Telescope Matching: A Flexible Approach to Estimating Direct Effects

Matthew Blackwell and Anton Strezhnev

International Methods Colloquium

October 12, 2018

direct effect

direct effect

effect of treatment not due to a particular downstream cause

direct effect

effect of treatment not due to a particular downstream cause

why do we care?

direct effect

effect of treatment not due to a particular downstream cause

why do we care?

causal mediation

direct effect

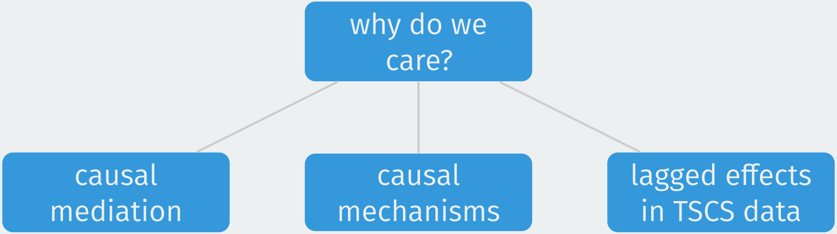
effect of treatment not due to a particular downstream cause

why do we care?

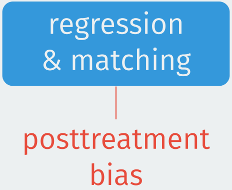
- causal mediation
- causal mechanisms

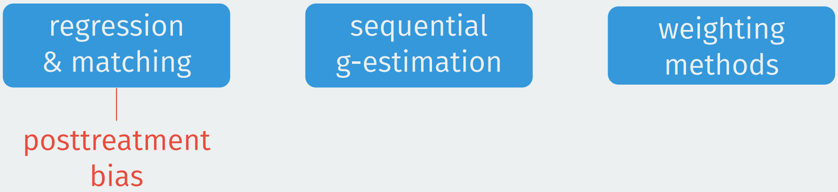
direct effect

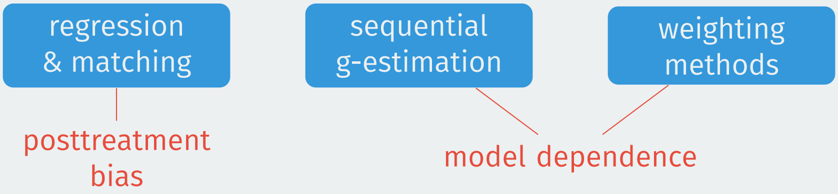
effect of treatment not due to a particular downstream cause

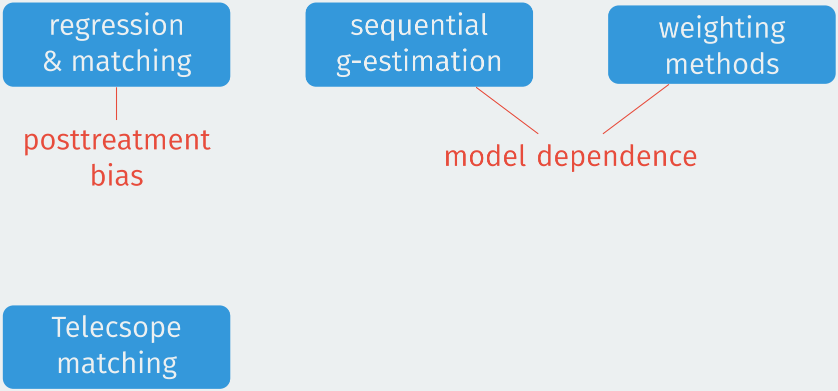


regression
& matching


regression
& matching
posttreatment
bias


regression & matching
sequential g-estimation
weighting methods
posttreatment bias


regression & matching
sequential g-estimation
weighting methods
posttreatment bias
model dependence


regression
& matching
posttreatment
bias
sequential
g-estimation
weighting
methods
model dependence
Telecsope
matching

regression & matching

posttreatment bias

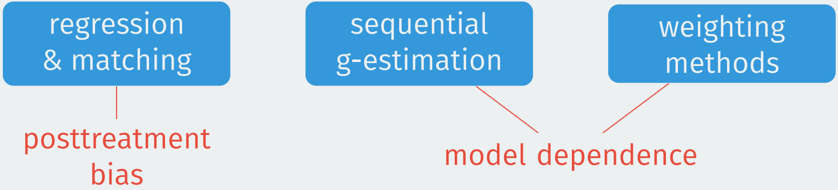
sequential g-estimation

weighting methods

model dependence

Telescope matching

- consistent for direct effects



Telescope matching

- consistent for direct effects
- avoids post-treatment bias

regression & matching

posttreatment bias

sequential g-estimation

weighting methods

model dependence

Telescope matching

- consistent for direct effects
- avoids post-treatment bias
- robust to (some) model misspecification

regression & matching

posttreatment bias

sequential g-estimation

weighting methods

model dependence

Telecsope matching

- consistent for direct effects
- avoids post-treatment bias
- robust to (some) model misspecification
- carries over logic from standard matching

Roadmap

1. The difficulty of direct effects
2. Our approach: telescope matching
3. Simulating misspecification
4. Application
5. Conclusion

1/ The difficulty of direct effects

Notation

Setting Effect of frame on immigration media accounts

Notation

Setting	Effect of frame on immigration media accounts
A_i	Binary treatment $\in \{\texttt{negative frame, positive frame}\}$

Notation

Setting	Effect of frame on immigration media accounts
A_i	Binary treatment $\in$ {`negative frame`, `positive frame`}
M_i	Binary mediator $\in$ {`high anxiety`, `low anxiety`}

Notation

Setting	Effect of frame on immigration media accounts
A_i	Binary treatment $\in \{\texttt{negative frame}, \texttt{positive frame}\}$
M_i	Binary mediator $\in \{\texttt{high anxiety}, \texttt{low anxiety}\}$
Y_i	Outcome (support for immigration)

Notation

Setting	Effect of frame on immigration media accounts
A_i	Binary treatment $\in \{\texttt{negative frame, positive frame}\}$
M_i	Binary mediator $\in \{\texttt{high anxiety, low anxiety}\}$
Y_i	Outcome (support for immigration)
$Y_i(a, m)$	Potential outcome

The Quantity of Interest

Definition (Average Controlled Direct Effect)

$$\tau(m) = E[Y_i(1, m) - Y_i(0, m)]$$

The Quantity of Interest

Definition (Average Controlled Direct Effect)

$$\tau(m) = E[Y_i(1, m) - Y_i(0, m)]$$

- Average effect of manipulating A_i while fixing M_i to level m

The Quantity of Interest

Definition (Average Controlled Direct Effect)

$$\tau(m) = E[Y_i(1, m) - Y_i(0, m)]$$

- Average effect of manipulating A_i while fixing M_i to level m
- Easily identified if A_i and M_i are randomized but...

The Quantity of Interest

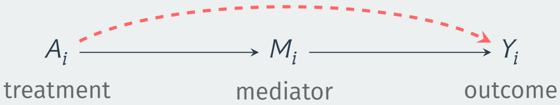
Definition (Average Controlled Direct Effect)

$$\tau(m) = E[Y_i(1, m) - Y_i(0, m)]$$

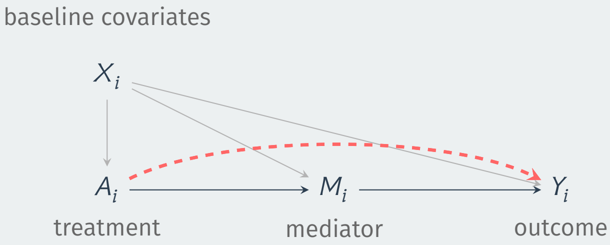
- Average effect of manipulating A_i while fixing M_i to level m
- Easily identified if A_i and M_i are randomized but...
- Lots of studies are observational in M_i or both.

Confounders

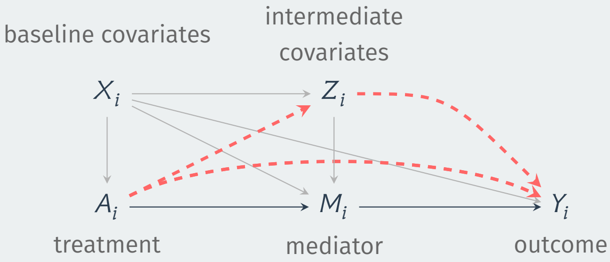
Confounders



Confounders



Confounders



Assumptions

Assumptions

Assumption (Sequential Ignorability)

$$\{Y_i(a,m), M_i(a), Z_i(a)\} \perp\!\!\!\perp A_i | X_i = x$$
$$Y_i(a,m) \perp\!\!\!\perp M_i | A_i = a, X_i = x, Z_i = z$$

No omitted variables for A_i given X_i.
No omitted variable for M_i given A_i, X_i, Z_i.

Assumptions

Assumption (Sequential Ignorability)

$$\{Y_i(a,m), M_i(a), Z_i(a)\} \perp\!\!\!\perp A_i | X_i = x$$
$$Y_i(a,m) \perp\!\!\!\perp M_i | A_i = a, X_i = x, Z_i = z$$

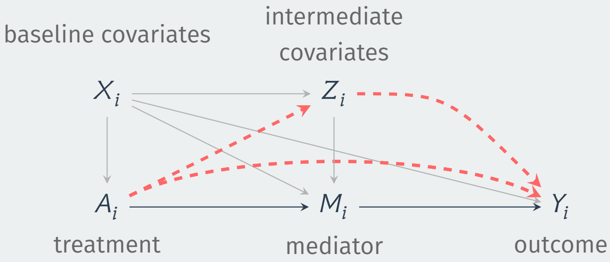
No omitted variables for A_i given X_i.
No omitted variable for M_i given A_i, X_i, Z_i.

Assumption (Positivity)

$$0 < P(A_i = 1 | X_i = x) < 1$$
$$0 < P(M_i = 1 | X_i = x, Z_i = z, A_i = a) < 1$$

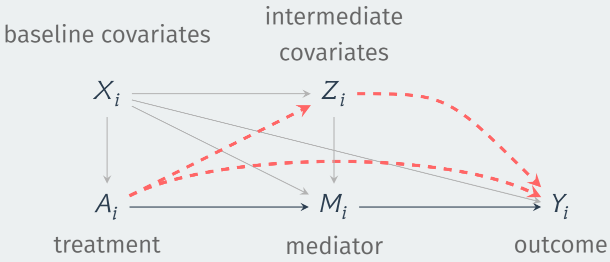
Overlap in the covariate distributions across levels of A_i and M_i

The Problem



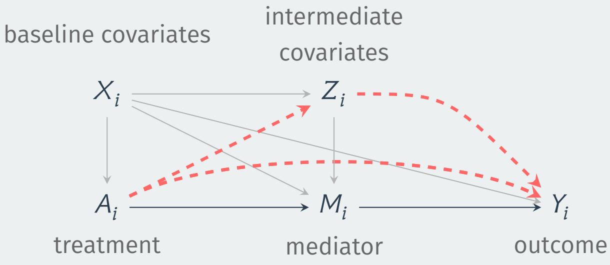
baseline covariates
intermediate covariates
X_i
Z_i
A_i
M_i
Y_i
treatment
mediator
outcome

The Problem



naive regression/matching of Y_i on X_i, A_i, M_i, and...

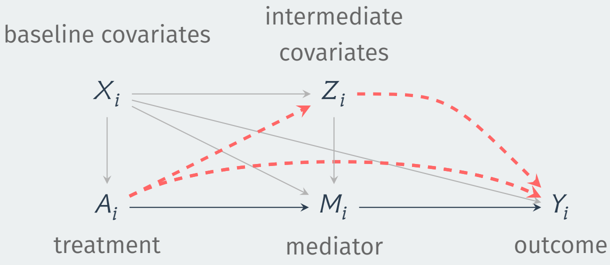
The Problem



naive regression/matching of Y_i on X_i, A_i, M_i, and...

Omit Z_i

The Problem

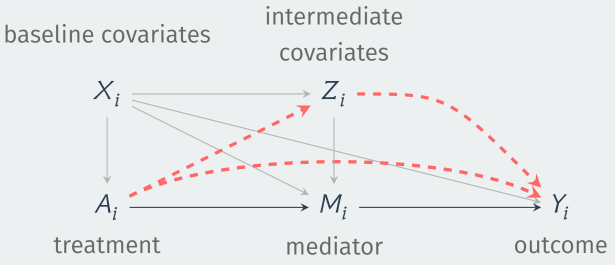


naive regression/matching of Y_i on X_i, A_i, M_i, and...

Omit Z_i

omitted variable bias for M_i

The Problem



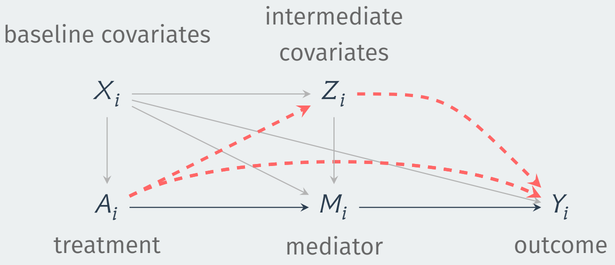
naive regression/matching of Y_i on X_i, A_i, M_i, and...

Omit Z_i

omitted variable bias for M_i

Control for Z_i

The Problem



naive regression/matching of Y_i on X_i, A_i, M_i, and...

Omit Z_i

omitted variable bias for M_i

Control for Z_i

post-treatment bias for A_i

Extant solutions are model dependent

Extant solutions are model dependent

Structural Nested Mean Models (SNMMs)

Need the correct model for $\mathbb{E}[Y_i | X_i, A_i, Z_i, M_i]$ and $\mathbb{E}[Y_i | X_i, A_i]$

Extant solutions are model dependent

Structural Nested Mean Models (SNMMs)

Need the correct model for $\mathbb{E}[Y_i | X_i, A_i, Z_i, M_i]$ and $\mathbb{E}[Y_i | X_i, A_i]$

Inverse probability of treatment weighting (IPTW)

Need the correct model for $\mathbb{P}[M_i | X_i, A_i, Z_i]$ and $\mathbb{P}[A_i | X_i]$

2/ Our approach: telescope matching

Telescope matching

Telescope matching

Two-stage matching procedure

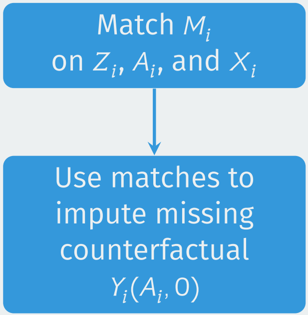
Telescope matching

Two-stage matching procedure

Match M_i
on Z_i, A_i, and X_i

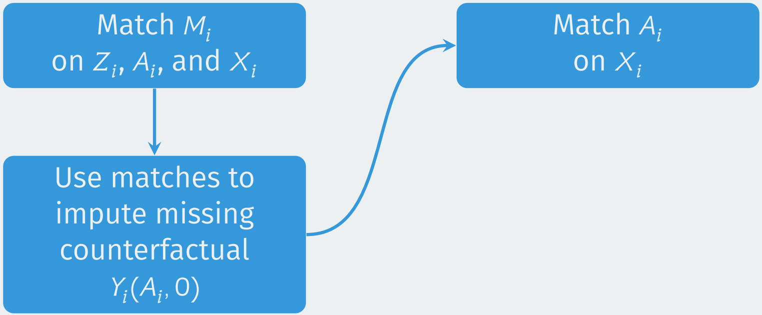
Telescope matching

Two-stage matching procedure



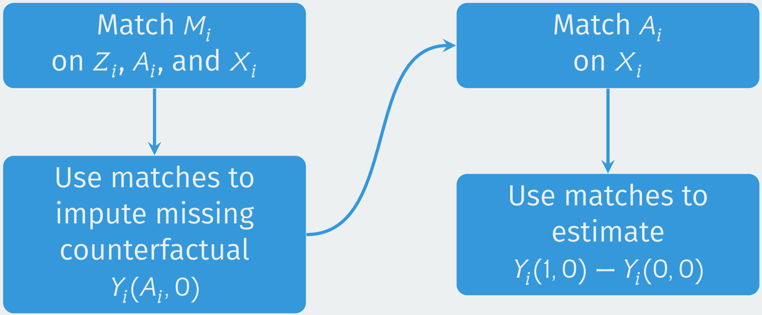
Telescope matching

Two-stage matching procedure



Telescope matching

Two-stage matching procedure



An imputation problem

Unit	Observed				Potential Outcomes			
	A_i	M_i	X_i	Z_i	$Y_i(1,1)$	$Y_i(1,0)$	$Y_i(0,1)$	$Y_i(0,0)$

An imputation problem

Unit	Observed				Potential Outcomes			
	A_i	M_i	X_i	Z_i	$Y_i(1,1)$	$Y_i(1,0)$	$Y_i(0,1)$	$Y_i(0,0)$
1	1	1	10	3	Y_1	?	?	?

An imputation problem

Unit	Observed				Potential Outcomes			
	A_i	M_i	X_i	Z_i	$Y_i(1,1)$	$Y_i(1,0)$	$Y_i(0,1)$	$Y_i(0,0)$
1	1	1	10	3	Y_1	?	?	?
2	1	0	9	2	?	Y_2	?	?

An imputation problem

Unit	Observed				Potential Outcomes			
	A_i	M_i	X_i	Z_i	$Y_i(1,1)$	$Y_i(1,0)$	$Y_i(0,1)$	$Y_i(0,0)$
1	1	1	10	3	Y_1	?	?	?
2	1	0	9	2	?	Y_2	?	?
3	1	0	8	1	?	Y_3	?	?
4	0	1	8	3	?	?	Y_4	?
5	0	0	9	2	?	?	?	Y_5
6	0	0	10	1	?	?	?	Y_6

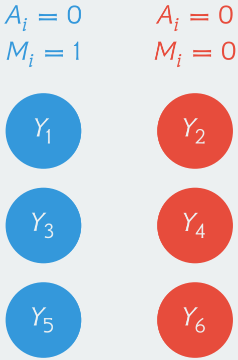
An imputation problem

Unit	Observed				Potential Outcomes			
	A_i	M_i	X_i	Z_i	$Y_i(1,1)$	$Y_i(1,0)$	$Y_i(0,1)$	$Y_i(0,0)$
1	1	1	10	3	Y_1	?	?	?
2	1	0	9	2	?	Y_2	?	?
3	1	0	8	1	?	Y_3	?	?
4	0	1	8	3	?	?	Y_4	?
5	0	0	9	2	?	?	?	Y_5
6	0	0	10	1	?	?	?	Y_6

$$\tau(0) = E[Y_i(1,0) - Y_i(0,0)]$$

First stage

First stage



First stage

1. Subset to a particular level of A_i

$A_i = 0$ $M_i = 1$

Y_1

Y_3

Y_5

$A_i = 0$ $M_i = 0$

Y_2

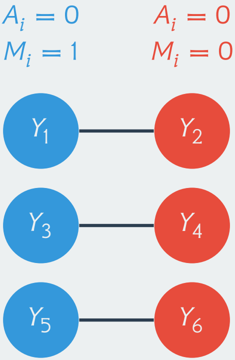
Y_4

Y_6

First stage

1. Subset to a particular level of A_i

2. Match each $M_i = 1$ to closest $M_i = 0$ unit in $\{X_i, Z_i\}$

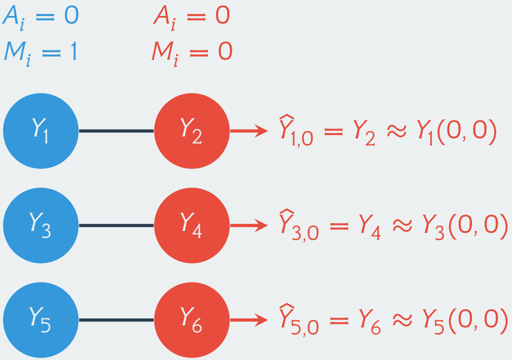


First stage

1. Subset to a particular level of A_i

2. Match each $M_i = 1$ to closest $M_i = 0$ unit in $\{X_i, Z_i\}$

3. Impute missing counterfactual with matched $\widehat{Y}_{i0}$

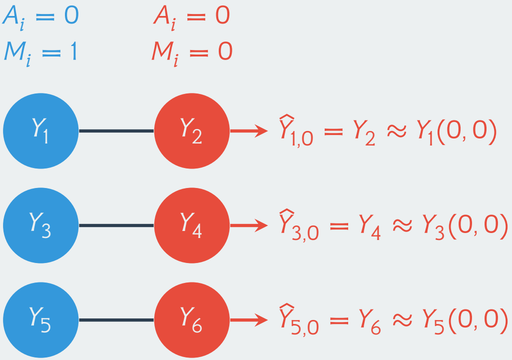


First stage

1. Subset to a particular level of A_i

2. Match each $M_i = 1$ to closest $M_i = 0$ unit in $\{X_i, Z_i\}$

3. Impute missing counterfactual with matched $\widehat{Y}_{i0}$



$$\widehat{Y}_{i0} = \begin{cases} Y_i & \text{if } M_i = 0 \\ Y_\ell & \text{if } M_i = 1, M_\ell = 0 \text{ and } \ell \text{ is matched to } i \end{cases}$$

1:1 matching example

Unit	Observed				Potential Outcomes			
	A_i	M_i	X_i	Z_i	$Y_i(1,1)$	$Y_i(1,0)$	$Y_i(0,1)$	$Y_i(0,0)$
1	1	1	10	3	Y_1	?	?	?
2	1	0	9	2	?	Y_2	?	?
3	1	0	8	1	?	Y_3	?	?
4	0	1	8	3	?	?	Y_4	?
5	0	0	9	2	?	?	?	Y_5
6	0	0	10	1	?	?	?	Y_6

1:1 matching example

Unit	Observed				Potential Outcomes			
	A_i	M_i	X_i	Z_i	$Y_i(1,1)$	$Y_i(1,0)$	$Y_i(0,1)$	$Y_i(0,0)$
1	1	1	10	3	Y_1	?	?	?
2	1	0	9	2	?	Y_2	?	?
3	1	0	8	1	?	Y_3	?	?
4	0	1	8	3	?	?	Y_4	?
5	0	0	9	2	?	?	?	Y_5
6	0	0	10	1	?	?	?	Y_6

1:1 matching example

Unit	Observed				Potential Outcomes			
	A_i	M_i	X_i	Z_i	$Y_i(1,1)$	$Y_i(1,0)$	$Y_i(0,1)$	$Y_i(0,0)$
1	1	1	10	3	Y_1	Y_2	?	?
2	1	0	9	2	?	Y_2	?	?
3	1	0	8	1	?	Y_3	?	?
4	0	1	8	3	?	?	Y_4	?
5	0	0	9	2	?	?	?	Y_5
6	0	0	10	1	?	?	?	Y_6

Second stage

Standard matching using $\widehat{Y}_{i0}$ as outcome completely ignoring M_i and Z_i

Second stage

Standard matching using $\widehat{Y}_{i0}$ as outcome completely ignoring M_i and Z_i

1. Match each $A_i = 0$ to closest $A_j = 1$ unit in X

Second stage

Standard matching using $\widehat{Y}_{i0}$ as outcome completely ignoring M_i and Z_i

1. Match each $A_i = 0$ to closest $A_j = 1$ unit in X

$$\widehat{Y}_i(1,0) = \begin{cases} \widehat{Y}_{j0} & \text{if } A_i = 0 \text{ \& } j \text{ is match for } i \\ \widehat{Y}_{i0} & \text{if } A_i = 1 \end{cases}$$

Second stage

Standard matching using $\widehat{Y}_{i0}$ as outcome completely ignoring M_i and Z_i

1. Match each $A_i = 0$ to closest $A_j = 1$ unit in X

2. Match each $A_i = 1$ to closest $A_j = 0$ unit in X

$$\widehat{Y}_i(1,0) = \begin{cases} \widehat{Y}_{j0} & \text{if } A_i = 0 \text{ \& } j \text{ is match for } i \\ \widehat{Y}_{i0} & \text{if } A_i = 1 \end{cases}$$

Second stage

Standard matching using $\widehat{Y}_{i0}$ as outcome completely ignoring M_i and Z_i

1. Match each $A_i = 0$ to closest $A_j = 1$ unit in X

$$\widehat{Y}_i(1,0) = \begin{cases} \widehat{Y}_{j0} & \text{if } A_i = 0 \text{ \& } j \text{ is match for } i \\ \widehat{Y}_{i0} & \text{if } A_i = 1 \end{cases}$$

2. Match each $A_i = 1$ to closest $A_j = 0$ unit in X

$$\widehat{Y}_i(0,0) = \begin{cases} \widehat{Y}_{i0} & \text{if } A_i = 0 \\ \widehat{Y}_{j0} & \text{if } A_i = 1 \text{ \& } j \text{ is match for } i \end{cases}$$

Second stage

Standard matching using $\widehat{Y}_{i0}$ as outcome completely ignoring M_i and Z_i

1. Match each $A_i = 0$ to closest $A_j = 1$ unit in X

$$\widehat{Y}_i(1,0) = \begin{cases} \widehat{Y}_{j0} & \text{if } A_i = 0 \text{ \& } j \text{ is match for } i \\ \widehat{Y}_{i0} & \text{if } A_i = 1 \end{cases}$$

2. Match each $A_i = 1$ to closest $A_j = 0$ unit in X

$$\widehat{Y}_i(0,0) = \begin{cases} \widehat{Y}_{i0} & \text{if } A_i = 0 \\ \widehat{Y}_{j0} & \text{if } A_i = 1 \text{ \& } j \text{ is match for } i \end{cases}$$

3. Take difference in means to estimate ACDE

$$\widehat{\tau} = \frac{1}{N}\sum_{i=1}^{N} \widehat{Y}_i(1,0) - \widehat{Y}_i(0,0)$$

1:1 matching, second stage

Unit	Observed				Potential Outcomes			
	A_i	M_i	X_i	Z_i	$Y_i(1,1)$	$Y_i(1,0)$	$Y_i(0,1)$	$Y_i(0,0)$
1	1	1	10	3	Y_1	Y_2	?	?
2	1	0	9	2	?	Y_2	?	?
3	1	0	8	1	?	Y_3	?	?
4	0	1	8	3	?	?	Y_4	?
5	0	0	9	2	?	?	?	Y_5
6	0	0	10	1	?	?	?	Y_6

1:1 matching, second stage

Unit	Observed				Potential Outcomes			
	A_i	M_i	X_i	Z_i	$Y_i(1,1)$	$Y_i(1,0)$	$Y_i(0,1)$	$Y_i(0,0)$
1	1	1	10	3	Y_1	Y_2	?	?
2	1	0	9	2	?	Y_2	?	?
3	1	0	8	1	?	Y_3	?	?
4	0	1	8	3	?	Y_3	Y_4	?
5	0	0	9	2	?	?	?	Y_5
6	0	0	10	1	?	?	?	Y_6

1:1 matching, second stage

Unit	Observed				Potential Outcomes			
	A_i	M_i	X_i	Z_i	$Y_i(1,1)$	$Y_i(1,0)$	$Y_i(0,1)$	$Y_i(0,0)$
1	1	1	10	3	Y_1	Y_2	?	?
2	1	0	9	2	?	Y_2	?	?
3	1	0	8	1	?	Y_3	?	?
4	0	1	8	3	?	Y_3	Y_4	?
5	0	0	9	2	?	Y_2	?	Y_5
6	0	0	10	1	?	?	?	Y_6

1:1 matching, second stage

Unit	Observed				Potential Outcomes			
	A_i	M_i	X_i	Z_i	$Y_i(1,1)$	$Y_i(1,0)$	$Y_i(0,1)$	$Y_i(0,0)$
1	1	1	10	3	Y_1	Y_2	?	?
2	1	0	9	2	?	Y_2	?	?
3	1	0	8	1	?	Y_3	?	?
4	0	1	8	3	?	Y_3	Y_4	?
5	0	0	9	2	?	Y_2	?	Y_5
6	0	0	10	1	?	Y_2	?	Y_6

1:1 matching, second stage

Unit	Observed				Potential Outcomes			
	A_i	M_i	X_i	Z_i	$Y_i(1,1)$	$Y_i(1,0)$	$Y_i(0,1)$	$Y_i(0,0)$
1	1	1	10	3	Y_1	Y_2	?	Y_6
2	1	0	9	2	?	Y_2	?	Y_5
3	1	0	8	1	?	Y_3	?	Y_5
4	0	1	8	3	?	Y_3	Y_4	Y_5
5	0	0	9	2	?	Y_2	?	Y_5
6	0	0	10	1	?	Y_2	?	Y_6

1:1 matching, second stage

Unit	Observed				Potential Outcomes			
	A_i	M_i	X_i	Z_i	$Y_i(1,1)$	$Y_i(1,0)$	$Y_i(0,1)$	$Y_i(0,0)$
1	1	1	10	3	Y_1	Y_2	?	Y_6
2	1	0	9	2	?	Y_2	?	Y_5
3	1	0	8	1	?	Y_3	?	Y_5
4	0	1	8	3	?	Y_3	Y_4	Y_5
5	0	0	9	2	?	Y_2	?	Y_5
6	0	0	10	1	?	Y_2	?	Y_6

$$\widehat{\tau} = \frac{1}{6}\Big[(Y_2 - Y_6) + (Y_2 - Y_5) + (Y_3 - Y_5) + (Y_3 - Y_5) + (Y_2 - Y_5) + (Y_2 - Y_6)\Big]$$

Properties of estimator

Properties of estimator

Simple

Properties of estimator

Simple

- two standard matching steps

Properties of estimator

Simple

- two standard matching steps
- both can be done without Y_i (avoid p-hacking)

Properties of estimator

Simple

- two standard matching steps
- both can be done without Y_i (avoid p-hacking)

Consistent

Properties of estimator

Simple

- two standard matching steps
- both can be done without Y_i (avoid p-hacking)

Consistent

- analysis similar to Abadie & Imbens (2006)

Properties of estimator

Simple

- two standard matching steps
- both can be done without Y_i (avoid p-hacking)

Consistent

- analysis similar to Abadie & Imbens (2006)
- under regularity conditions, $\hat{\tau}$ converges to ACDE

Properties of estimator

Simple

- two standard matching steps
- both can be done without Y_i (avoid p-hacking)

Consistent

- analysis similar to Abadie & Imbens (2006)
- under regularity conditions, $\hat{\tau}$ converges to ACDE

Biased

Properties of estimator

Simple

- two standard matching steps
- both can be done without Y_i (avoid p-hacking)

Consistent

- analysis similar to Abadie & Imbens (2006)
- under regularity conditions, $\hat{\tau}$ converges to ACDE

Biased

- Bias converges to 0 very slowly

Properties of estimator

Simple

- two standard matching steps
- both can be done without Y_i (avoid p-hacking)

Consistent

- analysis similar to Abadie & Imbens (2006)
- under regularity conditions, $\hat{\tau}$ converges to ACDE

Biased

- Bias converges to 0 very slowly
- $\rightsquigarrow$ doesn’t converge to normal

Properties of estimator

Simple

- two standard matching steps
- both can be done without Y_i (avoid p-hacking)

Consistent

- analysis similar to Abadie & Imbens (2006)
- under regularity conditions, $\hat{\tau}$ converges to ACDE

Biased

- Bias converges to 0 very slowly
- $\rightsquigarrow$ doesn't converge to normal
- follow Abadie and Imbens (2011) and develop bias correction

Properties of estimator

Simple

- two standard matching steps
- both can be done without Y_i (avoid p-hacking)

Consistent

- analysis similar to Abadie & Imbens (2006)
- under regularity conditions, $\hat{\tau}$ converges to ACDE

Biased

- Bias converges to 0 very slowly
- $\rightsquigarrow$ doesn't converge to normal
- follow Abadie and Imbens (2011) and develop bias correction

Robust

Properties of estimator

Simple

- two standard matching steps
- both can be done without Y_i (avoid p-hacking)

Consistent

- analysis similar to Abadie & Imbens (2006)
- under regularity conditions, $\hat{\tau}$ converges to ACDE

Biased

- Bias converges to 0 very slowly
- $\rightsquigarrow$ doesn't converge to normal
- follow Abadie and Imbens (2011) and develop bias correction

Robust

- more robust to model misspecification

Matching and the bootstrap

Matching and the bootstrap

Variance of $\hat{\tau}$ is *complicated*

Matching and the bootstrap

Variance of $\hat{\tau}$ is *complicated*

- each i could be matched multiple times at each stage

Matching and the bootstrap

Variance of $\hat{\tau}$ is *complicated*

- each i could be matched multiple times at each stage
- $\rightsquigarrow$ $\hat{\tau}$ is not a sum of i.i.d. variables.

Matching and the bootstrap

Variance of $\hat{\tau}$ is *complicated*

- each i could be matched multiple times at each stage
- $\rightsquigarrow$ $\hat{\tau}$ is not a sum of i.i.d. variables.

Nonparametric bootstrap?

Matching and the bootstrap

Variance of $\hat{\tau}$ is *complicated*

- each i could be matched multiple times at each stage
- $\rightsquigarrow$ $\hat{\tau}$ is not a sum of i.i.d. variables.

Nonparametric bootstrap?

- Abadie and Imbens (2008) show naively resampling rows is invalid for matching estimators

Matching and the bootstrap

Variance of $\widehat{\tau}$ is *complicated*

- each i could be matched multiple times at each stage
- $\rightsquigarrow$ $\widehat{\tau}$ is not a sum of i.i.d. variables.

Nonparametric bootstrap?

- Abadie and Imbens (2008) show naively resampling rows is invalid for matching estimators

Weighted bootstrap

Matching and the bootstrap

Variance of $\widehat{\tau}$ is *complicated*

- each i could be matched multiple times at each stage
- $\rightsquigarrow$ $\widehat{\tau}$ is not a sum of i.i.d. variables.

Nonparametric bootstrap?

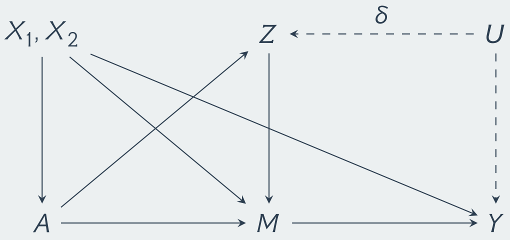
- Abadie and Imbens (2008) show naively resampling rows is invalid for matching estimators

Weighted bootstrap

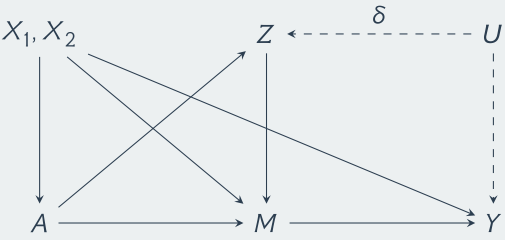
- we follow Otsu and Rai (2017) and resample each contribution to the estimator

3/ Simulating misspecification

Simulation set-up

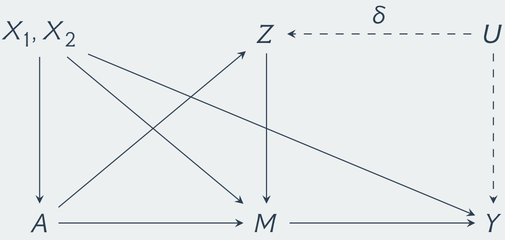


Simulation set-up



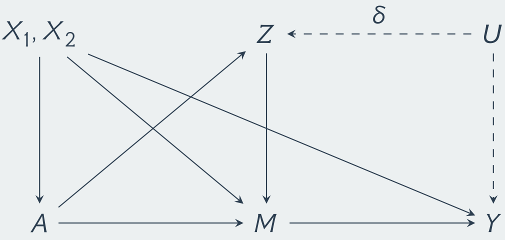
- All variables observed except U $\rightsquigarrow$ sequential ignorability holds

Simulation set-up



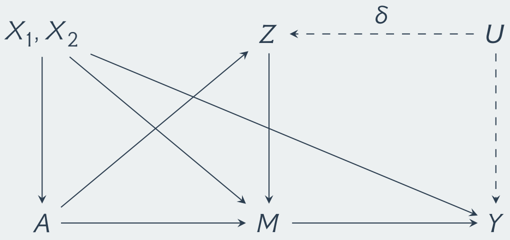
- All variables observed except U $\rightsquigarrow$ sequential ignorability holds
- Effect of A only through M so true ACDE: $\tau(0) = 0$

Simulation set-up



- All variables observed except $U \rightsquigarrow$ sequential ignorability holds
- Effect of A only through M so true ACDE: $\tau(0) = 0$
- δ controls magnitude of post-treatment confounding

Simulation set-up



- All variables observed except $U \rightsquigarrow$ sequential ignorability holds
- Effect of A only through M so true ACDE: $\tau(0) = 0$
- δ controls magnitude of post-treatment confounding
  - when $\delta \neq 0$, controlling for Z in a naive regression will induce post-treatment bias.

Simulation set-up

Simulation set-up

- Model misspecification as mismeasured confounders (Kang and Schafer, 2007)

Simulation set-up

- Model misspecification as mismeasured confounders (Kang and Schafer, 2007)
  - $X_1^* = \exp(X_1/2)$

Simulation set-up

- Model misspecification as mismeasured confounders (Kang and Schafer, 2007)
  - $X_1^* = \exp(X_1/2)$
  - $X_2^* = 1/(1+\exp(X_2))+10$

Simulation set-up

- Model misspecification as mismeasured confounders (Kang and Schafer, 2007)
  - $X_1^* = \exp(X_1/2)$
  - $X_2^* = 1/(1 + \exp(X_2)) + 10$
  - $Z_1^* = (Z_1/25 + 6)^3$

Simulation set-up

- Model misspecification as mismeasured confounders (Kang and Schafer, 2007)
  - $X_1^* = \exp(X_1/2)$
  - $X_2^* = 1/(1 + \exp(X_2)) + 10$
  - $Z_1^* = (Z_1/25 + 6)^3$
- Comparison methods:

Simulation set-up

- Model misspecification as mismeasured confounders (Kang and Schafer, 2007)
  - $X_1^* = \exp(X_1/2)$
  - $X_2^* = 1/(1 + \exp(X_2)) + 10$
  - $Z_1^* = (Z_1/25 + 6)^3$
- Comparison methods:
  - Naive regression conditioning on everything

Simulation set-up

- Model misspecification as mismeasured confounders (Kang and Schafer, 2007)
  - $X_1^* = \exp(X_1/2)$
  - $X_2^* = 1/(1 + \exp(X_2)) + 10$
  - $Z_1^* = (Z_1/25 + 6)^3$
- Comparison methods:
  - Naive regression conditioning on everything
  - Sequential g-estimation (SNMM with all linear CEFs)

Simulation set-up

- Model misspecification as mismeasured confounders (Kang and Schafer, 2007)
  - $X_1^* = \exp(X_1/2)$
  - $X_2^* = 1/(1 + \exp(X_2)) + 10$
  - $Z_1^* = (Z_1/25 + 6)^3$
- Comparison methods:
  - Naive regression conditioning on everything
  - Sequential g-estimation (SNMM with all linear CEFs)
  - Telescope matching with bias correction

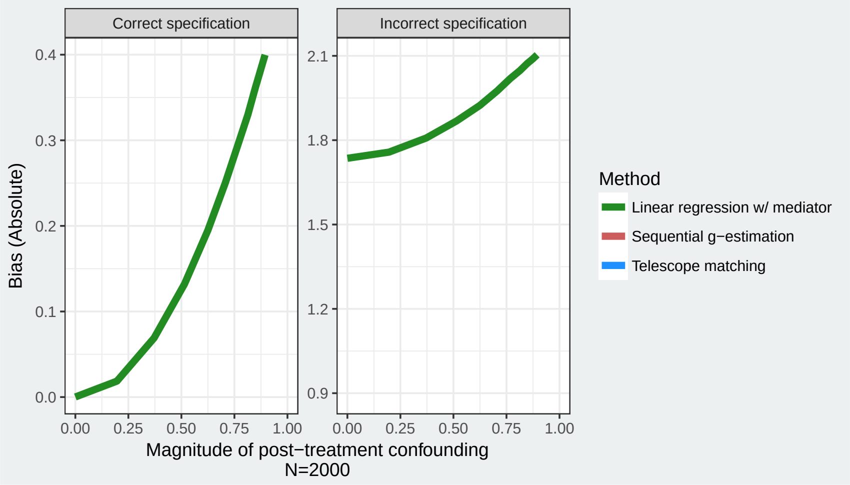
Simulation set-up

- Model misspecification as mismeasured confounders (Kang and Schafer, 2007)
    - $X_1^* = \exp(X_1/2)$
    - $X_2^* = 1/(1 + \exp(X_2)) + 10$
    - $Z_1^* = (Z_1/25 + 6)^3$
- Comparison methods:
    - Naive regression conditioning on everything
    - Sequential g-estimation (SNMM with all linear CEFs)
    - Telescope matching with bias correction

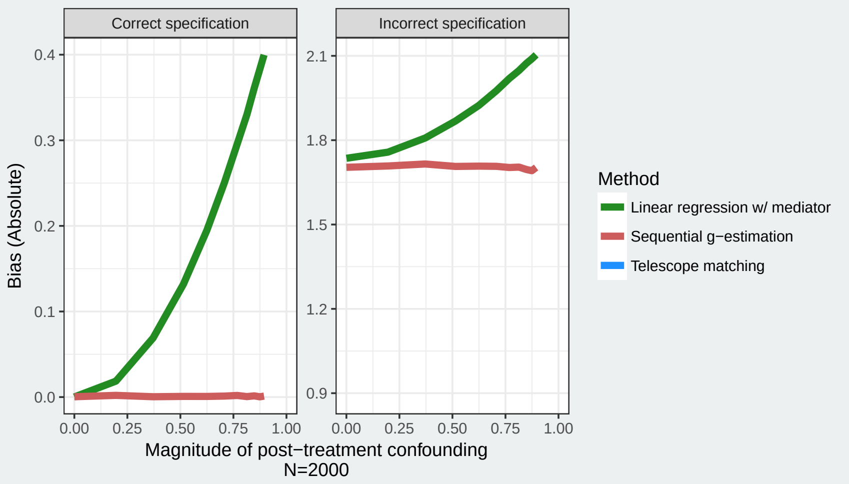
  Number of matches per stage = 3

Simulation results: Bias

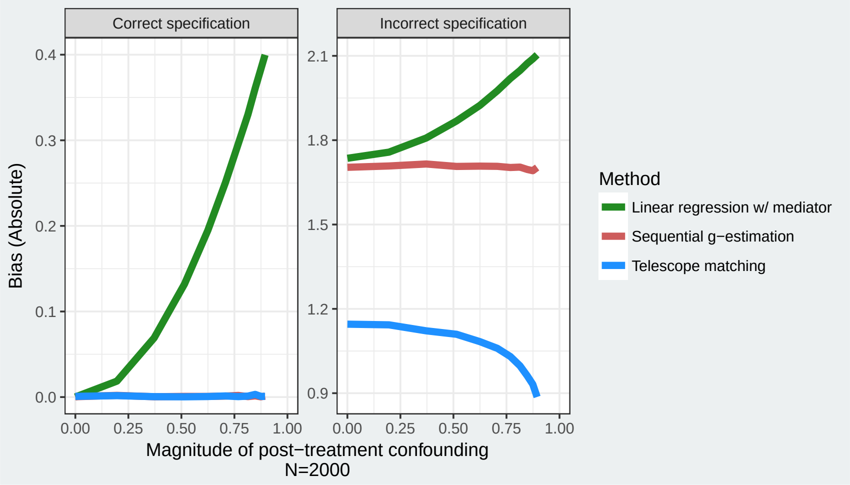
Simulation results: Bias



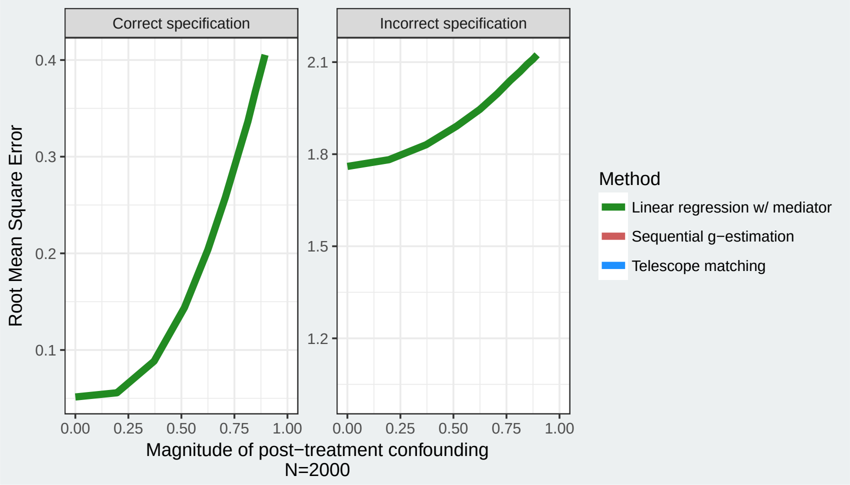
Simulation results: Bias



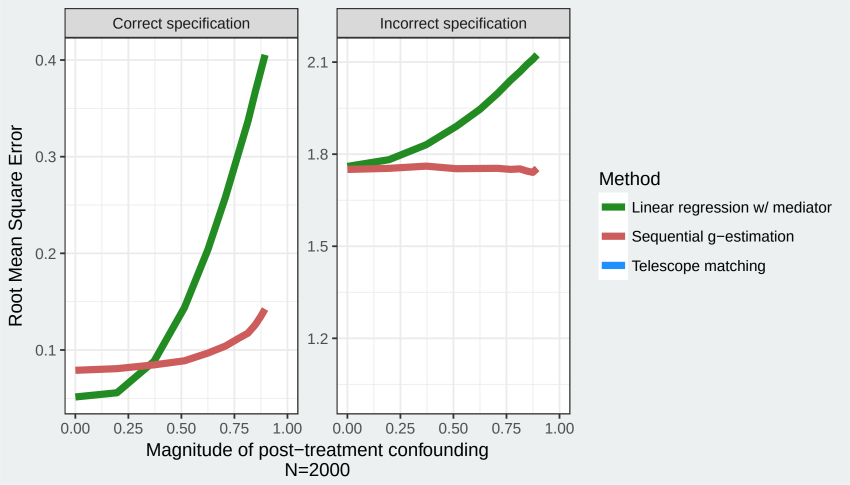
Simulation results: Bias



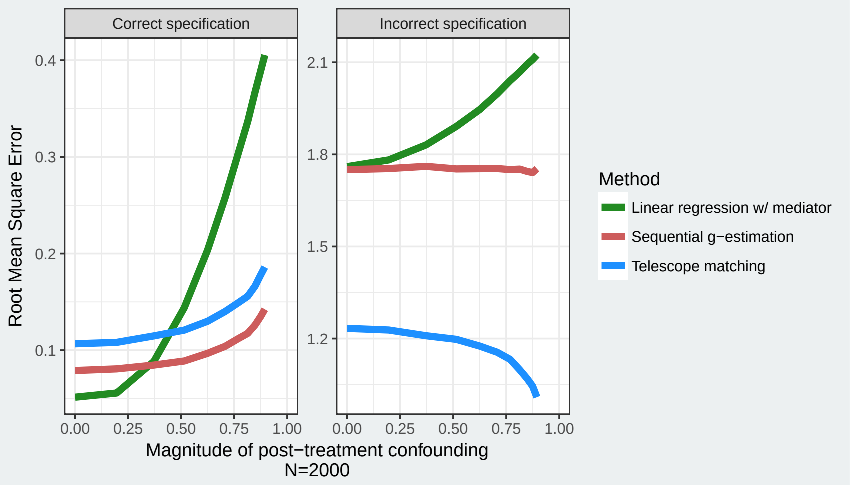
Simulation results: Root Mean Square Error



Simulation results: Root Mean Square Error



Simulation results: Root Mean Square Error



4/ Application

Brader, Suhay, Valentino (2008)

- Experiment on effect of media messages on support for immigration.

Brader, Suhay, Valentino (2008)

- Experiment on effect of media messages on support for immigration.
- Main effect: Story w/ negative tone + non-white immigrant reduced support for immigration.

Brader, Suhay, Valentino (2008)

- Experiment on effect of media messages on support for immigration.
- Main effect: Story w/ negative tone + non-white immigrant reduced support for immigration.
- Question: Treatment also affected levels of anxiety (M_i). Is there an effect of treatment that remains if the anxiety mediator is held fixed?

Brader, Suhay, Valentino (2008)

- Experiment on effect of media messages on support for immigration.
- Main effect: Story w/ negative tone + non-white immigrant reduced support for immigration.
- Question: Treatment also affected levels of anxiety (M_i). Is there an effect of treatment that remains if the anxiety mediator is held fixed?
    - Mediation assumption might be suspect.

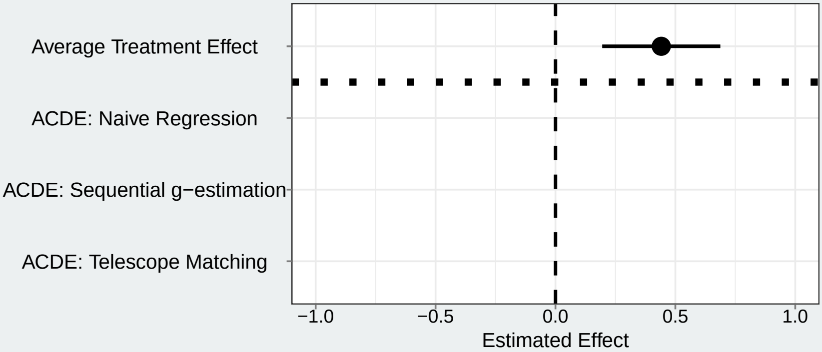
Brader, Suhay, Valentino (2008)

- Experiment on effect of media messages on support for immigration.
- Main effect: Story w/ negative tone + non-white immigrant reduced support for immigration.
- Question: Treatment also affected levels of anxiety (M_i). Is there an effect of treatment that remains if the anxiety mediator is held fixed?
  - Mediation assumption might be suspect.
- Pre-treatment confounders (X_i): Education, Gender, Income, Age.

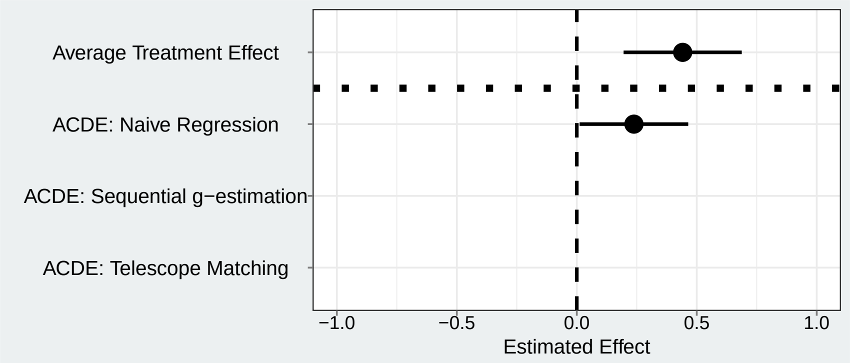
Brader, Suhay, Valentino (2008)

- Experiment on effect of media messages on support for immigration.
- Main effect: Story w/ negative tone + non-white immigrant reduced support for immigration.
- Question: Treatment also affected levels of anxiety (M_i). Is there an effect of treatment that remains if the anxiety mediator is held fixed?
  - Mediation assumption might be suspect.
- Pre-treatment confounders (X_i): Education, Gender, Income, Age.
- Post-treatment confounder (Z_i): Perceived harm due to immigration.

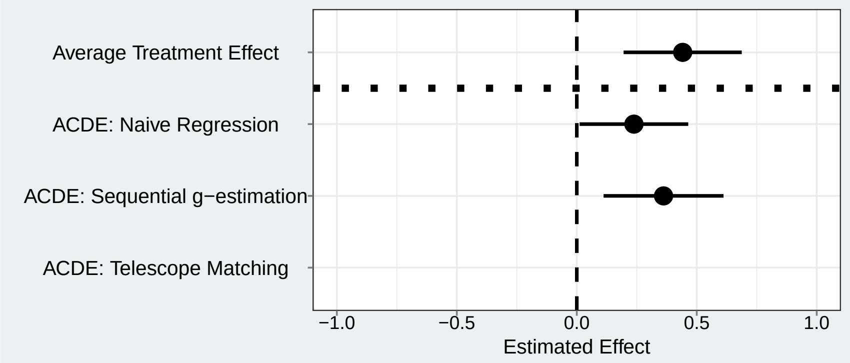
Brader, Suhay, Valentino (2008)



Brader, Suhay, Valentino (2008)

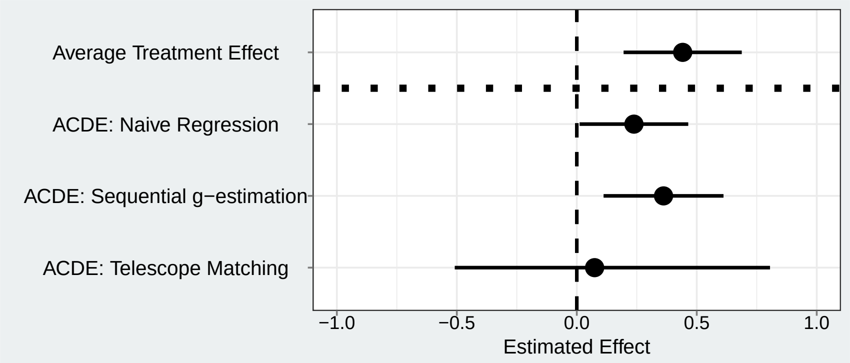


Brader, Suhay, Valentino (2008)



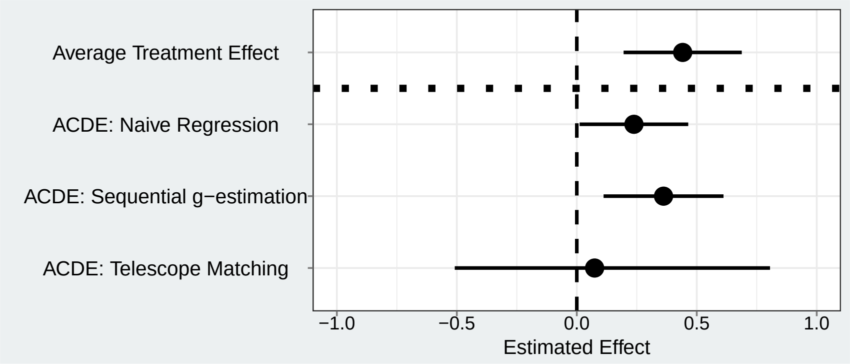
- Sequential g-estimation suggests a non-zero ACDE—there exists an effect even if we fix anxiety.

Brader, Suhay, Valentino (2008)



- Sequential g-estimation suggests a non-zero ACDE—there exists an effect even if we fix anxiety.
- Telescope matching shows ACDE closer to zero, high uncertainty.

Brader, Suhay, Valentino (2008)



- Sequential g-estimation suggests a non-zero ACDE—there exists an effect even if we fix anxiety.
- Telescope matching shows ACDE closer to zero, high uncertainty.
- ⇝ Fixing the mediator eliminates most of the treatment effect.

5/ Conclusion

Conclusion

- Standard matching doesn’t work for direct effects.

Conclusion

- Standard matching doesn't work for direct effects.
- Direct effects models such as SNMMs and IPTW are model dependent.

Conclusion

- Standard matching doesn't work for direct effects.
- Direct effects models such as SNMMs and IPTW are model dependent.
- We introduce two-stage matching procedure to close this gap.

Conclusion

- Standard matching doesn't work for direct effects.
- Direct effects models such as SNMMs and IPTW are model dependent.
- We introduce two-stage matching procedure to close this gap.
  - Estimator is consistent, but biased, so we use bias correction.

Conclusion

- Standard matching doesn't work for direct effects.
- Direct effects models such as SNMMs and IPTW are model dependent.
- We introduce two-stage matching procedure to close this gap.
  - Estimator is consistent, but biased, so we use bias correction.
  - Weighted bootstrap for uncertainty estimates.

Next steps

- Better variance estimators to handle undercoverage in smaller samples.

Next steps

- Better variance estimators to handle undercoverage in smaller samples.
- Apply ideas to mediation analysis where there are no Z_i.

Next steps

- Better variance estimators to handle undercoverage in smaller samples.
- Apply ideas to mediation analysis where there are no Z_i.
- How to handle dropping units in the first stage since it induces post-treatment bias?

Thanks!

- For more information, see:
- http://www.mattblackwell.org/files/papers/telescope_matchin
- http://www.mattblackwell.org
- https://www.antonstrezhnev.com/

SNMMs as imputation estimators

SNMMs as imputation estimators

1. Estimate the conditional effect of M_i

SNMMs as imputation estimators

1. Estimate the conditional effect of M_i

$$\gamma_m(x, z, a) = \mathbb{E}[Y_i(a, 1) - Y_i(a, 0) | X_i = x, A_i = a, Z_i = z, M_i = 1]$$

SNMMs as imputation estimators

1. Estimate the conditional effect of M_i

$$\gamma_m(x, z, a) = \mathbb{E}[Y_i(a, 1) - Y_i(a, 0) | X_i = x, A_i = a, Z_i = z, M_i = 1]$$

2. Impute $Y_i(A_i, 0)$ by subtracting effect of M_i

SNMMs as imputation estimators

1. Estimate the conditional effect of M_i

$$\gamma_m(x, z, a) = \mathbb{E}[Y_i(a, 1) - Y_i(a, 0) | X_i = x, A_i = a, Z_i = z, M_i = 1]$$

2. Impute $Y_i(A_i, 0)$ by subtracting effect of M_i

$$\overline{Y}_i(A_i, 0) \; = \; Y_i - M_i \times \widehat{\gamma}_m(X_i, Z_i, A_i)$$

SNMMs as imputation estimators

1. Estimate the conditional effect of M_i

$$\gamma_m(x, z, a) = \mathbb{E}[Y_i(a, 1) - Y_i(a, 0) | X_i = x, A_i = a, Z_i = z, M_i = 1]$$

2. Impute $Y_i(A_i, 0)$ by subtracting effect of M_i

$$\overline{Y}_i(A_i, 0) = Y_i - M_i \times \widehat{\gamma}_m(X_i, Z_i, A_i)$$

3. Regress imputations on treatment and baseline covariates to get ACDE

SNMMs as imputation estimators

1. Estimate the conditional effect of M_i

$$\gamma_m(x, z, a) = \mathbb{E}[Y_i(a, 1) - Y_i(a, 0) | X_i = x, A_i = a, Z_i = z, M_i = 1]$$

2. Impute $Y_i(A_i, 0)$ by subtracting effect of M_i

$$\overline{Y}_i(A_i, 0) \; = \; Y_i - M_i \times \widehat{\gamma}_m(X_i, Z_i, A_i)$$

3. Regress imputations on treatment and baseline covariates to get ACDE

$$\mathbb{E}[\overline{Y}_i(A_i, 0) | X_i, A_i]$$

SNMMs as imputation estimators

1. Estimate the conditional effect of M_i

$$\gamma_m(x, z, a) = \mathbb{E}[Y_i(a, 1) - Y_i(a, 0) | X_i = x, A_i = a, Z_i = z, M_i = 1]$$

Depends on correct model for $E[Y_i | X_i, Z_i, A_i, M_i]$

2. Impute $Y_i(A_i, 0)$ by subtracting effect of M_i

$$\overline{Y}_i(A_i, 0) \;=\; Y_i \,-\, M_i \,\times\, \widehat{\gamma}_m(X_i, Z_i, A_i)$$

3. Regress imputations on treatment and baseline covariates to get ACDE

$$\mathbb{E}[\overline{Y}_i(A_i, 0) | X_i, A_i]$$

SNMMs as imputation estimators

1. Estimate the conditional effect of M_i

$$\gamma_m(x, z, a) = \mathbb{E}[Y_i(a, 1) - Y_i(a, 0) | X_i = x, A_i = a, Z_i = z, M_i = 1]$$

Depends on correct model for $E[Y_i | X_i, Z_i, A_i, M_i]$

2. Impute $Y_i(A_i, 0)$ by subtracting effect of M_i

$$\overline{Y}_i(A_i, 0) = Y_i - M_i \times \widehat{\gamma}_m(X_i, Z_i, A_i)$$

Depends on correct model for $E[Y_i | X_i, Z_i, A_i, M_i]$

3. Regress imputations on treatment and baseline covariates to get ACDE

$$\mathbb{E}[\overline{Y}_i(A_i, 0) | X_i, A_i]$$

SNMMs as imputation estimators

1. Estimate the conditional effect of M_i

$$\gamma_m(x, z, a) = \mathbb{E}[Y_i(a, 1) - Y_i(a, 0) | X_i = x, A_i = a, Z_i = z, M_i = 1]$$

Depends on correct model for $E[Y_i | X_i, Z_i, A_i, M_i]$

2. Impute $Y_i(A_i, 0)$ by subtracting effect of M_i

$$\overline{Y}_i(A_i, 0) \ = \ Y_i \ - \ M_i \ \times \ \widehat{\gamma}_m(X_i, Z_i, A_i)$$

Depends on correct model for $E[Y_i | X_i, Z_i, A_i, M_i]$

3. Regress imputations on treatment and baseline covariates to get ACDE

$$\mathbb{E}[\overline{Y}_i(A_i, 0) | X_i, A_i]$$

Depends on correct model for $E[Y_i | X_i, Z_i, A_i, M_i]$ and $E[Y_i | X_i, A_i]$