

Multikollinearität in nicht-linearen latenten Strukturgleichungsmodellen

Dem Fachbereich Psychologie und Sportwissenschaften der
Goethe Universität Frankfurt am Main
zur Erlangung des akademischen Grades eines
Dr. phil. nat.

eingereichte Dissertation

von
Herrn Dipl.-Psych. Augustin Kelava
aus
Frankfurt am Main

Datum der Einreichung: 02.12.2008

Dekan: Prof. Dr. H. Moosbrugger

Erstgutachter: Prof. Dr. H. Moosbrugger

Zweitgutachter: Prof. Dr. V. Hodapp

Datum der Disputation: 11.03.2009

Zusammenfassung

Die vorliegende Arbeit befasst sich mit der übergeordneten Frage nach dem Einfluss der Multikollinearität bei der Spezifikation und Schätzung von latenten nicht-linearen Effekten in Strukturgleichungsmodellen.

Da das Multikollinearitätsproblem in latenten nicht-linearen Strukturgleichungsmodellen bisher in der Literatur kaum Beachtung gefunden hat, wird als erstes die Frage beantwortet, welche Relevanz das Vorliegen von Multikollinearität in nicht-linearen Strukturgleichungsmodellen hat. Dabei wird aufgezeigt, dass Multikollinearität von Prädiktoren nicht nur bei linearen Strukturgleichungsmodellen ein Problem darstellt, sondern vor allem bei nicht-linearen Strukturgleichungsmodellen aufgrund des Vorhandenseins nicht-linearer Terme. Da nicht-lineare Terme untereinander in einem höheren Zusammenhang stehen können als lineare Terme eines Strukturmodells, kann es zu Scheineffekten und zu verzerrten Schätzungen der nicht-linearen Effekte kommen. Ein Teststärkeverlust tritt immer ein.

Als zweites wird gezeigt, dass die am stärksten vertretene Ansatzklasse zur simultanen Schätzung von Interaktionseffekten und quadratischen Effekten, nämlich die Klasse der Produkt-Indikator-Ansätze, seit ihrer Erstpublikation durch Kenny und Judd (1984) *Spezifikationsfehler* aufweist, welche in ähnlicher Form auch in jüngeren Publikationen enthalten sind (z.B. bei Lee, Song & Poon, 2004). Spezifikationsfehler machen die ganze Ansatzklasse für Multikollinearität „anfällig“, so dass es theoretisch zu verzerrten Schätzungen oder Scheineffekten kommen muss. Es wird gezeigt, wie man die Spezifikationsfehler korrigieren kann, d.h. wie man die bestehenden Ansätze erweitern muss. *Da die latente Prädiktorkorrelation Bestandteil der Spezifikationsfehler ist, stellt sich die Frage, wie groß die Verzerrungen sind, die sich bei der Anwendung der (unkorrigierten) Produkt-Indikator-Ansätze ergeben, wenn die Multikollinearität steigt.* Um diese Frage zu beantworten, wird eine Simulationsstudie durchgeführt. Sie zeigt, dass der Zwei-Stufen-Ansatz von Ping (1996) verzerrte Schätzungen der nicht-linearen Effekte produziert. Dieser Ansatz ist ein Vertreter der Produkt-Indikator-Ansätze und enthält die besagten Spezifikationsfehler. In der Simulationsstudie wird außerdem gezeigt, dass die korrekte Erweiterung des Ansatzes von Jöreskog und Yang (1996) zu erwartungstreuen Schätzungen führt.

Als drittes wird der Frage nachgegangen werden, welche theoretischen und empirischen Unterschiede zwischen den neuen verteilungsanalytischen Ansätzen LMS (Klein & Moosbrugger, 2000) und QML (Klein & Muthén, 2007), einerseits, und den klassischen Produkt-Indikator-Ansätzen, andererseits, zu erwarten sind. Dazu werden die theoretischen Eigenschaften der verteilungsanalytischen Verfahren, LMS

und QML, untereinander und mit der Klasse der Produkt-Indikator-Ansätze verglichen. Ihre theoretischen Eigenschaften werden beim Vorliegen von Multikollinearität betrachtet. Illustriert werden die theoretischen Unterschiede anhand einer Simulationsstudie, die die Ansätze hinsichtlich ihrer Schätzeigenschaften bei steigender Multikollinearität miteinander vergleicht. In der Simulationsstudie werden die theoretisch herausgearbeiteten Vorteile der verteilungsanalytischen Verfahren, LMS und QML, bestätigt.

In der Zusammenschau der theoretischen Überlegungen und der Ergebnisse aus den Simulationsstudien zeigt sich, dass Multikollinearität für alle Ansätze zur simultanen Schätzung multipler nicht-linearer Effekte eine Herausforderung ist. Die forschungspraktische und inhaltliche Relevanz des Multikollinearitätsproblems wird anhand eines Beispiels veranschaulicht. Für in jüngster Zeit diskutierte komplexere Modelle (z.B. moderierte Mediatormodelle) wird Multikollinearität ein Kernproblem sein, dem sich alle Ansätze werden stellen müssen oder für das neue Ansätze entwickelt werden müssen.

Schlagwörter: Multikollinearität, Kollinearität, nicht-lineare Regressionsmodelle, Strukturgleichungsmodelle, SEM, latente Variablen, Interaktion, Wechselwirkung, Interaktionseffekte, Wechselwirkungseffekte, quadratischer Effekt, Restriktionen, Constraints.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Begriffsbestimmung	1
1.1.1	Interaktionseffekt	1
1.1.2	Quadratischer Effekt	3
1.1.3	Interaktionseffekt und zwei quadratische Effekte	3
1.2	Ein inhaltliches Beispiel für ein Modell mit einem Interaktionseffekt und zwei quadratischen Effekten	5
1.3	Nicht-lineare Strukturgleichungsmodelle	5
1.4	Fragestellungen der Arbeit	6
1.5	Übersicht über die Arbeit	9
2	Ansätze zur Schätzung nicht-linearer Effekte in latenten Struktur- gleichungsmodellen	10
2.1	Produkt-Indikator-Ansätze	10
2.1.1	Der Kenny und Judd-Ansatz	10
2.1.2	Weiterentwicklungen	12
2.2	Verteilungsanalytische Ansätze	16
2.2.1	Der Latent Moderated Structural Equations Ansatz (LMS)	16
2.2.2	Der Quasi-Maximum Likelihood Ansatz (QML)	17
2.3	Sonstige Ansätze	17
2.4	Schätzeigenschaften der Verfahren	19
3	Zum Problem der Multikollinearität und seiner Relevanz für nicht- lineare Strukturgleichungsmodelle	21
3.1	Zum Begriff der Multikollinearität	21
3.2	Folgen, Diagnostik und Abhilfen	22
3.2.1	Allgemeine Folgen der Multikollinearität	22

3.2.2	Multikollinearitätsdiagnostik	23
3.2.3	Abhilfen bei Vorliegen von Multikollinearität	23
3.3	Zur Bedeutung der Multikollinearität in nicht-linearen Strukturgleichungsmodellen	24
3.3.1	Nicht-essentielle Multikollinearität	25
3.3.2	Essentielle Multikollinearität	25
3.3.3	Implikationen des Multikollinearitätsproblems	26
4	Spezifikationsfehler der Produkt-Indikator-Ansätze	28
4.1	Zu den Spezifikationsfehlern der Produkt-Indikator-Ansätze und ihrer Korrektur	29
4.1.1	Zu den Spezifikationsfehlern bei Kenny und Judd (1984) und ihrer Korrektur	29
4.1.2	Zu den Spezifikationsfehlern bei Lee et al. (2004) und ihrer Korrektur	31
4.2	Eine Simulationsstudie zur Veranschaulichung des Effekts der Spezifikationsfehler bei steigender Multikollinearität	31
4.2.1	Methode	32
4.2.2	Ergebnisse	34
4.2.3	Implikationen	37
5	Zum theoretischen und empirischen Vergleich der verteilungsanalytischen Ansätze LMS und QML mit den Produkt-Indikator-Ansätzen	39
5.1	Theoretischer Vergleich der LMS und QML Ansätze	40
5.1.1	Funktionsweise von LMS	40
5.1.2	Funktionsweise von QML	40
5.1.3	Unterschiede zwischen den verteilungsanalytischen Ansätzen und ihre Konsequenzen	42
5.2	Theoretischer Vergleich der verteilungsanalytischen Ansätze und der Produkt-Indikator-Ansätze	43
5.2.1	Theoretische Unterschiede	43
5.2.2	Konsequenzen der Unterschiede	44
5.3	Eine Simulationsstudie zur Veranschaulichung der empirischen Leistungsfähigkeit von verteilungsanalytischen Ansätzen und Produkt-Indikator-Ansätzen bei Multikollinearität	46

5.3.1	Methode	46
5.3.2	Ergebnisse	47
5.3.3	Implikationen	50
6	Diskussion und Ausblick	51
6.1	Zusammenführung der Kapitel und Antworten auf die Fragestellungen	51
6.2	„Was sind schon 10 Prozent Verzerrung?“ – Ein Beispiel zur Veranschaulichung der Folgen von verzerrten Schätzungen nicht-linearer Effekte	53
6.3	Bewertung der eigenen Vorgehensweise und Ausblick	56
6.3.1	Mehrwert und Grenzen der Arbeit	56
6.3.2	Ausblick	58
	Literaturverzeichnis	59
	A Schrift (A)	65
	B Schrift (B)	66
	C Schrift (C)	67
	D Symbolverzeichnis	68
	E Selbständigkeitserklärung	70
	F Lebenslauf	71

Abbildungsverzeichnis

1.1	Pfaddiagramm einer einfachen Interaktion zwischen Variable X und Variable Z	2
1.2	Graph eines Interaktionseffektes mit $\beta_0 = 0$, $\beta_1 = \beta_2 = 0.3$ und $\omega_{12} = 0.175$	2
1.3	Pfaddiagramm eines einfachen quadratischen Effektes	3
1.4	Graph eines quadratischen Effektes mit $\beta_0 = 0$, $\beta_1 = 0.3$ und $\omega_{11} = 0.158$	4
1.5	Graph eines Interaktionseffektes und zweier quadratischer Effekte mit $\beta_0 = 0$, $\beta_1 = \beta_2 = 0.3$, $\omega_{12} = 0.175$ und $\omega_{11} = \omega_{22} = 0.158$	4
1.6	Pfaddiagramm eines latenten Strukturmodells mit einem Interaktionseffekt und zwei quadratischen Effekten (ohne Messmodelle)	7
2.1	Interaktionsmodell nach Kenny und Judd (1984) mit den linearen und nicht-linearen Messmodellen	11
3.1	Veranschaulichung der Multikollinearitätsproblematik im Kontext unreliabler Prädiktoren (Attenuation)	24
3.2	Multikollinearität im Strukturmodell: Korrelationen der latenten nicht-linearen Terme untereinander, wenn die Prädiktorkorrelation .50 beträgt	27
4.1	Simulationsstudie aus Schrift (B): Mittlere Parameterschätzungen eines Modells mit einem Interaktionseffekt und zwei quadratischen Effekten bei steigender Multikollinearität. Die horizontalen Linien entsprechen den wahren Parameterwerten.	37
5.1	LMS Ansatz: Veranschaulichung der Y_1 -Indikatorverteilung als finite Mischverteilung von drei bedingten Normalverteilungen (entnommen aus Schrift (C))	41

6.1	Bildungserwartungsbeispiel von Ganzach (1997): Modell mit einem Interaktionseffekt und zwei quadratischen Effekten	54
6.2	Bildungserwartungsbeispiel von Ganzach (1997): Modell mit einem Interaktionseffekt	55

Tabellenverzeichnis

2.1	Überblick über die nicht-linearen Parameterrestriktionen beim Produkt-Indikator-Ansatz.	13
4.1	Simulationsstudie aus Schrift (B): Wahre Populationsmodelle: Variation der erklärten Kriteriumsvarianz (in %) der nicht-linearen Effekte (ω_{12} , ω_{11} und ω_{22}) und der Multikollinearitätsbedingungen.	33
4.2	Simulationsstudie aus Schrift (B): Ergebnisse der wahren Modelle mit einem Interaktionseffekt und zwei quadratischen Effekten.	36
5.1	Simulationsstudie aus Schrift (C): Drei wahre nicht-lineare Modelle: Variation der Multikollinearität.	48
5.2	Simulationsstudie aus Schrift (C): Simulationsergebnisse für drei Modelle mit einem Interaktionseffekt und zwei quadratischen Effekten im Kontext ansteigender Multikollinearität.	49

Kapitel 1

Einleitung

1.1 Begriffsbestimmung

In den empirischen Sozialwissenschaften finden sich unzählige Beispiele für nicht-lineare Interaktionshypothesen, für nicht-lineare quadratische Hypothesen oder für beide Hypothesenformen (e.g. Ajzen, 1987; Cronbach & Snow, 1977; Karasek, 1979; Lusch & Brown, 1996; Snyder & Tanke, 1976).

1.1.1 Interaktionseffekt

Ein *Interaktionseffekt* liegt dann vor, wenn sich der quantifizierbare Zusammenhang zwischen zwei Variablen (z.B. X und Y) in Abhängigkeit von der Ausprägung einer Drittvariablen (z.B. Z) verändert. Die Drittvariable (hier Z) wird auch Moderatorvariable genannt¹. Abbildung 1.1 veranschaulicht einen Interaktionseffekt graphisch als ein einfaches Pfaddiagramm, bei dem der Zusammenhang von zwei Variablen X und Y (als Pfeil zwischen den Variablen X und Y) von der Variable Z verändert wird.

Es wird angenommen, dass sich der Zusammenhang zwischen X und Y sowie zwischen Z und Y in Form einer Regressionsgleichung beschreiben lässt:

$$Y = \beta_0 + \beta_1 X + \beta_2 Z + \omega_{12} XZ + \epsilon \quad (1.1)$$

wobei β_0 eine Regressionskonstante, β_1 und β_2 Regressionskoeffizienten der Variablen X und Z sowie ϵ der Vorhersagefehler der Regression sind. Es liegt dann ein

¹Da der statistische Begriff der Moderation (oder auch Wechselwirkung) konzeptuell identisch mit dem Begriff des Interaktion ist, wird im weiteren Verlauf dieser Arbeit nur der Begriff der Interaktion verwendet.

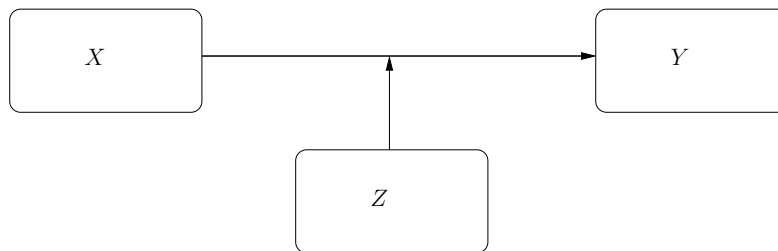


Abbildung 1.1: Pfaddiagramm einer einfachen Interaktion zwischen Variable X und Variable Z

Interaktionseffekt vor, wenn der Koeffizient $\omega_{12} \neq 0$ ist. Er beschreibt den Einfluss des nicht-linearen Produktterms XZ auf Y , der über die linearen Einflüsse β_1 und β_2 der beiden Prädiktoren X und Z hinausgeht. Neben der einfachen Veranschaulichung anhand eines Pfaddiagramms lässt sich dies auch räumlich illustrieren (s. Abbildung 1.2).

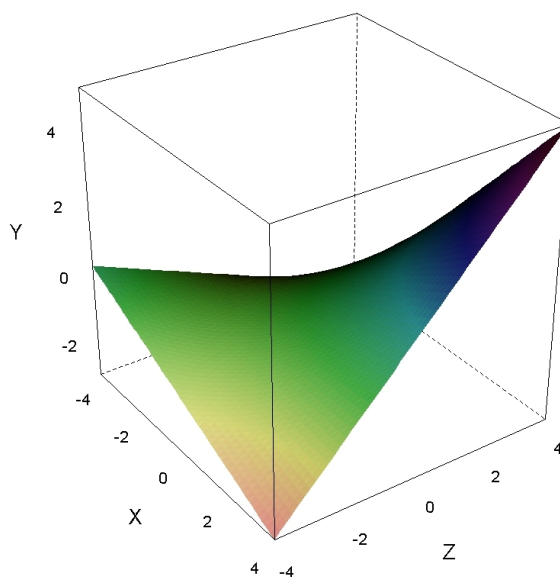


Abbildung 1.2: Graph eines Interaktionseffektes mit $\beta_0 = 0$, $\beta_1 = \beta_2 = 0.3$ und $\omega_{12} = 0.175$

Wie man Abbildung 1.2 entnehmen kann, wird die Zusammenhangsebene zwischen den Variablen X und Y sowie zwischen den Variablen Z und Y gekrümmt. ω_{12} repräsentiert das Ausmaß der Krümmung.

1.1.2 Quadratischer Effekt

Ein *quadratischer Effekt* ist als Spezialfall des Interaktionseffektes zu betrachten. Hierbei beeinflusst eine Variable X in Abhängigkeit von ihrer eigenen Ausprägung ihren Zusammenhang zu einer Variable Y (siehe Abbildung 1.3).

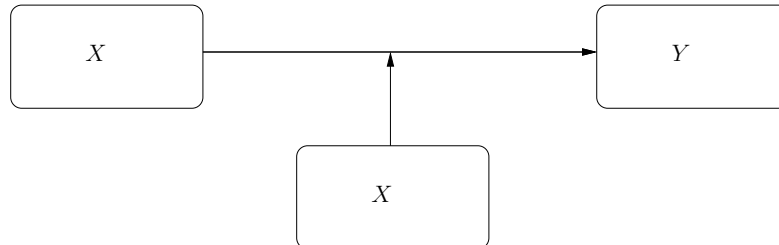


Abbildung 1.3: Pfaddiagramm eines einfachen quadratischen Effektes

Man beschreibt einen solchen Zusammenhang mit der folgenden Regressionsgleichung:

$$Y = \beta_0 + \beta_1 X + \omega_{11} X^2 + \epsilon \quad (1.2)$$

Es liegt ein quadratischer Effekt genau dann vor, wenn der Koeffizient $\omega_{11} \neq 0$ ist. ω_{11} repräsentiert den Einfluss des nicht-linearen Produktterms X^2 auf Y , der über den linearen Einfluss β_1 des Prädiktors X hinausgeht.

Neben der Darstellung anhand eines Pfaddiagramms lässt sich ein quadratischer Effekt als U-förmige Beziehung in einem Graphen veranschaulichen (s. Abbildung 1.4).

1.1.3 Interaktionseffekt und zwei quadratische Effekte

Liegt der Zusammenhang zwischen X und Y sowie zwischen Z und Y in Form eines Interaktionseffektes und zweier quadratischer Effekte vor, so folgt der Zusammenhang der Regressionsgleichung:

$$Y = \beta_0 + \beta_1 X + \omega_{12} XZ + \omega_{11} X^2 + \omega_{22} Z^2 + \epsilon \quad (1.3)$$

Hier kann man sich die im Raum aufgespannte Zusammenhangsebene wie eine „Hängematte“ vorstellen, die eine zusätzliche Rotation erfahren hat (s. Abbildung 1.5).

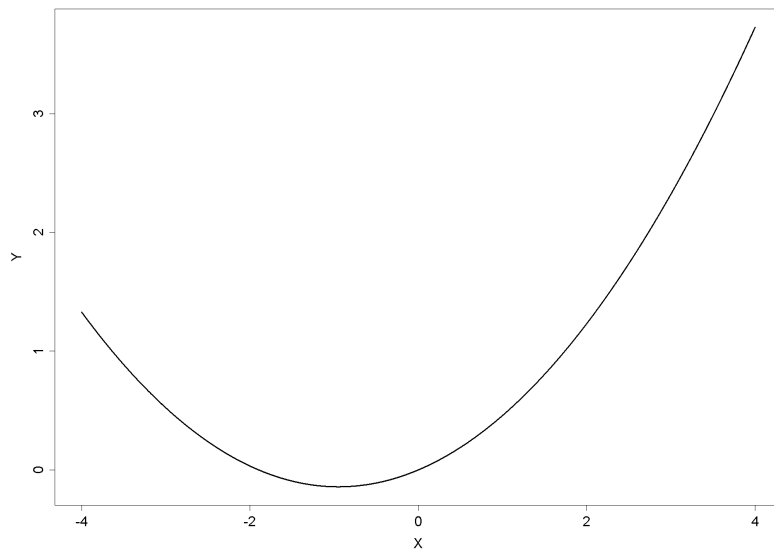


Abbildung 1.4: Graph eines quadratischen Effektes mit $\beta_0 = 0$, $\beta_1 = 0.3$ und $\omega_{11} = 0.158$

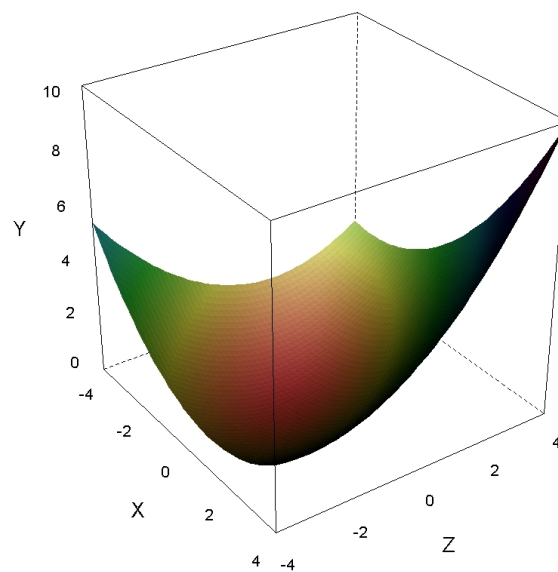


Abbildung 1.5: Graph eines Interaktionseffektes und zweier quadratischer Effekte mit $\beta_0 = 0$, $\beta_1 = \beta_2 = 0.3$, $\omega_{12} = 0.175$ und $\omega_{11} = \omega_{22} = 0.158$

1.2 Ein inhaltliches Beispiel für ein Modell mit einem Interaktionseffekt und zwei quadratischen Effekten

Ein anschauliches Beispiel für ein Modell mit einem Interaktionseffekt und zwei quadratischen Effekten findet sich bei Ganzach (1997). Er testet in seiner Arbeit die Hypothese einer komplexen interaktiven und quadratischen Beziehung zwischen der elterlichen Bildungserwartung an ihr Kind und den tatsächlichen Bildungsniveaus der Eltern (jeweils in Schuljahren). Ganzach zeigt darin, dass die Erwartungen der Eltern an das Bildungsniveau des Kindes hoch sind, wenn zumindest ein Elternteil ein hohes Bildungsniveau hat. Dabei ist es unerheblich, ob beide ein hohes Bildungsniveau haben. Für eine hohe elterliche Bildungserwartung an das Kind reicht es also aus, wenn ein Elternteil ein hohes Bildungsniveau hat. Im nicht-linearen Modell wird dieser kompensatorische Zusammenhang durch zwei positive quadratische Effekte (für das jeweilige Bildungsniveau jedes Elternteils) und einen negativen Interaktionseffekt repräsentiert. In einem manifesten Regressionsmodell lässt sich dies wie folgt ausdrücken (siehe Aiken & West, 1991):

$$EBK = \beta_0 + \beta_1 MB + \beta_2 VB + \omega_{12} MB \cdot VB + \omega_{11} MB^2 + \omega_{22} VB^2 + \epsilon \quad (1.4)$$

wobei EBK die elterlichen Bildungserwartungen an das Kind, MB das mütterliche Bildungsniveau, VB das väterliche Bildungsniveau und ϵ ein Regressionsresiduum sind. Dabei sind die Regressionskoeffizienten: $\omega_{12} < 0$ sowie $\omega_{11} > 0$ und $\omega_{22} > 0$ (auf dieses Beispiel wird in der Diskussion nochmals ausführlicher eingegangen).

1.3 Nicht-lineare Strukturgleichungsmodelle

Wenn man die oben genannten Zusammenhänge untersuchen möchte, so hat der empirisch arbeitende Forscher stets das Problem mangelnder Messgenauigkeit seiner Beobachtungen zu berücksichtigen (Reliabilitätsproblem). Liegt keine perfekte Reliabilität vor, so sind die Regressionskoeffizienten der nicht-linearen Effekte auf der Ebene manifester Regressionsmodelle verzerrt (Bohrnstedt & Marwell, 1978; MacCallum & Mar, 1995). Anhand von Strukturgleichungsmodellen lassen sich hingegen theoretisch unverzerrte, d.h. erwartungstreue Schätzungen der nicht-linearen Effekte von latenten Variablen produzieren (Marsh, Wen & Hau, 2006; Schumacker & Marcoulides, 1998). Obgleich man sich in der frühen Literatur zur Schätzung von

nicht-linearen latenten Effekten eher auf Modelle mit einzelnen Effekttypen (etwa Interaktionseffekte) beschränkt hat (e.g. Jöreskog & Yang, 1996; Kenny & Judd, 1984), so findet man in jüngerer Zeit eine Zunahme des Interesses für komplexere Modelle zur simultanen Schätzung von Interaktionseffekten und quadratischen Effekten (Kelava, Moosbrugger, Dimitruk & Schermelleh-Engel, 2008; Klein & Muthén, 2007; Lee et al., 2004; Lee, Song & Tang, 2007; MacCallum & Mar, 1995). Gleichung (1.5) drückt Ganzachs Bildungserwartungsmodell in der Notation der latenten Strukturgleichungsmodelle aus (vgl. Gleichung (1.4)):

$$\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_{12}\xi_1\xi_2 + \omega_{11}\xi_1^2 + \omega_{22}\xi_2^2 + \zeta \quad (1.5)$$

In Gleichung (1.5) ist η das latente Kriterium, ξ_1 und ξ_2 sind die latenten Prädiktoren, das Produkt $\xi_1\xi_2$ repräsentiert den Interaktionsterm, ξ_1^2 und ξ_2^2 sind quadratische Terme, α ist ein latentes Interzept, γ_1 und γ_2 sind lineare Effekte der Prädiktoren, ω_{12} ist der nicht-lineare Effekt des Interaktionstermes, ω_{11} und ω_{22} sind die nicht-linearen quadratischen Effekte und ζ ist der Modellfehler. Die allgemeine Matrixschreibweise des nicht-linearen Modells ist nachfolgend gegeben (siehe Gleichung (1.6)):

$$\begin{aligned} \eta &= \alpha + \mathbf{\Gamma}\xi + \xi'\mathbf{\Omega}\xi + \zeta \quad (1.6) \\ &= \alpha + \begin{pmatrix} \gamma_1 & \gamma_2 \end{pmatrix} \cdot \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} + \begin{pmatrix} \xi_1 & \xi_2 \end{pmatrix} \cdot \begin{pmatrix} \omega_{11} & \omega_{12} \\ 0 & \omega_{22} \end{pmatrix} \cdot \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} + \zeta \end{aligned}$$

In Gleichung (1.6) ist $\mathbf{\Gamma}$ der Koeffizienten-Vektor der linearen Effekte der n latenten Prädiktoren (zusammengefasst im ξ -Vektor) und $\mathbf{\Omega}$ ist die Koeffizienten-Matrix der nicht-linearen Effekte (mit den quadratischen Effekten in der Diagonalen und den Interaktionseffekten oberhalb der Diagonale)². Abbildung 1.6 veranschaulicht ein latentes Strukturgleichungsmodell mit einem Interaktionseffekt und zwei quadratischen Effekten (ohne Messmodelle).

1.4 Fragestellungen der Arbeit

Die übergeordnete Fragestellung der vorliegenden Arbeit fragt nach dem Einfluss der *Multikollinearität* bei der Spezifikation und Schätzung von latenten nicht-linearen Effekten in Strukturgleichungsmodellen. Dabei stehen drei Teilfragestellungen im Vordergrund:

²Die in dieser Arbeit verwendete Notation orientiert sich an der LISREL-Notation und für die nicht-linearen Effekte an der Notation von Klein und Moosbrugger (2000). Die genaue Symbolverwendung ist dem Symbolverzeichnis (Anhang D) zu entnehmen.

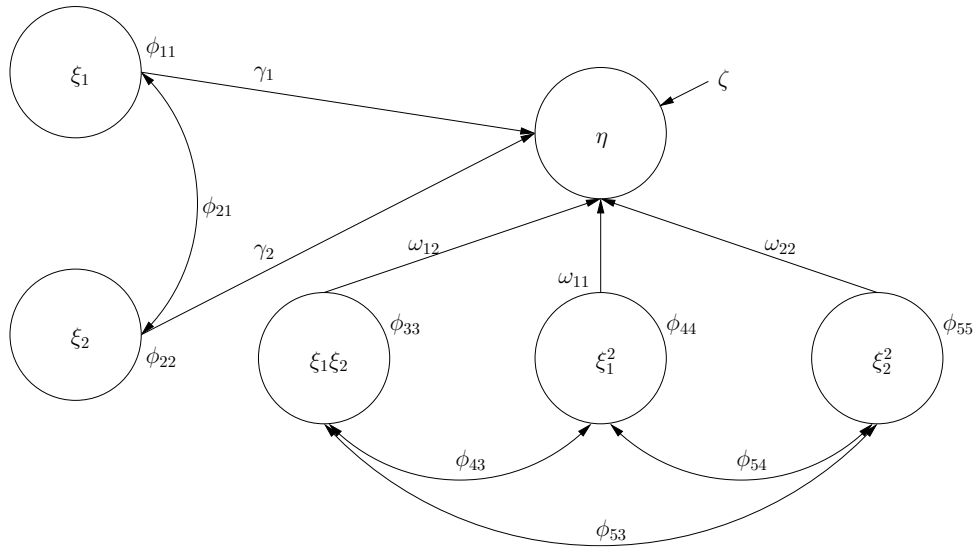


Abbildung 1.6: Pfaddiagramm eines latenten Strukturmodells mit einem Interaktionseffekt und zwei quadratischen Effekten (ohne Messmodelle)

1. Da das Multikollinearitätsproblem in latenten (nicht-linearen) Strukturgleichungsmodellen bisher in der Literatur kaum Beachtung gefunden hat, soll als erstes die Frage beantwortet werden, welche allgemeine Bedeutung Multikollinearität in nicht-linearen Strukturgleichungsmodellen hat.

In Kapitel 3 wird gezeigt, dass Multikollinearität von Prädiktoren nicht nur bei linearen Strukturgleichungsmodellen ein Problem darstellt, sondern vor allem bei nicht-linearen Strukturgleichungsmodellen aufgrund der Präsenz nicht-linearer Terme. Da nicht-lineare Terme untereinander in einem höheren Zusammenhang stehen können als lineare Terme eines Strukturmodells, kann es zu Scheineffekten kommen (im Sinne eines Fehlers erster Art) oder zu verzerrten Schätzungen der nicht-linearen Effekte.

2. In Kapitel 4 wird erstmalig gezeigt, dass die am stärksten vertretene Ansatzklasse zur *simultanen* Schätzung von Interaktionseffekten und quadratischen Effekten, nämlich die Klasse der Produkt-Indikator-Ansätze, seit ihrer Erstpublikation durch Kenny und Judd (1984) *Spezifikationsfehler* aufweist, welche in ähnlicher Form auch in jüngeren Publikationen enthalten sind (z.B. bei Lee et al., 2004). Spezifikationsfehler machen die ganze Ansatzklasse für Multikollinearität „anfällig“, so dass es theoretisch zu verzerrten Schätzungen oder Scheineffekten kommen muss. Es wird gezeigt, wie man die Spezifikationsfehler korrigieren kann, d.h. wie man die bestehenden Ansätze erweitern muss.

Da die latente Prädiktorkorrelation Bestandteil der Spezifikationsfehler ist, soll als zweites die Frage beantwortet werden, wie groß die Verzerrungen sind, die sich bei der Anwendung der (unkorrigierten) Produkt-Indikator-Ansätze ergeben, wenn die Multikollinearität steigt.

Um diese Frage zu beantworten, wird in Kapitel 4 anhand einer Simulationsstudie gezeigt, dass der Zwei-Stufen-Ansatz von Ping (1996) verzerrte Schätzungen der nicht-linearen Effekte produziert. Dieser Ansatz ist ein Vertreter der Produkt-Indikator-Ansätze und enthält die besagten Spezifikationsfehler. In der Simulationsstudie wird außerdem gezeigt, dass die korrekte Erweiterung des Ansatzes von Jöreskog und Yang (1996) zu erwartungstreuen Schätzungen führt.

3. *Als drittes soll der Frage nachgegangen werden, welche theoretischen und empirischen Unterschiede zwischen den neuen verteilungsanalytischen Ansätzen LMS (Klein & Moosbrugger, 2000) und QML (Klein & Muthén, 2007), einerseits, und den klassischen Produkt-Indikator-Ansätzen, andererseits, zu erwarten sind.*

Dazu werden in Kapitel 5 die *theoretischen Eigenschaften* der verteilungsanalytischen Verfahren, LMS und QML, untereinander und mit der Klasse der Produkt-Indikator-Ansätze verglichen. Ihre theoretischen Eigenschaften werden beim Vorliegen von Multikollinearität betrachtet. Illustriert werden die theoretischen Unterschiede anhand einer *Simulationsstudie*, die die Ansätze hinsichtlich ihrer Schätzeigenschaften bei steigender Multikollinearität miteinander vergleicht. Dabei treten die Vorteile der verteilungsanalytischen Verfahren, LMS und QML, hervor.

Für wesentliche theoretische Darstellungen in dieser Arbeit wird auf drei Schriften zurückgegriffen, nämlich

- (A) Dimitruk, P., Schermelleh-Engel, K., **Kelava, A.** & Moosbrugger, H. (2007). Challenges in nonlinear structural equation modeling. *Methodology*, 3, 100-114.
- (B) **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.
- (C) **Kelava, A.**, Werner, C., Schermelleh-Engel, K., Moosbrugger, H., Zapf, D., Ma, Y. & West, S. G. (zur Veröffentlichung eingereicht). Advanced nonlinear

structural equation modeling: Theoretical properties and empirical application of the distribution-analytic LMS and QML estimators. *Structural Equation Modeling*.

1.5 Übersicht über die Arbeit

Um den Leser inhaltlich weiter einzuführen, wird in *Kapitel 2* zunächst ein Überblick über die Ansätze zur Schätzung nicht-linearer Effekte in latenten Strukturgleichungsmodellen gegeben. In *Kapitel 3* wird das Multikollinearitätsproblem und dessen Relevanz betrachtet. Dazu wird auf Schrift (A) und Schrift (B) zurückgegriffen. *Kapitel 4* zeigt auf, dass die Produkt-Indikator-Ansätze *Spezifikationsfehler* aufweisen und korrigiert sie. Veranschaulicht wird dies anhand einer Simulationsstudie. Die Ausführungen sind Schrift (B) entnommen. In *Kapitel 5* werden die theoretischen Eigenschaften der verteilungsanalytischen Verfahren, LMS und QML, den Eigenschaften der Produkt-Indikator-Ansätze gegenübergestellt. Die Schätzeigenschaften bei steigender Multikollinearität werden anhand einer Simulationsstudie untersucht. Kapitel 5 ist eine Zusammenfassung von Schrift (C). *Kapitel 6* ordnet die gewonnenen Erkenntnisse in die Literatur zu nicht-linearen Strukturgleichungsmodellen ein und diskutiert sie kritisch. Abschließend gibt Kapitel 6 einen Ausblick auf zukünftige Fragestellungen.

Kapitel 2

Ansätze zur Schätzung nicht-linearer Effekte in latenten Strukturgleichungsmodellen

In den vergangenen zwei Jahrzehnten haben sich eine ganze Reihe von Ansätzen zur Schätzung nicht-linearer Effekte in latenten Strukturgleichungsmodellen herausgebildet (für einen Überblick Marsh, Wen & Hau, 2004; Schumacker & Marcoulides, 1998). Zu den zwei verbreitetsten Ansätzen gehören die sog. Produkt-Indikator-Ansätze sowie die verteilungsanalytischen Ansätze.

2.1 Produkt-Indikator-Ansätze

2.1.1 Der Kenny und Judd-Ansatz

Die Produkt-Indikator-Ansätze sind auf eine Arbeit von Kenny und Judd (1984) zurückzuführen. Darin beschreiben die Autoren zunächst die Spezifikation zweier latenter Strukturgleichungsmodelle: Das erste Modell ist ein Interaktionsmodell, das einen nicht-linearen Term $\xi_1\xi_2$ beinhaltet, der aus zwei latenten Prädiktoren ξ_1 und ξ_2 gebildet wird. Das zweite Modell ist ein quadratisches Modell, das einen nicht-linearen Term ξ_1^2 enthält¹.

$$Y = \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_{12}\xi_1\xi_2 + \zeta \quad (2.1)$$

$$Y = \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_{11}\xi_1^2 + \zeta \quad (2.2)$$

¹Aus Gründen der Konsistenz wurde die Originalnotation der Notation dieser Schrift angepasst.

Der wesentliche Beitrag dieser Publikation besteht darin, dass erstmalig die *Spezifikation des Messmodells der nicht-linearen Terme* (z.B. $\xi_1\xi_2$ oder ξ_1^2) vorgenommen wird. Als Indikatoren der nicht-linearen Terme werden die Produkte der Indikatoren des linearen Messmodells verwendet.

Im Folgenden wird die Spezifikation des Messmodells bei einem latenten Interaktionsmodell veranschaulicht: Seien X_1 und X_2 Indikatoren des latenten Prädiktors ξ_1 sowie X_3 und X_4 Indikatoren des Prädiktors ξ_2 , mit

$$X_1 = \lambda_{11}^X \xi_1 + \delta_1 \quad X_2 = \lambda_{21}^X \xi_1 + \delta_2 \quad (2.3)$$

$$X_3 = \lambda_{32}^X \xi_2 + \delta_3 \quad X_4 = \lambda_{42}^X \xi_2 + \delta_4 \quad (2.4)$$

dann sind die Produkte X_1X_3 , X_1X_4 , X_2X_3 und X_2X_4 Indikatoren des Interaktionstermes $\xi_1\xi_2$ (s. Abbildung 2.1).

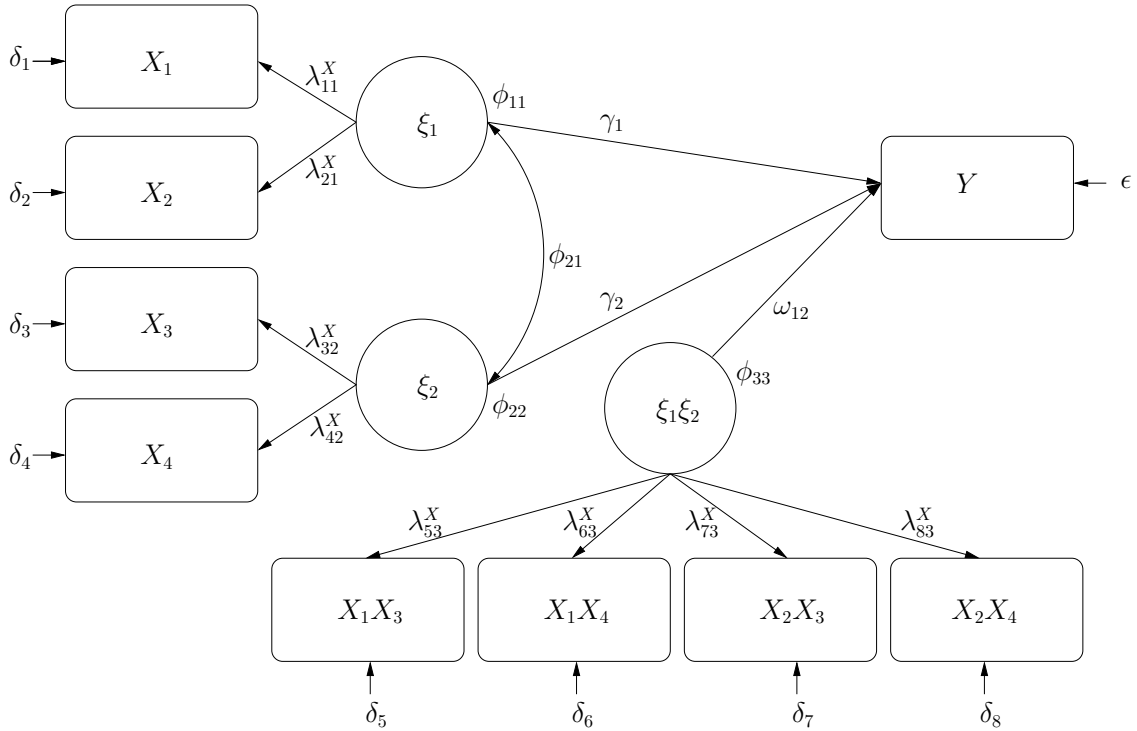


Abbildung 2.1: Interaktionsmodell nach Kenny und Judd (1984) mit den linearen und nicht-linearen Messmodellen

Der Produkt-Indikator X_1X_3 ergibt sich dann als:

$$X_1X_3 = \underbrace{\lambda_{11}^X \lambda_{32}^X}_{\lambda_{53}^X} \xi_1\xi_2 + \underbrace{\lambda_{32}^X \xi_2 \delta_1 + \lambda_{11}^X \xi_1 \delta_3 + \delta_1 \delta_3}_{\delta_5} \quad (2.5)$$

$$= \lambda_{53}^X \xi_1\xi_2 + \delta_5 \quad (2.6)$$

Die Varianzzerlegung des Produkt-Indikators X_1X_3 , die man zur Schätzung des Modells benötigt, ist dabei nicht trivial, da die Varianz des Produkt-Indikators eine nicht-lineare Funktion der Parameter des linearen Messmodells darstellt. Die Varianzzerlegung von X_1X_3 lautet wie folgt:

$$Var(X_1X_3) = (\lambda_{53}^X)^2 Var(\xi_1\xi_2) + Var(\delta_5) \quad (2.7)$$

wobei

$$\lambda_{53}^X = \lambda_{11}^X \lambda_{32}^X \quad (2.8)$$

$$Var(\xi_1\xi_2) = \phi_{33} = \phi_{11}\phi_{22} + \phi_{21}^2 \quad (2.9)$$

$$Var(\delta_5) = \theta_{55}^\delta = (\lambda_{11}^X)^2 \phi_{11}\theta_{33}^\delta + (\lambda_{32}^X)^2 \phi_{22}\theta_{11}^\delta + \theta_{11}^\delta \theta_{33}^\delta \quad (2.10)$$

Wie man erkennen kann, hat diese nicht-lineare Abhängigkeit zur Folge, dass zur Parameterschätzung nicht-lineare Restriktionen formuliert werden müssen. Da ihre Formulierung allerdings sehr fehleranfällig ist, ist dieser Ansatz bisher nur selten angewendet worden. Tabelle 2.1 gibt einen Überblick über die Parameterrestriktionen des Interaktionsmodells und des quadratischen Strukturgleichungsmodells, wenn man die Menge der Indikatoren der nicht-linearen Terme nicht vollständig miteinander einkreuzt.

Ein weiterer Grund für die seltene Anwendung dieses Ansatzes liegt außerdem darin, dass latente Produktterme (etwa $\xi_1\xi_2$ oder ξ_1^2) nicht normalverteilt sind. Dies ist selbst dann der Fall, wenn die manifesten und latenten Prädiktorvariablen normalverteilt sind (Aroian, 1944). In der Folge werden, wenn man die übliche Maximum Likelihood Schätzung der Parameter durchführt, die Standardfehlerschätzungen der nicht-linearen Effekte (z.B. ω_{12} im Falle eines Interaktionseffektes) unterschätzt, sodass darauf basierende Signifikanztests und Konfidenzintervalle nicht korrekt sind (Jöreskog & Yang, 1996). Bereits Kenny und Judd (1984) weisen darauf hin, dass der Verletzung der multivariaten Normalverteilungsannahme am besten dadurch zu begegnen ist, dass man Kleinstquadrateschätzungen vornimmt. Sie schlagen dazu GLS (generalized least squares) vor.

2.1.2 Weiterentwicklungen

In der Folgezeit hat der von Kenny und Judd (1984) vorgeschlagene Ansatz eine Vielzahl von Weiterentwicklungen erfahren. An dieser Stelle wird ein kurzer und nicht erschöpfender Überblick über diese Entwicklungen geben, da es sich bei den Produkt-Indikator-Ansätzen um die größte Ansatzklasse handelt.

Tabelle 2.1: Überblick über die nicht-linearen Parameterrestriktionen beim Produkt-Indikator-Ansatz.

Nicht-lineares Strukturmodell	Interaktionsmodell	Quadratisches Modell
Produkt-Indikatoren	$X_1 X_3, X_2 X_4$	X_1^2, X_2^2
Nicht-lineare Restriktionen	Faktorladungen $\lambda_{53}^X = \lambda_{11}^X \lambda_{32}^X, \lambda_{63}^X = \lambda_{21}^X \lambda_{42}^X$	Faktorladungen $\lambda_{53}^X = (\lambda_{11}^X)^2, \lambda_{63}^X = (\lambda_{21}^X)^2$
	Erwartungswert des latenten nicht-linearen Terms $E(\xi_1 \xi_2) = \phi_{21}$	Erwartungswert des latenten nicht-linearen Terms $E(\xi_1^2) = \phi_{11}$
	Varianz des latenten nicht-linearen Terms $Var(\xi_1 \xi_2) = \phi_{11} \phi_{22} + \phi_{21}^2$	Varianz des latenten nicht-linearen Terms $Var(\xi_1^2) = 2\phi_{11}^2$
	Kovarianz des latenten nicht-linearen Terms $Cov(\xi_1 \xi_2, \xi_1) = Cov(\xi_1 \xi_2, \xi_2) = 0$	Kovarianz des latenten nicht-linearen Terms $Cov(\xi_1^2, \xi_1) = Cov(\xi_1^2, \xi_2) = 0$
	Messfehlervarianzen $\theta_{55}^\delta = (\lambda_{11}^X)^2 \phi_{11} \theta_{33}^\delta + (\lambda_{32}^X)^2 \phi_{22} \theta_{11}^\delta + \theta_{11}^\delta \theta_{33}^\delta$ $\theta_{66}^\delta = (\lambda_{21}^X)^2 \phi_{11} \theta_{44}^\delta + (\lambda_{42}^X)^2 \phi_{22} \theta_{22}^\delta + \theta_{22}^\delta \theta_{44}^\delta$	Messfehlervarianzen $\theta_{55}^\delta = 4 (\lambda_{11}^X)^2 \phi_{11} \theta_{11}^\delta + 2 (\theta_{11}^\delta)^2$ $\theta_{66}^\delta = 4 (\lambda_{21}^X)^2 \phi_{11} \theta_{22}^\delta + 2 (\theta_{22}^\delta)^2$
	Messfehlerkovarianzen $\theta_{65}^\delta = 0$	Messfehlerkovarianzen $\theta_{65}^\delta = 0$

Anmerkungen. Die Publikation von Kenny und Judd (1984) enthält nicht das latente Interzept α . Es wurde von Kenny und Judd (1984) irrtümlicherweise nicht spezifiziert und daher ergänzt (Jöreskog & Yang, 1996). Zur Veranschaulichung wurden die Indikatoren nicht vollständig mit einander eingekreuzt. Es wurde z.B. auf die zusätzliche Aufnahme des Indikators $X_2 X_3$ in das Interaktionsmodell verzichtet, da die Zahl der Restriktionen unverhältnismäßig anwachsend würde. Dies gilt insbesondere für die Messfehlerkovarianzen.

Während Kenny und Judd (1984) die nicht-linearen Abhängigkeiten der Parameter untereinander nur unter großem Aufwand anhand der COSAN-Prozedur (Fraser, 1980) modellieren konnten, gelang Hayduk (1987) eine Implementierung des Ansatzes in die LISREL Version 7-Software (Jöreskog & Sörbom, 1989). Diese Implementierung macht allerdings die Modellierung einer ganzen Reihe von zusätzlichen latenten Variablen (sog. Phantomvariablen) notwendig. Jaccard und Wan (1995) erweiterten das Messmodell des latenten Kriteriums auf mehr als eine Indikatorvariable.

Mit der Einführung von LISREL Version 8 (Jöreskog & Sörbom, 1996) ließen sich nicht-lineare Restriktionen erstmalig direkt spezifizieren. Diesen Fortschritt nutzend konnten Jöreskog und Yang (1996) den Kenny und Judd-Ansatz für Strukturgleichungsmodelle mit einem Interaktionseffekt dahingehend verallgemeinern, dass die linearen Indikatoren nun nicht mehr als zentriert angenommen werden mussten. Dazu nahmen sie latente Interzepte (z.B. τ_1^X für X_1) in die Messmodellgleichungen der linearen Indikatoren auf:

$$X_1 = \tau_1^X + \lambda_{11}^X \xi_1 + \delta_1 \qquad X_2 = \tau_2^X + \lambda_{21}^X \xi_1 + \delta_2 \qquad (2.11)$$

$$X_3 = \tau_3^X + \lambda_{32}^X \xi_2 + \delta_3 \qquad X_4 = \tau_4^X + \lambda_{42}^X \xi_2 + \delta_4 \qquad (2.12)$$

Zudem führten Sie in das Strukturmodell das latente Interzept α ein (Gleichung (2.13)), das von Kenny und Judd (1984) zuvor „vergessen“ wurde (vgl. Gleichung (2.1)). Die Aufnahme des Interzepts α ist notwendig, da die Erwartungswerte der nicht-linearen Terme (z.B. $\xi_1 \xi_2$) nicht Null sein müssen. Selbst dann nicht, wenn alle latenten Prädiktoren zentriert und normalverteilt sind.

$$\eta = \alpha + \gamma_1 \xi_1 + \gamma_2 \xi_2 + \omega_{12} \xi_1 \xi_2 + \zeta \qquad (2.13)$$

Die Erweiterung des Ansatzes von Kenny und Judd (1984) zur Schätzung eines Interaktionseffektes bringt es mit sich, dass die Zahl der Restriktionen stark ansteigt. Einen Überblick über die Restriktionen geben Jöreskog und Yang (1996). Wie auch Kenny und Judd (1984) anmerken, empfehlen Jöreskog und Yang (1996) statt einer ML-Schätzung der Parameter eine Kleinstquadrateschätzung vorzunehmen. Hierzu verweisen sie auf die Alternativen WLS (weighted least squares) und WLSA (WLS based on augmented moment matrix). Die Eigenschaften dieser Schätzer werden weiter unten besprochen.

Algina und Moulder (2001) veränderten den Ansatz von Jöreskog und Yang (1996) dahingehend, dass die Indikatoren der latenten Prädiktoren (wie im Jaccard-Wan-Modell) zentriert werden, aber die Erwartungswerte innerhalb des latenten Struktur-

modells spezifiziert werden. Auch wenn ihr Ansatz im Wesentlichen gegenüber dem Ansatz von Jöreskog und Yang (1996) unverändert ist, fanden sie, dass er leichter konvergiert.

Eine weitere Erweiterung nahmen Wall und Amemiya (2001) vor. Vor dem Hintergrund, dass mit der Anwendung eines Produkt-Indikator-Ansatzes stets die Annahme der Normalverteilung der latenten Prädiktoren einhergeht, untersuchten sie die Verletzung dieser Annahme. In ihrer Publikation zeigen sie, dass im Falle nicht-normalverteilter Prädiktoren die Korrelation der linearen Prädiktoren und des Interaktionstermes (z.B. $Cov(\xi_1, \xi_1\xi_2)$) nicht Null ist. Zudem entspricht die Varianz des Interaktionstermes dann nicht mehr $\phi_{11}\phi_{22} + \phi_{21}^2$. Eine Nichtberücksichtigung dieses Umstandes führt – wie sie zeigen – zu verzerrten Parameterschätzungen des Interaktionseffektes. Um dieses Problem zu beheben, schlagen Wall und Amemiya (2001) den „generalized appended product indicator (partially constrained)“ Ansatz (GAPI) vor. Der GAPI-Ansatz beinhaltet die Aufhebung der Restriktionen innerhalb der Kovarianzmatrix der latenten Prädiktoren und eine freie Schätzung sämtlicher Elemente in ihr. Alle anderen Restriktionen des Messmodells bleiben von der Aufhebung unberührt. Durch diese Abänderung wird die Einsetzbarkeit des Ansatzes auf nicht-normalverteilte Konstrukte erweitert.

Marsh et al. (2004) gingen bei Ihrem Ansatz einen Schritt weiter und hoben alle Restriktionen auf. Das bedeutet, dass keine Restriktionen für die Parameter der Kovarianzmatrix der latenten Prädiktoren und des Interaktionstermes bestehen. Außerdem heben sie sämtliche Restriktionen der Messfehlervarianzen auf (vgl. im Gegensatz dazu Tabelle 2.1). Alle zuvor restringierten Parameter werden in diesem Ansatz frei geschätzt. Marsh et al. (2004) zeigen, dass ein einfaches Interaktionsmodell mit zwei Produkt-Indikatoren identifiziert ist. Um keine Messfehler-Kovarianz der beiden Produkt-Indikatoren schätzen zu müssen, bilden sie in ihrem Ansatz Produkt-Indikatoren, die aus sich nicht wiederholenden Indikatoren bestehen [Anmerkung: Wenn z.B. wie bei Kenny und Judd (1984) X_1 und X_2 die Indikatoren von ξ_1 sind sowie X_3 und X_4 die Indikatoren von ξ_2 , dann sind X_1X_3 und X_2X_4 nicht-redundante Produkt-Indikatoren. Während X_1X_4 in dieser Produkt-Indikatoren-Menge redundante Informationen enthalten würde.]

Little, Bovaird und Widaman (2006) stellten jüngst einen Ansatz vor, bei dem jeder nicht-lineare Produkt-Indikator (z.B. X_1X_4) auf alle linearen Indikatoren (hier X_1, X_2, \dots) regressiert wird. Danach werden die Residuen dieser Regression als Indikator des latenten Produkttermes ($\xi_1\xi_2$) verwendet. Dieser Ansatz besteht also aus zwei Schritten und gestaltet sich daher etwas aufwendiger. Er besteht aus ei-

nem Orthogonalisierungsschritt, bei dem die Residuen bestimmt werden und einem zweiten Schritt, der die Schätzung des Gesamtmodells auf Grundlage der Residuen beinhaltet.

Den Gedanken der zweistufigen Vorgehensweise findet man allerdings auch bei Ping (1996). Er schlägt ein Verfahren vor, das im Gegensatz zu den anderen Ansätzen der Produkt-Indikator-Ansätze für Anwender relativ einfach zu handhaben ist und als einziger Ansatz auch für die simultane Schätzung von Interaktionseffekten und quadratischen Effekten entwickelt wurde. Dabei wird im ersten Schritt zunächst ein lineares Messmodell der latenten Prädiktoren berechnet (in Form einer konfirmatorischen Faktorenanalyse). Danach werden im zweiten Schritt die geschätzten Parameter verwendet, um die Parameter des eigentlichen Strukturmodells zu schätzen. Dazu werden die Ladungsparameter und Messfehlervarianzen der nicht-linearen Indikatoren (z.B. X_1X_4) des nicht-linearen Messmodells auf Werte fixiert, die sich aus Gleichungen gemäß (2.7) - (2.10) ergeben. Man muss dann nicht mehr die Restriktionen im Gesamtmodell formulieren, sondern setzt die Messfehlervarianzen und Ladungen als bekannt in das Messmodell der nicht-linearen Terme ein.²

2.2 Verteilungsanalytische Ansätze

2.2.1 Der Latent Moderated Structural Equations Ansatz (LMS)

Allen Produkt-Indikator-Ansätzen ist gemein, dass sie aufgrund der Restriktionen für Anwender relativ kompliziert zu spezifizieren sind und ihre Standardfehler unterschätzt werden (Jöreskog & Yang, 1996). Um die Nachteile der Produkt-Indikator-Ansätze zu überwinden, entwickelten Klein und Moosbrugger (2000) einen verteilungsanalytischen Ansatz zur Schätzung nicht-linearer Interaktionseffekte und quadratischer Effekte. Der sogenannte Latent Moderated Structural Equations Ansatz (LMS) basiert auf der Analyse der Verteilung des multivariaten Indikatorvektors (\mathbf{X}, \mathbf{Y}) . Dieser Vektor ergibt sich, wenn man die q manifesten X -Indikatorvariablen der latenten Prädiktoren und die p manifesten Y -Indikatoren des latenten Kriteriums zusammenfasst: $(\mathbf{X}, \mathbf{Y}) = (X_1, \dots, X_q, Y_1, \dots, Y_p)$. Die multivariate Verteilung des

²Bereits jetzt sei angemerkt, dass dieser Ansatz wie auch alle anderen Ansätze der Produkt-Indikator-Ansatzklasse, die für die Schätzung nicht-linearer Terme höherer Ordnung konstruiert worden sind, Spezifikationsfehler aufweisen. Diese werden in Kapitel 4 beschrieben und Lösungsmöglichkeiten beschrieben.

Indikatorvektors (\mathbf{X}, \mathbf{Y}) ist nicht-normal. LMS approximiert diese durch eine finite Mischverteilung von bedingten Normalverteilungen. LMS beruht dabei auf dem Kerngedanken, dass man die nicht-linearen Effekte (z.B. den Effekt von $\xi_1\xi_2$) des Interaktionsmodells oder des quadratischen Modells durch lineare Effekte darstellen kann, wenn man die Verteilung von (\mathbf{X}, \mathbf{Y}) auf den einzelnen Stufen einer Komponente (z.B. auf den Stufen von ξ_1) betrachtet. Um die unbekanntes Modellparameter zu schätzen, wird in LMS der Expectation Maximization (EM) Algorithmus (Dempster, Laird & Rubin, 1977) angewendet, der Maximum Likelihood Schätzungen liefert. In späteren Abschnitten wird vertiefend auf diesen Ansatz eingegangen.

2.2.2 Der Quasi-Maximum Likelihood Ansatz (QML)

Da es sich bei LMS konstruktionsbedingt um ein sehr rechenintensives Verfahren handelt, haben Klein und Muthén (2007) den Quasi-Maximum Likelihood Ansatz (QML) entwickelt, um multiple Interaktionseffekte und quadratische Effekte schätzen zu können. Bei der Entwicklung von QML stand das Ziel im Vordergrund, ein theoretisch robusteres Verfahren zu schaffen, wenn die Prädiktorvariablen nicht-normalverteilt sind. Wie auch im Falle von LMS, trägt QML dem Gesichtspunkt Rechnung, dass das latente Kriterium und seine Indikatoren nicht-normalverteilt sind, indem es die Zahl der nicht-normalverteilten Komponenten der Indikatorvektorverteilung $f(\mathbf{X}, \mathbf{Y})$ reduziert. Dabei wird das Quasi-Likelihood Prinzip angewendet, welches die Verteilung $f(\mathbf{X}, \mathbf{Y})$ des Indikatorvektors (\mathbf{X}, \mathbf{Y}) durch ein Produkt einer normalverteilten und einer bedingt normalverteilten Dichtefunktion approximiert³. Durch diese Transformation der multivariaten Verteilung erreicht man, dass die Parameter des latenten nicht-linearen Modells sehr viel einfacher zu schätzen sind (z.B. mit dem Newton-Raphson Algorithmus).

2.3 Sonstige Ansätze

Wie bereits zu Beginn dieses Kapitels erwähnt, ist es nicht möglich, eine erschöpfende Zusammenfassung sämtlicher Ansätze zu bieten. Dennoch seien einige weitere Ansätze genannt, die sich nicht einfach in die beiden obigen Klassen einordnen lassen.

Von Bollen (1995) wurde ein zwei-stufiger Kleinstquadratesatz (two stages least

³Dem Leser sei in Erinnerung gerufen, dass im Unterschied zu QML bei LMS die multivariate Indikatorverteilung $f(\mathbf{x}, \mathbf{y})$ durch eine Mischverteilung approximiert wird.

squares, 2SLS) vorgeschlagen. Er basiert im Wesentlichen darauf, dass die latenten Prädiktoren durch ihre Indikatoren in der latenten Strukturmodellgleichung (vgl. Gleichung (2.13)) ersetzt werden. Die Indikatoren werden ihrerseits durch sog. instrumentelle Variablen geschätzt. Instrumentelle Variablen sind Indikatorvariablen, die mit den zuvor eingesetzten Indikatorvariablen zwar in Zusammenhang stehen, jedoch nicht mit ihren Fehlertermen korrelieren. Die resultierende Strukturgleichung erfüllt die Voraussetzungen einer OLS-Schätzung, sodass die Koeffizienten berechnet und beurteilt werden können. Allerdings hat sich der 2SLS-Ansatz bei kleinen Stichproben als weniger effizient im Vergleich zu anderen Verfahren erwiesen (z.B. Jöreskog, 1998; Klein, Moosbrugger & Schermelleh-Engel, unveröffentlicht). Moulder und Algina (2002) weisen nach, dass bei niedriger Reliabilität der Indikatoren die Parameterschätzungen nicht erwartungstreu geschätzt werden und die Teststärke vergleichsweise gering ist.

Wall und Amemiya (2000) entwickelten ebenfalls einen zweistufigen Ansatz, der als „two-step method of moments (2SMM)“ für ein polynomiales Strukturmodell höherer Ordnung entwickelt wurde. Im ersten Schritt des 2SMM-Ansatzes werden die Parameter des Messmodells anhand einer faktorenanalytischen Herangehensweise geschätzt. In einem zweiten Schritt werden bedingte Momente der Produkte der latenten Variablen berechnet. Diese werden dann zur Schätzung der unbekanntes Strukturmodellparameter verwendet. Bei diesem Ansatz kann insbesondere die Annahme der normalverteilten latenten Variablen gelockert werden, wenn verteilungsfreie Factor-Score-Schätzer im ersten Schritt berechnet werden. Theoretisch kann der 2SMM-Ansatz auf sehr viel komplexere nicht-lineare Modelle erweitert werden, allerdings wurde er bisher aufgrund seiner Komplexität kaum untersucht (Klein & Muthén, 2007).

Bayessche Ansätze wurden insbesondere von Arminger und Muthén (1998) und Lee et al. (2007) veröffentlicht. Sie erlauben auf Grundlage des Bayesschen Theorems flexible Modellierungen relativ komplexer Likelihoodfunktionen von nicht-linearen Zusammenhängen. Allerdings wird die Anwendung solcher Ansätze dadurch sehr erschwert, dass man fundiertes a-priori Wissen über die Verteilungen der unbekanntes Parameter braucht, um die Vorteile dieser Ansätze ausnutzen zu können. Dies gestaltet sich bei Interaktionsmodellen auf latenter Ebene schwierig, da nicht-lineare Modelle bis vor etwa zehn Jahren von Anwendern nicht schätzbar waren.⁴

⁴Mit dem Aufkommen sehr leistungsstarker Software in jüngster Zeit ist auch prinzipiell die Schätzung nicht-linearer Strukturgleichungsmodelle möglich geworden, so dass hier in Zukunft a priori Wissen zu erwarten ist.

2.4 Schätzeigenschaften der Verfahren

Die Ergebnisse größerer Simulationsstudien zur Schätzung von Interaktionseffekten (Marsh et al., 2004; Moulder & Algina, 2002; Schermelleh-Engel, Klein & Moosbrugger, 1998; Yang Jonsson, 1997) zu den Produkt-Indikator-Ansätzen lassen sich wie folgt zusammenfassen: Die zwei Hauptprobleme der Produkt-Indikator-Ansätze liegen einerseits darin, dass Konvergenz- und Schätzprobleme auftreten, andererseits aber auch, dass bei erfolgreicher Parameterschätzung die Standardfehler der Parameterschätzungen erheblich unterschätzt werden, weil die Normalverteilungsannahme der Indikatoren verletzt ist. Dies ist auf die Verwendung von Produkten von Variablen zurückzuführen (Aroian, 1944). Die Unterschätzung ist dabei umso gravierender, je größer der nicht-lineare Effekt ist. Konfidenzintervalle der Parameterschätzungen, die auf der Basis dieser unterschätzten Standardfehler bestimmt werden, sind folglich zu klein, was zu falschen Schlussfolgerungen führen kann (in Form von Scheineffekten). Die Verwendung von korrigierten Standardfehlern (Satorra & Bentler, 2001) führt dabei zu einer Überkorrektur (Moulder & Algina, 2002), so dass sich eher die Verwendung unkorrigierter Standardfehler empfiehlt.

Um dieses Problem der Verzerrung der Standardfehler bei ML-Schätzungen zu umgehen, kann eine Robust Maximum Likelihood-Schätzung (RML) durchgeführt werden, die eine asymptotische Kovarianzmatrix verwendet. Dabei erhält man bei nicht-normalverteilten Variablen korrekte Standardfehler und χ^2 -Werte (vgl. Jöreskog, Sörbom, Du Toit & Du Toit, 1999). Allerdings scheinen für unverzerrte RML-Parameterschätzungen relativ große Stichproben benötigt zu werden (Yang Wallentin, Schmidt, Davidov & Bamberg, 2003). Systematische Untersuchungen zu dieser Methode wurden bisher noch nicht durchgeführt.

Alternativ gibt es die Möglichkeit, das auf Produkt-Indikatoren basierende Strukturgleichungsmodell anhand verteilungsfreier Verfahren, wie z.B. Weighted Least Squares (WLS) oder Weighted Least Squares based on the Augmented Moment Matrix (WLSA) (Jöreskog & Yang, 1996) zu schätzen. WLS und WLSA Schätzungen bedürfen allerdings sehr großer Stichproben, um zu erwartungstreuen Parameterschätzungen zu gelangen (vgl. Schermelleh-Engel et al., 1998; Yang Jonsson, 1997). Ihre Parameterschätzer sind lediglich asymptotisch unverzerrt. Im Vergleich zu ML weisen sie eine geringere Effizienz auf, so dass prinzipiell der Effizienznachteil durch größere Stichproben ausgeglichen werden muss.

Klein und Moosbrugger (2000) zeigen, dass LMS verglichen mit 2SLS, ML und WLSA effizientere Parameterschätzungen von Interaktionseffekten und zuverlässi-

gere Modelldifferenztests (Likelihood-Ratio-Tests) berechnet. Zudem sind die Standardfehler von LMS relativ unverzerrt. Klein und Muthén (2007) zeigen in einer Simulationsstudie, dass QML- und LMS-Schätzungen sehr eng bei einander liegen und dass QML Schätzungen substantiell effizienter sind als Schätzungen auf Grundlage des 2SMM Ansatzes von Wall und Amemiya (2000). Bei nicht-normal verteilten Prädiktoren sind QML-Schätzungen nach einer Simulationsstudie von Marsh et al. (2004) leicht verzerrt, was von Klein und Muthén (2007) in der Form nicht bestätigt werden konnte. Klein und Muthén (2007) zeigen, dass QML höchst effizient bei nicht-normalen Prädiktoren und sein Modelldifferenztest der nicht-linearen Effekte robust ist. QML erweist sich im Vergleich zu anderen Produkt-Indikator-Ansätzen (GAPI, unconstrained Ansatz) als leistungsfähiger.

Simulationsstudien zur simultanen Schätzung von Interaktionseffekten und quadratischen Effekten mit QML und LMS werden auch in Schrift (C) durchgeführt, die dieser Arbeit zugrundeliegt. Die Ergebnisse werden in Kapitel 5 dargestellt.

Kapitel 3

Zum Problem der Multikollinearität und seiner Relevanz für nicht-lineare Strukturgleichungsmodelle

Die übergeordnete Fragestellung der vorliegenden Arbeit fragt nach dem Einfluss der Multikollinearität bei der Spezifikation und Schätzung von latenten nicht-linearen Effekten in Strukturgleichungsmodellen. *Da das Multikollinearitätsproblem in latenten (nicht-linearen) Strukturgleichungsmodellen bisher in der Literatur kaum Beachtung gefunden hat, soll als erstes die Frage beantwortet werden, welche allgemeine Bedeutung Multikollinearität in nicht-linearen Strukturgleichungsmodellen hat.* Zunächst soll auf den Begriff der Multikollinearität kurz eingegangen werden. Wesentliche Inhalte dieses Kapitels finden sich in Schrift (A) und Schrift (B).

3.1 Zum Begriff der Multikollinearität

Multikollinearität ist dann ein Problem, wenn Prädiktorvariablen in einem Modell stark korrelieren (z.B. zu .80). Eine hohe Prädiktorenkorrelation bedeutet, dass die Prädiktoren redundante Informationen enthalten. Im Extremfall liegt *Singularität* vor. Dann ist eine der Prädiktorvariablen eine (Linear-)Kombination von zwei oder mehreren anderen Variablen. Multikollinearität erweist sich bei der Schätzung von Modellparametern als störend, da die Schätzungen bei steigender Multikollinearität unpräziser werden. Singularität macht eine Schätzung der betreffenden Modellparameter sogar unmöglich.

Was die *Ursache für das Vorliegen von Multikollinearität* ist, kann auf verschie-

denen Ebenen diskutiert werden. Eine Ebene beinhaltet jedenfalls die empirisch-inhaltliche Auseinandersetzung mit der Konstruktvalidität der untersuchten Konstrukte. Hier sei lediglich angemerkt, dass beispielsweise die Verwendung von gleichartigen Messmethoden (etwa Selbsteinschätzungen anhand von Fragebögen) zu einer erhöhten Korrelation der Maße führen kann. Um eine Beurteilung der konvergenten und diskriminanten Validität der Konstrukte vorzunehmen, wurden in den letzten Jahrzehnten seit Campbell und Fiske (1959) sehr differenzierte Möglichkeiten geschaffen, die unter dem Stichpunkt der „Multitrait-Multimethod“-Ansätze ihren Eingang in die Literatur gefunden haben (für eine Übersicht s. Eid & Diener, 2006; für eine Einführung: Schermelleh-Engel & Schweizer, 2007). Auf diese Problemebene soll im weiteren Verlauf nicht eingegangen werden. Vielmehr soll jene Ebene im weiteren Verlauf betrachtet werden, die beinhaltet, dass Konstrukte oder andere Variablen zusammenhängen können und sich dieser Zustand in einer vorliegenden Forschungssituation nicht ohne weiteres beseitigen lässt (bis auf statistische Artefakte, die weiter unten dargestellt werden). Dies ist zum Beispiel dann denkbar, wenn man Angst und Depression als Konstrukte untersuchen möchte, die selbst dann einen „Restzusammenhang“ aufweisen, wenn sehr aufwändige Formen der Separierung von Varianzquellen vorgenommen worden sind (Nussbeck, Eid, Geiser, Courvoisier & Cole, 2007).

3.2 Folgen, Diagnostik und Abhilfen

3.2.1 Allgemeine Folgen der Multikollinearität

Prinzipiell hat das Vorliegen von Multikollinearität zur Folge, dass sich Regressionskoeffizienten in ihrem Betrag und in ihrem Vorzeichen – also der partiellen Zusammenhangsrichtung – ändern können (vgl. Cohen, Cohen, West & Aiken, 2003, für anschauliche Beispiele). Zudem wird der Standardfehler einer Parameterschätzung und das eine Schätzung umgebende Konfidenzintervall größer. Damit nimmt die Genauigkeit der Parameterschätzung ab, da sich der wahre Wert des Parameters in einem größeren Intervall befinden kann. Letztlich führt auch die Vergrößerung des Konfidenzintervalls dazu, dass Parameterschätzungen nicht (mehr) bedeutsam von Null verschieden sein können, wenn zusätzliche Prädiktoren aufgenommen werden, die mit den zuvor im Modell enthaltenen Prädiktoren korrelieren. Dies impliziert also einen Teststärkeverlust.

3.2.2 Multikollinearitätsdiagnostik

Damit es nicht zu solchen negativen Effekten von Multikollinearität kommt, sollte vor den eigentlichen Analysen eine intensive Multikollinearitätsdiagnostik betrieben werden. Diese umfasst z.B. für jeden Prädiktor X_j die Bestimmung seiner Varianz R_j^2 , die durch die anderen Prädiktoren erklärt wird, die Bestimmung der *Toleranz* $1 - R_j^2$ oder die Bestimmung des *Variance Inflation Factors* (VIF) $\frac{1}{1 - R_j^2}$ (Cohen et al., 2003). Ein $VIF \geq 10$ wird im allgemeinen als ernsthaftes Multikollinearitätsproblem betrachtet (Marquardt, 1970).

Desweiteren kann man die Eigenwerte der Prädiktorvariablenmatrix bestimmen und daraus die *condition number* κ ableiten (Belsley, Kuh & Welsh, 1980); wobei $\kappa = \sqrt{\frac{\lambda_{max}}{\lambda_{min}}}$. Dabei wird das Verhältnis des höchsten Eigenwertes λ_{max} zum niedrigsten Eigenwert λ_{min} bestimmt. Eine *condition number* von 1 bedeutet, dass die Prädiktoren gänzlich unabhängig voneinander sind. Je höher der Wert, desto größer ist die Multikollinearität. Als Faustregel kann man einen Wert zwischen 5 und 10 als akzeptabel betrachten. Eine *condition number* von größer 30 zeigt sehr starke Multikollinearität an.

3.2.3 Abhilfen bei Vorliegen von Multikollinearität

Gelegentlich kann Multikollinearität ein Artefakt der Stichprobenziehung sein. Das bedeutet im einfachsten Falle, dass die Stichprobe und die an ihr erhobenen Merkmale nicht die Merkmalsrelationen in der Population adäquat abbilden. Eine solche Ursache kann man – abgesehen von praktischen Hindernissen – am einfachsten dadurch beheben, indem man weitere Fälle erhebt.

Wenn die Stichprobe hingegen repräsentativ für die Population ist, bestehen weitere Möglichkeiten der Multikollinearitätsbeseitigung darin, dass man Prädiktorvariablen aus dem Modell entfernt (also ein anderes Modell spezifiziert) oder aber Variablen, die redundante Informationen beinhalten, zusammenfasst. Das kann durch die Bildung von Indices erreicht werden, indem man z.B. zwei Variablen zunächst z-transformiert und dann miteinander addiert oder aber Faktorscores berechnet, die eine Repräsentation der kollinearen Variablen darstellen.

In Situationen, in denen eine Respezifikation eines Modells nicht möglich ist, kann man auch die Erwartungstreue der Parameterschätzungen aufgeben und unter Inkaufnahme einer kleinen Verzerrung der Parameterschätzungen eine sog. Ridge Regression durchführen (Darlington, 1978). Diese beinhaltet, dass eine Konstante zu jeder Prädiktorenvarianz hinzuaddiert wird, so dass die Kollinearität (in der

Kovarianzmatrix) abgeschwächt wird. Der Vorteil der Ridge Regression liegt darin, dass die Standardfehler kleiner werden und somit die auf Konfidenzintervallen basierenden Inferenzen belastbarer sind.

3.3 Zur Bedeutung der Multikollinearität in nicht-linearen Strukturgleichungsmodellen

Betrachtet man die oben beschriebenen Gesichtspunkte auf der Ebene manifester Regressionsmodelle, so ist die Bedeutsamkeit des Multikollinearitätsproblems offensichtlich. *Weniger offensichtlich, aber umso schwerwiegender ist das Multikollinearitätsproblem auf der Ebene latenter Modelle.* Berücksichtigt man nämlich, dass Indikatoren in der Regel messfehlerbehaftet gemessen werden, sind selbst moderate Korrelationen der Indikatoren verschiedener Prädiktoren als problematisch anzusehen.

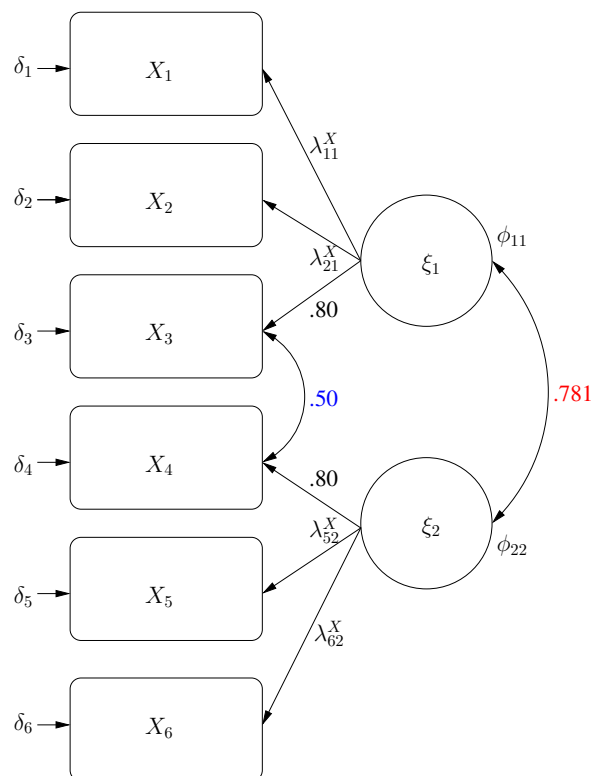


Abbildung 3.1: Veranschaulichung der Multikollinearitätsproblematik im Kontext unreliabler Prädiktoren (Attenuation)

Hierzu ein Beispiel: Seien zwei latente Merkmale ξ_1 und ξ_2 mit jeweils drei Indikatoren (s. Abbildung 3.1) gegeben; wobei X_1, X_2 und X_3 Operationalisierungen von

ξ_1 sowie X_4, X_5 und X_6 Operationalisierungen von ξ_2 seien. Dann entspricht eine (manifeste) Korrelation von .50 zwischen zwei Indikatoren X_3 und X_4 bei Faktorladungen von .80 einer latenten Korrelation von .781 (also etwa .80). Auf der Ebene latenter Variablen weisen die Prädiktoren ξ_1 und ξ_2 des Strukturmodells also weitaus größere Zusammenhänge auf.

In nicht-linearen Strukturgleichungsmodellen ist die Bedeutsamkeit der Multikollinearität noch größer, da hier prinzipiell zwei Quellen der Multikollinearität bestehen (Marquardt, 1980):

1. Nicht-essentielle Multikollinearität
2. Essentielle Multikollinearität

3.3.1 Nicht-essentielle Multikollinearität

Unter nicht-essentieller Multikollinearität versteht man jene Korrelation zwischen Variablen, die auf das Ausbleiben der Zentrierung der Prädiktoren vor der Produktbildung der nicht-linearen Terme (e.g. $\xi_1\xi_2$) zurückzuführen ist. Diese Form der Multikollinearität ist somit vermeidbar. Seien z.B. ξ_1 und ξ_2 normalverteilte latente Prädiktoren mit $E(\xi_1) = E(\xi_2) \neq 0$, dann sind

$$Cov(\xi_1, \xi_1\xi_2) = E(\xi_2)\phi_{11} + E(\xi_1)\phi_{21} \quad (3.1)$$

$$Cov(\xi_1, \xi_1^2) = 2E(\xi_1)\phi_{11} \quad (3.2)$$

Wie man erkennen kann, werden Bestandteile der Kovarianzen zwischen der linearen latenten Variablen ξ_1 und den nicht-linearen Termen ($\xi_1\xi_2$ oder ξ_1^2) auch dann nicht Null, wenn die Korrelation zwischen ξ_1 und ξ_2 Null ist. Diese dritten Momente werden bei normalverteilten linearen Variablen allerdings Null, wenn die Variablen vor der Produkttermbildung zentriert werden, so dass die Erwartungswerte der Variablen Null sind.

3.3.2 Essentielle Multikollinearität

Abgesehen von der nicht-essentiellen Multikollinearität gibt es Formen von Multikollinearität, die nicht durch lineare Transformationen beseitigt werden können. Diese Formen zählen zur essentiellen Multikollinearität. Eine Form resultiert daraus, dass die in die Produkte eingehenden Variablen nicht-normal verteilt sind. Sind die

Ausgangsvariablen (z.B. ξ_1 und ξ_2) nicht normalverteilt, so sind die dritten Momente (z.B. $Cov(\xi_1, \xi_1\xi_2)$) auch dann nicht Null, wenn die Korrelation der Ausgangsvariablen Null ist. Die Korrelation zwischen ξ_1 (resp. ξ_2) und den Produkttermen $\xi_1\xi_2$ und ξ_1^2 bleibt auch nach einer Zentrierung erhalten.

Eine weitere Form essentieller Multikollinearität besteht darin, dass *nicht-lineare Terme grundsätzlich untereinander in Zusammenhang stehen, sobald die linearen Ausgangsvariablen korrelieren*. Dies gilt auch für den Fall, dass die Ausgangsvariablen normalverteilt und zentriert sind. Erschwerend kommt hinzu, dass die Korrelation der nicht-linearen Terme untereinander sogar höher ist als die Korrelation der linearen Ausgangsvariablen.

Hierzu ein Beispiel: Seien ξ_1 und ξ_2 normalverteilte und zentrierte Variablen, dann beträgt die Kovarianz der nicht-linearen Terme untereinander (s. Gleichungen (3.3) - (3.5)):

$$Cov(\xi_1\xi_2, \xi_1^2) = 2\phi_{11}\phi_{21} \quad (3.3)$$

$$Cov(\xi_1\xi_2, \xi_2^2) = 2\phi_{22}\phi_{21} \quad (3.4)$$

$$Cov(\xi_1^2, \xi_2^2) = 2(\phi_{21})^2 \quad (3.5)$$

Dies ist in Abbildung 3.2 anhand einer latenten Prädiktorkorrelation von .50 veranschaulicht (vgl. Bohrnstedt & Goldberger, 1969). Wie man der Abbildung entnehmen kann, beträgt die Korrelation zwischen dem Interaktionsterm $\xi_1\xi_2$ und den quadratischen Termen ξ_1^2 und ξ_2^2 jeweils .63. Die quadratischen Terme ξ_1^2 und ξ_2^2 korrelieren zu .25 miteinander.

3.3.3 Implikationen des Multikollinearitätsproblems

Die Tatsache, dass nicht-lineare Terme zunehmend miteinander in Zusammenhang stehen, je höher die Korrelation der latenten Prädiktoren ist, bedeutet letztlich, dass sie zunehmend mehr *gemeinsame Varianz* aufweisen. Dies hat folgende Implikationen:

1. Je höher die Multikollinearität, desto schwieriger ist es, nicht-lineare Effekte getrennt voneinander aufzudecken, wenn sie gleichzeitig spezifiziert werden (Ganzach, 1997; MacCallum & Mar, 1995).
2. Als Konsequenz der gemeinsamen Varianz, können zudem konkurrierende nicht-spezifizierte „wahre“ nicht-lineare Effekte nicht aufgedeckt werden und statt dessen, spezifizierte „unwahre“ Effekte als Scheineffekte i.S. eines Fehlers erster

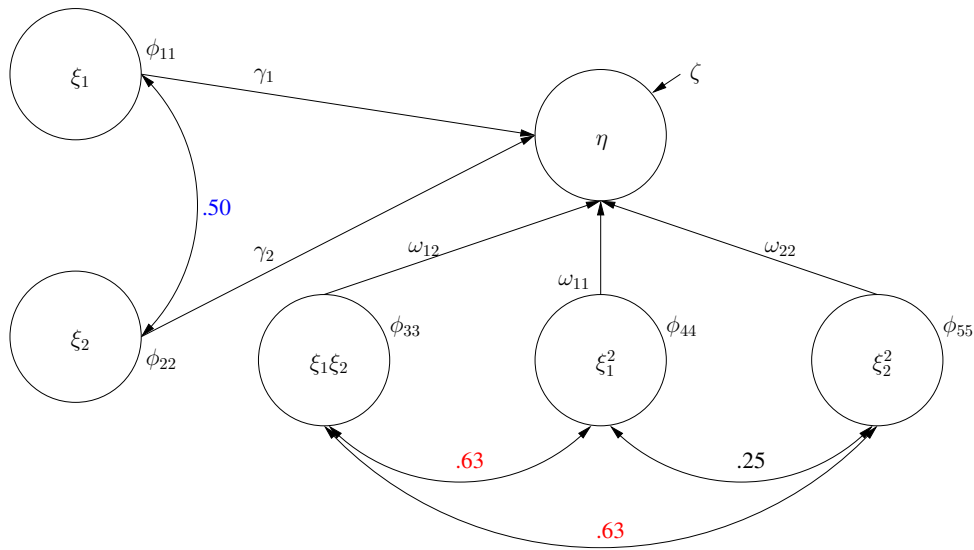


Abbildung 3.2: Multikollinearität im Strukturmodell: Korrelationen der latenten nicht-linearen Terme untereinander, wenn die Prädiktorkorrelation $.50$ beträgt

Art signifikant werden. Dies ist z.B. dann der Fall, wenn ein Interaktionseffekt anstelle eines wahren quadratischen Effektes spezifiziert wird und letzterer in der Folge nicht aufgedeckt wird.

Zusammenfassend lässt sich festhalten, dass Multikollinearität der Prädiktoren bei der Schätzung nicht-linearer Modelle per se ein größeres Problem darstellt als bei der Schätzung linearer Modelle. Daher müssen alle Verfahren zur Schätzung nicht-linearer Effekte daraufhin beurteilt werden, ob sie der Anforderung steigender Multikollinearität gerecht werden.

Wie im nachfolgenden Kapitel 4 gezeigt wird, enthält die Klasse der Produkt-Indikator-Ansätze Spezifikationsfehler, die zu Scheineffekten führen, wenn man nicht alle wahren nicht-linearen Effekte in ein Modell aufnimmt. Und selbst wenn alle potentiellen nicht-linearen Effekte in einem Modell enthalten sind, führt Multikollinearität – wie sich zeigen wird – zu Verzerrungen der Parameterschätzungen bei diesen Ansätzen. Multikollinearität erweist sich damit als grundsätzlich problematisch für diese Ansätze.

Kapitel 4

Spezifikationsfehler der Produkt-Indikator-Ansätze

Im vorhergehenden Kapitel 3 wurde gezeigt, dass Multikollinearität per se ein Problem darstellt, weil nicht-lineare Effekte (z.B. $\xi_1\xi_2$ und ξ_1^2) gemeinsame Varianz aufweisen, sobald die latenten Prädiktoren (z.B. ξ_1 und ξ_2) mit einander korrelieren. Dies erschwert die Aufdeckung der Effekte, wenn sie simultan geschätzt werden sollen, oder führt zu Scheineffekten, wenn statt der wahren nicht-linearen Effekte falsche nicht-lineare Effekte spezifiziert werden.

In diesem Kapitel soll darüber hinausgehend dargestellt werden, dass bereits die Ursprungspublikation von Kenny und Judd (1984) Spezifikationsfehler des Produkt-Indikator-Ansatzes enthält. Diese Spezifikationsfehler wurden nach Wissen des Autors bisher *nicht* entdeckt bzw. korrigiert. Vielmehr wurde der Ansatz in einer Erweiterung von Lee et al. (2004) zur simultanen Schätzung von Interaktionseffekten und quadratischen Effekten ebenfalls nicht korrekt spezifiziert. Spezifikationsfehler – und besonders solche aus der Ursprungsarbeit von Kenny und Judd (1984) – machen die ganze Ansatzklasse für Multikollinearität „anfällig“, so dass es zu verzerrten Schätzungen oder Scheineffekten kommen muss. Spezifikationsfehler erschweren über das allgemeine Multikollinearitätsproblem hinausgehend die Aufdeckung und die korrekte Schätzung der nicht-linearen Effekte.

Die Ziele dieses Kapitels lassen sich wie folgt zusammenfassen:

1. Zunächst wird aufgezeigt, welche Spezifikationsfehler in den Produkt-Indikator-Ansätzen enthalten sind und wie man die Ansätze korrigieren bzw. korrekt erweitern muss, wenn eine simultane Schätzung von nicht-linearen Interaktionseffekten und quadratischen Effekten erfolgen soll.

2. Da die latente Prädiktorkorrelation Bestandteil der Spezifikationsfehler ist, soll danach die zweite Fragestellung der Arbeit beantwortet werden, wie groß die Verzerrungen sind, die sich bei der Anwendung der (unkorrigierten) Produkt-Indikator-Ansätze ergeben, wenn die Multikollinearität steigt.

Diese Frage wird anhand einer Simulationsstudie mit dem Zwei-Stufen Ansatz von Ping (1996) und dem Ansatz von Jöreskog und Yang (1996) beantwortet, der vom Autor dieser Schrift korrekt erweitert wurde. Im Falle des Ansatzes von Ping ergeben sich verzerrte Schätzungen und Scheineffekte der nicht-linearen Effekte, da dieser die beschriebenen Spezifikationsfehler ausweist. Der korrekt erweiterte Ansatz von Jöreskog und Yang (1996) weist erwartungstreue Schätzungen auf. Die Schätzeigenschaften beider Ansätze werden vergleichend mit den Eigenschaften des verteilungsanalytischen LMS Ansatzes von Klein und Moosbrugger (2000) beurteilt.

Wesentliche Inhalte dieses Kapitels entstammen Schrift (B).

4.1 Zu den Spezifikationsfehlern der Produkt-Indikator-Ansätze und ihrer Korrektur

4.1.1 Zu den Spezifikationsfehlern bei Kenny und Judd (1984) und ihrer Korrektur

Wie man Kapitel 2 (Ansätze zur Schätzung latenter nicht-linearer Effekte) entnehmen kann, haben alle Produkt-Indikator-Ansätze gemein, dass sie Restriktionen über die Messfehlervarianzen formulieren. Dazu betrachten alle Autoren zunächst die linearen und die aus ihnen gebildeten nicht-linearen Indikatorvariablen.

Diesen Gedankengang wollen wir im Folgenden nachvollziehen: Seien im Kenny und Judd-Ansatz X_1 und X_2 normalverteilte, zentrierte Indikatoren des latenten Prädiktors ξ_1 sowie X_3 und X_4 normalverteilte, zentrierte Indikatoren des Prädiktors ξ_2 (s. Abbildung 2.1), dann dient X_1X_3 als ein Produkt-Indikator des latenten Interaktionstermes $\xi_1\xi_2$. X_1X_3 besteht aus folgenden Komponenten:

$$X_1X_3 = \underbrace{\lambda_{11}^X \lambda_{32}^X}_{\lambda_{53}^X} \xi_1 \xi_2 + \underbrace{\lambda_{32}^X \xi_2 \delta_1 + \lambda_{11}^X \xi_1 \delta_3 + \delta_1 \delta_3}_{\delta_5} \quad (4.1)$$

$$= \lambda_{53}^X \xi_1 \xi_2 + \delta_5 \quad (4.2)$$

Auf dieser Gleichung basierend werden folgende Restriktionen formuliert

$$Var(X_1X_3) = (\lambda_{53}^X)^2 Var(\xi_1\xi_2) + Var(\delta_5) \quad (4.3)$$

wobei

$$\lambda_{53}^X = \lambda_{11}^X \lambda_{32}^X \quad (4.4)$$

$$Var(\xi_1\xi_2) = \phi_{33} = \phi_{11}\phi_{22} + \phi_{21}^2 \quad (4.5)$$

$$Var(\delta_5) = \theta_{55}^\delta = (\lambda_{11}^X)^2 \phi_{11}\theta_{33}^\delta + (\lambda_{32}^X)^2 \phi_{22}\theta_{11}^\delta + \theta_{11}^\delta \theta_{33}^\delta \quad (4.6)$$

Diese Restriktionen der Messfehlervarianzen sind korrekt. Allerdings haben Kenny und Judd (1984) vergessen, Messfehlerkovarianzen bei der Schätzung eines erweiterten Strukturgleichungsmodells zu spezifizieren. Das von ihnen spezifizierte Strukturgleichungsmodell¹ enthält u.a. einen Interaktionsterm $\xi_1\xi_2$ und einen quadratischen Term ξ_1^2 . Die nicht-enthaltenen Messfehlerkovarianzen ergeben sich sofort, wenn man die zwei Indikatoren X_1X_3 und X_1^2 in ihre Komponenten zerlegt:

$$X_1X_3 = \lambda_{11}^X \lambda_{32}^X \xi_1\xi_2 + \lambda_{32}^X \xi_2\delta_1 + \lambda_{11}^X \xi_1\delta_3 + \delta_1\delta_3 \quad (4.7)$$

$$X_1^2 = (\lambda_{11}^X)^2 \xi_1^2 + 2\lambda_{11}^X \xi_1\delta_1 + \delta_1^2 \quad (4.8)$$

Wie man den Gleichungen (4.7) und (4.8) entnehmen kann, ist $\xi_2\delta_1$ Bestandteil von X_1X_3 und $\xi_1\delta_1$ Bestandteil von X_1^2 . Deren Kovarianz $Cov(\xi_2\delta_1, \xi_1\delta_1)$ ist im Allgemeinen nicht Null und damit die Nicht-Spezifikation dieser Kovarianz ein Spezifikationsfehler. Kenny und Judd haben übersehen, dass

$$Cov(\xi_2\delta_1, \xi_1\delta_1) = \phi_{21}\theta_{11}^\delta \quad (4.9)$$

ungleich Null ist, wenn $\phi_{21} \neq 0$ ist. *Wenn eine latente Prädiktorkorrelation (Multikollinearität) besteht, dann führt dieser Spezifikationsfehler zwangsläufig zu Verzerrungen in Form von nicht-erwartungstreuen Schätzungen der nicht-linearen Effekte, zu Scheineffekten oder zu nicht-signifikanten Effekten und damit potentiell zu gravierenden inhaltlichen Folgen.* Wenn man nach Kenny und Judd (1984) ein erweitertes nicht-lineares Strukturgleichungsmodell schätzen möchte, muss man also entgegen ihrer Anweisung auch die Kovarianzen der Messfehlerkomponenten berücksichtigen.

¹Das Originalmodell hatte bei Kenny und Judd (1984) auch einen kubischen Term, der aber für die weitere Betrachtung nicht relevant ist. Außerdem wurde erneut in der vorliegenden Schrift aus Konsistenzgründen eine Anpassung der Notation vorgenommen.

4.1.2 Zu den Spezifikationsfehlern bei Lee et al. (2004) und ihrer Korrektur

Schrift (B) kann man entnehmen, dass Lee et al. (2004) im Zuge der Durchführung einer ansatzvergleichenden Simulationsstudie ebenfalls ein Spezifikationsfehler unterläuft, der mit dem zuvor beschriebenen zusammenhängt². In ihrer Simulationsstudie vergleichen sie die Schätzeigenschaften diverser Ansätze bei der Schätzung des folgenden Strukturmodells (Gleichung (4.10)):

$$\eta = \alpha + \gamma_1\xi_1 + \gamma_2\xi_2 + \omega_{12}\xi_1\xi_2 + \omega_{11}\xi_1^2 + \omega_{22}\xi_2^2 + \zeta \quad (4.10)$$

Der LISREL-Notation folgend spezifizieren sie die Messfehlervarianzen der nicht-linearen Indikatoren (z.B. bei X_1X_3 siehe Gleichung (4.6)). In der sog. θ^δ -Matrix, die die Messfehlervarianzen und -kovarianzen enthält, spezifizieren sie jene Messfehlerkovarianzen nicht, die zu Produkt-Indikatoren unterschiedlicher nicht-linearer Indikatoren gehören. Seien z.B. wieder X_1X_3 ein Produkt-Indikator des Interaktionstermes $\xi_1\xi_2$ und X_1^2 ein Produkt-Indikator von ξ_1^2 , dann muss man die folgende Messfehlerkovarianz spezifizieren (Gleichung (4.11); siehe auch Schrift (B)):

$$Cov\left(\delta_{X_1^2}, \delta_{X_1X_3}\right) = 2\lambda_{11}^X\lambda_{32}^X\phi_{21}\theta_{11}^\delta \quad (4.11)$$

Auch hier führt die Nicht-Spezifikation der Messfehlerkovarianzen zwangsläufig zu verzerrten Schätzergebnissen bei der simultanen Schätzung von Interaktionseffekten und quadratischen Effekten, wenn die Multikollinearität ansteigt, weil die latente Prädiktorkorrelation Bestandteil dieses Kovarianzen ist. Die Ergebnisse von Lee et al. (2004) verlieren daher an Aussagekraft. Um die theoretischen Darstellungen dieses Abschnitts zu veranschaulichen, wurde eine Simulationsstudie durchgeführt.

4.2 Eine Simulationsstudie zur Veranschaulichung des Effekts der Spezifikationsfehler bei steigender Multikollinearität

Um zu veranschaulichen, wie sich die Spezifikationsfehler der Produkt-Indikator-Ansätze bei steigender Multikollinearität auf die Schätzungen der nicht-linearen Effekte auswirken, wurde eine Simulationsstudie durchgeführt. Die Durchführung

²Die in Abschnitt 4.1.1 beschriebenen Spezifikationsfehler bei Kenny und Judd (1984) sind nicht in Schrift (B) enthalten.

und ihre Ergebnisse sind in Schrift (B) detailliert beschrieben. In der vorliegenden Arbeit soll nur eine zusammenfassende Darstellung erfolgen.

4.2.1 Methode

In der Studie wurden drei Verfahren einander gegenübergestellt:

1. Der Zwei-Stufen Ansatz von Ping (1996)
2. Der korrekt erweiterte constrained Ansatz von Jöreskog und Yang (1996)
3. Der LMS Ansatz von Klein und Moosbrugger (2000)

Um die Auswirkung der Spezifikationsfehler zu überprüfen, wurde der Zwei-Stufen Ansatz verwendet. Dieser wurde von Ping (1996) mit dem Anspruch publiziert, auch für Terme höherer Ordnung (z.B. ξ_1^3) geeignet zu sein. Im Zwei-Stufen Ansatz sind die oben beschriebenen Spezifikationsfehler enthalten. Um eine Bezugsgröße innerhalb der Klasse der Produkt-Indikator-Ansätze zu haben, wurde der Ansatz von Jöreskog und Yang (1996) korrekt erweitert. Als Vergleich diente darüber hinausgehend der verteilungsanalytische LMS Ansatz von Klein und Moosbrugger (2000).

In der Studie wurden drei verschiedene Stufen von Multikollinearität untersucht, nämlich Prädiktorkorrelationen von .00, .50 und .80. Zudem wurden die wahren Modelle variiert. Es wurden innerhalb jeder Multikollinearitätsstufe Daten mit einer Stichprobengröße von $N = 400$ und eine Samplingrate von 500 generiert, die jeweils drei wahren Modellen folgten:

- einem Modell mit einem Interaktionsterm $\xi_1\xi_2$ oder
- einem Modell mit zwei quadratischen Termen ξ_1^2 und ξ_2^2 der latenten Prädiktoren ξ_1 und ξ_2 oder
- einem Modell mit einem Interaktionsterm $\xi_1\xi_2$ und zwei quadratischen Termen ξ_1^2 und ξ_2^2 .

Alle linearen Variablen (d.h. Indikatoren, Prädiktoren und Messfehler) wurden normalverteilt und zentriert generiert. Analysiert wurde stets ein Modell mit beiden Effekttypen, ungeachtet dessen, welche und wie viele nicht-lineare Effekte in dem wahren Modell enthalten waren. Tabelle 4.1 fasst die Bedingungen und die wahren Parameterausprägungen für alle Bedingungen zusammen.

4.2.2 Ergebnisse

Interaktionsmodell

Die Ergebnisse lassen sich für das *wahre Interaktionsmodell* wie folgt zusammenfassen (vgl. Tabelle 2 Schrift (B)): Solange die latenten Prädiktoren unkorreliert sind, wirkt sich die Nicht-Spezifikation der Messfehlerkovarianzen nicht nachteilig aus. Alle drei Ansätze, der Zwei-Stufen Ansatz, der erweiterte Jöreskog und Yang Ansatz und der LMS Ansatz kommen zu erwartungstreuen Schätzungen. Alle drei Ansätze sind teststark und halten die Typ-I-Fehlerrate von ca. 5 % ein.

Sobald die Korrelation der Prädiktoren ansteigt, erweisen sich die Schätzungen des Zwei-Stufen Ansatzes als instabil. Je nach Ausprägung der Prädiktorkorrelation kommt es zu Unter- oder Überschätzungen der nicht-linearen Effekte. Die Über- und Unterschätzungen der nicht-linearen Effekte sind dabei innerhalb des Modells entgegengerichtet. Werden z.B. die quadratischen Effekte unterschätzt, wird der Interaktionseffekt überschätzt (z.B. bei $\phi_{21} = .80$). In der Folge weist der Zwei-Stufen Ansatz eine erhöhte Rate von Scheineffekten auf. LMS und der erweiterte Jöreskog und Yang Ansatz zeigen hinsichtlich der Erwartungstreue vernachlässigbare Unterschiede. Beim erweiterten Jöreskog und Yang Ansatz kommt es infolge der verletzten multivariate Normalverteilungsannahme zu unterschätzten Standardfehlern der Parameterschätzungen. Die Unterschätzung bringt nur augenscheinliche, d.h. irreführende Vorteile hinsichtlich der Teststärke und ist nicht als Effizienzvorteil zu werten.

Quadratisches Modell

Bei der Schätzung des *wahren quadratischen Modells* kommt man zum selben Bild (vgl. Tabelle 3 Schrift (B)): Solange die latenten Prädiktoren unkorreliert sind, kommen alle drei Ansätze zu erwartungstreuen Schätzungen und sind teststark.

Bei steigender Korrelation der Prädiktoren werden beim Zwei-Stufen Ansatz allerdings die nicht-linearen Effekte stets entgegengesetzt über- oder unterschätzt. Dabei kommt es sogar zu unterschiedlichen Vorzeichen. Im Fall der Prädiktorkorrelation von .50 werden die quadratischen Effekte überschätzt und der (nicht-existierende) Interaktionseffekt mit einem negativen Vorzeichen geschätzt. In der Folge zeigt der Zwei-Stufen Ansatz also Scheineffekte mit falschem Vorzeichen oder günstigsten Falls verzerrte Effekte. LMS und der erweiterte Jöreskog und Yang Ansatz erweisen sich als gleichermaßen erwartungstreu. Allerdings findet man wieder zugunsten von LMS die präziseren Standardfehlerschätzungen.

Modell mit einem Interaktionseffekt und zwei quadratischen Effekten

Für das *wahre Modell mit allen Effekttypen* erbringen alle drei Ansätze erwartungstreue Schätzungen, wenn die latenten Prädiktoren unkorreliert sind (vgl. Tabelle 4.2). Die Teststärke der drei Ansätze ist hoch, nahezu 100 %. Dabei zeigt der Zwei-Stufen Ansatz die größten Unterschätzungen der Standardfehler.

Bei steigender Prädiktorkorrelation zeigen sich erneut verzerrte und instabile Parameterschätzungen für den Zwei-Stufen Ansatz. Während bei einer Korrelation von .50 der Interaktionseffekt beim Zwei-Stufen Ansatz unterschätzt und die quadratischen Effekte überschätzt werden, zeigen LMS und der erweiterte Jöreskog und Yang Ansatz erwartungstreue Schätzungen. Dabei sinkt die Teststärke des Zwei-Stufen Ansatzes stark im Vergleich zu den beiden anderen Verfahren ab, nicht-lineare Effekte zu entdecken.

Bei einer Korrelation von .80 kommt es zu einer hohen Überschätzung des Interaktionseffektes und zu einer starken Unterschätzung der quadratischen Effekte beim Zwei-Stufen Ansatz. Auch hier weisen LMS und der erweiterte Jöreskog und Yang Ansatz erwartungstreue Schätzungen auf. Der Zwei-Stufen Ansatz zeigt im Gegensatz zu den beiden anderen Verfahren verstärkt Scheineffekte, was sich in einer erhöhten Signifikanzrate der Interaktionseffekte äußert (keine korrekte Teststärke). LMS und der erweiterte Jöreskog und Yang Ansatz zeigen korrelationsbedingt Teststärkeverluste, wobei erneut die vom erweiterten Jöreskog und Yang Ansatz geschätzten Standardfehler zu klein ausfallen.

Die nachfolgende Abbildung 4.1 ist Schrift (B) entnommen und veranschaulicht die verzerrten Parameterschätzungen des Zwei-Stufen Ansatzes im Vergleich zu den Parameterschätzungen der beiden anderen Ansätze.

Tabelle 4.2: Simulationsstudie aus Schrift (B): Ergebnisse der wahren Modelle mit einem Interaktionseffekt und zwei quadratischen Effekten.

Parameter	wahrer Wert	mittlere Parameter-schätzung	Bias %	SD	SE	SE/SD	Typ-I-Fehlerrate	Teststärke
<i>Korrelation der latenten Prädiktoren: .00</i>								
Pings (1996) Zwei-Stufen Ansatz								
ω_{12}	0.224	0.222	0.90%	0.062	0.049 [†]	0.795 [†]	NA	100.00% [†]
ω_{11}	0.158	0.161	1.90%	0.036	0.035	0.969	NA	100.00%
ω_{22}	0.158	0.171	8.23%	0.041	0.035 [†]	0.851 [†]	NA	98.90% [†]
erweiterter Jöreskog & Yang (1996) Ansatz								
ω_{12}	0.224	0.216	-3.57%	0.049	0.046	0.935	NA	99.20%
ω_{11}	0.158	0.157	-0.63%	0.036	0.033	0.919	NA	99.40%
ω_{22}	0.158	0.156	-1.27%	0.038	0.033 [†]	0.853 [†]	NA	98.80% [†]
Klein & Moosbruggers (2000) LMS Ansatz								
ω_{12}	0.224	0.221	-1.34%	0.050	0.049	0.986	NA	98.80%
ω_{11}	0.158	0.161	1.90%	0.036	0.035	0.975	NA	99.40%
ω_{22}	0.158	0.160	1.27%	0.037	0.035	0.936	NA	98.80%
<i>Korrelation der latenten Prädiktoren: .50</i>								
Pings (1996) Zwei-Stufen Ansatz								
ω_{12}	0.200	0.166	-17.00%	0.112	0.112	0.992	NA	41.18%
ω_{11}	0.158	0.168	6.33%	0.060	0.060	1.005	NA	79.83%
ω_{22}	0.158	0.171	8.23%	0.058	0.062	1.061	NA	78.15%
erweiterter Jöreskog & Yang (1996) Ansatz								
ω_{12}	0.200	0.193	-3.50%	0.065	0.058 [†]	0.896 [†]	NA	88.82% [†]
ω_{11}	0.158	0.159	0.63%	0.041	0.037	0.905	NA	98.00%
ω_{22}	0.158	0.157	-0.63%	0.043	0.037 [†]	0.853 [†]	NA	97.21% [†]
Klein & Moosbruggers (2000) LMS Ansatz								
ω_{12}	0.200	0.196	-2.00%	0.067	0.063	0.942	NA	84.00%
ω_{11}	0.158	0.163	3.16%	0.042	0.041	0.971	NA	97.80%
ω_{22}	0.158	0.160	1.27%	0.043	0.040	0.948	NA	97.20%
<i>Korrelation der latenten Prädiktoren: .80</i>								
Pings (1996) Zwei-Stufen Ansatz								
ω_{12}	0.175	0.260	48.57%	0.084	0.074 [†]	0.883 [†]	NA	90.86% [†]
ω_{11}	0.158	0.105	-33.54%	0.053	0.050	0.940	NA	58.88%
ω_{22}	0.158	0.106	-32.91%	0.054	0.050	0.921	NA	58.38%
erweiterter Jöreskog & Yang (1996) Ansatz								
ω_{12}	0.175	0.172	-1.71%	0.143	0.121 [†]	0.848 [†]	NA	32.73% [†]
ω_{11}	0.158	0.157	-0.63%	0.078	0.067 [†]	0.854 [†]	NA	62.67% [†]
ω_{22}	0.158	0.154	-2.53%	0.080	0.066 [†]	0.835 [†]	NA	62.28% [†]
Klein & Moosbruggers (2000) LMS Ansatz								
ω_{12}	0.175	0.163	-6.86%	0.166	0.155	0.933	NA	22.80%
ω_{11}	0.158	0.168	6.33%	0.091	0.086	0.940	NA	50.00%
ω_{22}	0.158	0.163	3.16%	0.089	0.084	0.949	NA	51.60%

Note. SE = mittlerer Standardfehler; NA = nicht analysierbar; † = Ergebnisse basieren auf einem verzerrten Standardfehler: ($|Bias| > 10\%$)

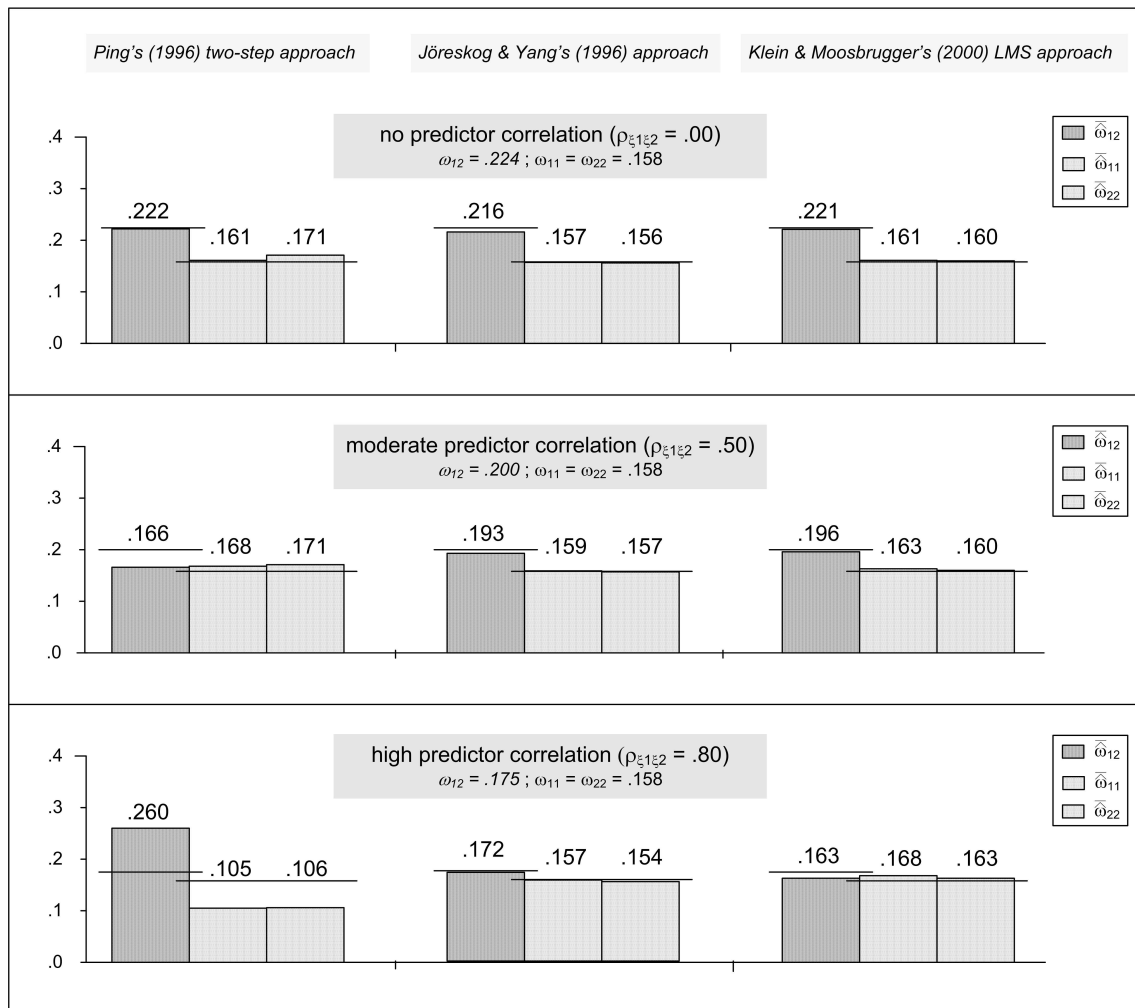


Abbildung 4.1: Simulationsstudie aus Schrift (B): Mittlere Parameterschätzungen eines Modells mit einem Interaktionseffekt und zwei quadratischen Effekten bei steigender Multikollinearität. Die horizontalen Linien entsprechen den wahren Parameterwerten.

4.2.3 Implikationen

Die Ergebnisse der Simulationsstudie verdeutlichen sehr stark die Relevanz der Spezifikation von Messfehlerkovarianzen über die Produkt-Indikatoren der latenten nicht-linearen Terme hinweg. Wenn man Interaktionseffekte und quadratische Effekte simultan in einem Modell anhand eines Produkt-Indikator-Ansatzes schätzen möchte, führt die Nicht-Spezifikation der Kovarianzen (wie im Falle des Zwei-Stufen Ansatzes von Ping (1996)) je nach Prädiktorkorrelation zu gravierenden Verzerrungen der nicht-linearen Effekte, zu Scheineffekten oder sogar zu umgekehr-

ten Vorzeichen der wahren Effekte. Sobald eine latente Prädiktorkorrelation nicht Null ist, was zweifellos für die meisten sozialwissenschaftlichen Untersuchungen gilt, sind verzerrte Schätzungen der nicht-linearen Effekte bei der Nicht-Spezifikation der Messfehlerkovarianzen zu erwarten. Dies gilt nicht nur für jene Varianten der Produkt-Indikator-Ansätze, die die Messfehlervarianzen in Form von Restriktionen spezifizieren, sondern auch für die Produkt-Indikator-Ansätze, die die Restriktionen aufheben und die zuvor restringierten Parameter frei schätzen lassen. So zeigen Moosbrugger, Schermelleh-Engel, Kelava und Klein (2008) für den unconstrained Ansatz von Marsh et al. (2004), dass es zu Verzerrungen der Parameterschätzungen kommt, wenn man diesen für die simultane Schätzung von Interaktionseffekten und quadratischen Effekten erweitert und die Messfehlerkovarianzen nicht frei schätzt.

Im Hinblick auf inhaltliche Anwendungen können solche Verzerrungen zu Ergebnissen führen, die der Wahrheit sogar widersprechen oder aber gesellschaftspolitisch schwierig sein können. Man stelle sich vor, dass im Rahmen von Bildungsstudien irrtümlicherweise herauskommen würde, dass ausschließlich die Schulbildung des Vaters (pos. quadratisch) zur guten Prognose der Schulbildung des Kindes beitrage, während die Schulbildung der Mutter (bei einem zusätzlichen neg. Interaktionseffekt) abträglich für die Bildungsprognose des Kindes sei.

Im nächsten Kapitel sollen die neuen verteilungsanalytischen Ansätze LMS (Klein & Moosbrugger, 2000) und QML (Klein & Muthén, 2007) den (korrekt erweiterten) klassischen Produkt-Indikator-Ansätzen gegenübergestellt werden. Dazu werden zunächst die theoretischen Eigenschaften der verteilungsanalytischen Verfahren, LMS und QML, untereinander und mit der Klasse der Produkt-Indikator-Ansätze verglichen. Es werden ihre theoretischen Eigenschaften beim Vorliegen von Multikollinearität betrachtet. Illustriert werden die Unterschiede anhand einer Simulationsstudie, die die Ansätze hinsichtlich ihrer Schätzeigenschaften bei steigender Multikollinearität miteinander vergleicht. Dabei treten die Vorteile der verteilungsanalytischen Verfahren, LMS und QML, hervor.

Kapitel 5

Zum theoretischen und empirischen Vergleich der verteilungsanalytischen Ansätze LMS und QML mit den Produkt-Indikator-Ansätzen

In diesem Kapitel soll die dritte Fragestellung der Arbeit beantwortet werden, welche theoretischen und empirischen Unterschiede zwischen den neuen verteilungsanalytischen Ansätzen LMS (Klein & Moosbrugger, 2000) und QML (Klein & Muthén, 2007), einerseits, und den klassischen Produkt-Indikator-Ansätzen, andererseits, bei steigender Multikollinearität zu erwarten sind.

Um diese Frage zu beantworten, ist es notwendig, sich die unterschiedlichen Funktionsweisen der verteilungsanalytischen Ansätze zu vergegenwärtigen. Daher werden erst die verteilungsanalytischen Ansätze untereinander verglichen. Anschließend werden die theoretischen Eigenschaften der verteilungsanalytischen Ansätze den theoretischen Eigenschaften der (korrekt erweiterten) Produkt-Indikator-Ansätze gegenübergestellt. Es werden dabei die theoretischen Unterschiede der Ansätze beim Vorliegen von Multikollinearität betrachtet. Anhand einer Simulationsstudie mit steigender Multikollinearität werden die Unterschiede illustriert, die sich aufgrund der theoretischen Überlegungen ergeben. Die Simulationsstudie bestätigt die dargestellten Vorteile der verteilungsanalytischen Verfahren.

Dieses Kapitel fasst die sehr technische und ausführliche Darstellung von Schrift (C) zusammen. Weitere Details sind Schrift (C) oder aber dem ihr zugehörigen technischen Appendix zu entnehmen.

5.1 Theoretischer Vergleich der LMS und QML Ansätze

5.1.1 Funktionsweise von LMS

Klein und Moosbrugger (2000) haben den verteilungsanalytischen LMS Ansatz zur simultanen Schätzung von Interaktionseffekten und quadratischen Effekten entwickelt. Diesem Ansatz liegt die Analyse des multivariaten Indikatorvektors $(\mathbf{X}, \mathbf{Y}) = (X_1, \dots, X_q, Y_1, \dots, Y_p)$ zugrunde, der alle q X -Indikatoren und alle p Y -Indikatoren zusammenfasst. Die multivariate Nicht-Normalverteilung von (\mathbf{X}, \mathbf{Y}) wird durch eine endliche Mischverteilung von bedingten Normalverteilungen angenähert. Dazu werden modellimplizite Mittelwerte und Kovarianzmatrizen jeder einzelnen bedingten Normalverteilung verwendet. Die Schlüsselidee dieses Ansatzes ist, dass man die nicht-linearen Terme (e.g., $\xi_1\xi_2$ und ξ_1^2) durch bedingte, lineare Effekte erklären kann. Dazu muss man z.B. die Gesamtverteilung von (\mathbf{X}, \mathbf{Y}) auf jeder möglichen Stufe von ξ_1 betrachten. Maximum-Likelihood-Schätzungen der unbekanntem Modellparameter werden unter Verwendung des Expectation Maximization (EM) Algorithmus (Dempster et al., 1977) berechnet.

Abbildung 5.1 veranschaulicht die endliche Mischverteilung der Indikatorvariable Y_1 anhand von drei bedingten Normalverteilungen, die nach einer Gewichtung als Summe die finite Mischverteilung repräsentieren. Die latente Prädiktorkorrelation beträgt .80. Wie man erkennen kann, haben die drei bedingten Normalverteilungen (farbig) jeweils unterschiedliche Mittelwerte und Varianzen. Die aus ihnen resultierende Mischverteilung (schwarz) ist dann nicht-normalverteilt. Die rechte Teil ihres Graphen nähert sich langsamer der Abszissenachse an, als der linke Teil. Dies veranschaulicht die Repräsentation der nicht-Normalverteilung der Indikatorvariable Y_1 .

5.1.2 Funktionsweise von QML

Klein und Muthén (2007) entwickelten den Quasi-Maximum Likelihood Ansatz zur simultanen Schätzung von multiplen nicht-linearen Interaktionseffekten und quadratischen Effekten. Dies geschah mit dem Anspruch, ein robusteres Verfahren zu entwickeln, wenn die Prädiktorvariablen nicht-normalverteilt sind. Wie LMS trägt QML dem Gesichtspunkt Rechnung, dass der multivariate Indikatorvektor (\mathbf{X}, \mathbf{Y}) nicht-normalverteilt ist. Das trifft insbesondere auf das latente Kriterium η und seine

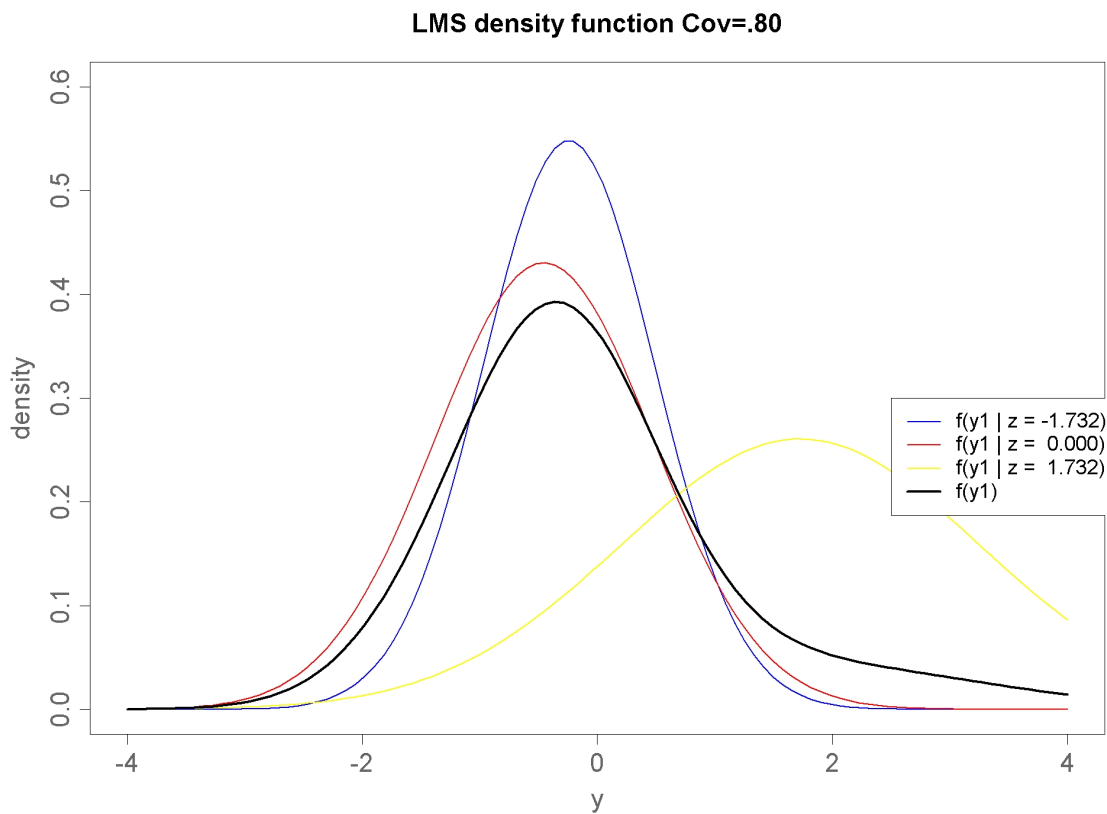


Abbildung 5.1: LMS Ansatz: Veranschaulichung der Y_1 -Indikatorverteilung als finite Mischverteilung von drei bedingten Normalverteilungen (entnommen aus Schrift (C))

Indikatoren zu. Um eine Vereinfachung der Verteilung $f(\mathbf{X}, \mathbf{Y})$ des Indikatorvektors (\mathbf{X}, \mathbf{Y}) vorzunehmen, transformiert QML zunächst den Indikatorvektor derartig, dass nur noch die Skalierungsvariable Y_1 nicht-normalverteilt ist und alle anderen Variablen im transformierten Indikatorvektor $(\mathbf{X}, \mathbf{Y}) = (\mathbf{X}, Y_1, \mathbf{Y}^*)$ normalverteilt sind. Danach werden der Mittelwert und die Varianz der nicht-normalen Verteilung von $(Y_1|\mathbf{X}, \mathbf{Y}^*)$ bestimmt. Anhand des Mittelwertes und der Varianz wird die nicht-normale Verteilung $f(Y_1|\mathbf{X}, \mathbf{Y}^*)$ durch eine Normalverteilung $f^*(Y_1|\mathbf{X}, \mathbf{Y}^*)$ mit dem selben Mittelwert und Varianz approximiert bzw. ersetzt. Diese Ersetzung ermöglicht es, die nicht-normale (Gesamt-)Verteilung $f(\mathbf{X}, \mathbf{Y})$ des Indikatorvektors durch das Produkt einer normalverteilten und einer bedingt normalverteilten Verteilung

zu approximieren:

$$f(\mathbf{X}, \mathbf{Y}) = f(\mathbf{X}, Y_1, \mathbf{Y}^*) = f(\mathbf{X}, \mathbf{Y}^*)f(Y_1|\mathbf{X}, \mathbf{Y}^*) \quad (5.1)$$

$$\approx f(\mathbf{X}, \mathbf{Y}^*)f^*(Y_1|\mathbf{X}, \mathbf{Y}^*) = f^*(\mathbf{X}, \mathbf{Y}) \quad (5.2)$$

Die approximierte Verteilung $f^*(\mathbf{X}, \mathbf{Y})$ ist nicht-normal. Allerdings besteht sie nun aus einer Normalverteilung $f(\mathbf{X}, \mathbf{Y}^*)$ und einer bedingten Normalverteilung $f^*(Y_1|\mathbf{X}, \mathbf{Y}^*)$, so dass für sie nun eine leichtere numerische Schätzung der Parameter möglich ist. Die Maximierung der Quasi-Likelihood $f^*(\mathbf{X}, \mathbf{Y})$ erfolgt anhand des Newton-Raphson Algorithmus (Freund & Hoppe, 2007; Isaacson & Keller, 1966).

5.1.3 Unterschiede zwischen den verteilungsanalytischen Ansätzen und ihre Konsequenzen

1. LMS basiert auf der Annahme, dass alle Prädiktor- und Residuumsvariablen normalverteilt sind. Die nicht-normale Indikatorverteilung $f(\mathbf{X}, \mathbf{Y})$ wird durch eine endliche Mischverteilung von bedingten Normalverteilungen angenähert. Im Gegensatz dazu beruht QML auf der Annahme der Normalverteilung nur insofern, dass die bedingte Verteilung von $(Y_1|\mathbf{X}, \mathbf{Y}^*)$ durch eine Normalverteilung mit dem selben Mittelwert und der derselben Varianz ersetzt werden kann. Der verbleibende Teil von $f(\mathbf{X}, \mathbf{Y})$, nämlich $f(\mathbf{X}, \mathbf{Y}^*)$, wird nicht angenähert¹.
2. Die Konsequenz dieses Unterschiedes ist, dass LMS und QML nahezu identische Schätzungen hervorbringen sollten, solange die *Prädiktor und Messfehlervariablen normalverteilt* sind. LMS sollte nur theoretisch einen vernachlässigbaren Vorteil haben, da QML eine Quasi-Likelihood mit einer approximierten Dichtefunktion $f^*(Y_1|\mathbf{X}, \mathbf{Y}^*)$ anstelle der „korrekten“ Likelihood maximiert. *Ein marginaler Nachteil sollte für QML vorhanden sein, wenn die Multikollinearität ansteigt*, da es zu einer höheren Nicht-Normalität des Y_1 -Indikators führt. Dies wiederum bedeutet, dass die normalverteilte Annäherung $f^*(Y_1|\mathbf{X}, \mathbf{Y}^*)$ stärker von der nicht-normalen Verteilung $f(Y_1|\mathbf{X}, \mathbf{Y}^*)$ abweicht. Bei LMS wird die Indikatorverteilung exakter durch die Mischverteilung modelliert (bei normalverteilten Prädiktorvariablen).

Wenn hingegen *die Prädiktor und Messfehlervariablen nicht-normalverteilt* sind, sollte QML robuster sein und genauere Schätzungen als LMS liefern, da LMS

¹Prinzipiell können auch andere Verteilungen als die Normalverteilung zur Approximation von $f(Y_1|\mathbf{X}, \mathbf{Y}^*)$ verwendet werden.

auf der Normalverteilungsannahme basiert, wenn die Mischverteilung modelliert wird (vgl. auch Schrift (C)). QML beruht nur insofern auf der Normalverteilungsannahme bzgl. der approximierten Dichte $f^*(Y_1|\mathbf{X}, \mathbf{Y}^*)$, so dass hier bessere Schätzungen zu erwarten sind, weil weniger Verteilungsvoraussetzungen verletzt werden. Aus dem selben Grunde sollten auch *Vorteile für QML bei steigender Multikollinearität und nicht-normalen Prädiktorverteilungen gegeben sein.*

3. Beim LMS Ansatz wird die finite Mischverteilungswahrscheinlichkeit unter Anwendung des Expectation Maximization (EM) Algorithmus maximiert (Dempster et al., 1977). Der EM Algorithmus maximiert eine Summe von gewichteten Normalverteilungen der Beobachtungen $(\mathbf{X}_i, \mathbf{Y}_i)$. In Gegensatz dazu wird die angenäherte (Log)Likelihood im QML Ansatz direkt maximiert, um Parameterschätzungen zu erhalten.
4. Die Konsequenz der unterschiedlichen Schätzmethode ist, dass der rechnerische Aufwand bei LMS stärker als bei QML ansteigt, wenn die Zahl der nicht-linearen Terme zunimmt. Da die Mischverteilung $f(\mathbf{X}, \mathbf{Y})$ im LMS Ansatz komplizierter wird, werden bei steigender Zahl der nicht-linearen Terme mehr sog. Stützstellen zur Approximation gebraucht. Je mehr Stützstellen gebraucht werden, umso höher ist der Rechenaufwand. Die direkte Maximierung der Quasi-Log-Likelihood in QML führt zu einer geringeren Steigerung des Rechenaufwandes.

5.2 Theoretischer Vergleich der verteilungsanalytischen Ansätze und der Produkt-Indikator-Ansätze

5.2.1 Theoretische Unterschiede

Es gibt im wesentlichen zwei gravierende Unterschiede zwischen den Produkt-Indikator-Ansätzen und den verteilungsanalytischen Ansätzen, woraus verschiedene Konsequenzen im Hinblick auf Multikollinearität resultieren.

1. Der erste Unterschied betrifft die Spezifikation des nicht-linearen Messmodells. Produkt-Indikator-Ansätze verwenden dem Namen nach Produkte von Indikatoren. LMS und QML brauchen zur Spezifikation des Messmodells keine

Indikator-Produkte. LMS umgeht die Spezifikation des nicht-linearen Messmodells, indem die Verteilung von (\mathbf{X}, \mathbf{Y}) als Mischverteilung repräsentiert wird. QML umgeht die Spezifikation dadurch, dass die Verteilung des Indikatorvektors (\mathbf{X}, \mathbf{Y}) unter Anwendung des Quasi-Likelihood Prinzips approximiert wird.

2. Der zweite Unterschied betrifft die Parameterschätzung des nicht-linearen Strukturgleichungsmodells. Da die Messmodellspezifikation verschieden ist, maximieren die Ansatzklassen gänzlich verschiedene Likelihoodfunktionen. Produkt-Indikator-Ansätze schätzen den Parametervektor Θ , indem sie eine Fitfunktion der folgenden Form minimieren, die primär bei linearen Strukturgleichungsmodellen zur Anwendung kommt:

$$F_{ML}(\Theta) = \log \|\Sigma\| + tr(\mathbf{S}\Sigma^{-1}) - \log \|\mathbf{S}\| - (p + q) + ((\bar{\mathbf{Y}}, \bar{\mathbf{X}}, \bar{\mathbf{X}}'\bar{\mathbf{X}})' - \mu)' \Sigma^{-1} ((\bar{\mathbf{Y}}, \bar{\mathbf{X}}, \bar{\mathbf{X}}'\bar{\mathbf{X}})' - \mu) \quad (5.3)$$

wobei Σ und \mathbf{S} die modellimpliziten und empirischen Kovarianzmatrizen der Indikatorvariablen und der Produkt-Indikatoren sind. Die Fitfunktion enthält daher Kovarianzen der nicht-linearen Indikatoren (e.g. X_1X_4 , X_2X_5 , ..., X_1^2 , ..., X_4^2 , ...). μ und $(\bar{\mathbf{Y}}, \bar{\mathbf{X}}, \bar{\mathbf{X}}'\bar{\mathbf{X}})$ sind die modellimpliziten und empirischen Mittelwerte der Indikatorvariablen. ML-Schätzungen erhält man, indem man die Fitfunktion aus Gleichung (5.3) minimiert. Diese basiert auf der (verletzten) Annahme der multivariaten Normalverteilung der Variablen.

5.2.2 Konsequenzen der Unterschiede

Aus diesen Unterschieden resultieren drei Konsequenzen im Hinblick auf Multikollinearität:

1. Die erste Konsequenz resultiert aus der Verletzung der Annahme einer multivariaten Normalverteilung bei den Produkt-Indikator-Ansätzen. Wenn man im Rahmen eines Produkt-Indikator-Ansatzes eine ML-Schätzung durchführt, ist diese Annahme stets verletzt. Selbst wenn nur ein (nicht-linearer) Interaktionseffekt $\xi_1\xi_2$ vorhanden ist, so ist $\xi_1\xi_2$ nicht-normalverteilt (Aroian, 1944; Moosbrugger, Schermelleh-Engel & Klein, 1997). *Diese Nicht-Normalität nimmt mit steigender Multikollinearität der Prädiktoren zu.* Die Nicht-Normalverteilung

des latenten Terms $\xi_1\xi_2$ führt wiederum zu nicht-normalverteilten Produkt-Indikatoren (z.B. X_1X_4) und nicht-normalverteilten Y -Indikatoren. Je mehr nicht-lineare Terme z.B. aufgrund von quadratischen Effekten (e.g., ξ_1^2 und/oder ξ_2^2) vorhanden sind, umso mehr wird diese Normalverteilungsannahme verletzt; erst recht bei Multikollinearität. Solange die nicht-linearen Effekte und die Stichprobengröße nicht zu groß sind, werden die nicht-linearen Effekte unverzerrt geschätzt (Schermelleh-Engel et al., 1998). Dies konnte man für den korrekte erweiterten Jöreskog und Yang (1996) Ansatz beobachten (vgl. Kapitel 4 oder Schrift (B)). *Allerdings werden dann die Standardfehler bei steigender Multikollinearität zunehmend unterschätzt.* Die Unterschätzung der Standardfehler führt bei den Produkt-Indikator-Ansätzen zu einer gesteigerten Typ-I-Fehlerrate und zu falschen Schlüssen aufgrund zu kleiner Konfidenzintervalle. Da LMS und QML die Nicht-Normalität explizit in ihren Likelihoods spezifizieren, sind ihre Standardfehlerschätzungen unverzerrt und effizient. Das gilt auch für den Fall steigender Multikollinearität (wie im nächsten Abschnitt anhand der Simulationsstudie belegt wird).

2. Die zweite Konsequenz der oben beschriebenen Unterschiede kann als Sampling Problem betrachtet werden. Produkt-Indikator-Ansätze beruhen darauf, dass die Kovarianzmatrizen lineare Indikatoren und nicht-lineare Produkt-Indikatoren enthalten. Die empirischen Kovarianzen und Varianzen der Produkt-Indikatoren sind allerdings instabil, da sie auf Momenten vierter Ordnung basieren (Jöreskog & Yang, 1996). Um stabile Kovarianzmatrizen zu erhalten, muss die Stichprobengröße relativ groß sein. Da LMS and QML nicht auf Momenten vierter Ordnung basieren, sind ihre Schätzungen reliabler bei kleineren Stichprobengrößen als die Schätzungen der Produkt-Indikator-Ansätze (Schermelleh-Engel et al., 1998). Da *Multikollinearität* auch aus einer zu kleinen Stichprobe resultieren kann (vgl. Kapitel 3), sind gerade Produkt-Indikator-Ansätze hierbei im Nachteil. Die Multikollinearität der linearen Indikatoren wird dann in noch stärkerer Form in die instabilen Kovarianzmatrizen übertragen, die die Produkt-Indikatoren enthalten. Die Produkt-Indikatoren weisen ihrerseits ohnehin schon essentielle Multikollinearität auf (vgl. Kapitel 3). Dieses Problem betrifft die verteilungsanalytischen Ansätze nicht, da sie keine Produkt-Indikatoren nutzen. *Die Schätzungen der Produkt-Indikator-Ansätze werden somit bei Multikollinearität unreliabler als die Schätzungen der verteilungsanalytischen Ansätze.*

3. Die dritte Konsequenz ist ein Effizienzproblem, welches in Zusammenhang mit den unterschätzten Standardfehlern der Produkt-Indikator-Ansätze steht. Um das Problem der nicht-normalverteilten Indikatoren zu umgehen, kann man robuste Schätzmethoden anwenden (z.B. WLSA, Jöreskog & Yang, 1996; alternativ eine Satorra-Bentler Korrektur der Standardfehler, Satorra & Bentler, 2001). Die Verwendung von robusten Schätzmethoden geht allerdings mit weniger effizienten Standardfehlerschätzungen bei Produkt-Indikator-Ansätzen und daher mit einer verminderten Teststärke einher (Schermelleh-Engel et al., 1998). Da LMS und QML ML-Schätzungen bieten, sind sie effizienter. Produkt-Indikator-Ansätze werden daher sehr viel größere Stichproben benötigen, um eine vergleichbare Teststärke zu haben. *Bei Vorliegen von Multikollinearität wären im Rahmen der psychologischen Forschung selten realisierbare Stichprobengrößen von jenseits 800 Fällen nötig, um eine ähnliche Teststärke zu erlangen und valide inferenzstatistische Schlüsse zu ermöglichen.*

5.3 Eine Simulationsstudie zur Veranschaulichung der empirischen Leistungsfähigkeit von verteilungsanalytischen Ansätzen und Produkt-Indikator-Ansätzen bei Multikollinearität

Um zu veranschaulichen, wie sich steigende Multikollinearität auf die theoretisch erwartete Leistungsfähigkeit der verteilungsanalytischen Ansätze und der Produkt-Indikator-Ansätze auswirkt, wurde eine Simulationsstudie durchgeführt. Ihre genaue Durchführung und ihre Ergebnisse sind in Schrift (C) detailliert beschrieben. In der vorliegenden Arbeit soll nur eine zusammenfassende Darstellung erfolgen.

5.3.1 Methode

In der Studie wurden drei Verfahren einander gegenübergestellt:

1. Der LMS Ansatz von Klein und Moosbrugger (2000)
2. Der QML Ansatz von Klein und Muthén (2007)
3. Der unconstrained Ansatz von Marsh et al. (2004)

Der unconstrained Ansatz ist der am einfachsten zu realisierende Produkt-Indikator-Ansatz, da er alle Parameterrestriktionen aufhebt und stattdessen eine freie Schätzung vornimmt. Der Ansatz wurde gemäß den Darstellungen im vorangegangenen Kapitel hinsichtlich der frei zu schätzenden Messfehlerkovarianzen erweitert, d.h. korrekt spezifiziert.

In der Studie wurden drei verschiedene Stufen von Multikollinearität untersucht, nämlich Prädiktorkorrelationen zwischen ξ_1 und ξ_2 von .00, .50 und .80. Auf jeder Stufe der Multikollinearität war das wahre Modell stets ein Modell, das einen Interaktionseffekt und zwei quadratische Effekte enthielt. Die linearen Effekte erklärten 10% der Kriteriumsvarianz; die nicht-linearen Effekte erklärten jeweils 5% Kriteriumsvarianz. Auf der Grundlage jedes wahren Modells wurden 500 Mal Daten mit einer Stichprobengröße von $N = 400$ generiert. Die latenten linearen Prädiktoren, Messfehler und Modellfehler waren normalverteilt und zentriert. Tabelle 5.1 gibt einen Überblick über die wahren Parameterwerte für die drei wahren Modelle. Analysiert wurde stets ein korrekt spezifiziertes Modell mit einem Interaktionseffekt und zwei quadratischen Effekten.

5.3.2 Ergebnisse

Wie man Tabelle 5.2 Teil A entnehmen kann, zeigen alle drei Ansätze erwartungstreue Schätzungen der nicht-linearen Effekte, wenn die latenten Prädiktoren unkorreliert sind. Die Teststärke zur Aufdeckung des Interaktionseffektes und der zwei quadratischen Effekte beträgt nahezu 100% für jeden Ansatz. LMS und QML bringen genaue Standardfehlerschätzungen hervor, während der unconstrained Ansatz eine leichte Unterschätzung des Standardfehlers des zweiten quadratischen Effektes aufweist.

Wenn die latente Prädiktorkorrelation auf .50 ansteigt, erhält man mit allen drei Verfahren akzeptable Parameterschätzungen (siehe Tabelle 5.2, Teil B). Der unconstrained Ansatz überschätzt leicht die beiden quadratischen Effekte zu 5.06% und 3.80%. Im Gegensatz dazu steigt der Bias der quadratischen Effekte bei LMS and QML nicht über 3.16%. Der unconstrained Ansatz zeigt auch die höchste empirische Standardabweichung dieser Parameterschätzungen. Keine Unterschiede zwischen den drei Ansätzen finden sich hinsichtlich der Schätzung des Interaktionseffektes oder seines Standardfehlers. Die Teststärke zur Aufdeckung des Interaktionseffektes ist für alle drei Ansätze größer als 80%. Die Teststärke zur Aufdeckung der quadratischen Effekte beträgt 96% - 98% für alle Ansätze.

Tabelle 5.1: Simulationsstudie aus Schrift (C): Drei wahre nicht-lineare Modelle: Variation der Multikollinearität.

	Keine Korrelation von ξ_1 und ξ_2 (.00)	Mittlere Korrelation von ξ_1 und ξ_2 (.50)	Hohe Korrelation von ξ_1 und ξ_2 (.80)
Parameter			
ω_{11}	0.158	0.158	0.158
ω_{12}	0.224	0.200	0.175
ω_{22}	0.158	0.158	0.158

Anmerkung: Die folgenden Parameter waren über die Bedingungen hinweg konstant: $Var(\xi_1) = Var(\xi_2) = 1$; $\gamma_1 = \gamma_2 = .316$; $\lambda_{11}^X = \lambda_{42}^X = \lambda_{11}^Y = 1$;
 $\lambda_{21}^X = \lambda_{31}^X = \lambda_{52}^X = \lambda_{62}^X = \lambda_{21}^Y = \lambda_{31}^Y = .894$; $\theta_{11}^\epsilon = \theta_{44}^\epsilon = \theta_{11}^\delta = \theta_{44}^\delta = \theta_{33}^\delta = \theta_{55}^\delta = \theta_{66}^\delta = \theta_{22}^\epsilon = \theta_{33}^\epsilon = .20$.

Tabelle 5.2: Simulationsstudie aus Schrift (C): Simulationsergebnisse für drei Modelle mit einem Interaktionseffekt und zwei quadratischen Effekten im Kontext ansteigender Multikollinearität.

Parameter	wahrer Wert	mittlere Parameter-schätzung	Bias %	SD	SE	SE/SD	Teststärke
<i>(A) Korrelation der latenten Prädiktoren: .00</i>							
Klein & Moosbruggers (2000) LMS Ansatz							
ω_{11}	0.158	0.161	1.90%	0.036	0.035	0.975	99.40%
ω_{12}	0.224	0.221	-1.34%	0.050	0.049	0.986	98.80%
ω_{22}	0.158	0.160	1.27%	0.037	0.035	0.936	98.80%
Klein & Muthéns (2007) QML Ansatz							
ω_{11}	0.158	0.162	2.53%	0.037	0.036	0.973	99.20%
ω_{12}	0.224	0.220	-1.79%	0.050	0.049	0.980	98.60%
ω_{22}	0.158	0.159	0.63%	0.036	0.035	0.972	99.80%
Marsh et al.s (2004) unconstrained Ansatz							
ω_{11}	0.158	0.164	3.80%	0.038	0.036	0.947	99.40%
ω_{12}	0.224	0.223	-0.45%	0.052	0.051	0.981	99.20%
ω_{22}	0.158	0.163	3.16%	0.040	0.036 [†]	0.900 [†]	99.80% [†]
<i>(B) Korrelation der latenten Prädiktoren: .50</i>							
Klein & Moosbruggers (2000) LMS Ansatz							
ω_{11}	0.158	0.163	3.16%	0.042	0.041	0.971	97.80%
ω_{12}	0.200	0.196	-2.00%	0.067	0.063	0.942	84.00%
ω_{22}	0.158	0.160	1.27%	0.043	0.040	0.948	97.20%
Klein & Muthéns (2007) QML Ansatz							
ω_{11}	0.158	0.163	3.16%	0.042	0.041	0.971	97.80%
ω_{12}	0.200	0.195	-2.50%	0.067	0.063	0.942	83.50%
ω_{22}	0.158	0.161	1.90%	0.043	0.040	0.948	97.00%
Marsh et al.s (2004) unconstrained Ansatz							
ω_{11}	0.158	0.166	5.06%	0.044	0.042	0.955	97.60%
ω_{12}	0.200	0.196	-2.00%	0.070	0.067	0.957	80.80%
ω_{22}	0.158	0.164	3.80%	0.046	0.043	0.935	96.00%
<i>(C) Korrelation der latenten Prädiktoren: .80</i>							
Klein & Moosbruggers (2000) LMS Ansatz							
ω_{11}	0.158	0.168	6.33%	0.091	0.086	0.940	50.00%
ω_{12}	0.175	0.163	-6.86%	0.166	0.155	0.933	22.80%
ω_{22}	0.158	0.163	3.16%	0.089	0.084	0.949	51.60%
Klein & Muthéns (2007) QML Ansatz							
ω_{11}	0.158	0.168	6.33%	0.091	0.086	0.945	50.02%
ω_{12}	0.175	0.164	-6.29%	0.166	0.156	0.940	23.70%
ω_{22}	0.158	0.163	3.16%	0.089	0.085	0.955	51.20%
Marsh et al.s (2004) unconstrained Ansatz							
ω_{11}	0.158	0.175	10.76%	0.107	0.104	0.972	42.60%
ω_{12}	0.175	0.159	-9.14%	0.194	0.188	0.969	17.20%
ω_{22}	0.158	0.171	8.23%	0.108	0.103	0.954	42.00%

Anmerkung. SE = mittlerer Standardfehler; NA = nicht analysierbar; † = Ergebnisse basieren auf einem verzerrten Standardfehler: ($|Verzerrung| > 10\%$)

Steigt die latente Prädiktorkorrelation auf .80, so weist der unconstrained Ansatz die größten Verzerrungen der Schätzungen der nicht-linearen Effekte auf (Tabelle 5.2, Teil C). LMS und QML überschätzen die beiden quadratischen Effekte mit 6.33% und 3.16%. Der unconstrained Ansatz hingegen überschätzt sie mit einem Bias von 10.76% und 8.23%. LMS und QML unterschätzen den Interaktionseffekt um 6.86% und 6.29%, während der unconstrained Ansatz den Interaktionseffekt um 9.14% unterschätzt. Der unconstrained Ansatz zeigt auch die größten empirischen Standardabweichungen der Parameterschätzungen. Die Teststärke zum Aufdecken der nicht-linearen Effekte fällt stark ab. LMS und QML zeigen die größte Teststärke von ca. 50% - 52% für die quadratischen Effekte, während der unconstrained Ansatz eine Teststärke von ca. 42% - 43% zeigt. Die Teststärke zur Aufdeckung eines Interaktionseffektes ist für LMS (22.80%) und QML (23.70%) sehr niedrig, aber für den unconstrained Ansatz sogar niedriger (17.20%).

5.3.3 Implikationen

Wie man den Ergebnissen der Simulationsstudie entnehmen kann, werden die theoretisch vorhergesagten Unterschiede hinsichtlich der Leistungsfähigkeit der Ansätze bestätigt.

Die verteilungsanalytischen Ansätze LMS und QML weisen bei der Schätzung eines nicht-linearen Modells mit einem Interaktionseffekt und zwei quadratischen Effekten unverzerrte Parameterschätzungen auf. Die Verzerrungen des unconstrained Ansatzes sind bei hoher Multikollinearität gemäß den theoretischen Erwartungen größer (siehe Abschnitt 5.2.2). Wie im theoretischen Teil dieses Kapitels beschrieben wurde, zeigen LMS und QML auch effizientere und im Durchschnitt weniger unterschätzte Standardfehlerschätzungen (Konsequenz 1; Abschnitt 5.2.2). Mit dem unconstrained Ansatz erhält man größere empirische Standardabweichungen von den wahren Parameterwerten und damit unreliaablere Schätzungen, wenn die Multikollinearität ansteigt (Konsequenz 2; Abschnitt 5.2.2). Die Schätzungen sind erwartungsgemäß instabiler, da sie auf einer empirischen Kovarianzmatrix beruhen, die Produkt-Indikatoren enthält. Die empirischen Kovarianzen der Produkt-Indikatoren sind höhere Momente. Daher bedarf es größerer Stichproben, damit zuverlässige Schätzungen der Populationskovarianzen möglich sind. Um mit LMS und QML vergleichbare Schätzungen zu erhalten, müsste also bei der Anwendung des unconstrained Ansatzes die Stichprobe größer sein. Der Teststärkeverlust beim unconstrained Ansatz ist höher (Konsequenz 3; Abschnitt 5.2.2).

Kapitel 6

Diskussion und Ausblick

6.1 Zusammenführung der Kapitel und Antworten auf die Fragestellungen

Da das Multikollinearitätsproblem in der Literatur zu nicht-linearen Strukturgleichungsmodellen nur in seltenen Fällen diskutiert wurde (z.B. Little et al., 2006), fragt die zentrale Fragestellung der vorliegenden Arbeit nach der Relevanz der Multikollinearität bei der Spezifikation und Schätzung von latenten nicht-linearen Effekten in Strukturgleichungsmodellen.

Kapitel 3 beschreibt daher zunächst allgemein (d.h. auch für lineare Strukturgleichungsmodelle) die Bedeutsamkeit und die potentiellen negativen Folgen der Multikollinearität. Danach wird aufgezeigt, warum gerade *in nicht-linearen Strukturgleichungsmodellen das Multikollinearitätsproblem schwerwiegender ist als in linearen Strukturgleichungsmodellen. Sobald nämlich latente lineare Prädiktoren korrelieren, ist die Multikollinearität im nicht-linearen Strukturmodell höher*, da die nicht-linearen Terme stärker untereinander in Zusammenhang stehen. Es kann in der Folge zu unter- oder überschätzten Effekten oder Scheineffekten kommen. Ein Teststärkeverlust tritt immer ein. Potentielle Suppressoreffekte können nicht nur schwer interpretierbar, sondern auch schwer kommunizierbar sein. Mit der Darstellung der Multikollinearitätsproblematik in Schrift (A) und in Teilen in Schrift (B) sollte die Fachöffentlichkeit für dieses Problem sensibilisiert werden.

Kapitel 4 beschreibt erstmalig, dass alle Erweiterungen der Produkt-Indikator-Ansätze zur simultanen Schätzung von Interaktions- und quadratischen Ef-

fekten seit der Erstpublikation von Kenny und Judd (1984) Spezifikationsfehler aufweisen (z.B. auch eine verfahrensvergleichende Publikation von Lee et al., 2004). Spezifikationsfehler machen die ganze Ansatzklasse für Multikollinearität anfällig, so dass es theoretisch zu verzerrten Schätzungen oder Scheineffekten kommen muss, da ein fehlspezifiziertes Modell an die Daten angepasst wird. Es wird gezeigt, wie man die Spezifikationsfehler korrigieren kann und wie die bestehenden Ansätze korrekt zu erweitern sind.

Da die latente Prädiktorkorrelation Bestandteil der Spezifikationsfehler ist, stellte sich die Frage, wie groß die Verzerrungen sind, die sich bei der Anwendung der (unkorrigierten) Produkt-Indikator-Ansätze ergeben, wenn die Multikollinearität steigt.

In Schrift (B) wird zu dieser Frage eine Simulationsstudie durchgeführt. Der in ihr verwendete Zwei-Stufen Ansatz von Ping (1996) produziert bei steigender Multikollinearität stark verzerrte Schätzungen der nicht-linearen Effekte, falschen Vorzeichen der Effekte und Scheineffekte. Verglichen werden die Schätzungen der Zwei-Stufen Ansatzes mit dem korrekt erweiterten Ansatz von Jöreskog und Yang (1996). Der Jöreskog und Yang (1996) Ansatz bringt nach einer korrekten Erweiterung erwartungstreue Schätzungen hervor. Damit lässt sich die Frage eindeutig beantworten, dass die nicht-linearen Effekte deutlich verzerrt geschätzt werden, wenn die Multikollinearität ansteigt und die Spezifikationsfehler nicht korrigiert werden.

Kapitel 5 geht der Frage nach, welche theoretischen und empirischen Unterschiede zwischen den neuen verteilungsanalytischen Ansätzen LMS (Klein & Moosbrugger, 2000) und QML (Klein & Muthén, 2007), einerseits, und den klassischen Produkt-Indikator-Ansätzen, andererseits, zu erwarten sind.

Dazu werden zunächst die theoretischen Eigenschaften der verteilungsanalytischen Verfahren, LMS und QML, untereinander und mit der Klasse der Produkt-Indikator-Ansätze verglichen. Ihre theoretischen Eigenschaften werden beim Vorliegen von Multikollinearität betrachtet. Illustriert werden die Unterschiede anhand einer *Simulationsstudie*, die die Ansätze hinsichtlich ihrer Schätzeigenschaften bei steigender Multikollinearität miteinander vergleicht. In der Simulationsstudie werden die erwarteten Vorteile der verteilungsanalytischen Verfahren bei Vorliegen von Multikollinearität bestätigt. Die Verzerrungen des in der Simulationsstudie untersuchten Produkt-Indikator-Ansatzes sind bei steigender Multikollinearität gemäß den theoretischen Überlegungen

in Kapitel 5 (bzw. Schrift (C)) größer als die Verzerrungen der verteilungsanalytischen Verfahren. Wenn die Multikollinearität steigt, sind die empirischen Standardabweichungen der Parameterschätzungen beim Produkt-Indikator-Ansatz größer und die Schätzungen somit unrelabler. Die Teststärke des gewählten Produkt-Indikator-Ansatzes ist geringer als die Teststärke der verteilungsanalytischen Verfahren.

6.2 „Was sind schon 10 Prozent Verzerrung?“ – Ein Beispiel zur Veranschaulichung der Folgen von verzerrten Schätzungen nicht-linearer Effekte

Die vorliegende Arbeit hat einen psychometrischen Charakter. Dennoch sollen Ergebnisse dieser Arbeit nicht nur psychometrische Forscher, sondern auch anwendungsorientierte Forscher erreichen. Aus diesem Grunde soll in einem *fiktiven Szenario* veranschaulicht werden, wie wichtig es ist, dass Schätzungen nicht-linearer Effekte erwartungstreu und reliabel sind. Dazu soll erneut das Bildungserwartungsbeispiel von Ganzach (1997) aus Kapitel 1 herangezogen werden. Ganzach (1997) analysiert Daten des groß angelegten „National Longitudinal Surveys of Youth (NLSY)“ (Center for Human Resource Research, 1995). Bei der Erklärung der elterlichen Bildungserwartungen an ihre Kinder wird ein nicht-lineares Modell mit einem Interaktionseffekt und zwei quadratischen Effekten geschätzt (s. Gleichung (6.1)). Der Einfachheit halber wird davon ausgegangen, dass das geschätzte Modell dem wahren Modell entspricht¹.

$$EBK = 13.708 + .191 MB + .167 VB - .012 MB \cdot VB + .018 MB^2 + .014 VB^2 + \epsilon \quad (6.1)$$

wobei EBK die elterlichen Bildungserwartungen an das Kind, MB das mütterliche Bildungsniveau, VB das väterliche Bildungsniveau und ϵ das Regressionsresiduum sind. Die linearen Prädiktoren MB und VB sind zentriert. Sie stellen die Abweichung des Elternteils vom durchschnittlichen Bildungsniveau (in Jahren) dar.

Wenn man zwei konkrete Fälle betrachtet, in denen das *mütterliche Bildungsniveau 5 Jahre unter* und *5 Jahre über dem durchschnittlichen Bildungsniveau der Mütter* liegt, dann gestalten sich die Regressionskurven (in Abhängigkeit des väterlichen Bildungsniveaus) wie in Abbildung 6.1 dargestellt.

¹Dieser Fall ist zwar sehr unwahrscheinlich, aber zu Veranschaulichungszwecken akzeptabel.

Modell mit Interaktionseffekt und quadratischen Effekten

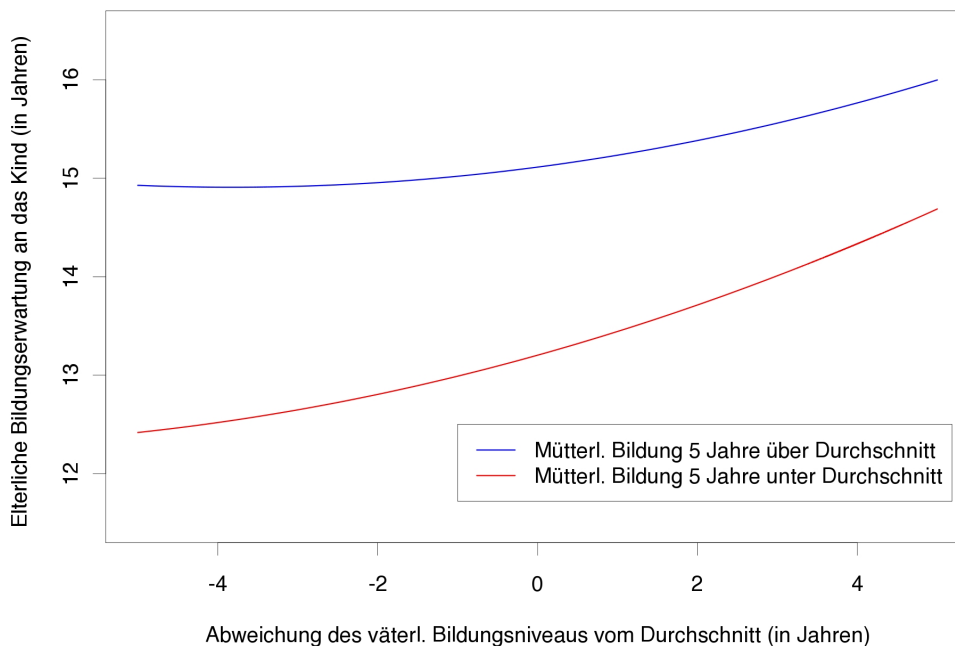


Abbildung 6.1: Bildungserwartungsbeispiel von Ganzach (1997): Modell mit einem Interaktionseffekt und zwei quadratischen Effekten

Wie man Abbildung 6.1 entnehmen kann, werden an das Kind auch dann hohe Bildungserwartungen gestellt, wenn das Bildungsniveau der Mutter niedrig ist, jedoch das Bildungsniveau des Vaters hoch (und vice versa). Mit wachsendem Bildungsniveau des Vaters nähern sich die Bildungserwartungen sehr stark jenen Erwartungen an, die man erhält, wenn auch die Mutter ein hohes Bildungsniveau hat (s. höchster Punkt im blauen Graphen, oben rechts). Es spielt also generell keine Rolle, ob beide Elternteile oder jeweils nur ein Elternteil ein hohes Bildungsniveau haben.

Nun könnte es sein, dass aufgrund der Wahl eines „ungünstigen“ Ansatzes zur Schätzung von nicht-linearen Effekten eine 10- oder mehr-prozentige Unterschätzung der quadratischen Effekte eintritt (beim Zwei-Stufen Ansatz von Ping (1996) waren es sogar bis zu 34 % Unterschätzung der quadratischen Effekte und ca. 50 % Überschätzung des Interaktionseffektes.). In der Folge würden die quadratischen Effekte nicht signifikant von Null verschieden sein und man würde nicht zuletzt der Modellsparsamkeit wegen ein Interaktionsmodell ohne quadratische Effekte beibehalten. Ein solches Modell wurde von Ganzach (1997) an die Daten angepasst (s.

Gleichung (6.2))

$$EBK = 13.866 + .185 MB + .159 VB + .017 MB \cdot VB + \epsilon \quad (6.2)$$

Abgesehen davon, dass nun der Interaktionseffekt ein anderes Vorzeichen aufweist, ergeben sich aus den Schätzungen des Modells andere Vorhersagen. Diese sind in Abbildung 6.2 veranschaulicht.

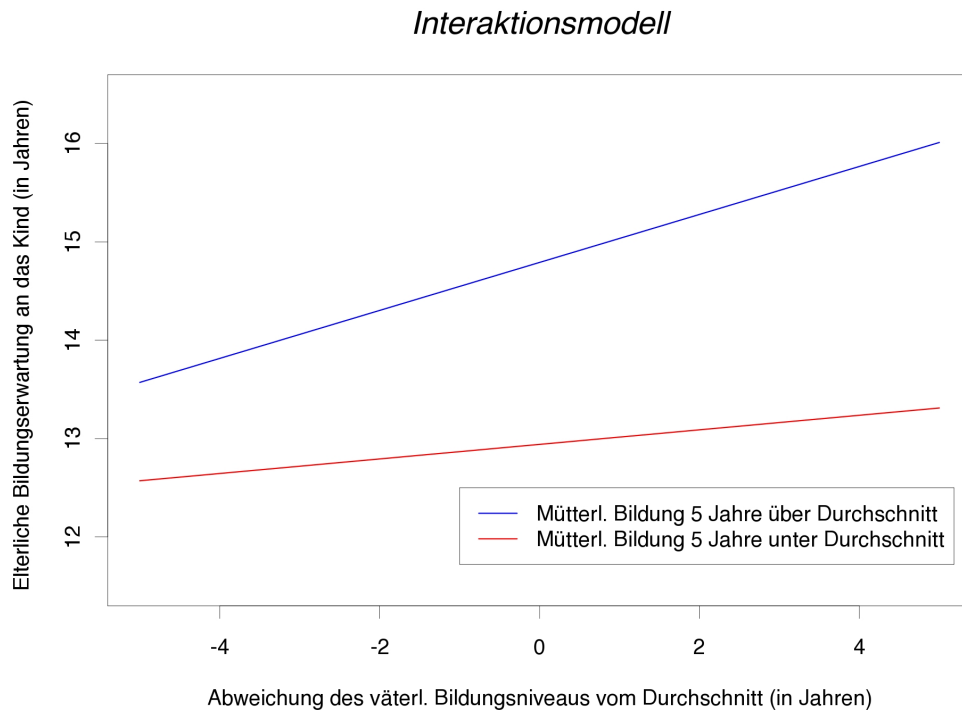


Abbildung 6.2: Bildungserwartungsbeispiel von Ganzach (1997): Modell mit einem Interaktionseffekt

Wie man Abbildung 6.2 entnehmen kann, sind die Bildungserwartungen an das Kind nur dann hoch, wenn beide Elternteile ein hohes Bildungsniveau haben. Es reicht nun nicht mehr aus, dass zumindest ein Elternteil ein hohes Bildungsniveau hat. Die Kompensation der elterlichen Bildungsniveaudifferenz ist nicht mehr möglich. Wie man sieht, können nicht-signifikante oder verzerrte Schätzungen inhaltlich schwerwiegende Konsequenzen haben. Der Leser braucht sich nur vorzustellen, wie groß menschliche Schäden sein könnten, wenn Wechselwirkungen von Medikamenten nicht adäquat geschätzt würden, die auf nicht-linearen Modellen beruhen.

6.3 Bewertung der eigenen Vorgehensweise und Ausblick

6.3.1 Mehrwert und Grenzen der Arbeit

Die vorliegende Arbeit folgt nicht dem typischen Schema einer Robustheitsuntersuchung eines Verfahrens oder einer Verfahrensklasse, da hier nicht ein Ansatz einer Vielzahl von Herausforderungen (z.B. variierender Multikollinearität oder variierenden Verteilungsformen) gegenübergestellt wird, sondern vielmehr ein Problem, nämlich das Multikollinearitätsproblem, auf eine Vielzahl von Ansätzen abgebildet wird. Die Stärken und Schwächen der Ansätze lassen sich so genauer beleuchten, indem die Unterschiede in den Funktionsweisen der Ansätze in den Vordergrund treten und untersucht werden müssen. Nur so gelang es dem Autor der vorliegenden Arbeit, einen 24 Jahre alten Fehler ausfindig zu machen und zu beheben, der schon in der Ursprungspublikation von Kenny und Judd (1984) enthalten ist und bisher in darauf aufbauenden Erweiterungen unentdeckt blieb. Über Jahre hinweg wurde nicht genau genug geprüft, ob die Darstellungen in der Publikation von Kenny und Judd (1984) korrekt sind und in der Folge auch nicht überprüft, wie anfällig die bereits entwickelten Produkt-Indikator-Ansätze bei der Schätzung von nicht-linearen Effekten für Multikollinearität sind. Da die Produkt-Indikator-Ansätze die verbreitetsten Ansätze sind, erscheint es umso wichtiger, dass ihre Anwendung zumindest zu erwartungstreuen Schätzungen führt und die Zahl der Fehlschlüsse durch die korrekte Spezifikation minimiert wird. Nachdem die Korrekturmöglichkeiten aufgezeigt wurden, ist nun auch ein „fairer“ Vergleich in Form von Simulationsstudien und theoretischen Auseinandersetzungen möglich. Beides erfolgte als erster Schritt in Schrift (B) mit dem erweiterten Ansatz von Jöreskog und Yang (1996) und in Schrift (C) unter Verwendung des erweiterten Ansatzes von Marsh et al. (2004).

In Kapitel 5 bzw. Schrift (C) wird erstmals explizit in einer Publikation auf die theoretischen Unterschiede zwischen den verteilungsanalytischen Ansätzen, einerseits, und den Produkt-Indikator-Ansätzen, andererseits, eingegangen. Der theoretische Vergleich erfolgt hierbei erstmals nicht nur innerhalb der verteilungsanalytischen Ansätze, sondern auch klassenübergreifend zwischen den verteilungsanalytischen Ansätzen und den Produkt-Indikator-Ansätzen. Im Kontext der Multikollinearität wurden explizit Vorhersagen zugunsten der verteilungsanalytischen Ansätze getroffen, die in der Simulationsstudie bestätigt werden. Die theoretischen und empirischen Vorteile der verteilungsanalytischen Ansätze wurden bisher in der Literatur

nicht aufgegriffen. Da die Vorteile der verteilungsanalytischen Ansätze in erwartungstreuen, effizienten, reliablen und teststarken Schätzungen der nicht-linearen Effekte bestehen, ist zu erwarten, dass sie in der Zukunft häufiger angewendet werden, wenn ihre Vorteile von der Fachöffentlichkeit stärker diskutiert werden. Mit Schrift (C) erfolgte ein Schritt in diese Richtung.

Die vorliegende Arbeit hat eindeutige Grenzen. Sämtliche Simulationsuntersuchungen wurden mit normalverteilten Variablen durchgeführt. Wie sich die Ansätze unterscheiden, wenn die Variablen nicht-normalverteilt sind, ist bei komplexeren als den hier untersuchten Modellen unklar und bedarf ausgiebiger *systematischer* Untersuchungen. Die Zahl solcher Untersuchungen ist relativ groß, da es vielfältige Verteilungsformen gibt, die alle zu unterschiedlichen Schätzeigenschaften der Ansätze führen können. Nicht-normalverteilte Prädiktoren wurden eher selten und dann zusammen mit einer Vielzahl von anderen Bedingungen untersucht (z.B. bei Marsh et al., 2004), so dass eine fokussierte Untersuchung notwendig wäre.

In der vorliegenden Arbeit wurde zudem darauf verzichtet, weiter auf die Unterschiede zwischen LMS und QML in Form von Simulationsstudien bei nicht-normalverteilten Variablen einzugehen. Es wurde zunächst der Vergleich mit den Produkt-Indikator-Ansätzen unter „regulären“ Bedingungen eine höhere Priorität eingeräumt. Klein und Muthén (2007) haben zwischenzeitlich die theoretischen Überlegungen, die in Kapitel 5 bezüglich der Unterschiede zwischen LMS und QML vorgenommen wurden, anhand erster Simulationsstudien zu Modellen mit mehreren Interaktionseffekten bestätigt.

Als weitere Grenze der Arbeit muss die Modellkomplexität genannt werden. Die vorliegende Arbeit befasst sich mit Modellen, in denen es max. drei nicht-lineare Terme und ein Kriterium gibt. Es wurden keine Modelle untersucht, die z.B. eine Verkettung von nicht-linearen Effekten und damit Effekte höherer Ordnung enthalten. Dazu gehören Kombinationen von Interaktionsmodellen und Mediatormodellen, die in jüngster Zeit verstärkt Betrachtung in der Literatur finden (Edwards & Schurrer Lambert, 2007; Preacher, Rucker & Hayes, 2007). Hier ist ebenfalls unklar, wie die verschiedenen Ansätze im Vergleich abschneiden und wie die theoretischen Eigenschaften zu bewerten sind.

Zuletzt wäre zu nennen, dass keine Untersuchungen zu der Robustheit der Modell-differenztests durchgeführt wurde. Auch hier wurde zunächst der Frage der korrekten Spezifikation der Vorrang gegeben, bevor überhaupt (korrekte) Modelle gegeneinander getestet werden können.

6.3.2 Ausblick

In jedem Fall werden die Eigenschaften von Modelldifferenztests und von Modellgütebeurteilungen ein wichtiges Forschungsgebiet der kommenden Jahre sein, da die übliche Bestimmung der Modellanpassung von linearen Strukturgleichungsmodellen nicht ohne weiteres auf nicht-lineare Strukturgleichungsmodelle übertragbar ist (Klein, zur Veröffentlichung eingereicht). Hierbei werden konstruktionsbedingt verteilungsanalytische Ansätze im Vorteil gegenüber Produkt-Indikator-Ansätzen sein, da erstere die präzisere Modellierung der Likelihood, d.h. ohne Verletzung der Verteilungsannahmen vornehmen.

Neben Simulationsstudien zu nicht-normalverteilten Variablen werden auch weiterhin neuartige Ansätze entwickelt werden. Sie werden allerdings problemorientierter als bisherige Ansätze sein. Betrachtet man in angemessener Distanz die Publikationen zur Schätzung von nicht-linearen Strukturgleichungsmodellen, so findet man einerseits entweder Veröffentlichungen, in denen Verfahren als Weiter- oder Neuentwicklungen publiziert werden (z.B. Jöreskog & Yang, 1996; Klein & Moosbrugger, 2000; Lee et al., 2004; Wall & Amemiya, 2000), andererseits Veröffentlichungen, die sich den Schätzeigenschaften der Ansätze widmen (z.B. Moulder & Algina, 2002; Schermelleh-Engel et al., 1998). Es findet sich keine problemorientierte Entwicklung eines Ansatzes, so dass ein Ansatz unter einer bestimmten Bedingungskonstellation, wie z.B. dem Vorliegen von Multikollinearität und kleinen Stichproben besonders einsetzbar ist.

Die Flexibilität neuerer Ansätze wird beinhalten, dass Prädiktorverteilungen „austauschbar“ werden. Das heißt, es werden Ansätze entwickelt werden, bei denen je nach empirisch-inhaltlicher Fragestellung auch andere Verteilungen als die Normalverteilung der Prädiktoren angenommen werden können (z.B. Poisson-artige Verteilungen), so dass trotzdem zuverlässig nicht-lineare Effekte schätzbar sind.

Multikollinearität wird besonders bei steigender Modellkomplexität *die* Herausforderung sein, an der sich zukünftige Ansätze werden messen müssen. Mit zunehmender Modellkomplexität steigt die Zahl der nicht-linearen Effekte (z.B. bei Kombinationen von Mediation und Moderation) und damit die Multikollinearität mit ihren negativen Folgen. Kaplan (1994, S. 211) betont zu Recht, dass „...exam[ining] the effects of increasing multicollinearity on [SEM] parameter estimates and standard errors ... is an important area for future research.“

Literaturverzeichnis

- Aiken, L. S. & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.
- 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.
- Algina, J. & Moulder, B. C. (2001). A note on estimating the Jöreskog-Yang model for latent variable interaction using LISREL 8.3. *Structural Equation Modeling*, 8, 40-52.
- Arminger, G. & Muthén, B. (1998). A Bayesian approach to nonlinear latent variable models using the Gibbs sampler and the Metropolis-Hastings algorithm. *Psychometrika*, 63, 271-300.
- Aroian, L. A. (1944). The probability function of the product of two normally distributed variables. *The Annals of Mathematical Statistics*, 18, 265-271.
- Belsley, D. A., Kuh, E. & Welsh, R. E. (1980). *Regression Diagnostics - Identifying Influential Data & Sources of Collinearity*. New York, NY: Wiley.
- Bohrnstedt, G. & Goldberger, A. (1969). On the exact covariance of products of random variables. *American Statistical Association Journal*, 64, 1439-1442.
- Bohrnstedt, G. & Marwell, G. (1978). The reliability of the products of two random variables. In K. Schuessler (Hrsg.), *Sociological methodology* (S. 254-273). San Francisco: Jossey-Bass.
- Bollen, K. A. (1995). Structural equation models that are nonlinear in latent variables: A least squares estimator. *Sociological Methodology*, 1995, 223-251.
- Campbell, D. & Fiske, D. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.

- Center for Human Resource Research. (1995). *National Longitudinal Study of Youth: User's guide 1995*. Columbus, OH: Ohio State University.
- Cohen, J., Cohen, P., West, S. G. & Aiken, L. S. (2003). *Applied multiple regression/correlation analyses for the behavioral sciences (3rd ed.)*. Hillsdale, NJ: Erlbaum.
- Cronbach, L. J. & Snow, R. E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions*. New York: Irvington.
- Darlington, R. B. (1978). Reduced-variance regression. *Psychological Bulletin*, 85, 1238-1255.
- 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.
- Edwards, J. & Schurer Lambert, L. (2007). Methods for integrating moderation and mediation: A general analytical framework using moderated path analysis. *Psychological Methods*, 12, 1-22.
- Eid, M. & Diener, E. (2006). *Handbook of multimethod measurement in psychology*. Washington, DC: American Psychological Association.
- Fraser, C. (1980). *Cosan User's Guide*. Toronto: Ontario Institute of Studies in Education.
- Freund, R. & Hoppe, R. H. W. (2007). *Stoer/Bulirsch: Numerische Mathematik 1 [Numerical Mathematics 1]*. Heidelberg: Springer.
- Ganzach, Y. (1997). Misleading interaction and curvilinear terms. *Psychological Methods*, 3, 235-247.
- Hayduk, L. (1987). *Structural equation modeling with LISREL: Essentials and advances*. Baltimore: John Hopkins University.
- Isaacson, E. & Keller, H. (1966). *Analysis of numerical methods*. New York, NY: Wiley.
- Jaccard, J. & Wan, C. (1995). Measurement error in the analysis of interaction effects between continuous predictors using multiple regression: Multiple indicator

- indicator and structural equation approaches. *Psychological Bulletin*, 117, 348-357.
- Jöreskog, K. G. (1998). Interaction and nonlinear modeling: issue and approaches. In R. E. Schumacker & G. A. Marcoulides (Hrsg.), *Interaction and nonlinear effects in structural equation modeling* (S. 239-250). Mahwah, NJ: Erlbaum.
- Jöreskog, K. G. & Sörbom, D. (1989). *LISREL 7: User's Reference Guide*. Lincolnwood, IL: Scientific Software International.
- Jöreskog, K. G. & Sörbom, D. (1996). *LISREL 8: User's Reference Guide*. Lincolnwood, IL: Scientific Software International.
- Jöreskog, K. G., Sörbom, D., Du Toit, S. & Du Toit, M. (1999). *LISREL 8: New Statistical Features*.
- 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: Lawrence Erlbaum Associates.
- Kaplan, D. (1994). Estimator conditioning diagnostics for covariance structure models. *Sociological Methods*, 23, 200-229.
- Karasek, R. (1979). Job demands, job decision latitude, and mental strain: Implication for job redesign. *Administrative Quarterly*, 24, 285-307.
- 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.
- Kenny, D. & Judd, C. M. (1984). Estimating the nonlinear and interactive effects of latent variables. *Psychological Bulletin*, 96, 201-210.
- Klein, A. (zur Veröffentlichung eingereicht). A saturated model and evaluation of model fit for multivariate heteroscedastic models. *Psychometrika*.
- Klein, A. & Moosbrugger, H. (2000). Maximum likelihood estimation of latent interaction effects with the LMS method. *Psychometrika*, 65, 457-474.

- Klein, A., Moosbrugger, H. & Schermelleh-Engel, K. (unveröffentlicht). *Die Robustheit der LMS-Methode bei Verletzung der Verteilungsannahmen [the robustness of the LMS-method when distribution assumptions are violated.]*. (Arbeiten aus dem Institut für Psychologie, Heft 9/2000. Frankfurt am Main: J. W. Goethe-University, Department for Psychology.)
- Klein, A. & Muthén, B. (2007). Quasi maximum likelihood estimation of structural equation models with multiple interaction and quadratic effects. *Multivariate Behavioral Research*, *42*, 647-674.
- Lee, S., Song, X. & Poon, W. (2004). Comparison of approaches in estimating interaction and quadratic effects of latent variables. *Multivariate Behavioral Research*, *39*, 37-67.
- Lee, S., Song, X. & Tang, N. (2007). Bayesian methods for analyzing structural equation models with covariates, interaction, and quadratic latent variables. *Structural Equation Modeling*, *14*, 404-434.
- Little, T., Bovaird, J. & Widaman, K. (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. & Mar, C. M. (1995). Distinguishing between moderator and quadratic effects in multiple regression. *Psychological Bulletin*, *118*, 405-421.
- Marquardt, D. W. (1970). Generalized Inverses, Ridge Regression and Biased Linear Estimation. *Technometrics*, *12*, 591-612.
- Marquardt, D. W. (1980). You should standardize the predictor variables in your regression models. *Journal of the American Statistical Association*, *75*, 87-91.
- Marsh, H., Wen, Z. & Hau, K. (2004). Structural equation models of latent interactions: Evaluation of alternative estimation strategies and indicator construction. *Psychological Methods*, *9*, 275-300.
- Marsh, H., Wen, Z. & Hau, K.-T. (2006). Structural equation modeling: A second course. In G. Hancock & R. Mueller (Hrsg.), (S. 225-265). Greenwich,CT: Information Age Publishing.

- Moosbrugger, H., Schermelleh-Engel, K. & Klein, A. G. (1997). Methodological problems of estimating latent interaction effects. *Methods of Psychological Research Online*, 2, 95-111.
- Moosbrugger, H., Schermelleh-Engel, K., Kelava, A. & Klein, A. G. (2008). Testing multiple nonlinear effects in structural equation modeling: A comparison of alternative estimation approaches. In T. Teo & M. Khine (Hrsg.), *Structural equation modeling in educational research: Concepts and applications*. 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.
- Nussbeck, F., Eid, M., Geiser, C., Courvoisier, D. & Cole, D. (2007). Konvergente und diskriminante Validität über die Zeit: Integration von Multitrait-multimethod-Modellen und der Latent-State-Trait-Theorie. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 361-388). Springer: Heidelberg.
- 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., Rucker, D. D. & Hayes, A. F. (2007). Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate Behavioral Research*, 42, 185-227.
- Satorra, A. & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66, 507-514.
- Schermelleh-Engel, K., Klein, A. & 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: Lawrence Erlbaum Associates.
- Schermelleh-Engel, K. & Schweizer, K. (2007). Multitrait-multimethod-Analysen. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 325-341). Springer: Heidelberg.

- Schumacker, R. & Marcoulides, G. (1998). *Interaction and nonlinear effects in structural equation modeling*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Snyder, M. & Tanke, E. D. (1976). Behavior and attitude: Some people are more consistent than others. *Journal of Personality*, *44*, 501-517.
- 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.
- Yang Jonsson, F. (1997). *Nonlinear structural equation models: Simulation studies of the kenny-judd model*. Unver"offentlichte Dissertation, Uppsala: University.
- Yang Wallentin, F., Schmidt, P., Davidov, E. & Bamberg, S. (2003). Is there an interaction effect between intention and perceived behavioral control? *Methods of Psychological Research - Online*, *8*, 127-157.

Anhang A

Schrift (A)

Dimitruk, P., Schermelleh-Engel, K., **Kelava, A.** & Moosbrugger, H. (2007). Challenges in nonlinear structural equation modeling. *Methodology*, 3, 100-114.

Anhang B

Schrift (B)

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.

Anhang C

Schrift (C)

Kelava, A., Werner, C., Schermelleh-Engel, K., Moosbrugger, H., Zapf, D., Ma, Y. & West, S. G. (zur Veröffentlichung eingereicht). Advanced nonlinear structural equation modeling: Theoretical properties and empirical application of the distribution-analytic LMS and QML estimators. *Structural Equation Modeling*.

Anhang D

Symbolverzeichnis

X	Indikatorvariable des latenten Prädiktors ξ
Y	Indikatorvariable des latenten Kriteriums η
(\mathbf{X}, \mathbf{Y})	Variablenvektor der p X -Indikatoren und q Y -Indikatoren
ξ	Latenter Prädiktor oder Vektor der n latenten ξ -Prädiktoren
η	Latentes Kriterium
λ	Faktorladung
$\mathbf{\Lambda}$	Faktorladungsmatrix
δ	Messfehlervariable eines Indikators X
θ	Parametervektor
θ^δ	Kovarianzmatrix der Messfehlervariablen der X -Indikatoren
ϵ	Messfehlervariable eines Indikators Y
θ^ϵ	Kovarianzmatrix der Messfehlervariablen der Y -Indikatoren
γ	Regressionskoeffizient im latenten Strukturmodell, Einfluss eines latenten Prädiktors ξ auf ein latentes Kriterium η
$\mathbf{\Gamma}$	Matrix der Regressionskoeffizienten der latenten Prädiktoren auf die latenten Kriterien
β	Regressionskoeffizient bei manifesten Modellen oder Einfluss eines latenten Kriteriums η_1 auf ein anderes latentes Kriterium η_2
\mathbf{B}	Matrix der Regressionskoeffizienten der latenten Kriterien untereinander
ϕ	Kovarianzmatrix der latenten Prädiktoren
ζ	Strukturmodellfehler, Vorhersagefehler innerhalb des Strukturmodells
ψ	Matrix der Strukturmodellfehlerkovarianzen und -varianzen

α	Latentes Interzept des Strukturmodells
τ	Latentes Interzept des Messmodells
$E(X)$	Erwartungswert der Variable X
$f(\mathbf{X}, \mathbf{Y})$	Dichtefunktion der Variablenvektors (\mathbf{X}, \mathbf{Y})
$f^*(\mathbf{X}, \mathbf{Y})$	Approximierte Dichtefunktion der Variablenvektors (\mathbf{X}, \mathbf{Y})
$f(\mathbf{X}, \mathbf{Y} Z)$	Z -Bedingte Dichtefunktion des Variablenvektors (\mathbf{X}, \mathbf{Y})

Anhang E

Selbständigkeitserklärung

Hiermit erkläre ich, Augustin Kelava, die vorliegende Dissertationsarbeit mit dem Titel „Multikollinearität in nicht-linearen latenten Strukturgleichungsmodellen“ eigenständig und ausschließlich unter Verwendung der angegebenen Hilfsmittel, angefertigt zu haben. Alle Quellen sind als solche kenntlich gemacht. Die vorliegende Arbeit ist in dieser oder anderer Form zuvor nicht als Dissertationsarbeit zur Begutachtung vorgelegt worden. Das beantragte Promotionsverfahren stellt das erste beantragte Promotionsverfahren dar.

Frankfurt am Main, 02.12.2008

Augustin Kelava

Anhang F

Lebenslauf