Causal mediation workshop

Nick Graetz[1]

[1] Princeton University, Eviction Lab

2/20/2023

Outline

1. Theory
2. Conventional mediation
3. Causal estimands
4. Mediation with post-treatment confounding
5. R examples: mediation using *CMAverse* and *paths*
6. Advanced topics in mediation: multiple mediators, path-specific effects, time-varying and cumulative estimands
7. Theory

Background

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

Background

- A central aim in sociological and demographic analysis is *explaining* the source of differences.
- We love to theorize mechanisms: Why is X related to Y?

Background

- A central aim in sociological and demographic analysis is *explaining* the source of differences.
- We love to theorize mechanisms: Why is X related to Y?
- Quant studies: *"It looks like there is an effect of X on Y, but it goes away when I control for M"*

Background

- But quantitative causal mediation is really hard and requires tons of assumptions - **theoretically, not just implementation.**

Background

- But quantitative causal mediation is really hard and requires tons of assumptions - **theoretically, not just implementation.**
- Most modern methodological developments explore relatively simple use cases.

Background

- But quantitative causal mediation is really hard and requires tons of assumptions - **theoretically, not just implementation.**
- Most modern methodological developments explore relatively simple use cases.
- Should we do it anyway?

Estimands

American Sociological Review
2021, Vol. 86(3) 532–565

DOI:10.1177/00031224211004187
journals.sagepub.com/home/asr
SAGE

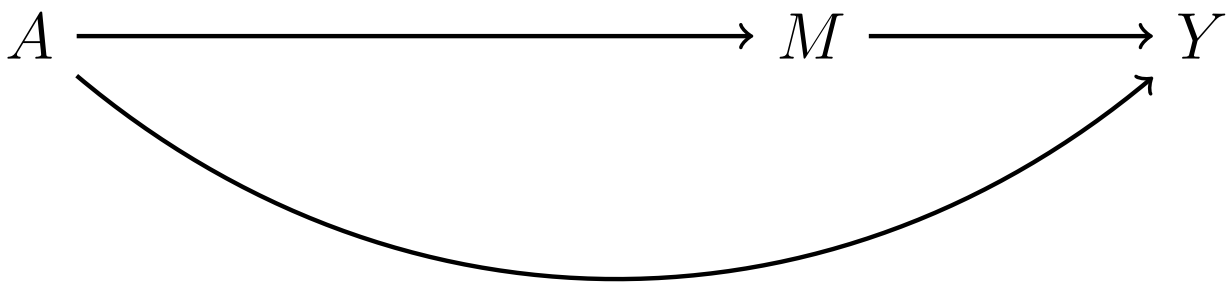
What Is Your Estimand? Defining the Target Quantity Connects Statistical Evidence to Theory

Ian Lundberg,[a] Rebecca Johnson,[b] and Brandon M. Stewart[a]

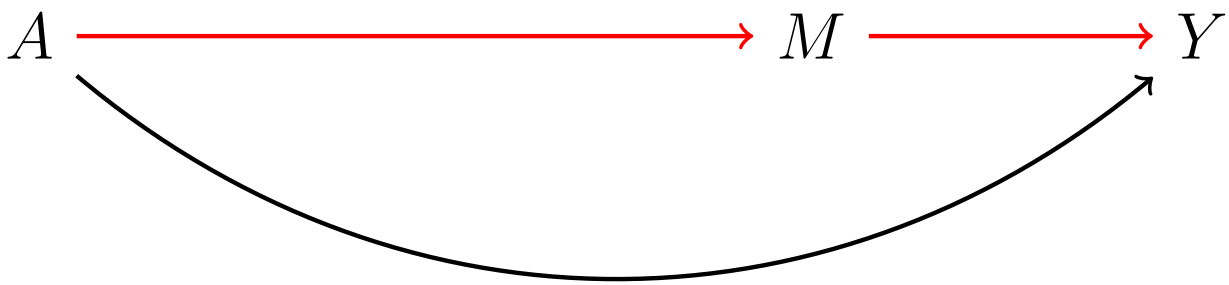
Abstract

We make only one point in this article. Every quantitative study must be able to answer the question: what is your estimand? The estimand is the target quantity—the purpose of the statistical analysis. Much attention is already placed on how to do estimation; a similar degree of care should be given to defining the thing we are estimating. We advocate that authors state the central quantity of each analysis—the theoretical estimand—in precise terms that exist outside of any statistical model. In our framework, researchers do three things: (1) set a theoretical estimand, clearly connecting this quantity to theory; (2) link to an empirical estimand, which is informative about the theoretical estimand under some identification assumptions; and (3) learn from data. Adding precise estimands to research practice expands the space of theoretical questions, clarifies how evidence can speak to those questions, and unlocks new tools for estimation. By grounding all three steps in a precise statement of the target quantity, our framework connects statistical evidence to theory.

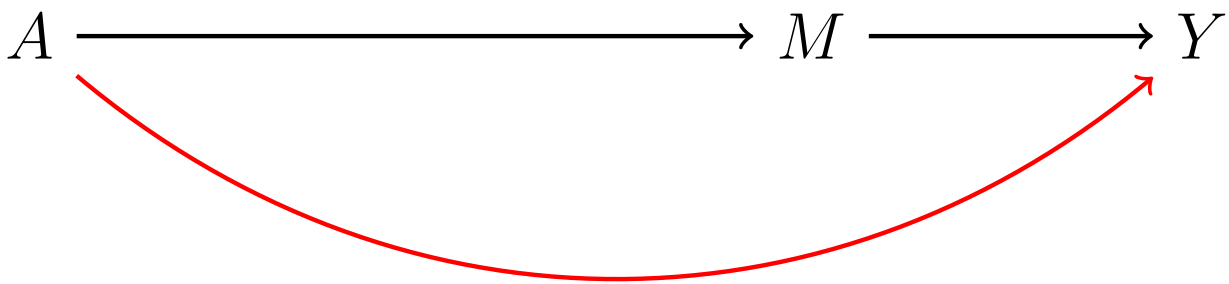
Direct effects



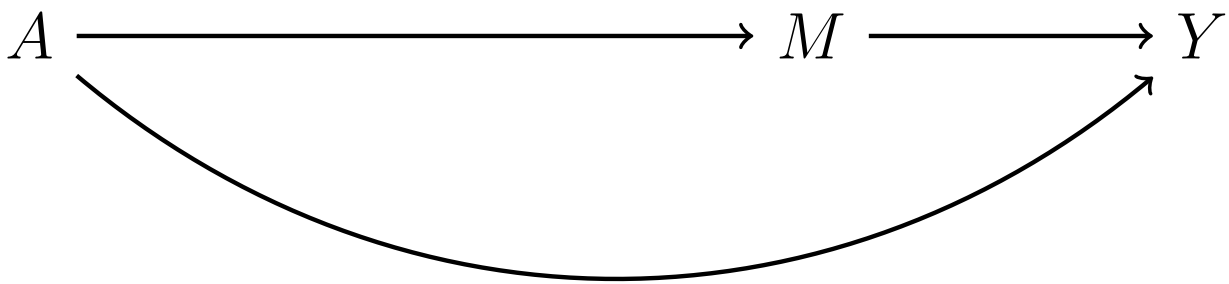
Direct effects



Direct effects

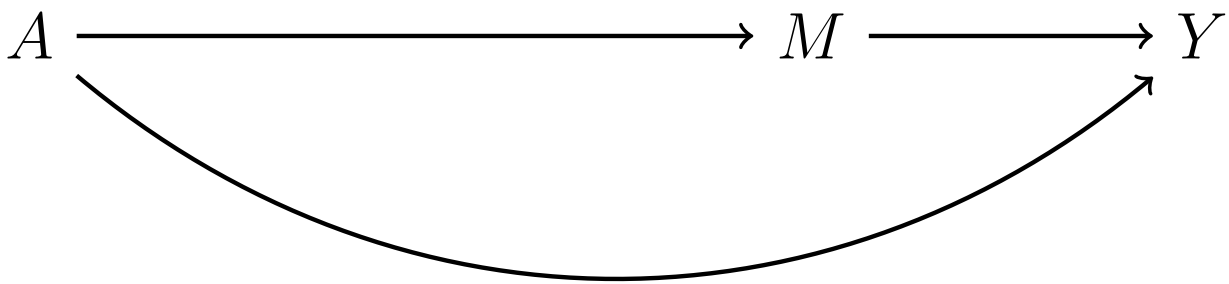


Direct effects



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

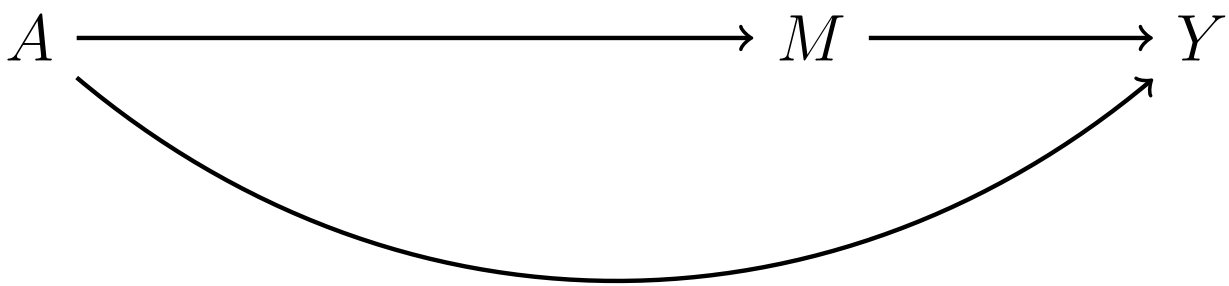
Direct effects



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

- This is important when thinking about complex exposures! (see Sen & Wasow 2016: Race as a Bundle of Sticks, Kohler-Hausmann 2019)

Direct effects

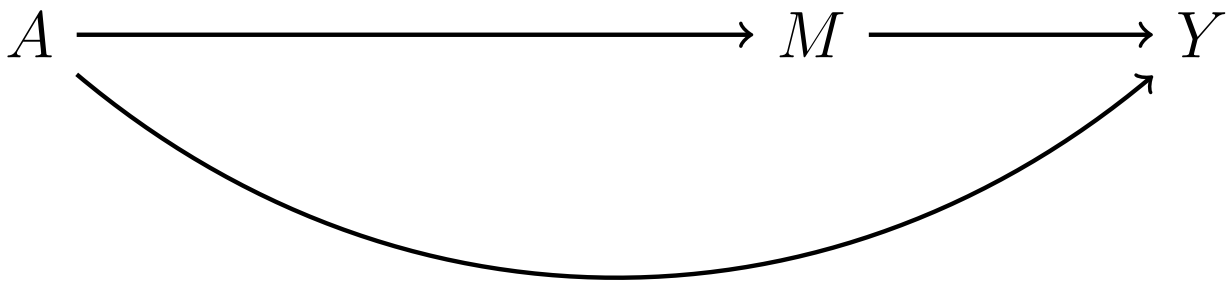


$$Y = \beta_0 + \beta_1 A$$

$$Y = \theta_0 + \theta_1 A + \theta_2 M$$

$$\text{Proportion mediated} = (\beta_1 - \theta_1)/\beta_1$$

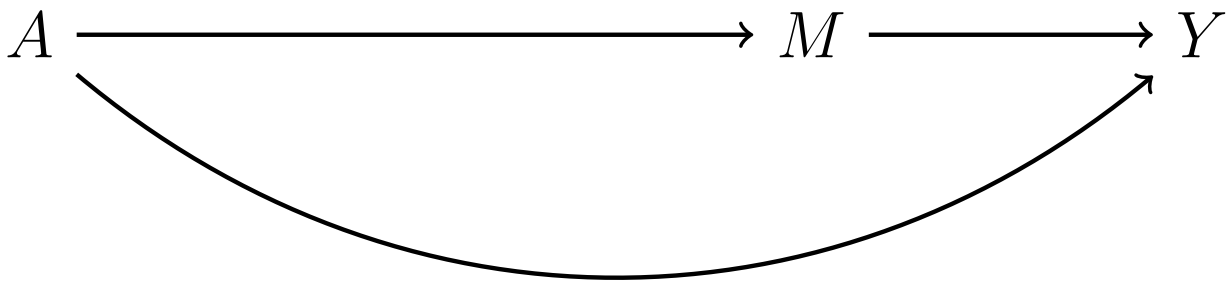
Direct effects



$$Y = \beta_0 + \beta_1 A$$
$$Y = \theta_0 + \theta_1 A + \theta_2 M$$
$$\text{Proportion mediated} = (\beta_1 - \theta_1)/\beta_1$$

- Baron-Kenny mediation (1986): 115,563 citations.

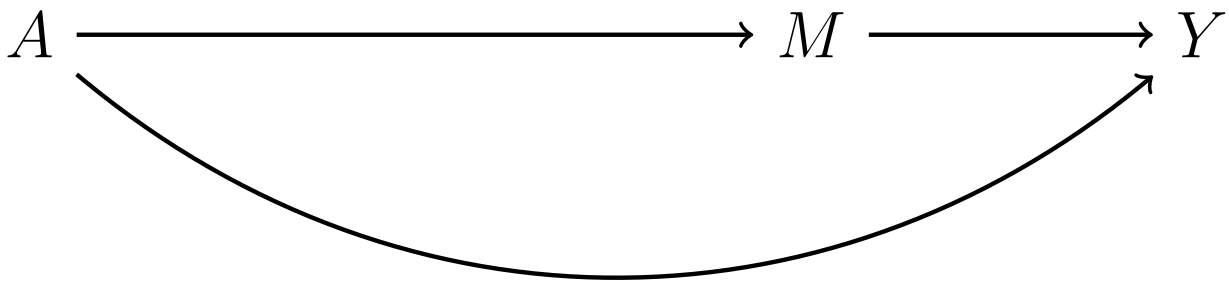
Direct effects



$$Y = \beta_0 + \beta_1 A$$
$$Y = \theta_0 + \theta_1 A + \theta_2 M$$
$$\text{Proportion mediated} = (\beta_1 - \theta_1)/\beta_1$$

- Baron-Kenny mediation (1986): 115,563 citations.
- Non-linear model? Use KHB adjustment.

Direct effects



$$Y = \beta_0 + \beta_1 A$$
$$Y = \theta_0 + \theta_1 A + \theta_2 M$$
$$\text{Proportion mediated} = (\beta_1 - \theta_1)/\beta_1$$

- When is this causal?

$$\text{CDE}_{M=m^*} = E[Y_{am^*}] - E[Y_{a^*m^*}]$$

Causal inference

- Causal inference in the potential outcomes tradition:

Causal inference

- Causal inference in the potential outcomes tradition:
- *Instrument-based*: assumptions revolve around validity of instrument; direct randomization to treatment is the most plausible instrument.

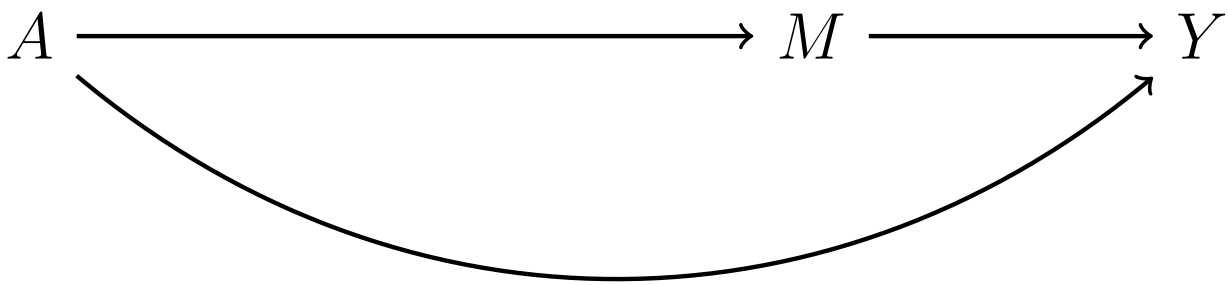
Causal inference

- Causal inference in the potential outcomes tradition:
- *Instrument-based*: assumptions revolve around validity of instrument; direct randomization to treatment is the most plausible instrument.
- *Confounder-control*: assumptions revolve around unmeasured confounding.

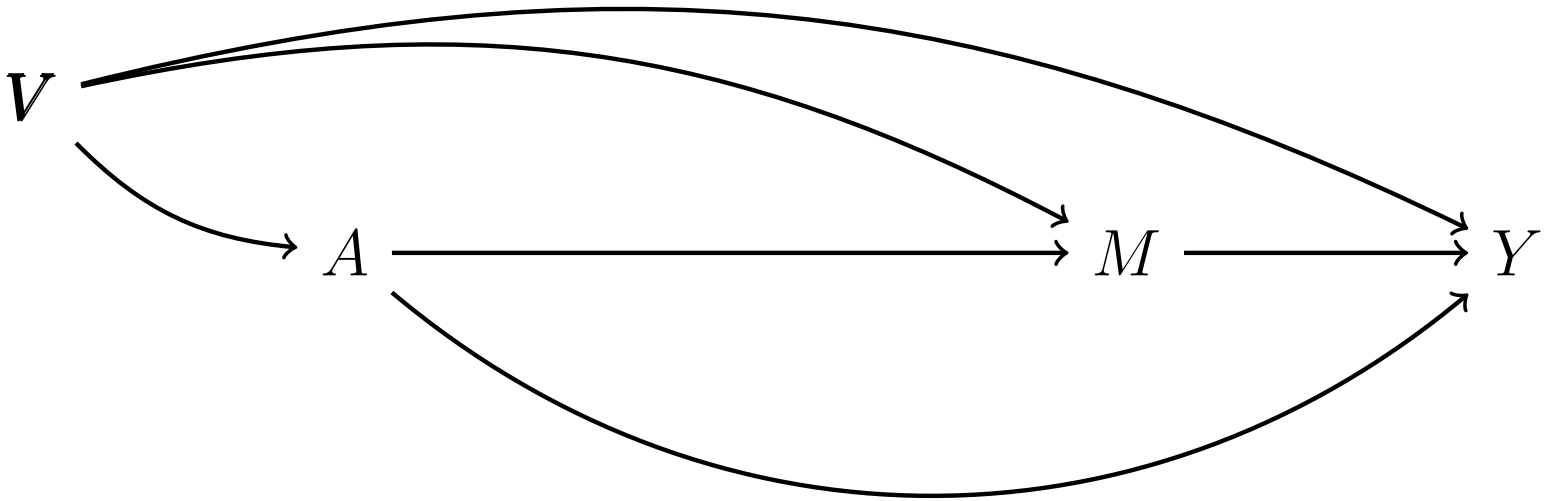
Causal inference

- Causal inference in the potential outcomes tradition:
- *Instrument-based*: assumptions revolve around validity of instrument; direct randomization to treatment is the most plausible instrument.
- *Confounder-control*: assumptions revolve around unmeasured confounding.
- Mediation analysis with instrument-based methods is very hard. There are people working on study designs ("crossover designs"), but I'm not going to talk about this. See Imai, Tingley and Yamamoto (2013, JRSS-A).

Causal mediation

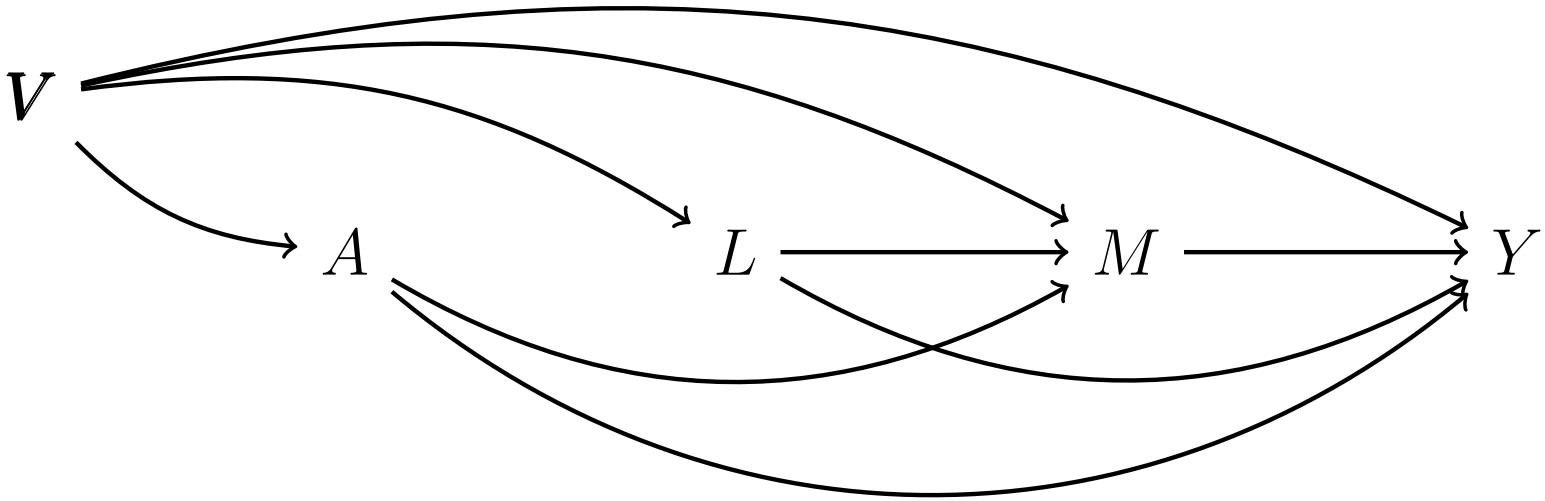


Causal mediation



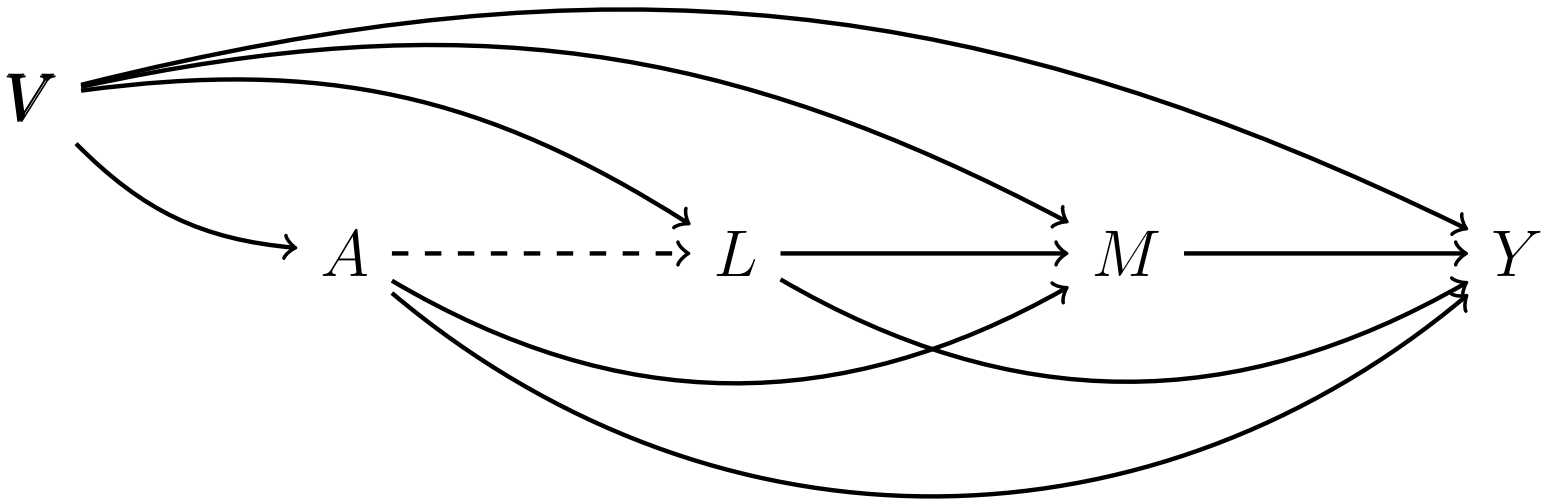
1. No unobserved confounding of $A \longrightarrow Y$ ($\boldsymbol{V}$ contains all relevant confounding variables of A and Y).
2. No unobserved confounding of $A \longrightarrow M$ ($\boldsymbol{V}$ contains all relevant confounding variables of A and M).

Causal mediation



1. No unobserved confounding of $A \longrightarrow Y$ ($\boldsymbol{V}$ contains all relevant confounding variables of A and Y).
2. No unobserved confounding of $A \longrightarrow M$ ($\boldsymbol{V}$ contains all relevant confounding variables of A and M).
3. No unobserved post-treatment confounding of $M \longrightarrow Y$ ($\boldsymbol{L}$ contains all relevant post-treatment confounding variables of M and Y).

Causal mediation



1. No unobserved confounding of $A \longrightarrow Y$ ($\boldsymbol{V}$ contains all relevant confounding variables of A and Y).
2. No unobserved confounding of $A \longrightarrow M$ ($\boldsymbol{V}$ contains all relevant confounding variables of A and M).
3. No unobserved post-treatment confounding of $M \longrightarrow Y$ ($\boldsymbol{L}$ contains all relevant post-treatment confounding variables of M and Y).
4. No post-treatment confounders ($\boldsymbol{L}$) are affected by A (the dotted arrow between A and $\boldsymbol{L}$ does not exist; the "cross-world independence assumption").

Causal mediation

- "Post-treatment confounding" or "exposure-induced mediator-outcome confounding."

Causal mediation

- "Post-treatment confounding" or "exposure-induced mediator-outcome confounding."
- Best description of this is Wodtke, Harding, & Elwert (2011, ASR).

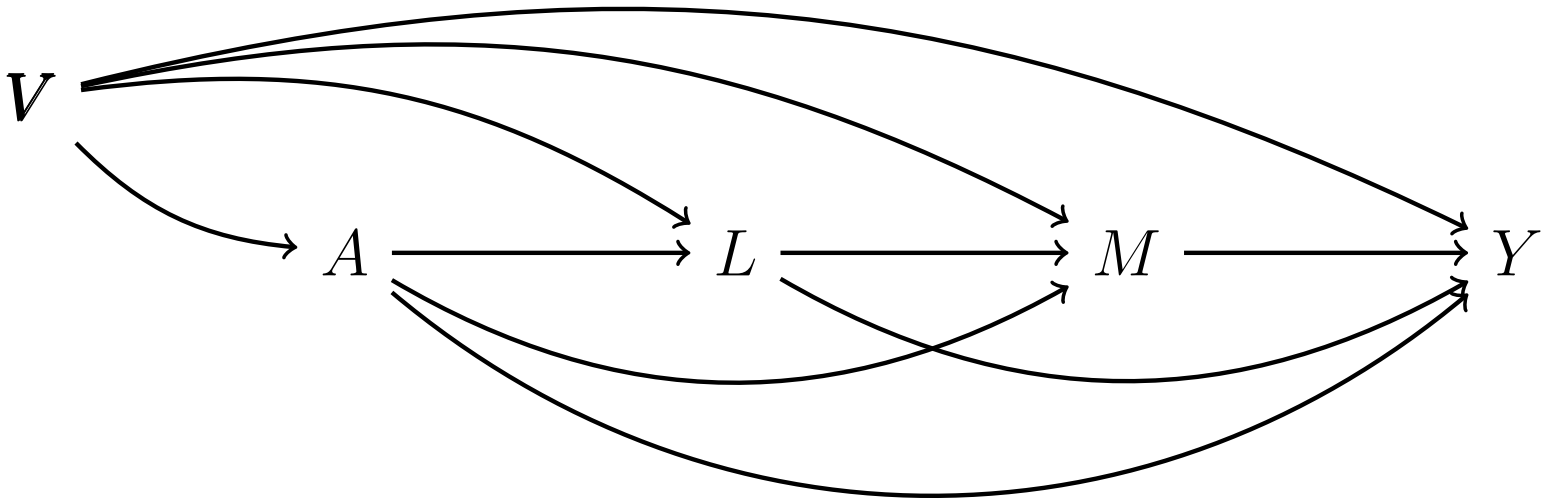
Causal mediation

- "Post-treatment confounding" or "exposure-induced mediator-outcome confounding."
- Best description of this is Wodtke, Harding, & Elwert (2011, ASR).
- This is impossible to handle with Baron-Kenny mediation. You'll likely always overestimate the proportion of the total effect that is mediated.

Causal mediation

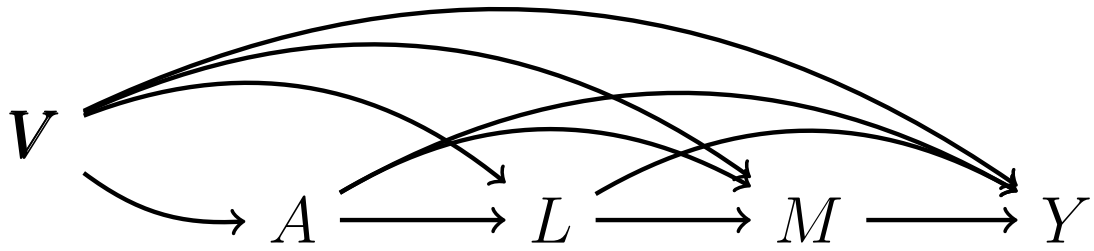
- "Post-treatment confounding" or "exposure-induced mediator-outcome confounding."
- Best description of this is Wodtke, Harding, & Elwert (2011, ASR).
- This is impossible to handle with Baron-Kenny mediation. You'll likely always overestimate the proportion of the total effect that is mediated.
- Requires "g-methods": IPTW (weighting) or g-formula (impute counterfactuals). See Naimi, Cole, & Kennedy (2016, IJE): An introduction to g-methods.

Causal mediation



1. The post-treatment variable L both *mediates* $A \longrightarrow Y$ and *confounds* $M \longrightarrow Y$, so we can't simply control for it in a single regression model.

G-FORMULA



$$E[Y] = \sum_m \sum_l \Bigg\{ P(Y|a,m,l,\boldsymbol{v}) \times P(m|a,l,\boldsymbol{v}) \times P(l|a,\boldsymbol{v}) \times P(\boldsymbol{v}) \Bigg\}$$

$$Y = \beta_0 + A \times (\beta_1 + \beta_2 M + \beta_3 L + \beta_4 \boldsymbol{V}) + \epsilon_Y$$

$$M = \beta_0 + A \times (\beta_1 + \beta_2 L + \beta_3 \boldsymbol{V}) + \epsilon_M$$

$$L = \beta_0 + A \times (\beta_1 + \beta_2 \boldsymbol{V}) + \epsilon_L$$

Mediation estimands

$$\text{NDE} = E[Y_{aM_{a^*}}] - E[Y_{a^*M_{a^*}}]$$

$$\text{NIE} = E[Y_{aM_a}] - E[Y_{aM_{a^*}}]$$

- New counterfactual quantities! These estimands contain the value the mediator would take "naturally" under treatment or no treatment.

Mediation estimands

$$\text{NDE} = E[Y_{aM_{a^*}}] - E[Y_{a^*M_{a^*}}]$$

$$\text{NIE} = E[Y_{aM_a}] - E[Y_{aM_{a^*}}]$$

- New counterfactual quantities! These estimands contain the value the mediator would take "naturally" under treatment or no treatment.
- This gets us closer to methods for "demographic decomposition", a connection I'll talk about at the end. For example, what would rates of college attainment look like for folks racialized as Black if they had the same distribution of parental income as those racialized as white? What would the crude mortality rate look like for Mexico if it had the same age structure as the US?

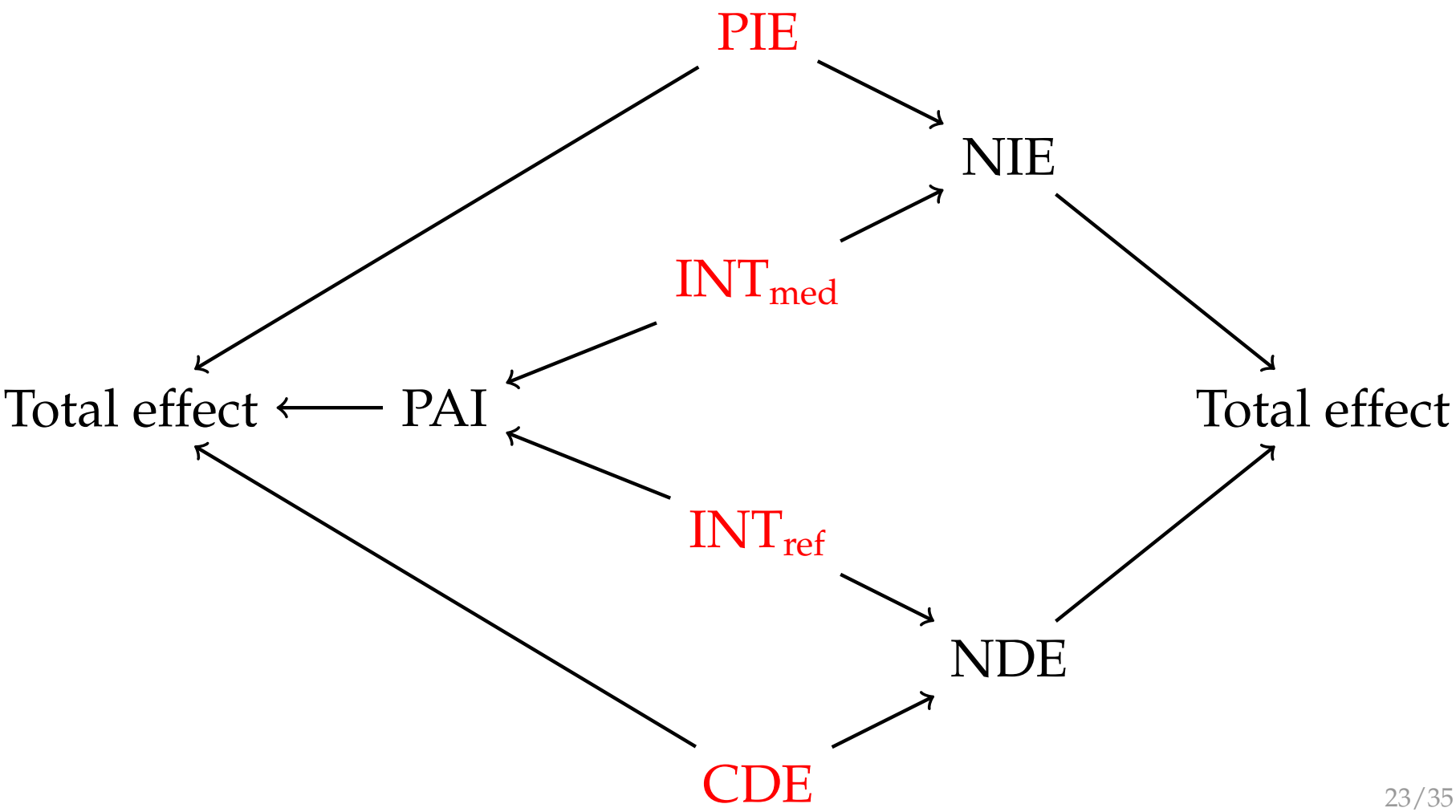
Interactions

- What about interactions? What if they effect of M on Y *depends* on treatment? Consider the Black-white income example I just gave.

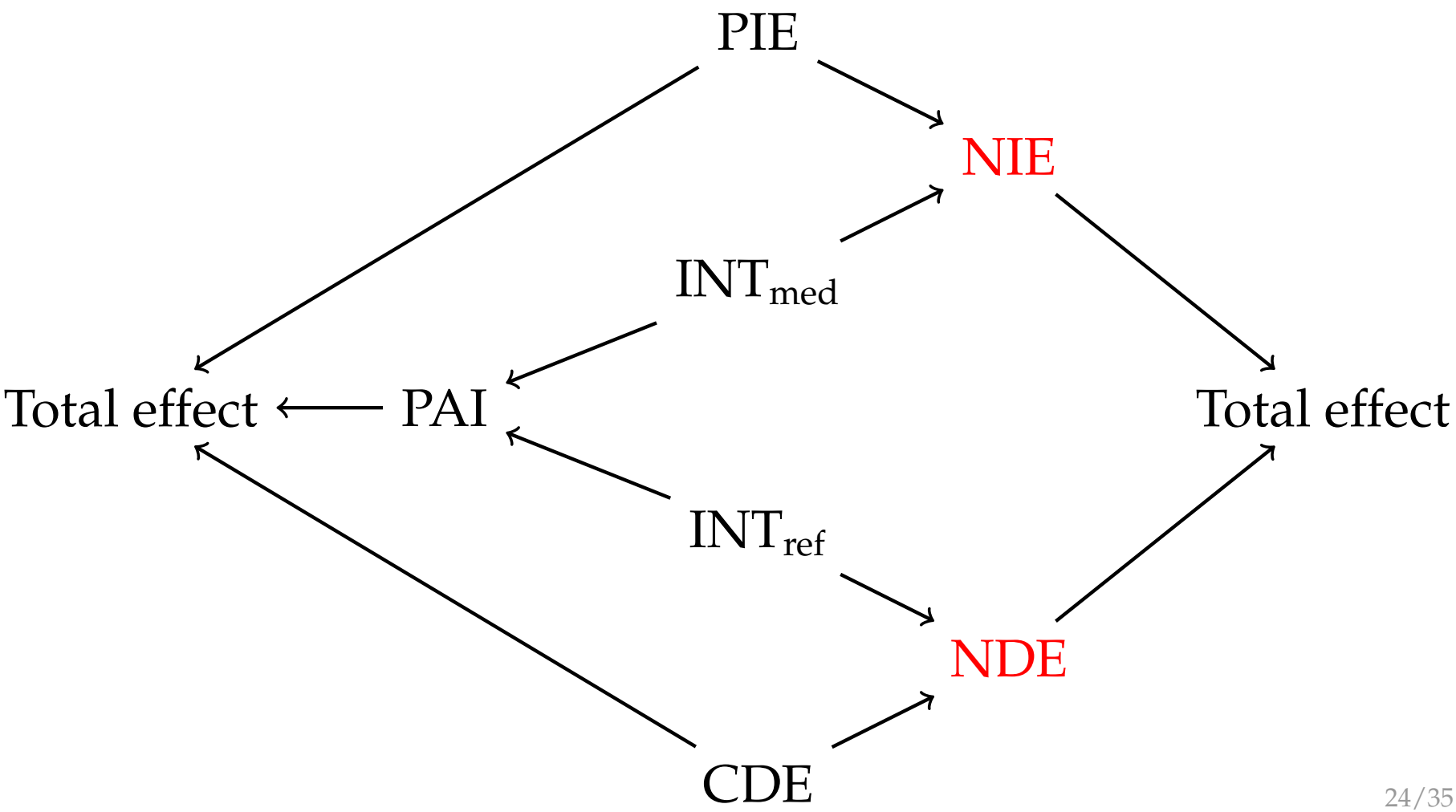
Interactions

- What about interactions? What if they effect of M on Y *depends* on treatment? Consider the Black-white income example I just gave.
- We can handle treatment-mediator interactions in VanderWeele's (2014) 4-way decomposition.

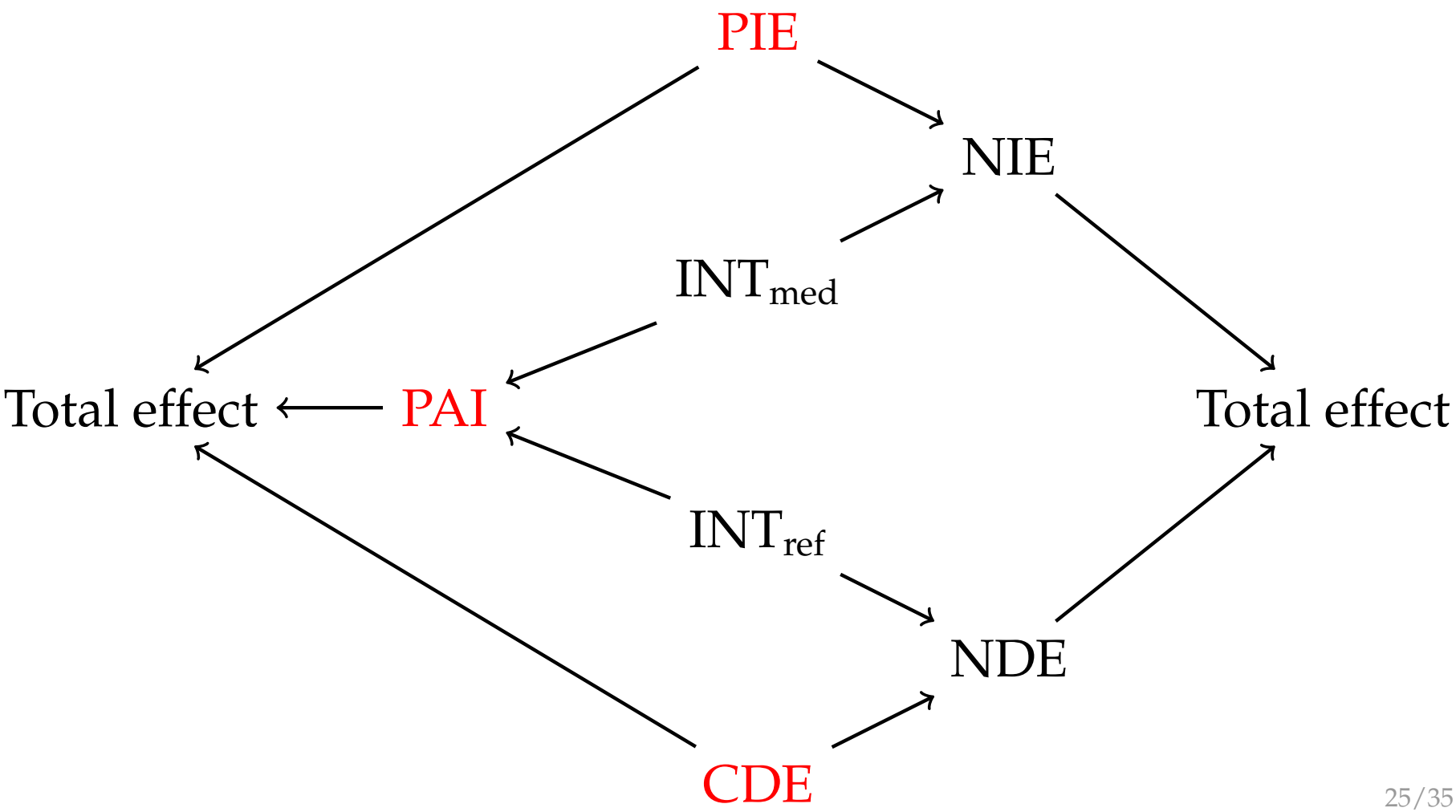
VanderWeele 2014: 4-way



VanderWeele 2014: 2-way



VanderWeele 2014: 3-way



Mediation estimands

$$\text{NDE} = E[Y_{xM_{a^*}}] - E[Y_{a^*M_{a^*}}]$$

$$\text{NIE} = E[Y_{aM_a}] - E[Y_{aM_{a^*}}]$$

$$\text{CDE}_{M=m^*} = E[Y_{am^*}] - E[Y_{a^*m^*}]$$

$$\text{PAI}^{(M)} = E[Y_{am} - Y_{a^*m} - Y_{am^*} + Y_{a^*m^*}](M_a)]$$

$$\text{PIE}^{(M)} = E[Y_{a^*M_a}] - E[Y_{a^*M_{a^*}}]$$

Estimands (Wang & Arah 2015)

Potential research questions related to various effects under effect decomposition

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*?

[a] TE: total effect, PDE: pure direct effect, TIE: total indirect effect, TDE: total direct effect, PIE: pure indirect effect, CDE: controlled direct effect (standard), CDE_{sto}: stochastic controlled direct effect, RIE: reference interaction effect (referred to as "INT_{ref}" by VanderWeele), MIE: mediated interaction effect (referred to as "INT_{med}" by VanderWeele), PAI: portion attributable to interaction.

Estimands (Wang & Arah 2015)

Definition and empirical analog of total effect and component effects[a]

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	$E[(Y_{xm} - Y_{xm^*} - Y_{x^*m} + Y_{x^*m^*})(M_x)]$	$\sum_z\sum_m\{E(Y\|x, m, z) - E(Y\|x, m^*, z) - E(Y\|x^*, m, z) + E(Y\|x^*, m^*, z)\}\,P(m\|x, z)\,P(z)$

[a] Y: outcome, X: exposure, M: mediator, Z: covariates; x and m represent the index values whereas x^* and m^* represent the reference values.

[b] Under the stable unit treatment value assumption, consistency, conditional exchangeability, positivity, different types of effect can be identified and estimated using the empirical analogs We use $E(Y|x, m, z)$ as a shorthand for $E(Y|X = x, M = m, \boldsymbol{Z} = z)$, and $P(m|x, z)$ as a shorthand for $P(M = m|X = x, \boldsymbol{Z} = z)$.

Implementation

- Baoyi Shi, Christine Choirat and Linda Valeri (2021). CMAverse: Causal Mediation Analysis.

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.

- Limitations
  - Multiple mediators can only be considered jointly
  - No cumulative effects via time-varying mediators
  - No survey weights

Paths package

- Zhou & Yamamoto (2023): Tracing Causal Paths from Experimental and Observational Data

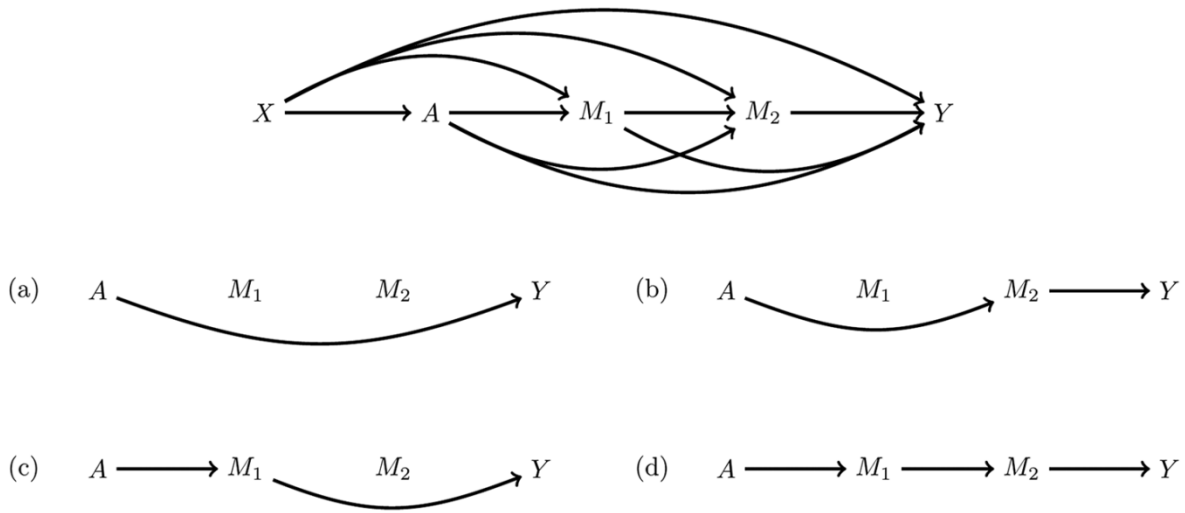


Figure 1. Causal relationships with two causally ordered mediators; A denotes the treatment, Y denotes the outcome of interest, X denotes a vector of pretreatment covariates, and M_1 and M_2 denote two causally ordered mediators. The confounding arcs between X and each of the other nodes are omitted in subgraphs a–d.

Estimation notes

- Randomized analogues

Estimation notes

- Randomized analogues
- Stochastic controlled direct effects

Estimation notes

- Randomized analogues
- Stochastic controlled direct effects
- The "cross world independence assumption": 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 (see Andrews & Didelez, 2021, Epidemiology).

Estimation notes

- But aren't we just imputing values to consider different counterfactual contrasts? Are we really doing anything *causal*? Is this just "surgery of equations?"

Estimation notes

- But aren't we just imputing values to consider different counterfactual contrasts? Are we really doing anything *causal*? Is this just "surgery of equations?"
- Is demographic decomposition ever causal? See Sudharsanan & Bijlsma (2022, IJE).

Estimation notes

- But aren't we just imputing values to consider different counterfactual contrasts? Are we really doing anything *causal*? Is this just "surgery of equations?"
- Is demographic decomposition ever causal? See Sudharsanan & Bijlsma (2022, IJE).
- *My take*: mechanisms are important, and we can use mediation analysis to rule out theoretical alternatives and use a holistic view of "inference to best explanation." Baron-Kenny mediation is bad, but "causal mediation" in the strict potential outcomes sense will probably never meet its own assumptions. Something in the middle, paying attention to complex post-treatment dynamics and combined with good theory, can be useful.

Advanced topics in mediation

- Loosening assumptions about model misspecification (e.g., machine learning).

Advanced topics in mediation

- Loosening assumptions about model misspecification (e.g., machine learning).
- Multiple mediators. See Zhou & Yamamoto (2023, The Journal of Politics), Graetz & Esposito (2022, Social Forces).

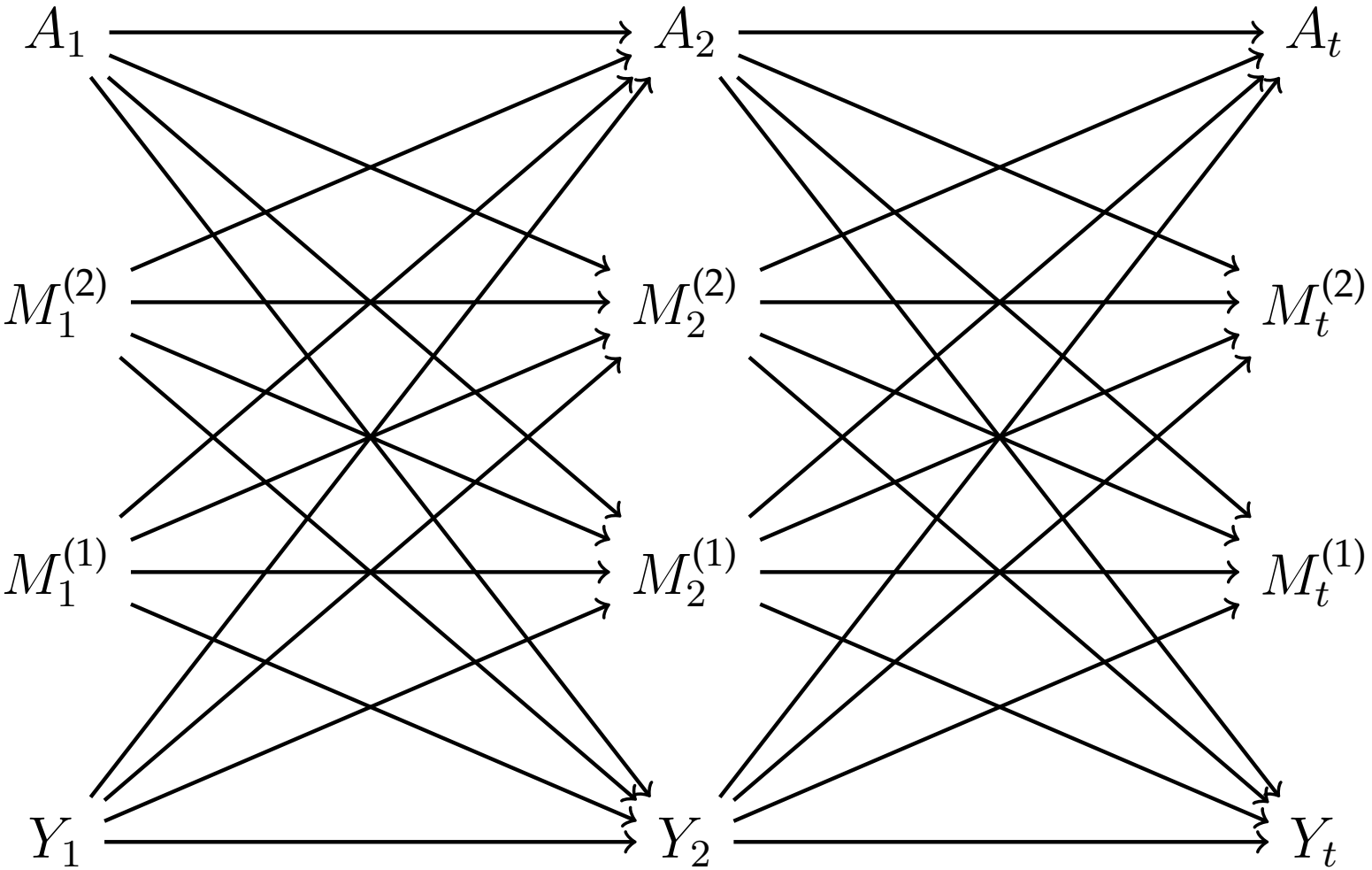
Advanced topics in mediation

- Loosening assumptions about model misspecification (e.g., machine learning).
- Multiple mediators. See Zhou & Yamamoto (2023, The Journal of Politics), Graetz & Esposito (2022, Social Forces).
- Time-varying, cumulative effects with single and multiple mediators. See Lin et al. (2018, Epidemiology).

Advanced topics in mediation

- Loosening assumptions about model misspecification (e.g., machine learning).
- Multiple mediators. See Zhou & Yamamoto (2023, The Journal of Politics), Graetz & Esposito (2022, Social Forces).
- Time-varying, cumulative effects with single and multiple mediators. See Lin et al. (2018, Epidemiology).
- Formal connections to demographic decomposition. See Sudharsanan & Bijlsma (2022, IJE).

Time-varying exposure, outcome, and mediators



Other theoretical considerations

- Drawing the causal field and the problem of *infinite regress*.

Other theoretical considerations

- Drawing the causal field and the problem of *infinite regress*.
- There are a lot of methods to investigate mechanisms, not just quantitative causal inference.