The **gap-closing estimand**

Slides at
ianlundberg.org

Software at
ilundberg.github.io/gapclosing

A causal approach to study **interventions** that **close disparities** across social categories

Ian Lundberg

Cornell University
Department of Information Science
ilundberg@cornell.edu

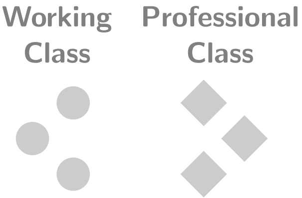
22 June 2023
Summer Institute in Computational Social Science at UCLA

Paper in *Sociological Methods and Research*. Replication code on Dataverse. R package `gapclosing` on CRAN.
Research reported in this publication was supported by The Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Number P2CHD047879 and by the National Science Foundation under Award Number 2104607.

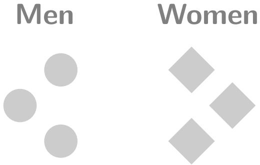
The Gap-Closing Estimand:

A Causal Approach to Study Interventions
That Close Disparities Across Social Categories

Ian Lundberg (Cornell)

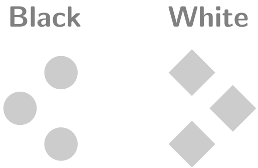


The Gap-Closing Estimand:
A Causal Approach to Study Interventions
That Close Disparities Across Social Categories



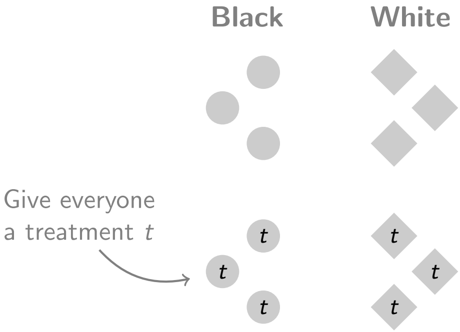
The Gap-Closing Estimand:
A Causal Approach to Study Interventions
That Close Disparities Across Social Categories

Ian Lundberg (Cornell)



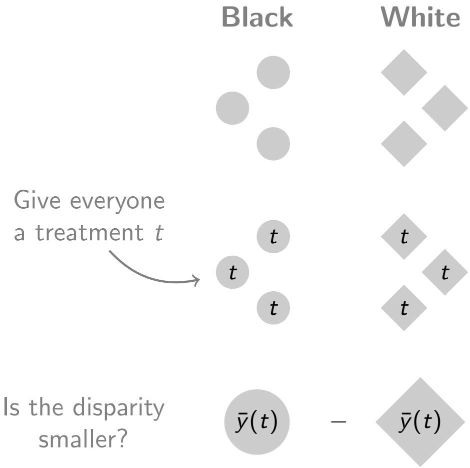
The Gap-Closing Estimand:
A Causal Approach to Study Interventions
That Close Disparities Across Social Categories

Ian Lundberg (Cornell)

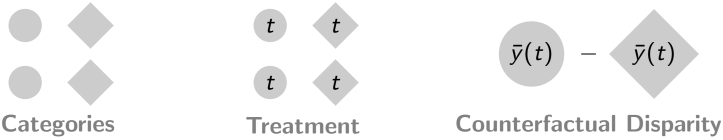


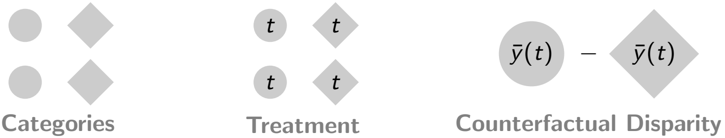
The Gap-Closing Estimand:

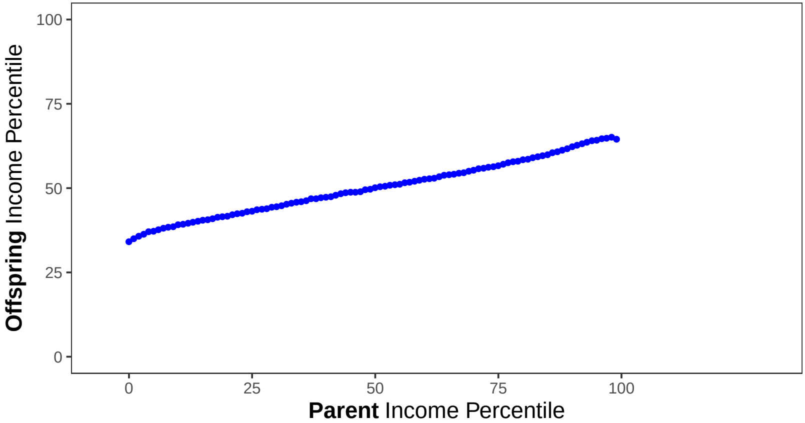
A Causal Approach to Study Interventions
That Close Disparities Across Social Categories



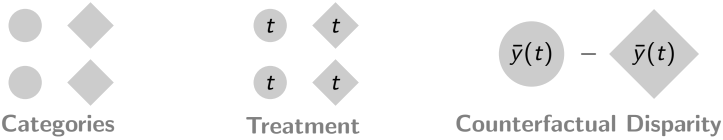
The Gap-Closing Estimand:
A Causal Approach to Study Interventions
That Close Disparities Across Social Categories

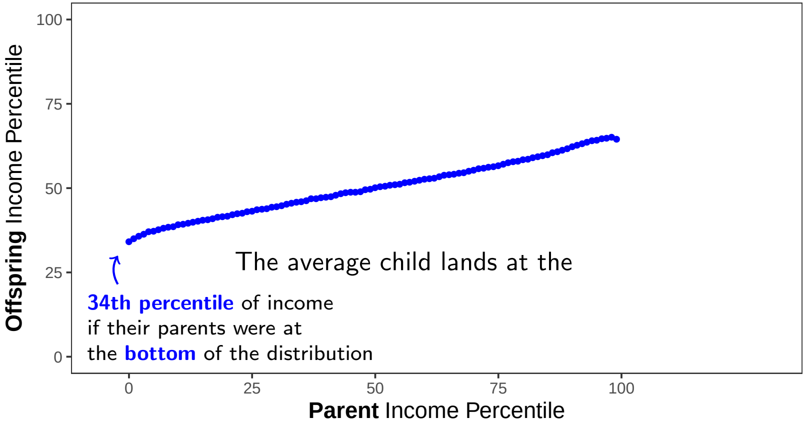

t
t
t
t
$\bar{y}(t)$
$-$
$\bar{y}(t)$
Categories
Treatment
Counterfactual Disparity



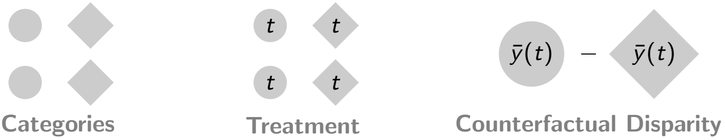


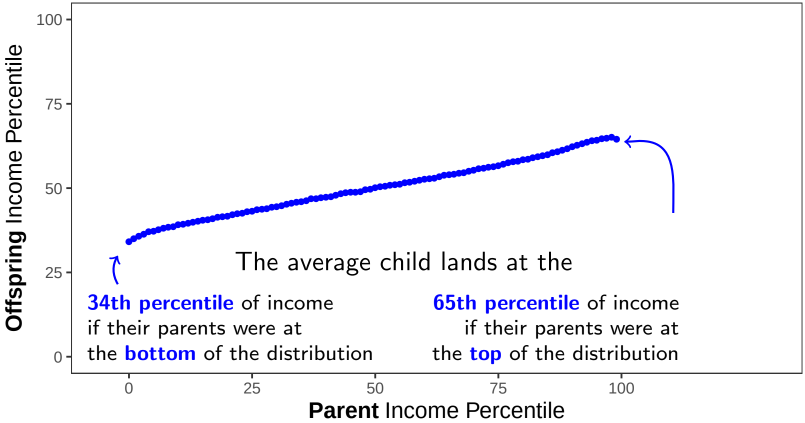
Chetty et al. 2017



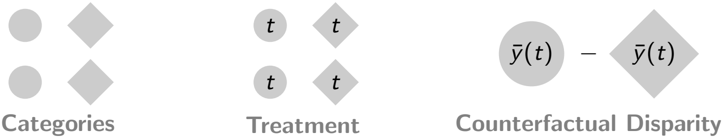


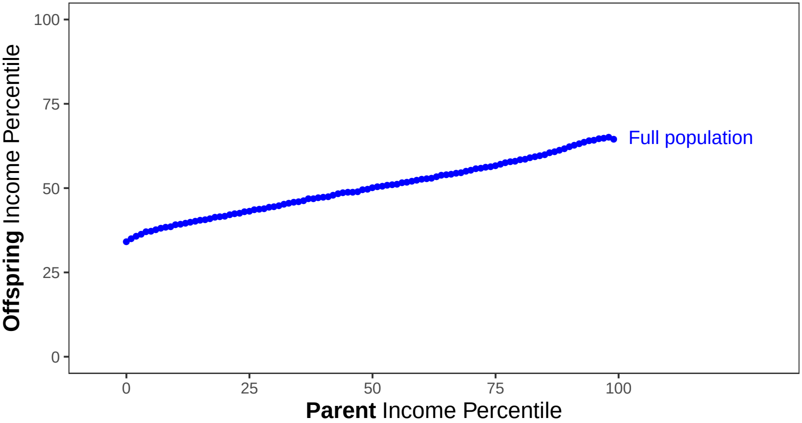
Chetty et al. 2017



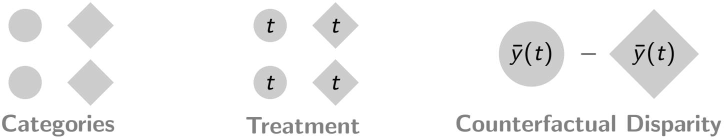


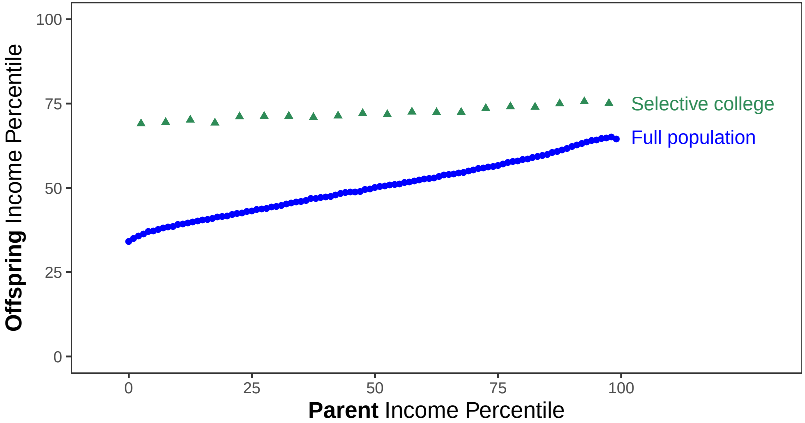
Chetty et al. 2017



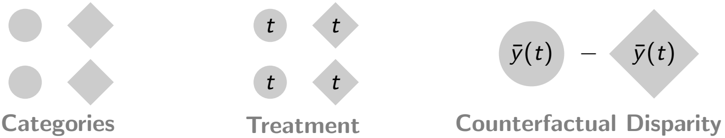


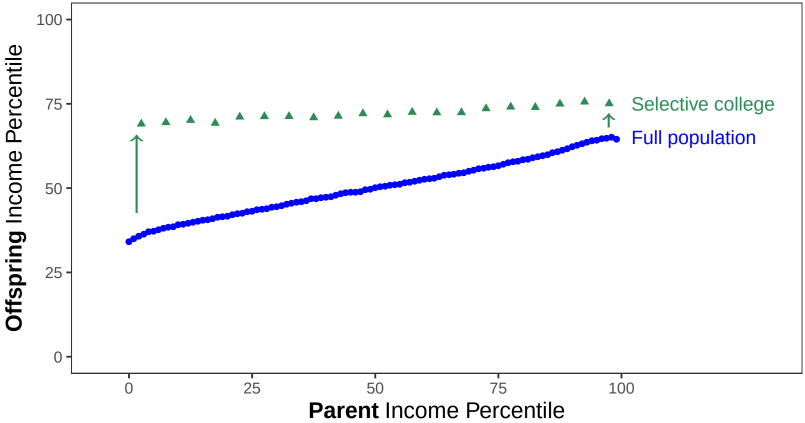
Chetty et al. 2017



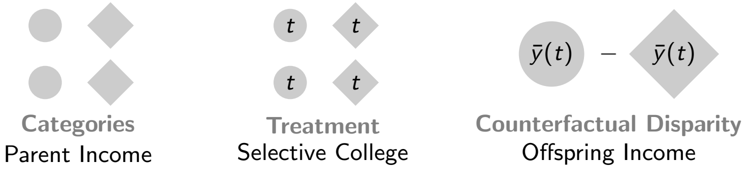


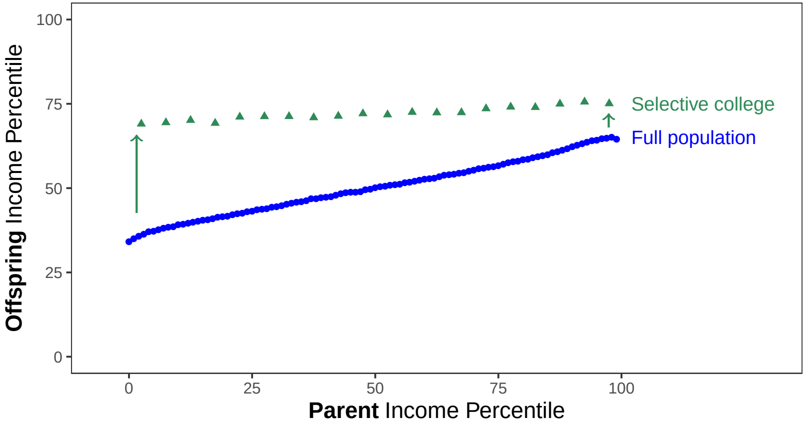
Chetty et al. 2017



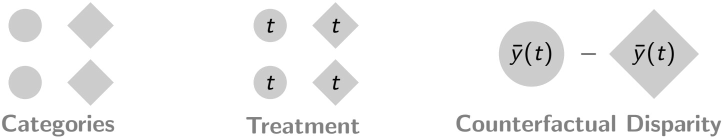


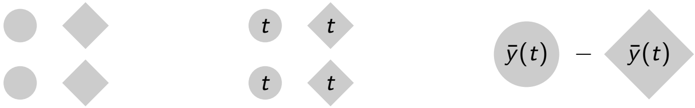
Chetty et al. 2017





Chetty et al. 2017


t
t
t
t
ȳ(t) − ȳ(t)
Categories
Treatment
Counterfactual Disparity

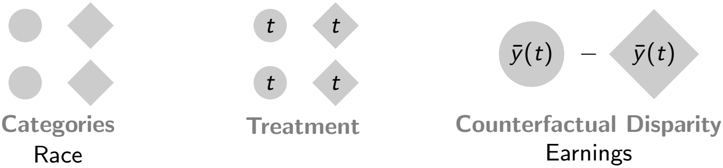


Categories **Treatment** **Counterfactual Disparity**

Race

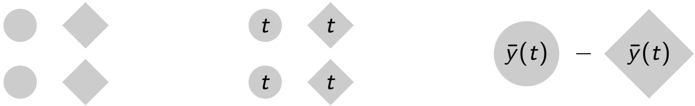
The difference in earnings
between blacks and whites

— Western 2006:12



The difference in earnings
between blacks and whites
would be reduced only by about 3 percent
if the incarceration rate were zero

— Western 2006:12



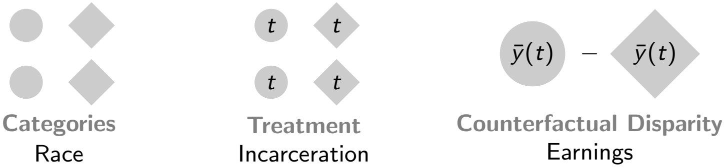
Categories
Race

Treatment
Incarceration

Counterfactual Disparity
Earnings

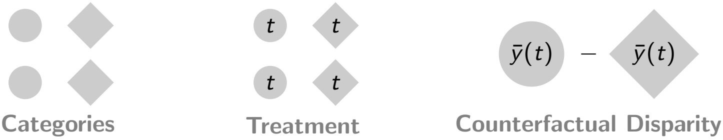
The difference in earnings
between blacks and whites
would be reduced only by about 3 percent
if the incarceration rate were zero

— Western 2006:12



The difference in earnings
between blacks and whites
would be reduced only by about 3 percent
if the incarceration rate were zero

— Western 2006:12



We often want to know
if **intervening** on a treatment variable
would close gaps

Causal decomposition analysis

Causal decomposition analysis

Unadjusted	Adjusted
$Y = \beta(\texttt{Black}) + \epsilon$	$Y = \gamma(\texttt{Black}) + \vec{X}'\vec{\eta} + \delta$

Causal decomposition analysis

Unadjusted	Adjusted
$Y = \beta(\texttt{Black}) + \epsilon$	$Y = \gamma(\texttt{Black}) + \vec{X}'\vec{\eta} + \delta$

Effect of race

Causal decomposition analysis

Unadjusted	Adjusted
$Y = \beta(\texttt{Black}) + \epsilon$	$Y = \gamma(\texttt{Black}) + \vec{X}'\vec{\eta} + \delta$

× Effect of race

Vanderweele & Robinson 2014

Causal decomposition analysis

Unadjusted	Adjusted
$Y = \beta(\texttt{Black}) + \epsilon$	$Y = \gamma(\texttt{Black}) + \vec{X}'\vec{\eta} + \delta$

- × Effect of race
- ✓ Disparity after intervention on $\vec{X}$

Vanderweele & Robinson 2014

Causal decomposition analysis

Unadjusted

$$Y = \beta(\texttt{Black}) + \epsilon$$

Adjusted

$$Y = \gamma(\texttt{Black}) + \vec{X}'\vec{\eta} + \delta$$

× Effect of race

✓ Disparity after intervention on $\vec{X}$

Vanderweele & Robinson 2014

Choice of intervention targets

Jackson & Vanderweele 2018

Causal decomposition analysis

Unadjusted

$$Y = \beta(\texttt{Black}) + \epsilon$$

Adjusted

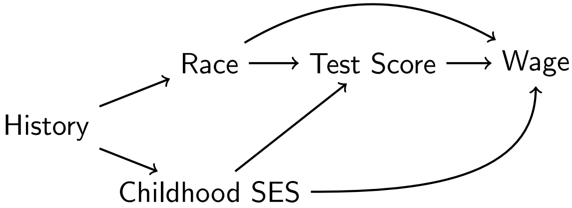
$$Y = \gamma(\texttt{Black}) + \vec{X}'\vec{\eta} + \delta$$

- × Effect of race
- ✓ Disparity after intervention on $\vec{X}$

Vanderweele & Robinson 2014

Choice of intervention targets

Jackson & Vanderweele 2018



Causal decomposition analysis

Unadjusted

$$Y = \beta(\texttt{Black}) + \epsilon$$

Adjusted

$$Y = \gamma(\texttt{Black}) + \vec{X}'\vec{\eta} + \delta$$

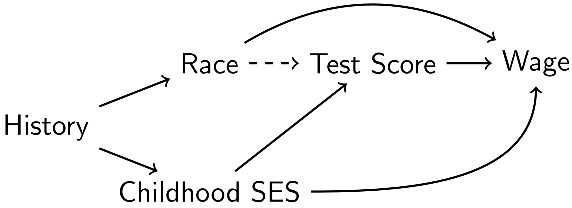
× Effect of race

✓ Disparity after intervention on $\vec{X}$

Vanderweele & Robinson 2014

Choice of intervention targets

Jackson & Vanderweele 2018



Causal decomposition analysis

Unadjusted

$$Y = \beta(\texttt{Black}) + \epsilon$$

Adjusted

$$Y = \gamma(\texttt{Black}) + \vec{X}'\vec{\eta} + \delta$$

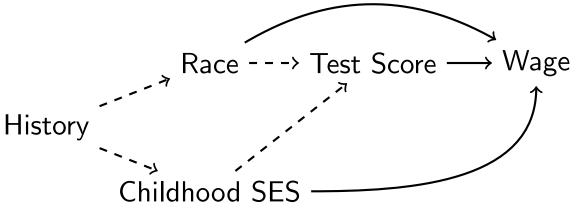
× Effect of race

✓ Disparity after intervention on $\vec{X}$

Vanderweele & Robinson 2014

Choice of intervention targets

Jackson & Vanderweele 2018



Causal decomposition analysis

Unadjusted	Adjusted
$Y = \beta(\texttt{Black}) + \epsilon$	$Y = \gamma(\texttt{Black}) + \vec{X}'\vec{\eta} + \delta$

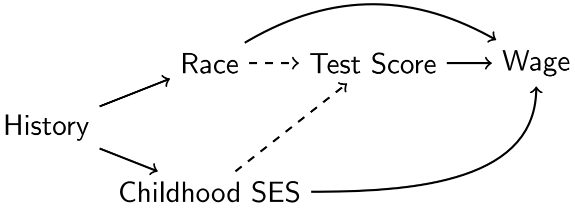
× Effect of race

✓ Disparity after intervention on $\vec{X}$

Vanderweele & Robinson 2014

Choice of intervention targets

Jackson & Vanderweele 2018



Causal decomposition analysis

Unadjusted

$$Y = \beta(\texttt{Black}) + \epsilon$$

Adjusted

$$Y = \gamma(\texttt{Black}) + \vec{X}'\vec{\eta} + \delta$$

- × Effect of race
- ✓ Disparity after intervention on $\vec{X}$

Vanderweele & Robinson 2014

Choice of intervention targets

Jackson & Vanderweele 2018

Equity: What should we equalize?

Jackson 2021

Causal decomposition analysis

Unadjusted	Adjusted
$Y = \beta(\texttt{Black}) + \epsilon$	$Y = \gamma(\texttt{Black}) + \vec{X}'\vec{\eta} + \delta$

× Effect of race

✓ Disparity after intervention on $\vec{X}$

Vanderweele & Robinson 2014

Choice of intervention targets

Jackson & Vanderweele 2018

Equity: What should we equalize?

Jackson 2021

Systems may adapt to maintain inequity

Jackson & Arah 2020

Causal decomposition analysis

Unadjusted

$$Y = \beta(\texttt{Black}) + \epsilon$$

Adjusted

$$Y = \gamma(\texttt{Black}) + \vec{X}'\vec{\eta} + \delta$$

× Effect of race

✓ Disparity after intervention on $\vec{X}$

Vanderweele & Robinson 2014

Choice of intervention targets

Jackson & Vanderweele 2018

Equity: What should we equalize?

Jackson 2021

Systems may adapt to maintain inequity

Jackson & Arah 2020

Present paper

— Local intervention interpretation

— Doubly robust estimator

— Software

Lundberg 2022

How this works

1. Define an intervention
2. Make causal assumptions
3. Estimate

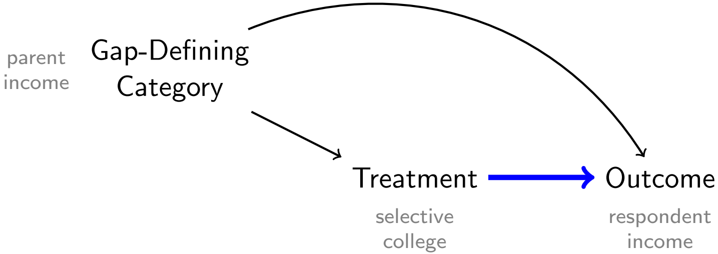
Define an intervention

Define an intervention

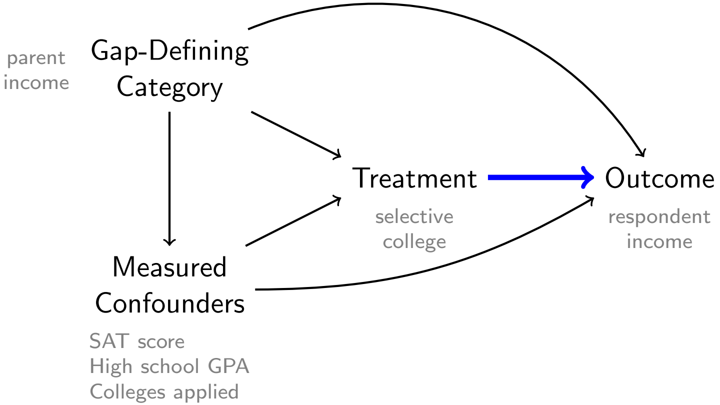
Using the Chetty et al. 2017 example,

What gap in respondent incomes would remain
across categories of parent income
if we intervened to send people to selective colleges?

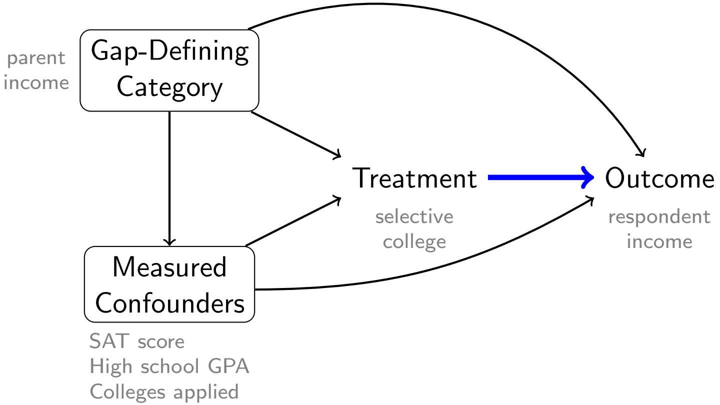
Make causal assumptions


parent
income
Gap-Defining
Category
Treatment
Outcome
selective
college
respondent
income

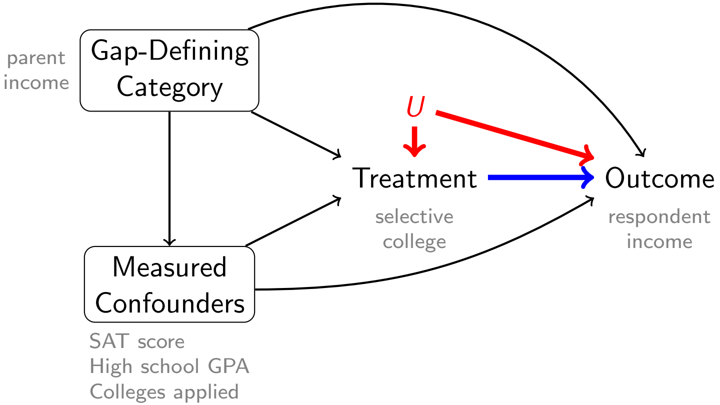
Pearl 2009


parent
income
Gap-Defining
Category
Treatment
Outcome
selective
college
respondent
income
Measured
Confounders
SAT score
High school GPA
Colleges applied

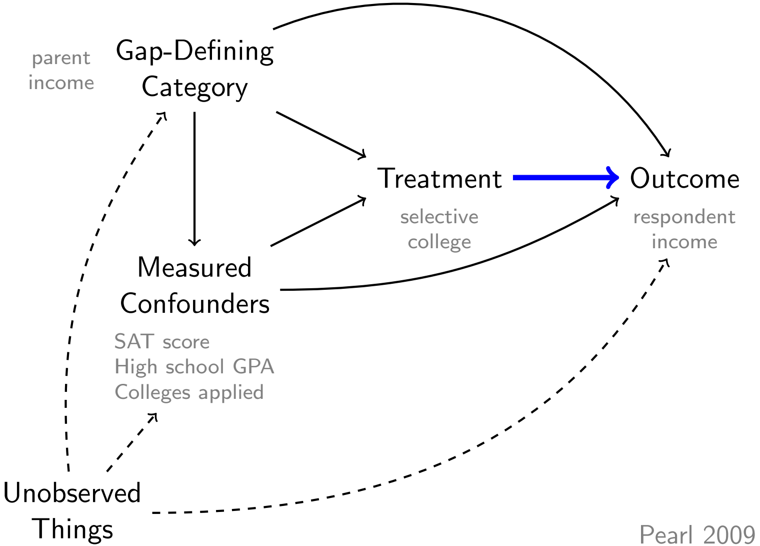
Pearl 2009


Gap-Defining Category
parent income
Treatment
selective college
Outcome
respondent income
Measured Confounders
SAT score
High school GPA
Colleges applied

Pearl 2009


Gap-Defining Category
parent income
U
Treatment
selective college
Outcome
respondent income
Measured Confounders
SAT score
High school GPA
Colleges applied

Pearl 2009



parent income
Gap-Defining Category
Treatment
selective college
Outcome
respondent income
Measured Confounders
SAT score
High school GPA
Colleges applied
Unobserved Things
Pearl 2009

Estimate

		Outcome under treatment	Outcome under control
People in category 1	Person 1	?	Y_1
	Person 2	Y_2	?
	Person 3	Y_3	?
People in category 2	Person 4	?	Y_4
	Person 5	Y_5	?
	Person 6	?	Y_6

		Outcome under treatment	Outcome under control
People in category 1	Person 1	?	Y_1
	Person 2	Y_2	?
	Person 3	Y_3	?
People in category 2	Person 4	?	Y_4
	Person 5	Y_5	?
	Person 6	?	Y_6

Learn a prediction function

		Outcome under treatment	Outcome under control
People in category 1	Person 1	?	Y_1
	Person 2	Y_2	?
	Person 3	Y_3	?
People in category 2	Person 4	?	Y_4
	Person 5	Y_5	?
	Person 6	?	Y_6

Learn a prediction function

		Outcome under treatment	Outcome under control
People in category 1	Person 1	?	Y_1
	Person 2	Y_2	?
	Person 3	Y_3	?
People in category 2	Person 4	?	Y_4
	Person 5	Y_5	?
	Person 6	?	Y_6

Predict the whole table

		Outcome under treatment	Outcome under control
People in category 1	Person 1	$\hat{Y}_1(1)$	$\hat{Y}_1(0)$
	Person 2	$\hat{Y}_2(1)$	$\hat{Y}_2(0)$
	Person 3	$\hat{Y}_3(1)$	$\hat{Y}_3(0)$
People in category 2	Person 4	$\hat{Y}_4(1)$	$\hat{Y}_4(0)$
	Person 5	$\hat{Y}_5(1)$	$\hat{Y}_5(0)$
	Person 6	$\hat{Y}_6(1)$	$\hat{Y}_6(0)$

Robins 1986
Hahn 1998

Learn a prediction function

		Outcome under treatment	Outcome under control
People in category 1	Person 1	?	Y_1
	Person 2	Y_2	?
	Person 3	Y_3	?
People in category 2	Person 4	?	Y_4
	Person 5	Y_5	?
	Person 6	?	Y_6

Predict the whole table

		Outcome under treatment	Outcome under control
People in category 1	Person 1	$\hat{Y}_1(1)$	$\hat{Y}_1(0)$
	Person 2	$\hat{Y}_2(1)$	$\hat{Y}_2(0)$
	Person 3	$\hat{Y}_3(1)$	$\hat{Y}_3(0)$
People in category 2	Person 4	$\hat{Y}_4(1)$	$\hat{Y}_4(0)$
	Person 5	$\hat{Y}_5(1)$	$\hat{Y}_5(0)$
	Person 6	$\hat{Y}_6(1)$	$\hat{Y}_6(0)$

Robins 1986
Hahn 1998

Learn a prediction function

		Outcome under treatment	Outcome under control
People in category 1	Person 1	?	Y_1
	Person 2	Y_2	?
	Person 3	Y_3	?
People in category 2	Person 4	?	Y_4
	Person 5	Y_5	?
	Person 6	?	Y_6

Predict the whole table

		Outcome under treatment	Outcome under control
People in category 1	Person 1	$\hat{Y}_1(1)$	$\hat{Y}_1(0)$
	Person 2	$\hat{Y}_2(1)$	$\hat{Y}_2(0)$
	Person 3	$\hat{Y}_3(1)$	$\hat{Y}_3(0)$
People in category 2	Person 4	$\hat{Y}_4(1)$	$\hat{Y}_4(0)$
	Person 5	$\hat{Y}_5(1)$	$\hat{Y}_5(0)$
	Person 6	$\hat{Y}_6(1)$	$\hat{Y}_6(0)$

Problem: Optimization for the wrong task

Learn a prediction function

		Outcome under treatment	Outcome under control
People in category 1	Person 1	?	Y_1
	Person 2	Y_2	?
	Person 3	Y_3	?
People in category 2	Person 4	?	Y_4
	Person 5	Y_5	?
	Person 6	?	Y_6

Predict the whole table

		Outcome under treatment	Outcome under control
People in category 1	Person 1	$\hat{Y}_1(1)$	$\hat{Y}_1(0)$
	Person 2	$\hat{Y}_2(1)$	$\hat{Y}_2(0)$
	Person 3	$\hat{Y}_3(1)$	$\hat{Y}_3(0)$
People in category 2	Person 4	$\hat{Y}_4(1)$	$\hat{Y}_4(0)$
	Person 5	$\hat{Y}_5(1)$	$\hat{Y}_5(0)$
	Person 6	$\hat{Y}_6(1)$	$\hat{Y}_6(0)$

Problem: Optimization for the wrong task

Prediction error over
observed
cases

Learn a prediction function

		Outcome under treatment	Outcome under control
People in category 1	Person 1	?	Y_1
	Person 2	Y_2	?
	Person 3	Y_3	?
People in category 2	Person 4	?	Y_4
	Person 5	Y_5	?
	Person 6	?	Y_6

Predict the whole table

		Outcome under treatment	Outcome under control
People in category 1	Person 1	$\hat{Y}_1(1)$	$\hat{Y}_1(0)$
	Person 2	$\hat{Y}_2(1)$	$\hat{Y}_2(0)$
	Person 3	$\hat{Y}_3(1)$	$\hat{Y}_3(0)$
People in category 2	Person 4	$\hat{Y}_4(1)$	$\hat{Y}_4(0)$
	Person 5	$\hat{Y}_5(1)$	$\hat{Y}_5(0)$
	Person 6	$\hat{Y}_6(1)$	$\hat{Y}_6(0)$

Problem: Optimization for the wrong task

Prediction error over **observed** cases

vs

Prediction error over **all** cases

Solution: Reweight errors to approximate the correct task

Solution: Reweight errors to approximate the correct task

		Prediction under treatment
People in category 1	Person 1	$\hat{Y}_1(1)$
	Person 2	$\hat{Y}_2(1)$
	Person 3	$\hat{Y}_3(1)$
People in category 2	Person 4	$\hat{Y}_4(1)$
	Person 5	$\hat{Y}_5(1)$
	Person 6	$\hat{Y}_6(1)$

Solution: Reweight errors to approximate the correct task

		Prediction under treatment	Outcome under treatment
People in category 1	Person 1	$\hat{Y}_1(1)$	?
	Person 2	$\hat{Y}_2(1)$	Y_2
	Person 3	$\hat{Y}_3(1)$	Y_3
People in category 2	Person 4	$\hat{Y}_4(1)$	?
	Person 5	$\hat{Y}_5(1)$	Y_5
	Person 6	$\hat{Y}_6(1)$	?

Solution: Reweight errors to approximate the correct task

		Prediction under treatment	Outcome under treatment	Error
People in category 1	Person 1	$\hat{Y}_1(1)$	?	?
	Person 2	$\hat{Y}_2(1)$	Y_2	$\hat{Y}_2(1) - Y_2$
	Person 3	$\hat{Y}_3(1)$	Y_3	$\hat{Y}_3(1) - Y_3$
People in category 2	Person 4	$\hat{Y}_4(1)$	?	?
	Person 5	$\hat{Y}_5(1)$	Y_5	$\hat{Y}_5(1) - Y_5$
	Person 6	$\hat{Y}_6(1)$	?	?

Solution: Reweight errors to approximate the correct task

		Prediction under treatment	Outcome under treatment	Error	Weight on error
People in category 1	Person 1	$\hat{Y}_1(1)$	?	?	
	Person 2	$\hat{Y}_2(1)$	Y_2	$\hat{Y}_2(1) - Y_2$	3 / 2
	Person 3	$\hat{Y}_3(1)$	Y_3	$\hat{Y}_3(1) - Y_3$	3 / 2
People in category 2	Person 4	$\hat{Y}_4(1)$	?	?	
	Person 5	$\hat{Y}_5(1)$	Y_5	$\hat{Y}_5(1) - Y_5$	3
	Person 6	$\hat{Y}_6(1)$	?	?	

Solution: Reweight errors to approximate the correct task

Estimated bias: Mean($\hat{Y}_i - Y_i$) with inverse probability of treatment weights

		Prediction under treatment	Outcome under treatment	Error	Weight on error
People in category 1	Person 1	$\hat{Y}_1(1)$	?	?	
	Person 2	$\hat{Y}_2(1)$	Y_2	$\hat{Y}_2(1) - Y_2$	3 / 2
	Person 3	$\hat{Y}_3(1)$	Y_3	$\hat{Y}_3(1) - Y_3$	3 / 2
People in category 2	Person 4	$\hat{Y}_4(1)$	?	?	
	Person 5	$\hat{Y}_5(1)$	Y_5	$\hat{Y}_5(1) - Y_5$	3
	Person 6	$\hat{Y}_6(1)$	?	?	

Solution: Reweight errors to approximate the correct task

Estimated bias: Mean($\hat{Y}_i - Y_i$) with inverse probability of treatment weights

New Estimate: (Original Estimate) $-$ (Estimated Bias)

Solution: Reweight errors to approximate the correct task

Estimated bias: Mean($\hat{Y}_i - Y_i$) with inverse probability of treatment weights

New Estimate: (Original Estimate) − (Estimated Bias)

Doubly Robust Estimation

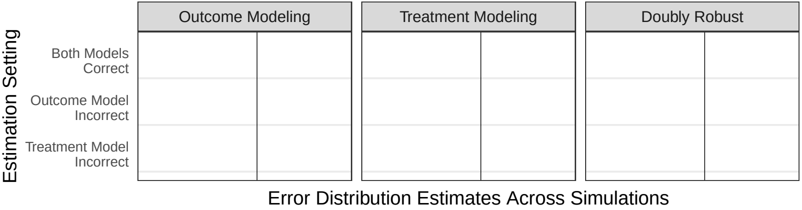
Robins, Rotnitzky, & Zhao 1994
Bang & Robins 2005

Solution: Reweight errors to approximate the correct task

Estimated bias: Mean($\hat{Y}_i - Y_i$) with inverse probability of treatment weights

New Estimate: (Original Estimate) − (Estimated Bias)

Doubly Robust Estimation



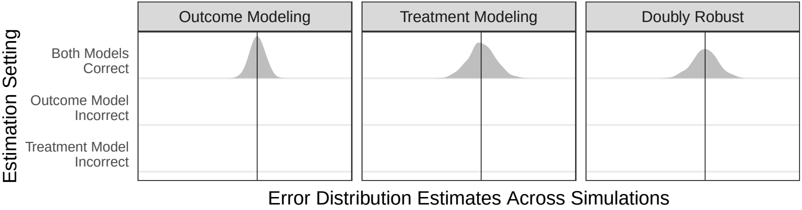
Robins, Rotnitzky, & Zhao 1994
Bang & Robins 2005

Solution: Reweight errors to approximate the correct task

Estimated bias: Mean($\hat{Y}_i - Y_i$) with inverse probability of treatment weights

New Estimate: (Original Estimate) − (Estimated Bias)

Doubly Robust Estimation



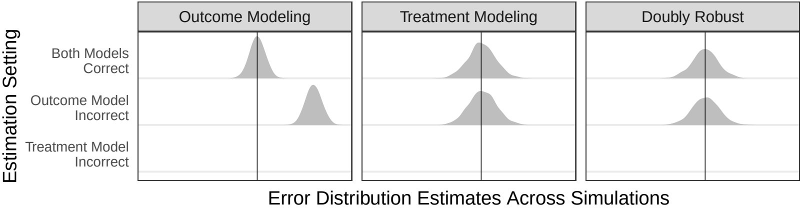
Robins, Rotnitzky, & Zhao 1994
Bang & Robins 2005

Solution: Reweight errors to approximate the correct task

Estimated bias: Mean($\hat{Y}_i - Y_i$) with inverse probability of treatment weights

New Estimate: (Original Estimate) − (Estimated Bias)

Doubly Robust Estimation



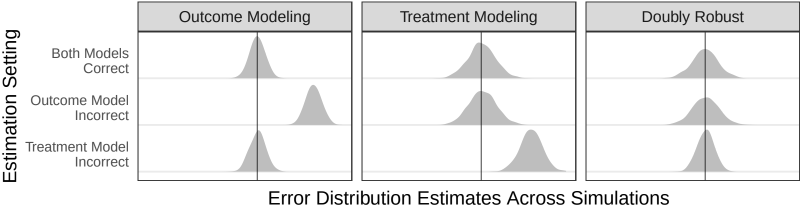
Robins, Rotnitzky, & Zhao 1994
Bang & Robins 2005

Solution: Reweight errors to approximate the correct task

Estimated bias: Mean($\hat{Y}_i - Y_i$) with inverse probability of treatment weights

New Estimate: (Original Estimate) − (Estimated Bias)

Doubly Robust Estimation



Robins, Rotnitzky, & Zhao 1994
Bang & Robins 2005

Solution: Reweight errors to approximate the correct task

Estimated bias: Mean($\hat{Y}_i - Y_i$) with
inverse probability of treatment weights

New Estimate: (Original Estimate) $-$ (Estimated Bias)

Doubly
Robust
Estimation

Even better:

Chernozhukov et al. 2018
Bickel 1982

Solution: Reweight errors to approximate the correct task

Estimated bias: Mean($\hat{Y}_i - Y_i$) with inverse probability of treatment weights

New Estimate: (Original Estimate) − (Estimated Bias)

Doubly Robust Estimation

Even better:

— Learn $\hat{Y}_i$ in sample A

— Estimate bias in sample B

Chernozhukov et al. 2018
Bickel 1982

Solution: Reweight errors to approximate the correct task

Estimated bias: Mean($\hat{Y}_i - Y_i$) with inverse probability of treatment weights

New Estimate: (Original Estimate) $-$ (Estimated Bias)

Doubly Robust Estimation

Even better:

— Learn $\hat{Y}_i$ in sample A

— Estimate bias in sample B

— Cross fit: Swap roles and average

Chernozhukov et al. 2018
Bickel 1982

Solution: Reweight errors to approximate the correct task

Estimated bias: Mean($\hat{Y}_i - Y_i$) with inverse probability of treatment weights

New Estimate: (Original Estimate) $-$ (Estimated Bias)

Doubly Robust Estimation

Even better:

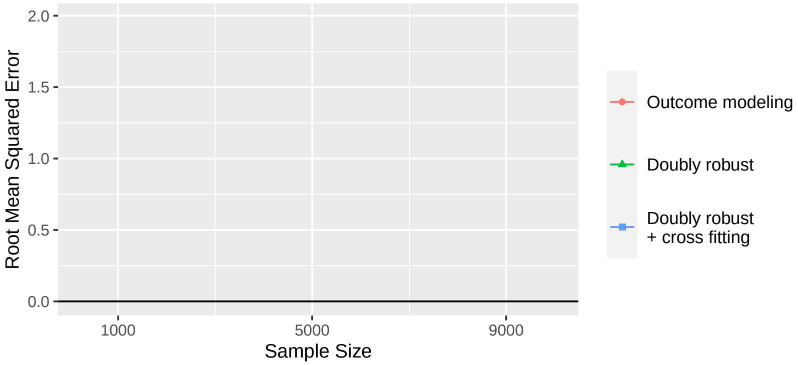
— Learn $\hat{Y}_i$ in sample A

— Estimate bias in sample B

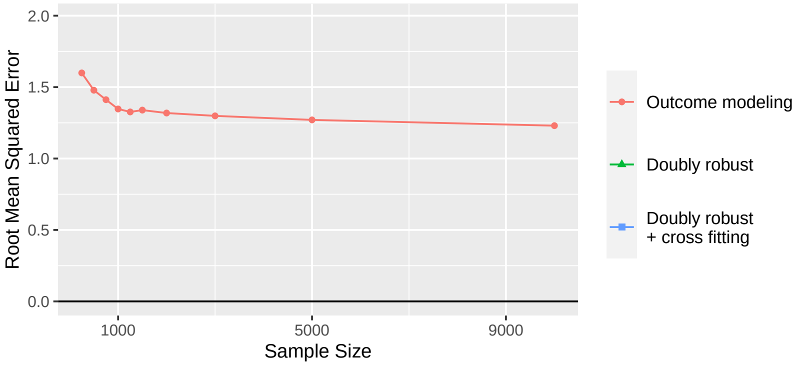
— Cross fit: Swap roles and average

Chernozhukov et al. 2018
Bickel 1982

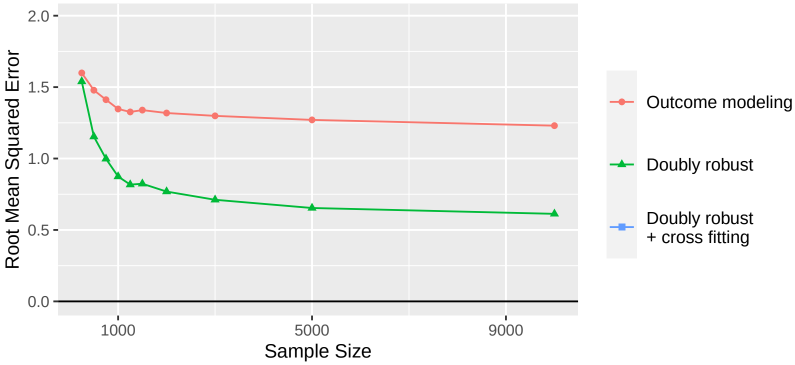
Solution: Reweight errors to approximate the correct task



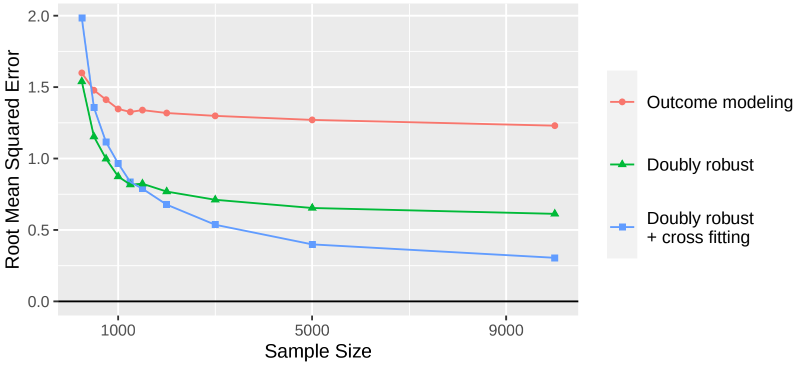
Solution: Reweight errors to approximate the correct task



Solution: Reweight errors to approximate the correct task



Solution: Reweight errors to approximate the correct task



Solution: Reweight errors to approximate the correct task

Estimated bias: Mean($\hat{Y}_i - Y_i$) with inverse probability of treatment weights

New Estimate: (Original Estimate) $-$ (Estimated Bias)

Doubly Robust Estimation

Even better:

— Learn $\hat{Y}_i$ in sample A

— Estimate bias in sample B

— Cross fit: Swap roles and average

Double Machine Learning

Chernozhukov et al. 2018
Bickel 1982

Solution: Reweight errors to approximate the correct task

Estimated bias: Mean($\hat{Y}_i - Y_i$) with inverse probability of treatment weights

New Estimate: (Original Estimate) $-$ (Estimated Bias)

Doubly Robust Estimation

Even better:

— Learn $\hat{Y}_i$ in sample A

— Estimate bias in sample B

— Cross fit: Swap roles and average

Double Machine Learning

So complicated!

gapclosing

An R package to estimate gap closing estimands. Install this package with the command
`devtools::install_github("ilundberg/gapclosing")`.

```
estimate <- gapclosing(
  data = simulated_data,
  outcome_formula = formula(outcome ~ category + confounder),
  treatment_formula = formula(treatment ~ category + confounder),
  category_name = "category",
  counterfactual_assignments = 1,
  outcome_algorithm = "ranger",
  treatment_algorithm = "ranger",
  sample_split = "cross_fit",
  se = T
)
```

```
       description estimate   se ci.min ci.max
       Factual gap     2.14 0.40   1.36    2.9
 Counterfactual gap     0.67 0.44  -0.19    1.5
```

How should one interpret results?

Interpret with respect to a **target trial** (Hernán & Robins 2016)

Interpret with respect to a **target trial** (Hernán & Robins 2016)

1. Sample $\mathcal{S}$ from the population

Interpret with respect to a **target trial** (Hernán & Robins 2016)

1. Sample $\mathcal{S}$ from the population
2. Assign treatment $T = 1$ to $\mathcal{S}$

Interpret with respect to a **target trial** (Hernán & Robins 2016)

1. Sample $\mathcal{S}$ from the population
2. Assign treatment $T = 1$ to $\mathcal{S}$
3. Observe the disparity across categories X

Interpret with respect to a **target trial** (Hernán & Robins 2016)

1. Sample $\mathcal{S}$ from the population
2. Assign treatment $T = 1$ to $\mathcal{S}$
3. Observe the disparity across categories X

Goal: Expected result over hypothetical samples $\mathcal{S}$

Interpret with respect to a **target trial** (Hernán & Robins 2016)

Local intervention

1. Sample $\mathcal{S}$ from the population
2. Assign treatment $T = 1$ to $\mathcal{S}$
3. Observe the disparity across categories X

Goal: Expected result over hypothetical samples $\mathcal{S}$

Interpret with respect to a **target trial** (Hernán & Robins 2016)

Local intervention

1. Sample $\mathcal{S}$ from the population
2. Assign treatment $T = 1$ to $\mathcal{S}$
3. Observe the disparity across categories X

Goal: Expected result over hypothetical samples $\mathcal{S}$

Global intervention

Interpret with respect to a **target trial** (Hernán & Robins 2016)

Local intervention

1. Sample $\mathcal{S}$ from the population
2. Assign treatment $T = 1$ to $\mathcal{S}$
3. Observe the disparity across categories X

Goal: Expected result over hypothetical samples $\mathcal{S}$

Global intervention

1. Take the entire population $\mathcal{P}$

Interpret with respect to a **target trial** (Hernán & Robins 2016)

Local intervention

1. Sample $\mathcal{S}$ from the population
2. Assign treatment $T = 1$ to $\mathcal{S}$
3. Observe the disparity across categories X

Goal: Expected result over hypothetical samples $\mathcal{S}$

Global intervention

1. Take the entire population $\mathcal{P}$
2. Assign treatment $T = 1$ to $\mathcal{P}$

Interpret with respect to a **target trial** (Hernán & Robins 2016)

Local intervention

1. Sample $\mathcal{S}$ from the population
2. Assign treatment $T = 1$ to $\mathcal{S}$
3. Observe the disparity across categories X

Goal: Expected result over hypothetical samples $\mathcal{S}$

Global intervention

1. Take the entire population $\mathcal{P}$
2. Assign treatment $T = 1$ to $\mathcal{P}$
3. Observe the disparity across categories X

Interpret with respect to a **target trial** (Hernán & Robins 2016)

Local intervention

1. Sample $\mathcal{S}$ from the population
2. Assign treatment $T = 1$ to $\mathcal{S}$
3. Observe the disparity across categories X

Goal: Expected result over hypothetical samples $\mathcal{S}$

Global intervention

1. Take the entire population $\mathcal{P}$
2. Assign treatment $T = 1$ to $\mathcal{P}$
3. Observe the disparity across categories X

Goal: Result of this procedure

Interpret with respect to a **target trial** (Hernán & Robins 2016)

Local intervention

1. Sample $\mathcal{S}$ from the population
2. Assign treatment $T = 1$ to $\mathcal{S}$
3. Observe the disparity across categories X

Goal: Expected result over hypothetical samples $\mathcal{S}$

Difficulty: Causal inference

Global intervention

1. Take the entire population $\mathcal{P}$
2. Assign treatment $T = 1$ to $\mathcal{P}$
3. Observe the disparity across categories X

Goal: Result of this procedure

Difficulty: Causal inference

Interpret with respect to a **target trial** (Hernán & Robins 2016)

Local intervention

1. Sample $\mathcal{S}$ from the population
2. Assign treatment $T = 1$ to $\mathcal{S}$
3. Observe the disparity across categories X

Goal: Expected result over hypothetical samples $\mathcal{S}$

Difficulty: Causal inference

Global intervention

1. Take the entire population $\mathcal{P}$
2. Assign treatment $T = 1$ to $\mathcal{P}$
3. Observe the disparity across categories X

Goal: Result of this procedure

Difficulty: Causal inference
Equilibrium dynamics

Interpret with respect to a **target trial** (Hernán & Robins 2016)

Local intervention

1. Sample $\mathcal{S}$ from the population
2. Assign treatment $T = 1$ to $\mathcal{S}$
3. Observe the disparity across categories X

Goal: Expected result over hypothetical samples $\mathcal{S}$

Difficulty: Causal inference

Global intervention

1. Take the entire population $\mathcal{P}$
2. Assign treatment $T = 1$ to $\mathcal{P}$
3. Observe the disparity across categories X

Goal: Result of this procedure

Difficulty: Causal inference
Equilibrium dynamics

Interpret with respect to a **target trial** (Hernán & Robins 2016)

Policy-relevant?

$\downarrow$

Local intervention	Global intervention
1. Sample $\mathcal{S}$ from the population 2. Assign treatment $T = 1$ to $\mathcal{S}$ 3. Observe the disparity across categories X	1. Take the entire population $\mathcal{P}$ 2. Assign treatment $T = 1$ to $\mathcal{P}$ 3. Observe the disparity across categories X
Goal: Expected result over hypothetical samples $\mathcal{S}$	Goal: Result of this procedure
Difficulty: Causal inference	Difficulty: Causal inference Equilibrium dynamics

Interpret with respect to a **target trial** (Hernán & Robins 2016)

Local intervention

1. Sample $\mathcal{S}$ from the population
2. Assign treatment $T = 1$ to $\mathcal{S}$
3. Observe the disparity across categories X

Goal: Expected result over hypothetical samples $\mathcal{S}$

Difficulty: Causal inference

Global intervention

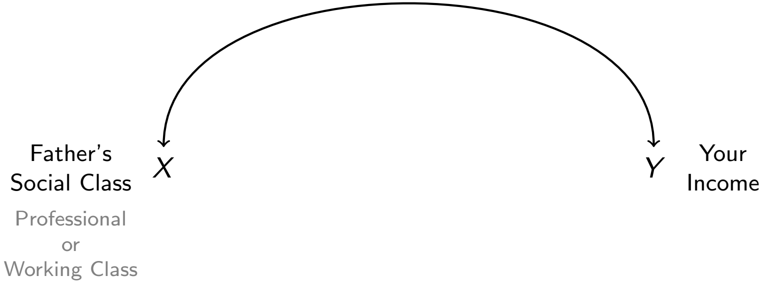
1. Take the entire population $\mathcal{P}$
2. Assign treatment $T = 1$ to $\mathcal{P}$
3. Observe the disparity across categories X

Goal: Result of this procedure

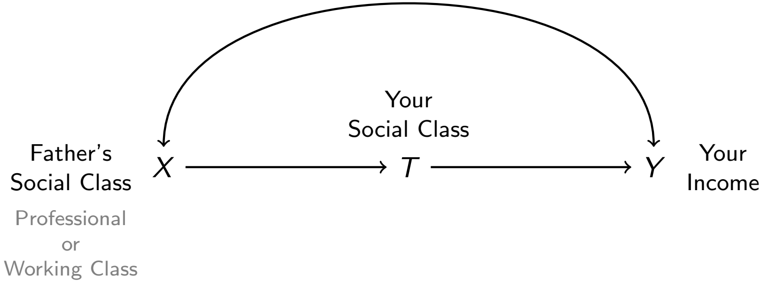
Difficulty: Causal inference
Equilibrium dynamics

Empirical Examples

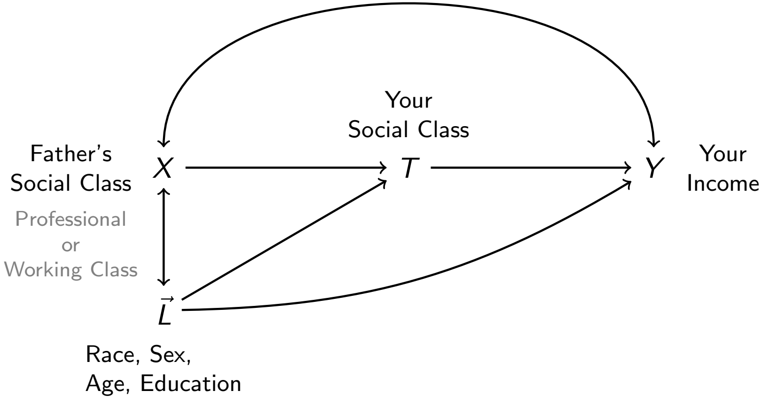
Empirical Example 1: Economic Mobility



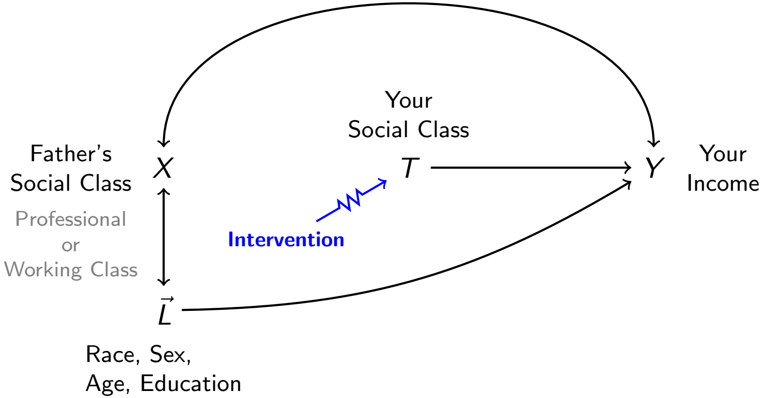
Empirical Example 1: Economic Mobility



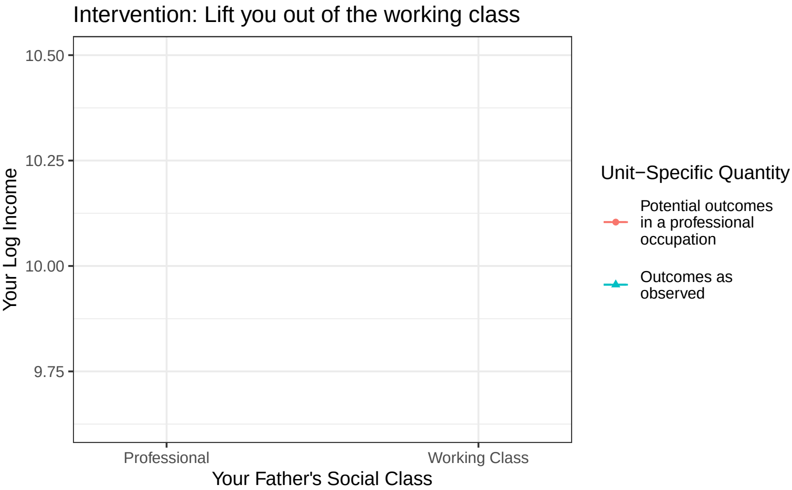
Empirical Example 1: Economic Mobility



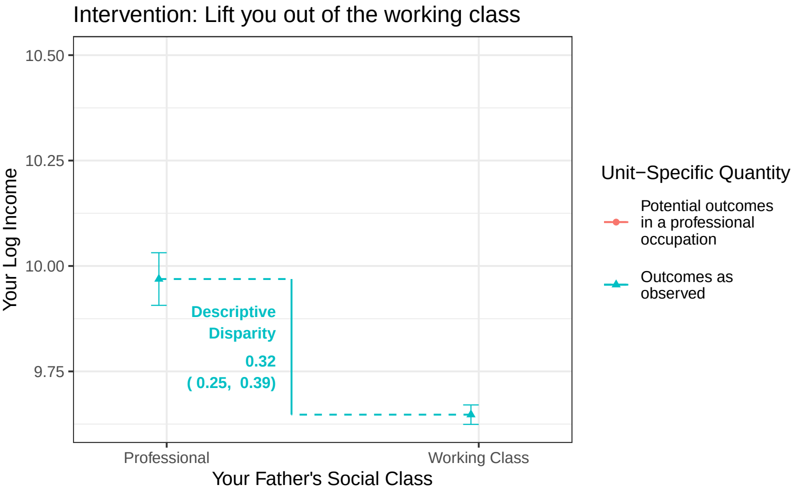
Empirical Example 1: Economic Mobility



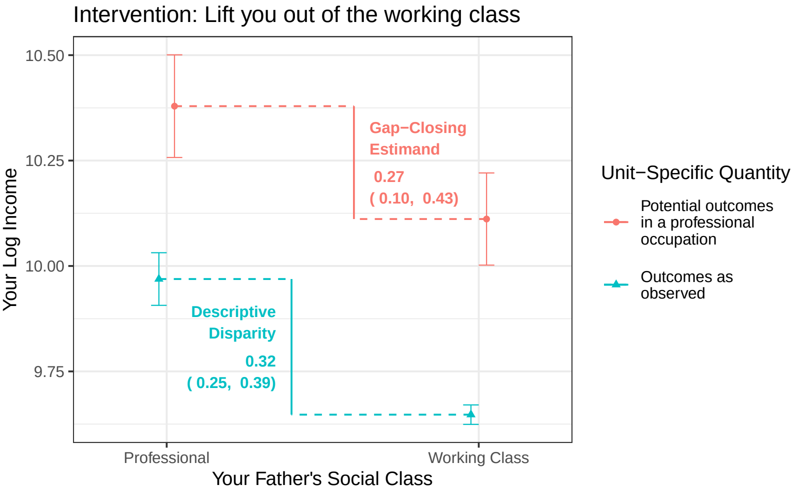
Empirical Example 1: Economic Mobility



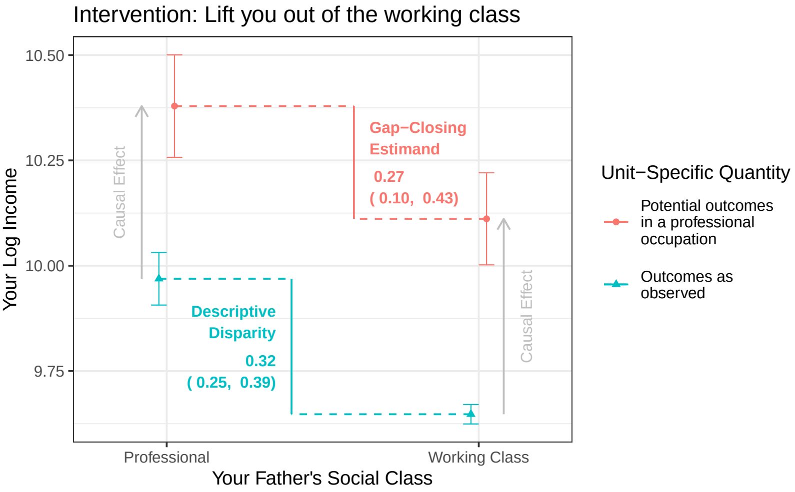
Empirical Example 1: Economic Mobility



Empirical Example 1: Economic Mobility

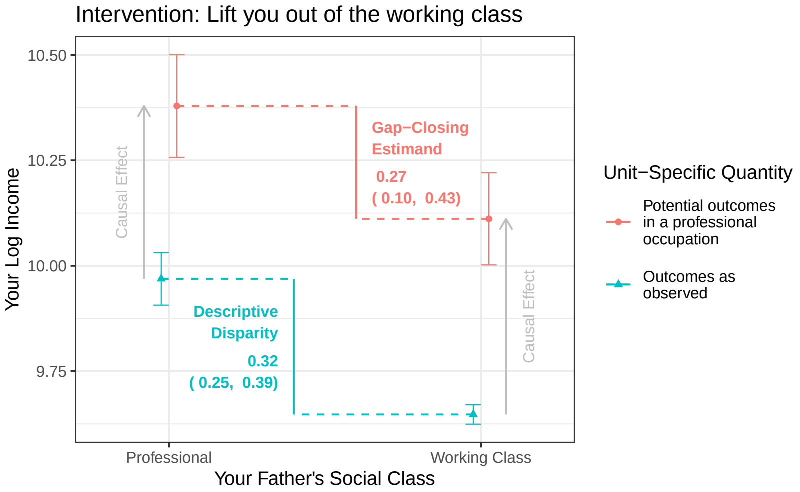


Empirical Example 1: Economic Mobility

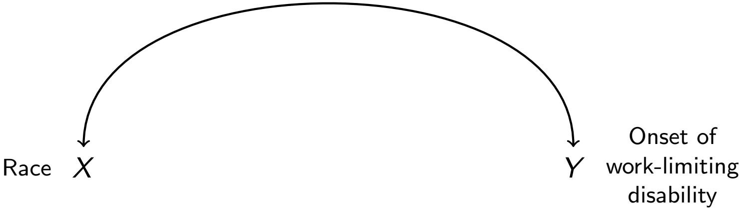


Empirical Example 1: Economic Mobility

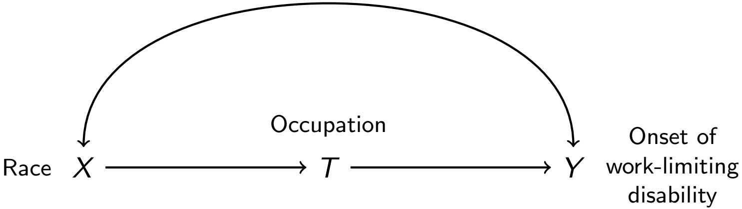
`plot_two_categories()`



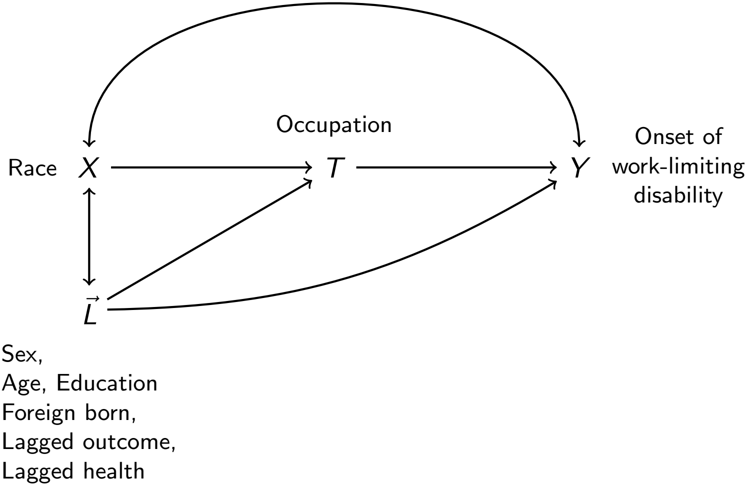
Empirical Example 2: Racial Disparities in Health



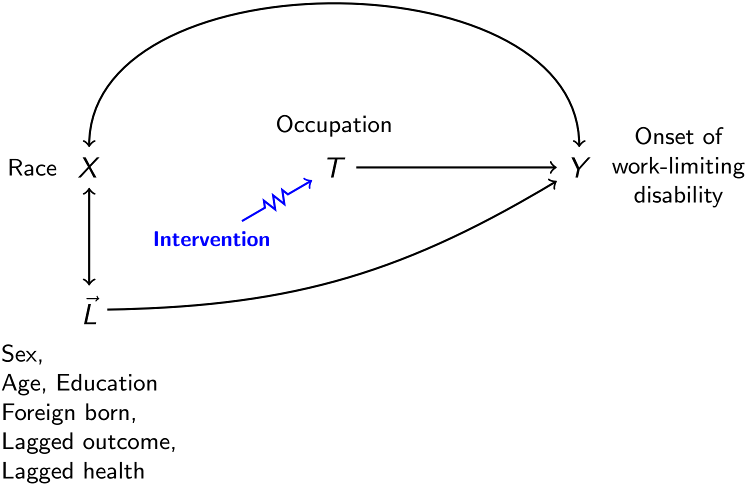
Empirical Example 2: Racial Disparities in Health



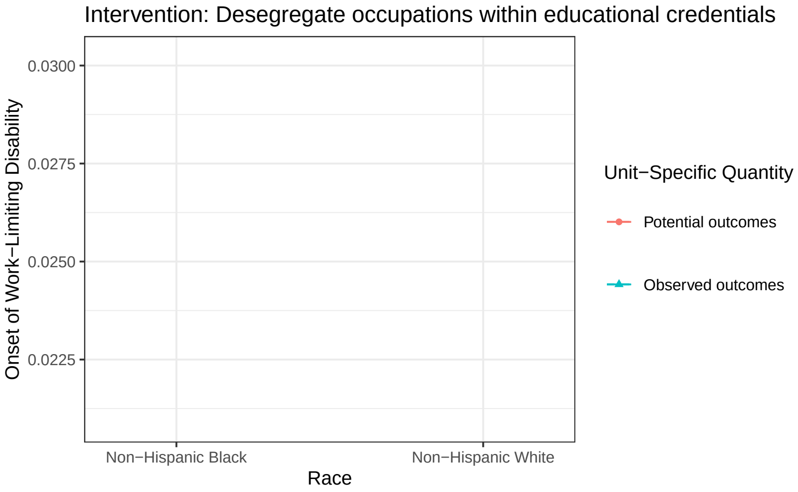
Empirical Example 2: Racial Disparities in Health



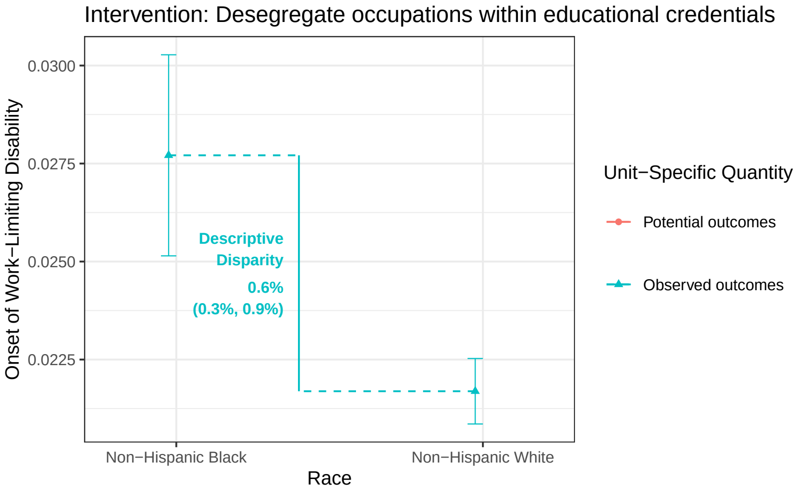
Empirical Example 2: Racial Disparities in Health



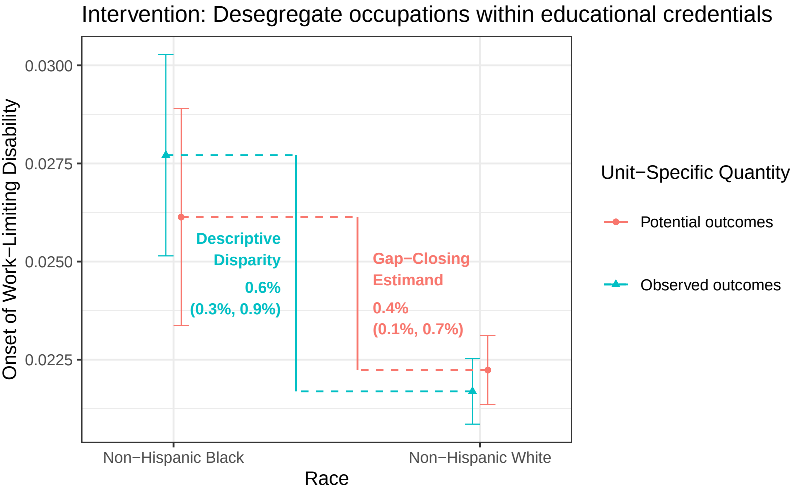
Empirical Example 2: Racial Disparities in Health



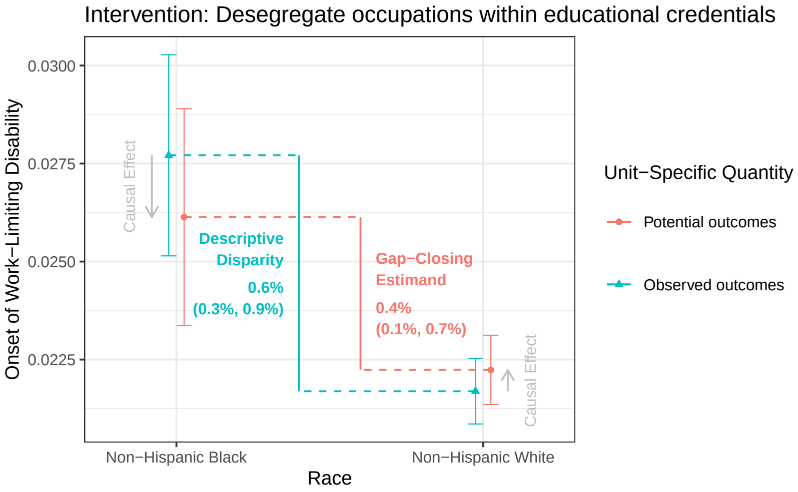
Empirical Example 2: Racial Disparities in Health



Empirical Example 2: Racial Disparities in Health



Empirical Example 2: Racial Disparities in Health



Discussion

Gap closing estimands

- Define the goal
  - two target populations
  - hypothetical intervention
- Make causal assumptions
  - draw a DAG
- Estimate
  - fit an outcome model
  - change the treatment
  - predict for everyone
  - average within the populations

Now try it!

ilundberg.github.io/gapclosing

Ian Lundberg
ianlundberg.org
ilundberg@cornell.edu