

Universität Trier

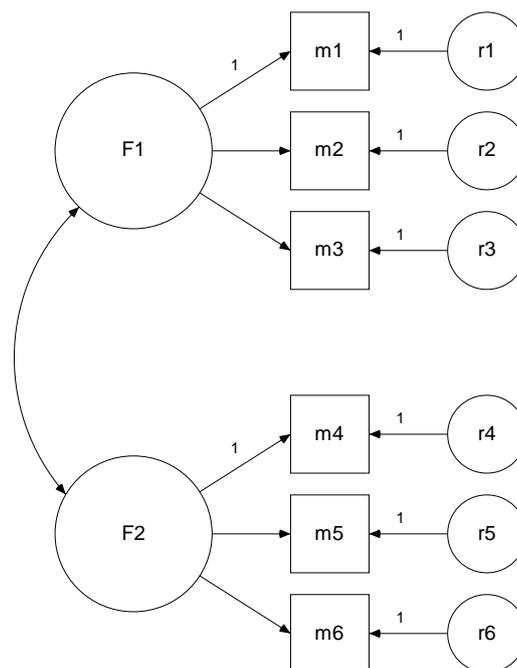
Zentrum für Informations-, Medien-  
und Kommunikationstechnologie  
(ZIMK)



Trier, den 01. 10. 2015

Bernhard Baltes-Götz

# Analyse von Strukturgleichungsmodellen mit Amos 18



<b>VORWORT .....</b>	<b>5</b>
<b>1 EINLEITUNG .....</b>	<b>6</b>
1.1 Strukturgleichungsmodelle mit latenten Variablen .....	6
1.2 Analyseoptionen in Amos 18 .....	7
<b>2 ALLER ANFANG IST LEICHT .....</b>	<b>8</b>
2.1 Pfaddiagramm zeichnen .....	9
2.2 Allgemeine Bedienungsmerkmale .....	11
2.2.1 Format der Zeichenfläche ändern und Pfaddiagramm neu einpassen .....	11
2.2.2 Objekte markieren .....	12
2.2.3 Objekte löschen .....	12
2.2.4 Objekte gestalten .....	12
2.2.5 Bearbeitungsschritte rückgängig machen oder wiederherstellen .....	13
2.2.6 Speichern, Sicherheitskopien .....	14
2.2.7 Pfaddiagramm in andere Anwendungen übertragen .....	15
2.3 Daten einlesen.....	15
2.4 Modellparameter schätzen .....	17
2.4.1 Textergebnisse.....	17
2.4.2 Erste Ergebnisinterpretationen .....	19
2.4.3 Das Ausgabe-Pfaddiagramm .....	20
2.5 Optionale Ausgaben.....	21
2.6 Amos in .NET - Programmen verwenden .....	22
2.7 Modellgültigkeitstest .....	26
2.7.1 Erforderliche Stichprobengröße .....	27
2.7.2 Verteilungsannahmen bei Strukturgleichungsanalysen.....	27
2.7.3 Normalverteilungs- und Ausreißerdiagnostik in Amos .....	28
2.7.4 Parameterrestriktionen formulieren .....	30
2.7.5 $\chi^2$ -Gültigkeitstest zum Modell mit den Parameterrestriktionen.....	32
2.8 Beschriftungen und andere Möglichkeiten zur Gestaltung des Pfaddiagramms.....	34
2.8.1 Diagrammbeschriftungen .....	35
2.8.2 Formatierungsoptionen.....	36
2.9 Parameterrestriktionen in AmosEngine-Programmen formulieren.....	37
2.10 Parameter per C.R. - oder Modellgültigkeitstest auf Signifikanz prüfen .....	38
<b>3 MODELLE FÜR MANIFESTE VARIABLEN.....</b>	<b>41</b>
3.1 Beispiel .....	41
3.2 Einlesen von Momentmatrizen.....	42
3.3 Pfaddiagramme mit gerichteten Pfeilen und latenten Variablen.....	43

3.4	Ein technisches Identifikationsproblem .....	45
3.5	Standardisierte Schätzwerte und Determinationskoeffizienten .....	47
3.6	Textausgaben zum Regressionsmodell .....	48
3.7	Pfad- bzw. Strukturgleichungsanalyse mit Amos.....	49
3.8	Kausalitätstheoretische Annahmen bei Strukturgleichungsmodellen .....	52
<b>4</b>	<b>KONFIRMATORISCHE FAKTORENANALYSE.....</b>	<b>56</b>
<b>5</b>	<b>GRUNDLAGEN UND VERTIEFUNGEN ZUR STRUKTURGLEICHUNGSANALYSE .....</b>	<b>62</b>
5.1	Einige Grundlagen aus der Wahrscheinlichkeitstheorie .....	62
5.1.1	Erwartungswert.....	62
5.1.2	Varianz.....	62
5.1.3	Kovarianz .....	63
5.1.4	Kovarianzmatrix eines Zufallsvektors .....	63
5.1.5	Korrelation .....	63
5.1.6	Kovarianzen von Linearkombinationen aus Zufallsvariablen.....	63
5.1.7	Stichprobenschätzer .....	64
5.2	Erklärung einer Kovarianzmatrix durch ein Strukturgleichungsmodell.....	64
5.3	Identifikation der Parameter .....	66
5.4	Schätzung der Parameter durch Minimierung einer Fit-Funktion .....	69
5.4.1	Unweighted Least Squares (ULS) .....	69
5.4.2	Maximum Likelihood (ML) .....	71
5.4.3	Hinweise zur praktischen Anwendung der Schätzverfahren.....	73
5.5	Flexible Hypothesentests durch den Vergleich von geschachtelten Modellen .....	74
5.6	Gütekriterien zur Modellbeurteilung .....	76
5.6.1	CMIN und die zugehörige Überschreitungswahrscheinlichkeit P .....	77
5.6.2	Relation aus CMIN und DF.....	77
5.6.3	Root Mean Square Error of Approximation (RMSEA).....	77
5.6.4	SRMR.....	79
5.6.5	Informationskriterien zum Vergleich von nicht-geschachtelten Modellen .....	80
<b>6</b>	<b>EIN NONREKURSIVES MODELL MIT LATENTEN VARIABLEN.....</b>	<b>82</b>
6.1	Messfehler und methodologische Gegenmaßnahmen .....	82
6.2	Ein nonrekursives Strukturmodell .....	83
6.3	Weitere Werkzeuge für die Erstellung von Pfaddiagrammen .....	84
6.4	Methodenstudie zum Nutzen von latenten Variablen .....	86
<b>7</b>	<b>EXPLORATIVE TECHNIKEN.....</b>	<b>89</b>
7.1	Modifikationsindikatoren.....	89
7.2	Spezifikationssuche .....	92

<b>8</b>	<b>DIE SIMULTANE ANALYSE MEHRERER GRUPPEN .....</b>	<b>97</b>
<b>9</b>	<b>BOOTSTRAPPING.....</b>	<b>104</b>
9.1	Prinzip und Nutzen .....	104
9.2	Einsatz bei verletzter Normalitätsannahme.....	105
9.3	Bootstrap-Signifikanztests für indirekte Effekte.....	109
9.4	Modellgültigkeitstest nach Bollen & Stine .....	109
9.5	Parametrisches Bootstrapping .....	110
<b>10</b>	<b>LITERATUR .....</b>	<b>112</b>
<b>11</b>	<b>STICHWORTVERZEICHNIS .....</b>	<b>114</b>

Herausgeber: Zentrum für Informations-, Medien- und Kommunikationstechnologie (ZIMK)  
an der Universität Trier  
Universitätsring 15  
D-54286 Trier  
WWW: <http://www.uni-trier.de/index.php?id=518>  
E-Mail: [zimk@uni-trier.de](mailto:zimk@uni-trier.de)  
Tel.: (0651) 201-3417, Fax.: (0651) 3921

Copyright © 2015; ZIMK  
Autor : Bernhard Baltes-Götz (E-Mail : [baltes@uni-trier.de](mailto:baltes@uni-trier.de))

## Vorwort

Das Manuskript entstand als Begleitlektüre zum Amos-Einführungskurs am Rechenzentrum der Universität Trier, was bei gelegentlichen Hinweisen zur Arbeitsumgebung auf den vom Rechenzentrum betreuten Pool-PCs deutlich wird.

Es können nur elementare Aspekte der Strukturgleichungsmethodologie angesprochen werden. Eine ausführliche Behandlung finden Sie z.B. in Bollen (1989), Kline (2005), Loehlin (2004) oder Schumacker & Lomax (2004). Im Internet sind weitere (meist kürzer gehaltene) Darstellungen der Strukturgleichungstechnologie zu finden, z.B. die ZIMK-Dokumentation zum Strukturgleichungsanalyseprogramm LISREL 7 (Baltes-Götz 1994).

Das Manuskript orientiert sich teilweise am sehr empfehlenswerten Handbuch zu Amos 18 (Arbuckle 2009a), das mit dem Programm ausgeliefert wird.

Von den zahlreichen Themen, die in diesem Manuskript überhaupt nicht oder zu knapp behandelt werden, sind besonders zu erwähnen:

- FIML-Schätzung (*Full Information Maximum Likelihood*) bei fehlenden Werten (siehe z.B. Baltes-Götz 2008a)
- Modellierung von Mittelwerten
- Modelle mit Interaktionen (siehe z.B. Baltes-Götz 2008b)
- Analyse latenter Klassen
- Behandlung zensierter und ordinaler Indikatoren über Bayes-Schätzverfahren

Die aktuelle Version des Manuskripts ist als PDF-Dokument zusammen mit allen im Kurs benutzen Dateien auf dem Webserver der Universität Trier von der Startseite (<http://www.uni-trier.de/>) ausgehend folgendermaßen zu finden:

[Rechenzentrum](#) > [Studierende](#) > [EDV-Dokumentationen](#) >  
[Statistik](#) > [Analyse von Strukturgleichungsmodellen mit Amos](#)

Leider sind in diesem Manuskript einige Teile mit heißer Nadel gestrickt, so dass Unzulänglichkeiten zu befürchten und entsprechende Hinweise der Leser(innen) zu erhoffen sind (z.B. an die Mail-Adresse [baltes@uni-trier.de](mailto:baltes@uni-trier.de)).

Trier, im Oktober 2010

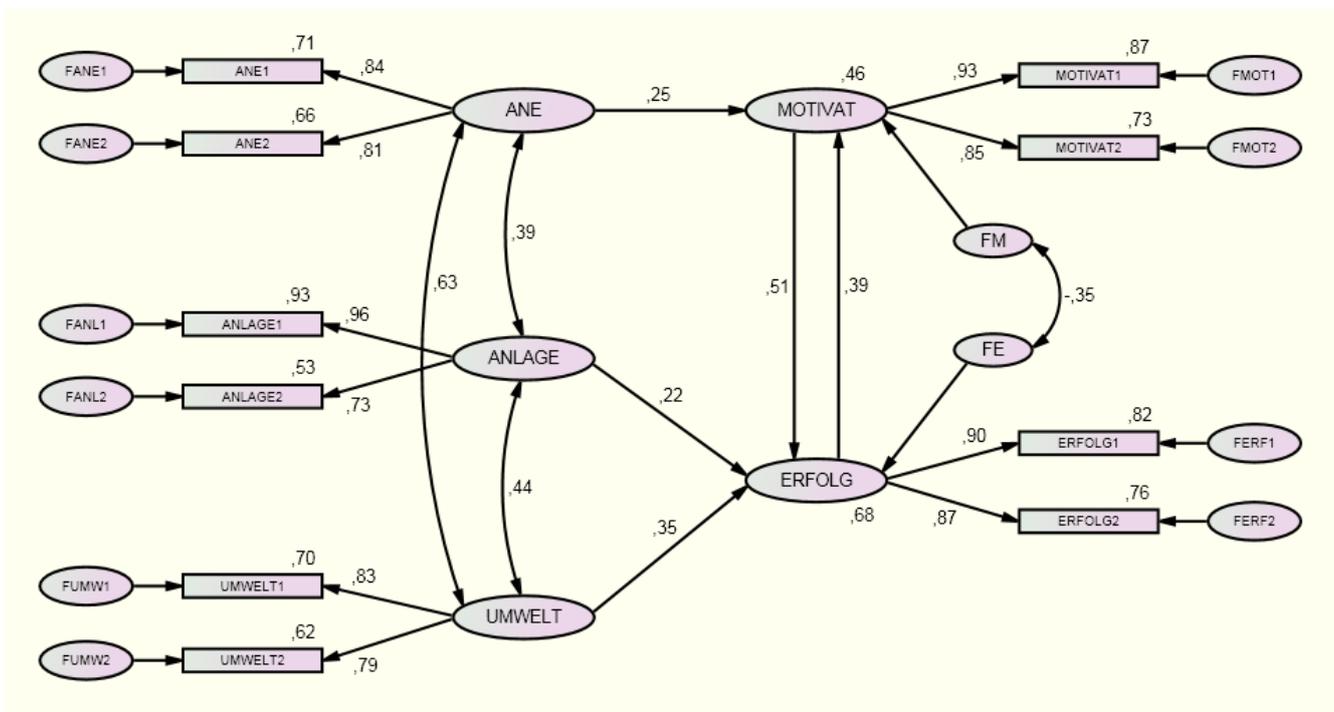
Bernhard Baltes-Götz

# 1 Einleitung

## 1.1 Strukturgleichungsmodelle mit latenten Variablen

Amos (Analysis of MOment Structures) ist ein modernes Werkzeug zur Analyse von Strukturgleichungsmodellen mit latenten Variablen. Diesen sehr flexiblen Analyseansatz, häufig auch als *Kovarianzstrukturanalyse* bezeichnet, kann man in erster Näherung als Kombination von Regressions- und Faktorenanalyse begreifen. Viele bekannte statistische Auswertungsverfahren können als Spezialfall der Strukturgleichungsanalyse aufgefasst werden (z.B. Regressions-, Varianz-, und Kovarianzanalyse). Darüber hinaus stellt die Möglichkeit zur Formulierung von Struktur- bzw. Regressionsgleichungen auf der Ebene *latenter* Variablen eine sehr attraktive Modellierungsoption dar, die ursprünglich vor allem von Karl Jöreskog entwickelt wurde (siehe z.B. Jöreskog & Sörbom 1989).

Das folgende Beispiel mit einer reziproken Abhängigkeitsbeziehung auf der Ebene latenter Variablen demonstriert, welche Flexibilität Amos bei der Formulierung von Strukturgleichungsmodellen bietet:



Im Vergleich zur linearen Regressionsanalyse (z.B. mit der SPSS-Prozedur REGRESSION realisierbar) bietet Amos vor allem die folgenden methodologischen Vorteile:

- **Keine Beschränkung auf rekursive Modelle**

Ein aus mehreren Gleichungen bestehendes Modell heißt **rekursiv**, eine die folgenden Bedingungen erfüllt sind:

- Die endogenen Variablen können so nummeriert werden, dass endogene Variablen keine kausalen Effekte auf andere endogene Variablen mit niedrigeren Nummern ausüben. Ist diese Bedingung verletzt, treten kausale Sequenzen auf, welche direkt (z.B.:  $Y_1$  beeinflusst  $Y_2$  und  $Y_2$  beeinflusst  $Y_1$ , siehe obiges Pfaddiagramm) oder indirekt (z.B.  $Y_1$  beeinflusst  $Y_2$ ,  $Y_2$  beeinflusst  $Y_3$  und  $Y_3$  beeinflusst  $Y_1$ ) zum Ausgangsort zurückführen.<sup>1</sup>

<sup>1</sup> Der Begriff *rekursiv* verursacht gelegentlich Missverständnisse, weil ein *Zurücklaufen* (lateinisch: *recurrere*) doch ausgerechnet bei den als *nonrekursiv* bezeichneten Modellen auftritt. Man denke jedoch daran, dass in der Logik und in der Mathematik eine rekursive Definition vorschreibt, wie spätere Ausdrücke aus unmittelbar vorausgehenden bestimmt werden können. Ist jedoch in einem Strukturgleichungsmodell eine Variable  $Y_2$  nur unter Verwendung der Variablen  $Y_1$  zu ermitteln,

- Die exogenen Variablen sind unkorreliert mit den Residuen.
- Die Residuen zu den Kriteriumsvariablen aus verschiedenen Gleichungen sind unkorreliert, sofern direkte Effekte zwischen den Kriteriumsvariablen bestehen.

Ist eine Bedingung verletzt, spricht man von einem **nonrekursiven** Modell. In diesem Fall ist eine *simultane* Analyse der Gleichungen erforderlich. Traditionelle Regressionsprozeduren mit einer separaten OLS-Schätzung (Ordinary Least Squares) für jede einzelne Gleichung liefern verzerrte Parameterschätzungen.

- **Latente Variablen**

Sind für theoretische Konzepte nur fehlerbehaftete Indikatoren vorhanden, erhält man geminderte bzw. verzerrte Effektschätzungen, wenn bei einer Analyse jedes Konstrukt durch einen *einzelnen* Indikator operationalisiert wird. Oft sind jedoch pro Konstrukt *mehrere* Indikatoren verfügbar. Amos kann diese nutzen, um die Konstrukte in der Beobachtungsbasis zu verankern und in das Gleichungssystem einzubeziehen. So lassen sich auf der Ebene latenter Variablen Effekte ohne messfehlerbedingte Verzerrungen schätzen.

## 1.2 Analyseoptionen in Amos 18

Die in diesem Kurs behandelte Amos-Version 18 bietet ein umfangreiches Spektrum von Analysetechniken, z.B. die Schätzmethoden

- Maximum Likelihood (ML), auch bei fehlenden Daten (FIML)
- generalisierte (GLS), ungewichtete (ULS) und skalenfreie (SLS) kleinste Quadrate
- asymptotisch verteilungsfreie Schätzer (ADF)
- Bootstrapping
- Bayes-Schätzung mit dem MCMC-Verfahren (Markoff Chain Monte Carlo)

Über die simultane Analyse von Daten aus mehreren Populationen lassen sich z.B. Interaktionseffekte mit kategorialen Moderatoren untersuchen.

Amos bietet etliche Hilfsmittel zur Modellsuche, so dass die traditionell eher für *konfirmatorische* Forschungsbemühungen konzipierte Strukturgleichungsmethodologie auch in explorativer Manier zu nutzen ist (z.B. für eine explorative Faktorenanalyse).

Bei aller Methodenvielfalt bemüht sich Amos um Bedienungsfreundlichkeit, wozu nicht zuletzt die Modellspezifikation durch den Aufbau eines Pfaddiagramms beiträgt. Dank exzellenter Gestaltungsmöglichkeiten taugen Amos-Pfaddiagramme nicht nur zur Modellspezifikation, sondern machen auch in Publikationen eine gute Figur.

Amos wird zwar von der Firma SPSS vertrieben, ist aber ein eigenständig (Windows-) Programm, das selbstverständlich SPSS-Datendateien (Extension: **.sav**) problemlos verarbeiten kann.

Das Programm bietet neben seiner grafikorientierten Benutzeroberfläche auch eine perfekte Unterstützung der .NET - Plattform, so dass Arbeitsabläufe mit .NET - Programmiersprachen wie VB.NET oder C# automatisiert werden können.

---

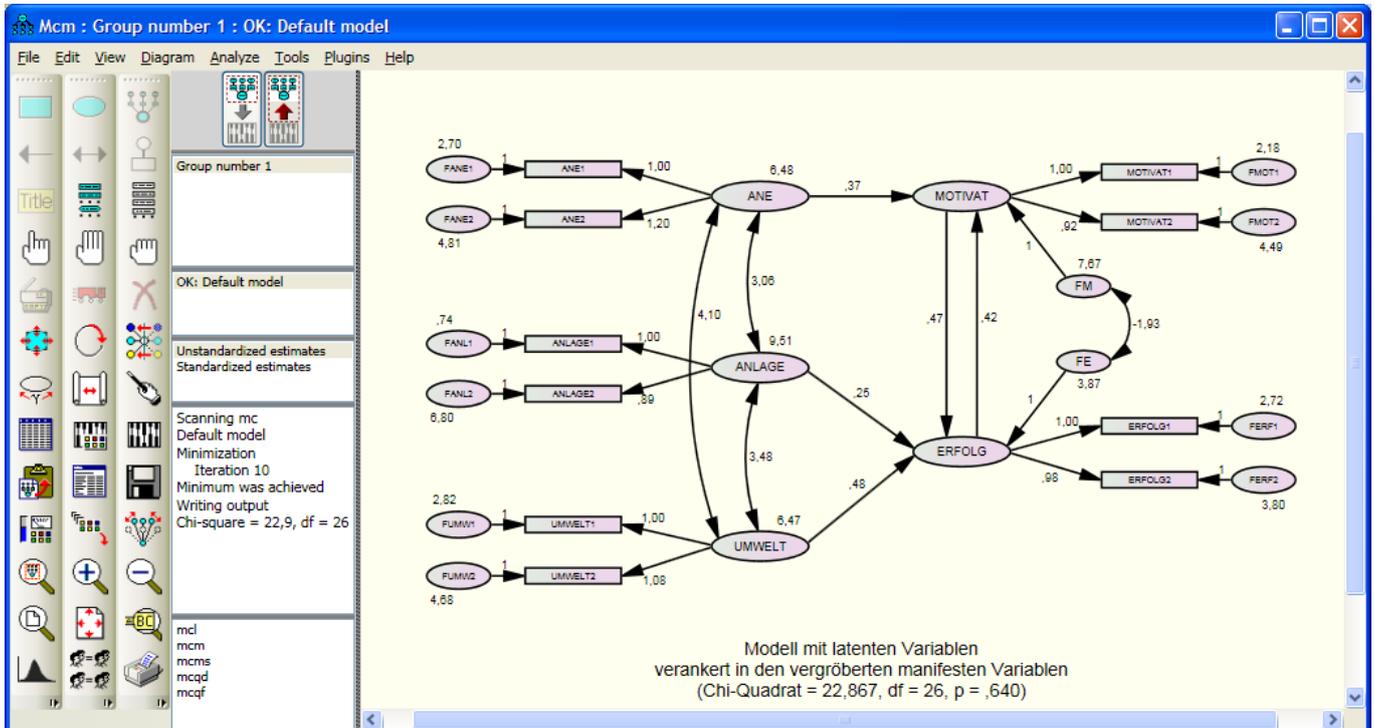
andererseits aber  $Y_1$  nur bei Bekanntheit von  $Y_2$  bestimmbar, dann ist offenbar ein systematischer Aufbau im Sinne der rekursiven Definition nicht möglich. Ein solches System wird daher zu Recht als *nonrekursiv* bezeichnet.

## 2 Aller Anfang ist leicht

Auf den vom Rechenzentrum betreuten Pool-PCs startet man mit

**Start > Alle Programme > Wissenschaftliche Programme >  
SPSS > Amos > Amos 18 Graphics**

die grafikorientierte Amos-Oberfläche. Diese bietet u.a. eine Zeichenfläche zur Aufnahme des Pfaddiagramms und (am linken Rand) drei nebeneinander gepackte streifenförmige Werkzeugpaletten:



Mit dem hier abgebildeten, recht komplexen, Modell werden wir uns erst später beschäftigen, wenn die technischen Aspekte der Amos-Bedienung geklärt sind.

Bei den ersten Übungsbeispielen kommen Daten aus einer Studie von Attig (1983) zum Einsatz.<sup>1</sup> Hier wurden vor und nach einem Gedächtnistraining bei 40 jungen Probanden drei Leistungsmaße erhoben:

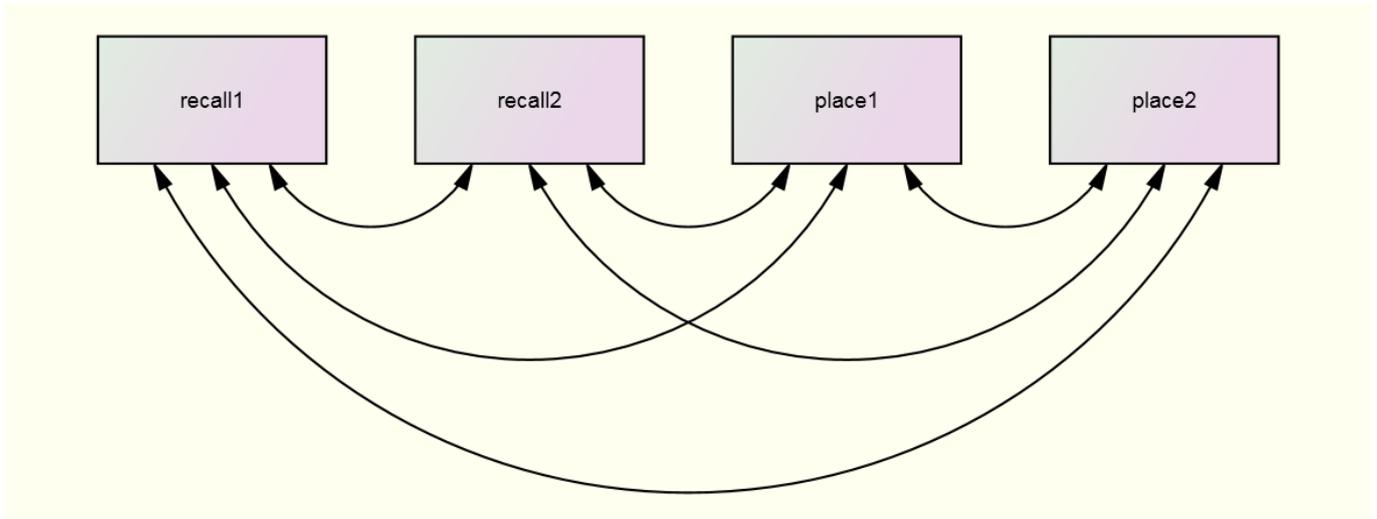
- Erinnerungsleistung (**recall1** bzw. **recall2**)
- Erinnerungsleistung nach unterstützenden Hinweisen (**cued1** bzw. **cued2**)
- Erinnerung der richtigen Präsentationsreihenfolge (**place1** bzw. **place2**)

Außerdem wurden noch weitere Merkmale erfasst: Leistung in einem Vokabeltest, Alter, Geschlecht, Ausbildungsniveau. Zunächst sollen die Varianzen und Kovarianzen der Variablen **recall1**, **recall2**, **place1** und **place2** geschätzt werden.

<sup>1</sup> Wo die mit Amos ausgelieferten Beispieldaten zu finden sind, ist gleich zu erfahren. In Abschnitt 8 kommen bei einem Gruppenvergleich auch noch analoge Daten einer älteren Stichprobe ins Spiel.

## 2.1 Pfaddiagramm zeichnen

Wir wollen das folgende Pfaddiagramm zeichnen:



Erstellen Sie zunächst vier Rechtecke für die zu untersuchenden manifesten Variablen auf der Amos-Zeichenfläche. Das erforderliche Werkzeug erhalten Sie mit

 oder **Diagram > Draw Observed** oder **F3**

Am besten erzeugen Sie das erste Rechteck, korrigieren dann nötigenfalls dessen Position mit

 oder **Edit > Move** oder **Strg+M**

und optimieren die Form (Seitenlängen) mit

 oder **Edit > Shape of Object**

Nun können Sie nach

 oder **Edit > Duplicate**

mit Klicken & Ziehen drei Kopien des Rechtecks anfertigen. Wenn Sie dabei die  $\uparrow$ -Taste gedrückt halten, wird (in Abhängigkeit von der anfänglichen Bewegungsrichtung) die horizontale oder die vertikale Koordinate beibehalten. So lassen sich z.B. leicht vier identische Rechtecke erzeugen, die exakt auf gleicher Höhe liegen.

Kehren Sie in Denk- und Ruhephasen zur Vermeidung unerwünschter Mauseffekte mit

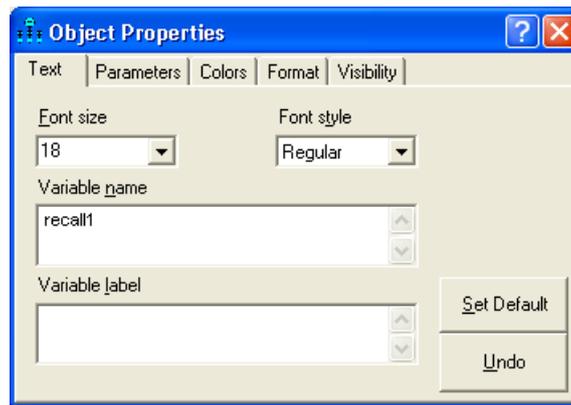
 oder **Edit > Select** oder **F2**

zum neutralen Markierungswerkzeug zurück. Selbstverständlich können Sie Missgriffe mit

 oder **Edit > Undo** oder **Strg+Z**

rückgängig machen (auch mehrstufig).

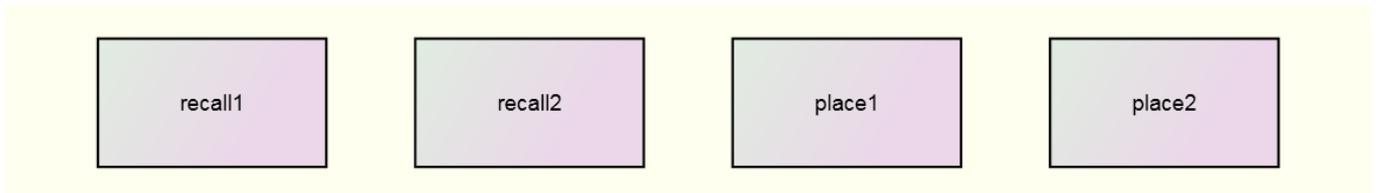
Wir verknüpfen nun jedes Rechteck mit einer manifesten Variablen, indem wir deren Namen in der **Object Properties** - Dialogbox eintragen, die via Kontextmenü-Eintrag geöffnet wird, z.B.:



Bei den Variablennamen in Amos sollte man sich an die SPSS-Namensregeln halten (max. 64 Zeichen, Buchstabe am Anfang, keine Leer- und Sonderzeichen, keine Schlüsselwörter der SPSS-Syntax). Momentan vereinbaren wir Namen für *beobachtete* Variablen, die aus einer Datei importiert werden, so dass die dort vorhandenen Variablennamen zu verwenden sind. Wenn dabei ein unschöner Name akzeptiert werden muss, kann dieser in Amos durch ein **Variable label** ergänzt werden, das im Pfaddiagramm an Stelle des Namens angezeigt wird.

Markieren Sie bei geöffneter **Object Properties** - Dialogbox nacheinander die übrigen Rechtecke, um auch diese mit Variablennamen (und bei Bedarf auch mit **Variable labels**) zu versorgen. Beenden Sie schließlich die **Object Properties** - Dialogbox über deren Schließfeld  am rechten Rand der Titelzeile.

Momentan beschreibt unser Pfaddiagramm *unkorrelierte* Variablen:



Weil unser Modell aber Kovarianzen erlaubt, müssen wir diese als Doppelpfeile in das Pfaddiagramm eintragen. Das erforderliche Werkzeug erhalten Sie mit

 oder **Diagram > Draw Covariance** oder **F6**

### Tipps:

- Um einen *unteren* Bogen zu erhalten, müssen Sie von *rechts nach links* zeichnen.
- Mit dem  -Werkzeug können Sie die Biegung der Pfeile beeinflussen.
- Mit dem  -Werkzeug können Sie die Pfeilspitzen verschieben.
- Wenn Sie mit dem Zauberstab  auf das Rechteck einer Variablen klicken, werden alle Verbindungspunkte normiert, was in der Regel zu einem attraktiven Ergebnis führt.

Nun sollte Ihr Pfaddiagramm ungefähr die eingangs abgebildete Gestalt haben.

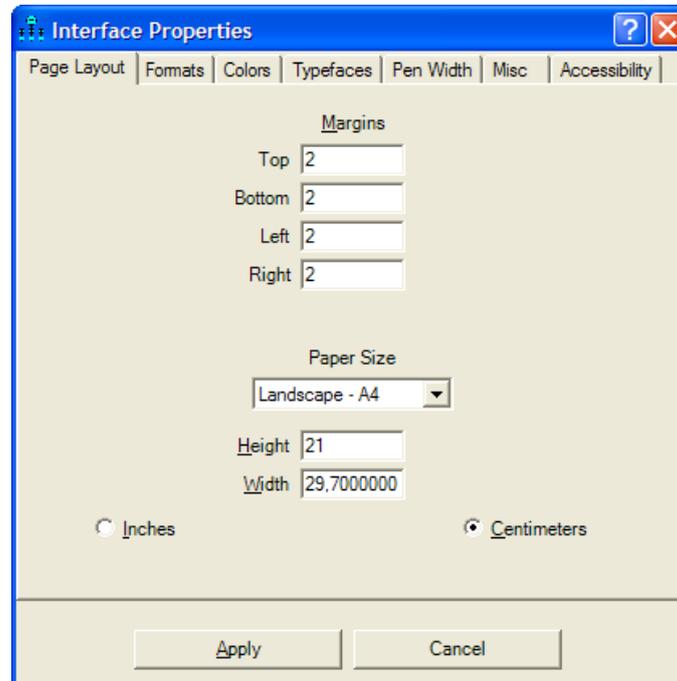
## 2.2 Allgemeine Bedienungsmerkmale

### 2.2.1 Format der Zeichenfläche ändern und Pfaddiagramm neu einpassen

Amos startet mit dem Seitenformat **Letter** in der Orientierung **Portrait** und einem allseitigen Rand von 1 Zoll (2,54 cm). Nach dem Menübefehl

**View > Interface Properties**

sind auf dem Registerblatt **Page Layout** der Dialogbox **Interface Properties**



alternative Einstellungen möglich:

- Frei wählbare Randbreiten
- **Landscape**-Orientierung statt **Portrait**
- Alternative bzw. individuelle Seitengröße

Man kann ein Standardformat wählen oder ein individuelles Format festlegen. Bei **Height** und **Width** akzeptiert Amos Werte von 0 bis 14 Zoll bzw. von 0 bis 35,56 cm.

Vorgenommene Änderungen werden mit **Apply** übernommen oder mit **Cancel** verworfen. Zum Beenden der Dialogbox dient das Schließkreuz  in der Titelzeile.

Nach einer Änderung des Seitenformats, der Seitenorientierung oder der Ränder kann man das Pfaddiagramm mit



oder **Edit >Fit to Page** oder **Strg+F**

neu einpassen lassen.

### 2.2.2 Objekte markieren

Sie haben folgende Möglichkeiten, Objekte auf der Zeichenfläche für die weitere Bearbeitung zu markieren:

- Einzelne Objekte kumulierend markieren  
Schalten Sie mit

 oder **Edit > Select** oder **F2**

den Markierungsmodus ein, und klicken Sie nacheinander auf die gewünschten Objekte. Eine vorhandene Markierung wird durch einen erneuten Mausklick aufgehoben.

- Alle Objekte markieren

 oder **Edit > Select All**

- Alle Markierungen aufheben

 oder **Edit > Deselect All**

Sind in unserem Beispiel alle vier Rechtecke markiert, kann man mit dem Menübefehl

**Edit > Space Horizontally**

für *horizontal* gleiche Abstände sorgen und bei Bedarf anschließend mit dem Zauberstab  die Ansatzpunkte der Kovarianz-Doppelpfeile wieder optimieren.

### 2.2.3 Objekte löschen

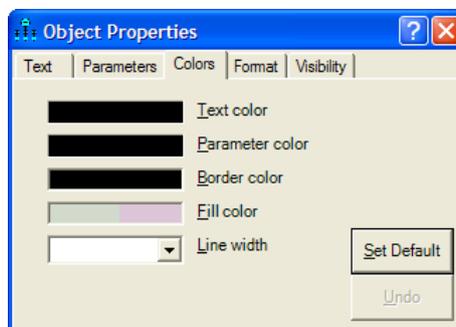
Mit

 oder **Edit > Erase** oder **Entf**

wird der Mauszeiger zum Stornierungswerkzeug, das jedes angeklickte Objekt entfernt.

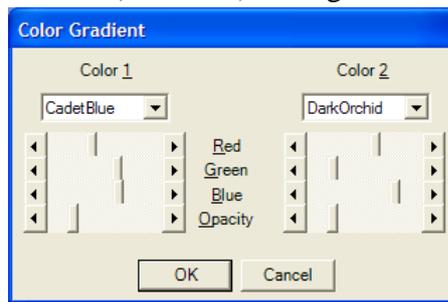
### 2.2.4 Objekte gestalten

Speziell für Rechtecke und Ellipsen bietet Amos einige Gestaltungsmöglichkeiten über die Registerkarte **Colors** der Dialogbox **Object Properties**:

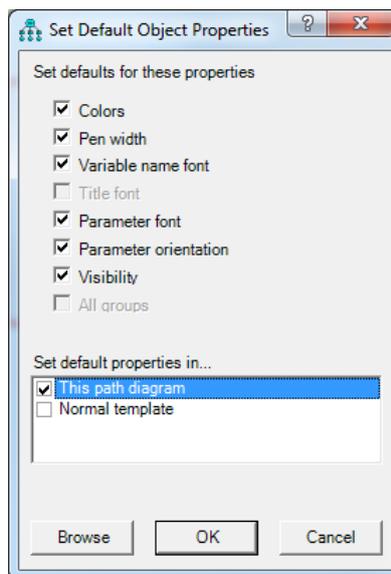


Hier lässt sich z.B. der voreingestellte Farbverlauf für den Hintergrund (siehe obige Beispiele) ändern oder abschalten:

- Mausklick auf die Zelle **Fill color**
- Im folgenden Dialog die beiden Farben (identisch) festlegen:



Über den Schalter **Set Default** in der Dialogbox **Object Properties** lässt sich die Voreinstellung für zukünftige Objekte im selben Projekt festlegen:



Eine Festlegung mit Gültigkeit für alle neuen Projekte (**Normal Template**) setzt Schreibrechte für die allgemeine Vorlagendatei voraus, bei einer typischen Amos-Installation also für

**C:\Program Files (x86)\SPSSInc\Amos18\Templates\English\Normal.amt**

### 2.2.5 Bearbeitungsschritte rückgängig machen oder wiederherstellen

Mit



oder **Edit > Undo** oder **Strg+Z**

können Sie bis zu vier Bearbeitungsschritte rückgängig machen. Mit



oder **Edit > Redo** oder **Strg+Y**

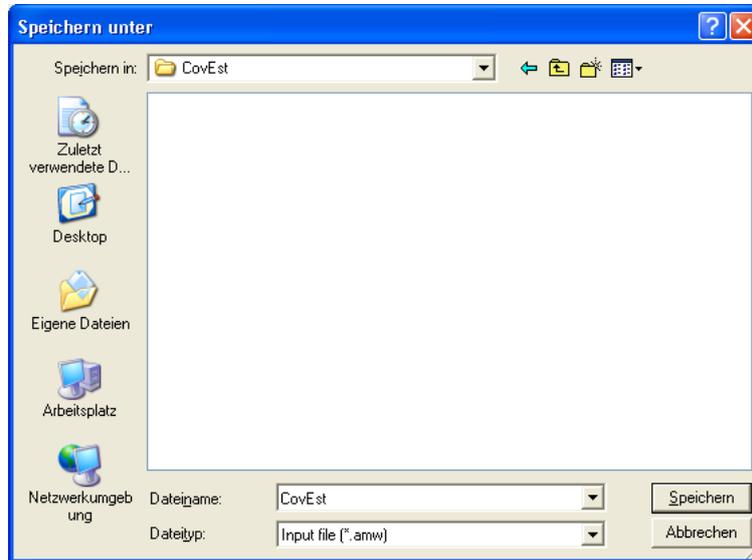
können Sie die zuletzt zurückgenommenen Bearbeitungsschritte wiederherstellen.

## 2.2.6 Speichern, Sicherheitskopien

Beim Speichern eines Projekts mit

**File > Save (As)** oder **Strg+S**

wird eine **Projektdatei** mit der Extension **.amw** angelegt bzw. aktualisiert, z.B.:



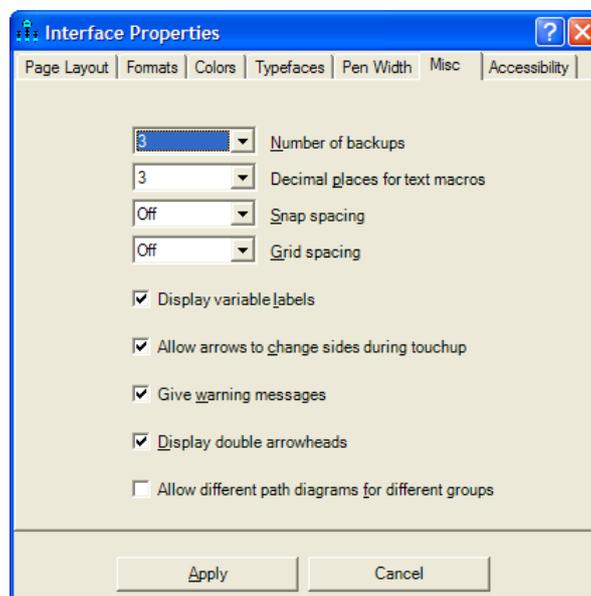
Während der Arbeit an einem Projekt entstehen noch weitere Dateien, die bei Gelegenheit vorgestellt werden. Alle Dateien zu einem Pfaddiagramm bzw. Modell haben denselben Namensstamm, in unserem Beispiel also **CovEst**. Diese Bezeichnung eignet sich auch für den Ordner, den man der Übersichtlichkeit halber für jedes Projekt anlegen sollte.

Wenn Sie später den Amos-Grafikmodus erneut starten, wird die zuletzt bearbeitete Projektdatei automatisch geöffnet. Natürlich kann man ein vorhandenes Projekt mit **File > Open** auch explizit öffnen.

Amos pflegt Sicherheitskopien mit älteren Versionen der Projektdatei, wobei deren Anzahl (Voreinstellung = 2) über den Menübefehl

**View > Interface Properties > Misc**

eingestellt werden kann, z.B.:



In unserem Beispiel heißen die Sicherheitskopien **CovEst.bk1**, **CovEst.bk2**, usw.

### 2.2.7 Pfaddiagramm in andere Anwendungen übertragen

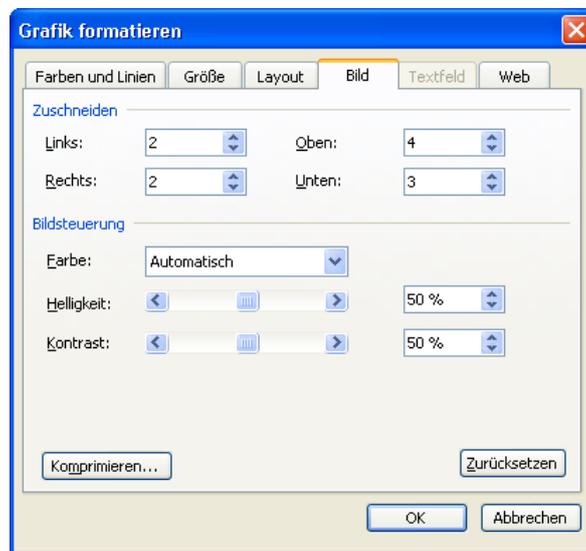
Sie können ein Amos-Pfaddiagramm ganz einfach über die Windows-Zwischenablage in andere Anwendungen übertragen, z.B. in Microsoft Word©

in Amos: **Edit > Copy (to clipboard)** oder **Strg+C**

in Word: **Bearbeiten > Einfügen** oder **Strg+V**

Auf diese Weise sind die obigen Pfaddiagramme in das vorliegende Manuskript gelangt.

Da Amos stets ein komplettes und oft nur mäßig gefülltes „Blatt“ überträgt, müssen Sie im Zielprogramm die eingefügte Grafik eventuell *beschneiden*, in Word 2003 z.B. mit Hilfe der Dialogbox **Grafik formatieren** (zu öffnen über den Menübefehl **Format > Grafik** oder per Kontextmenü zur Grafik):



### 2.3 Daten einlesen

Um unser Modell schätzen zu können, benötigt Amos noch Daten. Es werden zahlreiche Dateiformate unterstützt:

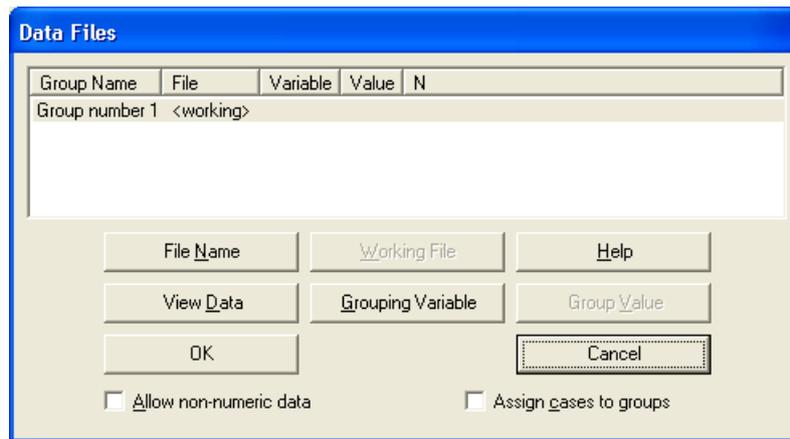
- dBase
- Microsoft Excel
- Microsoft Foxpro
- Lotus
- Microsoft Access
- SPSS
- Textdateien

Wir wollen die zum ersten Beispiel gehörigen Daten aus der SPSS-Datei **Attg\_yng.sav** laden, die sich im **Examples**-Unterverzeichnis zum Amos – Programmordner befindet. Nach

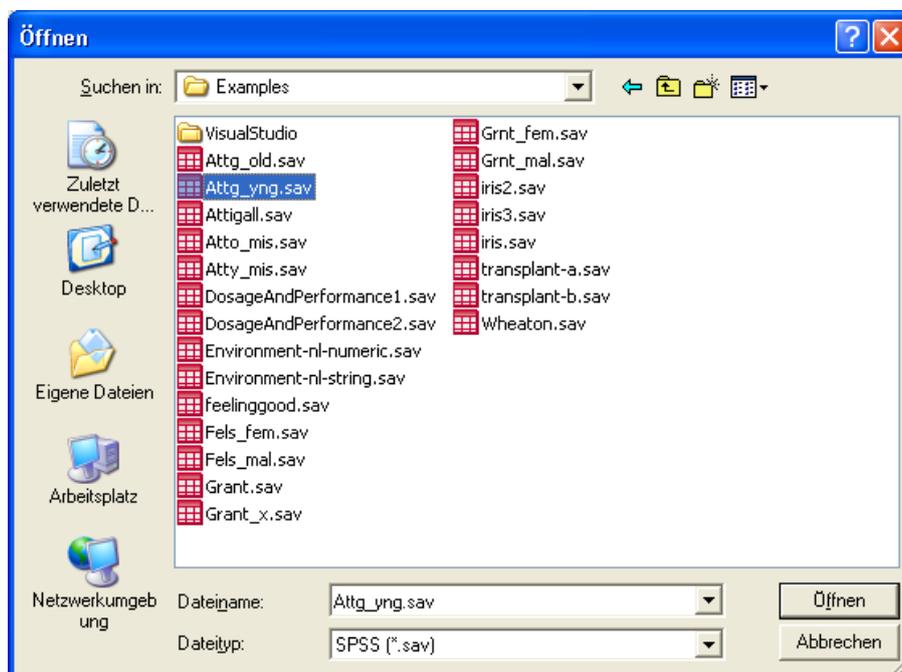


oder **File > Data Files** oder **Strg+D**

erscheint die Dialogbox **Data Files**:



Deren Schalter **File Name** öffnet schließlich den Standarddialog zur Dateiauswahl (hier unter Windows XP):



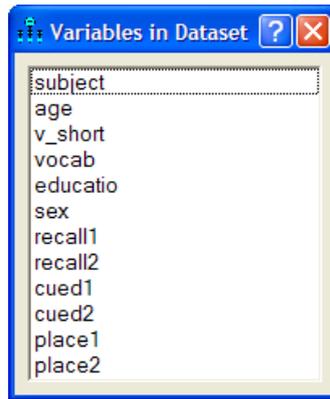
Nach erfolgreicher Dateizuordnung erlaubt die **Data Files** - Dialogbox mit dem Schalter **View Data** einen Blick auf die Daten, wobei in unserem Fall die Anzeige von SPSS selbst übernommen wird (, falls das Programm auf dem aktuellen Rechner installiert ist).

Wir schließen die **Data Files** – Dialogbox mit **OK** und besorgen uns dann mit



oder **View > Variables in Dataset** oder **Strg+⇧+D**

eine Liste mit den Variablen der geöffneten Datendatei:



Jetzt ist es möglich, im Pfaddiagramm ein Rechteck zu benennen, indem man aus der Liste einen Variablennamen per Maus packt und auf dieses Rechteck zieht. Wir hätten also durch frühzeitiges Zuordnen der Datendatei lästige Tipparbeit vermeiden können (siehe Abschnitt 2.1).

## 2.4 Modellparameter schätzen

Lassen Sie mit

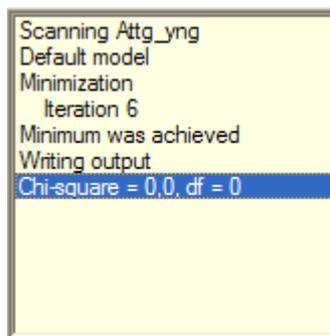


oder **Analyse > Calculate Estimates** oder **Strg+F9**

die Modellparameter schätzen.

Falls bislang noch keine Projektdatei (mit der Namenserweiterung **.amw**) angelegt worden ist (siehe Abschnitt 2.2.6), fordert Amos nun dazu auf.

Während Amos (in unserem Beispiel nur sehr kurz) mit der Parameterschätzung beschäftigt ist, werden wir über den Bearbeitungsfortschritt informiert.



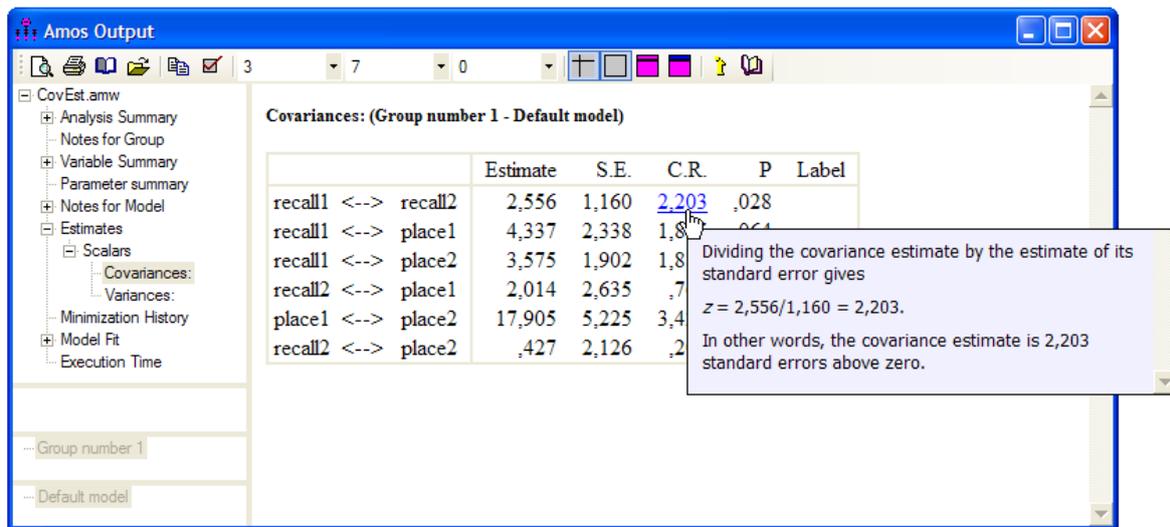
### 2.4.1 Textergebnisse

Amos erzeugt eine Ergebnisdatei mit der Namenserweiterung **AmosOutput**, die nach



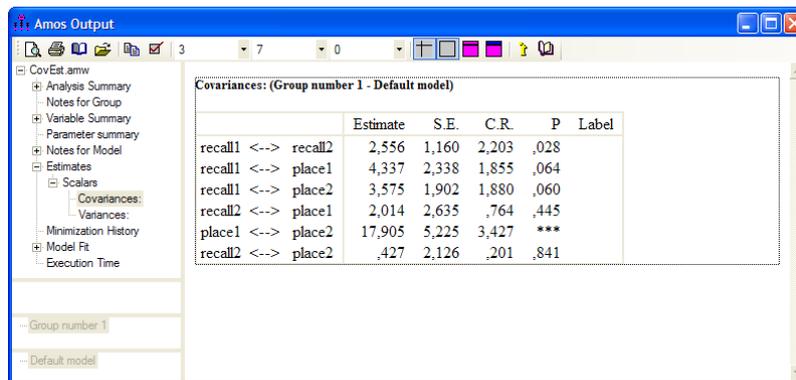
oder **View > Text Output** oder **F10**

von einem Hilfsprogramm angezeigt wird:



Sie besteht aus mehreren Tabellen, die bequem per Navigationszone zu erreichen sind. Zu vielen Tabellenbestandteilen können per Mausklick ausführliche Erläuterungen abgerufen werden (siehe Beispiel).

Eine per Mausklick markierte und daraufhin mit einem dünnen Extrarahmen versehene Tabelle



kann über den Symbolschalter  oder die Tastenkombination **Strg+C** in die Windows-Zwischenablage befördert und anschließend von dort in eine andere Anwendung übernommen werden, z.B.:

	Estimate	S.E.	C.R.	P	Label
recall1 <--> recall2	2,556	1,160	2,203	,028	
recall2 <--> place1	2,014	2,635	,764	,445	
place1 <--> place2	17,905	5,225	3,427	***	
recall2 <--> place2	,427	2,126	,201	,841	
recall1 <--> place1	4,337	2,338	1,855	,064	
recall1 <--> place2	3,575	1,902	1,880	,060	

In der Zielanwendung Microsoft Word© lässt sich die Tabelle beliebig formatieren bzw. modifizieren.

Das Anzeigeprogramm für Textausgaben bietet noch weitere Bedienungsmöglichkeiten:

-  Anzahl der Dezimalstellen wählen
-  Breite der Spalten einstellen
-  Gitterlinien aus- bzw. einschalten
-  Rahmen aus- bzw. einschalten
-  Hintergrundfarbe für die Tabelle wählen
-  Hintergrundfarbe für die Kopfzeile der Tabelle wählen

Man kann das Anzeigeprogramm auch per Amos-Programmgruppe (Verknüpfung **Text Output**) oder per Doppelklick auf eine **AmosOutput**-Datei starten.

### 2.4.2 Erste Ergebnisinterpretationen

Im Abschnitt **Notes for Model > Result** der Ergebnisdatei wird mit dem Eintrag

**Minimum was achieved**

signalisiert, dass die iterative Parameterschätzung gelungen ist.

Dem Abschnitt **Estimates** ist u.a. zu entnehmen, dass Amos per Voreinstellung bei der Schätzung nach dem **Maximum Likelihood (ML) - Prinzip** arbeitet. Wie oben erwähnt, unterstützt Amos auch einige alternative Schätzmethoden (wählbar über **View > Analysis Properties > Estimation**), doch erweist sich die voreingestellte ML-Methode als recht leistungsfähig, auch bei einer mäßigen Verletzung ihrer Voraussetzungen, mit denen wir uns in Abschnitt 2.7.2 beschäftigen werden.

In der Tabelle mit den Kovarianz-Schätzergebnissen (siehe oben) wird zu jedem Schätzwert in der Spalte **S.E.** sein approximativer Standardfehler angegeben, der unter gewissen Voraussetzungen (siehe Abschnitt 2.7.2) die Konstruktion eines approximativen **Vertrauensintervalls** erlaubt. Für die Kovarianz von **recall1** und **recall2** erhalten wir z.B. unter Verwendung des 97,5%-Fraktils der Standardnormalverteilung ( $\approx 1,96$ ) das folgende approximative 95%-Vertrauensintervall:

$$2,556 \pm 1,96 \cdot 1,16 = 2,556 \pm 2,274$$

Die mit **C.R.** (für *Critical Ratio*) überschriebene Spalte enthält für jeden Schätzwert den folgenden Quotienten:

$$\left| \frac{\text{Schätzwert}}{\text{Standardfehler}} \right|$$

Mit Hilfe dieses Quotienten lässt sich unter den in Abschnitt 2.7.2 beschriebenen Verteilungsvoraussetzungen z.B. das folgende (zweiseitige) Testproblem lösen:

$$H_0: \text{Cov}(\text{recall1}, \text{recall2}) = 0 \quad \text{versus} \quad H_1: \text{Cov}(\text{recall1}, \text{recall2}) \neq 0$$

Bei Gültigkeit der  $H_0$  ist der angegebene Quotient eine approximativ, d.h. für  $N \rightarrow \infty$ , standardnormalverteilte Zufallsvariable, so dass man bei folgender Entscheidungsregel approximativ einen Test zum Niveau  $\alpha = 0,05$  erhält:

$$\left| \frac{\text{Schätzwert}}{\text{Standardfehler}} \right| \begin{cases} \leq 1,96 & \Rightarrow H_0 \text{ beibehalten} \\ > 1,96 & \Rightarrow H_0 \text{ verwerfen} \end{cases}$$

In unserem Beispiel wird also die Nullhypothese verworfen (C.R. = 2,203).

Um die Testentscheidung durchzuführen, muss man nicht unbedingt den kritischen Wert für die CR-Statistik kennen, da Amos unter dem Spaltentitel **P** auch die empirischen Überschreitungswahrscheinlichkeiten für den zweiseitigen Test mitteilt. Die zugehörige Variante der Entscheidungsregel des Tests zum Niveau  $\alpha = 0,05$  lautet:

$$p \begin{cases} \geq 0,05 & \Rightarrow H_0 \text{ beibehalten} \\ < 0,05 & \Rightarrow H_0 \text{ verwerfen} \end{cases}$$

Im Beispiel erhalten wir  $p = 0,028$  und kommen zur selben Entscheidung wie beim Vergleich der CR-Statistik mit dem kritischen Wert 1,96.

Amos bietet bei erhöhtem Aufwand ein alternatives Verfahren zum selben Testproblem (Parameterrestriktion und Modellgültigkeitstest), das meist exaktere Ergebnisse liefert (siehe Abschnitt 2.10).

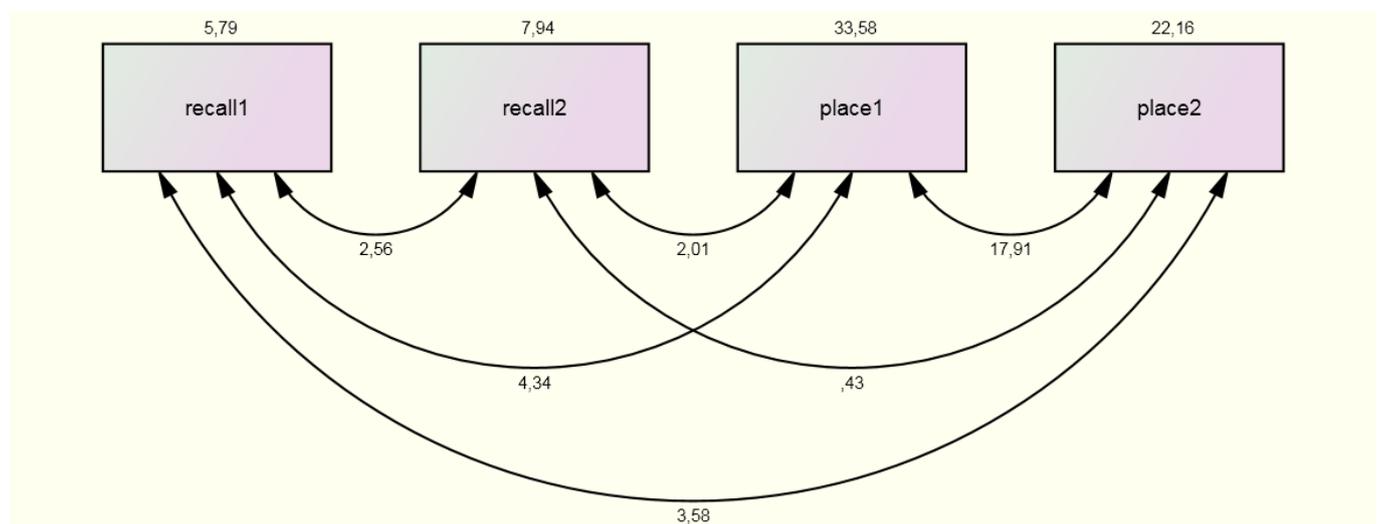
Beide Amos-Testverfahren sind nur *approximativ* gültig, d.h. bei ausreichend großen Stichproben. Für das eben behandelte simple Problem existiert mit dem t-Test für Korrelationen eine bei beliebigen Stichprobenumfängen (jedenfalls ab  $N \geq 3$ ) gültige Alternative, die im Beispiel eine Überschreitungswahrscheinlichkeit von  $p = 0,016$  liefert und zur selben Entscheidung kommt wie der CR-Test.

Beginnend mit Abschnitt 2.7 werden wir flexible Optionen zum Testen von Modellen und zusammengesetzten Hypothesen behandeln, zu denen *keine* exakte Alternative existiert.

### 2.4.3 Das Ausgabe-Pfaddiagramm

Amos kann seine Schätzergebnisse nicht nur in Tabellenform präsentieren, sondern auch in das Pfaddiagramm eintragen. Dazu ist nach erfolgreicher Modellschätzung der Schalter  einzurasten.

Im Beispiel sieht das komplettierte Pfaddiagramm so aus:



Wer im selbst erstellten Pfaddiagramm die Varianzschätzer ebenso zentriert sehen möchte wie die Exemplare in obiger Abbildung, muss nach



oder **Edit > Move Parameter**

das zuständige Werkzeug verwenden:

- Bei hinreichender Nähe zum Mauszeiger wird eine Variable rot angezeigt.
- Dann kann die Parameterposition durch Mausbewegungen bei gedrückter linker Maustaste verändert werden.
- Analog lassen sich auch die übrigen Parameter in einem Amos-Pfaddiagramm verschieben.
- Sind mehrere Variablen (oder auch sonstige „Parameterträger“) markiert, gelingt die synchrone Veränderung der Parameterpositionen.

Bei Anzeige eines Ausgabe-Pfaddiagramms (mit Parameterschätzungen) sind einige Werkzeuge zur Modifikation des Pfaddiagramms *nicht* verfügbar (z.B. ). Um diese Werkzeuge wieder nutzen zu können, muss man erst den Schalter



für den Modellspezifikationsmodus einrasten.

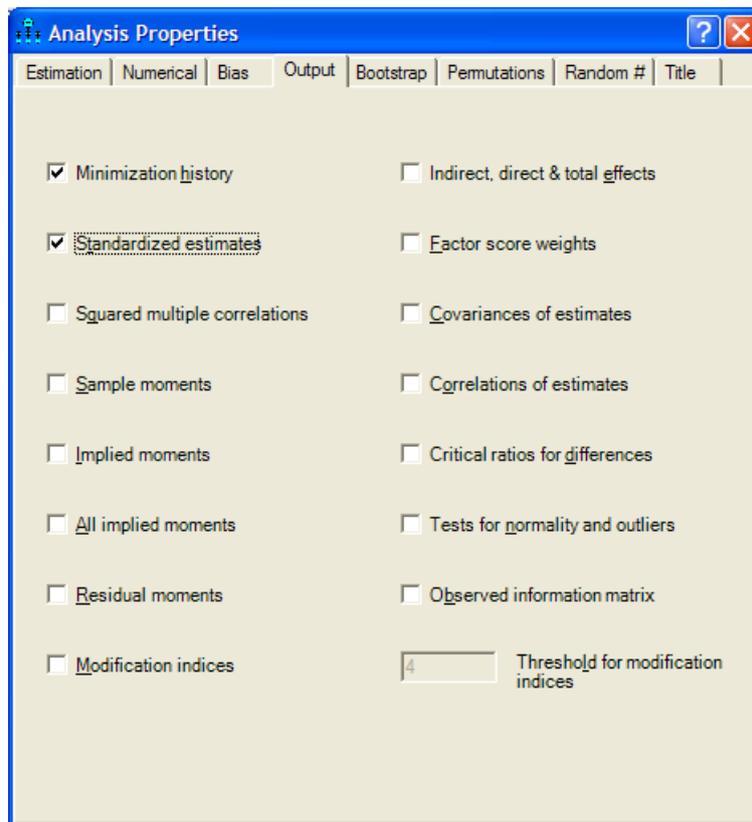
## 2.5 Optionale Ausgaben

Nach



oder **View > Analysis Properties > Output**

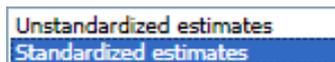
können in der folgenden Dialogbox



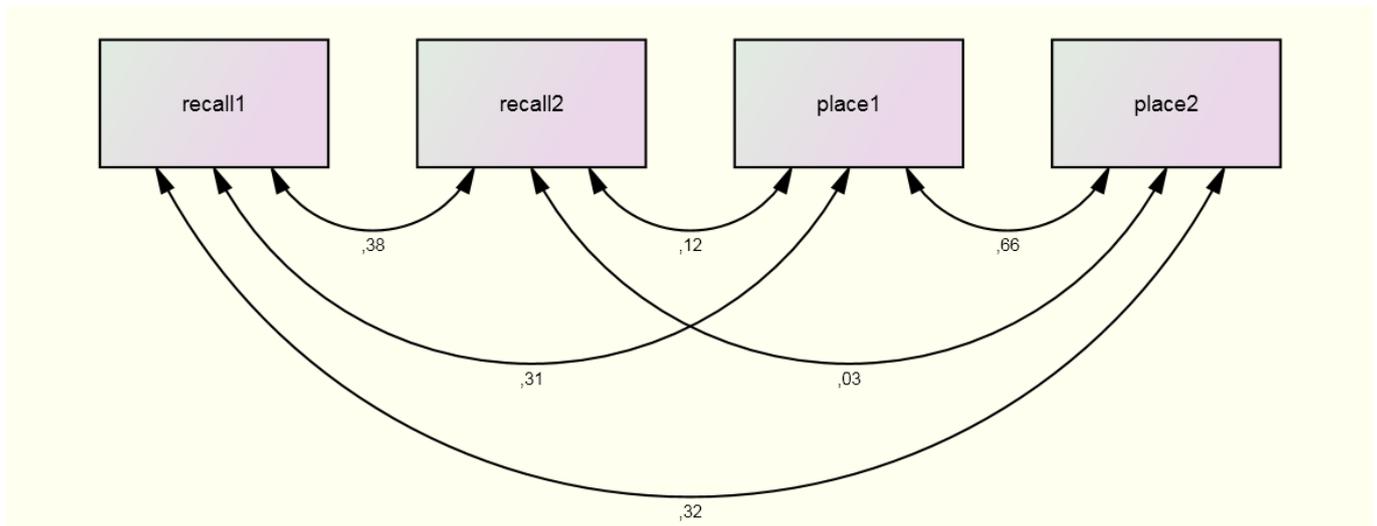
noch zahlreiche weitere Ergebnisse angefordert werden, z.B.:

### Standardized estimates

War diese Option bei der Modellschätzung aktiv, zeigt das Ausgabe-Pfaddiagramm nach dem Markieren der zugehörigen Option:



*standardisierte* Ergebnisse (z.B. Korrelationen statt Kovarianzen):



### Sample moments

Man erhält in der Textausgabe die Kovarianzmatrix und die Korrelationsmatrix der manifesten Variablen. Nach entsprechender Anforderung über das Kontrollkästchen **Means and intercepts** auf der Registerkarte **Estimates** der Dialogbox **Analysis Properties** liefert Amos auch die Mittelwerte der manifesten Variablen. Im Beispiel resultiert u.a. die folgende Tabelle mit den schon aus dem Pfaddiagramm bekannten Kovarianzen:

	place2	place1	recall2	recall1
place2	22,160			
place1	17,905	33,578		
recall2	,427	2,014	7,944	
recall1	3,575	4,338	2,556	5,788

### Implied moments

Man erhält eine Matrix mit den vom Modell implizierten Momenten, im Standardfall also mit implizierten Varianzen und Kovarianzen. Diese implizierten Varianzen und Kovarianzen unterscheiden sich von den beobachteten, sobald ein Modell Restriktionen enthält (siehe unten). In unserem trivialen Beispiel ohne Restriktionen gibt es allerdings (von Rundungsfehlern abgesehen) keine Unterschiede. Weitere Ausgabeoptionen werden bei Gelegenheit besprochen.

## 2.6 Amos in .NET - Programmen verwenden

Amos stellt sein gesamtes Leistungsspektrum über ein Objektmodell zur Verfügung, das in Programmiersprachen für das .NET - Framework (z.B. C#, VB.NET) genutzt werden kann. Es ist wohl im Wesentlichen an die beiden folgenden Einsatzzwecke gedacht:

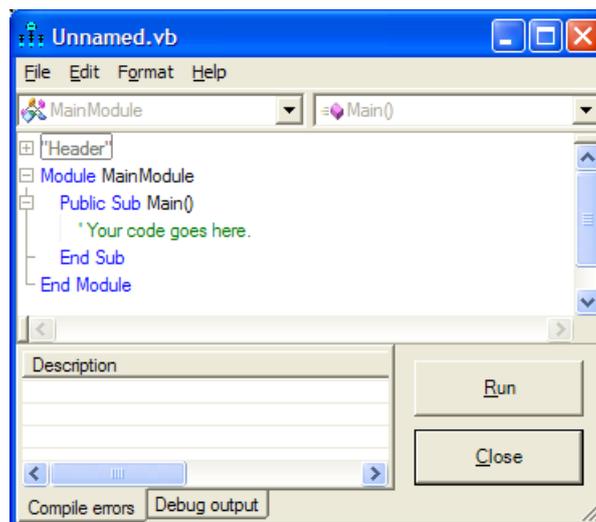
- Modellierung über *Gleichungen* an Stelle der graphischen Elemente (Rechtecke, Pfeile etc.) Dazu dient die Klasse **AmosEngine**, die wir gleich benutzen werden. Programme mit wesentlicher Beteiligung dieser Klasse werden wir später als *AmosEngine-Programme* bezeichnen.
- Automatisierung von Routinearbeiten mit der graphikorientierten Benutzerschnittstelle Dazu dient z.B. die Klasse **PathDiagrammer**, die in vielen vorgefertigten Plugins verwendet wird. So steht z.B. in Amos Graphics über

**Plugins > Draw Covariances**

eine Erweiterung bereit, die zu allen markierten exogenen Variablen die möglichen Kovarianzpfleile zeichnet und anschließend gleich das -Werkzeug für die meist erforderlichen Nachbesserungen aktiviert. Wir werden uns in diesem Kurs *nicht* damit beschäftigen, Plugins zum Automatisieren von Amos Graphics zu entwickeln.

Das Potential der Amos-Klassen ist im ca. 800 Seiten starken PDF-Dokument **Amos 18 Programming Reference Guide** beschrieben, das im **Documentation**-Unterordner einer Amos-Installation zu finden ist. Wer noch keine Erfahrung mit objektorientierter Programmierung hat, muss mit einigem Lernaufwand rechnen (siehe z.B. Baltes-Götz 2010) und sollte vielleicht besser bei der intuitiv bedienbaren Oberfläche von Amos Graphics bleiben.

Amos 18 enthält eine kleine Programmierungsumgebung, die über den Eintrag **Program Editor** in der Amos-Programmgruppe zu starten ist:



Alternativ kann man auch ein Standardwerkzeug wie Microsofts Visual Studio verwenden.

Wir wollen ein Programm in VB.NET (Visual Basic .NET) erstellen, das die Parameterschätzung in **Cov-Est** – Projekt automatisch ausführt. Nach dem Start der Entwicklungsumgebung ist bereits ein Programmrumpf vorhanden (siehe oben). Wir ersetzen in der Methode **Main()** den Kommentar

*'Your code goes here*

durch die folgende **Dim**-Anweisung, welche die Objektvariable **Sem** aus der Klasse **AmosEngine** deklariert:

**Dim Sem As New AmosEngine**

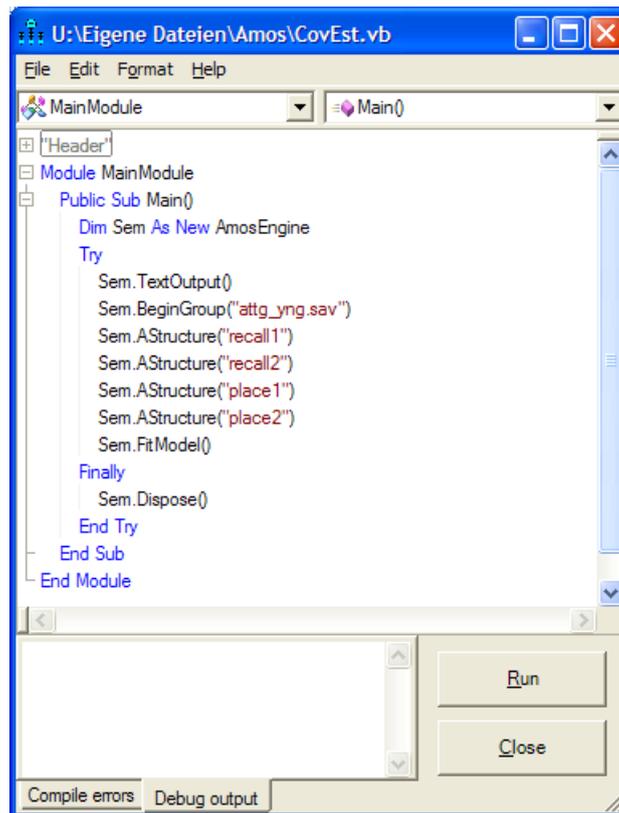
Das Objekt **Sem** steht für ein Strukturgleichungsmodell und besitzt zahlreiche Methoden, um seine Eigenschaften zu spezifizieren oder seine Leistungen abzurufen. Aufgrund günstig gewählter Voreinstellungen sind für unser Beispiel nur die in folgender Tabelle erläuterten Methodenaufrufe erforderlich:

Sem.TextOutput()	Ein erfolgreicher Programmablauf erzeugt die Ergebnisdatei <b>AmosScratch.AmosOutput</b> in einem Ordner für temporäre Dateien (z.B. <b>C:\Dokumente und Einstellungen\baltes\Lokale Einstellungen\Temp\AmosTemp</b> ). Durch Aufrufen seiner Methode <b>TextOutput()</b> teilen wir dem Objekt <b>Sem</b> mit, dass die Ausgabe automatisch vom zuständigen Hilfsprogramm geöffnet werden soll.
Sem.BeginGroup("attg_yng.sav")	Die Methode <b>BeginGroup()</b> leitet die Modellspezifikation für die erste (und hier auch einzige Gruppe) ein und nimmt den Namen der Eingabedatei entgegen. Dem Dateinamen müssen Sie eventuell noch eine Pfadangabe voranstellen.
Sem.AStructure("recall1") Sem.AStructure("recall2") Sem.AStructure("place1") Sem.AStructure("place2")	Mit der Methode <b>AStructure()</b> wird das Modell spezifiziert. Im Beispiel kann man sich auf die Deklaration der vier Variablen beschränken, weil die zugehörigen Kovarianzen auch ohne explizite Angabe in das Modell einbezogen werden. Einträge nach dem Muster  <p style="text-align: center;">Sem.AStructure("recall1 &lt;--&gt; recall1")</p> sind also überflüssig. Hier besteht ein Unterschied zu Amos <i>Graphics</i> , das bei einem fehlenden Doppelpfeil davon ausgeht, dass Sie die zugehörige Kovarianz auf Null fixieren wollen. Alternativ zur <b>AmosEngine</b> -Methode <b>AStructure</b> kann man auch die Methoden <b>Var</b> und <b>Cov</b> verwenden.
Sem.FitModel()	Mit der Methode <b>FitModel()</b> wird Amos aufgefordert, die Modellparameter zu schätzen und die Ausgaben zu erstellen.

Aus programmiertechnischen Gründen sollten noch die beiden folgenden Bestandteile ergänzt werden:

Sem.Dispose()	Die Methode <b>Dispose()</b> gibt alle mit dem <b>Sem</b> -Objekt verbundenen Ressourcen frei, beseitigt es also restlos. Dies ist wichtig, weil im Rahmen der Amos-Entwicklungsumgebung nur <i>ein AmosEngine</i> -Objekt existieren kann. Ohne die Aufräumaktion am Ende des Programms wäre kein zweiter Aufruf innerhalb derselben Instanz der Entwicklungsumgebung möglich.
Try ... Finally Sem.Dispose() End Try	Diese Konstruktion sorgt dafür, dass die wichtige Methode <b>Dispose()</b> auch dann noch ausgeführt wird, wenn das Programm wegen eines Ausnahmefehlers bei einem Methodenaufruf zwischen <b>Try</b> und <b>Finally</b> abgebrochen werden muss.

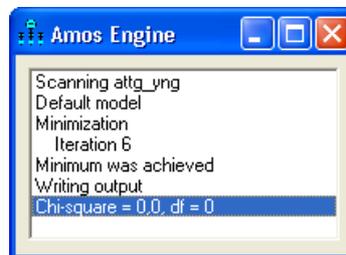
Wenn das fertige Programm



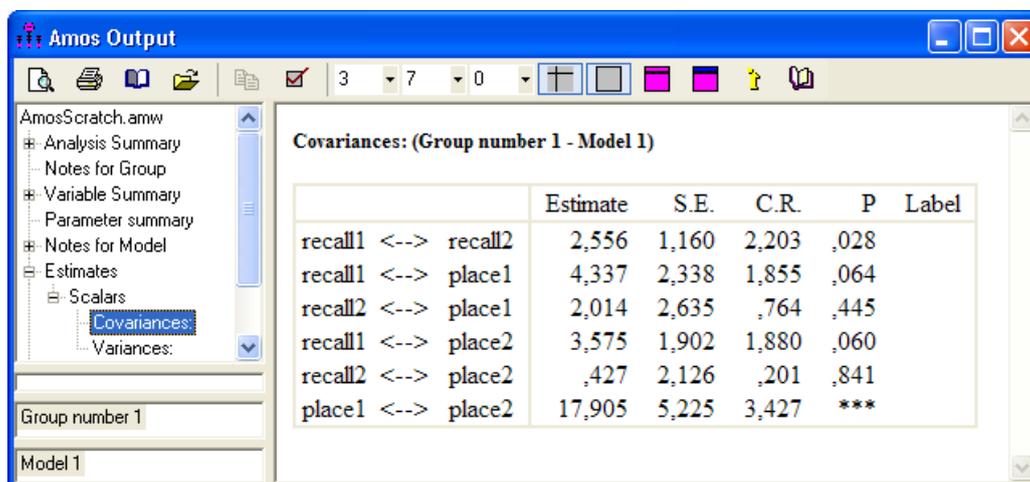
mit der Schaltfläche

### Run

gestartet wird, informiert Amos im folgenden Fenster über den Bearbeitungsfortschritt:



Wir erhalten dieselben Ergebnisbestandteile wie mit Amos Graphics:



Beim Speichern eines VB.NET - Programms entsteht eine einfache Textdatei mit der voreingestellten Namensweiterung **.vb**.

Es ist übrigens *nicht* möglich, zu einer Modellspezifikation über einen simplen Methodenaufwurf ein Pfaddiagramm erstellen zu lassen. Zwar kann man die vielfältigen Gestaltungsmöglichkeiten von Amos Graphics auch per Programm nutzen, doch erwartet die zuständige Klasse **PathDiagrammer** dazu präzise Einzelanweisungen.

Wer kein neues **AmosEngine**-Programm erstellen, sondern das zu einem vorhandenen Pfaddiagramm äquivalente Programm von Amos produzieren lassen möchte, wählt den Menübefehl

**Tools > Write a Program**

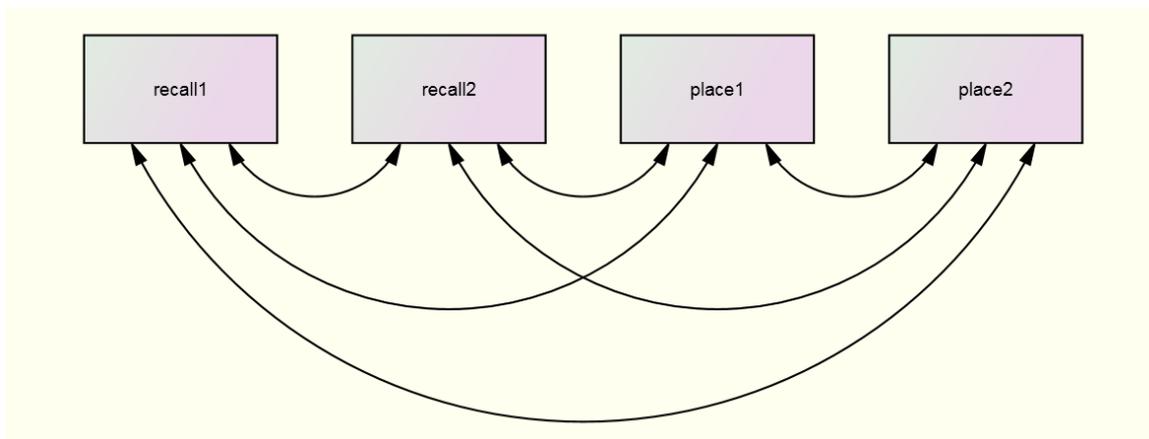
Dabei entsteht z.B. zur Projektdatei **CovEst.amw** die Programmdatei **CovEst\_converted.vb**.

**2.7 Modellgültigkeitstest**

Unter Verwendung der in Abschnitt 2 vorgestellten Daten zu zwei Gedächtnistests, die jeweils vor und nach einem Training erhoben wurden, wollen wir den  $\chi^2$ -Modellgültigkeitstest kennen lernen. Wir verwenden ihn zunächst anwendungsorientiert und verschieben die Behandlung seiner statistischen Grundlagen auf den Abschnitt 5.

Damit die Dateien zum ersten Beispiel im aktuellen Zustand erhalten bleiben, sichern wir das geöffnete Projekt mit **File > Save as** in die neue Projektdatei **Restrict.amw**.

Im bisherigen „Modell“



sind die vier Varianzen und die sechs Kovarianzen frei schätzbar, so dass keine prüfbare Behauptung vorliegt. Es liegt aber nahe, für die beiden Variablen **recall** und **place** zu beiden Messzeitpunkte (vor und nach dem Training) identische Kovarianzmatrizen anzunehmen:

- i)  $\text{Var}(\mathbf{recall1}) = \text{Var}(\mathbf{recall2})$
- ii)  $\text{Var}(\mathbf{place1}) = \text{Var}(\mathbf{place2})$
- iii)  $\text{Cov}(\mathbf{recall1}, \mathbf{place1}) = \text{Cov}(\mathbf{recall2}, \mathbf{place2})$

Durch diese Restriktionen resultiert ein Modell, das sich als falsch heraus stellen kann, z.B. weil die in einer Studie für **recall1** und **recall2** festgestellten Varianzunterschiede das durch zufällige Stichprobenschwankungen erklärbare Maß überschreiten.

Der von Amos bei Verwendung der Schätzmethoden ML, GLS oder ADF angebotene  $\chi^2$ -Modellgültigkeitstest beurteilt folgendes Hypothesenpaar:

- Nullhypothese  
Die Populationskovarianzmatrix hat die vom Modell behauptete bzw. implizierte Form.
- Alternativhypothese  
Die Populationskovarianzmatrix genügt nicht den vom Modell auferlegten Restriktionen.

Obwohl ein Modellgültigkeitstest methodologisch schon recht anspruchsvoll ist, verzichten wir vorläufig getreu dem Motto „Aller Anfang ist leicht“ aus der Abschnittsüberschrift noch auf eine Herleitung mit Hintergrundinformationen, die später in Abschnitt 5 nachgeliefert wird. Mit den Voraussetzungen der Strukturgleichungsanalyse (Stichprobengröße, Verteilung der manifesten Variablen) muss man sich aber auch bei ausgeprägter Anwendungsorientierung beschäftigen, was daher in den nächsten Abschnitten geschehen soll.

### 2.7.1 Erforderliche Stichprobengröße

Ein wichtiges Thema bei Verfahren, die auf asymptotischer Statistik beruhen, ist die **minimal erforderliche Stichprobengröße**. In der Literatur (z.B. Kline 2005, S. 111, Loehlin 2004, S. 55, Schumacker & Lomax 2004, S. 48ff) werden folgende Empfehlungen genannt:

- Mindestens  $N \geq 100$ , besser  $N \geq 200$
- Mindestens  $N \geq 5 \cdot t$ , wobei  $t$  die Anzahl der zu schätzenden Parameter ist.
- Beim ADF-Schätzverfahren sind noch größere Stichproben erforderlich ( $m$ : Anzahl der manifesten Variablen):

$$\min(200; 1,5 \cdot m \cdot (m + 1))$$

Im dem seit Abschnitt 2 betrachteten Gedächtnisprojekt, das aus der Amos-Beispielsammlung stammt, sind leider nur 40 Fälle vorhanden.

Eben wurde der erforderliche Stichprobenumfang für ein korrektes Verhalten der Signifikanztests bei *gültiger* Nullhypothese betrachtet, wobei vor allem eine erhöhte Fehlerrate *erster* Art zu vermeiden ist (falsche Entscheidung gegen eine gültige Nullhypothese). Das Planungsziel, die Wahrscheinlichkeit für einen Fehler *zweiter* Art (falsche Entscheidung gegen eine gültige Alternativhypothese) zu begrenzen, kann je nach angenommener Effektstärke eine erheblich größere Stichprobe erfordern (siehe z.B. Kline 2005, S. 156ff; Loehlin 2004, S. 70ff). Bei Tests zu einem *einzelnen* Parameter in einem Modell für manifeste Variablen) kann man sich an der Power-Analyse für die Korrelations- und Regressionsanalyse orientieren (siehe z.B. Cohen 1988). Mit der Power des Gültigkeitstests zu einem kompletten Modell werden wir uns in Abschnitt 5 kurz beschäftigen.

### 2.7.2 Verteilungsannahmen bei Strukturgleichungsanalysen

Hypothesentests, Vertrauensintervalle und Effizienzaussagen zu den **ML-** oder **GLS-**Schätzmethoden in Amos basieren auf folgenden **Verteilungsvoraussetzungen**:

- Die Beobachtungen (Fälle) sind stochastisch unabhängig voneinander.
- Die Kovarianzmatrix der manifesten Variablen hat vollen Rang, d.h. keine manifeste Variable ist von den anderen linear abhängig.
- Die manifesten Variablen sind gemeinsam multivariat normalverteilt.

Während die beiden ersten Voraussetzungen technischer Natur und meist erfüllt sind, stellt die Normalitätsannahme eine strenge Behauptung über die Population dar.

Wenn **manifeste exogene** Variablen als **fixiert** betrachtet werden können, genügen schwächere Verteilungsvoraussetzungen: Fixierte Variablen dürfen keinen Messfehler enthalten (siehe Abschnitt 6.1), was

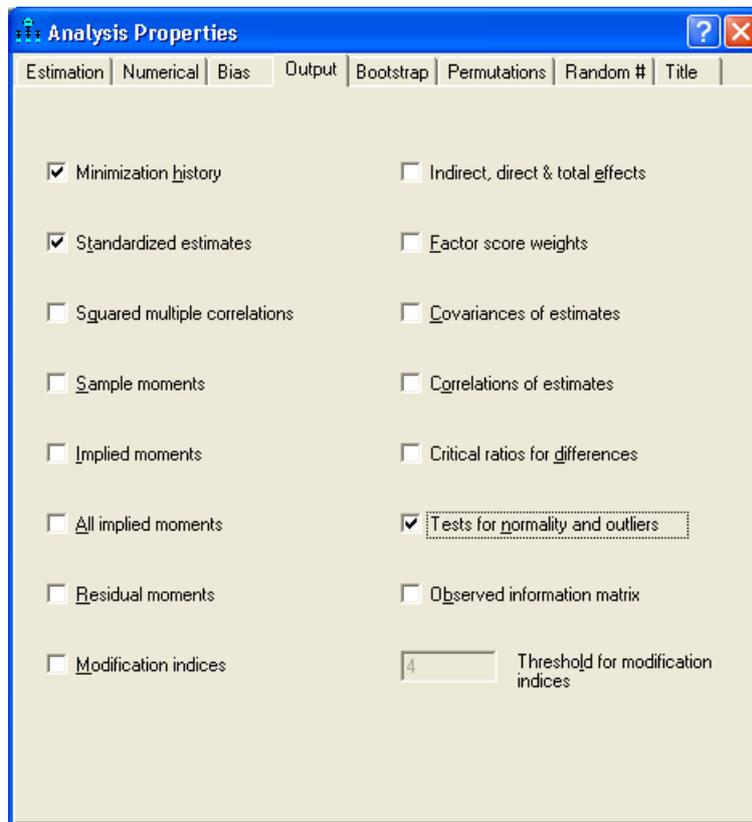
z.B. bei der von einem Experimentator kontrollierten Gruppenzugehörigkeit erfüllt ist. Die Verteilung der fixierten Variablen ist irrelevant, und für die übrigen Variablen muss bei Verwendung der ML- oder GLS-Schätzmethode gelten (Arbuckle 2009a, S. 35f):

- Für jede Wertekombination der fixierten Variablen ist die bedingte gemeinsame Verteilung der zufälligen Variablen multivariat normal.
- Die bedingten Varianz-Kovarianzmatrizen der Zufallsvariablen sind identisch für alle Wertekombinationen der fixierten Variablen.
- Die bedingten Erwartungswerte der Zufallsvariablen hängen linear von den Werten der fixierten Variablen ab.

Diese schwächeren Voraussetzungen sind identisch mit denjenigen der linearen Regressionsanalyse.

### 2.7.3 Normalverteilungs- und Ausreißerdiagnostik in Amos

Nach Abschnitt 2.7.2 setzt der geplante  $\chi^2$ -Modellgültigkeitstest für das restriktive Modell zu den Gedächtnisdaten voraus, dass die vier Variablen (**recall1**, **recall2**, **place1**, **place2**) gemeinsam multivariat normalverteilt sind. Wir fordern (über  oder **View > Analysis Properties > Output**) in der folgenden Dialogbox entsprechende Tests an



und erhalten nach Anforderung der Schätzung (mit  oder **Analyze > Calculate Estimates** oder **Strg+F9**) in der Textausgabe (zu öffnen mit  oder **View > Text Output** oder **F10**) folgende Ergebnisse:

**Assessment of normality (Group number 1)**

Variable	min	max	skew	c.r.	kurtosis	c.r.
place1	20,000	47,000	-,395	-1,021	,041	,053
recall1	5,000	16,000	,082	,212	-,139	-,179
recall2	4,000	16,000	-,481	-1,242	-,175	-,226
place2	30,000	48,000	,108	,278	-,750	-,968
Multivariate					-1,287	-,588

Bei den univariaten Ergebnissen zeigen sich beim Vergleich den CR - Statistiken mit dem kritischen Wert 1,96 (= 97,5%-Fraktile der Standardnormalverteilung) keine signifikanten Schiefe- bzw. Wölbungswerte (**skew** bzw. **kurtosis**). Auch die multivariate Wölbungsbeurteilung nach Mardia (1974) ist unauffällig. Allerdings sind die Testverfahren in der sehr kleinen Stichprobe ( $N = 40$ ) wenig sensibel gegenüber Abweichungen von der Normalitätsannahme.

Bei größeren Stichproben werden die (praktisch *immer* vorhandenen) Abweichungen von der (multivariaten) Normalität in der Regel aufgedeckt. Wie Arbuckle (2009b, S. 538) zu Recht feststellt, sind Signifikanztests wenig nützlich als Maße für den Abstand zur Normalverteilung und als Kriterien zur Unterscheidung zwischen harmlosen und ernsthaften Verletzungen der Verteilungsvoraussetzung. Aufgrund seiner Literaturdurchsicht empfiehlt Kline (2005, S. 50) Grenzwerte für die univariaten Verteilungsstatistiken: Die Beträge der Schiefe- bzw. Wölbungsstatistik sollten den Wert 3 bzw. 7 nicht überschreiben. Weil Amos wie viele andere Programme (z.B. SPSS) von der Wölbungsstatistik den Wert 3 subtrahiert, den eine Normalverteilung erreicht, wurde der von Kline angegebene Grenzwert 10 entsprechend korrigiert.

Zur Suche nach Ausreißern kann die folgendermaßen definierte **Mahalanobis-Distanz** dienen:

$$(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})$$

Sie quantifiziert die multivariate Distanz eines Falls mit den im Vektor  $\mathbf{x}$  gesammelten manifesten Variablenausprägungen vom korrespondierenden Vektor  $\boldsymbol{\mu}$  mit den Populationserwartungswerten, wobei die Elemente der invertierten Populationskovarianzmatrix  $\boldsymbol{\Sigma}$  zur Gewichtung verwendet werden. Amos berichtet in der Textausgabe die Schätzwerte und Signifikanzbeurteilungen:

Observation number	Mahalanobis d-squared	p1	p2
20	8,784	,067	,937
32	8,501	,075	,811
15	8,210	,084	,666
16	8,030	,090	,495
.	.	.	.

Für den Fall Nummer 20 mit der größten Mahalanobis-Distanz (8,784) erhalten wir folgende Überschreitungswahrscheinlichkeiten:

- Spalte **p1**

Bei einer zufällig aus einer multivariat normalverteilten Population gezogenen Beobachtung erhält man eine Mahalanobis-Distanz  $\geq 8,784$  mit einer Wahrscheinlichkeit von 0,067. In der Spalte **p1** sind Werte  $< 0,05$  speziell bei größeren Stichproben zu erwarten (eben mit der Wahrscheinlichkeit 0,05) und daher unkritisch.

- Spalte **p2**

Bei einer zufällig aus einer multivariat normalverteilten Population gezogenen Stichprobe mit demselben Umfang wie bei der aktuell analysierten erhält man eine *maximale* Mahalanobis-Distanz  $\geq 8,784$  mit einer Wahrscheinlichkeit von 0,937. In der Spalte **p2** sprechen Werte  $< 0,05$  für einen ernst zu nehmenden Ausreißer.

Während die Werte in der Spalte **p1** für alle Fälle identisch zu interpretieren sind, ändert sich die Bedeutung des **p2**-Werts in Abhängigkeit vom zugehörigen Rangplatz. Für den Fall Nummer 32 mit dem zweitgrößten Abstandswert erhalten wir in der **p2**-Spalte die Wahrscheinlichkeit dafür, dass in einer umfangsgleichen Stichprobe die zweitgrößte Mahalanobis-Distanz diesen Wert übersteigt.

Im Beispielprojekt gibt es keine Hinweise auf Ausreißer.

Die Normalitätsbeurteilungen erfordern einen kompletten Datensatz, stehen also fehlenden Werten nicht zur Verfügung.

### 2.7.4 Parameterrestriktionen formulieren

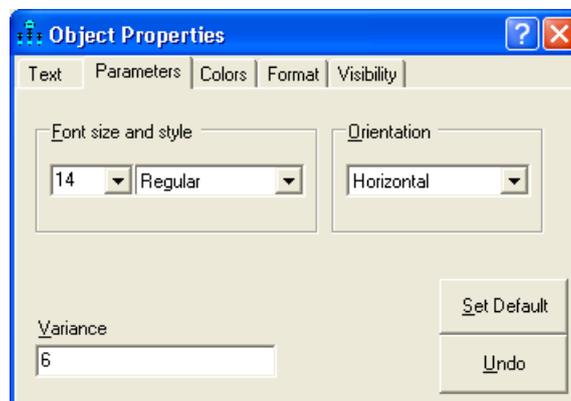
Amos und vergleichbare Programme erlauben eine flexible Modellformulierung mit Hilfe von Parameterrestriktionen:

- Parameter können auf feste Werte fixiert werden.  
Eine Fixierung auf den Wert Null lässt sich in Amos Graphics für viele Parameter durch Verzicht auf entsprechende Bestandteile im Modellspezifikations-Pfaddiagramm ausdrücken. Um z.B. eine Kovarianz auf Null zu setzen, lässt man den zugehörigen Doppelpfeil einfach weg.
- Zwei oder mehrere Parameter können auf einen identischen (schätzbaren) Wert gesetzt werden.

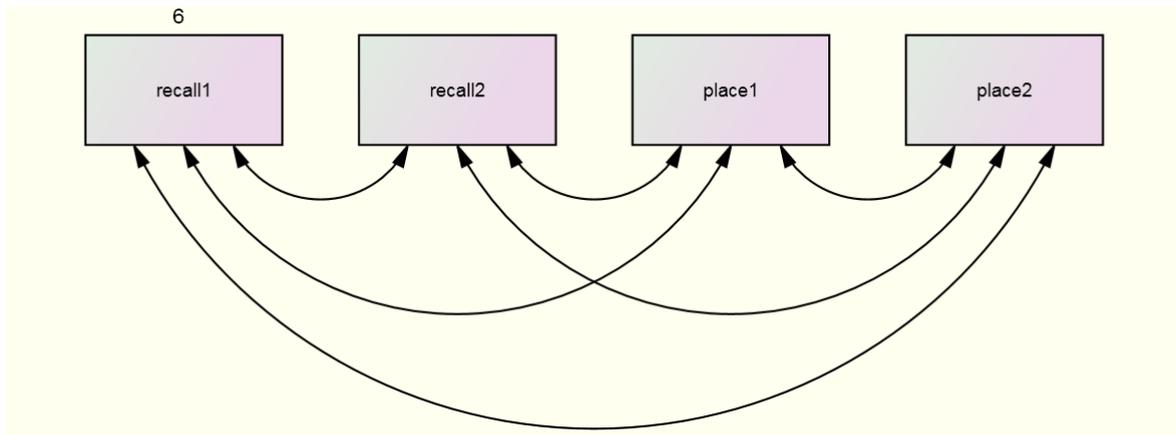
Wir werden in Abschnitt 2.7.5 mit dem  $\chi^2$ -Modellgültigkeitstest eine Möglichkeit kennen lernen, ein komplettes Modell mit allen darin enthaltenen Parameterrestriktionen zu prüfen. Durch den Vergleich von geschachtelten Modellen lässt sich eine einzelne Restriktion oder ein Bündel von Restriktionen prüfen (siehe Abschnitt 5.5).

Um in Amos Parameterrestriktionen zu formulieren, wechselt man nötigenfalls mit den Schalter  in

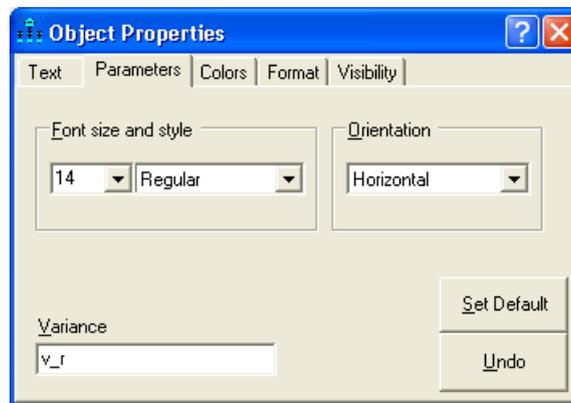
den Modellspezifikationsmodus und öffnet die Dialogbox **Object Properties** über das Kontextmenü zum betroffenen Objekt. Nun können auf der **Parameters**-Registerkarte die gewünschten Restriktionen eingetragen werden. Im folgenden Beispiel wird die Varianz der Variablen **recall1** auf den Wert 6 fixiert:



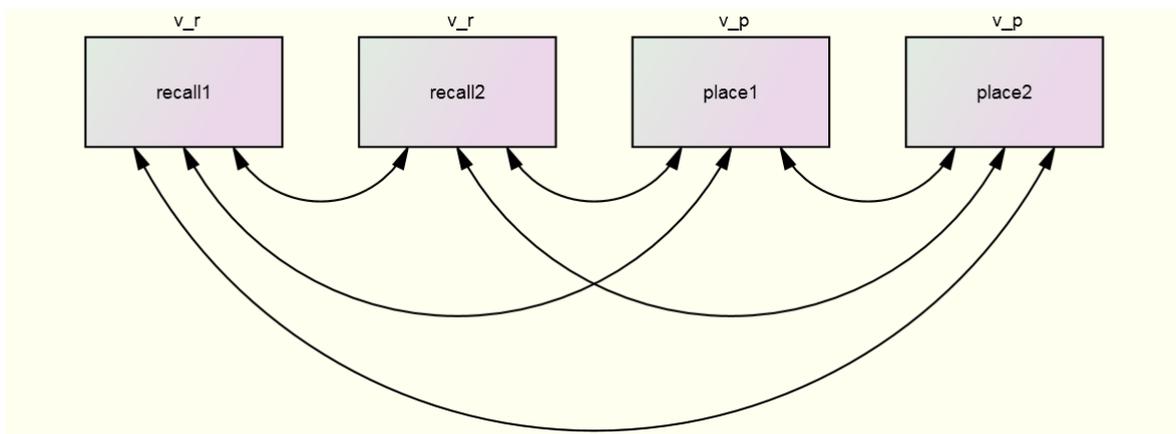
Amos zeigt die Werte der fixierten Parameter im Modellspezifikations-Pfaddiagramm an, z.B.:



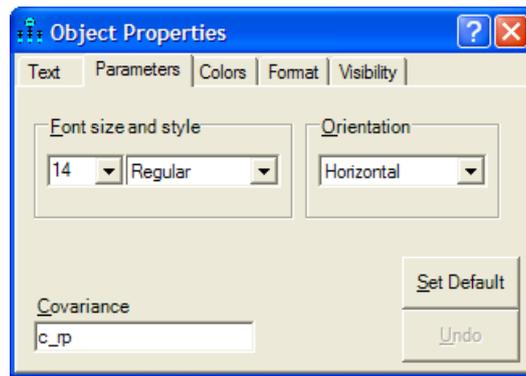
Um eine *Gleichheitsrestriktion* zu formulieren, trägt man für alle betroffenen Parameter eine gemeinsame symbolische Wertbezeichnung ein, z.B.:



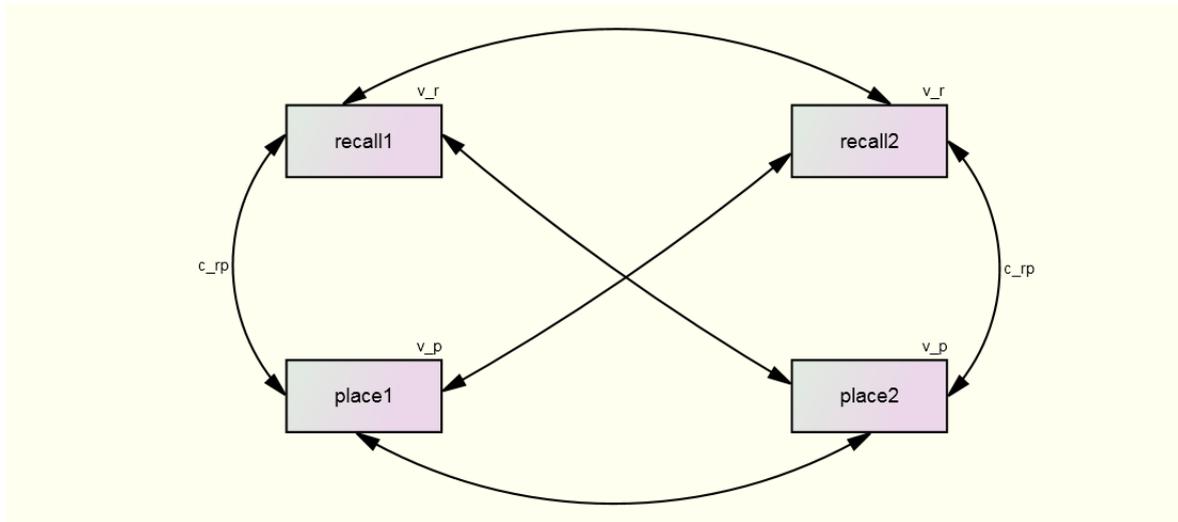
Die Bezeichnung ist weitgehend frei wählbar, sollte aber keine „syntaktisch bedeutsame“ Zeichen (z.B. Klammern) enthalten. Nachdem auf diese Weise in unserem Beispiel die Varianzen der beiden **recall**-Variablen sowie die Varianzen der beiden **place**-Variablen gleich gesetzt wurden, sieht das Modellspezifikations-Pfaddiagramm folgendermaßen aus:



Auch für Kovarianzen können auf analoge Weise Restriktionen formuliert werden. Im folgenden Modell wurde die **recall1-place1** - Kovarianz mit der **recall2-place2** - Kovarianz gleich gesetzt, indem beide Parameter über die Dialogbox **Object Properties** denselben symbolischen Namen



erhalten haben:



Außerdem wurde das Pfaddiagramm mit den **Shape**- und **Move**-Werkzeugen umgestaltet (vgl. Abschnitt 2.1).

### 2.7.5 $\chi^2$ -Gültigkeitstest zum Modell mit den Parameterrestriktionen

Das Modell aus Abschnitt 2.7.4 mit den drei Gleichheitsrestriktionen:

- i)  $\text{Var}(\mathbf{recall1}) = \text{Var}(\mathbf{recall2})$
- ii)  $\text{Var}(\mathbf{place1}) = \text{Var}(\mathbf{place2})$
- iii)  $\text{Cov}(\mathbf{recall1}, \mathbf{place1}) = \text{Cov}(\mathbf{recall2}, \mathbf{place2})$

ist **überidentifiziert** (siehe Abschnitt 5.3), weil ...

- es **identifiziert** ist, also für jeden Parameter eine eindeutige Schätzung erlaubt,
- und außerdem die Anzahl der Stichprobenmomente (im Beispiel: Varianzen und Kovarianzen) *größer* ist als die Anzahl der frei schätzbaren Parameter.

Auch bei perfekter Modellgültigkeit in der *Population* werden in einer *Stichprobe* die Varianzen und Kovarianzen im Allgemeinen *nicht* den Modellrestriktionen genügen. Für unser Beispiel liefert Amos bei Verwendung der voreingestellten Maximum Likelihood - Methode die folgenden modellgemäßen Parameterschätzer:

**Covariances**

	Estimate	S.E.	C.R.	P	Label
place2 <--> recall2	2,712	1,821	1,489	,136	c_rp
recall1 <--> place1	2,712	1,821	1,489	,136	c_rp
place2 <--> place1	17,149	5,155	3,327	***	
recall2 <--> place1	2,220	2,216	1,002	,316	
place2 <--> recall1	4,608	2,166	2,127	,033	
recall2 <--> recall1	2,872	1,208	2,377	,017	

**Variances**

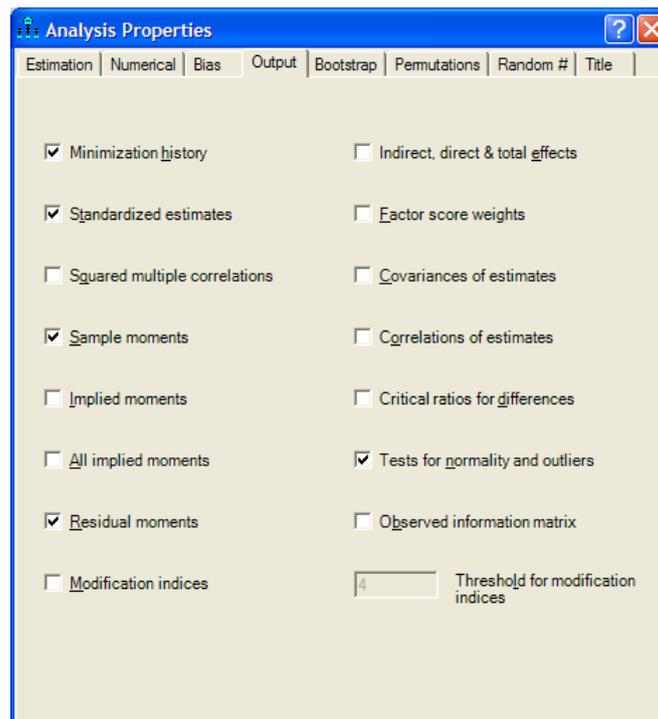
	Estimate	S.E.	C.R.	P	Label
recall1	7,055	1,217	5,798	***	v_r
recall2	7,055	1,217	5,798	***	v_r
place1	27,525	5,177	5,317	***	v_p
place2	27,525	5,177	5,317	***	v_p

Wir fordern nach



oder **View > Analysis Properties > Output**

in der folgenden Dialogbox



zusätzlich die **Sample moments** und die **Residual moments** an. Auf die **Implied moments** können wir verzichten, weil unser sehr elementares Modell die fraglichen (Ko)varianzen als Parameter enthält und deren Schätzwerte zur Standardausgabe gehören (siehe oben). Nach einer erneuten Modellschätzung enthält die Textausgabe auch die Stichprobenkovarianzmatrix:

**Sample Covariances (Group number 1)**

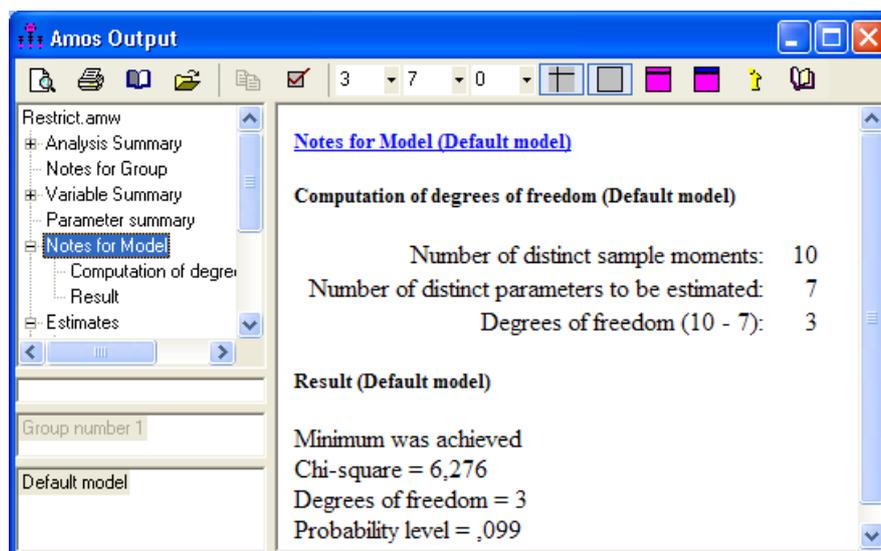
	place1	recall1	recall2	place2
place1	33,578			
recall1	4,338	5,788		
recall2	2,014	2,556	7,944	
place2	17,905	3,575	,427	22,160

und die Matrix mit den Abweichungen der Stichprobenkovarianzen von den Modellimplikationen:

**Residual Covariances (Group number 1 - Default model)**

	place1	recall1	recall2	place2
place1	6,052			
recall1	1,625	-1,267		
recall2	-,207	-,316	,890	
place2	,756	-1,033	-2,285	-5,365

Der  $\chi^2$ -**Modellgültigkeitstest** basiert auf der Wahrscheinlichkeit, solche oder stärkerer Abweichungen aufgrund von zufälligen Stichprobeneffekten bei einer Population mit gültigem Modell anzutreffen. Amos protokolliert das Ergebnis in der Textausgabe im Abschnitt **Notes for Model**:

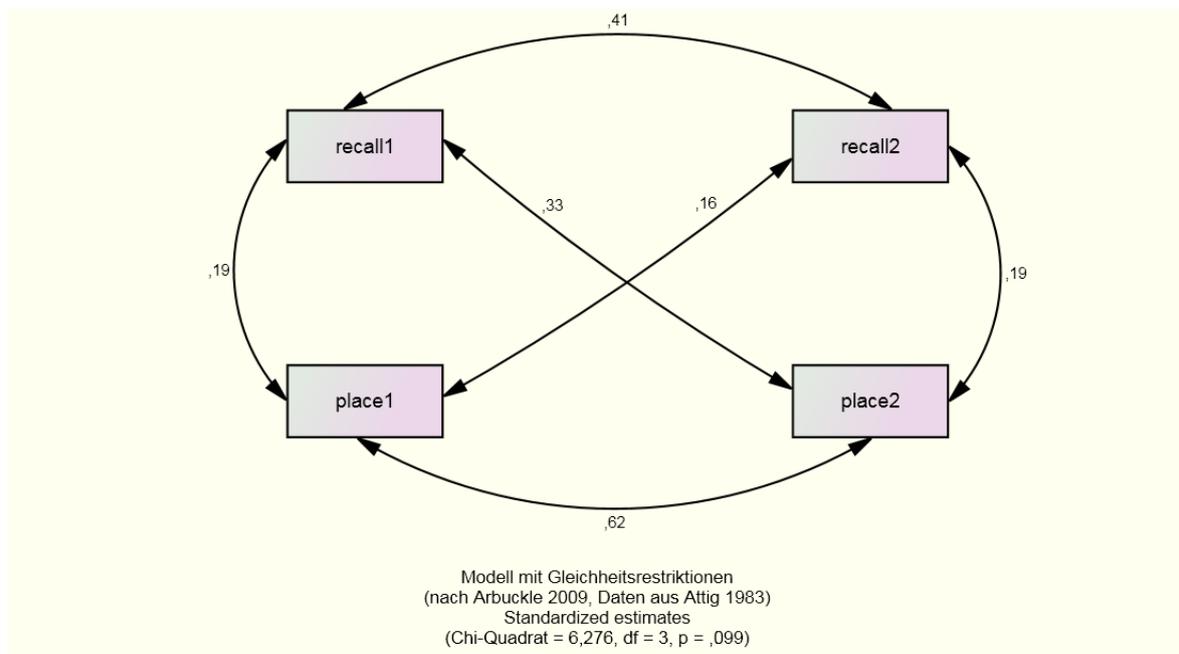


Amos ermittelt für die  $\chi^2$ -Prüfgröße den Wert 6,276, zu dem bei 3 Freiheitsgraden die Überschreitungswahrscheinlichkeit 0,099 gehört. Damit wird bei  $\alpha = 0,05$  die Nullhypothese, also unser Modell mit den drei Parameterrestriktionen, *nicht* abgelehnt.

Speziell bei größeren Stichproben neigt der  $\chi^2$ -Modellgültigkeitstest dazu, taugliche Modelle wegen vernachlässigbarer Schwächen zu verwerfen. Es sind zahlreiche weitere Kriterien zur Modellbeurteilung vorgeschlagen worden, die sich teilweise auch für größere Stichproben eignen (siehe Abschnitt 5.6).

## 2.8 Beschriftungen und andere Möglichkeiten zur Gestaltung des Pfaddiagramms

Das eben erfolgreich getestete Model sieht (nach einigen gleich zu besprechenden Nachbearbeitungen) im Ausgabe-Pfaddiagramm folgendermaßen aus:



### 2.8.1 Diagrammbeschriftungen

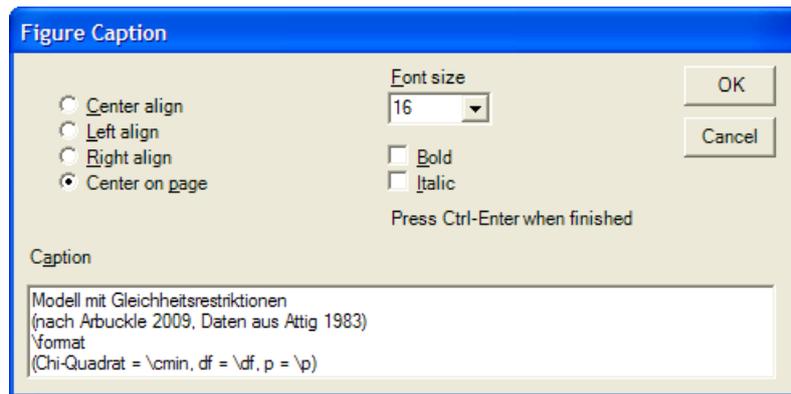
Gehen Sie folgendermaßen vor, um die in obigem Pfaddiagramm enthaltenen Beschriftungen einzufügen:

- Bei Anzeige eines *Ausgabe*-Pfaddiagramms (mit Parameterschätzungen) ist das Beschriftungs-Werkzeug nicht verfügbar. Wechseln Sie also nötigenfalls mit dem Schalter  in den Modellspezifikationsmodus.
- Aktivieren Sie das Beschriftungswerkzeug mit:  oder **Diagram > Figure Caption**
- Klicken Sie auf eine *freie Stelle* der Zeichenfläche während der Mauszeiger mit  beschriftet ist. Daraufhin erscheint die Dialogbox **Figure Caption**.
- Tragen Sie hier den gewünschten Text ein, wobei u.a. die folgenden **Textmakros** zur Verfügung stehen, die von Amos beim Zeichnen des Pfaddiagramms durch aktuell gültige Inhalte ersetzt werden:

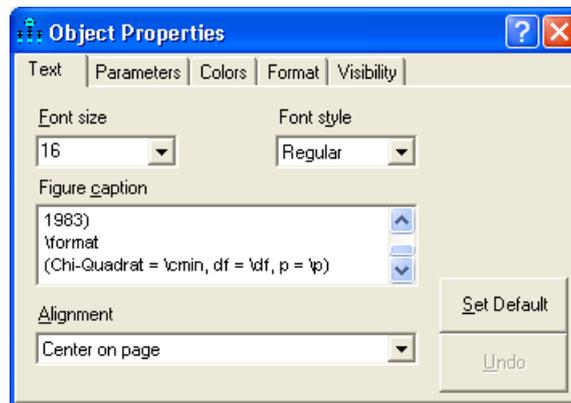
Textmakro	Es erscheint
<b>\format</b>	„Model Specification“ im Modellspezifikations-Pfaddiagramm, „Unstandardized estimates“ oder „Standardized estimates“ im Ausgabe-Pfaddiagramm
<b>\cmin</b>	$\chi^2$ -Statistik zum Modellgültigkeitstest
<b>\df</b>	Anzahl der Freiheitsgrade zum Modellgültigkeitstest
<b>\p</b>	Überschreitungswahrscheinlichkeit zum Modellgültigkeitstest

Eine vollständige Liste der Amos-Textmakros finden Sie im Hilfesystem über den Indexeintrag **Text macros**.

Die obige Beschriftung wurde mit folgender Dialogbox erzeugt:



- Beschriftungen lassen sich wie beliebige andere Elemente eines Pfaddiagramms mit dem  – Werkzeug verschieben und mit dem  –Werkzeug entfernen.
- Per  –Mausklick kann man im Modellspezifikationsmodus die Dialogbox **Figure Caption** zur Modifikation einer vorhandenen Beschriftung öffnen. Auch die per Standardwerkzeug  via Kontextmenü erreichbare Dialogbox **Object Properties** erlaubt eine Änderung der Beschriftung:



In den beiden genannten Dialogboxen steht die Positionierungsoption **Center on page** zur Verfügung, die für ein horizontales Zentrieren auf der Seite sorgt und das freie Bewegen auf die vertikale Richtung beschränkt.

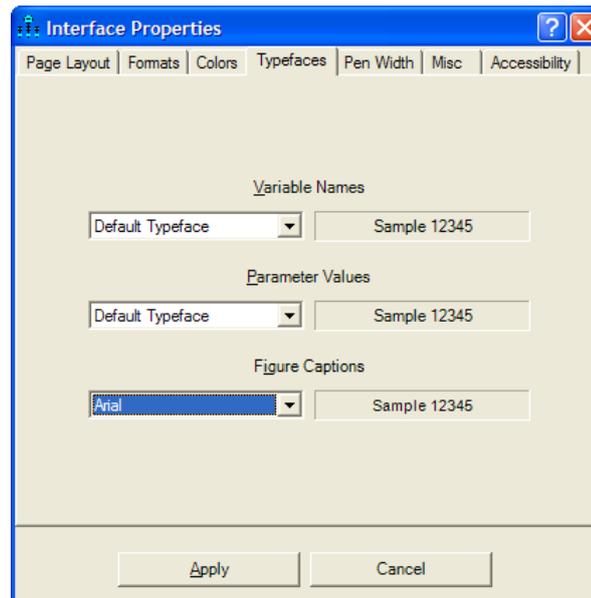
## 2.8.2 Formatierungsoptionen

Die Dialogbox **Interface Properties**, die nach

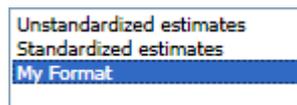
**View > Interface Properties** oder **Strg+I**

erscheint, bietet etliche Optionen zur Formatierung einzelner Ausgabebestandteile, z.B.:

- Auf dem Registerblatt **Typefaces** legt man fest, welche Schriftarten in Pfaddiagrammen für Variablenamen, Parameterwerte und Beschriftungen verwendet werden sollen:



- Auf dem Registerblatt **Pen Width** kann man Linienstärken und Pfeilspitzen verändern.
- Auf dem Registerblatt **Formats** kann man die beiden unveränderlichen Formatfamilien **Unstandardized estimates** und **Standardized estimates** einsehen sowie zusätzliche Familien definieren oder ändern. Diese stehen dann als Anzeigevarianten für das Pfaddiagramm zur Verfügung, z.B.:



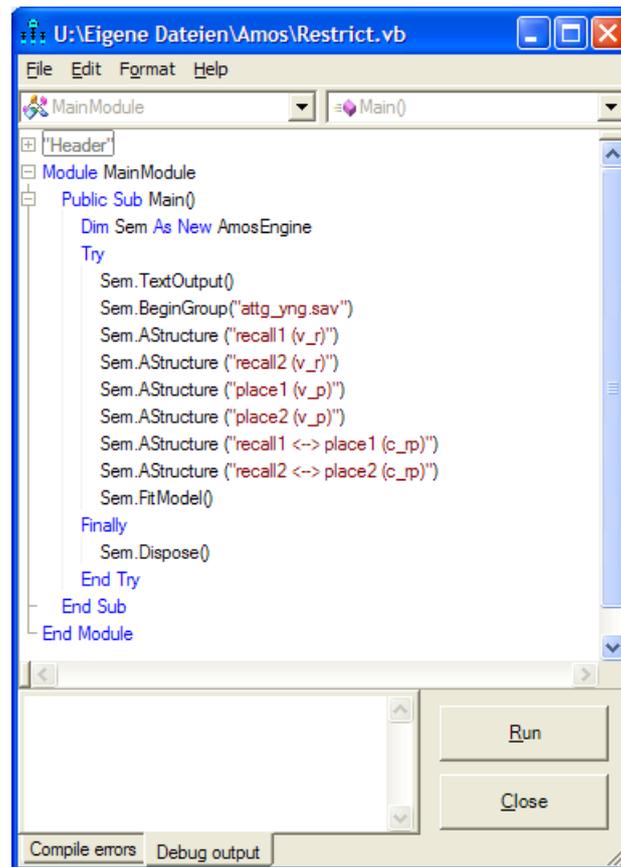
- Die von Textmakros verwendete Anzahl an Dezimalstellen lässt sich auf der Registerkarte **Misc** einstellen.
- Über die Optionen der Registerkarte **Page Layout** wurde schon in Abschnitt 2.2.1 berichtet.

## 2.9 Parameterrestriktionen in AmosEngine-Programmen formulieren

Mit Hilfe der folgenden Syntaxregeln lassen sich die in Abschnitt 2.7.4 behandelten Parameterrestriktionen in **AmosEngine**-Programmen (vgl. Abschnitt 2.6) formulieren:

- Für jede an einer Restriktion beteiligte Kovarianz ist ein **AStructure**-Methodenaufruf erforderlich. Dort setzt man zwischen die beiden Variablen einen Kovarianzpfel, dargestellt durch die Zeichenfolgen "<>" oder "<-->" (mit beliebig vielen Bindestrichen).
- Eine Restriktion wird durch einen in Klammern hinter die Parameterrepräsentation gesetzten Wert ausgedrückt. Dabei kann es sich um eine Konstante (bei der Fixierung auf einen festen Wert) oder um eine symbolische Wertbezeichnung handeln (bei der Gleichheitsrestriktion).

Diese etwas abstrakt formulierten Regeln werden in folgendem Beispielpogramm veranschaulicht, das die in Abschnitt 2.7.5 besprochene Textausgabe erzeugt:



Alternativ zur **AmosEngine**-Methode **AStructure** kann man auch die Methoden **Cov** und **Var** verwenden.

## 2.10 Parameter per C.R. - oder Modellgültigkeitstest auf Signifikanz prüfen

In Abschnitt 2.4.2 haben wir Parametertests über das **C.R.** (für *Critical Ratio*)

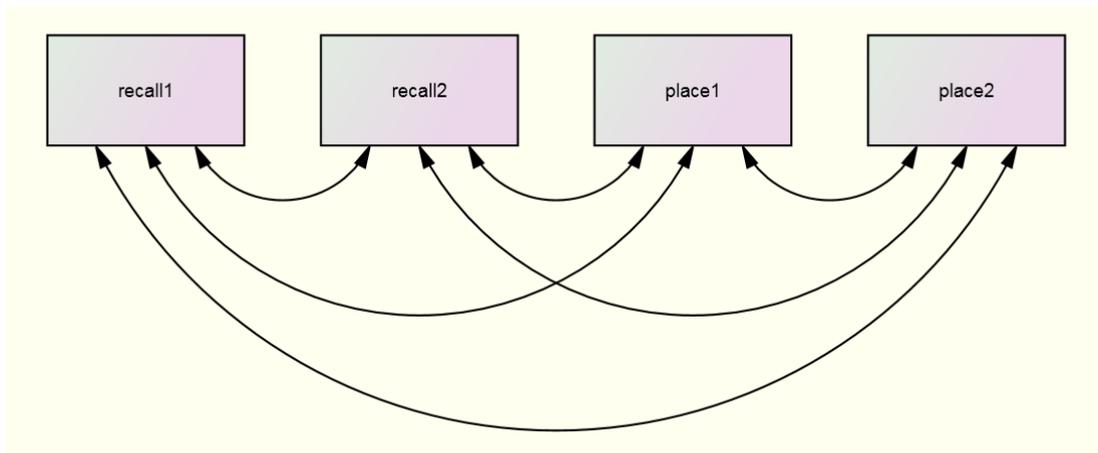
$$\left| \frac{\text{Schätzwert}}{\text{Standardfehler}} \right|$$

kennen gelernt und als Beispiel folgendes Testproblem behandelt:

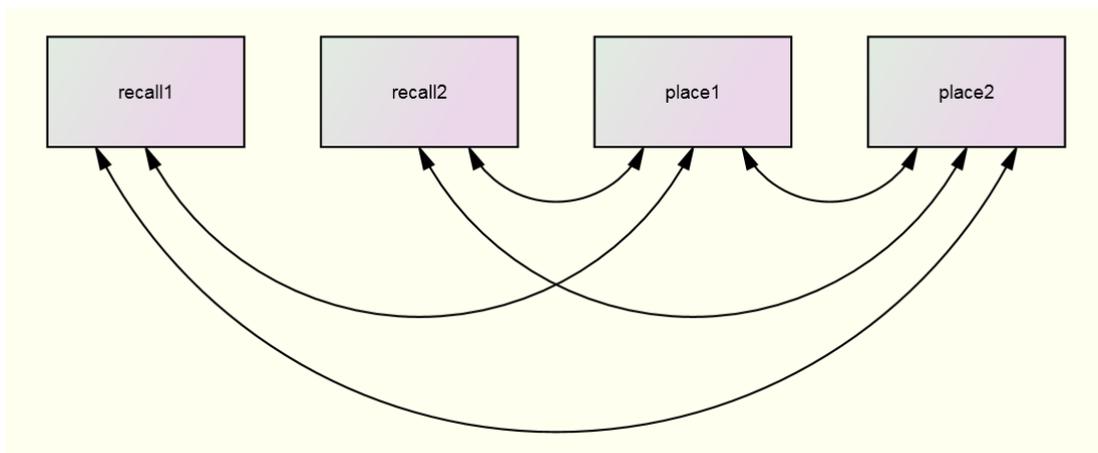
$$H_0: \text{Cov}(\text{recall1}, \text{recall2}) = 0 \quad \text{versus} \quad H_1: \text{Cov}(\text{recall1}, \text{recall2}) \neq 0$$

Es resultierte die Prüfgröße 2,203 und eine Überschreitungswahrscheinlichkeit von  $p = 0,028$ , so dass die Nullhypothese verworfen werden konnte.

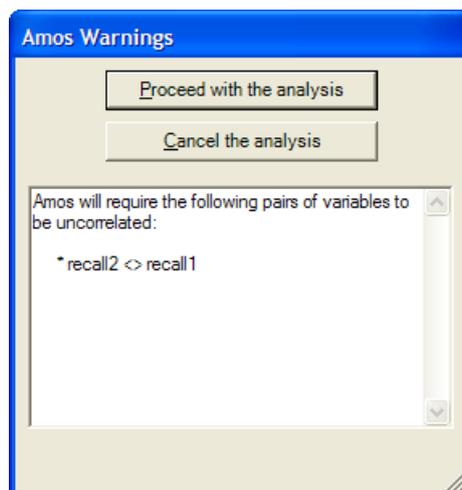
Eine alternative Signifikanzprüfung zu einem Parameter lässt sich über einen Modellgültigkeitstest (vgl. Abschnitt 2.7.5) realisieren, indem der fragliche Parameter aus einem **saturierten Modell** (vgl. Abschnitt 5.3) gestrichen wird, das jede Stichprobenkovarianzmatrix exakt reproduzieren kann und ebenso viele freie Parameter enthält, wie Stichprobenmomente vorhanden sind. Das simple „Modell“ aus Abschnitt 2.1 besitzt diese Eigenschaft:



Wir setzen nun die Kovarianz von RECALL1 und RECALL2 auf Null (z.B. durch Streichen des Kovarianzpfils),



und fordern Amos mit  zur Schätzung der Modellparameter auf. Das Programm fragt vorsichtshalber nach, ob wir die Fixierung auf Null ernst meinen oder vielleicht nur einen Pfeil vergessen haben:



Daher ist es vielleicht sinnvoller, die Fixierung mit den Methoden aus 2.7.4 explizit zu formulieren.

Im approximativen  $\chi^2$ -Test wird die Nullhypothese mit  $p = 0,014$  verworfen, wie der Textausgabe zu entnehmen ist (im Abschnitt **Notes for Model > Result**):

**Result (Default model)**

Minimum was achieved

Chi-square = 5,978

Degrees of freedom = 1

Probability level = ,014

Diese Überschreitungswahrscheinlichkeit liegt trotz der relativ kleinen Stichprobe ( $N = 40$ ) relativ nahe beim Ergebnis des exakten t-Tests, der  $p = 0,016$  liefert.

Asymptotisch werden die beiden Amos-Testverfahren miteinander und mit dem exakten t-Test übereinstimmen. Bei einer kleinen Stichprobe wie der unsrigen ist mit Abweichungen zu rechnen, die durchaus auch zu unterschiedlichen Testentscheidungen führen können. Unter den beiden von Amos angebotenen Testverfahren ist im Zweifelsfall das Verfahren per Parameterrestriktion und Modellgültigkeitstest zu bevorzugen. Dabei muss allerdings für *jeden* zu beurteilenden Parameter ein spezielles eingeschränktes Modell analysiert werden.

In Abschnitt 5.5 wird sich zeigen, dass ein Signifikanztest mit dem saturierten Modell als Referenz ein Spezialfall für das generelle Prinzip der Testkonstruktion durch Vergleich von zwei geschachtelten Modellen ist.

Wie Sie bereits wissen, besteht eine Stärke der Testkonstruktion über Parameterrestriktionen darin, dass sehr flexibel *zusammengesetzte* Hypothesen (über *mehrere* Parameter) formuliert und beurteilt werden können.

### 3 Modelle für manifeste Variablen

Nachdem wir schon einige Erfahrung bei der Bedienung des Computer-Programms Amos erworben haben, konzentrieren wir uns im Kurs nun zunehmend auf methodische Themen. Dabei bleiben wir zunächst anwendungsorientiert und vertagen die Beschäftigung mit den statistischen Grundlagen der Strukturgleichungsanalyse auf den Abschnitt 5. Vorher betrachten wir zwei wichtige Anwendungen, die man auch mit den beiden Bestandteilen eines „kompletten“ Strukturgleichungsmodells für latente Variablen assoziieren kann (Strukturmodell und Messmodell):

- Modelle für manifeste Variablen
- Konfirmatorische Faktorenanalyse

#### 3.1 Beispiel

In einer Untersuchung zu den Bedingungen der beruflichen Ambitionen von Jugendlichen (Kerchoff 1974) wurden bei 767 Jungen im Alter von 12 Jahren folgende Variablen erhoben:<sup>1</sup>

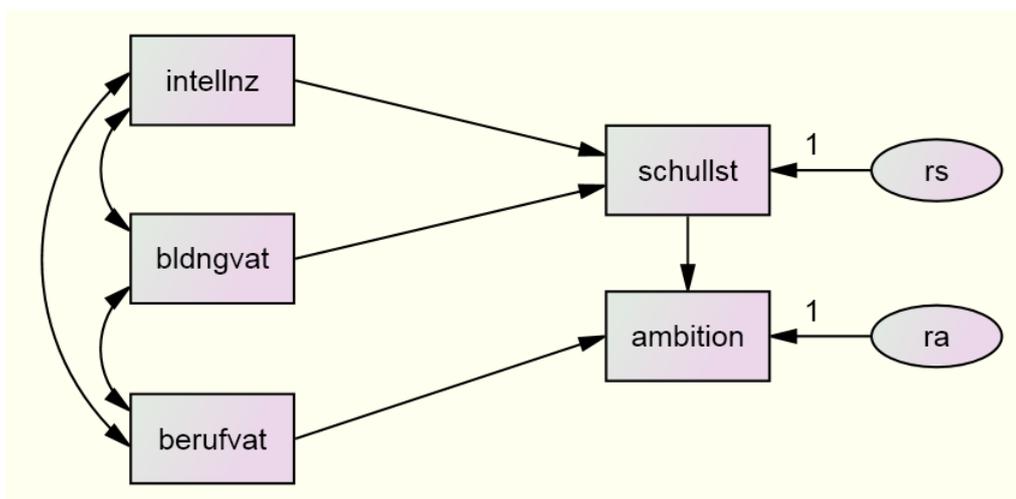
<b>INTELLNZ</b>	Intelligenz des Jungen
<b>BLDNGVAT</b>	Bildungsgrad des Vaters
<b>BERUFVAT</b>	Beruf des Vaters
<b>SCHULLST</b>	Schulleistung des Jungen
<b>AMBITION</b>	Berufliche Ambitionen des Jungen

Über das Zusammenwirken der Variablen wird angenommen:

$$\begin{aligned} \text{SCHULLST} &= \gamma_{11} \text{INTELLNZ} + \gamma_{12} \text{BLDNGVAT} + \text{RS} \\ \text{AMBITION} &= \gamma_{21} \text{BERUFVAT} + \gamma_{22} \text{SCHULLST} + \text{RA} \end{aligned}$$

In der ersten Gleichung wird u.a. behauptet, dass eine Erhöhung der Intelligenz um eine Maßeinheit bei Konstanthaltung des väterlichen Bildungsgrads im Mittel zu einer Veränderung der Schulleistung um  $\gamma_{11}$  Maßeinheiten führt. Analog sind die übrigen Strukturkoeffizienten zu interpretieren.

Das mit Amos bequem zu erstellende Pfaddiagramm zeigt das Modell anschaulich und präzise:



Wir beschränken uns zunächst auf die *erste* Gleichung, erproben Amos also bei einer Aufgabenstellung, die auch mit traditioneller Regressionstechnologie zu lösen ist.

<sup>1</sup> Die völlige Fixierung auf väterliche Karrieremerkmale ist wohl der im letzten Jahrtausend noch mangelhaften Emanzipation anzulasten.

Interessierte finden die Kerchhoff-Daten in der SPSS-Matrix-Datendatei **Ambition.sav** (Bezugsquelle: siehe Vorwort).

### 3.2 Einlesen von Momentmatrizen

Die eben genannte Datei **Ambition.sav** enthält *keine* Fälle  $\times$  Variablen - Datenmatrix, sondern eine Korrelationsmatrix mit einigen Zusatzinformationen, wie das folgende SPSS-Datenfenster zeigt:

	rowtype_	varname_	intellnz	bldngvat	berufvat	schullst	ambition
1	STDEV		1,0000	1,0000	1,0000	1,0000	1,0000
2	MEAN		,0	,0	,0	,0	,0
3	N		767,0000	767,0000	767,0000	767,0000	767,0000
4	CORR	INTELLNZ	1,0000	,2770	,2500	,5720	,3350
5	CORR	BLDNGVAT	,2770	1,0000	,6110	,2940	,3030
6	CORR	BERUFVAT	,2500	,6110	1,0000	,2480	,3310
7	CORR	SCHULLST	,5720	,2940	,2480	1,0000	,4780
8	CORR	AMBITION	,3350	,3030	,3310	,4780	1,0000

Amos benötigt für die meisten Analysen keine Rohdaten, sondern lediglich die Stichprobenmomente (d.h. Mittelwerte, Varianzen und Kovarianzen). Es genügt also, die Stichprobenmomente (z.B. in einer SPSS-Matrixdatendatei) anzubieten.

In der Regel sind jedoch Rohdaten vorhanden, und diesen können selbstverständlich direkt verwendet werden. Im aktuellen Abschnitt wird eine Lösung für spezielle Eingabedaten beschrieben und keinesfalls empfohlen, Rohdaten aufwändig auf dieses Format zu bringen.

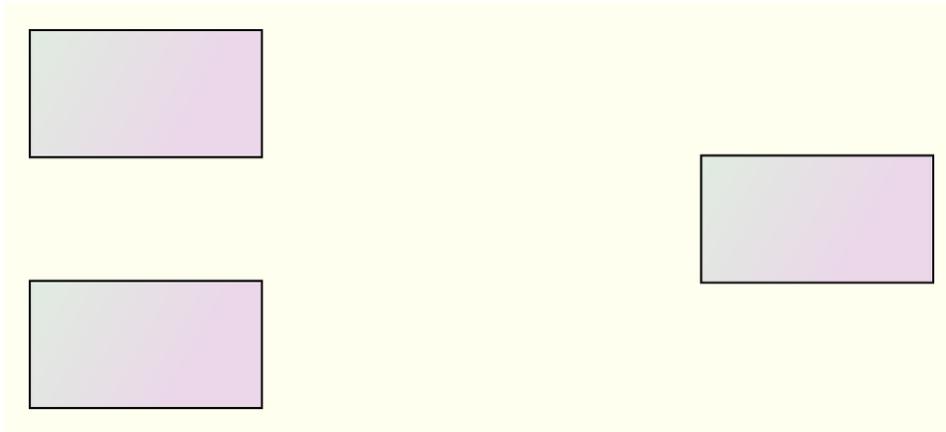
Die Verwendung der (gelegentlich ausschließlich publizierten) Korrelationsmatrix an Stelle der von Amos (und vergleichbaren Strukturgleichungsprogrammen) in der Regel verlangten Kovarianzmatrix ist allerdings nur unter folgenden Bedingungen erlaubt:

- Das Strukturgleichungsmodell ist skaleninvariant.  
Ein skaleninvariantes Modell ist nicht abhängig von den Maßeinheiten der manifesten Variablen und kann Änderungen der Maßeinheiten durch analoge Änderungen seiner Parameterwerte kompensieren (siehe Abschnitt 5.4.1).
- Das Modell kann die Varianzen der manifesten Variablen perfekt reproduzieren.
- Es wird eine skaleninvariante Schätzmethode (z.B. ML, GLS, ADF) verwendet.  
Sofern das untersuchte Modell seinerseits skaleninvariant ist, bleiben die Ergebnisse einer skaleninvarianten Schätzmethode im Wesentlichen invariant bei Veränderung der Maßeinheiten von manifesten Variablen (siehe Abschnitt 5.4.1).

Bei einem Regressionsmodell und der in Amos voreingestellten ML-Schätzmethode sind diese Bedingungen erfüllt. Weitere Informationen zur nicht ganz trivialen Frage der Äquivalenz von Kovarianz- und Korrelationsmatrizen bei der Analyse von Strukturgleichungsmodellen finden Sie z.B. bei Baltes-Götz (1994, S. 4-9), Jöreskog & Sörbom (1989, Abschnitt 1.21) und Lee (1985).

### 3.3 Pfaddiagramme mit gerichteten Pfeilen und latenten Variablen

Wir wollen die lineare Regression von Schulleistung auf die Intelligenz des Schülers und die Bildung des Vaters untersuchen und erstellen zunächst das Pfaddiagramm.



Beim Zeichnen der drei Rechtecke empfiehlt sich folgende Vorgehensweise:

- Malen Sie mit -Werkzeug ein Rechteck für die erste Variable, und korrigieren Sie seine Form bei Bedarf mit dem -Werkzeug.
- Erzeugen Sie mit -Werkzeug für die restlichen Variablen Kopien dieses Rechtecks. Halten Sie beim Verschieben der Kopie für **BLDNGVAT** die -Taste fest, damit das Rechteck seine horizontale Position behält.
- Mit

#### Edit > Space Vertically

können Sie dafür sorgen, dass die drei (zuvor markierten) Rechtecke *vertikal* gleiche Abstände haben.

Zeichnen Sie nun mit

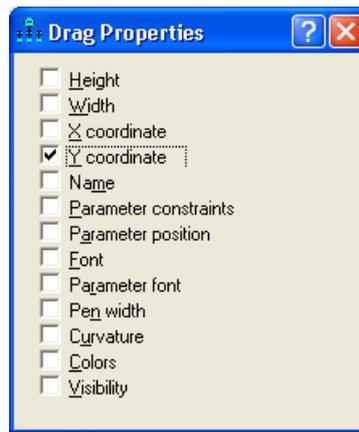
 oder **Diagram > Draw Unobserved** oder **F4**

eine Ellipse für die latente (!) Fehlervariable zum Kriterium.

Um die Y-(Höhen)Koordinate des Kriteriums-Rechtecks auf die zugehörige Fehler-Ellipse zu übertragen, aktivieren Sie am besten mit:

 oder **Edit > Drag Properties** oder **Strg+G**

den Eigenschaftstransfer, dessen Umfang in folgender Dialogbox festgelegt wird:

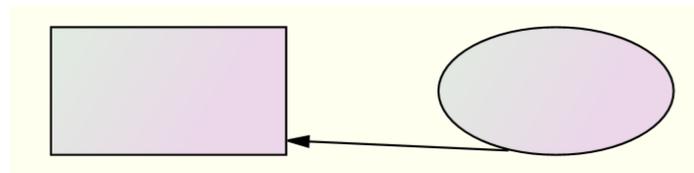


Klicken Sie dann auf das Rechteck und ziehen Sie mit gedrückter Maustaste die festgelegten Eigenschaften auf die Ellipse.

Ergänzen Sie den gerichteten Pfeil für den Effekt der Residualvariablen auf das Kriterium nach



ausnahmsweise nachlässig,

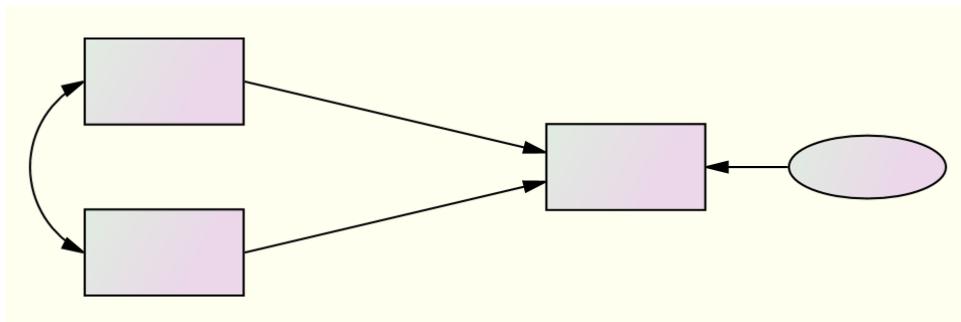


und lassen Sie ihn nach



oder **Edit > Touch Up** oder **Strg+H**

mit Mausklicks auf Rechteck und Ellipse von Amos in die richtige Position bringen. Auf analoge Weise lassen sich die gerichteten Pfeile von den Regressoren auf das Kriterium und der Kovarianz-Doppelpfeil zwischen den beiden Regressoren bei geringem Aufwand optimal platzieren, so dass folgendes Zwischenergebnis resultiert:



Zum bequemen Benennen der manifesten Variablen ordnen wir spätestens jetzt über



oder **File > Data Files** oder **Strg+D**

dem Projekt seine Eingabedatei zu. Dann öffnen wir mit

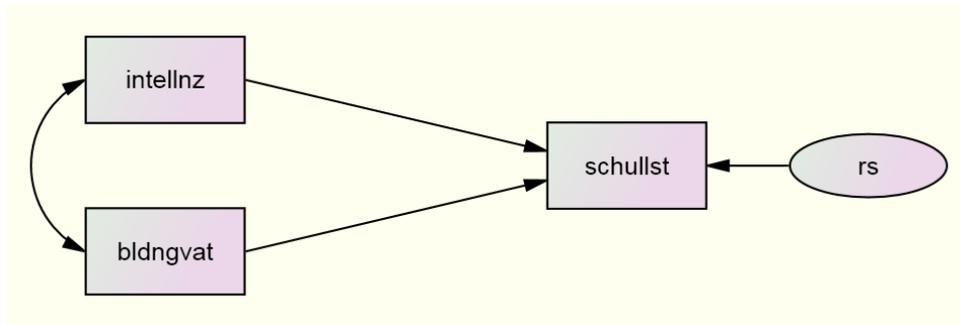


oder **View > Variables in Dataset** oder **Strg+⇧+D**

die folgende Variablenliste



und ziehen die benötigten Variablennamen per Maus auf das jeweils zugehörige Rechteck. Für die latente Residualvariable müssen wir per Eigenschaftsdialogbox einen Namen festlegen. So ergibt sich der nächste Entwicklungsstand des Pfaddiagramms:

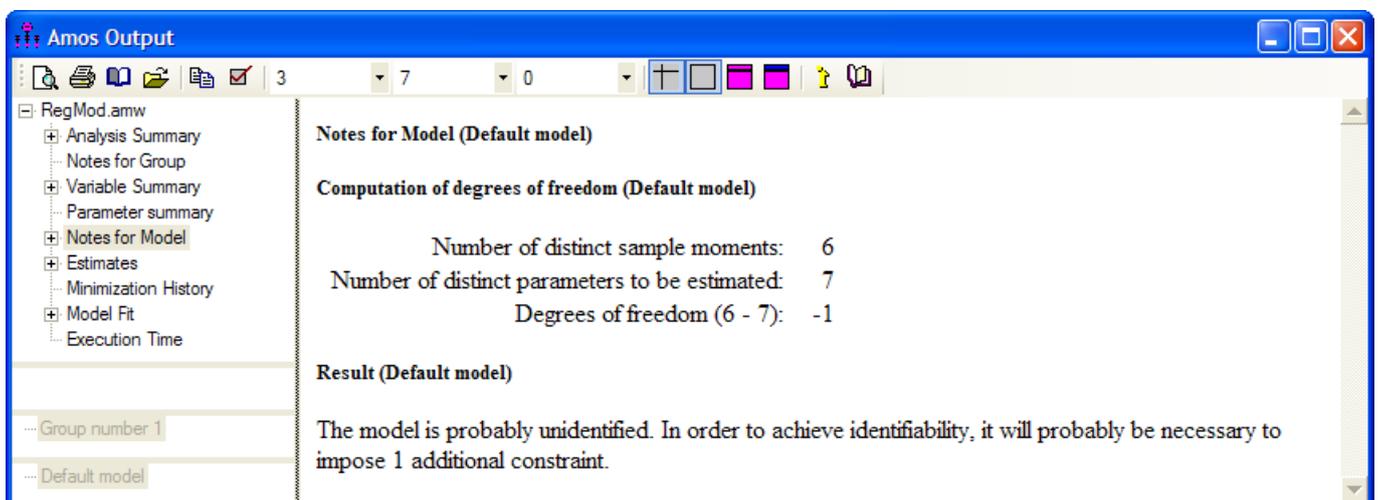


Wir verwenden die übliche, im Abschnitt 3.8 detailliert diskutierte Annahme, dass die Regressoren mit dem Residuum unkorreliert sind, so dass keine entsprechenden Doppelpfeile einzutragen sind. Kovarianzen zwischen den Regressoren und dem Kriterium sind natürlich erlaubt und erwartet, werden aber im Pfaddiagramm nicht dargestellt. Es sind keine Parameter, sondern *Implikationen* des Modells. Die Kovarianz zwischen den beiden Prädiktoren ist hingegen ein Modellparameter und muss durch einen Doppelpfeil repräsentiert werden, wenn sie nicht explizit auf den Wert Null fixiert werden soll.

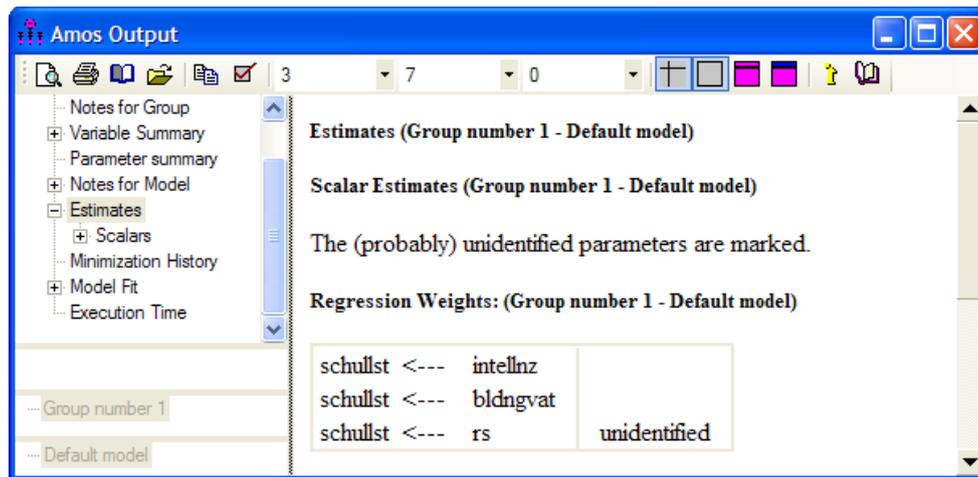
### 3.4 Ein technisches Identifikationsproblem

Wenn wir Amos nach obigen Vorarbeiten mit  zur Schätzung der Modellparameter auffordern, bleibt anschließend der  - Schalter zur Anzeige des Ausgabe-Pfaddiagramms passiv, und wir erhalten in der

Textausgabe im Abschnitt **Notes for Model** die folgende Informationen über ein Identifikationsproblem:



Amos informiert im Abschnitt **Estimates** über die Problemstelle:



Mit der in Amos per Voreinstellung frei schätzbaren Varianz der Fehlervariablen enthält das Modell sieben freie Parameter:

3 Varianzen:

**INTELLNZ**  
**BLDNGVAT**  
**RS**

1 Kovarianz:

**INTELLNZ <--> BLDNGVAT**

3 Strukturkoeffizienten (Pfadgewichte):

**SCHULLST <--- INTELLNZ**  
**SCHULLST <--- BLDNGVAT**  
**SCHULLST <--- RS**

Diesen stehen nur sechs Stichprobenmomente gegenüber (drei Varianzen und drei Kovarianzen), so dass keine Identifikation vorliegen kann.<sup>1</sup>

Insbesondere enthält das Modell mit dem Regressionsgewicht der Residualvariablen **RS** und mit deren Varianz zwei Parameter, die aufgrund der vorhandenen Information nicht simultan geschätzt werden können. Arbuckle (2009a, S. 70) illustriert die Situation mit folgender Analogie: Aus der Mitteilung „Ich habe Eintrittskarten im Wert von 50 € gekauft“ können wir nicht simultan die Anzahl der erworbenen Eintrittskarten und den Preis einer einzelnen Karte erschließen. Wenn wir aber Anzahl oder Preis kennen, ist der jeweils fehlende Wert leicht zu ermitteln. Bei unserem Identifikationsproblem ist allerdings nicht damit zu rechnen, dass wir einen der beiden unbekanntenen Werte irgendwie in Erfahrung bringen. Wir müssen daher *einen* Parameter *fixieren*, z.B. das Regressionsgewicht auf den Wert Eins. Damit wird keine empirische Behauptung aufgestellt, sondern lediglich eine Unbestimmtheit aus dem Modell beseitigt.

Zur Beseitigung dieser Unbestimmtheit kann man oft zwischen zwei Methoden wählen, die (abgesehen von Modellen mit einer so genannten *constraint interaction*, siehe unten) äquivalent sind:

- **Unit Loading Identification (ULI)**  
Das Regressionsgewicht wird auf den Wert Eins fixiert.
- **Unit Variance Identification (UVI)**  
Die Varianz der latenten Fehlervariablen wird auf den Wert Eins fixiert.

<sup>1</sup> Die Tatsache, dass wir an Stelle der Kovarianzmatrix eine Korrelationsmatrix verwenden mussten, ändert nichts an dieser Kalkulation.

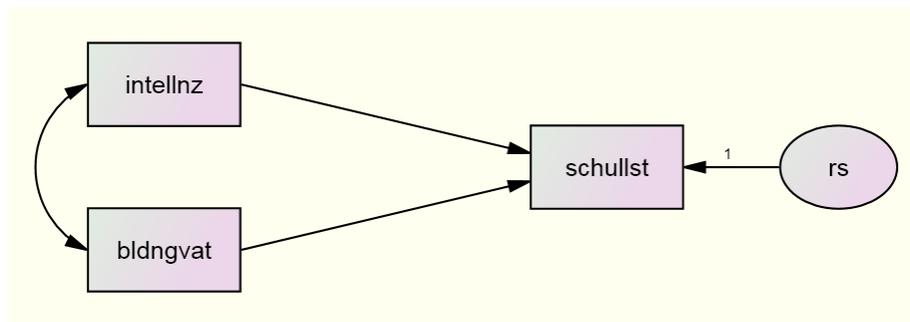
Das eben beschriebene Identifikationsproblem ist leicht zu verstehen, und eine Lösung ergibt sich praktisch von selbst. In komplexeren Modellen treten aber Identifikationsprobleme mit weniger trivialen Ursachen und Lösungen auf, so dass wir uns später noch mehrfach mit dem Thema beschäftigen müssen (siehe vor allem Abschnitt 5.3).

Weil es häufig nötig ist, eine endogene (abhängige) Variable mit einem Fehlerterm zu versorgen, bietet Amos über

 oder **Diagram > Draw Unique Variable**

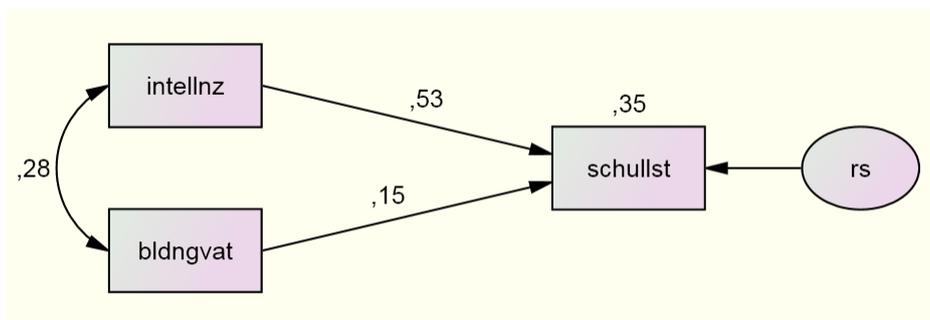
dazu ein spezielles Werkzeug an. Dabei wird das oben besprochene Identifikationsproblem automatisch per ULI-Technik behoben, d.h. das Regressionsgewicht der Fehlervariablen wird auf den Wert Eins fixiert. Um das Fehlerterm-Werkzeug ausprobieren zu können, müssen Sie die bereits vorhandene Fehlervariable mit dem  - Werkzeug wieder entfernen. Das  - Werkzeug produziert nach einem Klick auf das Kriterium ein kreisförmiges Fehlervariablensymbol und lässt dieses Symbol bei weiteren Mausklicks um das Kriterium wandern. Nachdem Sie eine gute Position gefunden haben, können Sie den Kreis bei Bedarf mit dem  - Werkzeug zu einer Ellipse verformen.

Nachdem die Residualvariable per Eigenschaftsdialog einen Namen erhalten hat, ist die Modellspezifikation komplett:



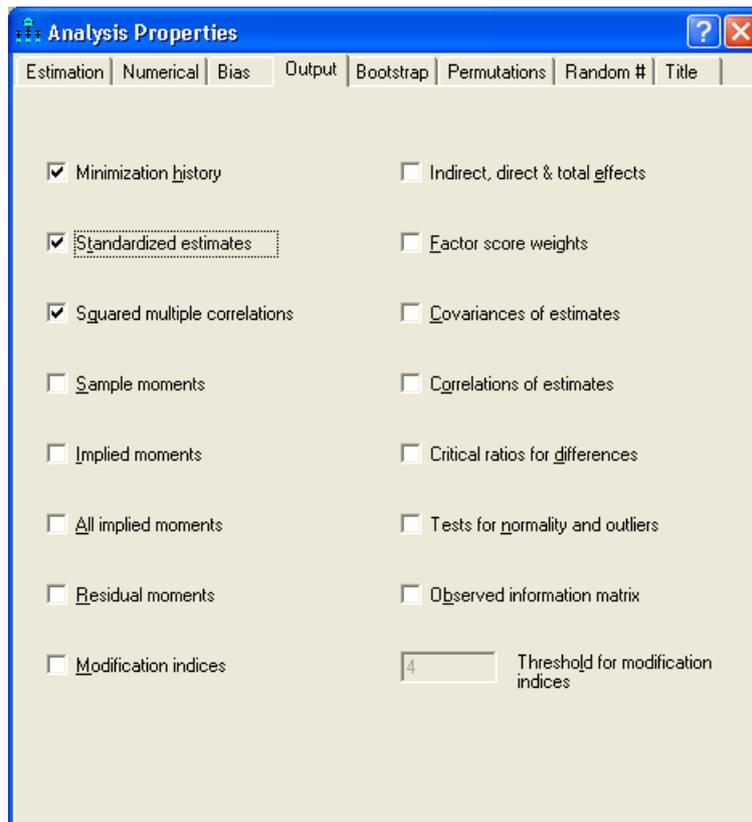
### 3.5 Standardisierte Schätzwerte und Determinationskoeffizienten

Nun steht einer erfolgreichen Schätzung nichts mehr im Weg. Beim Ausgabe-Pfaddiagramm wählen wir die *standardisierte* Variante, weil Amos in diesem Fall den **Determinationskoeffizienten** der Regressionsgleichung über dem Kriterium notiert:



Ansonsten zeigen die beiden Varianten des Ausgabe-Pfaddiagramms keine relevanten Unterschiede, weil wir eine *Korrelationsmatrix* als Datenbasis verwendet haben.

Wenn Sie in Ihrer Amos-Ausgabe die standardisierten Schätzer und/oder die quadrierte Multiple Korrelation des Kriteriums mit den Regressoren *nicht* finden, dann müssen Sie deren Ausgabe nach  oder **View > Analysis Properties > Output** anordnen (vgl. Abschnitt 2.5):



### 3.6 Textausgaben zum Regressionsmodell

Das eben formulierte Regressionsmodell ist identifiziert, erlaubt also für jeden Parameter eine eindeutige Schätzung (siehe Abschnitt 5.3). Weil das Modell gerade so viele freie Parameter besitzt, wie Stichprobenmomente vorhanden sind, nennt man es *genau identifiziert* oder *saturiert*. In dieser Situation liefert Amos im Wesentlichen dieselben Ergebnisse wie Programme zur multiplen linearen Regression, die mit dem OLS-Schätzverfahren (*Ordinary Least Squares*) arbeiten. In der Textausgabe finden sich u.a. die folgenden Tabellen:

**Scalar Estimates (Group number 1 - Default model)**

**Maximum Likelihood Estimates**

**Regression Weights: (Group number 1 - Default model)**

			Estimate	S.E.	C.R.	PLabel
schullst	<---	intellnz	,531	,030	17,487	***
schullst	<---	bldngvat	,147	,030	4,832	***

**Covariances: (Group number 1 - Default model)**

			Estimate	S.E.	C.R.	PLabel
intellnz	<-->	bldngvat	,277	,037	7,388	***

**Squared Multiple Correlations: (Group number 1 - Default model)**

	Estimate
schullst	,347

Ein saturiertes Modell impliziert *keine* Restriktionen für die Kovarianzmatrix der manifesten Variablen, so dass auch kein Modellgültigkeitstest möglich ist.

Mit der Regressionsprozedur von SPSS 18 erhalten wir nahezu identische Ergebnisse:<sup>1</sup>

**Modellzusammenfassung**

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	,589 <sup>a</sup>	,347	,345	,8090876

a. Einflußvariablen : (Konstante), intellnz, bldngvat

**Koeffizienten<sup>a</sup>**

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz
		B	Standardfehler	Beta		
1	(Konstante)	,000	,029		,000	1,000
	bldngvat	,147	,030	,147	4,826	,000
	intellnz	,531	,030	,531	17,464	,000

a. Abhängige Variable: schullst

Wir beobachten insbesondere, dass in der sehr großen Kerchoff-Stichprobe ( $N = 767$ ) die von Amos berechneten **Critical Ratios** annähernd mit den t-Werten der SPSS-Ausgabe übereinstimmen.

### 3.7 Pfad- bzw. Strukturgleichungsanalyse mit Amos

Das am Anfang von Abschnitt 3 formulierte Modell zu den beruflichen Ambitionen von Jugendlichen besteht aus *zwei* Regressionsgleichungen und behauptet (weniger auffällig), dass die Residualvariablen zu den beiden Kriterien unkorreliert sind. Damit liegt ein **rekursives** Modell vor (siehe Seite 6). Für die Modellierung empirischer Systeme durch rekursive Strukturgleichungsmodelle hat sich der Ausdruck **Pfadanalyse** eingebürgert. Unter diesen Voraussetzungen können alle Parameter mit zwei unabhängig voneinander durchgeführten Regressionsanalysen geschätzt und getestet werden, so dass scheinbar keine Strukturgleichungstechnologie benötigt wird.

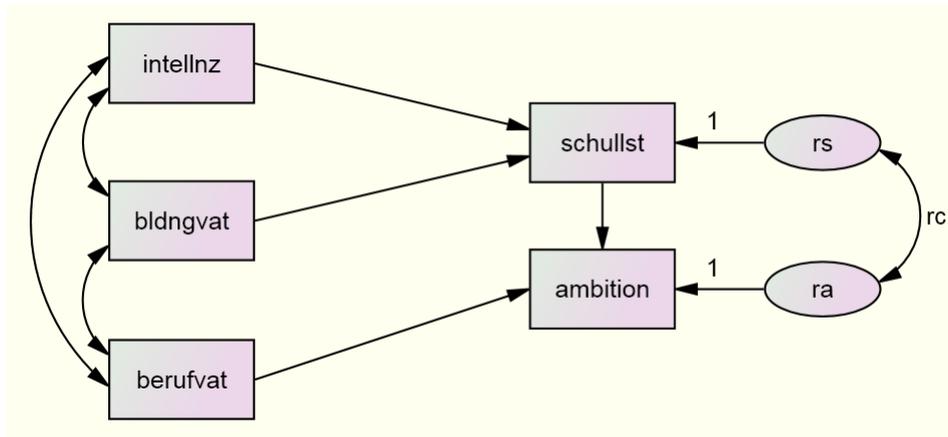
Bei der klassischen Regressionsanalyse haben wir aber keine Möglichkeit, die Annahme unabhängiger Residuen **RS** und **RA** zu überprüfen und ggf. eine vorhandene Residualkorrelation zu berücksichtigen. Dabei handelt es sich keinesfalls um eine nebensächliche Vereinbarung, sondern um die empirische Behauptung, dass für die Korrelation von **SCHULLST** und **AMBITION** ausschließlich der Effekt von **SCHULLST** auf **AMBITION** verantwortlich ist. Wir negieren damit z.B. die Existenz von exogenen Variablen mit Einfluss auf **SCHULLST** und **AMBITION**, die im Modell weggelassen wurden, und riskieren verzerrte Schätzungen für den Strukturkoeffizienten  $\gamma_{22}$ , falls wir uns irren.

In einem Strukturgleichungsmodell kann die Kovarianz von Residuen aus verschiedenen Gleichungen explizit berücksichtigt, d.h. geschätzt und getestet werden. Allerdings setzt die Berücksichtigung einer Fehlerkovarianz voraus, dass alle Parameter des Strukturgleichungsmodells identifiziert sind, also eindeutig geschätzt werden können (siehe Abschnitt 5.3). Im aktuellen Beispiel ist die Identifikation gegeben.

Wir vervollständigen das Pfaddiagramm, tragen den kritischen Kovarianzpfel ein und vergeben (im Hinblick auf eine spätere zu formulierende Restriktion) einen Namen für den zugehörigen Parameter:

<sup>1</sup> Zur Analyse der Matrix-Arbeitsdatei wurde folgende Syntax verwendet:

```
REGRESSION
MATRIX=IN(*)
/DEPENDENT schullst
/METHOD=ENTER bldngvat intellnz.
```



Durch eine von Null verschiedene Residuenkorrelation entsteht ein *nonrekursives* Modell. Aus dem zu Beginn von Abschnitt 3 formulierten Pfadmodell ist ein allgemeines Strukturgleichungsmodell geworden, und dessen Parameter können mit einer Regressionsprozedur nicht korrekt geschätzt werden.

Im Unterschied zum Regressionsmodell in Abschnitt 3.6 ist das jetzt betrachtete Strukturgleichungsmodell *nicht* saturiert, enthält also potentiell fehlerhafte Restriktionen. Es wird z.B. behauptet, dass **INTELLNZ** und **BLDNGVAT** keinen direkten Effekt auf **AMBITION** haben. Stattdessen wird ein vom Mediator **SCHULLST** vermittelter indirekter Effekt angenommen. Zwar können auch im Rahmen von Regressionsanalysen Nullhypothesen über die Abwesenheit direkter Effekte geprüft werden (einzeln oder auch blockweise), doch erlaubt die Strukturgleichungsanalyse mit Amos eine simultane Berücksichtigung beider Gleichungen.

Wir erhalten in der Textausgabe einen Modellgültigkeitstest mit zwei Freiheitsgraden, der das Pfadmodell (mit frei schätzbarer Kovarianz der beiden Residuen) freundlich beurteilt ( $p = 0,273$ ):

**Result (Default model)**

Minimum was achieved  
 Chi-square = 2,596  
 Degrees of freedom = 2  
 Probability level = ,273

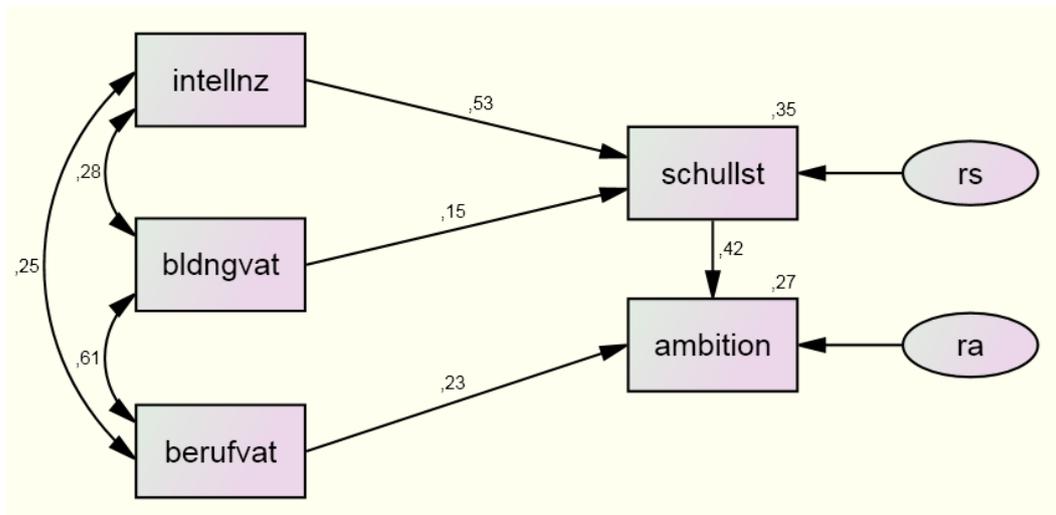
Damit steht einer Interpretation des approximativen Signifikanztests zur Fehlerkovarianz nichts mehr im Weg:

**Covariances: (Group number 1 - Default model)**

			Estimate	S.E.	C.R.	P	Label
intellnz	<-->	bldngvat	,277	,037	7,388	***	
bldngvat	<-->	berufvat	,610	,042	14,430	***	
intellnz	<-->	berufvat	,250	,037	6,713	***	
ra	<-->	rs	-,081	,045	-1,824	,068	rc

Zwar liegt die empirische Überschreitungswahrscheinlichkeit mit  $p = 0,068$  nur knapp über der Signifikanzgrenze, doch kommt dieses Ergebnis bei geringer Effektstärke (Korrelation(**RA**, **RS**) = -0,118) hauptsächlich aufgrund der enormen Stichprobengröße ( $N = 767$ ) zustande. Außerdem müsste für die *negative* Fehlerkorrelation eine inhaltliche Begründung gefunden werden. Daher kann die Nullhypothese unkorrelierter Residuen guten Gewissens beibehalten werden.

Für das nach obigen Überlegungen gültige und seiner Einfachheit wegen zu bevorzugende Pfadmodell ermittelt Amos die folgenden standardisierten Schätzwerte:



Zu den beiden Kriteriumsvariablen erhalten wir Determinationskoeffizienten (quadrierte multiple Korrelationen) in Höhe von 0,35 bzw. 0,27.

Die Regressionsprozedur von SPSS 18 kommt erwartungsgemäß bei der separaten Schätzung der beiden Regressionsgleichungen zu praktisch identischen Ergebnissen. Nachdem die SPSS-Ausgabe für die Kriteriumsvariable SCHULLST schon in Abschnitt 3.6 zu sehen war, folgt hier noch die Ausgabe für die Kriteriumsvariable AMBITION:

**Modellzusammenfassung**

Modell	R	R-Quadrat	Korrigiertes R-Quadrat	Standardfehler des Schätzers
1	,526 <sup>a</sup>	,277	,275	,8516538

a. Einflußvariablen : (Konstante), schullst, berufvat

**Koeffizienten<sup>a</sup>**

Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz
		B	Standardfehler	Beta		
1	(Konstante)	,000	,031		,000	1,000
	berufvat	,226	,032	,226	7,127	,000
	schullst	,422	,032	,422	13,281	,000

a. Abhängige Variable: ambition

Amos liefert den folgenden Gültigkeitstest für das komplette Modell, der mit  $p = 0,112$  seine Nullhypothese beibehält:

**Result (Default model)**

Minimum was achieved  
 Chi-square = 5,995  
 Degrees of freedom = 3  
 Probability level = ,112

Einen solchen Test kann eine Regressionsprozedur ebenso wenig liefern wie eine Prüfung der Annahme unkorrelierter Residuen.

In Amos steht zur Prüfung der Fehlerkovarianz neben dem CR-Test noch ein alternativer Test über eine  $\chi^2$ -Differenz zu zwei geschachtelten Modellen zur Verfügung (siehe Abschnitte 2.10 und 5.5), der im Allgemeinen eine höhere Präzision besitzt. Im konkreten Fall erhalten wir eine  $\chi^2$ -Differenz von

$$5,995 - 2,596 = 3,399$$

Unter der Nullhypothese folgt die Prüfgröße approximativ einer  $\chi^2$ -Verteilung mit einem Freiheitsgrad, so dass wir eine Überschreitungswahrscheinlichkeit von  $p = 0,065$  erhalten. Damit kommt der  $\chi^2$ -Differenz zum selben Ergebnis wie der oben vorgestellte CR-Test.

### 3.8 Kausalitätstheoretische Annahmen bei Strukturgleichungsmodellen

Da für ein Strukturgleichungsmodell meist die kausale Interpretierbarkeit seiner Parameter beansprucht wird, müssen wir auf das methodologisch und wissenschaftstheoretisch anspruchsvolle Thema der kausalen Interpretierbarkeit statistischer Ergebnisse eingehen. Für eine gründliche Behandlung der Thematik wird auf Bollen (1989, Kap. 3) und Steyer (1988, 1992) verwiesen.

Um die Strukturkoeffizienten der *exogenen* Variablen einer Modellgleichung **kausal** interpretieren zu können, wird meist für diese Variablen **Unkorreliertheit mit dem Residuum** angenommen.<sup>1</sup> Bollen (1989, S. 43) spricht in diesem Zusammenhang von der **Pseudo-Isolation**. Das Residuum enthält neben Messfehlern auch die Einflüsse von nicht einbezogenen Variablen, die oft als **potentielle Störvariablen** bezeichnet werden. Die Unkorreliertheit sorgt unter gewissen Zusatzannahmen für die **Invarianz** der geschätzten Strukturkoeffizienten bei einer Erweiterung der Modellgleichung um zusätzliche Einflussgrößen, die vorher Bestandteil des Residuums waren. Diese Invarianz ist das zentrale Kriterium moderner Kausalitätsbegriffe.

In der Experimentalmethodologie wird die Unkorreliertheitsvoraussetzung und damit die Invarianzbedingung durch **Randomisierung**, d.h. durch zufällige Auswahl der bei einem Fall zu realisierenden experimentellen Bedingung sichergestellt. In den Wirtschafts- und Sozialwissenschaften ist die Experimentalmethodologie häufig aus ethischen und/oder pragmatischen Gründen nicht einsetzbar. Bei der ersatzweise gewählten Strukturgleichungsanalyse von beobachteten Variablen muss dann die Unkorreliertheit der exogenen Variablen mit den potentiellen Störvariablen direkt angenommen werden. Geschieht dies zu Unrecht, so können verzerrte Schätzer für kausale Effekte resultieren. Dies wird nun an einem sehr einfachen Strukturgleichungsmodell demonstriert, dessen einzige Gleichung die Abhängigkeit einer endogenen Variablen  $Y$  von einer exogenen Variable  $X$  beschreibt: Ein nicht sehr aufgeklärter Forscher hat die Hypothese aufgestellt, dass die Geburtenrate (Anzahl der jährlichen Geburten pro 10000 Einwohner, Abk.:  $Y$ ) von der Anzahl der Klapperstörche (Abk.:  $X$ ) abhängt und möchte anhand der Daten aus schwedischen Landkreisen die Stärke  $\gamma$  des kausalen Effekts von  $X$  auf  $Y$  schätzen. Sein Modell lautet:

$$Y = \gamma X + R$$

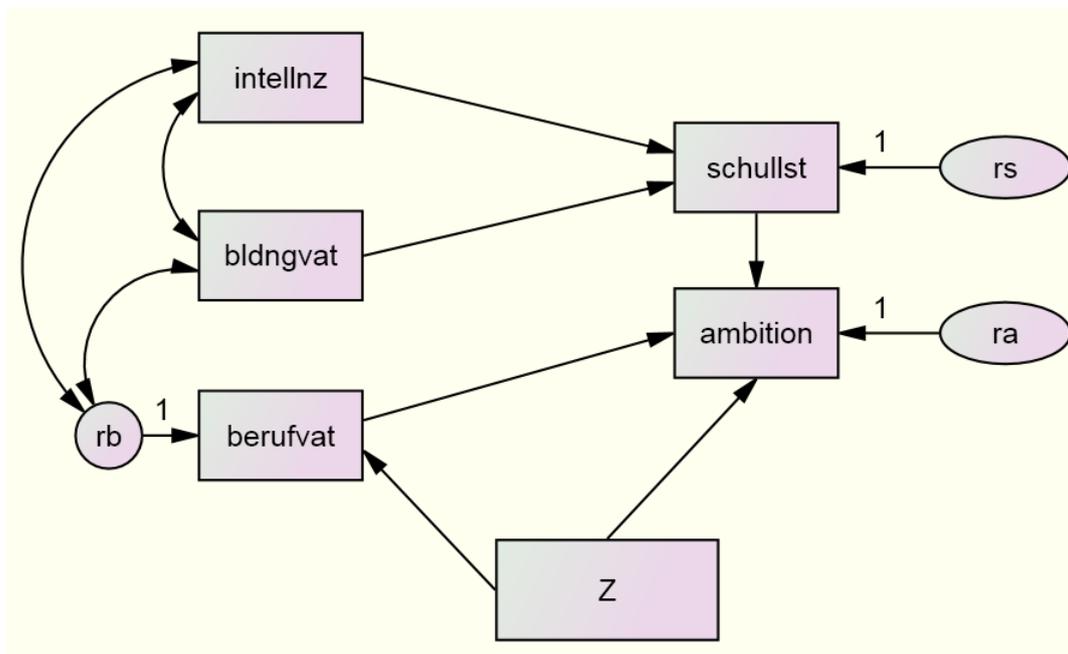
Unter der kritischen Annahme, dass die exogene Variable  $X$  nicht mit der Residualvariablen  $R$  korreliert, schätzt der Forscher aus einer Stichprobenkovarianzmatrix mit den beobachteten Variablen  $X$  und  $Y$  den Effekt  $\hat{\gamma} = 0,6$ . Er folgert, dass die Zunahme der Storchenpopulation um ein Exemplar eine Steigerung der Geburtenrate um ca. 0,6 bewirkt. Die Residualvariable  $R$  enthält in unserem Beispiel u.a. als bedeutende Einflussgröße den Industrialisierungsgrad, welcher die Geburtenrate negativ beeinflusst. Seine Nichtberücksichtigung allein gefährdet die Interpretierbarkeit von  $\hat{\gamma}$  noch nicht. Allerdings ist der Industrialisierungsgrad ebenfalls negativ mit der Anzahl der Störche korreliert. Die Unkorreliertheitsbedingung ist also verletzt, und damit darf  $\hat{\gamma}$  nicht als kausaler Effekt von  $X$  auf  $Y$  interpretiert werden. In der Tat ist  $\hat{\gamma}$  eine drastisch verzerrte Schätzung für den wahren Effekt, der vermutlich exakt Null beträgt. Offenbar führt der

<sup>1</sup> Mit *endogenen* Variablen in explanatorischer Position werden wir uns gleich beschäftigen

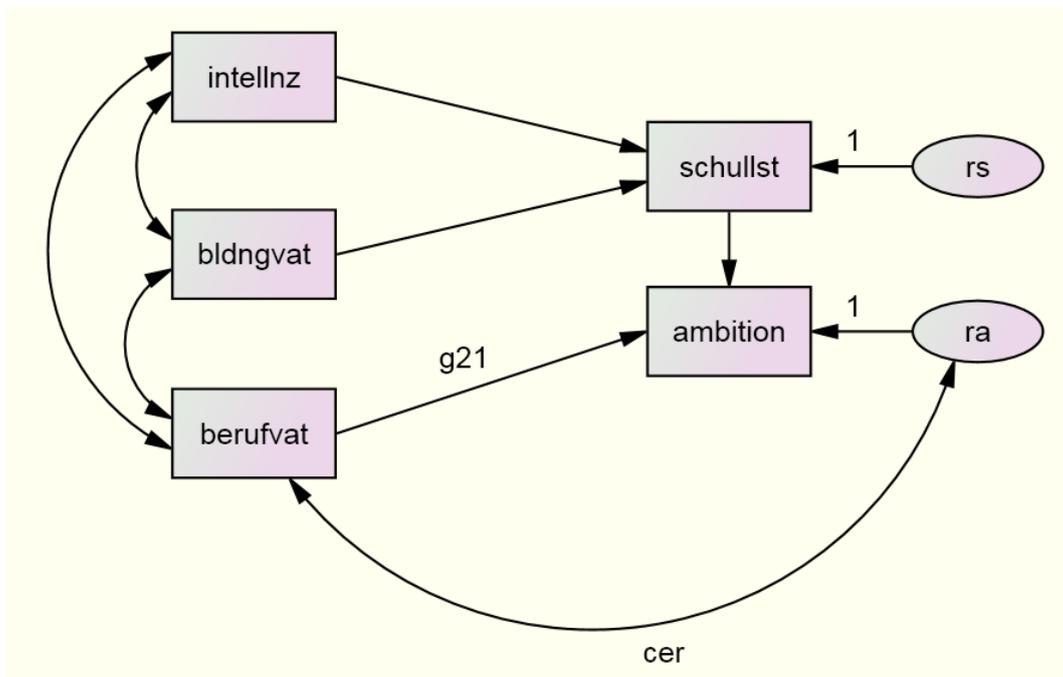
negative Effekt des Industrialisierungsgrades auf  $Y$  und  $X$  zu einer positiven Korrelation zwischen diesen Variablen. Daraus resultiert bei einer Analyse unter der falschen Annahme der Unkorreliertheit von  $X$  und  $R$  die unsinnige Effektschätzung  $\hat{\gamma}$ . Ihre Nichtigkeit erweist sich z.B. bei Erweiterung obiger Strukturgleichung um die unabhängige Variable Industrialisierungsgrad. Dann wird vermutlich für  $\gamma$  eine Schätzung nahe Null resultieren, d.h. die Invarianzbedingung gilt nicht.

Fehlt in einer Modellgleichung eine relevante Einflussgröße, die mit vorhandenen exogenen Variablen korreliert ist, dann liegt ein gravierender Spezifikationsfehler vor.

Unter bestimmten Voraussetzungen (siehe Kenny 1979, S. 93ff) können nicht gemessene Störvariablen in einem Strukturgleichungsmodell über eine Kovarianz zwischen dem Residuum und exogenen Variablen berücksichtigt werden. Hier wird die zu Beginn des Abschnitts genannte Generalvoraussetzung für die kausale Interpretierbarkeit der Strukturkoeffizienten von exogenen Variablen also *nicht* benötigt. Zur Lösung der dabei auftretenden Identifikationsprobleme werden aber im Modell zusätzliche exogene Variablen *ohne direkten* Einfluss auf das Kriterium benötigt (so genannte **Instrumentalvariablen**). Im Ambitionsmodell von Kerchoff (1974, siehe z.B. Abschnitt 3.7) könnte eine nicht gemessene Störvariablen **Z** mit Einfluss auf das Kriterium **AMBITION** und die ursprünglich exogene Variable **BERUFVAT**



über eine Kovarianz zwischen der Variablen **BERUFVAT** dem Residuum zum Kriterium **AMBITION** berücksichtigt werden:



Hier sorgen die Instrumentalvariablen **INTELLNZ** und **BILDNGVAT** für eine Identifikation aller Modellparameter. Bei Simulationsstudien mit künstlichen Daten habe ich den Eindruck gewonnen, dass die kritischen Parameter **g21** und **cer** nur in großen Stichproben (> 1000 Fälle) präzise geschätzt werden können. Offenbar erschwert die Berücksichtigung einer Korrelation zwischen Regressor und Residuum das Schätzproblem, wobei das Ignorieren einer vorhandenen Korrelation keine Alternative darstellt.

Bei einer *endogenen* Variablen (z.B. **SCHULLST**), die in einer Modellgleichung in explanatorischer Position auftritt (im Beispiel: mit dem Kriterium **AMBITION**), kann ohne große Probleme für die Identifikation und Schätzbarkeit eine von Null verschiedene Korrelation mit dem Residuum (im Beispiel: **RA**) zugelassen und in das Modell einbezogen werden, damit die kausale Interpretierbarkeit des Strukturkoeffizienten nicht durch ausgelassene Variablen gefährdet ist. Unter der üblicherweise akzeptierten Voraussetzung, dass alle *exogenen* Variablen des Modells mit allen Residuen des Modells unkorreliert sind, muss lediglich eine Korrelation zwischen dem Ambitions-Residuum **RA** und dem Residuum **RS** in der Gleichung zur endogenen Variablen **SCHULLST** zugelassen werden. Die praktische Umsetzung dieser Modellierungstechnik kennen Sie schon aus Abschnitt 3.7. Nun wird ihre Begründung erweitert und ein Zusammenhang mit der Kausalitätsdebatte aufgezeigt.

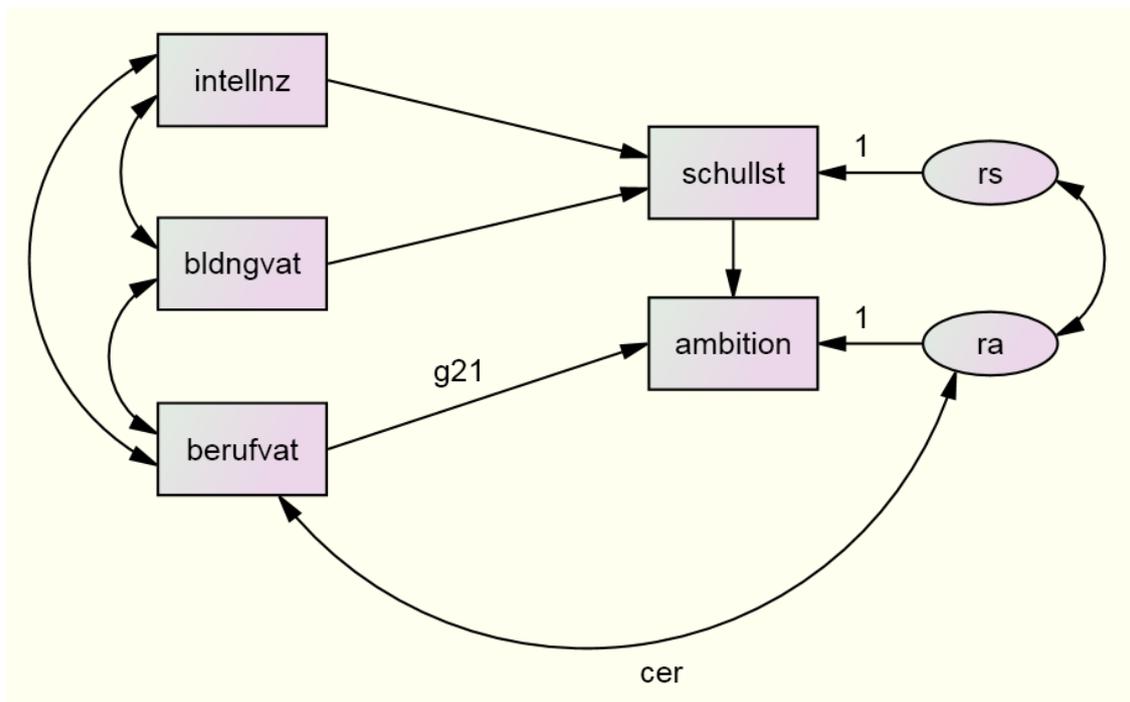
**SCHULLST** ist modellgemäß die Summe aus einer Linearkombination von exogenen Variablen und der Residualvariablen **RS**:

$$\text{SCHULLST} = \gamma_{11} \text{INTELLNZ} + \gamma_{12} \text{BLDNGVAT} + \text{RS}$$

Weil die exogenen Variablen voraussetzungsgemäß mit **RA** unkorreliert sind, ist **SCHULLST** genau dann mit dem Ambitions-Residuum **RA** korreliert, wenn die beiden Störvariablen **RS** und **RA** korreliert sind. Diese Aussagen folgen aus Theoremen über Varianzen, Kovarianzen und Korrelationen (siehe Abschnitt 5.1).

Wie Sie aus Abschnitt 3.7 wissen, kann die Kovarianz von **RS** und **RA** problemlos in ein Strukturgleichungsmodell als frei schätzbarer Parameter integriert und auf Signifikanz getestet werden. In der (sehr großen) Kerchhoff-Stichprobe wurde die Nullhypothese zur Fehlerkovarianz nicht verworfen, so dass wir auf den Parameter verzichtet haben.

Der Versuch, im Ambitionsbeispiel simultan für den exogenen Regressor BERUFVAT *und* für den endogenen Regressor SCHULLST eine Kovarianz mit dem Residuum zu AMBITION einzubeziehen, scheitert an einem Identifikationsproblem:



Bei einer **Regressionsanalyse mit OLS-Schätzung** (Ordinary Least Squares) ist zur Gewährleistung der kausalen Interpretierbarkeit der Koeffizienten zu fordern, dass alle unabhängigen Variablen einer Gleichung (gleichgültig ob exogen oder endogen in Bezug auf das Gesamtmodell) mit dem zugehörigen Residuum unkorreliert sind. Die im aktuellen Abschnitt diskutierten Modellerweiterungen (Kovarianz von **BERUFVAT** und **RA** bzw. Kovarianz von **RS** und **RA**) führen zu nonrekursiven Gleichungssystemen, so dass die OLS-Regressionen nicht anwendbar sind.

## 4 Konfirmatorische Faktorenanalyse

In einer Studie von Steyer (1988, S. 312) wurden bei 179 Studierenden im Abstand von zwei Monaten jeweils zwei Tests für die Zustandsangst erhoben. Zu jedem Zeitpunkt sind also zwei Indikatoren für die latente Variable Zustandsangst vorhanden. Diese wird als zeitlich variabel angenommen, so dass für die beiden Zeitpunkte unterschiedliche, gleichwohl korrelierte latente Variablen vorliegen. Insgesamt sind folgende Variablen zu modellieren:

- $X_1$  Erster Indikator, erhoben zum ersten Zeitpunkt
- $X_2$  Zweiter Indikator, erhoben zum ersten Zeitpunkt
- $X_3$  Erster Indikator, erhoben zum zweiten Zeitpunkt
- $X_4$  Zweiter Indikator, erhoben zum zweiten Zeitpunkt
- $\xi_1$  Latente Zustandsangst zum ersten Zeitpunkt (sprich: Ksi1)
- $\xi_2$  Latente Zustandsangst zum zweiten Zeitpunkt
- $\delta_1, \delta_2, \delta_3, \delta_4$  Residuen der Indikatorvariablen (sprich: Delta1 etc.)

Hier werden die zusammen mit dem Computerprogramm LISREL eingeführten Variablennamen verwendet (siehe z.B. Jöreskog & Sörbom 1989), um bei späteren Analysen des Beispiels kompakte Formeln zu erhalten.

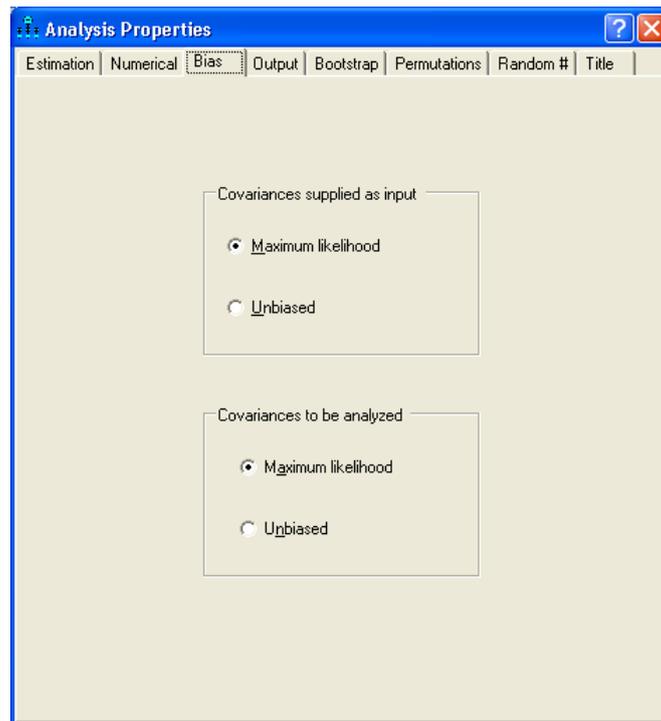
Die Kovarianzmatrix der manifesten Variablen finden Sie in der SPSS-Matrix-Datendatei **Steyer.sav** (Bezugsquelle: siehe Vorwort):

	rowtype_	varname_	X1	X2	X3	X4
1	N	X1	179,0000	179,0000	179,0000	179,0000
2	COV	X1	24,6700	21,8950	10,3530	11,6650
3	COV	X2	21,8950	25,1350	10,6240	12,6360
4	COV	X3	10,3530	10,6240	27,2390	25,2580
5	COV	X4	11,6650	12,6360	25,2580	28,6830

Eine Besonderheit dieser Datei besteht darin, dass sie *Stichprobenkovarianzen* enthält (ermittelt mit Stichprobengröße  $N$  als Nenner), also keine erwartungstreuen Schätzer der Populationsvarianzen (ermittelt mit dem Nenner  $N-1$ ). Amos verwendet für Maximum Likelihood – Schätzungen per Voreinstellung genau solche Stichprobenkovarianzen, interpretiert die Einträge in einer SPSS-Matrix-Datendatei aber als erwartungstreue Schätzer und konvertiert diese per Voreinstellung automatisch. In unserem Fall ist diese Automatik natürlich unerwünscht. Daher informieren wir Amos nach

### View > Analysis Properties > Bias

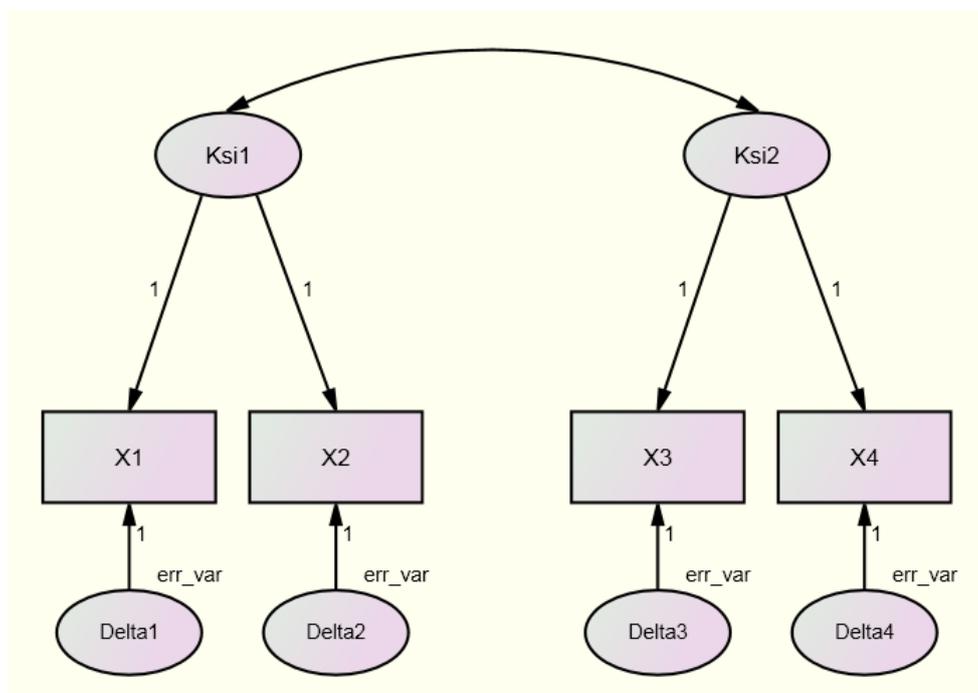
darüber, dass die von uns angelieferten **Covariances supplied as input** ohne Vorbehandlung für eine **Maximum Likelihood** - Schätzung geeignet sind:



Das zu prüfende Modell gehört zum Bereich der **konfirmatorischen Faktorenanalyse**, weil ...

- latente Variablen auftreten, die über Ladungspfade zur Erklärung von manifesten Indikatoren beitragen und auf diese Weise empirisch verankert werden,
- wobei die Beziehungen zwischen den latenten Variablen über eine frei schätzbare Kovarianzmatrix repräsentiert sind.

Das Modell



enthält zahlreiche Restriktionen:

- Es fehlen etliche potentiell mögliche Ladungspfade, und die Residuen sind unkorreliert. In der Terminologie von Kline (2005, S. 166f) liegt eine **Standard CFA** (*Confirmatory Factor Analysis*) vor:
  - Jeder Indikator wird nur von *einem* Faktor beeinflusst.
  - Die Residuen zu den Indikatoren sind unkorreliert.
- Die verbliebenen Ladungen sind alle auf den Wert Eins fixiert. Hier sind *zwei* Aussagen mit sehr verschiedenem Bedeutungsgehalt involviert:
  - In einer empirisch falsifizierbaren Behauptung werden alle Ladungen gleich gesetzt.
  - Durch die Wahl der Konstanten Eins (an Stelle eines Parameternamens) erhalten die latenten Variablen eine Maßeinheit. Hier wird keine empirisch falsifizierbare Behauptung aufgestellt, sondern ein technisches Identifikationsproblem gelöst (siehe Abschnitte 3.4 und 5.3).
- Von einer weiteren Gleichheitsrestriktion sind die vier Fehlervarianzen betroffen.

In Gleichungsform lautet das Modell:

$$\begin{array}{rcl}
 X_1 & = & \xi_1 \quad + \delta_1 \\
 X_2 & = & \xi_1 \quad + \delta_2 \\
 X_3 & = & \xi_2 \quad + \delta_3 \\
 X_4 & = & \xi_2 \quad + \delta_4 \\
 \text{Var}(\delta_1) & = & \text{Var}(\delta_2) = \text{Var}(\delta_3) = \text{Var}(\delta_4)
 \end{array}$$

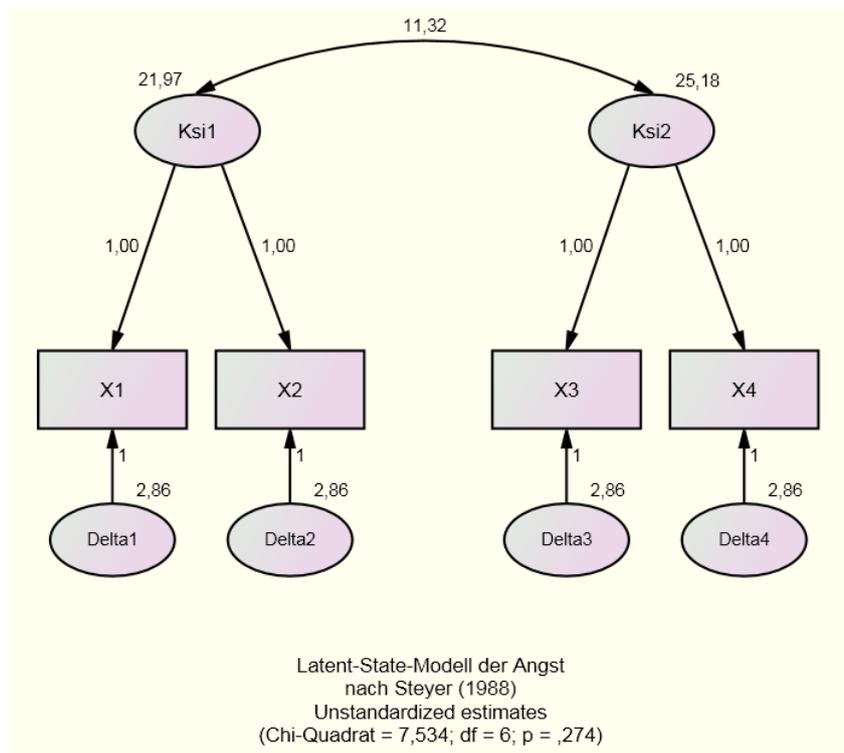
Das Modell enthält vier frei schätzbare Parameter:

$$\text{Var}(\xi_1), \text{Var}(\xi_2), \text{Cov}(\xi_1, \xi_2), \text{Var}(\delta_1)$$

Seine Aufgabe besteht darin, die zehn (Ko)varianzen zwischen den vier beobachteten Variablen ( $X_1, \dots, X_4$ ) zu erklären durch deren Abhängigkeit von zwei latenten Variablen und durch die (Ko)varianzen der latenten Variablen. Hier wird klar, warum die im Manuskript behandelte Methodologie auch als *Kovarianzstruktur-analyse* bezeichnet wird: Ein als empirische Gesetzmäßigkeit aufzufassende Kovarianzmatrix wird durch die Modellgleichungen erklärt, wobei die Anzahl der verwendeten Parameter möglichst klein sein sollte und nach oben durch die Anzahl der beobachteten (Ko)varianzen beschränkt ist.

Weil eine Standard-CFA - Modell mit zwei Faktoren und mit mindestens zwei Indikatoren pro Faktor vorliegt, sind alle Modellparameter *identifiziert*, können also eindeutig aus den Daten geschätzt werden (vgl. Abschnitt 5.3).

Wir fordern Amos über den Schalter  zur Berechnung der Schätzungen an und erhalten das folgende Ausgabe-Pfaddiagramm (unstandardisiert):

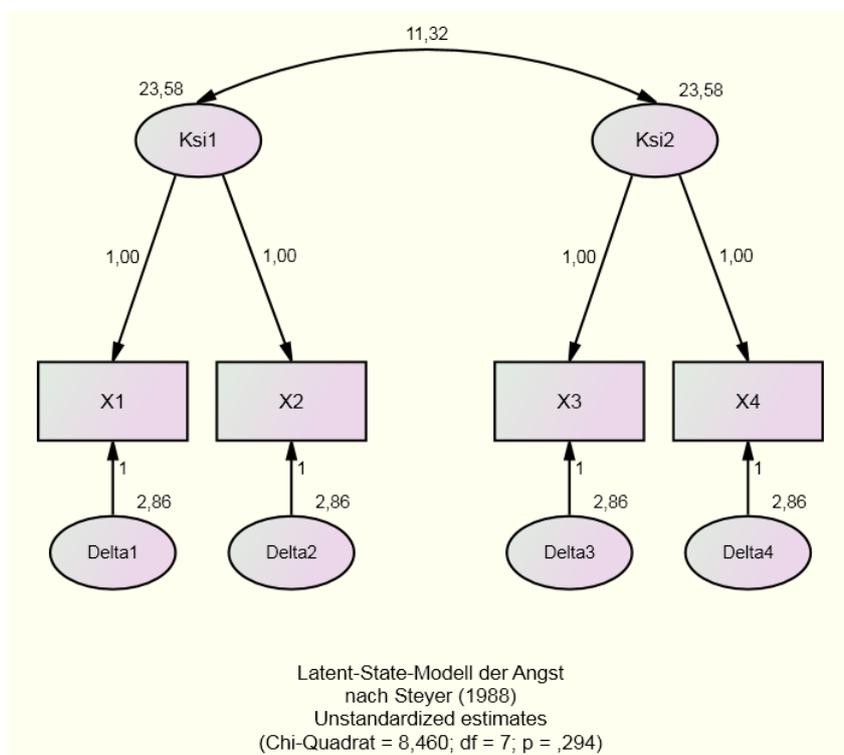


Der (bei  $N = 179$  relativ vertrauenswürdige)  $\chi^2$ -Test akzeptiert das Modell ( $p = 0,274$ ), so dass es als empirisch bewährt gelten kann.

Mit dem Ziel eines möglichst einfachen Modells könnte man die Varianzen der Faktoren gleich setzen:

$$\text{Var}(\xi_1) = \text{Var}(\xi_2)$$

Das nochmals vereinfachte Modell wird ebenfalls akzeptiert:



Nach dem in Abschnitt 5.5 vorzustellenden Prinzip der Testkonstruktion durch Vergleich von geschachtelten Modellen kann die Nullhypothese identischer Faktorvarianzen geprüft werden. Die  $\chi^2$ -Differenz aus beiden Modellen folgt unter der Nullhypothese einer  $\chi^2$ -Verteilung mit einem Freiheitsgrad (= *df*-Differenz). Das empirische Ergebnis

$$8,460 - 7,534 = 0,926$$

liegt deutlich unter dem kritischen Wert ( $\approx 4$  bei  $\alpha = 0,05$ ), so dass die Nullhypothese akzeptiert werden kann.

Wie ein Blick auf die standardisierte Lösung zeigt, ist die Korrelation der beiden Zustandsangstfaktoren theoriekonform mit 0,48 nur moderat, so dass die Annahme eines *einzigsten*, situationsunabhängigen bzw. zeitstabilen Faktors kaum plausibel ist. Nach einem Vorschlag von Kline (2005, S. 182f) lässt sich das Einfaktormodell als vereinfachte (geschachtelte) Variante des Zweifaktormodells betrachten, was einen Test der Einfaktor-Nullhypothese über die  $\chi^2$ -Differenz aus zwei geschachtelten Modellen erlaubt. Aus der Identität

$$\xi_1 \equiv \xi_2$$

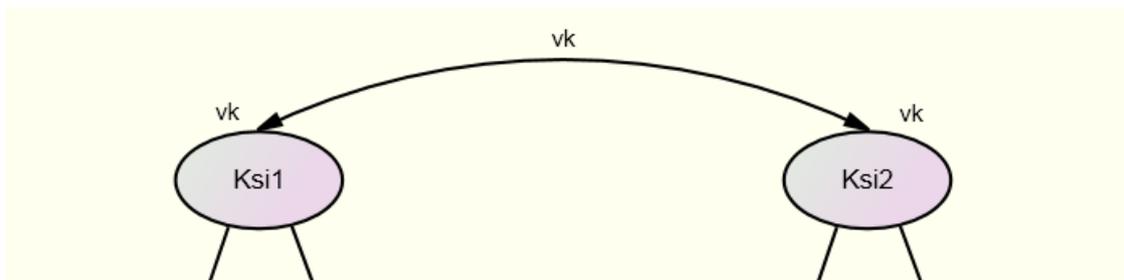
folgen die Parameterrestriktionen (vgl. Abschnitt 5.1):

$$\text{Cov}(\xi_1, \xi_2) = \text{Cov}(\xi_1, \xi_1) = \text{Var}(\xi_1) = \text{Var}(\xi_2)$$

Umgekehrt implizieren diese Restriktionen eine Korrelation von Eins zwischen den beiden Faktoren (siehe Abschnitt 5.1.5 wegen der Korrelationsdefinition):

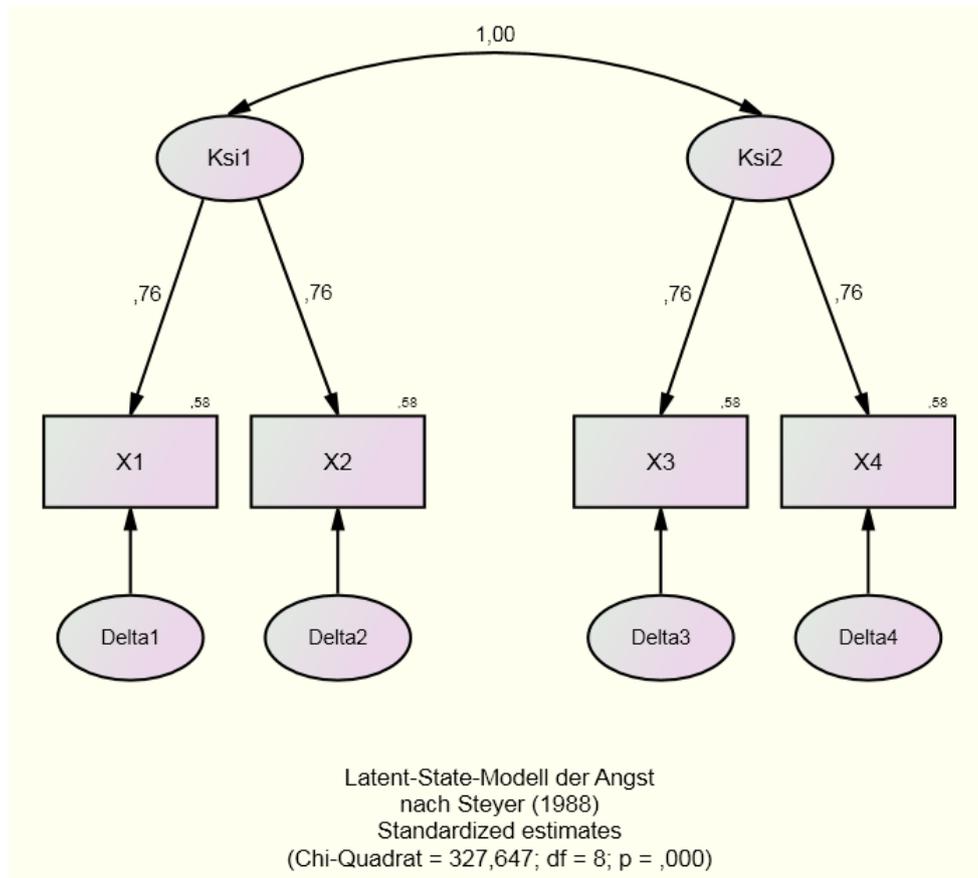
$$\text{Cor}(\xi_1, \xi_2) = \frac{\text{Cov}(\xi_1, \xi_2)}{\sqrt{\text{Var}(\xi_1) \text{Var}(\xi_2)}} = \frac{\text{Var}(\xi_1)}{\text{Var}(\xi_1)} = 1$$

Man kann also durch geschickte Parameterrestriktionen im Zweifaktormodell



ein Einfaktormodell als eingeschachtelte Variante formulieren.

Erwartungsgemäß wird im Beispiel das Einfaktormodell vom  $\chi^2$ -Test verworfen:



Im Angstbeispiel akzeptiert der  $\chi^2$ -Test das Zweifaktormodell und verwirft das Einfaktormodell, so dass die Frage nach der Faktorenzahl beantwortet ist. In der Praxis lehnt der  $\chi^2$ -Modellgültigkeitstest oft aber auch das Ausgangsmodell (hier: mit zwei Faktoren) ab, obwohl es eine sinnvolle Approximation der Empirie darstellt und von alternativen Modellgütekriterien (vgl. Abschnitt 5.6) günstig bewertet wird. Dann kann die Einfaktor-Nullhypothese über einen Vergleich geschachtelter Modelle gemäß Abschnitt 5.5 geprüft werden. Im Angstbeispiel erhalten wir bei einer *df*-Differenz von Eins eine extreme  $\chi^2$ -Differenz weit oberhalb des kritischen Werts ( $\approx 4$  bei  $\alpha = 0,05$ ), so dass die Einfaktor-Nullhypothese abzulehnen ist:

$$327,647 - 8,460 = 319,187$$

## 5 Grundlagen und Vertiefungen zur Strukturgleichungsanalyse

Bisher wurden die methodischen Grundlagen der Strukturgleichungsanalyse nur sehr zurückhaltend angesprochen, weil die Leser zunächst mit dem Programm Amos und mit typischen Anwendungsbeispielen vertraut werden sollten. In diesem Abschnitt wird unter Bezug auf die vorgestellten Beispiele ein vertieftes Verständnis für Theorie und Praxis der Strukturgleichungsanalyse vermittelt.

### 5.1 Einige Grundlagen aus der Wahrscheinlichkeitstheorie

Eine elementare Annahme der Strukturgleichungsanalyse lautet, dass die Verteilung der beobachteten Variablen hinreichend gut durch die **Momente erster und zweiter Ordnung (d.h. durch Erwartungswerte, Varianzen und Kovarianzen)** beschrieben werden kann. Diese Annahme ist im Fall multivariater Normalverteilung der manifesten Variablen perfekt erfüllt. Vermutlich erinnern Sie sich noch an die Bedeutung des Programmnamens Amos:

Analysis of **MO**ment Structures

Zur Wiederholung sollen die Definitionen dieser Momente und ihre wichtigsten Eigenschaften notiert werden, wobei auf mathematische Exaktheit verzichtet wird.

#### 5.1.1 Erwartungswert

Der Erwartungswert einer Zufallsvariablen wird in diesem Manuskript selten explizit betrachtet, weil durchweg für alle Variablen Zentriertheit (d.h. ein Erwartungswert von Null) angenommen wird. Jedoch basieren viele später einzuführende Definitionen auf dem Erwartungswert, so dass unser kleiner Exkurs in die mathematische Statistik mit diesem elementaren Begriff beginnen soll.

Für eine diskrete Zufallsvariable  $X$  mit der Verteilung  $\{x_i, P(X=x_i)\}_{i=1,2,\dots}$  ist der Erwartungswert definiert durch:

$$E(X) := \sum_i x_i P(X = x_i)$$

Für eine stetige (kontinuierliche) Zufallsvariable  $X$  mit der Dichte  $f$  ist der Erwartungswert definiert durch:

$$E(X) := \int_{-\infty}^{+\infty} x f(x) dx$$

Offenbar ist die Integralformel die natürliche Verallgemeinerung der Summenformel für den Fall überabzählbar unendlich vieler Variablenausprägungen. Mathematisch weniger interessierte Leser brauchen sich aber keine Sorgen zu machen: Die Integralrechnung ist für das Verständnis dieses Kurses irrelevant.

#### 5.1.2 Varianz

Die Varianz einer Zufallsvariablen  $X$  ist definiert durch:

$$\text{Var}(X) := E[(X - E(X))^2]$$

Aus der Definition folgt durch Ausmultiplizieren und wegen der Linearität des Erwartungswertes (d.h.:  $E(aX) = aE(X)$ ,  $a \in \mathbb{R}$ ;  $E(X+Y) = E(X) + E(Y)$ ) die so genannte **Verschiebungsformel**:

$$\text{Var}(X) = E(X^2) - E(X)^2$$

Bei  $E(X) = 0$  gilt also:

$$\text{Var}(X) = E(X^2)$$

### 5.1.3 Kovarianz

Die Kovarianz der beiden Zufallsvariablen  $X$  und  $Y$  ist definiert durch:

$$\text{Cov}(X, Y) := E[(X - E(X))(Y - E(Y))]$$

Aus der Definition folgt (mit denselben Argumenten wie bei der Varianz):

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y),$$

so dass bei  $E(X) = E(Y) = 0$  gilt:

$$\text{Cov}(X, Y) = E(XY).$$

Durch Einsetzen von  $X$  für  $Y$  sieht man außerdem sofort, dass die Varianz der Zufallsvariablen  $X$  gerade die Kovarianz von  $X$  mit sich selber ist.

### 5.1.4 Kovarianzmatrix eines Zufallsvektors

Ist  $\mathbf{X} := (X_1, X_2, \dots, X_q)^T$ ,  $q \in \mathbb{N}$  ein Spaltenvektor aus Zufallsvariablen, so wird die Kovarianzmatrix  $\text{Cov}(\mathbf{X})$  von  $\mathbf{X}$  wie folgt definiert:

$$\text{Cov}(\mathbf{X}) := \begin{bmatrix} \text{Var}(X_1) & & & & \\ \text{Cov}(X_2, X_1) & \text{Var}(X_2) & & & \\ \text{Cov}(X_3, X_1) & \text{Cov}(X_3, X_2) & \text{Var}(X_3) & & \\ \cdot & \cdot & \cdot & \cdot & \\ \text{Cov}(X_q, X_1) & \text{Cov}(X_q, X_2) & \text{Cov}(X_q, X_3) & \cdot & \text{Var}(X_q) \end{bmatrix}$$

Die obere Dreiecksmatrix ist dabei aus Symmetriegründen weggelassen.

### 5.1.5 Korrelation

Die Korrelation der beiden Zufallsvariablen  $X$  und  $Y$  ist definiert durch:

$$\text{Cor}(X, Y) := \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \text{Var}(Y)}}$$

Bei standardisierten Variablen (d.h.  $\text{Var}(X) = \text{Var}(Y) = 1$ ) gilt also:

$$\text{Cor}(X, Y) = \text{Cov}(X, Y)$$

Die Korrelationsmatrix eines Zufallsvektors ist analog zur Kovarianzmatrix definiert.

### 5.1.6 Kovarianzen von Linearkombinationen aus Zufallsvariablen

Die Beispiele in den Abschnitten 3 und 4 haben gezeigt, dass in einer Modellgleichung eine abhängige Variable als Linearkombination von unabhängigen Variablen dargestellt wird. Bei der Analyse von Strukturgleichungsmodellen (mit oder ohne latente Variablen) spielt die Berechnung der Varianzen und Kovarianzen von Linearkombinationen aus den Varianzen und Kovarianzen der Ausgangsvariablen eine zentrale Rolle. Für die Kovarianz von  $X := a_1X_1 + a_2X_2$  und  $Y := b_1Y_1 + b_2Y_2$  gilt:

$$\text{Cov}(X, Y) = a_1b_1 \text{Cov}(X_1, Y_1) + a_1b_2 \text{Cov}(X_1, Y_2) + a_2b_1 \text{Cov}(X_2, Y_1) + a_2b_2 \text{Cov}(X_2, Y_2) \quad (5.1)$$

Damit gilt speziell für die Varianz von  $X$ :

$$\text{Var}(X) = a_1^2 \text{Var}(X_1) + 2a_1a_2 \text{Cov}(X_1, X_2) + a_2^2 \text{Var}(X_2)$$

Für unkorrelierte Ausgangsvariablen  $X_1$  und  $X_2$  folgt:

$$\text{Var}(X) = a_1^2 \text{Var}(X_1) + a_2^2 \text{Var}(X_2)$$

Eine ausführliche Behandlung der Regeln für das Rechnen mit Kovarianzen findet sich z.B. in Kenny (1979, Kap. 2).

### 5.1.7 Stichprobenschätzer

Wir haben im aktuellen Abschnitt 5.1 bislang über Populationsparameter gesprochen. In einer konkreten Forschungssituation sind natürlich nur Stichprobenschätzer für die Parameter der manifesten Variablen verfügbar. Als Schätzer für den Erwartungswert der Variablen  $X$  verwendet man das Stichprobenmittel aus den Beobachtungswerten  $x_1, x_2, \dots, x_N$ :

$$\bar{x} := \frac{1}{N} \sum_{i=1}^N x_i$$

Die Kovarianz der Variablen  $X$  und  $Y$  kann erwartungstreu geschätzt werden durch:

$$s_{XY} := \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})$$

Die bei einer Strukturgleichungsanalyse in der Regel bevorzugten Kovarianzschätzer nach dem Maximum Likelihood - Prinzip sind allerdings *nicht* erwartungstreu:

$$s_{XY} := \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})$$

Alle oben vorgestellten Rechenregeln für Kovarianzen gelten erfreulicherweise uneingeschränkt auch für die Stichprobenschätzer.

Abschließend sollen noch zwei Bezeichnungen eingeführt werden. Für die Kovarianzmatrix des Zufallsvektors aus allen manifesten Variablen schreiben wir

$\Sigma$

Das Ziel einer Kovarianzstrukturanalyse besteht darin, eine möglichst gute Erklärung für diese Matrix  $\Sigma$  zu liefern. Bei der Suche nach einem guten Modell spielt die allein verfügbare Matrix mit den Stichprobenschätzern für die Varianzen und Kovarianzen der manifesten Variablen eine entscheidende Rolle. Von dieser zufallsabhängigen Matrix, die mit

$S$

bezeichnet wird, hängen alle Parameterschätzungen und alle Hypothesentests ab. Daher sollte diese Matrix vor Beginn der Analyse sehr sorgfältig auf mögliche Artefakte, z.B. durch Ausreißer, geprüft werden (siehe Abschnitt 2.7.3; vgl. Bollen 1989, S. 24ff).

## 5.2 Erklärung einer Kovarianzmatrix durch ein Strukturgleichungsmodell

Mit Hilfe der Rechenregeln für (Ko)varianzen aus Abschnitt 5.1 und der kausalitätstheoretisch begründeten Annahmen aus Abschnitt 3.8 kann die Formulierung aus Abschnitt 4, dass bei einer konfirmatorischen Faktorenanalyse die (Ko)varianzen zwischen den beobachteten Variablen durch deren Abhängigkeit von den latenten Variablen und durch die (Ko)varianzen der latenten Variablen erklärt würden, vollständig nachvollzogen werden.

Für die Varianzen und Kovarianzen der manifesten Variablen im Latent-State - Modell für vier Angst-Indikatoren lässt sich leicht herleiten:

$$\begin{aligned}\text{Var}(X_1) &= \text{Var}(\xi_1) + 2 \text{Cov}(\xi_1, \delta_1) + \text{Var}(\delta_1) \\ &= \text{Var}(\xi_1) + \text{Var}(\delta_1)\end{aligned}$$

$$\text{Var}(X_2) = \text{Var}(\xi_1) + \text{Var}(\delta_2)$$

$$\text{Var}(X_3) = \text{Var}(\xi_2) + \text{Var}(\delta_3)$$

$$\text{Var}(X_4) = \text{Var}(\xi_2) + \text{Var}(\delta_4)$$

$$\begin{aligned}\text{Cov}(X_1, X_2) &= \text{Cov}(\xi_1 + \delta_1, \xi_1 + \delta_2) \\ &= \text{Var}(\xi_1) + \text{Cov}(\xi_1, \delta_2) + \text{Cov}(\delta_1, \xi_1) + \text{Cov}(\delta_1, \delta_2) \\ &= \text{Var}(\xi_1)\end{aligned}$$

$$\begin{aligned}\text{Cov}(X_1, X_3) &= \text{Cov}(\xi_1 + \delta_1, \xi_2 + \delta_3) \\ &= \text{Cov}(\xi_1, \xi_2)\end{aligned}$$

$$\text{Cov}(X_1, X_4) = \text{Cov}(X_2, X_3) = \text{Cov}(X_2, X_4) = \text{Cov}(\xi_1, \xi_2)$$

$$\text{Cov}(X_3, X_4) = \text{Var}(\xi_2)$$

Nach einer Gleichheitsrestriktion des betrachteten faktorenanalytischen Modells (vgl. Abschnitt 4) sind die Residualvarianzen identisch, und wir schreiben ab jetzt stellvertretend jeweils  $\text{Var}(\delta_1)$ . Insgesamt impliziert das Modell die folgende Kovarianzmatrix der manifesten Variablen:

$$\Sigma = \begin{bmatrix} \text{Var}(\xi_1) + \text{Var}(\delta_1) & & & \\ \text{Var}(\xi_1) & \text{Var}(\xi_1) + \text{Var}(\delta_1) & & \\ \text{Cov}(\xi_2, \xi_1) & \text{Cov}(\xi_2, \xi_1) & \text{Var}(\xi_2) + \text{Var}(\delta_1) & \\ \text{Cov}(\xi_2, \xi_1) & \text{Cov}(\xi_2, \xi_1) & \text{Var}(\xi_2) & \text{Var}(\xi_2) + \text{Var}(\delta_1) \end{bmatrix}$$

Man kann die Kovarianzmatrix der manifesten Variablen als **empirische Gesetzmäßigkeit** auffassen, die durch die Entwicklung eines faktorenanalytischen Modells erklärt werden soll. Dieses Vorgehen entspricht dem Grundschemata des wissenschaftlichen Erklärens, wobei nach Überzeugung der Wissenschaftsphilosophie eine Erklärung mit Hilfe theoretischer Begriffe (latenter Variablen) besonders erfolgversprechend ist.

Die Darstellung der Kovarianzmatrix  $\Sigma$  der manifesten Variablen als Funktion der Modellparameter ist ein wesentliches Merkmal der Kovarianzstrukturanalyse. Neben den Pfad- bzw. Strukturkoeffizienten werden für diese Darstellung auch alle Varianzen und Kovarianzen der Quellvariablen (exogenen Variablen und Residualvariablen) benötigt. Der Vektor aller Modellparameter in diesem Sinn wird meist mit  $\theta$  bezeichnet (sprich: Theta). Im faktorenanalytischen Beispiel haben wir nach Auswertung der Modellgleichungen, Restriktionen und Identifikationsüberlegungen folgenden Vektor erhalten:

$$\theta = (\text{Var}(\xi_1), \text{Var}(\xi_2), \text{Cov}(\xi_1, \xi_2), \text{Var}(\delta_1))$$

Er enthält alle freien Parameter sowie einen Repräsentanten für die Gruppe gleich gesetzter Parameter.

Die Darstellung der Kovarianzmatrix  $\Sigma$  als Funktion dieses Parametervektors ist von der allgemeinen Gestalt:

$$\Sigma = f(\boldsymbol{\theta}) \quad (5.2)$$

Eine solche Darstellung werden wir später gelegentlich als die *Fundamentalthypothese* des jeweiligen Modells bezeichnen.

Weil die Elemente in der Kovarianzmatrix der latenten Variablen anschließend noch mehrmals in Formel auftauchen, verwenden wir die folgende, zusammen mit dem Computerprogramm LISREL (siehe z.B. Jöreskog & Sörbom 1989) eingeführte Symbolik:

$$\begin{bmatrix} \text{Var}(\xi_1) & \\ \text{Cov}(\xi_2, \xi_1) & \text{Var}(\xi_1) \end{bmatrix} = \begin{bmatrix} \phi_{11} & \\ \phi_{21} & \phi_{22} \end{bmatrix}$$

Wird außerdem für  $\text{Var}(\delta_1)$  das Symbol  $\sigma_\delta^2$  verwendet, lässt sich der Parametervektor sehr kompakt schreiben

$$\boldsymbol{\theta} = (\phi_{11}, \phi_{22}, \phi_{21}, \sigma_\delta^2)$$

### 5.3 Identifikation der Parameter

Wir haben im letzten Abschnitt gesehen, wie die Kovarianzmatrix  $\Sigma$  der manifesten Variablen als Funktion des Parametervektors  $\boldsymbol{\theta}$  ausgedrückt werden kann. Nun sind allerdings die Parameter unbekannt und wir möchten Informationen über sie aus den Daten gewinnen. Andererseits können wir die Varianzen und Kovarianzen der manifesten Variablen immerhin anhand der Stichprobe schätzen. Wenn es gelingt, die Gleichung

$$\Sigma = f(\boldsymbol{\theta})$$

nach den unbekanntem Parametern aufzulösen, kann man  $\Sigma$  durch die Matrix  $S$  der Stichprobenstatistiken ersetzen und Schätzer für die gesuchten Parameter berechnen.

Ein Parameter  $\theta_v$  heißt **identifiziert**, wenn für ihn das Auflösen gelingt, d.h. wenn er auf mindestens eine Weise als Funktion  $g_v$  der Elemente von  $\Sigma$  dargestellt werden kann:

$$\theta_v = g_v(\Sigma)$$

Ein Modell heißt **identifiziert**, wenn alle Parameter des Modells identifiziert sind.

Es gibt leider keine notwendige **und** hinreichende Bedingung für die Identifikation eines beliebigen Strukturgleichungsmodells. Allerdings sind für einige spezielle Modelle nützliche Identifikationsregeln bekannt:

- Rekursive Modelle für manifeste Variablen sind stets identifiziert (Kline 2005, S. 107).
- Bei einer *Standard-CFA* (vgl. Abschnitt 4) gelten für die Identifikation die folgenden hinreichenden Bedingungen (Kline 2005, S. 172):
  - Ein Modell der Standard CFA ist identifiziert, wenn jedem Faktor mindestens drei Indikatoren zugeordnet sind.
  - Bei einem Modell mit zwei oder mehr Faktoren genügen zwei Indikatoren pro Faktor.

Nach diesen Regeln ist das aktuell betrachtete Latent-State - Modell für vier Angst-Indikatoren identifiziert. Bei diesem Modell kann es allerdings zu einem so genannten **empirischen Identifikationsproblem** kommen, wenn die Korrelation zwischen den Faktoren gleich Null wird. Dann liegen nämlich zwei separate Teilmodelle vor, die jeweils einen Faktor und zwei Indikatoren enthalten, also nach obigen Regeln *nicht* identifiziert sind.

Prinzipiell muss zum Nachweis der Identifikation eines konkreten Modells für jeden Parameter seine Darstellbarkeit durch die Varianzen und Kovarianzen der manifesten Variablen gezeigt werden. Wichtige

Hilfsmittel bei diesem Nachweis sind die in Abschnitt 5.1 eingeführten Rechenregeln für Kovarianzen. Die Vorgehensweise soll beim aktuellen Beispiel vorgeführt werden, obwohl seine Identifikation bereits sichergestellt ist. Wir setzen die Kovarianzmatrix  $\text{Cov}(\mathbf{X})$  gleich der Modellimplikation  $f(\boldsymbol{\theta})$ :

$$\begin{bmatrix} \text{Var}(X_1) & & & & \\ \text{Cov}(X_2, X_1) & \text{Var}(X_2) & & & \\ \text{Cov}(X_3, X_1) & \text{Cov}(X_3, X_2) & \text{Var}(X_3) & & \\ \text{Cov}(X_4, X_1) & \text{Cov}(X_4, X_2) & \text{Cov}(X_4, X_3) & \text{Var}(X_4) & \end{bmatrix} = \begin{bmatrix} \phi_{11} + \sigma_{\delta}^2 & & & & \\ \phi_{11} & \phi_{11} + \sigma_{\delta}^2 & & & \\ \phi_{21} & \phi_{21} & \phi_{22} + \sigma_{\delta}^2 & & \\ \phi_{21} & \phi_{21} & \phi_{22} & \phi_{22} + \sigma_{\delta}^2 & \end{bmatrix}$$

Es sind zehn Gleichungen nach den vier Unbekannten in

$$\boldsymbol{\theta} = (\phi_{11}, \phi_{22}, \phi_{21}, \sigma_{\delta}^2)$$

aufzulösen. Wir erhalten mühelos:

$$\begin{aligned} \phi_{11} &= \text{Cov}(X_2, X_1) \\ \phi_{21} &= \text{Cov}(X_3, X_1) \\ \phi_{22} &= \text{Cov}(X_4, X_3) \\ \sigma_{\delta}^2 &= \text{Var}(X_1) - \text{Cov}(X_2, X_1) \end{aligned}$$

Damit steht also einer Schätzung der Parameter prinzipiell nichts mehr im Weg.

Bei genauerem Hinsehen fällt auf, dass unser System sogar noch sechs weitere Bestimmungsgleichungen für Parameter liefert:

$$\begin{aligned} \phi_{21} &= \text{Cov}(X_3, X_2) \\ \phi_{21} &= \text{Cov}(X_4, X_1) \\ \phi_{21} &= \text{Cov}(X_4, X_2) \\ \sigma_{\delta}^2 &= \text{Var}(X_2) - \text{Cov}(X_2, X_1) \\ \sigma_{\delta}^2 &= \text{Var}(X_3) - \text{Cov}(X_4, X_3) \\ \sigma_{\delta}^2 &= \text{Var}(X_4) - \text{Cov}(X_4, X_3) \end{aligned}$$

Für  $\phi_{21}$  und  $\sigma_{\delta}^2$  haben wir jeweils *mehrere* Bestimmungsgleichungen zur Verfügung. Parameter mit dieser Eigenschaft werden als **überidentifiziert** oder **überbestimmt** bezeichnet. Ein Modell wird so genannt, wenn es identifiziert ist und mindestens einen überidentifizierten Parameter enthält.

Bei einem überidentifizierten Modell ist die Anzahl  $t$  der Parameter kleiner als die Anzahl nicht-redundanter Elemente in der zu erklärenden Kovarianzmatrix und damit genügen diese Modelle dem wissenschaftstheoretischen Gebot der Sparsamkeit. Die Überbestimmtheit hat zwei Konsequenzen, die in späteren Abschnitten noch genauer beleuchtet werden:

- Verschiedene Bestimmungsgleichungen für einen Parameter liefern zufallsbedingt im Allgemeinen auch dann unterschiedliche Ergebnisse, wenn das Strukturgleichungsmodell stimmt. Zu eindeutigen Schätzern gelangt man durch folgendes Minimierungsprinzip: Man wählt den Vektor  $\hat{\boldsymbol{\theta}}$  der geschätzten Parameter so, dass der „Abstand“ zwischen der Stichprobenkovarianzmatrix  $\mathbf{S}$  und der durch das Modell postulierten Kovarianzmatrix  $f(\hat{\boldsymbol{\theta}})$  möglichst klein wird. Der Abstandsbegriff wird in Abschnitt 5.4 über Parameterschätzung präzisiert.
- Das Strukturgleichungsmodell impliziert für die Kovarianzmatrix der manifesten Variablen Restriktionen, die empirisch falsch sein können. Unter der Annahme multivariater Normalverteilung der manifesten Variablen erlaubt dies eine inferenzstatistische Beurteilung des Modells. Damit genügen überbestimmte Modelle der wissenschaftsphilosophischen Forderung nach Falsifizierbarkeit.

Ein Strukturgleichungsmodell heißt **genau identifiziert** oder **saturiert**, wenn alle Parameter identifiziert sind und kein Parameter überidentifiziert ist, d.h. wenn **genau ein** Lösungsvektor  $\hat{\theta}$  für die Gleichung (5.2) existiert. Für den Schätzer gilt dann:

$$S = f(\hat{\theta}),$$

d.h. die Stichprobenkovarianzmatrix  $S$  kann exakt reproduziert werden. Dies ist häufig der Fall, wenn die Anzahl  $t$  der Parameter in  $\theta$  identisch ist mit der Anzahl nicht-redundanter Elemente in  $S$ . Letztere Anzahl beträgt bei  $k$  manifesten Variablen  $\frac{1}{2}k(k+1)$ . Genau identifizierte Modelle sind nicht prüfbar. Sofern sie stimmen, können sie aber trotzdem sinnvoll sein, weil sie immerhin die Schätzung und (bei multivariater Normalverteilung der manifesten Variablen) die Testung der einzelnen Parameter erlauben.

Eine notwendige (aber nicht hinreichende!) Bedingung für die Identifikation eines Modells lautet

$$t \leq \frac{1}{2}k(k+1) \quad (5.3)$$

Bei einem **nicht-identifizierten (unterbestimmten)** Modell existieren unendlich viele verschiedene Vektoren  $\theta^* \in \mathbb{R}^t$ , welche dieselbe Kovarianzmatrix  $f(\theta^*)$  liefern und damit natürlich auch denselben Abstand zu einer Stichprobenkovarianzmatrix  $S$ . Bei der Parameterschätzung nach der oben erwähnten Abstand-Minimierungsmethode muss daher die iterative Suche nach dem globalen Minimum scheitern.

Wie Sie bereits wissen, wird auch bei der obligatorischen Definition einer Maßeinheit für jede latente Variable eines Strukturgleichungsmodells ein Identifikationsproblem gelöst (siehe Abschnitt 3.4). Dies soll im Zusammenhang mit dem Latent-State - Modell der Angst noch einmal erläutert werden. Nimmt man hier einen freien Parameter  $\lambda$  für die (als gleich angenommenen) Ladungen auf an Stelle der Fixierung auf Eins, so sind  $\lambda$ ,  $\phi_{11}$ ,  $\phi_{21}$  und  $\phi_{22}$  nicht identifiziert, obwohl fünf Parametern zehn Varianzen bzw. Kovarianzen gegenüberstehen, denn die Parametervektoren

$$(\lambda, \phi_{11}, \phi_{21}, \phi_{22}, \sigma_{\delta}^2) \quad \text{und} \quad (\alpha\lambda, \frac{1}{\alpha^2}\phi_{11}, \frac{1}{\alpha^2}\phi_{21}, \frac{1}{\alpha^2}\phi_{22}, \sigma_{\delta}^2), \alpha > 0$$

liefern dieselbe Kovarianzmatrix. Der Parameter  $\sigma_{\delta}^2$  ist hingegen nach wie vor überbestimmt. Dieses Beispiel zeigt, dass ein Modell sowohl identifizierte als auch nicht-identifizierte Parameter enthalten kann und dass eine erfüllte Ungleichung (5.3) keine Gewähr für die Identifikation bietet.

Amos entdeckt in vielen Fällen Identifikationsprobleme und gibt entsprechende Warnungen aus. Es ist jedoch anzustreben, die Identifikation eines Modells vor dem Einsatz von Amos abzuklären.

Kline (2005, S. 213) empfiehlt in Anlehnung an Bollen (19989), bei einem beliebigen Modell mit latenten Variablen die Identifikation separat für das Strukturmodell und das Messmodell (mit freien Beziehungen zwischen den latenten Variablen) zu prüfen. Sind beide Partialmodelle separat identifiziert, so ist die Identifikation auch für das Gesamtmodell gegeben. Gerade wurde eine hinreichende Bedingung genannt, die aber nicht notwendig ist. Ein Gesamtmodell kann also durchaus identifiziert sein, obwohl nicht beide Partialmodelle identifiziert sind.

In vielen Fällen kann man eine der Identifikationsregeln aus Bollen (1989) verwenden (siehe oben). Ist dies nicht möglich, so kontrolliere man zunächst die notwendige Bedingung (5.3). Ist sie erfüllt, versuche man, mit Hilfe der Rechenregeln für Kovarianzen das Gleichungssystem (5.2) nach den Parametern aufzulösen. Bei komplexen Modellen ist diese Aufgabe allerdings kaum zu bewältigen. Etwas leichter zu analysieren sind (siehe Kline 2005, S. 240):

- die notwendige Ordnungsbedingung
- die hinreichende Rangbedingung

Oft bleibt nichts anderes übrig, als eine Lösung empirisch auf Eindeutigkeit zu prüfen (Kline 2005, S.173).

## 5.4 Schätzung der Parameter durch Minimierung einer Fit-Funktion

Das Schätzen der Parameter setzt die Identifikation des Modells voraus. Dann gilt für je zwei verschiedene Parametervektoren  $\theta'$  und  $\theta''$ , dass sie verschiedene Kovarianzmatrizen implizieren, d.h.  $f(\theta') \neq f(\theta'')$ . Daher kann ein Schätzer  $\hat{\theta}$  nach folgendem Prinzip ermittelt werden:

*Bestimme  $\hat{\theta}$  so, dass der Abstand zwischen der Stichprobenkovarianzmatrix  $S$  und der von  $\hat{\theta}$  implizierten Kovarianzmatrix  $f(\hat{\theta})$  minimal wird.*

Weil bei den folgenden Darstellungen mehrere Thetas auftreten, soll deren Bedeutung vorab klargestellt werden:

- $\theta$  Der gesuchte Vektor der wahren Parameter. Dies ist ein festes Element aus dem  $\mathbb{R}^l$ .
- $\hat{\theta}$  Stichproben-Schätzer für  $\theta$
- $\theta^*$  Beliebiges Element aus der Menge aller potentiellen Parametervektoren. (Dies ist eine geeignet eingeschränkte Teilmenge des  $\mathbb{R}^l$ .)  $\theta^*$  ist also ein Platzhalter, der beliebige Werte aus einer sehr großen Menge von Vektoren annehmen kann, während  $\theta$  ein Vektor aus festen (aber unbekanntem) Zahlen ist.

Für ein beliebiges  $\theta^*$  soll  $f(\theta^*)$  im Folgenden durch  $\Sigma^*$  abgekürzt werden.

Aus dem dargestellten Grundprinzip wurden durch Definition einer speziellen **Fit-Funktion**

$$F(S, \Sigma^*)$$

zur Messung des Abstands zwischen  $S$  und  $\Sigma^*$  mehrere konkrete Schätzverfahren entwickelt.  $F$  ist dabei eine nicht-negative Funktion der Komponenten von  $\theta^*$ , die nur bei perfektem Fit ( $\Sigma^* = S$ ) den Wert Null annimmt. Sie ist im Allgemeinen nicht linear und so kompliziert, dass ihre Minimierung ein iteratives Verfahren erfordert. Dabei müht sich der Computer schrittweise durch einen oft sehr hochdimensionalen Parameterraum und sucht nach dem tiefsten Tal in der Funktionsoberfläche. Anschließend werden exemplarisch zwei Schätzverfahren betrachtet:

- Unweighted Least Squares (ULS)  
Es eignet sich aufgrund seiner Einfachheit zur Veranschaulichung der grundsätzlichen Arbeitsweise, ist aber in der Forschungspraxis selten die beste Wahl.
- Maximum Likelihood (ML)  
Dieses Verfahren dominiert in der Praxis der Strukturgleichungsanalyse und ist in Amos voreingestellt.

### 5.4.1 Unweighted Least Squares (ULS)

Bei der Parameterschätzung nach dem ULS-Verfahren wird die folgende Fit-Funktion minimiert:

$$F_{\text{ULS}}(S, \Sigma^*) := \frac{1}{2} \text{tr}[(S - \Sigma^*)^T (S - \Sigma^*)]$$

Mit  $\text{tr}$  wird dabei eine Funktion bezeichnet, die für eine quadratische Matrix die Summe ihrer Hauptdiagonalelemente liefert.  $F_{\text{ULS}}(S, \Sigma^*)$  ist nichts anderes als die Summe aller quadrierten Differenzen zwischen korrespondierenden Elementen aus  $S$  und  $\Sigma^*$ .  $F_{\text{ULS}}(S, \Sigma^*)$  wird offensichtlich umso kleiner, je ähnlicher sich  $S$  und  $\Sigma^*$  werden.

Beispiel: Bei

$$\mathbf{S} = \begin{bmatrix} 1.1 & 0.6 \\ 0.6 & 2.2 \end{bmatrix}, \quad \mathbf{\Sigma}^* = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 2 \end{bmatrix}$$

$$\Rightarrow \mathbf{S} - \mathbf{\Sigma}^* = \begin{bmatrix} 0.1 & 0.1 \\ 0.1 & 0.2 \end{bmatrix}$$

ergibt sich

$$F_{\text{ULS}}(\mathbf{S}, \mathbf{\Sigma}^*) = \frac{1}{2} (0.1^2 + 0.1^2 + 0.1^2 + 0.2^2) = 0.035.$$

Für die bei Minimierung von  $F_{\text{ULS}}(\mathbf{S}, \mathbf{\Sigma}^*)$  resultierenden Schätzer spricht, dass sie **ohne Voraussetzungen über die Verteilung** der manifesten Variablen **konsistent** sind, d.h. ( $N =$  Stichprobenumfang):

$$\lim_{N \rightarrow \infty} P\left(|\hat{\theta}_N - \theta| > \varepsilon\right) = 0 \quad \forall \varepsilon > 0$$

Die Parameter im Vektor  $\theta$  werden also asymptotisch richtig geschätzt. Die Konsistenz-Eigenschaft besitzen auch alle anderen von Amos berechneten Schätzer.

Nachteile der ULS-Schätzmethode:

i) Unter der klassischen Annahme multivariater Normalität der gemeinsamen Verteilung aller manifesten Variablen ist die ULS-Schätzmethode der GLS- und der ML-Schätzmethode unterlegen.

ii) Die Fit-Funktion  $F_{\text{ULS}}$  ist nicht skaleninvariant.

Zur Erläuterung dieser Aussage sind einige Definitionen erforderlich (vgl. Jöreskog & Sörbom 1989, Abschnitt 1.20):

- Ein Strukturgleichungsmodell mit der Fundamental-Hypothese  $\mathbf{\Sigma} = f(\theta)$  heißt skaleninvariant, wenn für jede Diagonalmatrix  $\mathbf{D}$  mit positiven Einträgen und für jeden Parametervektor  $\theta$  ein anderer Parametervektor  $\theta_D$  existiert mit:

$$f(\theta_D) = \mathbf{D} f(\theta) \mathbf{D}$$

Die Matrix  $\mathbf{D}$  enthält Skalenfaktoren für die manifesten Variablen. Sie bewirkt also Wechsel der Maßeinheiten (z.B. Gewichtsmessung in Gramm statt Kilogramm). Trägt man in die Hauptdiagonale von  $\mathbf{D}$  die Kehrwerte der Standardabweichungen der manifesten Variablen ein, so ist  $\mathbf{DSD}$  eine Korrelationsmatrix.

- Eine Fit-Funktion  $F(\mathbf{S}, \mathbf{\Sigma}^*)$  heißt skaleninvariant, wenn für jede Diagonalmatrix  $\mathbf{D}$  mit positiven Einträgen gilt:

$$F(\mathbf{DSD}, \mathbf{D}\mathbf{\Sigma}^*\mathbf{D}) = F(\mathbf{S}, \mathbf{\Sigma}^*)$$

- Eine Schätzmethode heißt skaleninvariant, wenn sie auf einer skaleninvarianten Fit-Funktion basiert. Eine skaleninvariante Schätzmethode verhält sich „vernünftig“, wenn bei der Analyse eines skaleninvarianten Modells Maßeinheiten für manifeste Variablen verändert werden, d.h. wenn statt  $\mathbf{S}$  die Matrix  $\mathbf{DSD}$  verwendet wird. In diesem Fall implizieren die neuen Parameterschätzungen in  $\hat{\theta}_D$  die Kovarianzmatrix  $\mathbf{D} f(\hat{\theta}) \mathbf{D}$ . Die neuen Schätzungen in  $\hat{\theta}_D$  ergeben sich also durch Berücksichtigung der Maßstabsänderungen unmittelbar aus den alten Schätzungen in  $\hat{\theta}$  (Jöreskog & Sörbom 1989).

Die Fit-Funktion  $F_{\text{ULS}}$  ist *nicht* skaleninvariant, und daher können durch Skalenfaktoren bei den ULS-Schätzungen Veränderungen auftreten, die nicht verträglich sind mit den Maßstabsänderungen. Weil der Maßstab sozialwissenschaftlicher Indikatoren meist willkürlich ist, wird daher oft empfohlen, ULS-Schätzer aus der Korrelationsmatrix statt aus der Kovarianzmatrix zu bestimmen (vgl. Long, 1983).

### 5.4.2 Maximum Likelihood (ML)

Die ML-Methode setzt voraus, dass die  $k$  manifesten Variablen multivariat normalverteilt sind. Unter dieser Voraussetzung hat Wishart (1928) die Wahrscheinlichkeitsdichte  $W(\mathbf{S}; \mathbf{\Sigma}, N-1)$  für eine Stichprobenkovarianzmatrix  $\mathbf{S}$  bei unabhängiger Ziehung von  $N$  Fällen aus einer Population mit Kovarianzmatrix  $\mathbf{\Sigma}$  hergeleitet. Das Maß für die Ähnlichkeit von  $\mathbf{S}$  und  $\mathbf{\Sigma}^*$  bzw. für den Fit eines potentiellen Parametervektors  $\boldsymbol{\theta}^*$  basiert auf dem Quotienten aus der Wahrscheinlichkeitsdichte von  $\mathbf{S}$  bei Annahme von  $\mathbf{\Sigma} = \mathbf{\Sigma}^*$  und der Wahrscheinlichkeitsdichte von  $\mathbf{S}$  bei Annahme von  $\mathbf{\Sigma} = \mathbf{S}$ :

$$\log\left(\frac{W(\mathbf{S}, \mathbf{\Sigma}^*, N-1)}{W(\mathbf{S}, \mathbf{S}, N-1)}\right) = -\frac{1}{2} (N-1) [\text{tr}(\mathbf{S}\mathbf{\Sigma}^{*-1}) + \log|\mathbf{\Sigma}^*| - \log|\mathbf{S}| - k]$$

Dabei wird vorausgesetzt, dass die Determinanten  $|\mathbf{S}|$  und  $|\mathbf{\Sigma}^*|$  positiv sind. Der Likelihood-Quotient zeigt an, inwiefern die Likelihood von  $\mathbf{S}$  unter Annahme von  $\mathbf{\Sigma}^*$  schlechter ist als die maximal mögliche Likelihood und kann folglich zwischen 0 und 1 variieren. Der logarithmierte Quotient nimmt daher Werte von  $-\infty$  bis 0 an und erreicht seinen maximalen Wert 0 genau dann, wenn  $\mathbf{\Sigma}^* = \mathbf{S}$ . Statt diesen Ausdruck zu maximieren kann man wegen des negativen Vorzeichens auch den Ausdruck in eckigen Klammern minimieren. So ergibt sich die folgende Fit-Funktion:

$$F_{\text{ML}}(\mathbf{S}, \mathbf{\Sigma}^*) = \text{tr}(\mathbf{S}\mathbf{\Sigma}^{*-1}) + \log|\mathbf{\Sigma}^*| - \log|\mathbf{S}| - k$$

$F_{\text{ML}}(\mathbf{S}, \mathbf{\Sigma}^*)$  wird genau dann minimal, wenn der Zähler des Likelihood-Quotienten die maximale Likelihood der beobachteten Kovarianzmatrix  $\mathbf{S}$  bei Gültigkeit des behaupteten Strukturgleichungsmodells, d.h. unter den Restriktionen, die das Modell für die Kovarianzmatrix der manifesten Variablen impliziert, annimmt. Dadurch ist der Name *Maximum-Likelihood - Schätzer* begründet.

Im Nenner des Likelihood-Quotienten steht die maximale Likelihood von  $\mathbf{S}$  in einem Normalverteilungsmodell mit beliebiger, d.h. unrestringierter, Kovarianzmatrix. Er bleibt während der Minimierungsprozedur konstant und ist daher für die Schätzung irrelevant. Zur Testung des Modells ist jedoch der komplette Likelihood-Quotient wesentlich (siehe unten).

Der ML-Schätzer hat **bei multivariater Normalverteilung der manifesten Variablen und Gültigkeit des Modells** über die auch beim ULS-Schätzer vorhandene Konsistenz (siehe Abschnitt 5.4.1) hinaus zahlreiche günstige Eigenschaften:

i) Er ist **asymptotisch erwartungstreu**, d.h.:

$$\lim_{N \rightarrow \infty} E(\hat{\boldsymbol{\theta}}_N) = \boldsymbol{\theta}$$

ii) Er ist **asymptotisch effizient**, d.h. mit wachsendem  $N$  wird seine Varianz, also sein Fehler, kleiner oder gleich der Varianz jedes beliebigen anderen Schätzers.

iii) Seine Verteilung konvergiert mit wachsendem  $N$  gegen eine **Normalverteilung**, d.h. die Stichprobenschätzer  $\hat{\theta}_i$  sind asymptotisch normalverteilt um den Erwartungswert  $\theta_i$ . Man kann ferner eine Schätzung für den asymptotischen **Standardfehler**  $SE(\hat{\theta}_i)$  des Schätzers  $\hat{\theta}_i$  berechnen. Das erlaubt die Konstruktion von Vertrauensintervallen und über die folgende, von Amos als **CR (Critical Ratio)** bezeichnete Prüfgröße

$$CR(\hat{\theta}_i) := \frac{\hat{\theta}_i}{SE(\hat{\theta}_i)}$$

einen (zweiseitigen) Test der Nullhypothese:

$$H_0^i : \theta_i = 0$$

Der Test zum Niveau  $\alpha = 0,05$  verwirft seine Nullhypothese, wenn  $|\text{CR}(\hat{\theta}_i)|$  größer ist als das 97.5% - Fraktile der Standardnormalverteilung ( $\approx 2$ ). Man kann auch allgemeiner die Hypothese:

$$H_0^i : \theta_i = c$$

prüfen, wobei  $c$  eine geeignete Konstante ist. Diese Nullhypothese ist in einem Test zum Niveau  $\alpha = 0,05$  zu verwerfen, wenn das 95% - Vertrauensintervall zu  $\hat{\theta}_i$  den Punkt  $c$  nicht enthält.

iv) Die zur ML-Schätzung verwendete Fit-Funktion  $F_{ML}$  erlaubt einen **globalen Modelltest**, d.h. eine Prüfung der Nullhypothese, dass die Populationskovarianzmatrix von der Form ist, die das Modell impliziert, gegen die Alternativhypothese, dass sie den vom Modell auferlegten Restriktionen nicht genügt.

Wie eben zu sehen war, ist die Fit-Funktion  $F_{ML}(S, \Sigma^*)$  bis auf den Vorfaktor  $\frac{1}{2}(N-1)$  mit dem logarithmierten Likelihood-Quotienten:

$$-2 \log \left( \frac{W(S; \Sigma^*, N-1)}{W(S; S, N-1)} \right)$$

identisch. Nach einer allgemeinen Theorie der mathematischen Statistik ist das Minimum dieses logarithmierten Likelihood-Quotienten bei Gültigkeit des Modells asymptotisch (d.h. für  $N \rightarrow \infty$ )  $\chi^2$ -verteilt mit

$$df = \frac{1}{2}k(k+1) - t$$

Freiheitsgraden, wobei wiederum  $k$  die Anzahl der manifesten Variablen und  $t$  die Anzahl der freien Parameter im Modell bedeutet. Damit erlaubt also die folgende Prüfgröße ( $\hat{\Sigma}$  sei die vom geschätzten Parametervektor implizierte Kovarianzmatrix):

$$\chi^2 := (N-1) \cdot F_{ML}(S, \hat{\Sigma})$$

eine Prüfung der Nullhypothese, dass die Populationskovarianzmatrix von der Form ist, die das Modell impliziert, gegen die Alternativhypothese, dass sie den vom Modell auferlegten Restriktionen nicht genügt.

Amos berechnet die  $\chi^2$ -Statistik und gibt die approximative Wahrscheinlichkeit für das Auftreten eines gleich großen oder größeren  $\chi^2$ -Wertes bei Gültigkeit des Modells an.

Es soll an dieser Stelle daran erinnert werden, dass ein  $\chi^2$ -Modelltest basierend auf der ML-Fitfunktion nur unter folgenden Voraussetzungen gültig ist (vgl. Abschnitt 2.7):

- Die manifesten Variablen sind multivariat normalverteilt.  
Bei manifesten exogenen Variablen kann auf diese Voraussetzung verzichtet werden.
- Die Stichprobe ist hinreichend groß.

Auch die Schätzmethoden GLS und ADF ermöglichen einen globalen Modellgültigkeitstest. Bei Verwendung der ADF-Schätzmethode kann auf die Normalitätsvoraussetzung verzichtet werden. Allerdings ist ein erheblich größerer Stichprobenumfang erforderlich.

v) Die ML-Fitfunktion ist **skaleninvariant**.

Sofern das untersuchte Modell seinerseits skaleninvariant ist, bleiben die ML-Schätzergebnisse im wesentlichen invariant bei Veränderung der Maßeinheiten von manifesten Variablen (vgl. Abschnitt 5.4.1). Während der globale Modelltest, die CR-Werte (siehe oben) und die Modifikationsindikatoren (vgl. Abschnitt 7.1) unverändert bleiben, werden die Parameterschätzer an die veränderten Maßeinheiten der manifesten Variablen angepasst (siehe Jöreskog & Sörbom 1989).

Der entscheidende Nachteil des ML-Schätzers ist die starke Voraussetzung der multivariaten Normalverteilung der manifesten Variablen. Allerdings wurde inzwischen gezeigt, dass diese Voraussetzung liberalisiert werden kann (vgl. Bollen 1989, S. 114), und in Simulationsstudien hat sich der ML-Schätzer als robust gegen mäßige Verletzungen der Normalität erwiesen. Die ML-Schätzmethode sollte jedoch in folgenden Fällen *nicht* verwendet werden:

- Die Verteilung der manifesten Variablen weicht erheblich von der Normalverteilung ab.
- Einige der manifesten Variablen sind ordinalskaliert, d.h. sie können nur wenige verschiedene, monoton geordnete Werte annehmen.

Eine besondere Attraktion der in Amos realisierten Schätzmethode **Full Information Maximum Likelihood** (FIML) besteht darin, dass bei Datensätzen mit fehlenden Werten alle vorhandenen Informationen genutzt werden, ohne fehlende Werte durch Imputationen ersetzen zu müssen. Man erhält konsistente Schätzung und korrekte Standardfehler unter der relativ liberalen MAR-Bedingung (*Missing At Random*) (siehe z.B. Baltés-Götz 2008a).

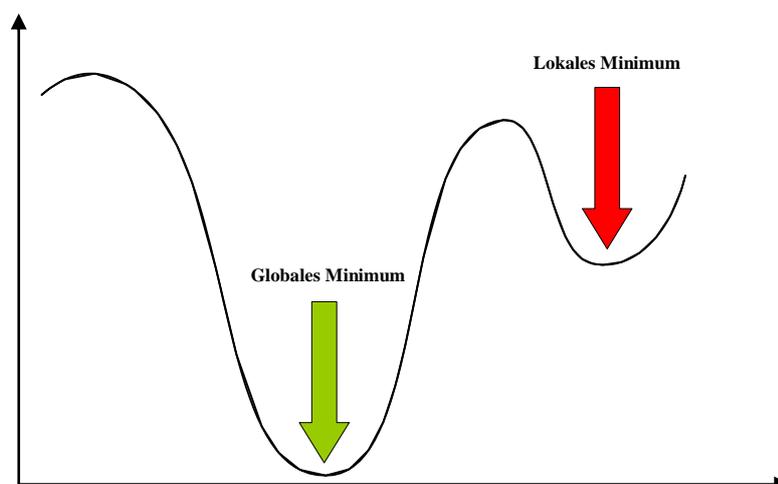
### 5.4.3 Hinweise zur praktischen Anwendung der Schätzverfahren

Für alle iterativen Schätzverfahren gelten folgende Anmerkungen zu möglichen Anwendungsproblemen:

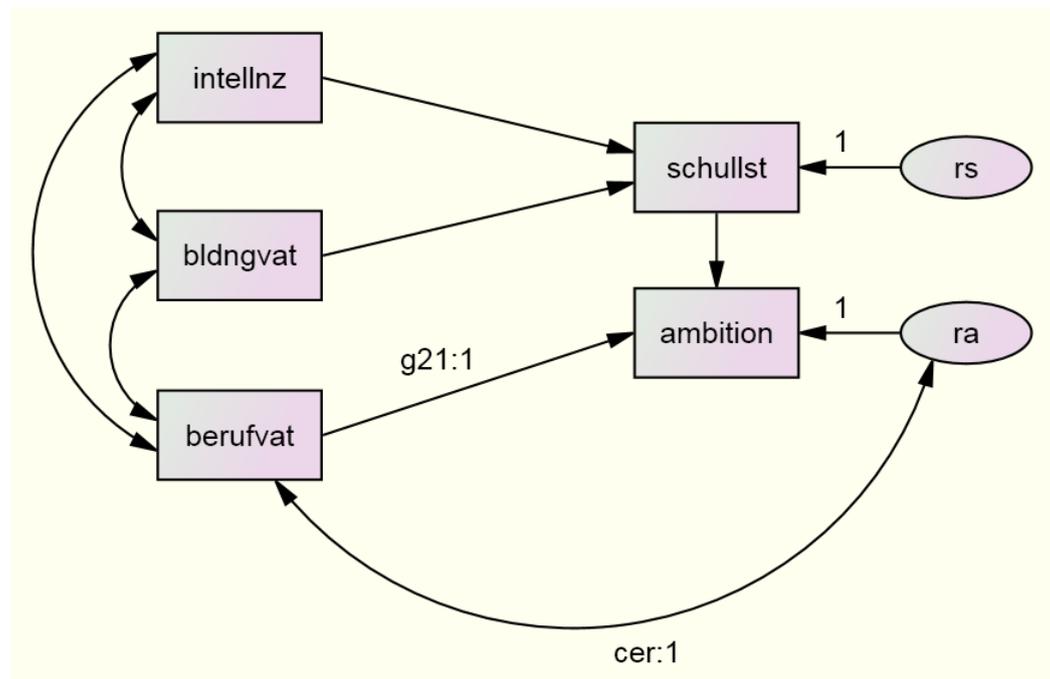
i) Wenn das Iterationsverfahren nicht konvergiert oder ein Minimum außerhalb des zulässigen Parameterraumes auftritt (z.B. negativer Varianzschätzer, so genannte **Heywood-Cases**) kommen vor allem folgende Ursachen in Frage:

- Das Modell ist nicht identifiziert.
- Bei einer konfirmatorischen Faktorenanalyse neigen Modelle mit lediglich zwei Indikatoren pro Faktor trotz Identifikation (vgl. Abschnitt 5.3) zu Schätzproblemen (Kline 2005, S. 178).
- Das Modell passt sehr schlecht zu den Daten.  
Ein schlechter Fit kann auch durch zu geringe Stichprobengröße entstehen (siehe Abschnitt 2.7).
- Die Daten sind schlecht konditioniert, z.B. bei einer Kovarianzmatrix mit stark mit stark unterschiedlichen Varianzen (vgl. Kline 2005, S.113).

ii) Es kann passieren, dass der Algorithmus in einem **lokalen Minimum** landet, das kein globales Minimum ist. Dies wird in folgender Abbildung für den einfachen Spezialfall eines einparametrischen Modells demonstriert.



Gegen diese Gefahr, die sich in Simulationsstudien allerdings als relativ gering erwiesen hat, kann man sich durch mehrfache Schätzversuche ausgehend von verschiedenen Startwerten absichern. Um mit der graphischen Benutzerschnittstelle von Amos für einen Parameter einen expliziten Startwert als Alternative zur automatischen Wahl durch Amos zu vereinbaren, benennt man den Parameter und setzt hinter den Namen durch Doppelpunkt abgetrennt den gewünschten Startwert, z.B.



### 5.5 Flexible Hypothesentests durch den Vergleich von geschachtelten Modellen

$M_u$  sei ein **gültiges** Strukturgleichungsmodell mit dem Parametervektor  $\theta_u$  und  $df_u$  Freiheitsgraden. Im Ambitionsmodell *mit* Fehlerkovarianz (vgl. Abschnitt 3.7) enthält  $\theta_u$  die Elemente

$$\theta_u = (\gamma_{11}, \gamma_{12}, \gamma_{21}, \gamma_{22}, \text{Var}(\mathbf{RS}), \text{Var}(\mathbf{RA}), \text{Cov}(\mathbf{RS}, \mathbf{RA}))$$

Eine prüfbare Nullhypothese  $H_0$  kann aus konjunktiv verknüpften Teilhypothesen  $H_0^i$  bestehen, die jeweils eine von der beiden folgenden Behauptungen aufstellen:

- Gewisse Komponenten von  $\theta_u$  sind gleich bestimmten Zahlen (Fixierung).
- Gewisse Komponenten von  $\theta_u$  sind untereinander gleich (Gleichheitsrestriktion).

$H_0$  führt zu einem Modell  $M_e$  mit  $df_e > df_u$  Freiheitsgraden. Wesentlich ist, dass  $M_e$  (eingeschränktes Modell) durch zusätzliche Forderungen aus  $M_u$  (**uneingeschränktes** Modell) entsteht. Man sagt auch,  $M_e$  sei in  $M_u$  geschachtelt. Ist  $\chi_u^2$  die  $\chi^2$ -Fitstatistik für  $M_u$  und  $\chi_e^2$  die  $\chi^2$ -Fitstatistik für  $M_e$ , dann ist:

$$D^2 := \chi_e^2 - \chi_u^2$$

unter der Nullhypothese  $H_0$  approximativ  $\chi^2$ -verteilt mit

$$d := df_e - df_u$$

Freiheitsgraden. Folglich ist die  $H_0$  bei einem Test zum Niveau  $\alpha$  genau dann abzulehnen, wenn  $D^2$  größer als das  $(1-\alpha)$ -Fraktile der  $\chi^2$ -Verteilung mit  $d$  Freiheitsgraden ist.

Die Voraussetzungen für den  $D^2$ -Test bei Verwendung der ML- oder GLS-Schätzmethode:

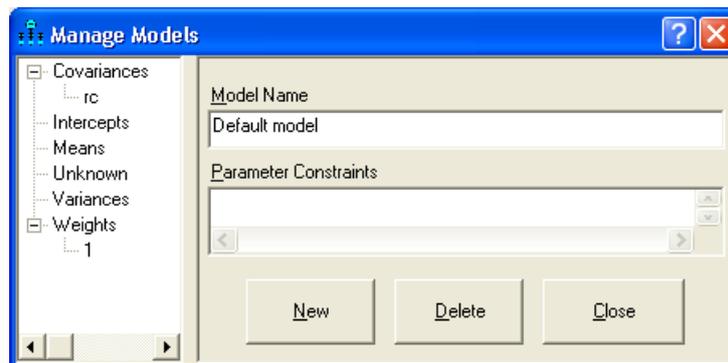
- Die manifesten Variablen sind multivariat normalverteilt (siehe Abschnitt 2.7.2).
- Das Modell  $M_u$  ist gültig.
- Die Stichprobe ist hinreichend groß (siehe Abschnitt 2.7.1).

Wie man leicht einsehen wird, ist z.B. auch der normale  $\chi^2$ -Modelltest ein Spezialfall für das  $D^2$ -Prinzip. Das uneingeschränkte Modell erlaubt in diesem Fall für jede empirische Kovarianz einen eigenen Parameter und hat daher Null Freiheitsgrade.

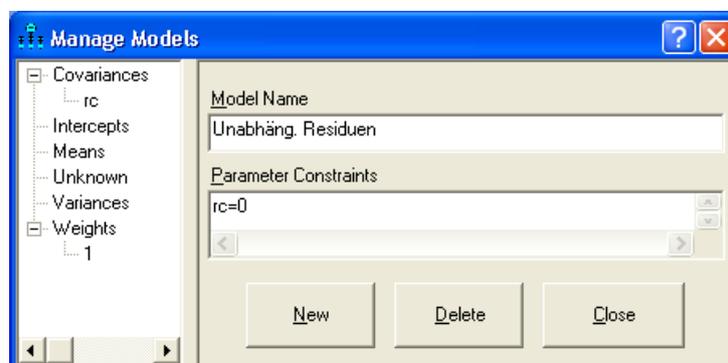
Für einen  $D^2$ -Test zwei Modelle spezifizieren und ihre Ergebnisse vergleichen zu müssen, ist auf Dauer etwas mühsam. Erfreulicherweise unterstützt Amos den Anwender beim Vergleich von Modellen samt  $D^2$ -Test. Nach

### Analyze > Manage Models

erscheint die Dialogbox **Manage Models**:

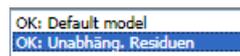


Hier legen wir mit **New** ein zusätzliches, eingeschränktes Modell ( $M_e$ ) an, das sich durch Parameterrestriktionen aus dem Standardmodell ( $M_u$ ) ergeben muss. In unserem Ambitionsbeispiel setzen wir die Fehlerkovarianz auf Null, wobei endlich die in Abschnitt 3.7 vereinbarte Bezeichnung **rc** zum Einsatz kommt:

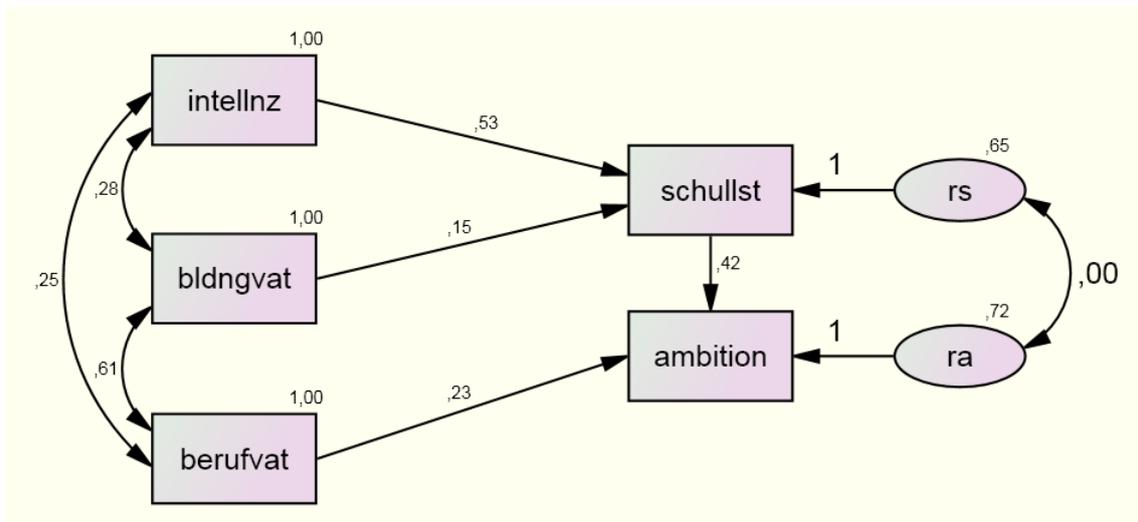


Die von Amos vorgegebene Modellbezeichnung kann beliebig geändert werden.

Nach der nächsten Schätzung mit  können wir per Mausklick auf die Modellbezeichnung zwischen zwei Ausgabe-Pfaddiagrammen wählen und die jeweiligen Parameterschätzungen vergleichen:



Hier ist das Pfadmodell (mit unkorrelierten Residuen) zu sehen:



Die Textausgabe beschreibt *beide* Modelle und liefert den  $D^2$ -Test (siehe Abschnitt **Model Comparison**):

**Nested Model Comparisons**

Assuming model Default model to be correct:

Model	DF	CMIN	P	NFI Delta-1	IFI Delta-2	RFI rho-1	TLI rho2
Unabhäng. Residuen	1	3,399	,065	,003	,003	,007	,007

Im unserem Beispiel setzt der  $D^2$ -Test die Gültigkeit des Modells mit frei schätzbare Fehlerkovarianz voraus und prüft, ob das restringierte Modell (mit der auf Null fixierten Fehlerkovarianz) beibehalten werden kann. Er endet wie der verwandte CR-Test (siehe Abschnitt 3.7) mit einer knappen Entscheidung für die Nullhypothese ( $p = 0,065$ ).

Die mit **CMIN** überschriebene Tabelle im Abschnitt **Model Fit**

**CMIN**

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	13	2,596	2	,273	1,298
Unabhäng. Residuen	12	5,995	3	,112	1,998
Saturated model	15	,000	0		
Independence model	5	1008,354	10	,000	100,835

liefert auch für  $M_e$  den  $\chi^2$ -Gültigkeitstest, der mit  $p = 0,112$  das sparsame und damit zu bevorzugende Modell akzeptiert.

**5.6 Gütekriterien zur Modellbeurteilung**

In diesem Abschnitt werden einige **Gütekriterien** vorgestellt, welche die bisher ausschließlich betrachtete Prüfgröße des  $\chi^2$ -Gültigkeitstests (samt Überschreitungswahrscheinlichkeit) bei der Modellbeurteilung wesentlich ergänzen können.

Bei den folgenden Ausführungen verwenden wir einige Bezeichnungen aus der Amos-Ausgabe:

- **CMIN, P**  
Prüfgröße des  $\chi^2$ -Modellgültigkeitstests mit der zugehörigen Überschreitungswahrscheinlichkeit **P**
- **NPAR**  
Anzahl der frei schätzbaren Parameter
- **N**  
Stichprobengröße
- **DF**  
Anzahl der Freiheitsgrade (= Anzahl der Stichprobenmomente – **NPAR**)

Praktisch alle Kriterien sind als Abweichungsindikatoren konzipiert, fallen also mit zunehmender Modellgüte kleiner aus.

### 5.6.1 CMIN und die zugehörige Überschreitungswahrscheinlichkeit P

In die Prüfgröße **CMIN** des  $\chi^2$ -Modellgültigkeitstests gehen ein:

- Die minimale Diskrepanz **FMIN** zwischen der Stichproben-Momentenmatrix und der vom geschätzten Modell implizierten Momentenmatrix, wobei die Diskrepanzdefinition von der Schätzmethode abhängt
- Die um Eins verminderte Stichprobengröße **N**

Speziell die starke Abhängigkeit von **N** beeinträchtigt die Nützlichkeit der Prüfgröße **CMIN** (und der zugehörigen Überschreitungswahrscheinlichkeit **P**) zur Modellbeurteilung bei größeren Stichproben:

$$\mathbf{CMIN} := (\mathbf{N} - 1) \cdot \mathbf{FMIN}$$

Bei gleicher Stichproben-Momentenmatrix und gleichem Modell führt ein doppelt so großes **N** zu einer ungefähr doppelt so großen **CMIN**-Statistik. Bei Gültigkeit des Modells ist dies eine sinnvolle Normierung, weil bei wachsendem **N** eine immer besserer Fit von Modell und Daten zu erwarten ist. Ist das Modell jedoch streng genommen falsch, was in den Sozial- und Wirtschaftswissenschaften als sicher angenommen werden kann, so konvergiert die Diskrepanz **FMIN** bei wachsendem **N** *nicht* gegen Null, und **CMIN** steigt drastisch an. Die Power des  $\chi^2$ -Tests zur Aufdeckung der als sicher anzunehmenden Modelldefekte steigt mit dem Stichprobenumfang. Damit geht auch für ein „gut passendes“ Modell die Wahrscheinlichkeit gegen Eins, dass es verworfen wird.

### 5.6.2 Relation aus CMIN und DF

Bei den Schätzmethoden mit approximativer  $\chi_{df}^2$ -Verteilung der **CMIN**-Statistik (ML, GLS, ADF) ist für *gültige* Modelle ein Quotient

$$\frac{\mathbf{CMIN}}{\mathbf{DF}}$$

nahe Eins zu erwarten. Empfehlungen zur Obergrenze für *akzeptable* Modelle schwanken zwischen 2 und 5 (Kline 2005, S. 137).

### 5.6.3 Root Mean Square Error of Approximation (RMSEA)

Als **F0** bezeichnet Amos die folgende Schätzung der Diskrepanz zwischen der *Populations*-Momentenmatrix und der Modellimplikation:

$$F0 := \max\left(\frac{CMIN - DF}{N - 1}, 0\right)$$

Während das **F0**-Kriterium keine Belohnung für sparsame Modelle (mit vielen Freiheitsgraden) enthält, ist diese im **RMSEA**-Maß (*Root Mean Square Error of Approximation*) enthalten:

$$RMSEA := \sqrt{\frac{F0}{DF}}$$

Amos liefert unter der Bezeichnung **LO90** bzw. **HI90** auch eine untere bzw. obere Schranke für das 90% - Vertrauensintervall zum RMSEA-Wert.

Browne & Cudeck (1993, S. 136ff) geben zum **RMSEA**-Maß folgende Bewertungshinweise:

<b>RMSEA</b> ≤ 0,05	gut
0,05 < <b>RMSEA</b> ≤ 0,08	akzeptabel
<b>RMSEA</b> > 0,08	inakzeptabel

Andere Autoren sehen die Grenze für akzeptable Modelle bei 0,10 (siehe Loehlin 2004, S. 69).

Amos liefert unter der Bezeichnung **PCLOSE** eine Größe, die tendenziell als Überschreitungswahrscheinlichkeit zur folgenden **Close-Fit - Nullhypothese** interpretiert werden kann:

$$H_0^c : RMSEA \leq 0,05$$

Diese Nullhypothese *geringfügiger Anpassungsfehler* ist zur Modellbeurteilung erheblich relevanter als die Nullhypothese der perfekten Modellgültigkeit, die im klassischen  $\chi^2$ -Anpassungstest begutachtet wird.

Kann diese Close-Fit - Nullhypothese nicht verworfen ist, ist sie aber keinesfalls bewiesen. Speziell bei Modellen mit wenigen Freiheitsgraden hat der Test zur Close-Fit – Nullhypothese oft eine geringe Power (Kline 2005, S. 158). Von großer Relevanz für die Modellbeurteilung ist daher ein Test zur folgendem **Poor-Fit - Nullhypothese**, die ein inakzeptables Modell behauptet:

$$H_0^p : RMSEA \geq 0,10$$

Nach Loehlin (2004, S. 69) kann diese Hypothese verworfen werden, wenn die obere Grenze das 90-prozentigen RMSEA-Vertrauensintervalls unter dem Kriterium für akzeptable Modell (0,10) liegt. Aus der Wahl des 90-prozentigen Vertrauensintervalls resultiert ein gerichteter (einseitiger) Test zum Niveau  $\alpha = 0,05$ .

Im Ambitionsbeispiel erreichen das nonrekursive Modell (mit korrelierten Residuen) und das Pfadmodell (mit unkorrelierten Residuen) eine gute RMSEA-Beurteilung (mit einem Schätzwert unter 0,05):

Model	RMSEA			
	RMSEA	LO 90	HI 90	PCLOSE
Default model	,020	,000	,077	,741
Unabhäng. Residuen	,036	,000	,078	,646
Independence model	,361	,342	,380	,000

Die Close-Fit - Nullhypothese kann jeweils beibehalten werden ( $p = 0,741$  bzw.  $0,646$ ), und die Poor-Fit - Nullhypothese kann verworfen werden.

Wie eine von Loehlin (2004, S. 264) berichtete Teststärketabelle zeigt, steigt die Wahrscheinlichkeit, bei einem wahren RMSEA-Populationswert von 0,05 die Poor-Fit - Nullhypothese bei einem  $\alpha$ -Niveau von

0,05 verwerfen zu können, mit der Anzahl der Freiheitsgrade eines Modells an. Für eine gewünschte Teststärke von 0,80 werden in Abhängigkeit vom *df*-Wert folgende Stichprobenumfänge benötigt:

<i>df</i>	benötigtes <i>N</i>
1	2475
5	569
10	320
50	99

Weil ein großer *df*-Wert am ehesten in komplexen Modellen zu realisieren ist, haben diese wohl insgesamt eine erheblich bessere Chance, den RMSEA-basierten Poor-Fit - Test zu bestehen.

### 5.6.4 SRMR

Der ursprünglich im Computer-Programm LISREL (siehe Jöreskog & Sörbom 1989) vorgestellte **RMR**-Index (*Root Mean Square Residual*) gibt das mittlere betragsmäßige Kovarianz-Residuum an. Im Eingruppenfall ist der RMR-Index folgendermaßen definiert:

$$RMR := \sqrt{\frac{1}{\frac{1}{2} k(k+1)} \sum_{i=1}^k \sum_{j=1}^i (s_{ij} - \hat{\sigma}_{ij})^2}$$

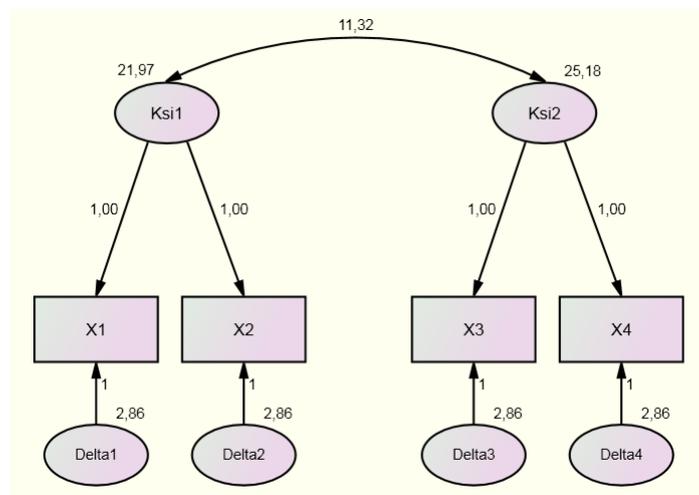
In der Formel steht  $s_{ij}$  bzw.  $\hat{\sigma}_{ij}$  für das Element ( $i, j$ ) der empirischen bzw. modellimplizierten Kovarianzmatrix der  $k$  manifesten Variablen.

Der RMR-Index ist nur dann sinnvoll interpretierbar, wenn alle manifesten Variablen dieselbe Maßeinheit besitzen, z.B. standardisiert sind. Im *standardisierten* RMR-Index (SRMR), der die beobachtete mit der modellimplizierten Korrelationsmatrix vergleicht, ist das Problem behoben. SRMR-Werte bis zur Grenze 0,10 gelten als akzeptabel (Kline 2005, S. 141).

Während Amos den RMR in der Textausgabe routinemäßig liefert, ist für den SRMR ein Plugin zuständig:

- Nach dem Menübefehl **Plugins > Standardized RMR** erscheint ein leeres Fenster.
- Läst man anschließend das Modell schätzen, wird dort der SRMR angezeigt.

Für das in Abschnitt 4 vorgestellte CFA-Modell der Zustandsangst

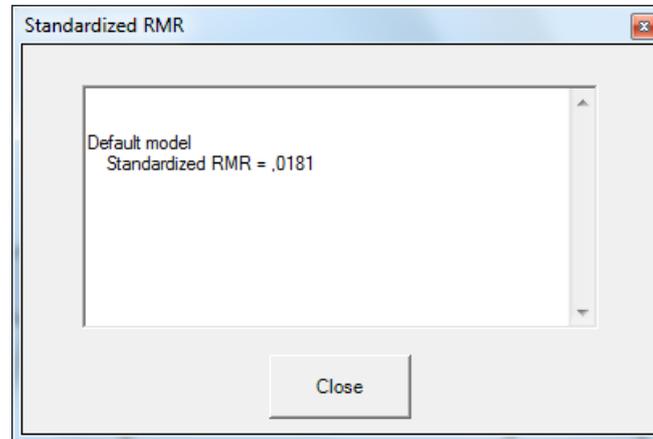


liefert die Textausgabe im Abschnitt **Model Fit** den wenig informativen RMR-Wert 0,668:

### RMR, GFI

Model	RMR	GFI	AGFI	PGFI
Default model	,668	,980	,966	,588
Saturated model	,000	1,000		
Independence model	12,779	,462	,104	,277

Den aussagekräftigen und im Beispiel sehr günstigen SRMR-Wert von 0,02 erhält man im Plugin-Fenster:



### 5.6.5 Informationskriterien zum Vergleich von nicht-geschachtelten Modellen

Die folgenden Gütemaße

- **Akaike Information Criterion (AIC)**
- **Browne-Cudeck Criterion (BCC)**
- **Bayes Information Criterion (BIC)**
- **Consistent AIC (CAIC)**

sind informationstheoretisch begründet und bestrafen komplexe Modelle (mit zahlreichen Parametern). Man verwendet sie oft zum Vergleich von nicht-geschachtelten Modellen, weil hier kein  $D^2$ -Test (vgl. Abschnitt 5.5) möglich ist.

Wir betrachten nur beim Akaike Information Criterion, welches die historische Vorreiterrolle gespielt hat, die hier angenehm einfache Definition:

$$\text{AIC} := \text{CMIN} + 2 \text{ NPAR}$$

Hinsichtlich der Tendenz zur Bevorzugung sparsamer Modelle bilden die Gütemaße folgende Rangreihe:

$$\text{AIC} < \text{BCC} < \text{CAIC} < \text{BIC}$$

Um den Modellvergleich zu vereinfachen, wird gelegentlich vom jeweiligen Gütemaß das Minimum der zum Vergleich anstehenden Modelle subtrahiert, so dass beim besten Modell der Wert Null resultiert. Amos präsentiert die entsprechenden Varianten der Informationsmaße unter den Bezeichnungen **AIC<sub>0</sub>**, **BCC<sub>0</sub>** und **BIC<sub>0</sub>**.

Im Ambitionsbeispiel (siehe Abschnitt 3.7) entscheiden sich die beiden Kriterien mit der stärksten Komplexitätsstrafe (BIC, CAIC) für das sparsamere Modell mit unabhängigen Residuen. Die Amos-Textausgabe liefert im Abschnitt **Model Fit** die folgende Tabelle:

**AIC**

Model	AIC	BCC	BIC	CAIC
Default model	28,596	28,801	88,948	101,948
Unabhäng. Residuen	29,995	30,184	85,705	97,705
Saturated model	30,000	30,237	99,637	114,637
Independence model	1018,354	1018,433	1041,567	1046,567

## 6 Ein nonrekursives Modell mit latenten Variablen

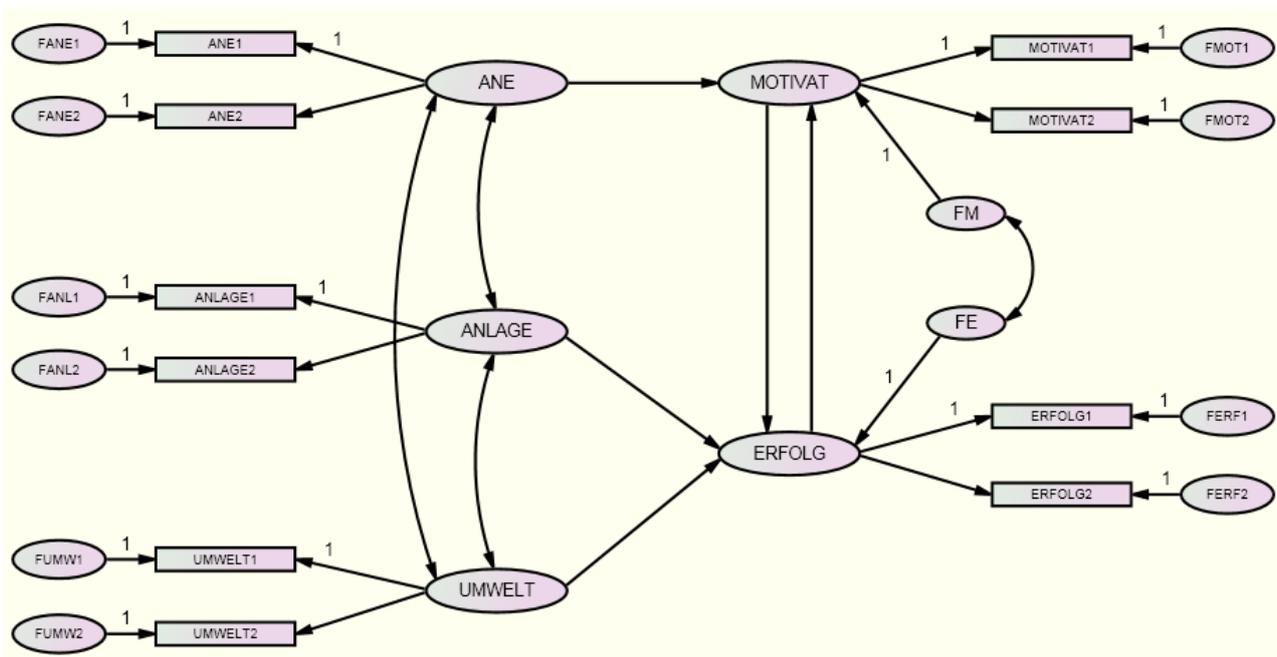
### 6.1 Messfehler und methodologische Gegenmaßnahmen

Bisher haben wir ein Strukturmodell für manifeste Variablen zur Beschreibung von Wirkungsbeziehungen (Ambitionsbeispiel, siehe Abschnitt 3.7) und ein Messmodell zur Verankerung von latenten Variablen in der Beobachtungsbasis behandelt (konfirmatorische Faktorenanalyse mit zwei Faktoren der Zustandsangst, siehe Abschnitt 4). Im nächsten Beispiel, dessen Pfaddiagramm schon in der Einleitung zu sehen war, kommen die beiden Partialmodelle zusammen zu einem Strukturmodell für latente (endogene und exogene) Variablen samt Messmodell. Damit lassen sich die folgenden, durch messfehlerbehaftete manifeste Variablen bedingten Probleme zu vermeiden (siehe Arbuckle 2009a, S. 81 und S. 145):

- Fehler in den *endogenen* Variablen führen zu geminderten Determinationskoeffizienten und zu vergrößerten Standardfehlern der Strukturkoeffizienten.
- Fehler in den *exogenen* Variablen mindern die Determinationskoeffizienten und verzerren die Schätzungen der Strukturkoeffizienten.

Um Effekte und Kovarianzen auf der Ebene von latenten, messfehlerfreien Variablen schätzen zu können, benötigt man im Allgemeinen für jede latente Variable mindestens zwei manifeste Indikatoren. Erst bei drei Indikatoren ist unter relativ allgemeinen Bedingungen garantiert, dass die Identifikation und erfolgreiche Schätzung des Modells nicht an der mangelhaften empirischen Verankerung der latenten Variablen scheitert (siehe Abschnitte 5.3 und 5.4.3). Grundsätzlich kann man aber auch *einfach* verankerte latente Variablen modellieren, sofern die Reliabilität des Einzelindikators bekannt ist (siehe z.B. Kline 2005, S.229 ff).

Beim dem bereits in der Einleitung gezeigte Modell<sup>1</sup> sind die latenten Variablen jeweils doppelt empirisch verankert:



Zur Identifikation der Varianzen der latenten Variablen, zu denen auch die Residuen gehören, wird in diesem Modell der üblichen Praxis folgend jeweils eine Ladung auf Eins fixiert (*Unit Loading Identification*, ULI). Bei den *exogenen* latenten Variablen könnte man alternativ auch die Varianzen auf Eins fixieren (*Unit*

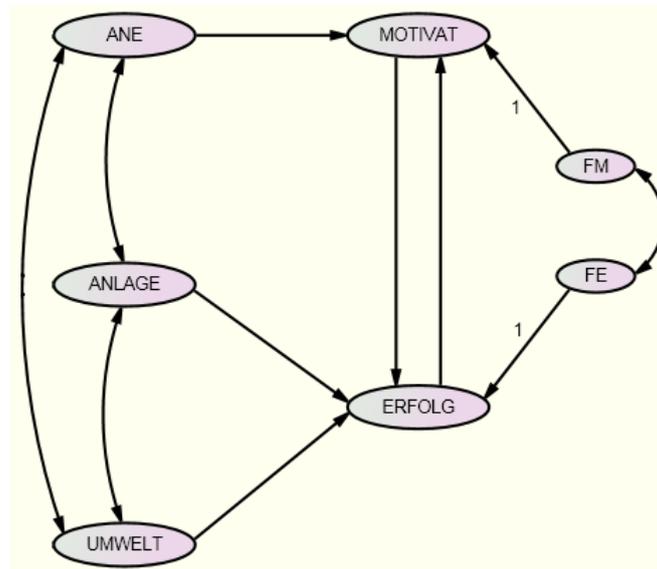
<sup>1</sup> Das Pfaddiagramm in der Einleitung enthält die *standardisierten* Schätzergebnisse.

*Variance Identification*, UVI). Beide Techniken sind in der Regel äquivalent, wobei aber bestimmte Konstellationen von Identifikationsrestriktionen einerseits und Gleichheitsrestriktionen für Varianzen oder Ladungen andererseits für Ausnahmen sorgen können (so genannte *constraint interactions*, siehe Kline 2005, S. 190).

Zur empirischen Verankerung der latenten Variablen in einem Strukturgleichungsmodell werden meist (wie in unseren Beispielen) **reflektive Indikatoren** verwendet, die Effekte von latenten Variablen erhalten. Gelegentlich werden so genannte **formative Indikatoren** diskutiert, die in umgekehrter Richtung Effekte auf latente Variablen ausüben und deren Bedeutung somit prägen. Ein so genanntes **MIMIC-Modell** (*Multiple Indicators Multiple Causes*) enthält latente Variablen, die einerseits Effekte auf reflektive Indikatoren ausüben und andererseits von manifesten Variablen beeinflusst werden.

## 6.2 Ein nonrekursives Strukturmodell

Im Unterschied zum faktorenanalytischen Beispiel aus Abschnitt 4 wird im aktuellen Beispiel das Messmodell mit den Beziehungen zwischen den latenten und manifesten Variablen ergänzt durch ein restriktives (empirisch gehaltvolles, falsifizierbares) Strukturmodell zu den Beziehungen zwischen den latenten Variablen:



Aus einer künstlichen Population mit der in obigem Pfaddiagramm beschriebenen Struktur wurde mit einem SPSS-Programm eine Zufallsstichprobe der Größe  $N = 384$  gezogen.<sup>1</sup> Neben stetigen und normalverteilten Indikatoren wurden auch vergrößerte (ordinale) Varianten mit ca. 25 Stufen erstellt, um die übliche Messpraxis nachzuempfinden.

Das obige Pfaddiagramm formuliert eine pädagogische „Theorie“ über Leistungsmotivation (**MOTIVAT**) und Schulerfolg (**ERFOLG**), die in der künstlichen Population perfekt gilt. Zur Erklärung werden folgende exogene Variablen herangezogen:

- Anspruchsniveau der Eltern (**ANE**)
- genetische Ausstattung (**ANLAGE**)
- Förderungspotential der Lernumgebung (**UMWELT**)

Weil es sich um ein *nonrekursives* Modell mit wechselseitiger Abhängigkeit von Leistungsmotivation und Erfolg handelt, sorgt das erzeugende SPSS-Programm in einer „Zeitschleife“ dafür, dass ein stabiler Zu-

<sup>1</sup> Sie finden das erzeugende SPSS-Programm **mc.sps** und die Datendatei **mc.sav** mit den resultierenden Zufallsdaten an der im Vorwort vereinbarten Stelle.

stand ausagierter Effekte eingekehrt ist, bevor die „Datenerhebung“ stattfindet. Jede Erforschung eines dynamischen Systems durch querschnittlich erhobene Daten setzt einen Gleichgewichtszustand ausagierter Effekte voraus (siehe z.B. Kline 2005, S. 239).

Wir „forschen“ mit dem korrekten Modell, machen also keine Spezifikationsfehler. Die in obigem Pfaddiagramm enthaltenen Fixierungen von Pfadkoeffizienten auf den Wert Eins stellen keine empirischen Behauptungen dar, sondern legen die Maßeinheiten der latenten Variablen fest (vgl. Abschnitt 3.4).

### 6.3 Weitere Werkzeuge für die Erstellung von Pfaddiagrammen

Die zur Erstellung eines Pfaddiagramms mit latenten Variablen im Amos-Grafikmodus erforderlichen Werkzeuge haben wir im Zusammenhang mit der Residualvariablen bereits kennen gelernt. Amos bietet darüber hinaus noch einige Arbeitserleichterungen, die beim Erstellen des obigen Pfaddiagramms sehr nützlich sind:

- Stellen Sie über

#### View > Interface Properties > Page Layout

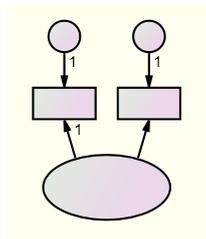
eine querformatige Zeichenfläche ein.

- Zeichnen Sie eine Ellipse für das latente Konstrukt ANE.
- Wählen Sie das Messmodell-Werkzeug mit



oder **Diagram > Draw Indicator Variable**

- Klicken Sie dann *zweimal* auf die Ellipse, um zwei Indikatoren einzeichnen zu lassen:



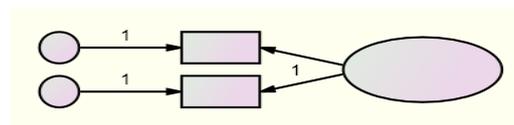
Amos hat die zur Definition der Maßeinheiten der latenten Variablen erforderlichen Parameterfixierungen automatisch vorgenommen.

- Wählen Sie mit



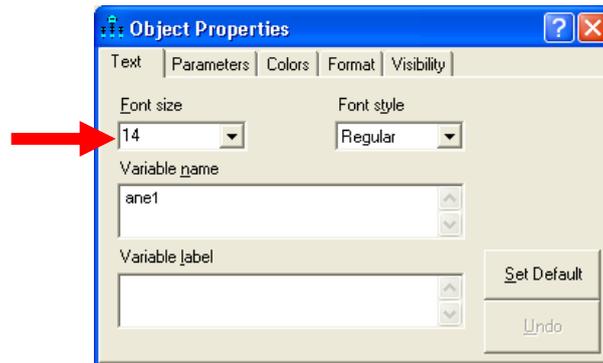
oder **Edit > Rotate**

das Rotationswerkzeug, und klicken Sie so lange auf die ANE-Ellipse, bis sich die Indikatoren auf der *linken* Seite befinden:



- Gehen Sie folgendermaßen vor, wenn Teile des Pfaddiagramms außerhalb der Zeichenfläche liegen:
  - Alles markieren mit dem -Werkzeug
  - Verschieben mit dem -Werkzeug
- Verfeinern Sie das Messmodell mit dem -Werkzeug.

- Ordnen Sie den fünf vorhandenen Variablen ihre Namen zu, um bei Platzproblemen frühzeitig die Rechtecke und Ellipsen vergrößern zu können, bevor Sie in die Serienproduktion einsteigen. Statt eine Form zu vergrößern kann auch die Beschriftung verkleinert werden, z.B.:



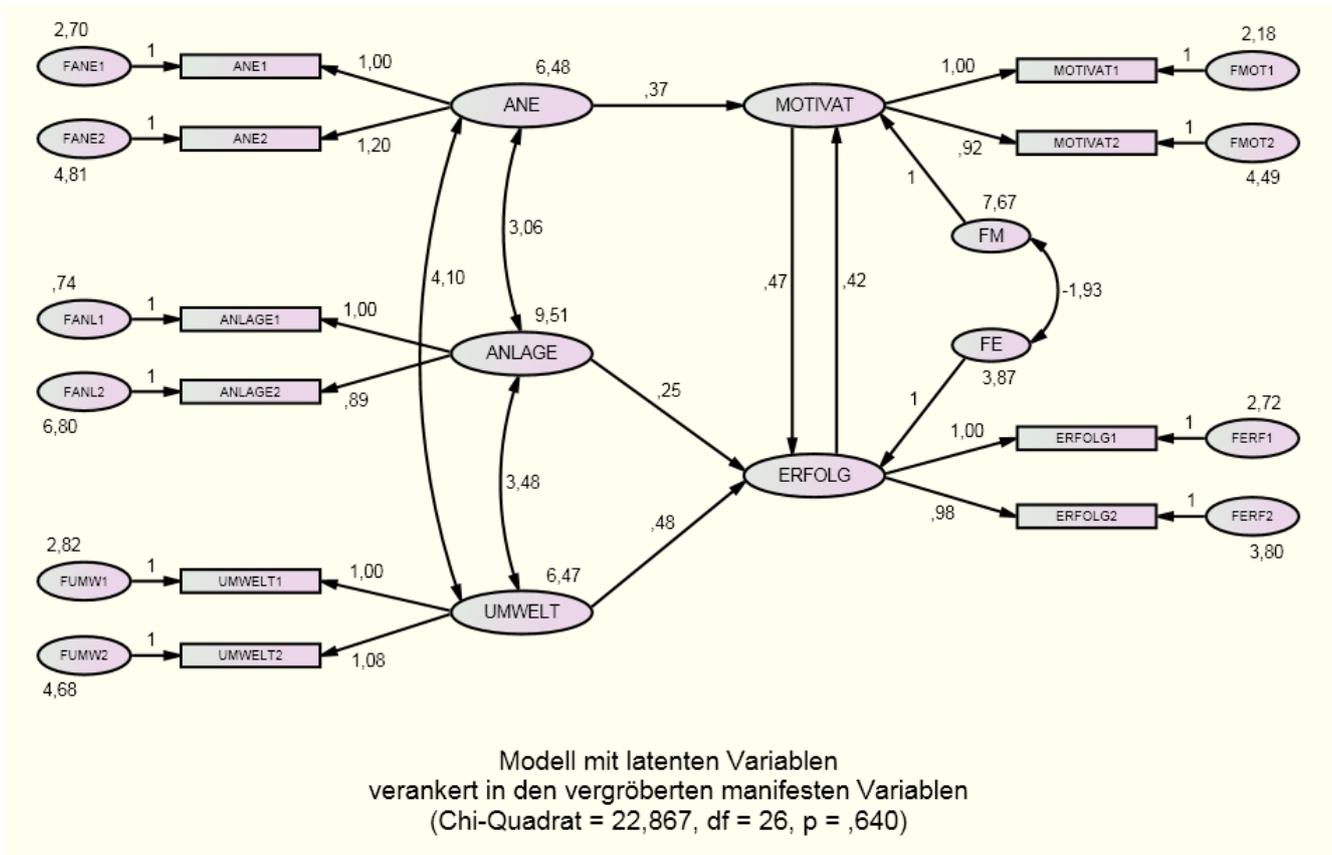
Die Namen für die latenten Variablen sind frei wählbar, soweit keine Namenskonflikte auftreten. Um die Namen der manifesten Variablen aus einer Liste übernehmen zu können, sollten Sie spätestens jetzt über **File > Data Files > File Name** die Eingabedatei festlegen. In der Datei **mc.sav** finden Sie die Indikatoren in stetiger (z.B. ANE1S) und in „vergrößerter“ Form (z.B. ANE1), wobei in diesem Manuskript die *vergrößerten* Varianten eingesetzt werden.

- Erzeugen Sie mit  für die beiden anderen exogenen latenten Variablen Kopien des (vollständig markierten) ANE-Messmodells. Halten Sie beim Verschieben der Kopien die -Taste fest, so dass die Messmodelle ihre horizontale Position behalten.
- Erzeugen Sie eine weitere Kopie für die erste *endogene* Variable, vorläufig möglichst weit vom rechten Rand der Zeichnungsfläche entfernt.
- Reflektieren Sie die Indikatoren der endogenen Variablen mit dem -Werkzeug, und bringen Sie das Messmodell in die gewünschte Position.
- Erzeugen Sie mit  eine Kopie für die zweite endogene Variable.
- Zeichnen Sie mit  je einen Fehlerterm für die endogenen latenten Variablen.
- Tragen Sie mit  bzw.  die Effekte bzw. Kovarianzen ein.  
Beim Optimieren der Pfeile leistet der Zauberstab  gute Dienste.
- Sorgen Sie für korrekte Variablennamen.

Nun sollte Ihr Pfaddiagramm dem oben wiedergegebenen Vorschlag ähneln.

### 6.4 Methodenstudie zum Nutzen von latenten Variablen

Lassen Sie Amos nun die Berechnungen zum fertig spezifizierten Modell ausführen. Es resultieren die folgenden (unstandardisierten) ML-Schätzergebnisse:



Wie es bei der Überprüfung eines *korrekten* Modells zu erwarten ist, kommt der  $\chi^2$ -Gültigkeitstest zu einem günstigen Ergebnis ( $p$ -Level = 0,64):

**Notes for Model (Default model)**

**Computation of degrees of freedom (Default model)**

Number of distinct sample moments: 55  
 Number of distinct parameters to be estimated: 29  
 Degrees of freedom (55 - 29): 26

**Result (Default model)**

Minimum was achieved  
 Chi-square = 22,867  
 Degrees of freedom = 26  
 Probability level = ,640

Auch die RMSEA-Ergebnisse sind ideal, was bei einem realen Modell kaum jemals passieren wird:

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	,000	,000	,034	,997
Independence model	,343	,331	,356	,000

Bei einem *nonrekursiven* Modell interessiert auch der **Stabilitätsindex**, der betragsmäßig kleiner als Eins sein muss und sich in unserem Beispiel als Produkt der beiden reziproken (unstandardisierten) Pfadkoeffizienten ergibt:

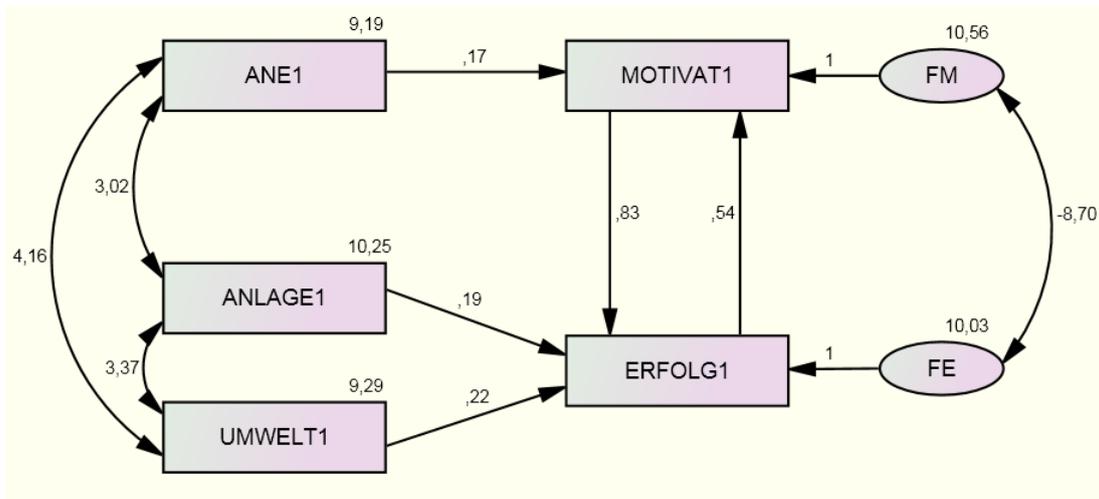
Notes for Group/Model (Group number 1 - Default model)

Stability index for the following variables is ,200

ERFOLG  
MOTIVAT

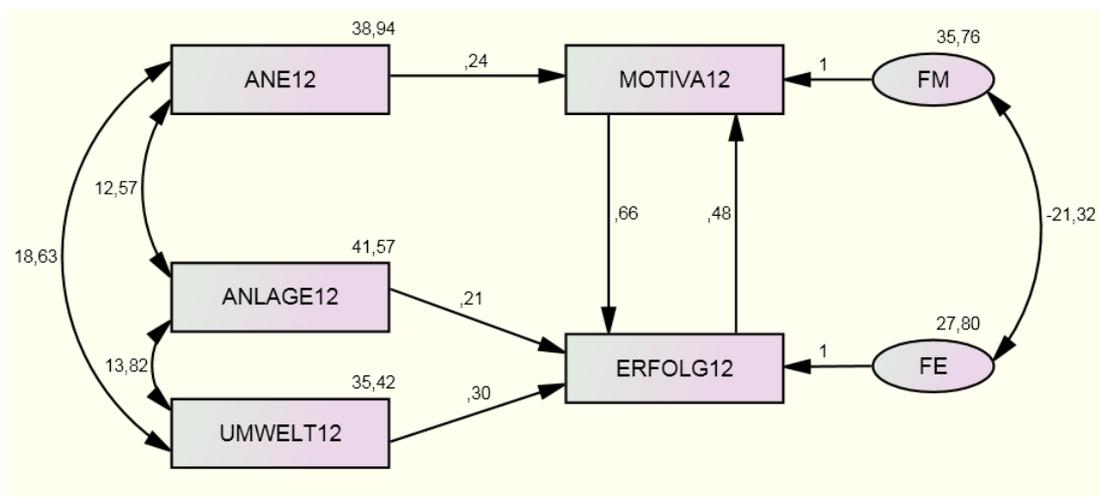
Wie ein Vergleich mit dem datengenerierenden SPSS-Programm zeigt, entsprechen die geschätzten Modellparameter recht gut den wahren Populationswerten. Dies gilt auch für die Reliabilitäten, die im Pfaddiagramm mit den *standardisierten* Schätzergebnissen (siehe Einleitung) über den Indikatoren eingetragen sind.

Zu Vergleichszwecken sollen auch die Ergebnisse einer rein *manifesten* Forschungsmethodologie betrachtet werden. Verwendet man an Stelle der latenten Variablen jeweils den *ersten* Indikator, ergibt sich das folgende Ausgabe-Pfaddiagramm mit unstandardisierten Schätzergebnissen:



Wir erhalten zwar ein vergleichbares Ergebnismuster, doch sind viele Schätzungen deutlich verzerrt.

Es ist allerdings unfair, Schätzungen auf der Basis von jeweils *zwei* Indikatoren pro Konstrukt mit Schätzungen auf der Basis der „halben“ Information zu vergleichen. Verwendet man an Stelle des ersten Indikators jeweils die Summe aus beiden Indikatoren, resultieren folgende Ergebnisse:



Durch die Testverlängerung und die damit verbundene Reliabilitätssteigerung gewinnt die manifeste Schätzmethodologie erheblich an Präzision, bleibt aber der latenten Methodologie unterlegen, wie die folgende Tabelle (mit unstandardisierten Schätzergebnissen) zeigt:

Parameter	Wahrer Wert	Schätzer für latente Variablen	Schätzer für manifeste Variablen (Summe der Ind.)	Schätzer für manifeste Variablen (erster Indikator)
MOTIVAT ← ANE	0,3	0,37	0,24	0,17
MOTIVAT ← ERFOLG	0,4	0,42	0,48	0,54
ERFOLG ← ANLAGE	0,3	0,25	0,21	0,19
ERFOLG ← UMWELT	0,4	0,48	0,30	0,22
ERFOLG ← MOTIVAT	0,5	0,47	0,66	0,83

Vermutlich resultiert die Überlegenheit der latenten Methodologie aus der Möglichkeit, die unterschiedlichen Reliabilitäten der Indikatoren zu berücksichtigen.

Bei den Vergleichen ist zu beachten, dass sie auf Ergebnissen einer einzigen Zufallsstichprobe basieren.

Zum Abschluss noch ein Hinweis zur Gestaltung der Pfaddiagramme: Wenn sich die von Amos eingetragenen Parameterschätzer zu nahe kommen, können Sie mit dem -Werkzeug deren Positionen verändern (bei aktivem Modellspezifikations-Pfaddiagramm).

## 7 Explorative Techniken

Trotz aller Argumente für eine primär konfirmatorische (hypothesenprüfende) Forschungsstrategie ist es legitim, aus der Empirie zu lernen, d.h. Daten explorativ (hypothesenbildend) zu analysieren. Beim Arbeiten mit Amos konfrontiert man ein Modell mit einer empirischen Kovarianzmatrix und erhält dabei eine Fülle diagnostischer Informationen, darunter u.a. auch Hinweise zu möglichen Ansatzpunkten für eine Verbesserung des Modells. Vor einer Diskussion von Techniken zur Verbesserung der Fit-Beurteilung muss jedoch nachdrücklich daran erinnert werden, dass ein mit Hilfe diagnostischer Informationen aus einer bestimmten Stichprobe optimiertes Modell nur an einer unabhängigen Stichprobe geprüft werden kann!

Als Strategie zur Überprüfung bzw. Modifikation eines Strukturgleichungsmodells empfiehlt Kline (2005, S. 216), zunächst das Messmodell (mit freien Beziehungen zwischen den latenten Variablen) zu analysieren. Hinter diesem Vorschlag steckt die Idee, Messung (Begriffsbildung) und Theorienbildung voneinander zu trennen, was wohl nicht regelmäßig gelingen kann und daher nicht unbedingt versucht werden sollte.

Beim Strukturmodell ist auch die begriffliche Eigenständigkeit der latenten Variablen zu beurteilen, wobei z.B. eine Korrelation von ca. 0,9 zwischen zwei latenten Variablen auf eine mangelhafte **divergente Validität** hindeutet (Kline 2005, S. 224).

Eine Verbesserung der Fit-Beurteilung kann auf zwei Wegen erreicht werden:

- **Aufhebung von Parameterfixierungen - bzw. Restriktionen**  
Die in Abschnitt 7.1 behandelten Modifikationsindikatoren geben Hinweise dazu, welche Parameterfixierungen - bzw. Restriktionen besonders stark zum Fehlermaß des  $\chi^2$ -Gültigkeitstest beitragen.
- **Fixieren bzw. Restringieren von bislang freien Parametern**  
Während durch eine solche Maßnahme die  $\chi^2$ -Statistik (hoffentlich nur wenig) ansteigt, wird jeweils ein Freiheitsgrad gewonnen, so dass diverse Gütekriterien (z.B. RMSEA, Überschreitungswahrscheinlichkeit des  $\chi^2$ -Tests) profitieren können.

Bei der Modifikation eines Modells sollte man keinesfalls blind empirischen Indizes folgen (und z.B. Zufälligkeiten der Stichprobe kapitalisieren), sondern unbedingt auf die theoretische Begründung achten. Kline (2005, S. 149) berichtet von Methodenstudien, die bei datenbasierten Verbesserungsvorschlägen kaum Fortschritte in Richtung auf das korrekte Modell fanden, und resümiert treffend:

*Learn from your data, but your data should not be your teacher.*

### 7.1 Modifikationsindikatoren

In den bisherigen Beispielen haben alle Modelle so gut gepasst, dass wir uns nicht für Hilfsmittel zur Diagnose von Spezifikationsfehlern interessieren mussten. Im aktuellen Abschnitt beschäftigen wir uns damit, wie *lokale* Modellschwächen mit Hilfe der so genannten *Modifikationsindikatoren* aufgespürt werden können. Diese Statistiken geben für einen fixierten oder restringierten Parameter an, welche Reduktion der  $\chi^2$ -Statistik zum Modellgültigkeitstest zu erwarten ist, wenn der Parameter freigesetzt und anschließend das (erweiterte) Modell neu geschätzt wird.

Da ein Modifikationsindikator annähernd der Differenz der  $\chi^2$ -Prüfgrößen zu den Modellen mit und ohne die fragliche Fixierung bzw. Gleichheitsrestriktion entspricht, und weil beim Aufheben der Fixierung bzw. Restriktion gerade ein Freiheitsgrad verloren geht, liegt unter Annahme der Gültigkeit des liberalisierten Modells ein Spezialfall für den in Abschnitt 5.5 vorgestellten  $D^2$ -Test mit  $df = 1$  vor.

Daher kann ein Modifikationsindikator bei Verwendung der ML-, GLS- oder ADF-Schätzmethode anhand der  $\chi^2$ -Verteilung mit einem Freiheitsgrad beurteilt werden. Je nach bevorzugter Regel wird man Werte größer 4 ( $\approx$  95% - Quantil) oder 7 ( $\approx$  99% - Quantil) als Kritik der Empirie an der zugrunde liegenden Fixierung oder Restriktion interpretieren. Amos schätzt außerdem für jeden fixierten bzw. restringierten Parameter seine erwartete Veränderung bei Freisetzung (Bezeichnung im Ergebnisprotokoll: **Par Change**).

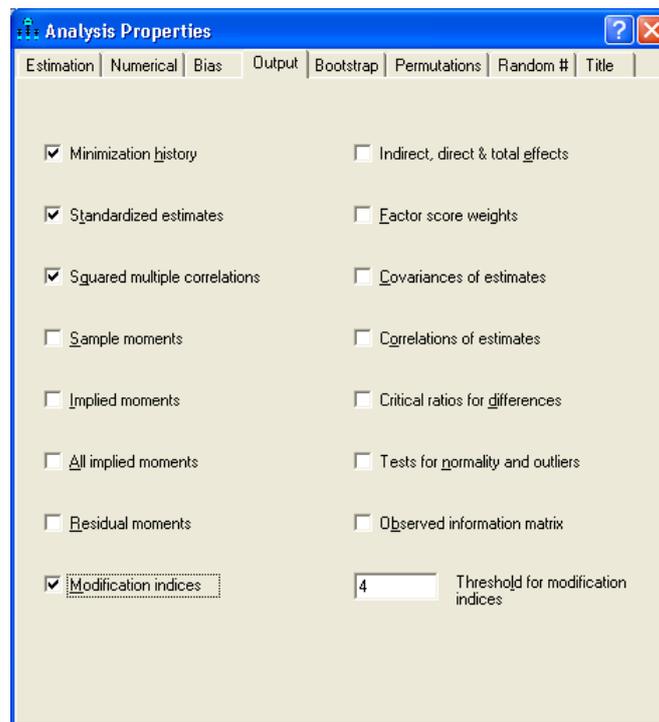
Es ist praktisch, dass Amos auf Anordnung Modifikationsindikatoren zu allen fixierten bzw. restringierten Parametern liefert, was die explizite Spezifikation von entsprechend vielen Vergleichsmodellen erspart.

Zur Demonstration der Modifikationsindikatoren greifen wir das in Abschnitt 4 vorgestellte faktorenanalytische Modell zur Zustandsangst auf, obwohl es den  $\chi^2$ -Modellgültigkeits gut bestanden hat ( $p = 0,274$ ). Allerdings lassen die zahlreichen Restriktionen des Modells lokale Anpassungsschwächen erwarten, die eventuell im globalen Test nicht zum Ausdruck kommen. Im Sinn des wissenschaftstheoretischen Sparsamkeitsgebots sollte man im konkreten Fall (beim relativ simplen Modell der Zustandsangst) wohl auf eine Fit-Optimierung verzichten. Bei einem komplexeren Modell geht aber eventuell ein gutes Ergebnis im globalen Modellgültigkeitstest auf exzellente Anpassung in nebensächlichen Modellregionen zurück, so dass man trotzdem auf Anpassungsschwächen an essentiellen Stellen achten sollte.

Wir verlangen nach

### View > Analysis Properties > Output

die Ausgabe von Modifikationsindikatoren:



Amos protokolliert nur Modifikationsindikatoren oberhalb der im Feld **Threshold for modification indices** angegebenen kritischen Grenze. Wir entscheiden uns für „Tests“ auf dem 5%-Niveau (siehe oben).

Leider kann Amos die Modifikationsindikatoren *nicht* berechnen, wenn wegen fehlender Werte die FIML-Schätzmethode (*Full Information Maximum Likelihood*) zum Einsatz kommt.

Wir fordern Amos über den Schalter  zur Berechnung der Schätzungen an und erhalten in der Textausgabe tatsächlich Hinweise auf lokale Schwächen des Modells durch Modifikationsindikatoren oberhalb der (bei  $\alpha = 0,05$ ) kritischen Grenze von 4:

**Modification Indices (Group number 1 - Default model)**

**Covariances: (Group number 1 - Default model)**

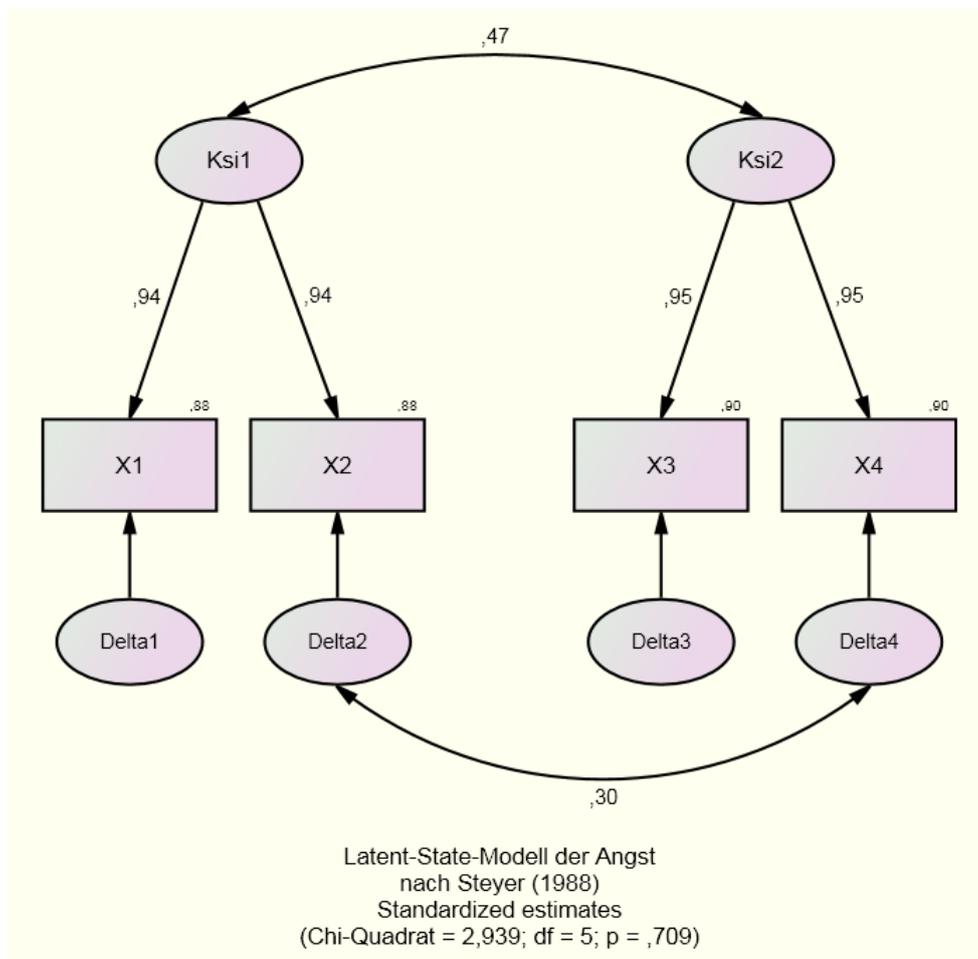
	M.I.	Par Change
Delta2 <--> Delta4	4,527	,849

**Regression Weights: (Group number 1 - Default model)**

	M.I.	Par Change
X3 <--- X2	4,600	-,075
X4 <--- X2	4,925	,077

Die Wahl der zuerst zu beseitigenden Restriktion sollte sich *nicht* strikt an der Modifikationsindikatoren-Rangreihe orientieren, sondern theoretisch begründbar sein. Im Beispiel ist der Modifikationsindex für die Kovarianz von  $\delta_2$  und  $\delta_4$  (im Modell auf Null fixiert) mit einem Wert von 4,53 signifikant und stellt eine durchaus nachvollziehbare Kritik an unserem Modell dar: Die beiden Messungen des zweiten Indikators haben vermutlich über den Einfluss der latenten Variablen hinaus eine gemeinsame Wurzel.

Durch Freisetzen der Kovarianz von  $\delta_2$  und  $\delta_4$  reduziert sich die  $\chi^2$ -Statistik zum Modellgültigkeitstest um 4,59 (etwas stärker als vom Modifikationsindikator prognostiziert) von 7,53 auf 2,94:



Für  $\delta_2$  und  $\delta_4$  wird eine nicht unerhebliche Korrelation von 0,3 geschätzt. Dass die bereits zuvor sehr günstige Beurteilung im  $\chi^2$ -Modellgültigkeitstest ( $p = 0,274$ ) noch zugelegt hat ( $p = 0,709$ ), ist nicht unbedingt entscheidungsrelevant. Beim RMSEA-Ergebnis zeigt sich jedoch eine argumentativ verwertbare Verbesserung, weil die Nullhypothese eines schlechten Fits (RMSEA  $\geq 0,10$ , siehe Abschnitt 5.6.3) vor der Modifikation *nicht* zurückgewiesen werden konnte (HO 90 = 0,11 > 0,10):

**RMSEA**

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	,038	,000	,110	,528
Independence model	,754	,704	,805	,000

Nach der Modifikation liegt die obere Grenze des Vertrauensintervalls sogar unter 0,08, der etwas strengen Grenze für akzeptable Modelle:

**RMSEA**

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	,000	,000	,078	,857
Independence model	,754	,704	,805	,000

In der Textausgabe zum erweiterten Modell finden sich *keine* signifikanten Modifikationsindikatoren mehr.

## 7.2 Spezifikationssuche

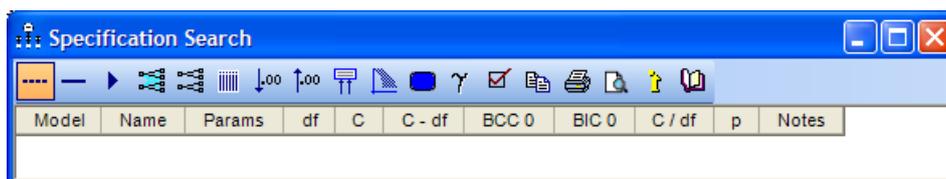
Amos kann einen durch *optionale* Parameter aufgespannten Raum von potentiellen Modellen nach einer „optimalen Lösung“ durchsuchen. Bei dieser Spezifikationssuche werden alle potentiellen Modelle geschätzt und nach diversen Kriterien beurteilt.

Wir greifen auf die in Abschnitt 6 beschriebenen Simulationsdaten und das dort vorgestellte Pfaddiagramm zurück und starten mit

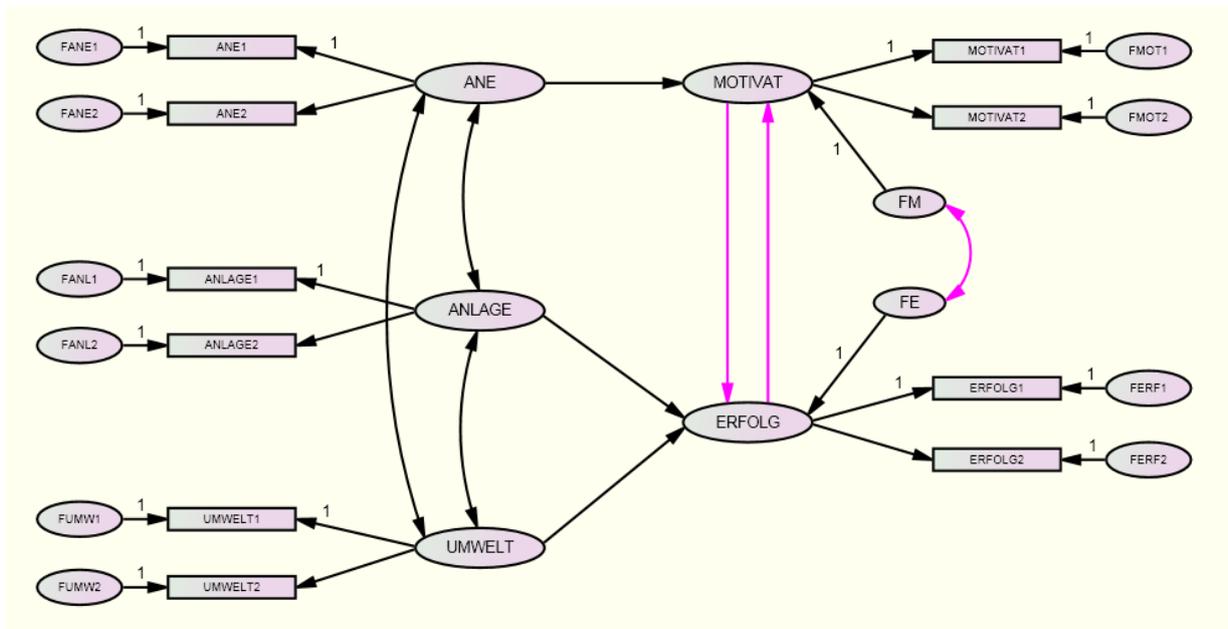


oder **Analyze > Specification Search**

das folgende Fenster:



Bei eingerastetem **----** - Schalter deklarieren wir im Pfaddiagramm per Mausklick drei Parameter als **optional** (erkennbar an einer alternativen Linienfarbe):



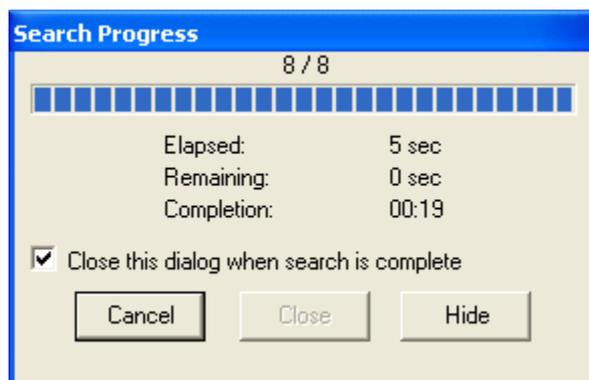
Bei eingerastetem  - Schalter kann man einen Parameter per Mausklick wieder in den Normalzustand zurückversetzen.

Die optionalen Parameter legen fest, welche Modelle bei der Spezifikationsuche zu berücksichtigen sind, wobei jeder einzelne optionale Parameter einen Faktor mit den folgenden Zuständen beisteuert:

- Parameter ist auf Null fixiert
- Parameter ist frei schätzbar

Gemeinsam ergeben  $k$  optionale Parameter also  $2^k$  Kombinationen bzw. Kandidatenmodelle, so dass im Beispiel acht Modelle zur Konkurrenz zugelassen sind.

Mit einem Mausklick auf den Schalter  der Spezifikations-Symbolleiste lässt man alle Modelle schätzen:



Anschließend erscheinen im Fenster der Spezifikationsuche die Ergebnisse aller Modelle, z.B.:

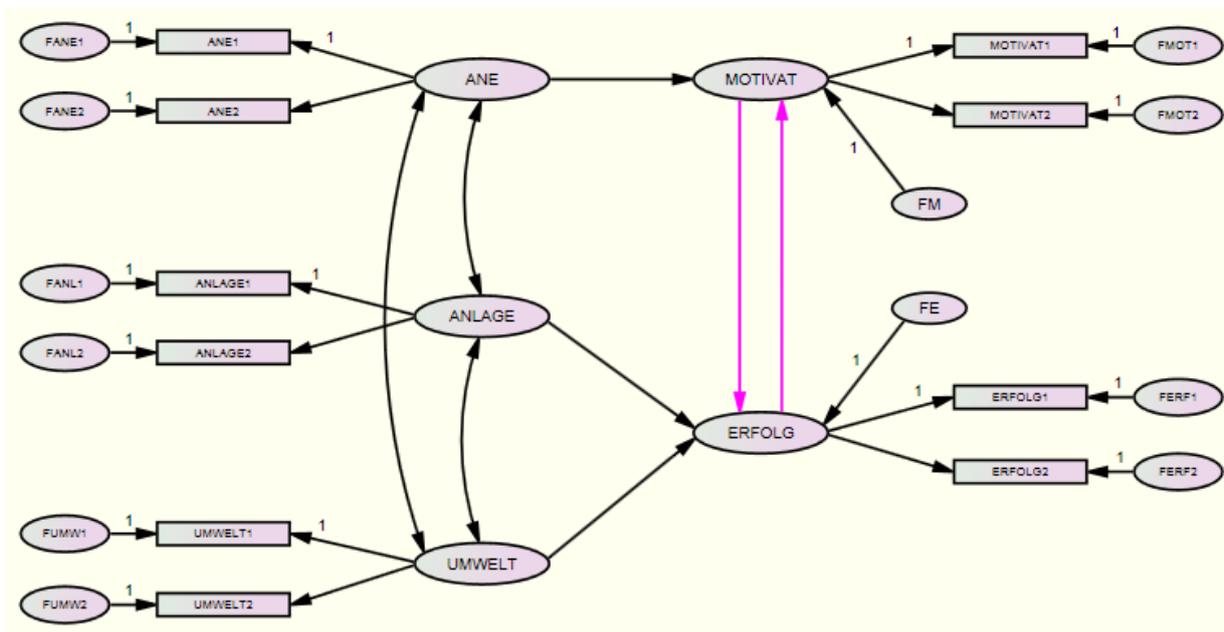
Model	Name	Params	df	C	C - df	BCC 0	BIC 0	C / df	p	Notes
1	Default model	26	29	140,177	111,177	112,016	104,233	4,834	0,000	
2	Default model	27	28	33,350	5,350	7,249	3,357	1,191	0,223	
3	Default model	27	28	35,390	7,390	9,289	5,398	1,264	0,159	
4	Default model	27	28	45,058	17,058	18,957	15,065	1,609	0,022	
5	Default model	28	27	24,042	-2,958	<u>0,000</u>	<u>0,000</u>	0,890	0,628	
6	Default model	28	27	28,024	1,024	3,981	3,981	1,038	0,410	
7	Default model	28	27	33,336	6,336	9,294	9,294	1,235	0,186	
8	Default model	29	26	22,867	<u>-3,133</u>	0,884	4,775	<u>0,879</u>	<u>0,640</u>	
Sat	[Saturated]	55	0	0,000	0,000	31,555	136,625			

In den vergleichsrelevanten Spalten ist der Wert des günstigsten Modells unterstrichen.

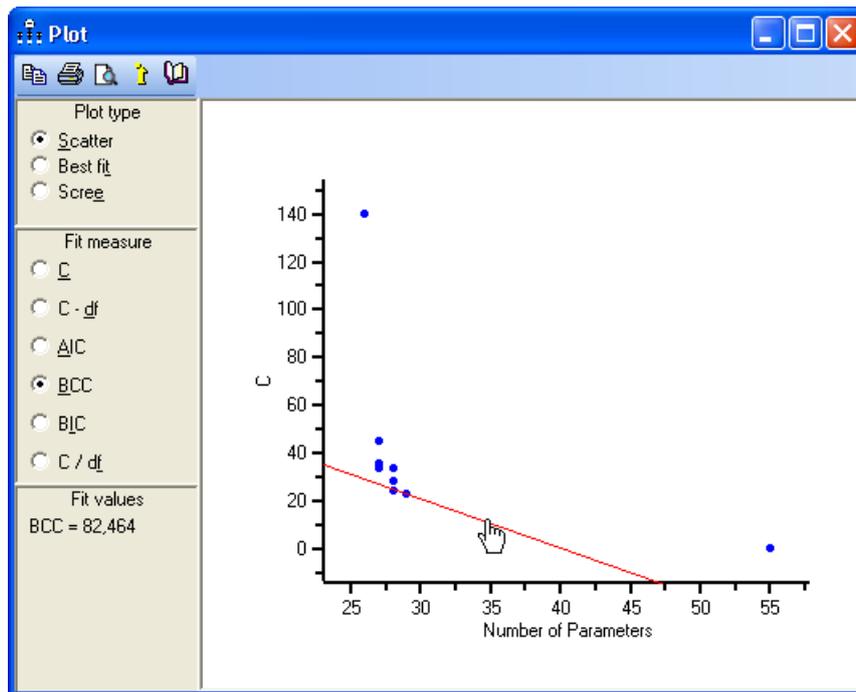
Bei Verwendung *mehrerer* Kriterien ist kaum damit zu rechnen, dass *ein* Modell *alle* Kriterien für sich entscheidet. Im Beispiel kommen zwei Modelle in die engere Wahl:

- Modell 8  
Dies ist das wahre Modell der Simulationsstudie (siehe Abschnitt 6). Es entscheidet die **CMIN** – abhängigen Disziplinen für sich.
- Modell 5  
Modell 5 gewinnt die informationstheoretischen Disziplinen. Hier ist im Vergleich zum Modell 8 die Kovarianz der Residuen zu den endogenen Variablen auf Null fixiert.

Per Doppelklick auf seine Zeile im Spezifikationsfenster erhält man zu einem Modell sein Pfaddiagramm, z.B. bei Modell 5:

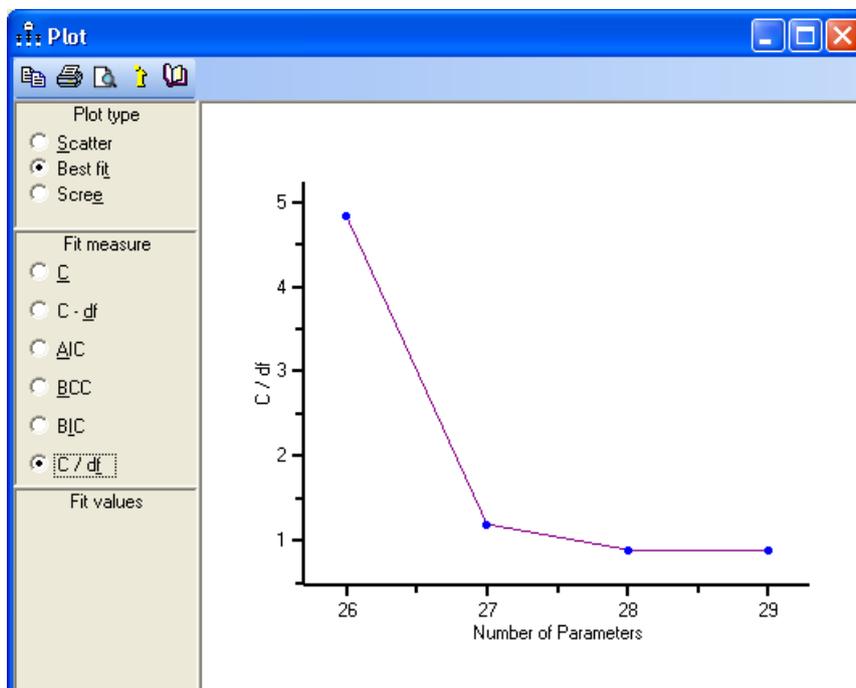


Über den Schalter  des Spezifikationsfensters ist eine Dialogbox zugänglich, die zahlreiche grafische Hilfen zur Modellauswahl bietet. Im **Scatter**-Plot



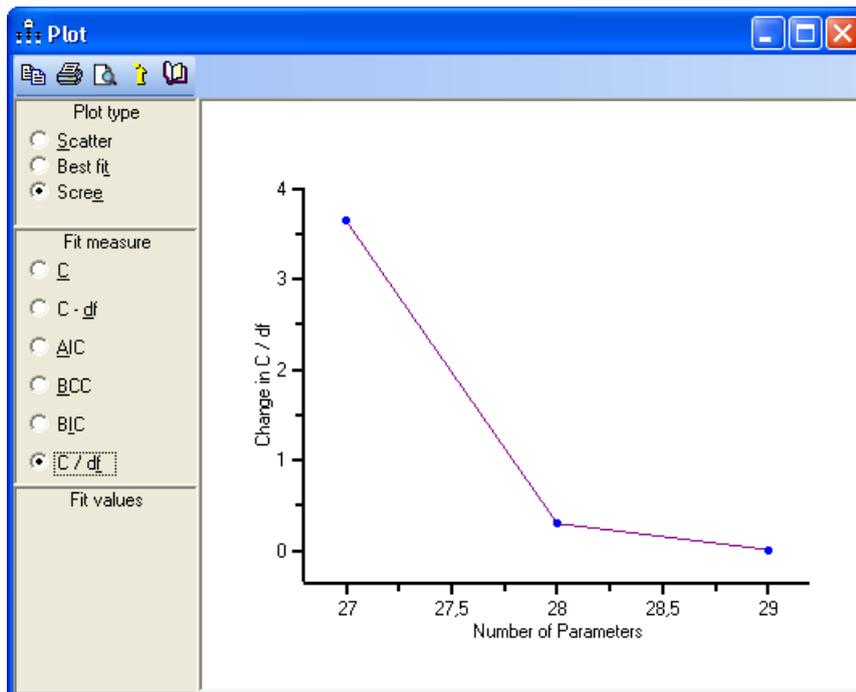
sind die **(CMIN, NPAR)** – Wertepaare aller Modelle zu sehen, wobei man per Mausklick auf einen Punkt die Nummer und den **CMIN**-Wert des zugehörigen Modells erhält. Eine verschiebbare Gerade kennzeichnet für ein **Fit measure** alle Modell mit konstantem Wert für dieses Kriterium. So zeigt sich in unserem Beispiel, dass die Modelle 5 und 8 praktisch denselben **BCC**-Wert besitzen.

In folgendem **Best fit** - Plot ist der beste **CMIN/DF** – Wert für ein Modell mit 26, 27, 28 oder 29 Parametern zu sehen:

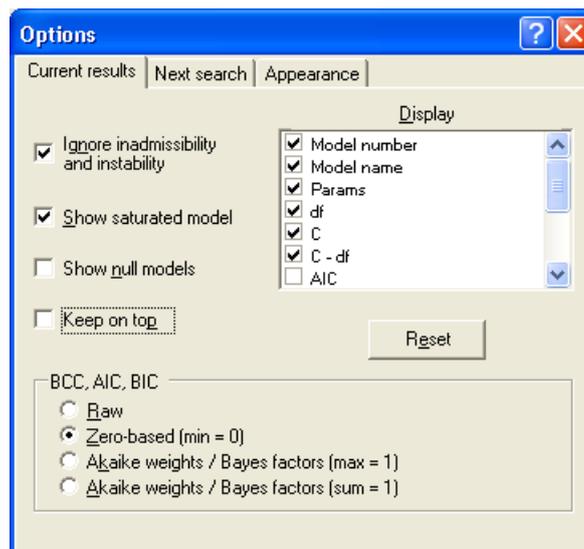


Offenbar empfiehlt es sich *nicht*, weniger als 27 Parameter zu verwenden.

Der folgende **Scree**-Plot dokumentiert für das beste Modell mit 27 Parametern eine deutliche Überlegenheit beim **CMIN/DF** – Kriterium gegenüber dem besten Modell mit 26 Parametern (Reduktion um 3,643), während bei einer weiteren Steigerung der Parameterzahl nur noch geringfügige Verbesserungen erreicht werden (0,301 bzw. 0,011):



Mit dem Schalter  erreicht man einen Optionsdialog:



Hier kann man z.B. über das Kontrollkästchen **Keep on top** festlegen, ob das Fenster der Spezifikationsuche ständig im Vordergrund bleiben soll.

## 8 Die simultane Analyse mehrerer Gruppen

Arbuckle (2009a, S. 159) nennt zwei wesentliche Vorteile der simultanen Analyse mehrerer Gruppen im Vergleich zu Einzelanalysen pro Gruppe:

- Es können Signifikanztests zu Gruppenunterschieden hinsichtlich diverser Verteilungseigenschaften durchgeführt werden (z.B. Vergleich von Kovarianzmatrizen, Ladungsmustern, Pfadkoeffizienten, Erwartungswerten). Dies ermöglicht insbesondere die Überprüfung von *Interaktionshypothesen*, allerdings mit der Einschränkung auf *kategoriale* Moderatoren (siehe Baltes-Götz 2008b). Man kann z.B. untersuchen, ob die Gruppenzugehörigkeit den Effekt einer latenten exogenen Variablen auf eine latente endogene Variable moderiert.
- Sind die Gruppen hinsichtlich wesentlicher Verteilungseigenschaften äquivalent, erlaubt die simultane Analyse präzisere Parameterschätzungen.

Im Demonstrationsbeispiel greifen wir wieder auf die Daten aus Attig (1983) zurück (vgl. Abschnitt 2), betrachten aber neben den jungen Probanden eine zusätzliche Stichprobe mit *alten* Teilnehmern. Dabei interessieren wir uns nur für die einfache und die unterstützte Erinnerungsleistung *vor* dem Gedächtnistraining (Variablen: **RECALL1**, **CUED1**).

Bei der Datenorganisation unterstützt Amos zwei Optionen:

- Die Daten jeder Gruppe befinden sich in einer eigenen Datei. In unserem Beispiel stehen die SPSS-Datendateien **Attg\_yng.sav** und **Attg\_old.sav** zur Verfügung, die sich im **Examples-**Unterverzeichnis zum Amos – Programmordner befinden.
- Eine einzelne Datei enthält die Daten aller Gruppen, und eine Variable definiert die Gruppenzugehörigkeit. Um diese Datenorganisation demonstrieren zu können, wurde die Datei **Attg.sav** mit den Daten beider Gruppen erstellt.<sup>1</sup> Die Variable **GROUP** hat folgende Werte:
  - 1 für die jungen Probanden
  - 2 für die alten Probanden

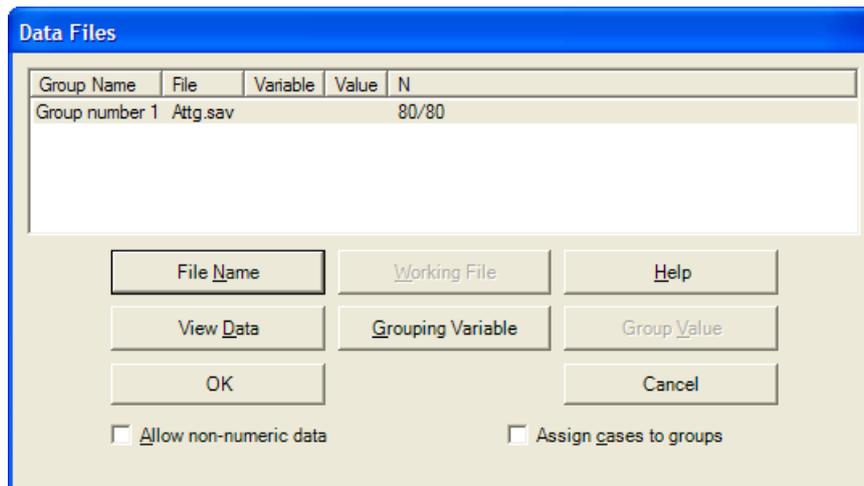
Weiterhin gelten folgende Regeln:

- Per Voreinstellung sind die Pfaddiagramme der einzelnen Gruppen identisch. Wenn es dabei bleibt, muss also nur *ein* Pfaddiagramm gezeichnet werden.
- Für unbenannte Parameter schätzt Amos in jeder Gruppe einen eigenen Wert. Die beiden ersten Regeln kann man also so zusammenfassen: Per Voreinstellung besitzen die Gruppen eine identische Modellstruktur, aber individuelle Parameterwerte.
- Parameter aus verschiedenen Gruppen können durch identische Benennung gleich gesetzt werden. Derartige Restriktionen erlauben sehr flexible Hypothesenprüfungen.

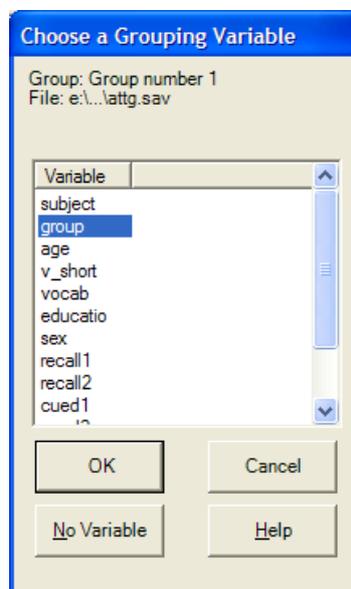
Öffnen Sie in einem neuen Amos-Projekt über **File > Data Files > File Name** die Datei **Attg.sav**:

---

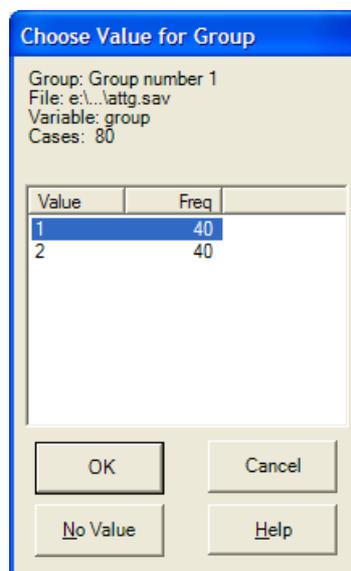
<sup>1</sup> Sie finden die Datei an der im Vorwort vereinbarten Stelle.



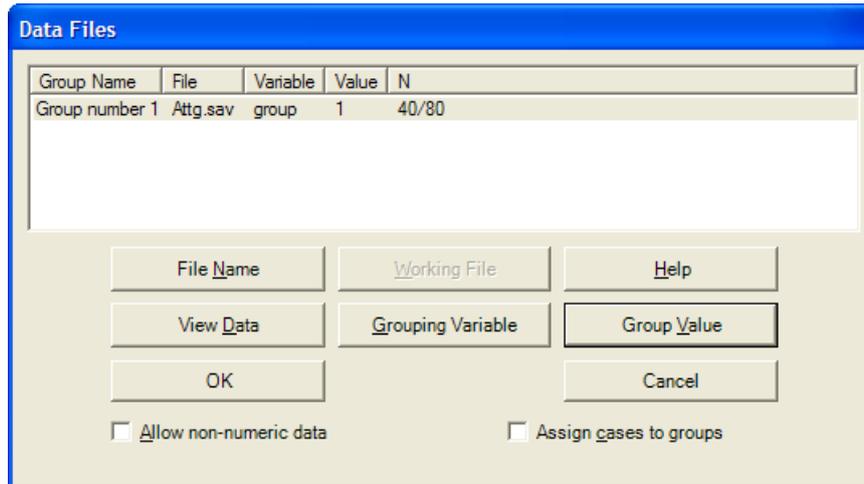
Wählen Sie nach einem Klick auf den Schalter **Grouping Variable** die Gruppierungsvariable **GROUP**:



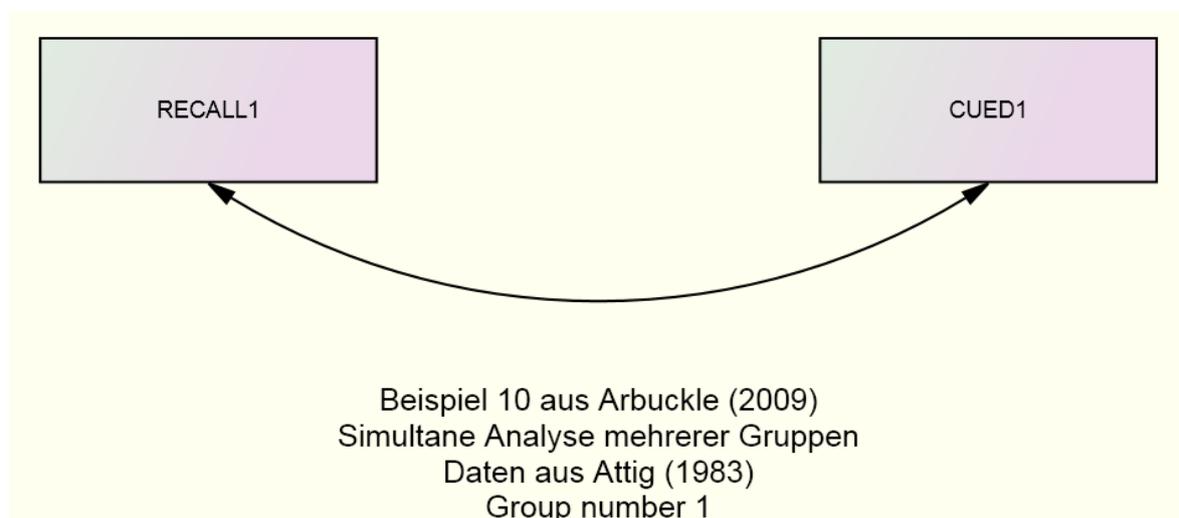
Klicken Sie im Dialog **Data Files** auf den Schalter **Group Value**, um die Gruppe der jungen Probanden (Wert 1) auszuwählen:



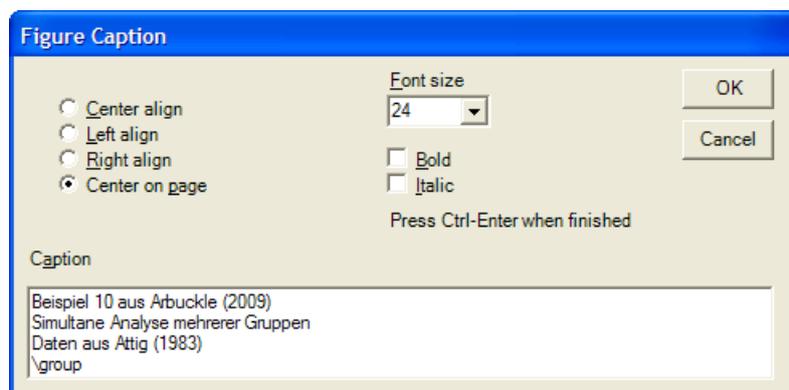
Quittieren Sie den Dialog **Data Files**:



Erstellen Sie das folgende Pfaddiagramm:



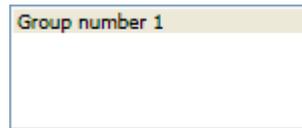
In der Diagrammbeschriftung (zu erstellen mit dem **Title**-Werkzeug) sorgt das Testmakro `\group` dafür, dass die aktuell dargestellte Gruppe stets erkennbar ist:



Nun wollen wir noch eine zweite Gruppe in die Analyse aufnehmen und außerdem für beide Gruppen eine informative Bezeichnung vereinbaren. Dazu starten wir den **Gruppenmanager** mit

**Analyze > Manage Groups**

Alternativ kann der Gruppenmanager auch per Doppelklick auf einen Eintrag im Gruppenauswahlbereich des Amos-Fensters gestartet werden:

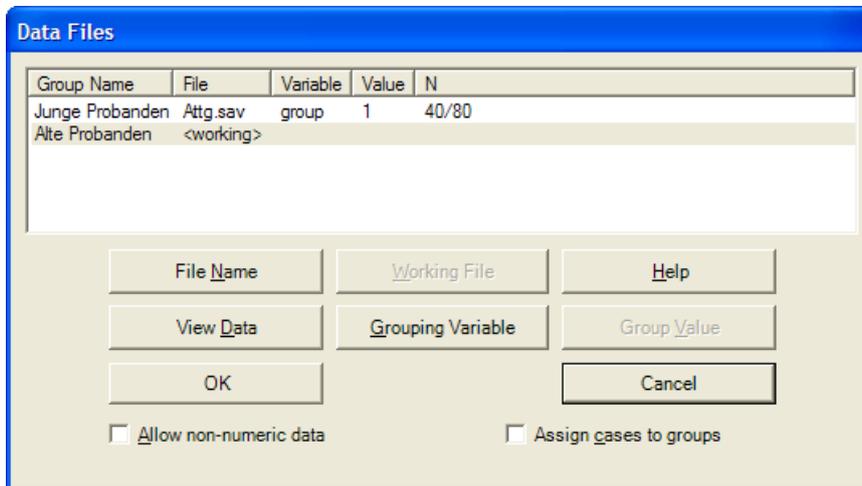


Zunächst ersetzen wir die automatisch vergebene Gruppenbezeichnung **Group number 1** durch **Junge Probanden**:

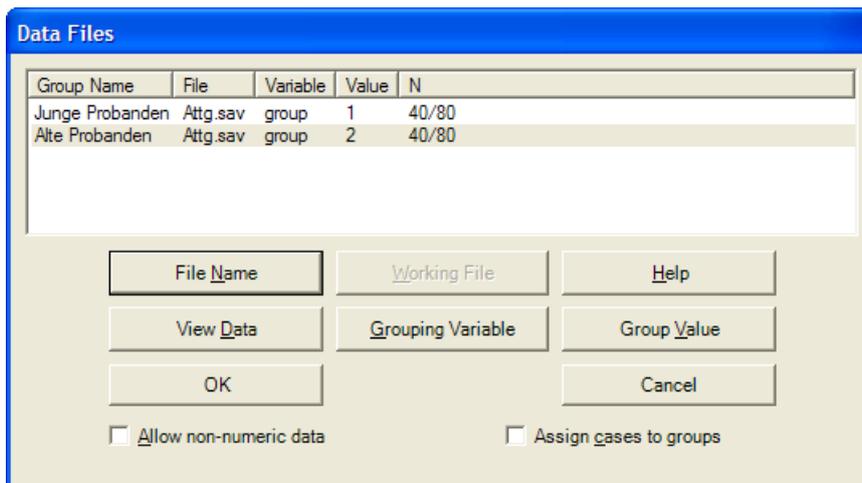


Dann legen wir mit **New** eine weitere Gruppe an und benennen diese gleich passend.

Schließlich muss über **File > Data Files** noch die Datenquelle zur zweiten, im Gruppenfenster aktuell markierten Gruppe vereinbart werden:



Nach Wahl der Datei **Attg.sav** entscheidet sich Amos spontan für die Fälle mit dem **GROUP**-Wert 2:



Nach unseren bisherigen Modellspezifikationen hat Amos in jeder Gruppe zwei Varianzen und eine Kovarianz zu schätzen. Es sind keine Restriktion und damit auch keine Freiheitsgrade für einen Modelltest vorhanden, wie die Textausgabe bestätigt:

**Computation of degrees of freedom (Default model)**

Number of distinct sample moments: 6  
 Number of distinct parameters to be estimated: 6  
 Degrees of freedom (6 - 6): 0

**Result (Default model)**

Minimum was achieved  
 Chi-square = ,000  
 Degrees of freedom = 0  
 Probability level cannot be computed

Die Schätzergebnisse für die jungen Probanden:

**Scalar Estimates (Junge Probanden - Default model)**

**Maximum Likelihood Estimates**

**Covariances: (Junge Probanden - Default model)**

		Estimate	S.E.	C.R.	P	Label
RECALL1	<--> CUED1	3,225	,944	3,416	***	

**Variances: (Junge Probanden - Default model)**

	Estimate	S.E.	C.R.	P	Label
RECALL1	5,788	1,311	4,416	***	
CUED1	4,210	,953	4,416	***	

Mit dem folgenden Werkzeug der Textausgabe



kann man zu den Schätzergebnissen für die alten Probanden wechseln:

**Scalar Estimates (Alte Probanden - Default model)**

**Maximum Likelihood Estimates**

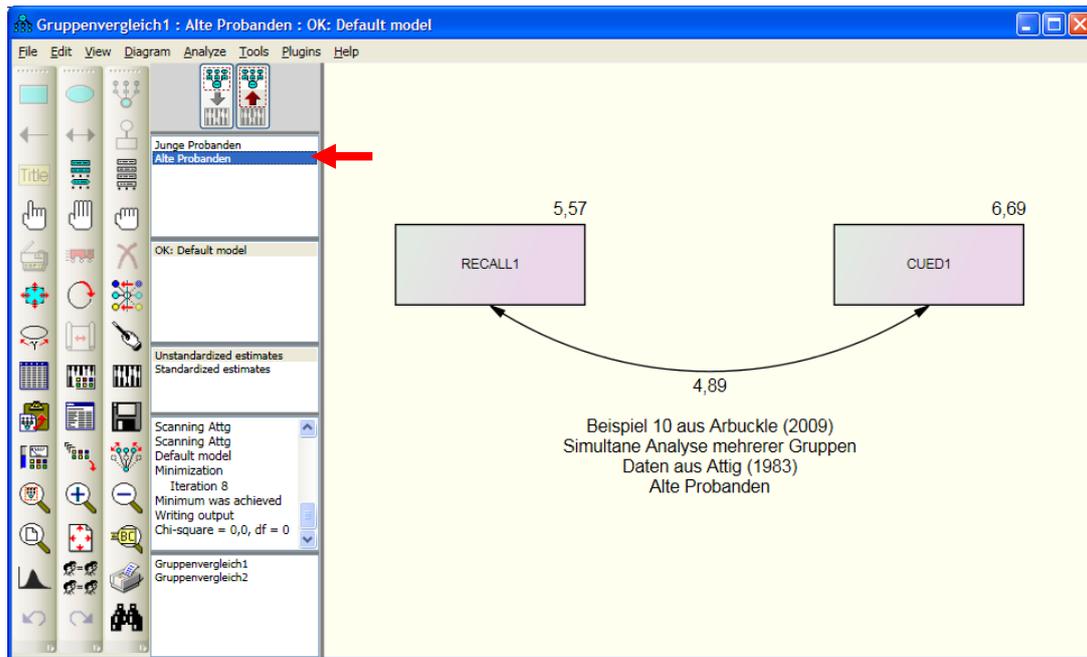
**Covariances: (Alte Probanden - Default model)**

		Estimate	S.E.	C.R.	PLabel
RECALL1	<--> CUED1	4,887	1,252	3,902	***

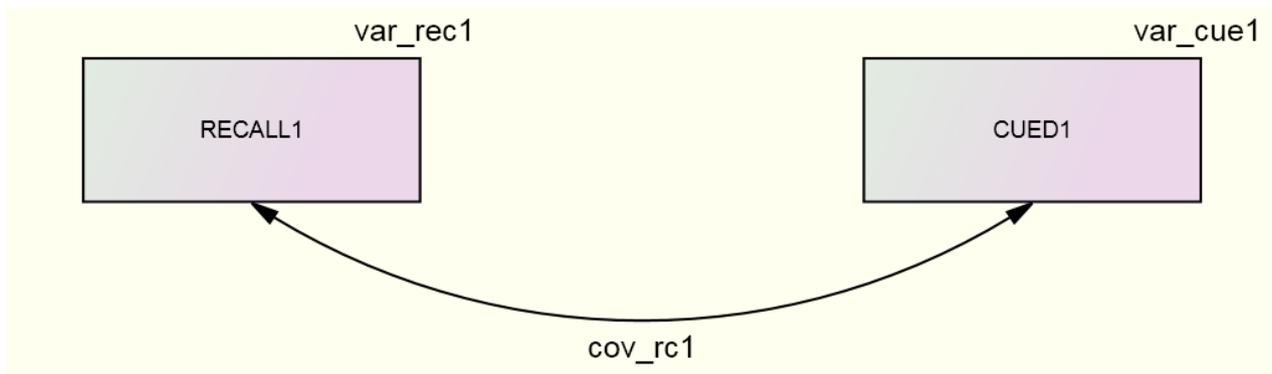
**Variances: (Alte Probanden - Default model)**

	Estimate	S.E.	C.R.	PLabel
RECALL1	5,569	1,261	4,416	***
CUED1	6,694	1,516	4,416	***

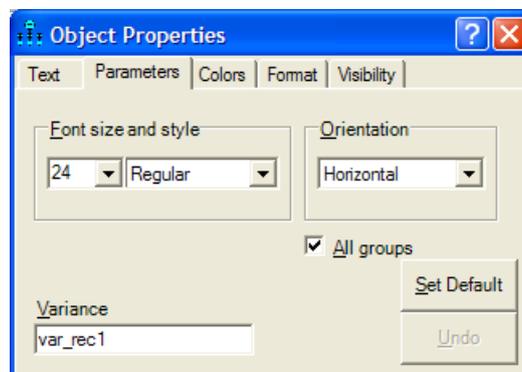
Das Ausgabe-Pfaddiagramm zeigt jeweils die Schätzergebnisse der im Gruppenauswahlbereich des Amos-Fensters markierten Gruppe, z.B.:



Nun soll die Hypothese identischer Kovarianzmatrizen in beiden Gruppen geprüft werden. Dazu geben wir korrespondierenden Parametern in beiden Gruppen identische Namen:



Auch bei offener **Object Properties** – Dialogbox sind Gruppenwechsel per Mausklick auf die Einträge im Gruppenauswahlbereich möglich. Soll für einen Parameter in *allen* Gruppen derselbe Wert oder Name festgelegt werden, kann man sich mit dem Kontrollkästchen **All groups** einigen Aufwand sparen:



Der  $\chi^2$ -Modelltest hat keine Einwände gegen die Gleichheitsrestriktionen, also gegen die Nullhypothese homogener Varianz-Kovarianzmatrizen:

**Notes for Model (Default model)**

**Computation of degrees of freedom (Default model)**

Number of distinct sample moments: 6  
 Number of distinct parameters to be estimated: 3  
 Degrees of freedom (6 - 3): 3

**Result (Default model)**

Minimum was achieved  
 Chi-square = 4,588  
 Degrees of freedom = 3  
 Probability level = ,205

Neben den Schätzern sind nun auch die Standardfehler in beiden Gruppen gleich und dabei kleiner als im unrestringierten Modell:

**Scalar Estimates (Junge Probanden - Default model)**

**Maximum Likelihood Estimates**

**Covariances: (Junge Probanden - Default model)**

		Estimate	S.E.	C.R.	PLabel
RECALL1	<--> CUED1	4,056	,780	5,202	***cov_rc1

**Variances: (Junge Probanden - Default model)**

	Estimate	S.E.	C.R.	PLabel
RECALL1	5,678	,909	6,245	***var_rec1
CUED1	5,452	,873	6,245	***var_cue1

Der eingangs erwähnte Präzisionsgewinn bei der simultanen Analyse mehrerer Gruppen stellt sich also genau dann ein, wenn durch Gleichheitsrestriktionen gemeinsame Parameterschätzungen möglich werden.

## 9 Bootstrapping

In der amerikanischen Variante der Münchhausen-Geschichte schafft es der Held, sich an den eigenen Stiefelriemen aus dem Sumpf zu ziehen, und dieses Verfahren liefert den Namen für eine prominente statistische Schätz- und Testmethodologie, die erstmals von Efron (1982) ausformuliert worden ist.

### 9.1 Prinzip und Nutzen

Man behandelt die Stichprobe als *Population*, ermittelt durch Ziehen mit Zurücklegen zahlreiche Sekundärstichproben (meist 500 – 1000) mit derselben Größe wie die Original- bzw. Primärstichprobe, wobei in der Regel etliche Fälle *mehrfach* in einer Sekundärstichprobe vertreten sind. Aus jeder Sekundärstichprobe wird mit den üblichen Methoden (z.B. ML) ein Schätzer für den interessierenden Parameter gewonnen, so dass man eine empirische Stichprobenkennwerteverteilung erhält. Diese ersetzt die theoretische Stichprobenkennwerteverteilung, die auf der Normalverteilungsannahme (vgl. Abschnitt 2.7.2) und Asymptotik basiert. Aus der empirischen Stichprobenkennwerteverteilung lassen sich Vertrauensintervalle und Testentscheidungen konstruieren, die nicht von der Normalverteilungsannahme abhängen.

Lange war die benötigte Rechenleistung ein Hindernis für die Weiterentwicklung und Anwendung der Bootstrap-Technologie, doch mittlerweile suchen die CPU-Hersteller nach relevanten Anwendungen für ihre Gigahertz- und Multicore-Boliden.

In folgenden Situationen ist die verteilungsvoraussetzungsfreie und sehr generell einsetzbare Ermittlung von Vertrauensintervallen zu Parameterschätzungen durch Bootstrap-Methoden von Interesse:

- Bei grob verletzter Normalverteilungsannahme sind die daraus abgeleiteten Standardfehler in der Regel zu klein und die zugehörigen Tests zu liberal. Mit Hilfe der Bootstrap-Technik gelangt man in vielen Fällen zu realistischeren Standardfehlern und vermeidet überhöhte  $\alpha$ -Fehlerraten.
- Für manche Statistiken (z.B. Determinationskoeffizienten, spezielle Modellgültigkeitskriterien, indirekte Effekte) lassen sich aus der Normalverteilungsannahme keine Standardfehler herleiten.

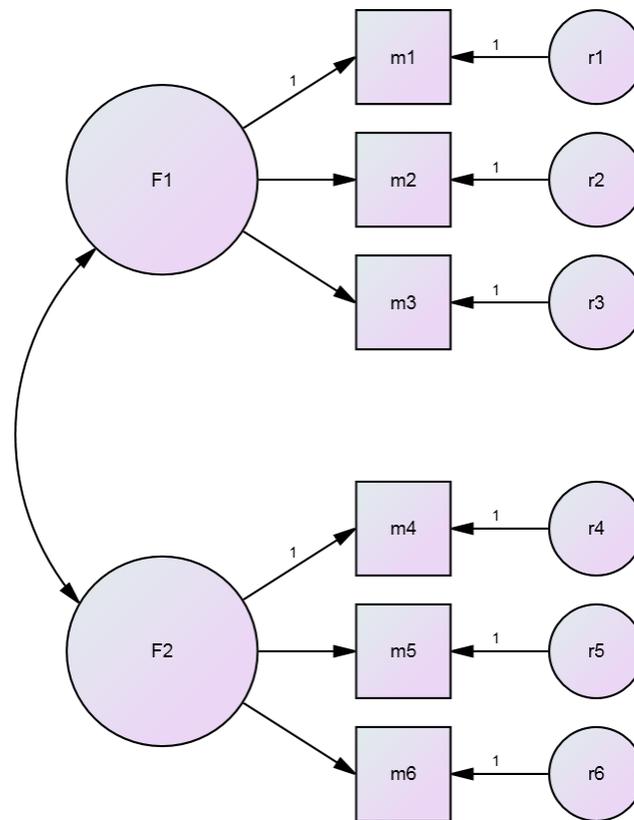
Damit eine Stichprobe als Pseudo-Population taugt, muss sie repräsentativ und hinreichend groß sein. Leider sind präzise Angaben zur erforderlichen Minimalgröße kaum zu finden. In einer von Kline (2005, S. 197) berichteten Studie wurden Bootstrap-Schätzer mit konventionellen ML-Schätzern bei einer konfirmatorischen Faktorenanalyse bei verletzter Normalitätsannahme verglichen. Während die Bootstrap-Methode bei Stichprobengrößen mit  $N \geq 200$  generell geringere Schätzfehler lieferte als die ML-Methode, waren die Ergebnisse bei  $N = 100$  weniger überzeugend: relativ große Standardfehler, viele indefinite (nicht analysierbare) Kovarianzmatrizen.

Bei sehr großen Stichproben ( $N > 1000$ ) kann man darauf setzen, dass sich Verletzungen der Normalverteilungsannahme nur geringfügig auf die letztlich relevanten Stichprobenkennwerteverteilungen auswirken (zentraler Grenzwertsatz). Hier sind Bootstrap-Methoden also wenig relevant als Remedium für eine verletzte Normalverteilungsannahme. Die Konstruktion von empirischen Prüfverteilungen für komplexe Statistiken sollte aber in dieser Situation besonders gut gelingen.

Die Standardabweichung der Parameterschätzungen aus den Sekundärstichproben liefert den **Bootstrap-Standardfehler**. Über das 2,5- und das 97,5 - Perzentil erhält man ein 95% - **Bootstrap-Vertrauensintervall**. Allerdings hat Efron (1987) eine verbesserte Variante vorgeschlagen, die Amos als *bias-corrected bootstrap confidence interval* anbietet.

### 9.2 Einsatz bei verletzter Normalitätsannahme

Als Beispiel betrachten wir eine künstliche Population mit der folgenden faktoriellen Struktur von manifesten Variablen **M1**, ..., **M6**:<sup>1</sup>

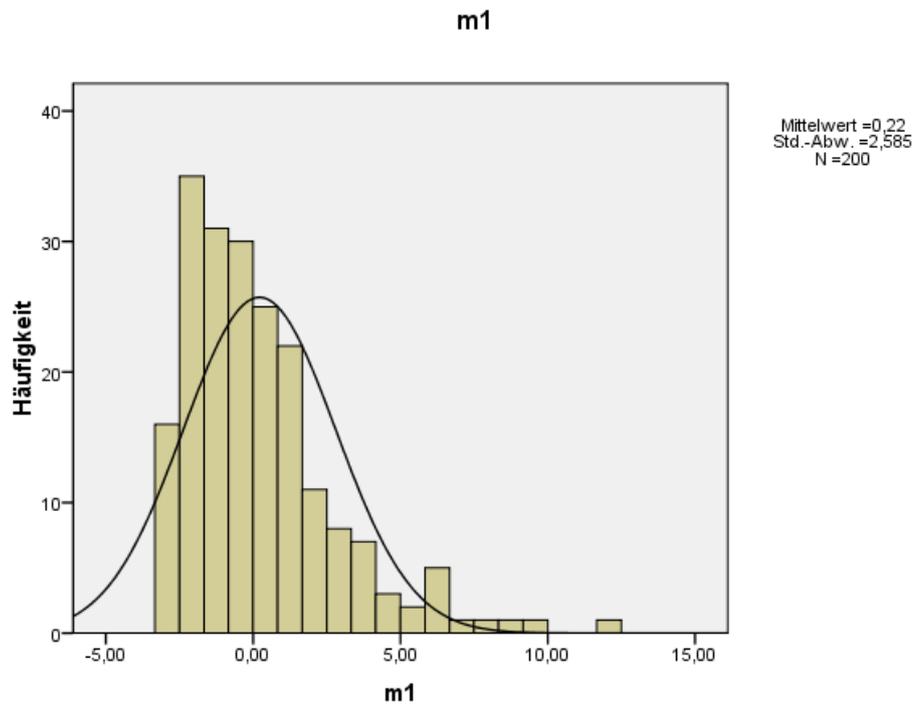


und den Ladungen:

Indikator	Faktor	
	F1	F2
M1	1	0
M2	0,5	0
M3	0	0
M4	0	1
M5	0	0,5
M6	0	0

Wie die beiden latenten Variablen sind auch alle manifesten Variablen (positiv) schief verteilt, z.B.:

<sup>1</sup> Sie finden das erzeugende SPSS-Programm **mcb.sps** und die Datendatei **mcb.sav** mit den resultierenden Zufallsdaten an der in Abschnitt **Fehler! Verweisquelle konnte nicht gefunden werden.** vereinbarten Stelle.



Die ML-Schätzergebnisse aus einer Stichprobe mit 200 Fällen zeigen trotz verletzter Normalitätsannahme einen guten Fit (vgl. Abschnitt 9.4):

**Result (Default model)**

Minimum was achieved  
Chi-square = 9,743  
Degrees of freedom = 8  
Probability level = ,284

Allerdings wird die Ladung von **M6** auf dem Faktor **F2** irrtümlich als signifikant beurteilt:

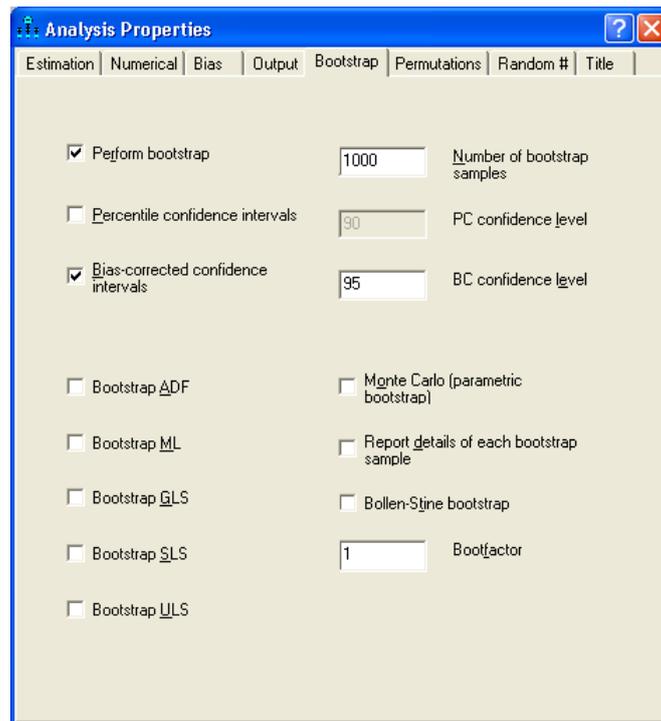
**Regression Weights: (Group number 1 - Default model)**

		Estimate	S.E.	C.R.	PLabel
m1 <---	F1	1,000			
m2 <---	F1	,419	,105	3,979	***
m3 <---	F1	-,066	,058	-1,137	,255
m4 <---	F2	1,000			
m5 <---	F2	,635	,150	4,240	***
m6 <---	F2	,220	,096	2,299	,021

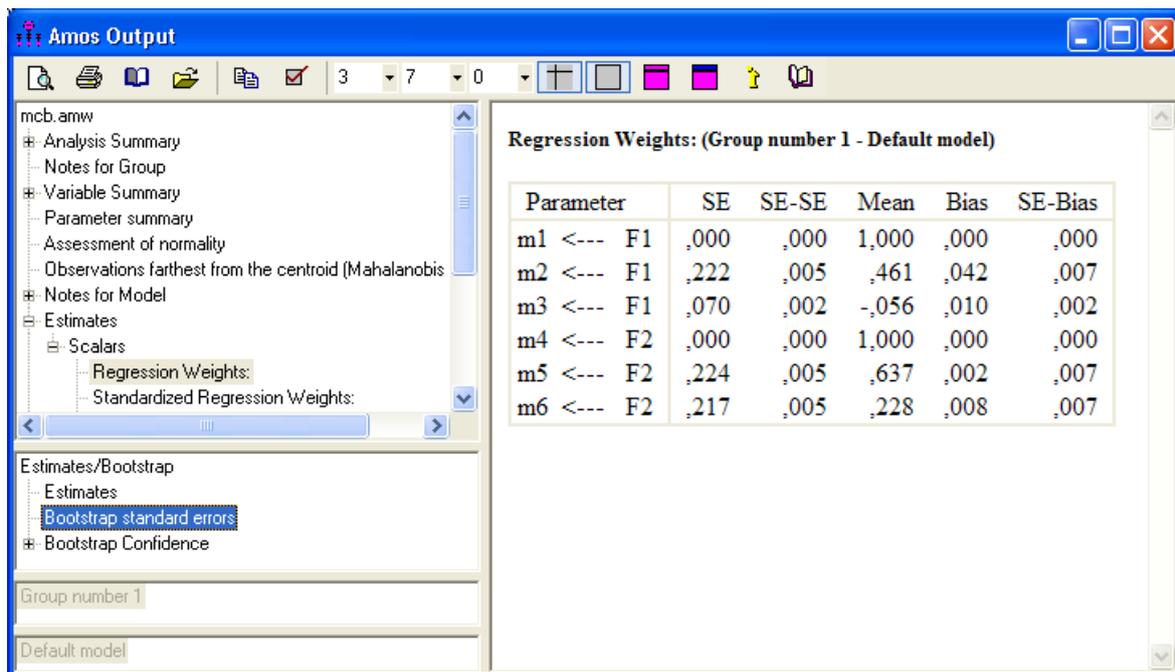
Wir fordern Amos nach

**View > Analysis Properties > Bootstrap**

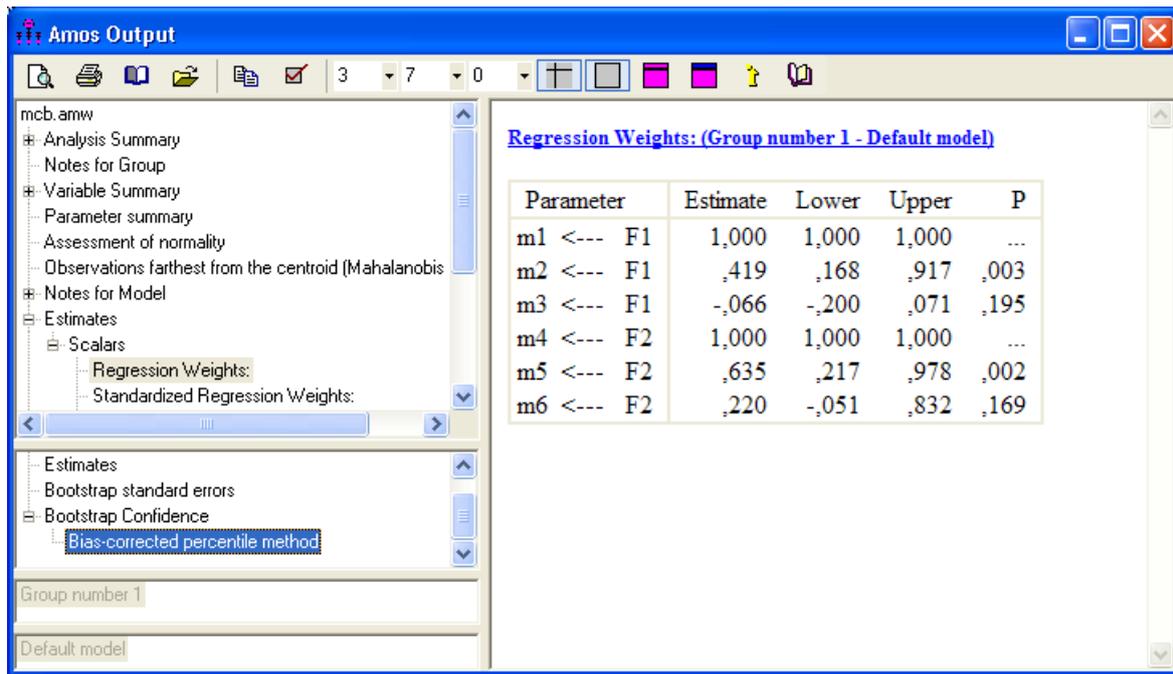
auf, 1000 Bootstrap-Sekundärstichproben zu ziehen und die korrigierten Vertrauensintervalle nach Efron (1987) zu berechnen:



In der Textausgabe fällt der Bootstrap-Standardfehler für die Ladung von **M6** auf dem Faktor **F2** erheblich größer aus (0,217) als der ML-Standardfehler (0,10):



Vom Bootstrap-Sinifkanztest wird die korrekte Nullhypothese zum betroffenen Parameter beibehalten:



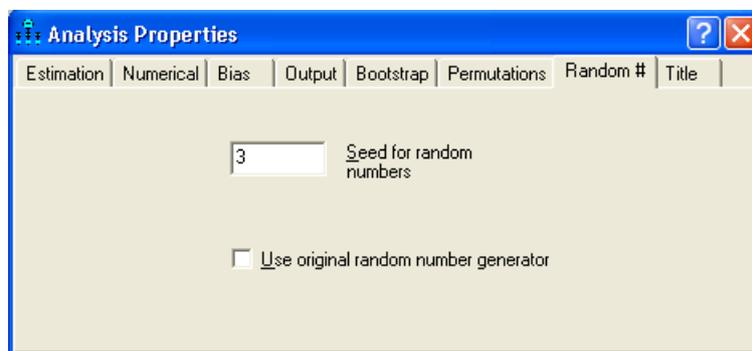
Im konstruierten Beispiel zeigen die von Amos berechneten Normalitätskriterien ein bedenkliches Bild:

**Assessment of normality (Group number 1)**

Variable	min	max	skew	c.r.	kurtosis	c.r.
m6	-1,000	13,172	3,632	20,970	19,658	56,748
m5	-1,978	5,550	1,152	6,653	1,173	3,387
m4	-2,969	9,202	,898	5,186	,544	1,571
m3	-1,000	10,016	2,948	17,022	12,817	36,999
m2	-1,897	4,905	1,097	6,334	,758	2,187
m1	-2,935	12,464	1,586	9,155	3,299	9,522
Multivariate					48,778	35,203

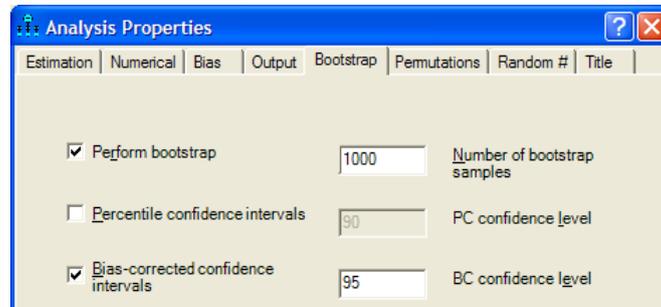
Die im Abschnitt 2.7.3 angegebenen Grenzwerte für univariate Verteilungen (Betrag 3 für die Schiefe und Betrag 7 für die Wölbung) sind bei mehreren Variablen überschritten, und auch der multivariate Wölbungstest lehnt seine Nullhypothese ab.

Bei allen Bootstrap-Methoden und anderen Techniken mit Zufallsbeteiligung hängen die Ergebnisse auch vom Startwert des Pseudozufallszahlengenerators ab, den man auf den Registerkarte **Random #** der Dialogbox **Analysis Properties** wählen kann:

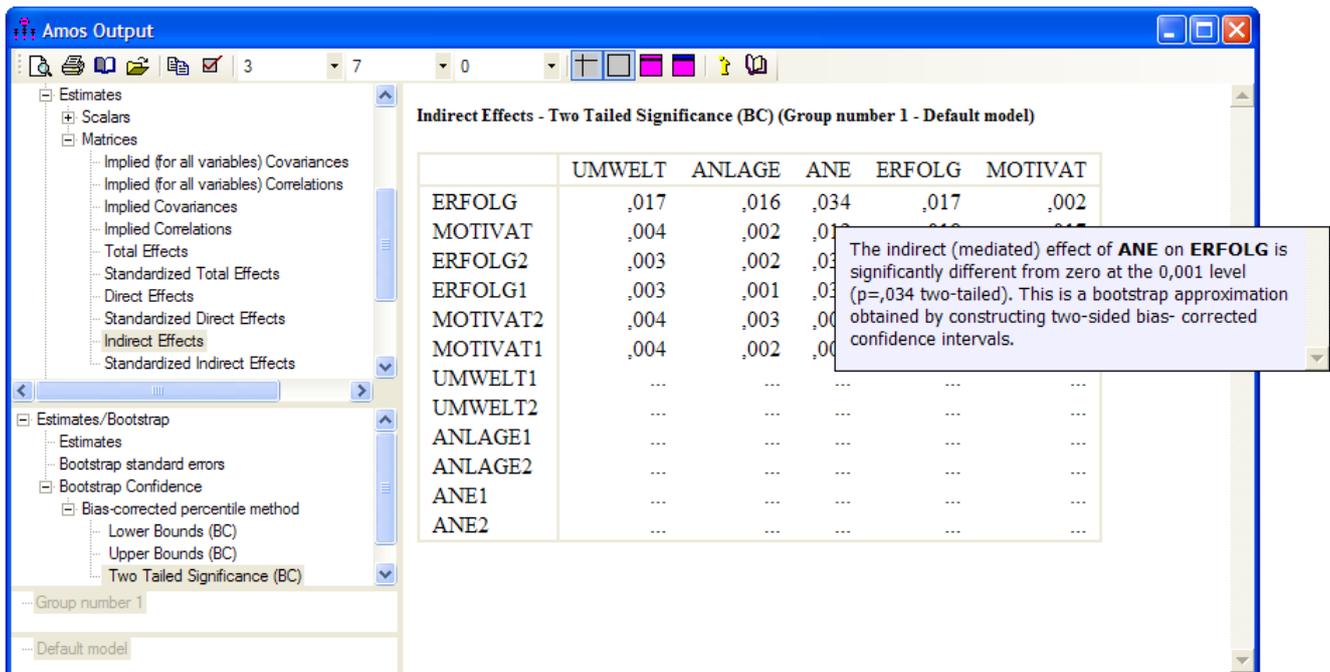


### 9.3 Bootstrap-Signifikanztests für indirekte Effekte

Wo die Maximum Likelihood - Theorie bei der Ableitung von Prüfverteilungen an ihre Grenzen stößt, kann die Bootstrap-Technik gelegentlich aushelfen. Zur Demonstration greifen wir das nonrekursive Modell mit latenten Variablen aus Abschnitt 6 auf und betrachten den indirekten Effekt der latenten Variablen **ANE** (Anspruchsniveau der Eltern) auf die Variable **ERFOLG** (vermittelt durch die Variable **MOTIVAT**). Ohne Bootstrap-Technik liefert Amos keine Signifikanzbeurteilung. Mit den folgenden, schon oben verwendeten Einstellungen im Dialog **Analysis Properties**



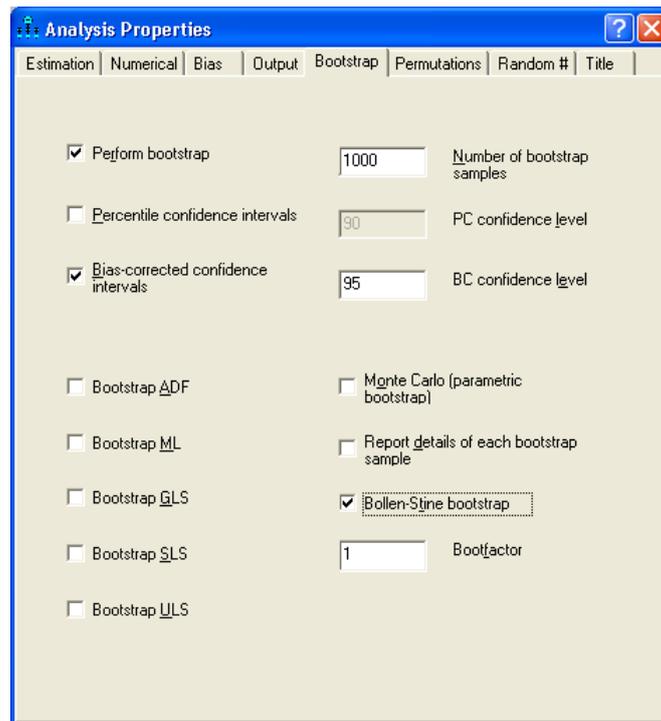
liefert Amos in der Textausgabe Bootstrap-Signifikanztests:



Der per Mausklick auf die per Bootstrap ermittelte Überschreitungswahrscheinlichkeit ( $p = 0,034$ ) angeforderte Infotext enthält klärende Erläuterungen und einen kleinen Fehler („0,001 level“).

### 9.4 Modellgültigkeitstest nach Bollen & Stine

Bei verletzter Normalverteilungsannahme sind nicht nur die Standardfehler zu den Parameterschätzungen betroffen, so dass Nullhypothesen zu den Parametern zu oft abgelehnt werden. Ein analoges Problem tritt auch beim Modellgültigkeitstest auf, so dass gültige Modell einer Wahrscheinlichkeit größer als 0,05 verworfen werden, weil die  $\chi^2$ -Prüfgröße tendenziell zu groß ausfällt (Kline 2005, S. 157). Als Korrektiv haben Bollen & Stine (1993) einen Bootstrap-basierten Modelltest entwickelt, den Amos in der Dialogbox **Analysis Properties** auf der **Bootstrap**-Registerkarte anbietet:



Unser Modell hatte schon den unkorrigierten Gültigkeitstest überstanden ( $p = 0,28$ ) und schneidet jetzt noch besser ab:

#### **Bollen-Stine Bootstrap (Default model)**

The model fit better in 588 bootstrap samples.

It fit about equally well in 0 bootstrap samples.

It fit worse or failed to fit in 412 bootstrap samples.

Testing the null hypothesis that the model is correct, Bollen-Stine bootstrap  $p = ,413$

Auch der korrigierte Test von Bollen & Stine (1993) tendiert dazu, in großen Stichproben Modelle „mit minimalen Fehlern“ zu verwerfen, so dass wir auf alternative Beurteilungsmethoden angewiesen bleiben.

Ist der Bollen-Stine Bootstrap-Modellgültigkeitstest gewählt, gibt Amos die oben beschriebenen Bootstrap-Konfidenzintervalle und -Tests *nicht* aus.

### **9.5 Parametrisches Bootstrapping**

Das bisher in Abschnitt 9 beschriebene Bootstrapping nennt man auch *nonparametrisch*, weil keine Normalverteilungsannahme im Spiel ist. Bei seiner Anwendung werden Rohdaten benötigt. Das so genannte *parametrische Bootstrapping* ist auch mit Momentenmatrizen (z.B. Kovarianzmatrizen) durchführbar. Um aus Momentenmatrizen eine Sekundärstichprobe zu erzeugen, muss allerdings ein Verteilungsmodell unterstellt werden. Man verwendet man die Momente der Primärstichprobe meist dazu, die Parameter einer Normalverteilung festzulegen, und zieht die Sekundärstichproben aus der so definierten Population. Damit eignet sich das parametrische Bootstrapping *nicht* bei fraglicher Normalitätsvoraussetzung. Es kann aber dazu verwendet werden, empirische Prüfverteilungen für komplexe Statistiken zu konstruieren. Die im Vergleich zum nonparametrischen Bootstrapping stärkeren Voraussetzungen sollten, sofern sie berechtigt sind, mit einer höheren Präzision (z.B. einer besseren Teststärke) belohnt werden.

Amos bietet das (fraglos als Monte Carlo - Technik zu beschreibende) parametrische Bootstrapping in der Dialogbox **Analysis Properties** auf der **Bootstrap**-Registerkarte an:

The screenshot shows the 'Analysis Properties' dialog box with the 'Bootstrap' tab selected. The dialog contains the following settings:

- Perform bootstrap: 1000 (Number of bootstrap samples)
- Percentile confidence intervals: 90 (PC confidence level)
- Bias-corrected confidence intervals: 95 (BC confidence level)
- Bootstrap ADF
- Monte Carlo (parametric bootstrap)
- Bootstrap ML
- Report details of each bootstrap sample
- Bootstrap GLS
- Bollen-Stine bootstrap
- Bootstrap SLS: 1 (Bootfactor)
- Bootstrap ULS

## 10 Literatur

- Arbuckle, J. L. (2009a). *Amos 18 User's Guide*. Chicago, IL: SPSS Inc.
- Arbuckle, J. L. (2009b). *Amos 18 Programming Reference Guide*. Amos Development Corporation.
- Attig, M. S. (1983). *The processing of spatial information by adults*. Presented at the annual meeting of The Gerontological Society, San Francisco.
- Baltes-Götz, B. (1994). *Einführung in die Analyse von Strukturgleichungsmodellen mit LISREL 7 und PRELIS unter SPSS*. Online-Dokumentation: <http://www.uni-trier.de/index.php?id=22734>
- Baltes-Götz, B. (2008a). *Behandlung fehlender Werte in SPSS und Amos*. Online-Dokumentation: <http://www.uni-trier.de/index.php?id=23239>
- Baltes-Götz, B. (2008b). *Interaktionseffekte in Strukturgleichungsmodellen*. Online-Dokumentation: <http://www.uni-trier.de/index.php?id=22631>
- Baltes-Götz, B. (2010). *Einführung in das Programmieren mit C# 3.0*. Online-Dokumentation: <http://www.uni-trier.de/index.php?id=22777>
- Bollen, K.A. (1989). *Structural equations with latent variables*. New York: Wiley.
- Bollen, K. A. & Stine, R. A. (1993). Bootstrapping goodness-of-fit measures in structural equation models. In: K. A. Bollen & J. S. Scott (S. 111-135), *Testing structural equation models*. Newbury Park, CA: Sage.
- Browne, M. & Cudeck, R. (1993). Alternative ways of assessing equation model fit. In: K. A. Bollen & J. S. Scott (S. 136-162), *Testing structural equation models*. Newbury Park, CA: Sage.
- Byrne, B. (2001). *Structural equation modeling with AMOS*. Mahwah: Lawrence Erlbaum Associates.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2<sup>nd</sup> ed). Hillsdale, NJ: Erlbaum.
- Efron, B. (1982). *The jackknife, the bootstrap and other resampling plans*. Philadelphia: SIAM-Monograph #38.
- Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American Statistical Association*, 82, 171-185.
- Efron, B. & Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. New York: Chapman & Hill.
- Jöreskog, K. G. & Sörbom, D. (1989). *LISREL 7. A Guide to the program and applications*. Chicago, IL: SPSS.
- Kerchoff, A.C. (1974). *Ambition and attainment*. Rose Monograph Series.
- Kenny, D.A. (1979). *Correlation and causality*. New York: Wiley.
- Kline, R. B. (2005). *Principles and Praxis of Structural Equation Modeling*. New York: Guilford Press.
- Lee, S.Y. (1985). Analysis of covariance and correlation structures. *Computational Statistics and Data Analysis*, 2, 279-295.
- Loehlin, J.C. (1987). *Latent variable models*. Hillsdale, NJ: Lawrence Erlbaum.
- Loehlin, J.C. (2004). *Latent variable models* (4<sup>th</sup> ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Mardia, K.V. (1974). Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies. *Sankhya*, Series B, 36, 115–128.
- Schumacker, R. E. & Lomax, R. G. (2004). *A Beginner's Guide to Structural Equation Modeling* (2<sup>nd</sup> ed.). Mahwah, NJ: Lawrence Erlbaum.
- Steyer, R. (1988). *Experiment, Regression und Kausalität: Die logische Struktur kausaler Regressionsmodelle*. Unveröff. Habilitationsschrift, Universität Trier.

Steyer, R. (1992). *Theorie kausaler Regressionsmodelle*. Stuttgart: Gustav Fischer.

Wishart, J. (1928). The generalised product moment distribution in samples from a normal multivariate population. *Biometrika*, 20A, 32–52.

## 11 Stichwortverzeichnis

.NET.....	22	Fehlervariablen	
ADF-Schätzer .....	89	automatisch erzeugen.....	47
AIC.....	80	FIML.....	73
Akaike Information Criterion .....	80	Fit-Funktion .....	69
AStructure		Fixierte exogene Variablen.....	27
Methode .....	24	Formative Indikatoren.....	82
Ausaben		Full Information Maximum Likelihood.....	73
optionale .....	21	Fundamentalthypothese .....	66
Bayes Information Criterion .....	80	Geschachtelte Modelle.....	74
BCC .....	80	Gleichheitsrestriktionen.....	30
BeginGroup		GLS-Schätzer.....	89
Methode .....	24	Gruppen-Manager.....	99
Bias .....	56	Gütekriterien .....	76
BIC.....	80	Heywood-Case.....	73
Bollen & Stine - Test .....	109	Hypothesentests	
Bootstrapping.....	104	Modellgültigkeit .....	26
parametrisches .....	110	zu den Parametern.....	19
Browne-Cudeck Criterion.....	80	zu Modellrestriktionen.....	30
C.R.....	49	Identifikation.....	58, 66
CAIC.....	80	Identifikationsproblem.....	45
CFA.....	57	identifiziert.....	33
Chiquadrat-Modelltest .....	26	Implizierte Momente .....	22
Close-Fit - Nullhypothese.....	78	Instrumentalvariablen .....	53
Consistent AIC.....	80	Interaktion.....	97
constraint interactions .....	82	Invarianz .....	52
CR.....	19, 71	Kausalität .....	52
Critical Ratio.....	19, 49, 71	Konfirmatorische Faktorenanalyse .....	57
Daten		Konsistente Schätzung.....	70
einlesen .....	16	Korrelation.....	63
Determinationskoeffizient .....	47	Kovarianz.....	63
Dezimalstellen		Kovarianzmatrix .....	56, 63
Anzahl ändern.....	18	Kurtosis.....	29
Diagrammbeschriftungen.....	36	LAST-Beispiel .....	56
Divergente Validität.....	88	Latente Variablen.....	81
Eigenschaften übertragen.....	43	Lineare Regressionsanalyse .....	28
Einlesen		Linienstärken .....	37
von Momentmatrizen.....	42	LISREL.....	56, 66
Ellipse .....	43	Lokalen Minimum .....	73
Empirisches Identifikationsproblem .....	66	Löschen.....	12
Ergebnisdatei .....	17	Mahalanobis-Distanz .....	29
Ergebnisse		MAR .....	73
berechnen lassen .....	17	Markieren.....	12
im Pfaddiagramm anzeigen .....	20	Maximum Likelihood - Prinzip .....	19
Erwartungstreue Schätzung .....	71	Mehrgruppenmodelle.....	97
Erwartungswert.....	62	Messfehler.....	81
F0 .....	77	Messmodell.....	82
Falsifizierbarkeit .....	67	Messmodell-Werkzeug.....	83
Fehlender Werte.....	73	MIMIC-Modell.....	82

Missing At Random.....	73	Seitenformat.....	11
ML-Schätzer .....	89	Sicherheitskopien.....	14
Modelle verwalten .....	75	Skaleninvarianz	
Modifikationsindikatoren .....	88	Modell.....	42, 70
Momente .....	62	Schätzmethode.....	42, 70, 72
Multivariate Normalität .....	28	Spezifikationssuche .....	92
Multivariate Normalverteilung .....	27	SPSS	
nonrekursiv .....	82	Matrix-Datendatei.....	42, 56
Nonrekursive Modelle .....	7	SRMR .....	79
Normal.amt .....	13	Stabilitätsindex .....	85
Normalitätstest.....	28, 108	Standard CFA .....	58
Normalverteilung		Standardfehler.....	19, 71
multivariate.....	27	Standardisierte Schätzungen.....	21
OLS-Schätzverfahren.....	48	Startwerte.....	74
optionale Ausaben .....	21	Stichprobengröße.....	27
Optionale Parameter .....	92	Stichprobenmomente .....	22
Ordinary Least Squares -Schätzung.....	48	Störvariablen.....	52
Parameter-Restriktionen		Strukturmodell .....	82
im Graphikmodus .....	30	Teststärke.....	78
in AmosEngine-Programmen .....	38	Textmakros .....	36, 38, 99
PathDiagrammer .....	26	TextOutput	
PCLOSE .....	78	Methode.....	24
Pfadanalyse .....	49	Touch Up - Werkzeug .....	44
Pfaddiagramm		t-Test.....	20
übertragen .....	15	t-Wert.....	19, 49
zeichnen .....	9	überidentifiziert.....	33
Pfeilspitzen .....	37	Überidentifizierte	
Plugins .....	22	Modelle.....	67
Poor-Fit - Nullhypothese .....	78	Parameter .....	67
Potentielle Störvariablen.....	52	ULI.....	46, 81
Projektdatei.....	14, 17	ULS.....	69
Pseudo-Isolation.....	52	Undo .....	13
Pseudozufallszahlengenerator.....	108	Unit Loading Identification.....	46, 81
Randomisierung.....	52	Unit Variance Identification .....	46, 82
Redo.....	13	Unweighted Least Squares.....	69
Reflektive Indikatoren .....	82	UVI.....	46, 82
Regression.....	41	Varianz.....	62
Rekursives Modell.....	6, 49	Verteilungsvoraussetzungen .....	27
Restriktionen.....	30	Vertikale Abstände ausgleichen.....	43
RMR .....	79	Vertrauensintervall .....	19
Root Mean Square Residual .....	79	Vorlagendatei.....	13
Rotationswerkzeug .....	83	Werkzeugpalette .....	8
Rückgängig machen.....	13	Wölbung .....	29
Saturiertes Modell.....	39, 48	Word .....	15
Schätzung		Zauberstab.....	10, 12, 44
anfordern.....	17	Zeichenfläche.....	11
Schiefe .....	29		