Causal Inference in Three-Wave Panel Designs

Joseph.Bulbulia@vuw.ac.nz // Victoria University

Section 1

The Fundamental Problem of Causal Inference

Causality Requires a Contrast of Exposures

$$Y_{\text{you}}(1) - Y_{\text{you}}(0)$$

But Individuals Experience Only **One** Exposure

$$Y_i|A_i = 1 \implies Y_i(0)|A_i = 1 \text{ is counterfactual}$$

Average Treatment Effect in Randomised Controlled Experiments Work From Assumptions

$$\text{Average Treatment Effect} = \left[\left(\underbrace{\mathbb{E}[Y(1)|A=1]}_{\text{observed}} + \underbrace{\mathbb{E}[Y(1)|A=0]}_{\text{unobserved}} \right) - \left(\underbrace{\mathbb{E}[Y(0)|A=0]}_{\text{observed}} + \underbrace{\mathbb{E}[Y(0)|A=1]}_{\text{unobserved}} \right) \right]$$

Section 2

The Three Fundemental Assumptions of Causal Inference

The Three Fundemental Assumptions of Causal Inference

Causal Consistency

$$Y_i^{observed}|A_i = \begin{cases} Y_i(a^*) & \text{if } A_i = a^* \\ Y_i(a) & \text{if } A_i = a \end{cases}$$

Conditional Exchangeabilty

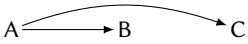
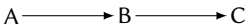
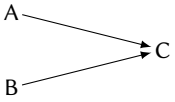
$$Y(a) \coprod A|L \quad \text{or equivalently} \quad A \coprod Y(a)|L$$

Positivity

$$0 < Pr(A = a|L = l) < 1$$

Section 3

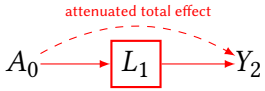
Causal Diagrams & Structures of Confounding Bias

		Two variables		
1	**Causality Absent**	A B	A and B have no causal effect on each other.	$A \coprod B$ (independent)
2	**Causality**	A → B	A causally affects B, and they are associated.	$A \not\coprod B$ (dependent)
		Three variables		
3	**Fork**		A causally affects both B and C; B and C are conditionally independent given A.	$B \coprod C \mid A$
4	**Chain**		C is affected by B which is, in turn, affected by A; A and C are conditionally independent given B.	$A \coprod C \mid B$
5	**Collider**		C is affected by both A and B, which are independent; conditioning on C induces association between A and B.	$A \not\coprod B \mid C$

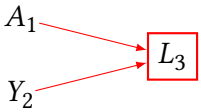
Counfounding

Reverse Causation

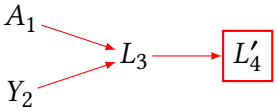
Mediator Bias



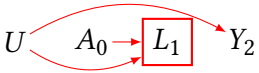
Collider Bias



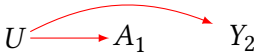
Collider Bias Proxy



Post Exposure Collider Bias



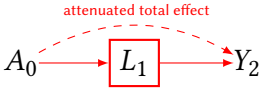
Unmeasured Common Cause

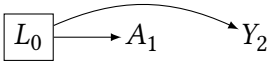


Reverse Causation Solution: Longitudional Hygiene

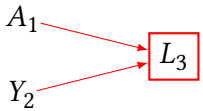
A_1 Y_2

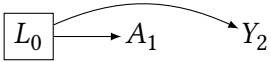
Mediator Bias Solution: Longitudional Hygiene





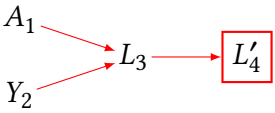
Collider Bias Solution: Longitudional Hygiene

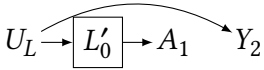




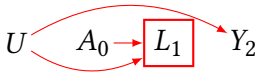
More Confounding Solutions: Longitudinal Hygiene

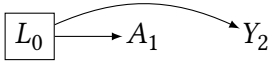
Collider Bias Proxy Solution: Longitudional Hygiene



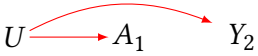


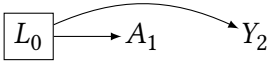
Post Exposure Collider Bias Solution: Longitudional Hygiene





Unmeasured Common Cause Solution: Longitudional Hygiene





Section 4

New Zealand Attitudes and Values Study

The New Zealand Attitudes and Values Study

- Planned 20-year longitudinal study, currently in its 14^{th} year.

The New Zealand Attitudes and Values Study

- Planned 20-year longitudinal study, currently in its 14^{th} year.
- Sample frame drawn randomly from NZ Electoral Roll.

The New Zealand Attitudes and Values Study

- Planned 20-year longitudinal study, currently in its 14^{th} year.
- Sample frame drawn randomly from NZ Electoral Roll.
- Postal questionnaire (coverage; retention ~ 80%)

The New Zealand Attitudes and Values Study

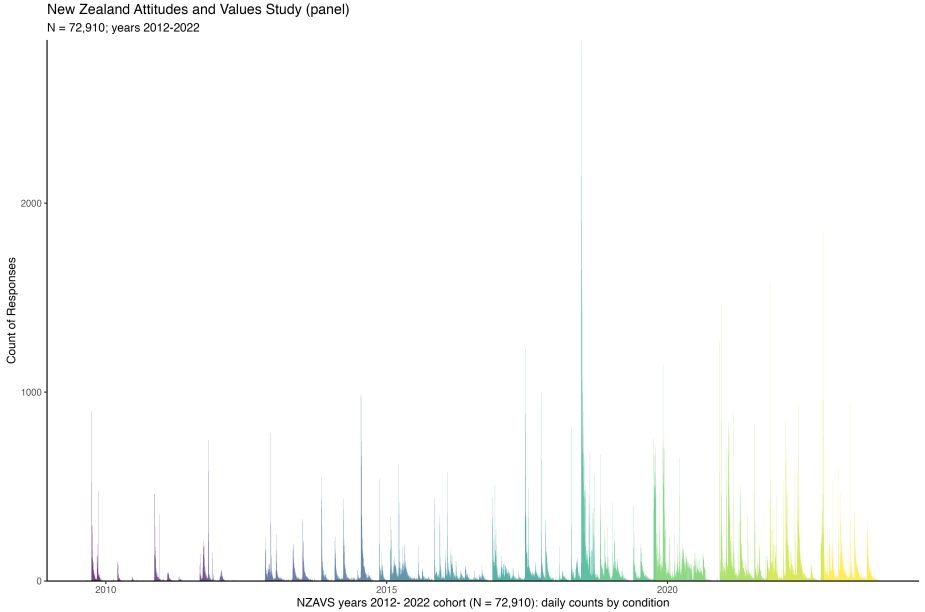
- Planned 20-year longitudinal study, currently in its 14^{th} year.
- Sample frame drawn randomly from NZ Electoral Roll.
- Postal questionnaire (coverage; retention ~ 80%)
- Large multidisciplinary research team (40 +)

The New Zealand Attitudes and Values Study

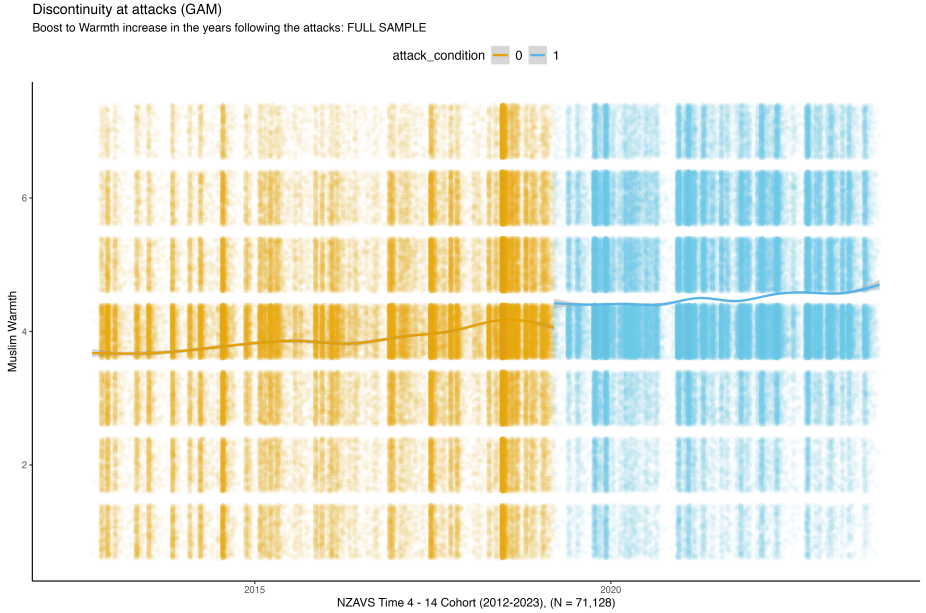
- Planned 20-year longitudinal study, currently in its 14^{th} year.
- Sample frame drawn randomly from NZ Electoral Roll.
- Postal questionnaire (coverage; retention ~ 80%)
- Large multidisciplinary research team (40 +)
- Focus on personality, social attitudes, values, religion, employment, prejudice ...

The New Zealand Attitudes and Values Study

- Planned 20-year longitudinal study, currently in its 14^{th} year.
- Sample frame drawn randomly from NZ Electoral Roll.
- Postal questionnaire (coverage; retention ~ 80%)
- Large multidisciplinary research team (40 +)
- Focus on personality, social attitudes, values, religion, employment, prejudice ...
- Current sample contains > 72,290,000 unique people, and ~ 38,000 in the longitudinal study


New Zealand Attitudes and Values Study (panel)
N = 72,910; years 2012-2022
Count of Responses
0
1000
2000
2010
2015
2020
NZAVS years 2012- 2022 cohort (N = 72,910): daily counts by condition

BIG EVENTS


Discontinuity at attacks (GAM)
Boost to Warmth increase in the years following the attacks: FULL SAMPLE
attack_condition 0 1
Muslim Warmth
6
4
2
2015
2020
NZAVS Time 4 - 14 Cohort (2012-2023), (N = 71,128)

Gradual Events

Religious Service Attendence

Socialising

Cash Value of Religious Service

Thanks

- Chris G. Sibley (NZAVS lead Investigator)

Thanks

- Chris G. Sibley (NZAVS lead Investigator)
- TRT Grant 0418

Thanks

- Chris G. Sibley (NZAVS lead Investigator)
- TRT Grant 0418
- Max Planck Institute for Evolutionary Anthropology

Thanks

- Chris G. Sibley (NZAVS lead Investigator)
- TRT Grant 0418
- Max Planck Institute for Evolutionary Anthropology
- Victoria University

Thanks

- Chris G. Sibley (NZAVS lead Investigator)
- TRT Grant 0418
- Max Planck Institute for Evolutionary Anthropology
- Victoria University
- University of Canterbury

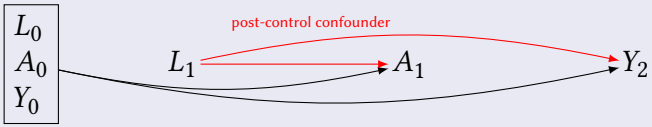
Thanks

- Chris G. Sibley (NZAVS lead Investigator)
- TRT Grant 0418
- Max Planck Institute for Evolutionary Anthropology
- Victoria University
- University of Canterbury
- 72,290 NZAVS participants

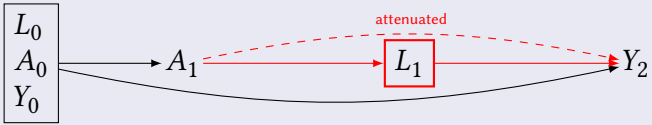
Extra Slides

Longitudinal Data Bring Their Own Problems

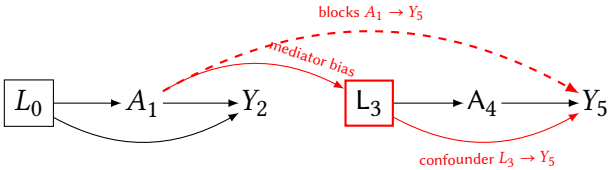
Timing of Confounder



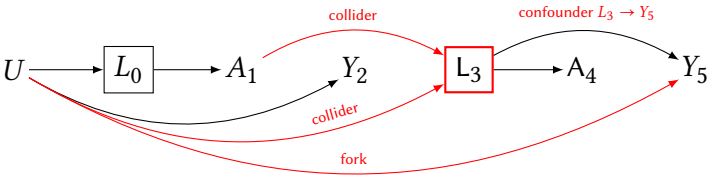
Timing of Mediator



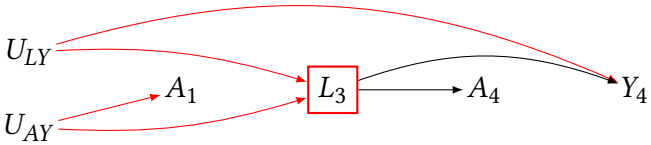
Treatment Confounder Bias



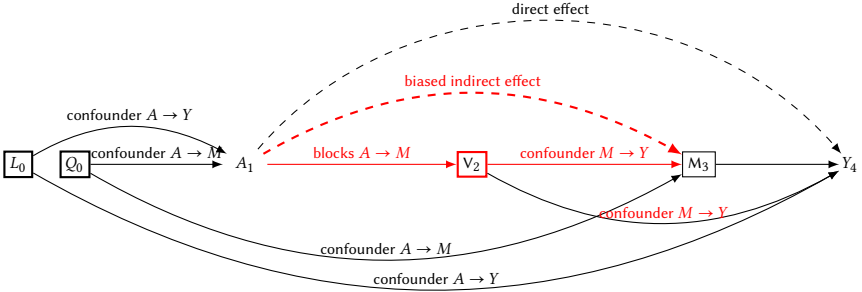
Treatment Confounder Feedback



Treatment Confounder Feedback Variation



Mediation


direct effect
biased indirect effect
confounder $A \rightarrow Y$
L_0
Q_0
confounder $A \rightarrow M$
A_1
blocks $A \rightarrow M$
V_2
confounder $M \rightarrow Y$
M_3
Y_4
confounder $M \rightarrow Y$
confounder $A \rightarrow M$
confounder $A \rightarrow Y$

Section 5

Mediation

Total Effect

$$TE = \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)]$$

Total Effect Considering Mediator

$$TE = \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)]$$

$$\mathbb{E}[Y(1)] = \mathbb{E}[Y(1, M(1))]$$

Natural Direct Effect

Natural Direct Effect (NDE) is the effect of the treatment on the outcome while maintaining the mediator at the level it would have been if the treatment had *not* been applied:

$$NDE = \mathbb{E}[Y(1, M(0))] - \mathbb{E}[Y(0, M(0))]$$

Natural Indirect Effect

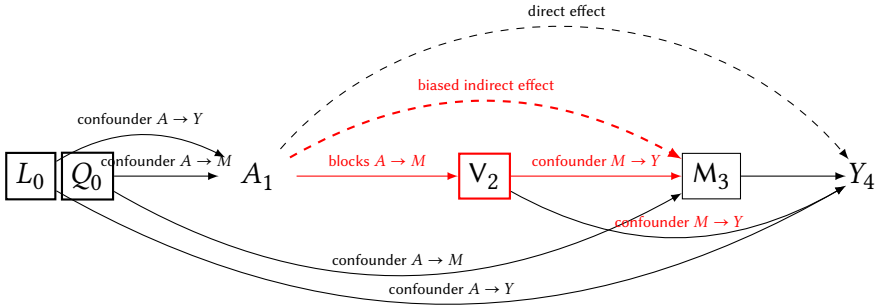
Natural Indirect Effect (NIE): is the effect of the exposure on the outcome that is mediated. To obtain these quantities we must compare the potential outcome Y under treatment, where the mediator assumes its natural level under treatment with the potential outcome when the mediator assumes its natural value under no treatment is given:

$$NIE = \mathbb{E}[Y(1, M(1))] - \mathbb{E}[Y(1, M(0))]$$

Decomposition

$$\text{Total Effect (TE)} = \underbrace{\Big\{\mathbb{E}[Y(1, M(1))] - \mathbb{E}[Y(1, M(0))]\Big\}}_{\text{Natural Indirect Effect (NIE)}} + \underbrace{\Big\{\mathbb{E}[Y(1, M(0))] - \mathbb{E}[Y(0, M(0))]\Big\}}_{\text{Natural Direct Effect (NDE)}}$$

Why Mediation is Difficult



Section 6

Interaction

Interaction: potential outcomes

$$\Big(\underbrace{\mathbb{E}[Y(1,1)]}_{\text{joint exposure}} - \underbrace{\mathbb{E}[Y(0,0)]}_{\text{neither exposed}} \Big) - \Big[\Big(\underbrace{\mathbb{E}[Y(1,0)]}_{\text{only A exposed}} - \underbrace{\mathbb{E}[Y(0,0)]}_{\text{neither exposed}} \Big) + \Big(\underbrace{\mathbb{E}[Y(0,1)]}_{\text{only B exposed}} - \underbrace{\mathbb{E}[Y(0,0)]}_{\text{neither exposed}} \Big) \Big] \neq 0$$

Interaction: simplifies to

$$\underbrace{\mathbb{E}[Y(1,1)]}_{\text{joint exposure}} - \underbrace{\mathbb{E}[Y(1,0)]}_{\text{only A exposed}} - \underbrace{\mathbb{E}[Y(0,1)]}_{\text{only B exposed}} + \underbrace{\mathbb{E}[Y(0,0)]}_{\text{neither exposed}} \neq 0$$

Key

Symbol	
X	Any variable.
A	The treatment or, equivalently, the exposure.
Y	The outcome.
$Y(a)$	The potential outcome when $A = a$.
L	Measured confounder(s): typically comprises a set of variables.
U	Unmeasured confounder.
Z	Effect-modifier (or 'moderator') of A on Y.
M	Mediator of A on Y.
$\bar{X}$	Sequential variables, e.g. $\bar{A} = \{A_1, A_2, A_3\}; \bar{L} = \{L_0, L_1, L_2\}$.
$\mathscr{R}$	Denotes randomisation into treatment event.