Causal Mediation in Natural Experiments

Senan Hogan-Hennessy
Economics Department, Cornell University
seh325@cornell.edu

Labour Work in Progress Seminar
6 March 2025

Plan

1. Start with explaining natural experiment, good for ATE $Z \to Y$. COnsider the Oregon health insurance experiment, or Vietnam draft instrument.
2. Does not illuminate how these causal effects came about.
3. You may have read epidemiology, medicine, or psychology and wondered what these claims are "mediated through."
4. These are mediation effect estimates, and they estimate "how much of the ATE goes through this channel? How much is left-over?"
5. Leads to my introduction page.

Introduction

Have you ever read an epidemiology/psychology/medicine paper's abstract, and seen claims of mediator effects **mediated** through some mechanism?

Family communication patterns, family environment, and the impact of parental alcoholism on offspring self-esteem [PDF] sagep

S Rangarajan, L Kelly - Journal of Social and Personal …, 2006 - journals.sagepub.com

This study examined the role of perceptions of family environment and family communication as mediators of the effects of parental alcoholism on the self-esteem of adult children of alcoholics. Participants (N= 227) completed self-reports of parental alcoholism, family environment, family communication patterns (FCP), and self-esteem. Results indicated a negative relationship between the seriousness of both maternal and paternal alcoholism and self-esteem. Paternal and maternal alcoholism were related to the two dimensions of …

Save Cite Cited by 122 Related articles All 3 versions

Introduction

Have you ever read an epidemiology/psychology/medicine paper's abstract, and seen claims of mediator effects **mediated** through some mechanism?

[HTML] Persistent depressive symptomatology and inflammation: to what extent do health behaviours and weight control mediate this relationship? [HTML] sciei

M Hamer, GJ Molloy, C de Oliveira... - Brain, Behavior, and ..., 2009 - Elsevier

We examined if persistent depressive symptoms are associated with markers of inflammation (C-Reactive Protein-CRP) and coagulation (fibrinogen), and if this association can be partly explained by weight control and behavioural risk factors (smoking, alcohol, physical activity). The study sample included 3609 men and women (aged 60.5±9.2 years) from The English Longitudinal Study of Ageing, a prospective study of community dwelling older adults. Depressive symptoms (using the 8-item CES-D scale), health behaviours ...

☆ Save ⁹⁹ Cite Cited by 111 Related articles All 6 versions

Introduction

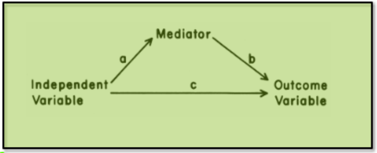
1980s: Psychometrics defined mediation (distinct from moderation).
1920s: Application of early econometric path analysis (Wright 1928).
2020s: Popular in epidemiology, medicine, psychology.

Figure: Baron Kelly (1986), p. 1176.

1176 REUBEN M. BARON AND DAVID A. KENNY

gression equasion, as described by Cohen and Cohen (1983) and Cleary and Kessler (1982). So if the independent variable is denoted as X, the moderator as Z, and the dependent variable as Y, Y is regressed on X, Z, and XZ. Moderator effects are indicated by the significant effect of XZ while X and Z are controlled. The simple effects of the independent variable for different levels of the moderator can be measured and tested by procedures described by Aiken and West (1986). (Measurement error in the moderator requires the same remedies as measurement error in the independent variable under Case 2.)

The quadratic moderation effect can be tested by dichotomizing the moderator at the point at which the function is presumed to accelerate. If the function is quadratic, as in Figure 2, the effect of the independent variable should be greatest for those who are high on the moderator. Alternatively, quadratic moderation can be tested by hierarchical regression procedures described by Cohen and Cohen (1983). Using the same notation as in the previous paragraph, Y is regressed on X, Z, XZ, Z^2, and XZ^2. The test of quadratic moderation is given by the test

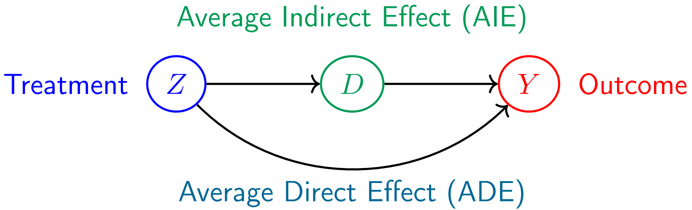


model, which recognizes that an active organism intervenes between stimulus and response, is perhaps the most generic formulation of a mediation hypothesis. The central idea in this model is that the effects of stimuli on behavior are mediated by various transformation processes internal to the organism. Theorists as diverse as Hull, Tolman, and Lewin shared a belief in the importance of postulating entities or processes that intervene between input and output. (Skinner's blackbox approach represents the notable exception.)

Introduction:

1. **[familiar]** Causal design to estimate a treatment effect.

2. **[unfamiliar]** CM decomposes ATE along a mechanism pathway.



3. **ATE** $\implies$ Average causal effect $Z \to Y$
   AIE $\implies$ How much $Z \to Y$ effect through mediator D?
   ADE $\implies$ How much $Z \to Y$ effect is left over?

Introduction— CM Examples:

1. Lottery military draft 1969 (Angrist 1990).

Drafted Service Earnings

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

Draft avoidance (education deferment)

Note: instrumental variables assumes ADE$= 0$ (exclusion restriction).

2. Oregon health insurance experiment (Finkelstein+ 2009).


The Oregon Health Insurance Experiment: What Did It Find and What Does that Mean?

Amy Finkelstein
November 2019

Medicaid Hospital use Health

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

All else (e.g., less worry)

Introduction

This project examines CM methods from an economic perspective:

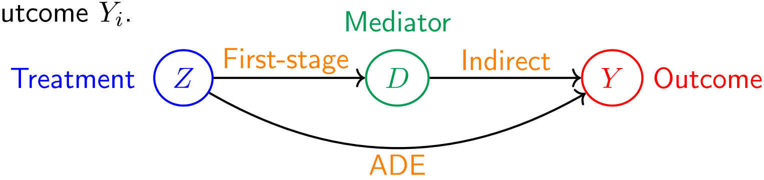
1. Problems with conventional, selection-on-observables, approach to CM in social science settings — including natural experiments.
   [Negative result]
2. Recovering valid CM effects under selection-into-mediator, using a selection model.
   [Positive result]

Brings together ideas from two different literatures:

- **Causal mediation.**
  Baron Kelly (1986), Imai Keele Yamamoto (2010), Flores Flores-Lagunes (2009), Frölich Huber (2017), Huber (2020), Kwon Roth (2024).
- **Selection-into-treatment, selection models/MTEs.**
  Roy (1951), Heckman (1979), Heckman Honore (1990), Florens Heckman Meghir Vytlacil (2008).

Direct & Indirect Effects — Model

Consider binary treatment $Z_i = 0, 1$, binary mediator $D_i = 0, 1$, and continuous outcome Y_i.



D_i is a function of Z_i :

$$D_i = Z_i D_i(1) + (1 - Z_i) D_i(0).$$

Y_i is a function of both Z_i, D_i :

$$Y_i = Z_i Y_i(1, D_i(1)) + (1 - Z_i) Y_i(0, D_i(0))$$

Assume Z_i is ignorable, conditional on $\boldsymbol{X}_i$.

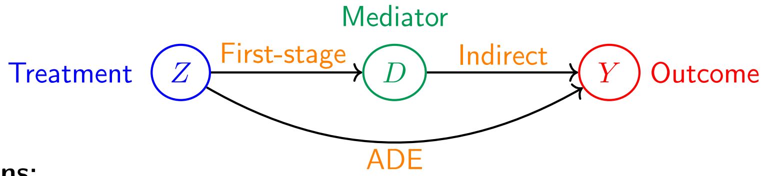
$$Z_i \perp\!\!\!\perp D_i(z), Y_i(z', d) \mid \boldsymbol{X}_i \text{ for } z, z', d = 0, 1.$$

Only two causal effects are identified so far.

ATE: $\mathbb{E}\left[Y_i(1, D_i(1)) - Y_i(0, D_i(0))\right] = \mathbb{E}\left[Y_i \mid Z_i = 1\right] - \mathbb{E}\left[Y_i \mid Z_i = 0\right]$

Average first-stage: $\mathbb{E}\left[D_i(1) - D_i(0)\right] = \mathbb{E}\left[D_i \mid Z_i = 1\right] - \mathbb{E}\left[D_i \mid Z_i = 0\right]$

Direct & Indirect Effects — Definitions



Definitions:

$$\text{Average Direct Effect (ADE)}: \quad \mathbb{E}\left[Y_i(1, D_i(Z_i)) - Y_i(0, D_i(Z_i))\right],$$
$$\text{Average Indirect Effect (AIE)}: \quad \mathbb{E}\left[Y_i(Z_i, D_i(1)) - Y_i(Z_i, D_i(0))\right].$$

- ADE is average effect $Z \to Y$, blocking the D path.
- AIE is causal effect of $D \to Y$, times number of $D(Z)$ compliers.[1]

$$\text{AIE} = \mathbb{E}\left[D_i(1) - D_i(0)\right] \mathbb{E}\left[Y_i(Z_i, 1) - Y_i(Z_i, 0) \,|\, D_i(1) = 1, D_i(0) = 0\right].$$

[1] Assume mediator monotonicity to simplify notation.

Direct & Indirect Effects — Identification

Sequential ignorability (**SI**, Imai Keele Yamamoto 2010):

Assume mediator D_i is *also* ignorable, conditional on $\boldsymbol{X}_i$ and Z_i realisation

$$D_i \perp\!\!\!\perp Y_i(z', d) \;\mid\; \boldsymbol{X}_i, Z_i = z', \text{ for } z', d = 0, 1.$$

If **SI** holds then ADE and AIE are identified by two-stage regression:

$$\text{ADE} = \mathbb{E}\left[\underbrace{\mathbb{E}\left[Y_i \,\middle|\, Z_i = 1, D_i = d', \boldsymbol{X}_i\right] - \mathbb{E}\left[Y_i \,\middle|\, Z_i = 0, D_i = d', \boldsymbol{X}_i\right]}_{\text{Second-stage regression, } Y_i \text{ on } Z_i \text{ holding } D_i, \boldsymbol{X}_i \text{ constant}} \right]$$

$$\text{AIE} = \mathbb{E}\left[\begin{array}{l} \underbrace{\Big(\mathbb{E}\left[D_i \,|\, Z_i = 1, \boldsymbol{X}_i\right] - \mathbb{E}\left[D_i \,|\, Z_i = 0, \boldsymbol{X}_i\right]\Big)}_{\text{First-stage regression, } D_i \text{ on } Z_i} \\ \times \underbrace{\Big(\mathbb{E}\left[Y_i \,\middle|\, Z_i = z', D_i = 1, \boldsymbol{X}_i\right] - \mathbb{E}\left[Y_i \,\middle|\, Z_i = z', D_i = 0, \boldsymbol{X}_i\right]\Big)}_{\text{Second-stage regression, } Y_i \text{ on } D_i \text{ holding } Z_i, \boldsymbol{X}_i \text{ constant}} \end{array} \right]$$

Direct & Indirect Effects — Identification

Sequential ignorability (**SI**, Imai Keele Yamamoto 2010):

Assume mediator D_i is *also* ignorable, conditional on $\boldsymbol{X}_i$ and Z_i realisation

$$D_i \perp\!\!\!\perp Y_i(z', d) \mid \boldsymbol{X}_i, Z_i = z', \text{ for } z', d = 0, 1.$$

E.g., OLS simultaneous regression (Imai Keele Yamamoto, 2010):

$Z_i \leftarrow$ Treatment First-stage: $D_i = \phi + \pi Z_i + \boldsymbol{\psi}_1' \boldsymbol{X}_i + \eta_i$

$D_i \leftarrow$ Mediator Second-stage: $Y_i = \alpha + \beta D_i + \gamma Z_i + \delta Z_i D_i + \boldsymbol{\psi}_2' \boldsymbol{X}_i + \varepsilon$

$Y_i \leftarrow$ Outcome $\implies$ ADE $= \gamma + \delta \mathbb{E}[D_i]$

AIE $= \pi(\beta + \delta \mathbb{E}[Z_i])$

i.e., a regression decomposition.

Other estimation methods do the same decomposition, avoiding linearity assumptions (see Huber 2020 for an overview).

Direct & Indirect Effects — Selection

$\implies$ Great, we can use the Imai Keele Yamamoto (2010) approach to CM all our respective applied projects.

$\implies$ Learn the mechanism pathways in causal research $\to$ big gain!

Before we join epidemiologists/psychologists/medical researchers in this conclusion, let us interrogate the **SI** assumption.

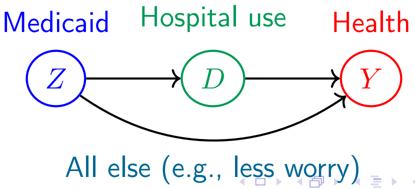
$$D_i \perp\!\!\!\perp Y_i(z', d) \mid \boldsymbol{X}_i, Z_i = z', \text{ for } z', d = 0, 1.$$

Would this assumption hold true in settings that social scientists consider?

Return to the Oregon health insurance experiment (Finkelstein+ 2009).


The Oregon Health Insurance Experiment: What Did It Find and What Does that Mean?

Amy Finkelstein
November 2019



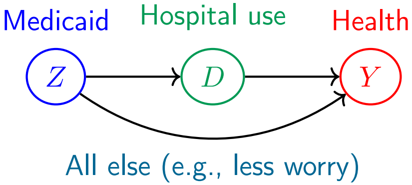
Direct & Indirect Effects — Selection

SI: $$D_i \perp\!\!\!\perp Y_i(z', d) \;\mid\; \boldsymbol{X}_i, Z_i = z', \text{ for } z', d = 0, 1.$$

Oregon health insurance experiment (Finkelstein+ 2009).


The Oregon Health Insurance Experiment: What Did It Find and What Does that Mean?

Amy Finkelstein
November 2019



SI in this setting:

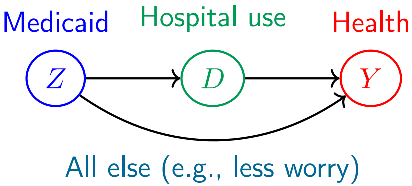
1. Health insurance assigned randomly (ensured by studying the 2008 Oregon waitlist lottery).
2. Hospital usage is quasi-random, conditional on Medicaid assignment Z_i and demographics $\boldsymbol{X}_i$.

Direct & Indirect Effects — Selection

SI: Hospital usage is quasi-random, conditional on Medicaid assignment Z_i and demographics $\boldsymbol{X}_i$.


The Oregon Health Insurance Experiment: What Did It Find and What Does that Mean?

Amy Finkelstein
November 2019



Consider the case **individuals go to the hospital** to maximise health.

$$D_i\left(z'\right) = \mathbb{1}\left\{ \underbrace{Y_i\left(z', 1\right) - Y_i\left(z', 0\right)}_{\text{Benefits}} \geq \underbrace{C_i}_{\text{Costs}} \right\}, \quad \text{for } z' = 0, 1.$$

i.e., Roy (1951) selection into D_i.

Direct & Indirect Effects — Selection

SI: Hospital usage is quasi-random, conditional on Medicaid assignment Z_i and demographics $\boldsymbol{X}_i$.

Consider the case **individuals go to the hospital** to maximise health.

$$D_i\left(z'\right) = \mathbb{1}\left\{ \underbrace{Y_i\left(z', 1\right) - Y_i\left(z', 0\right)}_{\text{Benefits}} \geq \underbrace{C_i}_{\text{Costs}} \right\}, \quad \text{for } z' = 0, 1.$$

i.e., Roy (1951) selection into D_i.

Theorem: If selection is Roy-style, and benefits are not 100% explained by $Z_i, \boldsymbol{X}_i$, then **SI** does not hold.

Proof sketch: suppose D_i is ignorable $\implies$ selection-into-D_i is explained 100% by $\{C_i, Z_i, \boldsymbol{X}_i\}$, while unobserved gains explain 0%.

Direct & Indirect Effects — Selection

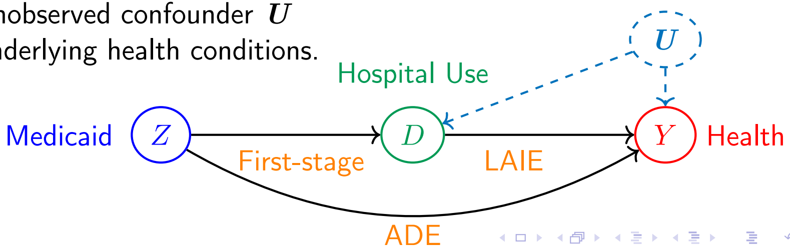
SI: Hospital usage is quasi-random, conditional on Medicaid assignment Z_i and demographics $\boldsymbol{X}_i$.

Consider the case **individuals go to the hospital** to maximise health.

$$D_i(z') = \mathbb{1}\left\{ \underbrace{Y_i(z', 1) - Y_i(z', 0)}_{\text{Benefits}} \geq \underbrace{C_i}_{\text{Costs}} \right\}, \quad \text{for } z' = 0, 1.$$

i.e., Roy (1951) selection into D_i.

$\Longrightarrow$ unobserved confounder $\boldsymbol{U}$
e.g., underlying health conditions.



Direct & Indirect Effects — Selection

In practice, the only way to believe the **SI** assumption (selection-on-observables) is to study a case with another natural experiment for D_i — in addition to the one that guaranteed Z_i is ignorable.

(a) Cells in a lab → **SI** believable.

(b) People choosing healthcare → **SI** not.

Direct & Indirect Effects — Selection Bias

- What happens if you go ahead and estimate CM anyway?
- Would this be problematic?