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

why do we care?

causal mediation

causal mechanisms

lagged effects in TSCS data
regression & matching

posttreatment bias
Regression & matching:
- posttreatment bias

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

Weighting methods:
regression & matching

posttreatment bias

sequential g-estimation

model dependence

weighting methods
Telecsope matching

regression & matching

posttreatment bias

sequential g-estimation

weighting methods

model dependence
regression & matching
posttreatment bias

sequential g-estimation
model dependence

weighting methods

Telecsope matching

• consistent for direct effects
regression & matching
posttreatment bias

Telescope matching

sequential g-estimation
model dependence

weighting methods

- consistent for direct effects
- avoids post-treatment bias
Telecsope matching

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

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

Binary treatment $i \in \{\text{negative frame, positive frame}\}$

Binary mediator $M_i \in \{\text{high anxiety, low anxiety}\}$

Outcome (support for immigration) $Y_i$
Setting  Effect of frame on immigration media accounts

$A_i$  Binary treatment $\in \{\text{negative frame, positive frame}\}$
Notation

Setting  Effect of frame on immigration media accounts

\( A_i \)  Binary treatment \( \in \{ \text{negative frame, positive frame} \} \)

\( M_i \)  Binary mediator \( \in \{ \text{high anxiety, low anxiety} \} \)
Notation

Setting: Effect of frame on immigration media accounts

\( A_i \): Binary treatment \( \in \{ \text{negative frame, positive frame} \} \)

\( M_i \): Binary mediator \( \in \{ \text{high anxiety, low anxiety} \} \)

\( Y_i \): Outcome (support for immigration)
Notation

Setting

Effect of frame on immigration media accounts

$A_i$ Binary treatment $\in \{\text{negative frame, positive frame}\}$

$M_i$ Binary mediator $\in \{\text{high anxiety, low anxiety}\}$

$Y_i$ Outcome (support for immigration)

$Y_i(a, m)$ Potential outcome
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 \( m \)
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

$A_i \rightarrow M_i \rightarrow Y_i$

treatment \hspace{1cm} mediator \hspace{1cm} outcome
Confounders

baseline covariates

\[ X_i \]

\[ A_i \]  \rightarrow  \[ M_i \]  \rightarrow  \[ Y_i \]

treatment  mediator  outcome
Confounders

Baseline covariates: $X_i$
Intermediate covariates: $Z_i$
Treatment: $A_i$
Mediator: $M_i$
Outcome: $Y_i$
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)
\[ 0 < P(A_i = 1 | X_i = x) < 1 \]
\[ 0 < P(M_i = 1 | X_i = x, Z_i = z, A_i = a) < 1 \]
Assumptions

Assumption (Sequential Ignorability)

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

\quad Y_i(a, m) \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

<table>
<thead>
<tr>
<th>Xi</th>
<th>Zi</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ai</td>
<td></td>
</tr>
</tbody>
</table>

treatment

mediator

outcome

intermediate covariates

naive regression/matching of Yi on Xi, Ai, Mi, and...

Omit Zi omitted variable bias for Mi

Control for Zi post-treatment bias for Ai
The Problem

baseline covariates  intermediate covariates

\( X_i \)  \( Z_i \)

\( A_i \)  \( M_i \)

treatment  mediator

outcome

naive regression/matching of \( Y_i \) on \( X_i, A_i, M_i, \) and...
The Problem

baseline covariates

$X_i$ → $A_i$ → $Y_i$

intermediate covariates

$Z_i$ → $M_i$ → $Y_i$

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 $Z_i$

Control for $Z_i$

omitted variable bias for $M_i$
The Problem

baseline covariates \rightarrow \text{mediator} \rightarrow \text{outcome}

\begin{align*}
X_i & \rightarrow Z_i \\
A_i & \rightarrow M_i \\
& \rightarrow Y_i
\end{align*}

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

\begin{itemize}
\item Omit $Z_i$: omitted variable bias for $M_i$
\item Control for $Z_i$: post-treatment bias for $A_i$
\end{itemize}
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 $\mathbb{P}[A_i | X_i]$
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

Two-stage matching procedure

\[ \tilde{Z}_i, \bar{A}_i, \bar{X}_i \]

Use matches to impute missing counterfactual \( \tilde{Y}_i(\bar{A}_i, 0) \)

Match \( \bar{A}_i \) on \( \bar{X}_i \)

Use matches to estimate \( \tilde{Y}_i(1, 0) - \tilde{Y}_i(0, 0) \)
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

Match $M_i$ on $Z_i, A_i$, and $X_i$

Use matches to impute missing counterfactual $Y_i(A_i, 0)$
Telescope matching

Two-stage matching procedure

Match $M_i$ on $Z_i, A_i,$ and $X_i$

Use matches to impute missing counterfactual $Y_i(A_i, 0)$

Match $A_i$ on $X_i$
Telescope matching

Two-stage matching procedure

Match $M_i$ on $Z_i, A_i,$ and $X_i$

Use matches to impute missing counterfactual $Y_i(A_i, 0)$

Match $A_i$ on $X_i$

Use matches to estimate $Y_i(1, 0) - Y_i(0, 0)$
An imputation problem

<table>
<thead>
<tr>
<th>Unit</th>
<th>Observed</th>
<th>Potential Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_i$</td>
<td>$M_i$</td>
<td>$X_i$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
An imputation problem

<table>
<thead>
<tr>
<th>Unit</th>
<th>Observed</th>
<th>Potential Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$A_i$</td>
<td>$M_i$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
An imputation problem

<table>
<thead>
<tr>
<th>Unit</th>
<th>$A_i$</th>
<th>$M_i$</th>
<th>$X_i$</th>
<th>$Z_i$</th>
<th>$Y_{i(1,1)}$</th>
<th>$Y_{i(1,0)}$</th>
<th>$Y_{i(0,1)}$</th>
<th>$Y_{i(0,0)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>10</td>
<td>3</td>
<td>$Y_1$</td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>9</td>
<td>2</td>
<td>?</td>
<td>$Y_2$</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>

$Y_{i(1,1)}$, $Y_{i(1,0)}$, $Y_{i(0,1)}$, $Y_{i(0,0)}$ represent the observed and potential outcomes for unit $i$. The question marks indicate missing values that need to be imputed.
## An imputation problem

<table>
<thead>
<tr>
<th>Unit</th>
<th>Observed</th>
<th>Potential Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$A_i$  $M_i$ $X_i$ $Z_i$</td>
<td>$Y_{i}(1, 1)$ $Y_{i}(1, 0)$ $Y_{i}(0, 1)$ $Y_{i}(0, 0)$</td>
</tr>
<tr>
<td>1</td>
<td>1 1 10 3</td>
<td>$Y_1$ ? ? ?</td>
</tr>
<tr>
<td>2</td>
<td>1 0 9 2</td>
<td>? $Y_2$ ? ?</td>
</tr>
<tr>
<td>3</td>
<td>1 0 8 1</td>
<td>? $Y_3$ ? ?</td>
</tr>
<tr>
<td>4</td>
<td>0 1 8 3</td>
<td>? ? $Y_4$ ?</td>
</tr>
<tr>
<td>5</td>
<td>0 0 9 2</td>
<td>? ? ? $Y_5$</td>
</tr>
<tr>
<td>6</td>
<td>0 0 10 1</td>
<td>? ? ? $Y_6$</td>
</tr>
</tbody>
</table>
### An imputation problem

<table>
<thead>
<tr>
<th>Unit</th>
<th>Observed</th>
<th>Potential Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$A_i$</td>
<td>$Y_i(1, 1)$</td>
</tr>
<tr>
<td>1</td>
<td>1 1 10 3</td>
<td>$Y_1$</td>
</tr>
<tr>
<td>2</td>
<td>1 0 9 2</td>
<td>?</td>
</tr>
<tr>
<td>3</td>
<td>1 0 8 1</td>
<td>?</td>
</tr>
<tr>
<td>4</td>
<td>0 1 8 3</td>
<td>?</td>
</tr>
<tr>
<td>5</td>
<td>0 0 9 2</td>
<td>?</td>
</tr>
<tr>
<td>6</td>
<td>0 0 10 1</td>
<td>?</td>
</tr>
</tbody>
</table>

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

\[ A_i = 0 \]
\[ M_i = 1 \]
\[ Y_1 \]
\[ Y_3 \]
\[ Y_5 \]
\[ A_i = 0 \]
\[ M_i = 0 \]
\[ Y_2 \]
\[ Y_4 \]
\[ Y_6 \]

1. Subset to a particular level of \( A_i \)
2. Match each \( M_i = 1 \) to closest \( M_i = 0 \) unit in \( \{X_i, Z_i\} \)
   \[ \hat{Y}_{1,0} = Y_2 \approx Y_1(0, 0) \]
   \[ \hat{Y}_{3,0} = Y_4 \approx Y_3(0, 0) \]
   \[ \hat{Y}_{5,0} = Y_6 \approx Y_5(0, 0) \]
3. Impute missing counterfactual 
   \[ \hat{Y}_{i0} = \begin{cases} Y_i & \text{if } M_i = 0 \\ Y_\ell & \text{if } M_i = 1, M_\ell = 0 \end{cases} \]
   and \( \ell \) is matched to \( i \).
First stage

\[ A_i = 0 \]
\[ M_i = 1 \]
\[ Y_1 \]
\[ Y_3 \]
\[ Y_5 \]

\[ A_i = 0 \]
\[ M_i = 0 \]
\[ Y_2 \]
\[ Y_4 \]
\[ Y_6 \]
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 $\{X_i, Z_i\}$

$A_i = 0, M_i = 1$

$Y_1 \approx Y_2 (0, 0)$

$Y_3 \approx Y_4 (0, 0)$

$Y_5 \approx Y_6 (0, 0)$
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 $\hat{Y}_{i0}$

<table>
<thead>
<tr>
<th>$A_i$</th>
<th>$M_i$</th>
<th>$Y_i$</th>
<th>Imputation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>$Y_1$</td>
<td>$\hat{Y}_{1,0} = Y_2 \approx Y_1(0,0)$</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>$Y_2$</td>
<td>$\hat{Y}_{1,0} = Y_2 \approx Y_1(0,0)$</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>$Y_3$</td>
<td>$\hat{Y}_{3,0} = Y_4 \approx Y_3(0,0)$</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>$Y_4$</td>
<td>$\hat{Y}_{3,0} = Y_4 \approx Y_3(0,0)$</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>$Y_5$</td>
<td>$\hat{Y}_{5,0} = Y_6 \approx Y_5(0,0)$</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>$Y_6$</td>
<td>$\hat{Y}_{5,0} = Y_6 \approx Y_5(0,0)$</td>
</tr>
</tbody>
</table>
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 $\hat{Y}_{i0}$

\[
\hat{Y}_{i0} = \begin{cases} 
Y_i & \text{if } M_i = 0 \\
Y_\ell & \text{if } M_i = 1, M_\ell = 0 \text{ and } \ell \text{ is matched to } i 
\end{cases}
\]
### 1:1 matching example

<table>
<thead>
<tr>
<th>Unit</th>
<th>$A_i$</th>
<th>$M_i$</th>
<th>$X_i$</th>
<th>$Z_i$</th>
<th>$Y_i(1,1)$</th>
<th>$Y_i(1,0)$</th>
<th>$Y_i(0,1)$</th>
<th>$Y_i(0,0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>10</td>
<td>3</td>
<td>$Y_1$</td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>9</td>
<td>2</td>
<td>?</td>
<td>$Y_2$</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>8</td>
<td>1</td>
<td>?</td>
<td>$Y_3$</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>1</td>
<td>8</td>
<td>3</td>
<td>?</td>
<td>?</td>
<td>$Y_4$</td>
<td>?</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>2</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>$Y_5$</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>1</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>$Y_6$</td>
</tr>
</tbody>
</table>
### 1:1 matching example

<table>
<thead>
<tr>
<th>Unit</th>
<th>Observed</th>
<th>Potential Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$A_i$</td>
<td>$Y_{i}(1, 1)$</td>
</tr>
<tr>
<td></td>
<td>$M_i$</td>
<td>$Y_{i}(1, 0)$</td>
</tr>
<tr>
<td></td>
<td>$X_i$</td>
<td>$Y_{i}(0, 1)$</td>
</tr>
<tr>
<td></td>
<td>$Z_i$</td>
<td>$Y_{i}(0, 0)$</td>
</tr>
<tr>
<td>1</td>
<td>1 1 10 3</td>
<td>$Y_1$ ? ? ?</td>
</tr>
<tr>
<td>2</td>
<td>1 0 9 2</td>
<td>? $Y_2$ ? ?</td>
</tr>
<tr>
<td>3</td>
<td>1 0 8 1</td>
<td>? $Y_3$ ? ?</td>
</tr>
<tr>
<td>4</td>
<td>0 1 8 3</td>
<td>? ? $Y_4$ ?</td>
</tr>
<tr>
<td>5</td>
<td>0 0 9 2</td>
<td>? ? ? $Y_5$</td>
</tr>
<tr>
<td>6</td>
<td>0 0 10 1</td>
<td>? ? ? $Y_6$</td>
</tr>
</tbody>
</table>
### 1:1 matching example

<table>
<thead>
<tr>
<th>Unit</th>
<th>Observed</th>
<th>Potential Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$A_i$  $M_i$ $X_i$ $Z_i$</td>
<td>$Y_i(1,1)$ $Y_i(1,0)$ $Y_i(0,1)$ $Y_i(0,0)$</td>
</tr>
<tr>
<td>1</td>
<td>1 1 10 3</td>
<td>$Y_1$ $Y_2$ ? ?</td>
</tr>
<tr>
<td>2</td>
<td>1 0 9 2</td>
<td>? $Y_2$ ? ?</td>
</tr>
<tr>
<td>3</td>
<td>1 0 8 1</td>
<td>? $Y_3$ ? ?</td>
</tr>
<tr>
<td>4</td>
<td>0 1 8 3</td>
<td>? ? $Y_4$ ?</td>
</tr>
<tr>
<td>5</td>
<td>0 0 9 2</td>
<td>? ? ? $Y_5$</td>
</tr>
<tr>
<td>6</td>
<td>0 0 10 1</td>
<td>? ? ? $Y_6$</td>
</tr>
</tbody>
</table>
Second stage

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

Standard matching using $\hat{Y}_{i0}$ 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 $X$

3. Take difference in means to estimate $\hat{\tau} = \frac{1}{N} \sum_{i=1}^{N} \hat{Y}_i(1,0) - \hat{Y}_i(0,0)$
Second stage

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

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

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

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

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

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

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

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

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

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

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

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

3. Take difference in means to estimate ACDE:

$$\hat{\tau} = \frac{1}{N} \sum_{i=1}^{N} \hat{Y}_{i}(1, 0) - \hat{Y}_{i}(0, 0)$$
Second stage

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

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

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

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

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

3. Take difference in means to estimate ACDE

$$\hat{\tau} = \frac{1}{N} \sum_{i=1}^{N} \hat{Y}_{i}(1, 0) - \hat{Y}_{i}(0, 0)$$
1:1 matching, second stage

<table>
<thead>
<tr>
<th>Unit</th>
<th>$A_i$</th>
<th>$M_i$</th>
<th>$X_i$</th>
<th>$Z_i$</th>
<th>$Y_{i(1,1)}$</th>
<th>$Y_{i(1,0)}$</th>
<th>$Y_{i(0,1)}$</th>
<th>$Y_{i(0,0)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>10</td>
<td>3</td>
<td>$Y_1$</td>
<td>$Y_2$</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>9</td>
<td>2</td>
<td>?</td>
<td>$Y_2$</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>8</td>
<td>1</td>
<td>?</td>
<td>$Y_3$</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>1</td>
<td>8</td>
<td>3</td>
<td>?</td>
<td>?</td>
<td>$Y_4$</td>
<td>?</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>2</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>$Y_5$</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>1</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>$Y_6$</td>
</tr>
</tbody>
</table>
1:1 matching, second stage

<table>
<thead>
<tr>
<th>Unit</th>
<th>Observed</th>
<th>Potential Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$A_i$</td>
<td>$Y_{i}(1,1)$</td>
</tr>
<tr>
<td></td>
<td>$M_i$</td>
<td>$Y_{i}(1,0)$</td>
</tr>
<tr>
<td></td>
<td>$X_i$</td>
<td>$Y_{i}(0,1)$</td>
</tr>
<tr>
<td></td>
<td>$Z_i$</td>
<td>$Y_{i}(0,0)$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>$Y_1$</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>?</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>?</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>?</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>?</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>?</td>
</tr>
</tbody>
</table>

$\hat{\tau} = \sum_{i=1}^{6} (Y_{2} - Y_{6}) + (Y_{2} - Y_{5}) + (Y_{3} - Y_{5}) + (Y_{3} - Y_{5}) + (Y_{2} - Y_{5}) + (Y_{2} - Y_{6})$
### 1:1 matching, second stage

<table>
<thead>
<tr>
<th>Unit</th>
<th>Observed</th>
<th>Potential Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$A_i$</td>
<td>$M_i$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
## 1:1 matching, second stage

<table>
<thead>
<tr>
<th>Unit</th>
<th>Observed</th>
<th>Potential Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$A_i$  $M_i$  $X_i$  $Z_i$</td>
<td>$Y_i(1, 1)$  $Y_i(1, 0)$  $Y_i(0, 1)$  $Y_i(0, 0)$</td>
</tr>
<tr>
<td>1</td>
<td>1        1        10       3</td>
<td>$Y_1$  $Y_2$  ?  ?</td>
</tr>
<tr>
<td>2</td>
<td>1        0        9        2</td>
<td>?  $Y_2$  ?  ?</td>
</tr>
<tr>
<td>3</td>
<td>1        0        8        1</td>
<td>?  $Y_3$  ?  ?</td>
</tr>
<tr>
<td>4</td>
<td>0        1        8        3</td>
<td>?  $Y_3$  $Y_4$  ?</td>
</tr>
<tr>
<td>5</td>
<td>0        0        9        2</td>
<td>?  $Y_2$  ?  $Y_5$</td>
</tr>
<tr>
<td>6</td>
<td>0        0        10       1</td>
<td>?  $Y_2$  ?  $Y_6$</td>
</tr>
</tbody>
</table>

\[ \hat{\tau} = \frac{1}{6} [(Y_2 - Y_6) + (Y_2 - Y_5) + (Y_3 - Y_5) + (Y_3 - Y_5) + (Y_2 - Y_5) + (Y_2 - Y_6)] \]
1:1 matching, second stage

<table>
<thead>
<tr>
<th>Unit</th>
<th>Observed</th>
<th>Potential Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$A_i$</td>
<td>$Y_i(1,1)$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>$Y_1$</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>$?\rightarrow Y_2$</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>$?\rightarrow Y_3$</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>$?\rightarrow Y_3$</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>$?\rightarrow Y_2$</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>$?\rightarrow Y_2$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$M_i$</th>
<th>$Y_i(1,0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>$Y_1$</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>$?\rightarrow Y_2$</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>$?\rightarrow Y_3$</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>$Y_3$</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>$Y_2$</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>$Y_2$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$X_i$</th>
<th>$Y_i(0,1)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>$Y_6$</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>$?\rightarrow Y_5$</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>$?\rightarrow Y_5$</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>$Y_5$</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td>$Y_5$</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>$Y_6$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$Z_i$</th>
<th>$Y_i(0,0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>$Y_6$</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>$Y_5$</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>$Y_5$</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>$Y_5$</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>$Y_5$</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>$Y_6$</td>
</tr>
</tbody>
</table>
1:1 matching, second stage

<table>
<thead>
<tr>
<th>Unit</th>
<th>Observed</th>
<th>Potential Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$A_i$</td>
<td>$M_i$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

\[
\hat{\tau} = \frac{1}{6} \left[ (Y_2 - Y_6) + (Y_2 - Y_5) + (Y_3 - Y_5) \\
+ (Y_3 - Y_5) + (Y_2 - Y_5) + (Y_2 - Y_6) \right]
\]
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
• $\Rightarrow$ doesn’t converge to normal
• follow Abadie and Imbens (2011) and develop bias correction

Robust
• more robust to model misspecification
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
- $\Rightarrow$ doesn't converge to normal
- follow Abadie and Imbens (2011) and develop bias correction

Robust

- more robust to model misspecification
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
- analysis similar to Abadie & Imbens (2006)
- under regularity conditions, $\hat{\tau}$ converges to ACDE
- Bias converges to 0 very slowly
- $\Rightarrow$ doesn't converge to normal
- follow Abadie and Imbens (2011) and develop bias correction

Robust
- more robust to model misspecification
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

<table>
<thead>
<tr>
<th>Simple</th>
<th>Consistent</th>
<th>Biased</th>
</tr>
</thead>
</table>
| • two standard matching steps  
  • both can be done without $Y_i$ (avoid p-hacking) | • analysis similar to Abadie & Imbens (2006)  
  • under regularity conditions, $\hat{\tau}$ converges to ACDE | • Bias converges to 0 very slowly  
  • $\Rightarrow$ doesn't converge to normal  
  • follow Abadie and Imbens (2011) and develop bias correction |
## Properties of estimator

<table>
<thead>
<tr>
<th>Simple</th>
<th>Consistent</th>
<th>Biased</th>
</tr>
</thead>
</table>
| • two standard matching steps  
  • both can be done without $Y_i$ (avoid p-hacking) | • analysis similar to Abadie & Imbens (2006)  
  • under regularity conditions, $\hat{\tau}$ converges to ACDE | • 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  
|              | • $\sim$ doesn’t converge to normal |
## Properties of estimator

<table>
<thead>
<tr>
<th>Simple</th>
<th>Consistent</th>
<th>Biased</th>
</tr>
</thead>
</table>
| • two standard matching steps  
  • both can be done without $Y_i$ (avoid p-hacking) | • analysis similar to Abadie & Imbens (2006)  
  • under regularity conditions, $\hat{\tau}$ converges to ACDE | • Bias converges to 0 very slowly  
  • $\leadsto$ doesn’t converge to normal  
  • follow Abadie and Imbens (2011) and develop bias correction |
<table>
<thead>
<tr>
<th>Properties of estimator</th>
<th>Simple</th>
<th>Consistent</th>
<th>Biased</th>
<th>Robust</th>
</tr>
</thead>
</table>
|                        | • two standard matching steps  
  • both can be done without $Y_i$ (avoid p-hacking)  
|                        | • analysis similar to Abadie & Imbens (2006)  
  • under regularity conditions, $\hat{\tau}$ converges to ACDE  
|                        | • Bias converges to 0 very slowly  
  • $\Rightarrow$ 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**
- more robust to model misspecification
Matching and the bootstrap

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

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

Nonparametric bootstrap?

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

Weighted bootstrap?

- We follow Otsu and Rai (2017) and resample each contribution to the estimator.
Matching and the bootstrap

Variance of $\tau$ is *complicated*.

- Each $i$ could be matched multiple times at each stage.
- $\tau$ 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.
Matching and the bootstrap

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

- each $i$ could be matched multiple times at each stage

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
Matching and the bootstrap

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

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

Nonparametric bootstrap?

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

Weighted bootstrap

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

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

- each $i$ could be matched multiple times at each stage
- $\Rightarrow \hat{\tau}$ 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 $\hat{\tau}$ is *complicated*

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

Nonparametric bootstrap?

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

Weighted bootstrap
Matching and the bootstrap

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

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

Nonparametric bootstrap?

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

Weighted bootstrap

- we follow Otsu and Rai (2017) and resample each contribution to the estimator
Simulating misspecification
Simulation set-up

\[ X_1, X_2 \]

\[ A \rightarrow M \rightarrow Y \]

\[ Z \leftarrow \delta \rightarrow U \]

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

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

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

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

- Model misspecification as mismeasured confounders (Kang and Schafer, 2007)
  
  $X^*_1 = \exp\left(\frac{X_1}{2}\right)$
  
  $X^*_2 = \frac{1}{1 + \exp\left(\frac{X_2}{2}\right)} + 10$
  
  $Z^*_1 = \left(\frac{Z_1}{25} + 6\right)^3$

- 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 set-up

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

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 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

Bias (Absolute) vs. Magnitude of post-treatment confounding

Correct specification

Incorrect specification

Method
- Linear regression w/ mediator
- Sequential g-estimation
- Telescope matching

N = 2000
Simulation results: Bias

- **Correct specification**
- **Incorrect specification**

### Magnitude of post-treatment confounding

- **N=2000**

**Bias (Absolute)**

- **Method**
  - Linear regression w/ mediator
  - Sequential g-estimation
  - Telescope matching

---
Simulation results: Bias

Magnitude of post-treatment confounding

N=2000

Bias (Absolute)

Method
- Linear regression w/ mediator
- Sequential g-estimation
- Telescope matching
Simulation results: Root Mean Square Error

Method
- Linear regression w/ mediator
- Sequential g−estimation
- Telescope matching

Correct specification Incorrect specification
0.00 0.25 0.50 0.75 1.00 0.00 0.25 0.50 0.75 1.00

Magnitude of post−treatment confounding
N=2000
Simulation results: Root Mean Square Error

Correct specification

Incorrect specification

<table>
<thead>
<tr>
<th>Method</th>
<th>Correct specification</th>
<th>Incorrect specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear regression w/ mediator</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sequential g–estimation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Telescope matching</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Root Mean Square Error

Magnitude of post–treatment confounding

N=2000
Simulation results: Root Mean Square Error

![Graph showing Root Mean Square Error for different methods and specification accuracies. The x-axis represents the magnitude of post-treatment confounding, ranging from 0.00 to 1.00. The y-axis represents the root mean square error, ranging from 0.00 to 2.1. The graph is divided into two panels: Correct specification and Incorrect specification. The methods tested include Linear regression w/ mediator, Sequential g-estimation, and Telescope matching. The graph shows how the root mean square error changes with different magnitudes of confounding and specification accuracies.]
4/ Application
• Experiment on effect of media messages on support for immigration.
Experiment on effect of media messages on support for immigration.

Main effect: Story w/ negative tone + non-white immigrant reduced support for immigration.
• 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?
• 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.
• 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.
• 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.
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.
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.
Sequential g-estimation suggests a non-zero ACDE—there exists an effect even if we fix anxiety.
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.
- 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
• 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.
• 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
- https://www.antonstrezhnev.com/
SNMMs as imputation estimators

1. Estimate the conditional effect of $\mathcal{M}_i\gamma_m(x, z, a) = \mathbb{E}[Y_i(a, 1) - Y_i(a, 0)|X_i = x, A_i = a, Z_i = z, \mathcal{M}_i = 1]$

2. Impute $Y_i(a, 0)$ by subtracting effect of $\mathcal{M}_i$: $Y_i(a, 0) = Y_i - \mathcal{M}_i \times \hat{\gamma}_m(X_i, Z_i, A_i)$

3. Regress imputations on treatment and baseline covariates to get ACDE $\mathbb{E}[Y_i(A_i, 0)|X_i, A_i]$.

Depends on correct model for $\mathbb{E}[Y_i|X_i, Z_i, A_i, \mathcal{M}_i]$.
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 $M_i$

$$Y_i(A_i, 0) = Y_i - M_i \times \hat{\gamma}^m(X_i, Z_i, A_i)$$

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

$$\mathbb{E}[Y_i(A_i, 0) | X_i, A_i]$$

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

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

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

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 $M_i$

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

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

$$\mathbb{E}[\bar{Y}_i(A_i, 0)|X_i, A_i]$$
SNMMs as imputation estimators

1. Estimate the conditional effect of $M_i$

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

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

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

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

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

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

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

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

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

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

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

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

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

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