The Front-Door Criterion in Theory and Practice

Adam Glynn
Emory University

Konstantin Kashin
Facebook

Washington University in St. Loius
August 15, 2019

Outline

Outline

Options for Assessing the Effects of a Program

Options for Assessing the Effects of a Program

- Compare pre- and post-program outcomes of participants (interrupted time series)

Options for Assessing the Effects of a Program

- Compare pre- and post-program outcomes of participants (interrupted time series)
- Compare participants to eligible non-participants (regression, matching, IPW, etc.)

Options for Assessing the Effects of a Program

- Compare pre- and post-program outcomes of participants (interrupted time series)
- Compare participants to eligible non-participants (regression, matching, IPW, etc.)
- Compare pre- and post-program outcomes of participants to pre- and post-program outcomes of eligible non-participants (DiD, generalized synthetic control, matrix completion, etc.)

Options for Assessing the Effects of a Program

- Compare pre- and post-program outcomes of participants (interrupted time series)
- Compare participants to eligible non-participants (regression, matching, IPW, etc.)
- Compare pre- and post-program outcomes of participants to pre- and post-program outcomes of eligible non-participants (DiD, generalized synthetic control, matrix completion, etc.)
- Compare barely-eligible participants to barely-not-eligible applicants (RDD)

Options for Assessing the Effects of a Program

- Compare pre- and post-program outcomes of participants (interrupted time series)
- Compare participants to eligible non-participants (regression, matching, IPW, etc.)
- Compare pre- and post-program outcomes of participants to pre- and post-program outcomes of eligible non-participants (DiD, generalized synthetic control, matrix completion, etc.)
- Compare barely-eligible participants to barely-not-eligible applicants (RDD)
- Randomly encourage the eligible to enter the program (instrumental variables)

Options for Assessing the Effects of a Program

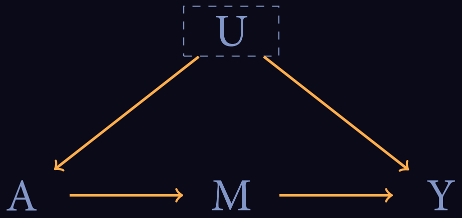
- Compare pre- and post-program outcomes of participants (interrupted time series)
- Compare participants to eligible non-participants (regression, matching, IPW, etc.)
- Compare pre- and post-program outcomes of participants to pre- and post-program outcomes of eligible non-participants (DiD, generalized synthetic control, matrix completion, etc.)
- Compare barely-eligible participants to barely-not-eligible applicants (RDD)
- Randomly encourage the eligible to enter the program (instrumental variables)
- Lottery the eligible applicants into the program (RCT)

Options for Assessing the Effects of a Program

- Compare pre- and post-program outcomes of participants (interrupted time series)
- Compare participants to eligible non-participants (regression, matching, IPW, etc.)
- Compare pre- and post-program outcomes of participants to pre- and post-program outcomes of eligible non-participants (DiD, generalized synthetic control, matrix completion, etc.)
- Compare barely-eligible participants to barely-not-eligible applicants (RDD)
- Randomly encourage the eligible to enter the program (instrumental variables)
- Lottery the eligible applicants into the program (RCT)

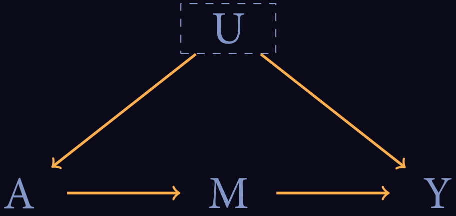
None of these techniques use information on the mechanisms by which the program is supposed to produce the effect (e.g. program attendance, learning, etc.).

The Front-door Approach



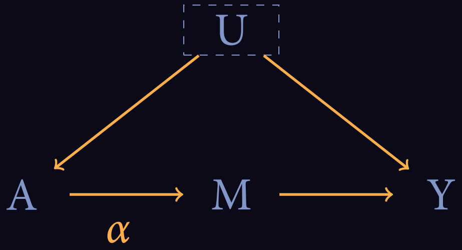
Pearl (1995) showed that <u>post-treatment mechanism</u> variables M can be used to non-parametrically identify causal effects when selection on observables doesn't hold.

Front-door Adjustment (easiest version)



With linear constant-effects models and singleton M, the front-door adjustment reduces to the following:

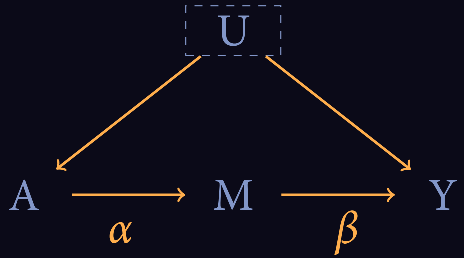
Front-door Adjustment (easiest version)



With linear constant-effects models and singleton M, the front-door adjustment reduces to the following:

- Get coefficient on A from regression of M on A.

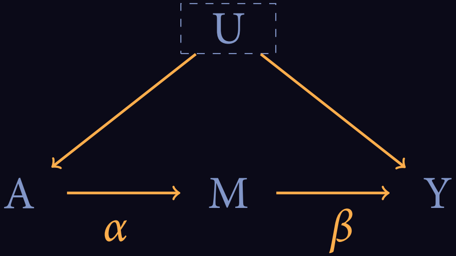
Front-door Adjustment (easiest version)



With linear constant-effects models and singleton M, the front-door adjustment reduces to the following:

- Get coefficient on A from regression of M on A.
- Get coefficient on M from regression of Y on M and A.

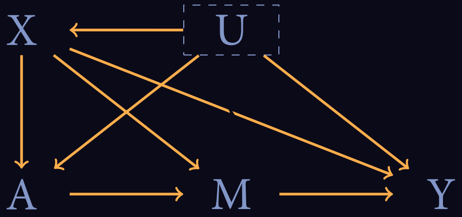
Front-door Adjustment (easiest version)



With linear constant-effects models and singleton M, the front-door adjustment reduces to the following:

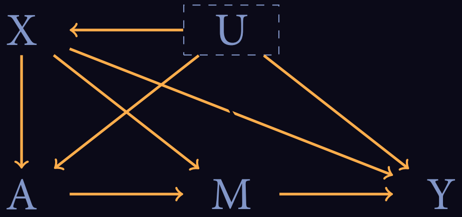
- Get coefficient on A from regression of M on A.
- Get coefficient on M from regression of Y on M and A.
- Multiply coeffients ($\alpha \cdot \beta$).

Front-door Adjustment (easiest, with covariates)



With linear constant-effects models and singleton M, the front-door adjustment reduces to the following:

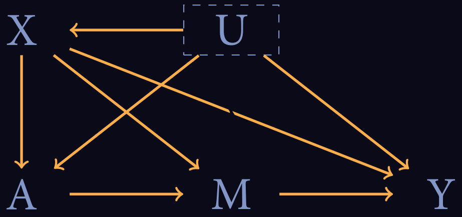
Front-door Adjustment (easiest, with covariates)



With linear constant-effects models and singleton M, the front-door adjustment reduces to the following:

- Get coefficient on A from regression of M on A and X.

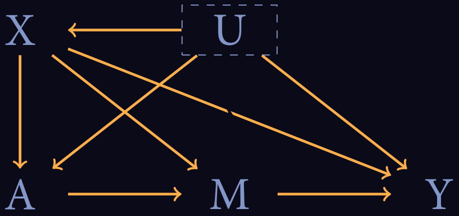
Front-door Adjustment (easiest, with covariates)



With linear constant-effects models and singleton M, the front-door adjustment reduces to the following:

- Get coefficient on A from regression of M on A and X.
- Get coefficient on M from regression of Y on M, A, and X.

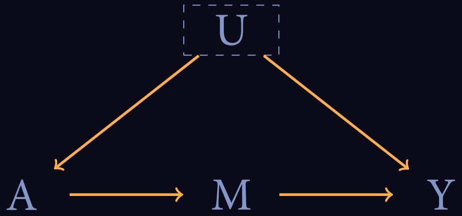
Front-door Adjustment (easiest, with covariates)



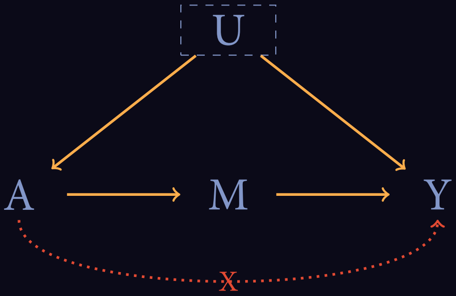
With linear constant-effects models and singleton M, the front-door adjustment reduces to the following:

- Get coefficient on A from regression of M on A and X.
- Get coefficient on M from regression of Y on M, A, and X.
- Multiply coeffients.

Front-door Criterion (Informal)

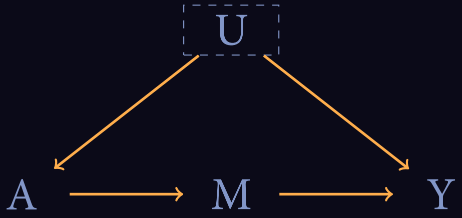


Front-door Criterion (Informal)

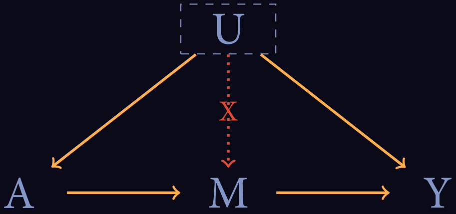


A affects *Y* only through *M* (e.g., *M* indicates program attendance).

Front-door Criterion (Informal)

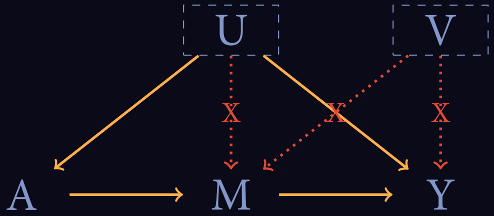


Front-door Criterion (Informal)



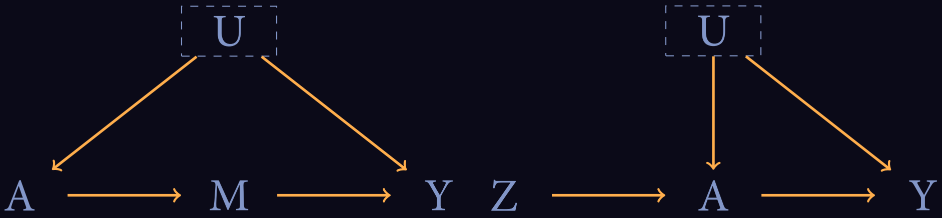
M as-if randomized conditional on *A* and *X*.
(e.g., non-attendance due to random factors).

Front-door Criterion (Informal)



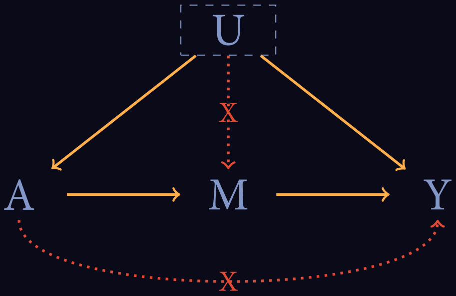
M as-if randomized conditional on *A* and *X*.
(e.g., non-attendance due to random factors).

Front-door vs Instrumental Variables

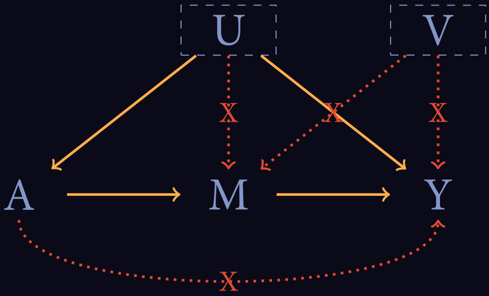


Why has front-door been so rarely used?

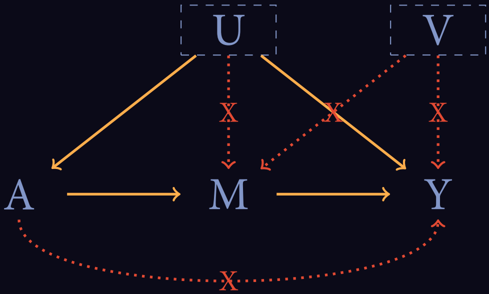
Why has front-door been so rarely used?



Why has front-door been so rarely used?

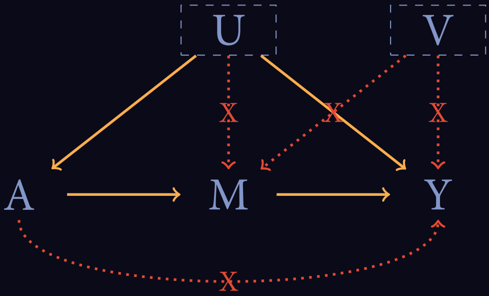


Why has front-door been so rarely used?



Often hard to find sets of covariates *X* and mediators *M* that satisfy these conditions.

Why has front-door been so rarely used?



Often hard to find sets of covariates X and mediators M that satisfy these conditions.

Glynn and Kashin 2017 and 2018 show that front-door can be informative, even when conditions don't hold exactly.

Outline

Motivation and Introduction

Front-door for ATT

Front-Door Difference-in-Differences

Average treatment effect on the treated (ATT)

Average treatment effect on the treated (ATT)

- Let a_1 denote active treatment (e.g., program)

Average treatment effect on the treated (ATT)

- Let a_1 denote active treatment (e.g., program)
- Let a_0 denote control (e.g., no program)

Average treatment effect on the treated (ATT)

- Let a_1 denote active treatment (e.g., program)
- Let a_0 denote control (e.g., no program)
- Y is the outcome (e.g., earnings)

Average treatment effect on the treated (ATT)

- Let a_1 denote active treatment (e.g., program)
- Let a_0 denote control (e.g., no program)
- Y is the outcome (e.g., earnings)
- $Y(a_0)$ is potential outcome under control (e.g., earnings without program)

Average treatment effect on the treated (ATT)

- Let a_1 denote active treatment (e.g., program)
- Let a_0 denote control (e.g., no program)
- Y is the outcome (e.g., earnings)
- $Y(a_0)$ is potential outcome under control (e.g., earnings without program); counterfactual for those in the program

Average treatment effect on the treated (ATT)

- Let a_1 denote active treatment (e.g., program)
- Let a_0 denote control (e.g., no program)
- Y is the outcome (e.g., earnings)
- $Y(a_0)$ is potential outcome under control (e.g., earnings without program); counterfactual for those in the program

$$ATT = \underbrace{\mathrm{E}[Y|a_1]}_{\text{Mean Outcome for Treated}} - \underbrace{\mathrm{E}[Y(a_0)|a_1]}_{\text{Mean Counterfactual Outcome for Treated}}$$

Average treatment effect on the treated (ATT)

- Let a_1 denote active treatment (e.g., program)
- Let a_0 denote control (e.g., no program)
- Y is the outcome (e.g., earnings)
- $Y(a_0)$ is potential outcome under control (e.g., earnings without program); counterfactual for those in the program

$$ATT = \underbrace{\mathrm{E}[Y|a_1]}_{\text{Mean Outcome for Treated}} - \underbrace{\mathrm{E}[Y(a_0)|a_1]}_{\text{Mean Counterfactual Outcome for Treated}}$$

Front-door Estimator for ATT

Let x be a set of observed covariates, and u be a set of unobserved covariates, such that

$$\underbrace{E[Y(a_0)|a_1]}_{\text{Counterfactual}} = \sum_x \sum_u \underbrace{E[Y|a_0, x, u]}_{\text{Unobserved}} \cdot \underbrace{P(u|x, a_1)}_{\text{Unobserved}} \cdot P(x|a_1)$$

Front-door Estimator for ATT

Let x be a set of observed covariates, and u be a set of unobserved covariates, such that

$$\underbrace{\mathrm{E}[Y(a_0)|a_1]}_{\text{Counterfactual}} = \sum_x \sum_u \underbrace{E[Y|a_0, x, u]}_{\text{Unobserved}} \cdot \underbrace{P(u|x, a_1)}_{\text{Unobserved}} \cdot P(x|a_1)$$

Let m be a set of mediating variables. The front-door estimator for this counterfactual can be written as

$$\widehat{E}_{fd}[Y(a_0)|a_1] = \sum_x \sum_m E[Y|a_1, x, m] \cdot P(m|x, a_0) \cdot P(x|a_1)$$

Identification

$$\sum_x \sum_u E[Y|a_0, x, u] \cdot P(u|x, a_1) \cdot P(x|a_1) =$$
$$\sum_x \sum_m E[Y|a_1, x, m] \cdot P(m|x, a_0) \cdot P(x|a_1)$$

when the following assumptions hold

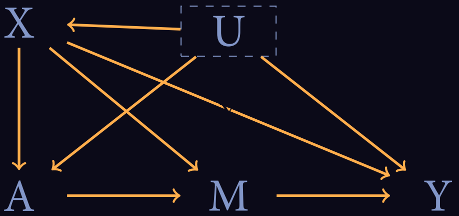
Assumption (1)

$E[Y|a_0, x, u, m] = E[Y|a_1, x, u, m]$

Assumption (2)

$P(m|x, a_0) = P(m|x, a_0, u)$ *and* $P(u|x, a_1) = P(u|x, a_1, m)$

Identification



Assumption (1)

$$E[Y|a_0, x, u, m] = E[Y|a_1, x, u, m]$$

Assumption (2)

$$P(m|x, a_0) = P(m|x, a_0, u) \text{ and } P(u|x, a_1) = P(u|x, a_1, m)$$

Develop intuition in a simple (canonical?) case

Develop intuition in a simple (canonical?) case

- M is binary variable that denotes receipt of treatment (e.g., m_1 denotes program attendance)

Develop intuition in a simple (canonical?) case

- M is binary variable that denotes receipt of treatment (e.g., m_1 denotes program attendance)
- Non-compliance for treated units

$$0 < P(m_0|a_1) < 1$$

Develop intuition in a simple (canonical?) case

- M is binary variable that denotes receipt of treatment (e.g., m_1 denotes program attendance)
- Non-compliance for treated units

$$0 < P(m_0|a_1) < 1$$

- No non-compliance for control units

$$P(m_1|a_0) = 0$$

Standard and front-door estimators under one-sided noncompliance

$$\text{ATT} = \mathrm{E}[Y|a_1] - \underbrace{\mathrm{E}[Y(a_0)|a_1]}_{\text{Counterfactual}}$$

Standard and front-door estimators under one-sided noncompliance

$$\text{ATT} = \mathrm{E}[Y|a_1] - \underbrace{\mathrm{E}[Y(a_0)|a_1]}_{\text{Counterfactual}}$$

$$\widehat{ATT}_{\text{Standard}} = \mathrm{E}[Y|a_1] - \sum_x \underbrace{\mathrm{E}[Y|a_0, x]}_{\text{Controls}} \cdot P(x|a_1)$$

Standard and front-door estimators under one-sided noncompliance

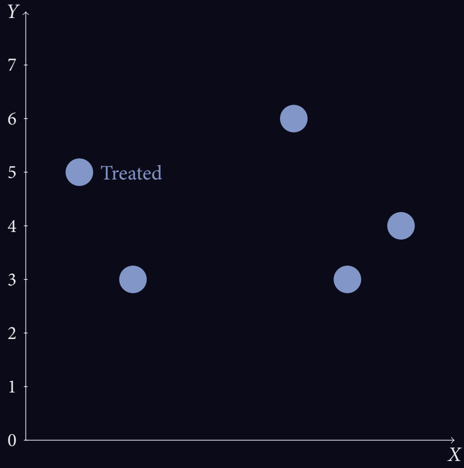
$$\text{ATT} = \text{E}[Y|a_1] - \underbrace{\text{E}[Y(a_0)|a_1]}_{\text{Counterfactual}}$$

$$\widehat{ATT}_{\text{Standard}} = \text{E}[Y|a_1] - \sum_x \underbrace{\text{E}[Y|a_0, x]}_{\text{Controls}} \cdot P(x|a_1)$$

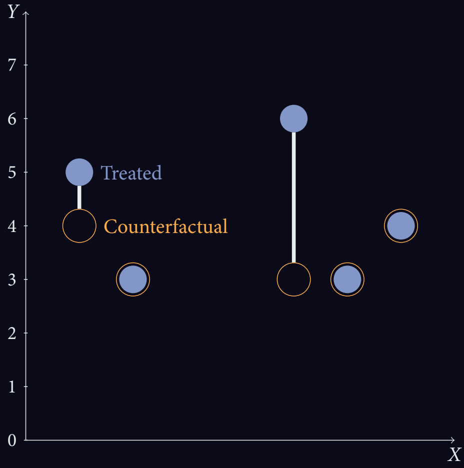
Glynn and Kashin (2018) shows that under these conditions the front-door estimator simplifies to the following:

$$\widehat{ATT}_{\text{Front-door}} = \text{E}[Y|a_1] - \sum_x \underbrace{\text{E}[Y|a_1, m_0, x]}_{\text{Treated non-compliers}} \cdot P(x|a_1)$$

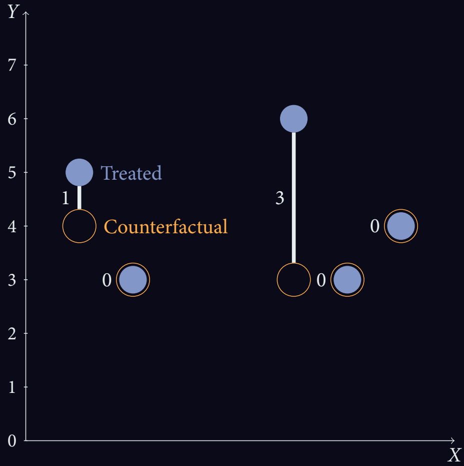
Visualizing the ATT



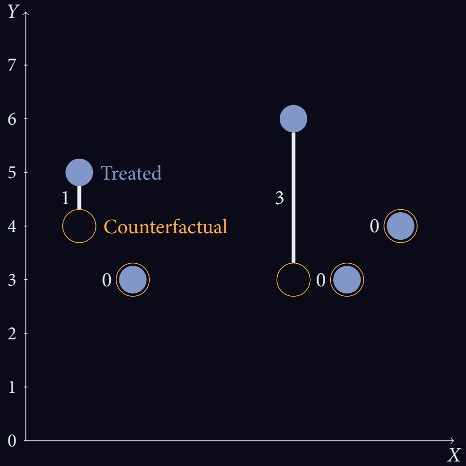
Each treated unit has an associated counterfactual



...and thus an individual causal effect

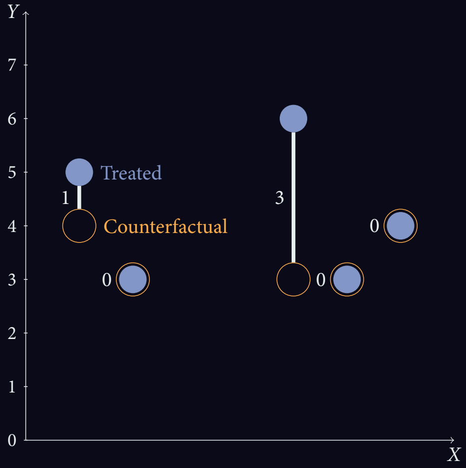


The ATT is an average of individual causal effects



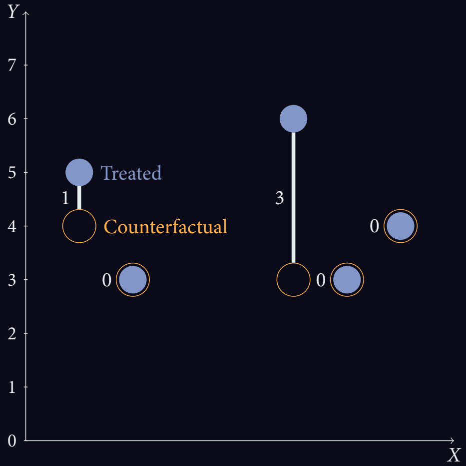
$$ATT = \frac{1+0+3+0+0}{5} = \frac{4}{5}$$

General expression for ATT



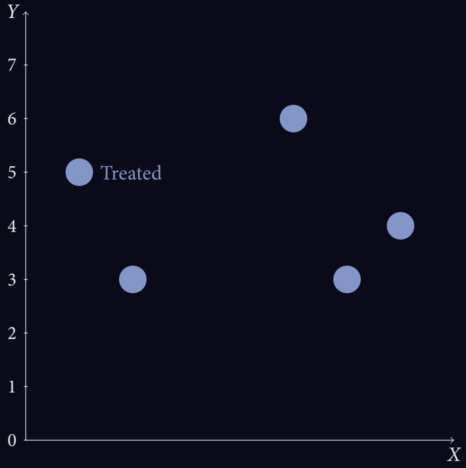
$$ATT = \mathrm{E}[Y - Y(a_0)|a_1]$$

General expression for ATT

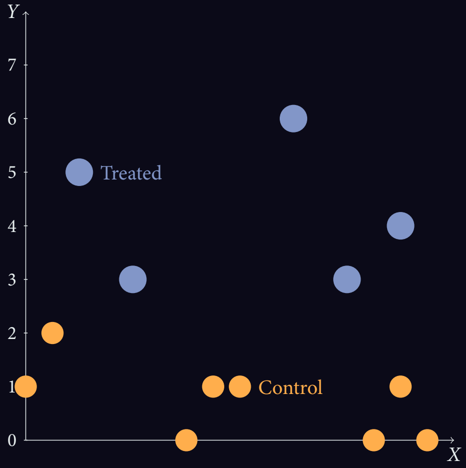


$$ATT = \mathrm{E}[Y - Y(a_0)|a_1] = \mathrm{E}[Y|a_1] - \mathrm{E}[Y(a_0)|a_1]$$

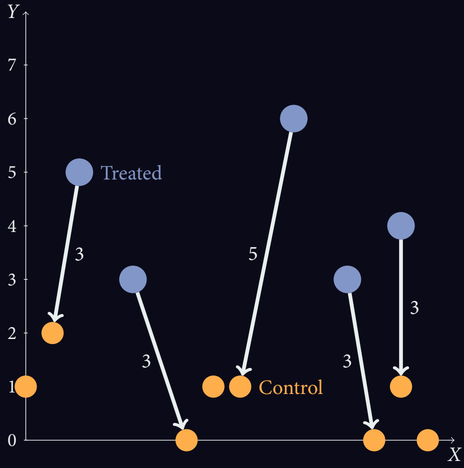
How does one usually estimate ATT?



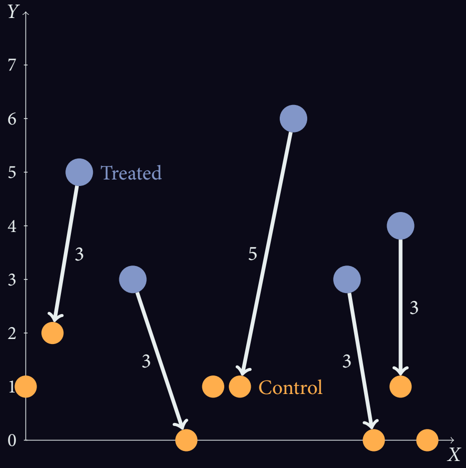
Find a group of control units



Perhaps match each treated unit to nearest control unit

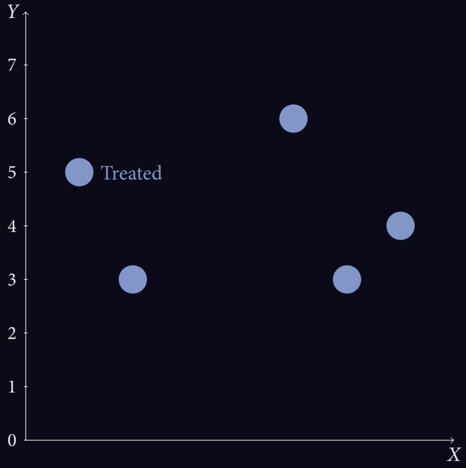


Perhaps match each treated unit to nearest control unit

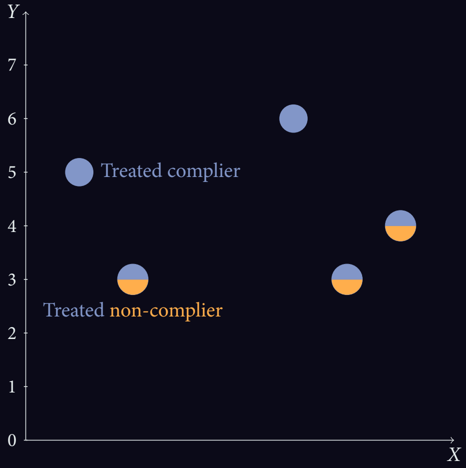


$$\widehat{ATT} = \frac{3+3+5+3+3}{5} = \frac{17}{5}$$

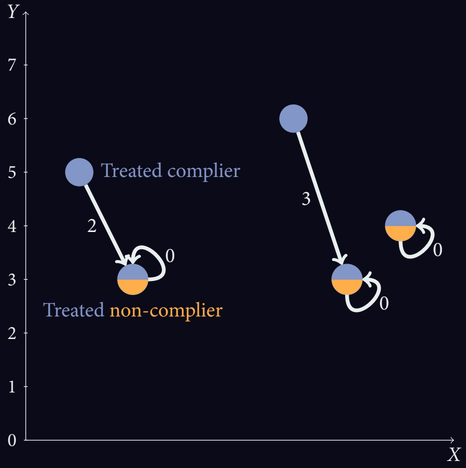
What if no (comparable) control group?



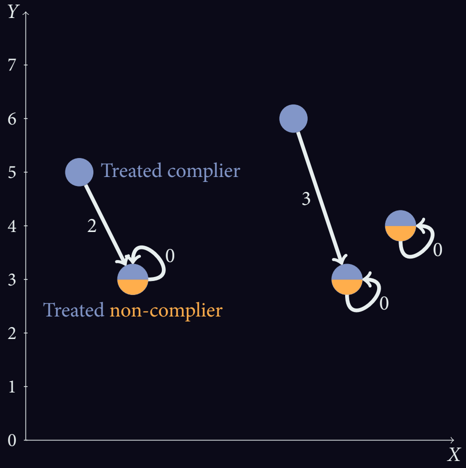
Observe compliance for each treated unit



Match each treated unit to nearest treated non-complier

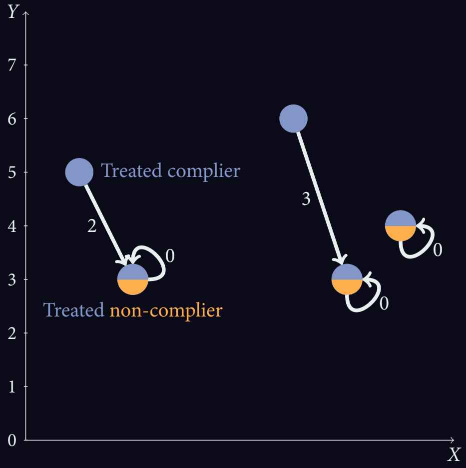


Front-door estimator (with one-sided noncompliance)



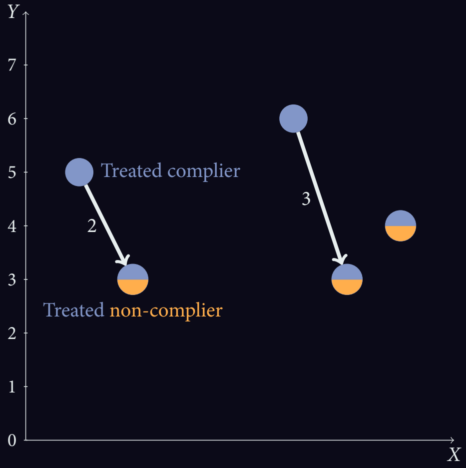
$$\widehat{ATT} = \frac{2+0+3+0+0}{5} = 1$$

Front-door estimator (with one-sided noncompliance)

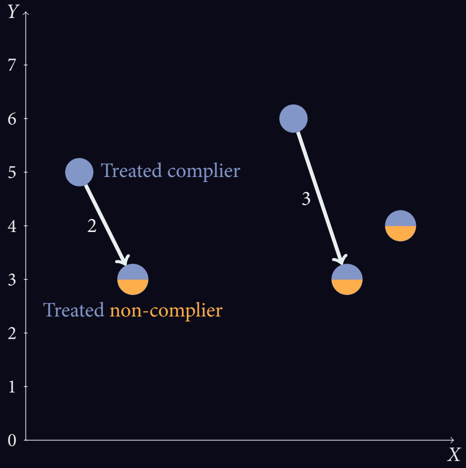


$$\widehat{ATT} = \mathrm{E}[Y|a_1] - \sum_x \mathrm{E}[Y|a_1, m_0, x] \cdot P(x|a_1)$$

"Effect of attendance" conceptualization

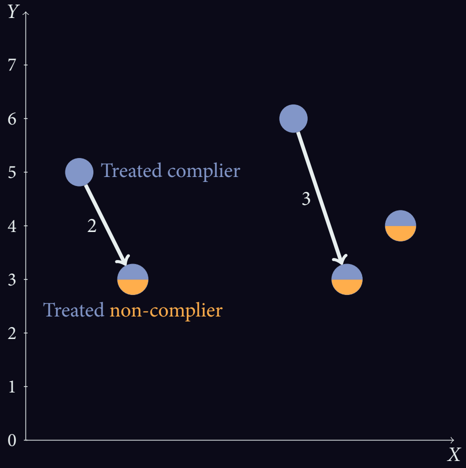


"Effect of attendance" conceptualization



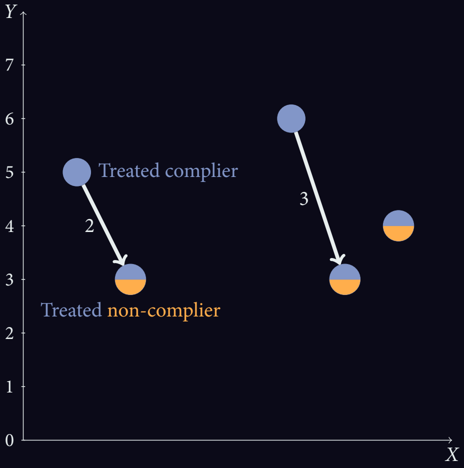
$$\widehat{ATT} = \frac{2+3}{2}$$

"Effect of attendance" conceptualization



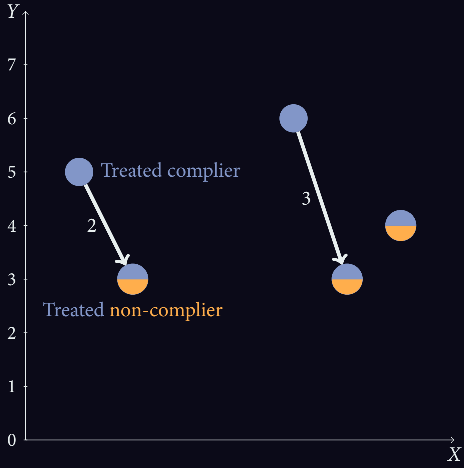
$$\widehat{ATT} = \frac{2}{5} \cdot \frac{2+3}{2}$$

"Effect of attendance" conceptualization



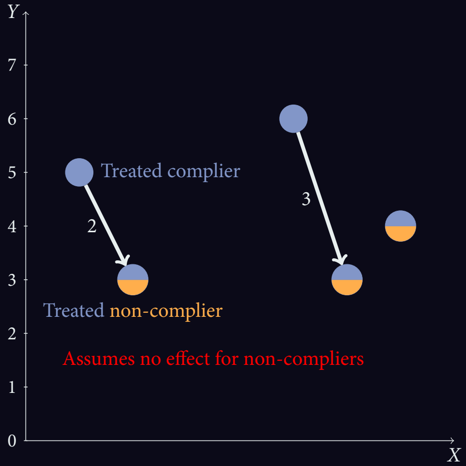
$$\widehat{ATT} = \frac{2}{5} \cdot \frac{2+3}{2} = 1$$

"Effect of attendance" conceptualization



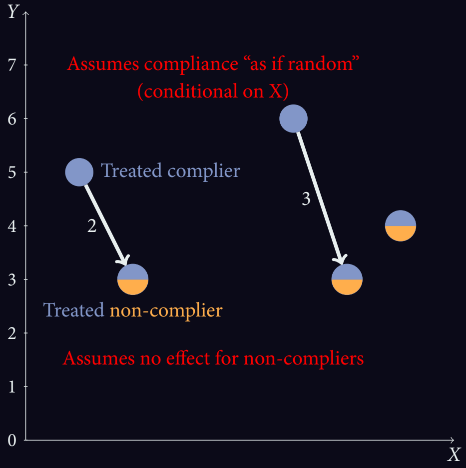
$$\widehat{ATT} = \sum_x P(x|a_1) \cdot P(m_1|a_1, x) \cdot \{E[Y|a_1, m_1, x] - E[Y|a_1, m_0, x]\}$$

"Effect of attendance" conceptualization



$$\widehat{ATT} = \sum_x P(x|a_1) \cdot P(m_1|a_1, x) \cdot \{E[Y|a_1, m_1, x] - E[Y|a_1, m_0, x]\}$$

"Effect of attendance" conceptualization



$$\widehat{ATT} = \sum_x P(x|a_1) \cdot P(m_1|a_1, x) \cdot \{E[Y|a_1, m_1, x] - E[Y|a_1, m_0, x]\}$$

Examples

See if we can recover experimental benchmark.

The National JTPA Study

Job training evaluation program with experimental treatments and controls, compliance information, covariates, and an observational control group:

The National JTPA Study

Job training evaluation program with experimental treatments and controls, compliance information, covariates, and an observational control group:

- Treatment: allowed to receive JTPA services or not

Sign up

The National JTPA Study

Job training evaluation program with experimental treatments and controls, compliance information, covariates, and an observational control group:

- Treatment: allowed to receive JTPA services or not
- Outcome: 18-month post-program earnings

$$\text{Sign up} \longrightarrow \text{Earnings}$$

The National JTPA Study

Job training evaluation program with experimental treatments and controls, compliance information, covariates, and an observational control group:

- Treatment: allowed to receive JTPA services or not
- Outcome: 18-month post-program earnings
- Compliance: program participation (one-sided noncompliance)

Sign up $\longrightarrow$ Show up $\longrightarrow$ Earnings

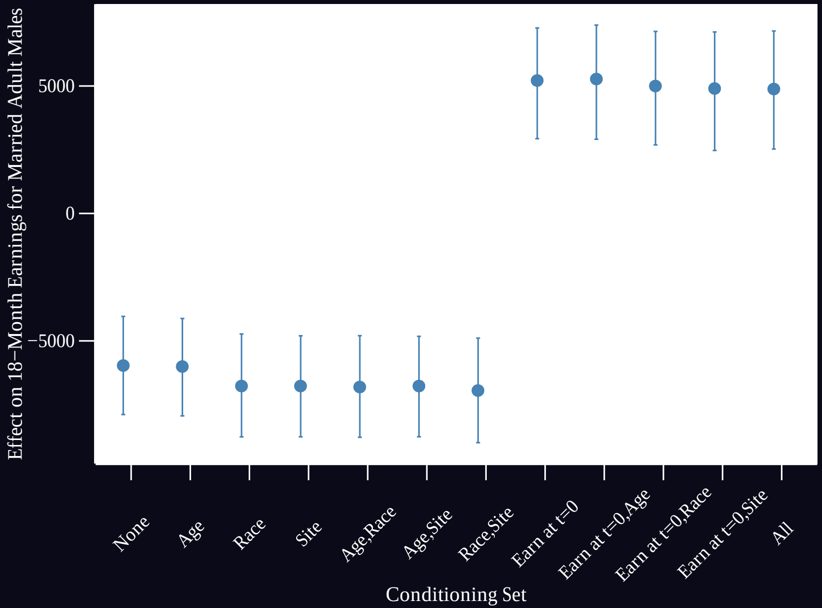
The National JTPA Study

Job training evaluation program with experimental treatments and controls, compliance information, covariates, and an observational control group:

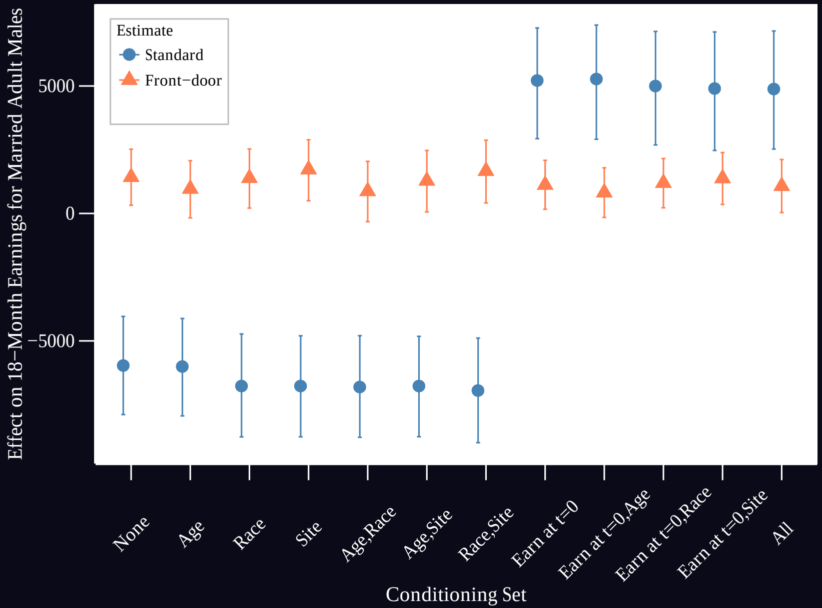
- Treatment: allowed to receive JTPA services or not
- Outcome: 18-month post-program earnings
- Compliance: program participation (one-sided noncompliance)
- Group of Interest: married adult males

$$\text{Sign up} \longrightarrow \text{Show up} \longrightarrow \text{Earnings}$$

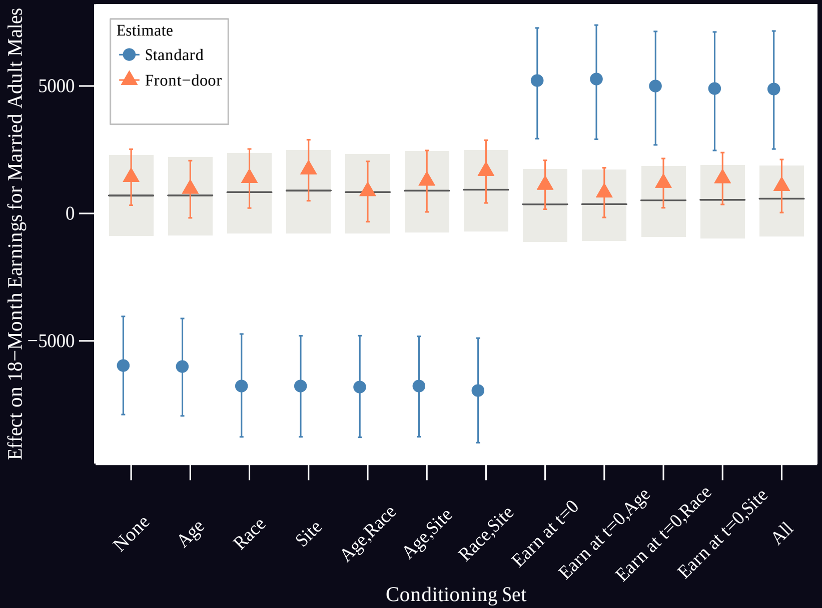
Are the standard estimates believable?



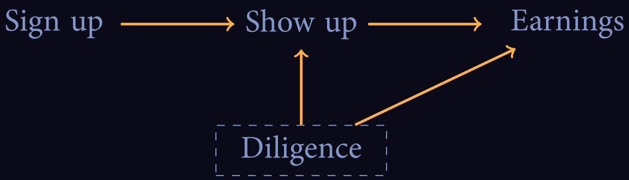
Which would you choose?



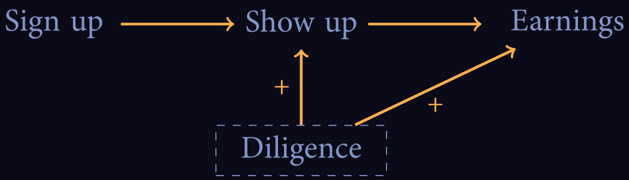
Front-door outperforms standard approach



Not surprising that front-door exhibits small positive bias



Not surprising that front-door exhibits small positive bias



Get-Out-the-Vote (GOTV) Phone Studies

Phone GOTV evaluation studies with experimental treatments and controls and compliance information:

Get-Out-the-Vote (GOTV) Phone Studies

Phone GOTV evaluation studies with experimental treatments and controls and compliance information:

- Treatment: phone call encouraging turnout

Call

Get-Out-the-Vote (GOTV) Phone Studies

Phone GOTV evaluation studies with experimental treatments and controls and compliance information:

- Treatment: phone call encouraging turnout
- Outcome: vote vs not vote

$$\text{Call} \longrightarrow \text{Vote}$$

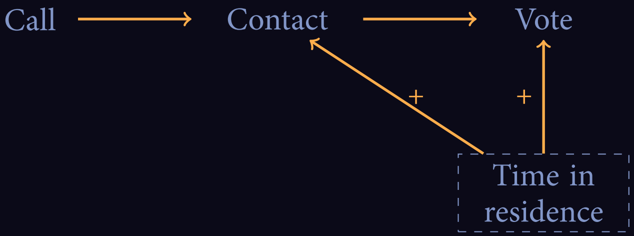
Get-Out-the-Vote (GOTV) Phone Studies

Phone GOTV evaluation studies with experimental treatments and controls and compliance information:

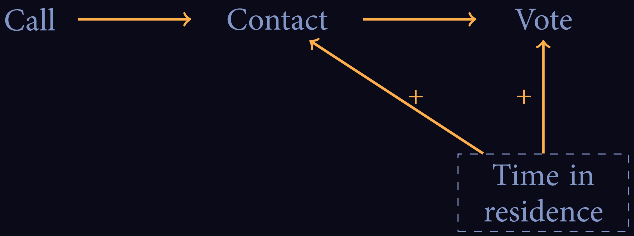
- Treatment: phone call encouraging turnout
- Outcome: vote vs not vote
- Compliance: contact (one-sided noncompliance)

$$\text{Call} \longrightarrow \text{Contact} \longrightarrow \text{Vote}$$

Front-door for GOTV studies

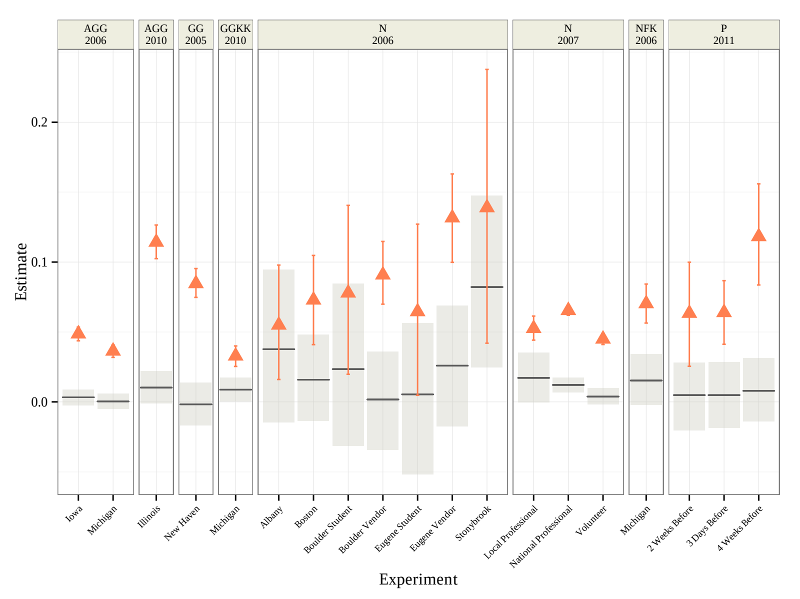


Front-door for GOTV studies



Front-door likely to exhibit positive bias.

Front-door estimates exhibit positive bias



Outline

Motivation and Introduction

Front-door for ATT

Front-Door Difference-in-Differences

General form of front-door difference-in-differences

Suppose we have a “differencing group” of observations for which we assume there is no treatment effect (e.g., pre-treatment outcomes for compliers and non-compliers)

General form of front-door difference-in-differences

Suppose we have a "differencing group" of observations for which we assume there is no treatment effect (e.g., pre-treatment outcomes for compliers and non-compliers)

Within levels of X:

1. "effect of attendance" for the group of interest
2. "effect of attendance" for the differencing group
3. Proportion of compliers in group of interest

$$\text{FD-DID} = (3) \cdot [(1) - (2)]$$

General form of front-door difference-in-differences

Suppose we have a “differencing group” of observations for which we assume there is no treatment effect (e.g., pre-treatment outcomes for compliers and non-compliers)

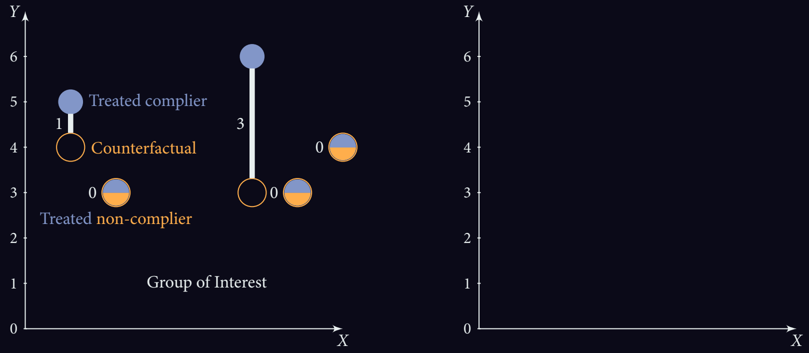
Within levels of X:

1. “effect of attendance” for the group of interest
2. “effect of attendance” for the differencing group
3. Proportion of compliers in group of interest

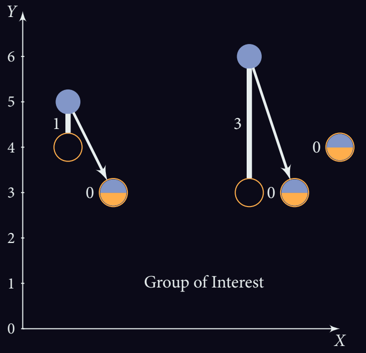
$$\text{FD-DID} = (3) \cdot [(1) - (2)]$$

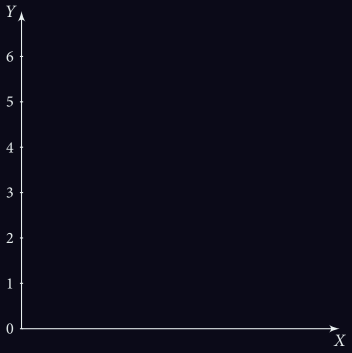
Assumes bias due to “not as if random” the same for group of interest and differencing group.

Revisiting the stylized example

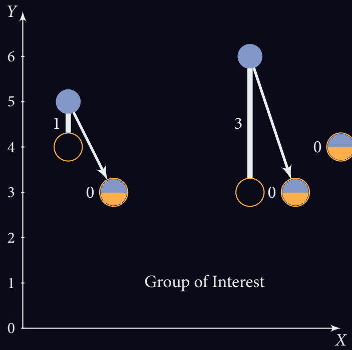


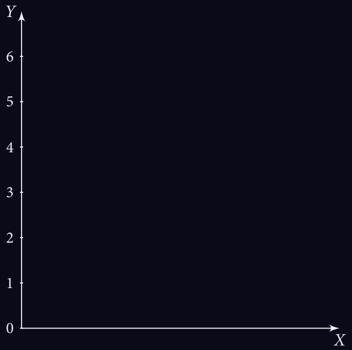
Positive bias in "effect of attendance"





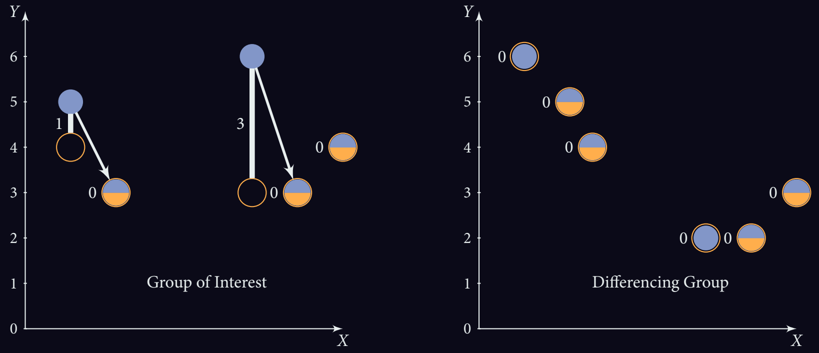
Positive bias in "effect of attendance"





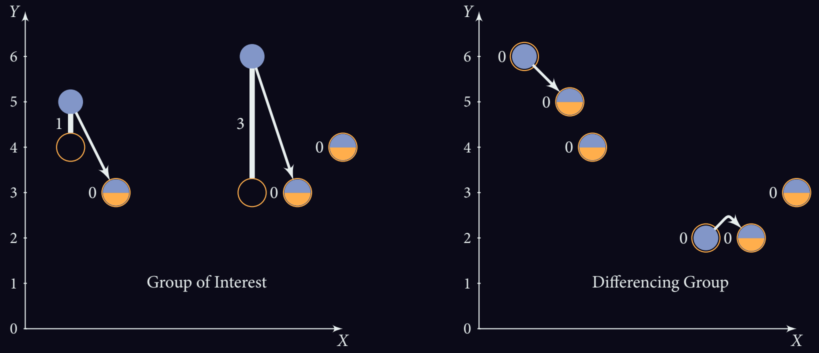
$$\frac{2+3}{2}$$

Find differencing group where believe effect is zero



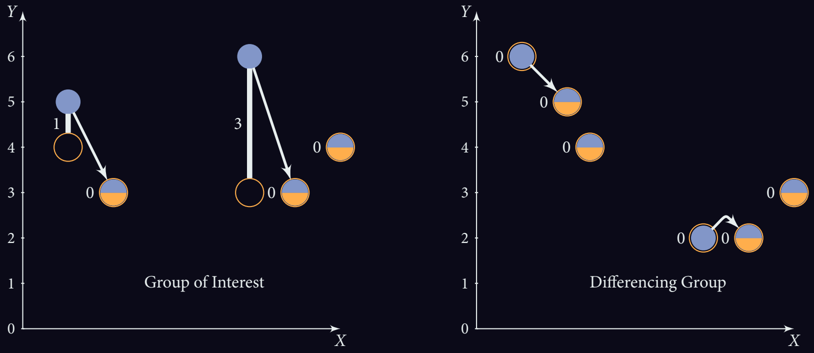
$$\frac{2+3}{2}$$

Non-zero "effect of attendance" estimate is evidence of bias



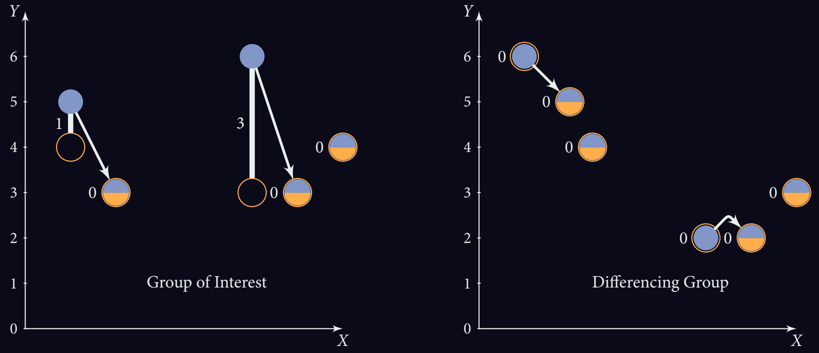
$$\frac{2+3}{2}$$

Non-zero "effect of attendance" estimate is evidence of bias



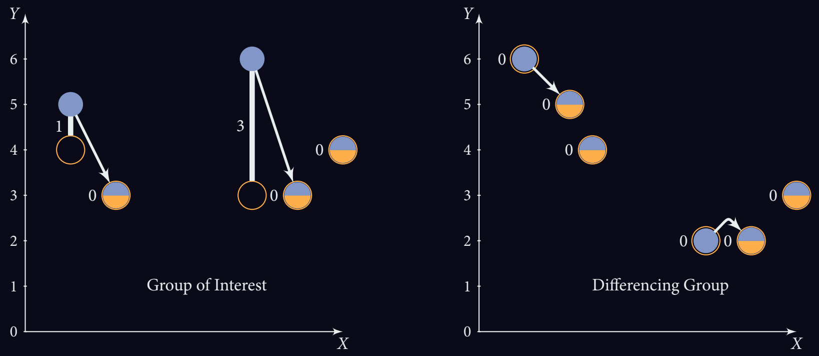
$$\frac{2+3}{2} \qquad \frac{1+0}{2}$$

Remove bias from group of interest



$$\frac{2}{5} \cdot \left[\frac{2+3}{2} - \frac{1+0}{2} \right]$$

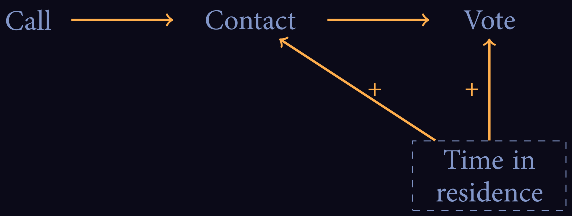
Front-door diff-in-diff enables point identification



$$\frac{2}{5} \cdot \left[\frac{2+3}{2} - \frac{1+0}{2} \right] = \frac{4}{5}$$

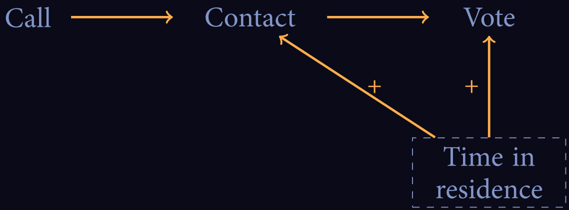
Want to remove bias due to unmeasured confounder

Group of interest:

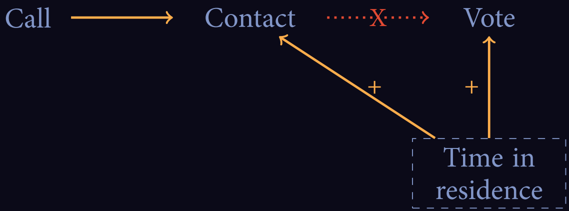


Find differencing group where believe effect is zero

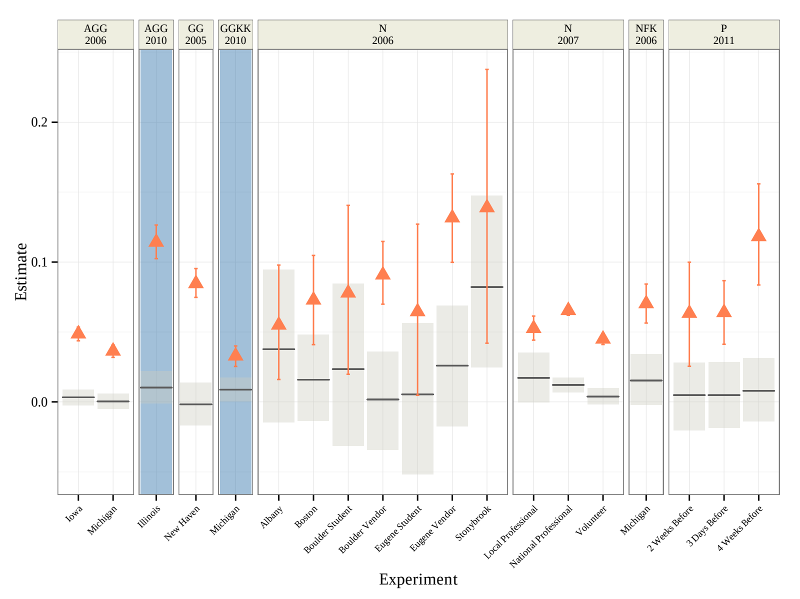
Group of interest:



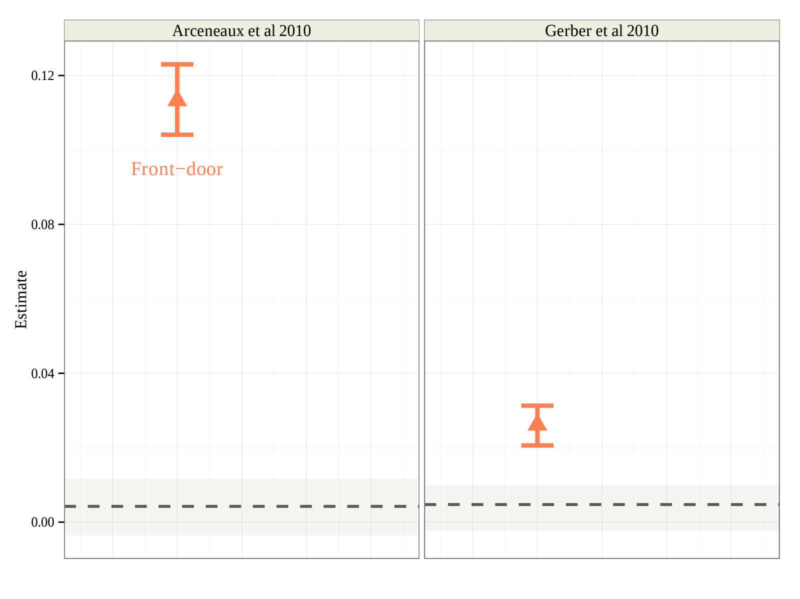
Differencing group:



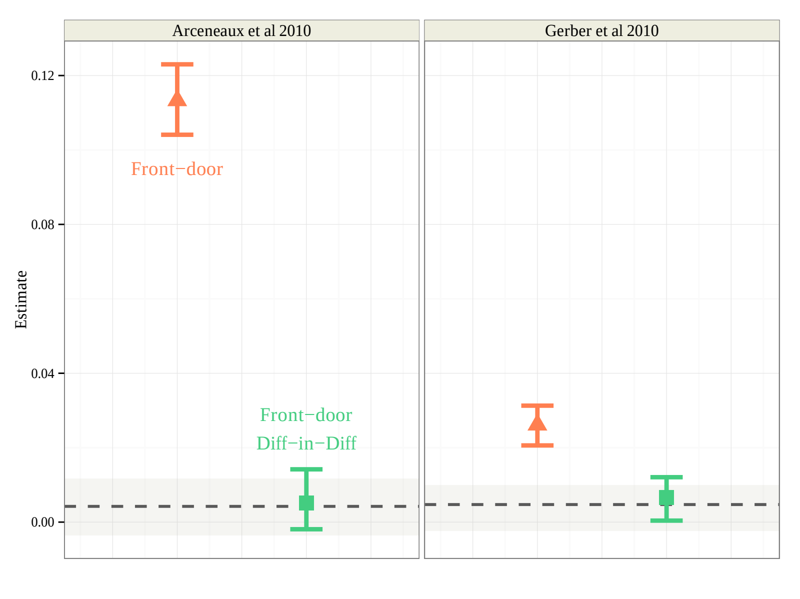
Placebo as prototypical differencing group



Prototypical FDDID as proof of concept



Prototypical FDDID as proof of concept



How to choose the differencing group

In practice, it may be difficult to choose a differencing group:

1. pre-treatment outcomes (does not allow repeated cross sections)

How to choose the differencing group

In practice, it may be difficult to choose a differencing group:

1. pre-treatment outcomes (does not allow repeated cross sections)
2. study of heterogeneous effects from prior experiments (e.g., Imai and Strauss (2011) find no effect for 18-19 year olds in a single text message GOTV campaign)

How to choose the differencing group

In practice, it may be difficult to choose a differencing group:

1. pre-treatment outcomes (does not allow repeated cross sections)
2. study of heterogeneous effects from prior experiments (e.g., Imai and Strauss (2011) find no effect for 18-19 year olds in a single text message GOTV campaign)
3. theory/observational studies

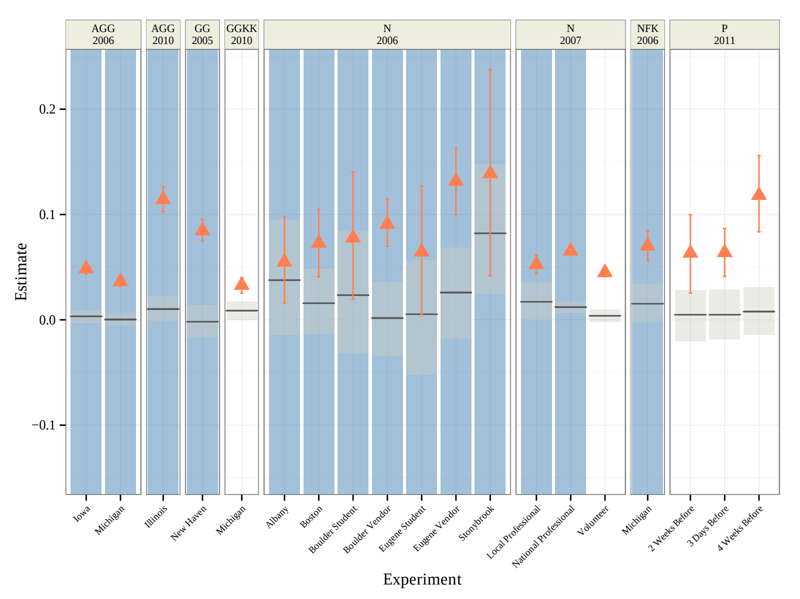
How to choose the differencing group

In practice, it may be difficult to choose a differencing group:

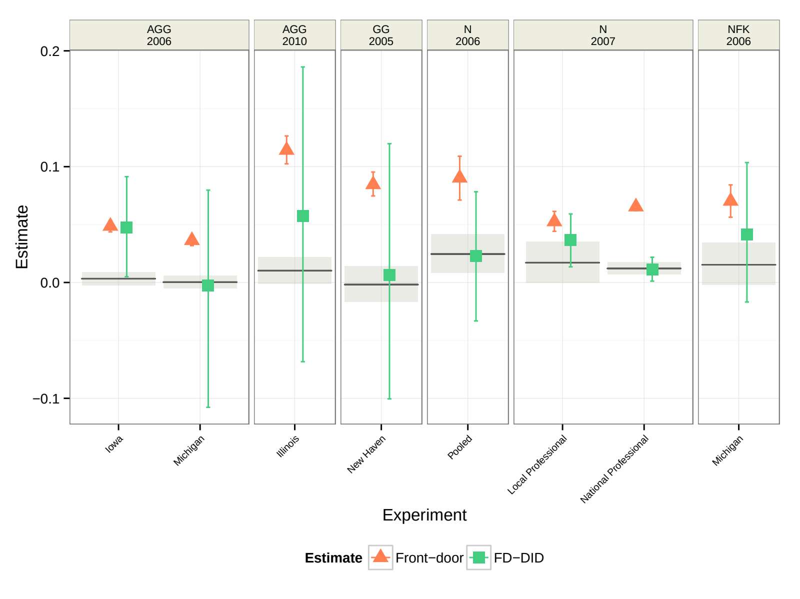
1. pre-treatment outcomes (does not allow repeated cross sections)
2. study of heterogeneous effects from prior experiments (e.g., Imai and Strauss (2011) find no effect for 18-19 year olds in a single text message GOTV campaign)
3. theory/observational studies

Assumption of no effect can be relaxed for bounding analysis.

FD-DID USING 18-19 yo differencing group



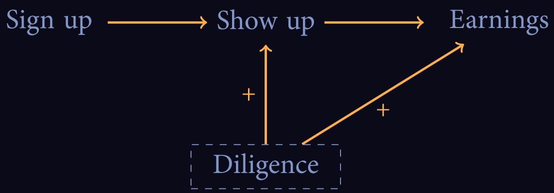
FD-DID using 18-19 yo differencing group



Revisiting the National JTPA Study

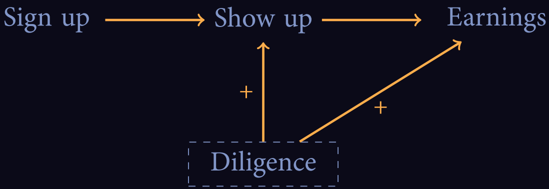
Revisiting the National JTPA Study

Married adult males (group of interest):

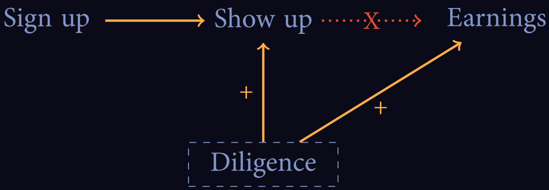


Revisiting the National JTPA Study

Married adult males (group of interest):

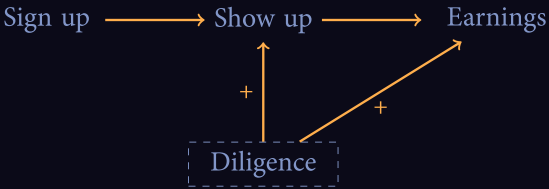


Single adult males (differencing group):

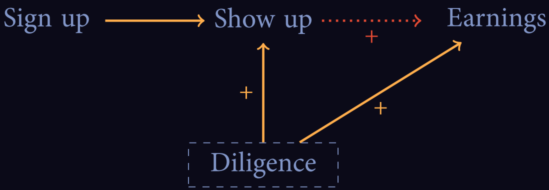


Lower bound if small positive effect for differencing group

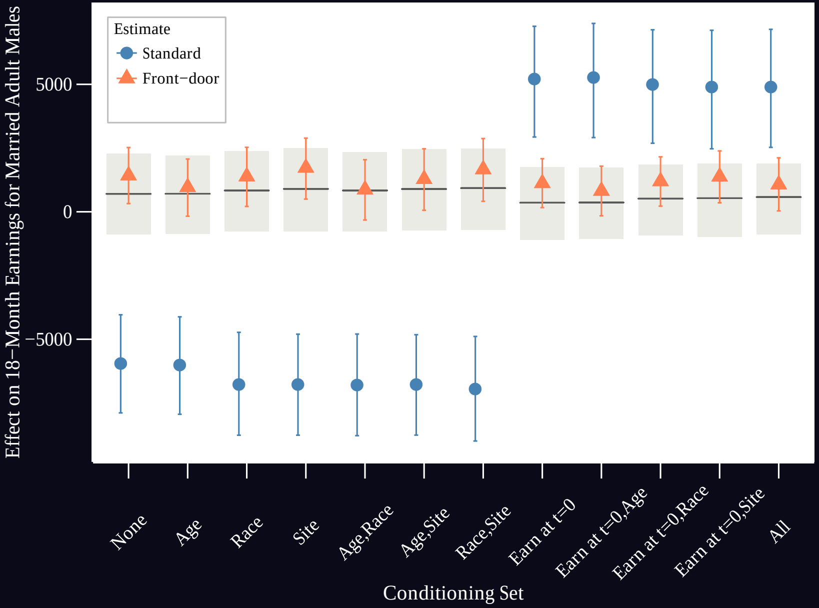
Married adult males (group of interest):



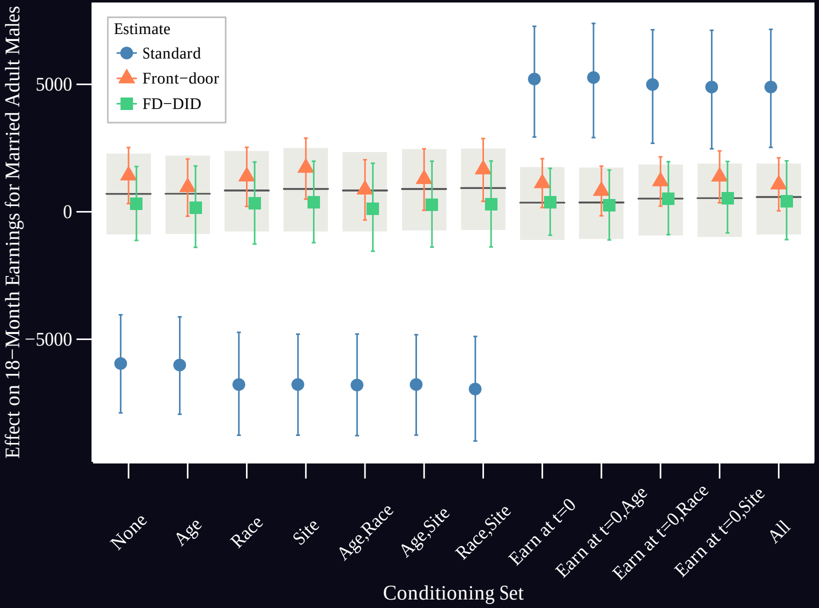
Single adult males (differencing group):



Front-door estimates exhibit slight positive bias



Front-door diff-in-diff provides lower bound



Summary

- Front-door approach uses mediating variables for identification when selection on observables does not hold.

Summary

- Front-door approach uses mediating variables for identification when selection on observables does not hold.
- Requires that M fully mediates the effect and that M is as-if random conditional on X.

Summary

- Front-door approach uses mediating variables for identification when selection on observables does not hold.
- Requires that M fully mediates the effect and that M is as-if random conditional on X.
- May provide useful information when assumptions don't hold exactly.

Summary

- Front-door approach uses mediating variables for identification when selection on observables does not hold.
- Requires that M fully mediates the effect and that M is as-if random conditional on X.
- May provide useful information when assumptions don't hold exactly.
- Can be used in concert with front-door difference-in-differences to bracket.

Summary

- Front-door approach uses mediating variables for identification when selection on observables does not hold.
- Requires that M fully mediates the effect and that M is as-if random conditional on X.
- May provide useful information when assumptions don't hold exactly.
- Can be used in concert with front-door difference-in-differences to bracket.
- In the canonical example (binary M with one-sided non-compliance), only treated units are used.

Thank You

Thank You

References:

Glynn, A.N. and Kashin, K. (2017) "Front-door Difference-in-Differences Estimators." *American Journal of Political Science*. 61 (4): 989–1002.

Glynn, A.N. and Kashin, K. (2018) "Front-door Versus Back-door Adjustment with Unmeasured Confounding: Bias Formulas for Front-door and Hybrid Adjustments with Application to a Job Training Program," *Journal of the American Statistical Association*. 113 (523): 1040–1049.

Proof

$$
\begin{aligned}
B_{a_1}^{fd} &= \mu_{0|a_1}^{fd} - \mu_{0|a_1} \\
&= \sum_x \sum_m P(m|a_0, x) \cdot E[Y|a_1, m, x] \cdot P(x|a_1) \\
&- \sum_x \sum_u E[Y|a_0, x, u] \cdot P(u|x, a_1) \cdot P(x|a_1) \\
&= \sum_x \sum_m P(m|a_0, x) \sum_u E[Y|a_1, m, x, u] \cdot P(u|a_1, m, x) \cdot P(x|a_1) \\
&- \sum_x \sum_u \sum_m E[Y|a_0, m, x, u] \cdot P(m|a_0, x, u) \cdot P(u|a_1, x) \cdot P(x|a_1) \\
&= \sum_x P(x|a_1) \sum_m \sum_u P(m|a_0, x) \cdot E[Y|a_1, m, x, u] \cdot P(u|a_1, m, x) \\
&- \sum_x P(x|a_1) \sum_m \sum_u P(m|a_0, x, u) \cdot E[Y|a_0, m, x, u] \cdot P(u|a_1, x)
\end{aligned}
$$

Program impact likely greater for married males

- Cross-sectional evidence that marriage increases male wages (Blackburn and Korenman, 1994; Korenman and Neumark, 1991; Krashinsky, 2004; Shoeni, 1995)
  - Studies generally estimate 10-40% marriage premium
- Some evidence that positive impact of marriage of productivity (Korenman and Neumark, 1991; Hellerstein et al., 1999; Mehay and Bowman, 2005)