The analysis of individual and average causal effects: Basic principles and some applications

EUROPEAN SCIENCE FOUNDATION
Programme 'Quantitative Methods in the Social Sciences' (QMSS)

Seminar:
'Theory-Driven Evaluation and Intervention Studies in the Social Sciences',

28-29 September 2006, Nicosia, Cyprus

Rolf Steyer

University of Jena (Germany)

Institute of Psychology
Department of Methodology and Evaluation Research
Email: rolf.steyer@uni-jena.de

www.uni-jena.de/svw/metheval

Overview

- Individual and average causal effects (Neyman, Rubin)
- Motivation: The Simpson Paradox
- Pre-Post Design with Control Group for the Analysis of Intervention Effects
- Designs for the Analysis of Individual Causal Effects
  - Example with an Intervention
  - Example with Method Effects
- Conclusions

The Simpson Paradox

Table 1. Total Sample

	treatment		
success	yes ($X = 1$)	no ($X = 0$)	total
yes ($Y = 1$)	500	600	1100
no ($Y = 0$)	500	400	900
	1000	1000	2000

The Simpson Paradox

Table 2. Males

	treatment		
success	yes $(X = 1)$	no $(X = 0)$	total
yes $(Y = 1)$	300	75	375
no $(Y = 0)$	450	175	625
	750	250	1000

The Simpson Paradox

Table 3. Women

	treatment		
success	yes $(X = 1)$	no $(X = 0)$	total
yes $(Y = 1)$	200	525	725
no $(Y = 0)$	50	225	275
	250	750	1000

The Simpson Paradox

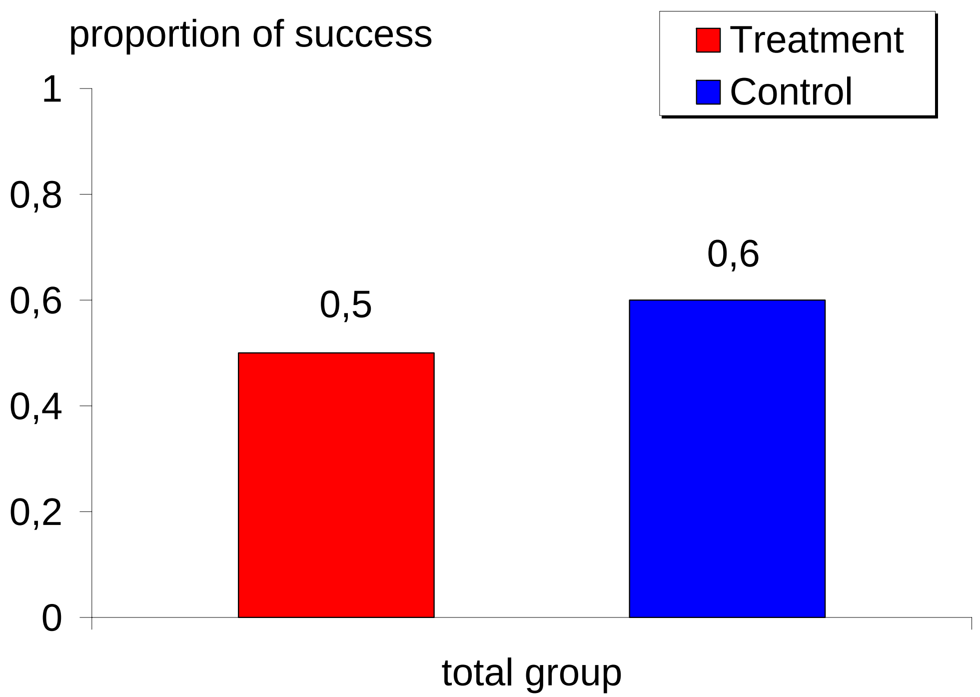
Table 2. Males

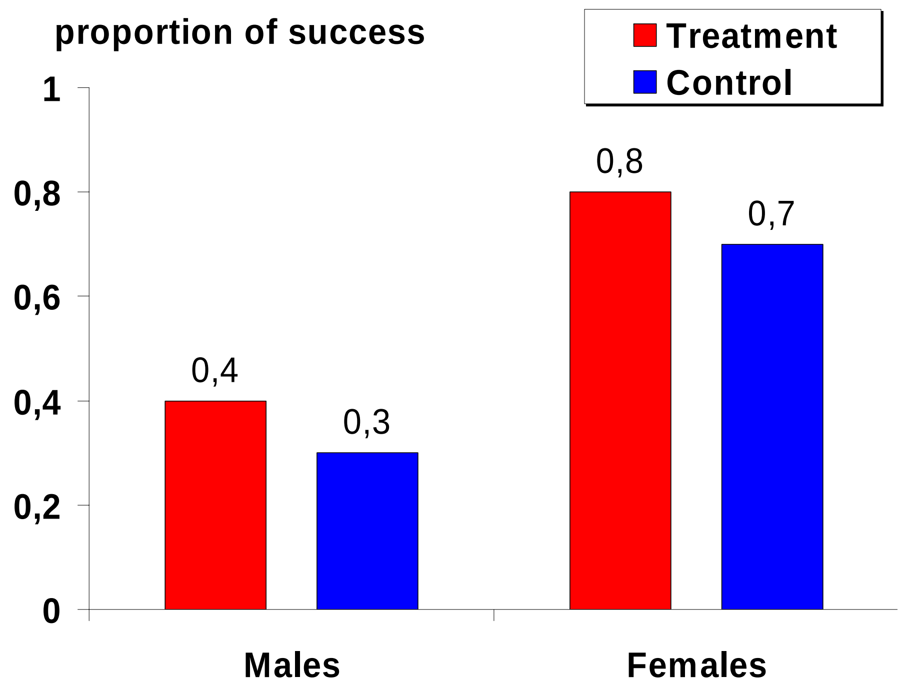
	treatment		
success	yes $(X = 1)$	no $(X = 0)$	total
yes $(Y = 1)$	300	75	375
no $(Y = 0)$	450	175	625
	750	250	1000

Table 3. Women

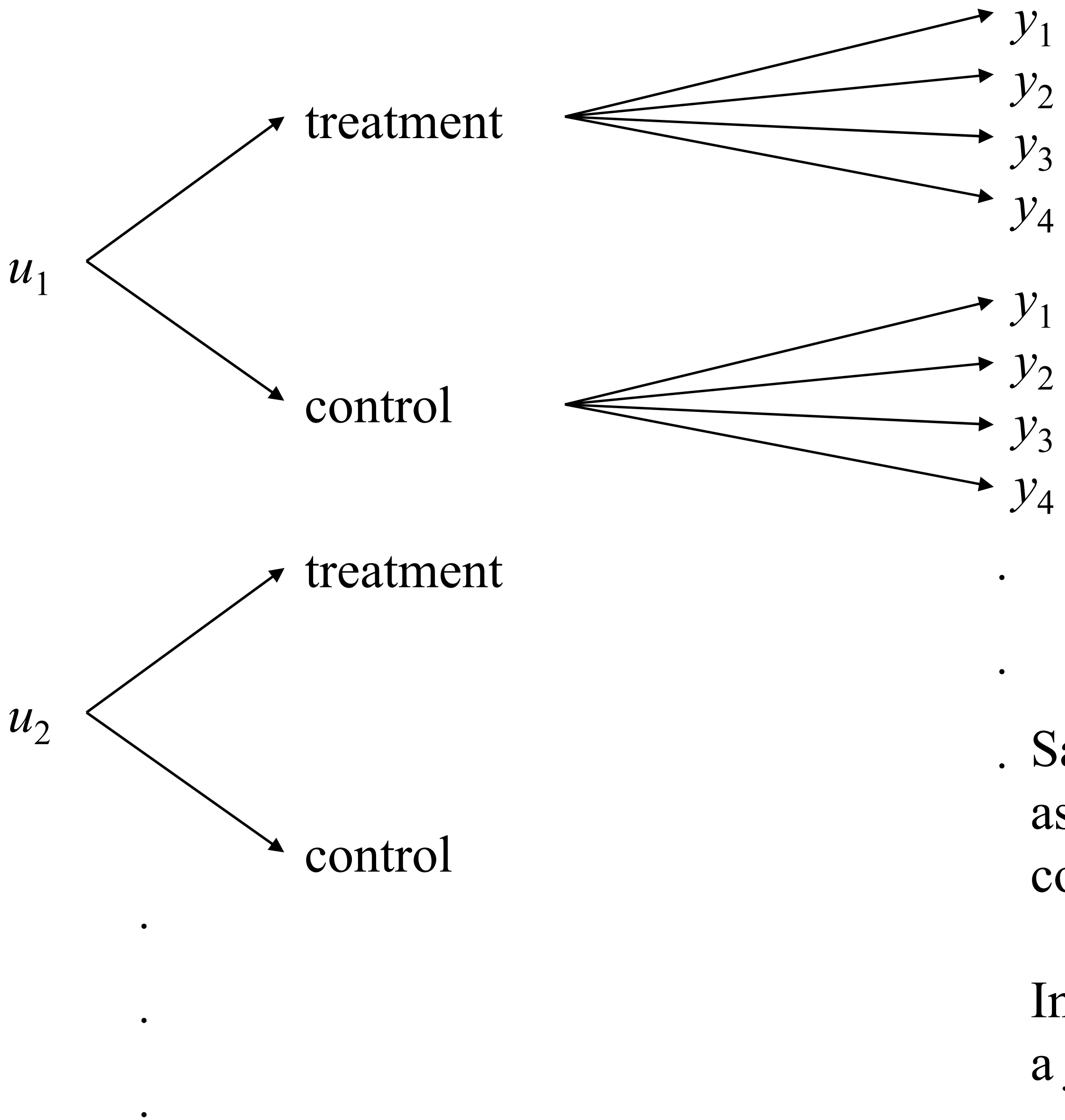
	treatment		
success	yes $(X = 1)$	no $(X = 0)$	total
yes $(Y = 1)$	200	525	725
no $(Y = 0)$	50	225	275
	250	750	1000

The Simpson Paradox





The single-unit trial



Sample a person u, register her assignment to one of the treatment conditions and observe her outcome y.

In this single-unit trial U, X, and Y have a joint distribution

Individual and average causal effects (Neyman, Rubin)

Unit	$P(U=u)$ Sampling probability	$\tau_0(u) = E(Y_0 \mid U = u)$ True outcome under control	$\tau_1(u) = E(Y_1 \mid U = u)$ True outcome under treatment	$ICE_{1\text{-}0}(u) = E(Y_1 \mid U = u) - E(Y_0 \mid U = u)$ Individual causal effect
u_1	1/8	68	82	14
u_2	1/8	81	89	8
u_3	1/8	89	101	12
u_4	1/8	102	108	6
u_5	1/8	112	118	6
u_6	1/8	119	131	12
u_7	1/8	131	139	8
u_8	1/8	138	152	14
		$E(\tau_0) = 105$	$E(\tau_1) = 115$	$ACE_{1\text{-}0} = 10$

Define the true-outcome variables as follows:

$$\tau_0(u) := E(Y_0 \mid U = u)$$

and

$$\tau_1(u) := E(Y_1 \mid U = u)$$

$$\tau_{1\text{-}0}(u) := \tau_1(u) - \tau_0(u)$$

$$= \textit{individual causal effect of unit } u$$

$$ACE_{1\text{-}0} = E(\tau_{1\text{-}0}) = E(\tau_1) - E(\tau_0)$$

$$=: \textit{average causal effect}$$

Individual, conditional and average causal effects (Neyman, Rubin)

Person	$P(U{=}u)$ Sampling probability	$\tau_0(u) = E(Y \mid X = 0, U = u)$ True outcome under control	$\tau_1(u) = E(Y \mid X = 1, U = u)$ True outcome unter treatment	$\tau_{1\text{-}0}(u) = E(Y \mid X = 1, U = u) - E(Y \mid X = 0, U = u)$ Individual causal effect	*Individual treatment Probability* $P(X{=}1 \mid U{=}u)$
u_1	1/8	68	82	14	8/9
u_2	1/8	81	89	8	7/9
u_3	1/8	89	101	12	6/9
u_4	1/8	102	108	6	5/9
u_5	1/8	112	118	6	4/9
u_6	1/8	119	131	12	3/9
u_7	1/8	131	139	8	2/9
u_8	1/8	138	152	14	1/9

$E(\tau_0) = 105$ $\qquad$ $115 = E(\tau_1)$

$ACE_{1\text{-}0} := E(\tau_{1\text{-}0}) = E(\tau_1) - E(\tau_0) = 10$

$PFE_{1\text{-}0} := E(Y \mid X = 1) - E(Y \mid X = 0) = -13.33$

$$E(Y \mid X{=}j) = \sum_u E(Y \mid X{=}j, U{=}u)\, P(U{=}u \mid X{=}j)$$

$$E(\tau_j) = \sum_u E(Y \mid X{=}j, U{=}u)\, P(U{=}u)$$

Bias Theorem

Bias Theorem. (i) Let X and Y be the random variables defined in the single-unit trial. Then

$$PFE_{j-0} = ACE_{j-0} + baseline\ bias_{j-0} + effect\ bias_{j-0}, \quad \text{for each } j = 1, \ldots, J,$$

where

$$baseline\ bias_{j-0} := E(\tau_0 \mid X = j) - E(\tau_0 \mid X = 0)$$

and

$$effect\ bias_{j-0} := E(\tau_{j-0} \mid X = j) - ACE_{j-0}.$$

Three Design Types

Between-Group Designs

Pre-post Designs (not used at all in the Neyman-Rubin tradition)

Between-Group Designs with Pre-Post Measures (only the between group comparisons are used in the Neyman-Rubin tradition)

Utilizing Pre-Post Designs for the Analysis of Individual Effects

Pre-post Designs and Between-Group Designs with Pre-Post Measures can be used to analyze not only average but also *individual causal effects*.

The crucial asssumption is that the *individual pretest distribution* is the same as the *individual posttest* (outcome variable) *distribution under control* (no treatment).

Theorem 1 (Sufficient conditions for unbiasedness of the (conditional) prima facie effects)

If $X \perp U$, then the regression $E(Y|X)$

and its values $E(Y|X = x)$ are unbiased

If $X \perp U \mid Z$, then the regression $E(Y|X, Z)$

and its values $E(Y|X = x, Z = z)$ are unbiased

(There are also other sufficient conditions for unbiasedness.)

Theorem 2

Suppose we have just 2 treatment conditions $X = 0$ and $X = 1$, then

$$E(Y \mid X, Z) = g_0(Z) + g_{1\text{-}0}(Z) \cdot X$$

If $E(Y \mid X, Z)$ is unbiased, then the values of $g_{1\text{-}0}(Z)$ are the average causal effects of X on Y given $Z = z$, and $E[g_{1\text{-}0}(Z)]$ is the average causal effect.

Standard research questions in the analysis of causal effects

The standard research questions are:

- What are the conditional effects of treatment as compared to the control given *Z*?

  Hence, we want to

  (a) **estimate the effect function** $g_{1\text{-}0}(Z)$ and

  (b) test H_0: $g_{1\text{-}0}(Z) = \gamma_{00}$ (constant) **no interaction**

- What is the average effect of treatment as compared to the control?

  Hence, we want

  (c) to **estimate the average effect** $E[g_{1\text{-}0}(Z)]$ and

  (d) test H_0: $E[g_{1\text{-}0}(Z)] = 0$ **no average effect**

Generalization to *J* + 1 Treatment Conditions

For $J + 1$ treatment conditions, the covariate treatment regression can always be written:

$$E(Y \mid X, Z) \;=\; g_0(Z) + g_{1\text{-}0}(Z) \cdot I_{X=1} + \ldots + g_{J\text{-}0}(Z) \cdot I_{X=J}$$

Average effects in the analysis of causal effects

Suppose $g_{1\text{-}0}(Z)$ is a linear function $\gamma_{10} + \gamma_{11} \cdot Z$

$$E\,[g_{1\text{-}0}(Z)] = E\,[\gamma_{10} + \gamma_{11} \cdot Z\,] = \gamma_{10} + \gamma_{11} \cdot E(Z\,)$$

H_0: $E[g_{1\text{-}0}(Z)] = \gamma_{10} + \gamma_{11} \cdot E(Z\,) = 0$ **no average effect**

If $E(Z\,)$ has to be estimated: this involves a *nonlinear hypothesis*

Types of covariates

The covariate Z can be:

- manifest discrete
- manifest continuous
- latent discrete
- latent continuous
- univariate or multivariate
- stochastic
- fixed

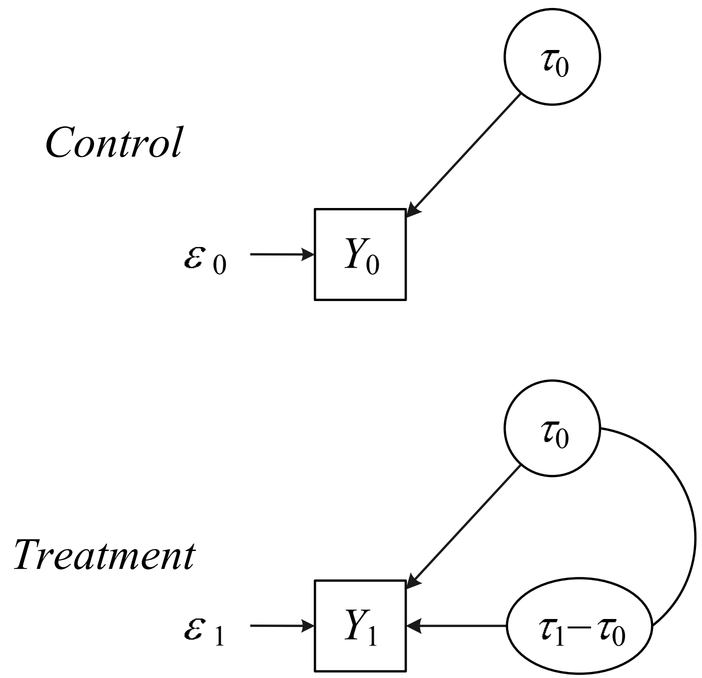
Example (using EffectLite): Intelligence training

Pre-Post Design with Control Group for the Analysis of Intervention Effects

Between-Group Design, no Pretest

$$Y_0 = \tau_0 + \varepsilon_0$$
$$Y_1 = \tau_0 + \tau_1 - \tau_0 + \varepsilon_1$$

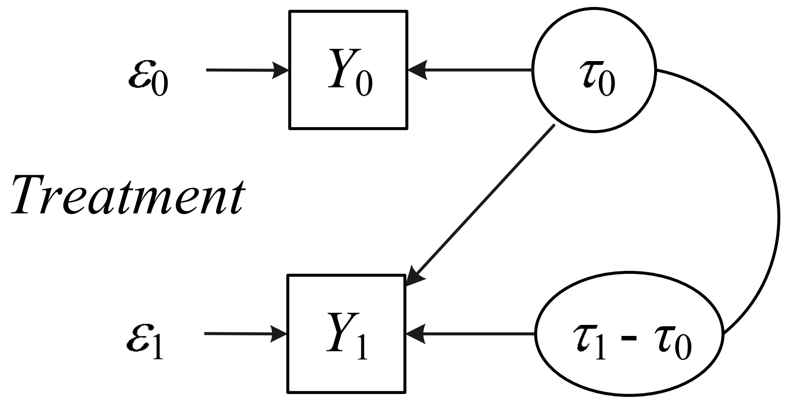
Unit	$P(U{=}u)$ Sampling probability	$\tau_0(u) = E(Y_0 \mid U = u)$ True outcome under control	$\tau_1(u) = E(Y_1 \mid U = u)$ True outcome under treatment	$ICE_{1\text{-}0}(u) = E(Y_1 \mid U = u) - E(Y_0 \mid U = u)$ Individual causal effect
u_1	1/8	68	82	14
u_2	1/8	81	89	8
u_3	1/8	89	101	12
u_4	1/8	102	108	6
u_5	1/8	112	118	6
u_6	1/8	119	131	12
u_7	1/8	131	139	8
u_8	1/8	138	152	14
		$E(\tau_0) = 105$	$E(\tau_1) = 115$	$ACE_{1\text{-}0} = 10$



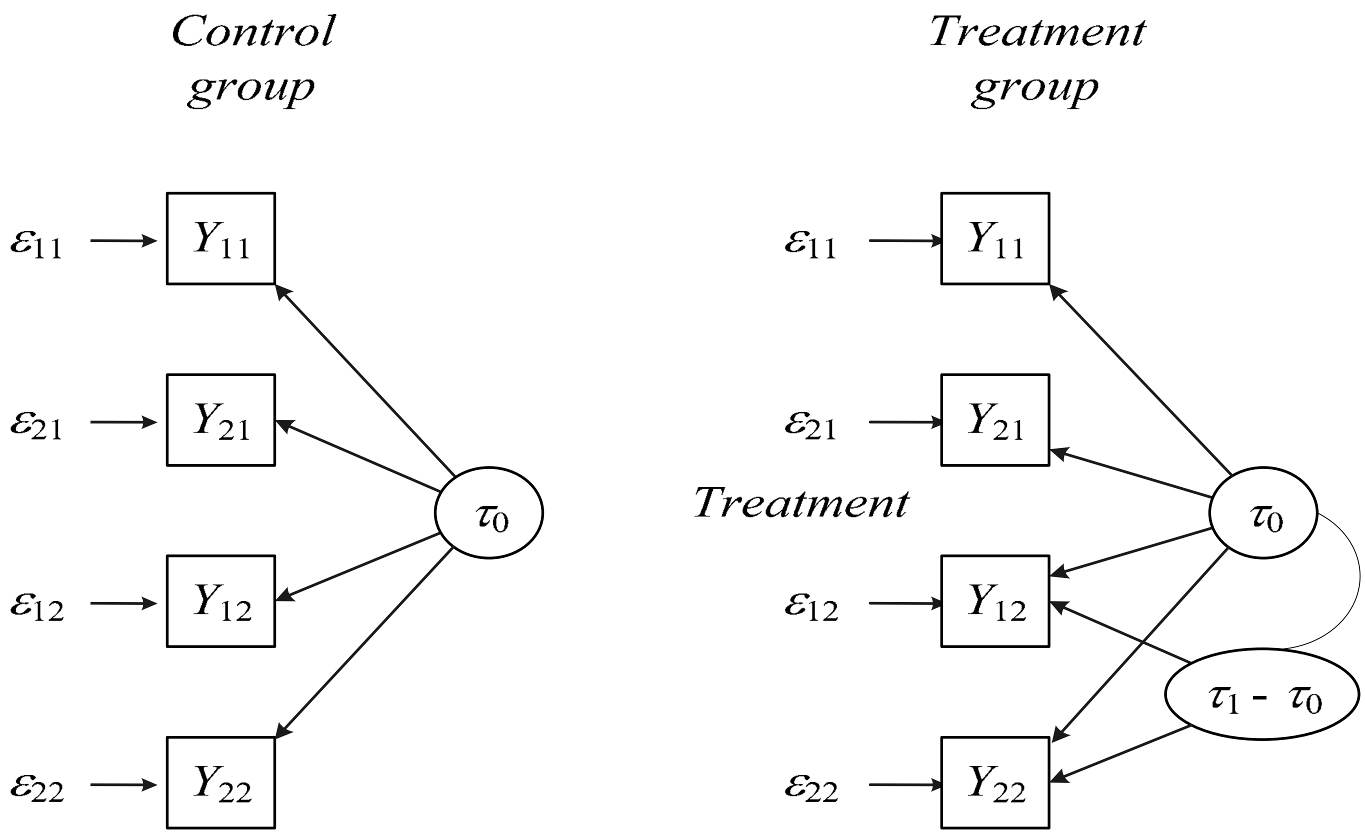
Pre-Post Design, no Control Group

Unit	$P(U=u)$ Sampling probability	$\tau_0(u) = E(Y_0 \mid U = u)$ True outcome under control	$\tau_1(u) = E(Y_1 \mid U = u)$ True outcome under treatment	$ICE_{1\text{-}0}(u) = E(Y_1 \mid U = u) - E(Y_0 \mid U = u)$ Individual causal effect
u_1	1/8	68	82	14
u_2	1/8	81	89	8
u_3	1/8	89	101	12
u_4	1/8	102	108	6
u_5	1/8	112	118	6
u_6	1/8	119	131	12
u_7	1/8	131	139	8
u_8	1/8	138	152	14
		$E(\tau_0) = 105$	$E(\tau_1) = 115$	$ACE_{1\text{-}0} = 10$

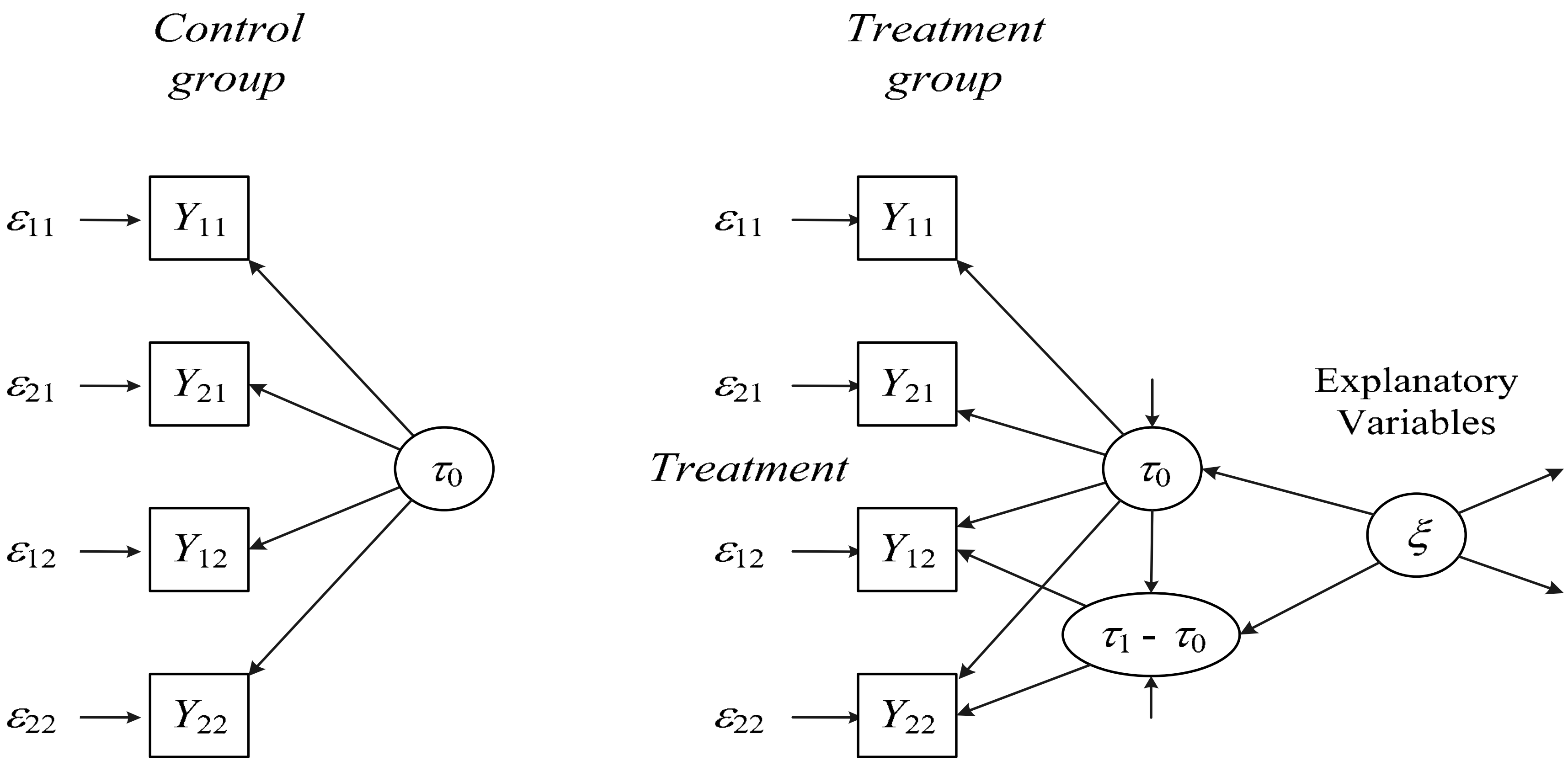
$$Y_0 = \tau_0 + \varepsilon_0$$
$$Y_1 = \tau_0 + (\tau_1 - \tau_0) + \varepsilon_1$$
$$= \tau_1 + \varepsilon_1$$



Identified Individual Effects Model with pretests Y_{11} and Y_{21}

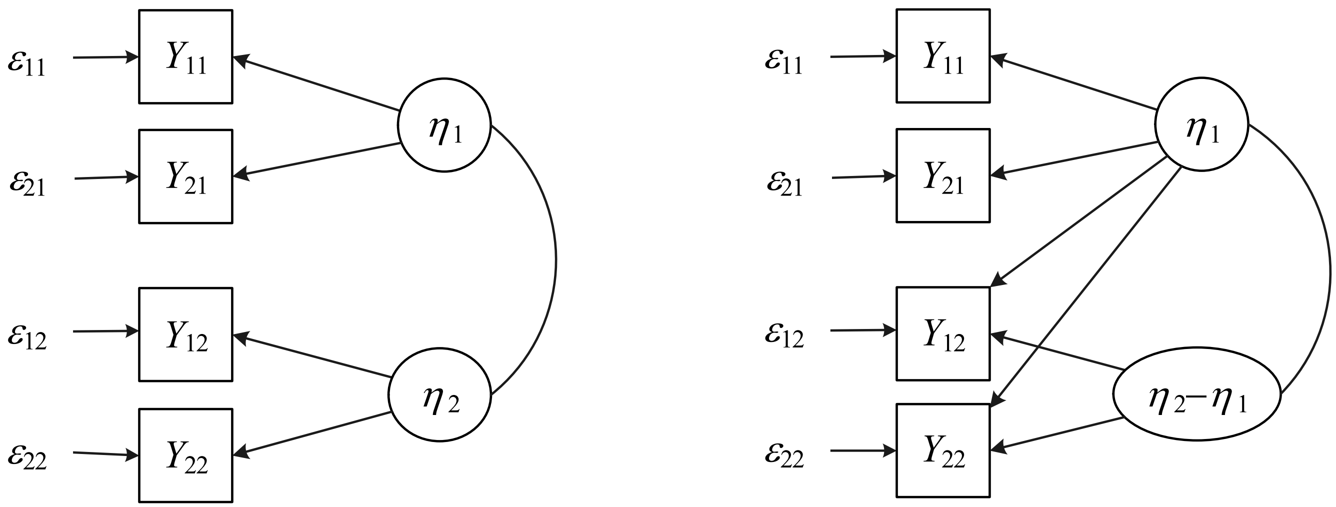


Identified Individual Effects Model with pretests Y_{11} and Y_{21}



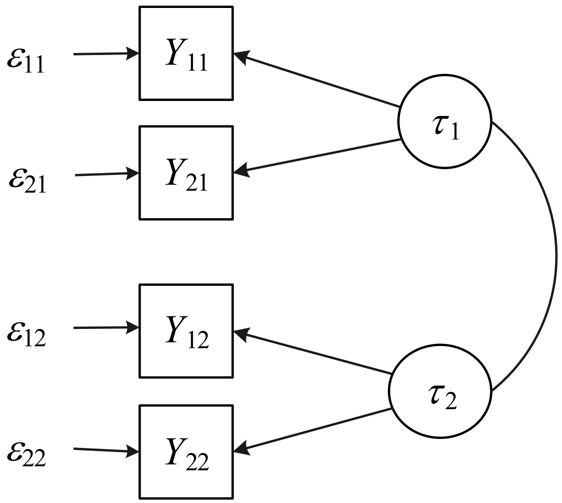
Design for the Analysis of Method Effects

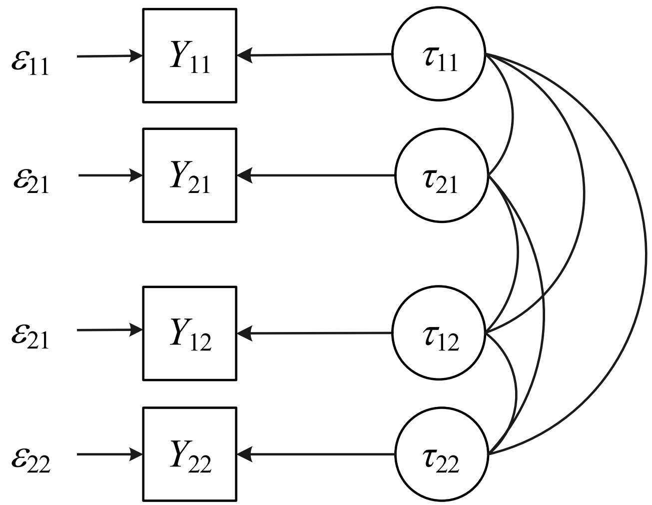
Introducing Individual Method Effects



Introducing Individual Method Effects

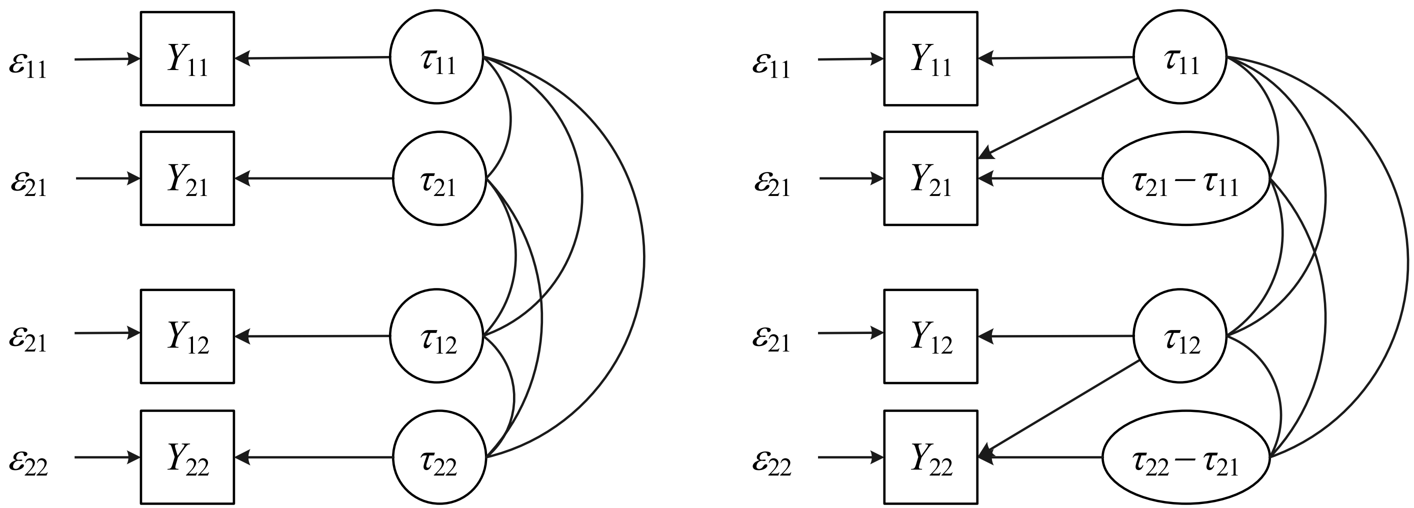
$$\tau_{22} - \tau_{12} = \tau_{21} - \tau_{11} = IME_{2\text{-}1}$$





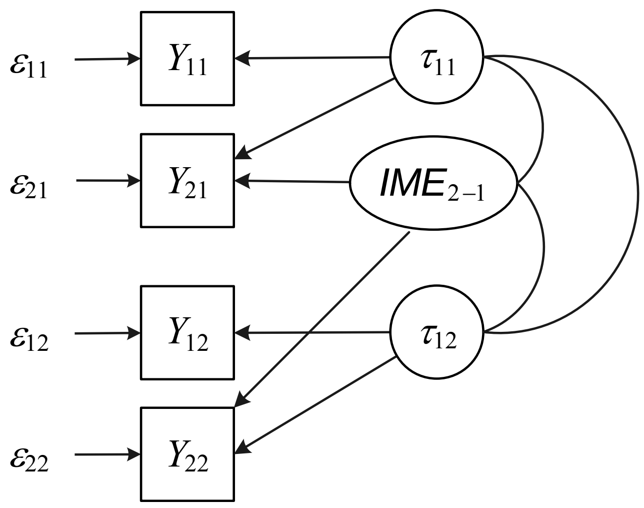
Introducing Individual Method Effects

$$\tau_{22} - \tau_{12} = \tau_{21} - \tau_{11} = IME_{2\text{-}1}$$

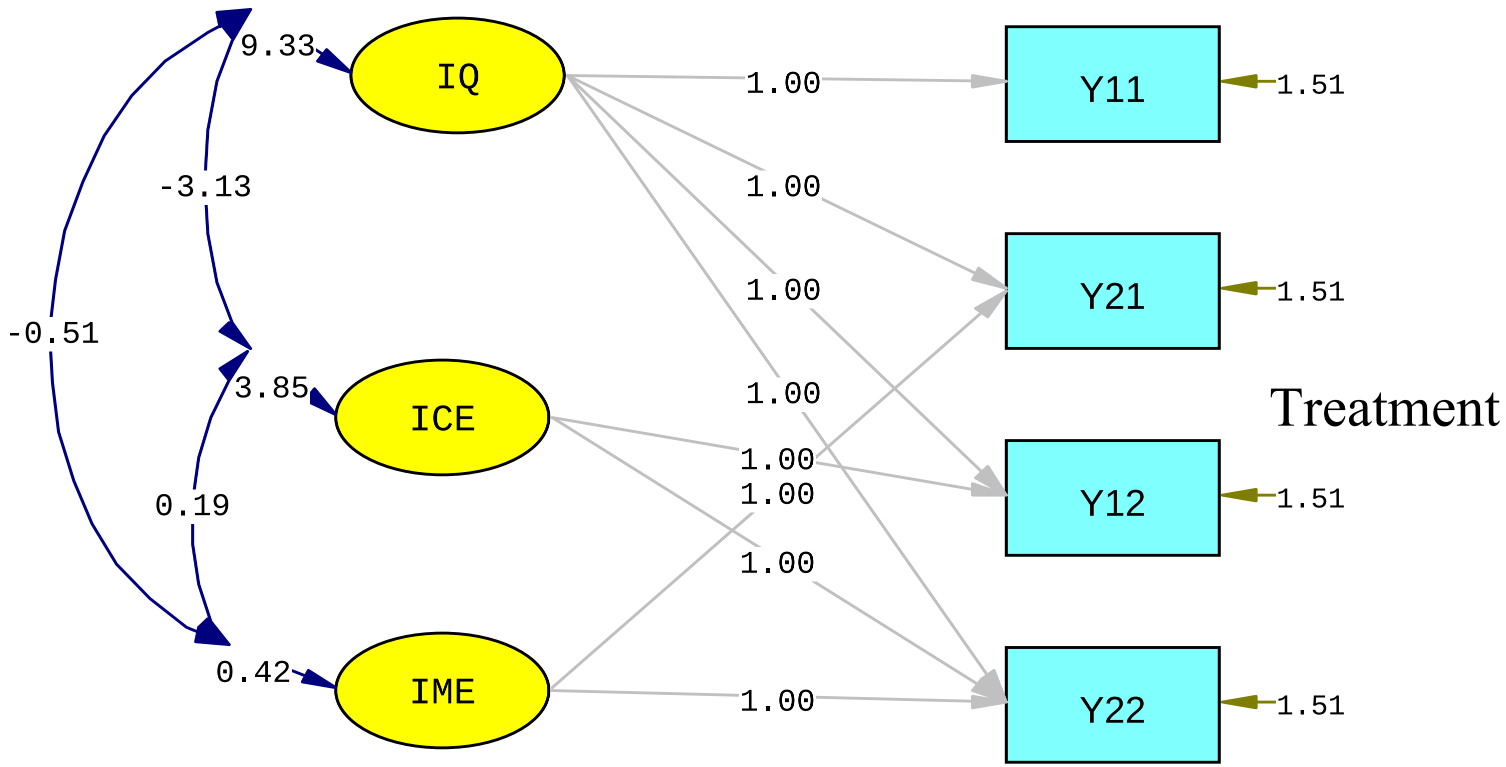


An identifíed Individual-Method-Effects Model

$$\tau_{22} - \tau_{12} = \tau_{21} - \tau_{11} = IME_{2\text{-}1}$$

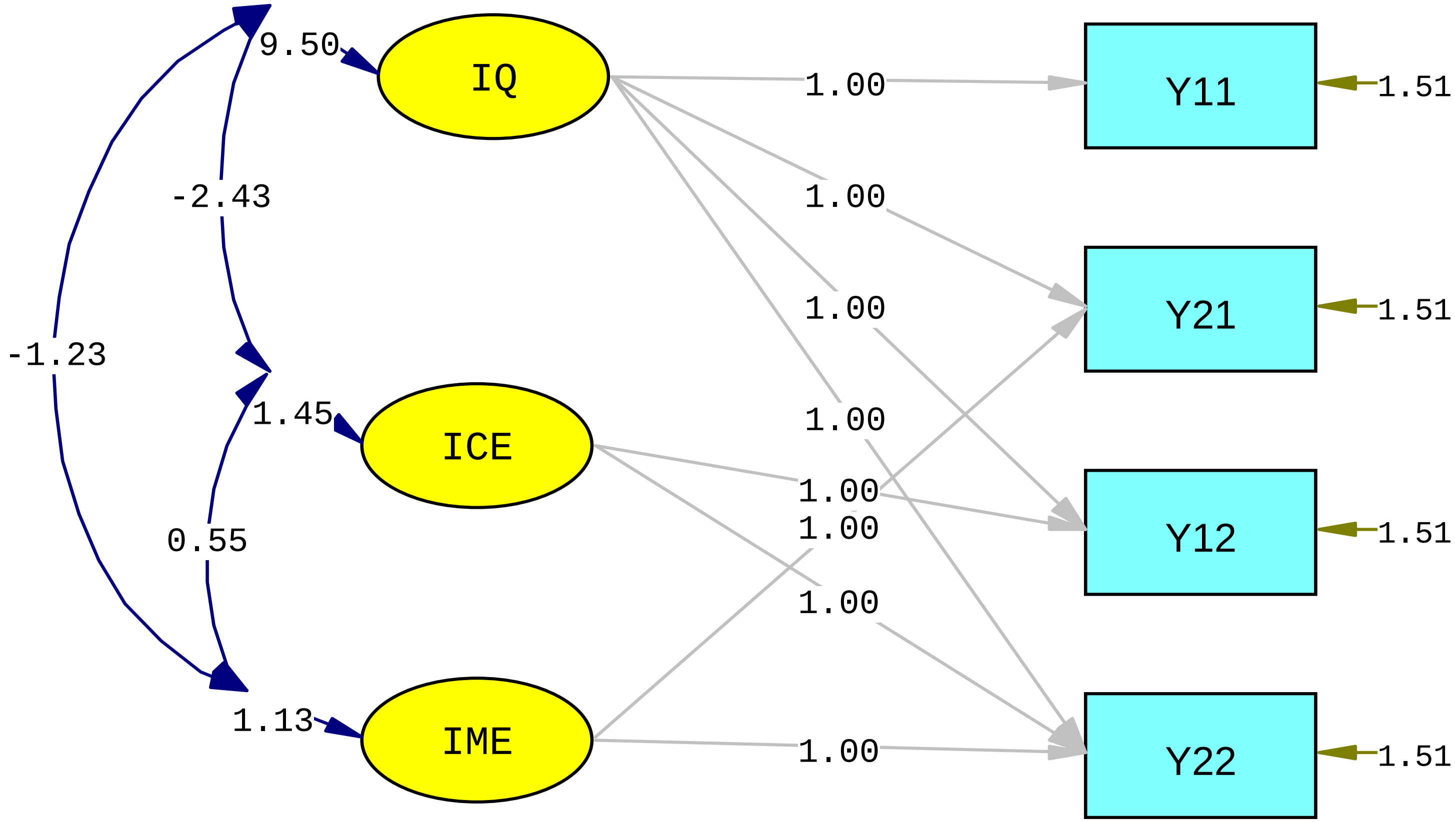


Model in treatment group



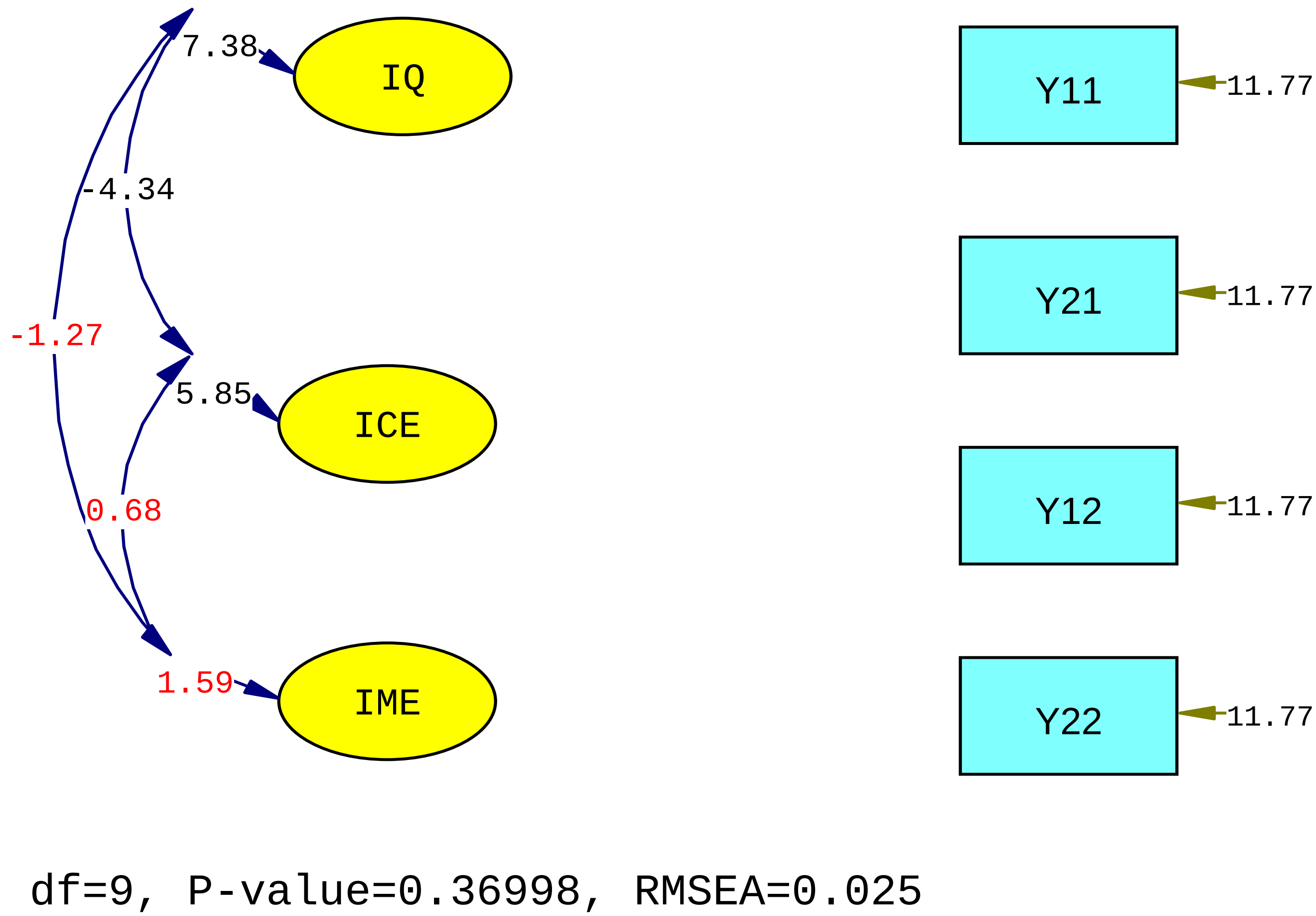
Chi-Square=9.76, df=9, P-value=0.36998, RMSEA=0.025

Model in control group

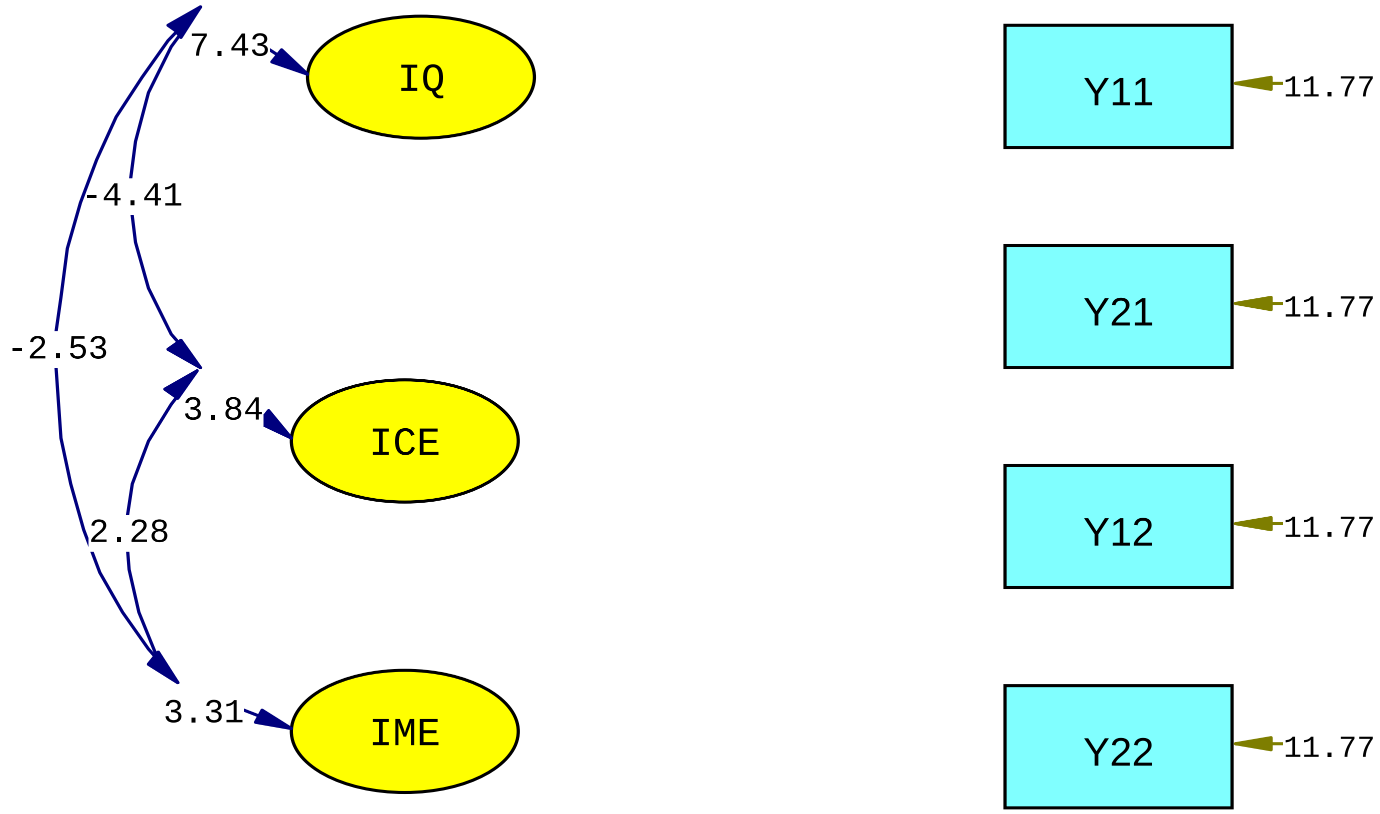


Chi-Square=9.76, df=9, P-value=0.36998, RMSEA=0.025

Model in treatment group (t-values)



Model in control group (t-values)



Chi-Square=9.76, df=9, P-value=0.36998, RMSEA=0.025

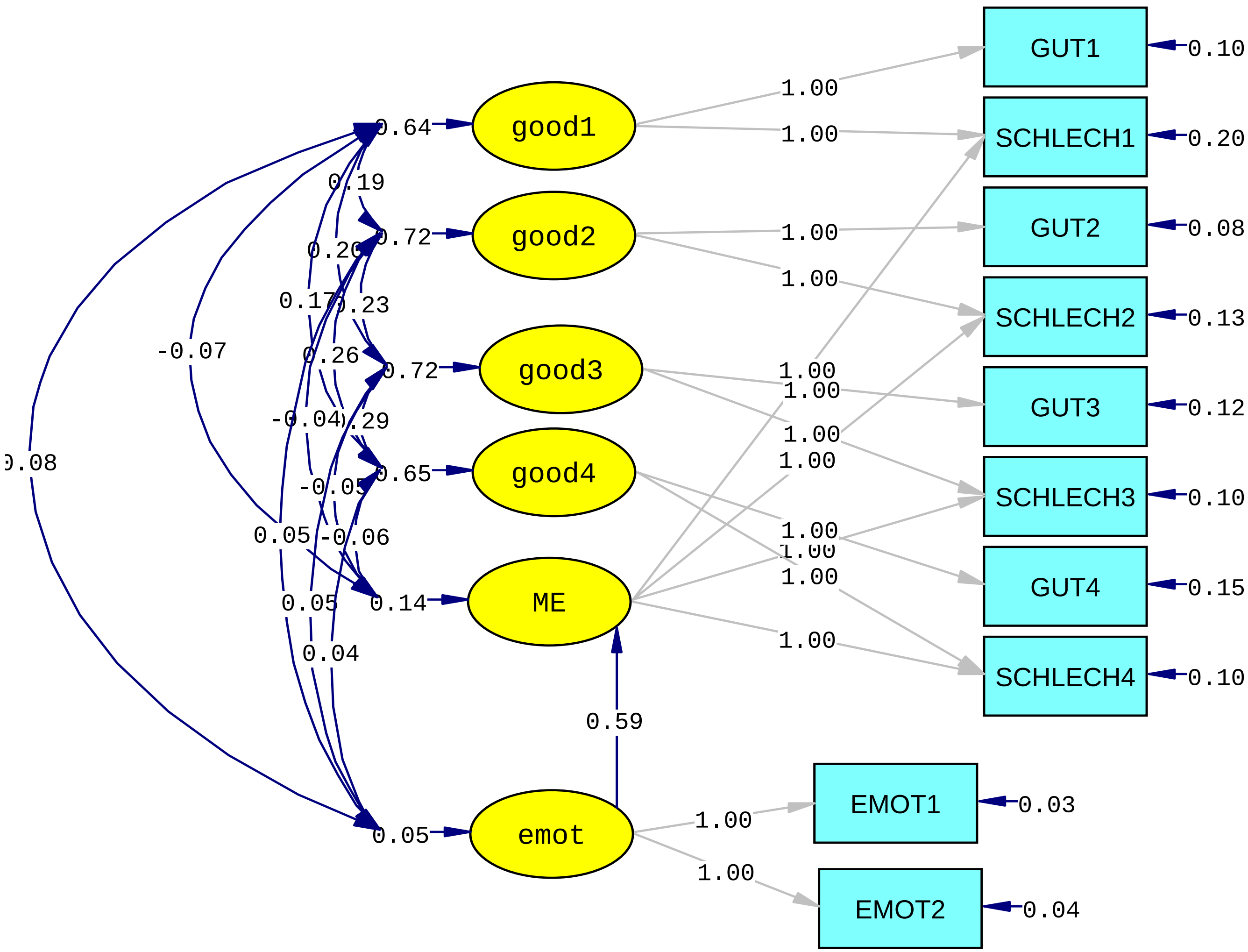
Correlation Matrix of ETA in Control group

	IQ	ICE	IME
IQ	1.00		
ICE	-0.65	1.00	
IME	-0.37	0.43	1.00

Correlation Matrix of ETA in Experimental Group

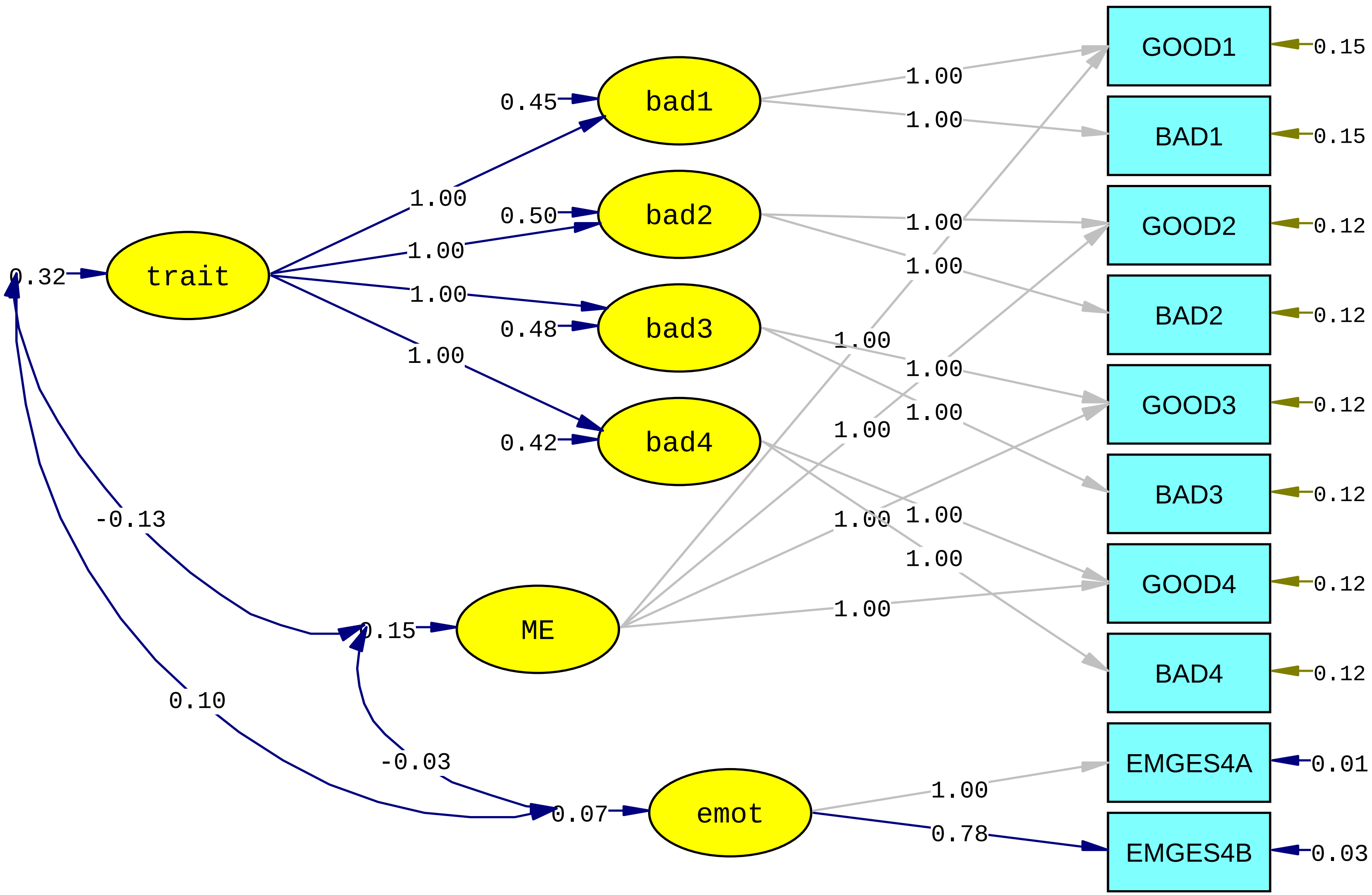
	IQ	ICE	IME
IQ	1.00		
ICE	-0.52	1.00	
IME	-0.26	0.15	1.00

The effects of negativ item formulation



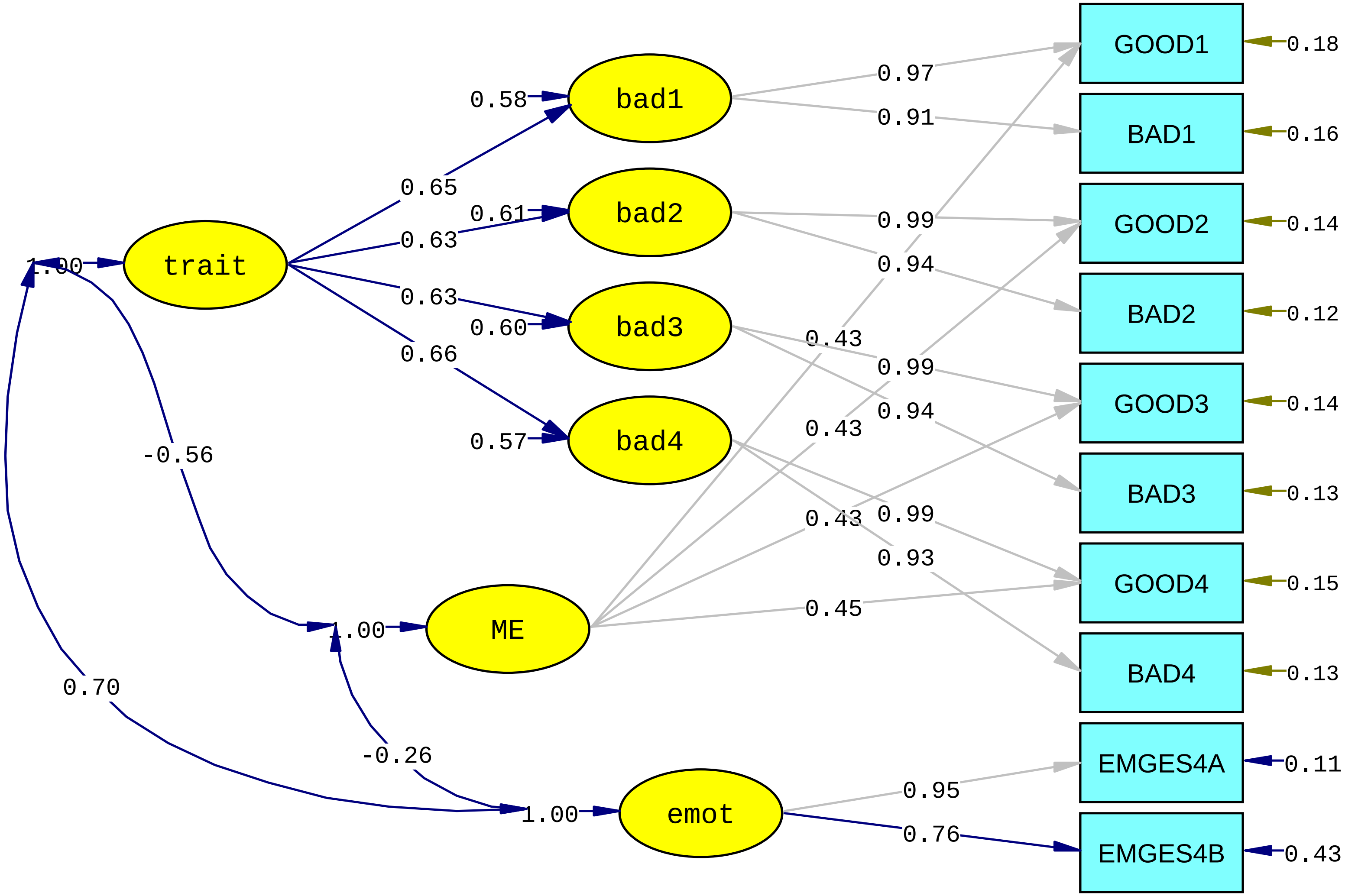
Chi-Square=53.09, df=28, P-value=0.00287, RMSEA=0.042

The effects of negativ item formulation



Chi-Square=76.70, df=43, P-value=0.00119, RMSEA=0.039

The effects of negativ item formulation (standardized)



Chi-Square=76.70, df=43, P-value=0.00119, RMSEA=0.039

Summary and Conclusion

- We can analyze individual (and average) causal effects in Pre-post Designs
- The causal interpretation rests on assumptions
- These assumptions can be tested
- Latent variables can be constructed from true-scores
- Not a single path in our SEM models represented a causal effect

Want More?

Steyer, R. & Partchev, I.. (2006). Causal Effects in Experiments and Quasi-Experiments: Theory (Chapters 1 -5 are available at www.causal-effects.de)

- Symposium on causality in Jena July 7 to 9, 2006 with Tom Cook, Steve West, Don Rubin … (videos available: see www.uni-jena.de/svw/metheval)
- Online video of workshop on the analysis of causal effects (same home page)
- Software „EffectLite“ (see: www.statlite.com)

Thanks to:

Sven Hartenstein
Ulf Kröhne
Benjamin Nagengast
Ivailo Partchev
Steffi Pohl