How to use Stata's sem with small samples? New corrections for the L. R. χ^2 statistics and fit indices

Meeting of the German Stata User Group at the Konstanz University, June 22nd, 2018

"All models are false, but some are useful."
(George E. P. Box)

Dr. Wolfgang Langer
Martin-Luther-Universität
Halle-Wittenberg
Institut für Soziologie

Assistant Professeur Associé
Université du
Luxembourg

Contents

- **What is the problem?**
- **What are the solutions for it?**
- **What do we know from Monte-Carlo simulation studies?**
- **How to implement it in Stata?**
- **Empirical example of Islamophobia in Germany 2016**
- **Conclusions**

What is the problem?

- **In empirical research more and more people estimate their SEM using a small sample (n<100) in psychology, marketing or business research**
- **When working with small samples we are confronted with a severe problem**
  - **The traditional Likelihood-Ratio χ^2 goodness-of-fit test and all fit-indices basing on it tend to overreject acceptable models. They are too conservative!**
  - **This is caused by the pure approximation of the χ^2 test statistics to the noncentral χ^2 distribution**

What are solutions for it?

- **Several correction procedures have been developed to improve the approximation of the L.R.χ^2-test statistics (T_{ML}) to the noncentral χ^2 distribution**

  - **The Bartlett correction**
  - **The Yuan correction**
  - **The Swain correction**

The Bartlett correction

- **Bartlett developed a small-sample correction to test the exact fit of exploratory factor models (1937, 1950, 1954) estimated by ML**

$$b = 1 - \frac{4k + 2p + 5}{6n}$$

$$Bartlett - corrected \ \ T_{MLb} = b \cdot T_{ML}$$

$Legend:$

$k:$ number of latent variables (factors)

$p:$ number of observed variables (indicators)

$n:$ sample size $N + 1$

The Yuan correction

- **Yuan (2005) proposed an "ad hoc" simplification of a Bartlett like correction formula developed by Wakaki, Eguchi & Fujikoshi (1990) for covariance structure models**

$$y = 1 - \frac{2k + 2p + 7}{6n}$$

$$Yuan-corrected\ \ T_{MLy} = y \cdot T_{ML}$$

The Swain correction

- **Swain (1975) proposed the following correction of the test statistics T_{ML}**

$$s = 1 - \frac{p \times \left(2p^2 + 3p - 1\right) - q\left(2q^2 + 3q - 1\right)}{12d \times n}$$

$$with \qquad q = \frac{\sqrt{1 + 4p\left(p+1\right) - 8d} - 1}{2}$$

$$Swain - corrected \;\; T_{MLs} = s \times T_{ML}$$

$Legend:$

$p:$ number of observed variables (indicators)

$d:$ degrees of freedom of actual model

n: sample size $N + 1$

What do we know from M. C. studies?

- A lot of Monte-Carlo simulation studies with small samples have been made to evaluate the shown corrections. They test systematically
    - Violations of the multivariate normal distribution assumption
    - Sample size
    - Number of indicators
    - Extend of model misspecification

What do we know ... ?

- **Fouladi (2000) and Newitt & Hancock (2004) recommended the Bartlett correction of the T_{ML} for normal data. For not normal distributed data they proposed the Bartlett correction of the Satorra-Bentler adjusted T_{ML}**
- **Herzog, Boomsma & Reinecke (2007) and Herzog & Boomsma (2009/13) showed that**
  - **Both Bartlett and Yuan corrections overestimate the type-I-error rate when sample size decreases**
  - **The Swain correction is the winner for small sample sizes and large models with many indicators**
    - **It reduces to a high extend the type-I-error rate**
    - **It works even to a sample size to estimated parameter ratio of 2:1**

What do we know ... ?

- **Herzog, Boomsma & Reinecke (2007) and Herzog & Boomsma (2009/13) also developed and tested a modified version of Tucker-Lewis-Index (TLI or NNFI) using the Swain-rescaled T_{ML} for the target model and usual T_{ML} for the baseline model**
  - **It clearly outperforms the TLI calculated by standard programs like MPLUS, EQS, LISREL**
  - **It reports correctly the misspecification of the SEM**
  - **They recommended this correction also for the Comparativ Fit Index developed by Bentler (1990) and Steiger's Root-Mean-Squared-Error of Approximation (RMSEA)**

Swain corrected Tucker-Lewis Index

- **Formulas**

For normal distributed data:

$$TLI = \frac{\dfrac{T_{ML_{bs}}}{df_{bs}} - \dfrac{s \cdot T_{ML_{ms}}}{df_{ms}}}{\dfrac{T_{ML_{bs}}}{df_{bs}} - 1}$$

For not normal distributed data:

$$TLI = \frac{\dfrac{sb \cdot T_{ML_{bs}}}{df_{bs}} - \dfrac{s \cdot sb \cdot T_{ML_{ms}}}{df_{ms}}}{\dfrac{sb \cdot T_{ML_{bs}}}{df_{bs}} - 1}$$

Swain corrected Comparative-Fit Index

- **Formulas**

For normal distributed data:

$$CFI = \frac{\left(T_{ML_{bs}} - df_{bs}\right) - \left(s \cdot T_{ML_{ms}} - df_{ms}\right)}{\left(T_{ML_{bs}} - df_{bs}\right)}$$

For not normal distributed data:

$$CFI = \frac{\left(sb \cdot T_{ML_{bs}} - df_{bs}\right) - \left(s \cdot sb \cdot T_{ML_{ms}} - df_{ms}\right)}{\left(sb \cdot T_{ML_{bs}} - df_{bs}\right)}$$

Swain corrected RMSEA

- **Formulas**

For normal distributed data:

$$RMSEA = \sqrt{\frac{s \cdot T_{ML_{ms}} - df_{ms}}{n \cdot df_{ms}}}$$

For not normal distributed data:

$$RMSEA = \sqrt{\frac{s \cdot sb \cdot T_{ML_{ms}} - df_{ms}}{n \cdot df_{ms}}}$$

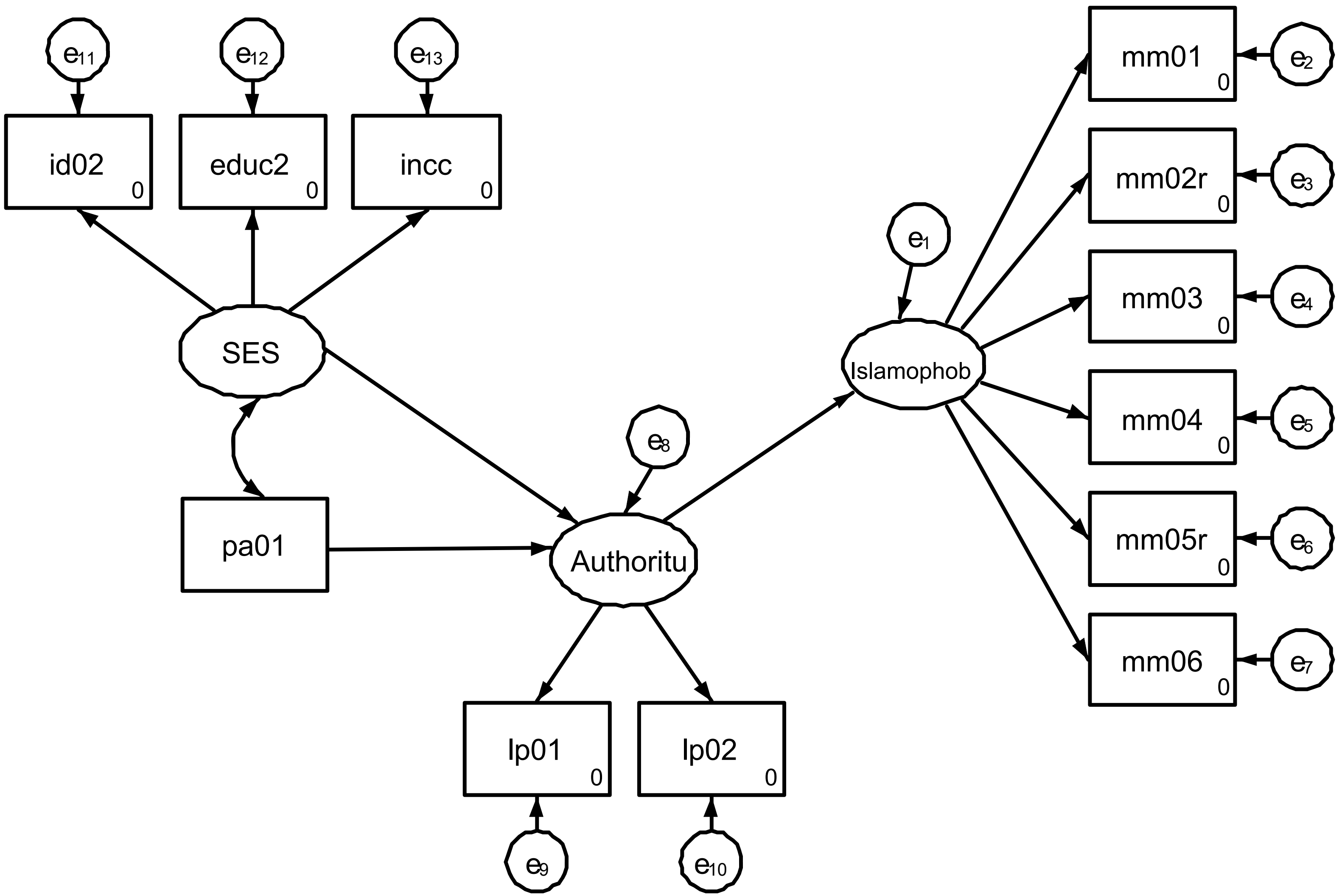
How to implement it in Stata ?

- **In 2013 John Antonakis and Nicolas Bastardoz, both from University of Lausanne, Switzerland, published their "swain.ado" calculating only the Swain-corrected T_{ML} value for comparison of the actual vs. saturated model**
- **I have modified this ado-file calculating now Swain-corrected T_{ML}, TLI, CFI and RMSEA**
  - **Under the assumption of multivariate normality (Jöreskog 1970, p. 239)**
  - **Under violation of the multivariate normality assumption (not normal distributed data) using the Satorra-Bentler-corrected T_{ML}**
  - **All calculated scalars are displayed and returned in r-containers**

Empirical example of Islamophobia

- **SEM explaining Islamophobia in West Germany 2016**
- **5% sample of the German General Social Survey 2016, subsample west: n=84**
- **Presentation of used indicators**
- **Test of multivariate normal distribution of observed indicators (mvtest in Stata)**
- **Estimated results from sembuilder**
- **Results of estat gof, stats(all)**
- **Output of my swain_gof.ado**

SEM to explain Islamophobia



Used indicators

- **Factor SES: Socio-economic status**
  - **id02: Self rating of social class**
    - **Underclass to upperclass [1;5]**
  - **educ2: educational degree**
    - **Without degree to grammar school [1;5]**
  - **incc: income class (quintiles) [1;5]**
- **Factor Authoritu: authoritarian submission**
  - **lp01: Thank to the leading heads saying us what to do [1;7]**
  - **lp02: It is good for a child to learn to obey its parents [1;7]**
- **Single indicator pa01: left-right self-rating**
  - **1) left .. 10) right**

Used indicators

- **Factor Islamophobia**
  - **Six items [1;7]**
    - **mm01 The religious practice of Islam should be restricted in Germany**
    - **mm02r The Islam does not belong to Germany**
    - **mm03 The presence of Muslims leads to conflicts**
    - **mm04 The Islamic communities should be supervised by the state**
    - **mm05r I object to have an Islamic mayor in my town**
    - **mm06 There are a lot of religious fanatics in the Islamic community**

Test of multivariate normality (n = 84)

```
. mvtest normality mm01 mm02r mm03 mm04 mm05r mm06 lp01 lp02 pa01 id02 educ2 incc, uni stats(all)
```

Test for univariate normality

Variable	Pr(Skewness)	Pr(Kurtosis)	joint adj chi2(2)	joint Prob>chi2
mm01	0.4475	0.0000	44.17	0.0000
mm02r	0.0086	0.4302	6.91	0.0317
mm03	0.4600	0.0040	7.89	0.0194
mm04	0.1012	0.0002	13.43	0.0012
mm05r	0.4737	0.0000	.	0.0000
mm06	0.5839	0.0000	24.28	0.0000
lp01	0.1037	0.0827	5.47	0.0648
lp02	0.0000	0.0174	19.23	0.0001
pa01	0.0280	0.9034	4.83	0.0893
id02	0.9191	0.4762	0.53	0.7685
educ2	0.0255	0.0142	9.47	0.0088
incc	0.8780	0.0000	20.93	0.0000

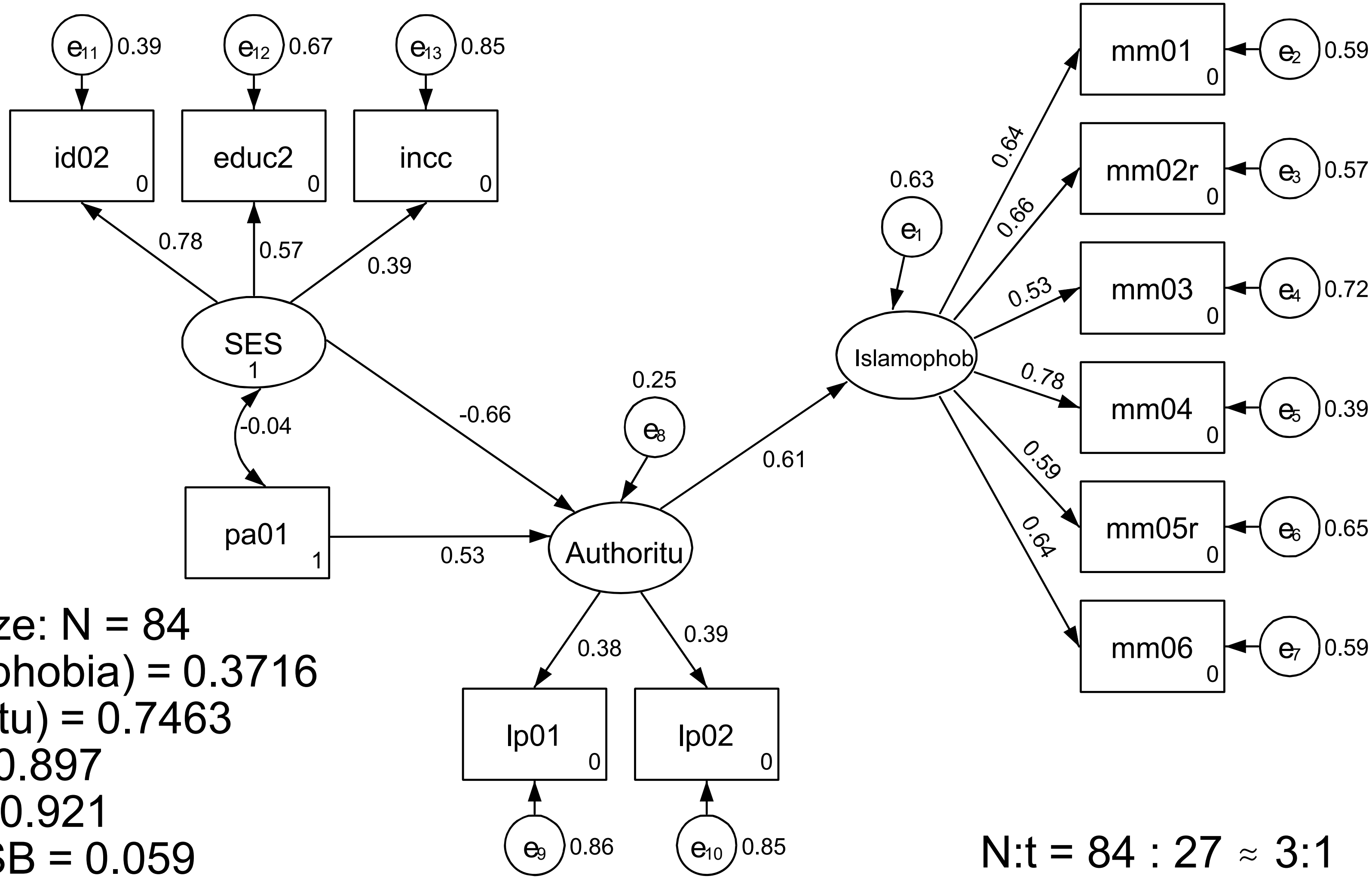
Except id02 all indicators violate the assumption of univariate nomality!

Test for multivariate normality

Mardia mSkewness =	31.04157	chi2(364) =	452.560	Prob>chi2 =	0.0011
Mardia mKurtosis =	173.1796	chi2(1) =	1.677	Prob>chi2 =	0.1954
Henze-Zirkler =	1.034168	chi2(1) =	40.558	Prob>chi2 =	0.0000
Doornik-Hansen		chi2(24) =	118.558	Prob>chi2 =	0.0000

All together violate the assumption of multivariate normality!

Standardized solution of the SEM with Satorra-Bentler corrections: vce(sbentler)



Sample size: N = 84

R^2(Islamophobia) = 0.3716

R^2(Authoritu) = 0.7463

TLI_SB = 0.897

CFI_SB = 0.921

RMSEA_SB = 0.059

N:t = 84 : 27 ≈ 3:1

Goodness of fit statistics: estat gof

Fit statistic	Value	Description
Likelihood ratio		
chi2_ms(51)	67.723	model vs. saturated
p > chi2	0.058	
chi2_bs(66)	264.488	baseline vs. saturated
p > chi2	0.000	
Satorra-Bentler		
chi2sb_ms(51)	65.876	
p > chi2	0.079	
chi2sb_bs(66)	253.809	
p > chi2	0.000	
Population error		
RMSEA	0.062	Root mean squared error of approximation
90% CI, lower bound	0.000	
upper bound	0.099	
pclose	0.292	Probability RMSEA <= 0.05
Satorra-Bentler		
RMSEA_SB	0.059	Root mean squared error of approximation
Baseline comparison		
CFI	0.916	Comparative fit index
TLI	0.891	Tucker-Lewis index
Satorra-Bentler		
CFI_SB	0.921	Comparative fit index
TLI_SB	0.897	Tucker-Lewis index
Size of residuals		
SRMR	0.083	Standardized root mean squared residual
CD	0.884	Coefficient of determination

Output of my swain_gof.ado

```
. swain_gof

Swain correction factor = 0.9391
Swain corrected chi-square = 63.597864
p-value of Swain corrected chi-square  = 0.1108

Satorra-Bentler-corrected statistics:
Swain-Satorra-Bentler corrected chi-square = 61.863491
p-value of Swain-Satorra-Bentler corrected chi-square = 0.1417

Fit indices under assumption of multivariate normal distribution

Swain-corrected Tucker-Lewis-Index = 0.9179
Swain-corrected Comparative-Fit-Index = 0.9365
Swain-correct RMSEA = 0.0542

Fit indices under violation of multivariate normal distribution

Swain-Satorra-Bentler-corrected Tucker-Lewis-Index = 0.9251
Swain-Satorra-Bentler-corrected Comparative-Fit-Index = 0.9422
Swain-Satorra-Bentler-correct RMSEA = 0.0504
```

r-containers of the swain_gof.ado

- **The swain_gof.ado returns the following r-containers**

```
. return list

scalars:
      r(swain_rmsea_sb) =  .0503570145572617
        r(swain_cfi_sb) =  .9421565802285176
        r(swain_tli_sb) =  .9251438097074932
         r(swain_rmsea) =  .0542280184201118
           r(swain_cfi) =  .936530799932064
           r(swain_tli) =  .917863388147377
          r(swain_sb_p) =  .1417421175619517
        r(swain_chi_sb) =  61.86349107237523
          r(swain_corr) =  .9390845490337976
           r(swain_chi) =  63.59786447391117
             r(swain_p) =  .1108283705582046
```

What do we see ?

- **Violation of the multivariate normality assumption. Therefore we look at the Satorra & Bentler (SB) corrected statistics of Stata**
  - **SB T_{ML} (Stata) = 65.876 df=51 p=0.079**
  - **Swain SB T_{ML}= 61.863 df=51 p=0.142 ☺**
  - **SB TLI (Stata) = 0.891**
  - **Swain SB TLI = 0.925 ☺**
  - **SB CFI (Stata) = 0.921**
  - **Swain SB CFI = 0.942 ☺**
  - **SB RMSEA (Stata)= 0.059**
  - **Swain SB RMSEA = 0.050 ☺**
- **The SB T_{ML} statistics is reduced by the Swain correction. Therefore all fit indices are improved!**

Conclusions

- The Monte-Carlo studies presented have proofed the advantage of the Swain correction for the SEM with small samples and many indicators
  - It works just to a sample size-parameter ratio of 2:1
- My swain_gof.ado calculates easily the Swain-corrected T_{ML} statistics and the fit indices TLI, CFI and RMSEA basing on it
  - Under the assumption of multivariate normality
  - Under violation of multivariate normality
- Therefore I recommend my swain_gof.ado to assess the fit of SEMs using small samples

Closing words

- **Thank you for your attention**
- **Do you have some questions?**

Contact

- **Affiliation**
  - **Dr. Wolfgang Langer**
    University of Halle
    Institute of Sociology
    D 06099 Halle (Saale)
  - **Email:**
    - **wolfgang.langer@soziologie.uni- halle.de**

References

- Antonakis, J., & Bastardoz, N. (2013) Swain: Stata module to correct the SEM chi-square overidentification test in small sample sizes or complex models. http://econpapers.repec.org/software/bocbocode/s457617.htm
- Bentler, P. M. & Chou, C.-P. (1987): Practical issues in structural equation modeling. Sociological Methods & Research, 16 (1), 78-117
- Bentler, P. M. & Yuan, K.-H. (1999): Structural equation modeling with small samples: Test statistics. Multivariate Behavioral Research, 34 (2), 181-197
- Bentler, P. M. (1990): Comparative fit indexes in structural equation models. Psychological Bulletin, 107, 238-246
- Boomsma, A. & Herzog, W.(2013): R function swain. Correcting structural equation model fit statistics and indexes unter small-sample and or large-model conditions. University of Groningen, The Netherlands, WHU Germany (http://www.gmw.rug.nl/~boomsma/swain.R)

References 2

- Curran, P. J., Bollen, K. A., Paxton, P., Kirby, J., & Chen, F. N. (2002). The noncentral chi-square distribution in misspecified structural equation models: Finite sample results from a Monte Carlo simulation. Multivariate Behavioral Research, 37(1), 1-36
- Fouladi, R. T. (2000): Performance of modified test statistics in covariance and correlation structure analysis under condition of multivariate nonnormality. Structural Equation Modeling, 7 (3), 356-410
- GESIS - Leibniz-Institut für Sozialwissenschaften (2017): Allgemeine Bevölkerungsumfrage der Sozialwissenschaften ALLBUS 2016. GESIS Datenarchiv, Köln. ZA5250 Datenfile Version 2.1.0, doi:10.4232/1.12796
- Herzog, W., Boomsma, A., & Reinecke, S. (2007): The model-size effect on traditional and modified tests of covariance structures. Structural Equation Modeling, 14(3), 361-90
- Herzog, W., & Boomsma, A. (2009): Small-sample robust estimators of noncentrality-based and incremental model fit. Structural Equation Modeling, 16(1), 1-27

References 3

- Jöreskog, K.G. (1970): A general method for analysis of covariance structures. Biometrika, 57(2), 239-251
- Nevitt, J. & Hancock, G. R. (2004): Evaluating small sample approaches for model test statistics in Structural Equation Modeling. Multivariate Behavioral Research, 39 (3), 439-478
- Satorra, A. & Bentler, P. M. (1994): Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (eds.), *Latent variables analysis: Applications for developmental research* (pp. 399-419). Newbury Park, CA: Sage
- Swain, A. J. (1975): Analysis of parametric structures for variance matrices (doctoral thesis). University of Adelaide, Adelaide
- Wakaki, H., Eguchi, S. & Fujikoshi, Y.(1990): A class of tests for a general covariance structure. Journal of Multivariate Analysis, 32, 313-325
- Yuan, K.-H. (2005): Fit indices versus test statistics. Multivariate Behavioral Research, 40 (1), 115-148

Appendix

Assumption of multivariate normality 1

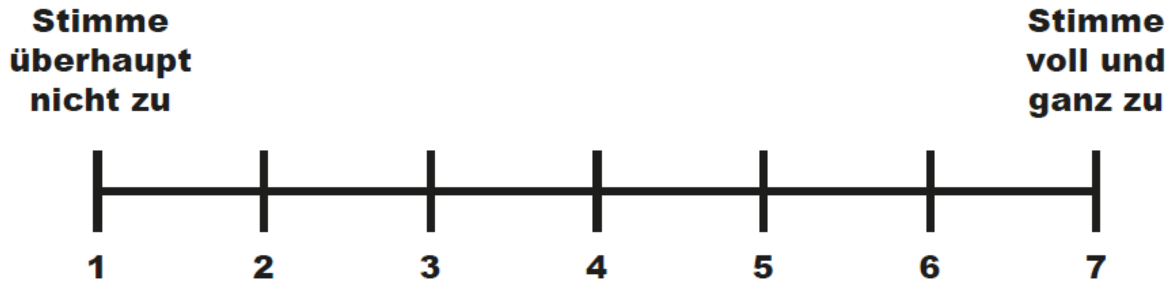
- Karl G. Jöreskog (1970) formulated this assumption in his article "A general method for analysis of covariance structures". Biometrika, 57 (2), p. 239-251

"It is assumed that observations on a set of variables have a multivariate normal distribution with a general parametric form of the mean vector and the variance-covariance parameters. Any parameter of the model may be fixed, free or constrained to be equal to other parameters. The free and constrained parameters are estimated by maximum likelihood." (p. 239)

Assumption of multivariate normality 2

- Jöreskog had reduced the whole data matrix X to the first and second moments of the observed variables ignoring the third and forth moments - their skewness and kurtosis. Therefore he needed a strict assumption of their distribution.
- That's why he introduced the multivariate normality assumption of the observed variables.

Items measuring Islamophobia



A Die Ausübung des islamischen Glaubens in Deutschland sollte eingeschränkt werden. +) mm01

B Der Islam passt in die deutsche Gesellschaft. -) mm02r

C Die Anwesenheit von Muslimen in Deutschland führt zu Konflikten. +) mm03

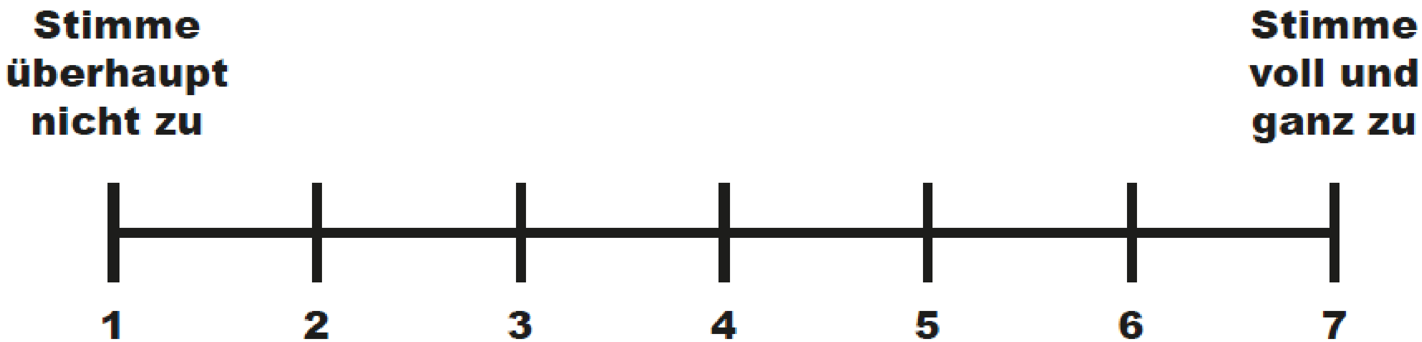
D Islamische Gemeinschaften sollten vom Staat beobachtet werden. +) mm04

E Ich hätte nichts gegen einen muslimischen Bürgermeister in meiner Gemeinde. -) mm05r

F Ich habe den Eindruck, dass unter den in Deutschland lebenden Muslimen viele religiöse Fanatiker sind. +) mm06

(GESIS 2017, Liste 54)

Items measuring authoritarian submission



A Wir sollten dankbar sein für führende Köpfe, die uns genau sagen können, was wir tun sollen und wie.

Ip01

B Im allgemeinen ist es einem Kind im späteren Leben nützlich, wenn es gezwungen wird, sich den Vorstellungen seiner Eltern anzupassen.

Ip02

(GESIS 2017, Liste 34)

Left-right-self rating

pa01

(GESIS 2017, Liste 46)