# Developmental Causality Medley

Julia M. Rohrer
Leipzig University
juliarohrer.com
www.the100.ci

Slides: Resources

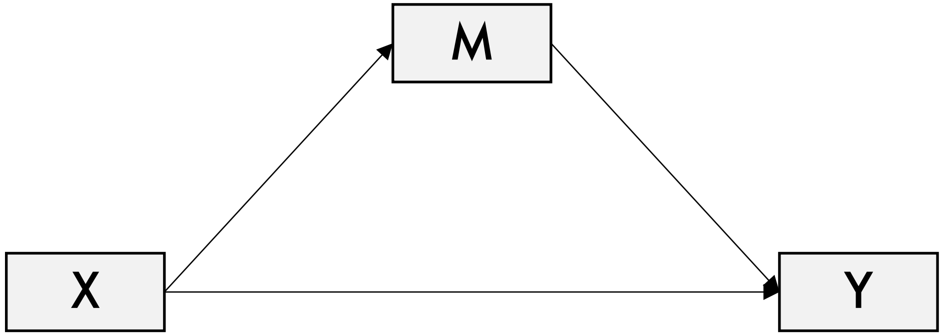
# Developmental Causality Medley

1. Mediation analysis
2. Causal inference with (observational) longitudinal data
3. The Age-Period-Cohort problem

# Mediation analysis

aka “Yes, but what’s the mechanism?”

# Mediation analysis is a causal inference problem




- Two central questions
  - To which extent does X affect Y via M? (Indirect effect)
  - To which extent does X affect Y apart from its effect via M? (direct effect)
- This is unambiguously a *causal* research question
  - Baron & Kenny, 1986; MacKinnon, 2008; Preacher, 2015
- That being said, there have been recent attempts to rebrand it as non-causal association analysis
  - Think of these papers as attempts of damage control

# Three steps of mediation analysis

## (or really any causal analysis), [Nguyen et al. (2020)]

1. Definition
   » Define the causal effects of interest
2. Identification
   » Figure out the assumptions under which the effects of interest can actually be identified from the data (assuming that infinite data were available)
3. Estimation
   » Actually estimate the effects of interest (using only the available finite data)

# Three steps of mediation analysis

## (or really any causal analysis), Nguyen et al. (2020)

In the „traditional“ mediation approach, this is a one-step procedure with a focus on estimation:

1. The effects of interest are defined by the parameters of the statistical model.
2. The identification assumption is essentially that the statistical model matches the actual model underlying reality.
3. Estimation via e.g. Baron & Kenny‘s (1986) three steps (not optimal), SEM (better), the regression-based approach implemented in PROCESS

# Three steps of mediation analysis

(or really any causal analysis), Nguyen et al. (2020)

In the „traditional“ mediation approach, this is a one-step procedure with a focus on estimation:

1. The effects of interest are defined by the parameters of the statistical model.
2. The identification assumption is essentially that the statistical model matches the actual model underlying reality.
3. Estimation via e.g. Baron & Kenny‘s (1986) three steps (not optimal), SEM (better), the regression-based approach implemented in PROCESS

# The central point of contention

- Mediation analysis rests on very strong identification assumptions – which are likely to be wrong in many applications
  - Bullock et al. (2010): Yes, but what's the mechanism? (Don't expect an easy answer)
  - Fiedler et al. (2011): What mediation analysis can (not) do
  - Rohrer et al. (2022): That's a lot to process! Pitfalls of popular paths models
  - Blog posts: [In psychology, everything mediates everything]; [Indirect Effect Ex Machina]; [Mediation analysis is counterintuitively invalid]; [That's a very nice mediation analysis you have there. It would be a shame if something happened to it].

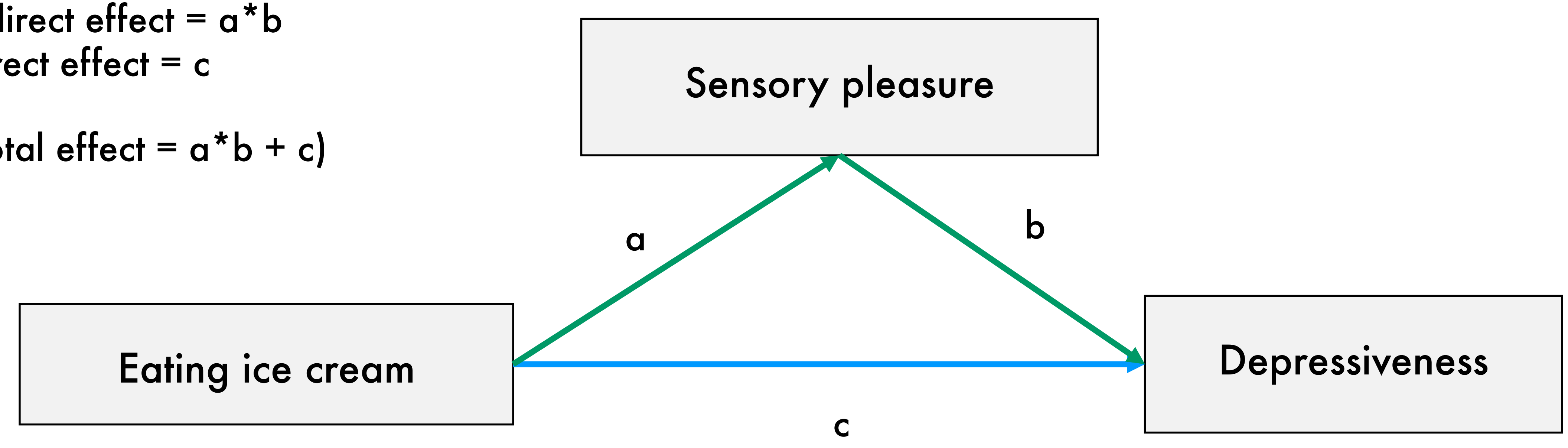
2. Identification:

We need to be able to estimate the causal effects a, b, and c *without any bias.*

1. Definition of the causal effects:

Indirect effect = a*b
Direct effect = c

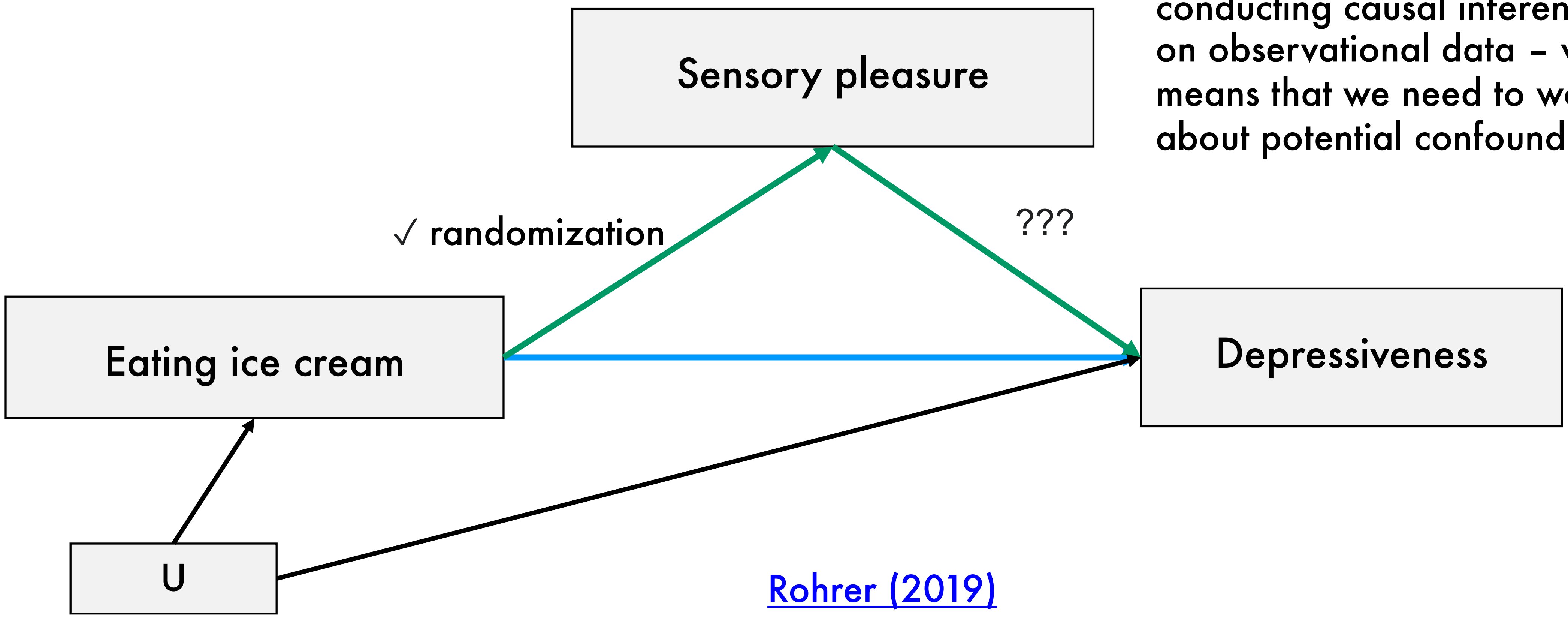
(Total effect = a*b + c)




Rohrer (2019)

Sensory pleasure has not been randomized.

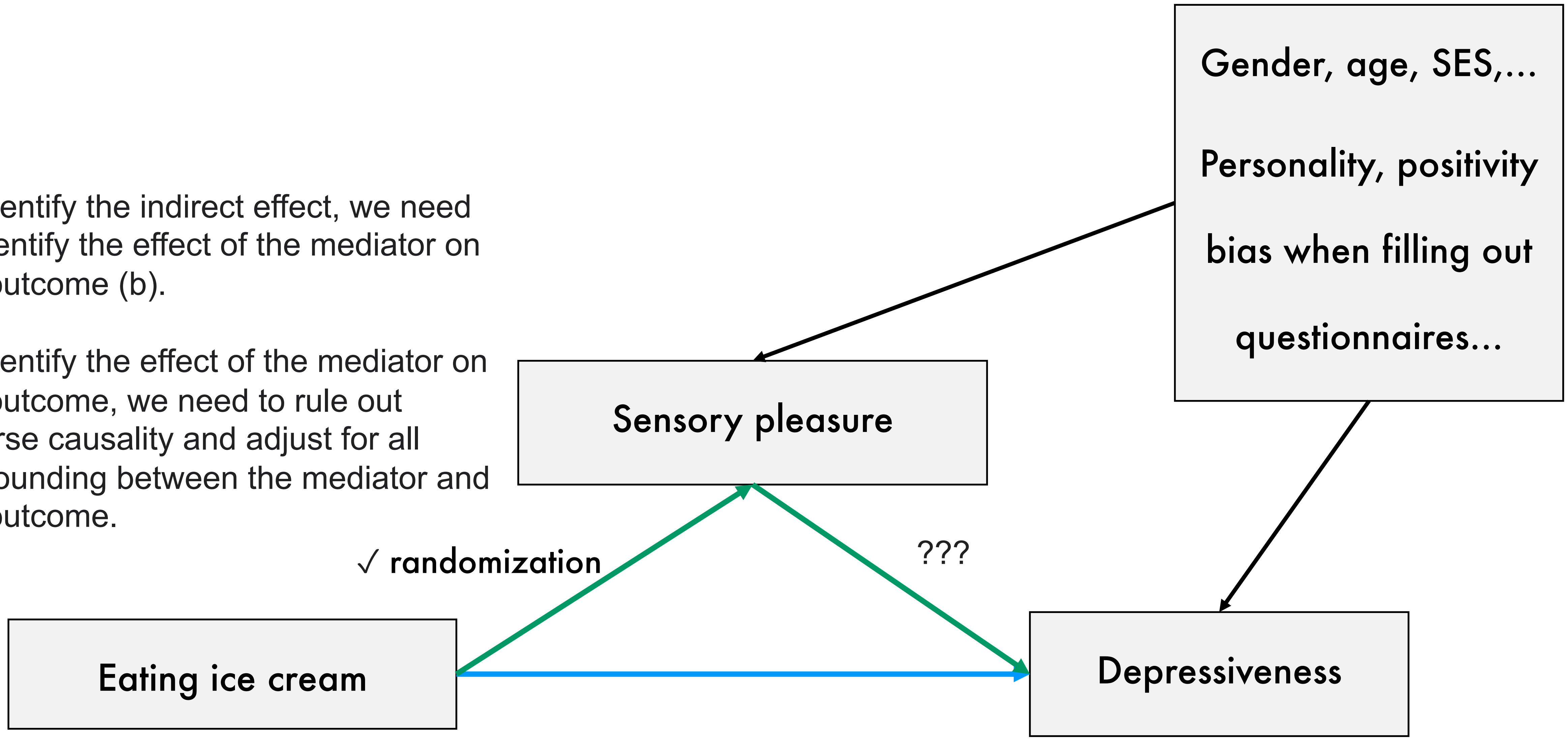
Thus, when we try to identify the b-path (Sensory Pleasure → Depressiveness), we are conducting causal inference based on observational data – which means that we need to worry about potential confounders.




Rohrer (2019)

To identify the indirect effect, we need to identify the effect of the mediator on the outcome (b).

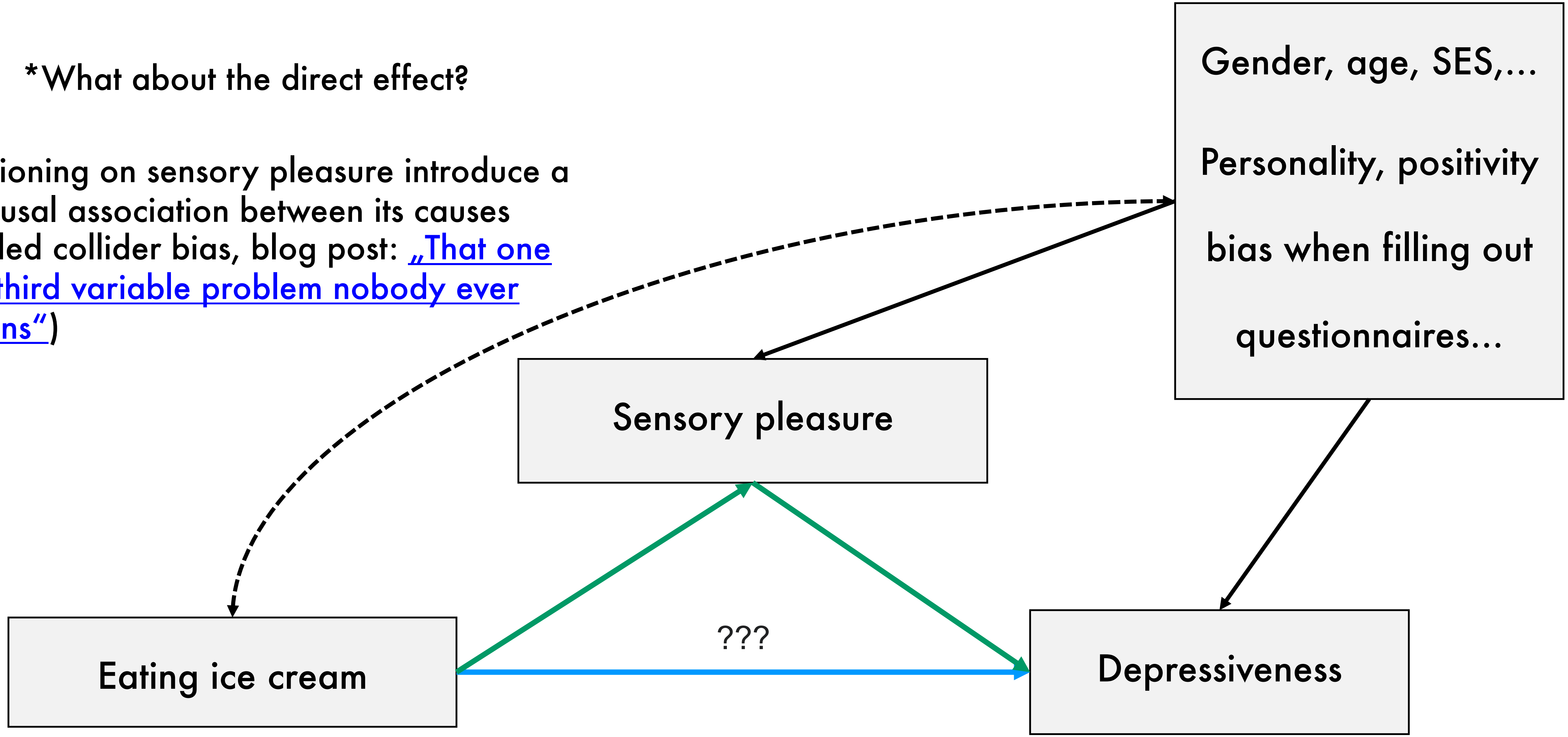
To identify the effect of the mediator on the outcome, we need to rule out reverse causality and adjust for all confounding between the mediator and the outcome.




Rohrer (2019)

*What about the direct effect?

Conditioning on sensory pleasure introduce a non-causal association between its causes (so-called collider bias, blog post: [„That one weird third variable problem nobody ever mentions“](#))



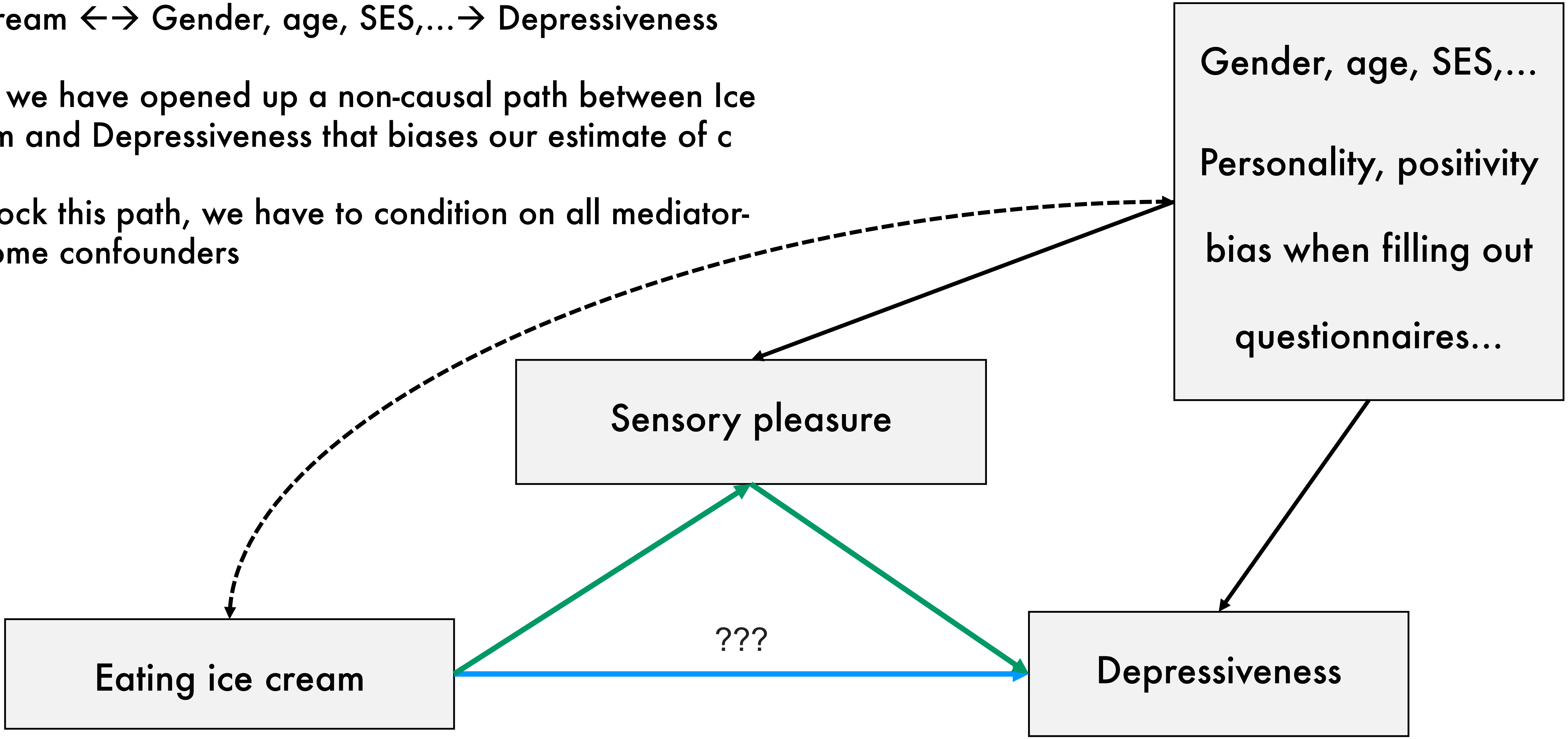

Ice Cream -> Sensory pleasure -> Depressiveness; block the indirect effect by conditioning on sensory pleasure

Ice Cream -> Sensory pleasure <- Positivity bias etc. -> Depressiveness

Ice cream ←→ Gender, age, SES,...→ Depressiveness

Now we have opened up a non-causal path between Ice cream and Depressiveness that biases our estimate of c

To block this path, we have to condition on all mediator-outcome confounders




Ice Cream -> Sensory pleasure -> Depressiveness; block the indirect effect by conditioning on sensory pleasure

Ice Cream -> Sensory pleasure <- Positivity bias etc. -> Depressiveness

# Identification assumptions of mediation analysis

- In a nutshell, to identify both the indirect and the direct effect, if X has been randomized, we still have to
  - *know* all mediator-outcome confounders
  - have measured them appropriately and
  - have statistically adjusted for them appropriately (e.g., by including them as a covariate in our regression analyses, by including them in our SEM)
- If the cause of interest (X) was not randomized, the same applies to any confounders between X and Y
  - And X and M (these are usually a subset of the confounders between X and Y)

# Identification assumptions of mediation analysis

- These assumptions are usually not realistic in many applications
  - For example, in psychology, both mediator and outcome are usually psychological variables which are potentially confounded by a *myriad* of other factors

# Identification assumptions of mediation analysis

- The assumptions about no unobserved confounding apply to both the traditional approach and more modern „causal" appraoches
- Additional assumptions in the traditional approach (indirect effect = a*b)
  - Effect of M on Y does not depend on X (no treatment-mediator interaction; MacKinnon et al., 2020)
  - The effects X → M and M → Y
    - Are the same for everyone or alternatively
    - May vary, but do so independently (otherwise, the multiplicative logic does not work out)

These may also often be implausible, in which case a more modern „causal" approach may make sense

# „Causal" mediation analysis

» *Causal* because they are explicitly (and actively) being developed in the causal inference literature (e.g., Imai et al., 2010)
  » The analysis goals are just as causal as the analysis goals of traditional mediation analysis
  » All problems concerning potentially unobserved confounding still very much apply

# „Causal“ mediation analysis: Three steps

- More focus on the precise definition of the effects of interest
- Usually more awareness of the necessary identifiation assumptions
  - (lack of unobserved confounding is still usually assumed)
- More flexible and varied estimation approaches
  - Allowing e.g. for non-linear effects, treatment-mediator interactions etc.
  - Implemented in various *R* packages

# Read more

- On causal mediation analysis:
  - Nguyen, Schmid, & Stuart (2020). Clarifying Causal Mediation Analysis for the Applied Researcher: Defining Effects Based on What We Want to Learn (Link)
- Some pragmatic advice from my perspective as an editor:
  - Blog post: “That’s a very nice mediation analysis you have there. It would be a shame if something happened to it” (Link)

# Causal inference with (observational) longitudinal data

„Our study was only observational, and thus no causal conclusions can be drawn.
Future **longitudinal** studies are needed to determine whether...“

# (Observational) longitudinal data

» Are not necessary for causal inference

» Are not sufficient for causal inference

» Can be very helpful for causal inference

[Rohrer & Murayama (2023)]

# Longitudinal data not necessary

- Common reasoning in psych: „Causal processes work within persons, thus within-person data are necessary to get at them"
- Causal effects are indeed defined on the individual level
- individual causal effect of reading an article about causal inference on your well-being (WB):
  - $WB^{\text{World in which you read article}}$ minus $WB^{\text{World in which you did not read article}}$

| | |
|---|---|
| $E[Y^1 - Y^0]$ | Expected value of the individual-level causal effects |
| $= E[Y^1] - E[Y^0]$ | Linearity of the expectation |
| $= E[Y^1 \mid X = 1] - E[Y^0 \mid X = 0]$ | To randomly sample from the distributions of the two potential outcomes, randomly assign people to two groups |
| $= E[Y \mid X = 1] - E[Y \mid X = 0]$ | In the two groups, after treatment, the respective potential outcomes are realized |

You can find an infographic for this on my website under "Resources"

# Longitudinal data not necessary

» Randomized between-subjects experiments can identify the average of the individual-level (within-subject!) causal effects

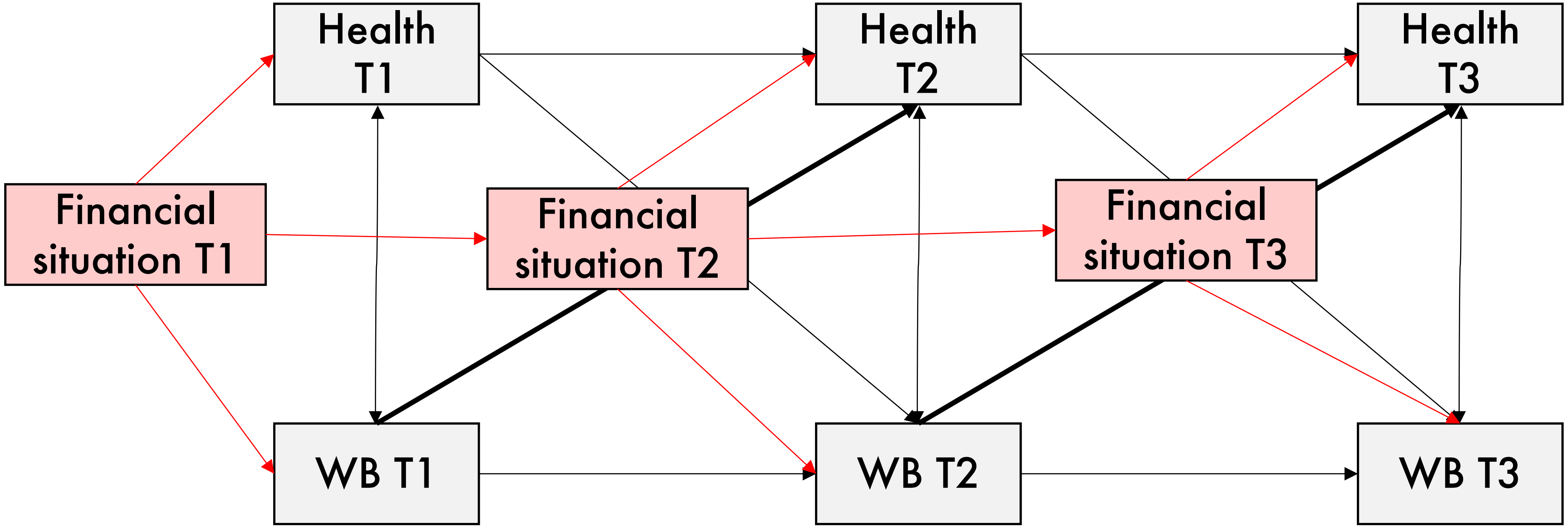
# Longitudinal data

» Are not necessary for causal inference

» Are not sufficient for causal inference

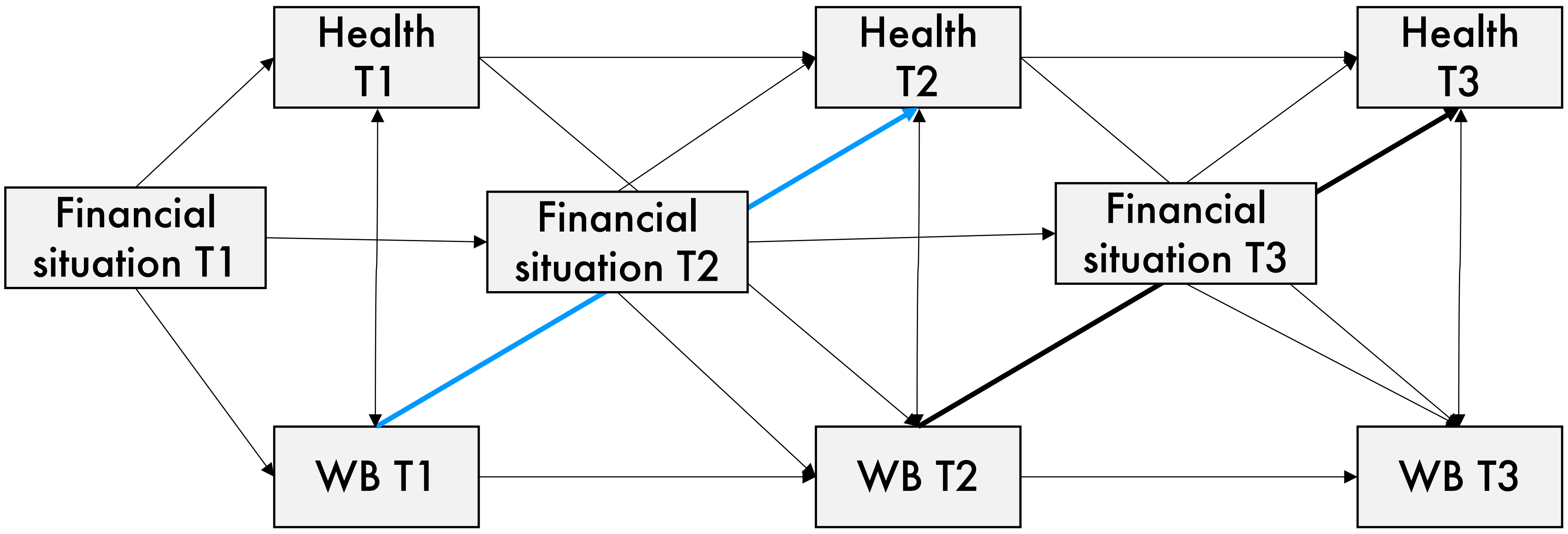
» Can be very helpful for causal inference

# Longitudinal data are not sufficient




Rohrer & Lucas (preprint)

# Longitudinal data are not sufficient

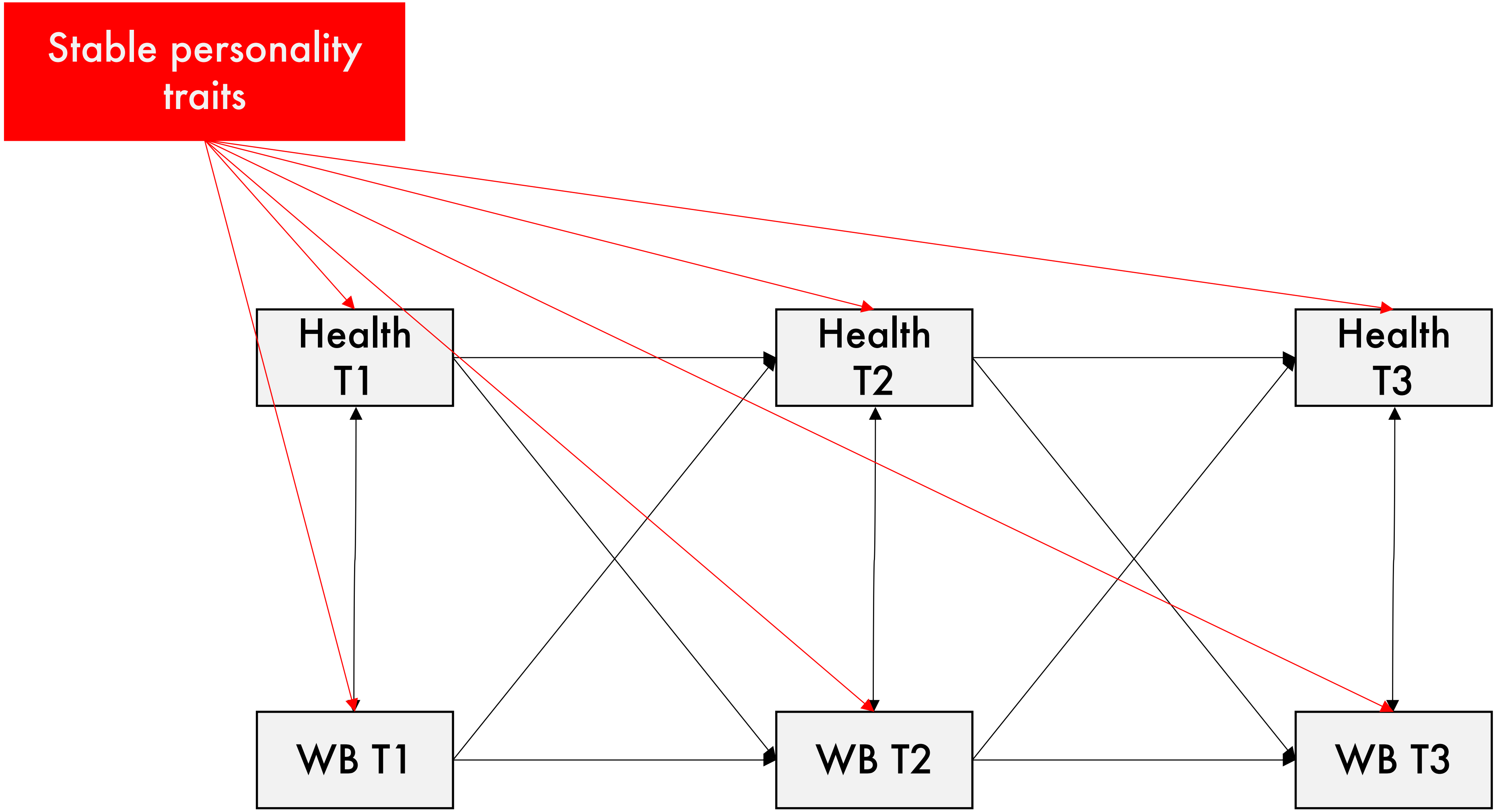



(Potentially unobserved) confounders can still open up non-causal backdoor paths between variables in longitudinal data

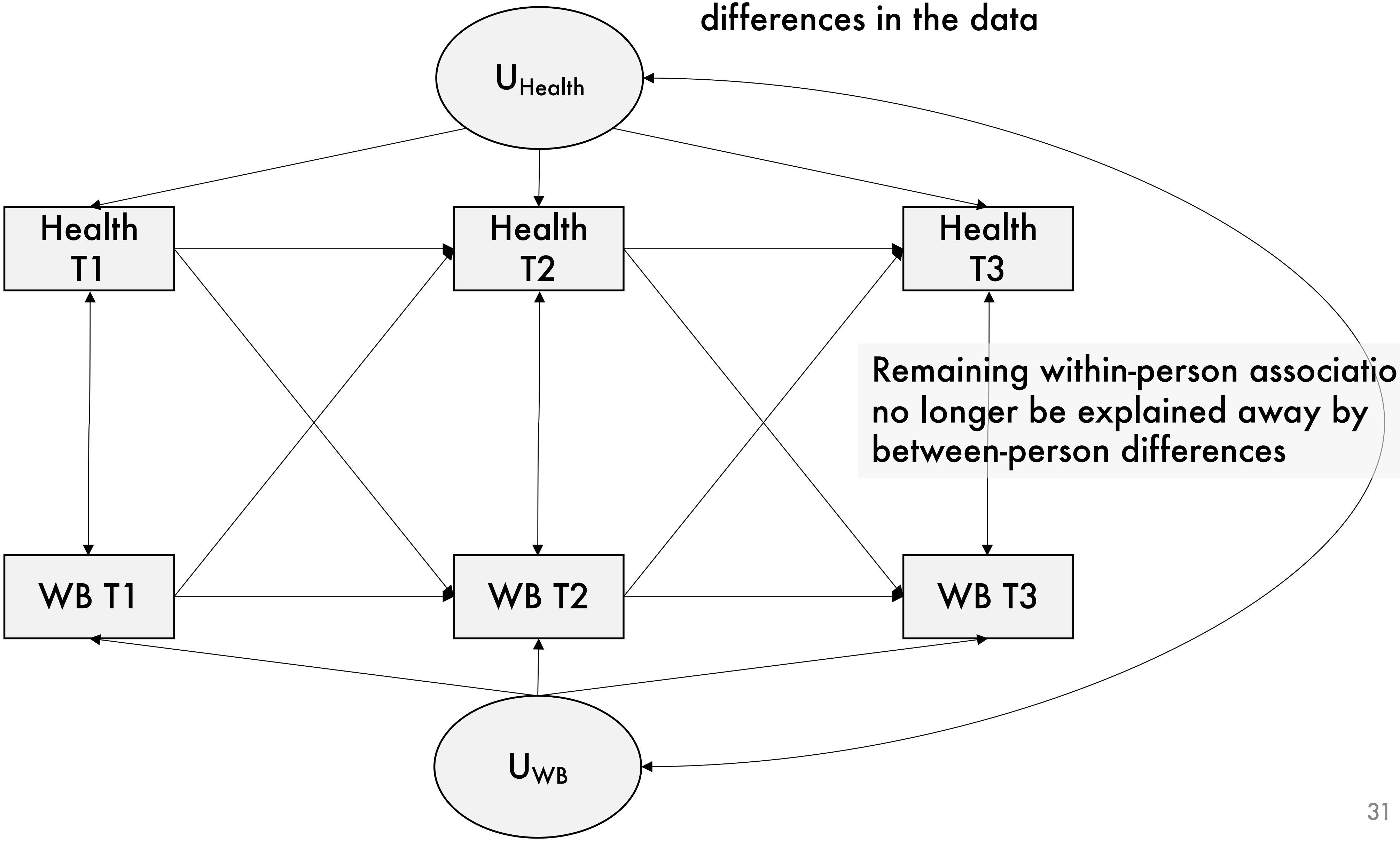
Rohrer & Lucas ([preprint](preprint))

# Longitudinal data

- » Are not necessary for causal inference
- » Are not sufficient for causal inference
- » Can be very helpful for causal inference


Stable personality traits
Health T1
Health T2
Health T3
WB T1
WB T2
WB T3

Intercept removes between-person differences in the data




Remaining within-person associations can no longer be explained away by between-person differences

# Longitudinal data can be very helpful

- Longitudinal data (analyzed properly) allow us to control for between-person differences
  - This means we no longer have to worry about so-called time-invariant confounders (variables that are stable within persons over the course of the study, e.g., gender, personality, childhood experiences…)
- But we still have to worry about any remaining *time-varying* confounders
  - e.g., stuff that happens in people's lives, development

# Read more

- The three points I just made
  - Rohrer & Murayama (2023). These are not the Effects you are Looking for: Causality and the Within-/Between-Persons Distinction in Longitudinal Data Analysis. (Link)
- Some nice hands-on device for the specific situation in which the outcome has been measured multiple times
  - VanderWeele, Mathur & Chen (2020). Outcome-Wide Longitudinal Designs for Causal Inference: A New Template for Empirical Studies. (Link)

# The Age-Period-Cohort Identification Problem

# What's an age/period/cohort effect?

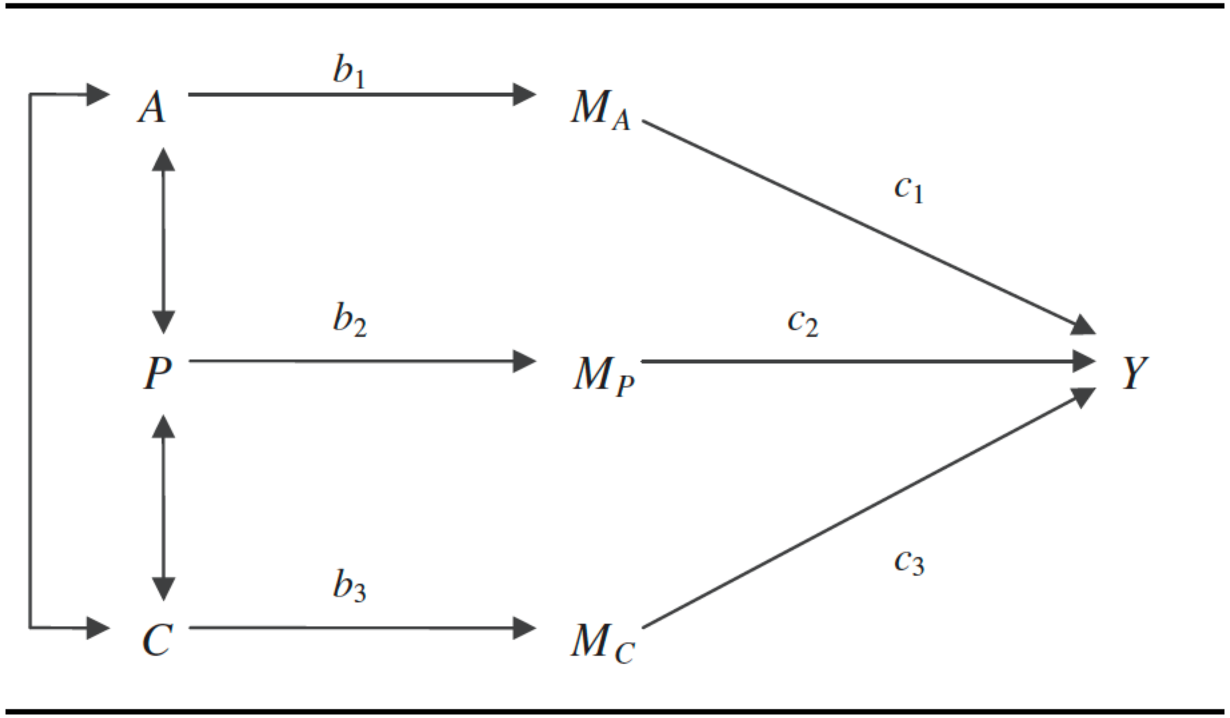
- A (causal) effect: contrast between two potential outcomes
  - the effect of time spent preparing this talk on its quality
    - Quality of talk$^{\text{Invested a whole day}}$ – Quality of talk$^{\text{Pasted together old slides in 30 minutes}}$, all else being equal
- But how does that work for age/period/cohort?
  - Life satisfaction$^{\text{At age 35}}$ – Life satisfaction$^{\text{At age 55}}$, *all else being equal?*
  - Life satisfaction$^{\text{Year 2026}}$ – Life satisfaction$^{\text{Year 2006}}$, *all else being equal?*
  - Life satisfaction$^{\text{1990 cohort}}$ – Life satisfaction$^{\text{1970 cohort}}$, *all else being equal?*

$$E(Y(a,p,c)) = \eta^* + \alpha^* \cdot a + \beta^* \cdot p + \theta^* \cdot c$$

$$E(Y\boxed{(c_a, c_p, c_c)}) = \eta' + \alpha' \cdot c_a + \beta' \cdot c_p + \theta' \cdot c_c,$$

The *consequences* of age, period, and cohort

**Figure 4**
**Hypothetical Age–Period–Cohort Model With Intervening Mechanisms**

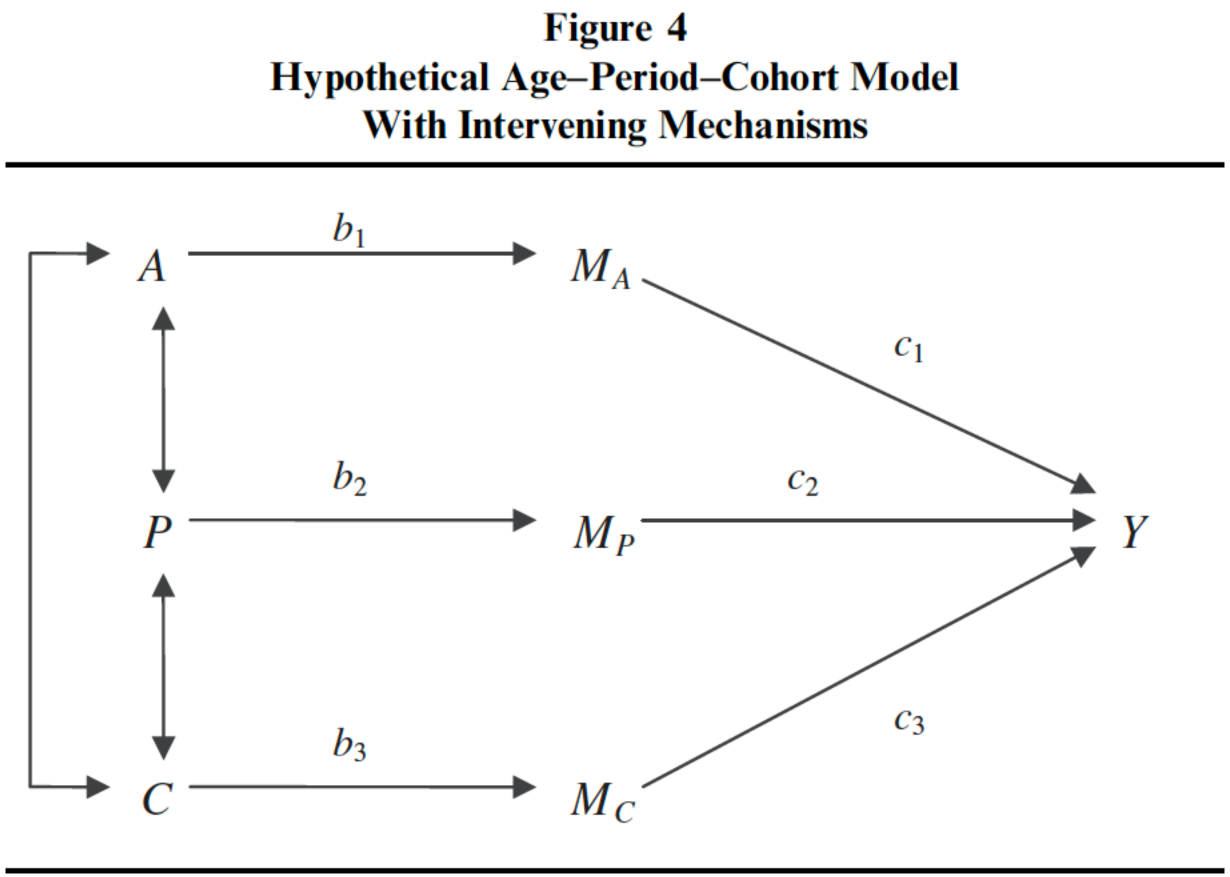



Bijlsma et al. (2017). An assessment and extension of the mechanism-based approach to the identification of age-period-cohort models. Link
Winship & Harding (2008). A mechanism-based approach to the identification of age-period-cohort models. Link

1. We cannot intervene on age, period, or cohort, so we need to rely on observational data.

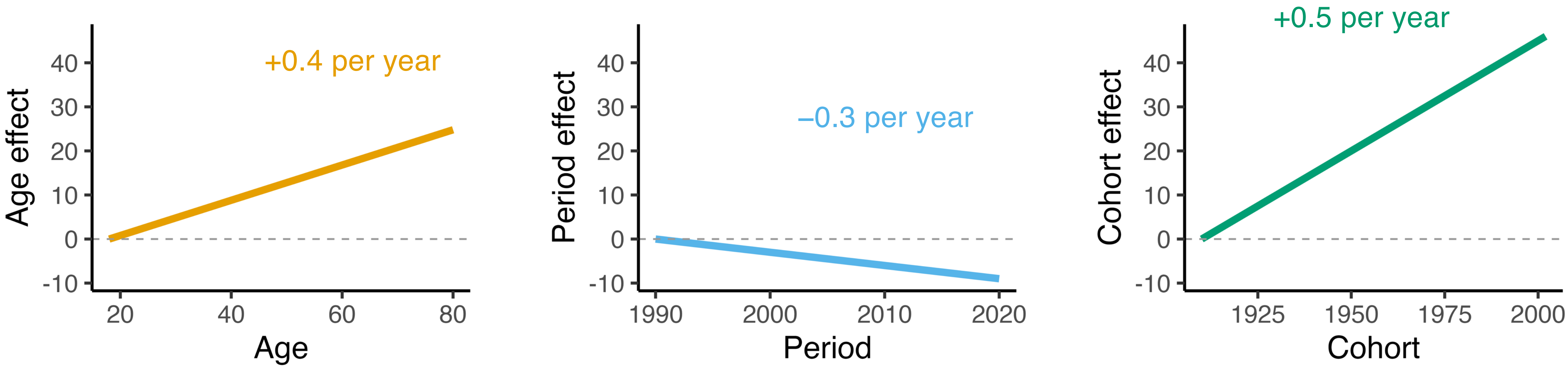
2. The variables of interest are in a deterministic relationship:

Age = Period − Cohort

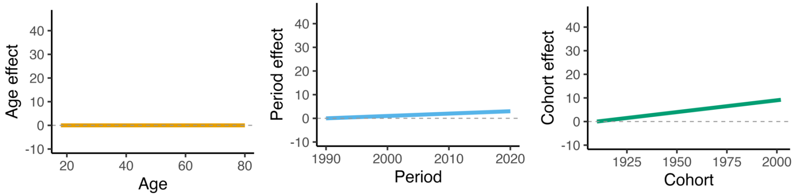



Bijlsma et al. (2017). An assessment and extension of the mechanism-based approach to the identification of age-period-cohort models. Link
Winship & Harding (2008). A mechanism-based approach to the identification of age-period-cohort models. Link

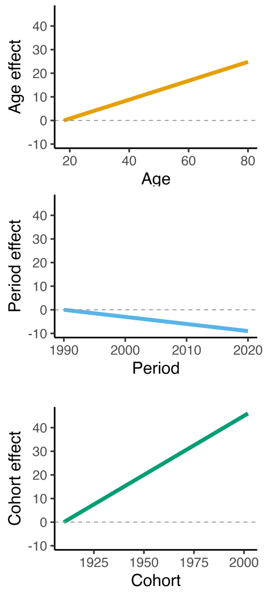
# A simulated example

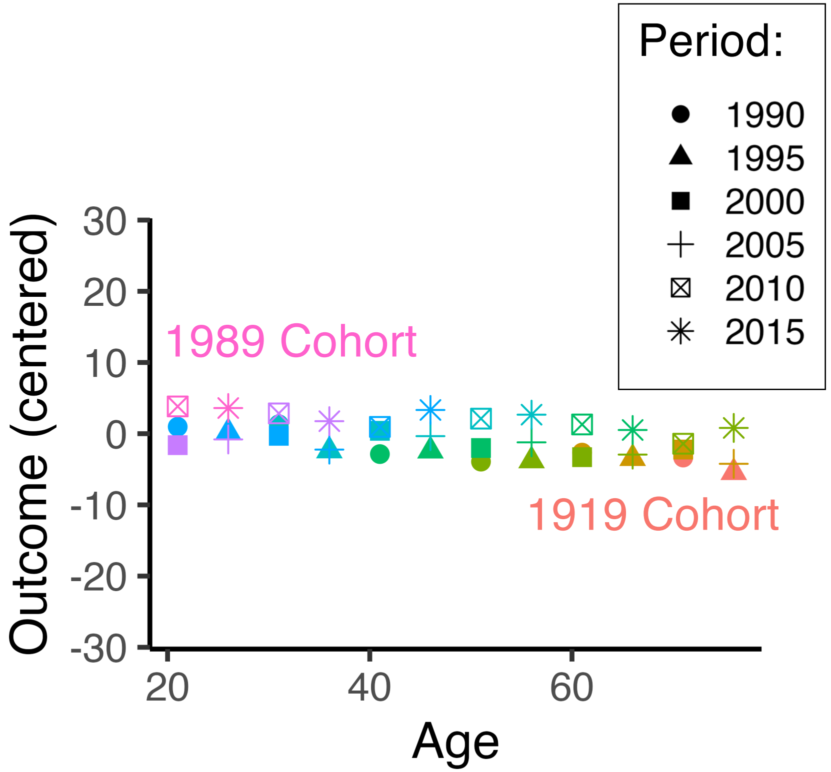


Age effect of 0
Period effect of 0.1
Cohort effect of 0.1

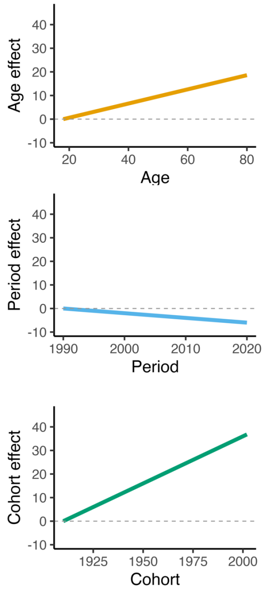



But the following effects are equally compatible with the observed mean pattern:



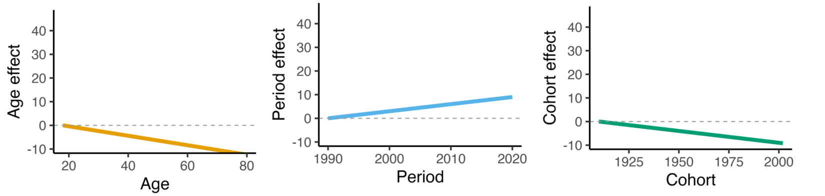




Age effect of +0.3
Period effect of −0.2
Cohort effect of +0.4




Effects that were used to generate the mean pattern:

Age effect of +0.4
Period effect of −0.3
Cohort effect of +0.5




Age effect of −0.2
Period effect of 0.3
Cohort effect of −0.1

Age effect of 0
Period effect of 0.1
Cohort effect of 0.1

**Age-Period-Cohort Identification Problem:**
The observed pattern of means fits an infinite number of combinations of age, period, and cohort effects equally well.

So, based on the data alone, it is impossible to identify which combination of effects generated the observed data.

Identification problem means that it is *not* „just" a problem of statistical estimation

→ It applies regardless of how much data are available to us
→ It even applies if we have the full population available

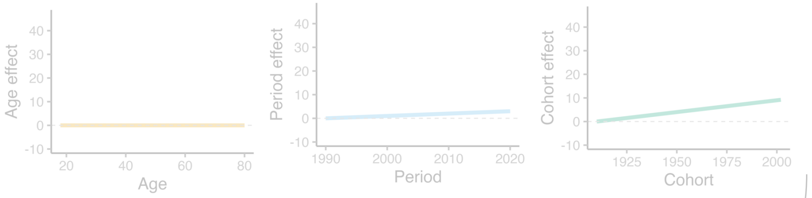
The problem applies regardless of the design with which we collected the data:
Cross-sectional design, repeated cross-sections, longitudinal design with a single cohort, longitudinal design with multiple cohorts („accelerated longitudinal design")...

The problem applies regardless of whether we are interested „only" in age (or cohort or period) effects, or two of them, or all three of them.

Effects that were used to generate the mean pattern:

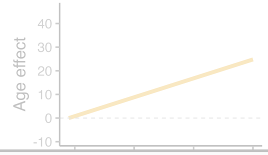
Age effect of +0.4
Period effect of −0.3
Cohort effect of +0.5

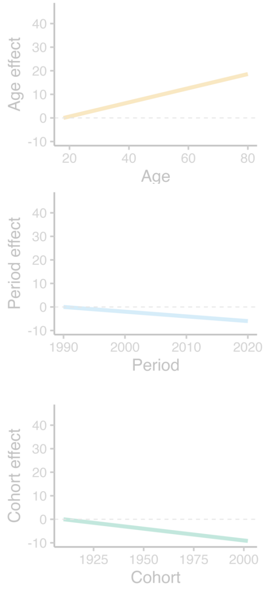
Age effect of 0
Period effect of 0.1
Cohort effect of 0.1




But the following effects are equally compatible with the observed mean pattern:

Age effect of +0.3
Period effect of −0.2
Cohort effect of +0.4




Period:
1990
1995
2000

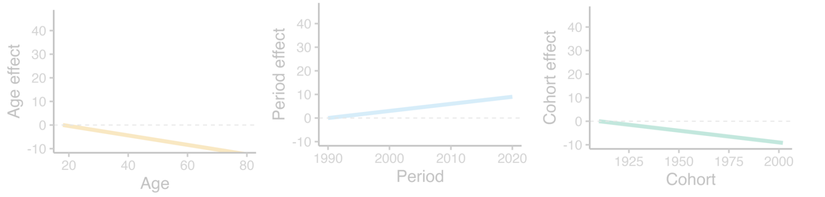
**Age-Period-Cohort Identification Problem:**
The observed pattern of means fi... number of combinations of age, p... cohort effects equally well.

So, based on the data alone, it is... identify which combination of effe... generated the observed data.

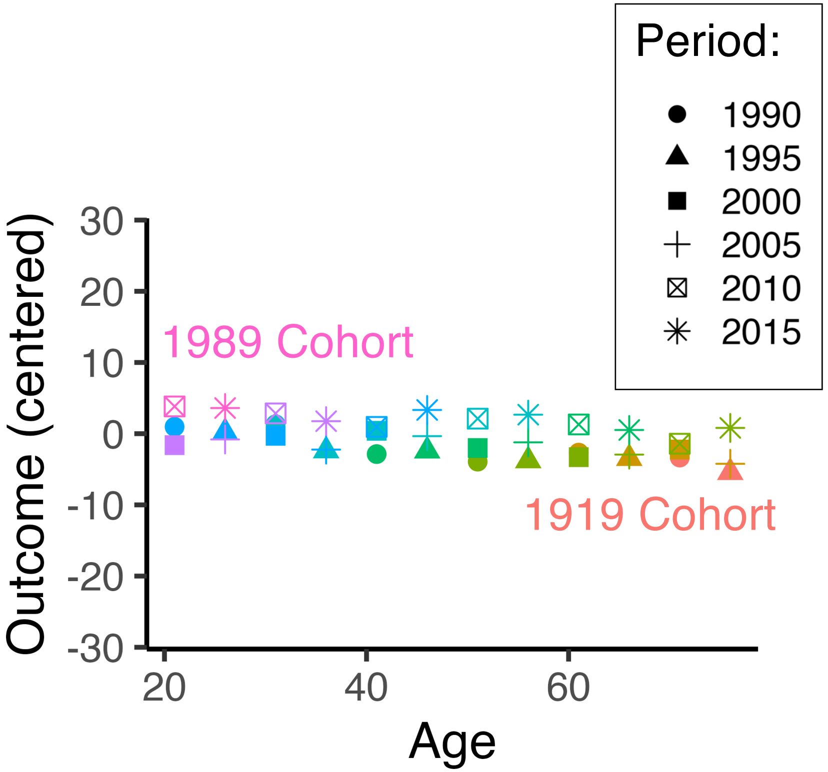
60
Age

Effects that were used to generate the mean pattern:

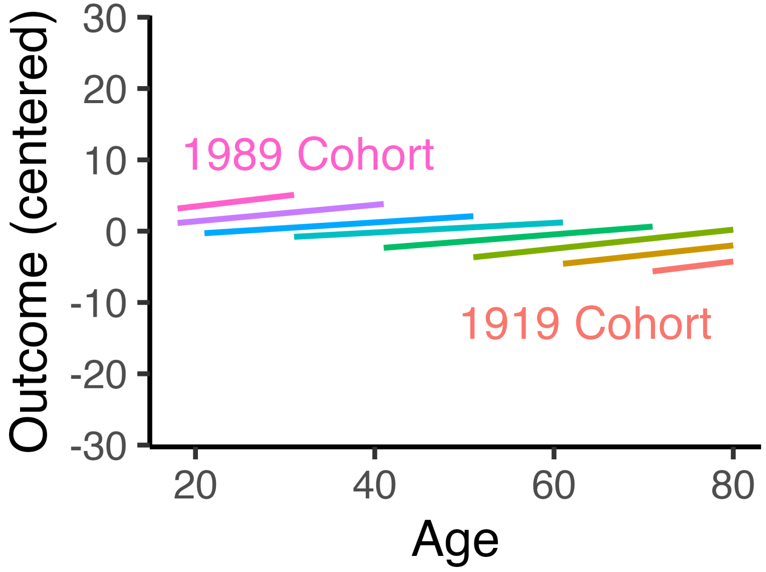
Age effect of +0.4
Period effect of −0.3
Cohort effect of +0.5




Age effect of −0.2
Period effect of 0.3
Cohort effect of −0.1




Connect cohortwise

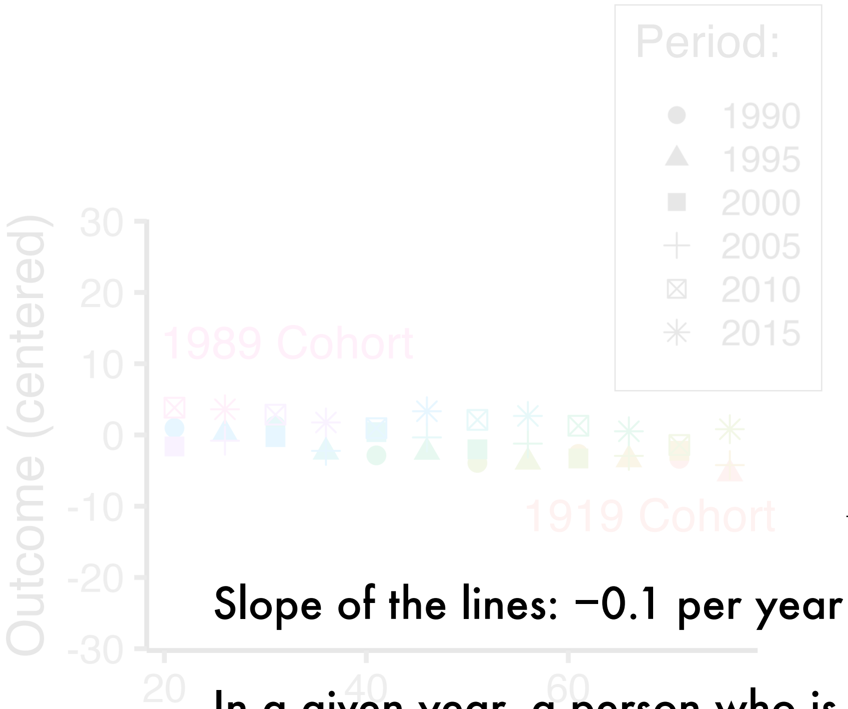



Age effect + period effect = 0.1

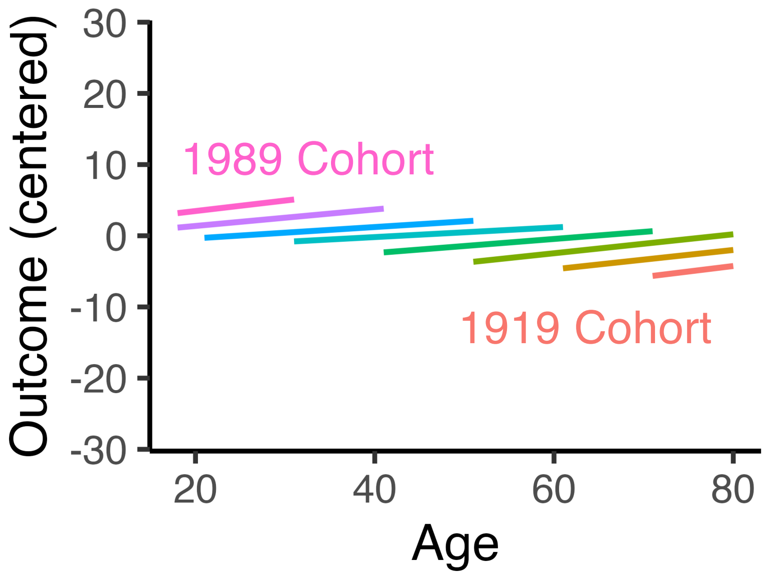
Slope of the lines: 0.1 per year

When a cohort ages by one year, both age and period increase by one.

So, the increase over one year is the sum of the age effect over one year and the period effect over one year.

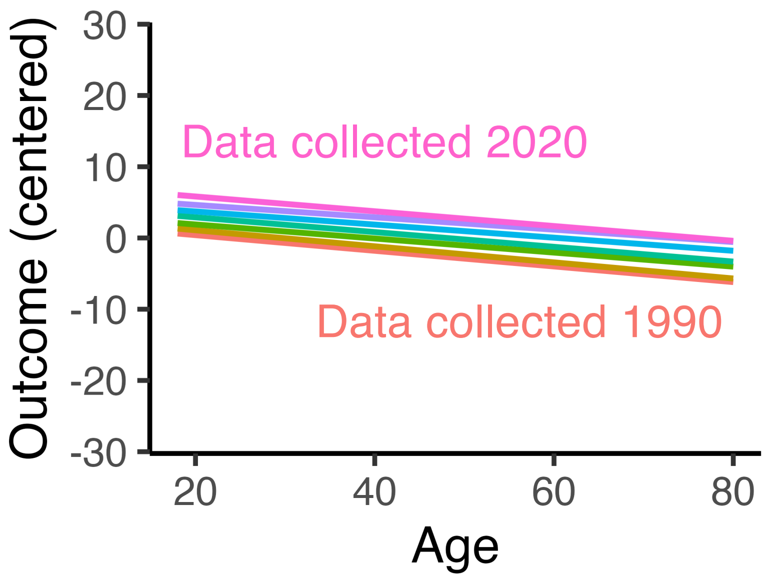


Connect cohortwise




Age effect + period effect = 0.1

Connect periodwise



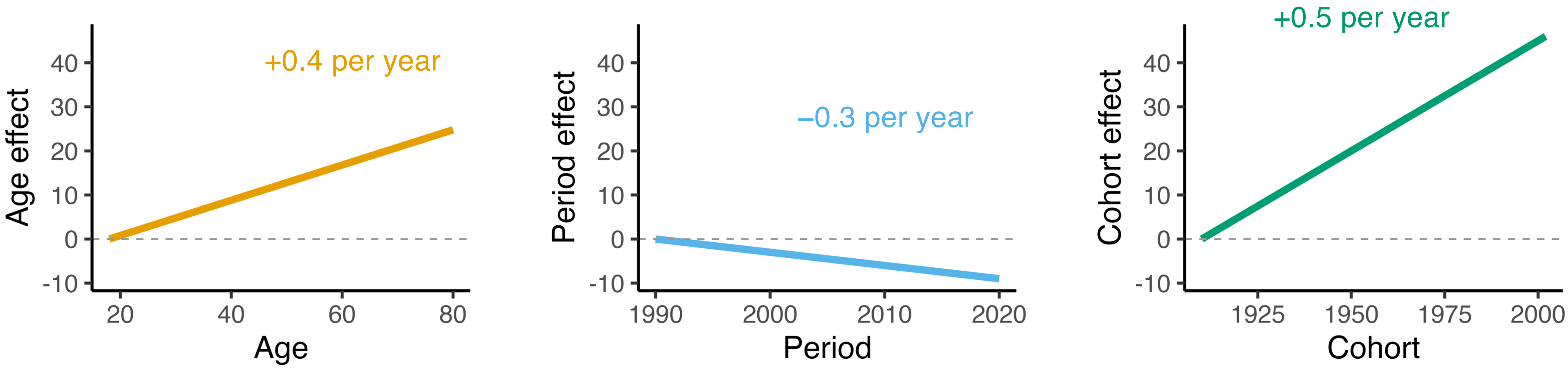

Age effect − cohort effect = −0.1

Slope of the lines: −0.1 per year

In a given year, a person who is one year older (age +1) was born one cohort earlier (cohort −1)

So, the increase over one year is the age effect over one year minus the cohort effect over one year

# Quick sanity check

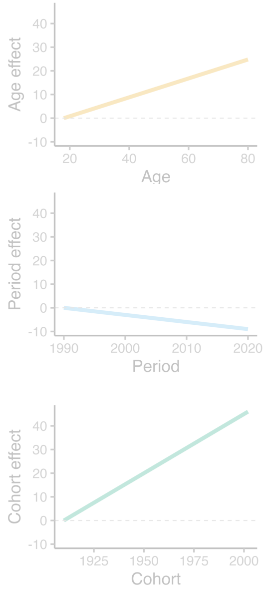



Age effect + period effect
= 0.1

Age effect − cohort effect
= −0.1

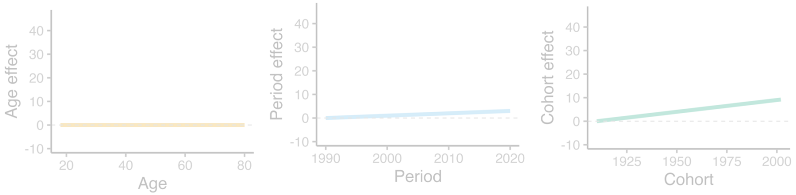
Effects that were used to generate the mean pattern:

Age effect of +0.4
Period effect of −0.3
Cohort effect of +0.5

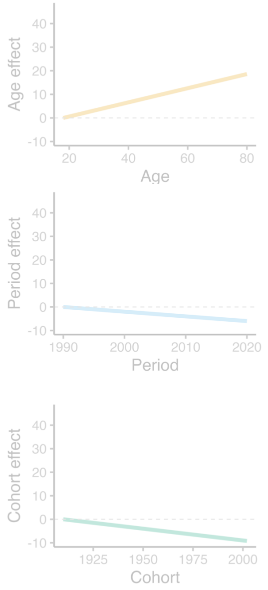



But the following effects are equally compatible with the observed mean pattern:

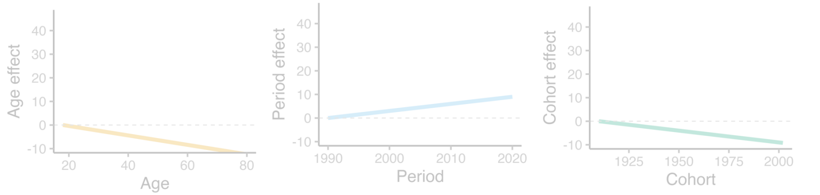
Age effect of 0
Period effect of 0.1
Cohort effect of 0.1




Age effect of +0.3
Period effect of −0.2
Cohort effect of +0.4




Age effect of −0.2
Period effect of 0.3
Cohort effect of −0.1




**This holds for all patterns that are equally compatible with the data:**

Age effect + period effect = 0.1

Age effect − cohort effect = −0.1

# Solutions to the age-period-cohort problem

(Fosse & Winship, 2018, 2019)

» We *cannot* learn the age, period, and cohort effects from the data (because an infinite number of combinations of them are equally compatible with the data)

» However, we can learn combinations of them

  » For example

    » Age effect + period effect = 0.1

    » Age effect − cohort effect = −0.1

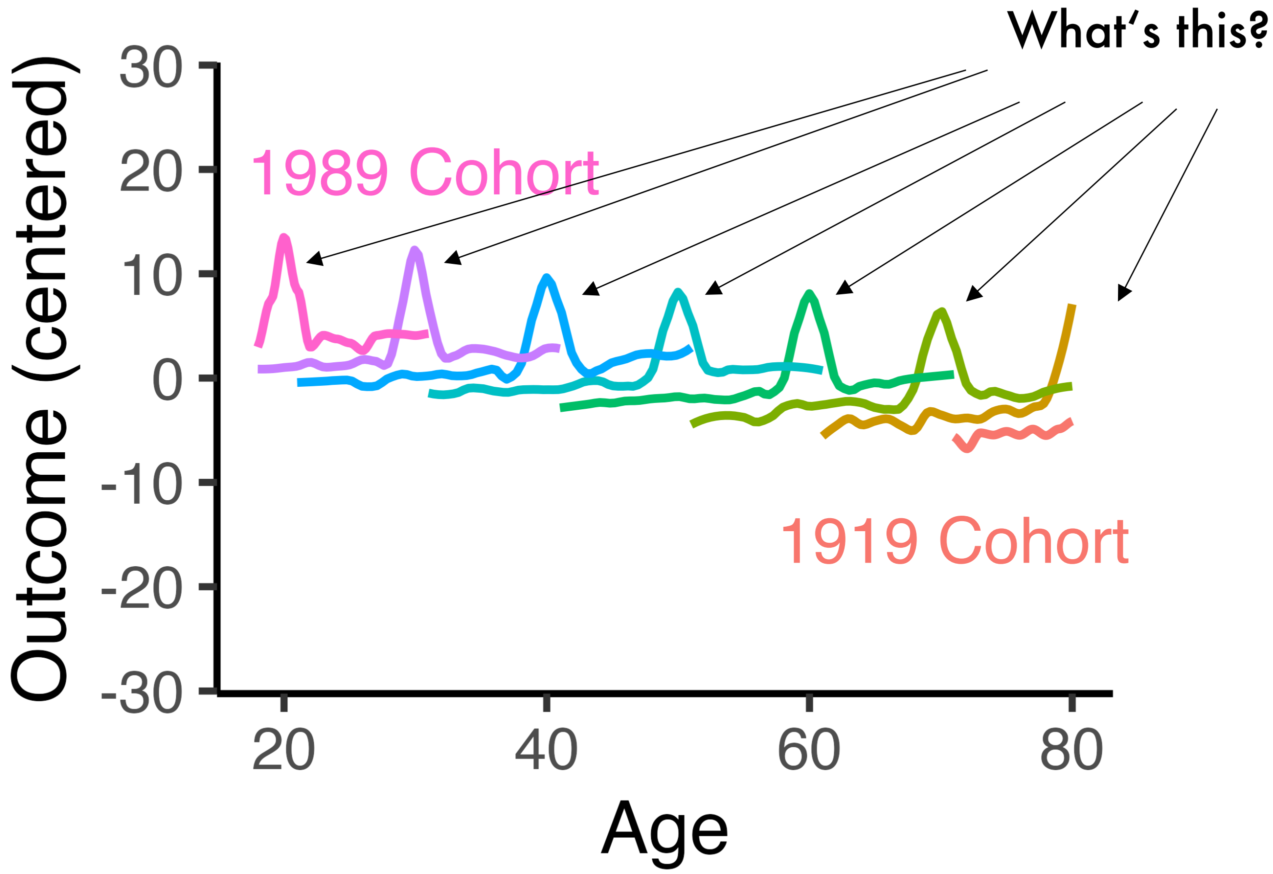
» This limits the space of solutions compatible with the data

# Solutions to the age-period-cohort problem

(Fosse & Winship, 2018, 2019)

» Additionally, under the assumption that age, period, and cohort do not interact, we can identify *nonlinearities* in their effects

# Identifying nonlinearities




If it were a non-linear age effect, it would occur at the same age in each cohort.

If it were a non-linear cohort effect, it would shift the line of one cohort but not affect the other cohorts.

→ It must be a period non-linearity!

(Period-peak from 2008 to 2010)

# What the data *do* tell us

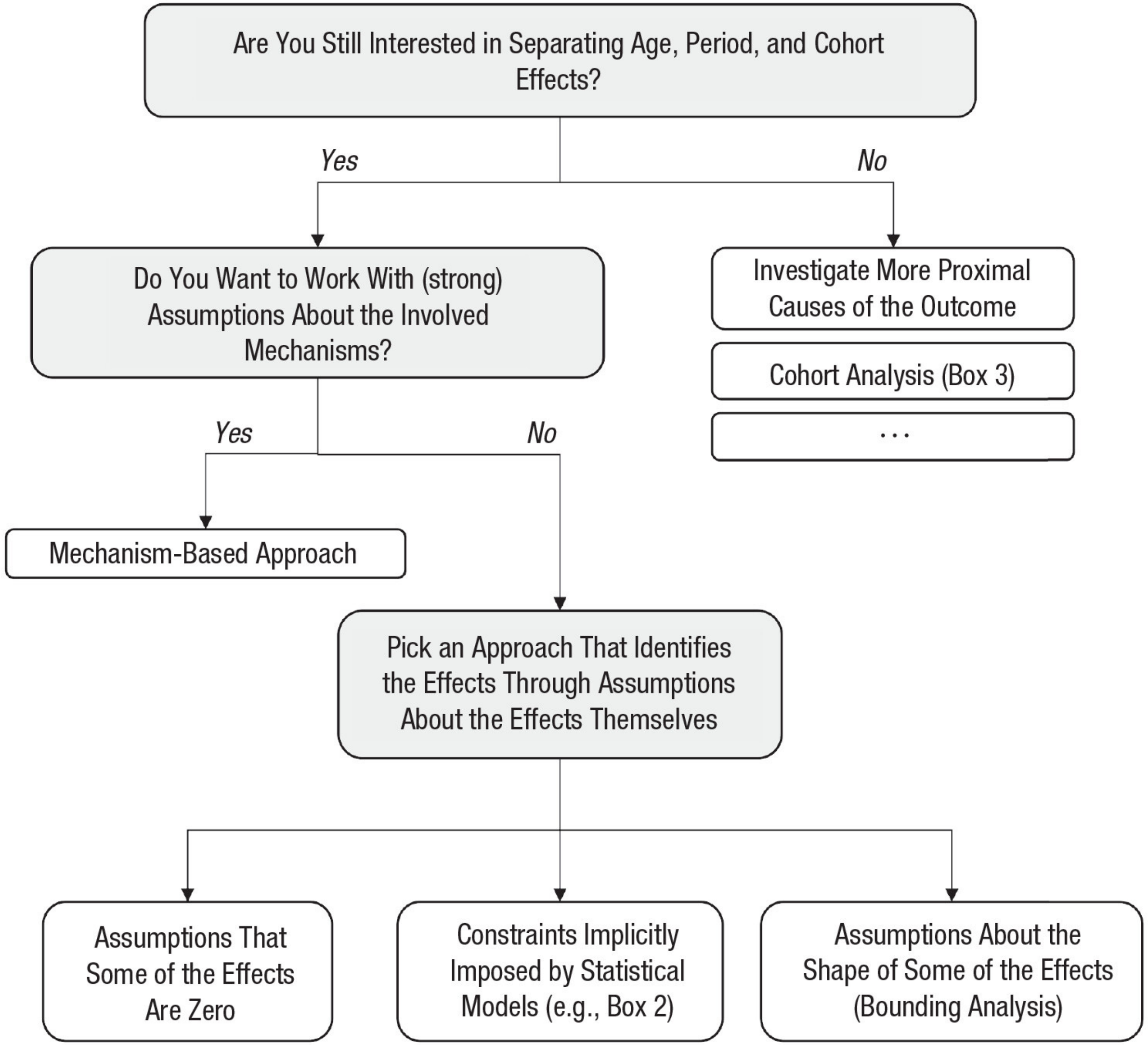
- *Combinations* of age, period, and cohort effects
- under the no-interactions assumptions: any *nonlinearities* in the effects of age, period, and cohort
  - but really *only* the nonlinearities (i.e., deviations from linear trends)

# What the data *don't* tell us

- Any conclusions beyond that – any conclusions about the effects of age, period, and cohort – rely on additional identification assumptions
  - e.g., assumption that some effects are zero (no period effects)
  - e.g., assumptions that effects are monotonous, don't exceed a certain magnitude
  - e.g., constraints imposed by the choice of the statistical model (random effects models, the intrinsic estimator)
    - these are usually obscure and *very* unlikely to be substantively plausible
    - don't believe *anybody* who claims to provide a general statistical solution to the problem
      - they are probably just hiding their very strange assumptions from you

# Read more

- Primer to "solutions" to the APC problem including a handy flowchart:
  - Rohrer (2025). Thinking Clearly About Age, Period, and Cohort Effects. (Link)
- Maybe you don't even need to solve the APC problem
  - Blog post: One approach to the age-period-cohort problem: Just don't. (Link)



# Thank you for your attention!

Julia M. Rohrer
Leipzig University
juliarohrer.com
www.the100.ci

Slides: Resources