meets these guidelines.

If you are writing a paper or doing a replication, here is the recommended structure:

1. Question and motivation (very brief)
2. Theory and/or competing hypotheses
3. Empirical strategy
 - a. Tests of the theory: empirical predictions; description of how to distinguish between alternative hypotheses
 - b. Causal identification strategy: rationale; identification assumptions; tests of assumptions, remaining identification concerns and their implications
 - c. Other statistical issues (e.g. non-independence, measurement error, etc)

4. Data and measurement
 - a. Description of data (sample frame, sampling strategy, methodology, potential sample selection, etc)
 - b. Measurement (key variables, possible measurement error, etc)
5. Results
6. Discussion
 - a. Conclusions
 - b. Scope conditions/external validity
 - c. Implications for theory, further research, etc
7. Next steps/research plan

Essential Reading

Required

The following papers and books are required reading:

- Angrist, Joshua D., and S. Pischke. 2008. "[Mostly Harmless Econometrics: An Empiricists' Companion](#)." Princeton, NJ: Princeton University Press.
- Guido W. Imbens and Jeffrey Wooldridge. 2008. "[Recent Developments in the Econometrics of Program Evaluation](#)." NBER Working Paper No. 14251.
- Chapters 1 and 2 of Deaton, Angus. 1997. [The Analysis of Household Surveys: A Microeconomic Approach to Development Policy](#). Baltimore: Johns Hopkins University Press & The World Bank:

All of the above are dense and detailed (and testable on the midterm). You may feel you understand just a fraction the first time you read them. If you plan to practice empirical work that addresses causality, you should go back and read these over and over again. Each time will be rewarding.

Strongly recommended

The essential econometric text, as far as I am concerned, is the following:

- Jeffrey Wooldridge. 2002. [Econometric Analysis of Cross-Section and Panel Data](#). Cambridge: MIT Press (especially, but not only, chapters 4, 5, 6, 10, 11, 17, and 18)

This is a very useful article for applied empiricists:

- Joshua D. Angrist and Alan B. Krueger. 1999. "[Empirical Strategies in Labor Economics](#)." In Handbook of Labor Economics, ed. O. Ashenfelter and D. Card: Elsevier Science.

The following resources may also prove helpful.

- [Handbook of Econometrics Vol. 1-6](#)
- [JEP Symposium on Econometric Tools](#)

Lectures

Reading marked with an asterisk (*) should be the focus of your weekly assignments

1. Introduction (Jan 12)

- Popper, Karl. 1953. "Science: Conjectures and Refutations." Lecture given at Peterhouse, Cambridge. (<http://philosophyfaculty.ucsd.edu/faculty/rarneson/Courses/popperphil1.pdf>) – **especially sections I and II, and VI and VII.**
- Holland, Paul W. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81 (396):945-60.
- Fearon, James D. 1991. "Counterfactuals and Hypothesis Testing in Political Science." *World Politics* 43 (2):169-95.
- Lewis, David. 1973. "Causation." *The Journal of Philosophy* 70 (17):556-67.

2. Panel and cross-country-time series analysis (Jan 19)

- Imbens & Wooldridge, sections 1-4
- Angrist & Pischke, Sections 1,2, 3.1, 3.2, and 5.1
- David Card. 1999. "The Causal Effect of Education on Earnings." In *Handbook of Labor Economics*, ed. O. A. D. Card. Amsterdam: Elsevier.
- *James D. Fearon and David D. Laitin. 2003. "Ethnicity, Insurgency and Civil War." *American Political Science Review* 97 (1):75-90.
- *Barbara F. Walter. 2006. "Information, Uncertainty, and the Decision to Secede." *International Organization* 60 (1):105-135.

3. Selection on observables and unobservables (Jan 26)

- Imbens & Wooldridge, Sections 5.1-5.4, 5.6, 6.1, 6.2
- *Christopher Blattman (2009). "From Violence to Voting: War and Political Participation in Uganda." *American Political Science Review* 103: 231-247.
- *Fortna, Virginia Page. 2003. "Inside and Out: Peacekeeping and the Duration of Peace after Civil and Interstate Wars." *International Studies Review* 5 (4):97-114.

4. Selection on observables and unobservables (Feb 2)

- Angrist & Pischke, Sections 3.3, 3.4, 5.3, 8.2
- *Jason Lyall. 2009. "Does Indiscriminate Violence Incite Insurgent Attacks? Evidence from Chechnya." *Journal of Conflict Resolution*, 53:3.
- *Michael J. Gilligan and Ernest J. Sergenti (2008) "Do UN Interventions Cause Peace? Using Matching to Improve Causal Inference", *Quarterly Journal of Political Science*: Vol. 3(2): 89-122.
- *Joseph Altonji, Todd E. Elder, and Christopher Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy* 113: 151-184.

5. Differences in differences (Feb 9)

- Angrist & Pischke, Section 5.2

- Imbens & Wooldridge, Section 6.5
- Marianne Bertrand, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics* (February):249-275.
- *Tom Bundervoet, Richard Akresh, and Philip Verwimp. 2009. "Health and Civil War in Rural Burundi." *Journal of Human Resources* 44(2): 536-563
- *Card, David. 1990. "The Impact of the Mariel Boatlift on the Miami Labor Market." *Industrial and Labor Relations Review* 43(2): 245-57.

6. Matching (Feb 16)

- Imbens & Wooldridge, Sections 5.5, 5.8
- Angrist & Pischke, Sections 3.3, 3.4
- Sekhon, Jasjeet S. 2009. "Opiates for the Matches: Matching Methods for Causal Inference." *Annual Review of Political Science* 12: 487–508.
- *Macartan Humphreys and Jeremy M. Weinstein (2007). "Demobilization and Reintegration." *Journal of Conflict Resolution* 51(4): 531-567.
- *Jasjeet S. Sekhon. 2008. "The Varying Role of Voter Information Across Democratic Societies." Unpublished working paper (v1.4).

7. Instrumental variables (Mar 2)

- Angrist & Pischke, Section 4
- Imbens & Wooldridge, Section 6.3
- *Edward Miguel, Shanker Satyanah, and Ernest Sergenti. 2004. "Economic Shocks and Civil Conflict: An Instrumental Variables Approach." *Journal of Political Economy* Vol. 112 No. 4 pp. 725-753.
- *Joshua A. Angrist 1990. Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records. *American Economic Review* 80(3): 313-36.

8. Instrumental variables & Synthetic control (Mar 9)

- Joshua Angrist and Alan Krueger. 2001. "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments." *Journal of Economic Perspectives* Vol. 15 No. 4 (Autumn): 69-85.
- Murray, M. P. 2006. "Avoiding Invalid Instruments and Coping with Weak Instruments." *Journal of Economic Perspectives* 20 (4):111-32.
- *Steven D. Levitt. 1997. "Using Electoral Cycles in Police Hiring to Estimate the Effect of Police on Crime." *The American Economic Review* 87 (3):270-90.
- *Alberto Abadie, Alexis Diamond, and Jens Hainmueller. 2009. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association*.
- *Oeindrila Dube and Suresh Naidu. 2009. "Bases, Bullets and Ballots: the Effect of U.S. Military Aid on Political Conflict in Colombia." Unpublished working paper, December 2009.

Spring break

9. Midterm (March 23)

10. Regression Discontinuity (March 30)

- Imbens & Wooldridge, Section 6.4 (RD)
- Angrist & Pischke, Section 6
- Imbens, Guido W., and Thomas Lemieux. 2008. "Regression discontinuity designs: A guide to practice." *Journal of Econometrics* 142 (2):615-35.
- *Andrew Eggers and Jens Hainmueller. 2009. "MPs for Sale? Returns to Office in Postwar British Politics." *American Political Science Review* 103(4).
- *Lee, David S. 2008. "Randomized experiments from non-random selection in US House elections." *Journal of Econometrics* 142 (2):675-97.

11. Student projects (Apr 6)

12. Student projects (Apr 13)

13. Student projects (Apr 20)

Recommended readings by topic

Causal identification using observational data

Chamberlain, Gary, Griliches Zvi, and D. Intriligator Michael. 1984. "Panel data." In *Handbook of Econometrics*: Elsevier.

Joshua D. Angrist and Alan B. Krueger. 1999. "Empirical Strategies in Labor Economics." In *Handbook of Labor Economics*, ed. O. Ashenfelter and D. Card: Elsevier Science.

Card, David. 1999. "The Causal Effect of Education on Earnings." In *Handbook of Labor Economics*, ed. O. A. D. Card. Amsterdam: Elsevier.

Imbens, Guido W. 2004. "Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review." *Review of Economics and Statistics* 86 (1):4-29.

Paul Holland. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81 (396):945-960.

James J. Heckman. 2000. "Causal Parameters and policy analysis in economics: A twentieth century retrospective." *Quarterly Journal of Economics* 115:45-97.

Gerber, Alan S., Donald P. Green, and Edward H. Kaplan. 2004. *The Illusion of Learning from Observational Research*. In Ian Shapiro, Rogers Smith, and Tarek Massoud, eds., *Problems and Methods in the Study of Politics*. New York: Cambridge University Press, pp. 251-73.

Donald Rubin. 1984. "William Cochran's Contributions to the Design, Analysis, and Evaluation of Observational Studies." In Donald Rubin. 2006. *Matched Sampling for Causal Effects*. Cambridge, UK: Cambridge University Press. Chapter 1.

William G. Cochran and Donald Rubin. 1973. "Controlling Bias in Observational Studies: A Review." In Donald Rubin. 2006. *Matched Sampling for Causal Effects*. Cambridge, UK: Cambridge University Press. Chapter 2.

Dunning, Thad. 2008. "Improving Causal Inference: Strengths and Limitations of Natural Experiments." *Political Research Quarterly* 61 (2):282.

Heckman, James J., Edward J. Vytlacil, J. Heckman James, and E. Leamer Edward. 2007. "Econometric Evaluation of Social Programs, Part I: Causal Models, Structural Models and Econometric Policy Evaluation." In *Handbook of Econometrics*: Elsevier.

Heckman, James J., Edward J. Vytlacil, J. Heckman James, and E. Leamer Edward. 2007. "Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Econometric Estimators to Evaluate Social Programs, and to Forecast their Effects in New Environments." In *Handbook of Econometrics*: Elsevier.

Abbring, Jaap H., James J. Heckman, J. Heckman James, and E. Leamer Edward. 2007. "Econometric Evaluation of Social Programs, Part III: Distributional Treatment Effects, Dynamic Treatment Effects, Dynamic Discrete Choice, and General Equilibrium Policy Evaluation." In *Handbook of Econometrics*: Elsevier.

Rubin, Donald B. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66 (5):688-701.

Measurement and Measurement Error

Griliches, Zvi, Griliches Zvi, and D. Intriligator Michael. 1986. "Economic data issues." In *Handbook of Econometrics*: Elsevier.

Bound, J., C. Brown, and N. Mathiowetz. 2001. "Measurement Error in Survey Data." *HANDBOOKS IN ECONOMICS* 2:3705-843.

Hyslop, Dean R., and Guido W. Imbens. 2001. "Bias From Classical and Other Forms of Measurement Error." *Journal of Business & Economic Statistics* 19 (4):475-81.

Bound, J., and Alan B. Krueger. 1991. "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?" *Journal of Labor Economics* 9 (1):1.

Difference in Differences

Marianne Bertrand, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics* (February):249-275.

Stephen Donald and Kevin Lang. 2007. "Inference with Difference-in-Differences and Other Panel Data." *The Review of Economics and Statistics* 89 (2):221-233.

Susan Athey and Guido Imbens. 2006. "Identification and Inference in Nonlinear Difference-in-Differences Models." *Econometrica* Vol. 74 No. 2 (March):431-497.

Alberto Abadie. 2005. "Semiparametric Difference-in-Differences Estimators." *Review of Economic Studies* 72, 1-19.

Instrumental variables

Joshua Angrist and Alan Krueger. 2001. "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments." *Journal of Economic Perspectives* Vol. 15 No. 4 (Autumn):69-85.

Murray, M. P. 2006. "Avoiding Invalid Instruments and Coping with Weak Instruments." *Journal of Economic Perspectives* 20 (4):111-32.

Murray, Michael P. 2006. "The Bad, the Weak, and the Ugly: Avoiding the Pitfalls of Instrumental Variables Estimation." Bates College unpublished working paper.

Larry Bartels. 1991. "Instrumental and "Quasi-Instrumental" Variables." *American Journal of Political Science* Vol. 35 No. 3 (August):777-800.

Tim Conley, Christian Hansen, and Peter E. Rossi. 2007. "Plausibly Exogenous." University of Chicago Working Paper.

Thad Dunning. 2008. Model Specification in Instrumental-Variables Regression. *Political Analysis* 16 (3): 290-302.

Joshua Angrist, Guido Imbens, and Donald Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of American Statistical Association* Vol. 91 No. 434 (June):444-455.

Guido Imbens and Joshua Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* Vol. 62 No. 2 (March):467-475.

James Heckman. 1997. "Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations." *Journal of Human Resources* Vol. 32 No. 3 (Summer):441-462.

John Bound, David Jaeger, and Regina Baker. 1995. "Problems with Instrumental Variable Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak." *Journal of the American Statistical Association* Vol. 90 No. 430 (June):443-450.

Hahn, J., and J. Hausman. 2003. "Weak Instruments: Diagnosis and Cures in Empirical Econometrics." *American Economic Review* 93 (2):118-25.

Douglas Staiger and James Stock. 1997. "Instrumental Variables Regression with Weak Instruments." *Econometrica* Vol. 65 No. 3 (May):557-586.

Regression Discontinuity

van der Klaauw, Wilbert 2008. "Regression-Discontinuity Analysis: A Survey of Recent Developments in Economics." *LABOUR* 22 (2):219-45.

Imbens, Guido W., and Thomas Lemieux. 2008. "Regression discontinuity designs: A guide to practice." *Journal of Econometrics* 142 (2):615-35.

Matching

Sekhon, Jasjeet S. "Opiates for the Matches: Matching Methods for Causal Inference." UC Berkeley, unpublished working paper.

Jasjeet Sekhon. Forthcoming. "The Neyman-Rubin Model of Causal Inference and Estimation via Matching Methods." In *The Oxford Handbook of Political Methodology*, Oxford, UK: Oxford University Press.

Paul Rosenbaum and Donald Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." In Donald Rubin. 2006. *Matched Sampling for Causal Effects*. Cambridge, UK: Cambridge University Press. Chapter 10.

Donald Rubin and Neal Thomas. 1996. "Matching Using Estimated Propensity Scores: Relating Theory to Practice." In Donald Rubin. 2006. *Matched Sampling for Causal Effects*. Cambridge, UK: Cambridge University Press. Chapter 17.

Stephen L. Morgan and Christopher Winship. 2007. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. Cambridge, UK: Cambridge University Press. Chapters 4-5.

Arceneaux, Kevin, Alan S. Gerber, and Donald P. Green. 2006. Comparing Experimental and Matching Methods using a Large-Scale Voter Mobilization Experiment. *Political Analysis* 14: 1-36.

Benson, Kjell, and Arthur Jr. Hartz. 2000. A Comparison of Observational Studies and Randomized Controlled Trials. *New England Journal of Medicine* 342(25): 1878-1886.

LaLonde, Robert J. 1986. Evaluating the Econometric Evaluations of Training Programs with Experimental Data. *American Economic Review* 76: 604-20.

Paul Rosenbaum and Donald Rubin. 1983. "Assessing Sensitivity to an Unobserved Binary Covariate in an Observation Study with Binary Outcome." In Donald Rubin. 2006. *Matched Sampling for Causal Effects*. Cambridge, UK: Cambridge University Press. Chapter 11.

Paul Rosenbaum and Donald Rubin. 1985. "The Bias Due to Incomplete Matching." In Donald Rubin. 2006. *Matched Sampling for Causal Effects*. Cambridge, UK: Cambridge University Press. Chapter 14.

Stephen Morgan and David Harding. 2006. "Matching Estimators of Causal Effects: Prospects and Pitfalls in Theory and Practice." *Sociological Methods Research* Vol. 35 No. 1 (August):3-60.

Alberto Abadie and Guido Imbens. 2006. "Large Sample Properties of Matching Estimators for Average Treatment Effects." *Econometrica* Vol. 74 No. 1 (January):235-267.

Sascha Becker and Andrea Ichino. 2002. "Estimation of average treatment effects based on propensity scores." *The Stata Journal* 2 (4):358-377.

James Heckman, Hidehiko Ichimura, and Petra Todd. 1998. "Matching as an Econometric Estimator." *The Review of Economic Studies* Vol. 65:261-294.

James Heckman, Hidehiko Ichimura, and Petra Todd. 1997. "Matching as an Econometric Estimator: Evidence from Evaluating a Job Training Programme." *The Review of Economic Studies* Vol. 64:605-654.

Alexis Diamond and Jasjeet Sekhon. 2005. "Genetic Matching for Estimating Causal Effects." Harvard/Berkeley Working Paper.

Sensitivity analysis

Guido W. Imbens. 2003). "Sensitivity to Exogeneity Assumptions in Program Evaluation." *The American Economic Review* 93(2): 126-32.

Joseph Altonji, Todd E. Elder, and Christopher Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy* Vol. 113:151-184.

Paul Rosenbaum and Donald Rubin. 1983. "Assessing Sensitivity to an Unobserved Binary Covariate in an Observational Study with Binary Outcome." *Journal of the Royal Statistical Society. Series B (Methodological)* 45(2): 212-18.

Attrition

Duncan E. Thomas, Elizabeth Frankenberg, et al. (2001). "Lost but Not Forgotten: Attrition and Follow-up in the Indonesia Family Life Survey." *The Journal of Human Resources* 36(3): 556-92.

Hirano, Keisuke, Guido W. Imbens, Geert Ridder, and Donald B. Rubin. 2001. "Combining panel data sets with attrition and refreshment samples." *Econometrica* 69 (6).

How to read and review an empirical paper

Questions you can/should address in your reports:

Research question and hypothesis:

- Is the researcher focused on well-defined questions?
- Is the question interesting and important?
- Are the propositions falsifiable?
- Has the alternative hypothesis been clearly stated?
- Is the approach inductive, deductive, or an exercise in data mining? Is this the right structure?

Research design:

- Is the author attempting to identify a causal impact?
- Is the “cause” clear? Is there a cause/treatment/program/first stage?
- Is the relevant counterfactual clearly defined? Is it compelling?
- Is the method for doing so clear and compelling? Has statistical inference been confused with causal inference?
- Does the research design identify a very narrow or a very general source of variation?
- Could the question be addressed with another approach?
- Useful trick: ask yourself, “What experiment would someone run to answer this question?”

Theory/Model:

- Is the theory/model clear, insightful, and appropriate?
- Could the theory benefit from being more explicit, developed, or formal?
- Are there clear predictions that can be falsified? Are these predictions “risky” enough? Does the theory generate any prohibitions that can be tested?
- Would an alternative theory/model be more appropriate?
- Could there be alternative models that produce similar predictions—that is, does evidence on the predictions necessarily weigh on the model or explanation?
- Is the theory a theory, or a list of predictions?
- Is the estimating equation clearly related to or derived from the model?

Data:

- Are the data clearly described?
- Is the choice of data well-suited to the question and test?
- Are there any worrying sources of measurement error or missing data? Are any proxies reasonable?
- Are there sample size or power issues?
- Could the data sources or collection method be biased?
- Are there better sources of data that you would recommend?

- Are there types of data that should have been reported, or would have been useful or essential in the empirical analysis?

Empirical analysis:

- Are the statistical techniques well suited to the problem at hand?
- What are the endogenous and exogenous variables?
- Has the paper adequately dealt with concerns about measurement error, simultaneity, omitted variables, selection, and other forms of bias and identification problems?
- Is there selection not just in who receives the “treatment”, but in who we observe, or who we measure?
- Is the empirical strategy convincing?
- Could differencing, or the use of fixed effects, exacerbate any measurement error?
- Did the author make any assumptions for identification (e.g. of distributions, exogeneity, etc)?
- Were these assumptions tested and, if not, how would you test them?
- Are the results demonstrated to be robust to alternative assumptions?
- Does the disturbance term have an interpretation, or is it just tacked on?
- Are the observations i.i.d., and if not, have corrections to the standard errors been made?
- What additional tests of the empirical strategy would you suggest for robustness and confidence in the research strategy?
- Are there any dangers in the empirical strategy (e.g. sensitivity to identification assumptions)?
- Can you imagine a better, or alternative, empirical strategy?

Results:

- Do the results adequately answer the question at hand?
- Are the conclusions convincing? Are appropriate caveats mentioned?
- What variation in the data identifies the elements of the model?
- Are there alternative explanations for the results, and can we test for them?
- Could the author have taken the analysis further, to look for impact heterogeneity, for causal mechanisms, for effects on other variables, etc?
- Is absence of evidence confused with evidence of absence?

Scope:

- Can we generalize these results?
- Has the author specified the scope conditions?
- Have casual mechanisms been explored?
- Are there further types of analysis that would illuminate the external validity, or the causal mechanism at work?
- Are there other data or approaches that would complement the current one?