Three Applications of Structural Equation Modeling as a Handy Tool:

A Network Comparison Test,
Testing Moderated Mediation, and
a Means to Correct for the Effects of Censoring

Kees-Jan Kan
Niels Smits

Why this title? (Part I)

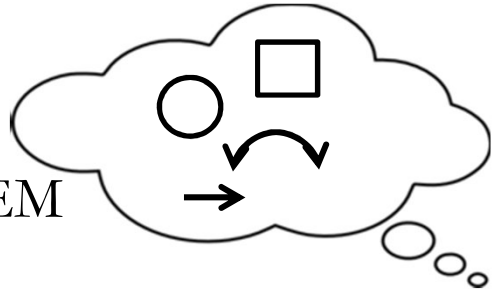
Sample of questions from consults:

- 'How can I test if a given psychometric network is invariant over males and females?'
- 'Is it correct I need to use a moderated mediation model in order answer my research question?'
  - i.e. Can I use SPSS macro PROCESS?
- 'How do I test if the genetic variance in a variabele is different across SES groups?'
  - while the variable was censored (to different extents)

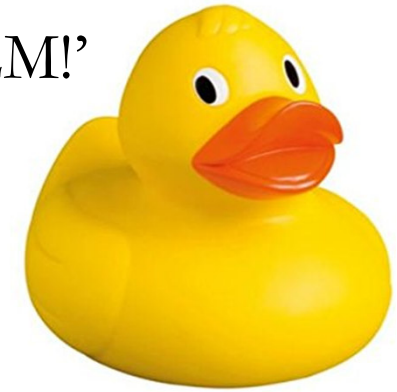
Why this title? (Part II)

Because I…

- look like a SEM-er
- quack like a SEM-er
- am influenced by Peter Molenaar
  - who translates almost everything to SEM



… I often think 'I believe this can be done using SEM!'



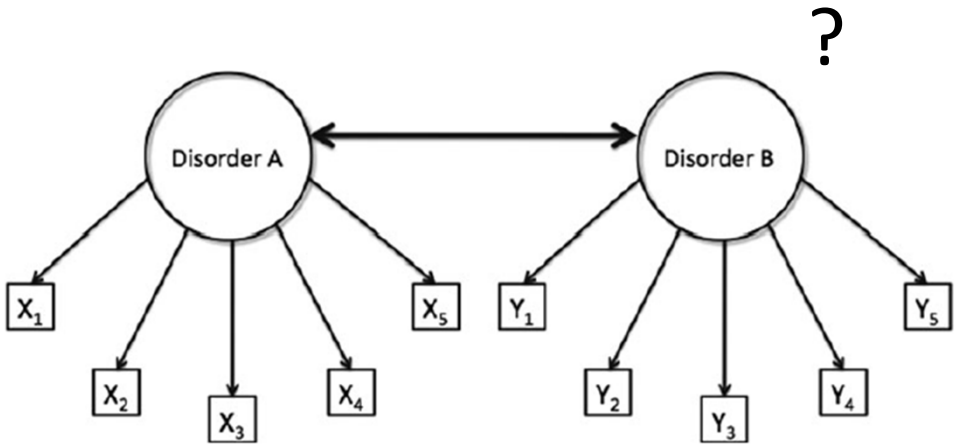
- Indeed SEM provided answers to all 3 questions
- In addition, SEM provided 'the best' solutions
  - That is, among the alternatives I encountered
- **SEM is such a handy tool!**

Why this title? (Part III)

Based on **recent reviewing experience**, I expect future questions:

Like: **'How do I test for age changes in comorbidity?'**

What pops up? (in the mind of a SEM-er)



Why this title? (Part III)

What If I told you

- Non-clinical samples are assessed using clinical instruments

  → **censored data**

  ML assumes normality (categorization?)

- The researcher who will pose that question is interested in **psychometric network models** instead of latent variable models

Why this title? (Part III)

Proposal:

- Implement the psychometric network in SEM
- Fit this model on the raw data so that the 'edges' be moderated by age
- Include correction for censoring effects

Why this title? (Part IV)

- Integration of the applications

- I thought 'better demonstrate them separately'
  - a network comparison test,
  - testing moderated mediation, and
  - a means to correct for the effects of censoring

Application 1

A Network Comparison Test

Psychometric network modeling

BEHAVIORAL AND BRAIN SCIENCES (2010) **33**, 137–193
doi:10.1017/S0140525X09991567

Comorbidity: A network perspective

Network Analysis: An Integrative Approach to the Structure of Psychopathology

Denny Borsboom and Angélique O.J. Cramer

Department of Psychology, University of Amsterdam, Amsterdam 1018 XA, The Netherlands;
email: D.Borsboom@uva.nl

Angélique O. J. Cramer
Department of Psychology, University of Amsterdam, 1018 WB Amsterdam, The Netherlands
A.O.J.Cramer@uva.nl
www.aojcramer.com

Lourens J. Waldorp
Department of Psychology, University of Amsterdam, 1018 WB Amsterdam, The Netherlands
L.J.Waldorp@uva.nl
http://users.fmg.uva.nl/lwaldorp

Han L. J. van der Maas
Department of Psychology, University of Amsterdam, 1018 WB Amsterdam, The Netherlands
H.L.J.vanderMaas@uva.nl
http://users.fmg.uva.nl/hvandermaas/

Denny Borsboom
Department of Psychology, University of Amsterdam, 1018 WB Amsterdam, The Netherlands
D.Borsboom@uva.nl
http://sites.google.com/site/borsboomdenny/dennyborsboom

An alternative to latent variable modeling

BEHAVIORAL AND BRAIN SCIENCES (2010) 33, 137–193
doi:10.1017/S0140525X09991567

Comorbidity: A network perspective

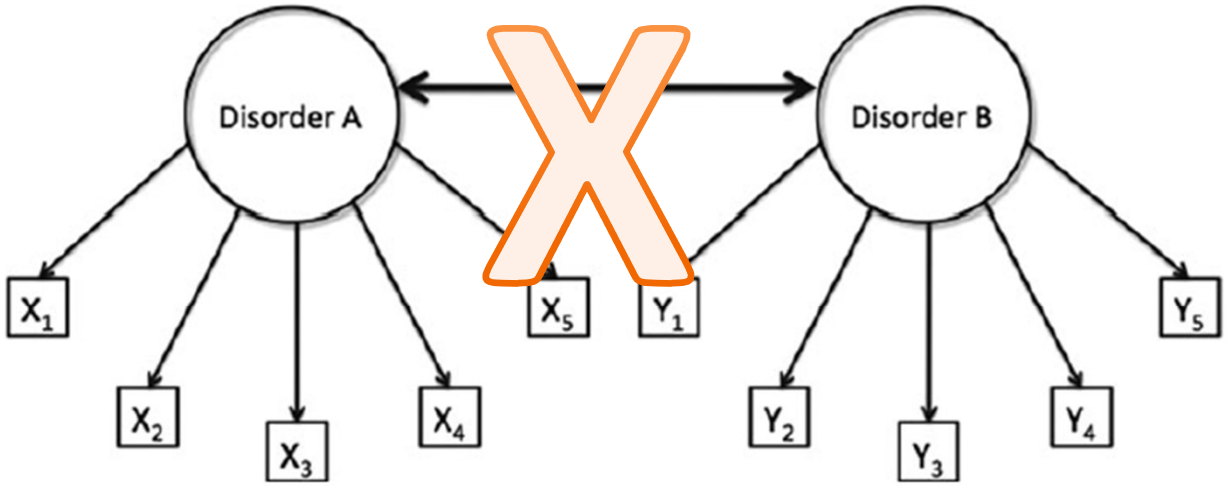


Figure 1. A model of comorbidity between disorders A and B, under the standard assumptions of latent variable modeling. The *circles* represent the disorders (i.e., latent variables) and the *rectangles* represent the observable core symptoms of those disorders (i.e., $X_1 - X_5$ for disorder A, and $Y_1 - Y_5$ for disorder B). In this model, comorbidity is viewed as a correlation between the latent variables, visualized by the *thick bidirectional edge* between disorders A and B.

An alternative to latent variable modeling

BEHAVIORAL AND BRAIN SCIENCES (2010) 33, 137–193
doi:10.1017/S0140525X09991567

Comorbidity: A network perspective

Angélique O. J. Cramer
Department of Psychology, University of Amsterdam, 1018 WB Amsterdam, The Netherlands
A.O.J.Cramer@uva.nl
www.ajcramer.com

Lourens J. Waldorp
Department of Psychology, University of Amsterdam, 1018 WB Amsterdam, The Netherlands
L.J.Waldorp@uva.nl
http://users.fmg.uva.nl/lwaldorp

Han L. J. van der Maas
Department of Psychology, University of Amsterdam, 1018 WB Amsterdam, The Netherlands
H.L.J.vanderMaas@uva.nl
http://users.fmg.uva.nl/hvandermaas/

Denny Borsboom
Department of Psychology, University of Amsterdam, 1018 WB Amsterdam, The Netherlands
D.Borsboom@uva.nl
http://sites.google.com/site/borsboomdenny/dennyborsboom

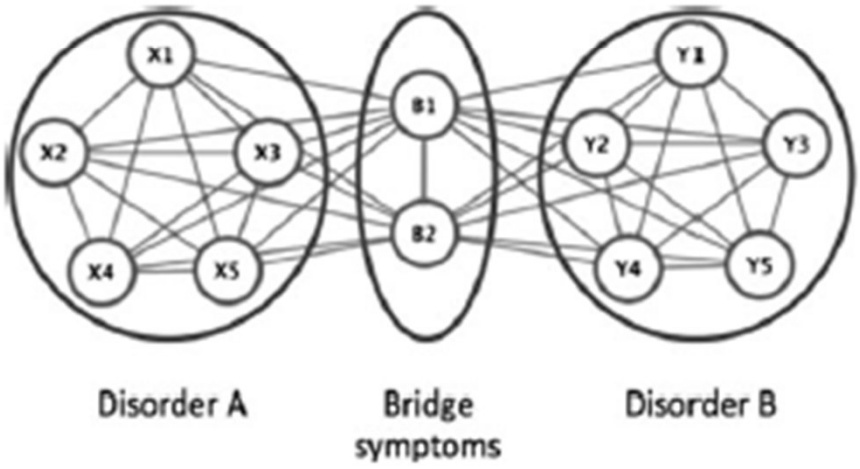
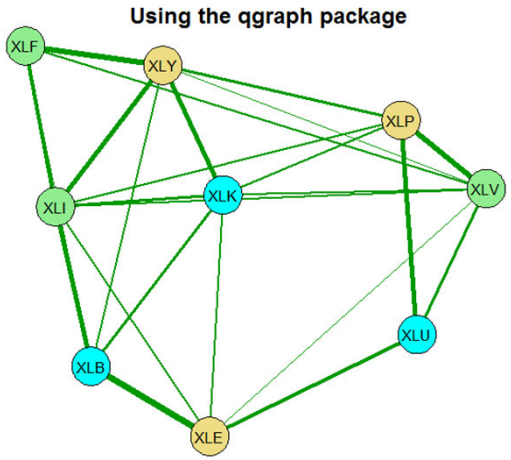


Figure 2. Comorbidity under a network approach. Disorder A consists of bidirectionally related symptoms $X_1 - X_5$, and disorder B consists of symptoms $Y_1 - Y_5$. Symptoms B_1 and B_2 are *bridge symptoms* that overlap between disorders A and B. In this model, comorbidity arises as a result of direct relations between the bridge symptoms of two disorders.

Commonly used analytical techniques



Journal of Statistical Software
May 2012, Volume 48, Issue 4. *http://www.jstatsoft.org/*

qgraph: Network Visualizations of Relationships in Psychometric Data

Sacha Epskamp
University of Amsterdam

Angélique O. J. Cramer
University of Amsterdam

Lourens J. Waldorp
University of Amsterdam

Verena D. Schmittmann
University of Amsterdam

Denny Borsboom
University of Amsterdam

R package qgraph

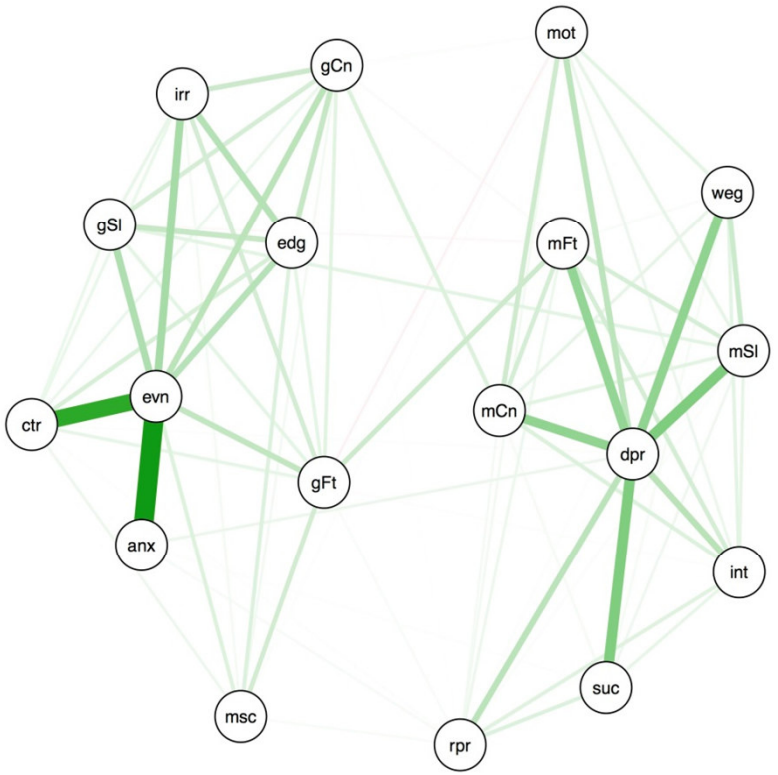
Networks based on (partial) correlations

Routines to come up with sparse matrices (glassso)

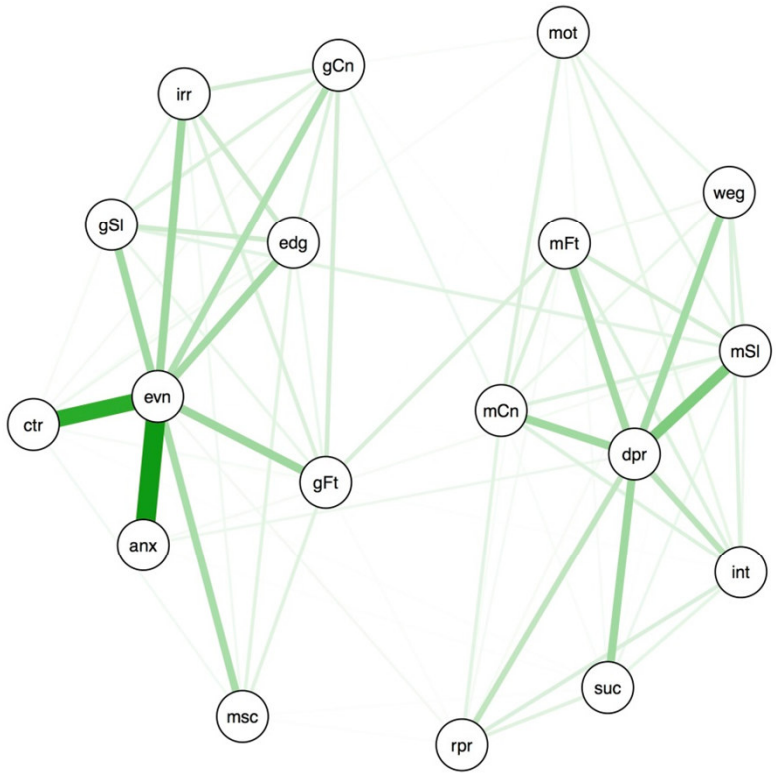
Next, these are vizualized

A network comparison test (visual comparison?)

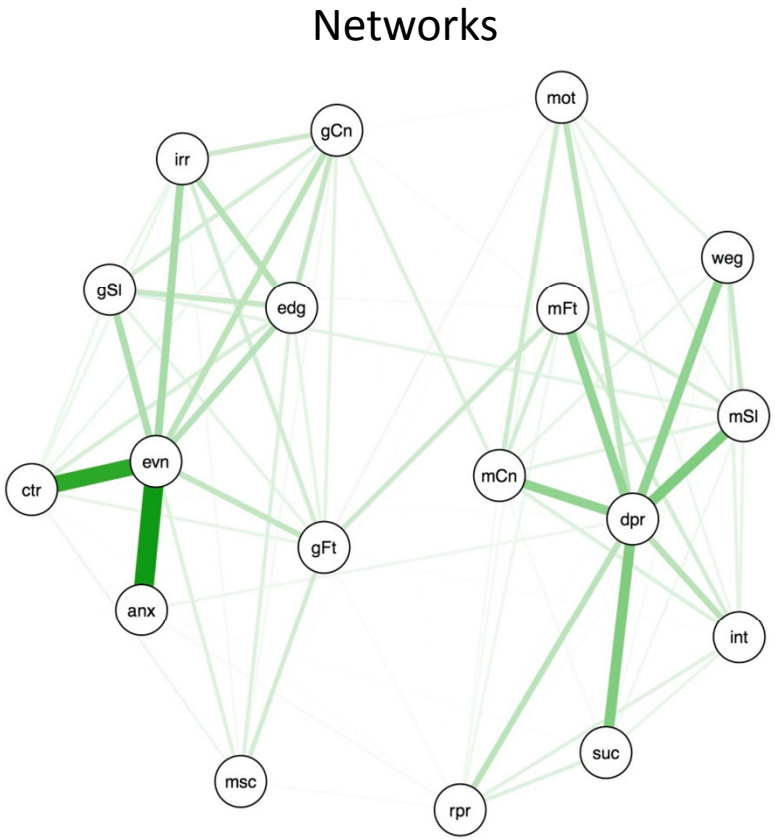
Network A

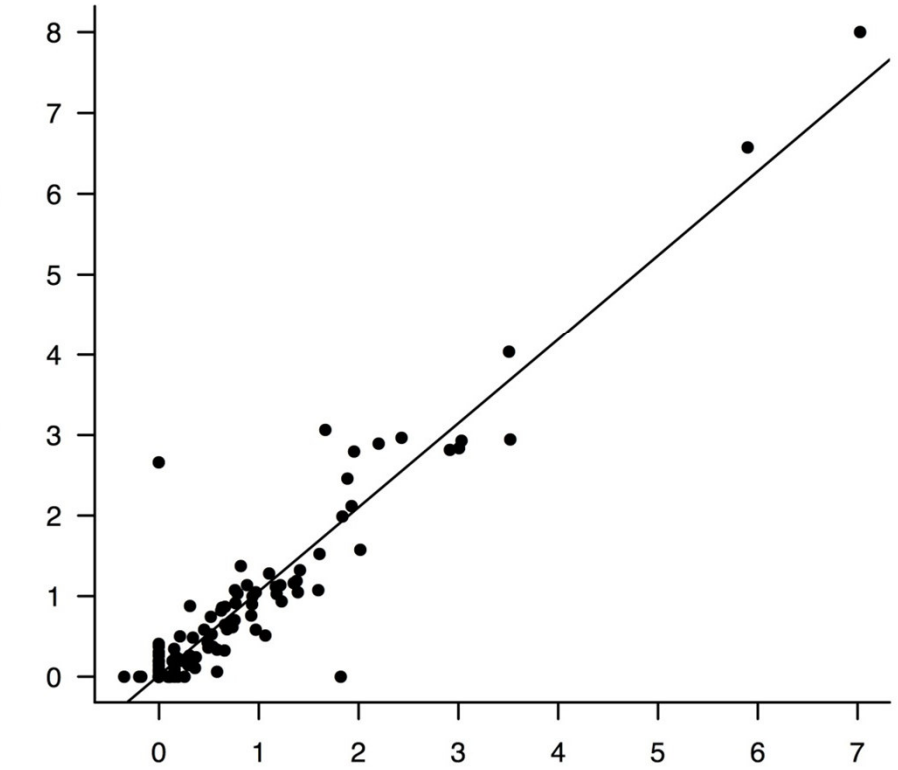


Network B



A network comparison test (correlating network properties?)





A network comparison test (permutation tests?)

Package 'NetworkComparisonTest'

October 29, 2016

Type Package

Title Statistical Comparison of Two Networks Based on Three Invariance Measures

Version 2.0.1

Date 2016-10-28

Author Claudia D. van Borkulo, with contributions from Sacha Epskamp and Alex Millner

Maintainer Claudia van Borkulo <cvborkulo@gmail.com>

1. Invariant network structure
2. Invariant edge strength
3. Invariant global strength

A network comparison test (permutation tests?)

Package 'NetworkComparisonTest'

October 29, 2016

Type Package

Title Statistical Comparison of Two Networks Based on Three Invariance Measures

Version 2.0.1

Date 2016-10-28

Author Claudia D. van Borkulo, with contributions from Sacha Epskamp and Alex Millner

Maintainer Claudia van Borkulo <cvborkulo@gmail.com>

Van Borkulo et al. "**The first invariance hypothesis […] is H0: A1=A2**, in which A1and A2 are the connection strength matrices of graphs (networks) G1and G2, respectively. **To test this hypothesis, we use a distance measure for symmetric n x n matrices: the maximum or L∞ norm.**"

A network comparison test
SEM!

Why not directly? (but how?)

This paper gave me the hint

PSYCHOMETRIKA—VOL. 82, NO. 4, 904–927
DECEMBER 2017
https://doi.org/10.1007/s11336-017-9557-x

GENERALIZED NETWORK PSYCHOMETRICS: COMBINING NETWORK AND LATENT VARIABLE MODELS

SACHA EPSKAMP, MIJKE RHEMTULLA AND DENNY BORSBOOM

UNIVERSITY OF AMSTERDAM

A network comparison test
SEM!

- Starting point
- CFA according to SEM-ers:

$$\widehat{\boldsymbol{\Sigma}} = \boldsymbol{\Lambda\Phi\Lambda}^{\mathrm{T}} + \boldsymbol{\Theta}$$

- Implementation:
  - As a system of regression equations
    - Lisrel, Mplus, lavaan, etc.
  - Drawing the model
    - Lisrel, Mplus, Amos, Onyx
  - **Matrix algebraic**
    - **Mx, OpenMx** (not) user friendly?

A network comparison test
SEM!

$$\hat{\boldsymbol{\Sigma}} = \boldsymbol{\Lambda\Phi\Lambda}^{\mathrm{T}} + \boldsymbol{\Theta}$$

- In OpenMx
  - Specify the matrices ($\boldsymbol{\Lambda}, \boldsymbol{\Phi}, \boldsymbol{\Theta}$)
  - Specify the algebraic expression (equation above)
  - Put them together with the data
  - Solve
  - Multi-group modeling is possible
  - (In)variance: test if the insertion of certain equality constraints (in $\boldsymbol{\Lambda}, \boldsymbol{\Phi}$, and/or $\boldsymbol{\Theta}$) reduce the model fit significantly

A network comparison test
SEM!

- Alternative expression (RAM model)

$$\widehat{\boldsymbol{\Sigma}} = \mathbf{F}(\mathbf{I} - \mathbf{A})^{-\mathbf{1}}\mathbf{S}(\mathbf{I} - \mathbf{A})^{-\mathbf{T}}\mathbf{F}^{\mathbf{T}}$$

- In OpenMx (not very different than before)
  - Specify the matrices (now $\mathbf{A}, \mathbf{S}, \mathbf{F}, \mathbf{I}$)
  - Specify the algebraic expression (equation above)
  - Put them together with the data
  - Solve
  - Multi-group modeling is possible
  - (In)variance: test if the insertion of certain equality constraints (in $\mathbf{A}, \mathbf{and/or}\ \mathbf{S}$) reduce the model fit significantly

A network comparison test
SEM!

- The network approach (just another alternative)!

$$\widehat{\boldsymbol{\Sigma}} = \boldsymbol{\Delta}(\mathbf{I} - \boldsymbol{\Omega})^{-1}\boldsymbol{\Delta}$$

Ω is a symmetric $n \times n$ matrix containing the edges **(partial/semi-partial relations)** among the variables

(but with 0's on the diagonal)

I is an $n \times n$ identity matrix

Δ is an $n \times n$ (diagonal) matrix containing scaling parameters

A network comparison test
SEM!

- The network approach (just yet another alternative!)

$$\widehat{\boldsymbol{\Sigma}} = \boldsymbol{\Delta}(\mathbf{I} - \boldsymbol{\Omega})^{-1}\boldsymbol{\Delta}$$

In OpenMx (not very different than before)

- Specify the matrices (now $\boldsymbol{\Omega}, \boldsymbol{\Delta}, \text{and } \mathbf{I},$)
- Specify the algebraic expressions (equation above)
- Put them together with the data
- Solve
- Multi-group modeling is possible
- (In)variance: test if the insertion certain equality constraints (in $\boldsymbol{\Omega}$ and/or $\boldsymbol{\Delta}$) reduce the model fit significantly

A network comparison test
SEM!

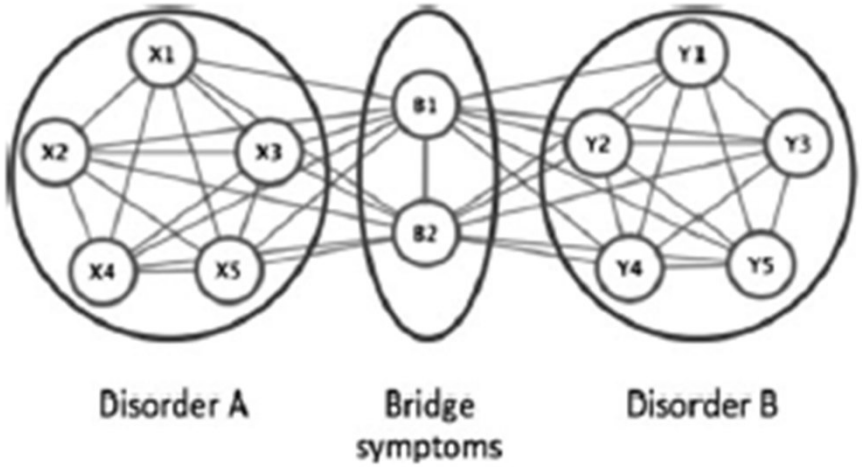


Figure 2. Comorbidity under a network approach. Disorder A consists of bidirectionally related symptoms $X_1 - X_5$, and disorder B consists of symptoms $Y_1 - Y_5$. Symptoms B_1 and B_2 are *bridge symptoms* that overlap between disorders A and B. In this model, comorbidity arises as a result of direct relations between the bridge symptoms of two disorders.

Network modeling in SEM

- Advantage: Confirmatory Network Modeling
  - Testing a priori theoretical networks
  - Replication research
  - Behavior genetic modelling

$$\widehat{\boldsymbol{\Sigma}} = \widehat{\boldsymbol{\Sigma}}_{\mathbf{genetic}} + \widehat{\boldsymbol{\Sigma}}_{\mathbf{environmetal}}$$
$$= \boldsymbol{\Lambda}(\boldsymbol{\Phi}_{\mathbf{genetic}} + \boldsymbol{\Phi}_{\mathbf{environmental}})\boldsymbol{\Lambda}^{T} + \boldsymbol{\Theta}_{\mathbf{genetic}} + \boldsymbol{\Theta}_{\mathbf{environmental}}$$
$$= \boldsymbol{\Delta}((\mathbf{I} - \boldsymbol{\Omega}_{\mathbf{genetic}})^{-1} + ((\mathbf{I} - \boldsymbol{\Omega}_{\mathbf{environmental}})^{-1})\boldsymbol{\Delta}$$

  - ….

Causal Networks

- In **causal** networks edges are directed paths

$$\widehat{\boldsymbol{\Sigma}} = (\mathbf{I} - \mathbf{B})^{-\mathbf{1}}\boldsymbol{\Phi}(\mathbf{I} - \mathbf{B})^{-\mathbf{1}}$$

That's simply path analyses!

Application 2

Testing Moderated Mediation

Testing Mediation

Mediation:

A is a cause of B, while B is a cause of C

Result: A is a cause of C

More precisely…

Testing Mediation

Mediation:

Differences in A give rise to differences in B, while Differences in B give rise tot differences in C.

Result: Differences in A give rise to differences in C.

To be even more precise replace ‘differences’ by:

- individual differences (subject of my talk)
- intraindivual differences (not subject of my talk)
- group differences (…. dunno….yes and no?)

Testing Mediation

Research question: Does interindividual variable M mediate between interindividual variables X and Y?

Tackle using **the scientific method**:

falsification (rather than verification):

+

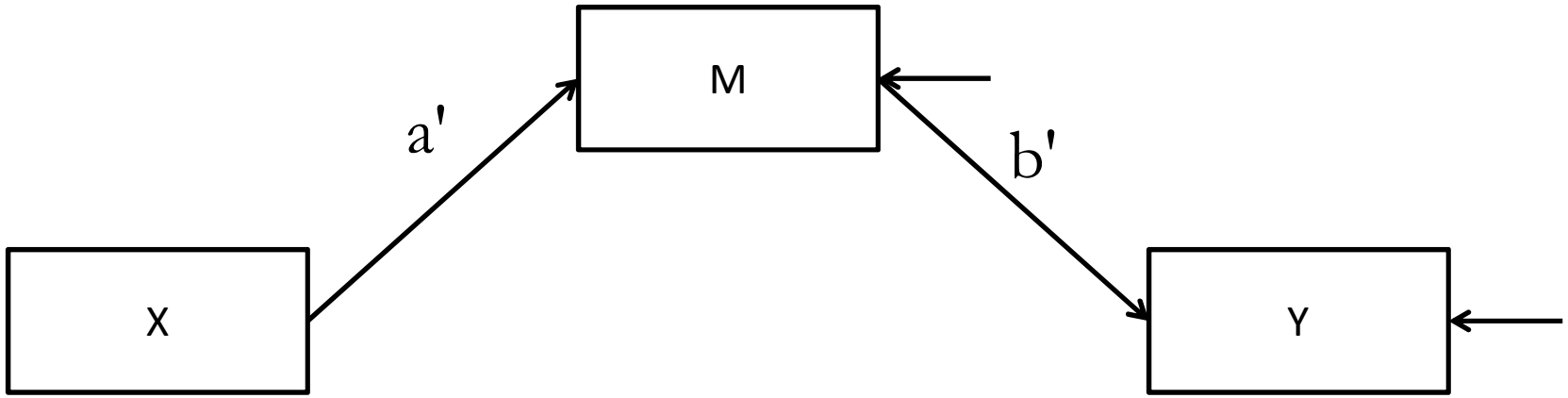
If two competing, falsifiable explanations cannot be rejected

parsimony

Basis of **model selection**!

Testing Mediation

Hypothesis (= stated in the form of a full mediation model)



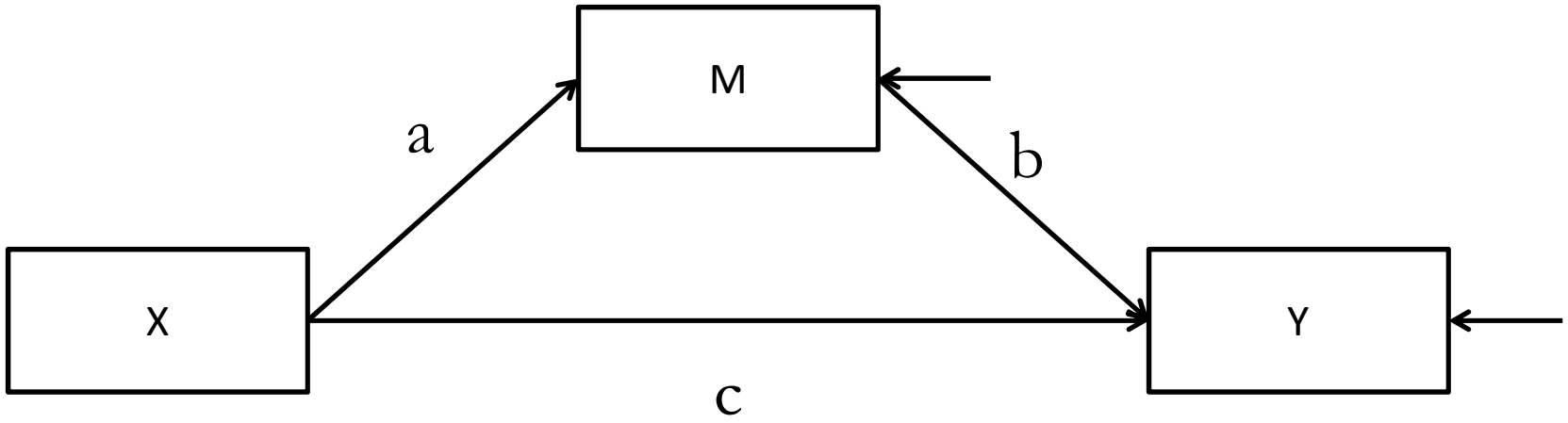
If true, the model explains there is a relation between IV and DV

If true, the indirect effect equals a'*b'

Falsifiable

Testing Mediation

Model used (partial mediation; weaker hypothesis)



If true, the presence of pathways X→M and M→Y give a partial explain of the relation between X and Y

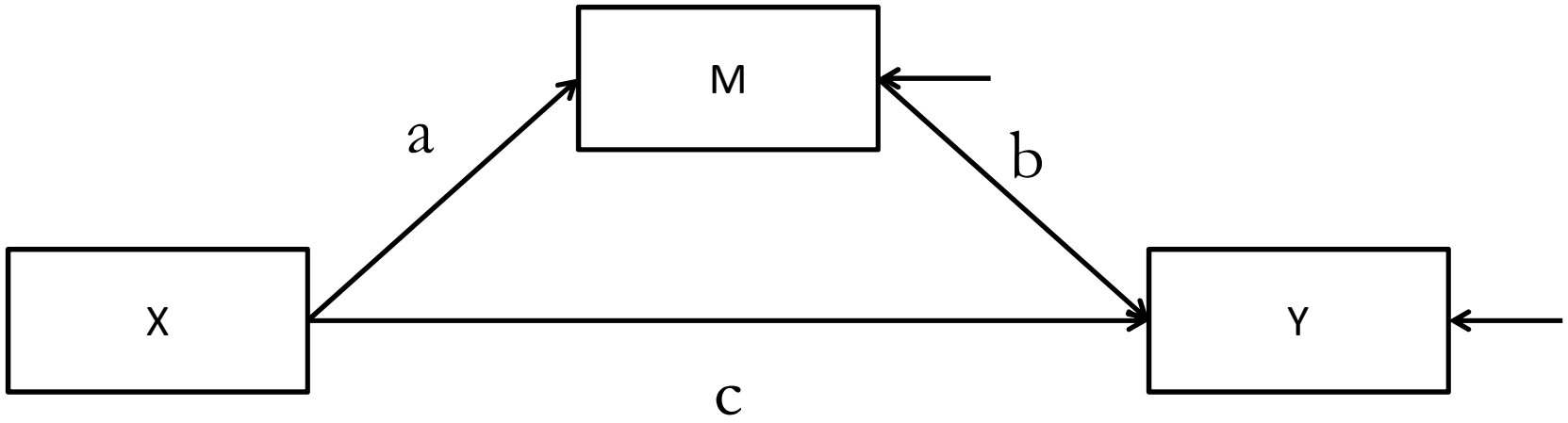
If true, the indirect effect equals a*b

Ok, does not run against the original hypothesis

But unfalsifiable! (saturated)

Testing Mediation

Typical test:



Provided with the unfalsifiable model: 'Determine if 0 lies inside or outside CI interval of a*b'

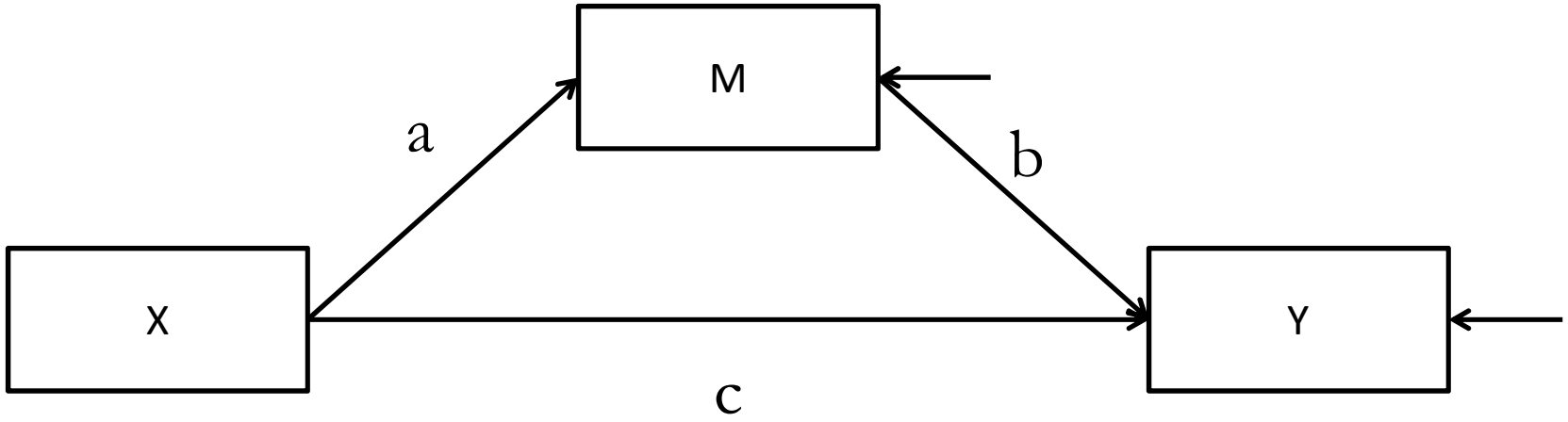
This is not the same as asking 'is my mediation theory (model) (in)correct?'
It is not even the same as asking 'Does a*b lie outside the CI of 0 ($\alpha\beta = 0$)'

Testing Mediation

- **PROVIDED there is mediation**, the result is product a*b
- The aim of traditional mediation tests is awkward
- Hypothesis test:
  - **GIVEN there is mediation**
    - **Assume there is no mediation(??)**
      - But mediaton was the given(!!)
- What I *do* understand:
- a*b is 0 when a = 0 or b = 0 (or both)
  - implying there would be no mediation
- Why not model these hypotheses explicitly??
- Test **WHETHER (we can assume) there is mediation**

Testing Mediation

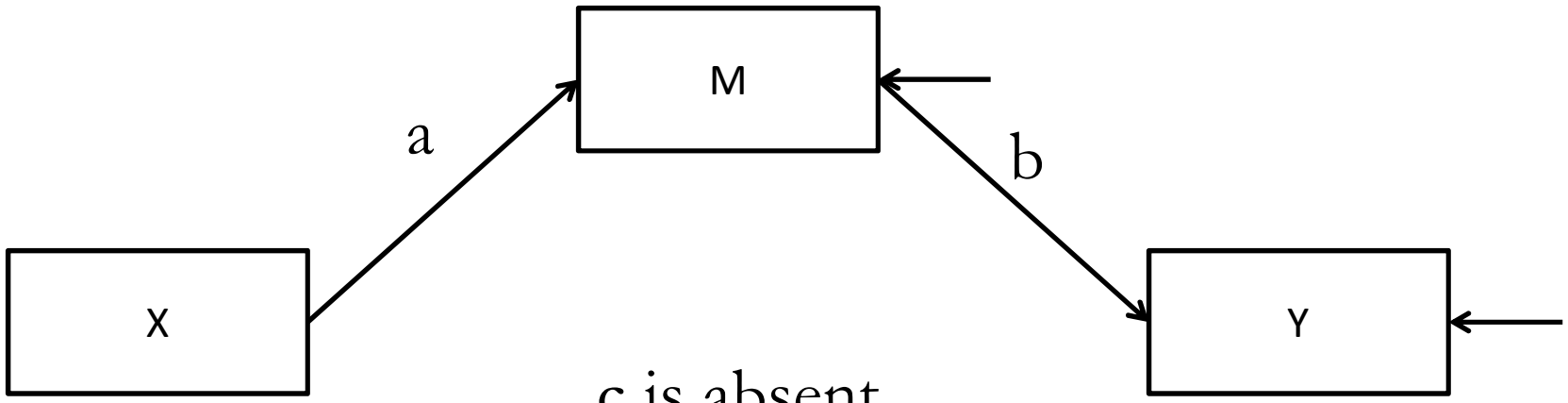
1) Fit the saturated model (partial mediation model)
2) Drop path(s) and/or specify other falsifiable models
3) Compare LL's and select model of preference



Testing Mediation

Competing (**falsifiable**) model:

Is there full mediation? (implying M does mediate)

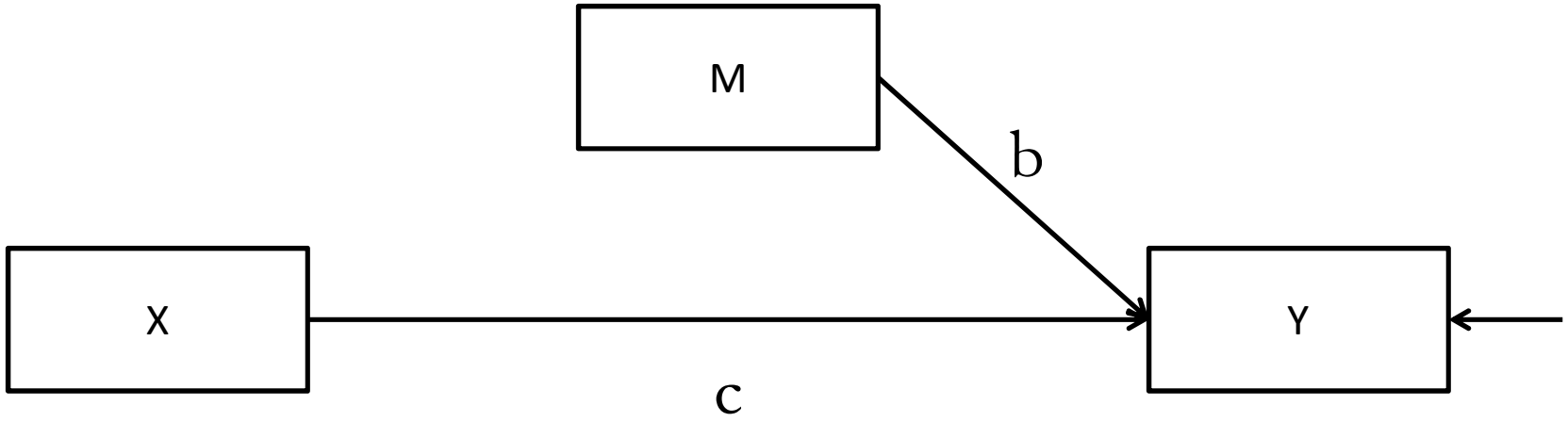


c is absent

Testing Mediation

Competing (**falsifiable**) model

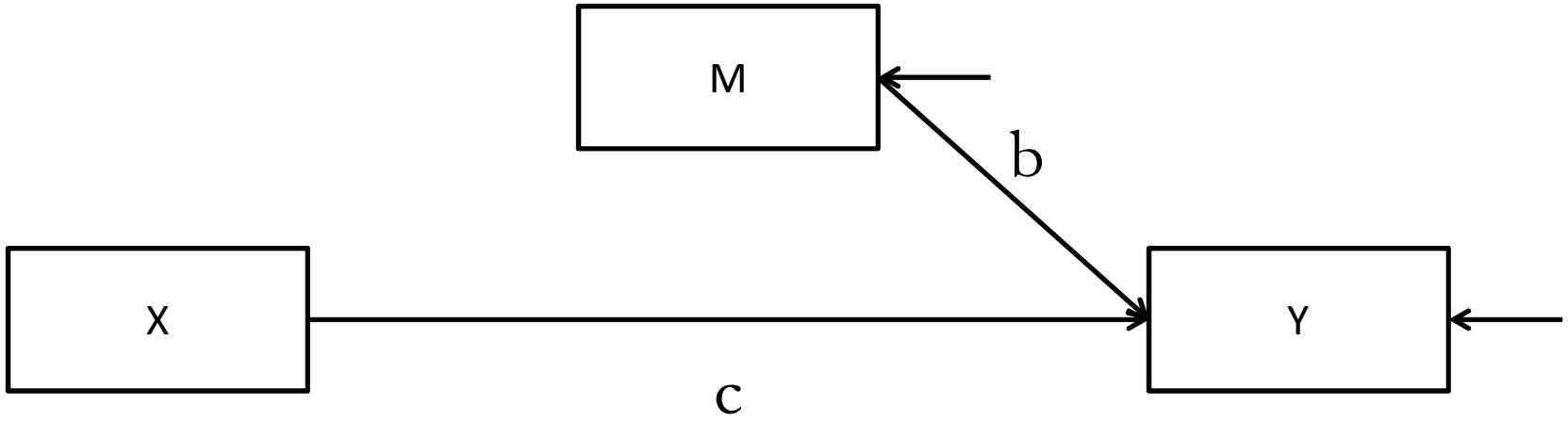
Is a 0? (implying M does not mediate)



Testing Mediation

Competing (**falsifiable**) model

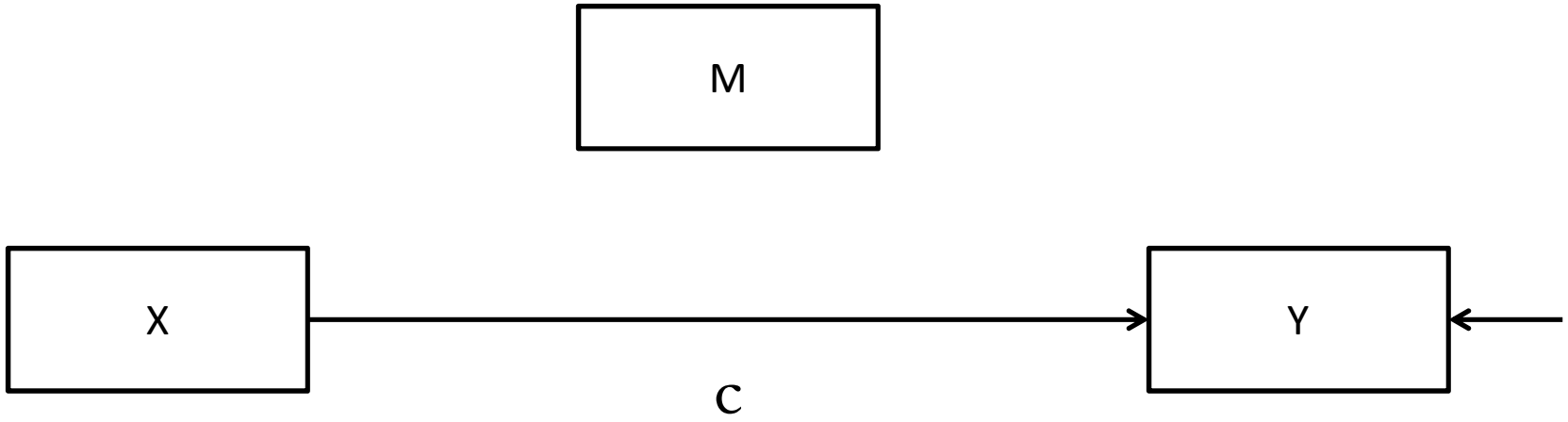
Is b 0? (also implying M does not mediate)



Testing Mediation

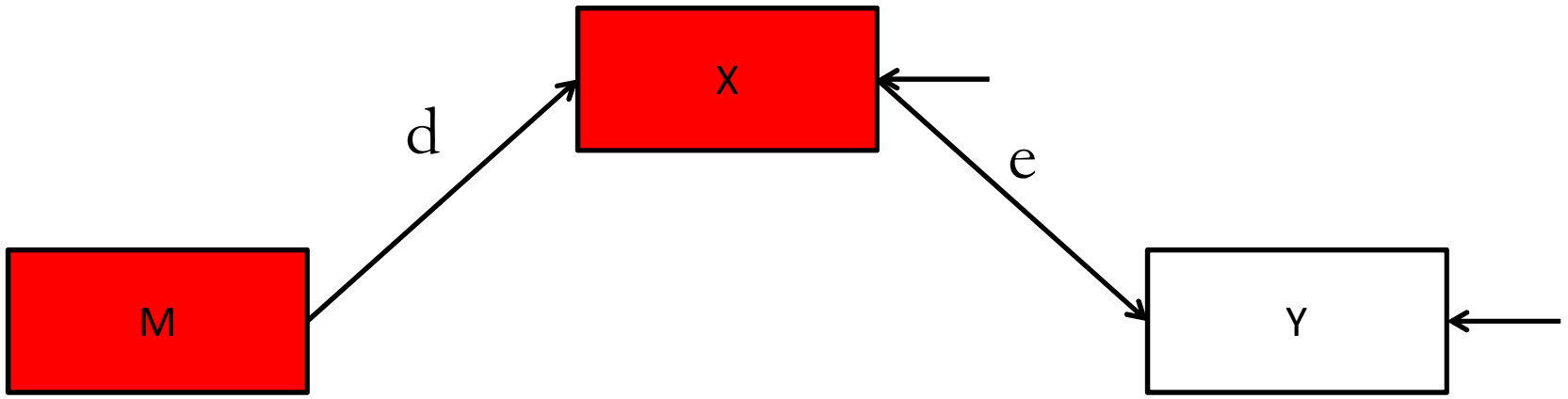
Competing (**falsifiable**) model:

Are both a and b 0? (implying M does not mediate)



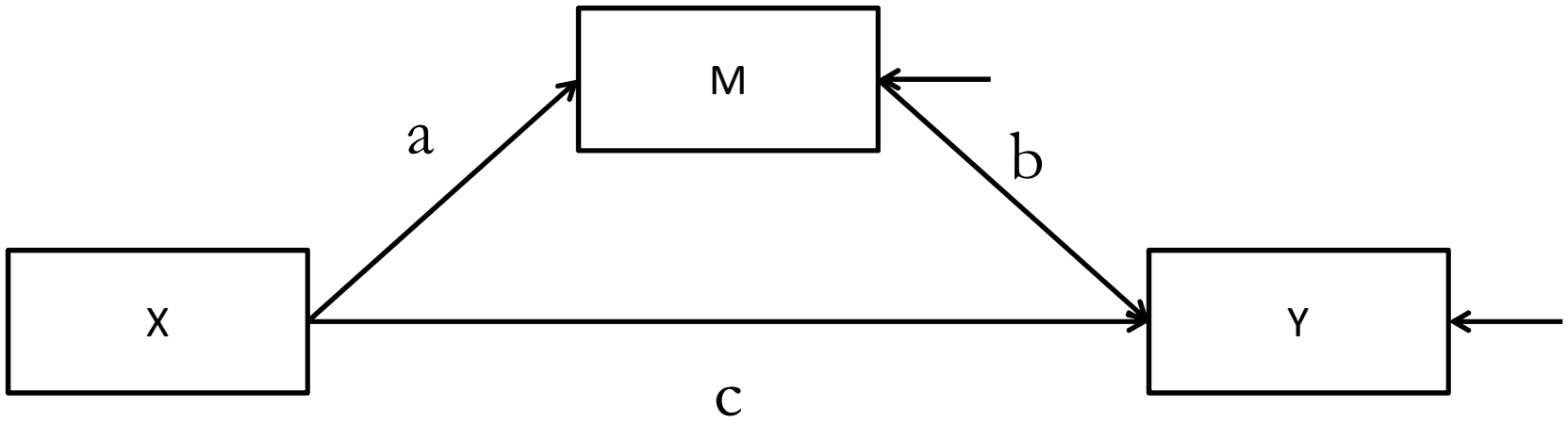
Testing Mediation

Example of another competing (**falsifiable**) model (implying M does not mediate)



Testing Mediation

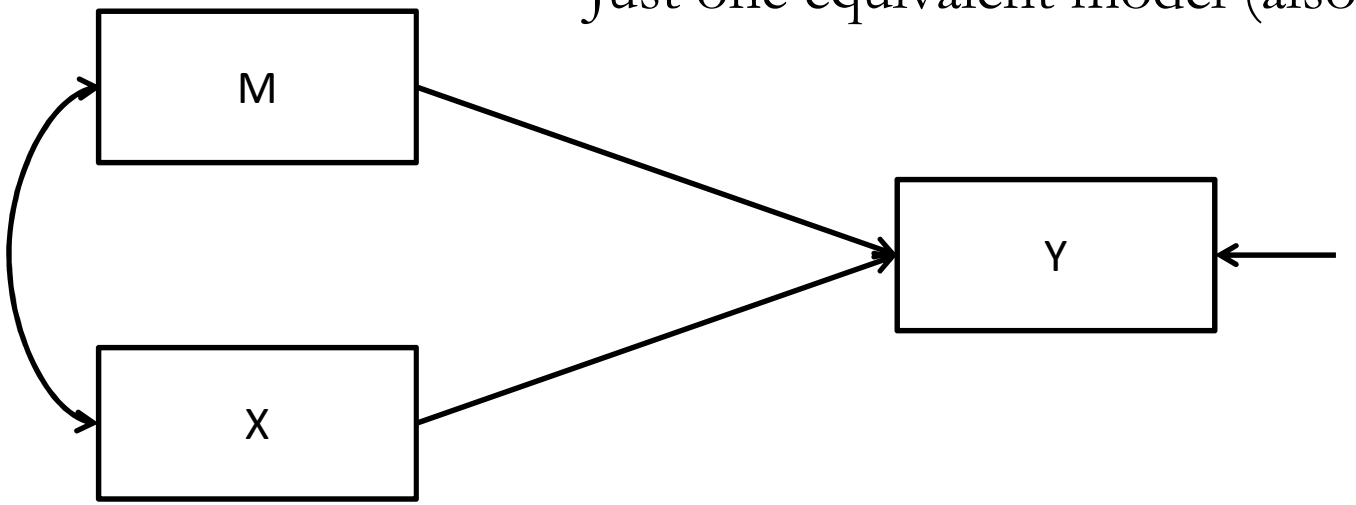
If we end up with the saturated model, we cannot conclude there is mediation(!)



Testing Mediation

If we end up with the saturated model, we cannot conclude there is mediation(!)

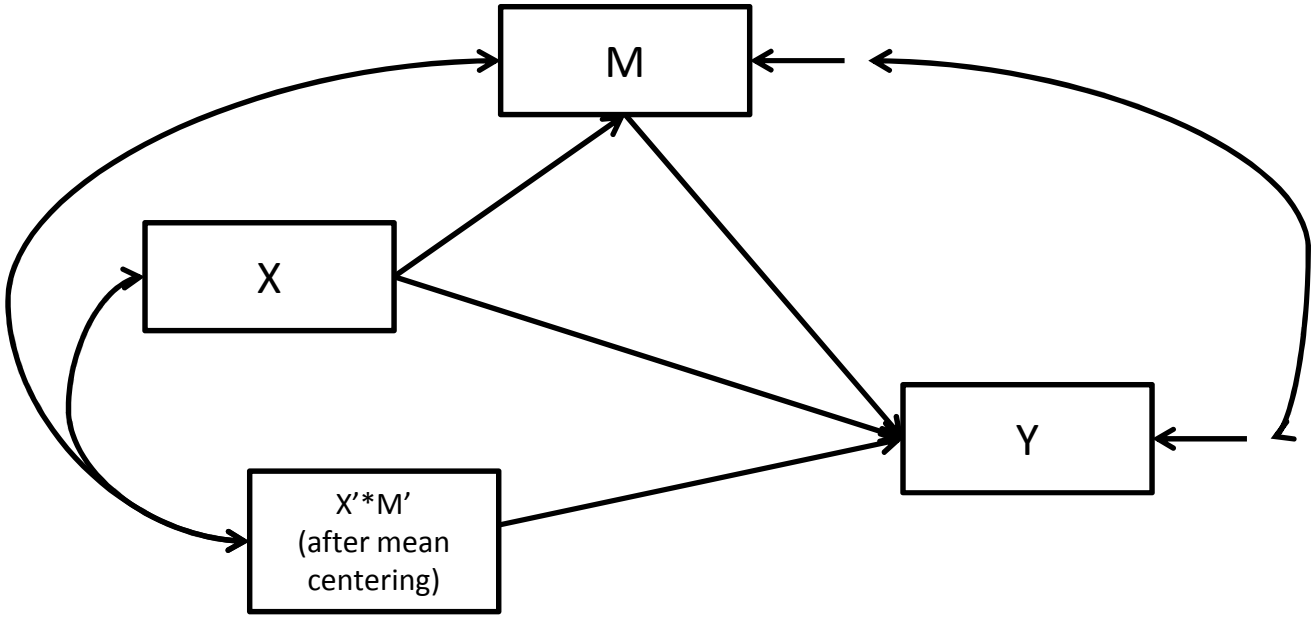
Just one equivalent model (also unfalsifiable)



Testing Moderated Mediation (SPSS macro PROCESS)

What is the substantial
interpretation of variable X*M?
Why does X*M have a direct effect?

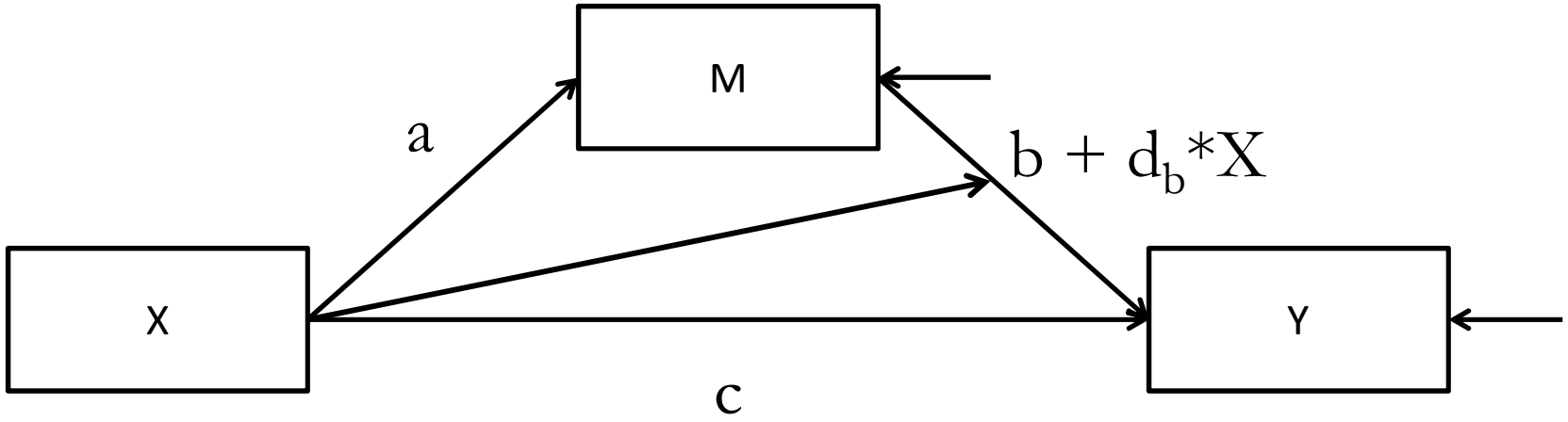
Saturated **(unfalsifiable!)**



Testing Moderated Mediation

1) Define the saturated model and e**xplicate which variable moderates which path**,

e.g. the independent variable X moderates path b

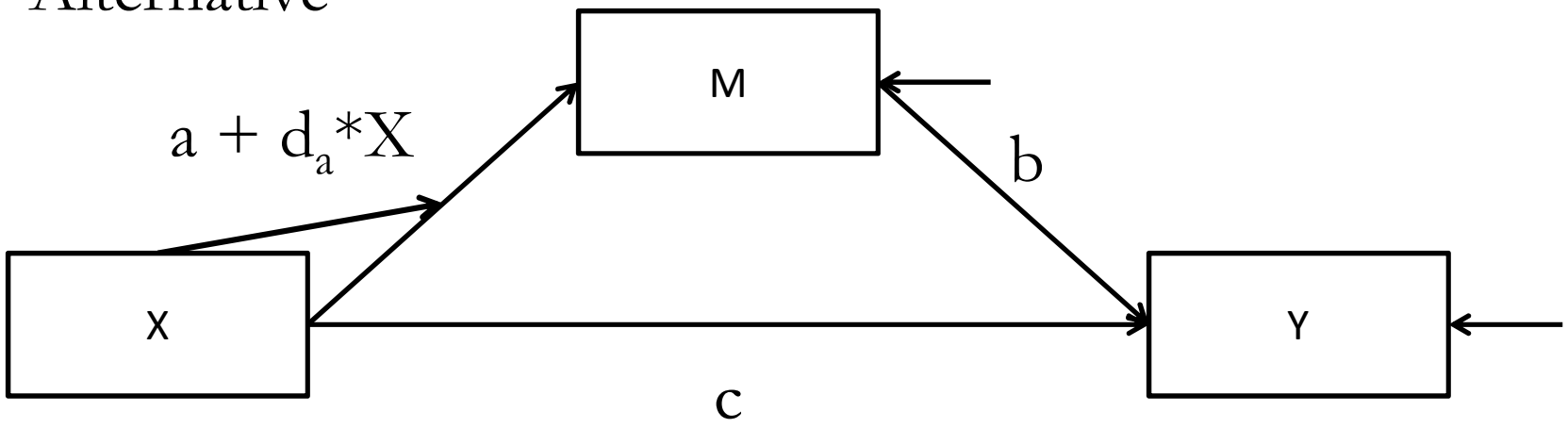


Works in OpenMx through the use of definition variables

Testing Moderated Mediation

1) Define the saturated model and e**xplicate which variable moderates which path**,

Alternative

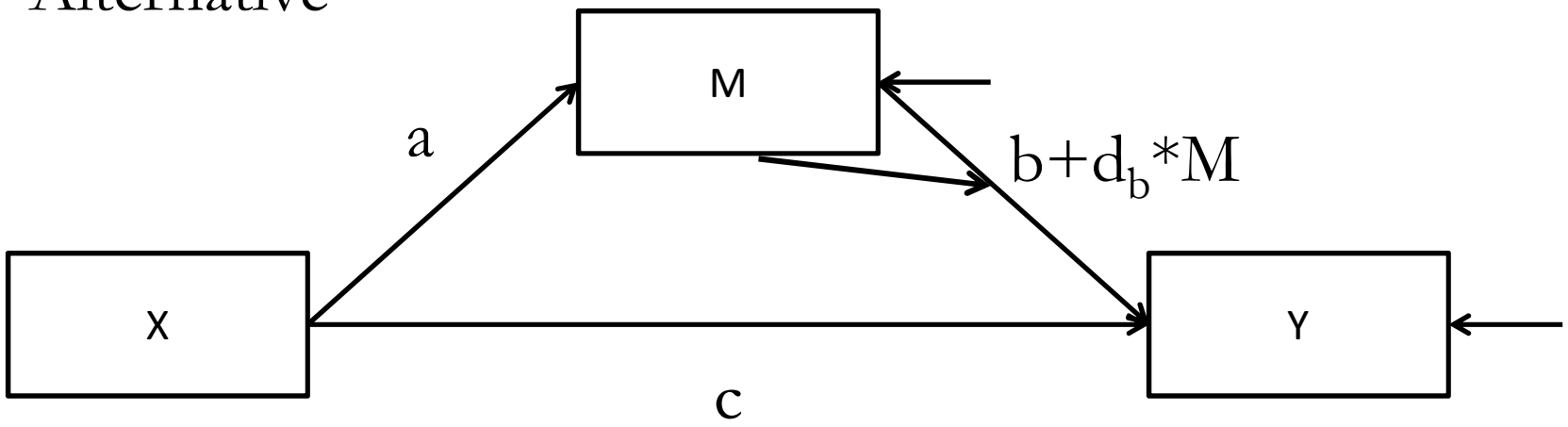


Works in OpenMx through the use of definition variables

Testing Moderated Mediation

1) Define the saturated model and e**xplicate which variable moderates which path**,

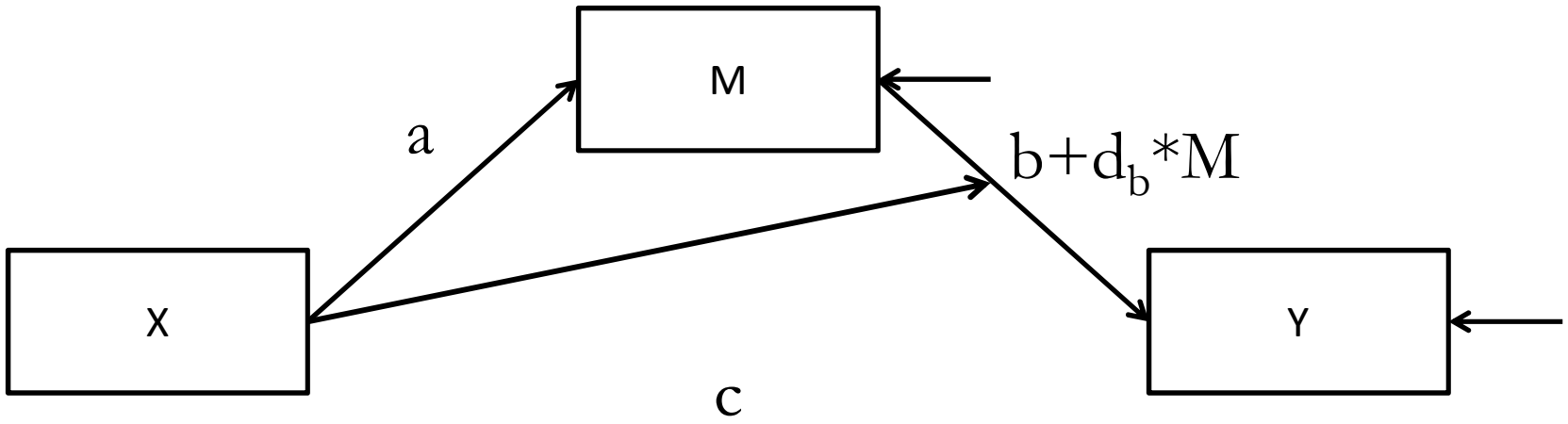
Alternative



Works in OpenMx through the use of definition variables

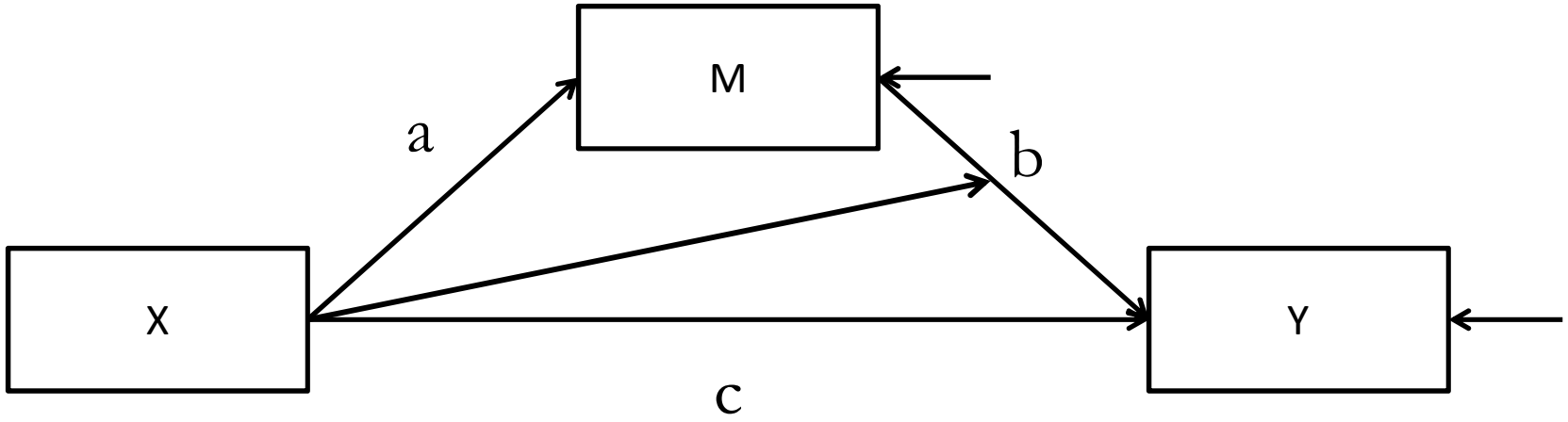
Testing Moderated Mediation

Suppose this reflects my hypothesis



Testing Moderated Mediation

1) Fit this saturated model
2) Drop path(s) and/or specify other falsifiable models
3) Compare LL's and select model of preference

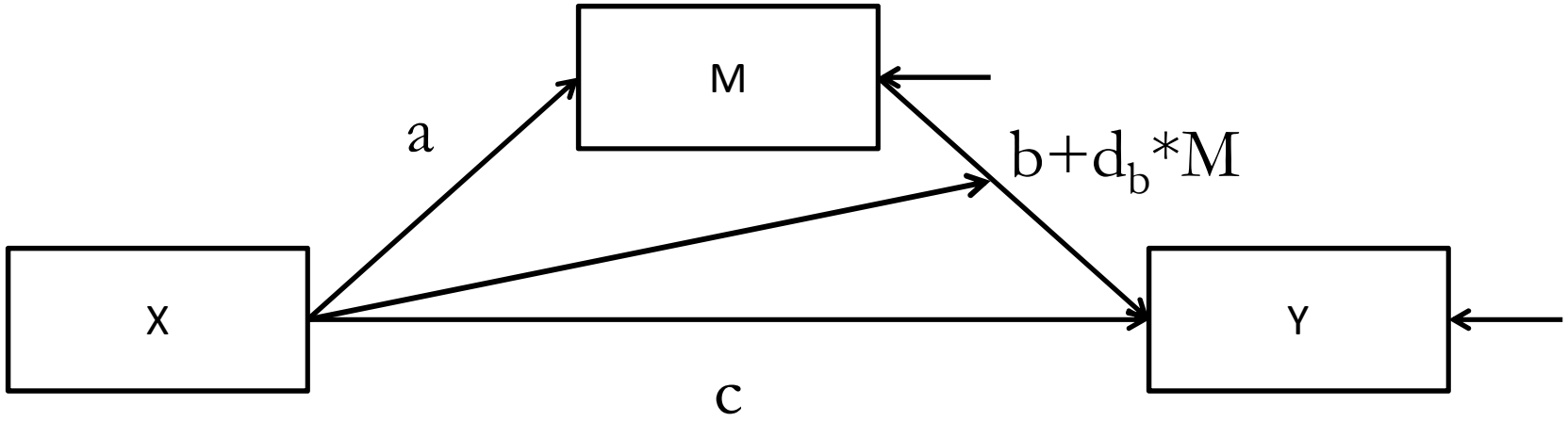


Hence the same steps as in normal mediation

Testing Moderated Mediation

Suppose this was the best model (we cannot conclude there is mediation though; model equivalence)

4) Drop the moderation to test if moderation is a necessary assumption



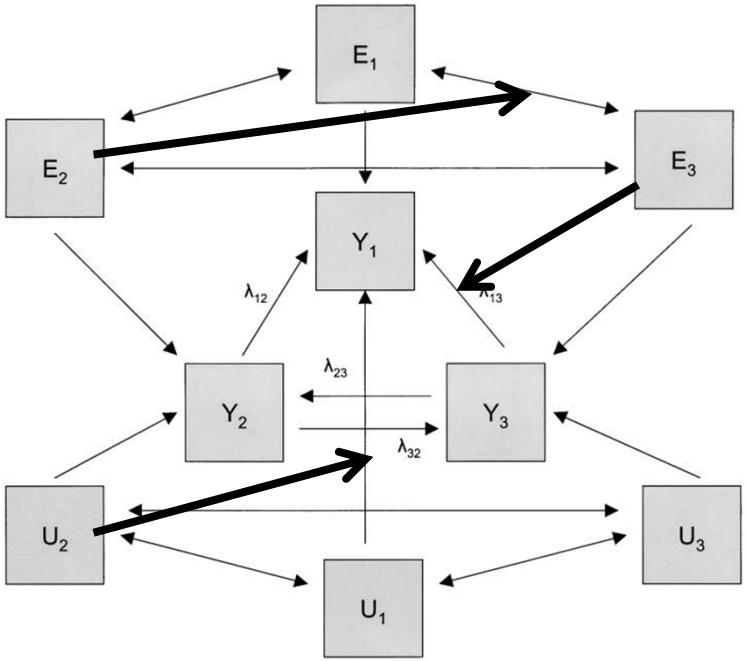
The saturated model wasn't that saturated!
With respect to variance-covariance, yes
With respect to homo/heteroskedasticity, no

Testing Moderated Mediation

Possible extentions

Moderation in more complex models

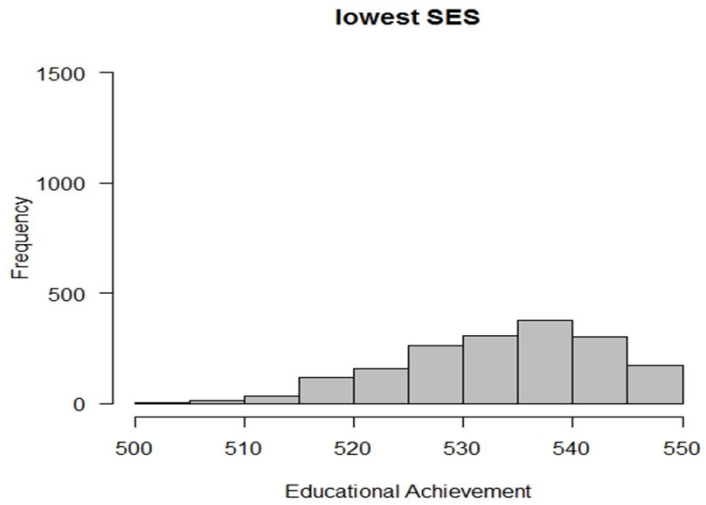
- e.g. (causal?) network models

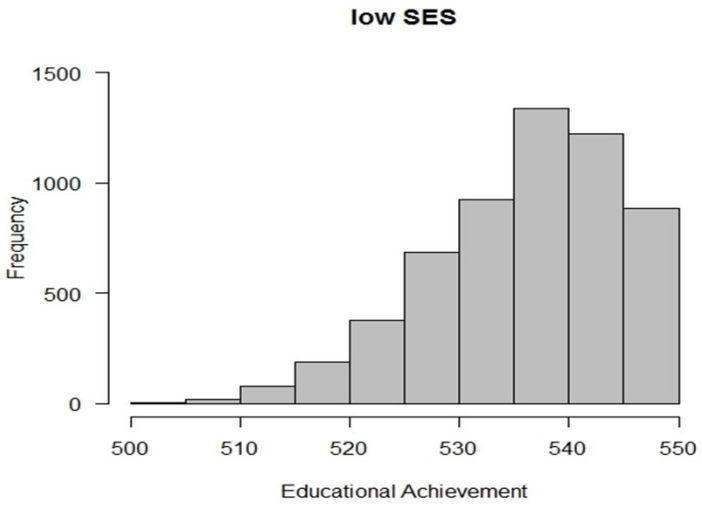


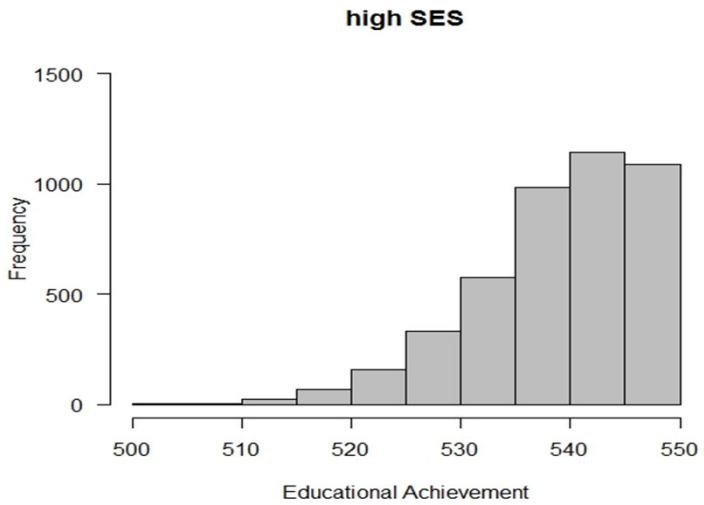
Application 3

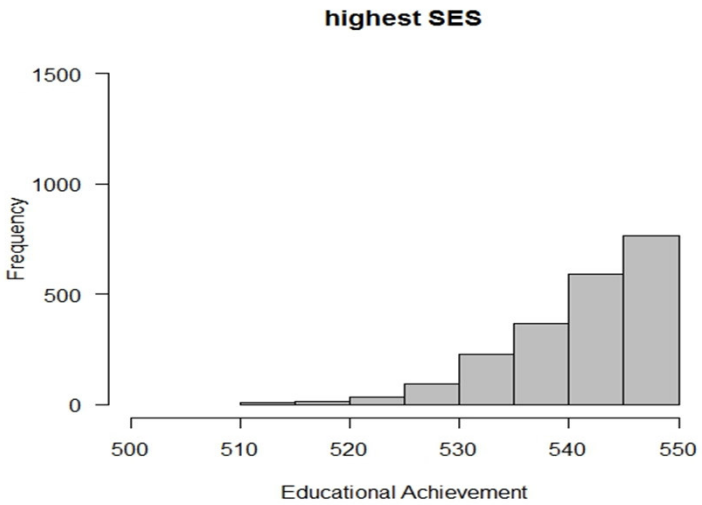
a Means to Correct for the Effects of Censoring

Correction for censoring

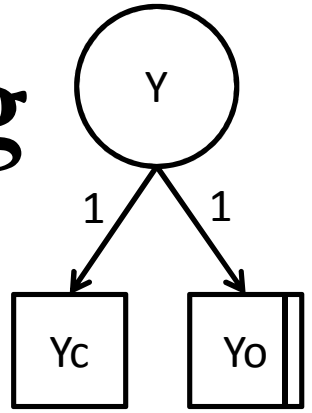








Correction for censoring

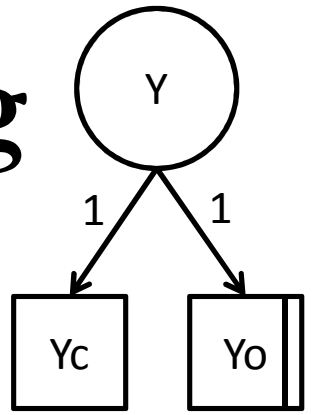


https://openmx.ssri.psu.edu/thread/1428

Suggestion by Mike Neale:

- Model a single latent variable which causes, with path=1, two variables
- one is ordinal and one is continuous
- For those individuals with censored scores, they are missing (NA) on the continuous variable, and vice-versa
- for those with the continuous variable scored (<[threshold]) - they are marked as missing on the ordinal variable.
- I think this is simplest but tbh I've **never tried** it."

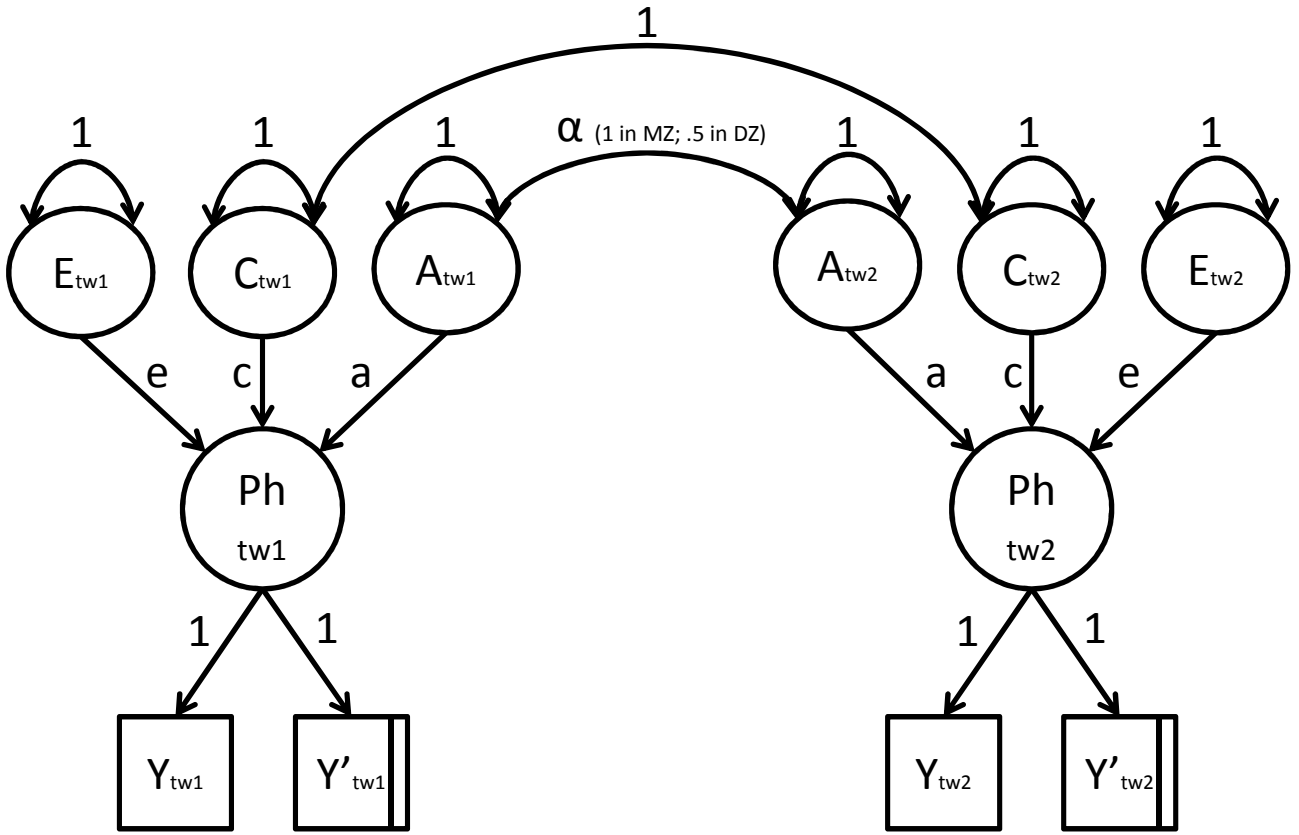
Correction for censoring



Checked performance

- By simulations
  - with and without extra missings
- With results from a FORTRAN tool once developed by prof. Conor Dolan
  - exact same results. i.e. LL, AIC etc.
- **Conclusion: Tried and works well!**

Correction for censoring



Extension: Implementation in behavior genetic modeling

Simulations told: The method works well!

Correction for censoring

My Research Master student, Maarten Schouten, continued the simulations and just finished his thesis

Eliminating the Biasing Effect of Censoring on Parameter Estimates in Biometrical Structural Equation Modeling

M. Schouten

11210729

1st assessor: dr. K.J. Kan

2nd assessor: dr. D. Molenaar

Universiteit van Amsterdam

Correction for censoring

Background:

- Standard twin models assume phenotypes are normally distributed.
- Fitting these models on censored data can yield biased parameter estimates
- Data transformation does not eliminate bias
- Categorical modeling reduces power to discriminate among nested models

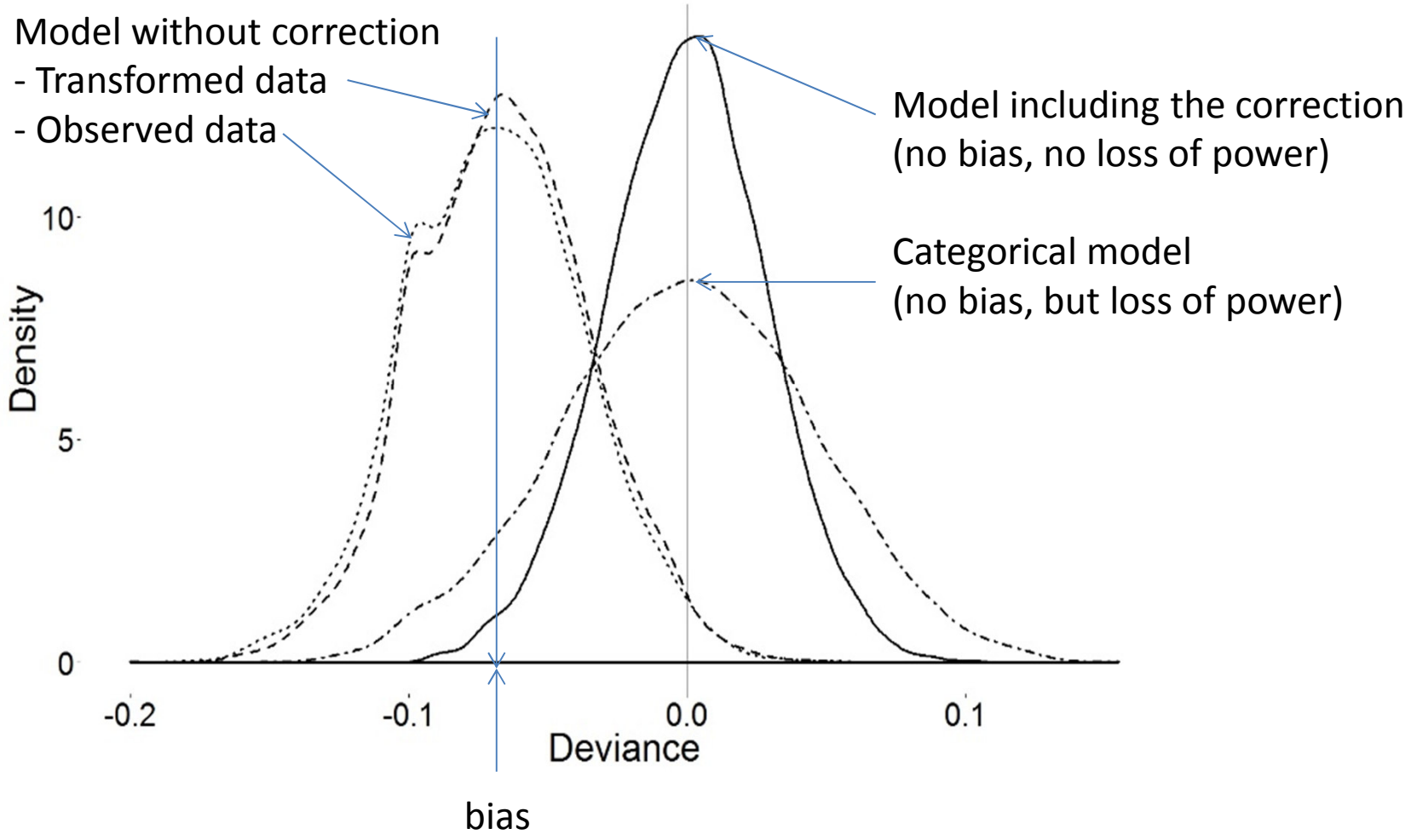
Correction for censoring

Objectives:

- To reassess the biasing effect of censored data on parameter estimates
- To assess the performance of the model that accounts for censoring

In comparison to the aforementioned models

Correction for censoring



Correction for censoring

Conclusions:

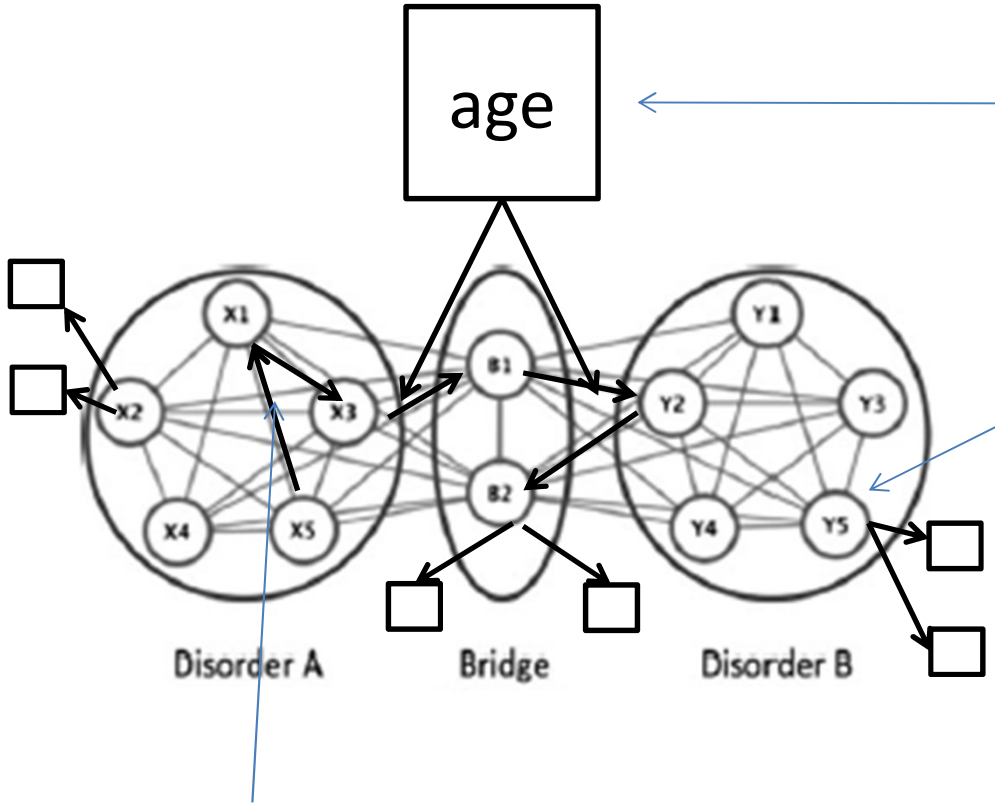
The censoring model outperforms both the continuous model, since it diminishes bias, and the categorical model, as it retains power.

Three applications

Conclusions

Conclusion

'How can I test for age changes in comorbidity?'



Moderator of pathways between disorder specific symtoms and bridge symptoms

Censored variables

(Preferably causal) pathways

Conclusion

Yves yesterday:

"We may not *need* SEM after all" (in certain cases)

Today:

"SEM is such a handy tool!" (in many cases)

SEM is very flexible

- To a large extent because it provides the opportunity to implement **(in)equality constraints**
- OpenMx provides the opportunity to use matrix algebraic expressions
  - OpenMx is not that user friendly though
  - Implementation of algebraic expressions in lavaan?