POL-GA.3202

Quantitative Field Methods

Fall 2016

Instructor: Professor Cyrus Samii
NYU Politics Department
19 West 4th Street, Office 424
cdsamii@gmail.com
Office hours: See online sign-up form.

Overview

This is a course about making data. Most of the methodological training that you have had up until now has been about analyzing pre-existing data. In this class, you will study methods for performing *ex ante* analyses of research design options to determine how to test hypotheses efficiently and with minimal bias, subject to practical constraints. You will study sampling, experimental design, observational study design, and measurement. The class will start of with a series of lectures on design-based inference and formal frameworks for analyzing research designs. Following that will be a series of design workshops in which students will lead presentations on state-of-the-art research designs. An end of term assignment involves writing an original research design for a study that you define and then presenting it to the class.

Prerequisites and Eligibility

The prerequisites for the class are that students have working knowledge of mathematical, programming, and other quantitative techniques at the level of POL-GA 1250 and POL-GA 1251 (Quant I and II). Because of resource constraints, only PhD students will be allowed to take the course for a grade. Priority access is given to political science PhD students beyond the first year. If space permits, PhD students from other departments are welcomed so long as they meet the prerequisites. Admission is at the discretion of the instructor based on consideration of prerequisites. Others may audit, space permitting. Students with special needs should come to office hours or schedule an appointment with the instructor to discuss possible accommodation.

Texts and Software

The course will draw on various texts and resources throughout the semester. The following texts are recommended for acquisition: Thompson (2012), Gerber and Green (2012), and Rosenbaum (2009, available in PDF from Springer). Sections from these books and other readings are indicated below. Soft copies of articles and other digital resources will be made available by the instructor through a Dropbox linked to the course webpage. For assignments, you will have to visit research web pages and data archives to obtain supplemental appendices and replication files. The one item that I cannot find digitally is Nisbett and Cohen (1996), and so you are requested to acquire a copy.

As for software, instructor support will be given for R and Stata.

Requirements

Leading In-Class Discussions on Published Papers

Students will be assigned to lead class discussions on published papers. Each discussion will present the paper in a manner that focuses on the hypotheses, research design, and the interrelation between the two. Presentations should use slides address the following:

- Present relevant theoretical framework(s) and state hypotheses in terms of theoretical model parameters, including a discussion of the "debate" to which the study is contributing;
- Explain the research design, including sample selection criteria, treatment operationalization, and method of assignment;
- Explain how key parameters, treatments, and outcomes are operationalized, how they are measured, and any complications that have to be overcome in obtaining the data;
- Show the data analysis specification, how to interpret the quantities to be estimated from data in terms of parameters in the theoretical hypotheses, any complications or ambiguities, and then explain testing procedures;
- Present a power analysis using either replication data from the study or a simulation based on values reported in the paper.

The presentation will typically require going beyond what is presented in the published paper and examining appendices and other supplementary information. Each presentation should include about 40 minutes worth of material, allowing for about 10-20 minutes of interruptions, questions, and discussion.

The number of presentations each student does depends on the number of people in the class and will account for 40% of your grade. Students can work in small groups to complete the assignments, although this must be approved by the instructor in advance.

Research Design Proposal

Each student will come up with a research design proposal for a field experiment or observational study. The design proposal will include a full discussion and justifications for the treatment assignment method or identification strategy, power analysis, sampling or data gathering plan, and measurement techniques. I will provide guidance on the structure of the proposal. It should be the caliber of a proposal that you would submit to a major scientific grant competition, such as the NSF. 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 (e.g., it may be for a dissertation prospectus) or "enhanced" replication of a study. 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 during a class session toward the end of the semester, and then the final proposal will be due one week after the last class.

The elements of the proposal assignment, including presentation of a preliminary proposal, presentation of a final proposal, and submission of a final proposal in proper format, will account for 50% of your grade. Students can work in small groups to complete the assignments, although this must be approved by the instructor in advance.

Attendance and participation

Attendance and participation in class discussions is required and counts toward 10% of your grade.

Sessions

1 Fundamentals I: design-based inference, sampling, and power

Date: 9/7
Readings: Thompson (2012, Ch. 1-4); Gerber and Green (2012, Ch. 1-3); Bloom (1995); Duflo et al. (2008); McConnell and Vera-Hernandez (2015).
Presenters: Instructor.

2 Fundamentals II: design effects, clustering, and covariates

Date: 9/14
Readings: Thompson (2012, Ch. 11-13); Gerber and Green (2012, Ch. 4); Bruhn and McKenzie (2009); Athey and Imbens (2016).
Presenters: Instructor.

3 Fundamentals III: observational study design

Date: 9/14
Readings: Rosenbaum (2009, Ch. 1, 3, 5, 7-8, 14-15, 19).
Presenters: Instructor.

NB: No class 9/28

4 Complex outcomes and credible inference

Date: 10/5
Readings: Casey et al. (2012, data available); Fearon et al. (2015, data available).
Presenters: TBD.

5 Mechanisms

Date: 10/12
Readings: Ashraf et al. (2010, data available); Gerber et al. (2010, data available); Olken (2007, data available); *for reference*: Gerber and Green (2012, Ch. 5-6, 9-10).
Presenters: TBD.

6 Spillover and networks I

Date: 10/19
Readings: Miguel and Kremer (2004, data available); Paluck et al. (2016, use simulation); *for reference*: Gerber and Green (2012, Ch. 8).
Presenters: TBD.

7 Spillover and networks II

Date: 10/26
Readings: Banerjee et al. (2013, data available); Nickerson (2008, use simulation).
Presenters: TBD.

8 Research design proposal presentations

Date: 11/2
Readings: None (prepare proposal presentation).
Presenters: All students.

9 Accessing hard-to-reach populations

Date: 11/9
Readings: Salganik et al. (2011, including appendix); Khabbazian et al. (2016); McKenzie and Mistiaen (2009); Paz-Bailey et al. (2011).
Presenters: TBD. NB: presentations will follow a different format for these.

10 Dynamic experiments

Date: 11/16
Readings: Gerber et al. (2011, data available); Hahn et al. (2011, replicate simulations).
Presenters: TBD.

NB: No class 11/23

11 Multi-mode research

Date: 11/30
Readings: Nisbett and Cohen (1996); Grimmer et al. (2012).
Presenters: TBD.

12 Research design presentations

Date: 12/7
Presenters: TBD.

13 Research design presentations

Date: 12/14
Presenters: TBD.

References

Ashraf, N., J. Berry, and J. M. Shapiro (2010). Can higher prices stimulate product use? evidence from a field experiment in Zambia. *American Economic Review 100*, 2383–2413.

Athey, S. and G. W. Imbens (2016). The econometrics of randomized experiments. Tyepscript, Stanford University. Available at https://www.povertyactionlab.org/handbook-field-experiments.

Banerjee, A. V., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2013). The diffusion of microfinance. *Science 341*(1236498), 363–373.

Bloom, H. S. (1995). Minimum detectable effects: A simple way to report the statistical power of experimental designs. *Evaluation Review 19*(5), 547–556.

Bruhn, M. and D. McKenzie (2009). In pursuit of balance: Randomization in practice in development field experiments. *American Economic Journal: Applied Economics 1*(4), 200–232.

Casey, K., R. Glennerster, and E. Miguel (2012). Reshaping institutions: Evidence on aid implacts using a pre-analysis plan. *Quarterly Journal of Economics 127*(4), 1755–1812.

Duflo, E., R. Glennerster, and M. Kremer (2008). Using randomization in development economics research: A toolkit. *Handbook of Development Economics 4*, 3895–3962.

Fearon, J. D., M. Humphreys, and J. M. Weinstein (2015). How does development assistance affect collective action capacity? results from a field experiment in post-conflict liberia. *American Political Science Review 109*(3), 450–469.

Gerber, A. S., J. G. Gimpel, D. P. Green, and D. R. Shaw (2011). How large and long-lasting are the persuasive effects of television campaign ads? Results from a randomized field experiment. *American Political Science Review 105*(1), 135–150.

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

Gerber, A. S., G. A. Huber, and E. Washington (2010). Party affiliation, partisanship, and political beliefs: A field experiment. *American Political Science Review 104*(4), 720–744.

Grimmer, J., S. Messing, and S. J. Westwood (2012). How words and money cultivate a personal vote: The effect of legislator credit claiming on constituent credit allocation. *American Political Science Review 106*(4), 703–719.

Hahn, J., K. Hirano, and D. Karlan (2011). Adaptive experimental design using the propensity score. *Journal of Business and Economic Statistics 29*(1), 96–108.

Khabbazian, M., B. Hanlon, Z. Russek, and K. Rohe (2016). Novel sampling designs for respondent-driven sampling. arXiv:1606.00387v2 [stata.ME].

McConnell, B. and M. Vera-Hernandez (2015). Going beyond simple sample size calculations: A practitioner's guide. IFS Working Paper W15/17.

McKenzie, D. J. and J. Mistiaen (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.

Miguel, E. and M. Kremer (2004). Worms: Identifying impacts on education and health in the presence of treatment externalities. *Econometrica 72*(1), 159–217.

Nickerson, D. W. (2008). Is voting contagious? evidence from two field experiments. *American Political Science Review 102*(1), 49–57.

Nisbett, R. E. and D. Cohen (1996). *Culture Of Honor: The Psychology Of Violence In The South*. Boulder, CO: Westview Press.

Olken, B. (2007). Monitoring corruption: Evidence from a field experiment in indonesia. *Journal of Political Economy 115*(2), 200–249.

Paluck, E. L., H. Shepherd, and P. M. Aronow (2016). Changing climates of conflict: A social network experiment in 56 schools. *Proceedings of the National Academy of Science 113*(3), 566–571.

Paz-Bailey, G., J. Jacobsen, M. Guardado, F. Hernandez, A. Nieto, M. Estrada, and J. Creswell (2011). How many men who have sex with men and female sex workers live in El Salvador? using respondent-driven sampling and capture recapture to estimate population sizes. *Sexually Transmitted Infections* (online first).

Rosenbaum, P. R. (2009). *Design of Observational Studies*. New York, NY: Springer.

Salganik, M. J., D. Fazito, N. Bertoni, A. H. Abdo, M. B. Mello, and F. I. Bastos (2011). 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.

Thompson, S. K. (2012). *Sampling, Third Edition*. Hoboken, NJ: Wiley.