# Political Science 15
# Introduction to Research in Political Science
## Lecture 9a: Benefits and Pitfalls of Experiments

Alice Lépissier
University of California Santa Barbara

Special thanks to Chad Hazlett and Allison Horst for select slides and images, used with permission.

# Housekeeping (recap from Lecture 8)

**Announcements**

- Problem Set 5 is extended to Wednesday 29 July at 11:55pm.

**To read**

- From week 4, make sure you have finished Chapter 5 on Multivariate OLS.
- This week, for Lecture 7 on Multivariate OLS in Research, you need to have read Bateson (2012).
- Read Chapter 6 (sections 6.1-6.3 only) for Lecture 8 on Types of Variables and Data Sets (this one).
  - Optional reading Chapters 7-8 for this lecture (more info on logged variables in Ch. 7 and fixed effects in Ch. 8).
- For the end of this week and the start of week 6, where we tackle Lecture 9 on Experiments, read Chapter 10 and Gerber, Green, and Larimer (2008).

# Recap

In Lecture 8, we discussed what types of variables and data sets we use in applied research.

**Variables**

- Dummy (AKA binary or dichotomous) variables
- Discrete versus continuous variables
- Ordinal variables
- Nominal variables
- Logged variables (which are one version of a transformed variable)

**Data sets**

- Cross-sectional data
- Time series data
- Panel (cross-sectional & time series) data

We also discussed how to use **dummy** variables and **logged** variables in regression.

# Experiments in political science research

- Experiments are becoming increasingly common in political science research.
- Example: The 2012 Obama campaign randomly assigned different email subject lines to see which ones would generate the most donations.

# Experiments in political science research

- What are some political science experiments we've read about/discussed in this class?
  - Kalla and Broockman (2016) find that campaign contributions facilitate access to congressional officials
  - Gerber, Green and Larimer (2008) find that social pressure increases electoral turnout

# Experiments: solving the endogeneity problem

- Randomized experiments deal with the endogeneity problem by creating exogeneity.
- Check. What factors do we need to control for, as covariates, in an experiment? None! If randomization was done properly, and you have a large enough sample size, in expectation we should have balance across all the covariates.

# OLS model for experiments

$$Y_i = \beta_0 + \beta_1 D_i + \epsilon_i$$

- In experiments, we often denote the treatment variable as $D$. When $D = 1$, the unit is treated; when $D = 0$, the unit is a control.
- We call $\beta_1$ the **average treatment effect (ATE)**.
- Why? Since $D$ can only take two values, a one unit shift in $D$ is the effect of a unit moving from control (0) to treated (1). It's an average effect calculated across the full sample.

- This is a relatively simple, bivariate regression. Why?

- Given proper balance in the covariates across the treated and control groups, we don't need to control for any covariates. In expectation (as the sample size grows) they are balanced across the two groups.

# Problems that can arise in experiments

- Although many issues are addressed with experiments, most prominently endogeneity, there are still some pitfalls with experiments. What are they? Read Chapter 10 in *Real Stats*.
- **Attrition**: Units drop out of our experiment. Therefore, we never observed their outcome variable.
- **Balance**: Do the covariates (control variables) have the same mean across the two groups? We expect balance, given proper random assignment and large sample sizes.
- **Compliance**: Whether units actually receive the treatment they were assigned to (control or treatment).

# Attrition

- Attrition is when units drop out of our experiment. For this reason, we never see the result of the experiment for this unit. In other words, the outcome variable is *censored*.
- Potential causes of attrition: frustration with experiment, busyness, moving, death.
- We might expect that the longer the experiment lasts, the more likely we are to see attrition.

- What is the problem with attrition? It's likely not random. Some part of our sample will be more likely to attrit than others. This undermines the exogeneity created through random assignment.

- Potential solutions to attrition?
  1. Trimming the data set to make the treated and control units look the same on observable covariates (remember: this might not address unobservables).
  2. Modeling the attrition to control for it.

# Balance

- Balance is when your treated and control units look the same on other variables (i.e. covariates).
- You can check this with a balance table: a table that describes the mean and variance for relevant control variables.
- If your experiment was properly randomly assigned, and you had a large enough sample size (e.g. Gerber, Green and Larimer) this shouldn't be a problem.

- Another common approach:
- Blocking: Designing your treatment and control units in advance to ensure you will end up with balance in your covariates, particularly important confounders. This is most important for small sample sizes.

# Compliance

- Compliance refers to whether experimental subjects *actually* take the treatment if they were assigned to (and don't if they weren't assigned to).
- How could non-compliance happen?
  Groups sneak into the treated group (e.g. the job training program) or decline to take the treatment (e.g. they are assigned to come to a class to learn about statistics, but sadly, do not show up, thereby missing out on an awesome field/career!)

- What's one solution if you find you have non-compliance?
  Run an intention-to-treat model – compare the groups by assignment, regardless of whether they took the treatment.
- What should we expect this to do to our ATE?
  Shrinks it towards zero (attenuation bias). As non-compliance grows, ATE gets closer to zero. This is like noise.

# Political Science 15
# Introduction to Research in Political Science

## Lecture 9b: Experiments in Practice

Alice Lépissier
University of California Santa Barbara

Special thanks to Chad Hazlett and Allison Horst for select slides and images, used with permission.

# Experiments in political science research

American Political Science Review Vol. 102, No. 1 February 2008
DOI: 10.1017/S000305540808009X

## Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment

ALAN S. GERBER *Yale University*
DONALD P. GREEN *Yale University*
CHRISTOPHER W. LARIMER *University of Northern Iowa*

*Voter turnout theories based on rational self-interested behavior generally fail to predict significant turnout unless they account for the utility that citizens receive from performing their civic duty. We distinguish between two aspects of this type of utility, intrinsic satisfaction from behaving in accordance with a norm and extrinsic incentives to comply, and test the effects of priming intrinsic motives and applying varying degrees of extrinsic pressure. A large-scale field experiment involving several hundred thousand registered voters used a series of mailings to gauge these effects. Substantially higher turnout was observed among those who received mailings promising to publicize their turnout to their household or their neighbors. These findings demonstrate the profound importance of social pressure as an inducement to political participation.*

Among the most striking features of democratic political systems is the participation of millions of voters in elections. Why do large numbers of people vote, despite the fact that, as Hegel once observed, "the casting of a single vote is of no significance where there is a multitude of electors"? One hypothesis is adherence to social norms. Voting is widely regarded as a citizen duty (Blais 2000), and citizens worry that others will think less of them if they fail to participate in elections. Voters' sense of civic duty has long been a leading explanation of voter turnout among both behavioral (Campbell, Gurin, and Miller 1954) and formal (Downs 1957; Riker and Ordeshook 1968) theories of voter turnout.

chology, which emphasizes the extent to which other-regarding behavior varies depending on whether people perceive their actions to be public (Cialdini and Goldstein 2004; Cialdini and Trost 1998; Lerner and Tetlock 1999).

The empirical literature on the effects of social norms on voting has not advanced much beyond the initial survey work on this topic during the 1950s. Researchers have frequently used cross-sectional survey data to show that people who report feeling a greater sense of civic duty are also more likely to report voting. However, such observational evidence is frequently a misleading guide to causality; it may be that espousing the virtue of voting is a symptom, not a cause, of

# Finding a research question

- What is the overarching puzzle that Gerber, Green and Larimer want to answer?
- "To what extent do social norms cause voter turnout?" (p. 33)

# Operationalizing our ideas as variables

- What is our independent variable? Social norms
- What is our dependent variable? Electoral turnout

- What are our hypotheses?
  - $H_0$: Social norms do not shape electoral turnout
  - $H_A$: Social norms increase (or decrease) electoral turnout

- Who make up the experimental sample?
  - 180,002 households in the State of Michigan

# Treatments

- How many treatment conditions are there in this experiment?
  - There were four separate treatments (CIVIC DUTY, HAWTHORNE, SELF, NEIGHBORS).
- Why is the CIVIC DUTY treatment a "baseline" measure? ...because it is common to all treatment mailings!

- What are the main differences between the HAWTHORNE, SELF, and NEIGHBORS treatments?
  - They progressively *increase* the level of social pressure. HAWTHORNE makes people aware they are being studied, SELF includes info on an individual's voting records, NEIGHBORS also includes information on the voting records of neighbors.

- What is the control condition?
  - The control group did not receive ANY mailers.
- Check. Why do we need a control group in this experiment?

# Treatment example: Hawthorne mailing

**Hawthorne mailing**

30424-1

For more information: (517) 351-1975
email: etov@grebner.com
Practical Political Consulting
P. O. Box 6249
East Lansing, MI 48826

PRSRT STD
U.S. Postage
PAID
Lansing, MI
Permit # 444

ECRLOT **C001
THE SMITH FAMILY
9999 PARK LANE
FLINT MI 48507

Dear Registered Voter:

YOU ARE BEING STUDIED!

Why do so many people fail to vote? We've been talking about this problem for years, but it only seems to get worse.

This year, we're trying to figure out why people do or do not vote. We'll be studying voter turnout in the August 8 primary election.

Our analysis will be based on public records, so you will not be contacted again or disturbed in any way. Anything we learn about your voting or not voting will remain confidential and will not be disclosed to anyone else.

DO YOUR CIVIC DUTY — VOTE!

# Balance

- What part of the paper checks to see if the sample is balanced? Table 1.

**TABLE 1. Relationship between Treatment Group Assignment and Covariates (Household-Level Data)**

| | Control | Civic Duty | Hawthorne | Self | Neighbors |
|---|---|---|---|---|---|
| | Mean | Mean | Mean | Mean | Mean |
| Household size | 1.91 | 1.91 | 1.91 | 1.91 | 1.91 |
| Nov 2002 | .83 | .84 | .84 | .84 | .84 |
| Nov 2000 | .87 | .87 | .87 | .86 | .87 |
| Aug 2004 | .42 | .42 | .42 | .42 | .42 |
| Aug 2002 | .41 | .41 | .41 | .41 | .41 |
| Aug 2000 | .26 | .27 | .26 | .26 | .26 |
| Female | .50 | .50 | .50 | .50 | .50 |
| Age (in years) | 51.98 | 51.85 | 51.87 | 51.91 | 52.01 |
| *N* = | 99,999 | 20,001 | 20,002 | 20,000 | 20,000 |

*Note*: Only registered voters who voted in November 2004 were selected for our sample. Although not included in the table, there were no significant differences between treatment group assignment and covariates measuring race and ethnicity.

- Does this table raise any concerns about the experimental sample? No! The balance is excellent across all groups.

# Results

- Which treatment had the largest effect on turnout (the largest ATE)? **Neighbors**.

**TABLE 2. Effects of Four Mail Treatments on Voter Turnout in the August 2006 Primary Election**

| | Experimental Group | | | | |
|---|---|---|---|---|---|
| | Control | Civic Duty | Hawthorne | Self | Neighbors |
| Percentage Voting | 29.7% | 31.5% | 32.2% | 34.5% | 37.8% |
| N of Individuals | 191,243 | 38,218 | 38,204 | 38,218 | 38,201 |

- Can we tell if these results are statistically significant from this table? No.

# More results

- Gerber, Green, and Larimer also analyze their experiment with an OLS estimator in Table 3.

**TABLE 3. OLS Regression Estimates of the Effects of Four Mail Treatments on Voter Turnout in the August 2006 Primary Election**

| | Model Specifications | | |
|---|---|---|---|
| | (a) | (b) | (c) |
| Civic Duty Treatment (Robust cluster standard errors) | .018* (.003) | .018* (.003) | .018* (.003) |
| Hawthorne Treatment (Robust cluster standard errors) | .026* (.003) | .026* (.003) | .025* (.003) |
| Self-Treatment (Robust cluster standard errors) | .049* (.003) | .049* (.003) | .048* (.003) |
| Neighbors Treatment (Robust cluster standard errors) | .081* (.003) | .082* (.003) | .081* (.003) |
| N of individuals | 344,084 | 344,084 | 344,084 |
| Covariates** | No | No | Yes |
| Block-level fixed effects | No | Yes | Yes |

*Note*: Blocks refer to clusters of neighboring voters within which random assignment occurred. Robust cluster standard errors account for the clustering of individuals within household, which was the unit of random assignment.
* $p < .001$.
** Covariates are dummy variables for voting in general elections in November 2002 and 2000, primary elections in August 2004, 2002, and 2000.

- Are their results statistically significant? Yes.

- What should we conclude from this experiment about the effect of social norms and pressures on voter turnout?

# Next week (final week!)

- Monday 27
  - Finish Lecture 9: Natural experiments
  - Strategies to study for the final
- Lecture 10: Data science
- Course review

# Political Science 15
# Introduction to Research in Political Science
## Lecture 9c: Generalizability and Natural Experiments

Alice Lépissier
University of California Santa Barbara

Special thanks to Chad Hazlett and Allison Horst for select slides and images, used with permission.

# Recap

Two desirable qualities for an experiment:

- **Internal validity**
- **External validity**

# Recap

Issues that may come up during the implementation of an experiment:

- (Lack of) **Balance**
  - Assess balance by looking at balance table
  - If lack of balance: use multivariate OLS to control for variables that are unbalanced across treatment and control groups
- (Non-)**Compliance**
  - Assess compliance by looking at what percent of those assigned to treatment *actually* take it
  - If non-compliance: run ITT model
- **Attrition**
  - Assess non-random attrition by looking for patterns of who drops out across treatment and control groups
  - If attrition: use multivariate OLS, or trim the data

$\Rightarrow$ If there is perfect balance and compliance, and no attrition, then you can use bivariate OLS. Multivariate OLS would also be OK because it provides more precise estimates.

# Generalizability and External Validity

- The Gerber, Green and Larimer (2008) experiment did not suffer from implementation problems. They found substantively and statistically significant effects.
  - Check. What is the difference between substantive and statistical significance?
- Imagine you were to run this experiment again on a new sample of 180,000 voters in Michigan. Would you expect to see similar results?
- Now imagine, you were to run the experiment on 180,000 voters in Indiana. Would you expect to see similar results?
- What about if you ran the experiment on 180,000 voters in the Canadian province of British Columbia?
- What about if you ran the experiment in Taiwan or Japan? Or in Uganda?

# Generalizability and External Validity

- We might expect the results to hold for similar political contexts. But as we move farther and farther away from the case where we found the initial evidence, we may encounter **generalizability** concerns.

- In general, experiments tend to score high on internal validity, but lower on external validity.

# Natural experiments

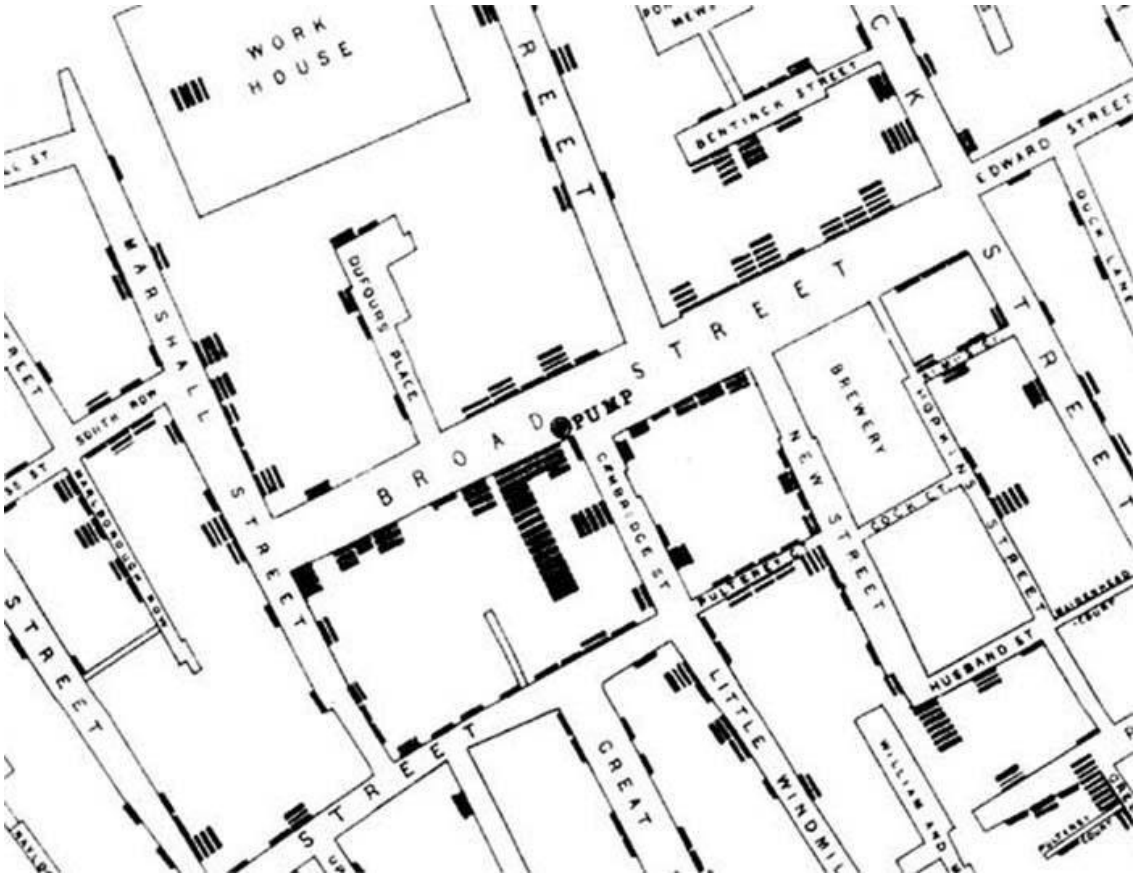
- We can't always conduct an experiment when we have a research question that interests us. Can you think of a political science research question that we can't answer with an experiment?
- Some experiments are impossible. Others are illegal. Others are unethical.
- In these cases, we can sometimes use a **natural experiment**. Natural experiments occur when a researcher identifies a situation in which values of the independent variable have been determined by a random, or at least exogenous, process.

# The first natural experiment

- Scene: mid-1800s, London. The germ theory of disease had not been developed yet. People believed that outbreaks of cholera were caused by "miasma" (bad air).
- John Snow was a skeptical of miasma theory.
- What caused the 1854 cholera outbreak in London?
- John Snow used a natural experiment to find out.

# The first natural experiment

Map of cholera outbreak in Soho district of London.




- Two main companies provided London water: Southwark and Vauxhall Waterworks Company, and the Lambeth Waterworks Company.
- Lambeth Company took water from the Thames further upriver; less exposed to sewage.
- Development of water supply routes was pretty haphazard.

# The first natural experiment

"In many cases a single house has a supply different from that on either side. Each company supplies both rich and poor, both large houses and small; there is no difference in the condition or occupation of the persons receiving the water of the different companies...As there is no difference whatever either in the houses or the people receiving the supply of the two Water Companies, or in any of the physical conditions with which they are surrounded, it is obvious that no experiment could have been devised which would more thoroughly test the effect of water supply on the progress of Cholera than this, which circumstances placed ready made before the observer."

John Snow, *On the Mode of Communication of Cholera*, 1855.

# Natural experiment: effect of Vietnam draft on political attitudes

- Erikson and Stoker (2011) study the Vietnam draft lottery. They argue that exposure to the draft is randomly assigned, and compare the political attitudes of individuals who had high and low draft lottery numbers.
- Males holding low lottery numbers (i.e. were more likely to be drafted) became more antiwar, more liberal and more Democratic in their voting behaviors.

# Natural experiment: effect of family size on labor market outcomes

- The effect of family size on the labor market outcomes of the mother is likely endogenous. There could be omitted variable bias (e.g. preferences) and reverse causality.
- Angrist and Evans (1998) observe that two-child families with 2 boys or 2 girls are more likely to have a third child, than those families with 1 boy and 1 girl. Distribution of sex is essentially random.
- They found that the labor market consequences of a third child were more severe for lower income and less educated women. And that having a third child had little impact on husbands' earnings.