Advances in mediation and decomposition for research on aging, health, and place

Nick Graetz

Department of Sociology,

Minnesota Population Center / Institute for Social Research and Data Innovation

9/8/2025

MINNESOTA POPULATION CENTER

UNIVERSITY OF MINNESOTA
Driven to Discover®

Climate & Community INSTITUTE

Goals

- So, you're thinking about doing a mediation or decomposition analysis…

- **Goal:** provide a framework with lots of citations and software for thinking through modern solutions to complex mediation/decomposition issues – but also *real theoretical limitations.*

Outline

1. **Review:** causal mediation
2. **Estimands:** post-treatment confounding
3. **Estimation:** g-methods
4. **Example:** g-computation
5. **Extensions** and other approaches
6. **Critiques**

Historical Redlining and Contemporary Racial Disparities in Neighborhood Life Expectancy

Nick Graetz[1] and Michael Esposito[2]

[1]Department of Sociology, Princeton University
[2]Department of Sociology, Washington University in St. Louis

Abstract

While evidence suggests a durable relationship between redlining and population health, we currently lack an empirical account of how this historical act of racialized violence produced contemporary inequities. In this paper, we use a mediation framework to evaluate how redlining grades influenced later life expectancy and the degree to which contemporary racial disparities in life expectancy between Black working-class neighborhoods and White professional-class neighborhoods can be explained by past Home Owners' Loan Corporation (HOLC) mapping. Life expectancy gaps between differently graded tracts are driven by economic isolation and disparate property valuation which developed within these areas in subsequent decades. Still, only a small percent of a total disparity between contemporary Black and White neighborhoods is explained by HOLC grades. We discuss the role of HOLC maps in analyses of structural racism and health, positioning them as only one feature of a larger public–private project conflating race with financial risk. Policy implications include not only targeting resources to formerly redlined neighborhoods but also the larger project of dismantling racist theories of value that are deeply embedded in the political economy of place.

Review: causal mediation

Explaining differences

- A central aim in sociological and demographic analysis is <u>**explaining the source of differences.**</u>

$$E[Y|x] - E[Y|x^*]$$

- **This is a lot of what we try to do!**

Mediation with regression

- "What are the possible mechanisms connecting X and Y?"
- "It looks like there is an effect of X on Y, but it goes away when I control for M."
- "It looks like there is an effect of X on Y, but I *explained* 70% of it by controlling for M."

Causal mediation analysis

- **Why does X cause Y?**
  - X causes M; M causes Y
- **Why does Y vary across levels of X?**
  - Why do health disparities exist?

Mediation analysis vs. decomposition

- **Why begin with causal mediation analysis rather than decomposition?**
  - Most published papers I read and papers I review try to explain differences/disparities using regression models.
  - The author interprets the coefficient on **X** and then adds **M** and interprets something about how the coefficient changes (i.e., tries to *explain* differences).
- I will tie in connections to decomposition throughout, especially Kitagawa-Oaxaca-Blinder.

Lundberg et al. (2021). What is your estimand?

Theoretical Estimand	
Empirical Estimand	
Estimator	

Lundberg et al. (2021). What is your estimand?

Theoretical Estimand	
Empirical Estimand	$E[Y\|x] - E[Y\|x^*]$ Stuff we can observe
Estimator	

Lundberg et al. (2021). What is your estimand?

Theoretical Estimand	$E[Y_x] - E[Y_{x^*}]$	Potential outcomes
Empirical Estimand	$E[Y \mid x] - E[Y \mid x^*]$	
Estimator		

Lundberg et al. (2021). What is your estimand?

Theoretical Estimand	$E[Y_x] - E[Y_{x^*}]$
Empirical Estimand	$E[Y\|x] - E[Y\|x^*]$
Estimator	

This is causal inference

Lundberg et al. (2021). What is your estimand?

Theoretical Estimand	**Average treatment effect** $E[Y_x] - E[Y_{x^*}]$
Empirical Estimand	**Association** $E[Y\|x] - E[Y\|x^*]$
Estimator	

Lundberg et al. (2021). What is your estimand?

Theoretical Estimand	$E[Y_x] - E[Y_{x^*}]$
Empirical Estimand	$E[Y\|x] - E[Y\|x^*]$
Estimator	

What is the assumption?

Lundberg et al. (2021). What is your estimand?

Theoretical Estimand	$E[Y_x] - E[Y_{x^*}]$
Empirical Estimand	$E[Y \mid x] - E[Y \mid x^*]$
Estimator	

Are units **exchangeable**?

Are potential outcomes **independent** of treatment?

Lundberg et al. (2021). What is your estimand?

Theoretical Estimand	$E[Y_x] - E[Y_{x^*}]$
Empirical Estimand	$E[Y \vert x] - E[Y \vert x^*]$
Estimator	$E[Y \vert x] - E[Y \vert x^*]$ Difference in means

What is the effect of college on health?

Theoretical Estimand	$E[Y_x] - E[Y_{x^*}]$
Empirical Estimand	$E[Y \mid x] - E[Y \mid x^*]$
Estimator	

What is the effect of college on health?

Theoretical Estimand	$E[Y_x] - E[Y_{x^*}]$
Empirical Estimand	$E[Y\|x] - E[Y\|x^*]$ Selection
Estimator	

What is the effect of college on health?

Theoretical Estimand	$E[Y_x] - E[Y_{x^*}]$
Empirical Estimand	$E[Y\|x] - E[Y\|x^*]$ Find an instrument
Estimator	

What is the effect of college on health?

Theoretical Estimand	$E[Y_x] - E[Y_{x^*}]$
	$E[Y \mid x] - E[Y \mid x^*]$ **Find an instrument**

Mediation analysis with instrument-based methods is hard. See Imai, Tingley and Yamamoto (2013, *JRSS-A*).

What is the effect of college on health?

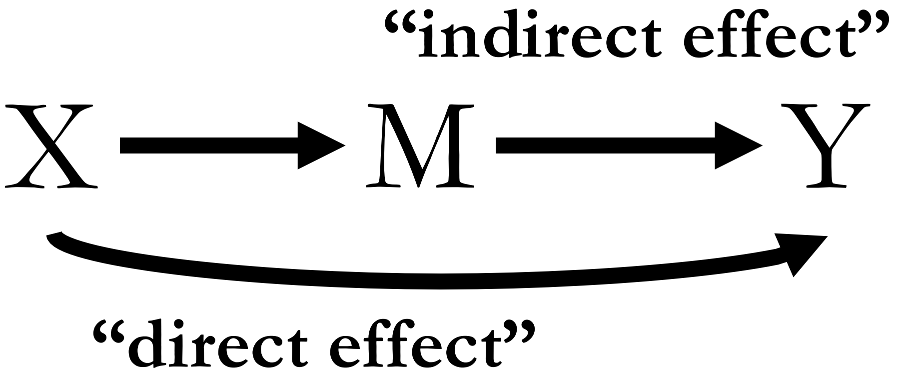
Theoretical Estimand	$E[Y_x] - E[Y_{x^*}]$
Empirical Estimand	$E[Y\|x, v] - E[Y\|x^*, v]$
Estimator	$Y = f(x, v)$ Regression adjustment Propensity weighting

$$X \longrightarrow Y$$

“indirect effect”

$$\begin{tikzcd} X \arrow[r] \arrow[rr, bend right] & M \arrow[r] & Y \end{tikzcd}$$

“direct effect”



Notice we are already committing to a theoretical framework where these "pathways" are **conceptually separable.**

Important when thinking about **complex exposures!**

$$\begin{tikzcd} X \arrow[r] \arrow[rr, bend right] & M \arrow[r] & Y \end{tikzcd}$$

$$Y = \beta_1 X$$
$$Y = \theta_1 X + \theta_2 M$$

$$\begin{tikzcd} X \arrow[r] \arrow[rr, bend right] & M \arrow[r] & Y \end{tikzcd}$$

$$Y = \beta_1 X$$
$$Y = \theta_1 X + \theta_2 M$$
$$\text{Proportion mediated} = (\beta_1 - \theta_1)/\beta_1$$

$$\begin{tikzcd}[column sep=huge] \text{Race} \arrow[r] \arrow[rr, bend right] & \text{Neighborhood} \arrow[r] & \text{Health} \end{tikzcd}$$

$$Y = 3(Race)$$
$$Y = 2(Race) + 6(Neighborhood)$$

$$
\begin{tikzcd}
\text{Race} \arrow[r] \arrow[rr, bend right] & \text{Neighborhood} \arrow[r] & \text{Health}
\end{tikzcd}
$$

$$Y = 3(Race)$$

$$Y = 2(Race) + 6(Neighborhood)$$

$$\text{Proportion mediated} = \frac{3-2}{3} = \frac{1}{3}$$

$$
\begin{tikzcd}
\text{Race} \arrow[r] \arrow[rr, bend right] & \text{Neighborhood} \arrow[r] & \text{Health}
\end{tikzcd}
$$

$$Y = 3(Race)$$

$$Y = 2(Race) + 6(Neighborhood)$$

$$\text{Proportion mediated} = \frac{3-2}{3} = \frac{1}{3}$$

"Neighborhood exposures *explain* one third of racial disparities in health."

$$
\begin{tikzcd}
\text{Race} \arrow[r] \arrow[rr, bend right] & \text{Neighborhood} \arrow[r] & \text{Health}
\end{tikzcd}
$$

$$Y = 3(Race)$$
$$Y = 2(Race) + 6(Neighborhood)$$
$$\text{Proportion mediated} = \frac{3-2}{3} = \frac{1}{3}$$

> "Neighborhood exposures *explain* one third of racial disparities in health."

Baron-Kenny (1986): 135,573 citations.

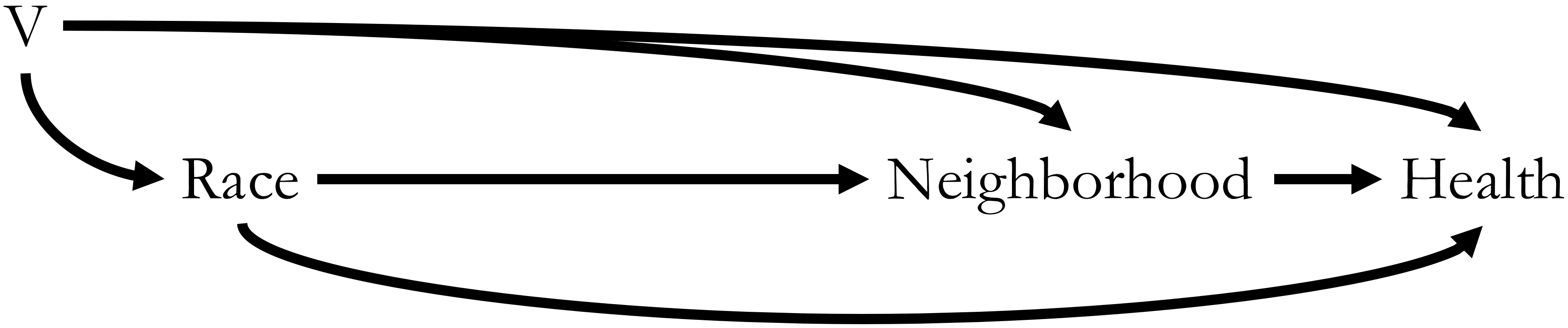
$$\begin{tikzcd} Race \arrow[rr] \arrow[rrr, bend right] & & Neighborhood \arrow[r] & Health \end{tikzcd}$$

$$Y = 3(Race)$$
$$Y = 2(Race) + 6(Neighborhood)$$
$$\text{Proportion mediated} = \frac{3-2}{3} = \frac{1}{3}$$

> "Neighborhood exposures *explain* one third of racial disparities in health."

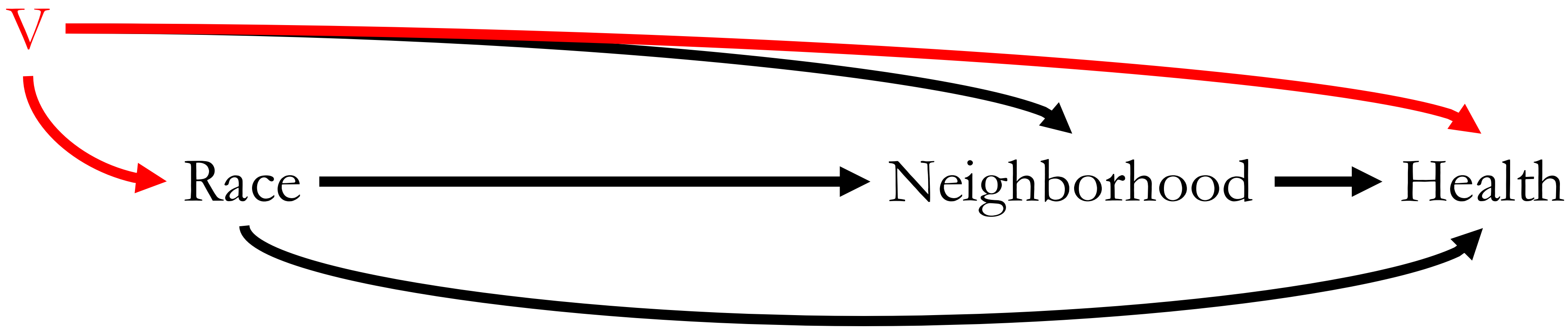
Baron-Kenny (1986): 135,573 citations.

When is this causal?

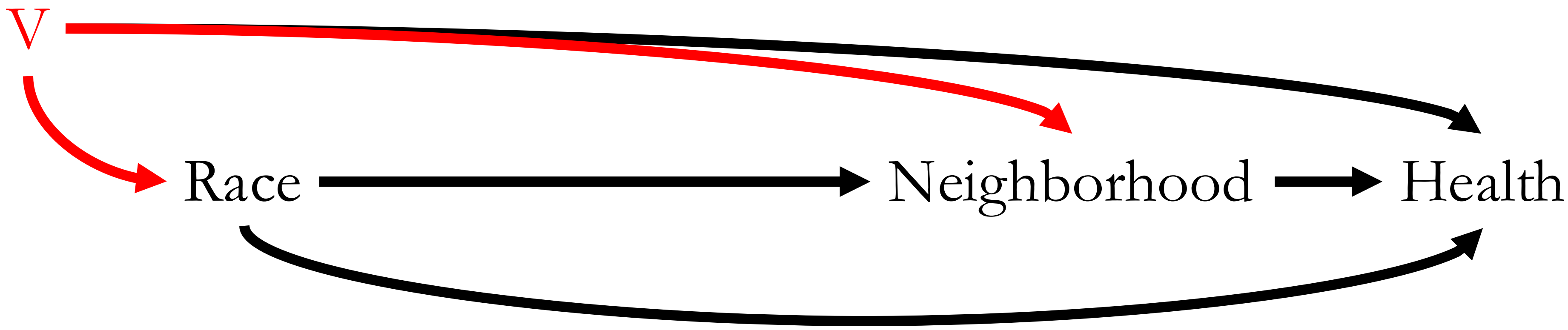


$$Y = 3(Race) + \boldsymbol{V}$$

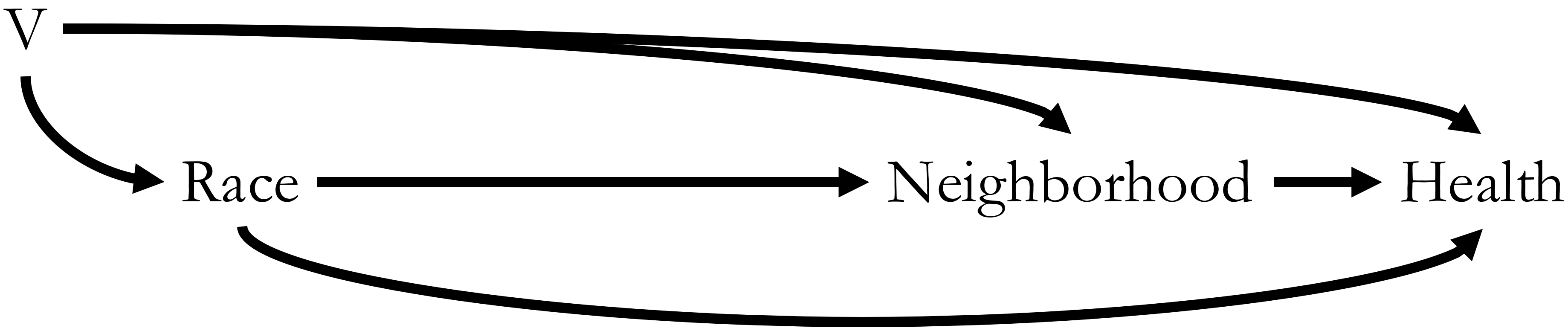
$$Y = 2(Race) + 6(Neighborhood) + \boldsymbol{V}$$



1. No unobserved confounding of X→Y (Race→Health).



1. No unobserved confounding of X→Y (Race→Health).
2. No unobserved confounding of X→M (Race→Neighborhood).



→ Many different mediation estimands can be considered within this simple DAG!

1. No unobserved confounding of X→Y (Race→Health).
2. No unobserved confounding of X→M (Race→Neighborhood).

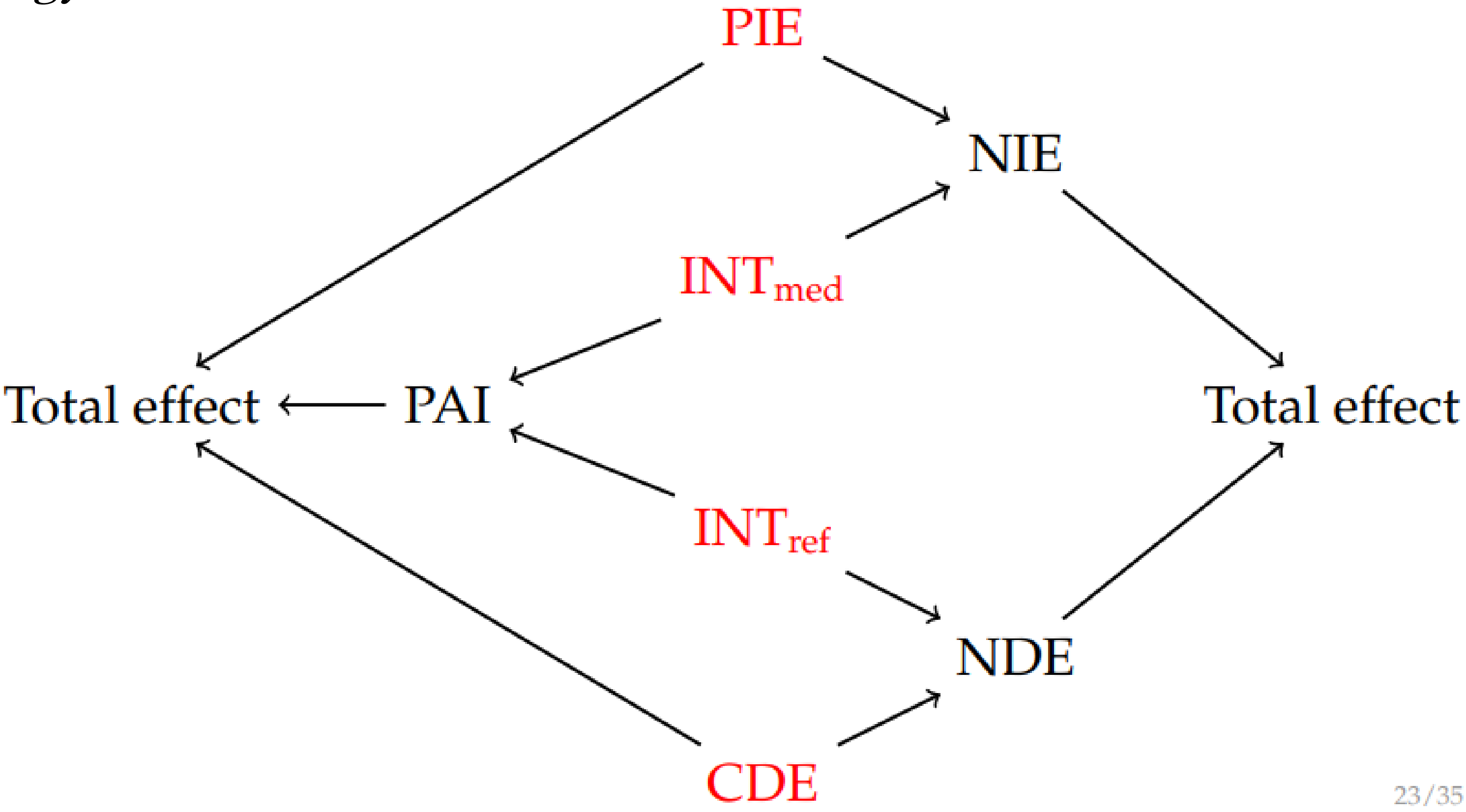
Wang & Arah (2014). "G-computation demonstration in causal mediation analysis." *European Journal of Epidemiology.*

Effect	Research question
TE[a]	Overall, to what extent does *X* cause *Y*?
PDE	In particular, to what extent does *X* cause *Y* via pathways other than through *M*?
TIE	In particular, to what extent does *X* cause *Y* via *M* (i.e. due to *X* affecting *M* and subsequently, *M* affecting *Y*) and the possible interaction between *X* and *M* in affecting *Y*? In other words, the effect of exposure that " would be prevented if the exposure did not cause the mediator" (i.e. the portion of the effect for which mediation is "necessary") [19,47].
TDE	In particular, to what extent does *X* cause *Y* via pathways other than through *M*, allowing *M* to boost up or tune down such effect at the same time?
PIE	In particular, to what extent does *X* cause *Y* via *M* only (i.e. due to *X* affecting *M* and subsequently, *M* affecting *Y*), not accounting for the possible interaction between *X* and *M*? In other words, the effect that the exposure would have had if "its only action were to cause the mediator" (i.e. the portion of the effect for which mediation is "sufficient") (i.e. the portion of [19,47].
CDE	What would be the effect of *X* on *Y*, when fixing *M* at a specific value for everyone in the population?
CDE_{sto}	What would be the effect of *X* on *Y*, when allowing *M* to attain certain controlled distribution (via intervention) in the population?
RIE	What would be the effect of *X* on *Y* that is due to interaction between *X* and *M* only?
MIE	What would be the effect of *X* on *Y* that is due to both interaction between *X* and *M* and the fact that *X* causes *M*?
PAI	What would be the effect of *X* on *Y* that is due to interaction between *X* and *M*, regardless whether *X* causes *M*?

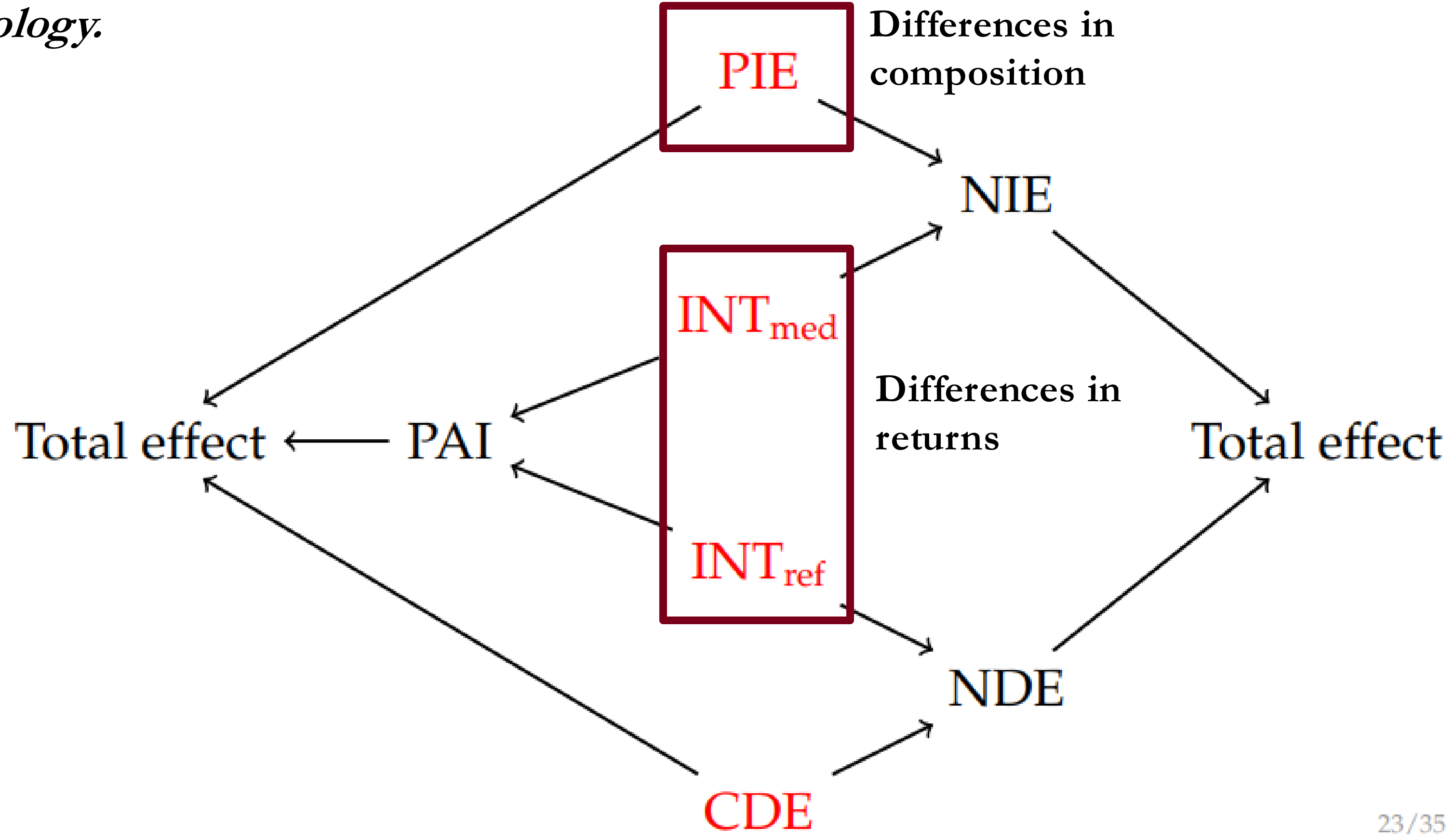
Wang & Arah (2014). "G-computation demonstration in causal mediation analysis." ***European Journal of Epidemiology.***

Effect	Counterfactual definition	Empirical analog[b]
TE[c]	$E[Y_x - Y_{x^*}]^d$	$\sum_z \sum_m \{E(Y\|x, m, z)\, P(m\|x, z) - E(Y\|x^*, m, z)\, P(m\|x^*, z)\}P(z)^e$
PDE	$E[Y_{xM_{x^*}} - Y_{x^*M_{x^*}}]$	$\sum_z \sum_m \{E(Y\|x, m, z) - (Y\|x^*, m, z)\}P(m\|x^*, z)\, P(z)^f$
TIE	$E[Y_{xM_x} - Y_{xM_{x^*}}]$	$\sum_z \sum_m E(Y\|x, m, z)\{P(m\|x, z) - P(m\|x^*, z)\}P(z)$
TDE	$E[Y_{xM_x} - Y_{x^*M_x}]$	$\sum_z \sum_m \{E(Y\|x, m, z) - (Y\|x^*, m, z)\}P(m\|x, z)\, P(z)$
PIE	$E[Y_{x^*M_x} - Y_{x^*M_{x^*}}]$	$\sum_z \sum_m E(Y\|x^*, m, z)\{P(m\|x, z) - P(m\|x^*, z)\}P(z)^f$
CDE$_{M=m^*}$	$E[Y_{xm^*} - Y_{x^*m^*}]$	$\sum_z \{E(Y\|x, m^*, z) - E(Y\|x^*, m^*, z)\}P(z)$
CDE$_{sto}$	$E[Y_{xM'} - Y_{x^*M'}]$	$\sum_z \sum_m \{E(Y\|x, m, z) - E(Y\|x^*, m, z)\}P(m')\, P(z)$
RIE	$E[(Y_{xm} - Y_{xm^*} - Y_{x^*m} + Y_{x^*m^*})(M_{x^*})]$	$\sum_z \sum_m \{E(Y\|x, m, z) - E(Y\|x, m^*, z) - E(Y\|x^*, m, z) + E(Y\|x^*, m^*, z)\}\, P(m\|x^*, z)\, P(z)$
MIE	$E[(Y_{xm} - Y_{xm^*} - Y_{x^*m} + Y_{x^*m^*})(M_x - M_{x^*})]$	$\sum_z \sum_m \{E(Y\|x, m, z) - E(Y\|x, m^*, z) - E(Y\|x^*, m, z) + E(Y\|x^*, m^*, z)\}\, \{P(m\|x, z) - P(m\|x^*, z)\}P(z)$
PAI	$E[(Y_{xm} - Y_{xm^*} - Y_{x^*m} + Y_{x^*m^*})(M_x)]$	$\sum_z \sum_m \{E(Y\|x, m, z) - E(Y\|x, m^*, z) - E(Y\|x^*, m, z) + E(Y\|x^*, m^*, z)\}\, P(m\|x, z)\, P(z)$

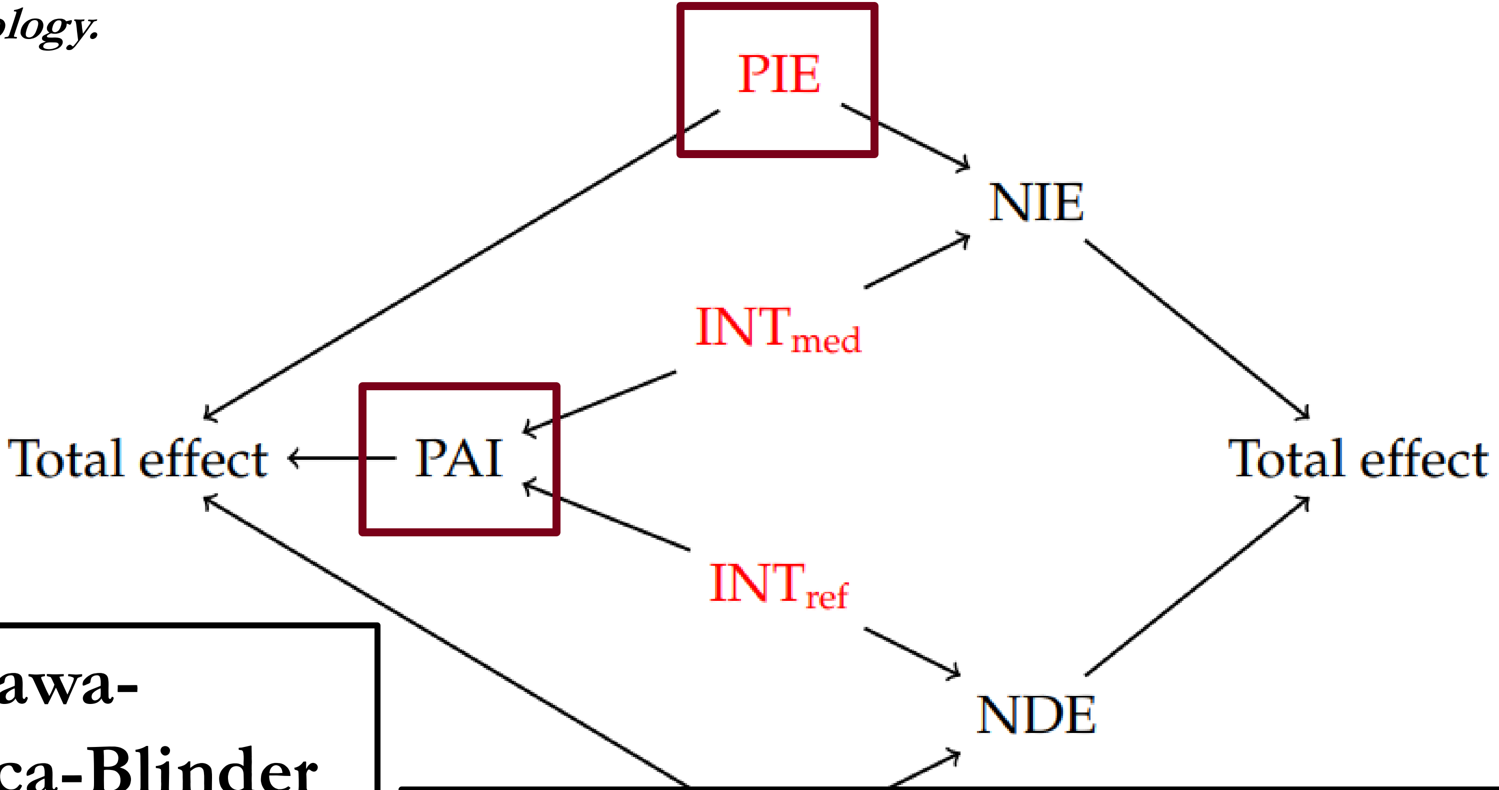
Vanderweele (2014). "A unification of mediation and interaction: A 4-way decomposition." *Epidemiology.*



Vanderweele (2014). "A unification of mediation and interaction: A 4-way decomposition." *Epidemiology.*



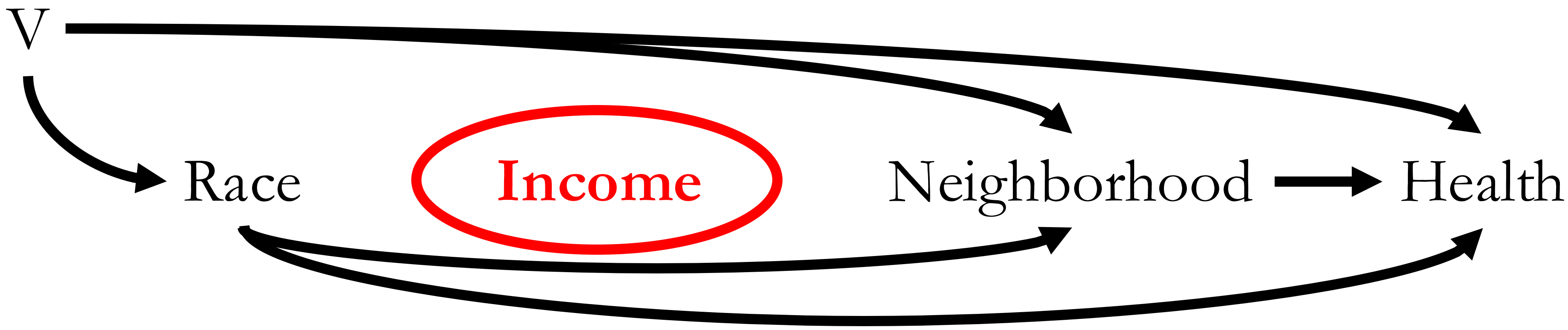
Vanderweele (2014). "A unification of mediation and interaction: A 4-way decomposition." *Epidemiology.*



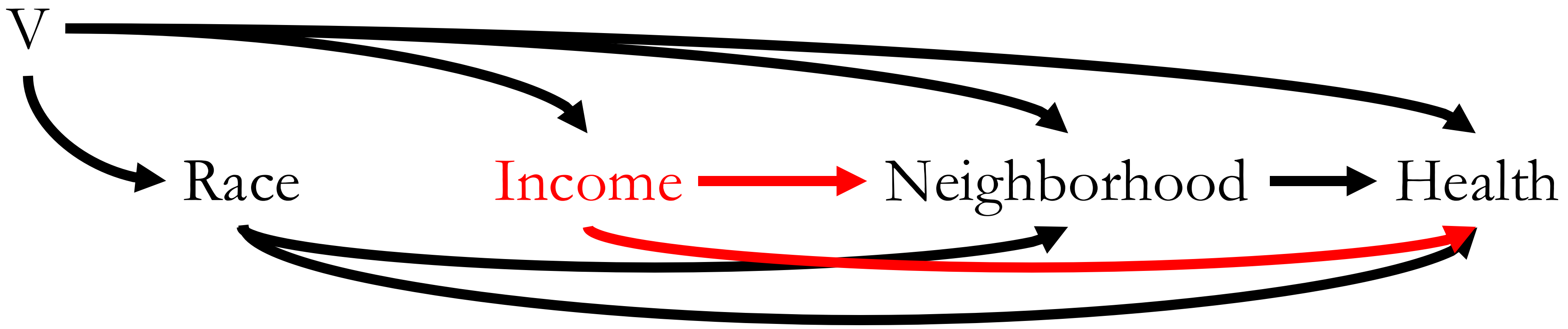
Kitagawa-Oaxaca-Blinder decomposition

Jackson & VanderWeele (2018). Decomposition analysis to identify intervention targets for reducing disparities. *Epidemiology.*

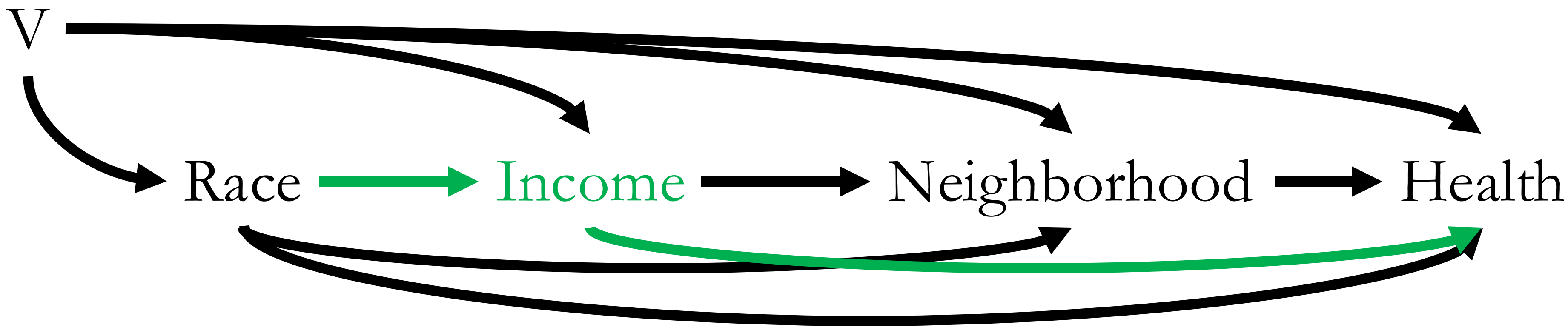
Estimands: post-treatment confounding



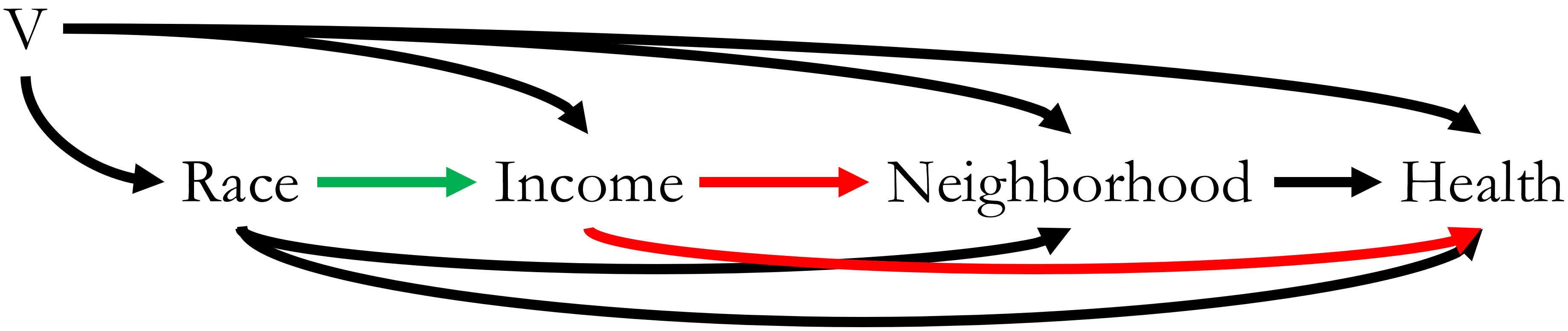
1. **No unobserved confounding of X→Y (Race→Health).**
2. **No unobserved confounding of X→M (Race→Neighborhood).**



1. No unobserved confounding of X→Y (Race→Health).
2. No unobserved confounding of X→M (Race→Neighborhood).
3. No unobserved post-treatment confounding of M→Y (Race→Neighborhood).

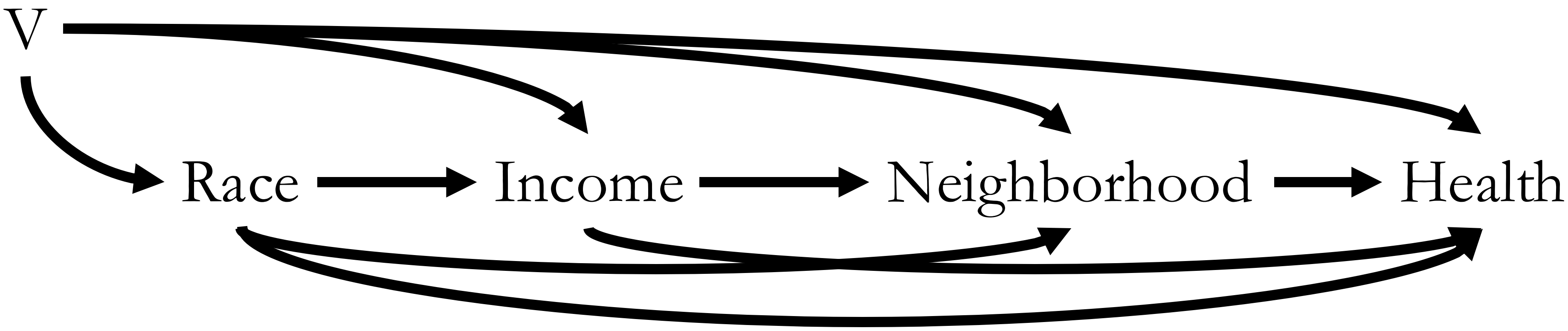


1. No unobserved confounding of X→Y (Race→Health).
2. No unobserved confounding of X→M (Race→Neighborhood).
3. No unobserved post-treatment confounding of M→Y (Race→Neighborhood).
4. No post-treatment confounders are affected by X.

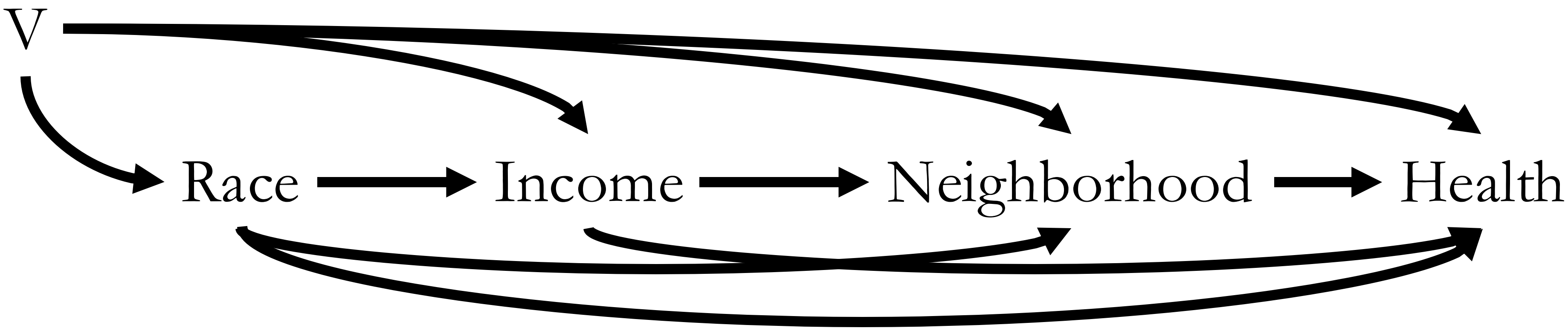


Should we control for Income?

1. No unobserved confounding of X→Y (Race→Health).
2. No unobserved confounding of X→M (Race→Neighborhood).
3. No unobserved post-treatment confounding of M→Y (Race→Neighborhood).
4. No post-treatment confounders are affected by X.



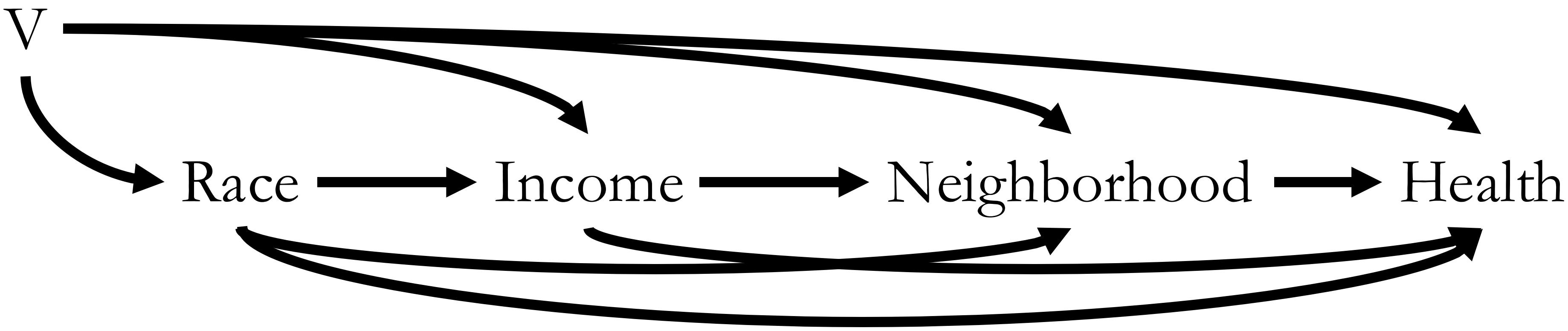
Baron-Kenny mediation **not controlling** for post-treatment confounders will produce an estimate **confounded** by those post-treatment confounders.



Baron-Kenny mediation **not controlling** for post-treatment confounders will produce an estimate **confounded** by those post-treatment confounders.

Baron-Kenny mediation **controlling** for post-treatment confounders will almost always **overestimate** the proportion mediated (i.e., underestimate the direct effect) because you're also inadvertently controlling for all mediating pathways through those post-treatment confounders.

Estimation: g-methods



"G-methods" can address theses issues for **observed post-treatment confounders.**

G-computation

→ Based on simulating

Marginal structural models

→ Based on weighting

Education Corner

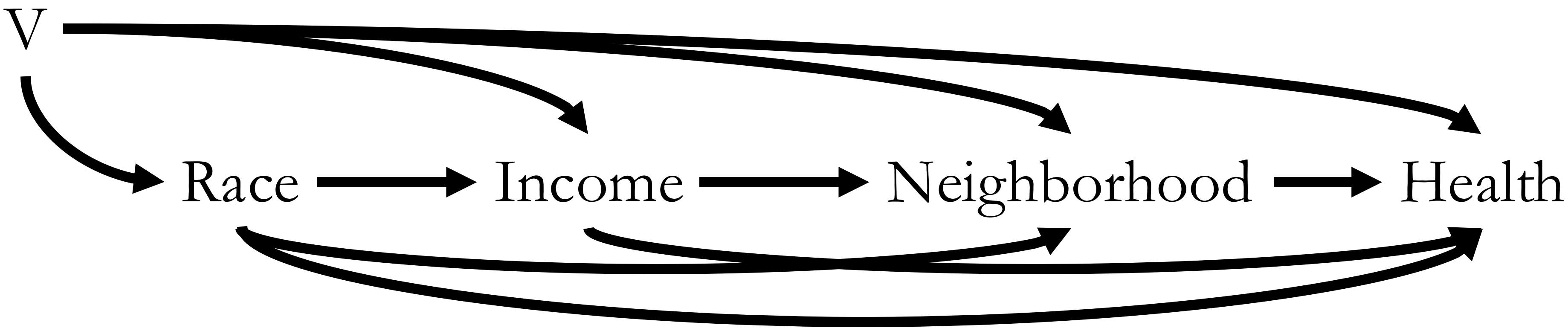
An introduction to g methods

Ashley I Naimi[1]*, Stephen R Cole[2] and Edward H Kennedy[3]

[1]Department of Epidemiology, University of Pittsburgh, [2]Department of Epidemiology, University of North Carolina at Chapel Hill and [3]Department of Statistics, Carnegie Mellon University

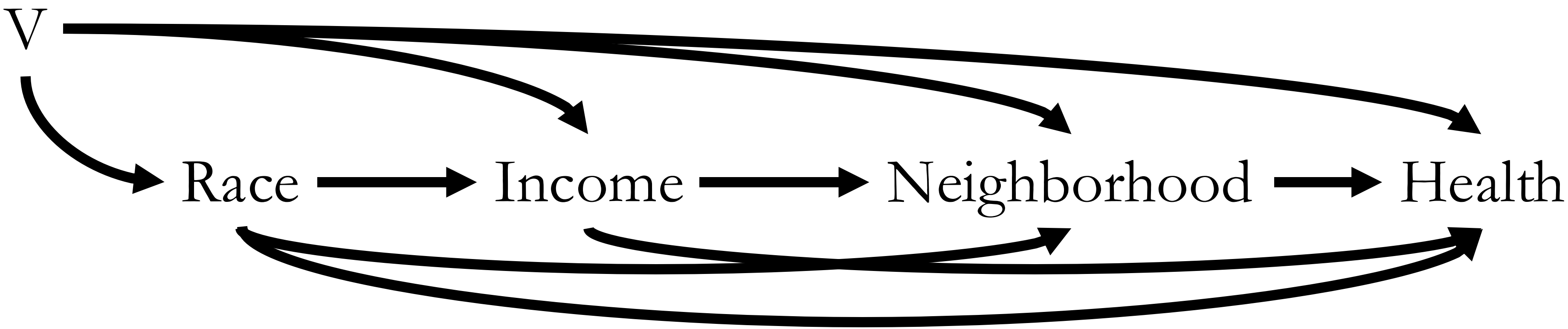
*Corresponding author. Department of Epidemiology University of Pittsburgh 130 DeSoto Street 503 Parran Hall Pittsburgh, PA 15261 ashley.naimi@pitt.edu

Accepted 17 October 2016



$$E[Y] = \sum_{m} \sum_{l} \{ P(y|x,m,l,v) \cdot P(m|x,l,v) \cdot P(l|x,v) \cdot P(v) \}$$

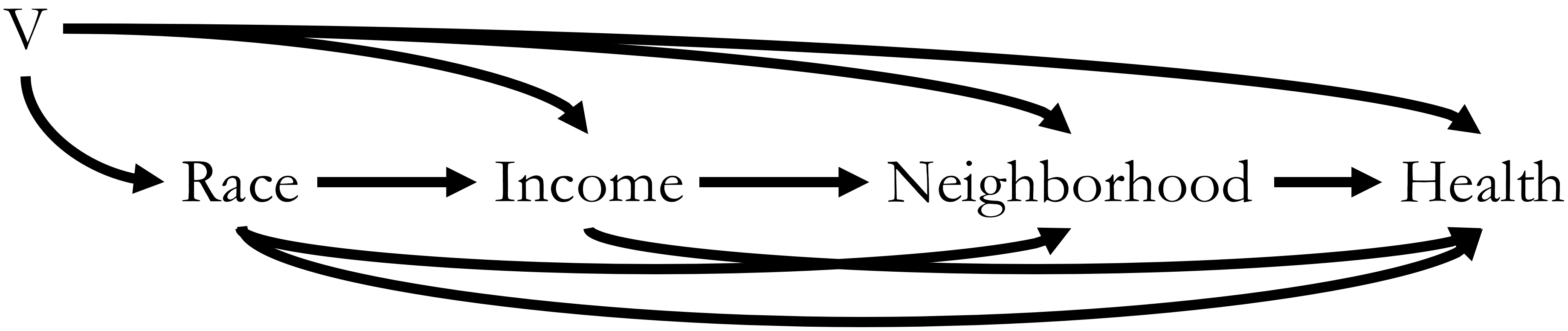
Step 1: Write out the expected value of Y in terms of each node in your DAG.



$$E[Y] = \sum_{m} \sum_{l} \{ P(y|x,m,l,v) \cdot P(m|x,l,v) \cdot P(l|x,v) \cdot P(v) \}$$

The g-formula

The *generalization* of standardization



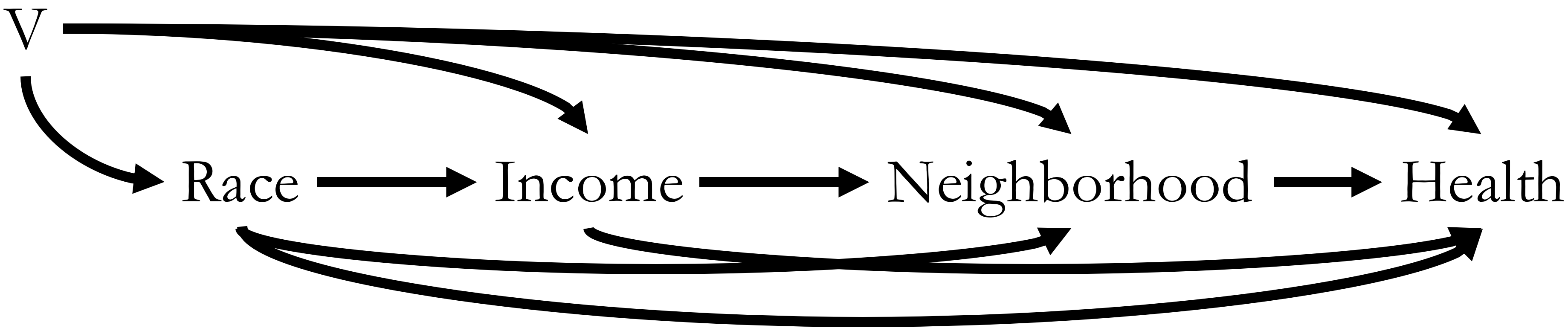
$$E[Y] = \sum_m \sum_l \{P(y|x,m,l,v) \cdot P(m|x,l,v) \cdot P(l|x,v) \cdot P(v)\}$$

The g-formula

The *generalization* of standardization

(Think about this like a multistate life table!)

Sudharsanan & Bijlsma (2021). "Educational note: causal decomposition of population health differences using Monte Carlo integration and the g-formula." *IJE*.



$$E[Y] = \sum_{m} \sum_{l} \{P(health|race, income, neighborhood, v) \cdot$$
$$P(neighborhood|race, income, v) \cdot$$
$$P(income|race, v) \cdot$$
$$P(v) \}$$

The parametric g-formula

We need model(s) to predict all post-treatment variables in our DAG.

$$E[Y] = \sum_m \sum_l \{P(y|x,m,l,v) \cdot$$

$$P(m|x,l,v) \cdot$$

$$P(l|x,v) \cdot$$

$$P(v)\}$$

$$\boldsymbol{Y} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{m}, \boldsymbol{l}, \boldsymbol{v})$$

$$\boldsymbol{M} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{l}, \boldsymbol{v})$$

$$\boldsymbol{L} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{v})$$

G-computation:

1. **Fit models.**

$$E[Y] = \sum_m \sum_l \{P(y|x,m,l,v) \cdot$$
$$P(m|x,l,v) \cdot$$
$$P(l|x,v) \cdot$$
$$P(v)\ \}$$

$$Y = f(x,m,l,v)$$
$$M = f(x,l,v)$$
$$L = f(x,v)$$

G-computation:

1. Fit models.
2. **Predict.**

$$E[Y] = \sum_m \sum_l \{P(y|x,m,l,v) \cdot$$

$$P(m|x,l,v) \cdot$$

$$P(l|x,v) \cdot$$

$$P(v)\}$$

$$Y = f(x,m,l,v)$$

$$M = f(x,l,v)$$

$$L = f(x,v)$$

G-computation:

1. Fit models.
2. Predict.
3. **Calculate target estimand.**

$$\text{PIE}^{(M)} = E\left[Y_{x^* L_{x^*} M_{xl}}\right] - E\left[Y_{x^* L_{x^*} M_{x^* l^*}}\right]$$

$$E[Y] = \sum_m \sum_l \{P(y|x,m,l,v) \cdot P(m|x,l,v) \cdot P(l|x,v) \cdot P(v)\}$$

$$Y = f(x,m,l,v)$$

$$M = f(x,l,v)$$

$$L = f(x,v)$$

G-computation:

1. **Bootstrap.**
2. Fit models.
3. Predict.
4. Calculate target estimand.

$$E[Y] = \sum_m \sum_l \{P(y|x,m,l,v) \cdot P(m|x,l,v) \cdot P(l|x,v) \cdot P(v)\}$$

$$Y = f(x,m,l,v)$$

$$M = f(x,l,v)$$

$$L = f(x,v)$$

G-computation:

1. Bootstrap.
2. Fit models.
3. Predict.
4. Calculate target estimand.

$$E[Y] = \sum_m \sum_l \{P(y|x,m,l,v) \cdot$$

$$P(m|x,l,v) \cdot$$

$$P(l|x,v) \cdot$$

$$P(v)\,\}$$

$$Y = f(x,m,l,v)$$

$$M = f(x,l,v)$$

$$L = f(x,v)$$

Simple example

Generalizable standardization

```
## OLS
ols <- lm(Sepal.Width~virginica, data=iris)
avg_slopes(ols, variables = 'virginica')


##
##  Estimate Std. Error     z Pr(>|z|)   S  2.5 % 97.5 %
##    -0.125      0.075 -1.67   0.0958 3.4 -0.272 0.0221
##
## Term: virginica
## Type: response
## Comparison: 1 - 0
```

Virginica → Width

Generalizable standardization

```r
## OLS
ols <- lm(Sepal.Width~virginica, data=iris)
avg_slopes(ols, variables = 'virginica')
```

```
##
##  Estimate Std. Error     z Pr(>|z|)   S  2.5 % 97.5 %
##    -0.125      0.075 -1.67   0.0958 3.4 -0.272 0.0221
##
## Term: virginica
## Type: response
## Comparison: 1 - 0
```

```r
## G-computation
## 1. Fit outcome model
y_model <- lm(Sepal.Width~virginica, data=iris)
## 2. Predict quantities needed for target estimand
counterfactuals <- as.data.table(avg_predictions(y_model,
                                                 by = "virginica",
                                                 variables = 'virginica'))
## 3. Calculate target estimand
estimand <- counterfactuals[virginica==1,estimate] - counterfactuals[virginica==0,estimate]
estimand
```

```
## [1] -0.125
```

Generalizable standardization

```r
## OLS
ols <- lm(Sepal.Width~virginica, data=iris)
avg_slopes(ols, variables = 'virginica')
```

```
##
##  Estimate Std. Error     z Pr(>|z|)   S  2.5 % 97.5 %
##    -0.125      0.075 -1.67   0.0958 3.4 -0.272 0.0221
##
## Term: virginica
## Type: response
## Comparison: 1 - 0
```

```r
## G-computation
## 1. Fit outcome model
y_model <- lm(Sepal.Width~virginica, data=iris)
## 2. Predict quantities needed for target estimand
counterfactuals <- as.data.table(avg_predictions(y_model,
                                                 by = "virginica",
                                                 variables = 'virginica'))
## 3. Calculate target estimand
estimand <- counterfactuals[virginica==1,estimate] - counterfactuals[virginica==0,estimate]
estimand
```

```
## [1] -0.125
```

$X \rightarrow Y$

DAG = how many models do I need?

Generalizable standardization

```
## OLS
ols <- lm(Sepal.Width~virginica, data=iris)
avg_slopes(ols, variables = 'virginica')
```

```
##
##  Estimate Std. Error     z Pr(>|z|)   S  2.5 % 97.5 %
##    -0.125      0.075 -1.67   0.0958 3.4 -0.272 0.0221
##
## Term: virginica
## Type: response
## Comparison: 1 - 0
```

```
## G-computation
## 1. Fit outcome model
y_model <- lm(Sepal.Width~virginica, data=iris)
## 2. Predict quantities needed for target estimand
counterfactuals <- as.data.table(avg_predictions(y_model,
                                                 by = "virginica",
                                                 variables = 'virginica'))
## 3. Calculate target estimand
estimand <- counterfactuals[virginica==1,estimate] - counterfactuals[virginica==0,estimate]
estimand
```

```
## [1] -0.125
```

X $\rightarrow$ Y

$E[Y|X=1]$

$E[Y|X=0]$

What quantities do I need to simulate?

Generalizable standardization

```
## OLS
ols <- lm(Sepal.Width~virginica, data=iris)
avg_slopes(ols, variables = 'virginica')


##
##  Estimate Std. Error     z Pr(>|z|)   S   2.5 % 97.5 %
##    -0.125      0.075 -1.67   0.0958 3.4 -0.272 0.0221
##
## Term: virginica
## Type: response
## Comparison: 1 - 0

## G-computation
## 1. Fit outcome model
y_model <- lm(Sepal.Width~virginica, data=iris)
## 2. Predict quantities needed for target estimand
counterfactuals <- as.data.table(avg_predictions(y_model,
                                                  by = "virginica",
                                                  variables = 'virginica'))
## 3. Calculate target estimand
estimand <- counterfactuals[virginica==1,estimate] - counterfactuals[virginica==0,estimate]
estimand

## [1] -0.125
```

$X \longrightarrow Y$

$E[Y|X=1]$

$E[Y|X=0]$

$E[Y|X=1] - E[Y|X=0]$

Simple arithmetic

Generalizable standardization

```r
## OLS
ols <- lm(Sepal.Width~virginica, data=iris)
avg_slopes(ols, variables = 'virginica')


##
##  Estimate Std. Error     z Pr(>|z|)   S  2.5 % 97.5 %
##    -0.125      0.075 -1.67   0.0958 3.4 -0.272 0.0221
##
## Term: virginica
## Type: response
## Comparison: 1 - 0

## G-computation
## 1. Fit outcome model
y_model <- lm(Sepal.Width~virginica, data=iris)
## 2. Predict quantities needed for target estimand
counterfactuals <- as.data.table(avg_predictions(y_model,
                                                 by = "virginica",
                                                 variables = 'virginica'))
## 3. Calculate target estimand
estimand <- counterfactuals[virginica==1,estimate] - counterfactuals[virginica==0,estimate]
estimand

## [1] -0.125
```

X → Y

$E[Y|X=1]$

$E[Y|X=0]$

$E[Y|X=1] - E[Y|X=0]$

Tutorial

Generalizable standardization

```
## OLS
ols <- lm(Sepal.Width~virginica, data=iris)
avg_slopes(ols, variables = 'virginica')


##
##  Estimate Std. Error     z Pr(>|z|)   S  2.5 % 97.5 %
##    -0.125      0.075 -1.67   0.0958 3.4 -0.272 0.0221
##
## Term: virginica
## Type: response
## Comparison: 1 - 0

## G-computation
## 1. Fit outcome model
y_model <- lm(Sepal.Width~virginica, data=iris)
## 2. Predict quantities needed for target estimand
counterfactuals <- as.data.table(avg_predictions(y_model,
                                                 by = "virginica",
                                                 variables = 'virginica'))
## 3. Calculate target estimand
estimand <- counterfactuals[virginica==1,estimate] - counterfactuals[virginica==0,estimate]
estimand


## [1] -0.125
```

Why use g-computation if I can get the same answer from a simple coefficient?

Generalizable standardization

```r
## OLS
ols <- lm(Sepal.Width~virginica, data=iris)
avg_slopes(ols, variables = 'virginica')
```

```
##
##  Estimate Std. Error     z Pr(>|z|)   S  2.5 % 97.5 %
##    -0.125      0.075 -1.67   0.0958 3.4 -0.272 0.0221
##
## Term: virginica
## Type: response
## Comparison: 1 - 0
```

```r
## G-computation
## 1. Fit outcome model
y_model <- lm(Sepal.Width~virginica, data=iris)
## 2. Predict quantities needed for target estimand
counterfactuals <- as.data.table(avg_predictions(y_model,
                                                 by = "virgi
                                                 variables =
## 3. Calculate target estimand
estimand <- counterfactuals[virginica==1,estimate] - counter
estimand
```

```
## [1] -0.125
```

1. **No relying on coefficients**
- Moving from *coefficient-based* inference to *estimand-based* inference: research question drives estimand, not model coefficient driving research question.

Generalizable standardization

```
## OLS
ols <- lm(Sepal.Width~virginica, data=iris)
avg_slopes(ols, variables = 'virginica')
```

```
##
##  Estimate Std. Error     z Pr(>|z|)   S  2.5 % 97.5 %
##    -0.125      0.075 -1.67   0.0958 3.4 -0.272 0.0221
##
## Term: virginica
## Type: response
## Comparison: 1 - 0
```

```
## G-computation
## 1. Fit outcome model
y_model <- lm(Sepal.Width~virginica, data=iris)
## 2. Predict quantities needed for target estimand
counterfactuals <- as.data.table(avg_predictions(y_model,
                                                 by = "virgi
                                                 variables =
## 3. Calculate target estimand
estimand <- counterfactuals[virginica==1,estimate] - counterfactuals[virginica==0,estimate]
estimand
```

```
## [1] -0.125
```

1. **No relying on coefficients**
2. **Isolate the prediction task**
- Not sensitive to model! Replace lm() with gbm(), etc.
- *marginaleffects* has incredibly flexible prediction functions.

Generalizable standardization

```r
## OLS
ols <- lm(Sepal.Width~virginica, data=iris)
avg_slopes(ols, variables = 'virginica')
```

```
##
##  Estimate Std. Error     z Pr(>|z|)   S  2.5 % 97.5 %
##    -0.125      0.075 -1.67   0.0958 3.4 -0.272 0.0221
##
## Term: virginica
## Type: response
## Comparison: 1 - 0
```

```r
## G-computation
## 1. Fit outcome model
y_model <- lm(Sepal.Width~virginica, data=iris)
## 2. Predict quantities needed for target estimand
counterfactuals <- as.data.table(avg_predictions(y_model,
                                                 by = "virginica",
                                                 variables = 'virginica'))
## 3. Calculate target estimand
estimand <- counterfactuals[virginica==1,estimate] - counterfactuals[virginica==0,estimate]
estimand
```

```
## [1] -0.125
```

1. No relying on coefficients
2. Isolate the prediction task
3. Generalizable to complex estimands

Generalizable standardization

```r
## OLS
ols <- lm(Sepal.Width~virginica, data=iris)
avg_slopes(ols, variables = 'virginica')
```

```
##
##  Estimate Std. Error     z Pr(>|z|)   S  2.5 % 97.5 %
##    -0.125      0.075 -1.67   0.0958 3.4 -0.272 0.0221
##
## Term: virginica
## Type: response
## Comparison: 1 - 0
```

```r
## G-computation
## 1. Fit outcome model
y_model <- lm(Sepal.Width~virginica, data=iris)
## 2. Predict quantities needed for target estimand
counterfactuals <- as.data.table(avg_predictions(y_model,
                                                 by = "virgi
                                                 variables =
## 3. Calculate target estimand
estimand <- counterfactuals[virginica==1,estimate] - counter
estimand
```

```
## [1] -0.125
```

More complicated DAG/estimand = more models, more predictions, more careful aggregation of predictions to get the quantities needed for estimand… but the framework is the same!

CMAverse: a suite of functions for causal mediation analysis

About the Package

The R package `CMAverse` provides a suite of functions for reproducible causal mediation analysis including `cmdag` for DAG visualization, `cmest` for statistical modeling and `cmsens` for sensitivity analysis.

See the package website for a quickstart guide, an overview of statistical modeling approaches and examples.

Cite the paper: CMAverse a suite of functions for reproducible causal mediation analyses

We welcome your feedback and questions:

- Email bs3141@columbia.edu for general questions
- Email zw2899@cumc.columbia.edu for questions related to `cmest_multistate`

CMAverse

Table 3: Causal Effects on the Difference Scale

Full Name	Abbreviation	Formula
Controlled Direct Effect	CDE	$E[Y_{am} - Y_{a^*m}]$
Randomized Analogue of $PNDE$	$rPNDE$	$E[Y_{aG_{a^*}} - Y_{a^*G_{a^*}}]$
Randomized Analogue of $TNDE$	$rTNDE$	$E[Y_{aG_a} - Y_{a^*G_a}]$
Randomized Analogue of $PNIE$	$rPNIE$	$E[Y_{a^*G_a} - Y_{a^*G_{a^*}}]$
Randomized Analogue of $TNIE$	$rTNIE$	$E[Y_{aG_a} - Y_{aG_{a^*}}]$
Total Effect	TE	$rPNDE + rTNIE$ or $rTNDE + rPNIE$
Randomized Analogue of INT_{ref}	$rINT_{ref}$	$rPNDE - CDE$
Randomized Analogue of INT_{med}	$rINT_{med}$	$rTNIE - rPNIE$
Proportion CDE	$prop^{CDE}$	CDE/TE
Proportion $rINT_{ref}$	$prop^{rINT_{ref}}$	$rINT_{ref}/TE$
Proportion $rINT_{med}$	$prop^{rINT_{med}}$	$rINT_{med}/TE$
Proportion $rPNIE$	$prop^{rPNIE}$	$rPNIE/TE$
Randomized Analogue of PM	rPM	$rTNIE/TE$
Randomized Analogue of INT	$rINT$	$(rINT_{ref} + rINT_{med})/TE$
Randomized Analogue of PE	rPE	$(rINT_{ref} + rINT_{med} + rPNIE)/TE$

Note:
a and a^* are the active and control values for A. m is the value at which M is controlled. G_a denotes a random draw from the distribution of M had $A = a$. Y_{am} denotes the counterfactual value of Y that would have been observed had A been set to be a, and M to be m. $Y_{aG_{a*}}$ denotes the counterfactual value of Y that would have been observed had A been set to be a, and M to be the counterfactual value G_{a*}.

The g-formula Approach

With the g-formula approach, `CMAverse` estimates causal effects through direct counterfactual imputation estimation by the following steps:

1. For $q = 1, \dots, s$, fit the regression model specified by `postcreg[q]` for the distribution of L_q given A and C.
2. For $q = 1, \dots, s$ and $i = 1, \dots, n$, simulate the counterfactuals $L_{a,q,i}$ and $L_{a^*,q,i}$ from the regression models in step 1.
   - Simulate $L_{a,q,i}$ by randomly drawing a value from the distribution of L_q given $A = a, C = C_i$. Denote $L_{a,i} = (L_{a,1,i}, \dots, L_{a,s,i})^T$.
   - Simulate $L_{a^*,q,i}$ by randomly drawing a value from the distribution of L_q given $A = a^*, C = C_i$. Denote $L_{a^*,i} = (L_{a^*,1,i}, \dots, L_{a^*,s,i})^T$.
3. For $p = 1, \dots, k$, fit the regression model specified by `mreg[p]` for the distribution of M_p given A, L and C.
4. For $p = 1, \dots, k$ and $i = 1, \dots, n$, simulate the counterfactuals $M_{a,p,i}$ and $M_{a^*,p,i}$ from the regression models in step 3.
   - Simulate $M_{a,p,i}$ by randomly drawing a value from the distribution of M_p given $A = a, L = L_{a,i}, C = C_i$. Denote $M_{a,i} = (M_{a,1,i}, \dots, M_{a,k,i})^T$.
   - Simulate $M_{a^*,p,i}$ by randomly drawing a value from the distribution of M_p given $A = a^*, L = L_{a^*,i}, C = C_i$. Denote $M_{a^*,i} = (M_{a^*,1,i}, \dots, M_{a^*,k,i})^T$.
5. Obtain $\{G_{a,i}\}_{i=1,\dots,n}$ by randomly permuting $\{M_{a,i}\}_{i=1,\dots,n}$ and obtain $\{G_{a^*,i}\}_{i=1,\dots,n}$ by randomly permuting $\{M_{a^*,i}\}_{i=1,\dots,n}$.
6. Fit the regression model specified by `yreg` for $E(Y|A, M, L, C)$.
7. For $i = 1, \dots, n$, obtain $E[Y_i|A = a^*, M = m, L = L_{a^*,i}, C = C_i]$, $E[Y_i|A = a, M = m, L = L_{a,i}, C = C_i]$, $E[Y_i|A = a^*, M = G_{a^*,i}, L = L_{a^*,i}, C = C_i]$, $E[Y_i|A = a^*, M = G_{a,i}, L = L_{a^*,i}, C = C_i]$, $E[Y_i|A = a, M = G_{a^*,i}, L = L_{a,i}, C = C_i]$ and $E[Y_i|A = a, M = G_{a,i}, L = L_{a,i}, C = C_i]$ from the regression model in step 5.
8. Impute the counterfactuals $E[Y_{a^*m}]$, $E[Y_{am}]$, $E[Y_{a^*Ga^*}]$, $E[Y_{aGa}]$, $E[Y_{aGa^*}]$ and $E[Y_{a^*Ga}]$.
   - Impute $E[Y_{a^*m}]$ by taking an average of $\{E[Y_i|A = a^*, M = m, L = L_{a^*,i}, C = C_i]\}_{i=1,\dots,n}$;
   - Impute $E[Y_{am}]$ by taking an average of $\{E[Y_i|A = a, M = m, L = L_{a,i}, C = C_i]\}_{i=1,\dots,n}$;
   - Impute $E[Y_{a^*Ga^*}]$ by taking an average of $\{E[Y_i|A = a^*, M = G_{a^*,i}, L = L_{a^*,i}, C = C_i]\}_{i=1,\dots,n}$;
   - Impute $E[Y_{aGa}]$ by taking an average of $\{E[Y_i|A = a, M = G_{a,i}, L = L_{a,i}, C = C_i]\}_{i=1,\dots,n}$;
   - Impute $E[Y_{aGa^*}]$ by taking an average of $\{E[Y_i|A = a, M = G_{a^*,i}, L = L_{a,i}, C = C_i]\}_{i=1,\dots,n}$;
   - Impute $E[Y_{a^*Ga}]$ by taking an average of $\{E[Y_i|A = a^*, M = G_{a,i}, L = L_{a^*,i}, C = C_i]\}_{i=1,\dots,n}$.
9. Calculate causal effects with formulas in table 3 or table 4.

KEY CITATIONS: G-computation with a single mediator

- VanderWeele, Vansteelandt, Robins (2013). "Effect decomposition in the presence of an exposure-induced mediator-outcome confounder." *Epidemiology*.
- Wang & Arah (2015). "G-computation demonstration in causal mediation analysis." *Eur J Epidemiol*.
- *marginaleffects* g-computation tutorials.
- *CMAverse* mediation tutorials.

KEY CITATIONS: Time-varying and multiple mediators

- VanderWeele, Tchetgen Tchetgen (2017). "Mediation analysis with time varying exposures and mediators." *J. R. Statist. Soc. B.*

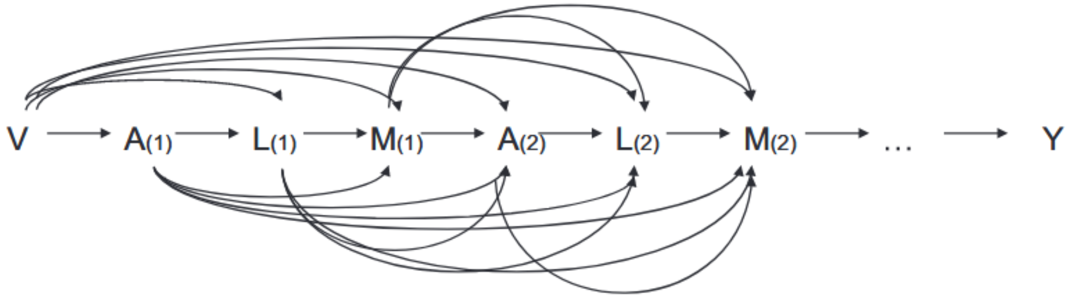


Fig. 5. Time varying mediation with variable ordering $A(t), L(t), M(t)$

$$E(Y_{\bar{a}\bar{G}_{\bar{a}^*|v}}|v) = \int_{\bar{m}} \int_{\bar{l}(T-1)} E(Y|\bar{a},\bar{m},\bar{l},v) \prod_{t=1}^{T-1} \mathrm{d}P\{\bar{l}(t)|\bar{a}(t),\bar{m}(t-1),\bar{l}(t-1),v\}$$

$$\times d\left[\int_{\bar{l}^\dagger(T-1)} \prod_{t=1}^{T} P\{m(t)|\bar{a}^*(t),\bar{m}(t-1),\bar{l}^\dagger(t),v\} \,\mathrm{d}P\{\bar{l}^\dagger(t)|\bar{a}^*(t),\bar{m}(t-1),\bar{l}^\dagger(t-1),v\}\right].$$

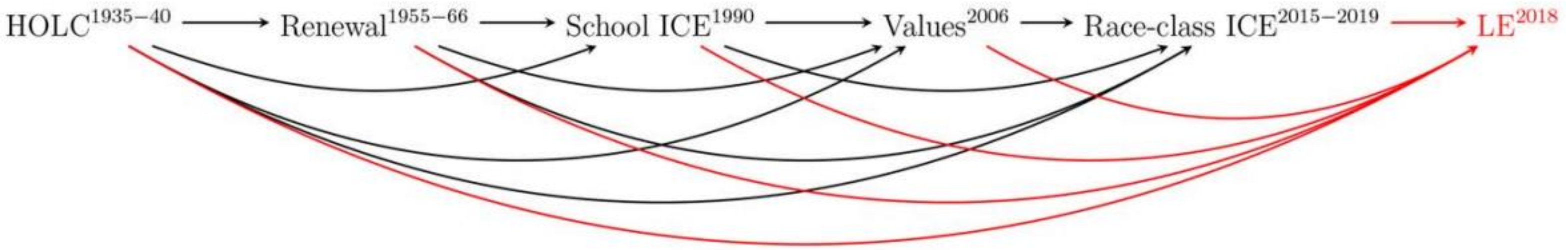
KEY CITATIONS: Comparisons to common decomposition methods

- Jackson, VanderWeele (2018). "Decomposition analysis to identify intervention targets for reducing disparities." *Epidemiology.*
- Sudharsanan, Bijlsma (2021). "Educational note: causal decomposition of population health differences using Monte Carlo integration and the g-formula." *International Journal of Epidemiology.*

Example: g-computation

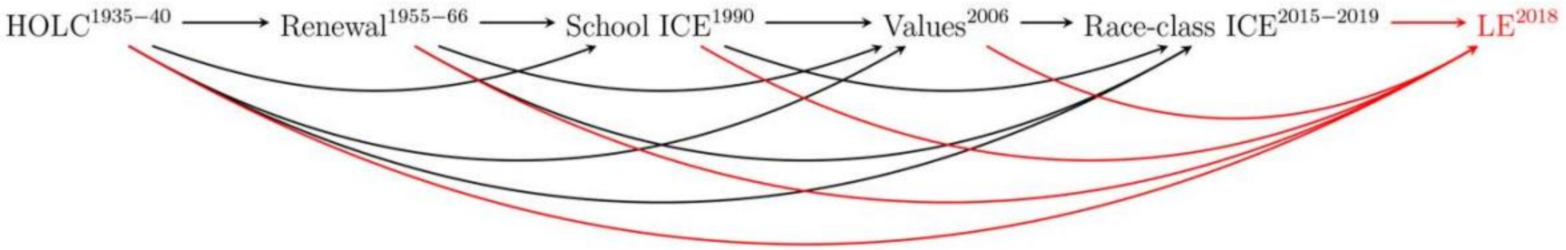
Step 1: Specify DAG

Graetz & Esposito (2023). Historical Redlining and Contemporary Racial Disparities in Neighborhood Life Expectancy. *Social Forces*.



Step 1: Specify DAG

Graetz & Esposito (2023). Historical Redlining and Contemporary Racial Disparities in Neighborhood Life Expectancy. *Social Forces.*



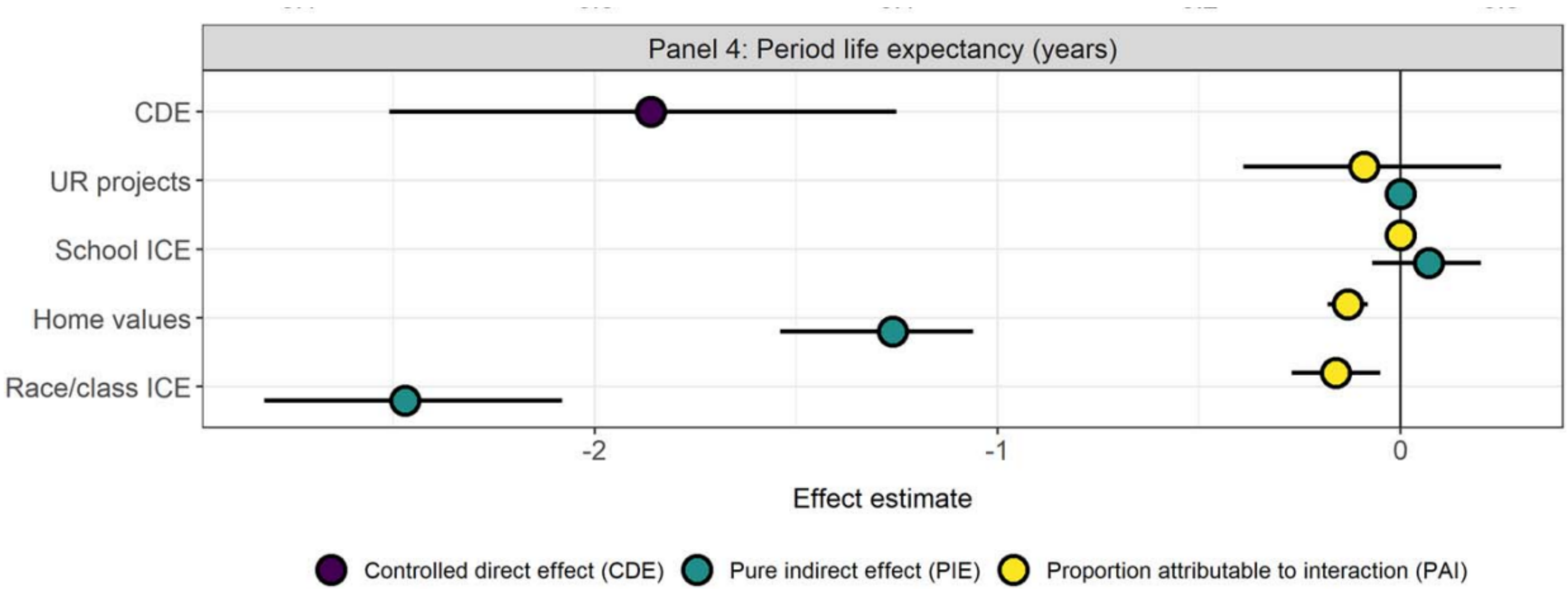
tc_vars	tv_vars	update_vars	family
V*X		M1	gaussian
V*X	M1*X	M2	gaussian
V*X	M1*X, M2*X	M3	gaussian
V*X	M1*X, M2*X, M3*X	M4	gaussian
V*X	M1*X, M2*X, M3*X, M4*X	Y	gaussian

G-computation:

1. Bootstrap.
2. Fit models.
3. Predict.
4. Calculate target estimand.

```
multmed_gformula(data=data,
                 tc_vars=tc_vars,
                 models=models,
                 sim_dir=sim_dir,
                 natural_courses=natural_courses,
                 intervention_rules=intervention_rules,
                 path_cb=path_cb,
                 treatment_var=treatment_var,
                 control_course=control_course,
                 total_sim_count=total_sim_count,
                 mc_replicates=mc_reps,
                 decomp_paths=decomp_paths,
                 parallelize=parallelize,
                 processors=processors,
                 windows=windows,
                 mean_betas=use_mean_betas,
                 decomp=do_decomp,
                 decomp_type=decomp_type,
                 dummy_vars=dummy_vars,
                 ordinal_levels=ordinal_levels,
                 ordinal_vars=ordinal_vars,
                 ordinal_refs=ordinal_refs,
                 duration_vars=duration_vars,
                 ever_vars=ever_vars,
                 cumcount_vars=cumcount_vars)
```

Step 3: Calculate estimands



Mediation 1

HOLC$^{1935-40}$ $\longrightarrow$ Renewal$^{1955-66}$ $\longrightarrow$ School ICE^{1990}

Mediation 2

HOLC$^{1935-40}$ $\longrightarrow$ Renewal$^{1955-66}$ $\longrightarrow$ School ICE^{1990} $\longrightarrow$ Values^{2006}

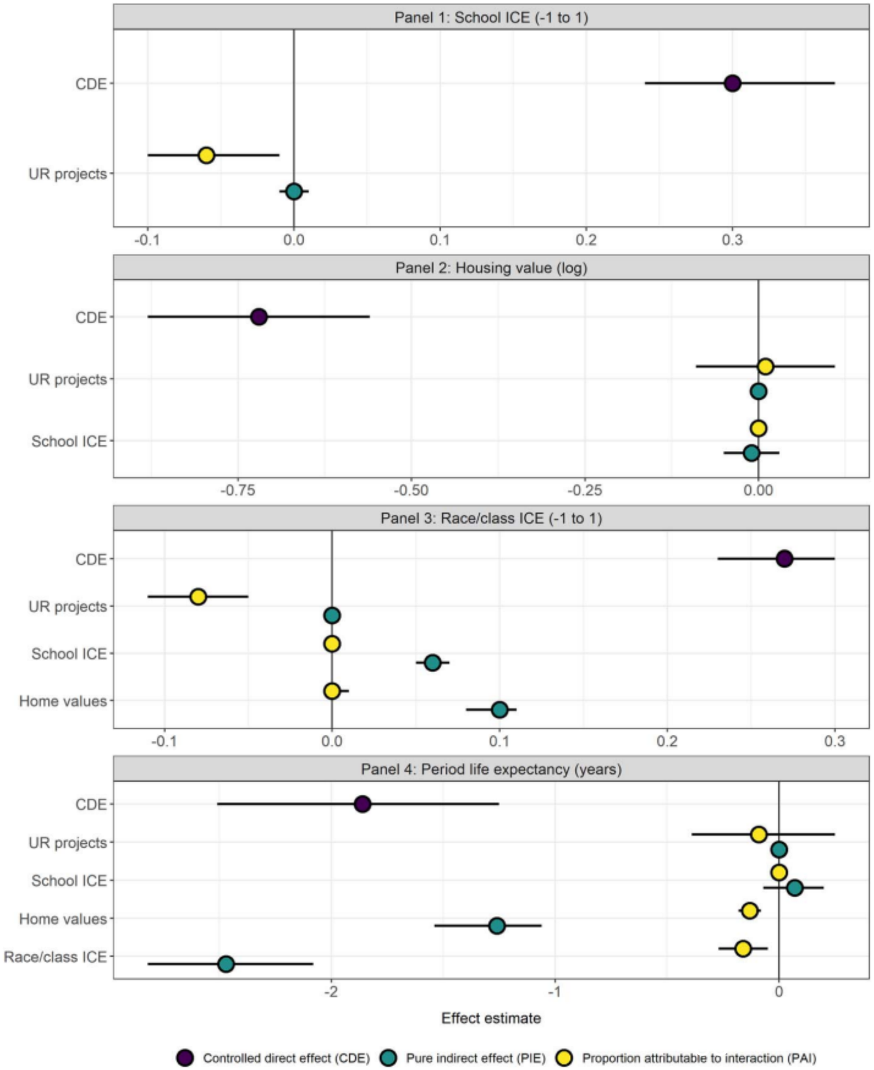
Mediation 3

HOLC$^{1935-40}$ $\longrightarrow$ Renewal$^{1955-66}$ $\longrightarrow$ School ICE^{1990} $\longrightarrow$ Values^{2006} $\longrightarrow$ Race-class ICE$^{2015-2019}$

Mediation 4

HOLC$^{1935-40}$ $\longrightarrow$ Renewal$^{1955-66}$ $\longrightarrow$ School ICE^{1990} $\longrightarrow$ Values^{2006} $\longrightarrow$ Race-class ICE$^{2015-2019}$ $\longrightarrow$ LE^{2018}

- A simple ATE of redlining is hard to interpret because of very different neighborhood trajectories.
- Urban renewal: many redlined neighborhoods look extremely different today.
- Counterfactual simulations on contemporary disparities.



G-computation software with multiple mediators

- We tried to code this up in a general format:
  - https://github.com/ngraetz/multmed_gcomp

- Cutting-edge:
  - Zhou & Wodtke (2025). "Causal mediation analysis with multiple mediators: A simulation approach." Working paper: https://arxiv.org/abs/2506.14019.

Extensions and other approaches

Path-specific effects

- Zhou, Yamamoto (2023). "Tracing causal paths from experimental and observational data." *The Journal of Politics.*
  - [R: ***paths***]

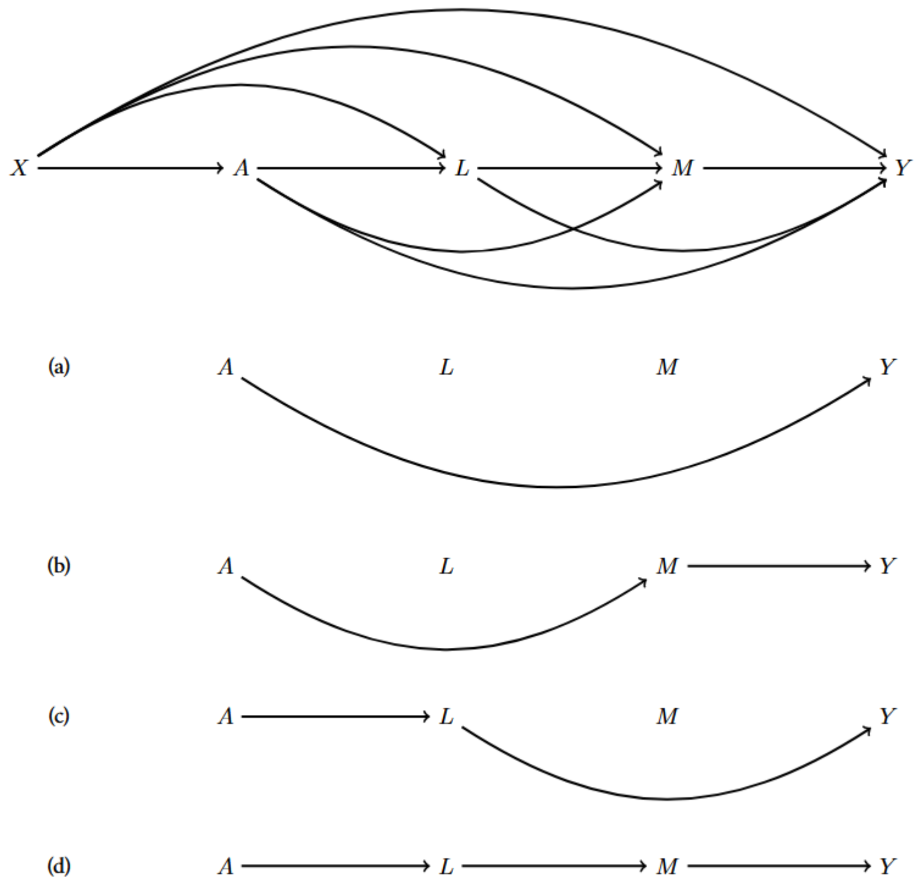


Figure 1: Causal Relationships with Two Causally Dependent Mediators Shown in a Directed Acyclic Graph (DAG).

Note: A denotes the treatment, Y denotes the outcome, X denotes pretreatment confounders, and L and M denote two causally dependent mediators.

Non-parametric decomposition

- Bohren, Hull, Imas (2022). "Systemic discrimination: Theory and measurement." *NBER*.

- Yu & Elwert (2024). "Nonparametric causal decomposition of group disparities." Working paper: https://arxiv.org/abs/2306.16591.
  - **[R: *cdgd*]**

Critiques

Critiques to potential outcomes: Inside the house

- **The cross-world independence assumption.**
- Andrews, Didelez (2021). "Insights into the cross-world independence assumption of causal mediation analysis." *Epidemiology*.

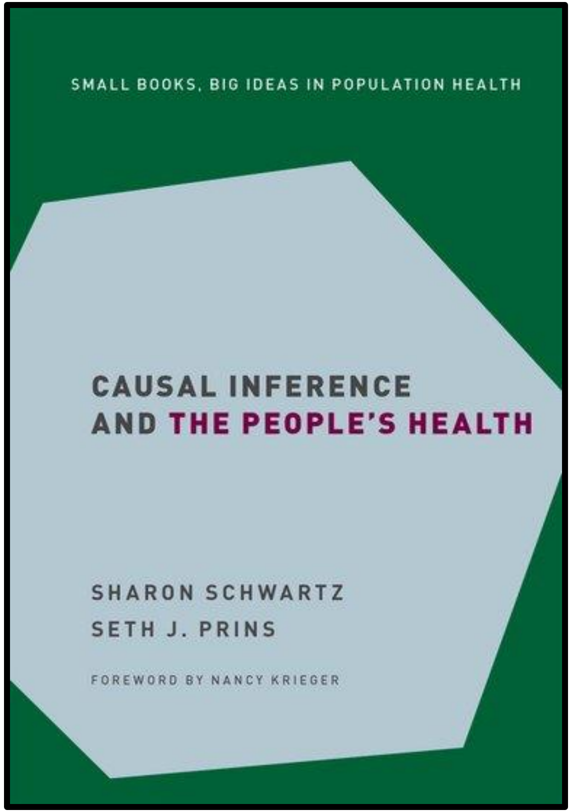
In the language of counterfactuals, the cross-world independence assumption is that

$$Y(a,m) \perp\!\!\!\perp M(a') \quad \forall m, \tag{1}$$

where the counterfactual $Y(a,m)$ is the value of the outcome Y that would be observed if, possibly counter to fact, exposure A were set to $A=a$ and the mediator M were set to m, and where $M(a')$ is the value of M under the assignment $A=a'$, with possibly $a' \neq a$ (for our purposes, we assume that $a \neq a'$). In words, this assumption is that there is an independence between counterfactual outcome and mediator values "across worlds," with one being a world in which the exposure is set to $A=a$ for the outcome and the other being a world in which it is set to $A=a'$ for the mediator. Such an exposure assignment cannot occur in real life, making the cross-world independence assumption impossible to verify, even in principle, without relying on other equally problematic assumptions.

Critiques to potential outcomes: Outside the house

- Do counterfactuals have to be defined with potential outcomes? What do we lose?
- Long philosophical history of counterfactual reasoning outside modern potential outcomes framework.



Critique #1: Are pathways separable?

- SUTVA problems

ARTICLE

Special Issue on New Perspectives on Empirical Methods and Critical Race Theory

What is perceived when race is perceived and why it matters for causal inference and discrimination studies

Lily Hu[1] and Issa Kohler-Hausmann[1,2]

[1]Department of Philosophy, Yale University, New Haven, CT, USA and [2]Yale Law School, Yale University, New Haven, CT, USA

Corresponding author: Lily Hu; Email: lily.hu@yale.edu

Critique #1: Are pathways separable?

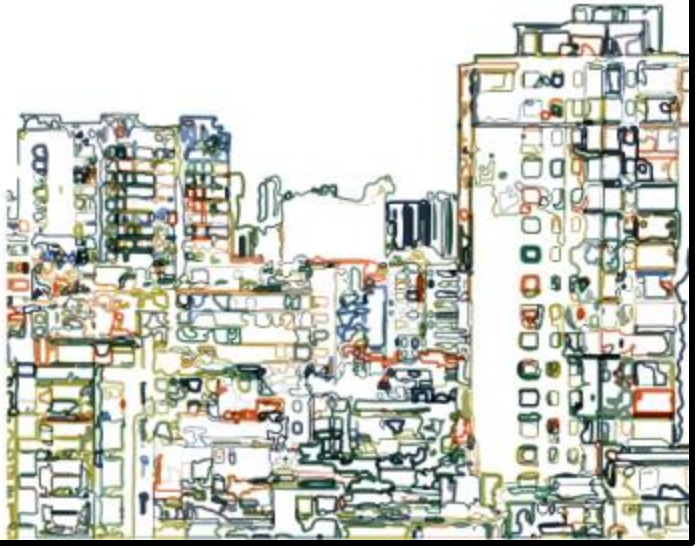
University of Michigan Institute for Social Research presents

Complexity in the Social World:
The Challenging Case of Structural Racism

October 9-10, 2025
Ann Arbor, Michigan

Registration opens August 2025

Travel awards available for early career scholars



Sponsored by:

ISR INSTITUTE FOR SOCIAL RESEARCH UNIVERSITY OF MICHIGAN

ISR SURVEY RESEARCH CENTER UNIVERSITY OF MICHIGAN

Critique #2: Focus more on measurement

- Including a lot of intermediate characteristics that are essentially *outcomes* is not the same as illuminating *processes.* Sometimes we need to study "…processes rather than processed people." (Desmond 2014).

Critique #2: Focus more on measurement

- Including a lot of intermediate characteristics that are essentially *outcomes* is not the same as illuminating *processes*. Sometimes we need to study "…processes rather than processed people." (Desmond 2014).
  - Krieger (1994). "Epidemiology and the web of causation: Has anyone seen the spider?" *Social Science & Medicine*.
  - Muntaner (2013). "Invited commentary: On the future of social epidemiology--a case for scientific realism." *American Journal of Epidemiology*.

Critique #2: Focus more on measurement

- Reynolds (2021). "Health power resources theory: A relational approach to the study of health inequalities." *Journal of Health and Social Behavior.*
- Creary (2021). "Bounded justice and the limits of health equity." *The Journal of Law, Medicine & Ethics.*
- Michener (2022). "Health justice through the lens of power." *Journal of Law, Medicine & Ethics.*

Concluding remarks

- Baron-Kenny mediation:
  - Makes a **lot of assumptions.**
  - Easy to be wrong (e.g., I explained 70% of the disparity/effect.)

Concluding remarks

- Baron-Kenny mediation:
  - Makes a **lot of assumptions.**
  - Easy to be wrong (e.g., I explained 70% of the disparity/effect.)
- G-computation:
  - Makes **fewer assumptions.**
  - Easy to be wrong in new ways that are not as obvious.

Concluding remarks

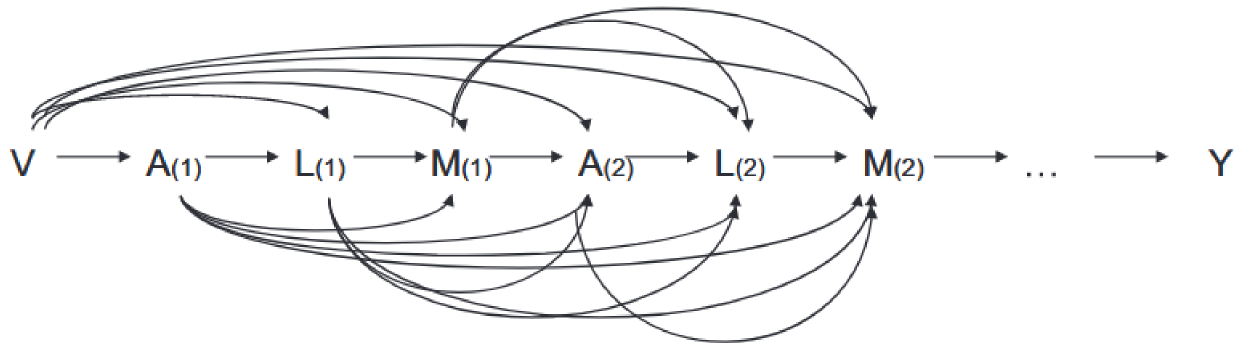


Fig. 5. Time varying mediation with variable ordering $A(t), L(t), M(t)$

$$E(Y_{\bar{a}\bar{G}_{\bar{a}^*|v}}|v) = \int_{\bar{m}}\int_{\bar{l}(T-1)} E(Y|\bar{a},\bar{m},\bar{l},v)\prod_{t=1}^{T-1} \mathrm{d}P\{\bar{l}(t)|\bar{a}(t),\bar{m}(t-1),\bar{l}(t-1),v\}$$
$$\times d\left[\int_{\bar{l}^{\dagger}(T-1)}\prod_{t=1}^{T} P\{m(t)|\bar{a}^*(t),\bar{m}(t-1),\bar{l}^{\dagger}(t),v\}\,\mathrm{d}P\{\bar{l}^{\dagger}(t)|\bar{a}^*(t),\bar{m}(t-1),\bar{l}^{\dagger}(t-1),v\}\right].$$

Concluding remarks

- Baron-Kenny mediation:
  – Makes a **lot of assumptions.**
  – Easy to be wrong (e.g., I *explained* 70% of the disparity/effect.)
- G-computation:
  – Makes **fewer assumptions.**
  – Easy to be wrong in new ways that are not as obvious.
- **Humility and interdisciplinary perspective is critical.**
  – We can't escape theory and deep contextual understanding of mechanisms, which often comes from qual/legal/historical work.

Email: ngraetz@umn.edu

MINNESOTA POPULATION CENTER

UNIVERSITY OF MINNESOTA
Driven to Discover®

Climate & Community
INSTITUTE