Overview

This is a course about making data for quantitative social science research. Most of the methodological training that you have had up until now has been about analyzing data. But where do you get the data? Options include surveys, administrative records, or other specialized observation techniques (e.g., remote sensing systems, computerized interaction systems, etc.). The data might simply describe prevailing conditions or record outcomes of researcher-controlled experiments. This class provides a foundation for designing research that generates the kind of data that you need to test a social science hypothesis. We consider sampling, experimental design, observational study design, measurement, and online data collection. We go back and forth between theory and applications. Assignments involve practicing quantitative research design techniques. An end of term assignment involves writing an original research design for a study that you define.

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 (http://cyrussamii.com/?page_id=1360). As for software, students are free to use whatever they want but instructor support will be given for R and Stata.
Requirements

Occasional Assignments
I will provide assignments every few weeks that have you practice research design tasks (e.g., power analyses or complex sampling). These assignments account for 40% of your grade. Students can work in small groups to complete the assignments.

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 proposal 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.

Topics

1 Fundamentals I: sampling and power
Readings: Thompson (2012, Ch. 1-4); Gerber and Green (2012, Ch. 1-3); Bloom (1995); Duflo et al. (2008, Sec. 1-5).

Other references: Cochran (1977, Ch. 1-4).

2 Fundamentals II: design effects
Readings: Thompson (2012, Ch. 11-13); Kalton et al. (2005); Gerber and Green (2012, Section 3.6.2); Duflo et al. (2008, Sec. 4.3, 7.1).

Other references: Cochran (1977, Ch. 5-5A,9-9A); Donner et al. (1981); Pettersson and do Nascimento Silva (2005).
3 Sampling hard to reach populations

Readings: Salganik et al. (2011, including appendix); Salganik and Heckathorn (2004).

Other references: Gile (2011); Gile and Handcock (2010); Goel and Salganik (2010); McKenzie and Mistiaen (2009); Paz-Bailey et al. (2011); Thompson (2012, Ch. 15-19).

4 Designing efficient experiments with covariates

Readings: Bruhn and McKenzie (2009); Duflo et al. (2008, 4.4.-4.5, 6.1, 7.3-7.4); Gerber and Green (2012, Ch. 4)

Other references: Barrios (2014); Basu (1980, with rejoinder); Gelman et al. (2004, Sec. 7.5-7.6); Greevy et al. (2004); Imai et al. (2009); Imbens (2011); Kasy (2013); Rubin (1980).

5 Designing experiments for causal explanation

Readings: Gerber and Green (2012, Ch. 5-6, 9-10); Imai et al. (2013).

Other references: Ashraf et al. (2010); Baird et al. (2011); Duflo et al. (2008, Sec. 6.2-6.3, 8.1); Gerber and Green (2000); Gerber et al. (2010); Hainmueller and Hopkins (2012); Jo (2002); Kramer et al. (2001); Ludwig et al. (2011); Manzi (2012).

6 Designing experiments to study spillover and social interaction

Readings: Gerber and Green (2012, 8); Aronow and Samii (2013).

Other references: Aral and Walker (2011); Banerjee et al. (2012); Bond et al. (2012); Chen et al. (2010); Duflo and Saez (2003); Hirano and Hahn (2010); Nickerson (2008); Paluck and Shepherd (2012); Sinclair et al. (2012).

7 Designing observational studies I: matching & case-control

Readings: Daly et al. (2014); Rosenbaum (2009, Ch. 1,3,7-8); Van der Laan and Rose (2011, Ch. 13-14, available in PDF from Springer)

Other references: Cochran (1983); Duflo et al. (2008, Sec. 2.3.1); Holland and Rubin (1988); Hong and Raudenbush (2006); Manski (1995, Ch. 5) Rosenbaum (1984); Young et al. (2009)

8 Designing observational studies II: efficient IV, DID, and RD

Readings: Baiocchi et al. (2010); Baiocchi et al. (2012); Zhang et al. (2011); Keele and Titiunik (2013).

Other references: Duflo et al. (2008, Sec. 2.3.2-2.3.3).
9  Measurement

Readings: Lecture notes and then skim as desired from other references below.

Other references: Beath et al. (2012); Bertrand et al. (2007); Bertrand and Mullainathan (2004); Deaton (1997); Delavande et al. (2011); Dovidio et al. (2010); Gilligan et al. (2014); Glaeser et al. (2000); Grosh and Glewwe (2000); Habyarimana et al. (2007); Henrich et al. (2005); Humphreys et al. (2012); King et al. (2004); Nisbett and Cohen (1996); Olken (2007, 2009); Olken and Barron (2009); Tourangeau and Yan (2005); van der Ark (2007).

10  Online research

Readings: Lecture notes and then skim as desired from other references below.

Other references: Aday et al. (2010); Bakshy et al. (2011); Berinsky et al. (2012); Bond et al. (2012); Cho et al. (2012); Farrell (2012); Grimmer et al. (2012); Rivers (2006, 2007); Rivers and Bailey (2009)

11  Research design proposal presentations I

12  Research design proposal presentations II
References


