Auswirkungen der Fehlspezifikation von Messfehlerkorrelationen auf die Parameterschätzung in linearen Strukturgleichungsmodellen in der Glaukomdiagnose

vorgelegt von

Dipl.-Soz.

Helmut Orawa

geb. in Erlangen

von der Fakultät VII – Wirtschaft und Management

der Technischen Universität Berlin

zur Erlangung des akademischen Grades

Doktor Public Health

- Dr. P. H. -

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Dr. med. Reinhard Busse

Gutachter: Prof. Dr. rer nat. Peter Martus

Gutachter: Prof. Dr. Michael Eid

Tag der wissenschaftlichen Aussprache: 26. Februar 2016

Berlin 2016

D 83

Danksagung und Widmung

Ich sage an erster Stelle großen Dank Herrn Univ.-Prof. Dr. rer. nat. Peter Martus, der mich damit betraute, seine Forschungen in der Glaukomdiagnostik fortzusetzen und mich stets und über eine lange Zeit intensiv bei der Anfertigung dieser Arbeit unterstützte.

An zweiter Stelle gilt mein Dank Herrn Univ.-Prof. Dr. Michael Eid aufgrund seiner wertvollen mehrfachen Anregungen zur multimethodalen Diagnostik, seiner Lehrveranstaltung in SEM und großzügigen Erlaubnis, in seinen Räumen zu recherchieren.

Drittens danke ich Herrn Univ.-Prof. Dr. med. habil., MSc Applied Statistics Peter Schlattmann, über den ganzen Forschungsprozess hinweg zum Gedankenaustausch bereit gewesen zu sein – insbesondere über das Softwarepaket SAS.

Last but not least bin ich Herrn Univ.-Prof. Dr. Horst Skarabis zu Dank dafür verpflichtet, mich über das Studium hinaus gefördert und mein Interesse an Epidemiologie und medizinischer Statistik geweckt zu haben.

Diese Arbeit ist meiner Tochter Hannah gewidmet.

Inhaltsverzeichnis

1 EINLEITUNG 12

2 LINEARE STRUKTURGLEICHUNGSMODELLE IN DER MEDIZINISCHEN DIAGNOSTIK 19

2.1 ÜBERBLICK 19
2.2 DAS KONZEPT DER LATENTEN VARIABLEN 20
2.3 MESSMODELL UND STRUKTURMODELL 22
2.4 MATHEMATISCHE GRUNDLAGE EINES SEM 25
2.5 DIE METHODE DER KONFIRMATORISCHEN FAKTORENANALYSE 26
2.6 BEWERTUNG VON DIAGNOSEVERFAHREN DURCH BESTIMMUNG VON PFADKOEFFIZIENTEN 27
2.7 IDENTIFIZIERBARKEIT 30
2.8 INDIVIDUELLE QUANTIFIZIERUNG DES ERKRANKUNGSSCHWEREGRADES VON PATIENTEN DURCH INDEXBILDUNG 33
2.9 GÜTE DER ANPASSUNG DES MODELLS AN DIE EMPIRISCHEN DATEN 33
2.10 KLASSISCHE TESTTHEORIE UND MESSMODELLE 34

3 AMOS-MESSMODELLE MIT EINEM FAKTOR UND DREI MESSVERFAHREN AUF BASIS KÜNSTLICH ERZEUGTER DATEN IN AMOS 39

3.1 ZUSAMMENFASSUNG 39
3.2 EINFÜHRUNG: ANWENDUNGSBEREICH, PROBLEMSTELLUNG UND LÖSUNGSANSATZ 39
3.2.1 Anwendungsbereich 39
3.2.2 Problemstellung 40
3.2.3 Lösungsansatz: Verwendung nichtempirischer Daten 41
3.3 METHODE ZUR ERZEUGUNG KÜNSTLICHER DATEN 42
3.3.1 Erzeugung von Zufallszahlen 42
3.3.2 Hauptkomponentenanalyse zur Herstellung unabhängiger Zufallszahlen 43
3.3.3 Konstruktion manifester und als latent betrachteter Variabler 43
3.3.4 Systematik der Variablenerzeugung 44
3.3.5 Parameter der konstruierten Variablen 46
3.4 ANALYSE KÜNSTLICHER DATEN MIT MESSMODELLEN IN AMOS – FEHLERKORRELATIONSSTRUKTUREN UND SPEZIFIKATIONSVARIANTEN 47
3.4.1 Warum ein Modell mit einem Faktor und genau drei Indikatorvariablen ? 47
3.4.2 Vergleich der in SPSS künstlich erzeugten mit der in AMOS geschätzten Korrelationsmatrix 48
3.4.3 Fehlerkorrrelationsstrukturen und Spezifikationsvarianten 48
3.4.3.1 Überlegung zu den Wirkungen der Modellspezifikation – Anpassungsgüte des Modells und Koeffizientenberechnung 50
3.5 ANALYSE 51

3.5.1 Das Grundmodell 51
3.5.2 Das Modell mit zwei positiv korrelierten Fehlervariablen 55
3.5.3 Das Modell mit zwei negativ korrelierten Fehlervariablen 58
3.5.4 Modell mit drei positiv korrelierten Fehlern 60
3.5.5 Fazit 62

4 LINEARE STRUKTURGLEICHUNGSMODELLE MIT ZWEI FAKTOREN UNTER VERWENDUNG VON AMOS 64

4.1 STRUKTURGLEICHUNGSMODELLE, DEREN FAKTOREN STATISTISCH UNABHÄNGIG SIND 66
4.1.1 Modell 1 67
4.1.2 Modell 2 68
4.1.3 Modell 3 70
4.1.4 Zusatzmodelle Z1–Z3 73
4.1.5 Z 1 73
4.1.6 Z 2 75
4.1.7 Z 3 76
4.2 STRUKTURGLEICHUNGSMODELLE, DEREN FAKTOREN STATISTISCH ABHÄNGIG SIND 79
4.2.1 Modell 4 79
4.2.2 Modell 5 80
4.2.3 Modell 6 82
4.2.4 Modell 7 84
4.2.5 Modell 8 86
4.2.6 Modell 9 88
4.2.7 Zusammenfassung 91

5 SIMULATIONEN ZUM EIN-FAKTOR-MODELL MIT DREI INDIKATOREN UNTER VERWENDUNG VON PROC CALIS (SAS) 93

5.1 EINLEITUNG 93
5.2 REPRODUKTION DER AMOS-ANALYSE MIT PROC CALIS 95
5.3 SIMULATION 95
5.3.1 Das Grundmodell 95
5.3.2 Das Modell mit zwei positiv korrelierten Fehlern 98
5.3.3 Das fehlspezifizierte Modell 99
5.3.4 Das Modell mit zwei negativ korrelierten Fehlertermen 101
5.3.5 Das fehlspezifizierte Modell 102
5.3.6 Das Modell mit drei positiv korrelierten Fehlern 103
5.3.7 Das fehlspezifizierte Modell 105

5.4 Zusammenhänge zwischen Fehlerkorrelationsstruktur, korrekter oder Fehlspezifikation, Standardfehler und Pfadkoeffizientenkorrelation 106
5.5 Fazit aus allen Modellen 107
6 SIMULATIONEN ZUM EIN-FAKTOR-MODELL MIT VIER ODER FÜNF INDIKATOREN UNTER VERWENDUNG VON PROC CALIS 109
6.1 Ein-Faktor-Modelle mit drei vs. vier und mehr Indikatoren 109
6.2 Ein-Faktor-Modelle mit vier oder fünf Indikatoren 109
6.3 Modelle mit vier Indikatoren 110
6.3.1 Modell 1 111
6.3.2 Modell 2 112
6.4 Modelle mit fünf Indikatoren 112
6.4.1 Modell 3 113
6.4.2 Modell 4 114
6.5 Fazit aus allen Modellen 115
7 SIMULATIONEN ZU LINEAREN STRUKTURGLEICHUNGSMODELLEN (SEMS) MIT ZWEI FAKTOREN UNTER VERWENDUNG VON PROC CALIS 116
7.1 Einleitung 116
7.2 Strukturgleichungsmodelle, deren Faktoren statistisch unabhängig sind 119
7.2.1 Modell 1 120
7.2.2 Modell 2 121
7.2.3 Modell 3 126
7.2.4 Fazit aus den Simulationen zu den Modellen mit zwei statistisch unabhängigen Faktoren 129
7.3 Strukturgleichungsmodelle, deren Faktoren statistisch abhängig sind 129
7.3.1 Modell 4 129
7.3.2 Modell 5 131
7.3.3 Modell 6 133
7.3.4 Modell 7 134
7.3.5 Modell 8 136
7.3.6 Modell 9 139
7.4 Zusatzmodell und korrekt spezifizierte Modellvarianten, in welchen Fehlerkorrelation als bekannt vorausgesetzt 141
7.4.1 Modell 5 k (korrekt spezifiziert) 142
7.4.2 Modell Z4 k (korrekt spezifiziert) 143
7.4.3 Modell Z4 (fehlspezifiziert) 145
7.4.4 Modell 8 korrekt spezifiziert 146

7.5 KORREKT SPEZIFIZIERTE MODELLVARIANTEN MIT UNABHÄNGIGEN FAKTOREN, IN WELCHEN FEHLERKORRELATIONEN GESCHÄTZT WERDEN 147
7.5.1 Modell 2 k_g 147
7.5.2 Modell 3 k_g 148
7.6 KORREKT SPEZIFIZIERTE MODELLVARIANTEN MIT LINEAR ABHÄNGIGEN FAKTOREN, IN WELCHEN FEHLERKORRELATIONEN GESCHÄTZT WERDEN 150
7.6.1 Modell 8 k_g) 150
7.6.2 Modell 5 k_g 151
7.6.3 Modell Z4 k_g 151
7.6.4 Modell 6 k_g 152
7.6.5 Modell 9 k_g 153
7.7 ZUSAMMENFASSUNG UND FAZIT 155
7.7.1 Fazit aus den Analysen zum Pfadkoeffizienten des Strukturmodells, gamma 155
7.7.2 Fazit aus den Analysen zu den simulierten Standardfehlern der Pfadkoeffizienten der Messmodelle, beta1 bis beta6 159
7.7.3 Fazit aus dem Vergleich der simulierten und mittleren geschätzten SE der Pfadkoeffizienten der Messmodelle 161
7.7.4 Fazit aus den korrekt spezifizierten Modellvarianten, in welchen Fehlerkorrelationen geschätzt werden 162

8 ADJUSTIERUNG DER STANDARDFEHLERSCHÄTZUNG NACH DER INTRAINDIVIDUELLEN ABHÄNGIGKEIT ZWISCHEN PAARIGEN ORGANEN – SIMULATION UND ANWENDUNG IN DER GLAUKOMDIAGNOSE 165

8.1.1 Einführung 165
8.1.2 Bisheriger Forschungstand 166
8.1.3 Verbesserung des Forschungsstandes 167
8.2 ERSTER ABSCHNITT: NICHT ADJUSTIERTE UND NACH ABHÄNGIGKEIT DER AUGENSEITEN ADJUSTIERTE BOOTSTRAPVERFAHREN ZUR SCHÄTZUNG VON STANDARDFEHLERN IN MODELLEN DER CFA 168
8.2.1 [A] Bootstrapverfahren auf Basis simulierter Daten anhand eines Ein-Faktor-Modells mit drei Indikatoren unter Verwendung von PROC CALIS 168
8.2.1.1 Erzeugung der Simulationsdaten 168
8.2.1.2 Analyse der Simulationsdaten 170
8.2.1.3 Ergebnisse 171
8.2.2 [B] Bootstrapverfahren auf Basis von Patientendaten anhand eines Zwei-Faktor-modells mit acht Diagnoseverfahren 176
8.2.2.1 Bisheriger Forschungsstand 176
8.2.2.1.1 Diagnoseverfahren für die Glaukomerkrankung 176
8.2.2.1.2 Modellbildung 177
8.2.2.2 Reproduktion der AMOS-Analyse durch PROC CALIS 180

8.2.3 Bootstrap mit PROC CALIS unter Ziehung einzelner Augen (nicht nach Augenseitenkorrelation adjustiert) *181*
8.2.3.1 Ergebnisse 181
8.2.3.1.1 Verteilungsform der Pfadkoeffizientenschätzungen 181
8.2.3.1.2 Statistische Maßzahlen 184
8.2.3.1.3 Goodness of Fit 187
8.2.4 Bootstrap mit PROC CALIS unter Ziehung von Augenpaaren (nach Augenseitenkorrelation adjustiert) *189*
8.2.5 Vergleich der adjustierten und nichtadjustierten Bootstrapanalyse *192*
8.2.5.1 Diskussion und Fazit 195
8.3 Zweiter Abschnitt: Das Verhältnis zwischen Augenseitenkorrelation und Korrekturfaktor – Überprüfung der Simulationsergebnisse (1. Abschn., Teil A) durch Empirische Patientendaten 198
8.3.1 M1: Modell der CFA in der Glaukomdiagnose zur Schätzung der Augenseitenkorrelation latenter Variabler (pro Augenseite ein Faktor und vier Messverfahren) *198*
8.3.2 M2: Modell zur Schätzung von Korrekturfaktoren für das Modell 1 (Ein Faktor und vier Messverfahren) *200*
8.3.3 Ergebnisse *200*
8.3.4 Aufbereitung der Ergebnisse aus der Analyse mit simulierten Daten nach Maßgabe der Patientendatenergebnisse *201*
8.3.5 Aggregierung simulierter Korrekturfaktoren *202*
8.3.6 Interpolation zwischen aggregierten Korrekturfaktoren *203*
8.3.7 Vergleich der empirischen Korrekturfaktoren aus M2 mit angemessenen simulierten Korrekturfaktoren *203*
9 GESAMTFAZIT AUS DEN ANALYSEN SIMULIERTER UND EMPIRISCHER DATEN (KAP. 3–8). 206
10 AUSWIRKUNGEN DER VERLETZUNG VON MODELLANNAHMEN IN SEM IM LICHTE DER FORSCHUNGSLITERATUR UND DISKUSSION 213
10.1 Arbeitsprogramm 213
10.2 Bewertung von Modellen durch Reliabilitätskoeffizienten 214
10.2.1 Reliabilität unter Einfluss von Dimensionalität und Homogenität *214*
10.3 Reliabilitätsrelevante Spezifika von Messinstrumenten 217
10.3.1 Multidimensionalität des Konstrukts (Itemheterogenität) *217*
10.3.2 Korreliertheit der Messfehler *222*
10.3.2.1 Differenzierung von Fehlerkorrelationen laut Green und Hershberger 222
10.3.2.2 Differenzierung von systematischen Messfehlerkorrelationen 224
10.3.2.2.1 Methodenvarianz – Fehlerkovariation aufgrund von Common Method Bias 225
10.3.2.2.2 Fehlerkovariation aufgrund von Eigenschaften der Person – Beispiele 227
10.3.2.3 Vorzeichen der Fehlerkorrelation 229
10.3.2.4 Nicht berücksichtigte Fehlerkorrelationen 232

10.3.2.5 Berücksichtigte Fehlerkorrelationen 235
10.3.2.5.1 Fehlerkorrelationen infolge von Common Method Variance 235
10.3.2.5.2 Fehlerkorrelationen infolge von Eigenschaften der Person (Traits) 244
10.3.3 Simultane Betrachtung von Multidimensionalität und Messfehlerkorrelation 246
10.3.3.1 Methode und Daten 246
10.3.3.2 Resultate 247
10.3.3.2.1 Verletzung ausschließlich der Annahme unkorrelierter Fehler 248
10.3.3.2.2 Verletzung ausschließlich der Annahme essentieller τ –Äquivalenz i. S. d. Unidimensionalität 249
10.3.3.2.3 Gleichzeitige Verletzung beider Annahmen 249
10.4 Bewertung von Modellen durch Fit-indizes (Heene et al.) 251
10.4.1 Simulation und Design 252
10.4.2 Resultate 254
10.4.3 Alternative Annäherungen an das Problem mangelnder Sensitivität (EPC in Komb. mit MIs) 257
10.5 Das Modell zur Glaukomdiagnose mit acht Indikatoren – Koeffizienten der CTT und des Fits 258
10.5.1 Strukturanalyse des Referenzmodells 258
10.5.2 Reliabilität, Reliabilitätskoeffizient alpha und weitere Koeffizienten der CTT 260
10.5.3 Bewertung des Referenzmodells und seiner Variationen durch Kenngrössen der CTT 261
10.5.3.1 Interpretation der Koeffizienten auf Basis der Erkenntnisse Conways (1998) 263
10.5.4 Schätzt Alpha die Reliabilität verzerrungsfrei? 264
10.5.5 Bewertung durch Fit-Indizes 266
10.5.5.1 Die Modellentwicklung im Hinblick auf Erkenntnisse von Heene et al. (2012) 267
10.6 Schlussbemerkungen und Ausblick 269
10.6.1 Modellannahmen und Schätzmethode 269
10.6.2 Analyse der Patientendaten und Forschungsliteratur 273
11 ANHANG 275
11.1 SPSS-Programmcode zur künstlichen Erzeugung von Daten für AMOS-Modelle mit einem Faktor und drei Indikatoren 275
11.1.1 Zusammenstellung der wichtigsten Befehle 275
11.1.2 Ausführliche Syntax 275
11.2 SPSS-Programmcode zur künstlichen Erzeugung von Daten für AMOS-Modelle mit zwei Faktoren und sechs Indikatoren 277
11.3 Proc CALIS Lineqs-Modelle mit einem Faktor und drei Indikatoren 280
11.3.1 SAS-Programmcode 280
11.3.1.1 Vollständiger kommentierter Programmcode des Grundmodells 280
11.3.1.2 Programmcodeausschnitt der anderen Modelle 282
11.3.1.2.1 Modell mit zwei positiv korrelierten Fehlern 282
11.3.1.2.2 Modell mit zwei negativ korrelierten Fehlern 283
11.3.1.2.3 Modell mit drei positiv korrelierten Fehlern 283
11.3.2 SAS-Output 284

11.3.2.1 Modell mit einem Faktor und drei unkorrelierten Fehlern; n=1000, 1000 Durchläufe 284
11.3.2.2 Modell mit einem Faktor und zwei positiv korrelierten Fehlern n=1000, 1000 Durchläufe 285
11.3.2.2.1 A korrekt spezifiziert 285
11.3.2.2.2 B fehlspezifiziert 286
11.3.2.3 Modell mit einem Faktor und zwei negativ korrelierten Fehlern n=1000, 1000 Durchläufe 287
11.3.2.3.1 A korrekt spezifiziert 287
11.3.2.3.2 B fehlspezifiziert 288
11.3.2.3.3 Korrelationsmatrix der PARMS aus MODEL mit 1 Faktor und neg. korrel.(r=-0,3) fehlspez. Fehler 288
11.3.2.4 Modell mit einem Faktor und drei positiv korrelierten Fehlern n=1000, 1000 Durchläufe 288
11.3.2.4.1 A korrekt spezifiziert 288
11.3.2.4.2 STDERR aus LINEQS MODEL mit 1 Faktor 3 pos. korr. korr. spez. Fehlern (1000 Iterationen) 289
11.3.2.4.3 Korrektur: Entfernung der Iteration 796 wegen extremen Ausreißern – SE (beta_f2) = 13,5, SE (beta2) = 7,5 290
11.3.2.4.4 B fehlspezifiziert 290
11.3.2.4.5 Korrelationsmatrix der PARMS aus LINEQS MODEL mit 1 Faktor u. 3 pos. korr. fehlspez. Fehlern. 291
11.4 PROC CALIS LINEQS-MODELLE MIT EINEM FAKTOR UND VIER ODER FÜNF INDIKATOREN 291
11.4.1 A SAS-Programmcode 291
11.4.1.1 Vollständiger kommentierter Programmcode des Modells 1 291
11.4.1.2 Programmcodeausschnitt der anderen Modelle 293
11.4.2 B SAS-Output 294
11.4.2.1 Modell 1 294
11.4.2.2 Modell 2 295
11.4.2.3 Modell 3 296
11.4.2.4 Modell 4 297
11.4.2.5 Ursprüngliches Modell 3 mit vier Indikatoren 298
11.4.2.6 Ursprüngliches Modell 4 mit vier Indikatoren 299
11.5 PROC CALIS LINEQS-MODELLE MIT ZWEI FAKTOREN UND SECHS INDIKATOREN 299
11.5.1 A Programmcode 299
11.5.1.1 Modelle mit unabhängigen Faktoren 299
11.5.1.1.1 Vollständiger kommentierter Programmcode des Modells 1 (Grundmodell) 299
11.5.1.1.2 Programmcodeausschnitt Modell 2 302
11.5.1.1.3 Programmcodeausschnitt Modell 3 303
11.5.1.2 Modelle mit abhängigen Faktoren 304
11.5.1.2.1 Variablenerzeugung und Bildung der Indikatorenkorrelationsmatrix 304
11.5.1.2.2 Vollständiger, komment. und modellspezifischer Programmcode des Modells 4 (Grundmodell) 306
11.5.1.3 Programmcodeausschnitt der Modelle 5 bis 9 308
11.5.1.3.1 Modell 5 308
11.5.1.3.2 Modell 6 308
11.5.1.3.3 Modell 7 309
11.5.1.3.4 Modell 8 309
11.5.1.3.5 Modell 9 309

11.5.1.4 Programmcodeausschnitte der Varianten bisheriger Modelle und Zusatzmodelle 309
11.5.1.4.1 Modell 5 k 309
11.5.1.4.2 Modell Z4 k 310
11.5.1.4.3 Modell Z4 310
11.5.1.4.4 Modell 8 k 310
11.5.1.4.5 Modell 2- und 3 k_g 310
11.5.1.4.6 Model 8-, 5-, Z 4-, 6- und 9 k_g 312
11.5.2 B SAS-Output 312
11.5.2.1 Modelle mit unabhängigen Faktoren 312
11.5.2.1.1 Modell 1 312
11.5.2.1.2 Modell 2 314
11.5.2.1.3 Modell 3 316
11.5.2.2 Modelle mit abhängigen Faktoren 319
11.5.2.2.1 Modell 4 319
11.5.2.2.2 Modell 5 321
11.5.2.2.3 Modell 6 322
11.5.2.2.4 Modell 7 323
11.5.2.2.5 Modell 8 324
11.5.2.2.6 Modell 9 325
11.5.2.3 Varianten bisheriger Modelle und Zusatzmodelle 326
11.5.2.3.1 Modell 5 k 326
11.5.2.3.2 Modell Z4 k 327
11.5.2.3.3 Modell Z4 328
11.5.2.3.4 Modell 8 k 329
11.5.2.3.5 Modell 2_k_g 330
11.5.2.3.6 Modell 3 k_g vor Ausschluss von Iteration 1560: 331
11.5.2.3.7 Modell 3 k_g nach Ausschluss von Iteration 1560: 334
11.5.2.3.8 Modell 8 k_g 337
11.5.2.3.9 Modell 5 k_g 338
11.5.2.3.10 Modell Z4 k_g 339
11.5.2.3.11 Modell 6 k_g 340
11.5.2.3.12 Modell 9 k_g nicht ausreißerbereinigt 341
11.5.2.3.13 Modell 9 k_g ausreißerbereinigt 342
11.6 ADJUSTIERUNG DER STANDARDFEHLERSCHÄTZUNG NACH DER INTRAINDIVIDUELLEN ABHÄNGIGKEIT ZWISCHEN PAARIGEN ORGANEN – SIMULATION UND ANWENDUNG IN DER GLAUKOMDIAGNOSE 343
11.6.1 Erster Abschnitt, Teil A: (Vergleich der) Bootstrapverfahren auf Basis simulierter Daten anhand eines Ein-Faktor-Modells mit drei Indikatoren 343
11.6.1.1 A Vollständiger kommentierter Programmcode des Modells zur Datensimulation der Korrelationsstruktur K6 (siehe 343
11.6.1.2 Tabelle 11; R Glaukomschweregrad = R Fehlerkorrelation = 0.5) 343

11.6.1.3 B Auszüge aus den Programmcodes der Modelle zur Datensimulation der Korrelationsstrukturen K1 bis K5 und K7 bis K16. 351
11.6.2 Erster Abschnitt, Teil B: (Vergleich der) Bootstrapverfahren auf Basis von Patientendaten anhand eines Zwei-Faktoren-Modells mit acht Diagnoseverfahren 353
11.6.2.1 Programmcode 353
11.6.2.1.1 PROC CALIS Lineqs-Bootstrap unter Ziehung einzelner Augen (nicht adjustierte Analyse) und unter Ziehung von Augenpaaren (adjustierte Analyse) 353
11.6.2.1.2 Vergleich des nichtadjustierten und adjustierten Bootstrapverfahrens. 362
11.6.2.2 Output zur adj. und nichtadj. Bootstrapanalyse und Vergleich beider Verfahren 371
11.6.2.2.1 empirische Standardfehler der Pfadkoeff. der adj. Analyse 371
11.6.2.2.2 empirische Standardfehler der Pfadkoeff. der nicht adj. Analyse 373
11.6.2.2.3 Quotienten aus den empirischen SE der Pfadkoeff. beider Verfahren 376
11.6.2.2.4 Empirische SE der Fehlerkorrelationen der winsorierten adj. und nicht adj. Analysen 377
11.6.2.2.5 Empirische SE der Fehlerkorr. d. winsorierten Analysen beider Verfahren: Quotienten 378
11.6.2.2.6 Pfadkoeff. des Faktors Konzentration: Median und weitere stat. Kennwerte der adj. Analyse 379
11.6.2.2.7 Pfadkoeff. des Faktors Konzentration: dito, nichtadjustierte Analyse 381
11.6.2.2.8 Verteilung der Chi-Quadrat-Statistik des globalen Goodness of Fit-Tests der nichtadj. Analyse 384
11.6.2.2.9 Verteilung der Chi-Quadrat-Statistik des globalen Goodness of Fit-Tests der adj. Analyse 385
11.6.3 Zweiter AbschNitt: Das Verhältnis zw. Augenseitenkorrelation und korrekturfaktor – Überprüfung der Simulationsergebnisse (1 Abschn., Teil A) durch empir. Patientendaten 387
11.6.3.1 A Programmcode für M2 387
11.6.3.2 B Dokumentation des SAS-Outputs von M2 394
11.6.4 Zu Teil A: Mögliche Kombinationen von Augenseitenkorrelation des Schweregrads und des Messfehlers bei Augenseitenkorrelation des Messverfahrens von 0,8 396
11.7 AUSWIRKUNGEN DER VERLETZUNG VON MODELLANNAHMEN IN SEM IM LICHTE DER FORSCHUNGSLITERATUR UND DISKUSSION 398
11.7.1 Programmcode analog Yang et al. (2012) zur Errechnung der Varianzanteile 398
11.7.2 Modelle zur Erfassung des habituellen Wohlbefindens, Eid et al. (2011) 401
11.7.2.1 Zweidimensionales Modell 401
11.7.2.2 Zweidimensionales Modell mit Methodenfaktor 402
11.7.3 Modellannahmen und Schätzmethoden – Nichtnormalität 403
12 ABBILDUNGSVERZEICHNIS 407
13 TABELLENVERZEICHNIS 410
14 LITERATURVERZEICHNIS 412
15 EIDESSTATTLICHE ERKLÄRUNG 416

1 Einleitung

Zusammenfassung

Hauptziel des Promotionsprojektes ist es, die Aussagekraft eines bisherigen Strukturgleichungsmodells (SEM) zur Glaukomdiagnose zu vergrößern. Mit Hilfe eines verbesserten Modells kann einerseits die Genauigkeit erhöht werden, mit welcher der Schweregrad der Glaukomerkrankung eines individuellen Patienten eingeschätzt wird. Die Einschätzung des Glaukomschweregrades ist für die Wahl der angemessenen Therapie, die Beschreibung des Erkrankungsverlaufs und schließlich die Prognose des Patienten von Bedeutung. Andererseits kann die statistische Bewertung der Diagnoseverfahren verbessert werden. Langfristig soll das Dissertationsprojekt dem Ziel dienen, Glaukomdiagnostik auf Basis von SEMs zu etablieren. Die Erweiterung des Standardinstrumentariums zur Diagnose dieser verhältnismäßig häufigen, schweren Augenerkrankung könnte einen Beitrag zur Verbesserung der Bevölkerungsgesundheit leisten.

Einführung

Chronische Glaukome sind Erkrankungen, bei denen typische Schädigungen des Sehnervs auftreten, die Gesichtsfeldeinschränkungen zur Folge haben und ohne Behandlung zur Erblindung führen können. Glaukome sind häufig, aber keineswegs bei jedem Patienten mit erhöhtem Augeninnendruck verbunden. Frühzeitige Diagnose ist beim Glaukom (Grüner Star) wichtig, da die Erkrankung vom Patienten erst in einer relativ fortgeschrittenen Phase wahrgenommen wird, aber einmal entstandene Schädigungen irreversibel sind.

Die statistische Bewertung von Diagnoseverfahren geschieht durch Vergleich mit i. A. mehreren Referenzverfahren. Dieses Vorgehen kann voneinander abweichende Bewertungen produzieren. Die Gründe hierfür können darin liegen, dass die Verfahren unterschiedliche Krankheitsaspekte messen (z. B. sensorische oder morphologische Schäden) oder dass nicht beobachtete Störgrößen (z. B. Konzentration des Patienten) die Verfahren unterschiedlich stark beeinflussen. Lineare Strukturgleichungsmodelle (SEMs) ermöglichen die einheitliche Bewertung von Referenz- und neuen Diagnoseverfahren unter Berücksichtigung von Störgrößen und zeigen auf, wie gut diese den tatsächlichen Glaukomschaden quantifizieren. Ein individueller Patientenindex für den Glaukomschweregrad auf Basis von SEMs ist die beste Annäherung an einen allgemeinen Goldstandard.

Bisheriger Forschungsstand

Letzter Stand der bisherigen Forschung und Ausgangspunkt dieses Dissertationsprojekts ist ein publiziertes Modell der konfirmatorischen Faktorenanalyse (CFA), die einen Spezialfall eines SEM darstellt.[1] Es wird hier nicht in Gleichungsform, sondern grafisch wiedergegeben (siehe Abb.).

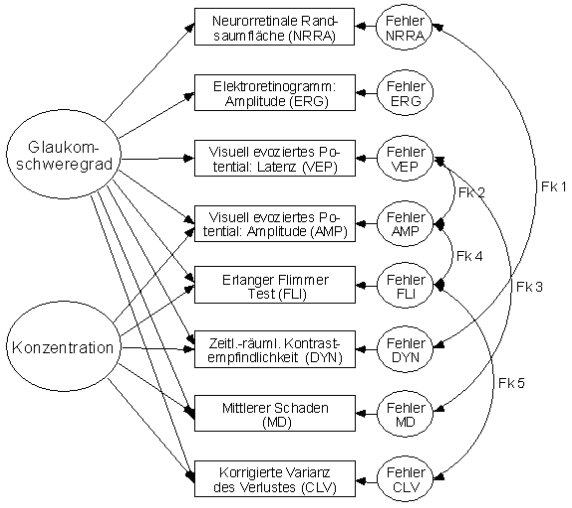


Abb.: Zwei-Faktoren-Modell für Diagnoseverfahren mit abhängigen Messfehlern

Dieses Modell wurde entwickelt anhand von Daten des Erlanger Glaukomregisters, das als klinisches Register zur Langzeitbeobachtung von Patienten mit manifestem Glaukom und Verdacht auf Glaukom angelegt ist (8.2.2). 237 Patienten durchliefen vollständig eine Reihe von Diagnoseverfahren, die durch unterschiedliche Formen der Reizgebung und Reaktionserfassung charakterisiert sind und zudem einen unterschiedlichen Grad von Objektivität aufweisen. In der Abbildung werden acht Diagnoseverfahren (Rechtecke) veranschaulicht, die (von oben nach unten) folgenden Gruppen zugeordnet sind: Verfahren 1: morphologische, Verfahren 2–4: elektrophysiologische und Verfahren 5–8: psychophysische Verfahren. Die Diagnoseinstrumente werden einerseits von zwei nichtbeobachtbaren Faktoren (Ellipsen) beeinflusst, dem Glaukomschweregrad und der Konzentrationsfähigkeit der Patienten bei der Absolvierung diagnostischer Tests. Andererseits beeinflussen Messfehlerkomponenten (Kreise) die Instrumente. Die Einflüsse werden durch einfache Pfeile (Pfade) dargestellt und angenommene Korrelationen zwischen den Messfehlerkomponenten durch Doppelpfeile symbolisiert. In diesem Modell werden fünf Fehlerkorrelationen (Fks) angenommen. Das Modell der CFA

[1] Martus (2001a), S. 269.

enthält für die beiden Faktoren jeweils ein Messmodell. Für das Modell wurden Pfadkoeffizienten geschätzt, die das Maß der Beeinflussung angeben. Die Richtigkeit einer Modellannahme über die Korrelationsstruktur von Messfehlern ist jedoch schwierig zu überprüfen. Bislang wurden auf Basis des abgebildeten Modells Bootstrapanalysen zur Schätzung der Standardfehler der Pfadkoeffizienten mit der Software AMOS durchgeführt. In der Analyse mit AMOS kann nicht nach der Abhängigkeit adjustiert werden, die zwischen den beiden Augen eines Patienten besteht. Daher ging man bisher von einer Ad-hoc-Lösung für die angemessene Schätzung für den SE durch den Korrekturfaktor 1–1,41 aus. Bei der Interpretation des globalen Tests für die Güte der Anpassung des Modells an die Daten wurde diese Abhängigkeit daher bislang in Form eines konservativen Korrekturfaktors von 1–0,5 berücksichtigt.

Forschungsvorhaben des Dissertationsprojektes

Im Dissertationsprojekt wird versucht, die möglichen verzerrenden Auswirkungen von fehlerhaften Annahmen (Fehlspezifikation) auf die Bewertung der Messverfahren in ihrer Richtung und Größenordnung abzuschätzen. Die Abschätzung des durch Modellfehlspezifikation induzierten Bias der interessierenden Pfadkoeffizienten erfolgt zunächst in einem dreistufigen Projekt, das dem Konzept von Peter Martus estspricht und die Gliederungspunkte A - C umfass. Im Anschluss (D) wird die Frage überprüft, ob anhand empirischer Daten ermittelte Korrekturfaktoren in gleichem Maß von dem empirischen Zusammenhang zwischen den Augen eines Patienten abhängen wie simulierte Korrekturfaktoren von dem entsprechenden simulierten Zusammenhang. Der abschliessende Gliederungspunkt (E) diskutiert Resultate der Datenanalyse vor dem Hintergrund von Forschungsliteratur. Die Gliederungspunkte D und E fußen auf einem Konzept des Autors.

(A) Nach bestimmten Modellvorgaben werden Daten mit festgelegter Korrelationsstruktur erzeugt und Auswirkungen diverser Fehlspezifikationen von Messfehlerkorrelationen untersucht. Hierbei handelte es sich um Korrelationen, die zwischen den Fehlertermen *verschiedener* Messverfahren bestehen. Fehlspezifikationen können sich in sogenannten Messmodellen auf die Höhe der Pfadkoeffizientenschätzung auswirken, welche die Stärke des Einflusses von latenten Variablen (z. B. Glaukomschweregrad) auf manifeste Variable (Messverfahren, z. B. Diagnoseverfahren) wiedergibt (Kap. 3). Werden Messmodelle (im Unterschied zum abgebildeten CFA) über ein sog. Strukturmodell miteinander verknüpft, das den Zusammenhang zwischen mindestens zwei latentenVariablen (z. B. die Schweregrade zweier Krankheiten) beschreibt, liegt ein SEM vor, in dem Fehlspezifikationen von Fehlertermen auch Pfadkoeffizienten des Strukturmodells beeinflussen (Kap. 4) können. Es wird die Software AMOS

(SPSS-kompatibel) verwendet. Bei der Simulation *einer Stichprobe* im Umfang von ca. 1500 werden in SPSS approximativ standardnormalverteilte Zufallsvariable erzeugt, aus welchen wiederum manifeste Variable konstruiert werden (11.1.2). Aus den manifesten Variablen werden je nach Modell empirische Kovarianzmatrizen erzeugt, die die Grundlage der jeweiligen Modellschätzungen bilden. Die Zufallsvariablen sind vollständig unabhängig voneinander, so dass die Korrelationsstruktur perfekt kontrollierbar ist.

(B) Auf Grundlage der Modellanalysen aus A werden Simulationsdaten erzeugt, die zusätzlich die *Variabilität der Korrelationen* berücksichtigen. Es wird die Software PROC CALIS aus dem Softwarepaket SAS verwendet, da die Leistungsgrenzen von AMOS auch hinsichtlich des Datenmanagements überschritten werden. Bei der Simulation einer Anzahl von mindestens n = 1000 Stichproben, die jeweils einem Umfang von 1000 aufweisen, wird grundsätzlich wie in A verfahren. Die Zufallsvariablen sind jedoch nicht perfekt unabhängig voneinander, sondern aufgrund von Zufallsschwankungen im Allgemeinen untereinander schwach korreliert. Die n simulierten Stichproben weichen also voneinander zufallsbedingt ab hinsichtlich des Mittelwertes, der Varianz und der Korrelationsstruktur der elementaren Zufallsvariablen (11.3.1.1). Unter diesen veränderten Bedingungen werden die an Messmodellen in Kap. 3 angestellten Untersuchungen (Ein-Faktor-Modelle) unter dem Variabilitätsaspekt in Kap. 5 also noch einmal durchgeführt. Entsprechend werden die Analysen zu SEMs in Kap. 4 (Zwei-Faktor-Modelle) in Kap. 7 fortgesetzt. Ein Zwischenkapitel (Kap. 6) beschäftigt sich mit Ein-Faktor-Modellen mit mehr als drei Variablen und ist als Erweiterung von Kap. 5 zu verstehen.

(C) Die ersten beiden Stufen dieses Forschungsvorhabens (A und B) behandelten die Auswirkungen der Fehlspezifikationen einer bestimmten Art von Messfehlerkorrelationen. Es handelte sich um Korrelationen, die zwischen den Fehlertermen *verschiedener* Messverfahren bestehen. In der dritten Stufe jedoch werden zur Beantwortung der Hauptfragestellung Messfehlerkorrelationen anderer Art behandelt: Es geht um Korrelationen zwischen den Ergebnissen *eines* jeweiligen Messverfahrens, das sowohl am linken und als auch rechten Auge eines Patienten durchgeführt wird (8.1.1). Dieser Fall ist immer gegeben, wenn es sich um vergleichbare Untersuchungen an *paarigen Organen* handelt.

Die Analyse wird *zunächst durchgeführt auf der Basis von simulierten Daten, die entsprechend dem Vorgehen in A erzeugt werden* (8.2.1). Hierbei wird ein Simulationsmodell verwendet, das Daten für paarige Organe erzeugt, wobei drei bestimmte Messverfahren in jedem der beiden Teilorgane vorgesehen sind (8.2.1.1). Die von einem bestimmten Messverfahren hervorgebrachten Ergebnisse des linken Teilorgans korrelieren somit mit jenen des entspre-

chenden rechten Teilorgans. Die Höhe dieser rechts-links-Korrelation (intraindividuelle Abhängigkeit der Messverfahren) hängt ab von zwei Komponenten ab, der entsprechenden Korrelation der latenten Variablen „Glaukomschweregrad“ und der latenten Variablen „Messfehler“. Beide Komponenten werden in dieser Untersuchung systematisch variiert, so dass 16 verschiedene Korrelationsstrukturen resultieren[2]. Die Simulationsdaten werden anschließend mit einem aus drei Messverfahren bestehenden Messmodell analysiert (8.2.1.2). Der Analyse zugeführt werden Stichproben, sogenannte Bootstrapsamples, die vielfach wiederholt im Ziehungsmodus „mit Zurücklegen“ aus einer simulierten Population (Sample) gezogen werden, die aus 400 virtuellen Individuen besteht. In einer Bootstrapvariante werden Samples aus den *800 einzelnen virtuellen linken und rechten Augen* (und zugehörigen Messdaten) gezogen, wobei die *intraindividuelle Abhängigkeit (Paarigkeit) der Augen ignoriert* wird. In der anderen Variante werden entsprechend Stichproben aus den *400 Augenpaaren* gezogen und damit der *intraindividuellen Abhängigkeit der Augen Rechnung getragen*. Die aus den beiden Bootstrapvarianten resultierenden Ergebnisse werden einander vergleichend gegenübergestellt, wobei das Augenmerk primär dem geschätzten Standardfehler (SE) der Pfadkoeffizienten gilt (8.2.1.3). Aus dem Verhältnis zwischen dem geschätzten SE, der die Datenpaarigkeit berücksichtigenden Analyse und dem geschätzten SE der Analyse, die die Paarigkeit ignoriert, wird für jede der 16 o. g. Korrelationsstrukturen ein sog. Korrekturfaktor ermittelt.

Diese „Bootstrapanalyse“ wird anschließend durchgeführt auf Basis des Zwei-Faktoren-Analysemodells anhand der Daten von 237 Patienten aus dem Erlanger Glaukomregisters. Der Ablauf dieser in 8.2.3 bis 8.2.5 dokumentierten Analyse gleicht dem Procedere in der vorausgehenden Bootstrapanalyse simulierter Daten. Es wird untersucht, ob die resultierenden Korrekturfaktoren in dem Bereich liegen, der auf Basis simulierter Daten ermittelten wurde. Untersuchungsgegenstand ist zudem, auf welche Weise die Korrekturfaktoren von der Augenseitenkorrelation der Messverfahren abhängen. Schließlich wird die mit der Veränderung der SE einhergehende Auswirkung auf die Chi-Quadrat-Statistiken analysiert.

(D) In 8.3 wird ein von Peter Martus publiziertes Modells vorgestellt, mit welchem u. a. *Korrelationen der latenten Variablen (Glaukomschweregrad und Messfehler) zwischen beiden Augenseiten* auf Basis von Patientendaten quantifiziert wurden. (8.3.1)[3]. Dieses Modell

[2] Für beide latente Variable wurden vier Seitenkorrelationskoeffizienten r simuliert (0, 0,5, 0,7 u. 0,9), deren Kombination $4^2 = 16$ Korrelationsstrukturen ergibt.
[3] Martus (2001b), S. 932.

gleicht strukturell dem Modell der Datensimulation in Kap. 8.2.1.1 und unterscheidet sich nur darin, dass es vier statt drei Messverfahren aufweist. (Die Hälfte der acht Diagnoseverfahren des oben abgebildeten Zwei-Faktoren-Modells sind implementiert.) Auf die mit diesem Modell analysierten Patientendaten wird ein Modell angewendet, das (bis auf ein viertes Messverfahren) strukturell mit dem Modell der Datenanalyse in 8.2.1.2 identisch ist. Nach dem in C vorgestellten Verfahren (Gegenüberstellung von adjustierter und nicht adjustierter Bootstrapanalyse) werden Korrekturfaktoren für die Schätzung des Standardfehlers der Pfadkoeffizienten von vier Messverfahren ermittelt (8.3.2). Diese an Patientendaten ermittelten Korrekturfaktoren (8.3.3) werden mit den auf Basis von simulierten Daten ermittelten Korrekturfaktoren verglichen. Hierbei wird die Frage überprüft, ob die anhand empirischer Daten ermittelten Korrekturfaktoren in gleichem Maß von den empirischen Augenseitenkorrelationen der latenten Variablen Glaukomschweregrad und Messfehler abhängen wie die simulierten Korrekturfaktoren von den entsprechenden simulierten Augenseitenkorrelationen (8.3.7). Dazu müssen zum Vergleich jene Kombinationen aus Schweregrad- und Messfehlerseitenkorrelation aus den 16 Korrelationsstrukturen ausgewählt werden, die den empirischen Augenseitenkorrelationen annäherungsweise entsprechen (8.3.4). Die simulierten, mehrfach geschätzten Korrekturfaktoren, die den ausgewählten Augenseiten-Korrelationsstrukturen entsprechen, werden jeweils aggregiert und gemittelt (8.3.5). Wenn empirische Augenseitenkorrelationen sich zwischen zwei verschiedenen simulierten Korrelationsstrukturen befinden, wird zwischen letzteren interpoliert (8.3.6). Ein Fazit der Analysen aus A bis D wird in Kap. 9 gezogen.

(E) Das Abschlusskapitel stellt Forschungsliteratur vor und diskutiert auf deren Basis Resultate der empirischen Datenanalyse. Es werden Modellannahmen in SEM behandelt, deren Verletzung gravierende Konsequenzen beinhaltet: Unidimensionalität des Konstrukts und Unkorreliertheit der Messfehler. Es wird analysiert, wie sich die Verletzung einer der beiden Annahmen u. a. auf die Reliabilität und die Anpassungsgüte eines Testmodells auswirkt. Neben dem Effekt der Multidimensionalität des Konstrukts (10.3.1) wird vorrangig der Effekt korrelierter Fehler auf die Reliabilität (10.3.2) untersucht und in diesem Zusammenhang eine Systematik von Messfehlern dargestellt. Anschließend wird analysiert, wie sich das simultane Auftreten von Multidimensionalität und Fehlerkorreliertheit auswirkt (10.3.3). Im zweiten Teil dieses Kapitels werden die Konsequenzen der Verletzung der Unkorreliertheitsannahme in Bezug auf Fit-Indizes untersucht. (10.4). Anschließend wird anhand von Forschungsliteratur analysiert, welche Folgen die Verletzung von Modellannahmen in dem bereits eingangs vorgestellten Modell zur Glaukomdiagnose nach sich zieht. Reliabilitäts- und weitere Koeffi-

zienten der klassischen Testtheorie sowie Cronbachs *alpha* und Fit-Indizes werden für das herkömmliche Modell und Modelle berechnet, die zur Reliabilitätssteigerung um weitere Methodenfaktoren erweitert wurden. Alle Modelle werden im Hinblick auf Forschungsliteratur diskutiert (10.5). Schließlich wird die angewandte ML-Schätzung im Hinblick auf eine mögliche Verletzung von Modellannahmen (Unkorreliertheit der Error Scores, multivariate Normalverteilung der Observed Scores) erörtert und die Verteilung der Chi-Quadrat-Statistik im adjustierten Bootstrapverfahren reflektiert (10.6).

2 Lineare Strukturgleichungsmodelle in der medizinischen Diagnostik

2.1 ÜBERBLICK

Lineare Strukturgleichungsmodelle haben in den letzten Jahrzehnten nicht nur in den Sozial- und Wirtschaftswissenschaften zu einer weiten Verbreitung geführt, sondern auch in der Medizin Anwendung gefunden. Mit solchen Modellen kann man anhand von empirischen Daten Hypothesen über den Zusammenhang zwischen Merkmalen prüfen. Die Zusammenhangshypothesen werden hierbei auf der Grundlage theoretischer Überlegungen aufgestellt, so dass bei ihrer Formulierung keine aus den empirischen Daten stammende Information verwendet wird.

In linearen Strukturgleichungsmodellen finden sich Elemente aus drei Bereichen statistischer Methodik – der Faktoren-, der Pfad- und der Regressionsanalyse.

Die Prüfung von a priori formulierten Kausalhypothesen entspricht der *Faktorenanalyse* in ihrer konfirmatorischen Variante. (Die konfirmatorische Faktorenanalyse unterscheidet sich von der explorativen im Wesentlichen darin, dass sie nicht als hypothesengenerierendes, sondern -prüfendes Verfahren eingesetzt wird.)

Wie in der *Pfadanalyse* ist mit linearen Strukturgleichungsmodellen die Überprüfung kausaler Beziehungen zwischen beobachtbaren Messgrößen möglich. Darüber hinaus können auch Zusammenhänge zwischen nicht direkt beobachtbaren Größen oder zwischen beobachtbaren und nicht beobachtbaren Merkmalen untersucht werden.

Lineare Strukturgleichungsmodelle sind insbesondere in der medizinischen Diagnostik gut anwendbar. Denn bei der Diagnose von Krankheiten handelt es sich um die Bestimmung von Phänomenen, die sich häufig unmittelbarer Messung (Beobachtung) entziehen und vielfach nur mittelbar über verschiedene Indikatoren (diagnostische Messverfahren) feststellbar sind.

Die Messfehler, die mit der Durchführung diagnostischer Messverfahren stets verbunden sind, werden in linearen Strukturgleichungsmodellen (wie die interessierenden Krankheiten) als nicht beobachtbare Größen modelliert.

Die Arbeit mit einem linearen Strukturgleichungsmodell im Bereich der Diagnose beginnt damit, gut begründete *Hypothesen über Ursache-Wirkungs-Zusammenhänge* zu formulieren. Ein solcher Kausalzusammenhang könnte beispielsweise eine Struktur beschreiben, die zwi-

schen zwei Krankheiten vermutet wird. Beispiel: „Diabetes mellitus kann zu sekundärem Glaukom führen."

Weiterhin wird festgelegt, durch welche diagnostischen Verfahren die jeweiligen interessierenden Phänomene (Krankheiten) erfasst werden sollen. Eine bestimmte Krankheit kann hierbei durch ein oder mehrere diagnostische Verfahren beschrieben werden. Beispiele: „Diabetes mellitus kann u. a. anhand der Höhe des Blutzuckerspiegels (mg/dl) festgestellt werden."; „das sekundäre Glaukom kann u. a. anhand der Höhe des Augeninnendrucks (mm Hg laut Applanationstonometer) bestimmt werden."

Die Zusammenhänge zwischen den Erkrankungen als nicht direkt beobachtbaren Variablen einerseits und zwischen diesen und den direkt beobachtbaren diagnostischen Messverfahren andererseits werden im nächsten Schritt in *Pfaddiagrammen* visualisiert. Diesen Diagrammen entsprechend können auch sogenannte *Modellgleichungen* aufgestellt werden, die die Form von Regressionsgleichungen aufweisen. Die Pfaddiagramme und Modellgleichungen enthalten Informationen sowohl über die Richtung als auch die Stärke der postulierten Zusammenhänge.

Zuletzt wird mittels eines *Tests* überprüft, ob sich das Modell an empirisch erhobenen Daten bewährt. Ist die Diskrepanz zwischen dem Modell und dem empirischen Datum gering, werden üblicherweise die aufgestellten Kausalhypothesen insofern als bestätigt angesehen, als sie durch die empirischen Daten nicht verworfen wurden.

Lineare Strukturgleichungsmodelle können mit Hilfe einer speziellen statistischen Software erstellt werden. In dieser Arbeit wurde AMOS (Analysis for Moment Structures) verwendet, das als SPSS-Produkt das bislang bekannteste spezifische Computerprogramm LISREL ablöste. Es bietet sowohl die Möglichkeit, Modelle durch eine Reihe von Gleichungen vollständig zu beschreiben (AMOSBasic) als auch die Option, Grafische Modellierung zu betreiben, die ohne die Erstellung von Modellgleichungen durch den Benutzer auskommt (AMOSGraphics).

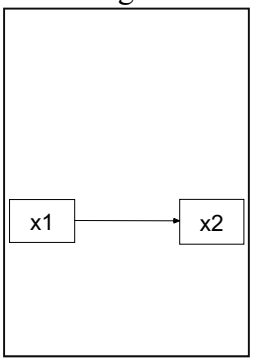
2.2 DAS KONZEPT DER LATENTEN VARIABLEN

In linearen Strukturgleichungsmodellen beschreiben Kausalhypothesen mögliche Zusammenhänge zwischen den Variablen des Modells. Eine Kausalhypothese kann in Form einer Strukturgleichung ausgedrückt oder graphisch dargestellt werden. Man stellt eine kausale Hypothe-

se grafisch dar, indem man zwei für bestimmte Variable stehende Symbole durch einen einfachen oder einen Doppelpfeil miteinander verbindet.[4]

Kausale Hypothesen können dabei beobachtbare, direkt messbare Variable beinhalten, die als *manifest* bezeichnet werden. In grafischen Darstellungen linearer Strukturgleichungsmodelle und in Pfadmodellen allgemein werden manifeste Variable durch Rechtecke symbolisiert. Beispielshypothese: „Die Dauer der praktischen Ausbildung eines Röntgenarztes beeinflusst den Prozentsatz der von ihm gestellten fehlerhaften Diagnosen." Hierbei wird „Ausbildungszeit" mit x1 und „Anteil der Fehldiagnosen" mit x2 bezeichnet (siehe Abbildung 1).

Abbildung 1: Kausalhypothese über zwei manifeste Variable

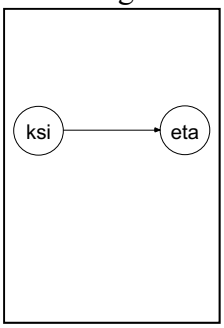


Kausale Beziehungen können auch formuliert werden für Variable, die nicht beobachtbar, also nur indirekt messbar sind und *latent* heißen. Das grafische Symbol für latente Größen ist der Kreis oder die Ellipse.

Bei latenten Variablen wird unterschieden zwischen endogenen, die im Modell erklärt werden sollen, und exogenen, die als erklärende Variable fungieren. Von den eingangs im Beispiel genannten Krankheiten ist Diabetes mellitus die exogene und das sekundäre Glaukom die endogene Variable. In den Termini der Regressionsanalyse können die endogenen Variablen als Kriteriums- und die exogenen als Prädiktorvariablen bezeichnet werden. In linearen Strukturgleichungen werden endogene Variable konventionell mit η (eta) und exogene Veränderliche mit ξ (ksi) benannt. In grafischen Darstellungen sind endogene Variable daran zu erkennen, dass Pfeile auf sie gerichtet sind, während exogene veränderliche, latente Größen sind, von welchen Pfeile ausgehen.

[4] In dieser einfachen Form einer graphisch dargestellten Kausalhypothese unterschlage ich sog. Fehlervariablen. Auf den Begriff „Fehlervariable" wird später eingegangen.

Abbildung 2: Kausalhypothese über zwei latente Variable



2.3 MESSMODELL UND STRUKTURMODELL

Ein Bestandteil eines linearen Strukturgleichungsmodells ist das sogenannte *Messmodell*. Es gibt im einfachsten Fall eine manifeste Variable an, durch die der Einfluss einer latenten Variablen gemessen wird, wobei λ die Stärke des Einflusses bezeichnet. Da die Varianz einer manifesten Variablen in der Praxis nicht vollständig durch eine latente Variable erklärt werden kann, beinhaltet ein Messmodell immer mindestens eine Fehlervariable, der der unerklärte Varianzanteil der manifesten Variablen zugeschrieben wird. Die folgenden beiden Abbildungen zeigen die Messmodelle für die latente endogene und die latente exogene Variable des einführenden Beispiels.

Abbildung 3: Messmodell der latenten exogenen Variablen ξ

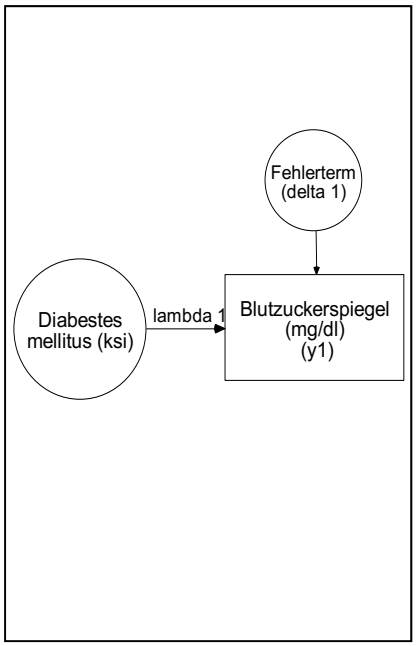
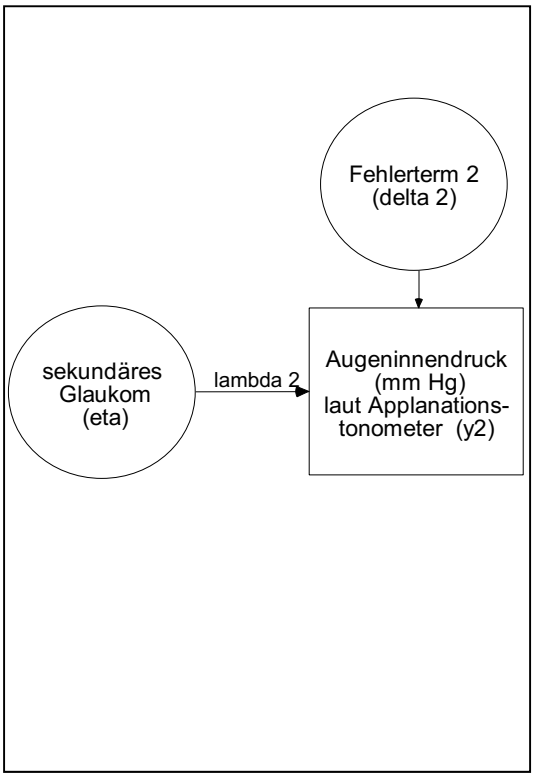


Abbildung 4: Messmodell der latenten endogenen Variablen η



In Gleichungsschreibweise lauten die Messmodelle

$$y1 = \lambda_1 \cdot \xi + \delta_1$$

$$y2 = \lambda_2 \cdot \eta + \delta_2$$

λ_1 und λ_2 heißen *Pfadkoeffizienten* und sind als *Korrelationen* zwischen der jeweiligen latenten und der zugehörigen Indikatorvariable aufzufassen (siehe auch 2.7.)

$$\lambda_1 = r_{y1\xi}$$

$$\lambda_2 = r_{y2\eta}$$

Im Sprachgebrauch der Faktorenanalyse kann eine latente Variable als *Faktor* und eine Korrelation zwischen ihr und der zugehörigen Indikatorvariablen als *Faktorladung* interpretiert werden.

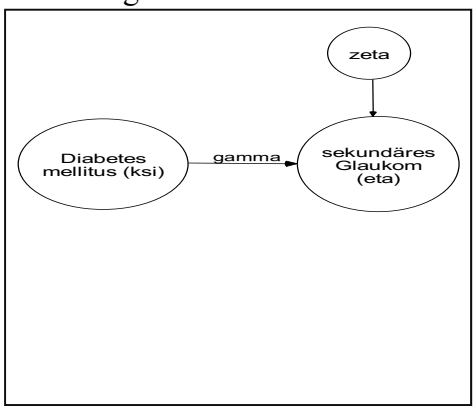
In einem sogenannten *Strukturmodell* werden die latenten Variablen miteinander verbunden.

Die latente endogene Variable sekundäres Glaukom (η_1) ist das Resultat der Wirkungen der latenten exogenen Variablen Diabetes mellitus (ξ) und eines Fehlerterms (ζ). Die Stärke des Einflusses der latenten exogenen Variablen wird durch den Pfadkoeffizienten γ bezeichnet.

Die Gleichung des Strukturmodells lautet somit $\eta = \gamma \cdot \xi + \zeta$.

In Abbildung 5 ist die grafische Darstellung des Strukturmodells zu sehen:

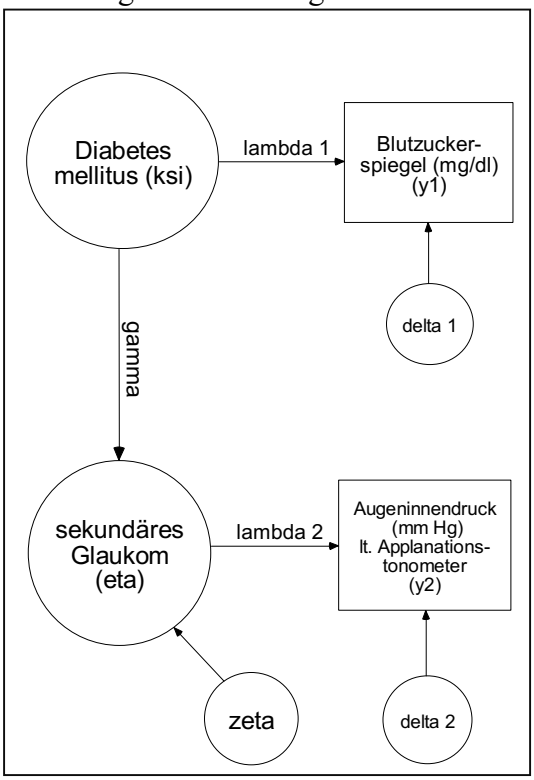
Abbildung 5: Strukturmodell mit einer endogenen und einer exogenen Variable



Indem man eine Verbindung zwischen der latenten exogenen und der latenten endogenen Variablen herstellt, werden die beiden obigen Messmodelle zu einem *vollständigen linearen Strukturgleichungsmodell* integriert.

In der medizinischen Praxis würde man versuchen, die beiden interessierenden Krankheiten Diabetes mellitus und sekundäres Glaukom jeweils durch mehrere Indikatorvariablen zu erfassen. Aus Gründen der einfachen Darstellung wird in Abbildung 6 das vollständige lineare Strukturgleichungsmodell jedoch mit nur zwei Indikatorvariablen skizziert:

Abbildung 6: Vollständiges lineares Strukturgleichungsmodell für zwei latente Variable



2.4 MATHEMATISCHE GRUNDLAGE EINES SEM

Strukturgleichungen sind einem System von Regressionsgleichungen ähnlich, die zur Spezifizierung, Schätzung und Testung eines Prozesses oder einer Theorie unter Beobachtung benutzt werden[5]. Die Gleichungen verbinden abhängige Variable mit einem Set von vorhersagenden Variablen, die wiederum abhängige Variable in anderen Gleichungen sein können. Diese Gleichungen und die assoziierten Varianz- und Kovarianzspezifikationen innerhalb der nichtabhängigen (= unabhängigen) Variablen sind das SEM-Modell. Die Implikationen dieser Spezifikationen für die Mittelwerte, Varianzen und Kovarianzen der gemessenen Variablen erhält man mit Hilfe von Matrixalgebra.

Im Rahmen der *Kovarianzstrukturanalyse* ist $\sum$ die Populationskovarianzmatrix und θ ein Vektor von Basisparametern wie Regressionskoeffizienten, Varianzen und Kovarianzen der unabhängigen Variablen. Im Kontext der *allgemeineren Strukturmodelle* sind die Mittelwerte der gemessenen Variablen ebenso eine Funktion von Basisparametern, die die Mittelwerte oder Intercepts von latenten, nicht gemessenen Variablen beinhalten können.

Gegeben die parametrischen Werte des unbekannten Parametervektors (θ), sind der Populationsparameter μ und die Kovarianzmatrix $\sum$ durch diese Parameter bestimmt, d. h.

$\mu = \mu(\theta)$ und $\sum = \Sigma(\theta)$. Das ist die zu testende „*Strukturhypothese*“.

In der Praxis werden der Vektor der Stichprobenmittelwerte $\bar{x}$ und die (empirische) Kovarianzmatrix S benutzt, um den unbekannten Parametervektor zu schätzen.

Wenn $\hat{\mu} = \mu(\hat{\theta})$ und $\hat{\Sigma} = \Sigma(\hat{\theta})$ nahe bei $\bar{x}$ bzw. S liegen[6], dann ist das SEM-Modell eine plausible Repräsentation der empirischen Daten.

In vielen Modellen sind die Mittelwerte nicht strukturiert als $\mu(\theta)$, $\hat{\mu} = \bar{x}$ und nur die Kovarianzmatrix ist von Interesse. Solche Modelle heißen Kovarianzstrukturmodelle.

Ein *Spezialfall von Kovarianzstrukturmodell*, der latente Faktoren wie in der (explorativen) Faktorenanalyse beinhaltet, ist die konfirmatorische Faktorenanalyse (CFA)

---

[5] Bentler & Stein (1992), S. 161.

[6] Verbalisiert heißt das, wenn die Schätzung für den wahren Stichprobenmittelwertsvektor (als Funktion des geschätzten wahren Parametervektors) nahe bei dem tatsächlichen Stichprobenmittelwertsvektor liegt und die Schätzung der wahren Kovarianzmatrix (als Funktion des geschätzten Parametervektors) nahe bei der empirischen Kovarianzmatrix liegt.

2.5 DIE METHODE DER KONFIRMATORISCHEN FAKTORENANALYSE

Die CFA ist, zusammenfassend formuliert, ein Spezialfall eines SEMs. Bestehen vollständige SEMs aus zwei Messmodellen und einem Strukturmodell, so beinhaltet die CFA mindestens ein Messmodell und kein Strukturmodell.

Da die CFA,ein Standardverfahren in anderen Anwendungsgebieten, in der Ophthalmologie mit wenigen Ausnahmen[7] kaum verbreitet ist, wird sie hier kurz vorgestellt. Zunächst wird hierbei von der unrealistischen Situation ausgegangen, dass ein vollkommener Goldstandard in der Glaukomdiagnose existiere, um einige Grundlagen zu erklären. Anschließend wird erklärt, was in der realistischen Situation eines nicht vorhandenen perfekten Goldstandards getan werden kann, um Diagnoseverfahren zu bewerten.

Man stelle sich vor, dass es einen Goldstandard gäbe, der fehlerfrei den Schweregrad einer Glaukomerkrankung erfassen könne. Unter dieser Voraussetzung kann man jedes beliebige quantitative Messverfahren zur Diagnose der Erkrankung bewerten, nachdem man die Korrelation zwischen dem Goldstandard und dem jeweiligen Messverfahren errechnet hat.

Wenn der Wert des Pearson'schen Korrelationskoeffizienten r größer als 0,9 beträgt, so liegt ein hoher positiver linearer Zusammenhang vor und das zu bewertende Verfahren wird als sehr gut brauchbar erachtet.

Beispiel: Angenommen, für zwei diagnostische Messverfahren sei bekannt, wie hoch sie mit dem Goldstandard korrelieren. Die Korrelation des Verfahrens 1 (x_1) mit dem Goldstandard G betrage $r_1 = 0{,}6$ und die des Verfahrens 2 (x_2) mit dem Goldstandard betrage $r_2 = 0{,}7$.

Dann lässt sich leicht zeigen[8], dass – unter einer Voraussetzung – die Korrelation der beiden Verfahren r_{12} untereinander gleich dem Produkt der Korrelationen zwischen dem jeweiligen Messverfahren und dem Goldstandard ist: $r_{12} = r_1 \cdot r_2 = 0{,}42$.

Die Voraussetzung lautet, dass neben dem Glaukomschaden kein weiterer *gemeinsamer* Faktor beide Messverfahren beeinflusst. Messverfahren, für die diese Bedingung zutrifft, heißen *bedingt unabhängig.* Das bedeutet, dass die Messverfahren nur vermittels des Schweregrades voneinander abhängig und sonst unabhängig sind.

[7] Martus et al. (2000), S. 1110, References 15-20.
[8] Bortz (1999), S. 463.

Wenn ein solcher weiterer gemeinsamer Faktor (Störfaktor), existiert, dann gilt, dass die Korrelation der beiden Verfahren untereinander größer als das Produkt ist: $r_{1,2} > r_1 \cdot r_2$.

Störfaktoren können in der *Person*, die sich den beiden diagnostischen Messverfahren unterzieht, angesiedelt sein. Dies wäre z. B. der Fall, wenn es sich bei x_1 und x_2 um verschiedene psychophysische diagnostische Tests handelte, wie die weiß auf weiß-Perimetrie und der Erlanger Flimmertest, bei welchen die *Konzentration* eine Rolle spielt.

Störfaktoren können auch in der *Messapparatur* begründet sein. Dies ist z. B. gegeben, wenn bei einer elektrophysiologischen Messung ein Potential abgeleitet wird und man als ein diagnostisches Verfahren die Amplitude und als ein zweites die zeitliche Latenz dieser Amplitude betrachtet. Beide abgeleitete Größen sind beeinflusst durch *Artefakte* beim speziellen Messvorgang.

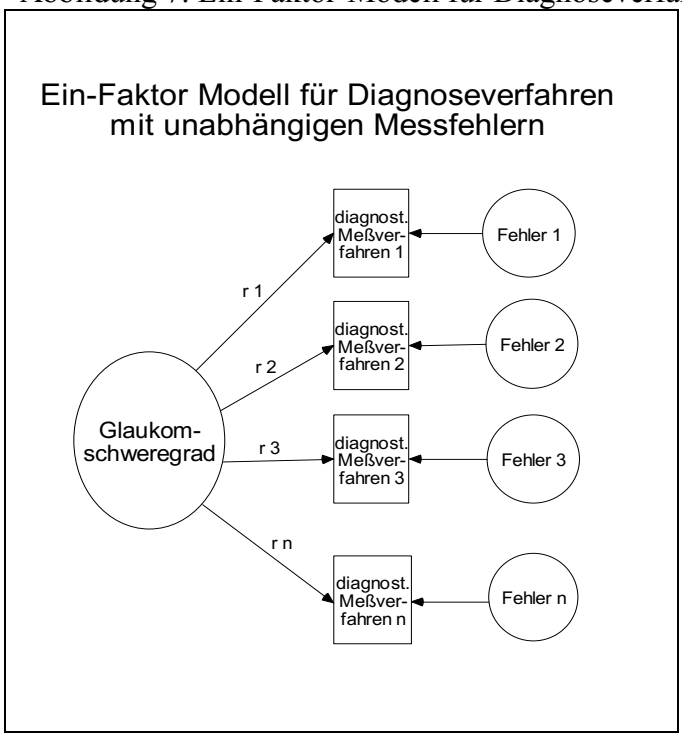
Detailliertere Betrachtung zu Störfaktoren, die ohne angemessene Modellierung zu Fehlerkovariation führen, finden sich in 10.3.2.

2.6 BEWERTUNG VON DIAGNOSEVERFAHREN DURCH BESTIMMUNG VON PFADKOEFFIZIENTEN

In der realistischen Situation eines *nicht vorhandenen perfekten Goldstandards* können nur die paarweisen Korrelationen der angewendeten Diagnoseverfahren untereinander bestimmt werden, in obigem Beispiel nur die Korrelation r_{12}.

Wenn jedoch mindestens drei verschiedene Diagnoseverfahren x_1, x_2 und x_3 angewandt wurden, die bedingt unabhängig sind, dann können aus den paarweisen Korrelationen r_{12} , r_{13} , r_{23} die Werte von r_1, r_2 und r_3 errechnet werden, ohne dass ein Goldstandard vorhanden ist (siehe Abbildung 7).

Abbildung 7: Ein-Faktor-Modell für Diagnoseverfahren mit unabhängigen Messfehlern



Das Gleichungssystem

$$r_{12} = r_1 \cdot r_2;\ r_{13} = r_1 \cdot r_3;\ r_{23} = r_2 \cdot r_3 \quad (1)$$

wird exakt gelöst durch[9]

$$r_1 = \sqrt{\frac{r_{12} \cdot r_{13}}{r_{23}}} \quad r_2 = \sqrt{\frac{r_{12} \cdot r_{23}}{r_{13}}} \quad r_3 = \sqrt{\frac{r_{13} \cdot r_{23}}{r_{12}}} \,.$$

Die Lösungen heißen im Sprachgebrauch der CFA Faktorladungen und im Sprachgebrauch der SEM Pfadkoeffizienten[10]. Ein Pfadkoeffizient von 0 zeigt an, dass keinerlei Zusammenhang zwischen der zugrundeliegenden Erkrankung und dem diagnostischen Verfahren besteht. Ein Pfadkoeffizient von 1 bedeutet, dass das Diagnoseinstrument die Erkrankung perfekt quantifiziert.

Die Annahme der bedingten Unabhängigkeit kann jedoch bei nur drei Messverfahren nicht überprüft werden[11]. Nur wenn mehr als drei Messverfahren zur Verfügung stehen, kann unter-

[9] Kenny (1979).

[10] Wenn der nicht vorhandene (nicht messbare) Goldstandard und das Messverfahren standardisiert und normalverteilt sind, dann ist der Pfadkoeffizient gleich der Korrelation.

[11] Da aus drei Gleichungen drei Unbekannte geschätzt werden, ist das Modell mit drei Messverfahren „genau identifiziert". Es können keine weiteren Parameter geschätzt werden, um die Annahme der bedingten Unabhängigkeit zu überprüfen.

sucht werden, ob ein Modell der CFA in dieser Hinsicht eine angemessene Beschreibung der Daten darstellt.

Wenn vier oder mehr Messverfahren gegeben sind, ist *keine exakte Bestimmung* der Pfadkoeffizienten analog obiger Formel mehr möglich. Stattdessen müssen die Pfadkoeffizienten nach einem zu bestimmenden Optimierungskriterium *geschätzt* werden.

Die Annahme der bedingten Unabhängigkeit kann bei mehr als drei Diagnoseverfahren bis zu einem gewissen Grad aufgeweicht werden. Stehen beispielsweise vier Diagnoseverfahren zur Verfügung, ergeben sich als beobachtete Parameter vier Varianzen und sechs Kovarianzen, aus denen nur acht Pfadkoeffizienten geschätzt werden müssen. Vorausgesetzt ist hierbei, dass zur Festlegung einer Metrik die Varianz der latenten Variablen und aller vier Fehlervariablen (auf den Wert 1) festgelegt sind. Es können daher zwei zusätzliche Parameter geschätzt werden, so dass das Modell (genau) identifiziert ist (Anzahl der beobachteten und zu schätzenden Parameter ist gleich). (Für weitere Informationen zur Identifizierbarkeit siehe 2.7.) In dieser Situation kann man untersuchen, ob Korrelationen zwischen Messverfahren bestehen, die nicht auf die Glaukomerkrankung zurückgehen. Auch wenn dies der Fall sein sollte, ist die CFA anwendbar. Die statistische Methode funktioniert auch dann noch, wenn die Annahme der bedingten Unabhängigkeit abgeschwächt wird, d. h, wenn einige der Messverfahren *bedingt abhängig* sind.

Die Annahme der bedingten Unabhängigkeit kann auf dreierlei Weise überprüft werden:

1. Durchführung einer Korrelationsanalyse in der einer Kontrollgruppe. Es wird hierbei erwartet, dass die paarweisen Korrelationen zwischen den diagnostischen Messverfahren einen Wert nahe Null aufweisen. Wenn jedoch wesentlich größere Korrelationen beobachtet werden, dann beeinflussen Faktoren, die nichts mit der Erkrankung zu tun haben (Störgrößen) die Messverfahren.

2. Alle paarweisen Korrelationen der diagnostischen Messverfahren, die im Patientenkollektiv beobachtet wurden, werden verglichen mit den Korrelationen, die auf Basis der Modellannahmen geschätzt werden[12]. Dieser Vergleich klärt die Frage, ob auf Basis eines kleineren Sets von Pfadkoeffizienten ein größeres Set von paarweisen Korrelationen geschätzt werden kann. Liegen die *geschätzten* paarweisen Korrelationen nahe an den jeweiligen *beobachteten*,

[12] Siehe 2.4. Der Unterschied zwischen geschätzter Korrelations- und geschätzter Kovarianzmatrix ist in diesem Zusammenhang unwesentlich.

so gelten die im Modell getroffenen Annahmen der bedingten Unabhängigkeit als vereinbar mit den empirischen Daten.

3. Sensitivitätsanalyse. Angenommen, man hat sich für ein konfirmatorisches Faktormodell mit einer bestimmten Anzahl von Diagnoseverfahren – mindestens vier – entschieden und die Pfadkoeffizienten geschätzt. In dieser Situation sollten Modelle untersucht werden, die nur eine Teilmenge dieser Diagnoseverfahren beinhalten. Die Pfadkoeffizienten dieser Untermodelle dürfen dann nicht wesentlich von den Pfadkoeffizienten abweichen, die im Modell der Wahl geschätzt wurden.

2.7 IDENTIFIZIERBARKEIT

Unter 2.4 wurde bereits dargelegt, dass die hier interessierenden SEM als Spezialfall der CFA und damit der Modellklasse der Kovarianzstrukturmodelle zuzurechnen sind. In diesen sind die (beobachteten) Stichprobenmittelwerte nicht von Belang, sondern nur die (beobachtete) Kovarianzmatrix. Unter 2.6 war von der Aufweichung der Annahme der bedingten Unabhängigkeit die Rede, wenn mehr als drei Messverfahren zur Verfügung stehen und die Identifizierbarkeit eines Modells mit vier Messverfahren wurde angeschnitten. Im Folgenden werden einige Aspekte zur Identifizierbarkeit in SEM benannt, die in unserem Zusammenhang relevant sind.

1. Die konkrete Schätzung von Parametern in SEM (allgemein: in Faktormodellen) und die grundsätzliche Bestimmbarkeit derselben, ihre Identifizierbarkeit, sind zu unterscheiden. Letztere betrifft die Frage, ob die Informationen, die der beobachteten Kovarianzmatrix entnommen werden können, *dem Grunde nach* zur Bestimmung der Parameter eines bestimmten Modells ausreichen. Die grundsätzliche Bestimmbarkeit geht hier von der Vorstellung aus, dass das Modell in der unbekannten Grundgesamtheit gelte und die Kovarianzmatrix bekannt sei. Die Parameterschätzung meint jedoch die Möglichkeit der Schätzung von Parametern in einer konkreten, aus dieser Grundgesamtheit gezogenen Stichprobe. Nur wenn Identifizierbarkeit theoretisch gegeben ist, ist auch eine Parameterschätzung praktisch möglich[13]. Wie oben bereits erwähnt, sind wir in SEM nur an der Kovarianzstruktur interessiert und nicht an den Mittelwerten der beobachteten Variablen und den Leichtigkeits- bzw. Schwierigkeitsparametern, da sie bei der Aufklärung des Zusammenhangs zwischen den latenten Variablen (Strukturmodell) und zwischen beobachteten und latenten Variablen (Mess-

[13] Eid et al. (2011), S. 862

modell) entbehrlich sind. Die Werte werden als Abweichungen der beobachteten Variablen vom Erwartungswert aufgefasst, $A_i = Y_i - E(Y_i)$, so dass die Erwartungswerte in diesen Modellen nicht mehr von Bedeutung sind[14]. Beispielsweise erhält man durch die Zentrierung statt $Y_1 = \alpha_1 + \lambda_1 \eta_1 + \varepsilon_1$ die Modellgleichung $A_1 = \lambda_1 \eta_1 + \varepsilon_1$, wobei E($\eta_1$) = 0 (vergleiche Modellgleichung 2 in 2.3).

Wir behandeln in unserem Zusammenhang stets Messmodelle kongenerischer Variabler, in welchen die unterschiedlichen Pfadkoeffizienten des Faktors und der Fehlervariablen frei geschätzt werden. Die Metrik eines jeden Faktors muss zur Identifizierbarkeit festgelegt werden.

Es gibt einige wesentliche Identifikationsregeln, die die Identifikation sicherstellen sollen, und für unseren Fall, in dem nur die Kovarianzstruktur interessiert, wie folgt wiedergegeben werden können:

1. notwendige Bedingung ist, dass die Anzahl der vorliegen Varianzen und Kovarianzen mindestens so groß sein muss wie die Zahl der zu bestimmenden Parameter. Wenn P die Anzahl der beobachteten Variablen wiedergibt, so gilt:

Anzahl der zu schätzenden Parameter $\leq \frac{p \cdot (p+1)}{2}$.

2. In der hier vorliegenden Untersuchung sind nicht nur die Pfadkoeffizienten des Faktors, sondern auch die der Fehlervariablen von Interesse. Um alle Pfadkoeffizienten frei schätzen zu können, wurde die Möglichkeit gewählt, die Varianz des Faktors und die Varianz der Fehlervariablen auf den Wert 1 festzulegen[15].

Anwendung der Regeln auf die hier gebildeten Modelle: Da in den einfachsten hier behandelten Modellen, den *Modellen mit einem Faktor und drei τ- kongenerischen Variablen*, die Varianzen der latenten Variablen so normiert werden, ist die Zahl der Varianzen und Kovarianzen, 3 + 3 = 6 und die Anzahl der zu bestimmenden Pfadkoeffizienten (drei Pfadkoeffizienten des gemeinsamen Faktors und drei Pfadkoeffizienten der Fehlervariablen) gerade gleich und die Modelle sind identifiziert (siehe 3.5).

3. Wenn jeweils drei manifeste Variable auf einen Faktor laden, sind Modelle unter Wahrung folgender Voraussetzungen identifiziert, wenn

---

[14] a. a. O., S. 863: Da die beobachteten Werte als Abweichungen vom Erwartungswert, oder anders ausgedrückt als Abweichungen vom Zentrum der Verteilung definiert werden, werden die Bildung von Abweichungsvariablen auch als Zentrierungs- und Abweichungsvariablen oder als zentrierte Variablen bezeichnet.

[15] a. a. O., S. 867: Die zweite Identifikationsregel gibt die Alternative wieder, entweder die Ladung einer manifesten Variablen zu fixieren oder die Varianz des Faktors.

a) jede Ladung ungleich Null ist,

b) jede Variable nur auf einen einzigen Faktor lädt und

c) die Residualvariablen unkorreliert sind[16].

Unter Bedingung c) ist zu verstehen, dass Fehlerkorrelationen nicht geschätzt werden müssen. Wenn Fehlerkorrelationen ungleich null vorhanden sind, aber nicht geschätzt werden müssen, sondern festgelegt werden (Modellrestriktion), dann ist Bedingung c) dennoch erfüllt, da sich die Zahl der zu schätzenden Parameter durch Restriktionen von Fehlerkorrelationen nicht erhöht.

Anwendung der Regeln auf die hier gebildeten Modelle:

In allen AMOS-Modellen mit *zwei Faktoren und sechs kongenerischen Variablen* (drei manifeste Variable pro Faktor sind a) und b) erfüllt und ist c) teilweise erfüllt.

In diesen Modellen stehen zur Verfügung $\frac{6 \cdot 7}{2} = 22$ Varianzen und Kovarianzen. Geschätzt werden sechs Pfadkoeffizienten der beiden Faktoren, sechs Pfadkoeffizienten der den Indikatoren zugehörigen Fehlervariablen, der Pfadkoeffizient des Strukturmodells und der Pfadkoeffizient der Fehlervariablen des zweiten Faktors, in Summe 14 Parameter. Daher ergibt sich in allen Modellen außer einem (fehlspezifiziertes Modell 7) eine Überidentifikation (acht Freiheitsgrade) und die hinreichende Identifikationsbedingung ist erfüllt.

Es ist mit einer Ausnahme für alle Modelle auch c) erfüllt, da höchstens zwei zusätzliche Fehlerkorrelationen angenommen werden, die in Form von Restriktionen in das Modell implementiert sind. Einzig im Modell 7 der fehlspezifizierten Variante wird die Korrelation zwischen den Fehlern von Indikator 1 und 2 geschätzt so dass sich insgesamt zu schätzende 15 Parameter ergeben (sieben Freiheitsgrade, siehe 4.2.4). Wegen der Überidentifikation ist trotz der Abweichung von c) auch die hinreichende Identifikationsbedingung erfüllt.

In einem zweiten Abschnitt dieser Arbeit werden mit PROC CALIS Simulationen zu den hier diskutierten AMOS-Modellen mit ein (siehe 5.3) oder zwei Faktoren (siehe 7.2 und 7.3) durchgeführt. Insofern Modellgleichheit besteht, treffen sowohl die notwendigen als auch die hinreichenden Identifikationsbedingungen zu. Die Modelle in 7.4 stellen hinsichtlich der Identifikationsbedingungen keine zusätzlichen Anforderungen. Nur in 7.5 weichen die Modelle von den bisherigen Modellen in einer die Identifizierbarkeit betreffenden Weise insofern

[16] a. a. O., vierte Identifikationsregel

ab, als bis zu zwei Fehlerkorrelationen geschätzt werden. Es sind jedoch aus 22 beobachteten Parametern höchstens 16 Parameter zu schätzen. Nur in einem Fall sind die groben Identifikationsregeln 1 und 4 verletzt (siehe im Vergleich 7.5.2 und 7.6.5).

2.8 Individuelle Quantifizierung des Erkrankungsschweregrades von Patienten durch Indexbildung

Bei der Entwicklung von Modellen versucht man nicht etwa, möglichst viele Diagnoseverfahren zu integrieren in der Absicht, jeden Aspekt abzubilden, durch den die Erkrankung in Erscheinung tritt. Man versucht, analog zum Prinzip der sparsamen Erklärung (Parsimonität), mit verhältnismäßig wenigen wichtigen Diagnoseverfahren den Schweregrad einer Erkrankung sehr gut anzunähern. Dies geschieht unter Berücksichtigung von Kosten und Nutzen, wobei mit ersteren nicht nur der wirtschaftliche, sondern auch der sonstige Aufwand gemeint ist, den Patient und Arzt bei einer Reihe von diagnostischen Tests hinzunehmen haben.

Nachdem also ein bestimmtes Modell als bestes ausgewählt wurde und eine Bewertung der Diagnoseverfahren erfolgte, können auf Basis dieses Modells Indizes für die Patienten berechnet werden. In einen individuellen Index gehen die individuellen Untersuchungsergebnisse ein, die der Patient in den verschiedenen Diagnoseverfahren erzielte. Die Gewichtung, mit der diese Testergebnisse einfließen, wird auf Grundlage der Pfadkoeffizienten ermittelt. Diese Quantifizierung des individuellen Schweregrad ist die *beste Annäherung an einen nicht vorhandenen Goldstandard* unter Berücksichtigung der ausgewählten Diagnoseverfahren.

2.9 Güte der Anpassung des Modells an die empirischen Daten

Die globale Anpassungsgüte des Modells (overall goodness of fit) wird auf Basis eines speziellen Chi-Quadrat-Tests ermittelt. Die zu testende Nullhypothese lautet, dass die empirischen Korrelationsmatrix und die auf Basis des Modells geschätzte Korrelationsmatrix identisch seien. Ist die Diskrepanz zwischen der empirischen und der geschätzten Korrelationsmatrix groß, wird die Nullhypothese abgelehnt und das Modell verworfen – andernfalls gilt das Modell als mit den empirischen Daten vereinbar. Dieser Test ist jedoch nicht hinreichend zur Beurteilung eines Modells. Es kann nämlich sein, dass eine Teilmenge der Messverfahren des Modells eine vorzügliche Anpassung aufweist und gleichzeitig eine Fehlspezifizierung des Modells zu einer schlechten Anpassung weiterer Messverfahren führt. Eine solche mangelhafte Anpassung nur einer Modellkomponente bleibt dem globalen Anpassungstest verborgen.

Um diesem Problem abzuhelfen, können Modifikationsindizes (Modifikation indices) benutzt werden. Diese Indizes sind Schätzungen einer Verbesserung der Anpassungsgüte für den Fall, dass neue Parameter ins Modell aufgenommen oder Modellrestriktionen fallengelassen werden. Die Aufnahme neuer Parameter kann z. B. erforderlich werden, wenn ein Störfaktor berücksichtigt werden muss. Eine Modellbeschränkung wird z. B. aufgehoben, wenn zwei Korrelationen zwischen Fehlertermen nicht mehr als gleich angenommen werden. Ist die geschätzte Verbesserung der Anpassungsgüte groß, wird die Modellanpassung vorgenommen und die Anpassungsgüte erneut getestet.

Da das Chi-Quadrat-Maß sehr empfindlich hinsichtlich des Stichprobenumfangs ist (Modelle mit einer großen Anzahl von Fällen haben oft keine gute Anpassung, obwohl nur triviale Differenzen zwischen $\hat{\Sigma}$ und **S** bestehen), werden üblicherweise *fit indices* zur Modellevaluation benutzt. Sie sind ein brauchbareres Instrument, da sie die Fähigkeit eines Modells wiedergeben, Kovarianzen zu erklären. Werte über 0,9 sind erwünscht, da sie anzeigen, dass über 90 % der Kovariation in den Daten durch das hypothetische Modell reproduziert werden können.

2.10 KLASSISCHE TESTTHEORIE UND MESSMODELLE

Im Prozess der Modellentwicklung kann erst im Anschluss an die Bewertung der Modellgüte eine weitere Bewertung von Modellen erfolgen. Hierzu sind zwei grundlegende Konzepte, Reliabilität und Validität, etabliert. Eine Darlegung dieser Konzepte erfordert die Darstellung der wichtigsten Aussagen der klassischen Testtheorie (CTT), auch Messfehlertheorie genannt.

Die Grundgleichung der CTT lautet $Y_i = \tau_i + \varepsilon_i$. Unter Y_i ist – über alle n Personen hinweg – eine Variable zu verstehen, die empirisch beobachtbar ist und durch die True-Score-Variable τ_i und die Error-Score-Variable (Messfehlervariable) ε_i beinflusst wird.

Aus obiger Gleichung ergeben sich vier wichtige Eigenschaften der Error-Score- und True-Score-Variablen[17]:

(1) Der Erwartungwert einer Messfehlervariablen ist für jede Ausprägung der True-Score-Variablen (bedingter Erwartungswert) gleich 0: $E(\varepsilon_i|\tau_j)$=0.

(2) Der unbedingte Erwartungswert einer Messfehlervariablen ist gleich 0: $E(\varepsilon_i)$=0.

(3) Messfehler- und True-Score-Variable sind unkorreliert: $Cov\ (\varepsilon_i, \tau_j)$=0.

[17] Eid et al. (2011), S. 818. Weitere Formulierungen im Zusammenhang mit der Darlegung von Definitionen oder Formeln der CTT sind ohne weitere Kennzeichnung daraus entnommen.

(4) Die Varianz einer beobachteten Messwertvariablen ist die Summe aus Varianz der True-Score-Variablen und der Varianz der Messfehlervariablen: $\mathrm{Var}(Y_i) = \mathrm{Var}(\tau_i) + \mathrm{Var}(\varepsilon_i)$.

Die Eigenschaft, dass die Messfehler der Variablen unkorreliert seien, kann nicht axiomatisch aus diesen abgeleitet werden. Sie ist nur dann gegeben, wenn „experimentelle Unabhängigkeit" angenommen wird, d. h. wenn die Messung jeder einzelnen beobachtbaren Variablen des Messinstruments in unabhängigen Experimenten erfolgt. Die Annahme der experimentellen Unabhängigkeit gilt als erfüllt, wenn die lineare Beziehungen messende Korrelation nach Pearson Null ist. Dies schließt das Vorhandensein nichtlinearer Assoziationen größer Null nicht aus.[18]

Die Reliabilität einer Variablen ist nun definiert als der Anteil der True-Score-Varianz an der Varianz einer beobachteten Variablen: $Rel(Y_i) = \frac{Var(\tau_i)}{Var(Y_i)}$.

Zimmerman betonte 1993, dass die Reliabilität auch betrachtet werden kann als die Korrelation zwischen geeignet definierten parallelen Messungen. Da es in den seltensten Fällen möglich ist, unabhängige Wiederholungen einer Messung zu erhalten, ist der Versuch der Forschung motiviert, Reliabilitätsschätzer auf Basis einer einzigen Testdurchführung zu erhalten.[19]

Die Items eines Messinstruments heißen:

(1) $\tau-$ parallel, wenn jeweils zwei Items i und j (i, j = 1, 2, ... k) eines Messinstruments identische True-Scores, $\tau_i = \tau_j$, und gleiche Fehlervarianzen aufweisen $\mathrm{Var}(\varepsilon_i) = \mathrm{Var}(\varepsilon_j)$.

(2a) Von $\tau-$ äquivalenten Items spricht man, wenn jeweils zwei Items

identische True-Scores, $\tau_i = \tau_j$, und ungleiche Fehlervarianzen besitzen $\mathrm{Var}(\varepsilon_i) \neq \mathrm{Var}(\varepsilon_j)$.

(2b) Essentiell $\tau-$ äquivalente Items (Spezialfall) sind dann gegeben, wenn sich die

True-Scores nur durch eine Konstante ungleich Null unterscheiden $\tau_i = \tau_j + a$.

$\tau-$ Parallelität und $\tau-$ Äquivalenz sind Spezialfälle der essentiellen $\tau-$ Äquivalenz.

(3) Kongenerische Items liegen vor, wenn True-Scores voneinander linear abhängig sind, $\tau_i = a + b \cdot \tau_j$ (a beliebig; b ungleich Null und ungleich 1), so dass $\mathrm{Var}(\tau_i) \neq \mathrm{Var}(\tau_j)$ und Fehlervarianzen verschieden sind: $\mathrm{Var}(\varepsilon_i) \neq \mathrm{Var}(\varepsilon_j)$. (Im Fall $\tau-$ paralleler oder $\tau-$ äqui-

[18] Komaroff (1997), S. 339 und Zimmerman (1993), S. 46.
[19] Kuder und Richardson, 1937, zitiert nach Zimmerman (1993), S. 45.

valenten Items ist $\tau_i = a + b \cdot \tau_j$ wobei a = 0 und b = 1, im Falle essentiell τ − äquivalenter Items gilt selbiges, wobei a ungleich null und b = 1.)

Auf Basis der Varianzen der True-Score- und der Messfehlervariablen lassen sich die Messmodelle in Kurzform wie folgt voneinander abgrenzen:

(1) Wenn Var(τ_i) = Var(τ_j) und Var(ε_i) = Var(ε_j), dann τ −Parallelität.

(2) Wenn Var(τ_i) = Var(τ_j) und Var(ε_i) ≠ Var(ε_j), dann essentielle τ − Äquivalenz oder τ − Äquivalenz.

(3) Wenn Var(τ_i) ≠ Var(τ_j) und Var(ε_i) ≠ Var(ε_j), dann Kongenerizität.

Auf Basis des Messmodells essentieller τ − Äquivalenz wird die Bedeutung einer latenten Variablen[20] eingeführt: Für ein aus drei Items oder Skalen bestehendes Messinstrument ergeben sich aus der Formel in 2b folgende Gleichungen:

$$\tau_1 = \tau_2 + \alpha_{12}$$

$$\tau_1 = \tau_3 + \alpha_{13}$$

$$\tau_2 = \tau_3 + \alpha_{23}$$

Die drei True-Score-Variablen sind perfekt korreliert und messen daher dasselbe. Man kann das Modell vereinfachen durch Wahl eines Vergleichsstandards, dessen Erwartungswert gleich Null oder einem anderen beliebigen Wert ist und als η (Eta) bezeichnet wird. Das Modell essentieller τ − äquivalenter Variabler kann unformuliert werden, so dass bei Bezug auf einen festgelegten Vergleichsstandard ein Index für alpha entfällt: $\tau_i = \eta + \alpha_i$. Für die drei Items lautet das Modell

$$\tau_1 = \eta + \alpha_1$$

$$\tau_2 = \eta + \alpha_2$$

$$\tau_3 = \eta + \alpha_3$$

Definition: Da es sich bei den True-Score-Variablen nur um Ableitungen derselben Variablen Eta handelt, kann man Eta auch als gemeinsame latente Variable bezeichnen. Wenn die Annahme der Unkorreliertheit der Messfehlervariablen zutrifft, dann ist die Kovarianz der beobachteten Variablen gleich der Varianz der gemeinsamen latenten Variablen: $Cov\left(Y_i, Y_j\right) =$

[20] Eine allgemeine Einführung des Begriffs der latenten Variablen erfolgte bereits in 2.2.

$Var(\eta); i \neq j$. Die Reliabilität einer beobachteten Y-Variablen (Item, Subskala) lässt sich daher folgendermaßen bestimmen:

$$Rel(Y_i) = \frac{Var(\tau_i)}{Var(Y_i)} = \frac{Var(\eta)}{Var(Y_i)} = \frac{Cov((Y_i, Y_j)}{Var(Y_i)}$$

Die Reliabilität einer beobachteten Variablen ist der Anteil der True-Score-Varianz an der Varianz der beobachteten Variablen und lässt sich bestimmen, indem man auf die Varianzen und Kovarianzen der beobachteten Variablen zurückgreift. Die Varianz der True-Score-Variablen ist gleich der Varianz der gemeinsamen latenten Variablen eta.

Die Unreliabilität einer beobachteten Y-Variablen ist definiert als $URel(Y_i) = \frac{Var(\varepsilon_i)}{Var(Y_i)} = \frac{Var(\varepsilon_i)}{Var(\tau_i) + Var(\varepsilon_i)} = 1 - Rel(Y_i)$.

Cronbachs alpha dient zur Berechnung der (Gesamt-)Reliabilität eines Messinstruments (einer Skala) und ist definiert als

$\alpha = \frac{p}{p-1} \cdot (1 - \frac{\sum_{i=1}^{p} Var(Y_i)}{Var(S)})$. Hierbei bezeichnet Y_i eine empirische beobachtbare Variable (Item eines Messinstruments), und S eine Summenvariable, wobei $S = \sum_{i=1}^{p} Y_i$. In die Summenvariable gehen neben den Varianzen der Subskalen die Kovarianzen der Subskalen ein.

Für das Verhältnis von Cronbachs alpha als Reliabilitätsschätzer und der „wahren Reliabilität" gilt folgende Fallunterscheidung: Unter den Voraussetzungen, dass (1) das Modell essentiell τ -äquivalenter Variabler gilt oder eines der Spezialfälle dieses Modelles vorliegt und dass (2) die Fehlerterme unkorreliert sind, ist Cronbachs alpha gleich der Reliabilität. Wenn die True-Score-Variablen eines Messinstruments voneinander linear abhängig sind, wie im Messmodell kongenerischer Items[21] der Fall, dann ist unter der Voraussetzung unkorrelierter Fehlerterme Cronbachs alpha eine Untergrenze der Reliabilität. In empirischen Testdaten ist die Annahme der Parallelität, aber auch nur die Annahme der essentiellen Äquivalenz wirklichkeitsfremd[22]. Ebenso ist in wirklichen Testsituationen experimentelle Unabhängigkeit

[21] In 2.3 ist ein Spezialfall des kongenerischen Messmodells, $\tau_i = \alpha_i + \boldsymbol{\lambda_i} \cdot \eta$, gegeben. In diesem wird ein Leichtigkeits- bzw. Schwierigkeitsparameter α_1 gleich Null angenommen, so dass sich das Modell auf $\tau_1 = \boldsymbol{\lambda_1} \cdot \eta$ verkürzt und das Testmodell $Y_1 = \alpha_1 + \lambda_1 \eta + \varepsilon_1$ auf $Y_1 = \lambda_1 \eta + \varepsilon_1$. (In 2.3 wird der Fehlerterm statt mit Epsilon äquivalent mit Delta bezeichnet.) Man kann den Wegfall des Leichtigkeitsparameters auch vor dem Hintergrund der zentrierten Variablen (siehe 2.7) interpretieren. Alle später erörterten, an die empirischen Glaukomdaten angepassten Messmodelle sind kongenerisch.

[22] Sijtsma (2009) zitiert nach Simsek et al (2013), S. 2009.

nicht gewährleistet, so dass auch die Bedingung der Unkorreliertheit der Fehler als unrealistisch und somit die Korreliertheit der Messfehler als Standardfall erscheint[23].

[23] Zimmerman (1993), S. 34. Er verweist auf Rozeboom (1966) und Zimmerman und Williams (1980) und zitiert auf S. 36 zur Begründung Rozeboom: „When the individual items on a test are administered at one sitting – it is inevitable or at least we have no reason to think otherwise – that a substantial proportion of the uncontrolled extraneous conditions which are presumable responsible for measurement error will persist from item to the next, thus inducing positive error correlation which may well be substantial."

3 AMOS-Messmodelle mit einem Faktor und drei Messverfahren auf Basis künstlich erzeugter Daten in AMOS

3.1 ZUSAMMENFASSUNG

Hauptgegenstand dieses Kapitels sind AMOS-Messmodelle auf der Basis von künstlich erzeugten Daten, in welchen Fehlergrößen der beobachtbaren Variablen tatsächlich korreliert sind. Es wird systematisch untersucht, wie sich hierbei die Modellannahme, es bestehe eine Fehlerkorrelation von Null (Fehlspezifikation), auf die Schätzung interessierender Pfadkoeffizienten und die Güte des Messmodells auswirkt.

3.2 EINFÜHRUNG: ANWENDUNGSBEREICH, PROBLEMSTELLUNG UND LÖSUNGSANSATZ

3.2.1 ANWENDUNGSBEREICH

In unserem Anwendungsbereich, der Glaukomforschung, liegen empirische Daten aus einem Patientenregister vor. Patienten dieses Registers durchliefen eine Reihe unterschiedlicher Verfahren, die der Diagnose des grünen Stars dienen sollen. Man befindet sich in einer Situation, in der kein fehlerfreier Goldstandard zur Messung des Schweregrades des Glaukoms vorliegt, sondern nur wenige Diagnoseverfahren als etabliert gelten. Diese Diagnoseverfahren weisen untereinander nur schwache bis mittlere Korrelationen auf. (In einer Stichprobe von 237 Patienten dieses Registers, die acht Diagnoseverfahren vollständig durchliefen, ergaben sich Korrelationskoeffizienten zwischen 0,20 und 0,82.)[24] In dieser Situation bieten Messmodelle die Möglichkeit, eine nicht messbare Variable „globaler Glaukomschweregrad“ zu modellieren, die die Ergebnisse einer Reihe von Diagnoseverfahren beeinflusst. Die Varianz eines jeden dieser Messverfahren wird durch den globalen Glaukomschweregrad in unterschiedlich hohem Maß statistisch erklärt. Auf Basis eines solchen Messmodells, das in der Grundform eine latente Variable „globaler Glaukomschweregrad“ und n Diagnoseverfahren enthält, kann ein Index berechnet werden, der für jeden Patienten einen individuellen allgemeinen Erkrankungsschweregrad angibt und gegenwärtig die beste Annäherung an einen allgemeinen Goldstandard sein dürfte.

[24] Martus (2001a), S. 267, Table 1.

Ein Messmodell beinhaltet i. A. neben den manifesten Variablen (Diagnoseverfahren) und der latenten exogenen[25] Variablen (globaler Glaukomschweregrad) die *Residualvariablen* der manifesten Variablen (siehe 2.3). In einem sehr einfachen (und stark restringierten) Modell wird angenommen, dass die Korrelationen zwischen diesen Residualvariablen gleich Null sind. In einem weniger restriktiven Modell geht man davon aus, dass zwischen zwei oder mehreren dieser Residual- oder Fehlervariablen Korrelationen ungleich Null bestehen.

3.2.2 PROBLEMSTELLUNG

In der Entwicklung von Messmodellen für unser Anwendungsgebiet bemüht man sich um die Etablierung von Modellen, die möglichst einfach sind, aber hinreichend komplex die empirische Realität abbilden (Prinzip der Parsimonität). Die Anzahl der zu schätzenden Parameter in einem Messmodell darf keinesfalls die Anzahl der beobachtbaren Parameter übersteigen. In dem Bestreben, die Anzahl der zu schätzenden Parameter möglichst klein zu halten und dabei eine ausreichend gute Anpassung an die empirischen Daten (Fit) zu erzielen, versucht man, Modelle zu vereinfachen. Eine dieser möglichen Vereinfachungen ist die Annahme unkorrelierter Fehlervariabler von Messverfahren.

Solche Annahmen können, wenn sie nicht zutreffen, u. a. zu verzerrten Schätzungen von Parametern führen, die im Fokus der Analyse stehen (z. B. die Stärke des Zusammenhangs zwischen dem globalen Glaukomschweregrad und den jeweiligen Diagnoseverfahren). Die tatsächliche Höhe des Koeffizienten eines Pfades zwischen einem Faktor und einem Messverfahren kann dadurch unter- oder überschätzt werden. Weitere Konsequenzen solcher nicht zutreffender Modellannahmen werden in Abschnitt 10 ausführlich erörtert.

Vollständige Strukturgleichungsmodelle, die mindestens zwei durch ein Strukturmodell verknüpfte Messmodelle enthalten, können mehr als eine falsche Modellannahme hinsichtlich der Fehlerkorrelation der manifesten Variablen enthalten. Hierbei kann der Fall vorliegen, dass die verzerrende Wirkung einer Modellannahme durch die der anderen *verstärkt* wird. Ebenso ist es denkbar, dass die Verzerrungen infolge der beiden Modellannahmen einander *abschwächen*. Der erste Fall ist insofern weniger problematisch, als fehlerhafte Modellannahmen nach angestellten Plausibilitätserwägungen entdeckt werden können: Weicht die Schätzung eines interessierenden Koeffizient wesentlich von der Erwartung des Forschers über den geschätzten Wert ab, so ist ein Anlass gegeben, das Modell zu überprüfen. Im zwei-

[25] In den hier diskutierten Messmodellen ist der globale Glaukomschweregrad *exogene* Variable, da sie in diesen Modellen manifeste Variable beeinflusst und nicht ihrerseits beeinflusst wird.

ten Fall hingegen kann sich der Modellentwickler auf Basis von Plausibilitätsüberlegungen fälschlicherweise in den Modellannahmen bestätigt sehen. (Im Extremfall heben die verzerrenden Wirkungen zweier verschiedener Modellannahmen auf einen interessierenden Koeffizienten einander auf. Ein Beispiel hierfür wird in der vorliegenden Untersuchung präsentiert.)

Ein Hauptaugenmerk dieser Arbeit besteht darin, die Wirkung von vereinfachenden Modellannahmen hinsichtlich der Fehlerkorrelationsstruktur auf die Schätzung von hauptsächlich interessierenden Parametern zu untersuchen. Eine solche Vereinfachung kann den wahren Sachverhalt korrekt beschreiben und wird dann eine korrekte Spezifikation genannt. Sie kann aber auch den wahren Sachverhalt verkennen und heißt dann Fehlspezifikation.

3.2.3 LÖSUNGSANSATZ: VERWENDUNG NICHTEMPIRISCHER DATEN

Auf Basis empirischer, aus einer Stichprobe gewonnener Daten wurde in Vorarbeiten zunächst versucht, die Wirkung von Fehlspezifikationen auf bestimmte interessierende Parameter zu analysieren.[26] Da man bei empirischen, stichprobenbasierten Daten häufig nur unter Unsicherheit entscheiden kann, ob eine vereinfachende Annahme zurecht oder fälschlicherweise getroffen werden kann, ist auch nicht mit Sicherheit zu klären ob, es sich bei dem Effekt einer solchen Modellannahme auf einen interessierenden Parameter um einen Bias handelt oder nicht. In dieser Situation bietet die *Erzeugung künstlicher Daten* einen Ausweg aus dem eben skizzierten Dilemma.

Wenn man in einem ersten Schritt Daten zu dem Zweck künstlich erzeugt, dass in einem zweiten Schritt ein Messmodell an diese Daten angepasst werden kann, so kann man Modelle erstellen, die einen perfekten Fit aufweisen (das Chi-Quadrat-Diskrepanzmaß zwischen der empirischen Korrelationsmatrix und der modellimmanenten Korrelationsmatrix ist gleich Null). Verändert man anschließend in diesem korrekt spezifizierten Modell genau eine der Modellannahmen, so dass es fehlspezifiziert ist, kann man sicher sein, dass es sich bei der Wirkung dieser Fehlspezifikation auf den interessierenden Parameter erstens tatsächlich um einen Bias handelt und dass dieser zweitens vollständig auf die Fehlspezifikation zurückzuführen ist.

Bei der Analyse von empirischen Daten hingegen ist es praktisch ausgeschlossen, Modelle mit perfektem Fit zu entwickeln. Man kann sich weder sicher sein, ob es sich bei einer Spezi-

[26] Diese Daten stammten nicht aus dem Glaukompatientenregister. Es handelte sich um geeignet erscheinende Musterdatensätze aus dem hier verwendeten Softwarepaket.

fikation einer Korrelation um eine korrekte oder ein Fehlspezifikation handelt und man kann nicht mit Sicherheit einen Bias von einem Nichtbias unterscheiden.[27]

In der hier vorliegenden Analyse gehen wir bei der Erzeugung nichtempirischer, künstlicher Daten von einer Reihe von Variablen mit bestimmten gewünschten Parametern aus, die voneinander statistisch unabhängig sind. Aus diesen werden „manifeste" und „latente" Variable so erzeugt, dass deren Kovarianzmatrix die Eigenschaften besitzt, die bei dem im nächsten Schritt konzipierten Modell in AMOS gewünscht sind. Im AMOS-Modell werden die Varianz und die Kovarianz der im Modell befindlichen Variablen so spezifiziert, dass sie exakt den zuvor erzeugten Daten entsprechen, und eine Kovarianzmatrix auf Basis der Modellannahmen geschätzt. Das auf diese Weise perfekt angepasste Modell kann anschließend mit dem Zweck modifiziert (fehlspezifiziert) werden, die Wirkung der Modifikation auf interessierende Parameter zu untersuchen. Es interessieren hierbei die Größe und Richtung des Bias, der bei einem oder mehreren Parametern beobachtet werden kann. Bei den in diesem Kapitel vorliegenden Modellen mit einem Faktor und drei Indikatoren kann jedoch die Güte der Anpassung an die künstlich erzeugten Daten nicht anhand des Chi-Quadrat-Diskrepanzmaßes überprüft werden, da es sich um genau identifizierte Modelle handelt, in welchen die Zahl der zu schätzenden Parameter mit der Zahl der beobachteten Parameter übereinstimmt (Modell ist genau identifiziert). Es resultiert dann – sowohl bei den korrekt- als auch bei den fehlspezifizierten Modellvarianten – ein Chi-Quadrat-Wert von 0, der das triviale Ergebnis einer perfekten Anpassung signalisiert (ausführlichere Erläuterung siehe unten). Bei den Modellen mit zwei Faktoren (bzw. den Modellen mit einem Faktor und mehr als drei Indikatoren) hingegen ist der häufig verwendete globale Chi-Quadrat-Anpassungstest bedeutungsvoll, sofern es sich um überidentifizierte Modelle (Anzahl der zu schätzenden Parameter größer als Anzahl der beobachteten) handelt.

3.3 METHODE ZUR ERZEUGUNG KÜNSTLICHER DATEN

3.3.1 ERZEUGUNG VON ZUFALLSZAHLEN

Erster Schritt ist die Erzeugung von standardnormalverteilten Zufallsvariablen ($\mu = 0$; $\sigma^2 = 1$). Es handelt sich genau genommen nicht um eine echte, sondern um eine Pseudo-Zufallszahlenerzeugung. Beginnend mit einem bestimmten Startwert (eine ganze Zahl zwi-

[27] Man kann allerdings durch den Vergleich der Werte für einen bestimmten interessierenden Parameter zweier empirischer Modelle, welche sich nur durch eine einzige Restriktion unterscheiden, die Wirkung dieser Restriktion auf den Parameter genau quantifizieren.

schen 1 und 2.000.000.000) wird eine bestimmte Zahlenfolge aus einer gegebenen Datendatei ausgewählt. Wird im Wiederholungsfall derselbe Startwert benutzt, so wird diese Zahlenfolge reproduziert.[28] Ziel ist, dass die *Zufallsvariablen perfekt unabhängig voneinander* sind. Die Unabhängigkeit ist erforderlich, um bei der Konstruktion einer Korrelationsstruktur genau die gewünschte Korrelationsmatrix zu erhalten. Eine angemessene Methode zur Erzeugung unabhängiger Zufallsvariabler ist die Hauptkomponentenanalyse.

3.3.2 HAUPTKOMPONENTENANALYSE ZUR HERSTELLUNG UNABHÄNGIGER ZUFALLSZAHLEN

Es wird die wichtigste Methode der Faktorenextraktion, die Hauptkomponentenanalyse (PCA), verwendet, da die Faktoren einer PCA voneinander unabhängig sind.

Der übliche Zweck der Faktorenanalyse, die Dimensionsreduktion, wird bei der Anwendung dieser Methode nicht verfolgt. Es werden nämlich so viele Faktoren extrahiert, wie standardnormalverteilte Zufallsvariable erzeugt wurden. Das Ziel ist in diesem Fall die Orthogonalität der Faktoren, d. h., dass die standardnormalverteilten und statistisch abhängigen Zufallsvariablen in statistisch unabhängige Faktoren überführt werden. (Die der Faktorenanalyse unterzogenen zehn standardnormalverteilten Variablen, die im Rahmen dieser Untersuchung verwendet werden, weisen beispielsweise trotz der hohen Fallzahl von über 1.500 Korrelationen von bis zu 0,05 auf.)

3.3.3 KONSTRUKTION MANIFESTER UND ALS LATENT BETRACHTETER VARIABLER

Bei der Erzeugung eines Datensatzes in SPSS, dem im zweiten Schritt ein AMOS-Modell perfekt angepasst werden soll, werden ausschließlich die aus der Hauptkomponentenanalyse hervorgegangenen standardnormalverteilten und unabhängigen Zufallsvariablen als Bauelemente verwendet. Der Datensatz wird bereits im Hinblick auf das spätere AMOS-Modell geschaffen, und zwar so, dass jeder Variablen des AMOS-Modells, sowohl beobachtbaren als auch nicht beobachtbaren, genau eine in SPSS konstruierte Variable zugeordnet werden kann.

In AMOS-Modellen gilt, dass alle als manifest modellierten Variablen in einem zugrundeliegenden Datensatz vorhanden sein müssen. In gewöhnlichen empirischen Datensätzen sind

[28] Es ist auch möglich, nicht den Startwert „SEED“ zu verwenden, sondern den Startwert wiederum dem Zufall zu überlassen, indem man – um den Preis der mangelnden Reproduzierbarkeit – stattdessen „RANDOM“ setzt.

latente Variable selbstverständlich nicht enthalten – deren Existenz und deren Eigenschaften sollen ja erst mit Hilfe von AMOS untersucht werden.

Bei der Erzeugung künstlicher Daten jedoch wird genau umgekehrt verfahren: In SPSS werden Variable mit bestimmten Eigenschaften produziert, die als latente Variable fungieren. Aus diesen werden wiederum manifeste Variable konstruiert. Daher gilt für alle AMOS-Modelle dieser Untersuchung: Nicht nur die manifesten, sondern auch die latenten Variablen in AMOS-Modellen sind in den jeweils zugrundeliegenden SPSS-Datendateien enthalten.

3.3.4 SYSTEMATIK DER VARIABLENERZEUGUNG

Die systematische Variablenerzeugung soll anhand eines später noch genauer analysierten Modells, dem Ein-Faktor-Modell mit drei Indikatoren und zwei positiv korrelierten Fehlervariablen veranschaulicht werden:

Abbildung 8: Konstruktion der beobachtbaren Variablen

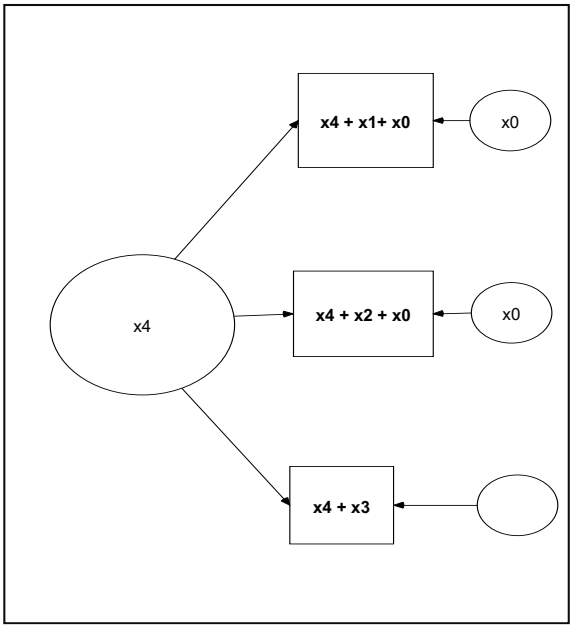
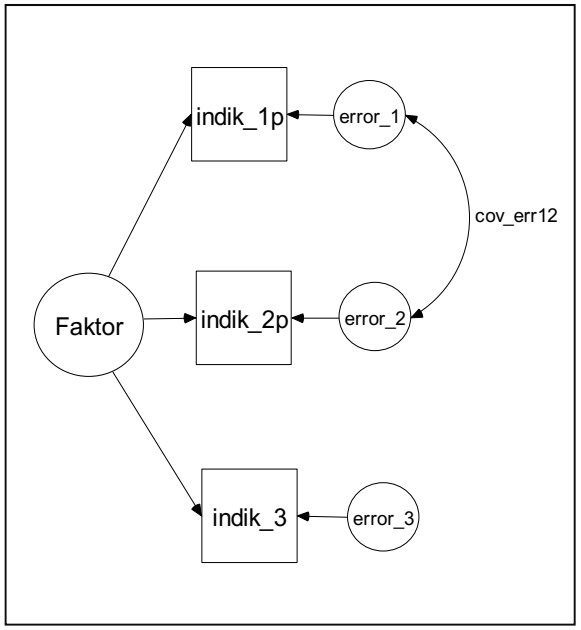


Abbildung 9: Der Konstruktion entsprechendes AMOS-Modell



In Abbildung 8, die die Ebene der Variablenkonstruktion in SPSS wiedergibt, sind die erzeugten *manifesten Variablen* als Rechtecke dargestellt. Sie bestehen jeweils aus zwei oder drei orthogonalen *Zufallsvariablen,* die als *latent* betrachtet werden und als Kreise dargestellt sind.

Die dritte manifeste Variable setzt sich aus der latenten Variablen x4 und dem Residualterm x3 zusammen. Dieser Fehlerterm besteht nur aus einer Zufallsvariablen, da er mit den Fehlertermen der anderen manifesten Variablen nicht korreliert ist.

Die bereits o. g. erste manifeste Variable besteht aus dem Faktor x4 und dem Residualterm x1 + x0. Letzterer setzt sich aus zwei Zufallsvariablen zusammen, da er mit dem Fehlerterm der zweiten manifesten Variablen korreliert ist.

In Abbildung 9, die die Ebene der Modellierung in AMOS wiedergibt, sind analog zu Abbildung 8 die auf SPSS-Ebene produzierten manifesten Variablen als Rechtecke dargestellt. Die Kreise symbolisieren hier die latenten Variablen, die aus ein oder zwei Zufallsvariablen konstruiert sein können. Die Berechnung der manifesten Variablen erfolgte gemäß den drei Gleichungen

```
COMPUTE indik_1p = x4 + x1 + x0.
COMPUTE indik_2p = x4 + x2 + x0.
COMPUTE indik_3   = x4 + x3.
```

Im Modell wird Folgendes angenommen:

1. Ein Faktor beeinflusst die drei beobachtbaren Variablen indik_1p, indik_2p und indik_3.

2. Die Fehlerterme der mit p indizierten Variablen sind (in einem noch zu spezifizierenden Maß) korreliert. Diese Korrelation wird durch einen Doppelpfeil zum Ausdruck gebracht.

3. Der Fehlerterm der verbleibenden Variablen ist von den anderen Fehlertermen unabhängig.

Ein Vergleich der beiden Diagramme macht deutlich, dass die drei oben genannten Modellannahmen des AMOS-Modells gerechtfertigt sind:

1. x4 ist gemeinsamer Bestandteil aller drei Indikatoren.

2. Die mit p indizierten Indikatoren besitzen einem Fehlerterm, der jeweils eine spezifische Komponente (x1 bzw. x2) und eine gemeinsam Komponente (x0) beinhaltet, die mit gleichem Gewicht in ihn eingehen.

3. Der Fehlerterm der verbleibenden Variablen besteht nur aus einer Zufallsvariablen, die in keiner anderen manifesten Variablen vorkommt.

Bei dem Modell in Abbildung 9 handelt es sich nur um eine schematische Darstellung. Es ist in der notierten Form nicht identifizierbar. Es fehlen darin die Spezifikation der Fehlerkorrelation und die Angabe der Varianz der latenten Variablen. Diese Details werden weiter unten erläutert.

3.3.5 PARAMETER DER KONSTRUIERTEN VARIABLEN

Jede der in SPSS erzeugten latenten Variablen hat eine Varianz von 1 und einen Mittelwert von 0, gleichgültig, ob es sich um eine orthogonale standardnormalverteilte Zufallsvariable oder um eine Variable handelt, die durch Addition oder Subtraktion solcher Zufallsvariabler hergestellt wurde. Diese Eigenschaft der *latenten Variablen* wurde durch eine *Standardisierung* der Variablen erreicht, die aus Operationen mit ursprünglichen Zufallsvariablen hervorgehen (siehe Beispiel unten).

Die Standardisierung berücksichtigt folgenden Umstand: Zwei Variable, die eine gleich große Korrelation mit einer dritten aufweisen aber unterschiedlich große Varianzen besitzen, führen im AMOS-Modell zu unterschiedlichen Schätzungen für die Gewichte der Pfade, die beide Variablen mit der dritten verbinden. Solch unterschiedliche Schätzungen für die beiden Pfadkoeffizienten sind unerwünscht. Schätzungen für Pfadkoeffizienten zwischen latenten und manifesten Variablen sollen ausschließlich auf die tatsächliche Korrelationsstruktur der künstlich erzeugten Daten und die im Modell getroffenen Annahmen über die Höhe der Fehlerkorrelationen zurückgeführt werden können. Wären die Varianzen der Fehlervariablen in obigem Modell verschieden, müssten die Schätzungen für die Pfadgewichte zwischen den spezifischen manifesten Variablen und den Fehlern bzw. dem Faktor auch auf deren Einfluss hin untersucht werden. Die Standardisierung der Varianz von Fehlervariablen erreicht man durch die Multiplikation des Fehlers mit einer geeigneten Konstante. Beispiel: Die manifeste Variable indik_1p des obigen Modells ist so konstruiert, dass sie mit der zweiten manifesten Variablen positiv korreliert. Ihre Fehlervariable y1 = x1 + x0 hat einen Mittelwert von 0 und eine Varianz von 2. Die Standardisierung geschieht durch Multiplikation mit $\frac{1}{\sqrt{2}}$, denn das gewünschte Resultat einer Varianz von 1 erfordert $\frac{1}{2} \cdot Var(x0 + x1) = Var[\frac{1}{\sqrt{2}}(x0 + x1)]$.[29] Durch die Standardisierung werden die Korrelationsstrukturen der erzeugten Variablen nicht verändert.

Es ist für die Zwecke der Modellkonstruktion nicht erforderlich, die *manifesten Variablen* zu standardisieren, da von ihnen als endogenen Variablen keine Wirkungen ausgehen, die durch die Varianz beeinflusst werden könnte.

[29] siehe 11.1.

3.4 ANALYSE KÜNSTLICHER DATEN MIT MESSMODELLEN IN AMOS – FEHLERKORRELATIONSSTRUKTUREN UND SPEZIFIKATIONSVARIANTEN

3.4.1 WARUM EIN MODELL MIT EINEM FAKTOR UND GENAU DREI INDIKATORVARIABLEN ?

In dieser Analyse wurde aus folgendem Grund ein Modell mit einem Faktor und genau drei Indikatorvariablen gewählt: Für die drei Indikatoren des Grundmodells, in welchem die Fehlerterme nicht korreliert sind, lassen sich die Pfadkoeffizienten $ß_i$ exakt aus den paarweisen Korrelationen der Indikatoren berechnen. Für Modelle mit korrelierten Fehlertermen gilt dies nicht. (Zum Symbol $ß_i$: siehe Abbildung 10, S. 53.) Ausgehend von der allgemeinen Beziehung $cor_{ind_i,ind_j} = \beta_i \cdot \beta_j$ bzw. von den spezifischen Gleichungen des Grundmodells

$$cor_{ind1,ind2} = \beta_1 \cdot \beta_2$$

$$cor_{ind1,ind3} = \beta_1 \cdot \beta_3$$

$$cor_{ind2,ind3} = \beta_2 \cdot \beta_3$$

gilt $\beta_i = \sqrt{\frac{cor_{ind_i,ind_j} \cdot cor_{ind_i,ind_k}}{cor_{ind_j,ind_k}}}$ für i, j, k = 1, 2, 3 und $i \neq j \neq k$.

unter der Voraussetzung, dass alle paarweisen Korrelationen größer Null sind.

Im Grundmodell errechnet sich z. B. der Koeffizient des Pfades zwischen Faktor und erstem Indikator als $\beta_1 = \sqrt{\frac{cor_{ind_1,ind_2} \cdot cor_{ind_1,ind_3}}{cor_{ind_2,ind_3}}} = \sqrt{\frac{0{,}5 \cdot 0{,}5}{0{,}5}} \sim 0{,}71$.

Aufgrund der perfekten Symmetrie des Modells sind alle paarweisen Korrelationen identisch und somit auch alle Pfadkoeffizienten.

Im *allgemeinen* Fall korrelierter Fehlerterme gilt die Beziehung:

$$cor(ind_1, ind_2) = cor(F, ind_1) \cdot cor(F, ind_2) + cor(error_1, error_2) \qquad \text{(abhängige Fehler),}$$

wobei cor ($error_1$, $error_2$) die Fehlerkorrelation der Indikatoren angibt. Im Spezialfall cor ($error_1$, $error_2$) = 0 reduziert sich die Beziehung auf

$$cor(ind_1, ind_2) = cor(F, ind_1) \cdot cor(F, ind_2) \qquad \text{(unabhängige Fehler).}$$

In Modellen mit mehr als drei Indikatorvariablen und nicht korrelierten Fehlertermen lassen sich die Pfadkoeffizienten nicht mehr exakt errechnen.

3.4.2 VERGLEICH DER IN SPSS KÜNSTLICH ERZEUGTEN MIT DER IN AMOS GESCHÄTZTEN KORRELATIONSMATRIX

Das AMOS-Modell ist perfekt den empirischen Daten angepasst, wenn die empirische Korrelationsmatrix und die auf Basis des Modells geschätzte Korrelationsmatrix genau übereinstimmen. In dieser Analyse künstlich erzeugter Daten ist nicht nur, wie bei empirischen Daten, die Korrelationsmatrix der beobachtbaren, sondern die der beobachtbaren und als nichtbeobachtbar angesehenen Variablen bekannt. Daher können die Schätzungen für bestimmte Korrelationen im AMOS-Modell zwischen latenten und latenten oder latenten und manifesten Variablen anhand der zugrundeliegenden Korrelationen der in SPSS erzeugten Variablen überprüft werden.

Ein Beispiel: Der Koeffizient des Pfades zwischen dem Faktor und der ersten manifesten Variablen indik_1 wird auf ~0,71 geschätzt. Zur Überprüfung errechnet man die Korrelation zwischen den entsprechenden Variablen x4 und der aus $x4 + \frac{1}{\sqrt{2}}(x1 + x0)$ bestehenden manifesten Variablen indik_1ps.

$$\text{Cor(x4, indi_1ps)} = \frac{\text{Cov(x4, indi_1ps)}}{\sqrt{Var(x4) \cdot Var(indi_1ps)}} = \frac{\text{E(x4 * indi_1ps) - E(x4) * E(indi_1ps)}}{\sqrt{Var(x4) \cdot Var(indi_1ps)}} =$$

$$= \frac{1 - 0 * 0}{\sqrt{1 \cdot 2}} \sim 0{,}71$$

.

Die Koeffizientenschätzung für den Pfad zwischen dem Faktor und der ersten Indikatorvariablen im AMOS-Modell und die Korrelation der zugrundeliegenden in SPSS erzeugten Variablen stimmen mit 0,71 also genau überein.

3.4.3 FEHLERKORRRELATIONSSTRUKTUREN UND SPEZIFIKATIONSVARIANTEN

Ausgangspunkt einer Serie von Modellen in dieser Analyse künstlich erzeugter Daten ist ein sogenanntes *Grundmodell*. Dabei handelt es sich um ein Modell, in dem nur manifeste Variable vorkommen, deren Fehlervariable als unkorreliert modelliert sind.

Das Grundmodell wird anschließend mehrfach variiert, so dass mehrere *Typen von Korrelationsbeziehungen* entstehen. Dabei wird die positive Korrelation zwischen zwei Fehlervariablen

(2), die negative Korrelation zwischen zwei Fehlervariablen (3) und die Situation dreier positiv korrelierter Fehlervariabler (4) bei der Erzeugung der Variablen in SPSS durchgespielt.

Das Grundmodell geht von dem einfachsten, unrealistischen Fall aus, dass jedes mit der manifesten Variablen erfasste Messverfahren eine eigene, von den anderen unterschiedene Fehlerquelle besitzt.

Die zweite Situation gibt den realistischen Fall wieder, in dem ein Teil der Messverfahren gemeinsamen Fehlerquellen unterliegt, die in der Person des untersuchten Probanden, der Messapparatur oder beidem gemeinsam begründet sein können .

Die dritte dürfte in der Realität kaum anzutreffen sein und ist der Vollständigkeit halber und aus theoretischen Gründen in die Typologie aufgenommen.

Die vierte Situation ist eine Variation der zweiten und geht davon aus, dass unkorrelierte Fehlerterme genauso wenig vorkommen wie negative.

Weitere Situationen als Kombinationen der genannten typischen vier sind darüber hinaus in Modellen mit mehr als drei Indikatoren denk- und in diesem Rahmen konstruierbar.

Mit jeder der vier beschriebenen, tatsächlichen Korrelationssituationen wird nun bei der Modellierung mit AMOS *auf mindestens zweierlei Weise* umgegangen. In der ersten Modellvariante, genannt *„korrekt spezifiziert“*, wird den tatsächlich vorhandenen Korrelationsbeziehungen exakt Rechnung getragen.

Beispiel: In dem oben vorgestellten Modell wird der Korrelationsparameter „cov_err12“ auf 0,5 gesetzt, den tatsächlichen Wert der Korrelation zwischen y1 und y2 (beziehungsweise standardisiert y1_s und y1_s). Hierbei gilt:

y1 = x1 + x0 und y2 = x2 + x0 sowie y1_s = $\frac{1}{\sqrt{2}}(x1+x0)$ und y2_s = $\frac{1}{\sqrt{2}}(x2+x0)$.

Im *„Modell fehlspezifiziert“* wird cov_err12 = 0 gesetzt (oder der Doppelpfeil wieder aus dem Modell entfernt).[30]

Der Effekt jeder Modellvariante wird anhand der *standardisierten Koeffizientenschätzungen (standardized estimates)* an den Pfaden zwischen dem Faktor und den manifesten Variablen beurteilt.

[30] In weiteren Varianten könnte der *Grad der Fehlspezifikation* variiert werden, indem cov_err12 auf einen Wert zwischen 0 und 0,5 festgelegt wird.

3.4.3.1 Überlegung zu den Wirkungen der Modellspezifikation – Anpassungsgüte des Modells und Koeffizientenberechnung

Die erste Modellvariante müsste erwartungsgemäß eine perfekte Anpassung an die Daten aufweisen. Für alle Pfadkoeffizienten müssten Werte geschätzt [31] werden, die, in ihrer standardisierten Form, mit den entsprechenden Werten der zugrunde liegenden, erzeugten Korrelationsmatrix übereinstimmen[32]. Es müssten weiterhin ein oder mehrere Pfadkoeffizienten von den entsprechenden Korrelationskoeffizienten der erzeugten Korrelationsmatrix abweichen.

Die Überprüfung der Anpassungsgüte des Modells erübrigt sich in beiden Modellvarianten, da die Anzahl der zu schätzenden Parameter gerade gleich der Anzahl der beobachtbaren Parameter ist (Modellvarianten sind genau identifiziert). Die aus den geschätzten sechs Pfadkoeffizienten rückgerechneten Korrelationen und Varianzen der Indikatoren entsprechen den künstlich erzeugten Korrelationen derselben natürlich exakt. Die Pfadkoeffizienten hingegen können in den jeweiligen Modellvarianten berechnet und die an den Pfadkoeffizienten auftretenden Effekte interpretiert werden.

[31] In dieser Analyse künstlich erzeugter Daten ist es streng genommen nicht korrekt, wie im gewöhnlichen Anwendungsfall – der Analyse empirischer Daten aus Stichproben – von Parameterschätzung zu sprechen. Vielmehr handelt es sich, da die Information der künstlich erzeugten Daten vollständig vorliegt, eher um eine Parameterberechnung. Dennoch wird – der Terminologie der Theorie der Strukturgleichungsmodelle und der entsprechenden Software AMOS folgend – weiterhin von Schätzung gesprochen.

[32] Siehe Beispiel in 3.4.2.

3.5 ANALYSE

3.5.1 DAS GRUNDMODELL

Ausgangspunkt der Modellreihe ist ein Modell, in welchem die Fehlerterme als unkorreliert angenommen werden. Die zugrunde liegenden Daten wurden entsprechend so erzeugt, dass die drei manifesten Variablen nur insofern korrelieren, als sie den Faktor (x4) beinhalten:

```
COMPUTE indik_1 = 1 * x4 + x1.
COMPUTE indik_2 = 1 * x4 + x2.
COMPUTE indik_3 = 1 * x4 + x3.
```

Die Kovarianzmatrix dieser drei manifesten Variablen ist in Tabelle 1 wiedergegeben.

Tabelle 1: Korrelationsmatrix der drei manifesten Variablen des Grundmodells

		INDIK_1	INDIK_2	INDIK_3
INDIK_1	Korrelation nach Pearson	1	,500(**)	,500(**)
	Signifikanz (2-seitig)	.	,000	,000
	N	1517	1517	1517
INDIK_2	Korrelation nach Pearson	,500(**)	1	,500(**)
	Signifikanz (2-seitig)	,000	.	,000
	N	1517	1517	1517
INDIK_3	Korrelation nach Pearson	,500(**)	,500(**)	1
	Signifikanz (2-seitig)	,000	,000	.
	N	1517	1517	1517

** Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Aus ihr gehen die sechs Stichprobenparameter Var(indik_1), Var(indik_1), Var(indik_1), r(indik_1, indik_2), r(indik_1, indik_3) und r(indik_2, indik_3) hervor, die die Schätzung einer gleich großen Anzahl von im Modell nicht spezifizierten Parametern erlauben.

Bei den zu schätzenden Parametern handelt sich um die sechs Koeffizienten der in der AMOS-Modellspezifikation (Abbildung 10)) mit ß$_1$, ß$_2$ und ß$_3$ sowie ß_f$_1$, ß_f$_2$ und ß_f$_3$ bezeichneten Pfade, die auch als Regressionsgewichte zu bezeichnen sind.

Sie sind so zu bestimmen, dass die folgenden Regressionsgleichungen erfüllt sind:

$$\text{indik_1} = \beta_1 \cdot Faktor + \beta_f_1 \cdot error_1$$

$$\text{indik_2} = \beta_2 \cdot Faktor + \beta_f_2 \cdot error_2$$

$$\text{indik_3} = \beta_3 \cdot Faktor + \beta_f_3 \cdot error_3$$

Die Modellspezifikation sieht weiterhin vor, dass die Varianzen aller latenten Variablen – entsprechend den als latent betrachten in SPPS erzeugten Variablen – gleich eins sind (siehe 3.3.5).[33]

Das Resultat dieser Parameterschätzung ist Abbildung 11 (standardized estimates) zu entnehmen.

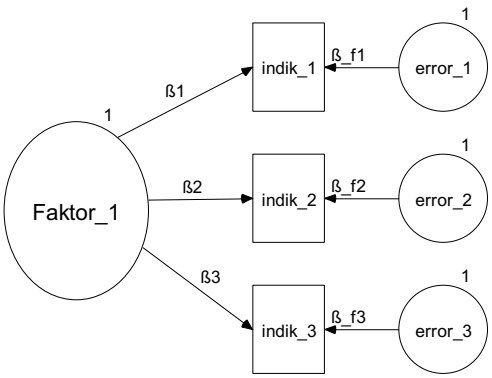
Der Vergleich der in AMOS geschätzten standardisierten Regressionsgewichte in folgender Tabelle 2 (oder in Abbildung 11) mit der *vollständigen* Korrelationsmatrix der in SPSS erzeugten Variablen in Tabelle 3 zeigt, dass die Schätzungen der Pfadkoeffizienten genau mit den betreffenden Korrelationskoeffizienten übereinstimmen.

Tabelle 2: In AMOS geschätzte standardisierte Regressionsgewichte des Grundmodells

Regression Weights:	Estimate	S.E.	C.R.	Standardized Regression Weights:	Label
indik_1 <----- Faktor_1	1,000	0,039	25,957	0,707	ß1
indik_1 <------ error_1	1,000	0,029	34,825	0,707	ß_f1
indik_2 <------ error_2	1,000	0,029	34,825	0,707	ß_f2
indik_3 <----- Faktor_1	1,000	0,039	25,957	0,707	ß3
indik_3 <------ error_3	1,000	0,029	34,825	0,707	ß_f3
indik_2 <----- Faktor_1	1,000	0,039	25,957	0,707	ß2

[33] Die Angabe, entweder der Varianz oder des Pfadgewichtes von latenten exogenen Variablen, ist zur Identifizierbarkeit erforderlich, da nicht beide Parameter zugleich geschätzt werden können. Alternativ könnte man auch Pfadgewichte von 1 bei allen Fehlervariablen spezifizieren, *einen* beliebigen der drei Pfade des Faktors mit 1 gewichten und die Varianz keiner der vier latenten Variablen spezifizieren. Die Gewichtung eines der drei Pfade des Faktors ist zur Normierung erforderlich.

Abbildung 10: Allgemeine Modellspezifikation des Grundmodells



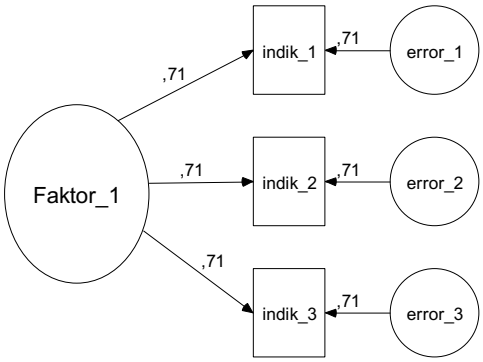
künstl. Erzeugung latenter standardisierter Variabler mittels orthogonaler Faktoren

Indikatoren mit unkorrelierten Fehlertermen

Model Specification

Abbildung 11: Schätzung des korrekt spezifizierten Grundmodells

Chi-square = ,000 (0 df)
p = \p



künstl. Erzeugung latenter standardisierter Variabler mittels orthogonaler Faktoren

Indikatoren mit unkorrelierten Fehlertermen

Standardized estimates

Tabelle 3: Korrelationsmatrix der in SPSS erzeugten Variablen des Grundmodells

		INDIK_1	INDIK_2	INDIK_3	X1	X2	X3	X4
INDIK_1	Korrelation nach Pearson	1	,500(**)	,500(**)	,707(**)	,000	,000	,707(**)
	Signifikanz (2-seitig)	.	,000	,000	,000	1,000	1,000	,000
	N	1517	1517	1517	1517	1517	1517	1517
INDIK_2	Korrelation nach Pearson	,500(**)	1	,500(**)	,000	,707(**)	,000	,707(**)
	Signifikanz (2-seitig)	,000	.	,000	1,000	,000	1,000	,000
	N	1517	1517	1517	1517	1517	1517	1517
INDIK_3	Korrelation nach Pearson	,500(**)	,500(**)	1	,000	,000	,707(**)	,707(**)
	Signifikanz (2-seitig)	,000	,000	.	1,000	1,000	,000	,000
	N	1517	1517	1517	1517	1517	1517	1517
X1	Korrelation nach Pearson	,707(**)	,000	,000	1	,000	,000	,000
	Signifikanz (2-seitig)	,000	1,000	1,000	.	1,000	1,000	1,000
	N	1517	1517	1517	1517	1517	1517	1517
X2	Korrelation nach Pearson	,000	,707(**)	,000	,000	1	,000	,000
	Signifikanz (2-seitig)	1,000	,000	1,000	1,000	.	1,000	1,000
	N	1517	1517	1517	1517	1517	1517	1517
X3	Korrelation nach Pearson	,000	,000	,707(**)	,000	,000	1	,000
	Signifikanz (2-seitig)	1,000	1,000	,000	1,000	1,000	.	1,000
	N	1517	1517	1517	1517	1517	1517	1517
X4	Korrelation nach Pearson	,707(**)	,707(**)	,707(**)	,000	,000	,000	1
	Signifikanz (2-seitig)	,000	,000	,000	1,000	1,000	1,000	.
	N	1517	1517	1517	1517	1517	1517	1517

** Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

Güte des Messmodells

Die Güte eines Messmodells wird daran abgelesen, ob die Summe der drei geschätzten Determinationskoeffizienten des Messmodells, z. B. $\hat{\beta}_1^{\,2} + \hat{\beta}_2^{\,2} + \hat{\beta}_3^{\,2}$ von der Summe der in SPSS erzeugten Determinationskoeffizienten, $\beta_1^{\,2} + \beta_2^{\,2} + \beta_3^{\,2}$, abweicht.[34] Die Summe der in AMOS geschätzten Determinationskoeffizienten beträgt $0{,}71^2 + 0{,}71^2 + 0{,}71^2 = 1{,}5$ und ist gleich der Summe der in SPSS festgelegten. Die Güte des Messmodells wird also korrekt geschätzt.

Modellfit

Für dieses Modell kann – so wie in allen folgenden Modellen – kein Maß für die Güte der Anpassung des AMOS-Modells an die künstlich erzeugten Daten geschätzt werden, da die Zahl der Freiheitsgrade gleich 0 ist.[35]

[34] Bei den geschätzten Regressionsgewichten in Tabelle 2 handelt es sich laut dieser Symbolik um die $\hat{\beta}_i$, nicht um die β_i. Der AMOS-Textoutput kennt diese Symbolik jedoch nicht.

[35] Vgl. Aarbuckle et al. (1999), S.75f.

3.5.2 DAS MODELL MIT ZWEI POSITIV KORRELIERTEN FEHLERVARIABLEN

Die Korrelationsstruktur dieses Modells wurde bereits in 3.4.3 ausführlich beschrieben (siehe Abbildung 8 und Abbildung 9). Die folgende Abbildung 12 zeigt die allgemeine[36] Modellspezifikation: Die Fehlerterme sind – abweichend von der schematischen Darstellung in Abbildung 9 – in diesem Modell standardisiert, was durch die Bezeichnungen Indi_...s angezeigt wird. Der Name eines Indikators zeigt den Korrelationsstatus seiner jeweiligen Fehlervariablen an: Enthält das Subskript eines Indikators den Buchstaben „p“, so handelt es sich um einen Indikator, der mit einem anderen *auch über die Fehlerterme* positiv korreliert ist. Die Korrelation zwischen error 1 und error 2 wird mit cov_err12 bezeichnet und in den beiden folgenden Modellvarianten unterschiedlich *spezifiziert.* Die Bezeichnungen für die Pfadgewichte lauten $ß_i$ bzw. $ß_f_i$ und werden in diesen Modellvarianten *nicht spezifiziert, sondern geschätzt*. Die Variablen wurden in SPSS wie folgt erzeugt:

```
COMPUTE indi_1ps = x4 + (x1 + x0) / SQRT(2).
COMPUTE indi_2ps = x4 + (x2 + x0) / SQRT(2).
COMPUTE indik_3 = 1 * x4 + x3.
```

In den folgenden Gleichungen werden die Koeffizienten geschätzt:

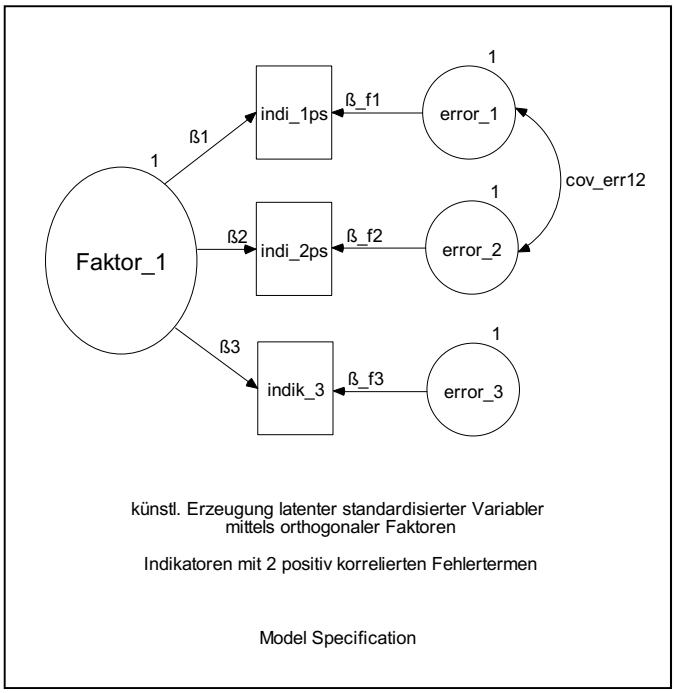
indi_1ps = $\beta_1 \cdot Faktor + \beta_f_1 \cdot error_1$

indi_2ps = $\beta_2 \cdot Faktor + \beta_f_2 \cdot error_2$

indik_3 = $\beta_3 \cdot Faktor + \beta_f_3 \cdot error_2$

[36] Die allgemeine Modellspezifikation (Most General Model) enthält Spezifikationen für eine Gruppe von noch näher zu spezifizierenden Modellvarianten. Sie enthält z. B. die Namen der zu schätzenden (und nicht zu schätzenden) Parameter. Sie enthält noch nicht die konkreten Modellrestriktionen (z. B. die Annahme eines bestimmten Wertes für die Fehlerkorrelation), die in den Modellvarianten (z. B. korrekt spezifiziertes Modell bzw. fehlspezifiziertes Modell) getroffen werden.

Abbildung 12: Modell mit zwei positiv korrelierten Fehlervariablen - allgemeine Modelspezifikation



Ergebnisse

Die Koeffizientenschätzungen des *korrekt spezifizierten* Pfadmodells, in dem *cov_err12 = 0,5* angenommen wird, sind Abbildung 13 zu entnehmen. Die Schätzungen sind mit jenen des Grundmodells mit der Ausnahme der negativen Vorzeichen der $ß_i$ identisch. Das negative Vorzeichen ist hier, wie in den folgenden Modellen, als ein numerisches Artefakt aufzufassen. Es kann, wie hier, eine ganze Gruppe von Pfadkoeffizienten betreffen, oder auch nur einzelne Pfadkoeffizienten eines Messmodells (siehe z. B. Abbildung 28), unter Umständen auch einen Pfadkoeffizienten des Strukturmodells (siehe unten). Dieses Phänomen beruht darauf, dass softwareintern aus korrelierten manifesten Variablen eine latente Variable z erzeugt wird, deren Ausprägungen auch negative Vorzeichen aufweisen können. Die Variable *z* ist hierbei äquivalent mit der Variablen -z. Das Phänomen softwareinterner Vorzeichenumkehr ist daher inhaltlich belanglos.

Die standardisierten Schätzungen des Pfadmodells, in dem die Fehlerkorrelation *fehlspezifiziert* wird (*cov_err12 = 0)*, werden in Abbildung 14 wiedergegeben.

Der mit ,00 beschriftete Doppelpfeil zwischen error 1 und error 2 bedeutet, dass auf den ersten Blick gravierende Abweichungen der Schätzungen für die Pfadgewichte $\hat{\beta}_1$, $\hat{\beta}_2$ und $\hat{\beta}_3$ vom für alle Pfade einheitlichen Gewicht 0,71 im Grundmodell erkennbar sind, wobei $ß_1$ und $ß_2$ über- und $ß_3$ unterschätzt werden. Umgekehrt verhält es sich bei den Gewichten der von den Fehlertermen ausgehenden Pfade β_f_1, β_f_2 und β_f_3 : Die ersten beiden werden unter- und das dritte überschätzt.

Abbildung 13: Korrekt spezifiziertes Modell mit zwei positiv korrelierten Fehlervariablen

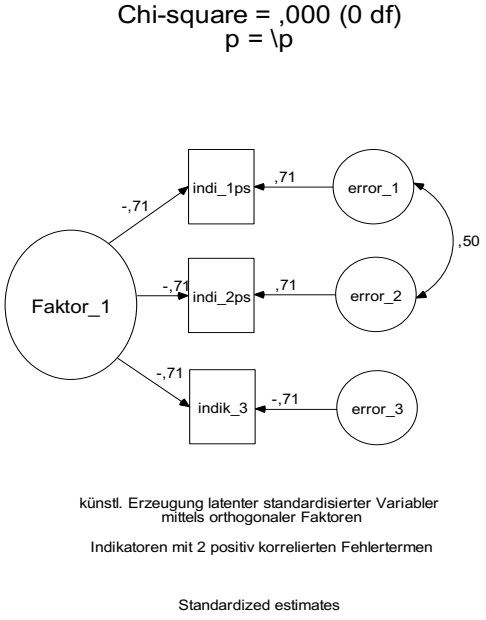
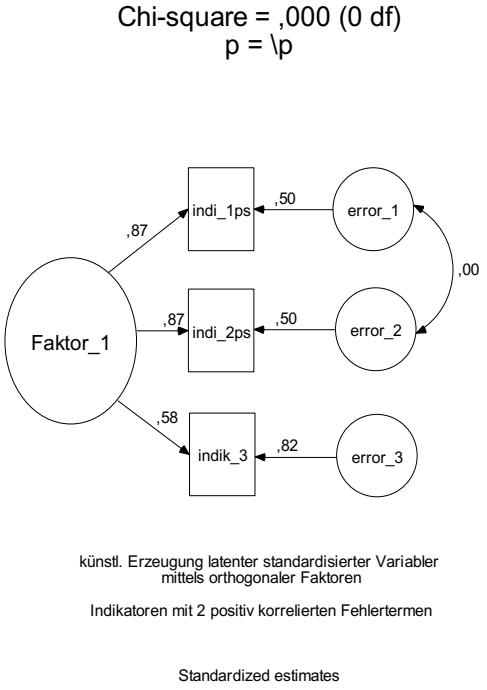


Abbildung 14: Fehlspezifiziertes Modell mit zwei positiv korrelierten Fehlervariablen



Die Bestimmung einer manifesten Variablen durch zwei latente Variable erfolgt nach der Gleichung Indikator_i = $ß_i$ * Faktor + $ß_{fi}$ * error_i unter der Restriktion $ß_i^2 + ß_{fi}^2 = 1$.

Im korrekt spezifizierten Modell beträgt der Determinationskoeffizient $\hat{\beta}_i^2$ aller Indikatoren jeweils 0,5, so dass die durch den Faktor erklärte und die unerklärte Varianz der Indikatoren gleich groß sind. Die Residuen der Indikatoren 1 und 2 setzen sich aus den spezifischen Komponenten x1 bzw. x2 und der gemeinsamen Komponente x0 zusammen, wobei die spezifische und die gemeinsame Komponente mit dem gleichen Gewicht eingehen. Da *infolge der Fehlspezifikation die Komponente x0 fälschlicherweise der Komponente x4 zugerechnet wird*, verringert sich die unerklärte Varianz der ersten beiden Indikatoren um die Hälfte (25 %) und

die *erklärte Varianz der ersten beiden Indikatoren nimmt* um genau diesen Betrag *zu*. $\hat{\beta}_f_1$ und $\hat{\beta}_f_2$ werden somit jeweils mit $\sqrt{0{,}25} = 0{,}5$ zu niedrig sowie $\hat{\beta}_1$ und $\hat{\beta}_2$ jeweils mit $\sqrt{0{,}75} = 0{,}87$ zu hoch geschätzt. Auch auf die Koeffizienten der mit dem dritten Indikator verbundenen Pfade wirkt sich die Erhöhung der Gewichte $\hat{\beta}_1$ und $\hat{\beta}_2$ aus. $\hat{\beta}_3$ wird im fehlspezifizierten Modell mit 0,58 geringer geschätzt als im korrekt spezifizierten Modell und entsprechend $\hat{\beta}_f_3$ mit 0,82 höher. Die die ersten beiden Indikatoren betreffende Fehlspezifikation wirkt sich somit indirekt auch in einer Reduktion der (durch den Faktor) erklärten Varianz des dritten Indikators von 50 auf 33 % aus.

Güte des Messmodells

Die Güte des Messmodells wird im Falle der Fehlspezifizierung mit $0{,}866^2 + 0{,}866^2 + 0{,}577^2 = 1{,}833$ überschätzt, da im korrekt spezifizierten Modell die Summe 1,5 resultiert.

3.5.3 DAS MODELL MIT ZWEI NEGATIV KORRELIERTEN FEHLERVARIABLEN

Die allgemeine Modellspezifikation dieses Modells entspricht mit folgender Ausnahme dem Vormodell und wird daher nicht graphisch wiedergegeben. Der zweite Indikator lautet in diesem Modell indi_2ns und dessen Fehlervariable ist mit jener des ersten Indikators negativ korreliert. Die negative Fehlertermkorrelation kommt zustande durch folgende Erzeugung des zweiten Indikators:

```
COMPUTE indi_2ns = x4 + (x2 - x0) / SQRT(2).
```

(Die anderen beiden Indikatoren werden wie entsprechend dem Vormodell erzeugt.)

Die Modellgleichungen sind bis auf die des zweiten Indikators ebenfalls identisch mit jenen des Modells mit zwei positiv korrelierten Fehlern und lauten:

indi_1ps = $\beta_1 \cdot Faktor + \beta_f_1 \cdot error_1$

indi_2ns = $\beta_2 \cdot Faktor + \beta_f_2 \cdot error_2$

indik_3 = $\beta_3 \cdot Faktor + \beta_f_3 \cdot error_3$

Ergebnis

Das korrekt spezifizierte Modell (Abbildung 15) enthält die Annahme cov_err12 = -0,5 und führt zu Schätzungen der Pfadkoeffizienten, die mit den beiden Vormodellen übereinstimmen. Das fehlspezifizierte Modell (Abbildung 16) beinhaltet die Restriktion cov_err12 = 0. Dessen Resultate sind in Anwendung derselben Logik erklärbar, die bereits im fehlspezifizierten Vormodell dargestellt wurde:

Abbildung 15: Korrekt spezifiziertes Modell mit zwei negativ korrelierten Fehlervariablen

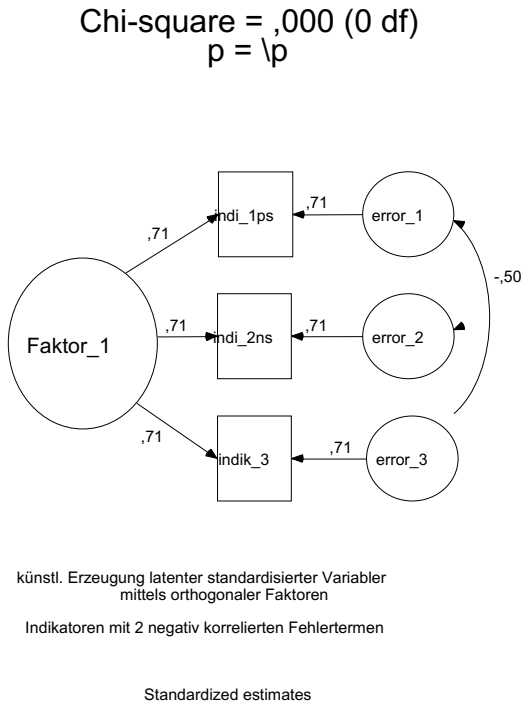
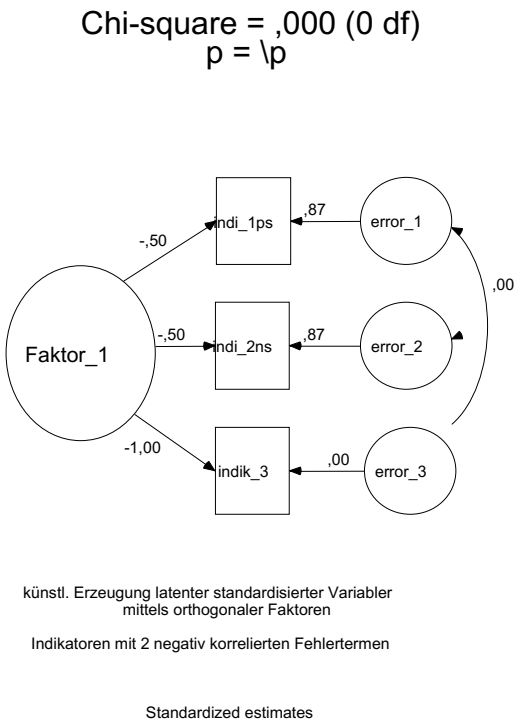


Abbildung 16: Fehlspezifiziertes Modell mit zwei negativ korrelierten Fehlervariablen



Die Modellannahme, dass die Fehlerkorrelation 0 betrage, bewirkt, dass die den Residuen der ersten beiden Indikatoren gemeinsame Komponente x0 fälschlicherweise der Komponente x4 zugerechnet wird. Dadurch *verringert sich die erklärte Varianz der ersten beiden Indikatoren* um die Hälfte (25 %) und die unerklärte Varianz der ersten beiden Indikatoren nimmt um genau diesen Betrag zu. Es werden also $\hat{\beta}_1$ und $\hat{\beta}_2$ jeweils mit $\sqrt{0.25}$=0,5 zu gering und $\hat{\beta}_f_1$ und $\hat{\beta}_f_2$ jeweils mit $\sqrt{0,75}$ = 0,87 zu hoch geschätzt. Die fehlspezifizierten Varianten des Modells mit positiv und mit negativ korrelierten Fehlern unterscheiden sich hinsichtlich der

Pfadkoeffizienten der ersten beiden Indikatoren dadurch, dass die *Schätzung der Faktorpfadkoeffizienten* $\hat{\beta}_i$ *und des Fehlerpfadkoeffizienten* $\hat{\beta}_f_i$ *vertauscht* werden.

Der Pfadkoeffizient $\hat{\beta}_3$ wird aufgrund der zu geringen Schätzung von sowohl $\hat{\beta}_1$ als auch $\hat{\beta}_2$ zu hoch geschätzt. Hierbei verdoppelt sich der Anteil der durch den Faktor erklärten Varianz des Indikators 3 auf 100 %, während sich die Anteile der durch den Faktor erklärten Varianz von Indikator 1 und 2 jeweils halbieren. (Der Anteil der Restvarianz des dritten Indikators wird im Gegenzug 0 %.) Die Wirkung der Fehlspezifikation auf die Pfadkoeffizienten des dritten Indikators ist wesentlich gravierender als im Modell mit positiv korrelierten Fehlern.

Güte des Messmodells

Die Güte des fehlspezifizierten Messmodells wird mit $0{,}5^2 + 0{,}5^2 + 1{,}0^2 = 1{,}5$ korrekt geschätzt. *An der Messmodellgüte ist die Fehlspezifikation in diesem Falle daher nicht ablesbar.*

3.5.4 MODELL MIT DREI POSITIV KORRELIERTEN FEHLERN

Dieses Modell gibt eine in der medizinischen Praxis realistische Situation dreier Diagnoseverfahren (Indikatoren) wieder, deren Fehler allesamt untereinander korreliert sind.

Die Erzeugung künstlicher Daten erfolgt so, dass alle Fehlerkorrelationen gleich hoch sind (r (error_i, error_j) = 0,5 für alle i, j und i ungleich j). Sie unterscheidet sich von jener des Modells 3a mit zwei positiv korrelierten Fehlern nur dadurch, dass der Fehlerterm des dritten Indikators nun auch mit den Fehlertermen der anderen positiv korreliert. Der dritte Indikator wird in SPSS folgendermaßen berechnet: `COMPUTE indi_3ps = x4 + (x3 + x0) / SQRT(2).`

Die Modellgleichungen sind bis auf die des dritten Indikators somit identisch mit jenen des Modells mit zwei positiv korrelierten Fehlern und lauten:

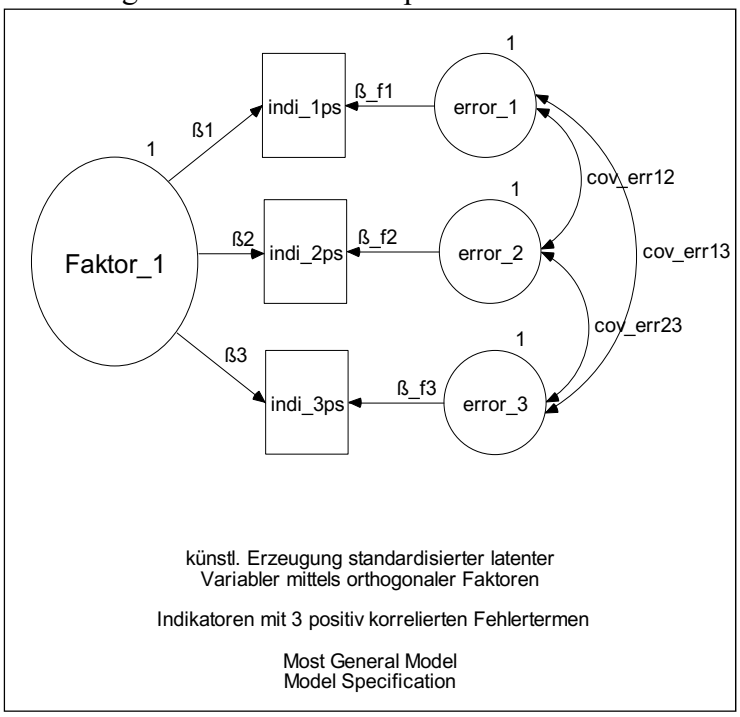
indi_1ps = $\beta_1 \cdot Faktor + \beta_f_1 \cdot error_1$

indi_2ps = $\beta_2 \cdot Faktor + \beta_f_2 \cdot error_2$

indi_3ps = $\beta_3 \cdot Faktor + \beta_f_3 \cdot error_3$

Die allgemeine Modellspezifikation ist in Abbildung 17 dargestellt.

Abbildung 17: Modell mit drei positiv korrelierten Fehlern - allgemeine Modellspezifikation



Das korrekt spezifizierte Modell (siehe Abbildung 18) enthält die Annahme, dass alle Fehlerkorrelationen 0,5 betragen und führt zu Koeffizientenschätzungen, die der tatsächlichen Korrelationsstruktur der erzeugten Variablen entsprechen. In der fehlspezifizierten Modellvariante (Abbildung 19) wird angenommen, dass alle Fehlerkorrelationen gleich Null sind. Daher werden im fehlspezifizierten Modell die wahren Pfadkoeffizienten $ß_i$ gleichmäßig über- und die wahren Pfadkoeffizienten $ß_{fi}$ gleichermaßen unterschätzt.

Abbildung 18: Korrekt spezifiziertes Modell mit drei positiv korrelierten Fehlern

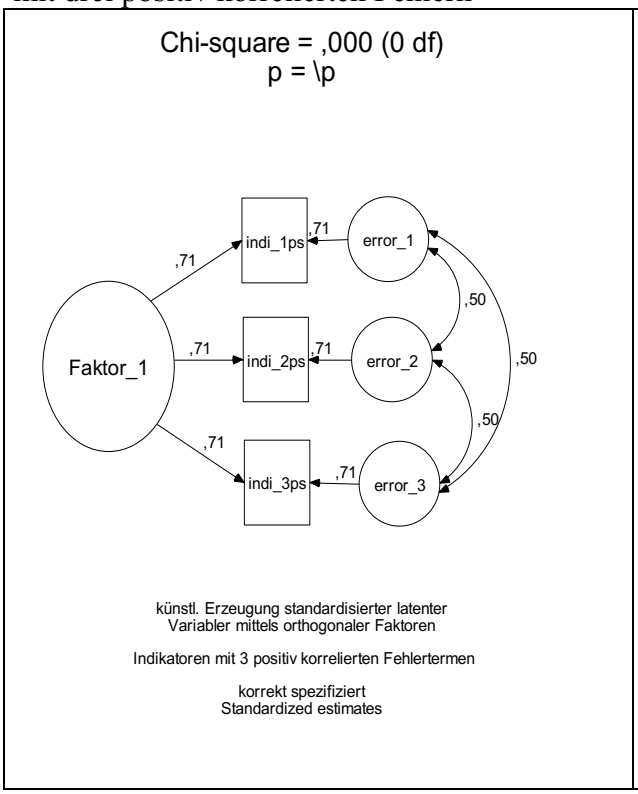
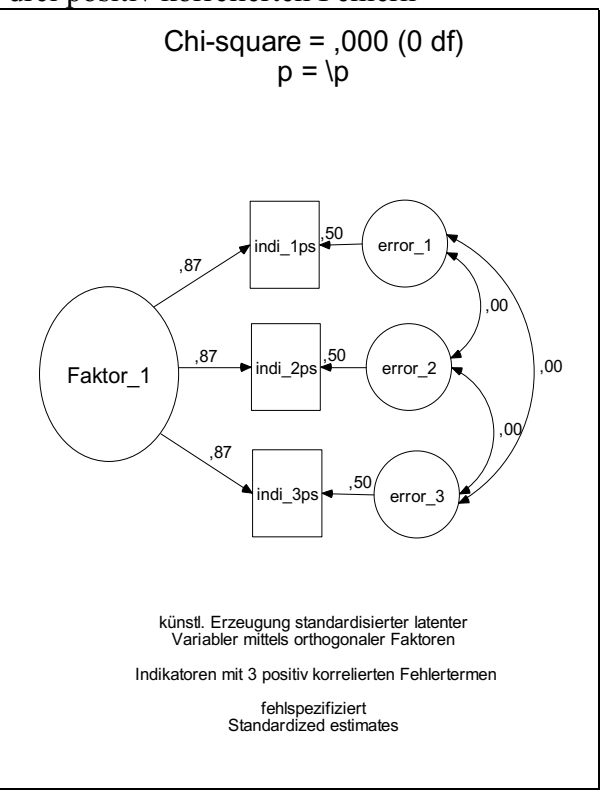


Abbildung 19: Fehlspezifiziertes Modell mit drei positiv korrelierten Fehlern



Güte des Messmodells

Die Güte des fehlspezifizierten Messmodells wird mit $0{,}866^2 + 0{,}866^2 + 0{,}866^2 = 2{,}250$ weit überschätzt.

3.5.5 FAZIT

In den Modellen mit *zwei korrelierten Fehlertermen* rührt die Korrelation der ersten beiden Indikatoren zu einem Teil daher, dass beide von einem gemeinsamen Faktor beeinflusst werden. Der andere Teil geht auf die Korrelation ihrer Fehlerterme zurück. Wenn man diese Fehlertermkorrelationen durch Annahme einer Korrelation von Null fehlspezifiziert, wird sie fälschlicherweise der Korrelation zwischen dem Faktor und den beiden Indikatoren zugerechnet. Dem Faktor wird qua Modell ein größeres (Abb.14) bzw. geringeres (Abb. 16) Potential zugeschrieben, die Varianzen der ersten beiden manifesten Variablen und deren Kovarianz zu erklären, als er tatsächlich besitzt. Die Fehlspezifikation senkt den Determinationskoeffizienten des dritten, nicht fehlerkorrelierten Indikators im Falle positiv korrelierter Fehler nur moderat. Im Modell mit negativ korrelierten Fehlern ist der Anstieg der erklärten Varianz des dritten Indikators aufgrund der Fehlspezifikation sehr stark. (Das Bestimmtheitsmaß wird

maximal). Die Fehlspezifikation führt somit in beiden Modellen zu einem Bias, der ein bzw. zwei Messverfahren als weniger geeignet erscheinen lässt, die nicht direkt messbare Variable (Faktor) zu erfassen. Im Gegenzug erscheinen die jeweils verbleibenden Messverfahren infolge des Bias als besonders geeignet.

In der Anwendungspraxis geht es um die Bewertung von Diagnoseverfahren mit dem Ziel, die Verfahren zu eruieren, welche die nicht direkt messbare Krankheit am besten annähern. Daher wäre bei einer solchen Fehlerkorrelationsstruktur (einige Fehler sind korreliert, andere nicht) die Annahme, alle Fehler seien unkorreliert, besonders problematisch, da dadurch bestimmte Diagnoseverfahren fälschlicherweise höher bewertet würden.

In der fehlspezifizierten Variante des Modells mit *drei positiv korrelierten Fehlertermen* scheinen alle drei Indikatoren in höherem Maß durch den Faktor erklärt zu werden, als es tatsächlich der Fall ist. In der Anwendungspraxis führte die Fehlspezifikation zwar zu einer besonders großen Überschätzung des Maßes, in dem die drei Diagnoseverfahren insgesamt die nicht direkt messbare Erkrankung anzunähern vermögen, aber keines der Diagnoseverfahren wird fälschlicherweise besser bewertet als ein konkurrierendes.

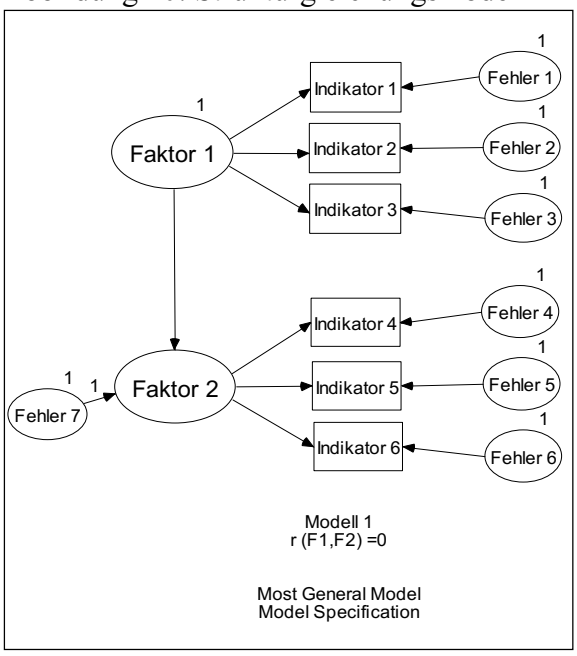
Insofern ist eine Fehlspezifikation weniger problematisch in Fällen, in welchen alle Fehlerkorrelationen fehlspezifiziert werden als in Fällen, in welchen nur ein Teil der Fehlerkorrelationen fehlspezifiziert wird.

Die Beurteilung der Frage, ob die Güte des Messmodells korrekt oder nicht korrekt geschätzt worden ist, ist kein hinreichendes Kriterium zur Entscheidung der Frage, ob ein Modell korrekt spezifiziert ist. Das Modell mit negativ korrelierten Fehlern zeigt, dass eine Fehlspezifikation Veränderungen im Bestimmtheitsmaß manifester Variabler bewirken kann, die als ein „Nullsummenspiel“ beschrieben werden können: Die Summe der „Verluste“ im Bestimmtheitsmaß zweier Messverfahren kann genau dem „Gewinn“ entsprechen, der im Bestimmtheitsmaß des dritten Koeffizienten zu verzeichnen ist.

4 Lineare Strukturgleichungsmodelle mit zwei Faktoren unter Verwendung von AMOS

Es werden lineare Strukturgleichungsmodelle mit zwei Faktoren und zugehörigen Messmodellen betrachtet. Im allgemeinen Fall ist der zweite Faktor hierbei von dem ersten Faktor statistisch abhängig. Ein Spezialfall liegt vor, wenn Faktor 2 von Faktor 1 unabhängig ist. In jedem Fall jedoch werden die hier betrachteten linearen Strukturgleichungsmodelle mit der Software AMOS graphisch wie in Abbildung 20 dargestellt.[37]

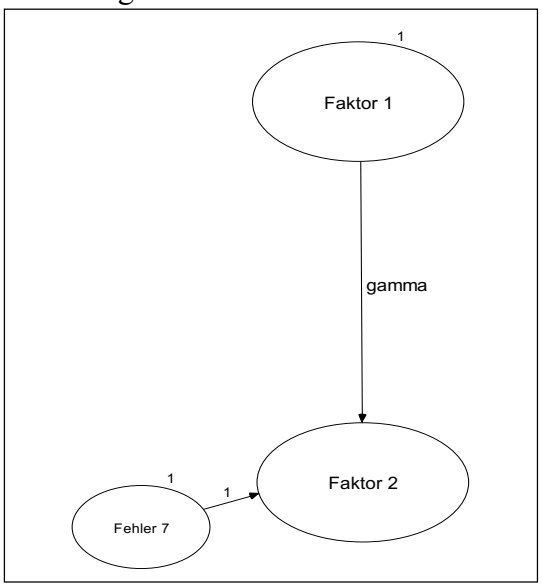
Abbildung 20: Strukturgleichungsmodell



Das Strukturmodell (Abbildung 21) zeigt, dass Faktor 2 mit dem Faktor 1 durch eine gerichtete Abhängigkeit verbunden ist.

[37] Die Zahl Eins an der Ellipse bedeutet, dass die latente Variable standardisiert ist. Auch die den Spezialfall behandelnden Modelle beinhalten einen von Faktor 1 ausgehenden und zu Faktor 2 führenden Pfad. Denn die in allen Modellen beabsichtigte Schätzung des Koeffizienten dieses Pfades ist technisch nur durchführbar, wenn die Graphik den Pfad enthält.

Abbildung 21: Strukturmodell



Aus didaktischen Gründen wird zunächst eine Reihe von Strukturgleichungsmodellen präsentiert, die dem Spezialfall der statistischen Unabhängigkeit der Faktoren entsprechen. Anschließend folgen lineare Strukturgleichungsmodelle, in welchen Faktor 2 von Faktor 1 statistisch abhängig ist.

Wenn lineare Strukturgleichungsmodelle fehlspezifiziert sind, kann dies zu einer verzerrten Schätzung des Pfadkoeffizienten gamma des Strukturmodells führen. Unter Fehlspezifikationen werden in diesem Zusammenhang falsche Modellannahmen über die Messfehlerkorrelationsstruktur verstanden. Grundsätzlich gehen wir – in Abweichung von dem vorhergehenden Kapitel über Ein-Faktor-Modelle – nur von dem Vorhandensein positiver Messfehlerkorrelationen aus.[38] Die betrachteten Strukturgleichungsmodelle weisen unterschiedliche Varianten von Fehlspezifikationen auf.

Hauptziel der Untersuchung ist es, zu beobachten, wie sich bestimmte Fehlspezifikationen auf die Schätzung des Pfadkoeffizienten[39] des Strukturmodells auswirken. Hierbei interessiert zum einen ein möglicher Zusammenhang zwischen der *Art der Fehlspezifikation* und der *Richtung der Abweichung* bei der Schätzung des Koeffizienten. Zum anderen interessiert man sich für den Zusammenhang zwischen *Ausmaß der Fehlspezifikationen* und *Stärke der Abweichung* bei der Schätzung von γ.

[38] Negative Messfehlerkorrelationen sind zwar denkbar, aber – da sie praktisch selten auftreten – nicht Gegenstand der Untersuchung.
[39] Wenn die Faktoren standardisiert und normalverteilt sind, dann ist der Pfadkoeffizient als Korrelation interpretierbar.

Es werden in der Situation, in welcher die Faktoren der zwei Messmodelle in Wirklichkeit unkorreliert sind (Spezialfall), nur bestimmte Modelle ausführlich präsentiert: Es handelt sich um Modelle, in denen Fehlerterme korreliert sind, die Indikatoren *verschiedener* Messmodelle zugehören. (In diesen Modellen können zusätzlich auch Fehlerterme korreliert sein, die Indikatoren desselben Messmodells zugehören.) Denn es zeigte sich, dass nur in solchen Modellen die Fehlspezifikation von Fehlertermkorrelationen Einfluss auf die Schätzung von γ nehmen kann.

Modelle hingegen, in welchen Fehlertermkorrelationen nur *innerhalb eines* Messmodells vorliegen, werden weniger ausführlich präsentiert. In diesen Modellen beeinflusst die Fehlspezifikation einer Fehlertermkorrelation wegen der Unabhängigkeit der Messmodellfaktoren untereinander die Schätzung von γ nicht, so dass sich eine ausführlichere Darstellung erübrigt.

Es werden in der Situation, in welcher die Faktoren der zwei Messmodelle in Wirklichkeit korreliert sind (allgemeiner Fall), Modelle aller oben genannten Kategorien ausführlich präsentiert. Die in der Situation unkorrelierter Faktoren weniger ausführlich präsentierten Modelle werden in diesem Zusammenhang eingehender dargestellt. Denn die Fehlspezifikation von Fehlertermkorrelationen dieser Modelle kann wegen der Korrelation der Messmodell-Faktoren untereinander die Schätzung von γ beeinträchtigen.

4.1 STRUKTURGLEICHUNGSMODELLE, DEREN FAKTOREN STATISTISCH UNABHÄNGIG SIND

In der folgenden Tabelle werden die drei Komponenten der linearen Strukturgleichungsmodelle, das Strukturmodell und die beiden Messmodelle, jeweils auf drei Ebenen beschrieben. Die erste Ebene „in SPSS gesetzt“ gibt die tatsächliche, in SPSS erzeugte Korrelationsstruktur der manifesten und der als latent betrachteten Variablen wieder. Die zweite Ebene „in AMOS gesetzt“ gibt an, ob Korrelationen vor der Modellschätzung in AMOS festgelegt worden sind, und wenn ja, in welcher Höhe. Die dritte Ebene „in AMOS geschätzt“ enthält die Ergebnisse der Schätzung der jeweiligen Korrelationen.

Im Mittelpunkt des Interesses steht der Bias des Pfadkoeffizienten des Strukturmodells γ.[40]

In den Messmodellen geben die Indizes der Pfadkoeffizienten Lambda die Ursprungs- und Zielvariablen an. Die erste Ziffer des Index bezeichnet die Variable, von der der Pfad ausgeht

[40] Der Bias gibt die verzerrende Wirkung einer bestimmten Fehlspezifikation auf die Schätzung des Einflusses an, den eine Krankheit oder Störgröße (Faktor 1) auf eine andere Krankheit (Faktor 2) ausübt.

(Ursprung). Die zweite Ziffer des Index bezeichnet die Variable, auf die der Pfad gerichtet ist (Ziel). Der Ursprung ist hierbei eine latente Variable (Faktor bzw. Fehlervariable ε) und das Ziel eine manifeste Variable.

Die tabellarische Beschreibung wird illustriert durch eine grafische Darstellung, die das *fehlspezifizierte Modell* wiedergibt. In dieser gibt die Zahl am Pfadkoeffizienten eine Schätzung in AMOS wieder. Die Zahl an dem die Fehlerkorrelation symbolisierenden Doppelpfeil informiert hingegen über eine in AMOS gesetzte Restriktion.

4.1.1 MODELL 1

Im ersten Modell sind alle Fehlervariablen untereinander unkorreliert.

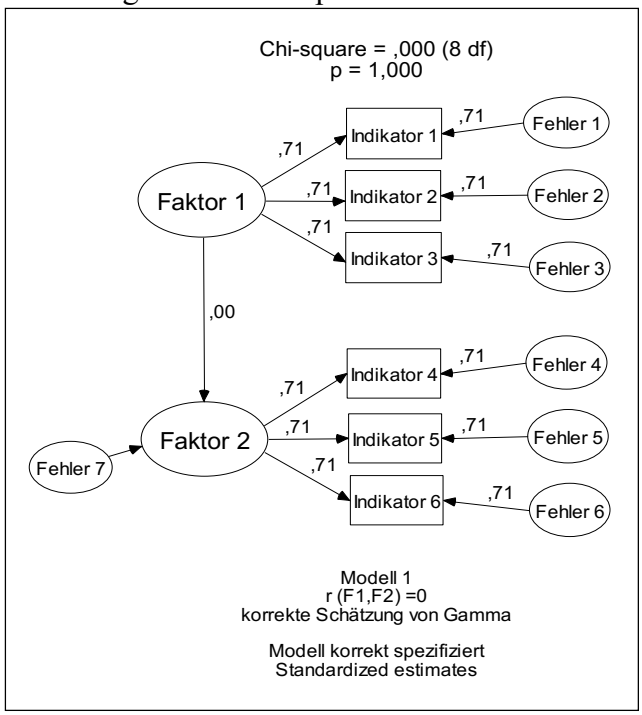
Beschreibung:

Tabelle 4: Tatsächlich erzeugte, vor der Modellschätzung festgelegte und geschätzte Korrelationen des Grundmodells

	In SPSS gesetzt	In AMOS gesetzt	In AMOS geschätzt
Strukturmodell	$\gamma = 0$	-	$\hat{\gamma} = 0$
Messmodelle	$\lambda_{11} = 0{,}71$	-	$\hat{\lambda}_{11} = 0{,}71$
	$\lambda_{12} = 0{,}71$	-	$\hat{\lambda}_{12} = 0{,}71$
	$\lambda_{13} = 0{,}71$	-	$\hat{\lambda}_{13} = 0{,}71$
	$\lambda_{24} = 0{,}71$	-	$\hat{\lambda}_{24} = 0{,}71$
	$\lambda_{25} = 0{,}71$	-	$\hat{\lambda}_{25} = 0{,}71$
	$\lambda_{26} = 0{,}71$	-	$\hat{\lambda}_{26} = 0{,}71$
	$\lambda_{\varepsilon_1 1} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_1 1} = 0{,}71$
	$\lambda_{\varepsilon_2 2} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_2 2} = 0{,}71$
	$\lambda_{\varepsilon_3 3} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_3 3} = 0{,}71$
	$\lambda_{\varepsilon_4 4} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_4 4} = 0{,}71$
	$\lambda_{\varepsilon_5 5} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_5 5} = 0{,}71$
	$\lambda_{\varepsilon_6 6} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_6 6} = 0{,}71$

Grafische Darstellung

Abbildung 22: Korrekt spezifiziertes Modell 1



Interpretation

Das Modell ist korrekt spezifiziert. Dies bedeutet im Einzelnen:

1. Die Pfadkoeffizienten aller Indikatoren werden der tatsächlichen Korrelation entsprechend geschätzt.

2. Die Güte eines Messmodells wird daran abgelesen, ob die Summe der drei geschätzten Determinationskoeffizienten des Messmodells, z. B. $\hat{\lambda}_{11}^{\,2} + \hat{\lambda}_{12}^{\,2} + \hat{\lambda}_{13}^{\,2}$ beim ersten Messmodell von der Summe der in SPSS erzeugten Determinationskoeffizienten, $\lambda_{11}^{\,2} + \lambda_{12}^{\,2} + \lambda_{13}^{\,2}$, abweicht. Die Summe der in AMOS geschätzten Determinationskoeffizienten beträgt $0{,}71^2 + 0{,}71^2 + 0{,}71^2 = 1{,}5$ und ist gleich der Summe der in SPSS festgelegten.

Die Güte der beiden Messmodelle des ersten linearen Strukturgleichungsmodells wird also korrekt geschätzt.

3. Der Pfadkoeffizient des Strukturmodells wird daher ebenso korrekt geschätzt ($\hat{\gamma} = 0.71$).

4.1.2 MODELL 2

Im zweiten Modell sind alle Fehlervariablen mit einer Ausnahme untereinander unkorreliert. Die Fehlervariablen der Indikatoren 1 und 4 weisen eine Korrelation von 0,5 auf. Falls die in

AMOS geschätzten Parameter von den in SPSS festgelegten abweichen, sind die geschätzten in den folgenden Tabellen durch Fettdruck hervorgehoben. Es werden hierbei nur die ersten sieben Zeilen der Tabelle berücksichtigt, welche gamma und die Koeffizienten der Pfade enthalten, die von den Faktoren zu den Indikatoren führen. (Abweichungen in den sog. Fehlerpfadkoeffizienten sind nicht hervorgehobenen.)

Beschreibung (des fehlspezifizierten Modells)

	In SPSS gesetzt	In AMOS gesetzt	In AMOS geschätzt
Strukturmodell	$\gamma = 0$	-	$\hat{\gamma} =$ **0,06**
Messmodelle	$\lambda_{11} = 0{,}71$	-	$\hat{\lambda}_{11} = 0{,}71$
	$\lambda_{12} = 0{,}71$	-	$\hat{\lambda}_{12} = 0{,}71$
	$\lambda_{13} = 0{,}71$	-	$\hat{\lambda}_{13} = 0{,}71$
	$\lambda_{24} = 0{,}71$	-	$\hat{\lambda}_{24} = 0{,}71$
	$\lambda_{25} = 0{,}71$	-	$\hat{\lambda}_{25} = 0{,}71$
	$\lambda_{26} = 0{,}71$	-	$\hat{\lambda}_{26} = 0{,}71$
	$\lambda_{\varepsilon_1 1} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_1 1} =$ **0,70**
	$\lambda_{\varepsilon_2 2} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_2 2} = 0{,}71$
	$\lambda_{\varepsilon_3 3} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_3 3} = 0{,}71$
	$\lambda_{\varepsilon_4 4} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_4 4} =$ **0,70**
	$\lambda_{\varepsilon_5 5} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_5 5} = 0{,}71$
	$\lambda_{\varepsilon_6 6} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_6 6} = 0{,}71$
	$r_{\varepsilon_1,\varepsilon_4} = 0{,}5$	$r_{\varepsilon_1,\varepsilon_4} = 0$	-

Grafische Darstellung

Abbildung 23: Fehlspezifiziertes Modell 2

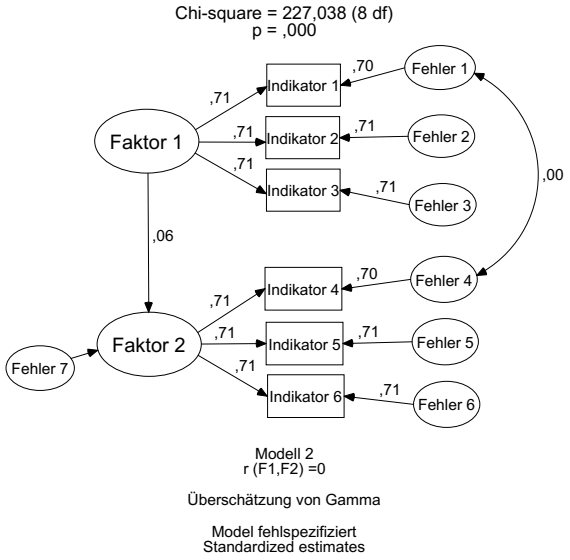
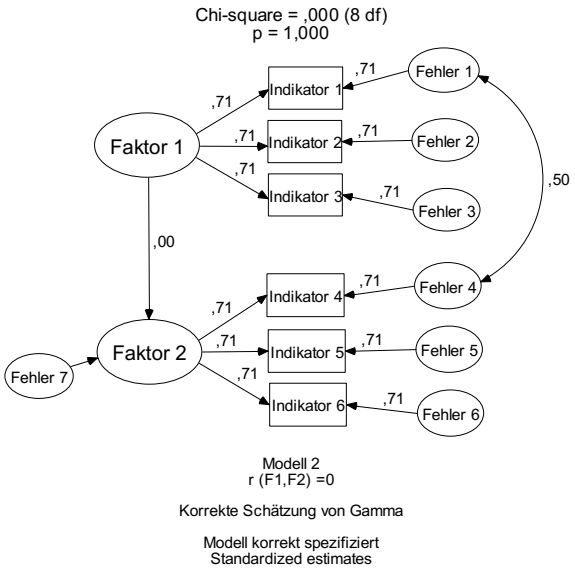


Abbildung 24: Korrekt spezifiziertes Modell 2



Interpretation von Abbildung 23

Die Fehlspezifikation in Form einer Ignorierung der Korrelation zwischen Fehler 1 und 4 betrifft beide Messmodelle und hat folgende Konsequenz:

1. Die Pfadkoeffizienten der beiden fehlerkorrelierten Indikatoren 1 und 4 werden – sowie die der anderen – der tatsächlichen Korrelation entsprechend geschätzt. (Nur die Pfadkoeffizienten $\lambda_{\varepsilon_1 1}$ und $\lambda_{\varepsilon_4 4}$ der zugehörigen Fehler werden minimal unterschätzt.)

2. Die Güte der beiden Messmodelle wird – ebenso wie die des ersten Modells – korrekt geschätzt.

3. Die Korrelation der Fehler 1 und 4 wird aufgrund der Restriktion, dass diese Fehlerkorrelation 0 sei, als eine Korrelation zwischen den Faktoren interpretiert. Der Pfadkoeffizient des Strukturmodells wird daher leicht überschätzt ($\hat{\gamma} = 0{,}06$).

4. Insgesamt reagiert das Modell also recht robust auf die gewählte Fehlspezifikation.

4.1.3 MODELL 3

Das dritte Modell weist zusätzlich zum zweiten eine tatsächliche Fehlerkorrelation zwischen zwei Indikatoren des ersten Messmodells auf.

Beschreibung (des fehlspezifizierten Modells)

	In SPSS gesetzt	In AMOS gesetzt	In AMOS geschätzt
Strukturmodell	$\gamma = 0$	-	$\hat{\gamma} =$ **0,07**
Messmodelle	$\lambda_{11} = 0{,}71$	-	$\hat{\lambda}_{11} =$ **0,87**
	$\lambda_{12} = 0{,}71$	-	$\hat{\lambda}_{12} =$ **0,86**
	$\lambda_{13} = 0{,}71$	-	$\hat{\lambda}_{13} =$ **0,58**
	$\lambda_{24} = 0{,}71$	-	$\hat{\lambda}_{24} = 0{,}71$
	$\lambda_{25} = 0{,}71$	-	$\hat{\lambda}_{25} =$ **0,70**
	$\lambda_{26} = 0{,}71$	-	$\hat{\lambda}_{26} =$ **0,70**
	$\lambda_{\varepsilon_1 1} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_1 1} = 0{,}49$
	$\lambda_{\varepsilon_2 2} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_2 2} = 0{,}51$
	$\lambda_{\varepsilon_3 3} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_3 3} = 0{,}82$
	$\lambda_{\varepsilon_4 4} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_4 4} = 0{,}70$
	$\lambda_{\varepsilon_5 5} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_5 5} = 0{,}71$
	$\lambda_{\varepsilon_6 6} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_6 6} = 0{,}71$
	$r_{\varepsilon_1,\varepsilon_4} = 0{,}5$	$r_{\varepsilon_1,\varepsilon_4} = 0$	-
	$r_{\varepsilon_1,\varepsilon_2} = 0{,}5$	$r_{\varepsilon_1,\varepsilon_2} = 0$	-

Grafische Darstellung

Abbildung 25: Fehlspezifiziertes Modell 3

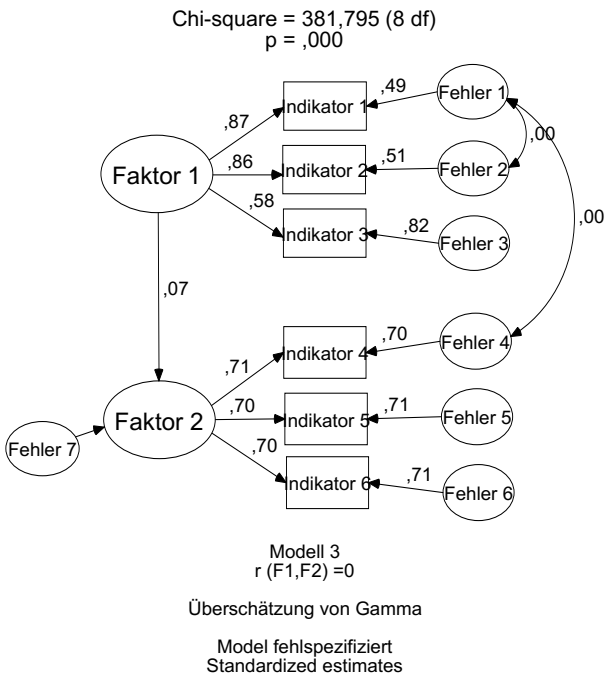
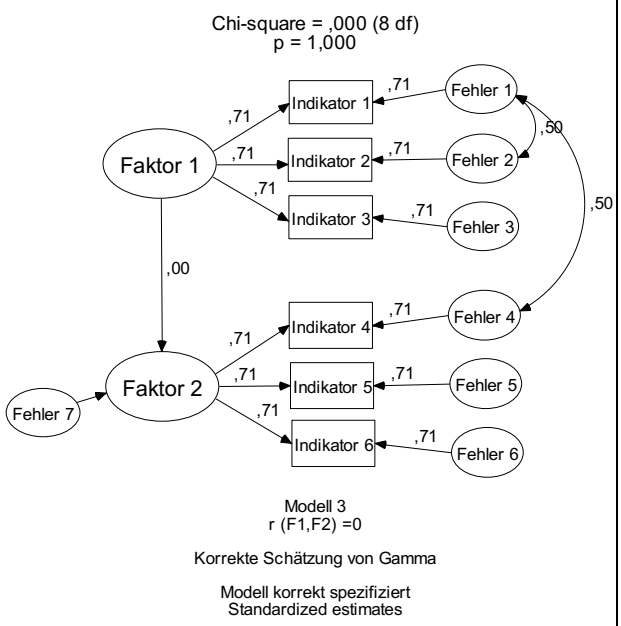


Abbildung 26: Korrekt spezifiziertes Modell 3



Interpretation von Abbildung 25

Die Fehlspezifikation des Messmodells für Faktor 1 hat folgende Konsequenz:

1. Die Pfadkoeffizienten der beiden fehlerkorrelierten Indikatoren 1 und 2, λ_{11} bzw. λ_{12} , werden überschätzt, der Pfadkoeffizient der dritten manifesten Variablen, λ_{13} wird unterschätzt.

2. Die Güte des Messmodells für Faktor 1 wird mit $0{,}87^2 + 0{,}86^2 + 0{,}58^2 = 1{,}83$ überschätzt.

Dies führte zu einem Absinken des Pfadkoeffizienten des Strukturmodells gamma, wenn Faktor 2 von Faktor 1 abhängig wäre. Da die Faktoren unabhängig voneinander sind, kann gamma nicht absinken.

3. Wie im Modell 2 wird die Korrelation der Fehler 1 und 4 ignoriert und als eine Korrelation zwischen den Faktoren interpretiert, so dass der entsprechende Pfadkoeffizient des Strukturmodells überschätzt wird $(\hat{\gamma} = 0{,}07)$.

4. Die Kombination beider Fehlspezifikationen hat weitere Konsequenzen:

Da der Indikator 1 aufgrund der Fehlspezifikation des ersten Messmodells als besonders guter Indikator geschätzt wird, wirkt sich die Ignorierung der Fehlerkorrelation zwischen diesem Indikator und dem Indikator 4 etwas stärker aus als im Vormodell. Dies äußert sich in zweierlei Form: Zum einen wird der Pfadkoeffizient des Strukturmodells γ mit 0,07 statt 0,06 geringfügig höher geschätzt, zum anderen wirkt sich die Überschätzung der beiden Indikatoren des ersten Messmodells beim ersten Indikator etwas stärker aus. Der Pfadkoeffizient λ_{11} wird mit 0,87 ein wenig höher geschätzt als λ_{12} mit 0,86. Auch im zweiten Messmodell wird der Indikator 4 leicht besser eingeschätzt als die beiden anderen Indikatoren, deren Pfadkoeffizienten $\hat{\lambda}_{25}$ und $\hat{\lambda}_{26}$ geringfügig hinter der tatsächlichen Korrelation zurückbleiben.[41]

Aus diesem Modell lässt sich folgende *qualitative Erkenntnis ableiten*: Wenn die Fehlervariablen zweier Indikatoren eines Messmodells positiv korreliert sind und diese Korrelation ignoriert wird, schneiden die Indikatoren nicht etwa als schlechte, sondern als besonders gute Indikatoren für eine zu messende Krankheit ab. Daher sollte vor einer Bewertung der Güte einer Messgröße deren Fehlervariable auf Korreliertheit überprüft werden.

4.1.4 ZUSATZMODELLE Z1–Z3

Zum Ende von 4.1 werden drei Zusatzmodelle in Kurzform dokumentiert.

Hierbei handelt es sich um Modelle, in denen Fehlertermkorrelationen nur *innerhalb eines* Messmodells vorliegen. Die Fehlspezifikation einer Fehlertermkorrelation beeinflusst wegen der Unabhängigkeit der Messmodellfaktoren untereinander die Schätzung des Strukturmodellpfadkoeffizienten γ nicht. Daher erscheinen sie in diesem Kapitel von geringerer Bedeutung, dessen primäres Ziel in der Untersuchung der Wirkung von Fehlspezifikation(en) auf gamma besteht.

4.1.5 Z 1

Dieses Modell weist eine tatsächliche Fehlerkorrelation zwischen zwei Indikatoren des ersten Messmodells auf.

---

[41] Das minimale Absinken der Pfadkoeffizienten λ_{25} und λ_{26} im zweiten Messmodell geht primär nicht auf die Fehlspezifikation der Korrelation zwischen Fehler 1 und 4 zurück, sondern auf die Fehlspezifikation des ersten Messmodells. Im Modell 2, das sich nur dadurch von Modell 3 unterscheidet, die fehlspezifizierte Korrelation zwischen Fehler 1 und 2 nicht zu enthalten, konnte ein solches Absinken nämlich nicht beobachtet werden.

Beschreibung des fehlspezifizierten Modells

	In SPSS gesetzt	In AMOS gesetzt	In AMOS geschätzt
Strukturmodell	$\gamma = 0$	-	$\hat{\gamma} = 0{,}00$
Messmodelle	$\lambda_{11} = 0{,}71$	-	$\hat{\lambda}_{11} = 0{,}87$
	$\lambda_{12} = 0{,}71$	-	$\hat{\lambda}_{12} = 0{,}87$
	$\lambda_{13} = 0{,}71$	-	$\hat{\lambda}_{13} = 0{,}58$
	$\lambda_{24} = 0{,}71$	-	$\hat{\lambda}_{24} = 0{,}71$
	$\lambda_{25} = 0{,}71$	-	$\hat{\lambda}_{25} = 0{,}71$
	$\lambda_{26} = 0{,}71$	-	$\hat{\lambda}_{26} = 0{,}71$
	$\lambda_{\varepsilon_1 1} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_1 1} = 0{,}50$
	$\lambda_{\varepsilon_2 2} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_2 2} = 0{,}50$
	$\lambda_{\varepsilon_3 3} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_3 3} = 0{,}82$
	$\lambda_{\varepsilon_4 4} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_4 4} = 0{,}71$
	$\lambda_{\varepsilon_5 5} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_5 5} = 0{,}71$
	$\lambda_{\varepsilon_6 6} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_6 6} = 0{,}71$
	$r_{\varepsilon_1,\varepsilon_2} = 0{,}5$	$r_{\varepsilon_1,\varepsilon_2} = 0$	-

Grafische Darstellung

Abbildung 27: Fehlspezifiziertes Modell Z1

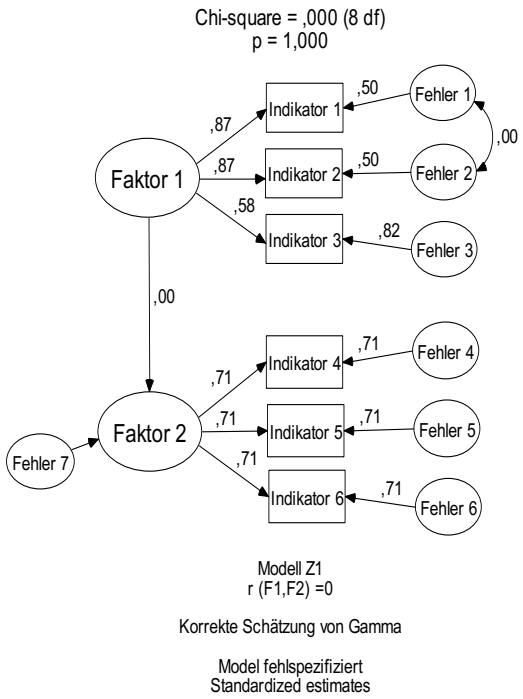
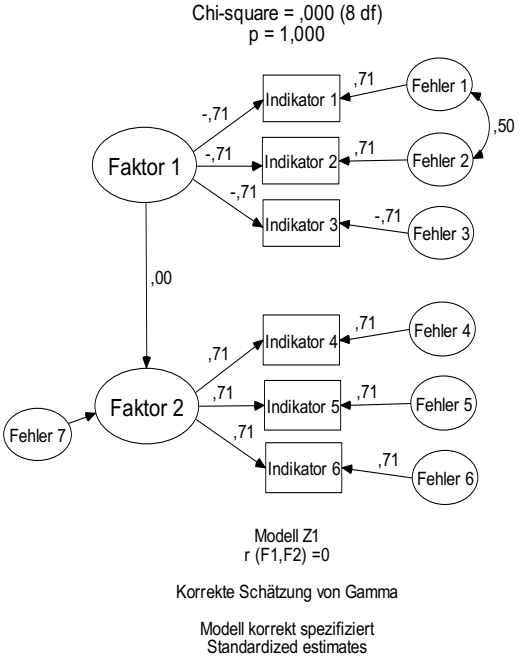


Abbildung 28: Korrekt spezifiziertes Modell Z1

negative Vorzeichen der Pfadkoeffizienten: siehe 3.5.2, Ergebnisse



Interpretation von Abbildung 27

Die Pfadkoeffizientenschätzungen des ersten Messmodells entsprechen exakt jenen des fehlspezifizierten Modells mit zwei positiv korrelierten Fehlervariablen (3.5.2.). Die Fehlspezifikation der Fehlerkorrelation im ersten Messmodell hat keinerlei Wirkung – weder auf das Strukturmodell, noch auf das zweite Messmodell. Sie beeinträchtigt auch nicht die perfekte Anpassung des Modells an die Daten.

4.1.6 Z 2

Dieses Modell weist sowohl eine tatsächliche Fehlerkorrelation zwischen zwei Indikatoren des ersten, also auch eine zwischen zwei Indikatoren des zweiten Messmodells auf.

Beschreibung des fehlspezifizierten Modells

	In SPSS gesetzt	In AMOS gesetzt	In AMOS geschätzt
Strukturmodell	$\gamma = 0$	-	$\hat{\gamma} = 0{,}00$
Messmodelle	$\lambda_{11} = 0{,}71$	-	$\hat{\lambda}_{11} = \mathbf{0{,}87}$
	$\lambda_{12} = 0{,}71$	-	$\hat{\lambda}_{12} = \mathbf{0{,}87}$
	$\lambda_{13} = 0{,}71$	-	$\hat{\lambda}_{13} = \mathbf{0{,}58}$
	$\lambda_{24} = 0{,}71$	-	$\hat{\lambda}_{24} = \mathbf{0{,}87}$
	$\lambda_{25} = 0{,}71$	-	$\hat{\lambda}_{25} = \mathbf{0{,}87}$
	$\lambda_{26} = 0{,}71$	-	$\hat{\lambda}_{26} = \mathbf{0{,}58}$
	$\lambda_{\varepsilon_1 1} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_1 1} = 0{,}50$
	$\lambda_{\varepsilon_2 2} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_2 2} = 0{,}50$
	$\lambda_{\varepsilon_3 3} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_3 3} = 0{,}82$
	$\lambda_{\varepsilon_4 4} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_4 4} = 0{,}71$
	$\lambda_{\varepsilon_5 5} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_5 5} = 0{,}71$
	$\lambda_{\varepsilon_6 6} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_6 6} = 0{,}71$
	$r_{\varepsilon_1,\varepsilon_2} = 0{,}5$	$r_{\varepsilon_1,\varepsilon_2} = 0$	-
	$r_{\varepsilon_4,\varepsilon_5} = 0{,}5$	$r_{\varepsilon_4,\varepsilon_5} = 0$	-

Grafische Darstellung

Abbildung 29: Fehlspezifiziertes Modell Z2

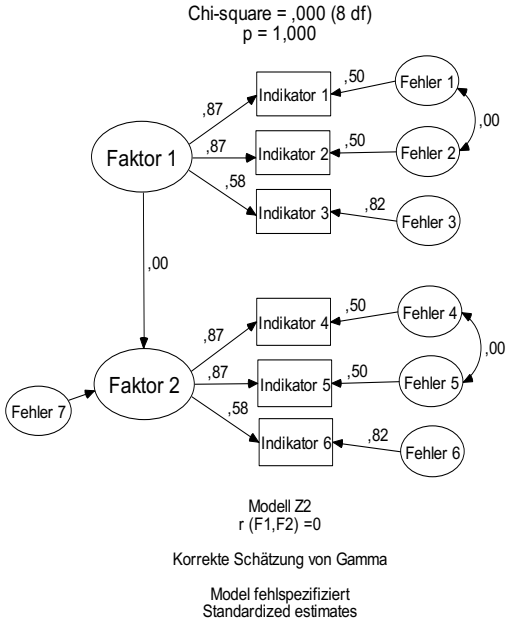
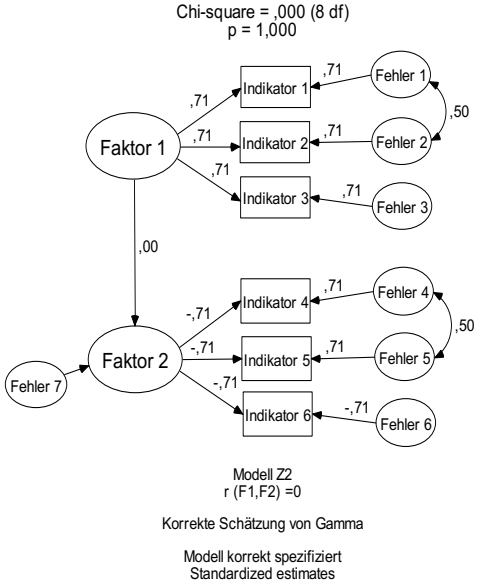


Abbildung 30: Korrekt spezifiziertes Modell Z2



negative Vorzeichen der Pfadkoeffizienten: siehe 3.5.2, Ergebnisse

Interpretation von Abbildung 29

Die Pfadkoeffizientenschätzungen sowohl des ersten also auch die des zweiten Messmodells entsprechen wie im Vormodell jenen des fehlspezifizierten Modells mit zwei positiv korrelierten Fehlervariablen (3.5.2). Die Fehlspezifikationen der Fehlerkorrelationen in den beiden Messmodellen haben keinerlei Wirkung auf das Strukturmodell und mindern auch nicht die optimale Anpassung des Modells an die Daten.

4.1.7 Z 3

Dieses Modell weist dieselben tatsächlichen Fehlerkorrelationen wie das Vormodell auf Z2 auf. Im Gegensatz zu letzteren ist jedoch die Fehlerkorrelation des ersten Messmodells nicht fehlspezifiziert

Beschreibung des fehlspezifizierten Modells

	In SPSS gesetzt	In AMOS gesetzt	In AMOS geschätzt
Strukturmodell	$\gamma = 0$	-	$\hat{\gamma} = 0{,}00$
Messmodelle	$\lambda_{11} = 0{,}71$	-	$\hat{\lambda}_{11} = \mathbf{0{,}87}$
	$\lambda_{12} = 0{,}71$	-	$\hat{\lambda}_{12} = \mathbf{0{,}87}$
	$\lambda_{13} = 0{,}71$	-	$\hat{\lambda}_{13} = \mathbf{0{,}58}$
	$\lambda_{24} = 0{,}71$	-	$\hat{\lambda}_{24} = \mathbf{0{,}87}$
	$\lambda_{25} = 0{,}71$	-	$\hat{\lambda}_{25} = \mathbf{0{,}87}$
	$\lambda_{26} = 0{,}71$	-	$\hat{\lambda}_{26} = \mathbf{0{,}58}$
	$\lambda_{\varepsilon_1 1} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_1 1} = 0{,}50$
	$\lambda_{\varepsilon_2 2} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_2 2} = 0{,}50$
	$\lambda_{\varepsilon_3 3} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_3 3} = 0{,}82$
	$\lambda_{\varepsilon_4 4} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_4 4} = 0{,}71$
	$\lambda_{\varepsilon_5 5} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_5 5} = 0{,}71$
	$\lambda_{\varepsilon_6 6} = 0{,}71$	-	$\hat{\lambda}_{\varepsilon_6 6} = 0{,}71$
	$r_{\varepsilon_1,\varepsilon_2} = 0{,}5$	$r_{\varepsilon_1,\varepsilon_2} = 0$	-
	$r_{\varepsilon_4,\varepsilon_5} = 0{,}5$	$r_{\varepsilon_4,\varepsilon_5} = 0$	-

Grafische Darstellung

Abbildung 31: Fehlspezifiziertes Modell Z3

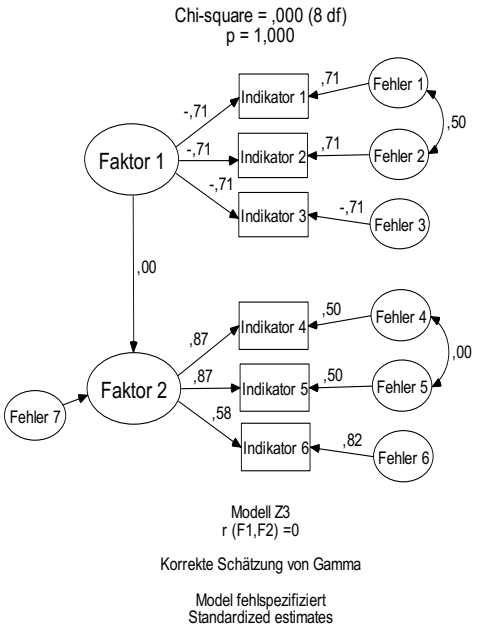
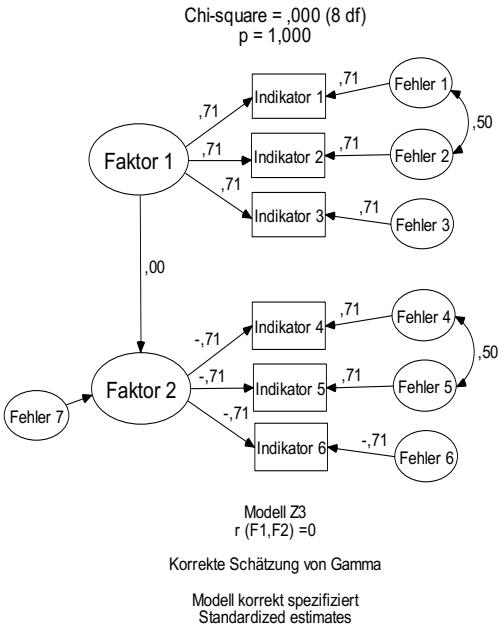


Abbildung 32: Korrekt spezifiziertes Modell Z3



Negative Vorzeichen der Pfadkoeffizienten: siehe 3.5.2, Ergebnisse.

Interpretation von Abbildung 31

Die Pfadkoeffizientenschätzungen des zweiten Messmodells entsprechen exakt jenen des fehlspezifizierten Modells mit zwei positiv korrelierten Fehlervariablen (3.5.2.). Die Fehlspezifikation der Fehlerkorrelation im zweiten Messmodell beeinflusst weder Struktur noch das erste Messmodell und mindert auch nicht die bestmögliche Anpassung des Modells an die Daten.

4.2 STRUKTURGLEICHUNGSMODELLE, DEREN FAKTOREN STATISTISCH ABHÄNGIG SIND

4.2.1 MODELL 4

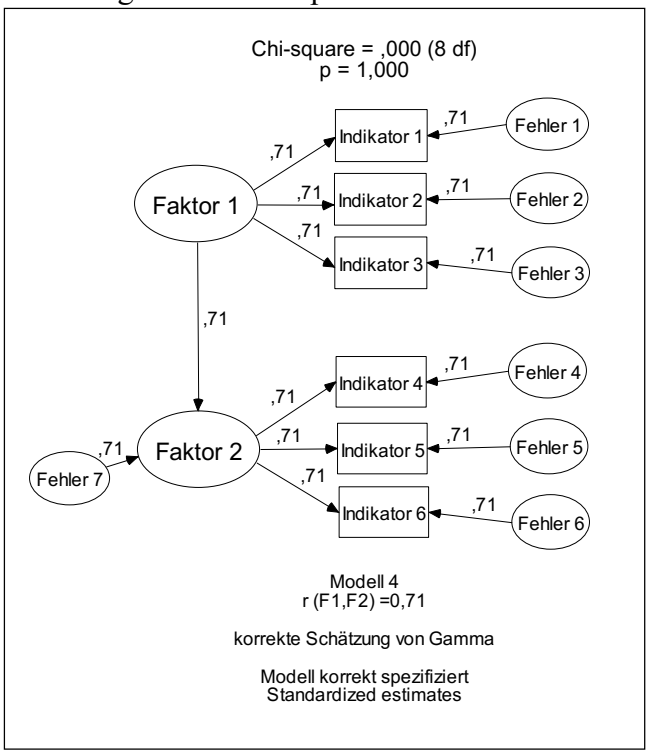
Im vierten Modell sind – wie im ersten Modell – alle Fehlervariablen untereinander unkorreliert. Es unterscheidet sich von letzterem nur dadurch, dass Faktor 2 von Faktor 1 linear abhängig ist.

Beschreibung

	In SPSS gesetzt	In AMOS gesetzt	In AMOS geschätzt
Strukturmodell	$\gamma = 0,71$	-	$\hat{\gamma} = 0,71$
Messmodelle	$\lambda_{11} = 0,71$	-	$\hat{\lambda}_{11} = 0,71$
	$\lambda_{12} = 0,71$	-	$\hat{\lambda}_{12} = 0,71$
	$\lambda_{13} = 0,71$	-	$\hat{\lambda}_{13} = 0,71$
	$\lambda_{24} = 0,71$	-	$\hat{\lambda}_{24} = 0,71$
	$\lambda_{25} = 0,71$	-	$\hat{\lambda}_{25} = 0,71$
	$\lambda_{26} = 0,71$	-	$\hat{\lambda}_{26} = 0,71$
	$\lambda_{\varepsilon_1 1} = 0,71$	-	$\hat{\lambda}_{\varepsilon_1 1} = 0,71$
	$\lambda_{\varepsilon_2 2} = 0,71$	-	$\hat{\lambda}_{\varepsilon_2 2} = 0,71$
	$\lambda_{\varepsilon_3 3} = 0,71$	-	$\hat{\lambda}_{\varepsilon_3 3} = 0,71$
	$\lambda_{\varepsilon_4 4} = 0,71$	-	$\hat{\lambda}_{\varepsilon_4 4} = 0,71$
	$\lambda_{\varepsilon_5 5} = 0,71$	-	$\hat{\lambda}_{\varepsilon_5 5} = 0,71$
	$\lambda_{\varepsilon_6 6} = 0,71$	-	$\hat{\lambda}_{\varepsilon_6 6} = 0,71$

Grafische Darstellung

Abbildung 33: Korrekt spezifiziertes Modell 4



Interpretation

Die Interpretation des Modells entspricht der des Modells 1.

4.2.2 MODELL 5

Das fünfte Modell weist eine Fehlerkorrelation zwischen zwei Indikatoren des ersten Messmodells auf; alle anderen Fehler sind untereinander unkorreliert.

Beschreibung

	In SPSS gesetzt	In AMOS gesetzt	In AMOS geschätzt
Strukturmodell	$\gamma = 0,71$	-	$\hat{\gamma} = \mathbf{0,60}$
Messmodelle	$\lambda_{11} = 0,71$	-	$\hat{\lambda}_{11} = \mathbf{0,86}$
	$\lambda_{12} = 0,71$	-	$\hat{\lambda}_{12} = \mathbf{0,86}$
	$\lambda_{13} = 0,71$	-	$\hat{\lambda}_{13} = \mathbf{0,60}$
	$\lambda_{24} = 0,71$	-	$\hat{\lambda}_{24} = 0,71$
	$\lambda_{25} = 0,71$	-	$\hat{\lambda}_{25} = 0,71$
	$\lambda_{26} = 0,71$	-	$\hat{\lambda}_{26} = 0,71$
	$\lambda_{\varepsilon_1 1} = 0,71$	-	$\hat{\lambda}_{\varepsilon_1 1} = 0,51$
	$\lambda_{\varepsilon_2 2} = 0,71$	-	$\hat{\lambda}_{\varepsilon_2 2} = 0,51$
	$\lambda_{\varepsilon_3 3} = 0,71$	-	$\hat{\lambda}_{\varepsilon_3 3} = 0,80$
	$\lambda_{\varepsilon_4 4} = 0,71$	-	$\hat{\lambda}_{\varepsilon_4 4} = 0,71$
	$\lambda_{\varepsilon_5 5} = 0,71$	-	$\hat{\lambda}_{\varepsilon_5 5} = 0,71$
	$\lambda_{\varepsilon_6 6} = 0,71$	-	$\hat{\lambda}_{\varepsilon_6 6} = 0,71$
	$r_{\varepsilon_1,\varepsilon_2} = 0,5$	$r_{\varepsilon_1,\varepsilon_2} = 0$	-

Grafische Darstellung

Abbildung 34: Fehlspezifiziertes Modell 5

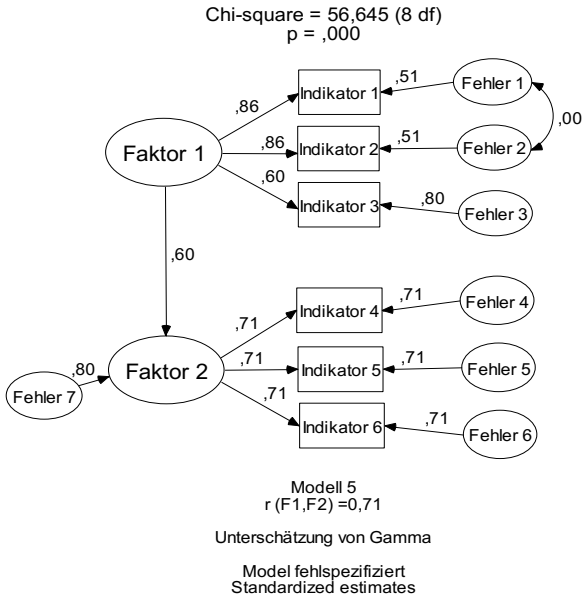
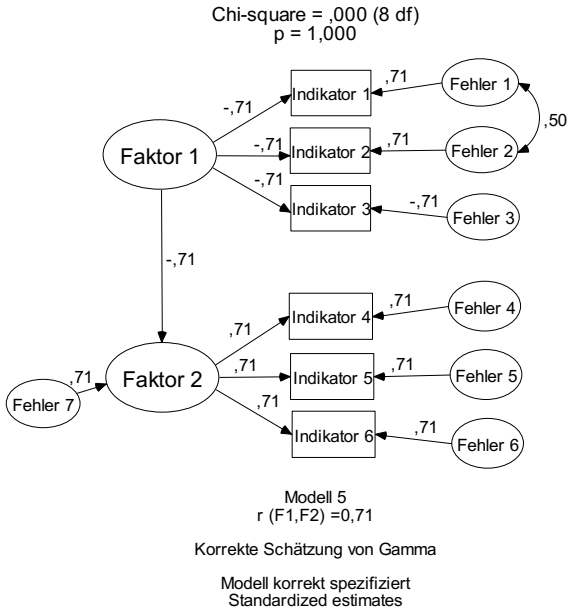


Abbildung 35: Korrekt spezifiziertes Modell 5



Negative Vorzeichen der Pfadkoeffizienten: siehe 3.5.2, Ergebnisse.

Interpretation von Abbildung 34

Die Fehlspezifikation des Messmodells für Faktor 1 hat folgende Konsequenzen:

1. Die Pfadkoeffizienten der beiden fehlerkorrelierten Indikatorvariablen 1 und 2 (λ_{11} bzw. λ_{12} **)** werden überschätzt, der Pfadkoeffizient der dritten Variablen (λ_{13}) wird unterschätzt.

2. Die Güte des Messmodells für Faktor 1 wird überschätzt.

3. Daher wird der Pfadkoeffizient des Strukturmodells (γ) unterschätzt.

Eine Fehlspezifikation innerhalb eines Messmodells wirkt sich also konservativ auf die Schätzung des interessierenden Pfadkoeffizienten aus.

4.2.3 MODELL 6

Das sechste Modell weist zusätzlich zur Fehlerkorrelation des ersten Messmodells (siehe Vormodell) eine Korrelation zwischen zwei Fehlern des zweiten Messmodells auf.

Beschreibung

	In SPSS gesetzt	In AMOS gesetzt	In AMOS geschätzt
Strukturmodell	$\gamma = 0,71$	-	$\hat{\gamma} = \mathbf{0,51}$
Messmodelle	$\lambda_{11} = 0,71$	-	$\hat{\lambda}_{11} = \mathbf{0,86}$
	$\lambda_{12} = 0,71$	-	$\hat{\lambda}_{12} = \mathbf{0,86}$
	$\lambda_{13} = 0,71$	-	$\hat{\lambda}_{13} = \mathbf{0,59}$
	$\lambda_{24} = 0,71$	-	$\hat{\lambda}_{24} = \mathbf{0,86}$
	$\lambda_{25} = 0,71$	-	$\hat{\lambda}_{25} = \mathbf{0,86}$
	$\lambda_{26} = 0,71$	-	$\hat{\lambda}_{26} = \mathbf{0,59}$
	$\lambda_{\varepsilon_1 1} = 0,71$	-	$\hat{\lambda}_{\varepsilon_1 1} = 0,51$
	$\lambda_{\varepsilon_2 2} = 0,71$	-	$\hat{\lambda}_{\varepsilon_2 2} = 0,51$
	$\lambda_{\varepsilon_3 3} = 0,71$	-	$\hat{\lambda}_{\varepsilon_3 3} = 0,81$
	$\lambda_{\varepsilon_4 4} = 0,71$	-	$\hat{\lambda}_{\varepsilon_4 4} = 0,51$
	$\lambda_{\varepsilon_5 5} = 0,71$	-	$\hat{\lambda}_{\varepsilon_5 5} = 0,51$
	$\lambda_{\varepsilon_6 6} = 0,71$	-	$\hat{\lambda}_{\varepsilon_6 6} = 0,81$
	$r_{\varepsilon_1,\varepsilon_2} = 0,5$	$r_{\varepsilon_1,\varepsilon_2} = 0$	-
	$r_{\varepsilon_4,\varepsilon_5} = 0,5$	$r_{\varepsilon_4,\varepsilon_5} = 0$	-

Grafische Darstellung

Abbildung 36: Fehlspezifiziertes Modell 6

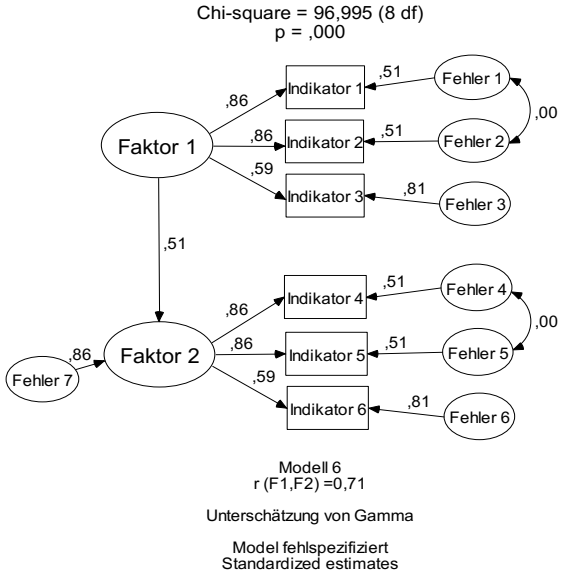
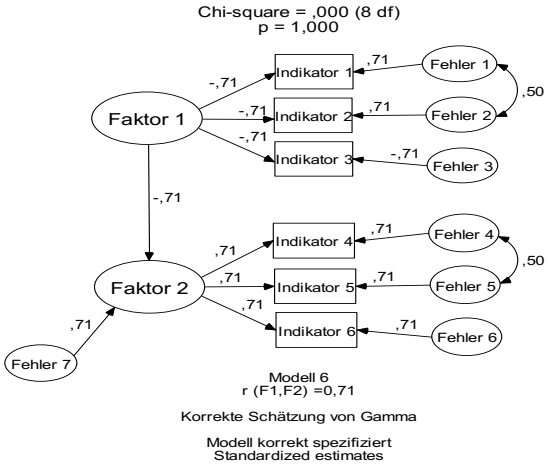


Abbildung 37: Korrekt spezifiziertes Modell 6



Negative Vorzeichen der Pfadkoeffizienten: siehe 3.5.2, Ergebnisse.

Interpretation von Abbildung 36

Die Fehlspezifikationen der Messmodelle für Faktor 1 und Faktor 2 haben folgende Konsequenzen:

1. Die Pfadkoeffizienten der beiden fehlerkorrelierten Indikatorvariablen 1 und 2 (λ_{11} bzw. λ_{12}) und der beiden fehlerkorrelierten Indikatorvariablen 4 und 5 (λ_{24} bzw. λ_{25}) werden überschätzt, die Pfadkoeffizienten der dritten und sechsten Indikatorvariablen (λ_{13} bzw. λ_{26}) werden unterschätzt.

2. Die Güte des Messmodells für Faktor 1 wird ebenso überschätzt wie die Güte des Messmodells für Faktor 2.

3. Die Überschätzungen der beiden Messmodelle führen zu einer Unterschätzung des Pfadkoeffizienten des Strukturmodells ($\hat{\gamma} = 0{,}51$). Die Abweichung von der korrekten Schätzung des Pfadkoeffizienten ist mit 0,71 - 0,51 = 0,20 fast doppelt so groß wie die Abweichung von dem korrekten Wert in Modell 5 (0,11). Das bedeutet, dass sich die Wirkungen der Fehlspezi-

fikationen in den beiden Messmodellen hinsichtlich des Bias der Pfadkoeffizientenschätzung des Strukturmodells nahezu addieren.

4.2.4 MODELL 7

Die Fehlerkorrelationsstruktur des siebenten Modells gleicht der des sechsten.

Die Besonderheit dieses Modells besteht darin, dass durch *Schätzung* einer Fehlertermkorrelation in AMOS die Fehlspezifikation im ersten Messmodell des Modells 6 korrigiert wurde, während die im zweiten Messmodell bestehen blieb.

Beschreibung

	In SPSS gesetzt	In AMOS gesetzt	In AMOS geschätzt
Strukturmodell	$\gamma = 0,71$	-	$\hat{\gamma} = \mathbf{0,60}$
Messmodelle	$\lambda_{11} = 0,71$	-	$\hat{\lambda}_{11} = 0,71$
	$\lambda_{12} = 0,71$	-	$\hat{\lambda}_{12} = 0,71$
	$\lambda_{13} = 0,71$	-	$\hat{\lambda}_{13} = 0,71$
	$\lambda_{24} = 0,71$	-	$\hat{\lambda}_{24} =$ **0,86**
	$\lambda_{25} = 0,71$	-	$\hat{\lambda}_{25} =$ **0,86**
	$\lambda_{26} = 0,71$	-	$\hat{\lambda}_{26} =$ **0,59**
	$\lambda_{\varepsilon_1 1} = 0,71$	-	$\hat{\lambda}_{\varepsilon_1 1} = 0,71$
	$\lambda_{\varepsilon_2 2} = 0,71$	-	$\hat{\lambda}_{\varepsilon_2 2} = 0,71$
	$\lambda_{\varepsilon_3 3} = 0,71$	-	$\hat{\lambda}_{\varepsilon_3 3} = 0,71$
	$\lambda_{\varepsilon_4 4} = 0,71$	-	$\hat{\lambda}_{\varepsilon_4 4} = 0,51$
	$\lambda_{\varepsilon_5 5} = 0,71$	-	$\hat{\lambda}_{\varepsilon_5 5} = 0,51$
	$\lambda_{\varepsilon_6 6} = 0,71$	-	$\hat{\lambda}_{\varepsilon_6 6} = 0,80$
	$r_{\varepsilon_1,\varepsilon_2} = 0,5$	-	$\hat{r}_{\varepsilon_1,\varepsilon_2} = 0,5$
	$r_{\varepsilon_4,\varepsilon_5} = 0,5$	$r_{\varepsilon_4,\varepsilon_5} = 0$	-

Grafische Darstellung

Abbildung 38: Fehlspezifiziertes Modell 7

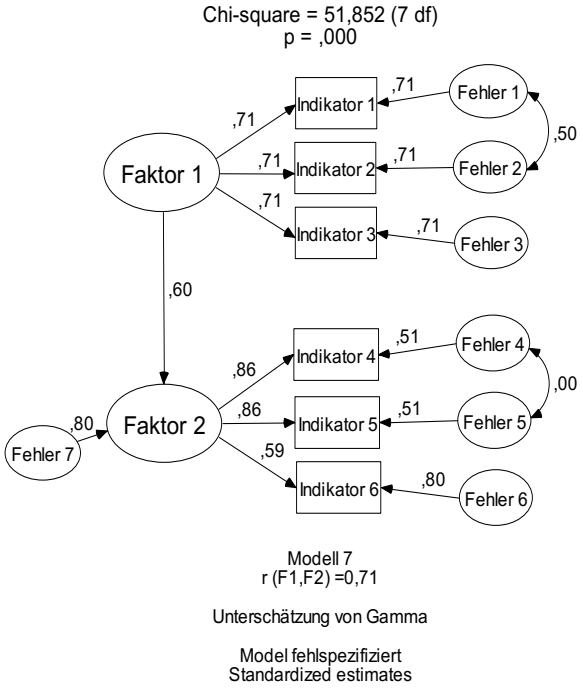
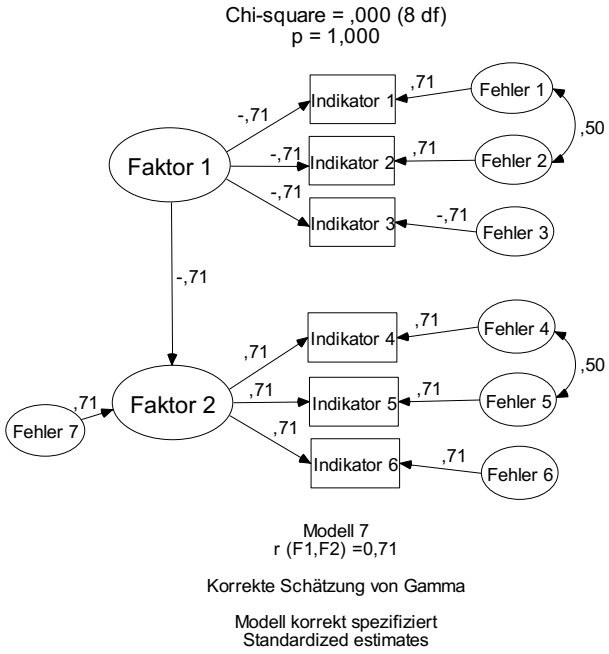


Abbildung 39: Korrekt spezifiziertes Modell 7



Negative Vorzeichen der Pfadkoeffizienten: siehe 3.5.2, Ergebnisse.

Interpretation von Abbildung 38

Die Spezifikationen der Messmodelle für Faktor 1 und Faktor 2 haben folgende Konsequenzen:

1. Die Pfadkoeffizienten der beiden fehlerkorrelierten Indikatorvariablen 1 und 2 werden (im Gegensatz zu Modell 5) nicht mehr überschätzt, da die in SPSS erzeugte Fehlerkorrelation auch in AMOS gesetzt ist. Der Pfadkoeffizient der dritten Indikatorvariablen wird entsprechend nicht mehr unterschätzt.

Die Pfadkoeffizienten der beiden fehlerkorrelierten Indikatorvariablen 4 und 5 werden über- und der Pfadkoeffizient der sechsten Indikatorvariablen unterschätzt.

2. Die Güte des Messmodells für Faktor 1 wird korrekt geschätzt und die Güte des Messmodells für Faktor 2 überschätzt.

3. Die Überschätzungen des zweiten Messmodells führen zu einer Unterschätzung des Pfadkoeffizienten des Strukturmodells. Diese Unterschätzung ist genau so groß wie die in Modell 5. Das bedeutet, dass eine Fehlspezifikation im ersten Messmodell (siehe Modell 5) auf den

Bias der Pfadkoeffizientenschätzung des Strukturmodells die gleiche Wirkung hat wie eine Fehlspezifikation derselben Art im zweiten Messmodell (siehe Modell 7).

4. Die Fehlerkorrelation zweier Indikatoren eines Messmodells führt nur dann zu einem Bias des Pfadkoeffizienten des Strukturmodells, wenn sie in AMOS ignoriert wird.

4.2.5 MODELL 8

Dieses Modell entspricht hinsichtlich seiner Fehlerkorrelationsstruktur dem Modell 2. Es weist eine Korrelation zwischen einem Indikator des ersten und einem Indikator des zweiten Messmodells auf.

Beschreibung

	In SPSS gesetzt	In AMOS gesetzt	In AMOS geschätzt
Strukturmodell	$\gamma = 0,71$	-	$\hat{\gamma} =$ **0,79**
Messmodelle	$\lambda_{11} = 0,71$	-	$\hat{\lambda}_{11} =$ **0,78**
	$\lambda_{12} = 0,71$	-	$\hat{\lambda}_{12} =$ **0,66**
	$\lambda_{13} = 0,71$	-	$\hat{\lambda}_{13} =$ **0,66**
	$\lambda_{24} = 0,71$	-	$\hat{\lambda}_{24} =$ **0,78**
	$\lambda_{25} = 0,71$	-	$\hat{\lambda}_{25} =$ **0,66**
	$\lambda_{26} = 0,71$	-	$\hat{\lambda}_{26} =$ **0,66**
	$\lambda_{\varepsilon_1 1} = 0,71$	-	$\hat{\lambda}_{\varepsilon_1 1} = 0,63$
	$\lambda_{\varepsilon_2 2} = 0,71$	-	$\hat{\lambda}_{\varepsilon_2 2} = 0,75$
	$\lambda_{\varepsilon_3 3} = 0,71$	-	$\hat{\lambda}_{\varepsilon_3 3} = 0,75$
	$\lambda_{\varepsilon_4 4} = 0,71$	-	$\hat{\lambda}_{\varepsilon_4 4} = 0,63$
	$\lambda_{\varepsilon_5 5} = 0,71$	-	$\hat{\lambda}_{\varepsilon_5 5} = 0,75$
	$\lambda_{\varepsilon_6 6} = 0,71$	-	$\hat{\lambda}_{\varepsilon_6 6} = 0,75$
	$r_{\varepsilon_1,\varepsilon_4} = 0,5$	$r_{\varepsilon_1,\varepsilon_4} = 0$	-

Grafische Darstellung

Abbildung 40: Fehlspezifiziertes Modell 8

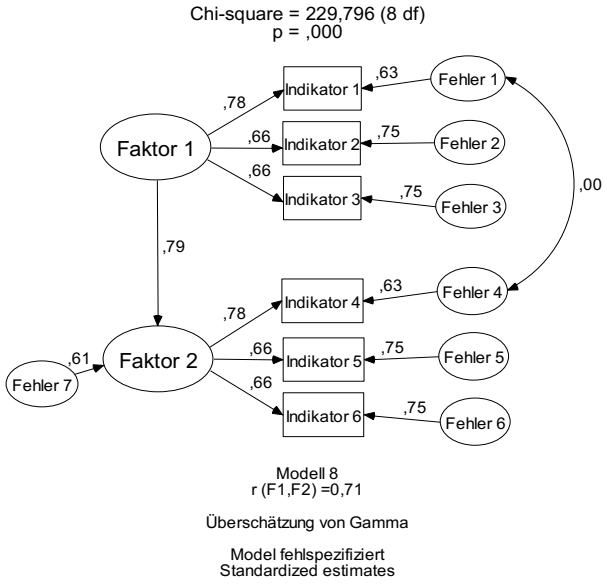
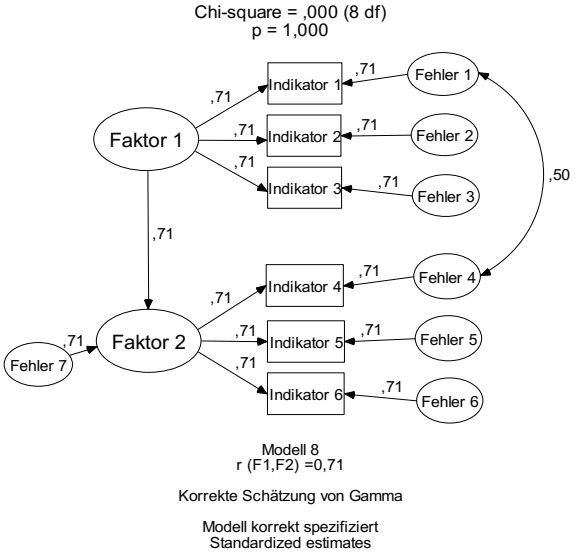


Abbildung 41: Korrekt spezifiziertes Modell 8



Interpretation von Abbildung 40

Die Fehlspezifikation in Form einer Ignorierung der Korrelation zwischen Fehler 1 und 4 betrifft beide Messmodelle. Einige Konsequenzen der Fehlspezifikation decken sich mit jenen, die in Modell 2 beobachtet wurden. Weitere Konsequenzen sind jedoch von den in Modell 2 beobachteten verschieden:

1. In Abweichung von Modell 2 werden die Pfadkoeffizienten der beiden fehlerkorrelierten Indikatoren 1 und 4 mit 0,78 höher geschätzt als die tatsächlich in SPSS gesetzte Korrelation von 0,71. Die Pfadkoeffizienten der vier nicht fehlerkorrelierten Indikatoren werden mit 0,66 leicht unterschätzt.

2. Die Güte beider Messmodelle wird leicht unterschätzt ($0{,}78^2 + 0{,}66^2 + 0{,}66^2 = 1{,}48$).

3. Die Korrelation der Fehler 1 und 4 wird aufgrund der Restriktion, dass diese Fehlerkorrelation 0 sei, als eine Korrelation zwischen den Faktoren interpretiert. Der Pfadkoeffizient des Strukturmodells wird daher überschätzt. Die Überschätzung ist mit 0,79 - 0,71 = 0,08 etwas größer als im Modell 2 (0,06).

4. Der Unterschied zwischen Modell 2 und 8 geht darauf zurück, dass im Gegensatz zu Modell 2 die Faktoren 1 und 2 und somit auch die Indikatoren 1 und 4 miteinander korreliert sind.

Im Fall der unabhängigen Faktoren und damit auch unabhängigen Messmodelle des Modells 2 beeinflusst eine ignorierte Fehlerkorrelation von Indikatoren, die verschiedenen Messmodellen angehören, nicht die Schätzung der Pfadkoeffizíenten λ_{11} und λ_{24} der Messmodelle. Infolge ihrer Ignorierung wird die Fehlertermkorrelation nur dem Pfadkoeffizienten γ des Strukturmodells zugeschrieben.

Im Fall der voneinander abhängigen Faktoren und damit auch abhängigen Messmodelle des Modells 8 wirkt sich die Ignorierung die Fehlertermkorrelation erhöhend sowohl auf den Pfadkoeffizienten γ des Strukturmodells als auch auf die Pfadkoeffizíenten λ_{11} und λ_{24} der Messmodelle aus.

4.2.6 MODELL 9

Dieses Modell entspricht hinsichtlich seiner Fehlerkorrelationsstruktur dem Modell 3. Die Fehler des ersten und zweiten Indikators sowie die Fehler des ersten und vierten Indikators sind korreliert.

Beschreibung

	In SPSS gesetzt	In AMOS gesetzt	In AMOS geschätzt
Strukturmodell	$\gamma = 0,71$	-	$\hat{\gamma} =$ **0,70**
Messmodelle	$\lambda_{11} = 0,71$	-	$\hat{\lambda}_{11} =$ **0,93**
	$\lambda_{12} = 0,71$	-	$\hat{\lambda}_{12} =$ **0,80**
	$\lambda_{13} = 0,71$	-	$\hat{\lambda}_{13} =$ **0,57**
	$\lambda_{24} = 0,71$	-	$\hat{\lambda}_{24} =$ **0,81**
	$\lambda_{25} = 0,71$	-	$\hat{\lambda}_{25} =$ **0,64**
	$\lambda_{26} = 0,71$	-	$\hat{\lambda}_{26} =$ **0,64**
	$\lambda_{\varepsilon_1 1} = 0,71$	-	$\hat{\lambda}_{\varepsilon_1 1} = 0,38$
	$\lambda_{\varepsilon_2 2} = 0,71$	-	$\hat{\lambda}_{\varepsilon_2 2} = 0,59$
	$\lambda_{\varepsilon_3 3} = 0,71$	-	$\hat{\lambda}_{\varepsilon_3 3} = 0,82$
	$\lambda_{\varepsilon_4 4} = 0,71$	-	$\hat{\lambda}_{\varepsilon_4 4} = 0,59$
	$\lambda_{\varepsilon_5 5} = 0,71$	-	$\hat{\lambda}_{\varepsilon_5 5} = 0,76$
	$\lambda_{\varepsilon_6 6} = 0,71$	-	$\hat{\lambda}_{\varepsilon_6 6} = 0,76$
	$r_{\varepsilon_1,\varepsilon_2} = 0,5$	$r_{\varepsilon_1,\varepsilon_2} = 0$	-
	$r_{\varepsilon_1,\varepsilon_4} = 0,5$	$r_{\varepsilon_1,\varepsilon_4} = 0$	-

Grafische Darstellung

Abbildung 42: Fehlspezifiziertes Modell 9

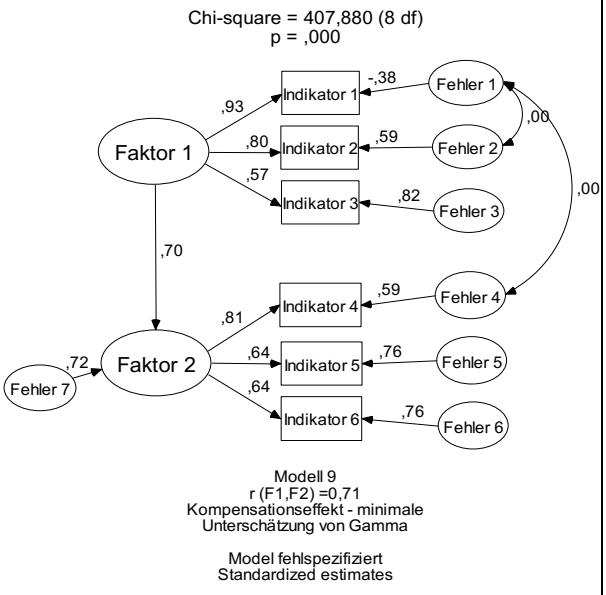
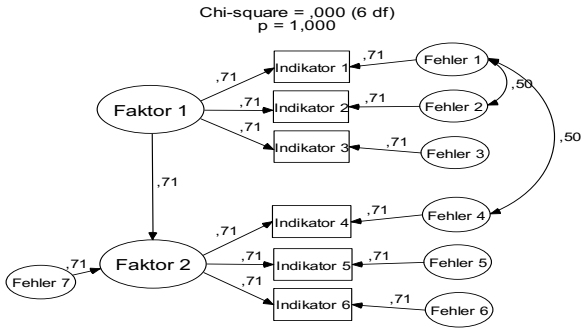


Abbildung 43: Korrekt spezifiziertes Modell 9



Interpretation von Abbildung 42

1. Die Pfadkoeffizienten der beiden fehlerkorrelierten Indikatoren 1 und 2 werden infolge der Fehlspezifikation (Ignorieren von $r_{\varepsilon_1,\varepsilon_2} = 0{,}5$) überschätzt, der Pfadkoeffizient der dritten Variablen wird unterschätzt.

2. Daher wird die Güte des Messmodells für Faktor 1 mit $0{,}93^2 + 0{,}80^2 + 0{,}57^2 = 1{,}83$ überschätzt – im gleichen Ausmaß wie in den Modellen 3 und 5. Diese Überschätzung der Messmodellgüte würde ohne weitere Fehlspezifikationen zur Unterschätzung des Pfadkoeffizienten des Strukturmodells $\hat{\gamma} = 0{,}60$ führen (siehe Modell 5).

3. Eine weitere Fehlspezifikation (Ignorieren von $r_{\varepsilon_1,\varepsilon_4} = 0{,}5$) führt zu einer minimalen Unterschätzung der Güte des zweiten Messmodells ($0{,}81^2 + 0{,}64^2 + 0{,}64^2 = 1{,}48$) – im gleichen Ausmaß wie in den Modellen 3 und 8. Diese Unterschätzung der Messmodellgüte würde ohne die erste Fehlspezifikationen zur Überschätzung des Pfadkoeffizienten des Strukturmodells $\hat{\gamma} = 0{,}79$ führen (siehe Modell 8).

4. Die Kombination beider Fehlspezifikationen führt dazu, dass einander entgegengesetzte Effekte auf die Schätzung von gamma sich nahezu aufheben. Der Unterschätzungseffekt aus Modell 5 (0,60) und der Überschätzungseffekt aus Modell 8 (0,79) münden in eine nahezu korrekte Schätzung von 0,70=(0,60 + 0,79)/2.

5. *Es lässt sich folgende Erkenntnis ableiten: Die Tatsache, dass der Pfadkoeffizient eines Strukturmodells korrekt geschätzt wird, erlaubt nicht den Rückschluss auf eine korrekte Messmodellfehlerspezifikation. Die Wirkungen mehrerer Fehlspezifikationen können einander aufheben.*

4.2.7 ZUSAMMENFASSUNG

Tabelle 5: Tatsächlich erzeugte, vor der Modellschätzung festgelegte und geschätzte Korrelationen der Ein-Faktor-Modelle mit drei Indikatoren (Übersicht)
a) Faktoren des Strukturmodells sind unabhängig ($\gamma = 0$)

Modell-nummer	In SPSS gesetzt	In AMOS gesetzt	$\hat{\gamma}$	Anpassungsgüte (Chi-Quadrat; df; p-Wert)	Interpretation des Effektes auf $\hat{\gamma}$
1	Alle Fehler unkorreliert	Alle Fehler unabhängig	**0**	0,000; (8df) p=1,000	----------
2	$r_{\varepsilon_1,\varepsilon_4} = 0{,}5$	$r_{\varepsilon_1,\varepsilon_4} = 0$	**0,06**	227 (8df) p= 0,000	Ignorierung einer die Messmodelle übergreifenden Fehlerkorrelation führt zu *leichter Überschätzung*
3	$r_{\varepsilon_1,\varepsilon_2} = 0{,}5$ $r_{\varepsilon_1,\varepsilon_4} = 0{,}5$	$r_{\varepsilon_1,\varepsilon_2} = 0$ $r_{\varepsilon_1,\varepsilon_4} = 0$	**0,07**	381,8 (8fd) p= 0,000	Ignorierung einer die Messmodelle übergreifenden Fehlerkorrelation (siehe Mod.2) und zugleich einer Korrelation zwischen Fehlern des ersten Messmodells führt *auch nur zu leichter Überschätzung*
Z1	$r_{\varepsilon_1,\varepsilon_2} = 0{,}5$	$r_{\varepsilon_1,\varepsilon_2} = 0$	**0**	0,000; (8df) p=1,000	Ignorierung einer Korrelation zwischen 2 Fehlern des ersten Messmodells führt zu keinem Bias
Z2	$r_{\varepsilon_1,\varepsilon_2} = 0{,}5$ $r_{\varepsilon_4,\varepsilon_5} = 0{,}5$	$r_{\varepsilon_1,\varepsilon_2} = 0$ $r_{\varepsilon_4,\varepsilon_5} = 0$	**0**	0,000; (8df) p=1,000	Ignorierung einer Korrelation zwischen 2 Fehlern des ersten und zugleich 2 Fehlern des zweiten Messmodells führen zu keinem Bias
Z3	$r_{\varepsilon_1,\varepsilon_2} = 0{,}5$ $r_{\varepsilon_4,\varepsilon_5} = 0{,}5$	$r_{\varepsilon_1,\varepsilon_2} = 0{,}5$ $r_{\varepsilon_4,\varepsilon_5} = 0$	**0**	0,000; (8df) p=1,000	Korrekte Spezifizierung einer Fehlerkorrelation innerhalb des ersten Messmodells und Ignorierung einer Fehlerkorrelation innerhalb des zweiten führen zu keinem Bias

b) Faktoren des Strukturmodells sind abhängig ($\gamma = 0,71$)

Modell-nummer	In SPSS gesetzt	In AMOS gesetzt	$\hat{\gamma}$	Anpassungs-güte (Chi-Quadrat; df; p-Wert)	Interpretation des Effektes auf $\hat{\gamma}$
4	Alle Fehler unkorreliert	Alle Fehler unabhängig	**0,71**	0,000 (8df) p=1,000	----------
5	$r_{\varepsilon_1,\varepsilon_2} = 0{,}5$	$r_{\varepsilon_1,\varepsilon_2} = 0$	**0,60**	56,6 (8df) p= 0,000	Ignorierung einer Fehlerkorrelation innerhalb des ersten Messmodells führt zu *Unterschätzung*
Z 4	$r_{\varepsilon_4,\varepsilon_5} = 0{,}5$	$r_{\varepsilon_4,\varepsilon_5} = 0$	**0,60**	56,6 (8df) p= 0,000	Ignorierung einer Fehlerkorrelation innerhalb des zweiten Messmodells führt zu *Unterschätzung in selber Höhe* wie Ignorierung einer Fehlerkorrelation innerhalb des ersten
6	$r_{\varepsilon_1,\varepsilon_2} = 0{,}5$ $r_{\varepsilon_4,\varepsilon_5} = 0{,}5$	$r_{\varepsilon_1,\varepsilon_2} = 0$ $r_{\varepsilon_4,\varepsilon_5} = 0$	**0,51**	97,0 (8df) p=0,000	Ignorierung einer Fehlerkorrelationen innerhalb *beider* Messmodelle führt zu *doppelt so großer Unterschätzung*
7	$r_{\varepsilon_1,\varepsilon_2} = 0{,}5$ $r_{\varepsilon_4,\varepsilon_5} = 0{,}5$	$r_{\varepsilon_4,\varepsilon_5} = 0$	**0,60**	51,8 (7df) p=0,000	Ignorierung einer Fehlerkorrelation innerhalb des einen Messmodells und Schätzung einer Fehlerkorrelation innerhalb des anderen führt zu *Unterschätzung, die genauso groß ist wie in Modell 5*
8	$r_{\varepsilon_1,\varepsilon_4} = 0{,}5$	$r_{\varepsilon_1,\varepsilon_4} = 0$	**0,79**	229,8 (8df) p=0,000	Ignorierung einer die Messmodelle übergreifenden Fehlerkorrelation führt zu *Überschätzung*
9	$r_{\varepsilon_1,\varepsilon_2} = 0{,}5$ $r_{\varepsilon_1,\varepsilon_4} = 0{,}5$	$r_{\varepsilon_1,\varepsilon_2} = 0$ $r_{\varepsilon_1,\varepsilon_4} = 0$	**0,70**	407,9 (8df) p=0,000	Ignorierung einer die Messmodelle übergreifenden Fehlerkorrelation (Überschätzung) und Ignorierung einer Fehlerkorrelation innerhalb eines Messmodells (Unterschätzung) heben sich in der Wirkung fast auf. Der Schluss von vernachlässigbarem Bias ($\hat{\gamma} = 0,70$) auf gute Anpassung wäre falsch.

5 Simulationen zum Ein-Faktor-Modell mit drei Indikatoren unter Verwendung von PROC CALIS (SAS)

5.1 EINLEITUNG

Ausgangspunkt der Simulationen ist die mit AMOS durchgeführte Analyse des Ein-Faktor-Modells mit drei Indikatoren. Mit AMOS wurden Daten analysiert, die zuvor in SPSS künstlich erzeugt worden waren. Zur Konstruktion manifester Variabler waren hierbei orthogonale Faktoren verwendet worden, die aus einer zu diesem Zweck durchgeführten Hauptkomponentenanalyse stammten. Diese orthogonalen Zufallsvariablen sind approximativ standardnormalverteilt und zeichnen sich insbesondere dadurch aus, vollkommen unabhängig voneinander zu sein. Auf diese Weise war es möglich gewesen, die Korrelationsstruktur der konstruierten manifesten Variablen exakt zu kontrollieren.

Es wurde in AMOS demonstriert, welche Werte für interessierende Pfadkoeffizienten bei exakt kontrollierter Korrelationsmatrix der manifesten Modellvariablen unter gegebenen Modellannahmen errechnet werden (siehe hierzu auch Tabelle 4). Als Basis aller in AMOS betrachteten Ein-Faktor-Modelle diente eine einzige, aus fünf orthogonalen Zufallsvariablen (elementare Zufallsvariablen) bestehende Datendatei mit 1517 Realisationen. Die Analyse wurde also auf Basis *einer einzigen* „Stichprobe“ durchgeführt.

Die Analyse in AMOS wird zunächst unter Verwendung der Software SAS mit der Prozedur PROC CALIS *reproduziert*, die später auch zur Simulation eingesetzt werden soll (siehe unten). Hierbei werden auf Basis der o. g. Datendatei unter identischen Modellannahmen Werte derselben interessierenden Pfadkoeffizienten errechnet. Die Reproduktion soll sicherstellen, dass bei Verwendung von AMOS und PROC CALIS unter identischen Randbedingungen identische Resultate erzielt werden.

Bei den *Simulationen* zum Ein-Faktor-Modell wird nun die Schätzung von Pfadkoeffizienten in einer *Anzahl von n Stichproben* analysiert, von welchen eine jede Stichprobe aus Zufallsvariablen mit jeweils 1000 Realisationen besteht. Hauptsächlich wird hierbei

a) zum einen untersucht, welchen Mittelwert – als Durchschnitt der geschätzten Pfadkoeffizienten aller n Stichproben – ein interessierender Pfadkoeffizient aufweist. Die Untersuchung gibt Auskunft darüber, ob der Schätzer „Mittelwert der Pfadkoeffizienten aus n Stichproben“ dem wahren Wert entspricht.

b) Zum anderen wird analysiert, ob der durchschnittliche Bias (durchschnittliche Abweichung vom wahren Wert) eines interessierenden Pfadkoeffizienten, der aus den n Stichproben eines bestimmten fehlspezifizierten Modells errechnet wird, gleich dem im entsprechenden AMOS-Modell errechneten Bias ist.

c) Schließlich wird der *Mittelwert aus* den in den n Stichproben *geschätzten Standardfehlern* eines interessierenden Pfadkoeffizienten verglichen mit der *Streuung des* in den n Stichproben ermittelten interessierenden *Pfadkoeffizienten* („empirischer" Standardfehler). Dieser Vergleich gibt darüber Auskunft, ob der empirische Standardfehler des Pfadkoeffizienten korrekt geschätzt wird.

Der empirische Standardfehler des Pfadkoeffizienten ist von Interesse, da auf seiner Basis das empirische Konfidenzintervall um den Schätzer „Mittelwert des Pfadkoeffizienten aus n Stichproben" ermittelt wird.

Die Simulationen zu dem Ein-Faktor-Modell können nicht mit AMOS durchgeführt werden, da die Anwendungsmöglichkeiten von SPSS eine Simulation innerhalb des Programmpaketes AMOS nicht vorsehen. Sie werden daher mit der Software SAS und der dem Programmpaket AMOS entsprechenden Prozedur PROC CALIS betrieben.

Bei der Simulation der n Stichproben in SAS werden – ebenso wie in SPSS – approximativ standardnormalverteilte Zufallsvariable erzeugt, aus welchen wiederum manifeste Variable konstruiert werden (siehe 11.3.1). Die Zufallsvariablen sind jedoch im Gegensatz zu den Zufallsvariablen der in SPSS erzeugten Stichprobe nicht perfekt unabhängig voneinander, sondern aufgrund von Zufallsschwankungen im Allgemeinen untereinander schwach korreliert. Die n simulierten Stichproben weichen also voneinander zufallsbedingt ab hinsichtlich des Mittelwertes, der Varianz und der Korrelationsstruktur der elementaren Zufallsvariablen.

Das Vorgehen bei der Konstruktion der manifesten Modellvariablen aus elementaren Zufallsvariablen in SAS entspricht – bis auf den Unterschied hinsichtlich der Orthogonalität – exakt dem Verfahren in SPSS.

In der SAS-Prozedur PROC CALIS stehen mehrere Möglichkeiten zur Verfügung, ein und dasselbe Modell zu spezifizieren. Die unterschiedlichen Modellspezifikationen gehen auf unterschiedliche Traditionen der Entwicklung von Kovarianzstrukturmodellen zurück. Eine be-

stimmte Modellspezifikationen wird in PROC CALIS über ein bestimmtes *Statement* vorgenommen.[42]

Die Anzahl der simulierten Stichproben bei der Analyse des Ein-Faktor-Modells mit drei Indikatoren beträgt in der Regel 1000**.** In besonderen Fällen (Auftreten von Ausreißern) wird eine größere Anzahl von Stichproben simuliert, um eine genauere Analyse zu ermöglichen.

5.2 REPRODUKTION DER AMOS-ANALYSE MIT PROC CALIS

Die Pfadkoeffizientenwerte aller AMOS-Modelle konnten unter Verwendung von PROC CALIS – gerundet auf drei Dezimalstellen – exakt reproduziert werden. Dies betrifft sowohl jeweils die richtig spezifizierten als auch die fehlspezifizierten Modellvarianten.

5.3 SIMULATION

5.3.1 DAS GRUNDMODELL

Es wird eine Simulation zu dem AMOS-Modell durchgeführt, dessen drei manifeste Variablen (Indikatoren) nur insofern untereinander korrelieren, als sie von einem Faktor gleichermaßen beeinflusst werden (r=0,5). Wie in allen anderen im Folgenden analysierten Modellen weisen sie, ebenso wie der Faktor, eine Varianz von 1 auf. Die Fehlerterme sind untereinander tatsächlich unkorreliert und es werden im Modell entsprechend auch keine Fehlerkorrelationen angenommen.

Die Erzeugung der Indikatoren dieses Modells sowie aller weiteren in diesem Kapitel beschriebenen Modelle aus den o. g. orthogonalen Zufallsvariablen wurde bereits im Anhang zu den AMOS-Modellen detailliert beschrieben und wird in 11.3.1 wiedergegeben.

Innerhalb der PROC CALIS Prozedur genügen im Falle des Grundmodells das LINEQS- und das STD-Statement zur Angabe der Modellgleichungen bzw. Fixierung der Varianzen. (Das

[42] Da in der SAS-Prozedur PROC CALIS eine grafische Modellkonstruktion nicht möglich ist, wurde die PROC CALIS-Modellspezifikation zur Nachbildung der AMOS-Modelle gewählt, die bei Vorliegen eines Sets von Strukturgleichungen angezeigt ist. Diese Modellspezifikation wird über das sogenannte LINEQS-Statement in Gleichungsform vorgenommen, der ein STD- und optional ein COV-Statement folgt. (Die STD- und COV-Statements legen fest, welche Varianzen bzw. Kovarianzen im Modell geschätzt werden sollen und welche fixiert sind.) Das LINEQS-Statement ähnelt jenem, das ursprünglich von P. M. Bentler 1985 entwickelt wurde. Siehe SAS Institute Inc (1999), The CALIS Procedure, S. 438. bzw. Bentler, P. M. (1985).
Die Entscheidung für das LINEQS-Statement erfolgte aus technischem Grund. Die alternativ in Frage kommenden Statements in Matrix- oder in Listenschreibweise, das COSAN- bzw. das RAM-Statement, würden im Wesentlichen zu denselben Ergebnissen geführt haben (siehe Institute Inc.,1999, S. 595). COSAN- und RAM-Statement kommen ohne ein STD- oder COV-Statement aus – letztere sind nur im Zusammenhang mit dem LINEQS-Statement möglich.

COV-Statement erfolgt nicht, da bezüglich der Fehlerkorrelation keine von der Unabhängigkeit abweichende Modellannahme gemacht wird.)

```
Proc Calis data = * ...;
Lineqs
indik_1 = beta1   Faktor_1          + beta_f1          Error_1,
indik_2 = beta2   Faktor_1          + beta_f2          Error_2,
indik_3 = beta3   Faktor_1          + beta_f3          Error_3;
Std
        Error_1 = 1,
        Error_2 = 1,
        Error_3 = 1,
        Faktor_1= 1;
Run;
```

In der AMOS-Analyse werden für die Faktorpfadkoeffizienten beta (Koeffizienten der drei Pfade, die vom Faktor zu den drei Indikatoren führen) identische Werte (0,707) errechnet. Dieser Wert kann als der *wahre Wert* bezeichnet werden, anhand dessen überprüft werden kann, ob die Schätzer „arithmetisches Mittel von beta1“, „arithmetisches Mittel von beta2“ und „arithmetisches Mittel von beta3“ erwartungstreu sind. Für die Fehlerpfadkoeffizienten beta_f (Koeffizienten der Pfade, die von den Fehlervariablen zu den manifesten Variablen führen), werden identische Werte in der AMOS Analyse errechnet.

Das arithmetische Mittel eines *Faktorpfadkoeffizienten* beta wird im Grund- sowie in allen Folgemodellen nicht aus dem in den n Stichproben ermittelten Wert für beta errechnet, sondern aus dem Betrag des ermittelten Werts. Diese Maßnahme erfolgt, da in einer nicht geringen Anzahl von Stichproben alle drei interessierenden Koeffizienten mit negativem Vorzeichen geschätzt werden (siehe 3.5.2). Das arithmetische Mittel eines *Fehlerpfadkoeffizienten* beta_f wird aus den in den n Stichproben ermittelten Werten für beta_f errechnet. Eine Berechnung aus den Beträgen der Werte ist nicht erforderlich, da das o. g. Phänomen in der Regel nicht auftritt.

Ergebnisse

Der Schätzer „arithmetisches Mittel des Betrags von beta1“ – abgekürzt im folgenden „Mittelwert von beta1“ – beträgt 0,7067, der Mittelwert von beta2 0,7072 und der Mittelwert von beta3 0,7066. Alle drei Schätzer für die *Faktor*pfadkoeffizienten sind somit erwartungstreu.

Die *empirischen* Standardfehler von beta1, beta2 und beta3 betragen 0,0244, 0,0258 bzw. 0,0262. Als Mittelwerte der *geschätzten* Standardfehler ergeben sich für beta1, beta2 und beta3 jeweils 0,0336. Die Mediane der geschätzten SE sind fast identisch (0,0335, 0,0336 und 0,0336). Mithin wird der empirischen Standardfehler von beta1 um den Faktor 0,0366 /

0,0244 ~ 1,38 überschätzt, der empirische Standardfehler von beta2 um den Faktor 0,0336 / 0,0258 ~ 1,30 und der empirische Standardfehler von beta3 um den Faktor 0,0336 / 0,0262 ~ 1,28 zu hoch geschätzt. Da die empirischen Standardfehler durch die geschätzten Standardfehler der Pfadkoeffizienten also um ungefähr ein Drittel überschätzt werden, kann man von einer deutlich *konservativen Schätzung* sprechen.[43]

Auch die Mittelwerte der drei Fehlerpfadkoeffizienten, beta_f1, beta_f2 und beta_f3, sind erwartungstreue Schätzungen des wahren Wertes der Fehlerpfadkoeffizienten. Die mittleren geschätzten Standardfehler der drei Fehlerpfadkoeffizienten schätzen die jeweils entsprechenden empirischen Standardfehler der Fehlerpfadkoeffizienten recht genau. Der Fehlerpfadkoeffizient der ersten manifesten Variablen (beta_f1) wird um ca. 3 % überschätzt (0,0251 / 0,0244 = 1,028), beta_f2 wird um ca. 3 % (0,0252 / 0,0259 = 0,9730) und beta_f3 um ca. 4 % (0,0251 / 0,0262 = 0,9580) unterschätzt. Diese Differenz zwischen der deutlichen Überschätzung der empirischen Standardfehler der Faktorpfadkoeffizienten und der recht genauen Schätzung der empirischen Standardfehler der Fehlerpfadkoeffizienten trifft auch auf alle weiteren Modelle zu. Die Fehlerpfadkoeffizienten sind in der Analyse jedoch nur von untergeordnetem Interesse, so dass die Untersuchung ihrer Standardfehler ohne große Relevanz ist.

Als Resultat der Analyse wird vorgeschlagen, den geschätzten Standardfehler mit einem Korrekturfaktor von 0,75 zu versehen, um dem unbekannten tatsächlichen Standardfehler nahezukommen.

In der Simulation zum Grundmodell mit 1000 Iterationen und 1000 Realisationen von Zufallsvariablen pro Iteration liegen die Minima der Pfadkoeffizienten beta1, beta2 und beta3 zwischen 0,61 und 0,63 und deren Maxima zwischen 0,78 und 0,79. Im Einzelfall können also selbst bei einer hohen Anzahl von 1000 Realisationen Pfadkoeffizienten geschätzt werden, die erheblich (also um fast 15 %) vom wahren Wert 0,707 abweichen. Unter der plausiblen Annahme normalverteilter Pfadkoeffizienten liegen jedoch ca. 95 % aller Pfadkoeffizienten in einem Bereich von ungefähr $0{,}707 \pm 1{,}96 \cdot 0{,}025$, also zwischen 0,658 und 0,756, so dass 95 % der Pfadkoeffizientenschätzungen höchstens 7 % vom wahren Wert abweichen.

[43] Die Standardabweichung des geschätzten Standardfehlers variiert für beta1, beta2 und beta3 zwischen 0,00077 und 0,00079 und ist damit sehr gering. Es kommt bei einem minimalen geschätzten Standardfehler von 0,0316 (und einem maximalen zwischen 0,0368 und 0,0370) nie vor, dass der empirischen Standardfehler durch den geschätzten Standardfehler unterschätzt wird. Minimal wird der empirische Standardfehler durch den geschätzten um 20 % überschätzt. Die Standardabweichung des geschätzten Standardfehlers soll in der Analyse der folgenden Modellvarianten außer Betracht gelassen werden.

5.3.2 DAS MODELL MIT ZWEI POSITIV KORRELIERTEN FEHLERN

In diesem Modell und allen weiteren Modellen mit einem Faktor und drei manifesten Variablen ist die Korrelationsstruktur zwischen den manifesten Variablen nicht nur, wie beim Grundmodell, durch den Einfluss des Faktors, sondern zusätzlich durch den Einfluss einer oder mehrerer Korrelationen zwischen Fehlervariablen bedingt.

Es werden in diesem (und den weiteren Modellen) stets zwei Modellspezifikationen unterschieden: zum einen ein *korrekt spezifiziertes* Modell, in dem die tatsächliche Fehlerkorrelationsstruktur sich auch in einer entsprechenden Modellannahme niederschlägt und zum anderen ein *fehlspezifiziertes* Modell, in dem die tatsächliche Fehlerkorrelationsstruktur ignoriert wird.

Diese beiden Modellspezifikationen erfordern neben dem LINEQS- und STD- zusätzlich das COV-Statement.

Die drei Statements lauten im korrekt spezifizierten Modell, in welchem die den ersten beiden manifesten Variablen zugehörigen Fehlervariablen mit r = 0,5 korreliert (und alle anderen Fehlervariablen unkorreliert) sind, wie folgt:

```
Lineqs
indi_1ps =      beta1   Faktor_1        + beta_f1       Error_1,
indi_2ps =      beta2   Faktor_1        + beta_f2       Error_2,
indik_3 =       beta3   Faktor_1        + beta_f3       Error_3;

Std
Error_1 = 1,
Error_2 = 1,
Error_3 = 1,
Faktor_1= 1;

Cov
Error_1 Error_2 = 0.5;
```

Ergebnisse

Die arithmetischen Mittel der Pfadkoeffizienten aus den n Stichproben sind, wie im Grundmodell, erwartungstreu und die empirischen Standardfehler von beta1 und beta2 entsprechen jenen im Grundmodell. Der empirische Standardfehler von beta3 hingegen ist mit 0,030 bezüglich des Grundmodells (0,026) deutlich erhöht.

Auch der mittlere geschätzte Standardfehler des dritten Faktorpfadkoeffizienten ist mit 0,038 den anderen gegenüber (beide 0,034) leicht erhöht. Die mittleren geschätzten Standardfehler der ersten beiden Faktorpfadkoeffizienten überschätzen die jeweiligen empirischen Standardfehler um ca. ein Drittel und somit im gleichen Maß wie im Grundmodell. Der mittlere ge-

schätzte Standardfehler des dritten Faktorpfadkoeffizienten überschätzt den empirischen Standardfehler in diesem Modell nur um ca. ein Viertel. (Die Mediane der geschätzten Standardfehler weichen von den arithmetischen Mittelwerten derselben – auf vier Stellen nach dem Komma genau – nicht ab.)

Die tatsächliche, korrekt spezifizierte Korrelation zwischen Fehler 1 und Fehler 2 schlägt sich nicht nur in einer Erhöhung der empirischen und der geschätzten Standardfehler von beta3, sondern auch in einer Vorzeichenumkehr und Erhöhung der *Korrelation der Schätzer der Faktorpfadkoeffizienten* beta1 und beta2 nieder: Die ersten beiden Faktorpfadkoeffizienten sind untereinander schwach positiv korreliert ($r(\hat{\beta}_1, \hat{\beta}_2)$ = 0,33), während sie im Grundmodell sehr schwach negativ korreliert sind ($r(\hat{\beta}_1, \hat{\beta}_2)$ = -0,13). Der dritte Faktorpfadkoeffizient hingegen ist mit den restlichen beiden schwach negativ korreliert (r = -0,27).[44] (Das Grundmodell wies hierbei nur sehr schwach negative Pfadkoeffizientenkorrelationen von maximal -0,15 auf.) Hierbei sind im Grundmodell Unterschiede zwischen $r(\hat{\beta}_1, \hat{\beta}_2)$, $r(\hat{\beta}_1, \hat{\beta}_3)$ und $r(\hat{\beta}_2, \hat{\beta}_3)$ zufällig.

Fazit

Im Modell, in dem die Fehler der ersten beiden manifesten Variablen positiv korreliert sind, gewährleistet die korrekte Spezifikation der Fehlerkorrelation, dass die arithmetischen Mittel der Pfadkoeffizienten erwartungstreue Schätzer der wahren Werte sind. Die Variabilität des von der Fehlerkorrelation nicht betroffenen Faktorpfadkoeffizienten beta3 jedoch nimmt zu. Da die Überschätzung des empirischen Standardfehlers durch den mittleren geschätzten Standardfehler dieses Faktorpfadkoeffizienten geringer ist als bei den anderen beiden, ist der im Grundmodell vorgeschlagene Korrekturfaktor in der Schätzung des Konfidenzintervalls für beta3 von 0,75 auf 0,8 zu erhöhen.

5.3.3 DAS FEHLSPEZIFIZIERTE MODELL

Hinsichtlich des LINEQS- und des STD-Statements unterscheiden sich fehl- und korrekt spezifiziertes Modell nicht. Die Fehlspezifikation der tatsächlichen Fehlerkorrelation $\lambda_{\varepsilon_1,\varepsilon_2}$ = 0,5, erfolgt im COV-Statement:

```
Cov
Error_1 Error_2 = 0;
```

[44] Die Korrelationsstruktur der Faktorpfadkoeffizienten wird in der Analyse nur dann erwähnt, wenn Standardfehler bestimmter Koeffizienten bezüglich des Grundmodells erhöht oder erniedrigt sind und daher zu weiteren Nachforschungen Anlass geben.

In der AMOS-Analyse werden für die Faktorpfadkoeffizienten beta1 und beta2 identische Werte von 0,866 errechnet. Für beta3 beträgt der Wert 0,577. Für beta1 und beta2 ist also ein identischer Bias (Überschätzung um ca. 22 %) zu verzeichnen, während beta3 um ca. 18 % unterschätzt wird.

Ergebnisse

Die arithmetischen Mittel der drei Faktorpfadkoeffizienten sind nicht erwartungstreu, insofern wird der für alle drei identische wahre Wert 0,707 deutlich verfehlt. Der durchschnittliche Bias der Koeffizienten entspricht sehr genau dem Bias, der sich in AMOS ergibt.

Die empirischen Standardfehler der Faktorpfadkoeffizienten, deren Fehlerterme korreliert sind (beta1 und beta2), sind bezüglich des korrekt spezifizierten Modells um ca. 30 % vermindert (0,018 bzw. 0,019). Der empirische Standardfehler von beta3 (0,023) entspricht in etwa dem des Grundmodells.

Der Mittelwert der geschätzten Standardfehler ist jedoch – auch im Unterschied zum korrekt spezifizierten Modell – bei allen Faktorpfadkoeffizienten fast gleich hoch (0,030 für die ersten beiden und 0,31 für den dritten) und damit etwas geringer als im Grundmodell. Somit überschätzen die Mittelwerte der geschätzten Standardfehler der Indikatoren, deren Fehlerterme korreliert sind, die jeweiligen empirischen Standardfehler um mehr als 50 %. Der empirische Standardfehler von beta3 wird nur um ca. ein Drittel zu hoch geschätzt und damit im selben Ausmaß wie im Grundmodell. Die Mediane der geschätzten Standardfehler der Faktorpfadkoeffizienten sind von den arithmetischen Mittelwerten derselben – auf vier Dezimalstellen genau – nicht verschieden.

Die Fehlspezifikation der Korrelation zwischen Fehler 1 und Fehler 2 geht einher mit einer mittelstarken negativen Korrelation der geschätzten Faktorpfadkoeffizienten beta1 und beta2 (r = -0,64). Diese gegenseitigen Determination der Koeffizienten beta1 und beta2 ist kaum problematisch, da der Streubereich, in dem sich die Pfadkoeffizienten bewegen, klein ist (0,80 bis 0,93) und sie daher nur zu verhältnismäßig geringen Abweichungen von dem in AMOS errechneten Wert (0,87) beitragen kann. Die restlichen Korrelationen, $r(\hat{\beta}_1, \hat{\beta}_3)$ und $r(\hat{\beta}_2, \hat{\beta}_3)$, sind vernachlässigbar gering.

Fazit

Der in AMOS errechnete, durch die Fehlspezifikation erzeugte Bias der Faktorpfadkoeffizienten wird durch die Simulation im Mittel exakt reproduziert. Die Schätzungen des empirischen

Standardfehlers der ersten beiden Faktorpfadkoeffizienten sind infolge der fehlspezifizierten Fehlerkorrelation sehr konservativ, die Schätzung des empirischen Standardfehlers des dritten Faktorpfadkoeffizienten ist konservativ.

5.3.4 DAS MODELL MIT ZWEI NEGATIV KORRELIERTEN FEHLERTERMEN

Das Modell weist eine negative Korrelation der Fehler der ersten und zweiten manifesten Variablen auf ($\lambda_{\varepsilon_1,\varepsilon_2} = -0.3$)[45]. Die Modellspezifikation des korrekt spezifizierten Modells lautet wie folgt:

```
Lineqs
indi_1ps =      beta1   Faktor_1          + beta_f1          Error_1,
indi_2ns =      beta2   Faktor_1          + beta_f2          Error_2,
indik_3 =       beta3   Faktor_1          + beta_f3          Error_3;

Std
Error_1 = 1,
Error_2 = 1,
Error_3 = 1,
Faktor_1= 1;

Cov
Error_1 Error_2 = -0,3;
```

In der Analyse mit AMOS werden für die Faktorpfadkoeffizienten beta1, beta2 und beta3 – übereinstimmend mit dem Grundmodell und dem korrekt spezifizierten Modells mit zwei positiv korrelierten Fehlern – identische Werte von 0,707 errechnet.

Ergebnisse

Die aus der SAS-Simulation resultierenden arithmetischen Mittel der Pfadkoeffizienten sind erwartungstreu und die empirischen Standardfehler der ersten beiden Pfadkoeffizienten sind so groß wie im Grundmodell. Der empirische Standardfehler des dritten Pfadkoeffizienten jedoch, der im korrekt spezifizierten Modell mit zwei positiv korrelierten Fehlern erhöht ist, ist hier auf 0,23 reduziert. Die mittleren geschätzten Standardfehler von beta1 und beta2 entsprechen ebenfalls jenen des Grundmodells, während der mittlere geschätzte Standardfehler von beta1 wiederum auf 0,031 verringert ist. Die mittleren geschätzten Standardfehler der

[45] Die Fehlerkorrelation wurde von, $\lambda_{\varepsilon_1,\varepsilon_2} = -0.5$ auf $\lambda_{\varepsilon_1,\varepsilon_2} = -0.3$ reduziert, da sich zum einen im ursprünglichen fehlspezifizierten AMOS-Referenzmodell für beta3 der Wert Eins und für beta_f3 der Wert Null ergab und die entsprechende Simulation aufgrund dieser wahren Werte eine große Instabilität der Schätzung hervorbrachte (von Iteration zu Iteration stark divergierende geschätzte Standardfehler von beta_f3). Zum anderen ist der extreme Fall, dass der Einfluss einer latenten Variable (Faktor) auf eine manifeste maximal wird (standardisierter Pfadkoeffizient von 1) und damit der Einfluss der anderen latenten Variable (Fehler) minimal wird (standardisierter Pfadkoeffizient von 0), unwahrscheinlich und somit kaum von praktischer Relevanz.

Faktorpfadkoeffizienten überschätzen die tatsächlichen Streuungen der Pfadkoeffizientenschätzer gleichermaßen um ca. 30 %. Es bestehen keine nennenswerten Differenzen zwischen Medianen und arithmetischen Mittelwerten der geschätzten Standardfehler.

Die geschätzten Faktorpfadkoeffizienten der Indikatoren, deren Fehler negativ korreliert sind, sind selbst schwach negativ korreliert ($r(\hat{\beta}_1, \hat{\beta}_2) = -0{,}32$), während die restlichen beiden Pfadkoeffizientenkorrelationen vernachlässigbar gering sind.

Fazit

Im Modell mit einem Faktor und drei Indikatoren, in welchem die Fehler zweier Indikatoren negativ korreliert sind ($\lambda_{\varepsilon_1,\varepsilon_2} = -0.3$), stellt man bei korrekter Spezifikation fest, dass der Standardfehler des Faktorpfadkoeffizienten des dritten Indikators, SE(beta3), um gut 10 % *geringer* ist als der SE(beta3) im Grundmodell. Da die Mittelwerte der geschätzten Standardfehler die empirischen Standardfehler gleichermaßen um ca. 30 % überschätzen, wird für die Praxis empfohlen, die geschätzten Standardfehler mit einem Korrekturfaktor von 0,75 zu versehen, um den unbekannten tatsächlichen Standardfehlern nahezukommen.[46] Die Diskussion der Ergebnisse wird nach der Präsentation aller Modellvarianten vertieft fortgesetzt.

5.3.5 DAS FEHLSPEZIFIZIERTE MODELL

Die Fehlspezifikation der tatsächlichen Fehlerkorrelation $\lambda_{\varepsilon_1,\varepsilon_2} = -0.3$ erfolgt im COV-Statement:

```
Cov
Error_1 Error_2 = 0;
```

In der AMOS-Analyse werden für die Faktorpfadkoeffizienten beta1 und beta2 identische Werte von 0,592 errechnet. Für beta3 beträgt der Wert 0,845. Für beta1 und beta2 ist also ein identischer Bias (Unterschätzung um ca. 16 %) zu verzeichnen, während beta3 um ca. 20 % überschätzt wird.

Ergebnisse

Die arithmetischen Mittel der Pfadkoeffizienten beta1 und beta2 betragen 0,591 und stimmen somit nahezu mit dem im AMOS-Referenzmodell errechneten Wert überein. Der Mittelwert von beta3 beträgt 0,845 und trifft den Referenzwert genau.

---

[46] Der Versuch der Generalisierung des Simulationsergebnisses erfolgt unter dem Vorbehalt, dass die Einflüsse des Faktors auf die manifesten Variablen annähernd gleich groß sind.

Die empirischen Standardfehler der Faktorpfadkoeffizienten sind bezüglich des Grundmodells um ca. 10–20 % erhöht (0,028, 0,029 und 0,032). Die Mittelwerte der geschätzten Standardfehler der ersten beiden Pfadkoeffizienten betragen 0,036 und sind damit ebenso etwas größer als im Grundmodell. Der Mittelwert des geschätzten Standardfehlers des dritten Pfadkoeffizienten ist mit 0,039 deutlich größer. Somit überschätzen die Mittelwerte der geschätzten Standardfehler der Indikatoren, deren Fehlerterme korreliert sind, die jeweiligen empirischen Standardfehler um 20–30 %. Die arithmetischen Mittel und Mediane der geschätzten Standardfehler der Pfadkoeffizienten sind jeweils gleich groß.

Die Fehlspezifikation der Korrelation zwischen Fehler 1 und Fehler 2 geht einher mit einer schwachen positiven *Korrelation der geschätzten Faktorpfadkoeffizienten* beta1 und beta2 (r= 0,28). Die restlichen Korrelationen zwischen den Faktorpfadkoeffizienten, $r(\hat{\beta}_1, \hat{\beta}_3)$ und $r(\hat{\beta}_2, \hat{\beta}_3)$, sind schwach negativ (-0,40 und -0,39).

Fazit

Das fehlspezifizierte Modell mit zwei negativ korrelierten Fehlertermen repräsentiert einen Fall, der in der Praxis selten anzutreffen sein dürfte. Negativ korrelierte Fehlerterme sind in der Analyse empirischer Daten zwar denkbar, aber nur wesentlich weniger häufig anzutreffen als positiv korrelierte. Wie im fehlspezifizierten Modell mit zwei positiv korrelierten Fehlern ergibt sich der in AMOS errechnete Bias der Faktorpfadkoeffizienten im Durchschnitt auch in der Simulation. Die empirischen Standardfehler der Faktorpfadkoeffizienten werden durch die mittleren geschätzten SE in verhältnismäßig geringem Maß (um ca. 25 %) überschätzt und sind daher nur leicht konservativ. Die geschätzten Standardfehler können also mit einem Korrekturfaktor von 0,8 versehen werden.

5.3.6 DAS MODELL MIT DREI POSITIV KORRELIERTEN FEHLERN

Das Modell weist positive Korrelationen zwischen den Fehlern aller manifesten Variablen auf ($\lambda_{\varepsilon_1,\varepsilon_2} = \lambda_{\varepsilon_1,\varepsilon_3} = \lambda_{\varepsilon_2,\varepsilon_3} = 0{,}5$). Die Modellspezifikation des korrekt spezifizierten Modells lautet wie folgt:

```
Lineqs
indi_1ps =      beta1   Faktor_1       + beta_f1       Error_1,
indi_2ps =      beta2   Faktor_1       + beta_f2       Error_2,
indi_3ps =      beta3   Faktor_1       + beta_f3       Error_3;

Std
Error_1 = 1,
Error_2 = 1,
Error_3 = 1,
```

```
Faktor_1= 1;

Cov
Error_1 Error_2 = 0.5,
Error_1 Error_3 = 0.5,
Error_2 Error_3 = 0.5;
```

In der Analyse mit AMOS werden für die Faktorpfadkoeffizienten beta1, beta2 und beta3 – übereinstimmend mit dem Grundmodell – identische Werte von 0,707 (wahrer Wert) errechnet.

Ergebnisse

Die aus der SAS-Simulation resultierenden arithmetischen Mittel der Pfadkoeffizienten sind erwartungstreu. (Die Schätzer sind nur um 0,001–0,002 höher als die in AMOS errechneten Werte des Grundmodells.) Die empirischen Standardfehler der Pfadkoeffizienten liegen im Bereich zwischen 0,028 und 0,029 und sind damit bezüglich des Grundmodells um ca. 10 % erhöht.

Die mittleren geschätzten Standardfehler von beta1 und beta3 überschätzen die empirischen Standardfehler um ca. ein gutes Drittel und damit etwas stärker als im Grundmodell. Dieses Phänomen beruht auch darauf, dass in Iteration 796 SEs für die Faktorpfadkoeffizienten geschätzt werden, die deutlich erhöht sind. Der mittlere geschätzte Standardfehler von beta2 überschätzt den empirischen Standardfehler sogar um fast 60 %, da in Iteration 796 für beta2 und beta_f2 besonders hohe Werte geschätzt werden:

Schätzungen der Pfadkoeffizienten in Iteration 796

beta1	beta_f1	beta2	beta_f2	beta3	beta_f3
0,31557	0,26709	7,81080	13,6565	0,44031	0,40286

Entfernt man Iteration 796, so überschätzten die mittleren geschätzten SEs der Faktorpfadkoeffizienten (alle 0,038) die empirischen SEs gleichermaßen um ein gutes Drittel. Die Mediane der geschätzten SEs betragen 0,037 und weichen somit nur unwesentlich von den arithmetischen Mittelwerten ab.

Die Korrelationsmatrix der geschätzten Faktorpfadkoeffizienten enthält nur sehr schwache Korrelationen, die jene des Grundmodells betragsmäßig sogar unterschreiten und bis auf $r(\hat{\beta}_2, \hat{\beta}_3)$ nicht signifikant verschieden von Null sind.

Fazit

Das korrekt spezifizierte Modell mit drei positiv korrelierten Fehlern entspricht aufgrund der Symmetrie der Fehlerkorrelationen hinsichtlich der analysierten Eigenschaften Erwartungstreue, Variabilität und Korrelationsstruktur der Pfadkoeffizientenschätzungen im wesentlichen dem Grundmodell.

5.3.7 DAS FEHLSPEZIFIZIERTE MODELL

Die Fehlspezifikation der tatsächlichen Fehlerkorrelation $\lambda_{\varepsilon_1,\varepsilon_2} = \lambda_{\varepsilon_1,\varepsilon_3} = \lambda_{\varepsilon_2,\varepsilon_3} = 0.5$ erfolgt im COV-Statement:

```
Cov
Error_1 Error_2 = 0,
Error_1 Error_3 = 0,
Error_2 Error_3 = 0;
```

In der AMOS-Analyse werden für die Faktorpfadkoeffizienten beta1, beta2 und beta3 identische Werte von 0,866 errechnet. Alle drei überschätzen also den wahren Wert 0,707 um ca. 22 %.

Ergebnisse

Die arithmetischen Mittel der drei Faktorpfadkoeffizienten stimmen mit dem in AMOS errechneten Wert genau überein. (Der durchschnittliche Bias eines jeden der drei Koeffizienten ist also genauso groß wie der im AMOS-Modell errechnete Bias.) Alle drei Faktorpfadkoeffizienten werden um ca. 22,5 % überschätzt, um dasselbe Maß also, wie die ersten beiden Pfadkoeffizienten im fehlspezifizierten Modell mit zwei positiv korrelierten Fehlern.

Die empirischen Standardfehler der drei Pfadkoeffizienten sind identisch (0,011) und bezüglich der empirischen Standardfehler des Grundmodells um mehr als die Hälfte reduziert. Die Mittelwerte der geschätzten Standardfehler lauten ebenfalls gleich (0,026) und überschätzen die empirischen Standardfehler um ca. 140 %. Die Mediane und die arithmetischen Mittelwerte der geschätzten SEs sind – auf vier Dezimalstellen gerundet – gleich groß.

Die Korrelationsmatrix der geschätzten Faktorpfadkoeffizienten weist nur nicht signifikant von Null verschiedene Korrelationen auf.

Fazit

Der in diesem Modell angelegte Effekt, dass die Einflüsse der Fehlerkorrelationen infolge der Fehlspezifikation gleichmäßig allen Faktorpfadkoeffizienten zugerechnet werden, hat eine die

Variation der Pfadkoeffizientenschätzung stark dämpfende Wirkung. Die softwarebasierte Schätzung der Standardfehler ist in diesem Modell besonders konservativ, da die geschätzten Standardfehler von dieser Varianzreduktion in deutlich schwächerem Maß betroffen sind als die empirischen.

5.4 ZUSAMMENHÄNGE ZWISCHEN FEHLERKORRELATIONSSTRUKTUR, KORREKTER ODER FEHLSPEZIFIKATION, STANDARDFEHLER UND PFADKOEFFIZIENTENKORRELATION

Zwei der in den oben analysierten Modellen bereits erwähnten Effekte werden unter Zuhilfenahme einer tabellarischen Übersicht (Tabelle 6) systematisch untersucht.

Tabelle 6: Simulation zu Ein-Faktor-Modellen mit drei Indikatoren: empirischer und mittlerer geschätzter Standardfehler und Korrelationsmatrix der geschätzten Koeffizienten

Modell	Spezifikation	Empirischer SE von (mittlerer geschätzter SE von)			Korrelation von		
Pfadkoeffizienten		$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$(\hat{\beta}_1, \hat{\beta}_2)$	$(\hat{\beta}_1, \hat{\beta}_3)$	$(\hat{\beta}_2, \hat{\beta}_3)$
Grund~	korrekte	,024 (,034)	,026 (,034)	,026 (,034)	-,13	-,04	-,15
zwei pos. korr. Fehler	korrekte	,023 (,034)	,024 (,034)	,030 (,038)	+,33	-,27	-,27
	Fehl~	,018 (,030)	,019 (,030)	,023 (,031)	-,64	+,03	+,05
zwei neg. korr. Fehler	korrekte	,024 (,034)	,025 (,034)	,023 (,031)	-,32	-,02	-,00
	Fehl~	,028 (,036)	,029 (,036)	,032 (,039)	+,28	-,40	-,39
drei pos. korr. Fehler	Korrekte*	,028 (,038)	,029 (,038)	,029 (,038)	-,05	+,02	-,08
	Fehl~	,011 (,026)	,011 (,026)	,011 (,026)	+,00	+,05	-,03

*n=999 infolge von Ausschluss einer Iteration

Die korrekte Spezifikation einer positiven Fehlerkorrelation geht einher mit der Erhöhung des SE des Faktorpfadkoeffizienten des nicht betroffenen Indikators (1).

Die korrekte Spezifikation einer negativen Fehlerkorrelation dagegen führt zu einer Verringerung des SE des Faktorpfadkoeffizienten des nicht betroffenen Indikators (2).

Es wird versucht, diese beiden Sachverhalte zu erklären:

Ad 1: Bei positiv korrelierten, korrekt spezifizierten Fehlern sind die zugehörigen geschätzten Faktorpfadkoeffizienten ebenfalls schwach positiv korreliert. Dies führt dazu, dass beispielsweise eine relativ hohe Schätzung eines zugehörigen Koeffizienten tendenziell mit einer relativ hohen Schätzung des anderen zugehörigen Koeffizienten einhergeht. Diese häufig vorkommende gleichzeitige Überschätzung beider zugehöriger Pfadkoeffizienten wird mit einer

Unterschätzung des nicht zugehörigen Faktorpfadkoeffizienten kompensiert. Dies äußert sich in Form der negativen Korrelationen $r(\hat{\beta}_1, \hat{\beta}_3)$ und $r(\hat{\beta}_2, \hat{\beta}_3)$. Die Variabilität von beta3 hängt daher auch von der der anderen Faktorpfadkoeffizienten ab und ist daher besonders hoch.

Ad 2: Bei negativ korrelierten, korrekt spezifizierten Fehlern sind die zugehörigen geschätzten Faktorpfadkoeffizienten ebenfalls leicht negativ korreliert. Dies bedeutet, dass eine relativ hohe Schätzung eines zugehörigen Koeffizienten tendenziell mit einer relativ niedrigen Schätzung des anderen zugehörigen Koeffizienten einhergeht. Da mithin Überschätzungen des einen zugehörigen Pfadkoeffizienten quasi mit Unterschätzung des anderen zugehörigen Pfadkoeffizienten kompensiert wird, findet ein Ausgleich durch den nicht zugehörigen Pfadkoeffizienten beta3 *nicht* statt. Die vernachlässigbaren Korrelationen $r(\hat{\beta}_1, \hat{\beta}_3)$ und $r(\hat{\beta}_2, \hat{\beta}_3)$ belegen dies. Die Variabilität von beta3 hängt nicht von der der anderen Faktorpfadkoeffizienten ab und ist daher relativ gering.[47] Dass sie sogar geringer ist als im Grundmodell, ist plausibel: Wenn im Grundmodell abweichende Pfadkoeffizientenschätzungen nicht innerhalb der zugehörigen Faktorpfadkoeffizienten kompensiert werden – $r(\hat{\beta}_1, \hat{\beta}_2)$ ist vernachlässigbar – dann findet eher ein Ausgleich durch den nicht zugehörigen Faktorpfadkoeffizienten statt, was dessen Variabilität erhöht.

5.5 FAZIT AUS ALLEN MODELLEN

Die PROC CALIS-Simulation korrekt spezifizierter AMOS-Modelle bringt Pfadkoeffizientenmittelwerte hervor, die erwartungstreue Schätzer der in AMOS (und PROC CALIS identisch) berechneten Koeffizienten sind. Der durchschnittliche Bias der Pfadkoeffizientenschätzungen, der in den Simulationen bei Fehlspezifikation auftritt, entspricht in fast allen Modellen dem in AMOS (bzw. PROC CALIS identisch) errechneten Bias (siehe Tabelle 7).

[47] Hierbei ist zu berücksichtigen, dass die negative Fehlerkorrelation nur r = -0,3 beträgt. Betrüge sie r = -0,5, wäre die Variabilität deutlich geringer (SE=0,021). Zur besseren Vergleichbarkeit der Modelleffekte (1) und (2) müssten die Beträge der Fehlerkorrelationen der Modelle mit positiv und negativ korrelierten Fehlern dieselbe Höhe aufweisen. Dies ist jedoch aufgrund der Instabilität der Schätzung nicht möglich (siehe hierzu Beginn von 3.5.3).

Tabelle 7: Pfadkoeffizienten der Ein-Faktor-Modelle mit drei Indikatoren laut AMOS- bzw. PROC CALIS-Berechnung und PROC CALIS-Simulation

Modell	Spezifikation	AMOS- bzw. PROC CALIS-Berechnung („n=1“)			PROC CALIS-Simulation (Mittelwert aus n=1000 Iterationen)		
Pfadkoeffizient		beta1	beta2	beta3	beta1	beta2	beta3
Grund~	korrekte	0,707	0,707	0,707	0,707	0,707	0,707
zwei pos. korr. Fehler	korrekte	0,707	0,707	0,707	0,707	0,708	0,707
	Fehl~	0,866	0,866	0,577	0,866	0,867	0,577
zwei neg. korr. Fehler	korrekte	0,707	0,707	0,707	0,707	0,707	0,707
	Fehl~	0,592	0,592	0,845	0,591	0,591	0,845
drei pos. korr. Fehler	korrekte	0,707	0,707	0,707	0,708*	0,709*	0,708*
	Fehl~	0,866	0,866	0,866	0,866	0,866	0,866

*n=999 infolge von Ausschluss einer Iteration

Die empirischen Standardfehler werden durch den Mittelwert der geschätzten Standardfehler stets überschätzt. Das Maß der Überschätzung beträgt in korrekt spezifizierten Modellen zwischen 23 und 43 Prozent. In fehlspezifizierten Modellen werden die empirischen Standardfehler um 24 bis 140 Prozent überschätzt. Innerhalb eines bestimmten Modells kommt es vor, dass die empirischen SE der drei Pfadkoeffizienten in unterschiedlich hohem Maß überschätzt werden. Fehlspezifikation führt mithin in einigen Fällen dazu, dass die empirischen Standardfehler besonders konservativ geschätzt werden, wobei die Überschätzung in umfassend fehlspezifizierten Modellen (Modell mit drei positiv korrelierten Fehlern) wesentlich größer ist. In den korrekt spezifizierten Modellen mit zwei positiv und zwei negativ korrelierten Fehlern ist der empirische Standardfehler des Pfadkoeffizienten beta3 erhöht bzw. verringert. Es handelt sich hierbei um den Koeffizienten jenes Indikators, dessen Fehlervariable mit keiner der anderen beiden Indikatoren korreliert. Da die Überschätzung des empirischen Standardfehlers eines Faktorpfadkoeffizienten im Mittel wenigstens ca. 25 Prozent beträgt, wird ein Korrekturfaktor von 0,8 für den geschätzten Standardfehler vorgeschlagen.

Korrelationen zwischen Faktorpfadkoeffizienten, die wesentlich verschieden von Null sind, treten nur in Modellen auf, in welchen tatsächlich nur eine Fehlerkorrelation vorliegt. Hierbei ist es gleichgültig, welches Vorzeichen die Fehlerkorrelation aufweist und ob das Modell korrekt spezifiziert ist. In den Modellen, in welchen entweder keine oder alle Fehler korreliert sind, kommen keine nennenswert von Null verschiedenen Koeffizientenkorrelationen vor. Korreliertheit von Pfadkoeffizienten ist in Anbetracht des kleinen Streubereichs der Pfadkoeffizientenschätzungen von untergeordneter Bedeutung.

6 Simulationen zum Ein-Faktor-Modell mit vier oder fünf Indikatoren unter Verwendung von PROC CALIS

6.1 EIN-FAKTOR-MODELLE MIT DREI VS. VIER UND MEHR INDIKATOREN

Ein-Faktor-Modelle mit drei Indikatoren (Messverfahren) wurden als Ausgangspunkt der Analyse gewählt, weil unter der Annahme der bedingten Unabhängigkeit der Messverfahren, die im Grundmodell zutrifft, die Pfadkoeffizienten *exakt errechnet* werden können.

Die Analyse der fehlspezifizierten Varianten der weiteren Modelle machte deutlich, dass die fälschlicherweise getroffene Annahme der bedingten Unabhängigkeit zu einem Bias in der Berechnung der Pfadkoeffizienten führt (siehe 2.6).

Stehen vier oder fünf Messverfahren zur Disposition, kann analysiert werden, ob eine bzw. zwei Korrelationen verschieden Null zwischen den Fehlern jeweils zweier Indikatoren gegeben sind.

Wenn auf diese Weise positive oder negative Fehlerkorrelationen detektiert und anschließend in den Modellannahmen berücksichtigt werden, können valide Schätzungen der Pfadkoeffizienten resultieren.

6.2 EIN-FAKTOR-MODELLE MIT VIER ODER FÜNF INDIKATOREN

Die Konstruktion der manifesten Variablen im Ein-Faktor-Modell mit drei Indikatoren wird übernommen und lediglich um eine bzw. zwei Indikatorvariable nebst zugehörigem Fehlerterm erweitert. Die manifesten Variablen werden also derart erzeugt, dass sich – *zunächst unter Verwendung perfekt unabhängiger Zufallsvariabler*[48] – für alle vier bzw. fünf Messverfahren identische standardisierte Pfadkoeffizienten (Faktor- und Fehlerpfadkoeffizienten) mit dem Wert 0,707 ergeben. Dieser Wert wird im Folgenden als „wahrer Wert eines Pfadkoeffizienten“ bezeichnet. Weiterhin werden – je nach Modellvariante – für die sechs verschiedenen Paare aus vier Messverfahren (bzw. für die zehn verschiedenen Paare aus fünf Messverfahren) Fehlerkorrelationen $r_{\varepsilon_i,\varepsilon_j}$ in der Höhe von 0, 0,5 oder -0,5 konstruiert. Auf diese drei möglichen Werte einer Fehlerkorrelation wird mit der Bezeichnung „wahrer Wert der Korrelation“

[48] Diese elementaren Zufallsvariablen sind orthogonale Faktoren, die aus einer Hauptkomponentenanalyse stammen (siehe auch 5.1)

im Folgenden Bezug genommen. Die wahren Fehlerkorrelationen $r_{\varepsilon_1,\varepsilon_3}$, $r_{\varepsilon_1,\varepsilon_4}$, $r_{\varepsilon_2,\varepsilon_3}$ und $r_{\varepsilon_2,\varepsilon_4}$ betragen in allen Modellen Null. In den beiden Modellen mit fünf Messverfahren betragen zusätzlich die wahren Werte aller Korrelationen Null, die zwischen dem Fehler des fünften Indikators und den Fehlern aller anderen Indikatoren bestehen $r_{\varepsilon_i,\varepsilon_5}$ $(i = 1, 2, ..., 4)$.

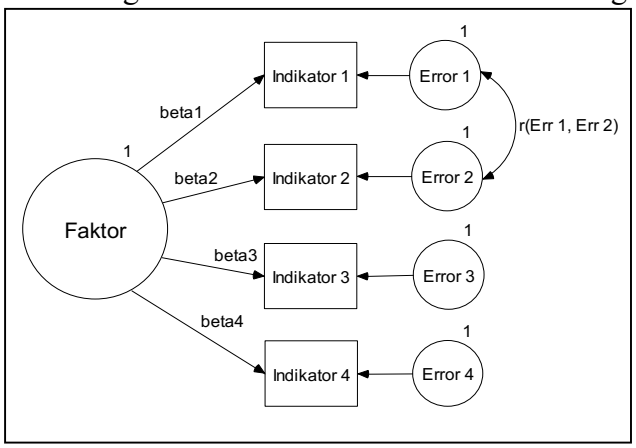
Bei der *anschließenden Simulation* der n Stichproben in SAS werden – wie beim Ein-Faktor-Modell mit drei Indikatoren – approximativ standardnormalverteilte, i. A. *nicht perfekt unabhängige Zufallsvariable* erzeugt, aus welchen wiederum manifeste Variable konstruiert werden (siehe 11.4.1). Die Simulation umfasst 1000 Iterationen (Stichproben), wobei jeweils Zufallsvariable mit 1000 Realisationen erzeugt werden.

In dieser Untersuchung werden Fehlerkorrelationen geschätzt. Hauptsächlich interessiert hierbei, ob die durchschnittlichen geschätzten Fehlerkorrelationen in der Simulation den konstruierten, erwarteten Korrelationen entsprechen. Weiterhin ist von Interesse, ob die durchschnittlichen, geschätzten Pfadkoeffizienten den erwarteten Pfadkoeffizienten entsprechen. Es soll dabei geklärt werden, inwiefern eine *Berücksichtigung einer tatsächlich vorhandenen Fehlerkorrelationsstruktur eine valide Schätzung der interessierenden Faktorpfadkoeffizienten* ermöglicht. In diesem Kapitel werden daher – im Unterschied zu 5.3 – hinsichtlich der Fehlerkorrelation nur korrekt spezifizierte Modelle untersucht.

6.3 MODELLE MIT VIER INDIKATOREN

Die unten stehende Abbildung 44 zeigt das Ein-Faktor-Modell mit vier Indikatoren, das die folgenden Modelle 1 und 2 charakterisiert, in grafischer Darstellung. In diesen *ersten beiden Modellen wird die Korrelation zwischen Fehler 1 und 2 geschätzt*. Im Vier-Indikatoren-Modell stehen $\frac{4 \cdot 5}{2} = 10$ Varianzen und Kovarianzen zur Verfügung. Somit können, um die notwendige Identifikationsbedingung zu erfüllen, acht Pfadkoeffizienten und eine Fehlerkorrelation, also neun Parameter geschätzt werden (df =1). Die Identifikationsregel 4 (siehe 2.7) ist erfüllt, da sie bereits bei Vorhandensein von nur drei manifesten Variablen erfüllt ist. Die vierte manifeste Variable, die – bezogen auf das Modell mit drei manifesten Variablen – die Zahl der Varianzen um eins und die der Kovarianzen um drei erhöht, gewährleistet die Erfüllung auch der hinreichenden Identifikationsbedingung.

Abbildung 44: Modell mit vier Indikatoren - allgemeine Modellspezifikation



6.3.1 MODELL 1

Die Erzeugung der manifesten Variablen dieses Modell ist dadurch gekennzeichnet, dass die Fehlervariablen der ersten beiden Indikatoren mit r = 0,5 (wahrer Wert) korreliert sind. Alle anderen fünf Korrelationen, die zwischen den Fehlertermen der vier Indikatoren gebildet werden können, weisen einen wahren Wert von 0 auf.

Ergebnisse

Die Mittelwerte von beta1 bis beta4 betragen 0,7067, 0,7068, 0,7082 und 0,7085. Alle vier Schätzer für die Faktorpfadkoeffizienten sind somit erwartungstreu. Die Mittelwerte der Koeffizienten der beiden Indikatoren mit korrelierten Fehlern liegen minimal unter den Mittelwerten der Koeffizienten mit nicht korrelierten Fehlern.

Der Schätzer „Mittelwert der Fehlerkorrelation der ersten beiden Indikatoren“ beträgt 0,4991 und ist daher ebenfalls erwartungstreu.

Die empirischen Standardfehler von beta1 bis beta4 betragen 0,0254, 0,0247, 0,0242 und 0,0232. Die Mittelwerte der geschätzten Standardfehler hingegen liegen bei 0,0335 (für beta1 und beta2) sowie 0,0326 (für beta3 und beta4) und überschätzen die empirischen Standardfehler ca. um ein Drittel – ein Phänomen, das aus den Modellen mit drei Indikatoren bereits bekannt ist. Die Mediane der geschätzten Standardfehler weichen für beta1 bis beta4 – auf vier Nachkommastellen gerundet – nicht von den arithmetischen Mittelwerten der geschätzten Standardfehler ab.

Der empirische Standardfehler von $r_{\varepsilon_1,\varepsilon_2}$ beträgt 0,0346 und wird durch den Mittelwert der geschätzten Standardfehler, 0,0348, fast exakt geschätzt.

6.3.2 *MODELL 2*

Die Konstruktion der manifesten Variablen dieses Modell unterscheidet sich von der des Vormodells nur dadurch, dass die Fehlervariablen der ersten beiden Indikatoren negativ korreliert sind (r = -0,5).

Ergebnisse

Die mittleren Werte der interessierenden Pfadkoeffizienten betragen 0,7068, 0,7059, 0,7083 und 0,7082 und sind somit erwartungstreue Schätzer. Sie weichen nur unwesentlich von jenen des Vormodells ab. Der Mittelwert der Fehlerkorrelation ist mit -0,4993 ebenfalls erwartungstreu.

Die empirischen Standardfehler der ersten beiden Pfadkoeffizienten (0,0255, 0,0253) entsprechen jenen des Vormodells, während die der letzten beiden deutlich verringert sind (0,0211, 0,0210).

Der empirische Standardfehler der Fehlerkorrelation hingegen ist bezüglich des Vormodells nahezu verdoppelt (0,0632).

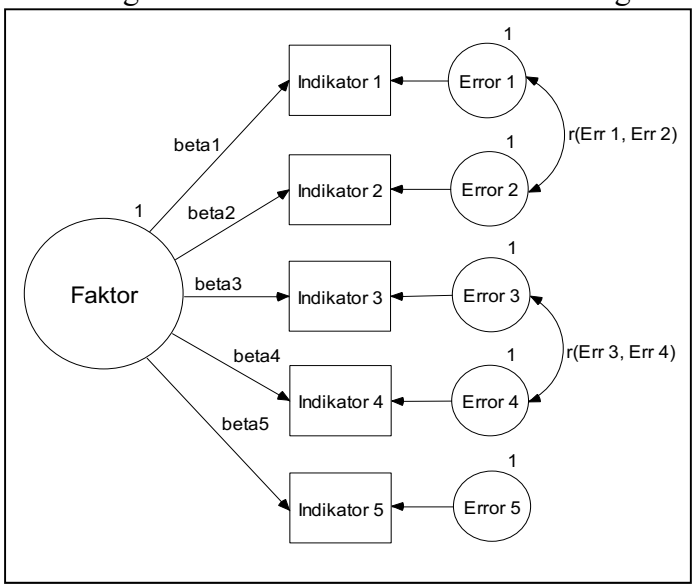
Die arithmetisch gemittelten geschätzten Standardfehler der ersten beiden Faktorpfadkoeffizienten überschätzen die empirischen Standardfehler im gleichen Maß wie im Vormodell. Die empirischen Standardfehler der letzten beiden Pfadkoeffizienten werden sogar um ca. 45 % überschätzt. Die medianen geschätzten Standardfehler weichen von den arithmetisch gemittelten geschätzten Standardfehlern – gleich dem Vormodell – nicht ab. Das arithmetische Mittel der geschätzten Standardfehler der negativen Fehlerkorrelation gibt ihren empirischen Standardfehler genau wieder.

6.4 MODELLE MIT FÜNF INDIKATOREN

Die nachfolgende Abbildung 45 gibt das Ein-Faktor-Modell mit fünf Indikatoren wieder, das die Struktur der Modelle 3 und 4 aufzeigt. In diesen letzten beiden Modellen kann aufgrund der Aufnahme einer fünften manifesten Variablen (unter Beibehaltung der bisherigen Modellstruktur) *zusätzlich* zur Fehlerkorrelation der Vormodelle *die Korrelation zwischen den Fehlern 3 und 4 geschätzt werden*. Im Fünf-Indikatoren-Modell stehen $\frac{5 \cdot 6}{2} = 15$ Varianzen und Kovarianzen zur Verfügung, aus welchen zehn Pfadkoeffizienten und zwei Fehlerkorrelationen geschätzt werden sollen (df =3), so dass die notwendige Bedingung zur Identifizierbarkeit des Modells erfüllt ist. Die hinreichende ist erfüllt, weil sie bereits im Modell mit drei manifesten Variablen erfüllt ist und im Unterschied zu letzterem zusätzlich zwei Varianzen und

sieben Kovarianzen zur Verfügung stehen und zusätzlich nur vier Pfadkoeffizienten und zwei Fehlerkorrelationen geschätzt werden sollen. (Zur Identifizierbarkeit siehe 2.7.)

Abbildung 45: Modell mit fünf Indikatoren - allgemeine Modellspezifikation



6.4.1 MODELL 3

Zusätzlich zur Fehlerkorrelation des Modells 1 sind die Fehlervariablen der Indikatoren 3 und 4 mit r = 0,5 (wahrer Wert) korreliert. Der Fehler des fünften Indikators ist mit den Fehlern aller anderen Indikatoren mit einem wahren Wert von Null korreliert.

Ergebnisse

Wie in den Vormodellen erweisen sich die arithmetischen Mittelwerte für beta1 bis beta4 als erwartungstreue Schätzer des wahren Wertes der Pfadkoeffizienten. Der Mittelwert der Fehlerkorrelation der ersten beiden Indikatoren beträgt 0,4959 und ist damit ebenso erwartungstreu wie der Mittelwert der Fehlerkorrelation der letzten beiden Indikatoren (0,5003).

Die empirischen Standardfehler der Pfadkoeffizienten gleichen in etwa jenen der Vormodelle. Die Mittelwerte der geschätzten Standardfehler der ersten vier Pfadkoeffizienten betragen gleichermaßen 0,0326, während der des Pfadkoeffizienten beta5 geringfügig geringer ist (0,0317).

Somit werden die arithmetischen Mittelwerte der geschätzten SE, den Vormodellen gleichend, um ca. ein Drittel überschätzt. Die Fehlerkorrelationen $r_{\varepsilon_1,\varepsilon_2}$ und $r_{\varepsilon_3,\varepsilon_4}$ weisen fast identische empirische Standardfehler (0,0333 bzw. 0,0321) auf, die etwas kleiner als in Modell 1 sind. Sie werden durch die arithmetischen Mittelwerte der geschätzten Standardfehler

(0,0333 bzw. 0,0329) fast exakt reproduziert. Die Mediane der geschätzten Standardfehler, sowohl der Pfadkoeffizienten als auch der Fehlerkorrelationen, weichen auch in diesem Modell nur minimal von den arithmetischen Mitteln derselben ab.

6.4.2 MODELL 4

Die Konstruktion der manifesten Variablen dieses Modell unterscheidet sich von der des Vormodells nur dadurch, dass die Fehlervariablen der ersten beiden Indikatoren negativ korreliert sind. Der wahre Wert von $r_{\varepsilon_1,\varepsilon_2}$ beträgt also -0,5 und der von $r_{\varepsilon_3,\varepsilon_4}$ lautet 0,5.

Ergebnisse

Entsprechend dem Vormodell stellen sich die arithmetischen Mittelwerte für beta1 bis beta5 als erwartungstreue Schätzer der Pfadkoeffizienten heraus. Die arithmetischen Mittel der beiden geschätzten Fehlerkorrelationen betragen -0,5079 und 0,5007 und sind ebenfalls erwartungstreu. Die empirischen Standardfehler der ersten beiden Pfadkoeffizienten (0,0232 und 0,0247) sind etwas größer als die der Pfadkoeffizienten 3 und 4 (0,0216 und 0,0214) und deutlich größer als die des Pfadkoeffizienten zum fünften Indikator, der keinen korrelierten Fehler besitzt (0,0206).

Die Mittelwerte der geschätzten Standardfehler der fünf Pfadkoeffizienten unterscheiden sich voneinander so wie die empirischen Standardfehler der fünf Pfadkoeffizienten: Der größte Wert ergibt sich für die ersten beiden Pfadkoeffizienten (beide 0,0326), der zweitgrößte für die Pfadkoeffizienten 3 und 4 (beide 0,0304) und der geringste für den Pfadkoeffizienten beta5 (0,0300).

Die arithmetischen Mittelwerte der geschätzten SE werden wie in allen Vormodellen mithin um ca. ein Drittel überschätzt. Die Fehlerkorrelationen $r_{\varepsilon_1,\varepsilon_2}$ und $r_{\varepsilon_3,\varepsilon_4}$ weisen jedoch – abweichend von Modell 3 – stark unterschiedlich große Standardfehler auf: Der empirische SE der negativen Fehlerkorrelation, 0,0600, ist mehr als das Doppelte so groß wie der der positiven Fehlerkorrelation (0,0289) und entspricht damit in etwa dem in Modell 2 beobachteten. Die Spannweite der simulierten negativen Fehlerkorrelationen ist mit -0,3073 - (-0,7040) = 0,3967 mehr als zweifach so groß wie die Spannweite der simulierten positiven (0,5799 – 0,4034 =) 0,1765. Der mittlere geschätzte SE der negativen Fehlerkorrelation, 0,5800, unterschätzt den empirischen SE minimal, während der mittlere geschätzte SE der positiven Fehlerkorrelation, 0,0293, den empirischen fast genau reproduziert. Die Mediane der geschätzten

Standardfehler, sowohl der Pfadkoeffizienten als auch der Fehlerkorrelationen, weichen auch in diesem Modell nur minimal von den arithmetischen Mitteln derselben ab.

6.5 FAZIT AUS ALLEN MODELLEN

In den Modellen wird die tatsächlich vorhandene Fehlerkorrelationsstruktur korrekt geschätzt und somit auch eine valide Schätzung der interessierenden Faktorpfadkoeffizienten ermöglicht (siehe Tabelle 8). Dies trifft sowohl für Modelle zu, die eine positive oder eine negative Fehlerkorrelation zwischen lediglich zwei Indikatoren aufweisen, als auch für Modelle, die zwei Fehlerkorrelationen besitzen. Die mittleren geschätzten Standardfehler der Fehlerkorrelationen geben die empirischen Standardfehler fast genau wieder. (Die Überschätzung der empirischen Standardfehler durch die mittleren geschätzten SE betrifft nur die Pfadkoeffizienten.) Die empirischen als auch die mittleren geschätzten Standardfehler der negativen Fehlerkorrelation sind ca. doppelt so hoch wie die der positiven Fehlerkorrelation. Dies gilt unabhängig davon, ob die negative Korrelation die einzig geschätzte Fehlerkorrelation in einem Modell ist (Modell 2) oder das Modell zusätzlich eine positive Fehlerkorrelation aufweist (Modell 4).

Tabelle 8: Ein-Faktor-Modelle mit vier oder fünf Indikatoren - Berechnung und Simulation mit PROC CALIS

Modellnummer	Berechnung (wahrer Wert) (Verwendung vollkommen unkorr. Zufallsvariabler; n=1)		Ergebnis der Schätzung in der Simulation (Verwendung schwach korrelierter Zufallsvariabler; Mittelwert aus n=1000)						
Fehlerkorrelation/ Pfadkoeffizient	$r_{\varepsilon_1,\varepsilon_2}$	$r_{\varepsilon_3,\varepsilon_4}$	$r_{\varepsilon_1,\varepsilon_2}$	$r_{\varepsilon_3,\varepsilon_4}$	beta1	beta2	beta3	beta4	beta5
1	0,5	0	0,499	-------	0,707	0,707	0,708	0,709	-------
2	-0,5	0	-0,499	-------	0,707	0,706	0,708	0,708	-------
3	0,5	0,5	0,496	0,500	0,709	0,710	0,707	0,707	0,707
4	-0,5	0,5	-0,508	0,501	0,709	0,708	0,707	0,707	0,707

Für die Modelle 1 und 2 gilt: Die berechneten (wahren Werte) der nicht angegebene Fehlerkorrelationen $r_{\varepsilon_1,\varepsilon_3}$, $r_{\varepsilon_1,\varepsilon_4}$ $r_{\varepsilon_2,\varepsilon_3}$ und $r_{\varepsilon_2,\varepsilon_4}$ betragen Null. Für die Modelle 3 und 4 gilt zusätzlich: Die wahren Werte der Korrelationen zwischen dem Fehler des fünften Indikators und den Fehlern aller anderen Indikatoren $r_{\varepsilon_i,\varepsilon_5}$ $(i = 1,2,...,4)$ betragen Null. Die wahren Werte der Pfadkoeffizienten beta1 – beta4 (bzw. – beta5) fehlen, da alle Pfadkoeffizienten in allen Modellen den Wert 0,707 aufweisen.

7 Simulationen zu linearen Strukturgleichungsmodellen (SEMs) mit zwei Faktoren unter Verwendung von PROC CALIS

7.1 EINLEITUNG

Die Simulationen zu linearen Strukturgleichungsmodellen (SEMs) mit zwei Faktoren bilden den Abschluss des systematischen Teils dieser Arbeit.

Zunächst wurden mit den Simulationen zu Ein-Faktor-Modellen mit drei Indikatoren *die Effekte verschiedener Fehlspezifikationen von Fehlerkorrelationen auf Pfadkoeffizienten* eines einfachen Messmodells untersucht. Es wurden hierbei hinsichtlich der Fehlerkorrelationsstruktur fehlspezifizierte Modelle korrekt spezifizierten gegenübergestell*t*. Diese und alle folgenden Simulationen erfolgten unter Verwendung von PROC CALIS (SAS) und berücksichtigen (im Unterschied zur jeweils vorausgehenden Analyse mit AMOS) zufällige Schwankungen in der Korrelationsstruktur der manifesten Variablen. Anschließend wurden mit den Simulationen zu Modellen mit vier oder fünf Indikatoren an umfangreicheren Messmodellen die Möglichkeit der Schätzung von Pfadkoeffizienten und zugleich Fehlerkorrelationen analysiert. Diese Simulationen untersuchten nur korrekt spezifizierte Modelle.

In dem hier vorliegenden Kapitel werden schließlich Simulationen zu vollständigen Strukturgleichungsmodellen (SEMs) vorgestellt, die zwei einfache Messmodelle mittels eines Strukturmodells verbinden. Das Kapitel behandelt drei Gruppen von Modellen. Im Focus steht hierbei die *Untersuchung des Effekts einer Fehlspezifikation der Fehlerkorrelationsstruktur auf die Schätzung des Pfadkoeffizienten gamma*[49] im Strukturmodell. (Gamma kann hierbei z. B. angeben, wie stark eine nicht direkt messbare Erkrankung eine andere ebenfalls nicht unmittelbar quantifizierbare Erkrankung beeinflusst.) Diese Untersuchung wird an der ersten Gruppe, bestehend aus neun Modellen, durchgeführt. Bis auf zwei sogenannte Grundmodelle (Modell 1 und Modell 4) sind alle Modelle dieser Gruppe fehlspezifiziert.

Es schließt sich die aus vier Modellen bestehende zweite Modellgruppe an. Hierbei handelt es sich um ein (in der Systematik der ersten neun Modelle fehlendes) Zusatzmodell und korrekt spezifizierte Varianten von Modellen mit statistisch abhängigen Faktoren, die eine einzige Fehlerkorrelation aufweisen. Die korrekte Spezifikation der jeweiligen Fehlerkorrelation wird

[49] siehe Abbildung 6

dadurch erzielt, dass man die (aus den entsprechenden AMOS-Modellen stammende) externe Information über die wahre Fehlerkorrelationsstruktur in das Modell aufnimmt.

Die dritte, sieben Modelle umfassende Gruppe von Modellen, enthält korrekt spezifizierte Varianten von Modellen mit einer oder zwei Fehlerkorrelationen, wobei die korrekte Spezifikation dadurch erzielt wird, dass die Fehlerkorrelationsstruktur im Modell geschätzt wird. Die korrekt spezifizierten Modelle der Modellgruppen zwei und drei dienen dazu, die in Modellgruppe eins untersuchte Wirkung der Fehlspezifikation der Fehlerkorrelationsstruktur auf gamma isoliert von den möglichen Wirkungen der Fehlerkorrelationsstruktur selbst quantifizieren zu können. Alle die Modelle mit zwei Faktoren und jeweils drei Indikatoren pro Faktor sind mit einer Ausnahme identifiziert.

Ausgangspunkt der Simulationen zu den SEMs mit PROC CALIS sind die mit AMOS analysierten Strukturgleichungsmodelle. Die Simulationen zu den SEMs unterscheiden sich von den mit AMOS untersuchten Strukturgleichungsmodellen in der gleichen Weise, wie die Simulationen zu den Ein-Faktor-Modellen mit PROC CALIS sich von den mit AMOS untersuchten Ein-Faktor-Modellen unterscheiden: Die Korrelationsstruktur der manifesten Variablen konnte in einem AMOS-Modell exakt kontrolliert werden, da diese Variablen unter Verwendung perfekt unabhängiger Zufallsvariabler konstruiert wurden. Jedes AMOS-Modell wurde auf Basis einer einzigen derart erzeugten Datendatei mit 1517 Realisationen von Zufallsvariablen errechnet. In den Simulationen hingegen unterliegt die Korrelationsstruktur der manifesten Variablen zufälligen Schwankungen, da die verwendeten Zufallsvariablen nicht perfekt unabhängig sind. Jede Simulation zu einem Modell geschieht auf Basis von n=3.000 Iterationen, wobei eine jede 1000 Realisationen von Zufallsvariablen umfasst. (Siehe 5.1.) Die Verdreifachung der in den vorausgehenden Simulationen verwendeten Anzahl von 1000 Iterationen folgt der bei Simulationen geltenden Konvention und trägt damit der Bedeutung dieses Kapitels Rechnung.

Zu jedem der neun mit AMOS untersuchten SEMs wird eine Simulation unter Verwendung von PROC CALIS durchgeführt.

In 5.3 wurde bereits für alle Modellvarianten dokumentiert, dass ein mit AMOS auf Basis einer einzigen Datei errechnetes Modell durch PROC CALIS (auf Basis derselben Datei) exakt reproduziert wird. Die Reproduzierbarkeit wurde vor Durchführung der Simulationen geprüft. Daher erübrigte sich die Prüfung der Reproduzierbarkeit der AMOS Ergebnisse

durch PROC CALIS im Fall einer einzigen „Stichprobe“ vor Durchführung der Simulationen zu den SEMs mit zwei Faktoren und kann auch hier als gegeben vorausgesetzt werden.

Das wichtigste Ziel der Simulationen zu den SEMs ist die *Analyse des Strukturmodells*.

Von Interesse (und völlig unabhängig von der Untersuchung der Erwartungstreue) ist zunächst die Frage, ob das aus der Simulation resultierende arithmetische Mittel von gamma den im entsprechenden AMOS-Modell errechneten Wert von gamma, im Folgenden genannt AMOS-Referenzwert, reproduziert. Das Hauptziel besteht dann darin, die Erwartungstreue des Schätzers für den Pfadkoeffizienten gamma in den neun Modellen zu überprüfen. Der Schätzer ist erwartungstreu, wenn er den bei korrekter Modellspezifikation im korrespondierenden AMOS errechneten Wert annähert. Insbesondere soll ermittelt werden, in welchen der fehlspezifizierten Modellvarianten ein Bias bei gamma festgestellt werden kann. Es werden das arithmetische Mittel von gamma und sein simulierter Standardfehler (SE) erhoben. Der mittlere geschätzten Standardfehler wird mit letzterem verglichen, um zu klären, wie gut der simulierte SE auf Basis des geschätzten vorhergesagt werden kann.

Ein weiteres Ziel der Simulationen zu den SEMs ist die *Analyse der der beiden Messmodelle*, die im SEM enthalten *sind*. Die meisten der neun SEMs beinhalten Messmodelle, in welchen die arithmetischen Mittel der Pfadkoeffizienten kaum von jenen abweichen, die sich bereits in den Simulationen zum Ein-Faktor-Modell mit drei Indikatoren ergaben und ausführlich erläutert wurden. Die Messmodelle dieser SEMs verdienen insofern wenig Aufmerksamkeit. (Abweichend verhalten sich nur die Messmodelle der beiden letzten SEMs.) Bei den simulierten Standardfehlern der Messmodellpfadkoeffizienten hingegen sind Phänomene zu beobachten, die in den Simulationen zu den Ein-Faktor-Modellen nicht auftraten, da sie sich erst aus der Integration zweier Ein-Faktor-Modelle in ein SEM ergeben.

In den ersten drei der neun SEMs, zu welchen Simulationen durchgeführt werden, ist der wahre Wert des Strukturmodellpfadkoeffizienten gamma gleich Null. In diesem *Spezialfall* sind die beiden Faktoren statistisch unabhängig. In den folgenden SEMs, zu welchen Simulationen durchgeführt werden, ist der wahre Wert des Pfadkoeffizienten gamma gleich 0,707. Die beiden Faktoren sind mithin statistisch abhängig. Der Spezialfall wurde im korrespondierenden AMOS-Kapitel 4 dem Normalfall aus didaktischen Gründen vorangestellt. Diese Reihenfolge soll auch in den Simulationen zu den SEMs aus gliederungstechnischen Gründen gewahrt bleiben.

Die Effekte der Fehlspezifikationen hinsichtlich der Schätzungen der Pfadkoeffizienten selbst wurden bereits im Kapitel 4 dokumentiert und dort bereits ausführlich interpretiert. Insofern keine qualitativ neuartigen Effekte in der Simulationsanalyse zu beobachten sind, gelten diese *Modellinterpretationen* auch für die folgende Untersuchung.

Hinsichtlich der Identifizierbarkeit der Modelle sind die in 2.7 dargelegten notwendigen und hinreichenden Identifikationsbedingungen erfüllt. Eine Besonderheit stellt das Modell in 7.5.2 dar, das aus Gründen der Identifizierbarkeit zu einer Erhöhung der Anzahl der manifesten Variablen zwang.

7.2 STRUKTURGLEICHUNGSMODELLE, DEREN FAKTOREN STATISTISCH UNABHÄNGIG SIND

Es wurden u. a. bereits Simulationen in PROC CALIS zu AMOS-Modellen mit drei Indikatoren durchgeführt. In diesen Simulationen fiel auf, dass in einer nicht geringen Anzahl der n Stichproben (Iterationen) alle interessierenden Koeffizienten des jeweiligen Modells mit negativem Vorzeichen geschätzt werden. Dieses Phänomen beruht darauf, dass softwareintern aus korrelierten manifesten Variablen eine latente Variable z erzeugt wird, deren Ausprägungen auch negative Vorzeichen aufweisen können. Die Variable z ist hierbei äquivalent mit der Variablen –z.

In den zugrunde liegenden AMOS-Modellen nahmen die interessierenden Pfadkoeffizienten stets Werte (AMOS-Referenzwerte) deutlich größer als Null an und man ging davon aus, dass die aus der Simulation resultierenden Schätzungen für die Pfadkoeffizienten die Referenzwerte annähern. Man konnte also folgern, dass ein simulierter negativer Parameterschätzwert mit überaus hoher Wahrscheinlichkeit ein falsches Vorzeichen habe. Um Verzerrungen der zu berechnenden statistischen Kennwerte zu vermeiden, wurde daher das arithmetische Mittel der Pfadkoeffizienten nicht aus dem in n Stichproben ermittelten Werts für beta, sondern aus dem *Betrag* des ermittelten Werts errechnet. Man musste hierbei nicht fürchten, bei der Betragsbildung einen Fehler zu begehen, indem richtigerweise negatives Vorzeichen tragende Parameterschätzungen in Schätzungen mit positivem umgewandelt werden.

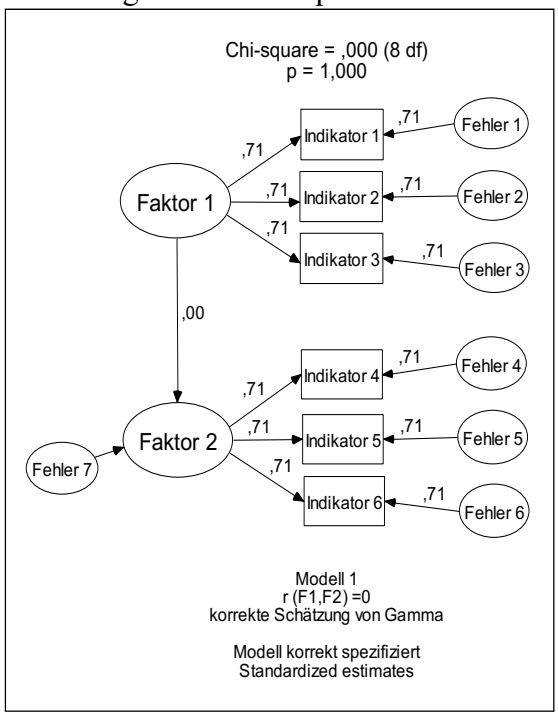
Dieses Verfahren ist bei Simulationen zu Modellen mit interessierenden Pfadkoeffizienten, deren AMOS-Referenzwert nicht deutlich größer als Null ist, oder genauer, deren angenommener Streubereich nicht gänzlich im Bereich positiver Werte liegt, nicht angemessen. Denn die Betragsbildung würde auch solche negativen Schätzwerte in positive transformieren, die nicht auf softwareinterne Vorzeichenumkehr zurückzuführen sind.

Was die Modelle betrifft, deren Faktoren statistisch unabhängig sind, so betragen die AMOS-Referenzwerte für gamma im korrekt spezifizierten Modell 1 Null und in den fehlspezifizierten Modellen 2 und Modell 3 0,057 bzw. 0,067. Daher wird auf die Betragsbildung gänzlich verzichtet (Modell 1) oder abhängig von Verteilungsannahmen über gamma nur dann der Betrag eines Schätzwertes gebildet, wenn der betreffende Wert in einem bestimmten Bereich liegt (Modell 2 und 3). Nähere Erläuterungen hierzu finden sich unter den einzelnen Modellen.

7.2.1 MODELL 1

Im ersten Modell, dem Grundmodell, sind alle Fehlervariablen untereinander unkorreliert und das Modell ist korrekt spezifiziert.

Abbildung 46: Korrekt spezifiziertes Modell 1



Ergebnisse

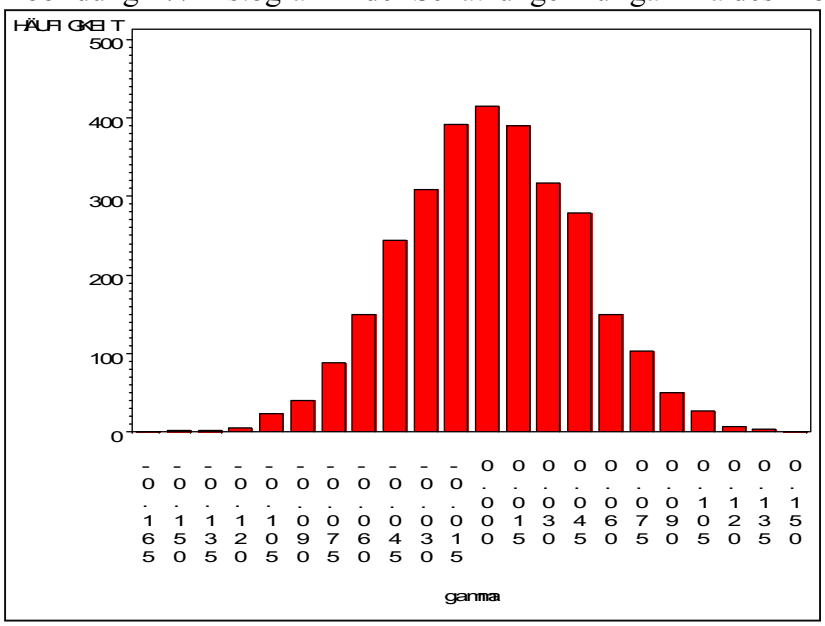
Die Pfadkoeffizientenschätzungen in jedem der beiden Messmodelle in Abbildung 46 entsprechen den Schätzungen, die sich im Grundmodell des Ein-Faktor-Modells mit drei Indikatoren ergaben (siehe 5.3.1).

Das arithmetische Mittel aus 3000 Iterationen für den Pfadkoeffizienten gamma des Strukturmodells beträgt 0,0015, reproduziert damit den im entsprechenden AMOS-Modell errechneten Wert und ist zugleich eine erwartungstreue Schätzung des wahren Wertes Null. Als simulierter Standardfehler und arithmetisches Mittel des geschätzten Standardfehlers von

gamma ergeben sich die Werte 0,0430 bzw. 0,0422, so dass die modellbasierten Standardfehler die simulierten im Mittel sehr gut schätzen. Abbildung 47 zeigt das Histogramm der Schätzungen für gamma und verdeutlicht, dass es sich um eine approximative Normalverteilung handelt.

Die simulierten SE der Pfadkoeffizienten beider Messmodelle betragen – entsprechend der Simulation zum Ein-Faktor-Grundmodell mit drei Indikatoren – ca. 0,025.

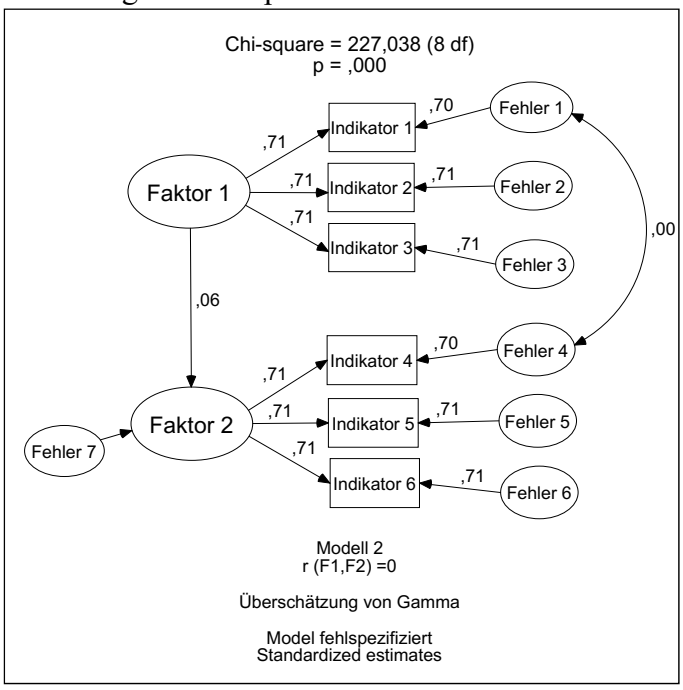
Abbildung 47: Histogramm der Schätzungen für gamma des Modells 1



7.2.2 MODELL 2

Das zweite Modell unterscheidet sich vom vorhergehenden nur dadurch, dass die Fehlervariablen der Indikatoren 1 und 4 eine Korrelation von 0,5 aufweisen ($r_{\varepsilon_1,\varepsilon_4} = 0{,}5$).

Abbildung 48: Fehlspezifiziertes Modell 2



(Die *korrekt spezifizierte* Variante ist in 4.1.2, Abbildung 24 zu sehen.)

Da, wie in Modell 1, nicht nur alle Indikatoren der beiden Messmodelle als (auf den jeweiligen Faktor) bedingt unabhängig angenommen werden, sondern auch für alle Korrelationen zwischen Fehlern verschiedener Messmodelle der Wert Null angenommen wird, ist das Modell fehlspezifiziert.

Im entsprechenden AMOS-Modell wird ein Referenzwert für gamma von 0,057 (dort gerundet mit 0,06 angegeben) errechnet (Abbildung 48).

Ergebnisse

Es wird versucht, entsprechend dem oben erläuterten Vorgehen, einen Wertebereich festzulegen, in dem der Betrag der für gamma resultierenden Schätzwerte zur Berechnung von statistischen Maßzahlen verwendet wird. (Schätzwerte, die nicht in diesen Bereich fallen, gehen als Originalwerte in die Berechnung ein.) Hierbei wird angenommen dass – so wie in Modell 1 – der simulierte Standardfehler von gamma ungefähr dem mittleren geschätzten SE gleicht. Daher wird zunächst statt des simulierten (von der Betragsbildung abhängigen) SE der (von Betragsbildung nicht abhängige) mittlere geschätzte SE in einem Simulationslauf ermittelt. Dieser mittlere geschätzte SE beträgt 0,0423. Es wird weiterhin angenommen, dass gamma, wie in Modell 1, ungefähr normalverteilt sei. Schließlich schätzt man, dass nur ca. 1 % der Werte der Verteilung von gamma unterhalb von 0,057 – 2,326*0,0423 = -0,0414 liegen, die nicht auf das Phänomen der softwareinternen Vorzeichenumkehr zurückzuführen sind. Werte,

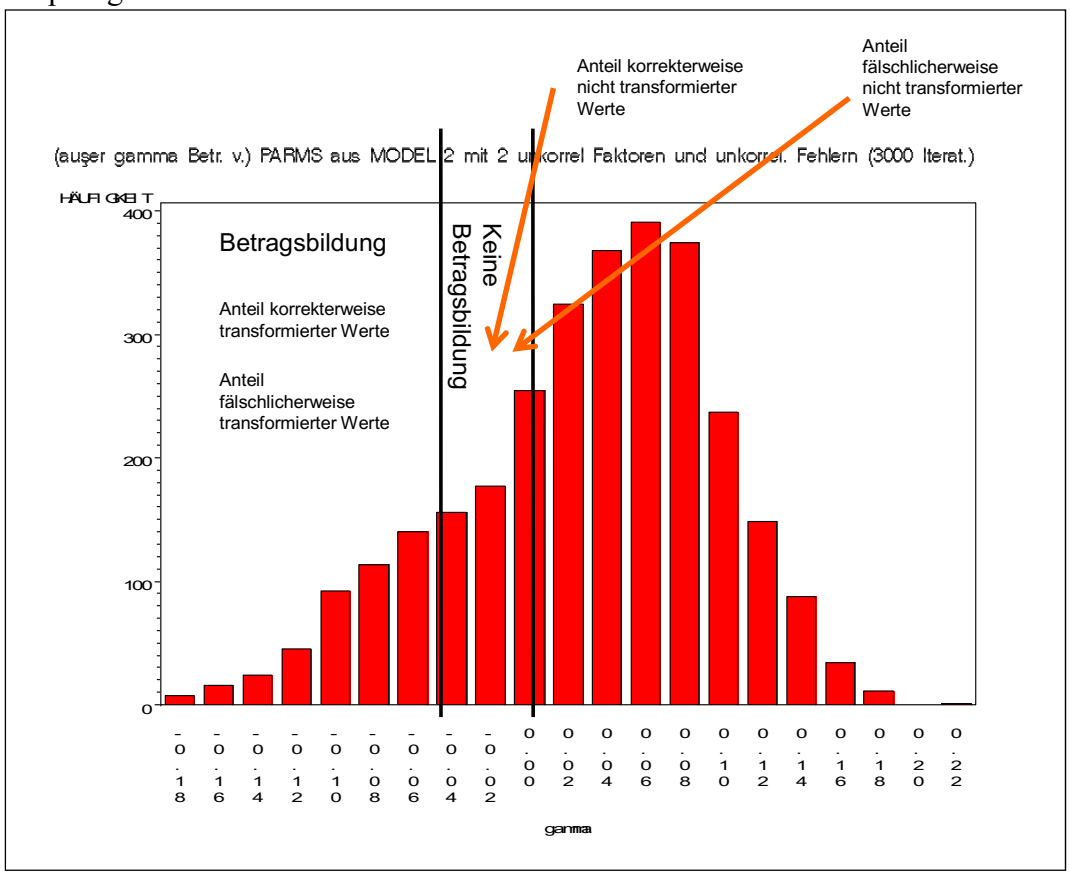
die kleiner als -0,0414 sind, werden dann in positive Werte umgewandelt und gehen zusammen mit den Ursprungswerten (Werte >= -0,0414) in die Berechnung der statistischen Maßzahlen (arithmetisches Mittel, simulierter SE) für gamma ein.

Sind die bei diesem Verfahren getroffenen Annahmen angemessen, so werden nur ca. ein Prozent der korrekten (von der Vorzeichenumkehr nicht betroffenen) Schätzungen für gamma fälschlicherweise mit positivem Vorzeichen versehen. Man kann zudem davon ausgehen, dass der Anteil fälschlicherweise transformierter negativer Werte durch den Teil der beibehaltenen negativen Werte größer als -0,0414 tendenziell ausgeglichen wird, der auf softwareinterne Vorzeichenumkehr zurückzuführen ist.

Neben den auf Grundlage bereichsabhängiger Betragsbildung gewonnenen statistischen Maßzahlen werden am Ende des Abschnitts zu Modell 2 zum Vergleich die statistischen Maßzahlen dokumentiert, die aus den ursprünglichen Schätzwerten für gamma (1) und aus der Verteilung der absoluten Werte von gamma (2) stammen.

Abbildung 49 zeigt das Histogramm der Verteilung von gamma unter Verwendung der *Ursprungswerte*. Dieses ist ergänzt durch den Wertebereich, in dem Betragsbildung vorgesehen ist (links der linken Vertikale) und den Bereich negativer Schätzwerte, die als Ursprungswerte beibehalten werden sollen (Bereich zwischen den Vertikalen).

Abbildung 49: Histogramm der Verteilung von gamma des Modells 2 unter Verwendung der Ursprungswerte

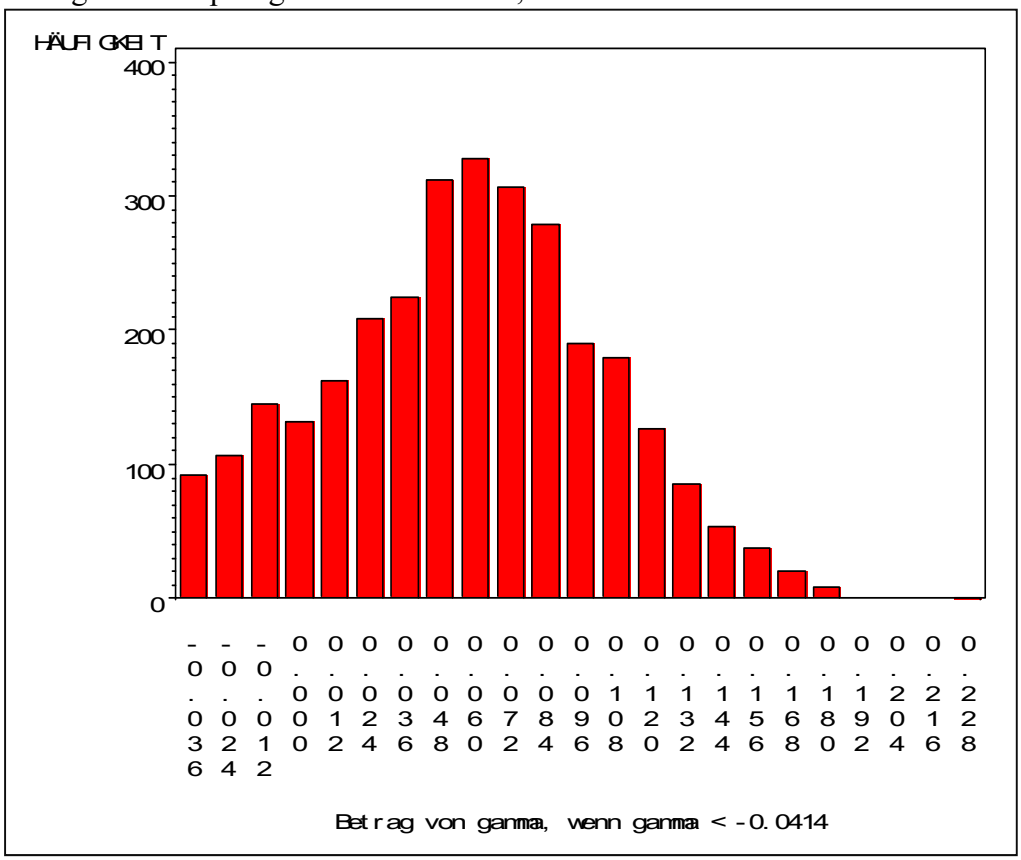


Links der Nulllinie (rechte Vertikale) befinden sich Schätzwerte, die korrekterweise ein negatives Vorzeichen tragen und solche, deren negatives Vorzeichen aufgrund softwareinterner Vorzeichenumkehr zustande gekommen ist. Die Werte unterhalb der linken Vertikalen werden in positive Werte transformiert, jene im Bereich zwischen den Vertikalen nicht. Der Anteil der fälschlicherweise transformierten Werte führt zu einer Erhöhung des arithmetischen Mittels der Schätzwerte. Der Anteil der fälschlicherweise nicht transformierten Werte führt zu einer Verringerung desselben. Diese beiden Effekte heben einander tendenziell auf. Die verzerrenden Auswirkungen auf das arithmetische Mittel, die sowohl von den fälschlicherweise transformierten als auch fälschlicherweise nicht transformierten Werten ausgehen, werden durch diese Egalisierung gedämpft.

Näheren Aufschluss über die Verteilungsform der Ursprungswerte vermittelt ein Histogramm auf Basis einer wesentlich höheren Anzahl von Iterationen (siehe 11.5.2.1.2).

Das Histogramm der Verteilung von gamma unter Verwendung des Betrags von Ursprungswerte kleiner -0,0414 zeigt Abbildung 50:

Abbildung 50: Histogramm der Verteilung von gamma des Modells 2 unter Verwendung des Betrags von Ursprungswerten kleiner -0,0414



Das arithmetische Mittel beträgt 0,0569 und gibt damit den im entsprechenden AMOS-Modell errechneten Wert, 0,057, genau wieder. Der wahre Wert von gamma im Falle des korrekt spezifizierten AMOS-Modells, das die Annahme $r_{\varepsilon_1,\varepsilon_4} = 0{,}5$ beinhaltet, lautet Null. Der Schätzer ist *nicht erwartungstreu*, insofern er den wahren Wert leicht überschätzt. Als simulierter Standardfehler und arithmetisches Mittel des geschätzten Standardfehlers von gamma ergeben sich die Werte 0,0465 bzw. 0,0423, so dass die modellbasierten Standardfehler den simulierten im Mittel leicht unterschätzen.[50]

Da angenommen werden muss, dass der simulierte SE, 0,0465, (aufgrund softwareinternen Vorzeichenumkehr und wertereichsabhängiger Betragsbildung) verzerrt ist, wird bei der Be-

[50] Der mittlere geschätzte SE ist weder vom Artefakt der softwareinternen Vorzeichenumkehr noch von Manipulation durch Betragsbildung in irgendeiner Weise betroffen. Geht man davon aus, dass der mittlere geschätzte SE den simulierten SE ebenso gut schätzt, wie dies in Modell 1 der Fall war, so kann man annehmen, dass der simulierte SE einer Verteilung, aus der das Phänomen der softwareinternen Vorzeichenumkehr vollständig eliminiert wäre, unterhalb des hier ermittelten simulierten SE läge.

rechnung des 95 %-Konfidenzintervalls für $\hat{\gamma}$ vom mittleren geschätzten SE, 0,0423, ausgegangen $\left(\hat{\gamma} \pm 1{,}96 \cdot \overline{\widehat{SE}}(\hat{\gamma})\right)$. Da der wahre Wert innerhalb des 95 %-KI [-0,0260; 0,1398] liegt, ist der Schätzer „arithmetisches Mittel des Pfadkoeffizienten gamma“ nicht verzerrt (kein Bias).

Zum Vergleich werden die statistischen Maßzahlen zu gamma für die anderen beiden o. g. Verfahrensweisen (ohne jegliche Betragsbildung und Betragsbildung bei allen Werten) angegeben: Das arithmetische Mittel und der simulierte SE von gamma betragen 0,0289 bzw. 0,0676, wenn die Werte nicht transformiert werden. Die Maßzahlen nach Betragsbildung über den gesamten Verteilungsbereich von gamma lauten 0,0620 bzw. 0,0395.

Der Mittelwert von gamma in der Variante ohne jegliche Betragsbildung beträgt nur die Hälfte des AMOS-Referenzwertes von 0,057 während der entsprechende simulierte SE von gamma stark über dem mittleren geschätzten SE der hier bevorzugten Variante bereichsabhängiger Betragsbildung liegt. Das arithmetische Mittel von gamma der Variante Betragsbildung über den gesamten Wertebereich, 0,0620, übertrifft den AMOS-Referenzwert nur geringfügig, während der entsprechende simulierte SE, 0,0395, den mittleren geschätzten SE der Variante bereichsabhängiger Betragsbildung geringfügig unterschreitet.

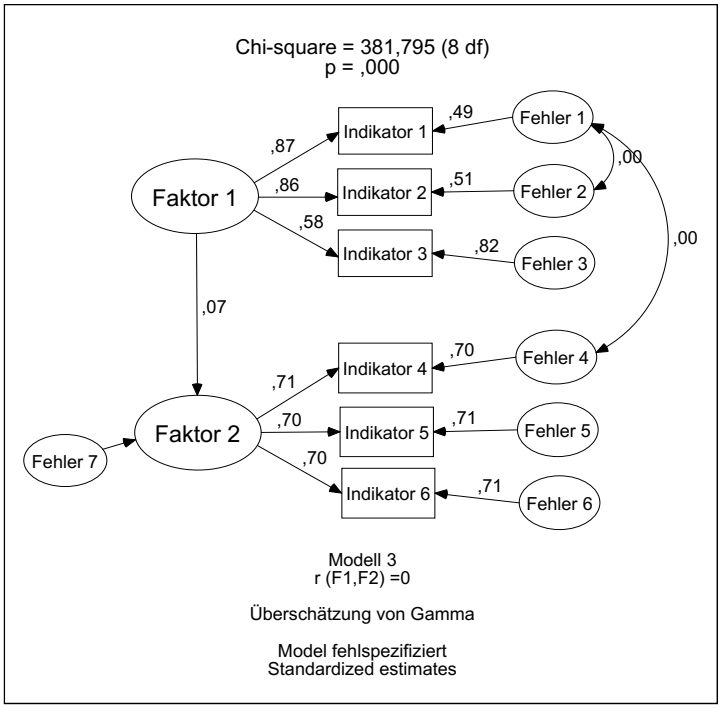
Die simulierten SE der Pfadkoeffizienten beider Messmodelle weisen ungefähr denselben Wert auf wie das Modell 1 (0,025). Ebenso verhält es sich beim Mittelwert der modellbasiert geschätzten SE – im Folgenden kurz „geschätzte SE“ genannt – die in beiden Modellen circa 0,33 betragen.

Eine ignorierte Korrelation zwischen Fehlervariablen, die Indikatoren aus *verschiedenen* Messmodellen angehören, hat mithin in der Situation statistisch unabhängiger Faktoren keinen Einfluss auf die Standardfehler.

7.2.3 MODELL 3

Das dritte Modell weist zusätzlich zum zweiten eine tatsächliche Korrelation der Fehler zweier Indikatoren des ersten Messmodells auf ($r_{\varepsilon_1,\varepsilon_4} = 0{,}5$ und $r_{\varepsilon_1,\varepsilon_2} = 0{,}5$).

Abbildung 51: Fehlspezifiziertes Modell 3
(Die *korrekt spezifizierte* Variante ist in 4.1.3, Abbildung 26 zu sehen.)



Für alle bivariaten Korrelationen zwischen Fehlervariablen der Indikatoren wird der Wert Null angenommen, so dass das Modell fehlspezifiziert ist.

Im entsprechenden AMOS-Modell wird für gamma der Wert 0,067 (dort gerundet angegeben mit 0,07) errechnet (Abbildung 51).

Ergebnisse

Analog zum Vorgehen in Modell 2 wird ein Wertebereich der Verteilung von gamma festgelegt, in dem statt des ursprünglichen Schätzwertes für gamma dessen Betrag zur Berechnung statistischer Kennwerte verwendet wird. Der mittlere geschätzte SE von gamma beträgt 0,0393. Es wird daher geschätzt, dass ca. 1 % der Werteverteilung unterhalb von 0,067 – 2,326 * 0,0393 = -0,0244 liegen, deren Entstehung nicht auf softwareinterne Vorzeichenumkehrung zurückgeht. Schätzungen von gamma < -0,0244 gehen damit betragsmäßig und alle anderen als Ursprungswerte in die Mittelwerts- und Standardfehlerberechnung ein.

Das arithmetische Mittel von gamma unter Anwendung der wertebereichsabhängigen Betragsbildung lautet 0,0701 und reproduziert damit den AMOS-Referenzwert 0,067 zufriedenstellend. Das arithmetische Mittel von gamma ist nicht erwartungstreu, da es den wahren Wert Null leicht und etwas deutlicher als im Vormodell überschätzt. Der simulierte SE, 0,0439, liegt etwas über dem mittleren geschätzten SE. Die arithmetischen Mittelwerte der Varianten

ohne Betragsbildung (0,0356) und mit Betragsbildung über den gesamten Wertebereich (0,0715) liegen weit unter bzw. leicht über dem AMOS-Referenzwert. Der simulierte SE der Variante ohne Betragsbildung, 0,0747, ist fast doppelt so groß wie der mittlere geschätzte SE und daher wenig plausibel. Der simulierte SE der Variante mit allgemeiner Betragsbildung, durch welche die Varianz am stärksten reduziert wird, 0,0416, nähert den mittleren geschätzten SE am besten an.

Wie im Vormodell wird bei der Untersuchung der Frage, ob das arithmetische Mittel von gamma in der Variante wertebereichsabhängiger Betragssetzung ein *unverzerrter Schätzer* des wahren Wertes von gamma, Null, sei, vom mittleren geschätzten SE, 0,0393, ausgegangen. *Da der Erwartungswert innerhalb des 95 %-KI [-0,0069; 0,1471] liegt, gilt der Schätzer (gerade noch) als unverzerrt.*

Die aus der AMOS-Analyse zum Modell 3 resultierenden Pfadkoeffizienten beta1 bis beta6 werden durch die entsprechenden Mittelwerte in der Simulationsanalyse genau getroffen.

In der Simulation zum fehlspezifizierten Ein-Faktor-Modell mit drei Indikatoren und zwei positiv korrelierten Fehlern stellten sich die arithmetischen Mittelwerte der Pfadkoeffizienten als nicht erwartungstreue (und verzerrte) Schätzer heraus. Das dritte hier untersuchte SEM integriert dieses Ein-Faktor-Modell als erstes Messmodell und beinhaltet zusätzlich die Annahme, dass die positive Korrelation zwischen einem Fehler des ersten und einem des zweiten Messmodells, $r_{\varepsilon_1,\varepsilon_4} = 0{,}5$, Null sei.

Die arithmetischen Mittelwerte der Pfadkoeffizienten des ersten Messmodells des hier untersuchten SEMs weichen von jenen im Ein-Faktor-Modell analysierten aufgrund dieser zusätzlichen fehlspezifizierten Korrelation nur geringfügig ab und sind ebenfalls als nicht erwartungstreue und verzerrte Schätzer zu bezeichnen.

Die simulierten SE der Pfadkoeffizienten, die den fehlerkorrelierten Indikatoren (Indikator 1 und 2) des ersten Messmodells zugehören, betragen ca. 0,020 und sind damit bezüglich der restlichen um ca. 20 % vermindert. Dieser Effekt entspricht den Simulationsergebnissen des dem ersten Messmodell entsprechenden Ein-Faktor-Modells. Die entsprechenden mittleren geschätzten SE dagegen betragen ca. 0,30 und sind bezüglich der anderen vier nur um ca. 10 Prozent vermindert. Wenn die Korrelationen $r_{\varepsilon_1,\varepsilon_4} = 0{,}5$ und $r_{\varepsilon_1,\varepsilon_2} = 0{,}5$ fehlspezifiziert

werden, hat dies in der Situation statistisch unabhängiger Faktoren keinen Einfluss auf die Standardfehler der Pfadkoeffizienten des zweiten Messmodells[51].

7.2.4 *Fazit aus den Simulationen zu den Modellen mit zwei statistisch unabhängigen Faktoren*

Man kann davon ausgehen, dass in den Simulationen zu den fehlspezifizierten Modellen 2 und 3 die arithmetischen Mittelwerte von gamma die in den entsprechenden AMOS-Modellen errechneten Abweichungen vom wahren Wert Null genau reproduzieren. Wenn in Modell 3 nur gezeigt werden konnte, dass der Mittelwert von gamma ungefähr mit dem AMOS-Referenzwert übereinstimmt, so ist dies auf das Phänomen der softwareinternen Vorzeichenvertauschung zurückzuführen. Die ergriffene Maßnahme zur Egalisierung dieses Phänomens (Betragsbildung bei einem definierten Teil der negativen Schätzwerte) ist zwar zielführend, aber notwendig fehlerbehaftet.

Die Analyse der SE der Messmodellpfadkoeffizienten bringt keine Ergebnisse hervor, die von den in korrespondierenden Ein-Faktor-Modellen beobachteten abweichen. Im Falle der ignorierten Fehlerkorrelation innerhalb eines Messmodells wurden keine Auswirkungen auf SEs der Pfadkoeffizienten des jeweils anderen Messmodells beobachtet.

7.3 Strukturgleichungsmodelle, deren Faktoren statistisch abhängig sind

Diese aus sechs verschiedenen Modellen bestehende Reihe behandelt den allgemeinen Fall des Strukturmodells, in dem Faktor 2 mit Faktor 1 durch eine gerichtete Abhängigkeit verbunden ist. Der in den entsprechenden AMOS-Analysen im Falle der korrekten Spezifikation der Fehlerkorrelationsstruktur errechnete wahre Wert für gamma beträgt in allen sechs Modellen 0,707. Bis auf das Grundmodell sind alle folgenden Modelle fehlspezifiziert. Da man sich sicher sein kann, dass das Auftreten einer negativen Schätzung für gamma auf softwareinterne Vorzeichenumkehrung zurückgeht, werden die statistischen Maßzahlen für gamma stets aus den Beträgen der geschätzten Werte errechnet.

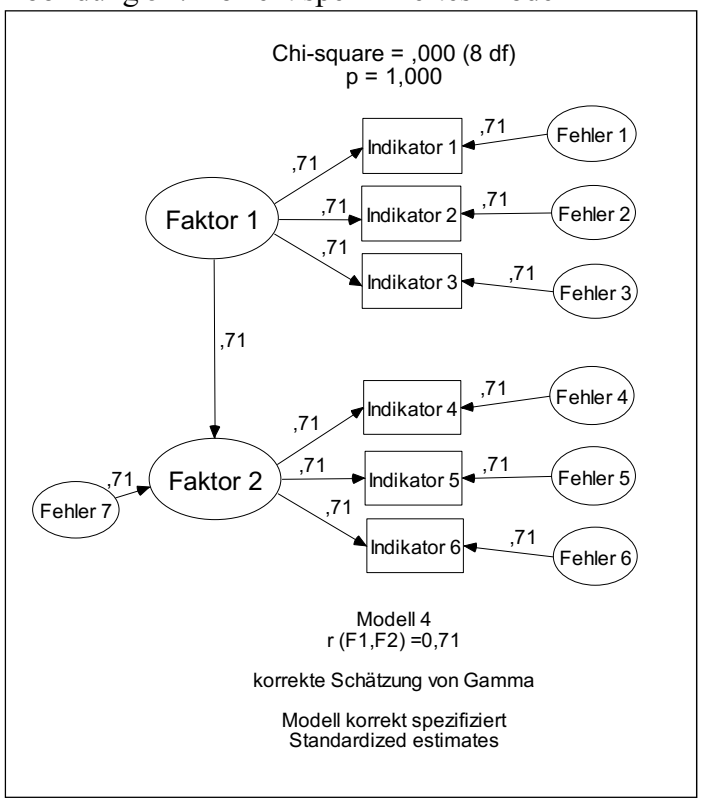
7.3.1 *Modell 4*

Im vierten Modell, dem Grundmodell der zweiten Modellreihe, sind – wie im Grundmodell der ersten – alle Fehlervariablen untereinander unkorreliert. Es unterscheidet sich von letzte-

[51] Es wird davon ausgegangen, dass ein fehlspezifizierte Modell, das nur $r_{\varepsilon_1,\varepsilon_2} = 0,5$ enthielte, keine anderen Ergebnisse hervorbrächte.

rem nur dadurch, dass Faktor 2 von Faktor 1 linear abhängig ist. Die Modellannahme lautet, dass alle bivariaten Korrelationen zwischen Fehlervariablen der Indikatoren Null seien, so dass das Modell korrekt spezifiziert ist.

Abbildung 52: Korrekt spezifiziertes Modell 4



Ergebnisse

Das arithmetische Mittel und der simulierte SE von gamma betragen 0,7109 bzw. 0,0571. Der mittlere geschätzte SE, 0,0562, unterschätzt den simulierten minimal. *Der arithmetische Mittelwert von gamma gibt den im entsprechenden AMOS-Modell (*Abbildung 52*) errechneten Wert genau wieder und ist zugleich eine erwartungstreue Schätzung des wahren Wertes 0,707.*

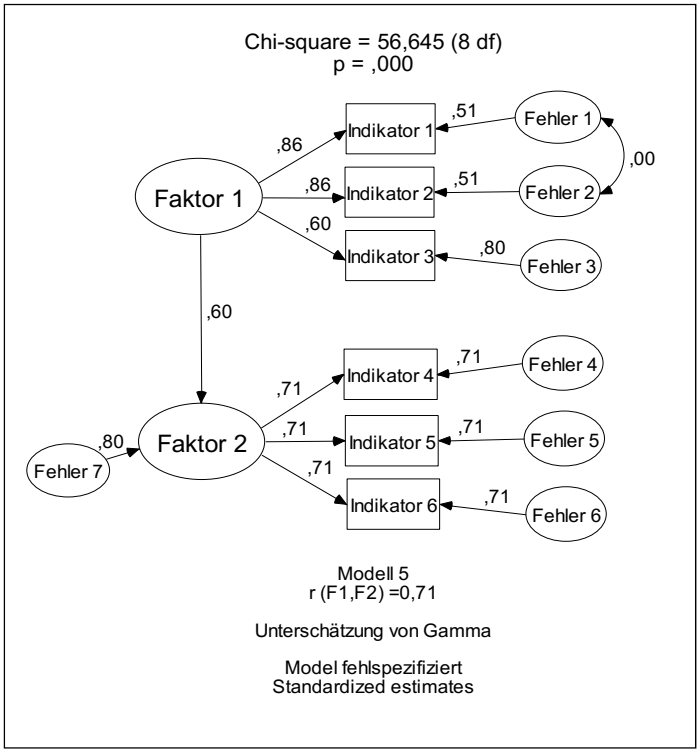
Der simulierte SE der Pfadkoeffizienten des ersten Messmodells beträgt ca. 0,022 und ist damit bezüglich des in den ersten beiden Modellen beobachteten SE um gut 10 % reduziert, während jener des zweiten Messmodells Werte zwischen 0,035 und 0,036 annimmt (Erhöhung um ca. 40 %). Der mittlere geschätzte SE der Pfadkoeffizienten des ersten Messmodells weist den Wert 0,0316 auf und ist somit nur um gut 5 Prozent reduziert, wobei im zweiten Messmodell eine Erhöhung um ca. 20 Prozent auf den Wert 0,0401 zu verzeichnen sind. Die bei den simulierten SE der Messmodellpfadkoeffizienten zu beobachteten Effekte sind in halbierter Stärke somit auch bei den mittleren geschätzten SE zu beobachten. In der Situation

eines SEM, in welchem der zweite Faktor vom ersten abhängig ist (aber der erste nicht vom zweiten[52]), sind also die SEs der Pfadkoeffizienten des ersten Messmodells leicht (simulierte SE) bzw. sehr leicht (geschätzte SE) vermindert und die des zweiten stark (simulierte SE) bzw. mäßig stark (geschätzte SE) erhöht.

7.3.2 MODELL 5

Das fünfte Modell weist eine Fehlerkorrelation zwischen zwei Indikatoren des ersten Messmodells $r_{\varepsilon_1,\varepsilon_2} = 0{,}5$ auf. Die Fehler dieser Indikatoren sind mit allen anderen und letztere untereinander unkorreliert.

Abbildung 53: Fehlspezifiziertes Modell 5
(Die *korrekt spezifizierte* Variante ist in 4.2.2, Abbildung 35 zu sehen)



Ergebnisse

Das arithmetische Mittel des Pfadkoeffizienten gamma, 0,6060, gibt den AMOS-Referenzwert 0,603 (siehe Abbildung 53), fast genau wieder. Der Schätzer ist nicht erwartungstreu, da er den wahren Wert 0,707 deutlich unterschätzt. Der simulierte Standardfehler von gamma beträgt 0,0464 und wird durch den Mittelwert der geschätzten SE (0,0449), wie in den Vormodellen, leicht unterschätzt. *Der Schätzer muss infolge der Fehlspezifikation der*

[52] Faktor 2 ist mit Faktor 1 durch eine gerichtete Abhängigkeit verbunden.

Fehlerkorrelationsstruktur zudem bereits als verzerrt bezeichnet werden, da der wahre Wert 0,707 außerhalb des 95 %-Konfidenzintervalls [0,515, 0,697] liegt.

In der Simulation zum fehlspezifizierten Ein-Faktor-Modell mit drei Indikatoren und zwei positiv korrelierten Fehlern stellten sich die arithmetischen Mittelwerte der Pfadkoeffizienten als nicht erwartungstreue (und verzerrte) Schätzer heraus. Das fünfte hier untersuchte SEM integriert dieses Ein-Faktor-Modell als erstes Messmodell. Die in der Simulation zum Ein-Faktor-Modell resultierenden arithmetischen Mittelwerte der Pfadkoeffizienten bleiben in der Simulation zum fünften SEM im Wesentlichen unverändert: Beta1 und beta2 sinken von 0,866 bzw. 0,867 etwas ab auf 0,862, während in einer Gegenbewegung beta3 von 0,577 auf 0,596 etwas ansteigt. Das Ausmaß der fehlspezifikationsbedingten Abweichung der drei Pfadkoeffizienten von ihren jeweiligen wahren Werten (0,707) wird also bei Vorliegen einer Abhängigkeit des Faktors 2 vom Faktor des fehlspezifizierten Messmodells leicht abgeschwächt.[53]

Der in Modell 3 beobachtete Effekt einer fehlspezifizierten Fehlerkorrelation zwischen zwei Indikatoren des ersten Messmodells auf die SE der zugehörigen Pfadkoeffizienten tritt hier in verstärktem Maß auf: Die simulierten SE dieser Pfadkoeffizienten sind bezogen auf Modell 4, das Grundmodell mit abhängigen Faktoren, um ein gutes Drittel auf 0,014 vermindert. Die entsprechenden mittleren geschätzten SE hingegen sind lediglich um gut 10 % auf 0,028 reduziert. Der Unterschied zwischen den simulierten SE der von der Fehlerkorrelation betroffenen Pfadkoeffizienten und dem SE des nicht betroffenen im erstem Messmodell ist beträchtlich (0,014 und 0,024), während der entsprechende Unterschied bei den mittleren geschätzten SE gering ist (0,028 und 0,031). Die simulierten SE der Pfadkoeffizienten des zweiten Messmodells werden um knapp 20 % auf Werte zwischen 0,29 und 0,030 verringert. Bei den mittleren geschätzten SE ist hier nur eine Reduktion um 12 % auf den Wert 0,0316 zu verzeichnen.

Tritt also in der Situation abhängiger Faktoren eine einzige Fehlerkorrelation innerhalb des Messmodells des unabhängigen Faktors auf, so sind mit Ausnahme des Pfadkoeffizienten des nicht fehlerkorrelierten Indikators desselben Messmodells die SE aller Pfadkoeffizienten deutlich (simulierte) bzw. leicht (mittlere geschätzte) reduziert. (Der Niveauunterschied zwischen

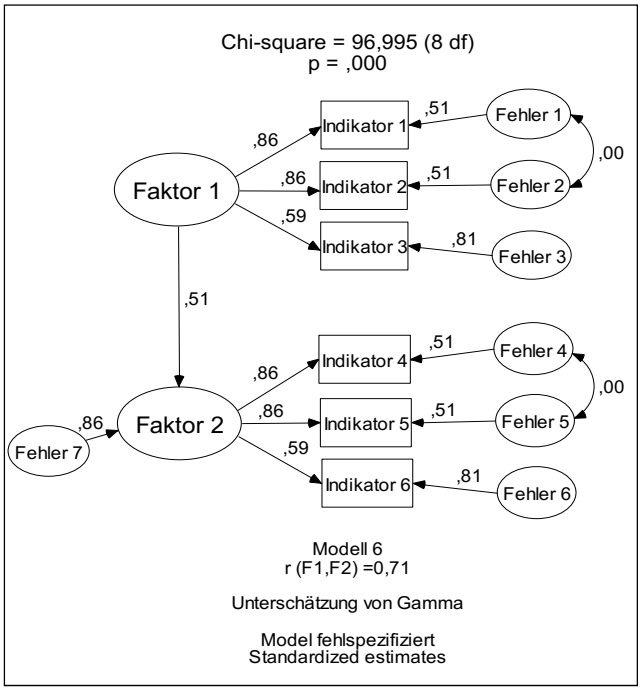
[53] Dieser Sachverhalt stellt sich schon beim Vergleich der entsprechenden AMOS-Modelle (fehlspezifiziertes Ein-Faktor-Modell mit zwei positiv korrelierten Fehlern und fünftes SEM) heraus.

den SE der Pfadkoeffizienten des abhängigen, zweiten Messmodells und jenen des ersten bleibt indes erhalten.)

7.3.3 MODELL 6

Dieses Modell beinhaltet, wie im Vormodell, eine Fehlerkorrelation im ersten Messmodell, $r_{\varepsilon_1,\varepsilon_2} = 0{,}5$ und zusätzlich eine Fehlerkorrelation innerhalb des zweiten, $r_{\varepsilon_4,\varepsilon_5} = 0{,}5$. Die beiden Messmodelle weisen also eine identische Struktur auch hinsichtlich ihrer Fehlerkorrelationen auf. Für alle 15 bivariaten Korrelationen zwischen Fehlervariablen der Indikatoren wird der Wert Null angenommen, so dass das Modell fehlspezifiziert ist.

Abbildung 54: Fehlspezifiziertes Modell 6
(Die *korrekt spezifizierte* Variante ist in 4.2.3, Abbildung 37 zu sehen)



Ergebnisse

Der arithmetische Mittelwert von gamma, 0,5146, reproduziert den in Abbildung 54 gerundet angegeben AMOS-Referenzwert, 0,513, sehr genau und erweist sich als nicht erwartungstreuer Schätzer des wahren Wertes 0,707. Der simulierte Standardfehler von gamma beträgt 0,0399 und wird durch den mittleren geschätzten SE, 0,0387, wie in den Vormodellen leicht unterschätzt. Der Schätzer „arithmetisches Mittel von gamma“ unterliegt einem starken Bias, da der wahre Wert einen großen Abstand zur rechten Grenze des 95 %-Konfidenzintervalls [0,4364; 0,5928] aufweist. Die arithmetischen Mittelwerte der Pfadkoeffizienten des ersten

Messmodells entsprechen sehr genau jenen des Vormodells. Die arithmetischen Mittelwerte der Pfadkoeffizienten des zweiten Messmodells gleichen jenen des ersten.

Die simulierten und mittleren geschätzten SE der Pfadkoeffizienten des ersten Messmodells weisen dieselbe Höhe wie im Vormodell auf, während alle SE des zweiten noch stärker reduziert sind. Im Unterschied zum ersten Messmodell ist im zweiten der SE des Pfadkoeffizienten des nicht fehlerkorrelierten Indikators auch um dasselbe Maß reduziert wie die SE der Pfadkoeffizienten der fehlerkorrelierten.

Bezüglich des Modells 4 wird bei den simulierten SE eine Verringerung um ca. 50 % auf Werte im Bereich zwischen 0,022 und 0,023 beobachtet, während die mittleren geschätzten SE des zweiten Messmodells lediglich eine Reduktion um ein Drittel auf 0,040 aufweisen.

Obwohl also die beiden Messmodelle vollkommen strukturgleich sind, ist es nicht der Fall, dass alle SE der Pfadkoeffizienten des einen Messmodells dieselbe Höhe wie die der korrespondierenden Pfadkoeffizienten des anderen aufweisen. Diese Beobachtung betrifft sowohl die simulierten als auch die mittleren geschätzten SE. (Im ersten Messmodell sind die SE der die fehlerkorrelierten Indikatoren betreffenden Pfadkoeffizienten erwartungsgemäß bezüglich des SE des dritten reduziert. Im zweiten Messmodell ist dieser Effekt nicht zu beobachten)

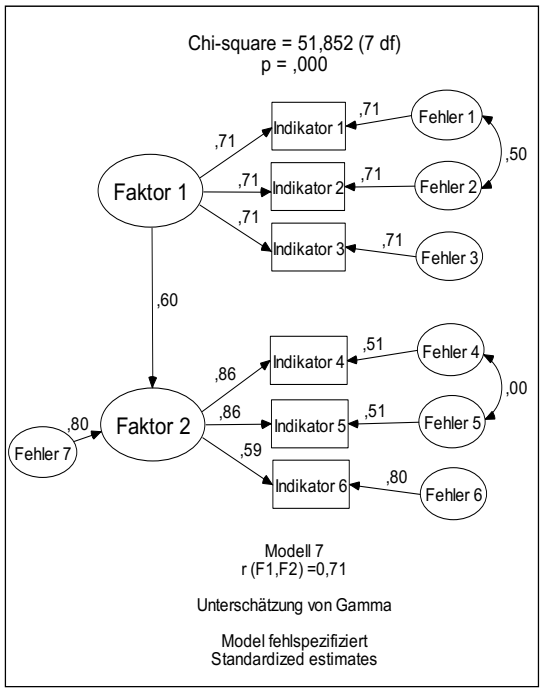
7.3.4 MODELL 7

Die Fehlerkorrelationsstrukturen des siebenten und des sechsten Modells unterscheiden sich nicht. Ersteres unterscheidet sich von dem Vormodell darin, dass für die tatsächliche Fehlerkorrelation des ersten Messmodells, $r_{\varepsilon_1,\varepsilon_2} = 0{,}5$, nicht der Wert Null angenommen wurde.

Diese Korrelation wurde wie die diverser Pfadkoeffizienten geschätzt[54]. Daher ist im siebenten Modell nur noch die Fehlerkorrelation des zweiten Messmodells, $r_{\varepsilon_4,\varepsilon_5} = 0{,}5$ **,** fehlspezifiziert.

[54] Der Schätzung von Fehlerkorrelationen ist bereits ein eigenes Kapitel gewidmet (6.2) und erfolgt in dem hier vorliegenden Kapitel nur in Modell 7.

Abbildung 55: Fehlspezifiziertes Modell 7
(Die *korrekt spezifizierte* Variante ist in 4.2.4, Abbildung 39 zu sehen)



Ergebnisse

Das arithmetische Mittel von gamma beträgt 0,6040 und gibt den in Abbildung 55 gerundet ablesbaren AMOS-Referenzwert 0,602 sehr gut wieder. Da dieser Schätzer den wahren Wert 0,707 deutlich unterschätzt, ist er nicht erwartungstreu. In diesem Modell stimmen der simulierte SE, 0,0483, und der mittlere geschätzte, 0,0481, sehr gut überein. Da der wahre Wert von gamma knapp außerhalb des 95 %-KI des Schätzers [0,5093; 0,6987] liegt, unterliegt der Schätzer einer Verzerrung. Das arithmetische Mittel der im ersten Messmodell geschätzten Fehlerkorrelation beträgt 0,4954 und reproduziert die tatsächliche Fehlerkorrelation[55] sehr gut.

Aus dem ersten Messmodell resultieren, da nicht mehr fehlspezifiziert, nur erwartungstreue und unverzerrte Pfadkoeffizientenschätzer. Die arithmetischen Mittelwerte der Pfadkoeffizienten des zweiten Messmodells hingegen gleichen denjenigen aus Modell 6 und sind weder erwartungstreu noch unverzerrt.

Die Aufhebung der Fehlspezifikation der Fehlerkorrelation im ersten Messmodell (Fehlerkorrelation wird frei geschätzt) zieht folgenden Effekt nach sich: Die simulierten SE aller Pfadkoeffizienten mit Ausnahme des sechsten (Pfadkoeffizient des nicht fehlerkorrelierten Indika-

[55] Der aus dem korrespondierenden AMOS-Modell bezogene wahre Wert der Fehlerkorrelation beträgt 0,5.

tors des zweiten Messmodells) steigen beträchtlich auf ca. 0,029 bzw 0,030 an. Die mittleren geschätzten SE steigen prozentual deutlich geringer an als die simulierten. Im zweiten Messmodell sind die simulierten SE der fehlerkorrelierten Indikatoren um ca. 30 % höher als im Vorgängermodell 6, obwohl das zweite Messmodell in Modell 6 und 7 identische Struktur aufweist. Die Erhöhung der mittleren geschätzten SE beträgt dagegen nur ca. knapp 15 %.

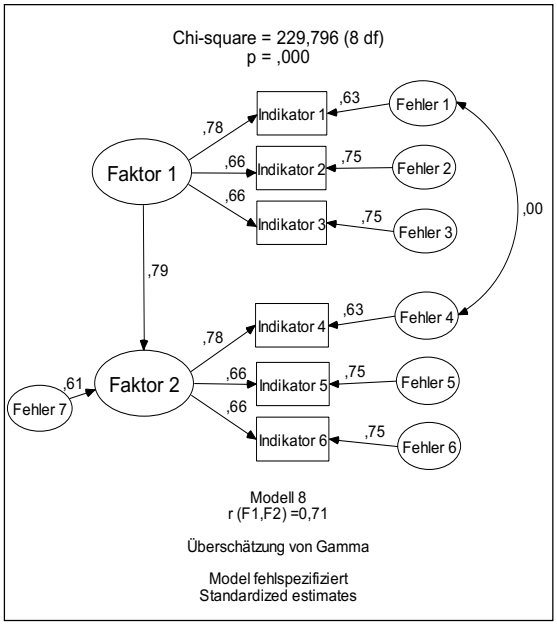
Im ersten Messmodell erhöhen sich die simulierten SE der Pfadkoeffizienten bezüglich des Grundmodells 4 um ca. ein Drittel. Bei den mittleren geschätzten SE ist nur eine Zunahme um ca. knapp 20 % zu beobachten.

Die simulierten SE der Pfadkoeffizienten des zweiten fehlspezifizierten Messmodells verhalten sich mithin gegensätzlich zu den SE im korrespondierenden fehlspezifizierten Ein-Faktor-Modell mit zwei positiv korrelierten Fehlern. Während dort die simulierten SE der die fehlerkorrelierten Indikatoren betreffenden Pfadkoeffizienten geringer sind als der simulierte SE des dritten, verhält es sich im zweiten Messmodell des hier betrachteten SEM genau umgekehrt. Die mittleren geschätzten SE der Pfadkoeffizienten des zweiten fehlspezifizierten Messmodells verhalten sich hingegen so wie die SE im korrespondierenden fehlspezifizierten Ein-Faktor-Modell mit zwei positiv korrelierten Fehlern: Die SE der drei Pfadkoeffizienten weisen jeweils fast dieselbe Höhe auf (Ein-Faktor-Modell: SE der fehlerkorrelierten Pfadkoeffizíenten 0,030 und SE des nicht fehlerkorrelierten Pfadkoeffizienten 0,031; Zwei-Faktor-Modell: SE der fehlerkorrelierten Pfadkoeffizíenten 0,034 und SE des nicht fehlerkorrelierten Pfadkoeffizienten 0,033).

7.3.5 MODELL 8

Dieses Modell unterscheidet sich von Modell 2 nur dadurch, dass die Faktoren der beiden Messmodelle statistisch abhängig sind. Es besteht eine Korrelation zwischen einem Indikator des ersten und einem Indikator des zweiten Messmodells ($r_{\varepsilon_1,\varepsilon_4} = 0{,}5$).

Abbildung 56: Fehlspezifiziertes Modell 8
(Die *korrekt spezifizierte* Variante ist in 4.2.5, Abbildung 41 zu sehen)



Ergebnisse

Das arithmetische Mittel von gamma, 0,7951, trifft den in Abbildung 56 ersichtlichen AMOS-Referenzwert, 0,789, recht genau und ist nicht erwartungstreu, indem es den wahren Wert deutlich überschätzt. Der simulierte SE, 0,0681, der größte unter allen untersuchten Modellen, wird durch den mittleren geschätzten SE, 0,0658, leicht unterschätzt. Da der wahre Wert von gamma innerhalb des 95 %-KI des Schätzers [0,6616; 0,9286] liegt, ist das arithmetische Mittel von gamma kein verzerrter Schätzer.

Die arithmetischen Mittelwerte der Pfadkoeffizienten beta1 bis beta6 der Messmodelle geben die im entsprechenden AMOS-Modell jeweils errechneten Pfadkoeffizienten genau wieder.

Die arithmetischen Mittelwerte der Pfadkoeffizienten des ersten Messmodells gleichen fast jenen des zweiten:

Die Pfadkoeffizienten beta1 und beta4, die die fehlerkorrelierten, jeweiligen ersten Indikatoren der Messmodelle betreffen, sind auf 0,7803, bzw. minimal geringer, auf 0,7783 erhöht. Beide arithmetischen Mittelwerte sind also keine erwartungstreuen Schätzer. Der simulierte SE von beta1 jedoch beträgt 0,0239, während der von beta4 mit 0,0425 fast doppelt so hoch ist. Dies hat zur Folge, dass der Schätzer von beta1 als verzerrt beurteilt werden muss, da das 95 % KI [0,7335; 0,8271] den Wert 0,707 nicht enthält, wohingegen der Schätzer von beta4 als unverzerrt zu gelten hat, da das entsprechende KI [0,6950; 0,8616] den Erwartungswert

einschließt! Der mittlere geschätzte SE von beta4 (0,0491) ist nur um das gut 1,6-fache höher als der von beta1 (0,0305).

Betrachtet man im ersten Messmodell die Pfadkoeffizienten der Indikatoren, deren Fehler unkorreliert sind, beta2 und beta3, und vergleicht sie mit jenen des zweiten, beta5 und beta6, so stellt man fest: Auch hier sind die Mittelwerte der Pfadkoeffizienten des ersten Messmodells, 0,6631 bzw. 0,6640, nur unwesentlich höher als die des zweiten, 0,6628 bzw. 0,6624. Die simulierten SE der Pfadkoeffizienten des ersten Messmodells, 0,0251 bzw. 0,0255, sind jedoch fast nur halb so groß wie die der Pfadkoeffizienten des zweiten, 0,0465 bzw. 0,0460. Die mittleren geschätzten SE der Pfadkoeffizienten des ersten Messmodells (0,0314) betragen dagegen gut zwei Drittel jener des zweiten Messmodells (0,0440). Diese unterschiedlich großen SE in beiden Messmodellen ändern zwar letztlich nichts daran, dass alle vier von der Fehlerkorrelation nicht betroffenen Pfadkoeffizienten keinem Bias unterliegen. Dennoch lässt sich für die Pfadkoeffizienten des ersten Messmodells sagen, dass die Entscheidung knapp ausfällt, da die 95 %-KIs der Mittelwerte [0,6139; 0,7123] bzw. [0,6140; 0,7140] den wahren Wert gerade noch enthalten . Die Entscheidung bei den Pfadkoeffizienten des zweiten Messmodells hingegen ist nicht als knapp zu bezeichnen, da der wahre Wert jeweils großen Abstand zur rechten KI-Grenze hält ([0,5717; 0,7539] bzw. [0,5722; 0,7526]).

Die Simulationsanalyse des achten SEMs führt zu einem auch qualitativ neuen Ergebnis: Die Varianz der Pfadkoeffizientenschätzungen eines Messmodells, dessen Faktor vom Faktor eines anderen Messmodells abhängig ist, ist deutlich höher als die Varianz der Pfadkoeffizientenschätzungen des anderen Messmodells, wenn eine positive Fehlerkorrelation *zwischen* den Messmodellen ignoriert wird. Die Schätzung von Pfadkoeffizienten des zweiten, „abhängigen“ Messmodells ist also deutlich unsicherer als die Schätzung der Pfadkoeffizienten des ersten.

Analyse der SE der Pfadkoeffizienten vor dem Hintergrund der Vormodelle:

Die simulierten SE der Pfadkoeffizienten des ersten Messmodells sind geringfügig größer (Werte zwischen 0,024 und 0,025) als die des ersten Messmodells im Grundmodell mit abhängigen Faktoren (Modell 4). Der Vergleich dieser beiden Teilmodelle hinsichtlich der mittleren geschätzten SE bringt dagegen keinen nennenswerten Unterschied hervor. Die simulierten SE der Pfadkoeffizienten des zweiten Messmodells sind erheblich größer (Werte zwischen 0,042 und 0,046) als die des zweiten Messmodells von Modell 4. Die simulierten SE der von der messmodellübergreifenden Fehlerkorrelation betroffenen Indikatoren sind etwas geringer

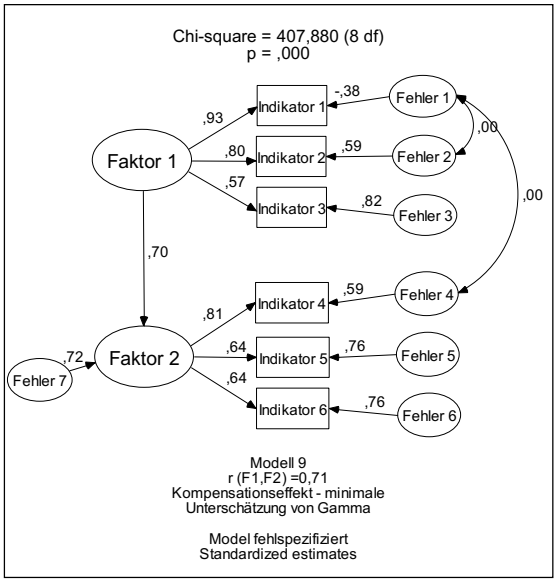
(Indikator 1: 0,024 und Indikator 4: 0,042) als die simulierten SE der Pfadkoeffizienten der nicht betroffenen Indikatoren des jeweiligen Messmodells. Die mittleren geschätzten SE der Pfadkoeffizienten des zweiten Messmodells (0,0440) übertreffen – mit Ausnahme des Pfadkoeffizienten des fehlerkorrelierten Indikators – jene des zweiten Messmodells von Modell 4 (0,0401) in geringerem Maß, als dies bei den simulierten SE der Fall ist. Im Unterschied zu den simulierten SE sind die mittleren geschätzten SE der Pfadkoeffizienten des ersten Messmodells etwa gleich hoch und der mittlere geschätzte SE des von der messmodellübergreifenden Fehlerkorrelation betroffenen Indikators 4 etwas größer (0,0491) als die mittleren geschätzten SE der Pfadkoeffizienten der nicht betroffenen Indikatoren des zweiten Messmodells (0,0440).

Schlussfolgerung: Die simulierten SE der Pfadkoeffizienten des zweiten Messmodells sind um ca. 80 % höher als die des zweiten Messmodells, da der Faktor des zweiten Messmodells abhängig vom Faktor des ersten ist (siehe Modell 4) und zusätzlich eine messmodellübergreifende Fehlerkorrelation ignoriert wird. Die mittleren geschätzten SE repräsentieren die simulierten SE auch in diesem Fall nur unzureichend.

7.3.6 MODELL 9

In diesem letzten Modell der ersten Modellgruppe (siehe 7.1) ist der Fehler des ersten Indikators sowohl mit dem Fehler des zweiten als auch des vierten Indikators korreliert. Daher entspricht es hinsichtlich seiner Fehlerkorrelationsstruktur dem Modell 3 und unterscheidet sich von diesem nur durch die Abhängigkeit der Faktoren. Alle bivariaten Korrelationen zwischen Fehlervariablen der Indikatoren betragen laut Modellannahme Null, so dass das Modell fehlspezifiziert ist.

Abbildung 57: Fehlspezifiziertes Modell 9
(Die *korrekt spezifizierte* Variante ist in 4.2.6, Abbildung 43 zu sehen.)



Ergebnisse

Die beiden Fehlspezifikationen der Fehlerkorrelationen üben einander entgegengesetzte Effekte auf die Schätzung von gamma aus. Die Fehlspezifikation der Fehlerkorrelation innerhalb des ersten Messmodells senkt gamma, während die der messmodellübergreifenden Fehlerkorrelation gamma erhöht. Da beide Effekte fast gleich stark sind, resultiert als arithmetisches Mittel von gamma der Wert 0,6986, der dem AMOS-Referenzwert 0,698 (gerundet ablesbar in Abbildung 57) fast exakt entspricht. Der Mittelwert unterschätzt den wahren Wert 0,707 nur minimal und ist somit ein erwartungstreuer Schätzer. Der simulierte Standardfehler von gamma beträgt 0,0474 und wird durch den Mittelwert der geschätzten SE exakt reproduziert. *Das arithmetische Mittel von gamma ist ein unverzerrter Schätzer, da der wahre Wert von gamma innerhalb des 95 %-KI des Schätzers liegt.*

Die arithmetischen Mittelwerte der sechs Pfadkoeffizienten der Messmodelle treffen die jeweiligen Werte aus dem korrespondierenden AMOS-Modell sehr genau. Die simulierten Standardfehler der Pfadkoeffizienten der ersten beiden Indikatoren sind bezüglich des Grundmodells 4 um ca. 20 % auf ca. 0,018 reduziert (bekannter Effekt), während der des dritten Indikators um dasselbe Maß erhöht ist (0,028). Ein solcher Erhöhungseffekt am Pfadkoeffizienten des von der Fehlerkorrelation nicht betroffenen Indikators konnte in keinem der entsprechenden Vormodelle mit ignorierter Fehlerkorrelation $r_{\varepsilon_1,\varepsilon_2} = 0{,}5$ (Modell 3, 5 und 6) beobachtet werden. Was die mittleren geschätzten SE betrifft, so ist bei den ersten beiden Indikatoren eine Reduktion um nur ca. 10 %, jedoch beim dritten Indikator statt einer Erhö-

hung sogar eine minimale Senkung zu beobachten. Die simulierten SE der Pfadkoeffizienten des zweiten Messmodells befinden sich mit Ausnahme des vierten Indikators (0,035) auf etwas höherem Niveau (0,038) als in Modell 4. Bei den mittleren geschätzten SE verhält es sich gerade umgekehrt: Sie befinden sich – ebenfalls mit Ausnahme des vierten Indikators (0,039) – auf etwas niedrigerem Niveau (0,036) als in Modell 4. Aus dem Vormodell ist bekannt, dass die simulierten SE der Pfadkoeffizienten der von der messmodellübergreifenden Fehlerkorrelation betroffenen Indikatoren bezüglich derjenigen der nicht betroffenen minimal (Messmodell 1) bzw. leicht (Messmodell 2) reduziert sind. Allerdings tritt – offensichtlich infolge einer egalisierenden Wirkung der Fehlerkorrelation $r_{\varepsilon_1,\varepsilon_2} = 0{,}5$ – dieser Effekt im Gegensatz zu Modell 8 hier nur im zweiten Messmodell auf.

Der Vergleich der fehlspezifikationsbedingten Effekte bei den simulierten SE und bei den mittleren geschätzten SE zeigt: Diese Effekte erfolgen in beiden Arten von SE nicht in gleichem Ausmaß und manchmal nicht einmal gleichförmig (Richtung der Veränderung der SE ist entgegengesetzt). Daher besitzt der für den Nutzer linearer Strukturgleichungsmodelle stets verfügbare geschätzte SE nur sehr geringen Informationswert.

Die arithmetischen Mittelwerte der ersten vier Pfadkoeffizienten sind verzerrte Schätzer des wahren Wertes 0,707, da die 95 %-Konfidenzintervalle der Schätzer letzteren nicht enthalten. Die arithmetischen Mittelwerte der letzten beiden Pfadkoeffizienten sind unverzerrt. Die identischen 95 %-KIs dieser Schätzer [0,570; 0,718] enthalten den Wert 0,707 knapp.

7.4 Zusatzmodell und korrekt spezifizierte Modellvarianten, in welchen Fehlerkorrelation als bekannt vorausgesetzt

Die Analyseergebnisse der bisherigen Modelle 1 bis 9 (erste Modellgruppe) führten zu weiteren Fragen: Welche Wirkung auf die SE ist in einem Modell zu beobachten, das eine nur das zweite Messmodell betreffende Fehlerkorrelation aufweist? Welche Effekte auf die SE interessierender Pfadkoeffizienten sind auf die Fehlerkorrelation selbst zurückzuführen und welche auf die Fehlspezifikation einer Fehlerkorrelation? Auf diese Fragen geben zunächst vier weitere Modelle (zweite Modellgruppe) aus der Reihe der Modelle mit abhängigen Faktoren eine Antwort. Es werden zwei Simulationen zu einem Modell durchgeführt, das in 4.2 nur tabellarisch dokumentiert wurde (Zusatzmodell Z4, Tabelle 5 b) und zwar in seiner korrekt (k) und seiner fehlspezifizierten Variante. Zwei weitere Simulationen betreffen korrekt spezifizierte Varianten bisheriger Modelle.

Die korrekte Spezifikation einer Fehlerkorrelation geschieht hier dadurch, dass der aus den AMOS-Referenzmodellen bekannte wahre Wert einer bestimmten Fehlerkorrelation als Information in das Modell aufgenommen wird (Fehlerkorrelationen als bekannt vorausgesetzt). Dieser Situation ist insofern unrealistisch, als eine *externe Information* über die Höhe einer Fehlerkorrelation dem Anwender meist nicht bekannt ist. Gewöhnlich muss eine Fehlerkorrelation aus den vorliegenden Daten geschätzt werden. Diese Modelle besitzen somit für die praktische Anwendung wenig Relevanz und wurden hauptsächlich in das Kapitel aufgenommen, um die in 5.3 angewandte Methode fortzusetzen. (In dem Kapitel über Modelle mit drei Messverfahren ist eine korrekte Spezifikation nur auf diese Weise möglich, da die Schätzung einer Fehlerkorrelation zu Nichtidentifizierbarkeit führen würde.)

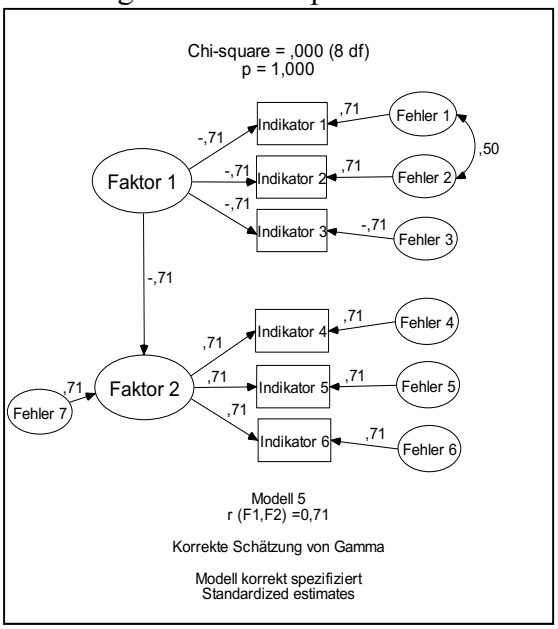
Diese zweite Modellgruppe enthält aufgrund ihrer mangelnden praktischen Relevanz nur drei korrekt spezifizierte Modelle, die a) den Normalfall beschrieben (Modelle mit statistisch abhängigen Faktoren) und b) nur jeweils eine Fehlerkorrelation beinhalten.

Die andere Möglichkeit eine Fehlerkorrelation korrekt zu spezifizieren, die Schätzung dieser Fehlerkorrelation im Modell, wird in 7.5 und 7.6 behandelt

7.4.1 MODELL 5 K (KORREKT SPEZIFIZIERT)

Dieses Modell gibt Antwort auf die Frage nach dem Effekt auf den SE (gamma), wenn ausschließlich eine korrekt spezifizierte messmodellspezifische Fehlerkorrelation im ersten Messmodell gegeben ist ($r_{\varepsilon_1,\varepsilon_2} = 0{,}5$). Das AMOS-Referenzmodell ist in Abbildung 58 wiedergegeben.

Abbildung 58: Korrekt spezifiziertes Modell 5



Ergebnisse

Der simulierte SE von gamma beträgt 0,0657 und übertrifft jenen der fehlspezifizierten Variante um fast 50 %. Die simulierten SE der Pfadkoeffizienten des ersten Messmodells betragen ca. 0,030 und werden von jenen des zweiten (ca. 0,039) um knapp 30 % übertroffen. Die mittleren geschätzten SE des ersten Messmodells (0,042) verhalten sich zu jenen des zweiten (0,032) entsprechend den simulierten SE. Der Vergleich der fehl- und der korrekt spezifizierten Variante von Modell 5 macht deutlich, dass die simulierten SE des ersten Messmodells durch die Fehlspezifikation enorm (teilweise um über 50 %) abfallen, während dieser Effekt bei den mittleren geschätzten SE nur sehr gedämpft (Abfall um nur 15 %) auftritt. Die geschätzten SE sind also nicht geeignet, die auf die Fehlspezifikation zurückgehende tatsächliche Reduktion der SE des direkt betroffenen Messmodells angemessen abzubilden.

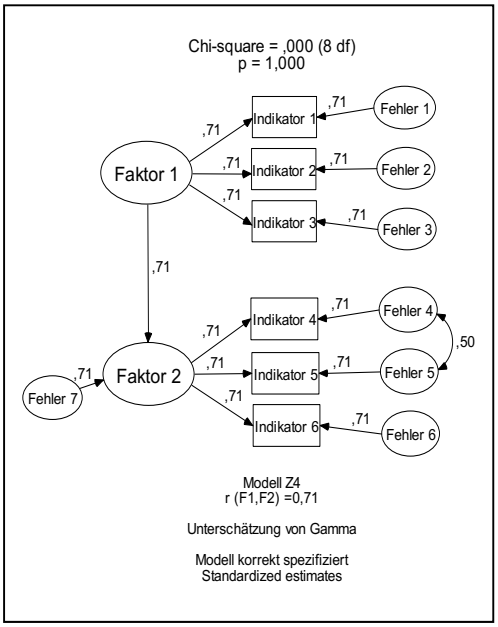
Eine korrekt spezifizierte Fehlerkorrelation zweier Indikatoren, die dem ersten, „unabhängigen“ Messmodell angehören, führt bezüglich des Grundmodells 4 zu einem ca. 25-prozentigen Anstieg der SE aller Pfadkoeffizienten dieses Messmodells und zu einem etwas geringeren Anstieg der SE der Pfadkoeffizienten des zweiten, „abhängigen“ Messmodells.

7.4.2 MODELL Z4 K (KORREKT SPEZIFIZIERT)

Dieses Modell beantwortet die Frage nach der Wirkung auf den SE (gamma), wenn lediglich eine korrekt spezifizierte messmodellspezifische Fehlerkorrelation im zweiten Messmodell vorhanden ist ($r_{\varepsilon_4,\varepsilon_5} = 0{,}5$). Das folgende AMOS-Referenzmodell der Abbildung 59 ist eine

Variation des Modells 5, die ihrerseits in der korrekt spezifizierten Variante im Kapitel 4.2 der AMOS-Modelle mit zwei Faktoren nicht enthalten ist.

Abbildung 59: Korrekt spezifiziertes Modell Z4



Ergebnisse

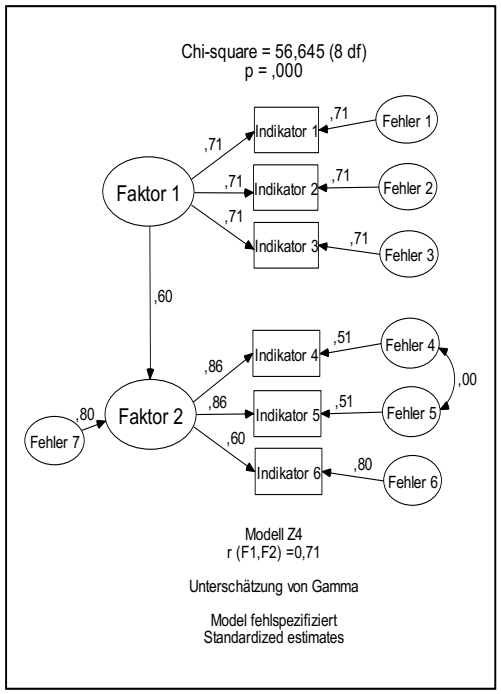
Der simulierte SE (gamma) lautet 0,0665 und ist damit nur unwesentlich größer als im korrekt spezifizierten Modell 5. Die simulierten SE der Pfadkoeffizienten des ersten Messmodells betragen ca. 0,022, während jene der von der Fehlerkorrelation betroffenen Indikatoren des zweiten Messmodells sehr hohe Werte von ungefähr 0,056 annehmen. Der Pfadkoeffizient des nicht betroffenen sechsten Indikators des zweiten Messmodells weist einen SE auf, der in etwa dem des korrespondierenden Pfadkoeffizienten im Vormodell 5 k entspricht (0,040). Die mittleren geschätzten SE weisen mit nur einer Ausnahme dieselbe Höhe auf wie im Grundmodell 4, das keinerlei Fehlerkorrelation enthält. Nur der mittlere geschätzte SE des sechsten, nicht von der Fehlerkorrelation betroffenen Pfadkoeffizienten ist minimal höher (0,044) als die der betroffenen und verhält sich damit entgegengesetzt zum simulierten SE (siehe oben).

Eine korrekt spezifizierte, messmodellspezifische und im „abhängigen“ Messmodell auftretende Fehlerkorrelation hat somit einen sehr starken Effekt auf die SE der betroffenen Pfadkoeffizienten. Der Effekt erstreckt sich aufgrund der gerichteten Abhängigkeit des Faktors des zweiten Messmodells nicht auf die SE der Pfadkoeffizienten des „unabhängigen“ Messmodells.

7.4.3 MODELL Z4 (FEHLSPEZIFIZIERT)

Die Fehlerkorrelation $r_{\varepsilon_4,\varepsilon_5} = 0{,}5$ ist in dieser Variante des Modells Z4 fehlspezifiziert. ($r_{\varepsilon_4,\varepsilon_5} = 0$). Das entsprechende in Abbildung 60 wiedergegebene AMOS-Referenzmodell wird im Kapitel 4.2 nur in der Zusammenfassung 4.2.7 abgehandelt.

Abbildung 60: Fehlspezifiziertes Modell Z4
(Das Vormodell gibt die korrekt spezifizierte Variante wieder.)



Ergebnisse

Der simulierte SE (gamma) dieses Modells ist mit dem des fehlspezifizierten Modells 5 identisch (0,046). Es ist daher unerheblich für die senkende Wirkung auf den SE (gamma), in welchem der beiden Messmodelle eine Fehlerkorrelation auftritt und fehlspezifiziert wird. Die simulierten als auch die mittleren geschätzten SE der Pfadkoeffizienten des ersten Messmodells entsprechen dem Grundmodell 4 und sind daher in keiner Weise betroffen. Beide Arten von SE der Pfadkoeffizienten des zweiten Messmodells entsprechen dem Modell 7, das ebenso die Fehlerkorrelation dieses Modells beinhaltet. Die simulierten SE des zweiten Messmodells sind (bezüglich des Grundmodells 4) deutlich und unterschiedlich stark abgesenkt: Der von der Fehlerkorrelation nicht betroffene Indikator 6 weist mit 0,024 einen SE kaum über dem Niveau der Pfadkoeffizienten des ersten Messmodells auf, während die Pfadkoeffizientenstandardfehler der betroffenen Indikatoren bezüglich des Grundmodells um ca. 20 % redu-

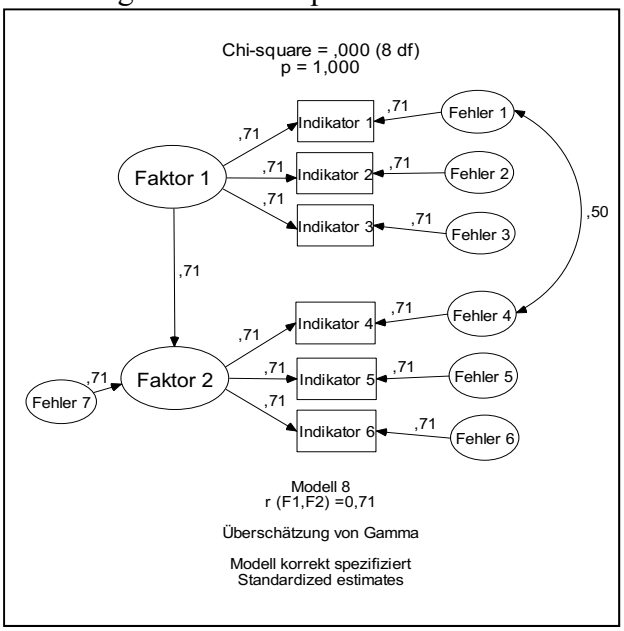
ziert sind (0,027 bzw. 0,028). Die mittleren geschätzten SE des zweiten Messmodells hingegen sind bezüglich des Grundmodells 4 gleichmäßig stark um ca. 20 % abgesenkt.

Der Vergleich der fehl- und der korrekt spezifizierten Variante von Modell Z4 zeigt, dass die simulierten SE des zweiten Messmodells aufgrund der Fehlspezifikation stark (um bis zu 50 %) abfallen, während diese Wirkung bei den mittleren geschätzten SE nur halb so stark (Reduktion um bis zu 25 %) zu Buche schlägt.

7.4.4 MODELL 8 KORREKT SPEZIFIZIERT

Mit dieser Modellvariante kann die Wirkung auf die SE des zweiten Messmodells festgestellt werden, die ausschließlich von einer korrekt spezifizierten messmodellübergreifenden Fehlerkorrelation ausgeht ($r_{\varepsilon_1,\varepsilon_4} = 0{,}5$). Abbildung 61 gibt das entsprechende AMOS-Referenzmodell wieder.

Abbildung 61: Korrekt spezifiziertes Modell 8



Ergebnisse

Der simulierte SE von gamma lautet 0,051 und ist bezüglich des Grundmodells 4 leicht verringert.

Das achte SEM in korrekt spezifizierter Variante bringt simulierte als auch mittlere geschätzte SE der Pfadkoeffizienten des ersten Messmodells hervor, die in etwa dem Modell 4 entsprechen. Lediglich die SE des zweiten Messmodells fallen etwas geringer aus (simulierte 0,031–0,033, mittlere geschätzte 0,035–0,038) als jene im zweiten Messmodell des Modells 4 (simulierte 0,035, mittlere geschätzte 0,040). Der SE des Pfadkoeffizienten des von der Fehlerkor-

relation betroffenen vierten Indikators ist bezüglich jener der nicht betroffenen Indikatoren leicht reduziert. *Daraus folgt, dass bei Messmodellen mit abhängigen Faktoren nicht die messmodellübergreifende Fehlerkorrelation selbst zu einem besonders hohen Anstieg der SE des zweiten Messmodells führt*. (Im Gegenteil: Die Fehlerkorrelation reduziert diese sogar geringfügig). *Vielmehr bewirkt erst die Fehlspezifikation derselben die Standardfehlererhöhung.*

7.5 KORREKT SPEZIFIZIERTE MODELLVARIANTEN MIT UNABHÄNGIGEN FAKTOREN, IN WELCHEN FEHLERKORRELATIONEN GESCHÄTZT WERDEN

In der dritten und letzten Modellgruppe werden alle Modelle der beiden ersten Gruppen mit zwei Faktoren präsentiert, in welchen ein oder zwei Fehlerkorrelationen geschätzt werden. Modelle, in denen Fehlerkorrelationen größer Null vorkommen, werden dadurch in einer der Anwendungssituation angemessenen Weise korrekt spezifiziert. Unter den Modellen mit statistisch unabhängigen Faktoren ist es in Modell 3 aus Gründen der Identifizierbarkeit erforderlich, im ersten Messmodell einen vierten Indikator einzuführen. (Das Pendant dieses Modells unter den Modellen mit statistisch abhängigen Faktoren, das Modell 9, ist ohne zusätzlichen Indikator identifiziert.)

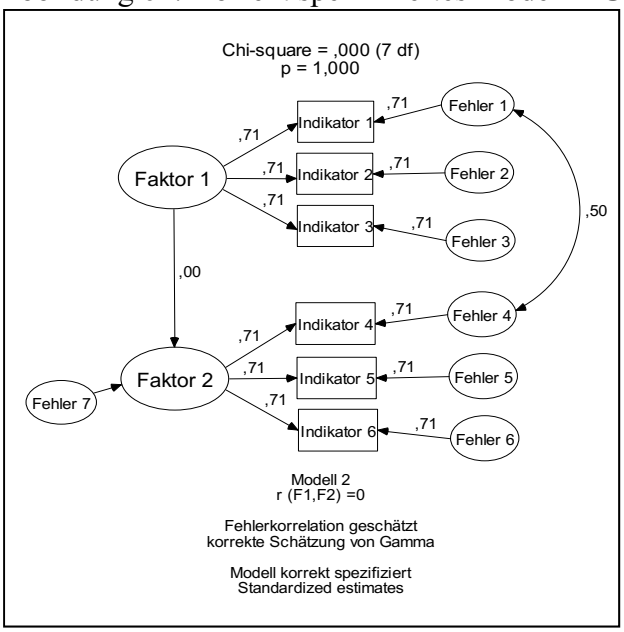
Um ein Modell dieser dritten Gruppe in seiner Kurzbezeichnung von den Modellen der anderen unterscheidbar zu machen, wird es in Kurzform nicht nur mit einem „k“ (für „korrekt spezifiziert“) sondern auch einem „g“ (für „Fehlerkorrelation geschätzt“) versehen. Für alle folgenden Modelle gilt: Die arithmetischen Mittelwerte von gamma, der geschätzten Fehlerkorrelationen und der Messmodellpfadkoeffizienten geben die jeweiligen wahren Werte fast genau wieder und können daher als erwartungstreue Schätzer bezeichnet werden. Im gesamten Kapitel über SEMs mit der Software AMOS (Kap. 4) kommen mit Ausnahme des Modells 7 keine Modelle vor, in welchen eine oder mehrere Fehlerkorrelationen geschätzt werden. Somit stammen sämtliche Grafik-Outputs aus AMOS-Modellen, die erst im Zusammenhang mit diesem Kapitel aus Gründen der Vollständigkeit und der Illustration der Simulationen in SAS konstruiert wurden.

7.5.1 MODELL 2 K_G

Das Modell unterscheidet sich vom fehlspezifiziert Modell 2 dadurch, dass die messmodellübergreifende Fehlerkorrelation ($r_{\varepsilon_1,\varepsilon_4} = 0{,}5$) nicht ignoriert, sondern geschätzt wird. Abbildung 62 gibt das entsprechende AMOS-Referenzmodell wieder. Der Grafik-Output unterscheidet sich von dem der korrekt spezifizierten Variante des Modells 2, in der die Fehlerkor-

relation $r_{\varepsilon_1,\varepsilon_4} = 0{,}5$ als bekannt vorausgesetzt wird, nur dadurch, dass der globale Chi-Quadrat-Anpassungstest mit sieben statt mit acht Freiheitsgraden durchgeführt wird. (In Modell 2 k_g ist die Anzahl der zu schätzenden Parameter um eins größer).

Abbildung 62: Korrekt spezifiziertes Modell 2 G



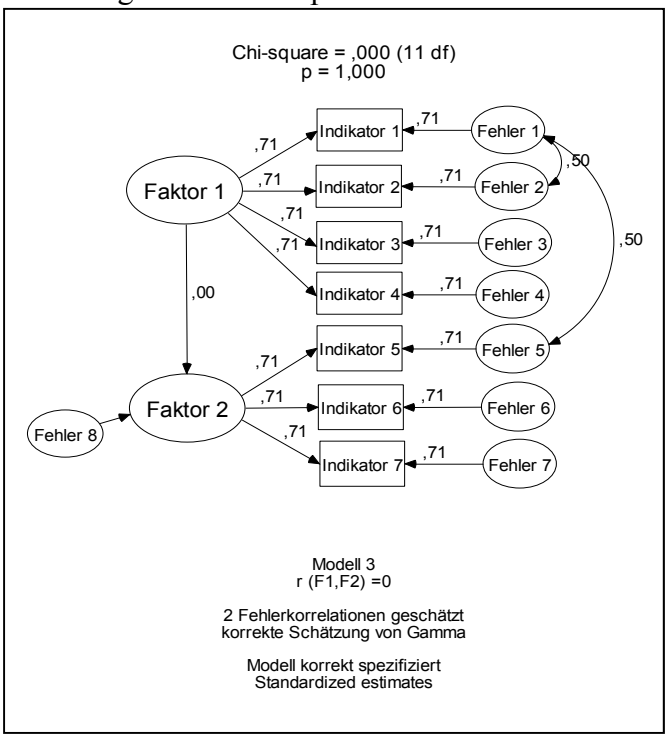
Der simulierte und geschätzte SE der Fehlerkorrelation lautet 0,038 bzw. 0,037.

Der arithmetische Mittelwert des Pfadkoeffizienten des Strukturmodells, gamma, beträgt 0,0004 und ist damit ebenso erwartungstreu. Modell 2 k_g, das Grundmodell 1, das keine Fehlerkorrelationen beinhaltet und das fehlspezifizierte Modell 2 unterscheiden sich hinsichtlich der SE von gamma und der Messmodellpfadkoeffizienten beta1 bis beta6 nur minimal – sowohl was die simulierten als auch die geschätzten SE betrifft.

7.5.2 MODELL 3 K_G

In diesem Modell wird neben der Fehlerkorrelation des Vormodells eine Fehlerkorrelationen innerhalb des ersten Messmodells, $r_{\varepsilon_1,\varepsilon_2}$, geschätzt (Abbildung 63).

Abbildung 63: Korrekt spezifiziertes Modell 3 G



Die Schätzung dieser zwei Fehlerkorrelationen ist in diesem Modell mit drei Indikatoren pro Messmodell nicht möglich, da das Modell nicht identifiziert ist. Zur Begründung sei auf die Ausführungen zum Modell mit drei manifesten Variablen und einem Faktor verwiesen, die auch hier gelten, da die beiden Messmodelle aufgrund unabhängiger Faktoren hinsichtlich der Identifizierbarkeitsfrage separat zu betrachten sind. Die beobachteten Varianzen und Kovarianzen, die zur Schätzung einer Korrelation zwischen Fehlern des ersten Messmodells erforderlich sind, können daher nur aus dem ersten Messmodell stammen. Da drei Varianzen und drei Kovarianzen zur Schätzung von sechs Pfadkoeffizienten und einer Fehlerkorrelation nicht ausreichen, ist das Messmodell unteridentifiziert und damit das Gesamtmodell nicht schätzbar (siehe 2.7). Fügt man einen vierten Indikator dem ersten Messmodell hinzu, so ist das erste Messmodell in einem Grad identifiziert, das dem Modell mit vier manifesten Variablen und einer zu schätzenden Fehlerkorrelation entspricht (siehe 6.3). Mit sieben manifesten Variablen sind beide Fehlerkorrelationen schätzbar. ($\frac{7 \cdot 8}{2} = 28$ beobachteten Parametern stehen 14 zu schätzenden Pfadkoeffizienten, ein Strukturmodellpfadkoeffizienten und zwei Fehlerkorrelationen, also insgesamt 17 Parameter gegenüber, so dass das Gesamtmodell mit 11 Freiheitsgraden überidentifiziert ist.) In einem von 3000 Simulationsdurchläufen resultierte eine Schätzung für gamma von 7,05, da die Kovarianzmatrix für die Schätzer nicht vollen Rang hatte. Nach Eliminierung dieses Durchlaufs betragen die simulierten und geschätzten

SE der Fehlerkorrelationen ca. 0,033 ($r_{\varepsilon_1,\varepsilon_2}$) und 0,030 ($r_{\varepsilon_1,\varepsilon_4}$). Die simulierten und geschätzten SE von gamma und den Messmodellpfadkoeffizienten weichen von jenen des Modells 3 und des Grundmodells nicht nennenswert ab. Die SE der ersten beiden Pfadkoeffizienten, die die messmodellspezifische Fehlerkorrelation betreffen, verhalten sich allerdings im Gegensatz zu Modell 3 so wie die der andern Pfadkoeffizienten.

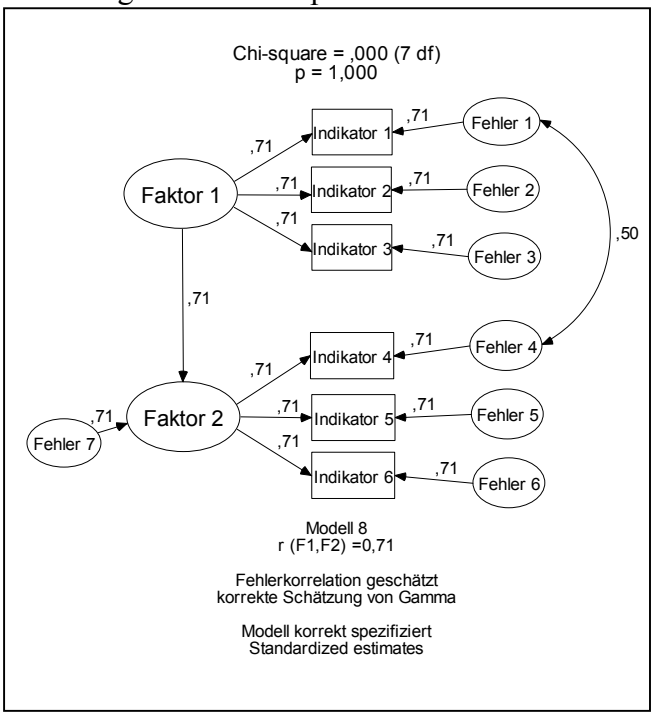
7.6 KORREKT SPEZIFIZIERTE MODELLVARIANTEN MIT LINEAR ABHÄNGIGEN FAKTOREN, IN WELCHEN FEHLERKORRELATIONEN GESCHÄTZT WERDEN

7.6.1 MODELL 8 K_G)

Das Modell beinhaltet eine messmodellübergreifende, geschätzte Fehlerkorrelation ($r_{\varepsilon_1,\varepsilon_4}$).

Analog zu Modell 2_k_g unterscheidet es sich der Grafik-Output in Abbildung 64 von jenem des korrekt spezifizierten Modells 8 nur durch die um 1 reduzierte Anzahl von Freiheitsgraden.

Abbildung 64: Korrekt spezifiziertes Modell 8 G



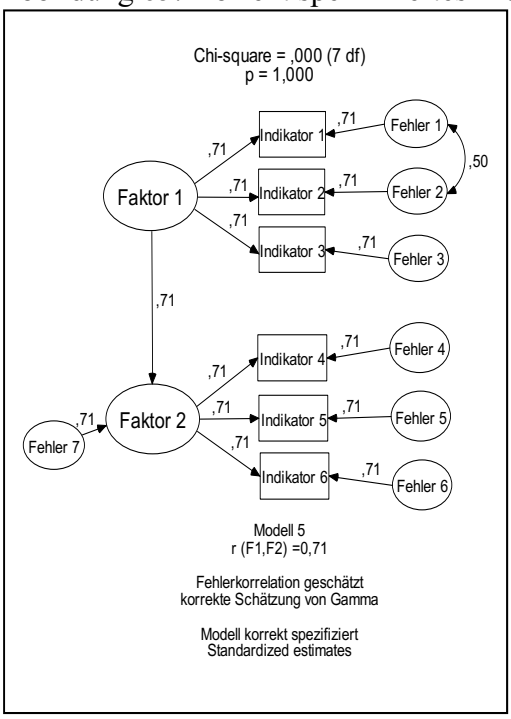
Der simulierte und geschätzte SE der Fehlerkorrelation liegt im Bereich der Vormodellwerte. Für beide Arten von SE gilt: Der SE von gamma (0,053) unterschreitet jenen des Grundmodells 4, das keine Fehlerkorrelation beinhaltet, leicht, und jenen des Modells 8 deutlich. Auch die SE der Pfadkoeffizíenten des zweiten, „abhängigen“ Messmodells weisen einen nur etwas geringeren Wert auf als jene im Grundmodell und unterschreiten jene des Modells 8 um ca.

30 %. Die SE des ersten „unabhängigen“ Modells hingegen verhalten sich so wie in Grundmodell und Modell 8.

7.6.2 MODELL 5 K_G

In diesem Modell wird eine das erste Messmodell betreffende Fehlerkorrelation geschätzt ($r_{\varepsilon_1,\varepsilon_2}$), deren simulierter und geschätzter SE 0,039 beträgt. In Abbildung 65 wird das entsprechende AMOS-Referenzmodell wiedergegeben.

Abbildung 65: Korrekt spezifiziertes Modell 5 G

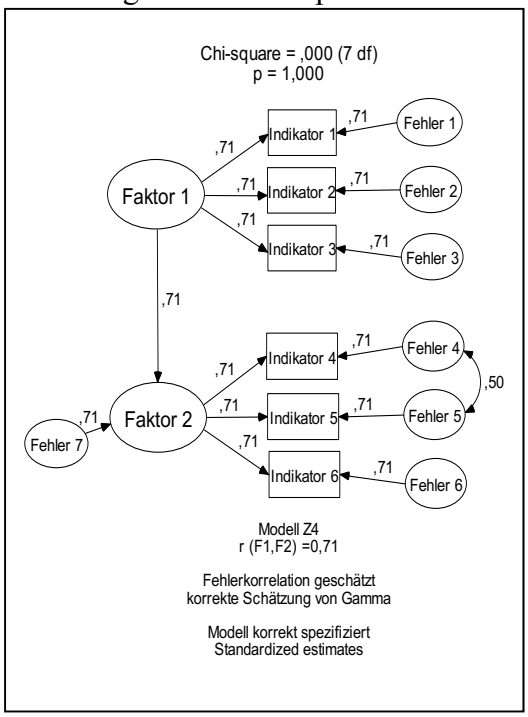


Die simulierten als auch geschätzten SE aller acht geschätzten Parameter liegen über jenen des Grundmodells. Bei den SE des ersten Messmodells ist dieser Unterschied ausgeprägter.

7.6.3 MODELL Z4 K_G

Der Grafik-Output des entsprechenden AMOS-Modells ist Abbildung 66 zu entnehmen. Das Modell beinhaltet eine das zweite Messmodell betreffende geschätzte Fehlerkorrelation ($r_{\varepsilon_4,\varepsilon_5}$), deren simulierte und geschätzte SE 0,040 bzw. 0,039 betragen.

Abbildung 66: Korrekt spezifiziertes Modell Z4 G

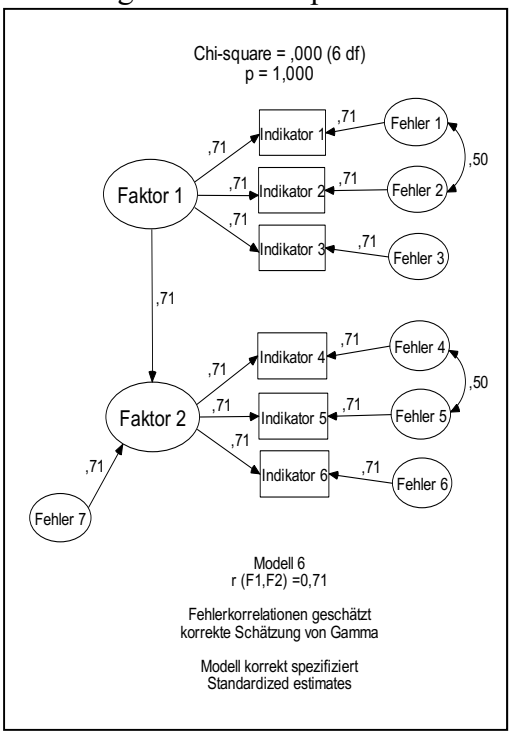


Hinsichtlich beider Arten von SE stimmen die SE der Pfadkoeffizienten des ersten Messmodells mit jenen des fehlspezifizierten Modells Z4 überein. Die simulierten und geschätzten SE von gamma als auch jene der Pfadkoeffizienten des zweiten Messmodells sind in diesem Modell deutlich höher als die im Grund- und fehlspezifizierten Modell gleichhohen SE.

7.6.4 MODELL 6 K_G

In diesem Modell, dessen AMOS-Referenz in Abbildung 67 wiedergegeben ist, werden die Fehlerkorrelationen beider Vormodelle geschätzt.

Abbildung 67: Korrekt spezifiziertes Modell 6 G

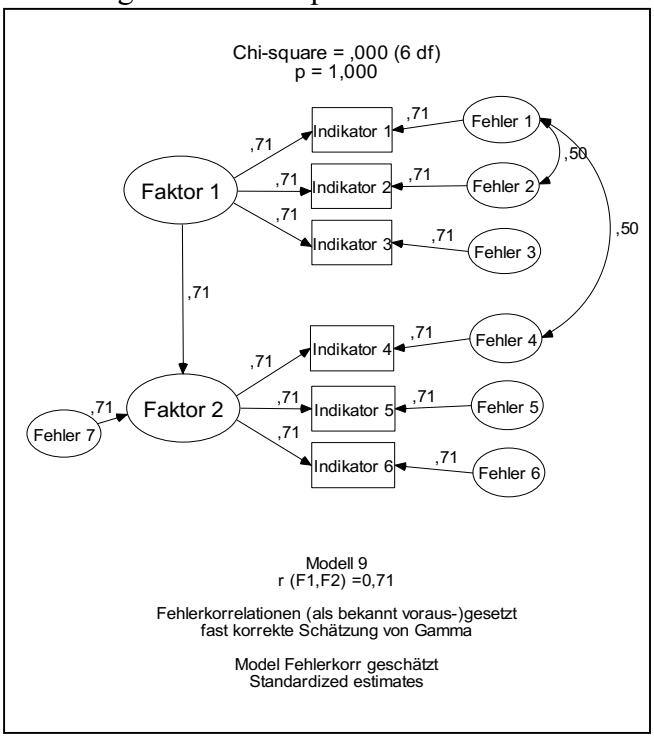


Die simulierten und geschätzten SE beider Fehlerkorrelation liegen zwischen 0,040 und 0,042. Die Untersuchung ergibt, dass die SE der Messmodellpfadkoeffizienten jene des Grundmodells um ca. 30 % (simulierte SE) bzw. 20 % (geschätzte SE) übertreffen. Hinsichtlich beider Arten von SE wird der SE von gamma um ca. 20% % übertroffen. (Im fehlspezifizierten Modell 6 hingegen kommt es bezüglich des Grundmodells – mit Ausnahme des nicht fehlerkorrelierten Indikators des ersten Messmodells – zu einer Absenkung der SE um ca. 30 %.)

7.6.5 MODELL 9 K_G

Dieses Modell beinhaltet eine messmodellspezifische Fehlerkorrelation im ersten Messmodell, $r_{\varepsilon_1,\varepsilon_2}$ **,** und eine messmodellübergreifende Fehlerkorrelation, $r_{\varepsilon_1,\varepsilon_4}$. Abbildung 68 zeigt das entsprechende AMOS-Referenzmodell.

Abbildung 68: Korrekt spezifiziertes Modell 9 G



Aufgrund der linearen Abhängigkeit der Faktoren der beiden Messmodelle ist dieses Modell – im Gegensatz zu Modell 3 k_g mit unabhängigen Faktoren (7.5.2) – identifiziert. Es stehen zur Schätzung der nur das erste Messmodell betreffenden Fehlerkorrelation mehr als nur die sechs Varianzen und Kovarianzen des ersten Messmodells zur Verfügung. Aus den in Modellen mit sechs manifesten Variablen bekannten 22 beobachteten Parametern sind zwölf Pfadkoeffizienten, ein Strukturmodellpfadkoeffizient, der Fehlertermpfadkoeffizient des zweiten Faktors und zwei Fehlerkorrelationen, also insgesamt 16 Parameter zu schätzen (df = 6).

In etlichen Simulationsdurchläufen erfolgt jedoch der Hinweis, dass die Kovarianzmatrix nicht vollen Rang habe. Dies führt in 44 Fällen dazu, dass keinerlei Parameter geschätzt werden. Unter den 2956 Simulationsdurchläufen mit gelungenen Parameterschätzungen kommt es in zahlreichen Fällen aus demselben Grund zu extremen Ausreißern, deren Beibehaltung die Statistik stark verzerren würden (siehe 11.5.2.3.12; nicht ausreißerbereinigte Statistik, (Betr. v.) PARMS). Wenn man 26 dieser Simulationsdurchläufe eliminiert, erhält man auf der Basis von genau 98 % aller Simulationsdurchläufe eine ausreißerbereinigte Statistik (siehe 11.5.2.3.13, (Betr. v.) PARMS).

Zur Ermittlung des Mittelwerts der *geschätzten SE der Parameter* stehen – abhängig vom spezifischen Parameter – nur 2953 bis 2956 Simulationsdurchläufe zur Verfügung (11.5.2.3.12, „STDERR“). Auch in diesen Simulationsdurchläufen sind etliche extreme Aus-

reißer enthalten. Nach Entfernung von zwölf Iterationen verbleiben je nach Parameter 2942 bis 2944 Durchläufe als Basis für die ausreißerbereinigte Statistik (11.5.2.3.13., STDERR).

Es ergeben sich auf Basis der 2930 bzw. 2942–2944 Iterationen für die Fehlerkorrelationen $r_{\varepsilon_1,\varepsilon_2}$ und $r_{\varepsilon_1,\varepsilon_4}$ SE von 0,035 bzw. 0,028, wobei sich simulierte und geschätzte SE nicht unterscheiden. Die SE Pfadkoeffizienten des ersten Messmodells liegen ca. 25 % (simulierten SE) bzw. 10 % (geschätzte SE) oberhalb der entsprechenden SE des Grundmodells. Die simulierten SE der Pfadkoeffizienten des zweiten Messmodells sind bzgl. jener des Grundmodells minimal abgesenkt. Die geschätzten SE der Pfadkoeffizienten des zweiten Messmodells hingegen sind bis auf den SE des fehlerkorrelierten Indikators (minimal abgesenkt) in diesem Modell und dem Grundmodell identisch. Die SE von gamma sind in beiden Modellen identisch. (Die Veränderungen des SE der Messmodellpfadkoeffizienten des *fehlspezifizierten* Modells 9 bezüglich des Grundmodells gingen gerade in die entgegengesetzte Richtung – mit Ausnahme des nicht fehlerkorrelierten Indikators des ersten Messmodells und des fehlerkorrelierten Indikators des zweiten).

7.7 ZUSAMMENFASSUNG UND FAZIT

7.7.1 FAZIT AUS DEN ANALYSEN ZUM PFADKOEFFIZIENTEN DES STRUKTURMODELLS, GAMMA

Unten stehende Tabelle 9 verschafft einen Überblick über die Ergebnisse der zum Pfadkoeffizienten gamma des Strukturmodells angestellten Untersuchungen in allen 20 Modellen

In den ersten neun Modellen (erste Modellgruppe) mit Ausnahme von Modell 7 werden diese Fehlerkorrelationen, soweit vorhanden, sämtlich fehlspezifiziert. Die vier weiteren Modelle (zweite Modellgruppe) sind mit Ausnahme von Modell Z4 korrekt spezifiziert. Die sieben letzten Modelle (dritte Modellgruppe) sind sämtlich in der Weise korrekt spezifiziert, als alle jeweils vorkommenden Fehlerkorrelationen frei geschätzt werden.

Tabelle 9: Simulationen zu Modellen mit zwei Faktoren - Analyseergebnisse zum Strukturmodellpfadkoeffizienten gamma

Modell	Wahrer Wert von gamma, $W(\gamma)$	Fehlerkorrelation	Fehlspezifikation	Arithmetisches Mittel von gamma, $\bar{\hat{\gamma}}$	Simul. SE (gamma), $SE(\hat{\gamma})$	Mittlerer geschätzter SE	$\frac{[\bar{\hat{\gamma}} - W(\gamma)]}{SE(\hat{\gamma})}$	$\bar{\hat{\gamma}}$ signifikant verschieden von $W(\gamma)$?
1	0	-	keine	0,002	0,043	0,042	0,035	Nein: $\bar{\hat{\gamma}}$ trifft $W(\gamma)$ fast
2	0	$r_{\varepsilon_1,\varepsilon_4} = 0{,}5$	$r_{\varepsilon_1,\varepsilon_4} = 0$	0,057	0,046	0,042	1,357*	Nein: $W(\gamma)$ im 95 %-KI
3	0	$r_{\varepsilon_1,\varepsilon_4} = 0{,}5$ $r_{\varepsilon_1,\varepsilon_2} = 0{,}5$	$r_{\varepsilon_1,\varepsilon_4} = 0$ $r_{\varepsilon_1,\varepsilon_2} = 0$	0,070	0,044	0,039	1,795*	Nein: $W(\gamma)$ im 95 %-KI
4	0,707	-	keine	0,711	0,057	0,056	0,070	Nein: $\bar{\hat{\gamma}}$ trifft $W(\gamma)$ fast
5	0,707	$r_{\varepsilon_1,\varepsilon_2} = 0{,}5$	$r_{\varepsilon_1,\varepsilon_2} = 0$	0,606	0,046	0,045	-2,196	Ja: $W(\gamma)$ ausserh. 95 %-KI
6	0,707	$r_{\varepsilon_1,\varepsilon_2} = 0{,}5$ $r_{\varepsilon_4,\varepsilon_5} = 0{,}5$	$r_{\varepsilon_1,\varepsilon_2} = 0$ $r_{\varepsilon_4,\varepsilon_5} = 0$	0,515	0,040	0,039	-4.800	Ja: $W(\gamma)$ ausserh. 95 %-KI
7	0,707	$r_{\varepsilon_1,\varepsilon_2} = 0{,}5$ $r_{\varepsilon_4,\varepsilon_5} = 0{,}5$	teilw. 56 $r_{\varepsilon_1,\varepsilon_2} = 0{,}5$ $r_{\varepsilon_4,\varepsilon_5} = 0$	0,604	0,048	0,048	-2,146	Ja: $W(\gamma)$ ausserh. 95 %-KI
8	0,707	$r_{\varepsilon_1,\varepsilon_4} = 0{,}5$	$r_{\varepsilon_1,\varepsilon_4} = 0$	0,795	0,068	0,066	1,294	Nein: $W(\gamma)$ im 95 %-KI
9	0,707	$r_{\varepsilon_1,\varepsilon_2} = 0{,}5$ $r_{\varepsilon_1,\varepsilon_4} = 0{,}5$	$r_{\varepsilon_1,\varepsilon_2} = 0$ $r_{\varepsilon_1,\varepsilon_4} = 0$	0,698	0,047	0,047	-0,191	Nein: $\bar{\hat{\gamma}}$ trifft $W(\gamma)$ fast
5 k	0,707	$r_{\varepsilon_1,\varepsilon_2} = 0{,}5$	keine	0,709	0,066	0,060	0,030	Nein: $\bar{\hat{\gamma}}$ trifft $W(\gamma)$ genau
Z4 k	0,707	$r_{\varepsilon_4,\varepsilon_5} = 0{,}5$	keine	0,708	0,066	0,059	0,015	Nein: $\bar{\hat{\gamma}}$ trifft $W(\gamma)$ fast
Z4	0,707	$r_{\varepsilon_4,\varepsilon_5} = 0{,}5$	$r_{\varepsilon_4,\varepsilon_5} = 0$	0,605	0,046	0,045	-2,217	Ja: $W(\gamma)$ ausserh. 95 %-KI

[56] Die Fehlerkorrelation $\varepsilon_1, \varepsilon_2$ ist, im Unterschied zu den Fehlerkorrelationen aller anderen Modelle, keine Modellannahme. Sie wird also nicht als bekannt vorausgesetzt, sondern im Modell geschätzt.

* statt [AM (gamma) – wahrer Wert (gamma)] / simulierter SE (gamma) wird in den Modellen mit unabhängigen Faktoren [AM (gamma) – wahrer Wert (gamma)] /*mittlerer geschätzter SE (gamma)* gesetzt, da der simulierte SE aufgrund des Phänomens der softwareinternen Vorzeichenumkehr nicht unverzerrt ist.

Forts. von Tabelle 9 (3. Modellgruppe)

Modell	Wahrer Wert von gamma, $W(\gamma)$	Fehlerkorrelation	Fehlspezifikation	Arithmetisches Mittel von gamma, $\bar{\hat{\gamma}}$	Simul. SE (gamma), $SE(\hat{\gamma})$	Mittlerer geschätzter SE	$\frac{[\bar{\hat{\gamma}} - W(\gamma)]}{SE(\hat{\gamma})}$	$\bar{\hat{\gamma}}$ signifikant verschieden von $W(\gamma)$?
8 k	0,707	$r_{\varepsilon_1,\varepsilon_4} = 0{,}5$	Keine *	0,711	0,051	0,051	0,078	Nein: $\bar{\hat{\gamma}}$ trifft $W(\gamma)$ fast
2 k_g	0	$r_{\varepsilon_1,\varepsilon_4} = 0$	Keine *	0,0004	0,043	0,043	0,009	Nein: $\bar{\hat{\gamma}}$ trifft $W(\gamma)$ fast
3 k_g	0	$r_{\varepsilon_1,\varepsilon_2} = 0{,}5$ $r_{\varepsilon_1,\varepsilon_4} = 0{,}5$	Keine *	0,002	0,042	0,042	0,057	Nein: $\bar{\hat{\gamma}}$ trifft $W(\gamma)$ fast
8 k_g	0,707	$r_{\varepsilon_1,\varepsilon_4} = 0{,}5$	Keine *	0,711	0,053	0,052	0,075	Nein: $\bar{\hat{\gamma}}$ trifft $W(\gamma)$ fast
5 k_g	0,707	$r_{\varepsilon_1,\varepsilon_2} = 0{,}5$	Keine *	0,711	0,062	0,061	0,066	Nein: $\bar{\hat{\gamma}}$ trifft $W(\gamma)$ fast
Z4 k_g	0,707	$r_{\varepsilon_4,\varepsilon_5} = 0{,}5$	Keine *	0,711	0,062	0,061	0,066	Nein: $\bar{\hat{\gamma}}$ trifft $W(\gamma)$ fast
6 k_g	0,707	$r_{\varepsilon_1,\varepsilon_2} = 0{,}5$ $r_{\varepsilon_4,\varepsilon_5} = 0{,}5$	Keine *	0,711	0,066	0,066	0,060	Nein: $\bar{\hat{\gamma}}$ trifft $W(\gamma)$ fast
9 k_g	0,707	$r_{\varepsilon_1,\varepsilon_2} = 0{,}5$ $r_{\varepsilon_1,\varepsilon_4} = 0{,}5$	Keine *	0,709	0,057	0,057	0,035	Nein: $\bar{\hat{\gamma}}$ trifft $W(\gamma)$ fast

Siehe umseitige Fußnote 56.

Der Vergleich der mittleren modellbasiert geschätzten Standardfehler mit den simulierten Standardfehlern, die eine höhere Aussagekraft besitzen, offenbart, dass man sich im Allgemeinen auf die Schätzung der modellbasiert geschätzten SE von gamma verlassen kann. Sie erlauben eine verlässliche Angabe über den Bereich, in dem sich der wahre Wert von gamma mit hoher Wahrscheinlichkeit befindet.

Alle folgenden Ausführungen (Punkt 1–6 unter 7.7.1, Punkt 1–8 unter 7.7.2) beziehen sich grundsätzlich auf die Analyse des *simulierten SE*.

1. Der simulierte SE von gamma im Grundmodell für abhängige Faktoren (Modell 4, 0,057) ist etwas höher als der des Grundmodells für unabhängige Faktoren (Modell 1, 0,043).
2. In *Modellen mit unabhängigen Faktoren* bewirken fehlspezifizierte Fehlerkorrelationen zwar eine deutliche, wenn auch nicht signifikante Abweichung vom wahren Wert des Strukturmodellpfadkoeffizienten gamma (Null). Sie bewirken jedoch keinen bemerkenswerten Anstieg der SE von gamma.

Zu den Modellen mit abhängigen Faktoren:

3. Eine *korrekt spezifizierte Fehlerkorrelation zweier Indikatoren, die demselben Messmodellen angehören* ($r_{\varepsilon_1,\varepsilon_2} = 0{,}5$ oder $r_{\varepsilon_4,\varepsilon_5} = 0{,}5$), lässt den SE (gamma) bezüglich des Grundmodells 4 *deutlich ansteigen (*Modelle 5 k und Z4 k). Das Ignorieren einer solchen messmodellspezifischen Fehlerkorrelation führt zu einer deutlichen Verringerung des SE. Hierbei ist es gleichgültig, ob die Fehlerkorrelation im ersten oder zweiten Messmodell vorkommt (Modelle 5 und Z4).
4. Wird nicht nur in einem Messmodell sondern *in beiden Messmodellen eine derartige Fehlerkorrelation ignoriert*, resultiert eine zusätzliche fast gleichgroße Abnahme des SE von gamma (siehe Modell 6 im Vergleich zu Modell 5). In Modell 6 ist der SE von gamma daher der geringste aller Modelle, einschließlich der mit unabhängigen Faktoren ($SE(\hat{\gamma}) = 0{,}040$).
5. Eine korrekt spezifizierte Fehlerkorrelation zweier Indikatoren, die verschiedenen Messmodellen angehören, führt zu einer leichten Verringerung des SE von gamma (Modell 8 k)
6. Das Ignorieren einer solchen messmodellübergreifenden Fehlerkorrelation führt zu einer deutlichen Erhöhung des SE (gamma) (Modell 8). Wird eine *messmodellspezifische mit einer messmodellübergreifenden Fehlerkorrelation kombiniert* (Modell 9), so wird die senkende Wirkung ersterer durch die erhöhende letzterer nicht egalisiert. Modell 5, das nur die messmodellspezifische Fehlerkorrelation enthält, weist nämlich nur einen minimal geringeren simulierten SE (gamma) auf (0,046) als Modell 9 (0,047). Das heißt, eine ignorierte messmodellspezifische Fehlerkorrelation ist hin-

sichtlich ihrer Wirkung auf den SE des Strukturmodellpfadkoeffizienten „dominant“, eine ignorierte messmodellübergreifende Fehlerkorrelation hingegen „rezessiv“.

Ein Bias des Strukturmodellpfadkoeffizienten ist nur in fehlspezifizierten Modellen gegeben, die ausschließlich eine oder zwei, messmodellspezifische Fehlerkorrelation aufweisen und daher gamma in hohem Maß absenken. In diesen Modellen ist der Betrag von $\frac{[\bar{\hat{\gamma}} - E(\gamma)]}{SE(\hat{\gamma})}$ deutlich größer als Zwei. Einen besonders großer Bias ist in dem Modell mit zwei messmodellspezifischen Fehlerkorrelationen zu beobachten (Modell 6; $\frac{[\bar{\hat{\gamma}} - E(\gamma)]}{SE(\hat{\gamma})} = -4.8$).

7.7.2 Fazit aus den Analysen zu den simulierten Standardfehlern der Pfadkoeffizienten der Messmodelle, beta1 bis beta6

Im Allgemeinen ist zu beobachten, dass die modellbasierten Standardfehler von beta 1 bis beta 6 eher zu hoch geschätzt, also konservativ sind.

Im Einzelnen gilt:

1. Die lineare Abhängigkeit des Faktors des zweiten Messmodells vom Faktor des ersten erhöht die SE der Pfadkoeffizienten des Messmodells des abhängigen Faktors um ca. 50 %, während es die des unabhängigen Faktors um ca. 10 % senkt (Vergleich der korrekt spezifizierten Modelle 1 und 4).

Zu den Modellen mit abhängigen Faktoren:

2. Wenn eine *korrekt spezifizierte Fehlerkorrelation innerhalb des ersten Messmodells* gegeben ist (Modell 5 k) führt dies bezüglich des Grundmodells 4 zu einer deutlichen Erhöhung der SE de Pfadkoeffizienten des ersten Messmodells (um ca. ein Drittel) und einer moderaten Erhöhung innerhalb des zweiten (um gut zehn Prozent). In der Simulationsanalyse des Ein-Faktor-Modells mit drei Indikatoren ergab sich im Gegensatz zum Effekt im ersten Messmodell, dass die SE der Pfadkoeffizienten fehlerkorrelierter Indikatoren bei korrekter Spezifikation (wahre Fehlerkorrelation als bekannt vorausgesetzt) nicht größer sind als die SE im Modell ohne Fehlerkorrelationen, während jener des nicht fehlerkorrelierten Indikators deutlich anstieg.
3. Das *Ignorieren dieser messmodellspezifischen Fehlerkorrelation* innerhalb des ersten Messmodells führt bezüglich des Grundmodells zu einer deutlichen Verringerung der SE der Pfadkoeffizienten der fehlerkorrelierten Indikatoren und zu einer minimalen Verringerung desjenigen des nicht fehlerkorrelierten Indikators. (Erstere werden mehr

als halbiert, während letzterer um gut 20 % reduziert wird.) Dieser Effekt entspricht tendenziell dem bereits aus den Simulationen zum Ein-Faktor-Modell bekannten Effekt. Weiterhin führt das Ignorieren dieser Fehlerkorrelation zu einer Verringerung der SE der Pfadkoeffizienten des zweiten Messmodells (Modell 5).

4. Eine *korrekt spezifizierte*[57]*, messmodellspezifische und im zweiten „abhängigen" Messmodell auftretende Fehlerkorrelation* zieht bezüglich des Grundmodells eine Erhöhung des SE der Pfadkoeffizienten der fehlerkorrelierten Indikatoren um ca. 60 % nach sich. (Diese Erhöhung ist die größte, die an den untersuchten 13 Modellen beobachtet wurde. Sie bedeutet eine erhebliche Verringerung der Präzision der Pfadkoeffizientenschätzung.) Der SE des Pfadkoeffizienten des nicht fehlerkorrelierten Indikators steigt lediglich um gut 10 Prozent an (Modell Z4 k). (Die SE der Pfadkoeffizienten des Messmodells 1 bleiben unverändert.)

5. Die *Fehlspezifikation dieser Fehlerkorrelation* (es wird fälschlicherweise eine Fehlerkorrelation von Null angenommen) halbiert zum einen die SE der Pfadkoeffizienten der fehlerkorrelierten Indikatoren, die in der korrekt spezifizierten Variante auftraten. (Bezüglich des Grundmodells 4 kommt die Halbierung einer Reduktion um ca. 25 % gleich). Die Fehlspezifikation reduziert zum anderen den aus der korrekt spezifizierten Variante resultierenden SE des Pfadkoeffizienten des nicht korrelierten Indikators um 40 %. (Bezogen auf das Grundmodell 4 entspricht dies einer Reduktion um ca. ein Drittel.)

6. Werden *im ersten und zugleich im zweiten Messmodell messmodellspezifische Fehlerkorrelationen ignoriert*, so zieht dies bezüglich des Grundmodells 4 eine Verringerung der SE aller Pfadkoeffizienten des zweiten Messmodells um ein gutes Drittel nach sich (Modell 6).

7. Das Auftreten einer *korrekt spezifizierten Fehlerkorrelation zwischen verschiedenen Messmodellen* erhöht die SE der Pfadkoeffizienten des zweiten Messmodells bezüglich des Grundmodells nicht – im Gegenteil, es reduziert die Pfadkoeffizientenstreuung des zweiten „abhängigen" Messmodells geringfügig (Modell 8 k).

8. Das *Ignorieren dieser messmodellübergreifenden Fehlerkorrelation* erhöht die SE der Pfadkoeffizienten des zweiten Messmodells bezüglich des Grundmodells 4 deutlich

[57] Korrekt spezifiziert in dem Sinn, dass die wahre Fehlerkorrelation im Modell als bekannt vorausgesetzt wird.

(Modell 8). (Die SE der Pfadkoeffizienten des zweiten Messmodells sind damit um 80 Prozent höher als die des ersten Messmodells.) Die Schätzung von Pfadkoeffizienten des zweiten, „abhängigen“ Messmodells ist also deutlich unpräziser als die Schätzung der Pfadkoeffizienten des ersten.

7.7.3 FAZIT AUS DEM VERGLEICH DER SIMULIERTEN UND MITTLEREN GESCHÄTZTEN SE DER PFADKOEFFIZIENTEN DER MESSMODELLE

Der Vergleich der beiden Arten von Standardfehlern der Pfadkoeffizienten, die in dieser Arbeit analysiert werden, der simulierten und die mittleren geschätzten SE, führte zu folgenden Resultaten:

1. Bestimmte Gruppen von Messmodellpfadkoeffizienten (zum Beispiel die des Messmodells 1 oder die des Messmodells 2, sofern diese nicht fehlerkorrelierten Indikatoren zugehören) weisen bei den *mittleren geschätzten SE eine wesentlich größere Homogenität* auf als bei den simulierten.
2. Beim Vergleich von Modellen wurde festgestellt, dass Änderungen in der Modellstruktur (Einführung von korrekt oder fehlspezifizierten Korrelationen sowie Herstellung linearer Abhängigkeit zwischen messmodellspezifischen Faktoren) zu Veränderungen in den SE bestimmter Pfadkoeffizienten führen. Es sind folgende Phänomene der Veränderung von SE zu unterscheiden:

a) Die Veränderungen, die bei den simulierten SE beobachtet werden können, gehen in die gleiche Richtung wie jene, die an den mittleren geschätzten zu beobachten sind und sind weitaus größer als letztere. Häufig ist die bei den mittleren geschätzten SE festgestellte Veränderung ziemlich genau halb so groß wie die bei den simulierten SE beobachtete. Beispiele für einen erheblichen Größenunterschied in der Veränderung der beiden SE-Arten finden sich in folgenden Modellen:

- Modell 3 (Fehlspezifikation)
- Modell 4 (Lineare Abhängigkeit der Faktoren)
- Modell 5 (Fehlspezifikation)
- Modell 6 (Fehlspezifikation)
- Modell 7 (Fehlspezifikation)
- Modell 8 (Fehlspezifikation)
- Modell 9 (Fehlspezifikation)
- Modell 5k (korrekte Spezifikation)

b) Veränderungen, die bei einer der beiden Arten von SE beobachtet werden können, werden bei der jeweils anderen nicht beobachtet oder finden in entgegengesetzter Richtung statt. Beispiele dafür finden sich in Modell:

- Modell 8
- Modell 9
- Modell Z4 k
- Modell Z4 (alle außer Z4 k enthalten fehlspezifizierte Fehlerkorrelationen)

c) Die in beiden Arten von SE beobachteten Veränderungen der SE entsprechen einander ungefähr.

- Ein Beispiel dafür findet sich nur in Modell 8k (korrekte Spezifikation).

Der Nutzer kann a priori nicht entscheiden, in welche der drei oben genannten Kategorien von Veränderung der SE das von ihm untersuchte Modell fällt. In den allermeisten Fällen ist jedoch davon auszugehen, dass der unter a) beschriebene Sachverhalt gegeben ist. Die mittleren geschätzten Standardfehler reagieren auf Fehlspezifikation häufig halb so stark wie die simulierten Standardfehler. Führt der Nutzer einen Vergleich zweier Modelle durch so kann er *in der Regel* davon ausgehen, dass der durch die Fehlspezifikation induzierte Veränderungseffekt tatsächlich deutlich größer ist. Die mittleren geschätzten SE können letztlich die simulierten nicht substituieren und bieten keine sichere Handlungsorientierung für den Nutzer.

7.7.4 Fazit aus den korrekt spezifizierten Modellvarianten, in welchen Fehlerkorrelationen geschätzt werden

Die jeweiligen wahren Werte von gamma, der Fehlerkorrelationen und der Messmodellpfadkoeffizienten stimmen mit den in den Simulationen sich ergebenden arithmetischen Mittelwerten dieser Parameter fast genau überein, so dass letztere als erwartungstreue Schätzer bezeichnet werden können.

Modelle mit unabhängigen Faktoren

Die Schätzung einer *messmodellübergreifenden* Fehlerkorrelation führt zu keinen anderen SE-Schätzungen als die Ignorierung einer solchen Fehlerkorrelation. Die SE unterscheiden sich nicht vom Modell 1, dem Grundmodell, das keine Fehlerkorrelation beinhaltet.

Wird zusätzlich eine *messmodellspezifische* Fehlerkorrelation des ersten Messmodells geschätzt, führt dies ebenfalls zu keiner Veränderung der im Grundmodell beobachteten SE. Im

fehlspezifizierten Modell, in dem nämliche Fehlerkorrelation ignoriert wird, wird eine Reduktion der SE der betroffenen Pfadkoeffizienten beobachtet[58].

Modelle mit abhängigen Faktoren

Die Schätzung einer *messmodellübergreifenden* Fehlerkorrelation (Modell 8 k_g) führt bezüglich des Grundmodells 4, das keine Fehlerkorrelationen beinhaltet, zu etwas geringeren SE von gamma und der Pfadkoeffizienten des zweiten Messmodells. (Die Fehlspezifikation dieser Fehlerkorrelation hingegen (Modell 8) erhöht diese SE deutlich.). Die SE der Pfadkoeffizienten des ersten Messmodells sind unverändert hoch.

Die Schätzung einer *messmodellspezifischen* Fehlerkorrelation im ersten Messmodell (Modell 5 k_g) bewirkt bezüglich des Referenzmodells hingegen leicht erhöhte SE von gamma und der Pfadkoeffizienten des zweiten Messmodells. Zum anderen werden die SE des ersten Messmodells wesentlich erhöht. (Die Fehlspezifikation dieser Fehlerkorrelation hingegen (Modell 5) hat – mit Ausnahme des dritten Pfadkoeffizienten – eine Senkung aller SEs zur Folge.)

Die Schätzung der messmodellübergreifenden und der -spezifischen Fehlerkorrelation des ersten Messmodells (Modell 9 k_g) egalisiert den in Modell 8 k_g beobachteten Senkungseffekt im zweiten Messmodell. Zum anderen wird – bezogen auf das Grundmodell – eine Erhöhung der SE der Pfadkoeffizienten des ersten Messmodells wie in Modell 5 k_g beobachtet. (Im fehlspezifizierten Modell (Modell 9) hingegen sind die SE der von der Fehlerkorrelation betroffenen Pfadkoeffizienten bezüglich des Grundmodells verringert.)

Die Schätzung einer *messmodellspezifischen* Fehlerkorrelation im zweiten Messmodell (Modell Z4 k_g) führt zu höheren SE von gamma und den Pfadkoeffizienten des zweiten Messmodells als im Grund- und fehlspezifizierten Modell.

Die Kombination aus der Schätzung von messmodellspezifischen Fehlerkorrelationen im ersten und zweiten Messmodell führt zu einer *Addition der Effekte* auf die SE des zweiten Messmodells, die in den Modellen mit nur jeweils einer von beiden Korrelationen (Modell 5 k_g und Modell Z4 k_g) zu beobachten waren. Der SE von gamma ist leicht bezüglich der zuletzt genannten Modelle erhöht, in welchen der SE von gamma identisch ist. (Die SE der Messmodellpfadkoeffizienten des ersten Messmodells weisen die Höhe des Modells 5 k_g

[58] Es ist davon auszugehen, dass entsprechende Beobachtungen im zweiten Messmodell gemacht werden könnten, wenn eine messmodellspezifische Fehlerkorrelation im zweiten Messmodell geschätzt würde.

auf, so dass bei den SE des ersten Messmodells der Additionseffekt nicht zu beobachten ist.) Alle Beobachtungen bezüglich der SE betreffen sowohl die simulierten als auch die geschätzten SE.

8 Adjustierung der Standardfehlerschätzung nach der intraindividuellen Abhängigkeit zwischen paarigen Organen – Simulation und Anwendung in der Glaukomdiagnose

8.1.1 EINFÜHRUNG

Im bisherigen systematisch-theoretischen Teil der Arbeit wurden auf Basis simulierter Daten Auswirkungen diverser Fehlspezifikationen von Messfehlerkorrelationen untersucht. Hierbei handelte es sich um Korrelationen, die zwischen den Fehlertermen *verschiedener* Messverfahren bestehen.

Das vorliegende Kapitel beschäftigt sich mit den Auswirkungen der Fehlspezifikation einer anderen Korrelationsstruktur. Es handelt sich um Korrelationen zwischen den Ergebnissen *eines* jeweiligen Messverfahrens, das sowohl am linken als auch am rechten Auge eines Patienten durchgeführt wird. Diese Korrelation zwischen der links- und rechtsseitig erhobenen Messung eines Diagnoseverfahrens kann in Rahmen von Strukturgleichungsmodellen in zwei Komponenten zerlegt werden: Ein Korrelationsbestandteil geht auf den nicht direkt messbaren, augenseitenspezifischen Schweregrad der Glaukomerkrankung zurück. Der zweite Korrelationsbestandteil ist auf den (ebenso nicht direkt zu erhebenden) Messfehler zurückzuführen, der bei der Messung in beiden Teilorganen begangen wird. Die Korrelation zwischen manifesten, linksseitig und rechtsseitig erhobenen Variablen, die beide demselben Messverfahren zuzuordnen sind, setzt sich also aus einer latenten Schweregrad- und einer latenten Messfehlerkorrelation zusammen.

Die in diesem Kapitel dokumentierte Untersuchung hat jedoch primär nicht zum Ziel, Schweregrad- und Messfehlerkorrelationen zwischen den beiden Teilorganen anhand von linearen Strukturgleichungsmodellen zu quantifizieren.[59] Es geht vielmehr um die Berücksichtigung dieser Korrelationen bei der Schätzung der Standardfehler interessierender Koeffizienten.

Die Standardfehler der Koeffizienten werden nämlich unter Vernachlässigung der intraindividuellen Korrelation zwischen Teilorganen fehlerhaft geschätzt, wenn beide Augen eines Patienten in die Analyse eingebracht werden. Dies hat Konsequenzen sowohl für die Angabe der Konfidenzintervalle der Koeffizienten als auch hinsichtlich der Entscheidung über Ablehnung bzw. Nichtablehnung eines getesteten Modells (siehe 8.2.2.1, 8.2.5.1, 9, 10.6.1).

[59] Dieser Versuch wurde unternommen in Martus (2001b), S. 932. Auf dieses Modell wird im zweiten Abschnitt dieses Kapitels eingegangen (8.3.1).

In keiner der hier verwendeten Softwarelösungen, weder in AMOS (SPSS) noch PROC CALIS (SAS) ist es ohne weiteres möglich, nach der Seitenkorrelation adjustierte Standardfehler zu schätzen: Die Software sieht die Möglichkeit nicht vor, einen Patienten als ein individuelles Cluster zu definieren, das die an beiden Teilorganen vorgenommenen Messungen umfasst. Alternativ bietet sich bei Verwendung der Software SAS eine Resampling-Methode zur Schätzung der Standardfehler an.[60] Mittels eines nonparametrischen Bootstrapverfahrens können Standardfehler interessierender Koeffizienten geschätzt werden, die nach der Seitenabhängigkeit adjustiert sind. Damit ist es möglich, adjustierte Konfidenzintervalle für die interessierenden Pfadkoeffizienten und die Chi-Quadrat-Statistik des globalen Goodness of Fit-tests anzugeben.[61]

8.1.2 BISHERIGER FORSCHUNGSTAND

Es sind von Peter Martus bereits mehrere Messmodelle zur Glaukomdiagnose publiziert worden, die mit der Software AMOS erstellt wurden. Ausgangspunkt der folgenden Überlegungen ist ein Modell, in dem die Schätzung der Standardfehler der Pfadkoeffizienten über Bootstrapverfahren mit n= 3000 Wiederholungen vorgenommen wurde[62]. Eine Bootstrapanalyse mit AMOS ist insofern nicht benutzerdefiniert möglich, als nicht der Patient, sondern das einzelne Auge als Untersuchungseinheit betrachtet wird, wenn man die an jeweils beiden Augen eines bestimmten Patienten vorgenommenen Messungen in die Analyse mit AMOS einbezieht. Man verfährt daher notgedrungen so, als handele es sich bei den individuellen Werten am linken und rechten Auge um Werte verschiedener Individuen. Zum einen wird damit die Abhängigkeit zwischen dem rechten und linken Auge eines jeweiligen Patienten ignoriert. Zum anderen wird fälschlicherweise davon ausgegangen, dass die Anzahl der unabhängigen Untersuchungseinheiten bei vollständigen Daten doppelt so groß sei wie die tatsächliche Anzahl. Werte zweier, einem Individuum zugehöriger Augen sind jedoch voneinander abhängig. Das Ignorieren dieser (intraindividuellen) Abhängigkeit der Augenseiten führt zu einer *Unterschätzung der Standardfehler*. Die Unterschätzung der SE bewirkt bei konfirmatorischen

---

[60] Diese Möglichkeit ist bei AMOS schon deshalb nicht gegeben, weil keine Output-Datendatei erzeugt werden kann, um bspw. das aus einem bestimmten Bootstrapsample resultierende Ergebnis zu speichern. Weiterhin sind in AMOS keine Simulationen möglich.

[61] Die Verteilung der P-Werte kann nicht angegeben werden. Denn zur Ermittlung der P-Wertverteilung benötigte man die Verteilung der Chi-Quadrat-Statistik unter der Nullhypothese, die nur durch parametrische Bootstrapverfahren gewonnen werden kann. Wie sich nach Durchführung der Bootstrapanalyse herausstellen sollte, ist eine Ermittlung adjustierter Konfidenzintervalle für die Chi-Quadrat-Statistik nicht ohne weiteres möglich (siehe 10.6.1).

[62] Martus (2001a), S. 269.

Vergleichen von Koeffizienten tendenziell, dass Koeffizienten irrtümlicherweise als signifikant verschieden von 0 erachtet werden. Weiterhin zieht sie eine Überschätzung der χ^2 –verteilten Diskrepanzfunktion nach sich, mit der häufig die globale Anpassungsgüte des Modells (Goodness of Fit) geprüft wird. Die fehlende Adjustierung nach der Augenseitenkorrelation macht somit die Ablehnung eines Messmodells wahrscheinlicher.

In Anlehnung an das „concept of effective number of eyes"[63] wurde daher ein Korrekturfaktor für die angemessene Schätzung der Standardfehler angenommen. Dieser ist vom Ausmaß der Korrelation zwischen den Augenseiten abhängig und nimmt Werte im Bereich zwischen 1 (keine Korrelation zwischen den Augenseiten) und $\sqrt{2}$ (perfekter Zusammenhang zwischen den Augenseiten) an. Der Korrekturfaktor für die χ^2 –verteilte Statistik liegt entsprechend im Bereich zwischen 1 (keine Korrelation) und 0,5 (perfekter Zusammenhang).

8.1.3 VERBESSERUNG DES FORSCHUNGSSTANDES

Die Software PROC CALIS ist besser zur Durchführung der erforderlichen Bootstrapanalysen geeignet als die Software AMOS. Hierbei werden die erweiterten Möglichkeiten von PROC CALIS hinsichtlich der Ausgabe von Analyseresultaten (Erzeugung von SAS-Datendateien als Outputdateien) und die bessere Einbindung der SAS-Prozedur in die SAS-Prozedurumgebung (Durchführbarkeit von Simulationen) genutzt.

Es können somit nicht nur – wie bisher – Bootstrapsamples durch Ziehung (mit Zurücklegen) von individuellen Augen, sondern durch Ziehung von individuellen Augenpaaren erzeugt werden. Die Ziehung von Augenpaaren trägt der Anforderung Rechnung, dass die Untersuchungseinheit der Patient und nicht das einzelne Auge sei.

Beiderlei Arten von Bootstrapsamples werden in dieser Untersuchung als Inputdateien für PROC CALIS verwendet. Die Resultate der verschiedenen Ziehungsvarianten in PROC CALIS können anschließend miteinander verglichen werden.

Im *ersten Abschnitt* dieses Kapitels (8.2) wird das herkömmliche Bootstrapverfahren (Ziehung einzelner Augen) dem verbesserten (Ziehung von Augenpaaren) zum Vergleich auf zweierlei Weise gegenübergestellt:

Der Vergleich wird einmal auf Basis von in SAS simulierten, approximativ normalverteilten Daten angestellt, die bestimmte Korrelationsstrukturen zwischen Messungen an linken und

[63] Rosner B. Statistical methods in ophthalmology: An adjustment for the intraclass correlation between eyes. Biometrics 1982 38: 105 -14, zitiert nach Martus (2001a), S. 267.

rechten Augen aufweisen. Die Datensimulation und Datenanalyse geschieht anhand zweier einfacher Modelle der CFA[64] mit drei *verschiedenen* Messverfahren und einem Faktor (Teil A).

Zum anderen wird einer AMOS-Bootstrapanalyse empirischer Daten (von Glaukompatienten), die auf Basis eines publizierten, komplexen Modells der CFA durchgeführt wurde (Paarigkeit der Daten ignoriert), eine Bootstrapanalyse mit PROC CALIS gegenübergestellt (Paarigkeit der Daten berücksichtigt). Das komplexe Modell beinhaltet acht verschiedene Messverfahren und zwei Faktoren (Teil B).

Im *zweiten Abschnitt* dieses Kapitels (8.3) wird versucht, die Ergebnisse aus Teil A durch Analysen an empirischen Daten nachzuvollziehen und zu überprüfen. Den einfachen Modellen aus Teil A entsprechend werden hierzu zwei in ihrer Komplexität reduzierte Modelle der CFA verwendet, von welchen eines zur Datenanalyse von Glaukompatienten entwickelt und ebenfalls publiziert wurde. Sie enthalten vier *verschiedene* Messverfahren und einen Faktor.

8.2 Erster Abschnitt: Nicht adjustierte und nach Abhängigkeit der Augenseiten adjustierte Bootstrapverfahren zur Schätzung von Standardfehlern in Modellen der CFA

8.2.1 [A] Bootstrapverfahren auf Basis simulierter Daten anhand eines Ein-Faktor-Modells mit drei Indikatoren unter Verwendung von PROC CALIS

8.2.1.1 Erzeugung der Simulationsdaten

Es werden in einem Ein-Faktor-Modell mit drei Indikatoren, dessen Struktur aus einem vorhergehenden Kapitel bekannt ist, Daten simuliert, die eine bestimmte Korrelationsstruktur aufweisen. Diese Korrelationsstruktur entspricht folgendem, in der Glaukomdiagnose als realistisch angenommenem Sachverhalt:

Drei verschiedene Messverfahren werden an einer Stichprobe von Patienten an jeweils beiden Augen durchgeführt. Man nimmt für jedes Messverfahren an, dass die Messfehler der Messung am linken und rechten Auge in einem bestimmten Ausmaß korreliert sind. Weiterhin geht man davon aus, dass die latente Variable „Schweregrad am linken Auge“, welche die drei linksseitig durchgeführten Messverfahren erklärt, mit der entsprechenden latenten Variablen „Schweregrad am rechten Auge“ in bestimmter Höhe korreliert ist.

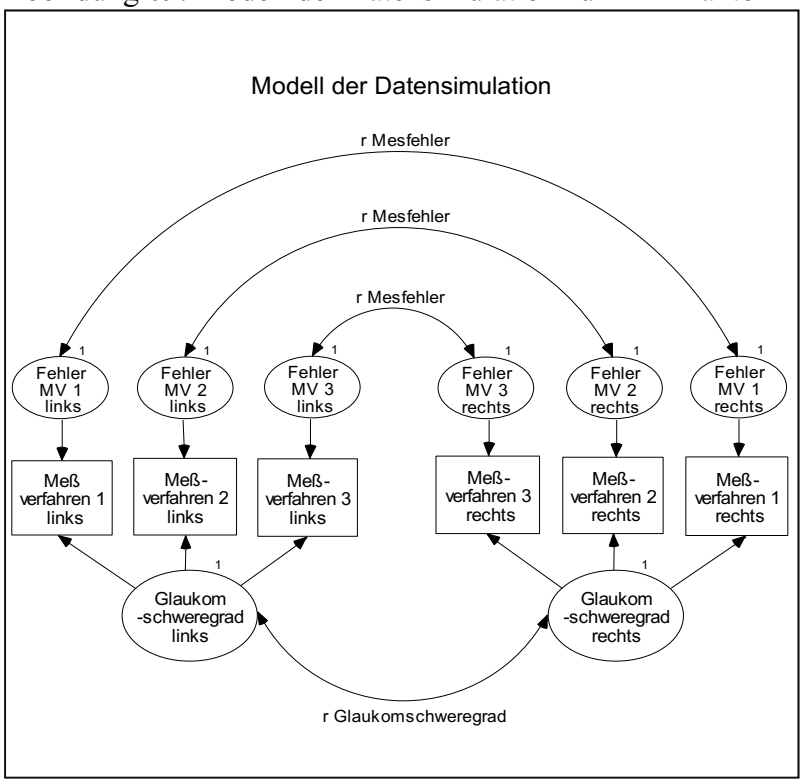
[64] Confirmative factor analysis, ein Spezialfall eines linearen Strukturgleichungsmodells (SEM; siehe 2.5)

Es ist anzunehmen, dass *real* die drei Fehlerkorrelationen zwischen den Augenseiten, r (Messfehler Verfahren 1 links, Messfehler Verfahren 1 rechts), r (Messfehler Verfahren 2 links, Messfehler Verfahren 2 rechts) und r (Messfehler Verfahren 3 links, Messfehler Verfahren 3 rechts), *unterschiedlich hoch* sind. Auch dürften tatsächlich die Koeffizienten der Pfade vom linken Schweregrad zu den linksseitig durchgeführten Messverfahren und jene vom rechten Schweregrad zu den linksseitig durchgeführten Messverfahren sämtlich verschieden hoch sein.

Bei der Erzeugung der simulierten Daten wird jedoch hierbei vereinfachend angenommen, dass die oben genannten drei augenseitenbezogenen Fehlerkorrelationen jeweils *gleich hoch* sind.

Einfacherweise nimmt man weiterhin an, dass die drei Pfadkoeffizienten zwischen dem latenten Schweregrad am linken Auge und den linksseitig durchgeführten Messverfahren zum einen untereinander identisch sind und mit jenen der anderen Augenseite übereinstimmen. Ebensolches gilt somit auch für die Pfadkoeffizienten zwischen den Fehlervariablen und den Messverfahren, so dass der wahre Wert aller zwölf Pfadkoeffizienten des folgenden Modells $1/\sqrt{2}$ =0,707 beträgt (siehe Abbildung 69).

Abbildung 69: Modell der Datensimulation zum Ein-Faktor-Modell in allgemeiner Form



Nach diesem Modell werden 16 verschiedene Korrelationsstrukturen simuliert, die sich als Kombination aus vier verschiedenen links-rechts-Korrelationen des Glaukomschweregrades einerseits und der Messfehler andererseits ergeben. Folgender Tabelle 10 zur Abbildung 69 können die 16 Korrelationsstrukturen entnommen werden.

Tabelle 10: 16 durch das Modell der Datensimulation erzeugte Korrelationsstrukturen

	R Messfehler 0	R Messfehler 0,5	R Messfehler 0,7	R Messfehler 0,9
R Glaukomschweregrad 0	K 1	K 2	K 3	K 4
R Glaukomschweregrad 0,5	K 5	K 6	K 7	K 8
R Glaukomschweregrad 0,7	K 9	K 10	K 11	K 12
R Glaukomschweregrad 0,9	K 13	K 14	K 15	K 16

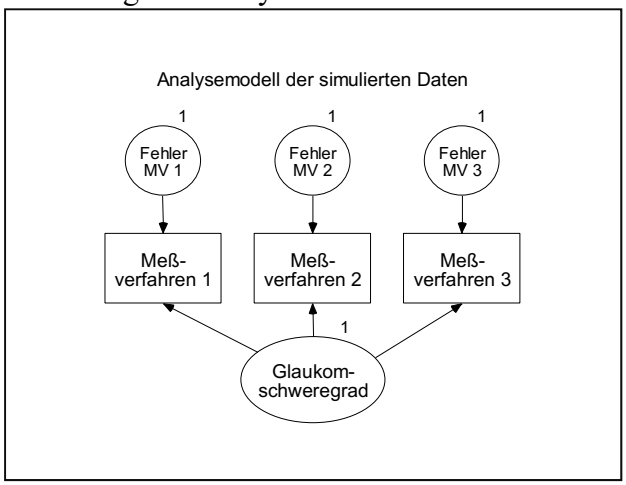
Somit kann bei Auswertung der korrelierten Daten anhand des unten beschriebenen Ein-Faktor-Modells mit drei Indikatoren der Effekt beurteilt werden, den *eine* der beiden Korrelationskomponenten (Schweregrad oder Fehler) eines bestimmten Ausmaßes auf die Standardfehlerschätzung ausübt. Gleichzeitig kann der Effekt ermittelt werden, der von der *jeweils anderen* Korrelationskomponente im Zusammenspiel mit ersterer ausgeht. Jede dieser 16 Korrelationsstrukturen wird – in Abhängigkeit vom Startwert bei der Zufallszahlenerzeugung – in zwei Varianten realisiert (simulierte „Population“ 1 und simulierte „Population“ 2). Dies geschieht, um durch Vergleich der Ergebnisse aus Variante 1 mit den Resultaten aus Variante 2 abzuschätzen, wie stark die Ergebnisse von der Zufallszahlenerzeugung abhängen. Jede der beiden simulierten Populationen besteht aus approximativ standardnormalverteilten Zufallsvariablen, aus welchen wiederum manifeste Variable (Messverfahren) konstruiert werden (siehe 11.6.1.1). Die Zufallsvariablen sind nicht perfekt unabhängig voneinander, sondern aufgrund von Zufallsschwankungen im Allgemeinen untereinander schwach korreliert. Die simulierten Populationen weichen also voneinander zufallsbedingt ab hinsichtlich des Mittelwertes, der Varianz und der Korrelationsstruktur der elementaren Zufallsvariablen. Alle latenten Variablen (Fehler und Glaukomschweregrad) sind approximativ standardnormalverteilt.

8.2.1.2 Analyse der Simulationsdaten

Die simulierten Daten werden anhand eines Messmodells (Ein-Faktor-Modell mit drei Indikatoren) analysiert (siehe Abbildung 70). Unter Annahme dieses Messmodells mit 3 Messverfahren (MV1, 2 und 3) werden auf Basis der aus den beiden Populationen (Samples) gezogenen Stichproben (Bootstrapsamples) die Mittelwerte der Pfadkoeffizienten des Glaukom-

schweregrades (PK1, 2 und 3) sowie deren Standardfehler [65] (SE) geschätzt. Der *wahre Wert* aller drei standardisierten Pfadkoeffizienten beträgt 0,707.[66]

Abbildung 70: Analysemodell der zum Ein-Faktor-Modell simulierten Daten



Nach Augenseitenkorrelation nicht adjustierter Bootstrap

Es werden 3000 Bootstrapsamples im Umfang von n=800 unter Ziehung einzelner Augen mit Zurücklegen gezogen. Die Stichprobenziehung erfolgt einmal auf Basis des Samples 1 (simulierte Population 1) und einmal auf Basis des Samples 2 (simulierte Population 2). Die simulierten Populationen 1 und 2 unterscheiden sich nur hinsichtlich des Startwertes bei der Zufallszahlenerzeugung.

8.2.1.3 Ergebnisse

Die nachfolgende Tabelle 11 gibt die Resultate der Bootstrapanalyse wieder, in der die intraindividuelle Augenseitenkorrelation nicht berücksichtigt wird. Die geschätzten Standardfehler der Pfadkoeffizienten sind daher zu gering.

65 Der Standardfehler eines Pfadkoeffizienten ergibt sich als die Standardabweichung der aus den Bootstrapsamples resultierenden geschätzten Pfadkoeffizienten.

66 Der wahre Wert wurde auf Basis des in zwei vorstehenden Kapiteln behandelten, identischen Messmodells unter Verwendung von AMOS auf Basis perfekt unabhängiger elementarer Zufallsvariabler ermittelt (siehe Kap. 3 sowie Kap. 5).

Tabelle 11: Pfadkoeffizienten und SE der nichtadjustierten Bootstrapanalyse über 16 simulierte Korrelationsstrukturen des Ein-Faktor-Modells mit drei Indikatoren

R: Korrelation zwischen linker und rechter Augenseite P: Population
MV: Messverfahren PK: Pfadkoeffizient SE: Standardfehler des PK

			R Messfehler 0			R Messfehler 0,5			R Messfehler 0,7			R Messfehler 0,9		
R Glaukomschweregrad 0			**MV1**	**MV2**	**MV3**	**MV1**	**MV2**	**MV3**	**MV1**	**MV2**	**MV3**	**MV1**	**MV2**	**MV3**
	P1	**PK**	,73	,71	,71	,74	,68	,71	,74	,66	,71	,74	. 66	,71
		SE	,027	,027	,727	,027	,027	,027	,027	,027	,028	,028	,027	,028
	P2	**PK**	,74	,68	,78	,75	,70	,76	,75	,71	,76	,74	,71	,77
		SE	,026	,027	,024	,024	,025	,025	,024	,025	,024	,024	,025	,024
R Glaukomschweregrad 0,5	**P1**	**PK**	,72	,68	,66	,72	,65	,70	,72	,64	,70	,71	,63	,71
		SE	,031	,030	,030	,029	,030	,030	,030	,030	,030	,030	,030	,030
	P2	**PK**	,73	,68	,78	,74	,70	,76	,74	,71	,76	,73	,71	,76
		SE	,027	,027	,026	,026	,026	,026	,025	,025	,025	,025	,025	,024
R Glaukomschweregrad 0,7	**P1**	**PK**	,71	,67	,65	,72	,64	,69	,71	,63	,70	,70	,63	,71
		SE	,032	,032	,031	,030	,030	,031	,030	,030	,031	,031	,030	,031
	P2	**PK**	,72	,68	,78	,73	,70	,76	,73	,71	,76	,72	,71	,76
		SE	,027	,028	,027	,026	,026	,027	,031	,030	,030	,026	,025	,025
R Glaukomschweregrad 0,9	**P1**	**PK**	,71	,66	,63	,71	,63	,68	,70	,63	,70	,69	,63	,71
		SE	,033	,033	,033	,030	,031	,032	,030	,030	,032	,031	,030	,033
	P2	**PK**	,71	,67	,77	,73	,70	,75	,72	,71	,75	,71	,71	,76
		SE	,028	,028	,028	,027	,026	,027	,027	,026	,027	,027	,025	,026

Die Schätzungen der Pfadkoeffizienten weisen geringfügige Abweichungen vom wahren Wert auf und es ist keine Abhängigkeit von der Korrelationsstruktur der Daten oder der Zufallszahlenerzeugung zu erkennen

Die Schätzung der Standardfehler für die drei Messverfahren auf Basis von simulierter Population 2 führt in allen 16 Korrelationsstrukturen zu sehr ähnlichen Resultaten. Die aus simulierter Population 1 geschätzten Standardfehler sind bei einer Schweregradkorrelation ungleich Null bezüglich jener bei Schweregradkorrelation gleich Null minimal erhöht (um ca. 10 Prozent). Mit Ausnahme letzterer Konstellation (Schweregradkorrelation gleich Null) sind jedoch in allen Kombinationen von Schweregrad- und Messfehlerkorrelation die SE der Population 1 geringfügig größer als die der Population 2 (um bis zu 25 %). Da die größten in dieser Tabelle zu beobachtenden Unterschiede zwischen SE-Schätzungen auf Varianten der Zufalls-

zahlenerzeugung zurückgehen, kann keine Abhängigkeit zwischen intraindividueller Seitenkorrelation und der Höhe der Standardfehlerschätzung behauptet werden.

Nach Augenseitenkorrelation adjustierter Bootstrap

Es werden 3000 Bootstrapsamples im Umfang von n=400 unter Ziehung individueller Augenpaare mit Zurücklegen gezogen. Die Stichproben stammen analog zur Stichprobenziehung einzelner Augen aus zwei Samples (simulierten Populationen).

Ergebnisse und Vergleich beider Bootstrapverfahren

Die untenstehende Tabelle 12 gibt die Ergebnisse der Bootstrapanalyse wieder, in der die Augenseitenkorrelation berücksichtigt wird. Die Tabelle enthält neben den Ergebnissen der nach Augenseitenabhängigkeit adjustierten Bootstrapanalyse auch das Verhältnis der Standardfehler dieser und der in Tabelle 11 dokumentierten nicht adjustierten Analyse (Q). Beispiel: In der nichtadjustierten Analyse (Tabelle 11) ergibt sich in der ersten Korrelationsstruktur (R Messfehler = 0, R Glaukomschweregrad = 0) in der Population 1 (P1) für den SE des ersten Messverfahrens der Wert 0,27. In der adjustierten Analyse (Tabelle 12) ergibt sich der entsprechende Wert 0,28. Der Quotient (Q) aus 0,28 und 0,27 beträgt 1,03 und wird auch als Korrekturfaktor bezeichnet.

Tabelle 12: Pfadkoeffizienten und SE der adjustierten Bootstrapanalyse über 16 simulierte Korrelationsstrukturen sowie Korrekturfaktoren

Q: Quotient aus SE des PK bei Ziehung von Augenpaaren und SE des PK bei Ziehung einzelner Augen

			R Messfehler 0			R Messfehler 0,5			R Messfehler 0,7			R Messfehler 0,9		
R Glaukom-schweregrad 0			**MV1**	**MV2**	**MV3**	**MV1**	**MV2**	**MV3**	**MV1**	**MV2**	**MV3**	**MV1**	**MV2**	**MV3**
	P1	**PK**	,73	,71	,71	,74	,67	,71	,74	,66	,71	,74	,66	,71
		SE	,028	,028	,026	,029	,028	,028	,031	,030	,030	,035	,032	,034
		Q	1,03	1,04	,97	1,08	1,04	1,02	1,16	1,10	1,10	1,28	1,18	1,22
	P2	**PK**	,74	,68	,78	,75	,71	,76	,75	,71	,76	,74	,71	,77
		SE	,027	,026	,025	,026	,026	,027	,028	,027	,028	,031	,030	,030
		Q	1,01	,98	1,02	1,07	1,03	1,07	1,15	1,10	1,14	1,27	1,20	1,23
R Glaukom-schweregrad 0,5	**P1**	**PK**	,72	,68	,66	,72	,65	,70	,72	,64	,70	,71	,63	,71
		SE	,033	,031	,030	,032	,031	,032	,035	,033	,035	,040	,036	,039
		Q	1,06	1,03	,97	1,12	1,05	1,06	1,22	1,12	1,16	1,32	1,22	1,28
	P2	**PK**	,73	,68	,78	,74	,70	,76	,74	,71	,76	,73	,71	,76
		SE	,028	,028	,027	,028	,028	,028	,030	,029	,029	,033	,033	,031
		Q	1,06	1,01	1,01	1,10	1,07	1,07	1,18	1,14	1,16	1,30	1,25	1,27
R Glaukom-schweregrad 0,7	**P1**	**PK**	,71	,67	,65	,72	,64	,69	,71	,63	,70	,70	,63	,71
		SE	,034	,033	,031	,033	,032	,034	,037	,034	,037	,041	,037	,041
		Q	1,07	1,03	,98	1,14	1,06	1,08	1,24	1,15	1,19	1,34	1,26	1,31
	P2	**PK**	,72	,68	,78	,73	,70	,76	,73	,71	,76	,72	,71	,76
		SE	,030	,029	,027	,030	,029	,029	,031	,030	,030	,034	,032	,032
		Q	1,07	1,03	1,01	1,12	1,11	1,09	1,20	1,18	1,17	1,32	1,29	1,29
R Glaukom-schweregrad 0,9	**P1**	**PK**	,71	,66	,63	,71	,63	,68	,70	,63	,70	,69	,62	,71
		SE	,036	,035	,033	,035	,034	,036	,038	,036	,040	,043	,039	,044
		Q	1,09	1,04	1,00	1,16	1,09	1,12	1,26	1,18	1,22	1,37	1,30	1,34
	P2	**PK**	,71	,67	,77	,73	,70	,75	,72	,70	,75	,71	,71	,75
		SE	,030	,030	,029	,031	,030	,030	,033	,031	,032	,036	,034	,034
		Q	1,09	1,05	1,03	1,14	1,15	1,11	1,23	1,22	1,19	1,35	1,33	1,31

1. Die Höhe der Pfadkoeffizientenschätzung entspricht der des nichtadjustierten Bootstrapverfahrens und hängt somit nicht von der der Ziehungsvariante beim Bootstrap ab.

2. Wie die nichtadjustierte Bootstrapanalyse, so bringt auch die adjustierte in allen Kombinationen von Schweregrad- und Messfehlerkorrelation Standardfehlerschätzungen der simulierten Population 1 hervor, die geringfügig größer als jene der simulierten Population 2 sind. Da die Zielgröße dieser Untersuchung der *populationsspezifische Quotient des SE aus adjustierter und nichtadjustierter Analyse* ist, ist dieser Unterschied zwischen den simulierten Populationen ohne Bedeutung: Es gibt keinen systematischen Größenunterschied zwischen den Quotienten der simulierten Population 1 und der 2 (zur Definition der beiden Populationen siehe 8.2.1.1)

3. Die nach Maßgabe der ersten von 16 Korrelationsstrukturen (siehe Tabelle 10 zu Abbildung 69) simulierten Populationen sind sowohl durch eine Glaukomschweregrad- als auch eine Messfehlerkorrelation von Null gekennzeichnet. Die adjustierte Bootstrapanalyse dieser

simulierten Populationen führt zu sehr ähnlichen Standardfehlern wie die nichtadjustierte. Damit wird gleichsam empirisch die dem vergleichenden Verfahren zugrundeliegende Annahme bestätigt, dass die adjustierte und nichtadjustierte Bootstrapanalyse identische Resultate hervorbringen, wenn keine Seitenabhängigkeit der Messverfahren vorliegt. M. a. W.: Unterschiedliche Resultate in den beiden Bootstrapvarianten sind ausschließlich auf Seitenabhängigkeit zurückzuführen.)

4. In allen anderen 15 Korrelationsstrukturen werden durch die adjustierte Bootstrapanalyse höhere Standardfehler geschätzt. (Die Ausnahme bildet die Standardfehlerschätzung des Pfadkoeffizienten von Messverfahren 3 in den Korrelationsstrukturen 1, 5 und 9 auf Basis der ersten simulierten Population, die dem Zufall geschuldet ist).

5. Die bisherige Ad-hoc-Lösung für die angemessene Schätzung der SE durch den Korrekturfaktor (im Bereich von 1 - $\sqrt{2}$) wird insofern bestätigt, als die Simulation Quotienten der Schätzungen des SE hervorgebracht hat, die sich fast über den ganzen angegebenen Wertebereich erstrecken (0,97–1,37). Wenn man künstlich Daten erzeugte, bei welchen die Schweregrad- als auch die Messfehlerkorrelation zwischen den beiden Augenseiten gleich 1 ist, erhöhte sich der Quotient aus dem adjustierten und dem nichtadjustierten SE auf genau 1,41.[67]

6. Die Messfehlerkorrelation erhöht die Standardfehler stärker (bisweilen mehr als doppelt so stark) als die entsprechende Schweregradkorrelation. Beispielsweise liegen die Quotienten aus SE der adjustierten und nicht adjustierten Analyse in der Korrelationsstruktur 14 (R Messfehler = 0,5, R Schweregrad = 0,9) im Bereich 1,09–1,16, während selbige in der Struktur 8 (R Messfehler = 0,9, R Schweregrad = 0,5) im Bereich 1,22–1,32 liegen.

---

[67] Diese Daten erhält man, indem man die für eine bestimmte Augenseite simulierten Daten verdoppelt, so dass die für die linken und rechten Augen simulierten Messwerte vollkommen identisch sind. Im Anhang lauteten dann die Gleichungen unter 2A für die linke Augenseite:

„indik_1 = g_li_5 + f1_li_5;
indik_2 = g_li_5 + f2_li_5;
indik_3 = g_li_5 + f3_li_5;“
und unter 2B für die rechte Augenseite statt
“indik_1 = g_re_5 + f1_re_5;
indik_2 = g_re_5 + f2_re_5;
indik_3 = g_re_5 + f3_re_5;”
ebenfalls
„indik_1 = g_li_5 + f1_li_5;
indik_2 = g_li_5 + f2_li_5;
indik_3 = g_li_5 + f3_li_5;“

8.2.2 [B] BOOTSTRAPVERFAHREN AUF BASIS VON PATIENTENDATEN ANHAND EINES ZWEI-FAKTOR-MODELLS MIT ACHT DIAGNOSEVERFAHREN

8.2.2.1 Bisheriger Forschungsstand

8.2.2.1.1 Diagnoseverfahren für die Glaukomerkrankung

Als etablierte Kriterien für die Diagnose der Glaukome gelten der intraokulare Druck, der Papillen- und der Gesichtsfeldschaden. Mittlerweile wird der morphometrisch erkennbare Papillenschaden als zentrales Diagnosekriterium angesehen. Aus diesen und weiteren Kriterien ergibt sich eine Vielzahl von quantitativen Diagnoseverfahren.

Neben den morphometrischen Verfahren gibt es psychophysische sowie elektrophysiologische Diagnoseerfahren. Die letzten beiden lassen sich gliedern nach der Methode der visuellen Reizung, die z. B. ungemustert oder gemustert, im schwarz-weiß- oder Farbkontrast erfolgen kann und nach der Methode der Signalgewinnung, die subjektiv-psychophysisch oder objektiv durchgeführt werden kann. Im Folgenden werden die Messverfahren erläutert, die das in der Einleitung erwähnte komplexe Modell der CFA beinhaltet.

Morphometrie

Die quantitative Komponente der Untersuchung des Sehnervkopfes wird als Fläche des neuroretinalen Randsaums (Area of the neuroretinal rim, **NRRA**) angegeben und korrespondiert mit der Anzahl der intakten Nervenfasern.

Ungemusterte Reizung und subjektive Signalgewinnung (psychophysische Verfahren)

a) Statische computergesteuerte automatische Projektionsperimetrie Weiß-auf-Weiß zur Messung des Gesichtsfeldschadens

Der Patient blickt in ein Gerät (Perimeter), das ein Gesichtsfeld von 180 Grad imitiert und muss seinen Blick auf einen festen Punkt richten, ohne die Augen zu bewegen. Nun werden in einem Gesichtsfeldbereich von 30 Grad Lichtpunkte projiziert, die der Patient erkennen soll. Diese Untersuchung kann mit verschiedenen Lichtstärken, die der Empfindlichkeit der verschiedenen Bereiche der Netzhaut entsprechen, durchgeführt werden. Bleibt die Reaktion des Patienten bei wiederholten Prüfungen mit verschiedenen Lichtstärken aus, so kann das bedeuten, dass ein Skotom (Ausfall eines Teiles des Gesichtsfeldes) vorliegt.

Sinnesphysiologischer Goldstandard in der Glaukomdiagnose ist die *konventionelle Weiß-auf-Weiß*-Perimetrie. In der vorliegenden Untersuchung wird automatische Perimetrie verwandt,

die an 59 Punkten des Sehfeldes Lichtempfindlichkeiten misst. Der mittlere Schaden (**M**ean **D**efect, **MD**) ist das arithmetische Mittel dieser 59 Differenzen zwischen individueller Lichtempfindlichkeit und altersadjustierten Normwerten. Die korrigierte Varianz des Verlustes (**C**orrected **L**oss **V**ariance, **CLV**) ist die Varianzkomponente dieser 59 Differenzen bezogen auf die räumliche Variation.

b) „Erlanger Flimmertest" (FLI)

Ungemusterte Reize werden in Form eines sinusförmig modulierten Ganzfeldreizes geboten und die zeitliche Kontrastempfindlichkeit im Ganzfeld subjektiv ermittelt. Im Gegensatz zur Perimetrie handelt es sich beim Flimmertest um ein schnelles Screeningverfahren. Gemeinsam ist die Reizung des Gesamtgesichtsfeldes und nicht ausgewählter Sektoren.

c) Ein Test der dynamischen, d. h. zeitlich-räumlichen Kontrastempfindlichkeit (DYN)

Bei diesem Test wird ein schwarz-weißes Schachbrettmuster in zeitlichem und räumlichem Kontrast in einen Bereich des Gesichtsfeldes projiziert, der für frühe Glaukomschäden bekannt ist.

Gemusterte Reizung und objektive Signalgewinnung (elektrophysiologische Verfahren)

Es wird mit verschiedenen Helligkeits- und Farbkontrasten gearbeitet und der Reiz nach dem schnellen (subjektiv nicht erkennbaren) Musterwechsel oder dem langsameren Muster Onset-offset-Verfahren angeboten. Zu den objektiven Methoden zählen vor allem die elektrophysiologischen Verfahren der Ableitung bioelektrischer Potentiale vom Auge in Form des Elektroretinogramms (ERG) und des visuell evozierten Potentials (VEP). Im Rahmen der hier vorliegenden Untersuchung werden die folgenden speziellen Verfahren verwendet: Die Amplitude eines schwarz-weiß-Musterwechsel-**E**lektro**r**etino**g**ramms (**ERG**) und die Amplitude eines Blau-auf-Gelb visuell evozierten Potentials (**AMP**) im Onset-offset-Modus sowie die Latenzzeit des Peaks (Stimulus-onset bis Peak) dieser Amplitude **(VEP).**

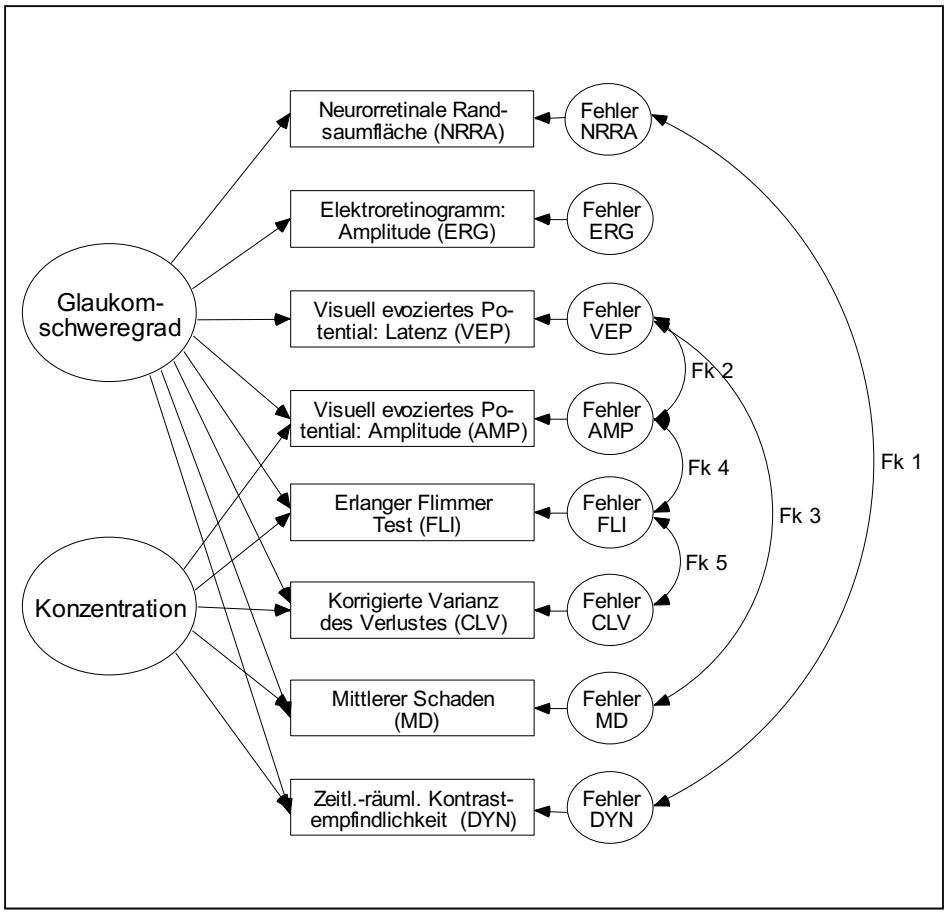
8.2.2.1.2 Modellbildung

Ausgangspunkt der folgenden Untersuchungen ist ein publiziertes Modell der CFA, das in folgender Abbildung 71 wiedergegeben wird.[68] Die Daten von 237 Patienten des Erlanger Glaukomregisters mit Verdacht auf Glaukom oder diagnostiziertem Glaukomschaden wurden

[68] Martus (2001a), S. 269.

anhand des Modells ausgewertet. Für alle im Modell enthaltene Diagnoseverfahren lagen Messungen an beiden Augen vollständig vor.

Abbildung 71: Messmodell zur Diagnose des allgemeinen Glaukomschweregrades mit acht Indikatoren



Veranschaulicht werden acht Diagnoseverfahren, die (von oben nach unten) den o. g. Gruppen wie folgt zugeordnet sind: Verfahren 1: morphologische, Verfahren 2–4: elektrophysiologische und Verfahren 5–8: psychophysische. Die Ergebnisse der Diagnoseinstrumente werden einerseits von zwei Faktoren beeinflusst, dem Glaukomschweregrad und der Konzentrationsfähigkeit der Patienten bei der Absolvierung diagnostischer Tests. Andererseits beeinflussen Messfehlerkomponenten die Instrumente. In diesem Modell werden fünf Fehlerkorrelationen angenommen, so dass, bis auf beim ERG, von abhängigen Messverfahren ausgegangen wird. Dieses Modell stellt eine Weiterentwicklung eines Modells dar, das ursprünglich konfirmatorisch zu bewerten war. Das Ursprungsmodell war deutlich restriktiver als das weiterentwickelte, da es für die Fehlerkorrelationen 1, 3, 4 und 5 den Wert Null annahm und unterstellte, dass die Konzentration das visuell evozierte Potential (AMP) nicht beeinflusse.

In einer Vorarbeit[69] wurden auf Basis des abgebildeten Modells Pfadkoeffizienten geschätzt und Bootstrapanalysen zur Schätzung der Standardfehler der Pfadkoeffizienten mit der Software AMOS durchgeführt. Der globale Goodness of Fit-Test ergab ein χ^2 von 7,957 (df=10). Da in der Analyse mit AMOS nicht nach der Abhängigkeit adjustiert werden kann, die zwischen den beiden Augen eines Patienten besteht (s. o.), wurde angenommen, dass zu geringe Standardfehler und entsprechend zu große χ^2-Werte geschätzt würden. Ein p-Wert wurde daher nicht angegeben. Um realistische Schätzungen der Standardfehler zu erhalten, wurde vorgeschlagen, die resultierenden SE mit einem Korrekturfaktor zu multiplizieren, der in Abhängigkeit von der Korrelation zwischen den an den beiden Augenseiten angewandten Messverfahren zwischen 1 und $\sqrt{2}$ liegt. Als Korrekturfaktor für die χ^2-verteilte Statistik des Goodness of Fit-Tests wurde entsprechend ein zwischen 1 und 0,5 liegender Wert vorgeschlagen.

Die folgende Tabelle 13 gibt in den mit „AMOS“ überschriebenen Spalten die standardisierten Schätzungen für die Pfadkoeffizienten zwischen den beiden Faktoren und den acht Diagnoseverfahren wieder. Die *Pfadkoeffizienten* wurden nur auf Basis der *Originalstichprobe* der 237 Patienten geschätzt, während die *Schätzungen der Standardfehler* durch eine *Bootstrapanalyse*[70] erfolgte, in der softwarebedingt die Korrelation zwischen den Augenseiten nicht berücksichtigt werden kann. Die Schätzung der Korrelationen zwischen den Fehlertermen dieser acht Messverfahren und ihrer Standardfehler erfolgte analog. Die Ergebnisse der Schätzungen zu den Korrelationskoeffizienten finden sich weiter unten in den mit „AMOS“ überschriebenen Spalten der Tabelle 16.

Die Ergebnisse der mit „PROC CALIS“ überschriebenen Spalten werden im folgenden Text unter Verweis auf Tabelle 13 sukzessive erläutert. Tabelle 13 fasst also auch bereits die Ergebnisse sowohl der nichtadjustierten als auch der adjustierten Bootstrapanalyse zu den Pfadkoeffizientenschätzungen zusammen, die weiter untern ausführlich besprochen werden und nimmt sie damit quasi vorweg.

[69] Martus, a. a. O.

[70] n=3.000 Resamples; 41dieser Bootstrapsamples wurden aus der Analyse ausgeschlossen wurden (siehe Martus (2001a), S. 270, Table 4, Fußnote 1).

Tabelle 13: Schätzungen der Pfadkoeffizienten zwischen den beiden Faktoren und den acht Diagnoseverfahren (via Originalsample) sowie der Standardfehler (via Bootstrapsamples)

Faktoren	Glaukomschweregrad			Konzentration		
Indikatoren	AMOS [1]	PROC CALIS [2] Bootstrap unter Ziehung einzelner Augen	PROC CALIS [2] Bootstrap unter Ziehung von Augenpaaren	AMOS [1]	PROC CALIS [3] Bootstrap unter Ziehung einzelner Augen	PROC CALIS [3] Bootstrap unter Ziehung von Augenpaaren
NRRA	0,63 (0,046)	0,63 (0,048)	0,63 (0,054)	_____	_____	_____
ERG	0,55 (0,037)	0,55 (0,038)	0,55 (0,044)	_____	_____	_____
VEP	0,50 (0,048)	0,50 (0,049)	0,49 (0,057)	_____	_____	_____
AMP	0,33 (0,057)	0,33 (0,057)	0,33 (0,073)	0,21 (0,086)	0,21 (0,092)	0,21 (0,120)
FLI	0,58 (0,050)	0,58 (0,051)	0,59 (0,065)	0,37 (0,077)	0,36 / 0,37 (0,082)	0,35 / 0,36 (0,108)
CLV	0,76 (0,045)	0,76 (0,047)	0,76 (0,055)	0,43 (0,094)	0,41 / 0,42 (0,096)	0,40 / 0,42 (0,120)
MD	0,81 (0,052)	0,81 (0,052)	0,81 (0,065)	0,47 (0,083)	0,45 / 0,46 (0,101)	0,44 / 0,46 (0,127)
DYN	0,66 (0,049)	0,66 (0,051)	0,67 (0,063)	0,45 (0,076)	0,45 (0,093)	0,44 / 0,45 (0,118)

1) standardisierter Pfadkoeffizient für den Glaukomschweregrad bzw. die Konzentration und SE; Pfadkoeffizientenschätzung auf Basis des Originalsamples, SE-Schätzung durch nicht nach Augenseitenkorrelation adjustierten Bootstrap

2) arithmetisches Mittel des standardisierter Pfadkoeffizient für den Glaukomschweregrad und SE auf Basis der Bootstrapanalyse

3) arithmetisches Mittel und – falls abweichend – Median des standardisierten Pfadkoeffizienten für die Konzentration sowie SE auf Basis der Bootstrapanalyse

8.2.2.2 Reproduktion der AMOS-Analyse durch PROC CALIS

In Vorbereitung der Bootstrapanalyse mit PROC CALIS wird zunächst versucht, die AMOS-Schätzung der Pfadkoeffizienten unter Anwendung von PROC CALIS nur auf Basis des Originalsamples zu reproduzieren. Unter Anwendung der Maximum Likelihood-Methode resultieren hierbei – gerundet auf zwei Nachkommastellen – identische Werte sowohl für die Pfadkoeffizienten beider Faktoren als auch die Fehlerkorrelationen. Die globale Anpassung des Modells an die Daten wird in beiden Modellen mit Hilfe einer unter Gültigkeit der Nullhypothese χ^2-verteilten Diskrepanzfunktion untersucht. In beiden Analysen ergibt sich ein identisches Diskrepanzmaß ($\chi^2 = 7{,}957$, df=10), das zur Nichtablehnung des vorgeschlagenen Modells führt (p=0,633). Da eine zufriedenstellende Gesamtanpassung nicht dafür Gewähr leistet, dass das Modell hinsichtlich aller Untergruppen der manifesten Variablen korrekt spezifiziert ist, werden sogenannte Modifikationsindizes (MIs) errechnet. MIs sind Schätzungen des Maßes, um das sich die Modellanpassung im Falle des Einschlusses neuer Parameter (z.

B. Fehlerkorrelationen) verbessern würde. Da alle MIs verhältnismäßig geringe Werte (< 3) annehmen, wird gefolgert, dass sich das vorgeschlagene Modell nicht mehr wesentlich verbessern lässt.

8.2.3 BOOTSTRAP MIT PROC CALIS UNTER ZIEHUNG EINZELNER AUGEN (NICHT NACH AUGENSEITENKORRELATION ADJUSTIERT)

Nachdem im ersten Schritt (Reproduktion) gezeigt wurde, dass die beiden Softwareanwendungen bei der Analyse des Originalsamples im wesentlichen gleiche Ergebnisse zeitigen, wird in einem zweiten Schritt versucht, die nicht nach der Augenseitenkorrelation adjustierte AMOS-Bootstrapanalyse zur Schätzung der SE mit PROC CALIS zu reproduzieren. Die nicht adjustierte PROC CALIS-Bootstrapanalyse bringt hierbei nicht nur Schätzungen der SE der Pfadkoeffizienten, sondern auch Schätzungen der Pfadkoeffizienten selbst hervor. In der AMOS-Analyse wurden nur die SE bootstrapbasiert ermittelt, die Pfadkoeffizienten selbst wurden auf Basis des Originalsamples – soll heißen ohne Anwendung eines Resamplingverfahrens – geschätzt.

8.2.3.1 Ergebnisse

8.2.3.1.1 Verteilungsform der Pfadkoeffizientenschätzungen

Entsprechend allen bisherigen Simulationsanalysen werden zur Berechnung von statistischen Kennzahlen grundsätzlich nicht die Rohwertschätzungen aus den Bootstrapsamples, sondern die Beträge derselben verwendet.

Bezüglich der acht standardisierten Pfadkoeffizienten für den Glaukomschweregrad gilt, dass die Beträge ungefähr normalverteilt sind, so dass hier ausschließlich das arithmetische Mittel der Beträge angegeben wird.

In Bezug auf die fünf standardisierten Pfadkoeffizienten für die Konzentration kann beobachtet werden, dass die Verteilung des Betrags der Schätzungen leicht linksschief ist (Ausnahme: Pfadkoeffizient von AMP[71]). Daher wird für diese neben dem arithmetischen Mittel der Median angegeben, sofern dieser vom arithmetischen Mittel abweicht.

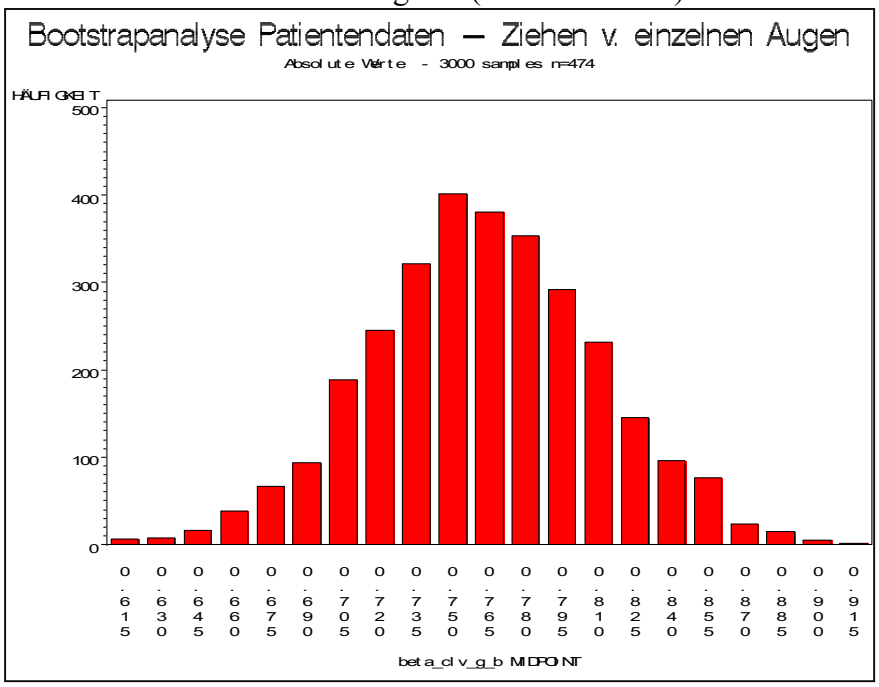
Typische Verteilungsformen der Pfadkoeffizienten der Faktoren Glaukomschweregrad und Konzentration im Vergleich

[71] Die Ausnahme geht auf das Vorkommen von Ausreißern im Bereich hoher Werte zurück.

Der Sachverhalt, dass die Schätzungen von Pfadkoeffizienten für den Glaukomschweregrad eher normal und solche für die Konzentration leicht linksschief verteilt sind, wird exemplarisch durch die folgenden vier Histogramme illustriert, die sich auf die manifeste Variable CLV beziehen.

Der Betrag der Schätzungen des Pfadkoeffizienten von CLV für den Glaukomschweregrad ist mit einer Schiefe von -0,008 nahezu symmetrisch verteilt (siehe Abbildung 72):

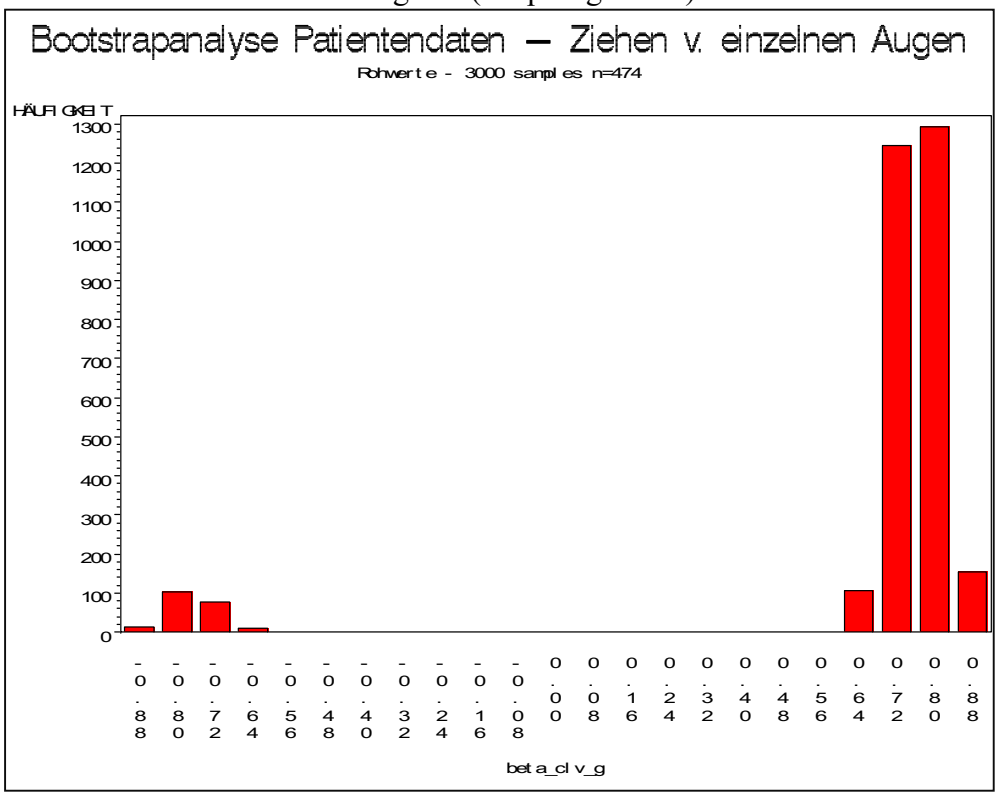
Abbildung 72: Verteilung der nichtadjustierten Pfadkoeffizientenschätzungen von CLV hinsichtlich des Glaukomschweregrads (absolute Werte)



Zum einen liegen die Modi der beiden zugrunde Rohwert-Subverteilungen[72] (siehe Abbildung 73) relativ weit voneinander entfernt, da dieses Messverfahren (so wie die anderen Diagnoseverfahren) den Faktor Krankheitsschweregrad verhältnismäßig gut annähert und somit für den entsprechenden Pfadkoeffizienten durchschnittlich relativ hohe Werte erzielt werden. Die große Distanz zwischen den Häufigkeitsgipfeln bedingt, dass sich die Verteilungen nicht überschneiden und somit die Betragsbildung nicht zu einer Häufung im Bereich der kleinen Werte der resultierenden Verteilung führt. Zum anderen weisen die beiden zugrunde Rohwert-Subverteilungen eine relativ gute Anpassung an die Normalverteilung auf.

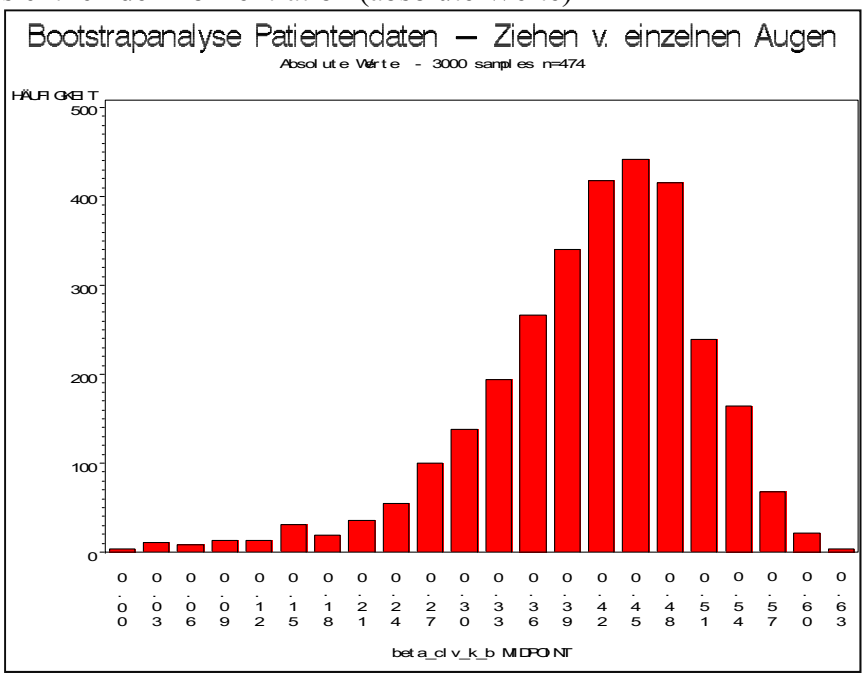
[72] Die größere der beiden Subverteilungen repräsentiert die vorzeichenmäßig korrekten, positiven Schätzungen, die zweite die aufgrund softwareinterner Vorzeichenumkehr zustande gekommenen negativen.

Abbildung 73 Verteilung der nichtadjustierten Pfadkoeffizientenschätzungen von CLV hinsichtlich des Glaukomschweregrads (Ursprungswerte)



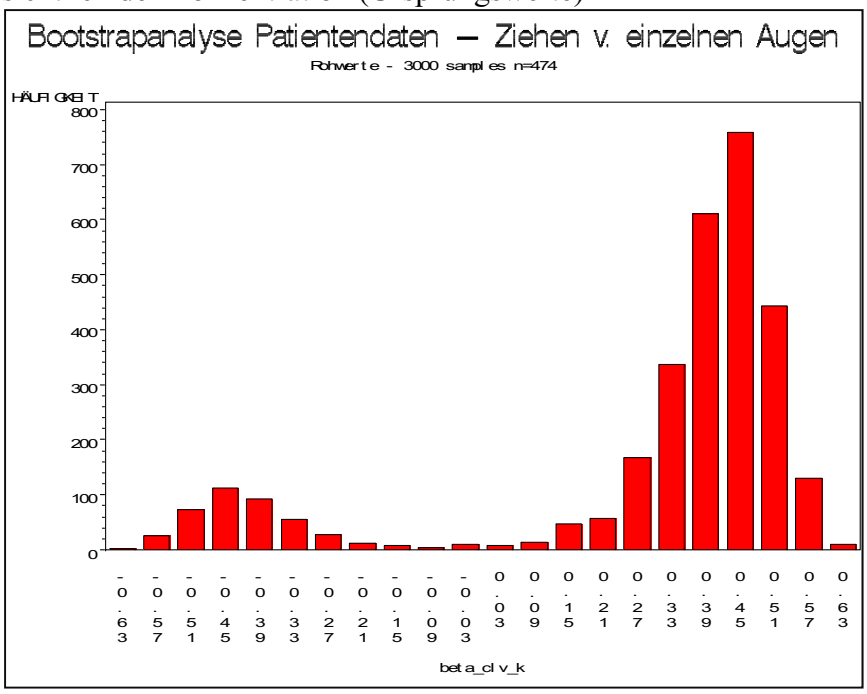
Für den Betrag der Schätzungen des Pfadkoeffizienten von CLV für die Konzentration hingegen gilt, dass er mit einer Schiefe von – 0,984 relativ linksschief verteilt ist (Abbildung 74).

Abbildung 74: Verteilung der nichtadjustierten Pfadkoeffizientenschätzungen von CLV hinsichtlich der Konzentration (absolute Werte)



Zum einen führt die verhältnismäßig geringe Distanz der Modi der beiden zugrunde Rohwert-Subverteilungen[73] (Abbildung 75)

Abbildung 75: Verteilung der nichtadjustierten Pfadkoeffizientenschätzungen von CLV hinsichtlich der Konzentration (Ursprungswerte)



zur Überschneidung derselben und nach Betragsbildung zu einer Häufung am linken Rand der resultierenden Verteilung. Zum anderen tendiert die Subverteilung der mit korrektem Vorzeichen versehenen Schätzungen zur Linksschiefe, während die Verteilung der Schätzungen mit falschem Vorzeichen entsprechend zur Rechtsschiefe neigt.

8.2.3.1.2 Statistische Maßzahlen

Pfadkoeffizienten des Faktors Glaukomschweregrad

Das arithmetische Mittel des Betrags der Pfadkoeffizientenschätzung aus der (nicht adjustierten) PROC CALIS-Bootstrapanalyse stimmt bezüglich aller acht Pfadkoeffizienten mit der AMOS-Schätzung nur auf Basis des Originalsamples überein. Die simulierten Standardfehler der Pfadkoeffizienten aus der PROC CALIS-Bootstrapanalyse stimmen auch überein oder liegen minimal (um maximal 0,002) über jenen, die im AMOS-Bootstrap ermittelt wurden (siehe Tabelle 13, Spalten 1 und 2).

Pfadkoeffizienten des Faktors Konzentration

[73] Die größere der beiden Subverteilungen repräsentiert die vorzeichenmäßig korrekten, positiven Schätzungen, die zweite die aufgrund softwareinterner Vorzeichenumkehr zustande gekommenen negativen.

Die arithmetischen Mittelwerte des Betrags der Pfadkoeffizienten von AMP (0,21) und von DYN (0,45) stimmen mit den AMOS-Schätzungen auf Basis der Originalstichprobe überein. Die arithmetischen Mittelwerte jener von FLI (0,36), CLV (0,41) und MD (0,45) liegen aufgrund der Linksschiefe geringfügig unterhalb der AMOS-Schätzungen. Die Mediane 0,37, 0,42 bzw. 0,46 nähern letztere etwas besser an.

Die mit PROC CALIS ermittelten simulierten Standardfehler der Pfadkoeffizienten liegen teils minimal (CLV), teils geringfügig (AMP und FLI) und teils deutlich (bei MD und DYN um gut 20 %) über den mit AMOS geschätzten SE (siehe

Tabelle 13, Spalten 4 und 5).[74]

Fehlertermkorrelationen

Die Analyse der Korrelationen zwischen einigen Fehlertermen der acht Messverfahren auf Basis des (nicht adjustierten) Bootstraps mit PROC CALIS geschieht teils auf Grundlage der Rohwertschätzungen und teils aufgrund des Betrags der Schätzungen. Der nach der AMOS-Analyse (nur auf Basis der Originalstichprobe) zu erwartende durchschnittliche Zusammenhang zwischen den Fehlern von NRRA und DYN, von VEP und AMP sowie von FLI und CLV beträgt ca. 0,2 oder mehr, so dass eine Betragsbildung angezeigt ist. Dahingegen sind bei den Fehlerkorrelationen zwischen AMP und FLI sowie VEP und MD Werte unter 0,1 zu erwarten, was eine Analyse der Rohwerte sinnvoll erscheinen lässt.

Die aus der Bootstrapanalyse mit PROC CALIS resultierenden Verteilungen weisen aufgrund von extremen Ausreißern große Variabilität und starke Asymmetrie auf. (Tabelle 14 und ausführlicher der entsprechende Output im Anhang).

Tabelle 14: Fünf Fehlerkorrelationen des Modells zur Diagnose des Glaukomschweregrads - Kennwerte der nichtadjustierten Bootstrapanalyse

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum
the1_b	Betr. v. Fehlerkorr nrra dyn	3000	0.2160446	0.2598212	0.000520494
the2_b	Betr. v. Fehlerkorr vep amp	3000	0.2065980	0.0753520	0.0267752
the3	Fehlerkorr vep md	3000	0.0527833	0.2265815	-3.2364501
the4	Fehlerkorr amp fli	3000	0.0830577	0.2476719	-8.3110187
the5_b	Betr. v. Fehlerkorr fli clv	3000	0.3310936	0.2436320	0.1164457

Variable	Etikett	Maximum
the1_b	Betr. v. Fehlerkorr nrra dyn	6.5484930
the2_b	Betr. v. Fehlerkorr vep amp	8.1133929
the4	Fehlerkorr amp fli	3.9459805

[74] Die Vergleichbarkeit der SEs ist insofern nur eingeschränkt möglich, als dass einerseits in der nicht adjustierten AMOS-Bootstrapanalyse aufgrund von Instabilität 41 Samples ausgeschlossen worden waren, wohingegen in der PROC CALIS-Bootstrapanalyse alle Samples verwandt wurden. In der adjustierten und nicht adjustierten PROC CALIS-Bootstrapanalyse wurde insgesamt in nur drei (von 6000) Bootstrapsamples darauf hingewiesen, dass lineare Abhängigkeit zwischen Parameterschätzern bestünde, was nicht zum Ausschluss der betroffenen Resamples führte.

the5_b	Betr. v. Fehlerkorr fli clv	7.7063156

Es kommen Ausreißer vor, die den – nicht repräsentativen – Mittelwert um bis zu 150-fach überschreiten und der Betrag der Schiefe der fünf Verteilungen liegt zwischen 13 und 19.

Vor Berechnung arithmetischer Mittelwerte und Standardfehler der Fehlerkorrelationen auf Basis der Bootstrapsamples ist es daher nötig, die Verteilungen zu *winsorieren*, um den Einfluss der Ausreißer zu eliminieren. Um die Vergleichbarkeit sowohl zwischen den fünf Fehlerkorrelationen als auch zwischen den beiden Ziehungsvarianten des Bootstrapverfahrens zu gewährleisten, werden die Verteilungen der Fehlerkorrelationen dabei auf dieselbe Weise behandelt. Als Maßstab dient die Verteilung mit der größten Zahl von Ausreißern, die Verteilung der Fehlerkorrelation zwischen NRRA und DYN in der Bootstrapanalyse mit Ziehung von Augenpaaren, die weiter unten behandelt wird. Diese Verteilung ist dadurch gekennzeichnet, dass die 60 größten Bootstrapschätzungen Werte über 0,53 annehmen, dem 2.8-fachen des Medians. Daher erscheint es sinnvoll, die zwei Prozent größten und zwei Prozent kleinsten Werte einer jeden Fehlerkorrelationsverteilung abzuschneiden und die Analyse der Verteilungen der Fehlerkorrelationen beider Ziehungsmodi fortan auf Basis von n=2880 Bootstrapsamples durchzuführen. Die winsorierten Verteilungen, die in Tabelle 15 charakterisiert sind, weisen Schiefen im Bereich zwischen –0.05 und 0,19 auf, so dass aussagekräftige arithmetische Mittelwerte und Standardfehler aus ihnen berechnet werden können.

Tabelle 15: Fünf Fehlerkorrelationen - Kennwerte der winsorierten Verteilungen

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum
the1_b	Betr. v. Fehlerkorr nrra dyn	2880	0.1968826	0.0614983	0.0616048
the2_b	Betr. v. Fehlerkorr vep amp	2880	0.2047535	0.0462231	0.0971547
the3	Fehlerkorr vep md	2880	0.0455524	0.0698502	-0.1198205
the4	Fehlerkorr amp fli	2880	0.0875532	0.0469262	-0.0261081
the5_b	Betr. v. Fehlerkorr fli clv	2880	0.3171670	0.0596536	0.1903539

Variable	Etikett	Maximum
the1_b	Betr. v. Fehlerkorr nrra dyn	0.3639634
the2_b	Betr. v. Fehlerkorr vep amp	0.3128790
the3	Fehlerkorr vep md	0.2149054
the4	Fehlerkorr amp fli	0.1947942
the5_b	Betr. v. Fehlerkorr fli clv	0.4687250

Der Vergleich der mit „AMOS“ und der benachbarten mit „PROC CALIS“ überschriebenen Spalte der Tabelle 16 zeigt Folgendes: Die arithmetischen Mittel der Korrelationskoeffizienten der winsorierten Verteilungen aus der (nicht adjustierten) PROC CALIS-Bootstrapanalyse stimmen mit den nur auf Basis des ursprünglichen Samples produzierten AMOS-Schätzungen überein bzw. überschreiten sie minimal (Fehlerkorrelation zwischen NRRA und DYN).

Die simulierten Standardfehler der Korrelationskoeffizienten auf Basis der mittleren 96 % der Datenverteilungen unterschätzen die SE-Schätzungen leicht (um 10–14 %), die aus der (nicht adjustierten) AMOS-Bootstrapanalyse hervorgingen.[75]

Tabelle 16: Vergleich der Kennwerte von fünf Fehlerkorrelationen auf Basis von Originalsample sowie nichtadjustierten und adjustierten Bootstrapanalysen

	AMOS [1]	PROC CALIS [2] Bootstrap unter Ziehung einzelner Augen	PROC CALIS [2, 3]. Bootstrap unter Ziehung von Augenpaaren
Fk1 (Nrra, Dyn)	0,19 [76] (0,071)	0,20 (0,061)	0,20 (0,074)
Fk2 (Vep, Amp)	0,20 (0,051)	0,20 (0,046)	0,20 (0,058)
Fk3 (Vep, Md)	0,05 (0,079)	0,05 (0,070)	0,05 (0,080)
Fk4 (Amp, Fli)	0,09 (0,053)	0,09 (0,047)	0,09 (0,063)
Fk5 (Fli, Clv)	0,32 (0,068)	0,32 (0,060)	0,31 (0,073)

1) Fehlerkorrelation und SE; Fehlerkorrelationsschätzung auf Basis des Originalsamples, SE-Schätzung durch nicht nach Augenseitenkorrelation adjustierten Bootstrap

2) Arithmetisches Mittel und SE der Fehlerkorrelation auf Basis der Bootstrapanalyse mit winsorierten Verteilungen (n=2880 Bootstrapsamples)

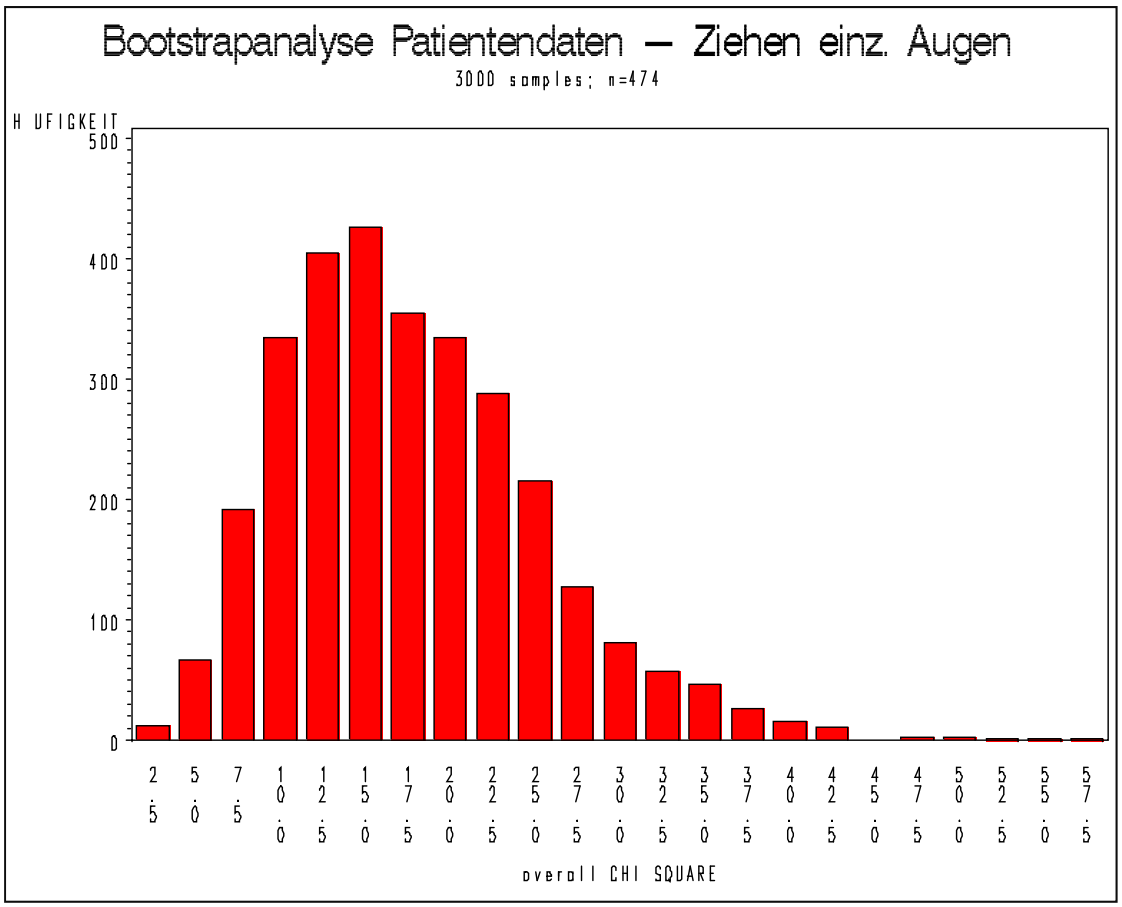
3) Diese Spalte ist inhaltlich dem unmittelbar folgenden Abschnitt 8.2.4 zuzuordnen und wurde nur zur tabellarischen Vergleichbarkeit beider Bootstrapvarianten in diesen Abschnitt *vorgezogen*.

8.2.3.1.3 Goodness of Fit

Für jedes der 3.000 Resamples wurde als globales Maß für die Güte der Anpassung des Modells an die Daten die Diskrepanzfunkion χ^2 bestimmt. Die Verteilung des χ^2 (siehe Abbildung 76) ist durch den Median 16,75 und die Schiefe 0,79 charakterisiert. Das (auf dem nicht adjustierten Bootstrap) basierende 95 %-Konfidenzintervall für das χ^2 ergibt sich aus dem 2,5. und dem 97,5. Perzentil dieser Verteilung (6,15; 35,13). Das 95 %-Quantil der χ^2-Verteilung unter der Nullhypothese bei $\nu = 10$, $\chi^2_{0,95,10}$, lautet 18,31. (1722 oder 57,4 % der 3000 auf Resamples basierenden Modelle weisen ein geringeres χ^2 auf und wären somit bei $\alpha = 0{,}05$ nicht abgelehnt worden.)

[75] Die SE-Schätzungen der AMOS-Bootstrapanalyse beruhen auf 98,7 % der Bootstrapsamples.
[76] Abweichend hiervon wird laut publiziertem Modell 0,20 geschätzt. Siehe Martus (2001a), S. 270, Table 4.

Abbildung 76: χ^2 -Verteilung des nichtadjustierten globalen Goodness of Fit-Tests



Fazit

1. Die Verteilungen der Pfadkoeffizienten des Faktors Glaukomschweregrad weisen eine annähernd normale Gestalt auf, während jene des Faktors Konzentration leicht linkschief sind. Dies hat folgende Konsequenzen:

- Die arithmetischen Mittelwerte der Glaukomschweregrad-Pfadkoeffizienten aus der PROC CALIS-Bootstrapanalyse reproduzieren die nur auf Basis der Originalstichprobe (unter Anwendung von PROC CALIS als auch AMOS identisch) ermittelten Pfadkoeffizientenschätzungen sehr genau. Bei einigen *auch* von der Konzentration abhängigen Messverfahren unterschreitet der Median der aus den Resamples resultierenden Pfadkoeffizientenschätzungen die Schätzungen leicht, die nur auf dem Sample beruhen.

- Die in der Bootstrapanalyse mit PROC CALIS ermittelten SE der Pfadkoeffizienten des Faktors Glaukomschweregrad reproduzieren die entsprechenden SE der AMOS-Bootstrapanalyse fast genau oder überschätzen sie minimal, während die SE der Pfadkoeffizienten des Faktors Konzentration hingegen stärker überschätzt werden.

2. Das aufgrund einer hohen Zahl teils extremer Ausreißer nötige Winsorieren der Verteilungen der Fehlertermkorrelationen *ermöglicht erst* eine angemessene Berechnung der Standardfehler aus den Verteilungen der Bootstrapsamples. Der Verzicht auf vier Prozent der Werte an den Verteilungsrändern erklärt hinreichend, dass die SE der AMOS-Bootstrapanalyse[77] leicht und relativ gleichmäßig unterschritten werden.

3. Die nichtadjustierte Bootstrapanalyse mit AMOS wurde durch die nichtadjustierte Bootstrapanalyse mit PROC CALIS um die Information erweitert, wie die χ^2 -verteilte Statistik des globalen Goodness of Fit-Tests verteilt ist. Die linke Grenze des 95 %-Konfidenzintervalls dieser Statistik ist deutlich von der 0 entfernt (6,15)

8.2.4 BOOTSTRAP MIT PROC CALIS UNTER ZIEHUNG VON AUGENPAAREN (NACH AUGENSEITENKORRELATION ADJUSTIERT)

Nach Durchführung der nicht nach Augenseitenkorrelation adjustierten Bootstrapanalysen mit den Softwaretools AMOS und PROC CALIS erfolgt eine adjustierte Bootstrapanalyse mit PROC CALIS. Einige im vorausgehenden Abschnitt über die nichtadjustierte PROC CALIS-Bootstrapanalyse beschriebenen Sachverhalte treffen auch für die adjustierte Analyse zu. Dies betrifft die unterschiedliche Verteilungsform von Pfadkoeffizienten des Faktors Glaukomschweregrad und Konzentration sowie das Vorkommen von Ausreißern bei den Bootstrapsamples zur Analyse der Fehlerkorrelationen. Die diesen Sachverhalten Rechnung tragenden Maßnahmen (Wahl der angemessenen Maßzahl für die zentrale Tendenz und das Winsorieren) erfolgen daher auch in der hier beschriebenen Analyse. Die Verwendung von absoluten Werten vs. Rohwerten bei der Berechnung statistischer Maßzahlen zur Charakterisierung der Fehlerkorrelationsverteilungen folgt ebenso dem vorausgehenden Abschnitt.

Daher werden diese auch in dieser Untersuchung gegebenen Sachverhalte und angewandten Methoden nicht näher erläutert.

Ergebnisse

Pfadkoeffizienten des Faktors Glaukomschweregrad

Der Vergleich der mit PROC CALIS überschriebenen Spalten 2 und 3 der obenstehenden Tabelle 13 ergibt folgende Resultate: Das arithmetische Mittel der Beträge der Pfadkoeffi-

[77] In der AMOS-Bootstrapanalyse wurden laut Publikation 41 von 3000 Resamples ausgeschlossen. Nähere Angaben, um welche es sich dabei handelt, geht aus der Publikation nicht hervor. Siehe Martus (2001a), S. 270, Table 4, Fußnote 1.

zientenschätzungen stimmt bezüglich aller acht Pfadkoeffizienten mit jenem der nichtadjustierten PROC CALIS-Schätzung überein, unterschreitet (VEP) oder überschreitet es (FLI und DYN) minimal. Die adjustierten Standardfehler aller Pfadkoeffizienten liegen deutlich über jenen, die im nicht adjustierten PROC CALIS-Bootstrapverfahren ermittelt wurden.[78]

Pfadkoeffizienten des Faktors Konzentration

Nur der arithmetische Mittelwert des Betrags der Pfadkoeffizientenschätzungen von AMP stimmt mit dem der nicht adjustierten PROC CALIS-Bootstrapanalyse nahezu überein. Die arithmetischen Mittelwerte (des Betrags) der anderen Pfadkoeffizientenschätzungen unterschreiten geringfügig jene der nichtadjustierten PROC CALIS-Bootstrapanalyse. Bis auf bei FLI stimmen die Mediane der Pfadkoeffizientenschätzungen in beiden Ziehungsvarianten der PROC CALIS-Bootstrapanalyse genau überein.

Die simulierten Standardfehler der Pfadkoeffizienten überschreiten jene der nicht adjustierten PROC CALIS-Bootstrapanalyse deutlich (

Tabelle 13, Spalten 5 und 6). Die Details werden im folgenden Abschnitt dargestellt.

Fehlertermkorrelationen

Aus dem Vergleich der Spalten 2 und 3 obiger Tabelle 16 geht hervor, dass die arithmetischen Mittel der adjustierten mit jenen aus der nicht adjustierten PROC CALIS-Bootstrapanalyse übereinstimmen bzw. sie minimal überschreiten (Fehlerkorrelation zwischen FLI und CLV).

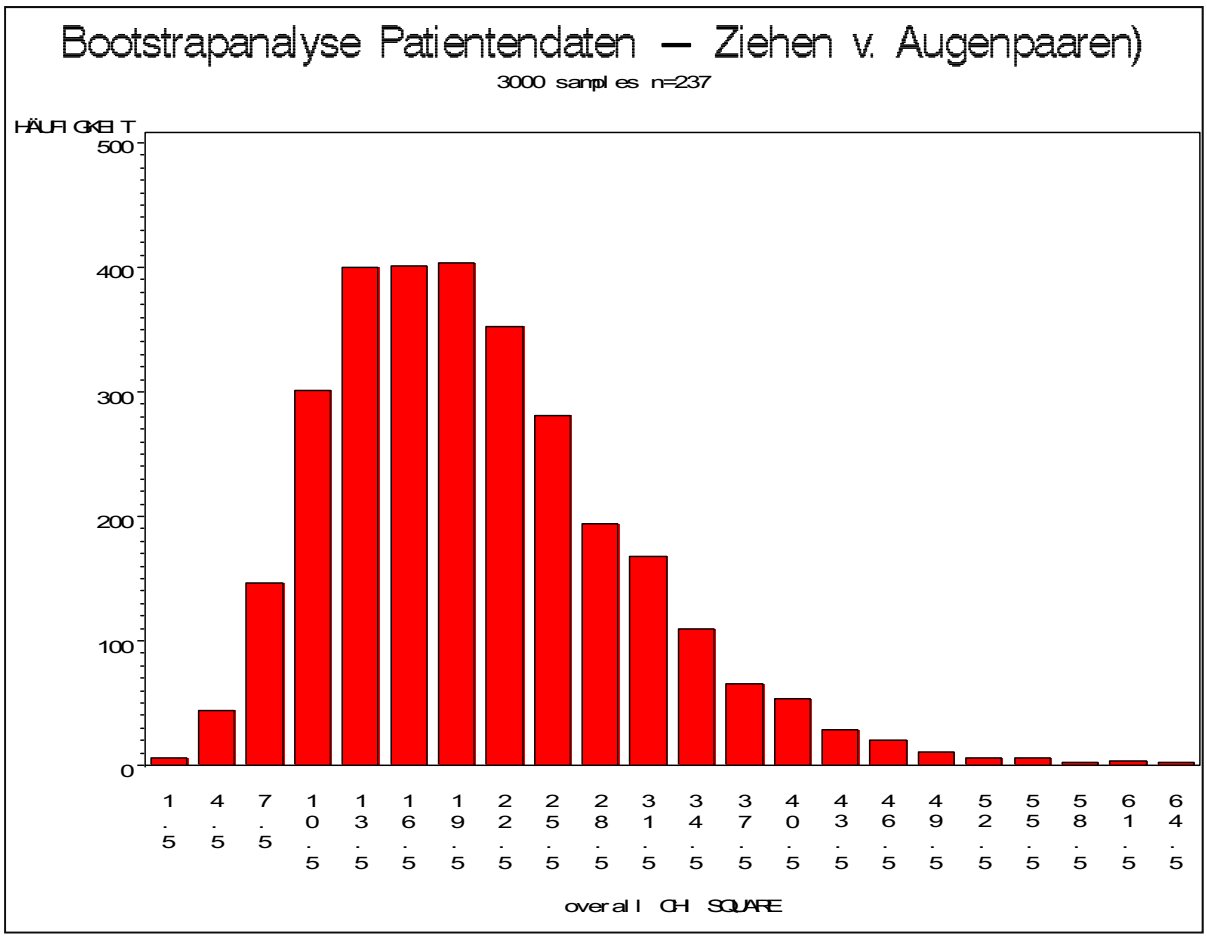
Die simulierten Standardfehler sind deutlich höher als die aus der nicht adjustierten PROC CALIS-Bootstrapanalyse.

Goodness of Fit

Die Verteilung des χ^2 des globalen Goodness of Fit-Tests ist in Abbildung 77 dargestellt. Der Median liegt bei 19,53 und die Schiefe beträgt 0,86. Das 95 %-KI für das χ^2 schließt den Bereich von 6,68 bis 42,34 ein. (1342 oder 44,7 % der 3000 auf Resamples beruhenden Modelle wären bei $\alpha = 0{,}05$ nicht abgelehnt worden.)

[78] Dieses pauschale Ergebnis wird in 8.2.5 quantifiziert.

Abbildung 77: χ^2 -Verteilung des adjustierten globalen Goodness of Fit-Tests



8.2.5 *VERGLEICH DER ADJUSTIERTEN UND NICHTADJUSTIERTEN BOOTSTRAPANALYSE*

Eine den ersten Abschnitt des Kapitels abschließende Untersuchung setzt die Ergebnisse aus den beiden vorausgehenden PROC CALIS-Bootstrapanalysen ins Verhältnis. Die Standardfehler der geschätzten Parameter aus der adjustierten, die Augenseitenkorrelation der Messverfahren berücksichtigenden Bootstrapanalyse werden durch die entsprechenden SE aus der nicht adjustierten Variante dividiert. Damit erhält man einen Faktor zur Korrektur von SE-Schätzungen, die aus einer nicht adjustierten Analyse hervorgehen.

Tabelle 17: Quotienten aus simulierten SE der Pfadkoeffizienten der adjustierten und nicht adjustierten Bootstrapanalyse (Korrekturfaktoren) sowie links-rechts-Korrelation der Messverfahren

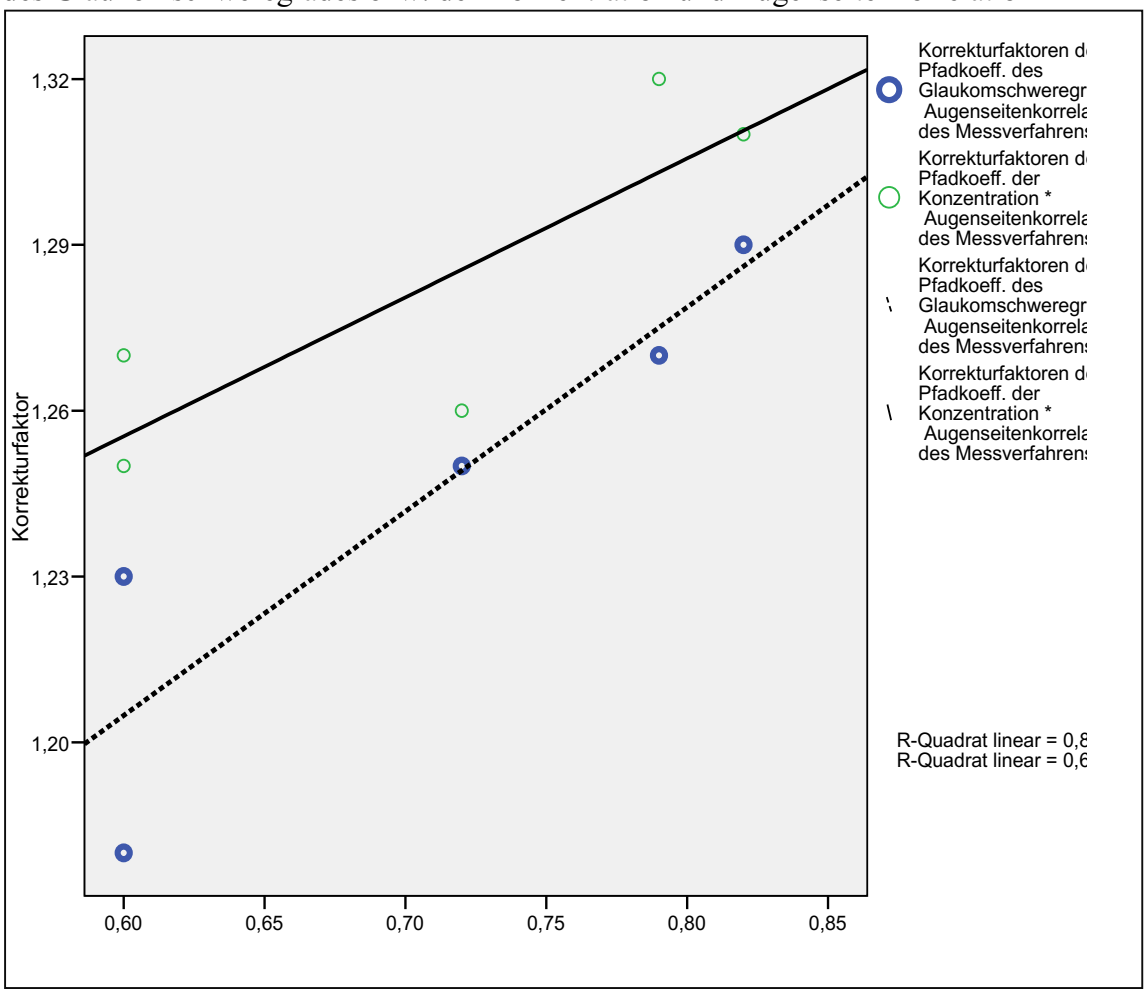
	Quotienten aus simulierten SE der Pfadkoeffizienten des Glaukomschweregrads	Quotienten aus simulierten SE der Pfadkoeffizienten der Konzentration	Augenseitenkorrelation des Messverfahrens
NRRA	1,13	-----	0,62
ERG	1,18	-----	0,66
VEP	1,18	-----	0,62
AMP	1,29	1,31	0,82
FLI	1,27	1,32	0,79
CLV	1,18	1,25	0,60
MD	1,25	1,26	0,72
DYN	1,23	1,27	0,60

Obige Tabelle 17 zeigt, dass die *simulierten Standardfehler der Pfadkoeffizienten* des Faktors Glaukomschweregrad in der nichtadjustierten Bootstrapanalyse mit PROC CALIS um das 1,13 bis 1,27-fache zu korrigieren sind, um Standardfehlerschätzungen zu erhalten, die die Augenseitenkorrelation der Messverfahren berücksichtigen. Die Standardfehler der Pfadkoeffizienten des Faktors Konzentration aus der nichtadjustierten Analyse sind sogar um das 1,25- bis 1,32-fache höher anzusetzen.

Daraus kann man schließen, dass die Komponente der links-rechts-Korrelation der Messverfahren, die auf den Faktor Konzentration zurückgeht, höher ist als jene, die auf den Glaukomschweregrad zurückzuführen ist. Dieser Schluss ist insofern medizinisch plausibel, als das Maß der Konzentration des Patienten bei Durchführung der Messverfahren an beiden Augenseiten stärker übereinstimmen dürfte als der Schweregrad des Glaukomschadens an beiden Augenseiten. Die Verschiedenheit der Messergebnisse ein- und desselben Verfahrens am linken und rechten Auge dürfte somit eher auf den unterschiedlichen Glaukomschweregrad oder in der Messapparatur liegende Fehlerquellen zurückzuführen sein.

Trägt man die Werte der Tabelle 17 in einem Streudiagramm (siehe Abbildung 78) an, so ergibt sich ein hoher positiver linearer Zusammenhang (R^2=0,82) zwischen der Augenseitenkorrelation der Messverfahren und dem jeweiligen Korrekturfaktor für die Pfadkoeffizienten des Glaukomschweregrades (untere Regressionsgerade). Nicht ganz so hoch (R^2 = 0,70) ist der lineare Zusammenhang zwischen der Augenseitenkorrelation der Messverfahren und dem jeweiligen Korrekturfaktor für die Pfadkoeffizienten der Konzentration (obere Regressionsgerade).Der Shift der oberen Regressionsgerade (bezüglich der unteren) verdeutlicht das oben beschriebene höhere Niveau der Korrekturfaktoren der Pfadkoeffizienten des Glaukomschweregrades.[79]

Abbildung 78: Linearer Zusammenhang zwischen Korrekturfaktor für die Pfadkoeffizienten des Glaukomschweregrades bzw. der Konzentration und Augenseitenkorrelation



Im Teil A (Vergleich der Bootstrapverfahren auf Basis *simulierter* Daten) wurde gezeigt, dass die *Höhe des Korrekturfaktors von der Augenseitenkorrelation der Fehlerterme und des*

[79] Der Steigungsunterschied der beiden Geraden ist so gering, dass auf eine Interpretation dieses Unterschieds verzichtet wird.

Glaukomschweregrades abhängt, wobei erstere einen größeren Einfluss besitzt. In der Analyse auf Basis der simulierten Daten liegt allerdings der Sonderfall vor, dass für alle drei Messverfahren standardisierte (Faktor- und Fehler-)Pfadkoeffizienten gegeben sind, deren Höhe zufallsbedingt um den Wert 0,71 streut. Daher ist in der Simulationsanalyse die Höhe der Pfadkoeffizienten als Einflussgröße auf den Korrekturfaktor ohne Bedeutung. In der *Analyse der Patientendaten* ist die Höhe der Pfadkoeffizienten der Messverfahren verschieden und daher bedeutungsvoll. Dass die Korrekturfaktoren des SE der Pfadkoeffizienten für den Faktor Konzentration größer sind als jene für den Faktor Glaukomschweregrad, ist auch dadurch erklärbar, dass der Einfluss der (Störgröße) Konzentration auf die Messverfahren geringer ist als der Einfluss der Glaukomschweregrades: Die Fehlertermkomponente eines von beiden Faktoren beeinflussten Diagnoseverfahrens, die dem Faktor Konzentration zuzuordnen ist, ist größer als die Fehlertermkomponente, die dem Faktor Glaukomschweregrad zuzuordnen ist. Da nun die Augenseitenkorrelation der Fehlerterme einen größeren Einfluss auf den Korrekturfaktor des SE des Pfadkoeffizienten besitzt als jene des Faktors, sind die Korrekturfaktoren für den SE der Pfadkoeffizienten des Faktors Konzentration höher.

Daten über *die Augenseitenkorrelation der Faktoren* Glaukomschweregrad und Konzentration als auch die Augenseitenkorrelation der *Fehlerterme* liegen in dieser Untersuchung der Patientendaten nicht vor.

Daher wird im Teil B (Vergleich der Bootstrapverfahren auf Basis von Patientendaten) ersatzweise der Einfluss der *Augenseitenkorrelationen der Messverfahren* auf die Korrekturfaktoren untersucht. Man geht hierbei davon aus, dass die Augenseitenkorrelation zweier Messverfahren durch drei Komponenten erklärt werden kann: die Augenseitenkorrelation des Glaukomschweregrades, die Augenseitenkorrelation der Fehlerterme und die Höhe der Pfadkoeffizienten.

Da die ersten beiden dieser drei Bestimmungsstücke unbekannt sind, ist nur der Vergleich der Augenseitenkorrelationen der Messverfahren und der Korrekturfaktoren für die SE der Pfadkoeffizienten des Glaukomschweregrades in Tabelle 17 möglich. Ein Vergleich (ohne Berücksichtigung der Größe der Pfadkoeffizienten) zeigt, dass relativ geringe Augenseitenkorrelationen (im Bereich zwischen 0,60 und 0,62) mit relativ geringen Korrekturfaktoren (im Bereich zwischen 1,13 und 1,18) einhergehen. Nur im Fall des Messverfahrens DYN geht eine verhältnismäßig geringe Augenseitenkorrelation (0,60) mit einem mittelhohen Korrekturfaktor (1,23) einher. Relativ hohe Augenseitenkorrelationen (im Bereich zwischen 0,72 und 0,82) führen zu hohen Korrekturfaktoren (zwischen 1,25 und 1,29). Ein ähnlicher Zusammenhang

besteht zwischen der Augenseitenkorrelation der Messverfahren und den Korrekturfaktoren für die Pfadkoeffizienten der Konzentration.[80]

Aus Tabelle 18 ist ersichtlich, dass die *simulierten Standardfehler der Fehlerkorrelationen* in der nichtadjustierten Bootstrapanalyse mit PROC CALIS um das 1,15 bis 1,33-fache zu korrigieren sind.

Tabelle 18: Quotienten aus SE der Fehlerkorrelationen der adjustierten und nicht adjustierten Bootstrapanalyse

Fk1 (Nrra, Dyn)	1,21
Fk2 (Vep, Amp)	1,25
Fk3 (Vep, Md)	1,15
Fk4 (Amp, Fli)	1,33
Fk5 (Fli, Clv)	1,22

8.2.5.1 Diskussion und Fazit

1. *Methodische Anmerkung:* Die Winsorierung der Bootstrapsampleverteilung der fünf Fehlerkorrelationen ist eine Möglichkeit, der Verzerrung der statistischen Maßzahlen entgegenzuwirken, die zunächst auf Basis der unbehandelten Bootstrapsampleverteilungen ermittelt wurden. In Tabelle 14 sind die aus der *nichtadjustierten* Bootstrapvariante resultierenden Standardfehler der fünf Fehlerkorrelationen verzeichnet, die sich auf Basis aller 3000 Bootstrapsamples ergeben. Diese simulierten Standardfehler sind offensichtlich um ein vielfaches zu hoch. (Die analog simulierten Standardfehler der Fehlerkorrelationen der *adjustierten* Bootstrapvariante sind nicht angegeben worden.) Berechnet man die Quotienten aus den simulierten Standardfehlern beider Bootstrapvarianten, zeigt sich, dass auch die Quotienten deutlich oberhalb des Bereichs liegen, der zum einen laut bisherigem Forschungsstand für den Korrekturfaktor des SE angenommen wurde. Zum anderen übertreffen diese SE-Quotienten für die Fehlerkorrelationen auch die Quotienten deutlich, die sich für die Pfadkoeffizienten ergeben (siehe Tabelle 19).

Tabelle 19: Korrekturfaktoren für die nichtadjustierten SE der fünf Fehlerkorrelationen aus dem Modell zur Diagnose des Glaukomschwergrades mit acht Indikatoren

Quot_cor_err_nrr-dyn	Quot_cor_ err_vep_ amp	Quot_cor_err_vp_md	Quot_cor_err_amp_fli	Quot_cor_err_fli_clv
1.54936	1.74529	1.73342	1.41473	1.45876

Auch wenn das Verfahren der Winsorierung als suboptimal kritisiert werden kann, insofern der simulierte SE dadurch unterschätzt werde, kann das Verfahren als unproblematisch erachtet werden. Eine mögliche SE-Unterschätzung durch das Winsorieren ist nicht problematisch, da durch die Quotientenbildung der Standardfehler beider Ziehungsmodi zur Ermittlung von Korrekturfaktoren die Unterschätzung egalisiert wird.

[80] Siehe 11.6.4

2. Die Bootstrapanalysen auf Basis der Patientendaten unter Anwendung des Modells mit acht Messverfahren bestätigen die Ergebnisse der Untersuchungen auf Grundlage der simulierten Daten *tendenziell*. Je größer die Augenseitenkorrelation der Messverfahren, desto größer der notwendige Korrekturfaktor zur Adjustierung. Ob die Ergebnisse der Bootstrapanalyse der Patientendaten jene der simulierten Daten *en detail* bestätigt, ist noch ungeklärt, da die Augenseitenkorrelation der Faktoren Glaukomschweregrad und Konzentration sowie der Fehlerterme im Falle des Modells mit acht Diagnoseverfahren unbekannt und möglicherweise auch nicht oder nur unter großem Aufwand ermittelbar sind. Weitere Untersuchungen zu dieser Frage werden im zweiten Abschnitt des Kapitels (8.3) behandelt werden.

3. Bislang wurde zur Adjustierung des SE von Pfadkoeffizienten, die in nicht nach Seitenkorrelation der Messverfahren adjustierten Modellen geschätzt wurden, ein Faktor im Bereich von 1-$\sqrt{2}$ angenommen. Was die Höhe des notwendigen Korrekturfaktors im Modell mit acht Messverfahren betrifft, so kann man im Wesentlichen folgende, den bisherigen Forschungsstand erweiternde Resultate festhalten:

a) Die Augenseitenkorrelation des Messverfahrens, die leicht unmittelbar (ohne Modelle der CFA) errechnet werden können, liefert keine zureichende Information über die notwendige Höhe des Korrekturfaktors. (Dieses Resultat war zwar durchaus zu erwarten, aber lag bislang noch nicht vor, da keine Berechnungen zu Korrekturfaktoren vorlagen.)

b) Auf Basis der hier analysierten Stichprobe ergeben sich für die simulierten SE der Pfadkoeffizienten des Glaukomschweregrades Korrekturfaktoren im Bereich zwischen 1,13 (NRRA) und 1,29 (AMP). Für jeden einzelnen Parameter (es wurde hier die Auswahl von acht Pfadkoeffizienten des Faktors Glaukomschweregrad, fünf Pfadkoeffizienten des Faktors Konzentration und fünf Fehlerkorrelationen präsentiert) kann ein spezifischer Korrekturfaktor zur Adjustierung des Standardfehlers angegeben werden (Tabelle 17 u. Tabelle 18).

c) Für bestimmte *Gruppen von Diagnoseverfahren* können nun Größenordnungen für den Korrekturfaktor angegeben werden. Die SE der Pfadkoeffizienten des *morphologischen Verfahren*s NRRA sind im geringsten Maße korrekturbedürftig. Die Korrekturfaktoren für die anderen beiden Gruppen, *elektrophysiologische und psychophysischen Verfahren,* liegen zwischen 1,18 und 1,29 bzw. 1,27 und sind mithin, bezogen auf die jeweilige Gruppe, nicht homogen. Die Korrekturfaktoren für die SE der Pfadkoeffizienten des *Faktors Konzentration* liegen zwischen 1,25 und 1,32. Jene für den SE der *Fehlerkorrelationen* lagen im Bereich zwischen 1,15 und 1,33. Letzterer betrifft die Fehlerkorrelation zwischen den Fehlern der bei-

den Messverfahren AMP und FLI. Für die Pfadkoeffizienten-SE dieser beiden Messverfahren ergaben sich, bezüglich aller acht Messverfahren, die höchsten Korrekturfaktoren.

d) Im Falle des vorliegenden Modells wurden Korrekturfaktoren für den SE für insgesamt 18 zu schätzende Parameter (siehe 3b) bestimmt, von welchen der kleinste 1,13 und der größte 1,33 beträgt[81]. Die Verwendung eines Korrekturfaktors für den SE von 1,41 [82] geht von einer maximalen Augenseitenkorrelation der Messverfahren aus und wäre in diesem Fall also deutlich zu konservativ. Die erhöhende Korrektur der Standardfehler führt dazu, dass im Test der Nullhypothese, ein interessierender Pfadkoeffizient betrage 0, unter sonst gleichen Bedingungen eher mit Nichtablehnung der Nullhypothese zu rechnen ist. Die in diesem Modell durchgeführten Hypothesentests der Pfadkoeffizienten kommen jedoch auch bei adjustierten SEs nicht zu abweichenden Ergebnissen.

4. Bislang wurde zur Adjustierung der χ^2-basierten Statistik, wie sie für den globalen Goodness of Fit-Test verwendet wird, ein Faktor im Bereich von 1–0,5 angenommen. (Diese Adjustierung der χ^2-basierten Statistik entspricht der Adjustierung der SE im Bereich von 11,41).

Gemäß dieser Annahme liegt die Wahrscheinlichkeit, die Nullhypothese, dass das unterstellte Modell zutrifft, fälschlicherweise abzulehnen, nach Adjustierung in einem höheren Bereich. Man geht daher davon aus, dass sich qua Adjustierung die Evidenz des unterstellten Modells verbessert. Der Vergleich der aus der adjustierten und nicht adjustierten Bootstrapanalyse resultierenden χ^2-Verteilungen führt zu einem Ergebnis, das im Widerspruch zu dieser Annahme steht: Die linke Grenze des 95 %-KI entfernt sich durch die Adjustierung um 6,68 - 6,14 = 0,54 Einheiten von der 0 und das mediane χ^2 nimmt um 19,53 – 16,75 = 2,78 zu. Daher ist sogar von einer leichten *Verschlechterung der Evidenz des Modells* nach Adjustierung nach der Seitenabhängigkeit des Auges auszugehen. Der adjustierte p-Wert würde den nichtadjustierten, auf Basis des Originalsamples ermittelten unterschreiten.

5. Man könnte sich als Anwender wünschen, dass eine weniger aufwendige Methode zur unverzerrten Schätzung der SE zur Verfügung stünde, als eine nach Seitenabhängigkeit adjus-

[81] Die acht Korrekturfaktoren für den SE der Pfadkoeffizienten der *Fehlerterme*, die auch zu berücksichtigen wären, sind hierbei außer Acht gelassen worden, da sie nicht Gegenstand der bisherigen Untersuchung waren. Sie sind teilweise deutlich größer als 1,33, aber bislang noch keiner Prüfung (Inspektíon der zugrunde liegenden Verteilungen) unterzogen worden.
[82] Martus et al. (2000) S. 1108, Table 3.

tierte Bootstrapanalyse. Wünschenswert wäre eine Option unter PROC CALIS, Cluster von individuellen Patienten (oder allgemein: Untersuchungseinheiten) zu definieren, die jeweils aus diversen Messungen am linken und rechten Auge (oder allgemein aus dieser Untersuchungseinheit zugehörigen Messwerten) bestehen.[83] Die in Punkt 3 und 4 besprochenen Probleme der Schätzung von Korrekturfaktoren für den SE von geschätzten Modellparametern und die χ^2-Statistik würden sich gar nicht stellen, wenn im Rahmen der Standardsoftware Cluster durch den Benutzer definiert werden könnten.

8.3 ZWEITER ABSCHNITT: DAS VERHÄLTNIS ZWISCHEN AUGENSEITENKORRELATION UND KORREKTURFAKTOR – ÜBERPRÜFUNG DER SIMULATIONSERGEBNISSE (1. ABSCHN., TEIL A) DURCH EMPIRISCHE PATIENTENDATEN

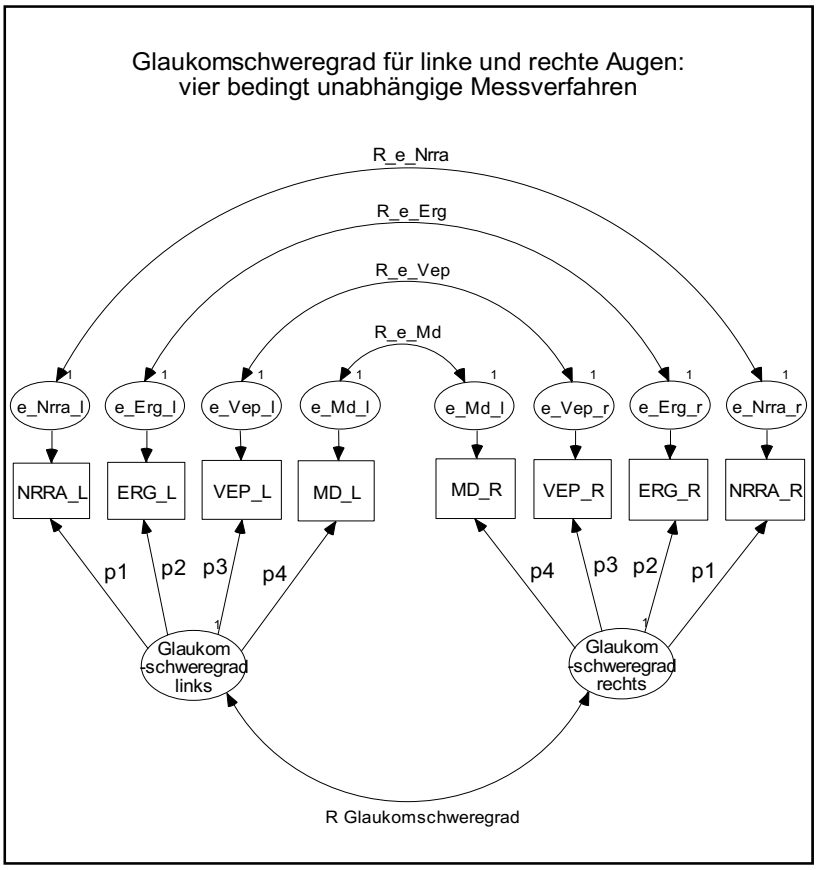
Ob die Ergebnisse der Bootstrapanalyse der Patientendaten jene der simulierten Daten *en detail* bestätigen, ist bislang noch nicht geklärt worden (siehe Punkt 2 in 8.2.5.1) Eine genauere Aussage hinsichtlich der Bestätigung der Analyseergebnisse simulierter Daten durch Untersuchungsresultate empirischer (Patienten-)Daten wird nur unter Durchführung eines Vergleichs zweier Modelle möglich, die im folgenden beschrieben werden (M1 und M2):

8.3.1 M1: MODELL DER CFA IN DER GLAUKOMDIAGNOSE ZUR SCHÄTZUNG DER AUGENSEITENKORRELATION LATENTER VARIABLER (PRO AUGENSEITE EIN FAKTOR UND VIER MESSVERFAHREN)

Modell 1 (M1): Ein publiziertes Modell der CFA, das – für jede Augenseite separat – von einer latenten Variablen (Glaukomschweregrad) und vier manifesten Variablen (Messverfahren) ausgeht, die bedingt (auf den Glaukomschweregrad) unabhängig sind.[84] (Siehe Abbildung 79):

Abbildung 79: M1: Messmodell zur Schätzung der Augenseitenkorrelation latenter Variabler (pro Augenseite ein Faktor und vier Messverfahren)

[83] Die Definition solcher Cluster ist in manchen statistischen Prozeduren möglich, die bei der Analyse paariger Organe anwendbar sind. Siehe z. B.: SAS Institute Inc (2003), The PHREG Procedure. Example 54.9: Analysis of Clustered Data.
[84] Martus (2001b), S. 932.



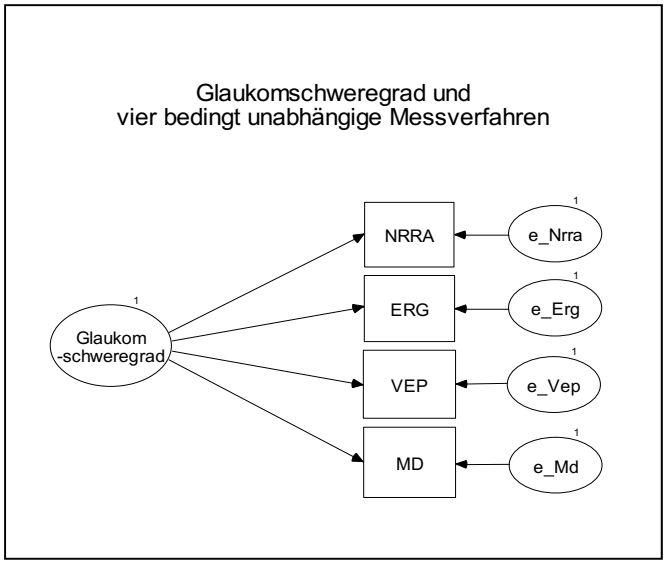
Das Modell gleicht strukturell dem *Modell der Datensimulation* (Abbildung 69) in Teil A des ersten Abschnitts (8.2.1.1). Es unterscheidet sich von jenem nur dadurch, dass ein weiteres Messverfahren hinzutritt und sowohl die vier Augenseitenkorrelationen der Messfehler als auch die Pfadkoeffizienten p1, p2, p3 und p4 hier selbstverständlich unterschiedliche Werte annehmen können. (Diese Parameter werden im Modell geschätzt.) Das Modell der Datensimulation hingegen sah vor, dass sowohl die drei Augenseitenkorrelationen der Messfehler als auch alle sechs Pfadkoeffizienten (bis auf zufallsbedingte Abweichungen) jeweils identische Werte annehmen. Es wird in diesem Modell vereinfachend angenommen, dass die Pfadkoeffizienten des Glaukomschweregrades p1, …, p4 für das linke und rechte Auge jeweils identisch sind.

Die vier Messverfahren sind aus dem Modell mit acht Diagnoseverfahren bereits bekannt. Es handelt es sich um die drei Diagnoseverfahren, welche im o. g. Modell durch die Konzentration nicht beeinflusst werden, neuroretinale Randsaumfläche (NRRA), Elektroretinogramm (ERG) und das visuell evozierte Potential (VEP), sowie um das Verfahren mit der größten Assoziation zum Glaukomschweregrad, dem mittleren Schaden (MD).

8.3.2 M2: MODELL ZUR SCHÄTZUNG VON KORREKTURFAKTOREN FÜR DAS MODELL 1 (EIN FAKTOR UND VIER MESSVERFAHREN)

Das Modell 2 (M2), welches die latente Variable Glaukomschweregrad und die vier manifesten bedingt unabhängigen Diagnoseverfahren beinhaltet, gleicht strukturell dem *Analysemodell der simulierten Daten* (Abbildung 70) in 8.2.1 und unterscheidet sich von diesem nur dadurch, dass es ein viertes Messverfahren umfasst. (Siehe Abbildung 80:)

Abbildung 80: M2: Messmodell mit einem Faktor und vier Messverfahren zur Schätzung von Korrekturfaktoren für M1



Die Parameter beider Modelle wurden auf Basis exakt derselben Stichprobe von Patienten des Glaukomregisters geschätzt, die bei der Parameterschätzung für das Modell mit acht Diagnoseverfahren verwendet wurde (n=237).

8.3.3 ERGEBNISSE

In folgender Tabelle 20 sind zunächst ein Maß für die Güte der Anpassung des Modells 1 an die Daten (p = 0,71) und die Augenseitenkorrelation der latenten Variablen Glaukomschweregrad dokumentiert, die auf Basis des Modells 1 geschätzt wurde (Zeilen 2 und 3). Sowohl die beobachteten als auch die durch das Modell 1 vorhergesagten Augenseitenkorrelationen der einzelnen Messverfahren finden sich in Spalte 1. Bezüglich der manifesten Variablen des Modells 1 wurden Pfadkoeffizienten (p1 bis p4) und SE geschätzt, die in der zweiten Spalte verzeichnet sind (Zeile 1 einer jeden Zelle). In Spalte 2 sind ebenso die Augenseitenkorrelationen der Fehlerterme dieser manifesten Variablen vermerkt (Zeile 2 einer jeden Zelle).

Diesen letztgenannten Parameterschätzungen auf Basis des Modells 1 werden gegenübergestellt die Pfadkoeffizientenschätzungen nebst SE, die sich aus den beiden Ziehungsvarianten beim Bootstrap anhand von Modell 2 ergeben (Spalten 3 und 4). Die fünfte Spalte schließlich enthält den Quotienten aus den SE der nach der Augenseitenkorrelation adjustierten und der nicht adjustierten Variante des Modells 2.

Tabelle 20: Aus M1 und M2 resultierende Pfadkoeffizientenschätzungen und SE (adjustierte sowie nicht adjustierte Bootstrapanalyse und Korrekturfaktoren)

	Modell 1 [1]	Modell 2 [7]		Korrekturfaktor [8]
		Nicht adjustiert	Adjustiert	
$\chi^2(DF)$	15,2 (19)	----	----	----
R (Gl, Gr) [2]	0,67 (0,056)	----	----	----
NRRA 0,62 [3] 0,65 [4]	0,59 (0,049) [5] 0,55 (0,069) [6]	0,62 (0,045) [5]	0,62 (0,050)	1,12
ERG 0,66 0,65	0,54 (0,049) 0,64 (0,043)	0,56 (0,037)	0,56 (0,043)	1,17
VEP 0,62 0,66	0,56 (0,050) 0,65 (0,054)	0,51 (0,047)	0,51 (0,057)	1,20
MD 0,72 0,73	0,82 (0,062) 0,84 (0,077)	0,82 (0,044)	0,82 (0,055)	1,24

1) Schätzungen unter Verwendung von AMOS, SE auf Basis von 3000 Bootstrapsamples
2) Augenseitenkorrelation (SE) des (latenten) Glaukomschweregrades
3) Beobachtete Augenseitenkorrelationen des (manifesten) Messverfahrens
4) Vorhergesagte Augenseitenkorrelationen des (manifesten) Messverfahrens
5) Pfadkoeffizient (SE)
6) Augenseitenkorrelationen des (latenten) Fehlerterms (SE)
7) Schätzungen unter Verwendung von PROC CALIS, Pfadkoeffizienten und SE derselben auf Basis von 3000 Bootstrapsamples
8) Quotient aus SE der adjustierten und nicht adjustierten Analyse mit Modell 2

Die aus Modell 2 resultierenden Schätzungen für die Pfadkoeffizienten weichen geringfügig von jenen aus Modell 1 ab[85].

8.3.4 AUFBEREITUNG DER ERGEBNISSE AUS DER ANALYSE MIT SIMULIERTEN DATEN NACH MAßGABE DER PATIENTENDATENERGEBNISSE

Die in Tabelle 20 präsentierten Ergebnisse sollen mit den Resultaten verglichen werden, die sich aus den Bootstrapanalysen auf Basis simulierter Daten ergaben und in

[85] Die größte Abweichung zeigt sich beim Pfadkoeffizienten des Verfahrens VEP. In einer hier nicht präsentierten Variante des Modells 1, das separate Schätzungen für die linke und rechte Augenseite im Falle des Pfadkoeffizienten von VEP (p2) erlaubte, resultierten stark abweichende Schätzungen von 0,49 (links) und 0,62 (rechts). Möglicherweise ist die schlechtere Bewertung des Verfahrens VEP durch das Modell 2 auf diese starke rechts-links-Divergenz zurückzuführen.

Tabelle 11 und Tabelle 12 dokumentiert wurden. Letztere wurden angestellt, um bei gegebenen Pfadkoeffizienten (zufallsbedingte geringe Abweichungen vom Wert 0,71) und für bestimmte Kombinationen von Glaukomschweregrad- und Fehlerterm-Seitenkorrelationen (Tab. zu Abbildung 69 in den Endnoten) Korrekturfaktoren für den SE des Pfadkoeffizienten zu ermitteln.Da nur für vier verschiedene Korrelationskoeffizienten (0, 0,5, 0,7 und 0,9) Daten simuliert wurden, sollen zur *Herstellung der Vergleichbarkeit mit Ergebnissen in* Tabelle 19 die Simulationsergebnisse aggregiert und anschließend Interpolationen durchgeführt werden.

8.3.5 AGGREGIERUNG SIMULIERTER KORREKTURFAKTOREN

Die Augenseitenkorrelation des Schweregrades beträgt laut Modell 1 0,67 und jene der Fehlerterme 0,55 (bezüglich NRRA), 0,64 und 0,65 (für ERG bzw. VEP) sowie 0,84 (MD betreffend). Diese *aus den Patientendaten ermittelten* Korrelationsstrukturen liegen somit im Bereich drei verschiedener Kombinationen von *simulierten* Glaukomschweregrad- bzw. Fehlerterm-Seitenkorrelationen: (0,7, 0,5), (0,7, 0,7) und (0,7, 0,9). (Es handelt sich hierbei um die Korrelationsstrukturen K10, K11 und K12 in der Tabelle zur Abbildung 70.) Für jede dieser drei Korrelationsstrukturen werden laut Tabelle 12 sechs Korrekturfaktoren geschätzt, die in folgender Tabelle 21 nochmals wiedergegeben sind. (Sechs Korrekturfaktoren, da drei Messverfahren mit identisch angenommenen Pfadkoeffizienten zweier simulierter Populationen vorliegen.)

Tabelle 21: Schätzung von Korrekturfaktoren für drei ausgewählte Korrelationsstrukturen der Tabelle 11

K10	K11	K12
1,14	1,24	1,34
1,06	1,15	1,26
1,08	1,19	1,31
1,12	1,20	1,32
1,11	1,18	1,29
1,09	1,17	1,29

Der aus jeweils sechs Schätzungen berechnete Mittelwert[86] ist eine Schätzung für den bei gegebener simulierter Korrelationsstruktur zu erwartenden Korrekturfaktor und lautet für die drei Korrelationsstrukturen wie folgt:

K 10 (0,7, 0,5): 1,10

K 11 (0,7, 0,7): 1,185

[86] An Stelle des arithmetischen Mittelwertes wird im Folgenden der Median verwendet, wobei dieser vom Mittelwert kaum abweicht.

K 12 (0,7, 0,9): 1,30

8.3.6 INTERPOLATION ZWISCHEN AGGREGIERTEN KORREKTURFAKTOREN

Um auf Basis der simulierten Daten Korrekturfaktoren für Korrelationsstrukturen zu ermitteln, die den in der Analyse der Patientendaten sich ergebenden, empirischen Korrelationsstrukturen entsprechen, wird zwischen den drei oben angegebenen mittleren Korrekturfaktoren interpoliert.

Beispiel: Ein Korrekturfaktor auf Basis simulierter Daten für die empirische Korrelationsstruktur bezüglich des Messverfahrens NRRA (0,67, 0,55) ergibt sich, indem *zwischen dem Korrekturfaktor für K10 und K11 interpoliert* wird. Hierbei wird angenommen, dass die simulierte Augenseitenkorrelation für den Glaukomschweregrad, 0,70, nur unwesentlich von der Schweregradseitenkorrelation abweicht, die sich im Modell 1 auf Basis der Patientendaten herausstellte (0,67), zumal die Augenseitenkorrelation des Schweregrades die Höhe des Korrekturfaktors wesentlich geringer beeinflusst als die Augenseitenkorrelation der Fehlerterme.

Der *simulierte* Korrekturfaktor für die empirische Korrelationsstruktur bezüglich des Messverfahrens NRRA lautet also 1,10 + (1,185 - 1,10): 4 ~ 1,12125 ~1,12.

Da die Seitenkorrelationen der Fehlerterme der Messverfahren ERG und VEP (0,64 und 0,65) fast identisch sind, wird nur ein simulationsbasierter Korrekturfaktor berechnet, der den empirischen Korrelationsstrukturen beider Messverfahren entspricht: 1,10 + (1,185 - 1,10) : 4 * 3 = 1,16375 ~1,16.

Der simulierte Korrekturfaktor für die das Messverfahren MD betreffende empirische Korrelationsstruktur (0,67, 0,85) wir durch *Interpolation zwischen K11 und K12* gewonnen:

1,185 + (1,30 - 1,185) : 4 * 3 = 1,27125 ~ 1,27.

8.3.7 VERGLEICH DER EMPIRISCHEN KORREKTURFAKTOREN AUS M2 MIT ANGEMESSENEN SIMULIERTEN KORREKTURFAKTOREN

In folgender Übersicht werden die auf Basis des Modells 2 anhand der Patientendaten ermittelten *empirischen* Korrekturfaktoren den Korrekturfaktoren gegenübergestellt, die eben durch Interpolation zwischen mittleren *simulierten* Korrelationsstrukturen gewonnen wurden.

Tabelle 22: Pfadkoeffizienten und empirische Korrekturfaktoren aus M2 sowie angemessene simulierte Korrekturfaktoren

	Pfadkoeffizient (Determinationskoeffizient)	Empirischer Korrekturfaktor aus Modell 2	Durch Interpolation ermittelter Korrekturfaktor aus simulierten Daten
NRRA	0,59 (0,35)	1,12	1,12
ERG	0,54 (0,29)	1,17	1,16
VEP	0,56 (0,30)	1,20	1,16
MD	0,82 (0,67)	1,24	1,27

Die beiden gegenübergestellten Korrekturfaktoren für das Diagnoseverfahren NRRA stimmen überein. Dieses Ergebnis erstaunt aufgrund der Tatsache, dass der empirische Korrekturfaktor bei einem Faktorpfadkoeffizienten von 0,59 resultiert, während die Simulationsbedingungen vorsehen, dass alle Pfadkoeffizienten zufällig und geringfügig um den Wert 0,71 streuen. In der empirischen Situation determinieren der Faktor und der Fehlerterm die Varianz der manifesten Variablen zu 35 bzw. 65 %, während in der simulierten Situation die beiden latenten Variablen die Varianz der manifesten stets im gleichen Maße bestimmen. Daher war zu erwarten, dass der (ohnehin stärkere) Einfluss der Augenseitenkorrelation der Fehlerterme (0,55) stärker zu Buche schlage als der Einfluss der Augenseitenkorrelation des Glaukomschweregrades (0,67) und daher der empirisch ermittelte Korrekturfaktor größer als der simulierte sei.

Dieses erwartete, aber nicht eingetroffene Phänomen könnte zur Erklärung des Sachverhalts herangezogen werden, dass die empirischen Korrekturfaktoren für das Messverfahren ERG (1,17) und VEP (120) den interpolierten simulierten Korrekturfaktor (1,16) minimal bzw. geringfügig übertreffen. Der Unterschied zwischen Übereinstimmung der Korrekturfaktoren im Falle von NRRA und geringer Nichtübereinstimmung im Falle von ERG und VEP kann wiederum dadurch erklärt werden, dass der durch den NRRA-Fehlerterm erklärte Varianzanteil nur 65 % beträgt, während er bei ERG bzw. VEP auf 71 bzw. 70 % ansteigt.

Der hier gewählte Erklärungsansatz wird dadurch bestätigt, dass einzig beim besten Diagnoseverfahren MD der empirische Korrekturfaktor (1,24) den simulierten (1,27) unterschreitet: *Nur hier ist der durch den Fehlerterm erklärte Varianzanteil geringer als 50 %*, so dass der Einfluss der Augenseitenkorrelation des Fehlers auf die Höhe des Korrekturfaktors schwächer ist als unter Simulationsbedingungen (Fehlervarianzanteil gleich 50 %).

Zusammenfassend kann man konstatieren, dass die Größe des quadrierten standardisierten Pfadkoeffizienten, also die Reliabilität des Messverfahrens, einen zwar nennenswerten, aber verhältnismäßig geringen Einfluss auf die Höhe des Korrekturfaktors zu haben scheint. Am Beispiel des Messverfahrens NRRA zeigt sich:

Unter Simulationsbedingungen besitzen Faktor und Fehlerterm gleichgroßen Einfluss (der Determinationskoeffizient beträgt 50 %). Selbst wenn der empirische Determinationskoeffizient (0,5 * 0,7= 0,35) um 30 % geringer ist als der simulierte, ist keine Abweichung des empirischen Korrekturfaktors vom simulierten zu beobachten. Hierbei ist allerdings zu berücksichtigen, dass die die Höhe des Korrekturfaktors hauptsächlich bestimmende Augenseitenkorrelation des *Fehlers* nur gering ist (0,55). Erst bei noch größeren Abweichungen von dieser Simulationsbedingung (ERG, VEP) bzw. höheren Augenseitenkorrelationen des Fehlerterms (ERG, VEP und MD) weicht der empirische Korrekturfaktor vom simulierten ab. Die hier beobachteten Abweichungen betragen jedoch maximal gut 3 % (VEP: 1,20/1,16) und sind daher möglicherweise in der praktischen Anwendung zu vernachlässigen. Diese Aussage erfolgt unter Vorbehalt, da sie sich nur auf wenige obige Fallbeobachtungen stützt und noch nicht durch eine systematische Analyse belegt wurde.

9 Gesamtfazit aus den Analysen simulierter und empirischer Daten (Kap. 3–8)

Im ersten – systematisch-theoretischen – Teil der Arbeit werden auf Basis künstlich erzeugter und Daten und darauf aufbauender Simulationen Auswirkungen diverser Fehlspezifikationen von Messfehlerkorrelationen untersucht. Hierbei handelte es sich um Korrelationen, die zwischen den Fehlertermen *verschiedener* Messverfahren bestehen.

Die wichtigsten Ergebnisse der *Analysen künstlich erzeugter Daten mit AMOS,* die im Folgenden dargestellt werden*,* beziehen sich nur auf Modelle mit positiven Fehlerkorrelationen.

1. Aus der Analyse der Ein-Faktor-Modelle auf Basis künstlich erzeugter Daten ergibt sich als wichtigste qualitative Erkenntnis: Wenn die Fehlervariablen zweier Indikatoren innerhalb eines Messmodells positiv korreliert sind und diese Korrelation ignoriert wird, werden diese Indikatoren fälschlicherweise als besonders gute Indikatoren für eine zu messende Krankheit bewertet (3.5.2, Abbildung 14).

2. Die wesentlichen Resultate der Analyse der Zwei-Faktor-Modelle sind:

a) Ist ein fehlspezifiziertes Messmodell wie in 1. und ein weiteres, korrekt spezifiziertes Messmodell in ein Strukturmodell integriert, das die Abhängigkeit der Faktoren dieser beiden Messmodelle voneinander beschreibt, so wird diese Abhängigkeit unterschätzt. Die Dependenz zwischen den Faktoren wird durch den Pfadkoeffizienten des Strukturmodells beschrieben. Hierbei ist es einerlei, ob die Fehlspezifikation im Messmodell des exogenen oder endogenen Faktors auftritt. Die Fehlspezifikation einer positiven Fehlerkorrelation *in einem vollständigen SEM* [87]zieht also einen Bias in zweierlei Hinsicht nach sich (4.2.2, Modell 5 und 7.4.3, Z4).

b) Werden in einem vollständigen SEM in beiden Messmodellen positive Fehlerkorrelationen ignoriert, addieren sich die Effekte der beiden Messmodelle hinsichtlich der Unterschätzung der Abhängigkeit zwischen den Faktoren. (4.2.3, Modell 6, Abbildung 36).

c) Wenn die Fehlerkorrelationen zweier Indikatoren positiv korrelieren, die *unterschiedlichen* Messmodellen zugehören, und diese messmodellübergreifende Fehlerkorrelation in einem

[87] Siehe Abbildung 6.

vollständigen SEM ignoriert wird, resultiert eine Überschätzung des Pfadkoeffizienten des Strukturmodells (siehe 7.3.5).

d) Wenn in einem vollständigen SEM sowohl eine Fehlerkorrelation innerhalb eines Messmodells (Modell 5, Z4), also auch eine messmodellübergreifenden Fehlerkorrelation (Modell 8) enthalten sind, dann heben sich Unterschätzung und Überschätzung des Strukturmodellpfadkoeffizienten tendenziell auf. Dieser Sonderfall zeigt, dass es unzulässig ist, auf gute Anpassung eines Modells an die Daten zu schließen, wenn der Bias des Strukturmodellpfadkoeffizienten vernachlässigbar ist.

Nach der *Analyse der Simulationen mit PROC CALIS* auf Basis der oben besprochenen AMOS-Modelle mit einem Faktor ist im Wesentlichen festzuhalten:

a) Die mit PROC CALIS durchgeführten Simulationen auf Basis korrekt spezifizierter, in AMOS analysierter Modelle produzieren Pfadkoeffizientenmittelwerte, die erwartungstreue Schätzer der in AMOS berechneten Koeffizienten sind. In Simulationen ist in den fehlspezifizierten Modellvarianten bei den Pfadkoeffizientenschätzungen ein Bias zu beobachten, der im Durchschnitt dem in AMOS errechneten Bias gleicht.

b) Der Mittelwert der modellbasiert geschätzten Standardfehler überschätzt immer die empirischen Standardfehler. Die geschätzten Standardfehler sind mithin konservativ. Besonders konservativ sind sie in fehlspezifizierten Modellen, wobei das Ausmaß der Überschätzung mit der Anzahl der fehlspezifizierten Fehlerkorrelationen zunimmt. In den korrekt spezifizierten Modellen mit zwei positiv und jenen mit zwei negativ korrelierten Fehlern ist der empirische Standardfehler des Pfadkoeffizienten des nicht fehlerkorrelierten Indikators erhöht bzw. verringert.

Es wird ein Korrekturfaktor von 0,8 für den geschätzten Standardfehler vorgeschlagen, da die empirischen Standardfehler eines Faktorpfadkoeffizienten durch den geschätzten SE durchschnittlich um mindestens 25 Prozent überschätzt werden. Da Pfadkoeffizientenschätzungen innerhalb eines relativ kleinen Bereichs streuen, kann die beobachtete Korreliertheit der Pfadkoeffizienten untereinander vernachlässigt werden.

In *Ein-Faktor-Modellen mit vier oder fünf Fehlerkorrelationen* können nicht nur Pfadkoeffizienten, sondern auch ein bzw. zwei Fehlerkorrelationen geschätzt werden.[88] Die Fehlspezifikation von Fehlerkorrelationen wird in dieser Analyse nicht untersucht.

Mit den Simulationen zu diesen Modellen konnte zum einen gezeigt werden, dass durchschnittliche geschätzte *Fehlerkorrelationen* in der Simulation *den wahren Korrelationen* entsprechen, die auf Basis künstlich erzeugter Daten in AMOS berechnet wurden. Zum anderen stellte sich heraus, dass hierbei auch die durchschnittlichen geschätzten Pfadkoeffizienten die wahren Werte der Pfadkoeffizienten reproduzieren. Dies betrifft Modelle mit ein oder zwei zu schätzenden Fehlerkorrelationen auf Basis von Daten, die tatsächlich positive und/oder negative Fehlerkorrelationen aufweisen. Eine Abweichung der simulierten SE von den mittleren (modellbasiert) geschätzten SE tritt bei Schätzung von Fehlerkorrelationen im Gegensatz zur Schätzung von Pfadkoeffizienten nicht auf. Die SE negativer Fehlerkorrelationen sind ca. doppelt so groß wie jene positiver Fehlerkorrelationen.

Die wesentlichen Resultate der *Modelle mit zwei Faktoren* lauten: Simulierte und modellbasiert geschätzte SE des *Pfadkoeffizienten des Strukturmodells, gamma*, stimmen fast genau überein, so dass letztere verlässlich sind. In den Modellen mit abhängigen Faktoren erhöht eine korrekt spezifizierte Fehlerkorrelation zweier Indikatoren, die demselben Messmodell angehören, deutlich den simulierten SE der Pfadkoeffizienten des Strukturmodells, gamma. (Es werden die SE eines Modells stets mit jenen des Grundmodells verglichen, das keine Fehlerkorrelationen erhält). Wird eine solche messmodellspezifische Fehlerkorrelation fehlspezifiziert, verringert sich der SE von gamma ebenso deutlich. Eine *korrekt spezifizierte Fehlerkorrelation zweier Indikatoren, die verschiedenen Messmodellen angehören*, führt zu einer leichten Verringerung des SE von gamma. Die Fehlspezifikation einer solchen zieht eine nennenswerte Erhöhung des SE (gamma) nach sich. Die beiden Typen von Fehlerkorrelationen (messmodellspezifisch und -übergreifend) haben auf die Variation von gamma gegensätzliche Effekte. Treten beide Typen in einem Modell auf und werden fehlspezifiziert, so wirkt sich fast nur ersterer auf den SE von gamma aus, indem sie ihn deutlich senkt. Nur in Modellen mit mindestens einer fehlspezifizierten messmodellspezifischen Fehlerkorrelation wird ein

[88] Der Analyse derartiger Modelle in AMOS wurde kein eigenes Kapitel gewidmet. Die Ergebnisse der AMOS-Analysen bilden aber hier wie in der gesamten Arbeit die Grundlage der nachfolgenden Simulationen.

Bias von gamma beobachtet. Von Bias wird in diesem Kapitel dann gesprochen, wenn der Betrag von $\frac{[\bar{\hat{\gamma}} - E(\gamma)]}{SE(\hat{\gamma})}$ größer als Zwei ist.[89]

Die *SE der Pfadkoeffizienten der beiden Messmodelle* verhalten sich aufgrund der Abhängigkeit der jeweiligen Faktoren in einigen Modellen deutlich abweichend von den SE, die in den Simulationen zu den separaten Ein-Faktor-Modellen mit drei Indikatoren beobachtet wurden.[90] Dies betrifft das Modell mit einer korrekt spezifizierten messmodellspezifischen Fehlerkorrelation im ersten Messmodell und in noch stärkerem Maß das Modell mit einer korrekt spezifizierten (messmodellspezifischen) Korrelation im zweiten Messmodell. Die Fehlspezifikation der Fehlerkorrelationen in diesen beiden Modellen hingegen bewirkt tendenziell die bereits aus den Ein-Faktor-Modellen mit drei Indikatoren bekannten Veränderungen. Die Fehlspezifikation einer messmodellspezifischen Fehlerkorrelation im ersten „unabhängigen" Messmodell reduziert auch die SE der Pfadkoeffizienten des zweiten „abhängigen" Messmodells. Korrekt spezifizierte messmodellübergreifende Fehlerkorrelationen reduzieren die SE der Pfadkoeffizienten des abhängigen Messmodells, fehlspezifizierte messmodellübergreifende Fehlerkorrelationen erhöhen letztere deutlich.

Die modellbasiert geschätzten SE der *Pfadkoeffizienten der Messmodelle* sind – abhängig vom jeweiligen Modell in unterschiedlichem Ausmaß – größer als die simulierten und mithin konservativ.

Die *Veränderungen* (bezüglich des Grundmodells ohne Fehlerkorrelationen), die an den simulierten SE der Pfadkoeffizienten des Messmodells infolge der Fehlerkorrelationen und deren Spezifikation beobachtet werden, gehen in den meisten Modellen in die gleiche Richtung wie jene, die an den mittleren geschätzten SE beobachtbar sind, aber sind weitaus (oft um Faktor 2) größer. Da in nicht wenigen Modellen die beiden Arten von SE aber nicht einmal hinsichtlich der Veränderungsrichtung übereinstimmen, können zur Analyse der Veränderung die modellbasierten SE die simulierten nicht ersetzen.

In den Modellen, in welchen die *Fehlerkorrelationen geschätzt* werden, werden die wahren Werte von gamma, der Fehlerkorrelationen und der Messmodellpfadkoeffizienten durch die

---

[89] In dem den Simulationsanalysen mit PROC CALIS vorausgehenden entsprechenden Kapitel mit AMOS-Modellen (mit zwei Faktoren) wird von Bias hingegen dann gesprochen, wenn der geschätzte Pfadkoeffizient vom wahren Wert des Pfadkoeffizienten (auf zwei Dezimalstellen gerundet) abweicht.

[90] Ein Modell mit zwei Faktoren und sechs Indikatoren integriert zwei Messmodelle mit je einem Faktor und drei Indikatoren).

aus den Simulationen errechneten arithmetischen Mittelwerte dieser Parameter fast genau reproduziert. Bezogen auf das Grundmodell gleichen die Veränderung der SE des Strukturmodellpfadkoeffizienten gamma und der Messmodellpfadkoeffizienten infolge der Schätzung von Fehlerkorrelationen ungefähr jenen, die aus den entsprechenden Modellen bekannt sind, in welchen die Fehlerkorrelationen als bekannt vorausgesetzt werden[91] – mit einer Ausnahme: Im Modell mit einer geschätzten messmodellspezifischen Fehlerkorrelation im zweiten Messmodell sind die SE höher als in dem Modell, in dem diese Korrelation als bekannt vorausgesetzt wird. Werden messmodellspezifische Fehlerkorrelationen in *beiden* Messmodellen eines Strukturgleichungsmodells geschätzt, ist die Erhöhung der SE im zweiten Messmodell besonders hoch.

Der zweite, empirisch und an bisher publizierter Forschung orientierte Teil dieser Arbeit beschäftigt sich mit den Auswirkungen der Fehlspezifikation einer Korrelationsstruktur, die sich von jener im ersten Arbeitsteil analysierten unterscheidet. Es handelt sich um Korrelationen zwischen den Ergebnissen *eines* jeweiligen Messverfahrens, das sowohl am linken als auch am rechten Auge eines Patienten angewandt wird. Zur Analyse werden zunächst Bootstrapanalysen durchgeführt, die auf künstlich erzeugten Daten beruhen. Abschließend folgen Bootstrapanalysen, die auf empirischen Patientendaten basieren und den Kenntnisstand bisheriger Publikationen zu SEMs in der Glaukomdiagnose erweitern.

Die zentralen Ergebnisse der Bootstrapanalysen eines einfachen Messmodells auf Basis künstlich erzeugter Daten lauten:

Das Bootstrapverfahren, das die Korrelation zwischen den Ergebnissen eines linksseitig und rechtsseitig angewandten Messverfahrens berücksichtigt (nach Augenseitenabhängigkeit adjustiertes Bootstrapverfahren), bringt höhere Standardfehlerschätzungen der Pfadkoeffizienten hervor als jenes Bootstrapverfahren, das diese Korrelation ignoriert. Die Standardfehler sind umso größer, je höher die Augenseitenkorrelation ist. Man geht nun davon aus, dass sich die Korrelation zwischen den manifesten Ergebnissen eines linksseitig und rechtsseitig erhobenen Messverfahrens aus einer Korrelation des latenten (links- und rechtseitigen) Schweregrades und einer Korrelation der entsprechenden latenten Messfehler zusammensetzt. Die Untersuchungen zeigten klar, dass die Messfehlerkorrelation die Standardfehler wesentlich stärker erhöht als die Schweregradkorrelation. Werden Schweregrad als auch Messfehlerkorrelatio-

[91] In den Modellen, in welchen eine Fehlerkorrelation als bekannt vorausgesetzt wird, enthält das Modell eine Restriktion. Bsp.: $r(\varepsilon_1, \varepsilon_2)$ = 0,5 (wahrer Wert).

nen von 0,9 angenommen, so ergeben sich in der adjustierten Bootstrapanalyse SE-Schätzungen, die um den Faktor 1,33 höher sind als in der nichtadjustieren. Diese Beobachtung ist mit der bisherigen Annahme vereinbar, dass die Standardfehler um bis zu Faktor 1,41 bei maximaler Korrelation erhöht seien.

Die Analyse von Daten des Erlanger Glaukomregisters mittels Resamplingverfahren unter Verwendung von PROC CALIS zeigt auf, wie stark die empirische Augenseitenkorrelation der Messverfahren die Standardfehlerschätzung der Pfadkoeffizienten tatsächlich beeinflusst. Eine publizierte Bootstrapanalyse eines Modells mit AMOS [92], die die Seitenkorrelation nicht berücksichtigen konnte, kann somit hinsichtlich adjustierter SE-Schätzungen korrigiert werden: Es wurden Korrekturfaktoren für die Standardfehler der geschätzten Pfadkoeffizienten und Fehlerkorrelationen dieses Modells im Bereich zwischen 1,13 und 1,33 ermittelt. Mit dem Gewinn einer adjustierten SE-Schätzung ist eine Prüfung der Hypothese, der jeweils interessierende Koeffizient sei Null, unter Angabe einer korrekten Irrtumswahrscheinlichkeit möglich. Die vorläufigen, auf nichtadjustierten SE-Schätzungen basierenden Signifikanzaussagen können demnach bestätigt werden: Alle im o. g. Modell enthaltenen Koeffizienten sind signifikant verschieden von Null.

Mit der durch diese Untersuchung empirisch überprüften Annahme, die Standardfehler der Koeffizienten nähmen qua Adjustierung nach der Seitenkorrelation der Messergebnisse um bis zu 1,41 zu, ging die Annahme einher, die χ^2 -Statistik zur Prüfung der Anpassungsgüte des Modells würde sich verringern und im Extremfall (maximale Seitenkorrelation der Messverfahren) halbieren. Der Vergleich der der Verteilungen der χ^2 -Statistik, die sich aus der adjustierten und der nichtadjustierten Bootstrapanalyse des o. g. Modells ergeben, bestätigt jedoch letztere Annahme *nicht*. Die linke Grenze des 95 %-Konfidenzintervalls der χ^2 -Verteilung aus der adjustierten Bootstrapanalyse ist weiter von der Null entfernt als jene aus der nicht adjustierten.

Die Ergebnisse der Bootstrapanalyse eines einfachen Messmodells auf Basis künstlich erzeugter Daten (16 verschiedene Korrelationsstrukturen) wurden mit den Resultaten der Bootstrapanalyse eines fast identischen Messmodells auf der Grundlage von Daten des Erlanger Glaukomregisters (eine bestimmte Korrelationsstruktur) verglichen, das lediglich ein zusätzliches Messverfahren enthielt. In der Bootstrapanalyse auf Basis künstlich erzeugter Daten

[92] Martus (2001a), S. 269.

zeigte sich ein quantitativ gut beschreibbarer Zusammenhang zwischen der Korrelationsstruktur (Korrelation aus links- und rechtsseitigem Glaukomschweregrad in Kombination mit Korrelation aus links- und rechtsseitigen Messfehlern) und dem Korrekturfaktor (Faktor, um den sich die Standardfehler der geschätzten Koeffizienten erhöhen). Die auf Basis der Bootstrapanalyse der Daten des Erlanger Glaukomregisters gefundenen Werte für die Schweregrad- und die Messfehlerkorrelationen der Augenseiten sowie die Korrekturfaktoren stimmten recht gut mit jenen überein, die anhand der Analyseergebnisse auf Basis künstlich erzeugter Daten für diese spezifische empirische Korrelationsstruktur mittels Interpolation abgeschätzt werden konnten. Die Größe des quadrierten standardisierten Schweregrads-Pfadkoeffizienten, also die Reliabilität des Messverfahrens, scheint einen zwar nennenswerten, aber verhältnismäßig geringen Einfluss auf die Höhe des Korrekturfaktors zu haben. Dieses Ergebnis ist vorläufig, da es sich nur auf einige wenige Beobachtungen stützt.

10 Auswirkungen der Verletzung von Modellannahmen in SEM im Lichte der Forschungsliteratur und Diskussion

In diesem Kapitel sollen Modellannahmen behandelt werden, die in SEM eine bedeutende Rolle spielen, insofern deren (häufig vorkommende) Verletzung gravierende Konsequenzen beinhaltet. Es geht, wichtige Annahmen herausgreifend, um die Unidimensionalität des Konstrukts und die Unkorreliertheit der Messfehler. Aufgrund des Themas der gesamten Arbeit (Fehlspezifikation von Fehlerkorrelationen) steht letztere im Vordergrund. Es gibt eine große Anzahl von Veröffentlichungen zu der Frage, wie sich die Verletzung der Annahme unkorrelierter Fehlerterme in SEM auswirkt. Wesentliche Auswirkungen betreffen die Reliabilität des Testmodells, die Testkriteriums-Validität im Zusammenhang mit der Correction for attenuation, und die Anpassungsgüte eines Testmodells[93]. Jüngere Forschung hebt auf die Auswirkung von korrelierten Fehlern in SEM mit dem Schwerpunkt auf Itemparcelling ab.

10.1 ARBEITSPROGRAMM

Der Effekt der Multidimensionalität des Konstrukts auf die Reliabilität – eine der unten erwähnten Kenngrössen in der Modellbewertung – wird eingangs behandelt (10.3.1). Die Auswirkungen korrelierter Fehler auf die Reliabilität sollen eine zentrale Rolle einnehmen (10.3.2). Unter diesen Gliederungspunkt fällt auch die (am Ende des folgenden Abschnitts angekündigte) Darstellung einer Systematik von Messfehlern. Anschließend wird analysiert, wie sich das simultane Auftreten von Multidimensionalität und Fehlerkorreliertheit auswirkt (10.3.3). Im Prozess der Modellentwicklung hat die Einschätzung der Anpassungsgüte eines Modells (Fit) den größten Stellenwert[94]. Die Bewertung eines Modells anhand von Kenngrössen und die Feststellung von theoretischen und praktischen Erkenntnissen können erst im Anschluss an diesen Schritt erfolgen. Dennoch soll erst im zweiten Teil dieses Kapitels die Auswirkung der Verletzung der Unkorreliertheitsannahme auf Fit-Indizes untersucht werden (10.4). Anschließend wird anhand von Forschungsliteratur untersucht, wie sich Verletzungen von Modellannahmen in einem bereits vorgestellten, publizierten Modell zur Glaukomdiagnose auswirken (10.5). Die Eröffnung bildet hierbei eine Strukturanalyse dieses Referenzmodells mit zwei Faktoren und acht Diagnoseverfahren (10.5.1), die in einen Versuch mündet, zur Reliabilitätssteigerung weitere Methodenfaktoren zu implementieren (10.5.2). Es werden

[93] Heene et al (2012), S. 38.
[94] a. a. O., S. 36.

Reliabilitäts- und weitere Koeffizienten der klassischen Testtheorie sowie Cronbachs alpha in den resultierenden vier Modellvariationen berechnet (10.5.3) und die im Modell zum Glaukomschweregrad diagnostizierten Effekte im Hinblick auf eine mögliche Verzerrung von alpha diskutiert (10.5.4). Für die o. g. fünf Modelle werden Fit-Indizes berechnet und im Hinblick auf Forschungsliteratur diskutiert (10.5.5). Den Abschluss des Kapitels und der gesamten Arbeit bildet 10.6: Erstens wird die schon im publizierten Modell angewandte ML-Schätzung im Hinblick auf eine mögliche Verletzung von Modellannahmen (Unkorreliertheit der error scores; multivariate Normalverteilung der observed scores) diskutiert und gegebenenfalls ein alternativer globaler Modelltest vorgeschlagen. Zweitens wird die unerwartete Verteilung der Chi-Quadrat-Statistik im adjustierten Bootstrapverfahren (Kap.9) reflektiert und eine Datentransformation nahegelegt.

Bei „Correction for attenuation“ geht es um die Frage, wie die Korrelation zwischen zwei Variablen unter dem Einfluss korrelierter Error Scores verzerrt geschätzt werden kann und welche Möglichkeiten bestehen, diesen Bias zu korrigieren. Da die Arbeit nicht die Validierung von hier diskutierten Modellen durch ein Kriterium zum Gegenstand hat, weder durch ein äußeres noch ein inneres, ist Correction for attenuation in unserem Zusammenhang, obgleich theoretisch von Interesse, praktisch ohne Bedeutung. Itemparcelling, also das Kombinieren individueller items und die Verwendung dieser Kombinationen als beobachtete Variable (z. B. Summenscore) spielt in unserem Forschungskontext keine Rolle und ist daher auch kein Arbeitsprogrammpunkt.

10.2 BEWERTUNG VON MODELLEN DURCH RELIABILITÄTSKOEFFIZIENTEN

10.2.1 RELIABILITÄT UNTER EINFLUSS VON DIMENSIONALITÄT UND HOMOGENITÄT

Zwei grundlegende Konzepte der Bewertung von Modellen, allgemeiner ausgedrückt, von Messinstrumenten, liegen vor: Validität und Reliabilität[95]. Ein Messinstrument ist valide (gültig), genauer gesagt konstruktvalide, in dem Ausmaß, in dem es das misst, was es messen soll. Ein Messinstrument (Test, Experiment) ist reliabel (zuverlässig) in dem Grad, mit dem durch die Messung hervorgebrachte Ergebnisse bei Wiederholung stabil sind. Insofern die Variabilität der Messwerte über die Zeit hinweg auf systematische Unterschiede und nicht lediglich auf

[95] Simsek et al (2013), S. 2008.

messfehlerbestimmte Unterschiede zurückführbar sein muss, ist die Reliabilität eine Voraussetzung der Konstruktvalidität[96].

Hauptgegenstand der folgenden Darstellung ist die Reliabilität. Denn zum einen bildete die Frage, inwiefern die Fehlspezifikation von Fehlerkorrelationen die Pfadkoeffizienten des Glaukomschweregrades beeinflusst – eine Frage, die unmittelbar die Reliabilität des jeweiligen Messmodells betrifft – einen Schwerpunkt der bisherigen Analysen. Zum anderen geht aus der einschlägigen Forschungsliteratur hervor, dass u. a. die Verletzung der Annahme unkorrelierter Messfehler erhebliche Auswirkung auf die Reliabilität von SEM haben kann. Die Berücksichtigung dieser Effekte ist im Rahmen der Weiterentwickelung publizierter Messmodelle zur Glaukomdiagnose daher unerlässlich.

Reliabilität kann auf vier verschiedene Weisen bestimmt werden. Es kann eine Wiederholung der Prozedur vorgenommen werden (retest method), man kann ein paralleles (vergleichbares) Instrument anwenden (alternative form method), die Items eines Messinstruments können in zwei Gruppen aufgeteilt werden (split half method) und es kann erhoben werden, wie die Items eines Messinstruments miteinander zusammenhängen (internal consistency method). Die interne Konsistenz erfordert als einzige keine Wiederholung oder die Schwierigkeit der Itemgruppierung sondern nur eine Testdurchführung und sieht eine einzige Schätzung der Reliabilität vor (siehe unten). Die Reliabilität kann durch zahlreiche konkurrierende Koeffizienten geschätzt werden, deren bekanntester Cronbachs alpha ist.

Vor Auswahl des angemessenen Koeffizienten ist die Frage der Dimensionalität und der Grad der Homogenität des Messinstruments zu klären. Unidimensionale Instrumente weisen Items auf, die eine einzige latente Variable (Faktor) messen. Multidimensionale Messinstrumente bestehen aus Items, die zwei oder mehr latente Variablen messen. Bei den eindimensionalen unterscheidet man τ −parallele, kongenerische und (essentiell) τ −äquivalente[97] Messinstrumente, die sich durch den Grad der Itemhomogenität unterscheiden und Messmodelle genannt werden. Auch multidimensionale Messinstrumente können unter Bezugnahme auf diese Messmodelle definiert werden. M. B. Miller legte 1995 eine Formel zur Berechnung von Cronbachs alpha für ein multidimensionales Messinstrument mit orthogonalen Faktoren und

[96] Eid et al. (2011), S. 815. Siehe hierzu auch 10.3.2.5.1, Fußnote 154.
[97] Wenn die Instrumente τ −äquivalent oder essentiell τ −äquivalent sind, dann ist Cronbachs alpha exakt gleich der Reliabilität. Daher reicht es in diesem Zusammenhang aus, nur die Bezeichnung „τ −äquivalent“ zu verwenden, statt zu differenzieren. Siehe 2.10 und insbes. Osburn (2000), S. 345 und Komaroff (1997), S. 338.

(essentiell) τ −äquivalenten Items vor[98]. Die Definition der Reliabilität und die Charakterisierung dieser Messmodelle erfolgt auf Basis der CTT (siehe 2.10).

Die oben angesprochene Eigenschaft der Uni- oder Multidimensionalität eines Messinstruments hat ebenso Auswirkungen auf die Reliabilität. Einen Versuch, Koeffizienten zur Schätzung der Reliabilität auch im Hinblick auf die Dimensionalität des Konstrukts auszuwählen, unternahm Osburn (2000). Er führt zwei verschiedene Settings an, die zur Überschätzung der Reliabilität durch den klassischen Koeffizienten alpha führen können[99] und stellt dann fest: „Although it is true that coefficient alpha may sometimes be inflated because of the conditions mentioned above, a potentially more serious problem is the tendency of coefficient alphas to underestimate the true reliability when the data are multidimensional." Daher benennt er als Hauptziel seines Artikels die Untersuchung der Bedingung, unter welchen Cronbachs alpha die wahre Reliabilität unterschätzt bei Anwendung eines nicht τ −äquivalenten Messinstruments, also eines solchen, das entweder kongenerisch oder multidimensional ist. Die Tatsache, dass Koeffizient alpha bei Vorliegen von heterogenen Items erheblich unterschätzt werden könne, sei eng verknüpft mit der Multidimensionalität der latenten Variablen, die im Hintergrund der Komponenten (Items) stehen. Zusammenfassend könne man sagen, dass bei Eindimensionalität alpha eine angemessene Schätzung der Reliabilität liefere, während eine ernsthafte Unterschätzung resultiere, wenn ein oder mehr mäßig stark korrelierte Faktoren (im Bereich 0,2 bis 0,4) die Komponenten charakterisierten.[100]

In diesem Zusammenhang ist darauf hinzuweisen, dass die latenten Faktorwerte und die Error Scores, von welchen im Rahmen der Faktorenanalyse die Rede ist, nicht gleichzusetzen sind mit den True Scores und Messfehlern der klassischen Testtheorie[101]. Unten wird der Unterschied zwischen diesen Begriffen weiter herausgearbeitet werden.

---

[98] zitiert nach Simsek (2013), S. 2009

[99] Osburn 2000, S. 343: (1) coefficients *assign transient error* due to *differences in test administration*, temporary changes in the examinee, and so on *to True-Score-Variance* (Guio, 1965; Schmidt and Hunter, 1996) (2). In addition, if the errors made in responding to test items are positively correlated, coefficient alpha may be inflated.

[100] Osburn (2000) S. 344 f.

[101] Komaroff (1997) S. 340.

Im Einzelnen geht es um folgende Gesichtspunkte:
a) Dimensionalität des Konstrukts (wie viel Faktoren besitzt das Messinstrument?) – anders formuliert: Homogenität bzw. Heterogenität der Items eines Messinstruments
b) Qualität des zugrundeliegenden Messmodells (essentielle τ −Äquivalenz oder Kongenerizität?)
c) Einteilung der Items eines Messinstruments in Subgruppen (gruppierte oder ungruppierte Instrumente?)
d) Anzahl der Items eines Messinstruments
e) Annahme der Unkorreliertheit der Messfehler
(Hierbei ist eine Darstellung der verschiedenen Typen von Messfehlerkorrelationen nötig.)
f) Bedeutung der Methodenvarianz (method variance).

Die Verletzung der Annahme der Unkorreliertheit der Messfehler soll aufgrund des thematischen Schwerpunktes der vorliegenden Arbeit den größten Stellenwert besitzen. Die Frage der Messfehlerkorrelation wird nicht nur hinsichtlich des Effekts auf die Reliabilität, sondern auch hinsichtlich der Fit-Indizes bedacht werden. Hinsichtlich des letzten o. g. Gesichtspunkts, der Methodenvarianz, benannte Osburn (2000) mit „transient error" einen Aspekt des method bias, der dazu führt, dass Koeffizienten die Reliabilität überschätzen. Aufgrund seiner Allgemeinheit wird method bias in einem Unterpunkt ausführlich behandelt werden (10.3.2.2). Dabei wird vor allem auf einen sehr umfangreichen und grundlegenden Forschungsbeitrag (Podsakoff et al., 2003) Bezug genommen. Zentraler Ausgangspunkt ist dabei eine Darstellung der Systematik der Messfehler.

10.3 RELIABILITÄTSRELEVANTE SPEZIFIKA VON MESSINSTRUMENTEN

10.3.1 MULTIDIMENSIONALITÄT DES KONSTRUKTS (ITEMHETEROGENITÄT)

In der Literatur zum prominentesten Schätzer für die Reliabilität eines Messinstruments –

Cronbachs alpha (kurz: coefficient alpha oder alpha) – wird mitunter, so auch in Osburn (2000), der Verwendungszweck des Koeffizienten klargestellt: Es handelt sich um einen Schätzer für die Reliabilität von „composite measures", also solchen Messinstrumenten, die aus mehreren Bestandteilen (multiple components) zusammengesetzt sind. Homogen heißen solche zusammengesetzten Messinstrumente, die eine einzige latente Variable messen. Heterogen entsprechend solche, die zwei oder mehr latente Variable erfassen. Osburn setzt sich

zum Ziel, die Bedingungen zu untersuchen, unter welchen der Koeffizient alpha die wahre Reliabilität von nicht τ − äquivalenten Messmodellen[102] unterschätzt. Es handelt sich dabei um solche, die entweder kongenerisch oder multidimensional sind[103]. Er untersucht neben Koeffizient alpha neun Reliabilitätskoeffizienten, die alle mit Ziel entwickelt wurden, das Maß der Unterschätzung zu reduzieren, wenn die Komponenten eines Tests *nicht τ −äquivalent* sind. In diesem Zusammenhang ist bekannt: Wenn True-Score- und Fehlervarianzen verschieden sind, so wie bei Kongenerizität der Fall, und das Konstrukt gleichzeitig eindimensional ist, so liegt „Koeffizient alpha einigermaßen nah an der wahren Reliabilität".[104]

Zur Diskussion der Effekte ist es nützlich, den Schätzer Cronbachs alpha im Hinblick auf die Varianzkomponenten zu formulieren, wobei alpha eine Funktion der Anzahl der Komponenten (items) n, der durchschnittlichen Kovarianz der i Komponenten, Ave (σ_{ij}^2) und der durchschnittlichen Varianz der Komponenten des Messinstruments Ave(σ_i^2) ist.

$$\alpha = \frac{n^2 \text{Ave}\ (\sigma_{ij}^2)}{\text{n Ave}\ (\sigma_i^2) + n(n-1)\text{Ave}\ (\sigma_{ij})}$$

Es gelten nach Osburn folgende Zusammenhänge:

- wenn die Kovarianz, verglichen mit der Varianz, groß ist, dann wird alpha groß,

- wenn die Anzahl der Komponenten groß ist, dann wird alpha groß,

- wenn die Kovarianz relativ zur Varianz klein ist – entweder, weil Kovarianzen einheitlich klein sind oder weil die Kovarianzen sehr unterschiedlich sind –, dann wird alpha klein.

Eine weitere, die Struktur von Messinstrumenten betreffende Eigenschaft hat ebenso Einfluss auf die Reliabilität: Die mögliche Anordnung von Items in ein, zwei oder mehr Gruppen (Subsets).[105]

---

[102] Weiter unten wird betont, dass in unserem Zusammenhang τ −äquivalente Messmodelle und essentiell τ −äquivalente Messmodelle gleichwertig behandelt werden können, da der in der CTT ausgeführte Unterschied zwischen beiden auf die Varianz der True Score Variablen keinen Einfluss hat.

[103] Simsek et al. erwähnen in diesem Zusammenhang, dass Miller (1995) einen Koeffizienten für τ − äquivalente und *gleichzeitig* multidimensionale Konstrukte entwickelt habe. D. h., τ − Äquivalenz und Multidimensionalität schließen sich demnach nicht notwendigerweise aus. Siehe hierzu Simsek et al. (2013), S. 2009.

[104] Osburn (2000), S. 345.

[105] Prominentes Beispiel ist der SF 36 mit den beiden Subtests „physische und psychische Gesundheit".

Zur Untersuchung der Frage, wie stark die Heterogenität der Daten die Reliabilität beeinflusst, simulierte Osburn u. a. Daten eines Messinstruments mit acht ungruppierten kongenerischen Indikatoren (components), das wegen der formalen Vergleichbarkeit mit dem publizierten Messmodell zur Erhebung des Glaukomschweregrades (Martus, 2001a) beispielhaft herausgegriffen wird. Hierbei wurden Reliabilitäten für die acht Komponenten zwischen 0,22 und 0,28 simuliert, die True-Score-Varianz wurde systematisch variiert und die True-Score-Korrelation zwischen den ersten und den letzten vier Items je nach Heterogenitätsgrad variiert.[106] Die Vergleichbarkeit ist dadurch nicht wesentlich eingeschränkt, dass die beobachteten Korrelationen im Falle der Daten der Glaukompatienten im Mittel etwas größer sind als in der Simulation.[107]

Zunächst ist festzustellen, dass die wahre Reliabilität des Gesamttests (Reliabilität des Messinstruments) bei der größeren Anzahl von Komponenten auf geringerem Niveau ist und in beiden Datensituationen mit zunehmender Heterogenität abnimmt (Tabelle 23 und 24).

[106] Details der Simulation: Für kongenerische Daten wurden die wahren Reliabilitäten auf 0,22 (Komponente 1 und 2), 0,24 (Komponente 3 und 4), 0,26 (Komponente 5 und 6) und 0,28 (Komponente 7 und 8) gesetzt. Für die Komponenten 1–4 wurde die True-Score-Varianz mit 2, für die der Komponenten 5–8 mit 1 festgelegt. Zur Simulation schwacher Heterogenität wurde die True-Score-Korrelation zwischen den Komponenten 1 bis 4 und den Komponenten 5 bis 8 auf 0,8 gesetzt, bei mittelstarker Heterogenität auf 0,6 und bei starker auf 0,2. Da die Reliabilität definiert ist als „True-Score-Varianz geteilt durch Observed-Score-Varianz“, ergibt sich z. B. die Observed-Score-Varianz der Komponente 1, indem man die True-Score-Varianz durch die Reliabilität dividiert (2,0/0,22 = 9,1). Die beobachtete Korrelation zwischen Komponente 1 und 5 wird bspw. berechnet als das Produkt aus der True-Score-Korrelation bei schwacher Heterogenität (0,8) und der Quadratwurzel aus dem Produkt der betreffenden Reliabilitäten: 0,8 $\sqrt{0{,}22 \cdot 0{,}26} = 0{,}19$.

[107] von den (n ▪ n - n) /2 = 28 Korrelationen liegen 17 im Bereich zwischen 0,20 und 0,36, sieben im Bereich zwischen 0,41 und 0,55 und nur vier im Bereich zwischen 0,64 und 0,82. Die Simulation mit acht Komponenten weist beobachtete Korrelationen in der Größenordnung 0,25 auf. Die entsprechende Simulation eines Messinstruments mit nur vier Komponenten und beobachteten Korrelationen zwischen 0,45 und 0,55 zeitigt Resultate, die mit jenen der acht Komponenten Simulation vergleichbar sind – die Reliabilität wird in etwas größerem Ausmaß unterschätzt. Die Variabilität der Korrelationskoeffizienten zwischen den Items zur Glaukomdiagnose (Minimum 0,20; Maximum 0,82) ist in etwa vergleichbar mit dem Simulationsfall starker Heterogenität (Minimum: beobachtete Korrelation zw. Komponente 1 und 5 = 0,2 $\sqrt{0{,}22 \cdot 0{,}26} = 0{,}05$; Maximum: beobachtete Korrelation zw. Komponente 7 und 8 = 1,0 $\sqrt{0{,}28 \cdot 0{,}28} = 0{,}28$).

Unterschätzung der Gesamtreliabilität

Tabelle 23: Heterogenität der Daten, wahre Reliabilität und Coeffizient alpha in four component data (Osburn)

	Degree of Heterogeneity					
	Parallel	τ −equiv.	Congen.	Slight	Moderate	Strong
True reliability	.800	,798	,786	,781	,760	,703
Coefficient alpha	.800	,798	,778 (99 %)	,752 (96%)	,692 (92%)	,547 (78 %)

Tabelle 24: Heterogenität der Daten, wahre Reliabilität und coeffizient alpha in eight component data (Osburn)

	Degree of Heterogeneity					
	Parallel	τ −equiv.	Congen.	Slight	Moderate	Strong
True reliability	,727	,725	,712	,704	,679	,613
Coefficient alpha	,727	,725	,709 (100%)	,692 (98%)	,654 (96%)	,554 (90%)

Es wurde hauptsächlich untersucht, wie die durch Koeffizienten geschätzte Reliabilität sich in beiden Modellen verhält. Koeffizient alpha betrug in dem hier relevanten Acht-Indikatoren-Modell mit starker Heterogenität 90 % der wahren Reliabilität und unterschätzte nur in einem moderaten Ausmaß (bei schwacher und mittlerer Heterogenität lauteten die Ergebnisse 98 bzw. 96 %). Der Vergleich mit dem Vier-Komponenten-Modell, das unter entsprechenden Bedingungen Reliabilitätsunterschätzungen durch alpha im Ausmaß von 78 bis 96 % hervorbrachte, zeigt, dass das Ausmaß der Unterschätzung mit zunehmender Zahl von Komponenten abnimmt[108]. Die Vergleichbarkeit des Simulationsmodells mit dem publizierten Modell mit acht Diagnoseverfahren ist jedoch nur hinsichtlich der Heterogenität der Items bzw. der Dimensionalität der Faktoren zulässig, da eine herausragende Eigenschaft des publizierten Modells – das Vorhandensein einer beträchtlichen Zahl von Fehlerkorrelationen – hier nicht berücksichtigt wird und folgenden Abschnitten (10.3.2 und 10.3.3) vorbehalten bleibt.

Zimmerman et al. (1993) führten eine Simulationsstudie durch, deren Ablauf weiter unten im Abschnitt 10.3.2.4 vertieft behandelt wird, die sich u. a. mit der Verletzung der Annahme der essentiellen τ −Äquivalenz beschäftigt. Für die folgende aus der Veröffentlichung stammende Tabelle 25 gilt, dass die erste Spalte der Stichprobenmittelwerte diese Annahme erfüllt, da Subtest-True-Scores voneinander linear abhängig sind (Subtest-True-Scores unterscheiden sich nur durch eine additive Konstante) und die Spalten 1 bis 5 nicht. In Spalte 1 sind Subtest-True-Scores untereinander multiplikativ verknüpft, in Spalte 2 exponential und in den weiteren Spalten in zunehmendem Ausmaß zufällig.

[108] Osburn (2000), S 343.

Tabelle 25: Wahre Reliabilität und gemittelte Alpha-Koeffizienten bei additiven und nichtadditiven Subtest-True-Scores (Zimmerman)

Mean of Sample Values of Coefficient Alpha as a Function of Reliability for Additive and Nonadditive Subtest True Score Matrices

ADDITIVE		NONADDITIVE				
$\rho_{xx'}$		1	2	3	4	5
.50	.467	.448	−.015	.460	.422	.334
.60	.580	.567	.023	.573	.526	.417
.70	.692	.666	.051	.684	.632	.501
.80	.802	.770	.081	.787	.730	.590
.90	.903	.868	.114	.890	.823	.667

Note. $M = 20$, $N = 8$.

Alpha unterschätzt die Reliabilität in allen nichtadditiven Fällen, wobei das Ausmaß der Unterschätzung beim exponentiellen Verknüpfungsmodus enorm ist und die Ergebnisse des multiplikativen in etwa jenen des gemäßigt zufallsbeeinflussten Modus entsprechen. Die Ergebnisse dieser Studie unterstützen die Resultate von Osburn, insofern die Reliabilität durch alpha mit zunehmender Heterogenität der Items (bzw. Subtests True Scores) zunehmend unterschätzt wird.

Oben wurden mit „Multidimensionalität“ (Itemheterogenität) und „ geringer Anzahl von Indikatoren“ zwei Spezifika von Messinstrumenten genannt, die sich auf das Ausmaß der Unterschätzung der Reliabilität verstärkend auswirken. Wenn weitere Spezifika untersucht werden, dann geschieht dies (entsprechend dem gewöhnlichen experimentellen Vorgehen) zunächst derart, dass beim Vergleich von Modellen nur jeweils eine potentielle Einflussgröße verändert wird, während die anderen konstant bleiben (Ceteris-paribus-Prinzip). Der „klassische“ Aufsatz, der die Verletzung der Annahmen „Essentielle τ −Äquivalenz“ und „Unkorreliertheit der Messfehler“ *separat* untersucht, ist von Zimmermann et al. (1993). Aufsätze über dieses Thema beziehen sich häufig auf diesen Artikel und haben meist Reliabilitätskoeffizienten, aber auch Fit-Indizes zum Untersuchungsgegenstand. In einem weiteren Schritt sind jedoch die simultanen Wirkungen mehrerer Einflussgrößen zu analysieren, da die simultane Verletzung von zwei oder mehr Modellannahmen in der Forschungswirklichkeit häufig anzutreffen ist. Komaroff (1997) demonstriert *gleichzeitig* die Verletzung die Annahme unkorrelierter Fehler und die Verletzung der τ −Äquivalenz.

Hervorzuheben ist aufgrund der Vergleichbarkeit mit dem zu beurteilenden publizierten Modell der Glaukomdiagnose die Arbeit von Yang und Green (2011), die ein sieben Items beinhaltendes Messinstrument hinsichtlich der Messfehlerkorrelation untersuchen. Relevant er-

scheint auch die Studie von Green und Hershberger (2000), deren nur sechs Items umfassendes Modell auf die zeitliche Abfolge der Erhebung unterschiedlicher Items eingeht und im folgenden Abschnitt behandelt wird. Beide werden im folgenden Abschnitt (10.3.2) behandelt werden.

10.3.2 KORRELIERTHEIT DER MESSFEHLER

10.3.2.1 Differenzierung von Fehlerkorrelationen laut Green und Hershberger[109]

Wie oben schon erwähnt, sind Messfehler im Rahmen der CTT nicht gleichzusetzen mit Messfehlern im Rahmen der SEM (oder der Faktorenanalyse i. A.). Green und Hershberger führen aus, dass Messfehler in SEM im Gegensatz zum Theorierahmen der CTT, wo sie immer als „Zufallsfehler" interpretiert werden, Residualgrößen seien, also Itemwerte nach erfolgtem Herauspartialisieren der indirekten und direkten Effekte von allen Variablen, die das Item betreffen[110]. In SEM können Messfehler zufällige oder systematische Fehler sein, wobei Fehlerkorrelationen im Allgemeinen letzterer Kategorie zugerechnet werden, da sie auf (nicht ins Modell aufgenommene) Faktoren zurückgehen. Im Rahmen der CTT sind Fehlerkorrelationen hingegen nicht erlaubt. Die nicht im Modell enthaltenen und daher Fehlerkorrelationen hervorrufenden Faktoren stellen eine Verletzung der essentiellen τ −Äquivalenz dar, so dass Koeffizient alpha die wahre Reliabilität unterschätzen müsste. Eingedenk der Tatsache, dass jedoch das Vorhandensein von (positiven) Fehlerkorrelationen zur Überschätzung der Reliabilität führen kann, ziehen die Autoren daher den Schluss, „neither classical test theory nor SEM offers a framework for interpreting correlated errors as a unreliable variance and for understanding why coefficient alpha might be an overestimate of reliability."

Daher befürworten Green und Hershberger eine Modifizierung der Annahmen der CTT. Ausgangspunkt ist die Tatsache, dass die Items eines Messinstruments in einer zeitlichen Abfolge angewendet werden, so dass es früher und später durchgeführte Testungen der entsprechenden Items gibt. Wenn es nun im Rahmen der CTT gestattet ist, dass die Zufallsfehler früher gemessener Items die Werte später gemessener Items auf indirekte oder direkte Weise beein-

---

[109] Es ist auch eine alternative (nicht widersprechende) Differenzierung der Messfehlerkorrelationen denkbar, die auf eine inhaltliche Ordnung abhebt. Diese schlugen Rubio und Gillespie 1995 vor: Sie gibt die Kategorien „Autocorrelation", „Same Source Correlation" und „Spatial Correlation" vor. Zufallsfehler werden dort u. a. mit dem Hinweis bedacht, dass manche Fehlerkorrelationen nur in Stichproben auftauchen, nicht aber in der Population, aus der Stichproben gezogen wurden.

[110] Green & Hershberger (2000), S. 254.

flussen können, dann leisten diese Fehler einen Beitrag zur Kovarianz zwischen den Messfehlern der Items[111]. Dieser Beitrag ist – im Gegensatz zur o. g. allgemeinen Zurechnung korrelierter Fehler – eine Funktion zufälliger Messfehler, stellt „unreliable Variance“ dar und führt zu einer Erhöhung des Koeffizienten alpha. Korrelierte Messfehler können demnach wie folgt unterschieden werden (Tabelle 26):

Tabelle 26: Typologie korrelierter Fehler – Ursachen und Effekte

	Typ korrelierter Fehler	
	durch Faktoren hervorgerufen	durch zufällige Messfehler hervorgerufen
Fehlerquelle	im Hintergrund vorhandene jedoch unspezifizierte[112] Faktoren	zeitabhängige unspezifizierte[113] Komponenten
Art der Varianz	reliabel	unreliabel
Effekt auf Reliabilitätskoeffizienten	Unterschätzung	Überschätzung (Voraussetzung: pos. Fehlerkovarianzen)

In zwei Beispielen demonstrieren die Autoren, dass die o. g. Modifikation der CTT angemessen ist[114]: Wenn Daten vorliegen, in welchen zeitabhängige Fehlerkorrelationen gegeben sind und Modelle τ −äquivalenter Variabler an diese Daten angepasst werden sollen, dann weisen im Sinne zeitabhängiger True Score Models modifizierte Modelle τ −äquivalenter Variabler einen wesentlich besseren Fit auf. (Ein drittes Modell zeitabhängiger True Scores unter Berücksichtigung eines Memory-Effektes zur Erklärung der Ursache dafür, dass frühere Items spätere beeinflussen, kann im Rahmen dieser Abhandlung nicht erläutert werden.)

---

[111] *Direkter Effekt:* $X_i=T+ E_i+\sum_{j-1}^{i-1}\lambda_{X_iE_j} E_j$, wobei E_j ein Fehlerscore irgendeines früher gemessenen Items und $\lambda_{X_iE_j}$ der direkte Effekt eines früheren Fehlers E_j auf X_i.
Dieses Modell kann vereinfacht werden, indem nur die Effekte des unmittelbar vorhergehenden Items berücksichtigt werden (moving average component of order 1). Hierbei sind alle $\lambda_{X_iE_j} = 0$, wenn i - j Grösser als 1.
Indirekter Effekt: $X_i=T+ E_i+\sum_{j-1}^{i-1}\lambda_{X_iX_j} X_j$. Ein vereinfachtes Modell schließt eine autoregressive Komponente der Ordnung 1 ein und weist Lambdas größer 0 nur zwischen unmittelbar aufeinanderfolgenden Items auf: $X_i=T+ E_i+\lambda_{X_iX_{i-1}} X_{i-1}$.

[112] Unspezifiziert heißt: „nicht ins Modell aufgenommen“.

[113] dito

[114] Auf Basis eines jeden der zwei vorgestellten Modelle, des True-Score-Modells mit einer gleitender Durschschnitt-Komponente (moving average component) und des True-Score-Modells mit einer autoregressiven Komponente der Ordnung 1 werden Parameter festgesetzt und nach diesen Kovarianzmatrizen erzeugt. Versucht man ein τ −äquivalentes Modell jeweils an diese Daten anzupassen, so gelingt es auf Basis dieses Modells nicht, die jeweilige Kovarianzmatrix zu reproduzieren. Erweitert man das Modell und gestattet Fehlerkovarianzen zwischen aufeinanderfolgenden Items, so verbessert sich die Anpassung beträchtlich. Im Falle des ersten Modells wird eine perfekte Anpassung erzielt, im Falle des letzteren eine Steigerung des CFI von 0,78 auf 0,94. Diese nicht perfekte Anpassung ist dem Umstand geschuldet, dass in den nach Maßgabe des Modells mit der autoregressiven Komponente erzeugten Daten auch Fehler von nicht direkt aufeinanderfolgenden Items korreliert sind, während das simplifizierte True-Score-Modell mit einer autoregressiven Komponente der Ordnung 1 für nicht benachbarte Items eine Fehlerkorrelation von Null annimmt. Die Reliabilität des ersten zeitabhängigen True-Score-Modells (moving average-Component), 0,82, wird durch coefficient alpha (0,98) ebenso stark überschätzt wie im Falle des zweiten True-Score-Modells mit der autoregressiven Komponente (0,70 vs. 0,90).

Green und Hershberger resümieren, dass es vonnöten ist, die Quelle der Fehlerkorrelation – unspezifizierte zugrundeliegende Faktoren oder unspezifizierte zeitabhängige Komponenten – zu bestimmen: Ein Vorgehen, das unterschiedslos beide Fälle berücksichtigt, existiert nicht. Es sei, so die Autoren, aber nicht immer möglich, via SEM zu klären, welches Modell das angemessene ist. Die weitverbreitete Annahme, dass alpha eine Untergrenze der Reliabilität sei, falls keine Geschwindigkeitsmessungen[115] vorliegen, sei irrig. Alpha könne bei Vorliegen von durch Zeitabhängigkeit induzierten Fehlerkorrelationen auch eine Obergrenze der Reliabilität sein.

10.3.2.2 Differenzierung von systematischen Messfehlerkorrelationen

Die oben vorgenommene Typologie korrelierter Fehler lässt sich weiter verfeinern durch eine Differenzierung des Typs von Fehlerkorrelationen, der auf im Hintergrund vorhandene – jedoch im Modell außer Acht gelassene – Faktoren zurückzuführen ist. „SEM makes it possible to separate unsystematic measurement error from systematic individual differences that are due to trait and method effects.“[116] „Unsystematischer Messfehler“ ist hierbei mit der Bezeichnung „zufälliger Messfehler“ gleichzusetzen und „systematische individuelle Unterschiede aufgrund von trait- und Methodeneffekten“ korrespondiert mit „Einflüssen von im Hintergrund befindlicher Faktoren“, wie oben bezeichnet.

Podsakoff et al. (2004) gliedern entsprechend systematische Messfehler hinsichtlich ihrer Ursachen in solche auf, die der Methodenvarianz geschuldet sind und solche, die der Varianz aufgrund eines im Modell ignorierten Konstrukts zuzurechnen sind. Letzteres betrifft die Frage der Uni- oder Multidimensionalität des Modells. Eine Definition dessen, was unter Methodenvarianz zu verstehen ist, liefert Fiske 1982:[117] „Method Variance refers to variance, that is attributable to the measurement method rather than to the construct of interest. The term method refers to the form of measurement at different levels of abstraction.“ Im Konkreten ist unter „method“ ein zur Messung gewähltes Skalenniveau oder der Kontext eines Items zu verstehen; in einem abstrakteren Sinne ist Methode z. B. ein Erhebungsverfahren, das soziale Erwünschtheit zur Wirkung bringt und der entsprechende „Methodeneffekt“ ein dadurch verzerrtes Antwortverhalten.

---

[115] Der Ausschluss von Geschwindigkeitsmessungen hebt auf die von Green und Hershberger auf S. 253 erwähnten weiter unten angeführten Erkenntnisse von Rozeboom (1966) ab.
[116] Eid et al. (2006), S. 283.
[117] Zitiert nach Podsakoff et al., S. 879.

Korrelationen systematischer Messfehler im eigentlichen Sinne sind solche, die auf den Messvorgang, die Erhebungsmethode, zurückzuführen sind. Daneben aber können systematisch korrelierte Fehler auch dadurch entstehen, dass eine latente Variable außer Acht gelassen wurde, die Eigenschaften (traits) der Untersuchungseinheit (i. d. R. die Person) erfasst.

10.3.2.2.1 Methodenvarianz – Fehlerkovariation aufgrund von Common Method Bias

Der Schwerpunkt von Podsakoffs Untersuchung ist „Common method bias", der wie folgt erläutert werden kann: Wenn die Beziehung zwischen zwei hypothetischen Konstrukten A und B untersucht werden soll, kann man annehmen, dass die Indikatoren (Measures) des Konstrukts A mit jenen von B korreliert sind. Wenn jedoch die Indikatoren von A der gleichen Methode angehören wie die Indikatoren von B, so ist zu vermuten, dass diese Methoden einen systematischen Effekt auf die beobachtete Korrelation zwischen den Indikatoren ausüben: Somit üben Verzerrungen aufgrund der (Anwendung der) gleichen Methode (Common method bias) zumindest teilweise eine konkurrierende Erklärung für die Entstehung der beobachteten Indikatorenkorrelation aus.[118] [119] Von Forschern wird die von Fiske eingeführte Unterscheidung eines konkreteren und abstrakteren Verständnisses von method effects begrifflich auch gefasst mit der Unterscheidung von Methodeneffekten auf der primären Ebene des Items aber auch sekundär auf der Ebene des Konstrukts.[120] Podsakoff et al. stellten eine Übersicht der möglichen Quellen für common method bias zusammen, von welchen ich nur den jeweiligen Oberbegriff und zur Illustration einen von mehreren Unterpunkten in Klammern nenne, falls die Art des Bias auf meinen Anwendungsbereich (Glaukomdiagnose) nicht zutrifft:

Common rater effects (z. B. Aquiescence Bias (Tendenz zur inhaltsunabhängigen Zustimmung oder Ablehnung)),

---

[118] a. a. O, S. 879.

[119] Versuch der begrifflichen Klarstellung zu *systematischem Messfehler* und *systematischer Messfehlerkorrelation***,** jeweils im Zusammenhang mit der *Erhebungsmethode***,** sowie *method bias* und *common method bias*: Ein systematischer Messfehler, der auf den Einfluss einer Methode zurückzuführen ist, engl. auch als method bias bezeichnet, kann auch einen einzigen Indikator betreffen. Betrifft ein und derselbe systematische Messfehler infolge eines Methodeneffekts zwei Indikatoren, so entsteht eine (methodenbedingte) systematische Messfehlerkorrelation und wird engl. als common method bias bezeichnet. (Ich unterscheide in diesem Klarstellungsversuch nicht zwischen „Vorgang" (Verzerrung) und „Produkt" (Fehler).)

[120] a. a. O, S. 889: „Generally speaking, the techniques used to control for common method variance should reflect the fact, that it is expected to have effects on the item level. However for certain types of biases (e. g. social desireability, negative affectivity), it may make theoretical sense to also model the effects of method variance at the construct level. "

Item characteristic effects (z. B. common scale formats (artifizielle Kovariation durch gleichförmiges Antwortverhalten bei gleichen Skalenformat, beispielsweise nur Likertskala)),

Item Context effects (z. B. Intermixing or grouping of items (Gruppierung von zu verschiedenen Konstrukten gehörenden Items kann die Intra-Konstrukt-Korrelationen verringern und die Inter-Konstrukt-Korrelationen vergrößern)).

Im Rahmen der Datenanalyse zur Glaukomdiagnose sind jedenfalls Methodeneffekte zu diskutieren, die unter den Oberbegriff „measurement context effects" fallen: Zwei oder mehr Messungen wurden zur selben Zeit, am selben Ort oder unter Benutzung desselben Mediums gemessen.[121] Dieser Methodeneffekt bezieht sich auf die Beobachtung, dass Indikatoren verschiedener Konstrukte, die zur selben Zeit am selben Ort oder durch das selbe Medium erhoben wurden, Artefakte von Kovariation produzieren, also Kovariation, die unabhängig von Inhalt des Konstrukts auftritt. Mit Gewissheit trifft auf einige der hier untersuchten Verfahren der mögliche Verzerrungseffekt „identisches Medium" zu, da sie aus ein und derselben Messung rechnerisch *abgeleitet* wurden. Dies betrifft die Verfahren AMP und VEP sowie MD und CLV.[122] Die Korrelation dieser Verfahren innerhalb der Kontrollen (Gesunde) betragen 0,08 und 0,23. (Siehe zur Übersicht und zu Details 10.3.2.3.) Letzterer Korrelationskoeffizient weist darauf hin, dass ein Methodeneffekt vorliegt, der nichts mit der Erkrankung zu tun hat. Für ersteren ergibt sich kein statistisch abgesicherter Beleg, dass eine Korrelation größer Null vorliegt[123]. Es ist davon auszugehen, dass aus Gründen der Praktikabilität (Kosten, Personalaufwand, Effizienz für Patient und Klinik) zumindest die einer bestimmten Diagnoseverfah-

[121] Es ist nicht auszuschließen, dass auch in dem hier gegebenen Anwendungsbereich die Anordnung der Diagnoseverfahren im Zeitablauf die Intra-Konstrukt-Korrelation erhöht haben könnte. Denn es ist anzunehmen, dass Testdurchläufe an ein und derselben Messapparatur aus praktischen Gründen unmittelbar aufeinander folgend durchgeführt worden sind. (Siehe 8.2.2.1).

[122] Es ist davon auszugehen, dass daher ein spezifischer ausschließlich AMP und VEP betreffender Methodeneffekt existiert, da diese beiden Verfahren korrelierte Fehler haben. Ein weiterer Methodenfaktor „Blau auf Gelb evoziertes visuelles Potential" könnte somit eingeführt werden und die Fehlerkorrelation zwischen AMP und VEP ersetzen, die beide aus ein und derselben Messung abgeleitet wurden. Die zu vermutende Assoziation zwischen den ebenfalls aus derselben Messung abgeleiteten MD und CLV wird bislang nur durch den Faktor „Konzentration" modelliert, der als bislang einziger Methodenfaktor bezeichnet werden kann. Er beeinflusst alle sinnesphysiologischen Verfahren außer VEP. Die Korrelationen zwischen VEP und AMP einerseits und VEP und MD andererseits machen die Zugehörigkeit von VEP zur sinnesphysiologischen Verfahrensgruppe jedoch deutlich. Weitere Ausführungen zum Thema „Methodenfaktor" finden sich unten.

[123] Die Hypothese, dass r = 0, kann bei zweiseitiger Testung und Signifikanzniveau 0,05 nicht abgelehnt werden. (95%-KI für r: [-0,04; 0,20]; gegeben: n = 140 gesunde Augenpaare, also 280 abhängige Augen; KI wurde ohne Korrekturfaktor für SE berechnet und stellt daher eine untere Grenze des wahren, breiteren KI dar.)

rensgruppe zugehörigen Verfahren am selben Ort erhoben wurden und vermutlich auch in derselben Untersuchungssitzung, also zur selben Zeit[124].

10.3.2.2.2 Fehlerkovariation aufgrund von Eigenschaften der Person – Beispiele

Neben den zahlreichen Arten von Methodeneffekten, die für systematische Messfehler und Fehler-Kovariation im eigentlichen Sinne verantwortlich sind, gibt es Fehler-Kovariation, deren Ursache in Eigenschaften der Person (oder allgemein der Untersuchungseinheit) zu suchen ist. Fehlerkovariationen dieser Art entstehen, weil eine oder mehrere latente Variable, deren Integration in ein SEM zur Vermeidung von Fehlerkovariation angemessen wäre, nicht in ein SEM aufgenommen wurden.

Ich möchte diesen Sachverhalt am Beispiel der sozialen Erwünschtheit darlegen, die einen Methodeneffekt auf der Ebene des Konstrukts darstellt. Soziale Erwünschtheit wird von Podsakoff folgendermaßen gefasst: „Social desireability refers to the tendency of some people to respond items more as a result to their social acceptability than their true feelings." Die Formulierung der Items oder die Gestaltung der Befragungssituation derart, dass gesellschaftlich erwartetes Antwortverhalten nahegelegt ist, ist eine Frage der Erhebungsmethode. Beispielsweise ist die soziale Erwünschtheit eines Sachverhaltes wirkmächtiger, wenn ein Interviewer als Person im Raum in Erscheinung tritt, als wenn möglicherweise anonymisiert lediglich ein Fragebogen ohne Beisein eines Interviewers auszufüllen ist. Die in der Definition vorgenommene Einschränkung auf „some people" verweist darauf, dass die Tendenz von sozial erwünschtem Antwortverhalten bei einigen Personen besteht und bei anderen nicht. D. h. in Eigenschaften / Dispositionen der jeweiligen Person ist begründet, ob, und wenn ja, in welchem Maß das individuelle Antwortverhalten durch – zweifellos unabhängig von der jeweiligen Person vorhandene – soziale Erwünschtheit verzerrt wird. Das Pendant zur sozialen Erwünschtheit auf der Ebene der Person könnte formuliert werden als die individuelle Konformität. Unter Konformität verstehe ich das Bedürfnis, so zu handeln, dass man einer (stark verbreiteten) gesellschaftlichen Norm entspricht. Ich stelle daher die Hypothese auf, dass sozial erwünschtes Antwortverhalten bedingt auf bestimmte Persönlichkeitsmerkmale auftritt, von welchen eines versuchsweise begrifflich gefasst wird als individueller Grad der Konformität: Je größer die Neigung einer Person ist, sozial angepasst zu handeln, desto grösser kann

[124] Eine empirische Verifizierung der Koinzidenz der Untersuchungskontexte (wann wurde welcher Patient mit welchem Verfahren wo untersucht) ist – wenn überhaupt – nur mit erheblichem Aufwand möglich, der durch den mutmaßlichen Erkenntnisgewinn nicht gerechtfertigt zu sein scheint.

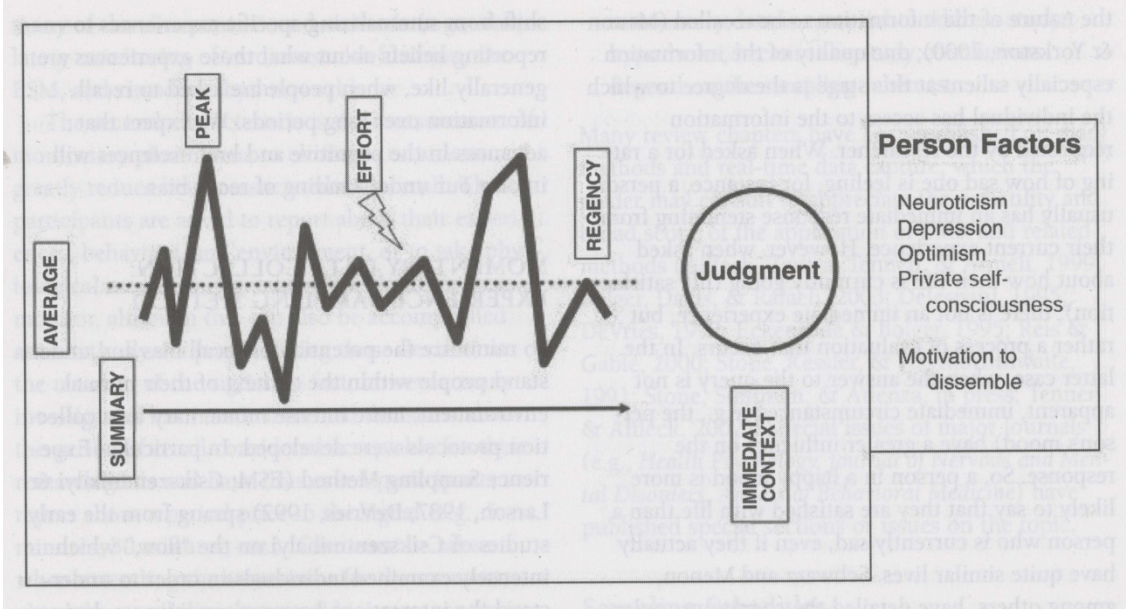
sich ein Methodeneffekt der sozialen Erwünschtheit auswirken. Die individuelle Konformitätsneigung ist die Bedingung der Möglichkeit des Methodeneffekts der sozialen Erwünschtheit.

Somit sollte in ein SEM, das beobachtete Variable beinhaltet, die der Verzerrungsquelle Social desireability unterliegen, zur Vermeidung von systematischer Fehler-Kovarianz, sowohl ein entsprechender Methodenfaktor implementiert werden als auch ein Faktor auf der Ebene der Traits, der die persönliche Konformitätsneigung erfasst.

Ein zweites Beispiel: Die Autoren Stone und Litcher-Kelly beschäftigen sich im Rahmen der Forschung über Momentaufnahmen der mentalen Zustände, Verhaltensweisen und Umgebungsqualitäten von Individuen im Selbstbericht mit den Einflüssen, denen ein Prozess der Beurteilung (Judgement) unterliegt. Judgement ist die Bezeichnung für das Wiedererinnern von Information, ein dynamischer Prozess, der unterschiedliche Gewichtung von Information umfasst und in folgendem Schaubild (Abbildung 81) dargestellt ist.[125] Der Forschungsfokus liegt auf der Entwicklung von EMA (Ecological Momentary Assessment), was auf das Prinzip abhebt, dass alle Messungen in der gewohnten Umgebung stattfinden. Wenn derartige ökologische Messungen in einem SEM als beobachtbare Variable modelliert werden, dann müssen die Person Factors (rechter Bildrand) Berücksichtigung finden. Neben den aufgezählten Traits und Konzepten (von Neurotizismus bis persönliches Selbstbewusstsein) sind noch andere denkbar, deren Außerachtlassung unweigerlich zu Korrelationen zwischen den Fehlern der Indikatoren eines SEMs führen würde.

[125] Stone A. A. & Litcher-Kelly L., (2006), S. 63. Die gestrichelte horizontale Linie („Average“) gibt in dem hier gewählten Beispielsfall – Beurteilung des Schmerzes – dessen durchschnittlichen Level an; „Regency“ meint, dass die Beurteilung des Schmerzes einer aktuellen Periode durch die Erfahrung der kürzlich erlittenen Schmerzen beeinflusst ist; „Effort“ ist eine Abkürzung für Effort-aftermeaning und bezieht sich auf den Prozess, dass die Erfahrung eines frühen Schmerzereignisses durch später stattfindende Ereignisse überlagert ist; „Summary“ bezieht sich auf ein einziges Rating, das der Schmerzleidende aus der Fülle der Information innerhalb der Zeitperiode erstellt (hier bestand 2006 noch Forschungsbedarf). „Immediate Context“ bezieht sich auf den Prozess, in dem kognitive Heuristiken im Moment der Beurteilung angewandt werden, der wiederum von der Art der Information abhängt, die es ins Gedächtnis zu rufen gilt. Wird ein aktuelles Gefühl abgefragt, so ist eine aus der Erfahrung gespeiste unmittelbare Antwort möglich. Bei der Frage nach der Lebenszufriedenheit hingegen ist die eher bilanzierende Antwort in hohem Masse abhängig von dem momentanen Empfinden. „So a person in a happy mood is more likely to say that they are satisfied with life than a person who is currently sad.“

Abbildung 81: Judgement: dynamischer Prozess des Wiedererinnerns von Information (Stone & Litcher-Kelly)



10.3.2.3 Vorzeichen der Fehlerkorrelation

Grundsätzlich ist in der Literatur von „Fehlerkorrelationen“ ohne Angabe des Vorzeichens die Rede, wobei stillschweigend häufig davon ausgegangen wird, dass sie positiv seien. In der Tat dürften die meisten empirisch auftretenden Fehlerkorrelationen mit positivem Vorzeichen versehen sein. Dies ist jedoch nicht immer der Fall, wie auch die obige Einschränkung, alpha sei eine Untergrenze der Reliabilität, falls keine Geschwindigkeitsmessungen vorlägen, belegt. Denn die oben genannten Autoren zitieren Rozeboom (1966) mit den Worten, dass alpha bei „speeded tests“ nicht verwendet werden soll, da diese sowohl positive als auch negative Fehlerkorrelationen erzeugen würden. Das Vorzeichen bestimmt darüber, wie sich die Fehlerkorrelation auf die Reliabilität auswirkt. Wenn positive Fehlerkorrelationen infolge von Fehlspezifikation der True-Score-Varianz zugerechnet werden, so steigt die Reliabilität; wenn hingegen negative Fehlerkorrelationen der True-Score-Varianz zugerechnet werden, so sinkt sie. Wie könnte in diesem Forschungskontext eine negative Fehlerkorrelation entstehen? Diese Frage führt zu einer vorgeordneten Frage: Gibt es sowohl negative Fehlerkorrelationen, die auf Zufallsmessfehler zurückführen sind, als auch negative Fehlerkorrelationen, die aufgrund nicht spezifizierter im Hintergrund befindlicher Faktoren entstehen? Es gibt Fehlerkorrelationen die auf zufällige Variation zurückzuführen sind und solche nicht zufälliger Art.

Zufällige Fehlerkorrelationen sollen an einem empirischen Beispiel erläutert werden: Peter Martus untersuchte Daten von 237 Glaukompatienten oder Personen mit Verdacht auf Glaukomerkrankung und 140 Kontrollen. Die folgende Korrelationsmatrix (Tabelle 27) zeigt in

der unteren Hälfte die Korrelationskoeffizienten der acht Diagnoseverfahren der 140 Kontrollen[126]. Zehn der 28 Korrelationskoeffizienten tragen ein negatives Vorzeichen. Da die Korrelation zwischen jeweils zwei Diagnoseverfahren nicht durch die Korrelation der True-Score-Variablen „Glaukomerkrankung" zustande gekommen sein kann, muss es sich bei den Korrelationen um solche handeln, deren Herkunft außerhalb der Erkrankung liegt. Für den Teil der Korrelationen, deren Betrag signifikant von Null abweicht (für sechs von 28 gilt |r |> 0,14) erscheint zufällige Variation als Erklärung unplausibel[127]. Diese schwachen Korrelationen treten mit Ausnahme von VEP und AMP[128] auch nur unter den psychophysischen Verfahren auf und sind vermutlich auf einen Einfluss zurückzuführen, der nur in diesen Verfahren wirksam wird, mutmaßlich der im Ein-Faktor-Modell nicht erfassten, latenten Variable Konzentration (Martus (2001a), S. 267). Die vier größten dieser sechs schwachen Korrelationen betreffen ausschließlich psychophysische Verfahren. Die geringste ist negativ: r (AMP, FLI) = -0,15.

Korrelationen, deren Betrag nicht signifikant von Null abweicht, die weder durch die Erkrankung noch durch einen Methodenfaktor erklärt werden können und die auch auf keine weitere bekannte latente Variable zurückgeführt werden können sind vermutlich der *zufälligen Variation* („Rauschen") geschuldet. Wenn innerhalb der Kontrollgruppe zwischen den Diagnoseverfahren Korrelationskoeffizienten von Null erwartet werden, da sie der zufälligen Variation geschuldet sind, ist auch zu erwarten, dass Abweichungen gleichermaßen in den negativen als auch positiven Bereich gehen.

---

[126] Martus (2001a), S. 267.

[127] Der z-transformierte Korrelationskoeffizient r lautet $z = 0{,}5 \cdot \ln\left(\frac{1+r}{1-r}\right)$. Um jenes r zu ermitteln, dessen untere Konfidenzintervallgrenze 0 ist, wird folgende Gleichung für die Untergrenze $z_1 = 0$ nach z aufgelöst: $0 = \mathrm{z} - 1{,}96 \cdot \text{Korrekturfaktor} \cdot 1/\sqrt{\mathrm{n} - 3}$. Hierbei gleicht der Korrekturfaktor die Unterschätzung des SE durch die nicht nach Paarigkeit adjustierte Analyse aus und wird arbeitshypothetisch mit 1,20 angesetzt, n = $2 \cdot 140 = 280$Augen und $1/\sqrt{\mathrm{n} - 3}$ gibt approximativ den SE von z wieder. $0 = \mathrm{z} - 1{,}96 \cdot 1{,}2 \cdot 0{,}0601$ $\leftrightarrow z = 0{,}14$. (Retransformation von z_1 ergibt die untere Konfidenzintervallgrenze von r: $\frac{e^{2 \cdot z_1} - 1}{e^{2 \cdot z_1} + 1} = \frac{1-1}{1+1} = 0$.) Zur Rechtfertigung der Größe des Korrekturfaktors: In 8.2.5 wurden Korrekturfaktoren für den SE von 5 modellierten der Fehlerkorrelationen vorgestellt, die zwischen 1,15 und 1,33 liegen (Tabelle 18). Als konservative Herangehensweise schlage ich daher einen Korrekturfaktor von 1,2 vor.

[128] Siehe 10.3.2.2.1, Stichwort „measurement context effects".

Tabelle 27: Korrelationskoeffizienten der acht Diagnoseverfahren getrennt nach Kontroll- und Patientengruppe

Correlation analysis within the control and patient groups.

	NRR	ERG	VEP	AMP	FLI	CLV	MD	DYN
NRR	---	0.34	0.31	0.20	0.35	0.50	0.51	0.51
ERG	-0.08	---	0.31	0.21	0.36	0.39	0.45	0.37
VEP	-0.04	0.10	---	0.32	0.29	0.36	0.41	0.31
AMP	0.03	-0.01	0.08	---	0.33	0.32	0.37	0.32
FLI	-0.01	0.06	0.03	-0.15	---	0.48	0.64	0.55
CLV	-0.05	0.03	-.05	-0.02	0.19	---	0.82	0.69
MD	-0.04	0.07	0.17	0.09	0.25	0.23	---	0.74
DYN	0.08	0.11	0.03	-0.01	0.12	0.12	0.30	---

Upper right entry: correlation in the patient group, lower left entry: correlation in the control group; NRR neuroretinal rim area, ERG electroretinogram, VEP peak latency of the visual evoked potential, AMP amplitude of the visual evoked potential, FLI flicker test, CLV perimetric corrected loss variance, MD perimetric mean defect, DYN spatiotemporal contrast sensitivity.

Negative Fehlerkorrelationen infolge nicht spezifizierter Faktoren

Eine nicht spezifizierte Einflussgröße, die den Messwert eines Indikators (Items) vergrößert und den Messwert eines anderen Indikators abschwächt, würde eine nichtzufällige negative Fehlerkorrelation hervorrufen. Es wird angenommen, dass bei der Erhebung eines bestimmten Merkmals Störgeräusche die Leistung des Probanden mindert. Denkbar ist z. B. die Konstellation einer überdurchschnittlich starken Geräuschbeeinträchtigung bei der Erhebung eines Merkmals, die zu einem bestimmten Zeitpunkt oder an einem bestimmten Ort stattfindet und eine fehlende bzw. unterdurchschnittlich starke Geräuschbeeinträchtigung bei der Erhebung eines anderen Merkmals, die zu einem anderen Zeitpunkt oder an einem anderen Ort angesetzt ist. D. h. die Probanden werden in der einen Messung im Schnitt schlechtere Werte erzielen und in der anderen Messung bessere Werte. Dadurch wird die Korrelation zwischen diesen beiden Variablen abgeschwächt. (Die positive True-Score-Korrelation und die als schwächer angenommene negative Error-Score-Korrelation führen zu einer reduzierten positiven Observed-Score-Korrelation.) Wenn man jedoch die Präsenz des Störfaktors Geräusch-

pegel als latente Variable einbezieht (d. h. nach dieser Störvariablen kontrolliert), wird die Stärke der Beziehung zwischen den beiden Variablen ansteigen. Diese Vorstellung entspricht der Vorgehensweise, mit der im Simulationsteil dieser Arbeit negative Fehler programmiert wurden[129]. In der Literatur wird der mögliche Umstand, dass Fehlerkorrelationen auch negativ sein können, in Simulationsbedingungen berücksichtigt. Heene et al. (2012) liefern dafür ein Beispiel: In einem SEM, in dem zwei miteinander korrelierte Faktoren jeweils zwölf Items beeinflussen, wird von vier Szenarien ausgegangen, wobei zwei der vier Szenarien auch negative Fehlerkorrelationen vorsehen.[130]

10.3.2.4 Nicht berücksichtigte Fehlerkorrelationen

Eine grundsätzliche Unterscheidung zwischen folgenden zwei Fällen ist sinnvoll, um in einer ersten Annäherung aufzuzeigen, in welchem Verhältnis Reliabilitätsschätzer und zu schätzende wahre Reliabilität stehen können:

1) Fehlerkorrelationen sind zwar vorhanden, aber werden im Modell nicht berücksichtigt.

2) Fehlerkorrelationen sind vorhanden und werden im Modell berücksichtigt.

Der erste Fall wird sogleich, der zweite im Anschluss behandelt werden

Vorhandene, im Modell nicht berücksichtigte Fehlerkorrelationen lassen sich anhand der häufig zitierten Studie von Zimmerman et al. (1993) illustrieren. Ich werde auf diese etwas näher auch deswegen eingehen, weil die Programmierung der von mir durchgeführten Simulationen zum selben Ergebnis kommt wie die im Folgenden von Zimmerman dargestellte Programmlogik. Die künstlich erzeugten Variablen unterscheiden sich bei Zimmerman et al. und in dieser Arbeit darin, dass erstere Normalverteilungen und auch andere Verteilungen mit Hilfe von

---

[129] Es geht um die zwei folgenden beobachteten Variablen ind_1_ps und ind_2_ns, deren True Scores positiv und deren Error Scores negativ korreliert sind:
ink_1_ps = x4 + (x1 + x0) / SQRT (2)
ink_2_ns = x4 + (x2 - x0) /SQRT (2)
x4 steht für die latente gemeinsame (True Score) Variable.
x1 ist der Zufallsfehler von Var 1.
x2 ist der Zufallsfehler von Var 2.
x0 ist der gemeinsame Störterm (Geräuschpegel).
SQRT (2) dient der Standardisierung.

[130] (1) Die ersten drei der dem ersten Faktor zugehörigen Items sind mit den ersten drei der dem zweiten Faktor zugehörigen Items positiv korreliert.
(2) Die ersten sechs der dem ersten Faktor zugehörigen Items sind mit den ersten sechs der dem zweiten Faktor zugehörigen Items positiv korreliert.
(3) Fall 1 wird dahingehend abgewandelt, dass zwei der drei Korrelationen als negativ angenommen werden.
(4) Fall 2 wird dahingehend abgewandelt, dass drei der sechs Korrelationen als negativ angenommen werden.

Pseudozufallszahlen erzeugten und eine jeweils vorgegebene Reliabilität des Gesamtscore durch ein geeignetes Verhältnis zwischen True-Score- und Error-Score-Varianz produzierten, während ich nur approx. Normalverteilungen generierte und die Reliabilität nicht steuerte. Weiterhin darin, dass in meinen Ein-Faktor-Modellen die Additivitätsbedingung immer zutraf (s. u.). Unterschiede hinsichtlich der Anzahl der Durchläufe und der Anzahl der Personen sind eher sekundär und hierbei nicht berücksichtigt.

Zimmerman simulierte auf Basis der CTT für mehrere Personen Subtest Scores, die zusammengenommen einen Gesamtscore ergeben. Diese Subtest-Scores bestehen einerseits aus wahren Werten, die über die Messwiederholungen konstant sind, aber sich von Person zu Person unterscheiden und andererseits aus Fehlerwerten, die sowohl über die Wiederholungen als auch die Personen variabel sind. Die Variabilität der Error-Scores wurde nach Maßgabe der Variabilität der True-Scores so gesteuert, dass sich eine vordefinierte Reliabilität nach CTT ergab. Error- und True-Scores wurden addiert und ergaben die Observed-Scores. Für eine bestimmte Person wurde ein True Score, eine Konstante, und ein über die Messwiederholungen normalverteilter Error-Score addiert (propensity distribution)[131]. Fehlerkorrelationen wurden erzeugt, indem ein Vielfaches einer normalverteilten Zufallsvariable zu zwei Error-Scores addiert wurde, so dass über den Multiplikator die Höhe der Korrelation gesteuert werden konnte. Zwei Simulationsbedingungen wurden eingeführt: Die erzeugten True-Scores entsprachen der Additivitätsbedingung der Subtest-True-Score-Matrix 1, die die Linearitätsbedingung der (essentiellen) τ −Äquivalenz sicherstellt, oder sie entsprachen ihr nicht.[132] Matrix 1 ist die erste von vier Matrizen in folgender Tabelle 28 und additiv: Aufeinanderfolgende Reihen (i) oder Spalten (k) entstehen dadurch, dass 1 zu jedem Eintrag in der vorhergehenden Reihe oder Spalte addiert wird: $T(i,k) = \mathrm{i} + \mathrm{k}$. Die True-Scores unterscheiden sich demnach nur durch eine additive Konstante.

[131] Lord und Novick (1968), zitiert nach Zimmerman et al. (1993), S. 35.

[132] Es wurden drei weitere nonadditive Matrizen erzeugt, eine multiplikative, eine exponentiale und eine random-Matrix, für die die Beziehung $T(i,k) = \mathrm{i}\cdot\mathrm{k}$, $T(i,k) = \mathrm{k}^{i}$ bzw. $T(i,k) = \mathrm{i}+\mathrm{k}$ + random number galten. Novick and Lewis (1967) formulierten die Additivitätsbedingung und führten zusammen mit Lord and Novick (1968) die Terminologie der (essentiellen)) τ −Äquivalenz ein; zitiert a. a. O.

Tabelle 28: Beispiele für additive und nichtadditive Subtest-True-Score-Matrizen (Zimmerman)

Examples of Additive and Nonadditive Subtest True Score Matrices

	ADDITIVE					NONADDITIVE (MULTIPLICATIVE)			
i^k	1	2	3	4	i^k	1	2	3	4
1	2	3	4	5	1	1	2	3	4
2	3	4	5	6	2	2	4	6	8
3	4	5	6	7	3	3	6	9	12
4	5	6	7	8	4	4	8	12	16
NONADDITIVE (EXPONENTIAL)						NONADDITIVE (RANDOM)			
i^k	1	2	3	4	i^k	1	2	3	4
1	1	2	3	4	1	4	6	4	7
2	1	4	9	16	2	8	11	5	6
3	1	8	27	64	3	11	9	11	12
4	1	16	81	256	4	10	12	9	15

Die dichteste Information über die Effekte korrelierter Fehler lässt sich folgender Tabelle 29 entnehmen:[133]

Tabelle 29: Alpha in Abhängigkeit von wahrer Reliabilität, Korrelation zwischnen Subtest-Error-Scores und Zahl der paarweise korrelierten Subtest-Error-Scores (Zimmerman)

Mean of Sample Values of Coefficient Alpha as a Function of Correlation Between Subtest Error Scores and Number of Subtest Error Scores That are Pairwise Correlated

	ALPHA					
	$\rho_{xx'} = .80$			$\rho_{xx'} = .50$		
	$\rho(E_i, E_j)$					
	.10	.30	.50	.10	.30	.50
C						
2	.807	.815	.826	.485	.520	.541
3	.817	.842	.862	.521	.582	.630
4	.828	.871	.898	.551	.662	.729
5	.546	.899	.927	.594	.730	.806
6	.864	.921	.949	.636	.792	.866
7	.880	.940	.965	.677	.840	.909
8	.894	.955	.978	.721	.881	.940

Note. $M = 20$, $N = 8$, Replications = 2000. Also, $i = 1, \ldots, C; j = 1, \ldots, C, i \neq j, 2 \leq C \leq N$.

Die vorgegebenen Reliabilitäten betragen 0,50 und 0,80. Für jede dieser Reliabilitäten wurden drei Fehlerkorrelationen in zunehmendem Ausmaß simuliert. In der ersten Spalte ist die Zahl der paarweise korrelierten Subtest-Fehlerscores angetragen. Der Reliabilitätsschätzer alpha überschätzt die jeweilige wahre Reliabilität in nahezu allen Fällen[134], wobei zunehmende Feh-

[133] Zimmerman (1993), S. 43.

[134] Als einzige, nicht ohne weiteres erklärbare Ausnahme ist die Kombination „Reliabilität 0,8, C = 5 und Fehlerkorrelation 0,1“ anzuführen (0,546) und als einzige durch die Zufallsstreuung erklärbare die Kombination „Reliabilität 0,5, C = 2 und Fehlerkorrelation 0,1“ (0,485).

lerkorrelation und zunehmende Anzahl der paarweise korrelierten Fehler jeweils für sich genommen die Reliabilitätsschätzung erhöhen und in Kombination diese Wirkung verstärken. Da die wahre Korrelation 0,8 am oberen Ende der Skala des Korrelationskoeffizienten angesiedelt ist, ist eine denkbare Überschätzung nur um 0,2 Korrelationspunkte auf das theoretische Maximum 1,0 vorstellbar. Bei der Simulationsbedingung „wahre Korrelation = 0,5“ ist eine Überschätzung um bis zu 0,5 denkbar. Der wesentlich größere Möglichkeitsraum in der letzten Bedingung macht es plausibel, dass der Überschätzungseffekt enorm sein kann.[135]

Ein großer Vorteil der Arbeit mit SEM besteht darin, dass die Fehlerkorrelationen modelliert werden können und das Ausmaß der Fehlerkorrelationen geschätzt werden kann.[136] Im Rahmen meiner Simulationen auf Basis künstlich erzeugter Daten und der Untersuchung dieser Daten mittels SEM wurde standardmäßig ein Fall berücksichtigt, in dem tatsächlich vorhandene Fehlerkorrelation angemessen modelliert wurden und ein Fall, in dem diese Modellierung unterblieb. Der Effekt der Nichtberücksichtigung von Fehlerkorrelationen auf wesentliche Modelparameter konnte somit analysiert werden.

10.3.2.5 Berücksichtigte Fehlerkorrelationen

10.3.2.5.1 Fehlerkorrelationen infolge von Common Method Variance

In der Literatur zu SEM werden mehrere Möglichkeiten diskutiert, common method variance zu modellieren. Eine der Möglichkeiten besteht darin, die Annahme der Unkorreliertheit der Fehler für bestimmte Paare von Items aufzugeben und Fehlerkorrelationen ins Modell zu implementieren. Hierbei lässt man sich weniger von theoretischen Überlegungen leiten als von Modifikationsindizes (MIs), die besagen, dass die Annahme der Unkorreliertheit für bestimmte Paare von Fehlerkorrelationen nicht haltbar sei. Eine andere Möglichkeit besteht darin, neben einem oder mehreren Faktoren, die die eigentlich interessierende zu messende Dimension (Theoretisches Konstrukt) darstellt, einen oder mehrere sogenannte Methodenfaktoren zu bilden, die die gemeinsame Methodenvarianz messen sollen. Dies ist wiederum auf prinzipiell dreierlei Arten denkbar. 1. Die Verwendung *eines allgemeinen direkt gemessenen* Methodenfaktors in einem SEM wurde u. a. von Williams 1996 demonstriert. Der Methodenfaktor wird

---

[135] Sieben oder acht fehlerkorrelierte Subtest-Score-Paare und Fehlerkorrelation >= 0,30 (Situation 1) oder Fehlerkorrelation = 0,50 und mindestens fünf fehlerkorrelierte Subtest-Score-Paare (Situation 2).

[136] Komaroff benennt im Rahmen seines Arbeitsprogramms zum Artikel von 1997 (S. 339f), dass trotz der Tatsache der theoretischen Untersuchung von Fehlerkorrelationen schon ab 1953 deren Bedeutung ziemlich ignoriert wurde. Dies leite sich aus der (falschen, aber) vorherrschenden Meinung ab, dass Fehlerkorrelationen zu messen schwer, wenn nicht sogar unmöglich sei.

typischerweise durch einen Surrogatparameter wie soziale Erwünschtheit oder positive / negative Affektivität geschätzt, von dem man annimmt, dass er allgemeine Methodenvarianz repräsentiert. Die Gesamtvarianz der beobachtbaren Variablen kann dann auf Einflüsse des theoretischen Konstrukts, der Methode und des Zufalls zurückgeführt werden. 2. Daneben besteht die Möglichkeit*, einen allgemeinen nicht direkt* gemessenen Methodenfaktor ins Modell zu implementieren, wie z. B. durchgeführt von Podsakoff et al. (1990), wobei eine Operationalisierung durch Surrogatparameter entfällt. Schließlich können 3. *mehrere (nicht direkt gemessene)* Methodenfaktoren eingeführt werden, um die Methodenvarianz zu erfassen.

Der allgemeinste und bekannteste Versuch wurde von Campel und Fiske 1959 mit dem Multitrait-Multimethod (MTMM-)Modell unternommen. Die Gesamtvarianz kann in diesem auf drei Varianzquellen – Trait[137], Methode und Zufallsfehler – zurückgeführt werden, so dass interessierende (das Strukturmodell betreffende) Parameter unter gleichzeitiger Kontrolle von Methoden- und Zufallsfehlervarianz schätzbar sind. Zwei oder mehr Methoden – beispielsweise der Fremdbericht oder der Selbstbericht in der Messung von bestimmten Emotionen (Traits) – können vorteilhafterweise gleichzeitig gemessen werden. Im Rahmen von MTMM-Modellen ist es möglich, bestimmte Pfadkoeffizienten, die den Einfluss einer Methode auf eine bestimmte beobachtete Variable anzeigen, zu restringieren, also z. B. auf Null zu setzen und dadurch den Effekt einer anderen in den Focus genommenen Methode quasi isoliert zu betrachten. Ein weiterer Vorteil besteht darin, dass Methodenfaktoren nicht direkt gemessen werden müssen (s. o.). Spector und Brannick zählen 1995 auch Probleme auf, die mit Nichtidentifizierbarkeit, Spezifikations- und Stichprobenfehlern zu tun haben[138]. Zwei weitere Nachteile bestehen darin, dass Interaktionen zwischen Traits und Methodenfaktoren nicht vorgesehen sind und verschiedene Methoden identifiziert worden sein müssen, um überhaupt festlegen zu können, welche beobachtbaren Variablen von welchen Methoden beeinflusst werden.

Anhand von vier verschiedenen Varianten eines Modells mit sieben Indikatoren (Yang & Green, 2011; Abbildung 82) soll im Folgenden demonstriert werden, wie die Einführung eines Methodenfaktors Einfluss nimmt a) auf die Anpassungsgüte des Modells, als dem zentra-

[137] Trait = Eigenschaft, die i. d. Regel durch eine primär interessierende latente Variable gemessen wird und beispielsweise eine Emotion wie Ärger oder Wut sein kann.
[138] Die Darstellung der Vor- und Nachteile geschieht auf Basis von Podsakoff et al. (2003). Es ist im Zusammenhang mit Identifikationsproblemen auf die Verbesserung hinzuweisen, die Eid et al. durch ihr CTC(M-1) erreicht haben; siehe Eid et al. (2006), S. 237.

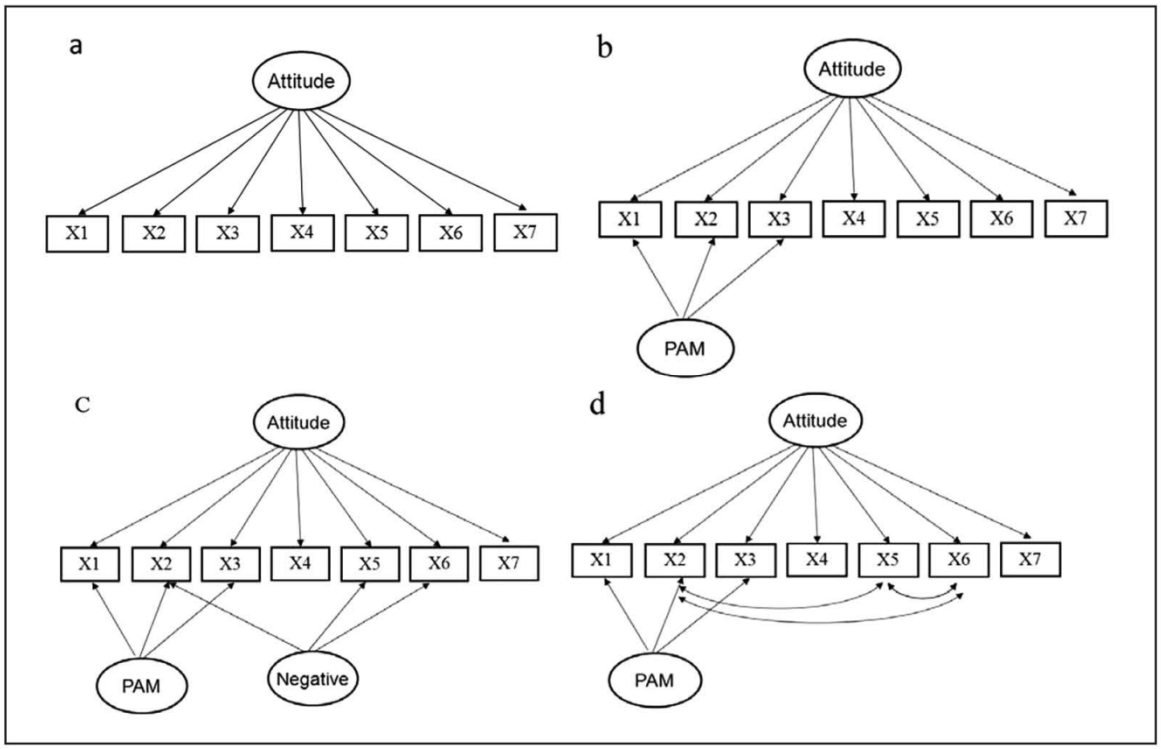
len Schritt in der Modellentwicklung und b) auf die Reliabilität. Die Daten stammen aus einer Studie über die Einstellung zum Mathematiklernen in der Schule (Tabelle 30):

Tabelle 30: Korrelationsmatrix, Mittelwert und Standardabweichung der Items einer Studie über die Einstellung zum Mathematiklernen (Yang & Green)

	Items						
	1	2	3	4	5	6	7
1. I enjoy learning mathematics[a]		.679	.860	.421	.297	.471	.410
2. Mathematics is boring	.559		.645	.317	.323	.485	.212
3. I like mathematics[a]	.778	.520		.482	.330	.422	.426
4. I usually do well in mathematics[a]	.351	.238	.401		.516	.684	.565
5. Mathematics is harder for me than for many of my classmates	.254	.272	.275	.409		.662	.403
6. I'm just not good at mathematics	.376	.390	.330	.538	.547		.455
7. I learn things quickly in mathematics[a]	.358	.171	.375	.464	.324	.351	
M	3.119	3.130	3.106	3.362	2.911	3.369	3.034
SD	0.995	1.081	1.059	0.749	1.113	0.926	0.921

a. Items that were positively worded but reverse scaled.

Abbildung 82: Implementation von Methodenfaktoren in ein Modell über die Einstellung zum Mathematiklernen (Yang & Green)



Das erste Modell (Abb. a) beinhaltet als einzige latente Variable nur das theoretische Konstrukt und weist eine schlechte Anpassung auf ($\chi^2(14) = 201{,}18; p < 0.01$; CFI = 0,76; RMSEA = 0,21; SRMR = 0,11; Das zweite Modell (Abb. b) enthält neben dem interessieren-

den Konstrukt noch eine von Yang et al. als Gruppenfaktor bezeichnete und „PAM“ genannte latente Variable, die sich auf die Items 1–3 [139] bezieht und positive Affekte bezüglich des Mathematiklernens (**P**ositiv **A**ffect towards **M**athematics) repräsentiert. Da PAM somit lediglich einen Aspekt der Variablen Attitude wiedergibt, quasi eine Subdimension von Attitude und keine von Attitude verschiedene Dimension, kann man das Modell 2 als *essentiell unidimensional* bezeichnen. Die anderen Items beinhalten dagegen Art und Weise des Umgangs mit dem Lernfach oder eine Bezeichnung des Lernerfolgs und könnten daher unter die Subdimension „Erfassung der Lernpraxis“ subsumiert werden. Die Faktorladungen auf das theoretische Konstrukt „Attitude“ reichen von 0,45 bis 0,71 und die Ladungen auf den Gruppenfaktor betragen 0,77, 0,49 und 0,72. Die Anpassung des Modells ist einigermaßen zufriedenstellend ($\chi^2(14) = 48{,}07; p < 0.01$ [140]; CFI = 0,95; RMSEA = 0,11; SRMR = 0,04) und der Anteil der Varianz des allgemeinen Faktors „Attitude“ an der Varianz aller latenten Variablen, die Konsistenz[141], beträgt 0,80, wohingegen die Reliabilität mit 0,86 angegeben wird.[142] Modifikationsindices legen nahe, einen weiteren Gruppenfaktor „Negative“ aufzunehmen, der dem Umstand Rechnung trägt, dass die Items 2, 5 und 6 negativ formulierte Aussagen enthalten.[143] Dieses dritte Modell (Modell c) weist einen guten Fit auf ($\chi^2(8) = 14{,}83; p = .063$; CFI = 0,991; RMSEA = 0,05; SRMR = 0,02), die Konsistenz beträgt 0,71 und die Reliabilität 0,88. Das vierte Modell (Modell d) unterscheidet sich von Modell 3 dadurch, dass der zweite Gruppenfaktor Negative durch drei Fehlerkorrelationen ersetzt wird und daher einen – verglichen mit Modell 3 – erheblich verringerten Wert für die Reliabilität aufweist (0,81). Begründung: Die korrelierten Fehler werden in Modell 4 der Error-Score-Varianz zugerechnet, wohingegen sie in Modell 3 durch den Gruppenfaktor „Negative“ erfasst werden, was die True-Score-Varianz des Modells 3 erhöht. Die Reliabilität ist in Modell 3 deshalb von allen drei gefitteten Modellen am höchsten, da in ihm einzig drei Faktoren einen Teil der Gesamtvarianz erklären und die Bedeutung der Items umfänglicher erklärbar machen. Die Autoren konstatieren zwar, dass der Fit von Mod. 3 und 4 gleich sei, betonen jedoch, dass sie der Modellspezifikation mit einem zweiten Gruppenfaktor („Negative“, Mod. 3) den Vorzug geben, da letzterer auf Basis

---

[139] Die restlichen Items repräsentieren den Aspekt Selbstvertrauen in die Lernfähigkeit (self confidence in learning mathematics, SCM).
[140] Die Fitdiagnose der Autoren ist in Frage zu stellen, da bei $p < 0{,}05$ das Modell eigentlich verworfen werden muss.
[141] Konsistenzkoeffizient oder konvergente Validität des Messinstruments
[142] Dieser Prozentsatz ist nicht gleichzusetzen mit der Reliabilität, die als Anteil der True-Score-Varianz an der Gesamtvarianz verstanden wird.
[143] Verglichen mit anderen Fächern wird dieses mit negativer Emotion assoziiert; verhältnismäßig große Schwierigkeiten mit dem Fach und schlechte Leistung im Fach.

der Inhalte seiner Items interpretierbar sei und fahren fort: „In other words we may view it [gemeint ist „Negative"] as an irrelevant source of variance and detracting from the scales validity but not due to measurement error". Der entsprechende Varianzanteil ist demnach ebenso wie der dem ersten Gruppenfaktor PAM zugehörige zwar True-Score-Variance (da es sich nicht um Error-Score-Variance handelt), aber der Konstruktvalidität abträglich, da mit dem 7-Item-Messinstrument das Konstrukt „Haltung zum schulischen Lernfach Mathematik" gemessen werden soll und nicht auch die aufgrund positiver oder negativer Formulierungsvarianten entstehende Varianz (Tabelle 31).

Tabelle 31: Modellvarianten der Einstellung zum Mathematiklernen - konvergente Validität und Reliabilität

	Faktoren und Fehlerkorrelationen				Konvergente Validität und Reliabilität	
	Attitude	„PAM" (Item 1, 2, 3)	„Negative" (Item 2, 5, 6)	Error Correl. (Item 2, 5, 6)	Konsistenz:Var (attitude) / Var (True Score)	Reliability
Modell 2 (b)	X	X			,80	,86
Modell 3 (c)	X	X	X		,71	,88
Modell 4 (d)	X	X		X	,77*[144]	,81

Forts. Tabelle 31

	Fit und Fintindizes				
	χ^2	p	CFI	RMSEA	SRMR
Modell 2 (b)	$48{,}07(df = 14)$	$< 0{,}01$	,95	,11	,04
Modell 3 (c)	$14{,}83(df = 8)$ [145]	$0{,}063$	,991	,05	,02
Modell 4 (d)	$14{,}83(df = 8)$	$0{,}063$	,991*	,05*	,02*

*eigene Ergänzungen

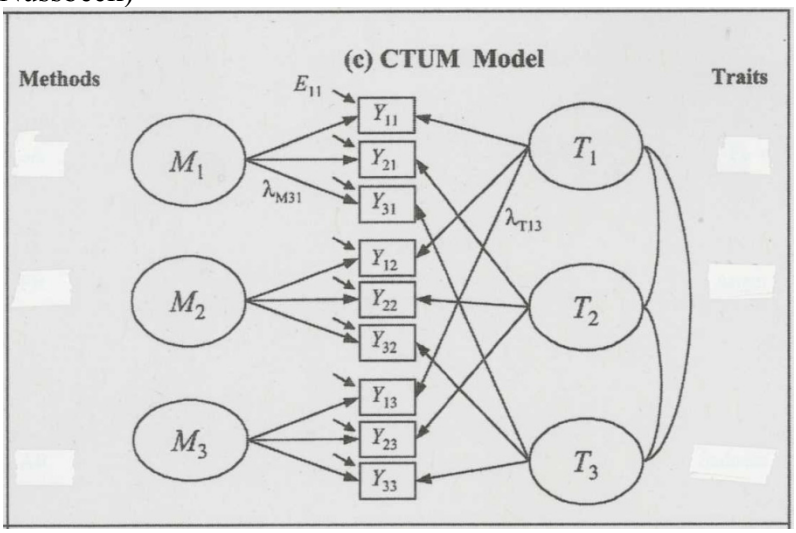
Vorhandenen Fehlerkorrelationen in den Daten dadurch gerecht zu werden, dass statt der Spezifizierung korrelierter Fehler im Modell Methoden- oder Gruppenfaktoren ins Modell aufgenommen werden, wurde als notwendiges Vorgehen betrachtet, um ein Verringern der wahren Reliabilität und eine Überschätzung derselben durch den Koeffizienten alpha zu verhindern. Im sog. Correlated Uniqueness (CU-)Modell[146], das keine Methodenfaktoren, sondern Fehlerkorrelationen aufweist, wurde die Schätzung des Anteils der Methodenvarianz historisch als prinzipiell unmöglich betrachtet.

Conway legte jedoch 1998 dar, dass eben dies doch möglich sei, indem man den Mittelwert der Fehlerkorrelationen im CU-Modell bildet. Um dies aufzuzeigen wurden künstlichen Daten

[144] selbst nachgerechnet anhand 11.7.1.
[145] Nachberechnung mit PROC CALIS und Originalprogrammcode: χ^2= 14,78; p= 0,064; Fitindices identisch.
[146] In der Nomenklatur von Eid et al. (2006) wird dies auch als CTCU (Correlated trait/Correlated Uniqueness)-Modell bezeichnet.

erzeugt, die einer MTMM-Matrix entsprechen. An diese Daten wurde ein CFA-UM-Model[147] als auch ein CU-Modell angepasst.[148] Das CFA-UM-Modell ist ein konfirmatorisches Faktormodell, in dem jede beobachtete Variable sowohl auf den Trait-Faktor (Trait), den Methodenfaktor als auch die latente Fehlervariable lädt und in dem Traits untereinander korreliert sind. Im Gegensatz zum gewöhnlichen CFA-Modell sind die Methoden jedoch untereinander *nicht* korreliert (**U**ncorrelated **M**ethods, Abbildung 83).

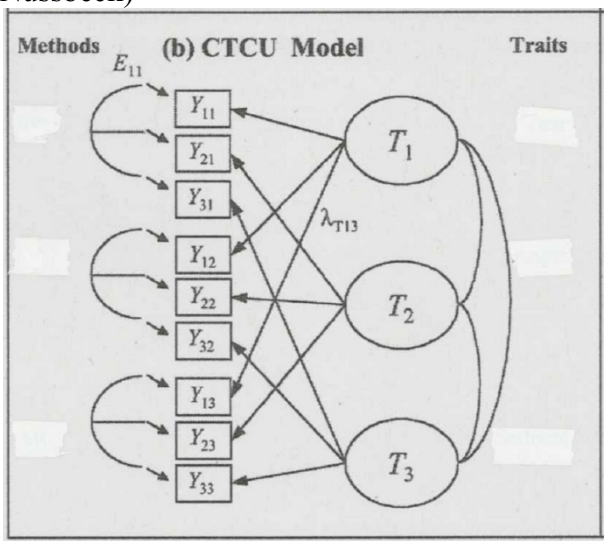
Abbildung 83: Correlated trait/uncorrelated method model (CTUM; Eid, Lischetzke & Nussbeck)



Um den Anteil der Trait-Varianz und der Methodenvarianz einer bestimmten beobachteten Variablen zu schätzen, werden die Faktorladungen der beiden standardisierten Pfadkoeffizienten quadriert. Das CU-Modell beinhaltet Traits, latente Fehlerterme und Methodeneffekte. An Stelle der entsprechenden Methodenfaktoren sind Fehlerkorrelationen spezifiziert, wobei die Fehlerterme der Variablen, die durch dieselbe Methode beeinflusst wurden, miteinander korrelieren dürfen (Abbildung 84).

[147] a. a. O.: wird auch als CTUM (Correlated trait/Uncorrelated Method)-Modell geführt.
[148] Conway (1998), S. 214.

Abbildung 84: Correlated trait/correlated uniqueness model (CTCU; Eid, Lischetzke & Nussbeck)



Laut Marsh, 1989, sind die beiden Modelle bei einer Anzahl von drei Traits im Wesentlichen unterschiedliche Parametrisierungen ein und desselben Modells[149]. Obwohl das CFA-UM häufig eine bessere Performance aufweist als das CFA-Modell, lassen sich durch Verwendung des CU-Modells Schätzprobleme vermeiden, die beim CFA-UM-Modell manchmal auftreten.

Conway kann mathematisch zeigen, dass die quadrierten Ladungen der jeweiligen Methodenfaktoren dem Mittel der entsprechenden unquadrierten Fehlerkorrelationen entsprechen. Die Äquivalenz kann auch anhand der a) künstlich erzeugten und b) realen MTMM-Daten gezeigt werden. In ersteren haben alle korrelierten Fehler des CU-Modells den Wert 0,2 und die Faktorladungen des CFA-UM-Modells den Wert 0,45. Die quadrierten Faktorladungen (0,2) gleichen somit den nicht quadrierten Fehlerkorrelationen. Was letztere betrifft, so wird hier nur eine von insgesamt drei Datenerhebungen besprochen, die in der Studie herangezogen wurden. Es handelt sich um Daten aus einem strukturierten Interview , durchgeführt an 102 Angestellten, die acht Fragen umfassten, wobei sieben Fragen von vorherein drei Dimensionen (Traits) zugeordnet wurden:

- Arbeitsmotivation (Mittelwert über zwei Fragen)

- Grundwissen im Job (Mittelwert über drei Fragen) und

- Befürchtungen über den Job (Mittelwert über zwei Fragen).

Drei Rater führten die Interviews durch und definieren die drei Methoden.

Die folgende Tabelle 32 zeigt die drei Traits-/drei Methoden-Korrelationsmatrix:

[149] Marsh (1989), S. 341; zitiert nach Conway (1998), S. 213.

Tabelle 32: MTMM-Matrix mit drei Traits und drei Methoden für ein strukturiertes Interview mit Angestellten (Conway)

Three-Trait, Three-Method MTMM Matrix for a Structured Interview

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*
Rater 1								
1. Motivation	—							
2. Job knowledge	.60							
3. Apprehensiveness	.51	.64						
Rater 2								
4. Motivation	*.51*	.54	.40					
5. Job knowledge	.22	*.58*	.34	.50				
6. Apprehensiveness	.32	.44	*.59*	.54	.60			
Rater 3								
7. Motivation	*.54*	.32	.34	*.53*	.15	.26		
8. Job knowledge	.35	*.60*	.40	.41	*.62*	.41	.43	
9. Apprehensiveness	.37	.43	*.63*	.37	.31	*.58*	.51	.56

Note. **MTMM = multitrait-multimethod. Italicized values are convergent validities. *N* = 102. All correlations were significant, $p < .05$.**

Sowohl das CU-Modell also auch das CFA-UM-Modell konnten sehr gut an die Daten angepasst werden: $\chi^2(15) = 22{,}12; p > 0{,}05$. Der Tucker-Lewis-Index betrug 0,97[150] und beide Modelle resultierten in identischen Trait-Faktorladungen und Korrelationen zwischen den Traits. Für die Fehlerkorrelationen des CU-Modells wurden ein Gesamtmittelwert und ein Mittelwert über jede der Methoden hinweg ermittelt. Für das CFA-UM-Modell können entsprechend ein Gesamtmittelwert der Faktorladungen angegeben werden und ein Mittelwert der zur jeweiligen Methode gehörenden Faktorladungen. Die Gesamtdurchschnitte waren identisch (0,20) und die methodenspezifischen Mittelwerte fast identisch (CU: 0,21; 0,21; 0,17; CFA-UM: 0,22; 0,21; 0,16).

Das für diesen Abschnitt relevante Ergebnis der Studie von Conway besagt, dass die Verwendung von Correlated Uniqueness- statt Multitrade-Multimethod-Modellen nicht bedeutet, dass die Reliabilität eines Testinstruments notwendigerweise sinkt. Wenn Conways Methode angewendet wird, dann erhöhen korrelierte Fehler nicht die Error-Score-Varianz, sondern werden der True-Score-Varianz zurechenbar. Mittelt man die Fehlerkorrelationen, die mit einer bestimmten Methode erhobenen Items zuzuordnen sind, so kann – mit folgenden Einschränkungen – der Anteil der Methodenvarianz in CU-Modellen ebenso ermittelt werden wie bei Anwendung von MTMM-Modellen: Es wird angenommen, dass – analog dem CFA-UM-Modell – die Methoden untereinander nicht korreliert sind. Diese *erste Einschränkung* ist in

[150] Ein TLI größer 0,9 zeigt eine gute Anpassung an.

meinem Kontext relevant, weil hier angewandte instrumenten- oder, genauer, apparatur-basierte Erhebungsmethoden eher zu Korreliertheit neigen als rater-basierte. Sie hat jedoch insofern keine einschneidenden Konsequenzen, da das CU-Modell in Hinblick auf Verletzungen der Annahme unkorrelierter Methoden ziemlich robust ist[151]. Man kann erwarten, dass eine Verletzung dieser Annahme zu einer Unterschätzung der Methodenvarianz führt, die jedoch moderat ausfallen dürfte[152].

Eine *zweite Einschränkung*, wonach es nicht möglich sei, den Anteil der Methodenvarianz für einzelne Items mittels CU-Modellen zu schätzen, ist nach Scullen (1999), der Conways Ansatz weiterverfolgte, nicht gegeben.

Scullen (1999) entwickelte eine Methode, die konfirmatorische Faktorenanalyse der Kovarianzmatrix von CU-Modellen, die erstmals eine Schätzung der Methodenvarianz für jedes einzelne Item erlaubt und zudem unverzerrter und mit größerer Präzision als bisher möglich operiert.[153] Damit wurde ein wichtiger Beitrag zur Verbesserung der *Konstruktvalidität* geleistet: „For a given trait, measurement methods, that show very little evidence of method variance, are likely to be more construct valid (all else equal) than other methods, that exhibit a larger proportion of method variance.“[154] Er konnte weiterhin zeigen, dass die durchschnittliche Fehlerkorrelation eine *Untergrenze* des korrekten Methodenvarianzanteils ist. Schließlich lieferte er eine Erklärung für Conways kontraintuitives Ergebnis, dass Fehlerkorrelationen bei Ermittlung der Methodenvarianz nicht zu quadrieren seien, um den quadrierten Ladungen der Methodenfaktoren zu entsprechen[155]: Fehlerkorrelationen, besser gesagt, korre-

---

[151] Conway (1998), S. 220

[152] Williams et al. (1989) bzw. Marsh and Bailey (1996), zitiert nach Conway (1998).

[153] Die Darstellung von Scullens Methode unterbleibt aus Platzgründen, obwohl sie eine ernsthafte Konkurrenz zum MTMM-Ansatz liefert.

[154] Scullen (1999), S. 289.

[155] Conway zeigte dieses Ergebnis auf, indem er einen Spezialfall benutzte: Gegeben sei eine MTMM-Matrix, die drei Traits einschließt, von welchen ein jeder denselben Anteil an Methodenvarianz auf sich vereint. Die algebraische Beziehung zwischen der Formulierung für Methodenvarianz im CTUM -Modell (CFA-UM) und der Formulierung für Methodenvarianz im CTCU -Modell (CU) lautet unter Verwendung der Originalnotation folgendermaßen:
Für das CFA-UM-Modell gilt:
$DT,\ SM_{Correlation} = t_1 t_2 r + m_1 m_2,$
wobei $DT,\ SM_{Correlation}$ die beobachtete Korrelation zwischen zwei Traits, die durch dieselbe Methode gemessen werden; t_1: Traitfaktorladung der ersten gemessenen Variablen; t_2: Traitfaktorladung der zweiten gemessenen Variablen; r: Korrelation zwischen den beiden Traitfaktoren (Hervorhebung hinzugefügt); m_1: Methodenfaktorladung der ersten gemessenen Variablen: m_2: Methodenfaktorladung der zweiten gemessenen Variablen.
Die entsprechende Beziehung für das CTCU-Modell lautet, nur im letzten Summanden abweichend:
$DT, SM_{Correlation} = t_1 t_2 r + cu$, wobei cu: Korrelation zwischen den Fehlertermen der zwei gemessenen Variablen.
Laut Marsh, 1989, (siehe oben) müssen die rechten Gleichungsseiten der beiden Gleichungen identisch sein. und daher $m_1 m_2 = cu$. Da im Spezialfall alle Variablen identische Methodenvarianzanteile haben, folgt, dass

lierte Fehlerterme sind keine Korrelationen sondern Kovarianzen, denn Fehlerterme sind – auch in komplett standardisierten LISREL-Lösungen – nicht standardisiert[156].

10.3.2.5.2 Fehlerkorrelationen infolge von Eigenschaften der Person (Traits)

Neben den Studien, die sich mit durch *method variance* induzierten Fehlerkorrelationen beschäftigen (Yang et al. (2011) durch Einführung von Gruppen- oder Methodenfaktoren – auch als Alternative zu Fehlerkorrelationen – Conway (1998) und Scullen (1999) durch Vergleich von Modellen mit unkorrelierten Methodenfaktoren und correlated Uniqueness-Modellen) soll auch eine Studie von Rubio und Gillespie (1995) angeführt werden, die sich mit Fehlerkorrelationen infolge der (Fehl-)Spezifikation von *Traits* beschäftigt[157]. Ausgangspunkt ist ein von Bollen und Hoyle (1990) publiziertes zweifaktorielles SEM, das die Traits „belonging" und „morale" enthält, welche beide sich jeweils in drei beobachteten Variablen zum Ausdruck bringen (Abbildung 86). Die Kovarianzmatrix dieser insgesamt sechs beobachteten Variablen, gemessen an 110 Probanden, ist in folgender Tabelle 33 wiedergegeben:

Tabelle 33: Kovarianzmatrix der Studie über wahrgenommenen Zusammenhalt (Bollen & Hoyle)

Covariance Matrix From Bollen and Hoyle (1990)

Covariance Matrix					
6.43					
5.43	7.26				
4.56	4.47	8.03			
5.57	4.94	4.48	7.93		
5.38	5.58	4.94	6.41	7.69	
4.92	4.82	4.37	5.47	5.88	6.46

Dieses Modell wird durch Entfernung eines Traits (belonging) neu spezifiziert und getestet. Das Ergebnis ist in Zeile 1 der Tabelle 34 ablesbar und weist auf eine schlechte Anpassung hin.

$m_1 = m_2$ und daher auch $m_1 m_2 = {m_1}^2 = {m_2}^2 = cu$.
Im CFA-UM Modell schätzt die quadrierte Methodenfaktorladung den Anteil der Varianz, der auf die Methode zurückgeht, so dass ${m_1}^2$ und ${m_2}^2$ Schätzungen der Methodenvarianzanteile im CFA-UM Modell sind und der Methodenvarianzanteil im CU-Modell ist.

[156] Mathematische Darlegung, siehe Scullen (1999), S. 280. *Correlated uniquenesses*, also z. B. Cov (δ_1, δ_2) sind im LISREL-Modell (1996) Elemente der Residualmatrix Θ_δ fern der Hauptdiagonale (off-diagonal elements). Fußnote 1 stellt fest, dass die Hauptdiagonale von Θ_δ die Fehlervarianzen enthält und weist auf zwei Möglichkeiten hin: 1. Die Ladungen der Stör- oder Fehlerterme sind gleich eins und die Varianzen sind kleiner als eins. 2. Die Varianzen der Fehlerterme sind gleich 1 (standardisierte Fehlerterme) und die Ladungen der Fehlerterme sind dann Standardabweichungen von ihren jeweiligen Fehlertermen.

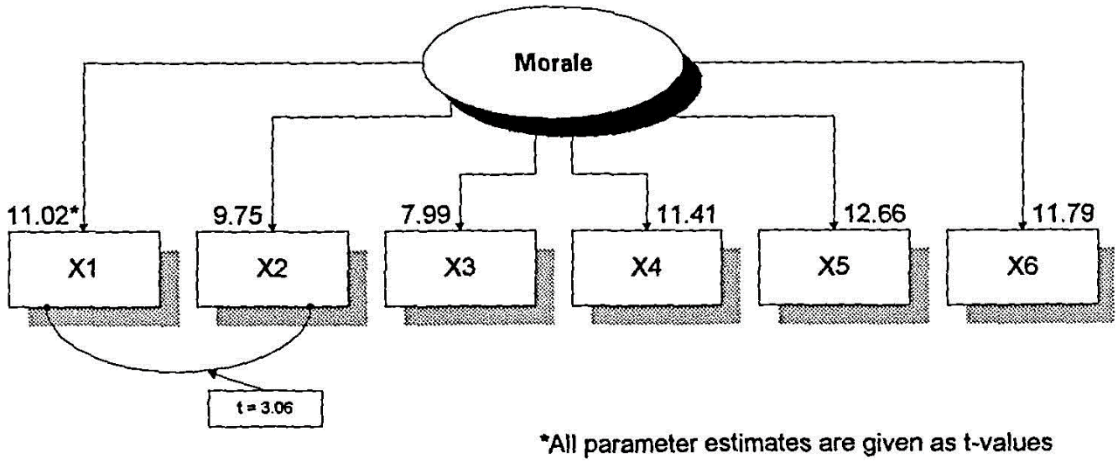
[157] Rubio & Gillespie (1995), S. 373.

Tabelle 34: Fit und Fitindizes von Modellvarianten des wahrgenommenen Zusammenhalts (Rubio & Gillespie)

	Chi-Quadrat	Df	p	RMSEA	CN158
1 Faktor-Modell	28,37	9	< ,001	,14	84,23
1 Faktor-Modell, 1 Fehlerkorrelation	15,92	8	,04	,095	97
2 Faktoren-Modell	13,89	8	,085	,08	158,62

Modifikationsindices geben nun den Hinweis, dass eine Korrelation zwischen den Fehlern der Variablen x1 und x2 besteht, die dem eben entfernten Trait „belonging“ zugehörten (Abbildung 85).

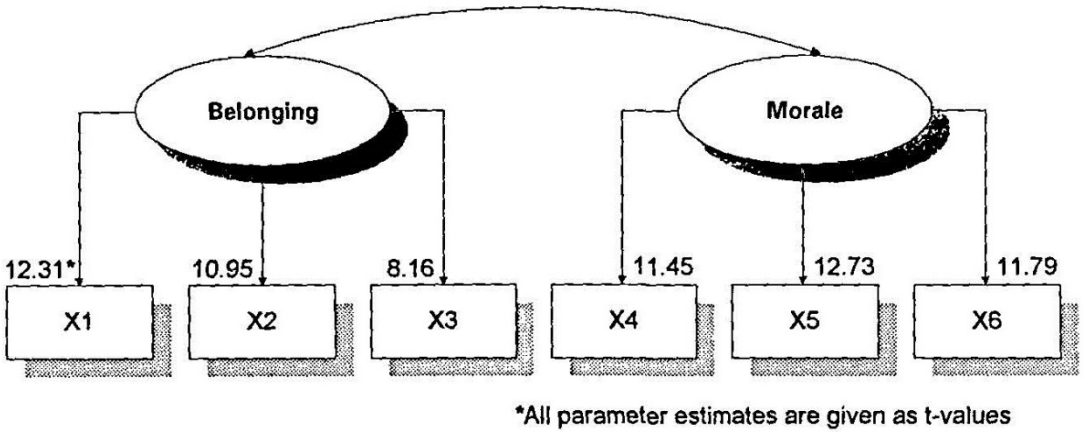
Abbildung 85: Messmodell zu „morale“ nach Entfernten des Traits „belonging“ (Rubio & Gillespie)



Die Aufhebung der Annahme der Unkorreliertheit der Fehler für dieses Variablenpaar führt zu einer halbwegs guten Anpassung des Modells (Tabelle 34, Zeile 2). Der signifikante Chi-Quadrat-Test auf Unterschiede zwischen den beiden Modellen weist eine überlegene Anpassung des durch Fehlerkorrelation ergänzten Modells nach ($\chi^2 = 12{,}24$, $df = 1$, $p < 0{,}001$). Aus dem Modell ist jedoch nicht ersichtlich, warum das o. g. Fehlerpaar korreliert ist und erst das Wissen, dass dieses einen weiteren Faktor misst, macht die Fehlerkorrelation plausibel. Die Wiedereinführung des ursprünglich bei Bollen und Hoyle vorhandenen und aus experimentellen Gründen entfernten zweiten Faktors „belonging“ bringt die Fehlerkorrelation zum Verschwinden und verbessert das Modell ein weiteres Mal (Tabelle 34, Zeile 3 und Abbildung 86).

[158] Siehe Hu, L.-T. & Bentler, P. M. (1995), S. 86: Holters CN, ein Index zur Schätzung der adäquaten Stichprobengröße. Wenn der Index den Wert 200 überschreitet, dann bedeutet dies, dass das gegebene Modell eine angemessene Repräsentation der Stichprobendaten ist. Offenbar ist dies hier nicht der Fall.

Abbildung 86: SEM zur Studie über wahrgenommenen Zusammenhalt mit den Faktoren „belonging“ und „morale“ (Bollen & Hoyle)



Letzteres ist zwar nicht signifikant besser als das Modell mit der Fehlerkorrelation, aber angemessener, auch weil es der Theorie der Kohäsion der Autoren eher entspricht.

10.3.3 Simultane Betrachtung von Multidimensionalität und Messfehlerkorrelation

Die oben angestellten Betrachtungen über Multidimensionalität zum einen und Messfehlerkorrelationen zum anderen als reliabilitätsrelevante Spezifika von Messinstrumenten reichen in isolierter Form nicht aus, da sie in der Forschungsrealität nicht nur isoliert auftreten. Komaroff stellt 1997 fest, dass Zimmerman et al. 1993 die Effekte dieser beiden Phänomene auf die Reliabilität in Simulationsstudien jeweils separat untersuchten und macht sich die Analyse der gleichzeitigen Wirkung beider zum Programm. Ich werde im Folgenden die wichtigsten Resultate seiner Computersimulation wiedergeben. Die ergänzenden mathematischen Ausführungen finden sich im selben Artikel und werden hier nicht behandelt. Unter anderem wird hier ein modifizierter Koeffizient alpha $_K$ vorgestellt, der zur Berechnung eines verzerrungsfreien Reliabilitätskoeffizienten geeignet ist, falls Fehlerkorrelationen nicht durch das experimentelle Design kontrollierbar sind und empirische Schätzungen über die Fehlerkorrelationshöhe vorliegen.

10.3.3.1 Methode und Daten

Die künstlichen Daten wurden erzeugt mit Prozeduren und Funktionen aus SAS[159] Datendateien. Es wurden insgesamt 150 zusammengesetzte Tests (Composite tests) oder Skalen pro-

[159] RANNOR function zur Herstellung von normalverteilten Pseudozufallszahlen, Proc IML zur Manipulation von Korrelationen sowohl zwischen t_i (True Scores) als auch e_i (Error Scores) und um alpha als auch alpha$_K$ zu

duziert, wobei diese wahlweise aus 6, 12 oder 18 Items bestanden und für jedes Item 10.000 Werte simuliert wurden. Die 150 Tests entstanden, indem eine best. Anzahl von Items (s. o.), darunter unterschiedliche Teilmengen korrelierter Items (**6** Varianten), unterschiedliche Korrelationsgrade der True-Score-Variablen (**5** Varianten r = 0, 0,2, 0,5, 0,7 oder 1) und Grade von Fehlerkorrelation der betroffenen Items (**5** Varianten r = 0, 0,2, 0,5, 0,7 oder 1) miteinander kombiniert wurden 6 * 5 * 5 = 150).[160]

10.3.3.2 Resultate

Die folgende Tabelle 35 enthält für jede der drei Itemanzahlen (6, 12, 18) Referenzwerte für unverzerrte alphas in einer Situation, in der keine Fehlerkorrelation vorliegt und essentielle τ −Äquivalenz gegeben ist. In den ersten beiden Spalten und der jeweils ersten Zeile, in der $\rho(t_i, t_j) = 1$, betragen alpha 0,53, 0,69 und 0,78 und die wahre Reliabilität 0,54, 0,70 und 0,78.

berechnen. Weitere Verwendung fanden PROC CORR mit den Funktionen „Cov“ und „alpha“ zur Berechnung der Kovarianzmatrizen für alpha $_K$ und zur Kontrolle der Berechnungen von alpha in IML von sowie CALIS, das auch in meiner Analyse von SEM mit künstlich erzeugten Daten die Hauptrolle spielte. Weitere Erläuterungen finden sich in Komaroff (1996).

[160] Die Details zur *Datenerzeugung* sind umfänglich und deren Darstellung würde den Rahmen sprengen. Im Aufsatz finden sich ebenso genaue Angaben zu *Manipulation der Annahme der Essentiellen τ −Äquivalenz, Manipulation der unkorrelierten Fehler, CFA-Messmodellen, alpha$_K$ und der Reliabilität.*

Tabelle 35: Alpha, wahre Reliabilität und modifizierter Koeffizient alpha $_K$ bei unterschiedlicher Itemanzahl sowie Zahl und Stärke von Fehlerkorrelationen (Komaroff)

α, $\rho_{XX'}$, and α_K for Composite Tests or Scales of Different Lengths, and $\rho(t_i, t_j)$ and $\rho(e_i, e_j)$ of 0 to 1

	$\rho(e_i, e_j) = 0$		$\rho(e_i, e_j) = .2$		$\rho(e_i, e_j) = .5$		$\rho(e_i, e_j) = .7$		$\rho(e_i, e_j) = 1$	
$\rho(t_i, t_j)$	α	$\rho_{XX'}$	α	α_K	α	α_K	α	α_K	α	α_K
6 Items; 3 Items With Correlated Errors										
1.0	.53	.54	.60	.55	.66	.55	.69	.54	.74	–
.7	.42	.47	.50	.44	.58	.43	.63	.43	.69	–
.5	.34	.41	.43	.35	.53	.34	.58	.34	.65	–
.2	.15	.29	.27	.14	.41	.14	.48	.14	.58	–
0.0	0.00	.17	.13	–	.31	–	.40	–	.53	–
12 Items; 3 Items With Correlated Errors										
1.0	.69	.70	.71	.70	.72	.70	.73	.70	.75	–
.7	.60	.63	.62	.60	.64	.60	.65	.60	.68	–
.5	.50	.56	.53	.51	.56	.51	.58	.51	.62	–
.2	.26	.38	.31	.27	.37	.27	.40	.27	.48	–
0.0	0.00	.17	.05	–	.15	–	.20	–	.32	–
12 Items; 6 Items With Correlated Errors										
1.0	–	–	.75	.71	.80	.71	.83	.71	.87	–
.7	–	–	.68	.62	.75	.62	.79	.61	.84	–
.5	–	–	.61	.52	.71	.53	.75	.53	.82	–
.2	–	–	.45	.28	.61	.28	.68	.28	.78	–
0.0	–	–	.29	–	.52	–	.62	–	.74	–
18 Items; 3 Items With Correlated Errors										
1.0	.78	.78	.79	.78	.79	.78	.79	.78	.80	–
.7	.69	.72	.71	.70	.72	.70	.72	.71	.73	–
.5	.61	.65	.63	.62	.64	.62	.65	.62	.66	–
.2	.37	.47	.41	.39	.44	.39	.45	.39	.48	–
0.0	0.00	.17	.04	–	.10	–	.14	–	.23	–
18 Items; 6 Items With Correlated Errors										
1.0	–	–	.80	.79	.83	.79	.84	.79	.85	–
.7	–	–	.74	.71	.77	.71	.79	.71	.81	–
.5	–	–	.67	.63	.72	.63	.74	.63	.78	–
.2	–	–	.49	.40	.58	.40	.63	.40	.70	–
0.0	–	–	.21	–	.41	–	.49	–	.62	–
18 Items; 9 Items With Correlated Errors										
1.0	–	–	.83	.79	.87	.79	.89	.79	.91	–
.7	–	–	.77	.71	.83	.71	.86	.71	.89	–
.5	–	–	.72	.63	.79	.63	.83	.63	.88	–
.2	–	–	.58	.38	.72	.40	.78	.40	.84	–
0.0	–	–	.40	–	.64	–	.72	–	.82	–

10.3.3.2.1 Verletzung ausschließlich der Annahme unkorrelierter Fehler

Je größer die Fehlerkorrelation, desto stärker ist die Zunahme des Wertes von Koeffizient alpha. Der adjustierte Koeffizient alpha$_K$ hingegen reagiert nicht auf die Höhe der Fehlerkorrelation und ist bis auf geringe Abweichungen identisch mit dem jeweils entsprechenden unadjustierten alpha bzw. der wahren Reliabilität (siehe sechs o. g. Werte). Wenn mindestens essentielle τ −Äquivalenz gegeben ist, aber korrelierte Fehler vorhanden, so wird der wahre Wert durch alpha über- und durch alpha$_K$ korrekt geschätzt. Z. B. beträgt bei sechs Items, davon drei korreliert (Zeile 1, Spalte 7) und Fehlerkorrelation in Höhe von 0,7, die wahre Reliabilität 0,54, alpha 0,69 und alpha$_K$ 0,54. Der nichtadjustierte Reliabilitätskoeffizient überschätzt somit die wahre Reliabilität in diesem Extremfall um 0,14; der adjustierte trifft sie

hingegen genau. Je größer die Höhe der Fehlerkorrelation, desto größer ist die resultierende Überschätzung.[161]

10.3.3.2.2 Verletzung ausschließlich der Annahme essentieller τ −Äquivalenz i. S. d. Unidimensionalität

Je größer die Abweichung von der essentiellen τ −Äquivalenz im Sinne der Unidimensionalität, desto größer ist die Abnahme der wahren Reliabilität und des nichtadjustierten Alphas.[162] Wenn keine Fehlerkorrelation gegeben ist, aber die True-Score-Variablen nicht perfekt korreliert sind, $\rho(t_i, t_j) < 1$, so wird der wahre Wert durch alpha unterschätzt. Z. B. beträgt bei sechs Items, davon drei korrelierten (Zeile 5, Spalte 1) und $\rho\ (t_i, t_j) =$,0 alpha 0,00 und die wahre Reliabilität 0,17 (Spalte 2). Der nichtadjustierte Reliabilitätskoeffizient unterschätzt mithin die wahre Reliabilität in diesem Extremfall um 0,17. Je weniger die True-Score-Variablen miteinander korreliert sind, desto größer ist das Maß der Unterschätzung der wahren Reliabilität.

10.3.3.2.3 Gleichzeitige Verletzung beider Annahmen

Ausgangpunkt der Betrachtung sei eine beliebige Zeile, in der die essentielle τ −Äquivalenz im Sinne der Unidimensionalität verletzt ist, d. h. $\rho(t_i, t_j) < 1$. Wenn die Fehlerkorrelationen gleich 0 sind, dann unterschätzt alpha die wahre Reliabilität. Wenn Fehlerkorrelationen größer gleich 0,2 gegeben und damit beide Annahmen gleichzeitig verletzt sind, dann überschätzt alpha mit folgenden Ausnahmen das adjustierte alpha: Wenn Fehlerkorrelationen mit 0,2 sehr gering, und der Anteil der korrelierten Items an den Items im composite test ebenfalls gering (3/18 =17 % oder 3/12 = 25 %) sind, dann kompensiert die überschätzende Wirkung der Feh-

---

[161] Eine Ausnahme stellt das Modell mit 18 Items, drei davon mit korrelierten Fehlern, dar: alpha bleibt unter der Bedingung einer True-Score-Korrelation von geringstenfalls 0,7 über die Grade von Fehlerkorrelation hinweg nahezu konstant.

[162] Die Korrelation von True-Score-Variablen des kongenerischen Modells (und aller seiner Spezialfälle (essentielle τ −Äquivalenz, τ −Äquivalenz, τ −Parallelität) ist stets eins, da Corr(x, y) =1, für y = a + b* x wobei a und b ungleich 0. (Lineare Transformation). Daher entspricht die Angabe, „Abweichung von essentieller τ −Äquivalenz" zwar Komaroffs Wortlaut und ist auch zutreffend, aber nicht hinreichend. Denn die Verletzung der Annahme essentieller τ −Äquivalenz ist auch bei kongenerischen Variablen gegeben. Hinzu kommt hier das Kriterium: Corr(t_i, t_j) < 1, also Multidimensionalität. Kurz gesagt: Essentielle τ −Äquivalenz impliziert Unidimensionalität aber Unidimensionalität impliziert nicht essentielle τ −Äquivalenz. Der Unterschied zwischen mindestens essentiellen τ −äquivalenten Variablen (mindestens bedeutet essentielle τ −Äquivalenz und ihre Speziallfälle) und kongenerischen Variablen beruht darin, dass die Varianzen der True-Score-Variablen bei ersteren gleich und bei letzteren ungleich sind. Genauere Definitionen hinsichtlich der Auswirkungen einer Lineartransformation auf die Varianzen von t_i und t_j finden sich bei Komaroff (1997), S. 338. Es ist nicht bekannt, ob bei Verletzung der essentiellen τ −Äquivalenz im Sinne von Corr(t_i, t_j) < 1 *auch* gegen die Annahme verstoßen wurde, dass die Varianzen der True Scores gleich seien. Es ist zwar zu vermuten, jedoch wäre Corr(t_i, t_j) < 1 unter Wahrung der Varianzengleichheit theoretisch programmierbar.

lerkorrelation die unterschätzende Wirkung der Verletzung der τ −Äquivalenz. Ist die Fehlerkorrelation größer als 0,2 oder der Anteil der korrelierten Items an den Items des zusammengesetzten Tests größer als 25 %, so ist die Wirkung der Fehlerkorrelation überkompensativ.

$Alpha_K$ bei beliebiger Fehlerkorrelation ist eine gute Schätzung für alpha unter Gültigkeit der Annahme unkorrelierter Fehler. Dies ist daran erkennbar, dass die $Alpha_K$ eines Modells mit einer bestimmten Itemanzahl über alle Grade von Fehlerkorrelation zunächst einander gleichen, wenn eine bestimmte True-Score-Korrelation herausgegriffen wird. (Bsp.: 12 Items. Wenn die True-Score-Korrelation 0,7 beträgt, dann nimmt $alpha_K$ einen Wert von 0,60 bis 0,62 an, sowohl bei 25 % als auch 50 % fehlerkorrelierten Items.) Weiterhin gleichen diese fast identischen $alpha_K$, gegeben eine bestimmte True-Score-Korrelation, zusätzlich den Alphas unter der Bedingung einer Fehlerkorrelation von 0. Bsp.: die $alpha_K$ für 12 Items, davon sechs fehlerkorreliert und mit einer Fehlerkorrelation von 0,7, die bei den vier True-Score-Korrelationsstufen > 0 auftreten (0,71, 0,61, 0,53, 0,28), sind nur unwesentlich verschieden von jenen bei drei fehlerkorrelierten Items (0,70, 0,60, 0,51, 0,27) als auch von den entsprechenden alphas bei einer Fehlerkorrelation von 0 (0,69, 0,60, 0,50, 0,26). Die Subtraktion der in der konfirmatorischen Faktorenanalyse ermittelten positiven Summe der Fehlervarianz vom Koeffizienten alpha resultiert daher in vertrauenswürdigen adjustierten Koeffizienten $alpha_K$. Ein relativ hoher Wert des unadjustierten Koeffizienten alpha lässt nicht auf eine hohe Reliabilität schließen. Dies ist beispielsweise im Modell mit 18, zu 50 % fehlerkorrelierten Items an dem Wert für alpha ablesbar, der in der Situation „Fehlerkorrelation 0,7 und True-Score-Korrelation von 0,2“ mit 0,78 ermittelt wird. Das adjustierte alpha beträgt nur 0,40 und die wahre Reliabilität dürfte ca. bei 0,47 liegen, wenn man den entsprechenden Wert unter der Bedingung unkorrelierter Fehler zugrundelegt (Modell mit 18 Items, Spalte 2).

Mit zunehmendem Anteil von Items mit korrelierten Fehlern überschätzt alpha die Reliabilität zunehmend. Zum Nachvollzug kann für ein Modell mit bestimmter Itemanzahl und geringem Anteil von fehlerkorrelierten Items ein beliebiges alpha herausgegriffen und mit dem entsprechenden alpha eines Modells mit höherem Anteil verglichen werden.

An der Höhe des unadjustierten alpha kann nicht abgelesen werden, ob ein Verstoß gegen die Unkorreliertheit der Fehler gegeben ist oder Multidimensionalität vorliegt. Alpha nimmt beispielsweise im Modell mit 18 Items den Wert 0,78 für den Fall an, dass beide Annahmen vollständig erfüllt sind. Derselbe Wert allerdings resultiert auch aus einem Modell, in dem 33 Prozent dieser Items perfekt korreliert sind und die True-Score-Korrelation 0,5 beträgt.

Komaroff folgend ist festzuhalten:

1. Alpha ist eine Funktion, die u. a. abhängt

- von der Zahl der Items mit korrelierten Fehlern

- vom interaktiven Effekt des gleichzeitigen Verstoßes gegen die Unkorreliertheit der Fehler zum einen und die essentiellen τ-Äquivalenz i. S. der Unidimensionalität zum anderen.

2. Alpha kann nicht differenzieren zwischen den Quellen der beobachteten Inter-Item-Kovarianz und reagiert in gleicher Richtung auf Veränderung einer der beiden Varianzquellen: Nimmt die True-Score-Kovarianz oder die Fehlerkovarianz zu, steigt alpha[163].

3. Eine geringfügige positive Fehlerkorrelation führt schon zur Überschätzung der Reliabilität, auch wenn gleichzeitig Multidimensionalität vorliegt, es sei denn, der Anteil der Items mit Fehlerkorrelationen ist klein.

4. Die Inflation von alpha kann reduziert oder gar eliminiert werden, indem man alpha um die geschätzte positive Summe der Fehlerkovarianz dezimiert ($alpha_K$).

10.4 BEWERTUNG VON MODELLEN DURCH FIT-INDIZES (HEENE ET AL.)

Dieser Abschnitt ist orientiert an dem Aufsatz von Heene et al. (2012), die darauf hinweisen, dass trotz des Auftretens korrelierter Fehlerterme in SEM zur Modellbewertung globale Fit-Indizes verwendet werden. Heene et al. untersuchen in ihrer Studie die Sensitivität von normalerweise gewählten Cut-Off-Values globaler Fit-Indizes unter folgenden Bedingungen:

a) verschiedene Arten von Fehlertermkorrelationen (genauer: Verletzungen der Annahme unkorrelierter Fehler),

b) verschiedene Grade von Fehlertermkorrelationen,

c) verschiedene realistische Bedingungen von Faktorladungen.

Sie richten ihr Augenmerk dabei auf die Fit-Indizes: RMSEA, SRMR und CFI und verfolgen damit das Ziel, Ratschläge für die Identifikation von Fehlpezifikationen zu geben. Als Grenzwerte für Fit-Indizes sind, orientiert an Hu und Bentler (1999) und Beauducel und

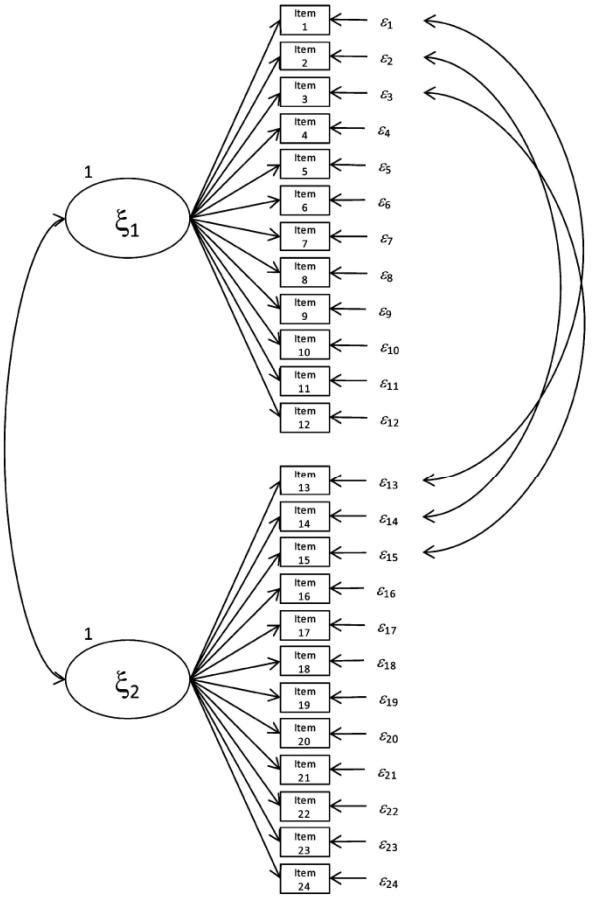
[163] „Alpha absorbs both true and Error Score covariance into observed Covariance equally“ (Komaroff (1997), S. 347). Siehe auch Osburn, a. a. O. und 10.3.1 hinsichtlich der Analyse des Koeffizienten alpha. Die Zunahme der True-Score-Kovarianz mündet in einer Erfüllung der Annahme der essentiellen τ −Äquivalenz; die Zunahme der Error-Score-Kovarianz ist hingegen mit Verletzung der Annahme der Unkorreliertheit verbunden.

Wittman (2005) und abweichend von Eid et al. (2011), die durchgehend strengere Grenzen vorschlagen,[164] folgende Werte gesetzt worden: SRMR < 0,11, RMSEA < 0,06 bei N >= 250 und RMSEA < 0,08 bei N < 250; CFI > 0,95. (Gute Anpassung ist gegeben, wenn der jeweilige Index in den empfohlenen Bereich fällt.)

10.4.1 SIMULATION UND DESIGN

Zwei Modelle A und B wurden entwickelt, die aus je 24 gemessenen Variablen bestehen, wobei zwölf Variable einem allgemeinen Faktor zugeordnet sind, weitere zwölf einem anderen und die beiden Faktoren untereinander korreliert sind. Modell A (Abbildung 87) weist drei korrelierte Fehlerpaare auf und Modell B (Abbildung 88) sechs, wobei nur Korrelationen zwischen Fehlern bestehen, die Items unterschiedlicher Variablengruppen zugeordnet sind.

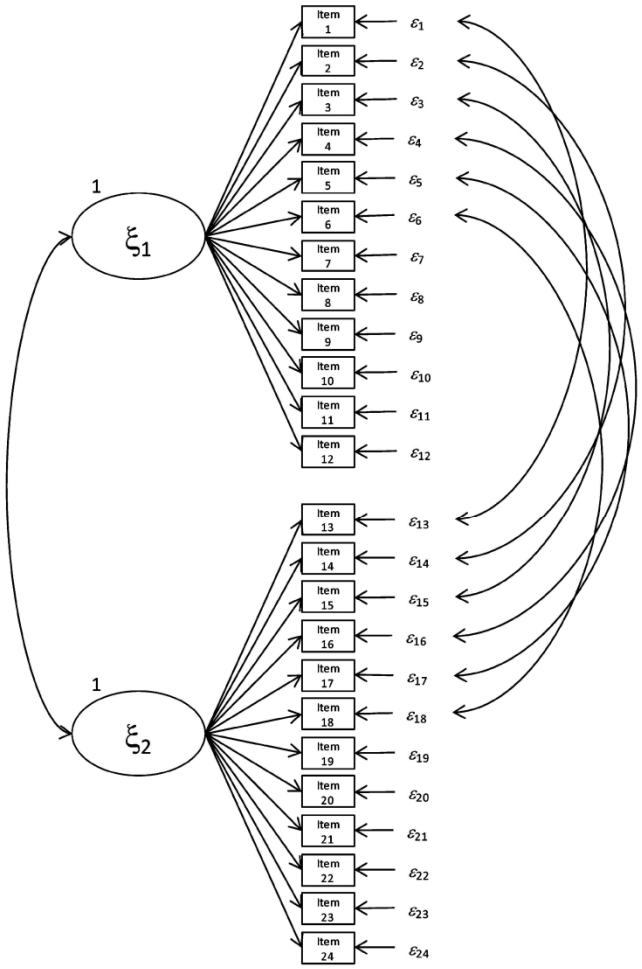
Abbildung 87: SEM-Schema A mit 2 Faktoren und je zwölf Indikatoren (Heene et al.)



Schematic representation of Model A with three correlated errors.

164 Eid et al. (2011), S. 880 ff.

Abbildung 88 SEM-Schema B mit 2 Faktoren und je zwölf Indikatoren (Heene et al.)



Schematic representation of Model B with six correlated errors.

Es wurden Daten zu folgenden Bedingungen simuliert: zwei Intervalle von Faktorladungen, fünf verschiedene Stichprobengrößen (150, 250, 500, 1000, 2500) und vier verschiedene Fehlerkorrelationsstrukturen. Anschließend wurden alle Kombinationen aus diesen Bedingungen gebildet. Für jede der resultierenden 40 Datenstrukturen wurden 1000 Wiederholungen durchgeführt. Bei der Festsetzung der Höhe der Faktorladungen orientierte man sich an empirischen Faktorladungen aus der deutschen Version des Big Five Inventory (NEO-FFI). Die Faktorladungen wurden in Anbetracht der Erkenntnis, dass die Größe der Modellparameter a) auf die Sensitivität des Fitindex und b) auf die Power des Chi-Quadrat-Modelltests zur Entdeckung von Fehlspezifikation Einfluss nimmt[165], zur Herstellung der Vergleichbarkeit mit vor-

[165] Heene et al. (2011), Saris et al. (2009) und Schonemann (1981a, 1981b), zitiert nach Heene et al. (2012), S. 40.

liegenden Forschungsergebnissen um eine Konstante (0,20) erhöht, so dass sich als Kategorien Faktorladungen von 0,31 bis 59 (NEO-FFI) und 0,51 bis 0,79 (Higher) ergeben. Die Korrelation zwischen den Faktoren wurde auf 0,30 gesetzt. Die Werte der Fehlerkorrelationen wurden zufällig aus zwei stetigen Gleichverteilungen entnommen, deren Begrenzungen in der letzten Zeile folgender Tabelle 36 wiedergegeben sind, so dass sich bei verschiedenen Anzahlen (drei und sechs) und Vorzeichen von Fehlerkorrelationen die vier folgenden Spezifikationen ergaben.

Tabelle 36 Spezifikationen von Fehlerkorrelationen unter verschiedenen Simulationsbedingungen (Heene et al.)

Error Correlation Specifications Under Different Simulation Conditions

Weakest Model Deviation	*Weak Model Deviation*	*Moderate Model Deviation*	*Strong Model Deviation*
Three positive and negative $\rho_{\varepsilon_i \varepsilon_j}$	Six positive and negative $\rho_{\varepsilon_i \varepsilon_j}$	Three positive $\rho_{\varepsilon_i \varepsilon_j}$	Six positive $\rho_{\varepsilon_i \varepsilon_j}$
Uniform [−.3, .3]		Uniform [.3, .5]	

Bei der Simulation nach oben genannten Spezifikationen wurden folgende Fehlerkorrelationen realisiert (Tabelle 37):

Tabelle 37: Simulierte Fehlerkorrelationen unter verschiedenen Simulationsbedingungen

Simulated Error Correlations Under Different Simulation Conditions

$\rho_{e_i \varepsilon_j}$	*Three Positive*	*Six Positive*	*Three Positive and Negative*	*Six Positive and Negative*
$\rho_{\varepsilon_{item1}\varepsilon_{item13}}$	.356	.356	−.236	−.236
$\rho_{\varepsilon_{item2}\varepsilon_{item14}}$	.399	.399	.101	.101
$\rho_{\varepsilon_{item3}\varepsilon_{item15}}$	.417	.417	−.215	−.215
$\rho_{\varepsilon_{item4}\varepsilon_{item16}}$		.448		.189
$\rho_{\varepsilon_{item5}\varepsilon_{item17}}$		.445		.223
$\rho_{\varepsilon_{item6}\varepsilon_{item18}}$		.355		−.080

10.4.2 RESULTATE

Für beide Faktorladungssituationen wurden nach der Simulation über 1000 Replikationen Mittelwert und Standardabweichung für den ML-basierten Chi-Quadrat-Wert sowie die Fitindices RMSEA, SRMR und CFI errechnet. Nur die Ergebnisse für die höhere Faktorladungssituation (0,51–0,79) in Tabelle 39 sind für Vergleich mit dem Modell zur Schätzung des Glaukomschweregrades geeignet, da nur sie in etwa dessen Faktorladungen entsprechen (0,32 für AMP; alle anderen im Bereich 0,50–0,81). Tabelle 38 ist dennoch berücksichtigt worden, da

der Effekt der Faktorladungshöhe (über den Einzelfall der Glaukomdatenanwendung hinausgehend) von Interesse ist.

Tabelle 38: χ^2 und Fitindizes in Abhängigkeit von Fehlerkorrelationsspezifikation und Stichprobenumfang bei geringen Faktorladungen

Maximum Likelihood-Based Mean Chi-Square and Fit Index Values, NEO–FFI Loadings Results

Sample Size	*Type of Population Error Correlation* $\rho_{\varepsilon_i \varepsilon_j}$	$M\,\chi^2$ *(df = 251)*	*RMSEA*		*SRMR*		*CFI*	
			M	*SD*	*M*	*SD*	*M*	*SD*
$N = 150$								
	Three positive and negative	284.404 (24.916)	.027	.013	.067	.003	.896	.068
	Six positive and negative	297.308 (25.953)	.033	.011	.069	.003	.862	.070
	Three positive	336.889 (28.855)	.047	.008	.072	.003	.771	.069
	Six positive	417.36 (33.934)	.066	.007	.079	.004	.635	.065
$N = 250$								
	Three positive and negative	287.739 (25.655)	.022	.010	.053	.003	.928	.045
	Six positive and negative	308.766 (27.422)	.029	.008	.055	.003	.892	.047
	Three positive	375.394 (31.012)	.044	.006	.060	.003	.793	.048
	Six positive	508.517 (38.074)	.064	.005	.068	.003	.650	.046
$N = 500$								
	Three positive and negative	306.892 (26.642)	.020	.006	.039	.002	.945	.025
	Six positive and negative	349.244 (30.141)	.028	.004	.042	.002	.907	.027
	Three positive	483.489 (36.398)	.043	.003	.048	.002	.804	.029
	Six positive	749.625 (49.336)	.063	.003	.058	.002	.659	.031
$N = 1{,}000$								
	Three positive and negative	358.742 (30.559)	.02	.003	.030	.001	.947	.015
	Six positive and negative	443.728 (36.063)	.028	.003	.033	.001	.908	.017
	Three positive	708.947 (45.620)	.043	.002	.041	.001	.807	.019
	Six positive	1,241.091 (66.596)	.063	.002	.053	.001	.661	.021
$N = 2{,}500$								
	Three positive and negative	516.699 (38.680)	.021	.002	.023	.001	.947	.007
	Six positive and negative	727.056 (48.281)	.028	.001	.027	.001	.909	.009
	Three positive	1,390.987 (66.097)	.043	.001	.036	.001	.808	.011
	Six positive	2,718.732 (99.567)	.063	.001	.049	.001	.661	.013

Note. RMSEA = root mean squared error of approximation; SRMR = standardized root mean residual; CFI = comparative fit index.

Tabelle 39: χ^2 und Fitindizes in Abhängigkeit von Fehlerkorrelationsspezifikation und Stichprobenumfang bei hohen Faktorladungen

Maximum Likelihood-Based Mean Chi-Square and Fit Index Values, Higher Loadings Results

Sample Size	*Type of Population Error Correlation* $\rho_{\varepsilon_i \varepsilon_j}$	$M\,\chi^2$ *(df = 251)*	*RMSEA*		*SRMR*		*CFI*	
			M	*SD*	*M*	*SD*	*M*	*SD*
$N = 150$								
	Three positive and negative	284.321 (25.085)	.027	.013	.062	.004	.954	.031
	Six positive and negative	296.779 (26.055)	.033	.011	.063	.004	.939	.033
	Three positive	335.903 (28.940)	.047	.008	.066	.004	.893	.036
	Six positive	415.517 (33.937)	.066	.007	.071	.004	.811	.038
$N = 250$								
	Three positive and negative	287.199 (25.754)	.022	.010	.048	.003	.970	.020
	Six positive and negative	307.46 (27.410)	.029	.008	.050	.003	.954	.021
	Three positive	373.311 (31.107)	.044	.006	.054	.003	.906	.024
	Six positive	505.018 (38.012)	.063	.005	.061	.003	.822	.027
$N = 500$								
	Three positive and negative	305.694 (26.542)	.020	.006	.035	.002	.977	.011
	Six positive and negative	346.643 (29.996)	.027	.005	.038	.002	.961	.012
	Three positive	478.971 (36.176)	.042	.003	.042	.002	.912	.014
	Six positive	742.392 (48.869)	.062	.003	.051	.002	.828	.018
$N = 1{,}000$								
	Three positive and negative	356.325 (3.382)	.020	.003	.027	.001	.978	.006
	Six positive and negative	438.602 (35.926)	.027	.003	.030	.001	.962	.007
	Three positive	699.467 (45.271)	.042	.002	.036	.001	.913	.009
	Six positive	1,226.306 (65.885)	.062	.002	.046	.001	.829	.012
$N = 2{,}500$								
	Three positive and negative	51.342 (38.373)	.020	.002	.020	.001	.978	.003
	Six positive and negative	714.106 (47.680)	.027	.001	.024	.001	.962	.004
	Three positive	1,367.042 (65.919)	.042	.001	.031	.001	.914	.005
	Six positive	2,681.635 (99.057)	.062	.001	.042	.001	.830	.007

Note. RMSEA = root mean squared error of approximation; SRMR = standardized root mean residual; CFI = comparative fit index.

RMSEA und SRMR reagierten nicht (oder nur ungenügend) auf die Fehlspezifikation, wenn man die oben angeführten Cut-Off-Values anlegt, d. h., eigentlich abzulehnende Modelle wurden nicht abgelehnt.[166] CFI reagierte a) auf die mangelnde Anpassung und b) auf die Größe der Faktorladungen. Kleine Faktorladungen führen zu kleineren CFI Werten, da das Zielmodell bei kleineren Faktorladungen näher am Nullmodell ist. Allerdings reagierte CFI nicht bei schwachen oder sehr schwachen Fehlspezifikationen, wenn größere Faktorladungen gegeben waren. Daher reagiert CFI nicht nur auf Fehlspezifikation: Der Zusammenhang zwischen letzteren beiden ist konfundiert durch die Höhe der Faktorladungen. SRMR und CFI weisen zusätzlich eine leichte aber unterschiedliche Beziehung zur Fallzahl auf.[167] Resümee: Alle Fit-Indizes reagieren ungenügend auf Fehlspezifikationen im Hinblick auf die Annahme unkorrelierter Fehlerterme.

[166] Die Autoren behaupten, dass die beiden Indices *nicht* reagierten, da die resultierenden Werte entweder unter oder genau auf den Cut-off-Values lagen. Legt man die Ablehnungsvorschrift jedoch genau aus, dann reagierte wenigstens RMSEA bei Stichprobengrössen von 250 oder mehr auf die stärkste Modellabweichung, da ein Wert von 0,06 bereits zur Modellablehnung führt. Nach Eid et al. (2011) ist der Cut-off bei 0,05, so dass man bei groben Modellverletzungen durchaus auf den RMSEA zählen kann.

[167] Mittlere CFI-Werte stehen in einem leichten positiven monotonen Zusammenhang, mittlere SRMR-Werte stehen in einem leichten negativen monotonen Zusammenhang mit der Fallzahl.

Zum Zusammenhang zwischen CFI und Chi-Quadrat-Modelltest: Wie im Chi-Quadrat-Modelltest reagieren die Chi-Quadrat-Werte der Fit-Indizes auf die Fehlspezifikation insofern angemessen, da der Chi-Quadrat-Wert mit einer höheren Anzahl von Fehlerkorrelationen (6) größer ist als bei einer geringeren (3). Die Sensitivität des Chi-Quadrat-Modelltests hinsichtlich Fehlspezifikation unterliegt allerdings der Einschränkung, dass bei schwächster Fehlspezifikation (drei positive drei negative Fehlerkorrelationen) und einer Fallzahl von 150 der Test an Power verlor (p = 0,07, d. h. keine Ablehnung).

10.4.3 Alternative Annäherungen an das Problem mangelnder Sensitivität (EPC in Komb. mit MIs)

Saris (2009)[168] schlägt eine Evaluation von Modellen auf Basis des Expected Parameter Change (EPC) in Kombination mit Modifikation Indices (MIs) und der erwarteten statistischen Power des MI-Tests vor. EPC schätzt das Ausmaß der Fehlspezifikation für alle Constraint Parameters. Ein Modell ist schlecht, wenn es ein oder mehrere fehlspezifizierte relevante Parameter enthält. Was ein relevanter Parameter ist, ist im Einzelfall zu diskutieren. Wenn der MI durch die Fallzahl und die Größe der Faktorladungen beeinflusst ist, dann muss die Power des Signifikanztests einbezogen werden, um die Fehlspezifikation festzustellen.

Es werden Fehlerkorrelationen von 0,1 bereits als substantiell angesehen, da a) andere Parameter ernsthaft verzerrt sein können, wenn auch nur kleine Modellverletzungen beobachtet werden und b) kleine Werte in der Residual-Kovarianzmatrix nicht notwendigerweise bedeuten, dass das Ausmaß der Fehlspezifikation gering ist.[169] Das Ausmaß, in dem die Schätzung anderer relevanter Parameter durch eine solch mögliche kleine Modellverletzung betroffen ist, wird bewertet, indem eine iterative Prozedur durchgeführt wird:[170] Die Fehlerkorrelation mit dem größten MI kann frei geschätzt werden und die anderen werden auf 0 gesetzt. Wenn die resultierenden Modellparameter sich nicht gravierend verändern, wird die Prozedur gestoppt; falls sie es tun, wird der Prozess wiederholt, bis die Fehlerkorrelation mit dem nächstgrößten MI identifiziert ist. Es ist eine Software namens Jrule zur Durchführung dieser Prozedur verfügbar.[171] Heene et al. führten die Inspektion anhand des o. g. Beispiels simulierter Daten,

168 zitiert nach Heene et al. (2012), S. 45.
169 a. a. O.: Hayduk et al. (2007) und weitere Autoren belegen diesen Sachverhalt.
170 a. a. O.: Stuive et al. (2009).
171 Oberski (2009).

eingeschränkt auf die Stichprobengrössen 150 und 500, mit folgenden Ergebnissen durch:[172] Modifikationsindices aller wahren Fehlspezifikationen nehmen allesamt Werte größer als 3,84 (df = 1) an, d. h. Fehlspezifikationen werden mit Sicherheit identifiziert, sofern deren Wert nicht unterhalb von 0,1 liegt. Eine Ausnahme bildet die Kombination aus niedrigen Faktorladungen, schwächster Modellabweichung (Korrelationsstruktur 1) und N = 150, die unentdeckt bleibt. Die Power für N = 500 ist größer als die für N = 150, aber immer unterhalb der erwünschten 80 %. Man kann also schließen, dass trotz der geringen Power eine Fehlspezifikation vorliegt, sobald MI signifikant sind. Die Sensitivität des Tests ist mithin groß. Da aber auch in Wahrheit nicht fehlspezifizierte Korrelationen bei beiden Stichprobenumfängen- und beiden Faktorladungen größer als 3,84 annehmen können, kann man von einem Wert größer 3,84 nicht auf eine wahre Fehlspezifikation schließen und der positive Vorhersagewert der MI-Prozedur liegt unter 100 %. Die MIs von in Wahrheit fehlspezifizierten und daher signifikanten Fehlerkorrelationen sind größer als jene von Wahrheit nicht fehlspezifizierten signifikanten Fehlerkorrelationen. Die mit den falsch positiven MIs assoziierten EPCs nehmen hierbei unter N = 500 geringere Werte an als unter N = 150. Aus dem Beispiel lässt sich resümieren, dass die Untersuchung von lokalen Modellverletzungen unter Anwendung des iterativen MI-Ansatzes zur Entdeckung von korrelierten Fehlertermen in der Lage ist, die bei alleiniger Anwendung des globalen Chi-Quadrat-Modelltests nicht entdeckt worden wären, so dass eine Kombination beider Ansätze empfehlenswert ist.

10.5 Das Modell zur Glaukomdiagnose mit acht Indikatoren – Koeffizienten der CTT und des FITS

10.5.1 Strukturanalyse des Referenzmodells

Das 2001 publizierte Referenzmodell (Abbildung 71) beinhaltet zwei unkorrelierte Faktoren, von welcher der erste alle Indikatoren beeinflusst und der zweite einen Großteil derselben. Seine Struktur gleicht wesentlich dem in Eid et al. (2011) dargestellten Modell (11.7.2.2, Abbildung 94)[173]. Der zweite „Konzentration“ genannte Faktor ist derjenige Faktor, der (mit Ausnahme von AMP) die psychophysischen Verfahren beeinflusst. Daher können unter diesem Faktor arbeitshypothetisch „Aspekte psychophysischer Messung des Glaukoms“ zusammengefasst werden. Der zweite Faktor kennzeichnet somit die Abweichung der wahren Werte

[172] Bei der Inspektion von möglichen Fehlerkorrelationen wird ein großes Set von MIs berücksichtigt. Die Autoren bieten an, die Tabellen über diese exemplarische Suche von Fehlspezifikationen via Anfrage an den Erstautor zu beziehen.
[173] Eid et al. (2011), S. 853.

des durch psychophysische Verfahren ermittelten Glaukomschweregrades von den wahren Werten des allgemeinen Glaukomschweregrades. *Das Referenzmodell ist mithin ein Modell mit einem Methodenfaktor*. An ihm fällt auf, dass ERG der einzige Indikator ist, dessen Fehlerterm nicht mit den Fehlern anderer Indikatoren korreliert ist und ERG ausschließlich vom Faktor Glaukomschweregrad beeinflusst wird. Die Variabilität des Glaukomschweregrads wird über ERG definiert und somit ist der Glaukomschweregrad gleichsam das ERG. Die Pfadkoeffizienten, die vom so definierten Glaukomschweregrad zu den restlichen sieben Indikatoren führen, stellen dar, was jeweils an diesen sieben Indikatoren durch das ERG vorhergesagt wird. *Indem man zeigt, dass ERG der einzige Indikator ist, der keinen Methodenanteil hat, wird der Faktor Glaukomschweregrad zum von systematischen Messfehlern bereinigten ERG-Faktor.*

Das Referenzmodell zur Messung der Glaukomdiagnose und alle seine Variationen sind als Modelle mit jeweils einem einzigen Trait-Faktor (allgemeiner Glaukomschweregrad) angelegt. In dem vorliegenden Modell ist aufgrund der oben dargelegten Struktur der zweite Faktor klar als Methodenfaktor und nicht als zweiter Trait-Faktor ausgewiesen. Grundsätzlich gilt jedoch, dass es eine vom forschungsleitenden Interesse abhängige Definitionssache des Forschers ist, welcher Kategorie Kovariation (von Fehlern) zugeordnet wird, der substanziellen, also Trait-Faktor-Variation oder Methodenfaktor-Variation.[174] Die in 10.5.3 dargestellten Variationen entstehen durch die Hinzunahme von Methodenfaktoren, die nach den in 10.3.2.2.1 angestellten Überlegungen plausibel erscheinen. Neben der dort bereits erörterten Tatsache, dass jeweils zwei Diagnoseverfahren messtechnisch aus derselben Quelle abgeleitet wurden, spricht für die Etablierung eines zweiten Methodenfaktors „Visuell evoziertes Potential“ auch das Resultat einer ergänzend durchgeführten Hauptkomponentenanalyse (siehe Einzelergebnis d) in folgender Erläuterung. Bei einer Analyse mit drei extrahierten Komponenten[175] fallen nach der rotierten Varimax-Lösung 39 % der Varianz auf die erste Komponente, 18 % auf die zweite und 15 % auf die dritte. Folgende Tabelle 40 gibt Diagnoseverfahren mit Ladungen größer 0,4 wieder.

---

[174] Conway (1998) zitiert einen unbekannten Reviewer: „The increased covariation among uniqueness could be due to a method variable or a substantive variable not included in the model.“ Er bemerkt im Anschluss, „However it is up to the researcher to decide what constitutes method variance versus substantive variance, and the distinction is relative to the researcher's purpose.“ (Note 1 auf S. 221).

[175] Abweichend von der allgemeinen Konvention wurden aus heuristischen Gründen bereits Komponenten mit Eigenvalue größer 0,75 extrahiert; die Eigenvalues der Komponenten lauten 4,06, 0,91 und 0,78.

Tabelle 40: Hauptkomponentenanalyse über 8 Diagnoseverfahren mit angezeigten Faktorladungen > 0,4
Rotated Component Matrix[a]

	Component		
	1	2	3
md	,853		
dyn	,851		
clv	,818		
fli	,696		
nrr	,592	,433	
erg		,737	
vep		,732	,463
amp			,885

Es zeigt sich, dass

a) prioritär die psychophysischen Verfahren und nachrangig NRR der ersten Komponente zugeordnet werden können,

b) die verbleibenden Verfahren sich zwei weiteren Komponenten zuordnen lassen,

c) wobei die zweite Komponente sich hierbei durch gleich große Ladungen von ERG und VEP und eine schwächere von NRR auszeichnet,

d) und die dritte Komponente einzig die Verfahren VEP und AMP auf sich vereinigt.

10.5.2 RELIABILITÄT, RELIABILITÄTSKOEFFIZIENT ALPHA UND WEITERE KOEFFIZIENTEN DER CTT

Vorbemerkung: Die in der Literatur genannten Koeffizienten Konsistenz-, Reliabilitäts- und Methodenspezifitätskoeffizient beziehen sich konventionell jeweils auf einzelne Items [Eid et al. (2006, 2011)]. In Zusammenhang mit der hier vorliegenden Arbeit ist die Verwendung des Reliabilitätskoeffizienten im Sinne von Yang et al. (2011) nützlich, die die Begrifflichkeit auf Gruppen von Items erweitern und sich auf das von McDonald (1978) eingeführten *coefficient omega* ω beziehen – von Revelle & Zinnbar (2009) in ω_T umbenannt.[176]

Ein Softwareprogramm zu SEM, EQS, hat den entsprechenden Koeffizienten *reliability coefficient rho* implementiert; das Analogon in MPLUS und SAS muss jedoch, orientiert an Raykov, T. (1997a, 2002), „per Hand" berechnet werden: Es werden zur Berechnung in dem von

[176] zitiert nach Yang et al. (2011), S. 384: „With the linear Model the reliability coefficient is a ratio of a scale's estimated true score variance to the scale's estimated total score variance." Sie betonen: „These estimated variances are a function of the fitted estimated linear model." Die Literaturrecherche ergibt, dass McDonalds Omega schon 1970 entwickelt worden ist.

Yang et al. zur Verfügung gestellten Programmcode[177] zwei sog. „Phantomfaktoren", die den Scale-True-Score und den Scale-Total-Score (Skalengesamtwert) wiedergeben, definiert.:

Der Phantomfaktor für den Scale-True-Score berechnet sich wie folgt:

a) Die Ladungen eines jeden Faktors werden summiert.

b) Die summierten Ladungen eines jeden Faktors werden mit dem Faktor multipliziert.

c) Die in (b) erhaltenen Produkte werden summiert.

Der Phantomfaktor für den Skalengesamtwert besteht aus der Summe der beobachteten Items.

Zwei weitere wichtige Koeffizienten lassen sich daraus berechnen: In unidimensionalen Modellen, wie dem hier vorliegenden, gleicht der Anteil der Varianz des allgemeinen Faktors (Trait Faktor) an der True-Score-Varianz dem Konsistenzkoeffizienten, auch *konvergente Validität*[178] genannt. Der Anteil der Varianz des Methodenfaktors (oder der Methodenfaktoren) an der True-Score-Varianz wird der *Methodenspezifitätskoeffizient* genannt[179]. Die Summe beider Koeffizienten beträgt 1. Im Sonderfall des Correlated-Trait-, (CT-)Models, das weder korrelierte Fehler, noch einen oder mehrere Methodenfaktoren, sondern nur (unsystematische) Messfehler enthält, ist der Methodenspezifitätskoeffizient gleich Null und daher die Konsistenz gleich der Reliabilität.

Die Programmcodes nach Yang et al. (2012) zur Errechnung der o. g. Varianzanteile im Grundmodell und seiner vier Variationen finden sich im Anhang 11.7.1. Alle Varianzanteile dieser fünf Modelle lassen sich Tabelle 41 entnehmen[180]:

10.5.3 Bewertung des Referenzmodells und seiner Variationen durch Kenngrössen der CTT

Tabelle 41: Variationen des Modells zur Glaukomdiagnose mit acht Indikatoren – Varianzanteile

---

[177] http://myweb.fsu.edu/yyang3/yang.html

[178] Eid et al. (2006), S.289: „The consistency coefficient is the degree of true variance of an observed variable that is explained by the respective trait factor"; Yang et al. (2011) erläutern anhand eines Beispiels: „The percentage of estimated common variance (i. e. variance due to all factors), accounted for by the general factor were (...) percent for Models (...), indicating a moderately strong contribution of the general factor". „Variance due to all factors" entspricht hierbei exakt „true variance" und „general factor" korrespondiert mit „respective trait factor".

[179] Eid a. a. O.: „The method specificity coefficient indicates the degree of true score variance that is explained by the respective method factor."

[180] Wenn die Fehlerkorrelationen zwischen VEP und AMP durch den Methodenfaktor visuell evoziertes Potential substituiert wird, dann ist der Constraint „Fehlerkorrelation (AMP, VEP) = 0" impliziert und daher nicht in der Tabelle vermerkt.

				Methodenvarianz				
Modell	Total - Score - Varianz	True - Score- Varianz	Trait Faktor- Varianz (Schwere- grad)	Konzen- tration	Visuell evoziertes Potential (VEP, AMP)	Automati- sche Peri- meterie (MD, CLV)	Total	Veränderung v. Modellre- striktionen
Referenz	31,44	26,86	23,15	3,70	-	-	3,70	-
Variation 1	31,44	26,88	23,22	3,66	-	-	3,66	Fehlerkorr. (VEP, MD) =0
Variation 2	31,44	27,58	23,15	3,70	0,73	-	4,43	-
Variation 3	31,44	27,70	23,22	3,66	0,82	-	4,48	Fehlerkorr. (VEP, MD) =0
Variation 4	31,50	28,31	23,23	3,59	1,30	0,20	5,09	-

Es fällt auf, dass die Implementierung des zweiten Methodenfaktors „Automatische Perimetrie“, der seinerseits keine bedeutende Varianzquelle darstellt, nahezu zu einer Verdopplung des auf den ersten Methodenfaktor „Visuell evoziertes Potential“ entfallenden Varianzanteils führt.

Die *Schätzung der Methodenvarianz berücksichtigt nicht die Varianz, die laut Conway* (siehe 10.3.2.5.1; Fußnote 148) *aus den Fehlerkorrelationen geschätzt werden kann*. Dennoch ist es von Interesse, den Methodenvarianzanteil der – je nach Modellvariation – vier oder fünf Fehlerkorrelationen zu schätzen. Zum einen möchte man eine Vorstellung von deren Größenordnung besitzen: Von der Höhe des verborgenen Methodenvarianzanteils hängt die Relevanz der Frage ab, ob Anstrengungen lohnen, Fehlerkorrelationen durch Methodenfaktoren zu substituieren. Zum anderen möchte man den Varianz-Impact der einzigen negativen Fehlerkorrelation zwischen FLI und CLV erfassen. Zu diesem Zweck wurden alle in der Modellvariation 4 zu schätzenden fünf Fehlerkorrelationen jeweils durch einen Methodenfaktor ersetzt. Die True-Score-Varianz des resultierenden Modells beträgt 28,69 und die Reliabilität 0,91. Die Varianzen aller fünf zusätzlichen Methodenfaktoren summieren sich auf 28,69 - 28,31 = 0,38. Diese Varianzsumme macht nur 0,38 / (28,69 - 23,23) = 7 Prozent der maximal schätzbaren Methodenvarianz aus. Da 93 % der maximal schätzbaren Methodenvarianz durch entsprechende Faktoren erfasst werden, fällt die entgangene Methodenvarianz wenig ins Gewicht. Es fällt auf, dass hierbei die Fehlerkorrelation mit dem geringsten Methodenvarianzanteil (0,05) jene ist, die im Referenzmodell betragsmäßig den höchsten Wert aufweist (Fehler von FLI und CLV, -0,32). Dies besagt, dass in diesem Einzelfall einer wenn auch relativ starken negativen Fehlerkorrelation ein besonders schwacher Varianzanteil zuzurechnen ist.

In Tabelle 42 wird der Reliabilitätskoeffizient *Cronbachs Alpha* wiedergegeben, der auf Basis der acht z-standardisierten Items mit der Software SPSS ermittelt wurde. Sämtliche angeführte Koeffizienten ergeben sich aus den in Tabelle 41 präsentierten Varianzanteilen.

Tabelle 42 Modellvariationen zur Glaukomdiagnose – (Cronbachs) Alpha und Koeffizienten der klassischen Testtheorie

Modell	Cronbachs Alpha	Reliabilität	Konvergente Validität (Konsistenzkoeffizient)	Methodenspezifität	Unreliabilität
Referenz	0,85	0,85	0,86	0,14	0,15
Variation 1		0,85	0,86	0,14	0,15
Variation 2		0,88	0,84	0,16	0,12
Variation 3		0,88	0,84	0,16	0,12
Variation 4		0,90	0,82	0,18	0,10

Ergebnisse: Cronbachs Alpha nimmt für das Referenzmodell und dessen erste Variation denselben Wert an wie die Reliabilität, unterschätzt aber die Reliabilität aller weiteren Modelle. Das Hinzunehmen eines Methodenfaktors (in Variation 2 und 4) führt jeweils zur Erhöhung der Unterschätzung. Je mehr Methodenfaktoren ins Modell implementiert werden, desto mehr steigt – bei gleichbleibender Trait-Faktor-Varianz – die Gesamtvarianz der Methodenfaktor(en), so dass auch die True-Score-Varianz zunimmt. Diese Zunahme führt zu einer Erhöhung der Reliabilität und Abnahme der konvergenten Validität. Das Produkt aus Reliabilität und konvergenter Validität, also der Anteil der Trait-Faktor-Varianz an der Gesamtvarianz, hat in Modellvariation 4 den gleichen Wert wie im Ausgangsmodell (0,74). Die Zuverlässigkeit des eigentlich zu messenden Merkmals „Glaukomschweregrad“ bleibt mithin konstant.

Die Pfadkoeffizienten des Trait-Faktors sind (bis auf minimale, unbedeutende Schwankungen) in den vier Variationen konstant, so dass der auf den Diskriminationsparametern beruhende Patientenindex zur Ermittlung des Glaukomschweregrades von den Modellvariationen unberührt bleibt.

10.5.3.1 Interpretation der Koeffizienten auf Basis der Erkenntnisse Conways (1998)

Bei Conway wurde die Übereinstimmung zwischen a) dem Wert der quadrierten Pfadkoeffizienten eines Methodenfaktors in einem Modell und b) dem durchschnittlichen Wert der entsprechenden Fehlerkorrelationen im Vergleichsmodell für den Fall überprüft, dass entweder nur Methodenfaktoren (a) oder nur Fehlerkorrelationen (b) in die zu vergleichenden Modelle implementiert waren. Im Fall des Modells zur Diagnose des Glaukomschweregrades beinhaltet aber die Gruppe der konzentrationssensiblen Indikatoren sowohl Methodenfaktoren als auch Fehlerkorrelationen, so dass der Varianzanteil der Fehlerkorrelationen nicht analog zu Conway abgeschätzt werden kann.

Dennoch ist angesichts der von Conway angeführten grundsätzlichen Möglichkeit, Fehlerkorrelationen in die Berechnung der Reliabilität einzubeziehen, bei der Interpretation der Koeffizienten zu beachten: Wenn man die in den Fehlerkorrelationen zum Ausdruck kommende method variance berücksichtigte, würde bei konstanter Trait-Faktor-Varianz die True-Score-Varianz ansteigen und daher der Konsistenzkoeffizient sinken. Daher gilt: *Der* geschätzte *Konsistenzkoeffizient ist aufgrund* des *Methodenvarianzpotentials der Fehlerkorrelationen eine Obergrenze der wahren konvergenten Validität.* Aus demselben Grund gilt: *Die geschätzten Methodenspezifitäts- und Reliabilitätskoeffizienten sind Untergrenzen der wahren Methodenspezifität bzw. Reliabilität.*

10.5.4 SCHÄTZT ALPHA DIE RELIABILITÄT VERZERRUNGSFREI?

Gegenstand der Diskussion sind zunächst Effekte, die aus Verletzungen von Modellannahmen resultieren und den Reliabilitätsschätzer Cronbachs Alpha (kurz: alpha) im Falle des publizierten Modells zur Glaukomdiagnose beeinflussen. Diesem wird die auf Basis des gefitteten Modells errechnete wahre Reliabilität (kurz: Reliabilität) gegenübergestellt. Die aus Verletzungen von Modellannahmen resultierenden Effekte mit Auswirkung auf den Modellfit werden im Anschluss behandelt.

These: Die Gleichheit von alpha und Reliabilität in Referenzmodell und Modellvariation 1 bedeutet nicht, dass alpha unverzerrt ist. Sie rührt im Gegenteil daher, dass die zu Über- und Unterschätzung führenden Verzerrungsmomente in diesem Fall gerade einander aufheben. Dieser Egalisierungseffekt ist aus der Literatur (Komaroff, 1997) bekannt und wurde S. 249 bereits erörtert.

Das die Reliabilität *unterschätzende* Moment geht auf die Verletzung der Annahme essentieller $\tau -$ Äquivalenz zurück. Diese Annahme ist in unserem Fall doppelt verletzt, wegen der Kongenerizität der Daten (1) und der Multidimensionalität des Konstrukts (2). Es handelt sich bei den acht diagnostischen Verfahren um kongenerische Variable, da sich die Diskriminationsparameter $\boldsymbol{\lambda_i}$ der True-Score-Variablen unterscheiden, $\tau_i = \alpha_i + \boldsymbol{\lambda_i} \cdot \eta$, so dass Var($\boldsymbol{\tau}_i$) $\neq$ Var($\boldsymbol{\tau}_j$) und Var(ε_i) $\neq$ Var(ε_j). Das Ausmaß der Multidimensionalität der Daten kann man dadurch leicht zum Vorschein bringen, in dem man das Modell mit Methodenfaktor (siehe 10.5.1) umformuliert, so dass es die Form des gewöhnlichen zweidimensionalen Modells hat (Vgl. Eid et al. (2011) 11.7.2.1, Abbildung 93)[181]: Zum einen werden die fünf zusätzlich

[181] Eid et al. (2011), S. 851.

vom ersten und wichtigsten Methodenfaktor „Konzentration“ beeinflussten Indikatoren im umformulierten Modell ausschließlich durch diesen (und die Fehlerterme) beeinflusst. Damit verliert der Faktor Konzentration seinen Status als Methodenfaktor und wird zum zweiten Trait-Faktor. Zum anderen wird die Restriktion, dass der Trait-Faktor „Glaukomschweregrad“ und der Faktor „Konzentration“ unkorreliert seien, aufgehoben.[182] Im Zuge der Reformulierung wird die Fehlerkorrelationsstruktur unwesentlich verändert – lediglich die Korrelation zwischen den Fehlern von NRR und DYN nimmt um ca. 20 % auf 0,16 ab. Die Korrelation zwischen den Faktoren „Glaukomschweregrad“ und „Aspekte psychophysischer Messung des Glaukoms (Konzentration)“ wird auf 0,86 geschätzt. Osburn (2000) äußert sich zur Auswirkung einer Korrelation von Faktoren auf die Verzerrung des Reliabilitätskoeffizienten alpha (siehe S. 219) und bezeichnet in seiner Simulationsstudie eine True-Score-Korrelation zwischen den Komponenten 1 bis 4 und den Komponenten 5 bis 8 in Höhe von 0,8 als schwache Heterogenität. Das Simulationsmodell unterscheidet sich vom umformulierten Modell nur unwesentlich dadurch, dass der erste Faktor nur Komponente (Indikator) 1 bis 3 und der zweite Faktor Komponente 4 bis 8 beeinflusst. Das umformulierte zweidimensionale Modell zur Schätzung des Glaukomschweregrades ist aufgrund der geschätzten True-Score-Korrelation zwischen den beiden Faktoren von 0,86 durch schwache bis sehr schwache Heterogenität gekennzeichnet. Es ist daher davon auszugehen, dass aufgrund von sehr geringer bis geringer Multidimensionalität die wahre Gesamtreliabilität durch alpha um wenige Prozent unterschätzt wird (Tabelle 24: Eight Component Data).

Den beiden zur Unterschätzung von alpha führenden Momenten steht ein *zur Überschätzung führendes Moment* gegenüber, das auf die mit einer Ausnahme positiv korrelierten Fehler zurückgeht. Diese betreffen zwar einen Großteil der Items (im Referenzmodell sieben und in Modellvariation 1 sechs von insgesamt acht Items) aber sind sehr schwach (die in Modellvariation 1 wegfallende Fehlerkorrelation in Höhe von 0,05) bis schwach (zwei Fehlerkorrelationen in Höhe von 0,2). Daher kann die überschätzende Wirkung durch die positiv korrelierten Fehler am besten durch die Studie von Zimmerman (1993) (siehe S. 232) angenähert werden, der in Tabelle 29 aufzeigt, dass bei wahrer Reliabilität von 0,8 und sechs von acht fehlerkorrelierten Items mit einer Korrelation von 0,1 eine Überschätzung von 0,86 resultiere (Szenario

[182] Diese Umformulierung erscheint nur dann unplausibel, wenn man sich auf den Namen „Konzentration“ festlegt, da es prima facie keinen vernünftigen Grund gibt, warum Konzentration, ein starken temporären Schwankungen unterworfener Zustand, und der Glaukomschweregrad, eine als permanent angenommener irreversibler Zustand, korreliert sein könnten (siehe 260). So betrachtet ist eine Korrelation zwischen den beiden hinsichtlich ihres Varianzanteils wichtigsten Faktoren des Modells plausibel.

1), während beim sonst selben Setting mit Fehlerkorrelation 0,3 eine Überschätzung von 0,92 zu verzeichnen sei (Szenario 2). Ignoriert man die negative, betragsmäßig größte Fehlerkorrelation zwischen FLI und CLV, da keine der einbezogenen Studien Auskunft darüber gibt, auf welche Weise negative Fehlerkorrelationen die Schätzung von alpha beeinflusst, so kann man Folgendes schließen: In den beiden o. g. Szenarien von Zimmermann wird wahre Reliabilität um 7,5 bis 15 % überschätzt. Unser Referenzmodell wird durch das erste Szenario wesentlich besser angenähert als durch das zweite, da der Durchschnitt der vier positiven Fehlerkorrelationen 0,14 beträgt, so dass die wahre Reliabilität von 0,85 infolge der Fehlerkorrelationen um mehr als 7,5 % aber deutlich unter 15 % überschätzt werden dürfte. Sollte die negative Fehlerkorrelation die verzerrenden Effekte der positiven Fehlerkorrelationen abmildern, wie die Klassifikation von Heene et al. (Tabelle 36) vermuten lässt[183], so dürfte die überschätzende Wirkung der korrelierten Fehler deutlich unter 7,5 % und somit minimal sein. Die Klärung dieser Frage muss weiterer Forschung vorbehalten bleiben.

Die beiden einander entgegenlaufenden Effekte der Über- und Unterschätzung heben sich gerade auf, da wir davon ausgehen, dass eine Überschätzung von deutlich unter 7,5 % und eine Unterschätzung von ca. 2 % gleichzeitig zur Wirkung kommen.

10.5.5 BEWERTUNG DURCH FIT-INDIZES

Der relative Fitindex CFI und der Closeness to Fit-Koeffizient RMSEA betragen in allen Modellen 1,000 bzw. 0,000 und zeigen das hohe Niveau der Anpassungsgüte der Modellreihe auf.

Es wurden eine Reihe absoluter Fit-Indizes (χ^2 , RMR, CAIC, ECVI) und – aufgrund der Modellindifferenz des RMSEA – ein spezielles Maß (RMSEA, obere 90 %-Vertrauensgrenze) zur differenzierenden Beurteilung der Anpassungsgüte ausgewählt (Tabelle 43). Die ersten beiden und der letzte dienen der Überprüfung der Gesamtanpassung und die informationstheoretischen Maße CAIC und ECVI, die auch die Modellsparsamkeit berücksichtigen, sind für Modellvergleiche nützlich. Vergleiche auf Basis des Chi-Quadrat-Differenzentests sind, obwohl die Modelle ineinander verschachtelt sind, wegen teilweise nicht erfüllter Voraussetzungen[184] unstatthaft.

---

[183] Der Durchschnitt der fünf Fehlerkorrelationen des Referenzmodells, 0,2; 0,2; 0,09; 0,05 und -0,32, beträgt 0,04.

[184] Parameter des unrestringierten Modells wurden auf Grenzwerte (Null) gesetzt. Siehe hierzu auch Eid et al. (2011), S. 883.

Tabelle 43: Modellvariationen zur Glaukomdiagnose – Fit und Fitindizes

	Fit und Fit-Indizes					
	χ^2	p	RMR	RMSEA obere 90 % Vertrauensgrenze	CAIC	ECVI
Referenz	$7{,}96(df = 10)$	0,63	0,0131	0,042	-63,66	0,129
Variation 1	$8{,}35(df = 11)$	0,68	0,0131	0,038	-70,43	0,125
Variation 2	$7{,}95(df = 9)$	0,54	0,0131	0,048	-56,50	0,133
Variation 3	$8{,}35(df = 10)$	0,59	0,0131	0,044	-63,26	0,130
Variation 4	$6{,}91(df = 7)$	0,44	0,0125	0,056	-43,22	0,140

Aus den Ergebnissen soll nur das sparsamste Modell (Variation 1) und das am wenigsten sparsame (Variation 4) herausgegriffen werden: Variation 1 ist das lediglich um eine Fehlerkorrelation zwischen VEP und MD verringerte Referenzmodell und ist hinsichtlich aller Fitkriterien (außer Root Mean Square Residual, RMR) überlegen. (Allerdings ist dessen Reliabilität suboptimal.) Variation 4 übertrifft Variation 1 im verbesserten Chi-Quadrat-Maß, wenn die Anzahl der Freiheitsgrade unberücksichtigt bleibt, und im RMR; in allen anderen Maßen steht es hinter den Konkurrenzmodellen zurück. Der Reliabilitätskoeffizient hingegen wird von keinem anderen Modell erreicht.

Fazit zu Fit-Indizes des variierten Referenzmodells: Wenn die Anzahl der zu schätzenden Parameter als Gütekriterium nicht den Ausschlag geben sollen, dann kann man das Modell Variation 4 als bestes ansehen, da die Konstruktion seiner Methodenfaktoren einer inhaltlichen Logik folgt: Die Methodenfaktoren sind gerechtfertigt, da die jeweils durch sie beeinflussten Paare von Diagnoseverfahren (AMP und VEP bzw. MD und CLV) aus derselben gemeinsamen Datenquelle „Visuell evoziertes Potential“ bzw. „Automatische Perimetrie“ abgeleitet sind. Der grundsätzliche Einwand Eids, dass bei Modellvergleichen anhand von Chi-Quadrat-Werten durch Lockerung von Restriktionen in einem Modell der Chi-Quadrat-Wert immer verbessert werden könne, die damit einhergehende Parameterfreisetzung jedoch inhaltlich möglicherweise nicht zu rechtfertigen sei[185], trifft daher hier nicht zu. Eine Überspezifikation scheint durch Implementierung dieser zwei zusätzlichen Methodenfaktoren nicht gegeben.

10.5.5.1 Die Modellentwicklung im Hinblick auf Erkenntnisse von Heene et al. (2012)

Von den vier simulierten Fehlerkorrelationsstrukturen kommt die Struktur „Weak Model Deviation“ (sechs positive und negative Korrelationen) der Fehlerkorrelationsstruktur des Mo-

[185] Eid et al. (2011) S. 883.

dells zur Schätzung des allgemeinen Glaukomschweregrades am nächsten. Jedoch kann der Anteil der Items mit Fehlerkorrelationen (sieben von acht bzw. 87,5 %) und das Verhältnis von positiven zu negativen Fehlerkorrelationen (vier zu eins) im Referenzmodell durch keine simulierte Struktur angenähert werden. Denn der maximale Anteil der fehlerkorrelierten Items beträgt in Heenes Modell B 12 von 24 bzw. 50 %.[186] Grundsätzlich unterscheiden sich das zum gewöhnlichen zweidimensionalen Modell umformulierte Referenzmodell zur Schätzung des allgemeinen Glaukomschweregrades und die beiden Modelle von Heene et al. darin, dass die Faktoren in ersterem stark (0,86) und in letzteren schwach assoziiert sind (r = 0,3), sowie darin, dass in ersterem nur drei von fünf Fehlerkorrelationen Items unterschiedlicher Faktoren zugehören und in letzteren nur Fehlerkorrelationen zwischen Items unterschiedlicher Faktoren auftreten.[187]

Zur *Sensitivität* der Fit-Indizes: Es kann nach Heene et al. bei allen o. g. Verschiedenheiten vorsichtig vermutet werden, dass der Chi-Quadrat-Modelltest im Modell zur Schätzung des allgemeinen Glaukomschweregrades auf die Fehlerkorrelationsstruktur mindestens mit einer leichten Erhöhung reagiert hätte, wenn die Annahme der Unkorreliertheit für alle Fehlerpaare aufrechterhalten worden wäre. Denn zwar sind die positiven Fehlerkorrelationen des Glaukomschweregradmodells von geringerer Höhe als die der simulierten Korrelationsstrukturen „Moderate und Strong Model Deviation“ und entsprechen nur jenen der beiden schwachen Modellabweichungen, aber der Anteil der fehlerkorrelierten Items ist im Glaukomschweregradmodell weitaus höher.

Andere Fit-Indizes wurden zur Modelloptimierung im Rahmen der Publikation 2001 nicht hinzugezogen, so dass deren mangelnde Sensitivität im Hinblick auf Fehlerkorrelationen sich bei der Suche nach dem angemessenen Modell auch nicht negativ auswirken konnte.

Die von Heene et al. vorgeschlagene Prozedur des Expected Parameter Change (EPC) auf Basis der Modifikationsindices (MIs) ist im Rahmen der Modellentwicklung 2001 nicht softwaregesteuert angewendet worden, zumal die Software JRULE erst 2009 veröffentlich wurde. Es wurden bei der Modellentwicklung lediglich MIs verwendet. Allerdings ist davon auszugehen, dass die Veränderung relevanter Parameter im Zuge der Aufhebung der Unkorreliertheitsannahme für einzelne Fehlerpaare (auf Basis der MIs) beobachtet wurde. Eine aktuelle Revision des Referenzmodells bestätigt, dass alle relevanten Fehlspezifikationen (falsche Un-

[186] Von den sechs Fehlerkorrelationen sind insgesamt zwölf Items betroffen. Im Modell zum Glaukomschweregrad sind im Unterschied hierzu einzelne Items von mehr als einer Fehlerkorrelation betroffen.
[187] Über die Bedeutung der Unterschiede gibt Kapitel 4 Aufschluss.

korreliertheitsannahmen) eliminiert worden sind und daher Fitindices (nach erfolgter Aufhebung der Unabhängigkeitsannahme für fünf korrelierte Fehlerpaare) keiner spezifischen Verzerrung unterliegen dürften. Der Powerverlust im Simulationsbeispiel (schwächste Fehlerkorrelation, geringste Fallzahl (n=150)) trat im Modellentwicklungsprozess zur Schätzung des allgemeinen Glaukomschweregrades vermutlich nicht ein, da die Fallzahl (n=2 mal 237) in etwa mit der Fallzahl 500 in der Simulation von Heene et al. gleichzusetzen ist, die vom Powerverlust nicht betroffen war.

10.6 SCHLUSSBEMERKUNGEN UND AUSBLICK

10.6.1 MODELLANNAHMEN UND SCHÄTZMETHODE

Diese Arbeit stellt die Fortsetzung einer publizierten Analyse dar, die mit der Maximum-Likelihood-Schätzung operierte. Die ML-Methode besitzt unter der Bedingung multivariat normalverteilter Daten optimale asymptotische Eigenschaften: Normalität, Unverzerrtheit und Effizienz der Schätzer.[188] In Wahrung der Kontinuität mit der publizierten Vorarbeit wurde die Schätzmethode beibehalten.

Korreliertheit der Fehler

Laut Rubio und Gillespie (1995) sollten Daten, die korrelierte Fehler enthalten, nicht mit der ML-Methode geschätzt werden, sondern mit der General-Least-Square (GLS)-Methode (S. 372). Wenn die Annahme der Normalverteilung für einige oder alle Variablen nicht erfüllt ist, wird die Weighted-Least-Square- (WLS)-Methode unter Verweis auf eine eigene Studie (Alsup & Gillespie, 1994) empfohlen. Größere Fehlerkorrelationen wurden schon im Rahmen der Vorarbeit weitgehend eliminiert, geringere ($r < 0,1$) jedoch nicht ausgeschlossen. In Anbetracht des sehr geringen Wertes der betreffenden Korrelationskoeffizienten scheint die Verwendung von ML-Estimation vertretbar.

Nichtnormalität

Die Überprüfung der Annahme der univariaten Normalverteilung aller acht Variablen erfolgt getrennt nach Augenseite und ergibt in der grafischen Analyse, dass sämtliche Parameter, außer CLV (zweigipflig, Abbildung 95f. im Anhang), eingipflig sind und alle eine Schiefe im Bereich zwischen -1 und +1 aufweisen. Die Schiefe von fünf Variablen ist bzgl. beider Augenseiten positiv im Bereich zwischen 0,13 und 0,62 (Tabelle 47 im Anhang).

[188] Kaplan (2000), S. 80.

Die Überprüfung wird ergänzt durch entsprechende Tests nach Shapirpo-Wilk W und Tests der multivariaten Normalverteilung nach den in SAS PROC MODEL verfügbaren Verfahren Mardia Skewness, Mardia Kurtosis sowie nach Henze-Zirkler T (Tabelle 48 dito)[189]. Die Verteilungsprüfungen werden mittels der Software STATA durchgeführt, die statt des Shapirpo-Wilk W den sogenannten sktest einsetzt (Tabelle 49 dito). Die Ergebnisse der unvariaten Verteilungsprüfungen in beiden Softwareprodukten teilweise erheblich voneinander ab; die Ergebnisse der Signifikanztests auf multivariate Normverteilung hingegen stimmen ausnahmslos überein. Die auf der rechten Augenseite erhobenen Werte entsprechen deutlich schlechter der Normalverteilungsannahme, als die auf der linken Augenseite erhobenen. Nur unter den Messwerten der linken Augenseite ist nach dem Kriterium Mardia Kurtosis die Annahme der multivariaten Normalverteilung nicht zu verwerfen (p=0,0840).

Kaplan (2000) legt den (mehrfach in der Literatur angeführten) Sachverhalt dar, dass Standardfehler unter Verletzung der Normalitätsannahme – verglichen mit der empirischen Standardabweichung der (Parameter-)schätzungen – unterschätzt würden. Was die Anpassungsgüte des Modells betrifft, so führe Nichtnormalität zu einer Überschätzung der Likelihood-Ratio-Chi-Quadrat-Statistiken, welche wiederum verknüpft sei mit der Anzahl der Freiheitsgrade des zu testenden Modells. Er empfiehlt in diesem Fall die Anwendung der WLS-Schätzmethode[190]. Die Resultate der Studien hierzu seien uneinheitlich; dennoch seien in allen Fällen das ADF-Chi-Quadrat kleiner als jenes unter Anwendung der ML-Methode.

Anhand der Analyse der Glaukomdaten kann in einem allerersten Schritt der von Kaplan oben berichtete Sachverhalt über die SE nicht nachvollzogen werden: Die bei Anwendung der ML-Methode hervorgebrachten modellbasierten SE liegen im Durchschnitt geringfügig (7 %) über den empirischen (via Bootstrapanalyse ermittelten) SE. Die bei Anwendung des ADF-Algorithmus modellbasiert geschätzten Standardfehler sind um ca. 10 % geringer als die empirischen SE (Tabelle 44). Diese erste Untersuchung gibt, obgleich sie nur auf der Analyse einer einzigen Datendatei beruht, schon einen fundierten Hinweis darauf, dass die Verletzung der Annahme der multivariaten Normalverteilung nicht gravierend sein kann. Denn schwer-

[189] Die Hypothesen in den Testverfahren lauten nach der Information im SAS-Manual stets, dass die *Residuen* normalverteilt seien. Der Vergleich mit den Prüfungen auf multivariate Normalverteilung mittels der Software STATA, in der die beobachteten Werte getestet werden, bringt jedoch bis auf die 4 Nachkommastelle identische Werte hervor.

[190] Laut PROC CALIS-Manual ist die Schätzmethode WLS im Falle einer nicht spezifizierten Weight-Matrix äquivalent mit der von Browne's (1984, 1982) asymptotisch verteilungsfreien Schätzmethode (ADF).

wiegende Verletzung würde die durch die Literatur gedeckte Unterschätzung der SE nach sich ziehen.[191]

Die Ausführungen Kaplans zur Testung der Modellgüte lassen sich in einem ersten Schritt anhand der Glaukomdaten nachvollziehen:

Tabelle 44: Über- und Unterschätzung der empirischen Standardfehler durch modellbasierte SE nach ML- bzw. WLS-Methode

Ver-fah-ren	*modellbasiert geschätzte SE*		*empirischer SE*	*Quotienten: modellbas. gesch. SE / empir. SE*	
	Ausgangsdatensatz, ML-Estimation	Ausgangsdatensatz, WLS-Estimation	Bootstrap,unadjustiert, ML-Estimation	Ausgangsd.(ML)/ Bootstrap (ML)	Ausgangsd.(WLS)/ Bootstrap (ML)
nrra	0,049	0,043	0,048	1,017	0,894
fli	0,055	0,046	0,051	1,071	0,892
vep	0,050	0,046	0,049	1,020	0,933
erg	0,049	0,035	0,038	1,287	0,908
md	0,051	0,045	0,052	0,975	0,873
amp	0,059	0,054	0,057	1,040	0,954
dyn	0,056	0,044	0,051	1,104	0,871
clv	0,051	0,040	0,047	1,081	0,843
Durchschnittl. Abweichung von Bootstrapanalyse				**1,074**	**0,896**

Der Chi-Quadrat-Wert des globalen Modelltests unter Anwendung von ADF (χ^2 =6,51; df= 10; p= 0,77) ist deutlich geringer als jener unter Anwendung von ML (χ^2 =7,96; df=10; p=0,63).

Eine in welche Richtung auch immer *verzerrte Schätzung der modellbasierten Standardfehler infolge der Anwendung der ML-Methode hat keinen Einfluss auf die zentrale Forschungsfrage*, die Ermittlung von geeigneten Korrekturfaktoren für die SE von nicht nach Paarigkeit adjustierten Analysen. Sowohl die nach Paarigkeit der Daten korrigierte als auch die nicht korrigierte Analyse wurde mittels ML-Schätzung durchgeführt, wobei jeweils der *empirische Standardfehler* via Bootstrapping ermittelt wurde. Der empirische SE, die Standardabweichung der Parameterschätzungen, ist nicht von dem Überschätzungseffekt betroffen. Die Ermittlung der Korrekturfaktoren und die Hypothesentestung bei der Parameterschätzung wurden auf Basis des unverzerrten empirischen SE durchgeführt.

[191] In dieser Arbeit wurden im Rahmen der Bootstrapanalysen simulierter Daten stets die Mittelwerte des modellbasiert geschätzten Standardfehlers mit dem empirischen Standardfehler verglichen. Dabei wurden mindestens 1000 Bootstrapsamples zugrundegelegt. Im Rahmen der Bootstrapanalysen empirischer Patientendaten wurden die modellbasierten SE nicht untersucht, sondern nur die empirischen SE betrachtet. An dieser Stelle werden in einer ersten Untersuchung lediglich die modellbasierten SE der Pfadkoeffizienten des Glaukomschweregrads in der Ausgangsdatendatei betrachtet. Diese Untersuchung bringt dennoch Erkenntnisgewinn, da die *Standardabweichung der geschätzten SE* erfahrungsgemäß minimal (häufig unter 0,001) ist und somit eine Analyse der geschätzten SE auf Basis nur einer Datendatei eine grobe, aber recht zuverlässige Schätzung liefert. Siehe Fußnote 43.

Ausblick: Eine gründlichere Untersuchung der modellbasiert geschätzten SE auf Basis von Bootstrapanalysen anhand von Patientendaten (analog den durchgeführten Analysen simulierter Daten) würde genaueren Aufschluss darüber geben, ob die beobachtete Verletzung der Annahme multivariat normalverteilter Patientendaten sich tatsächlich in einer in der Literatur behaupteten Unterschätzung der SE. niederschlagen würde. Wenn nicht, so könnte der Schluss gezogen werden, dass das Ausmaß der Verletzung in Hinblick auf den SE-Effekt als geringfügig zu betrachten ist.

Eid et al. gehen davon aus, dass schon bei moderater Abweichung von der multivariaten Normalverteilung, insbesondere bei positiver Schiefe, der Fehler erster Art höher ist, als man aufgrund der Chi-Quadrat-Verteilung erwarten würde. Dies führe dazu, dass korrekte Modelle eher verworfen würden.[192] Neuere Literatur (Finney und Di Stefano, 2006) schränkt Kaplans Empfehlung der WLS-Schätzmethode (s. o.) ein, da diese erst ab einer Stichprobengröße von n=5000 akzeptable Fehlerraten hervorbringe. Die Anwendung der auf verteilungsfreien Methoden basierenden Chi-Quadrat-Tests scheint zusätzlich zur o. g. tendenziellen Verwerfung korrekter Modelle auch dazu zu führen, dass falsch spezifizierte Modelle nicht detektiert würden. Daher sei besser der Satorra-Bentler-korrigierte Chi-Quadrat-Test anzuwenden (Satorra & Bentler, 1994).

Ausblick: Es wird daher vorgeschlagen, in nachfolgenden Forschungen sich einer Software zu bedienen, welche die Modellgüte mit Hilfe der (in der verwendeten PROC CALIS-Version nicht verfügbaren) Satorra-Bentler-korrigierten Chi-Quadrat-Teststatistik untersucht.

Bootstrapping

Ein Ergebnis der Bootstrapanalyse bestand darin, dass zwar erwartungsgemäß die *nach Paarigkeit adjustierte Analyse* zu höheren Standardfehlern führt, jedoch keine ebenso erwartete Verringerung der Chi-Quadrat-Statistik hervorbringt (9, Gesamtfazit). Eid et al. erläutern hierzu, dass der naive Bootstrap zwar in einer angemessenen Schätzung der Standardfehler resultiere, aber nicht in probaten Ergebnissen bei der Modelltestung, da „die Stichprobenkovarianzmatrix nicht perfekt dem Modell entspricht und man daher aus nicht modellkonformen Daten Stichproben zieht“. Bollen und Stine (1992) gemäß müssten hierzu die Daten vor

[192] Eid et al. (2011), S. 881.

Durchführung von Bootstrapanalysen erst so transformiert werden, dass sie dem Modell entsprechen[193].

Ausblick: Die Tatsache, dass der „naive", nach Augenseitenkorrelation (= nach Paarigkeit der Daten) adjustierte Bootstrap erhöhte SE hervorbringt, bestätigt eine zentrale Forschungshypothese dieser Doktorarbeit. Die „andere der Seite der Medaille", der Effekt verringerter Chi-Quadrat-Werte des globalen Modelltests, stellte sich aus nunmehr geklärten Gründen nicht ein. Daher ist die Wiederholung der Bootstrapanalysen auf Basis von Daten, die der Bollen-Stine-Transformation unterzogen wurden, die zentrale zukünftige Forschungsaufgabe.

Sonstiges

Weitere Forschungsaktivität könnte auf die systematische Klärung der Frage gerichtet werden, ob die Höhe der Pfadkoeffizienten des allgemeinen Glaukomschweregrades einen Einfluss auf die Höhe der Korrekturfaktoren haben, die zur angemessenen Schätzung der Pfadkoeffizienten SE ermittelt wurden. Die vorläufige Antwort, gestützt auf wenige Beobachtungen, lautet: Je größer – ceteris paribus – ein Pfadkoeffizient des Schweregrades, desto geringer der angemessene Korrekturfaktor.

Die Methode zur Adjustierung der SE interessierender Pfadkoeffizienten lässt sich auch auf Daten mit komplexeren Clusterstrukturen übertragen (z. B. Messungen von Patienten mit einem und mit zwei erkrankten Augen oder – allgemeiner formuliert – einem oder zwei erkrankten paarigen Organen).

10.6.2 Analyse der Patientendaten und Forschungsliteratur

Die Forschungsliteratur bestätigt ein wichtiges, aus der Analyse der Patientendaten resultierendes Faktum: In Messmodellen mit korrelierten (positiven)[194] Fehlertermen, in welchen die Fehlerkorrelation nicht angemessen modelliert wird, wird der Einfluss des allgemeinen (Trait-) Faktors überschätzt. Mit anderen Worten: Die Reliabilität des Messinstruments, wird überschätzt[195], was beispielweise in überhöhten Werten für Coefficient alpha resultiert. Es ist gelungen, Fehlerkorrelationen im publizierten Modell durch Methodenfaktoren zu substituieren und dadurch den Anteil der True-Score-Varianz an der Gesamtvarianz, mit anderen Wor-

---

[193] Genauere Darstellung des Bollen-Stine-Bootstrap, der in häufig in SEM-Software vorgesehen sei, findet sich ebenda.

194 Der umgekehrte Fall, die Unterschätzung der Reliabilität infolge fehlspezifizierter negativer Fehlerkorrelationen, wird als Ausnahmeerscheinung in der Literatur nicht beschrieben.

195 In den Abschnitten des Kapitels 3.5 ist, korrespondierend mit dem Begriff der Reliabilität, von der „Güte des Messmodells" die Rede.

ten die wahre Reliabilität auf 90 % zu steigern, womit die systematische Fehlervarianz weitestgehend erfasst ist. (Dass die theoretische Möglichkeit der Schätzung der restlichen True-score Varianz im Rahmen eines CU-Modells in unserem Fall praktisch nicht umsetzbar war, fällt daher kaum ins Gewicht.) Die Anwendung eines 2012 veröffentlichten Programmcodes ermöglichte die Schätzung der wahren Reliabilität des so verbesserten publizierten Modells zur Schätzung des allgemeinen Glaukomschweregrades. Im ursprünglichen, publizierten Modell heben die Effekte, die zu Überschätzung der Reliabilität (Fehlerkorreliertheit) führen und die Effekte, die für Unterschätzung verantwortlich sind (Kongenerizität), einander gerade auf.

11 ANHANG

11.1 SPSS-PROGRAMMCODE ZUR KÜNSTLICHEN ERZEUGUNG VON DATEN FÜR AMOS-MODELLE MIT EINEM FAKTOR UND DREI INDIKATOREN

11.1.1 ZUSAMMENSTELLUNG DER WICHTIGSTEN BEFEHLE

```
COMPUTE y1_s = (x1 + x0) / SQRT(2).
COMPUTE y2_s = (x2 + x0) / SQRT(2).
COMPUTE y2b_s = (x2 - x0) / SQRT(2).
COMPUTE y3_s = (x3 + x0) / SQRT(2).
COMPUTE indik_1 = 1 * x4 + x1.
COMPUTE indik_2 = 1 * x4 + x2.
COMPUTE indik_3 = 1 * x4 + x3.
COMPUTE indi_1ps = x4 + y1_s.
COMPUTE indi_2ps = x4 + y2_s.
COMPUTE indi_2ns = x4 + y2b_s.
COMPUTE indi_3ps = x4 + y3_s.
```

11.1.2 AUSFÜHRLICHE SYNTAX

Der folgende Befehl zur Ladung einer Datendatei namens US Sozialerhebung dient lediglich dazu, eine Matrix mit einer ausreichenden Anzahl von Zeilen für die zu erzeugenden Zufallszahlen anlegen zu können. Die Datendatei hat inhaltlich also keinerlei Belang und ist jederzeit gegen eine andere austauschbar, sofern sie eine ähnlich hohe Zeilenanzahl aufweist.

```
*Aufruf einer Dummydatei mit 1517 Fällen unter Beibehaltung einer einzigen Variablen zum Zweck der Erzeugung reproduzierbarer, approximativ standardnormalverteilter Zufallszahlen.
GET
  FILE='C:\Programme\SPSS\1991 US Sozialerhebung.sav' /keep geschl.
execute.
set seed 20000.
compute w1 = normal(1).
compute w2 = normal(1).
compute w3 = normal(1).
compute w4 = normal(1).
compute w5 = normal(1).
execute.

FACTOR
  /VARIABLES w1 w2 w3 w4 w5
  /MISSING LISTWISE
  /ANALYSIS w1 w2 w3 w4 w5
  /PRINT INITIAL EXTRACTION
  /CRITERIA FACTORS(5) ITERATE(25)
  /EXTRACTION PC
  /ROTATION NOROTATE
  /SAVE REG(ALL)
   /METHOD=CORRELATION .
VARIABLE LABELS
w1 'erste simulierter standardnormalverteilte Variable'
w2 'zweite simulierte standardnormalverteilte Variable'
w3 'dritte simulierte standardnormalverteilte Variable'
w4 'vierte simulierte standardnormalverteilte Variable'
w5 'fünfte simulierte standardnormalverteilte Variable'
fac1_1 'erster simulierter orthogonaler standardnormalverteilter Faktor'
fac2_1 'zweiter simulierter orthogonaler standardnormalverteilter Faktor'
fac3_1 'dritter simulierter orthogonaler standardnormalverteilter Faktor'
fac4_1 'vierter simulierter orthogonaler standardnormalverteilter Faktor'
fac5_1 'fünfter simulierter orthogonaler standardnormalverteilter Faktor'.
EXE.
```

```
*Umbenennung der Faktoren.
RENAME VARIALES
(fac1_1 fac2_1 fac3_1 fac4_1 fac5_1= x0 x1 x2 x3 x4).
EXE.
*Errechnung der 3 manifesten Variablen aus den 4 latenten Variablen.
*Dazu Regressionssgewichte stets gleich 1.

*GRUNDMODELL ohne korrelierte Fehler:
VARIABLE LABELS
x0 'Variable zur Konstruktion der pos. und neg. korrelierten Fehlerterme y... (=fac1_1)'
x1 'error von indik_1 (=fac2_1)'
x2 'error von indik_2 (=fac3_1)'
x3 'error von indik_3 (=fac4_1)'
x4 'latente Variable (=fac5_1)'.
EXE.
COMPUTE indik_1 = 1 * x4 + x1.
COMPUTE indik_2 = 1 * x4 + x2.
COMPUTE indik_3 = 1 * x4 + x3.
EXE.
VARIABLE LABELS
indik_1 'erste manifeste Variable, die über die Fehlerterme nicht korreliert ist'
indik_2 'zweite manifeste Variable, die über die Fehlerterme nicht korreliert ist'
indik_3 'dritte manifeste Variable, die über die Fehlerterme nicht korreliert ist'.
EXE.

*Erzeugung der positiv oder negativ korrelierten Fehlerterme für alle weiteren Modelle
*A
COMPUTE y1 = (x1 + x0).
COMPUTE y2 = (x2 + x0).
COMPUTE y2b = (x2 - x0).
COMPUTE y3 = (x3 + x0).

*B
*Varianz der pos. oder neg. korrelierten Fehlerterme auf 1 setzen: Standardisierung
COMPUTE y1_s = y1 / SQRT(2).
COMPUTE y2_s = y2 / SQRT(2).
COMPUTE y2b_s = y2b / SQRT(2).
COMPUTE y3_s = y3 / SQRT(2).
EXE.
VARIABLE LABELS
y1_s 'error von indi_1ps mit Var. 1'
y2_s 'error von indi_2ps mit Var. 1'
y2b_s 'error von indi_2ns mit Var. 1'
y3_s 'error von indi_3ps mit Var. 1'.
EXE.

* MODELL MIT 2 POS. KORRELIERTEN FEHLERVARIABLEN
COMPUTE indi_1ps = x4 + y1_s.
COMPUTE indi_2ps = x4 + y2_s.
EXE.
VARIABLE LABELS
indi_1ps 'erste manif. Variable, die mit zweiter über die standard. Fehler pos. korreliert ist'
indi_2ps 'zweite manif. Variable, die mit erster über die standard. Fehler pos. korreliert ist'.
EXE.
*****

* MODELL MIT 2 NEG. KORRELIERTEN FEHLERVARIABLEN
COMPUTE indi_2ns = x4 + y2b_s.
VARIABLE LABELS
indi_2ns 'zweite manif. Variable, die mit erster über die standard. Fehler neg. korreliert ist '.
EXE.
*******
* MODELL MIT 3 POS. KORRELIERTEN FEHLERVARIABLEN
COMPUTE indi_3ps = x4 + y3_s.
VARIABLE LABELS
indi_3ps 'dritte manif. Variable, die mit den anderen über die standard. Fehler pos. korreliert ist'.
EXE.
```

11.2 SPSS-PROGRAMMCODE ZUR KÜNSTLICHEN ERZEUGUNG VON DATEN FÜR AMOS-MODELLE MIT ZWEI FAKTOREN UND SECHS INDIKATOREN

```
GET
  FILE='C:\Programme\SPSS\1991 US Sozialerhebung.sav' /keep geschl.
execute.
set seed 20000.
compute w1 = normal(1).
compute w2 = normal(1).
compute w3 = normal(1).
compute w4 = normal(1).
compute w5 = normal(1).
compute w6 = normal(1).
compute w7 = normal(1).
compute w8 = normal(1).
compute w9 = normal(1).
compute w10 = normal(1).
compute w11 = normal(1).
execute.

*Durchführung einer Hauptkomponentenanalyse zur Erzeugung orthogonaler Faktoren
FACTOR
  /VARIABLES w1 w2 w3 w4 w5 w6 w7 w8 w9 w10 w11
  /MISSING LISTWISE
  /ANALYSIS w1 w2 w3 w4 w5 w6 w7 w8 w9 w10 w11
  /PRINT INITIAL EXTRACTION
  /CRITERIA FACTORS(11) ITERATE(25)
  /EXTRACTION PC
  /ROTATION NOROTATE
  /SAVE REG(ALL)
   /METHOD=CORRELATION .

*Umbenennung der Faktoren.
RENAME VARIABLES
(fac1_1 fac2_1 fac3_1 fac4_1 fac5_1 fac6_1 fac7_1 fac8_1 fac9_1 fac10_1 fac11_1 = x0 x1 x2 x3 x4 x5 x6 x7 x8 x9 x10).
EXE.
VARIABLE LABELS
x0 'nullter simulierter orthogonaler standardnormalverteilter Faktor'
x1 'erster simulierter orthogonaler standardnormalverteilter Faktor'
x2 'zweiter simulierter orthogonaler standardnormalverteilter Faktor'
x3 'dritter simulierter orthogonaler standardnormalverteilter Faktor'
x4 'vierter simulierter orthogonaler standardnormalverteilter Faktor'
x5 'fünfter simulierter orthogonaler standardnormalverteilter Faktor'
x6 'sechster simulierter orthogonaler standardnormalverteilter Faktor'
x7 'siebenter simulierter orthogonaler standardnormalverteilter Faktor'
x8 'achter simulierter orthogonaler standardnormalverteilter Faktor'
x9 'neunter simulierter orthogonaler standardnormalverteilter Faktor'
x10 'zehnter simulierter orthogonaler standardnormalverteilter Faktor'.

*Strukturgleichungsmodelle, deren Faktoren statistisch unabhängig sind
*******************************************************************************
* Berechnung der Modellvariablen
*Modell 1:
COMPUTE indi_1 = x1 + x3.
COMPUTE indi_2 = x1 + x4.
COMPUTE indi_3 = x1 + x5.
COMPUTE indi_4 = x2 + x6.
COMPUTE indi_5 = x2 + x7.
COMPUTE indi_6 = x2 + x8.
EXE.

*Modell Z 1:
COMPUTE ind_1sp2 = x1 + (x3 + x9) /SQRT(2).
COMPUTE ind_2sp2 = x1 + (x4 + x9) /SQRT(2).
COMPUTE indi_3 = x1 + x5.
COMPUTE indi_4 = x2 + x6.
COMPUTE indi_5 = x2 + x7.
COMPUTE indi_6 = x2 + x8.
EXE.

*Modelle Z 2 und Z 3:
```

```
COMPUTE ind_1sp2 = x1 + (x3 + x9) /SQRT(2).
COMPUTE ind_2sp2 = x1 + (x4 + x9) /SQRT(2).
COMPUTE indi_3 = x1 + x5.
COMPUTE ind_4sp3 = x2 + (x6 + x10)/SQRT(2).
COMPUTE ind_5sp3 = x2 + (x7 + x10)/SQRT(2).
COMPUTE indi_6 = x2 + x8.

*Modell 2:
COMPUTE ind_1sp5 = x1 + (x3 + x9) /SQRT(2).
COMPUTE indi_2 = x1 + x4.
COMPUTE indi_3 = x1 + x5.
COMPUTE ind_4sp5 = x2 + (x6 + x9) /SQRT(2).
COMPUTE indi_5 = x2 + x7.
COMPUTE indi_6 = x2 + x8.
EXE.

*Modell 3:
COMPUTE ind_1sp6 = x1 + (x3 + x9 + x10) /SQRT(3).
COMPUTE ind_2sp6 = x1 + (x4 + 1.73*x10) /SQRT(4).
COMPUTE indi_3 = x1 + x5.
COMPUTE ind_4sp6 = x2 + (x6 + 1.73*x9) /SQRT(4).
COMPUTE indi_5 = x2 + x7.
COMPUTE indi_6 = x2 + x8.
EXE.

*Strukturgleichungsmodelle, deren Faktoren statistisch abhängig sind
*********************************************************************************
*Strukturmodell
COMPUTE y1_ = x0 + x1.
COMPUTE y1s_= y1_ / SQRT(2).

*Modell 4
COMPUTE indi_1 = x1 + x3.
COMPUTE indi_2 = x1 + x4.
COMPUTE indi_3 = x1 + x5.
COMPUTE indi_4s_ = y1s_ + x6.
COMPUTE indi_5s_ = y1s_ + x7.
COMPUTE indi_6s_ = y1s_ + x8.
EXE.

* Modell 5
COMPUTE ind_1sp2 = x1 + (x3 + x9) /SQRT(2).
COMPUTE ind_2sp2 = x1 + (x4 + x9) /SQRT(2).
COMPUTE indi_3 = x1 + x5.
COMPUTE indi_4s_ = y1s_ + x6.
COMPUTE indi_5s_ = y1s_ + x7.
COMPUTE indi_6s_ = y1s_ + x8.
EXE.

*Modelle 6 und 7
COMPUTE ind_1sp2 = x1 + (x3 + x9) /SQRT(2).
COMPUTE ind_2sp2 = x1 + (x4 + x9) /SQRT(2).
COMPUTE indi_3 = x1 + x5.
COMPUTE ind_4sp3 = y1s_+ (x6 + x10)/SQRT(2).
COMPUTE ind_5sp3 = y1s_+ (x7 + x10)/SQRT(2).
COMPUTE indi_6s_ = y1s_ + x8.
EXE.

*Modell 8
COMPUTE ind_1sp2 = x1 + (x3 + x9) /SQRT(2).
COMPUTE indi_2 = x1 + x4.
COMPUTE indi_3 = x1 + x5.
COMPUTE ind_4sp5 = y1s_ + (x6 + x9)/SQRT(2).
COMPUTE indi_5s_ = y1s_ + x7.
COMPUTE indi_6s_ = y1s_ + x8.
EXE.

*Modell 9
COMPUTE ind_1sp6 = x1 + (x3 + x9 + x10) /SQRT(3).
COMPUTE ind_2sp6 = x1 + (x4 + 1.73*x10) /SQRT(4).
COMPUTE indi_3 = x1 + x5.
COMPUTE ind_4sp6 = y1s_ + (x6 + 1.73*x9) /SQRT(4).
COMPUTE indi_5s_ = y1s_ + x7.
COMPUTE indi_6s_ = y1s_ + x8.
EXE.
```

```
*Modell Z 4
COMPUTE indi_1 = x1 + x3.
COMPUTE indi_2 = x1 + x4.
COMPUTE indi_3 = x1 + x5.
COMPUTE ind_4sp3 = y1s_+ (x6 + x10)/SQRT(2).
COMPUTE ind_5sp3 = y1s_+ (x7 + x10)/SQRT(2).
COMPUTE indi_6s_ = y1s_ + x8.
EXE.
```

11.3 PROC CALIS LINEQS-MODELLE MIT EINEM FAKTOR UND DREI INDIKATOREN

Die Anhänge 11.3 bis 11.7 beziehen sich stets auf die SAS Prozedur CALIS, in der u. a. Strukturgleichungsmodelle des Typs Lineqs durchgeführt werden. Daher wird die Bezeichnung PROC CALIS Lineqs in diesen Kapiteln auf Lineqs verkürzt.

11.3.1 SAS-PROGRAMMCODE

11.3.1.1 Vollständiger kommentierter Programmcode des Grundmodells

```
*1.;
*Es werden 1000 mal 5 standardnormalverteilte Zufallsvariable mit n=1000 Realisationen erzeugt;
%MACRO machdat5 (i, seedx0, seedx1, seedx2, seedx3, seedx4, dateinam=);
%GLOBAL sds;
%LET sds=&dateinam;
DATA &sds;
Do i = 1 to &i;
x0 = Rannor (&seedx0);
x1 = Rannor (&seedx1);
x2 = Rannor (&seedx2);
x3 = Rannor (&seedx3);
x4 = Rannor (&seedx4);
OUTPUT;
END;
RUN;
%MEND machdat5;
%macro create9(howmany);
   %do i=1 %to &howmany;
      %machdat5 (1000, &i, &i, &i, &i, &i, dateinam=iterat&i);
       run;
   %end;
%mend create9;
%create9(1000)
*2.;
*Diese Zufallsvariablen werden weiterverarbeitet, indem 3 manifeste Variablen aus ihnen konstruiert werden,
* entsprechend dem in AMOS gebildeten Grundmodell;
%Macro arit_op2 (dateinam=);
%GLOBAL sds;
%LET sds=&dateinam;
DATA b_&sds;
SET &sds;
 indik_1 = 1 * x4 + x1;
 indik_2 = 1 * x4 + x2;
 indik_3 = 1 * x4 + x3;
RUN;
%MEND arit_op2;

%macro create7(howmany);
   %do i=1 %to &howmany;
      %arit_op2 (dateinam=iterat&i);
       run;
  %end;
%mend create7;
%create7(1000)
*3.;
*Es werden Korrelationsmatrizen aus den erzeugten Rohdatensätzen gebildet;
%Macro corrmatr (dateinam=);
%GLOBAL sds;
%LET sds=&dateinam;
PROC CORR DATA=&sds noprint outp=cm_&sds;
var indik_1 indik_2 indik_3;
TITLE;
   run;
%MEND corrmatr;
%macro create8(howmany);
   %do i=1 %to &howmany;
      %corrmatr (dateinam=b_iterat&i);
       run;
   %end;
%mend create8;
%create8(1000)
*4.;
```

```
*Erzeugung von (1000) OUTEST-Datasets bei den (1000) Durchläufen von PROC *CALIS / LINEQS auf Basis der oben
*erzeugten (1000) Korrelationsmatrizen;
%Macro proccalis (dateinam=);
%GLOBAL sds;
%LET sds=&dateinam;
proc calis data=&sds OUTEST=OUTE_&sds tech=nr edf=999 nop;
         Lineqs
                  indik_1 =beta1    Faktor_1 + beta_f1Error_1,
                  indik_2 =beta2    Faktor_1 + beta_f2Error_2,
                  indik_3 =beta3    Faktor_1 + beta_f3Error_3;

         Std
                  Error_1 = 1,
                  Error_2 = 1,
                  Error_3 = 1,
                  Faktor_1= 1;
         run;
%MEND proccalis;
%macro create10(howmany);
   %do i=1 %to &howmany;
      %proccalis (dateinam=cm_b_iterat&i);
       run;
   %end;
%mend create10;
%create10(1000)
*5.;
*Von den (1000) OUTEST Sas-data-sets wird nur jeweils die interessierende *Zeile (Beobachtung) mit dem _TYPE_ "PARMS"
(=final Parameter Estimates) *bewahrt;
%Macro parms_isol (dateinam=);
%GLOBAL sds;
%LET sds=&dateinam;
DATA PARMS_&sds;
SET &sds;
IF _TYPE_ NE "PARMS" THEN DELETE;
RUN;
TITLE "PARMS von LINEQS-Modell mit 1 Faktor und unkorrelierten Fehlern &sds";
run;
%MEND parms_isol;
%macro create11(howmany);
   %do i=1 %to &howmany;
      %parms_isol (dateinam=OUTE_CM_B_ITERAT&i);
       run;
   %end;
%mend create11;
%create11(1000)
*6.;
*Die (1000) einzeiligen Datensätze addieren;
%macro create12(howmany);
  %do i=2 %to &howmany;
   proc APPEND BASE=PARMS_OUTE_CM_B_ITERAT1
           DATA=PARMS_OUTE_CM_B_ITERAT&i;
      run;
   %end;
%mend create12;
%create12(1000)
*7.;
*Umbenennen der Datei;
DATA PARMS_ALL_ITERAT;
SET PARMS_OUTE_CM_B_ITERAT1;
RUN;
8.;
*Die negativen Werte von beta werden in positive umgewandelt;
*Schließlich werden statistische Maßzahlen und Histogramme für die inte-
*ressierenden Pfadkoeffizienten erstellt sowie Pfadkoeffizientenkorrela-
*tionen berechnet;
DATA PARMS_ALL_ITERAT_B;
SET PARMS_ALL_ITERAT;
beta1_b=ABS(beta1);
beta2_b=ABS(beta2);
beta3_b=ABS(beta3);
DROP beta1 beta2 beta3;
LABEL beta1_b='Betrag von beta1'
           beta2_b='Betrag von beta2'
           beta3_b='Betrag von beta3';
TITLE "(Betrag von) PARMS aus LINEQS MODEL mit 1 Faktor und unkorrel. Fehlern (1000 Iterationen)";
RUN;
PROC MEANS DATA=PARMS_ALL_ITERAT_B;
TITLE "(Betrag von)PARMS aus LINEQS MODEL mit 1 Faktor und unkorrel. Fehlern (1000 Iterationen)";
RUN;
PROC GCHART DATA=PARMS_ALL_ITERAT_B;
VBAR beta1_b beta2_b beta3_b;
TITLE "Betrag von PARMS aus LINEQS MODEL mit 1 Faktor und unkorrel. Fehlern (1000 Iterationen)";
RUN;
```

```
PROC CORR DATA=PARMS_ALL_ITERAT_B NOSIMPLE;
VAR beta1_b beta2_b beta3_b;

TITLE "Korrelationsmatrix der PARMS aus LINEQS MODEL mit 1 Faktor und unkorrel. Fehlern (1000 Iterationen)";
RUN;
*9.;
*Die geschätzten Standardfehler der 1000 PROC CALIS OUTEST-Outputdateien *isolieren;
%Macro se_isol (dateinam=);
%GLOBAL sds;
%LET sds=&dateinam;
DATA Stderr_&sds;
SET &sds;
IF _TYPE_ NE "STDERR" THEN DELETE;
RUN;
TITLE "STDERR von LINEQS-Modell mit 1 Faktor und unkorrelierten Fehlern &sds";
run;
%MEND se_isol;
%macro create13(howmany);
   %do i=1 %to &howmany;
      %se_isol (dateinam=OUTE_CM_B_ITERAT&i);
       run;
   %end;
%mend create13;
%create13(1000)
*10.;
*Die (1000) isolierten einzeiligen Datensätze addieren;
%macro create14(howmany);
   %do i=2 %to &howmany;
   proc APPEND BASE=STDERR_OUTE_CM_B_ITERAT1
            DATA=STDERR_OUTE_CM_B_ITERAT&i;
      run;
   %end;
%mend create14;
%create14(1000)
*11.;
*Umbenennen der Datei;
DATA STDERR_ALL_ITERAT;
SET STDERR_OUTE_CM_B_ITERAT1;
RUN;
*12.;
*statistische Maßzahlen für den SE ausgeben;
PROC MEANS DATA=STDERR_ALL_ITERAT;
TITLE "STDERR aus LINEQS MODEL mit 1 Faktor und unkorrel. Fehlern (1000 Iterationen)";
RUN;
```

11.3.1.2 Programmcodeausschnitt der anderen Modelle

Im Folgenden werden für jedes Modell nur jeweils die Gleichungen zur Konstruktion der Indikatoren aufgeführt, die innerhalb des Makros „arit_op2“ unter *2.[196] zu finden und mit jenen der AMOS-Modelle identisch sind. Anschließend erfolgten jeweils das Makro „proccalis“ mit allen notwendigen Statements. Das Statement COV ist in den Modellvarianten der korrekten und fehlerhaften Spezifikation wiedergegeben.

11.3.1.2.1 Modell mit zwei positiv korrelierten Fehlern

```
 y1 = (x1 + x0);
 y2 = (x2 + x0);
 y1_s = y1 / SQRT(2);
 y2_s = y2 / SQRT(2);
 indi_1ps = x4 + y1_s;
 indi_2ps = x4 + y2_s;
 indik_3 = 1 * x4 + x3;
%Macro proccalis (dateinam=);
%GLOBAL sds;
%LET sds=&dateinam;
proc calis data=&sds OUTEST=OUTE_&sds edf=999 nop;
Lineqs
                  indi_1ps =                beta1    Faktor_1          + beta_f1          Error_1,
```

[196] Selbstverständlich variieren entsprechend der Variablenkonstruktion in „arit_op2“ auch die Variablenbezeichnungen bei der Erzeugung der Korrelationsmatrizen unter Schritt 3 des Programmcodes.

```
                indi_2ps =                beta2    Faktor_1          + beta_f2          Error_2,
                indik_3 =         beta3    Faktor_1          + beta_f3          Error_3;


        Std
                Error_1 = 1,
                Error_2 = 1,
                Error_3 = 1,
                Faktor_1= 1;
*Korrekt spezifizierte Variante;
        Cov
                Error_1 Error_2 = 0.5;
*fehlspezifizierte Variante;
        Cov
                Error_1 Error_2 = 0;
        run;
%MEND proccalis;
```

11.3.1.2.2 Modell mit zwei negativ korrelierten Fehlern

```
 y1 = (x1*1.528 + x0);
 y2b = (x2*1.528 - x0);
*Vorfaktor 1.528 erhöht das Gewicht des jeweiligen Fehlerterms x1 und x2 *(entspricht Reduzierung des Einflusses des
*gemeinsamen Fehlerterms x0), so daß die Fehlerkorrelation auf r=-0,3 absinkt.
 y1_s = y1 / 1.82614;
 y2b_s = y2b / 1.82614;
*1.82614 ist die Standardabweichung von y1 bzw. y2b;
 indi_1ps = x4 + y1_s;
 indi_2ns = x4 + y2b_s;
 indik_3 = 1 * x4 + x3;
%Macro proccalis (dateinam=);
%GLOBAL sds;
%LET sds=&dateinam;
proc calis data=&sds OUTEST=OUTE_&sds tech=nr edf=999 nop;
Lineqs
                indi_1ps =                beta1    Faktor_1          + beta_f1          Error_1,
                indi_2ns =                beta2    Faktor_1          + beta_f2          Error_2,
                indik_3 =         beta3    Faktor_1          + beta_f3          Error_3;

        Std
                Error_1 = 1,
                Error_2 = 1,
                Error_3 = 1,
                Faktor_1= 1;
*Korrekt spezifizierte Variante;
        Cov
                Error_1 Error_2 = -0.3;
*fehlspezifizierte Variante;
        Cov
                Error_1 Error_2 = 0;
        run;
%MEND proccalis;
```

11.3.1.2.3 Modell mit drei positiv korrelierten Fehlern

```
 y1 = (x1 + x0);
 y2 = (x2 + x0);
 y3 = (x3 + x0);
 y1_s = y1 / SQRT(2);
 y2_s = y2 / SQRT(2);
 y3_s = y3 / SQRT(2);
 indi_1ps = x4 + y1_s;
 indi_2ps = x4 + y2_s;
 indi_3ps = x4 + y3_s;

%Macro proccalis (dateinam=);
%GLOBAL sds;
%LET sds=&dateinam;
proc calis data=&sds OUTEST=OUTE_&sds tech=nr edf=999 nop;
Lineqs
                indi_1ps =                beta1    Faktor_1          + beta_f1          Error_1,
                indi_2ps =                beta2    Faktor_1          + beta_f2          Error_2,
                indi_3ps =                beta3    Faktor_1          + beta_f3          Error_3;

        Std
                Error_1 = 1,
                Error_2 = 1,
                Error_3 = 1,
                Faktor_1= 1;
*Korrekt spezifizierte Variante;
```

```
        Cov
Error_1 Error_2 = 0.5;
Error_1 Error_3 = 0.5;
Error_2 Error_3 = 0.5;


*fehlspezifizierte Variante;
        Cov
Error_1 Error_2 = 0;
Error_1 Error_3 = 0;
Error_2 Error_3 = 0;
        run;
%MEND proccalis;
```

11.3.2 SAS-OUTPUT

Die Mediane der geschätzten Standardfehler der Pfadkoeffizienten wurden über PROC UNIVARIATE erzeugt, deren Output wegen seines großen Umfangs hier nicht wiedergegeben wird.

Der Standard ist n = 1000 Realisationen von Zufallsvariablen und 1000 Iterationen.

11.3.2.1 Modell mit einem Faktor und drei unkorrelierten Fehlern; n=1000, 1000 Durchläufe

(Betrag von)PARMS aus LINEQS MODEL mit 1 Faktor und unkorrel. Fehlern (1000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		1000	2.408008E-12	3.97297E-12	0	2.117709E-11
ITER	Iteration	0	.	.	.	.
beta_f1	beta_f1	1000	0.7066814	0.0243907	0.6105454	0.7834463
beta_f2	beta_f2	1000	0.7060015	0.0259436	0.6250233	0.7763183
beta_f3	beta_f3	1000	0.7066783	0.0262148	0.6103369	0.7896106
beta1_b	Betrag von beta1	1000	0.7066912	0.0244067	0.6214595	0.7919812
beta2_b	Betrag von beta2	1000	0.7072658	0.0257996	0.6303411	0.7806061
beta3_b	Betrag von beta3	1000	0.7065645	0.0262078	0.6136083	0.7921419

STDERR aus LINEQS MODEL mit 1 Faktor und unkorrel. Fehlern (1000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta1	beta1	1000	0.0336124	0.000781834	0.0316459	0.0369522
beta_f1	beta_f1	1000	0.0251363	0.0011860	0.0223452	0.0300587
beta2	beta2	1000	0.0336162	0.000785290	0.0316454	0.0367903
beta_f2	beta_f2	1000	0.0251845	0.0013079	0.0225018	0.0312678
beta3	beta3	1000	0.0336078	0.000771251	0.0316459	0.0369405
beta_f3	beta_f3	1000	0.0251375	0.0012024	0.0224750	0.0303832

Pearsonsche Korrelationskoeffizienten, N = 1000

Prob > |r| under H0: Rho=0

	beta1_b	beta2_b	beta3_b
beta1_b Betrag von beta1			
beta2_b Betrag von beta2	-0.13091 <.0001		
beta3_b Betrag von beta3	-0.04131 0.1918	-0.14549 <.0001	

11.3.2.2 Modell mit einem Faktor und zwei positiv korrelierten Fehlern n=1000, 1000 Durchläufe

11.3.2.2.1 A korrekt spezifiziert

(Betrag von)PARMS aus LINEQS MODEL mit 1 Faktor und 2 pos. korr. korr. spez. Fehlern (1000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		1000	5.9268E-13	2.724477E-12	0	1.892652E-11
ITER	Iteration	0	.	.	.	.
beta_f1	beta_f1	1000	0.7059867	0.0231546	0.6217091	0.7747406
beta_f2	beta_f2	1000	0.7053106	0.0242564	0.6258195	0.7877899
beta_f3	beta_f3	1000	0.7063250	0.0304269	0.5834961	0.8028577
beta1_b	Betrag von beta1	1000	0.7074706	0.0230889	0.6322792	0.7832483
beta2_b	Betrag von beta2	1000	0.7080708	0.0241960	0.6159440	0.7799679
beta3_b	Betrag von beta3	1000	0.7065874	0.0302556	0.5961710	0.8121160

STDERR aus LINEQS MODEL mit 1 Faktor und 2 pos. korr. korr. spez. Fehlern (1000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta1	beta1	1000	0.0340737	0.0010327	0.0313732	0.0388323
beta_f1	beta_f1	1000	0.0218025	0.0013052	0.0187645	0.0270128
beta2	beta2	1000	0.0340904	0.0010124	0.0313594	0.0376152
beta_f2	beta_f2	1000	0.0218661	0.0013734	0.0186767	0.0293402
beta3	beta3	1000	0.0376324	0.0013289	0.0342350	0.0430242
beta_f3	beta_f3	1000	0.0304278	0.0026711	0.0253056	0.0432161

Korrelationsmatrix der PARMS aus LINEQS MODEL mit 1 Faktor und 2 pos. korr. korr. spez. Fehlern

Die Prozedur CORR

3 Variablen: beta1_b beta2_b beta3_b

Pearsonsche Korrelationskoeffizienten, N = 1000

Prob > |r| under H0: Rho=0

	beta1_b	beta2_b	beta3_b
beta1_b	1.00000	0.32683	-0.26657
Betrag von beta1		<.0001	<.0001
beta2_b	0.32683	1.00000	-0.26806
Betrag von beta2	<.0001		<.0001
beta3_b	-0.26657	-0.26806	1.00000
Betrag von beta3	<.0001	<.0001	

11.3.2.2.2 B fehlspezifiziert

(Betrag von) PARMS aus LINEQS MODEL mit 1 Faktor und 2 pos. korr. fehlspez. Fehlern (1000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		1000	9.708124E-13	3.106399E-12	0	1.799009E-11
ITER	Iteration	0	.	.	.	.
beta_f1	beta_f1	1000	0.4988761	0.0315846	0.3903880	0.5948786
beta_f2	beta_f2	1000	0.4975301	0.0331170	0.3695747	0.5993613
beta_f3	beta_f3	1000	0.8161919	0.0165349	0.7557794	0.8699026
beta1_b	Betrag von beta1	1000	0.8659069	0.0182137	0.8038155	0.9206509
beta2_b	Betrag von beta2	1000	0.8666096	0.0188800	0.8004786	0.9292015
beta3_b	Betrag von beta3	1000	0.5770696	0.0234286	0.4932236	0.6548264

STDERR aus LINEQS MODEL mit 1 Faktor und 2 pos. korr. fehlspez. Fehlern (1000 Iterationen) 6

17:02 Wednesday, June 18, 2008

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta1	beta1	1000	0.0299825	0.000682796	0.0283377	0.0324571
beta_f1	beta_f1	1000	0.0309523	0.0028862	0.0250794	0.0419510
beta2	beta2	1000	0.0299782	0.000689214	0.0282660	0.0325152
beta_f2	beta_f2	1000	0.0310733	0.0030333	0.0244500	0.0465506
beta3	beta3	1000	0.0308981	0.000333446	0.0299923	0.0319495
beta_f3	beta_f3	1000	0.0198478	0.000246781	0.0191524	0.0205644

Korrelationsmatrix der PARMS aus LINEQS MODEL mit 1 Faktor und 2 pos. korr. fehlspez. Fehlern

Die Prozedur CORR

3 Variablen: beta1_b beta2_b beta3_b

Pearsonsche Korrelationskoeffizienten, N = 1000

Prob > |r| under H0: Rho=0

	beta1_b	beta2_b	beta3_b
beta1_b	1.00000	-0.64312	0.03567
Betrag von beta1		<.0001	0.2597
beta2_b	-0.64312	1.00000	0.05381
Betrag von beta2	<.0001		0.0890
beta3_b	0.03567	0.05381	1.00000
Betrag von beta3	0.2597	0.0890	

11.3.2.3 Modell mit einem Faktor und zwei negativ korrelierten Fehlern n=1000, 1000 Durchläufe

11.3.2.3.1 A korrekt spezifiziert

(Betrag von)PARMS aus MODEL mit 1 Faktor und neg. korrel. (r=-0.3) korr. spez. Fehlern (1000 Iterat.)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		1000	2.60021E-13	1.405239E-12	0	1.675383E-11
ITER	Iteration	0	.	.	.	.
beta_f1	beta_f1	1000	0.7063370	0.0244070	0.6267281	0.7839751
beta_f2	beta_f2	1000	0.7065445	0.0253707	0.6241416	0.7797832
beta_f3	beta_f3	1000	0.7069045	0.0230744	0.6266802	0.7761479
beta1_b	Betrag von beta1	1000	0.7070318	0.0244864	0.6207922	0.7792380
beta2_b	Betrag von beta2	1000	0.7067638	0.0252485	0.6260498	0.7813113
beta3_b	Betrag von beta3	1000	0.7065569	0.0230682	0.6305509	0.7792766

STDERR aus MODEL mit 1 Faktor und mit r=-0.3 korr. spez. Fehlern (1000 Iterat.)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta1	beta1	1000	0.0337710	0.000790162	0.0319146	0.0366433
beta_f1	beta_f1	1000	0.0263707	0.0011259	0.0232809	0.0313148
beta2	beta2	1000	0.0337709	0.000799801	0.0317337	0.0364878
beta_f2	beta_f2	1000	0.0263611	0.0011143	0.0235703	0.0318087
beta3	beta3	1000	0.0314377	0.000538096	0.0300798	0.0331979
beta_f3	beta_f3	1000	0.0221062	0.000594316	0.0205517	0.0241784

Korrelationsmatrix der PARMS aus MODEL mit 1 Faktor und neg. korrel.(r=-0.3) korr. spez. Fehl.

Die Prozedur CORR

3 Variablen: beta1_b beta2_b beta3_b

Pearsonsche Korrelationskoeffizienten, N = 1000

Prob > |r| under H0: Rho=0

	beta1_b	beta2_b	beta3_b
beta1_b	1.00000	-0.32109	-0.02289
Betrag von beta1		<.0001	0.4696
beta2_b	-0.32109	1.00000	-0.00405
Betrag von beta2	<.0001		0.8983
beta3_b	-0.02289	-0.00405	1.00000
Betrag von beta3	0.4696	0.8983	

11.3.2.3.2 B fehlspezifiziert

(Betrag von)PARMS aus MODEL mit 1 Faktor und neg. korrel. (r=-0.3) fehlspez. Fehlern (1000 Iterat.)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		1000	3.666729E-13	1.686957E-12	0	1.727157E-11
ITER	Iteration	0	.	.	.	.
beta_f1	beta_f1	1000	0.8056520	0.0202686	0.7394513	0.8707364
beta_f2	beta_f2	1000	0.8057495	0.0211055	0.7374919	0.8694859
beta_f3	beta_f3	1000	0.5308245	0.0520038	0.3401202	0.6659843
beta1_b	Betrag von beta1	1000	0.5913933	0.0277340	0.4917501	0.6732100
beta2_b	Betrag von beta2	1000	0.5911876	0.0286484	0.4939577	0.6753559
beta3_b	Betrag von beta3	1000	0.8452884	0.0318183	0.7459658	0.9403820

STDERR aus MODEL mit 1 Faktor und mit r=-0.3 fehlspez. Fehlern (1000 Iterat.)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta1	beta1	1000	0.0355413	0.000866479	0.0333154	0.0385429
beta_f1	beta_f1	1000	0.0237027	0.000649991	0.0219314	0.0264380
beta2	beta2	1000	0.0355357	0.000857753	0.0334980	0.0390106
beta_f2	beta_f2	1000	0.0236969	0.000642976	0.0221427	0.0270750
beta3	beta3	1000	0.0393121	0.0021946	0.0343023	0.0489972
beta_f3	beta_f3	1000	0.0503052	0.0103008	0.0319718	0.1024109

11.3.2.3.3 Korrelationsmatrix der PARMS aus MODEL mit 1 Faktor und neg. korrel.(r=-0,3) fehlspez. Fehler

Die Prozedur CORR

3 Variablen: beta1_b beta2_b beta3_b

Pearsonsche Korrelationskoeffizienten, N = 1000

Prob > |r| under H0: Rho=0

	beta1_b	beta2_b	beta3_b
beta1_b	1.00000	0.28219	-0.40000
Betrag von beta1		<.0001	<.0001
beta2_b	0.28219	1.00000	-0.39263
Betrag von beta2	<.0001		<.0001
beta3_b	-0.40000	-0.39263	1.00000
Betrag von beta3	<.0001	<.0001	

11.3.2.4 Modell mit einem Faktor und drei positiv korrelierten Fehlern n=1000, 1000 Durchläufe

11.3.2.4.1 A korrekt spezifiziert

(Betrag von)PARMS aus LINEQS MODEL mit 1 Faktor u. 3 pos. korr. korr. spez. Fehlern (1000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		1000	1.797936E-10	5.6544212E-9	0	1.7880939E7
ITER	Iteration	0	.	.	.	.
beta_f1	beta_f1	1000	0.7048309	0.0281556	0.5846060	0.7796581
beta_f2	beta_f2	1000	0.7039527	0.0303445	0.4964767	0.7773060
beta_f3	beta_f3	1000	0.7049266	0.0294998	0.5360107	0.7848326
beta1_b	Betrag von beta1	1000	0.7082781	0.0276481	0.6262055	0.8113178
beta2_b	Betrag von beta2	1000	0.7089952	0.0292816	0.6291227	0.8680498
beta3_b	Betrag von beta3	1000	0.7080814	0.0288244	0.6197078	0.8442121

11.3.2.4.2 STDERR aus LINEQS MODEL mit 1 Faktor 3 pos. korr. korr. spez. Fehlern (1000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta1	beta1	1000	0.0381069	0.0093736	0.0324880	0.3155712
beta_f1	beta_f1	1000	0.0266035	0.0097622	0.0178088	0.2670937
beta2	beta2	1000	0.0458451	0.2458538	0.0329091	7.8108039
beta_f2	beta_f2	1000	0.0404647	0.4311281	0.0187964	13.6564733
beta3	beta3	1000	0.0382390	0.0131738	0.0329551	0.4403072
beta_f3	beta_f3	1000	0.0267813	0.0136445	0.0184136	0.4028597

Korrelationsmatrix der PARMS aus LINEQS MODEL mit 1 Faktor u. 3 pos. korr. korr. spez. Fehlern

Die Prozedur CORR

Pearsonsche Korrelationskoeffizienten, N = 1000

Prob > |r| under H0: Rho=0

	beta1_b	beta2_b	beta3_b
beta1_b	1.00000	-0.04982	0.02468
Betrag von beta1		0.1154	0.4357
beta2_b	-0.04982	1.00000	-0.07575
Betrag von beta2	0.1154		0.0166
beta3_b	0.02468	-0.07575	1.00000
Betrag von beta3	0.4357	0.0166	

11.3.2.4.3 Korrektur: Entfernung der Iteration 796 wegen extremen Ausreißern – SE (beta_f2) = 13,5, SE (beta2) = 7,5

(Betrag von)PARMS aus LINEQS MODEL mit 1 Faktor u. 3 pos. korr. korr. spez. Fehlern (999 Iterationen

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
beta_f1	beta_f1	999	0.7047699	0.0281035	0.5846060	0.7796581
beta_f2	beta_f2	999	0.7041604	0.0296401	0.5122009	0.7773060
beta_f3	beta_f3	999	0.7048968	0.0294994	0.5360107	0.7848326
beta1_b	Betrag von beta1	999	0.7083434	0.0275849	0.6262055	0.8113178
beta2_b	Betrag von beta2	999	0.7088360	0.0288599	0.6291227	0.8588658
beta3_b	Betrag von beta3	999	0.7081111	0.0288235	0.6197078	0.8442121

STDERR aus LINEQS MODEL mit 1 Faktor 3 pos. korr. korr. spez. Fehlern (999 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta1	beta1	999	0.0378292	0.0032763	0.0324880	0.0766114
beta_f1	beta_f1	999	0.0263628	0.0061144	0.0178088	0.0946142
beta2	beta2	999	0.0380724	0.0053650	0.0329091	0.1637847
beta_f2	beta_f2	999	0.0268350	0.0102214	0.0187964	0.2711819
beta3	beta3	999	0.0378365	0.0034028	0.0329551	0.0944881
beta_f3	beta_f3	999	0.0264048	0.0066707	0.0184136	0.1430050
Iteration		999	500.2042042	288.8125130	1.0000000	1000.00

Korrelationsmatrix der PARMS aus LINEQS MODEL mit 1 Faktor u. 3 pos. korr. korr. spez. Fehler

Die Prozedur CORR

3 Variablen: beta1_b beta2_b beta3_b

Pearsonsche Korrelationskoeffizienten, N = 999

Prob > |r| under H0: Rho=0

	beta1_b	beta2_b	beta3_b
beta1_b	1.00000	-0.03765	0.02231
Betrag von beta1		0.2344	0.4812
beta2_b	-0.03765	1.00000	-0.07123
Betrag von beta2	0.2344		0.0244
beta3_b	0.02231	-0.07123	1.00000
Betrag von beta3	0.4812	0.0244	

11.3.2.4.4 B fehlspezifiziert

(Betrag von)PARMS aus LINEQS MODEL mit 1 Faktor u. 3 pos. korr. fehlspez. Fehlern (1000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		1000	3.977387E-13	1.076813E-12	0	5.941218E12
ITER	Iteration	0	.	.	.	.
beta_f1	beta_f1	1000	0.4994213	0.0189733	0.4323884	0.5584488
beta_f2	beta_f2	1000	0.4992270	0.0195370	0.4301934	0.5550054
beta_f3	beta_f3	1000	0.4996493	0.0198207	0.4309272	0.5590509
beta1_b	Betrag von beta1	1000	0.8660826	0.0109450	0.8295390	0.9016876
beta2_b	Betrag von beta2	1000	0.8661782	0.0112509	0.8318467	0.9027369
beta3_b	Betrag von beta3	1000	0.8659261	0.0114147	0.8291334	0.9023869

STDERR aus LINEQS MODEL mit 1 Faktor 3 pos. korr. fehlspez. Fehlern (1000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta1	beta1	1000	0.0262322	0.000262468	0.0254342	0.0270152
beta_f1	beta_f1	1000	0.0177084	0.000522198	0.0160289	0.0196265
beta2	beta2	1000	0.0262309	0.000263281	0.0254795	0.0270672
beta_f2	beta_f2	1000	0.0177167	0.000549856	0.0162182	0.0198297
beta3	beta3	1000	0.0262341	0.000268702	0.0253282	0.0270957
beta_f3	beta_f3	1000	0.0177054	0.000521306	0.0162415	0.0194568

11.3.2.4.5 Korrelationsmatrix der PARMS aus LINEQS MODEL mit 1 Faktor u. 3 pos. korr. fehlspez. Fehlern

```
                    Die Prozedur CORR
        3  Variablen:    beta1_b  beta2_b  beta3_b

     Pearsonsche Korrelationskoeffizienten, N = 1000
                Prob > |r| under H0: Rho=0
                       beta1_b       beta2_b       beta3_b
beta1_b                1.00000       0.00242       0.05371
Betrag von beta1                      0.9390        0.0896
beta2_b                0.00242       1.00000      -0.03214
Betrag von beta2        0.9390                      0.3100
beta3_b                0.05371      -0.03214       1.00000
Betrag von beta3        0.0896        0.3100
```

11.4 PROC CALIS LINEQS-MODELLE MIT EINEM FAKTOR UND VIER ODER FÜNF INDIKATOREN

11.4.1 A SAS-PROGRAMMCODE

11.4.1.1 Vollständiger kommentierter Programmcode des Modells 1

```
*1.;
*Es werden 1000 mal 7 standardnormalverteilte Zufallsvariablen mit n=1000 erzeugt;
%MACRO machdat5 (i, seedx0, seedx1, seedx2, seedx3, seedx4, seedx5 ,seedx6
 , dateinam=, lib=);
DATA &lib..&dateinam;
Do i = 1 to &i;
x0 = Rannor (&seedx0);
x1 = Rannor (&seedx1);
x2 = Rannor (&seedx2);
x3 = Rannor (&seedx3);
x4 = Rannor (&seedx4);
x5 = Rannor (&seedx5);
x6 = Rannor (&seedx6);
OUTPUT;
END;
RUN;
%MEND machdat5;
%macro create9(howmany);
   %do i=1 %to &howmany;
      %machdat5 (1000, &i, &i, &i, &i, &i, &i, &i,  dateinam=iterat&i,
                 lib=base_mak);
       run;
   %end;
%mend create9;
%create9(1000)
*2.;
*diese Zufallsvariablen werden weiterverarbeitet, indem manifeste Variablen aus Ihnen konstruiert werden;
%Macro arit_op2 (dateinam=, lib=);
DATA &lib..b_&dateinam;
SET &lib..&dateinam;
*Modell 1:
*2 pos. korrel. Fehler, die von allen anderen unabhängig sind;
 y1_s = (x1 + x0)/ SQRT(2);
 y2_s = (x2 + x0)/ SQRT(2);
 indi_1ps = x4 + y1_s;
 indi_2ps = x4 + y2_s;
 indik_3 = x4 + x3;
 indik_4 = x4 + x5;
TITLE
"Sieben_unkorrel_normalvert_plus_resultierende_Var_standard_korr_Fehler &lib..b_&dateinam";
RUN;
%MEND arit_op2;
%macro create7(howmany);
   %do i=1 %to &howmany;
      %arit_op2 (dateinam=iterat&i, lib=base_mak);
       run;
   %end;
%mend create7;
%create7(1000)

*3.;
*Es wird eine Korrelationsmatrix aus dieser eben erzeugten Rohdatendatei gebildet;
%Macro corrmatr (dateinam=, lib=);
```

```
PROC CORR DATA=&lib..b_&dateinam noprint outp=&lib..cm_b_&dateinam;
var indi_1ps indi_2ps indik_3 indik_4;
TITLE;
   run;
%MEND corrmatr;
%macro create8(howmany);
   %do i=1 %to &howmany;
      %corrmatr (dateinam=iterat&i, lib=base_mak);
       run;
   %end;
%mend create8;
%create8(1000)

*4.;
*Erzeugung von (1000) OUTEST-datasets bei den (1000) Durchläufen von PROC *CALIS / LINEQS auf Basis der oben
*erzeugten (1000) Korrelationsmatrizen;
%Macro proccalis (dateinam=, lib=);
proc calis data=&lib..cm_b_&dateinam OUTEST=&lib..OUTE_cm_b_&dateinam tech=nr edf=999 nop;
         Lineqs
                  indi_1ps =         beta1    Faktor_1 + beta_f1 Error_1,
                  indi_2ps =         beta2    Faktor_1 + beta_f2 Error_2,
                  indik_3 =          beta3    Faktor_1 + beta_f3 Error_3,
                  indik_4 =          beta4    Faktor_1 + beta_f4 Error_4;

         Std
                  Error_1 = 1,
                  Error_2 = 1,
                  Error_3 = 1,
                  Error_4 = 1,
                  Faktor_1= 1;
          Cov
                  Error_1 Error_2 = cov_err12;
%MEND proccalis;
%macro create10(howmany);
   %do i=1 %to &howmany;
      %proccalis (dateinam=iterat&i, lib=base_mak);
       run;
   %end;
%mend create10;
%create10(1000)

*5.;
*Von den (1000) OUTEST-SAS-datasets wird nur jeweils die interessierende *Zeile (Beobachtung) mit dem _TYPE_ "PARMS"
(=final Parameter Estimates) bewahrt;
%Macro parms_isol (dateinam=, lib=);
DATA &lib..PARMS_OUTE_cm_b_&dateinam;
SET &lib..OUTE_cm_b_&dateinam;
IF _TYPE_ NE "PARMS" THEN DELETE;
RUN;
TITLE;
   run;
%MEND parms_isol;
%macro create11(howmany);
   %do i=1 %to &howmany;
      %parms_isol (dateinam=ITERAT&i, lib=base_mak);
       run;
   %end;
%mend create11;
%create11(1000)
*6.;
*die (1000) einzeiligen Datensätze addieren;
%macro create12(howmany);
   %do i=2 %to &howmany;
   proc APPEND BASE=BASE_MAK.PARMS_OUTE_CM_B_ITERAT1
            DATA=BASE_MAK.PARMS_OUTE_CM_B_ITERAT&i;
      run;
   %end;
%mend create12;
%create12(1000)

*7.;
*Umbenennen der Datei;
DATA BASE_MAK.PARMS_ALL_ITERAT;
SET BASE_MAK.PARMS_OUTE_CM_B_ITERAT1;
RUN;

8.;
*Die negativen Werte von beta werden in positive umgewandelt; Schließlich werden statistische Maßzahlen und Histog-
ramme für die Fehlerkorrelationen und interessierenden Pfadkoeffizienten erstellt;
DATA BASE_MAK.PARMS_ALL_ITERAT_B;
SET BASE_MAK.PARMS_ALL_ITERAT;
beta1_b=ABS(beta1);
beta2_b=ABS(beta2);
beta3_b=ABS(beta3);
```

```
beta4_b=ABS(beta4);
DROP beta1 beta2 beta3 beta4;
LABEL beta1_b='Betrag von beta1'
         beta2_b='Betrag von beta2'
         beta3_b='Betrag von beta3'
         beta4_b='Betrag von beta4'
         cov_err_12='Korrelation zwischen Fehler 1 und 2';
TITLE "(Betrag von) PARMS aus MODEL 1 mit 4 Indikat. u. 1 pos korrel. Fehlerpaar (1000 Iterationen)";
RUN;
*Kommentar: Es kommen keine negativen Werte für cov_er_12 vor, daher ist *Betragssetzung nicht nötig;
PROC MEANS DATA=BASE_MAK.PARMS_ALL_ITERAT_B;
TITLE "(Betrag von) PARMS aus MODEL 1 mit  4 Indikat. u. 1 pos korrel. Fehlerpaar (1000 Iterationen)";
RUN;
PROC GCHART DATA=BASE_MAK.PARMS_ALL_ITERAT_B;
VBAR beta1_b beta2_b beta3_b beta4_b cov_err12;
TITLE "(Betrag von) PARMS aus MODEL 1 mit 4 Indikat. u. 1 pos korrel. Fehlerpaar (1000 Iterationen)";
RUN;

*9.;
*Die geschätzten Standardfehler (SE) der 1000 PROC CALIS OUTEST-
*Outputdateien isolieren;
%Macro se_isol (dateinam=, lib=);
DATA &lib..Stderr_OUTE_cm_b_&dateinam;
SET &lib..OUTE_cm_b_&dateinam;
IF _TYPE_ NE "STDERR" THEN DELETE;
RUN;
TITLE "STDERR von MODEL 1 mit 4 Indikat. u. 1 pos korrel. Fehlerpaar &lib..OUTE_cm_b_&dateinam";
   run;
%MEND se_isol;
%macro create13(howmany);
   %do i=1 %to &howmany;
      %se_isol (dateinam=iterat&i, lib=base_mak);
       run;
   %end;
%mend create13;
%create13(1000)
*10.;
*die (1000) isolierten einzeiligen Datensätze addieren;
%macro create14(howmany);
   %do i=2 %to &howmany;
   proc APPEND BASE=base_mak.STDERR_OUTE_CM_B_ITERAT1
            DATA=base_mak.STDERR_OUTE_CM_B_ITERAT&i;
      run;
   %end;
%mend create14;
%create14(1000)
*11.;
*Umbenennen der Datei;
DATA base_mak.STDERR_ALL_ITERAT;
SET base_mak.STDERR_OUTE_CM_B_ITERAT1;
RUN;
*12.;
*statistische Maßzahlen für den SE ausgeben;
*a) arithmetisches Mittel, Standardabweichung sowie Minimum und Maximum;
PROC MEANS DATA=base_mak.STDERR_ALL_ITERAT;
TITLE "STDERR aus MODEL 1 mit 4 Indikat. u. 1 pos. korrel. Fehlerpaar (1000 Iterationen)";
RUN;
*b)Zur Errechnung des Medians;
PROC UNIVARIATE DATA=base_mak.STDERR_ALL_ITERAT;
TITLE "STDERR aus MODEL 1 mit 4 Indikat. u. 1 pos. korrel. Fehlerpaar (1000 Iterationen)";
RUN;
```

11.4.1.2 Programmcodeausschnitt der anderen Modelle

Der Programmcode der einzelnen Lineqs-Modelle unterscheidet sich im Wesentlichen nur durch die Indikatorenkonstruktion und das COV-Statement. Daher sind im Folgenden nur die Gleichungen zur Konstruktion der Indikatoren aufgeführt, die innerhalb des Makros „arit_op2“ unter *2.[197] zu finden sind (siehe 11.4.1.1).

[197] Selbstverständlich variieren entsprechend der Variablenkonstruktion in „arit_op2“ auch die Variablenbezeichnungen bei der Erzeugung der Korrelationsmatrizen unter Schritt 3 des Programmcodes.

```
*Modell 2
*2 neg. korrel. Fehler, die von allen anderen unabhängig sind;
 y1_s = (x1 + x0) / SQRT(2);
 y2b_s = (x2 - x0) / SQRT(2);
 indi_1ps = x4 + y1_s;
 indi_2ns = x4 + y2b_s;
 indik_3 = x4 + x3;
 indik_4 = x4 + x5;

*Modell 3
*2 pos. korrel. Fehlerpaare; die beteiligten Fehler sind jeweils von allen
*anderen unabhängig;

 y1_s = (x1 + x0)/ SQRT(2);
 y2_s = (x2 + x0)/ SQRT(2);
 indi_1ps = x4 + y1_s;
 indi_2ps = x4 + y2_s;
 y3_s = (x3 + x6)/SQRT(2);
 y4_s = (x5 + x6)/SQRT(2);
 indi_3ps = x4 + y3_s;
 indi_4ps = x4 + y4_s;
 indik_5 = x4 + x7;

*Modell 4
*1 pos. und 1 neg. korreliertes Fehlerpaar; die beteiligten Fehler sind je
*weils von allen anderen unabhängig;
 y1_s = (x1 + x0)/ SQRT(2);
 y2b_s = (x2 - x0)/ SQRT(2);
 indi_1ps = x4 + y1_s;
 indi_2ns = x4 + y2b_s;

 y3_s = (x3 + x6)/SQRT(2);
 y4_s = (x5 + x6)/SQRT(2);
 indi_3ps = x4 + y3_s;
 indi_4ps = x4 + y4_s;
 indik_5 = x4 + x7;

*ursprüngl. Modell 3
*2 pos. korrel. Fehlerpaare; die beteiligten Fehler sind jeweils von allen anderen unabhängig;
 y1_s = (x1 + x0)/SQRT(2);
 y2_s = (x2 + x0)/SQRT(2);
 y3_s = (x3 + x6)/SQRT(2);
 y4_s = (x5 + x6)/SQRT(2);
 indi_1ps = x4 + y1_s;
 indi_2ps = x4 + y2_s;
 indi_3ps = x4 + y3_s;
 indi_4ps = x4 + y4_s;

* ursprüngl. Modell 4
*1 pos. und 1 neg. korreliertes Fehlerpaar; die beteiligten Fehler sind jeweils von allen anderen unabhängig;
 y1_s = (x1 + x0)/SQRT(2);
 y2b_s = (x2 - x0)/SQRT(2);
 y3_s = (x3 + x6)/SQRT(2);
 y4_s = (x5 + x6)/SQRT(2);
 indi_1ps = x4 + y1_s;
 indi_2ns = x4 + y2b_s;
 indi_3ps = x4 + y3_s;
 indi_4ps = x4 + y4_s;
```

11.4.2 B SAS-OUTPUT

Der Standard ist n = 1000 Realisationen von Zufallsvariablen und 1000 Iterationen.

11.4.2.1 Modell 1

(Betrag von) PARMS aus MODELL 1 mit vier Indikat. u. 1 pos korrel. Fehlerpaar (1000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		1000	0.000961993	0.0012913	1.423491E-10	0.0088954
ITER	Iteration	0	.	.	.	.
beta_f1	beta_f1	1000	0.7065401	0.0254107	0.6227154	0.7781718
beta_f2	beta_f2	1000	0.7065330	0.0247133	0.6244645	0.7837705
beta_f3	beta_f3	1000	0.7051906	0.0243283	0.6085865	0.7902444
beta_f4	beta_f4	1000	0.7049237	0.0232743	0.6391660	0.7716030

cov_err12	cov_err12	1000	0.4991150	0.0345595	0.3266595	0.5922352
beta1_b	Betrag von beta1	1000	0.7067596	0.0254480	0.6280514	0.7824484
beta2_b	Betrag von beta2	1000	0.7068162	0.0247478	0.6210505	0.7810532
beta3_b	Betrag von beta3	1000	0.7081873	0.0242147	0.6127918	0.7934876
beta4_b	Betrag von beta4	1000	0.7085216	0.0232176	0.6361044	0.7690688

STDERR aus MODEL 1 mit 4 Indikat. u. 1 pos. korrel. Fehlerpaar (1000 Iterationen)
Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta1	beta1	1000	0.0335386	0.000775999	0.0314216	0.0367045
beta_f1	beta_f1	1000	0.0250495	0.0012277	0.0223923	0.0312920
beta2	beta2	1000	0.0335382	0.000770628	0.0315186	0.0364790
beta_f2	beta_f2	1000	0.0250475	0.0012167	0.0223523	0.0315714
beta3	beta3	1000	0.0325860	0.000663323	0.0308050	0.0352382
beta_f3	beta_f3	1000	0.0237880	0.000911106	0.0217525	0.0282623
beta4	beta4	1000	0.0325892	0.000667805	0.0308147	0.0349093
beta_f4	beta_f4	1000	0.0237998	0.000910068	0.0216686	0.0273399
cov_err12	cov_err12	1000	0.0347573	0.0039023	0.0262829	0.0612445

11.4.2.2 Modell 2

(Betrag von) PARMS aus MODELL 2 mit vier Indikat. u. 1 neg korrel. Fehlerpaar (1000 Iterationen)
Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		1000	0.0010173	0.0014349	6.083605E-12	0.0113713
ITER	Iteration	0	.	.	.	.
beta_f1	beta_f1	1000	0.7064782	0.0254194	0.6226792	0.7789501
beta_f2	beta_f2	1000	0.7074183	0.0253588	0.6205219	0.7785053
beta_f3	beta_f3	1000	0.7052912	0.0211735	0.6321016	0.7732076
beta_f4	beta_f4	1000	0.7053337	0.0210855	0.6372411	0.7775051
cov_err12	cov_err12	1000	-0.4993387	0.0631622	-0.7984121	-0.3044216
beta1_b	Betrag von beta1	1000	0.7068208	0.0254568	0.6270859	0.7824772
beta2_b	Betrag von beta2	1000	0.7058862	0.0253427	0.6276380	0.7841892
beta3_b	Betrag von beta3	1000	0.7082892	0.0210576	0.6341531	0.7748855
beta4_b	Betrag von beta4	1000	0.7082497	0.0210484	0.6288766	0.7706645

STDERR aus MODEL 2 mit 4 Indikat. u. 1 neg. korrel. Fehlerpaar (1000 Iterationen)
Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta1	beta1	1000	0.0335322	0.000775857	0.0314195	0.0367008
beta_f1	beta_f1	1000	0.0250438	0.0012298	0.0223963	0.0312920
beta2	beta2	1000	0.0335467	0.000787992	0.0315855	0.0363128
beta_f2	beta_f2	1000	0.0250196	0.0012077	0.0222077	0.0322292
beta3	beta3	1000	0.0306129	0.000500469	0.0291232	0.0326293
beta_f3	beta_f3	1000	0.0209409	0.000452278	0.0198995	0.0230428
beta4	beta4	1000	0.0306134	0.000497125	0.0293485	0.0324873
beta_f4	beta_f4	1000	0.0209415	0.000457636	0.0194875	0.0225424
cov_err12	cov_err12	1000	0.0632313	0.0075716	0.0477342	0.1060461

11.4.2.3 Modell 3

(Betrag von) PARMS aus MODELL 1 mit fünf Indikat. u. zwei pos korrel. Fehlerpaaren (1000) Iteration 69
11:36 Wednesday, July 25, 2007

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		1000	0.0029994	0.0025468	0.000027548	0.0187687
ITER	Iteration	0	.	.	.	.
beta_f1	beta_f1	1000	0.7044476	0.0233510	0.6211125	0.7804807
beta_f2	beta_f2	1000	0.7037755	0.0231244	0.6274597	0.7735718
beta_f3	beta_f3	1000	0.7060626	0.0245363	0.6291471	0.7828032
beta_f4	beta_f4	1000	0.7067437	0.0241144	0.6297886	0.7928322
beta_f5	beta_f5	1000	0.7068062	0.0230524	0.6315751	0.7743209
cov_err12	cov_err12	1000	0.4958575	0.0333547	0.3817924	0.5970870
cov_err34	cov_err34	1000	0.5003449	0.0321932	0.3925850	0.5879921
beta1_b	Betrag von beta1	1000	0.7089905	0.0232771	0.6251799	0.7837214
beta2_b	Betrag von beta2	1000	0.7096755	0.0229554	0.6337086	0.7786490
beta3_b	Betrag von beta3	1000	0.7073035	0.0244246	0.6222694	0.7772863
beta4_b	Betrag von beta4	1000	0.7066474	0.0241328	0.6094400	0.7767666
beta5_b	Betrag von beta5	1000	0.7066539	0.0231269	0.6327931	0.7753147

STDERR aus MODELL 1 mit vier Indikat. u. zwei pos. korrel. Fehlerpaaren (1000) Iterationen) 154
11:36 Wednesday, July 25, 2007

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
beta1	beta1	1000	0.0325880	0.000671430	0.0307911	0.0349977
beta2	beta2	1000	0.0325775	0.000666931	0.0303437	0.0350214
beta3	beta3	1000	0.0325612	0.000649691	0.0305598	0.0350800
beta4	beta4	1000	0.0325715	0.000661272	0.0306233	0.0351000
beta5	beta5	1000	0.0317334	0.000571409	0.0300226	0.0336698
beta_f1	beta_f1	1000	0.0238207	0.000977630	0.0215103	0.0280394
beta_f2	beta_f2	1000	0.0238321	0.000980568	0.0214471	0.0280128
beta_f3	beta_f3	1000	0.0237245	0.000921780	0.0215542	0.0281646
beta_f4	beta_f4	1000	0.0237083	0.000893872	0.0215425	0.0273983
beta_f5	beta_f5	1000	0.0225299	0.000687582	0.0206030	0.0267773
cov_err12	cov_err12	1000	0.0334551	0.0032472	0.0242729	0.0484110
cov_err34	cov_err34	1000	0.0329943	0.0031722	0.0259668	0.0477428

11.4.2.4 Modell 4

(Betrag von) PARMS aus MODELL 1 mit 5 Indikat., 1 pos u. 1 neg. korrel. Fehlerpaar (1000) It 226
11:36 Wednesday, July 25, 2007

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		1000	0.0030226	0.0024785	7.9485491E-6	0.0171030
ITER	Iteration	0	.	.	.	.
beta_f1	beta_f1	1000	0.7044481	0.0232649	0.6204455	0.7791982
beta_f2	beta_f2	1000	0.7055268	0.0247738	0.6232741	0.7809336
beta_f3	beta_f3	1000	0.7061587	0.0216346	0.6290084	0.7699770
beta_f4	beta_f4	1000	0.7070008	0.0213125	0.6426452	0.7851580
beta_f5	beta_f5	1000	0.7070939	0.0204913	0.6357191	0.7652091
cov_err12	cov_err12	1000	-0.5078748	0.0600015	-0.7039524	-0.3072683
cov_err34	cov_err34	1000	0.5007064	0.0288771	0.4034075	0.5798770
beta1_b	Betrag von beta1	1000	0.7089958	0.0231882	0.6267776	0.7842496
beta2_b	Betrag von beta2	1000	0.7078213	0.0246654	0.6246140	0.7820034
beta3_b	Betrag von beta3	1000	0.7073950	0.0215635	0.6380716	0.7773985
beta4_b	Betrag von beta4	1000	0.7065691	0.0213706	0.6192955	0.7661639
beta5_b	Betrag von beta5	1000	0.7065243	0.0205561	0.6437818	0.7719205

STDERR aus MODEL 1 mit 4 Indikat., 1 pos u. 1 neg. korrel. Fehlerpaar (1000) Iterationen 225
11:36 Wednesday, July 25, 2007

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
beta1	beta1	1000	0.0325501	0.000667299	0.0307282	0.0349640
beta2	beta2	1000	0.0325677	0.000683722	0.0308234	0.0350985
beta3	beta3	1000	0.0303840	0.000455538	0.0289342	0.0318488
beta4	beta4	1000	0.0303949	0.000463135	0.0290580	0.0320074
beta5	beta5	1000	0.0299829	0.000439065	0.0288390	0.0314708
beta_f1	beta_f1	1000	0.0237684	0.000977357	0.0213766	0.0279839
beta_f2	beta_f2	1000	0.0237548	0.000972843	0.0212839	0.0282189
beta_f3	beta_f3	1000	0.0205897	0.000455941	0.0194298	0.0221258
beta_f4	beta_f4	1000	0.0205823	0.000447020	0.0194281	0.0224569
beta_f5	beta_f5	1000	0.0199599	0.000382933	0.0187859	0.0214877
cov_err12	cov_err12	1000	0.0579888	0.0061151	0.0448619	0.0878434
cov_err34	cov_err34	1000	0.0293458	0.0018843	0.0242595	0.0362360

11.4.2.5 Ursprüngliches Modell 3 mit vier Indikatoren

Iteration Nr. 24 wurde wegen extrem hoher Werte für die vier Faktor-Pfadkoeffizienten beta (und die vier Fehler-Pfadkoeffizienten beta_f) entfernt:

(Betrag von) PARMS aus MODEL 3 mit vier Indikat. u. zwei pos korrel. Fehlerpaaren (999 Iterat.)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		999	0.000960658	0.0012849	4.7638038E-9	0.0085730
ITER	Iteration	0	.	.	.	.
beta_f1	beta_f1	999	0.7055537	0.0255402	0.6107508	0.7782059
beta_f2	beta_f2	999	0.7055839	0.0251550	0.5919885	0.7787651
beta_f3	beta_f3	999	0.7049019	0.0267363	0.5958599	0.7921428
beta_f4	beta_f4	999	0.7049462	0.0255779	0.6183901	0.7871057
cov_err12	cov_err12	999	0.4980239	0.0345614	0.3577274	0.6041534
cov_err34	cov_err34	999	0.4983408	0.0365260	0.3378236	0.5857207
beta1_b	Betrag von beta1	999	0.7077429	0.0253568	0.6280092	0.7918229
beta2_b	Betrag von beta2	999	0.7077380	0.0250341	0.6273157	0.8059464
beta3_b	Betrag von beta3	999	0.7083094	0.0264150	0.6103357	0.8030884
beta4_b	Betrag von beta4	999	0.7083514	0.0252231	0.6168181	0.7858713

STDERR aus MODEL 3 mit 4 Indikat. u. 2 pos. korrel. Fehlerpaaren (999 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta1	beta1	999	0.0290913	0.000592092	0.0271866	0.0309821
beta_f1	beta_f1	999	0.0196701	0.000423808	0.0185071	0.0209440
beta2	beta2	999	0.0290915	0.000586828	0.0272108	0.0311171
beta_f2	beta_f2	999	0.0196688	0.000422687	0.0183567	0.0211097
beta3	beta3	999	0.0290767	0.000611599	0.0271384	0.0314639
beta_f3	beta_f3	999	0.0196648	0.000434782	0.0183447	0.0210306
beta4	beta4	999	0.0290768	0.000584506	0.0272639	0.0314765
beta_f4	beta_f4	999	0.0196632	0.000429853	0.0183046	0.0209735
cov_err12	cov_err12	999	0.0236727	0.000719221	0.0207477	0.0254570
cov_err34	cov_err34	999	0.0236168	0.000744969	0.0204349	0.0254633

11.4.2.6 Ursprüngliches Modell 4 mit vier Indikatoren

In Iteration 710 und 966 traten für $r_{\varepsilon_1,\varepsilon_2}$ Werte größer Null auf (0,23 bzw. 0,19), Diese Iterationen wurden zur Vermeidung einer Verzerrung der statistischen Maßzahlen aus der Analyse ausgeschlossen.

(Betrag von) PARMS aus MODEL 4 mit vier Indikatoren mit 1 pos u. 1 neg. korrel. Fehlerpaar (998 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
beta1_b	Betrag von beta1	998	0.6544551	0.0226899	0.5693130	0.7145775
beta2_b	Betrag von beta2	998	0.6524618	0.0232799	0.5703426	0.7182909
beta3_b	Betrag von beta3	998	0.7658341	0.0179766	0.7021060	0.8185191
beta4_b	Betrag von beta4	998	0.7656469	0.0174954	0.7076340	0.8188287
beta_f1	beta_f1	998	0.7555082	0.0195443	0.6995563	0.8221209
beta_f2	beta_f2	998	0.7572007	0.0199966	0.6957429	0.8214069
beta_f3	beta_f3	998	0.6424271	0.0215206	0.5744792	0.7120724
beta_f4	beta_f4	998	0.6426832	0.0209252	0.5740379	0.7065792
cov_err12	cov_err12	998	-0.3084928	0.0370420	-0.4036654	-0.1811790
cov_err34	cov_err34	998	0.3968474	0.0323768	0.2867406	0.4803804

STDERR aus MODEL 4 mit 4 Indikatoren mit 1 pos u. 1 neg. korrel. Fehlerpaar (998 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
beta1	beta1	998	0.0316769	0.000566124	0.0302169	0.0337367
beta2	beta2	998	0.0317131	0.000583540	0.0299650	0.0333981
beta3	beta3	998	0.0274189	0.000347026	0.0264160	0.0285538
beta4	beta4	998	0.0274222	0.000343306	0.0264445	0.0286010
beta_f1	beta_f1	998	0.0215661	0.000396043	0.0204552	0.0229644
beta_f2	beta_f2	998	0.0215703	0.000398998	0.0201734	0.0230181
beta_f3	beta_f3	998	0.0177362	0.000377797	0.0163632	0.0189390
beta_f4	beta_f4	998	0.0177365	0.000375122	0.0163563	0.0189295
cov_err12	cov_err12	998	0.0274247	0.0011251	0.0238302	0.0316868
cov_err34	cov_err34	998	0.0243605	0.000699389	0.0214048	0.0258639

11.5 PROC CALIS LINEQS-MODELLE MIT ZWEI FAKTOREN UND SECHS INDIKATOREN

11.5.1 A PROGRAMMCODE

11.5.1.1 Modelle mit unabhängigen Faktoren

11.5.1.1.1 Vollständiger kommentierter Programmcode des Modells 1 (Grundmodell)

```
*Unterdrückung des Logfiles;
options nonotes nosource nosource2 errors=0;
*1.;
*Es werden 3000 mal acht standardnormalverteilte Zufallsvariable mit n=1000 Realisationen erzeugt;
%MACRO machdat5 (i, seedx1, seedx2, seedx3, seedx4, seedx5 ,seedx6 ,seedx7 ,seedx8 , dateinam=);
%GLOBAL sds;
%LET sds=&dateinam;
DATA &sds;
Do i = 1 to &i;
x1 = Rannor (&seedx1);
x2 = Rannor (&seedx2);
x3 = Rannor (&seedx3);
x4 = Rannor (&seedx4);
x5 = Rannor (&seedx5);
x6 = Rannor (&seedx6);
x7 = Rannor (&seedx7);
x8 = Rannor (&seedx8);
OUTPUT;
```

```
END;
RUN;
%MEND machdat5;
%macro create9(howmany);
   %do i=1 %to &howmany;
      %machdat5 (1000, &i, &i, &i, &i, &i, &i, &i, &i, dateinam=iterat&i);
       run;
   %end;
%mend create9;
%create9(3000)
*2.;
*Diese Zufallsvariablen werden weiterverarbeitet, indem sechs manifeste Variable aus ihnen konstruiert werden – ent-
sprechend dem
in AMOS gebildeten Grundmodell;
%Macro arit_op2 (dateinam=);
%GLOBAL sds;
%LET sds=&dateinam;
DATA b_&sds;
SET &sds;
indi_1 = x1 + x3;
indi_2 = x1 + x4;
indi_3 = x1 + x5;
indi_4 = x2 + x6;
indi_5 = x2 + x7;
indi_6 = x2 + x8;
RUN;
%MEND arit_op2;
%macro create7(howmany);
   %do i=1 %to &howmany;
      %arit_op2 (dateinam=iterat&i);
       run;
   %end;
%mend create7;
%create7(3000)
*3.;
*Es wird je eine Korrelationsmatrix aus diesen eben erzeugten Rohdatendateien gebildet;
%Macro corrmatr (dateinam=);
%GLOBAL sds;
%LET sds=&dateinam;
PROC CORR DATA=&sds noprint outp=cm_&sds;
var indi_1 indi_2 indi_3 indi_4 indi_5 indi_6;
TITLE;
   run;
%MEND corrmatr;
%macro create8(howmany);
   %do i=1 %to &howmany;
      %corrmatr (dateinam=b_iterat&i);
       run;
   %end;
%mend create8;
%create8(3000)
*4.;
*Erzeugung von (3000) OUTEST-Datasets bei den (3000) Durchläufen von PROC CALIS / LINEQS auf Basis der oben erzeugten
*(3000) Korrelationsmatrizen;
%Macro proccalis (dateinam=);
%GLOBAL sds;
%LET sds=&dateinam;
proc calis data=&sds OUTEST=OUTE_&sds tech=nr edf=999 nop;
         Lineqs
                  indi_1 = beta1     Faktor_1          + beta_f1          Error_1,
                  indi_2 = beta2     Faktor_1          + beta_f2          Error_2,
                  indi_3 = beta3     Faktor_1          + beta_f3          Error_3,
                  indi_4 = beta4     Faktor_2          + beta_f4          Error_4,
                  indi_5 = beta5     Faktor_2          + beta_f5          Error_5,
                  indi_6 = beta6     Faktor_2          + beta_f6          Error_6,
                  Faktor_2 = gamma   Faktor_1          +                  Error_7;
         Std
                  Error_1 = 1,
                  Error_2 = 1,
                  Error_3 = 1,
                  Error_4 = 1,
                  Error_5 = 1,
                  Error_6 = 1,
                  Error_7 = 1,
                  Faktor_1 = 1;
                  run;
%MEND proccalis;
%macro create10(howmany);
   %do i=1 %to &howmany;
      %proccalis (dateinam=cm_b_iterat&i);
       run;
   %end;
%mend create10;
%create10(3000)
```

```
*5.;
*Von den (3000) OUTEST-sas-data-sets wird nur jeweils die interessierende Zeile (Beobachtung) mit dem _TYPE_ "PARMS"
*(=final parameter estimates) bewahrt;
%Macro parms_isol (dateinam=);
%GLOBAL sds;
%LET sds=&dateinam;
DATA PARMS_&sds;
SET &sds;
IF _TYPE_ NE "PARMS" THEN DELETE;
RUN;
TITLE "PARMS von LINEQS-Modell mit 1 Faktor und unkorrelierten Fehlern &sds";
   run;
%MEND parms_isol;
%macro create11(howmany);
   %do i=1 %to &howmany;
      %parms_isol (dateinam=OUTE_CM_B_ITERAT&i);
       run;
   %end;
%mend create11;
%create11(3000)
*6.;
*die (3000) einzeiligen Datensätze addieren;
%macro create12(howmany);
   %do i=2 %to &howmany;
   proc APPEND BASE=PARMS_OUTE_CM_B_ITERAT1
            DATA=PARMS_OUTE_CM_B_ITERAT&i;
     run;
   %end;
%mend create12;
%create12(3000)
*7.;
*Umbenennen der Datei;
DATA PARMS_ALL_ITERAT;
SET PARMS_OUTE_CM_B_ITERAT1;
RUN;

8.;
*die negativen Werte von beta werden in positive umgewandelt; Schließlich werden statistische Maßzahlen und
*Histogramme für die interessierenden Pfadkoeffizienten erstellt sowie Pfadkoeffizientenkorrelationen berechnet;

DATA PARMS_ALL_ITERAT_B;
SET PARMS_ALL_ITERAT;
beta1_b=ABS(beta1);
beta2_b=ABS(beta2);
beta3_b=ABS(beta3);
beta4_b=ABS(beta4);
beta5_b=ABS(beta5);
beta6_b=ABS(beta6);
DROP beta1 beta2 beta3 beta4 beta5 beta6 ;
LABEL beta1_b='Betrag von beta1'
           beta2_b='Betrag von beta2'
           beta3_b='Betrag von beta3'
        beta4_b='Betrag von beta4'
           beta5_b='Betrag von beta5'
           beta6_b='Betrag von beta6'
           gamma='gamma';
TITLE "(Betrag von) PARMS aus MODEL 1 mit 2 unkorrel Faktoren und unkorrel. Fehlern (3000 Iterationen)";
RUN;
PROC MEANS DATA=PARMS_ALL_ITERAT_B;
TITLE "(außer gamma Betrag v.) PARMS MODEL 1 m. 2 unkorr. Faktoren und unkorrel. Fehlern (3000 Iterationen)";
RUN;
PROC GCHART DATA=PARMS_ALL_ITERAT_B;
VBAR beta1_b beta2_b beta3_b beta4_b beta5_b beta6_b gamma;
TITLE "(außer gamma Betrag v.) PARMS MODEL 1 m. 2 unkorr. Faktoren und unkorrel. Fehlern (3000 Iterationen)";
RUN;
*9.;
*Die geschätzten Standardfehler der 3000 PROC CALIS OUTEST-Outputdateien isolieren;
%Macro se_isol (dateinam=);
%GLOBAL sds;
%LET sds=&dateinam;
DATA Stderr_&sds;
SET &sds;
IF _TYPE_ NE "STDERR" THEN DELETE;
RUN;
TITLE "STDERR von MODEL 1 mit 2 unkorrel Faktoren und unkorrelierten Fehlern &sds";
   run;
%MEND se_isol;
%macro create13(howmany);
   %do i=1 %to &howmany;
      %se_isol (dateinam=OUTE_CM_B_ITERAT&i);
       run;
   %end;
%mend create13;
%create13(3000)
```

```
*10.;
*die (3000) isolierten einzeiligen Datensätze addieren;
%macro create14(howmany);
   %do i=2 %to &howmany;
   proc APPEND BASE=STDERR_OUTE_CM_B_ITERAT1
           DATA=STDERR_OUTE_CM_B_ITERAT&i;
      run;
   %end;
%mend create14;
%create14(3000)
*11.;
*Umbenennen der Datei;
DATA STDERR_ALL_ITERAT;
SET STDERR_OUTE_CM_B_ITERAT1;
RUN;
*12.;
*statistische Maßzahlen für den SE ausgeben;
PROC MEANS DATA=STDERR_ALL_ITERAT;
TITLE "STDERR aus MODEL 1 mit 2 unkorrel Faktoren und unkorrel. Fehlern (3000 Iterationen)";
RUN;
```

11.5.1.1.2 Programmcodeausschnitt Modell 2

Der Programmcode der einzelnen Modelle unterscheidet sich im Wesentlichen nur durch die notwendige Anzahl der standardnormalverteilten Zufallsvariablen, die Indikatorenkonstruktion, die im COV-Statement der Prozedur CALIS fehlspezifizierten Fehlerkorrelationen und die Betragsbildung der Schätzwerte für den Pfadkoeffizienten gamma. Daher sind im Folgenden nur die entsprechenden Schritte 1., 2., 4. und 8. des Programmcodes aufgeführt. Das STD-Statement in Schritt 4 entfällt, da es in allen Modellen gleich und mit jenem in Modell 1 identisch ist.

```
*1.;
*Es werden 3000 mal neun standardnormalverteilte Zufallsvariablen mit n=1000 Realisationen erzeugt;
%MACRO machdat5 (i, seedx1, seedx2, seedx3, seedx4, seedx5, seedx6 ,seedx7 ,seedx8 , seedx9, dateinam=);
%GLOBAL sds;
%LET sds=&dateinam;
DATA &sds;
Do i = 1 to &i;
x1 = Rannor (&seedx1);
x2 = Rannor (&seedx2);
x3 = Rannor (&seedx3);
x4 = Rannor (&seedx4);
x5 = Rannor (&seedx5);
x6 = Rannor (&seedx6);
x7 = Rannor (&seedx7);
x8 = Rannor (&seedx8);
x9 = Rannor (&seedx9);
OUTPUT;
END;
RUN;
%MEND machdat5;

*2.;
*Diese Zufallsvariablen werden weiterverarbeitet, indem sechs manifeste Variable
*aus ihnen konstruiert werden – entsprechend dem in AMOS gebildeten Modell 2;
%Macro arit_op2 (dateinam=);
%GLOBAL sds;
%LET sds=&dateinam;
DATA b_&sds;
SET &sds;
ind_1sp5 = x1 + (x3 + x9) /SQRT(2);
indi_2 = x1 + x4;
indi_3 = x1 + x5;
ind_4sp5 = x2 + (x6 + x9) /SQRT(2);
indi_5 = x2 + x7;
indi_6 = x2 + x8;
RUN;
%MEND arit_op2;
*4.;
%Macro proccalis (dateinam=);
%GLOBAL sds;
%LET sds=&dateinam;
proc calis data=&sds OUTEST=OUTE_&sds tech=nr edf=999 nop;
```

```
        Lineqs
                ind_1sp5 = beta1   Faktor_1          + beta_f1          Error_1,
                indi_2 =   beta2   Faktor_1          + beta_f2          Error_2,
                indi_3 =   beta3   Faktor_1          + beta_f3          Error_3,
                ind_4sp5 = beta4   Faktor_2          + beta_f4          Error_4,
                indi_5 =   beta5   Faktor_2          + beta_f5          Error_5,
                indi_6 =   beta6   Faktor_2          + beta_f6          Error_6,
                Faktor_2 = gamma   Faktor_1          +                  Error_7;
        Std
                ...
Cov
                Error_1 Error_4 = 0;
        run;
%MEND proccalis;
*8.;
*Die negativen Werte von beta werden in positive umgewandelt; Es werden drei Varianten von gamma erstellt:
*Originalwert, Betrag von gamma (bereichsunabhängig) und Betrag von gamma (bereichsabhängig)

DATA PARMS_ALL_ITERAT_B;
SET PARMS_ALL_ITERAT;
beta1_b=ABS(beta1);
beta2_b=ABS(beta2);
beta3_b=ABS(beta3);
beta4_b=ABS(beta4);
beta5_b=ABS(beta5);
beta6_b=ABS(beta6);
gamma_b=ABS(gamma);
IF gamma < -0.0414 THEN gamma_b2=ABS(gamma);
IF gamma >=-0.0414 THEN gamma_b2=gamma;
DROP beta1 beta2 beta3 beta4 beta5 beta6;
LABEL beta1_b='Betrag von beta1'
           beta2_b='Betrag von beta2'
           beta3_b='Betrag von beta3'
        beta4_b='Betrag von beta4'
           beta5_b='Betrag von beta5'
           beta6_b='Betrag von beta6'
           gamma='gamma'
           gamma_b='Betrag von gamma'
           gamma_b2='Betrag von gamma, wenn gamma < -0.0414';
RUN;
```

11.5.1.1.3 Programmcodeausschnitt Modell 3

```
*1.;
%MACRO machdat5 (i, seedx1, seedx2, seedx3, seedx4, seedx5 ,seedx6 ,seedx7 ,seedx8 , seedx9,
seedx10, dateinam=);
%GLOBAL sds;
%LET sds=&dateinam;
DATA &sds;
Do i = 1 to &i;
x1 = Rannor (&seedx1);
x2 = Rannor (&seedx2);
x3 = Rannor (&seedx3);
x4 = Rannor (&seedx4);
x5 = Rannor (&seedx5);
x6 = Rannor (&seedx6);
x7 = Rannor (&seedx7);
x8 = Rannor (&seedx8);
x9 = Rannor (&seedx9);
x10 = Rannor (&seedx10);
OUTPUT;
END;
RUN;
%MEND machdat5;

*2.;
%Macro arit_op2 (dateinam=);
%GLOBAL sds;
%LET sds=&dateinam;
DATA b_&sds;
SET &sds;
ind_1sp6 = x1 + (x3 + x9 + x10) /SQRT(3);
ind_2sp6 = x1 + (x4 + 1.73*x10) /SQRT(4);
indi_3 = x1 + x5;
ind_4sp6 = x2 + (x6 + 1.73*x9) /SQRT(4);
indi_5 = x2 + x7;
indi_6 = x2 + x8;
RUN;
%MEND arit_op2;
*4.;
%Macro proccalis (dateinam=);
%GLOBAL sds;
```

```
%LET sds=&dateinam;
proc calis data=&sds OUTEST=OUTE_&sds tech=nr edf=999 nop;

Lineqs
                  ind_1sp6 = beta1   Faktor_1          + beta_f1          Error_1,
                  ind_2sp6 = beta2   Faktor_1          + beta_f2          Error_2,
                  indi_3 =   beta3   Faktor_1          + beta_f3          Error_3,
                  ind_4sp6 = beta4  Faktor_2           + beta_f4          Error_4,
                  indi_5 =   beta5   Faktor_2          + beta_f5          Error_5,
                  indi_6 =   beta6   Faktor_2          + beta_f6          Error_6,
                  Faktor_2 = gamma   Faktor_1          +                  Error_7;
         Std
                  ...
Cov
                  Error_1 Error_4 = 0,
                  Error_1 Error_2 = 0;
         run;
%MEND proccalis;
8.;
DATA PARMS_ALL_ITERAT_B;
SET PARMS_ALL_ITERAT;
beta1_b=ABS(beta1);
beta2_b=ABS(beta2);
beta3_b=ABS(beta3);
beta4_b=ABS(beta4);
beta5_b=ABS(beta5);
beta6_b=ABS(beta6);
gamma_b=ABS(gamma);
IF gamma < -0.0244 THEN gamma_b2=ABS(gamma);
IF gamma >=-0.0244 THEN gamma_b2=gamma;
DROP beta1 beta2 beta3 beta4 beta5 beta6;
LABEL beta1_b='Betrag von beta1'
         beta2_b='Betrag von beta2'
         beta3_b='Betrag von beta3'
      beta4_b='Betrag von beta4'
         beta5_b='Betrag von beta5'
         beta6_b='Betrag von beta6'
         gamma='gamma'
         gamma_b='Betrag von gamma'
         gamma_b2='Betrag von gamma, wenn gamma < -0.0244';
RUN;
```

11.5.1.2 Modelle mit abhängigen Faktoren

11.5.1.2.1 Variablenerzeugung und Bildung der Indikatorenkorrelationsmatrix

Dieser Schritt liefert die Grundlage aller Modelle mit abhängigen Faktoren (einschließlich der korrekt spezifizierten Varianten und Zusatzmodelle). Im Unterschied zum Programmcode der Modelle 1 bis 3 legt dieser Programmcode nicht temporäre, sondern permanente SAS-Dateien als Grundlage aller Modelle an. Auf diese SAS-Dateien bauen die folgenden modellspezifischen Programme (siehe 11.5.1.2.2 und 11.5.1.3) auf. Die Makros des Programmcodes kommen ohne die bislang verwendete globale Makrovariable „SDS“ aus.

```
*Unterdrückung der Angaben im log;
options nonotes nosource nosource2 errors=0;
*1.;
*Es werden 3000 mal 11 standardnormalverteilte Zufallsvariablen mit n=1000
*Realisationen erzeugt;
%MACRO machdat5 (i, seedx0, seedx1, seedx2, seedx3, seedx4, seedx5 ,seedx6 ,seedx7, seedx8, seedx9, seedx10,
dateinam=, lib=);
DATA &lib..&dateinam;
Do i = 1 to &i;
x0 = Rannor (&seedx0);
x1 = Rannor (&seedx1);
x2 = Rannor (&seedx2);
x3 = Rannor (&seedx3);
x4 = Rannor (&seedx4);
x5 = Rannor (&seedx5);
x6 = Rannor (&seedx6);
x7 = Rannor (&seedx7);
x8 = Rannor (&seedx8);
x9 = Rannor (&seedx9);
x10 = Rannor (&seedx10);
OUTPUT;
END;
RUN;
%MEND machdat5;
%macro create9(howmany);
   %do i=1 %to &howmany;
      %machdat5 (1000, &i, &i, &i, &i, &i, &i, &i, &i, &i, &i, &i, dateinam=iterat&i, lib=base_mak);
       run;
   %end;
%mend create9;
%create9(10000)
*2.;
*diese Zufallsvariablen werden weiterverarbeitet, indem die für alle Modelle  benötigten  manifesten Variablen aus
*ihnen konstruiert werden – entsprechend den in AMOS gebildeten Modellen;
%Macro arit_op2 (dateinam=, lib=);
DATA &lib..b_&dateinam;
SET &lib..&dateinam;

*Mod. 4 (Grundmodell zur Erzeugung einer Korrelation zwischen den Faktoren von *0,707);
 y1_ = x0 + x1;
 y1s_= y1_ / SQRT(2);
 indi_1 = x1 + x3;
 indi_2 = x1 + x4;
 indi_3 = x1 + x5;
 indi_4s_ = y1s_ + x6;
 indi_5s_ = y1s_ + x7;
 indi_6s_ = y1s_ + x8;

*Mod. 5;
 ind_1sp2 = x1 + (x3 + x9) /SQRT(2);
 ind_2sp2 = x1 + (x4 + x9) /SQRT(2);

*Mod. 6 und 7;
 ind_4sp3 = y1s_+ (x6 + x10)/SQRT(2);
 ind_5sp3 = y1s_+ (x7 + x10)/SQRT(2);

*Mod. 8;
 ind_4sp5 = y1s_ + (x6 + x9)/SQRT(2);

*Mod. 9;
 ind_1sp6 = x1 + (x3 + x9 + x10) /SQRT(3);
 ind_2sp6 = x1 + (x4 + 1.73*x10) /SQRT(4);
 ind_4sp6 = y1s_ + (x6 + 1.73*x9) /SQRT(4);
RUN;

%MEND arit_op2;
%macro create7(howmany);
   %do i=1 %to &howmany;
      %arit_op2 (dateinam=iterat&i, lib=base_mak);
       run;
   %end;
%mend create7;
%create7(3000)
*Wiederaufhebung der Unterdrückung der Angaben im log;
options notes source source2 errors=20;
```

11.5.1.2.2 Vollständiger, komment. und modellspezifischer Programmcode des Modells 4 (Grundmodell)

```
*Dieses Programm setzt auf das Programm "Grundlage aller Modelle mit abhängigen *Faktoren“ auf, das die Erzeugung von
*Zufallszahlen (Schritt 1) für alle 6 Modelle dieser Reihe und die Verarbeitung derselben (Schritt 2) leistet. Es
*setzt somit bei Schritt 3 ein;
*Unterdrückung von der Angaben im log;
options nonotes nosource nosource2 errors=0;
*3.;
*für Mod. 4
**********
*Es wird eine Korrelationsmatrix aus diesen eben erzeugten Rohdatendateien gebildet;
%Macro corrmatr (dateinam=, lib=);
PROC CORR DATA=&lib..b_&dateinam noprint outp=&lib..cm_b_&dateinam;
var indi_1 indi_2 indi_3 indi_4s_ indi_5s_ indi_6s_ ;
TITLE;
   run;
%MEND corrmatr;
%macro create8(howmany);
   %do i=1 %to &howmany;
      %corrmatr (dateinam=iterat&i, lib=base_mak);
       run;
   %end;
%mend create8;
%create8(3000)

*4.;
*Erzeugung von (3000) OUTEST-datasets bei den (3000) Durchläufen von PROC CALIS / *LINEQS auf Basis der oben
*erzeugten (3000) Korrelationsmatrizen;
%Macro proccalis (dateinam=, lib=);
proc calis data=&lib..cm_b_&dateinam OUTEST=&lib..OUTE_cm_b_&dateinam tech=nr edf=999 nop;
         Lineqs
                  indi_1   = beta1  Faktor_1          + beta_f1          Error_1,
                  indi_2   = beta2  Faktor_1          + beta_f2          Error_2,
                  indi_3   = beta3  Faktor_1          + beta_f3          Error_3,
                  indi_4s_ = beta4  Faktor_2          + beta_f4          Error_4,
                  indi_5s_ = beta5  Faktor_2          + beta_f5          Error_5,
                  indi_6s_ = beta6  Faktor_2          + beta_f6          Error_6,
                  Faktor_2 = gamma  Faktor_1          + 0.707            Error_7;
         Std
                  Error_1 = 1,
                  Error_2 = 1,
                  Error_3 = 1,
                  Error_4 = 1,
                  Error_5 = 1,
                  Error_6 = 1,
                  Error_7 = 1,
                  Faktor_1 = 1;
         run;
%MEND proccalis;
%macro create10(howmany);
   %do i=1 %to &howmany;
      %proccalis (dateinam=iterat&i, lib=base_mak);
       run;
   %end;
%mend create10;
%create10(3000)

*5.;
*Von den OUTEST-datasets wird nur jeweils die interessierende Zeile Beobachtung) *mit dem _TYPE_ "PARMS" (=final
*Parameter Estimates) bewahrt;
%Macro parms_isol (dateinam=, lib=);
DATA &lib..PARMS_OUTE_cm_b_&dateinam;
SET &lib..OUTE_cm_b_&dateinam;
IF _TYPE_ NE "PARMS" THEN DELETE;
RUN;
TITLE;
   run;
%MEND parms_isol;
%macro create11(howmany);
   %do i=1 %to &howmany;
      %parms_isol (dateinam=ITERAT&i, lib=base_mak);
       run;
   %end;
%mend create11;
%create11(3000)
*6.;
*die einzeiligen Datensätze addieren;
%macro create12(howmany);
   %do i=2 %to &howmany;
   proc APPEND BASE=BASE_MAK.PARMS_OUTE_CM_B_ITERAT1
           DATA=BASE_MAK.PARMS_OUTE_CM_B_ITERAT&i;
```

```
      run;
   %end;
%mend create12;
%create12(3000)
*7.;
*Umbenennen der Datei;
DATA BASE_MAK.PARMS_ALL_ITERAT;
SET BASE_MAK.PARMS_OUTE_CM_B_ITERAT1;
RUN;
8.;
*die negativen Werte von beta und gamma werden in positive umgewandelt. Schließlich werden statistische Maßzahlen und
*Histogramme für dieinteressierenden Pfadkoeffizienten erstellt;
DATA BASE_MAK.PARMS_ALL_ITERAT_B;
SET BASE_MAK.PARMS_ALL_ITERAT;
beta1_b=ABS(beta1);
beta2_b=ABS(beta2);
beta3_b=ABS(beta3);
beta4_b=ABS(beta4);
beta5_b=ABS(beta5);
beta6_b=ABS(beta6);
gamma_b=ABS(gamma);
DROP beta1 beta2 beta3 beta4 beta5 beta6 gamma;
LABEL beta1_b='Betrag von beta1'
           beta2_b='Betrag von beta2'
           beta3_b='Betrag von beta3'
           beta4_b='Betrag von beta4'
           beta5_b='Betrag von beta5'
           beta6_b='Betrag von beta6'
           gamma_b='Betrag von gamma';
RUN;
PROC MEANS DATA=BASE_MAK.PARMS_ALL_ITERAT_B;
TITLE "(Betrag von)PARMS aus MODEL 4 mit 2 korrel Faktoren und unkorrel. Fehlern (3000 Iterationen)";
RUN;
PROC GCHART DATA=BASE_MAK.PARMS_ALL_ITERAT_B;
VBAR beta1_b beta2_b beta3_b beta4_b beta5_b beta6_b gamma_b;
TITLE "Betrag von PARMS aus MODEL 4 mit 2 korrel Faktoren und unkorrel. Fehlern (3000 Iterationen)";
RUN;
*9.;
*Die geschätzten Standardfehler der 3000 PROC CALIS OUTEST-Outputdateien
*isolieren, indem man nur die interessierende Zeile mit dem _TYPE_ "STDERR" be*wahrt;
%Macro se_isol (dateinam=, lib=);
DATA &lib..Stderr_OUTE_cm_b_&dateinam;
SET &lib..OUTE_cm_b_&dateinam;
IF _TYPE_ NE "STDERR" THEN DELETE;
RUN;
TITLE "STDERR von MODEL 4 mit 2 korrel Faktoren und unkorrelierten Fehlern &lib..OUTE_cm_b_&dateinam";
   run;
%MEND se_isol;
%macro create13(howmany);
   %do i=1 %to &howmany;
      %se_isol (dateinam=iterat&i, lib=base_mak);
       run;
   %end;
%mend create13;
%create13(3000)
*10.;
*die (3000) einzeiligen Datensätze addieren;
%macro create14(howmany);
   %do i=2 %to &howmany;
   proc APPEND BASE=base_mak.STDERR_OUTE_CM_B_ITERAT1
           DATA=base_mak.STDERR_OUTE_CM_B_ITERAT&i;
      run;
   %end;
%mend create14;
%create14(3000)
*11.;
*Umbenennen der Datei;
DATA base_mak.STDERR_ALL_ITERAT;
SET base_mak.STDERR_OUTE_CM_B_ITERAT1;
RUN;
*
*12.;
*statistische Maßzahlen für den SE ausgeben;
PROC MEANS DATA=base_mak.STDERR_ALL_ITERAT;
TITLE "STDERR aus MODEL 4 mit 2 korrel Faktoren und unkorrel. Fehlern (3000 Iterationen)";
RUN;
*Wiederaufhebung der Unterdrückung von – u. a. – den Angaben im log;
options notes source source2 errors=20;
```

11.5.1.3 Programmcodeausschnitt der Modelle 5 bis 9

Die folgenden Ausschnitte betreffen nur den *Schritt 4* des Programmcodes, in dem das Strukturgleichungsmodell definiert wird. Aus ihm lassen sich u. a. die im jeweiligen Modell verwendeten Indikatorvariablen und im COV Statement vorgenommene Fehlspezifikationen von Fehlerkorrelationen entnehmen. Da das STD Statement aller Modelle identisch mit jenem des Modells 4 ist, wird es nicht dokumentiert.

11.5.1.3.1 Modell 5

```
%Macro proccalis (dateinam=, lib=);
proc calis data=&lib..cm_b_&dateinam OUTEST=&lib..OUTE_cm_b_&dateinam tech=nr edf=999 nop;

Lineqs
                ind_1sp2 = beta1   Faktor_1          + beta_f1          Error_1,
                ind_2sp2 = beta2   Faktor_1          + beta_f2          Error_2,
                indi_3 =   beta3   Faktor_1          + beta_f3          Error_3,
                indi_4s_   beta4   Faktor_2          + beta_f4          Error_4,
                indi_5s_ = beta5   Faktor_2          + beta_f5          Error_5,
                indi_6s_ = beta6   Faktor_2          + beta_f6          Error_6,
                Faktor_2 = gamma   Faktor_1          + 0.798198         Error_7;
        Std
                ...
        Cov
                Error_1 Error_2 = 0;
        run;
        %MEND proccalis;
```

11.5.1.3.2 Modell 6

```
%Macro proccalis (dateinam=, lib=);
proc calis data=&lib..cm_b_&dateinam OUTEST=&lib..OUTE_cm_b_&dateinam tech=nr edf=999 nop;

Lineqs
                ind_1sp2 =         beta1    Faktor_1          + beta_f1          Error_1,
                ind_2sp2 =         beta2    Faktor_1          + beta_f2          Error_2,
                indi_3 =           beta3    Faktor_1          + beta_f3          Error_3,
                ind_4sp3 =         beta4    Faktor_2          + beta_f4          Error_4,
                ind_5sp3 =         beta5    Faktor_2          + beta_f5          Error_5,
                indi_6s_ =         beta6    Faktor_2          + beta_f6          Error_6,
                Faktor_2 =         gamma    Faktor_1          + 0.859            Error_7;
        Std
                ...
        Cov
                Error_1 Error_2 = 0,
                Error_4 Error_5 = 0;
        run;
%MEND proccalis;
```

[198] 0,798 ist das standardisierte Regressionsgewicht und entspricht dem unstandardisierten Regressionsgewicht 0,707.

11.5.1.3.3 Modell 7

```
%Macro proccalis (dateinam=, lib=);
proc calis data=&lib..cm_b_&dateinam OUTEST=&lib..OUTE_cm_b_&dateinam tech=nr edf=999 nop;
        Lineqs
                ind_1sp2 = beta1   Faktor_1          + beta_f1          Error_1,
                ind_2sp2 = beta2   Faktor_1          + beta_f2          Error_2,
                indi_3 =   beta3   Faktor_1          + beta_f3          Error_3,
                ind_4sp3 = beta4   Faktor_2          + beta_f4          Error_4,
                ind_5sp3 = beta5   Faktor_2          + beta_f5          Error_5,
                indi_6s_ = beta6   Faktor_2          + beta_f6          Error_6,
                Faktor_2 = gamma   Faktor_1          + 0.798            Error_7;
        Std
                ...
        Cov
                Error_1 Error_2 = 0.5,
                Error_4 Error_5 = 0;
        run;
%MEND proccalis;
```

11.5.1.3.4 Modell 8

```
%Macro proccalis (dateinam=, lib=);
proc calis data=&lib..cm_b_&dateinam OUTEST=&lib..OUTE_cm_b_&dateinam tech=nr edf=999 nop;
        Lineqs
                ind_1sp2 = beta1   Faktor_1          + beta_f1          Error_1,
                indi_2 =   beta2   Faktor_1          + beta_f2          Error_2,
                indi_3 =   beta3   Faktor_1          + beta_f3          Error_3,
                ind_4sp5 = beta4   Faktor_2          + beta_f4          Error_4,
                indi_5s_ = beta5   Faktor_2          + beta_f5          Error_5,
                indi_6s_ = beta6   Faktor_2          + beta_f6          Error_6,
                Faktor_2 = gamma   Faktor_1          + 0.614            Error_7;
        Std
                ...
        Cov
                Error_1 Error_4 = 0;
        run;
%MEND proccalis;
```

11.5.1.3.5 Modell 9

```
%Macro proccalis (dateinam=, lib=);
proc calis data=&lib..cm_b_&dateinam OUTEST=&lib..OUTE_cm_b_&dateinam tech=nr edf=999 nop;

Lineqs
                ind_1sp6 = beta1   Faktor_1          + beta_f1          Error_1,
                ind_2sp6 = beta2   Faktor_1          + beta_f2          Error_2,
                indi_3 =   beta3   Faktor_1          + beta_f3          Error_3,
                ind_4sp6 = beta4   Faktor_2          + beta_f4          Error_4,
                indi_5s_ = beta5   Faktor_2          + beta_f5          Error_5,
                indi_6s_ = beta6   Faktor_2          + beta_f6          Error_6,
                Faktor_2 = gamma   Faktor_1          + 0.716            Error_7;
        Std
                ...
        Cov
                Error_1 Error_2 = 0,
                Error_1 Error_4 = 0;
        run;
%MEND proccalis;
```

11.5.1.4 Programmcodeausschnitte der Varianten bisheriger Modelle und Zusatzmodelle

11.5.1.4.1 Modell 5 k

```
%Macro proccalis (dateinam=, lib=);
proc calis data=&lib..cm_b_&dateinam OUTEST=&lib..OUTE_cm_b_&dateinam tech=nr edf=999 nop;
        Lineqs
                ind_1sp2 = beta1   Faktor_1          + beta_f1          Error_1,
                ind_2sp2 = beta2   Faktor_1          + beta_f2          Error_2,
                indi_3 =   beta3   Faktor_1          + beta_f3          Error_3,
                indi_4s_ = beta4   Faktor_2          + beta_f4          Error_4,
                indi_5s_ = beta5   Faktor_2          + beta_f5          Error_5,
                indi_6s_ = beta6   Faktor_2          + beta_f6          Error_6,
                Faktor_2 = gamma   Faktor_1          + 0.707            Error_7;
        Std
                ...
```

```
        Cov
                Error_1 Error_2 = 0.50;
        run;
        %MEND proccalis;
```

11.5.1.4.2 Modell Z4 k

```
%Macro proccalis (dateinam=, lib=);
proc calis data=&lib..cm_b_&dateinam OUTEST=&lib..OUTE_cm_b_&dateinam tech=nr edf=999 nop;

Lineqs
                indi_1 = beta1     Faktor_1          + beta_f1          Error_1,
                indi_2 = beta2     Faktor_1          + beta_f2          Error_2,
                indi_3 = beta3     Faktor_1          + beta_f3          Error_3,
                ind_4sp3 = beta4   Faktor_2          + beta_f4          Error_4,
                ind_5sp3 = beta5   Faktor_2          + beta_f5          Error_5,
                indi_6s_ = beta6   Faktor_2          + beta_f6          Error_6,
                Faktor_2 =gamma    Faktor_1          + 0.707            Error_7;
        Std
                ...
        Cov
                Error_4 Error_5 = 0.5;
        run;
%MEND proccalis;
```

11.5.1.4.3 Modell Z4

```
%Macro proccalis (dateinam=, lib=);
proc calis data=&lib..cm_b_&dateinam OUTEST=&lib..OUTE_cm_b_&dateinam tech=nr edf=999 nop;
        Lineqs
                indi_1 = beta1     Faktor_1          + beta_f1          Error_1,
                indi_2 = beta2     Faktor_1          + beta_f2          Error_2,
                indi_3 = beta3     Faktor_1          + beta_f3          Error_3,
                ind_4sp3 = beta4   Faktor_2          + beta_f4          Error_4,
                ind_5sp3 = beta5   Faktor_2          + beta_f5          Error_5,
                indi_6s_ = beta6   Faktor_2          + beta_f6          Error_6,
                Faktor_2 = gamma   Faktor_1          + 0.798            Error_7;
        Std
        ...
        Cov
                Error_4 Error_5 = 0;
        run;
%MEND proccalis;
```

11.5.1.4.4 Modell 8 k

```
%Macro proccalis (dateinam=, lib=);
proc calis data=&lib..cm_b_&dateinam OUTEST=&lib..OUTE_cm_b_&dateinam tech=nr edf=999 nop;

Lineqs
                ind_1sp2 = beta1   Faktor_1          + beta_f1          Error_1,
                indi_2 = beta2     Faktor_1          + beta_f2          Error_2,
                indi_3 = beta3     Faktor_1          + beta_f3          Error_3,
                ind_4sp5 = beta4   Faktor_2          + beta_f4          Error_4,
                indi_5s_ = beta5   Faktor_2          + beta_f5          Error_5,
                indi_6s_ = beta6   Faktor_2          + beta_f6          Error_6,
                Faktor_2 = gamma   Faktor_1          + 0.707            Error_7;
        Std
                ...
        Cov
                Error_1 Error_4 = 0.5;
        run;
%MEND proccalis;
```

11.5.1.4.5 Modell 2- und 3 k_g

LINEQS-Statement identisch mit Modell 2, bis auf:

```
        Cov
                Error_1 Error_4 = Corr_14;
```

Modell 3 k_g

Dieses Modell erfordert der Identifizierbarkeit halber eine 7. Indikatorvariable, daher wird der Programmcode zu 1., 2., 4. und 8 dokumentiert.

```
*1.;

%MACRO machdat5 (i, seedx1, seedx2, seedx3, seedx4, seedx5 ,seedx6 ,seedx7 ,seedx8 , seedx9,
seedx10, seedx11, dateinam=);
%GLOBAL sds;
%LET sds=&dateinam;
DATA &sds;
Do i = 1 to &i;
x1 = Rannor (&seedx1);
x2 = Rannor (&seedx2);
x3 = Rannor (&seedx3);
x4 = Rannor (&seedx4);
x5 = Rannor (&seedx5);
x6 = Rannor (&seedx6);
x7 = Rannor (&seedx7);
x8 = Rannor (&seedx8);
x9 = Rannor (&seedx9);
x10 = Rannor (&seedx10);
x11 = Rannor (&seedx11);
OUTPUT;
END;
RUN;
*%PRINTITL;
%MEND machdat5;

*2.;

%Macro arit_op2 (dateinam=);
%GLOBAL sds;
%LET sds=&dateinam;
DATA b_&sds;
SET &sds;
indi_3 = x1 + x5;
indi_7 = x1 + x11;
indi_5 = x2 + x7;
indi_6 = x2 + x8;
ind_1sp6 = x1 + (x3 + x9 + x10) /SQRT(3);
ind_2sp6 = x1 + (x4 + 1.73*x10) /SQRT(4);
ind_4sp6 = x2 + (x6 + 1.73*x9) /SQRT(4);
RUN;
%MEND arit_op2;

*4.;

%Macro proccalis (dateinam=);
%GLOBAL sds;
%LET sds=&dateinam;
proc calis data=&sds OUTEST=OUTE_&sds tech=nr edf=999 nop;
         Lineqs
         ind_1sp6 =         beta1    Faktor_1          + beta_f1          Error_1,
         ind_2sp6 =         beta2    Faktor_1          + beta_f2          Error_2,
         indi_3 =           beta3    Faktor_1          + beta_f3          Error_3,
         Indi_7 =           beta7    Faktor_1          + beta_f7          Error_7,
         ind_4sp6 =         beta4    Faktor_2          + beta_f4          Error_4,
         indi_5 =           beta5    Faktor_2          + beta_f5          Error_5,
         indi_6 =           beta6    Faktor_2          + beta_f6          Error_6,
         Faktor_2 =         gamma    Faktor_1          +                  Error_8;

         Std
                  ...
         Cov
                  Error_1 Error_4 = Corr_14,
                  Error_1 Error_2 = Corr_12;
8.;

DATA PARMS_ALL_ITERAT_B;
SET PARMS_ALL_ITERAT;
Iteration=_N_;
beta1_b=ABS(beta1);
beta2_b=ABS(beta2);
beta3_b=ABS(beta3);
beta4_b=ABS(beta4);
beta5_b=ABS(beta5);
beta6_b=ABS(beta6);
beta7_b=ABS(beta7);
corr_14_b=ABS(corr_14);
corr_12_b=ABS(corr_12);
DROP beta1 beta2 beta3 beta4 beta5 beta6 beta7;
LABEL beta1_b='Betrag von beta1'
           beta2_b='Betrag von beta2'
           beta3_b='Betrag von beta3'
           beta4_b='Betrag von beta4'
           beta5_b='Betrag von beta5'
           beta6_b='Betrag von beta6'
```

```
           beta7_b='Zusatzindikator Messmodell 1'
           gamma='gamma'
           corr_14_b='Betr. d. Korr (Err 1, Err 4)'
           corr_12_b='Betr. d. Korr (Err 1, Err 2)';
TITLE "(außer gamma Betr. v.) PARMS aus MODEL 3 mit 2 unkorrel Faktoren und korrel. Fehlern (10000 Iterationen)";
RUN;
```

11.5.1.4.6 Model 8-, 5-, Z 4-, 6- und 9 k_g

```
LINEQS-Statement identisch mit Modell 8 k, bis auf:

Cov
                  Error_1 Error_4 = corr_14;
```

Modell 5_k_g

```
LINEQS-Statement identisch mit Modell 5 k, bis auf:
         Cov
                  Error_1 Error_2 = Corr_12;
```

Modell Z 4 k_g

```
LINEQS-Statement identisch mit Modell Z4 k, bis auf:

         Cov
                  Error_4 Error_5 = Corr_45;
```

Modell 6_k_g

```
Lineqs-Statement identisch mit Modell 6, bis auf:

         Faktor_2 = gamma Faktor_1   + 0.707 Error_7;

         Cov
                  Error_1 Error_2 = Corr_12,
                  Error_4 Error_5 = Corr_45;
```

Modell 9 k_g

```
Lineqs-Statement identisch mit Modell 9, bis auf:

         Faktor_2 = gamma Faktor_1   + 0.707 Error_7;

         Cov
                  Error_1 Error_2 = Corr_12,
                  Error_1 Error_4 = Corr_14;
```

11.5.2 B SAS-OUTPUT

11.5.2.1 Modelle mit unabhängigen Faktoren

11.5.2.1.1 Modell 1

(außer gamma Betrag v.) PARMS MODELL 1 m. 2 unkorr. Faktoren und unkorrel. Fehlern (3000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		3000	0.0078976	0.0039190	0.000605712	0.0270266
ITER	Iteration	0	.	.	.	.
beta_f1	beta_f1	3000	0.7035378	0.0631842	-0.7220447	0.8204324
beta_f2	beta_f2	3000	0.7048560	0.0510517	-0.7202284	0.7944140
beta_f3	beta_f3	3000	0.7040458	0.0677547	-0.7169598	0.7943725
beta_f4	beta_f4	3000	0.7003686	0.0958910	-0.7255636	0.7904586
beta_f5	beta_f5	3000	0.7004822	0.0961258	-0.7244252	0.7785816

beta_f6	beta_f6	3000	0.7018649	0.0809251	-0.7161624	0.7872295
gamma	gamma	3000	0.0015267	0.0430440	-0.1666160	0.1512420
beta1_b	Betrag von beta1	3000	0.7073604	0.0261732	0.5717435	0.8188455
beta2_b	Betrag von beta2	3000	0.7070898	0.0244299	0.6073767	0.7833479
beta3_b	Betrag von beta3	3000	0.7064528	0.0255915	0.6074310	0.7881797
beta4_b	Betrag von beta4	3000	0.7062190	0.0250299	0.6118980	0.7948298
beta5_b	Betrag von beta5	3000	0.7060858	0.0246892	0.6270561	0.7969553
beta6_b	Betrag von beta6	3000	0.7066084	0.0248590	0.6165655	0.7920979

STDERR aus MODELL 1 mit 2 unkorrel Faktoren und unkorrel. Fehlern (3000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta4	beta4	3000	0.0335763	0.000742589	0.0312237	0.0366960
beta5	beta5	3000	0.0335777	0.000747475	0.0312207	0.0370901
beta6	beta6	3000	0.0335796	0.000746272	0.0312038	0.0376275
beta1	beta1	3000	0.0335986	0.000793352	0.0313553	0.0375152
beta_f1	beta_f1	3000	0.0251594	0.0012457	0.0223371	0.0327358
beta2	beta2	3000	0.0336002	0.000802005	0.0313213	0.0372639
beta_f2	beta_f2	3000	0.0251358	0.0012061	0.0220324	0.0320572
beta3	beta3	3000	0.0335940	0.000793467	0.0313356	0.0368004
beta_f3	beta_f3	3000	0.0251076	0.0012006	0.0224022	0.0303735
beta_f4	beta_f4	3000	0.0251281	0.0011856	0.0224359	0.0310116
beta_f5	beta_f5	3000	0.0251245	0.0012081	0.0220707	0.0302862
beta_f6	beta_f6	3000	0.0251504	0.0012090	0.0223519	0.0320190
gamma	gamma	3000	0.0421751	0.000558690	0.0405556	0.0450090

11.5.2.1.2 Modell 2

(Betr. v.) PARMS aus MODELL 2 mit 2 unkorrel Faktoren und unkorrel. Fehlern (3000 Iterat.)
Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung
RHS		3000	0.1578149	0.0241197
ITER	Iteration	0	.	.
beta_f1	beta_f1	3000	0.6987937	0.0762600
beta_f2	beta_f2	3000	0.6990944	0.1171598
beta_f3	beta_f3	3000	0.7045548	0.0808894
beta_f4	beta_f4	3000	0.6818560	0.1733405
beta_f5	beta_f5	3000	0.7051406	0.0718408
beta_f6	beta_f6	3000	0.7035245	0.0848332
gamma	gamma	3000	0.0289106	0.0675556
beta1_b	Betrag von beta1	3000	0.7107871	0.0256069
beta2_b	Betrag von beta2	3000	0.7049126	0.0253703
beta3_b	Betrag von beta3	3000	0.7045727	0.0252825
beta4_b	Betrag von beta4	3000	0.7082751	0.0251944
beta5_b	Betrag von beta5	3000	0.7030916	0.0255532
beta6_b	Betrag von beta6	3000	0.7032477	0.0258838
gamma_b	Betrag von gamma	3000	0.0619747	0.0394679
gamma_b2	Betrag von gamma, wenn gamma < -0.0414	3000	0.0569121	0.0464737

Variable	Etikett	Minimum	Maximum
RHS		0.0882004	0.2523867
ITER	Iteration	.	.
beta_f1	beta_f1	-0.7214833	0.7852833
beta_f2	beta_f2	-0.7339374	0.7811571
beta_f3	beta_f3	-0.7166677	0.7886330
beta_f4	beta_f4	-0.7327439	0.7808693
beta_f5	beta_f5	-0.7092154	0.7777580
beta_f6	beta_f6	-0.7106210	0.7916312
gamma	gamma	-0.1831522	0.2290601
beta1_b	Betrag von beta1	0.6191366	0.8026644
beta2_b	Betrag von beta2	0.6243346	0.7818466
beta3_b	Betrag von beta3	0.6148642	0.8069238
beta4_b	Betrag von beta4	0.6241827	0.8033044
beta5_b	Betrag von beta5	0.6240558	0.8017968
beta6_b	Betrag von beta6	0.6070769	0.7957303
gamma_b	Betrag von gamma	0.000015582	0.2290601
gamma_b2	Betrag von gamma, wenn gamma < -0.0414	-0.0413876	0.2290601

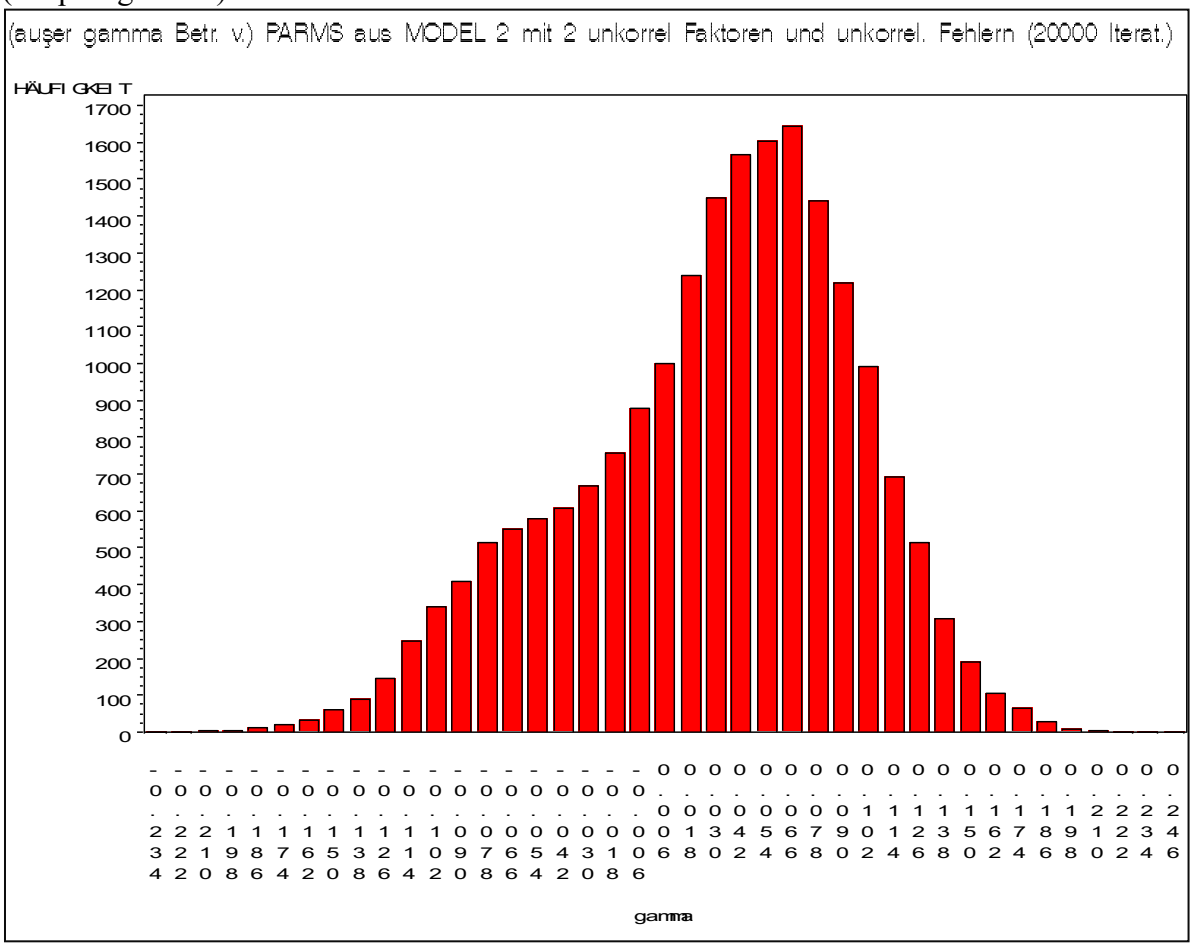
STDERR von fehlspez. MODELL 2 mit 2 unkorrel. Faktoren und korrel. Fehlern (3000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta4	beta4	3000	0.0335640	0.000788553	0.0313632	0.0373580
beta5	beta5	3000	0.0335356	0.000785732	0.0313853	0.0368195
beta6	beta6	3000	0.0335322	0.000769904	0.0313712	0.0368791
beta1	beta1	3000	0.0336175	0.000788997	0.0312227	0.0374219
beta_f1	beta_f1	3000	0.0253370	0.0012922	0.0223874	0.0331529
beta2	beta2	3000	0.0335861	0.000771547	0.0312191	0.0362899
beta_f2	beta_f2	3000	0.0250330	0.0011918	0.0221423	0.0326377
beta3	beta3	3000	0.0335844	0.000769308	0.0312464	0.0368164
beta_f3	beta_f3	3000	0.0250154	0.0011797	0.0221037	0.0309912
beta_f4	beta_f4	3000	0.0253072	0.0012990	0.0223056	0.0337416
beta_f5	beta_f5	3000	0.0250423	0.0012499	0.0220338	0.0326994
beta_f6	beta_f6	3000	0.0250391	0.0011571	0.0216812	0.0309779
gamma	gamma	3000	0.0423019	0.000600963	0.0406293	0.0446465

Das Histogramm der Originalwerte von gamma auf Basis von 20000 Iterationen (Abbildung 89) verdeutlicht, dass die Verteilung von gamma aus zwei Subpopulationen zusammengesetzt ist: Eine Verteilung von gamma im positiven Wertebereich, die den weitaus größten Anteil der Schätzungen umfasst, weist einen Häufigkeitsgipfel bei 0,66 auf. Für die zweite kleinere Verteilung von gamma im negativen Wertebereich, die auf softwareinterner Vorzeichenumkehr beruht, kann einen Häufigkeitsgipfel bei -0,66 *unterstellt* werden. Dieser ist nicht klar erkennbar, da die kleinere Subverteilung durch die größere überlagert ist.

Abbildung 89: Verteilung von gamma des Modells 2 auf Basis von 20000 Iterationen (Ursprungswerte)



11.5.2.1.3 Modell 3

(außer gamma Betr. v.) PARMS aus MODELL 3 mit 2 unkorrel Faktoren und korrel. Fehlern (3000 Iterat. Betrag, wenn gamma < -0.0244

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung
RHS		3000	0.2608524	0.0309281
ITER	Iteration	0	.	.
beta_f1	beta_f1	3000	0.4868713	0.0404885
beta_f2	beta_f2	3000	0.5090317	0.0381544
beta_f3	beta_f3	3000	0.8160366	0.0335856
beta_f4	beta_f4	3000	0.6704051	0.2001647
beta_f5	beta_f5	3000	0.6953914	0.1456212
beta_f6	beta_f6	3000	0.6970292	0.1344098
gamma	gamma	3000	0.0355895	0.0746556
beta1_b	Betrag von beta1	3000	0.8723062	0.0199888
beta2_b	Betrag von beta2	3000	0.8596741	0.0197974
beta3_b	Betrag von beta3	3000	0.5765088	0.0243796
beta4_b	Betrag von beta4	3000	0.7115382	0.0259495
beta5_b	Betrag von beta5	3000	0.7009150	0.0254590
beta6_b	Betrag von beta6	3000	0.7015195	0.0259487
gamma_b	Betrag von gamma	3000	0.0714981	0.0415549
gamma_b2	Betrag von gamma, wenn gamma < -0.0244	3000	0.0700850	0.0438971

Variable	Etikett	Minimum	Maximum
RHS		0.1628055	0.3863744
ITER	Iteration	.	.
beta_f1	beta_f1	-0.4957836	0.5925295
beta_f2	beta_f2	-0.4893761	0.6348492
beta_f3	beta_f3	-0.7648509	0.8682200
beta_f4	beta_f4	-0.7515477	0.7887474
beta_f5	beta_f5	-0.7793560	0.7930836
beta_f6	beta_f6	-0.7468047	0.7844722
gamma	gamma	-0.2070914	0.2414438
beta1_b	Betrag von beta1	0.8055488	0.9534320
beta2_b	Betrag von beta2	0.7726361	0.9178530
beta3_b	Betrag von beta3	0.4961795	0.6580429
beta4_b	Betrag von beta4	0.6147113	0.8009447
beta5_b	Betrag von beta5	0.6004399	0.8014360
beta6_b	Betrag von beta6	0.6134231	0.8118044
gamma_b	Betrag von gamma	0.000064444	0.2414438
gamma_b2	Betrag von gamma, wenn gamma < -0.0244	-0.0242902	0.2414438

STDERR aus MODEL 3 mit 2 unkorrel Faktoren und korrel. Fehlern (3000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta4	beta4	3000	0.0335306	0.000775558	0.0311002	0.0366487
beta5	beta5	3000	0.0334703	0.000755868	0.0311084	0.0363911
beta6	beta6	3000	0.0334753	0.000764748	0.0311084	0.0368978
beta1	beta1	3000	0.0299453	0.000731108	0.0279742	0.0327793
beta_f1	beta_f1	3000	0.0319647	0.0036274	0.0239881	0.0662975
beta2	beta2	3000	0.0299925	0.000717512	0.0278099	0.0329886
beta_f2	beta_f2	3000	0.0300801	0.0028168	0.0243007	0.0437385
beta3	beta3	3000	0.0309086	0.000354362	0.0297137	0.0320932
beta_f3	beta_f3	3000	0.0198545	0.000261687	0.0189214	0.0206486
beta_f4	beta_f4	3000	0.0254857	0.0014117	0.0220603	0.0344541
beta_f5	beta_f5	3000	0.0249148	0.0011420	0.0218645	0.0317474
beta_f6	beta_f6	3000	0.0249534	0.0011989	0.0221150	0.0308299
gamma	gamma	3000	0.0393344	0.000442750	0.0379002	0.0415070

Das erste folgende Histogramm (Abbildung 90) verdeutlicht auch schon bei regulärer Iterationsanzahl, dass es sich auch bei der Verteilung von gamma des Modells 3 um zwei Subverteilungen handelt – um eine größere mit dem Häufigkeitsgipfel bei 0,08 und eine kleinere mit Modus bei -0,08.

Das zweite (Abbildung 91) zeigt, dass es nicht angemessen ist, den Betrag von gamma zu verwenden, da die resultierende Verteilung links abgeschnitten und stark rechtsschief ist sowie einen zu hohen Mittelwert aufweist.

Das dritte (Abbildung 92) demonstriert, dass die bereichsabhängige Betragsbildung ein angemessenes Vorgehen darstellt. Die resultierende Verteilung nähert jene hypothetische Verteilung an, in der eine softwarebedingte Vorzeichenumkehr bei der Schätzung nicht vorkommt (größere Subverteilung in Abbildung 90).

Abbildung 90: Verteilung von gamma des Modells 3 (Ursprungswerte)

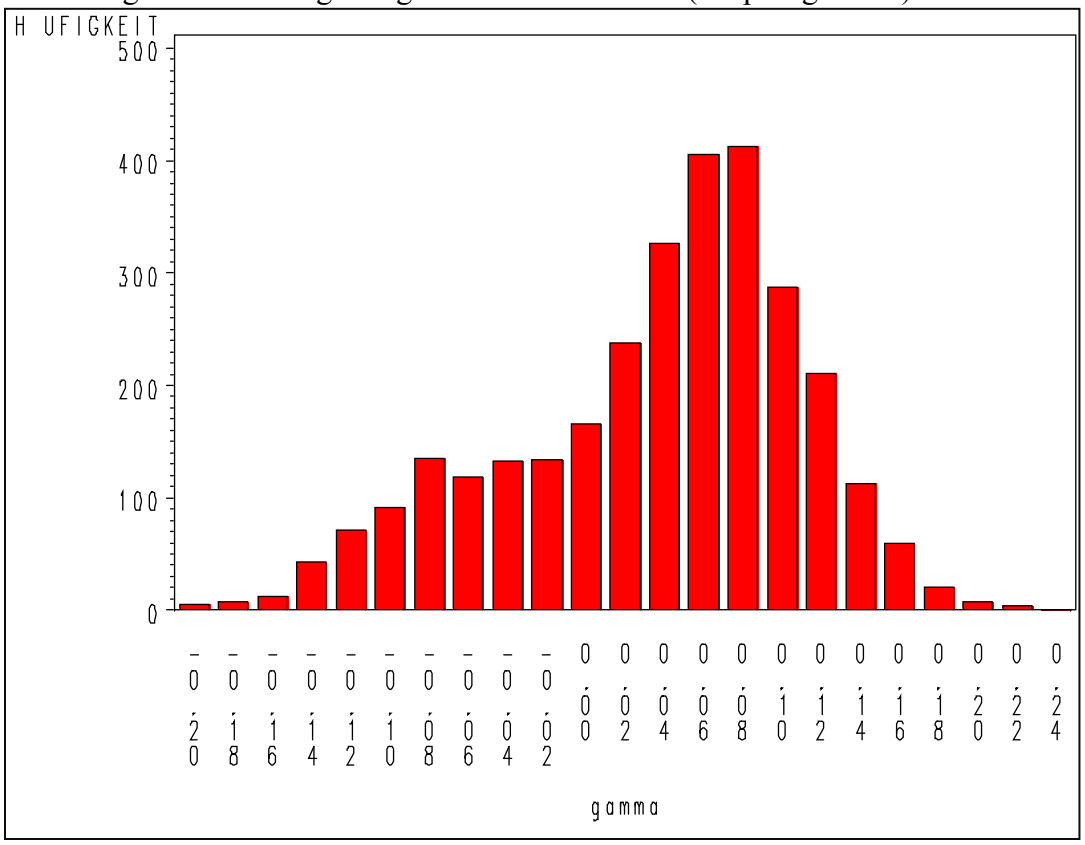


Abbildung 91: Verteilung von gamma des Modells 3 (absolute Werte)

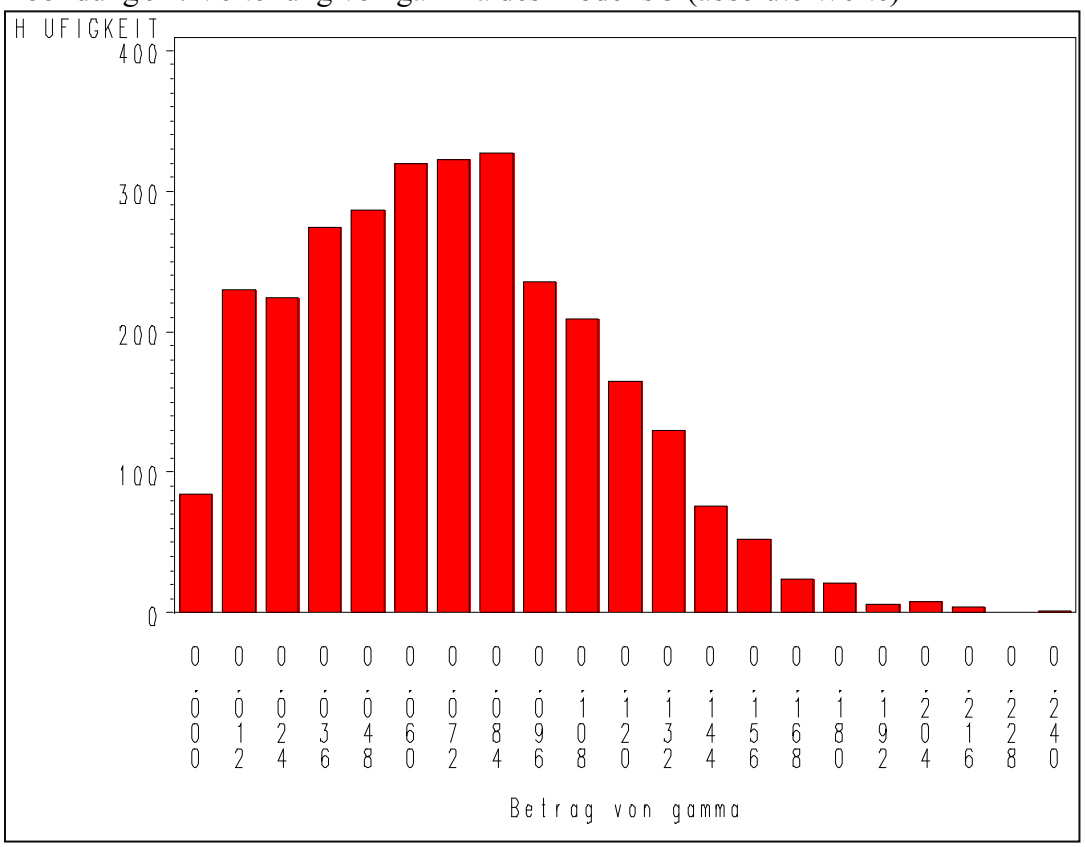
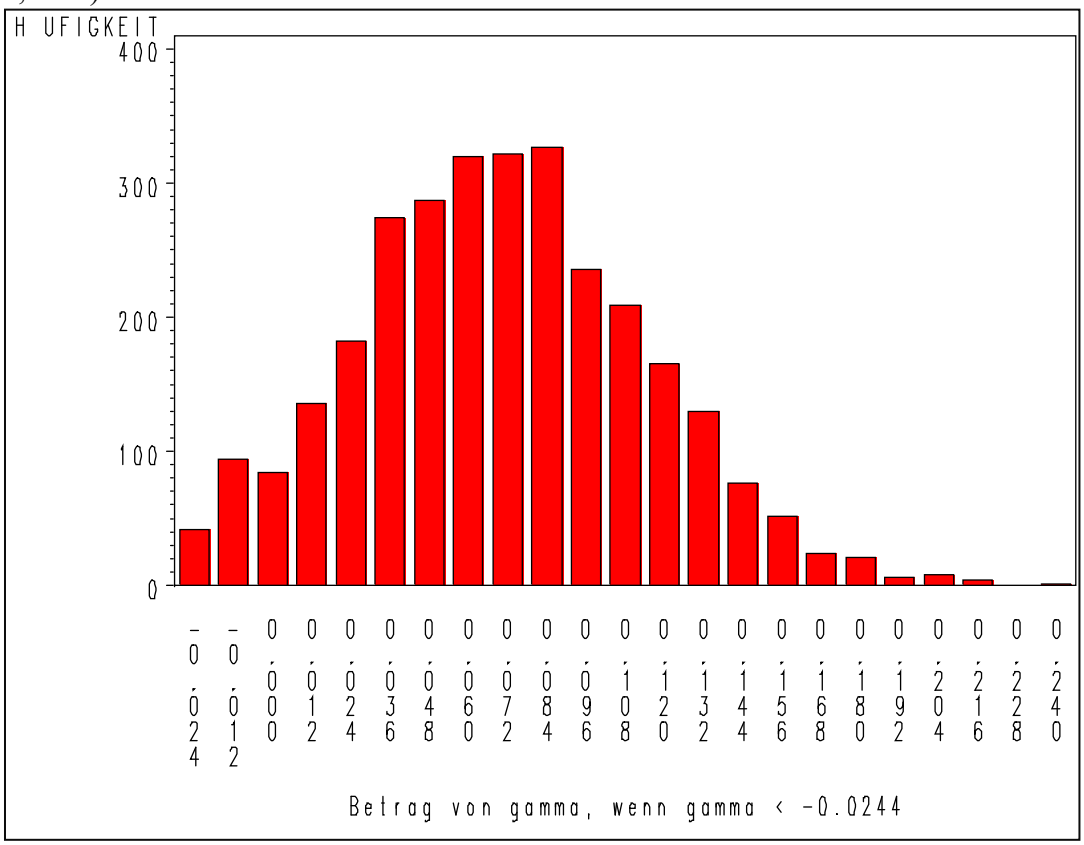


Abbildung 92: Verteilung von gamma des Modells 3 (absolute Werte, wenn gamma < -0,0244)



11.5.2.2 Modelle mit abhängigen Faktoren

11.5.2.2.1 Modell 4

(Betrag von)PARMS aus MODEL 4 mit 2 korrel Faktoren und unkorrel. Fehlern (3000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		3000	0.0080933	0.0039923	0.000266387	0.0275543
ITER	Iteration	0	.	.	.	.
beta_f1	beta_f1	3000	0.7060131	0.0223290	0.6291505	0.7722130
beta_f2	beta_f2	3000	0.7071232	0.0218877	0.6273635	0.7854143
beta_f3	beta_f3	3000	0.7059617	0.0224853	0.6348849	0.7771148
beta_f4	beta_f4	3000	0.7056560	0.0219939	0.6363160	0.7799888
beta_f5	beta_f5	3000	0.7060516	0.0224733	0.6263722	0.7755763
beta_f6	beta_f6	3000	0.7061134	0.0225685	0.6233106	0.7819770
beta1_b	Betrag von beta1	3000	0.7074948	0.0223264	0.6353638	0.7772835
beta2_b	Betrag von beta2	3000	0.7064141	0.0218436	0.6189704	0.7787265
beta3_b	Betrag von beta3	3000	0.7075374	0.0224434	0.6293589	0.7726067
beta4_b	Betrag von beta4	3000	0.7065567	0.0351689	0.5694674	0.8252799
beta5_b	Betrag von beta5	3000	0.7061726	0.0361763	0.5516722	0.8281429
beta6_b	Betrag von beta6	3000	0.7060458	0.0351637	0.5834647	0.8467732
gamma_b	Betrag von gamma	3000	0.7109483	0.0570929	0.5201038	0.9424668

STDERR aus MODEL 4 mit 2 korrel Faktoren und unkorrel. Fehlern (3000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta4	beta4	3000	0.0401483	0.000912424	0.0367784	0.0437311
beta5	beta5	3000	0.0401359	0.000906045	0.0370833	0.0436952
beta6	beta6	3000	0.0401389	0.000904574	0.0371958	0.0441075
beta1	beta1	3000	0.0316153	0.000536001	0.0296341	0.0336239
beta_f1	beta_f1	3000	0.0223823	0.000559620	0.0205063	0.0257072
beta2	beta2	3000	0.0316232	0.000531774	0.0296600	0.0335837
beta_f2	beta_f2	3000	0.0223581	0.000560744	0.0208677	0.0247260
beta3	beta3	3000	0.0316139	0.000541400	0.0297408	0.0336215
beta_f3	beta_f3	3000	0.0223801	0.000540296	0.0205941	0.0246989
beta_f4	beta_f4	3000	0.0223717	0.000557901	0.0204941	0.0249378
beta_f5	beta_f5	3000	0.0223661	0.000566225	0.0208770	0.0252134
beta_f6	beta_f6	3000	0.0223622	0.000554574	0.0208511	0.0245831
gamma	gamma	3000	0.0562408	0.0049152	0.0443892	0.0820992

11.5.2.2.2 Modell 5

(Betrag von)PARMS aus MODEL 5 mit 2 korrel Faktoren und korrel. Fehlern (3000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		3000	0.0451961	0.0129395	0.0119167	0.0982815
ITER	Iteration	0	.	.	.	.
beta_f1	beta_f1	3000	0.5053777	0.0357718	-0.5123213	0.5785463
beta_f2	beta_f2	3000	0.5065445	0.0297934	-0.4726077	0.5821700
beta_f3	beta_f3	3000	0.8024713	0.0176081	0.7432124	0.8566473
beta_f4	beta_f4	3000	0.7056769	0.0223240	0.6318753	0.7789959
beta_f5	beta_f5	3000	0.7059051	0.0229831	0.6303596	0.7785629
beta_f6	beta_f6	3000	0.7061338	0.0230904	0.6250124	0.7846643
beta1_b	Betrag von beta1	3000	0.8620391	0.0142443	0.8156496	0.9183394
beta2_b	Betrag von beta2	3000	0.8615860	0.0139649	0.8130671	0.9158980
beta3_b	Betrag von beta3	3000	0.5959571	0.0237725	0.5159025	0.6690555
beta4_b	Betrag von beta4	3000	0.7064710	0.0289485	0.5910375	0.7994956
beta5_b	Betrag von beta5	3000	0.7062270	0.0299954	0.5830470	0.8064882
beta6_b	Betrag von beta6	3000	0.7059450	0.0290484	0.6078146	0.8102388
gamma_b	Betrag von gamma	3000	0.6059518	0.0463635	0.4560241	0.7758776

STDERR aus MODEL 5 mit 2 korrel Faktoren und korrel. Fehlern (3000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta4	beta4	3000	0.0352934	0.000905487	0.0321608	0.0385723
beta5	beta5	3000	0.0352907	0.000908377	0.0320374	0.0387546
beta6	beta6	3000	0.0352907	0.000899654	0.0323081	0.0385956
beta1	beta1	3000	0.0281070	0.000448395	0.0266627	0.0300140
beta_f1	beta_f1	3000	0.0244499	0.0013643	0.0205733	0.0331619
beta2	beta2	3000	0.0281116	0.000444944	0.0265203	0.0300532
beta_f2	beta_f2	3000	0.0244082	0.0013669	0.0208235	0.0317955
beta3	beta3	3000	0.0305980	0.000342349	0.0294834	0.0319130
beta_f3	beta_f3	3000	0.0196217	0.000262611	0.0186585	0.0205345
beta_f4	beta_f4	3000	0.0228290	0.000637232	0.0208388	0.0257360
beta_f5	beta_f5	3000	0.0228267	0.000647525	0.0211512	0.0256792
beta_f6	beta_f6	3000	0.0228178	0.000633606	0.0210568	0.0256873
gamma	gamma	3000	0.0448671	0.0022285	0.0389800	0.0539124

11.5.2.2.3 Modell 6

(Betrag von)PARMS aus fehlspez. MODELL 6 mit 2 korrel Faktoren und korrel. Fehlern (3000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		3000	0.0719424	0.0161890	0.0229711	0.1512491
ITER	Iteration	0	.	.	.	.
beta_f1	beta_f1	3000	0.5042993	0.0253129	0.4003682	0.5862832
beta_f2	beta_f2	3000	0.5052583	0.0248668	0.4072871	0.5828079
beta_f3	beta_f3	3000	0.8052670	0.0175081	0.7468767	0.8611424
beta_f4	beta_f4	3000	0.5046404	0.0242487	0.4194684	0.5805891
beta_f5	beta_f5	3000	0.5044235	0.0250940	0.4192483	0.5885359
beta_f6	beta_f6	3000	0.8054596	0.0170889	0.7436859	0.8648250
beta1_b	Betrag von beta1	3000	0.8630318	0.0147619	0.8101062	0.9163544
beta2_b	Betrag von beta2	3000	0.8624882	0.0144923	0.8126100	0.9133002
beta3_b	Betrag von beta3	3000	0.5921731	0.0238711	0.5083637	0.6649625
beta4_b	Betrag von beta4	3000	0.8615647	0.0223735	0.7782857	0.9400141
beta5_b	Betrag von beta5	3000	0.8616692	0.0230601	0.7763520	0.9388036
beta6_b	Betrag von beta6	3000	0.5909032	0.0225836	0.5022516	0.6643123
gamma_b	Betrag von gamma	3000	0.5146033	0.0398809	0.3900390	0.6544650

STDERR aus fehlspez. MODELL 6 mit 2 korrel Faktoren und korrel. Fehlern (3000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta4	beta4	3000	0.0301858	0.000699159	0.0279073	0.0333644
beta5	beta5	3000	0.0301854	0.000698500	0.0277510	0.0332051
beta6	beta6	3000	0.0313603	0.000669474	0.0291291	0.0336540
beta1	beta1	3000	0.0283468	0.000480171	0.0267089	0.0304450
beta_f1	beta_f1	3000	0.0253787	0.0015559	0.0211391	0.0351317
beta2	beta2	3000	0.0283519	0.000476953	0.0266477	0.0306134
beta_f2	beta_f2	3000	0.0253243	0.0015553	0.0215582	0.0334766
beta3	beta3	3000	0.0306581	0.000343564	0.0295386	0.0319477
beta_f3	beta_f3	3000	0.0196676	0.000261627	0.0187040	0.0205556
beta_f4	beta_f4	3000	0.0253454	0.0015122	0.0216525	0.0318238
beta_f5	beta_f5	3000	0.0253614	0.0015380	0.0217579	0.0324358
beta_f6	beta_f6	3000	0.0196678	0.000254231	0.0187357	0.0205270
gamma	gamma	3000	0.0387314	0.0011670	0.0352259	0.0434996

11.5.2.2.4 Modell 7

Betrag v. PARMS aus fehlspez. MODELL 7 mit 2 korrel Faktoren und korrel. Fehlern (3000 Iteration

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		3000	0.0414651	0.0124265	0.0092720	0.1005457
ITER	Iteration	0	.	.	.	.
beta_f1	beta_f1	3000	0.7043427	0.0304772	0.5738364	0.7955292
beta_f2	beta_f2	3000	0.7049009	0.0304695	0.5714167	0.7905211
beta_f3	beta_f3	3000	0.7017847	0.0812298	-0.7598350	0.7913231
beta_f4	beta_f4	3000	0.5055776	0.0237827	0.4244692	0.5786031
beta_f5	beta_f5	3000	0.5053904	0.0244944	0.4239233	0.5838653
beta_f6	beta_f6	3000	0.8038594	0.0171687	0.7429398	0.8641362
r_err_12	Korrel. Fehler 1 u. 2	3000	0.4954012	0.0428734	0.3014435	0.6113875
beta1_b	Betrag von beta1	3000	0.7085679	0.0300774	0.6059153	0.8189699
beta2_b	Betrag von beta2	3000	0.7080119	0.0301004	0.6124347	0.8206601
beta3_b	Betrag von beta3	3000	0.7071585	0.0287950	0.6113982	0.8073904
beta4_b	Betrag von beta4	3000	0.8617500	0.0286790	0.7563501	0.9642084
beta5_b	Betrag von beta5	3000	0.8618387	0.0290384	0.7619696	0.9639498
beta6_b	Betrag von beta6	3000	0.5935048	0.0244299	0.4936338	0.6759606
gamma_b	Betrag von gamma	3000	0.6039793	0.0483378	0.4389651	0.8047550

STDERR aus fehlspez. MODELL 6 mit 2 korrel Faktoren und korrel. Fehlern (3000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta4	beta4	3000	0.0341723	0.000725942	0.0319701	0.0374420
beta5	beta5	3000	0.0341736	0.000718384	0.0316857	0.0371845
beta6	beta6	3000	0.0332548	0.000609844	0.0309360	0.0356130
beta1	beta1	3000	0.0376315	0.0014081	0.0336727	0.0463491
beta_f1	beta_f1	3000	0.0305428	0.0027274	0.0248115	0.0479881
beta2	beta2	3000	0.0376375	0.0014058	0.0333594	0.0463410
beta_f2	beta_f2	3000	0.0305156	0.0027089	0.0248550	0.0528814
beta3	beta3	3000	0.0364271	0.0012102	0.0326937	0.0420818
beta_f3	beta_f3	3000	0.0288840	0.0020829	0.0243488	0.0457075
beta_f4	beta_f4	3000	0.0247905	0.0014045	0.0211399	0.0307165
beta_f5	beta_f5	3000	0.0248032	0.0014191	0.0210465	0.0310092
beta_f6	beta_f6	3000	0.0196413	0.000254939	0.0187043	0.0205086
gamma	gamma	3000	0.0481016	0.0029337	0.0409262	0.0606593
r_err_12	r_err_12	3000	0.0427778	0.0079038	0.0271451	0.0994024

11.5.2.2.5 Modell 8

Betrag v. PARMS aus fehlspez. MODELL 7 mit 2 korrel Faktoren und korrel. Fehlern (3000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		3000	0.1593577	0.0252246	0.0854051	0.2722005
ITER	Iteration	0	.	.	.	.
beta_f1	beta_f1	3000	0.6242356	0.0299000	0.5116122	0.7228928
beta_f2	beta_f2	3000	0.7476975	0.0223617	0.6644349	0.8200462
beta_f3	beta_f3	3000	0.7469485	0.0225784	0.6722740	0.8251087
beta_f4	beta_f4	3000	0.6237588	0.0296350	0.4998763	0.7119183
beta_f5	beta_f5	3000	0.7467562	0.0225003	0.6601967	0.8169890
beta_f6	beta_f6	3000	0.7470053	0.0226466	0.6563477	0.8213671
beta1_b	Betrag von beta1	3000	0.7802971	0.0239321	0.6909602	0.8592165
beta2_b	Betrag von beta2	3000	0.6631854	0.0251775	0.5722973	0.7473461
beta3_b	Betrag von beta3	3000	0.6640101	0.0254788	0.5649740	0.7403025
beta4_b	Betrag von beta4	3000	0.7783951	0.0424620	0.6135553	0.9283769
beta5_b	Betrag von beta5	3000	0.6627503	0.0465103	0.4913388	0.8177801
beta6_b	Betrag von beta6	3000	0.6624100	0.0459930	0.5079773	0.8291437
gamma_b	Betrag von gamma	3000	0.7951177	0.0680550	0.5865776	1.1207509

STDERR aus fehlspez. MODELL 6 mit 2 korrel Faktoren und korrel. Fehlern (3000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta4	beta4	3000	0.0491448	0.0018578	0.0441527	0.0572666
beta5	beta5	3000	0.0440326	0.0012575	0.0393257	0.0491703
beta6	beta6	3000	0.0440272	0.0012513	0.0399363	0.0484658
beta1	beta1	3000	0.0304731	0.000570607	0.0285849	0.0328488
beta_f1	beta_f1	3000	0.0237052	0.0010311	0.0210382	0.0315674
beta2	beta2	3000	0.0314305	0.000467415	0.0297710	0.0333283
beta_f2	beta_f2	3000	0.0210736	0.000381925	0.0198471	0.0228550
beta3	beta3	3000	0.0314217	0.000475134	0.0297457	0.0332464
beta_f3	beta_f3	3000	0.0210769	0.000376309	0.0198272	0.0224608
beta_f4	beta_f4	3000	0.0236969	0.0010287	0.0208689	0.0299078
beta_f5	beta_f5	3000	0.0210701	0.000382732	0.0198517	0.0228256
beta_f6	beta_f6	3000	0.0210665	0.000378911	0.0199627	0.0225195
gamma	gamma	3000	0.0658374	0.0084460	0.0477217	0.1155092

11.5.2.2.6 Modell 9

Betrag v. PARMS aus fehlspez. MODELL 9 mit 2 korrel Faktoren und korrel. Fehlern (3000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		3000	0.2762747	0.0329854	0.1813444	0.4207012
ITER	Iteration	0	.	.	.	.
beta_f1	beta_f1	3000	0.3684730	0.0480740	0.0259779	0.5055675
beta_f2	beta_f2	3000	0.5947301	0.0243522	0.5044396	0.6833900
beta_f3	beta_f3	3000	0.8229564	0.0194893	0.7399304	0.8956459
beta_f4	beta_f4	3000	0.5857266	0.0385444	0.4366583	0.7043382
beta_f5	beta_f5	3000	0.7638907	0.0244865	0.6702477	0.8460731
beta_f6	beta_f6	3000	0.7640638	0.0247085	0.6739692	0.8446301
beta1_b	Betrag von beta1	3000	0.9282133	0.0183708	0.8627870	0.9996626
beta2_b	Betrag von beta2	3000	0.8033536	0.0180664	0.7300535	0.8634470
beta3_b	Betrag von beta3	3000	0.5670588	0.0284183	0.4447679	0.6726834
beta4_b	Betrag von beta4	3000	0.8091107	0.0349776	0.7037590	0.9387018
beta5_b	Betrag von beta5	3000	0.6444476	0.0378330	0.5133918	0.7807199
beta6_b	Betrag von beta6	3000	0.6442007	0.0378134	0.5205127	0.7926747
gamma_b	Betrag von gamma	3000	0.6985761	0.0473518	0.5680933	0.8715809

STDERR aus fehlspez. MODELL 9 mit 2 korrel Faktoren und korrel. Fehlern (3000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta4	beta4	3000	0.0391363	0.0012285	0.0354148	0.0442737
beta5	beta5	3000	0.0356093	0.0010891	0.0320248	0.0393062
beta6	beta6	3000	0.0356088	0.0010883	0.0320220	0.0390556
beta1	beta1	3000	0.0268574	0.000470228	0.0253724	0.0285929
beta_f1	beta_f1	3000	0.0330091	0.0109683	0.0226021	0.5396792
beta2	beta2	3000	0.0282131	0.000369479	0.0268816	0.0296100
beta_f2	beta_f2	3000	0.0198609	0.000719317	0.0179997	0.0234218
beta3	beta3	3000	0.0305513	0.000307390	0.0293761	0.0316162
beta_f3	beta_f3	3000	0.0195986	0.000267511	0.0186310	0.0204972
beta_f4	beta_f4	3000	0.0260246	0.0019776	0.0219294	0.0376206
beta_f5	beta_f5	3000	0.0210643	0.000387805	0.0198951	0.0227253
beta_f6	beta_f6	3000	0.0210624	0.000385393	0.0199641	0.0227467
gamma	gamma	3000	0.0474205	0.0030075	0.0400664	0.0606285

11.5.2.3 Varianten bisheriger Modelle und Zusatzmodelle

11.5.2.3.1 Modell 5 k

(Betr. v.)PARMS aus MODELL 5 mit zwei korrel Faktoren und korr. spez. korrel. Fehlern (3000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		3000	0.0087618	0.0071981	0.000485282	0.0868493
ITER	Iteration	0	.	.	.	.
beta_f1	beta_f1	3000	0.6951987	0.1063944	-0.4746952	0.7733780
beta_f2	beta_f2	3000	0.6993378	0.0787285	-0.4434691	0.7753615
beta_f3	beta_f3	3000	0.7075886	0.0284998	0.6247693	0.8682450
beta_f4	beta_f4	3000	0.7056858	0.0220990	0.6395088	0.7799811
beta_f5	beta_f5	3000	0.7059403	0.0226888	0.6275194	0.7774472
beta_f6	beta_f6	3000	0.7061515	0.0228458	0.6254436	0.7840354
beta1_b	Betrag von beta1	3000	0.7103412	0.0304696	0.6322500	0.9303208
beta2_b	Betrag von beta2	3000	0.7098687	0.0306997	0.6282123	0.9394393
beta3_b	Betrag von beta3	3000	0.7054045	0.0301851	0.4961357	0.7808094
beta4_b	Betrag von beta4	3000	0.7078706	0.0389107	0.5685228	0.8989320
beta5_b	Betrag von beta5	3000	0.7076042	0.0395784	0.5627639	0.9098914
beta6_b	Betrag von beta6	3000	0.7073462	0.0391207	0.5709224	0.8948305
gamma_b	Betrag von gamma	3000	0.7085555	0.0656957	0.4119851	0.9755922

Stderr aus MODELL 5 mit zwei korrel Faktoren und korr. spez. korrel. Fehlern (3000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta4	beta4	3000	0.0415932	0.0010113	0.0380102	0.0458902
beta5	beta5	3000	0.0415845	0.0010012	0.0377928	0.0453093
beta6	beta6	3000	0.0415815	0.0010045	0.0380380	0.0457368
beta1	beta1	3000	0.0315993	0.000725867	0.0266645	0.0341816
beta_f1	beta_f1	3000	0.0195119	0.0013771	0.0175306	0.0342617
beta2	beta2	3000	0.0315985	0.000734483	0.0264641	0.0341507
beta_f2	beta_f2	3000	0.0194968	0.0014096	0.0177283	0.0357918
beta3	beta3	3000	0.0332388	0.000799205	0.0293043	0.0363477
beta_f3	beta_f3	3000	0.0246380	0.0011150	0.0185058	0.0291213
beta_f4	beta_f4	3000	0.0225618	0.000590919	0.0206261	0.0252411
beta_f5	beta_f5	3000	0.0225586	0.000596351	0.0208997	0.0249846
beta_f6	beta_f6	3000	0.0225522	0.000590554	0.0208716	0.0250239
gamma	gamma	3000	0.0598678	0.0065671	0.0319410	0.0970265

11.5.2.3.2 Modell Z4 k

(Betr. v.) PARMS aus korr. spez. MODELL Z4 mit 2 korrel Faktoren und korrel. Fehlern (3000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		3000	0.0087985	0.0077072	0.000565197	0.1064884
ITER	Iteration	0	.	.	.	.
beta_f1	beta_f1	3000	0.7059434	0.0226005	0.6252812	0.7763088
beta_f2	beta_f2	3000	0.7071222	0.0221544	0.6248687	0.7863774
beta_f3	beta_f3	3000	0.7059765	0.0225856	0.6348203	0.7781136
beta_f4	beta_f4	3000	0.6973891	0.0888522	-0.4932623	0.7744162
beta_f5	beta_f5	3000	0.6943158	0.1101662	-0.4694834	0.7797940
beta_f6	beta_f6	3000	0.7077193	0.0287955	0.6112209	0.8588511
beta1_b	Betrag von beta1	3000	0.7075477	0.0225766	0.6303528	0.7803996
beta2_b	Betrag von beta2	3000	0.7063984	0.0221114	0.6177464	0.7807299
beta3_b	Betrag von beta3	3000	0.7075161	0.0225494	0.6281235	0.7726598
beta4_b	Betrag von beta4	3000	0.7114500	0.0555162	0.5820548	1.1521582
beta5_b	Betrag von beta5	3000	0.7114983	0.0559847	0.5560595	1.1061130
beta6_b	Betrag von beta6	3000	0.7054937	0.0398296	0.5415516	0.8306373
gamma_b	Betrag von gamma	3000	0.7075580	0.0664788	0.3948319	0.9733004

STDERR aus korr. spez. MODELL Z4 mit 2 korrel Faktoren und korrel. Fehlern (3000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta4	beta4	3000	0.0417919	0.0013054	0.0338735	0.0463842
beta5	beta5	3000	0.0417916	0.0012894	0.0342040	0.0459200
beta6	beta6	3000	0.0439151	0.0015574	0.0341905	0.0504234
beta1	beta1	3000	0.0317524	0.000550032	0.0297243	0.0337718
beta_f1	beta_f1	3000	0.0225807	0.000598527	0.0206182	0.0262125
beta2	beta2	3000	0.0317600	0.000546428	0.0297594	0.0337708
beta_f2	beta_f2	3000	0.0225526	0.000597129	0.0210331	0.0255031
beta3	beta3	3000	0.0317510	0.000554983	0.0298298	0.0337979
beta_f3	beta_f3	3000	0.0225747	0.000568443	0.0207167	0.0250384
beta_f4	beta_f4	3000	0.0195223	0.0014584	0.0177767	0.0345045
beta_f5	beta_f5	3000	0.0195273	0.0014564	0.0178054	0.0355503
beta_f6	beta_f6	3000	0.0246349	0.0011444	0.0181522	0.0297839
gamma	gamma	3000	0.0598055	0.0066903	0.0316624	0.0927653

11.5.2.3.3 Modell Z4

(Betr. v.) PARMS aus fehlspez. MODELL Z4 mit 2 korrel Faktoren und korrel. Fehlern (3000 Iterationen

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		3000	0.0455673	0.0130190	0.0116857	0.0982189
ITER	Iteration	0	.	.	.	.
beta_f1	beta_f1	3000	0.7059199	0.0228765	0.6185511	0.7779945
beta_f2	beta_f2	3000	0.7071497	0.0223906	0.6217537	0.7851641
beta_f3	beta_f3	3000	0.7059018	0.0228749	0.6323954	0.7769833
beta_f4	beta_f4	3000	0.5061995	0.0235456	0.4287429	0.5774093
beta_f5	beta_f5	3000	0.5061463	0.0241944	0.4184106	0.5824565
beta_f6	beta_f6	3000	0.8026952	0.0172291	0.7403809	0.8627402
beta1_b	Betrag von beta1	3000	0.7075538	0.0228404	0.6282711	0.7857446
beta2_b	Betrag von beta2	3000	0.7063562	0.0223407	0.6192877	0.7832129
beta3_b	Betrag von beta3	3000	0.7075723	0.0228307	0.6295212	0.7746458
beta4_b	Betrag von beta4	3000	0.8610202	0.0272787	0.7644522	0.9589567
beta5_b	Betrag von beta5	3000	0.8610447	0.0280130	0.7424228	0.9584395
beta6_b	Betrag von beta6	3000	0.5948181	0.0237529	0.4904882	0.6749008
gamma_b	Betrag von gamma	3000	0.6046712	0.0457713	0.4553274	0.7903117

STDERR aus fehlspez. MODELL Z4 mit 2 korrel Faktoren und korrel. Fehlern (3000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta4	beta4	3000	0.0328591	0.000654347	0.0303709	0.0358535
beta5	beta5	3000	0.0328574	0.000648046	0.0305432	0.0356886
beta6	beta6	3000	0.0326358	0.000669283	0.0301784	0.0349840
beta1	beta1	3000	0.0319429	0.000567752	0.0297826	0.0341653
beta_f1	beta_f1	3000	0.0228526	0.000648629	0.0209265	0.0270344
beta2	beta2	3000	0.0319497	0.000564646	0.0299663	0.0341612
beta_f2	beta_f2	3000	0.0228202	0.000646972	0.0211720	0.0262958
beta3	beta3	3000	0.0319407	0.000573005	0.0300710	0.0341798
beta_f3	beta_f3	3000	0.0228467	0.000612177	0.0208560	0.0256388
beta_f4	beta_f4	3000	0.0244453	0.0013379	0.0211940	0.0302213
beta_f5	beta_f5	3000	0.0244520	0.0013645	0.0205994	0.0303283
beta_f6	beta_f6	3000	0.0196226	0.000255429	0.0186843	0.0204870
gamma	gamma	3000	0.0448229	0.0022143	0.0384608	0.0555026

11.5.2.3.4 Modell 8 k

(Betrag von) PARMS aus MODELL 8 korr. spez. mit 2 korrel Faktoren und 2 pos. korrel. Fehlern (e1, e4) (3000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		3000	0.0080720	0.0039943	0.000513011	0.0270166
ITER	Iteration	0	.	.	.	.
beta_f1	beta_f1	3000	0.7057076	0.0226987	0.6210937	0.7758116
beta_f2	beta_f2	3000	0.7071069	0.0224134	0.6129461	0.7834007
beta_f3	beta_f3	3000	0.7060039	0.0227938	0.6280439	0.7837073
beta_f4	beta_f4	3000	0.7052379	0.0223795	0.6258875	0.7711367
beta_f5	beta_f5	3000	0.7059330	0.0226615	0.6259972	0.7801840
beta_f6	beta_f6	3000	0.7062275	0.0229752	0.6169501	0.7877147
beta1_b	Betrag von beta1	3000	0.7075271	0.0227062	0.6222946	0.7850884
beta2_b	Betrag von beta2	3000	0.7063976	0.0223644	0.6215169	0.7901247
beta3_b	Betrag von beta3	3000	0.7074742	0.0227932	0.6211303	0.7781780
beta4_b	Betrag von beta4	3000	0.7064910	0.0312778	0.5992052	0.8324256
beta5_b	Betrag von beta5	3000	0.7061169	0.0333668	0.5747302	0.8127062
beta6_b	Betrag von beta6	3000	0.7057666	0.0329030	0.5943090	0.8239747
gamma_b	Betrag von gamma	3000	0.7109760	0.0514000	0.5384944	0.9249620

Stderr v. PARMS aus MODELL 8 mit 2 korrel Faktoren und 2 pos. korrel. Fehlern (e1, e4) (3000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta4	beta4	3000	0.0349453	0.000793086	0.0317975	0.0380634
beta5	beta5	3000	0.0381907	0.0010441	0.0344375	0.0421539
beta6	beta6	3000	0.0381866	0.0010318	0.0348626	0.0421140
beta1	beta1	3000	0.0312076	0.000532618	0.0294250	0.0333939
beta_f1	beta_f1	3000	0.0218899	0.000646417	0.0199063	0.0252991
beta2	beta2	3000	0.0317941	0.000554553	0.0298329	0.0340384
beta_f2	beta_f2	3000	0.0226011	0.000601325	0.0208504	0.0249964
beta3	beta3	3000	0.0317863	0.000559286	0.0298199	0.0340316
beta_f3	beta_f3	3000	0.0226257	0.000593116	0.0208782	0.0249317
beta_f4	beta_f4	3000	0.0218812	0.000644578	0.0198180	0.0251871
beta_f5	beta_f5	3000	0.0226094	0.000605600	0.0206690	0.0258308
beta_f6	beta_f6	3000	0.0226012	0.000599429	0.0210137	0.0248923
gamma	gamma	3000	0.0511329	0.0033757	0.0426620	0.0676525

11.5.2.3.5 Modell 2_k_g

(außer gamma Betr. v.) PARMS aus MODELL 2 mit 2 unkorrel Faktoren Fehlerkorr. geschätzt
(3000 Iterat.)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		3000	0.0069911	0.0038316	0.000261735	0.0309436
ITER	Iteration	0	.	.	.	.
beta_f1	beta_f1	3000	0.7048196	0.0502040	-0.6924621	0.7857713
beta_f2	beta_f2	3000	0.7015817	0.0887574	-0.7398622	0.7749316
beta_f3	beta_f3	3000	0.7030579	0.0767434	-0.7363176	0.7855710
beta_f4	beta_f4	3000	0.7063132	0.0355373	-0.7218516	0.7751297
beta_f5	beta_f5	3000	0.6999765	0.0986301	-0.7249111	0.7775029
beta_f6	beta_f6	3000	0.7032894	0.0716455	-0.7376080	0.7848760
gamma	gamma	3000	0.000400221	0.0433397	-0.1443226	0.1733169
beta1_b	Betrag von beta1	3000	0.7074302	0.0244121	0.6208624	0.8031495
beta2_b	Betrag von beta2	3000	0.7066242	0.0242949	0.6320451	0.7827386
beta3_b	Betrag von beta3	3000	0.7065665	0.0242058	0.6187715	0.7904302
beta4_b	Betrag von beta4	3000	0.7061055	0.0240993	0.6286147	0.7925948
beta5_b	Betrag von beta5	3000	0.7062557	0.0240188	0.6277791	0.7942744
beta6_b	Betrag von beta6	3000	0.7061975	0.0245690	0.6185752	0.7917791
corr_14_b	Betrag von corr_14	3000	0.5016387	0.0378482	0.3802722	0.6486587

STDERR von korr. spez. MODELL 2 mit 2 unkorrel. Faktoren Fehlerkorr. geschätzt (3000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta4	beta4	3000	0.0316357	0.000780264	0.0292049	0.0351073
beta5	beta5	3000	0.0328692	0.000699347	0.0307692	0.0362114
beta6	beta6	3000	0.0328653	0.000689872	0.0308440	0.0363563
beta1	beta1	3000	0.0316560	0.000766199	0.0296321	0.0348044
beta_f1	beta_f1	3000	0.0248653	0.0011523	0.0221386	0.0306776
beta2	beta2	3000	0.0328825	0.000694137	0.0308112	0.0352768
beta_f2	beta_f2	3000	0.0241407	0.0010008	0.0217665	0.0299064
beta3	beta3	3000	0.0328825	0.000693519	0.0308788	0.0355739
beta_f3	beta_f3	3000	0.0241377	0.000993999	0.0217041	0.0280408
beta_f4	beta_f4	3000	0.0248349	0.0011520	0.0220753	0.0316275
beta_f5	beta_f5	3000	0.0241577	0.0010426	0.0214754	0.0306113
beta_f6	beta_f6	3000	0.0241476	0.000970304	0.0214550	0.0291032
gamma	gamma	3000	0.0424585	0.000606937	0.0407369	0.0455090
Corr_14	Corr_14	3000	0.0372793	0.0017240	0.0323316	0.0462765

11.5.2.3.6 Modell 3 k_g vor Ausschluss von Iteration 1560:

```
Iteration 1560:
                                  The CALIS Procedure
               Covariance Structure Analysis: Maximum Likelihood Estimation
HINWEIS: Covariance matrix for the estimates is not full rank.

HINWEIS: Die Varianz einiger Parameterschätzer ist Null oder einige Parameterschätzer sind
         linear abhängig zu anderen Parameterschätzern wie in folgenden Gleichungen
         dargestellt:
gamma      =       7.045234   +       0.014574   *   beta4      +       0.013666   *   beta5
                              +       0.014260   *   beta6      +       0.000165   *   beta1
                              -    0.000009789   *   beta_f1    +       0.000126   *   beta2
                              -    0.000005649   *   beta_f2    +    0.000059275   *   beta3
                              -    0.000001239   *   beta_f3    -   2.3332338E-8   *   beta_f4
                              +   3.0622834E-8   *   beta_f5    +   4.3409152E-8   *   beta_f6
                              -    0.000082084   *   beta7      -    0.000002376   *   beta_f7
                              -    0.000002998   *   Corr_12    +    0.000003355   *   Corr_14
```

Die Werte dieser Iteration werden daher nach erstmaliger Kennzahlenberechnung ausgeschlossen.

(außer gamma Betr. v.) PARMS aus MODELL 3 mit 2 unkorrel Faktoren und korrel. Fehlern (3000 Iterat.)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum
RHS		3000	0.0110574	0.0152584	0.0015921
ITER	Iteration	0	.	.	.
beta_f1	beta_f1	3000	0.7063279	0.0250791	0.6166131
beta_f2	beta_f2	3000	0.7052510	0.0451556	-0.7251985
beta_f3	beta_f3	3000	0.7066611	0.0351746	-0.6899470
beta_f4	beta_f4	3000	0.7059346	0.0238831	0.6158052
beta_f5	beta_f5	3000	0.6847115	0.1728769	-0.7540168
beta_f6	beta_f6	3000	0.6821141	0.1852708	-0.7326221
beta_f7	beta_f7	3000	0.7051861	0.0513090	-0.7348930
gamma	gamma	3000	0.0047577	0.1353220	-0.1435824
Corr_12	Corr_12	3000	0.4976908	0.0428181	-0.5577727
Corr_14	Corr_14	3000	0.5007101	0.0306223	0.3326061
beta1_b	Betrag von beta1	3000	0.7072670	0.0270540	0.0591347
beta2_b	Betrag von beta2	3000	0.7069569	0.0281925	0.0449287
beta3_b	Betrag von beta3	3000	0.7061703	0.0267766	0.0208992
beta4_b	Betrag von beta4	3000	0.7066645	0.0263106	0.1048064
beta5_b	Betrag von beta5	3000	0.7068243	0.0255426	0.0982785
beta6_b	Betrag von beta6	3000	0.7061758	0.0258795	0.1025479
beta7_b	Zusatzindikator Messmodell 1	3000	0.7066515	0.0269205	0.0289342
corr_14_b	Betr. d. Korr (Err 1, Err 4)	3000	0.5007101	0.0306223	0.3326061
corr_12_b	Betr. d. Korr (Err 1, Err 2)	3000	0.4983978	0.0335993	0.3578394

Variable	Etikett	Maximum
RHS		0.8070635
ITER	Iteration	.
beta_f1	beta_f1	1.0013342
beta_f2	beta_f2	0.9989902
beta_f3	beta_f3	0.9997816
beta_f4	beta_f4	0.7897320
beta_f5	beta_f5	0.7766720
beta_f6	beta_f6	0.7837796
beta_f7	beta_f7	0.9995813
gamma	gamma	7.0495656
Corr_12	Corr_12	0.7703458
Corr_14	Corr_14	0.6172324
beta1_b	Betrag von beta1	0.7978804
beta2_b	Betrag von beta2	0.7903359
beta3_b	Betrag von beta3	0.7886351
beta4_b	Betrag von beta4	0.7842722
beta5_b	Betrag von beta5	0.7811600
beta6_b	Betrag von beta6	0.7822156
beta7_b	Zusatzindikator Messmodell 1	0.7889910
corr_14_b	Betr. d. Korr (Err 1, Err 4)	0.6172324
corr_12_b	Betr. d. Korr (Err 1, Err 2)	0.7703458

STDERR aus MODELL 3 mit 2 unkorrel Faktoren und korrel. Fehlern (3000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum
RHS		0	.	.	.
ITER	Iteration	0	.	.	.
beta4	beta4	3000	0.0299943	0.000888691	0.0045718
beta5	beta5	3000	0.0323710	0.000816328	0.0044574
beta6	beta6	3000	0.0323698	0.000816319	0.0044584
beta1	beta1	3000	0.0316571	0.000782629	0.0295168
beta_f1	beta_f1	3000	0.0238212	0.0011749	0.0209621
beta2	beta2	3000	0.0335961	0.000779715	0.0314344
beta_f2	beta_f2	3000	0.0251419	0.0012191	0.0223152
beta3	beta3	3000	0.0324138	0.000653994	0.0305250
beta_f3	beta_f3	3000	0.0234761	0.000858320	0.0211437
beta_f4	beta_f4	3000	0.0245979	0.0011242	0.0217025
beta_f5	beta_f5	3000	0.0234972	0.000850324	0.0211465
beta_f6	beta_f6	3000	0.0234762	0.000870027	0.0212164
beta7	Zusatzindikator Messmodell 1	3000	0.0324154	0.000656296	0.0305792
beta_f7	beta_f7	3000	0.0234994	0.000890570	0.0211563
gamma	gamma	3000	0.0416759	0.000918041	0.000129049
Corr_12	Corr_12	3000	0.0330366	0.0037140	0.0123299
Corr_14	Corr_14	3000	0.0302950	0.0015584	0.0229031

Variable	Etikett	Maximum
RHS		.
ITER	Iteration	.
beta4	beta4	0.0332984
beta5	beta5	0.0349573
beta6	beta6	0.0352532
beta1	beta1	0.0374291
beta_f1	beta_f1	0.0295713
beta2	beta2	0.0366715
beta_f2	beta_f2	0.0308770
beta3	beta3	0.0359415
beta_f3	beta_f3	0.0276635
beta_f4	beta_f4	0.0293173
beta_f5	beta_f5	0.0279528
beta_f6	beta_f6	0.0283088
beta7	Zusatzindikator Messmodell 1	0.0359380
beta_f7	beta_f7	0.0285990
gamma	gamma	0.0436796
Corr_12	Corr_12	0.0499629
Corr_14	Corr_14	0.0370043

11.5.2.3.7 Modell 3 k_g nach Ausschluss von Iteration 1560:

(außer gamma Betr. v.) PARMS aus MODELL 3 mit 2 unkorrel Faktoren und korrel. Fehlern (3000)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum
RHS		2999	0.0107920	0.0046344	0.0015921
ITER	Iteration	0	.	.	.
beta_f1	beta_f1	2999	0.7062296	0.0244976	0.6166131
beta_f2	beta_f2	2999	0.7051531	0.0448432	-0.7251985
beta_f3	beta_f3	2999	0.7065634	0.0347706	-0.6899470
beta_f4	beta_f4	2999	0.7059481	0.0238757	0.6158052
beta_f5	beta_f5	2999	0.6847017	0.1729049	-0.7540168
beta_f6	beta_f6	2999	0.6821138	0.1853017	-0.7326221
beta_f7	beta_f7	2999	0.7050879	0.0510351	-0.7348930
gamma	gamma	2999	0.0024087	0.0419342	-0.1435824
Corr_12	Corr_12	2999	0.4975999	0.0425346	-0.5577727
Corr_14	Corr_14	2999	0.5007662	0.0304731	0.3960459
Iteration		2999	1500.48	866.3134934	1.0000000
beta1_b	Betrag von beta1	2999	0.7074831	0.0243310	0.6174458
beta2_b	Betrag von beta2	2999	0.7071777	0.0254724	0.5803646
beta3_b	Betrag von beta3	2999	0.7063988	0.0236756	0.6170372
beta4_b	Betrag von beta4	2999	0.7068652	0.0239085	0.6154035
beta5_b	Betrag von beta5	2999	0.7070272	0.0230017	0.6292809
beta6_b	Betrag von beta6	2999	0.7063771	0.0234178	0.6206918
beta7_b	Zusatzindikator Messmodell 1	2999	0.7068775	0.0239103	0.6278811
corr_14_b	Betr. d. Korr (Err 1, Err 4)	2999	0.5007662	0.0304731	0.3960459
corr_12_b	Betr. d. Korr (Err 1, Err 2)	2999	0.4983072	0.0332357	0.3578394

Variable	Etikett	Maximum
RHS		0.0413439
ITER	Iteration	.
beta_f1	beta_f1	0.7881458
beta_f2	beta_f2	0.8143567
beta_f3	beta_f3	0.7869339
beta_f4	beta_f4	0.7897320
beta_f5	beta_f5	0.7766720
beta_f6	beta_f6	0.7837796
beta_f7	beta_f7	0.7783092
gamma	gamma	0.1386615
Corr_12	Corr_12	0.6021740
Corr_14	Corr_14	0.6172324
Iteration		3000.00
beta1_b	Betrag von beta1	0.7978804
beta2_b	Betrag von beta2	0.7903359
beta3_b	Betrag von beta3	0.7886351
beta4_b	Betrag von beta4	0.7842722
beta5_b	Betrag von beta5	0.7811600
beta6_b	Betrag von beta6	0.7822156
beta7_b	Zusatzindikator Messmodell 1	0.7889910
corr_14_b	Betr. d. Korr (Err 1, Err 4)	0.6172324
corr_12_b	Betr. d. Korr (Err 1, Err 2)	0.6021740

STDERR aus MODELL 3 mit 2 unkorrel Faktoren und korrel. Fehlern (3000 Iterationen)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum
RHS		0	.	.	.
ITER	Iteration	0	.	.	.
beta4	beta4	2999	0.0300028	0.000757882	0.0275649
beta5	beta5	2999	0.0323803	0.000637676	0.0302425
beta6	beta6	2999	0.0323791	0.000637696	0.0301169
beta1	beta1	2999	0.0316552	0.000775626	0.0295168
beta_f1	beta_f1	2999	0.0238220	0.0011742	0.0209621
beta2	beta2	2999	0.0335952	0.000778185	0.0314344
beta_f2	beta_f2	2999	0.0251428	0.0012183	0.0223152
beta3	beta3	2999	0.0324126	0.000650921	0.0305250
beta_f3	beta_f3	2999	0.0234765	0.000858226	0.0211437
beta_f4	beta_f4	2999	0.0245976	0.0011243	0.0217025
beta_f5	beta_f5	2999	0.0234976	0.000850184	0.0211465
beta_f6	beta_f6	2999	0.0234762	0.000870171	0.0212164
beta7	Zusatzindikator Messmodell 1	2999	0.0324142	0.000653244	0.0305792
beta_f7	beta_f7	2999	0.0234998	0.000890479	0.0211563
gamma	gamma	2999	0.0416897	0.000516841	0.0401220
Corr_12	Corr_12	2999	0.0330435	0.0036953	0.0234399
Corr_14	Corr_14	2999	0.0302974	0.0015528	0.0258693
Iteration		2999	1500.48	866.3134934	1.0000000

Variable	Etikett	Maximum
RHS		.
ITER	Iteration	.
beta4	beta4	0.0332984
beta5	beta5	0.0349573
beta6	beta6	0.0352532
beta1	beta1	0.0346139
beta_f1	beta_f1	0.0295713
beta2	beta2	0.0366715
beta_f2	beta_f2	0.0308770
beta3	beta3	0.0347366
beta_f3	beta_f3	0.0276635
beta_f4	beta_f4	0.0293173
beta_f5	beta_f5	0.0279528
beta_f6	beta_f6	0.0283088
beta7	Zusatzindikator Messmodell 1	0.0348399
beta_f7	beta_f7	0.0285990
gamma	gamma	0.0436796
Corr_12	Corr_12	0.0499629
Corr_14	Corr_14	0.0370043
Iteration		3000.00

11.5.2.3.8 Modell 8 k_g

(Betrag von) PARMS aus MODELL 8 mit 2 korrel Faktoren und 2 pos. korrel. Fehlern (e1, e4 gesch.)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		3000	0.0071043	0.0037773	0.000236362	0.0266911
ITER	Iteration	0	.	.	.	.
beta_f1	beta_f1	3000	0.7058968	0.0225699	0.6182335	0.7789370
beta_f2	beta_f2	3000	0.7070896	0.0224230	0.6132718	0.7835610
beta_f3	beta_f3	3000	0.7059926	0.0227934	0.6286058	0.7834089
beta_f4	beta_f4	3000	0.7054408	0.0223385	0.6250741	0.7755973
beta_f5	beta_f5	3000	0.7059074	0.0226887	0.6253896	0.7802486
beta_f6	beta_f6	3000	0.7062074	0.0229706	0.6168369	0.7876811
beta1_b	Betrag von beta1	3000	0.7075585	0.0226096	0.6224594	0.7847733
beta2_b	Betrag von beta2	3000	0.7064143	0.0223747	0.6213149	0.7898719
beta3_b	Betrag von beta3	3000	0.7074855	0.0227910	0.6215067	0.7777241
beta4_b	Betrag von beta4	3000	0.7066802	0.0312565	0.6074454	0.8204467
beta5_b	Betrag von beta5	3000	0.7063388	0.0340065	0.5630196	0.8155260
beta6_b	Betrag von beta6	3000	0.7059789	0.0334048	0.5939737	0.8270303
gamma_b	Betrag von gamma	3000	0.7106395	0.0525717	0.5315151	0.9037947
corr_14_b		3000	0.5009071	0.0317013	0.3938263	0.6126303

Stderr v. PARMS aus MODELL 8 mit 2 korrel Faktoren und 2 pos. korrel. Fehlern (e1, e4 geschätzt)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta4	beta4	3000	0.0355675	0.000881551	0.0319478	0.0390128
beta5	beta5	3000	0.0386898	0.000997418	0.0349611	0.0423237
beta6	beta6	3000	0.0386865	0.000988104	0.0353069	0.0425011
beta1	beta1	3000	0.0312238	0.000514829	0.0294024	0.0333569
beta_f1	beta_f1	3000	0.0225651	0.000573938	0.0208208	0.0259331
beta2	beta2	3000	0.0317951	0.000555205	0.0298495	0.0340279
beta_f2	beta_f2	3000	0.0226032	0.000603800	0.0208391	0.0250556
beta3	beta3	3000	0.0317875	0.000560016	0.0298378	0.0340212
beta_f3	beta_f3	3000	0.0226277	0.000595291	0.0208749	0.0249499
beta_f4	beta_f4	3000	0.0225554	0.000571398	0.0207557	0.0255285
beta_f5	beta_f5	3000	0.0226122	0.000609788	0.0206729	0.0258543
beta_f6	beta_f6	3000	0.0226035	0.000601374	0.0210459	0.0250736
gamma	gamma	3000	0.0524202	0.0037701	0.0428349	0.0718086
corr_14	corr_14	3000	0.0322371	0.0015171	0.0270084	0.0390571

11.5.2.3.9 Modell 5 k_g

(Betr. v.)PARMS aus MODELL 5 mit 2 korrel Faktoren und korrel. Fehlern (Corr_12 geschätzt)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		3000	0.0071420	0.0038379	0.000214009	0.0291873
ITER	Iteration	0	.	.	.	.
beta_f1	beta_f1	3000	0.7045917	0.0285837	0.5983531	0.7941930
beta_f2	beta_f2	3000	0.7050727	0.0286490	0.5892013	0.7894884
beta_f3	beta_f3	3000	0.7061224	0.0268496	0.6165492	0.7829451
beta_f4	beta_f4	3000	0.7056924	0.0221119	0.6398868	0.7798800
beta_f5	beta_f5	3000	0.7059546	0.0226710	0.6271993	0.7772357
beta_f6	beta_f6	3000	0.7061330	0.0228304	0.6252287	0.7840063
beta1_b	Betrag von beta1	3000	0.7084703	0.0283516	0.6076656	0.8012325
beta2_b	Betrag von beta2	3000	0.7079875	0.0283888	0.6137655	0.8079863
beta3_b	Betrag von beta3	3000	0.7070747	0.0267595	0.6220909	0.7873164
beta4_b	Betrag von beta4	3000	0.7065390	0.0371280	0.5769526	0.8241065
beta5_b	Betrag von beta5	3000	0.7062648	0.0377450	0.5636890	0.8223922
beta6_b	Betrag von beta6	3000	0.7060368	0.0371799	0.5678700	0.8364680
gamma_b	Betrag von gamma	3000	0.7111018	0.0615399	0.5154722	0.9840513
corr_12_b	Betrag von corr_12	3000	0.4959707	0.0397329	0.3334931	0.6177021

Stderr aus MODELL 5 mit 2 korrel Faktoren und korrel. Fehlern (Corr_12 geschätzt) (3000 Iterat.)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta4	beta4	3000	0.0420085	0.0010884	0.0384746	0.0478974
beta5	beta5	3000	0.0420000	0.0010823	0.0387250	0.0466356
beta6	beta6	3000	0.0419978	0.0010787	0.0384957	0.0463522
beta1	beta1	3000	0.0359108	0.0010938	0.0323979	0.0410305
beta_f1	beta_f1	3000	0.0283054	0.0020240	0.0236773	0.0391912
beta2	beta2	3000	0.0359174	0.0010906	0.0320077	0.0408720
beta_f2	beta_f2	3000	0.0282894	0.0020322	0.0235743	0.0395458
beta3	beta3	3000	0.0347769	0.000944352	0.0314318	0.0390251
beta_f3	beta_f3	3000	0.0267230	0.0015670	0.0233083	0.0357317
beta_f4	beta_f4	3000	0.0225527	0.000587902	0.0206204	0.0252620
beta_f5	beta_f5	3000	0.0225492	0.000592239	0.0208996	0.0249761
beta_f6	beta_f6	3000	0.0225434	0.000585809	0.0208712	0.0250111
gamma	gamma	3000	0.0612692	0.0064626	0.0459287	0.1002876
Corr_12	Corr_12	3000	0.0394928	0.0060614	0.0265494	0.0755376

11.5.2.3.10 Modell Z4 k_g

(Betr. v.) PARMS aus korr. spez. MODELL Z4 m. 2 korrel Faktoren u. korrel. Fehlern (Corr_45 gesch.)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		3000	0.0070174	0.0036726	0.000520978	0.0265269
ITER	Iteration	0	.	.	.	.
beta_f1	beta_f1	3000	0.7059447	0.0225894	0.6255772	0.7762204
beta_f2	beta_f2	3000	0.7071240	0.0221677	0.6264172	0.7863483
beta_f3	beta_f3	3000	0.7059745	0.0225963	0.6349041	0.7786358
beta_f4	beta_f4	3000	0.7058109	0.0276466	0.5650439	0.7934591
beta_f5	beta_f5	3000	0.7057718	0.0279523	0.5784837	0.7878660
beta_f6	beta_f6	3000	0.7054144	0.0273833	0.6038606	0.7915606
beta1_b	Betrag von beta1	3000	0.7075471	0.0225667	0.6304616	0.7801623
beta2_b	Betrag von beta2	3000	0.7063959	0.0221235	0.6177834	0.7794880
beta3_b	Betrag von beta3	3000	0.7075173	0.0225605	0.6274762	0.7725909
beta4_b	Betrag von beta4	3000	0.7061975	0.0429393	0.5609859	0.9141349
beta5_b	Betrag von beta5	3000	0.7062482	0.0436407	0.5395483	0.8730309
beta6_b	Betrag von beta6	3000	0.7064537	0.0401504	0.5424901	0.8712605
gamma_b	Betrag von gamma	3000	0.7110675	0.0618994	0.5290702	0.9870177
corr_45_b	betrag von corr_45	3000	0.4979166	0.0402898	0.3061179	0.6122569

STDERR aus korr. spez. MODELL Z4 m. 2 korrel Faktoren u. korrel. Fehlern (Corr_45 geschätzt)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta4	beta4	3000	0.0475628	0.0023665	0.0413802	0.0573908
beta5	beta5	3000	0.0475598	0.0023692	0.0415530	0.0580818
beta6	beta6	3000	0.0444312	0.0020093	0.0384298	0.0529008
beta1	beta1	3000	0.0317450	0.000547749	0.0297124	0.0337596
beta_f1	beta_f1	3000	0.0225702	0.000592777	0.0206150	0.0261778
beta2	beta2	3000	0.0317528	0.000543947	0.0297610	0.0337700
beta_f2	beta_f2	3000	0.0225427	0.000594288	0.0210337	0.0255093
beta3	beta3	3000	0.0317437	0.000552732	0.0298210	0.0337854
beta_f3	beta_f3	3000	0.0225648	0.000566347	0.0207121	0.0250390
beta_f4	beta_f4	3000	0.0282470	0.0020380	0.0238797	0.0450770
beta_f5	beta_f5	3000	0.0282531	0.0020510	0.0235180	0.0403728
beta_f6	beta_f6	3000	0.0267836	0.0015971	0.0230654	0.0377522
gamma	gamma	3000	0.0613165	0.0065144	0.0469484	0.0981627
Corr_45	Corr_45	3000	0.0392590	0.0061610	0.0268821	0.0935944

11.5.2.3.11 Modell 6 k_g

(Betr. v.) PARMS aus korr. spez. MODELL 6 m. 2 korrel. Faktoren Corr_12 u. Corr_45 geschätzt)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		3000	0.0060882	0.0035654	0.000165679	0.0284789
ITER	Iteration	0	.	.	.	.
beta_f1	beta_f1	3000	0.7045284	0.0293561	0.5848039	0.7907555
beta_f2	beta_f2	3000	0.7050509	0.0293791	0.5883814	0.7882058
beta_f3	beta_f3	3000	0.7060078	0.0275085	0.6127648	0.7892159
beta_f4	beta_f4	3000	0.7057135	0.0283860	0.5850061	0.7907791
beta_f5	beta_f5	3000	0.7055952	0.0288060	0.5888687	0.7854433
beta_f6	beta_f6	3000	0.7053672	0.0281902	0.6072299	0.7922593
beta1_b	Betrag von beta1	3000	0.7084740	0.0290387	0.6121321	0.8111747
beta2_b	Betrag von beta2	3000	0.7079499	0.0291135	0.6154117	0.8085836
beta3_b	Betrag von beta3	3000	0.7071383	0.0274258	0.6141158	0.7902654
beta4_b	Betrag von beta4	3000	0.7064554	0.0453306	0.5524473	0.8787099
beta5_b	Betrag von beta5	3000	0.7065611	0.0457576	0.5363052	0.8687281
beta6_b	Betrag von beta6	3000	0.7066125	0.0419716	0.5486205	0.8612753
gamma_b	Betrag von gamma	3000	0.7108859	0.0662357	0.5190978	1.0270013
corr_12_b	Betrag von corr_12	3000	0.4958076	0.0410356	0.3110863	0.6120618
corr_45_b	Betrag von corr_45	3000	0.4975961	0.0415389	0.2986411	0.6087016

STDERR aus korr. spez. MODELL 6 m. 2 unkorrel. Faktoren Corr_12 u. Corr_45 geschätzt (3000 Iter.)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta4	beta4	3000	0.0497604	0.0026123	0.0429584	0.0594068
beta5	beta5	3000	0.0497605	0.0026157	0.0426887	0.0594202
beta6	beta6	3000	0.0464909	0.0022870	0.0397460	0.0548337
beta1	beta1	3000	0.0365284	0.0011918	0.0327285	0.0422893
beta_f1	beta_f1	3000	0.0291121	0.0022542	0.0241522	0.0404937
beta2	beta2	3000	0.0365349	0.0011883	0.0322709	0.0425340
beta_f2	beta_f2	3000	0.0290913	0.0022421	0.0240400	0.0428766
beta3	beta3	3000	0.0353675	0.0010368	0.0317073	0.0402110
beta_f3	beta_f3	3000	0.0275056	0.0017552	0.0236906	0.0374718
beta_f4	beta_f4	3000	0.0290430	0.0022724	0.0237521	0.0421180
beta_f5	beta_f5	3000	0.0290522	0.0022831	0.0235857	0.0409586
beta_f6	beta_f6	3000	0.0275508	0.0017707	0.0235460	0.0386680
gamma	gamma	3000	0.0663186	0.0081000	0.0482223	0.1172032
Corr_12	Corr_12	3000	0.0406587	0.0066914	0.0271008	0.0791664
Corr_45	Corr_45	3000	0.0404283	0.0067692	0.0261682	0.0780058

11.5.2.3.12 Modell 9 k_g nicht ausreißerbereinigt

nicht ausreißerbereinigte Statistik (n=2956)

(Betr. v.) PARMS aus Modell 9 mit 2 unkorrel Faktoren u. korrel. Fehlern (corr_12 u. _14 gesch.)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		2956	0.0621480	1.2032992	0.000306890	30.5395964
ITER	Iteration	0	.	.	.	.
beta_f1	beta_f1	2956	0.7015602	0.0882155	-0.7099163	1.1814545
beta_f2	beta_f2	2956	0.6132253	0.3530656	-0.7713664	1.0072478
beta_f3	beta_f3	2956	0.7063390	0.0490446	-0.7150144	1.0000000
beta_f4	beta_f4	2956	0.7053679	0.0362017	-0.6907698	1.1662022
beta_f5	beta_f5	2956	0.7070700	0.0424214	0.6252397	2.5494871
beta_f6	beta_f6	2956	0.7127141	0.3590201	-0.7050792	20.1253247
beta1_b	Betrag von beta1	2956	0.7532294	1.1259192	0.0456079	30.8325779
beta2_b	Betrag von beta2	2956	0.7567123	1.2414861	0.0627777	40.4588439
beta3_b	Betrag von beta3	2956	0.7297917	0.7515282	0.0351284	28.5879375
beta4_b	Betrag von beta4	2956	0.9970478	9.4256176	0.000117423	468.5182068
beta5_b	Betrag von beta5	2956	0.9929508	8.3880312	0.000117204	396.0642512
beta6_b	Betrag von beta6	2956	0.9747671	8.3040618	0.000120856	398.6411500
gamma_b	Betrag von gamma	2956	24.8454595	500.7042833	0.0158721	17493.71
corr_12_b	Betrag von corr_12	2956	0.6589741	7.8489973	0.0468110	425.6803639
corr_14_b	Betrag von corr_14	2956	0.5801956	3.8814330	0.2557286	210.3035493

STDERR aus Mod. 9 m. 2 unkorrel Faktoren u. korrel. Fehlern (corr_12 u. _14 geschätzt)(3000 Iter.)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta4	beta4	2953	0.0679720	1.6303066	5.4245425E-6	88.6003144
beta5	beta5	2953	0.0659664	1.3782474	5.4247759E-6	74.8982905
beta6	beta6	2953	0.0660997	1.3871797	5.3864808E-6	75.3855985
beta1	beta1	2955	0.0350019	0.0051150	0.0219573	0.2317288
beta_f1	beta_f1	2956	0.0269647	0.0592440	0.000572420	3.1864328
beta2	beta2	2955	0.0364439	0.0198294	0.0235508	1.0956408
beta_f2	beta_f2	2956	0.0293129	0.0299197	0.0155412	1.3237082
beta3	beta3	2956	0.0351745	0.0011795	0.0211491	0.0655160
beta_f3	beta_f3	2956	0.0272238	0.0016574	0.0161429	0.0347682
beta_f4	beta_f4	2956	0.0226049	0.0011978	0.0207028	0.0677280
beta_f5	beta_f5	2956	0.0223349	0.0034348	0.0205284	0.1581481
beta_f6	beta_f6	2956	0.0224136	0.0082650	0.0205131	0.4555839
gamma	gamma	2954	3.3762937	180.3799554	1.21451E-12	9803.84
Corr_12	Corr_12	2956	2.9351768	150.4129473	0.0129418	8170.29
Corr_14	Corr_14	2956	1.8521765	97.7341458	0.0217418	5313.20

11.5.2.3.13 Modell 9 k_g ausreißerbereinigt

ausreißerbereinigte Statistik (n=2930)

(Betr. v.) PARMS aus Mod. 9 m. 2 unkorrel Faktoren u. korrel. Fehlern (corr_12 u. _14 gesch.)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		2930	0.0061209	0.0038579	0.000306890	0.0763619
ITER	Iteration	0	.	.	.	.
beta_f1	beta_f1	2930	0.7018497	0.0820971	-0.7099163	0.7957261
beta_f2	beta_f2	2930	0.6126468	0.3529349	-0.7713664	0.7937045
beta_f3	beta_f3	2930	0.7048896	0.0461601	-0.7150144	0.8308379
beta_f4	beta_f4	2930	0.7050077	0.0340050	-0.6907698	0.7798105
beta_f5	beta_f5	2930	0.7060504	0.0221934	0.6252397	0.7810737
beta_f6	beta_f6	2930	0.7062687	0.0224629	0.6183278	0.7926403
beta1_b	Betrag von beta1	2930	0.7071492	0.0270429	0.6119629	0.9153540
beta2_b	Betrag von beta2	2930	0.7066357	0.0285844	0.6083035	0.9330365
beta3_b	Betrag von beta3	2930	0.7072809	0.0274970	0.5567207	0.7979655
beta4_b	Betrag von beta4	2930	0.7078842	0.0339786	0.5921388	0.8459866
beta5_b	Betrag von beta5	2930	0.7073424	0.0352205	0.5832218	0.8273630
beta6_b	Betrag von beta6	2930	0.7070649	0.0345729	0.5875073	0.8164346
gamma_b	Betrag von gamma	2930	0.7085446	0.0565344	0.5075562	0.9433181
corr_12_b	Betrag von corr_12	2930	0.4970873	0.0350247	0.3127403	0.5921485
corr_14_b	Betrag von corr_14	2930	0.5015487	0.0279811	0.4182304	0.8054102
Iteration		2930	1466.50	845.9624696	2.0000000	2931.00

STDERR aus Mod. 9 m. 2 unkorrel Faktoren u. korrel. Fehlern (corr_12 u. _14 geschätzt) (3000 Iter.)

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		0	.	.	.	.
ITER	Iteration	0	.	.	.	.
beta4	beta4	2942	0.0371749	0.0026550	5.4245425E-6	0.0412902
beta5	beta5	2942	0.0397932	0.0027686	5.4247759E-6	0.0446119
beta6	beta6	2942	0.0397795	0.0028511	5.3864808E-6	0.0439790
beta1	beta1	2943	0.0348578	0.0011686	0.0235785	0.0684422
beta_f1	beta_f1	2944	0.0255460	0.0018743	0.0112212	0.0350068
beta2	beta2	2943	0.0360009	0.0012876	0.0235508	0.0679781
beta_f2	beta_f2	2944	0.0282901	0.0020930	0.0155412	0.0392940
beta3	beta3	2944	0.0351775	0.0011384	0.0228359	0.0655160
beta_f3	beta_f3	2944	0.0272401	0.0016381	0.0161429	0.0347682
beta_f4	beta_f4	2944	0.0225908	0.0010839	0.0207936	0.0677280
beta_f5	beta_f5	2944	0.0223295	0.0034337	0.0205284	0.1581481
beta_f6	beta_f6	2944	0.0223725	0.0080072	0.0205131	0.4555839
gamma	gamma	2942	0.0567486	0.0063512	1.21451E-12	0.0820896
Corr_12	Corr_12	2944	0.0347145	0.0052651	0.0129418	0.0595552
Corr_14	Corr_14	2944	0.0275535	0.0015800	0.0217418	0.0357859
Iteration		2944	1480.50	850.0039216	9.0000000	2952.00

11.6 Adjustierung der Standardfehlerschätzung nach der intraindividuellen Abhängigkeit zwischen paarigen Organen – Simulation und Anwendung in der Glaukomdiagnose

11.6.1 Erster Abschnitt, Teil A: (Vergleich der) Bootstrapverfahren auf Basis simulierter Daten anhand eines Ein-Faktor-Modells mit drei Indikatoren

11.6.1.1 A Vollständiger kommentierter Programmcode des Modells zur Datensimulation der Korrelationsstruktur K6 (siehe

11.6.1.2 Tabelle 11; R Glaukomschweregrad = R Fehlerkorrelation = 0.5)

```
**********:

/* Ad Teil A "Vergleich der Bootstrapverfahren auf Basis simulierter Daten anhand eines
Ein-Faktor-Modells mit drei Indikatoren"
A
Vollständiger kommentierter Programmcode des Modells zur Datensimulation
der Korrelationsstruktur K6 (siehe
Tabelle 11; R Glaukomschweregrad = R Fehlerkorrelation = 0.5)
*/
*1. DATA Step zur Ziehung von personenbezogenen, also nur augenpaarweisen Stichproben ;
*Erzeugung von Zufallszahlen von 1 bis 400 zur Generierung der Bootstrapamples;
*dazu zunächst Erzeugung von Zufallszahlen im Bereich zwischen 0.5 und 400.5 ;
%MACRO bootstrap (i, seed, dateinam=, lib=);
DATA &lib..&dateinam;
         DO i=1 to &i;
         random_variate=400*ranuni(&seed)+0.5;
         OUTPUT;
         END;
RUN;
%MEND bootstrap;
%macro create_1(howmany);
   %do i=1 %to &howmany;
      %bootstrap (400, &i, dateinam=ranuni&i, lib=boot_str);
   %end;
%mend create_1;
%create_1(3000)
*Rundung der o. g. Zufallszahlen auf null Dezimalstellen;
%MACRO bootstrap2 (dateinam=, lib=);
DATA &lib..r_&dateinam;
SET &lib..&dateinam;
         rand_var_r=round(random_variate);
RUN;
%MEND bootstrap2;
%macro create_2(howmany);
   %do i=1 %to &howmany;
      %bootstrap2 (dateinam=ranuni&i, lib=boot_str);
   %end;
%mend create_2;
%create_2(3000)

*Sortieren der gerundeten Zufallszahlen;
%MACRO bootstrap3 (dateinam=, lib=);
PROC SORT DATA=&lib..r_&dateinam OUT=&lib..rs_&dateinam;
BY rand_var_r;
%MEND bootstrap3;
%macro create_3(howmany);
   %do i=1 %to &howmany;
      %bootstrap3 (dateinam=ranuni&i, lib=boot_str);
   %end;
%mend create_3;
%create_3(3000)
*1. DATA Step zur augenbezogenen Ziehung von Stichproben, d. h. zur Ziehung von einzelnen Augen;
*Zufallszahlen von 1 bis 800 erzeugen;
%MACRO bootstrap (i, seed, dateinam=, lib=);
DATA &lib..&dateinam;
         DO i=1 to &i;
         random_variate=800*ranuni(&seed)+0.5;
         OUTPUT;
```

```
         END;
RUN;
%MEND bootstrap;
%macro create_1(howmany);
   %do i=1 %to &howmany;
      %bootstrap (800, &i, dateinam=ranuni_aug&i, lib=boot_str);
   %end;
%mend create_1;
%create_1(3000)
*Runden;
%MACRO bootstrap2 (dateinam=, lib=);
DATA &lib..r_&dateinam;
SET &lib..&dateinam;
         rand_var_r=round(random_variate);
RUN;
%MEND bootstrap2;
%macro create_2(howmany);
   %do i=1 %to &howmany;
      %bootstrap2 (dateinam=ranuni_aug&i, lib=boot_str);
   %end;
%mend create_2;
%create_2(3000)
*Sortieren;
%MACRO bootstrap3 (dateinam=, lib=);
PROC SORT DATA=&lib..r_&dateinam OUT=&lib..rs_&dateinam;
BY rand_var_r;
%MEND bootstrap3;
%macro create_3(howmany);
   %do i=1 %to &howmany;
      %bootstrap3 (dateinam=ranuni_aug&i, lib=boot_str);
   %end;
%mend create_3;
%create_3(3000)
*Erzeugen der ersten von insgesamt zwei Populationen, aus der  Stichproben mit Zurücklegen
*(Bootstrapsamples) gezogen werden;
%MACRO machdat5 (i, seedx1, seedx2, seedx3, seedx4, seedx5, seedx6,
seedx7, seedx8, seedx9, seedx10, seedx11, seedx12, dateinam=, lib=);
DATA &lib..&dateinam;
Do i = 1 to &i;
x1 = Rannor (&seedx1);
x2 = Rannor (&seedx2);
x3 = Rannor (&seedx3);
x4 = Rannor (&seedx4);
x5 = Rannor (&seedx5);
x6 = Rannor (&seedx6);
x7 = Rannor (&seedx7);
x8 = Rannor (&seedx8);
x9 = Rannor (&seedx9);
x10 = Rannor (&seedx10);
x11 = Rannor (&seedx11);
x12 = Rannor (&seedx12);
g_li_5=(x4 + x5)/SQRT(2);
g_re_5= (x4 + x6)/SQRT(2);
f1_li_5=(x1 + x7)/SQRT(2);
f1_re_5= (x1 + x8)/SQRT(2);
f2_li_5=(x2 + x9)/SQRT(2);
f2_re_5= (x2 + x10)/SQRT(2);
f3_li_5=(x3 + x11)/SQRT(2);
f3_re_5= (x3 + x12)/SQRT(2);
OUTPUT;
END;
RUN;
%MEND machdat5;
%macro create9(howmany);
   %do i=1 %to &howmany;
      %machdat5 (400, &i, &i, &i, &i, &i, &i, &i, &i, &i, &i, &i, &i, &i, dateinam=iterat&i, lib=boot_str);;
       run;
   %end;
%mend create9;
%create9(1)
/*Kommentar zum Macro "create9" : Die Zeile "%do i=1 %to &howmany;" wird bei Erzeugung der zweiten
von insgesamt zwei Populationen in "%do i=2 %to &howmany;" abgewandelt. Entsprechend wird die das
Macro 9 auslösende Zeile "%create9(1);" zu "%create9(2);" verändert. Es handelt sich hierbei um eine
Variation des Startwerts bei der Zufallszahlenerzeugung die auch für das unten stehende Macro
create 7 (für das Ziehen einzelner Augen) zutrifft*/

*2 A.;
*diese Zufallsvariablen werden weiterverarbeitet, indem 3 manifeste Variablen für die linke Seite
*aus Ihnen konstruiert werden – entsprechend einem bestimmten in AMOS gebildeten Grundmodell;
*Erzeugen einer Variable in dieser Population, die bei späterem Mergen mit den Zufallszahlendateien als BY-Variable
*dient;
*Erzeugen einer Variable für die linke Augenseite;
%Macro arit_op2 (dateinam=, lib=);
DATA &lib..bl_&dateinam;
```

```
SET &lib..&dateinam;
indik_1 = g_li_5 + f1_li_5;
indik_2 = g_li_5 + f2_li_5;
indik_3 = g_li_5 + f3_li_5;
rand_var_r=i;
seite=1;
DROP x1 x2 x3 x4 x5 x6 x7 x8 x9 x10 x11 x12 g_li_5 g_re_5 f1_li_5 f1_re_5
f2_li_5 f2_re_5 f3_li_5 f3_re_5;
RUN;
%MEND arit_op2;
%macro create7(howmany);
   %do i=1 %to &howmany;
      %arit_op2 (dateinam=iterat&i, lib=boot_str);
       run;
   %end;
%mend create7;
%create7(1)

*2 B.;
*dieselben Zufallsvariablen werden weiterverarbeitet, indem 3 manifeste Variablen für die rechte Seite
*aus Ihnen konstruiert werden – entsprechend einem bestimmten in AMOS gebildeten Grundmodell;
*Erzeugen einer Variable in dieser Population, die bei späterem Mergen mit den Zufallszahlendateien als BY-Variable
*dient;
*Erzeugen einer Variable für die rechte Augenseite;
%Macro arit_op2 (dateinam=, lib=);
DATA &lib..br_&dateinam;
SET &lib..&dateinam;
indik_1 = g_re_5 + f1_re_5;
indik_2 = g_re_5 + f2_re_5;
indik_3 = g_re_5 + f3_re_5;
rand_var_r=i;
seite=2;
DROP x1 x2 x3 x4 x5 x6 x7 x8 x9 x10 x11 x12 g_li_5 g_re_5 f1_li_5 f1_re_5
f2_li_5 f2_re_5 f3_li_5 f3_re_5;
RUN;
%MEND arit_op2;
%macro create7(howmany);
   %do i=1 %to &howmany;
      %arit_op2 (dateinam=iterat&i, lib=boot_str);
       run;
   %end;
%mend create7;
%create7(1)

*II
***
*800 A u g e n von 400 Patienten ziehen aus dem gleichen Datensatz, aus dem 400 P a t i e n t e n
*mit ihren jeweiligen Augenpaaren gezogen wurden (jeweils mit Zurücklegen);
/*Addieren der Datendatei linker Augen, aus der Bootstrapsamples gezogen werden
und der Datendatei rechter Augen, aus der Bootstrapsamples gezogen werden*/
DATA BOOT_STR.BLR_ITERAT1;
SET BOOT_STR.BL_ITERAT1 BOOT_STR.BR_ITERAT1;
rand_var_r=_N_;
RUN;
PROC SORT DATA=BOOT_STR.BLR_ITERAT1 OUT=BOOT_STR.BLRS_ITERAT1;
BY rand_var_r;
RUN;
PROC CORR DATA=BOOT_STR.BLRS_ITERAT1;
VAR indik_1 indik_2 indik_3;
RUN;

***Anfang*****************************************************************************
*Die Korrelation zwischen dem linken Wert und dem rechten Wert eines jeweiligen Indikators
*Ausgangsdatensatz umstellen
*a) seitenneutrale Variablen in seitenanzeigende Variablen umbenennen;
DATA BOOT_STR.BL_UK_ITERAT1;
SET BOOT_STR.BL_ITERAT1;
indik1_li = indik_1;
indik2_li = indik_2;
indik3_li = indik_3;
DROP indik_1 indik_2 indik_3 seite;
RUN;

DATA BOOT_STR.BR_UK_ITERAT1;
SET BOOT_STR.BR_ITERAT1;
indik1_re = indik_1;
indik2_re = indik_2;
indik3_re = indik_3;
DROP indik_1 indik_2 indik_3 seite;
RUN;

*b) match-mergen der umbenannten Datendateien;
DATA BOOT_STR.BLR_UK_ITERAT1;
MERGE BOOT_STR.BL_UK_ITERAT1 BOOT_STR.BR_UK_ITERAT1;
```

```
BY rand_var_r;
RUN;
PROC CORR DATA=BOOT_STR.BLR_UK_ITERAT1;
VAR INDIK1_li INDIK1_re;
TITLE1 "Analysen m. Ziehen v. Augenpaaren und einz. Augen";
TITLE2 "3000 samples n=400 bzw. 800 li-re-Fehler- und Glaukomschweregrad-Korr = 0.5";
RUN;
PROC CORR DATA=BOOT_STR.BLR_UK_ITERAT1;
VAR INDIK2_li INDIK2_re;
RUN;
PROC CORR DATA=BOOT_STR.BLR_UK_ITERAT1;
VAR INDIK3_li INDIK3_re;
RUN;
PROC CORR DATA=BOOT_STR.BLR_UK_ITERAT1;
VAR INDIK1_li INDIK2_li INDIK3_li;
RUN;
PROC CORR DATA=BOOT_STR.BLR_UK_ITERAT1;
VAR INDIK1_re INDIK2_re INDIK3_re;
RUN;
***Ende*************************************************************************
************************************************************************
*hier geht das Mergen für beide Ziehungsarten und alle weiteren Prozesse los
************************************************************************
*zunächst das Personenbezogene paarweise Ziehen von Augen, das in 3 A und 3 B und
*"Addieren der li. (3 A) und re (3 B) Augen" durchgeführt wird
*3 A. DATA STEP für die linke Seite;
*********************************;
*Mergen;
*entspricht nach SAS-Language Concept
*dem Beispiel 'Match-Merged with Duplicate BY Values';
%MACRO bootstrap4 (dateinam=, lib=);
DATA &lib..merged_l_&dateinam;
***************;
*Hier Hochzählen;
***************;
MERGE &lib..rs_&dateinam &lib..bl_iterat1;
BY rand_var_r;
%MEND bootstrap4;
%macro create_4(howmany);
   %do i=1 %to &howmany;
      %bootstrap4 (dateinam=ranuni&i, lib=boot_str);
   %end;
%mend create_4;
%create_4(3000)

*Löschen der 'nicht gezogenen' Records in den "gemergeten" Datendateien;
%MACRO bootstrap5 (dateinam=, lib=);
DATA &lib..sample_l_&dateinam;
SET &lib..merged_l_&dateinam;
IF random_variate EQ "." THEN DELETE;
DROP random_variate i;
RUN;
%MEND bootstrap5;
%macro create_5(howmany);
   %do i=1 %to &howmany;
      %bootstrap5 (dateinam=ranuni&i, lib=boot_str);
   %end;
%mend create_5;
%create_5(3000)

*3 B. DATA STEP für die rechte Seite;
*********************************;
%MACRO bootstrap4 (dateinam=, lib=);
DATA &lib..merged_r_&dateinam;
***************;
*Hier Hochzählen;
***************;
MERGE &lib..rs_&dateinam &lib..br_iterat1;
BY rand_var_r;
%MEND bootstrap4;
%macro create_4(howmany);
   %do i=1 %to &howmany;
      %bootstrap4 (dateinam=ranuni&i, lib=boot_str);
   %end;
%mend create_4;
%create_4(3000)

*Löschen der 'nicht gezogenen' Records in den "gemergeten" Datendateien;
%MACRO bootstrap5 (dateinam=, lib=);
DATA &lib..sample_r_&dateinam;
SET &lib..merged_r_&dateinam;
IF random_variate EQ "." THEN DELETE;
DROP random_variate i;
RUN;
```

```
%MEND bootstrap5;
%macro create_5(howmany);
   %do i=1 %to &howmany;
      %bootstrap5 (dateinam=ranuni&i, lib=boot_str);
   %end;
%mend create_5;
%create_5(3000)
*Addieren der linken und rechten Augen;
****************************************;
%MACRO bootstrap6 (dateinam=, lib=);
DATA &lib..sample_&dateinam;
SET &lib..sample_l_&dateinam &lib..sample_r_&dateinam;
LABEL seite='Augenseite';
RUN;
%MEND bootstrap6;
%macro create_6(howmany);
   %do i=1 %to &howmany;
      %bootstrap6 (dateinam=ranuni&i, lib=boot_str);
   %end;
%mend create_6;
%create_6(3000)

*Sortieren nach Patient;
%MACRO bootstrap7 (dateinam=, lib=);
PROC SORT DATA=&lib..sample_&dateinam;
         BY rand_var_r seite;
%MEND bootstrap7;
%macro create_7(howmany);
   %do i=1 %to &howmany;
      %bootstrap7 (dateinam=ranuni&i, lib=boot_str);
   %end;
%mend create_7;
%create_7(3000)
**********************************************************************;
%Macro proccalis (dateinam=, lib=);
proc calis DATA=&lib..sample_&dateinam OUTEST=&lib..oute_sample_&dateinam tech=nr nop edf=799;
         Lineqs
                  indik_1 =          beta1     Faktor_1           + beta_f1           Error_1,
                  indik_2 =          beta2     Faktor_1           + beta_f2           Error_2,
                  indik_3 =          beta3     Faktor_1           + beta_f3           Error_3;

         Std
                  Error_1 = 1,
                  Error_2 = 1,
                  Error_3  = 1,
                  Faktor_1= 1;
                  run;
         %MEND proccalis;
%macro create_7(howmany);
   %do i=1 %to &howmany;
      %proccalis (dateinam=ranuni&i, lib=boot_str);
       run;
   %end;
%mend create_7;
%create_7(3000)
*5.;
*Von den (1000) OUTEST-sas-data-sets wird nur jeweils die interessierende Zeile (Beobachtung)
*mit dem _TYPE_ "PARMS" (=final Parameter Estimates) bewahrt (und der Rest gelöscht);
%Macro parms_isol (dateinam=, lib=);
DATA &lib..PARMS_sample_&dateinam;
SET &lib..oute_sample_&dateinam;
IF _TYPE_ NE "PARMS" THEN DELETE;
RUN;
%MEND parms_isol;
%macro create_8(howmany);
   %do i=1 %to &howmany;
      %parms_isol (dateinam=ranuni&i, lib=boot_str);
       run;
   %end;
%mend create_8;
%create_8(3000)

*6.;
*die (1000) einzeiligen Datensätze addieren (unter Verwendung eines „iterativen Do-Makros“);
%macro create_9(howmany);
   %do i=2 %to &howmany;
   proc APPEND BASE= BOOT_STR.PARMS_SAMPLE_RANUNI1
            DATA= BOOT_STR.PARMS_SAMPLE_RANUNI&i;
      run;
   %end;
%mend create_9;
%create_9(3000)
*7.;
*Umbenennen, Ausdrucken der Datei;
```

```
DATA BOOT_STR.PARMS_ALL_RANUNI;
SET BOOT_STR.PARMS_SAMPLE_RANUNI1;
RUN;
*Es stellt sich heraus, dass – wie in den AMOS-Modellen auch schon – mitunter negative Werte für
*die Koeffizienten der Pfade geschätzt werden, die vom Faktor zu den manifesten Variablen führen
*(beta1, beta2 und beta3).
*dieser Tatsache soll keine Beachtung geschenkt werden und hat etwas mit dem Schätzalgorithmus
*zu tun. Daher folgende Datenmanipulation:
8.;
*die negativen Werte von beta* werden nun in positive umgewandelt in dem man neue Variablen
*beta*_b erzeugt, die den Betrag von beta* angeben;
*die alten Variablen werden gelöscht;
*Schließlich werden statistische Maßzahlen und Histogramme für die sechs interessierenden
*Pfadkoeffizienten erstellt;
DATA BOOT_STR.PARMS_ALL_RANUNI_B;;
SET BOOT_STR.PARMS_ALL_RANUNI;
beta1_b=ABS(beta1);
beta2_b=ABS(beta2);
beta3_b=ABS(beta3);
DROP beta1 beta2 beta3;
LABEL beta1_b='Betrag von beta1'
           beta2_b='Betrag von beta2'
           beta3_b='Betrag von beta3';
TITLE "(Betrag von) PARMS aus LINEQS MODEL m. unkorr. Fehlern (Bootstrap Ziehen v. Augenpaaren)";
RUN;
PROC MEANS DATA=BOOT_STR.PARMS_ALL_RANUNI_B; OUTPUT OUT=BOOT_STR.OUT_PARMS_ALL_RANUNI_B;
TITLE1 "Betr. v. PARMS aus LINEQS MODEL m. unkorr. Fehlern (Bootstrap Ziehen v. Augenpaaren)";
TITLE2 "n=400 li-re-F- und G-Korr = 0.5";
RUN;
DATA BOOT_STR.OUT_B_PARMS_ALL_RANUNI_B;
SET BOOT_STR.OUT_PARMS_ALL_RANUNI_B;
IF _STAT_ NE "STD" THEN DELETE;
DROP _TYPE_ _RHS_ _ITER_ beta_f1 beta_f2 beta_f3;
RUN;
/*PROC GCHART DATA=BOOT_STR.PARMS_ALL_RANUNI_B;
VBAR beta1_b beta2_b beta3_b;
TITLE1 "Betr. v. PARMS aus LINEQS MODEL m. unkorr. Fehlern (Bootstrap Ziehen v. Augenpaaren)";
TITLE2 "n=400 li-re-Indikator-Korr = 0.9";
RUN;
*/
PROC CORR DATA=BOOT_STR.PARMS_ALL_RANUNI_B NOSIMPLE;
VAR beta1_b beta2_b beta3_b;
TITLE1 "Betr. v. PARMS aus LINEQS MODEL m. unkorr. Fehlern (Bootstrap Ziehen v. Augenpaaren)";
TITLE2 "n=400 li-re-F- und G-Korr = 0.5";
RUN;
*************************************************************************************************
*3. DATA STEP ;
*zur Erzeugung einer Datei, aus der  e i n z e l n e A u g e n  gezogen werden – unabhängig von ihrer
*Zugehörigkeit zu einem Patienten
*********************************;
*Bei dem Datensatz boot_str.blrs_iterat1 (siehe unten) handelt es sich um eine Datei,
*in der die 400 li. Augen und 400 re. Augen addiert wurden.
*Aus dieser sollen nun a u g e n b e z o g e n
*Stichproben im Umfang n gezogen werden.
*Mergen;
%MACRO bootstrap4 (dateinam=, lib=);
DATA &lib..merged_&dateinam;
***************;
*Hier Hochzählen;
***************;
MERGE &lib..rs_&dateinam &lib..blrs_iterat1;
BY rand_var_r;
%MEND bootstrap4;
%macro create_4(howmany);
   %do i=1 %to &howmany;
      %bootstrap4 (dateinam=ranuni_aug&i, lib=boot_str);
   %end;
%mend create_4;
%create_4(3000)
*Löschen der 'nicht gezogenen' Records in den "gemergeten" Datendateien;
%MACRO bootstrap5 (dateinam=, lib=);
DATA &lib..sample_&dateinam;
SET &lib..merged_&dateinam;
IF random_variate EQ "." THEN DELETE;
DROP random_variate i;
RUN;
%MEND bootstrap5;
%macro create_5(howmany);
   %do i=1 %to &howmany;
      %bootstrap5 (dateinam=ranuni_aug&i, lib=boot_str);
   %end;
%mend create_5;
%create_5(3000)
```

```
%Macro proccalis (dateinam=, lib=);
proc calis DATA=&lib..sample_&dateinam OUTEST=&lib..oute_sample_&dateinam tech=nr nop edf=799;
         Lineqs
                  indik_1 =          beta1     Faktor_1          + beta_f1          Error_1,
                  indik_2 =          beta2     Faktor_1          + beta_f2          Error_2,
                  indik_3 =          beta3     Faktor_1          + beta_f3          Error_3;

         Std
                  Error_1 = 1,
                  Error_2 = 1,
                  Error_3 = 1,
                  Faktor_1= 1;
         run;
%MEND proccalis;
%macro create_7(howmany);
   %do i=1 %to &howmany;
      %proccalis (dateinam=ranuni_aug&i, lib=boot_str);
       run;
   %end;
%mend create_7;
%create_7(3000)
*5.;
*Von den (1000) OUTEST-sas-data-sets wird nur jeweils die interessierende Zeile (Beobachtung)
*mit dem _TYPE_ "PARMS" (=final Parameter Estimates) bewahrt (und der Rest gelöscht);
%Macro parms_isol (dateinam=, lib=);
DATA &lib..PARMS_sample_&dateinam;
SET &lib..oute_sample_&dateinam;
IF _TYPE_ NE "PARMS" THEN DELETE;
RUN;
TITLE "LINEQS-Modell mit unkorrelierten Fehlern &lib..PARMS_sample_&dateinam";
   run;
%MEND parms_isol;
%macro create_8(howmany);
   %do i=1 %to &howmany;
      %parms_isol (dateinam=ranuni_aug&i, lib=boot_str);
       run;
   %end;
%mend create_8;
%create_8(3000)

*6.;
*die (3000) einzeiligen Datensätze addieren (unter Verwendung eines "iterativen Do-Makros");
%macro create_9(howmany);
   %do i=2 %to &howmany;
   proc APPEND BASE= BOOT_STR.PARMS_SAMPLE_ranuni_aug1
            DATA= BOOT_STR.PARMS_SAMPLE_ranuni_aug&i;
      run;
   %end;
%mend create_9;
%create_9(3000)
*7.;
*Umbenennen, Ausdrucken der Datei;
DATA BOOT_STR.PARMS_ALL_ranuni_aug;
SET BOOT_STR.PARMS_SAMPLE_ranuni_aug1;
RUN;
*Es stellt sich heraus, dass – wie in den AMOS-Modellen auch schon – mitunter negative Werte für
*die Koeffizienten der Pfade geschätzt werden, die vom Faktor zu den manifesten Variablen führen
*(beta1, beta2 und beta3).
*dieser Tatsache soll keine Beachtung geschenkt werden und hat etwas mit dem Schätzalgorithmus
*zu tun. Daher folgende Datenmanipulation:
8.;
*die negativen Werte von beta* werden nun in positive umgewandelt in dem man neue Variablen
*beta*_b erzeugt, die den Betrag von beta* angeben; die alten Variablen werden gelöscht;
*Schließlich werden statistische Maßzahlen und Histogramme für die sechs interessierenden
*Pfadkoeffizienten erstellt;
DATA BOOT_STR.PARMS_ALL_ranuni_aug_B;;
SET BOOT_STR.PARMS_ALL_ranuni_aug;
beta1_b=ABS(beta1);
beta2_b=ABS(beta2);
beta3_b=ABS(beta3);
DROP beta1 beta2 beta3;
LABEL    beta1_b='Betrag von beta1'
         beta2_b='Betrag von beta2'
         beta3_b='Betrag von beta3';
TITLE "Betr. v. PARMS aus LINEQS MODEL m. unkorr. Fehlern (Bootstrap Ziehen v. n=800 einz. Augen)";
RUN;
PROC MEANS DATA=BOOT_STR.PARMS_ALL_ranuni_aug_B;
OUTPUT OUT=BOOT_STR.OUT_PARMS_ALL_RANUNI_aug_B;
TITLE1 "Betr. v. PARMS aus LINEQS MODEL m. unkorr. Fehlern (Bootstrap Ziehen einzelner Augen)";
TITLE2 "n=400 li-re-F- und G-Korr = 0.5";
RUN;
/*PROC PRINT DATA=BOOT_STR.OUT_PARMS_ALL_RANUNI_AUG_B;
RUN;
*/
```

```
DATA BOOT_STR.OUT_B_PARMS_ALL_RANUNI_AUG_B;
SET BOOT_STR.OUT_PARMS_ALL_RANUNI_AUG_B;
IF _STAT_ NE "STD" THEN DELETE;
DROP _TYPE_ _RHS_ _ITER_ beta_f1 beta_f2 beta_f3;
RUN;
/*PROC PRINT DATA=BOOT_STR.OUT_B_PARMS_ALL_RANUNI_AUG_B;
RUN;
PROC GCHART DATA=BOOT_STR.PARMS_ALL_ranuni_aug_B;
VBAR beta1_b beta2_b beta3_b;
TITLE "Betr. v. PARMS aus LINEQS MODEL m. unkorr. Fehlern (Bootstrap Ziehen einzelner Augen)";
TITLE2 "n=800 li-re-Indikator-Korr = 0.9";
RUN;
PROC GCHART DATA=PARMS_ALL_ITERAT_B;
VBAR beta_f1 beta_f2 beta_f3;
TITLE "PARMS aus LINEQS MODEL mit 1 Faktor und unkorrel. Fehlern (1200 Iterationen)";
RUN;
*/
PROC CORR DATA=BOOT_STR.PARMS_ALL_ranuni_aug_B NOSIMPLE;
VAR beta1_b beta2_b beta3_b;
TITLE "Betr. v. PARMS aus LINEQS MODEL m. unkorr. Fehlern (Bootstrap Ziehen einzelner Augen)";
TITLE2 "n=400 li-re-F- und G-Korr = 0.5";
RUN;
*************************************************************************************************
*ad 3.;
******;
/*Umbenennen und Mergen der OUTPUT-Dateien aus PROC MEANS der
Bootstrapanalyse unter Ziehung einzelner Augen und unter Ziehung von Augenpaaren (SD der Faktor-Pfadkoeefizienten)*/
DATA BOOT_STR.OUT_B2_PARMS_ALL_RANUNI_AUG_B;
SET BOOT_STR.OUT_B_PARMS_ALL_RANUNI_AUG_B;
beta1_B_A=beta1_B;
beta2_B_A=beta2_B;
beta3_B_A=beta3_B;
DROP beta1_B beta2_B beta3_B;
RUN;
DATA BOOT_STR.OUT_B2_PARMS_ALL_RANUNI_B;
SET BOOT_STR.OUT_B_PARMS_ALL_RANUNI_B;
beta1_B_AP=beta1_B;
beta2_B_AP=beta2_B;
beta3_B_AP=beta3_B;
DROP beta1_B beta2_B beta3_B;
RUN;
DATA BOOT_STR.OUT_PARMS_ALL_RANUNI_AUG_AUGPAAR;
MERGE BOOT_STR.OUT_B2_PARMS_ALL_RANUNI_AUG_B BOOT_STR.OUT_B2_PARMS_ALL_RANUNI_B;
RUN;
PROC PRINT DATA = BOOT_STR.OUT_PARMS_ALL_RANUNI_AUG_AUGPAAR;
TITLE1 "SD der Pfadkoeff. d. Analysen m. Ziehen v. Augenpaaren und einz. Augen";
TITLE2 "3000 samples n=400 bzw. 800 li-re-F- und G-Korr = 0.5";
RUN;

DATA BOOT_STR.QUOT_PARMS_ALL_RANUNI;
SET BOOT_STR.OUT_PARMS_ALL_RANUNI_AUG_AUGPAAR;
Quot1=beta1_b_ap/beta1_b_a;
Quot2=beta2_b_ap/beta2_b_a;
Quot3=beta3_b_ap/beta3_b_a;
DROP beta1_b_ap beta1_b_a beta2_b_ap beta2_b_a beta3_b_ap beta3_b_a;
RUN;
PROC PRINT DATA = BOOT_STR.QUOT_PARMS_ALL_RANUNI;
TITLE1 "Quotient aus SD der Pfadkoeff. d. Analysen m. Ziehen v. Augenpaaren und einz. Augen";
TITLE2 "3000 samples n=400 bzw. 800 li-re-F- und G-Korr = 0.5";
RUN;
options notes source source2 errors=20;


*************************************************************************************************************************
**********************************
```

11.6.1.3 B Auszüge aus den Programmcodes der Modelle zur Datensimulation der Korrelationsstrukturen K1 bis K5 und K7 bis K16.

```
*K 1

g_li_0 = x5;
 g_re_0 = x6;
 f1_li_0 = x7;
 f1_re_0 = x8;
 f2_li_0 = x9;
 f2_re_0 = x10;
 f3_li_0 = x11;
 f3_re_0= x12;

*K 2

g_li_0=x5;
g_re_0=x6;
f1_li_5=(x1 + x7)/SQRT(2);
f1_re_5= (x1 + x8)/SQRT(2);
f2_li_5=(x2 + x9)/SQRT(2);
f2_re_5= (x2 + x10)/SQRT(2);
f3_li_5=(x3 + x11)/SQRT(2);
f3_re_5= (x3 + x12)/SQRT(2);

*K 3

 g_li_0=x5;
 g_re_0=x6;
 f1_li_7=(x1 + x7*0.6542)/1.19498;
 f1_re_7= (x1 + x8*0.6542)/1.19498;
 f2_li_7=(x2 + x9*0.6542)/1.19498;
 f2_re_7= (x2 + x10*0.6542)/1.19498;
 f3_li_7=(x3 + x11*0.6542)/1.19498;
 f3_re_7= (x3 + x12*0.6542)/1.19498;

*K 4

 g_li_0=x5;
 g_re_0=x6;
 f1_li_9=(x1 + x7*0.333)/1.05399;
 f1_re_9= (x1 + x8*0.333)/1.05399;
 f2_li_9=(x2 + x9*0.333)/1.05399;
 f2_re_9= (x2 + x10*0.333)/1.05399;
 f3_li_9=(x3 + x11*0.333)/1.05399;
 f3_re_9= (x3 + x12*0.333)/1.05399;

*K 5

 g_li_5=(x4 + x5)/SQRT(2);
 g_re_5= (x4 + x6)/SQRT(2);
 f1_li_0 = x7;
 f1_re_0 = x8;
 f2_li_0 = x9;
 f2_re_0 = x10;
 f3_li_0 = x11;
 f3_re_0= x12;

*K 7

  g_li_5=(x4 + x5)/SQRT(2);
 g_re_5= (x4 + x6)/SQRT(2);
 f1_li_7=(x1 + x7*0.6542)/1.19498;
 f1_re_7= (x1 + x8*0.6542)/1.19498;
 f2_li_7=(x2 + x9*0.6542)/1.19498;
 f2_re_7= (x2 + x10*0.6542)/1.19498;
 f3_li_7=(x3 + x11*0.6542)/1.19498;
 f3_re_7= (x3 + x12*0.6542)/1.19498;

*K 8

 g_li_5=(x4 + x5)/SQRT(2);
 g_re_5= (x4 + x6)/SQRT(2);
```

```
 f1_li_9=(x1 + x7*0.333)/1.05399;
 f1_re_9= (x1 + x8*0.333)/1.05399;
 f2_li_9=(x2 + x9*0.333)/1.05399;
 f2_re_9= (x2 + x10*0.333)/1.05399;
 f3_li_9=(x3 + x11*0.333)/1.05399;
 f3_re_9= (x3 + x12*0.333)/1.05399;

*K 9

 g_li_7=(x4 + x5*0.6542)/1.19498;
 g_re_7= (x4 + x6*0.6542)/1.19498;
 f1_li_0 = x7;
 f1_re_0 = x8;
 f2_li_0 = x9;
 f2_re_0 = x10;
 f3_li_0 = x11;
 f3_re_0= x12;

*K 10

g_li_7=(x4 + x5*0.6542)/1.19498;
g_re_7= (x4 + x6*0.6542)/1.19498;
f1_li_5=(x1 + x7)/SQRT(2);
f1_re_5= (x1 + x8)/SQRT(2);
f2_li_5=(x2 + x9)/SQRT(2);
f2_re_5= (x2 + x10)/SQRT(2);
f3_li_5=(x3 + x11)/SQRT(2);
f3_re_5= (x3 + x12)/SQRT(2);


*K 11

g_li_7=(x4 + x5*0.6542)/1.19498;
 g_re_7= (x4 + x6*0.6542)/1.19498;
 f1_li_7=(x1 + x7*0.6542)/1.19498;
 f1_re_7= (x1 + x8*0.6542)/1.19498;
 f2_li_7=(x2 + x9*0.6542)/1.19498;
 f2_re_7= (x2 + x10*0.6542)/1.19498;
 f3_li_7=(x3 + x11*0.6542)/1.19498;
 f3_re_7= (x3 + x12*0.6542)/1.19498;

*K 12

 g_li_7=(x4 + x5*0.6542)/1.19498;
 g_re_7= (x4 + x6*0.6542)/1.19498;
 f1_li_9=(x1 + x7*0.333)/1.05399;
 f1_re_9= (x1 + x8*0.333)/1.05399;
 f2_li_9=(x2 + x9*0.333)/1.05399;
 f2_re_9= (x2 + x10*0.333)/1.05399;
 f3_li_9=(x3 + x11*0.333)/1.05399;
 f3_re_9= (x3 + x12*0.333)/1.05399;

*K 13

 g_li_9=(x4 + x5*0.333)/1.05399;
 g_re_9= (x4 + x6*0.333)/1.05399;
 f1_li_0 = x7;
 f1_re_0 = x8;
 f2_li_0 = x9;
 f2_re_0 = x10;
 f3_li_0 = x11;
 f3_re_0= x12;


*K 14

g_li_9=(x4 + x5*0.333)/1.05399;
g_re_9= (x4 + x6*0.333)/1.05399;
f1_li_5=(x1 + x7)/SQRT(2);
f1_re_5= (x1 + x8)/SQRT(2);
f2_li_5=(x2 + x9)/SQRT(2);
f2_re_5= (x2 + x10)/SQRT(2);
f3_li_5=(x3 + x11)/SQRT(2);
f3_re_5= (x3 + x12)/SQRT(2);
```

```
*K 15

g_li_9=(x4 + x5*0.333)/1.05399;
 g_re_9= (x4 + x6*0.333)/1.05399;
 f1_li_7=(x1 + x7*0.6542)/1.19498;
 f1_re_7= (x1 + x8*0.6542)/1.19498;
 f2_li_7=(x2 + x9*0.6542)/1.19498;
 f2_re_7= (x2 + x10*0.6542)/1.19498;
 f3_li_7=(x3 + x11*0.6542)/1.19498;
 f3_re_7= (x3 + x12*0.6542)/1.19498;

*K 16

 g_li_9=(x4 + x5*0.333)/1.05399;
 g_re_9= (x4 + x6*0.333)/1.05399;
 f1_li_9=(x1 + x7*0.333)/1.05399;
 f1_re_9= (x1 + x8*0.333)/1.05399;
 f2_li_9=(x2 + x9*0.333)/1.05399;
 f2_re_9= (x2 + x10*0.333)/1.05399;
 f3_li_9=(x3 + x11*0.333)/1.05399;
 f3_re_9= (x3 + x12*0.333)/1.05399;
```

11.6.2 Erster Abschnitt, Teil B: (Vergleich der) Bootstrapverfahren auf Basis von Patientendaten anhand eines Zwei-Faktoren-Modells mit acht Diagnoseverfahren

11.6.2.1 Programmcode

11.6.2.1.1 PROC CALIS Lineqs-Bootstrap unter Ziehung einzelner Augen (nicht adjustierte Analyse) und unter Ziehung von Augenpaaren (adjustierte Analyse)

```
*BOOTSTRAPprogramm;
*********************************;
*********************************;
LIBNAME boot_str 'D:\AMOS\237 Pat 140 Kontr\data';
options nonotes nosource nosource2 errors=0;

*1. DATA Step für Ziehung unter Berücksichtigung der Paarigkeit;
*****************************************************************;
/*The RANUNI function returns a number that is generated from the uniform distribution
on the interval (0,1) using a prime modulus multiplicative generator with modulus 231-
and multiplier 397204094 (Fishman and Moore 1982) (See References).
You can use a multiplier to change the length of the interval and an added constant to
move the interval. For example,
random_variate=a*ranuni(seed)+b;
returns a number that is generated from the uniform distribution on the interval (b,a+b).
*/
*Zufallszahlen von 1 bis 237 erzeugen;
%MACRO bootstrap (i, seed, dateinam=, lib=);
DATA &lib..&dateinam;
          DO i=1 to &i;
          random_variate=237*ranuni(&seed)+0.5;
          OUTPUT;
          END;
RUN;
%MEND bootstrap;
%macro create_1(howmany);
   %do i=1 %to &howmany;
      %bootstrap (237, &i, dateinam=ranuni&i, lib=boot_str);
   %end;
%mend create_1;
%create_1(3000)

*Runden;
%MACRO bootstrap2 (dateinam=, lib=);
DATA &lib..r_&dateinam;
SET &lib..&dateinam;
          rand_var_r=round(random_variate);
RUN;
%MEND bootstrap2;
%macro create_2(howmany);
   %do i=1 %to &howmany;
```

```
      %bootstrap2 (dateinam=ranuni&i, lib=boot_str);
   %end;
%mend create_2;
%create_2(3000)

*Sortieren;
%MACRO bootstrap3 (dateinam=, lib=);
PROC SORT DATA=&lib..r_&dateinam OUT=&lib..rs_&dateinam;
BY rand_var_r;
%MEND bootstrap3;
%macro create_3(howmany);
   %do i=1 %to &howmany;
      %bootstrap3 (dateinam=ranuni&i, lib=boot_str);
   %end;
%mend create_3;
%create_3(3000)
*PROC PRINT DATA=boot_str.rs_ranuni30;
*RUN;

*1. DATA Step für Ziehung unter Missachtung der Paarigkeit;
*********************************************************;
*Zufallszahlen von 1 bis 474 erzeugen;
%MACRO bootstrap (i, seed, dateinam=, lib=);
DATA &lib..&dateinam;
         DO i=1 to &i;
         random_variate=474*ranuni(&seed)+0.5;
         OUTPUT;
         END;
RUN;
%MEND bootstrap;
%macro create_1(howmany);
   %do i=1 %to &howmany;
      %bootstrap (474, &i, dateinam=ranuni_aug&i, lib=boot_str);
   %end;
%mend create_1;
%create_1(3000)

*Runden;
%MACRO bootstrap2 (dateinam=, lib=);
DATA &lib..r_&dateinam;
SET &lib..&dateinam;
         rand_var_r=round(random_variate);
RUN;
%MEND bootstrap2;
%macro create_2(howmany);
   %do i=1 %to &howmany;
      %bootstrap2 (dateinam=ranuni_aug&i, lib=boot_str);
   %end;
%mend create_2;
%create_2(3000)

*Sortieren;
%MACRO bootstrap3 (dateinam=, lib=);
PROC SORT DATA=&lib..r_&dateinam OUT=&lib..rs_&dateinam;
BY rand_var_r;
%MEND bootstrap3;
%macro create_3(howmany);
   %do i=1 %to &howmany;
      %bootstrap3 (dateinam=ranuni_aug&i, lib=boot_str);
   %end;
%mend create_3;
%create_3(3000)
*PROC PRINT DATA=BOOT_STR.RS_RANUNI_AUG30;
*RUN;


*II;
***
*800 A u g e n von 400 Patienten ziehen aus dem gleichen Datensatz, aus dem 400 P a t i e n t e n
*mit ihren jeweiligen Augenpaaren gezogen wurden (jeweils mit Zurücklegen);
/*Addieren der Datendatei linker Augen, aus der Bootstrapsamples gezogen werden
und der Datendatei rechter Augen, aus der Bootstrapsamples gezogen werden*/
*******************;
*Ab hier Hochzählen;
*******************;

*EINLESEN der SPSS-DATEI von PM vergl_3_2 ;
*****************************************;
LIBNAME in SPSS 'D:\AMOS\237 Pat 140 Kontr\Patienten 2 mal 237.por';
DATA Boot_str.Patienten_2_mal_237;
SET in.spssdat;
KEEP PATNR AUGSEIT NRR ERG VEP AMP FLI CLV MD DYN;
RUN;
*Teilen in Dateien linker und rechter Augen;
```

```
DATA Boot_str.Patienten_237_l;
SET Boot_str.Patienten_2_mal_237;
IF AUGSEIT= 'L';
RUN;
DATA Boot_str.Patienten_237_r;
SET Boot_str.Patienten_2_mal_237;
IF AUGSEIT= 'R';
RUN;
*Zuordnen einer fortlaufenden Nummer;
DATA Boot_str.Patienten_237_lb;
SET Boot_str.Patienten_237_l;
rand_var_r=_N_;
RUN;
DATA Boot_str.Patienten_237_rb;
SET Boot_str.Patienten_237_r;
rand_var_r=_N_;
RUN;
**********************;
/*
DATA BOOT_STR.BLR_ITERAT2;
SET BOOT_STR.BL_ITERAT2 BOOT_STR.BR_ITERAT2;
rand_var_r=_N_;
RUN;
PROC SORT DATA=BOOT_STR.BLR_ITERAT2 OUT=BOOT_STR.BLRS_ITERAT2;
BY rand_var_r;
RUN;
PROC CORR DATA=BOOT_STR.BLRS_ITERAT2;
VAR indik_1 indik_2 indik_3;
RUN;

*Die Korrelation zwischen dem linken Wert und dem rechten Wert eines jeweiligen Indikators
*Ausgangsdatensatz umstellen
*a) seitenneutrale Variablen in seitenanzeigende Variablen umbenennen;
DATA BOOT_STR.BL_UK_ITERAT2;
SET BOOT_STR.BL_ITERAT2;
indik1_li = indik_1;
indik2_li = indik_2;
indik3_li = indik_3;
DROP indik_1 indik_2 indik_3 seite;
RUN;
DATA BOOT_STR.BR_UK_ITERAT2;
SET BOOT_STR.BR_ITERAT2;
indik1_re = indik_1;
indik2_re = indik_2;
indik3_re = indik_3;
DROP indik_1 indik_2 indik_3 seite;
RUN;
*b) match-mergen der umbenannten Datendateien;
DATA BOOT_STR.BLR_UK_ITERAT2;
MERGE BOOT_STR.BL_UK_ITERAT2 BOOT_STR.BR_UK_ITERAT2;
BY rand_var_r;
RUN;
PROC CORR DATA=BOOT_STR.BLR_UK_ITERAT2;
VAR INDIK1_li INDIK1_re;
TITLE1 "Analysen m. Ziehen v. Augenpaaren und einz. Augen";
TITLE2 "3000 samples n=400 bzw. 800 li-re-Indikator-Korr = 0.5";
RUN;
PROC CORR DATA=BOOT_STR.BLR_UK_ITERAT2;
VAR INDIK2_li INDIK2_re;
RUN;
PROC CORR DATA=BOOT_STR.BLR_UK_ITERAT2;
VAR INDIK3_li INDIK3_re;
RUN;
PROC CORR DATA=BOOT_STR.BLR_UK_ITERAT2;
VAR INDIK1_li INDIK2_li INDIK3_li;
RUN;
PROC CORR DATA=BOOT_STR.BLR_UK_ITERAT2;
VAR INDIK1_re INDIK2_re INDIK3_re;
RUN;
*/

*****************************************************************************
*hier geht das Mergen für augenpaarbezogene Ziehung und alle weiteren Prozesse los

*****************************************************************************
*3 A. DATA STEP für die linke Seite;
*********************************;
*Mergen;
*entspricht nach SAS-Language Concept dem Beispiel 'Match-Merged with Duplicate BY Values';
%MACRO bootstrap4 (dateinam=, lib=);
DATA &lib..merged_l_&dateinam;
MERGE &lib..rs_&dateinam &lib..Patienten_237_lb;
BY rand_var_r;
%MEND bootstrap4;
```

```
%macro create_4(howmany);
   %do i=1 %to &howmany;
      %bootstrap4 (dateinam=ranuni&i, lib=boot_str);
   %end;
%mend create_4;
%create_4(3000)
*Löschen der 'nicht gezogenen' Records in den "gemergeten" Datendateien;
%MACRO bootstrap5 (dateinam=, lib=);
DATA &lib..sample_l_&dateinam;
SET &lib..merged_l_&dateinam;
IF random_variate EQ "." THEN DELETE;
DROP random_variate i;
RUN;
%MEND bootstrap5;
%macro create_5(howmany);
   %do i=1 %to &howmany;
      %bootstrap5 (dateinam=ranuni&i, lib=boot_str);
   %end;
%mend create_5;
%create_5(3000)

*3 B. DATA STEP für die rechte Seite;
*********************************;
%MACRO bootstrap4 (dateinam=, lib=);
DATA &lib..merged_r_&dateinam;
MERGE &lib..rs_&dateinam &lib..Patienten_237_rb;
BY rand_var_r;
%MEND bootstrap4;
%macro create_4(howmany);
   %do i=1 %to &howmany;
      %bootstrap4 (dateinam=ranuni&i, lib=boot_str);
   %end;
%mend create_4;
%create_4(3000)
*Löschen der 'nicht gezogenen' Records in den "gemergeten" Datendateien;
%MACRO bootstrap5 (dateinam=, lib=);
DATA &lib..sample_r_&dateinam;
SET &lib..merged_r_&dateinam;
IF random_variate EQ "." THEN DELETE;
DROP random_variate i;
RUN;
%MEND bootstrap5;
%macro create_5(howmany);
   %do i=1 %to &howmany;
      %bootstrap5 (dateinam=ranuni&i, lib=boot_str);
   %end;
%mend create_5;
%create_5(3000)

*Addieren der linken und rechten Augen;
***************************************;
%MACRO bootstrap6 (dateinam=, lib=);
DATA &lib..sample_&dateinam;
SET &lib..sample_l_&dateinam &lib..sample_r_&dateinam;
LABEL augseit='Augenseite';
RUN;
%MEND bootstrap6;
%macro create_6(howmany);
   %do i=1 %to &howmany;
      %bootstrap6 (dateinam=ranuni&i, lib=boot_str);
   %end;
%mend create_6;
%create_6(3000)

*Sortieren nach Patient;
%MACRO bootstrap7 (dateinam=, lib=);
PROC SORT DATA=&lib..sample_&dateinam;
        BY rand_var_r augseit;
%MEND bootstrap7;
%macro create_7(howmany);
   %do i=1 %to &howmany;
      %bootstrap7 (dateinam=ranuni&i, lib=boot_str);
   %end;
%mend create_7;
%create_7(3000)
*********************************************************************;
%Macro proccalis (dateinam=, lib=);
proc calis DATA=&lib..sample_&dateinam OUTEST=&lib..oute_sample_&dateinam OUTRAM=&lib..outra_sample_&dateinam tech=nr
nop edf=473;
         Lineqs
         NRR=              beta_nrr_g        Fakt_G   + beta_nrr_e       Err_nrr,
         ERG=              beta_erg_g        Fakt_G   + beta_erg_e       Err_erg,
         VEP=              beta_vep_g        Fakt_G   + beta_vep_e       Err_vep,
         AMP=              beta_amp_g        Fakt_G   + beta_amp_k       Fakt_K   + beta_amp_e      Err_amp,
```

```
         FLI=              beta_fli_g        Fakt_G   + beta_fli_k      Fakt_K   + beta_fli_e       Err_fli,
         CLV=              beta_clv_g        Fakt_G   + beta_clv_k      Fakt_K   + beta_clv_e       Err_clv,
         MD =              beta_md_g         Fakt_G   + beta_md_k       Fakt_K   + beta_md_e        Err_md,
         DYN=              beta_dyn_g        Fakt_G   + beta_dyn_k      Fakt_K   + beta_dyn_e       Err_dyn;

         Std
                  Err_nrr = 1,
                  Err_erg = 1,
                  Err_vep  = 1,
                  Err_amp = 1,
                  Err_fli  = 1,
                  Err_clv  = 1,
                  Err_md   = 1,
                  Err_dyn  = 1,
                  Fakt_G= 1,
                  Fakt_K= 1;
         Cov
                  Err_nrr Err_dyn = The1,
                  Err_vep Err_amp = The2,
                  Err_vep Err_md = The3,
                  Err_amp Err_fli =The4,
                  Err_fli Err_clv = The5;
                  run;
         %MEND proccalis;
%macro create_7(howmany);
   %do i=1 %to &howmany;
      %proccalis (dateinam=ranuni&i, lib=boot_str);
       run;
   %end;
%mend create_7;
%create_7(3000)
*5.;
*Von den (3000) OUTEST-sas-data-sets wird nur jeweils die interessierende Zeile (Beobachtung)
*mit dem _TYPE_ "PARMS" (=final Parameter Estimates) bewahrt (und der Rest gelöscht);
%Macro parms_isol (dateinam=, lib=);
DATA &lib..PARMS_sample_&dateinam;
SET &lib..oute_sample_&dateinam;
IF _TYPE_ NE "PARMS" THEN DELETE;
RUN;
%MEND parms_isol;
%macro create_8(howmany);
   %do i=1 %to &howmany;
      %parms_isol (dateinam=ranuni&i, lib=boot_str);
       run;
   %end;
%mend create_8;
%create_8(3000)
*6.;
*die (3000) einzeiligen Datensätze addieren (unter Verwendung eines „iterativen Do-Makros“);
%macro create_9(howmany);
   %do i=2 %to &howmany;
   proc APPEND BASE= BOOT_STR.PARMS_SAMPLE_RANUNI1
            DATA= BOOT_STR.PARMS_SAMPLE_RANUNI&i;
      run;
   %end;
%mend create_9;
%create_9(3000)
*7.;
*Umbenennen, Ausdrucken, Labeln der Datei;
*                 Err_nrr Err_dyn = The1,
                  Err_vep Err_amp = The2,
                  Err_vep Err_md = The3,
                  Err_amp Err_fli =The4,
                  Err_fli Err_clv = The5;
DATA BOOT_STR.PARMS_ALL_RANUNI;
SET BOOT_STR.PARMS_SAMPLE_RANUNI1;
LABEL
 The1='Corr Err_nrr Err_dyn'
 The2='Corr Err_vep Err_amp'
 The3='Corr Err_vep Err_md'
 The4='Corr Err_amp Err_fli'
 The5='Corr Err_fli Err_clv';
RUN;
PROC MEANS DATA=BOOT_STR.PARMS_ALL_RANUNI; OUTPUT OUT=BOOT_STR.OUT_PARMS_ALL_RANUNI;
TITLE1 "Bootstrapanalyse Patientendaten – Ziehen v. Augenpaaren)";
TITLE2 "30 samples n=237";
RUN;
PROC GCHART DATA=BOOT_STR.PARMS_ALL_RANUNI;
VBAR beta_nrr_g
          beta_erg_g
          beta_vep_g
          beta_amp_g
          beta_fli_g
          beta_clv_g
```

```
          beta_md_g
          beta_dyn_g
          beta_amp_k
          beta_fli_k
          beta_clv_k
          beta_md_k
          beta_dyn_k
          the1
          the2
          the3
          the4
          the5;
TITLE1 "Bootstrapanalyse Patientendaten – Ziehen v. Augenpaaren)";
TITLE2 "3000 samples n=237";
RUN;
PROC UNIVARIATE DATA=BOOT_STR.PARMS_ALL_RANUNI;
VAR the3
          the4;
TITLE1 "Bootstrapanalyse Patientendaten - Ziehen v. Augenpaaren)";
TITLE2 "3000 samples n=237; Fehlerkorrelationen";
RUN;

*Es stellt sich heraus, dass – wie in den AMOS-Modellen auch schon – mitunter negative Werte für
*die Koeffizienten der Pfade geschätzt werden, die vom Faktor zu den manifesten Variablen führen
*(beta1, beta2 und beta3).
*dieser Tatsache soll keine Beachtung geschenkt werden und hat etwas mit dem Schätzalgorithmus
*zu tun. Daher folgende Datenmanipulation:
8.;
*die negativen Werte von beta* werden nun in positive umgewandelt, in dem man neue Variablen
*beta*_b erzeugt, die den Betrag von beta* angeben;
*die alten Variablen werden gelöscht;
*Schließlich werden statistische Maßzahlen und Histogramme für die sechs interessierenden Pfadkoeffizienten erstellt;
DATA BOOT_STR.PARMS_ALL_RANUNI_B;
SET BOOT_STR.PARMS_ALL_RANUNI;
                  beta_nrr_g_b =ABS(beta_nrr_g);
                  beta_erg_g_b =ABS(beta_erg_g);
                  beta_vep_g_b =ABS(beta_vep_g);
                  beta_amp_g_b =ABS(beta_amp_g);
                  beta_fli_g_b =ABS(beta_fli_g);
                  beta_clv_g_b =ABS(beta_clv_g);
                  beta_md_g_b =ABS(beta_md_g);
                  beta_dyn_g_b =ABS(beta_dyn_g);
                  beta_amp_k_b =ABS(beta_amp_k);
                  beta_fli_k_b =ABS(beta_fli_k);
                  beta_clv_k_b =ABS(beta_clv_k);
                  beta_md_k_b  =ABS(beta_md_k);
                  beta_dyn_k_b =ABS(beta_dyn_k);
                  beta_nrr_e_b =ABS(beta_nrr_e);
                  beta_erg_e_b =ABS(beta_erg_e);
                  beta_vep_e_b =ABS(beta_vep_e);
                  beta_amp_e_b =ABS(beta_amp_e);
                  beta_fli_e_b =ABS(beta_fli_e);
                  beta_clv_e_b =ABS(beta_clv_e);
                  beta_md_e_b =ABS(beta_md_e);
                  beta_dyn_e_b =ABS(beta_dyn_e);
                  the1_b=ABS(the1);
                  the2_b=ABS(the2);
                  the5_b=ABS(the5);

DROP
beta_nrr_g
beta_erg_g
beta_vep_g
beta_amp_g
beta_fli_g
beta_clv_g
beta_md_g
beta_dyn_g
beta_amp_k
beta_fli_k
beta_clv_k
beta_md_k
beta_dyn_k
beta_nrr_e
beta_erg_e
beta_vep_e
beta_amp_e
beta_fli_e
beta_clv_e
beta_md_e
beta_dyn_e
the1
the2
the5;
```

```
LABEL
the1_b ='Betr. v. Fehlerkorr nrr dyn'
the2_b='Betr. v. Fehlerkorr vep amp'
the5_b='Betr. v. Fehlerkorr fli clv';
TITLE ;
RUN;
PROC MEANS DATA=BOOT_STR.PARMS_ALL_RANUNI_B; OUTPUT OUT=BOOT_STR.OUT_PARMS_ALL_RANUNI_B;
TITLE1 "Bootstrapanalyse Patientendaten – Ziehen v. Augenpaaren";
TITLE2 "3000 samples n=237";
RUN;
PROC UNIVARIATE DATA=BOOT_STR.PARMS_ALL_RANUNI_B;
VAR beta_amp_k_b
         beta_fli_k_b
         beta_clv_k_b
         beta_md_k_b
    beta_dyn_k_b;
TITLE1 "Bootstrapanalyse Patientendaten - Ziehen v. Augenpaaren";
TITLE2 "3000 samples n=237 – Pfadkoeff. des Faktors Konzentration";
RUN;
PROC UNIVARIATE DATA=BOOT_STR.PARMS_ALL_RANUNI_B;
VAR the1_b
the2_b
the5_b;
TITLE1 "Bootstrapanalyse Patientendaten – Ziehen v. Augenpaaren";
TITLE2 "3000 samples n=237 – Fehlerkorrelationen";
RUN;
DATA BOOT_STR.OUT_B_PARMS_ALL_RANUNI_B;
SET BOOT_STR.OUT_PARMS_ALL_RANUNI_B;
IF _STAT_ NE "STD" THEN DELETE;
DROP _TYPE_ _RHS_ _ITER_;
RUN;

*dieser Output STD (Empirische Standardfehler bzw. neuerdings 'simulierter Standardfehler')
*dient dem späteren Vergleich Quotient aus augenbez. und personenbez. Bootstrap-SE;
PROC GCHART DATA=BOOT_STR.PARMS_ALL_RANUNI_B;
VBAR beta_nrr_g_b
          beta_erg_g_b
          beta_vep_g_b
          beta_amp_g_b
          beta_fli_g_b
          beta_clv_g_b
          beta_md_g_b
          beta_dyn_g_b
          beta_amp_k_b
          beta_fli_k_b
          beta_clv_k_b
          beta_md_k_b
          beta_dyn_k_b
          the1_b
          the2_b
          the3
          the4
          the5_b;
TITLE1 "Bootstrapanalyse Patientendaten – Ziehen v. Augenpaaren)";
TITLE2 "3000 samples n=237";
RUN;

*************************************************************************************************
*3. DATA STEP für die Ziehung von einzelnen Augen ohne Berücksichtigung der Paarigkeit;
*************************************************************************************;
*Vorbereiten der Empir. Datendatei:
*Zuordnen einer Fortlaufenden Nummer;
DATA Boot_str.Patienten_2_mal_237_b;
SET Boot_str.Patienten_2_mal_237;
rand_var_r=_N_;
RUN;
*Mergen;
%MACRO bootstrap4 (dateinam=, lib=);
DATA &lib..merged_&dateinam;
MERGE &lib..rs_&dateinam &lib..Patienten_2_mal_237_b;
BY rand_var_r;
%MEND bootstrap4;
%macro create_4(howmany);
   %do i=1 %to &howmany;
      %bootstrap4 (dateinam=ranuni_aug&i, lib=boot_str);
   %end;
%mend create_4;
%create_4(3000)
*Löschen der 'nicht gezogenen' Records in den "gemergeten" Datendateien;
%MACRO bootstrap5 (dateinam=, lib=);
DATA &lib..sample_&dateinam;
SET &lib..merged_&dateinam;
IF random_variate EQ "." THEN DELETE;
DROP random_variate i;
```

```
RUN;
%MEND bootstrap5;
%macro create_5(howmany);
   %do i=1 %to &howmany;
      %bootstrap5 (dateinam=ranuni_aug&i, lib=boot_str);
   %end;
%mend create_5;
%create_5(3000)
%Macro proccalis (dateinam=, lib=);
proc calis DATA=&lib..sample_&dateinam OUTEST=&lib..oute_sample_&dateinam OUTRAM=&lib..outra_sample_&dateinam tech=nr
nop edf=473;
         Lineqs
         NRR=     beta_nrr_g        Fakt_G   + beta_nrr_e       Err_nrr,
         ERG=     beta_erg_g        Fakt_G   + beta_erg_e       Err_erg,
         VEP=     beta_vep_g        Fakt_G   + beta_vep_e       Err_vep,
         AMP=     beta_amp_g        Fakt_G   + beta_amp_k      Fakt_K   + beta_amp_e      Err_amp,
         FLI=     beta_fli_g        Fakt_G   + beta_fli_k      Fakt_K   + beta_fli_e      Err_fli,
         CLV=     beta_clv_g        Fakt_G   + beta_clv_k      Fakt_K   + beta_clv_e      Err_clv,
         MD =     beta_md_g         Fakt_G   + beta_md_k       Fakt_K   + beta_md_e       Err_md,
         DYN=     beta_dyn_g        Fakt_G   + beta_dyn_k      Fakt_K   + beta_dyn_e      Err_dyn;

         Std
                  Err_nrr = 1,
                  Err_erg = 1,
                  Err_vep  = 1,
                  Err_amp = 1,
                  Err_fli  = 1,
                  Err_clv  = 1,
                  Err_md   = 1,
                  Err_dyn  = 1,
                  Fakt_G= 1,
                  Fakt_K= 1;
         Cov
                  Err_nrr Err_dyn = The1,
                  Err_vep Err_amp = The2,
                  Err_vep Err_md = The3,
                  Err_amp Err_fli =The4,
                  Err_fli Err_clv = The5;
                  run;
         %MEND proccalis;

%macro create_7(howmany);
   %do i=1 %to &howmany;
      %proccalis (dateinam=ranuni_aug&i, lib=boot_str);
       run;
   %end;
%mend create_7;
%create_7(3000)
*5.;
*Von den (3000) OUTEST-sas-data-sets wird nur jeweils die interessierende Zeile (Beobachtung)
*mit dem _TYPE_ "PARMS" (=final Parameter Estimates) bewahrt (und der Rest gelöscht);
%Macro parms_isol (dateinam=, lib=);
DATA &lib..PARMS_sample_&dateinam;
SET &lib..oute_sample_&dateinam;
IF _TYPE_ NE "PARMS" THEN DELETE;
RUN;
TITLE;
   run;
%MEND parms_isol;
%macro create_8(howmany);
   %do i=1 %to &howmany;
      %parms_isol (dateinam=ranuni_aug&i, lib=boot_str);
       run;
   %end;
%mend create_8;
%create_8(3000)

*6.;
*die (3000) einzeiligen Datensätze addieren (unter Verwendung eines „iterativen Do-Makros“);
%macro create_9(howmany);
   %do i=2 %to &howmany;
   proc APPEND BASE= BOOT_STR.PARMS_SAMPLE_ranuni_aug1
            DATA= BOOT_STR.PARMS_SAMPLE_ranuni_aug&i;
      run;
   %end;
%mend create_9;
%create_9(3000)
*7.;
*Umbenennen, Ausdrucken, Labeln der Datei;
DATA BOOT_STR.PARMS_ALL_ranuni_aug;
SET BOOT_STR.PARMS_SAMPLE_ranuni_aug1;
LABEL
 The1='Corr Err_nrr Err_dyn'
 The2='Corr Err_vep Err_amp'
```

```
 The3='Corr Err_vep Err_md'
 The4='Corr Err_amp Err_fli'
 The5='Corr Err_fli Err_clv';
RUN;
PROC GCHART DATA=BOOT_STR.PARMS_ALL_ranuni_aug;
VBAR beta_nrr_g
          beta_erg_g
          beta_vep_g
          beta_amp_g
          beta_fli_g
          beta_clv_g
          beta_md_g
          beta_dyn_g
          beta_amp_k
          beta_fli_k
          beta_clv_k
          beta_md_k
          beta_dyn_k
          the1
          the2
          the3
          the4
          the5;
TITLE1 "Bootstrapanalyse Patientendaten – Ziehen v. einzelnen Augen";
TITLE2 "Rohwerte – 3000 samples n=474";
RUN;
PROC UNIVARIATE DATA=BOOT_STR.PARMS_ALL_ranuni_aug;
VAR the3
          the4;
TITLE1 "Bootstrapanalyse Patientendaten – Ziehen v. einzelnen Augen";
TITLE2 "Rohwerte – 3000 samples n=474 – Fehlerkorrelationen";
RUN;

*Es stellt sich heraus, dass – wie in den AMOS-Modellen auch schon – mitunter negative Werte für
*die Koeffizienten der Pfade geschätzt werden, die vom Faktor zu den manifesten Variablen führen
*(beta1, beta2 und beta3).
*dieser Tatsache soll keine Beachtung geschenkt werden und hat etwas mit dem Schätzalgorithmus
*zu tun. Daher folgende Datenmanipulation:
8.;
*die negativen Werte von beta* werden nun in positive umgewandelt, in dem man neue Variablen
*beta*_b erzeugt, die den Betrag von beta* angeben;
*die alten Variablen werden gelöscht;
*Schließlich werden statistische Maßzahlen und Histogramme für die sechs interessierenden
*Pfadkoeffizienten erstellt;
DATA BOOT_STR.PARMS_ALL_ranuni_aug_B;
SET BOOT_STR.PARMS_ALL_ranuni_aug;
                  beta_nrr_g_b =ABS(beta_nrr_g);
                  beta_erg_g_b =ABS(beta_erg_g);
                  beta_vep_g_b =ABS(beta_vep_g);
                  beta_amp_g_b =ABS(beta_amp_g);
                  beta_fli_g_b =ABS(beta_fli_g);
                  beta_clv_g_b =ABS(beta_clv_g);
                  beta_md_g_b  =ABS(beta_md_g);
                  beta_dyn_g_b =ABS(beta_dyn_g);
                  beta_amp_k_b =ABS(beta_amp_k);
                  beta_fli_k_b =ABS(beta_fli_k);
                  beta_clv_k_b =ABS(beta_clv_k);
                  beta_md_k_b  =ABS(beta_md_k);
                  beta_dyn_k_b =ABS(beta_dyn_k);
                  beta_nrr_e_b =ABS(beta_nrr_e);
                  beta_erg_e_b =ABS(beta_erg_e);
                  beta_vep_e_b =ABS(beta_vep_e);
                  beta_amp_e_b =ABS(beta_amp_e);
                  beta_fli_e_b =ABS(beta_fli_e);
                  beta_clv_e_b =ABS(beta_clv_e);
                  beta_md_e_b =ABS(beta_md_e);
                  beta_dyn_e_b =ABS(beta_dyn_e);
                  the1_b=ABS(the1);
                  the2_b=ABS(the2);
                  the5_b=ABS(the5);

DROP
beta_nrr_g
beta_erg_g
beta_vep_g
beta_amp_g
beta_fli_g
beta_clv_g
beta_md_g
beta_dyn_g
beta_amp_k
beta_fli_k
beta_clv_k
beta_md_k
```

```
beta_dyn_k
beta_nrr_e
beta_erg_e
beta_vep_e
beta_amp_e
beta_fli_e
beta_clv_e
beta_md_e
beta_dyn_e
the1
the2
the5;
LABEL
the1_b ='Betr. v. Fehlerkorr nrr dyn'
the2_b='Betr. v. Fehlerkorr vep amp'
the5_b='Betr. v. Fehlerkorr fli clv';
TITLE ;
RUN;

PROC MEANS DATA=BOOT_STR.PARMS_ALL_ranuni_aug_B;
OUTPUT OUT=BOOT_STR.OUT_PARMS_ALL_RANUNI_aug_B;
TITLE1 "Bootstrapanalyse Patientendaten – Ziehen v. einzelnen Augen)";
TITLE2 "3000 samples n=474";
RUN;
PROC UNIVARIATE DATA=BOOT_STR.PARMS_ALL_RANUNI_aug_B;
VAR beta_amp_k_b
         beta_fli_k_b
         beta_clv_k_b
         beta_md_k_b
         beta_dyn_k_b;
TITLE1 "Bootstrapanalyse Patientendaten – Ziehen v. einzelnen Augen";
TITLE2 "3000 samples n=474 – Pfadkoeff. des Faktors Konzentration";
RUN;
PROC UNIVARIATE DATA=BOOT_STR.PARMS_ALL_RANUNI_aug_B;
VAR the1_b
the2_b
the5_b
;
TITLE1 "Bootstrapanalyse Patientendaten – Ziehen v. einzelnen Augen";
TITLE2 "3000 samples n=474 – Fehlerkorrelationen";
RUN;

*Zur Vorbereitung der SE-Auswertung;
DATA BOOT_STR.OUT_B_PARMS_ALL_RANUNI_AUG_B;
SET BOOT_STR.OUT_PARMS_ALL_RANUNI_AUG_B;
IF _STAT_ NE "STD" THEN DELETE;
DROP _TYPE_ _RHS_ _ITER_;
RUN;
PROC GCHART DATA=BOOT_STR.PARMS_ALL_ranuni_aug_B;
VBAR beta_nrr_g_b
          beta_erg_g_b
          beta_vep_g_b
          beta_amp_g_b
          beta_fli_g_b
          beta_clv_g_b
          beta_md_g_b
          beta_dyn_g_b
          beta_amp_k_b
          beta_fli_k_b
          beta_clv_k_b
          beta_md_k_b
          beta_dyn_k_b
          the1_b
          the2_b
          the5_b;
TITLE1 "Bootstrapanalyse Patientendaten – Ziehen v. einzelnen Augen";
TITLE2 "Absolute Werte – 3000 samples n=474";
RUN;
*************************************************************************************************
```

11.6.2.1.2 Vergleich des nichtadjustierten und adjustierten Bootstrapverfahrens

```
*ad 3.;
******;
/*Umbenennen und Mergen der OUTPUT DATEIEN aus PROC MEANS der
Bootstrapanalyse unter Ziehung einzelner Augen und
unter Ziehung von Augenpaaren
(SD der Faktor-Pfadkoeefizienten)*/
DATA BOOT_STR.OUT_B2_PARMS_ALL_RANUNI_AUG_B;
SET BOOT_STR.OUT_B_PARMS_ALL_RANUNI_AUG_B;
```

```
        beta_nrr_g_b_A =beta_nrr_g_b;
                  beta_erg_g_b_A=beta_erg_g_b;
                  beta_vep_g_b_A=beta_vep_g_b;
                  beta_amp_g_b_A= beta_amp_g_b;
                  beta_fli_g_b_A= beta_fli_g_b;
                  beta_clv_g_b_A= beta_clv_g_b ;
                  beta_md_g_b_A= beta_md_g_b   ;
                  beta_dyn_g_b_A= beta_dyn_g_b;
                  beta_amp_k_b_A= beta_amp_k_b;
                  beta_fli_k_b_A= beta_fli_k_b;
                  beta_clv_k_b_A= beta_clv_k_b;
                  beta_md_k_b_A= beta_md_k_b;
                  beta_dyn_k_b_A= beta_dyn_k_b;
                  beta_nrr_e_b_A= beta_nrr_e_b;
                  beta_erg_e_b_A= beta_erg_e_b;
                  beta_vep_e_b_A= beta_vep_e_b;
                  beta_amp_e_b_A= beta_amp_e_b;
                  beta_fli_e_b_A= beta_fli_e_b;
                  beta_clv_e_b_A= beta_clv_e_b;
                  beta_md_e_b_A= beta_md_e_b;
                  beta_dyn_e_b_A= beta_dyn_e_b;
                  the1_b_A= the1_b;
                  the2_b_A= the2_b;
                  the3_A= the3;
                  the4_A= the4;
                  the5_b_A= the5_b;
DROP      beta_nrr_g_b
                  beta_erg_g_b
                  beta_vep_g_b
                  beta_amp_g_b
                  beta_fli_g_b
                  beta_clv_g_b
                  beta_md_g_b
                  beta_dyn_g_b
                  beta_amp_k_b
              beta_fli_k_b
                  beta_clv_k_b
                  beta_md_k_b
                  beta_dyn_k_b
              beta_nrr_e_b
                  beta_erg_e_b
                  beta_vep_e_b
                  beta_amp_e_b
                  beta_fli_e_b
                  beta_clv_e_b
              beta_md_e_b
                  beta_dyn_e_b
                  the1_b
                  the2_b
                  the3
                  the4
                  the5_b;
RUN;
DATA BOOT_STR.OUT_B2_PARMS_ALL_RANUNI_B;
SET BOOT_STR.OUT_B_PARMS_ALL_RANUNI_B;
                  beta_nrr_g_b_AP = beta_nrr_g_b;
                  beta_erg_g_b_AP= beta_erg_g_b;
                  beta_vep_g_b_AP= beta_vep_g_b;
                  beta_amp_g_b_AP= beta_amp_g_b;
                  beta_fli_g_b_AP= beta_fli_g_b;
                  beta_clv_g_b_AP=beta_clv_g_b ;
                  beta_md_g_b_AP= beta_md_g_b  ;
                  beta_dyn_g_b_AP= beta_dyn_g_b;
                  beta_amp_k_b_AP= beta_amp_k_b;
                  beta_fli_k_b_AP= beta_fli_k_b;
                  beta_clv_k_b_AP= beta_clv_k_b;
                  beta_md_k_b_AP= beta_md_k_b;
                  beta_dyn_k_b_AP= beta_dyn_k_b;
                  beta_nrr_e_b_AP= beta_nrr_e_b;
                  beta_erg_e_b_AP= beta_erg_e_b;
                  beta_vep_e_b_AP= beta_vep_e_b;
                  beta_amp_e_b_AP= beta_amp_e_b;
                  beta_fli_e_b_AP= beta_fli_e_b;
                  beta_clv_e_b_AP= beta_clv_e_b;
                  beta_md_e_b_AP= beta_md_e_b;
                  beta_dyn_e_b_AP= beta_dyn_e_b;
                  the1_b_AP= the1_b;
                  the2_b_AP= the2_b;
                  the3_AP= the3;
                  the4_AP= the4;
                  the5_b_AP= the5_b;
DROP      beta_nrr_g_b
                  beta_erg_g_b
                  beta_vep_g_b
```

```
                    beta_amp_g_b
                    beta_fli_g_b
                    beta_clv_g_b
                    beta_md_g_b
                    beta_dyn_g_b
                    beta_amp_k_b
               beta_fli_k_b
                    beta_clv_k_b
                    beta_md_k_b
                    beta_dyn_k_b
               beta_nrr_e_b
                    beta_erg_e_b
                    beta_vep_e_b
                    beta_amp_e_b
                    beta_fli_e_b
                    beta_clv_e_b
               beta_md_e_b
                    beta_dyn_e_b
                    the1_b
                    the2_b
                    the3
                    the4
                    the5_b;
RUN;
DATA BOOT_STR.OUT_PARMS_ALL_RANUNI_AUG_AUGPAAR;
MERGE BOOT_STR.OUT_B2_PARMS_ALL_RANUNI_AUG_B BOOT_STR.OUT_B2_PARMS_ALL_RANUNI_B;
RUN;
PROC PRINT DATA = BOOT_STR.OUT_PARMS_ALL_RANUNI_AUG_AUGPAAR;
TITLE1 "SD der Pfadkoeff. bzw. einer Fehlerkorr. d. BS-Analysen";
TITLE2 "Ziehen v. Augenpaaren und einz. Augen – 3000 samples n=237 bzw. 474";
RUN;

DATA BOOT_STR.QUOT_PARMS_ALL_RANUNI;
SET BOOT_STR.OUT_PARMS_ALL_RANUNI_AUG_AUGPAAR;
Quot_beta_nrr_g =  beta_nrr_g_b_AP / beta_nrr_g_b_A;
Quot_beta_erg_g =  beta_erg_g_b_AP / beta_erg_g_b_A;
Quot_beta_vep_g =  beta_vep_g_b_AP / beta_vep_g_b_A;
Quot_beta_amp_g =  beta_amp_g_b_AP / beta_amp_g_b_A;
Quot_beta_fli_g =  beta_fli_g_b_AP / beta_fli_g_b_A;
Quot_beta_clv_g =  beta_clv_g_b_AP / beta_clv_g_b_A;
Quot_beta_md_g =   beta_md_g_b_AP / beta_md_g_b_A;
Quot_eta_dyn_g =   beta_dyn_g_b_AP / beta_dyn_g_b_A;
Quot_beta_amp_k =  beta_amp_k_b_AP / beta_amp_k_b_A;
Quot_beta_fli_k    =        beta_fli_k_b_AP / beta_fli_k_b_A;
Quot_beta_clv_k    =        beta_clv_k_b_AP / beta_clv_k_b_A;
Quot_beta_md_k     =        beta_md_k_b_AP / beta_md_k_b_A;
Quot_beta_dyn_k    =        beta_dyn_k_b_AP / beta_dyn_k_b_A;
Quot_beta_nrr_e    =        beta_nrr_e_b_AP / beta_nrr_e_b_A;
Quot_beta_erg_e    =        beta_erg_e_b_AP / beta_erg_e_b_A;
Quot_beta_vep_e    =        beta_vep_e_b_AP / beta_vep_e_b_A;
Quot_beta_amp_e    =        beta_amp_e_b_AP / beta_amp_e_b_A;
Quot_beta_fli_e    =        beta_fli_e_b_AP / beta_fli_e_b_A;
Quot_beta_clv_e    =        beta_clv_e_b_AP / beta_clv_e_b_A;
Quot_beta_md_e     =        beta_md_e_b_AP / beta_md_e_b_A;
Quot_beta_dyn_e    =        beta_dyn_e_b_AP / beta_dyn_e_b_A;
Quot_cor_err_nrr_dyn =      the1_b_AP / the1_b_A;
Quot_cor_err_vep_amp =      the2_b_AP / the2_b_A;
Quot_cor_err_vep_md =       the3_AP / the3_A;
Quot_cor_err_amp_fli =      the4_AP / the4_A;
Quot_cor_err_fli_clv =      the5_b_AP / the5_b_A;
DROP
          beta_nrr_g_b_AP beta_nrr_g_b_A
          beta_erg_g_b_AP beta_erg_g_b_A
          beta_vep_g_b_AP beta_vep_g_b_A
          beta_amp_g_b_AP beta_amp_g_b_A
          beta_fli_g_b_AP beta_fli_g_b_A
          beta_clv_g_b_AP beta_clv_g_b_A
          beta_md_g_b_AP beta_md_g_b_A
          beta_dyn_g_b_AP beta_dyn_g_b_A
          beta_amp_k_b_AP beta_amp_k_b_A
          beta_fli_k_b_AP beta_fli_k_b_A
          beta_clv_k_b_AP beta_clv_k_b_A
          beta_md_k_b_AP beta_md_k_b_A
          beta_dyn_k_b_AP beta_dyn_k_b_A
          beta_nrr_e_b_AP beta_nrr_e_b_A
          beta_erg_e_b_AP beta_erg_e_b_A
          beta_vep_e_b_AP beta_vep_e_b_A
          beta_amp_e_b_AP beta_amp_e_b_A
          beta_fli_e_b_AP beta_fli_e_b_A
          beta_clv_e_b_AP beta_clv_e_b_A
          beta_md_e_b_AP beta_md_e_b_A
          beta_dyn_e_b_AP beta_dyn_e_b_A
          the1_b_AP the1_b_A
          the2_b_AP the2_b_A
```

```
         the3_AP the3_A
         the4_AP the4_A
         the5_b_AP the5_b_A;
RUN;
PROC PRINT DATA = BOOT_STR.QUOT_PARMS_ALL_RANUNI;
TITLE1 "Quotient aus SD der Pfadkoeff. bzw. einer Fehler-Korr. d. BS-Analysen ";
TITLE2 "Ziehen v. Augenpaaren und einz. Augen – 3000 samples n=237 bzw. 474";
RUN;
options notes source source2 errors=20;

*Winsorieren der Verteilungen für die Schätzungen der 5 Messfehlerkorrelationen in beiden Bootstrapverfahren;
/*Für die Augenziehung und die Augenpaarziehung getrennt:
- jede Theta 1–5-Datei einzeln sortieren und eine Sortierungsvariable bilden und abspeichern;
- alle Theta 1–5 wieder zusammenmergen;
- alle zusammen gleichermaßen abschneiden;
- die Outputdateinen bilden;
- diese umbenennennen so dass A- und AP-index
- mergen undQuotientenbildung*/
*1a.Ziehung von Augen
*teta1;
DATA BOOT_STR.PARMS_The1_b_ranuni_aug_B;
SET BOOT_STR.PARMS_ALL_ranuni_aug_B;
KEEP the1_b;
ITERATION=_N_;
RUN;
PROC SORT DATA=BOOT_STR.PARMS_The1_b_ranuni_aug_B;
BY The1_b;
RUN;
DATA BOOT_STR.PARMS_The1_b_ranuni_aug_B;
SET BOOT_STR.PARMS_The1_b_ranuni_aug_B;
Sortierung = _N_;
run;
*teta2;
DATA BOOT_STR.PARMS_The2_b_ranuni_aug_B;
SET BOOT_STR.PARMS_ALL_ranuni_aug_B;
KEEP the2_b;
ITERATION=_N_;
RUN;
PROC SORT DATA=BOOT_STR.PARMS_The2_b_ranuni_aug_B;
BY The2_b;
RUN;
*Anmerkung: Hier wären nur ca. die fünf größten und kleinsten Werte abzuschneiden;
DATA BOOT_STR.PARMS_The2_b_ranuni_aug_B;
SET BOOT_STR.PARMS_The2_b_ranuni_aug_B;
Sortierung = _N_;
run;
*teta3;
DATA BOOT_STR.PARMS_The3_ranuni_aug;
SET BOOT_STR.PARMS_ALL_ranuni_aug;
KEEP the3;
ITERATION=_N_;
RUN;
PROC SORT DATA=BOOT_STR.PARMS_The3_ranuni_aug;
BY The3;
RUN;
DATA BOOT_STR.PARMS_The3_ranuni_aug;
SET BOOT_STR.PARMS_The3_ranuni_aug;
Sortierung = _N_;
run;
*Anmerkung: Hier wären nur ca. die zehn größten und kleinsten Werte abzuschneiden;
*teta4;
DATA BOOT_STR.PARMS_The4_ranuni_aug;
SET BOOT_STR.PARMS_ALL_ranuni_aug;
KEEP the4;
ITERATION=_N_;
RUN;
PROC SORT DATA=BOOT_STR.PARMS_The4_ranuni_aug;
BY The4;
RUN;
DATA BOOT_STR.PARMS_The4_ranuni_aug;
SET BOOT_STR.PARMS_The4_ranuni_aug;
Sortierung = _N_;
run;
*Anmerkung: hier wären nur ca. die 20 größten und kleinsten Werte abzuschneiden;
*teta5;
DATA BOOT_STR.PARMS_The5_b_ranuni_aug_B;
SET BOOT_STR.PARMS_ALL_ranuni_aug_B;
KEEP the5_b;
ITERATION=_N_;
RUN;
PROC SORT DATA=BOOT_STR.PARMS_The5_b_ranuni_aug_B;
BY The5_b;
RUN;
DATA BOOT_STR.PARMS_The5_b_ranuni_aug_B;
```

```
SET BOOT_STR.PARMS_The5_b_ranuni_aug_B;
Sortierung = _N_;
run;
*Anmerkung: Hier wären nur ca. die 15 größten und kleinsten Werte abzuschneiden;
*Vorbild;
*2a;
*Hinweis: Die folgende neu benannte Datei endet mit "_B", obwohl nicht alle fünf dateien, die
zu letzterer zusammengemerged werden, mit "_B" enden;
DATA BOOT_STR.PARMS_Thetas_ranuni_aug_B;
MERGE BOOT_STR.PARMS_The1_b_ranuni_aug_B
BOOT_STR.PARMS_The2_b_ranuni_aug_B
BOOT_STR.PARMS_The3_ranuni_aug
BOOT_STR.PARMS_The4_ranuni_aug
BOOT_STR.PARMS_The5_b_ranuni_aug_B;
RUN;
* 1b. Ziehung von Augenpaaren
*teta1;
DATA BOOT_STR.PARMS_The1_b_ranuni_B;
SET BOOT_STR.PARMS_ALL_ranuni_B;
KEEP the1_b;
ITERATION=_N_;
RUN;
PROC SORT DATA=BOOT_STR.PARMS_The1_b_ranuni_B;
BY The1_b;
RUN;
DATA BOOT_STR.PARMS_The1_b_ranuni_B;
SET BOOT_STR.PARMS_The1_b_ranuni_B;
Sortierung = _N_;
run;
*Anmerkung: hier wären die 60 (!) größten und kleinsten Werte abzuschneiden;
*teta2;
DATA BOOT_STR.PARMS_The2_b_ranuni_B;
SET BOOT_STR.PARMS_ALL_ranuni_B;
KEEP the2_b;
ITERATION=_N_;
RUN;
PROC SORT DATA=BOOT_STR.PARMS_The2_b_ranuni_B;
BY The2_b;
RUN;
*Anmerkung: Hier wären nur ca. die 15 größten und kleinsten Werte abzuschneiden;
DATA BOOT_STR.PARMS_The2_b_ranuni_B;
SET BOOT_STR.PARMS_The2_b_ranuni_B;
Sortierung = _N_;
run;
*teta3;
DATA BOOT_STR.PARMS_The3_ranuni;
SET BOOT_STR.PARMS_ALL_ranuni;
KEEP the3;
ITERATION=_N_;
RUN;
PROC SORT DATA=BOOT_STR.PARMS_The3_ranuni;
BY The3;
RUN;
DATA BOOT_STR.PARMS_The3_ranuni;
SET BOOT_STR.PARMS_The3_ranuni;
Sortierung = _N_;
run;
*Anmerkung: Hier wären nur ca. die 25 größten und kleinsten Werte abzuschneiden;
*teta4;
DATA BOOT_STR.PARMS_The4_ranuni;
SET BOOT_STR.PARMS_ALL_ranuni;
KEEP the4;
ITERATION=_N_;
RUN;
PROC SORT DATA=BOOT_STR.PARMS_The4_ranuni;
BY The4;
RUN;
DATA BOOT_STR.PARMS_The4_ranuni;
SET BOOT_STR.PARMS_The4_ranuni;
Sortierung = _N_;
run;
*Anmerkung: Hier wären ca. die 50 größten und kleinsten Werte abzuschneiden;
*teta5;
DATA BOOT_STR.PARMS_The5_b_ranuni_B;
SET BOOT_STR.PARMS_ALL_ranuni_B;
KEEP the5_b;
ITERATION=_N_;
RUN;
PROC SORT DATA=BOOT_STR.PARMS_The5_b_ranuni_B;
BY The5_b;
RUN;
DATA BOOT_STR.PARMS_The5_b_ranuni_B;
SET BOOT_STR.PARMS_The5_b_ranuni_B;
Sortierung = _N_;
```

```
run;
*Anmerkung: Hier wären nur ca. die 45 größten und kleinsten Werte abzuschneiden;
*Vorbild;
*2b.;
*Hinweis: Die folgende neu benannte Datei endet mit "_B", obwohl nicht alle fünf dateien, die
*zu letzterer zusammengemerged werden, mit "_B" enden;
DATA BOOT_STR.PARMS_Thetas_ranuni_B;
MERGE BOOT_STR.PARMS_The1_b_ranuni_B
BOOT_STR.PARMS_The2_b_ranuni_B
BOOT_STR.PARMS_The3_ranuni
BOOT_STR.PARMS_The4_ranuni
BOOT_STR.PARMS_The5_b_ranuni_B;
RUN;
*3a;
DATA BOOT_STR.PARMS_Thetas_ranuni_aug_B_wins ;
SET BOOT_STR.PARMS_Thetas_ranuni_aug_B;
Sortierung = _N_;
IF (sortierung le 60 or sortierung > 2940) THEN delete;
DROP sortierung;
run;
*3b;
DATA BOOT_STR.PARMS_Thetas_ranuni_B_wins ;
SET BOOT_STR.PARMS_Thetas_ranuni_B;
Sortierung = _N_;
IF (sortierung le 60 or sortierung > 2940) THEN delete;
DROP sortierung;
run;
*4a;
*Hinweis: Da der Name der OUTPUT-Datei zu lang wäre, wird der Bestandteil PARMS
*einfach wegggelassen.;
PROC MEANS DATA=BOOT_STR.PARMS_Thetas_ranuni_aug_B_wins ;
OUTPUT OUT=BOOT_STR.OUT_Thetas_ranuni_aug_B_wins;
TITLE1 "Bootstrapanalyse Patientendaten - Ziehen v. einzelnen Augen";
TITLE2 "nach Winsorierung 2880 samples n=474";
RUN;
PROC GCHART DATA=BOOT_STR.PARMS_Thetas_ranuni_aug_B_wins ;
VBAR the1_b
          the2_b
          the3
          the4
          the5_b;
TITLE1 "Bootstrapanalyse Patientendaten - Ziehen v. einzelnen Augen";
TITLE2 "nach Winsorierung 2880 samples n=474";
RUN;
PROC UNIVARIATE DATA=BOOT_STR.PARMS_Thetas_ranuni_aug_B_wins ;
VAR the1_b
          the2_b
          the3
          the4
          the5_b;
TITLE1 "Bootstrapanalyse Patientendaten - Ziehen v. einzelnen Augen";
TITLE2 "nach Winsorierung 2880 samples n=474";
RUN;
*4b;
PROC MEANS DATA=BOOT_STR.PARMS_Thetas_ranuni_B_wins ;
OUTPUT OUT=BOOT_STR.OUT_PARMS_Thetas_ranuni_B_wins;
TITLE1 "Bootstrapanalyse Patientendaten - Ziehen v. Augenpaaren";
TITLE2 "nach Winsorierung 2880 samples n=474";
RUN;

PROC GCHART DATA=BOOT_STR.PARMS_Thetas_ranuni_B_wins ;
VBAR the1_b
          the2_b
          the3
          the4
          the5_b;
TITLE1 "Bootstrapanalyse Patientendaten - Ziehen v. Augenpaaren";
TITLE2 "nach Winsorierung 2880 samples n=237";
RUN;
PROC UNIVARIATE DATA=BOOT_STR.PARMS_Thetas_ranuni_B_wins ;
VAR the1_b
          the2_b
          the3
          the4
          the5_b;
TITLE1 "Bootstrapanalyse Patientendaten - Ziehen v. Augenpaaren";
TITLE2 "nach Winsorierung 2880 samples n=237 - Fehlerkorrelationen";
RUN;
*5a;
*Zur Vorbereitung der SE-Auswertung;
DATA BOOT_STR.OUT_B_Thetas_ranuni_aug_B_wins;
SET BOOT_STR.OUT_Thetas_ranuni_aug_B_wins;
IF _STAT_ NE "STD" THEN DELETE;
DROP _TYPE_;
```

```
RUN;
DATA BOOT_STR.OUT_B2_Thetas_ranuni_aug_B_wins;
SET BOOT_STR.OUT_B_Thetas_ranuni_aug_B_wins;
                    the1_b_A= the1_b;
                    the2_b_A= the2_b;
                    the3_A= the3;
                    the4_A= the4;
                    the5_b_A= the5_b;
DROP
                    the1_b
                    the2_b
                    the3
                    the4
                    the5_b;
RUN;
*5b;
*Zur Vorbereitung der SE-Auswertung;
DATA BOOT_STR.OUT_B_PARMS_Thetas_ranuni_B_wins;
SET BOOT_STR.OUT_PARMS_Thetas_ranuni_B_wins;
IF _STAT_ NE "STD" THEN DELETE;
DROP _TYPE_;
RUN;
*Hinweis: da der Name der OUTPUT-datei zu lang wäre, wird der Bestandteil PARMS
*einfach um einen Buchstaben auf PARM gekürzt;
DATA BOOT_STR.OUT_B2_PARM_Thetas_ranuni_B_wins;
SET BOOT_STR.OUT_B_PARMS_Thetas_ranuni_B_wins;
                    the1_b_AP= the1_b;
                    the2_b_AP= the2_b;
                    the3_AP= the3;
                    the4_AP= the4;
                    the5_b_AP= the5_b;
DROP
                    the1_b
                    the2_b
                    the3
                    the4
                    the5_b;
RUN;
******
6;
DATA BOOT_STR.OUT_Thetas_RANUNI_AUG_u_PAAR_win;
MERGE BOOT_STR.OUT_B2_Thetas_ranuni_aug_B_wins BOOT_STR.OUT_B2_PARM_Thetas_ranuni_B_wins;
RUN;
PROC PRINT DATA = BOOT_STR.OUT_Thetas_RANUNI_AUG_u_PAAR_win;
TITLE1 "SD der Fehlerkorr. d. winsorierten BS-Analysen";
TITLE2 "Ziehen v. Augenpaaren und einz. Augen - 2880 samples n=237 bzw. 474";
RUN;
DATA BOOT_STR.QUOT_Thetas_RANUNI_wins;
SET BOOT_STR.OUT_Thetas_RANUNI_AUG_u_PAAR_win;
Quot_cor_err_nrr_dyn=       the1_b_AP / the1_b_A;
Quot_cor_err_vep_amp=       the2_b_AP / the2_b_A;
Quot_cor_err_vep_md=        the3_AP / the3_A;
Quot_cor_err_amp_fli=       the4_AP / the4_A;
Quot_cor_err_fli_clv=       the5_b_AP / the5_b_A;
DROP
         the1_b_AP the1_b_A
         the2_b_AP the2_b_A
         the3_AP the3_A
         the4_AP the4_A
         the5_b_AP the5_b_A;
RUN;
PROC PRINT DATA = BOOT_STR.QUOT_Thetas_RANUNI_wins;
TITLE1 "Quotient aus SD der Fehler-Korr. d. winsorierten BS-Analysen ";
TITLE2 "Ziehen v. Augenpaaren und einz. Augen - 2880 samples n=237 bzw. 474";
RUN;
********
*Berechnung der Vedrteilung der Chi-Quadrat-Statistik und des 95%-KI für selbige
*a)Auf Basis des nichtadjustierten Bootstrap
*******************************************
*5.;
*Von den (3000) OUTRAM-sas-data-sets wird nur jeweils die interessierende Zeile (Beobachtung)
*mit dem _NAME_ "CHISQ" (=overall CHI SQUARE) bewahrt (und der Rest gelöscht);

%Macro chisq_isol1 (dateinam=, lib=);
DATA &lib..CHISQ_sample_&dateinam;
SET &lib..outra_sample_&dateinam;
IF _NAME_ NE "CHISQUAR" THEN DELETE;
KEEP _ESTIM_;
RUN;
%MEND chisq_isol1;
%macro create_10(howmany);
   %do i=1 %to &howmany;
      %chisq_isol1 (dateinam=ranuni_aug&i, lib=boot_str);
       run;
```

```
   %end;
%mend create_10;
%create_10(3000)

*6.;
*die (3000) einzeiligen Datensätze addieren (unter Verwendung eines „iterativen Do-Makros“);
%macro create_11(howmany);
   %do i=2 %to &howmany;
   proc APPEND BASE= BOOT_STR.CHISQ_SAMPLE_RANUNI_aug1
            DATA= BOOT_STR.CHISQ_SAMPLE_RANUNI_aug&i;
      run;
   %end;
%mend create_11;
%create_11(3000)
*7.;
*Umbenennen, Ausdrucken, Labeln der Datei;
DATA BOOT_STR.CHISQ_ALL_RANUNI_aug;
SET BOOT_STR.CHISQ_SAMPLE_RANUNI_aug1;
LABEL _ESTIM_='overall CHI SQUARE';
RUN;
PROC SORT DATA=BOOT_STR.CHISQ_ALL_RANUNI_aug;
BY _ESTIM_;
RUN;
DATA BOOT_STR.CHISQ_ALL_s_RANUNI_aug;
SET BOOT_STR.CHISQ_ALL_RANUNI_aug;
Sortierung = _N_;
RUN;
PROC MEANS DATA=BOOT_STR.CHISQ_ALL_s_RANUNI_aug;
TITLE1 "Bootstrapanalyse Patientendaten - Ziehen einz. Augen";
TITLE2 "3000 samples; n=474 ; Verteilung des Chi-Quadrat";
RUN;
PROC UNIVARIATE DATA=BOOT_STR.CHISQ_ALL_s_RANUNI_aug;
VAR _ESTIM_;
TITLE1 "Bootstrapanalyse Patientendaten - Ziehen v. einzelnen Augen";
TITLE2 "3000 samples n=474 - Verteilung des Chi-Quadrat";
RUN;
PROC GCHART DATA=BOOT_STR.CHISQ_ALL_s_RANUNI_aug;
VBAR _ESTIM_;
TITLE1 "Bootstrapanalyse Patientendaten - Ziehen einz. Augen";
TITLE2 "3000 samples; n=474";
RUN;
*Berechnung des 95%-KI durch Winsorierung;
*li KI-Grenze:*0.025 *3000 = 75
*re KI-Grenze:*0.975 *3000 = 2925;
DATA BOOT_STR.CHISQ_95_PERC_RANUNI_aug;
SET BOOT_STR.CHISQ_ALL_s_RANUNI_aug;
IF (Sortierung LT 75 or Sortierung GT 2925) THEN delete;
RUN;
PROC PRINT DATA=BOOT_STR.CHISQ_95_PERC_RANUNI_aug;
RUN;
*Minimum bzw. Maximum bestimmen die Vertrauensgrenzen;
PROC MEANS DATA=BOOT_STR.CHISQ_95_PERC_RANUNI_aug;
TITLE1 "Bootstrapanalyse Patientendaten - Ziehen einz. Augen";
TITLE2 "3000 samples; n=474";
RUN;
*Resultat: das KI geht von 6.15 bis 35.13
**********

b)*Auf BASIS des ADJUSTIERTEN BOOTSTRAP
************************************
*1.;
*Von den (3000) OUTRAM-sas-data-sets wird nur jeweils die interessierende Zeile (Beobachtung)
*mit dem _NAME_ "CHISQ" (=overall CHI SQUARE) bewahrt (und der Rest gelöscht);

%Macro chisq_isol1 (dateinam=, lib=);
DATA &lib..CHISQ_sample_&dateinam;
SET &lib..outra_sample_&dateinam;
IF _NAME_ NE "CHISQUAR" THEN DELETE;
KEEP _ESTIM_;
RUN;
%MEND chisq_isol1;
%macro create_10(howmany);
   %do i=1 %to &howmany;
      %chisq_isol1 (dateinam=ranuni&i, lib=boot_str);
       run;
   %end;
%mend create_10;
%create_10(3000)

*2.;
*die (3000) einzeiligen Datensätze addieren (unter Verwendung eines „iterativen Do-Makros“);
%macro create_11(howmany);
   %do i=2 %to &howmany;
   proc APPEND BASE= BOOT_STR.CHISQ_SAMPLE_RANUNI1
```

```
            DATA= BOOT_STR.CHISQ_SAMPLE_RANUNI&i;
      run;
   %end;
%mend create_11;
%create_11(3000)
*3;
*Umbenennen, Ausdrucken, Labeln der Datei;
DATA BOOT_STR.CHISQ_ALL_RANUNI;
SET BOOT_STR.CHISQ_SAMPLE_RANUNI1;
LABEL _ESTIM_='overall CHI SQUARE';
RUN;
PROC SORT DATA=BOOT_STR.CHISQ_ALL_RANUNI;
BY _ESTIM_;
RUN;
DATA BOOT_STR.CHISQ_ALL_s_RANUNI;
SET BOOT_STR.CHISQ_ALL_RANUNI;
Sortierung = _N_;
RUN;
PROC PRINT DATA=BOOT_STR.CHISQ_ALL_s_RANUNI;
RUN;
*4. Maßzahlen;
PROC MEANS DATA=BOOT_STR.CHISQ_ALL_s_RANUNI;
TITLE1 "Bootstrapanalyse Patientendaten - Ziehen v. Augenpaaren)";
TITLE2 "3000 samples n=237";
RUN;
PROC GCHART DATA=BOOT_STR.CHISQ_ALL_s_RANUNI;
VBAR _ESTIM_;
TITLE1 "Bootstrapanalyse Patientendaten - Ziehen v. Augenpaaren)";
TITLE2 "3000 samples n=237";
RUN;
*Berechnung des 95%-KI durch Winsorierung;
*li KI-Grenze:*0.025 *3000 = 75
*re KI-Grenze:*0.975 *3000 = 2925;
DATA BOOT_STR.CHISQ_95_PERC_RANUNI;
SET BOOT_STR.CHISQ_ALL_s_RANUNI;
IF (Sortierung LT 75 or Sortierung GT 2925) THEN delete;
RUN;
PROC PRINT DATA=BOOT_STR.CHISQ_95_PERC_RANUNI;
RUN;
*Minimum bzw. Maximum bestimmen die Vertrauensgrenzen;
PROC MEANS DATA=BOOT_STR.CHISQ_95_PERC_RANUNI;
TITLE1 "Bootstrapanalyse Patientendaten - Ziehen v. Augenpaaren)";
TITLE2 "30 samples n=237";
RUN;
*Resultat: das KI geht von 6.68 bis 42.34
*************************************************************************
```

11.6.2.2 Output zur adj. und nichtadj. Bootstrapanalyse und Vergleich beider Verfahren

11.6.2.2.1 empirische Standardfehler der Pfadkoeff. der adj. Analyse

Bootstrapanalyse Patientendaten - Ziehen v. Augenpaaren

3000 samples n=237

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum
RHS		3000	0.0439492	0.0195834	0.0036773
ITER	Iteration	0	.	.	.
The4	Corr Err_amp Err_fli	3000	0.0798259	0.3531750	-4.8926859
The3	Corr Err_vep Err_md	3000	0.0633779	0.3868926	-5.5901334
beta_nrr_g_b		3000	0.6301255	0.0542559	0.4320223
beta_erg_g_b		3000	0.5476069	0.0444626	0.3924202
beta_vep_g_b		3000	0.4919926	0.0574835	0.2845598
beta_amp_g_b		3000	0.3336102	0.0733739	0.0678917
beta_fli_g_b		3000	0.5869666	0.0649891	0.3776325
beta_clv_g_b		3000	0.7637655	0.0553766	0.5553261
beta_md_g_b		3000	0.8143818	0.0646534	0.5664406
beta_dyn_g_b		3000	0.6654344	0.0632214	0.3982785
beta_amp_k_b		3000	0.2148474	0.1201493	0.000181913
beta_fli_k_b		3000	0.3519505	0.1081323	0.0018216
beta_clv_k_b		3000	0.4023841	0.1199450	0.0044773
beta_md_k_b		3000	0.4422309	0.1273382	0.0040619
beta_dyn_k_b		3000	0.4372316	0.1183018	0.000196903
beta_nrr_e_b		3000	0.7733322	0.0442458	0.5898813
beta_erg_e_b		3000	0.8350434	0.0292132	0.7197108
beta_vep_e_b		3000	0.8679995	0.0323813	0.7483087
beta_amp_e_b		3000	0.9051659	0.0659400	0.0737825
beta_fli_e_b		3000	0.7142537	0.0688510	0.0297940
beta_clv_e_b		3000	0.4854051	0.0410730	0.0375181
beta_md_e_b		3000	0.3447681	0.0471833	0.000631650
beta_dyn_e_b		3000	0.5832370	0.0853125	0.0211197
the1_b	Betr. v. Fehlerkorr nrr dyn	3000	0.2422195	0.4025560	0.0012081
the2_b	Betr. v. Fehlerkorr vep amp	3000	0.2108981	0.1315113	0.000660791
the5_b	Betr. v. Fehlerkorr fli clv	3000	0.3452474	0.3554017	0.0741360

Variable	Etikett	Maximum
RHS		0.1374031
ITER	Iteration	.
The4	Corr Err_amp Err_fli	5.4654217
The3	Corr Err_vep Err_md	11.6820861
beta_nrr_g_b		0.8074900
beta_erg_g_b		0.6942740
beta_vep_g_b		0.6638373
beta_amp_g_b		0.6426784
beta_fli_g_b		0.8240058
beta_clv_g_b		0.9189457
beta_md_g_b		0.9908783
beta_dyn_g_b		0.8806814
beta_amp_k_b		0.9936078
beta_fli_k_b		0.8608164
beta_clv_k_b		0.7018916
beta_md_k_b		0.7499756
beta_dyn_k_b		0.8128849

3000 samples n=237

Die Prozedur MEANS

Variable	Etikett	Maximum
beta_nrr_e_b		0.9018629
beta_erg_e_b		0.9197861
beta_vep_e_b		0.9586854
beta_amp_e_b		0.9740921
beta_fli_e_b		0.8412158
beta_clv_e_b		0.5837040
beta_md_e_b		0.4618838
beta_dyn_e_b		0.7200847
the1_b	Betr. v. Fehlerkorr nrr dyn	5.7543842
the2_b	Betr. v. Fehlerkorr vep amp	4.0533128
the5_b	Betr. v. Fehlerkorr fli clv	8.7774170

11.6.2.2.2 empirische Standardfehler der Pfadkoeff. der nicht adj. Analyse

Bootstrapanalyse Patientendaten - Ziehen v. einzelnen Augen
3000 samples n=474
Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum
RHS		3000	0.0374620	0.0158812	0.0042303
ITER	Iteration	0	.	.	.
The4	Corr Err_amp Err_fli	3000	0.0830577	0.2476719	-8.3110187
The3	Corr Err_vep Err_md	3000	0.0527833	0.2265815	-3.2364501
beta_nrr_g_b		3000	0.6334549	0.0479848	0.3795249
beta_erg_g_b		3000	0.5484688	0.0375851	0.3864874
beta_vep_g_b		3000	0.4934134	0.0486702	0.3379535
beta_amp_g_b		3000	0.3318520	0.0569295	0.1254431
beta_fli_g_b		3000	0.5833757	0.0511357	0.3856382
beta_clv_g_b		3000	0.7628482	0.0467384	0.6122073
beta_md_g_b		3000	0.8123586	0.0518518	0.6318389
beta_dyn_g_b		3000	0.6602652	0.0512651	0.4730572
beta_amp_k_b		3000	0.2104929	0.0916754	0.000056941
beta_fli_k_b		3000	0.3605883	0.0818954	0.0038049
beta_clv_k_b		3000	0.4104932	0.0959190	0.0015970
beta_md_k_b		3000	0.4510279	0.1014189	0.0054344
beta_dyn_k_b		3000	0.4452332	0.0934656	0.0022456
beta_nrr_e_b		3000	0.7712869	0.0393753	0.6228587
beta_erg_e_b		3000	0.8349620	0.0246672	0.7317271
beta_vep_e_b		3000	0.8678809	0.0277827	0.7453143
beta_amp_e_b		3000	0.9126734	0.0434029	0.0716441
beta_fli_e_b		3000	0.7191083	0.0481720	0.0363596
beta_clv_e_b		3000	0.4871119	0.0299902	0.0070586
beta_md_e_b		3000	0.3503147	0.0353441	0.0014946
beta_dyn_e_b		3000	0.5919098	0.0591897	0.0248989
the1_b	Betr. v. Fehlerkorr nrr dyn	3000	0.2160446	0.2598212	0.000520494
the2_b	Betr. v. Fehlerkorr vep amp	3000	0.2065980	0.0753520	0.0267752
the5_b	Betr. v. Fehlerkorr fli clv	3000	0.3310936	0.2436320	0.1164457

Variable	Etikett	Maximum
RHS		0.1208423
ITER	Iteration	.
The4	Corr Err_amp Err_fli	3.9459805
The3	Corr Err_vep Err_md	8.1133929
beta_nrr_g_b		0.7823344
beta_erg_g_b		0.6815977
beta_vep_g_b		0.6665850
beta_amp_g_b		0.5596934
beta_fli_g_b		0.8065662
beta_clv_g_b		0.9107302
beta_md_g_b		0.9661713
beta_dyn_g_b		0.8563431
beta_amp_k_b		0.9545373
beta_fli_k_b		0.8042996
beta_clv_k_b		0.6403799
beta_md_k_b		0.6797351
beta_dyn_k_b		0.8052914

Bootstrapanalyse Patientendaten - Ziehen v. einzelnen Augen
3000 samples n=474
Die Prozedur MEANS

Variable	Etikett	Maximum
beta_nrr_e_b		0.9251815
beta_erg_e_b		0.9222947
beta_vep_e_b		0.9410961
beta_amp_e_b		0.9769631
beta_fli_e_b		0.8213136
beta_clv_e_b		0.5723454
beta_md_e_b		0.4451387
beta_dyn_e_b		0.7136765
the1_b	Betr. v. Fehlerkorr nrr dyn	6.5484930
the2_b	Betr. v. Fehlerkorr vep amp	2.6884315
the5_b	Betr. v. Fehlerkorr fli clv	7.7063156

SD der Pfadkoeff. bzw. einer Fehlerkorr. d. BS-Analysen
Ziehen v. Augenpaaren und einz. Augen - 3000 samples n=237 bzw. 474

Beob.	_FREQ_	_STAT_	beta_ nrr_g_ b_A	beta_ erg_g_ b_A	beta_ vep_g_ b_A	beta_ amp_g_ b_A	beta_ fli_g_ b_A	beta_ clv_g_ b_A	beta_md_ g_b_A	beta_ dyn_g_ b_A
1	3000	STD	0.047985	0.037585	0.048670	0.056930	0.051136	0.046738	0.051852	0.051265

Beob.	beta_ amp_k_ b_A	beta_ fli_k_ b_A	beta_ clv_k_ b_A	beta_md_ k_b_A	beta_ dyn_k_ b_A	beta_ nrr_e_ b_A	beta_ erg_e_ b_A	beta_ vep_e_ b_A	beta_ amp_e_ b_A
1	0.091675	0.081895	0.095919	0.10142	0.093466	0.039375	0.024667	0.027783	0.043403

Beob.	beta_ fli_e_ b_A	beta_ clv_e_ b_A	beta_md_ e_b_A	beta_ dyn_e_ b_A	the1_b_A	the2_b_A	the3_A	the4_A	the5_b_A
1	0.048172	0.029990	0.035344	0.059190	0.25982	0.075352	0.22658	0.24767	0.24363

Beob.	beta_ nrr_g_ b_AP	beta_ erg_g_ b_AP	beta_ vep_g_ b_AP	beta_ amp_g_ b_AP	beta_ fli_g_ b_AP	beta_ clv_g_ b_AP	beta_md_ g_b_AP	beta_ dyn_g_ b_AP	beta_ amp_k_ b_AP
1	0.054256	0.044463	0.057484	0.073374	0.064989	0.055377	0.064653	0.063221	0.12015

Beob.	beta_ fli_k_ b_AP	beta_ clv_k_ b_AP	beta_md_ k_b_AP	beta_ dyn_k_ b_AP	beta_ nrr_e_ b_AP	beta_ erg_e_ b_AP	beta_ vep_e_ b_AP	beta_ amp_e_ b_AP	beta_ fli_e_ b_AP
1	0.10813	0.11995	0.12734	0.11830	0.044246	0.029213	0.032381	0.065940	0.068851

Beob.	beta_ clv_e_ b_AP	beta_md_ e_b_AP	beta_ dyn_e_ b_AP	the1_ b_AP	the2_ b_AP	the3_AP	the4_AP	the5_ b_AP
1	0.041073	0.047183	0.085312	0.40256	0.13151	0.38689	0.35317	0.35540

11.6.2.2.3 Quotienten aus den empirischen SE der Pfadkoeff. beider Verfahren

Quotient aus SD der Pfadkoeff. bzw. einer Fehler-Korr. d. BS-Analysen
Ziehen v. Augenpaaren und einz. Augen - 3000 samples n=237 bzw. 474

Beob.	_FREQ_	_STAT_	Quot_ beta_ nrr_g	Quot_ beta_ erg_g	Quot_ beta_ vep_g	Quot_ beta_ amp_g	Quot_ beta_ fli_g	Quot_ beta_ clv_g	Quot_ beta_ md_g	Quot_ eta_dyn_ g
1	3000	STD	1.13069	1.18299	1.18108	1.28885	1.27091	1.18482	1.24689	1.23323

Beob.	Quot_ beta_ amp_k	Quot_ beta_ fli_k	Quot_ beta_ clv_k	Quot_ beta_ md_k	Quot_ beta_ dyn_k	Quot_ beta_ nrr_e	Quot_ beta_ erg_e	Quot_ beta_ vep_e	Quot_ beta_ amp_e	Quot_ beta_ fli_e
1	1.31060	1.32037	1.25048	1.25557	1.26573	1.12369	1.18429	1.16552	1.51925	1.42928

Beob.	Quot_ beta_ clv_e	Quot_ beta_ md_e	Quot_ beta_ dyn_e	Quot_cor_ err_nrr_ dyn	Quot_cor_ err_vep_ amp	Quot_cor_ err_vep_md	Quot_cor_ err_amp_ fli	Quot_cor_ err_fli_ clv
1	1.36955	1.33497	1.44134	1.54936	1.74529	1.70752	1.42598	1.45876

11.6.2.2.4 Empirische SE der Fehlerkorrelationen der winsorierten adj. und nicht adj. Analysen

Bootstrapanalyse Patientendaten - Ziehen v. Augenpaaren
nach Winsorierung 2880 samples n=474

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum
the1_b	Betr. v. Fehlerkorr nrr dyn	2880	0.1961628	0.0743229	0.0431497
the2_b	Betr. v. Fehlerkorr vep amp	2880	0.2047416	0.0579189	0.0705051
The3	Corr Err_vep Err_md	2880	0.0468372	0.0802085	-0.1442192
The4	Corr Err_amp Err_fli	2880	0.0876318	0.0624967	-0.0736075
the5_b	Betr. v. Fehlerkorr fli clv	2880	0.3147133	0.0729558	0.1597591

Variable	Etikett	Maximum
the1_b	Betr. v. Fehlerkorr nrr dyn	0.5343875
the2_b	Betr. v. Fehlerkorr vep amp	0.3454042
The3	Corr Err_vep Err_md	0.2569062
The4	Corr Err_amp Err_fli	0.2358648
the5_b	Betr. v. Fehlerkorr fli clv	0.5251527

Bootstrapanalyse Patientendaten - Ziehen v. einzelnen Augen
nach Winsorierung 2880 samples n=474

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum
the1_b	Betr. v. Fehlerkorr nrr dyn	2880	0.1968826	0.0614983	0.0616048
the2_b	Betr. v. Fehlerkorr vep amp	2880	0.2047535	0.0462231	0.0971547
The3	Corr Err_vep Err_md	2880	0.0455524	0.0698502	-0.1198205
The4	Corr Err_amp Err_fli	2880	0.0875532	0.0469262	-0.0261081
the5_b	Betr. v. Fehlerkorr fli clv	2880	0.3171670	0.0596536	0.1903539

Variable	Etikett	Maximum
the1_b	Betr. v. Fehlerkorr nrr dyn	0.3639634
the2_b	Betr. v. Fehlerkorr vep amp	0.3128790
The3	Corr Err_vep Err_md	0.2149054
The4	Corr Err_amp Err_fli	0.1947942
the5_b	Betr. v. Fehlerkorr fli clv	0.4687250

11.6.2.2.5 Empirische SE der Fehlerkorr. d. winsorierten Analysen beider Verfahren: Quotienten

SD der Fehlerkorr. d. winsorierten BS-Analysen
Ziehen v. Augenpaaren und einz. Augen - 2880 samples n=237 bzw. 474

Beob.	_FREQ_	_STAT_	the1_b_A	the2_b_A	the3_A	the4_A	the5_b_A
1	2880	STD	0.061498	0.046223	0.069850	0.046926	0.059654

Beob.	the1_ b_AP	the2_ b_AP	the3_AP	the4_AP	the5_ b_AP
1	0.074323	0.057919	0.080208	0.062497	0.072956

Quotient aus SD der Fehler-Korr. d. winsorierten BS-Analysen
Ziehen v. Augenpaaren und einz. Augen - 2880 samples n=237 bzw. 474

Beob.	_FREQ_	_STAT_	Quot_cor_ err_nrr_ dyn	Quot_cor_ err_vep_ amp	Quot_cor_ err_vep_md	Quot_cor_ err_amp_ fli	Quot_cor_ err_fli_ clv
1	2880	STD	1.20854	1.25303	1.14829	1.33181	1.22299

11.6.2.2.6 Pfadkoeff. des Faktors Konzentration: Median und weitere stat. Kennwerte der adj. Analyse

Berechnung des *Medians* dieser Pfadkoeffizienten, da bei den dem Faktor Konzentration zugehörigen Pfadkoeffizienten häufig Abweichungen zwischen arithmetischen Mittel und Median bestehen

Bootstrapanalyse Patientendaten - Ziehen v. Augenpaaren
3000 samples n=237 –

Die Prozedur UNIVARIATE
Variable: beta_amp_k_b

Momente

N	3000	Summe Gewichte	3000
Mittelwert	0.21484739	Summe Beobacht.	644.54216
Std.abweichung	0.12014933	Varianz	0.01443586
Schiefe	1.55419027	Kurtosis	6.6012318
Unkorr. Qu.summe	181.771345	Korr. Quad.summe	43.2931465
Variationskoeff.	55.9231038	Stdfeh. Mittelw.	0.00219362

Grundlegende Statistikmaße

Lage		Streuung	
Mittelwert	0.214847	Std.abweichung	0.12015
Median	0.207217	Varianz	0.01444
Modalwert	.	Spannweite	0.99343
		Interquartilsabstand	0.14127

Die Prozedur UNIVARIATE
Variable: beta_fli_k_b

Momente

N	3000	Summe Gewichte	3000
Mittelwert	0.35195051	Summe Beobacht.	1055.85152
Std.abweichung	0.1081323	Varianz	0.01169259
Schiefe	-0.280494	Kurtosis	1.4741758
Unkorr. Qu.summe	406.673566	Korr. Quad.summe	35.066088
Variationskoeff.	30.7237223	Stdfeh. Mittelw.	0.00197422

Grundlegende Statistikmaße

Lage		Streuung	
Mittelwert	0.351951	Std.abweichung	0.10813
Median	0.359561	Varianz	0.01169
Modalwert	.	Spannweite	0.85899
		Interquartilsabstand	0.12758

Bootstrapanalyse Patientendaten - Ziehen v. Augenpaaren
3000 samples n=237 - Pfadkoeff. des Faktors Konzentration

Die Prozedur UNIVARIATE
Variable: beta_clv_k_b

Momente

N	3000	Summe Gewichte	3000
Mittelwert	0.40238414	Summe Beobacht.	1207.15241
Std.abweichung	0.11994501	Varianz	0.01438681
Schiefe	-0.9352493	Kurtosis	0.92111873
Unkorr. Qu.summe	528.885009	Korr. Quad.summe	43.1460284
Variationskoeff.	29.8085828	Stdfeh. Mittelw.	0.00218989

Grundlegende Statistikmaße

Lage		Streuung	
Mittelwert	0.402384	Std.abweichung	0.11995
Median	0.421281	Varianz	0.01439
Modalwert	.	Spannweite	0.69741
		Interquartilsabstand	0.14116

Die Prozedur UNIVARIATE
Variable: beta_md_k_b

Momente

N	3000	Summe Gewichte	3000
Mittelwert	0.4422309	Summe Beobacht.	1326.69271
Std.abweichung	0.12733823	Varianz	0.01621503
Schiefe	-0.9804038	Kurtosis	1.19959248
Unkorr. Qu.summe	635.333376	Korr. Quad.summe	48.6288631
Variationskoeff.	28.794513	Stdfeh. Mittelw.	0.00232487

Grundlegende Statistikmaße

Lage		Streuung	
Mittelwert	0.442231	Std.abweichung	0.12734
Median	0.459711	Varianz	0.01622
Modalwert	.	Spannweite	0.74591
		Interquartilsabstand	0.14584

Bootstrapanalyse Patientendaten - Ziehen v. Augenpaaren
3000 samples n=237 - Pfadkoeff. des Faktors Konzentration

Die Prozedur UNIVARIATE
Variable: beta_dyn_k_b

Momente

N	3000	Summe Gewichte	3000
Mittelwert	0.43723164	Summe Beobacht.	1311.69493
Std.abweichung	0.11830182	Varianz	0.01399532
Schiefe	-0.955253	Kurtosis	2.43943099
Unkorr. Qu.summe	615.486497	Korr. Quad.summe	41.9719644
Variationskoeff.	27.0570117	Stdfeh. Mittelw.	0.00215989

Grundlegende Statistikmaße

Lage		Streuung	
Mittelwert	0.437232	Std.abweichung	0.11830
Median	0.450123	Varianz	0.01400
Modalwert	.	Spannweite	0.81269
		Interquartilsabstand	0.11863

11.6.2.2.7 Pfadkoeff. des Faktors Konzentration: dito, nichtadjustierte Analyse

Bootstrapanalyse Patientendaten - Ziehen v. einzelnen Augen

3000 samples n=474

Die Prozedur UNIVARIATE
Variable: beta_amp_k_b

Momente

N	3000	Summe Gewichte	3000
Mittelwert	0.21049295	Summe Beobacht.	631.478839
Std.abweichung	0.09167535	Varianz	0.00840437
Schiefe	1.45837809	Kurtosis	9.37059907
Unkorr. Qu.summe	158.126549	Korr. Quad.summe	25.2047072
Variationskoeff.	43.5526964	Stdfeh. Mittelw.	0.00167376

Grundlegende Statistikmaße

Lage		Streuung	
Mittelwert	0.210493	Std.abweichung	0.09168
Median	0.208770	Varianz	0.00840
Modalwert	.	Spannweite	0.95448
		Interquartilsabstand	0.10824

Die Prozedur UNIVARIATE

Variable: beta_fli_k_b

Momente

N	3000	Summe Gewichte	3000
Mittelwert	0.36058828	Summe Beobacht.	1081.76483
Std.abweichung	0.08189541	Varianz	0.00670686
Schiefe	-0.3468369	Kurtosis	1.97578295
Unkorr. Qu.summe	410.185587	Korr. Quad.summe	20.1138689
Variationskoeff.	22.7116125	Stdfeh. Mittelw.	0.0014952

Grundlegende Statistikmaße

Lage		Streuung	
Mittelwert	0.360588	Std.abweichung	0.08190
Median	0.365877	Varianz	0.00671
Modalwert	.	Spannweite	0.80049
		Interquartilsabstand	0.09837

Die Prozedur UNIVARIATE

Variable: beta_clv_k_b

Momente

N	3000	Summe Gewichte	3000
Mittelwert	0.41049324	Summe Beobacht.	1231.47971
Std.abweichung	0.09591895	Varianz	0.00920045
Schiefe	-0.9838568	Kurtosis	1.62501458
Unkorr. Qu.summe	533.106229	Korr. Quad.summe	27.5921352
Variationskoeff.	23.3667556	Stdfeh. Mittelw.	0.00175123

Grundlegende Statistikmaße

Lage		Streuung	
Mittelwert	0.410493	Std.abweichung	0.09592
Median	0.423878	Varianz	0.00920
Modalwert	.	Spannweite	0.63878
		Interquartilsabstand	0.11352

Bootstrapanalyse Patientendaten - Ziehen v. einzelnen Augen
3000 samples n=474 - Pfadkoeff. des Faktors Konzentration

Die Prozedur UNIVARIATE
Variable: beta_md_k_b

Momente

N	3000	Summe Gewichte	3000
Mittelwert	0.45102794	Summe Beobacht.	1353.08382
Std.abweichung	0.10141895	Varianz	0.0102858
Schiefe	-1.1718685	Kurtosis	2.57202331
Unkorr. Qu.summe	641.125728	Korr. Quad.summe	30.8471219
Variationskoeff.	22.4861781	Stdfeh. Mittelw.	0.00185165

Grundlegende Statistikmaße

Lage		Streuung	
Mittelwert	0.451028	Std.abweichung	0.10142
Median	0.464208	Varianz	0.01029
Modalwert	.	Spannweite	0.67430
		Interquartilsabstand	0.11324

Die Prozedur UNIVARIATE
Variable: beta_dyn_k_b

Momente

N	3000	Summe Gewichte	3000
Mittelwert	0.44523321	Summe Beobacht.	1335.69964
Std.abweichung	0.09346556	Varianz	0.00873581
Schiefe	-1.0561757	Kurtosis	4.34675257
Unkorr. Qu.summe	620.896539	Korr. Quad.summe	26.1986985
Variationskoeff.	20.9924957	Stdfeh. Mittelw.	0.00170644

Grundlegende Statistikmaße

Lage		Streuung	
Mittelwert	0.445233	Std.abweichung	0.09347
Median	0.452712	Varianz	0.00874
Modalwert	.	Spannweite	0.80305
		Interquartilsabstand	0.09813

11.6.2.2.8 Verteilung der Chi-Quadrat-Statistik des globalen Goodness of Fit-Tests der nichtadj. Analyse

Bootstrapanalyse Patientendaten - Ziehen einz. Augen
3000 samples; n=474 –

Verteilung der χ^2 -Statistik des globalen Goodness of Fit-Tests

Die Prozedur MEANS
Analysis Variable : _ESTIM_ overall CHI SQUARE

N	Mittelwert	Std. abweichung	Minimum	Maximum
3000	17.7195166	7.5117960	2.0009115	57.1583872

Die Prozedur UNIVARIATE
Variable: _ESTIM_ (overall CHI SQUARE)

Momente

N	3000	Summe Gewichte	3000
Mittelwert	17.7195166	Summe Beobacht.	53158.5498
Std.abweichung	7.51179598	Varianz	56.4270788
Schiefe	0.79399194	Kurtosis	0.93552687
Unkorr. Qu.summe	1111168.61	Korr. Quad.summe	169224.809
Variationskoeff.	42.3927817	Stdfeh. Mittelw.	0.137146

Grundlegende Statistikmaße

Lage		Streuung	
Mittelwert	17.71952	Std.abweichung	7.51180
Median	16.75227	Varianz	56.42708
Modalwert	.	Spannweite	55.15748
		Interquartilsabstand	10.12134

Tests auf Lageparameter: Mu0=0

Test	-Statistik-		------p-Wert------	
Studentsches t	t	129.2018	Pr > |t|	<.0001
Vorzeichen	M	1500	Pr >= |M|	<.0001
Vorzeichen-Rang	S	2250750	Pr >= |S|	<.0001

Bootstrapanalyse Patientendaten - Ziehen einz. Augen
3000 samples; n=474 - Verteilung des Chi-Quadrat

Quantile (Definition 5)

Quantil	Schätzwert
100% Max	57.15839
99%	39.19679
95%	31.78949
90%	27.44616
75% Q3	22.32168
50% Median	16.75227
25% Q1	12.20034
10%	8.92883
5%	7.24541
1%	4.95221
0% Min	2.00091

Die Prozedur UNIVARIATE
Variable: _ESTIM_ (overall CHI SQUARE)
Extreme Beobachtungen

-------Kleinste-------		--------Größte--------	
Wert	Beobachtung	Wert	Beobachtung
2.00091	1	49.8711	2996
2.46865	2	50.6124	2997
2.55155	3	52.5684	2998
2.61489	4	54.8502	2999
2.70001	5	57.1584	3000

11.6.2.2.9 Verteilung der Chi-Quadrat-Statistik des globalen Goodness of Fit-Tests der adj. Analyse

Bootstrapanalyse Patientendaten - Ziehen v. Augenpaaren
3000 samples n=237

Die Prozedur MEANS
Analysis Variable : _ESTIM_ overall CHI SQUARE

N	Mittelwert	Std. abweichung	Minimum	Maximum
3000	20.7879665	9.2629520	1.7393399	64.9916483

Bootstrapanalyse Patientendaten - Ziehen v. Augenpaaren
3000 samples n=237 - Verteilung des Chi-Quadrat

Die Prozedur UNIVARIATE
Variable: _ESTIM_ (overall CHI SQUARE)

Momente

N	3000	Summe Gewichte	3000
Mittelwert	20.7879665	Summe Beobacht.	62363.8994
Std.abweichung	9.26295201	Varianz	85.80228
Schiefe	0.86441407	Kurtosis	1.04737536
Unkorr. Qu.summe	1553739.69	Korr. Quad.summe	257321.038
Variationskoeff.	44.5592022	Stdfeh. Mittelw.	0.16911759

Grundlegende Statistikmaße

Lage		Streuung	
Mittelwert	20.78797	Std.abweichung	9.26295
Median	19.52587	Varianz	85.80228
Modalwert	.	Spannweite	63.25231
		Interquartilsabstand	12.13864

Tests auf Lageparameter: Mu0=0

Test	-Statistik-		------p-Wert------	
Studentsches t	t	122.9202	Pr > |t|	<.0001
Vorzeichen	M	1500	Pr >= |M|	<.0001
Vorzeichen-Rang	S	2250750	Pr >= |S|	<.0001

Quantile (Definition 5)

Quantil	Schätzwert
100% Max	64.99165
99%	47.93514
95%	38.15711
90%	33.12950
75% Q3	26.09811
50% Median	19.52587
25% Q1	13.95947
10%	10.16343
5%	8.38454
1%	4.97665
0% Min	1.73934

Bootstrapanalyse Patientendaten - Ziehen v. Augenpaaren

3000 samples n=237 - Verteilung des Chi-Quadrat

Die Prozedur UNIVARIATE
Variable: _ESTIM_ (overall CHI SQUARE)

Extreme Beobachtungen

-------Kleinste-------		--------Größte--------	
Wert	Beobachtung	Wert	Beobachtung
1.73934	1	60.3628	2996
2.33049	2	60.5114	2997
2.50280	3	62.8019	2998
2.79600	4	63.0187	2999
2.82558	5	64.9916	3000

11.6.3 Zweiter AbschNitt: Das Verhältnis zw. Augenseitenkorrelation und korrekturfaktor – Überprüfung der Simulationsergebnisse (1 Abschn., Teil A) durch empir. Patientendaten

11.6.3.1 A Programmcode für M2

```
*BOOTSTRAPprogramm;
*******************************;
*******************************;
LIBNAME boot_str 'D:\AMOS\237 Pat 140 Kontr\data';
options nonotes nosource nosource2 errors=0;
*1. DATA Step für Ziehung unter Berücksichtigung der Paarigkeit;
*****************************************************************;
/*The RANUNI function returns a number that is generated from the uniform distribution
on the interval (0,1) using a prime modulus multiplicative generator with modulus 231-
and multiplier 397204094 (Fishman and Moore 1982) (See References).
You can use a multiplier to change the length of the interval and an added constant to
move the interval. For example,
random_variate=a*ranuni(seed)+b;
returns a number that is generated from the uniform distribution on the interval (b,a+b).
*/
*Zufallszahlen von 1 bis 237 erzeugen;
%MACRO bootstrap (i, seed, dateinam=, lib=);
DATA &lib..&dateinam;
         DO i=1 to &i;
         random_variate=237*ranuni(&seed)+0.5;
         OUTPUT;
         END;
RUN;
%MEND bootstrap;
%macro create_1(howmany);
   %do i=1 %to &howmany;
      %bootstrap (237, &i, dateinam=ranuni&i, lib=boot_str);
   %end;
%mend create_1;
%create_1(3000)

*Runden;
%MACRO bootstrap2 (dateinam=, lib=);
DATA &lib..r_&dateinam;
SET &lib..&dateinam;
         rand_var_r=round(random_variate);
RUN;
%MEND bootstrap2;
%macro create_2(howmany);
   %do i=1 %to &howmany;
      %bootstrap2 (dateinam=ranuni&i, lib=boot_str);
   %end;
%mend create_2;
%create_2(3000)

*Sortieren;
%MACRO bootstrap3 (dateinam=, lib=);
PROC SORT DATA=&lib..r_&dateinam OUT=&lib..rs_&dateinam;
BY rand_var_r;
%MEND bootstrap3;
%macro create_3(howmany);
   %do i=1 %to &howmany;
      %bootstrap3 (dateinam=ranuni&i, lib=boot_str);
   %end;
%mend create_3;
%create_3(3000)

*1. DATA Step für Ziehung unter Missachtung der Paarigkeit;
*********************************************************;
*Zufallszahlen von 1 bis 474 erzeugen;
%MACRO bootstrap (i, seed, dateinam=, lib=);
DATA &lib..&dateinam;
         DO i=1 to &i;
         random_variate=474*ranuni(&seed)+0.5;
         OUTPUT;
         END;
RUN;
%MEND bootstrap;
%macro create_1(howmany);
   %do i=1 %to &howmany;
```

```
      %bootstrap (474, &i, dateinam=ranuni_aug&i, lib=boot_str);
   %end;
%mend create_1;
%create_1(3000)

*Runden;
%MACRO bootstrap2 (dateinam=, lib=);
DATA &lib..r_&dateinam;
SET &lib..&dateinam;
        rand_var_r=round(random_variate);
RUN;
%MEND bootstrap2;
%macro create_2(howmany);
   %do i=1 %to &howmany;
      %bootstrap2 (dateinam=ranuni_aug&i, lib=boot_str);
   %end;
%mend create_2;
%create_2(3000)

*Sortieren;
%MACRO bootstrap3 (dateinam=, lib=);
PROC SORT DATA=&lib..r_&dateinam OUT=&lib..rs_&dateinam;
BY rand_var_r;
%MEND bootstrap3;
%macro create_3(howmany);
   %do i=1 %to &howmany;
      %bootstrap3 (dateinam=ranuni_aug&i, lib=boot_str);
   %end;
%mend create_3;
%create_3(3000)
*II;
***
*800 A u g e n von 400 Patienten ziehen aus dem gleichen Datensatz, aus dem 400 P a t i e n t e n
*mit ihren jeweiligen Augenpaaren gezogen wurden (jeweils mit Zurücklegen);
/*Addieren der Datendatei linker Augen, aus der Bootstrapsamples gezogen werden
und der Datendatei rechter Augen, aus der Bootstrapsamples gezogen werden*/
******************;
*Ab hier Hochzählen;
******************;

*EINLESEN der SPSS DATEI von PM vergl_3_2 ;
*****************************************;
LIBNAME in SPSS 'D:\AMOS\237 Pat 140 Kontr\Patienten 2 mal 237.por';
*PROC PRINT DATA= in.spssdat;
*RUN;
DATA Boot_str.Patienten_2_mal_237;
SET in.spssdat;
KEEP PATNR AUGSEIT NRR ERG VEP MD ;
RUN;

*Teilen in  Dateien linker und rechter Augen;
DATA Boot_str.Patienten_237_l;
SET Boot_str.Patienten_2_mal_237;
IF AUGSEIT= 'L';
RUN;
DATA Boot_str.Patienten_237_r;
SET Boot_str.Patienten_2_mal_237;
IF AUGSEIT=  'R';
RUN;
*Zuordnen einer Fortlaufenden Nummer;
DATA Boot_str.Patienten_237_lb;
SET Boot_str.Patienten_237_l;
rand_var_r=_N_;
RUN;
DATA Boot_str.Patienten_237_rb;
SET Boot_str.Patienten_237_r;
rand_var_r=_N_;
RUN;


*************************************************************************
*hier geht das Mergen für augenpaarbezogene Ziehung und alle weiteren Prozesse los
*************************************************************************
*3 A. DATA STEP für die linke Seite;
*********************************;
*Mergen;
*entspricht nach SAS-Language Concept (~-Sprachkonzepte)
dem Beispiel 'Match-Merged with Duplicate BY Values';
%MACRO bootstrap4 (dateinam=, lib=);
DATA &lib..merged_l_&dateinam;
MERGE &lib..rs_&dateinam &lib..Patienten_237_lb;
BY rand_var_r;
%MEND bootstrap4;
%macro create_4(howmany);
```

```
   %do i=1 %to &howmany;
      %bootstrap4 (dateinam=ranuni&i, lib=boot_str);
   %end;
%mend create_4;
%create_4(3000)

*Löschen der 'nicht gezogenen' Records in den "gemergeten" Datendateien;
%MACRO bootstrap5 (dateinam=, lib=);
DATA &lib..sample_l_&dateinam;
SET &lib..merged_l_&dateinam;
IF random_variate EQ "." THEN DELETE;
DROP random_variate i;
RUN;
%MEND bootstrap5;
%macro create_5(howmany);
   %do i=1 %to &howmany;
      %bootstrap5 (dateinam=ranuni&i, lib=boot_str);
   %end;
%mend create_5;
%create_5(3000)

*3 B. DATA STEP für die rechte Seite;
**********************************;
%MACRO bootstrap4 (dateinam=, lib=);
DATA &lib..merged_r_&dateinam;
MERGE &lib..rs_&dateinam &lib..Patienten_237_rb;
BY rand_var_r;
%MEND bootstrap4;
%macro create_4(howmany);
   %do i=1 %to &howmany;
      %bootstrap4 (dateinam=ranuni&i, lib=boot_str);
   %end;
%mend create_4;
%create_4(3000)
*Löschen der 'nicht gezogenen' Records in den "gemergeten" Datendateien;
%MACRO bootstrap5 (dateinam=, lib=);
DATA &lib..sample_r_&dateinam;
SET &lib..merged_r_&dateinam;
IF random_variate EQ "." THEN DELETE;
DROP random_variate i;
RUN;
%MEND bootstrap5;
%macro create_5(howmany);
   %do i=1 %to &howmany;
      %bootstrap5 (dateinam=ranuni&i, lib=boot_str);
   %end;
%mend create_5;
%create_5(3000)

*Addieren der linken und rechten Augen;
**************************************;
%MACRO bootstrap6 (dateinam=, lib=);
DATA &lib..sample_&dateinam;
SET &lib..sample_l_&dateinam &lib..sample_r_&dateinam;
LABEL augseit='Augenseite';
RUN;
%MEND bootstrap6;
%macro create_6(howmany);
   %do i=1 %to &howmany;
      %bootstrap6 (dateinam=ranuni&i, lib=boot_str);
   %end;
%mend create_6;
%create_6(3000)

*Sortieren nach Patient;
%MACRO bootstrap7 (dateinam=, lib=);
PROC SORT DATA=&lib..sample_&dateinam;
        BY rand_var_r augseit;
%MEND bootstrap7;
%macro create_7(howmany);
   %do i=1 %to &howmany;
      %bootstrap7 (dateinam=ranuni&i, lib=boot_str);
   %end;
%mend create_7;
%create_7(3000)
%Macro proccalis (dateinam=, lib=);
proc calis DATA=&lib..sample_&dateinam OUTEST=&lib..oute_sample_&dateinam tech=nr nop edf=473;
         Lineqs
         NRR=              beta_nrr  Fakt_G   + beta_nrr_e      Err_nrr,
         ERG=              beta_erg  Fakt_G   + beta_erg_e      Err_erg,
         VEP=              beta_vep  Fakt_G   + beta_vep_e      Err_vep,
         MD =              beta_md   Fakt_G   + beta_md_e       Err_md;
```

```
Std
                  Err_nrr = 1,
                  Err_erg = 1,
                  Err_vep  = 1,
                  Err_md   = 1,
                  Fakt_G = 1;
         run;
         %MEND proccalis;
%macro create_7(howmany);
   %do i=1 %to &howmany;
      %proccalis (dateinam=ranuni&i, lib=boot_str);
       run;
   %end;
%mend create_7;
%create_7(3000)
*5.;
*Von den (3000) OUTEST-sas-data-sets wird nur jeweils die interessierende Zeile (Beobachtung)
*mit dem _TYPE_ "PARMS" (=final Parameter Estimates) bewahrt (und der Rest gelöscht);
%Macro parms_isol (dateinam=, lib=);
DATA &lib..PARMS_sample_&dateinam;
SET &lib..oute_sample_&dateinam;
IF _TYPE_ NE "PARMS" THEN DELETE;
RUN;
%MEND parms_isol;
%macro create_8(howmany);
   %do i=1 %to &howmany;
      %parms_isol (dateinam=ranuni&i, lib=boot_str);
       run;
   %end;
%mend create_8;
%create_8(3000)

*6.;
*die (3000) einzeiligen Datensätze addieren (unter Verwendung eines "iterativen Do-Makros");
%macro create_9(howmany);
   %do i=2 %to &howmany;
   proc APPEND BASE= BOOT_STR.PARMS_SAMPLE_RANUNI1
           DATA= BOOT_STR.PARMS_SAMPLE_RANUNI&i;
      run;
   %end;
%mend create_9;
%create_9(3000)
*7.;
*Umbenennen, Ausdrucken, Labeln der Datei;
DATA BOOT_STR.PARMS_ALL_RANUNI;
SET BOOT_STR.PARMS_SAMPLE_RANUNI1;
RUN;
PROC MEANS DATA=BOOT_STR.PARMS_ALL_RANUNI; OUTPUT OUT=BOOT_STR.OUT_PARMS_ALL_RANUNI;
TITLE1 "Bootstrapanalyse Patientendaten - Ziehen v. Augenpaaren)";
TITLE2 "30 samples n=237";
RUN;
PROC GCHART DATA=BOOT_STR.PARMS_ALL_RANUNI;
VBAR beta_nrr
          beta_erg
          beta_vep
          beta_md
          beta_nrr_e
          beta_erg_e
          beta_vep_e
          beta_md_e;
TITLE1 "Bootstrapanalyse Patientendaten - Ziehen v. Augenpaaren)";
TITLE2 "3000 samples n=237";
RUN;
*Es stellt sich heraus, dass - wie in den AMOS-Modellen auch schon - mitunter negative Werte für
*die Koeffizienten der Pfade geschätzt werden, die vom Faktor zu den manifesten Variablen führen
*(beta1, beta2 und beta3). Dieser Tatsache soll keine Beachtung geschenkt werden und hat etwas mit dem
*Schätzalgorithmus zu tun. Daher folgende Datenmanipulation:
8.;
*die negativen Werte von beta* werden nun in positive umgewandelt in dem man neue Variablen
*beta*_b erzeugt, die den Betrag von beta* angeben;
*die alten Variablen werden gelöscht;
*Schließlich werden statistische Maßzahlen und Histogramme für die sechs interessierenden
*Pfadkoeffizienten erstellt;
DATA BOOT_STR.PARMS_ALL_RANUNI_B;
SET BOOT_STR.PARMS_ALL_RANUNI;
                  beta_nrr_b =ABS(beta_nrr);
                  beta_erg_b =ABS(beta_erg);
                  beta_vep_b =ABS(beta_vep);
                  beta_md_b  =ABS(beta_md);
DROP
beta_nrr
beta_erg
beta_vep
beta_md;
```

```
TITLE ;
RUN;
PROC MEANS DATA=BOOT_STR.PARMS_ALL_RANUNI_B; OUTPUT OUT=BOOT_STR.OUT_PARMS_ALL_RANUNI_B;
TITLE1 "Bootstrapanalyse Patientendaten - Ziehen v. Augenpaaren";
TITLE2 "3000 samples n=237";
RUN;
DATA BOOT_STR.OUT_B_PARMS_ALL_RANUNI_B;
SET BOOT_STR.OUT_PARMS_ALL_RANUNI_B;
IF _STAT_ NE "STD" THEN DELETE;
DROP _TYPE_ _RHS_ _ITER_;
RUN;

*dieser Output STD (empirischer Standardfehler bzw. nach neuer Nomenklatur 'simulierter Standardfehler')
*dient dem späeteren Vergleich (Quotient aus augenbez. und personenbez. Bootstrap-SE);
PROC GCHART DATA=BOOT_STR.PARMS_ALL_RANUNI_B;
VBAR beta_nrr_b
          beta_erg_b
          beta_vep_b
          beta_md_b;
TITLE1 "Bootstrapanalyse Patientendaten - Ziehen v. Augenpaaren)";
TITLE2 "3000 samples n=237";
RUN;

*************************************************************************************************
*3. DATA STEP für die Ziehung von einzelnen Augen ohne Berücksichtigung der Paarigkeit;
*************************************************************************************;
*Vorbereiten der Empir. Datendatei:
*Zuordnen einer Fortlaufenden Nummer;
DATA Boot_str.Patienten_2_mal_237_b;
SET Boot_str.Patienten_2_mal_237;
rand_var_r=_N_;
RUN;
*Mergen;
%MACRO bootstrap4 (dateinam=, lib=);
DATA &lib..merged_&dateinam;
MERGE &lib..rs_&dateinam &lib..Patienten_2_mal_237_b;
BY rand_var_r;
%MEND bootstrap4;
%macro create_4(howmany);
   %do i=1 %to &howmany;
      %bootstrap4 (dateinam=ranuni_aug&i, lib=boot_str);
   %end;
%mend create_4;
%create_4(3000)
*Löschen der 'nicht gezogenen' Records in den "gemergeten" Datendateien;
%MACRO bootstrap5 (dateinam=, lib=);
DATA &lib..sample_&dateinam;
SET &lib..merged_&dateinam;
IF random_variate EQ "." THEN DELETE;
DROP random_variate i;
RUN;
%MEND bootstrap5;
%macro create_5(howmany);
   %do i=1 %to &howmany;
      %bootstrap5 (dateinam=ranuni_aug&i, lib=boot_str);
   %end;
%mend create_5;
%create_5(3000)
%Macro proccalis (dateinam=, lib=);
proc calis DATA=&lib..sample_&dateinam OUTEST=&lib..oute_sample_&dateinam tech=nr nop edf=473;
         Lineqs
         NRR=              beta_nrr  Fakt_G   + beta_nrr_e      Err_nrr,
         ERG=              beta_erg  Fakt_G   + beta_erg_e      Err_erg,
         VEP=              beta_vep  Fakt_G   + beta_vep_e      Err_vep,
         MD =              beta_md   Fakt_G   + beta_md_e       Err_md;

         Std
                  Err_nrr = 1,
                  Err_erg = 1,
                  Err_vep  = 1,
                  Err_md   = 1,
                  Fakt_G  = 1;
                  run;
         %MEND proccalis;

%macro create_7(howmany);
   %do i=1 %to &howmany;
      %proccalis (dateinam=ranuni_aug&i, lib=boot_str);
       run;
   %end;
%mend create_7;
%create_7(3000)
*5.;
*Von den (3000) OUTEST-sas-data-sets wird nur jeweils die interessierende Zeile (Beobachtung)
```

```
*mit dem _TYPE_ "PARMS" (=final Parameter Estimates) bewahrt (und der Rest gelöscht);
%Macro parms_isol (dateinam=, lib=);
DATA &lib..PARMS_sample_&dateinam;
SET &lib..oute_sample_&dateinam;
IF _TYPE_ NE "PARMS" THEN DELETE;
RUN;
TITLE;
   run;
%MEND parms_isol;
%macro create_8(howmany);
   %do i=1 %to &howmany;
      %parms_isol (dateinam=ranuni_aug&i, lib=boot_str);
       run;
   %end;
%mend create_8;
%create_8(3000)

*6.;
*die (3000) einzeiligen Datensätze addieren (unter Verwendung eines „iterativen Do-Makros“);
%macro create_9(howmany);
   %do i=2 %to &howmany;
   proc APPEND BASE= BOOT_STR.PARMS_SAMPLE_ranuni_aug1
            DATA= BOOT_STR.PARMS_SAMPLE_ranuni_aug&i;
      run;
   %end;
%mend create_9;
%create_9(3000)
*7.;
*Umbenennen, Ausdrucken, Labeln der Datei;
DATA BOOT_STR.PARMS_ALL_ranuni_aug;
SET BOOT_STR.PARMS_SAMPLE_ranuni_aug1;
RUN;
PROC MEANS DATA=BOOT_STR.PARMS_ALL_RANUNI_AUG; OUTPUT OUT=BOOT_STR.OUT_PARMS_ALL_RANUNI_AUG;
TITLE1 "Bootstrapanalyse Paperdaten - Ziehen v. einzelnen Augen";
TITLE2 "30 samples n=474";
RUN;

PROC GCHART DATA=BOOT_STR.PARMS_ALL_ranuni_aug;
VBAR beta_nrr
          beta_erg
          beta_vep
          beta_md
          beta_nrr_e
          beta_erg_e
          beta_vep_e
          beta_md_e;
TITLE1 "Bootstrapanalyse Patientendaten - Ziehen v. einzelnen Augen";
TITLE2 "Rohwerte - 3000 samples n=474";
RUN;

*Es stellt sich heraus, dass – wie in den AMOS-Modellen auch schon – mitunter negative Werte für
*die Koeffizienten der Pfade geschätzt werden, die vom Faktor zu den manifesten Variablen führen
*(beta1, beta2 und beta3). Dieser Tatsache soll keine Beachtung geschenkt werden, da sie auf den
*Schätzalgorithmus zurückgeht. Daher folgende Datenmanipulation:
8.;
*die negativen Werte von beta* werden nun in positive umgewandelt, in dem man neue Variablen
*beta*_b erzeugt, die den Betrag von beta* angeben;
*die alten Variablen werden gelöscht;
*Schließlich werden statistische Maßzahlen und Histogramme für die sechs interessierenden
*Pfadkoeffizienten erstellt;
DATA BOOT_STR.PARMS_ALL_ranuni_aug_B;
SET BOOT_STR.PARMS_ALL_ranuni_aug;
                  beta_nrr_b =ABS(beta_nrr);
                  beta_erg_b =ABS(beta_erg);
                  beta_vep_b =ABS(beta_vep);
                  beta_md_b =ABS(beta_md);
DROP
beta_nrr
beta_erg
beta_vep
beta_md;
TITLE ;
RUN;

PROC MEANS DATA=BOOT_STR.PARMS_ALL_ranuni_aug_B;
OUTPUT OUT=BOOT_STR.OUT_PARMS_ALL_RANUNI_aug_B;
TITLE1 "Bootstrapanalyse Patientendaten - Ziehen v. einzelnen Augen)";
TITLE2 "3000 samples n=474";
RUN;
*Zur Vorbereitung der SE-Auswertung;
DATA BOOT_STR.OUT_B_PARMS_ALL_RANUNI_AUG_B;
SET BOOT_STR.OUT_PARMS_ALL_RANUNI_AUG_B;
IF _STAT_ NE "STD" THEN DELETE;
DROP _TYPE_ _RHS_ _ITER_;
```

```
RUN;
PROC GCHART DATA=BOOT_STR.PARMS_ALL_ranuni_aug_B;
VBAR beta_nrr_b
          beta_erg_b
          beta_vep_b
          beta_md_b;
TITLE1 "Bootstrapanalyse Patientendaten - Ziehen v. einzelnen Augen";
TITLE2 "Absolute Werte - 3000 samples n=474";
RUN;
*************************************************************************************************
*ad 3.;
******;
/*Umbenennen und Mergen der OUTPUT DATEIEN aus PROC MEANS der Bootstrapanalyse unter Ziehung einzelner Augen und
unter Ziehung von Augenpaaren (SD der Faktor-Pfadkoeefizienten)*/
DATA BOOT_STR.OUT_B2_PARMS_ALL_RANUNI_AUG_B;
SET BOOT_STR.OUT_B_PARMS_ALL_RANUNI_AUG_B;
        beta_nrr_b_A =beta_nrr_b;
                  beta_erg_b_A=beta_erg_b;
                  beta_vep_b_A=beta_vep_b;
                  beta_md_b_A= beta_md_b;
          beta_nrr_e_A=beta_nrr_e;
     beta_erg_e_A=beta_erg_e;
          beta_vep_e_A=beta_vep_e;
          beta_md_e_A=beta_md_e;
DROP      beta_nrr_b
                  beta_erg_b
                  beta_vep_b
                  beta_md_b
                  beta_nrr_e
        beta_erg_e
             beta_vep_e
             beta_md_e;
RUN;
DATA BOOT_STR.OUT_B2_PARMS_ALL_RANUNI_B;
SET BOOT_STR.OUT_B_PARMS_ALL_RANUNI_B;
                  beta_nrr_b_AP =beta_nrr_b;
                  beta_erg_b_AP=beta_erg_b;
                  beta_vep_b_AP=beta_vep_b;
                  beta_md_b_AP= beta_md_b;
     beta_nrr_e_AP=beta_nrr_e;
     beta_erg_e_AP=beta_erg_e;
          beta_vep_e_AP=beta_vep_e;
          beta_md_e_AP=beta_md_e;
DROP      beta_nrr_b
          beta_erg_b
          beta_vep_b
          beta_md_b
          beta_nrr_e
          beta_erg_e
          beta_vep_e
          beta_md_e;
RUN;
DATA BOOT_STR.OUT_PARMS_ALL_RANUNI_AUG_AUGPAAR;
MERGE BOOT_STR.OUT_B2_PARMS_ALL_RANUNI_AUG_B BOOT_STR.OUT_B2_PARMS_ALL_RANUNI_B;
RUN;
PROC PRINT DATA = BOOT_STR.OUT_PARMS_ALL_RANUNI_AUG_AUGPAAR;
TITLE1 "SD der Pfadkoeff. bzw. einer Fehlerkorr. d. BS-Analysen";
TITLE2 "Ziehen v. Augenpaaren und einz. Augen - 3000 samples n=237 bzw. 474";
RUN;

DATA BOOT_STR.QUOT_PARMS_ALL_RANUNI;
SET BOOT_STR.OUT_PARMS_ALL_RANUNI_AUG_AUGPAAR;
Quot_beta_nrr =    beta_nrr_b_AP / beta_nrr_b_A;
Quot_beta_erg =    beta_erg_b_AP / beta_erg_b_A;
Quot_beta_vep =    beta_vep_b_AP / beta_vep_b_A;
Quot_beta_md =     beta_md_b_AP / beta_md_b_A;
Quot_beta_nrr_e    =         beta_nrr_e_AP / beta_nrr_e_A;
Quot_beta_erg_e    =         beta_erg_e_AP / beta_erg_e_A;
Quot_beta_vep_e    =         beta_vep_e_AP / beta_vep_e_A;
Quot_beta_md_e     =         beta_md_e_AP / beta_md_e_A;
DROP
          beta_nrr_b_AP beta_nrr_b_A
          beta_erg_b_AP beta_erg_b_A
          beta_vep_b_AP beta_vep_b_A
          beta_md_b_AP beta_md_b_A
          beta_nrr_e_AP beta_nrr_e_A
          beta_erg_e_AP beta_erg_e_A
          beta_vep_e_AP beta_vep_e_A
          beta_md_e_AP beta_md_e_A;
RUN;
PROC PRINT DATA = BOOT_STR.QUOT_PARMS_ALL_RANUNI;
TITLE1 "Quotient aus SD der Pfadkoeff. bzw. einer Fehler-Korr. d. BS-Analysen ";
TITLE2 "Ziehen v. Augenpaaren und einz. Augen - 3000 samples n=237 bzw. 474";
RUN;
```

11.6.3.2 B Dokumentation des SAS-Outputs von M2

Bootstrapanalyse Patientendaten - Ziehen v. Augenpaaren
3000 samples n=237

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		3000	0.0075620	0.0073377	3.8385973E-6	0.0603514
ITER	Iteration	0	.	.	.	.
beta_nrr	beta_nrr	3000	0.4676540	0.4128120	-0.7530968	0.7916865
beta_nrr_e	beta_nrr_e	3000	0.7806105	0.0398823	0.6109275	0.8951561
beta_vep	beta_vep	3000	0.3822178	0.3377424	-0.6653785	0.6674087
beta_vep_e	beta_vep_e	3000	0.8595219	0.0331684	0.7446916	0.9606611
beta_erg	beta_erg	3000	0.4185490	0.3683518	-0.7075138	0.6977221
beta_erg_e	beta_erg_e	3000	0.8296550	0.0291843	0.7066995	0.9237581
beta_md	beta_md	3000	0.6139797	0.5395673	-0.9726299	1.0000002
beta_md_e	beta_md_e	3000	0.5704857	0.0808787	-0.000060561	0.8049296

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		3000	0.0075620	0.0073377	3.8385973E-6	0.0603514
ITER	Iteration	0	.	.	.	.
beta_nrr_e	beta_nrr_e	3000	0.7806105	0.0398823	0.6109275	0.8951561
beta_vep_e	beta_vep_e	3000	0.8595219	0.0331684	0.7446916	0.9606611
beta_erg_e	beta_erg_e	3000	0.8296550	0.0291843	0.7066995	0.9237581
beta_md_e	beta_md_e	3000	0.5704857	0.0808787	-0.000060561	0.8049296
beta_nrr_b		3000	0.6217542	0.0497967	0.4457527	0.7916865
beta_erg_b		3000	0.5558252	0.0433589	0.3829765	0.7075138
beta_vep_b		3000	0.5068788	0.0565447	0.2777232	0.6674087
beta_md_b		3000	0.8154698	0.0549255	0.5933703	1.0000002

Bootstrapanalyse Paperdaten - Ziehen v. einzelnen Augen

3000 samples n=474

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		3000	0.0058762	0.0055326	5.1560058E-6	0.0407866
ITER	Iteration	0	.	.	.	.
beta_nrr	beta_nrr	3000	0.4592587	0.4243331	-0.7582537	0.7812598
beta_nrr_e	beta_nrr_e	3000	0.7796205	0.0357100	0.6242060	0.9070843
beta_vep	beta_vep	3000	0.3738888	0.3462739	-0.6302855	0.6753332
beta_vep_e	beta_vep_e	3000	0.8599753	0.0280037	0.7375128	0.9414667
beta_erg	beta_erg	3000	0.4091270	0.3779671	-0.6618377	0.6942557
beta_erg_e	beta_erg_e	3000	0.8301765	0.0247132	0.7197284	0.9204155
beta_md	beta_md	3000	0.6002876	0.5534631	-0.9327884	0.9458431
beta_md_e	beta_md_e	3000	0.5739014	0.0638278	0.3246241	0.7489077

Die Prozedur MEANS

Variable	Etikett	N	Mittelwert	Std. abweichung	Minimum	Maximum
RHS		3000	0.0058762	0.0055326	5.1560058E-6	0.0407866
ITER	Iteration	0	.	.	.	.
beta_nrr_e	beta_nrr_e	3000	0.7796205	0.0357100	0.6242060	0.9070843
beta_vep_e	beta_vep_e	3000	0.8599753	0.0280037	0.7375128	0.9414667
beta_erg_e	beta_erg_e	3000	0.8301765	0.0247132	0.7197284	0.9204155
beta_md_e	beta_md_e	3000	0.5739014	0.0638278	0.3246241	0.7489077
beta_nrr_b		3000	0.6236422	0.0445888	0.4209491	0.7812598
beta_erg_b		3000	0.5557277	0.0369264	0.3909416	0.6942557
beta_vep_b		3000	0.5073806	0.0471608	0.3371059	0.6753332
beta_md_b		3000	0.8152330	0.0442752	0.6626743	0.9458431

SD der Pfadkoeff. bzw. einer Fehlerkorr. d. BS-Analysen
Ziehen v. Augenpaaren und einz. Augen - 3000 samples n=237 bzw. 474

Beob.	_FREQ_	_STAT_	beta_ nrr_b_A	beta_ erg_b_A	beta_ vep_b_A	beta_md_ b_A	beta_ nrr_e_A	beta_ erg_e_A	beta_ vep_e_A	beta_md_ e_A
1	3000	STD	0.044589	0.036926	0.047161	0.044275	0.035710	0.024713	0.028004	0.063828

Beob.	beta_ nrr_b_AP	beta_ erg_b_AP	beta_ vep_b_AP	beta_md_ b_AP	beta_ nrr_e_AP	beta_ erg_e_AP	beta_ vep_e_AP	beta_md_ e_AP
1	0.049797	0.043359	0.056545	0.054926	0.039882	0.029184	0.033168	0.080879

Quotient aus SD der Pfadkoeff. bzw. einer Fehler-Korr. d. BS-Analysen
Ziehen v. Augenpaaren und einz. Augen - 3000 samples n=237 bzw. 474

Beob.	_FREQ_	_STAT_	Quot_ beta_nrr	Quot_ beta_erg	Quot_ beta_vep	Quot_ beta_md	Quot_ beta_ nrr_e	Quot_ beta_ erg_e	Quot_ beta_ vep_e	Quot_ beta_ md_e
1	3000	STD	1.11680	1.17420	1.19898	1.24055	1.11684	1.18092	1.18443	1.26714

11.6.4 Zu Teil A: Mögliche Kombinationen von Augenseitenkorrelation des Schweregrads und des Messfehlers bei Augenseitenkorrelation des Messverfahrens von 0,8

Dass die Augenseitenkorrelation der Messverfahren nicht hinreicht, um die Höhe des Korrekturfaktors für den SE der Pfadkoeffizienten zu erklären, kann mit Hilfe der Ergebnisse aus 8.2.1 (Erster Abschnitt, Teil A: Bootstrapanalysen an simulierten Daten) illustriert werden. In den Analysen zu Teil A werden standardisierte Pfadkoeffizienten mit wahrem Wert 0,71 vorausgesetzt. In den Analysen zu 8.2.2 (Teil B: Bootstrapanalysen an Patientendaten) entfällt selbst diese vereinfachende, standardisierende Annahme, so dass die Höhe der Augenseitenkorrelation der Messverfahren noch weniger Rückschlüsse auf den nötigen Korrekturfaktor zulässt, als in der Analyse an simulierten Daten.

Betrachten wir den Fall einer Augenseitenkorrelation zweier Messverfahren von 0,80. Eine Korrelation in dieser Höhe kann zum einen dadurch zustande kommen, dass eine Augenseitenkorrelation des Schweregrades von 0,7 mit einer Seitenkorrelation der Messfehler von 0,9 kombiniert wird (siehe Tabelle 45).

Tabelle 45: Augenseitenkorrelation zweier Messverfahren von 0,80 bei Kombination einer Seitenkorrelation des Schweregrades von 0,7 mit einer Seitenkorrelation der Messfehler von 0,9

Korrelationen

		indik1_li_79	indik2_li_79	indik3_li_79
indik1_li_79	Korrelation nach Pearson	1	,500(**)	,500(**)
	Signifikanz (2-seitig)		.000	.000
	N	1517	1517	1517
indik2_li_79	Korrelation nach Pearson	,500(**)	1	,500(**)
	Signifikanz (2-seitig)	.000		.000
	N	1517	1517	1517
indik3_li_79	Korrelation nach Pearson	,500(**)	,500(**)	1
	Signifikanz (2-seitig)	.000	.000	
	N	1517	1517	1517
indik1_re_79	Korrelation nach Pearson	,800(**)	,350(**)	,350(**)
	Signifikanz (2-seitig)	.000	.000	.000
	N	1517	1517	1517
indik2_re_79	Korrelation nach Pearson	,350(**)	,800(**)	,350(**)
	Signifikanz (2-seitig)	.000	.000	.000
	N	1517	1517	1517
indik3_re_79	Korrelation nach Pearson	,350(**)	,350(**)	,800(**)
	Signifikanz (2-seitig)	.000	.000	.000
	N	1517	1517	1517

** Die Korrelation ist auf dem Niveau von 0.01 (2-seitig) signifikant.

R (Indik1_li_79, Indik1_re_79): Korrelation des am linken und rechten Auge durchgeführten ersten Messverfahrens, das durch eine Seitenkorrelation des Schweregrades von 0,7 und eine Seitenkorrelation der Messfehler von 0,9 charakterisiert ist.

Eine Seitenkorrelation des Messverfahrens von 0,80 kann aber beispielsweise auch umgekehrt durch Kombination einer Seitenkorrelation des Schweregrades von 0,9 mit einer Seitenkorrelation der Messfehler von 0,7 entstehen (siehe Tabelle 46).

Tabelle 46: Augenseitenkorrelation zweier Messverfahren von 0,80 bei Kombination einer Seitenkorrelation des Schweregrades von 0,9 mit einer Seitenkorrelation der Messfehler von 0,7

Korrelationen

		indik1_li_97	indik2_li_97	indik3_li_97
indik1_li_97	Korrelation nach Pearson	1	,500(**)	,500(**)
	Signifikanz (2-seitig)		.000	.000
	N	1517	1517	1517
indik2_li_97	Korrelation nach Pearson	,500(**)	1	,500(**)
	Signifikanz (2-seitig)	.000		.000
	N	1517	1517	1517
indik3_li_97	Korrelation nach Pearson	,500(**)	,500(**)	1
	Signifikanz (2-seitig)	.000	.000	
	N	1517	1517	1517
indik1_re_97	Korrelation nach Pearson	,800(**)	,450(**)	,450(**)
	Signifikanz (2-seitig)	.000	.000	.000
	N	1517	1517	1517
indik2_re_97	Korrelation nach Pearson	,450(**)	,800(**)	,450(**)
	Signifikanz (2-seitig)	.000	.000	.000
	N	1517	1517	1517
indik3_re_97	Korrelation nach Pearson	,450(**)	,450(**)	,800(**)
	Signifikanz (2-seitig)	.000	.000	.000
	N	1517	1517	1517

** Die Korrelation ist auf dem Niveau von 0.01 (2-seitig) signifikant.

R (Indik1_li_97, Indik1_re_97): Korrelation des am linken und rechten Auge durchgeführten ersten Messverfahrens, das durch eine Seitenkorrelation des Schweregrades von 0,9 und eine Seitenkorrelation der Messfehler von 0,7 charakterisiert ist.

Aus Tabelle 12 in 8.2.1 geht hervor, dass der geschätzte Korrekturfaktor für die SE der Pfadkoeffizienten von Messverfahren, deren Augenseitenkorrelation Tabelle 45 entspricht, im Bereich zwischen 1,26 und 1,34 liegt. Für Messverfahren mit Augenseitenkorrelationen entsprechend Tabelle 46 hingegen werden Korrekturfaktoren nur in einer Höhe zwischen 1,18 und 1,26 ermittelt.

11.7 Auswirkungen der Verletzung von Modellannahmen in SEM im Lichte der Forschungsliteratur und Diskussion

11.7.1 Programmcode analog Yang et al. (2012) zur Errechnung der Varianzanteile

Im bisherigen Programmcode werden Pfadkoeffizienten für die Fehlervariablen geschätzt, in Yang et al. nicht. Stattdessen werden die Varianzen der Fehlervariablen auf 1 gesetzt, während sie in Yang et al. geschätzt werden.

```
/* Anpassung des bisherigen Programmcodes zum publizierten Modell [Martus (2001a), S. 269)] an die Schreibweise von
Yang et al. zur Schätzung der TRUE SCORE VARIANCE und TOTAL VARIANCE.

Hierbei entspricht strukturell genau Yangs Model 4 (1 Trait, 1 Gruppen- oder Methodenfaktor, mehrere Fehlerkorrelati-
on)*/

proc calis DATA=Boot_str.Patienten_2_mal_237 edf=473;
        Lineqs

NRR =           a1      F1                                      +       e1,
ERG =           a2      F1                                      +       e2,
VEP =           a3      F1                                      +       e3,
AMP =           a4      F1              + b1    F2              +       e4,
FLI =           a5      F1              + b2    F2              +       e5,
CLV =           a6      F1              + b3    F2              +       e6,
MD =            a7      F1              + b4    F2              +       e7,
DYN =           a8      F1              + b5    F2              +       e8,
Ftrue1 = g1 F1 + d1,
Ftrue2 = g2 F2 + d2,
Ftrue=Ftrue1+Ftrue2 +d3,
Ftotal=NRR + ERG + VEP + AMP + FLI + CLV + MD + DYN + d4;

        Std
                e1 - e8 = ve1 - ve8,
                F1 = 1,
                F2 = 1,
                d1 = 0,
                d2 = 0,
                d3 = 0,
                d4 = 0;
        Cov
                f1 f2 = 0,
                e1 e8 = The1,
                e3 e4 = The2,
                e3 e7 = The3,
                e4 e5 = The4,
                e5 e6 = The5;
        lincon
                g1-a1-a2-a3-a4-a5-a6-a7-a8=0,
                g2-b1-b2-b3-b4-b5=0;
run;
```

Variation 1 (Entfernung von Theta 3) in allen Statements identisch mit Referenzmodell bis auf

```
**********
Cov
                f1 f2=0,
                e1 e8 = The1,
                e3 e4 = The2,
                e4 e5 = The4,
                e5 e6 = The5;
```

Variation 2

```
**********
proc calis DATA=Boot_str.Patienten_2_mal_237 tech=nr edf=473;
        Lineqs

NRR=            a1      F1                                      +       e1,
ERG=            a2      F1                                      +       e2,
VEP=            a3      F1                                      +       c1      F3  +  e3,
AMP=            a4      F1              + b1    F2              +       c2      F3  +  e4,
FLI=            a5      F1              + b2    F2              +       e5,
CLV=            a6      F1              + b3    F2              +       e6,
MD =            a7      F1              + b4    F2              +       e7,
DYN=            a8      F1              + b5    F2              +       e8,
Ftrue1=g1 F1 +d1,
Ftrue2=g2 F2 +d2,
Ftrue3=g3 F3 +d3,
Ftrue=Ftrue1+Ftrue2+Ftrue3 +d4,
Ftotal=NRR + ERG + VEP + AMP + FLI + CLV + MD + DYN +d5;

        Std
                e1 - e8 = ve1 - ve8,
                F1 - F3 = 1,
                d1=0,
                d2=0,
                d3=0,
                d4=0,
```

```
                d5=0;
        Cov
                f1 f2=0,
                f1 f3=0,
                f3 f2=0,
                e1 e8 = The1,
                e3 e7 = The3,
                e4 e5 = The4,
                e5 e6 = The5;
        lincon
                g1-a1-a2-a3-a4-a5-a6-a7-a8=0,
                g2-b1-b2-b3-b4-b5=0,
                g3-c1-c2=0;
run;
```

Variation 3 (Entfernung von Theta 3) in allen Statements identisch mit Variation 2 bis auf

```
***********
Cov
                f1 f2=0,
                f1 f3=0,
                f3 f2=0,
                e1 e8 = The1,
                e4 e5 = The4,
                e5 e6 = The5;
```

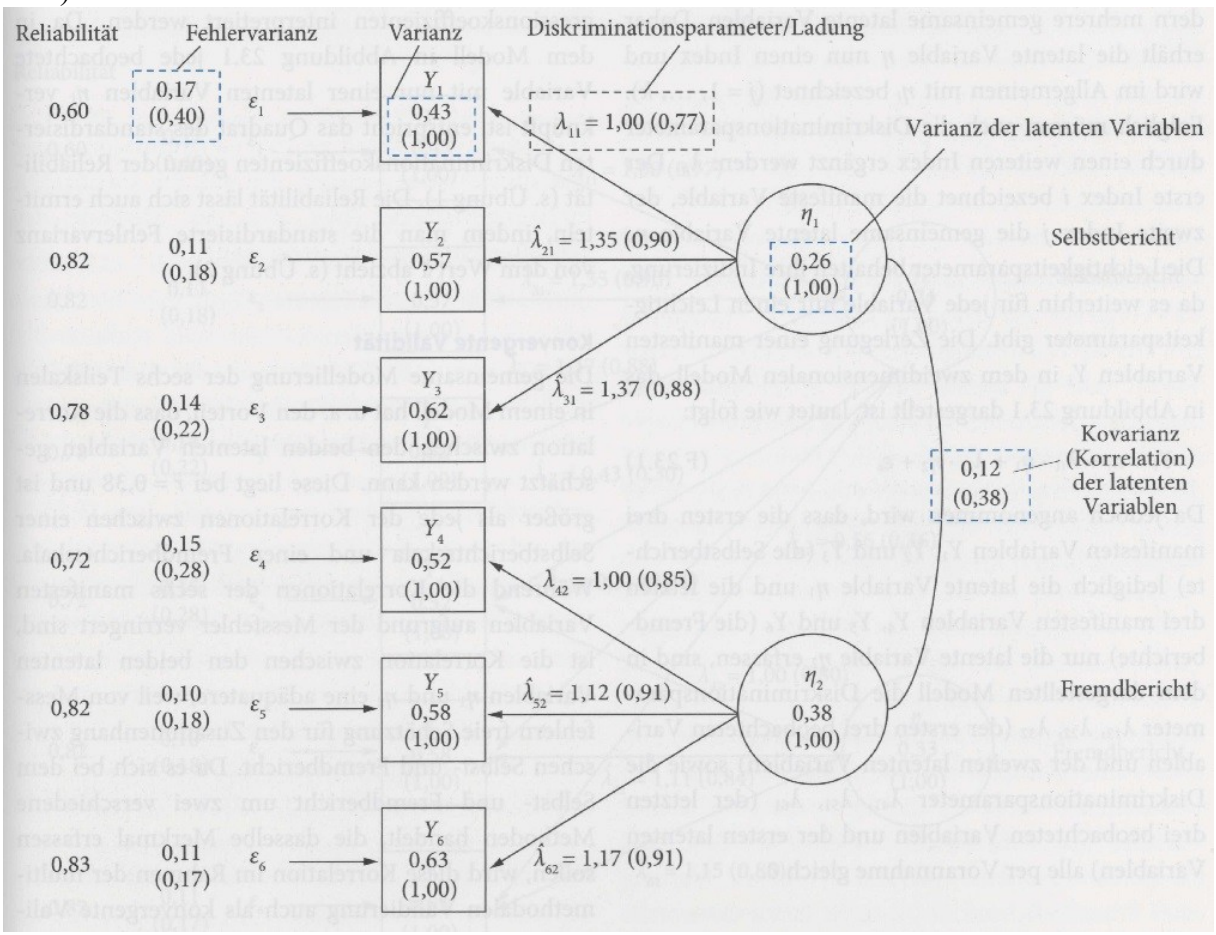
Variation 4

```
***********
proc calis DATA=Boot_str.Patienten_2_mal_237 method=max edf=473;
        Lineqs

NRR=            a1      F1                                              +       e1,
ERG=            a2      F1                                              +       e2,
VEP=            a3      F1                                              +        c1 F3  + e3,
AMP=            a4      F1              + b1            F2              +        c2 F3  + e4,
FLI=            a5      F1              + b2            F2              +       e5,
CLV=            a6      F1              + b3            F2              +        h1 F4  + e6,
MD =            a7      F1              + b4            F2              +        h2 F4  + e7,
DYN=            a8      F1              + b5            F2              +       e8,
Ftrue1=g1 F1 +d1,
Ftrue2=g2 F2 +d2,
Ftrue3=g3 F3 +d3,
Ftrue4=g4 F4 +d4,
Ftrue=Ftrue1+Ftrue2+Ftrue3+Ftrue4 +d5,
Ftotal=NRR + ERG + VEP + AMP + FLI + CLV + MD + DYN +d6;

        Std
                e1 - e8 = ve1 - ve8,
                F1 - F4 = 1,
                d1=0,
                d2=0,
                d3=0,
                d4=0,
                d5=0,
                d6=0;
        Cov
                f1 f2=0,
                f1 f3=0,
                f1 f4=0,
                f2 f3=0,
                f2 f4=0,
                f3 f4=0,
                e1 e8 = The1,
                e3 e7 = The3,
                e4 e5 = The4,
                e5 e6 = The5;
        lincon
                g1-a1-a2-a3-a4-a5-a6-a7-a8=0,
                g2-b1-b2-b3-b4-b5=0,
                g3-c1-c2=0,
                g4-h1-h2=0;
run;
```

11.7.2 Modelle zur Erfassung des habituellen Wohlbefindens, Eid et al. (2011)

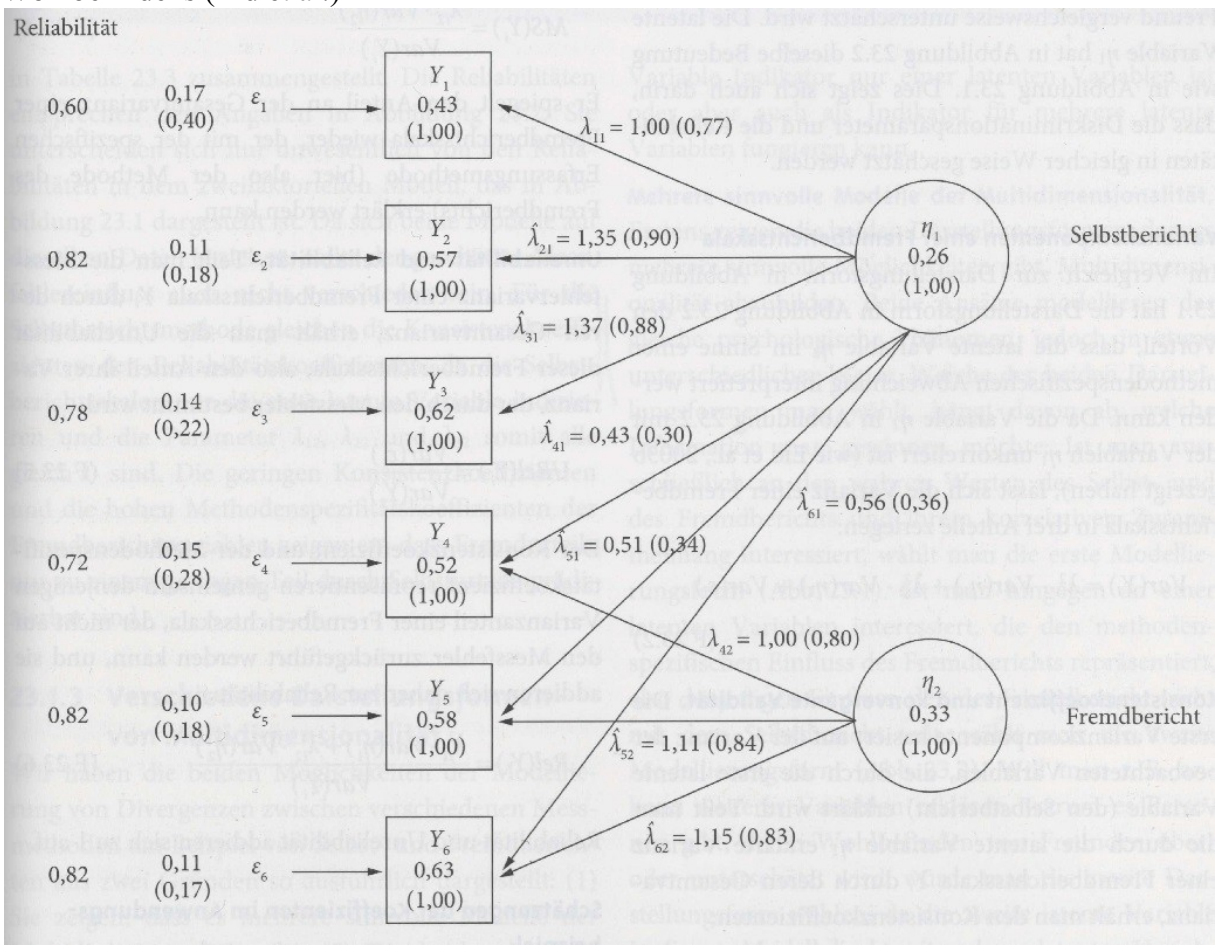
11.7.2.1 Zweidimensionales Modell

Abbildung 93: Zweidimensionales Modell zur Erfassung des habituellen Wohlbefindens (Eid et al.)



11.7.2.2 Zweidimensionales Modell mit Methodenfaktor

Abbildung 94: Zweidimensionales Modell mit Methodenfaktortzur Erfassung des habituellen Wohlbefindens (Eid et al.)



11.7.3 MODELLANNAHMEN UND SCHÄTZMETHODEN – NICHTNORMALITÄT

Tabelle 47: Statistische Kennwerte der acht Diagnoseverfahren getrennt nach Augenseite in der Gruppe der Patienten mit (Verdacht auf) Glaukom

Statistics

Augenseite			nrr	fli	vep	erg	md	amp	dyn	clv
L	N	Valid	237	237	237	237	237	237	237	237
		Missing	0	0	0	0	0	0	0	0
	Mean		-1,1544	-1,6654	1,5202	-3.0524	,8410	-,8327	-35,7510	,2291
	Median		-1,1400	-1,7135	1,5064	-2,8974	,8172	-,8228	-38,8299	,1851
	Skewness		-,938	,419	,196	-,900	,622	-,298	,390	,421
	Std. Error of Skewness		,158	,158	,158	,158	,158	,158	,158	,158
	Kurtosis		3,444	-.041	-,484	1,659	.027	,179	-,418	-1.027
	Std. Error of Kurtosis		,315	,315	,315	,315	,315	,315	,315	,315
R	N	Valid	237	237	237	237	237	237	237	237
		Missing	0	0	0	0	0	0	0	0
	Mean		-1,1206	-1,6842	1,5272	-3.0608	,8550	-,8380	-36,9613	,2250
	Median		-1,1200	-1,7343	1,5156	-3.0053	,8260	-,8274	-40,3637	,1851
	Skewness		-,347	,257	,132	-,607	,444	-,334	,423	,428
	Std. Error of Skewness		,158	,158	,158	,158	,158	,158	,158	,158
	Kurtosis		,278	-,156	-,479	,554	,245	-,151	-,498	-1,115
	Std. Error of Kurtosis		,315	,315	,315	,315	,315	,315	,315	,315

Abbildung 95: Werteverteilung der korrigierten Varianz des Verlustes (CLV) am linken Auge

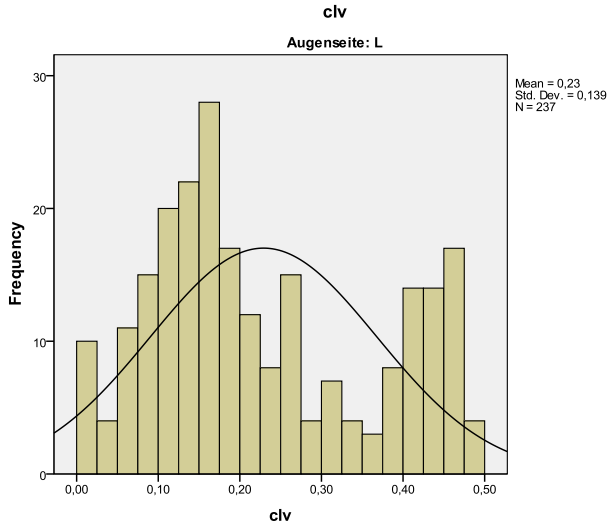


Abbildung 96: Werteverteilung der korrigierten Varianz des Verlustes (CLV) am rechten Auge

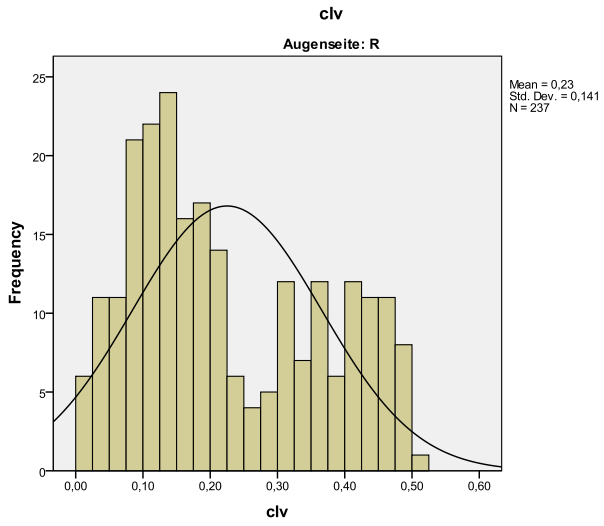


Tabelle 48: Test auf uni- und multivariate Normalverteilung der acht Diagnoseverfahren in Patienten mit SAS - beide Augenseiten gemeinsam sowie getrennt nach Augenseite

The MODEL Procedure

Anzahl der Beobachtungen		Statistiken für System	
Used	474	Objective	291.2044
Missing	0	Objective*N	138031

Normalverteilungstest

Gleichung	Teststatistik	Wert	Wahr. scheinl.
NRR	Shapiro-Wilk W	0.98	0.0144
ERG	Shapiro-Wilk W	0.96	<.0001
VEP	Shapiro-Wilk W	0.97	0.0019
AMP	Shapiro-Wilk W	0.98	0.0036
FLI	Shapiro-Wilk W	0.97	<.0001
CLV	Shapiro-Wilk W	0.90	<.0001
MD	Shapiro-Wilk W	0.96	<.0001
DYN	Shapiro-Wilk W	0.95	<.0001
System	Mardia Skewness	643.4	<.0001
	Mardia Kurtosis	7.14	<.0001
	Henze-Zirkler T	11.99	<.0001

---------- Augenseite=L ----------

The MODEL Procedure

Anzahl der Beobachtungen		Statistiken für System	
Used	237	Objective	294.5584
Missing	0	Objective*N	69810

Normalverteilungstest

Gleichung	Teststatistik	Wert	Wahr. scheinl.
NRR	Shapiro-Wilk W	0.97	0.0014
ERG	Shapiro-Wilk W	0.96	<.0001
VEP	Shapiro-Wilk W	0.97	0.0243
AMP	Shapiro-Wilk W	0.98	0.1496
FLI	Shapiro-Wilk W	0.96	0.0006
CLV	Shapiro-Wilk W	0.90	<.0001
MD	Shapiro-Wilk W	0.95	<.0001
DYN	Shapiro-Wilk W	0.94	<.0001
System	Mardia Skewness	484.1	<.0001
	Mardia Kurtosis	6.30	<.0001
	Henze-Zirkler T	7.23	<.0001

---------- Augenseite=R ----------

The MODEL Procedure

Anzahl der Beobachtungen		Statistiken für System	
Used	237	Objective	287.1170
Missing	0	Objective*N	68047

Normalverteilungstest

Gleichung	Teststatistik	Wert	Wahr. scheinl.
NRR	Shapiro-Wilk W	0.98	0.2728
ERG	Shapiro-Wilk W	0.97	0.0019
VEP	Shapiro-Wilk W	0.97	0.0171
AMP	Shapiro-Wilk W	0.97	0.0364
FLI	Shapiro-Wilk W	0.98	0.1316
CLV	Shapiro-Wilk W	0.90	<.0001
MD	Shapiro-Wilk W	0.96	0.0007
DYN	Shapiro-Wilk W	0.95	<.0001
System	Mardia Skewness	275.3	<.0001
	Mardia Kurtosis	1.73	0.0840
	Henze-Zirkler T	3.70	0.0002

Tabelle 49: Test auf uni- und multivariate Normalverteilung der acht Diagnoseverfahren in Patienten mit STATA - beide Augenseiten gemeinsam sowie getrennt nach Augenseite

beide Augenseiten

Test for univariate normality

Variable	Pr(Skewness)	Pr(Kurtosis)	joint adj chi2(2)	joint Prob>chi2
nrr	0.0000	0.0000	48.68	0.0000
fli	0.0028	0.7110	8.54	0.0140
vep	0.1474	0.0111	8.14	0.0171
erg	0.0000	0.0008	38.03	0.0000
md	0.0000	0.5849	17.69	0.0001
amp	0.0058	0.8841	7.39	0.0249
dyn	0.0004	0.0083	16.66	0.0002
clv	0.0003	0.0000	.	0.0000

Test for multivariate normality

Mardia mSkewness = 8.082041	chi2(120) = 643.425	Prob>chi2 = 0.0000
Mardia mKurtosis = 88.29593	chi2(1) = 50.972	Prob>chi2 = 0.0000
Henze-Zirkler = 1.219557	chi2(1) = 143.735	Prob>chi2 = 0.0000
Doornik-Hansen	chi2(16) = 239.644	Prob>chi2 = 0.0000

Forts.: linke Augenseite

Test for univariate normality

Variable	Pr(Skewness)	Pr(Kurtosis)	joint adj chi2(2)	joint Prob>chi2
nrr_l	0.0000	0.0000	39.46	0.0000
fli_l	0.0094	0.9965	6.49	0.0389
vep_l	0.2109	0.0553	5.26	0.0719
erg_l	0.0000	0.0008	28.91	0.0000
md_l	0.0002	0.8214	12.01	0.0025
amp_l	0.0598	0.4946	4.04	0.1325
dyn_l	0.0152	0.1177	7.76	0.0206
clv_l	0.0091	0.0000	42.12	0.0000

Test for multivariate normality

Mardia mSkewness	=	12.06845	chi2(120) =	484.094	Prob>chi2 = 0.0000
Mardia mKurtosis	=	90.35684	chi2(1) =	39.721	Prob>chi2 = 0.0000
Henze-Zirkler	=	1.133362	chi2(1) =	52.295	Prob>chi2 = 0.0000
Doornik-Hansen			chi2(16) =	158.617	Prob>chi2 = 0.0000

Forts.: rechte Augenseite

Test for univariate normality

Variable	Pr(Skewness)	Pr(Kurtosis)	joint adj chi2(2)	joint Prob>chi2
nrr_r	0.0297	0.3424	5.59	0.0612
fli_r	0.1037	0.6826	2.84	0.2418
vep_r	0.3972	0.0592	4.31	0.1158
erg_r	0.0003	0.1102	13.48	0.0012
md_r	0.0061	0.3891	7.70	0.0213
amp_r	0.0358	0.6968	4.60	0.1005
dyn_r	0.0088	0.0460	9.71	0.0078
clv_r	0.0080	0.0000	60.63	0.0000

Test for multivariate normality

Mardia mSkewness	=	6.864098	chi2(120) =	275.335	Prob>chi2 = 0.0000
Mardia mKurtosis	=	82.83946	chi2(1) =	2.986	Prob>chi2 = 0.0840
Henze-Zirkler	=	1.04892	chi2(1) =	13.691	Prob>chi2 = 0.0002
Doornik-Hansen			chi2(16) =	109.810	Prob>chi2 = 0.0000

12 Abbildungsverzeichnis

Abbildung 1: Kausalhypothese über zwei manifeste Variable 21
Abbildung 2: Kausalhypothese über zwei latente Variable 22
Abbildung 3: Messmodell der latenten exogenen Variablen ξ 22
Abbildung 4: Messmodell der latenten endogenen Variablen η 23
Abbildung 5: Strukturmodell mit einer endogenen und einer exogenen Variable..... 24
Abbildung 6: Vollständiges lineares Strukturgleichungsmodell für zwei latente Variable 24
Abbildung 7: Ein-Faktor-Modell für Diagnoseverfahren mit unabhängigen Messfehlern..... 28
Abbildung 8: Konstruktion der beobachtbaren Variablen 44
Abbildung 9: Der Konstruktion entsprechendes AMOS-Modell..... 44
Abbildung 10: Allgemeine Modellspezifikation des Grundmodells..... 53
Abbildung 11: Schätzung des korrekt spezifizierten Grundmodells..... 53
Abbildung 12: Modell mit zwei positiv korrelierten Fehlervariablen - allgemeine Modelspezifikation..... 56
Abbildung 13: Korrekt spezifiziertes Modell mit zwei positiv korrelierten Fehlervariablen .. 57
Abbildung 14: Fehlspezifiziertes Modell mit zwei positiv korrelierten Fehlervariablen 57
Abbildung 15: Korrekt spezifiziertes Modell mit zwei negativ korrelierten Fehlervariablen . 59
Abbildung 16: Fehlspezifiziertes Modell mit zwei negativ korrelierten Fehlervariablen 59
Abbildung 17: Modell mit drei positiv korrelierten Fehlern - allgemeine Modellspezifikation 61
Abbildung 18: Korrekt spezifiziertes Modell mit drei positiv korrelierten Fehlern 62
Abbildung 19: Fehlspezifiziertes Modell mit drei positiv korrelierten Fehlern..... 62
Abbildung 20: Strukturgleichungsmodell 64
Abbildung 21: Strukturmodell 65
Abbildung 22: Korrekt spezifiziertes Modell 1..... 68
Abbildung 23: Fehlspezifiziertes Modell 2..... 70
Abbildung 24: Korrekt spezifiziertes Modell 2..... 70
Abbildung 25: Fehlspezifiziertes Modell 3..... 72
Abbildung 26: Korrekt spezifiziertes Modell 3..... 72
Abbildung 27: Fehlspezifiziertes Modell Z1 74
Abbildung 28: Korrekt spezifiziertes Modell Z1 74
Abbildung 29: Fehlspezifiziertes Modell Z2 76
Abbildung 30: Korrekt spezifiziertes Modell Z2 76
Abbildung 31: Fehlspezifiziertes Modell Z3 77
Abbildung 32: Korrekt spezifiziertes Modell Z3 77
Abbildung 33: Korrekt spezifiziertes Modell 4..... 80
Abbildung 34: Fehlspezifiziertes Modell 5..... 81
Abbildung 35: Korrekt spezifiziertes Modell 5..... 81
Abbildung 36: Fehlspezifiziertes Modell 6..... 83
Abbildung 37: Korrekt spezifiziertes Modell 6..... 83
Abbildung 38: Fehlspezifiziertes Modell 7..... 85
Abbildung 39: Korrekt spezifiziertes Modell 7..... 85
Abbildung 40: Fehlspezifiziertes Modell 8..... 87
Abbildung 41: Korrekt spezifiziertes Modell 8..... 87
Abbildung 42: Fehlspezifiziertes Modell 9..... 89
Abbildung 43: Korrekt spezifiziertes Modell 9..... 89
Abbildung 44: Modell mit vier Indikatoren - allgemeine Modellspezifikation 111

Abbildung 45: Modell mit fünf Indikatoren - allgemeine Modellspezifikation 113
Abbildung 46: Korrekt spezifiziertes Modell 1 120
Abbildung 47: Histogramm der Schätzungen für gamma des Modells 1 121
Abbildung 48: Fehlspezifiziertes Modell 2 122
Abbildung 49: Histogramm der Verteilung von gamma des Modells 2 unter Verwendung der Ursprungswerte 124
Abbildung 50: Histogramm der Verteilung von gamma des Modells 2 unter Verwendung des Betrags von Ursprungswerten kleiner -0,0414 125
Abbildung 51: Fehlspezifiziertes Modell 3 127
Abbildung 52: Korrekt spezifiziertes Modell 4 130
Abbildung 53: Fehlspezifiziertes Modell 5 131
Abbildung 54: Fehlspezifiziertes Modell 6 133
Abbildung 55: Fehlspezifiziertes Modell 7 135
Abbildung 56: Fehlspezifiziertes Modell 8 137
Abbildung 57: Fehlspezifiziertes Modell 9 140
Abbildung 58: Korrekt spezifiziertes Modell 5 143
Abbildung 59: Korrekt spezifiziertes Modell Z4 144
Abbildung 60: Fehlspezifiziertes Modell Z4 145
Abbildung 61: Korrekt spezifiziertes Modell 8 146
Abbildung 62: Korrekt spezifiziertes Modell 2 G 148
Abbildung 63: Korrekt spezifiziertes Modell 3 G 149
Abbildung 64: Korrekt spezifiziertes Modell 8 G 150
Abbildung 65: Korrekt spezifiziertes Modell 5 G 151
Abbildung 66: Korrekt spezifiziertes Modell Z4 G 152
Abbildung 67: Korrekt spezifiziertes Modell 6 G 153
Abbildung 68: Korrekt spezifiziertes Modell 9 G 154
Abbildung 69: Modell der Datensimulation zum Ein-Faktor-Modell in allgemeiner Form .. 169
Abbildung 70: Analysemodell der zum Ein-Faktor-Modell simulierten Daten 171
Abbildung 71: Messmodell zur Diagnose des allgemeinen Glaukomschweregrades mit acht Indikatoren 178
Abbildung 72: Verteilung der nichtadjustierten Pfadkoeffizientenschätzungen von CLV hinsichtlich des Glaukomschweregrads (absolute Werte) 182
Abbildung 73 Verteilung der nichtadjustierten Pfadkoeffizientenschätzungen von CLV hinsichtlich des Glaukomschweregrads (Ursprungswerte) 183
Abbildung 74: Verteilung der nichtadjustierten Pfadkoeffizientenschätzungen von CLV hinsichtlich der Konzentration (absolute Werte) 183
Abbildung 75: Verteilung der nichtadjustierten Pfadkoeffizientenschätzungen von CLV hinsichtlich der Konzentration (Ursprungswerte) 184
Abbildung 76: χ^2-Verteilung des nichtadjustierten globalen Goodness of Fit-Tests 188
Abbildung 77: χ^2-Verteilung des adjustierten globalen Goodness of Fit-Tests 191
Abbildung 78: Linearer Zusammenhang zwischen Korrekturfaktor für die Pfadkoeffizienten des Glaukomschweregrades bzw. der Konzentration und Augenseitenkorrelation 193
Abbildung 79: M1: Messmodell zur Schätzung der Augenseitenkorrelation latenter Variabler (pro Augenseite ein Faktor und vier Messverfahren) 198
Abbildung 80: M2: Messmodell mit einem Faktor und vier Messverfahren zur Schätzung von Korrekturfaktoren für M1 200
Abbildung 81: Judgement: dynamischer Prozess des Wiedererinnerns von Information (Stone & Litcher-Kelly) 229

Abbildung 82: Implementation von Methodenfaktoren in ein Modell über die Einstellung zum Mathematiklernen (Yang & Green) 237
Abbildung 83: Correlated trait/uncorrelated method model (CTUM; Eid, Lischetzke & Nussbeck) 240
Abbildung 84: Correlated trait/correlated uniqueness model (CTCU; Eid, Lischetzke & Nussbeck) 241
Abbildung 85: Messmodell zu „morale“ nach Entfernten des Traits „belonging“ (Rubio & Gillespie) 245
Abbildung 86: SEM zur Studie über wahrgenommenen Zusammenhalt mit den Faktoren „belonging“ und „morale“ (Bollen & Hoyle) 246
Abbildung 87: SEM-Schema A mit 2 Faktoren und je zwölf Indikatoren (Heene et al.) 252
Abbildung 88 SEM-Schema B mit 2 Faktoren und je zwölf Indikatoren (Heene et al.) 253
Abbildung 89: Verteilung von gamma des Modells 2 auf Basis von 20000 Iterationen (Ursprungswerte) 316
Abbildung 90: Verteilung von gamma des Modells 3 (Ursprungswerte) 318
Abbildung 91: Verteilung von gamma des Modells 3 (absolute Werte) 318
Abbildung 92: Verteilung von gamma des Modells 3 (absolute Werte, wenn gamma < -0,0244) 319
Abbildung 93: Zweidimensionales Modell zur Erfassung des habituellen Wohlbefindens (Eid et al.) 401
Abbildung 94: Zweidimensionales Modell mit Methodenfaktortzur Erfassung des habituellen Wohlbefindens (Eid et al.) 402
Abbildung 95: Werteverteilung der korrigierten Varianz des Verlustes (CLV) am linken Auge 403
Abbildung 96: Werteverteilung der korrigierten Varianz des Verlustes (CLV) am rechten Auge 404

13 Tabellenverzeichnis

Tabelle 1: Korrelationsmatrix der drei manifesten Variablen des Grundmodells 51
Tabelle 2: In AMOS geschätzte standardisierte Regressionsgewichte des Grundmodells 52
Tabelle 3: Korrelationsmatrix der in SPSS erzeugten Variablen des Grundmodells 54
Tabelle 4: Tatsächlich erzeugte, vor der Modellschätzung festgelegte und geschätzte Korrelationen des Grundmodells 67
Tabelle 5: Tatsächlich erzeugte, vor der Modellschätzung festgelegte und geschätzte Korrelationen der Ein-Faktor-Modelle mit drei Indikatoren (Übersicht) 91
Tabelle 6: Simulation zu Ein-Faktor-Modellen mit drei Indikatoren: empirischer und mittlerer geschätzter Standardfehler und Korrelationsmatrix der geschätzten Koeffizienten 106
Tabelle 7: Pfadkoeffizienten der Ein-Faktor-Modelle mit drei Indikatoren laut AMOS- bzw. PROC CALIS-Berechnung und PROC CALIS-Simulation 108
Tabelle 8: Ein-Faktor-Modelle mit vier oder fünf Indikatoren - Berechnung und Simulation mit PROC CALIS 115
Tabelle 9: Simulationen zu Modellen mit zwei Faktoren - Analyseergebnisse zum Strukturmodellpfadkoeffizienten gamma 156
Tabelle 10: 16 durch das Modell der Datensimulation erzeugte Korrelationsstrukturen 170
Tabelle 11: Pfadkoeffizienten und SE der nichtadjustierten Bootstrapanalyse über 16 simulierte Korrelationsstrukturen des Ein-Faktor-Modells mit drei Indikatoren 172
Tabelle 12: Pfadkoeffizienten und SE der adjustierten Bootstrapanalyse über 16 simulierte Korrelationsstrukturen sowie Korrekturfaktoren 174
Tabelle 13: Schätzungen der Pfadkoeffizienten zwischen den beiden Faktoren und den acht Diagnoseverfahren (via Originalsample) sowie der Standardfehler (via Bootstrapsamples) 180
Tabelle 14: Fünf Fehlerkorrelationen des Modells zur Diagnose des Glaukomschweregrads - Kennwerte der nichtadjustierten Bootstrapanalyse 185
Tabelle 15: Fünf Fehlerkorrelationen - Kennwerte der winsorierten Verteilungen 186
Tabelle 16: Vergleich der Kennwerte von fünf Fehlerkorrelationen auf Basis von Originalsample sowie nichtadjustierten und adjustierten Bootstrapanalysen 187
Tabelle 17: Quotienten aus simulierten SE der Pfadkoeffizienten der adjustierten und nicht adjustierten Bootstrapanalyse (Korrekturfaktoren) sowie links-rechts-Korrelation der Messverfahren 192
Tabelle 18: Quotienten aus SE der Fehlerkorrelationen der adjustierten und nicht adjustierten Bootstrapanalyse 195
Tabelle 19: Korrekturfaktoren für die nichtadjustierten SE der fünf Fehlerkorrelationen aus dem Modell zur Diagnose des Glaukomschwergrades mit acht Indikatoren 195
Tabelle 20: Aus M1 und M2 resultierende Pfadkoeffizientenschätzungen und SE (adjustierte sowie nicht adjustierte Bootstrapanalyse und Korrekturfaktoren) 201
Tabelle 21: Schätzung von Korrekturfaktoren für drei ausgewählte Korrelationsstrukturen der Tabelle 11 201
Tabelle 22: Pfadkoeffizienten und empirische Korrekturfaktoren aus M2 sowie angemessene simulierte Korrekturfaktoren 203
Tabelle 23: Heterogenität der Daten, wahre Reliabilität und Coeffizient alpha in four component data (Osburn) 220
Tabelle 24: Heterogenität der Daten, wahre Reliabilität und coeffizient alpha in eight component data (Osburn) 220
Tabelle 25: Wahre Reliabilität und gemittelte Alpha-Koeffizienten bei additiven und nichtadditiven Subtest-True-Scores (Zimmerman) 221

Tabelle 26: Typologie korrelierter Fehler – Ursachen und Effekte 223
Tabelle 27: Korrelationskoeffizienten der acht Diagnoseverfahren getrennt nach Kontroll- und Patientengruppe 231
Tabelle 28: Beispiele für additive und nichtadditive Subtest-True-Score-Matrizen (Zimmerman) 234
Tabelle 29: Alpha in Abhängigkeit von wahrer Reliabilität, Korrelation zwischnen Subtest-Error-Scores und Zahl der paarweise korrelierten Subtest-Error-Scores (Zimmerman) 234
Tabelle 30: Korrelationsmatrix, Mittelwert und Standardabweichung der Items einer Studie über die Einstellung zum Mathematiklernen (Yang & Green) 237
Tabelle 31: Modellvarianten der Einstellung zum Mathematiklernen - konvergente Validität und Reliabilität 239
Tabelle 32: MTMM-Matrix mit drei Traits und drei Methoden für ein strukturiertes Interview mit Angestellten (Conway) 242
Tabelle 33: Kovarianzmatrix der Studie über wahrgenommenen Zusammenhalt (Bollen & Hoyle) 244
Tabelle 34: Fit und Fitindizes von Modellvarianten des wahrgenommenen Zusammenhalts (Rubio & Gillespie) 245
Tabelle 35: Alpha, wahre Reliabilität und modifizierter Koeffizient alpha $_K$ bei unterschiedlicher Itemanzahl sowie Zahl und Stärke von Fehlerkorrelationen (Komaroff) .. 248
Tabelle 36 Spezifikationen von Fehlerkorrelationen unter verschiedenen Simulationsbedingungen (Heene et al.) 254
Tabelle 37: Simulierte Fehlerkorrelationen unter verschiedenen Simulationsbedingungen .. 254
Tabelle 38: χ^2 und Fitindizes in Abhängigkeit von Fehlerkorrelationsspezifikation und Stichprobenumfang bei geringen Faktorladungen 255
Tabelle 39: χ^2 und Fitindizes in Abhängigkeit von Fehlerkorrelationsspezifikation und Stichprobenumfang bei hohen Faktorladungen 256
Tabelle 40: Hauptkomponentenanalyse über 8 Diagnoseverfahren mit angezeigten Faktorladungen > 0,4 260
Tabelle 41: Variationen des Modells zur Glaukomdiagnose mit acht Indikatoren – Varianzanteile 261
Tabelle 42 Modellvariationen zur Glaukomdiagnose – (Cronbachs) Alpha und Koeffizienten der klassischen Testtheorie 263
Tabelle 43: Modellvariationen zur Glaukomdiagnose – Fit und Fitindizes 267
Tabelle 44: Über- und Unterschätzung der empirischen Standardfehler durch modellbasierte SE nach ML- bzw. WLS-Methode 271
Tabelle 45: Augenseitenkorrelation zweier Messverfahren von 0,80 bei Kombination einer Seitenkorrelation des Schweregrades von 0,7 mit einer Seitenkorrelation der Messfehler von 0,9 397
Tabelle 46: Augenseitenkorrelation zweier Messverfahren von 0,80 bei Kombination einer Seitenkorrelation des Schweregrades von 0,9 mit einer Seitenkorrelation der Messfehler von 0,7 398
Tabelle 47: Statistische Kennwerte der acht Diagnoseverfahren getrennt nach Augenseite in der Gruppe der Patienten mit (Verdacht auf) Glaukom 403
Tabelle 48: Test auf uni- und multivariate Normalverteilung der acht Diagnoseverfahren in Patienten mit SAS - beide Augenseiten gemeinsam sowie getrennt nach Augenseite 404
Tabelle 49: Test auf uni- und multivariate Normalverteilung der acht Diagnoseverfahren in Patienten mit STATA - beide Augenseiten gemeinsam sowie getrennt nach Augenseite 405

14 Literaturverzeichnis

Aarbuckle, J. L & Wothke, W. (1999). AMOS 4.0 User's Guide. Chicago, Ill.: SmallWaters Corperation.

Backhaus, K., Erichson, B., Plinke, W., Weber, R. (2000). Multivariate Analysemethoden. Berlin: Springer

Bentler, P. M. (1985) Theory and Implementation of EQS. A structural equation program. Manual for program version 2.0, Los Angeles: BMDP Statistical Software, Inc.

Bentler, P. M., Stein, J. A. (1992). Structural equation models in medical research. Statistical Methods in Medical Research 1: 159–81

Bollen, K. A. & Stine R. A. (1992). Bootstrapping goodness-of-fit measures in structural equation models. Sociological Methods & Research. 21: 205–229

Bortz, J. (1999). Statistik für Sozialwissenschaftler. Berlin: Springer.

Brown, Timothy A. (2006). Confirmatory factor analysis for applied research. New York: The Guilford Press.

Conway, J. M. (1998). Estimation and uses of the proportion of method variance for multitrait-multimethod data. Organizational Research Methods. 1: 209–222.

Eid M., Lischetzke T. & Nussbeck W.F. (2006). Structural equation methods for multitrait-multimethod data. In Eid, Diener (Hrsg.), Multimethod Measurement in Psychology. Washington: American Psychological Association.

Eid, M., Gollwitzer, M. Schmitt, M. (2011). Statistik und Forschungsmethoden. Weinheim: Beltz.

Finney, S. J. & DiStefano, C. (2006). Nonnormal and categorical data in structural equation models. In G. R. Hancock & R. O. Mueller (Eds.), A second course in structural equation modeling (pp. 269–314). Greenwich, CT: Information Age.

Green S. B. & Hershberger S. L. (2000). Correlated errors in true score models and their effect on coefficient alpha. *Structural equation modeling*, 7 (2) 251–270.

Heene M., Hilbert S, Freudenthaler H. H.& Bühner M. (2012). Sensitivity of SEM fit indexes with respect to violations of uncorrelated errors. Structural equation modeling: 19, 36–50.

Hu L.-T. & Bentler, P. M. (1995). Evaluating model fit. In Hoyle, Rick H. (Hrsg.), Structural equation modeling. Concepts, Issues and Applications. Thousand Oaks, CA.: Sage Publications.

Kaplan, D. (2000). Structural Equation Modeling – Foundations and Extensions (Advandced Quantitative Techniques in the Social Sciences Series 10). Thousand Oaks, CA.: Sage Publications.

Kenny, D. A. (1979). Correlation and Causality. New York: Wiley-Science.

Komaroff E. (1997). Effect of simultaneous violations of essential τ-equivalence and uncorrelated error on coefficient $\propto$. Applied Psychological Measurement. 21: 337–348.

Lord, F. M. and Novick, M.R. (1968). Statistical theories of mental test scores, Reading, Massach.: Addison-Wesley.

Marsh, H. W. (1989) Confirmatory factor analysis of multitrait-multimethod data: many problems and a few solutions. Applied Psychological Measurement. 13: 335–361.

Martus P. et al. (2000). Multivariate approach for quantification of morphologic and functional damage in glaukoma. Investigative Ophtalmology & Visual Science. 41: 1099–1110.

Martus, P. (2001a). A measurement model of disease severity in absence of a gold standard. Methods of information in medicine. 40: 265–271.

Martus, P. (2001b). A measurement model of disease damage in paired organs. Biometrical Journal; 43 (8): 927–940.

Miller M. B. (1995). Coefficient alpha: A basic introduction from the perspectives of classical test theorie and structural equation modelling. Structural equation modeling. 2: 255–273.

Novick, M. R. and Lewis, C. (1967). Coefficient alpha and the reliability of composite measures. *Psychometrica*. 32: 1–13.

Oberski, D. L. (2009). Jrule for Mplus. Retrieved from https://github.com/daob/JruleMplus/wiki

Osburn, H.G. (2000). Coefficient alpha and related internal consistency reliable coefficients. Psychological Methods. 5: 343–355.

Podsakoff P. M., MacKenzie S. B., Lee, J.-Y. & Podsakoff N. P. (2003). Common method biases in behavioral research – a critical review of the literature and recommended remedies. Journal of applied Psychology. 88: 879–903.

Raykov T . & Shrout, P. E. (2002). Reliability of scales with general structure. Point and interval estimations using a structural equation modelling approach. *Structural equation modelling*. 9: 195–2012.

Raykov, T (1997a). Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement.* 21: 173–184.

Reinecke, J, Pöge, A. Strukturgleichungsmodelle. In C. Wolf & und H. Best (Hrsg.), Handbuch der sozialwissenschaftlichen Datenanalyse, (S. 775–804). Wiesbaden: VS-Verlag für Sozialwissenschaften.

Reinecke, J. (2005). Strukturgleichungsmodelle in den Sozialwissenschaften. München: Oldenburg-Verlag.

Rozeboom, W. W. (1966). Foundations of the theorie of prediction. Homewood, Ill.: Dorsey.

Rubio, D. M. & Gillespie, D. F. (1995). Problems with errors in structural equation models. Structural equation modeling, 2: 367–378.

Saris, W. E., Satorra, A. & van der Veld, W.M. (2009). Testing structural equation models or detection of misspecification? Structural equation modeling. 16: 561–582.

SAS Institute Inc (1999). SAS/STAT® User's Guide, Version 8. Cary, NC: SAS Institute Inc.

SAS Institute Inc (2003). SAS 9.1 Help and Documentation. Cary, NC: SAS Institute Inc.

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 variable analysis: Applications for developmental research (pp. 399–419). Thousand Oaks, CA: Sage.

Scullen, S. E. (1999). Using confirmatory factor analysis of correlated uniquenesses to estimate method variance in multitrait-multimethod matrices. Organizational Research Methods. 2: 275–292.

Simsek, G. G. & Noyan, F. (2013). Mc Donalds ω_t Cronbach's α generalized $\theta$for composite reliability of common factors structures. Communications in Statistics – Simulation and Computation, 42: 2008–2025.

Stone, A. A. & Litcher-Kelly, L., (2006). Momentary capture of real world data. In: M. Eid & E. Diener (Hrsg.), Handbook of multimethod measurement in psychology (pp. 61–72). Washington, DC: American Psychological Association.

Yang, Y. & Green, S. (2011). Coefficient alpha: A reliability coefficient of the 21 century? Journal of Psychoeducational Assessment. 29: 377–392.

Zimmerman, D. W. (2007). Correction for attenuation with biased reliability estimates and correlated errors in populations and samples. Educational and Psychological Measurement. 67: 920–939.

Zimmerman D. W. Zumbo B. D. & Lalonde C. (1993). Coefficient alpha as an estimate of test reliability under violation of two assumptions. Educational and Psychological Measurement. 53: 33 – 49.

15 Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, dass ich die Dissertation selbständig verfasst habe. Die benutzten Hilfsmittel und Quellen sind aufgeführt.

Berlin, 03.08.2016