Lectures on causal inference and experimental methods

Macartan Humphreys

Section 1

Topics

Topics

- Diff in Diff
- RDD
- Survey experiments
- LATE
- Spillovers
- Mediation

Subsection 1

The list experiment

The list experiment

- Survey experiments are used to *measure* things: nothing (except answers) should be *changed*!
- If the experiment in the survey is changing things then it is a field experiment in a survey, not a survey experiment

Design

```
declaration_17.3 <-
  declare_model(
    N = 500,
    control_count = rbinom(N, size = 3, prob = 0.5),
    Y_star = rbinom(N, size = 1, prob = 0.3),
    potential_outcomes(Y_list ~ Y_star * Z + control_count)
  ) +
  declare_inquiry(prevalence_rate = mean(Y_star)) +
  declare_assignment(Z = complete_ra(N)) +
  declare_measurement(Y_list = reveal_outcomes(Y_list ~ Z)) +
  declare_estimator(Y_list ~ Z, .method = difference_in_means,
                    inquiry = "prevalence_rate")

diagnosands <- declare_diagnosands(
  bias = mean(estimate - estimand),
  mean_CI_width = mean(conf.high - conf.low)
)
```

Negatively correlated items

- How would estimates be affected if the items selected for the list were *negatively* correlated?
- How would subject protection be affected?

Diagnosis

```
diagnose_design(declaration_17.3, diagnosands = diagnosands)
```

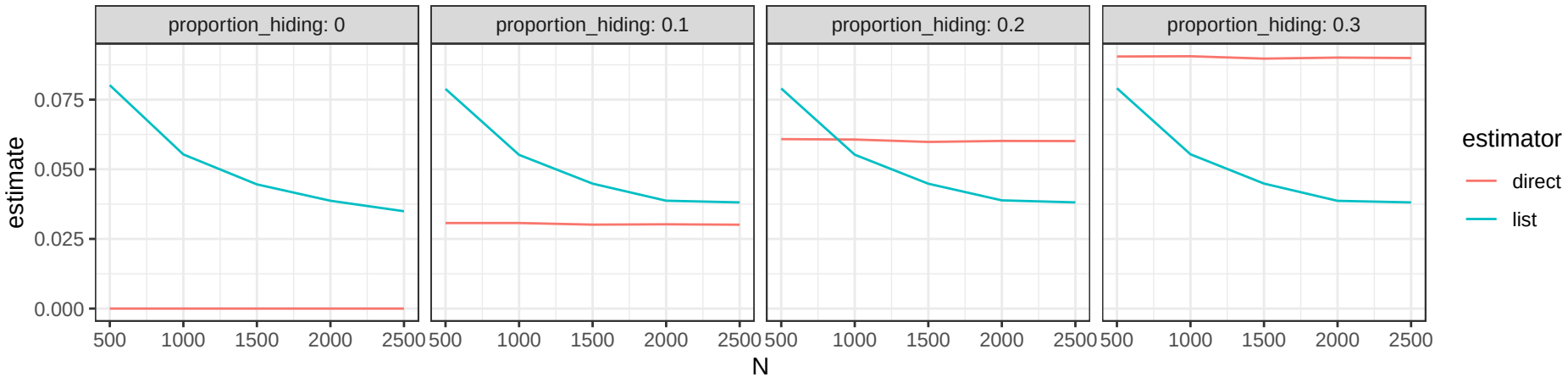
Design	Inquiry	Bias	Mean CI Width
declaration_17.3	prevalence_rate	0.00	0.32
		(0.00)	(0.00)

Tradeoffs: is the question really sensitive?

```
declaration_17.4 <-
  declare_model(
    N = N,
    U = rnorm(N),
    control_count = rbinom(N, size = 3, prob = 0.5),
    Y_star = rbinom(N, size = 1, prob = 0.3),
    W = case_when(Y_star == 0 ~ OL,
                  Y_star == 1 ~ rbinom(N, size = 1, prob = pro
    potential_outcomes(Y_list ~ Y_star * Z + control_count)
  ) +
  declare_inquiry(prevalence_rate = mean(Y_star)) +
  declare_assignment(Z = complete_ra(N)) +
  declare_measurement(Y_list = reveal_outcomes(Y_list ~ Z),
                      Y_direct = Y_star - W) +
  declare_estimator(Y_list ~ Z, inquiry = "prevalence_rate", I
  declare_estimator(Y_direct ~ 1, inquiry = "prevalence_rate",
```

Diagnosis

```
declaration_17.4 |>
  redesign(proportion_hiding = seq(from = 0, to = 0.3, by = 0.
           N = seq(from = 500, to = 2500, by = 500)) |>
  diagnose_design()
```



Hiders and liars

- Note that here we looked at "hiders" – people not answering the direct question truthfull
- See @li2019relaxing on bounds when the "no liars" assumption is threatened — this is about whether people respond truthfully to the list experimental question

Subsection 2

Noncompliance and the LATE estimand

LATE—Local Average Treatment Effects

Sometimes you give a medicine but only a non random sample of people actually try to use it. Can you still estimate the medicine's effect?

	X=0	X=1
T=0	$\overline{y}_{00}$ (n_{00})	$\overline{y}_{01}$ (n_{01})
T=1	$\overline{y}_{10}$ (n_{10})	$\overline{y}_{11}$ (n_{11})

Say that people are one of 3 types:

1. n_a "always takers'' have $X = 1$ no matter what and have average outcome $\overline{y}_a$
2. n_n never takers have $X = 0$ no matter what with outcome $\overline{y}_n$
3. n_c compliers have $X = T$ and average outcomes $\overline{y}_c^1$ if treated and $\overline{y}_c^0$ if not.

LATE—Local Average Treatment Effects

Sometimes you give a medicine but only a non random sample of people actually try to use it. Can you still estimate the medicine's effect?

	X=0	X=1
T=0	$\overline{y}_{00}$ (n_{00})	$\overline{y}_{01}$ (n_{01})
T=1	$\overline{y}_{10}$ (n_{10})	$\overline{y}_{11}$ (n_{11})

We can figure something about types:

	$X = 0$	$X = 1$
$T = 0$	$\frac{\frac{1}{2}n_c}{\frac{1}{2}n_c+\frac{1}{2}n_n}\overline{y}_c^0 + \frac{\frac{1}{2}n_n}{\frac{1}{2}n_c+\frac{1}{2}n_n}\overline{y}_n$	$\overline{y}_a$
$T = 1$	$\overline{y}_n$	$\frac{\frac{1}{2}n_c}{\frac{1}{2}n_c+\frac{1}{2}n_a}\overline{y}_c^1 +$

LATE—Local Average Treatment Effects

You give a medicine to 50% but only a non random sample of people actually try to use it. Can you still estimate the medicine's effect?

	$X=0$	$X=1$
$T=0$	$\frac{n_c}{n_c+n_n}\overline{y}_c^0 + \frac{n_n}{n_c+n_n}\overline{y}_n$	$\overline{y}_a$
(n)	$(\frac{1}{2}(n_c+n_n))$	$(\frac{1}{2}n_a)$
$T=1$	$\overline{y}_n$	$\frac{n_c}{n_c+n_a}\overline{y}_c^1 + \frac{n_a}{n_c+n_a}\overline{y}_a$
(n)	$(\frac{1}{2}n_n)$	$(\frac{1}{2}(n_a+n_c))$

Average in $T=0$ group: $\frac{n_c\overline{y}_c^0+(n_n\overline{y}_n+n_a\overline{y}_a)}{n_a+n_c+n_n}$

Average in $T=1$ group: $\frac{n_c\overline{y}_c^1+(n_n\overline{y}_n+n_a\overline{y}_a)}{n_a+n_c+n_n}$

Difference: $ITT = (\overline{y}_c^1 - \overline{y}_c^0)\frac{n_c}{n}$

The good and the bad of LATE

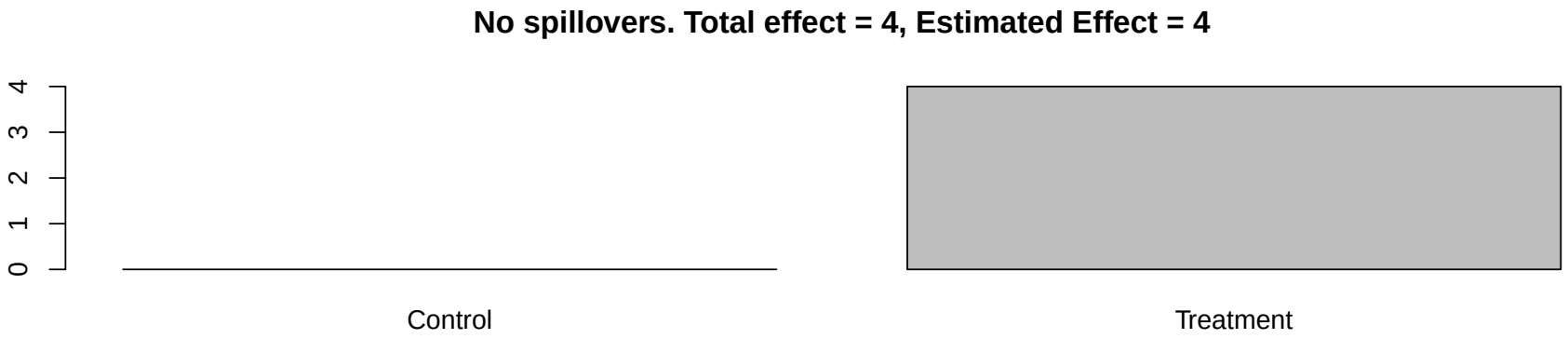
- You get a well-defined estimate even when there is non-random take-up
- May sometimes be used to assess mediation or knock-on effects
- But:
  - You need assumptions (monotonicity and the exclusion restriction – *where were these used above*?)
  - Your estimate is only for a subpopulation
  - The subpopulation is not chosen by you and is unknown
  - Different encouragements may yield different estimates since they may encourage different subgroups

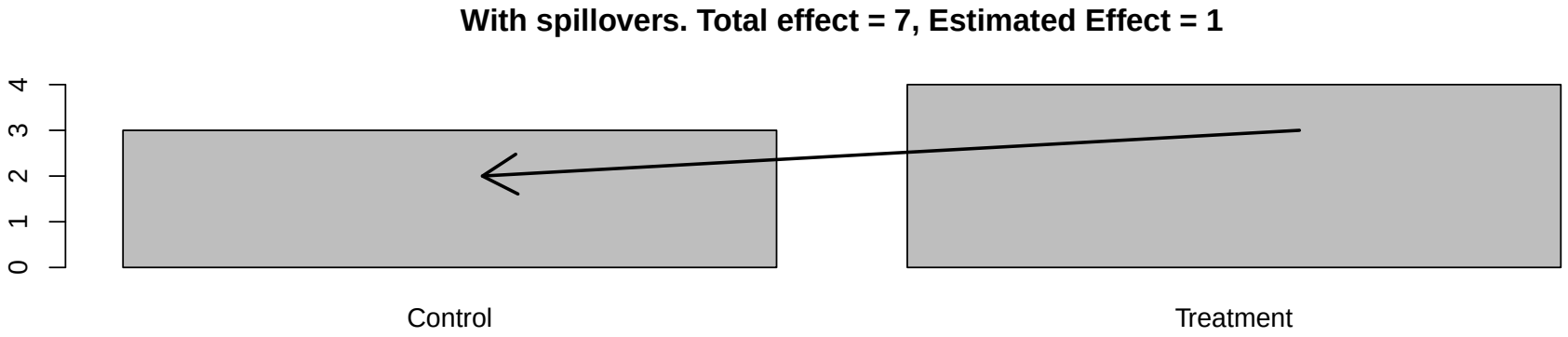
Subsection 3

Spillovers

SUTVA violations (Spillovers)

Spillovers can result in the estimation of weaker effects when effects are actually stronger.





SUTVA violations

More completely specified potential outcomes (and estimands)

			0		**1**		**2**		**3**		**4**
Unit	Location	$D_\emptyset$	$y(D_\emptyset)$	D_1	$y(D_1)$	D_2	$y(D_2)$	D_3	$y(D_3)$	D_4	$y(D_4)$
A	1	0	0	1	3	0	1	0	0	0	0
B	2	0	0	0	3	1	3	0	3	0	0
C	3	0	0	0	0	0	3	1	3	0	3
D	4	0	0	0	0	0	0	0	1	1	3

SUTVA violations

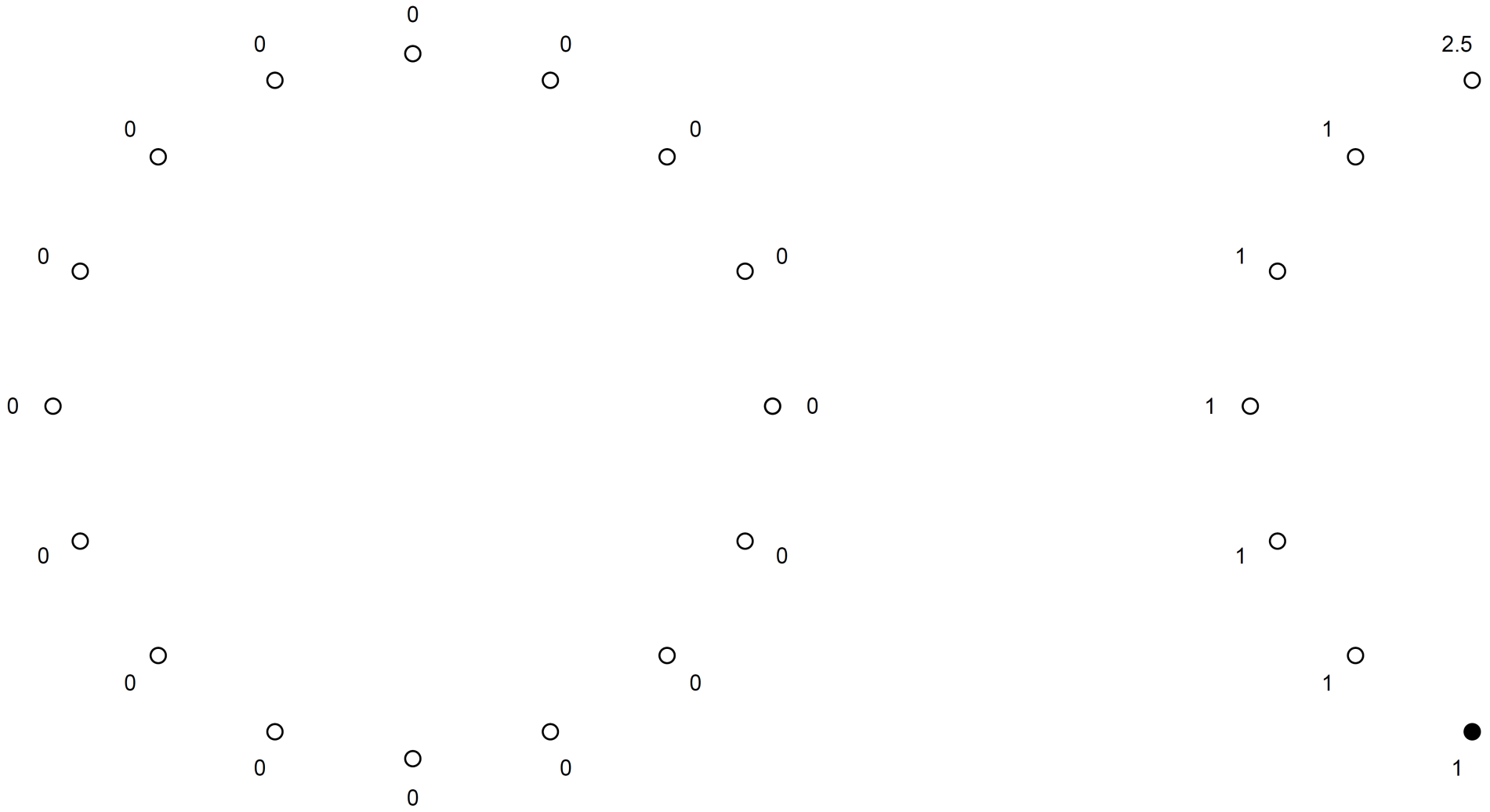
Unit	Location	$D_\emptyset$	$y(D_\emptyset)$	D_1	$y(D_1)$	D_2	$y(D_2)$	D_3	$y(D_3)$	D_4	$y(D_4)$
A	1	0	0	1	3	0	1	0	0	0	0
B	2	0	0	0	3	1	3	0	3	0	0
C	3	0	0	0	0	0	3	1	3	0	3
D	4	0	0	0	0	0	0	0	1	1	3

- The key is to think through the structure of spillovers.
- Here immediate neighbors are exposed
- In this case we can **define a direct treatment** (being exposed) and **an indirect treatment** (having a neighbor exposed) and we can work out *the **propensity** for each unit of receiving each type of treatment*
- These may be non uniform (here central types are more likely to have teated neighbors); but we can still use the randomization to assess effects

Idea: You can use the design to get a handle on spillovers

SUTVA violations

Even still, to estimate effects you need some SUTVA like assumption.

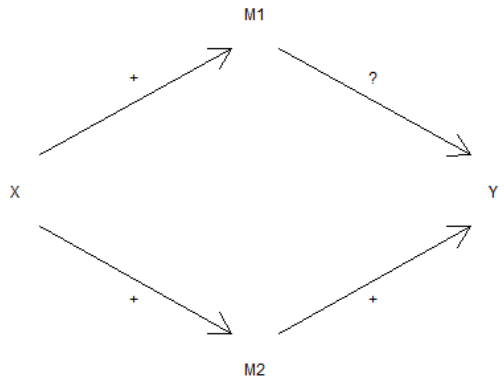


Subsection 4

Mediation

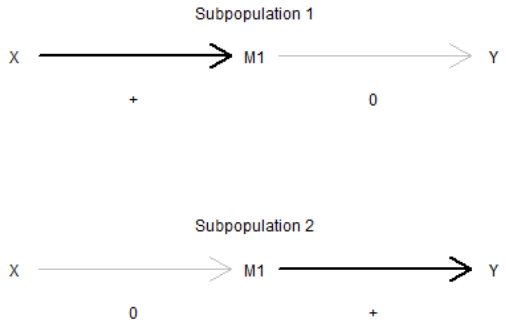
The problem of unidentified mediators

- Consider a causal system like the below.
- The effect of X on M1 and M2 can be measured in the usual way.
- But unfortunately, if there are multiple mediators, the effect of M1 (or M2) on Y is not identified.
- The 'exclusion restriction' is obviously violated when there are multiple mediators (unless you can account for them all).



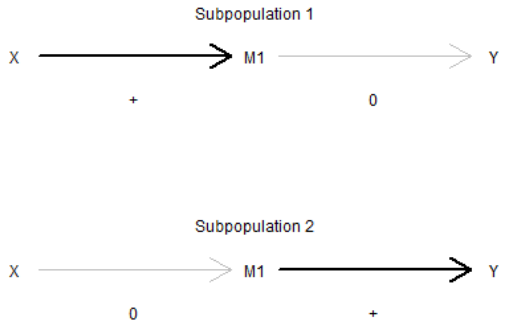
The problem of unidentified mediators

An obvious approach is to first examine the (average) effect of X on M1 and then use another manipulation to examine the (average) effect of M1 on Y. But **both of these average effects may be positive (for example) even if there is no effect of X on Y through M1**.



The problem of unidentified mediators

An obvious approach is to first examine the (average) effect of X on M1 and then use another manipulation to examine the (average) effect of M1 on Y. Similarly **both of these average effects may be zero even if X affects on Y through M1 for every unit!**.



The problem of unidentified mediators

Another somewhat obvious approach is to see how the effect of X on Y in a regression is reduced when you control for M. If the effect of X on Y passes through M then surely there should be no effect of X on Y after you control for M. But this common strategy is also not guaranteed to produce reliable results.

- See Imai on better ways to think about this problem and designs to address it.

The problem of unidentified mediators: Quantities

- In the potential outcomes framework we can describe a **mediation effect** as (see Imai et al):

$$\delta_i(t) = Y_i(t, M_i(1)) - Y_i(t, M_i(0)) \textbf{ for } t = 0, 1$$

- The **direct effect** is:

$$\psi_i(t) = Y_i(1, M_i(t)) - Y_i(0, M_i(t)) \textbf{ for } t = 0, 1$$

- This is a **decomposition**, since:

$$Y_i(1, M_i(1)) - Y_1(0, M_i(0)) = \frac{1}{2}(\delta_i(1) + \delta_i(0) + \psi_i(1) + \psi_i(0))$$

- If (and a big if), there are no interaction effects—ie $\delta_i(1) = \delta_i(0), \psi_i(1) = \psi_i(0)$, then

$$Y_i(1, M_i(1)) - Y_1(0, M_i(0)) = \delta_i + \psi_i$$

- The bad news is that although a single experiment might identify the total effect, it can not identify these elements of the direct effect.

The problem of unidentified mediators: Solutions?

- Check **formal requirement** for identification under single experiment design ("sequential ignorability"—that, conditional on actual treatment, it is as if the value of the mediation variable is randomly assigned relative to potential outcomes). But this is strong (and in fact unverifiable) and if it does not hold, bounds on effects always include zero (Imai et al)
- You can use **interactions** with covariates **if you are willing to make assumptions on no heterogeneity of direct treatment effects** over covariates. eg you think that money makes people get to work faster because they can buy better cars; you look at the marginal effect of more money on time to work for people with and without cars and find it higher for the latter. This might imply mediation through transport but only if there is no direct effect heterogeneity (eg people with cars are less motivated by money).

The problem of unidentified mediators: Solutions?

- Weaker assumptions justify **parallel design**
  - Group A: T is randomly assigned, M left free.
  - Group B: divided into four groups $T \times M$ (requires two more assumptions (1) that the **manipulation** of the mediator only affects outcomes through the mediator (2) **no interaction**, for each unit, $Y(1, m) - Y(0, m) = Y(1, m') - Y(0, m')$.)

Idea 5: Understanding mechanisms is harder than you think. Figure out what assumptions fly.

Subsection 5

Differences in differences

Differences in differences

New challenges, new developments

Subsection 6

Regression discontintuity

Regression discontintuity

Errors and diagnostics