POL-GA.3200.002
Quantitative Field Methods
Spring 2012
Working Draft – Subject to Revision – Comments Welcome

Instructor: Cyrus Samii
NYU Politics Department
19 West 4th Street, Office 424
cds2083@nyu.edu

Class time/location: Thursdays, 4:00pm – 6:00pm
19 W 4th St, Rm. 212

Office hours: Wednesdays, 1-4pm, or by appointment

I. Overview

This is a graduate course on statistical methods for designing quantitative social science field research, including sample surveys, field experiments, and observational (quasi-experimental) studies. The purpose of this course is to train graduate students in the social sciences to design rigorous quantitative micro-level fieldwork for their research. The learning goals are (i) to understand why some sampling, experimental, or measurement techniques are to be preferred over others, (ii) to be able to analyze design alternatives and implement sampling, treatment assignment, and measurement algorithms in the R statistical computing environment, and (iii) to develop an ability to take meaningful social science questions and translate them into hypotheses and research designs that can address the questions in a compelling manner.

II. Prerequisites

Prerequisites are introductory graduate-level statistics such as a first-year sequence covering probability, inference, and regression. The *Survey Research Methods* (RESCH-GE.2139) course taught through the Steinhardt School (or equivalent) is highly recommended though not required. The presentation will be math intensive. Weekly assignments will involve statistical programming in R.

III. Texts

In order to provide adequate clarity, depth, and breadth, the course will draw on both classic texts as well as recent literature in statistics and social sciences. Texts covering each of the topics are shown below and will be made available in PDF format. All of the texts are recommended, but for weeks with exceptionally heavy reading lists, students may choose to focus on only a few of the readings for that week while skimming the others. Texts that contain key concepts and formula and therefore require special attention will be flagged in class.

IV. Requirements and policies

Programming assignments: Developing fluency with R as a statistical research design tool is a core goal of the class. You will have regular R programming exercises to carry out analyses and simulations based on topics covered in class. Data for the assignments will be provided and relevant packages and functions will be indicated. For each assignment, students must submit their own write-ups, with all graphs and tables. These will typically be due at the next class, although late assignments may be accepted under approval of the instructor. We may discuss assignments in class, and so each student should be prepared to do so. These assignments will account for 50% of your grade.

Research design proposal: Each student will each come up with a research design proposal for either a field experiment or observational study. The design proposal will include a full discussion and justifications for the treatment assignment method or identification strategy, sample size and sampling plan, and measurement techniques. The last few weeks of class will be reserved for students to present their research designs. The design may be based on either an original research topic (including thesis or dissertation research) or "enhanced" replication of a study archived on either the ISPS or JPAL experiments archives (ISPS experiments archive: http://bit.ly/sV3yxu; JPAL political economy experiments archive: http://bit.ly/sK7lAB). By "enhanced" replication I mean that the design allows for rigorous testing of refined or additional hypotheses that go beyond the main hypothesis that was tested in the original research. For example, the enhanced replication may use blocking or a factorial design in order to unpack the meaning of an interaction effect or a presumed mediation effect that was identified in the original study. A draft version of the proposal will be presented before the class, and then the final proposal will be due one week after the last class. The proposal will account for 30% of your grade.

Attendance and participation: Attendance is required. Participation in discussions of assignments is required. Presentation of a draft research design proposal is required. Attendance and participation will account for 20% of your grade.

Eligibility for the class: The class is primarily meant to serve PhD students in politics, however if availability permits, the class is open to graduate students from other programs/schools. The cap is 20 students. If we reach the cap, priority will be given to NYU politics PhD students beyond first year, NYU PhD students in other departments beyond first year, and then NYU master's students beyond first year. Other types of students will be admitted based on consideration of their preparation, potential to benefit from the course, and possibility of taking the course in the future.

Auditing: The nature of the class is interactive and requires full participation. Thus, no auditing will be permitted.

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

V. Topics and session schedule

The course is broken down into three themes: (1) Sampling Designs, (2) Field Experiment and Observational Study Designs, and (3) Measurement Techniques. The ordering reflects the cumulative nature of the material, with sampling theory being most foundational. The texts and topics will include the following (subject to revision):

A. Sampling Designs

1. Introduction to randomization theory, simple random sampling theory, large sample theory, and sample size determination.

Readings: Cochran (1977), Ch. 1-4; Thompson (2002), Ch. 1-4; United Nations (2005), Ch. V.

Assignment: Write R programs to graph asymptotic sampling distributions of the following statistics for population distributions that exhibit various degrees of departure from normality: (i) mean, (ii) standardized mean, (iii) difference in means, (iv) standardized difference in means, (v) median, (vi) difference in medians, (vii) maximum, (viii) ratio, and (ix) proportion with values above the population median. Determine sample sizes needed to estimate these quantities with these populations with a pre-specified margin of error.

2. Cluster and stratified sampling designs, self-weighting and non-self-weighting designs, optimal allocation, and design effects.

   Readings: Cochran (1977), Ch. 5, 5A, 9-11; Deaton (1997), Ch. 1; Thompson (2002), Ch. 11-13; United Nations (2005), Ch. VI-VII.

   Assignment: Write an R program to (i) carry out stratification via exact matching and minimum-distance clustering with auxiliary data; (ii) compute design effects for various sampling designs using census datasets; and (iii) carry out an optimal allocation analysis.

3. Designs for rare, hidden, or elusive subpopulations.

   Readings: Bernard et al. (2010); Goel and Salganik (2010); Kalton (2009); Landry and Shen (2005); McKenzie and Mistiaen (2010); Salganik and Heckathorn (2004); Salganik et al (2011a); Salganik et al (2011b); Thompson (2002), Ch. 15-19.

   Assignment: Write an R program to examine the performance of different methods of estimating the conditional prevalence of a rare attribute based on networked population data.

4. Longitudinal designs.

   Readings: Duncan and Kalton (1987); Kalton and Brick (1995); Lynn (2009), Ch. 1-2, 11-13; Rivers (2007).

   Assignment: Write R programs to (i) evaluate design effects of various longitudinal designs and (ii) implement a matching algorithm to construct a matched sample. Draft a memo describing two-three possibilities for the research design proposal assignment.

B. Field Experiment and Observational Study Designs

5. Complete unit randomized designs, minimal detectable effects, and power function.

   Readings: Bloom (1995, 2006); Cox and Reid (2000), Ch. 1-2; Duflo et al. (2008); Gerber and Green (2012), Ch. 1-3; Imbens and Rubin (2011), Ch. 1, 4, 6; List et al. (2010); Neyman (1923); Rubin (1990).

   Assignment: Write R functions to produce Ω for small experiments and random draws from Ω for large experiments, then make graphs of randomization distributions of ATEs and standardized effects for different types of outcome data. Relate these results to asymptotic power approximations.

6. Block-randomized designs including matched pairs, stratified designs, and Latin squares.

   Readings: Bruhn and McKenzie (2009); Cox and Reid (2000), Ch. 3-4; Gerber and Green (2012), Ch. 4; Greevy et al. (2004).

   Assignment: Write an R program to block via stratification and multivariate clustering (non-bipartite matching) and then carry out block randomization based on these results.

7. Cluster randomized designs and associated design effects.

   Readings: Bloom (2004); Bloom et al. (1999); Imai et al (2009); Imbens (2011); Raudenbush (1997); Raudenbush and Liu (2000); Schochet (2007).

   Assignment: Derive expression for the power function under non-balanced cluster randomized design and write an R program to examine power based on this design. Write an R program to compute design effects for cluster randomization with and without blocking for a set of population outcomes.

8. Designs for interference, spill-over, and social network effects.

   Readings: Aronow and Samii (2012); Duflo and Saez (2003); Gerber and Green (2012), Ch. 8; Halloran et al. (2009), Ch. 2, 13; Hirano and Hahn (2010); Levitt and List (2011); Sinclair et al. (2011).

   Assignment: Write an R program to analyze implications of alternative exposure models for power and bias given different types of "dependent happening" data.

9. Designs for mechanisms or mediated effects, factorial designs, and encouragement designs.

   Readings: Angrist et al. (1996); Ashraf et al. (2007); Baird et al. (2011); Cox and Reid (2000), Ch. 5-6; Frangakis and Baker (2001); Frangakis et al. (2002); Gerber and Green (2012), Ch. 6, 9-10; Gerber et al. (2010); Hirano et al. (2000); Imai et al (2012); Imbens and Rubin (2011), Ch. 24-25; Jo (2002).

   Assignment: Propose a factorial design to assess the mechanisms that may link a treatment to outcomes for a set of experiments that are provided. Write and implement R programs to analyze power for a factorial design given clustering and blocking. Write and implement R programs based on (i) Jo (2002) and (ii) Imai et al (2012).

10. Observational study designs I: matched samples, multilevel/multistage observational studies, and case-control designs.

    Readings: Hong and Raudenbush (2006); Imbens and Rubin (2011), Ch. 15; Kalton and Piesse (2007)*download; Korn and Graubard (DATE), Ch. X; Rosenbaum (2010), Ch. 1,3, 7-13; Sekhon (2009).

Assignment: Write R programs for power analysis to determine cluster sample size and construction of matched clusters for studies with 2 and then 3 treatment groups, using population data provided.

11. Observational study designs II: regression discontinuity designs, repeated observations, and downstream experiments.

    Readings: Dinardo et al. (2006); Gerber and Green (2012), Ch. 6; McKenzie (2011); Schochet (2009).

    Assignment: Write R programs to evaluate power for different sampling strategies for an RD design based on Schochet and to use simulation to evaluate McKenzie's claims. Propose a downstream analysis for an experiment from the ISPS and JPAL archives and use data from that experiment to propose a well-powered design.

C. Measurement Techniques

12. Concepts and principles for measurement; behavioral measurement techniques, including indirect observation, experiments, games, activities, and audits.

    Readings: Bertrand and Mullainathan (2005); Bertrand et al. (2007); Casey et al. (2011); Fearon et al. (2009); Habyarimana et al. (2007); Levitt and List (2007); Olken (2007); Samii (2011); Webb (2000), skim all; Zwane et al. (2011).

    Assignment: Work on your research design proposal, including meeting with instructor.

13. Index techniques, including axiomatic index construction (poverty/welfare) and latent index construction via scaling and item-response modeling (psychometric scales).

    Readings: Deaton (1997), Ch. 3; Heckman (2011); Jensen and Miller (2010); Revelle (2011), Ch. 6-8.

    Assignments: Write R code to construct poverty and psychometric indices from raw data. Study power given indices that vary in the number of items and validity levels.

14. Specialized questionnaire techniques, including list experiments, anchoring vignettes, and expectation elicitation.

    Readings: Delavande et al. (2010; 2011); Glynn (2010); King et al. (2004); Warner (1965).

    Assignment: Work on research design proposal.

D. Student Presentations

The remaining sessions will be committed to student presentations of research design proposals.

Sampling References

Bernard HR, et al. 2010. "Counting Hard-to-Reach Populations: The Network Scale-Up Method for Public Health." *Sexually Transmitted Infections*. 86(Supp 2): ii11-ii15.

Cochran WG. 1977. *Sampling Techniques* (3rd Ed.). New York: Wiley.

Deaton A. 1997. *The Analysis of Household Surveys*. Baltimore: Johns Hopkins Press.

Duncan GJ, Kalton G. 1987. "Issues of Design and Analysis of Surveys Across Time." *International Statistical Review*. 55(1):97-117.

Goel S, Salganik MJ. 2010. "Assessing Respondent Driven Sampling." *Proceedings of the National Academy of Sciences*. 107(15): 6743-6747.

Kalton G. 2009. "Methods for Oversampling Rare Subpopulations in Social Surveys." *Survey Methodology*. 35(2):125-141.

Kalton G, Brick JM. 1995. "Weighting Schemes for Household Panel Surveys." US Census Working Paper, No. 199.

Kasprzyk D, et al. 1989. *Panel Surveys*. New York: Wiley.

Landry P, Shen M. 2005. "Reaching Migrants in Survey Research: The Use of Global Positioning System to Reduce Coverage Bias in China." *Political Analysis*. 13:1:22. [Rare/Spatial]

Lynn P, ed. 2009. *Methodology of Longitudinal Surveys*. New York: Wiley.

McKenzie D, Mistiaen J. 2009. "Surveying Migrant Households: A Comparison of Census-Based, Snowball, and Intercept Point Surveys." *Journal of the Royal Statistical Society, Series A*. 172(2):339-360.

Rivers D. 2007. "Sampling for Web Surveys." Paper Presented at the 2007 ASA Joint Statistical Meeting.

Salganik MJ, Heckathorn DD. 2004. "Sampling and Estimation in Hidden Populations Using Respondent-Driven Sampling." *Sociological Methodology*. 34(1):193-239.

Salganik MJ, et al. 2011a. "Assessing Network Scale-Up Estimates for Groups Most at Risk of HIV/AIDS: Evidence from a Multiple-Method Study of Heavy Drug Users in Curitiba, Brazil." *American Journal of Epidemiology*. 174(10):1190-1196.

Salganik MJ, et al. 2011b. "The Game of Contacts: Estimating the Social Visibility of Groups." *Social Networks*. 33(1):70-78. http://www.sciencedirect.com/science/article/pii/S0378873310000559

Thompson SK. 2002. *Sampling, 2nd Edition*. New York: Wiley.

United Nations. 2005. *Household Sample Surveys in Developing and Transition Countries*. New York: United Nations Department of Social and Economic Affairs.

Field Experiments and Observational Studies References

Angrist J, et al. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association*. 91(434):444-455.

Aronow P, Samii C. 2012. "Estimating Average Causal Effects Under General Interference." Manuscript, Yale University and New York University.

Asraf N. et al. 2007. "Can Higher Prices Stimulate Product Use? Evidence from a Field Experiment in Zambia." NBER Working Paper, 13247.

Baird S, et al. 2011. "Cash or Condition? Evidence from a Cash Transfer Experiment." *Quarterly Journal of Economics*. 32(1):*download.

Bloom HS. 1995 "Minimum Detectable Effects: A Simple Way to Report the Statistical Power of Experimental Designs." *Evaluation Review*. 19(5):547-556.

Bloom HS. 2004. "Randomized Groups to Evaluate Place-Based Programs." MDRC Working Paper.

Bloom HS. 2006. "The Core Analytics of Randomized Experiments for Social Research." MDRC Working Paper.

Bloom HS, et al. 1999. "Using Cluster Random Assignment to Measure Program Impacts: Statistical Implications for the Evaluation of Education Programs." Manuscript, MDRC.

Bruhn M, McKenzie D. 2009. "In Pursuit of Balance: Randomization in Practice in Development Field Experiments." *American Economic Journal: Applied Economics*. 1(4):200-232.

Cox DR, Reid N. 2000. *Theory of the Design of Experiments*. Boca Raton: CRC Press.

Dinardo J, et al. 2006. "Constructive Proposals for Dealing with Attrition: An Empirical Example." Working paper, University of Michigan and NBER.

Duflo E, Saez E. 2003. "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiments. *Quarterly Journal of Economics*. 118:815-842.

Duflo E, et al. 2008. "Using Randomization in Development Economics Research: A Toolkit." *Handbook of Development Economics*. 4:3895-3962.

Frangakis CE, Baker SG. 2001. "Compliance Subsampling Designs for Comparative Research: Estimation and Optimal Planning." *Biometrics*. 57(3):899-908.

Frangakis CE, et al. 2002. "Clustered Encouragement Designs with Individual Noncompliance: Bayesian Inference with Randomization, and Application to Advance Directive Forms." *Biostatistics*. 3(2):147-164.

Gerber AS, et al. 2010. "Baseline, Placebo, and Treatment: Efficient Estimation for Three-Group Experiments." *Political Analysis*. 18:297-315.

Gerber, Alan, and Donald Green. 2012. *Field Experiments: Design, Analysis, and Interpretation*. Manuscript, Yale University.

Greevy R, et al. 2004. "Optimal Multivariate Matching Before Randomization." *Biostatistics*. 5(2):263-272.

Hirano K, Hahn J. 2010. "Design of randomized experiments to measure social interaction effects." *Economics Letters*. 106(1):51-53.

Halloran ME, et al. 2009. *Design and Analysis of Vaccine Studies*. New York: Springer.

Hirano K, et al. 2000. "Assessing the Effect of an Influenza Vaccine in an Encouragements Design." *Biostatistics*. 1(1):69-88.

Hong G, Raudenbush SW. 2006. "Evaluating Kindergarten Retention Policy: A Case Study of Causal Inference for Multilevel Observational Data." *Journal of the American Statistical Association*. 101(475):901-910.

Imai K, et al. 2009. "The Essential Role of Pair Matching in Cluster-Randomized Experiments, with Application to the Mexican Universal Health Insurance Evaluation." *Statistical Science*. 24(1):29-53.

Imai K. et al. 2012. "Experimental Designs for Identifying Causal Mechanisms." *Journal of the Royal Statistical Society, Series A*. Forthcoming.

Imbens GW. 2011. "Experimental Designs for Unit and Cluster Randomized Trials." Manuscript, Harvard University.

Imbens GW, Rubin DB. 2011. *Causal Inference in Statistics and the Social Sciences*. Manuscript, Harvard University..

Jacob R, et al. 2009. "New Empirical Evidence for the Design of Group Randomized Trials in Education." MDRC Working Paper.

Jo B. 2002. "Statistical Power in Randomized Interventions Studies with Noncompliance." *Pscyhological Methods*. 7(2):178-193.

Kalton G, Piesse A. 2007. "Survey Research Methods in Evaluation and Case-Control Studies." *Statistics in Medicine*. 26(8):1675-1687.

Levitt SD, List JA. 2011. "Was There Really a Hawthorne Effect at the Hawthorne Plant? An Analysis of the Original Illumination Experiments." *American Economic Journal: Applied Economics*. 3:224-238.

List JA, et al. 2010. "So You Want to Run an Experiment, Now What? Some Simple Rules of Thumb for Optimal Experimental Design." *NBER Working Paper Series*. No. 15701.

McKenzie D. 2011. "Beyond Baseline in Experiments: The Case for More T in Experiments." *World Bank Policy Research Working Papers*, 5639.

Neyman J. 1923. "On the Application of Probability Theory to Agricultural Experiments: Essay on Principles, Section 9." Reprinted in *Statistical Science*. 5(4)[1990]:465-472.

Raudenbush SW. 1997. "Statistical Analysis and Optimal Design of Cluster Randomized Trials." *Psychological Methods*. 2(2):173-185.

Raudenbush SW, Liu X. 2000. "Statistical Power and Optimal Design for Multisite Randomized Trials." *Psychological Methods*. 5(2):199-213.

Rosenbaum PR. 2010. *Design of Observational Studies*. New York: Springer.

Rubin DB. 1990. "Comment: Neyman (1923) and Causal Inference in Experiments and Observational Studies." *Statistical Science*. 5(4):472-480.

Schochet PZ. 2007. "Statistical Power for Random Assignment Evaluation of Educational Programs." *Journal of Educational and Behavioral Statistics*. 33(1):62-87.

Schochet PZ. 2009. "Statistical Power for Regression Discontinuity Designs in Education." *Journal of Educational and Behavioral Statistics*. 34(2):238-266.

Sekhon J. 2009. "Opiates for the Matches." *Annual Review of Political Science*. 12:487-508.

Sinclair B, et al. 2011. "Detecting Spillover Effects: Design and Analysis of Multi-level Experiments." Manuscript, University of Chicago, Harvard University, and Yale University.

Measurement References

Bertrand M, Mullainathan S. 2005. "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *American Economic Review*. 94(4):991-1013.

Bertrand M. et al. 2007. "Obtaining a Driver's License in India: An Experimental Approach to Studying Corrpution." *Quarterly Journal of Economics*. 122(4):1639-1676. [Audit]

Casey K, et al. 2011. "Reshaping Institutions: Evidence on Aid Impacts Using a Pre-Analysis Plan." Manuscript, Stanford University, MIT, and University of California-Berkeley.

Deaton A. 1997. *The Analysis of Household Surveys*. Baltimore: Johns Hopkins Press.

Delavande A. et al. 2010. "Eliciting Probabilistic Expectations With Visual Aids in Developing Countries: How Sensitive Are Answers to Variations in Elicitation Design?" World Bank Policy Research Working Paper, 5458.

Delavande A. et al. 2011. "Measuring Subjective Expectations in Developing Countries: A Critical Review." *Journal of Development Economics*. 94:151-163.

Fearon JD, et al. 2009. "Can Development Aid Contribute to Social Cohesion and Civil War? Evidence from a Field Experiment in Post-Conflict Liberia." *American Economic Review: Papers and Proceedings*. 99(2):287-291.

Glynn AN. 2010. "What Can We Learn with Statistical Truth Serum? Design and Analysis of the List Experiment." Manuscript, Harvard University.

Habyarimana J, et al. 2007. "Why Does Ethnic Diversity Undermine Public Goods Provision?" *American Political Science Review*. 101(4):709-725.

Heckman J. 2011. "Integrating Personality into Economics." NBER Working Paper, 17378.

Jensen RT, Miller NH. 2010. "A Revealed Preference Approach to Measuring Hunger and Undernutrition." *NBER Working Paper Series*. No. 16555.

King G, et al. 2004. "Enhancing the Validity and Cross-Cultural Comparability of Measurement in Survey Research." *American Political Science Review*. 98(1):191-207.

Levitt SD, List JA. 2007. "What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World?" *Journal of Economic Perspectives*. 21(2):153-174.

Olken B. 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy*. 115(2):200-249.

Revelle W. 2011. *An Introduction to Psychometric Theory with Applications in R*. Manuscript, Northwestern University.

Samii C. 2011. "Do Quotas Reduce or Exacerbate Ethnic Conflict? Evidence from Burundi's Military." Manuscript, New York University.

Warner SL. 1965. "Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias." *Journal of the American Statistical Association*. 60:63-69.

Webb EJ, et al. 2000. *Unobtrusive Measures, Revised Edition*. Thousand Oaks: Sage.

Zwane AP, et al. 2011. "Being Surveyed Can Change Later Behavior and Related Parameter Estimates." *Proceedings of the National Academy of Sciences*. 108(5):1821-1826.