PATHWAYS

PATHWAYS
biosocial influences on health

Current and new approaches to estimating causal pathways from observational data

Rhian Daniel and Bianca De Stavola

ESRC Research Methods Festival, 4th July 2012, 9.15am

Website http://pathways.lshtm.ac.uk
Email pathways@lshtm.ac.uk
Twitter @pathwaysNCRM

UNIVERSITY OF CAMBRIDGE
Department of Geography

NCRM National Centre for Research Methods

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

PATHWAYS

PATHWAYS biosocial influences on health

Old and new approaches to estimating causal pathways from observational data

Rhian Daniel and Bianca De Stavola

ESRC Research Methods Festival, 4th July 2012, 9.15am

Website http://pathways.lshtm.ac.uk
Email pathways@lshtm.ac.uk
Twitter @pathwaysNCRM

University of Cambridge Department of Geography

NCRM National Centre for Research Methods

London School of Hygiene & Tropical Medicine

Outline

Outline

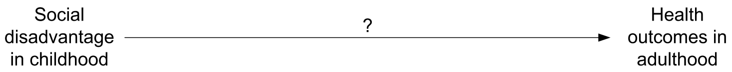
Why are we interested in pathways?

An example

- Being socially disadvantaged in childhood is associated with having poorer health outcomes in adulthood.

Why are we interested in pathways?

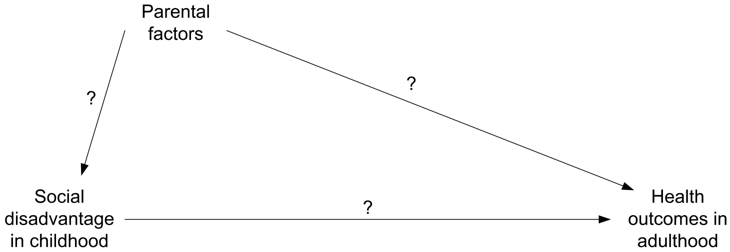
An example



- Natural first question: is this causal?

Why are we interested in pathways?

An example



- Or explained by other things? (Confounding).

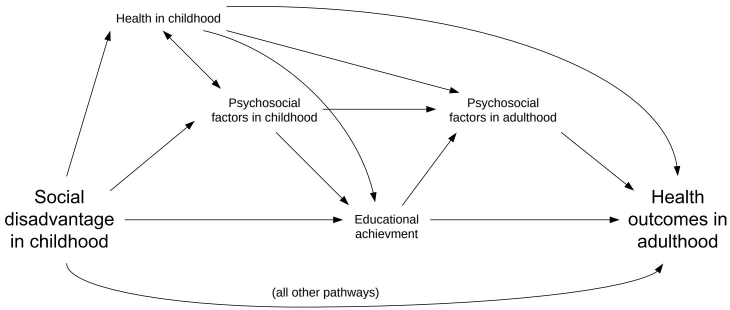
Why are we interested in pathways?

An example

- Suppose a causal effect can be established.

Why are we interested in pathways?

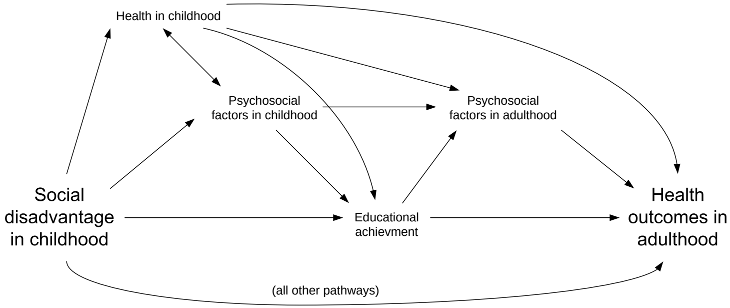
An example



- Natural next step: how does this causal effect act?
- How important are the different pathways?
- Where should interventions be targeted?

Why are we interested in pathways?

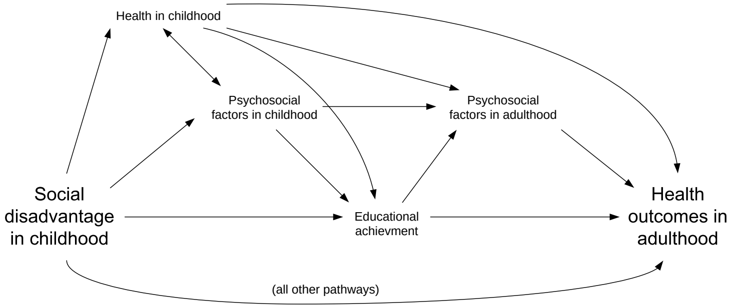
An example



- Natural next step: how does this causal effect act?
- How important are the different pathways?
- Where should interventions be targeted?

Why are we interested in pathways?

An example

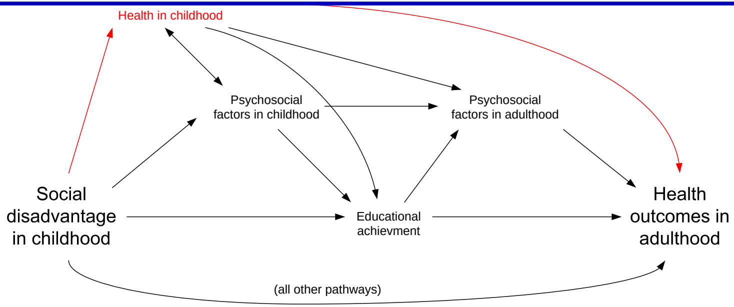


- Natural next step: how does this causal effect act?
- How important are the different pathways?
- Where should interventions be targeted?

Why are we interested in pathways?

An example

There are many possible pathways to explore. . .

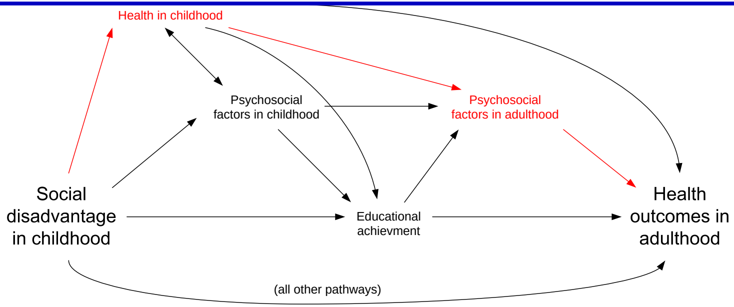


- Natural next step: how does this causal effect act?
- How important are the different pathways?
- Where should interventions be targeted?

Why are we interested in pathways?

An example

There are many possible pathways to explore. . .

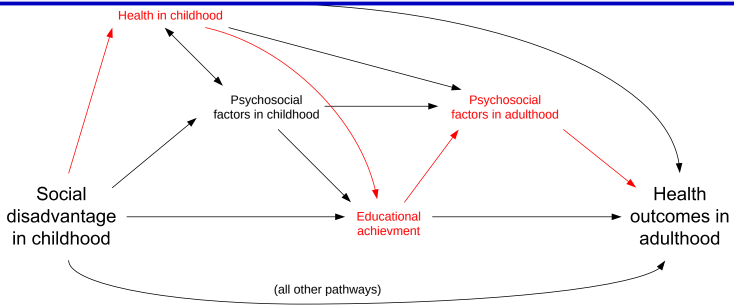


- Natural next step: how does this causal effect act?
- How important are the different pathways?
- Where should interventions be targeted?

Why are we interested in pathways?

An example

There are many possible pathways to explore. . .

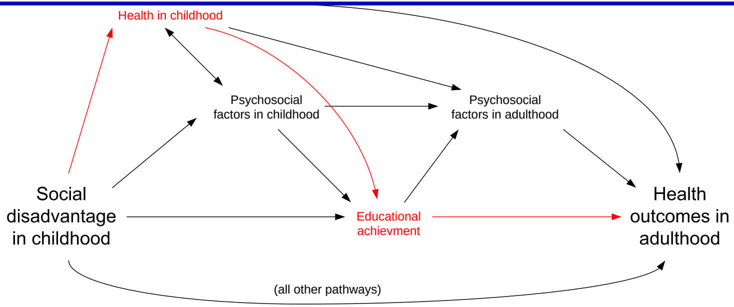


- Natural next step: how does this causal effect act?
- How important are the different pathways?
- Where should interventions be targeted?

Why are we interested in pathways?

An example

There are many possible pathways to explore. . .

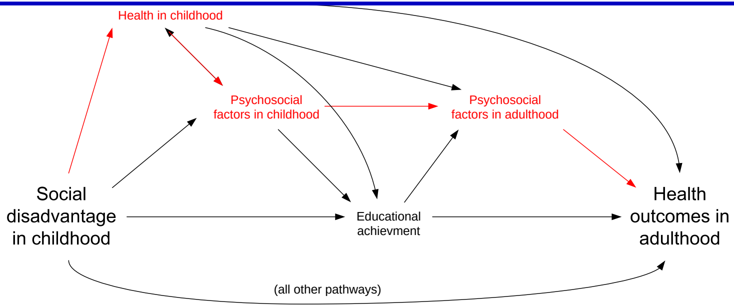


- Natural next step: how does this causal effect act?
- How important are the different pathways?
- Where should interventions be targeted?

Why are we interested in pathways?

An example

There are many possible pathways to explore. . .

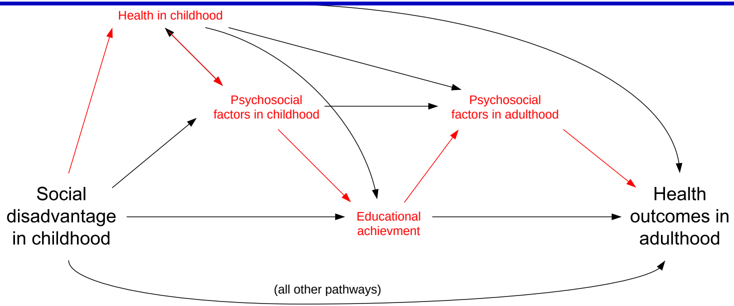


- Natural next step: how does this causal effect act?
- How important are the different pathways?
- Where should interventions be targeted?

Why are we interested in pathways?

An example

There are many possible pathways to explore. . .

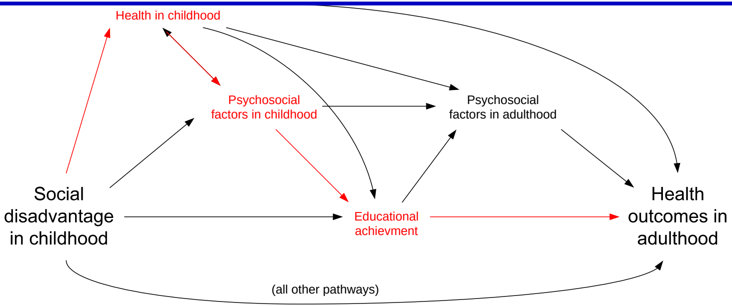


- Natural next step: how does this causal effect act?
- How important are the different pathways?
- Where should interventions be targeted?

Why are we interested in pathways?

An example

There are many possible pathways to explore. . .

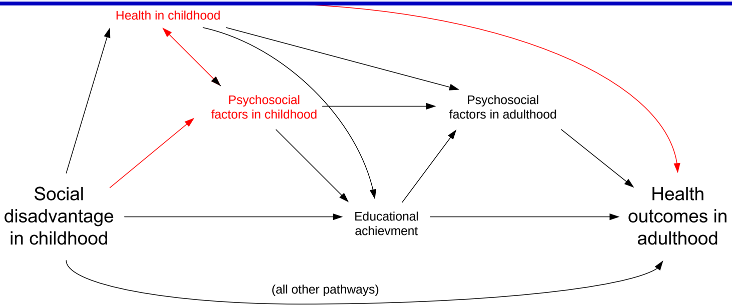


- Natural next step: how does this causal effect act?
- How important are the different pathways?
- Where should interventions be targeted?

Why are we interested in pathways?

An example

There are many possible pathways to explore. . .

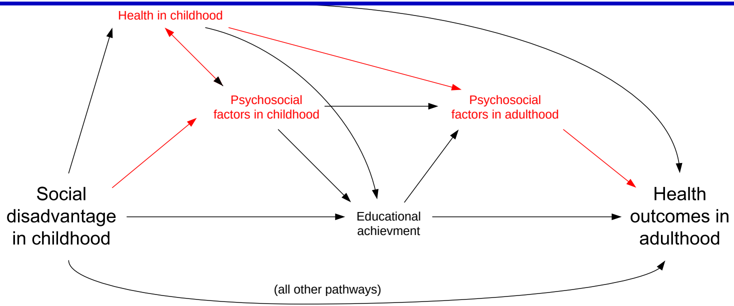


- Natural next step: how does this causal effect act?
- How important are the different pathways?
- Where should interventions be targeted?

Why are we interested in pathways?

An example

There are many possible pathways to explore. . .

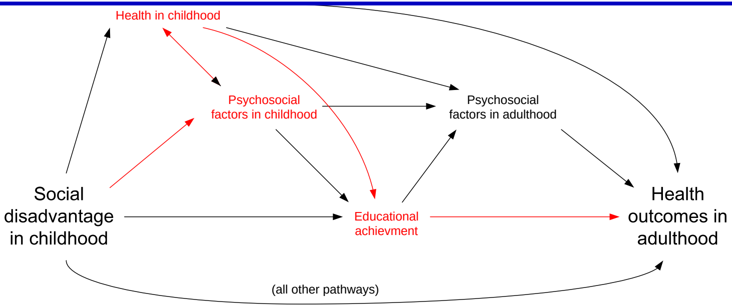


- Natural next step: how does this causal effect act?
- How important are the different pathways?
- Where should interventions be targeted?

Why are we interested in pathways?

An example

There are many possible pathways to explore. . .

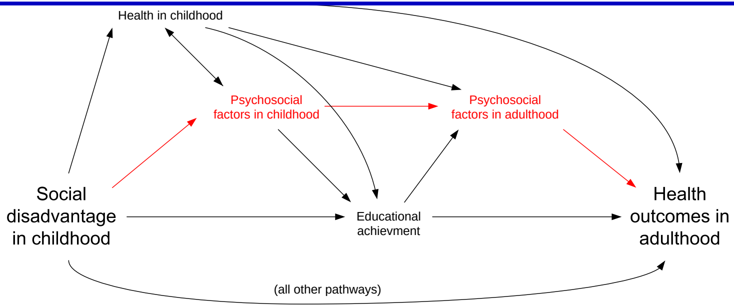


- Natural next step: how does this causal effect act?
- How important are the different pathways?
- Where should interventions be targeted?

Why are we interested in pathways?

An example

There are many possible pathways to explore. . .

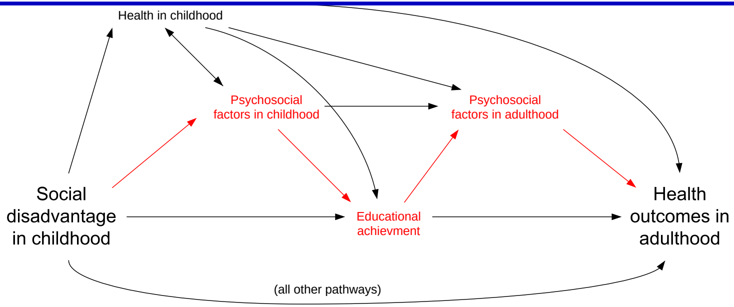


- Natural next step: how does this causal effect act?
- How important are the different pathways?
- Where should interventions be targeted?

Why are we interested in pathways?

An example

There are many possible pathways to explore. . .

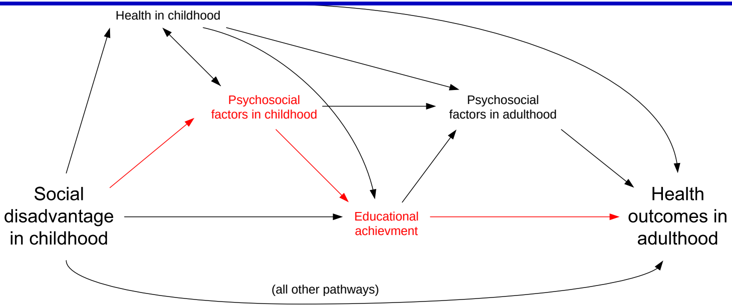


- Natural next step: how does this causal effect act?
- How important are the different pathways?
- Where should interventions be targeted?

Why are we interested in pathways?

An example

There are many possible pathways to explore. . .

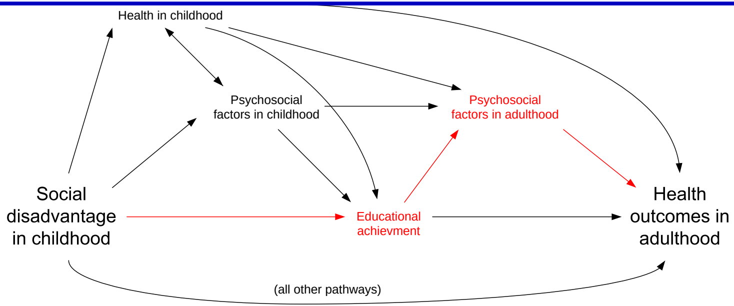


- Natural next step: how does this causal effect act?
- How important are the different pathways?
- Where should interventions be targeted?

Why are we interested in pathways?

An example

There are many possible pathways to explore. . .

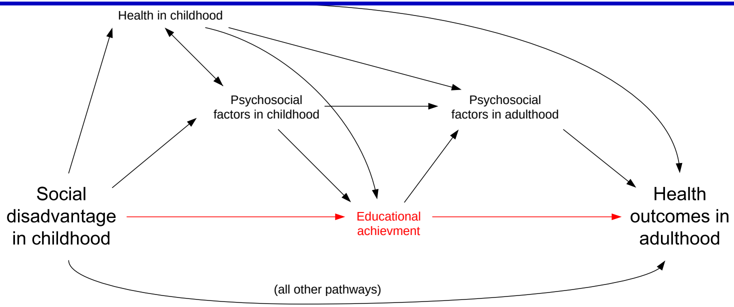


- Natural next step: how does this causal effect act?
- How important are the different pathways?
- Where should interventions be targeted?

Why are we interested in pathways?

An example

There are many possible pathways to explore. . .

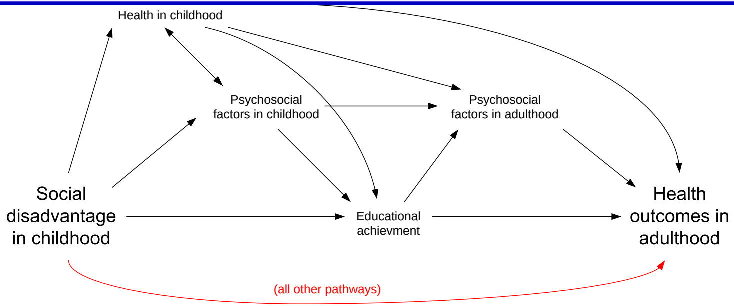


- Natural next step: how does this causal effect act?
- How important are the different pathways?
- Where should interventions be targeted?

Why are we interested in pathways?

An example

There are many possible pathways to explore. . .

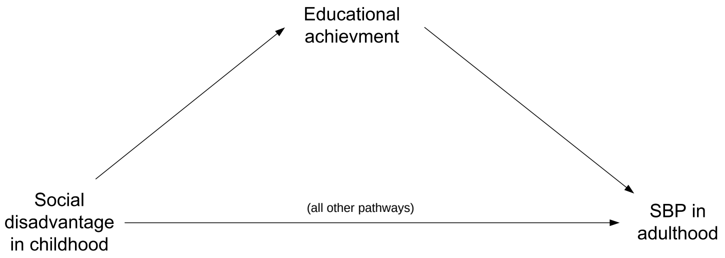


- Natural next step: how does this causal effect act?
- How important are the different pathways?
- Where should interventions be targeted?

Outline

A simpler question

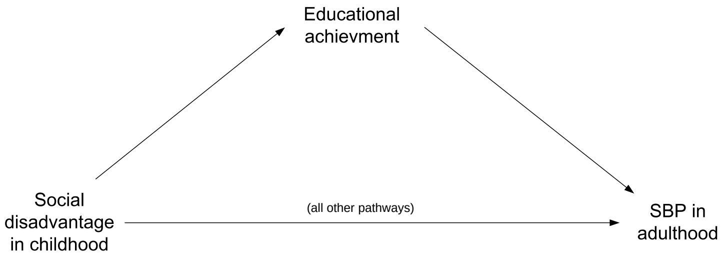
Two pathways



- Consider a much simpler (but still very challenging!) question.
- How much of the effect of social disadvantage in childhood on, say, systolic blood pressure in adulthood, is mediated by educational achievement?
- Only two pathways ('direct' and 'indirect').

A simpler question

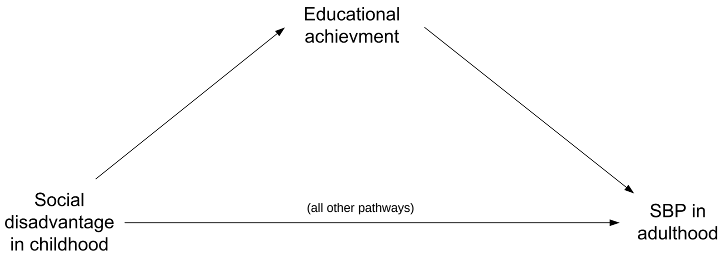
Two pathways



- Consider a much simpler (but still very challenging!) question.
- How much of the effect of social disadvantage in childhood on, say, systolic blood pressure in adulthood, is mediated by educational achievement?
- Only two pathways ('direct' and 'indirect').

A simpler question

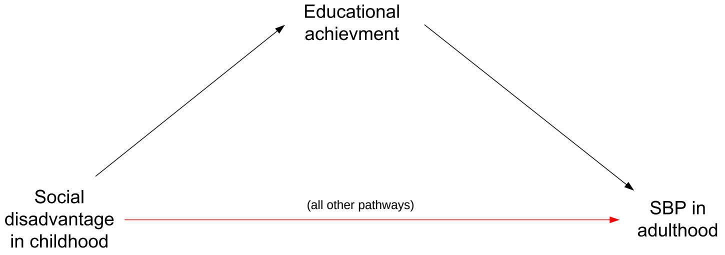
Two pathways



- Consider a much simpler (but still very challenging!) question.
- How much of the effect of social disadvantage in childhood on, say, systolic blood pressure in adulthood, is mediated by educational achievement?
- Only two pathways ('direct' and 'indirect').

A simpler question

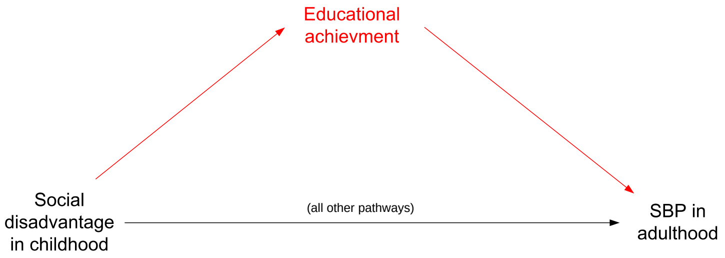
Two pathways



- Consider a much simpler (but still very challenging!) question.
- How much of the effect of social disadvantage in childhood on, say, systolic blood pressure in adulthood, is mediated by educational achievement?
- Only two pathways ('direct' and 'indirect').

A simpler question

Two pathways

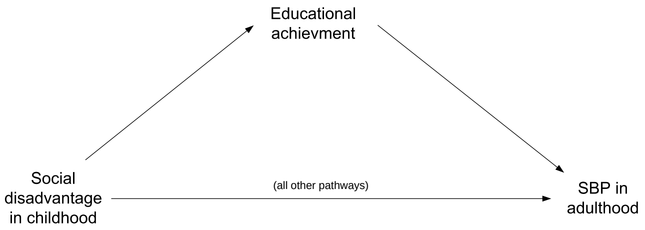


- Consider a much simpler (but still very challenging!) question.
- How much of the effect of social disadvantage in childhood on, say, systolic blood pressure in adulthood, is mediated by educational achievement?
- Only two pathways ('direct' and 'indirect').

Outline

The curruent/old approach to estimating pathways

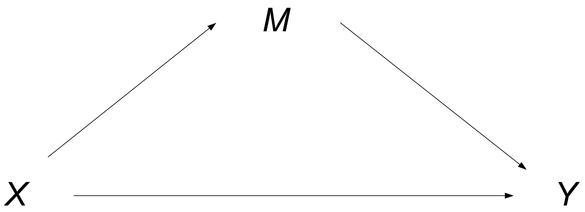
Combination of simple least squares regressions



- Suppose social disadvantage and educational achievement are each measured using a univariate continuous score.
- Write X for the exposure, M for the mediator and Y for the outcome.
- Let's explicitly include confounders C.

The curruent/old approach to estimating pathways

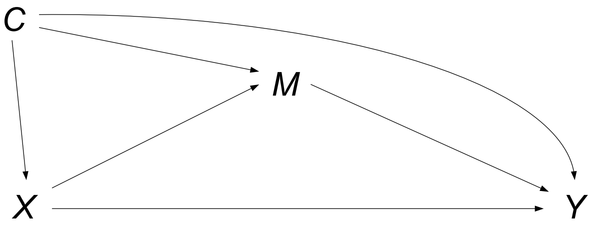
Combination of simple least squares regressions



- Suppose social disadvantage and educational achievement are each measured using a univariate continuous score.
- Write X for the exposure, M for the mediator and Y for the outcome.
- Let's explicitly include confounders C.

The curruent/old approach to estimating pathways

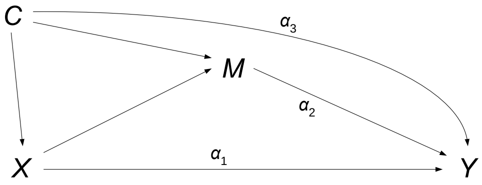
Combination of simple least squares regressions



- Suppose social disadvantage and educational achievement are each measured using a univariate continuous score.
- Write X for the exposure, M for the mediator and Y for the outcome.
- Let's explicitly include confounders C.

'Difference' method

Baron and Kenny, 1986

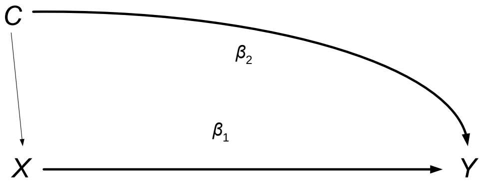


Consider two regression models:

$$E(Y|C,X,M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$
$$E(Y|C,X) = \beta_0 + \beta_1 X + \beta_2^T C$$

‘Difference’ method

Baron and Kenny, 1986

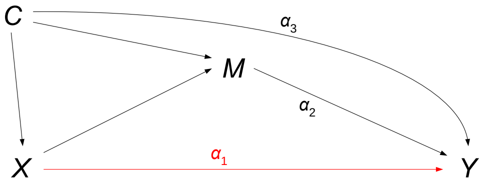


Consider two regression models:

$$E(Y\,|\,C,X,M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$
$$E(Y\,|\,C,X) = \beta_0 + \beta_1 X + \beta_2^T C$$

'Difference' method

Baron and Kenny, 1986



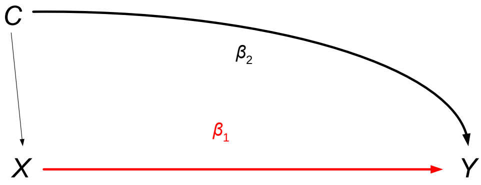
Consider two regression models:

$$E(Y|C,X,M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$
$$E(Y|C,X) = \beta_0 + \beta_1 X + \beta_2^T C$$

- α_1 is interpreted as the direct effect (not via M),
- β_1 is interpreted as the total effect,
- and thus $\beta_1 - \alpha_1$ is the indirect effect (via M).

'Difference' method

Baron and Kenny, 1986



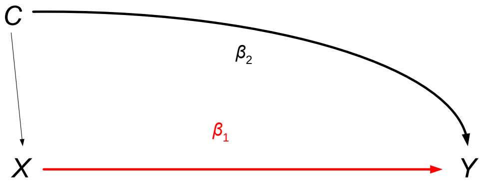
Consider two regression models:

$$E(Y \mid C, X, M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$
$$E(Y \mid C, X) = \beta_0 + \beta_1 X + \beta_2^T C$$

- α_1 is interpreted as the direct effect (not via M),
- β_1 is interpreted as the total effect,
- and thus $\beta_1 - \alpha_1$ is the indirect effect (via M).

'Difference' method

Baron and Kenny, 1986



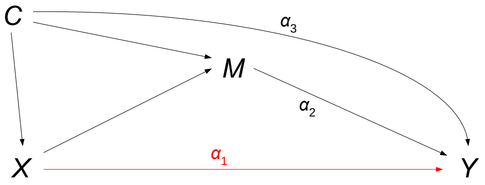
Consider two regression models:

$$E(Y|C,X,M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$
$$E(Y|C,X) = \beta_0 + \beta_1 X + \beta_2^T C$$

- α_1 is interpreted as the direct effect (not via M),
- β_1 is interpreted as the total effect,
- and thus $\beta_1 - \alpha_1$ is the indirect effect (via M).

'Difference' method

Baron and Kenny, 1986



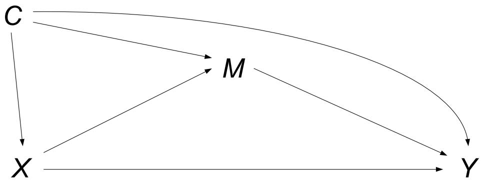
Consider two regression models:

$$E(Y|C,X,M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$
$$E(Y|C,X) = \beta_0 + \beta_1 X + \beta_2^T C$$

- α_1 is interpreted as the direct effect (not via M),
- β_1 is interpreted as the total effect,
- and thus $\beta_1 - \alpha_1$ is the indirect effect (via M).

'Difference' method

Baron and Kenny, 1986



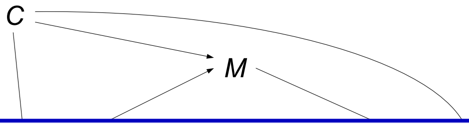
Consider two regression models:

$$E(Y|C,X,M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$
$$E(Y|C,X) = \beta_0 + \beta_1 X + \beta_2^T C$$

- α_1 is interpreted as the direct effect (not via M),
- β_1 is interpreted as the total effect,
- and thus $\beta_1 - \alpha_1$ is the indirect effect (via M).

'Difference' method

Baron and Kenny, 1986



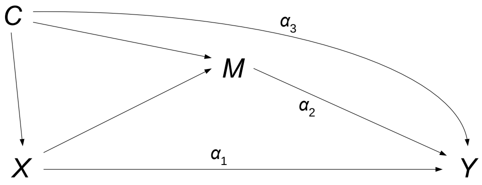
- Estimation via ordinary least squares.
- Various options (delta method, bootstrapping) to obtain SE for the indirect effect.

$$E(Y|C, X, M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$
$$E(Y|C, X) = \beta_0 + \beta_1 X + \beta_2^T C$$

- α_1 is interpreted as the direct effect (not via M),
- β_1 is interpreted as the total effect,
- and thus $\beta_1 - \alpha_1$ is the indirect effect (via M).

'Product' method

Wright, 1921



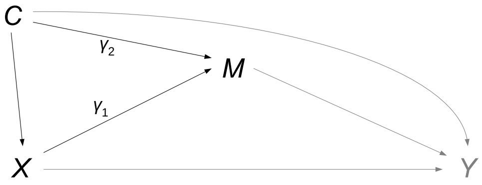
Alternatively, consider the two regression models:

$$E(Y \mid C, X, M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$

$$E(M \mid C, X) = \gamma_0 + \gamma_1 X + \gamma_2^T C$$

'Product' method

Wright, 1921

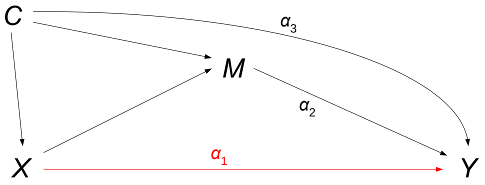


Alternatively, consider the two regression models:

$$E(Y \mid C, X, M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$
$$E(M \mid C, X) = \gamma_0 + \gamma_1 X + \gamma_2^T C$$

‘Product’ method

Wright, 1921



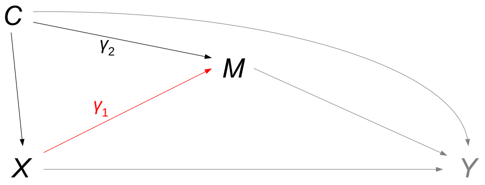
Alternatively, consider the two regression models:

$$E(Y\,|\,C,X,M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$
$$E(M\,|\,C,X) = \gamma_0 + \gamma_1 X + \gamma_2^T C$$

- α_1 is (as before) the direct effect,
- and now $\gamma_1\alpha_2$ is the indirect effect.
- These (‘difference’, ‘product’) are equivalent: $\beta_1 - \alpha_1 = \gamma_1\alpha_2$.

'Product' method

Wright, 1921



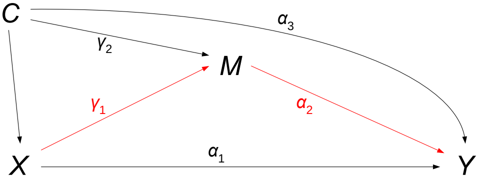
Alternatively, consider the two regression models:

$$E(Y|C,X,M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$
$$E(M|C,X) = \gamma_0 + \gamma_1 X + \gamma_2^T C$$

- α_1 is (as before) the direct effect,
- and now $\gamma_1\alpha_2$ is the indirect effect.
- These ('difference', 'product') are equivalent: $\beta_1 - \alpha_1 = \gamma_1\alpha_2$.

‘Product’ method

Wright, 1921



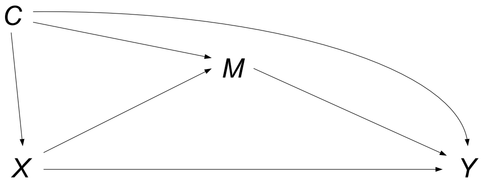
Alternatively, consider the two regression models:

$$E(Y\,|\,C,X,M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$
$$E(M\,|\,C,X) = \gamma_0 + \gamma_1 X + \gamma_2^T C$$

- α_1 is (as before) the direct effect,
- and now $\gamma_1\alpha_2$ is the indirect effect.
- These (‘difference’, ‘product’) are equivalent: $\beta_1 - \alpha_1 = \gamma_1\alpha_2$.

'Product' method

Wright, 1921



Alternatively, consider the two regression models:

$$E(Y \mid C, X, M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$
$$E(M \mid C, X) = \gamma_0 + \gamma_1 X + \gamma_2^T C$$

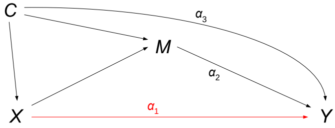
- α_1 is (as before) the direct effect,
- and now $\gamma_1 \alpha_2$ is the indirect effect.
- These ('difference', 'product') are equivalent: $\beta_1 - \alpha_1 = \gamma_1 \alpha_2$.

Outline

Outline

Problem 1: estimands and assumptions

Specific to this (associational) model; correspondence to direct/indirect vague

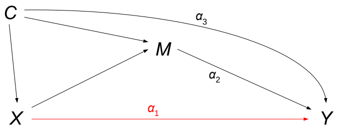


Consider the regression model that includes α_1:

$$E(Y \mid C, X, M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$

Problem 1: estimands and assumptions

Specific to this (associational) model; correspondence to direct/indirect vague



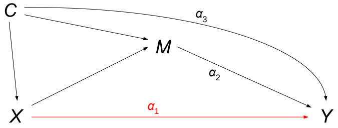
Consider the regression model that includes α_1:

$$E(Y \mid C, X, M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$

- α_1 is just a parameter in an associational model.
- For two subjects with the same values of C and M, but values of X that differ by 1 unit, it is the expected difference in their values of Y.
- Under what conditions does this correspond to a direct effect?
- Does this question even make sense without a definition (other than α_1) of 'direct effect'?

Problem 1: estimands and assumptions

Specific to this (associational) model; correspondence to direct/indirect vague



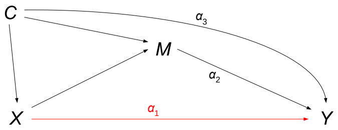
Consider the regression model that includes α_1:

$$E(Y \mid C, X, M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$

- α_1 is just a parameter in an associational model.
- For two subjects with the same values of C and M, but values of X that differ by 1 unit, it is the expected difference in their values of Y.
- Under what conditions does this correspond to a direct effect?
- Does this question even make sense without a definition (other than α_1) of 'direct effect'?

Problem 1: estimands and assumptions

Specific to this (associational) model; correspondence to direct/indirect vague



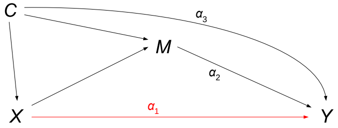
Consider the regression model that includes α_1:

$$E(Y \mid C, X, M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$

- α_1 is just a parameter in an associational model.
- For two subjects with the same values of C and M, but values of X that differ by 1 unit, it is the expected difference in their values of Y.
- Under what conditions does this correspond to a direct effect?
- Does this question even make sense without a definition (other than α_1) of 'direct effect'?

Problem 1: estimands and assumptions

Specific to this (associational) model; correspondence to direct/indirect vague



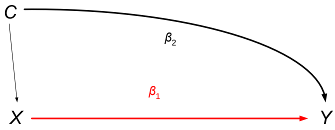
Consider the regression model that includes α_1:

$$E(Y \mid C, X, M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$

- α_1 is just a parameter in an associational model.
- For two subjects with the same values of C and M, but values of X that differ by 1 unit, it is the expected difference in their values of Y.
- Under what conditions does this correspond to a direct effect?
- Does this question even make sense without a definition (other than α_1) of 'direct effect'?

Problem 1: estimands and assumptions

Specific to this (associational) model; correspondence to direct/indirect vague

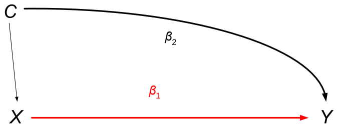


Consider the regression model that includes X and C but not M:

$$E(Y|C,X) = \beta_0 + \beta_1 X + \beta_2^T C$$

Problem 1: estimands and assumptions

Specific to this (associational) model; correspondence to direct/indirect vague



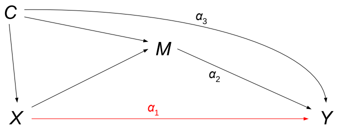
Consider the regression model that includes X and C but not M:

$$E(Y \mid C, X) = \beta_0 + \beta_1 X + \beta_2^T C$$

- β_1 is the total causal effect of X on Y only if C is sufficient to control for all confounding between X and Y.
- There must surely be similar conditions for α_1 to be interpreted as a direct effect.
- No unmeasured confounding of M and Y? Of X and M? In addition to no unmeasured confounding of X on Y?
- Without a (model-free) definition of direct effect, this is impossible to establish.

Problem 1: estimands and assumptions

Specific to this (associational) model; correspondence to direct/indirect vague



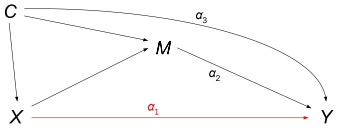
Consider the regression model that includes X and C but not M:

$$E(Y|C,X) = \beta_0 + \beta_1 X + \beta_2^T C$$

- β_1 is the total causal effect of X on Y only if C is sufficient to control for all confounding between X and Y.
- There must surely be similar conditions for α_1 to be interpreted as a direct effect.
- No unmeasured confounding of M and Y? Of X and M? In addition to no unmeasured confounding of X on Y?
- Without a (model-free) definition of direct effect, this is impossible to establish.

Problem 1: estimands and assumptions

Specific to this (associational) model; correspondence to direct/indirect vague



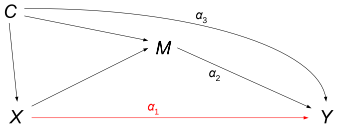
Consider the regression model that includes X and C but not M:

$$E(Y \mid C, X) = \beta_0 + \beta_1 X + \beta_2^T C$$

- β_1 is the total causal effect of X on Y only if C is sufficient to control for all confounding between X and Y.
- There must surely be similar conditions for α_1 to be interpreted as a direct effect.
- No unmeasured confounding of M and Y? Of X and M? In addition to no unmeasured confounding of X on Y?
- Without a (model-free) definition of direct effect, this is impossible to establish.

Problem 1: estimands and assumptions

Specific to this (associational) model; correspondence to direct/indirect vague



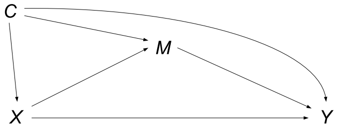
Consider the regression model that includes X and C but not M:

$$E(Y\,|\,C,X) = \beta_0 + \beta_1 X + \beta_2^T C$$

- β_1 is the total causal effect of X on Y only if C is sufficient to control for all confounding between X and Y.
- There must surely be similar conditions for α_1 to be interpreted as a direct effect.
- No unmeasured confounding of M and Y? Of X and M? In addition to no unmeasured confounding of X on Y?
- Without a (model-free) definition of direct effect, this is impossible to establish.

Outline

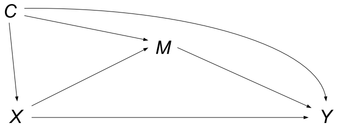
Problem 2: model (in)flexibility



Consider the regression models again:

$$E(Y \mid C, X, M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$
$$E(Y \mid C, X) = \beta_0 + \beta_1 X + \beta_2^T C$$
$$E(M \mid C, X) = \gamma_0 + \gamma_1 X + \gamma_2^T C$$

Problem 2: model (in)flexibility

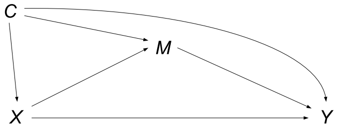


Consider the regression models again:

$$E(Y \mid C, X, M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$
$$E(Y \mid C, X) = \beta_0 + \beta_1 X + \beta_2^T C$$
$$E(M \mid C, X) = \gamma_0 + \gamma_1 X + \gamma_2^T C$$

- These are flexible for C.
- We can include interactions between confounders, and non-linearities for the confounders.
- But the flexibility ends here.

Problem 2: model (in)flexibility



Consider the regression models again:

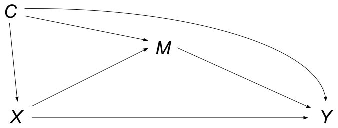
$$E(Y \mid C, X, M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_{31} C_1 + \alpha_{32} C_2 + \alpha_{33} C_1 C_2$$

$$E(Y \mid C, X) = \beta_0 + \beta_1 X + \beta_{31} C_1 + \beta_{32} C_2 + \beta_{33} C_1 C_2$$

$$E(M \mid C, X) = \gamma_0 + \gamma_1 X + \gamma_{31} C_1 + \gamma_{32} C_2 + \alpha_{33} C_1 C_2$$

- These are flexible for C.
- We can include interactions between confounders, and non-linearities for the confounders.
- But the flexibility ends here.

Problem 2: model (in)flexibility

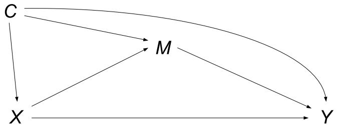


Consider the regression models again:

$$E(Y \mid C, X, M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_{31} C_1 + \alpha_{32} \exp(C_2)$$
$$E(Y \mid C, X) = \beta_0 + \beta_1 X + \beta_{31} C_1 + \beta_{32} C_1^2 + \beta_{33} C_2$$
$$E(M \mid C, X) = \gamma_0 + \gamma_1 X + \gamma_{31} C_1 + \gamma_{32} \log(C_2) + \gamma_{33} \sqrt{C_1 C_2}$$

- These are flexible for C.
- We can include interactions between confounders, and non-linearities for the confounders.
- But the flexibility ends here.

Problem 2: model (in)flexibility

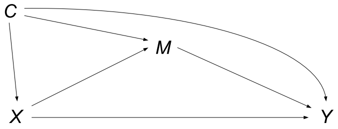


Consider the regression models again:

$$E(Y|C,X,M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$
$$E(Y|C,X) = \beta_0 + \beta_1 X + \beta_2^T C$$
$$E(M|C,X) = \gamma_0 + \gamma_1 X + \gamma_2^T C$$

- These are flexible for C.
- We can include interactions between confounders, and non-linearities for the confounders.
- But the flexibility ends here.

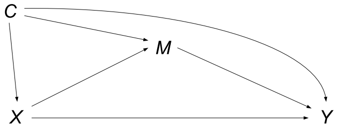
Problem 2: model (in)flexibility



Consider the regression models again:

$$E(Y \mid C, X, M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C$$
$$E(Y \mid C, X) = \beta_0 + \beta_1 X + \beta_2^T C$$
$$E(M \mid C, X) = \gamma_0 + \gamma_1 X + \gamma_2^T C$$

Problem 2: model (in)flexibility

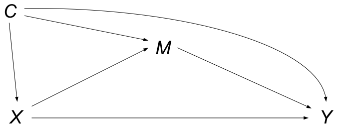


Consider the regression models again:

$$E(Y|C,X,M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C + \alpha_4 XM$$
$$E(Y|C,X) = \beta_0 + \beta_1 X + \beta_2^T C$$
$$E(M|C,X) = \gamma_0 + \gamma_1 X + \gamma_2^T C$$

- What about interactions between X and other terms?
- Non-linearities for X?
- Non-linear link functions (eg for binary Y)?
- Not clear how 'difference' and 'product' methods extend.

Problem 2: model (in)flexibility

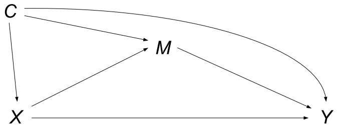


Consider the regression models again:

$$E(Y \mid C, X, M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C + \alpha_4^T XC$$
$$E(Y \mid C, X) = \beta_0 + \beta_1 X + \beta_2^T C + \beta_3^T XC$$
$$E(M \mid C, X) = \gamma_0 + \gamma_1 X + \gamma_2^T C + \gamma_3^T XC$$

- What about interactions between X and other terms?
- Non-linearities for X?
- Non-linear link functions (eg for binary Y)?
- Not clear how 'difference' and 'product' methods extend.

Problem 2: model (in)flexibility

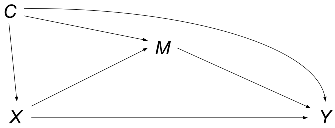


Consider the regression models again:

$$E(Y\,|\,C,X,M) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C + \alpha_4 X^2$$
$$E(Y\,|\,C,X) = \beta_0 + \beta_1 X + \beta_2^T C + \beta_3 X^2$$
$$E(M\,|\,C,X) = \gamma_0 + \gamma_1 X + \gamma_2^T C + \gamma_3 X^2$$

- What about interactions between X and other terms?
- Non-linearities for X?
- Non-linear link functions (eg for binary Y)?
- Not clear how 'difference' and 'product' methods extend.

Problem 2: model (in)flexibility

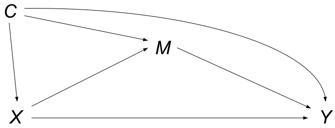


Consider the regression models again:

$$E(Y \mid C, X, M) = \text{logit}^{-1}(\alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C)$$
$$E(Y \mid C, X) = \text{logit}^{-1}(\beta_0 + \beta_1 X + \beta_2^T C)$$
$$E(M \mid C, X) = \gamma_0 + \gamma_1 X + \gamma_2^T C$$

- What about interactions between X and other terms?
- Non-linearities for X?
- Non-linear link functions (eg for binary Y)?
- Not clear how 'difference' and 'product' methods extend.

Problem 2: model (in)flexibility



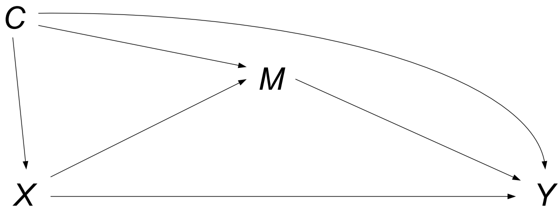
Consider the regression models again:

$$E(Y|C,X,M) = \text{logit}^{-1}(\alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3^T C + \alpha_4 XM)$$
$$E(Y|C,X) = \text{logit}^{-1}(\beta_0 + \beta_1 X + \beta_2^T C + \beta_3^T XC)$$
$$E(M|C,X) = \gamma_0 + \gamma_1 X + \gamma_2^T C + \gamma_3 X^2$$

- What about interactions between X and other terms?
- Non-linearities for X?
- Non-linear link functions (eg for binary Y)?
- Not clear how 'difference' and 'product' methods extend.

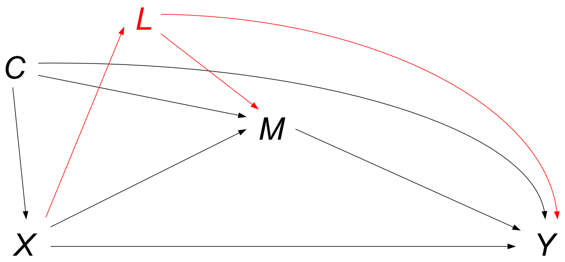
Outline

Problem 3: intermediate confounding



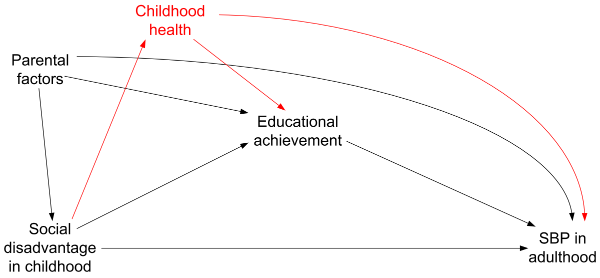
- Thus far, we have considered confounders C that are (potentially) common causes of X, M and Y.
- Intermediate confounders L are common causes of M and Y that are affected by X.
- eg childhood health in our example.

Problem 3: intermediate confounding



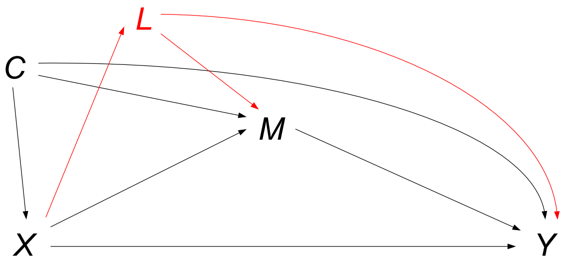
- Thus far, we have considered confounders C that are (potentially) common causes of X, M and Y.
- Intermediate confounders L are common causes of M and Y that are affected by X.
- eg childhood health in our example.

Problem 3: intermediate confounding



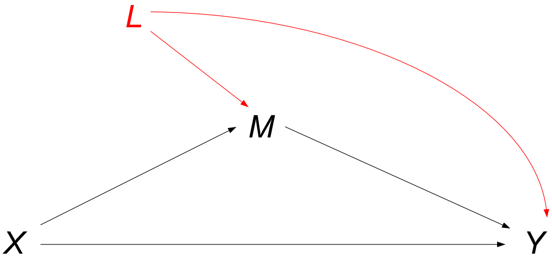
- Thus far, we have considered confounders C that are (potentially) common causes of X, M and Y.
- Intermediate confounders L are common causes of M and Y that are affected by X.
- eg childhood health in our example.

Problem 3: intermediate confounding



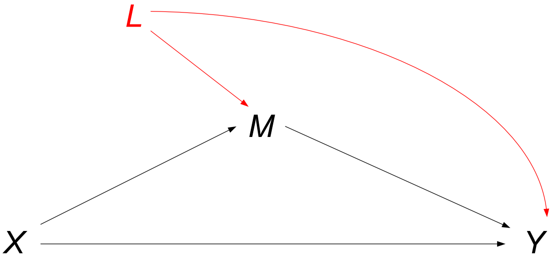
- Such L are problematic as we now show.
- Let us ignore C for simplicity, and, let us even ignore the arrow from X to L at first, ie L is NOT an intermediate confounder in this diagram for now...

Problem 3: intermediate confounding



- Such L are problematic as we now show.
- Let us ignore C for simplicity, and, let us even ignore the arrow from X to L at first, ie L is NOT an intermediate confounder in this diagram for now...

Problem 3: intermediate confounding

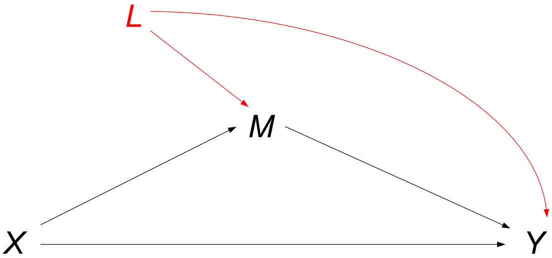


- Why can't we ignore L altogether?
- Recall that the regression model (forgetting C) that defines the direct effect α_1 is:

$$E(Y|X,M) = \alpha_0 + \alpha_1 X + \alpha_2 M.$$

- This model conditions on M.

Problem 3: intermediate confounding

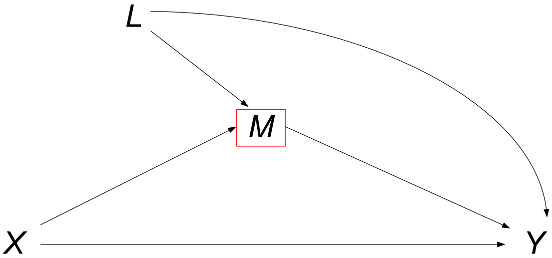


- Why can't we ignore L altogether?
- Recall that the regression model (forgetting C) that defines the direct effect α_1 is:

$$E(Y|X,M) = \alpha_0 + \alpha_1 X + \alpha_2 M.$$

- This model conditions on M.

Problem 3: intermediate confounding

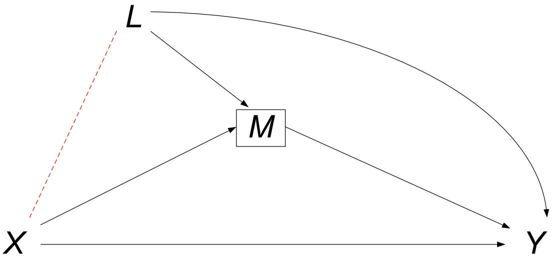


- Why can't we ignore L altogether?
- Recall that the regression model (forgetting C) that defines the direct effect α_1 is:

$$E(Y|X, M) = \alpha_0 + \alpha_1 X + \alpha_2 M.$$

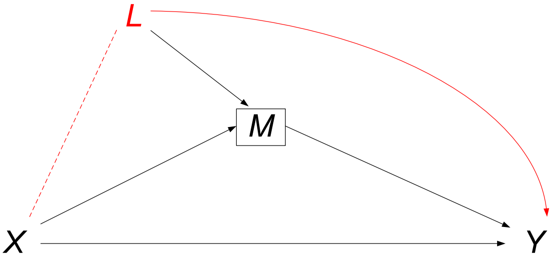
- This model conditions on M.

Problem 3: intermediate confounding



- Conditioning on M induces an association between X and L even if there was none there before (and would alter an existing association)—why?

Problem 3: intermediate confounding



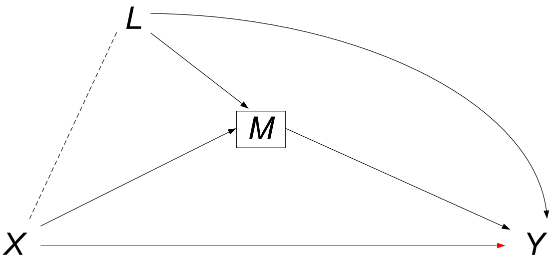
- Thus α_1 in:

$$E(Y|X, M) = \alpha_0 + \alpha_1 X + \alpha_2 M$$

represents this spurious association...

- ... as well as this direct effect.

Problem 3: intermediate confounding



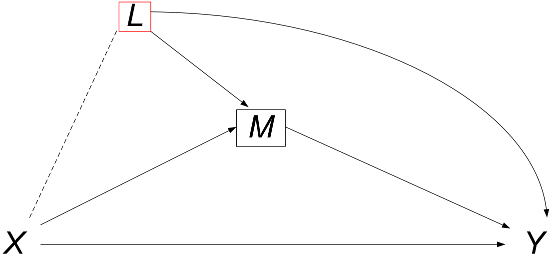
- Thus α_1 in:

$$E(Y|X, M) = \alpha_0 + \alpha_1 X + \alpha_2 M$$

represents this spurious association...

- ...as well as this direct effect.

Problem 3: intermediate confounding

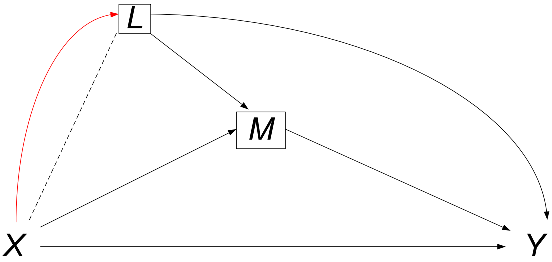


- A solution would be to include L in the model:

$$E(Y|X, M, L) = \alpha_0 + \alpha_1 X + \alpha_2 M + \alpha_3 L$$

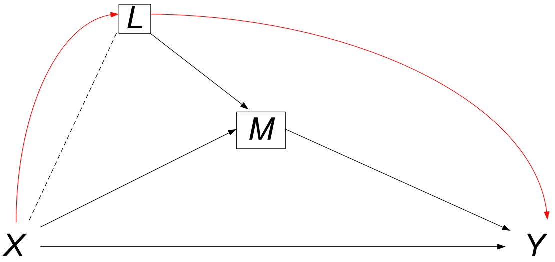
—blocking the spurious association.

Problem 3: intermediate confounding



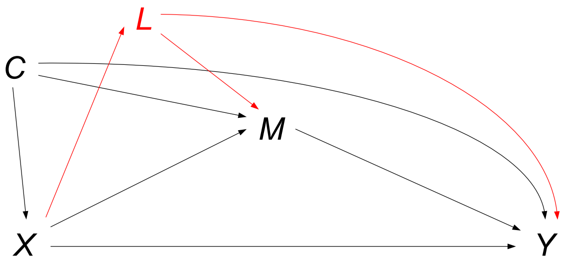
- But this is NOT a solution when L is affected by X.
- Since we block part of the direct effect (unmediated by M).
- Thus it is unclear how the 'old' approaches can be used when there is intermediate confounding.

Problem 3: intermediate confounding



- But this is NOT a solution when L is affected by X.
- Since we block part of the direct effect (unmediated by M).
- Thus it is unclear how the 'old' approaches can be used when there is intermediate confounding.

Problem 3: intermediate confounding



- But this is NOT a solution when L is affected by X.
- Since we block part of the direct effect (unmediated by M).
- Thus it is unclear how the ‘old’ approaches can be used when there is intermediate confounding.

Summary so far

- Traditional methods for estimating direct and indirect effects (more generally pathways) suffer from some limitations:
  1. They give no model-free definitions of direct/indirect effect.
  2. It is therefore unclear under what assumptions the parameters being estimated can be interpreted as direct/indirect effects.
  3. The models are restricted to be linear.
  4. It is unclear how intermediate confounders can be dealt with.
- More recent contributions from the causal inference literature have brought clarity to these issues, and greater flexibility to the modelling.
- Issues 1–3, and, to a certain extent, 4, have been addressed.

Summary so far

- Traditional methods for estimating direct and indirect effects (more generally pathways) suffer from some limitations:
  1. They give no model-free definitions of direct/indirect effect.
  2. It is therefore unclear under what assumptions the parameters being estimated can be interpreted as direct/indirect effects.
  3. The models are restricted to be linear.
  4. It is unclear how intermediate confounders can be dealt with.
- More recent contributions from the causal inference literature have brought clarity to these issues, and greater flexibility to the modelling.
- Issues 1–3, and, to a certain extent, 4, have been addressed.

Summary so far

- Traditional methods for estimating direct and indirect effects (more generally pathways) suffer from some limitations:
  1. They give no model-free definitions of direct/indirect effect.
  2. It is therefore unclear under what assumptions the parameters being estimated can be interpreted as direct/indirect effects.
  3. The models are restricted to be linear.
  4. It is unclear how intermediate confounders can be dealt with.
- More recent contributions from the causal inference literature have brought clarity to these issues, and greater flexibility to the modelling.
- Issues 1–3, and, to a certain extent, 4, have been addressed.

Summary so far

- Traditional methods for estimating direct and indirect effects (more generally pathways) suffer from some limitations:
  1. They give no model-free definitions of direct/indirect effect.
  2. It is therefore unclear under what assumptions the parameters being estimated can be interpreted as direct/indirect effects.
  3. The models are restricted to be linear.
  4. It is unclear how intermediate confounders can be dealt with.
- More recent contributions from the causal inference literature have brought clarity to these issues, and greater flexibility to the modelling.
- Issues 1–3, and, to a certain extent, 4, have been addressed.

Summary so far

- Traditional methods for estimating direct and indirect effects (more generally pathways) suffer from some limitations:
  1. They give no model-free definitions of direct/indirect effect.
  2. It is therefore unclear under what assumptions the parameters being estimated can be interpreted as direct/indirect effects.
  3. The models are restricted to be linear.
  4. It is unclear how intermediate confounders can be dealt with.
- More recent contributions from the causal inference literature have brought clarity to these issues, and greater flexibility to the modelling.
- Issues 1–3, and, to a certain extent, 4, have been addressed.

Summary so far

- Traditional methods for estimating direct and indirect effects (more generally pathways) suffer from some limitations:
  1. They give no model-free definitions of direct/indirect effect.
  2. It is therefore unclear under what assumptions the parameters being estimated can be interpreted as direct/indirect effects.
  3. The models are restricted to be linear.
  4. It is unclear how intermediate confounders can be dealt with.
- More recent contributions from the causal inference literature have brought clarity to these issues, and greater flexibility to the modelling.
- Issues 1–3, and, to a certain extent, 4, have been addressed.

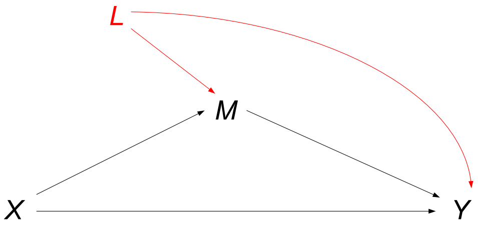
Summary so far

- Traditional methods for estimating direct and indirect effects (more generally pathways) suffer from some limitations:
  1. They give no model-free definitions of direct/indirect effect.
  2. It is therefore unclear under what assumptions the parameters being estimated can be interpreted as direct/indirect effects.
  3. The models are restricted to be linear.
  4. It is unclear how intermediate confounders can be dealt with.
- More recent contributions from the causal inference literature have brought clarity to these issues, and greater flexibility to the modelling.
- Issues 1–3, and, to a certain extent, 4, have been addressed.

Outline

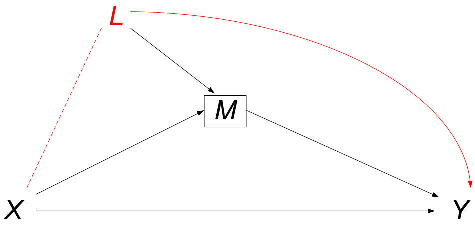
Outline

Direct and indirect effects are about causality



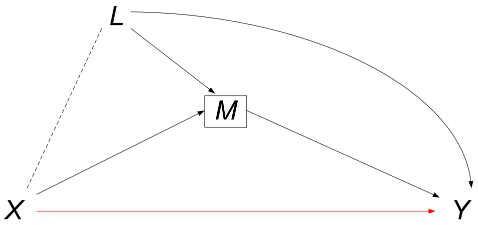
- Recall that if there are ignored common causes L of M and Y, the 'direct effect' we would naïvely estimate includes an association via this pathway...
- ...as well as this one.
- Thus the whole enterprise makes sense only if we are talking about path-specific causal effects.
- There's no such thing as a direct association.

Direct and indirect effects are about causality



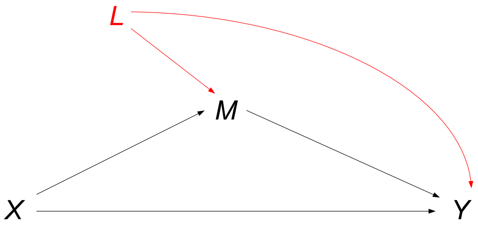
- Recall that if there are ignored common causes L of M and Y, the 'direct effect' we would naïvely estimate includes an association via this pathway. . .
- . . . as well as this one.
- Thus the whole enterprise makes sense only if we are talking about path-specific causal effects.
- There's no such thing as a direct association.

Direct and indirect effects are about causality



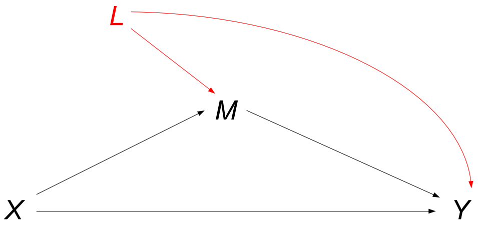
- Recall that if there are ignored common causes L of M and Y, the 'direct effect' we would naïvely estimate includes an association via this pathway. . .
- . . . as well as this one.
- Thus the whole enterprise makes sense only if we are talking about path-specific causal effects.
- There's no such thing as a direct association.

Direct and indirect effects are about causality



- Recall that if there are ignored common causes L of M and Y, the 'direct effect' we would naïvely estimate includes an association via this pathway. . .
- . . . as well as this one.
- Thus the whole enterprise makes sense only if we are talking about path-specific causal effects.
- There's no such thing as a direct association.

Direct and indirect effects are about causality



- Recall that if there are ignored common causes L of M and Y, the 'direct effect' we would naïvely estimate includes an association via this pathway. . .
- . . . as well as this one.
- Thus the whole enterprise makes sense only if we are talking about path-specific causal effects.
- There's no such thing as a direct association.

Counterfactuals

- Causal, unlike associational, quantities are not just about describing this world, but involve a notion of how the world would have been had something been different.
- The causal quantities we will define thus require counterfactuals (or equivalent).
- So first let's define the counterfactuals we'll need.

Counterfactuals

- Causal, unlike associational, quantities are not just about describing this world, but involve a notion of how the world would have been had something been different.
- The causal quantities we will define thus require counterfactuals (or equivalent).
- So first let's define the counterfactuals we'll need.

Counterfactuals

- Causal, unlike associational, quantities are not just about describing this world, but involve a notion of how the world would have been had something been different.
- The causal quantities we will define thus require counterfactuals (or equivalent).
- So first let's define the counterfactuals we'll need.

Counterfactuals

- Let $Y(x)$ be the value that Y would take if we intervened on X and set it (possibly counter to fact) to the value x.
- Let $Y(x, m)$ be the value that Y would take if we intervened simultaneously on both X and M and set them to the values x and m.
- Let $M(x)$ be the value that M would take if we intervened on X and set it to x.
- Let $Y\{x, M(x^*)\}$ be the value that Y would take if we intervened on X and set it to x whilst simultaneously intervening on M and setting it to $M(x^*)$, the value that M would take under an intervention setting X to x^*, where x and x^* are not necessarily equal.

Counterfactuals

- Let $Y(x)$ be the value that Y would take if we intervened on X and set it (possibly counter to fact) to the value x.
- Let $Y(x, m)$ be the value that Y would take if we intervened simultaneously on both X and M and set them to the values x and m.
- Let $M(x)$ be the value that M would take if we intervened on X and set it to x.
- Let $Y\{x, M(x^*)\}$ be the value that Y would take if we intervened on X and set it to x whilst simultaneously intervening on M and setting it to $M(x^*)$, the value that M would take under an intervention setting X to x^*, where x and x^* are not necessarily equal.

Counterfactuals

- Let $Y(x)$ be the value that Y would take if we intervened on X and set it (possibly counter to fact) to the value x.
- Let $Y(x, m)$ be the value that Y would take if we intervened simultaneously on both X and M and set them to the values x and m.
- Let $M(x)$ be the value that M would take if we intervened on X and set it to x.
- Let $Y\{x, M(x^*)\}$ be the value that Y would take if we intervened on X and set it to x whilst simultaneously intervening on M and setting it to $M(x^*)$, the value that M would take under an intervention setting X to x^*, where x and x^* are not necessarily equal.

Counterfactuals

- Let $Y(x)$ be the value that Y would take if we intervened on X and set it (possibly counter to fact) to the value x.
- Let $Y(x, m)$ be the value that Y would take if we intervened simultaneously on both X and M and set them to the values x and m.
- Let $M(x)$ be the value that M would take if we intervened on X and set it to x.
- Let $Y\{x, M(x^*)\}$ be the value that Y would take if we intervened on X and set it to x whilst simultaneously intervening on M and setting it to $M(x^*)$, the value that M would take under an intervention setting X to x^*, where x and x^* are not necessarily equal.

Counterfactuals

- Let $Y(x)$ be the value that Y would take if we intervened on X and set it (possibly counter to fact) to the value x.
- Let $Y(x, m)$ be the value that Y would take if we intervened simultaneously on both X and M and set them to the values x and m.
- Let $M(x)$ be the value that M would take if we intervened on X and set it to x.
- Let $Y\{x, M(x^*)\}$ be the value that Y would take if we intervened on X and set it to x whilst simultaneously intervening on M and setting it to $M(x^*)$, the value that M would take under an intervention setting X to x^*, where x and x^* are not necessarily equal.

These counterfactuals are central to the (model-free) definitions of direct/indirect effects in causal inference.

Estimands

- Many (subtly different) counterfactual definitions of direct/indirect effects have been proposed.
- Direct effects:
  - Controlled direct effect (Pearl, 2001),
  - Natural direct effect (Pearl, 2001), also called Pure direct effect (Robins and Greenland, 1992),
  - Total direct effect (Robins and Greenland, 1992),
  - Direct effect in the exposed (Vansteelandt and VanderWeele, 2012),
  - Principal stratum direct effect (Rubin, 2004).
- Indirect effects:
  - Natural indirect effect (Pearl, 2001), also called Total indirect effect (Robins and Greenland, 1992),
  - Pure indirect effect (Robins and Greenland, 1992),
  - Indirect effect in the exposed (Vansteelandt and VanderWeele, 2012).
- We focus on these today (and also the total causal effect).

Estimands

- Many (subtly different) counterfactual definitions of direct/indirect effects have been proposed.
- Direct effects:
  - Controlled direct effect (Pearl, 2001),
  - Natural direct effect (Pearl, 2001), also called Pure direct effect (Robins and Greenland, 1992),
  - Total direct effect (Robins and Greenland, 1992),
  - Direct effect in the exposed (Vansteelandt and VanderWeele, 2012),
  - Principal stratum direct effect (Rubin, 2004).
- Indirect effects:
  - Natural indirect effect (Pearl, 2001), also called Total indirect effect (Robins and Greenland, 1992),
  - Pure indirect effect (Robins and Greenland, 1992),
  - Indirect effect in the exposed (Vansteelandt and VanderWeele, 2012).
- We focus on these today (and also the total causal effect).

Estimands

- Many (subtly different) counterfactual definitions of direct/indirect effects have been proposed.
- Direct effects:
  - Controlled direct effect (Pearl, 2001),
  - Natural direct effect (Pearl, 2001), also called Pure direct effect (Robins and Greenland, 1992),
  - Total direct effect (Robins and Greenland, 1992),
  - Direct effect in the exposed (Vansteelandt and VanderWeele, 2012),
  - Principal stratum direct effect (Rubin, 2004).
- Indirect effects:
  - Natural indirect effect (Pearl, 2001), also called Total indirect effect (Robins and Greenland, 1992),
  - Pure indirect effect (Robins and Greenland, 1992),
  - Indirect effect in the exposed (Vansteelandt and VanderWeele, 2012).
- We focus on these today (and also the total causal effect).

Estimands

- Many (subtly different) counterfactual definitions of direct/indirect effects have been proposed.
- Direct effects:
  - Controlled direct effect (Pearl, 2001),
  - Natural direct effect (Pearl, 2001), also called Pure direct effect (Robins and Greenland, 1992),
  - Total direct effect (Robins and Greenland, 1992),
  - Direct effect in the exposed (Vansteelandt and VanderWeele, 2012),
  - Principal stratum direct effect (Rubin, 2004).
- Indirect effects:
  - Natural indirect effect (Pearl, 2001), also called Total indirect effect (Robins and Greenland, 1992),
  - Pure indirect effect (Robins and Greenland, 1992),
  - Indirect effect in the exposed (Vansteelandt and VanderWeele, 2012).
- We focus on these today (and also the total causal effect).

Estimands

- Many (subtly different) counterfactual definitions of direct/indirect effects have been proposed.
- Direct effects:
  - Controlled direct effect (Pearl, 2001),
  - Natural direct effect (Pearl, 2001), also called Pure direct effect (Robins and Greenland, 1992),
  - Total direct effect (Robins and Greenland, 1992),
  - Direct effect in the exposed (Vansteelandt and VanderWeele, 2012),
  - Principal stratum direct effect (Rubin, 2004).
- Indirect effects:
  - Natural indirect effect (Pearl, 2001), also called Total indirect effect (Robins and Greenland, 1992),
  - Pure indirect effect (Robins and Greenland, 1992),
  - Indirect effect in the exposed (Vansteelandt and VanderWeele, 2012).
- We focus on these today (and also the total causal effect).

Total causal effect

- The total causal effect of X on Y, conditional on $C = c$, expressed as a mean difference comparing x^* vs x is

$$\text{TCE}(c, x, x^*) = E\{Y(x^*)|C = c\} - E\{Y(x)|C = c\}.$$

- Note that this can also be written as

$$\text{TCE}(c, x, x^*) = E[Y\{x^*, M(x^*)\}|C = c] - E[Y\{x, M(x)\}|C = c].$$

Total causal effect

- The total causal effect of X on Y, conditional on $C = c$, expressed as a mean difference comparing x^* vs x is

$$\text{TCE}\left(c, x, x^*\right) = E\left\{Y\left(x^*\right)|C = c\right\} - E\left\{Y\left(x\right)|C = c\right\}.$$

- Note that this can also be written as

$$\text{TCE}\left(c, x, x^*\right) = E\left[Y\left\{x^*, M\left(x^*\right)\right\}|C = c\right] - E\left[Y\left\{x, M\left(x\right)\right\}|C = c\right].$$

Controlled direct effect

Pearl, 2001

- The controlled direct effect of X on Y, conditional on $C = c$, when M is controlled at m, expressed as a mean difference comparing x^* vs x is

$$\text{CDE}\left(c, x, x^*, m\right) = E\left\{Y\left(x^*, m\right) | C = c\right\} - E\left\{Y\left(x, m\right) | C = c\right\}.$$

Controlled direct effect

Pearl, 2001

- The controlled direct effect of X on Y, conditional on $C = c$, when M is controlled at m, expressed as a mean difference comparing x^* vs x is

$$\text{CDE}\,(c, x, x^*, m) = E\,\{Y\,(x^*, m)\,|C = c\} - E\,\{Y\,(x, m)\,|C = c\}\,.$$

- This (as always with a causal contrast) is a comparison of two hypothetical worlds.
- In the first, X is set to x^*, and in the second X is set to x. In both worlds, M is set to m.
- By keeping M fixed at m, we are getting at the direct effect of X, unmediated by M.

Controlled direct effect

Pearl, 2001

- The controlled direct effect of X on Y, conditional on $C = c$, when M is controlled at m, expressed as a mean difference comparing x^* vs x is

$$\text{CDE}(c, x, x^*, m) = E\{Y(x^*, m)|C = c\} - E\{Y(x, m)|C = c\}.$$

- This (as always with a causal contrast) is a comparison of two hypothetical worlds.
- In the first, X is set to x^*, and in the second X is set to x. In both worlds, M is set to m.
- By keeping M fixed at m, we are getting at the direct effect of X, unmediated by M.

Controlled direct effect
Pearl, 2001

- The controlled direct effect of X on Y, conditional on $C = c$, when M is controlled at m, expressed as a mean difference comparing x^* vs x is

$$\text{CDE}(c, x, x^*, m) = E\{Y(x^*, m)|C = c\} - E\{Y(x, m)|C = c\}.$$

- This (as always with a causal contrast) is a comparison of two hypothetical worlds.
- In the first, X is set to x^*, and in the second X is set to x. In both worlds, M is set to m.
- By keeping M fixed at m, we are getting at the direct effect of X, unmediated by M.

Controlled indirect effect?

- Ideally, we would express the total causal effect as the sum of a direct and an indirect effect.
- But this is tricky using our definition of a controlled direct effect.
- The controlled direct effect could be different for each value of m.
- But the total causal effect is independent of m.

Controlled indirect effect?

- Ideally, we would express the total causal effect as the sum of a direct and an indirect effect.
- But this is tricky using our definition of a controlled direct effect.
- The controlled direct effect could be different for each value of m.
- But the total causal effect is independent of m.

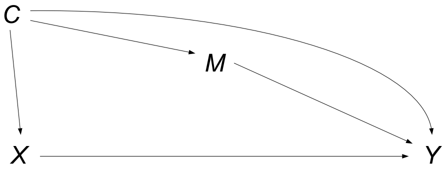
Controlled indirect effect?

- Ideally, we would express the total causal effect as the sum of a direct and an indirect effect.
- But this is tricky using our definition of a controlled direct effect.
- The controlled direct effect could be different for each value of m.
- But the total causal effect is independent of m.

Controlled indirect effect?

- Ideally, we would express the total causal effect as the sum of a direct and an indirect effect.
- But this is tricky using our definition of a controlled direct effect.
- The controlled direct effect could be different for each value of m.
- But the total causal effect is independent of m.

Controlled indirect effect?



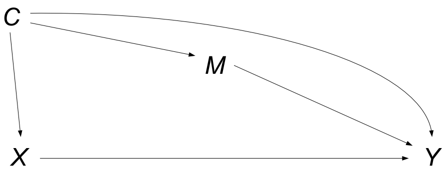
- Even with no mediation, an interaction would mean that

$$\mathrm{TCE}\,(c, x, x^*) \neq \mathrm{CDE}\,(c, x, x^*, m)\,.$$

- But in this case, the indirect effect would intuitively be zero.
- And so we cannot hope to find a definition of a controlled indirect effect (CIE) such that

$$\mathrm{TCE}\,(c, x, x^*) = \mathrm{CDE}\,(c, x, x^*, m) + \mathrm{CIE}\,(c, x, x^*, m)\,.$$

Controlled indirect effect?



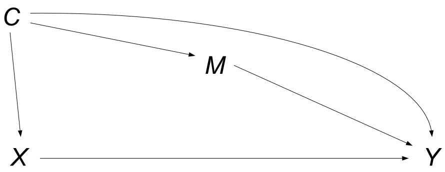
- Even with no mediation, an interaction would mean that

$$\text{TCE}(c, x, x^*) \neq \text{CDE}(c, x, x^*, m).$$

- But in this case, the indirect effect would intuitively be zero.
- And so we cannot hope to find a definition of a controlled indirect effect (CIE) such that

$$\text{TCE}(c, x, x^*) = \text{CDE}(c, x, x^*, m) + \text{CIE}(c, x, x^*, m).$$

Controlled indirect effect?



- Even with no mediation, an interaction would mean that

$$\text{TCE}\,(c, x, x^*) \neq \text{CDE}\,(c, x, x^*, m)\,.$$

- But in this case, the indirect effect would intuitively be zero.
- And so we cannot hope to find a definition of a controlled indirect effect (CIE) such that

$$\text{TCE}\,(c, x, x^*) = \text{CDE}\,(c, x, x^*, m) + \text{CIE}\,(c, x, x^*, m)\,.$$

Natural direct effect

Pearl, 2001; Robins and Greenland, 1992

- For this reason, it is useful to have a different definition of a direct effect.
- The natural direct effect of X on Y, conditional on $C = c$, expressed as a mean difference comparing x^* vs x is

$$\begin{aligned} \text{NDE}(c, x, x^*) = {} & E\left[Y\left\{x^*, M(x)\right\} | C = c\right] \\ & - E\left[Y\left\{x, M(x)\right\} | C = c\right]. \end{aligned}$$

Natural direct effect

Pearl, 2001; Robins and Greenland, 1992

- For this reason, it is useful to have a different definition of a direct effect.
- The natural direct effect of X on Y, conditional on $C = c$, expressed as a mean difference comparing x^* vs x is

$$\begin{aligned} \text{NDE}(c, x, x^*) = {} & E[Y\{x^*, M(x)\} | C = c] \\ & - E[Y\{x, M(x)\} | C = c]. \end{aligned}$$

Natural direct effect

Pearl, 2001; Robins and Greenland, 1992

- For this reason, it is useful to have a different definition of a direct effect.
- The natural direct effect of X on Y, conditional on $C = c$, expressed as a mean difference comparing x^* vs x is

$$\text{NDE}(c, x, x^*) = E\left[Y\left\{x^*, M(x)\right\}|C = c\right] - E\left[Y\left\{x, M(x)\right\}|C = c\right].$$

- This is a comparison of two hypothetical worlds.
- In the first, X is set to x^*, and in the second X is set to x. In both worlds, M is set to $M(x)$, the value it would take if X were set to x.
- Since M is the same (*within* subject) in both worlds, we are still getting at the direct effect of X.
- In the absence of individual-level interaction between X and M, $\text{CDE}(c, x, x^*, m) = \text{NDE}(c, x, x^*) \quad \forall m$.

Natural direct effect

Pearl, 2001; Robins and Greenland, 1992

- For this reason, it is useful to have a different definition of a direct effect.
- The natural direct effect of X on Y, conditional on $C = c$, expressed as a mean difference comparing x^* vs x is

$$\text{NDE}(c, x, x^*) = E[Y\{x^*, M(x)\} | C = c] - E[Y\{x, M(x)\} | C = c].$$

- This is a comparison of two hypothetical worlds.
- In the first, X is set to x^*, and in the second X is set to x. In both worlds, M is set to $M(x)$, the value it would take if X were set to x.
- Since M is the same (*within* subject) in both worlds, we are still getting at the direct effect of X.
- In the absence of individual-level interaction between X and M, $\text{CDE}(c, x, x^*, m) = \text{NDE}(c, x, x^*) \quad \forall m$.

Natural direct effect

Pearl, 2001; Robins and Greenland, 1992

- For this reason, it is useful to have a different definition of a direct effect.
- The natural direct effect of X on Y, conditional on $C = c$, expressed as a mean difference comparing x^* vs x is

$$\text{NDE}(c, x, x^*) = E[Y\{x^*, M(x)\}|C = c] - E[Y\{x, M(x)\}|C = c].$$

- This is a comparison of two hypothetical worlds.
- In the first, X is set to x^*, and in the second X is set to x. In both worlds, M is set to $M(x)$, the value it would take if X were set to x.
- Since M is the same (*within* subject) in both worlds, we are still getting at the direct effect of X.
- In the absence of individual-level interaction between X and M, $\text{CDE}(c, x, x^*, m) = \text{NDE}(c, x, x^*) \quad \forall m$.

Natural direct effect

Pearl, 2001; Robins and Greenland, 1992

- For this reason, it is useful to have a different definition of a direct effect.
- The natural direct effect of X on Y, conditional on $C = c$, expressed as a mean difference comparing x^* vs x is

$$\text{NDE}\left(c, x, x^*\right) = E\left[Y\left\{x^*, M\left(x\right)\right\}|C = c\right] - E\left[Y\left\{x, M\left(x\right)\right\}|C = c\right].$$

- This is a comparison of two hypothetical worlds.
- In the first, X is set to x^*, and in the second X is set to x. In both worlds, M is set to $M\left(x\right)$, the value it would take if X were set to x.
- Since M is the same (*within* subject) in both worlds, we are still getting at the direct effect of X.
- In the absence of individual-level interaction between X and M, $\text{CDE}\left(c, x, x^*, m\right) = \text{NDE}\left(c, x, x^*\right) \quad \forall m$.

Natural indirect effect

Pearl, 2001; Robins and Greenland, 1992

- The advantage of defining the natural direct effect in this way, is that it leads to a natural *in*direct effect.
- The natural indirect effect of X on Y, conditional on $C = c$, expressed as a mean difference comparing x^* vs x is

$$\begin{aligned}\mathrm{NIE}(c, x, x^*) = E\left[Y\left\{x^*, M(x^*)\right\}|C = c\right] \\ - E\left[Y\left\{x^*, M(x)\right\}|C = c\right].\end{aligned}$$

Natural indirect effect

Pearl, 2001; Robins and Greenland, 1992

- The advantage of defining the natural direct effect in this way, is that it leads to a natural *in*direct effect.
- The natural indirect effect of X on Y, conditional on $C = c$, expressed as a mean difference comparing x^* vs x is

$$\text{NIE}(c, x, x^*) = E\left[Y\{x^*, M(x^*)\} | C = c\right] - E\left[Y\{x^*, M(x)\} | C = c\right].$$

Natural indirect effect

Pearl, 2001; Robins and Greenland, 1992

- The advantage of defining the natural direct effect in this way, is that it leads to a natural *in*direct effect.
- The natural indirect effect of X on Y, conditional on $C = c$, expressed as a mean difference comparing x^* vs x is

$$\text{NIE}(c, x, x^*) = E\left[Y\{x^*, M(x^*)\}|C = c\right] - E\left[Y\{x^*, M(x)\}|C = c\right].$$

- This is a comparison of two hypothetical worlds.
- In the first, M is set to $M(x^*)$, the value it would take if X were set to x^* and in the second M is set to $M(x)$, the value it would take if X were set to x. In both worlds, X is set to x^*.
- X is allowed to influence Y only through its influence on M. Thus it is an indirect effect through M.

Natural indirect effect

Pearl, 2001; Robins and Greenland, 1992

- The advantage of defining the natural direct effect in this way, is that it leads to a natural *in*direct effect.
- The natural indirect effect of X on Y, conditional on $C = c$, expressed as a mean difference comparing x^* vs x is

$$\text{NIE}(c, x, x^*) = E\left[Y\{x^*, M(x^*)\}|C = c\right] - E\left[Y\{x^*, M(x)\}|C = c\right].$$

- This is a comparison of two hypothetical worlds.
- In the first, M is set to $M(x^*)$, the value it would take if X were set to x^* and in the second M is set to $M(x)$, the value it would take if X were set to x. In both worlds, X is set to x^*.
- X is allowed to influence Y only through its influence on M. Thus it is an indirect effect through M.

Natural indirect effect

Pearl, 2001; Robins and Greenland, 1992

- The advantage of defining the natural direct effect in this way, is that it leads to a natural *in*direct effect.
- The natural indirect effect of X on Y, conditional on $C = c$, expressed as a mean difference comparing x^* vs x is

$$\text{NIE}(c, x, x^*) = E\left[Y\left\{x^*, M\left(x^*\right)\right\}|C = c\right] - E\left[Y\left\{x^*, M\left(x\right)\right\}|C = c\right].$$

- This is a comparison of two hypothetical worlds.
- In the first, M is set to $M(x^*)$, the value it would take if X were set to x^* and in the second M is set to $M(x)$, the value it would take if X were set to x. In both worlds, X is set to x^*.
- X is allowed to influence Y only through its influence on M. Thus it is an indirect effect through M.

Effect decomposition

Now we see that the sum of the natural direct and indirect effects is

$$
\begin{aligned}
&\text{NDE}(c, x, x^*) + \text{NIE}(c, x, x^*) \\
&\quad = E[Y\{x^*, M(x)\}|C = c] - E[Y\{x, M(x)\}|C = c] \\
&\quad + E[Y\{x^*, M(x^*)\}|C = c] - E[Y\{x^*, M(x)\}|C = c] \\
&\quad = E[Y\{x^*, M(x^*)\}|C = c] - E[Y\{x, M(x)\}|C = c] \\
&\quad = \text{TCE}(c, x, x^*),
\end{aligned}
$$

the total causal effect, as desired.

Effect decomposition

Now we see that the sum of the natural direct and indirect effects is

$$
\begin{aligned}
&\mathrm{NDE}\left(c,x,x^{*}\right)+\mathrm{NIE}\left(c,x,x^{*}\right)\\
&\quad= E\left[Y\left\{x^{*},M\left(x\right)\right\}|C=c\right]-E\left[Y\left\{x,M\left(x\right)\right\}|C=c\right]\\
&\quad+E\left[Y\left\{x^{*},M\left(x^{*}\right)\right\}|C=c\right]-E\left[Y\left\{x^{*},M\left(x\right)\right\}|C=c\right]\\
&\quad= E\left[Y\left\{x^{*},M\left(x^{*}\right)\right\}|C=c\right]-E\left[Y\left\{x,M\left(x\right)\right\}|C=c\right]\\
&\quad= \mathrm{TCE}\left(c,x,x^{*}\right),
\end{aligned}
$$

the total causal effect, as desired.

Effect decomposition

Now we see that the sum of the natural direct and indirect effects is

$$
\begin{aligned}
&\text{NDE}(c, x, x^*) + \text{NIE}(c, x, x^*) \\
&\quad = E[Y\{x^*, M(x)\} | C = c] - E[Y\{x, M(x)\} | C = c] \\
&\quad + E[Y\{x^*, M(x^*)\} | C = c] - E[Y\{x^*, M(x)\} | C = c] \\
&\quad = E[Y\{x^*, M(x^*)\} | C = c] - E[Y\{x, M(x)\} | C = c] \\
&\quad = \text{TCE}(c, x, x^*),
\end{aligned}
$$

the total causal effect, as desired.

Controlled vs natural

- The choice between controlled and natural direct effect partly depends on the question.

Controlled vs natural

- The choice between controlled and natural direct effect partly depends on the question.

Controlled direct effect

What effect does intervening on social disadvantage have on later SBP if we also intervened on everyone's educational achievement and set it to a particular level?

- A hypothetical world in which educational achievement does not vary at all from child to child is strange. . .
- And how would we choose a particular set level? 7 GCSEs or 10?
- In other contexts, this is not so strange (eg a hypothetical world in which a disease is eradicated).

Controlled vs natural

- The choice between controlled and natural direct effect partly depends on the question.

Controlled direct effect

What effect does intervening on social disadvantage have on later SBP if we also intervened on everyone's educational achievement and set it to a particular level?

- A hypothetical world in which educational achievement does not vary at all from child to child is strange. . .
- And how would we choose a particular set level? 7 GCSEs or 10?
- In other contexts, this is not so strange (eg a hypothetical world in which a disease is eradicated).

Controlled vs natural

- The choice between controlled and natural direct effect partly depends on the question.

Controlled direct effect

What effect does intervening on social disadvantage have on later SBP if we also intervened on everyone's educational achievement and set it to a particular level?

- A hypothetical world in which educational achievement does not vary at all from child to child is strange. . .
- And how would we choose a particular set level? 7 GCSEs or 10?
- In other contexts, this is not so strange (eg a hypothetical world in which a disease is eradicated).

Controlled vs natural

- The choice between controlled and natural direct effect partly depends on the question.

Controlled direct effect

What effect does intervening on social disadvantage have on later SBP if we also intervened on everyone's educational achievement and set it to a particular level?

- A hypothetical world in which educational achievement does not vary at all from child to child is strange. . .
- And how would we choose a particular set level? 7 GCSEs or 10?
- In other contexts, this is not so strange (eg a hypothetical world in which a disease is eradicated).

Controlled vs natural

Natural direct effect

What effect does intervening on social disadvantage have on later SBP in a world in which there is no effect of social disadvantage on educational achievement?

- More meaningful in this context?
- Allows effect decomposition.
- But requires stronger conditions for identification (see later).

Controlled vs natural

Natural direct effect

What effect does intervening on social disadvantage have on later SBP in a world in which there is no effect of social disadvantage on educational achievement?

- More meaningful in this context?
- Allows effect decomposition.
- But requires stronger conditions for identification (see later).

Controlled vs natural

Natural direct effect

What effect does intervening on social disadvantage have on later SBP in a world in which there is no effect of social disadvantage on educational achievement?

- More meaningful in this context?
- Allows effect decomposition.
- But requires stronger conditions for identification (see later).

Controlled vs natural

Natural direct effect

What effect does intervening on social disadvantage have on later SBP in a world in which there is no effect of social disadvantage on educational achievement?

- More meaningful in this context?
- Allows effect decomposition.
- But requires stronger conditions for identification (see later).

Effects on alternative scales

Risk ratio scale (NB these decompose multiplicatively)

$$\mathrm{TCE}^{\mathrm{RR}}(c,x,x^*) = \frac{E\{Y(x^*)|C=c\}}{E\{Y(x)|C=c\}},$$

$$\mathrm{CDE}^{\mathrm{RR}}(c,x,x^*,m) = \frac{E\{Y(x^*,m)|C=c\}}{E\{Y(x,m)|C=c\}},$$

$$\mathrm{NDE}^{\mathrm{RR}}(c,x,x^*) = \frac{E[Y\{x^*,M(x)\}|C=c]}{E[Y\{x,M(x)\}|C=c]},$$

$$\mathrm{NIE}^{\mathrm{RR}}(c,x,x^*) = \frac{E[Y\{x^*,M(x^*)\}|C=c]}{E[Y\{x^*,M(x)\}|C=c]}.$$

Effects on alternative scales

Odds ratio scale (NB these decompose multiplicatively)

$$\text{TCE}^{\text{OR}}(c,x,x^*) = \frac{E\{Y(x^*)|C=c\}/[1-E\{Y(x^*)|C=c\}]}{E\{Y(x)|C=c\}/[1-E\{Y(x)|C=c\}]},$$

$$\text{CDE}^{\text{OR}}(c,x,x^*,m) = \frac{E\{Y(x^*,m)|C=c\}/[1-E\{Y(x^*,m)|C=c\}]}{E\{Y(x,m)|C=c\}/[1-E\{Y(x,m)|C=c\}]},$$

$$\text{NDE}^{\text{OR}}(c,x,x^*) = \frac{E[Y\{x^*,M(x)\}|C=c]/(1-E[Y\{x^*,M(x)\}|C=c])}{E[Y\{x,M(x)\}|C=c]/(1-E[Y\{x,M(x)\}|C=c])},$$

$$\text{NIE}^{\text{OR}}(c,x,x^*) = \frac{E[Y\{x^*,M(x^*)\}|C=c]/(1-E[Y\{x^*,M(x^*)\}|C=c])}{E[Y\{x^*,M(x)\}|C=c]/(1-E[Y\{x^*,M(x)\}|C=c])}.$$

What next?

- Given clear definitions of the estimands we would like to estimate, we can give assumptions under which they can be identified from data and methods for doing so.
- Whenever counterfactual quantities are to be estimated from actual data, assumptions are needed to link the two.
- The assumptions come in three flavours:
  1. Consistency assumptions: allow linking of counterfactual outcomes such as $Y(x, m)$ with the actual outcome Y, for certain subjects.
  2. Exchangeability assumptions: allow linking certain subjects with certain other subjects so that counterfactuals not identified by consistency, can be estimated by borrowing information across subjects.
  3. Modelling assumptions: allow this sharing of information to happen more efficiently. (See models and methods).

What next?

- Given clear definitions of the estimands we would like to estimate, we can give assumptions under which they can be identified from data and methods for doing so.
- Whenever counterfactual quantities are to be estimated from actual data, assumptions are needed to link the two.
- The assumptions come in three flavours:
  1. Consistency assumptions: allow linking of counterfactual outcomes such as $Y(x, m)$ with the actual outcome Y, for certain subjects.
  2. Exchangeability assumptions: allow linking certain subjects with certain other subjects so that counterfactuals not identified by consistency, can be estimated by borrowing information across subjects.
  3. Modelling assumptions: allow this sharing of information to happen more efficiently. (See models and methods).

What next?

- Given clear definitions of the estimands we would like to estimate, we can give assumptions under which they can be identified from data and methods for doing so.
- Whenever counterfactual quantities are to be estimated from actual data, assumptions are needed to link the two.
- The assumptions come in three flavours:
  1. Consistency assumptions: allow linking of counterfactual outcomes such as $Y(x, m)$ with the actual outcome Y, for certain subjects.
  2. Exchangeability assumptions: allow linking certain subjects with certain other subjects so that counterfactuals not identified by consistency, can be estimated by borrowing information across subjects.
  3. Modelling assumptions: allow this sharing of information to happen more efficiently. (See models and methods).

What next?

- Given clear definitions of the estimands we would like to estimate, we can give assumptions under which they can be identified from data and methods for doing so.
- Whenever counterfactual quantities are to be estimated from actual data, assumptions are needed to link the two.
- The assumptions come in three flavours:
  1. Consistency assumptions: allow linking of counterfactual outcomes such as $Y(x, m)$ with the actual outcome Y, for certain subjects.
  2. Exchangeability assumptions: allow linking certain subjects with certain other subjects so that counterfactuals not identified by consistency, can be estimated by borrowing information across subjects.
  3. Modelling assumptions: allow this sharing of information to happen more efficiently. (See models and methods).

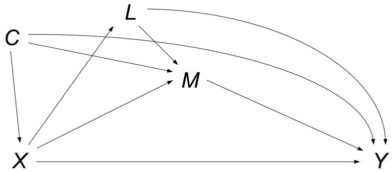
What next?

- Given clear definitions of the estimands we would like to estimate, we can give assumptions under which they can be identified from data and methods for doing so.
- Whenever counterfactual quantities are to be estimated from actual data, assumptions are needed to link the two.
- The assumptions come in three flavours:
  1. Consistency assumptions: allow linking of counterfactual outcomes such as $Y(x, m)$ with the actual outcome Y, for certain subjects.
  2. Exchangeability assumptions: allow linking certain subjects with certain other subjects so that counterfactuals not identified by consistency, can be estimated by borrowing information across subjects.
  3. Modelling assumptions: allow this sharing of information to happen more efficiently. (See models and methods).

What next?

- Given clear definitions of the estimands we would like to estimate, we can give assumptions under which they can be identified from data and methods for doing so.
- Whenever counterfactual quantities are to be estimated from actual data, assumptions are needed to link the two.
- The assumptions come in three flavours:
  1. Consistency assumptions: allow linking of counterfactual outcomes such as $Y(x, m)$ with the actual outcome Y, for certain subjects.
  2. Exchangeability assumptions: allow linking certain subjects with certain other subjects so that counterfactuals not identified by consistency, can be estimated by borrowing information across subjects.
  3. Modelling assumptions: allow this sharing of information to happen more efficiently. (See models and methods).

Assumptions for identification: TCE



- Consistency for $Y(x)$:

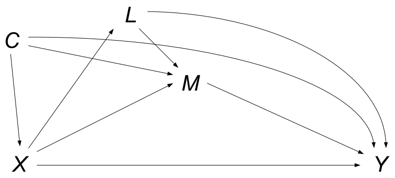
$$Y = Y(x) \text{ if } X = x$$

- Conditional exchangeability given C for X wrt Y:

$$Y(x) \perp\!\!\!\perp X \mid C \quad \forall x$$

- Essentially, this means no unmeasured confounding of the X–Y relationship.

Assumptions for identification: TCE



- Consistency for $Y(x)$:

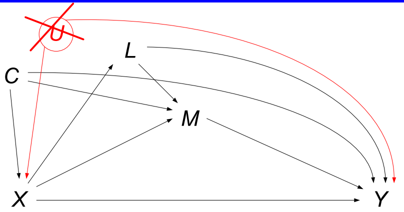
$$Y = Y(x) \text{ if } X = x$$

- Conditional exchangeability given C for X wrt Y:

$$Y(x) \perp\!\!\!\perp X \mid C \quad \forall x$$

- Essentially, this means no unmeasured confounding of the X–Y relationship.

Assumptions for identification: TCE



- Consistency for $Y(x)$:

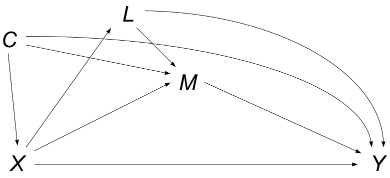
$$Y = Y(x) \text{ if } X = x$$

- Conditional exchangeability given C for X wrt Y:

$$Y(x) \perp\!\!\!\perp X \mid C \quad \forall x$$

- Essentially, this means no unmeasured confounding of the X–Y relationship.

Assumptions for identification: CDE



- Consistency for $Y(x, m)$:

$$Y = Y(x, m) \text{ if } X = x \text{ and } M = m$$

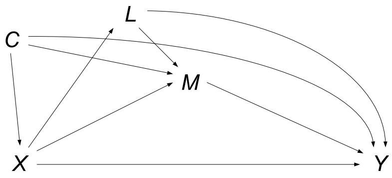
- Sequential conditional exchangeability given C for X wrt Y and given C, X, L for M wrt Y:

$$Y(x) \perp\!\!\!\perp X \mid C \quad \forall x$$

$$Y(x, m) \perp\!\!\!\perp M \mid C, X, L \quad \forall x, m$$

- Essentially, this means no unmeasured confounding of the X–Y or M–Y relationships.

Assumptions for identification: CDE



- Consistency for $Y(x, m)$:

$$Y = Y(x, m) \text{ if } X = x \text{ and } M = m$$

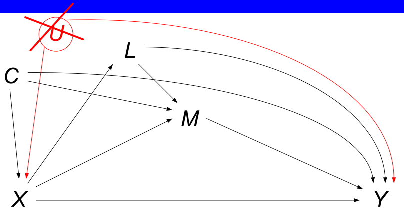
- Sequential conditional exchangeability given C for X wrt Y and given C, X, L for M wrt Y:

$$Y(x) \perp\!\!\!\perp X \mid C \quad \forall x$$

$$Y(x, m) \perp\!\!\!\perp M \mid C, X, L \quad \forall x, m$$

- Essentially, this means no unmeasured confounding of the X–Y or M–Y relationships.

Assumptions for identification: CDE



- Consistency for $Y(x, m)$:

$$Y = Y(x, m) \text{ if } X = x \text{ and } M = m$$

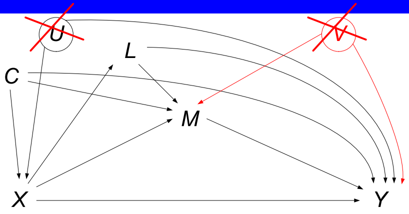
- Sequential conditional exchangeability given C for X wrt Y and given C, X, L for M wrt Y:

$$Y(x) \perp\!\!\!\perp X \,|\, C \quad \forall x$$

$$Y(x, m) \perp\!\!\!\perp M \,|\, C, X, L \quad \forall x, m$$

- Essentially, this means no unmeasured confounding of the X–Y or M–Y relationships.

Assumptions for identification: CDE



- Consistency for $Y(x, m)$:

$$Y = Y(x, m) \text{ if } X = x \text{ and } M = m$$

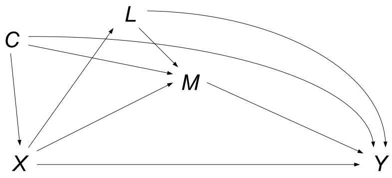
- Sequential conditional exchangeability given C for X wrt Y and given C, X, L for M wrt Y:

$$Y(x) \perp\!\!\!\perp X \mid C \quad \forall x$$

$$Y(x, m) \perp\!\!\!\perp M \mid C, X, L \quad \forall x, m$$

- Essentially, this means no unmeasured confounding of the X–Y or M–Y relationships.

Assumptions for identification: NDE, NIE



- Consistency for $Y(x,m)$, $M(x)$ and $Y\{x,M(x^*)\}$:

$$Y = Y(x,m) \text{ if } X = x \text{ and } M = m, \quad M = M(x) \text{ if } X = x,$$

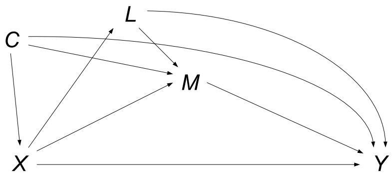
$$Y = Y\{x, M(x^*)\} \text{ if } X = x \text{ and } M = M(x^*).$$

- Sequential conditional exchangeability given C for X wrt Y, given C, X, L for M wrt Y, and given C for X wrt M:

$$Y(x) \perp\!\!\!\perp X | C \quad \forall x, \quad Y(x,m) \perp\!\!\!\perp M | C, X, L \quad \forall x, m,$$

$$M(x) \perp\!\!\!\perp X | C \quad \forall x.$$

Assumptions for identification: NDE, NIE



- Consistency for $Y(x,m)$, $M(x)$ and $Y\{x, M(x^*)\}$:

$$Y = Y(x,m) \text{ if } X = x \text{ and } M = m, \quad M = M(x) \text{ if } X = x,$$

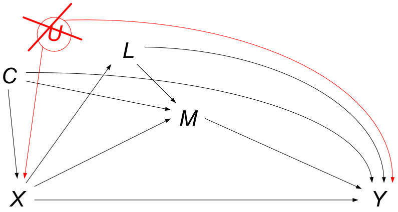
$$Y = Y\{x, M(x^*)\} \text{ if } X = x \text{ and } M = M(x^*).$$

- Sequential conditional exchangeability given C for X wrt Y, given C, X, L for M wrt Y, and given C for X wrt M:

$$Y(x) \perp\!\!\!\perp X \,|\, C \quad \forall x, \quad Y(x,m) \perp\!\!\!\perp M \,|\, C, X, L \quad \forall x, m,$$

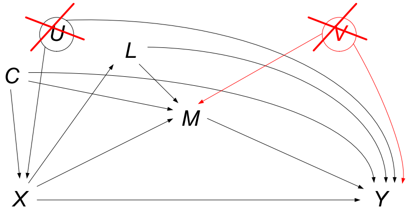
$$M(x) \perp\!\!\!\perp X \,|\, C \quad \forall x.$$

Assumptions for identification: NDE, NIE



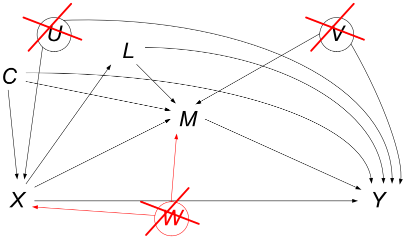
- Essentially, this means no unmeasured confounding of the X–Y, M–Y, or X–M relationships.
- Finally, in addition, either:
  - No intermediate confounding, or
  - Some restriction on the extent to which X and M interact in their effect on Y (Petersen et al. 2006).

Assumptions for identification: NDE, NIE



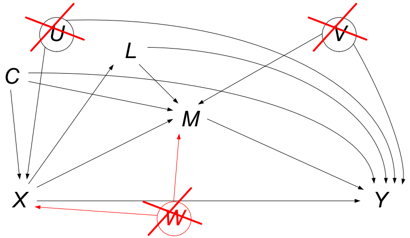
- Essentially, this means no unmeasured confounding of the X–Y, M–Y, or X–M relationships.
- Finally, in addition, either:
  - No intermediate confounding, or
  - Some restriction on the extent to which X and M interact in their effect on Y (Petersen et al. 2006).

Assumptions for identification: NDE, NIE



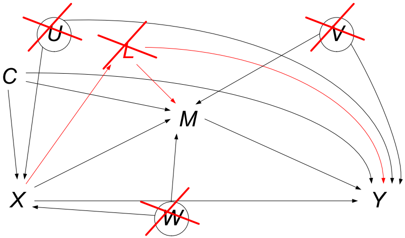
- Essentially, this means no unmeasured confounding of the X–Y, M–Y, or X–M relationships.
- Finally, in addition, either:
  - No intermediate confounding, or
  - Some restriction on the extent to which X and M interact in their effect on Y (Petersen et al. 2006).

Assumptions for identification: NDE, NIE



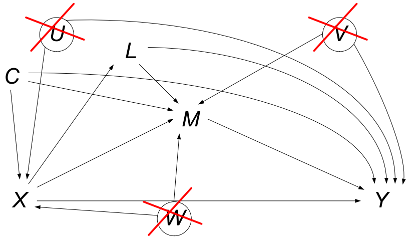
- Essentially, this means no unmeasured confounding of the X–Y, M–Y, or X–M relationships.
- Finally, in addition, either:
  - No intermediate confounding, or
  - Some restriction on the extent to which X and M interact in their effect on Y (Petersen et al, 2006).

Assumptions for identification: NDE, NIE



- Essentially, this means no unmeasured confounding of the X–Y, M–Y, or X–M relationships.
- Finally, in addition, either:
  - No intermediate confounding, or
  - Some restriction on the extent to which X and M interact in their effect on Y (Petersen et al, 2006).

Assumptions for identification: NDE, NIE



- Essentially, this means no unmeasured confounding of the X–Y, M–Y, or X–M relationships.
- Finally, in addition, either:
  - No intermediate confounding, or
  - Some restriction on the extent to which X and M interact in their effect on Y (Petersen et al, 2006).

Outline

G-computation formula for the CDE

Robins 1986

- Let's look at how the CDE is estimated:

$$\begin{aligned}
&\text{CDE}\,(c, x, x^*, m) = \\
&\qquad E\{Y(x^*, m)\,|\,C = c\} - E\{Y(x, m)\,|\,C = c\} \\
&= \int E(Y\,|\,C = c, X = x^*, L = l, M = m)\, f_{L|C,X}(l\,|\,c, x^*)\, dl \\
&- \int E(Y\,|\,C = c, X = x, L = l, M = m)\, f_{L|C,X}(l\,|\,c, x)\, dl
\end{aligned}$$

- This is the g-computation formula.
- It requires correct specification of these parametric associational models for $Y\,|\,C, X, L, M$ and $L\,|\,C, X$.
- Both models can be completely flexible: they can include non-linearities and interactions.
- By marginalising over $L\,|\,C, X$, intermediate confounding is appropriately dealt with.

G-computation formula for the CDE

Robins 1986

- Let's look at how the CDE is estimated:

$$\text{CDE}\,(c, x, x^*, m) = E\left\{Y\left(x^*, m\right)|C = c\right\} - E\left\{Y\left(x, m\right)|C = c\right\}$$

$$= \int E\left(Y\,|C = c, X = x^*, L = l, M = m\right) f_{L|C,X}\left(l\,|c, x^*\right) dl$$

$$- \int E\left(Y\,|C = c, X = x, L = l, M = m\right) f_{L|C,X}\left(l\,|c, x\right) dl$$

- This is the g-computation formula.
- It requires correct specification of these parametric associational models for $Y\,|C, X, L, M$ and $L\,|C, X$.
- Both models can be completely flexible: they can include non-linearities and interactions.
- By marginalising over $L\,|C, X$, intermediate confounding is appropriately dealt with.

G-computation formula for the CDE

Robins 1986

- Let's look at how the CDE is estimated:

$$\text{CDE}(c, x, x^*, m) = E\{Y(x^*, m) | C = c\} - E\{Y(x, m) | C = c\}$$

$$= \int E(Y | C = c, X = x^*, L = l, M = m) f_{L|C,X}(l | c, x^*) \, dl$$

$$- \int E(Y | C = c, X = x, L = l, M = m) f_{L|C,X}(l | c, x) \, dl$$

- This is the g-computation formula.
- It requires correct specification of these parametric associational models for $Y | C, X, L, M$ and $L | C, X$.
- Both models can be completely flexible: they can include non-linearities and interactions.
- By marginalising over $L | C, X$, intermediate confounding is appropriately dealt with.

G-computation formula for the CDE

Robins 1986

- Let's look at how the CDE is estimated:

$$\text{CDE}\,(c, x, x^*, m) = E\left\{Y\left(x^*, m\right)|C=c\right\} - E\left\{Y\left(x, m\right)|C=c\right\}$$

$$= \int E\left(Y\,|C=c, X=x^*, L=l, M=m\right) f_{L|C,X}\left(l\,|c, x^*\right) dl$$

$$- \int E\left(Y\,|C=c, X=x, L=l, M=m\right) f_{L|C,X}\left(l\,|c, x\right) dl$$

- This is the g-computation formula.
- It requires correct specification of these parametric associational models for $Y\,|C, X, L, M$ and $L\,|C, X$.
- Both models can be completely flexible: they can include non-linearities and interactions.
- By marginalising over $L\,|C, X$, intermediate confounding is appropriately dealt with.

G-computation formula for the CDE

Robins 1986

- Let's look at how the CDE is estimated:

$$\text{CDE}(c, x, x^*, m) = E\{Y(x^*, m)|C = c\} - E\{Y(x, m)|C = c\}$$
$$= \int E(Y|C = c, X = x^*, L = l, M = m) f_{L|C,X}(l|c, x^*)\, dl$$
$$- \int E(Y|C = c, X = x, L = l, M = m) f_{L|C,X}(l|c, x)\, dl$$

- This is the g-computation formula.
- It requires correct specification of these parametric associational models for $Y|C, X, L, M$ and $L|C, X$.
- Both models can be completely flexible: they can include non-linearities and interactions.
- By marginalising over $L|C, X$, intermediate confounding is appropriately dealt with.

G-computation formula for the CDE

Robins 1986

- Let's look at how the CDE is estimated:

$$\text{CDE}\,(c, x, x^*, m) = E\{Y(x^*, m) \mid C = c\} - E\{Y(x, m) \mid C = c\}$$

$$= \int E(Y \mid C = c, X = x^*, L = l, M = m)\, f_{L|C,X}(l \mid c, x^*)\, dl$$

$$- \int E(Y \mid C = c, X = x, L = l, M = m)\, f_{L|C,X}(l \mid c, x)\, dl$$

- This is the g-computation formula.
- It requires correct specification of these parametric associational models for $Y \mid C, X, L, M$ and $L \mid C, X$.
- Both models can be completely flexible: they can include non-linearities and interactions.
- By marginalising over $L \mid C, X$, intermediate confounding is appropriately dealt with.

G-computation formula for the CDE

Robins 1986

- In the absence of intermediate confounding, non-linearities and interactions, this becomes α_1 as earlier.
- The g-computation formula thus generalises the earlier approaches to allow felxible modelling, interactions and intermediate confounding.
- The associational models can be estimated using usual regression-fitting techniques (OLS, ML).
- If analytically intractable, the integration over L can be done by Monte Carlo simulation.
- SEs can be obtained either by the delta method or by bootstrapping.
- This can be carried out in Stata (using the `gformula` command).

G-computation formula for the CDE

Robins 1986

- In the absence of intermediate confounding, non-linearities and interactions, this becomes α_1 as earlier.
- The g-computation formula thus generalises the earlier approaches to allow felxible modelling, interactions and intermediate confounding.
- The associational models can be estimated using usual regression-fitting techniques (OLS, ML).
- If analytically intractable, the integration over L can be done by Monte Carlo simulation.
- SEs can be obtained either by the delta method or by bootstrapping.
- This can be carried out in Stata (using the `gformula` command).

G-computation formula for the CDE

Robins 1986

- In the absence of intermediate confounding, non-linearities and interactions, this becomes α_1 as earlier.
- The g-computation formula thus generalises the earlier approaches to allow felxible modelling, interactions and intermediate confounding.
- The associational models can be estimated using usual regression-fitting techniques (OLS, ML).
- If analytically intractable, the integration over L can be done by Monte Carlo simulation.
- SEs can be obtained either by the delta method or by bootstrapping.
- This can be carried out in Stata (using the `gformula` command).

G-computation formula for the CDE

Robins 1986

- In the absence of intermediate confounding, non-linearities and interactions, this becomes α_1 as earlier.
- The g-computation formula thus generalises the earlier approaches to allow felxible modelling, interactions and intermediate confounding.
- The associational models can be estimated using usual regression-fitting techniques (OLS, ML).
- If analytically intractable, the integration over L can be done by Monte Carlo simulation.
- SEs can be obtained either by the delta method or by bootstrapping.
- This can be carried out in Stata (using the `gformula` command).

G-computation formula for the CDE

Robins 1986

- In the absence of intermediate confounding, non-linearities and interactions, this becomes α_1 as earlier.
- The g-computation formula thus generalises the earlier approaches to allow felxible modelling, interactions and intermediate confounding.
- The associational models can be estimated using usual regression-fitting techniques (OLS, ML).
- If analytically intractable, the integration over L can be done by Monte Carlo simulation.
- SEs can be obtained either by the delta method or by bootstrapping.
- This can be carried out in Stata (using the `gformula` command).

G-computation formula for the CDE

Robins 1986

- In the absence of intermediate confounding, non-linearities and interactions, this becomes α_1 as earlier.
- The g-computation formula thus generalises the earlier approaches to allow felxible modelling, interactions and intermediate confounding.
- The associational models can be estimated using usual regression-fitting techniques (OLS, ML).
- If analytically intractable, the integration over L can be done by Monte Carlo simulation.
- SEs can be obtained either by the delta method or by bootstrapping.
- This can be carried out in Stata (using the `gformula` command).

G-computation formula for the NDE/NIE

- The g-computation formula can similarly be used to estimate the NDE and NIE, with further modelling and assumptions.
- A model for $M|C, X, L$ is now required.
- Either there must be no intermediate confounding, or the Petersen et al interaction restriction assumption is required:

$$E\{Y(x^*, m) - Y(x, m)|C = c, M(x) = m\}$$
$$= E\{Y(x^*, m) - Y(x, m)|C = c\}.$$

- This can also be carried out in Stata's `gformula` command.
- Muthén (2011) also shows how this estimation can be done (with almost as much flexibility) in Mplus.

G-computation formula for the NDE/NIE

- The g-computation formula can similarly be used to estimate the NDE and NIE, with further modelling and assumptions.
- A model for $M \mid C, X, L$ is now required.
- Either there must be no intermediate confounding, or the Petersen et al interaction restriction assumption is required:

$$E\left\{Y(x^*, m) - Y(x, m) \mid C = c, M(x) = m\right\} = E\left\{Y(x^*, m) - Y(x, m) \mid C = c\right\}.$$

- This can also be carried out in Stata's `gformula` command.
- Muthén (2011) also shows how this estimation can be done (with almost as much flexibility) in Mplus.

G-computation formula for the NDE/NIE

- The g-computation formula can similarly be used to estimate the NDE and NIE, with further modelling and assumptions.
- A model for $M \mid C, X, L$ is now required.
- Either there must be no intermediate confounding, or the Petersen et al interaction restriction assumption is required:

$$E\left\{Y(x^*, m) - Y(x, m) \mid C = c, M(x) = m\right\} \\ = E\left\{Y(x^*, m) - Y(x, m) \mid C = c\right\}.$$

- This can also be carried out in Stata's `gformula` command.
- Muthén (2011) also shows how this estimation can be done (with almost as much flexibility) in Mplus.

G-computation formula for the NDE/NIE

- The g-computation formula can similarly be used to estimate the NDE and NIE, with further modelling and assumptions.
- A model for $M\,|\,C, X, L$ is now required.
- Either there must be no intermediate confounding, or the Petersen et al interaction restriction assumption is required:

$$E\left\{Y\left(x^{*}, m\right) - Y\left(x, m\right) \middle| C = c, M\left(x\right) = m\right\} = E\left\{Y\left(x^{*}, m\right) - Y\left(x, m\right) \middle| C = c\right\}.$$

- This can also be carried out in Stata's `gformula` command.
- Muthén (2011) also shows how this estimation can be done (with almost as much flexibility) in Mplus.

G-computation formula for the NDE/NIE

- The g-computation formula can similarly be used to estimate the NDE and NIE, with further modelling and assumptions.
- A model for $M \mid C, X, L$ is now required.
- Either there must be no intermediate confounding, or the Petersen et al interaction restriction assumption is required:

$$E\{Y(x^*, m) - Y(x, m) \mid C = c, M(x) = m\} \\ = E\{Y(x^*, m) - Y(x, m) \mid C = c\}.$$

- This can also be carried out in Stata's `gformula` command.
- Muthén (2011) also shows how this estimation can be done (with almost as much flexibility) in Mplus.

Alternative semiparametric approches

- One drawback of the g-computation formula is its fully-parametric nature and hence heavy reliance on parametric modelling assumptions.
- In particular, the necessity to model $L|C,X$ can be problematic if L is high-dimensional.
- Alternative semiparametric methods from the causal inference literature do not require a model for $L|C,X$:
  - inverse probability weighted estimation of a marginal structural model (VanderWeele, 2009),
  - g-estimation of a structural nested model (Robins, 1999),
  - other flavours of g-estimation (Joffe and Greene, 2009; Vansteelandt, 2009).

Alternative semiparametric approches

- One drawback of the g-computation formula is its fully-parametric nature and hence heavy reliance on parametric modelling assumptions.
- In particular, the necessity to model $L \mid C, X$ can be problematic if L is high-dimensional.
- Alternative semiparametric methods from the causal inference literature do not require a model for $L \mid C, X$:
  - inverse probability weighted estimation of a marginal structural model (VanderWeele, 2009),
  - g-estimation of a structural nested model (Robins, 1999),
  - other flavours of g-estimation (Joffe and Greene, 2009; Vansteelandt, 2009).

Alternative semiparametric approches

- One drawback of the g-computation formula is its fully-parametric nature and hence heavy reliance on parametric modelling assumptions.
- In particular, the necessity to model $L \mid C, X$ can be problematic if L is high-dimensional.
- Alternative semiparametric methods from the causal inference literature do not require a model for $L \mid C, X$:
  - inverse probability weighted estimation of a marginal structural model (VanderWeele, 2009),
  - g-estimation of a structural nested model (Robins, 1999),
  - other flavours of g-estimation (Joffe and Greene, 2009; Vansteelandt, 2009).

Alternative semiparametric approches

- One drawback of the g-computation formula is its fully-parametric nature and hence heavy reliance on parametric modelling assumptions.
- In particular, the necessity to model $L \mid C, X$ can be problematic if L is high-dimensional.
- Alternative semiparametric methods from the causal inference literature do not require a model for $L \mid C, X$:
  - inverse probability weighted estimation of a marginal structural model (VanderWeele, 2009),
  - g-estimation of a structural nested model (Robins, 1999),
  - other flavours of g-estimation (Joffe and Greene, 2009; Vansteelandt, 2009).

Alternative semiparametric approches

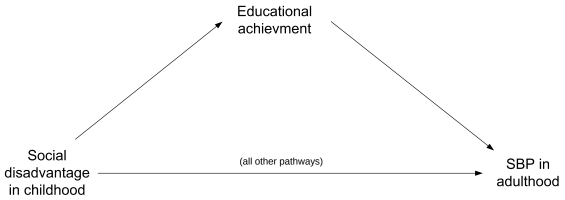
- One drawback of the g-computation formula is its fully-parametric nature and hence heavy reliance on parametric modelling assumptions.
- In particular, the necessity to model $L \mid C, X$ can be problematic if L is high-dimensional.
- Alternative semiparametric methods from the causal inference literature do not require a model for $L \mid C, X$:
  - inverse probability weighted estimation of a marginal structural model (VanderWeele, 2009),
  - g-estimation of a structural nested model (Robins, 1999),
  - other flavours of g-estimation (Joffe and Greene, 2009; Vansteelandt, 2009).

Alternative semiparametric approches

- One drawback of the g-computation formula is its fully-parametric nature and hence heavy reliance on parametric modelling assumptions.
- In particular, the necessity to model $L \mid C, X$ can be problematic if L is high-dimensional.
- Alternative semiparametric methods from the causal inference literature do not require a model for $L \mid C, X$:
  - inverse probability weighted estimation of a marginal structural model (VanderWeele, 2009),
  - g-estimation of a structural nested model (Robins, 1999),
  - other flavours of g-estimation (Joffe and Greene, 2009; Vansteelandt, 2009).

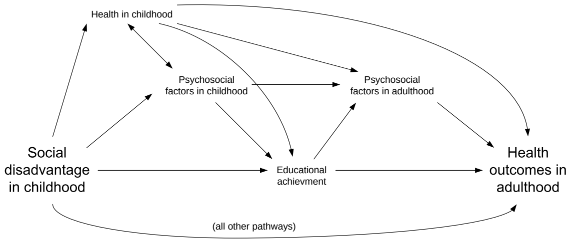
Outline

Multiple mediators



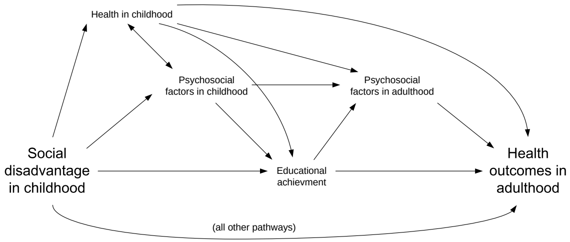
- How can we take all this discussion of direct and indirect effects, and apply it to the complicated multiple pathway setting we started with?
- The definitions of direct and indirect effects extend easily to path-specific effects.
- But defining effects that sum to the total effect is not quite as trivial (Daniel and De Stavola, submitted).

Multiple mediators



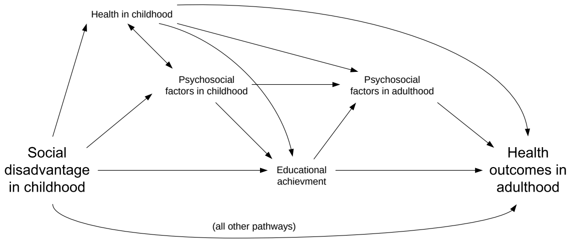
- How can we take all this discussion of direct and indirect effects, and apply it to the complicated multiple pathway setting we started with?
- The definitions of direct and indirect effects extend easily to path-specific effects.
- But defining effects that sum to the total effect is not quite as trivial (Daniel and De Stavola, submitted).

Multiple mediators



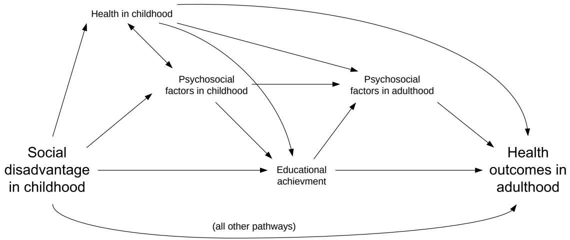
- How can we take all this discussion of direct and indirect effects, and apply it to the complicated multiple pathway setting we started with?
- The definitions of direct and indirect effects extend easily to path-specific effects.
- But defining effects that sum to the total effect is not quite as trivial (Daniel and De Stavola, submitted).

Multiple mediators



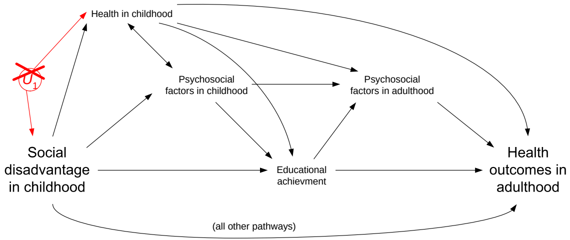
- How can we take all this discussion of direct and indirect effects, and apply it to the complicated multiple pathway setting we started with?
- The definitions of direct and indirect effects extend easily to path-specific effects.
- But defining effects that sum to the total effect is not quite as trivial (Daniel and De Stavola, submitted).

Multiple mediators



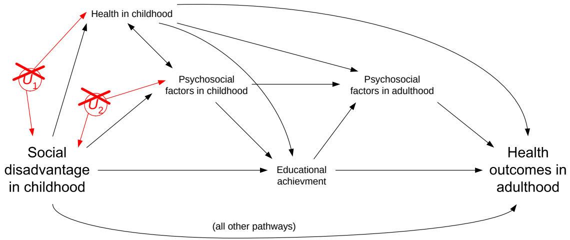
- Identification assumptions are along the same lines, but require more and more assumptions of no unmeasured confounding!
- And, whenever pathways 'interfere', intermediate confounding is inevitable.
- Thus assumptions abound, and sensitivity analyses are paramount (George: next talk!).

Multiple mediators



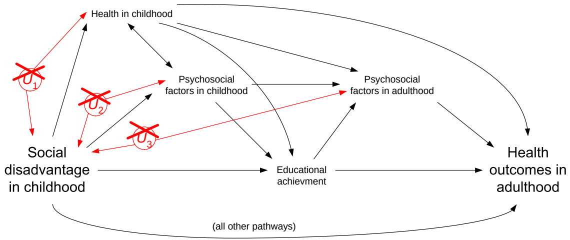
- Identification assumptions are along the same lines, but require more and more assumptions of no unmeasured confounding!
- And, whenever pathways 'interfere', intermediate confounding is inevitable.
- Thus assumptions abound, and sensitivity analyses are paramount (George: next talk!).

Multiple mediators



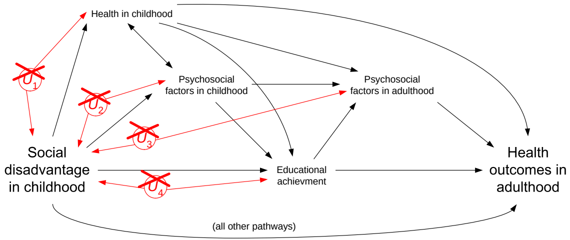
- Identification assumptions are along the same lines, but require more and more assumptions of no unmeasured confounding!
- And, whenever pathways 'interfere', intermediate confounding is inevitable.
- Thus assumptions abound, and sensitivity analyses are paramount (George: next talk!).

Multiple mediators



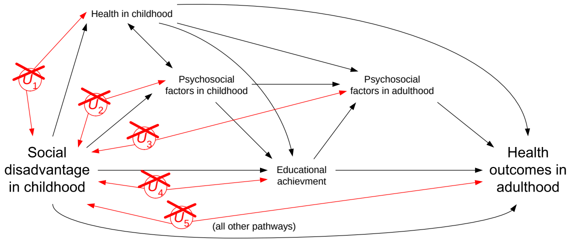
- Identification assumptions are along the same lines, but require more and more assumptions of no unmeasured confounding!
- And, whenever pathways 'interfere', intermediate confounding is inevitable.
- Thus assumptions abound, and sensitivity analyses are paramount (George: next talk!).

Multiple mediators



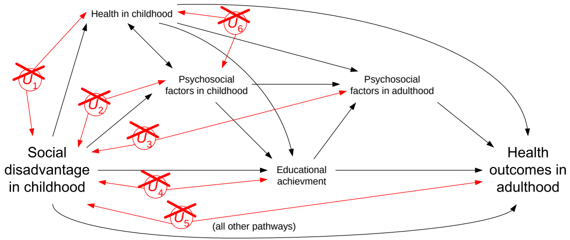
- Identification assumptions are along the same lines, but require more and more assumptions of no unmeasured confounding!
- And, whenever pathways 'interfere', intermediate confounding is inevitable.
- Thus assumptions abound, and sensitivity analyses are paramount (George: next talk!).

Multiple mediators



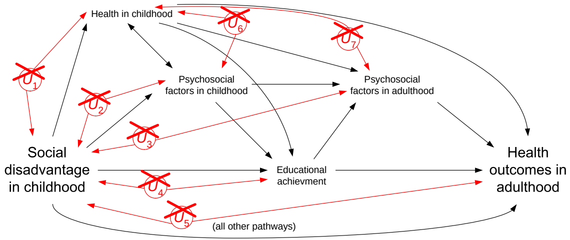
- Identification assumptions are along the same lines, but require more and more assumptions of no unmeasured confounding!
- And, whenever pathways 'interfere', intermediate confounding is inevitable.
- Thus assumptions abound, and sensitivity analyses are paramount (George: next talk!).

Multiple mediators



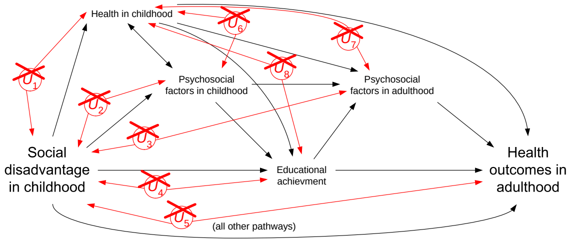
- Identification assumptions are along the same lines, but require more and more assumptions of no unmeasured confounding!
- And, whenever pathways 'interfere', intermediate confounding is inevitable.
- Thus assumptions abound, and sensitivity analyses are paramount (George: next talk!).

Multiple mediators



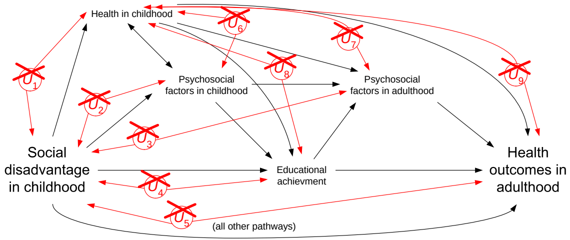
- Identification assumptions are along the same lines, but require more and more assumptions of no unmeasured confounding!
- And, whenever pathways 'interfere', intermediate confounding is inevitable.
- Thus assumptions abound, and sensitivity analyses are paramount (George: next talk!).

Multiple mediators



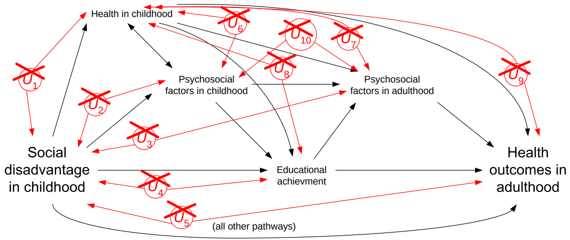
- Identification assumptions are along the same lines, but require more and more assumptions of no unmeasured confounding!
- And, whenever pathways 'interfere', intermediate confounding is inevitable.
- Thus assumptions abound, and sensitivity analyses are paramount (George: next talk!).

Multiple mediators



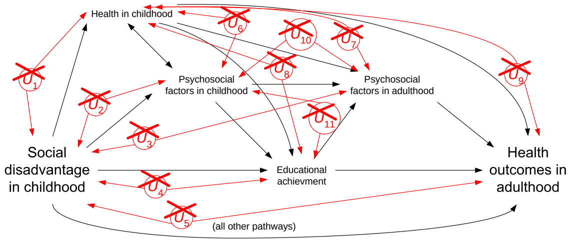
- Identification assumptions are along the same lines, but require more and more assumptions of no unmeasured confounding!
- And, whenever pathways 'interfere', intermediate confounding is inevitable.
- Thus assumptions abound, and sensitivity analyses are paramount (George: next talk!).

Multiple mediators



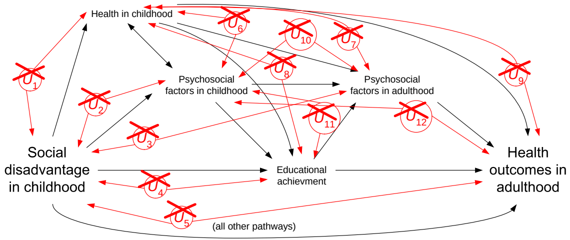
- Identification assumptions are along the same lines, but require more and more assumptions of no unmeasured confounding!
- And, whenever pathways 'interfere', intermediate confounding is inevitable.
- Thus assumptions abound, and sensitivity analyses are paramount (George: next talk!).

Multiple mediators



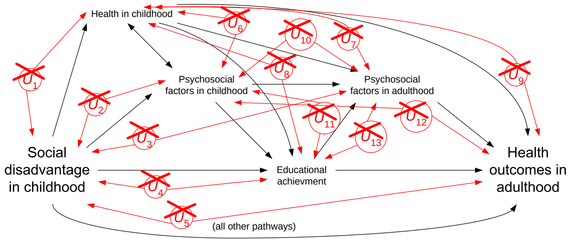
- Identification assumptions are along the same lines, but require more and more assumptions of no unmeasured confounding!
- And, whenever pathways 'interfere', intermediate confounding is inevitable.
- Thus assumptions abound, and sensitivity analyses are paramount (George: next talk!).

Multiple mediators



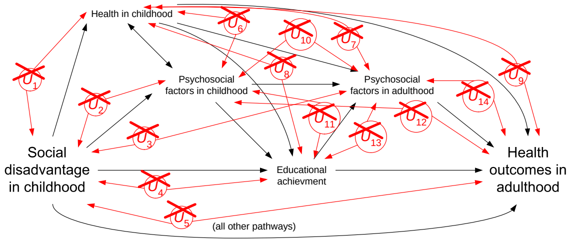
- Identification assumptions are along the same lines, but require more and more assumptions of no unmeasured confounding!
- And, whenever pathways 'interfere', intermediate confounding is inevitable.
- Thus assumptions abound, and sensitivity analyses are paramount (George: next talk!).

Multiple mediators



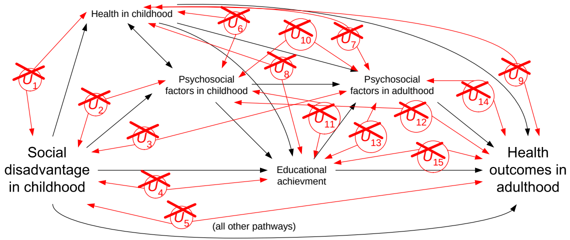
- Identification assumptions are along the same lines, but require more and more assumptions of no unmeasured confounding!
- And, whenever pathways 'interfere', intermediate confounding is inevitable.
- Thus assumptions abound, and sensitivity analyses are paramount (George: next talk!).

Multiple mediators



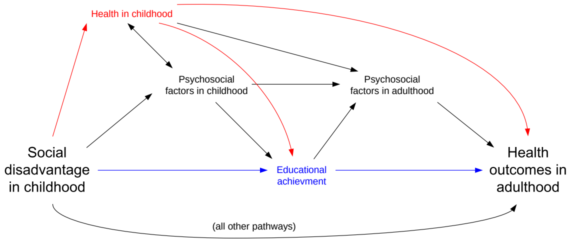
- Identification assumptions are along the same lines, but require more and more assumptions of no unmeasured confounding!
- And, whenever pathways 'interfere', intermediate confounding is inevitable.
- Thus assumptions abound, and sensitivity analyses are paramount (George: next talk!).

Multiple mediators



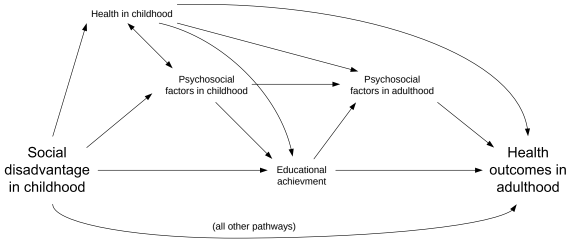
- Identification assumptions are along the same lines, but require more and more assumptions of no unmeasured confounding!
- And, whenever pathways 'interfere', intermediate confounding is inevitable.
- Thus assumptions abound, and sensitivity analyses are paramount (George: next talk!).

Multiple mediators



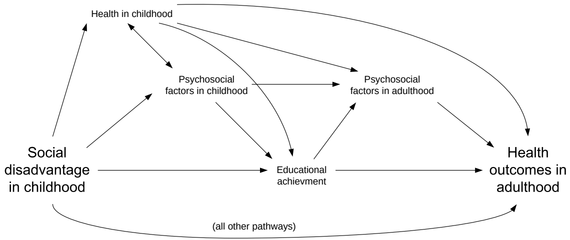
- Identification assumptions are along the same lines, but require more and more assumptions of no unmeasured confounding!
- And, whenever pathways 'interfere', intermediate confounding is inevitable.
- Thus assumptions abound, and sensitivity analyses are paramount (George: next talk!).

Multiple mediators



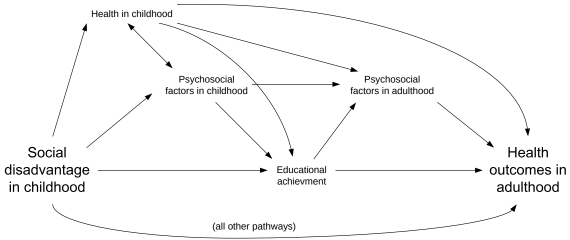
- Identification assumptions are along the same lines, but require more and more assumptions of no unmeasured confounding!
- And, whenever pathways 'interfere', intermediate confounding is inevitable.
- Thus assumptions abound, and sensitivity analyses are paramount (George: next talk!).

Multiple mediators



- The same methods can be used.
- The extensions are straightforward but need implementing in software.

Multiple mediators



- The same methods can be used.
- The extensions are straightforward but need implementing in software.

Outline

Quick summary

- The ‘old’ approaches are somewhat limited, and too vague for these limitations to be always apparent.
- Newer contributions have led to more hygienic thinking on these issues and more flexible methods.
- But there can be no panacea.
- Very strong assumptions are required for such an ambitious causal endeavour.
- But these (and more) were needed in the ‘old’ approach even if we didn’t realise it.
- Hygienic thinking keeps us honest, and aids sensitivity analyses. . .

Quick summary

- The ‘old’ approaches are somewhat limited, and too vague for these limitations to be always apparent.
- Newer contributions have led to more hygienic thinking on these issues and more flexible methods.
- But there can be no panacea.
- Very strong assumptions are required for such an ambitious causal endeavour.
- But these (and more) were needed in the ‘old’ approach even if we didn’t realise it.
- Hygienic thinking keeps us honest, and aids sensitivity analyses. . .

Quick summary

- The 'old' approaches are somewhat limited, and too vague for these limitations to be always apparent.
- Newer contributions have led to more hygienic thinking on these issues and more flexible methods.
- But there can be no panacea.
- Very strong assumptions are required for such an ambitious causal endeavour.
- But these (and more) were needed in the 'old' approach even if we didn't realise it.
- Hygienic thinking keeps us honest, and aids sensitivity analyses. . .

Quick summary

- The 'old' approaches are somewhat limited, and too vague for these limitations to be always apparent.
- Newer contributions have led to more hygienic thinking on these issues and more flexible methods.
- But there can be no panacea.
- Very strong assumptions are required for such an ambitious causal endeavour.
- But these (and more) were needed in the 'old' approach even if we didn't realise it.
- Hygienic thinking keeps us honest, and aids sensitivity analyses. . .

Quick summary

- The ‘old’ approaches are somewhat limited, and too vague for these limitations to be always apparent.
- Newer contributions have led to more hygienic thinking on these issues and more flexible methods.
- But there can be no panacea.
- Very strong assumptions are required for such an ambitious causal endeavour.
- But these (and more) were needed in the ‘old’ approach even if we didn’t realise it.
- Hygienic thinking keeps us honest, and aids sensitivity analyses. . .

Quick summary

- The 'old' approaches are somewhat limited, and too vague for these limitations to be always apparent.
- Newer contributions have led to more hygienic thinking on these issues and more flexible methods.
- But there can be no panacea.
- Very strong assumptions are required for such an ambitious causal endeavour.
- But these (and more) were needed in the 'old' approach even if we didn't realise it.
- Hygienic thinking keeps us honest, and aids sensitivity analyses. . .

Outline

References

Baron, R. M. and Kenny, D. A. (1986)
The moderator-mediator variable distinction in social psychological research: conceptual, strategic and statistical considerations
Journal of Personality and Social Psychology, 51:1173–1182.

Wright, S. (1921)
Correlation and causation: Part I - Method of Path Coefficients
Journal of Agriculture Research, 20:557-575.

Pearl, J. (2001)
Direct and indirect effects
Proceedings of the 17th Conference in Uncertainty in Artificial Intelligence, 411-420.

References

Robins, J. M. and Greenland, S. (1992):
Identifiability and exchangeability for direct and indirect effects
Epidemiology, 3:143-155.

VanderWeele, T. J. and Vansteelandt, S. (2010)
Odds ratios for mediation analysis for a dichotomous outcome
American Journal of Epidemiology, 172:1339-1348.

Petersen, M. L., Sinisi, S. E. and van der Laan, M. J. (2006)
Estimation of direct causal effects
Epidemiology, 17:276-284.

References

Robins, J. M. (1986)
A new approach to causal inference in mortality studies with a sustained exposure period - application to control of the healthy worker survivor effect
Mathematical Modelling, 7:1393-1512.

Daniel, R. M., De Stavola, B. L. and Cousens, S. N. (2011)
gformula: Estimating causal effects in the presence of time-varying confounding or mediation using the g-computation formula
The Stata Journal, 11:479-517.

Muthén, B. (2011)
Applications of causally defined direct and indirect effects in mediation analysis using SEM in Mplus.
Technical report, Graduate School of Education and Information Studies, University of California, Los Angeles, CA.

References

- Ten Have, T. R. and Joffe, M. (2012)
  A review of causal estimation of effects in mediation analyses
  Statistical Methods in Medical Research, 21:77-107.
- Imai, K., Keele, L. and Yamamoto, T. (2010)
  Identification, inference, and sensitivity analysis for causal mediation effects
  Statistical Science, 25:51-71.
- Daniel, R. M. and De Stavola, B. L. (submitted)
  Causal mediation estimands that permit decomposition with multiple causally-ordered mediators.