Political Science 590
Political Experiments: Design & Analysis, Part I

Jake Bowers and Corrine McConnaughy

Fall 2012

General Information

Overview   Randomized interventions allow political scientists to claim that comparisons are causal: randomization allows us the ability to characterize counterfactual comparisons - to say how the treated group would have responded had treatment been withheld. Randomization also allows us to test hypotheses about causal comparisons without requiring large samples or probability models of outcomes. Randomized experiments thus promise to simplify our lives and to enable clear answers to important substantive and theoretical questions without requiring much in the way of extraneous justification. Yet, there is an art in designing and analyzing a randomized experiment. Only the simplest of experiments can be analyzed simply. And, in political science, we may not be able to directly require subjects to be exposed to a dose of our treatment and may need to work indirectly, using the randomization as an instrument to manipulate a dose, perhaps at a distance, perhaps weakly. In this course, we will practice designing experiments and making statistical inferences from the resulting data. Building on the most basic foundations (what is an experiment? why randomize?) we will engage with questions about clustered assignment, blocking, multi-valued treatments, and instrumental variables. Our hope is to prepare students to face their own design and analysis decisions of lab and field experiments with creativity and with a strong foundation in the design and statistical literatures on these topics.

Over the two terms of this course and the winter break, we anticipate that you will design, field, and analyze data from, a randomized experiment.

This course is brand new. Expect changes to the syllabus and format of the course as we go along. We welcome constructive comments and suggestions.

Where/When   Part I will meet Fridays 11:00–1:00 PM Central Time / 12:00–2:00 PM Eastern Time on Oct. 26, Nov. 2, 9, 16, 30, Dec. 7.

Part II will meet Wednesdays 2:30pm-4:30 PM Eastern Time / 1:30pm-3:30 PM Central Time on Jan. 23, 30, Feb. 6, 13, 20, 27, March 6.

Who   This course is co-designed by Corrine McConnaughy [1] and Jake Bowers[2]. Bowers will lead the in-class discussions and work in the Fall term and McConnaughy will lead the Spring term.

Office Hours   Bowers' Fall 2012 office hours are 3–4:30 on Tuesdays by appointment. I am very happy to meet with you. If you know in advance that you want to talk during office hours, please email me to reserve a 20 minute slot. I am happy to video chat, too. I do not have a camera on my desktop at the office, but I can arrange to have a camera if we schedule a video chat in advance.

McConnaughy's office hours are TBA.

---

[1] mailto:mcconnaughy.3@osu.edu, http://politicalscience.osu.edu/faculty/mcconnaughy.3/index.htm
[2] mailto:jwbowers@illinois.edu, http://jakebowers.org.

Goals & Expectations

In-Class Most class meetings will involve practice analyzing data from randomized studies in addition to short lectures from the professors. So, you will need to bring laptops to class. We encourage you to work in groups.

Final Paper We will help you complete your final paper by asking for parts of it as the course goes along.

First, we will ask you to write no more than about two pages stating and motivating a substantive and/or theoretical question that you'd like to pursue with a randomized experiment. In this first assignment, we will also ask you to explain why a randomized experiment offers you "real leverage" for your effort to get answers to your question.

Second, we will ask you to add no more than about five pages to your revised question motivation section specifying the manipulation and the outcome and specifically showing how said manipulation and outcome directly address your substantive question.

Third, over the winter break we will ask you to add a few more pages completing the details of the design: are there background variables that would enable you to use moderating effects to answer your substantive question more conclusively? Are there background variables that would seriously enhance your precision of estimation? Are you concerned about non-compliance or attrition or interference? If so, how do you plan to address such problems?

Fourth, also over break, we will ask you to write down your hypotheses and analysis plan. This will protect you from charges of data snooping and also, perhaps, cause you to change your design as you discover that you require different variables or design features given your analysis plans.

Fifth, we will ask you to fill out an IRB form. [Since this paper is for a class and not for publication we don't think you'll need to formally go through an IRB procedure.]

Sixth and finally, we will ask you to combine the previous work with a report of the results themselves, including, perhaps a new section in which you explore the data for the purposes of suggesting future hypotheses, theories, and designs.

General We expect you to engage with this class with an open-mind, curiosity and energy.

We expect that students at each site will meet at least one other time per week to finish in-class work, discuss readings, and projects.

We expect you to have completed the assigned readings before coming to class.

We expect you to be in charge of your own learning: you should ask questions when you do not understand something.

We expect you to attend the class.

We require that you give us written work in pdf format.

We expect that all papers written in this class will assume familiarity with the principles of good writing in Becker (1986).

Late Work If you'd prefer to spend more time using the paper assignment in this class to learn more, I am happy for you to take that time as long as you let me know in advance.

Incompletes   Incompletes are fine in theory but terrible at the University of Illinois in practice. I urge you to avoid an incomplete in this class. If you must take an incomplete, you must give me *at least* 2 months from the time of turning in an incomplete before you can expect a grade from me. This means that if your fellowship, immigration status, or job depends on erasing an incomplete in this class, you should not leave this incomplete until the last minute.

Grading   Because moments of evaluation are also moments of learning in this class (and not moments of competition or ranking), we do not curve. If you all perform at 100%, then you will all receive As. We will provide rubrics so that you can check to see whether you have completed all of the work that we expect; work that we will be grading.

Books

Required: Gerber, A. and Green, D. (2012). *Field experiments: Design, analysis, and interpretation*. WW Norton

Recommended: Rosenbaum, P. R. (2010). *Design of Observational Studies*. Springer (pdf free to download from campus ip addresses or via campus library springerlink subscriptions: http://www.springerlink.com/content/978-1-4419-1212-1/contents/)

Becker, H. S. (1986). *Writing for Social Scientists: How to Start and Finish Your Thesis, Book, or Article*. University of Chicago Press

Becker, H. S. (1998). *Tricks of the trade : how to think about your research while you're doing it*. University of Chicago Press, Chicago, Ill

Angrist, J. and Pischke, J. (2009). *Mostly harmless econometrics: an empiricist's companion*. Princeton Univ Pr (This seems to be the main source for discussions of the analysis of experiments in economics)

Computing

In this class, we will be using the R statistical language. You are free to use other languages although we suspect you will find it easier to learn R unless you are already a code ninja in some other language and enjoy the thought of writing your own statistical routines.

Schedule

Note: This schedule is preliminary and subject to change. If you miss a class make sure you contact me or one of your colleagues to find out about changes in the lesson plans or assignments.

Important: Come to class with your laptops (if you have them) with R installed.

Oct 26 – Why experiment? What experiment? Context and Concepts

Introduction to the idea of potential outcomes and the use of counterfactuals as heuristics.

Randomization for "statistical control" as opposed to "scientific control".

Read: Kinder and Palfrey (1993)

Rosenbaum (2002, Chap 2) (especially 2.4.1; 2.4.5; 2.5.1; 2.5.2)

Rubin (2004)

Due in Class: Come to class prepared to tell us about how you think a randomized experiment can help you answer an important substantive question or shed light on an important substantive theory.

Do: Engage with a short introduction to the course. Iron out technical problems.

Recommended: Gerber and Green (2012, Chap 1)

Fisher (1935, Chap 2)

Sekhon (2008)

Holland (1986) (including discussion) — statistical versus scientific control

(Rosenbaum, 2010, Chap 2)

Nov 2 – Treatments and Outcomes

How do we know what a manipulation means? How do we know what a relationship between manipulation and outcome means? The importance of concepts and measurement and theory: treatment as concept, and conceptualizing the causal relationship between treatment and outcomes.

Reading:

Application: White (2007)

Gerber and Green (2012, Chap 2) (on average treatment effects, excludability)

Adcock and Collier (2001) (esp their discussion of their Figure 1)

Answer: What is your question? Why would an experiment help you answer your question?

Do: Show that the difference of means in the sample effectively estimates the difference of means of potential outcomes. Understand a sampling perspective on in-sample statistical inference about causal effects using the ATE.

Recommended: Neyman (1990) and Rubin (1990)

Lohr (1999, Section 2.7)

Nov 9 – Populations and Subjects

What is the population? Why choose the subjects you choose? Concepts of sampling. Local versus Global (or Sample versus Population) treatment effects? Introduction to the idea of an instrument.

Reading: Deaton (2010)

Imbens (2010)

Sears (1986)

Henrich et al. (2010)

Recommended: A discussion about populations and laboratories in the context of an experiment on racial framing:

Huber and Lapinski (2006)

Huber and Lapinski (2008)

Mendelberg (2008a)

Mendelberg (2008b)

The third article in the debate about field experiments in economics:

Heckman and Urzua (2010)

Due: Question statement and explanation for why an experiment will help you answer your question. What is it about experiments that will help you answer your question?

Nov 16 – Power

What unobserved quantities do we want to make inferences about? What are we testing and why? What is power? How do pre-treatment covariates help us enhance power?

Reading:

Application: Panagopoulos and Green (2008)

Gerber and Green (2012, Chap 3)

Moore (2012)

Duflo et al. (2007, esp Section 4)

Kaplan (2012, Chap 13)

Recommended: Nickerson (2005)

Bowers (2011)

Nov 30 – Moderation and Heterogeneous Treatment Effects

Design to assess interactions of treatment with pre-treatment covariates (aka "moderators"). Challenges of statistical inference for moderated effects. Moderation is not mediation.

Reading: Gerber and Green (2012, Chap 4 & 9)

Recommended: Gerber and Green (2012, Chap 10)

bullock and shang e. ha (2011)

Imai et al. (2010a,b)

Dec 7 – Encouragement Designs and Instruments

Random assignment and non-random compliance with the treatment may still allow well-defined causal effects and statistical inferences about same.

Reading:

Application: Gerber and Green (2000)

Angrist et al. (1996)

Gerber and Green (2012, Chap 5 & 6)

Recommended: Hansen and Bowers (2009)

Imbens and Rosenbaum (2005)

Dec 14 – No Class Meeting

Jake will be available for (virtual) meetings this day. Feel free to email him to schedule time if you would like to discuss your project.

Dec 21 – No Class Meeting – Research Designs Due

We will be handing out a rubric with the elements of a research design in advance of this date.

Meanwhile see

Gerber and Green (2012, Chap 12)

References

Adcock, R. and Collier, D. (2001). Measurement validity: A shared standard for qualitative and measurement validity: A shared standard for qualitative and quantitative research. *American Political Science Review*, 95(3):529–546.

Angrist, J. and Pischke, J. (2009). *Mostly harmless econometrics: an empiricist's companion*. Princeton Univ Pr.

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

Becker, H. S. (1986). *Writing for Social Scientists: How to Start and Finish Your Thesis, Book, or Article*. University of Chicago Press.

Becker, H. S. (1998). *Tricks of the trade : how to think about your research while you're doing it*. University of Chicago Press, Chicago, Ill.

Bowers, J. (2011). Making effects manifest in randomized experiments. In Druckman, J. N., Green, D. P., Kuklinski, J. H., and Lupia, A., editors, *Cambridge Handbook of Experimental Political Science*, chapter 32. Cambridge University Press, New York, NY.

bullock, j. G. and shang e. ha (2011). Mediation analysis is harder than it looks. In Druckman, J. N., Green, D. P., Kuklinski, J. H., and Lupia, A., editors, *Cambridge Handbook of Experimental Political Science*, chapter 35. Cambridge University Press, New York, NY.

Deaton, A. (2010). Instruments, randomization, and learning about development. *Journal of Economic Literature*, 48(2):424–455.

Duflo, E., Glennerster, R., and Kremer, M. (2007). Using randomization in development economics research: A toolkit. *Handbook of development economics*, 4:3895–3962.

Fisher, R. (1935). *The design of experiments. 1935*. Oliver and Boyd, Edinburgh.

Gerber, A. and Green, D. (2012). *Field experiments: Design, analysis, and interpretation*. WW Norton.

Gerber, A. S. and Green, D. P. (2000). The effects of canvassing, telephone calls, and direct mail on voter turnout: A field experiment. *American Political Science Review*, 94:653–663.

Hansen, B. B. and Bowers, J. (2009). Attributing effects to a cluster randomized get-out-the-vote campaign. *Journal of the American Statistical Association*, 104(487):873—885.

Heckman, J. and Urzua, S. (2010). Comparing iv with structural models: What simple iv can and cannot identify. *Journal of Econometrics*, 156(1):27–37.

Henrich, J., Heine, S. J., and Norenzayan, A. (2010). The weirdest people in the world? *BEHAVIORAL AND BRAIN SCIENCES*, 33:61–135.

Holland, P. W. (1986). Statistics and causal inference (with discussion). *Journal of the American Statistical Association*, 81:945–970.

Huber, G. and Lapinski, J. (2006). The race card revisited: Assessing racial priming in policy contests. *American Journal of Political Science*, 50(2):421–440.

Huber, G. and Lapinski, J. (2008). Testing the implicit-explicit model of racialized political communication. *Perspectives on Politics*, pages 125–134.

Imai, K., Keele, L., and Tingley, D. (2010a). A general approach to causal mediation analysis. *Psychological methods*, 15(4):309.

Imai, K., Keele, L., and Yamamoto, T. (2010b). Identification, inference and sensitivity analysis for causal mediation effects. *Statistical Science*, 25(1):51–71.

Imbens, G. (2010). Better late than nothing: Some comments on deaton (2009) and heckman and urzua (2009). *Journal of Economic Literature*, 48(2):399–423.

Imbens, G. and Rosenbaum, P. (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.

Kaplan, D. (2012). *Statistical Modeling A Fresh Approach*. Daniel Kaplan, Macalester College, St. Paul, MN, second edition.

Kinder, D. and Palfrey, T. (1993). On behalf of an experimental political science. *Experimental foundations of political science*, pages 1–39.

Lohr, S. (1999). *Sampling: Design and Analysis.* Brooks/Cole.

Mendelberg, T. (2008a). Racial priming: Issues in research design and interpretation. *Perspectives on Politics*, 6(1):135.

Mendelberg, T. (2008b). Racial priming revived. *Perspectives on Politics*, 6(01):109–123.

Moore, R. (2012). Multivariate continuous blocking to improve political science experiments. *Political Analysis*.

Neyman, J. (1923 [1990]). On the application of probability theory to agricultural experiments. essay on principles. section 9 (1923). *Statistical Science*, 5:463–480. reprint. Transl. by Dabrowska and Speed.

Nickerson, D. (2005). Scalable protocols offer efficient design for field experiments. *Political Analysis*, 13(3):233.

Panagopoulos, C. and Green, D. (2008). Field experiments testing the impact of radio advertisements on electoral competition. *American Journal of Political Science*, 52(1):156–168.

Rosenbaum, P. R. (2002). *Observational Studies*. Springer-Verlag, second edition.

Rosenbaum, P. R. (2010). *Design of Observational Studies*. Springer.

Rubin, D. (2004). Teaching Statistical Inference for Causal Effects in Experiments and Observational Studies. *Journal of Educational and Behavioral Statistics*, 29(3):343.

Rubin, D. B. (1990). [on the application of probability theory to agricultural experiments. essay on principles. section 9.] comment: Neyman (1923) and causal inference in experiments and observational studies. *Statistical Science*, 5(4):472–480.

Sears, D. (1986). College sophomores in the laboratory: Influences of a narrow data base on social psychology's view of human nature. *Journal of Personality and Social Psychology*, 51(3):515.

Sekhon, J. (2008). The neyman-rubin model of causal inference and estimation via matching methods.

White, I. (2007). When race matters and when it doesn't: Racial group differences in response to racial cues. *American Political Science Review*, 101(02):339–354.