Experiments

Causality and Ethics

Understanding Political Numbers

April 3, 2019

In-class presentations

Sign-up sheet is **INCOMING** (pick week 1 or week 2)

Sign-up sheet is **INCOMING** (pick week 1 or week 2)

Five-minute slideshow presentation on your project

Sign-up sheet is **INCOMING** (pick week 1 or week 2)

Five-minute slideshow presentation on your project

What's the question, theory/hypotheses, data, variables

Sign-up sheet is **INCOMING** (pick week 1 or week 2)

Five-minute slideshow presentation on your project

What's the question, theory/hypotheses, data, variables

Any results? (if you're ready)

Sign-up sheet is INCOMING (pick week 1 or week 2)

Five-minute slideshow presentation on your project

What's the question, theory/hypotheses, data, variables

Any results? (if you're ready)

Email slides to Michael before section

Sign-up sheet is **INCOMING** (pick week 1 or week 2)

Five-minute slideshow presentation on your project

What's the question, theory/hypotheses, data, variables

Any results? (if you're ready)

Email slides to Michael before section

Vote to advance best presentation to the "finals"

Sign-up sheet is INCOMING (pick week 1 or week 2)

Five-minute slideshow presentation on your project

What's the question, theory/hypotheses, data, variables

Any results? (if you're ready)

Email slides to Michael before section

Vote to advance best presentation to the "finals"

Don't get EC if you don't show up

Using data to learn about "cause and effect"

$$y = \alpha + \beta_1 x + \beta_2 z + \epsilon$$

Under what conditions can we learn the effect of x (β_1) if we *don't* control for z?

$$y = \alpha + \beta_1 x + \beta_2 z + \epsilon$$

Under what conditions can we learn the effect of x (β_1) if we *don't* control for z?

If x and z are *independent*...

```
# fake data
obs_data <- tibble(
  z = random_noise(),
  x = random_noise(),
  y = (2*x) + (-4*z) + random_noise()
)

# estimate regression
lm(y ~ x, data = obs_data)
```

```
##
## Call:
## lm(formula = y ~ x, data = obs_data)
##
## Coefficients:
## (Intercept)            x
##    -0.06811      2.28935
```

$$y = \alpha + \beta_1 x + \beta_2 z + \epsilon$$

Under what conditions can we learn the effect of x (β_1) if we *don't* control for z?

If x and z are *independent*...

```
# fake data
obs_data <- tibble(
  z = random_noise(),
  x = random_noise(),
  y = (2*x) + (-4*z) + random_noise()
)

# estimate regression
lm(y ~ x, data = obs_data)
```

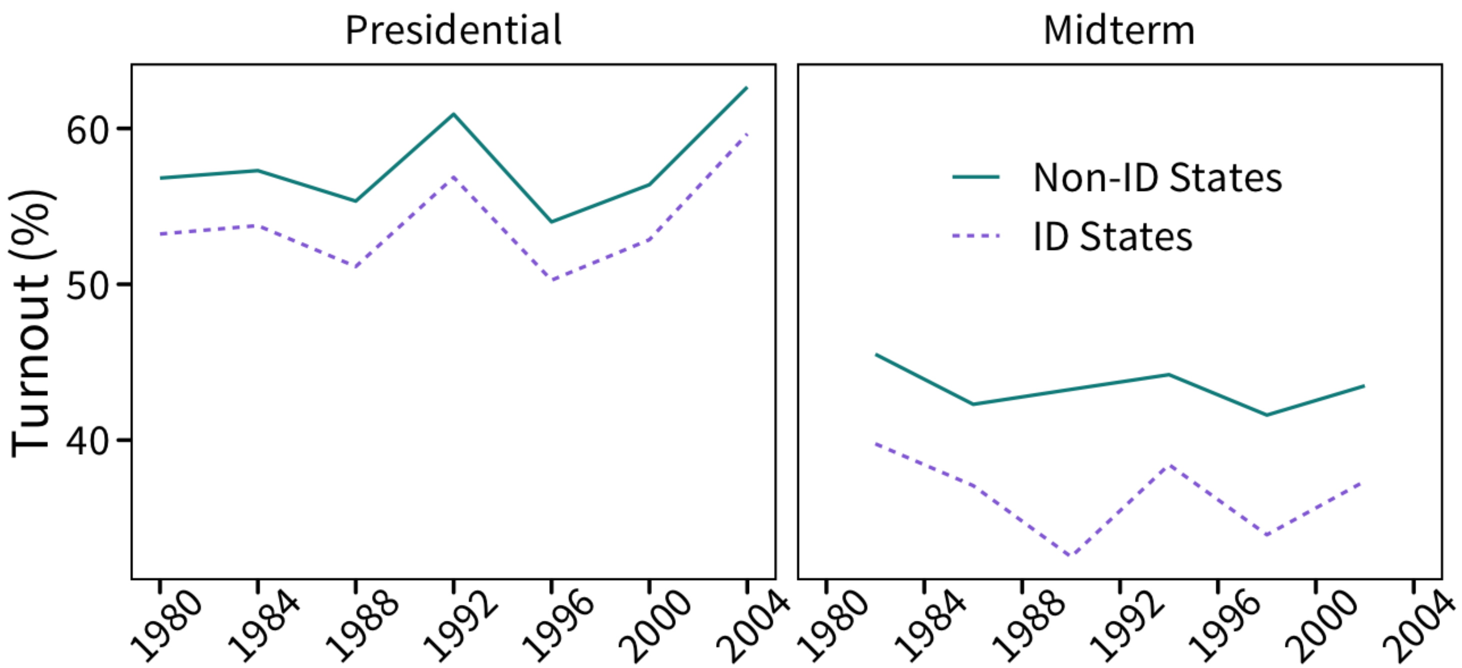
```
##
## Call:
## lm(formula = y ~ x, data = obs_data)
##
## Coefficients:
## (Intercept)            x
##    -0.06811      2.28935
```

If x and z are not independent...

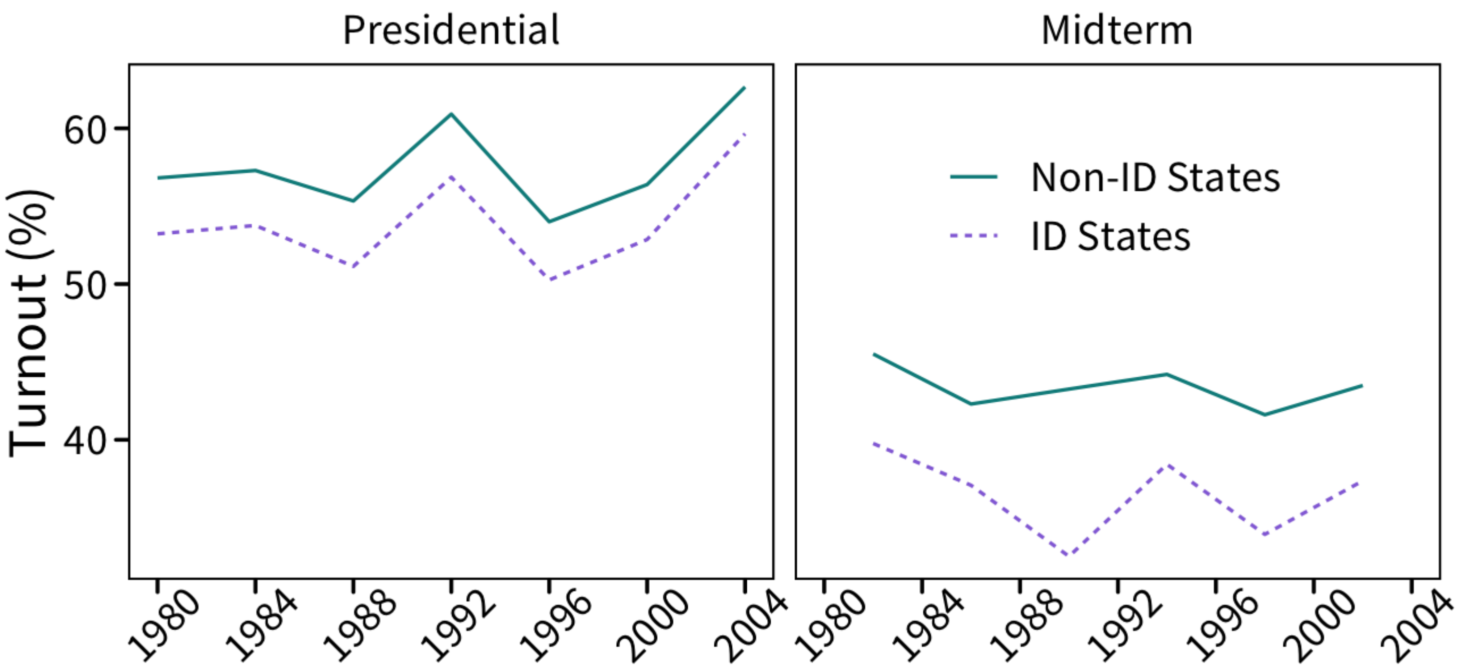
```
# x is now f(z)
obs_data <- tibble(
  z = random_noise(),
  x = (-1*z) + random_noise(),
  y = (2*x) + (-4*z) + random_noise()
)

# x effect will be biased
lm(y ~ x, data = obs_data)
```

```
##
## Call:
## lm(formula = y ~ x, data = obs_data)
##
## Coefficients:
## (Intercept)            x
##      0.1389       3.9516
```

Selection bias and confounding



Selection bias and confounding



How to fix? Control for everything that influences *both* X and Y

(which is really hard to do)

Endogenous vs. Exogenous causes

Independent variable comes from within the system (endogenous)

Endogenous vs. Exogenous causes

Independent variable comes from within the system (endogenous)

- Includes "reverse causality," y causes x

Endogenous vs. Exogenous causes

Independent variable comes from within the system (endogenous)

- Includes "reverse causality," y causes x

Value of the independent variable is influenced by external forces (exogenous)

- Sudden shocks (rain $\rightarrow$ lower turnout)
- Unanticipated events (surprise terror attacks)

Endogenous vs. Exogenous causes

Independent variable comes from within the system (endogenous)

- Includes "reverse causality," y causes x

Value of the independent variable is influenced by external forces (exogenous)

- Sudden shocks (rain $\rightarrow$ lower turnout)
- Unanticipated events (surprise terror attacks)

It's always easier if X is an exogenous shock

Experiments

- Controlled setting
- Independent variable is *randomly assigned*, no selection biases
- Confounding variables either *controlled* or *evenly balanced* between "treatment" and "control" groups (thanks to randomization)

Experiments

- Controlled setting
- Independent variable is *randomly assigned*, no selection biases
- Confounding variables either *controlled* or *evenly balanced* between "treatment" and "control" groups (thanks to randomization)

Observational data

- Uncontrolled setting
- Independent variable affected by various social forces, not all of them observed
- Confounding variables are everywhere, up to the researcher to measure and control

Anatomy of an experiment

What's the effect of watching MSNBC on Trump approval?

What's the effect of watching MSNBC on Trump approval?

1. Recruit participants into the study

What's the effect of watching MSNBC on Trump approval?

1. Recruit participants into the study

2. ***RANDOMLY*** divide them into two (or more) groups

What's the effect of watching MSNBC on Trump approval?

1. Recruit participants into the study

2. ***RANDOMLY*** divide them into two (or more) groups

3. Groups get specially designed *treatments* (or *control*)

What's the effect of watching MSNBC on Trump approval?

1. Recruit participants into the study

2. *RANDOMLY* divide them into two (or more) groups

3. Groups get specially designed *treatments* (or *control*)

4. Measure the dependent variable

What's the effect of watching MSNBC on Trump approval?

1. Recruit participants into the study
2. *RANDOMLY* divide them into two (or more) groups
3. Groups get specially designed *treatments* (or *control*)
4. Measure the dependent variable
5. Did *treatment assignment* (IV) affect the DV?

What's the effect of watching MSNBC on Trump approval?

1. Recruit participants into the study
2. *RANDOMLY* divide them into two (or more) groups
3. Groups get specially designed *treatments* (or *control*)
4. Measure the dependent variable
5. Did *treatment assignment* (IV) affect the DV?

Practical and ethical concerns

What's the effect of watching MSNBC on Trump approval?

1. Recruit participants into the study
2. *RANDOMLY* divide them into two (or more) groups
3. Groups get specially designed *treatments* (or *control*)
4. Measure the dependent variable
5. Did *treatment assignment* (IV) affect the DV?

Practical and ethical concerns

1. Imagine replacing "watching MSNBC" with "gender"

What's the effect of watching MSNBC on Trump approval?

1. Recruit participants into the study
2. *RANDOMLY* divide them into two (or more) groups
3. Groups get specially designed *treatments* (or *control*)
4. Measure the dependent variable
5. Did *treatment assignment* (IV) affect the DV?

Practical and ethical concerns

1. Imagine replacing "watching MSNBC" with "gender"
2. Imagine replacing "watching MSNBC" with "unemployment"

Why randomization?

What is causality?

Causality can be a statement about *potential outcomes*.

What is causality?

Causality can be a statement about *potential outcomes*.

$$y_i = \begin{cases} y_i\,(\text{Treatment}) & \text{if treatment} \\ y_i\,(\text{Control}) & \text{if control} \end{cases}$$

What is causality?

Causality can be a statement about *potential outcomes*.

$$y_i = \begin{cases} y_i\,(\text{Treatment}) & \text{if treatment} \\ y_i\,(\text{Control}) & \text{if control} \end{cases}$$

Does the treatment influence y_i?

$$\text{Treatment effect}_i = y_i\,(\text{Treatment}) - y_i\,(\text{Control})$$

What is causality?

Causality can be a statement about *potential outcomes*.

$$y_i = \begin{cases} y_i\,(\text{Treatment}) & \text{if treatment} \\ y_i\,(\text{Control}) & \text{if control} \end{cases}$$

Does the treatment influence y_i?

$$\text{Treatment effect}_i = y_i\,(\text{Treatment}) - y_i\,(\text{Control})$$

Fundamental problem of causal inference: can never observe $y_i\,(\text{Treatment})$ and $y_i\,(\text{Control})$ simultaneously

What is causality?

Causality can be a statement about *potential outcomes*.

$$y_i = \begin{cases} y_i\,(\text{Treatment}) & \text{if treatment} \\ y_i\,(\text{Control}) & \text{if control} \end{cases}$$

Does the treatment influence y_i?

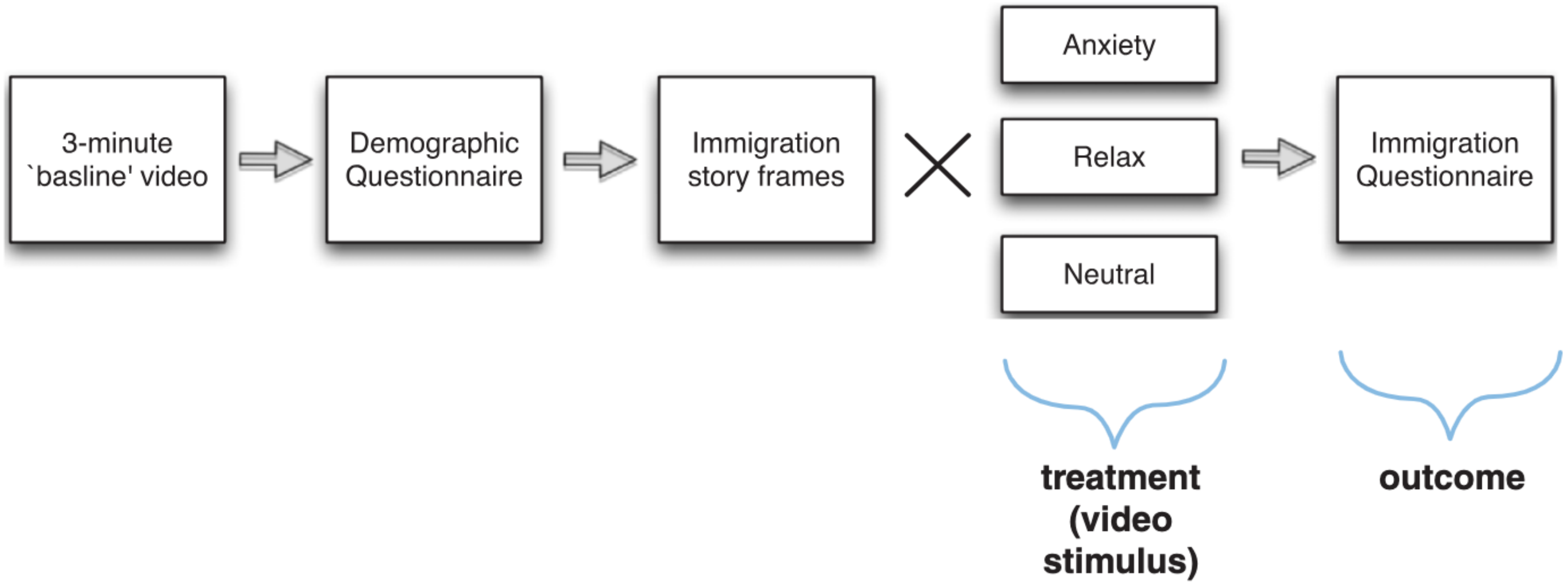
$$\text{Treatment effect}_i = y_i\,(\text{Treatment}) - y_i\,(\text{Control})$$

Fundamental problem of causal inference: can never observe $y_i\,(\text{Treatment})$ and $y_i\,(\text{Control})$ simultaneously

Randomly assign treatments, measure *average* $y_i\,(\text{Treatment})$ and *average* $y_i\,(\text{Control})$

Examples

Lab experiments



Exposure, emotion, communication

Sample quality

Financial cost

Survey experiments

Public opinion, willingness to act

Financial cost

Please read the descriptions of the potential immigrants carefully. Then, please indicate which of the two immigrants you would personally prefer to see admitted to the United States.

	Immigrant 1	**Immigrant 2**
Prior Trips to the U.S.	Entered the U.S. once before on a tourist visa	Entered the U.S. once before on a tourist visa
Reason for Application	Reunite with family members already in U.S.	Reunite with family members already in U.S.
Country of Origin	Mexico	Iraq
Language Skills	During admission interview, this applicant spoke fluent English	During admission interview, this applicant spoke fluent English
Profession	Child care provider	Teacher
Job Experience	One to two years of job training and experience	Three to five years of job training and experience
Employment Plans	Does not have a contract with a U.S. employer but has done job interviews	Will look for work after arriving in the U.S.
Education Level	Equivalent to completing two years of college in the U.S.	Equivalent to completing a college degree in the U.S.
Gender	Female	Male

	Immigrant 1	Immigrant 2
If you had to choose between them, which of these two immigrants should be given priority to come to the United States to live?	○	○

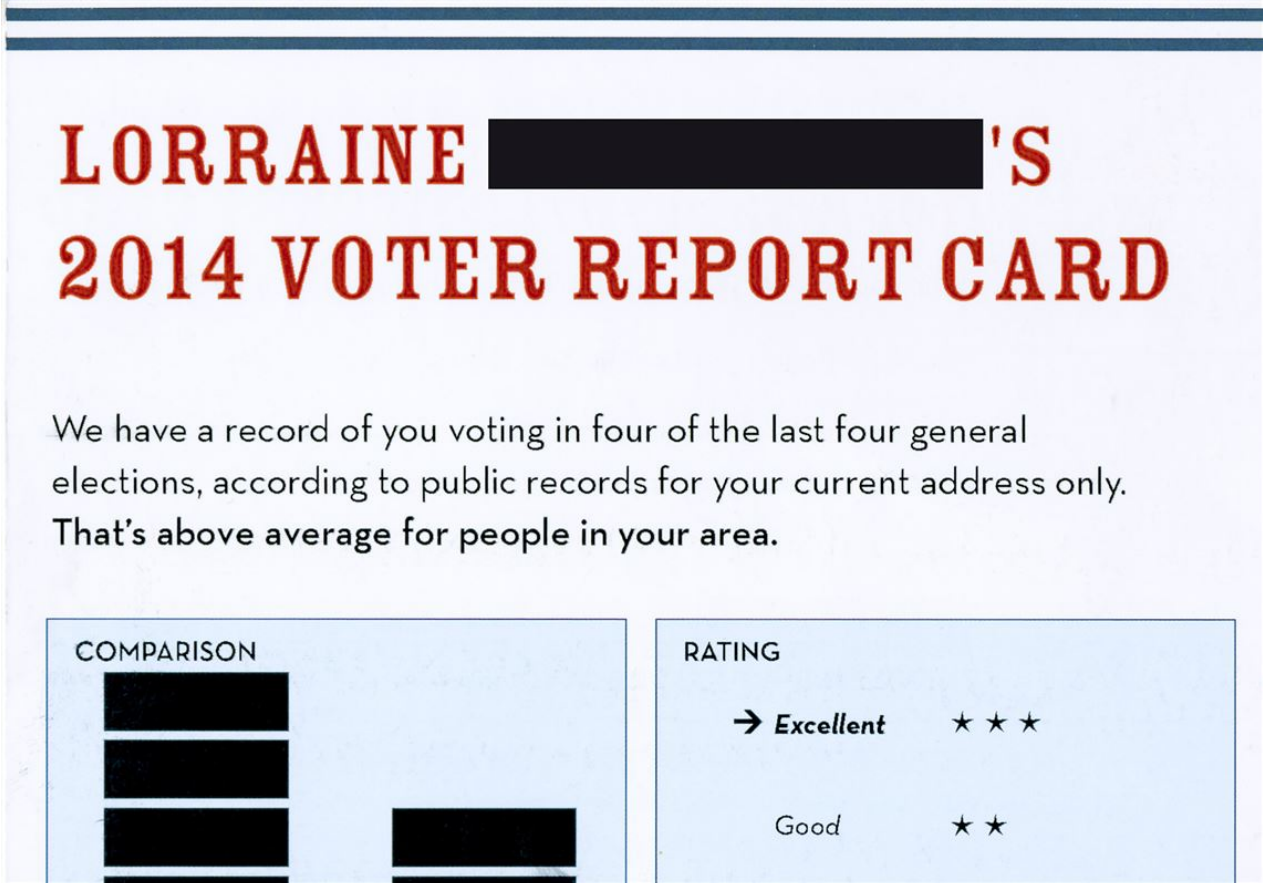
Field experiments

Stimulate turnout, economic activity, government responsiveness

More realistic

Practically difficult

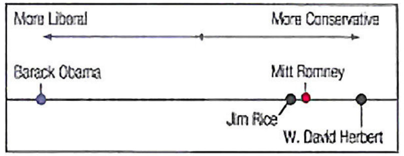
Even more ethically dicey



2014 Montana General Election Voter Information Guide

Election Date: November 4, 2014

Nonpartisan Supreme Court Justice #1 Race

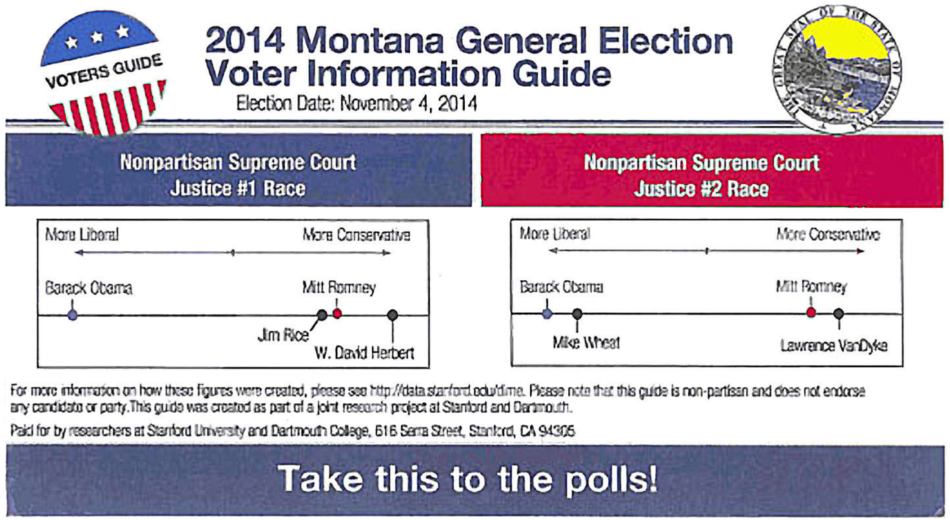


Nonpartisan Supreme Court Justice #2 Race

For more information on how these figures were created, please see http://data.stanford.edu/dime. Please note that this guide is non-partisan and does not endorse any candidate or party. This guide was created as part of a joint research project at Stanford and Dartmouth.

Paid for by researchers at Stanford University and Dartmouth College, 616 Serra Street, Stanford, CA 94305

Take this to the polls!



Professors' Research Project Stirs Political Outrage in Montana

By Derek Willis

Oct. 28, 2014

31

The only thing that three political scientists wanted to do was send mailers to thousands of Montana voters as part of a study of nonpartisan elections. What could possibly go wrong?

A lot, judging from the outrage and a state investigation. It has also raised thorny questions about political science field research, which isn't uncommon, and its ability to affect an election.

The experiment, by the political scientists Adam Bonica and Jonathan Rodden of Stanford University and Kyle Dropp of Dartmouth College, sent mail to 100,000 Montana registered voters about two elections for the state's supreme court. The Montana mailer, labeled "2014 Montana General Election Voter Information Guide," featured the official state seal. It also placed the four judicial candidates on an ideological spectrum that included Barack Obama and Mitt Romney as reference

Everything's got issues

Internal vs. external validity

Internal vs. external validity

Institutional Review Boards (IRB)

Internal vs. external validity

Institutional Review Boards (IRB)

Informed consent, confidentiality, sensitive issues

Internal vs. external validity

Institutional Review Boards (IRB)

Informed consent, confidentiality, sensitive issues

Deception (and debriefing)

Internal vs. external validity

Institutional Review Boards (IRB)

Informed consent, confidentiality, sensitive issues

Deception (and debriefing)

Balancing risks and rewards

Audit studies

Tyrone Washington To: decrescenzo@wisc.edu 3/11/18, 11:28 AM

Dear Michael,

Hello! My name is Tyrone Washington, and I am a senior at UW-Madison. I'm hoping to apply to graduate school next fall. I am interested in pursuing a graduate degree in Political Science and saw that you are one of the graduate students in that department. I know you are very busy, but I was hoping you could tell me a little bit about your experience with UW-Madison's Political Science graduate program and what it is like to be a graduate student at UW-Madison. Specifically, what led you to choose your program for your graduate studies, what is your day-to-day life like, and what are the most rewarding or challenging parts of your graduate education?

Thank you for your time. I look forward to your responses.

Sincerely,
Tyrone Washington
...

Audit studies

Tyrone Washington To: decrescenzo@wisc.edu 3/11/18, 11:28 AM

Dear Michael,

Hello! My name is Tyrone Washington, and I am a senior at UW-Madison. I'm hoping to apply to graduate school next fall. I am interested in pursuing a graduate degree in Political Science and saw that you are one of the graduate students in that department. I know you are very busy, but I was hoping you could tell me a little bit about your experience with UW-Madison's Political Science graduate program and what it is like to be a graduate student at UW-Madison. Specifically, what led you to choose your program for your graduate studies, what is your day-to-day life like, and what are the most rewarding or challenging parts of your graduate education?

Thank you for your time. I look forward to your responses.

Sincerely,
Tyrone Washington

...

Don't always need informed consent (public purpose, typical interaction)

Government services, resume callbacks...

Natural experiments, Quasi experiments, "As-if-random" assignment

Can Contact Cure Prejudice: A Natural Experiment in Israeli Medical Clinics *

Chagai M. Weiss *University of Wisconsin - Madison*

In many societies, even when segregation or conflict are pronounced, brief intergroup contact in busses, markets, shops and hospitals is prevalent. Such contact is often theorized as a force influencing intergroup attitudes as well as voting behavior and violence. Despite the prevalence of such intergroup contact, and despite the prominent role of contact in multiple theoretical frameworks of ethnic politics, there is little evidence regarding its causal effects. Exploiting the random assignment of patients to doctors in medical clinics in Israel, and leveraging a treatment evaluation survey, I introduce a natural experiment suited to identify the causal effects of intergroup contact between Jewish patients and Palestinian doctors. I further explore how doctor and patient characteristics moderate the effects of contact.

Keywords: contact, prejudice, intergroup relations, identity

The Election Timing Effect: Evidence from a Policy Intervention in Texas*

Sarah F. Anzia

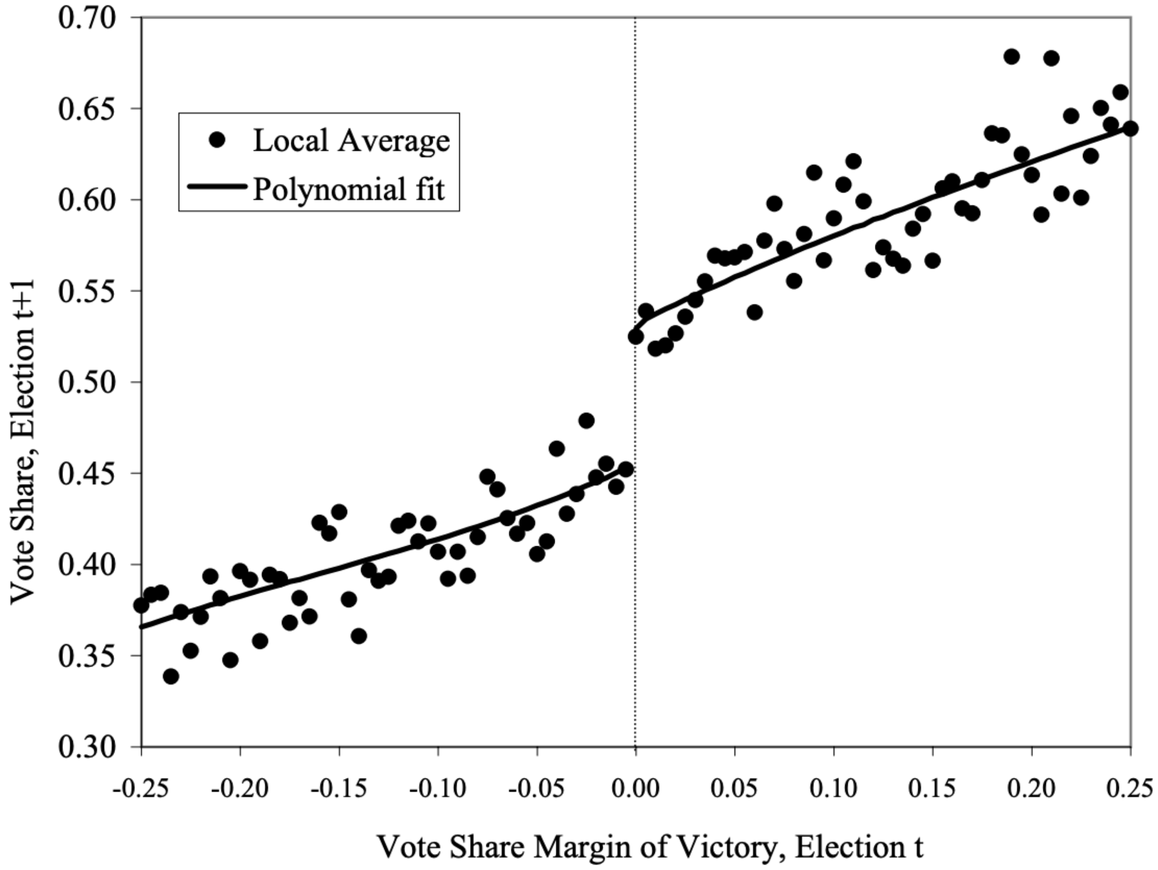
Goldman School of Public Policy, University of California, Berkeley, USA; sanzia@berkeley.edu

---

ABSTRACT

Many governments in the United States hold elections on days other than national Election Day. Recent studies have argued that the low voter turnout that accompanies such off-cycle elections could create an advantage for interest groups. However, the endogeneity of election timing makes it difficult to estimate its causal effect on political outcomes. In this paper, I develop a theoretical framework that explains how changes to election timing affect the electoral fortunes of organized interest groups. I test the theory by examining the effects of a 2006 Texas law that forced approximately 20 percent of the state's school districts to move their elections to the same day as national elections. Using matching as well as district fixed effects regression, I estimate the causal effect of the switch to on-cycle election timing on district teacher salaries, since teachers and their unions tend to be the dominant

Figure IVa: Democrat Party's Vote Share in Election t+1, by Margin of Victory in Election t: local averages and parametric fit



Looking ahead

In section: regression practice

On Monday: Big data (listen to podcast)

On Wednesday: Elections and campaigns (do reading)