Sensitivity Analyses

Ryan T. Moore

American University

The Lab @ DC

2024-07-16

Table of contents I

Sensitivity

What is “sensitivity”?

When inputs change, do outputs change?

What is “sensitivity”?

When inputs change, do outputs change?

- With different variables in model, does parameter of interest change?

What is “sensitivity”?

When inputs change, do outputs change?

- With different variables in model, does parameter of interest change?
- With different assumptions about error structures, does causal mediation estimate change?

What is “sensitivity”?

When inputs change, do outputs change?

- With different variables in model, does parameter of interest change?
- With different assumptions about error structures, does causal mediation estimate change?
- With different data collected, would causal conclusion change?

Sensitivity to Model Specification

Should we trust our model?

Suppose I present observational results:

	Coefficient
$\geq$ 1000 auto workers	0.87
	(0.39)
DW-NOMINATE	-5.04
	(0.53)
Ford/Chrysler/GM PAC Contribs (log)	0.15
	(0.05)
AFL-CIO PAC Contribs (log)	0.09
	(0.04)
Intercept	-0.14
	(0.30)
N	406
AIC	258.59
BIC	338.72
log L	-109.30

Should we trust our model?

Suppose I present observational results:

	Coefficient
$\geq$ 1000 auto workers	0.87
	(0.39)
DW-NOMINATE	-5.04
	(0.53)
Ford/Chrysler/GM PAC Contribs (log)	0.15
	(0.05)
AFL-CIO PAC Contribs (log)	0.09
	(0.04)
Intercept	-0.14
	(0.30)
N	406
AIC	258.59
BIC	338.72
$\log L$	-109.30

What would be your questions?

Multiplicity of Models

- For k predictors, there are 2^k possible models

Multiplicity of Models

- For k predictors, there are 2^k possible models
- For 30 predictors, 10^9 possible models

Multiplicity of Models

- For k predictors, there are 2^k possible models
- For 30 predictors, 10^9 possible models
- ($7\times$ as many km to the sun!)

Multiplicity of Models

- For k predictors, there are 2^k possible models
- For 30 predictors, 10^9 possible models
- ($7\times$ as many km to the sun!)
- So, typically, we show

Multiplicity of Models

	Model 1	Model 2	Model 3	Model 4
Intercept	2.12	0.61	1.26	-0.14
	(0.24)	(0.20)	(0.31)	(0.30)
$\geq$ 1000 auto workers	0.98	1.14	0.74	0.87
	(0.34)	(0.37)	(0.36)	(0.39)
Republican				
	(0.33)		(0.42)	
DW-NOMINATE		-4.87		-5.04
		(0.42)		(0.53)
Ford/Chrysler/GM PAC Contribs (log)			0.14	0.15
			(0.05)	(0.05)
AFL-CIO PAC Contribs (log)			0.14	0.09
			(0.04)	(0.04)
N	407	406	407	406
AIC	301.48	268.76	284.11	258.59
BIC	349.58	316.83	364.29	338.72
$\log L$	-138.74	-122.38	-122.06	-109.30

Multiplicity of Models

	Model 1	Model 2	Model 3	Model 4
Intercept	2.12	0.61	1.26	-0.14
	(0.24)	(0.20)	(0.31)	(0.30)
$\geq$ 1000 auto workers	0.98	1.14	0.74	0.87
	(0.34)	(0.37)	(0.36)	(0.39)
Republican				
	(0.33)		(0.42)	
DW-NOMINATE		-4.87		-5.04
		(0.42)		(0.53)
Ford/Chrysler/GM PAC Contribs (log)			0.14	0.15
			(0.05)	(0.05)
AFL-CIO PAC Contribs (log)			0.14	0.09
			(0.04)	(0.04)
N	407	406	407	406
AIC	301.48	268.76	284.11	258.59
BIC	349.58	316.83	364.29	338.72
$\log L$	-138.74	-122.38	-122.06	-109.30

Rather than small set of substantively-informed models, just show them all!

Example 1: Congressional Auto Industry Support

- “Great Recession” following global financial crisis of 2008-2009 (“subprime mortage crisis”)

Example 1: Congressional Auto Industry Support

- “Great Recession” following global financial crisis of 2008-2009 (“subprime mortage crisis”)
- Two big bills in US Congress to shore up US auto industry

Example 1: Congressional Auto Industry Support

- "Great Recession" following global financial crisis of 2008-2009 ("subprime mortage crisis")
- Two big bills in US Congress to shore up US auto industry
  - Auto Bailout: $80bn to GM and Chrysler

Example 1: Congressional Auto Industry Support

- “Great Recession” following global financial crisis of 2008-2009 (“subprime mortage crisis”)
- Two big bills in US Congress to shore up US auto industry
  - Auto Bailout: $80bn to GM and Chrysler
  - Cash for Clunkers: $3bn to consumer rebates

Example 1: Congressional Auto Industry Support

- “Great Recession” following global financial crisis of 2008-2009 (“subprime mortage crisis”)
- Two big bills in US Congress to shore up US auto industry
  - Auto Bailout: $80bn to GM and Chrysler
  - Cash for Clunkers: $3bn to consumer rebates

Example 1: Congressional Auto Industry Support

- “Great Recession” following global financial crisis of 2008-2009 (“subprime mortage crisis”)
- Two big bills in US Congress to shore up US auto industry
  - Auto Bailout: \$80bn to GM and Chrysler
  - Cash for Clunkers: \$3bn to consumer rebates

Moore, Powell, and Reeves (2013): two quasi-private, particularistic bills.

Estimate relationship

(presence of auto factories) $\Rightarrow$ (Congressional votes)

Example 1: Congressional Auto Industry Support

- “Great Recession” following global financial crisis of 2008-2009 (“subprime mortage crisis”)
- Two big bills in US Congress to shore up US auto industry
  - Auto Bailout: \$80bn to GM and Chrysler
  - Cash for Clunkers: \$3bn to consumer rebates

Moore, Powell, and Reeves (2013): two quasi-private, particularistic bills.

Estimate relationship

(presence of auto factories) $\Rightarrow$ (Congressional votes)

Claim: **Local econ interests** at least on par w/ corporate campaign contributions, lobbying, public positions.

Moore, Powell, and Reeves (2013)

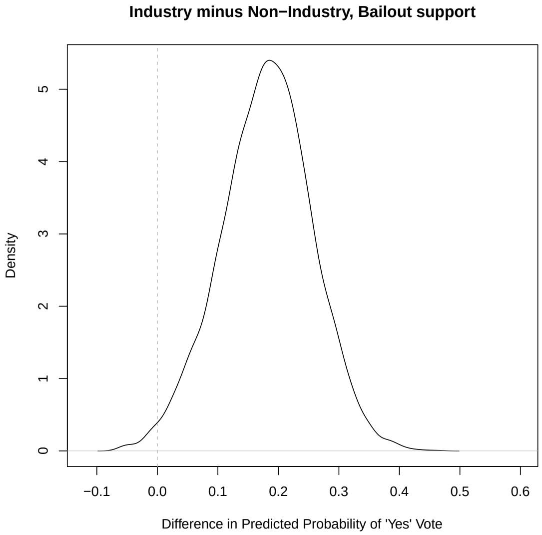


Figure 1: First diffs, predicted prob MoC supports auto bailout, member from industry v. non-industry district, other vars at means.

Moore, Powell, and Reeves (2013)

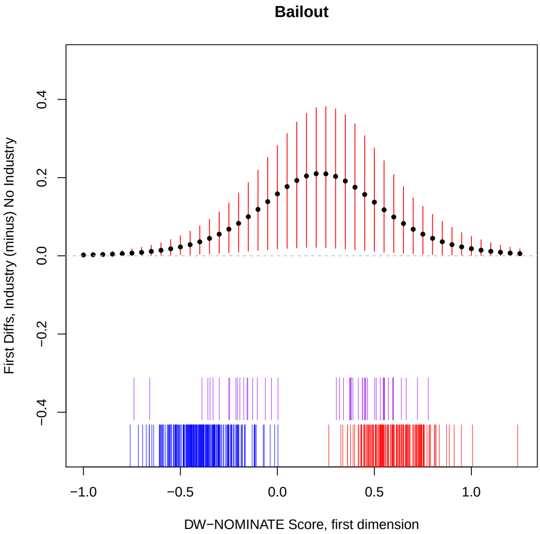


Figure 2: First diffs, industry v. non-industry district member prob of supporting bailout positive at any value of DW-NOMINATE score.

Insensitivity to Specification

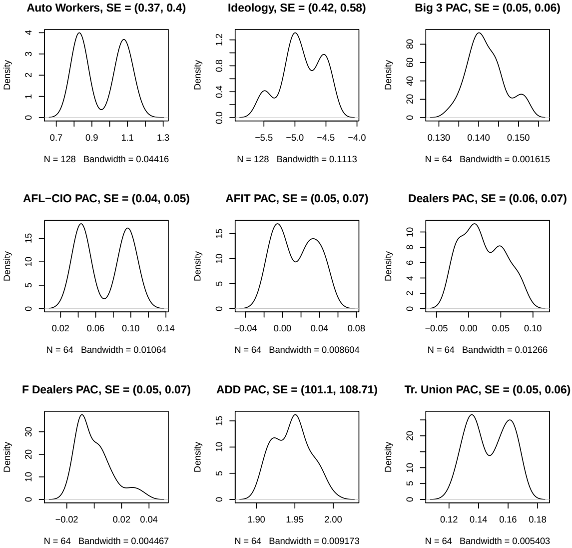


Figure 3: Industry presence coef always positive in Bailout logistic regressions. Coef densities w/ industry presence and

Insensitivity to Specification

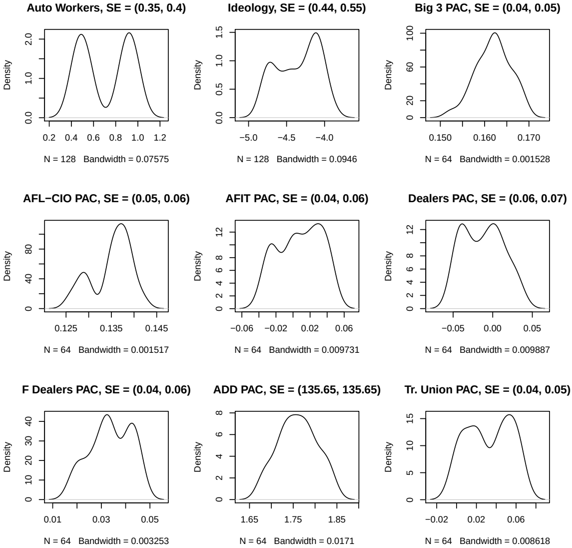


Figure 4: Industry presence coef always positive in Cash for Clunkers logistic regressions. Coef densities w/ industry presence and

Implementation

```
library(olsrr)
```

- Estimate (all) linear models

Implementation

```r
library(olsrr)
```

- Estimate (all) linear models
- Provide model fit stats

Implementation

```r
library(olsrr)
```

- Estimate (all) linear models
- Provide model fit stats
- Provide coefs

Implementation

```
library(olsrr)
```

- Estimate (all) linear models
- Provide model fit stats
- Provide coefs

Implementation

```r
library(olsrr)
```

- Estimate (all) linear models
- Provide model fit stats
- Provide coefs

Hebbali (2024)

Example 2: Social Pressure Mailers

Example 2: Social Pressure Mailers

Dear Registered Voter:

WHAT IF YOUR NEIGHBORS KNEW WHETHER YOU VOTED?

Why do so many people fail to vote? We've been talking about the problem for years, but it only seems to get worse. This year, we're taking a new approach. We're sending this mailing to you and your neighbors to publicize who does and does not vote.

The chart shows the names of some of your neighbors, showing which have voted in the past. After the August 8 election, we intend to mail an updated chart. You and your neighbors will all know who voted and who did not.

DO YOUR CIVIC DUTY — VOTE!

MAPLE DR	Aug 04	Nov 04	Aug 06
9995 JOSEPH JAMES SMITH	Voted	Voted	______
9995 JENNIFER KAY SMITH		Voted	______
9997 RICHARD B JACKSON		Voted	______
9999 KATHY MARIE JACKSON		Voted	______
9999 BRIAN JOSEPH JACKSON		Voted	______
9991 JENNIFER KAY THOMPSON		Voted	______
9991 BOB R THOMPSON		Voted	______
9993 BILL S SMITH			______
9989 WILLIAM LUKE CASPER		Voted	______
9989 JENNIFER SUE CASPER		Voted	______
9987 MARIA S JOHNSON	Voted	Voted	

Example 2: Social Pressure Mailers

```r
library(qss)
data(social)
```

```r
social |> select(-yearofbirth) |> head()
```

```
     sex primary2004    messages primary2006 hhsize age
1   male           0 Civic Duty           0      2  65
2 female           0 Civic Duty           0      2  59
3   male           0   Hawthorne           1      3  55
4 female           0   Hawthorne           1      3  56
5 female           0   Hawthorne           1      3  24
6   male           0     Control           0      3  25
```

Example 2: Social Pressure Mailers

```
lm_out <- lm(primary2006 ~ messages + sex + age +
               primary2004 + hhsize, data = social)

all_lm_social <- ols_step_all_possible(lm_out)$result

dim(all_lm_social)
```

```
[1] 31 15
```

```
head(all_lm_social)
```

```
   mindex n       predictors      rsquare         adjr
4       1 1      primary2004 0.0261502651 0.0261470812 0.457
3       2 1              age 0.0167659386 0.0167627240 0.459
1       3 1         messages 0.0032825640 0.0032727879 0.462
5       4 1           hhsize 0.0025142362 0.0025109749 0.462
2       5 1              sex 0.0001863186 0.0001830498 0.463
13      6 2  age primary2004 0.0409175309 0.0409112596 0.453
```

Example 2: Social Pressure Mailers

```r
all_lm_social_coefs <- ols_step_all_possible_betas(lm_out)

all_lm_social_coefs
```

```
   model         predictor         beta
1      1       (Intercept)  0.2966383083
2      1 messagesCivic Duty  0.0178993441
3      1  messagesHawthorne  0.0257363121
4      1  messagesNeighbors  0.0813099129
5      2       (Intercept)  0.3059095493
6      2           sexmale  0.0126509479
7      3       (Intercept)  0.1055564253
8      3               age  0.0041515670
9      4       (Intercept)  0.2508820413
10     4       primary2004  0.1528795252
11     5       (Intercept)  0.3763534949
12     5            hhsize -0.0293482475
13     6       (Intercept)  0.2902800648
```

Example 2: Social Pressure Mailers

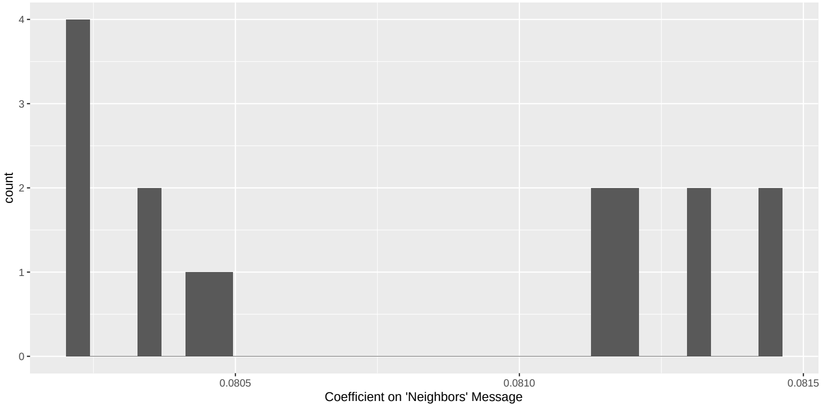


Figure 5: 'Neighbors' Coefs from All Possible Regressions

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.08023 0.08032 0.08081 0.08080 0.08122 0.08145
```

All Coefficients

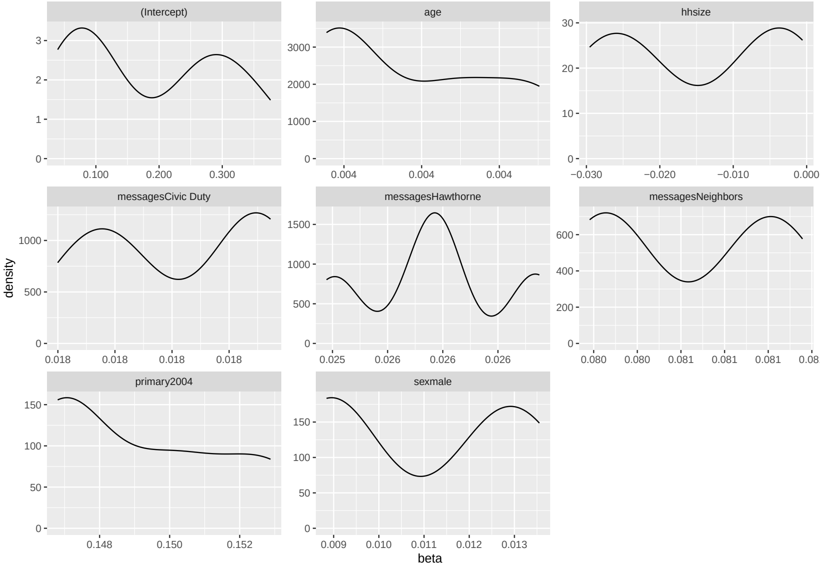


Figure 6: Coefs from All Possible Regressions

Preprocessing to Control Sensitivity

- So far, “show all the models”

Preprocessing to Control Sensitivity

- So far, "show all the models"
- Better: preprocess data to minimize effects of model-based adjustment

Preprocessing to Control Sensitivity

- So far, "show all the models"
- Better: preprocess data to minimize effects of model-based adjustment
- Match, subclassify

Preprocessing to Control Sensitivity

- So far, "show all the models"
- Better: preprocess data to minimize effects of model-based adjustment
- Match, subclassify

Preprocessing to Control Sensitivity

- So far, “show all the models”
- Better: preprocess data to minimize effects of model-based adjustment
- Match, subclassify

“model-based adjustments ...will give basically the same point estimates”

Matching

X	T	$Y(0)$	$Y(1)$	Y^{obs}
1	1	1	2	2
1	0	1	2	1
1	0	1	2	1
2	1	2	3	3
2	1	2	3	3
2	0	2	3	2

Matching

X	T	$Y(0)$	$Y(1)$	Y^{obs}
1	1	1	2	2
1	0	1	2	1
1	0	1	2	1
2	1	2	3	3
2	1	2	3	3
2	0	2	3	2

- $\tau_i = 1 \quad \forall i$

Matching

X	T	$Y(0)$	$Y(1)$	Y^{obs}
1	1	1	2	2
1	0	1	2	1
1	0	1	2	1
2	1	2	3	3
2	1	2	3	3
2	0	2	3	2

- $\tau_i = 1 \quad \forall i$
- $ATE = \overline{Y(1) - Y(0)} = 1$

Matching

X	T	$Y(0)$	$Y(1)$	Y^{obs}
1	1	1	2	2
1	0	1	2	1
1	0	1	2	1
2	1	2	3	3
2	1	2	3	3
2	0	2	3	2

- $\tau_i = 1 \quad \forall i$
- $ATE = \overline{Y(1) - Y(0)} = 1$
- $\widehat{ATE} = \left(\overline{Y(1)}|T = 1\right) - \left(\overline{Y(0)}|T = 0\right) = \frac{8}{3} - \frac{4}{3} = \frac{4}{3}$

Matching

Suppose we 1:1 exact match on X:

X	T	$Y(0)$	$Y(1)$	Y^{obs}
1	1	1	2	2
1	0	1	2	1
~~1~~	~~0~~	~~1~~	~~2~~	~~1~~
2	1	2	3	3
~~2~~	~~1~~	~~2~~	~~3~~	~~3~~
2	0	2	3	2

Matching

Suppose we 1:1 exact match on X:

X	T	$Y(0)$	$Y(1)$	Y^{obs}
1	1	1	2	2
1	0	1	2	1
~~1~~	~~0~~	~~1~~	~~2~~	~~1~~
2	1	2	3	3
~~2~~	~~1~~	~~2~~	~~3~~	~~3~~
2	0	2	3	2

$$\widehat{ATE}_m = \left(\overline{Y_m(1)}|T=1\right) - \left(\overline{Y_m(0)}|T=0\right) = \frac{5}{2} - \frac{3}{2} = 1$$

Matching

Suppose we 1:1 exact match on X:

X	T	$Y(0)$	$Y(1)$	Y^{obs}
1	1	1	2	2
1	0	1	2	1
~~1~~	~~0~~	~~1~~	~~2~~	~~1~~
2	1	2	3	3
~~2~~	~~1~~	~~2~~	~~3~~	~~3~~
2	0	2	3	2

$$\widehat{ATE}_m = \left(\overline{Y_m(1)}|T=1\right) - \left(\overline{Y_m(0)}|T=0\right) = \frac{5}{2} - \frac{3}{2} = 1$$

Not just coincidence; matching removes $X \to T$.

Ho et al. (2007)

"Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference"

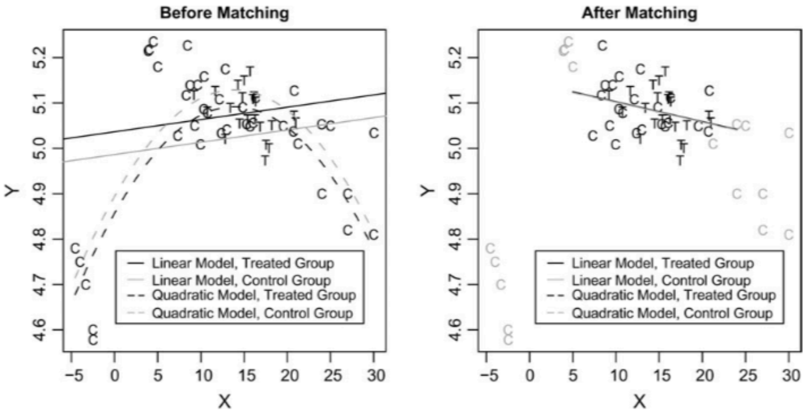


Figure 7: Before: Direction of Effect depends on Model. After: Effect indendent of Model.

Reducing Sensitivity in FDA Example

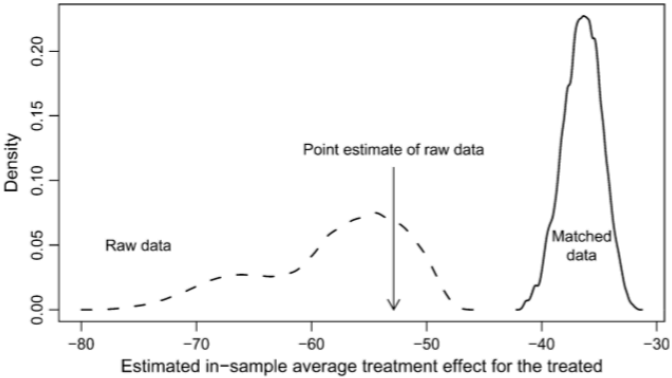


Fig. 2 Kernel density plot (a smoothed histogram) of point estimates of the in-sample ATT of the Democratic Senate majority on FDA drug approval time across 262,143 specifications. The solid line

How to Identify Sensitivity?

Different distributions; non-overlap

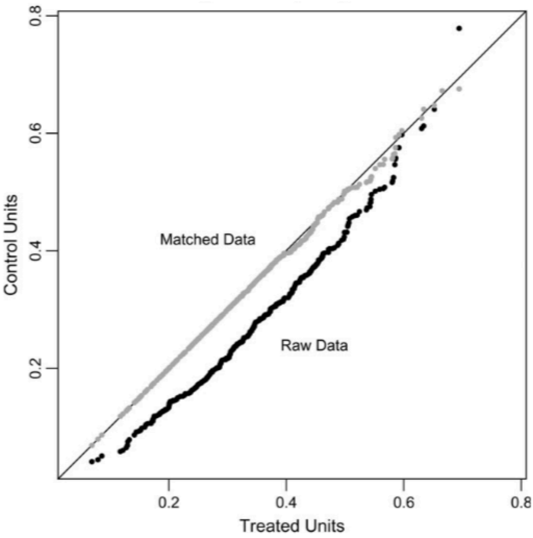


Fig. 3 QQ plot of propensity score for candidate visibility. The black dots represent empirical QQ

Reducing Sensitivity in Candidate Visibility Example

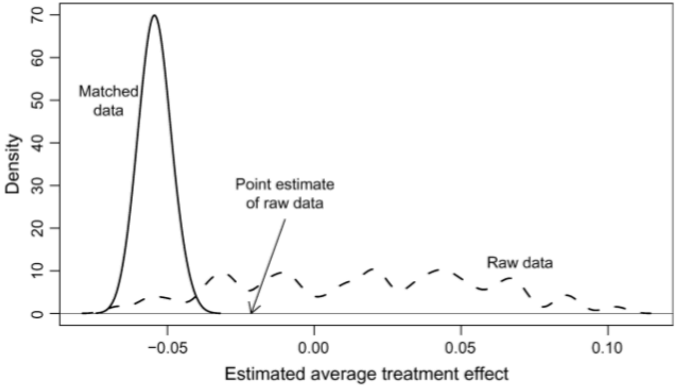


Fig. 4 Kernel density plot of point estimates of the effect of being a less visible male Republican candidate across 63 possible specifications with the Koch data. The dashed line presents estimates for

Paradox of Regression for causal inference?

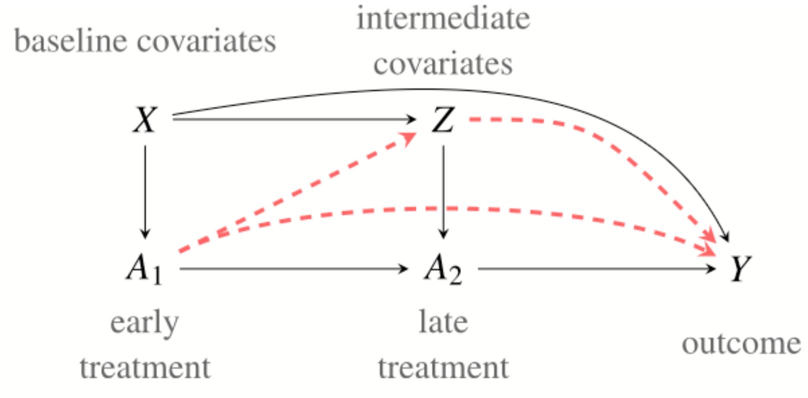
- If large diffs in distn's,
  $\rightsquigarrow$ regression not enough, very sensitive

Paradox of Regression for causal inference?

- If large diffs in distn's,
  $\rightsquigarrow$ regression not enough, very sensitive
- If small diffs in distn's,
  $\rightsquigarrow$ regression won't matter much

Dynamic Treatment Regimes

Dynamic Treatment Regimes



Blackwell and Strezhnev (2022)

Telescope Matching

Preprocessing for Dynamic Treatment Regimes:

- Match across early Tr (A_1) on baseline covariates X

Telescope Matching

Preprocessing for Dynamic Treatment Regimes:

- Match across early Tr (A_1) on baseline covariates X
- Match across late Tr (A_2) on early Tr (exact), baseline + intermediate covariates (A_1, X, Z [or X_1, X_2])

Telescope Matching

Preprocessing for Dynamic Treatment Regimes:

- Match across early Tr (A_1) on baseline covariates X
- Match across late Tr (A_2) on early Tr (exact), baseline + intermediate covariates (A_1, X, Z [or X_1, X_2])
- Use matches to impute "paths not taken"

Telescope Matching

Preprocessing for Dynamic Treatment Regimes:

Diff-in-means estimator for effect of “early treatment”:

$$\hat{\tau} \equiv \frac{1}{N} \sum_{i=1}^{N} \left(\hat{Y}_i(1,0) - \hat{Y}_i(0,0) \right)$$

Telescope Matching Example

```
library(DirectEffects)
data(jobcorps)
```

- Y: self-reported good health (0/1)
- $X1$: school/training/job before Job Corps
- $A1$: Job Corps program
- $X2$: employment in Q4 after assg
- $A2$: employment in Q just before outcome

```
# Formula: Y ~ X1 | A1 | X2 | A2

tm_form <- exhealth30 ~ schobef + trainyrbef + jobeverbef
  treat | emplq4 + emplq4full | work2year2q

tm_out <- telescope_match(tm_form, data = jobcorps, verbose
```

Telescope Matching Example

```
tm_out
```

```
Telescope matching output

Call:
telescope_match(formula = tm_form, data = jobcorps, verbose

Active treatment:  treat
Controlled treatment(s): work2year2q

Estimated controlled direct effects of treat:
  work2year2q     estimate
1           0 -0.006569616
2           1  0.029847086
```

Telescope Matching Example

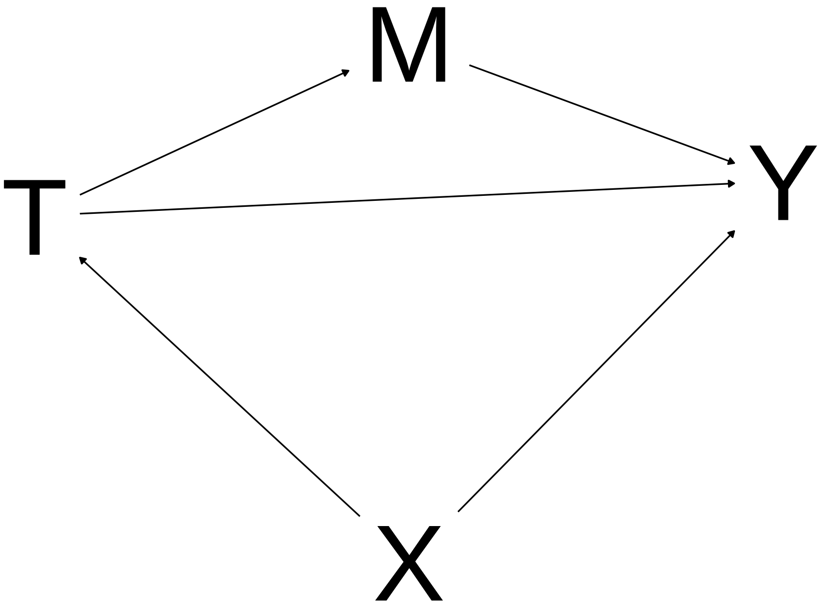
```r
summary(tm_out)
```

```
Telescope matching results
----------------------------

Call:
telescope_match(formula = tm_form, data = jobcorps, verbose = FALSE)

Active treatment:  treat
Controlled treatment(s): work2year2q

Matching summary:
         Term Matching Ratio L:1 N == 1 N == 0 Matched == 1 Matched == 0
1       treat                  5   6034   3991         5800         3989
2 work2year2q                  5   6207   3818         3655         3659


Summary of units matching contributions:
                  Min. 1st Qu. Median Mean 3rd Qu.  Max.
treat                0     0.6    0.8    1    1.40  3.80
treat:work2year2q    0     0.0    0.4    1    1.04 90.08
work2year2q          0     0.0    0.4    1    1.20 68.00


Estimated controlled direct effects of treat:
                  work2year2q Estimate Estimate (no BC) Std. Err.
(1, 0) vs. (0, 0)           0 -0.00657        -0.005897   0.03767
(1, 1) vs. (0, 1)           1  0.02985         0.029813   0.01421
```

Sensitivity to an Unidentifiable Parameter

Mediation Analysis

Confounding in Observational Studies



Mediation Effects

- If interest is $M \to Y$, seek experiment-like M

Mediation Effects

- If interest is $M \to Y$, seek experiment-like M
  - random M

Mediation Effects

- If interest is $M \to Y$, seek experiment-like M
  - random M
  - subclassify/match for M

Mediation Effects

- If interest is $M \to Y$, seek experiment-like M
  - random M
  - subclassify/match for M
  - instrumented M

Mediation Effects

- If interest is $M \to Y$, seek experiment-like M
  - random M
  - subclassify/match for M
  - instrumented M
  - RDD, synthetic control for M

Mediation Effects

- If interest is $M \to Y$, seek experiment-like M
  - random M
  - subclassify/match for M
  - instrumented M
  - RDD, synthetic control for M
- If interest is $T \to Y$, seek experimental T

Mediation Effects

- If interest is $M \to Y$, seek experiment-like M
  - random M
  - subclassify/match for M
  - instrumented M
  - RDD, synthetic control for M
- If interest is $T \to Y$, seek experimental T
  - random T

Mediation Effects

- If interest is $M \to Y$, seek experiment-like M
  - random M
  - subclassify/match for M
  - instrumented M
  - RDD, synthetic control for M
- If interest is $T \to Y$, seek experimental T
  - random T
  - subclassify/match for T

Mediation Effects

- If interest is $M \rightarrow Y$, seek experiment-like M
  - random M
  - subclassify/match for M
  - instrumented M
  - RDD, synthetic control for M
- If interest is $T \rightarrow Y$, seek experimental T
  - random T
  - subclassify/match for T
  - instrumented T

Mediation Effects

- If interest is $M \to Y$, seek experiment-like M
  - random M
  - subclassify/match for M
  - instrumented M
  - RDD, synthetic control for M
- If interest is $T \to Y$, seek experimental T
  - random T
  - subclassify/match for T
  - instrumented T
  - RDD, synthetic control for T

Mediation Effects

- If interest is $M \rightarrow Y$, seek experiment-like M
  - random M
  - subclassify/match for M
  - instrumented M
  - RDD, synthetic control for M
- If interest is $T \rightarrow Y$, seek experimental T
  - random T
  - subclassify/match for T
  - instrumented T
  - RDD, synthetic control for T
- In <u>mediation</u>, interest is $T \rightarrow M \rightarrow Y$

Mediation Effects

- If interest is $M \to Y$, seek experiment-like M
  - random M
  - subclassify/match for M
  - instrumented M
  - RDD, synthetic control for M
- If interest is $T \to Y$, seek experimental T
  - random T
  - subclassify/match for T
  - instrumented T
  - RDD, synthetic control for T
- In mediation, interest is $T \to M \to Y$
  - (and maybe $T \to (\neg M) \to Y$)

Mediation Effects

Condition on /control for M?

Mediation Effects

Condition on /control for M?

- No: how to estimate $M \to Y$?

Mediation Effects

Condition on /control for M?

- No: how to estimate $M \to Y$?
- Yes: induces post-treatment bias in estimate of $T \to Y$

Mediation Effects

Condition on /control for M?

- No: how to estimate $M \rightarrow Y$?
- Yes: induces post-treatment bias in estimate of $T \rightarrow Y$

Mediation Effects

Condition on /control for M?

- No: how to estimate $M \rightarrow Y$?
- Yes: induces post-treatment bias in estimate of $T \rightarrow Y$

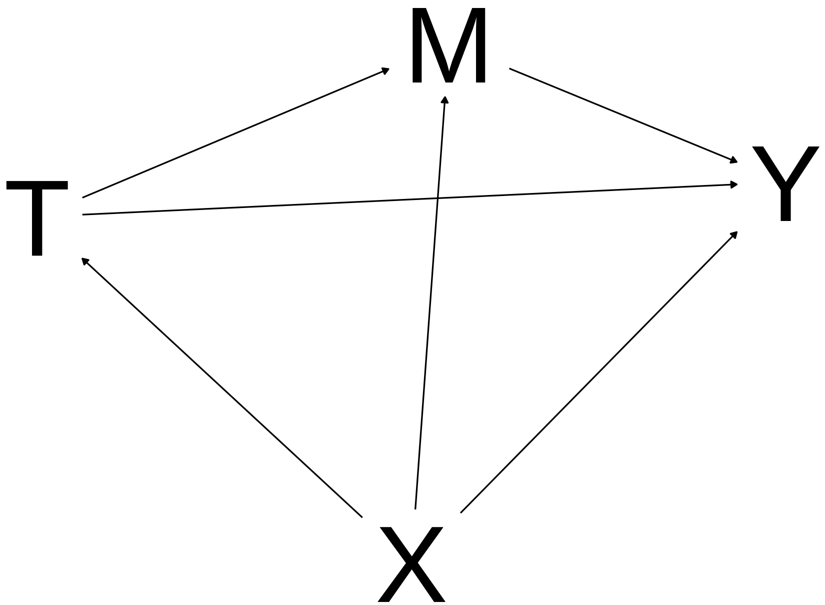
- And if $X \rightarrow M$, too?

Mediation Effects

Condition on /control for M?

- No: how to estimate $M \rightarrow Y$?
- Yes: induces post-treatment bias in estimate of $T \rightarrow Y$

- And if $X \rightarrow M$, too?
- Even worse ...


M
T
Y
X

Addressing Confounding

To break confounding,

- can't break $X \to Y$

Addressing Confounding

To break confounding,

- can't break $X \to Y$
- break $X \to T$

Addressing Confounding

To break confounding,

- can't break $X \rightarrow Y$
- break $X \rightarrow T$
- but $X \rightarrow M$ may still remain!

Post-Treatment Bias

- Interest in effect of news on attitude.

Post-Treatment Bias

- Interest in effect of news on attitude. Randomly assign news:

```r
n <- 200
news <- sample(0:1, n, replace = TRUE)
```

Post-Treatment Bias

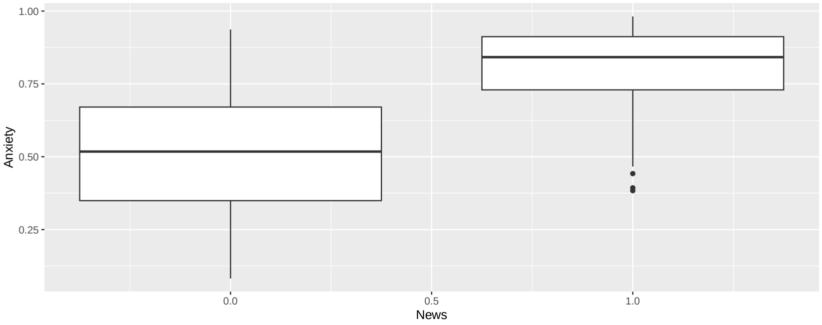
- News status greatly affects Anxiety:

```
pr.anx <- 1/(1 + exp(-(news * 2 + rnorm(n))))
```

Post-Treatment Bias

- News status greatly affects Anxiety:

```
pr.anx <- 1/(1 + exp(-(news * 2 + rnorm(n))))
```



Post-Treatment Bias

- News status greatly affects Anxiety:

```r
summary(lm(pr.anx ~ news))$coef |> round(3)
```

```
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.519      0.018  29.196        0
news           0.294      0.025  11.692        0
```

Post-Treatment Bias

- Anxiety greatly increases (negative) attitude

Post-Treatment Bias

- Anxiety greatly increases (negative) attitude
  - (but news also has other ways to increase negative attitude)

Post-Treatment Bias

- Anxiety greatly increases (negative) attitude
  - (but news also has other ways to increase negative attitude)

Post-Treatment Bias

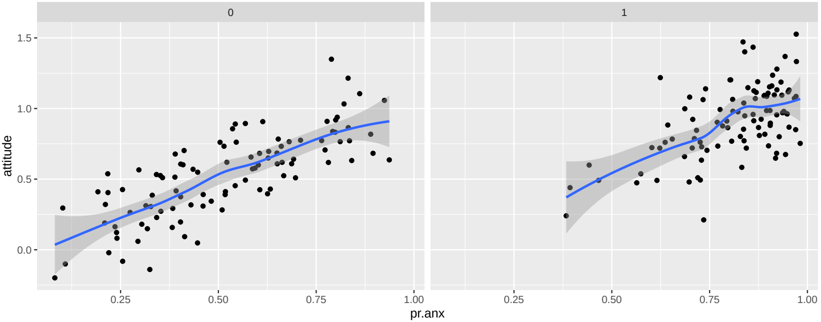
- Anxiety greatly increases (negative) attitude
  - (but news also has other ways to increase negative attitude)

```
attitude <- .1 * news + pr.anx + rnorm(n, sd = 0.2)
```

Post-Treatment Bias

- Anxiety greatly increases (negative) attitude
  - (but news also has other ways to increase negative attitude)

```
attitude <- .1 * news + pr.anx + rnorm(n, sd = 0.2)
```



Post-Treatment Bias

- Interested in causal effect of news on attitude

Post-Treatment Bias

- Interested in causal effect of news on attitude
- Analysis 1: Adjust for anxiety status:

Post-Treatment Bias

- Interested in causal effect of news on attitude
- Analysis 1: Adjust for anxiety status:

Post-Treatment Bias

- Interested in causal effect of news on attitude
- Analysis 1: Adjust for anxiety status:

```r
summary(lm(attitude ~ news + pr.anx))$coef |> round(4)
```

```
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.0438     0.0461 -0.9511   0.3427
news          0.0655     0.0368  1.7807   0.0765
pr.anx        1.0995     0.0800 13.7513   0.0000
```

Post-Treatment Bias

- Interested in causal effect of news on attitude
- Analysis 1: Adjust for anxiety status:

```r
summary(lm(attitude ~ news + pr.anx))$coef |> round(4)
```

```
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.0438     0.0461 -0.9511   0.3427
news          0.0655     0.0368  1.7807   0.0765
pr.anx        1.0995     0.0800 13.7513   0.0000
```

- Great! Now, say, multiply coefs, to get $T \to M \to Y$?

Post-Treatment Bias

- Interested in causal effect of news on attitude
- Analysis 1: Adjust for anxiety status:

```r
summary(lm(attitude ~ news + pr.anx))$coef |> round(4)
```

```
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.0438     0.0461 -0.9511   0.3427
news          0.0655     0.0368  1.7807   0.0765
pr.anx        1.0995     0.0800 13.7513   0.0000
```

- Great! Now, say, multiply coefs, to get $T \to M \to Y$?
- Problem: This doesn't work.

Post-Treatment Bias

- Interested in causal effect of news on attitude

Post-Treatment Bias

- Interested in causal effect of news on attitude
- Analysis 2: Don't control for anxiety status:

Post-Treatment Bias

- Interested in causal effect of news on attitude
- Analysis 2: Don't control for anxiety status:

Post-Treatment Bias

- Interested in causal effect of news on attitude
- Analysis 2: Don't control for anxiety status:

```r
summary(lm(attitude ~ news))$coef |> round(4)
```

```
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.5272     0.0279 18.8638        0
news          0.3889     0.0395  9.8400        0
```

Post-Treatment Bias

- Interested in causal effect of news on attitude
- Analysis 2: Don't control for anxiety status:

```r
summary(lm(attitude ~ news))$coef |> round(4)
```

```
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.5272     0.0279 18.8638        0
news          0.3889     0.0395  9.8400        0
```

- Great! Now, say, subtract coef Analysis 1 from this, to get $T \to M \to Y$?

Post-Treatment Bias

- Interested in causal effect of news on attitude
- Analysis 2: Don't control for anxiety status:

```r
summary(lm(attitude ~ news))$coef |> round(4)
```

```
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   0.5272     0.0279 18.8638        0
news          0.3889     0.0395  9.8400        0
```

- Great! Now, say, subtract coef Analysis 1 from this, to get $T \to M \to Y$?
- Problem: This doesn't work.

Post-Treatment Bias

If you want the effect of T on Y,

- Match/adjust for (pre-treatment) covariates
- Don't match/adjust for other (post-treatment) quantities

Post-Treatment Bias

If you want the effect of T on Y,

- Match/adjust for (pre-treatment) covariates
- Don't match/adjust for other (post-treatment) quantities

If I'm not sure about the causal ordering, can I just try both?

Post-Treatment Bias

If you want the effect of T on Y,

- Match/adjust for (pre-treatment) covariates
- Don't match/adjust for other (post-treatment) quantities

If I'm not sure about the causal ordering, can I just try both?

Idea:

- Estimate $Y_i = \beta_0 + \beta_1 T_i + \epsilon_i$

Post-Treatment Bias

If you want the effect of T on Y,

- Match/adjust for (pre-treatment) covariates
- Don't match/adjust for other (post-treatment) quantities

If I'm not sure about the causal ordering, can I just try both?

Idea:

- Estimate $Y_i = \beta_0 + \beta_1 T_i + \epsilon_i$
- Estimate $Y_i = \delta_0 + \delta_1 T_i + \delta_2 X_i + \nu_i$

Post-Treatment Bias

If you want the effect of T on Y,

- Match/adjust for (pre-treatment) covariates
- Don't match/adjust for other (post-treatment) quantities

If I'm not sure about the causal ordering, can I just try both?

Idea:

- Estimate $Y_i = \beta_0 + \beta_1 T_i + \epsilon_i$
- Estimate $Y_i = \delta_0 + \delta_1 T_i + \delta_2 X_i + \nu_i$

Post-Treatment Bias

If you want the effect of T on Y,

- Match/adjust for (pre-treatment) covariates
- Don't match/adjust for other (post-treatment) quantities

If I'm not sure about the causal ordering, can I just try both?

Idea:

- Estimate $Y_i = \beta_0 + \beta_1 T_i + \epsilon_i$
- Estimate $Y_i = \delta_0 + \delta_1 T_i + \delta_2 X_i + \nu_i$

Problem:

- These 2 estimates of TE **don't** bound the truth.

Post-Treatment Bias

If you want the effect of T on Y,

- Match/adjust for (pre-treatment) covariates
- Don't match/adjust for other (post-treatment) quantities

If I'm not sure about the causal ordering, can I just try both?

Idea:

- Estimate $Y_i = \beta_0 + \beta_1 T_i + \epsilon_i$
- Estimate $Y_i = \delta_0 + \delta_1 T_i + \delta_2 X_i + \nu_i$

Problem:

- These 2 estimates of TE **don't** bound the truth.
- $ATE \overset{?}{\in} [\hat{\beta}_1, \hat{\delta}_1]$

Post-Treatment Bias

If you want the effect of T on Y,

- Match/adjust for (pre-treatment) covariates
- Don't match/adjust for other (post-treatment) quantities

If I'm not sure about the causal ordering, can I just try both?

Idea:

- Estimate $Y_i = \beta_0 + \beta_1 T_i + \epsilon_i$
- Estimate $Y_i = \delta_0 + \delta_1 T_i + \delta_2 X_i + \nu_i$

Problem:

- These 2 estimates of TE **don't** bound the truth.
- $ATE \overset{?}{\in} [\hat{\beta}_1, \hat{\delta}_1]$

Post-Treatment Bias

If you want the effect of T on Y,

- Match/adjust for (pre-treatment) covariates
- Don't match/adjust for other (post-treatment) quantities

If I'm not sure about the causal ordering, can I just try both?

Idea:

- Estimate $Y_i = \beta_0 + \beta_1 T_i + \epsilon_i$
- Estimate $Y_i = \delta_0 + \delta_1 T_i + \delta_2 X_i + \nu_i$

Problem:

- These 2 estimates of TE **don't** bound the truth.
- $ATE \overset{?}{\in} [\hat{\beta}_1, \hat{\delta}_1]$ We don't know!

Mediation

Mediation analysis tries to estimate <u>how much</u> effect of T on Y goes through M.

Notation

- $M_i(t)$: value of the mediator (function of treatment)

Notation

- $M_i(t)$: value of the mediator (function of treatment)
- $Y_i(t, m)$: potential outcome under some combination of t, m

Notation

- $M_i(t)$: value of the mediator (function of treatment)
- $Y_i(t, m)$: potential outcome under some combination of t, m
- $Y_i(T_i, M_i(T_i))$: potential outcome under
  - $T_i = t$
  - M_i you would get with $T_i = t$

Notation

- $M_i(t)$: value of the mediator (function of treatment)
- $Y_i(t, m)$: potential outcome under some combination of t, m
- $Y_i(T_i, M_i(T_i))$: potential outcome under
  - $T_i = t$
  - M_i you would get with $T_i = t$
- Quiz:

Notation

- $M_i(t)$: value of the mediator (function of treatment)
- $Y_i(t, m)$: potential outcome under some combination of t, m
- $Y_i(T_i, M_i(T_i))$: potential outcome under
  - $T_i = t$
  - M_i you would get with $T_i = t$
- Quiz: In news/anxiety/attitude example,
  - what's $Y_i(1, M_i(1))$?
  - what's $Y_i(0, M_i(0))$?
  - what's $Y_i(1, M_i(1)) - Y_i(0, M_i(0))$?
  - what's $Y_i(1, M_i(0))$?

Notation: Causal Effects

For individuals:

- $Y_i(1, M_i(1)) - Y_i(0, M_i(0))$: Total effect

Notation: Causal Effects

For individuals:

- $Y_i(1, M_i(1)) - Y_i(0, M_i(0))$: Total effect
- $Y_i(0, M_i(1)) - Y_i(0, M_i(0)) \equiv \delta_i(0)$: Indirect/Mediation effect under Co
- $Y_i(1, M_i(1)) - Y_i(1, M_i(0)) \equiv \delta_i(1)$: Indirect/Mediation effect under Tr

Notation: Causal Effects

For individuals:

- $Y_i(1, M_i(1)) - Y_i(0, M_i(0))$: Total effect
- $Y_i(0, M_i(1)) - Y_i(0, M_i(0)) \equiv \delta_i(0)$: Indirect/Mediation effect under Co
- $Y_i(1, M_i(1)) - Y_i(1, M_i(0)) \equiv \delta_i(1)$: Indirect/Mediation effect under Tr
- ACMEs: $\bar{\delta}(1)$ and $\bar{\delta}(0)$

Notation: Causal Effects

For individuals:

- $Y_i(1, M_i(1)) - Y_i(0, M_i(0))$: Total effect
- $Y_i(0, M_i(1)) - Y_i(0, M_i(0)) \equiv \delta_i(0)$: Indirect/Mediation effect under Co
- $Y_i(1, M_i(1)) - Y_i(1, M_i(0)) \equiv \delta_i(1)$: Indirect/Mediation effect under Tr
- ACMEs: $\bar{\delta}(1)$ and $\bar{\delta}(0)$
- $Y_i(1, M_i(0)) - Y_i(0, M_i(0)) \equiv \zeta_i(0)$: Direct effect of Tr on Y, under mediator value as if control
- $Y_i(1, M_i(1)) - Y_i(0, M_i(1)) \equiv \zeta_i(1)$: Direct effect of Tr on Y, under mediator value as if treated

Notation: Causal Effects

For individuals:

- $Y_i(1, M_i(1)) - Y_i(0, M_i(0))$: Total effect
- $Y_i(0, M_i(1)) - Y_i(0, M_i(0)) \equiv \delta_i(0)$: Indirect/Mediation effect under Co
- $Y_i(1, M_i(1)) - Y_i(1, M_i(0)) \equiv \delta_i(1)$: Indirect/Mediation effect under Tr
- ACMEs: $\bar{\delta}(1)$ and $\bar{\delta}(0)$
- $Y_i(1, M_i(0)) - Y_i(0, M_i(0)) \equiv \zeta_i(0)$: Direct effect of Tr on Y, under mediator value as if control
- $Y_i(1, M_i(1)) - Y_i(0, M_i(1)) \equiv \zeta_i(1)$: Direct effect of Tr on Y, under mediator value as if treated
- ADEs: $\bar{\zeta}(1)$ and $\bar{\zeta}(0)$

Notation: Causal Effects

For individuals:

- $Y_i(1, M_i(1)) - Y_i(0, M_i(0))$: Total effect
- $Y_i(0, M_i(1)) - Y_i(0, M_i(0)) \equiv \delta_i(0)$: Indirect/Mediation effect under Co
- $Y_i(1, M_i(1)) - Y_i(1, M_i(0)) \equiv \delta_i(1)$: Indirect/Mediation effect under Tr
- ACMEs: $\bar{\delta}(1)$ and $\bar{\delta}(0)$
- $Y_i(1, M_i(0)) - Y_i(0, M_i(0)) \equiv \zeta_i(0)$: Direct effect of Tr on Y, under mediator value as if control
- $Y_i(1, M_i(1)) - Y_i(0, M_i(1)) \equiv \zeta_i(1)$: Direct effect of Tr on Y, under mediator value as if treated
- ADEs: $\bar{\zeta}(1)$ and $\bar{\zeta}(0)$

Moderation and Interaction

- Moderators and mediators are both “third variables”

Moderation and Interaction

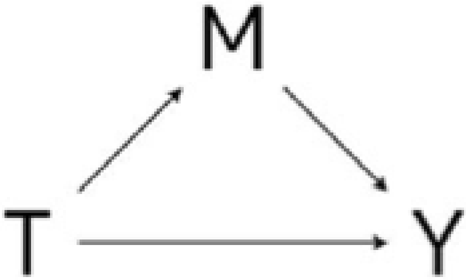
- Moderators and mediators are both “third variables”
- First, our DAG model for mediation:

Moderation and Interaction

- Moderators and mediators are both “third variables”
- First, our DAG model for mediation:

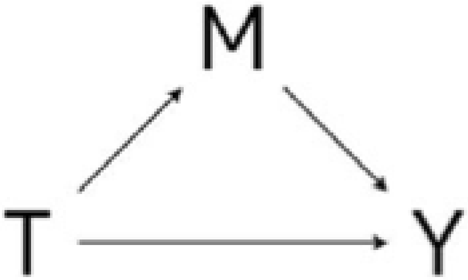
Moderation and Interaction

- Moderators and mediators are both “third variables”
- First, our DAG model for mediation:



Moderation and Interaction

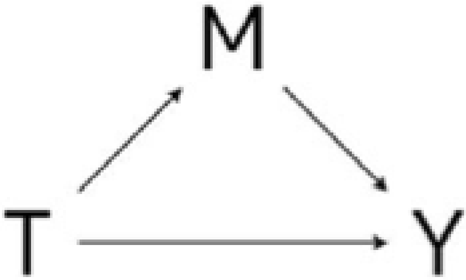
- Moderators and mediators are both "third variables"
- First, our DAG model for mediation:



- What's "moderation"?

Moderation and Interaction

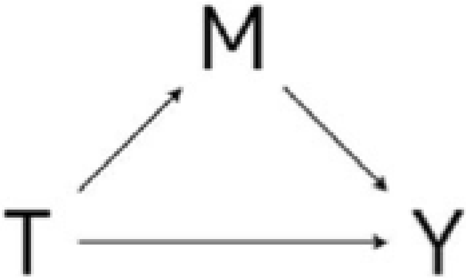
- Moderators and mediators are both “third variables”
- First, our DAG model for mediation:



- What’s “moderation”?
  - When $E(Y_i(1) - Y_i(0)|X = x_1) \neq E(Y_i(1) - Y_i(0)|X = x_2)$

Moderation and Interaction

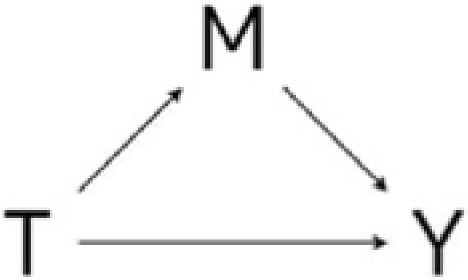
- Moderators and mediators are both "third variables"
- First, our DAG model for mediation:



- What's "moderation"?
  - When $E(Y_i(1) - Y_i(0)|X = x_1) \neq E(Y_i(1) - Y_i(0)|X = x_2)$
  - When there are "heterogeneous treatment effects"

Moderation and Interaction

- Moderators and mediators are both “third variables”
- First, our DAG model for mediation:



- What’s “moderation”?
  - When $E(Y_i(1) - Y_i(0)|X = x_1) \neq E(Y_i(1) - Y_i(0)|X = x_2)$
  - When there are “heterogeneous treatment effects”
  - When there is an “interaction between T and X”

Are 2 Experiments Enough for Mediation CEs?

- Exp. 1: Randomize T_i, measure M_i, get "ACE of T on M"

Are 2 Experiments Enough for Mediation CEs?

- Exp. 1: Randomize T_i, measure M_i, get "ACE of T on M"
- Exp. 2: Randomize M_i, measure Y_i, get "ACE of M on Y"

Are 2 Experiments Enough for Mediation CEs?

- Exp. 1: Randomize T_i, measure M_i, get "ACE of T on M"
- Exp. 2: Randomize M_i, measure Y_i, get "ACE of M on Y"
- Then, combine somehow, get ACME/Indir. effect of T on Y via M?

Are 2 Experiments Enough for Mediation CEs?

- Exp. 1: Randomize T_i, measure M_i, get "ACE of T on M"
- Exp. 2: Randomize M_i, measure Y_i, get "ACE of M on Y"
- Then, combine somehow, get ACME/Indir. effect of T on Y via M?
- But, this doesn't get you
  - Unbiased estimate

Are 2 Experiments Enough for Mediation CEs?

- Exp. 1: Randomize T_i, measure M_i, get "ACE of T on M"
- Exp. 2: Randomize M_i, measure Y_i, get "ACE of M on Y"
- Then, combine somehow, get ACME/Indir. effect of T on Y via M?
- But, this doesn't get you
  - Unbiased estimate
  - Sign of ACME

Are 2 Experiments Enough for Mediation CEs?

- Exp. 1: Randomize T_i, measure M_i, get "ACE of T on M"
- Exp. 2: Randomize M_i, measure Y_i, get "ACE of M on Y"
- Then, combine somehow, get ACME/Indir. effect of T on Y via M?
- But, this doesn't get you
  - Unbiased estimate
  - Sign of ACME
  - Informative bounds for ACME!

Usual (Best-Case?) Way to "Combine Somehow"

"Baron & Kenny Procedure"

$$M_i = \alpha_1 + aT_i + \epsilon_{i1} \tag{1}$$

$$Y_i = \alpha_2 + cT_i + \epsilon_{i2} \tag{2}$$

$$Y_i = \alpha_3 + dT_i + bM_i + \epsilon_{i3} \tag{3}$$

Usual (Best-Case?) Way to "Combine Somehow"

"Baron & Kenny Procedure"

$$
\begin{aligned}
M_i &= \alpha_1 + aT_i + \epsilon_{i1} && (1)\\
Y_i &= \alpha_2 + cT_i + \epsilon_{i2} && (2)\\
Y_i &= \alpha_3 + dT_i + bM_i + \epsilon_{i3} && (3)
\end{aligned}
$$

(Can add $+\mathbf{e}_1 X_i$, $+\mathbf{e}_2 X_i$, $+\mathbf{e}_3 X_i$.)

Usual (Best-Case?) Way to "Combine Somehow"

"Baron & Kenny Procedure"

$$
\begin{aligned}
M_i &= \alpha_1 + aT_i + \epsilon_{i1} && (1)\\
Y_i &= \alpha_2 + cT_i + \epsilon_{i2} && (2)\\
Y_i &= \alpha_3 + dT_i + bM_i + \epsilon_{i3} && (3)
\end{aligned}
$$

(Can add $+\mathbf{e}_1 X_i$, $+\mathbf{e}_2 X_i$, $+\mathbf{e}_3 X_i$.)

Then, call effect of

$$
\begin{aligned}
T \to M &= a & \\
T \to Y &= c & \text{(Total)}\\
T \to Y &= d & \text{(Direct)}\\
M \to Y &= b & \\
T \to M \to Y &= c - d = ab & \text{(Mediation)}
\end{aligned}
$$

Usual (Best-Case?) Way to "Combine Somehow"

"Baron & Kenny Procedure"

$$
\begin{aligned}
M_i &= \alpha_1 + aT_i + \epsilon_{i1} && (1)\\
Y_i &= \alpha_2 + cT_i + \epsilon_{i2} && (2)\\
Y_i &= \alpha_3 + dT_i + bM_i + \epsilon_{i3} && (3)
\end{aligned}
$$

(Can add $+\mathbf{e}_1 X_i$, $+\mathbf{e}_2 X_i$, $+\mathbf{e}_3 X_i$.)

Then, call effect of

$$
\begin{aligned}
T \to M &= a \\
T \to Y &= c && \text{(Total)}\\
T \to Y &= d && \text{(Direct)}\\
M \to Y &= b \\
T \to M \to Y &= c - d = ab && \text{(Mediation)}
\end{aligned}
$$

Problem: This doesn't work.

Why Aren't 2 Experiments Enough?

TABLE 1. The Fallacy of the Causal Chain Approach

Population Proportion	Potential Mediators and Outcomes $M_i(1)$	$M_i(0)$	$Y_i(t, 1)$	$Y_i(t, 0)$	Treatment Effect on Mediator $M_i(1) - M_i(0)$	Mediator Effect on Outcome $Y_i(t, 1) - Y_i(t, 0)$	Causal Mediation Effect $Y_i(t, M_i(1)) - Y_i(t, M_i(0))$
0.3	1	0	0	1	1	−1	−1
0.3	0	0	1	0	0	1	0
0.1	0	1	0	1	−1	−1	1
0.3	1	1	1	0	0	1	0
Average	0.6	0.4	0.6	0.4	0.2	0.2	−0.2

Notes: The left five columns of the table show a hypothetical population proportion of "types" of units defined by the values of potential mediators and outcomes. Note that these values can never be jointly observed. The last row of the table shows the population average value of each column. In this example, the average causal effect of the treatment on the mediator (the sixth column) is positive and equal to 0.2. Moreover, the average causal effect of the mediator on the outcome (the seventh column) is also positive and equals 0.2. And yet the average causal mediation effect (ACME; final column) is negative and equals −0.2.

Why Aren't 2 Experiments Enough?

TABLE 1. The Fallacy of the Causal Chain Approach

Population Proportion	Potential Mediators and Outcomes $M_i(1)$	$M_i(0)$	$Y_i(t, 1)$	$Y_i(t, 0)$	Treatment Effect on Mediator $M_i(1) - M_i(0)$	Mediator Effect on Outcome $Y_i(t, 1) - Y_i(t, 0)$	Causal Mediation Effect $Y_i(t, M_i(1)) - Y_i(t, M_i(0))$
0.3	1	0	0	1	1	−1	−1
0.3	0	0	1	0	0	1	0
0.1	0	1	0	1	−1	−1	1
0.3	1	1	1	0	0	1	0
Average	0.6	0.4	0.6	0.4	0.2	0.2	−0.2

Notes: The left five columns of the table show a hypothetical population proportion of "types" of units defined by the values of potential mediators and outcomes. Note that these values can never be jointly observed. The last row of the table shows the population average value of each column. In this example, the average causal effect of the treatment on the mediator (the sixth column) is positive and equal to 0.2. Moreover, the average causal effect of the mediator on the outcome (the seventh column) is also positive and equals 0.2. And yet the average causal mediation effect (ACME; final column) is negative and equals −0.2.

$$
\begin{aligned}
T \rightarrow M &= a = 0.2 \\
M \rightarrow Y &= b = 0.2 \\
T \rightarrow M \rightarrow Y &= ab = 0.04
\end{aligned}
$$

Why Aren't 2 Experiments Enough?

TABLE 1. The Fallacy of the Causal Chain Approach

Population Proportion	Potential Mediators and Outcomes $M_i(1)$	$M_i(0)$	$Y_i(t, 1)$	$Y_i(t, 0)$	Treatment Effect on Mediator $M_i(1) - M_i(0)$	Mediator Effect on Outcome $Y_i(t, 1) - Y_i(t, 0)$	Causal Mediation Effect $Y_i(t, M_i(1)) - Y_i(t, M_i(0))$
0.3	1	0	0	1	1	−1	−1
0.3	0	0	1	0	0	1	0
0.1	0	1	0	1	−1	−1	1
0.3	1	1	1	0	0	1	0
Average	0.6	0.4	0.6	0.4	0.2	0.2	−0.2

Notes: The left five columns of the table show a hypothetical population proportion of "types" of units defined by the values of potential mediators and outcomes. Note that these values can never be jointly observed. The last row of the table shows the population average value of each column. In this example, the average causal effect of the treatment on the mediator (the sixth column) is positive and equal to 0.2. Moreover, the average causal effect of the mediator on the outcome (the seventh column) is also positive and equals 0.2. And yet the average causal mediation effect (ACME; final column) is negative and equals −0.2.

$$
\begin{aligned}
T \rightarrow M &= a &= 0.2 \\
M \rightarrow Y &= b &= 0.2 \\
T \rightarrow M \rightarrow Y &= ab &= 0.04
\end{aligned}
$$

But, true $\bar{\delta}(t)$, ACME, $= -0.2$!

What Else Do You Need?

Consistency assumption: $T_i = t$, $M_i = m$ have same effect regardless of how they came to have those values.

What Else Do You Need?

Consistency assumption: $T_i = t$, $M_i = m$ have same effect regardless of how they came to have those values.

(Using lottery to estimate effect of income on attitude requires **lottery income** to have same effect as **regular income**.)

What Else Do You Need?

Consistency assumption: $T_i = t$, $M_i = m$ have same effect regardless of how they came to have those values.

(Using lottery to estimate effect of income on attitude requires **lottery income** to have same effect as **regular income**.)

The ACME, e.g., is an estimate of the effect of changes in M due to changing T (but without changing T).

What Else Do You Need?

Consistency assumption: $T_i = t$, $M_i = m$ have same effect regardless of how they came to have those values.

(Using lottery to estimate effect of income on attitude requires **lottery income** to have same effect as **regular income**.)

The ACME, e.g., is an estimate of the effect of changes in M due to changing T (but without changing T).

(Other manipulations of M rely on consistency.)

What Else Do You Need?

Big picture: to get more detailed estimates from same data, need more assumptions

Assumption 1 [Sequential Ignorability (Imai, Keele, and Yamamoto 2010)].

$$\{Y_i(t', m), M_i(t)\} \perp\!\!\!\perp T_i \mid X_i = x, \quad \textbf{(3)}$$

$$Y_i(t', m) \perp\!\!\!\perp M_i(t) \mid T_i = t, X_i = x, \quad \textbf{(4)}$$

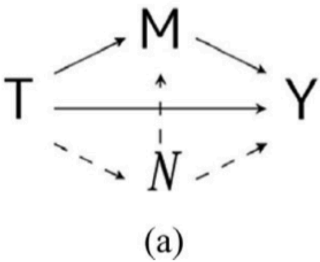
where $0 < \Pr(T_i = t \mid X_i = x)$ and $0 < p(M_i = m \mid T_i = t, X_i = x)$ for $t = 0, 1$, and all x and m in the support of X_i and M_i, respectively.

What Else Do You Need?

- Eqn 3: Conditional independence of PotOut's from Tr, given X (pretreatment!)
  - Ok, for random T, or balanced obs design. T as good as random, exog., etc.
  - (t' is just saying, for each $t = 0, 1$, must have Y's from both $t = 0, 1$ must be indep.)
- Eqn 4: Hard. Mediator is as good as random, given particular Tr status
- Problem: can't randomize <u>both</u> T and M in same experiment
  - (if want effect of T through M)
- You're getting 2 different QoI's if you randomize both: $T \to M, Y$ and $M \to Y$.
  - Showed can't combine those into $T \to M \to Y$

When Can You Get It?

FIGURE 8. Second Mediator Causing Serious Problem



Estimating Mediation Effects

- Subtlety: $\exists$ causal DAG's for which sensitivity of ACME can be est'd; even some with ACME identified.

Estimating Mediation Effects

- Subtlety: $\exists$ causal DAG's for which sensitivity of ACME can be est'd; even some with ACME identified.
- However, must convince that $\nexists\ N \rightarrow M$, whether or not N observed.

Estimating Mediation Effects

- Subtlety: $\exists$ causal DAG's for which sensitivity of ACME can be est'd; even some with ACME identified.
- However, must convince that $\nexists$ $N \to M$, whether or not N observed.
- Practical advice: start there.

Estimating Mediation Effects

- Subtlety: $\exists$ causal DAG's for which sensitivity of ACME can be est'd; even some with ACME identified.
- However, must convince that $\nexists\ N \rightarrow M$, whether or not N observed.
- Practical advice: start there. Then, formal mediation.

Sensitivity for Mediation Effects

Given

$$\begin{aligned} M_i &= \alpha_1 + aT_i + \epsilon_{i1} && (4) \\ Y_i &= \alpha_2 + cT_i + \epsilon_{i2} && (5) \\ Y_i &= \alpha_3 + dT_i + bM_i + \epsilon_{i3} && (6) \end{aligned}$$

- Q: How much covariance ρ is there between ϵ_{i1} and ϵ_{i3}?

Sensitivity for Mediation Effects

Given

$$
\begin{aligned}
M_i &= \alpha_1 + aT_i + \epsilon_{i1} && (4) \\
Y_i &= \alpha_2 + cT_i + \epsilon_{i2} && (5) \\
Y_i &= \alpha_3 + dT_i + bM_i + \epsilon_{i3} && (6)
\end{aligned}
$$

- Q: How much covariance ρ is there between ϵ_{i1} and ϵ_{i3}?
- A: If Seq. Ig. is true, then none (X-adjustment does its job)

Sensitivity for Mediation Effects

Given

$$
\begin{aligned}
M_i &= \alpha_1 + aT_i + \epsilon_{i1} && (4)\\
Y_i &= \alpha_2 + cT_i + \epsilon_{i2} && (5)\\
Y_i &= \alpha_3 + dT_i + bM_i + \epsilon_{i3} && (6)
\end{aligned}
$$

- Q: How much covariance ρ is there between ϵ_{i1} and ϵ_{i3}?
- A: If Seq. Ig. is true, then none (X-adjustment does its job)

  (If P, then Q.)

Sensitivity for Mediation Effects

Given

$$
\begin{aligned}
M_i &= \alpha_1 + aT_i + \epsilon_{i1} && (4) \\
Y_i &= \alpha_2 + cT_i + \epsilon_{i2} && (5) \\
Y_i &= \alpha_3 + dT_i + bM_i + \epsilon_{i3} && (6)
\end{aligned}
$$

- Q: How much covariance ρ is there between ϵ_{i1} and ϵ_{i3}?
- A: If Seq. Ig. is true, then none (X-adjustment does its job)

  (If P, then Q.)
- If $\rho \neq 0$, then Seq. Ig. is false (likely hidden confounder)

Sensitivity for Mediation Effects

Given

$$\begin{aligned} M_i &= \alpha_1 + aT_i + \epsilon_{i1} && (4) \\ Y_i &= \alpha_2 + cT_i + \epsilon_{i2} && (5) \\ Y_i &= \alpha_3 + dT_i + bM_i + \epsilon_{i3} && (6) \end{aligned}$$

- Q: How much covariance ρ is there between ϵ_{i1} and ϵ_{i3}?
- A: If Seq. Ig. is true, then none (X-adjustment does its job)

  (If P, then Q.)
- If $\rho \neq 0$, then Seq. Ig. is false (likely hidden confounder)

  (If $\neg Q$, then $\neg P$.)

Sensitivity for Mediation Effects

Given

$$
\begin{aligned}
M_i &= \alpha_1 + aT_i + \epsilon_{i1} \qquad (4)\\
Y_i &= \alpha_2 + cT_i + \epsilon_{i2} \qquad (5)\\
Y_i &= \alpha_3 + dT_i + bM_i + \epsilon_{i3} \qquad (6)
\end{aligned}
$$

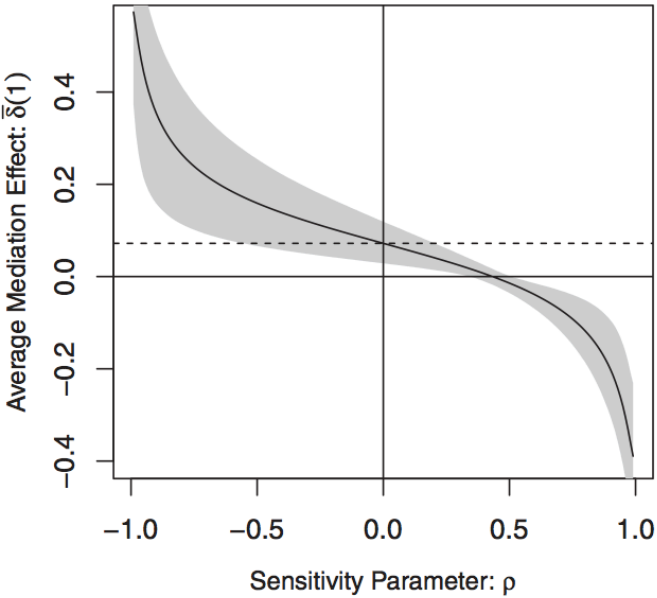
- Q: How much covariance ρ is there between ϵ_{i1} and ϵ_{i3}?
- A: If Seq. Ig. is true, then none (X-adjustment does its job)

  (If P, then Q.)
- If $\rho \neq 0$, then Seq. Ig. is false (likely hidden confounder)

  (If $\neg Q$, then $\neg P$.)

  (I.e., From freq. standpoint, you can find "evidence of problem", or "no evidence of problem", but not "evidence of no problem".)

Sensitivity for Mediation Effects



Summary

- Be careful.

Summary

- Be careful. If you estimate, you must do sensitivity.

Summary

- Be careful. If you estimate, you must do sensitivity.
  - A serious case of “don’t just get an answer”

Summary

- Be careful. If you estimate, you must do sensitivity.
  - A serious case of “don’t just get an answer”
  - (Do `plot(lm_out)`, too ...)

Summary

- Be careful. If you estimate, you must do sensitivity.
  - A serious case of "don't just get an answer"
  - (Do `plot(lm_out)`, too ...)
- Imai et al. (2011) thorough on assumptions, when trouble, when sensitivity is OK, when identification can be done

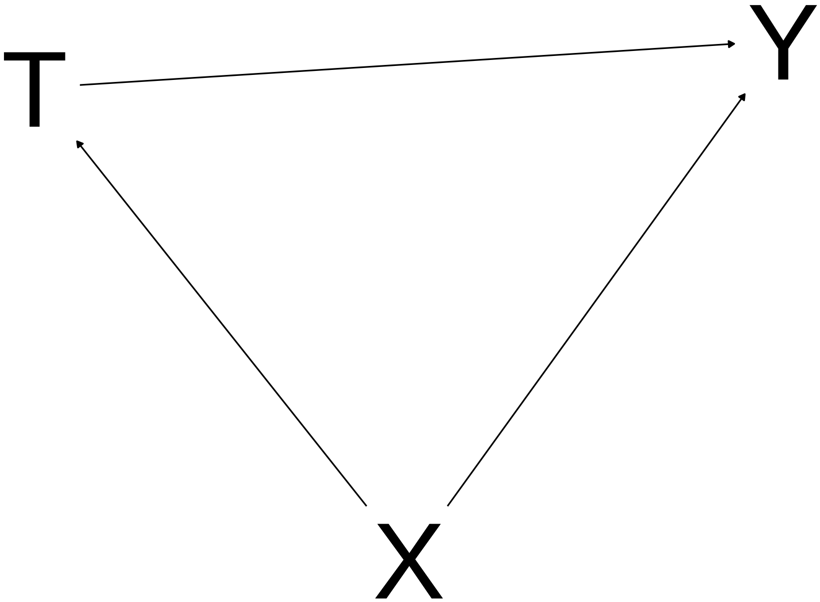
Summary

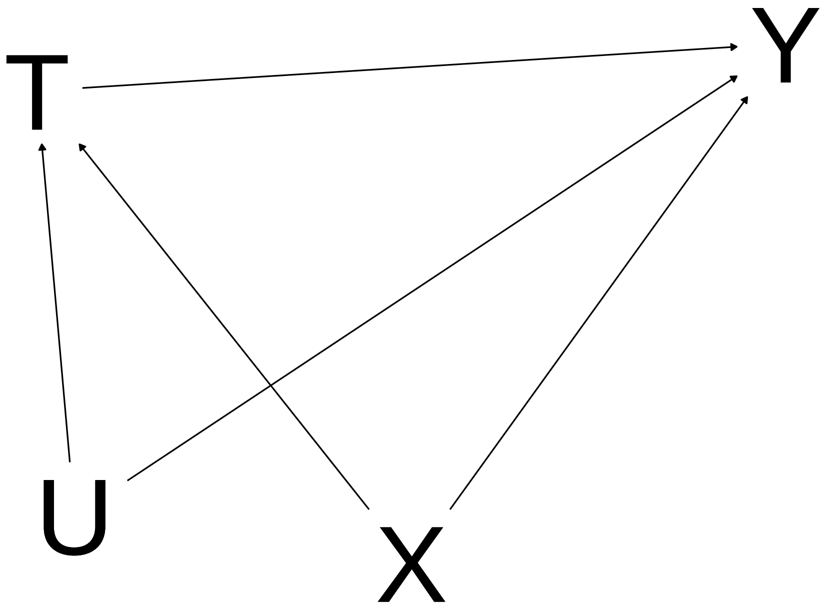
- Be careful. If you estimate, you must do sensitivity.
  - A serious case of "don't just get an answer"
  - (Do `plot(lm_out)`, too ...)
- Imai et al. (2011) thorough on assumptions, when trouble, when sensitivity is OK, when identification can be done
- From Bullock, Green, and Ha (2010):

> a cumulative enterprise. Persuasive conclusions about mediation are difficult to reach under any circumstances, but they are most likely to be reached when they derive from an experimental research program that addresses the particular challenges of mediation analysis—challenges that we describe here.

Sensitivity to an Unobserved Covariates

Confounding in Observational Studies




T
Y
U
X

Addressing Confounding

To break confounding,

- can't break $X \to Y$
- break $X \to T$
- I.e., make $X \perp\!\!\!\perp T$
- But this doesn't address $U \to T$ (or $U \to Y$).

Addressing Confounding

To break confounding,

- can't break $X \to Y$
- break $X \to T$
- I.e., make $X \perp\!\!\!\perp T$
- But this doesn't address $U \to T$ (or $U \to Y$).

(Of course, if no causal effect of $U \to Y$, no problem.)

Hidden Bias

Where there is $U \to T$ and $U \to Y$, there is <u>hidden bias</u>.

Hidden Bias

Where there is $U \to T$ and $U \to Y$, there is <u>hidden bias</u>.

Formally, i and j appear similar:

$$\mathbf{x}_i = \mathbf{x}_j$$

Hidden Bias

Where there is $U \to T$ and $U \to Y$, there is hidden bias.

Formally, i and j appear similar:

$$\mathbf{x}_i = \mathbf{x}_j$$

but are different in prop score:

$$\pi_i \neq \pi_j$$

Example

We are interested in the effect of phone calls on turnout.

Example

We are interested in the effect of phone calls on turnout.

Two voters look identical on observed predictors of whether called (that might affect turnout, too): age, education, income, party ID.

Example

We are interested in the effect of phone calls on turnout.

Two voters look identical on observed predictors of whether called (that might affect turnout, too): age, education, income, party ID.

However, **different** probabilities of being called, due to unobserved confounder, sociability.

Example

We are interested in the effect of phone calls on turnout.

Two voters look identical on observed predictors of whether called (that might affect turnout, too): age, education, income, party ID.

However, **different** probabilities of being called, due to unobserved confounder, sociability.

Sociability affects whether called (know more people) and turnout.

Example

We are interested in the effect of phone calls on turnout.

Two voters look identical on observed predictors of whether called (that might affect turnout, too): age, education, income, party ID.

However, **different** probabilities of being called, due to unobserved confounder, sociability.

Sociability affects whether called (know more people) and turnout.

Sensitivity: how strong must sociability be to invalidate inference about phone calls?

Odds

The odds of A_1 vs. A_2 is

$$A_1 : A_2 = \frac{p(A_1)}{p(A_2)}$$

Odds

The odds of A_1 vs. A_2 is

$$A_1 : A_2 = \frac{p(A_1)}{p(A_2)}$$

Odds often expressed as

- integers: $3 : 2$

Odds

The odds of A_1 vs. A_2 is

$$A_1 : A_2 = \frac{p(A_1)}{p(A_2)}$$

Odds often expressed as

- integers: $3 : 2$ Know $p(\Omega) = 1$, so

$$3 : 2 = \frac{.6}{.4}$$

Odds

The odds of A_1 vs. A_2 is

$$A_1 : A_2 = \frac{p(A_1)}{p(A_2)}$$

Odds often expressed as

- integers: $3:2$ Know $p(\Omega) = 1$, so

$$3:2 = \frac{.6}{.4}$$

- base $= 1$: $1.5 : 1$.

Odds

The odds of A_1 vs. A_2 is

$$A_1 : A_2 = \frac{p(A_1)}{p(A_2)}$$

Odds often expressed as

- integers: $3 : 2$ Know $p(\Omega) = 1$, so

$$3 : 2 = \frac{.6}{.4}$$

- base $= 1$: $1.5 : 1$. Know $p(\Omega) = 1$, so

$$1.5 : 1 = \frac{.6}{.4}$$

Odds Ratios

An <u>odds ratio</u> is

Odds Ratios

An odds ratio is a ratio of odds:

Odds Ratios

An odds ratio is a ratio of odds:

$$OR = \frac{\left(\frac{p(A_1)}{p(A_2)}\right)}{\left(\frac{p(A_3)}{p(A_4)}\right)}$$

Odds Ratios

An odds ratio is a ratio of odds:

$$OR = \frac{\left(\frac{p(A_1)}{p(A_2)}\right)}{\left(\frac{p(A_3)}{p(A_4)}\right)}$$

The strength, and weakness, is comparing changes from different base rates.

Odds Ratios

An odds ratio is a ratio of odds:

$$OR = \frac{\left(\frac{p(A_1)}{p(A_2)}\right)}{\left(\frac{p(A_3)}{p(A_4)}\right)}$$

The strength, and weakness, is comparing changes from different base rates.

From base odds of $1:1$, say a change of condition produces odds ratio of 3.

Odds Ratios

An odds ratio is a ratio of odds:

$$OR = \frac{\left(\frac{p(A_1)}{p(A_2)}\right)}{\left(\frac{p(A_3)}{p(A_4)}\right)}$$

The strength, and weakness, is comparing changes from different base rates.

From base odds of $1:1$, say a change of condition produces odds ratio of 3.

$$\frac{\frac{.03}{.01}}{\frac{.01}{.01}}$$

Odds Ratios

An odds ratio is a ratio of odds:

$$OR = \frac{\left(\frac{p(A_1)}{p(A_2)}\right)}{\left(\frac{p(A_3)}{p(A_4)}\right)}$$

The strength, and weakness, is comparing changes from different base rates.

From base odds of $1:1$, say a change of condition produces odds ratio of 3.

$$\frac{\frac{.03}{.01}}{\frac{.01}{.01}} = \frac{\frac{.06}{.02}}{\frac{.02}{.02}}$$

Odds Ratios

An odds ratio is a ratio of odds:

$$OR = \frac{\left(\frac{p(A_1)}{p(A_2)}\right)}{\left(\frac{p(A_3)}{p(A_4)}\right)}$$

The strength, and weakness, is comparing changes from different base rates.

From base odds of $1:1$, say a change of condition produces odds ratio of 3.

$$\frac{\frac{.03}{.01}}{\frac{.01}{.01}} = \frac{\frac{.06}{.02}}{\frac{.02}{.02}} = \frac{\frac{.9}{.3}}{\frac{.3}{.3}}$$

Odds Ratios

An <u>odds ratio</u> is a ratio of odds:

$$OR = \frac{\left(\frac{p(A_1)}{p(A_2)}\right)}{\left(\frac{p(A_3)}{p(A_4)}\right)}$$

The strength, and weakness, is comparing changes from different base rates.

From base odds of $1:1$, say a change of condition produces odds ratio of 3.

$$\frac{\frac{.03}{.01}}{\frac{.01}{.01}} = \frac{\frac{.06}{.02}}{\frac{.02}{.02}} = \frac{\frac{.9}{.3}}{\frac{.3}{.3}} = \frac{\frac{.9}{.3}}{\frac{.9}{.9}}$$

Odds Ratios

An odds ratio is a ratio of odds:

$$OR = \frac{\left(\frac{p(A_1)}{p(A_2)}\right)}{\left(\frac{p(A_3)}{p(A_4)}\right)}$$

The strength, and weakness, is comparing changes from different base rates.

From base odds of $1:1$, say a change of condition produces odds ratio of 3.

$$\frac{\frac{.03}{.01}}{\frac{.01}{.01}} = \frac{\frac{.06}{.02}}{\frac{.02}{.02}} = \frac{\frac{.9}{.3}}{\frac{.3}{.3}} = \frac{\frac{.9}{.3}}{\frac{.9}{.9}} = \ldots$$

Application: Measuring Group Differences (JP Scanlon)

	% Below Pov Line			% Above Pov Line		
	B	W	$\frac{B}{W}$	B	W	$\frac{B}{W}$
t_1	90	80	1.1	10	20	0.5

Application: Measuring Group Differences (JP Scanlon)

	% Below Pov Line			% Above Pov Line		
	B	W	$\frac{B}{W}$	B	W	$\frac{B}{W}$
t_1	90	80	1.1	10	20	0.5
t_2	15	5	3.0	85	95	0.89

Application: Measuring Group Differences (JP Scanlon)

	% Below Pov Line			% Above Pov Line		
	B	W	$\frac{B}{W}$	B	W	$\frac{B}{W}$
t_1	90	80	1.1	10	20	0.5
t_2	15	5	3.0	85	95	0.89

Application: Measuring Group Differences (JP Scanlon)

	% Below Pov Line			% Above Pov Line		
	B	W	$\frac{B}{W}$	B	W	$\frac{B}{W}$
t_1	90	80	1.1	10	20	0.5
t_2	15	5	3.0	85	95	0.89

- At t_1: More blacks below, whites above PovLine

Application: Measuring Group Differences (JP Scanlon)

	% Below Pov Line			% Above Pov Line		
	B	W	$\frac{B}{W}$	B	W	$\frac{B}{W}$
t_1	90	80	1.1	10	20	0.5
t_2	15	5	3.0	85	95	0.89

- At t_1: More blacks below, whites above PovLine
- At t_2: are things getting better or worse for Blacks relative to Whites?

Application: Measuring Group Differences (JP Scanlon)

	% Below Pov Line			% Above Pov Line		
	B	W	$\frac{B}{W}$	B	W	$\frac{B}{W}$
t_1	90	80	1.1	10	20	0.5
t_2	15	5	3.0	85	95	0.89

- At t_1: More blacks below, whites above PovLine
- At t_2: are things getting better or worse for Blacks relative to Whites?

Application: Measuring Group Differences (JP Scanlon)

	% Below Pov Line			% Above Pov Line		
	B	W	$\frac{B}{W}$	B	W	$\frac{B}{W}$
t_1	90	80	1.1	10	20	0.5
t_2	15	5	3.0	85	95	0.89

- At t_1: More blacks below, whites above PovLine
- At t_2: are things getting better or worse for Blacks relative to Whites?

Clearly, worse (odds of below pov line):

Odds Ratios: $\frac{1.1}{.5} = 2.2$, $\frac{3}{.89} = 3.4$

Application: Measuring Group Differences (JP Scanlon)

	% Below Pov Line			% Above Pov Line		
	B	W	$\frac{B}{W}$	B	W	$\frac{B}{W}$
t_1	90	80	1.1	10	20	0.5
t_2	15	5	3.0	85	95	0.89

- At t_1: More blacks below, whites above PovLine
- At t_2: are things getting better or worse for Blacks relative to Whites?

Clearly, worse (odds of below pov line):

Odds Ratios: $\frac{1.1}{.5} = 2.2$, $\frac{3}{.89} = 3.4$

Clearly, no change:

Absolute Differences: 10, 10, 10, 10

Application: Measuring Group Differences (JP Scanlon)

	% Below Pov Line			% Above Pov Line		
	B	W	$\frac{B}{W}$	B	W	$\frac{B}{W}$
t_1	90	80	1.1	10	20	0.5
t_2	15	5	3.0	85	95	0.89

- At t_1: More blacks below, whites above PovLine
- At t_2: are things getting better or worse for Blacks relative to Whites?

Clearly, worse (odds of below pov line):

Odds Ratios: $\frac{1.1}{.5} = 2.2$, $\frac{3}{.89} = 3.4$

Clearly, no change:

Absolute Differences: 10, 10, 10, 10

Clearly, huge absolute improvements.

Application: Measuring Group Differences (JP Scanlon)

- Key: it's not clear whether relative disparities getting better/worse/neither by below/above measures.

Application: Measuring Group Differences (JP Scanlon)

- Key: it's not clear whether relative disparities getting better/worse/neither by below/above measures.
- (Easy to produce examples of OR's same and AbsDiffs slightly diff.)

Application: Measuring Group Differences (JP Scanlon)

- Key: it's not clear whether relative disparities getting better/worse/neither by below/above measures.
- (Easy to produce examples of OR's same and AbsDiffs slightly diff.)
- (Diffs betwn groups real, importnt, but how we meas. changes is tricky)

King's Conjecture

17 October 2012

Modeling Hidden Bias

Odds of treatment for i and j:

$$\frac{\pi_i}{1-\pi_i}, \frac{\pi_j}{1-\pi_j}$$

Modeling Hidden Bias

Odds of treatment for i and j:

$$\frac{\pi_i}{1 - \pi_i}, \frac{\pi_j}{1 - \pi_j}$$

OR of i versus j:

$$\begin{aligned} OR &= \frac{\pi_i}{1 - \pi_i} \div \frac{\pi_j}{1 - \pi_j} \\ &= \frac{\pi_i(1 - \pi_j)}{\pi_j(1 - \pi_i)} \end{aligned}$$

Modeling Hidden Bias

Let Γ be upper bound on OR of treatment.

$$\frac{1}{\Gamma} \leq \frac{\pi_i(1-\pi_j)}{\pi_j(1-\pi_i)} \leq \Gamma \qquad \forall i, j \text{ s.t. } \mathbf{x}_i = \mathbf{x}_j$$

Modeling Hidden Bias

Let Γ be upper bound on OR of treatment.

$$\frac{1}{\Gamma} \leq \frac{\pi_i(1-\pi_j)}{\pi_j(1-\pi_i)} \leq \Gamma \qquad \forall i, j \text{ s.t. } \mathbf{x}_i = \mathbf{x}_j$$

By what factor does the odds of treatment differ? (No more than Γ)

Modeling Hidden Bias

Rosenbaum (2020) shows that this is same as

$$\begin{aligned}\log\left(\frac{\pi_i}{1-\pi_i}\right) &= \kappa(\mathbf{x}_i) + \gamma u_i \\ \log\left(\frac{\pi_j}{1-\pi_j}\right) &= \kappa(\mathbf{x}_j) + \gamma u_j\end{aligned}$$

s.t. $0 \leq u_i \leq 1$.

Modeling Hidden Bias

Rosenbaum (2020) shows that this is same as

$$
\begin{aligned}
\log\left(\frac{\pi_i}{1-\pi_i}\right) &= \kappa(\mathbf{x}_i) + \gamma u_i \\
\log\left(\frac{\pi_j}{1-\pi_j}\right) &= \kappa(\mathbf{x}_j) + \gamma u_j
\end{aligned}
$$

s.t. $0 \leq u_i \leq 1$.

Interpretation: first rewrite

$$
\log\left(\frac{\pi_j}{1-\pi_j}\right) = \kappa(\mathbf{x}_i) + \gamma u_j
$$

Exponentiate:

$$\begin{aligned}\left(\frac{\pi_i}{1-\pi_i}\right) &= e^{\boldsymbol{\kappa}(\mathbf{x}_i)+\gamma u_i} \\ \left(\frac{\pi_j}{1-\pi_j}\right) &= e^{\boldsymbol{\kappa}(\mathbf{x}_i)+\gamma u_j}\end{aligned}$$

Exponentiate:

$$\begin{aligned}
\left(\frac{\pi_i}{1-\pi_i}\right) &= e^{\kappa(\mathbf{x}_i)+\gamma u_i} \\
\left(\frac{\pi_j}{1-\pi_j}\right) &= e^{\kappa(\mathbf{x}_i)+\gamma u_j}
\end{aligned}$$

Calculate OR:

$$\begin{aligned}
OR &= \frac{\pi_i(1-\pi_j)}{\pi_j(1-\pi_i)} \\
&= \frac{e^{\kappa(\mathbf{x}_i)+\gamma u_i}}{e^{\kappa(\mathbf{x}_i)+\gamma u_j}} \\
&= e^{(\kappa(\mathbf{x}_i)+\gamma u_i)-(\kappa(\mathbf{x}_i)+\gamma u_j)} \\
&= e^{(\gamma u_i-\gamma u_j)} \\
&= e^{\gamma(u_i-u_j)}
\end{aligned}$$

Interpreting Γ

$$OR = e^{\gamma(u_i - u_j)}$$

Interpreting Γ

$OR = e^{\gamma(u_i - u_j)}$

Log odds differ by factor of γ times diff in unobs confounder.

Interpreting Γ

$OR = e^{\gamma(u_i - u_j)}$

Log odds differ by factor of γ times diff in unobs confounder.

Shows $\Gamma = e^{\gamma}$.

TABLE 4.1. Sensitivity Analysis for Hammond's Study of Smoking and Lung Cancer: Range of Significance Levels for Hidden Biases of Various Magnitudes.

Γ	Minimum	Maximum
1	< 0.0001	< 0.0001
2	< 0.0001	< 0.0001
3	< 0.0001	< 0.0001
4	< 0.0001	0.0036
5	< 0.0001	0.03
6	< 0.0001	0.1

TABLE 4.1. Sensitivity Analysis for Hammond's Study of Smoking and Lung Cancer: Range of Significance Levels for Hidden Biases of Various Magnitudes.

Γ	Minimum	Maximum
1	< 0.0001	< 0.0001
2	< 0.0001	< 0.0001
3	< 0.0001	< 0.0001
4	< 0.0001	0.0036
5	< 0.0001	0.03
6	< 0.0001	0.1

- Groups: smokers/nonsmokers
- Outcome: lung cancer
- Something must increase smoking by $6\times$ to change inference.
- If exists, maybe it's that factor, not smoking directly.

(Bias from $U \to T$; effectively, $U \to Y$ nearly perfect.)

Γ	Minimum	Maximum
1	≤ 0.0001	≤ 0.0001
2	≤ 0.0001	0.0018
3	≤ 0.0001	0.0136
4	≤ 0.0001	0.0388
4.25	≤ 0.0001	0.0468
5	≤ 0.0001	0.0740

Table 4.2: Signed-Rank Statistic p-value Sensitivity for Lead in Children's Blood

- Groups: parents occupationally exposed/unexposed
- Outcome: children's levels
- Something must increase parents' exposure by $5\times$ to change inference.
- If exists, maybe it's that, not parental exposure directly.

Γ	Minimum	Maximum
1	≤ 0.0001	≤ 0.0001
2	≤ 0.0001	0.0018
3	≤ 0.0001	0.0136
4	≤ 0.0001	0.0388
4.25	≤ 0.0001	0.0468
5	≤ 0.0001	0.0740

Table 4.2: Signed-Rank Statistic p-value Sensitivity for Lead in Children's Blood

- Groups: parents occupationally exposed/unexposed
- Outcome: children's levels
- Something must increase parents' exposure by $5\times$ to change inference.
- If exists, maybe it's that, not parental exposure directly.

(one-sided)

Γ	Minimum	Maximum
1	15	15
2	10.25	19.5
3	8	23
4	6.5	25
5	5	26.5

Table 4.3: Point Estimate Sensitivity for Lead in Children's Blood

Γ	Minimum	Maximum
1	15	15
2	10.25	19.5
3	8	23
4	6.5	25
5	5	26.5

Table 4.3: Point Estimate Sensitivity for Lead in Children's Blood

- HL point estimate: 15 (median of all $m \times n$ possible matched pairs)
- With confounding, wider range of possible effects.

Γ	95% CI
1	(9.5, 20.5)
2	(4.5, 27.5)
3	(1.0, 32.0)
4	(-1.0, 36.5)
5	(-3.0, 41.5)

Table 4.4: Confidence Interval Sensitivity for Lead in Children's Blood

Γ	95% CI
1	(9.5, 20.5)
2	(4.5, 27.5)
3	(1.0, 32.0)
4	(-1.0, 36.5)
5	(-3.0, 41.5)

Table 4.4: Confidence Interval Sensitivity for Lead in Children's Blood

- Inverted NHST CI's
- If something increases parental exposure by $4\times$, negative estimates of parents on children are reasonable.

(two-sided)

Implementation

Packages

- Frank et al. (2013): `konfound`
- Keele (2022): `rbounds`
- `sensitivitymw`
- `sensitivitymv`

Example

```r
anes <- read_csv("../data/anes_pilot_2016.csv")
dim(anes)
```

```
[1] 1200  594
```

```r
anes <- anes |> mutate(age = 2016 - birthyr,
                       pid_rep = as.numeric(pid3 == 3),
                       pid_dem = as.numeric(pid3 == 1))
```

```r
lm_out <- lm(turnout12 ~ pid_rep, data = anes)
summary(lm_out)
```

```

Call:
lm(formula = turnout12 ~ pid_rep, data = anes)

Residuals:
    Min      1Q  Median      3Q     Max
-0.3395 -0.2451 -0.2451 -0.2451  1.7549

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.24512    0.01868  66.641  < 2e-16 ***
pid_rep      0.09435    0.03320   2.842  0.00456 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '

Residual standard error: 0.535 on 1198 degrees of freedom
```

```r
library(konfound)
konfound(lm_out, pid_rep)
```

```r
library(konfound)
konfound(lm_out, pid_rep)
```

```
Robustness of Inference to Replacement (RIR):
To invalidate an inference, 30.959 % of the estimate would
have to be due to bias.
This is based on a threshold of 0.065 for statistical
significance (alpha = 0.05).

To invalidate an inference, 372 observations would have to
be replaced with cases
for which the effect is 0 (RIR = 372).

See Frank et al. (2013) for a description of the method.

Citation: Frank, K.A., Maroulis, S., Duong, M., and Kelcey,
B. (2013).
What would it take to change an inference?
Using Rubin's causal model to interpret the
robustness of causal inferences
```

```r
lm_out <- lm(turnout12 ~ pid_rep + age, data = anes)
summary(lm_out)
```

```

Call:
lm(formula = turnout12 ~ pid_rep + age, data = anes)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.5825 -0.3388 -0.1711  0.0301  1.9831 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept)  1.678649   0.045960  36.524  < 2e-16 ***
pid_rep      0.082685   0.031870   2.594  0.00959 ** 
age         -0.008943   0.000873 -10.244  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '

Residual standard error: 0.5132 on 1197 degrees of freedom
```

```r
konfound(lm_out, pid_rep)
```

```
Robustness of Inference to Replacement (RIR):
To invalidate an inference, 24.379 % of the estimate would
have to be due to bias.
This is based on a threshold of 0.063 for statistical
significance (alpha = 0.05).

To invalidate an inference, 293 observations would have to
be replaced with cases
for which the effect is 0 (RIR = 293).

See Frank et al. (2013) for a description of the method.

Citation: Frank, K.A., Maroulis, S., Duong, M., and Kelcey,
B. (2013).
What would it take to change an inference?
Using Rubin's causal model to interpret the
robustness of causal inferences.
Education, Evaluation and
```

```r
cor(anes[,c("pid_rep", "turnout12", "econnow")])
```

```
            pid_rep   turnout12     econnow
pid_rep   1.00000000 0.081825966 0.141257803
turnout12 0.08182597 1.000000000 0.008599061
econnow   0.14125780 0.008599061 1.000000000
```

```r
lm_out <- lm(turnout12 ~ pid_rep + age + econnow, data = ar
summary(lm_out)
```

```
Call:
lm(formula = turnout12 ~ pid_rep + age + econnow, data = ar

Residuals:
     Min       1Q   Median       3Q      Max
-0.60257 -0.33748 -0.17138  0.04458  1.96702

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.6290966  0.0565381  28.814   <2e-16 ***
pid_rep      0.0755031  0.0322095   2.344   0.0192 *
age         -0.0091496  0.0008833 -10.358   <2e-16 ***
econnow      0.0202398  0.0134633   1.503   0.1330
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '
```

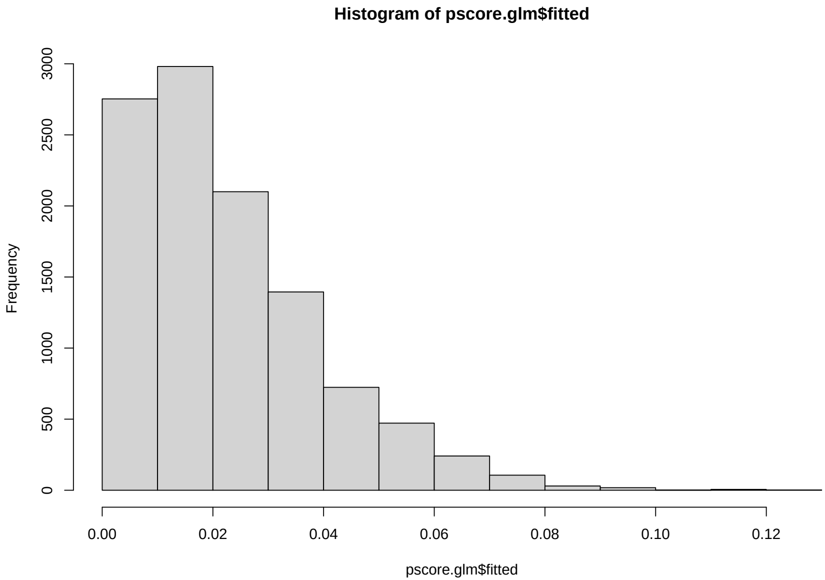
```r
konfound(lm_out, pid_rep)
```

```
Robustness of Inference to Replacement (RIR):
To invalidate an inference, 16.303 % of the estimate would
have to be due to bias.
This is based on a threshold of 0.063 for statistical
significance (alpha = 0.05).

To invalidate an inference, 196 observations would have to
be replaced with cases
for which the effect is 0 (RIR = 196).

See Frank et al. (2013) for a description of the method.

Citation: Frank, K.A., Maroulis, S., Duong, M., and Kelcey,
B. (2013).
What would it take to change an inference?
Using Rubin's causal model to interpret the
robustness of causal inferences.
Education, Evaluation and
```

Implementation in `rbounds`

```r
library(Matching)
data(GerberGreenImai)

# Estimate Propensity Score
pscore.glm <- glm(PHN.C1 ~ PERSONS + VOTE96.1 +
                    NEW + MAJORPTY + AGE + WARD +
                    PERSONS:VOTE96.1 + PERSONS:NEW +
                    AGE2, family = binomial(logit),
                  data = GerberGreenImai)
```

```r
hist(pscore.glm$fitted)
```



Implementation in rbounds

```
# Match - without replacement
m.obj <- Match(Y = GerberGreenImai$VOTED98,
               Tr = GerberGreenImai$PHN.C1,
               X = fitted(pscore.glm), M = 1, replace = FA

summary(m.obj)
```

```
Estimate...  0.036437
SE.........  0.041421
T-stat.....  0.87968
p.val......  0.37903

Original number of observations..............  10829
Original number of treated obs...............  247
Matched number of observations...............  247
Matched number of observations  (unweighted).  247
```

Implementation in `rbounds`

```r
library(rbounds)

# Sensitivity Test
# binarysens(m.obj, Gamma = 2, GammaInc = .1)
```

Implementation in `rbounds`

```
#hlsens(m.obj, Gamma = 5, GammaInc = 1)
```

Thanks!

rtm@american.edu
www.ryantmoore.org

References I

Blackwell, Matthew, and Anton Strezhnev. 2022. "Telescope Matching for Reducing Model Dependence in the Estimation of the Effects of Time-Varying Treatments: An Application to Negative Advertising." Journal of the Royal Statistical Society, Series A 185 (1): 377–99. https://doi.org/10.1111/rssa.12759.

Bullock, John G., Donald P. Green, and Shang E. Ha. 2010. "Yes, but What's the Mechanism? (Don't Expect an Easy Answer)." Journal of Personality and Social Psychology 98 (4): 550–58.

Frank, Kenneth A., Spiro J. Maroulis, Minh Q. Duong, and Benjamin M. Kelcey. 2013. "What Would It Take to Change an Inference? Using Rubin's Causal Model to Interpret the Robustness of Causal Inferences." Educational Evaluation and Policy Analysis 35 (4): 437–60.

Hebbali, Aravind. 2024. olsrr: Tools for Building OLS Regression Models. https://CRAN.R-project.org/package=olsrr.

Ho, Daniel, Kosuke Imai, Gary King, and Elizabeth Stuart. 2007. "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." Political Analysis 15: 199–236.

References II

Imai, Kosuke, Luke Keele, Dustin Tingley, and Teppei Yamamoto. 2011. "Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies." American Political Science Review 105 (4): 765–89.

Keele, Luke J. 2022. rbounds: Perform Rosenbaum Bounds Sensitivity Tests for Matched and Unmatched Data. https://CRAN.R-project.org/package=rbounds.

Moore, Ryan T., Eleanor Neff Powell, and Andrew Reeves. 2013. "Driving Support: Workers, PACs, and Congressional Support of the Auto Industry." Business and Politics 15 (2): 137–62.

Rosenbaum, Paul. 2020. Design of Observational Studies. Second. New York, NY: Springer.