Causal Mediation in Natural Experiments

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

Cornell, Labor Economics Seminar
17 November 2025

Introduction

Natural experiments are settings with credible estimates of causal effects

- Little information on the mechanisms through which they operate
- Limits understanding of the decisions and underlying economic system
- Causal Mediation (CM) is a framework for sufficiently analysing a causal effect along an observed mechanism, which is not widely used in applied economics.

This paper:

1. Develop selection bias concept for CM when we do not believe its assumptions, which can be large in practice
2. Build an MTE-based approach to tackle the identification problem
3. Illustrate my methods with decomposing causal effects in the Oregon Health Insurance Experiment.

Introduction — Contributions

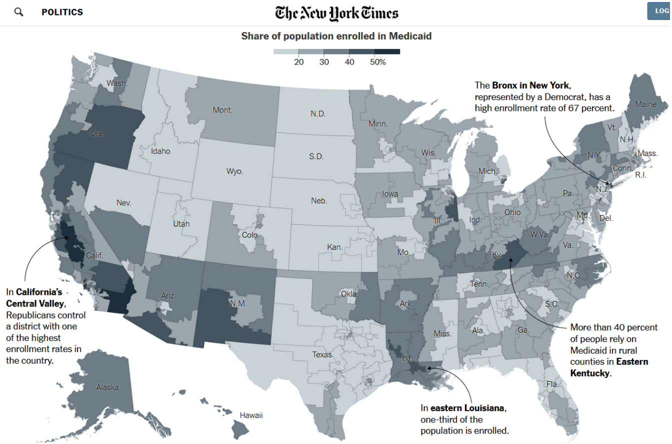
1. Problems with conventional approach to CM in observational settings
   → makes explicit the folk-style reasoning for economics not engaging in CM.
   [Negative result]
2. Recovering CM effects, via Marginal Treatment Effect (MTE) model
   → Causal mediation from a quasi-experimental economist approach.
   [Positive result]

New insights from intersection of two fields:

- **Causal Mediation (CM).**
  Imai Keele Yamamoto (2010), Frölich Huber (2017), Deuchert Huber Schelker (2019), Huber (2020), Kwon Roth (2024).
- **Labour theory, Selection-into-treatment, MTEs.**
  Roy (1951), Heckman (1979), Heckman Honoré (1990), Vycatil (2002), Heckman Vycatil (2005), Brinch Mogstad Wiswall (2017), Kline Walters (2019).

Oregon Health Insurance Experiment

In the USA, healthcare is only provided by the government in special cases $\rightarrow$ Medicaid is the government programme which provides health insurance for those close to the poverty line (> 70 million people in 2025).



Oregon Health Insurance Experiment

In the USA, healthcare is only provided by the government in special cases $\rightarrow$ Medicaid is the government programme which provides health insurance for those close to the poverty line (> 70 million people in 2025).

2004: Oregon Medicaid enrolment frozen

↓

2008: Enrolment reopens...
90,000 sign up (far exceeding funding allocated)

↓

Wait-list Lottery:

35,000 Won
(Treatment)

55,000 Lost
(Control)

Oregon Health Insurance Experiment

In the USA, healthcare is only provided by the government in special cases $\rightarrow$ Medicaid is the government programme which provides health insurance for those close to the poverty line (> 70 million people in 2025).

2004: Oregon Medicaid enrolment frozen

↓

2008: Enrolment reopens...
90,000 sign up (far exceeding funding allocated)

↓

Wait-list Lottery:

35,000 Won (Treatment) | 55,000 Lost (Control)

Oregon Health Insurance Experiment

In the USA, healthcare is only provided by the government in special cases
$\rightarrow$ Medicaid is the government programme which provides health insurance for those close to the poverty line (> 70 million people in 2025).

2004: Oregon Medicaid enrolment frozen

↓

2008: Enrolment reopens...
90,000 sign up (far exceeding funding allocated)

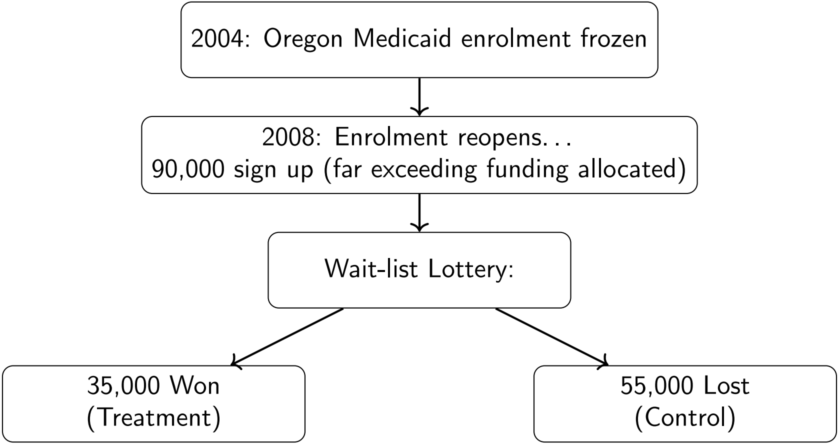
↓

Wait-list Lottery:

35,000 Won (Treatment) | 55,000 Lost (Control)

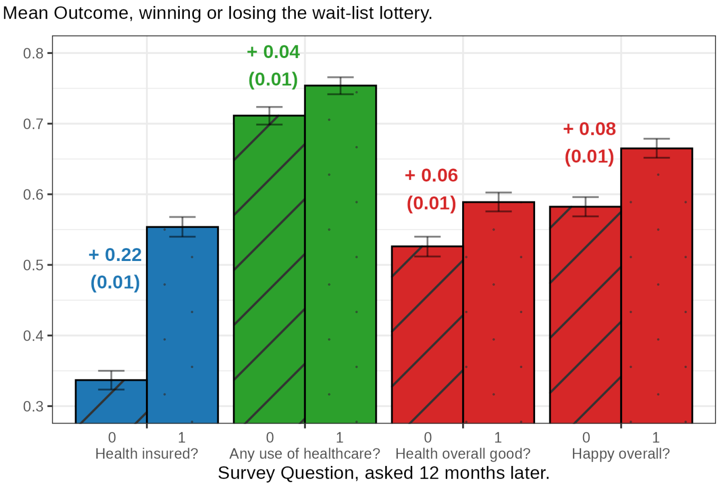
Oregon Health Insurance Experiment

In the USA, healthcare is only provided by the government in special cases $\rightarrow$ Medicaid is the government programme which provides health insurance for those close to the poverty line (> 70 million people in 2025).



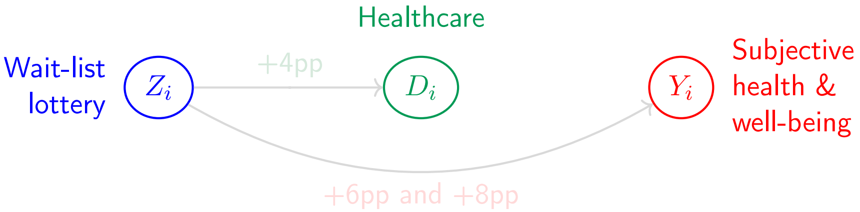
Oregon Health Insurance Experiment

Winning this wait-list lottery significantly increased healthcare usage, plus subjective health and well-being (Finkelstein et al, 2012).



Oregon — Suggestive Evidence

Winning this wait-list lottery significantly increased healthcare usage, plus subjective health and well-being (Finkelstein et al, 2012).

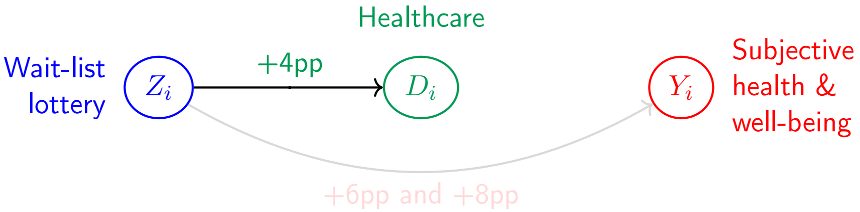


Suggestive evidence:

- If first-stage $\neq 0$, then healthcare may be a mediating mechanism
- This gives **suggestive evidence for healthcare as mechanism.**

Oregon — Suggestive Evidence

Winning this wait-list lottery significantly increased healthcare usage, plus subjective health and well-being (Finkelstein et al, 2012).

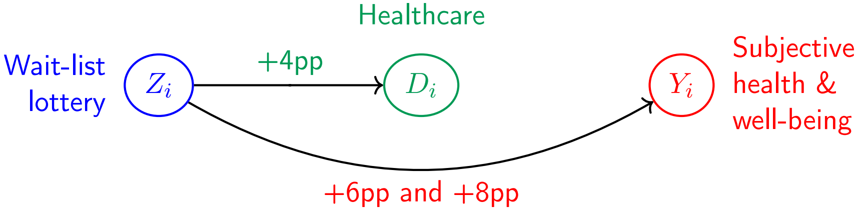


Suggestive evidence:

- If first-stage $\neq 0$, then healthcare may be a mediating mechanism
- This gives **suggestive evidence for healthcare as mechanism.**

Oregon — Suggestive Evidence

Winning this wait-list lottery significantly increased healthcare usage, plus subjective health and well-being (Finkelstein et al, 2012).

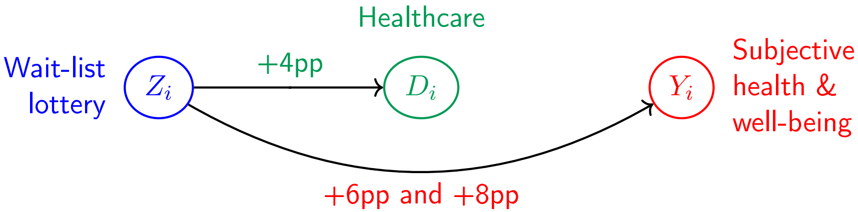


Suggestive evidence:

- If first-stage $\neq 0$, then healthcare may be a mediating mechanism
- This gives **suggestive evidence for healthcare as mechanism**.

Oregon — Suggestive Evidence

Winning this wait-list lottery significantly increased healthcare usage, plus subjective health and well-being (Finkelstein et al, 2012).



Suggestive evidence:

- If first-stage $\neq 0$, then healthcare may be a mediating mechanism
- This gives **suggestive evidence for healthcare as mechanism**.

Oregon — Suggestive Evidence

Suggestive evidence is primarily how economics investigates mechanisms.

Abstract — Lundborg Rooth Alex-Petersen (2022, ReStud).

"...Exposure to the [free school meals] programme also had substantial effects on educational attainment and health, which can explain a large part of the effect of the programme on lifetime income."

Abstract — Bloom Mahajan McKenzie Roberts (2013, QJE).

"...We find that adopting these management practices had three main effects. First, it raised average productivity by 11% through improved quality and efficiency and reduced inventory [...]."

Abstract — Carvhalo (2025, JPE Micro).

"... Evidence suggests fluid intelligence and self-control partly mediate the relationship between the [education polygenic index] and education."

Oregon — Suggestive Evidence

Suggestive evidence is primarily how economics investigates mechanisms.

Abstract — Lundborg Rooth Alex-Petersen (2022, ReStud).

"…Exposure to the [free school meals] programme also had substantial effects on educational attainment and health, which can explain a large part of the effect of the programme on lifetime income."

Abstract — Bloom Mahajan McKenzie Roberts (2013, QJE).

"…We find that adopting these management practices had three main effects. First, it raised average productivity by 11% through improved quality and efficiency and reduced inventory […]."

Abstract — Carvhalo (2025, JPE Micro).

"… Evidence suggests fluid intelligence and self-control partly mediate the relationship between the [education polygenic index] and education."

Oregon — Suggestive Evidence

Suggestive evidence is primarily how economics investigates mechanisms.

Abstract — Lundborg Rooth Alex-Petersen (2022, ReStud).

"...Exposure to the [free school meals] programme also had substantial effects on educational attainment and health, which can explain a large part of the effect of the programme on lifetime income."

Abstract — Bloom Mahajan McKenzie Roberts (2013, QJE).

"...We find that adopting these management practices had three main effects. First, it raised average productivity by 11% through improved quality and efficiency and reduced inventory [...]."

Abstract — Carvhalo (2025, JPE Micro).

"... Evidence suggests fluid intelligence and self-control partly mediate the relationship between the [education polygenic index] and education."

Oregon — Suggestive Evidence

Suggestive evidence is primarily how economics investigates mechanisms.

Abstract — Lundborg Rooth Alex-Petersen (2022, ReStud).

"…Exposure to the [free school meals] programme also had substantial effects on educational attainment and health, which can explain a large part of the effect of the programme on lifetime income."

Abstract — Bloom Mahajan McKenzie Roberts (2013, QJE).

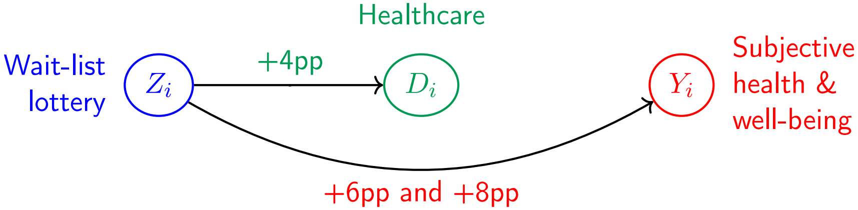
"…We find that adopting these management practices had three main effects. First, it raised average productivity by 11% through improved quality and efficiency and reduced inventory […]."

Abstract — Carvhalo (2025, JPE Micro).

"… Evidence suggests fluid intelligence and self-control partly mediate the relationship between the [education polygenic index] and education."

Oregon — Suggestive Evidence

Winning this wait-list lottery significantly increased healthcare usage, plus subjective health and well-being (Finkelstein et al, 2012).

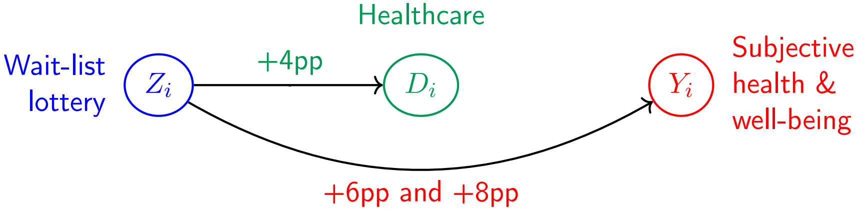


What about direct effects?

- Winning access to Medicaid means you can file for free health insurance (income effect)
- Less stress from no longer having to be uninsured (psychological gains).

Oregon — Suggestive Evidence

Winning this wait-list lottery significantly increased healthcare usage, plus subjective health and well-being (Finkelstein et al, 2012).

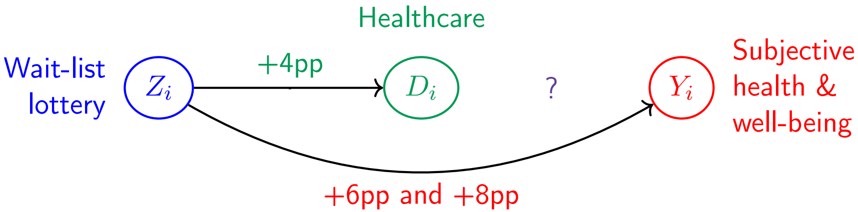


What about direct effects?

- Winning access to Medicaid means you can file for free health insurance (income effect)
- Less stress from no longer having to be uninsured (psychological gains).

Oregon — Suggestive Evidence

Winning this wait-list lottery significantly increased healthcare usage, plus subjective health and well-being (Finkelstein et al, 2012).



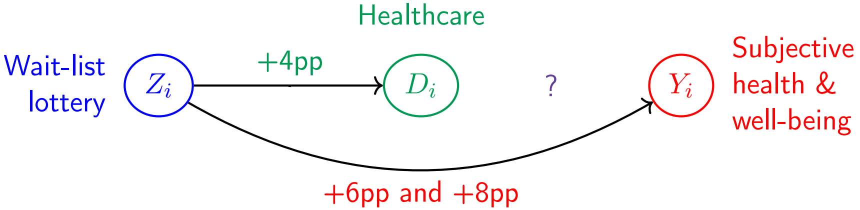
There is one missing piece to make a **definitive conclusion**:

Size of causal effect $D_i \to Y_i \ldots$

- If large, then healthcare explains all the lottery effect
- If small/zero then, then all direct (e.g., psychological) gains.

Oregon — Suggestive Evidence

Winning this wait-list lottery significantly increased healthcare usage, plus subjective health and well-being (Finkelstein et al, 2012).



There is one missing piece to make a **definitive conclusion**:

Size of causal effect $D_i \rightarrow Y_i \ldots$

- If large, then healthcare explains all the lottery effect
- If small/zero then, then all direct (e.g., psychological) gains.

Causal Mediation (CM)

CM is an alternative framework to studying mechanisms, giving sufficient evidence on the mediating mechanism.

Define

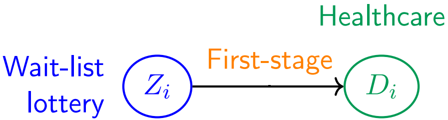
- Treatment $Z_i = 0, 1$, wait-list lottery
- Mediator mechanism $D_i = 0, 1$, healthcare usage
- Outcome Y_i, subjective health and well-being.

CM aims to decompose the ATE in two channels, direct and indirect effects

$$\text{ATE} = \text{ADE} + \text{AIE}.$$

Causal Mediation (CM)

CM is an alternative framework to studying mechanisms, giving sufficient evidence on the mediating mechanism.



Write $D_i(z')$ and $Y_i(z', d')$ for the potential outcomes.

Two average causal effects are identified, with Z_i randomly assigned:

1. Average first-stage

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

2. Average Treatment Effect (ATE)

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

Causal Mediation (CM)

CM is an alternative framework to studying mechanisms, giving sufficient evidence on the mediating mechanism.

Write $D_i(z')$ and $Y_i(z', d')$ for the potential outcomes.

Two average causal effects are identified, with Z_i randomly assigned:

1. Average first-stage

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

2. Average Treatment Effect (ATE)

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

Causal Mediation (CM)

CM is an alternative framework to studying mechanisms, giving sufficient evidence on the mediating mechanism.

Wait-list lottery $Z_i \xrightarrow{\text{Treatment Effect (ATE)}} Y_i$ Subjective health & well-being

Write $D_i(z')$ and $Y_i(z', d')$ for the potential outcomes.

Two average causal effects are identified, with Z_i randomly assigned:

1. Average first-stage

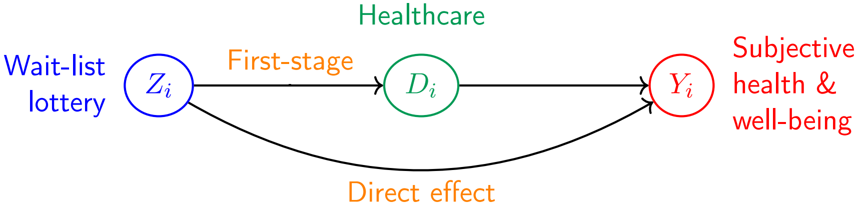
$$\mathbb{E}\left[D_i(1) - D_i(0)\right] = \mathbb{E}\left[D_i \mid Z_i = 1\right] - \mathbb{E}\left[D_i \mid Z_i = 0\right]$$

2. Average Treatment Effect (ATE)

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

Causal Mediation (CM)

CM is an alternative framework to studying mechanisms, giving sufficient evidence on the mediating mechanism.



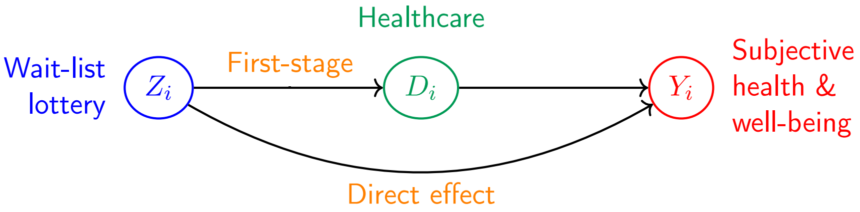
CM decomposes the ATE into components

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

AIE represents the average effect going through healthcare.

Causal Mediation (CM)

CM is an alternative framework to studying mechanisms, giving sufficient evidence on the mediating mechanism.



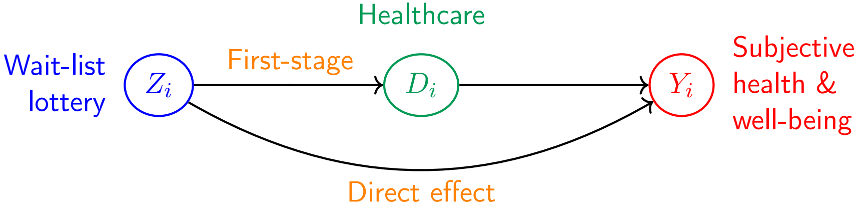
CM decomposes the ATE into components

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

ADE represents the average effect going absent healthcare.

Causal Mediation (CM)

ADE + AIE are not separately identified without further assumptions.

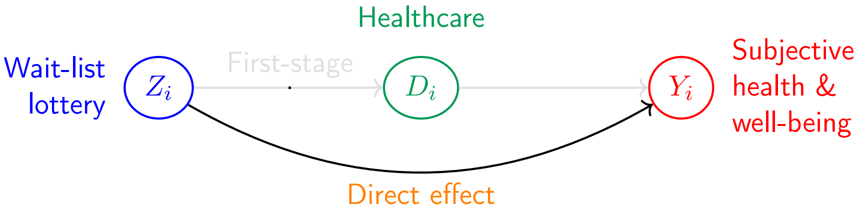


Conventional CM relies on two identifying assumptions,

1. Treatment Z_i is (quasi-)randomly assigned
2. Mediator D_i is (quasi-)randomly assigned, conditional on Z_i realisation (and covariates $\boldsymbol{X}_i$).

Causal Mediation (CM)

Under assumptions (1) + (2), the ADE + AIE are separately identified by two-stage regression (Imai Keele Yamamoto 2010).

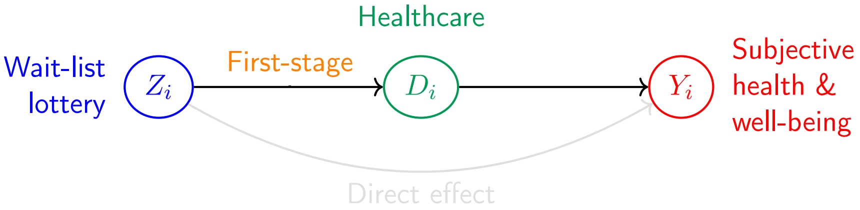


ADE is the effect of Z_i after controlling for D_i

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

Causal Mediation (CM)

Under assumptions (1) + (2), the ADE + AIE are separately identified by two-stage regression (Imai Keele Yamamoto 2010).



AIE is the effect of D_i after controlling for Z_i, times average first-stage.

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

Causal Mediation (CM)

This approach (conventional CM) is used heavily in epidemiology and medicine to give evidence for the channels of a treatment effect, but there is a reason why this is not prominent in economics.

Identifying assumptions:

1. Treatment Z_i is (quasi-)randomly assigned
2. Mediator D_i is (quasi-)randomly assigned, conditional on Z_i realisation (and covariates $\boldsymbol{X}_i$).

Translation: Healthcare is a random choice, conditional on wait-list lottery realisation and demographic controls.

Would this be plausible in settings economists study?

Causal Mediation (CM) — Roy Model

Consider the case that people, after the lottery, choose to visit the doctor in the next 12 months based on subjective costs and benefits,

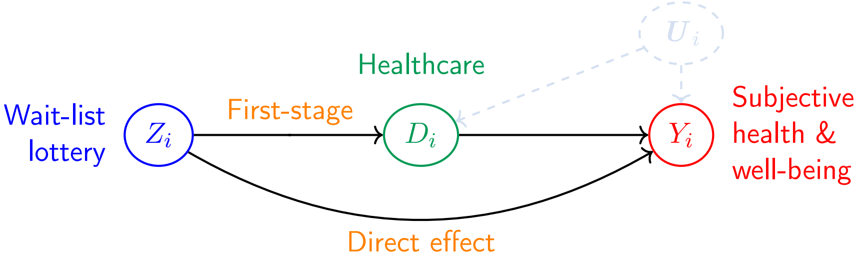
$$D_i\left(z'\right) = \mathbb{1}\left\{ \underbrace{C_i}_{\text{Costs}} \leq \underbrace{Y_i\left(z', \boxed{1}\right) - Y_i\left(z', \boxed{0}\right)}_{\text{Benefits}} \right\}.$$

The wait-list lottery has no strategic selection, but visiting healthcare after is an unconstrained choice.

Theorem: If choice to attend healthcare is unconstrained, based on costs and benefits (Roy model) and demographics do not explain all benefits $\implies$ mediator mechanism is not random, there is unobserved confounding.

Causal Mediation (CM) — Selection Bias

Individual unobserved benefits are an unobserved confounder $\boldsymbol{U}_i$ here,



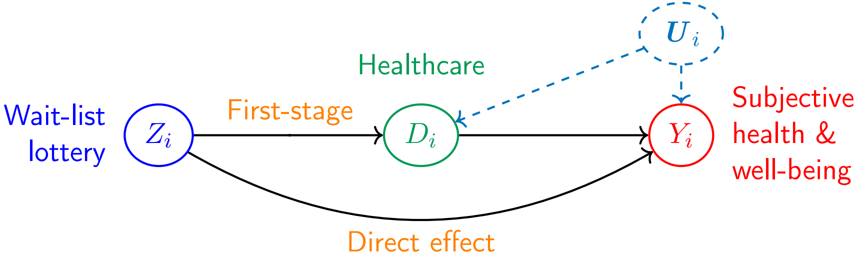
In economic settings, Conventional CM analyses have bias similar to classical selection bias (Heckman Ichimura Smith Todd 1998).

- Direct: $\text{CM Estimand} = \text{ADE} + \Big(\text{Selection Bias} + \text{Group difference bias}\Big)$
- Indirect: $\text{CM Estimand} = \text{AIE} + \Big(\text{Selection Bias} + \text{Group difference bias}\Big)$

ADE biases AIE biases

Causal Mediation (CM) — Selection Bias

Individual unobserved benefits are an unobserved confounder $\boldsymbol{U}_i$ here,



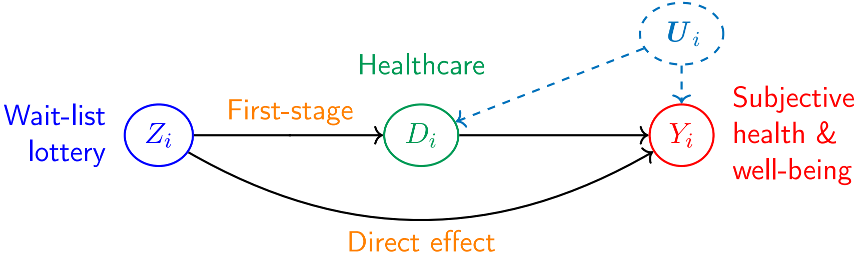
In economic settings, Conventional CM analyses have bias similar to classical selection bias (Heckman Ichimura Smith Todd 1998).

- Direct: CM Estimand = ADE + (Selection Bias + Group difference bias)
- Indirect: CM Estimand = AIE + (Selection Bias + Group difference bias)

ADE biases AIE biases

Causal Mediation (CM) — Selection Bias

Individual unobserved benefits are an unobserved confounder $\boldsymbol{U}_i$ here,



In economic settings, Conventional CM analyses have bias similar to classical selection bias (Heckman Ichimura Smith Todd 1998).

- Direct: CM Estimand $=$ ADE $+\Big($Selection Bias $+$ Group difference bias$\Big)$
- Indirect: CM Estimand $=$ AIE $+\Big($Selection Bias $+$ Group difference bias$\Big)$

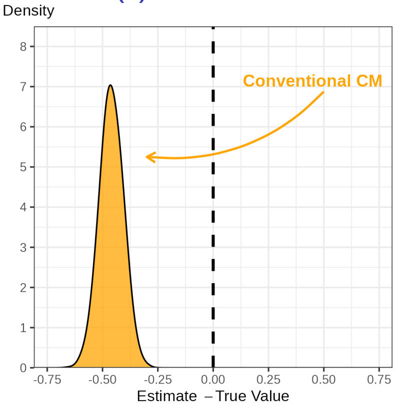
ADE biases AIE biases

Causal Mediation (CM) — Selection Bias

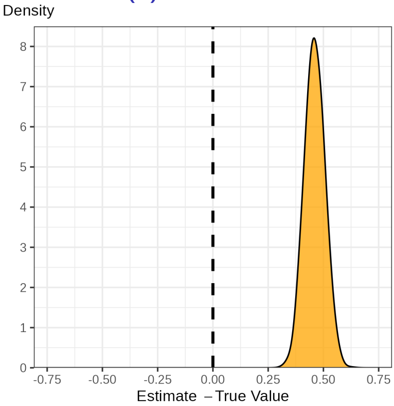
With strategic selection, the bias terms can be large and mislead inference on how much goes through the mediating channel.

Figure: Simulated Distribution of CM Effect Estimates from 10,000 DGPs.

(a) $\widehat{\text{ADE}} - \text{ADE}$.



(b) $\widehat{\text{AIE}} - \text{AIE}$.



CM with Selection

Conventional CM does not identify ADE + AIE in economic settings, so I build a structural model for natural experiment settings.

Take as given that Z_i is quasi-randomly assigned, but D_i is not:

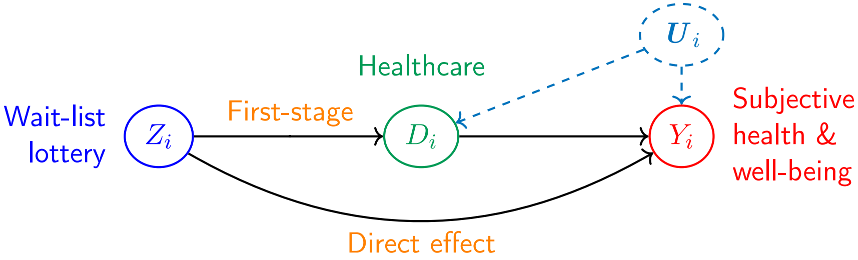
1. Average first-stage, $Z_i \to D_i$, is identified
2. Average second-stage, $Z_i, D_i \to Y_i$, is not — represented by $\boldsymbol{U}_i$.

Intuition: model $\boldsymbol{U}_i$ via mediator MTE to identify ADE + AIE.

CM with Selection

Conventional CM does not identify ADE + AIE in economic settings, so I build a structural model for natural experiment settings.

Take as given that Z_i is quasi-randomly assigned, but D_i is not:



1. Average first-stage, $Z_i \to D_i$, is identified
2. Average second-stage, $Z_i, D_i \to Y_i$, is not — represented by U_i.

Intuition: model U_i via mediator MTE to identify ADE + AIE.

CM with Selection

Conventional CM methods do not identify ADE + AIE in settings, so I build a structural model for natural experiment settings.

Write potential outcomes as mean + unobserved, in choosing to visit healthcare or not, $D_i = 0, 1$:

$$Y_i(z', 0) = \mathbb{E}\left[Y_i(z', 0) \,\middle|\, \boldsymbol{X}_i\right] + U_{0,i}, \quad Y_i(z', 1) = \mathbb{E}\left[Y_i(z', 1) \,\middle|\, \boldsymbol{X}_i\right] + U_{1,i}.$$

CM has two-stage regression equations:

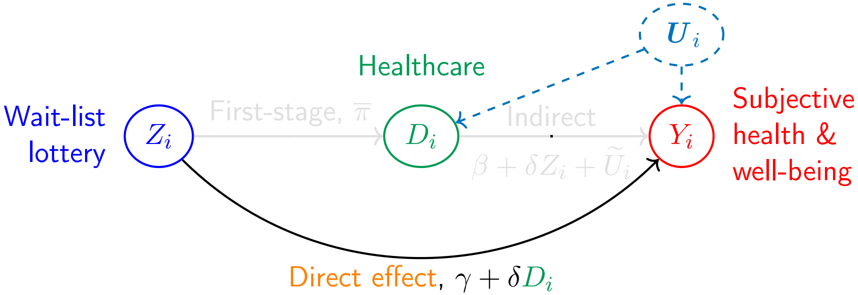
$$D_i = \phi + \overline{\pi} Z_i + \varphi(\boldsymbol{X}_i) + V_i$$

$$Y_i = \alpha + \beta D_i + \gamma Z_i + \delta Z_i D_i + \zeta(\boldsymbol{X}_i) + \underbrace{(1 - D_i) U_{0,i} + D_i U_{1,i}}_{\text{Correlated error term}}$$

1. $\overline{\pi}$ is average first-stage, effect $Z_i \to D_i$
2. β, γ, δ are separated effects of Z_i, D_i.

CM with Selection

Conventional CM methods do not identify ADE + AIE in settings, so I build a structural model for natural experiment settings.

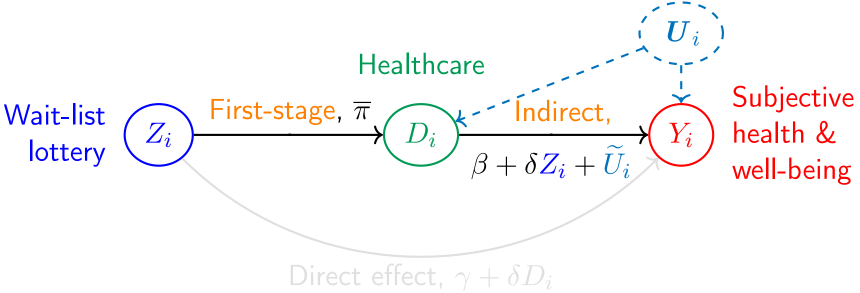


ADE composes effects of Z_i, holding D_i constant:

$$\text{ADE} = \mathbb{E}\left[\gamma + \delta D_i\right].$$

CM with Selection

Conventional CM methods do not identify ADE + AIE in settings, so I build a structural model for natural experiment settings.



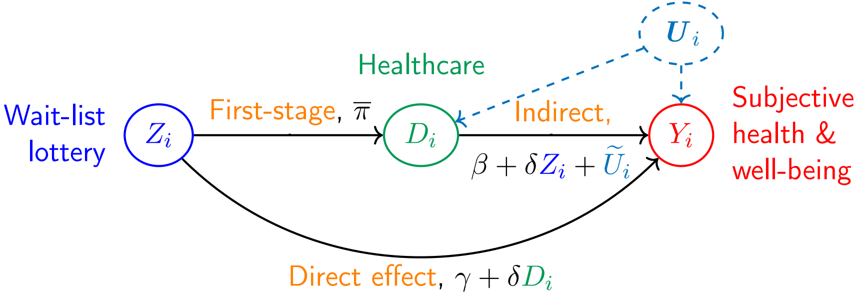
AIE composes effects of D_i, holding Z_i constant, times average first-stage:

$$\text{AIE} = \mathbb{E}\left[\overline{\pi}\left(\beta + \delta Z_i + \widetilde{U}_i\right)\right],$$

where $\widetilde{U}_i = \mathbb{E}\left[U_{1,i} - U_{0,i} \mid \boldsymbol{X}_i, D_i(0) \neq D_i(1)\right]$ unobserved complier gains.

CM with Selection

Conventional CM methods do not identify ADE + AIE in settings, so I build a structural model for natural experiment settings.

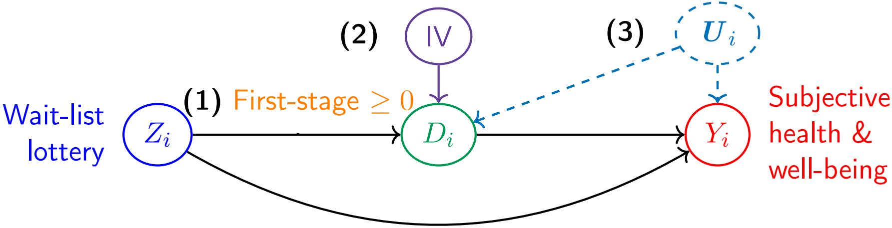


Structural model must solve the following issues:

1. β, γ, δ are not identified (see: selection bias)
2. $\widetilde{U}_i$ is also not known (unobserved complier healthcare gains).

MTE Model

The structural model is based on 3 assumptions.



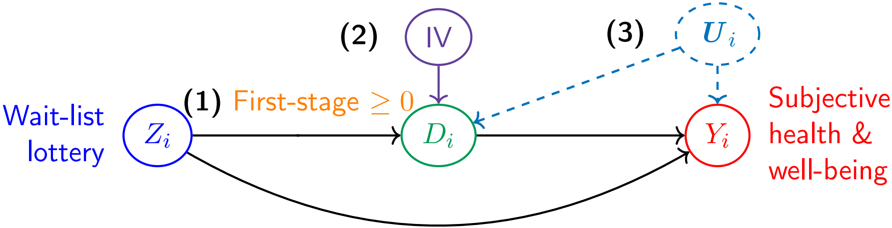
(1) First-stage monotonicity,

$$\Pr\left(D_i(0) \leq D_i(1)\right) = 1.$$

Intuition: No defiers — no one visits healthcare less if winning wait-list lottery, relative to losing.

MTE Model

The structural model is based on 3 assumptions.



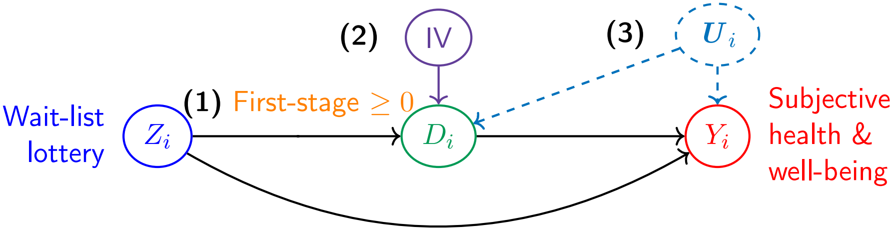
(2) Mediator take-up cost IV

Requires an IV, which affects Y_i only via D_i.

Key example: Cost-shifting IV — random variation in healthcare take-up (not gains), e.g. different healthcare costs.

MTE Model

The structural model is based on 3 assumptions.



(3) Selection on benefits — unobserved selection is relevant

$$\text{Cov}\left(V_i, U_{0,i}\right), \ \text{Cov}\left(V_i, U_{1,i}\right) \neq 0.$$

Key example: Roy model, people choose to take healthcare if internal subjective gains exceed costs.

MTE Model — Identification

Proposition: Under assumptions **(1)**, **(2)**, **(3)** mediator MTE is identified

$$\text{MTE} = \mathbb{E}\left[Y_i\left(z', \boxed{1}\right) - Y_i\left(z', \boxed{0}\right) \,\middle|\, Z_i = z', \boldsymbol{X}_i, V_i = p' \right]$$
$$= \beta + \delta z' + \underbrace{\mathbb{E}\left[U_{1,i} - U_{0,i} \,\middle|\, \boldsymbol{X}_i, V_i = p' \right]}_{=\rho_1 \lambda_1(p') - \rho_0 \lambda_0(p')}, \qquad \text{for } p' \in (0,1).$$

Mediator MTE is the causal effect of healthcare, relative to likelihood of visiting healthcare, $\Pr\left(D_i = 1 \,|\, \boldsymbol{X}_i, Z_i\right)$.

Outline:

(1) Gives a selection model by Vycatil (2002)

(2) IV separates first-stage identification from second

(3) Correlated errors connect D_i take-up with unobserved selection.

MTE Model — Identification

Theorem: Under assumptions **(1)**, **(2)**, **(3)** ADE + AIE are identified.

$$\text{ADE} = \mathbb{E}\left[\gamma + \delta D_i\right],$$

$$\text{AIE} = \mathbb{E}\left[\overline{\pi}\Big(\beta + \delta Z_i + \underbrace{(\rho_1 - \rho_0)\,\Gamma\big(\pi(0; \boldsymbol{X}_i),\, \pi(1; \boldsymbol{X}_i)\big)}_{=\widetilde{U}_i\text{, Mediator compliers}}\Big)\right].$$

where $\pi(z'; \boldsymbol{X}_i) = \Pr\left(D_i = 1 \,|\, \boldsymbol{X}_i, Z_i = z'\right)$ and $\Gamma(.,.)$ is a function that depends on the Mediator MTE.

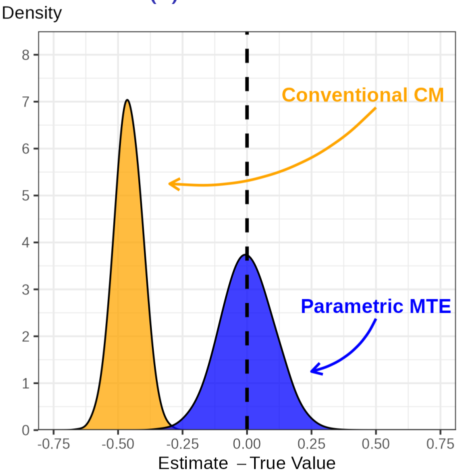
ADE Intuition:
Control for unobserved confounding via Mediator MTE.

AIE Intuition:
Extrapolate indirect effects across Mediator MTE.

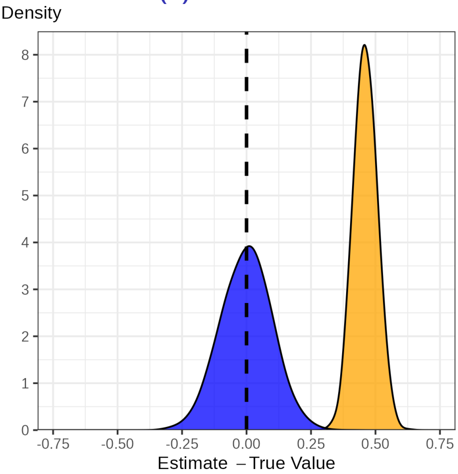
MTE Model — Estimation

Figure: CM Estimates from 10,000 DGPs with **Normal** Errors.

(a) $\widehat{\text{ADE}} - \text{ADE}$.

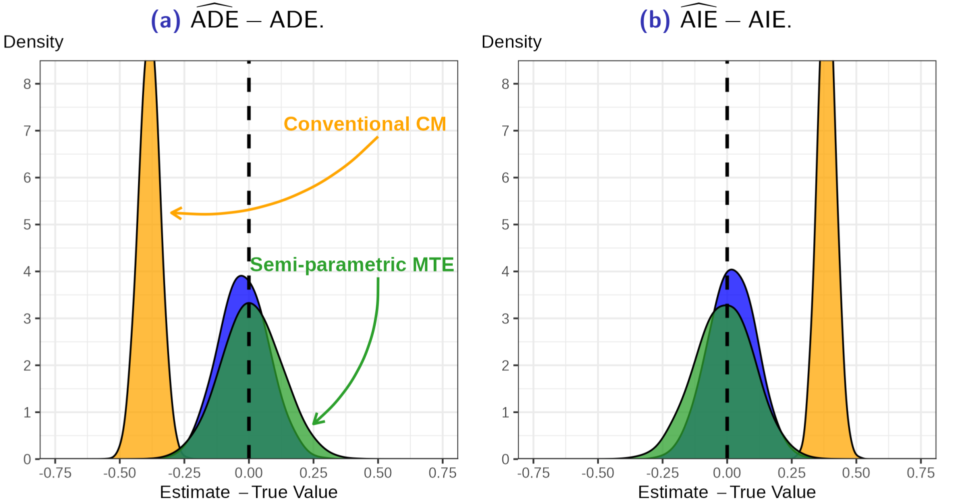


(b) $\widehat{\text{AIE}} - \text{AIE}$.



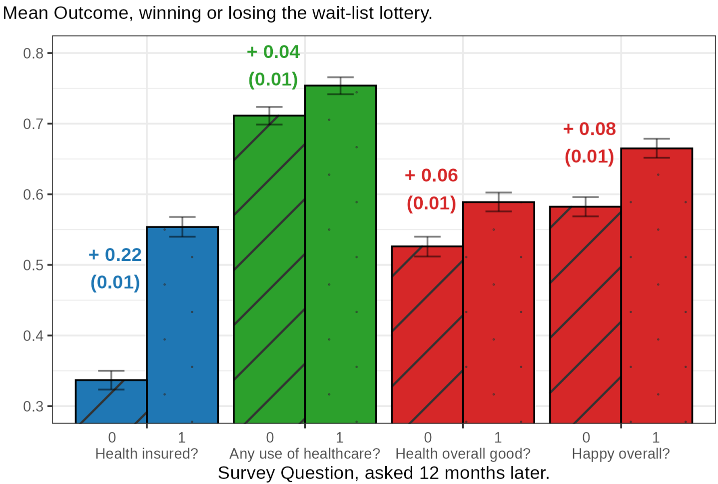
MTE Model — Estimation

Figure: CM Estimates from 10,000 DGPs with **Uniform** Errors.



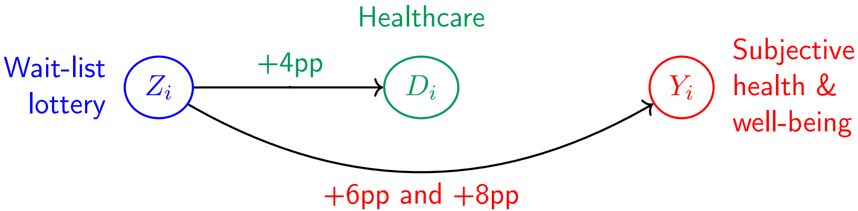
Return to Oregon

Winning the wait-list lottery significantly increased healthcare usage, plus subjective health and well-being (Finkelstein et al, 2012).



Return to Oregon

Winning the wait-list lottery significantly increased healthcare usage, plus subjective health and well-being (Finkelstein et al, 2012).

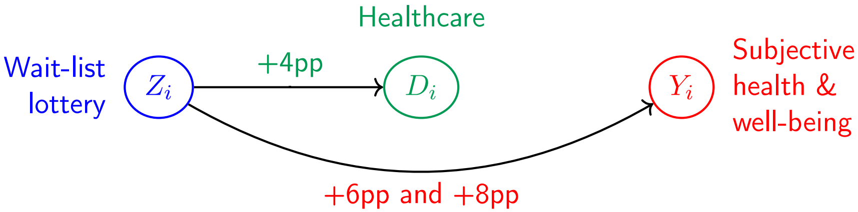


Suggestive evidence: Healthcare is a mechanism.

Plausible direct effects: Stress reduction and psychological gains.

Return to Oregon

Winning the wait-list lottery significantly increased healthcare usage, plus subjective health and well-being (Finkelstein et al, 2012).

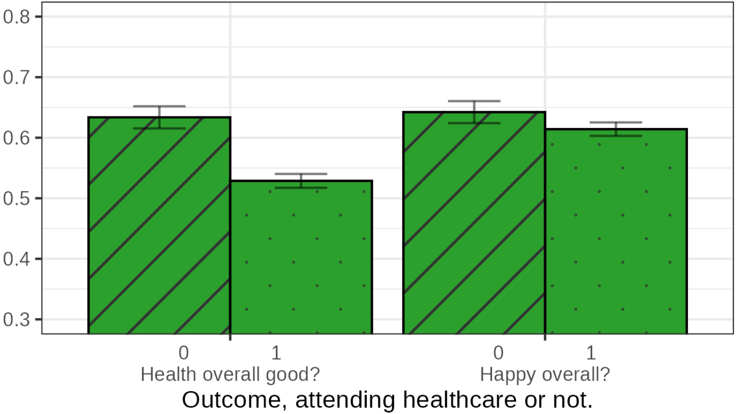


Suggestive evidence: Healthcare is a mechanism.

Plausible direct effects: Stress reduction and psychological gains.

Oregon — Conventional CM

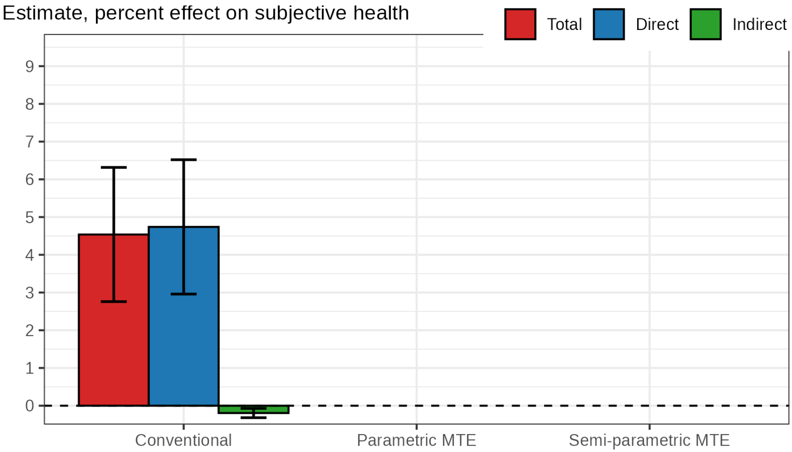
Does using healthcare improve subjective health and well-being?



- OLS estimate of $D_i \to Y_i$ is -10pp (1.1) and -2.5pp (1.1).
- Controls for initial health conditions gives -2.7pp (1.1) and +2.8 (1.1).

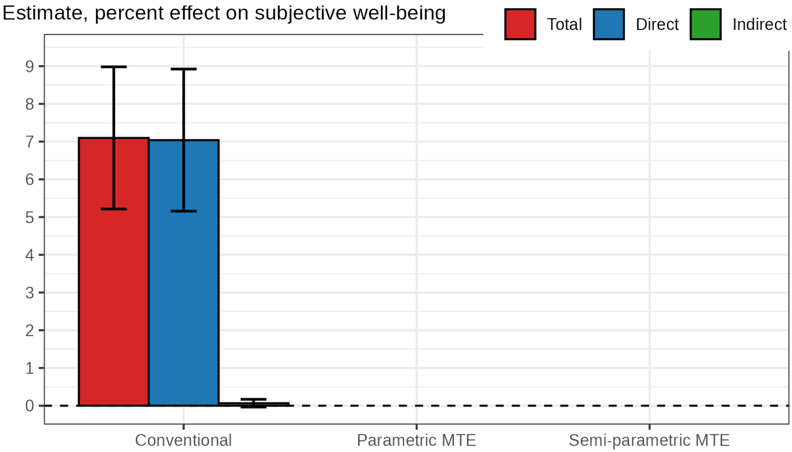
Oregon — Conventional CM

Conventional CM estimates lottery subjective health effects as mostly direct, ≈ 0 healthcare.



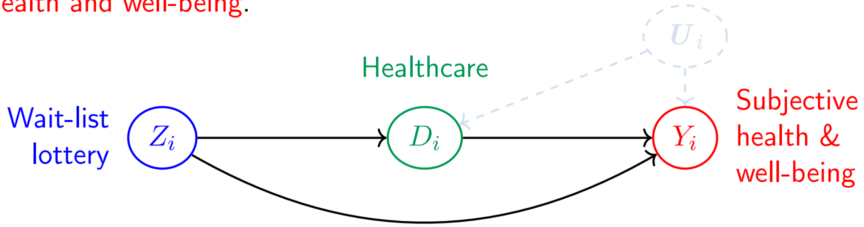
Oregon — Conventional CM

Conventional CM estimates lottery subjective well-being effects as mostly direct, ≈ 0 healthcare.



Oregon — Selection Bias

OLS Estimates had negative or little effects of healthcare → subjective health and well-being.

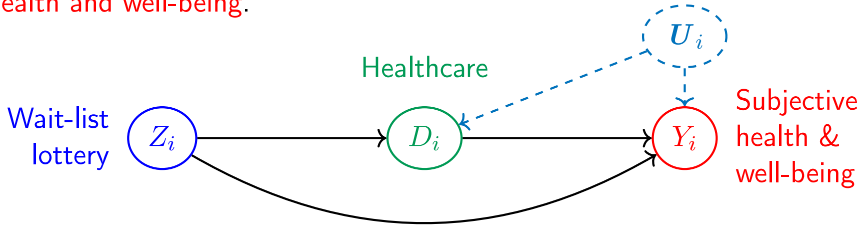


Example confounder: undiagnosed conditions among those without health insurance, near or below the poverty line.

Implication: negative selection bias in OLS estimates, and Conventional CM underestimated indirect healthcare channel.

Oregon — Selection Bias

OLS Estimates had negative or little effects of healthcare → subjective health and well-being.

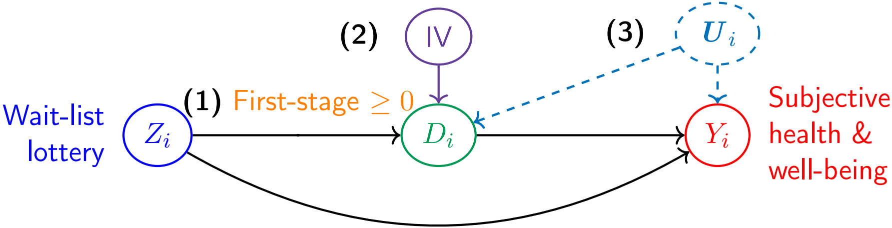


Example confounder: undiagnosed conditions among those without health insurance, near or below the poverty line.

Implication: negative selection bias in OLS estimates, and Conventional CM underestimated indirect healthcare channel.

Oregon — MTE Model

I bring the MTE model to these data instead.



Healthcare IV: pre-lottery healthcare location.

Intuition: Different locations charge different prices for similar healthcare. Heading to A&E costs more than a local doctor's office.

Oregon — MTE Model

Oregon Health Insurance applicants asked pre-lottery healthcare location.

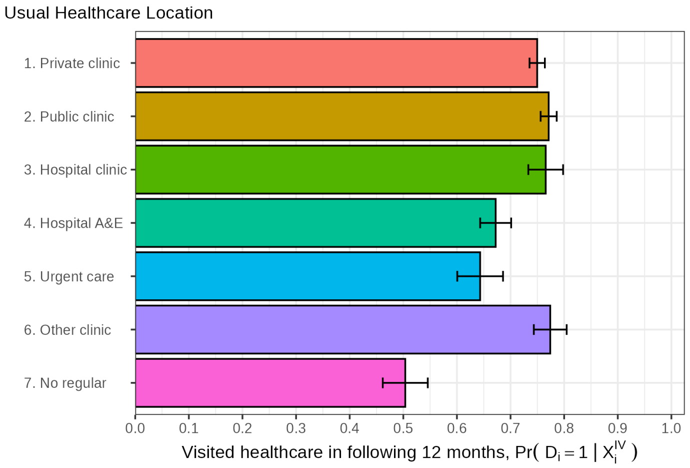
Survey question: Where do you usually go to receive medical care?

• A private doctor's office or clinic	35.5%
• A public health clinic, or community health centre	30.3%
• A hospital-based clinic	6.6%
• A hospital emergency room	10.2%
• An urgent care clinic	4.9%
• Other place not listed here	7.2%
• I don't have a usual place.	5.5%

IV assumption: where the uninsured, near poverty line, participants visits is indicative of their local healthcare access and cost.

Oregon — MTE Model

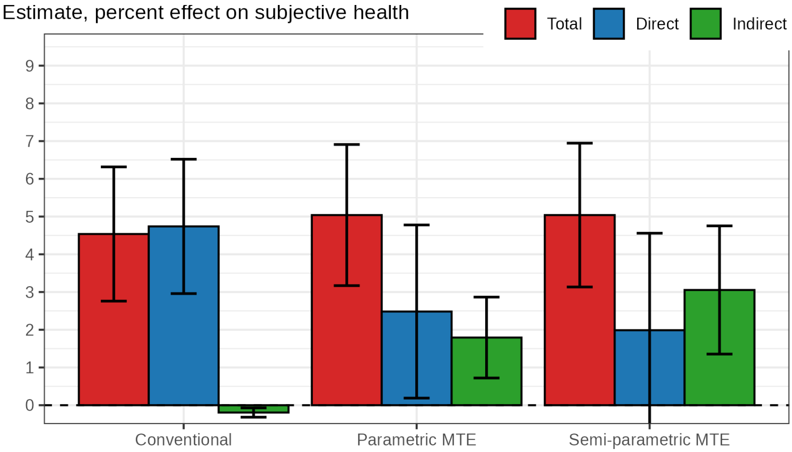
IV first-stage F stat. is 38.4, for all categories (minus base).



MTE Estimates of $D_i \to Y_i$ are +19.4pp (7.6) and +27.0pp (7.5).

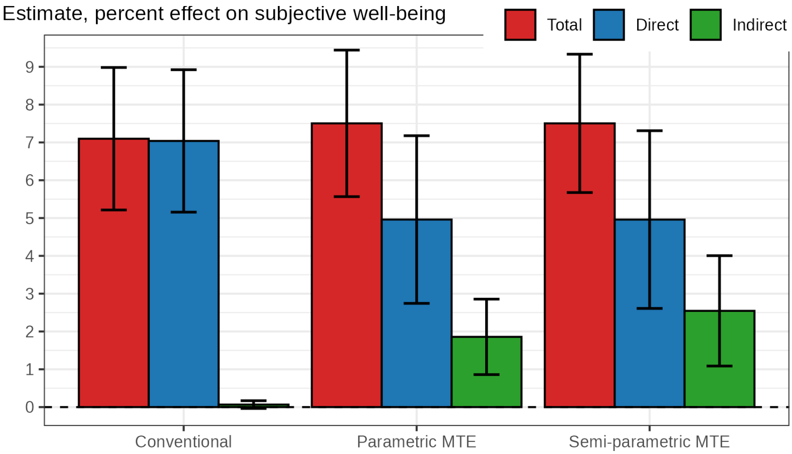
Oregon — MTE Model

Using MTE approach, with regular healthcare location IV, restores indirect effect through increasing healthcare visitation.



Oregon — MTE Model

Using MTE approach, with regular healthcare location IV, restores indirect effect through increasing healthcare visitation.



Conclusion

Overview:

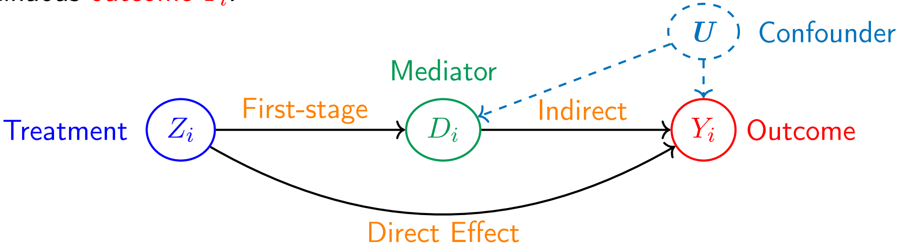
1. Selection bias in conventional CM analyses with no case for mediator (quasi-)random assignment.
2. Connect CM with labour theory + selection-into-treatment + MTEs.

Caveats and points to remember:

- Structural assumptions and IV for identification + estimation (not ideal)
- Application to Oregon Health Insurance Experiment, showing subjective health + well-being effects mediated by healthcare take-up
- **Credible** analyses of mechanisms are hard in practice, and wide confidence intervals show true uncertainty.

Appendix: CM Guiding Model

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



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

- ADE is causal effect $Z \to Y$, blocking the indirect D_i path.

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

- AIE is causal effect of $D_i(Z_i) \to Y_i$, blocking the direct Z_i path.

Group Difference — ADE

CM effects contaminated by (less interpretable) bias terms.

$$\text{CM Estimand} = \text{ADEM} + \text{Selection Bias}$$

$$\underbrace{\mathbb{E}_{D_i}\Big[\mathbb{E}\left[Y_i \,|\, Z_i = 1, D_i\right] - \mathbb{E}\left[Y_i \,|\, Z_i = 0, D_i\right]\Big]}_{\text{Estimand, Direct Effect}}$$

$$= \underbrace{\mathbb{E}_{D_i = d'}\left[\mathbb{E}\left[Y_i(1, D_i(Z_i)) - Y_i(0, D_i(Z_i)) \,\middle|\, D_i(1) = d'\right]\right]}_{\text{Average Direct Effect on Mediator (ADEM) take-up — i.e., } D_i(1) \text{ weighted}}$$

$$+ \underbrace{\mathbb{E}_{D_i}\Big[\mathbb{E}\left[Y_i(0, D_i(Z_i)) \,\middle|\, D_i(1) = d'\right] - \mathbb{E}\left[Y_i(0, D_i(Z_i)) \,\middle|\, D_i(0) = d'\right]\Big]}_{\text{Selection Bias}}$$

The weighted ADE you get here is a positive weighted sum of local ADEs, but with policy irrelevant weights $D_i(1) = d'$.

$\implies$ consider this group bias, noting difference from true ADE. Back

Selection Bias — Direct Effect

CM Effects + contaminating bias.

$$\text{CM Estimand} = \text{ADE} + \Big(\text{Selection Bias} + \text{Group difference bias}\Big)$$

Model

$$\underbrace{\mathbb{E}_{D_i=d'}\Big[\mathbb{E}\left[Y_i \,\middle|\, Z_i = 1, D_i = d'\right] - \mathbb{E}\left[Y_i \,\middle|\, Z_i = 0, D_i = d'\right]\Big]}_{\text{Estimand, Direct Effect}}$$

$$= \underbrace{\mathbb{E}\left[Y_i(1, D_i(Z_i)) - Y_i(0, D_i(Z_i))\right]}_{\text{Average Direct Effect}}$$

$$+ \underbrace{\mathbb{E}_{D_i=d'}\Big[\mathbb{E}\left[Y_i(0, D_i(Z_i)) \,\middle|\, D_i(1) = d'\right] - \mathbb{E}\left[Y_i(0, D_i(Z_i)) \,\middle|\, D_i(0) = d'\right]\Big]}_{\text{Selection Bias}}$$

$$+ \underbrace{\mathbb{E}_{D_i=d'}\left[\begin{array}{l}\Big(1 - \Pr\left(D_i(1) = d'\right)\Big) \\ \times \begin{pmatrix}\mathbb{E}\left[Y_i(1, D_i(Z_i)) - Y_i(0, D_i(Z_i)) \,\middle|\, D_i(1) = 1 - d'\right] \\ - \mathbb{E}\left[Y_i(1, D_i(Z_i)) - Y_i(0, D_i(Z_i)) \,\middle|\, D_i(0) = d'\right]\end{pmatrix}\end{array}\right]}_{\text{Group difference bias}}$$

Group-diff

Group Difference — AIE

CM effects contaminated by (less interpretable) bias terms.

$$\text{CM Estimand} = \text{AIEM} + \Big(\text{Selection Bias} + \text{Group difference bias}\Big)$$

$$\underbrace{\mathbb{E}_{Z_i}\left[\Big(\mathbb{E}\left[D_i \mid Z_i = 1\right] - \mathbb{E}\left[D_i \mid Z_i = 0\right]\Big) \times \Big(\mathbb{E}\left[Y_i \mid Z_i, D_i = 1\right] - \mathbb{E}\left[Y_i \mid Z_i, D_i = 0\right]\Big)\right]}_{\text{Estimand, Indirect Effect}}$$

$$= \underbrace{\mathbb{E}\left[Y_i(Z_i, D_i(1)) - Y_i(Z_i, D_i(0)) \mid D_i = 1\right]}_{\text{Average Indirect Effect on Mediated (AIEM) — i.e., } D_i = 1 \text{ weighted}}$$

$$+ \underbrace{\overline{\pi}\Big(\mathbb{E}\left[Y_i(Z_i, 0) \mid D_i = 1\right] - \mathbb{E}\left[Y_i(Z_i, 0) \mid D_i = 0\right]\Big)}_{\text{Selection Bias}}$$

$$+ \underbrace{\overline{\pi}\left[\left(\frac{1 - \Pr\left(D_i(1) = 1, D_i(0) = 0\right)}{\Pr\left(D_i(1) = 1, D_i(0) = 0\right)}\right)\Big(\mathbb{E}\left[Y_i(Z_i, 1) - Y_i(Z_i, 0) \mid D_i(1) = 0 \text{ or } D_i(0) \right. \\ \left. - \mathbb{E}\left[Y_i(Z_i, 1) - Y_i(Z_i, 0)\right]\right.}_{\text{Groups difference Bias}}$$

The weighted AIE you get here is not a positive weighted sum of local AIEs, because the AIE is only about $D(Z)$ compliers. Model.

$\implies$ consider this group bias, noting difference from true AIE. Back

Selection Bias — Indirect Effect

CM Effects + contaminating bias, where $\overline{\pi} = \Pr\left(D_i(0) \neq D_i(1)\right)$.

$$\text{CM Estimand} = \text{AIE} + \Big(\text{Selection Bias} + \text{Group difference bias}\Big)$$

Model

$$\underbrace{\mathbb{E}_{Z_i}\left[\Big(\mathbb{E}\left[D_i \mid Z_i = 1\right] - \mathbb{E}\left[D_i \mid Z_i = 0\right]\Big) \times \Big(\mathbb{E}\left[Y_i \mid Z_i, D_i = 1\right] - \mathbb{E}\left[Y_i \mid Z_i, D_i = 0\right]\Big)\right]}_{\text{Estimand, Indirect Effect}}$$

$$= \underbrace{\mathbb{E}\left[Y_i(Z_i, D_i(1)) - Y_i(Z_i, D_i(0))\right]}_{\text{Average Indirect Effect}}$$

$$+ \underbrace{\overline{\pi}\Big(\mathbb{E}\left[Y_i(Z_i, 0) \mid D_i = 1\right] - \mathbb{E}\left[Y_i(Z_i, 0) \mid D_i = 0\right]\Big)}_{\text{Selection Bias}}$$

$$+ \underbrace{\overline{\pi}\left[\Big(1 - \Pr\left(D_i = 1\right)\Big)\begin{pmatrix}\mathbb{E}\left[Y_i(Z_i, 1) - Y_i(Z_i, 0) \mid D_i = 1\right] \\ - \mathbb{E}\left[Y_i(Z_i, 1) - Y_i(Z_i, 0) \mid D_i = 0\right]\end{pmatrix} + \left(\frac{1 - \Pr\left(D_i(1) = 1, D_i(0) = 0\right)}{\Pr\left(D_i(1) = 1, D_i(0) = 0\right)}\right)\begin{pmatrix}\mathbb{E}\left[Y_i(Z_i, 1) - Y_i(Z_i, 0) \mid D_i(Z_i) \neq Z_i\right] \\ - \mathbb{E}\left[Y_i(Z_i, 1) - Y_i(Z_i, 0)\right]\end{pmatrix}\right]}_{\text{Groups difference Bias}}$$

Group-diff

Semi-parametric Control Functions

Semi-parametric specifications for the CFs λ_0, λ_1 bring some complications to estimating the AIE.

$$\mathbb{E}\left[Y_i \mid Z_i, D_i = 0, \boldsymbol{X}_i\right] = \alpha + \gamma Z_i + \varphi\big(\boldsymbol{X}_i\big) + \rho_0 \lambda_0\big(\pi(Z_i; \boldsymbol{X}_i)\big),$$

$$\mathbb{E}\left[Y_i \mid Z_i, D_i = 1, \boldsymbol{X}_i\right] = (\alpha + \beta) + (\gamma + \delta) Z_i + \varphi\big(\boldsymbol{X}_i\big) + \rho_1 \lambda_1\big(\pi(Z_i; \boldsymbol{X}_i)\big)$$

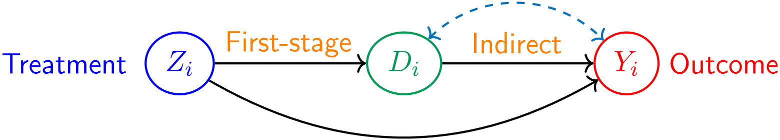
Intercepts, $\alpha, (\alpha + \beta)$, and relevance parameters ρ_0, ρ_1 are not separately identified from the CFs $\lambda_0(.), \lambda_1(.)$ so CF extrapolation term $(\rho_1 - \rho_0)\Gamma\big(\pi(0; \boldsymbol{X}_i),\, \pi(1; \boldsymbol{X}_i)\big)$ is not directly identified or estimable.

These problems can be avoided by estimating the AIE using its relation to the ATE, $\widehat{\text{AIE}}^{\text{CF}} =$

$$\widehat{\text{ATE}} - (1 - \overline{Z}) \underbrace{\left(\frac{1}{N} \sum_{i=1}^{N} \widehat{\gamma} + \widehat{\delta}\, \widehat{\pi}(1; \boldsymbol{X}_i) \right)}_{\widehat{\text{ADE}} \text{ given } Z_i = 1} - \overline{Z} \underbrace{\left(\frac{1}{N} \sum_{i=1}^{N} \widehat{\gamma} + \widehat{\delta}\, \widehat{\pi}(0; \boldsymbol{X}_i) \right)}_{\widehat{\text{ADE}} \text{ given } Z_i = 0}.$$

Appenidx: CM with Selection

Suppose Z_i is ignorable, D_i is not, so we have the following causal model.



Then this system has the following random coefficient equations:

$$D_i = \phi + \overline{\pi} Z_i + \varphi(\boldsymbol{X}_i) + U_i$$

$$Y_i = \alpha + \beta D_i + \gamma Z_i + \delta Z_i D_i + \zeta(\boldsymbol{X}_i) + \underbrace{(1 - D_i)\, U_{0,i} + D_i U_{1,i}}_{\text{Correlated error term}}$$

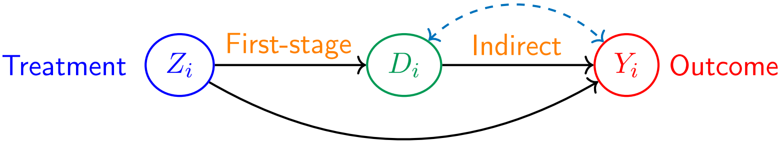
where β, γ, δ are functions of $\mu_{d'}(z'; \boldsymbol{X}_i)$.

$$\text{ADE} = \mathbb{E}\left[\gamma + \delta D_i\right], \quad \text{AIE} = \mathbb{E}\left[\overline{\pi}\left(\beta + \delta Z_i + \widetilde{U}_i\right)\right]$$

with $\widetilde{U}_i = \mathbb{E}\left[U_{1,i} - U_{0,i} \,|\, \boldsymbol{X}_i, D_i(0) \neq D_i(1)\right]$ unobserved complier gains.

Appenidx: CM with Selection

Suppose Z_i is ignorable, D_i is not, so we have the following causal model.



Main problem, second-stage is not identified:

$$
\begin{aligned}
\mathbb{E}\left[Y_i \,|\, Z_i, D_i, \boldsymbol{X}_i\right] &= \alpha + \beta D_i + \gamma Z_i + \delta Z_i D_i + \varphi(\boldsymbol{X}_i) \\
&\quad \underbrace{+ \left(1 - D_i\right) \mathbb{E}\left[U_{0,i} \,|\, D_i = 0, \boldsymbol{X}_i\right] + D_i \mathbb{E}\left[U_{1,i} \,|\, D_i = 1, \boldsymbol{X}_i\right]}_{\text{Unobserved } D_i \text{ confounding.}}
\end{aligned}
$$

Identification intuition: Identify second-stage via MTE control function.

Appenidx: CM with Selection — Identification

Assume:

1. Mediator monotonicity, $\Pr\left(D_i(0) \leq D_i(1) \,|\, \boldsymbol{X}_i\right) = 1$

$$\implies D_i(z') = \mathbb{1}\left\{U_i \leq \pi(z'; \boldsymbol{X}_i)\right\}, \quad \text{for } z' = 0, 1 \text{ (Vycatil 2002).}$$

2. Selection on mediator benefits, $\text{Cov}\left(U_i,\, U_{0,i}\right),\ \text{Cov}\left(U_i,\, U_{1,i}\right) \neq 0$

$\implies$ First-stage take-up informs second-stage confounding.

3. There is an IV for the mediator, $\boldsymbol{X}_i^{\text{IV}}$ among control variables $\boldsymbol{X}_i$.

$$\implies \pi(Z_i; \boldsymbol{X}_i) = \Pr\left(D_i = 1 \,|\, Z_i, \boldsymbol{X}_i\right) \text{ is separately identified.}$$

Proposition:

$$\begin{aligned} &\mathbb{E}\left[Y_i(z', 1) - Y_i(z', 0) \,\middle|\, Z_i = z', \boldsymbol{X}_i, U_i = p'\right] \\ &= \beta + \delta z' + \mathbb{E}\left[U_{1,i} - U_{0,i} \,\middle|\, \boldsymbol{X}_i, U_i = p'\right], \qquad \text{for } p' \in (0, 1). \end{aligned}$$

Appenidx: CM with Selection — Identification

The marginal effect has corresponding Control Functions (CFs), describing unobserved selection-into-D_i,

$$\rho_0 \lambda_0(p') = \mathbb{E}\left[U_{0,i} \,\middle|\, p' \leq U_i\right], \quad \rho_1 \lambda_1(p') = \mathbb{E}\left[U_{1,i} \,\middle|\, U_i \leq p'\right].$$

These CFs restore second-stage identification, by extrapolating from $\boldsymbol{X}_i^{\text{IV}}$ compliers to $D_i(Z_i)$ mediator compliers,

$$\begin{aligned} \mathbb{E}\left[Y_i \,|\, Z_i, D_i, \boldsymbol{X}_i\right] \;&=\; \alpha + \beta D_i + \gamma Z_i + \delta Z_i D_i + \varphi(\boldsymbol{X}_i) \\ &\quad + \underbrace{\rho_0 \left(1 - D_i\right) \lambda_0\big(\pi(Z_i; \boldsymbol{X}_i)\big) + \rho_1 D_i \lambda_1\big(\pi(Z_i; \boldsymbol{X}_i)\big)}_{\text{CF adjustment.}} \end{aligned}$$

This adjusted second-stage re-identifies the ADE and AIE,

$$\text{ADE} = \mathbb{E}\left[\gamma + \delta D_i\right], \;\; \text{AIE} = \mathbb{E}\left[\overline{\pi}\left(\beta + \delta Z_i + \underbrace{\left(\rho_1 - \rho_0\right) \Gamma\big(\pi(0; \boldsymbol{X}_i),\, \pi(1; \boldsymbol{X}_i)\big)}_{\text{Mediator compliers extrapolation.}}\right)\right]$$

Appenidx: CM with Selection — Estimation

Will explain how estimation works, with simulation evidence.

1. Random treatment $Z_i \sim \text{Binom}(0.5)$, for $n = 5,000$.
2. $(U_{0,i}, U_{1,i}) \sim \text{BivariateNormal}(0, 0, \sigma_0, \sigma_1, \rho)$, Costs $C_i \sim N(0, 0.5)$.

Roy selection-into-D_i, with constant partial effects + interaction term.

$$D_i(z') = \mathbb{1}\left\{C_i \leq Y_i(z', 1) - Y_i(z', 0)\right\},$$
$$Y_i(z', d') = \left(z' + d' + z'd'\right) + U_{d'} \qquad \text{for } z', d' = 0, 1.$$

Following the previous, these data have the following first and second-stage equations, where $\boldsymbol{X}_i^{\text{IV}}$ is an additive cost IV:

$$D_i = \mathbb{1}\left\{C_i - \left(U_{1,i} - U_{0,i}\right) \leq Z_i - \boldsymbol{X}_i^{\text{IV}}\right\}$$
$$Y_i = Z_i + D_i + Z_i D_i + (1 - D_i) U_{0,i} + D_i U_{1,i}.$$

$\implies$ unobserved confounding by BivariateNormal $\left(U_{0,i}, U_{1,i}\right)$.

Appenidx: CM with Selection — Estimation

Errors are normal, so system is Heckman (1979) selection model.

CFs are the inverse Mills ratio, with $\phi(.)$ normal pdf and $\Phi(.)$ normal cdf,

$$\lambda_0(p') = \frac{\phi(-\Phi^{-1}(p'))}{\Phi(-\Phi^{-1}(p'))}, \quad \lambda_1(p') = \frac{\phi(\Phi^{-1}(p'))}{\Phi(\Phi^{-1}(p'))}, \quad \text{for } p' \in (0,1).$$

Parametric Estimation Recipe:

1. Estimate first-stage $\pi(Z_i; \boldsymbol{X}_i)$ with probit, including $\boldsymbol{X}_i^{\text{IV}}$.
2. Include λ_0, λ_1 CFs in second-stage OLS estimation.
3. Compose CM estimates from two-stage plug-in estimates.

$\rightarrow$ Same as conventional CM estimates (two-stages), with CFs added.

$$\widehat{\text{ADE}} = \mathbb{E}\left[\widehat{\gamma} + \widehat{\delta} D_i\right], \quad \widehat{\text{AIE}} = \mathbb{E}\left[\widehat{\overline{\pi}}\left(\widehat{\beta} + \widehat{\delta} Z_i + \underbrace{\left(\widehat{\rho}_1 - \widehat{\rho}_0\right)\Gamma\left(\widehat{\pi}(0; \boldsymbol{X}_i), \widehat{\pi}(1; \boldsymbol{X}_i)\right)}_{\text{Mediator compliers extrapolation.}}\right)\right]$$

Appenidx: CM with Selection — Estimation

If errors are not normal, then CFs do not have a known form, so semi-parametrically estimate them (e.g., splines).

$$\mathbb{E}\left[Y_i \,|\, Z_i, D_i = 0, \boldsymbol{X}_i\right] = \alpha + \gamma Z_i + \varphi\big(\boldsymbol{X}_i\big) + \rho_0 \lambda_0\big(\pi(Z_i; \boldsymbol{X}_i)\big),$$

$$\mathbb{E}\left[Y_i \,|\, Z_i, D_i = 1, \boldsymbol{X}_i\right] = (\alpha + \beta) + (\gamma + \delta) Z_i + \varphi\big(\boldsymbol{X}_i\big) + \rho_1 \lambda_1\big(\pi(Z_i; \boldsymbol{X}_i)\big)$$

Semi-parametric Estimation Recipe:

1. Estimate first-stage $\pi(Z_i; \boldsymbol{X}_i)$, including $\boldsymbol{X}_i^{\mathsf{IV}}$.
2. Estimate second-stage separately for $D_i = 0$ and $D_i = 1$, with regressors $\lambda_0(p'), \lambda_1(p')$, semi-parametric in $\widehat{\pi}(Z_i; \boldsymbol{X}_i)$.
3. Compose CM estimates from two-stage plug-in estimates.

$\rightarrow$ Same as conventional CM estimates, with semi-parametric CFs. ▸ CFs.

$$\widehat{\mathsf{ADE}} = \mathbb{E}\left[\widehat{\gamma} + \widehat{\delta} D_i\right], \quad \widehat{\mathsf{AIE}} = \mathbb{E}\left[\widehat{\overline{\pi}}\left(\widehat{\beta} + \widehat{\delta} Z_i + \left(\widehat{\rho}_1 - \widehat{\rho}_0\right)\Gamma\big(\widehat{\pi}(0; \boldsymbol{X}_i), \,\widehat{\pi}(1; \boldsymbol{X}_i)\big)\right)\right]$$

Appenidx: CM with Selection — Estimation

Figure: CF Adjusted Estimates Work with Different Error Term Parameters.

(a) ADE. (b) AIE.