Gov 50: 6. Causality

Matthew Blackwell

Harvard University

Roadmap

1. What is causality?
2. Randomized experiments
3. Calculating effects

1/ What is causality?

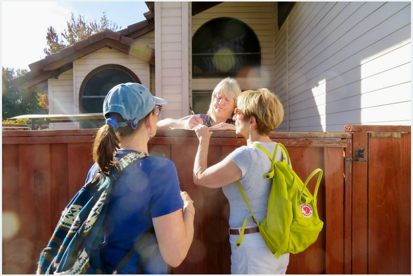
Two roads diverged in a yellow wood,
And sorry I could not travel both
And be one traveler, long I stood
And looked down one as far as I could
To where it bent in the undergrowth;

What is a causal effect?

factual vs. counterfactual

- Does increasing the minimum wage increase the unemployment rate?
  - Unemployment rate went up after the minimum wage increased
  - Would it have gone up if the minimum wage increase not occurred?
- Does having girls affect a judge's rulings in court?
  - A judge with a daughter gave a pro-choice ruling.
  - Would they have done that if had a son instead?
- **Fundamental problem of causal inference**:
  - Can never observe counterfactuals, must be inferred.

Political canvassing study



POLITICAL SCIENCE

Durably reducing transphobia: A field experiment on door-to-door canvassing

David Broockman[1]* and Joshua Kalla[2]

Existing research depicts intergroup prejudices as deeply ingrained, requiring intense intervention to lastingly reduce. Here, we show that a single approximately 10-minute conversation encouraging actively taking the perspective of others can markedly reduce prejudice for at least 3 months. We illustrate this potential with a door-to-door canvassing intervention in South Florida targeting antitransgender prejudice. Despite declines in homophobia, transphobia remains pervasive. For the intervention, 56 canvassers went door to door encouraging active perspective-taking with 501 voters at voters' doorsteps. A randomized trial found that these conversations substantially reduced transphobia, with decreases greater than Americans' average decrease in homophobia from 1998 to 2012. These effects persisted for 3 months, and both transgender and nontransgender canvassers were effective. The intervention also increased support for a nondiscrimination law, even after exposing voters to counterarguments.

- Can canvassers change minds about topics like transgender rights?
- Experimental setting:
  - Randomly assign canvassers to have a conversation about transgender right or a conversation about recycling.
  - Trans rights conversations focused on “perspective taking”
- Outcome of interest: support for trans rights policies.

Credit: Fabrice Florian via Flickr

A tale of two respondents

	Conversation Script	Support for Nondiscrimination Law
Respondent 1	Recycling	No
Respondent 2	Trans rights	Yes

Did the second respondent support the law **because** of the perspective-taking conversation?

Translating into math

Useful to have **compact** notation for referring to **treatment variable**:

$$T_i = \begin{cases} 1 & \text{if respondent } i \text{ had trans rights conversation} \\ 0 & \text{if respondent } i \text{ had recycling conversation} \end{cases}$$

Similar notation for the **outcome variable**:

$$Y_i = \begin{cases} 1 & \text{if respondent } i \text{ supports trans nondiscrimination laws} \\ 0 & \text{if respondent } i \text{ doesn't support nondiscrimination laws} \end{cases}$$

i is a placeholder to refer to a generic unit/respondent: Y_{42} is the outcome for the 42nd unit.

A tale of two respondents (redux)

	Conversation Script	Support for Nondiscrimination Law
Respondent 1	Recycling	No
Respondent 2	Trans rights	Yes

becomes...

i	T_i	Y_i
Respondent 1	0	0
Respondent 2	1	1

Causal effects & counterfactuals

- What does "T_i causes Y_i" mean? $\rightsquigarrow$ **counterfactuals**, "what if"
- Would respondent change their support based on the conversation?
- Two **potential outcomes**:
  - $Y_i(1)$: would respondent *i* support ND laws if they had trans rights script?
  - $Y_i(0)$: would respondent *i* support ND laws if they had recycling script?
- **Causal effect**: $Y_i(1) - Y_i(0)$
  - $Y_i(1) - Y_i(0) = 0 \rightsquigarrow$ script has no effect on policy views
  - $Y_i(1) - Y_i(0) = -1 \rightsquigarrow$ trans rights script lower support for laws
  - $Y_i(1) - Y_i(0) = +1 \rightsquigarrow$ trans rights script increases support for laws

Potential outcomes

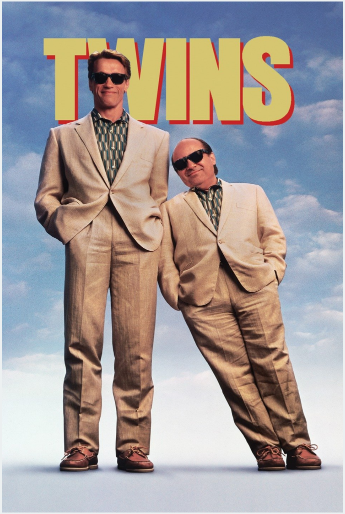
i	T_i	Y_i	$Y_i(1)$	$Y_i(0)$
Respondent 1	0	0	???	0
Respondent 2	1	1	1	???

- **Fundamental problem of causal inference**:
  - We only observe one of the two potential outcomes.
  - Observe $Y_i = Y_i(1)$ if $T_i = 1$ or $Y_i = Y_i(0)$ if $T_i = 0$
- To infer causal effect, we need to infer the missing counterfactuals!

How can we figure out counterfactuals?

- Find a similar unit! $\rightsquigarrow$ **matching**
  - Mill's method of difference
- Does respondent support law because of the trans rights script?
  - $\rightsquigarrow$ find a identical respondent who got the recycling script.
- NJ increased the minimum wage. Causal effect on unemployment?
  - $\rightsquigarrow$ find a state similar to NJ that didn't increase minimum wage.

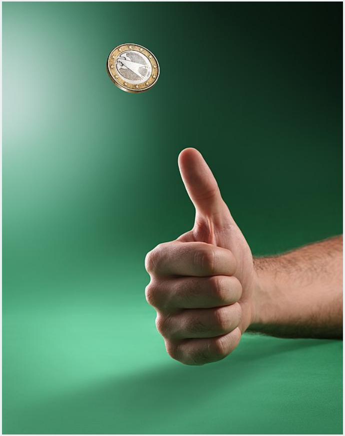
Imperfect matches



- The problem: imperfect matches!
- Say we match i (treated) and j (control)
- **Selection Bias:** $Y_i(1) \neq Y_j(1)$
- Those who take treatment may be different that those who take control.
- How can we correct for that?

2/ Randomized experiments

Match groups not individuals



- **Randomized control trial**: each unit's treatment assignment is determined by chance.
    - Flip a coin; draw red and blue chips from a hat; etc
- Randomization ensures **balance** between treatment and control group.
    - Treatment and control group are identical **on average**
    - Similar on both observable and unobservable characteristics.

A little more notation

- We will often refer to the **sample size** (number of units) as n.
- We often have n measurements of some variable: $(Y_1, Y_2, \dots, Y_n)$
- How many in our sample support nondiscrimination laws?

$$Y_1 + Y_2 + Y_3 + \cdots + Y_n$$

- Notation is a bit clunky, so we often use the **Sigma notation**:

$$\sum_{i=1}^{n} Y_i = Y_1 + Y_2 + Y_3 + \cdots + Y_n$$

- $\Sigma_{i=1}^{n}$ means sum each value from Y_1 to Y_n

Averages

- The **sample average** or **sample mean** is simply the sum of all values divided by the number of values.
- Sigma notation allows us to write this in a compact way:

$$\overline{Y} = \frac{1}{n} \sum_{i=1}^{n} Y_i$$

- Suppose we surveyed 6 people and 3 supported nondiscrim. laws:

$$\overline{Y} = \frac{1}{6} (1 + 1 + 1 + 0 + 0 + 0) = 0.5$$

Quantity of interest

- We want to estimate the average causal effects over all units:

$$\text{Sample Average Treatment Effect (SATE)} = \frac{1}{n}\sum_{i=1}^{n}\{Y_i(1) - Y_i(0)\}$$
$$= \frac{1}{n}\sum_{i=1}^{n} Y_i(1) - \frac{1}{n}\sum_{i=1}^{n} Y_i(0)$$

- Why can't we just calculate this quantity directly?
- What we can estimate instead:

$$\text{Difference in means} = \overline{Y}_{\text{treated}} - \overline{Y}_{\text{control}}$$

  - $\overline{Y}_{\text{treated}}$: sample average outcome for treated group
  - $\overline{Y}_{\text{control}}$: sample average outcome for control group
- When will the difference-in-means is a good estimate of the SATE?

Why randomization works

- Under an RCT, treatment and control groups are random samples.
- Average in the treatment group will be similar to average if all treated:

$$\overline{Y}_{\text{treated}} \approx \frac{1}{n}\sum_{i=1}^{n} Y_i(1)$$

- Average in the control group will be similar to average if all untreated:

$$\overline{Y}_{\text{control}} \approx \frac{1}{n}\sum_{i=1}^{n} Y_i(0)$$

- Implies difference-in-means should be close to SATE:

$$\overline{Y}_{\text{treated}} - \overline{Y}_{\text{control}} \approx \frac{1}{n}\sum_{i=1}^{n} Y_i(1) - \frac{1}{n}\sum_{i=1}^{n} Y_i(0) = \frac{1}{n}\sum_{i=1}^{n}\{Y_i(1) - Y_i(0)\} = \text{SATE}$$

Some potential problems with RCTs

- **Placebo effects**:
    - Respondents will be affected by any intervention, even if they shouldn't have any effect.
    - Reason to have control group be recycling script
- **Hawthorne effects**:
    - Respondents act differently just knowing that they are under study.

Balance checking

- Can we determine if randomization "worked"?
- If it did, we shouldn't see large differences between treatment and control group on **pretreatment variable**.
    - Pretreatment variable are those that are unaffected by treatment.
- We can check in the actual data for some pretreatment variable X
    - $\overline{X}_{\text{treated}}$: average value of variable for treated group.
    - $\overline{X}_{\text{control}}$: average value of variable for control group.
    - Under randomization, $\overline{X}_{\text{treated}} - \overline{X}_{\text{control}} \approx 0$

Multiple treatments

- Instead of 1 treatment, we might have multiple **treatment arms**:
    - Control condition
    - Treatment A
    - Treatment B
    - Treatment C, etc
- In this case, we will look at multiple comparisons:
    - $\overline{Y}_{\text{treated, A}} - \overline{Y}_{\text{control}}$
    - $\overline{Y}_{\text{treated, B}} - \overline{Y}_{\text{control}}$
    - $\overline{Y}_{\text{treated, A}} - \overline{Y}_{\text{treated, B}}$
- If treatment arms are randomly assigned, these differences will be good estimators for each causal contrast.

3/ Calculating effects

Transphobia study data

```r
## reinstall gov50data if necessary
library(gov50data)
```

Variable Name	Description
age	Age of the R in years
female	1=R marked “Female” on voter reg., 0 otherwise
voted_gen_14	1 if R voted in the 2014 general election
vote_gen_12	1 if R voted in the 2012 general election
treat_ind	1 if R assigned to trans rights script, 0 for recycling
racename	name of racial identity indicated on voter file
democrat	1 if R is a registered Democrat
nondiscrim_pre	1 if R supports nondiscrim. law at baseline
nondiscrim_post	1 if R supports nondiscrim. law after 3 months

Peak at the data

```
trans
```

```
## # A tibble: 565 x 9
##      age female voted_gen_14 voted_g~1 treat~2 racen~3 democ~4 nondi~5
##    <dbl>  <dbl>        <dbl>     <dbl>   <dbl> <chr>     <dbl>   <dbl>
##  1    29      0            0         1       0 Africa~       1       1
##  2    59      1            1         0       1 Africa~       1       1
##  3    35      1            1         1       1 Africa~       1       0
##  4    63      1            1         1       1 Africa~       1       0
##  5    65      0            1         1       1 Africa~       0       1
##  6    51      1            1         1       0 Caucas~       0       1
##  7    26      1            1         1       0 Africa~       1       1
##  8    62      1            1         1       1 Africa~       1       1
##  9    37      0            1         1       0 Caucas~       0       1
## 10    51      1            1         1       0 Caucas~       0       0
## # ... with 555 more rows, 1 more variable: nondiscrim_post <dbl>, and
## #   abbreviated variable names 1: voted_gen_12, 2: treat_ind,
## #   3: racename, 4: democrat, 5: nondiscrim_pre
```

Calculate the average outcomes in each group

```
treat_mean <- trans |>
  filter(treat_ind == 1) |>
  summarize(nondiscrim_mean = mean(nondiscrim_post))
treat_mean
```

```
## # A tibble: 1 x 1
##   nondiscrim_mean
##             <dbl>
## 1           0.687
```

```
control_mean <- trans |>
  filter(treat_ind == 0) |>
  summarize(nondiscrim_mean = mean(nondiscrim_post))
control_mean
```

```
## # A tibble: 1 x 1
##   nondiscrim_mean
##             <dbl>
## 1           0.648
```

Calculating the difference in means

```r
treat_mean - control_mean
```

```
##   nondiscrim_mean
## 1           0.039
```

We'll see more ways to do this throughout the semester.

Checking balance on numeric covariates

We can use group_by to see how the mean of covariates varies by group:

```
trans |>
  group_by(treat_ind) |>
  summarize(age_mean = mean(age))
```

```
## # A tibble: 2 x 2
##   treat_ind age_mean
##       <dbl>    <dbl>
## 1         0     48.2
## 2         1     48.3
```

Checking balance on categorical covariates

Or we can group by treatment and a categorical control:

```
trans |>
  group_by(treat_ind, racename) |>
  summarize(n = n())
```

```
## # A tibble: 9 x 3
## # Groups:   treat_ind [2]
##   treat_ind racename             n
##       <dbl> <chr>            <int>
## 1         0 African American    58
## 2         0 Asian                2
## 3         0 Caucasian           77
## 4         0 Hispanic           150
## 5         1 African American    68
## 6         1 Asian                4
## 7         1 Caucasian           75
## 8         1 Hispanic           130
## 9         1 Native American      1
```

Hard to read!

pivot_wider

`pivot_wider()` takes data from a single column and moves it into multiple columns based on a grouping variable:

```
trans |>
  group_by(treat_ind, racename) |>
  summarize(n = n()) |>
  pivot_wider(
    names_from = treat_ind,
    values_from = n
  )
```

`names_from` tells us what variable will map onto the columns

`values_from` tell us what values should be mapped into those columns

```r
trans |>
  group_by(treat_ind, racename) |>
  summarize(n = n()) |>
  pivot_wider(
    names_from = treat_ind,
    values_from = n
  )
```

```
## # A tibble: 5 x 3
##   racename           `0`   `1`
##   <chr>            <int> <int>
## 1 African American    58    68
## 2 Asian                2     4
## 3 Caucasian           77    75
## 4 Hispanic           150   130
## 5 Native American     NA     1
```

Calculating diff-in-means by group

```
trans |>
  mutate(
    treat_ind = if_else(treat_ind == 1, "Treated", "Control"),
    party = if_else(democrat == 1, "Democrat", "Non-Democrat")
  ) |>
  group_by(treat_ind, party) |>
  summarize(nondiscrim_mean = mean(nondiscrim_post)) |>
  pivot_wider(
    names_from = treat_ind,
    values_from = nondiscrim_mean
  ) |>
  mutate(
    diff_in_means = Treated - Control
  )
```

```
## # A tibble: 2 x 4
##   party        Control Treated diff_in_means
##   <chr>          <dbl>   <dbl>         <dbl>
## 1 Democrat       0.704   0.754        0.0498
## 2 Non-Democrat   0.605   0.628        0.0234
```

```
## # A tibble: 2 x 4
##   party        Control Treated diff_in_means
##   <chr>          <dbl>   <dbl>         <dbl>
## 1 Democrat       0.704   0.754        0.0498
## 2 Non-Democrat   0.605   0.628        0.0234
```