

Schätzung latenter nichtlinearer Effekte in Quer- und Längsschnittstudien

Dissertation

zur Erlangung des Doktorgrades
der Naturwissenschaften

vorgelegt beim Fachbereich 05
der Goethe-Universität
in Frankfurt am Main

von

Holger Brandt
aus Köthen (Anhalt)

Frankfurt am Main, 04. März 2013

Dekan: Prof. Dr. Rolf van Dick

Erstgutachter: Prof. Dr. Andreas Klein

Zweitgutachter: Prof. Dr. Karin Schermelleh-Engel

Drittgutachter: Prof. Dr. Jost Reinecke

Datum der Disputation: 16. Oktober 2013

Zusammenfassung

Strukturgleichungsmodelle (SEM) werden in den letzten Jahren vermehrt zur Aufdeckung von nichtlinearen Effekten wie Interaktionseffekten oder quadratischen Effekten in der empirischen Forschung verwendet. Daher kommt der Bereitstellung von effizienten und robusten Schätzverfahren für die Analyse von nichtlinearen SEM, die simultan multiple nichtlineare Effekte schätzen können, eine wichtige Bedeutung in der methodologischen Forschung zu. Bisher wurde jedoch nur ungenügend die Problematik untersucht, dass zwar die üblicherweise verwendeten Schätzverfahren aus der Klasse der Produktindikator-(PI)-Ansätze (z.B. der Unconstrained-Ansatz; Kelava & Brandt, 2009; Marsh, Wen & Hau, 2004) und der Klasse der verteilungsanalytischen Verfahren (z.B. LMS oder QML; Klein & Moosbrugger, 2000; Klein & Muthén, 2007) auf der Annahme einer multivariaten Normalverteilung für einen Großteil der im Modell enthaltenen Variablen beruhen, diese jedoch in der Empirie fast nie gegeben ist. Andere Ansätze, wie die momentbasierten Verfahren des 2SMM- oder des MM-Ansatzes (Wall & Amemiya, 2000, 2003; Mooijaart & Bentler, 2010), die die Normalitätsannahme deutlich abschwächen können, finden in der Literatur hingegen nur geringe Berücksichtigung. Im ersten Teil dieser Arbeit werden diese momentbasierten Verfahren zur Schätzung von multiplen nichtlinearen Effekten erweitert und hinsichtlich ihrer Schätzigenschaften bei nicht-normalverteilten Daten im Vergleich zu den PI- und den verteilungsanalytischen Ansätzen sowohl theoretisch als auch anhand einer umfangreichen Simulationsstudie untersucht (Brandt, Kelava & Klein, in press).

Zusammenfassend zeigt sich, dass LMS und QML bei normalverteilten Indikatoren die effizientesten Schätzungen liefern und in diesem Fall eingesetzt werden sollten. Bei nicht-normalverteilten Daten ist jedoch ein Parameterbias zu beobachten. Der erweiterte Unconstrained-Ansatz liefert zwar sowohl für normalverteilte als auch für nicht-normalverteilte Indikatoren erwartungstreue Parameterschätzungen, die Standardfehler werden jedoch stets unterschätzt (auch bei der Verwendung eines robusten Schätzers für die Standardfehler), was sich in einem erhöhten α -Fehler widerspiegelt. Der 2SMM-Ansatz liefert sehr gute Schätzergebnisse für normalverteilte und nicht-normalverteilte Indikatoren und kann insbesondere bei nicht-normalverteilten Indikatoren verwendet werden; bei normalverteilten Indikatoren ist das Verfahren etwas weniger effizient als LMS. Ein Nachteil des Verfahrens ist jedoch seine schwierige Erweiterung für andere als das hier untersuchte Querschnittsmodell (z.B. für Latente Wachstumskurvenmodelle). Der MM-Ansatz zeigt deutliche Schwächen in seinen Schätzungen, die sowohl bei einer simultanen Schätzung von mehr als einem nichtlinearen Effekt als auch bei nicht-normalverteilten Indikatoren auftreten, und stellt daher keine Alternative zu den anderen Schätzverfahren dar.

Um ein Verfahren bereitzustellen, das bei nicht-normalverteilten Daten zuverlässige und effiziente Schätzungen für nichtlineare Effekte liefern kann, wurde der NSEMM-Ansatz entwickelt (Kelava, Nagengast & Brandt, in press). Der NSEMM-Ansatz ist ein allgemeinerer Ansatz als der 2SMM-Ansatz und kann flexibler auch für andere Modelle genutzt werden, da er direkt in Mplus spezifizierbar ist. Der NSEMM-Ansatz verwendet zur Approximation der Verteilung der (latenten) Prädiktoren ein Mischverteilungsmodell, und stellt damit eine Erweiterung der SEMM (Structural Equation Mixture Models) dar. Im Gegensatz zu den SEMM, die nichtlineare Effekte semi-parametrisch modellieren, d.h. keine Funktion für die Nichtlinearität annehmen, verwendet der NSEMM-Ansatz eine parametrische Funktion für die nichtlinearen Effekte (z.B. einen quadratischen Effekt), wodurch eine Aussage über die Effektstärke des nichtlinearen Effekts möglich ist. Neben einer Darstellung des Konzepts der Mischverteilungsmodelle und des NSEMM-Ansatzes werden auch erste Ergebnisse zu den Schätzeigenschaften des NSEMM-Ansatzes im Vergleich zu LMS und dem erweiterten Unconstrained-Ansatz berichtet.

Im letzten Teil der Arbeit wird auf Interaktionseffekte in Längsschnittstudien eingegangen. Hierbei wird das Problem von heterogenen Entwicklungs- und Wachstumsprozessen aufgegriffen, das in der Empirie häufig auftritt, wenn die Streuung der Entwicklungsverläufe von den Ausgangsbedingungen abhängt. Wird diese Heterogenität nicht adäquat berücksichtigt, wie es der Fall bei latenten Wachstumskurvenmodellen (LGM) ist, sind Prognosen basierend auf den Ausgangsbedingungen einer Person nicht akkurat, da die Konfidenzintervalle für die Prognose auf einer fehlspezifizierten Varianz beruhen. Das heterogene Wachstumskurvenmodell (HGM; Klein & Muthén, 2006) erweitert das Standard-LGM um eine heterogene Varianzkomponente, die mit einem spezifischen Interaktionseffekt modelliert wird. Das HGM stellt eine Alternative zu den Growth Curve Mixture Modellen (GMM) dar, wenn keine Annahme über diskrete Klassen getroffen werden soll, sondern eine feinstufigere, kontinuierliche Modellierung der Heterogenität von Interesse ist. In dieser Arbeit wird für das HGM eine Implementierung in Mplus basierend auf dem LMS-Schätzer entwickelt, deren Anwendung und Interpretation an einem empirischen Datensatz aus der AIDS-Forschung demonstriert wird (Brandt, Klein & Gäde, submitted). Die Implementierung ermöglicht eine anwenderfreundliche Verwendung des Verfahrens und erlaubt die Berücksichtigung von Kovariaten zur Erklärung der Heterogenität in den Entwicklungsverläufen. Die Ergebnisse einer Simulationsstudie zeigen, dass das Verfahren auch bei moderater Verletzung der Verteilungsannahmen eine zuverlässige Parameterschätzung liefert.

Schlüsselwörter: Interaktionseffekte, Quadratische Effekte, Nichtnormalität, Nichtlineare Strukturgleichungsmodelle, Mischverteilungsmodelle, Latente Wachstumskurvenmodelle, Heterogene Wachstumsprozesse

Inhaltsverzeichnis

1	Einleitung	5
1.1	Differentielle Effekte in Quer- und Längsschnittstudien	5
1.2	Nichtlineare Effekte in Strukturgleichungsmodellen	7
1.2.1	Modellformulierung	8
1.2.2	Anforderungen an Ansätze für die Schätzung von nichtlinearen SEM	9
1.2.3	Mischverteilungsmodelle	11
1.2.4	Nichtlineare Effekte in latenten Wachstumskurvenmodellen . . .	11
1.3	Aufbau & Schriften	12
2	Ein Vergleich aktueller Verfahren zur Analyse latenter nichtlinearer Effekte	13
2.1	Nichtlineare SEM: Ein theoretischer Vergleich	13
2.1.1	Der (erweiterte) Unconstrained-Ansatz	16
2.1.2	Verteilungsanalytische Ansätze: LMS und QML	19
2.1.3	Der 2SMM-Ansatz	20
2.1.4	Der MM-Ansatz	23
2.2	Bisherige Simulationsstudien	25
2.3	Simulationsstudie	27
2.3.1	Design der Simulationsstudie	27
2.3.2	Zusammenfassung der wichtigsten Ergebnisse	28
2.4	Zusammenfassung und Schlussfolgerung	28
3	Ein Schätzverfahren für nichtlineare Effekte bei nicht-normalverteilten Prädiktoren	30
3.1	Idee der Mischverteilungsmodelle	31
3.2	Structural Equation Mixture Modeling	33
3.3	Interpretation der Klassen und Variablen in SEMM	35
3.4	Nonlinear Structural Equation Mixture Modeling	35
3.5	Diskussion des NSEMM-Ansatzes	37

4 Implementierung und praktische Anwendung eines Heterogenen Wachstumskurvenmodells	39
4.1 Latente Wachstumskurvenmodelle	39
4.2 Interaktionseffekte in LGM	41
4.3 Heterogene Wachstumsprozesse	42
4.4 Das HGM nach Klein und Muthén (2006)	43
4.5 Implementierung des HGM in Mplus	44
4.6 Empirisches Beispiel	46
4.7 Zusammenfassung und Diskussion zum HGM	49
5 Diskussion	52
5.1 Zusammenfassung	52
5.2 Bedeutung und Einordnung der Arbeit	53
5.3 Grenzen der Arbeit	54
5.4 Ausblick	56
Literaturverzeichnis	58
A Erweiterung des 2SMM-Ansatzes	69
A.1 Schätzung der Parameter	69
A.2 Schätzung der Standardfehler	72
B Konfidenzintervalle für die Prognose von \hat{y}_t	76
C Schrift (A)	78
D Schrift (B)	79
E Schrift (C)	80
F Selbstständigkeitserklärung	81

Kapitel 1

Einleitung

1.1 Differentielle Effekte in Quer- und Längsschnittstudien

Die zuverlässige Schätzung von Interaktionseffekten und quadratischen Effekten ist in der empirischen Forschung im Bereich der Differentiellen, Arbeits- und Klinischen Psychologie, den Verhaltenswissenschaften, als auch in weiten Bereichen sozialwissenschaftlicher Studien von besonderer Bedeutung (z.B. Ajzen, 1987; Cronbach & Snow, 1977; Karasek, 1979; Lusch & Brown, 1996; Snyder & Tanke, 1976). Dies soll anhand der folgenden Beispiele illustriert werden.

Ein *Interaktions-* oder auch *Moderatoreffekt* modelliert einen differentiellen Effekt: In Abhängigkeit einer Drittvariablen, der Moderatorvariablen, verändert sich der Zusammenhang zwischen einem Prädiktor und einer abhängigen Variablen. Eine der bekanntesten Theorien in der Psychologie, in der Interaktionseffekte postuliert werden, ist die *Theorie des geplanten Verhaltens* (Ajzen & Madden, 1986), in der angenommen wird, dass die Intention ein Verhalten zu zeigen einen umso höheren Einfluss auf das tatsächlich gezeigte Verhalten hat, je stärker bei einer Person die wahrgenommene Verhaltenskontrolle ausgeprägt ist. Nur dann, wenn die Person glaubt, dass sie eine Handlung auch wirklich durchführen kann, wird sie die Intention in eine Handlung umsetzen. Wie Ajzen und Fishbein (2005) feststellen, muss zudem davon ausgegangen werden, dass auch der Zusammenhang zwischen der Einstellung gegenüber einem Verhalten und dem tatsächlich gezeigten Verhalten von einer Reihe von potentiellen personen- und kontextspezifischen Moderatorvariablen abhängt, wie der Selbstwahrnehmung, dem Self-Monitoring oder dem Zeitdruck einer Handlung (Gangestad & Snyder, 1985; Jamieson & Zanna, 1989; Kline, 1987).

Auch in der Motivationsforschung wird bei der *Erwartungs-mal-Wert-Theorie* ein Interaktionseffekt formuliert (Atkinson, 1964). In ihr wird postuliert, dass die Motivati-

on ein Verhalten zu zeigen von der Erfolgserwartung und dem Wert eines erfolgreichen Verhaltens abhängt. Hierbei wird angenommen, dass es sich um einen kompensatorischen Effekt handelt, dass also sowohl eine alleinige hohe Erfolgsaussicht als auch der alleinige hohe Wert einer Handlung zu einer hohen Motivation führen. Dieser Effekt wird als Produkt beider Variablen (Erwartung \times Wert) zur Vorhersage der Motivation operationalisiert. In einem konkreten Beispiel zur Prognose der Berufswahl von Jugendlichen leiten Eccles et al. (1983) sowie Wigfield und Eccles (2000) daraus folgende Hypothese ab: Für Jugendliche kann das akademische Selbstkonzept in Naturwissenschaften (Erfolgserwartung) die Karrierewünsche im Bereich der Naturwissenschaften (Motivation) prognostizieren und zwar umso stärker, als dass ein Jugendlicher Freude an Naturwissenschaften besitzt (Wert der Handlung). Für Jugendliche mit einer großen Freude ist der Zusammenhang zwischen dem Selbstkonzept und den Karrierewünschen sehr hoch; für Jugendliche mit nur wenig Freude an den Naturwissenschaften hingegen gibt es nur einen geringen Zusammenhang zwischen den beiden Variablen, da auch naturwissenschaftlich begabte Jugendliche ihren Beruf unabhängig von dieser Eigenschaft wählen.

Ein *quadratischer Effekt* stellt einen kurvilinearen Zusammenhang zwischen einem Prädiktor und einer abhängigen Variablen dar, wie er im Yerkes-Dodson-Gesetz (Yerkes & Dodson, 1908) oder der darauf aufbauenden Theorie zum hedonischen Tonus (Eysenck, 1994) postuliert wird. In ihr wird ein umgekehrt U-förmiger Zusammenhang zwischen dem Grad der Stimulation und dem hedonischen Tonus („Wohlbefinden“) angenommen, der ein Maximum – und damit ein hohes Wohlbefinden – bei einer mittleren Stimulusintensität aufweist. Sowohl eine sehr geringe (im Extremfall eine Deprivation) als auch eine sehr starke Stimulusintensität (bis hin zu einem Schmerzempfinden) werden hingegen als unangenehm empfunden.

Auch im Bereich der Entwicklungs- und pädagogischen Psychologie stellt die Modellierung von Wachstums- und Entwicklungsprozessen, bei denen Interaktionseffekte berücksichtigt werden, einen wichtigen Forschungsgegenstand dar. Beispielsweise wird die Entwicklung kognitiver Fähigkeiten in der Kindheit zumeist als ein differentieller Wachstumsprozess modelliert, dessen Zuwachs je nach anfänglichen Eigenschaften der Kinder oder ihrer Eltern (z.B. Wissensstand, sozioökonomischer Status, Bildungsgrad) stärker oder schwächer ausgeprägt ist (empirische Beispiele s. z.B. in Goodnight, Bates, Dawn Staples, Pettit & Dodge, 2007; Larsen, Overbeek, Vermulst, Granic & Engels, 2010; Parrila, Aunola, Leskinen, Nurmi & Kirby, 2005). In einem konkreten Beispiel von Sénéchal und LeFevre (2002) weisen Kinder bei gleichen anfänglichen Lesefähigkeiten einen umso größeren Zuwachs in ihrer Lesefähigkeit im Laufe der ersten Schuljahre auf, je stärker die häusliche Literatheit ausgeprägt ist.

Parrila et al. (2005) zeigen in einer Studie zur Entwicklung der Leseleistung, dass die Entwicklungsverläufe der untersuchten Schüler eine deutliche Heterogenität aufweisen. Kinder mit hohen anfänglichen Leseleistungen entwickelten sich sehr konsistent, was sich in parallelen Entwicklungsverläufen zeigte. Hingegen entwickelten sich Kinder mit anfänglich niedriger Leseleistung sehr verschieden. Einige dieser Kinder erwarben im Laufe der Grundschule eine sehr gute Leseleistung, andere jedoch nicht. Eine Konsequenz dieser Heterogenität der Entwicklungsverläufe spiegelt sich in einer unterschiedlichen Prognosegenauigkeit der langfristigen Leseleistungen wider. Eine Prognose kann umso genauer erfolgen, je ausgeprägter die anfänglichen Leseleistungen waren. Die Berücksichtigung einer solchen Heterogenität erfolgt ebenfalls mittels einer spezifischen Modellierung von Interaktionseffekten (Genauerer folgt in Kapitel 4).

1.2 Nichtlineare Effekte in Strukturgleichungsmodellen

Strukturgleichungsmodelle („Structural Equation Models“, SEM), die Interaktionseffekte oder quadratische Effekte berücksichtigen, werden in der letzten Dekade in den empirischen Wissenschaften immer häufiger verwendet (z.B. in Beierlein, Werner, Preiser & Wermuth, 2011; Brody et al., 2006; Goodnight et al., 2007; Hoyer & Klein, 2000; Klein & Stoolmiller, 2003; Larsen et al., 2010; Milan, Snow & Belay, 2009; Wei, Malinckrodt, Russell & Abraham, 2004). Dies ist durch zwei Argumente zu begründen. Zum einen ermöglicht der Einsatz von SEM die Berücksichtigung von Messfehlern, die explizit im Messmodell aufgenommen werden, sodass die Effekte zwischen latenten, theoretisch messfehlerfreien Konstrukten geschätzt werden können. Dies ist insbesondere bei nichtlinearen Effekten relevant, da die Reliabilität der für die Modellierung von Interaktions- und quadratischen Effekten eingesetzten Produktterme geringer ist als die der Prädiktoren, aus denen sie gebildet werden (Bohrnstedt & Marwell, 1978; Dimitruk, Schermelleh-Engel, Kelava & Moosbrugger, 2007; MacCallum & Mar, 1995). Werden Messfehler – wie in einer Regressionsanalyse bei den Prädiktoren – nicht berücksichtigt, kann es zu einer Verzerrung der Schätzungen der Effekte kommen. Zum anderen führt die Berücksichtigung von mehr als einem Indikator in SEM zu einer höheren Inhaltsvalidität, da mehrere Aspekte simultan berücksichtigt werden können. Ebenso ist eine flexible Modellierung von latenten Variablen möglich, was eine adäquatere (validere) Darstellung komplexerer latenter Konstrukte ermöglicht, beispielsweise durch die Verwendung von hierarchischen Faktorenstrukturen oder der Berücksichtigung von Methodeneffekten.

Für die Analyse von SEM mit Interaktions- und quadratischen Effekten, für die häufig der Begriff „nichtlineare SEM“ verwendet wird¹, existieren eine Reihe verschiedener Schätzverfahren (Jöreskog & Yang, 1996; Kelava et al., 2011; Klein & Moosbrugger, 2000; Klein & Muthén, 2007; Little, Bovaird & Widaman, 2006; Marsh et al., 2004; Marsh, Wen & Hau, 2006; Mooijaart & Bentler, 2010; Moulder & Algina, 2002; Schumacker & Marcoulides, 1998; Wall & Amemiya, 2000, 2003). Diese Verfahren unterscheiden sich in der Konzeptualisierung der Modelle und damit auch in ihren Schätzeigenschaften. Im folgenden Abschnitt soll jedoch zuerst die Modellformulierung für nichtlineare SEM eingeführt werden, auf der alle Ansätze gleichermaßen beruhen.

1.2.1 Modellformulierung

Nichtlineare SEM stellen eine Erweiterung von linearen SEM dar und berücksichtigen – neben den linearen Einflüssen der exogenen Variablen (Prädiktoren) auf die endogenen (abhängigen) Variablen – auch nichtlineare Effekte, die durch die Hinzunahme von Produkten der Prädiktoren modelliert werden. Allgemein wird ein nichtlineares Modell für eine einzelne endogene Variable η dargestellt durch

$$\eta = \alpha + \mathbf{\Gamma}\boldsymbol{\xi} + \boldsymbol{\xi}'\boldsymbol{\Omega}\boldsymbol{\xi} + \zeta, \quad (1.1)$$

wobei α das latente Interzept von η ist, $\boldsymbol{\xi} = (\xi_1, \dots, \xi_v)'$ der Vektor der v Prädiktoren, $\mathbf{\Gamma}$ die $(1 \times v)$ Koeffizientenmatrix der linearen Effekte von $\boldsymbol{\xi}$ auf η , $\boldsymbol{\Omega}$ eine obere $(v \times v)$ Dreiecksmatrix mit den Koeffizienten der nichtlinearen Effekte (vgl. Klein & Moosbrugger, 2000)², und ζ das latente Residuum von η . Die exogenen Variablen $\boldsymbol{\xi}$ besitzen einen $(v \times 1)$ Mittelwertsvektor $\boldsymbol{\mu}^\xi$ und eine $(v \times v)$ Kovarianzmatrix $\boldsymbol{\Phi}$; das Residuum ζ ist unkorreliert mit $\boldsymbol{\xi}$ und besitzt einen Mittelwert von null und eine Varianz Ψ .

Für die latenten Konstrukte der endogenen Variable η und der Prädiktoren $\boldsymbol{\xi}$ wird ein Messmodell der Form

$$\begin{aligned} \mathbf{y} &= \boldsymbol{\tau}^y + \boldsymbol{\Lambda}^y \eta + \boldsymbol{\epsilon} \\ \mathbf{x} &= \boldsymbol{\tau}^x + \boldsymbol{\Lambda}^x \boldsymbol{\xi} + \boldsymbol{\delta} \end{aligned} \quad (1.2)$$

spezifiziert, wobei $\mathbf{y} = (y_1, \dots, y_p)'$ und $\mathbf{x} = (x_1, \dots, x_q)'$ die Vektoren der Indikatorvariablen darstellen, $\boldsymbol{\tau}^y, \boldsymbol{\tau}^x$ die $(p \times 1)$ bzw. $(q \times 1)$ Vektoren der latenten Interzepte, $\boldsymbol{\Lambda}^y, \boldsymbol{\Lambda}^x$ die $(p \times 1)$ bzw. $(q \times v)$ Faktorladungsmatrizen und $\boldsymbol{\epsilon}, \boldsymbol{\delta}$ die $(p \times 1)$ bzw. $(q \times 1)$ Vektoren der Residuen mit Mittelwertsvektoren $\mathbf{0}_p$ bzw. $\mathbf{0}_q$ und Kovarianzmatrizen $\boldsymbol{\Theta}^\epsilon$

¹Genau genommen sind die nichtlinearen SEM lineare Modelle, die Polynome der Prädiktoren enthalten. Der Begriff „nichtlinear“ hat sich jedoch für diese Modelle in der Literatur durchgesetzt.

² $\boldsymbol{\Omega}$ enthält in den oberen Nebendiagonalelementen die Interaktionseffekte, in der Diagonale die quadratischen Effekte und in den unteren Nebendiagonalelementen Nullen.

bzw. Θ^δ . Für ein einfaches Modell mit zwei Prädiktoren ξ_1 und ξ_2 , das neben den linearen Effekte ausschließlich einen Interaktionseffekt enthält, vereinfacht sich Gleichung 1.1 zu

$$\begin{aligned}\eta &= \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_{12}\xi_1\xi_2 + \zeta \\ &= \alpha + (\gamma_1 + \omega_{12}\xi_2)\xi_1 + \gamma_2\xi_2 + \zeta,\end{aligned}\tag{1.3}$$

wobei γ_1 und γ_2 die linearen Einflussgewichte sind und ω_{12} der Interaktionseffekt. Wie man erkennen kann, verändert sich der Zusammenhang zwischen ξ_1 und η auf den einzelnen Stufen von ξ_2 linear, wenn $\omega_{12} \neq 0$ ist. Die um zwei quadratische Effekte ω_{11} und ω_{22} erweiterte Modellgleichung ist gegeben als

$$\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_{12}\xi_1\xi_2 + \omega_{11}\xi_1^2 + \omega_{22}\xi_2^2 + \zeta.\tag{1.4}$$

Sie wird als „vollständiges nichtlineares Modell“ für zwei Prädiktoren bezeichnet.

1.2.2 Anforderungen an Ansätze für die Schätzung von nichtlinearen SEM

Für die Schätzung von nichtlinearen SEM stehen verschiedene Ansätze zur Verfügung, die sich in ihren Konzeptualisierungen und Annahmen unterscheiden, wodurch sie auch verschiedene Eigenschaften aufweisen. Sie lassen sich größtenteils in drei Klassen einteilen: Produkt-Indikator-Ansätze (Jöreskog & Yang, 1996; Kelava & Brandt, 2009; Kenny & Judd, 1984; Little et al., 2006; Marsh et al., 2004, 2006; Wall & Amemiya, 2001), verteilungsanalytische Ansätze (Klein & Moosbrugger, 2000; Klein & Muthén, 2007) und momentbasierte Ansätze (Mooijaart & Bentler, 2010; Wall & Amemiya, 2000, 2003). Gemeinsam ist ihnen jedoch das Ziel einer zuverlässigen Schätzung der nichtlinearen Effekte. Die Güte der Schätzungen der einzelnen Verfahren hängt insbesondere von zwei Aspekten ab, die in der empirischen Forschung bei der Analyse von nichtlinearen Effekten berücksichtigt werden müssen.

Der erste Aspekt betrifft die Testung multipler nichtlinearer Effekte. Während inhaltliche Hypothesen sich zumeist auf Interaktionseffekte beziehen, hat sich die Notwendigkeit einer Berücksichtigung quadratischer Effekte in den letzten Jahren durch methodische Argumente verdeutlicht. Ganzach (1997) sowie MacCallum und Mar (1995) zeigten, dass die Analyse eines Interaktionseffekts ohne eine gleichzeitige Schätzung von quadratischen Effekten zu Scheineffekten führen kann. Dies ist immer dann der Fall, wenn in der Population ein quadratischer Effekt vorliegt, die Prädiktoren miteinander korrelieren, aber im Analysemodell ausschließlich ein Interaktionseffekt getestet wird. Durch die aus der Korrelation der Prädiktoren resultierende Korrelation der Pro-

duktterme (z.B. zwischen $\xi_1\xi_2$ und ξ_1^2) kann es vorkommen, dass ein Interaktionseffekt fälschlicherweise signifikant wird, obwohl tatsächlich ein quadratischer Effekt vorliegt. Klein, Schermelleh-Engel, Moosbrugger und Kelava (2009) konnten zudem zeigen, dass quadratische Terme ausreichend sind, um das Aufdecken von Scheininteraktionseffekten zu vermeiden, und nicht noch weitere Potenzen der Prädiktoren (beispielsweise ξ_1^3) notwendig sind. Als Empfehlung kann daher aus konzeptioneller Sicht gelten, dass auch zur Aufdeckung eines Interaktionseffekts ein vollständiges nichtlineares Modell getestet werden sollte. Dadurch steigt die Anforderung an die Verfahren zur Schätzung nichtlinearer Effekte, nicht nur einzelne Interaktionseffekte sondern simultan multiple nichtlineare Effekte zuverlässig zu schätzen.

Der zweite Aspekt bei der Testung latenter nichtlinearer Effekte betrifft die Voraussetzungen für die einzelnen Verfahren, insbesondere die Verteilungsvoraussetzungen. Fast alle Verfahren, die im Rahmen der nichtlinearen SEM entwickelt wurden, verwenden Annahmen über die Normalverteilung eines Teils oder aller im Modell enthaltenen Variablen. Eine (multivariate) Normalverteilung der Daten ist jedoch in der Empirie nur sehr selten gegeben (Micceri, 1989), weshalb untersucht werden muss, inwieweit die Schätzungen noch zuverlässig sind, wenn die den Daten unterliegenden Annahmen verletzt sind. Bisher ist dieser Aspekt jedoch nur in begrenztem Umfang untersucht worden und Aussagen können nur eingeschränkt für Interaktionsmodelle für einige der Verfahren getroffen werden (Cham, West, Ma & Aiken, 2012; Marsh et al., 2004; Wall & Amemiya, 2001).

In Kapitel 2 soll daher die Frage beantwortet werden, **welches der derzeit aktuellen Verfahren zur Schätzung multipler latenter nichtlinearer Effekte die zuverlässigsten Ergebnisse bei einer Verletzung der Normalverteilungsannahme liefern kann.** Zu dieser Fragestellung sollen die fünf als relevant eingestuften Verfahren – LMS (Klein, 2000; Klein & Moosbrugger, 2000), QML (Klein & Muthén, 2007), der erweiterte Unconstrained-Ansatz (Kelava & Brandt, 2009; Marsh et al., 2004), der 2SMM-Ansatz (Wall & Amemiya, 2000, 2003) und der MM-Ansatz (Mooijart & Bentler, 2010) – hinsichtlich ihrer Eigenschaften eingehend analysiert werden. Zudem soll eine Simulationsstudie durchgeführt werden, in der die Schätzeigenschaften der Verfahren bei Verletzung der Normalverteilungsannahme untersucht werden. Basierend auf diesen Informationen werden Empfehlungen für den Einsatz der Verfahren gegeben. Einen Schwerpunkt stellen der 2SMM- und der MM-Ansatz dar, da für diese bisher nur wenige Untersuchungen existieren. Hierfür werden beide Ansätze für die Analyse von vollständigen nichtlinearen Modellen erweitert.

1.2.3 Mischverteilungsmodelle

In Kapitel 3 wird das Problem aufgegriffen, dass Daten in empirischen Studien häufig nicht-normalverteilt vorliegen, was in den meisten Schätzverfahren der nichtlinearen SEM nicht berücksichtigt wird. **Es wird ein Verfahren vorgestellt, das eine zuverlässige Schätzung der nichtlinearen Effekte auch dann erlaubt, wenn die (latenten) Prädiktoren nicht normalverteilt sind.** Das NSEMM-Verfahren („Non-linear Structural Equation Mixture Model“; Kelava et al., in press) verwendet als Approximation der Verteilung der Prädiktoren ein Mischverteilungsmodell, das eine sehr flexible Modellierung von beliebigen Verteilungen ermöglicht. Dadurch kann die sonst häufig notwendige Annahme der Normalverteilung für die Schätzung der nichtlinearen Effekte (wie in LMS oder dem erweiterten Unconstrained-Ansatz) abgeschwächt werden und die Nichtnormalität wird explizit im Modell berücksichtigt.

Bisher wurden Mischverteilungsmodelle im Rahmen von Strukturgleichungsmodellen, sogenannte Structural Equation Mixture Models (SEMM), vor allem eingesetzt, um eine semi-parametrische Schätzung von nichtlinearen Effekten zu erhalten (Arminger & Stein, 1997; Arminger, Stein & Wittenberg, 1999; Bauer & Curran, 2004; Bauer, 2005; Dolan & van der Maas, 1998; Jedidi, Jagpal & DeSarbo, 1997a, 1997b; McLachlan & Peel, 2000; Pek, Losardo & Bauer, 2011). Der Vorteil der semi-parametrischen Schätzung ist, dass keine Annahme über den Funktionstyp (z.B. quadratischer Effekt) getroffen werden muss; dies ist gerade dann ein Vorteil, wenn man nicht weiß, welcher Art die Nichtlinearität ist oder man keine Annahme dazu treffen möchte. Der Nachteil ist jedoch, dass keine Größenaussage zu den nichtlinearen Effekte möglich ist (Kelava et al., in press). Das NSEMM-Verfahren, das in Schrift (B) entwickelt wurde, erweitert die SEMM um eine parametrische Modellierung der nichtlinearen Effekte. Die Hintergründe zu diesem Verfahren und eine erste Beurteilung seiner Schätzeigenschaften werden in Kapitel 3 vorgestellt.

1.2.4 Nichtlineare Effekte in latenten Wachstumskurvenmodellen

Eine häufig verwendete Methode für die Analyse von Längsschnittdaten stellen die latenten Wachstumskurvenmodelle dar, die individuelle Entwicklungsverläufe im Rahmen der SEM modellieren (Latent Growth Curve Models, LGM; Bollen & Curran, 2006; Duncan, Duncan & Strycker, 2006; Reinecke, 2012). Auch bei diesen Modellen stellt, wie in Abschnitt 1.1 deutlich wurde, die Formulierung und Testung von differentiellen Effekten anhand von Interaktionshypothesen eine wichtige Erweiterung dar, um empirische Sachverhalte adäquat zu modellieren. Werden Interaktionseffekte zwischen (beobachteten) Kovariaten und dem individuellen Startpunkt zur Vorhersage des individuellen

Entwicklungsverlaufs verwendet, so erfolgt eine Modellierung der Interaktion ebenfalls über Produktterme dieser Variablen. Die Operationalisierung kann daher zum Teil analog zu den Verfahren für nichtlineare SEM erfolgen. Obwohl nur eine begrenzte Literatur zur Schätzung von nichtlinearen Effekten in LGM existiert (Bollen & Curran, 2006; Khoo, 2001; Muthén, 2001; Preacher, Curran & Bauer, 2006), ist zu erwarten, dass die Eigenschaften der meisten Ansätze für nichtlineare SEM (wie LMS) übertragbar sind und diese Ansätze direkt für ein LGM verwendet werden können.

Eine besondere Herausforderung implizieren jedoch heterogene Wachstumsprozesse, wie sie am Beispiel der Entwicklung der Leseleistungen in Abschnitt 1.1 dargestellt wurde. Heterogene Wachstumsprozesse treten dann auf, wenn sich die Varianz der Entwicklungskurven für Personen mit verschiedenen Ausgangsbedingungen ändert. In diesen Fällen ist der Einsatz von Nicht-Standardverfahren notwendig wie das heterogene Wachstumskurvenmodell von Klein und Muthén (2006). **In Kapitel 4 wird eine auf dem Schätzprinzip von LMS basierende Mplus-Implementation für dieses Modell entwickelt und ihre Anwendung demonstriert.**

1.3 Aufbau & Schriften

In Kapitel 2 werden aktuelle Ansätze zur Schätzung nichtlinearer Effekte in SEM vorgestellt und ihre Vor- und Nachteile besprochen. In Kapitel 3 wird dann auf den NSEMM-Ansatz eingegangen, der spezifisch zum Einsatz bei nicht-normalverteilten Daten entwickelt wurde. Das heterogene Wachstumskurvenmodell und seine Implementierung in Mplus werden anschließend in Kapitel 4 vorgestellt. Im letzten Kapitel werden die bearbeiteten Themen gemeinsam in einen breiteren Rahmen gestellt und die Grenzen der Arbeit kritisch diskutiert. Für wesentliche theoretische Darstellungen in den Kapiteln wurde auf folgende Schriften zurückgegriffen, die in Anhang C - E zu finden sind:

- (A) **Brandt, H.**, Kelava, A. & Klein, A. G. (*in Druck*). A Simulation study comparing recent approaches for the estimation of nonlinear effects in SEM under the condition of non-normality. *Structural Equation Modeling*.
- (B) Kelava, A., Nagengast, B., & **Brandt, H.** (*in Druck*). A nonlinear structural equation mixture modeling approach for non-normally distributed latent predictor variables. *Structural Equation Modeling*.
- (C) **Brandt, H.**, Klein, A. G. & Gäde, J. (*in Revision*). Implementation and practical application of a heterogeneous growth curve model. *Multivariate Behavioral Research*.

Kapitel 2

Ein Vergleich aktueller Verfahren zur Analyse latenter nichtlinearer Effekte

In diesem Kapitel werden aktuelle Ansätze zur Schätzung latenter nichtlinearer Effekte im Rahmen der SEM vorgestellt und zueinander in Bezug gesetzt. Neben einer theoretischen Auseinandersetzung mit den Konzepten und den Verteilungsannahmen, auf denen die einzelnen Verfahren beruhen, werden die Ergebnisse einer Simulationsstudie berichtet, die die Schätzeigenschaften der Verfahren bei Verletzung dieser Verteilungsannahmen untersucht.

Anmerkung: Dieses Kapitel basiert zu großen Teilen auf folgendem Artikel, der in Anhang C enthalten ist und in dem auch die detaillierten Ergebnisse der Simulationsstudie zu finden sind:

Brandt, H., Kelava, A. & Klein, A. G. (*im Druck*). A Simulation study comparing recent approaches for the estimation of nonlinear effects in SEM under the condition of non-normality. *Structural Equation Modeling*.

2.1 Nichtlineare SEM: Ein theoretischer Vergleich

Eine erwartungstreue und effiziente Schätzung von nichtlinearen SEM ist seit etwa zwei Jahrzehnten das Ziel neuer Verfahren (u.a. Bollen, 1995; Jaccard & Wan, 1995; Jöreskog & Yang, 1996; Kelava, Moosbrugger, Dimitruk & Schermelleh-Engel, 2008; Kelava & Brandt, 2009; Kelava et al., 2011, in press; Klein & Moosbrugger, 2000; Klein & Muthén, 2007; Lee, Song & Poon, 2004; Little et al., 2006; Marsh et al., 2004, 2006; Mooijaart & Bentler, 2010; Ping, 1995, 1996; Wall & Amemiya, 2000, 2001, 2003). Die Gemeinsamkeit dieser Verfahren ist ein parametrisches Modell für nichtlineare Effekte,

d.h. es wird z.B. für den Interaktionseffekt angenommen, dass sich der Zusammenhang zwischen dem Prädiktor und dem Kriterium über die Stufen des Moderators linear verändert (s. Gleichung 1.3)¹. Die Verfahren unterscheiden sich hauptsächlich in der Operationalisierung bzw. Konzeptualisierung der latenten Produktterme (beispielsweise $\xi_1\xi_2$ in Gleichung 1.4) und in den Annahmen, auf denen die Schätzungen der jeweiligen Verfahren beruhen. Diese Annahmen beinhalten je nach Verfahren die (multivariate) Normalverteilung aller oder zumindest einiger Variablen im Modell. Die Verfahren unterscheiden sich in ihren Schätzeigenschaften, was auf die Stärke dieser Annahmen und der Konsequenzen bei ihrer Verletzung zurückzuführen ist. Grundsätzlich man kann die Stärke einer Annahme für ein Verfahren anhand dreier Aspekte einordnen (vgl. Kane, 1992):

1. Die Annahme sollte erfahrungswissenschaftlich plausibel sein.
2. Die Gültigkeit einer Annahme sollte auf der Grundlage der Daten entscheidbar oder testbar sein.
3. Die Schätzungen und inferenzstatistischen Schlüsse des Verfahrens sollten gegenüber einer Verletzung der Annahme robust sein.

Eine Annahme ist *plausibel*, wenn sie in empirischen Kontexten zumeist gegeben ist, d.h. sie sollte in den meisten Situationen nicht verletzt sein. Diese Entscheidung kann auf der Basis von Vorwissen zu dem Gebiet getroffen werden, in dem das Verfahren eingesetzt werden soll. Weiterhin sollten die Annahmen so formuliert sein, dass ein Forscher, der das Verfahren verwendet, mit einiger Sicherheit heuristisch entscheiden kann, ob konkret für seine Studie und seine Daten die dem Verfahren unterliegenden Annahmen erfüllt sein könnten.²

Für eine Annahme ist es weiterhin stets vorteilhaft, wenn sie *testbar* ist. In einem solchen Fall muss nicht ausschließlich anhand von Plausibilitätsüberlegungen argumentiert werden, sondern auf der Basis eines Tests kann eine empirisch belastbarere Aussage getroffen werden.³ Eine solche Testung kann sowohl inferenzstatistische als auch deskrip-

¹Auf semi-parametrische Modelle wird in Kapitel 3 kurz eingegangen.

²Ein Beispiel für eine unplausible Annahme stellt die sequentielle Ignorabilität im Rahmen der Mediatoranalyse dar (Gilbert & Hudgens, 2008; Holland, 1988; Frangakis & Rubin, 2002; Ten Have et al., 2007). Diese Annahme beinhaltet die (bedingte) Unabhängigkeit zwischen dem Mediator und der Kriteriumsvariable, die immer verletzt ist, sobald eine (unbeobachtete) Drittvariable beide Variablen beeinflusst. Dies ist zum einen fast immer der Fall und zum anderen kann ein Anwender nur extrem schwer über ihre Verletzung in einem konkreten Beispiel eine Entscheidung treffen.

³Auch hier soll als Beispiel die sequentielle Ignorabilität dienen: Sie ist nicht testbar, denn hierfür müssten alle Drittvariablen bekannt sein. Dies kann aber forschungslogisch fast nie der Fall sein, weshalb eine Testung faktisch unmöglich ist. In dem einzigen (theoretischen) Fall, in dem alle Drittvariablen bekannt sind, müsste man zugleich den gesamten dem Mediatormodell unterliegenden Kausalmechanismus kennen. In diesem Fall benötigt man aber trivialerweise keine Mediatoranalyse mehr, da alle Fragen zum Mediator a priori durch den schon bekannten Kausalmechanismus beantwortet wären.

tivstatistische Verfahren berücksichtigen. Für die Testung einer univariaten Normalverteilung stehen z.B. als inferenzstatistischer Test der Kolmogorov-Smirnoff-Test zur Verfügung und deskriptivstatistisch die empirische Schiefe und Kurtosis einer Variablen sowie graphische Verfahren (Histogramm, ...). Für die multivariate Normalverteilung existieren ebenfalls eine Reihe verschiedener Tests⁴ (Cox & Small, 1978; Mardia, 1970; Shapiro & Wilk, 1965; Small, 1980; s. auch Rencher, 2002).

Die *Robustheit* eines Verfahrens gibt Auskunft über seine Sensibilität gegenüber der Verletzung seiner Annahmen. Ein Hinweis auf Robustheit findet sich, wenn ein Verfahren auch dann erwartungstreue (unverzerrte) Schätzungen liefert, wenn die Annahmen (in einem empirisch relevanten Ausmaß) verletzt sind (z.B. Fleishman, 1978). Während manche Verfahren auch zuverlässige Ergebnisse liefern, obwohl ihre Annahmen verletzt sind – wie beispielsweise die Normalverteilungsvoraussetzung beim *t*-Test (Boneau, 1960) – existieren andere Verfahren, bei denen auch schon kleine Verletzungen der Annahmen zu verzerrten Schätzungen führen.⁵

Im Folgenden sollen nun die Konzepte zu relevanten Verfahren für die Schätzung latenter nichtlinearer Effekte vorgestellt und ihre Verteilungsannahmen unter den drei Aspekten Plausibilität, Testbarkeit und Robustheit diskutiert werden. Konkret werden folgende Ansätze vorgestellt:

- Der (erweiterte) Unconstrained-Ansatz (Kelava & Brandt, 2009; Marsh et al., 2004) aus der Klasse der PI-Ansätze;
- LMS (Klein & Moosbrugger, 2000) und QML (Klein & Muthén, 2007) aus der Klasse der verteilungsanalytischen Ansätze;
- Der 2SMM-Ansatz (Wall & Amemiya, 2000, 2003) und der MM-Ansatz (Mooijaart & Bentler, 2010) aus der Klasse der momentbasierten Ansätze.

Der Schwerpunkt der konzeptuellen Darstellung wird hierbei auf die momentbasierten Verfahren gelegt. Ein theoretischer Vorteil dieser Ansätze besteht darin, dass Verteilungsannahmen nur für einen Teil der im Modell enthaltenen Variablen getroffen werden

⁴Die Vielzahl an Tests deutet auf die Schwierigkeit einer adäquaten Testung der multivariaten Normalverteilung hin, die ein eigenes Problemfeld darstellt und auf die daher in dieser Arbeit nicht näher eingegangen werden kann (s. z.B. in Joanes & Gill, 1998; Mecklin & Mundfrom, 2004). Neben den ca. 50 Tests der multivariaten Normalverteilung, die in einem Review von Mecklin und Mundfrom (2004) beschrieben werden, existieren auch eine Reihe verschiedener Operationalisierungen der Schiefe und Kurtosis, da die üblichen momentbasierten Maße beliebig groß und daher nur schwer interpretierbar werden können (Groeneveld & Meeden, 1984; Oja, 1981).

⁵Wiederum liefert die Mediatoranalyse mit der sequentiellen Ignorabilität hier ein besonders deutliches Beispiel: Der Bias sowohl des indirekten als auch des direkten Effekts ist proportional zu der Korrelation, die durch eine Nichtberücksichtigung einer Drittvariablen entsteht (Holland, 1988). Ein Bias ist von der Größenordnung her nicht abschätzbar, da man die Drittvariablen nicht kennt, und produziert auch schon bei harmlosen Verletzungen der Annahme eine α -Inflation von über 90% (Klein & Brandt, in Arbeit).

müssen, während für die anderen Verfahrensklassen eine Verteilungsannahme für alle im Modell enthaltenen Variablen notwendig ist. Während zudem eine theoretische Auseinandersetzung mit den PI-Ansätzen und den verteilungsanalytischen Ansätzen existiert (z.B. Kelava & Brandt, 2009; Kelava et al., 2008, 2011; Klein & Muthén, 2007; Moosbrugger, Schermelleh-Engel, Kelava & Klein, 2009; Moulder & Algina, 2002; Wall & Amemiya, 2001), wurden die momentbasierten Verfahren bisher fast ausschließlich in der jeweiligen Originalliteratur dargestellt (Wall & Amemiya, 2000, 2003 bzw. Mooijaart & Bentler, 2010). Dies soll hier erweitert und präzisiert werden.

2.1.1 Der (erweiterte) Unconstrained-Ansatz

Der erweiterte Unconstrained-Ansatz gehört zu der Klasse der PI-Ansätze. Für alle Verfahren innerhalb dieser Klasse werden für die latenten Produktterme Produkte aus den Indikatoren gebildet, die neben den normalen Indikatoren in das Modell aufgenommen werden. Abbildung 2.1 zeigt ein Pfaddiagramm für ein vollständiges nichtlineares Modell mit drei Produktindikatoren für jeden latenten Produktterm. Neben dem ursprünglichen PI-Ansatz, dem Constrained-Ansatz (Kenny & Judd, 1984), wurden weitere Verfahren entwickelt, die die Probleme des Constrained-Ansatzes lösen sollten, die durch die Modell-Constraints für die Faktorladungen (z.B. λ_{73}^z) und Fehler(ko-)varianzen der Produktindikatoren (z.B. θ_{77}^{δ}) entstehen (Marsh et al., 2004; Wall & Amemiya, 2001; Kelava et al., 2008; Kelava & Brandt, 2009).

Von den bisherigen Variationen der PI-Ansätze besitzt der Unconstrained-Ansatz (Marsh et al., 2004) die theoretisch besten Schätzeigenschaften, da bei ihm ein Großteil dieser Constraints nicht mehr gesetzt werden muss, sondern die Parameter frei geschätzt werden können, wodurch weniger Annahmen getroffen werden müssen. Die Erweiterung für vollständige nichtlineare Modelle, die eine Reihe von korrekt spezifizierten Residualkovarianzen beinhaltet, wurde von Kelava und Brandt (2009) entwickelt.

Für den erweiterten Unconstrained-Ansatz werden Verteilungsvoraussetzungen für zwei Aspekte benötigt: Für die Modellformulierung und für die auf das Modell angewendete Schätzung. Für die Darstellung der Annahmen kann man zwischen einem linearen und einem nichtlinearen Teilmodell der Prädiktoren unterscheiden (s. Abbildung 2.1). Im linearen Teilmodell sind die Messmodelle der Prädiktoren (ξ_1, ξ_2) enthalten, im nichtlinearen Teilmodell die Messmodelle der Produktterme. Die *Modellformulierung* beinhaltet eine Reihe von Kovarianzen sowohl innerhalb des nichtlinearen Teilmodells als auch zwischen dem linearen und dem nichtlinearen Teilmodell. Innerhalb des nichtlinearen Teilmodells betrifft dies neben den zu schätzenden Kovarianzen zwischen den Produkttermen auch die Kovarianzen zwischen den Residuen der Produktindikatoren (z.B. $Cov(\delta_7, \delta_{10})$, s. Abbildung 2.1). Zwischen den beiden Teilmodellen betrifft es die auf null fixierten Kovarianzen (z.B. $Cov(\delta_1, \delta_7)$). Während für die Gültigkeit der Cons-

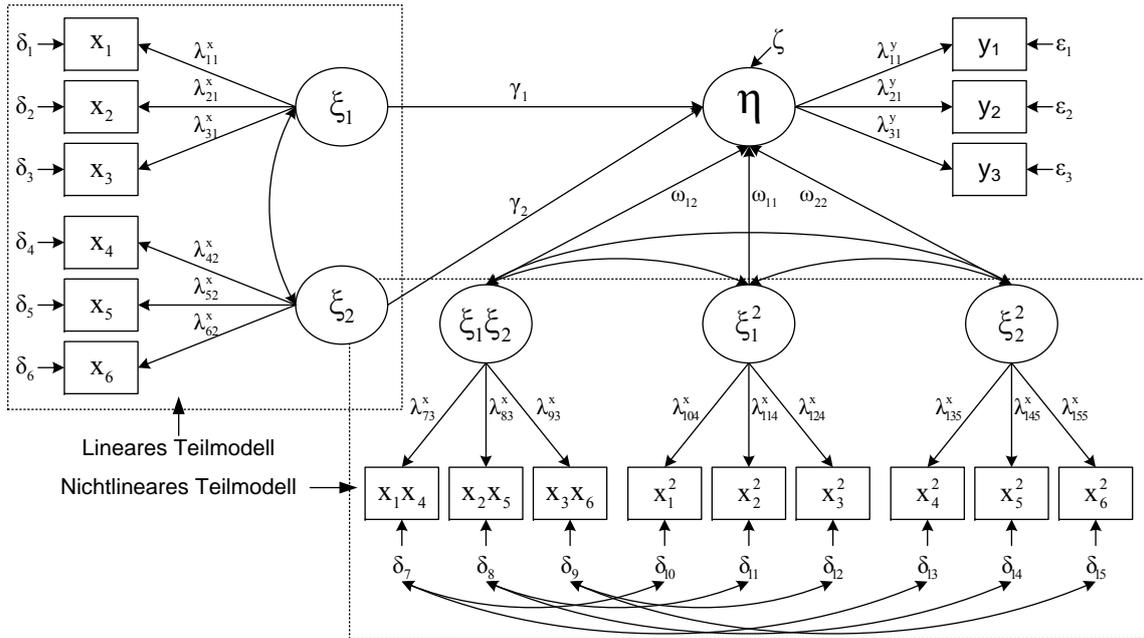


Abbildung 2.1: Nichtlineares Strukturgleichungsmodell mit einem Interaktionseffekt und zwei quadratischen Effekten. Die Produktindikatoren ($x_1 x_4, x_1^2, x_4^2$, etc.) werden nur für die PI-Ansätze benötigt.

traints im Constrained-Ansatz eine multivariate Normalverteilung aller exogenen Variablen im linearen Teilmodell notwendig ist (vgl. Kelava et al., 2008; Wall & Amemiya, 2001), kann diese Annahme für den Unconstrained-Ansatz abgeschwächt werden. Alle ursprünglich durch die Constraints restringierten Faktorladungen, Residualvarianzen und (Ko)-Varianzen der latenten Produktterme werden frei geschätzt (Kelava & Brandt, 2009). Jedoch ist die multivariate Normalverteilung der Variablen des linearen Teilmodells notwendig, damit die (auf null) fixierten Kovarianzen zwischen den latenten Variablen des linearen und des nichtlinearen Teilmodells gelten. Beispielsweise gilt $Cov(\delta_1, \delta_7) = 0$ bzw. $Cov(\xi_1, \xi_1 \xi_2) = 0$ nur dann, wenn δ_1 bzw. ξ_1 normalverteilt sind (und einen Mittelwert von 0 besitzen; vgl. Kelava & Brandt, 2009; Wall & Amemiya, 2001). Wird zudem zur Schätzung des Modells ein ML-Schätzer verwendet, so ist hierfür die multivariate Normalverteilung aller Indikatoren und Produktindikatoren Voraussetzung.⁶

⁶Verteilungsfreie Schätzer wie WLSA, die augenscheinlich eine plausible Alternative wären, zeigten eine schlechtere Performanz als der ML-Schätzer bei empirisch relevanten Stichprobengrößen zwischen $N = 200$ und $N = 800$ (Schermelleh-Engel, Klein & Moosbrugger, 1998) und werden hier nicht weiter diskutiert.

Diskussion der Annahmen

Die Annahme der multivariaten Normalverteilung aller exogenen Variablen des linearen Teilmodells ($\xi_1, \xi_2, \delta_1, \dots, \delta_q$) ist in den Fällen plausibel, in denen die Indikatoren (x_1, \dots, x_q) normalverteilt sind. Da sich die Indikatoren aus den Faktoren und den Residuen zusammensetzen (die jeweils unkorreliert sind), kann man allgemein den Schluss ziehen, dass bei Normalität der Indikatoren auch eine Normalität der unterliegenden Variablen gegeben ist (Satz von Cramér; Cramér, 1936, 1999). Die für die ML-Schätzung verwendete multivariate Normalverteilung aller Variablen auch der Produktindikatoren ist hingegen nicht plausibel, denn sie ist stets verletzt, da Produkte von normalverteilten Variablen nicht-normalverteilt sind (Aroian, 1944). Die Annahme ist umso stärker verletzt, je größer die nichtlinearen Effekte sind, da dann ebenso das Kriterium und seine Indikatoren (y_1, \dots, y_p) nicht-normalverteilt sein müssen. Eine Testung der Normalität der Indikatoren ist möglich, auch wenn die Verfahren zur Testung einer multivariaten Verteilung eine Reihe von Nachteilen und Schwierigkeiten aufweisen (Mecklin & Mundfrom, 2004).

Solange die Indikatoren x_1, \dots, x_q normalverteilt sind, so sind ausschließlich die Voraussetzungen für den ML-Schätzer durch die Nichtnormalität der Produktindikatoren verletzt, nicht aber die Voraussetzungen für die Modellformulierung. In diesem Fall ist hinsichtlich der Robustheit des Verfahrens zu erwarten, dass die Standardfehler unterschätzt werden, wie in verschiedenen Studien zu linearen SEM mit nicht-normalverteilten Variablen gezeigt wurde (Bollen, 1989; Boomsma, 1983; West, Finch & Curran, 1995). Die Parameterschätzungen selbst sind jedoch zumeist weiterhin erwartungstreu, wenn das Modell (bis auf die Verteilungsannahmen) korrekt spezifiziert ist. Sind neben den Produktindikatoren auch die Indikatoren des linearen Teilmodells nicht-normalverteilt, sind notwendigerweise auch zumindest ein Teil der latenten exogenen Variablen des linearen Teilmodells nicht-normalverteilt. Dadurch liegt eine Fehlspezifikation im Modell vor, da die Kovarianzen zwischen dem linearen und dem nichtlinearen Teilmodell nicht mehr alle null sind, was zumindest theoretisch zu einem Bias in der Parameterschätzung führen kann.

Aufgrund der Konzeptualisierung mit Produktindikatoren ist das Verfahren nicht sehr sparsam, da viele Parameter im nichtlinearen Teilmodell geschätzt werden müssen. Daher ist damit zu rechnen, dass das Verfahren eine geringere Effizienz als andere, sparsamer parametrisierte Verfahren aufweist.

2.1.2 Verteilungsanalytische Ansätze: LMS und QML

Zur Gruppe der verteilungsanalytischen Ansätze gehören vor allem die Ansätze LMS („Latent Moderated Structures“; Klein, 2000; Klein & Moosbrugger, 2000) und QML („Quasi-ML“; Klein & Muthén, 2007). In diesen Verfahren werden die latenten Produktterme nicht durch manifeste Indikatoren operationalisiert, sondern die nichtlinearen Effekte durch eine Verteilungsanalyse der Kriteriumsvariablen geschätzt. Diese Verteilungsanalyse nutzt die Eigenschaft nichtlinearer Modelle aus, dass das Kriterium modelltheoretisch einer Mischung bedingter Normalverteilungen auf den einzelnen Stufen der Moderatorvariablen folgt.

Die Hauptannahme für LMS und QML ist die multivariate Normalverteilung aller exogenen Variablen $(\xi_1, \xi_2, \delta_1, \dots, \delta_q, \epsilon_1, \dots, \epsilon_p, \zeta)$ und damit implizit die Normalverteilung aller Indikatoren der Prädiktoren $\mathbf{x} = (x_1, \dots, x_q)'$, nicht aber die der Indikatoren des Kriteriums $\mathbf{y} = (y_1, \dots, y_p)'$. LMS modelliert die nicht-normale Verteilung der Indikatoren (\mathbf{x}, \mathbf{y}) als eine bedingte Mischverteilung aus multivariaten Normalverteilungen, deren bedingte Mittelwerte und (Ko)-Varianzen eine komplexe Funktion beinhalten, die nicht analytisch lösbar ist (Klein & Moosbrugger, 2000).⁷ Daher wird das Integral durch numerische Methoden approximiert, beispielsweise mit dem Hermite-Gauss-Verfahren (Klein & Moosbrugger, 2000; Stoer, 1989). Während das Hermite-Gauss-Verfahren das Integral durch eine endliche Mischung von bedingten Normalverteilungen approximiert, werden in der Mplus-Implementierung von LMS andere (simplere) numerische Methoden verwendet, wie die numerische Integration anhand der Trapezregel (Muthén & Muthén, 1998–2012; Stoer, 1989)⁸. Für die approximierte Verteilung kann eine ML-Parameterschätzung anhand des EM-Algorithmus (Dempster, Laird & Rubin, 1977) vorgenommen werden.

Für QML werden im Vergleich zu LMS zwei Vereinfachungen für das Modell vorgenommen (Klein & Muthén, 2007). Erstens wird die Nichtnormalität des Indikatorvektors auf einen einzigen Indikator reduziert (y_1); aus den weiteren Indikatoren des Kriteriums wird die Nichtnormalität extrahiert, sodass die neuen Indikatoren $\mathbf{y}^* = (y_2^*, \dots, y_p^*)'$ normalverteilt sind. Zweitens wird die nicht-normale Verteilungsfunktion $f(y_1, \mathbf{y}^*, \mathbf{x})$ als Produkt einer normalen Verteilungsfunktion $f(\mathbf{y}^*, \mathbf{x})$ und einer bedingt nicht-normalen Verteilungsfunktion $f(y_1 | \mathbf{y}^*, \mathbf{x})$ dargestellt. Dadurch wird das Problem der Lösung des Integrals für die Verteilungsfunktion $f(y_1, \mathbf{y}^*, \mathbf{x})$, die eine multivariate nicht-normale Verteilung beinhaltet, auf die einfachere Lösung für eine multivariate Normalverteilung und für eine univariate bedingt nicht-normale Verteilung reduziert. Diese bedingt nicht-normale Verteilungsfunktion $f^*(y_1 | \mathbf{y}^*, \mathbf{x})$ wird durch eine bedingt normale Verteilungs-

⁷Die technischen Details, auch die zur Cholesky-Zerlegung der Prädiktoren ξ_1, \dots, ξ_v in unabhängige Komponenten z_1, \dots, z_v , auf die bedingt wird, sind in Klein und Moosbrugger (2000) zu finden.

⁸Diese technischen Aspekte wirken sich jedoch im Allgemeinen nicht auf die Schätzgüte aus, sondern nur auf den Rechenaufwand.

funktion $f^*(y_1|\mathbf{y}^*, \mathbf{x})$ mit gleichem Erwartungswert und gleicher Varianz wie $f(y_1|\mathbf{y}^*, \mathbf{x})$ approximiert. Durch diese Approximation liefert QML keine echten ML-Schätzungen, denn die Parameterschätzung basiert auf einer nicht korrekt spezifizierten Verteilungsfunktion.

Der Unterschied zwischen LMS und QML liegt darin, dass bei LMS die korrekt spezifizierte Verteilung (Likelihood) numerisch approximiert wird, während bei QML die vereinfachte Verteilungsfunktion numerisch approximiert wird. Dadurch ist die Approximation bei QML deutlich weniger aufwendig (und Analysen mit QML sind daher schneller). Zudem ist die Verteilung des Kriteriums (η) bei QML nicht mehr eine exakte Mischung von Normalverteilungen, sondern es genügt, wenn die bedingten Varianzen und Mittelwerte der Verteilung $f^*(y_1|\mathbf{y}^*, \mathbf{x})$ erwartungstreue Schätzer für die der Verteilungsfunktion $f(y_1|\mathbf{y}^*, \mathbf{x})$ sind.

Weitere technische Details zu LMS und QML können der angegebenen Originalliteratur entnommen werden (Klein, 2000; Klein & Moosbrugger, 2000; Klein & Muthén, 2007). Ein weiterer theoretischer Vergleich dieser beiden Verfahren ist bei Kelava et al. (2011) zu finden.

Diskussion der Annahmen

Eine direkte Testung der Verteilung der latenten exogenen Variablen kann nicht erfolgen; aber der oben erwähnte Schluss gilt, dass eine Normalität der Indikatoren auch auf eine Normalität der unterliegenden Konstrukte zurückzuführen ist. In diesem Sinne sind die Annahmen auch beschränkt testbar (s. oben), d.h. eine Nichtnormalität in den Indikatoren weist auf eine Verletzung der Annahmen hin.

Für LMS und QML ist zu erwarten, dass eine Nichtnormalität der Prädiktoren (oder der Residuen) zu einem Bias in der Parameterschätzung führt, da die Verteilungsannahme direkt in die Modellspezifikation einfließt. Da QML eine Approximation der nicht-normalen Verteilungsfunktion verwendet, ist jedoch zumindest theoretisch zu erwarten, dass QML in diesem Fall einen geringeren Bias als LMS aufweist, aber dass QML bei normalverteilten Variablen etwas weniger effizient ist als LMS (Kelava et al., 2011), da LMS ein ML-Verfahren ist und damit die effizientesten Schätzungen liefern muss (Kendall, 1948; White, 1982).

2.1.3 Der 2SMM-Ansatz

Der „2 Step Method of Moments“- (2SMM)-Ansatz (Wall & Amemiya, 2000, 2003) ist ein zweistufiges Verfahren und basiert auf einer Error-in-Variables-Regression. In einem ersten Schritt werden in einem Standard-CFA-Modell die (Bartlett)-Faktorwerte für die latenten Variablen (η, ξ_1, ξ_2) geschätzt. In einem zweiten Schritt werden die gewonnenen

Faktorwert-Schätzungen in einer Regression für das Strukturmodell verwendet. Bei dieser Schätzung wird berücksichtigt, dass die im ersten Schritt für die $i = 1 \dots n$ Personen erhaltenen Faktorwerte $(\hat{\eta}_i, \hat{\xi}_{1i}, \hat{\xi}_{2i})$ Schätzungen für die wahren Werte $(\eta_i, \xi_{1i}, \xi_{2i})$ sind und somit ein Residuum besitzen. Konkret ergeben sich die geschätzten Faktorwerte modelltheoretisch als

$$\begin{aligned}\hat{\xi}_{1i} &= \xi_{1i} + e_{1i} \\ \hat{\xi}_{2i} &= \xi_{2i} + e_{2i} \\ \hat{\eta}_i &= \eta_i + e_{3i}.\end{aligned}\tag{2.1}$$

Die (Ko)-Varianzen und höheren Momente der Residuen e_1, e_2, e_3 werden für die Herleitung der erwartungstreuen, momentbasierten Gleichungen verwendet. Die Parameter eines vollständigen nichtlinearen Modells nach Gleichung 1.4 werden analog zur Standardregression als Lösung der Gleichung $\mathbf{M}\boldsymbol{\theta} = \mathbf{m}$ für $\boldsymbol{\theta} = (\alpha, \gamma_1, \gamma_2, \omega_{12}, \omega_{11}, \omega_{22})'$ definiert, wobei \mathbf{M} und \mathbf{m} auf Populationsebene gegeben sind durch

$$\mathbf{M} = \frac{1}{n} \sum_{i=1}^n \mathbf{M}_i = \frac{1}{n} \sum_{i=1}^n \begin{pmatrix} 1 & \xi_{1i} & \xi_{2i} & \xi_{1i}\xi_{2i} & \xi_{1i}^2 & \xi_{2i}^2 \\ \xi_{1i} & \xi_{1i}^2 & \xi_{1i}\xi_{2i} & \xi_{1i}^2\xi_{2i} & \xi_{1i}^3 & \xi_{1i}\xi_{2i}^2 \\ \xi_{2i} & \xi_{1i}\xi_{2i} & \xi_{2i}^2 & \xi_{1i}\xi_{2i}^2 & \xi_{1i}^2\xi_{2i} & \xi_{2i}^3 \\ \xi_{1i}\xi_{2i} & \xi_{1i}^2\xi_{2i} & \xi_{1i}\xi_{2i}^2 & \xi_{1i}^2\xi_{2i}^2 & \xi_{1i}^3\xi_{2i} & \xi_{1i}\xi_{2i}^3 \\ \xi_{1i}^2 & \xi_{1i}^3 & \xi_{1i}^2\xi_{2i} & \xi_{1i}^3\xi_{2i} & \xi_{1i}^4 & \xi_{1i}^2\xi_{2i}^2 \\ \xi_{2i}^2 & \xi_{1i}\xi_{2i}^2 & \xi_{2i}^3 & \xi_{1i}\xi_{2i}^3 & \xi_{1i}^2\xi_{2i}^2 & \xi_{2i}^4 \end{pmatrix}, \tag{2.2}$$

$$\mathbf{m} = \frac{1}{n} \sum_{i=1}^n \mathbf{m}_i = \frac{1}{n} \sum_{i=1}^n \begin{pmatrix} \eta_i \\ \eta_i\xi_{1i} \\ \eta_i\xi_{2i} \\ \eta_i\xi_{1i}\xi_{2i} \\ \eta_i\xi_{1i}^2 \\ \eta_i\xi_{2i}^2 \end{pmatrix}. \tag{2.3}$$

\mathbf{M} enthält die Quadrat- und Kreuzproduktsummen der Prädiktoren und ihrer Produkte und \mathbf{m} die Kreuzproduktsummen von η mit den Prädiktoren und ihrer Produkte. Der Momentschätzer $\hat{\boldsymbol{\theta}}$ ergibt sich folglich als

$$\hat{\boldsymbol{\theta}} = \hat{\mathbf{M}}^{-1}\hat{\mathbf{m}}. \tag{2.4}$$

Die Matrizen $\hat{\mathbf{M}}$ und $\hat{\mathbf{m}}$ müssen so spezifiziert werden, dass der Momentschätzer erwartungstreu ist, was der Fall ist, wenn für alle $i = 1 \dots n$ Personen gilt:

$$E[\hat{\mathbf{M}}_i | \xi_{1i}, \xi_{2i}] = \mathbf{M}_i \text{ und } E[\hat{\mathbf{m}}_i | \xi_{1i}, \xi_{2i}, \eta_i] = \mathbf{m}_i. \tag{2.5}$$

Die für das vollständige nichtlineare Modell mit einem Interaktionseffekt und zwei quadratischen Effekten erweiterten momentbasierten Gleichungen für die Matrizen $\hat{\mathbf{M}}$ und $\hat{\mathbf{m}}$ sowie die Schätzungen der Standardfehler sind in Anhang A zu finden. An dieser Stelle soll an einem Beispiel die Logik der Herleitung der Momentgleichungen demonstriert werden. Um einen erwartungstreuen Schätzer für das Moment $\xi_{1i}\xi_{2i}^2$ (Element M_{26} aus der Matrix \mathbf{M}_i , s. Gleichung 2.2) zu definieren, muss die Kovarianz der Residuen e_1 und e_2 aus Gleichung 2.1 (σ_{12}), die Varianz von e_2 (σ_{22}) und das dritte Moment $E[e_1e_2e_2] = \mu_{122}$ berücksichtigt werden. Als erwartungstreuer Schätzer ergibt sich $\hat{M}_{26} = \hat{\xi}_{1i}\hat{\xi}_{2i}^2 - 2\hat{\sigma}_{12}\hat{\xi}_{2i} - \hat{\sigma}_{22}\hat{\xi}_{1i} - \hat{\mu}_{122}$, da Gleichung 2.5 für jede Person i erfüllt ist:

$$\begin{aligned}
E[\hat{M}_{26}|\xi_{1i}, \xi_{2i}] &= E[\hat{\xi}_{1i}\hat{\xi}_{2i}^2 - 2\hat{\sigma}_{12}\hat{\xi}_{2i} - \hat{\sigma}_{22}\hat{\xi}_{1i} - \hat{\mu}_{122}|\xi_{1i}, \xi_{2i}] \\
&= E[(\xi_{1i} + e_{1i})(\xi_{2i} + e_{2i})^2 - 2\hat{\sigma}_{12}(\xi_{2i} + e_{2i}) - \hat{\sigma}_{22}(\xi_{1i} + e_{1i}) - \hat{\mu}_{122}|\xi_{1i}, \xi_{2i}] \\
&= E[\xi_{1i}\xi_{2i}^2 + 2\xi_{1i}\xi_{2i}e_{2i} + \xi_{1i}e_{2i}^2 + e_{1i}\xi_{2i}^2 + 2\xi_{2i}e_{1i}e_{2i} + e_{1i}e_{2i}^2 \\
&\quad - 2\hat{\sigma}_{12}\xi_{2i} - 2\hat{\sigma}_{12}e_{2i} - \hat{\sigma}_{22}\xi_{1i} - \hat{\sigma}_{22}e_{1i} - \hat{\mu}_{122}|\xi_{1i}, \xi_{2i}] \\
&= \underbrace{E[\xi_{1i}\xi_{2i}^2|\xi_{1i}, \xi_{2i}]}_{\xi_{1i}\xi_{2i}^2} + \underbrace{E[2\xi_{1i}\xi_{2i}e_{2i}|\xi_{1i}, \xi_{2i}]}_0 + \underbrace{E[\xi_{1i}e_{2i}^2|\xi_{1i}, \xi_{2i}]}_{\sigma_{22}\xi_{1i}} \\
&\quad + \underbrace{E[e_{1i}\xi_{2i}^2|\xi_{1i}, \xi_{2i}]}_0 + \underbrace{E[2\xi_{2i}e_{1i}e_{2i}|\xi_{1i}, \xi_{2i}]}_{2\sigma_{12}\xi_{2i}} + \underbrace{E[e_{1i}e_{2i}^2|\xi_{1i}, \xi_{2i}]}_{\mu_{122}} \\
&\quad - \underbrace{E[2\hat{\sigma}_{12}\xi_{2i}|\xi_{1i}, \xi_{2i}]}_{2\sigma_{12}\xi_{2i}} - \underbrace{E[2\hat{\sigma}_{12}e_{2i}|\xi_{1i}, \xi_{2i}]}_0 - \underbrace{E[\hat{\sigma}_{22}\xi_{1i}|\xi_{1i}, \xi_{2i}]}_{\sigma_{22}\xi_{1i}} \\
&\quad - \underbrace{E[\hat{\sigma}_{22}e_{1i}|\xi_{1i}, \xi_{2i}]}_0 - \underbrace{E[\hat{\mu}_{122}|\xi_{1i}, \xi_{2i}]}_{\mu_{122}} \\
&= \xi_{1i}\xi_{2i}^2, \tag{2.6}
\end{aligned}$$

wenn $E[e_1] = E[e_2] = 0$ gilt, e_1, e_2 untereinander und mit ξ_1, ξ_2 unkorreliert und die Schätzer $\hat{\sigma}_{12}, \hat{\sigma}_{22}, \hat{\mu}_{122}$ erwartungstreu sind.

Allgemein werden keine Verteilungsannahmen für die Herleitung der Momentgleichungen benötigt. Wird eine Einfachstruktur des Messmodells mit normalverteilten Residuen angenommen, so können jedoch die dritten und vierten Momente vereinfacht werden ($\mu_{jkl} = 0$ und $\nu_{jklm} = \sigma_{jk}\sigma_{lm}$).⁹ Dies ist für die Schätzung von Vorteil, da die Schätzung höherer Momente erst bei größeren Stichproben zuverlässig erfolgen kann (Joanes & Gill, 1998). Weiterhin ist die Annahme normalverteilter Residuen eine Voraussetzung für die Konsistenz des Standardfehlerschätzers.

⁹Vierte Momente kommen ebenfalls in $\hat{\mathbf{M}}$ vor und zwar für die Momente aus jeweils zwei Produkttermen (s. Anhang A).

Diskussion der Annahmen

Die Hauptannahme für die Konsistenz des Parameter- und des Standardfehlerschätzers ist die Normalverteilung der Residuen des Messmodells $(\delta_1, \dots, \delta_q, \epsilon_1, \dots, \epsilon_p)$. Die latenten Prädiktoren und das latente Kriterium hingegen können einer beliebigen Verteilung folgen. Wenn die Indikatoren nicht-normalverteilt sind und die Indikatoren reliable Messungen darstellen, so ist davon auszugehen, dass die Nichtnormalität hauptsächlich durch die Faktoren und nicht durch die Residuen verursacht werden. In diesem Fall wäre die Annahme immer noch plausibel und nicht verletzt. Bei einer geringen Reliabilität und gleichzeitiger Nichtnormalität der Indikatoren ist jedoch zu erwarten, dass die Annahme verletzt ist. Inhaltlich kann die Plausibilität anhand von Vorwissen über Konstrukte oder der Messung eingeschätzt werden (Molenaar, Dolan & Verhelst, 2010). Eine nicht-normale Verteilung der Residuen kann beispielsweise auf Boden-/Deckeneffekte oder Antwortformate eines Messinstruments zurückzuführen sein. Sind die Indikatoren nicht-normalverteilt, so kann empirisch nicht entschieden werden, ob dies auf die latenten Prädiktoren oder die Residuen zurückzuführen ist (Molenaar et al., 2010). Eine Aussage über die Nichtnormalität latenter Variablen ist daher nur im Sinne der Plausibilität möglich.

Wenn die Indikatoren reliable Messungen darstellen, ist zu erwarten, dass der 2SMM-Schätzer robust ist. Neben der Hauptannahme der Normalverteilung der Residuen hängt jedoch die Robustheit des Verfahrens auch von der Schätzung der Faktorwerte ab. Wird hierfür ein ML-Schätzer verwendet und sind die Indikatoren nicht-normalverteilt, so kann in Analogie zu den Ergebnissen zu linearen SEM (s. oben) geschlussfolgert werden, dass die Parameterschätzungen erwartungstreu sind, jedoch die Standardfehler unterschätzt werden. Die erhaltenen Faktorwerte, die auf den ML-Parameterschätzungen basieren, und die Parameterschätzungen des Strukturmodells wären somit ebenfalls erwartungstreu, die Standardfehler des 2SMM-Ansatzes aber möglicherweise unterschätzt.

Da für die Schätzung des Strukturmodells nicht die Information zur spezifischen Verteilungsfunktion von η verwendet wird, sondern eine konstante Varianz angenommen wird, sollte das Verfahren weniger effizient sein als LMS und QML (s. hierzu auch Klein & Muthén, 2007).

2.1.4 Der MM-Ansatz

Der MM-Ansatz („Method of Moments“; Mooijaart & Bentler, 2010) berücksichtigt für die Schätzung der nichtlinearen Effekte Momente höherer Ordnung der Prädiktoren. Die zweiten Momente (Varianzen und Kovarianzen) und dritten Momente der Produktterme werden als Funktionen der (Ko)-Varianzen und dritten Momente der Prädiktoren

(ξ_1, ξ_2) formuliert. Hierfür wird für die Prädiktoren (ξ_1, ξ_2) eine Normalverteilung und ein Mittelwert von null vorausgesetzt, wodurch ungerade Momente null sind und vierte bzw. sechste Momente als einfache Funktionen der zweiten Momente dargestellt werden können. Beispielsweise kann unter dieser Annahme die Kovarianz zwischen dem latenten Interaktionsterm $\xi_1\xi_2$ und dem quadratischen Term ξ_1^2 ausgedrückt werden als

$$\begin{aligned} Cov(\xi_1\xi_2, \xi_1^2) &= E[\xi_1\xi_2\xi_1\xi_1] - E[\xi_1\xi_2]E[\xi_1\xi_1] \\ &= \phi_{1211} - \phi_{12}\phi_{11} \\ &= \phi_{12}\phi_{11} + \phi_{11}\phi_{12} + \phi_{11}\phi_{12} - \phi_{12}\phi_{11} \\ &= 2\phi_{11}\phi_{12}, \end{aligned} \tag{2.7}$$

wobei ϕ_{jk} das zweite Moment der Prädiktoren ξ_j und ξ_k darstellt und ϕ_{jklm} das vierte Moment (vgl. Mooijaart & Bentler, 2010). In der modelltheoretischen Kovarianzmatrix vereinfachen sich dadurch die Momente, die auf die latenten Produktterme zurückzuführen sind. Beispielsweise ist die modelltheoretische Kovarianz zwischen y_1 und y_2 für ein vollständiges nichtlineares Modell gegeben durch

$$Cov(y_1, y_2) = \lambda_1^y \lambda_2^y [\gamma_1^2 \phi_{11} + \gamma_2^2 \phi_{22} + 2\gamma_1 \gamma_2 \phi_{12} + 2(\omega_{12} \phi_{12} + \omega_{11} \phi_{11} + \omega_{22} \phi_{22})^2]. \tag{2.8}$$

Zur Identifikation der nichtlinearen Effekte müssen zudem dritte Momente der Indikatoren in der Schätzung berücksichtigt werden (näheres hierzu s. Mooijaart & Satorra, 2011). Anhand einer Weighted-Least-Squares-(WLS)-Diskrepanzfunktion werden die Parameter $(\boldsymbol{\theta})$ geschätzt:

$$F_{WLS}(\mathbf{S}, \boldsymbol{\Sigma} | \boldsymbol{\theta}) = [\mathit{vech}(\mathbf{S}) - \mathit{vech}(\boldsymbol{\Sigma}(\boldsymbol{\theta}))]' \mathbf{W}^{-1} [\mathit{vech}(\mathbf{S}) - \mathit{vech}(\boldsymbol{\Sigma}(\boldsymbol{\theta}))], \tag{2.9}$$

wobei $\mathit{vech}(\mathbf{S})$ der Vektor ist, der die nichtredundanten ersten, zweiten und ausgewählten dritten Momente der Indikatorvariablen beinhaltet und $\mathit{vech}(\boldsymbol{\Sigma}(\boldsymbol{\theta}))$ die entsprechenden modelltheoretischen Momente (vgl. Gleichung 3, Mooijaart & Bentler, 2010; s. auch Gleichung 1.15 in Jöreskog & Sörbom, 1996). Für die Standardfehlerschätzung werden sechste Momente benötigt, deren Schätzungen sehr instabil sind, weshalb für die Gewichtungsmatrix \mathbf{W} nach Mooijaart und Bentler (2010) die Identitätsmatrix \mathbf{I} verwendet werden sollte.

Da der WLS-Schätzer ein asymptotisch verteilungsfreier Schätzer ist (Browne, 1984), müssen neben den Modellannahmen für die Herleitung der Momente der Produktterme keine Verteilungsannahmen getroffen werden, um die Konsistenz der Schätzung zu gewährleisten. Es müssen also explizit keine Annahmen über die Verteilung der latenten Residuen getroffen werden.

Diskussion der Annahmen

Ebenso wie beim 2SMM-Ansatz ist nicht testbar, ob eine Nichtnormalität der Indikatoren auf die latenten Prädiktoren oder die Residuen zurückzuführen ist. Auch hier ist nur eine Plausibilitätsüberlegung in dem Sinne möglich, dass Kenntnisse über die Verteilungen von Konstrukten vorliegen (z.B. die Normalverteilung des IQ; Wechsler, 1944). Jedoch ist davon auszugehen, dass das Verfahren sensibel auf die Verletzung der Verteilungsannahme reagiert. Sind die latenten Prädiktoren nicht-normalverteilt, so sind die Gleichungen für die Momente fehlspezifiziert, was sich direkt auf die Schätzungen der Parameter auswirkt. Anders als beim 2SMM-Ansatz wird die Annahme der Normalverteilung für die latenten Prädiktoren verwendet, die üblicherweise einen starken Einfluss auf die Indikatoren haben. In der Empirie kann bei nicht-normalen Indikatoren und reliablen Messungen daher vermutet werden, dass der Schätzer nicht mehr erwartungstreu ist, da die Nichtnormalität vermutlich auf die Prädiktoren zurückzuführen ist. In den verwendeten dritten Momenten sind in diesem Fall die Nichtnormalität und die Nichtlinearität konfundiert. Zudem kann von einer zuverlässigen Schätzung der Momente höherer Ordnung erst bei größeren Stichproben ausgegangen werden (Joanes & Gill, 1998; Mecklin & Mundfrom, 2004). Dies hat insbesondere einen Einfluss auf die Effizienz des Verfahrens, die geringer sein sollte als z.B. bei LMS. Bei zusätzlich geringer Reliabilität ist die Schätzung der Momente der latenten Faktoren noch stärker beeinträchtigt und könnte zu weiteren Problemen führen.

2.2 Bisherige Simulationsstudien

Bisher wurden nur die theoretischen Überlegungen zur Schätzgüte und Effizienz der Verfahren berichtet. Um eine quantitative Aussage darüber treffen zu können, werden üblicherweise Simulationsstudien durchgeführt. Die meisten bisherigen größeren Studien konzentrierten sich auf die Verfahren der PI-Ansätze und verglichen diese miteinander (Kelava et al., 2008; Marsh et al., 2004; Moulder & Algina, 2002). Studien, die die momentbasierten Verfahren, die PI-Ansätze und die verteilungsanalytischen Ansätze hinsichtlich ihrer Robustheit und Effizienz vergleichend untersuchen, existieren bisher nicht.

Aus den bisherigen Simulationsstudien kann man für Modelle mit einem einzelnen Interaktionseffekt und normalverteilten Variablen zusammenfassend schließen, dass QML und LMS ähnliche Schätzergebnisse liefern und effizienter sind als der Unconstrained-Ansatz (Klein & Muthén, 2007). Bei nicht-normalverteilten Prädiktoren tritt jedoch ein Bias in der Parameterschätzung bei QML und LMS auf, während der Unconstrained-Ansatz erwartungstreue Parameterschätzungen, aber unterschätzte Standardfehler liefert (Klein & Muthén, 2007; Marsh et al., 2004). Der MM-Ansatz liefert bei einer hohen

Tabelle 2.1: Zentrale Simulationsstudien, in denen die fünf relevanten Verfahren untersucht wurden, mit der Angabe der Anzahl der getesteten nichtlinearen Effekte und der Angabe, ob nicht-normalverteilte Variablen verwendet wurden.

Artikel	verwendete Verfahren	# ω	NNV
Cham et al. (2012)	LMS, Unconstrained	1	ja
Klein und Moosbrugger (2000)	LMS	1	nein
Klein und Muthén (2007), Studie 1	QML, Unconstrained	1	ja
Klein und Muthén (2007), Studie 2	QML, LMS	3	nein
Marsh et al. (2004)	QML, Unconstrained	1	ja
Mooijaart und Bentler (2010)	LMS, MM	1	nein
Moosbrugger et al. (2009)	QML, LMS, Unconstrained	3	nein
Wall und Amemiya (2003)	2SMM	1	ja

Anmerkung: # ω – Anzahl der getesteten nichtlinearen Effekte, NNV – Verwendung von nicht-normalverteilten Variablen

Reliabilität ähnliche Ergebnisse für normalverteilte Variablen wie LMS (Mooijaart & Bentler, 2010). Der 2SMM-Ansatz ist robust gegenüber Nichtnormalität (Wall & Amemiya, 2003).

Für ein vollständiges nichtlineares Modell existieren nur Untersuchungen für normalverteilte Variablen und zwar ausschließlich für den erweiterten Unconstrained-Ansatz, LMS und QML (Moosbrugger et al., 2009). Diese Ergebnisse bestätigen die theoretisch zu erwartende Effizienz der verteilungsanalytischen Verfahren.

In Tabelle 2.1 sind die zentralen Studien aufgeführt, die die fünf aktuell relevanten Verfahren in Simulationsstudien untersuchten¹⁰. Wie man ihr entnehmen kann, sind zwei Fragestellungen zu diesen Verfahren aufgrund des Designs der Studien bisher nicht beantwortbar. Erstens wurden zumeist nur Modelle mit Interaktionseffekten analysiert (z.B. Marsh et al., 2004; Wall & Amemiya, 2003), während Simulationsstudien zu Modellen mit vollständigen nichtlinearen Modellen kaum durchgeführt wurden. Daher kann nicht beurteilt werden, inwieweit insbesondere die Effizienz momentbasierter Verfahren für multiple nichtlineare Effekte vergleichbar ist zur Effizienz anderer Verfahren. Zweitens wurden die Verfahren zumeist mit normalverteilten Variablen untersucht (z.B. Moosbrugger et al., 2009; Mooijaart & Bentler, 2010). Daher ist eine Einschätzung über die Robustheit der Verfahren wie dem MM-Ansatz gegenüber der Verletzung von Verteilungsannahmen bisher nicht möglich.

Eine Ursache für die fehlenden Simulationsstudien, insbesondere für die momentbasierten Verfahren, ist die fehlende Implementierung in Standard-Software. So liegt der MM-Ansatz nur in einer bisher nicht verfügbaren Experimentalversion von EQS vor

¹⁰Andere Verfahren wie der Constrained-Ansatz wurden für diese Tabelle nicht berücksichtigt.

(EQS 6.12; Bentler, 2000–2008) und für den 2SMM-Ansatz existiert ausschließlich eine SAS-Syntax für einen einzelnen quadratischen, kubischen oder einen einzelnen Interaktionseffekt¹¹. Eine Dokumentierung für die konkrete Spezifizierung des 2SMM-Ansatzes existierte bisher zudem ausschließlich für ein Interaktionsmodell, nicht jedoch für ein vollständiges nichtlineares Modell (Wall & Amemiya, 2003).

2.3 Simulationsstudie

In Schrift (A) wird eine Simulationsstudie präsentiert, die folgende offene Fragen beantworten soll:

1. Welches der Verfahren liefert die effizientesten Schätzungen für multiple nichtlineare Effekte?
2. Welche der Verfahren liefern auch dann erwartungstreue Schätzungen für die Parameter und Standardfehler, wenn die Verteilungsannahmen verletzt sind?

Alle Details zu dieser Studie sind in Anhang C zu finden (Brandt et al., in press). An dieser Stelle sollen die wichtigsten Aspekte der Studie zusammengefasst werden.

2.3.1 Design der Simulationsstudie

Die Simulationsstudie wurde unter empirisch relevanten Bedingungen durchgeführt, bei der die Nichtnormalität der latenten Prädiktoren (und damit auch die der Indikatoren) variiert wurde. Für den erweiterten Unconstrained-Ansatz, LMS, QML, den 2SMM-Ansatz und den MM-Ansatz wurden die Erwartungstreue der Parameter- und der Standardfehlerschätzungen sowie die Power und der α -Fehler untersucht. Hierfür wurde sowohl ein einfaches Modell mit Interaktionseffekt als auch ein vollständiges nichtlineares Modell mit einem Interaktions- und zwei quadratischen Effekten bei einer Stichprobengröße von $N = 400$ und einer Indikatorreliabilität von .80 spezifiziert.

Der 2SMM- und der MM-Ansatz wurden für vollständige nichtlineare Modelle erweitert (für die Erweiterung des 2SMM-Ansatzes s. Anhang A). Im Rahmen der Erstellung von Schrift (A) wurde der 2SMM-Ansatz in R (R Development Core Team, 2011) und der MM-Ansatz in `Matlab` (MATLAB, 2010) implementiert. Für LMS wurde die Implementation in `Mplus v6.11` (Muthén & Muthén, 1998–2012) verwendet, für QML die Standalone Software `QML v3.11`. Der erweiterte Unconstrained-Ansatz wurde mit dem `lavaan`-Package in R und einem robusten Standardfehlerschätzer (Sandwich Estimator; White, 1982; Yuan & Bentler, 2000) spezifiziert.

¹¹<http://www.columbia.edu/~mmw2177/NONLINEARSEM/index.html>

2.3.2 Zusammenfassung der wichtigsten Ergebnisse

Die verteilungsanalytischen Verfahren LMS und QML liefern für normalverteilte Prädiktoren die effizientesten Schätzungen mit der höchsten Power. Sie unterscheiden sich in ihren Schätzungen nur marginal. Bei vorhandener Nichtnormalität wurde jedoch ein absoluter Bias von etwa 10% – 20% für die Parameterschätzungen gefunden.

Der 2SMM- und der (erweiterte) Unconstrained-Ansatz liefern die robustesten Parameterschätzungen. Sie sind auch dann noch erwartungstreu, wenn die Prädiktoren nicht-normalverteilt sind. Der 2SMM-Ansatz kann aber das gesetzte α -Fehlerniveau besser einhalten, als der erweiterte Unconstrained-Ansatz. Dies ist auf die Unterschätzung der Standardfehler beim erweiterten Unconstrained-Ansatz – trotz des Sandwich Estimators – zurückzuführen.

Der MM-Ansatz liefert nur für das einfache Interaktionsmodell bei normalverteilten Prädiktoren erwartungstreue Schätzungen, die jedoch wesentlich weniger effizient sind als beispielsweise bei LMS. Zudem sind sowohl die Parameter- als auch die Standardfehlerschätzungen bei mehr als einem nichtlinearen Effekt und/oder bei nicht-normalverteilten Prädiktoren sehr stark verzerrt (teilweise über 200% Bias) und das nominelle α -Fehlerniveau wird mit teilweise über 60% nicht eingehalten.

2.4 Zusammenfassung und Schlussfolgerung

In diesem Abschnitt wurden die derzeit relevanten Ansätze zur Schätzung latenter nichtlinearer Effekte diskutiert. Stärken und Schwächen der Ansätze sowie ihre Funktionsweise wurden vorgestellt und eine Simulationsstudie bestätigte die Erwartungen hinsichtlich der Performanz der Verfahren.

LMS und QML sind wie theoretisch erwartet die effizientesten Verfahren. Jedoch sind die Verteilungsannahmen vergleichsweise stärker, da bei Verletzung der Annahmen ein Bias in den Parameterschätzungen entsteht. Die Annahmen können jedoch getestet werden, da sie sich zumindest indirekt in den Verteilungen der Indikatoren widerspiegeln. In den Situationen, in denen die Indikatoren normalverteilt sind, sollte eines dieser Verfahren verwendet werden. Insbesondere die Implementierung des LMS-Schätzers in Mplus stellt für den Anwender ein sehr einfach verwendbares und erweiterbares Verfahren dar, das mit anderen Aspekten von Mplus kombinierbar ist (z.B. die Berücksichtigung von Missing Data; Schafer, 1997).

Die Schätzungen des (erweiterten) Unconstrained-Ansatzes sind robust und die Verteilungsannahmen dieses Ansatzes vergleichsweise schwächer. Die Verletzung der Annahmen wirkt sich hauptsächlich auf die Schätzung der Standardfehler aus. Aufgrund seiner aufwendigen Parametrisierung ist der Unconstrained-Ansatz nicht so praktikabel wie LMS oder QML.

Der 2SMM-Ansatz ist ein effizienter und robuster Schätzer. Seine Annahmen über die Verteilung der Residuen kann nicht getestet werden, da diese latent sind. Jedoch sollte im Allgemeinen auch eine Verletzung dieser Annahmen von untergeordneter Bedeutung sein, da die Residuen nur einen kleinen Varianzanteil der Indikatorvariablen erklären sollten. Ist dies im Falle von unreliaiblen Messungen nicht gegeben, so sollte von schlechteren Schätzeigenschaften des 2SMM-Ansatzes ausgegangen werden. Der 2SMM-Ansatz kann insbesondere dann verwendet werden, wenn die Indikatoren nicht-normalverteilt sind. Die Implementation des Verfahrens in R erlaubt nun auch eine Anwendung des Verfahrens für vollständige nichtlineare Modelle in einer Standard-Software. Ein Einsatz des 2SMM-Ansatzes für andere Modelle, z.B. für latente Wachstumskurvenmodelle (s. Kapitel 4), ist jedoch sehr aufwendig, da nicht nur die Momentgleichungen sondern auch der Schätzer für die Standardfehler hergeleitet werden müssen (s. Anhang A).

Der MM-Ansatz besitzt sehr starke (und nicht direkt testbare) Annahmen, die in der Praxis schnell verletzt sind und die dann zu stark verzerrten Schätzungen führen. Die Verwendung der höheren Momente, deren Schätzung bei empirisch relevanten Stichprobengrößen mit großer Unsicherheit behaftet ist, führt sogar ohne Verletzung der Voraussetzung bei mehr als einem nichtlinearen Effekt zu nicht zufriedenstellenden Ergebnissen.

Im folgenden Abschnitt soll nun ein neues Verfahren vorgestellt werden, das speziell für den Einsatz bei nicht-normalverteilten Prädiktoren Verwendung findet. Für dieses Verfahren wird die Annahme der Normalverteilung aller Variablen – ähnlich wie beim 2SMM-Ansatz – abgeschwächt. Die Konzeptualisierung dieses Verfahrens basiert auf den Mischverteilungsmodellen und ist im Gegensatz zum 2SMM-Ansatz leicht zu implementieren und zu erweitern.

Kapitel 3

Ein Schätzverfahren für nichtlineare Effekte bei nicht-normalverteilten Prädiktoren

Im vorherigen Kapitel wurde gezeigt, dass die derzeit relevanten Verfahren zur Schätzung latenter nichtlinearer Effekte auf zum Teil strengen Verteilungsannahmen beruhen. LMS und QML wiesen bei nicht-normalverteilten Prädiktoren keine erwartungstreuen Parameterschätzungen auf, der MM-Ansatz lieferte insgesamt nicht zufriedenstellende Ergebnisse und der erweiterte Unconstrained-Ansatz wies eine Unterschätzung der Standardfehler und damit einhergehend eine leichte α -Fehlerinflation auf. Der 2SMM-Ansatz wies zwar sehr gute Schätzeigenschaften auf, ist aber nicht leicht in Standard-Software zu implementieren.

In diesem Abschnitt soll nun ein neues Verfahren vorgestellt werden, dass die Voraussetzung der Normalverteilung der Prädiktoren abschwächt und direkt in eine Standard-Software implementierbar ist. Der „Nonlinear Structural Equation Mixture Modeling“- (NSEMM)-Ansatz erweitert die bisherigen Verfahren, als dass er die Verteilung der Prädiktoren durch eine Mischverteilung approximiert. Somit können die Prädiktoren einer beliebigen Verteilung folgen. Das Konzept und die Eigenschaften sollen in den folgenden Abschnitten dargestellt werden.

Anmerkung: Dieses Kapitel basiert zu großen Teilen auf folgendem Artikel, der in Anhang D zu finden ist:

Kelava, A., Nagengast, B., & **Brandt, H.** (*im Druck*). A nonlinear structural equation mixture modeling approach for non-normally distributed latent predictor variables. *Structural Equation Modeling*.

3.1 Idee der Mischverteilungsmodelle

Mischverteilungsmodelle basieren auf der Idee, dass eine beobachtete Nichtnormalität in einer Variablen auf zwei oder mehrere, nicht-beobachtete Subpopulationen (latente Klassen) mit jeweils eigenen Verteilungen zurückgeht. Innerhalb jeder latenten Klasse ist die Variable normalverteilt, weist jedoch gruppenspezifische Mittelwerte und Varianzen auf. Bei Nichtberücksichtigung dieser Gruppierung werden diese Normalverteilungen über die latenten Klassen hinweg vermischt, wodurch die nicht-normale Verteilung entsteht.

Ein Beispiel hierfür ist die leicht nicht-normale Verteilung der Körpergröße in den USA, die eine Mischung aus den Verteilungen der Körpergröße für Frauen und Männer darstellt (Schiefe=.028; Kurtosis=2.84). Sowohl die Körpergröße der Frauen ist annähernd normalverteilt mit Mittelwert 63.86 Zoll ($SD = 4.39$ Zoll), als auch die der Männer mit Mittelwert 69.41 Zoll ($SD = 4.48$ Zoll; McDowell, Fryar, Ogden & Flegal, 2008). Unter Berücksichtigung des Geschlechterverhältnisses Männer zu Frauen von .97 ergibt sich die Verteilung der Körpergröße f_{height} als Mischverteilung

$$f_{height} = .5076 \cdot N(63.86, 4.39) + .4924 \cdot N(69.41, 4.48). \quad (3.1)$$

Abbildung 3.1 verdeutlicht das Beispiel anhand der (gewichteten) Verteilungen der Subpopulationen und der entstehenden (leicht kurtotischen) Mischvariable ohne Berücksichtigung des Geschlechts. Die Verwendung von Mischverteilungen für die Approximation von verschiedenen Verteilungen mit unterschiedlichen Bedingungen für Schiefe und Kurtosis sowie für mehrgipflige Verteilungen sind in McLachlan und Peel (2000) zu finden.

Neben der Approximation von Verteilungen, können Mischverteilungsmodelle auch für die Approximation von a priori nicht näher spezifizierten, nichtlinearen Zusammenhängen zwischen zwei oder mehreren Variablen verwendet werden. Hierbei ist anders als in den zuvor dargestellten Interaktions- und quadratischen Effekten keine Annahme über die Art der Nichtlinearität notwendig. In Abbildung 3.2 ist ein nichtlinearer Zusammenhang der Form

$$y = .2x^3 - .25x^2 + e \quad (3.2)$$

dargestellt (durchgezogene Linie). Anhand dieses Modells wurde ein Zufallsdatensatz für $N = 500$ Fälle mit $Var(e) = 1$, $Var(x) = 3$ und $E(x) = 0$ generiert. Der Datensatz wurde mit vier Mischverteilungsmodellen mit $G = 2$ bis $G = 5$ latenten Klassen analysiert, die einen linearen Zusammenhang der Form $y = \alpha + \gamma_1 x + \epsilon$ innerhalb jeder latenten Klasse annahm. Die Funktion kann mit zunehmender Klassenzahl G genauer durch den vorhergesagten Verlauf von $\hat{y} = E(y|x)$ approximiert werden.¹

¹Der bedingte Erwartungswert ergibt sich als gewichtete Summe der Zusammenhänge in den latenten Klassen durch $E(y|x) = \sum_{g=1}^G P(g|x)E_g(y|x)$, wobei $P(g|x)$ die bedingte Klassenwahrscheinlichkeit

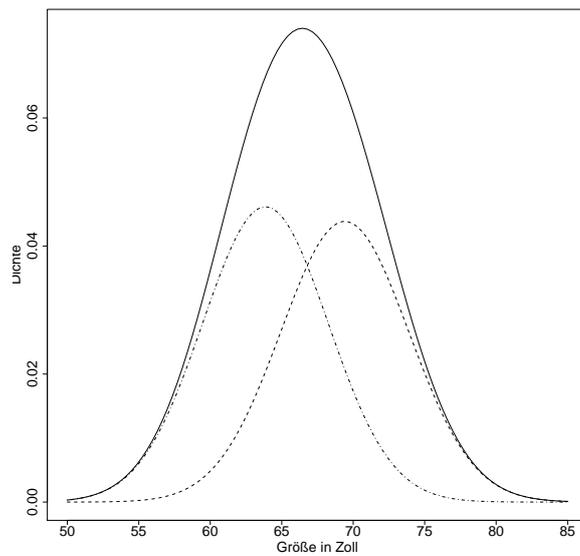


Abbildung 3.1: Verteilung der Körpergröße in den USA insgesamt (durchgezogene Linie) sowie die (gewichteten) Verteilungen getrennt für Frauen (gepunkt-strichelte Linie) und Männer (gepunktete Linie).

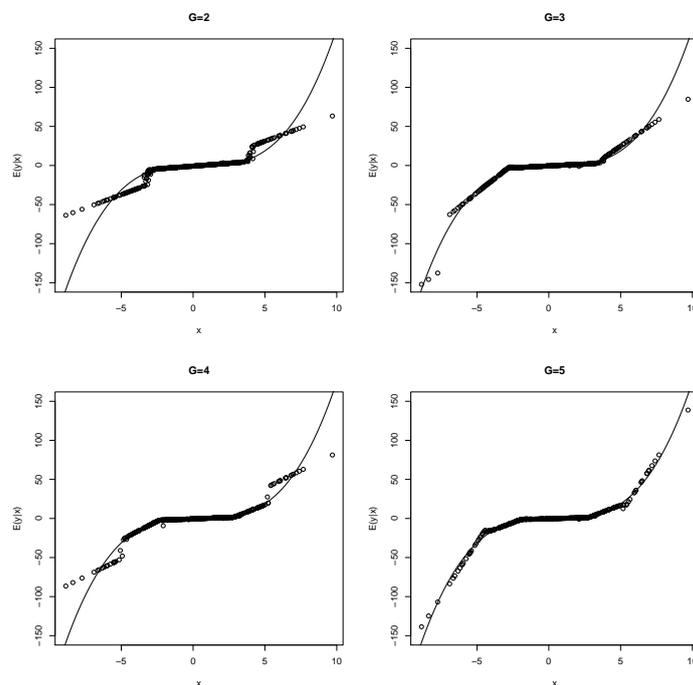


Abbildung 3.2: Streudiagramme für einen Zufallsdatensatz mit nichtlinearem Zusammenhang zwischen x und y nach Gleichung 3.2 (durchgezogene Linie). Die Werte für $E(y|x)$ (Punkte) basieren auf Mischverteilungsmodellen mit $G = 2$ bis $G = 5$ latenten Klassen.

3.2 Structural Equation Mixture Modeling

Die Verwendung von Mischverteilungen für SEM findet in zunehmenden Maße Interesse in der methodischen und empirischen Forschung (Arminger & Stein, 1997; Arminger et al., 1999; Bauer, 2005; Bauer & Curran, 2004; Dolan & van der Maas, 1998; Jedidi et al., 1997a, 1997b; Muthén, 2001; Pek, Sterba, Kok & Bauer, 2009; Pek et al., 2011). In Standard-SEM werden multivariate Zusammenhänge zwischen beobachteten und latenten Variablen modelliert, wobei zur Schätzung der Parameter des Modells eine Likelihood-Funktion formuliert wird, die auf einer multivariaten Normalverteilung der beobachteten Variablen basiert (s. Kapitel 1; vgl. Bollen, 1989; Reinecke, 2005). Die modelltheoretische Verteilungsfunktion der Indikatoren der abhängigen Variable $\mathbf{y} = (y_1, \dots, y_p)'$ und der der Prädiktoren $\mathbf{x} = (x_1, \dots, x_q)'$ ist gegeben durch

$$f((\mathbf{y}', \mathbf{x}')') = N(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \quad (3.3)$$

wobei N eine $(p + q)$ -dimensionale Normalverteilung mit modellimpliziertem Mittelwertsvektor $\boldsymbol{\mu}$ und modellimplizierter Kovarianzmatrix $\boldsymbol{\Sigma}$ ist. Für die Schätzung eines SEM wird also für alle Personen dieselbe multivariate Verteilung verwendet.

Die „Structural Equation Mixture Models“ (SEMM) unterscheiden sich nun durch die Annahme, dass die Population aus einer Anzahl von G latenten Klassen besteht. Die Verteilung der Variablen in einer latenten Klasse wird mit dem Begriff „Mixture (component)“ bezeichnet; die Verteilung der Variablen in der Population Mischverteilung („Mixture distribution“). Für jede der G latenten Klassen wird eine separate multivariate Normalverteilung mit spezifischem Mittelwertsvektor $\boldsymbol{\mu}_g$ und Kovarianzmatrix $\boldsymbol{\Sigma}_g$ angenommen. Die beobachtete (potentiell nicht-normale) multivariate Verteilung $f((\mathbf{y}', \mathbf{x}')')$ wird durch eine endliche, gewichtete Mischung der G multivariaten Normalverteilungen modelliert:

$$f((\mathbf{y}', \mathbf{x}')') = \sum_{g=1}^G w_g N(\boldsymbol{\mu}_g, \boldsymbol{\Sigma}_g), \quad (3.4)$$

wobei w_g die Gewichtungen der einzelnen Verteilungen sind (mit $w_g > 0$ und $\sum_{g=1}^G w_g = 1$; McLachlan & Peel, 2000).

Der Mittelwertsvektor $\boldsymbol{\mu}_g$ und die Kovarianzmatrix $\boldsymbol{\Sigma}_g$ implizieren ein gruppenspezifisches Mess- und Strukturmodell mit einem eigenen Parametervektor. Dieser Parametervektor kann sich sowohl in der Anzahl seiner Elemente als auch in der Größe der einzelnen Parameter unterscheiden. So könnten in einer latenten Klasse die Indikatoren auf nur einen Faktor zurückzuführen sein, während in einer anderen Klasse mehrere

ist und $E_g(y|x)$ der bedingte Erwartungswert in der latenten Klasse g (vgl. Pek et al., 2011).

Faktoren adäquat sind. Ebenso könnten sich beispielsweise die latenten Klassen nur in der Höhe der Korrelationen zwischen den Faktoren unterscheiden, während das Modell selbst in allen latenten Klassen strukturell dasselbe ist, also dieselben Parameter besitzt. Dies ermöglicht eine flexible Modellierung für die einzelnen latenten Klassen.

Konkret erweitert sich das Messmodell aus Gleichung 1.2 für die in der jeweiligen Gruppe g beobachteten Variablen $\mathbf{y}|g$ und $\mathbf{x}|g$ zu

$$\mathbf{y}|g = \boldsymbol{\tau}_g^y + \boldsymbol{\Lambda}_g^y \eta_g + \boldsymbol{\epsilon}_g \quad (3.5)$$

$$\mathbf{x}|g = \boldsymbol{\tau}_g^x + \boldsymbol{\Lambda}_g^x \boldsymbol{\xi}_g + \boldsymbol{\delta}_g, \quad (3.6)$$

wobei $\boldsymbol{\tau}_g^y, \boldsymbol{\tau}_g^x$ die Vektoren der latenten Interzepte in der Gruppe g darstellen und $\boldsymbol{\Lambda}_g^y, \boldsymbol{\Lambda}_g^x$ die entsprechende Faktorladungsmatrizen; $\boldsymbol{\epsilon}_g$ und $\boldsymbol{\delta}_g$ sind die Vektoren der Residuen mit gruppenspezifischen Kovarianzmatrizen $\boldsymbol{\Theta}_g^\epsilon$ bzw. $\boldsymbol{\Theta}_g^\delta$. $\boldsymbol{\xi}_g$ ist der Vektor der latenten exogenen Variablen mit gruppenspezifischem Mittelwertsvektor $\boldsymbol{\mu}_g^\xi$ und Kovarianzmatrix $\boldsymbol{\Phi}_g^2$. η_g ergibt sich aus einem gruppenspezifischen Strukturmodell (analog zum Strukturmodell in SEM; s. Bollen, 1989) der Form

$$\eta_g = \alpha_g + \boldsymbol{\Gamma}_g \boldsymbol{\xi}_g + \zeta_g, \quad (3.7)$$

mit gruppenspezifischem Interzept α_g , den Effektkoeffizienten zwischen der endogenen und den exogenen Variablen in der Matrix $\boldsymbol{\Gamma}_g$ sowie dem gruppenspezifischen Residuum ζ_g mit Varianz Ψ_g .³ Für die Schätzung eines SEMM wird die Annahme einer multivariaten Normalverteilung von $\boldsymbol{\xi}_g, \boldsymbol{\delta}_g, \boldsymbol{\epsilon}_g$ und ζ_g innerhalb jeder der G latenten Klassen getroffen.

Die Modellierung nichtlinearer, semi-parametrischer Effekte der Prädiktoren auf die endogenen Variablen erfolgt durch die Parameter α_g und $\boldsymbol{\Gamma}_g$, die sich zwischen den latenten Klassen unterscheiden können. Innerhalb jeder latenten Klasse beschreiben diese Parameter zwar einen linearen Zusammenhang zwischen den Prädiktoren und den endogenen Variablen; sie variieren jedoch über die latenten Klassen hinweg, wenn tatsächlich eine Nichtlinearität in den Daten vorliegen sollte (vgl. Abbildung 3.2).

²Die Anzahl v der Prädiktoren ξ_1, \dots, ξ_v kann sich theoretisch in jeder Klasse unterscheiden.

³Die Verallgemeinerung für u endogene Variablen ist in Schrift (B) zu finden. Aus Gründen der Konsistenz mit den vorherigen Kapiteln und der Anschaulichkeit wird hier nur auf den einfachen Fall $u = 1$ eingegangen.

3.3 Interpretation der Klassen und Variablen in SEMM

Ein wichtiger Punkt für die Anwendung der SEMM stellt die Interpretation der latenten Klassen dar. Hier unterscheidet man zwischen einer direkten und einer indirekten Interpretation. Eine direkte Interpretation der Ergebnisse erfolgt dann, wenn die identifizierten latenten Klassen (Mixture Components) als substantiell angesehen werden. Die einzelnen Klassen werden also als tatsächlich existierende Subpopulationen angesehen. Diese Art der Interpretation ist jedoch als kritisch zu betrachten, da die datengetriebene Identifikation von latenten Klassen nur sehr bedingt „echte“ Subpopulationen identifizieren kann (Borsboom, Mellenbergh & van Heerden, 2003; Dolan & van der Maas, 1998; Titterington, Smith & Makov, 1985). In der indirekten Interpretation werden den extrahierten latenten Klassen hingegen keine inhaltliche Bedeutung zugemessen, sondern sie werden ausschließlich als methodisches Werkzeug gesehen, um Verteilungen und Zusammenhänge statistisch zu approximieren (z.B. McLachlan & Peel, 2000). Eine Interpretation des Modells erfolgt entsprechend nur für die Gesamtpopulation.

Trotz der theoretisch sehr hohen Flexibilität des Verfahrens ist es in der Praxis sinnvoll, zumindest bestimmte Strukturaspekte (wie die Anzahl der Prädiktoren) und einen Teil der Parameter über die Gruppen hinweg zu restringieren, um eine Identifikation und Stabilität der Schätzungen der SEMM zu gewährleisten (Bauer & Curran, 2004; Bauer, 2005; Pek et al., 2009, 2011). Um insbesondere eine eindeutige Interpretation der latenten Variablen über die latenten Klassen hinweg zu gewährleisten, ist es notwendig, die Parameter der Messmodelle ($\tau_g^y, \tau_g^x, \Lambda_g^y, \Lambda_g^x, \Theta_g^\delta, \Theta_g^\epsilon$) über die latenten Klassen hinweg zu restringieren (Bauer, 2005; Pek et al., 2011). Nur in diesem Fall ist es sinnvoll, einen variierenden Parameter aus Γ_g als nichtlinearen Effekt zwischen zwei Variablen zu interpretieren.

3.4 Nonlinear Structural Equation Mixture Modeling

Ein Nachteil der semi-parametrischen Definition eines nichtlinearen Effekts in SEMM ist, dass keine Aussage über seine Effektstärke möglich ist. Man kann einen gefundenen nichtlinearen Effekt nicht hinsichtlich seiner empirischen Bedeutsamkeit einschätzen, sondern nur seiner Form nach beschreiben (Bauer, 2005). Ist es jedoch von Interesse, wie groß der Effekt ist, so muss eine Annahme über die Funktion getroffen werden, die der nichtlineare Effekt darstellen soll.

Der NSEMM-Ansatz, der in Schrift (B) entwickelt wurde, führt ein parametrisches

Modell für nichtlineare Effekte für die Mixture-Modelle ein. Aus Sicht der SEMM stellt das Verfahren eine Erweiterung in dem Sinne dar, als dass es eine explizite funktionale Form der Nichtlinearität modelliert, womit eine Größenaussage zu einem nichtlinearen Effekt möglich ist. Aus Sicht der nichtlinearen SEM ist das Verfahren eine Erweiterung, die eine erwartungstreue (ML)-Schätzung der nichtlinearen Effekte ermöglicht, wenn die Indikatoren nicht-normalverteilt sind. Durch die Verwendung der Mischverteilungen können auch sehr allgemeine Verteilungstypen wie z.B. mehrgipflige Verteilungen berücksichtigt werden.

Das Strukturmodell aus Gleichung 3.7 wird für den NSEMM-Ansatz folgendermaßen erweitert (vgl. Gleichung 1.1):

$$\eta_g = \alpha_g + \mathbf{\Gamma}_g \boldsymbol{\xi}_g + \boldsymbol{\xi}_g' \boldsymbol{\Omega}_g \boldsymbol{\xi}_g + \zeta_g, \quad (3.8)$$

wobei die Matrix $\boldsymbol{\Omega}_g$ die Parameter der nichtlinearen Effekte enthält (z.B. ω_{12g} für einen Interaktionseffekt, vgl. Gleichung 1.4). Durch die Verwendung der latenten Produktterme ist die Verteilungsfunktion $f((\mathbf{y}', \mathbf{x}')')$ der Indikatoren keine Mischung von Normalverteilungen wie in Gleichung 3.4, sondern einer Mischung von spezifischen nicht-normalen Verteilungen $f_g((\mathbf{y}', \mathbf{x}')')$ innerhalb jeder latenten Klasse (s. Abschnitt 2.1.2; Klein & Moosbrugger, 2000). Die Verteilung $f((\mathbf{y}', \mathbf{x}')')$ kann innerhalb jeder latenten Klasse durch eine bedingte Mischverteilung approximiert werden und besitzt eine komplexe Funktion für die Mittelwerte und Kovarianzmatrizen der Indikatorvariablen $\mathbf{y}|g$. Die Approximation kann mit verschiedenen numerischen Methoden wie in LMS (Hermite-Gauss), QML (Gauss-Newton) oder Mplus (Trapezregel; Stoer, 1989) erfolgen. Die Parameterschätzung des NSEMM-Ansatzes basiert somit auf der Analyse einer höherdimensionalen Mischverteilung. Mplus (Muthén & Muthén, 1998–2012) ist in der Lage solche Mischverteilungsmodelle zu schätzen, wodurch eine sehr einfach zu handhabende Schätzung des NSEMM-Ansatzes für den Anwender verfügbar ist. Dies erlaubt eine flexible und direkte Verwendung des Verfahrens.

Um eine (parametrische) Interpretation der Effektkoeffizienten in $\mathbf{\Gamma}_g$ und $\boldsymbol{\Omega}_g$ als lineare bzw. nichtlineare Effekte von $\boldsymbol{\xi}_g$ auf η_g treffen zu können, müssen neben der Strukturinvarianz (gleiche Ladungsmuster, gleiche Anzahl latenter Variablen) Constraints über die latenten Klassen hinweg für die Schätzung gesetzt werden. Die Constraints beinhalten neben den Parametern des Messmodells ($\boldsymbol{\tau}_g^y, \boldsymbol{\tau}_g^x, \boldsymbol{\Lambda}_g^y, \boldsymbol{\Lambda}_g^x, \boldsymbol{\Theta}_g^\delta, \boldsymbol{\Theta}_g^\epsilon$) auch die des Strukturmodells ($\alpha_g, \mathbf{\Gamma}_g, \boldsymbol{\Omega}_g, \Psi_g$), die alle invariant über die G latenten Klassen hinweg gesetzt werden müssen. Um eine Approximation der (beliebigen) Verteilung der v latenten Prädiktoren zu ermöglichen, wird der Prädiktorvektor als eine Mischverteilung aus G multivariaten Normalverteilungen mit gruppenspezifischen Mittelwertsvektoren

Tabelle 3.1: Zusammenfassung der Ergebnisse zur Power bei $\omega_{12} = \omega_{11} = \omega_{22} = .2$ und $N = 800$ einer Simulationsstudie zum NSEMM-Ansatz (Ergebnisse entnommen aus Schrift (B)). Als Populations- und Analysemodell diente Gleichung 1.4.

	ω_{12}	ω_{11}	ω_{22}
<i>Schiefe=0, Kurtosis=0</i>			
NSEMM, 2 latente Klassen	66.50%	97.40%	97.80%
NSEMM, 3 latente Klassen	66.00%	97.10%	97.40%
NSEMM, 4 latente Klassen	65.30%	96.90%	97.00%
LMS	69.20%	98.20%	97.90%
Erweiterter Unconstrained	62.60%	96.80%	97.00%
<i>Schiefe=2, Kurtosis=7</i>			
NSEMM, 2 latente Klassen	48.00%	84.50%	87.00%
NSEMM, 3 latente Klassen	47.00%	82.30%	85.10%
NSEMM, 4 latente Klassen	46.50%	81.10%	84.00%
LMS	32.40%•	88.50%•	88.80%•
Erweiterter Unconstrained	48.40%°	83.90%°	85.20%
<i>Schiefe=3, Kurtosis=21</i>			
NSEMM, 2 latente Klassen	40.30%	81.20%	80.70%
NSEMM, 3 latente Klassen	39.80%	81.40%	78.50%
NSEMM, 4 latente Klassen	38.90%	78.30%	77.20%
LMS	27.40%•	88.30%•	88.70%•
Erweiterter Unconstrained	48.70%°	87.10%°	86.70%°

Anmerkung: • – absoluter Parameterbias $\geq 10\%$, ° – $SE/SD \leq .90$

μ_g^ξ und Kovarianzmatrizen Φ_g modelliert.

$$\xi \sim \sum_{g=1}^G w_g N(\mu_g^\xi, \Phi_g). \quad (3.9)$$

Ist $G = 1$ so resultiert in Mplus ein zu LMS äquivalenter Schätzer. Für jede weitere latente Gruppe erhöht sich die Anzahl der zu schätzenden Parameter um v latente Mittelwerte und $v(v+1)/2$ (Ko)-Varianzen. Mit steigender Anzahl von latenten Klassen wird der Ansatz daher weniger sparsam.

3.5 Diskussion des NSEMM-Ansatzes

Der NSEMM-Ansatz erlaubt eine effiziente Schätzung von latenten nichtlinearen Effekten bei nicht-normalverteilten Indikatoren. Er verbindet somit die Vorteile von parametrischen nichtlinearen SEM und semi-parametrischen Mischverteilungsmodellen.

Während die Schätzung der meisten Verfahren im Rahmen der nichtlinearen SEM auf einer multivariaten Normalverteilung beruhen, kann der NSEMM-Ansatz diese Annahme abschwächen. Gleichzeitig ist es aber im Gegensatz zu den SEMM möglich, eine Größenaussage zu den nichtlinearen Effekten zu treffen.

In einer Simulationsstudie, die in Schrift (B) zu finden ist, konnte gezeigt werden, dass die theoretischen Vorteile auch unter empirisch relevanten Bedingungen zu Unterschieden zwischen dem NSEMM-Ansatz, LMS und dem erweiterten Unconstrained-Ansatz führen. In Tabelle 3.1 sind einige wichtige Ergebnisse dieser Simulationsstudie zusammengefasst. Wie man erkennt, liefert LMS die effizientesten Schätzungen, solange die Indikatoren normalverteilt sind, was auf die geringere Komplexität des Verfahrens (weniger Parameter) zurückzuführen ist. In diesem Fall ist der NSEMM-Ansatz überparametrisiert, da eine Mischverteilung zur Approximation der Verteilung der latenten Prädiktoren nicht notwendig ist. Bei Nichtnormalität sind jedoch nur für den NSEMM-Ansatz sowohl die Schätzungen der Parameter als auch der Standardfehler erwartungstreu, während der erweiterte Unconstrained-Ansatz unterschätzte Standardfehler (mit einem Verhältnis der mittleren Standardfehler zu der Monte-Carlo-Standardabweichung von $SE/SD \leq .90$) produziert und LMS einen Parameter-Bias von mehr als 10% aufweist. Es zeigt sich zudem, dass die Verwendung von drei bis vier latenten Klassen beim NSEMM-Ansatz zu guten Schätzungen führt und dass die Power zur Aufdeckung der nichtlinearen Effekte nicht zu sehr mit der Anzahl der Klassen abnimmt.

Eine wichtige Frage, die im Rahmen des NSEMM-Ansatzes stets beantwortet werden muss, ist die nach der Anzahl der latenten Klassen. Bauer (2005) empfiehlt drei bis fünf latente Klassen, was auch mit den Erfahrungen aus den in Schrift (B) berichteten Simulationsstudien übereinstimmt. In der Praxis hat sich der BIC als Entscheidungskriterium etabliert (Nylund, Asparouhov & Muthén, 2007). Die Entscheidung muss aber stets unter Berücksichtigung der Stichprobengröße erfolgen, da der NSEMM-Ansatz ein komplexes Verfahren darstellt, der bei kleinen Stichproben nur bedingt einsatzfähig ist. Auch sollte stets darauf Acht gegeben werden, ein Modell nicht zu überparametrisieren, sondern ein möglichst sparsames Modell zu verwenden.

Kapitel 4

Implementierung und praktische Anwendung eines Heterogenen Wachstumskurvenmodells

In diesem Kapitel wird eine Implementierung des Heterogenen Wachstumskurvenmodells (HGM) von Klein und Muthén (2006) in Mplus vorgestellt. Hierzu wird in einem ersten Schritt das Latente Wachstumskurvenmodell dargestellt und die Verwendung von Interaktionseffekten in diesem Modell kurz diskutiert. Anschließend wird das HGM und seine Implementierung beschrieben und seine Anwendung an einem empirischen Datensatz aus der AIDS-Forschung demonstriert.

Anmerkung: Dieses Kapitel basiert zu großen Teilen auf folgendem Artikel, der in Anhang E zu finden ist:

Brandt, H., Klein, A. G. & Gäde, J. (*zur Veröffentlichung eingereicht*). Implementation and practical application of a heterogeneous growth curve model. *Multivariate Behavioral Research*.

4.1 Latente Wachstumskurvenmodelle

Im Rahmen der Analyse von Längsschnittdaten, hat sich als eine wichtige Methode das Latente Wachstumskurvenmodell etabliert (Latent Growth Curve Model, LGM), das Wachstumsverläufe anhand des SEM-Frameworks modelliert. Eine Einführung zum LGM ist in Bollen und Curran (2006), Duncan et al. (2006) oder Reinecke (2012) zu finden. In einem LGM werden die individuellen Wachstumskurven durch ein individuelles Interzept (η_{0i}) und einen individuellen Slope (η_{1i}) für jede Person i dargestellt. Die durchschnittlichen Wachstumsprozesse (einer Stichprobe bzw. einer Population) werden entsprechend durch die Mittelwerte und Varianzen der latenten Variablen η_0

(Interzept-Faktor) und η_1 (Slope-Faktor) charakterisiert. Das Messmodell für eine beobachtete Variable y zum Messzeitpunkt t ist gegeben durch:

$$y_t = \eta_0 + \lambda_t \eta_1 + \epsilon_t, \quad (4.1)$$

wobei λ_t einen sogenannten „Time Score“ und ϵ_t das Residuum von y_t darstellen. Die Time Scores modellieren den Verlauf des Wachstumsprozesses. Sie können einen spezifischen Wachstumsverlauf modellieren, indem sie auf entsprechende Werte a priori fixiert werden, z.B. für einen linearen Verlauf $\lambda_1 = 0, \lambda_2 = 1, \lambda_3 = 2$, etc. Oder sie können zum Teil frei geschätzt werden und somit einen datengetriebenen Verlauf modellieren, z.B. indem λ_3 frei geschätzt wird. Für die Identifikation eines LGM ist es notwendig, zwei Faktorladungen zu fixieren (zumeist $\lambda_1 = 0, \lambda_2 = 1$). Die Residuen ϵ_t werden als untereinander und mit η_0 unkorreliert angenommen; sie sind normalverteilt mit Mittelwert null und Varianz θ_{tt} .

Die Zusammenhänge des Interzept- und Slope-Faktors mit einer Kovariaten w können modelliert werden als

$$\begin{aligned} \eta_0 &= \beta_{00} + \beta_{02}w + \zeta_0 \\ \eta_1 &= \beta_{10} + \beta_{11}\eta_0 + \beta_{12}w + \zeta_1, \end{aligned} \quad (4.2)$$

wobei β_{00}, β_{10} die Interzepte von η_0 und η_1 sind, β_{11} das Einflussgewicht von η_0 auf η_1 und β_{02} bzw. β_{12} die Einflussgewichte von w auf η_0 bzw. η_1 . ζ_0, ζ_1 sind normalverteilte Residuen mit Mittelwert null und Varianzen ψ_{00} und ψ_{11} . Die Kovariate w soll hier eine beobachtete Baseline-Kovariate sein, die zum ersten Messzeitpunkt erhoben wurde ($t = 1$) und zeitlich invariant ist¹.

Von Bedeutung für die Analyse von Längsschnittdaten sind drei Aspekte:

1. Die Mittelwerte der Faktoren geben Auskunft über den durchschnittlichen Entwicklungsprozess einer Population.
2. Die Kovarianzen bzw. die Einflussgewichte in Gleichung 4.2 dienen dazu, Entwicklungsprozesse näher zu erklären und Kovariaten zu identifizieren, die die Entwicklungsverläufe vorhersagen können.
3. Eine individuelle Prognose basierend auf den individuellen Startbedingungen kann anhand der Schätzungen des Modells erfolgen. Die individuellen Konfidenzintervalle zur Prognose hängen hierbei insbesondere von der Varianz des Slope-Faktors η_1 ab.

¹Eine Verwendung von latenten Kovariaten oder mehreren Kovariaten ist durch direkte Erweiterung des obigen Modells möglich.

4.2 Interaktionseffekte in LGM

Die Verwendung von Interaktionseffekten zwischen dem Interzept η_0 und einer Kovariaten w zur Vorhersage des Slopes η_1 kann als eine direkte Anwendung der Methoden für nichtlineare Effekte in SEM modelliert werden. Hierbei erweitert sich Gleichung 4.2 für den Slope-Faktor um einen (latenten) Produktterm $\eta_0 w$, der den differentiellen Entwicklungsverlauf modelliert:

$$\eta_1 = \beta_{10} + \beta_{11}\eta_0 + \beta_{12}w + \beta_{13}\eta_0 w + \zeta_1 \quad (4.3)$$

mit dem nichtlinearen Effekt β_{13} . Khoo (2001) sowie Muthén und Curran (1997) diskutieren Interaktionseffekte für kategoriale Kovariaten w , wie sie insbesondere in Treatmentstudien von Bedeutung sind. Sie empfehlen hierfür insbesondere Mehrgruppenmodelle zu verwenden (s. auch Bollen, 1989; Bollen & Curran, 2006). Ist die Kovariate jedoch kontinuierlich, so kann eine Operationalisierung des nichtlinearen Modells in Gleichung 4.3 beispielsweise durch LMS oder den NSEMM-Ansatz (s. Kapitel 2 oder 3) erfolgen (für empirische Beispiele mit LMS s. Goodnight et al., 2007; Larsen et al., 2010). Alle Schlüsse zu den Vorteilen und den Nachteilen dieser Ansätze sowie zu ihren Voraussetzungen lassen sich direkt auf die LGM mit Interaktionseffekten übertragen. Eine Schätzung anhand des 2SMM-Ansatzes ist ebenso möglich, obwohl eine komplexe Erweiterung des Ansatzes für LGM notwendig ist. Obwohl eine Schätzung mit QML theoretisch erfolgen kann, ist eine praktische Anwendung für das LGM derzeit nur bedingt möglich, da QML nur in einer Standalone-Software implementiert ist. Die Verwendung von PI-Ansätzen stellt sich jedoch für dieses Modell als schwierig dar, da die Indikatoren im LGM sowohl für den Prädiktor η_0 als auch zum Teil für die endogene Variable η_1 verwendet werden. Die Bildung der Produktindikatoren kann somit nur einen Teil der Information berücksichtigen (y_1). Die Interaktion von zwei Kovariaten zur Vorhersage der Wachstumsfaktoren anhand eines PI-Ansatzes stellen Li, Duncan und Acock (2000) vor.

Ein nichtlineares Modell, das einen spezifischen Aspekt von Längsschnittdaten modelliert, ist das Heterogene Wachstumskurvenmodell (Heterogeneous Growth Curve Model, HGM; Klein & Muthén, 2006). In ihm wird eine systematische Veränderung der Varianz des Slope-Faktors in Abhängigkeit des Interzept-Faktors und/oder einer Kovariaten modelliert. Dies ermöglicht eine modellimplizierte variable Prognosegenauigkeit. In den folgenden Abschnitten wird dieses Modell und seine Implementierung in Mplus vorgestellt (Brandt et al., submitted). Um den Einsatzbereich des Modells zu verdeutlichen und zu motivieren, sollen heterogene Wachstumsprozesse jedoch zuerst charakterisiert werden.

4.3 Heterogene Wachstumsprozesse

Parrila et al. (2005) untersuchten, wie sich die Leseleistung von Englisch- und Finnischsprachigen Kindern im Laufe der ersten Schuljahre entwickelt. Es zeigte sich insgesamt ein negativer Zusammenhang zwischen den individuellen Startwerten der Schüler und den Slopes. Je besser ein Kind anfangs lesen konnte, desto geringer veränderte sich diese Leseleistung im Laufe der ersten Schuljahre. Es wurde eine vergleichsweise kleine Varianz der Entwicklungsverläufe für Kinder mit einer hohen anfänglichen Leseleistung beobachtet, während die für Kinder mit mittlerer oder geringer anfänglicher Leseleistung eine große Varianz besaßen. Eine Prognose der Leseleistung am Ende der Grundschule basierend auf den anfänglichen Leseleistungen muss folglich umso ungenauer sein, je geringer die anfänglichen Leseleistungen ausgeprägt sind.

Diese Heterogenität der Varianzen der Entwicklungsverläufe impliziert eine Heterogenität der Varianz des Slope-Faktors. Sie ist in empirischen Studien häufig zu beobachten, gerade auch dann, wenn Boden- oder Deckeneffekte in den Messungen auftreten. Wird diese Heterogenität nicht im Modell berücksichtigt, d.h. wird eine homogene Varianz für den Slope-Faktor angenommen, so liegt eine Fehlspezifikation vor. Als Konsequenz sind Konfidenzintervalle für individuelle Prognosen nicht mehr akkurat, denn sie beruhen auf der fehlspezifizierten Varianz des Slope-Faktors. Im angeführten Beispiel wären die Konfidenzintervalle für eine Prognose der Leseleistung im vierten Schuljahr basierend auf den Leistungen im ersten Schuljahr für Kinder mit hohen anfänglichen Leseleistungen zu weit und für Kinder mit geringen Leseleistungen zu schmal. In einem Standard-LGM (mit einer Gruppe) wird stets implizit eine Homogenität angenommen ($Var(\zeta_i) = \psi_{11}$ für alle Personen). Daher können Aussagen basierend auf einem LGM in manchen Fällen zu falschen oder ungenauen Schlussfolgerungen führen.

Es existieren zwei Verfahren, die eine Heterogenität des Slope-Faktors berücksichtigen: Das Growth Curve Mixture Model (GMM; Meredith & Tisak, 1990; Muthén, 2001, 2004) und das Heterogene Wachstumskurvenmodell (HGM; Klein & Muthén, 2006). Das GMM ist ein Spezialfall der Mischverteilungsmodelle, wie sie im vorangegangenen Kapitel 3 vorgestellt wurden. Bei einem GMM werden separate LGM in latenten, diskreten Klassen geschätzt, die sich in ihren spezifischen Wachstumsprozessen (Mittelwert des Interzepts, Korrelation zwischen Interzept und Slope, etc.) unterscheiden. Wichtig bei den GMM ist zum einen die Identifikation der latenten Klassen und zum anderen ihre Interpretation. Die damit verbundenen Schwierigkeiten einer direkten und indirekten Interpretation der latenten Klassen existieren ebenso wie bei den Mischverteilungsmodellen (s. Abschnitt 3.3).

Möchte man keine Annahme über diskrete Klassen treffen, sondern eine feinstufigere Heterogenität modellieren, so eignet sich das HGM. In ihm wird eine Abhängigkeit zwischen den Anfangsbedingungen und der Varianz des Slope-Faktors modelliert, die

eine Berücksichtigung einer Heterogenität auf Individualebene erlaubt. Die Varianz des Slope-Faktors verändert sich kontinuierlich für Personen mit verschiedenen Anfangsbedingungen. Damit stellt das HGM eine Alternative zum GMM dar, wenn die Annahme von diskreten Klassen unplausibel oder nicht adäquat ist. Ein Vorteil des HGM ist seine Sparsamkeit; im einfachsten Fall (ohne Kovariate) besitzt das Modell nur zwei Parameter mehr als ein Standard-LGM mit einer Gruppe.

Klein und Muthén (2006) stellten für das HGM eine Standalone-Software bereit, die für ein spezielles Modell (ohne Kovariate) in Delphi Pascal programmiert war. Die in den folgenden Abschnitten vorgestellte Implementation des Modells in Mplus erlaubt eine einfache Anwendung des Modells, die ebenfalls eine Verwendung von Kovariaten für die Modellierung heterogener Varianzen erlaubt.

4.4 Das HGM nach Klein und Muthén (2006)

Für das HGM wird das in Gleichungen 4.1 und 4.2 formulierte Modell erweitert, indem das Residuum ζ_1 des Slope-Faktors η_1 in zwei Varianzkomponenten aufgeteilt wird, einen heterogenen Varianzanteil (repräsentiert durch ζ_2) und einen homogenen Varianzanteil (repräsentiert durch ζ_3):

$$\zeta_1 = (\gamma_0 + \gamma_1\eta_0 + \gamma_2w)\zeta_2 + \zeta_3. \quad (4.4)$$

Die Parameter γ_0 , γ_1 und γ_2 spezifizieren die Größe der Heterogenität und geben zudem Aufschluss über die Variable, auf die die Heterogenität zurückzuführen ist (γ_1 , γ_2). Zur Identifikation des HGM muss die Varianz von ζ_2 auf eins fixiert werden ($\psi_{22} = 1$). Das Modell ist zudem nur dann identifiziert, wenn γ_0 und einer der beiden Parameter γ_1 oder γ_2 von null verschieden sind. Gleichung 4.4 beinhaltet Produktterme von η_0 , w und ζ_2 , die das HGM zu einem Nicht-Standardmodell machen. Im Unterschied zu den bisher vorgestellten Modellen mit latenten Produkttermen in Kapitel 2 und 3 besitzt ζ_2 keine Indikatoren (im Gegensatz zu ξ_1 oder ξ_2 in Gleichung 1.4, vgl. Abbildung 2.1).

Ein Pfaddiagramm für ein HGM, spezifiziert für vier Messzeitpunkte und einen linearen Slope-Faktor, ist in Abbildung 4.1 dargestellt. Aufgrund der Produktterme in Gleichung 4.4 folgt η_1 nicht mehr einer Normalverteilung (Aroian, 1944, vgl. Abschnitt 2.1.2). Jedoch ist die bedingte Verteilung von η_1 gegeben η_0 , w normalverteilt mit einer bedingten Varianz von

$$Var(\eta_1|\eta_0, w) = (\gamma_0 + \gamma_1\eta_0 + \gamma_2w)^2 + \psi_{33} \quad (4.5)$$

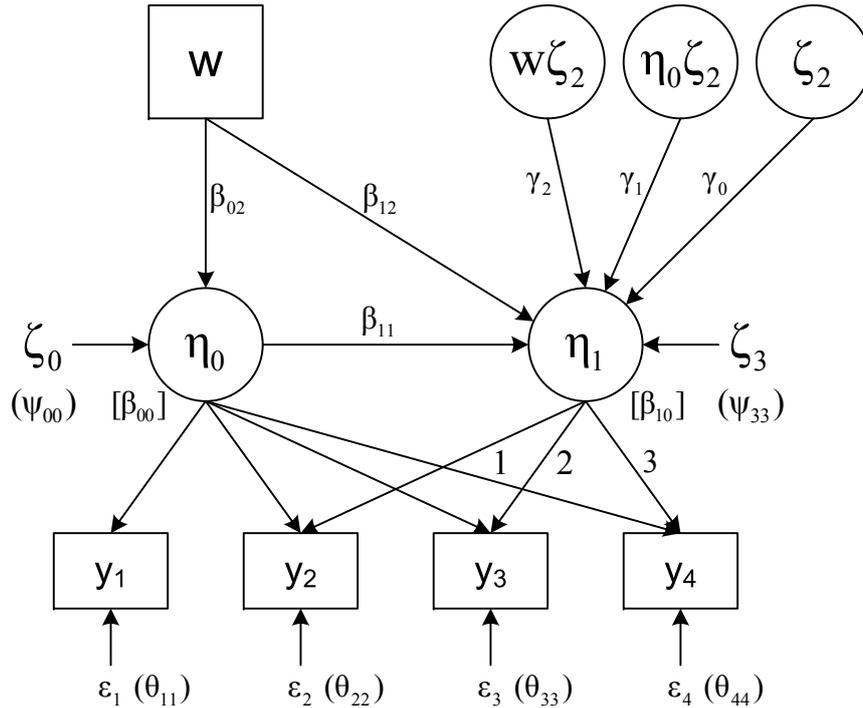


Abbildung 4.1: Heterogenes Wachstumskurvenmodell (HGM) mit vier Messzeitpunkten und einer Kovariate w .

und einem bedingten Erwartungswert von

$$E(\eta_1|\eta_0, w) = \beta_{10} + \beta_{11}\eta_0 + \beta_{12}w. \quad (4.6)$$

Für die Schätzung des HGM wird diese Eigenschaft der bedingten Normalverteilung von η_1 ausgenutzt. Klein und Muthén (2006) verwendeten hierfür einen Quasi-ML-Schätzer ähnlich wie in QML (Klein & Muthén, 2007), der die nicht-normale Verteilungsfunktion durch ein Produkt einer normalen Verteilungsfunktion und einer bedingt normalen Verteilungsfunktion approximiert. Im nächsten Abschnitt wird im Gegensatz dazu eine direkte numerische Approximation der nicht-normalen Verteilungsfunktion im Sinne von LMS (s. Abschnitt 2.1.2; Klein & Moosbrugger, 2000) angewandt. Diese Implementierung besitzt den Vorteil, dass sie ML-Schätzungen liefert und dass sie leicht umsetzbar ist, da sie in Mplus (Muthén & Muthén, 1998–2012) spezifiziert werden kann.

4.5 Implementierung des HGM in Mplus

Die Implementierung des HGM in Mplus erfolgt anhand der Option zur Modellierung von latenten Produkttermen. Das Besondere am HGM im Vergleich zu anderen Modellen, die latente Produktterme enthalten, ist das Fehlen von Indikatoren für die latente

Variable ζ_2 , die Teil der Produktterme ist. Trotzdem ist die Statistik-Software so flexibel, dass eine Spezifikation möglich ist. Der Syntax-Code für ein Modell, das durch die Gleichungen 4.1, 4.2 und 4.4 für vier Messzeitpunkte und einen linearen Wachstumsprozess spezifiziert wird (vgl. Abbildung 4.1), ist gegeben durch:

```
(1) ANALYSIS: TYPE=RANDOM;
(2)          ALGORITHM=INTEGRATION;
              ! Messmodell für einen linearen Slope
(3) MODEL:   eta0 eta1 | y1@0 y2@1 y3@2 y4@3;
              ! Strukturmodell
(4)          eta1 ON eta0 w;
(5)          eta0 ON w;
              ! zeta2
(6)          zeta2@1; [zeta2@0];
(7)          zeta2 WITH eta0@0;
(8)          zeta2 WITH w@0;
(9)          zeta2 BY eta1;           ! gamma0
              ! Spezifikation der heterogenen Komponenten
(10)         int1 | eta0 XWITH zeta2; ! Produktterm eta0*zeta2
(11)         int2 | w XWITH zeta2;   ! Produktterm w*zeta2
(12)         eta1 ON int1 int2;      ! gamma1 und gamma2
```

In den Zeilen (3) bis (5) wird ein Standard-LGM spezifiziert. In den Zeilen (6) bis (9) wird anschließend ζ_2 als latente Variable mit Mittelwert null und Varianz eins implementiert, die unkorreliert mit den exogenen Variablen im Modell ist und auf η_1 lädt. Durch den XWITH-Befehl in den Zeilen (10) und (11) wird in Kombination mit den Zeilen (1) und (2) die Option für die Spezifikation von Produkttermen aufgerufen. Zeile (12) implementiert zuletzt die Koeffizienten γ_1 und γ_2 . Weitere Kovariaten können in das Modell integriert werden, indem entsprechend die Zeilen (4), (5), (8), (11) und (12) erweitert werden.

Eine technische Schwierigkeit für die Schätzung des HGM in Mplus ist die Wahl der Startwerte. Die Verwendung der automatischen Startwerte in Mplus resultiert in der Praxis häufig in nicht-konvergierten Lösungen, weshalb es sinnvoll erscheint, individuelle Startwerte für jeden Datensatz zu verwenden. Hierfür können die Schätzungen für die Parameter aus einem zuvor analysierten Standard-LGM, wie es anhand der Zeilen (3) bis (5) spezifiziert würde, als Startwerte für das HGM verwendet werden.

Für die Schätzung des Modells in Mplus wird angenommen, dass sich die exogenen Variablen multivariat normalverteilen. Für das oben spezifizierte Modell beinhaltet dies die Variablen $\zeta_0, \zeta_2, \zeta_3$ und die Messfehlerresiduen $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4$. Diese Annahme ist

in empirischen Situationen nicht immer erfüllt und kann zu einem gewissen Parameterbias führen (s. die ausführliche Diskussion dieses Aspekts für verteilungsanalytische Verfahren in Kapitel 2). In Längsschnittstudien kann manchmal eine Nichtnormalität des Indikators zum ersten Messzeitpunkt zu beobachten sein und somit eine Nichtnormalität des Interzept-Faktors η_0 . Dies kann z.B. dann auftreten, wenn Fähigkeiten gemessen werden, die bei den meisten Probanden nur sehr gering zum ersten Messzeitpunkt vorhanden sind, beispielsweise die Lesefähigkeiten von Schülern zu Beginn der ersten Klasse. In einer Simulationsstudie (s. Schrift (C) in Anhang E) konnte jedoch gezeigt werden, dass die Implementierung des HGM zuverlässige Schätzungen der Parameter der heterogenen Varianzkomponente erlaubt, auch wenn die Daten moderat nicht-normalverteilt sind.

4.6 Empirisches Beispiel

In diesem Beispiel zur Entwicklung der T-Helferzellen bei HIV-Patienten soll die Anwendung des HGM und der Unterschied zu einem Standard-LGM demonstriert werden. T-Helferzellen sind zentral für die Funktionsweise des menschlichen Immunsystems. Anhand ihrer Konzentration kann das Fortschreiten einer AIDS-Erkrankung beschrieben werden (Zeger & Diggle, 1994). Zudem wird vermutet, dass Medikamente, die die T-Helferzellen anvisieren, auch eine Verbesserung des allgemeinen Gesundheitszustandes bei HIV-Patienten erzielen (Burzykowski, Molenberghs & Buyse, 2006). Damit stellt die Konzentration der T-Helferzellen auch ein „Surrogat“ für den Entwicklungsverlauf der Krankheit dar (Fleming & DeMets, 1996). Als Surrogat wird in Therapiestudien ein Indikator für den langfristigen Therapieerfolg bezeichnet. Surrogate sind daher von besonderem Interesse bei der Behandlung von langfristigen Erkrankungen wie AIDS, Krebs oder auch Herzkreislauferkrankungen.

Die Analyse des Datensatzes mit dem HGM soll von zwei Fragen geleitet werden. Erstens: Gibt es Personen für die ein sehr ähnlicher Krankheitsverlauf vorliegt? Diese Information kann für Erststudien zur Wirksamkeitstestung neuer Medikamente relevant sein, denn eine Veränderung des Verlaufs der T-Helferzellenkonzentration bei Personen, deren Verlauf üblicherweise sehr konsistent ist, kann auf das Medikament zurückgeführt werden. Zweitens: Muss bei der Vorhersage einer langfristigen T-Helferzellenkonzentration basierend auf der anfänglichen Konzentration eine Heterogenität der Prognosegenauigkeit berücksichtigt werden? Diese Information ist relevant für individuelle Verlaufsprognosen.

Der Datensatz entstammt einer randomisierten, kontrollierten Doppelblindstudie zur Effektivität verschiedener AIDS-Medikamente (Henry et al., 1998; Daten aus Fitzmaurice, Laird & Ware, 2004 entnommen). Für die Analyse wurden vier Wiederholungs-

Tabelle 4.1: Deskriptive Statistiken für die vier Messungen der T-Helferzellen.

	Mittelwert	SD	Schiefe	Kurtosis	Korrelationen		
					y_2	y_3	y_4
y_1	.000	2.751	1.330	5.616	.792	.773	.633
y_2	.042	2.847	1.429	6.181		.838	.786
y_3	-.595	2.462	1.360	5.803			.832
y_4	-.420	2.708	1.389	5.383			

messungen der T-Helferzellen in Intervallen von ca. 8 Wochen ausgewählt. Es wurden insgesamt $N = 821$ Personen in die Auswertung aufgenommen, für die Daten zu mindestens einem der Messzeitpunkte verfügbar waren. Fehlende Daten wurden als zufällig fehlend angenommen (Missing at Random)² und wurden bei der Schätzung des HGM und des LGM durch die FIML-Option („Full Information Maximum Likelihood“) in Mplus berücksichtigt. Um die starke Nichtnormalität der Verteilung der T-Helferzellen zu reduzieren, wurde eine Wurzeltransformation auf die Daten angewandt. Anschließend wurde für alle Messzeitpunkte der mittlere T-Helferzellenkonzentration des ersten Messzeitpunkts subtrahiert.

In Tabelle 4.1 sind die deskriptiven Statistiken für die T-Helferzellen zu den vier Messzeitpunkten (y_1, y_2, y_3, y_4) dargestellt. Der Verlauf der Mittelwerte zeigt, dass sich die Konzentration der T-Helferzellen im Laufe der Messung im Mittel verschlechterten (von .000 bis $-.420$). Auch nach der Transformation waren die Daten leicht nicht-normalverteilt mit einer Schiefe zwischen 1.330 und 1.429. Die Korrelationen zwischen den Messzeitpunkten waren sich ähnlich und lagen zwischen .633 und .838.

In Abbildung 4.2 ist ein Streudiagramm für die T-Helferzellenkonzentration zu Messzeitpunkt 1 und 4 dargestellt. Wie man erkennen kann, ist eine systematische Heterogenität in den Daten vorhanden: Für Personen mit einer niedrigen Konzentration der T-Helferzellen zum ersten Messzeitpunkt (y_1) ist die Streuung der Konzentration zum vierten Messzeitpunkt (y_4) gering. Die Streuung vergrößert sich jedoch mit steigender anfänglicher Konzentration. Bei Personen mit anfänglichen Werten über $y_1 = 3$ existieren einige Patienten, die sich deutlich verschlechtern, und einige, die eine gleichbleibend hohe Konzentration zum vierten Messzeitpunkt aufweisen.

Die Ergebnisse der Auswertung mit dem HGM und dem LGM sind in Tabelle 4.2 präsentiert. Da die Varianz von ζ_3 nahe null und nicht signifikant war, wurde sie für das finale Modell auf null restringiert. Für beide Modelle wurde ein linearer Verlauf angenommen. Die Modellgüte des LGM ($\chi_{SB}^2 = 5.802$, $df = 5$, $p = .326$) lässt an-

²Für die Adäquatheit der Annahme von Missing at Random für diesen Datensatz s. Vallejo, Fernández, Livacic-Rojas und Tuero-Herrero (2011).

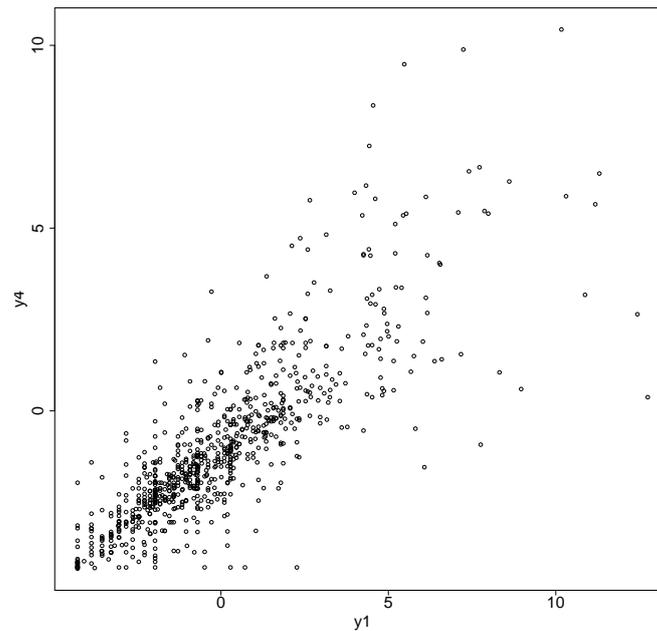


Abbildung 4.2: Streudiagramm für die Konzentration der T-Helferzellen zum ersten Messzeitpunkt (y_1) und zum vierten Messzeitpunkt (y_4).

nehmen, dass eine gute Modell-Datenpassung vorliegt und insbesondere der lineare Verlauf adäquat ist.³ Alle Parameterschätzungen waren signifikant (bei einer Irrtumswahrscheinlichkeit von $\alpha = .05$) bis auf den Mittelwert β_{00} des Interzept-Faktors für das HGM.

Die Schätzungen für das Messmodell und für die linearen Zusammenhänge sind in beiden Modellen sehr ähnlich. Die Reliabilitäten der Indikatoren lagen zwischen .805 und .912. Der Interzept-Faktor hatte einen Mittelwert nahe null ($\hat{\beta}_{00} = .193$ für das HGM und $\hat{\beta}_{00} = .208$ für das LGM) und eine Varianz von etwa sieben ($\hat{\psi}_{00} = 7.028$ für das HGM und $\hat{\psi}_{00} = 7.072$ für das LGM). Der Einfluss des Interzept-Faktors auf den Slope-Faktor war signifikant und negativ ($\hat{\beta}_{11} = -.100$ für das HGM und $\hat{\beta}_{11} = -.081$ für das LGM). Die Standardfehler waren für die Schätzungen des HGM im Vergleich zum LGM wie erwartet zum Teil leicht größer, da für das HGM ein Parameter mehr geschätzt werden musste.

Die zusätzlichen Parameter des HGM, γ_0 und γ_1 , waren signifikant mit Schätzungen $\hat{\gamma}_0 = .356$ und $\hat{\gamma}_1 = .147$. Dies weist auf eine Heterogenität des Slope-Faktors hin. Die bedingte Varianz des Slope-Faktors ergibt sich als $Var(\eta_1|\eta_0) = (.356 + .147\eta_0)^2$ (s. Gleichung 4.5); die Varianz des Slope-Faktors verändert sich quadratisch in Abhängigkeit von η_0 . Für Personen mit einer anfänglichen T-Helferzellenkonzentration von 2 SD

³Für das HGM kann keine Modellgüte angegeben werden, da das hierfür nötige saturierte Modell nicht spezifiziert werden kann. Dies stellt eine allgemeine Problematik in nichtlinearen SEM dar.

Tabelle 4.2: Ergebnisse der Parameterschätzungen und Standardfehler (in Klammern) für das HGM und das LGM (Parameterbezeichnungen wie in Abbildung 4.1).

	HGM		LGM	
γ_0	.356	(.079)	–	
γ_1	.147	(.022)	–	
β_{00}	.193	(.102)	.208	(.103)
β_{10}	–.329	(.027)	–.324	(.028)
β_{11}	–.100	(.018)	–.081	(.027)
ψ_{00}	7.028	(.618)	7.072	(.698)
ψ_{33}	.000		.222	(.070)
θ_{11}	1.079	(.293)	1.397	(.346)
θ_{22}	1.518	(.310)	1.491	(.306)
θ_{33}	.523	(.096)	.563	(.138)
θ_{44}	.798	(.191)	.874	(.282)

unterhalb der mittleren T-Helferzellenkonzentration betrug die bedingte Varianz des Slope-Faktors $Var(\eta_1|\eta_0 = -5.109) = .156$; für Personen mit einem anfänglichen Wert von 2 SD oberhalb der mittleren T-Helferzellenkonzentration ergab sich eine bedingte Varianz von $Var(\eta_1|\eta_0 = 5.495) = 1.354$. Im LGM betrug hingegen die bedingte Varianz des Slope-Faktors $\hat{\psi}_{33} = .222$ für alle Personen.

Die Konfidenzintervalle für eine Vorhersage der T-Helferzellen zum vierten Messzeitpunkt (\hat{y}_4) durch die T-Helferzellen zum ersten Messzeitpunkt (y_1) zeigen die Konsequenzen für die Prognose anhand des HGM im Vergleich zum LGM (s. Abbildung 4.3; die verwendeten Gleichungen für die Konfidenzintervalle sind in Anhang B zu finden). Während das LGM eine Vorhersagegenauigkeit modelliert, die für alle Personen gleich ist (gepunktete Linie), verändert sich die Konfidenzintervallbreite für das HGM in Abhängigkeit von y_1 (durchgezogene Linie). Für Personen nahe dem Mittelwert von y_1 ist eine hohe Prognosegenauigkeit möglich, die sogar genauer ist, als durch das LGM. Diese Personen entwickeln sich sehr konsistent. Mit steigender anfänglicher Konzentration der T-Helferzellen vergrößern sich die Konfidenzintervalle und Personen entwickeln sich inkonsistent. Das HGM kann somit die beobachtete Heterogenität modellbasiert in die Prognose aufnehmen und führt damit hier zu einer adäquateren Schätzung der Konfidenzintervalle als das LGM.

4.7 Zusammenfassung und Diskussion zum HGM

In diesem Kapitel wurde die Implementierung des HGM vorgestellt und seine Anwendung an einem empirischen Datensatz demonstriert. Das HGM modelliert heteroge-

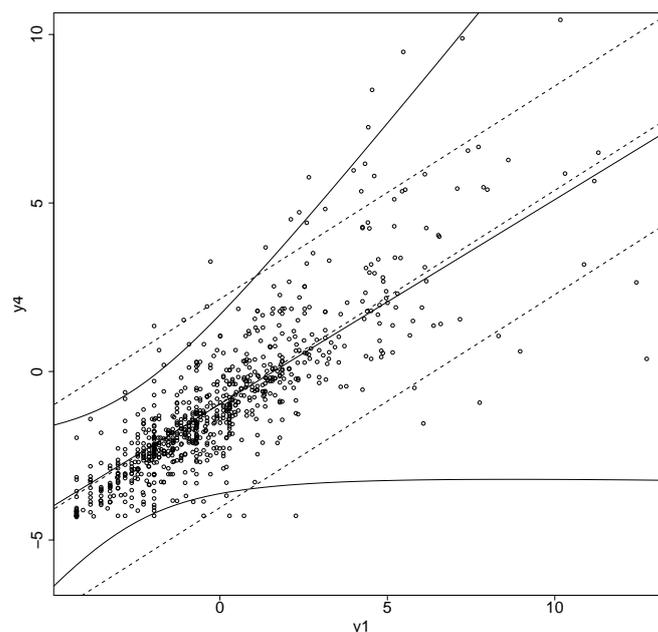


Abbildung 4.3: Konfidenzintervalle für die Prognose der T-Helferzellenkonzentration zum vierten Messzeitpunkt (\hat{y}_4) gegeben der Konzentration zum ersten Messzeitpunkt (y_1) für die Schätzungen des HGM (durchgezogene Linie) und des LGM (gepunktete Linie).

ne Wachstumsprozesse, wie sie in der Empirie häufig vorkommen. Im Gegensatz zum GMM ist hierbei eine feinstufigere Modellierung der Heterogenität möglich. In dem empirischen Beispiel, das den Entwicklungsverlauf der T-Helferzellen für HIV-Patienten untersuchte, zeigte sich ein deutlich heterogenes Wachstum: Während Patienten mit niedriger Anfangskonzentration sehr konsistent auf einem gleichbleibend niedrigen Niveau blieben, veränderte sich die Konzentration bei Patienten mit einer hohen Anfangskonzentration sehr verschieden und diese Patienten wiesen bei der letzten Messung zum Teil sehr hohe und zum Teil sehr niedrige Konzentrationen der T-Helferzellen auf.

Das HGM kann zur Identifikation von Substichproben mit konsistenten Entwicklungsverläufen verwendet werden und liefert eine verbesserte Schätzung der Konfidenzintervalle im Vergleich zum LGM, wenn die Entwicklungsverläufe heterogen sind. Im Gegensatz zum GMM ist die Spezifikation des HGM sparsamer und sollte aus diesem Grund kleinere Konfidenzintervalle erzeugen, wodurch auch eine genauere Prognose als durch das GMM möglich ist. Die Implementierung des HGM in Mplus erweitert den Einsatzbereich des Modells und die Anwenderfreundlichkeit deutlich. Neben einer einfachen Spezifikation des Modells ist auch eine Kombination mit weiteren Optionen in Mplus möglich, z.B. für den Umgang mit Missing Data (Schafer, 1997) oder individuellen Time Scores (Mehta & West, 2000).

Eine Grenze des HGM ist die Voraussetzung der multivariaten Normalverteilung der exogenen Variablen. Wie in einer Simulationsstudie in Schrift (C) berichtet wird, liefert das HGM auch bei moderater Schiefe der Variablen jedoch noch zuverlässige Schätzungen für die Parameter der heterogenen Varianzkomponente. Die Restriktion des Modells auf Mittelwertzusammenhänge zwischen dem Interzept-Faktor, der Kovariate und dem Slope-Faktor (β_{11}, β_{12}) , die für alle Personen gleich sind, ist keine zwingende Restriktion. Die in Abschnitt 4.2 vorgestellte Verwendung von Interaktionseffekten im LGM zur Modellierung von differentiellen Effekten ist auch in das HGM integrierbar.

Kapitel 5

Diskussion

Im letzten Teil der Arbeit sollen die vorangegangenen Kapitel sowie die Schriften (A), (B) und (C) in einen gemeinsamen Rahmen gefasst und kritisch diskutiert werden. Hierzu folgt auf die kurze Zusammenfassung eine Diskussion der Bedeutung und der Grenzen der Arbeit. Am Ende wird ein Ausblick auf weitere, offene Forschungsfragen gegeben.

5.1 Zusammenfassung

In Kapitel 2 wurden die Eigenschaften von Verfahren zur Schätzung multipler nichtlinearer Effekten in SEM bei normal- und nicht-normalverteilten Indikatoren analysiert. Hierbei wurden neben LMS, QML und dem (erweiterten) Unconstrained-Ansatz zwei bisher kaum untersuchte momentbasierte Verfahren verwendet, die erweitert und in Standard-Software implementiert wurden. Während der 2SMM-Ansatz von Wall und Amemiya (2003) sich als sehr robust erwies, zeigte der MM-Ansatz von Mooijaart und Bentler (2010) deutliche Schwächen. Der erweiterte Unconstrained-Ansatz (Marsh et al., 2004; Kelava & Brandt, 2009) führte zwar insgesamt zu guten Schätzeigenschaften, unterschätzte jedoch die Standardfehler, was zu einem leicht erhöhten α -Fehler führte. Die verteilungsanalytischen Verfahren LMS und QML (Klein & Moosbrugger, 2000; Klein & Muthén, 2007) zeigten bei normalverteilten Indikatoren erwartungsgemäß die besten Schätzeigenschaften; bei nicht-normalverteilten Indikatoren – wie sie häufig in der Empirie vorliegen – wiesen die Schätzer jedoch einen Parameterbias auf.

In Kapitel 3 wurde der NSEMM-Ansatz vorgestellt, der speziell dafür entwickelt wurde, latente nichtlineare Effekte zu schätzen, wenn die Prädiktoren und damit auch ihre Indikatoren keiner Normalverteilung folgen. Der Ansatz approximiert die nicht-normale Verteilung der Prädiktoren anhand von Mischverteilungen aus einzelnen Normalverteilungen mit verschiedenen Mittelwerten und Varianzen. Dies ermöglicht selbst bei schief oder bimodal verteilten Prädiktoren eine erwartungstreue Schätzung der nichtlinearen Effekte.

In Kapitel 4 wurde auf die Schätzung von nichtlinearen Effekten in latenten Wachstumskurvenmodellen eingegangen. Die Schätzung von üblichen Interaktionseffekten in diesen Modellen, z.B. zwischen Interzept-Faktor und Kovariate zur Vorhersage des Slope-Faktors, ist zum größten Teil eine direkte Anwendung der Schätzverfahren für nichtlineare SEM, wie sie in den vorangegangenen Kapiteln vorgestellt wurden. Kapitel 4 befasste sich insbesondere mit dem heterogenen Wachstumskurvenmodell von Klein und Muthén (2006), das einen Nicht-Standard-Interaktionseffekt verwendet, um eine Heterogenität der Entwicklungsverläufe zu modellieren. Für diesen Ansatz wurde eine Implementierung in Mplus entwickelt, die auf dem LMS-Schätzer basiert. Neben der detaillierten Demonstration und Interpretation des Modells an einem empirischen Datensatz aus der AIDS-Forschung, wurden die Eigenschaften des Ansatzes anhand einer Simulationsstudie untersucht. Zudem wurde der Umgang mit technischen Aspekten erläutert, wie der Wahl der Startwerte oder die Erweiterung mit mehreren Kovariaten, die für die Anwendung des Modells relevant sind.

5.2 Bedeutung und Einordnung der Arbeit

Ein Schwerpunkt dieser Arbeit lag auf der Bedeutung von Annahmen für die Schätzgüte verschiedener Verfahren. Jedes statistische Verfahren basiert auf Annahmen, deren Stärke sich jedoch von einem Verfahren zum anderen unterscheidet. Nichtlineare SEM sind Verfahren, deren Schätzgüte insbesondere von Verteilungsannahmen – häufig einer multivariaten Normalverteilung der Variablen – abhängt. Jedoch sind Daten in der Empirie nur sehr selten normalverteilt, weshalb es für den Anwender notwendig ist zu wissen, wann ein eingesetztes Verfahren noch zuverlässige Ergebnisse liefert. Diesen Aspekt kann man anhand von drei Fragen konkreter formulieren, die sich jeder empirische Forscher vor einer Auswertung beantworten sollte: Sind die Annahmen heuristisch plausibel in dem empirischen Kontext? Sind die Annahmen testbar? Und wenn sie verletzt sind, in welchem Rahmen erhält man trotzdem noch valide Schlüsse? Berücksichtigt man diese Aspekte nicht bei der Interpretation der Ergebnisse, so können die statistischen Schlüsse Scheineffekte beinhalten, die keinerlei empirischen Gehalt aufweisen. Daher sollte ein Anwender kritisch mit den verwendeten Verfahren umgehen.

In dieser Arbeit wurde ein Beitrag dazu geleistet, um entscheiden zu können, wann eine erwartungstreue Schätzung von nichtlinearen Effekte in SEM möglich ist und welches Verfahren dafür verwendet werden sollte. Hierzu wurden neben einer Simulationsstudie zur Robustheitsuntersuchung die drei formulierten Fragen zu den Verfahren auch versucht theoretisch zu beantworten und es wurden Empfehlungen zur Verwendung der Verfahren ausgesprochen, anhand derer ein empirischer Forscher entscheiden kann, welches Verfahren er in seinem Kontext einsetzen möchte.

Zudem wurde ein neues Verfahren für die Schätzung nichtlinearer SEM in Schrift (B) entwickelt, das in dem Sinne plausiblere Annahmen trifft, als dass es nicht verlangt, dass die Indikatoren normalverteilt sind. Die Konzeptionalisierung anhand eines Mischverteilungsmodells erlaubt eine flexible Approximation von Verteilungen. Damit erweitern sich die Möglichkeiten zur Schätzung von nichtlinearen Effekten deutlich. Während zwar LMS bei normalverteilten Indikatoren die effizientesten Schätzungen liefert, muss bei nicht-normalverteilten Indikatoren ein Parameterbias in den Schätzungen von LMS erwartet werden. Dieser tritt beim NSEMM-Verfahren nicht auf, weshalb dieses Verfahren bei Nichtnormalität eingesetzt werden kann. Der NSEMM-Ansatz besitzt gegenüber dem 2SMM-Ansatz, der ebenfalls keine Annahmen über die Verteilung der Prädiktoren trifft, den Vorteil einer einfachen Implementierung in Mplus. Eine Erweiterung und Implementierung des 2SMM-Ansatzes, der sehr gute Schätzeigenschaften aufweist, für andere Modelle, wie z.B. für ein LGM mit Interaktionseffekten, ist hingegen komplex. Insoweit besteht eine Bedeutung dieser Arbeit darin, sowohl mit einer Implementierung des 2SMM-Ansatzes als auch mit dem NSEMM-Ansatz zwei Verfahren zur Verfügung gestellt zu haben, die in vielen empirisch relevanten Situationen zuverlässige Schätzungen liefern.

Zuletzt wurde die Implementierung eines speziellen nichtlinearen Modells für heterogene Wachstumsprozesse in Schrift (C) entwickelt. Eine Heterogenität tritt in Entwicklungsprozessen häufig auf, nämlich dann wenn die Varianzen der Entwicklungsverläufe von den Ausgangsbedingungen abhängen. In diesen Fällen ist die Annahme eines homogenen Wachstumsprozesses wie im LGM unplausibel. Durch die Implementierung des HGM in Mplus ist es nun für Anwender möglich eine aufgetretene Heterogenität in den Daten adäquat zu berücksichtigen. Daraus resultieren insbesondere eine akkuratere Schätzung der Konfidenzintervalle für modellbasierte Individualprognosen. Zudem erlaubt es Variablen zu identifizieren, auf die die Heterogenität zurückzuführen ist, oder Personen zu finden, die konsistente Entwicklungsverläufe aufweisen. Beide Aspekte können z.B. für die Planung und Durchführung von Erststudien zu neuen Therapien verwendet werden, da sie die Rekrutierung von Probanden mit bestimmten vorteilhaften Eigenschaften erlauben, wie beispielsweise Personen, die ohne Behandlung einen sehr ähnlichen Verlauf aufweisen.

5.3 Grenzen der Arbeit

Eine der Grenzen dieser Arbeit muss im Rahmen der Simulationsstudien diskutiert werden, wie sie für jedes der Verfahren in Kapitel 2, 3 und 4 durchgeführt wurden. Die Aussagekraft jeder Simulationsstudie hängt von der Repräsentativität des gewählten Populationsmodells und den gewählten Simulationsbedingungen ab. In dieser Arbeit

wurde versucht ein Populationsmodell zu wählen, das den durchschnittlichen empirischen Gegebenheiten im Sinne von Stichprobengrößen, Effektstärken und Reliabilitäten möglichst gut entspricht. Kritisch mag hier die gewählte vergleichsweise hohe Reliabilität von .80 erscheinen (in Schrift (A) und (B)), da in einigen empirischen Kontexten niedrigere Reliabilitäten zu finden sind. Doch zugleich sollte man berücksichtigen, dass die Aufdeckung von nichtlinearen Effekten eine komplexe statistische Auswertung erfordert. Hierfür ist es notwendig, dass der empirische Forscher Messinstrumente verwendet, die gute psychometrische Eigenschaften aufweisen und somit die Effekte potentiell am Besten aufdecken können.

Weiterhin wurde die Nichtnormalität der Indikatoren in allen Bedingungen durch eine Nichtnormalität der latenten Prädiktoren erzeugt. In der Empirie kann man zumindest statistisch nicht entscheiden, ob eine Nichtnormalität der Indikatoren auf die latenten Prädiktoren oder die Residuen zurückzuführen ist (Molenaar et al., 2010). Trotzdem sollte eine Nichtnormalität der Residuen in den meisten Kontexten nicht das zentrale Problem darstellen, denn Residuen setzen sich üblicherweise aus einem Konglomerat verschiedener Fehlerquellen zusammen, die zwar theoretisch jeweils nicht-normalverteilt sein können; ihre Summe sollte jedoch (asymptotisch) normalverteilt sein (zentraler Grenzwertsatz; s. z.B. in Cramér, 1999). Zudem sollten die Residuen nur einen verhältnismäßig kleineren Anteil am Indikator erklären als die Prädiktoren, wenn eine ausreichende Reliabilität der Indikatoren sichergestellt ist. In Situationen, in denen deutliche Decken- oder Bodeneffekte bei den Messungen auftreten oder die Indikatoren eine sehr geringe Reliabilität aufweisen, sollten die aus den vorgestellten Simulationsstudien gezogenen Schlüsse nur mit Vorsicht verwendet werden.

Auch wenn die hier präsentierten Simulationsstudien nicht exhaustiv sein können, stellen sie einen wichtigen Beitrag zu der gewählten Forschungsfrage dar, da sie durch die Wahl der Bedingungen deutliche Unterschiede zwischen den Verfahren unter empirisch relevanten Bedingungen demonstrieren konnten. Die gewählten Bedingungen der Nichtnormalität beinhalten hierbei ein empirisch relevantes Ausmaß an Nichtnormalität (Curran, West & Finch, 1996). Deutlich nicht-normalere Verteilungen würden vermutlich zu einer schlechteren Performanz zumindest bei einigen der eingesetzten Verfahren führen. Wie jedoch Fleishman (1978) betont, sollte die Robustheit eines Verfahrens überwiegend bei einer Verletzung der Annahmen, wie sie auch in der Empirie üblicherweise auftreten kann, untersucht werden. Der Performanz eines Verfahrens unter deutlich extremeren Bedingungen misst er nur eine untergeordnete Rolle zu, da sie für das vorgesehene Einsatzgebiet des Verfahrens nicht relevant ist.

Eine weitere Grenze sollte in Bezug auf die Anwendung des NSEMM-Ansatzes kritisch diskutiert werden muss. Zum einen muss man für das Verfahren annehmen, dass die Nichtnormalität der Indikatoren auf die Prädiktoren zurückzuführen ist. Wie oben

ausgeführt, ist diese Annahme in manchen Fällen unplausibel, nämlich dann, wenn man nicht-normalverteilte Indikatoren mit einer niedrigen Reliabilität im Modell verwendet. Zum anderen ist die Wahl der Anzahl der Mischverteilungskomponenten (latenten Klassen) in der Praxis ein wichtiger Punkt für die Schätzgüte des Verfahrens. Zu wenige Klassen (z.B. zwei latente Klassen) könnten zu einer nicht ausreichend guten Approximation der Verteilungen der Prädiktoren führen, die wiederum in einer leicht verzerrten Schätzung der nichtlinearen Effekte resultieren kann (s. Abschnitt 3.5). Zu viele Klassen (z.B. mehr als sechs) könnten eine Überparametrisierung zur Folge haben, die in einer niedrigen Konvergenzrate und/oder einer niedrigen Power für die Aufdeckung der nichtlinearen Effekte resultieren könnte. Neben statistischen Entscheidungskriterien wie dem AIC oder dem BIC (Nylund et al., 2007), gibt es kaum Möglichkeiten, die Adäquatheit der Anzahl der latenten Klassen statistisch zu überprüfen. Inwieweit jedoch der AIC und der BIC bei nichtlinearen Modellen verwendbar ist, stellt eine noch nicht untersuchte Forschungsfrage dar. Als heuristische Alternative zur Entscheidung über die Anzahl der Klassen können zwei Informationsquellen herangezogen werden. Erstens erscheint eine Einschränkung der Anzahl der Klassen auf drei bis fünf zumindest auf der Basis von verschiedenen Simulationsstudien als sinnvoll (Bauer, 2005; Kelava et al., in press). Zweitens beschreiben McLachlan und Peel (2000), wie man optische Hinweise der (univariaten) Verteilungsform verwenden kann, um eine heuristische Entscheidung über die Zahl der Komponenten zu treffen. Anhand dieser Aspekte und unter Berücksichtigung der Stichprobengröße, sollte in der Praxis eine Entscheidung getroffen werden.

5.4 Ausblick

In den letzten 20 Jahren fand eine deutliche Entwicklung und Verbesserung der Verfahren zur Schätzung von nichtlinearen SEM statt. Trotzdem sind für die Zukunft drei Aspekte von Bedeutung. Der erste wichtige Aspekt, der in dieser Arbeit nicht diskutiert wurde, ist die Modellgüte von nichtlinearen SEM. Es existiert derzeit kein Modellgütemaß, das eine Gesamtbeurteilung eines nichtlinearen SEM erlaubt. Dies ist darauf zurückzuführen, dass kein saturiertes Modell wie für lineare SEM existiert. Ein erster Ansatz für ein eingeschränktes nichtlineares SEM mit nur einer einzigen Indikatorvariablen für die endogene Variable basiert auf einem Maß für die Heteroskedastizität der Residuen (Klein & Schermelleh-Engel, 2010).

Die vorgestellte Implementierung des HGM in Mplus weist einen Parameterbias auf, wenn die Indikatoren (und das latente Interzept) deutlich nicht-normalverteilt sind. Jedoch treten gerade in Daten von Entwicklungsprozessen solche Boden- oder Deckeneffekte auf, die beispielsweise auf anfänglich fehlende Ausprägungen in den Variablen zurückzuführen sind (z.B. eine sehr geringe Leseleistung fast aller Schüler zu Beginn

des ersten Schuljahres) oder auf Grenzen des Messinstruments. Es bleibt abzuwarten, ob eine Erweiterung der Implementierung anhand des NSEMM-Ansatzes oder auch anhand von Bayes-Schätzverfahren für nichtlineare Modelle (z.B. Kelava & Nagengast, 2012) diese Problematik verringern kann.

Zuletzt wurden in dieser Arbeit vornehmlich Interaktionseffekte und quadratische Effekte untersucht, die eine spezifische Funktion der Nichtlinearität modellieren, deren Adäquatheit im Allgemeinen nicht testbar ist. Neue Ansätze, die auch andere Funktionen approximieren und testen können, ohne zu strenge Annahmen über den Funktionstyp zu treffen, könnten eine Verallgemeinerung der nichtlinearen SEM erlauben. Dies wiederum ermöglicht eine verbesserte modellbasierte Prognose.

Literaturverzeichnis

- Ajzen, I. (1987). Attitudes, traits, and actions: Dispositional prediction of behavior in personality and social psychology. In L. Berkowitz (Hrsg.), *Advances in experimental social psychology* (Bd. 20, S. 1–63). New York: Academic Press.
- Ajzen, I. & Fishbein, M. (2005). The influence of attitudes on behavior. In D. Albarracín, B. T. Johnson & M. P. Zanna (Hrsg.), *The handbook of attitudes* (S. 173–221). Mahwah, NJ: Erlbaum.
- Ajzen, I. & Madden, T. J. (1986). Prediction of goal-directed behavior: Attitudes, intentions, and perceived behavioral control. *Journal of Experimental Social Psychology*, *22*, 453–474.
- Arminger, G. & Stein, P. (1997). Finite mixtures of covariance structure models with regressors. *Sociological Methods and Research*, *26*, 148–182.
- Arminger, G., Stein, P. & Wittenberg, J. (1999). Mixtures of conditional mean- and covariance-structure models. *Psychometrika*, *64*, 475–494.
- Aroian, L. A. (1944). The probability function of the product of two normally distributed variables. *The Annals of Mathematical Statistics*, *18*, 265–271.
- Atkinson, J. (1964). *An introduction to motivation*. Oxford: Van Nostrand.
- Bauer, D. J. (2005). A semiparametric approach to modeling nonlinear relations among latent variables. *Structural Equation Modeling*, *12*, 513–535.
- Bauer, D. J. & Curran, P. J. (2004). The integration of continuous and discrete latent variable models: Potential problems and promising opportunities. *Psychological Methods*, *9*, 3–29.
- Beierlein, C., Werner, C. S., Preiser, S. & Wermuth, S. (2011). Are just-world-beliefs compatible with justifying inequality? Collective political efficacy as a moderator. *Social Justice Research*, *24*, 278–296.

- Bentler, P. M. (2000–2008). *EQS 6 structural equations program manual*. Encino, CA: Multivariate Software, Inc.
- Bohrstedt, G. W. & Marwell, G. (1978). The reliability of the products of two random variables. In K. Schüssler (Hrsg.), *Sociological methodology* (S. 254–273). San Francisco: Jossey-Bass.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.
- Bollen, K. A. (1995). Structural equation models that are nonlinear in latent variables: A least squares estimator. In P. V. Marsden (Hrsg.), *Sociological methodology 1995* (S. 223–251). Oxford: Blackwell.
- Bollen, K. A. & Curran, P. J. (2006). *Latent curve models: A structural equation perspective*. Hoboken, NJ: Wiley.
- Boneau, C. A. (1960). The effects of violations of assumptions underlying the t test. *Psychological Bulletin*, *57*, 49–64.
- Boomsma, A. (1983). *On the robustness of LISREL (maximum likelihood estimation) against small sample size and nonnormality*. Amsterdam: Sociometric Research Foundation.
- Borsboom, D., Mellenbergh, G. J. & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, *110*, 203–218.
- Brandt, H., Kelava, A. & Klein, A. G. (in press). A simulation study comparing recent approaches for the estimation of nonlinear effects in sem under the condition of non-normality. *Structural Equation Modeling*.
- Brandt, H., Klein, A. G. & Gäde, J. (submitted). Implementation and practical application of a heterogeneous growth curve model. *Multivariate Behavioral Research*.
- Brody, G. H., Chen, Y.-F., McBride Murry, V., Simons, R. L., Ge, X., Gibbons, F. X., Gerrard, M. & Cutrona, C. E. (2006). Perceived discrimination and the adjustment of african american youths: A five-year longitudinal analysis with contextual moderation effects. *Child Development*, *77*, 1170–1189.
- Browne, M. W. (1984). Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, *37*, 62–83.
- Burzykowski, T., Molenberghs, G. & Buyse, M. (2006). *The evaluation of surrogate endpoints*. New York: Springer.

- Cham, H., West, S. G., Ma, Y. & Aiken, L. S. (2012). Estimating latent variable interactions with nonnormal observed data: A comparison of four approaches. *Multivariate Behavioral Research*, *47*, 840–876.
- Cox, D. R. & Small, N. J. H. (1978). Testing multivariate normality. *Biometrika*, *65*, 263–272.
- Cramér, H. (1936). Über eine Eigenschaft der normalen Verteilungsfunktion. *Mathematische Zeitschrift*, *41*, 405–414.
- Cramér, H. (1999). *Mathematical methods of statistics* (19. Aufl.). Princeton, NJ: Princeton University Press.
- Cronbach, L. J. & Snow, R. E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions*. New York: Irvington.
- Curran, P. J., West, S. G. & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, *1*, 16–29.
- Dempster, A. P., Laird, N. M. & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Ser. B*, *39*, 1–38.
- Dimitruk, P., Schermelleh-Engel, K., Kelava, A. & Moosbrugger, H. (2007). Challenges in nonlinear structural equation modeling. *Methodology*, *3*, 100–114.
- Dolan, C. V. & van der Maas, H. L. J. (1998). Fitting multivariate normal finite mixtures subject to structural equation modeling. *Psychometrika*, *63*, 227–253.
- Duncan, T. E., Duncan, S. C. & Strycker, L. A. (2006). *An introduction to latent variable growth curve modeling: Concepts, issues, and applications* (2. Aufl.). Mahwah, NJ: Erlbaum.
- Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L. & Midgley, C. (1983). Expectancies, values and academic behaviors. In J. T. Spence (Hrsg.), *Achievement and achievement motives. Psychological and sociological approaches* (S. 75–146). San Francisco, CA: Freeman.
- Eysenck, H. J. (1994). Personality: Biological foundations. In P. A. Vernon (Hrsg.), *The neuropsychology of individual differences*. London: Academic Press.
- Fitzmaurice, G. M., Laird, N. M. & Ware, J. H. (2004). *Applied longitudinal analysis*. New York: Wiley.

- Fleishman, A. I. (1978). A method for simulating non-normal distributions. *Psychometrika*, *43*, 521–532.
- Fleming, T. R. & DeMets, D. L. (1996). Surrogate end points in clinical trials: Are we being misled? *Annals of Internal Medicine*, *125*, 605–613.
- Frangakis, C. E. & Rubin, D. B. (2002). Principal stratification in causal inference. *Biometrics*, *58*, 21–29.
- Gangestad, S. & Snyder, M. (1985). “To carve nature at its joints“: On the existence of discrete classes in personality. *Psychological Review*, *92*, 317–349.
- Ganzach, Y. (1997). Misleading interaction and curvilinear terms. *Psychological Methods*, *3*, 235–247.
- Gilbert, P. B. & Hudgens, M. G. (2008). Evaluating candidate principal surrogate endpoints. *Biometrics*, *64*, 1146–1154.
- Goodnight, J. A., Bates, J. E., Dawn Staples, A., Pettit, G. S. & Dodge, K. A. (2007). Temperamental resistance to control increases the association between sleep problems and externalizing behavior development. *Journal of Family Psychology*, *21*, 39–48.
- Groeneveld, R. A. & Meeden, G. (1984). Measuring skewness and kurtosis. *The Statistician*, *33*, 391–399.
- Henry, K., Erice, A., Tierney, C., Balfour, H. H. J., Fischl, M., Kmack, A., Liou, S. H., Kenton, A., Hirsch, M. S., Phair, J., Martinez, A. & Kahn, J. O. (1998). A randomized, controlled, double-blind study comparing the survival benefit of four different reverse transcriptase inhibitor therapies (three-drug, two-drug, and alternating drug) for the treatment of advanced aids. *Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology*, *19*, 339–349.
- Holland, P. W. (1988). Causal inference, path analysis, and recursive structural equations models. *Sociological Methodology*, *18*, 449–484.
- Hoyer, J. & Klein, A. G. (2000). Self-reflection and well-being: Is there a healthy amount of introspection? *Psychological Reports*, *86*, 135–141.
- Jaccard, J. & Wan, C. K. (1995). Measurement error in the analysis of interaction effects between continuous predictors using multiple regression: Multiple indicator and structural equation approaches. *Psychological Bulletin*, *117*, 348–357.

- Jamieson, D. & Zanna, M. P. (1989). Need for structure in attitude formation and expression. In A. R. B. Pratkanis & S. J. Breckler (Hrsg.), *Attitude structure and function* (S. 383–406). Hillsdale, NJ: Erlbaum.
- Jedidi, K., Jagpal, H. S. & DeSarbo, W. S. (1997a). Finite-mixture structural equation models for response based segmentation and unobserved heterogeneity. *Marketing Science*, *16*, 39–59.
- Jedidi, K., Jagpal, H. S. & DeSarbo, W. S. (1997b). STEMM: A general finite mixture structural equation model. *Journal of Classification*, *14*, 23–50.
- Joanes, D. N. & Gill, C. A. (1998). Comparing measures of sample skewness and kurtosis. *The Statistician*, *47*, 183–189.
- Jöreskog, K. G. & Sörbom, D. (1996). *LISREL 8: User's reference guide*. Lincolnwood, IL: Scientific Software International.
- Jöreskog, K. G. & Yang, F. (1996). Nonlinear structural equation models: The Kenny-Judd model with interaction effects. In G. A. Marcoulides & R. E. Schumacker (Hrsg.), *Advanced structural equation modeling: Issues and techniques* (S. 57–87). Mahwah, NJ: Erlbaum.
- Kane, M. T. (1992). An argument-based approach to validity. *Quantitative Methods in Psychology*, *112*, 527–535.
- Karasek, R. A. (1979). Job demands, job decision latitude, and mental strain: Implication for job redesign. *Administrative Quarterly*, *24*, 285–307.
- Kelava, A. & Brandt, H. (2009). Estimation of nonlinear latent structural equation models using the extended unconstrained approach. *Review of Psychology*, *16*, 123–131.
- Kelava, A., Moosbrugger, H., Dimitruk, P. & Schermelleh-Engel, K. (2008). Multicollinearity and missing constraints: A comparison of three approaches for the analysis of latent nonlinear effects. *Methodology*, *4*, 51–66.
- Kelava, A. & Nagengast, B. (2012). A bayesian model for the estimation of latent interaction and quadratic effects when latent variables are non-normally distributed. *Multivariate Behavioral Research*, *47*, 717–742.
- Kelava, A., Nagengast, B. & Brandt, H. (in press). A nonlinear structural equation mixture modeling approach for non-normally distributed latent predictor variables. *Structural Equation Modeling*.

- Kelava, A., Werner, C., Schermelleh-Engel, K., Moosbrugger, H., Zapf, D., Ma, Y., Cham, H., Aiken, L. S. & West, S. G. (2011). Advanced nonlinear structural equation modeling: Theoretical properties and empirical application of the distribution-analytic LMS and QML estimators. *Structural Equation Modeling*, *18*, 465–491.
- Kendall, M. G. (1948). *The advanced theory of statistics* (Bd. II). London: Charles Griffin & Company Limited.
- Kenny, D. & Judd, C. M. (1984). Estimating the nonlinear and interactive effects of latent variables. *Psychological Bulletin*, *96*, 201–210.
- Khoo, S. T. (2001). Assessing program effects in the presence of treatment-baseline interactions: A latent curve approach. *Psychological Methods*, *6*, 234–257.
- Klein, A. G. (2000). *Verfahren zur Analyse von Moderatoreffekten in Strukturgleichungsmodellen*. Hamburg: Verlag Kovac.
- Klein, A. G. & Brandt, H. (in Arbeit). A causal model for the estimation of unbiased direct and indirect effects in mediation analysis.
- Klein, A. G. & Moosbrugger, H. (2000). Maximum likelihood estimation of latent interaction effects with the LMS method. *Psychometrika*, *65*, 457–474.
- Klein, A. G. & Muthén, B. O. (2006). Modeling heterogeneity of latent growth depending on initial status. *Journal of Educational and Behavioral Statistics*, *31*, 357–375.
- Klein, A. G. & Muthén, B. O. (2007). Quasi maximum likelihood estimation of structural equation models with multiple interaction and quadratic effects. *Multivariate Behavioral Research*, *42*, 647–674.
- Klein, A. G. & Schermelleh-Engel, K. (2010). Introduction of a new measure for detecting poor fit due to omitted nonlinear terms in SEM. *Advanced Statistical Analysis*, *94*, 157–166.
- Klein, A. G., Schermelleh-Engel, K., Moosbrugger, H. & Kelava, A. (2009). Assessing spurious interaction effects. In T. Teo & M. Khine (Hrsg.), *Structural equation modeling in educational research: Concepts and applications* (S. 13–28). Rotterdam, NL: Sense Publishers.
- Klein, A. G. & Stoolmiller, M. (2003). Detecting latent interaction effects in behavioral data. *Methods of Psychological Research Online*, *8*, 113–126.

- Kline, S. L. (1987). Self-monitoring and attitude-behavior correspondence in cable television subscription. *Journal of Social Psychology, 127*, 605–609.
- Larsen, H., Overbeek, G., Vermulst, A. A., Granic, I. & Engels, R. C. M. E. (2010). Initiation and continuation of best friends and adolescents' alcohol consumption: Do self-esteem and self-control function as moderators? *International Journal of Behavioral Development, 34*, 406–416.
- Lee, S. Y., Song, X. Y. & Poon, W. Y. (2004). Comparison of approaches in estimating interaction and quadratic effects of latent variables. *Multivariate Behavioral Research, 39*, 37–67.
- Li, F., Duncan, T. E. & Acock, A. (2000). Modeling interaction effects in latent growth curve models. *Structural Equation Modeling, 7*, 497–533.
- Little, T. D., Bovaird, J. A. & Widaman, K. F. (2006). On the merits of orthogonalizing powered and interaction terms: Implications for modeling interactions among latent variables. *Structural Equation Modeling, 13*, 497–519.
- Lusch, R. F. & Brown, J. R. (1996). Interdependency, contracting, and relational behavior in marketing channels. *Journal of Marketing, 60*, 19–38.
- MacCallum, R. C. & Mar, C. M. (1995). Distinguishing between moderator and quadratic effects in multiple regression. *Psychological Bulletin, 118*, 405–421.
- Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika, 57*, 519–530.
- Marsh, H. W., Wen, Z. & Hau, K.-T. (2004). Structural equation models of latent interactions: Evaluation of alternative estimation strategies and indicator construction. *Psychological Methods, 9*, 275–300.
- Marsh, H. W., Wen, Z. & Hau, K.-T. (2006). Structural equation models of latent interaction and quadratic effects. In G. R. Hancock & R. O. Müller (Hrsg.), *Structural equation modeling: A second course*. (S. 225–265). Greenwich, CT: Information Age Publishing.
- MATLAB. (2010). *Version 7.10.0*. Natick, Massachusetts: The MathWorks Inc.
- McDowell, M. A., Fryar, C. D., Ogden, C. L. & Flegal, K. M. (2008). Anthropometric reference data for children and adults: United States, 2003–2006. *National Health Statistics Report, 10*, 1–45.
- McLachlan, G. J. & Peel, D. (2000). *Finite mixture models*. New York: Wiley.

- Mecklin, C. J. & Mundfrom, D. J. (2004). An appraisal and bibliography of tests for multivariate normality. *International Statistical Review*, *72*, 123–138.
- Mehta, P. D. & West, S. G. (2000). Putting the individual back into individual growth curves. *Psychological Methods*, *5*, 23–43.
- Meredith, W. & Tisak, J. (1990). Latent curve analysis. *Psychometrika*, *55*, 107–122.
- Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin*, *105*, 156–166.
- Milan, S., Snow, S. & Belay, S. (2009). Depressive symptoms in mothers and children: Preschool attachment as a moderator of risk. *Developmental Psychology*, *45*, 1019–1033.
- Molenaar, D., Dolan, C. V. & Verhelst, N. D. (2010). Testing and modelling non-normality within the one-factor model. *British Journal of Mathematical and Statistical Psychology*, *63*, 293–317.
- Mooijaart, A. & Bentler, P. M. (2010). An alternative approach for nonlinear latent variable models. *Structural Equation Modeling*, *17*, 357–573.
- Mooijaart, A. & Satorra, A. (2011). MM versus ML estimates of structural equation models with interaction terms: robustness to non-normality of the consistency. *Department of Statistics Papers, Department of Statistics, UCLA*.
- Moosbrugger, H., Schermelleh-Engel, K., Kelava, A. & Klein, A. G. (2009). Testing multiple nonlinear effects in structural equation modeling: A comparison of alternative estimation approaches. In T. Teo & M. S. Khine (Hrsg.), *Structural equation modeling in educational research: Concepts and applications* (S. 103–136). Rotterdam, NL: Sense Publishers.
- Moulder, B. C. & Algina, J. (2002). Comparison of methods for estimating and testing latent variable interactions. *Structural Equation Modeling*, *9*, 1–19.
- Muthén, B. O. (2001). Second-generation structural equation modeling with a combination of categorical and continuous latent variables: New opportunities for latent-class growth modeling. In L. M. Collins & A. Sayer (Hrsg.), *New methods for the analysis of change* (S. 291–322). Washington DC: American Psychological Association.
- Muthén, B. O. (2004). Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (Hrsg.), *Handbook of quantitative methodology for the social sciences* (S. 345–368). Newbury Park, CA: Sage.

- Muthén, B. O. & Curran, J. P. (1997). General longitudinal modeling of individual differences in experimental designs: A latent variable framework for analysis and power estimation. *Psychological Methods*, *2*, 371–402.
- Muthén, L. K. & Muthén, B. O. (1998–2012). *Mplus user's guide* (7. Aufl.). Los Angeles, CA: Muthén & Muthén.
- Nylund, K. L., Asparouhov, T. & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling*, *14*, 535–569.
- Oja, H. (1981). On location, scale, skewness and kurtosis of univariate distributions. *Scandinavian Journal of Statistics*, *8*, 154–168.
- Parrila, R., Aunola, K., Leskinen, E., Nurmi, J.-E. & Kirby, J. R. (2005). Development of individual differences in reading: Results from longitudinal studies in English and Finnish. *Journal of Educational Psychology*, *97*, 299–319.
- Pek, J., Losardo, D. & Bauer, D. J. (2011). Confidence intervals for a semiparametric approach to modeling nonlinear relations among latent variables. *Structural Equation Modeling*, *18*, 537–553.
- Pek, J., Sterba, S. K., Kok, B. E. & Bauer, D. J. (2009). Estimating and visualizing nonlinear relations among latent variables: A semiparametric approach. *Multivariate Behavioral Research*, *44*, 407–436.
- Ping, R. A. (1995). A parsimonious estimating technique for interaction and quadratic latent variables. *Journal of Marketing Research*, *32*, 336–347.
- Ping, R. A. (1996). Latent variable interaction and quadratic effect estimation: A two-step technique using structural equation analysis. *Psychological Bulletin*, *119*, 166–175.
- Preacher, K. J., Curran, P. J. & Bauer, D. J. (2006). Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics*, *31*, 437–448.
- R Development Core Team. (2011). *R: A language and environment for statistical computing*. Vienna, Austria.
- Reinecke, J. (2005). *Strukturgleichungsmodelle in den Sozialwissenschaften*. München: Oldenbourg.
- Reinecke, J. (2012). *Wachstumsmodelle*. Mehring: Rainer Hampp Verlag.

- Rencher, A. C. (2002). *Methods of multivariate analysis* (2. Aufl.). New York: Wiley.
- Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. Boca Raton: Chapman and Hall/CRC.
- Schermelleh-Engel, K., Klein, A. G. & Moosbrugger, H. (1998). Estimating nonlinear effects using a Latent Moderated Structural Equations Approach. In R. E. Schumacker & G. A. Marcoulides (Hrsg.), *Interaction and nonlinear effects in structural equation modeling* (S. 203–238). Mahwah, NJ: Erlbaum.
- Schumacker, R. E. & Marcoulides, G. A. (1998). *Interaction and nonlinear effects in structural equation modeling*. Mahwah, NJ: Erlbaum.
- Sénéchal, M. & LeFevre, J.-A. (2002). Parental involvement in the development of children's reading skill: A five-year longitudinal study. *Child Development*, *73*, 445–460.
- Shapiro, S. S. & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, *52*, 591–611.
- Small, N. J. H. (1980). Marginal skewness and kurtosis in testing multivariate normality. *Applied Statistics*, *29*, 85–87.
- Snyder, M. & Tanke, E. D. (1976). Behavior and attitude: Some people are more consistent than others. *Journal of Personality*, *44*, 501–517.
- Stoer, J. (1989). *Numerische Mathematik 1* (5. Aufl.). Heidelberg: Springer.
- Ten Have, T. R., Joffe, M. M., Lynch, K. G., Brown, G. K., Maisto, S. A. & Beck, A. T. (2007). Causal mediation analyses with rank preserving models. *Biometrics*, *63*, 926–934.
- Titterton, D. M., Smith, A. F. M. & Makov, U. E. (1985). *Statistical analysis of finite mixture distributions*. Chichester, England: Wiley.
- Vallejo, G., Fernández, M. P., Livacic-Rojas, P. E. & Tuero-Herrero, E. (2011). Comparison of modern methods for analyzing repeated measures data with missing values. *Multivariate Behavioral Research*, *46*, 900–937.
- Wall, M. M. & Amemiya, Y. (2000). Estimation for polynomial structural equation models. *Journal of the Statistical American Association*, *95*, 929–940.
- Wall, M. M. & Amemiya, Y. (2001). Generalized appended product indicator procedure for nonlinear structural equation analysis. *Journal of Educational and Behavioral Statistics*, *26*, 1–29.

- Wall, M. M. & Amemiya, Y. (2003). A method of moments technique for fitting interaction effects in structural equation models. *British Journal of Mathematical and Statistical Psychology*, *56*, 47–63.
- Wechsler, D. (1944). *The measurement of adult intelligence* (3. Aufl.). Baltimore: Williams & Wilkins.
- Wei, M., Mallinckrodt, B., Russell, D. W. & Abraham, W. T. (2004). Maladaptive perfectionism as a mediator and moderator between adult attachment and depressive mood. *Journal of Counseling Psychology*, *51*, 201–212.
- West, S. G., Finch, J. F. & Curran, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Hrsg.), *Structural equation modeling: Concepts, issues and applications*. (S. 56–75). Newbury Park, CA: Sage.
- White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica*, *50*, 1–25.
- Wigfield, A. & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, *25*, 68–81.
- Yerkes, R. M. & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, *18*, 459–482.
- Yuan, K. H. & Bentler, P. M. (2000). Three likelihoodbased methods for mean and covariance structure analysis with nonnormal missing data. *Sociological Methodology*, *30*, 165–200.
- Zeger, S. L. & Diggle, P. J. (1994). Semiparametric models for longitudinal data with application to CD4 cell numbers in HIV seroconverters. *Biometrics*, *50*, 689–699.

Anhang A

Erweiterung des 2SMM-Ansatzes

Im folgenden wird zuerst der Parameterschätzer $\hat{\alpha}$ für ein vollständiges nichtlineares Modell vorgestellt. Anschließend wird auf die Erweiterung der Gleichungen für die Standardfehlerschätzer eingegangen. Es wird ein direkter Bezug zu den Gleichungen von Wall und Amemiya (2003) (WA) gegeben, die ein Interaktionsmodell spezifizierten. Aus Gründen der Einfachheit und Vergleichbarkeit wird hier im Gegensatz zu Abschnitt 2.1.3 die Notation von WA übernommen (bis auf die Bezeichnung der linearen und nichtlinearen Einflussgewichte), d.h. $f_1 = \eta$, $f_2 = \xi_1$ und $f_3 = \xi_2$.

Die Schätzung der Faktorwerte erfolgt wie in dem in WA Stage 1 (i) und (ii) beschriebenen Vorgehen. In der Notation werden Schätzungen für einen Parameter θ als $\hat{\theta}$ dargestellt.

A.1 Schätzung der Parameter

Sei ein vollständiges nichtlineares Modell für die i -te Person ($i = 1 \dots n$) gegeben durch

$$f_{1i} = \alpha_0 + \gamma_1 f_{2i} + \gamma_2 f_{3i} + \omega_{12} f_{2i} f_{3i} + \omega_{11} f_{2i}^2 + \omega_{22} f_{3i}^2 + \zeta_i \quad (\text{A.1})$$

mit einem Messmodell

$$\begin{pmatrix} y_{2i} \\ y_{3i} \\ x_{2i} \\ x_{3i} \\ x_{5i} \\ x_{6i} \\ y_{1i} \\ x_{1i} \\ x_{4i} \end{pmatrix} = \begin{pmatrix} \beta_{01} \\ \beta_{02} \\ \beta_{03} \\ \beta_{04} \\ \beta_{05} \\ \beta_{06} \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \beta_{11} & 0 & 0 \\ \beta_{12} & 0 & 0 \\ 0 & \beta_{13} & 0 \\ 0 & \beta_{14} & 0 \\ 0 & 0 & \beta_{15} \\ 0 & 0 & \beta_{16} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} f_{1i} \\ f_{2i} \\ f_{3i} \end{pmatrix} + \begin{pmatrix} \epsilon_{1i} \\ \epsilon_{2i} \\ \epsilon_{3i} \\ \epsilon_{4i} \\ \epsilon_{5i} \\ \epsilon_{6i} \\ \epsilon_{7i} \\ \epsilon_{8i} \\ \epsilon_{9i} \end{pmatrix}. \quad (\text{A.2})$$

Für weitere Schritte sei

$$\boldsymbol{\beta}_0 := \begin{pmatrix} \beta_{01} \\ \beta_{02} \\ \beta_{03} \\ \beta_{04} \\ \beta_{05} \\ \beta_{06} \end{pmatrix} \quad \text{und} \quad \boldsymbol{\beta}_1 := \begin{pmatrix} \beta_{11} & 0 & 0 \\ \beta_{12} & 0 & 0 \\ 0 & \beta_{13} & 0 \\ 0 & \beta_{14} & 0 \\ 0 & 0 & \beta_{15} \\ 0 & 0 & \beta_{16} \end{pmatrix} \quad (\text{A.3})$$

sowie $\text{vec}(\boldsymbol{\beta}_1)$ der Vektor, der die Spalten von $\boldsymbol{\beta}_1$ enthält. Die später gebrauchte Matrix $\hat{\boldsymbol{\Gamma}}$ ist für die neun Indikatorvariablen definiert als (s. Gleichung WA 6)

$$\hat{\boldsymbol{\Gamma}} = (\mathbf{0}_{3 \times 6}, \mathbf{I}_3) \hat{\boldsymbol{\Psi}} \begin{pmatrix} \mathbf{I}_6 \\ -\hat{\boldsymbol{\beta}}_1' \end{pmatrix} \left((\mathbf{I}_6, -\hat{\boldsymbol{\beta}}_1) \hat{\boldsymbol{\Psi}} \begin{pmatrix} \mathbf{I}_6 \\ -\hat{\boldsymbol{\beta}}_1' \end{pmatrix} \right)^{-1} \quad (\text{A.4})$$

mit der geschätzten Kovarianzmatrix $\hat{\boldsymbol{\Psi}}$ der Residuen $(\epsilon_1, \dots, \epsilon_9)$ der Indikatoren, der Identitätsmatrix \mathbf{I}_k mit Dimension $(k \times k)$ und der Nullmatrix $\mathbf{0}_{k \times l}$ mit Dimension $(k \times l)$. Sei weiterhin die Kovarianzmatrix $\hat{\boldsymbol{\Sigma}}_{ee}$ für die Schätzfehler e_{ji} der geschätzten Faktorwerte \hat{f}_{ji} (aus Gleichung WA 6), $e_{ji} = \hat{f}_{ji} - f_{ji}$, $j = 1, 2, 3$, gegeben durch (s. Gleichung WA 7)

$$\hat{\boldsymbol{\Sigma}}_{ee} = \begin{pmatrix} \hat{\sigma}_{11} & \hat{\sigma}_{12} & \hat{\sigma}_{13} \\ & \hat{\sigma}_{22} & \hat{\sigma}_{23} \\ & & \hat{\sigma}_{33} \end{pmatrix} \quad (\text{A.5})$$

und $\text{vech}\hat{\boldsymbol{\Sigma}}_{ee}' = (\hat{\sigma}_{11}, \hat{\sigma}_{12}, \hat{\sigma}_{13}, \hat{\sigma}_{22}, \hat{\sigma}_{23}, \hat{\sigma}_{33})$ der Vektor, der die nichtredundanten Elemente von $\hat{\boldsymbol{\Sigma}}_{ee}$ enthält. Sei $\mathbf{f}_i = (1, f_{2i}, f_{3i}, f_{2i}f_{3i}, f_{2i}^2, f_{3i}^2)'$ der (erweiterte) Scorevector der Prädiktoren f_{2i}, f_{3i} und ihrer Produkte sowie $\hat{\mathbf{f}}_i$ der entsprechende Schätzer des Scorevectors mit $\hat{f}_{ji} = f_{ji} + e_{ji}$, $j = 1, 2, 3$ (Gleichung WA 8).

Dann ist der erwartungstreue Momentschätzer $\hat{\boldsymbol{\alpha}} = (\hat{\alpha}_0, \hat{\gamma}_1, \hat{\gamma}_2, \hat{\omega}_{12}, \hat{\omega}_{11}, \hat{\omega}_{22})'$ nach Gleichung WA 11 gegeben durch

$$\hat{\boldsymbol{\alpha}} = \hat{\mathbf{M}}^{-1} \hat{\mathbf{m}} \quad (\text{A.6})$$

mit $\hat{\mathbf{M}}$ und $\hat{\mathbf{m}}$ so definiert, dass sie erwartungstreue Schätzer für \mathbf{M} und \mathbf{m} sind, was der Fall ist, wenn für alle $i = 1 \dots n$ Personen gilt:

$$E[\hat{\mathbf{M}}_i | \mathbf{f}_i] = \mathbf{M}_i \quad \text{und} \quad E[\hat{\mathbf{m}}_i | \mathbf{f}_i, f_{1i}] = \mathbf{m}_i. \quad (\text{A.7})$$

Für das vollständige nichtlineare Modell ergibt sich für $\hat{\mathbf{M}}$ die (symmetrische) (6×6) Matrix

$$\hat{\mathbf{M}} = \frac{1}{n} \sum_{i=1}^n \hat{\mathbf{M}}_i = \frac{1}{n} \sum_{i=1}^n \begin{pmatrix} 1 & \hat{f}_{2i} & \hat{f}_{3i} & M_{41} & M_{51} & M_{61} \\ & M_{22} & M_{32} & M_{42} & M_{52} & M_{62} \\ & & M_{33} & M_{43} & M_{53} & M_{63} \\ & & & M_{44} & M_{54} & M_{64} \\ & & & & M_{55} & M_{65} \\ & & & & & M_{66} \end{pmatrix} \quad (\text{A.8})$$

mit den Elementen

$$\begin{aligned} M_{41} &= \hat{f}_{2i} \hat{f}_{3i} - \hat{\sigma}_{23} \\ M_{51} &= M_{22} = \hat{f}_{2i}^2 - \hat{\sigma}_{22} \\ M_{61} &= M_{33} = \hat{f}_{3i}^2 - \hat{\sigma}_{33} \\ M_{32} &= \hat{f}_{2i} \hat{f}_{3i} - \hat{\sigma}_{23} \\ M_{42} &= M_{53} = \hat{f}_{2i}^2 \hat{f}_{3i} - 2\hat{\sigma}_{23} \hat{f}_{2i} - \hat{\sigma}_{22} \hat{f}_{3i} - \hat{\mu}_{223} \\ M_{52} &= \hat{f}_{2i}^3 - 3\hat{\sigma}_{22} \hat{f}_{2i} - \hat{\mu}_{222} \\ M_{62} &= M_{43} = \hat{f}_{2i} \hat{f}_{3i}^2 - 2\hat{\sigma}_{23} \hat{f}_{3i} - \hat{\sigma}_{33} \hat{f}_{2i} - \hat{\mu}_{233} \\ M_{63} &= \hat{f}_{3i}^3 - 3\hat{\sigma}_{33} \hat{f}_{3i} - \hat{\mu}_{333} \\ M_{44} &= M_{65} = \hat{f}_{2i}^2 \hat{f}_{3i}^2 - \hat{\sigma}_{33} \hat{f}_{2i}^2 - \hat{\sigma}_{22} \hat{f}_{3i}^2 - 4\hat{\sigma}_{23} \hat{f}_{2i} \hat{f}_{3i} + 2\hat{\sigma}_{22} \hat{\sigma}_{33} + 4\hat{\sigma}_{23}^2 - 2\hat{\mu}_{233} \hat{f}_{2i} \\ &\quad - 2\hat{\mu}_{223} \hat{f}_{3i} - \hat{\nu}_{2233} \\ M_{54} &= \hat{f}_{2i}^3 \hat{f}_{3i} - 3\hat{\sigma}_{23} \hat{f}_{2i}^2 - 3\hat{\sigma}_{22} \hat{f}_{2i} \hat{f}_{3i} + 6\hat{\sigma}_{22} \hat{\sigma}_{23} - 3\hat{\mu}_{223} \hat{f}_{2i} - \hat{\mu}_{222} \hat{f}_{3i} - \hat{\nu}_{2223} \\ M_{64} &= \hat{f}_{2i} \hat{f}_{3i}^3 - 3\hat{\sigma}_{23} \hat{f}_{3i}^2 - 3\hat{\sigma}_{33} \hat{f}_{2i} \hat{f}_{3i} + 6\hat{\sigma}_{23} \hat{\sigma}_{33} - 3\hat{\mu}_{233} \hat{f}_{3i} - \hat{\mu}_{333} \hat{f}_{2i} - \hat{\nu}_{2333} \\ M_{55} &= \hat{f}_{2i}^4 - 6\hat{\sigma}_{22} \hat{f}_{2i}^2 + 6\hat{\sigma}_{22}^2 - 4\hat{\mu}_{222} \hat{f}_{2i} - \hat{\nu}_{2222} \\ M_{66} &= \hat{f}_{3i}^4 - 6\hat{\sigma}_{33} \hat{f}_{3i}^2 + 6\hat{\sigma}_{33}^2 - 4\hat{\mu}_{333} \hat{f}_{3i} - \hat{\nu}_{3333}. \end{aligned}$$

Hierbei sind μ_{jkl} die dritten Momente aus den Elemente e_j, e_k, e_l , also ist z.B. das dritte Moment $E[e_2 e_2 e_3] = \mu_{223}$; ν_{jklm} sind entsprechend die vierten Momente. Für die Schätzer $\hat{\mu}_{jkl}$ und $\hat{\nu}_{jklm}$ können, wie in WA Stage 1 (iv) beschrieben, Annahmen über das Messmodell getroffen werden oder sie können aus den Daten geschätzt werden. Weiterhin ist

$$\hat{\mathbf{m}} = \frac{1}{n} \sum_{i=1}^n \hat{\mathbf{m}}_i = \frac{1}{n} \sum_{i=1}^n \begin{pmatrix} \hat{f}_{1i} \\ \hat{f}_{1i} \hat{f}_{2i} - \hat{\sigma}_{12} \\ \hat{f}_{1i} \hat{f}_{3i} - \hat{\sigma}_{13} \\ \hat{f}_{1i} \hat{f}_{2i} \hat{f}_{3i} - \hat{\sigma}_{23} \hat{f}_{1i} - \hat{\sigma}_{13} \hat{f}_{2i} - \hat{\sigma}_{12} \hat{f}_{3i} - \hat{\mu}_{123} \\ \hat{f}_{1i} \hat{f}_{2i}^2 - \hat{\sigma}_{22} \hat{f}_{1i} - 2\hat{\sigma}_{12} \hat{f}_{2i} - \hat{\mu}_{122} \\ \hat{f}_{1i} \hat{f}_{3i}^2 - \hat{\sigma}_{33} \hat{f}_{1i} - 2\hat{\sigma}_{13} \hat{f}_{3i} - \hat{\mu}_{133} \end{pmatrix}. \quad (\text{A.9})$$

A.2 Schätzung der Standardfehler

Nach Gleichung WA 12 ergibt sich der Schätzer für die (6×6) asymptotische Kovarianzmatrix $\hat{\mathbf{V}}$ als

$$\hat{\mathbf{V}}(\hat{\boldsymbol{\alpha}}) = \frac{1}{n} \hat{\mathbf{M}}^{-1} \hat{\boldsymbol{\Omega}} \hat{\mathbf{M}}^{-1}, \quad (\text{A.10})$$

aus deren Diagonalelementen sich die Standardfehler von $\hat{\boldsymbol{\alpha}}$ ergeben. Für $\hat{\boldsymbol{\Omega}}$ wird nach Gleichung WA 13 folgender korrigierter Schätzer verwendet:

$$\hat{\boldsymbol{\Omega}} = \frac{1}{n} \sum_{i=1}^n \boldsymbol{\lambda}_i(\hat{\boldsymbol{\alpha}}) \boldsymbol{\lambda}_i(\hat{\boldsymbol{\alpha}})' + \hat{\mathbf{C}} \hat{\mathbf{Y}} \hat{\mathbf{C}}' \quad (\text{A.11})$$

Hierbei ist $\boldsymbol{\lambda}_i(\hat{\boldsymbol{\alpha}}) = \hat{\mathbf{m}}_i - \hat{\mathbf{M}}_i \hat{\boldsymbol{\alpha}}$ und $\hat{\boldsymbol{\Omega}}^* = \frac{1}{n} \sum_{i=1}^n \boldsymbol{\lambda}_i(\hat{\boldsymbol{\alpha}}) \boldsymbol{\lambda}_i(\hat{\boldsymbol{\alpha}})'$ der unkorrigierte Schätzer für die Standardfehler. Die zweite Matrix $\hat{\mathbf{C}} \hat{\mathbf{Y}} \hat{\mathbf{C}}'$ vom Format (6×6) beinhaltet die Korrektur dafür, dass die Faktorwerte Schätzungen sind und bezieht sich konkret auf die Ableitungen in der Matrix $\hat{\mathbf{C}}$ sowie auf die asymptotische Kovarianzmatrix $\hat{\mathbf{Y}}$ der Schätzungen aus dem ersten Schritt für das Messmodell. In den folgenden Abschnitten wird zuerst die Matrix $\hat{\mathbf{C}}$ beschrieben, dann die asymptotische Kovarianzmatrix $\hat{\mathbf{Y}}$.

Die Matrix $\hat{\mathbf{C}}$

$\hat{\mathbf{C}}$ beinhaltet überwiegend Ableitungen von $\lambda_i(\boldsymbol{\alpha})$ nach $\hat{\mathbf{f}}_i$ und $\text{vech} \hat{\boldsymbol{\Sigma}}_{ee}$ an den Stellen $\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\alpha}}$. Die Erweiterung von $\hat{\mathbf{C}}$ aus Gleichung WA 14 ist gegeben durch:

$$\hat{\mathbf{C}} = \frac{1}{n} \sum_{i=1}^n \left(\left. \frac{\partial \boldsymbol{\lambda}_i(\boldsymbol{\alpha})}{\partial \hat{\mathbf{f}}_i'} \right|_{\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\alpha}}} \hat{\boldsymbol{\Gamma}}, \left. \frac{\partial \boldsymbol{\lambda}_i(\boldsymbol{\alpha})}{\partial \hat{\mathbf{f}}_i'} \right|_{\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\alpha}}} \hat{\boldsymbol{\Gamma}} (\mathbf{x}_i \otimes \mathbf{I}_r)', \mathbf{0}_{6 \times 3r}, \left. \frac{\partial \boldsymbol{\lambda}_i(\boldsymbol{\alpha})}{\partial \text{vech} \hat{\boldsymbol{\Sigma}}_{ee}'} \right|_{\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\alpha}}} \right) \quad (\text{A.12})$$

mit $\mathbf{x}_i = (y_{1i}, x_{1i}, x_{4i})'$, dem Cronecker-Produkt \otimes , $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}_0', \text{vec}(\hat{\boldsymbol{\beta}}_1)', \text{vec}(\hat{\boldsymbol{\Gamma}})')$, $r = p + q - 3$ und allen weiteren Spezifikationen von $\hat{\mathbf{C}}$ wie in WA. $\hat{\mathbf{C}}$ ist vom Format (6×48) .

Ableitung nach \mathbf{f}_i

Man erhält folgende (3×6) -Matrix der Ableitungen nach \mathbf{f}_i

$$\frac{\partial \boldsymbol{\lambda}_i(\boldsymbol{\alpha})}{\partial \hat{\mathbf{f}}_i'} = \left(\frac{\partial \boldsymbol{\lambda}_i(\boldsymbol{\alpha})}{\partial \hat{f}_{1i}}, \frac{\partial \boldsymbol{\lambda}_i(\boldsymbol{\alpha})}{\partial \hat{f}_{2i}}, \frac{\partial \boldsymbol{\lambda}_i(\boldsymbol{\alpha})}{\partial \hat{f}_{3i}} \right), \quad (\text{A.13})$$

mit

$$\frac{\partial \boldsymbol{\lambda}_i(\boldsymbol{\alpha})}{\partial \hat{f}_{1i}} = \begin{bmatrix} 1 \\ \hat{f}_{2i} \\ \hat{f}_{3i} \\ \hat{f}_{2i}\hat{f}_{3i} - \sigma_{23} \\ \hat{f}_{2i}^2 - \sigma_{22} \\ \hat{f}_{3i}^2 - \sigma_{33} \end{bmatrix}, \quad (\text{A.14})$$

$$\frac{\partial \boldsymbol{\lambda}_i(\boldsymbol{\alpha})}{\partial \hat{f}_{2i}} = \begin{bmatrix} 0 \\ \hat{f}_{1i} \\ 0 \\ \hat{f}_{1i}\hat{f}_{3i} - \hat{\sigma}_{13} \\ 2\hat{f}_{1i}\hat{f}_{2i} - 2\hat{\sigma}_{12} \\ 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 & 0 & \hat{f}_{3i} & 2\hat{f}_{2i} & 0 \\ 1 & 2\hat{f}_{2i} & \hat{f}_{3i} & \mathbf{d2}_{24} & \mathbf{d2}_{25} & \mathbf{d2}_{26} \\ 0 & \hat{f}_{3i} & 0 & \mathbf{d2}_{34} & \mathbf{d2}_{35} & 0 \\ \hat{f}_{3i} & \mathbf{d2}_{42} & \mathbf{d2}_{43} & \mathbf{d2}_{44} & \mathbf{d2}_{45} & \mathbf{d2}_{46} \\ 2\hat{f}_{2i} & \mathbf{d2}_{52} & \mathbf{d2}_{53} & \mathbf{d2}_{54} & \mathbf{d2}_{55} & \mathbf{d2}_{56} \\ 0 & \mathbf{d2}_{62} & 0 & \mathbf{d2}_{64} & \mathbf{d2}_{65} & 0 \end{bmatrix} \boldsymbol{\alpha} \quad (\text{A.15})$$

wobei

$$\begin{aligned} \mathbf{d2}_{24} &= 2\hat{f}_{2i}\hat{f}_{3i} - 2\hat{\sigma}_{23} & \mathbf{d2}_{25} &= 3\hat{f}_{2i}^2 - 3\hat{\sigma}_{22} \\ \mathbf{d2}_{26} &= \hat{f}_{3i}^2 - \hat{\sigma}_{33} & \mathbf{d2}_{34} &= \hat{f}_{3i}^2 - \hat{\sigma}_{33} \\ \mathbf{d2}_{35} &= 2\hat{f}_{2i}\hat{f}_{3i} - 2\hat{\sigma}_{23} & \mathbf{d2}_{42} &= 2\hat{f}_{2i}\hat{f}_{3i} - 2\hat{\sigma}_{23} \\ \mathbf{d2}_{43} &= \hat{f}_{3i}^2 - \hat{\sigma}_{33} & \mathbf{d2}_{44} &= 2\hat{f}_{2i}\hat{f}_{3i}^2 - 2\hat{\sigma}_{33}\hat{f}_{2i} - 4\hat{\sigma}_{23}\hat{f}_{3i} - 2\hat{\mu}_{233} \\ \mathbf{d2}_{45} &= 3\hat{f}_{2i}^2\hat{f}_{3i} - 6\hat{\sigma}_{23}\hat{f}_{2i} - 3\hat{\sigma}_{22}\hat{f}_{3i} - 3\hat{\mu}_{223} & \mathbf{d2}_{46} &= \hat{f}_{3i}^3 - 3\hat{f}_{3i}\hat{\sigma}_{33} - \hat{\mu}_{333} \\ \mathbf{d2}_{52} &= 3\hat{f}_{2i}^2 - 3\hat{\sigma}_{22} & \mathbf{d2}_{53} &= 2\hat{f}_{2i}\hat{f}_{3i} - 2\hat{\sigma}_{23} \\ \mathbf{d2}_{54} &= 3\hat{f}_{2i}^2\hat{f}_{3i} - 6\hat{\sigma}_{23}\hat{f}_{2i} - 3\hat{\sigma}_{22}\hat{f}_{3i} - 3\hat{\mu}_{223} & \mathbf{d2}_{55} &= 4\hat{f}_{2i}^3 - 12\hat{f}_{2i}\hat{\sigma}_{22} - 4\hat{\mu}_{222} \\ \mathbf{d2}_{56} &= 2\hat{f}_{2i}\hat{f}_{3i}^2 - 2\hat{\sigma}_{33}\hat{f}_{2i} - 4\hat{\sigma}_{23}\hat{f}_{3i} - 2\hat{\mu}_{233} & \mathbf{d2}_{62} &= \hat{f}_{3i}^2 - \hat{\sigma}_{33} \\ \mathbf{d2}_{64} &= \hat{f}_{3i}^3 - 3\hat{f}_{3i}\hat{\sigma}_{33} - \hat{\mu}_{333} & \mathbf{d2}_{65} &= 2\hat{f}_{2i}\hat{f}_{3i}^2 - 2\hat{\sigma}_{33}\hat{f}_{2i} - 4\hat{\sigma}_{23}\hat{f}_{3i} - 2\hat{\mu}_{233}, \end{aligned}$$

$$\frac{\partial \boldsymbol{\lambda}_i(\boldsymbol{\alpha})}{\partial \hat{f}_{3i}} = \begin{bmatrix} 0 \\ 0 \\ \hat{f}_{1i} \\ \hat{f}_{1i}\hat{f}_{2i} - \hat{\sigma}_{12} \\ 0 \\ 2\hat{f}_{1i}\hat{f}_{3i} - 2\hat{\sigma}_{13} \end{bmatrix} - \begin{bmatrix} 0 & 0 & 1 & \hat{f}_{2i} & 0 & 2\hat{f}_{3i} \\ 0 & 0 & \hat{f}_{2i} & \mathbf{d3}_{24} & 0 & \mathbf{d3}_{26} \\ 1 & \hat{f}_{2i} & 2\hat{f}_{3i} & \mathbf{d3}_{34} & \mathbf{d3}_{35} & \mathbf{d3}_{36} \\ \hat{f}_{2i} & \mathbf{d3}_{42} & \mathbf{d3}_{43} & \mathbf{d3}_{44} & \mathbf{d3}_{45} & \mathbf{d3}_{46} \\ 0 & 0 & \mathbf{d3}_{53} & \mathbf{d3}_{54} & 0 & \mathbf{d3}_{56} \\ 2\hat{f}_{3i} & \mathbf{d3}_{62} & \mathbf{d3}_{63} & \mathbf{d3}_{64} & \mathbf{d3}_{65} & \mathbf{d3}_{66} \end{bmatrix} \boldsymbol{\alpha} \quad (\text{A.16})$$

wobei

$$\begin{aligned}
\mathbf{d3}_{24} &= \hat{f}_{2i}^2 - \hat{\sigma}_{22} & \mathbf{d3}_{26} &= 2\hat{f}_{2i}\hat{f}_{3i} - 2\hat{\sigma}_{23} \\
\mathbf{d3}_{34} &= 2\hat{f}_{2i}\hat{f}_{3i} - 2\hat{\sigma}_{23} & \mathbf{d3}_{35} &= \hat{f}_{2i}^2 - \hat{\sigma}_{22} \\
\mathbf{d3}_{36} &= 3\hat{f}_{3i}^2 - 3\hat{\sigma}_{33} & \mathbf{d3}_{42} &= \hat{f}_{2i}^2 - \hat{\sigma}_{22} \\
\mathbf{d3}_{43} &= 2\hat{f}_{2i}\hat{f}_{3i} - 2\hat{\sigma}_{23} & \mathbf{d3}_{44} &= 2\hat{f}_{2i}^2\hat{f}_{3i} - 2\hat{\sigma}_{22}\hat{f}_{3i} - 4\hat{\sigma}_{23}\hat{f}_{2i} - 2\hat{\mu}_{223} \\
\mathbf{d3}_{45} &= \hat{f}_{2i}^3 - 3\hat{f}_{2i}\hat{\sigma}_{22} - \hat{\mu}_{222} & \mathbf{d3}_{46} &= 3\hat{f}_{2i}\hat{f}_{3i}^2 - 6\hat{\sigma}_{23}\hat{f}_{3i} - 3\hat{\sigma}_{33}\hat{f}_{2i} - 3\hat{\mu}_{233} \\
\mathbf{d3}_{53} &= \hat{f}_{2i}^2 - \hat{\sigma}_{22} & \mathbf{d3}_{54} &= \hat{f}_{2i}^3 - 3\hat{f}_{2i}\hat{\sigma}_{22} - \hat{\mu}_{222} \\
\mathbf{d3}_{56} &= 2\hat{f}_{2i}^2\hat{f}_{3i} - 2\hat{\sigma}_{22}\hat{f}_{3i} - 4\hat{\sigma}_{23}\hat{f}_{2i} - 2\hat{\mu}_{223} & \mathbf{d3}_{62} &= 2\hat{f}_{2i}\hat{f}_{3i} - 2\hat{\sigma}_{23} \\
\mathbf{d3}_{63} &= 3\hat{f}_{3i}^2 - 3\hat{\sigma}_{33} & \mathbf{d3}_{64} &= 3\hat{f}_{2i}\hat{f}_{3i}^2 - 6\hat{\sigma}_{23}\hat{f}_{3i} - 3\hat{\sigma}_{33}\hat{f}_{2i} - 3\hat{\mu}_{233} \\
\mathbf{d3}_{65} &= 2\hat{f}_{2i}^2\hat{f}_{3i} - 2\hat{\sigma}_{22}\hat{f}_{3i} - 4\hat{\sigma}_{23}\hat{f}_{2i} - 2\hat{\mu}_{223} & \mathbf{d3}_{66} &= 4\hat{f}_{3i}^3 - 12\hat{f}_{3i}\hat{\sigma}_{33} - 4\hat{\mu}_{333}.
\end{aligned}$$

Ableitung nach $\text{vech}\hat{\Sigma}_{ee}$

Als nächstes ergibt sich für die Ableitung nach $\text{vech}\hat{\Sigma}_{ee}$ eine (6×6) -Matrix

$$\begin{aligned}
\frac{\partial \boldsymbol{\lambda}_i(\boldsymbol{\alpha})}{\partial \text{vech}\hat{\Sigma}'_{ee}} &= \left(\frac{\partial \boldsymbol{\lambda}_i(\boldsymbol{\alpha})}{\partial \hat{\sigma}_{11}}, \frac{\partial \boldsymbol{\lambda}_i(\boldsymbol{\alpha})}{\partial \hat{\sigma}_{12}}, \frac{\partial \boldsymbol{\lambda}_i(\boldsymbol{\alpha})}{\partial \hat{\sigma}_{13}}, \frac{\partial \boldsymbol{\lambda}_i(\boldsymbol{\alpha})}{\partial \hat{\sigma}_{22}}, \frac{\partial \boldsymbol{\lambda}_i(\boldsymbol{\alpha})}{\partial \hat{\sigma}_{23}}, \frac{\partial \boldsymbol{\lambda}_i(\boldsymbol{\alpha})}{\partial \hat{\sigma}_{33}} \right) \\
&= \begin{pmatrix} 0 & 0 & 0 & \omega_{11} & \omega_{12} & \omega_{22} \\ 0 & -1 & 0 & \mathbf{d4}_{24} & \mathbf{d4}_{25} & \omega_{22}\hat{f}_{2i} \\ 0 & 0 & -1 & \omega_{11}\hat{f}_{3i} & \mathbf{d4}_{35} & \mathbf{d4}_{36} \\ 0 & -\hat{f}_{3i} & -\hat{f}_{2i} & \mathbf{d4}_{44} & \mathbf{d4}_{45} & \mathbf{d4}_{46} \\ 0 & -2\hat{f}_{2i} & 0 & \mathbf{d4}_{54} & \mathbf{d4}_{55} & \mathbf{d4}_{56} \\ 0 & 0 & -2\hat{f}_{3i} & \mathbf{d4}_{64} & \mathbf{d4}_{65} & \mathbf{d4}_{66} \end{pmatrix} \quad (\text{A.17})
\end{aligned}$$

mit

$$\begin{aligned}
\mathbf{d4}_{24} &= \gamma_1 + \omega_{12}\hat{f}_{3i} + 3\omega_{11}\hat{f}_{2i} \\
\mathbf{d4}_{25} &= \gamma_2 + 2\omega_{12}\hat{f}_{2i} + 2\omega_{22}\hat{f}_{3i} \\
\mathbf{d4}_{35} &= \gamma_1 + 2\omega_{12}\hat{f}_{3i} + 2\omega_{11}\hat{f}_{2i} \\
\mathbf{d4}_{36} &= \gamma_2 + \omega_{12}\hat{f}_{2i} + 3\omega_{22}\hat{f}_{3i} \\
\mathbf{d4}_{44} &= \gamma_1\hat{f}_{3i} + \omega_{12}\hat{f}_{3i}^2 - 2\omega_{12}\hat{\sigma}_{33} + 3\omega_{11}\hat{f}_{2i}\hat{f}_{3i} - 6\omega_{11}\hat{\sigma}_{23} \\
\mathbf{d4}_{45} &= \alpha_0 - \hat{f}_{1i} + 2\gamma_1\hat{f}_{2i} + 2\gamma_2\hat{f}_{3i} + 4\omega_{12}\hat{f}_{2i}\hat{f}_{3i} - 8\omega_{12}\hat{\sigma}_{23} + 3\omega_{11}\hat{f}_{2i}^2 - 6\omega_{11}\hat{\sigma}_{22} \\
&\quad + 3\omega_{22}\hat{f}_{3i}^2 - 6\omega_{22}\hat{\sigma}_{33} \\
\mathbf{d4}_{46} &= \gamma_2\hat{f}_{2i} + \omega_{12}\hat{f}_{2i}^2 - 2\omega_{12}\hat{\sigma}_{22} + 3\omega_{22}\hat{f}_{2i}\hat{f}_{3i} - 6\omega_{22}\hat{\sigma}_{23}
\end{aligned}$$

$$\begin{aligned}
\mathbf{d4}_{54} &= \alpha_0 - \hat{f}_{1i} + 3\gamma_1 \hat{f}_{2i} + \gamma_2 \hat{f}_{3i} + 3\omega_{12} \hat{f}_{2i} \hat{f}_{3i} - 6\omega_{12} \hat{\sigma}_{23} + 6\omega_{11} \hat{f}_{2i}^2 - 12\omega_{11} \hat{\sigma}_{22} \\
&\quad + \omega_{22} \hat{f}_{3i}^2 - 2\omega_{22} \hat{\sigma}_{33} \\
\mathbf{d4}_{55} &= 2\gamma_2 \hat{f}_{2i} + 3\omega_{12} \hat{f}_{2i}^2 - 6\omega_{12} \hat{\sigma}_{22} + 4\omega_{22} \hat{f}_{2i} \hat{f}_{3i} - 8\omega_{22} \hat{\sigma}_{23} \\
\mathbf{d4}_{56} &= \omega_{22} (\hat{f}_{2i}^2 - 2\hat{\sigma}_{22}) \\
\mathbf{d4}_{64} &= \omega_{11} (\hat{f}_{3i}^2 - 2\hat{\sigma}_{33}) \\
\mathbf{d4}_{65} &= 2\gamma_1 \hat{f}_{3i} + 3\omega_{12} \hat{f}_{3i}^2 - 6\omega_{12} \hat{\sigma}_{33} + 4\omega_{11} \hat{f}_{2i} \hat{f}_{3i} - 8\omega_{11} \hat{\sigma}_{23} \\
\mathbf{d4}_{66} &= \alpha_0 - \hat{f}_{1i} + \gamma_1 \hat{f}_{2i} + 3\gamma_2 \hat{f}_{3i} + 3\omega_{12} \hat{f}_{2i} \hat{f}_{3i} - 6\omega_{12} \hat{\sigma}_{23} + \omega_{11} \hat{f}_{2i}^2 - 2\omega_{11} \hat{\sigma}_{22} \\
&\quad + 6\omega_{22} \hat{f}_{3i}^2 - 12\omega_{22} \hat{\sigma}_{33}.
\end{aligned}$$

Ein Schätzer für Υ

Die asymptotische Kovarianzmatrix $\hat{\Upsilon}$ vom Format (48×48) ist vollständig im Appendix von WA beschrieben und benötigt keine Erweiterung, da sie auf demselben Messmodell wie in WA beruht (s. Gleichung A.2) und ist gegeben als

$$\hat{\Upsilon} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ \frac{\partial \text{vech} \hat{\Sigma}_{ee}}{\partial \hat{\beta}_0} & \frac{\partial \text{vech} \hat{\Sigma}_{ee}}{\partial (\text{vec} \hat{\beta}_1)'} & \frac{\partial \text{vech} \hat{\Sigma}_{ee}}{\partial (\text{diag} \hat{\psi})'} \end{bmatrix} \widehat{\text{Cov}} \begin{bmatrix} \hat{\beta}_0 \\ \text{vec} \hat{\beta}_1 \\ \text{diag} \hat{\psi} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & \frac{\partial \text{vech} \hat{\Sigma}_{ee}}{\partial \hat{\beta}_0} \\ 0 & 1 & 0 & \frac{\partial \text{vech} \hat{\Sigma}_{ee}}{\partial (\text{vec} \hat{\beta}_1)'} \\ 0 & 0 & 0 & \frac{\partial \text{vech} \hat{\Sigma}_{ee}}{\partial (\text{diag} \hat{\psi})'} \end{bmatrix} \quad (\text{A.18})$$

Die Kovarianzmatrix $\widehat{\text{Cov}}(\hat{\beta}_0', \text{vec} \hat{\beta}_1', \text{diag} \hat{\psi}')'$ erhält man als asymptotische Schätzung aus der CFA im ersten Schritt. Die Kovarianzmatrix ist vom Format (33×33) . Die einzelnen Ableitungen sind auf S. 63 in WA beschrieben¹. Alle Ableitungen nach $\hat{\beta}_0$ sind 0. Jede Ableitung (nach den 18 Elementen von $\text{vec} \hat{\beta}_1$ und den einzelnen 9 Elementen von $\hat{\psi}$) ergeben einen Vektor vom Format (6×1) (für jedes Element von $\text{vech} \hat{\Sigma}_{ee}$).

¹Die Ableitung in WA von $\hat{\Sigma}_{ee}$ nach $\hat{\psi}_{22l}$ enthält einen Tippfehler (unterstrichen); korrekt wäre

$$\frac{\partial \hat{\Sigma}_{ee}}{\partial \hat{\psi}_{22l}} = \mathbf{E}_{ll}^{k \times k} - \mathbf{E}_{ll}^{k \times k} \hat{\beta}_1' \hat{\Sigma}_{vv}^{-1} \hat{\beta}_1 \hat{\psi}_{22} - \hat{\psi}_{22} \hat{\beta}_1' \hat{\Sigma}_{vv}^{-1} \hat{\beta}_1 \mathbf{E}_{ll}^{k \times k} + \hat{\psi}_{22} \hat{\beta}_1' \hat{\Sigma}_{vv}^{-1} \hat{\mathbf{b}}_l \hat{\mathbf{b}}_l' \hat{\Sigma}_{vv}^{-1} \hat{\beta}_1 \hat{\psi}_{22}. \quad (\text{A.19})$$

Anhang B

Konfidenzintervalle für die Prognose von \hat{y}_t

Konfidenzintervalle für das HGM

Das Adhoc-Konfidenzintervall für die Prognose von \hat{y}_t gegeben y_1 basiert auf der bedingten Varianz $\sigma_{y_t|y_1}^2$ und dem bedingten Erwartungswert $E(y_t|y_1)$. Für ein HGM *ohne Kovariate* w ergibt sich dies als

$$\sigma_{y_t|y_1}^2 = [(1 + \lambda_t\beta_{11})^2 + (\lambda_t\gamma_1)^2]\sigma_{\bullet}^2 + \lambda_t^2(\gamma_0 + \gamma_1 E_{\bullet})^2 + \lambda_t^2\psi_{33} + \theta_{tt} \quad (\text{B.1})$$

$$E(y_t|y_1) = \lambda_t\beta_{10} + (1 + \lambda_t\beta_{11})E_{\bullet} \quad (\text{B.2})$$

mit der bedingten Varianz und dem bedingten Erwartungswert des Slope-Faktors

$$\begin{aligned} \sigma_{\bullet}^2 &= V(\eta_0|y_1) = \frac{\psi_{00}\theta_{11}}{\psi_{00} + \theta_{11}} \\ E_{\bullet} &= E(\eta_0|y_1) = \beta_{00} + \frac{\psi_{00}}{\psi_{00} + \theta_{11}}(y_1 - \beta_{00}). \end{aligned} \quad (\text{B.3})$$

Für ein HGM *mit einer Kovariaten* w erweitert sich die Gleichung für die bedingte Varianz und den bedingten Erwartungswert für \hat{y}_t gegeben y_1 und w zu

$$\begin{aligned} \sigma_{y_t|y_1,w}^2 &= [(1 + \lambda_t\beta_{11})^2 + (\lambda_t\gamma_1)^2]\sigma_{\circ}^2 + \lambda_t^2(\gamma_0 + \gamma_1 E_{\circ} + \gamma_2 w)^2 \\ &\quad + \lambda_t^2\psi_{33} + \theta_{tt} \end{aligned} \quad (\text{B.4})$$

$$E(y_t|y_1, w) = \lambda_t\beta_{10} + (1 + \lambda_t\beta_{11})E_{\circ} + \lambda_t\beta_{12}w \quad (\text{B.5})$$

mit der bedingten Varianz und dem bedingten Erwartungswert des Slope-Faktors

$$\sigma_{\circ}^2 = V(\eta_0|y_1, w) = \frac{(2\beta_{02}^2 V(w) + \psi_{00})\theta_{11}}{\psi_{00} + \theta_{11}} \quad (\text{B.6})$$

$$E_{\circ} = E(\eta_0|y_1, w) = \beta_{00} + \frac{\psi_{00}}{\psi_{00} + \theta_{11}}(y_1 - \beta_{00}) + \frac{\beta_{02}\theta_{11}}{\psi_{00} + \theta_{11}}w. \quad (\text{B.7})$$

Konfidenzintervalle für das LGM

Die bedingte Varianz für ein LGM *ohne Kovariate* ($\sigma_{y_t|y_1}^2$) bzw. *mit einer Kovariaten* ($\sigma_{y_t|y_1, w}^2$) ergibt sich als

$$\sigma_{y_t|y_1}^2 = (1 + \lambda_t\beta_{11})^2\sigma_{\bullet}^2 + \lambda_t^2\psi_{11} + \theta_{tt} \text{ bzw.} \quad (\text{B.8})$$

$$\sigma_{y_t|y_1, w}^2 = (1 + \lambda_t\beta_{11})^2\sigma_{\circ}^2 + \lambda_t^2\psi_{11} + \theta_{tt} \quad (\text{B.9})$$

mit der homogenen (Residual)-Varianz des Slope-Faktors ψ_{11} . Der bedingte Erwartungswert ist identisch zu dem des HGM in Gleichung B.2 bzw. B.5. Das 90%-Konfidenzintervall in Abbildung 4.3 wurden berechnet als $E(y_t|y_1) \pm 1.645 \cdot \sigma_{y_t|y_1}$.

Anhang C

Schrift (A)

Brandt, H., Kelava, A. & Klein, A. G. (*in Druck*). A Simulation study comparing recent approaches for the estimation of nonlinear effects in SEM under the condition of non-normality. *Structural Equation Modeling*.

Anhang D

Schrift (B)

Kelava, A., Nagengast, B., & **Brandt, H.** (*in Druck*). A nonlinear structural equation mixture modeling approach for non-normally distributed latent predictor variables. *Structural Equation Modeling*.

Anhang E

Schrift (C)

Brandt, H., Klein, A. G. & Gäde, J. (*in Revision*). An ML estimator for heterogeneous growth curve models. *Multivariate Behavioral Research*.

Anhang F

Selbstständigkeitserklärung

Hiermit erkläre ich, Holger Brandt, die vorliegende Dissertation mit dem Titel „Schätzung latenter nichtlinearer Effekte in Quer- und Längsschnittstudien“ eigenständig und ausschließlich unter Verwendung der angegebenen Hilfsmittel angefertigt zu haben. Alle Quellen sind als solche kenntlich gemacht. Die vorliegende Arbeit wurde zuvor weder in dieser noch anderer Form als Dissertationsarbeit zur Begutachtung vorgelegt. Das beantragte Promotionsverfahren stellt das erste beantragte Promotionsverfahren dar.

Frankfurt am Main, 04.03.2013

Holger Brandt