Causal Inference

Vanessa Didelez and Robin Evans
BIPS, University of Bremen (Germany), and University of Oxford (UK)

September 2024

APTS — Oxford

Appendix I:

Causal Mediation Analysis

Notation

$Y =$ outcome

$M =$ mediating variable(s)

$A =$ exposure / treatment

$Y(a, m)$ potential response under intervention in A and M

cf. $p(y \mid \text{do}(A = a, M = m))$

$M(a)$ pot. outcome of mediator under intervention in A

Mediational Research Questions

Setting: ‘important’ events occur between exposure and outcome $\Rightarrow$ want to understand their (causal) role

Causal mediation:
– *very* special & strange estimand
– often no target trial possible (not even hypothetically)

$\Rightarrow$ Must understand meaning / assumptions to decide if causal mediational / (in)direct effects relevant to question at hand!

Mediational Research Questions

Example: Randomised placebo-controlled trial

Wanted: effect of a new drug over and above the placebo effect; i.e. want the 'direct' effect of the drug, not its indirect effect via 'patient's (or doctor's) expectation'.

Note: in such a trial, we investigate the target of inference, the direct effect, *by design*.

Can use similar ideas to investigate indirect placebo effect.

Often, such trials not possible

$\Rightarrow$ need suitable assumptions and methods.

Example: Attitudes to immigration

Typical social science experiment: (Brader et al., 2008)

$A =$ exposure (randomised) to new report emphasising positive ($A = 0$) or negative ($A = 1$) aspects of immigration

$M =$ anxiety, measured via questionnaire (quasi-continuous scale)

$Y =$ feelings towards immigration ($0 =$ pro, $1 =$ con)

$X =$ typical covariates: gender, age, income, education etc.

Research question: 'role' of anxiety in translating 'information' into political attitude?

Aside

Note:

‘The’ direct or ‘the’ indirect effect do not exist...

– always relative to the (set of) mediator(s) considered.

– even with given mediators, may depend on other choices.

Terminology: Background

- Traditionally (in some fields): mediation = path analysis, based on linear structural equation model (LSEMs).
- Advantage: LSEMs simple parameterisation with (apparently) 'intuitive' meaning of parameters in terms of direct effects.
- Disadvantage: LSEMs overly simplistic, do not carry over to non-linear settings (e.g. interactions, binary variables,...).

Terminology: Background

'Non-parametric' definition of (in)direct effects:

Wanted: notions of (in)direct effects that do not pre-suppose a certain parametric model.

$\Rightarrow$ 'target trial' for target of inference, e.g. placebo-controlled

$\Rightarrow$ & use do$(\cdot)$ or potential responses to define our target!

Controlled Direct Effect?

$$CDE = E(Y|\text{do}(A = 1, M = 0)) - E(Y|\text{do}(A = 0, M = 0))$$

Causal effect of A on Y while *intervening* to hold M constant at baseline ($M = 0$).

Advantage: CDE conceptually simple; identifying conditions straightforward; can be related to parameters of variety of regression models; will suffice in many applications.

Disadvantage: no corresponding notion of indirect effect
— in fact: M could be prior / post A or both could be independent of each other with same CDE.
$\Rightarrow$ does not fully capture what we might mean by 'mediation'.

‘Natural’ (In)Direct Effects

Motivation

In placebo trial, M is not controlled

$\rightarrow$ instead ‘pretend’ A has different value:

control (placebo) group will think they receive treatment, but they do not receive active ingredient.

$\Rightarrow$ mediator is $M(a')$, while actual treatment is different $A = a$.

Natural (In)Direct Effects

Definition (Robins & Greenland, 1992; Pearl 2001)

$$NDE = E(Y(a', M(a')) - Y(a, M(a')))$$

$$NIE = E(Y(a, M(a')) - Y(a, M(a)))$$

Or: other contrasts, e.g. relative risks.

Note: NDE, NIE can be different if a, a' reversed
— interactions!

Effect decomposition

Assuming only consistency; no particular parametric model.

Total effect $=$

$$\begin{aligned} E(Y(a') - Y(a)) &= E(Y(a', M(a')) - Y(a, M(a))) \\ &= E(Y(a', M(a')) - Y(a, M(a'))) \\ &\quad + E(Y(a, M(a')) - Y(a, M(a))) \\ &= NDE + NIE \end{aligned}$$

Interactions

Note: if (outcome) model non-linear / with interactions, typically:

$$\underbrace{E(Y(1, M(1)) - Y(0, M(1)))}_{total\ DE\ (NDE)} \neq \underbrace{E(Y(1, M(0)) - Y(0, M(0)))}_{pure\ DE}$$

and similar for indirect effects.

Nested Counterfactual

Key quantity: nested counterfactual $Y(a, M(a'))$

In words: the outcome Y we would observe if exposure were set to a while the mediator be set to the value it would take under exposure setting a'

— genuinely *counterfactual* ('cross-world',
cf. Andrews & Didelez, 2021)

Separable Effects

Re-interpretation of nested counterfactuals

...in terms of do$(\cdot)$ based on extended model:

Assume A can be separated into an aspect A^M affecting only M and another aspect A^Y affecting only Y:

$\Rightarrow$ target of inference $E(Y \mid \text{do}(A^Y = a, A^M = a'))$.

(Robins & Richardson, 2011; Didelez, 2019)

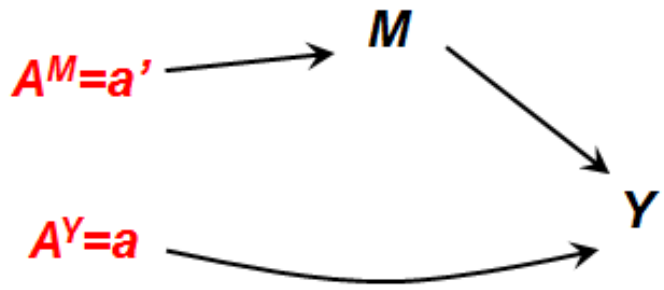
Separable Effects

$\Rightarrow$ Can make sense of $Y(a, M(a'))$ in terms of augmented system (DAG) and do-interventions

Target trial: e.g. placebo controlled trial,

$A^M =$ awareness of receiving treatment

$A^Y =$ actual receiving active ingredient



Observational data: always $A \equiv A^M \equiv A^Y$; identification?

(Robins & Richardson, 2011; Didelez, 2019)

Mediational G-Formula

X observed covariates, not affected by A or M (non-descendants)

Under identifying assumptions:

$$E(Y(a, M(a')) \mid x) = \sum_m E(Y \mid A = a, M = m, x) \times p(m \mid A = a', x)$$

(or marginalise over X)

NDE/NIE: Identifying Assumptions

As before: consistency, positivity

No unmeasured confounding

$$Y(a,m) \perp\!\!\!\perp A \mid X, \qquad M(a) \perp\!\!\!\perp A \mid X,$$

$$Y(a,m) \perp\!\!\!\perp M(a) \mid (A = a, X)$$

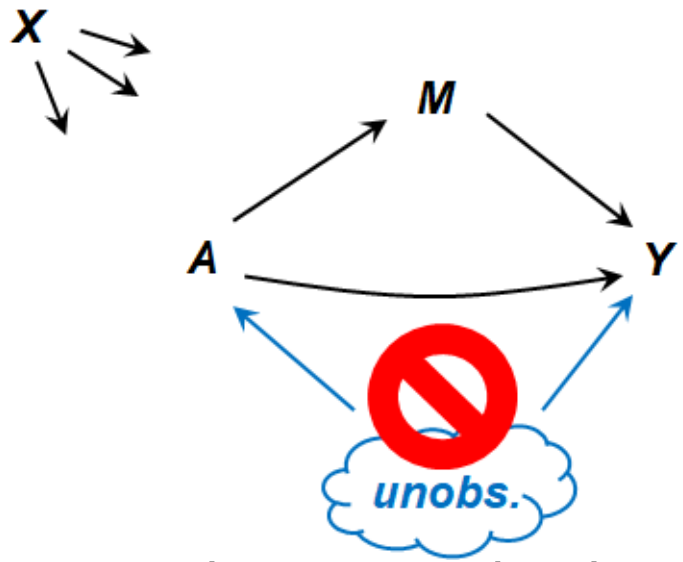
Cross-world independence

$$Y(a,m) \perp\!\!\!\perp M(a') \mid X$$

Or: assume extended causal DAG with separable effects.

Key Assumptions – Graphically

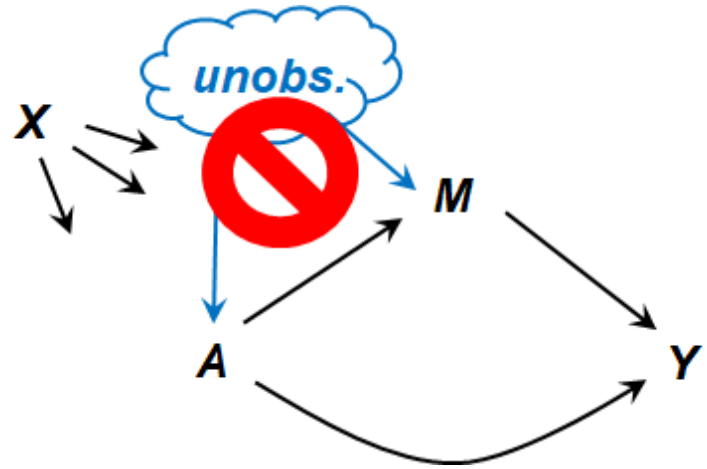
No unobserved A-Y confounding given X, i.e. $Y(a, m) \perp\!\!\!\perp A \mid X$:



Note: automatically true when A randomised.

Key Assumptions – Graphically

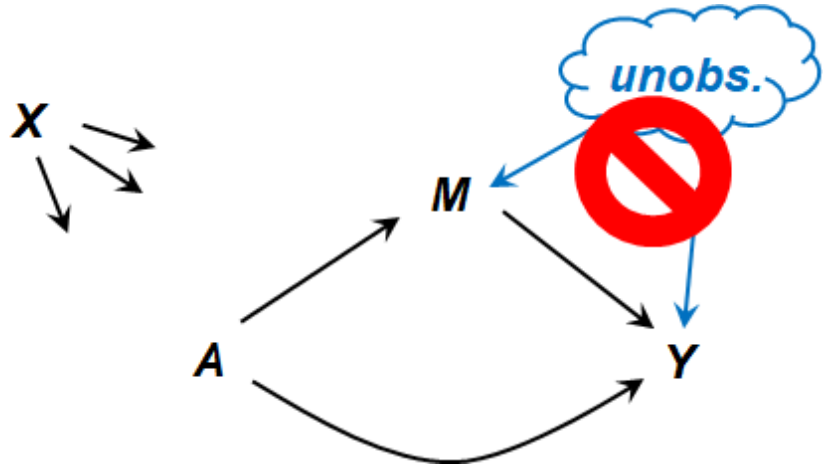
No unobserved A-M confounding given X, i.e. $M(a) \perp\!\!\!\perp A \mid X$:



Note: automatically true when A randomised.

Key Assumptions – Graphically

No unobserved M-Y confounding given X, i.e.
$Y(a, m) \perp\!\!\!\perp M \mid (A = a, X)$:

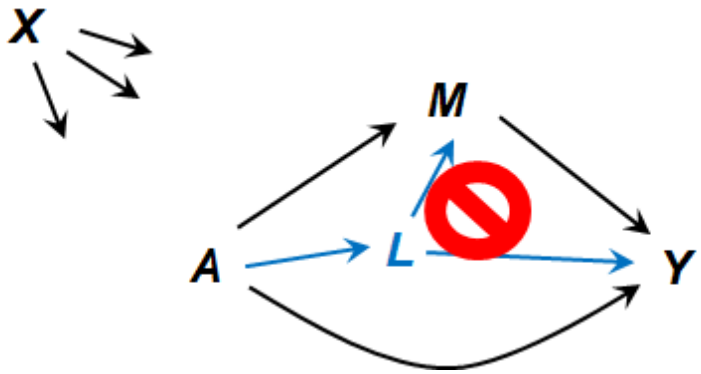


Note: *NOT* automatically true even when A randomised!
Cannot randomise M in same experiment.

Key Assumptions – Graphically

Cross-world independence: $Y(a, m) \perp\!\!\!\perp M(a') \mid X$

e.g. no treatment-induced M-Y confounding by some L, observed nor unobserved!

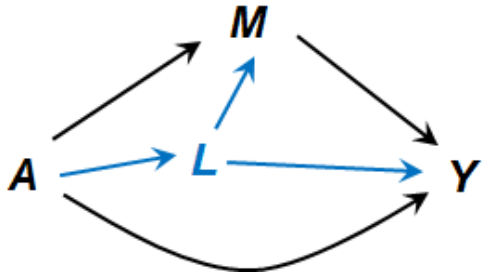


Note: Cannot be verified in ANY experiment!

Treatment-Induced Confounding

of M and Y

Why is treatment-induced confounding a problem?



$$Y(a, M(a')) = Y(a, L(a), M(a', L(a')))$$

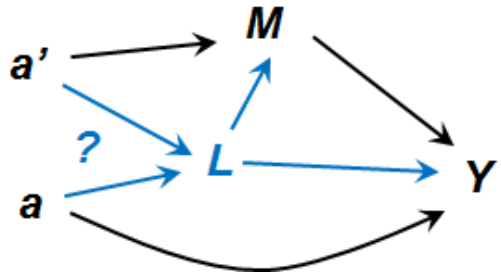
$\Rightarrow$ no empirical joint information on $(L(a), L(a'))$!

Note: under LSEM, problem resolved by assumption of *constant individual-level* effects.

But: under NPSEM-IE, problem only avoided when no treatment-induced confounding.

Treatment-Induced Confounding

Why is treatment-induced confounding a problem?



$$Y(a, M(a')) = Y(a, L(a), M(a', L(a')))$$

$\Rightarrow$ separation of paths due to L unclear

L also called 'recanting witness' (Avin et al, 2005)

Target of inference may not be meaningful / of any practical relevance. Instead: methods for multiple mediators.

Approaches to Inference

(1) For certain parametric models, analytic expressions for NDE and NDE can be derived, e.g. LSEM, or see VanderWeele (2015)

(2) Fit 'pieces' of mediational g-formula and plug-in or use MC-methods

$\Rightarrow$ R package `mediation` by Imai et al (2010)
see also *Stata* Command `gformula` Daniel et al. 2011

Approaches to Inference

(3) Specify model for $E(Y(a, M(a')))$ with explicit parameters for direct / indirect effect, possibly with interaction effect (use suitable / desired link function); fitting requires 'imputing' of missing information using auxiliary (working) models for either mediator or outcome;

$\Rightarrow$ R package `medflex` (Steen et al., 2017)

(4) Other more robust approaches exist but are complicated to implement (Tchetgen Tchetgen & Shpitser, 2012).

Linear SEMs (LSEMs)

Reminder: SEMs — assignments assumed invariant to how input comes about.

$\Rightarrow$ can generate joint distribution on all potential responses.

Now, functional dependence **linear** in inputs.

Background on LSEM

$\mathbf{Y} = (Y_1, \ldots, Y_K)$ set of endogenous variables

$\mathbf{A} = (A_1, \ldots, A_L)$ set of exogenous variables

General structure: (Bollen, 1989)

$$\mathbf{Y} = B\mathbf{Y} + \Gamma\mathbf{A} + \xi$$

B, Γ conformable matrices of parameters (coefficients)

$\xi =$ noise, $\xi \perp\!\!\!\perp \mathbf{A}$

Endogenous: (interrelated) outcomes we are interested in

Exogenous: fixed by design, randomised or always conditioned

Background on LSEM

$$\mathbf{Y} = B\mathbf{Y} + \Gamma\mathbf{A} + \xi$$

If B lower triangular $\Rightarrow$ representable by DAG on $(Y_1, \ldots, Y_K)$

If $\Psi = Var(\xi)$ diag. $\Rightarrow$ causal sufficiency / no unobserved conf.

If both $\Rightarrow$ recursive model.

Further, let $\Phi = Var(\mathbf{A})$.

Background on LSEM

$$\mathbf{Y} = B\mathbf{Y} + \Gamma\mathbf{A} + \xi$$

Identification:

place restrictions on B, Γ, Ψ, Φ so that unique solutions in terms of $\Sigma = Var(\mathbf{Y})$ exist.

$\Rightarrow$ every recursive model is identified.

Various sufficient rules for other models.

Generally no necessary & sufficient rules (Drton, 2016).

Background on LSEM

LSEM encompass

- path analyses
- measurement error models
- measurement models for latent constructs (e.g. IQ)
- growth curves
- factor analyses
- instrumental variables $\rightarrow$ later.

Causal Mediation and LSEMs

Assume simple LSEM:

$$M = \beta_0 + \beta_1 A + \beta_2 X + \epsilon_M$$

$$Y = \theta_0 + \theta_1 A + \theta_2 M + \theta_3 X + \epsilon_Y$$

Hence:

$$Y(a, M(a')) = \theta_0 + \theta_1 a + \theta_2 (\underbrace{\beta_0 + \beta_1 a' + \beta_2 X + \epsilon_M}_{M(a')}) + \theta_3 X + \epsilon_Y$$

re-arranging:

$$Y(a, M(a')) = \underbrace{\theta_0 + \theta_2\beta_0}_{\text{const.}} + \theta_1 x + \theta_2\beta_1 x' + \underbrace{(\theta_2\beta_2 + \theta_3)}_{\text{coeff. of } X} X + \underbrace{\theta_2\epsilon_M + \epsilon_Y}_{\text{noise}}$$

$\Rightarrow NDE$ will be in terms of θ_1, NIE in terms of $\theta_2\beta_1$

Causal Mediation and LSEMs

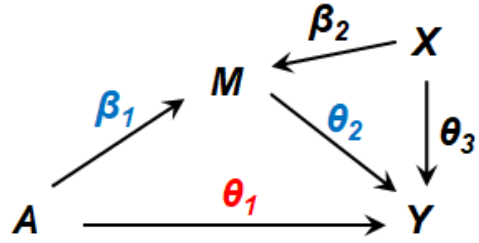
$$Y(a, M(a')) = \underbrace{\theta_0 + \theta_2\beta_0}_{\text{const.}} + \theta_1 a + \theta_2\beta_1 a' + \underbrace{(\theta_2\beta_2 + \theta_3)}_{\text{coeff. of } X} X + \underbrace{\theta_2\epsilon_M + \epsilon_Y}_{\text{noise}}$$

$\Rightarrow$ path-tracing formula

known from Baron & Kenny (1986)

total effect: $\theta_1 + \beta_1\theta_2$.

Generalises to more complex graphs.



Limitations of LSEMs

Simplicity breaks down when using more complex models, e.g. when

$$Y = \theta_0 + \theta_1 A + \theta_2 M + \theta^* AM + \theta_3 X + \epsilon_Y$$

Then $Y(a, M(a'))$ = const. + noise...

$$+(\theta_1 + \theta^* \beta_0)a + \theta_2 \beta_1 a' + \underbrace{\theta^* \beta_1 a a'}_{\text{interact.}} + (\theta_2 \beta_2 + \theta_3)X + \underbrace{(\theta^* \beta_2) a X}_{\text{interact.}}$$

Limitations of LSEMs

Assume M or Y or both binary:

LSEM not sensible (does not constrain $M, Y \in \{0, 1\}$).

Instead: try e.g. logistic for each of $p(m|a, x)$ and $p(y|m, a, x)$

$\Rightarrow$ *NO simple* (logistic) model for $E(Y(a, M(a')))$!

Using Mediational G-Formula

Reminder:

$$E(Y(a, M(a')) \mid X = x) = \sum_m E(Y \mid A = a, M = m, x) \times p(m \mid A = a', x)$$

Idea: assume parametric models for $E(Y \mid a, m, x)$ and $p(m \mid a', x)$ and combine.

Inference: bootstrap, or MC based on sampling distributions of parameters of both models.

$\Rightarrow$ reliance on correct specification of both models.

(Imai et al, 2010; Daniel et al, 2011)

Mediational G-Formula – Example

Example:

Attitudes to immigration (Brader et al, 2008; Tingley et al, 2014)

`treat`= news report on pos/neg aspects of immigration;
`anxiety`= anxiety (on scale 1-4);
`immigr bin`= attitude towards immigration (binary: pro/con);

$\Rightarrow$ linear model $p(m|a,x)$, logistic model $p(y|m,a,x)$

```
imai_m <- lm(anxiety ~ treat + gender + age + educ + income,
             data=framing)
imai_y <- glm(immigr_bin ~ treat + anxiety + gender + age + educ + income,
              family = binomial(link="logit"),
              data=framing)
```

Mediational G-Formula – Example

Output: mean differences!

```
## Nonparametric Bootstrap Confidence Intervals with the Percentile Method
##
##                            Estimate 95% CI Lower 95% CI Upper p-value
## ACME (control)             0.069929     0.031781          0.12   0.002 **
## ACME (treated)             0.053625     0.020445          0.10   0.002 **
## ADE (control)              0.125458     0.000975          0.24   0.050 *
## ADE (treated)              0.109155     0.000878          0.21   0.050 *
## Total Effect               0.179083     0.066660          0.28   0.008 **
## Prop. Mediated (control)   0.390481     0.162717          0.96   0.006 **
## Prop. Mediated (treated)   0.299444     0.101226          0.95   0.006 **
## ACME (average)             0.061777     0.026817          0.10   0.002 **
## ADE (average)              0.117306     0.000927          0.22   0.050 *
## Prop. Mediated (average)   0.344962     0.134809          0.95   0.006 **
```

Suggests: a considerable proportion of the effect of immigration reporting on attitude is mediated by anxiety.

Some indication for treatment-mediator interaction.

Notes on `mediation`

(Tingley et al, 2014)

- Allows for survival outcomes
- Includes tools for sensitivity analysis
- Only outputs mean-differences
- Nothing to prevent *g-null paradox*...

G-Null Paradox

(Robins & Wasserman, 1997)

Note:

choice of models for $p(y|a, m, x)$ and $p(m|a, x)$ will implicitly restrict $E(Y(a, M(a')))$.

Example: Combine linear (for Y) and logistic regression (for M)

$\Rightarrow$ total effect can only be zero if *both* NDE and NIE are zero

— there is no canceling out of NDE and NIE possible.

$\Rightarrow$ might inadvertently impose undesirable restrictions!

Natural Effects Models

(Lange et al, 2012)

Model for $E(Y(a, M(a')))$ (or suitable link-function), e.g.

$$E(Y(a, M(a'))) = \eta_0 + \eta_1 a + \eta_2 a' \qquad a, a' \in \mathcal{A}$$

or conditional on baseline covariates X

$$E(Y(a, M(a'))|X = x) = \eta_0 + \eta_1 a + \eta_2 a' + \eta_3 x$$

$\Rightarrow \eta_1, \eta_2$ explicit parameters for direct/indirect effects.

We never observe *different* values a, a' together, so how on Earth should we ever be able to fit such a model???

Fitting NE Models (1)

First trick: note that expectation is wrt.

$$p(y|a,m,x)p(m|a',x) = p(y,m|a,x)\frac{p(m|a',x)}{p(m|a,x)}$$

$\Rightarrow$ 'clone' observations with $A = a$, assign $A = a'$ and give weight

$$\text{weight} = \frac{p(m \mid A = a', X)}{p(m \mid A = a, X)}$$

obtained from separate model for $p(m \mid a, x)$.

$\Rightarrow$ extended data set $\Rightarrow$ can consistently estimate η's providing $p(m|a,x)$-model correctly specified.

NE Models – Reweighting

with `medflex` ***(Steen et al, 2017)***

Anxiety – immigration example: cloning and weighting

```
weightData <- neWeight(anxiety ~ factor(treat) + gender + age + educ + income,
                       data = framing)

head(data.frame(subset(weightData,
                       select=c('id', 'treat0', 'treat1', 'immigr', 'anxiety')),
                weights = weights(weightData)))
```

```
##   id treat0 treat1 immigr anxiety   weights
## 1  1      0      0      4       3 1.0000000
## 2  1      0      1      4       3 1.1897101
## 3  2      0      0      3       2 1.0000000
## 4  2      0      1      3       2 0.9799741
## 5  3      0      0      3       3 1.0000000
## 6  3      0      1      3       3 1.1476039
```

Fitting NE Models (2)

Second trick:

impute $\hat{Y}(a, M(a'))$ from model for $E(Y \mid a, m, x)$.

Here: $E(Y \mid a, m, x)$ imputation ('working') model.

$\Rightarrow$ 'clone' observations with $A = a'$, assign $A = a$, generate $\hat{Y}(a, M(a'))$.

$\Rightarrow$ extended data set $\Rightarrow$ can consistently estimate η's providing $p(y \mid a, m, x)$-model correctly specified.

NE Models – Imputing

with `medflex` ***(Steen et al, 2017)***

Anxiety – immigration example: imputing

```
impData <- neImpute(immigr_bin ~ factor(treat) + anxiety + gender + income + age + e
                    family = binomial,
                    data = framing)

head(subset(impData,
            select=c('id', 'treat0', 'treat1', 'immigr', 'anxiety', 'immigr_bin')))
```

```
##   id treat0 treat1 immigr anxiety immigr_bin
## 1  1      0      0      4       3  0.9406477
## 2  1      1      0      4       3  0.9714502
## 3  2      0      0      3       2  0.7453235
## 4  2      1      0      3       2  0.8626985
## 5  3      0      0      3       3  0.4441374
## 6  3      1      0      3       3  0.6317367
```

NE Model – Example

Example: Attitudes to immigration

NE model: logistic without interaction; imputation: linear / main effects

Output: log-odds-ratios

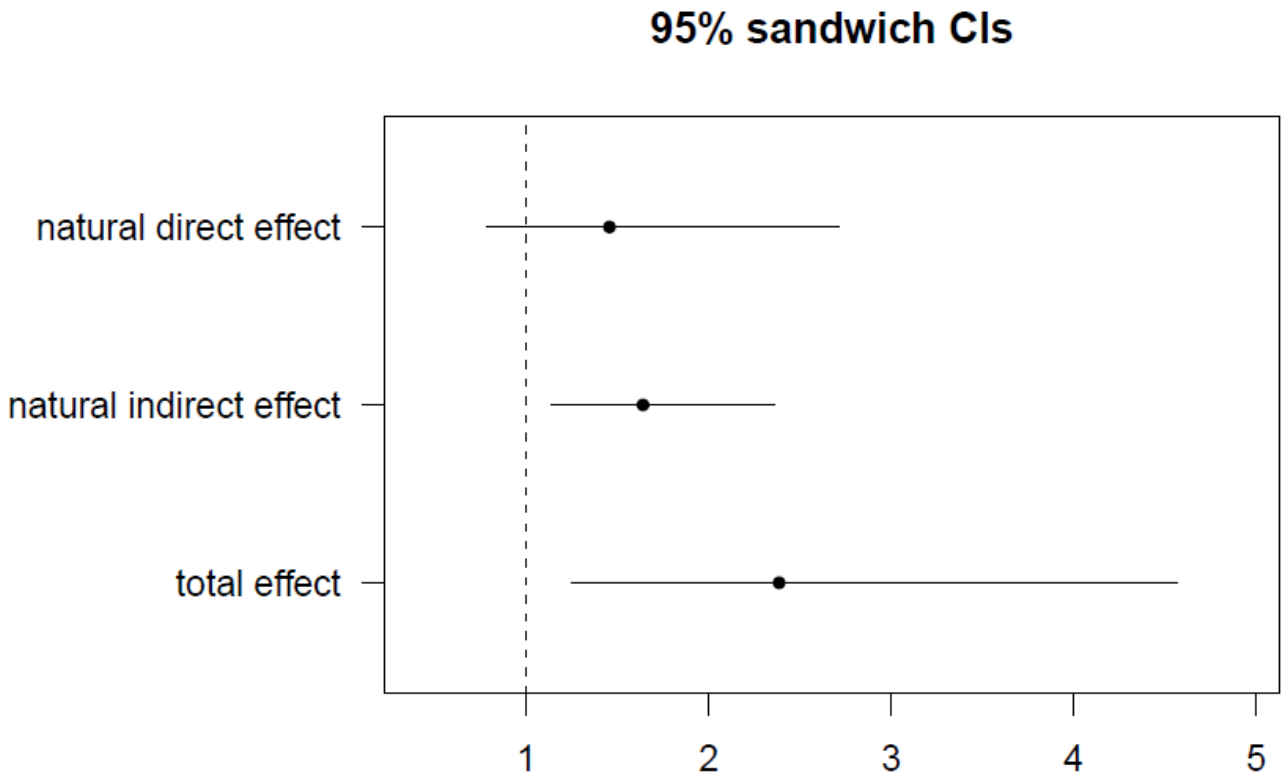
```
neModOR <- neModel(immigr_bin ~ treat0 + treat1, family=binomial, expData = impData,
summary(neModOR)

## Natural effect model
## with robust standard errors based on the sandwich estimator
## ---
## Exposure: treat
## Mediator(s): anxiety
## ---
## Parameter estimates:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   0.5738     0.1510   3.800 0.000145 ***
## treat01       0.3753     0.3178   1.181 0.237668
## treat11       0.4941     0.1857   2.661 0.007787 **
```

NE Model – Example

Example: Attitudes to immigration

Graphical display for odds-ratios no interaction:



Congeniality

NE model is for $E(Y(a, M(a')))$,

imputation model is for $E(Y|a, m, x)$

Congeniality: violated if these two models are not compatible with each other, e.g. both logistic (non-collapsibility)

Problem known from missing data theory

$\Rightarrow$ Advice: choose sufficiently rich imputation model, ideally saturated.

NE Model – Example

Example: Attitudes to immigration

NE model: logistic with interaction; rich imputation model

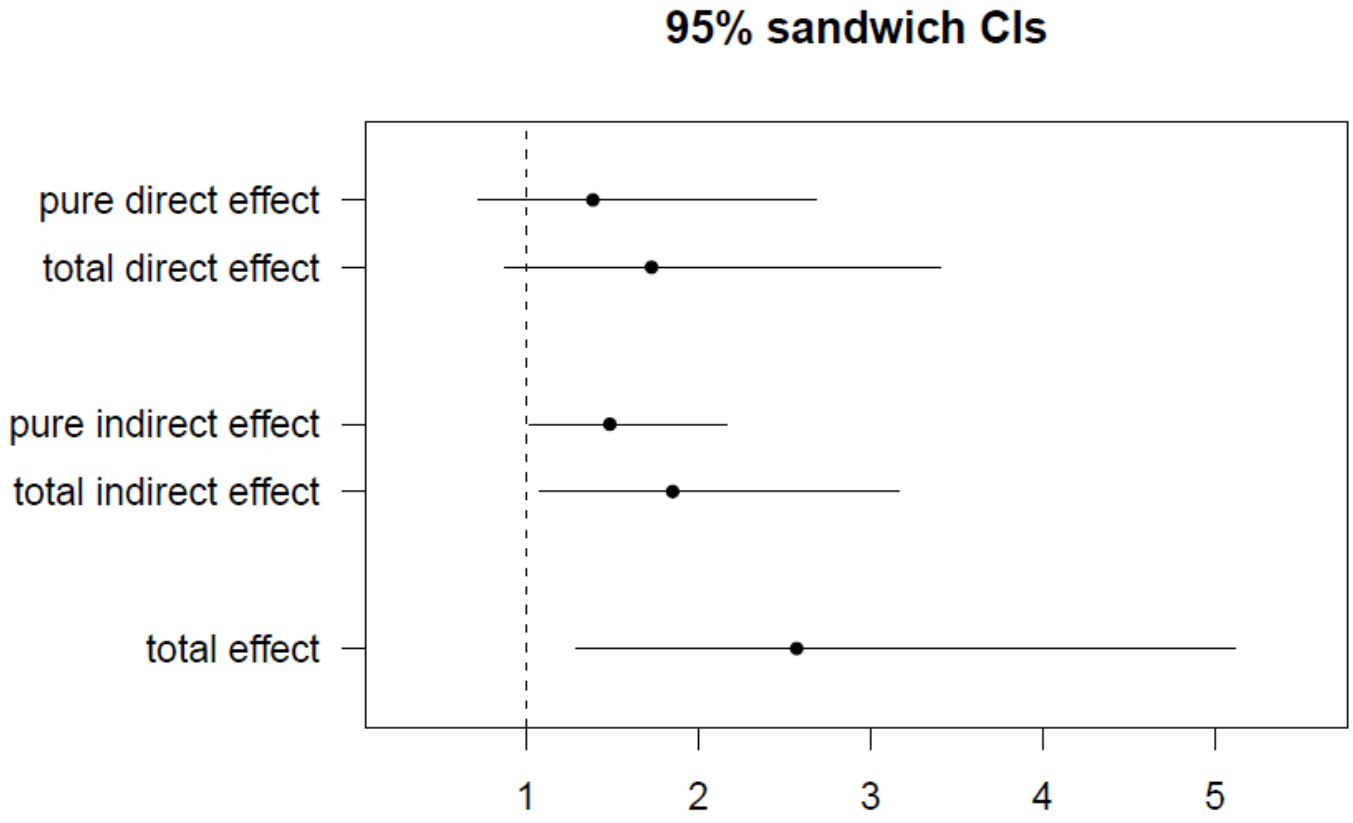
Output: log-odds-ratios

```
##                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)       0.5957     0.1489   4.002 6.29e-05 ***
## treat01           0.3285     0.3357   0.978   0.3279
## treat11           0.3970     0.1907   2.082   0.0374 *
## treat01:treat11   0.2193     0.2779   0.789   0.4300
```

NE Model – Example

Example: Attitudes to immigration

Graphical display for odds-ratios with interaction:



Notes on `medflex`

(Steen et al, 2017)

- Allows for multiple mediators
- No sensitivity analysis included yet
- Outputs on scale as determined by NE model
- Offers machine learning models for fitting flexible imputation models.

Natural Effects Models

Weighting versus imputing?

- In principle: the same (e.g. with saturated models).
- In practice: different... must decide which aspect, $M|A, X$ or $Y|M, A, X$, can be modelled more plausibly
- ... but weighting often less stable (extreme weights), especially when M continuous, requires density estimation.
- Weighting avoids inadvertent extrapolation — small weights appropriately result in large standard errors.
- Imputation: specify sufficiently rich imputation model, for mean only.

G-Formula versus NE Models

- In principle: the same (e.g. with saturated models)
- NE models avoid g-null paradox, and less parametric modelling altogether
- NE models use immediately interpretable parameters / less computationally intensive than MC methods
- NE models fit elegantly with separable effects interpretation in terms of $E(Y|\text{do}(A^Y = a, A^M = a'))$!

Causal Mediation Analysis

Outlook

Seperable effects approach of Robins & Richardson (2011) has been extended to

- survival settings with time-varying mediator (Didelez, 2019)
- ... using additive hazards model (Aalen et al, 2019)
- to competing risks (Stensrud et al, 2019)

Yet another **alternative approach** is based on 'randomised interventions' in the mediator

(Didelez et al, 2006; Vansteelandt & Daniel, 2016)

Causal Mediation Analysis

Summary

For realistic / plausible data analyses: LSEMs too simplistic.

Over many technical issues, don't forget most important points:

- What is the research question / target of inference and is it adequately addressed by causal mediation approaches?

  Do we believe at least hypothetically in separable effects?

  Is research question better addressed by joint effects?

- Are the identifying assumptions plausibly met?

  – no unobserved confounding especially of Y and M?

  – no treatment-induced confounding of Y and M?

Appendix II:

Instrumental Variables (IVs)

IVs: Motivation

Unobserved cofounding present

Often in observational studies: assumption of sufficient covariates (or 'no unmeasured confounders') not realistic.

Alternative: can sometimes use an **instrumental variable (IV)** to identify, at least partially, desired causal effect

IV: similar to 'nature is randomising' (or some other external source of randomness)

IVs still rely on **assumptions**, but different ones...

Instrumental Variables

Notation

$G =$ instrumental variable (e.g. genetic marker)

$A =$ exposure of interest (e.g. alcohol consumption)

$U =$ unobserved confounders (e.g. life-style)

$Y =$ outcome of interest (e.g. cardiovascular disease)

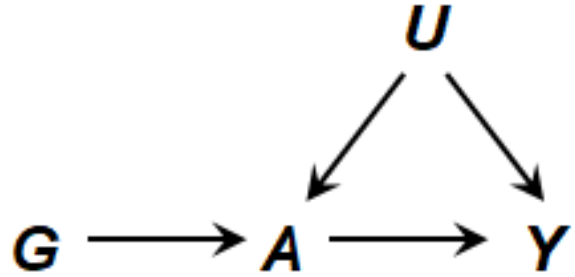
Mostly: target in terms of $E(Y|\text{do}(A = a))$

Sometimes in terms of potential outcomes $Y(a)$

Assumptions of IV

G is IV for the effect of A on Y if there is a U with

1. $G \perp\!\!\!\perp U$
2. $G \not\perp\!\!\!\perp A$
3. $G \perp\!\!\!\perp Y \mid (A, U)$.



Structural assumptions:

$$p(y|u,a) = p(y|u; \mathsf{do}(a)), \quad p(g) = p(g|\mathsf{do}(a)), \quad p(u) = p(u|\mathsf{do}(a))$$

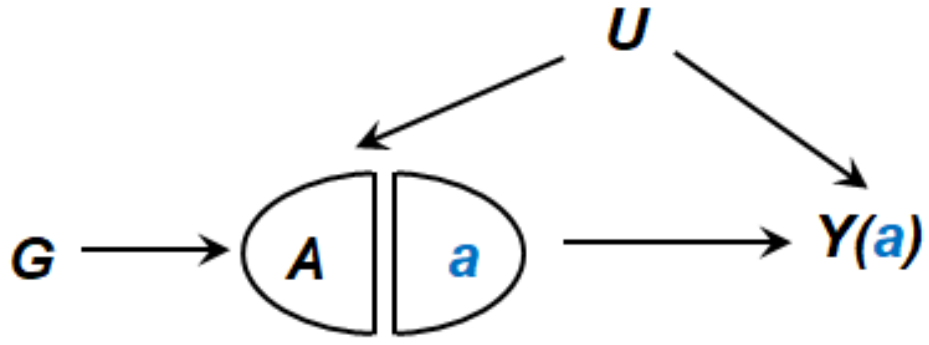
i.e. (cond.) distributions not changed by intervention in A.

(Greenland, 2000; Hernán & Robins, 2006, Didelez & Sheehan, 2007)

Assumptions of IV – with SWIG

Alternatively write assumptions as

1.' $Y(a) \perp\!\!\!\perp G$
2. $G \not\perp\!\!\!\perp A$



Note: many other versions of IV
$\rightarrow$ subtle differences in estimands

Examples:
'exclusion restriction' $Y(a, g) = Y(a)$;
'monotonicity' $A(G = 1) \geq A(G = 0)$ (where $A(g)$ PO of A under setting of G).

Examples for IVs

- In randomised trials with partial compliance:

  IV = treatment assignment, $A =$ actual treatment taken, $Y =$ health outcome.

- In epidemiology: IV = genetic variant, $A =$ exposure (often phenotype), $Y =$ health outcome

  $\Rightarrow$ **Mendelian randomisation**

- In observational studies / econometrics:
  - physicians drug preference,
  - accessibility of facilities,
  - birth date for years of education,
  - weather conditions for availability of fish / cereal etc.
  - lottery situations etc.

Use of IVs?

Testing:

check if $Y \perp\!\!\!\perp G$ — this is (roughly) testing whether there is a causal effect at all.

Estimation:

(1) when all observable variables are discrete, we can obtain bounds on causal effects without further assumptions.

(2) for point estimates need some (semi–)parametric / structural assumptions, as well as clear definition of target causal parameter.

But first, will discuss IV assumptions.

'Untestable' Assumptions

The assumptions

1. $G \perp\!\!\!\perp U$
3. $G \perp\!\!\!\perp Y \mid (A, U)$.

do not imply that $G \perp\!\!\!\perp Y | A$ or $G \perp\!\!\!\perp Y$ — check!!.

However, when all variables discrete, they impose inequality restrictions on the joint distribution $p(y, a|g)$ — these can easily be checked and provide a test against *gross violations* of the above assumptions. (Balke & Pearl, 1994)

Structural assumptions cannot be tested and may even depend on the particular intervention you have in mind.

$\Rightarrow$ Justify IV assumptions with **expert background** knowledge!

Justifying IV Assumptions

Example: Partial Compliance

In randomised trials with partial compliance: G = treatment assignment, $A =$ actual treatment taken, $Y =$ health outcome.

Treatment assignment is randomised $\Rightarrow G \perp\!\!\!\perp U$ seems very plausible.

Most subjects comply with treatment assignment $\Rightarrow G \not\perp\!\!\!\perp A$.

$Y \perp\!\!\!\perp G | (U, A)$ usually only plausible in a double–blind randomised trial!

Justifying IV Assumptions

Example: Effect of alcohol consumption

Genotype: ALDH2 determines blood acetaldehyde, the principal metabolite for alcohol.

Two alleles/variants: wildetype *1 and “null” variant *2.

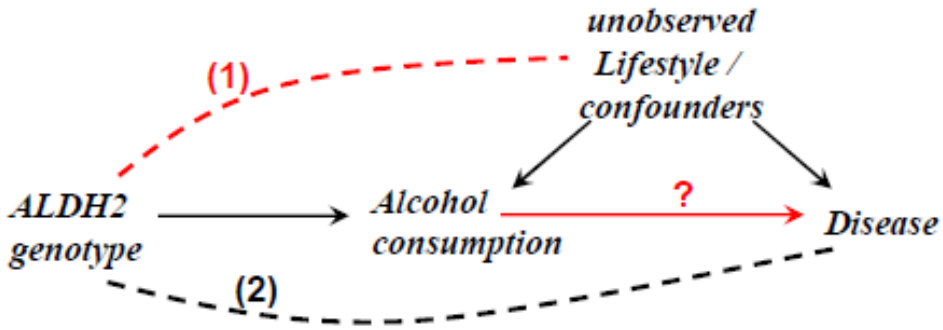
*2*2 homozygous individuals suffer facial flushing, nausea, drowsiness and headache after alcohol consumption.

$\Rightarrow$ *2*2 homozygous individuals have low alcohol consumption *regardless* of other lifestyle behaviours – Mendelian randomisation

IV–idea: check if these individuals have a different risk than others for alcohol related health problems!

Justifying IV Assumptions

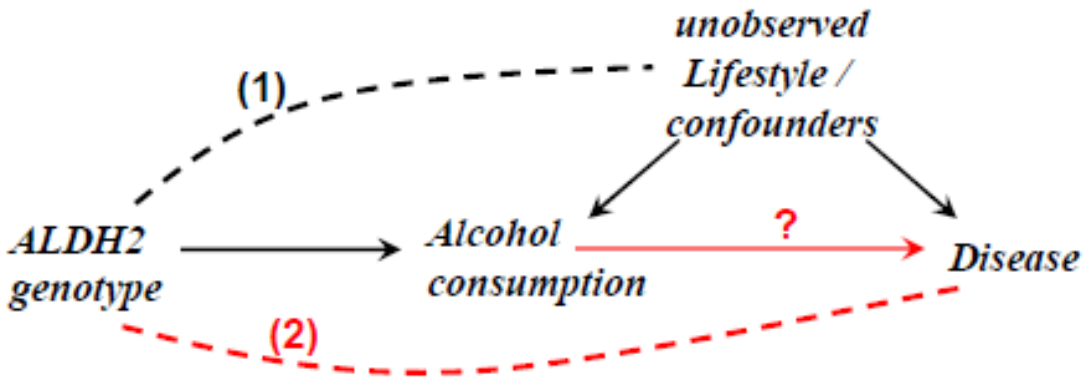
Example: Effect of alcohol consumption



Note 1: due to random allocation of genes at conception, can be fairly confident that genotype is not associated with unobserved confounders (subpopulation structure can be a problem).
Further evidence: in extensive studies no evidence for association with *observed* confounders, e.g. age, smoking, BMI, cholesterol. (see e.g. Davey Smith et al., 2007)

Justifying IV Assumptions

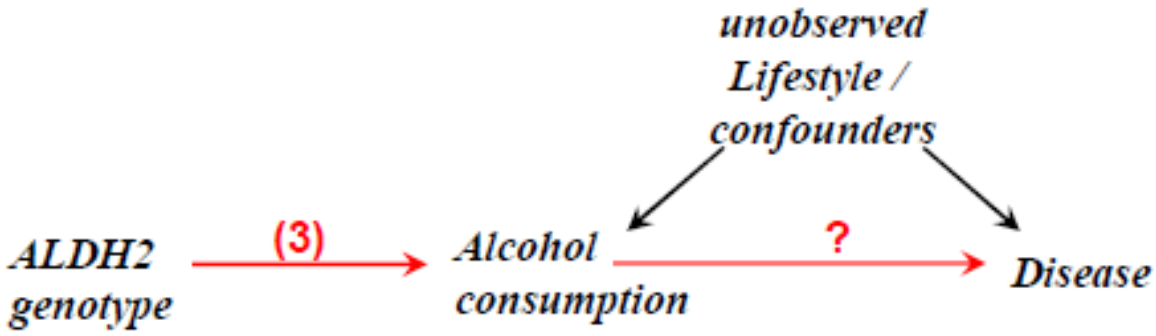
Example: Effect of alcohol consumption



Note 2: due to known 'functionality' of ALDH2 gene, we can exclude that it affects the typical diseases considered by *another* route than through alcohol consumption.
$\Rightarrow$ important to use well studied genes as instruments!

Justifying IV Assumptions

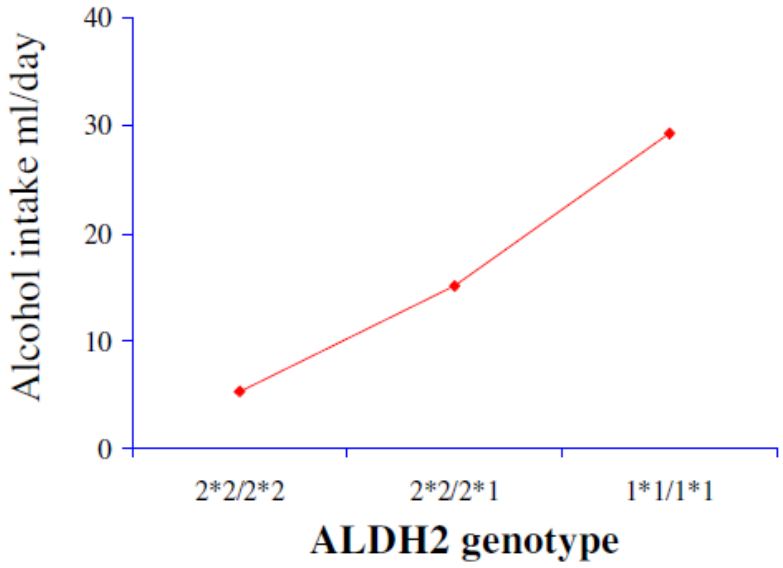
Example: Effect of alcohol consumption



Note 3: association of ALDH2 with alcohol consumption well established, strong, and underlying biochemistry well understood.

Justifying IV Assumptions

Example: Effect of alcohol consumption



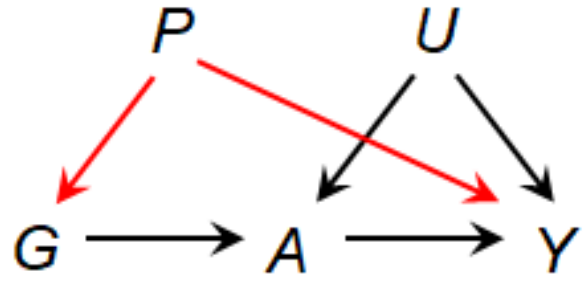
Note 3: association of ALDH2 with alcohol consumption well established, strong, and underlying biochemistry well understood.

Violation of IV Assumptions

Example: Mendelian randomisation

Population stratification occurs when there exist population subgroups that experience both, different disease rates (or different distributions of phenotypes) and have different frequencies
of alleles of interest.

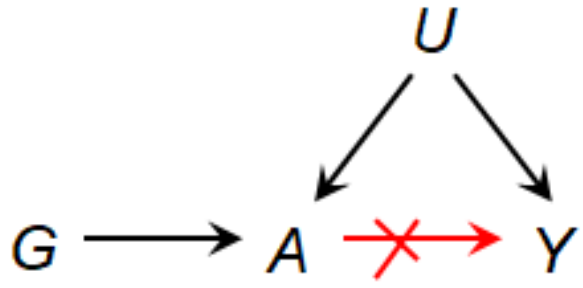
$\Rightarrow$ might violate condition
$Y \perp\!\!\!\perp G | (A, U)$.



Solution?

Testing for Causal Effect with IV

No causal effect "$\Leftrightarrow$" G independent of Y.



Here: take causal null-hypothesis as 'no $A \to Y$ edge'

Note: not the same as $ACE = 0$.

Example: Alcohol Consumption

Findings: (Meta-analysis by Chen et al., 2008)

Blood pressure on average 7.44mmHg higher and
risk of hypertension 2.5 higher
for ALDH2*1*1 than for ALDH2*2*2 carriers (only males).
$\Rightarrow$ mimics the effect of *large versus low* alcohol consumption.

Blood pressure on average 4.24mmHg higher and
risk of hypertension 1.7 higher
for ALDH2*1*2 than for ALDH2*2*2 carriers (only males).
$\Rightarrow$ mimics the effect of *moderate versus low* alcohol consumption.

Testing for Causal Effect with IV

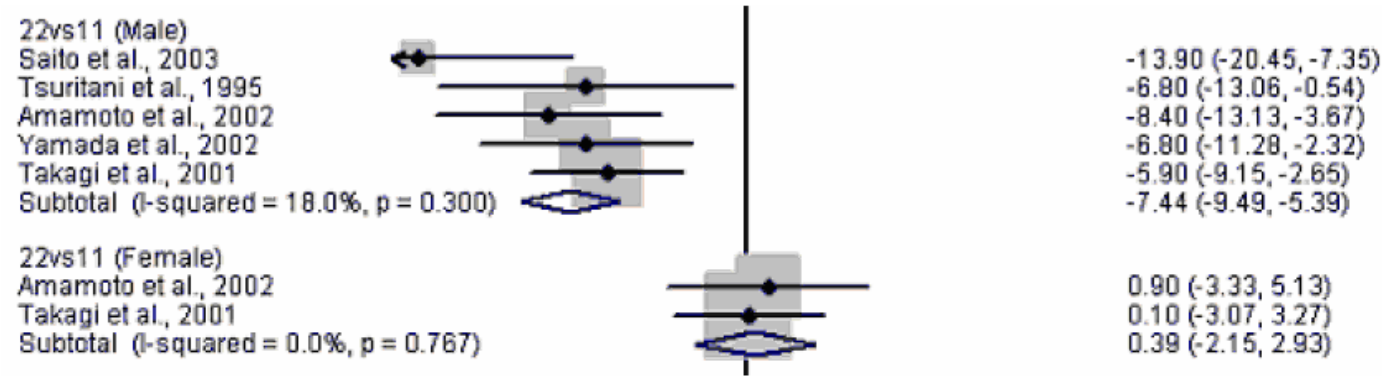
Results of Meta-analysis by Chen et al. (2008) suggest that **even moderate** alcohol consumption is **harmful**.

Note: studies mostly in Japanese populations (where ALDH2*2*2 is common), where women drink only little alcohol in general.

$\Rightarrow$ use women as 'negative control' group.

Example: Alcohol Consumption

Is condition $Y \perp\!\!\!\perp G|(A, U)$ satisfied?



Some indication

Women in Japanese study population did not drink. ALDH2 genotype in women not associated with blood pressure $\Rightarrow$ there does not seem to be another pathway creating a G–Y association here.

Bounds on Causal Effect

The all–binary case

Without parametric assumptions we cannot normally identify any *population* causal effect parameters.

But with A, Y, G all binary (or all discrete) we can derive upper and lower **bounds** on the causal effect (e.g. ACE).

(Balke & Pearl, 1994)

The derivation exploits restrictions on joint distribution of A, Y, G due to the conditional indepedencies involving U.

Interpretation of bounds: for a given observed frequency table on A, Y, G there exist different causal models that agree with these frequencies and can give causal effects anywhere within these bounds.

Bounds on Causal Effect

Y, A and G are **all binary**;
$ACE = E(Y|\text{do}(A = 1)) - E(Y|\text{do}(A = 0))$.

Let $p_{yx.g} = p(y, a|g)$. Then we have

$$\left.\begin{array}{c} p_{11.1} + p_{00.0} - 1 \\ p_{11.0} + p_{00.1} - 1 \\ p_{11.0} - p_{11.1} - p_{10.1} - p_{01.0} - p_{10.0} \\ p_{11.1} - p_{11.0} - p_{10.0} - p_{01.1} - p_{10.1} \\ -p_{01.1} - p_{10.1} \\ -p_{01.0} - p_{10.0} \\ p_{00.1} - p_{01.1} - p_{10.1} - p_{01.0} - p_{00.0} \\ p_{00.0} - p_{01.0} - p_{10.0} - p_{01.1} - p_{00.1} \end{array}\right\} \leq \text{ACE} \leq \left\{\begin{array}{c} 1 - p_{01.1} - p_{10.0} \\ 1 - p_{01.0} - p_{10.1} \\ -p_{01.0} + p_{01.1} + p_{00.1} + p_{11.0} + p_{00.0} \\ -p_{01.1} + p_{11.1} + p_{00.1} + p_{01.0} + p_{00.0} \\ p_{11.1} + p_{00.1} \\ p_{11.0} + p_{00.0} \\ -p_{10.1} + p_{11.1} + p_{00.1} + p_{11.0} + p_{10.0} \\ -p_{10.0} + p_{11.0} + p_{00.0} + p_{11.1} + p_{10.1} \end{array}\right.$$

$\Rightarrow$ can easily be estimated by observed relative frequencies.

Note: bounds are *sharp* — for given frequencies on (A, Y, G) there always exists joint distributions on (A, Y, U, G) for which the bounds are attained.

Notes on Bounds

- In most realistic scenarios: bounds are very wide and include 'no–causal–effect,' i.e. include $ACE = 0$.
- **Interpretation:** if $ACE = 0$ included, there is always another model, where A has no causal effect on Y, that could generate the same data.
- Width of bounds depends on strength of IV and amount of confounding.
- Still, bounds should always be calculated to assess how informative the data 'alone' are.
- Stata / R package `bpbounds` (Palmer et al., 2011, 2018); various IV methods: R package `ivtools` (Sjølander, 2018)

IV Estimation

ETT

The binary case — Effect of treatment on the treated (ETT)

With a key parametric assumption, we can identify the causal effect within a *subgroup* of the population, the treated

Assume structural mean model (SMM)

$$E(Y(1) - Y(0)|A = 1, G = g) = \psi$$

IV Estimation

ETT

Assume: **no effect modification** by the IV G (NEM)

$$E(Y(1) - Y(0)|A = 1, G = g) = \psi$$

It can then be shown that

$$E(Y(1) - Y(0)|A = 1) = \frac{E(Y|G = 1) - E(Y|G = 0)}{E(A|G = 1) - E(A|G = 0)}.$$

$\Rightarrow$ 'ratio estimator' (Wald-estimator)

With $E(Y(1) - Y(0)|U) = E(Y(1) - Y(0))$, i.e. no effect modification by U (on additive scale), the above equals population ACE.

IV Estimation

All binary: other target parameters, e.g.

'Wald type' IV estimators for RR and OR (Y and G binary)

$$WaldRR = \hat{RR}(Y|G)^{1/\Delta} \qquad WaldOR = \hat{OR}(Y|G)^{1/\Delta}$$

where $\Delta = \hat{E}(A|G=1) - \hat{E}(A|G=0)$.

WaldRR consistent for CRR if (Didelez et al, 2010)
– log-linearity of Y in A
– no A-U interaction on Y on log-linear scale
– $A|(G,U)$ normally distributed.

WaldOR approximation to WaldRR for rare disease.

Advantages: WaldOR can be used in case–control studies.

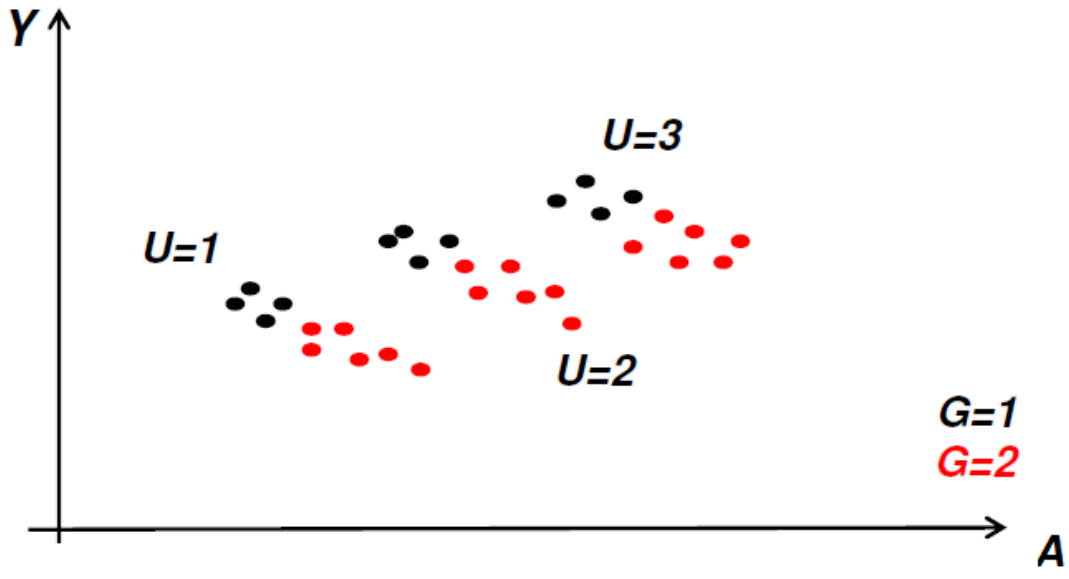
IV Estimation: 2SLS

The linear case: two-stage-least-squares (2SLS)

Some intuition first!

IV Estimation: 2SLS

Linear models



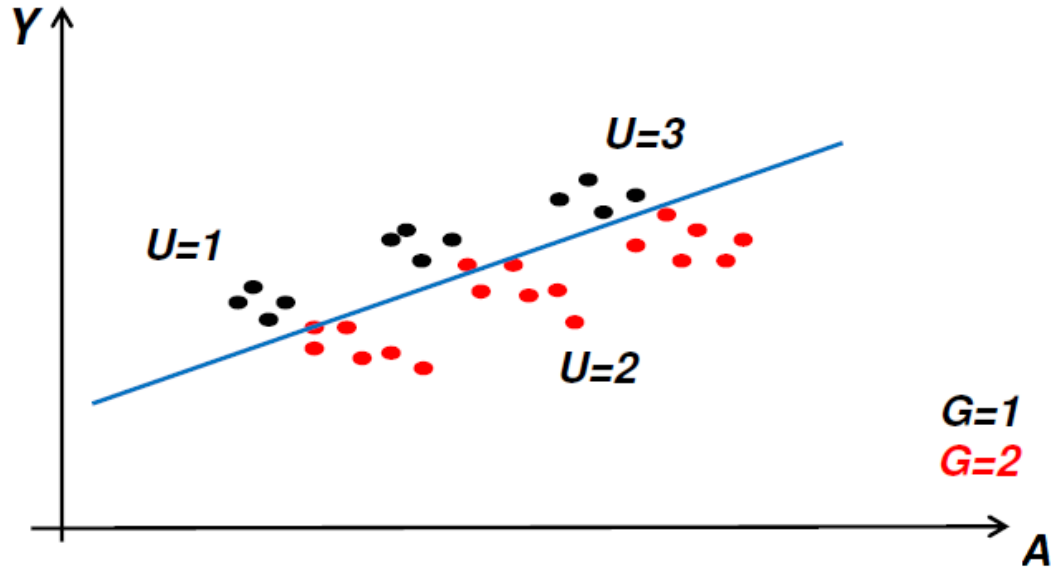
Positive confounding: larger values of U induce larger Y and larger A.

But conditional on (unobservable) U we have that Y and A have negative association.

Different colours = different values of IV.

IV Estimation: 2SLS

Linear models

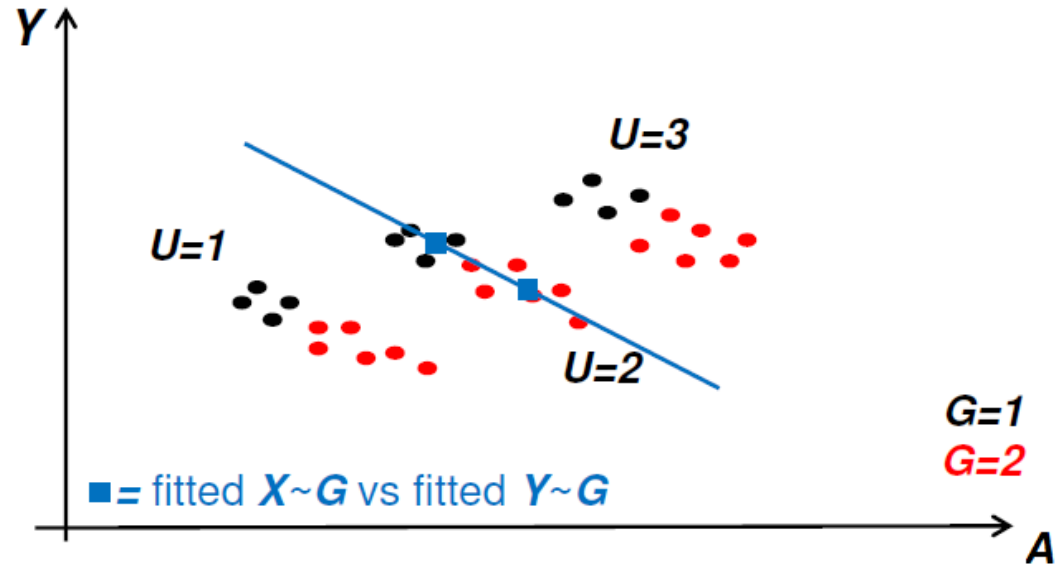


Regression of Y on A (ignoring U and G) results in positive slope.

Due to role of (unobservable!) U, biased estimate of causal effect.

IV Estimation: 2SLS

Linear models



With instrument: regress of A on G and Y on G and divide slopes.

This recovers the negative slope without knowing U.

IV Estimation in LSEM

Written as LSEM:

two endogenous variables

$$Y = \beta_0 + \beta A + \xi_Y(U)$$

$$A = \alpha_0 + \alpha G + \xi_A(U)$$

where $\xi_Y(U), \xi_A(U)$ are **correlated errors**.

G as IV is exogenous variable.

Econometrics: various approaches to estimating such (and more general) systems of equations (Bowden & Turkington, 1984)

IV Estimation in LSEM

$$Y = \beta_0 + \beta A + \xi_Y(U)$$

$$A = \alpha_0 + \alpha G + \xi_A(U)$$

Path-tracing results in:

– total effect of G on A is α

– total effect of G on Y is $\alpha\beta$

$\Rightarrow \beta =$ ratio of coefficients from OLS regr. Y on G and A on G.

Or: regress Y on $\hat{A}$, predicted from OLS $A|G$.

IV Estimation: 2SLS

Alternative: **weaker** model assumption — linear SMM.

A, Y, G arbitrary scale. Assume

$$E(Y|U = u; \text{do}(A = a)) = E(Y|u, a) = \mu_Y + \beta a + h(u)$$

Note, no (A, U)-interaction on linear scale.

Then $ACE = \beta$.

Can show

$$\beta = \frac{Cov(Y, G)}{Cov(A, G)}$$

i.e. β is identified from obs. data on A, Y, G.

IV Estimation: 2SLS

The linear case

Hence, consistent estimator for β given by ratio of estimated coefficients from regression of Y on G and from A on G

$\Rightarrow$ called IV–estimator or two–stage–least–squares (2SLS):

$$\hat{\beta}_{IV} = \frac{\hat{\beta}_{Y|G}}{\hat{\beta}_{A|G}}$$

where $\hat{\beta}_{Y|G}$, $\hat{\beta}_{A|G}$ least squares regression coefficients.

Note: denominator: weak IV (weak G-A association)
$\Rightarrow$ unstable *and also biased* IV estimators.

Notes on 2SLS

- popular, very simple to implement (many softwares)
- surprisingly robust towards misspecification

  (Vansteelandt & Didelez, 2017)
- can be generalised to multiple IVs, multiple exposures, multiple outcomes; but **weak** IV problem quickly becomes more serious in higher dimensions
- can also be used in 2-samples situation with separate (A, G)-data and (Y, G)-data.

Instrumental Variables

Summary

In presence of unobserved confounding: hope to find IV

- 'natural' experiment – genes, year of birth etc;
- can be used for testing for causal effect, or bounds;
- *estimation* requires more assumptions (e.g. NEM, linearity or other);
- recent work: inference with multiple instruments, some of which may be *invalid* (Bowden et al., 2015; Guo et al, 2018)

Natural Experiments

When suspicious of unobserved confounding: look for ‘natural experiments’

- regression discontinuity designs (RDD);
- interrupted time-series (e.g. policy changes) / difference-in-differences / before-after-design;
- negative controls;
- differences in difference;
- twin / sibling studies etc.

Unobserved Confounding

Absence of instruments / natural experiments?
$\Rightarrow$ sensitivity analysis!

See book: Lash et al (2009)

– needs some assumption on plausible confounding

– ad-hoc adjustment formulas or

– MC methods or

– Bayesian approaches (Gustafson et al, 2010)

Thank You!

www.leibniz-bips.de/en

Contact
Vanessa Didelez
Leibniz Institute for Prevention Research
and Epidemiology – BIPS
Achterstraße 30
D-28359 Bremen
didelez@leibniz-bips.de