POL-GA 1251

Quantitative Political Analysis II

Spring 2012

Instructor:	Professor Cyrus Samii
	NYU Politics Department
	19 West 4th Street, Office 424
	cdsamii@gmail.com
Class time/location:	Wednesdays & Fridays, 10am-noon
	19 West 4th Street, Room 212
Office hours:	Wednesdays, 1-3pm or by appointment
Teaching Assistant:	Omar Garcia Ponce
	ogp200@nyu.edu
Recitation:	Mondays, 6-8pm
	19 West 4th Street, 3rd Floor Lab

Overview

This course provides a contemporary perspective on estimating causal effects in social science research. We will ground ourselves in methods related to ordinary least squares (OLS) regression, considering it as a fitting mechanism that can be meaningfully used to estimate conditional expectations with many kinds of data. We will emphasize research design, causal identification, and robust inference. Other methods of estimation will be introduced as well and related to OLS-based methods.

Prerequisites

The course requires that students have working knowledge of probability theory, matrix algebra, and calculus at the level of POL-GA 1250, "Quant I." The course is intended to provide foundational methodological training to PhD students in politics as part of their required sequence of courses toward their degree. The course is not available to students from other programs nor is it open to student auditors.

Texts and software

The course will draw mostly from the following two textbooks:

1. Angrist, Joshua, and Steffan Jorg Pischke. 2008. *Mostly Harmless Econometrics*. Princeton: Princeton University Press. (Referred to as MHE.)

2. Morgan, Stephen L., and Christopher Winship. 2007. *Counterfactuals and Causal Inference: Methods and Principles for Social Research.* Cambridge, UK: Cambridge University Press. (Referred to as CCI.)

The course will mostly follow MHE, using CCI to provide more intuitive background and illustrations. I will also supplement the textbooks with notes, sections from other textbooks. Occasionally we will read methodological papers and journal articles on special topics. I have listed "further reading" for each topic, and my lectures will sometimes draw on these. Readings will be available in a public Dropbox (current url is http://db.tt/Axav2LI7).

The books are each detailed, up to date, and they complement each other well. The books are not terribly expensive, and so students are recommended to acquire both. MHE can be mathematically difficult at times, but you are strongly encouraged to dive in, replicate proofs, and work hard to understand it. CCI provides intuition to keep you sane and grounded.

While these two textbooks are technically sound, they do gloss over some details on mechanics. Unfortunately there is no single textbook that I think covers all the bases on mechanics in a manner that is sufficiently clear and detailed. Nonetheless, textbooks that I recommend include:

- Davidson, Russell, and James G. MacKinnon. 2003. *Econometric Theory and Methods*. Oxford: Oxford University Press. (Rather technical but also comprehensive treatment of linear regression for both micro- and macro-econometric applications. If you can only pick up one reference book, I guess this would be it, because of its breadth.)

- DeGroot, Morris H., and Mark J. Schervish. 2002. *Probability and Statistics* (Third Edition). Boston: Addison Wesley. (Authoritative and technical textbook on foundational topics in statistics.)

- Freedman, David A. 2009. *Statistical Models: Theory and Practice*. Cambridge, UK: Cambridge University Press. (Step-by-step, workbook style development of rudiments for regression, statistical inference, and bootstrapping, with a critical eye. If you feel that there are "holes" in your understanding of regression mechanics, use this book to fill those in.)

- Goldberger, Arthur S. 1991. *A Course in Econometrics*. Cambridge, MA: Harvard University Press. (Develops the theory of "agnostic" linear regression step-by-step from first principles. A bit dated, but a classic. This book leads most naturally into what MHE covers.)

- Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross-Section and Panel Data*. Cambridge, MA: MIT Press. (Authoritative and technical treatment of linear regression for micro-econometric applications. If you plan to work on advanced methodology problems that go beyond the scope of this course, then you will want to get to know this book.)

If you can afford it, it would be a good idea to have at least one or two of these on your shelf as a reference.

As for software, we will work in both R and Stata. Each program has its advantages, and so it is useful to obtain fluency in both. R is great for programming estimators and algorithms "from scratch," programming simulations, and making graphics. Some assignments will ask you to do that. However many of the packaged estimation routines are flawed. Stata's pre-programmed estimation routines are more reliable, and some assignments could be done using them; however Stata is very inconvenient for programming, simulations, or graphics.

Requirements and policies

Homework

You will receive homework about every week and then you will have to submit your completed assignment by the following week before the end of class; exact deadlines will be made clear on the assignment. You can work with others to complete the homework. But to receive credit for a homework assignment, you must turn in a *hard copy* of your own homework by the stated deadline. Completed homework assignments can be submitted as a hand written or typed document, but they must be clearly legible. Estimates obtained from data analysis programs (e.g., Stata or R) must be *formatted into tables* resembling journal presentation styles. Be sure to use a reasonable (2 or at most 3) number of digits after decimal points, report standard errors along with coefficients, and explain in footnotes what kinds of estimators or adjustments have been used. Print outs of “raw” screen output or logs will not receive credit. Mathematical derivations should include all key steps with annotation or explanation for steps that apply important techniques. Homework will count toward 40% of your grade.

Mid-term exam

An in-class mid-term exam is provisionally scheduled during class on March 2. The mid-term serves the purpose of evaluating individual progress, which in turn helps me to understand where to place emphasis for the remainder of the semester. If you are unable to make it to the exam, you must provide notice a week prior so that we can arrange an alternative time. The mid-term will count toward 20% of your grade.

Final exam

An in-class final exam will be scheduled during the final examination period. The final also serves the purpose of evaluating individual progress, which in turn allows me to provide individualized recommendations on where students should apply effort to strengthen their methodological foundations. If you are unable to make it to the exam, you must provide notice a week prior so that we can arrange an alternative time. The final will count toward 30% of your grade.

Attendance and participation

Attendance is required. Participation in class discussions and activities is required. Attendance and participation will count toward 10% of your grade.

Special needs

Students with special needs should come to office hours or schedule an appointment with the instructor to discuss possible accommodation.

Topics

Topics listed below will be covered in around 1-3 class sessions each. Required reading sometimes corresponds directly to material covered in the sessions and sometimes builds up background needed for future sessions. Most of the required reading comes from MHE and CCI, although the topics covered toward the end of the semester will draw on other texts that will be made available.

1 Identification, estimation, and inference

Identification concepts: randomization and the experimental ideal, manipulability, potential outcomes and other models of causality, average treatment effects, dose-response functions. *Estimation concepts*: estimands and estimators, bias, consistency, and efficiency. *Inference concepts*: finite and infinite populations, data generating process, implications of randomization and sampling, stochasticity, choosing an “error term," exact distributions, asymptotic distributions and central limit theorem.

Required reading: MHE Ch. 1-2; CCI Ch. 1-2.

Further reading: Angrist and Krueger (1999); Angrist and Pischke (2010); DiNardo and Lee (2011); Freedman (1991); Holland (1986); Imbens and Wooldridge (2009); Imbens and Rubin (2011, Ch. 1-3); Manski (1995, Ch. 1); Pearl (2009, Ch. 1); Rothman (1976); Rosenbaum (1999); Rosenbaum (2002, Ch. 1-3); Rubin (1978); Rubin (1986).

2 Regression mechanics

Core decompositions and their relation to regression; Frisch-Waugh-Lovell; conditional independence/exogeneity; omitted variable bias formula; effect heterogeneity and nonlinearity; regression and matching; regression and simple contrasts; influence and leverage; testing restrictions on coefficients.

Required reading: MHE Ch. 3; CCI Ch. 3 and 5.

Further reading: Freedman (2008a); Freedman (2008b); Imbens and Rubin (2011, Ch. 7); Lin (2011); Samii and Aronow (2012); Schochet (2010).

3 Notions of bias

Confounding; post-treatment bias; sample selection bias; aggregation/effect heterogeneity biases; misspecification/extrapolation bias and model dependence; measurement error biases.

Required reading: CCI Ch. 6; Bound et al. (2000, pp. 1-39); Heckman (1979); Humphreys (2009); Imai et al. (2008); King and Zeng (2006); Lalonde (1986); Rosenbaum (1984).

Further reading: Frangakis and Rubin (2002); Heckman et al. (1998); Hyslop and Imbens (2001); Pearl (2009, Ch. 3, 6).

4 Control and balance via matching and weighting

Identification under conditional exogeneity; alternative matching and weighting algorithms; estimation and inference after matching; matching and weighting for multivalued treatments

Required reading: CCI Ch. 4.

Further reading: Busso et al. (2011); Dehejia and Wahba (2002); Hainmueller (2011); Hirano and Imbens (2004); Ho et al. (2007); Iacus et al. (2011); Imai and van Dyk (2004); Imbens (2000); King et al. (2011); Lu et al. (2001); Rosenbaum and Rubin (1983); Sekhon (2009); Todd (2006).

5 Robust inference

Clustering, autocorrelation, and spatial dependence; Moulton's problem; heteroskedasticity and cluster robust standard errors; bootstrapping; estimating the exact randomization variance.

Required reading: MHE Ch. 8.

Further reading: Barrios et al. (2010); Cameron et al. (2008); Cameron et al. (2009); Conley (1999); Davidson and MacKinnon (2004, Ch. 4-5); Freedman (2009, Ch. 8); Hansen (2007); Mosteller and Tukey (1977, Ch. 7); ; Moulton (1986).

6 Identification with instrumental variables

Exclusion restriction; valid first stage; principal strata; local average treatment effect (LATE); weak instrument; sensitivity analysis.

Required reading: MHE Ch. 4; CCI Ch. 7.

Further reading: Angrist et al. (1996); Baum et al. (2003); Bound et al. (1995); Conley et al. (2010); Deaton (2010); Heckman and Urzua (2009); Imbens (2010); Imbens and Rosenbaum (2005); Kolesar et al. (2011); Sovey and Green (2011); Staiger and Stock (1997); Stock et al. (2002).

7 Identification with repeated observations

Adjusting for fixed, unobserved heterogeneity via fixed effects and difference-in-differences; synthetic control.

Required reading: MHE Ch. 5; CCI Ch. 9.

Further reading: Abadie et al. (2010); Beck and Katz (2001); Bertrand et al. (2004); Green et al. (2001); Imai and Kim (2012).

8 Regression discontinuity (RD)

Forcing variables; sharp and fuzzy RD; conditional average treatment effect (CATE); local linearity, bandwidth, and non-parametric regression; kernel weighting; multiway discontinuities; checks for sorting around cut-points; endogenous forcing variables; measurement error in forcing variables.

Required reading: MHE Ch. 6; CCI Ch. 9.

Further reading: Caughy and Sekhon (2011); Green et al. (2009); Imbens and Kalyanaraman (2009); Imbens and Lemieux (2008); Lee and Card (2008); Lee and Lemieux (2010); McCrary (2008); Papay et al. (2011); Urquiola and Verhoogen (2009).

9 Distributional effects and quantile regression

Effect heterogeneity; quantile treatment effect; minimum absolute deviations; rank invariance.

Required reading: MHE Ch. 7.

Further reading: Bitler et al. (2006); Chernozhukov and Hansen (2005); Heckman et al. (1997); Koenker and Hallock (2000).

10 Multiple endpoints

Mean effects and weighted mean effects; multiple comparisons adjustments.

Required reading: Anderson (2008); Kling and Liebman (2004); Shaffer (1995).

Further reading: Casey et al. (2011); Clingingsmith et al. (2009); Farcomeni (2008); Gibson et al. (2011); O'Brien (1984).

11 Missing data and attrition

Bounds; inverse-propensity weighting; imputation.

Required reading: Gerber and Green (2012, Ch. 7); Horton and Kleinman (2007); King et al. (2001); Lee (2009); Manski (1995, Ch. 2); Vansteelandt et al. (2010).

Further reading: Jones (1996); Puma et al. (2009); Samii (2011).

12 Limited dependent variable effects

(*Material to be covered depends on number of sessions that remain at end of semester.*) Overview of maximum likelihood estimation for models for binary, multinomial, count, and duration data; comparison to results from OLS estimation and other modes of estimation.

Required reading: Davidson and MacKinnon (2004, Ch. 10-11).

Further reading: Angrist (2001); Fox (2002); Freedman (2006); Greene (2004); Hubbard et al. (2010); Liang and Zeger (1986); Wooldridge (2002, Ch. 15, 19-20).

References

Abadie, A., A. Diamond, and J. Hainmueller (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association 105*(490), 493–505.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervetion: A reevaluation of the Abecedarian, Perry Preschool and Early Training projects. *Journal of the American Statistical Association 103*(484), 1481–1495.

Angrist, J. D. (2001). Estimation of limited dependent variable models with dummy endogenous regressors: Simple strategies for empirical practice. *Journal of Business and Economic Statistics 19*(1), 2–16.

Angrist, J. D., G. W. Imbens, and D. B. Rubin (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association 91*(434), 444–455.

Angrist, J. D. and A. B. Krueger (1999). Empirical strategies in labor economics. In O. C. Ahsenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3. Amsterdam: North Holland.

Angrist, J. D. and J.-S. Pischke (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspectives 24*(2), 3–30.

Barrios, T., R. Diamond, G. W. Imbens, and M. Kolesar (2010). Clustering, spatial correlations, and randomization inference. *NBER Working Paper Series 15760*.

Baum, C. F., M. E. Schaffer, and S. Stillman (2003). Instrumental variables and GMM: Estimation and testing. *Stata Journal 3*(1), 1–31.

Beck, N. and J. N. Katz (2001). Throwing out the baby with the bath water: A comment on Green, Kim, and Yoon. *International Organization 55*(2), 487–495.

Bertrand, M., E. Duflo, and S. Mullainathan (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics 119*(1), 249–275.

Bitler, M. P., J. B. Gelbach, and H. W. Hoynes (2006). What mean impacts miss: Distributional effects of welfare reform experiments. *American Economic Review 96*(4), 988–1012.

Bound, J., C. Brown, and N. Mathiowetz (2000). Measurement error in survey data. Technical Report Report No. 00-450, Population Studies Center, University of Michigan.

Bound, J., D. A. Jaeger, and R. M. Baker (1995). Problems with instrumental variables estimation when correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association 90*(430), 443–450.

Busso, M., J. DiNardo, and J. McCrary (2011). New evidence on the finite sample properties of propensity score reweighting and matching estimators. Manuscript, IDB, University of Michigan, and University of California, Berkeley.

Cameron, A., J. Gelbach, and D. L. Miller (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics 90*(3), 414–427.

Cameron, A., J. Gelbach, and D. L. Miller (2009). Robust inference with multi-way clustering. Manuscript, University of California, Berkeley.

Casey, K., R. Glennerster, and E. Miguel (2011). Reshaping institutions: Evidence on aid implacts using a pre-analysis plan. *NBER Working Paper Series 17012*.

Caughy, D. and J. S. Sekhon (2011). Elections and the regression discontinuity design: Lessons from close US House races, 1942-2008. *Political Analysis 19*(4), 385–408.

Chernozhukov, V. and C. B. Hansen (2005). An IV model of quantile treatment effects. *Econometrica 73*(1), 245–261.

Clingingsmith, D., A. I. Khwaja, and M. Kremer (2009). Estimating the impact of the hajj: Religion and tolerance in Islam's global gathering. *Quarterly Journal of Economics 124*(3), 1133–1170.

Conley, T. G. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics 92*, 1–45.

Conley, T. G., C. B. Hansen, and P. E. Rossi (2010). Plausibly exogenous. *The Review of Economics and Statistics forthcoming*.

Davidson, R. and J. G. MacKinnon (2004). *Econometric Theory and Methods*. Oxford: Oxford University Press.

Deaton, A. (2010). Instruments, randomization, and learning about development. *Journal of Economic Literature 48*, 424–455.

Dehejia, R. H. and S. Wahba (2002). Propensity score-matching methods for nonexperimental causal studies. *The Review of Economics and Statistics 84*(1), 151–161.

DiNardo, J. and D. S. Lee (2011). Program evaluation and research designs. *Handbook of Labor Economics 4a*(463-536).

Farcomeni, A. (2008). A review of modern multiple hypothesis testing, with particular attention to the false discovery proportion. *Statistical Methods in Medical Research 17*, 347–388.

Fox, J. (2002). Cox proportional-hazards regression for survival data. Online appendix to *An R and S-PLUS Companion to Applied Regression*.

Frangakis, C. E. and D. B. Rubin (2002). Principal stratification in causal inference. *Biometrics 58*, 21–29.

Freedman, D. A. (1991). Statistical models and shoe leather. *Sociological Methodology 21*, 291–313.

Freedman, D. A. (2006). On the so-called "Huber sandwich estimator" and "robust standard errors". *The American Statistician 60*(4), 299–302.

Freedman, D. A. (2008a). On regression adjustments in experiments with several treatments. *The Annals of Applied Statistics 2*(1), 176–196.

Freedman, D. A. (2008b). On regression adjustments to experimental data. *Advances in Applied Mathematics 40*(1), 180–193.

Freedman, D. A. (2009). *Statistical Models: Theory and Practice*. Cambridge: Cambridge University Press.

Gerber, A. S. and D. P. Green (2012). *Field Experiments: Design and Analysis*. New York: Norton.

Gibson, J., D. McKenzie, and S. Stillman (2011). The impacts of international migration on remaining household members: Omnibus results from a migration lottery program. *The Review of Economics and Statistics 93*(4), 1297–1318.

Green, D. P., S. Y. Kim, and D. H. Yoon (2001). Dirty pool. *International Organization 55*(2), 441–468.

Green, D. P., T. Y. Leong, H. L. Kern, A. S. Gerber, and C. W. Larimer (2009). Testing the accuracy of regression discontinuity analysis using experimental benchmarks. *Political Analysis 17*(4), 400–417.

Greene, W. (2004). The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects. *Econometrics Journal 7*, 98–119.

Hainmueller, J. (2011). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis 17*(4), 400–417.

Hansen, C. B. (2007). Asymptotic properties of a robust variance matrix estimator for panel data when *t* is large. *Journal of Econometrics 141*, 597–620.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica 47*(1), 153–161.

Heckman, J. J., H. Ichimura, J. Smith, and P. Todd (1998). Characterizing selection bias using experimental data. *Econometrica 66*(5), 1017–1098.

Heckman, J. J., J. Smith, and N. Clements (1997). Makings the most out of programme evaluations and social experiments: Accounting for heterogeneity in programme impacts. *Review of Economic Studies 64*(4), 487–535.

Heckman, J. J. and S. Urzua (2009). Comparing IV with structural models: What simple IV can and cannot identify. *NBER Working Paper Series 14706*.

Hirano, K. and G. W. Imbens (2004). The propensity score with continuous treatments. Manuscript, University of Miami and University of California, Berkeley.

Ho, D. E., K. Imai, G. King, and E. A. Stuart (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis 15*(3), 199–236.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association 81*(396), 945–960.

Horton, N. J. and K. P. Kleinman (2007). Much ado about nothing: A comparison of missing data methods and software to fit incomplete data regression models. *American Statistician 61*(1), 79–90.

Hubbard, A. E., J. Ahern, N. L. Fleischer, M. Van der Laan, S. A. Lippman, N. Jewell, T. Bruckner, and W. A. Satariano (2010). To GEE or not to GEE. *Epidemiology 21*(4), 467–474.

Humphreys, M. (2009). Bounds on least squares estimates of causal effects in the presence of heterogenous assignment probabilities. Manuscript, Columbia University.

Hyslop, D. R. and G. W. Imbens (2001). Bias from classical and other forms of measurement error. *Journal of Business and Economic Statistics 19*(4), 475–481.

Iacus, S. M., G. King, and G. Porro (2011). Causal inference without balance checking: coarsened exact matching. *Political Analysis 19*(4), 1–24.

Imai, K. and I. Kim (2012). Understanding and improving linear fixed effects regression models for causal inference. Manuscript, Princeton University.

Imai, K., G. King, and E. A. Stuart (2008). Misunderstandings between experimentalists and observationalists about causal inference. *Journal of the Royal Statistical Society, Series A 171*(2), 481–502.

Imai, K. and D. A. van Dyk (2004). Causal inference with general treatment regimes: Generalizing the propensity score. *Journal of the American Statistical Association 99*(467), 845–866.

Imbens, G. W. (2000). The role of the propensity score in estimating dose-response functions. *Biometrika 87*, 706–710.

Imbens, G. W. (2010). Better LATE than nothing: Some comments on Deaton (2009) and Heckman and Urzua (2009). *Journal of Economic Literature 48*, 399–423.

Imbens, G. W. and K. Kalyanaraman (2009). Optimal bandwidth choice for the regression discontinuity estimator. *NBER Working Paper Series 14726*.

Imbens, G. W. and T. Lemieux (2008, 635). Regression discontinuity designs: A guide to practice. *Journal of Econometrics 142*(2), 615.

Imbens, G. W. and P. R. Rosenbaum (2005). Robust, accurate confidence intervals with a weak instrument: quarter of birth and education. *Journal of the Royal Statistical Society, Series A 168*(1), 109–126.

Imbens, G. W. and D. B. Rubin (2011). Causal inference in statistics and social sciences. Manuscript, Harvard University.

Imbens, G. W. and J. M. Wooldridge (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature 47*(1), 5–86.

Jones, M. P. (1996). Indicator and stratification methods for missing explanatory variables in multiple linear regression. *Journal of the American Statistical Association 91*(433), 222–230.

King, G., J. Honaker, A. Joseph, and K. Scheve (2001). Analyzing incomplete political science data: An alternative algorithm for multiple imputation. *American Political Science Review 95*(1), 49–69.

King, G., R. Nielsen, C. Coberley, J. E. Pope, and A. Wells (2011). Comparative effectiveness of matching methods for causal inference. Manuscript, Harvard University.

King, G. and L. Zeng (2006). The dangers of extreme counterfactuals. *Political Analysis 14*(2), 131–159.

Kling, J. R. and J. B. Liebman (2004). Experimental analysis of neighborhood effects on youth. Manuscript, Princeton University and Harvard University.

Koenker, R. and K. F. Hallock (2000). Quantile regression: An introduction. Manuscript, University of Illinois at Urbana-Champaign.

Kolesar, M., R. Chetty, J. Friedman, E. Glaeser, and G. W. Imbens (2011). Identification and inference with many invalid instruments. Manuscript, Harvard University.

Lalonde, R. J. (1986). Evaluating the econometric evaluations of training programs with experimental data. *American Economic Review 76*(4), 604–620.

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *Review of Economic Studies 76*, 1071–1102.

Lee, D. S. and D. Card (2008). Regression discontinuity inference with specification error. *Journal of Econometrics 142*(2), 655–674.

Lee, D. S. and T. Lemieux (2010). Regression discontinuity designs in economics. *Journal of Economic Literature 48*, 281–355.

Liang, K. and S. L. Zeger (1986). Longitudinal data analysis using generalized linear models. *Biometrika 73*(1), 13–22.

Lin, W. (2011). Agnostic notes on regression adjustments to experimental data: Reexamining freedman's critique. Manuscript, Department of Statistics, University of California, Berkeley.

Lu, B., E. Zanutto, R. Hornik, and P. R. Rosenbaum (2001). Matching with doses in an observational study of a media campaign against drug abuse. *Journal of the American Statistical Association 96*(456), 1245–1253.

Manski, C. F. (1995). *Identification Problems in the Social Sciences*. Cambridge, MA: Harvard University Press.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics 142*(2), 698–714.

Mosteller, F. and J. W. Tukey (1977). *Data Analysis and Regression: A Second Course in Statistics*. Boston: Addison Wesley.

Moulton, B. R. (1986). Random group effects and the precision of regression estimates. *Journal of Econometrics 32*, 385–397.

O'Brien, P. C. (1984). Procedures for comparuing samples with multiple endpoints. *Biometrics 40*(4), 1079–1087.

Papay, J. P., J. B. Willett, and R. J. Murnane (2011). Extending the regression-discontinuity approach to multiple assignment variables. *Journal of Econometrics 161*(2), 203–207.

Pearl, J. (2009). *Causality: Models, Reasoning, and Inference, Second Edition*. Cambridge: Cambridge University Press.

Puma, M. J., R. B. Olsen, S. H. Bell, and C. Price (2009). What to do when data are missing in group randomized trials. Technical report 2009-0049, National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, Washington D.C.

Rosenbaum, P. R. (1984). The consequences of adjustment for a concomitant variable that has been affected by the treatment. *Journal of the Royal Statistical Society, Series A 147*(5), 656–666.

Rosenbaum, P. R. (1999). Choice as an alternative to control in observational studies. *Statistical Science 14*(3), 259–304.

Rosenbaum, P. R. (2002). *Observational Studies, 2nd Ed.* New York: Springer.

Rosenbaum, P. R. and D. B. Rubin (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika 70*(1), 41–55.

Rothman, K. J. (1976). Causes. *American Journal of Epidemiology 104*(6), 587–592.

Rubin, D. B. (1978). Bayesian inference for causal effects: The role of randomization. *The Annals of Statistics 6*(1), 34–58.

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

Samii, C. (2011). Weighting and augmented weighting for causal inference with missing data: new directions. Manuscript, Columbia University and Yale University.

Samii, C. and P. M. Aronow (2012). On equivalencies between design-based and regression-based variance estimators for randomized experiments. *Statistics and Probability Letters 82*(2), 365–370.

Schochet, P. Z. (2010). Is regression adjustment supported by the Neyman model for causal inference? *Journal of Statistical Planning and Inference 140*(1), 246–259.

Sekhon, J. S. (2009). Opiates for the matches: Matching methods for causal inference. *Annual Review of Political Science 12*(1), 487–508.

Shaffer, J. P. (1995). Multiple hypothesis testing. *Annual Review of Psychology 46*, 561–584.

Sovey, A. J. and D. P. Green (2011). Instrumental variables estimation in political science: A reader's guide. *American Journal of Political Science 55*(1), 188–200.

Staiger, D. and J. H. Stock (1997). Instrumental variables regression with weak instruments. *Econometrica 65*(3), 557–586.

Stock, J. H., J. H. Wright, and M. Yogo (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business and Economic Statistics 20*(4), 518–529.

Todd, P. (2006). Matching estimators. Manuscript, University of Pennsylvania.

Urquiola, M. and E. Verhoogen (2009). Class-size caps, sorting, and the regression-discontinuity design. *American Economic Review 99*(1), 179–215.

Vansteelandt, S., J. Carpenter, and M. G. Kenward (2010). Analysis of incomplete data using inverse probability weighting and doubly robust estimators. *Methodology 6*(1), 37–48.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.