

Hildebrandt, Lutz; Görz, Nicole

Working Paper

Strukturgleichungsmodelle mit latenten Variablen zur Analyse heterogener Daten

SFB 373 Discussion Paper, No. 1999,55

Provided in Cooperation with:

Collaborative Research Center 373: Quantification and Simulation of Economic Processes, Humboldt University Berlin

Suggested Citation: Hildebrandt, Lutz; Görz, Nicole (1999) : Strukturgleichungsmodelle mit latenten Variablen zur Analyse heterogener Daten, SFB 373 Discussion Paper, No. 1999,55, Humboldt University of Berlin, Interdisciplinary Research Project 373: Quantification and Simulation of Economic Processes, Berlin, <https://nbn-resolving.de/urn:nbn:de:kobv:11-10046458>

This Version is available at:

<https://hdl.handle.net/10419/61716>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Strukturgleichungsmodelle mit latenten Variablen zur Analyse heterogener Daten

Nicole Görz
Lutz Hildebrandt

Humboldt-Universität zu Berlin

Institut für Marketing II

Spandauer Str. 1

10178 Berlin

Tel.: +49/(0)30 2093-5691

Fax: +49/(0)30 2093-5675

E-Mail: hildebr@wiwi.hu-berlin.de

Zusammenfassung

Die Anwendung der Strukturgleichungsmethodologie im Marketing geht im allgemeinen von homogenen Daten bei der Analyse aus, sofern nicht a priori Informationen über Unterschiede in den Daten vorliegen. Latente Strukturen in den Daten bleiben jedoch meist unbeachtet.

Dieser Beitrag befaßt sich mit Möglichkeiten der Aufdeckung von latenten Strukturen in Kovarianzstrukturanalysen durch Anwendung von Finite Mixture Models. Es wird mit simulierten Daten aufgezeigt, wie mit dem Programmsystem MECOSA 3.01 (Arminger, Wittenberg & Schepers, 1996) eine bessere Anpassung der Daten durch einen Finite Mixture Ansatz erzeugt wird als mit a priori Clusterung und der stufenweisen Analyse mit vorgeschalteter Clusterung und Mehrgruppenanalyse. Für die Standardanalysen wird das Programm AMOS 3.6 (Arbuckle, 1997) eingesetzt.

1 Einleitung

In der Marktforschung erhobene Daten sind im allgemeinen durch heterogene Strukturen gekennzeichnet. So wird in Studien zum Kaufverhalten häufig ein differenziertes Kaufmuster festgestellt, oder Studien zum Entscheidungsverhalten decken unterschiedliche Präferenzmuster auf, so daß zwangsläufig für das Marketing eine Aufteilung des Marktes in einzelne Marktsegmente notwendig wird, um diese mit unterschiedlichen Mitteln bearbeiten zu können. Aus Unternehmensperspektive kann dies z. B. zum verstärkten Einsatz von Kommunikationsinstrumenten führen, wenn Produkte impositioniert werden sollen oder Preise gesenkt werden, um neue Käuferschaften zu gewinnen. Erfahrungsgemäß reagieren unterschiedliche Konsumenten unterschiedlich auf den Einsatz des Marketinginstrumentariums.

Diese Tatsache ist bislang in der Strukturgleichungsmethodologie nur unzureichend berücksichtigt worden. Die kausalanalytischen Studien im Marketing (vgl. z. B. Homburg & Baumgartner, 1995 oder Hildebrandt & Homburg, 1998) gehen im allgemeinen von homogenen Daten aus, oder es erfolgt eine a priori Clusterung vor der Analyse eines Strukturgleichungsmodells. Ist das Sample nicht homogen, so kann, wie u. a. Jedidi, Jagpal & DeSarbo (1994) nachweisen, die undifferenzierte Analyse heterogener Daten zu falschen inhaltlichen Schlußfolgerungen führen. So ist es beispielsweise vorstellbar, daß eine starke Beziehung zwischen zwei Variablen bei der Schätzung des Modells nicht nachgewiesen wird, da die entsprechenden Parameter in den Gruppen entgegengesetzte Vorzeichen aufweisen.

Weiterhin tritt bei empirischen Studien häufig das Problem auf, daß die Variablen, in denen sich die Gruppen unterscheiden, d. h. die Quellen der Heterogenität der Daten, nicht bekannt sind und/oder keine Meßmodelle für sie verfügbar sind. Während für die Analyse von Strukturgleichungsmodellen mit latenten Variablen (im folgenden als SEM-Modelle bezeichnet) bei a-priori klar abgrenzbaren Gruppen eine ausgereifte Methode zur Verfügung steht, gestaltet sich die Analyse bei latenten Strukturen, d. h. a-priori nicht klar identifizierbarer Gruppenzugehörigkeit, problematisch.

Die vorliegende Studie ist wie folgt aufgebaut. Nach einer kurzen Darstellung der grundlegenden Spezifikation von SEM-Modellen wird zunächst auf mögliche Ausprägungsformen von Heterogenität eingegangen. Im Anschluß daran werden dann Mehrgruppen SEM-Modelle vorgestellt, in denen die Gruppenzugehörigkeit der ein-

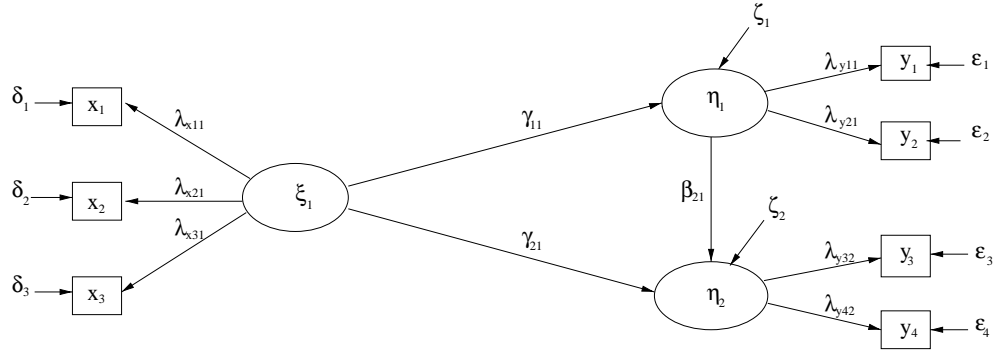
zelenen Beobachtungen bekannt ist. Daran schließt sich die Darstellung der Analysemöglichkeiten von SEM-Modellen bei unbekannter Gruppenzugehörigkeit an. Hierbei wird auf zwei verschiedene Analyseansätze eingegangen: einen traditionellen, bei dem die Clusteranalyse und Mehrgruppen SEM-Modelle kombiniert werden, und eine innovative Modellklasse, die mit Mischverteilungen bei der Schätzung arbeitet. Nachfolgend werden die Ergebnisse einer Simulationsstudie zum Vergleich der verschiedenen Methoden beschrieben und diskutiert. Den Abschluß bildet ein Ausblick auf weitergehende Fragestellungen, die bei der Analyse latenter Strukturen mit SEM-Modellen entstehen.

2 Strukturgleichungsmodelle mit latenten Variablen (SEM-Modelle)

Die Strukturgleichungsmethodik vereint mehrere multivariate Analysetechniken, deren Kopplung die Analyse von Beziehungsstrukturen zwischen nicht beobachtbaren (latenten) Variablen und beobachtbaren Variablen (Indikatoren) ermöglicht. Sie kann als Kombination von regressions- und faktoranalytischen Techniken angesehen werden (Ullmann, 1996) und ist den konfirmatorischen Analyseverfahren zuzuordnen. Bei ihrer Anwendung stellt die Modellbildung, die theoretische Herleitung von Hypothesen über die Beziehungsstrukturen in der betrachteten Variablenmenge, den ersten Schritt der Analyse dar.

Die vermuteten Beziehungen werden mit Hilfe eines Pfaddiagramms graphisch abgebildet. Diese Visualisierung dient einerseits der leichteren Erfassbarkeit der Problemstellung bei komplexeren Modellstrukturen. Andererseits kann aus der graphischen Darstellung auch die formale Gleichungsstruktur des Modells abgeleitet werden, die bei der Modellschätzung von Bedeutung ist. Abbildung 1 zeigt das Pfaddiagramm und das daraus abgeleitete Gleichungssystem für ein einfaches SEM-Modell unter Verwendung der weit verbreiteten LISREL-Notation.

Kennzeichnend für SEM-Modelle ist die explizite Unterscheidung zwischen der Meßtheorie (Meßmodell) und der Substanztheorie (Strukturmodell), wie sie auch in Abbildung 1 erkennbar ist (vgl. Homburg & Hildebrandt, 1998). Das Meßmodell beschreibt die Beziehungen zwischen den unbeobachtbaren theoretischen Konstrukten



Meßmodell

$$\begin{aligned} x_1 &= \lambda_{x11} \xi_1 + \delta_1 \\ x_2 &= \lambda_{x21} \xi_1 + \delta_2 \\ x_3 &= \lambda_{x31} \xi_1 + \delta_3 \end{aligned}$$

Strukturmodell

$$\begin{aligned} \eta_1 &= \gamma_{11} \xi_1 + \zeta_1 \\ \eta_2 &= \gamma_{21} \xi_1 + \beta_{21} \eta_1 + \zeta_2 \end{aligned}$$

Meßmodell

$$\begin{aligned} y_1 &= \lambda_{y11} \eta_1 + \epsilon_1 \\ y_2 &= \lambda_{y21} \eta_1 + \epsilon_2 \\ y_3 &= \lambda_{y32} \eta_2 + \epsilon_3 \\ y_4 &= \lambda_{y42} \eta_2 + \epsilon_4 \end{aligned}$$

Abbildung 1: Aufbau eines Strukturgleichungsmodells mit latenten Variablen

(latenten Variablen) und ihren Indikatoren (beobachtbaren Variablen), die als Faktorstruktur modelliert werden. Das Strukturmodell hingegen bildet die vermuteten Kausalbeziehungen auf der Ebene der theoretischen Konstrukte ab. Ein vollständiges SEM-Modell kann allgemeingültig mit den folgenden Gleichungen beschrieben werden:

Strukturmodell:

$$\eta = B\eta + \Gamma\xi + \zeta \quad (1)$$

Meßmodell:

$$y = \Lambda_y \eta + \epsilon \quad (2)$$

$$x = \Lambda_x \xi + \delta \quad (3)$$

wobei η und ξ respektive die Vektoren der endogenen und exogenen latenten Variablen beschreiben. y und x sind Vektoren der beobachtbaren Indikatoren, und ζ, ϵ und δ repräsentieren Fehlervariablen. Die Parametermatrizen B und Γ bilden die Beziehungen im Strukturmodell, und Λ_y und Λ_x die Beziehungen im Meßmodell ab. Es werden folgende Annahmen getroffen:

- ζ ist unkorreliert mit ξ ,

- ϵ ist unkorreliert mit η ,
- δ ist unkorreliert mit ξ ,
- ζ, ϵ und δ sind untereinander unkorreliert,
- $I - B$ ist nicht singulär.

Insgesamt sind zur Spezifikation eines vollständigen SEM-Modells acht Parametermatrizen notwendig. B, Γ, Λ_y und Λ_x enthalten Regressionsgewichte, wohingegen $\Phi, \Psi, \Theta_\epsilon$ und Θ_δ die Kovarianzen der exogenen latenten Variablen, die Kovarianzen der Residuen der endogenen latenten Variablen und der Residuen der beobachtbaren Variablen beschreiben.

Die Datengrundlage für die Analyse von SEM-Modellen bilden die Varianzen und Kovarianzen der Indikatorvariablen, weshalb die Strukturgleichungsmethodik häufig auch als Kovarianzstrukturanalyse bezeichnet wird. Zur Schätzung der Parameter wird aus den Modellgleichungen eine Kovarianzstrukturgleichung hergeleitet, die die Populationskovarianzmatrix der Indikatoren Σ zum Parametervektor des Modells (ϑ) in Beziehung setzt:

$$\Sigma = \Sigma(\vartheta) \tag{4}$$

SEM-Modelle bilden so eine Klasse von Modellen, die viele im Marketing bekannte Modelltypen als Spezialfälle umfaßt. Zu nennen sind hier beispielsweise konfirmatorische Faktormodelle, Pfadmodelle ohne latente Variablen, multiple Regressionsmodelle und Panelmodelle. Heterogenität wird in den meisten Anwendungen jedoch vernachlässigt.

3 Heterogenität

Eine Grundgesamtheit wird als heterogen bezeichnet, wenn sie sich aus mehreren Untergruppen zusammensetzt, die sich in der Ausprägung bestimmter Merkmale voneinander unterscheiden. Im Marketingkontext können solche Untergruppen z. B. Marktsegmente repräsentieren. Diese Marktsegmente können sich auch in Bezug auf ihre Modellstruktur unterscheiden, d. h. die Beziehungen zwischen den Modellvariablen können unterschiedlich stark ausgeprägt sein.

Im allgemeinen können zwei Typen von Heterogenität unterschieden werden: beobachtbare und unbeobachtbare Heterogenität. Beobachtbare Heterogenität liegt immer dann vor, wenn die Quellen der Gruppenunterschiede aufgrund theoretischer Vorüberlegungen bekannt sind und über Indikatoren gemessen wurden. Eine a priori Zuordnung der Beobachtungspunkte zu den einzelnen Gruppen ist dann möglich. Klassifizierungsprobleme treten allenfalls auf, wenn eine Vielzahl von Indikatoren mit unterschiedlicher Ausprägung vorliegen. Für den Fall der beobachtbaren Heterogenität steht dem Forscher mit den Mehrgruppen SEM-Modellen eine etablierte Analysetechnik zur Verfügung (vgl. z. B. Jöreskog & Sörbom, 1993).

Unbeobachtbare Heterogenität ist dagegen dadurch gekennzeichnet, daß a-priori keine Informationen über die Gruppenstruktur bekannt sind, und/oder die Indikatoren der Gruppenzugehörigkeit nicht gemessen wurden. Es liegen dann latente Strukturen vor, d. h. eine a priori Klassifikationen der Beobachtungspunkte aufgrund theoretischer Erkenntnisse kann nicht vorgenommen werden. Zur Analyse von SEM-Modellen bei unbeobachtbarer Heterogenität existieren zwei Modellansätze. In der traditionellen Analyse wird eine sequentielle Vorgehensweise gewählt, bei der Clusteranalyse und Mehrgruppen SEM-Modelle nacheinander angewendet werden.

Aufgrund der Probleme des traditionellen Ansatzes, die aus der Vernachlässigung der Beziehungen zwischen den Variablen bei der Gruppenbildung resultieren, ist in den letzten Jahren ein modellbasierter Ansatz zur Analyse latenter Strukturen in heterogenen Daten entwickelt worden, die sogenannten Finiten Mischungen von Strukturgleichungsmodellen. Hier erfolgt die Gruppenbildung und die Schätzung der Modellstruktur simultan (vgl. z. B. Arminger, Stein & Wittenberg, 1998, Jedidi, Jagpal & DeSarbo, 1997).

Im folgenden Abschnitt werden zunächst die Mehrgruppen SEM-Modelle zur Analyse bei beobachtbarer Heterogenität vorgestellt. Das Folgekapitel widmet sich den Modellansätzen bei unbeobachtbarer Heterogenität.

4 Mehrgruppen SEM-Modelle bei beobachtbarer Heterogenität

Mehrgruppen SEM-Modelle stellen eine Erweiterung der allgemeinen SEM-Modelle dar, wie sie in Kapitel 2 vorgestellt wurden. Eine Voraussetzung für ihre Anwendbarkeit besteht in der klaren Abgrenzbarkeit der einzelnen Gruppen voneinander.

Da sich die Gruppen häufig nicht nur in den Beziehungsstrukturen voneinander unterscheiden sondern auch und vor allem in der unterschiedlich starken Ausprägung der latenten Variablen, wird die Annahme des allgemeinen SEM-Modells, daß alle Variablen als Abweichungen von ihren Mittelwerten gemessen werden, gelockert. Neben der Kovarianzstruktur wird somit in den Mehrgruppenmodellen i. d. R. auch die Mittelwertstruktur der betrachteten Variablen analysiert. Ein Mehrgruppen SEM-Modell wird unter Verwendung der LISREL-Notation wie folgt spezifiziert:

Strukturmodell:

$$\eta^{(g)} = \alpha^{(g)} + B^{(g)}\eta^{(g)} + \Gamma^{(g)}\xi^{(g)} + \zeta^{(g)} \quad (5)$$

Meßmodell:

$$y^{(g)} = \tau_y^{(g)} + \Lambda_y^{(g)}\eta^{(g)} + \epsilon^{(g)} \quad (6)$$

$$x^{(g)} = \tau_x^{(g)} + \Lambda_x^{(g)}\xi^{(g)} + \delta^{(g)} \quad (7)$$

Zusätzlich zu den acht Parametermatrizen des allgemeinen SEM-Modells beinhalten Mehrgruppen SEM-Modelle Parameter für die Absolutglieder in den Gleichungen $(\alpha^{(g)}, \tau_y^{(g)}, \tau_x^{(g)})$ und für die Mittelwerte der exogenen latenten Variablen $(E(\xi^{(g)}) = \kappa)$. Die Mittelwerte der endogenen latenten Variablen sowie der Indikatoren lassen sich über die Modellgleichungen (5-7) berechnen:

$$E(\eta^{(g)}) = (I - B^{(g)})^{-1}(\alpha^{(g)} + \Gamma^{(g)}\kappa^{(g)}) \quad (8)$$

$$E(y^{(g)}) = \mu_y^{(g)} = \tau_y^{(g)} + \Lambda_y^{(g)}(I - B^{(g)})^{-1}(\alpha^{(g)} + \Gamma^{(g)}\kappa^{(g)}) \quad (9)$$

$$E(x^{(g)}) = \mu_x^{(g)} = \tau_x^{(g)} + \Lambda_x^{(g)}\kappa^{(g)} \quad (10)$$

Aus Gründen der Identifizierbarkeit können keine absoluten Mittelwerte für die latenten Variablen geschätzt werden, sondern nur Mittelwertdifferenzen zwischen den Gruppen. Dazu werden bei der Modellspezifikation die Mittelwerte der exogenen latenten Variablen in einer der Gruppen auf 0 restringiert. Die gleiche Vorgehensweise

kann für die Absolutglieder im Strukturmodell angewendet werden (vgl. Sörbom, 1981).

Die Mehrgruppenanalyse ist einer separaten Modellschätzung für die einzelnen Gruppen immer dann überlegen, wenn Parameter über die Gruppen hinweg restringiert werden. Es können sowohl Gleichsetzungsrestriktionen als auch Bereichs- und sonstige lineare Restriktionen spezifiziert werden (vgl. Jöreskog & Sörbom, 1993). In diesen Fällen liefert nur die Mehrgruppenanalyse effiziente Parameterschätzwerte. Ein weiterer Vorteil besteht in der Anwendbarkeit des χ^2 -Differenzentests, mit dem die Signifikanz jeglicher vermuteter Gruppenunterschiede geprüft werden kann.

5 SEM-Modelle bei unbeobachtbarer Heterogenität

Das Hauptproblem bei unbeobachtbarer Heterogenität besteht in der Klassifikation der einzelnen Beobachtungspunkte. Nachfolgend werden zwei Analyseverfahren von SEM-Modellen bei unbeobachtbarer Heterogenität vorgestellt, die sich in der Art und Weise der Zuordnung der einzelnen Beobachtungen zu den Gruppen unterscheiden.

5.1 Der traditionelle Analyseansatz

Die in dieser Arbeit als “traditioneller Analyseansatz“ bezeichnete Methode ist durch eine sequentielle Vorgehensweise gekennzeichnet. Um die heterogenen Daten in weitgehend homogene Gruppen aufzuteilen, wird zunächst eine Clusteranalyse eingesetzt, z. B. K-Means Clustering. Für die so identifizierten Teilmengen an Beobachtungen wird in einem zweiten Schritt jeweils separat eine Kovarianzmatrix erzeugt. Diese Kovarianzmatrizen dienen dann als Input für eine Mehrgruppenanalyse, wie sie oben beschrieben wurde.

Der Hauptkritikpunkt an dieser stufenweisen Vorgehensweise ist die mangelnde Berücksichtigung der postulierten Datenstruktur, d. h. der angenommenen Beziehungen zwischen den Variablen, bei der Gruppenbildung. Traditionelle Clusterverfahren gehen von einer Unabhängigkeit der betrachteten Variablen aus. Sie unter-

stellen somit eine diagonale Kovarianzmatrix. Vorliegende Informationen für die Aufteilung der Daten bleiben unbeachtet. Vermutet der Forscher eine hohe Korrelation zwischen den beobachteten Variablen, so kann er diese zwar durch die Vorschaltung einer Technik zur Datenreduktion, z. B. einer Faktoranalyse oder einer Hauptkomponentenanalyse, berücksichtigen und die Gruppenbildung auf Basis der reduzierten Daten vornehmen. Dieses Vorgehen ist jedoch aufgrund seiner logischen Inkonsistenz kritisch zu beurteilen. Die Datenreduktionstechniken unterstellen eine homogene Grundgesamtheit, während die Clusteranalyse von Heterogenität ausgeht (Jedidi, Jagpal & DeSarbo, 1997).

Die inhaltlichen Kritikpunkte am traditionellen Ansatz schlagen sich auch in statistischen Problemen nieder, beispielsweise in verzerrten Schätzwerten für die Gruppenanteile an der Gesamtheit der Daten und die Modellparameter für einzelne Gruppen.

Zusammenfassend kann festgestellt werden, daß der traditionelle Analyseansatz eher unbefriedigende Ergebnisse liefert, da das den Daten zugrundeliegende Modell bei der Gruppenbildung nicht berücksichtigt wird. Aufgrund dessen erscheint es notwendig, einen innovativen, modellbasierten Analyseansatz einzusetzen, wie ihn z. B. Modelle auf Basis von Mischverteilungen repräsentieren.

5.2 Analyseansätze auf Basis von Mischverteilungen

Bei der Arbeit mit Mischverteilungen wird unterstellt, daß sich die Daten aus zwei oder mehreren Grundgesamtheiten zusammensetzen, die in verschiedenen Anteilen miteinander vermischt wurden (McLachlan & Basford, 1988). Für die Dichtefunktionen innerhalb der Gruppen wird eine bestimmte parametrische Form unterstellt, häufig wird mit der Normalverteilungsannahme gearbeitet. Weiterhin wird davon ausgegangen, daß die Daten eine Zufallsstichprobe repräsentieren, d. h. identisch und unabhängig voneinander verteilt sind.

Die gemischte Dichtefunktion wird als gewichtete Summe der Dichtefunktionen in den einzelnen Gruppen berechnet:

$$f(y|v) = \sum_{g=1}^G \omega_g f_g(y|\theta_g) \quad (11)$$

wobei sich der Parametervektor der gemischten Dichtefunktion v aus den Gruppenanteilen an der Gesamtheit der Daten ω_g und den Parametern der Dichtefunktionen

in den Gruppen θ_g zusammensetzt:

$$v = (\omega_g, \theta_g). \quad (12)$$

Wird in den Gruppen eine Normalverteilung angenommen, so enthält θ_g den Erwartungswert und die Varianz.

Folgende Restriktionen werden im Zusammenhang mit ω_g getroffen:

$$\sum_{g=1}^G \omega_g = 1 \quad (13)$$

$$\omega_g \geq 0 \quad (g = 1, \dots, G). \quad (14)$$

Auf dem Gebiet der psychometrischen Forschung wurden im letzten Jahrzehnt verschiedene Modelle entwickelt, die auf Mischverteilungen beruhen (vgl. Tabelle 1).

DeSarbo/Cron	1988	Clusterwise linear regression
Jones/McLachlan	1992	Multivariate clusterwise regression
Jedidi/Ramaswamy/DeSarbo/Wedel	1996	Finite mixtures of multivariate regression and simultaneous equation systems
Yung	1997	Finite mixtures in confirmatory factor analysis
Jedidi/Jagpal/DeSarbo	1997	Finite mixture structural equation models
Arminger/Stein/Wittenberg	1998	Mixtures of conditional mean and covariance structure models

Tabelle 1: Modelle auf Basis von Mischverteilungen im Marketing

Die vorliegende Arbeit konzentriert sich im weiteren auf die Erläuterung des Modells von Arminger, Stein & Wittenberg (1998), da dieses die allgemeinste Form der Modellspezifikation darstellt und die anderen Modelle als Spezialfälle beinhaltet.

Finite Mischungen bedingter Mittelwert- und Kovarianzstrukturmodelle

Die grundlegende Theorie über Mischverteilungen, wie sie in den vorangegangenen Abschnitten beschrieben wurde, wird bei den finiten Mischungen bedingter Mittelwert- und Kovarianzstrukturmodelle in zwei Richtungen erweitert:

1. Die Menge der beobachtbaren Variablen wird aufgeteilt in abhängige Variablen y und unabhängige Variablen x . Die Erwartungswerte der abhängigen Variablen innerhalb jeder Gruppe seien von den möglicherweise nicht normalverteilten Regressorvariablen linear abhängig.
2. Die gruppenspezifischen Erwartungswerte und Kovarianzmatrizen können parametrisiert werden, z. B. in Form eines SEM-Modells.

Finite Mischungen bedingter Mittelwert- und Kovarianzstrukturmodelle werden dann wie folgt spezifiziert (vgl. auch Arminger & Stein, 1997, S. 4). Das Modell beruht auf beobachtbaren kontinuierlichen abhängigen Zufallsvariablen y_i und einem Vektor von Regressoren x_i , die kontinuierliche oder auch Dummy Variable sein können,

$$h(y_i, x_i) = f(y_i|x_i)g(x_i) \quad (15)$$

wobei $h(y_i, x_i)$ die gemeinsame Dichtefunktion der Daten repräsentiert, die sich als Produkt der bedingten Dichtefunktion $f(y_i|x_i)$ und der Dichtefunktion der Regressorvariablen $g(x_i)$ ergibt. Über $g(x_i)$ werden keine Annahmen getroffen. Die bedingte Dichtefunktion $f(y_i|x_i)$ wird als Mischverteilung spezifiziert:

$$f(y_i|x_i) = \sum_{g=1}^G \omega_g f_g(y_i|x_i) \quad (16)$$

wobei für die gruppenspezifischen bedingten Dichtefunktionen Normalverteilung angenommen wird:

$$f_g(y_i|x_i) = \Phi(y_i; \mu_{ig}, \Sigma_g). \quad (17)$$

Für die bedingte Mittelwertstruktur in den Gruppen wird ein Regressionsmodell spezifiziert,

$$E(y_i|x_i, g) = \mu_{ig} = \gamma_g + \Pi_g x_i \quad (18)$$

in dem γ_g die Regressionskonstante und Π_g die Matrix der Regressionskoeffizienten ist. Die Kovarianzstruktur wird durch

$$V(y_i|x_i, g) = \Sigma_g \quad (19)$$

repräsentiert. Aus den Gleichungen kann dann die reduzierte Form des Modells der Finiten Mischungen bedingter Mittelwert- und Kovarianzstrukturmodelle abgeleitet

werden:

$$y_{ig} = \gamma_g + \Pi_g x_i + \epsilon_{ig}^* \quad (20)$$

$$V(\epsilon_{ig}^*) = \Sigma_g \quad (21)$$

mit ϵ_{ig}^* als Residuen. Die Parametermatrizen der reduzierten Form $(\gamma_g, \Pi_g, \Sigma_g)$ können als beliebige Funktionen eines strukturellen Parametervektors (v) parametrisiert werden. So besteht beispielsweise für die Struktur von Σ_g die Möglichkeit der Spezifikation eines SEM-Modells.

Das allgemeine Modell finiter Mischungen bedingter Mittelwert- und Kovarianzstrukturmodelle wird zum besseren Verständnis in Abbildung 2 noch einmal graphisch dargestellt. Wie dort erkennbar ist, werden die Quellen der Heterogenität im

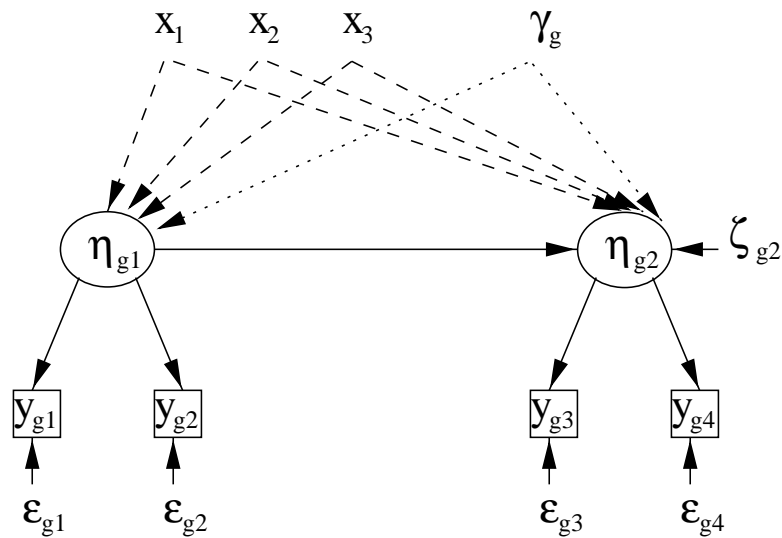


Abbildung 2: Finite Mischungen bedingter Mittelwert- und Kovarianzstrukturmodelle (Modellstruktur)

Modell auf zwei verschiedene Variablentypen zurückgeführt. Die mit x bezeichneten Regressorvariablen repräsentieren die bekannten Merkmale, die zu Gruppenunterschieden führen. Ihr Einfluß wird durch gestrichelte Pfeile dargestellt. Die unbekannt, latenten Ursachen der Heterogenität werden in γ_g modelliert, und durch die gepunkteten Pfeile abgebildet.

Geht man davon aus, daß in den meisten Anwendungsfällen überhaupt keine Informationen über die Quellen der Heterogenität vorhanden sind bzw. keine Meßmo-

delle für die Indikatoren der Gruppenunterschiede vorliegen, d. h. keine x Variablen vorhanden sind, reduziert sich das allgemeine Modell finiter Mischungen bedingter Mittelwert- und Kovarianzstrukturmodelle auf eine unbedingte Mischverteilung:

$$f(y_i) = \sum_{g=1}^G \omega_g f_g(y_i) \quad (22)$$

$$f_g(y_i) = \Phi(y_i; \mu_{ig}, \Sigma_g) \quad (23)$$

Die reduzierte Form des Modells vereinfacht sich zu

$$y_{ig} = \gamma_g + \epsilon_{ig}^* \quad (24)$$

da

$$E(y_i|g) = \mu_{ig} = \gamma_g \quad (25)$$

Bei einer Strukturierung von $V(\epsilon_{ig}^*) = \Sigma_g$ als SEM-Modell enthält der strukturelle Parametervektor die Matrizen des allgemeinen SEM-Modells (vgl. Kapitel 2). Dieser Spezialfall des Modellansatzes von Arminger, Stein & Wittenberg (1998) entspricht den finiten Mischungen von SEM-Modellen, wie sie bei Jedidi, Jagpal & DeSarbo (1997) beschrieben werden.

Bevor im folgenden die Schätzung Finiter Mischungen von SEM-Modellen vorgestellt wird, muß auf die Identifikationsproblematik eingegangen werden. Unter der Voraussetzung, daß die Struktur des spezifizierten SEM-Modells in den einzelnen Gruppen identifiziert ist, bzw. die Identifikation durch gruppenübergreifende Parameterrestriktionen gesichert wird, sind Finite Mischungen von SEM-Modellen dann identifiziert, wenn die Daten für die unbekanntenen Gruppen einer multivariaten Normalverteilung folgen (vgl. den Beweis in Appendix 1 in Jedidi, Jagpal & DeSarbo, 1997). Ausführliche Erläuterungen zu diesem Thema finden sich auch bei Titterton, Smith & Makov (1985) und McLachlan & Basford (1988).

Die Schätzung von Mischverteilungsmodellen beruht auf der Anwendung des EM-Algorithmus (Dempster, Laird & Rubin, 1977). Der EM-Algorithmus ist ein iteratives Verfahren zur Berechnung der ML-Schätzer für die Parameter einer Mischverteilung. In der praktischen Anwendung wird die Struktur des Algorithmus an die jeweilige Problemstellung angepaßt. So verwenden die in Tabelle 1 aufgeführten Modellansätze unterschiedlich spezifizierte EM-Algorithmen.

Arminger, Stein & Wittenberg (1998) stellen drei alternative Strategien zur Schätzung der Parameter finiter Mischungen bedingter Mittelwert- und Kovarianzstrukturmodelle vor: eine zweistufige Schätzprozedur, den direkten und den gradienten EM-Algorithmus.

Bei der zweistufigen Schätzung werden in der ersten Stufe zunächst die Parameter der reduzierten Form des Modells $(\gamma_g, \pi_g, \Sigma_g)$ und ihre asymptotische Kovarianzmatrix mittels des EM-Algorithmus von Jones & McLachlan (1992) geschätzt. Darauf aufbauend erfolgt in der nächsten Stufe die Schätzung des strukturellen Parametervektors mit der Kleinste-Quadrate-Methode.

Alternativ kann der strukturelle Parametervektor des Modells auch direkt mit dem EM-Algorithmus geschätzt werden. Diese Variante hat den Vorteil, daß die Modellrestriktionen bereits bei der Gruppenbildung beachtet werden. Im Fall der zweistufigen Schätzung werden lediglich die Kovarianzstruktur, jedoch nicht die Modellrestriktionen, in die Gruppenbildung einbezogen. Nachteilig wirkt sich bei der direkten Schätzung aber aus, daß die Modellspezifikation nicht durch die Betrachtung der Varianz/Kovarianzmatrix unterstützt werden kann und damit ausschließlich auf Basis theoretischer Überlegungen vorgenommen werden muß. Weiterhin benötigt die direkte Schätzung weit längere Rechenzeiten als das zweistufige Verfahren, was bei komplexen Modellstrukturen zu einem entscheidenden Auswahlkriterium werden kann. Die Rechenzeiten der direkten Schätzung können jedoch durch die Anwendung des gradienten EM-Algorithmus (Becker, Yang & Lange, 1997) verkürzt werden.

Für die praktische Anwendung empfehlen Arminger, Stein & Wittenberg (1998) zuerst die Stufe 1 der zweistufigen Schätzung durchzuführen, d. h. die Parameter der reduzierten Form des Modells zu schätzen, um die Anzahl der Gruppen und mögliche Modellstrukturen zu identifizieren. Im Anschluß daran sollte der gradienten EM-Algorithmus genutzt werden, um die endgültigen Schätzwerte zu generieren.

Abschließend kann festgehalten werden, daß finite Mischungen von Strukturgleichungsmodellen verglichen mit dem traditionellen Analyseansatz die bessere Alternative zur Analyse heterogener Daten mit SEM-Modellen darstellen, da sie durch die simultane Gruppenbildung und Kovarianzstrukturschätzung die Beziehungen zwischen den Variablen in die Gruppenbildung einbeziehen. Damit wird die restriktive Annahme einer diagonalen Kovarianzmatrix überwunden und eine modellbasierte Gruppenbildung vorgenommen, wie sie u. a. von McLachlan & Basford (1988) ge-

fordert wird.

6 Eine Simulationsstudie zur Analyse heterogener Daten mit SEM-Modellen

6.1 Zum Aufbau der Studie

Da im Bereich des Marketing nur sehr wenige vergleichende Studien zur Anwendung der vorgestellten Analysemethoden existieren, wird im folgenden eine Simulationsstudie durchgeführt, die die Ergebnisse der verschiedenen Verfahren zur Analyse heterogener Daten mit SEM-Modellen gegenüberstellt.

Dabei werden die verschiedenen Methoden auf drei simulierte Datensätze von unterschiedlicher Größe angewendet, wobei jeweils verschiedene Annahmen über die Heterogenität der Daten getroffen werden (vgl. Tabelle 2), um einen besseren Einblick in die Leistungsfähigkeit der Verfahren zu erhalten.

Annahme	Modell
Homogenität	SEM-Modell
Beobachtbare Heterogenität	Mehrgruppen SEM-Modell
Unbeobachtbare Heterogenität	Traditioneller Analyseansatz Finite Mischungen bedingter Mittelwert- und Kovarianzstrukturmodelle

Tabelle 2: Heterogenitätsannahmen in der Simulationsstudie

Daten

Die drei simulierten Datensätze weisen jeweils Stichprobengrößen von 50.000, 20.000 und 5.000 Beobachtungen auf. Es wird mit solch großen Stichprobenumfängen gearbeitet, um den Stichprobenfehler beim Vergleich der Ergebnisse der verschiedenen Analyseverfahren vernachlässigen zu können. Die Daten setzen sich jeweils aus zwei Gruppen zusammen, die mit einem Anteil von je 50% in der Stichprobe vertreten sind. Weiterhin wird Normalverteilung der Variablen in den Gruppen unterstellt. Die Simulation der Daten erfolgt mit Hilfe des Softwarepaketes TETRAD (vgl. Scheines,

Spirtes, Glymour & Meek, 1994), das eine Routine für die Erzeugung von Kovarianzstrukturdaten enthält. Die zunächst separat für jede Gruppe erzeugten Datensätze werden in einem zweiten Schritt zu einer heterogenen Datendatei zusammengefügt.

Modell Für das der Simulation zugrundeliegende Modell wird unterstellt, daß keine Informationen über die Quellen der Heterogenität verfügbar sind, d. h. es handelt sich um eine unbedingte Mischverteilung. Die Kovarianzstruktur der beobachtbaren Variablen wird durch die SEM-Modelle in Abbildung 3 dargestellt. Die Gruppen

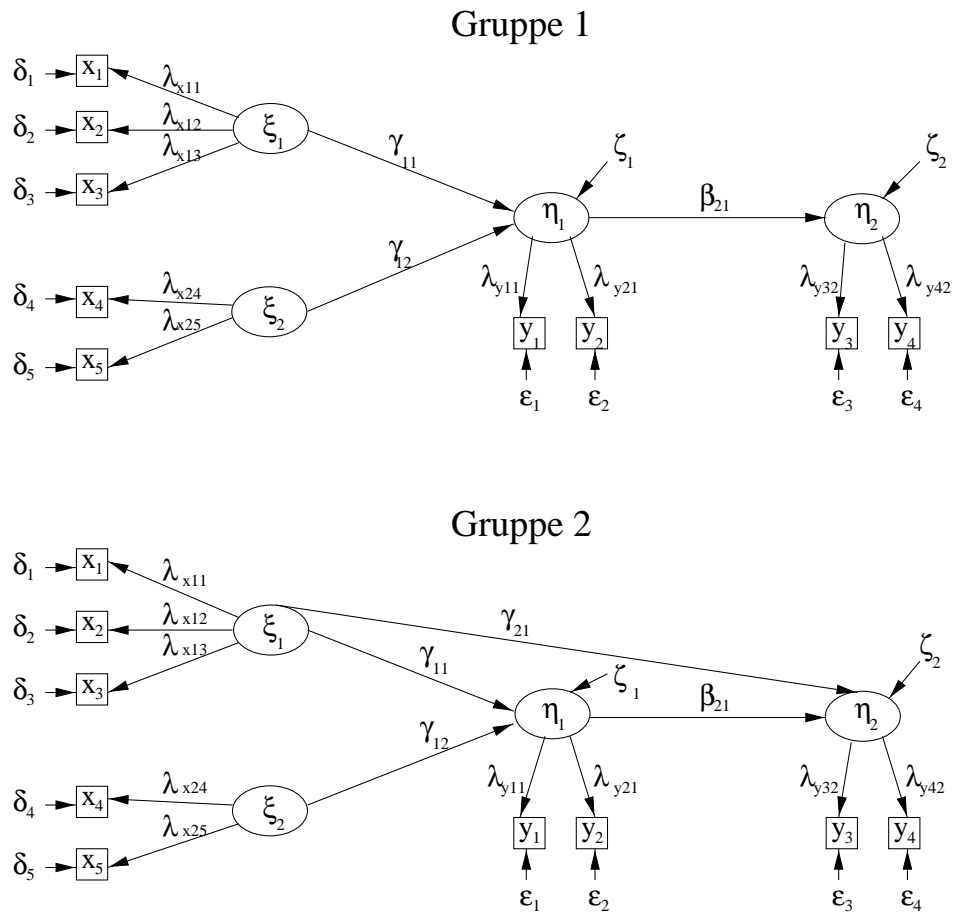


Abbildung 3: Simulierte Kovarianzstruktur der Daten

unterscheiden sich sowohl in Hinblick auf ihre Mittelwerte als auch in der Beziehungsstruktur. Einen Überblick über die der Simulation zugrundeliegenden Parameterwerte gibt Tabelle 3.

Pfad- koeffizienten	Gruppe 1	Gruppe 2	Mittel- werte	Gruppe 1	Gruppe 2
γ_{11}	0,70	0,20	$\mu_{\xi 1}$	0,00	1,00
γ_{12}	0,60	-0,30	$\mu_{\xi 2}$	0,00	1,00
γ_{21}	0,00	0,80	$\mu_{\eta 1}$	0,00	1,00
β_{21}	0,60	0,30	$\mu_{\eta 2}$	0,00	1,00
λ_{x1}	1,00	1,00	μ_{x1}	0,00	1,00
λ_{x2}	0,90	0,90	μ_{x2}	0,00	1,00
λ_{x3}	0,80	0,80	μ_{x3}	0,00	1,00
λ_{x4}	1,00	1,00	μ_{x4}	0,00	1,00
λ_{x5}	0,90	0,90	μ_{x5}	0,00	1,00
λ_{y1}	1,00	1,00	μ_{y1}	0,00	1,00
λ_{y2}	0,90	0,90	μ_{y2}	0,00	1,00
λ_{y3}	1,00	1,00	μ_{y3}	0,00	1,00
λ_{y4}	0,80	0,80	μ_{y4}	0,00	1,00

Tabelle 3: Simulierte Modellparameter

Analyse

Die simulierten Daten werden unter verschiedenen Heterogenitätsannahmen ausgewertet, um die Ergebnisse der vorgestellten Analyseverfahren zu vergleichen. Zunächst werden homogene Daten unterstellt und ein SEM-Modell geschätzt, um die Gefahren nicht erkannter Heterogenität der Stichprobe zu verdeutlichen. Daran anschließend erfolgt unter der Annahme beobachtbarer Heterogenität die Schätzung eines Mehrgruppen SEM-Modells. In einem weiteren Schritt wird unbeobachtbare Heterogenität angenommen. Für diese Bedingung wird sowohl der traditionelle stufenweise Analyseansatz angewendet, als auch ein Modell auf Basis von Mischverteilungen geschätzt (vgl. Tabelle 2). Zur Anwendung kommen die SEM Software Programme AMOS 3.6 (Arbuckle, 1997) für die einfachen Mehrgruppenanalysen und MECOSA3.01 (Arminger, Wittenberg & Schepers, 1996).

6.2 Ergebnisse der Studie

Die Ergebnisse der Modellschätzungen sind in den Tabellen 4 bis 6 dokumentiert. Es ist zu erkennen, daß sich die Vermutungen über die Eignung der vorgestellten Verfahren zur Analyse heterogener Daten mit SEM-Modellen weitgehend bestätigt haben.

Zur Beurteilung der Anpassung des Modells an die Daten werden der χ^2 -Test sowie eine Auswahl allgemein gebräuchlicher deskriptiver Fit-Indizes (AIC, RMSEA, AGFI) aufgeführt (zur Diskussion vgl. Bagozzi & Baumgartner, 1994). Sowohl der χ^2 -Test als auch das AIC weisen nur für das Mehrgruppen SEM-Modell und die finiten Mischungen von Strukturgleichungsmodellen einen akzeptablen Fit aus. Das SEM-Modell bei unterstellter Homogenität der Daten sowie das traditionelle Modell bei unbeobachtbarer Heterogenität werden eindeutig abgelehnt.

Der RMSEA und der AGFI differenzieren die verschiedenen Analysemodelle weniger deutlich. Während der RMSEA beim allgemeinen SEM-Modell zwar Grenzwerte annimmt, aber zu keiner eindeutigen Ablehnung führt, weist der AGFI bei allen Modellen akzeptable Werte auf. Diese Ergebnisse bestätigen die Aussagen von Jedidi, Jagpal & DeSarbo (1994), wonach viele der häufig verwendeten Fit-Indizes, wie z. B. der GFI und der AGFI, nicht geeignet sind, um eine Heterogenität der Daten zu erkennen. Die Autoren schlagen vor, statt dessen das AIC bzw. ähnliche globale Fitmaße, wie z. B. CAIC oder BIC, zu untersuchen.

Vergleicht man die Ergebnisse der Schätzungen mit unterschiedlichen Stichprobengrößen, so liefert die Modellschätzung mit Homogenitätsannahme erwartungsgemäß die schlechteste Anpassung. Es zeigt sich aber auch, daß bei der Benutzung von üblichen Schwellenwerten zur Prüfung der Modellanpassung, wie etwa in Homburg & Giering (1998) mit $AGFI \geq 0,9$, ein geschätztes Modell angenommen wird, das auf falschen Annahmen beruht. Die Modelle mit beobachtbarer Heterogenität werden in allen Samples akzeptabel angepaßt, gleiches gilt für den Finite Mixture Ansatz. Die traditionelle stufenweise Vorgehensweise kann dagegen weniger gute Modellanpassungen leisten, auch werden die Gruppen weniger gut reproduziert.

Nachdem in den vorangegangenen Abschnitten die Anpassung des Gesamtmodells an die Daten diskutiert wurde, wird nun die Güte der einzelnen Parameterschätzwerte beurteilt. In den Tabellen 4 bis 6 werden lediglich die Parameterschätzwerte des

Strukturmodells dokumentiert. Da sich die Struktur der Meßbeziehungen in dem der Simulation zugrundeliegenden Modell zwischen den Gruppen nicht unterscheidet, weisen die Parameterschätzwerte des Meßmodells keine nennenswerten Unterschiede auf. Die spezifizierten Faktorladungen λ_{x1} bis λ_{x5} und λ_{y1} bis λ_{y4} werden in allen Modellen durchweg gut reproduziert. Auf eine detaillierte Darstellung dieser Parameterschätzwerte wird aus Gründen der Übersichtlichkeit verzichtet.

Für die Parameter des Strukturmodells kann festgehalten werden, daß die verschiedenen Modelle zu abweichenden Schätzergebnissen gelangen. Das allgemeine SEM-Modell, das eine homogene Grundgesamtheit unterstellt, liefert die schlechtesten Schätzwerte, was jedoch durch die schlechte Anpassung des Gesamtmodells an die Daten bereits zu erwarten war. Insbesondere γ_{12} und γ_{21} , d. h. die Parameterwerte, in denen sich die Gruppen wesentlich unterscheiden, werden verzerrt geschätzt. Das Mehrgruppen SEM-Modell bei beobachtbarer Heterogenität reproduziert die Parameterwerte für die beiden Gruppen gut. Dabei wird auch erkannt, daß γ_{21} in Gruppe 1 nur zufällig von Null verschieden ist. Eine gleichermaßen akzeptable Schätzung der Parameter liefern auch die finiten Mischungen von Strukturgleichungsmodellen bei unbeobachtbarer Heterogenität. Zusätzlich zu den Pfadkoeffizienten des Modells werden hier die Gruppenanteile an der Grundgesamtheit geschätzt. Diese können mit einer Fehlzuordnungsquote von weniger als 1% als äußerst zuverlässig bewertet werden. Bei einer Verringerung des Stichprobenumfangs ist zwar die Tendenz zu einer leichten Erhöhung des Zuordnungsfehlers zu erkennen, die Ergebnisse sind jedoch immer noch deutlich besser als beim traditionellen Ansatz.

Auch bei den Parameterschätzwerten für die Pfadkoeffizienten des Strukturmodells liefert der traditionelle Ansatz schlechtere Ergebnisse als die finiten Mischungen von Strukturgleichungsmodellen. In allen drei Stichproben wird für γ_{21} in Gruppe 1 ein signifikanter Wert geschätzt, obwohl der Parameter in den Daten nicht von Null verschieden ist. Weiterhin wird bei einem Stichprobenumfang von 50.000 γ_{11} in Gruppe 2 mit einem negativen Vorzeichen geschätzt, obwohl im Modell eine positive Beziehung spezifiziert ist. Bei einem Stichprobenumfang von 20.000 Beobachtungen wird γ_{12} in Gruppe 2 in seiner Stärke deutlich unterschätzt.

7 Schlußbemerkungen

Nach den vorliegenden Simulationsergebnissen erweisen sich die Schätzungen mit dem Finite Mixture Ansatz (hier mit dem Programm MECOSA3.01) gegenüber der traditionellen stufenweisen Schätzmethode als überlegen. Allerdings ist der Modellansatz in der vorgestellten Form an die relativ strengen Annahmen der Normalverteilung der Daten gebunden. Mit der Möglichkeit simultan latente Gruppen zu identifizieren und für diese Gruppen Strukturgleichungsmodelle zu schätzen, liefert der Finite Mixture Ansatz jedoch eine an die Bedürfnisse der Marketingforschung mit seinen Segmentierungsaufgaben besser angepaßte Methodik. Allerdings ist die Schätzung von Finite Mixture Modellen mit den vorhandenen Programmen an große Stichproben gebunden, um konsistente Schätzwerte zu erzielen. In der Marktforschung, sieht man von den Konsumentenpaneln ab, wird meist mit kleinen Stichproben gearbeitet, so daß weitere Erkenntnisse darüber benötigt werden, wie reliabel die Schätzungen von latenten Strukturen und "kausalen" Beziehungen bei kleineren Stichproben bleiben.

Literatur

- Arbuckle, J. (1997). *AMOS User's Guide, Version 3.6*. SmallWaters Corporation.
- Arminger, G. and Stein, P. (1997, April). Finite Mixtures of Covariance Structure Models with Regressors. Technical report, Bergische Universität Wuppertal and Gerhard Mercator Universität Duisburg.
- Arminger, G., Stein, P., and Wittenberg, J. (1998, May). Mixtures of Conditional Mean- and Covariance Structure Models. Technical report, Bergische Universität Wuppertal and Gerhard Mercator Universität Duisburg.
- Arminger, G., Wittenberg, J., and Schepers, A. (1996). *MECOSA 3 User Guide*. ADDITIVE GmbH.
- Bagozzi, R. and Baumgartner, H. (1994). The Evaluation of Structural Equation Models and Hypothesis Testing. In *Principles of Marketing Research*, pp. 386–422. Bagozzi, R.
- Becker, M., Yang, I., and Lange, K. (1997). EM Algorithms without Missing Data. *Statistical Methods in Medical Research* (6), 37–53.

- Dempster, A., Laird, N., and Rubin, D. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm (with discussion). *Journal of the Royal Statistical Society, B* 39, 1–38.
- Hildebrandt, L. and Homburg, C. H. (1998). *Die Kausalanalyse: Instrument der empirischen betriebswirtschaftlichen Forschung*. Stuttgart: Schäffer-Poeschel.
- Homburg, C. and Baumgartner, H. (1995). Die Kausalanalyse als Instrument der Marketingforschung. *Zeitschrift für Betriebswirtschaft* 65, 1091–1108.
- Homburg, C. and Giering, A. (1998). Konzeptualisierung und Operationalisierung komplexer Konstrukte - Ein Leitfaden für die Marketingforschung. In L. Hildebrandt & C. Homburg (Eds.), *Die Kausalanalyse*, pp. 111–148. Schäffer-Poeschel.
- Homburg, C. and Hildebrandt, L. (1998). Die Kausalanalyse: Bestandsaufnahme, Entwicklungsrichtungen, Problemfelder. In L. Hildebrandt & C. Homburg (Eds.), *Die Kausalanalyse*, pp. 16–43. Schäffer Poeschel Verlag.
- Jedidi, K., Jagpal, H., and DeSarbo, W. (1994, April). Latent Class Structural Equation Models for Marketing Research. Working paper.
- Jedidi, K., Jagpal, H., and DeSarbo, W. (1997). Finite-Mixture Structural Equation Models for Response-Based Segmentation and Unobserved Heterogeneity. *Marketing Science* 16(1), 39–59.
- Jones, P. and McLachlan, G. (1992). Fitting Finite Mixture Models in a Regression Context. *Australian Journal of Statistics* 32(2), 233–240.
- Jöreskog, K. and Sörbom, D. (1993). *LISREL 8: User's Reference Guide*. Scientific Software International, Inc.
- McLachlan, G. and Basford, K. (1988). *Mixture Models: Inference and Applications to Clustering*. Number 84 in Statistics: textbooks and monographs. Marcel Dekker, Inc.
- Scheines, R., Spirtes, P., Glymour, C., and Meek, C. (1994). *TETRAD II: Tools for Causal Modelling*. Pittsburgh.
- Sörbom, D. (1981). *Systems under Indirect Observation*, Chapter Structural Equation Models with Structured Means, pp. 184–195. North-Holland.
- Titterton, D., Smith, A., and Makov, U. (1985). *Statistical Analysis of Finite Mixture Distributions*. Wiley, Chichester.

Ullmann, J. (1996). Structural Equation Modeling. In B. Tabachnik & L. Fidell (Eds.), *Using Multivariate Statistics* (3 ed.), Chapter 14, pp. 709–811. HarperCollins College Publishers.

n=50.000				
	SEM- Modell	Mehrgruppen SEM-Modell	Traditioneller Analyseansatz	Finite Mischungen
Modell-Fit				
AIC				
- geschätztes Modell	3.315,87	124,74	1.340,15	121,22
- gesättigtes Modell	90,00	180,00	180,00	180,00
- Unabhängig- keitsmodell	362.285,85	216.835,64	188.615,51	216.712,49
χ^2 -Test	3.269,87	42,74	1.258,15	39,22
- df	22	49	49	49
- p	0,00	0,72	0,00	0,84
RMSEA (< 0,05)	0,05	0,00	0,02	0,00
AGFI (> 0,90)	0,97	1,00	0,99	1,00
Parameter- schätzwerte				
Gruppenanteile				
- ω_1			0,5975	0,5006
- ω_2			0,4025	0,4994
Modellparameter Gruppe 1				
- γ_{11}	0,55	0,69	0,56	0,69
- γ_{12}	0,28	0,60	0,41	0,59
- γ_{21}	0,72	-0,07 (n.s.)	0,12	-0,05 (n.s.)
- β_{21}	0,55	0,60	0,60	0,60
Modellparameter Gruppe 2				
- γ_{11}		0,20	-0,03	0,20
- γ_{12}		-0,30	-0,30	-0,30
- γ_{21}		0,79	0,62	0,79
- β_{21}		0,30	0,21	0,30

Tabelle 4: Simulationsergebnisse für n=50.000

n=20.000				
	SEM- Modell	Mehrgruppen SEM-Modell	Traditioneller Analyseansatz	Finite Mischungen
Modell-Fit				
AIC				
- geschätztes Modell	1.434,06	126,20	837,79	125,40
- gesättigtes Modell	90,00	180,00	180,00	180,00
- Unabhängig- keitsmodell	144.245,45	86.623,66	73.350,96	86.656,23
χ^2 -Test	1.388,06	44,20	755,59	43,40
- df	22	49	49	49
- p	0,00	0,67	0,00	0,70
RMSEA (< 0,05)	0,06	0,00	0,03	0,00
AGFI (> 0,90)	0,97	1,00	0,98	1,00
Parameter- schätzwerte				
Gruppenanteile				
- ω_1			0,3684	0,5011
- ω_2			0,6316	0,4989
Modellparameter Gruppe 1				
- γ_{11}	0,56	0,69	0,64	0,69
- γ_{12}	0,30	0,59	0,54	0,59
- γ_{21}	0,72	-0,01 (n.s.)	-0,08	-0,01 (n.s.)
- β_{21}	0,55	0,59	0,42	0,59
Modellparameter Gruppe 2				
- γ_{11}		0,19	0,27	0,18
- γ_{12}		-0,31	-0,06	-0,31
- γ_{21}		0,79	0,85	0,79
- β_{21}		0,30	0,39	0,31

Tabelle 5: Simulationsergebnisse für n=20.000

n=5.000				
	SEM-Modell	Mehrgruppen SEM-Modell	Traditioneller Analyseansatz	Finite Mischungen
Modell-Fit				
AIC				
- geschätztes Modell	435,19	124,04	275,26	123,79
- gesättigtes Modell	90,00	180,00	180,00	180,00
- Unabhängigkeitsmodell	36.232,09	22.087,55	18.278,80	22.231,77
χ^2 -Test	389,19	42,04	193,26	41,79
- df	22	49	49	49
- p	0,00	0,75	0,00	0,76
RMSEA (< 0,05)	0,06	0,00	0,02	0,00
AGFI (> 0,90)	0,96	1,00	0,98	1,00
Parameter-schätzwerte				
Gruppenanteile				
- ω_1			0,3908	0,5044
- ω_2			0,6092	0,4956
Modellparameter Gruppe 1				
- γ_{11}	0,56	0,68	0,61	0,69
- γ_{12}	0,28	0,62	0,59	0,63
- γ_{21}	0,71	-0,04 (n.s.)	-0,11	-0,03 (n.s.)
- β_{21}	0,55	0,66	0,53	0,64
Modellparameter Gruppe 2				
- γ_{11}		0,23	0,30	0,25
- γ_{12}		-0,33	-0,14	-0,33
- γ_{21}		0,82	0,89	0,81
- β_{21}		0,27	0,29	0,27

Tabelle 6: Simulationsergebnisse für n=5.000