

# Modelle zur Aufdeckung unbeobachteter Heterogenität bei der Erklärung von Lebenszufriedenheit

Petra Stein

Gerhard Mercator Universität Duisburg, Fachbereich 1/Soziologie, Lotharstraße 65, D-47048 Duisburg

**Zusammenfassung:** Ausgehend von dem zentralen Problem der unbeobachteten Heterogenität in der empirischen Sozialforschung wird zur Lösung des Problems das statistische Verfahren der finiten Mischung von Mittelwert- und Kovarianzstrukturmodellen vorgestellt. In diesem Modell kann der Erwartungswert jeder Komponente der Mischung sowohl von normalverteilten als auch von nicht normalverteilten Regressorvariablen abhängen. Der Erwartungswert und die Kovarianzmatrix jeder Komponente der Mischung werden als konditionale Mittelwert- und Kovarianzstrukturmodelle (z. B. vom Typ LISREL) parametrisiert. Zur Schätzung der Parameter werden drei Schätzstrategien vorgestellt, Teststatistiken für die geschätzten Parameter, Fit-Indices zur Modellprüfung sowie zwei Testverfahren auf die Anzahl der Komponenten werden angegeben. Modellspezifikation und Schätzverfahren werden in der Form von Mischungen konditionaler LISREL-Modelle zur Identifikation heterogener Personengruppen bezüglich der Lebenszufriedenheit an den Wohlfahrtssurvey 1993 angewendet. Als Ergebnis werden drei Typen identifiziert, die sowohl einen unterschiedlichen Grad an Lebenszufriedenheit als auch eine unterschiedliche Beziehungsstruktur zwischen den Zufriedenheiten in einzelnen Lebensbereichen, allgemeiner Lebenszufriedenheit und sozialstrukturellen Determinanten aufweisen.

## 1. Problembeschreibung

In der empirischen Sozialforschung werden bei der Analyse multivariater Daten immer häufiger komplexe Modelle angewandt, die zum einen differenzierte Beziehungen zwischen abhängigen Variablen untereinander und zwischen abhängigen und unabhängigen Variablen abbilden und zum anderen Beziehungen zwischen Indikatoren von latenten Variablen zu latenten Variablen durch ein Messmodell herstellen. Der erste Modelltyp umfasst die multivariate Regression, die multivariate Varianz- und Kovarianzanalyse sowie die simultanen Gleichungssysteme der Ökonometrie. Der zweite Modelltyp umfasst die konfirmatorische Faktorenanalyse. Kombiniert man den ersten mit dem zweiten Modelltyp, indem latente Variablen durch ein faktorenanalytisches Messmodell mit beobachtbaren Indikatoren verknüpft werden und auf die latenten Variablen Regressionsmodelle angewandt werden, erhält man die Klasse der Mittelwert- und Kovarianzstrukturmodelle. Aus historischen Gründen werden für diese Modelle auch die Begriffe „Strukturgleichungsmodelle“ und „Kausalmodelle“ synonym benutzt, die jedoch etwas irreführend sind, da einerseits die Klasse der Mittelwert- und Kovarianzstrukturmodelle über die Strukturgleichungsmodelle hinausgeht und andererseits Kausalbeziehungen nur vermutet und falsifiziert, jedoch nicht durch diese Modelle bewiesen werden können. Eine

Darstellung des allgemeinen Modells sowie wichtige Submodelle, wie etwa das LISREL-Modell findet man in Browne/Arminger (1995). Anwendungen lassen sich z. B. in Hildebrand (1983), Faulbaum (1987) und Jagodzinski et al. (1987) finden.

Zentrale Voraussetzung einer konsistenten Schätzung der Parameter eines Mittelwert- und Kovarianzstrukturmodells ist die Annahme, dass die Stichprobenelemente aus der selben Population stammen (Populationshomogenität). Ist die Homogenitätsbedingung verletzt, aber die kritische Variable, die die Inhomogenität erzeugt, bekannt, lässt sich die Gesamtstichprobe aus der heterogenen Population in unabhängige Teilstichproben aus bekannten Subpopulationen aufteilen, die in sich homogen sind. In diesem Fall lassen sich mit Hilfe der simultanen Analyse mehrerer Stichproben, die in Programmen zur Analyse von Mittelwert- und Kovarianzstrukturen wie LISREL (Jöreskog/Sörbom 1993) oder EQS (Bentler 1995) als Multiple-Gruppen-Option implementiert sind, getrennte Analysen für jede Merkmalsausprägung der kritischen Variablen berechnen und Hypothesen über die Gleichheit von Strukturen in den einzelnen Subpopulationen für jede Merkmalsausprägung der kritischen Variablen testen.

In der sozialwissenschaftlichen Forschung liegt jedoch häufig das Problem vor, dass sowohl die Zugehörigkeit der Elemente zu den einzelnen Subpopulationen als auch die Anzahl der Subpopulationen

unbekannt sind. Im Unterschied zur Voraussetzung der Mehrgruppenanalyse ist keine kritische Variable bekannt, die die Heterogenität der Stichprobe erzeugt. In diesem Fall, also im Fall der unbeobachteten Heterogenität, werden üblicherweise zunächst Klassifikationsverfahren in Form von explorativen clusteranalytischen Verfahren eingesetzt, mit deren Hilfe eine Population derart in Teilmengen zerlegt wird, dass die Elemente (Personen) innerhalb einer Teilmenge bezüglich der in die Klassifikation einbezogenen Variablen eine homogene Subpopulation bilden. Nach erfolgter Klassifikation wird für jede Subpopulation ein eigenes Mittelwert- und Kovarianzstrukturmodell formuliert, das in gleicher Weise unter Verwendung der Multiplen-Gruppen-Option der herkömmlichen Mittelwert- und Kovarianzstrukturanalyse wie im Fall einer Population mit bekannter Heterogenität analysiert wird. Es handelt sich daher um eine sequentielle Verbindung von Methoden zur Auffindung einer Typologie und zur Analyse von Ursache-Wirkungszusammenhängen.

Eine derartige sequentielle Vorgehensweise ist mit einer Reihe statistischer und forschungstechnischer Problemen verbunden. Diese Probleme betreffen zum einen die Anwendung der Clusteranalyse und zum anderen die Anwendung der Multiplen-Gruppen-Option in heterogenen Populationen, in denen die kritische Variable, die die Heterogenität erzeugt, unbekannt ist.

Probleme, die sich bei der Anwendung der Clusteranalyse ergeben, sind ausführlich in Stein (1997) dargestellt und werden im Folgenden stichpunktartig zusammengefasst: Ein Problem liegt in der generellen Schwierigkeit der Berechnung von statistischen Kennzahlen für die einzelnen Cluster. Werden in der Clusteranalyse mit Hilfe eines nicht überlappenden Zuordnungsverfahrens Individuen einem Cluster zugeordnet, kann es zu Fehlklassifikationen kommen, die umso häufiger auftreten, je schlechter die Cluster voneinander getrennt sind. Werden nun Verteilungskennzahlen wie Erwartungswerte und Varianzen der Cluster nach erfolgter Klassifikation aus den Daten der einzelnen Cluster geschätzt, kann es aufgrund der Fehlklassifikationen zu inkonsistenten Schätzungen dieser Kennzahlen kommen. Diese fehlerhaften Schätzungen führen bei nachfolgenden Analysen, wie Mittelwert- und Kovarianzstrukturanalysen, wiederum zu fehlerhaften Schätzungen. Ein weiteres Problem ist darin zu sehen, dass bei der Bestimmung der Cluster keine Konditionierung auf erklärende Variablen vorgenommen werden kann. Erst nach erfolgter Klassifikation können die extrahierten Cluster mit Variablen

in Beziehung gesetzt werden, die als Prädiktoren für die Zugehörigkeit zu einem Cluster fungieren (z. B. durch Diskriminanzanalysen oder Regressionsanalysen). Das nächste Problem ist die Bestimmung der Anzahl der Cluster, da es keinen formalen Test für die Anzahl der Cluster gibt und damit die angemessene Clusterzahl nur auf explorativen Weg ermittelt werden kann. Als letztes Problem ist festzuhalten, dass die Clusteranalyse ein exploratives Verfahren ist, das keine Schätzung der asymptotischen Kovarianzmatrix der univariaten und multivariaten Kennzahlen eines Clusters liefert.

Die Anwendung der Multiplen-Gruppen-Option der herkömmlichen Mittelwert- und Kovarianzstrukturanalyse in Populationen mit unbeobachteter Heterogenität ist aus folgenden Gründen fehlerhaft: Eine zentrale Annahme der Mehrgruppenanalyse ist, dass die Stichproben in den Gruppen voneinander statistisch unabhängig sind. Dies bedeutet, dass die Kovarianzmatrix des Parametervektors, der die zu schätzenden Parameter enthält, blockdiagonal ist. Wie Arminger et al. (1999) zeigen, ist dies beim Vorliegen unbeobachteter Heterogenität nicht der Fall, da die Schätzungen der Parameter der verschiedenen Gruppen korrelieren. Dementsprechend kann nicht von einer statistischen Unabhängigkeit der Gruppen ausgegangen werden. Daher liefert die Verwendung der Multiplen-Gruppen-Option von Programmen für Mittelwert- und Kovarianzstrukturen im Fall der unbeobachteten Heterogenität zwar konsistente Parameterschätzer, jedoch inkonsistente Schätzer der asymptotischen Kovarianzmatrix der Parameter und damit falsche Standardfehler sowie falsche Teststatistiken.

Ausgehend von diesen zentralen methodischen Problemen wird im Folgenden als mögliche Verbesserung der Analysetechniken zur Verbindung von Methoden zur Auffindung von Typen und zur Analyse von Ursache-Wirkungsmodellen ein Verfahren vorgestellt, das technisch als finite Mischung konditionaler Mittelwert- und Kovarianzstrukturmodelle konstruiert ist. Dabei wird eine Synthese von zwei Modellklassen durchgeführt, indem die statistischen Verfahren, Subpopulationen zu identifizieren, mit parametrischen Mittelwert- und Kovarianzstrukturmodellen für bekannte Subpopulationen verbunden werden. Mit der Entwicklung dieses neuen statistischen Verfahrens wird der empirischen Sozialforschung ein Analyseinstrument zur Verfügung gestellt, das es ermöglicht, eine Population mit unbeobachteter Heterogenität in homogene Subpopulationen konditional auf erklärende Variablen zu zerlegen und komplexe Modelle über einseitige und wechselseitige Beziehungen von Va-

riablen in den Subpopulationen in Form von konditionalen Mittelwert- und Kovarianzstrukturmodellen zu spezifizieren und zu überprüfen.

In diesem Artikel wird versucht, das Verfahren in anschaulicher Weise auch dem technisch weniger interessierten Leser zugänglich zu machen. Dies geschieht erstens dadurch, dass vor allem das Modell ausführlich dargestellt und auf technische Einzelheiten wie Schätztechniken, Testverfahren und Beweise verzichtet wird. Diese werden in Arminger et al. (1999) dargestellt und von Stein (1997) und Arminger/Stein (1997) ausführlich erläutert. Zweitens geschieht dies durch die Anwendung des Verfahrens auf ein empirisches Beispiel im Bereich der Lebensqualitätsforschung, das dem Leser verdeutlicht, in welcher Weise das Verfahren in der inhaltlichen Forschung eingesetzt werden kann.

Der Artikel enthält nachfolgend zwei Hauptpunkte: In Abschnitt 2 erfolgt zunächst eine ausführliche Darstellung des statistischen Modells und seiner Entwicklung, eine kurze verbale Präsentation der Schätzmethoden sowie die Diskussion über die Wahl der spezifischen Schätzmethode in Abhängigkeit von der jeweiligen Problemstellung des Forschers. Dieser Abschnitt endet mit einer Skizze über die Testverfahren. In Abschnitt 3 wird das Verfahren auf die Modellierung der Lebenszufriedenheit basierend auf dem Datensatz des Wohlfahrts-surveys 1993 angewendet. Ziel ist die Identifikation heterogener Personengruppen, die eine unterschiedliche Beziehungsstruktur zwischen den Zufriedenheiten in einzelnen Lebensbereichen, allgemeiner Lebenszufriedenheit und sozialstrukturellen Determinanten aufweisen. Damit wird an diesem Beispiel deutlich, dass die übliche Behandlung eines Datensatzes als Stichprobe aus einer homogenen Population zu irreführenden Ergebnissen führen kann.

## 2. Finite Mischungen heterogener Populationen

Ein Standardmodell zur Behandlung unbeobachteter Heterogenität ist das statistische Modell der finiten Mischung. Das Modell geht von der Überlegung aus, dass eine Population nicht homogen ist, sondern aus einer unbekanntem Anzahl  $K$  heterogener Subpopulationen besteht. Diese Subpopulationen sind unbekannt. In der Begriffssprache der Statistik über finite Mischungen werden sie als Komponenten und der Anteil  $\pi_k$  jeder Komponente  $k$  an der Population als mischende Wahrscheinlichkeit bezeichnet (vgl. Everitt/Hand 1981).

Finite Mischungen als Klassifikationsverfahren sind in der empirischen Sozialforschung fast ausschließlich in der Form der latenten Klassenanalyse für dichotome und qualitativ abhängige Variablen (Goodman 1974; Clogg 1995) bekannt.<sup>1</sup> Im Unterschied zur latenten Klassenanalyse wird in dem hier vorgestellten Modell von einer finiten Mischung metrischer Zufallsvariablen ausgegangen. Es wird angenommen, dass jedes Stichprobenelement mit einer bestimmten Wahrscheinlichkeit aus einer Subpopulation stammt, die eine spezielle Verteilung der Zufallsvariablen aufweist. Die Schwierigkeit besteht darin, die Verteilungen der Subpopulationen zu identifizieren und die Parameter der einzelnen Verteilungen zu schätzen. Die Möglichkeit, den Anteil jeder Komponente an der Population (die mischenden Wahrscheinlichkeiten) und die Parameter der Verteilungen der Subpopulationen zu schätzen, wird erst durch spezielle Annahmen über diese Verteilung möglich. Daher wird in finiten Mischungen von metrischen Zufallsvariablen von der Standardannahme ausgegangen, dass die Stichprobenwerte innerhalb jeder Komponente der Mischung einer Normalverteilung folgen (vgl. McLachlan/Basford 1988).

### 2.1 Finite Mischungen von multivariaten Normalverteilungen

Das Modell einer finiten Mischung multivariater Normalverteilungen ist formal durch die Dichtefunktion

$$f(\mathbf{y}_i) = \sum_{k=1}^K \pi_k \phi(\mathbf{y}_i; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \quad (1)$$

definiert. In dem Modell ist  $\mathbf{y}_i$  ein Vektor von Zufallsvariablen,  $K$  die Anzahl der Komponenten der Mischverteilung,  $\pi_k$  gibt die Wahrscheinlichkeit an, daß  $\mathbf{y}_i$  einer Komponente  $k$  entstammt, wobei  $\pi_k \geq 0$  und  $\sum_{k=1}^K \pi_k = 1$  ist. Die Wahrscheinlichkeiten lassen sich als Anteile der heterogenen Subpopulationen an der Grundgesamtheit interpretieren.  $\phi(\mathbf{y};$

<sup>1</sup> Die latente Profilanalyse von Gibson (1966) und Lazarsfeld/Henry (1968) für multivariate stetige Variablen hat sich wegen der restriktiven Modellspezifikation kaum durchgesetzt. In der latenten Profilanalyse werden die Subpopulationen nur durch unterschiedliche Mittelwerte charakterisiert, wobei in Analogie zur explorativen Faktorenanalyse die Kovarianzmatrix jeder Komponente als Diagonalmatrix spezifiziert wird. Dadurch lassen sich keine Hypothesen über Zusammenhänge zwischen Variablen in den einzelnen Subpopulationen spezifizieren. Dieses Verfahren ist ein Spezialfall des hier vorgestellten Modells.

$\mu_k; \Sigma_k$ ) ist die Dichte einer multivariat normalverteilten Zufallsvariable mit dem Erwartungswert  $\mu_k$  und der Kovarianzmatrix  $\Sigma_k$ .

Bei einer Zufallsstichprobe vom Umfang  $n$  aus einer Mischung von  $K$  multivariaten Normalverteilungen liegt das statistische Problem vor, den Erwartungswert  $\mu_k$  und die Kovarianzmatrix  $\Sigma_k$  jeder Komponentenverteilung, die Anzahl der Komponenten der Mischverteilung sowie die mischenden Wahrscheinlichkeiten  $\pi_k$ ,  $k=1, \dots, K-1$  und damit die Größe der Komponenten und die Zugehörigkeit der Stichprobenelemente zu einer Komponente zu schätzen. Die Schätzung der Parameter erfolgt mit Hilfe der Maximum-Likelihood-Methode und dem EM-Algorithmus und wird detailliert diskutiert in Everitt/Hand (1981), McLachlan/Basford (1988) und Stein (1997).

## 2.2 Konditionale Mischungen multivariater Normalverteilungen

Das Modell finiter Mischungen multivariater Normalverteilungen beruht auf der Annahme, dass alle Variablen in jeder Komponente der Mischverteilung multivariat normalverteilt sind. Bei der Modellierung inhaltlicher Fragestellungen ist im Allgemeinen davon auszugehen, dass bestimmte Variablen als abhängige, unabhängige oder sowohl als abhängige als auch als unabhängige Variablen fungieren. In diesem Fall impliziert die Annahme der multivariaten Normalverteilung, dass sowohl die Kriteriumsvariablen als auch die Prädiktorvariablen in den einzelnen Komponenten der Mischverteilung gemeinsam normalverteilt sind. Die Annahme der multivariaten Normalverteilung der Variablen in den einzelnen Komponenten der Mischverteilung erweist sich jedoch häufig für soziologische Fragestellungen als unrealistisch, da typische Prädiktorvariablen, z. B. soziodemographische Variablen wie Geschlecht, Alter und Beruf nicht normalverteilt sind.

Stein (1997) sowie Arminger et al. (1999) erweitern daher die traditionelle Analyse von Mischverteilungen, so dass die Einbeziehung von qualitativen Regressoren in Form von Dummyvariablen und nicht normalverteilten metrischen Regressoren möglich wird. Sie analysieren Mischungen von multivariaten Normalverteilungen, bei denen der Erwartungswert für jede Komponente der Mischverteilung linear von Regressoren abhängt, die selbst nicht normalverteilt sein müssen. Dies geschieht, indem der Erwartungswert jeder Komponente der Mischverteilung  $\mu_{ik}$  als ein multivariates Regressionsmodell

$$\mu_{ik} = \gamma_k + \Pi_k \boldsymbol{x}_i \quad (2)$$

spezifiziert wird.<sup>2</sup> Hier bezeichnet  $\gamma_k$  den Vektor der Regressionskonstanten und  $\Pi_k$  die Matrix der Regressionskoeffizienten, die für jede Komponente  $k$  spezifiziert wird.

Diese Verallgemeinerung des Modells finiter Mischungen von multivariaten Normalverteilungen auf Mischungen konditionaler multivariater Normalverteilungen hat neben dem Vorteil der Einbeziehung von Regressoren zwei weitere Vorzüge. Erstens wird die Möglichkeit geschaffen, die in den Regressoren enthaltene Information explizit in das Mischverteilungsmodell einzubeziehen, was zum Beispiel in der Clusteranalyse nicht möglich ist. Zweitens wird die restriktive und häufig unrealistische Annahme der un konditionalen Normalverteilung der abhängigen Variablen in jeder Komponente der Mischung durch die wesentlich schwächere Annahme der konditionalen Normalverteilung ersetzt.

## 2.3 Mischungen von Mittelwert- und Kovarianzstrukturmodellen

Neben der Verallgemeinerung finiter Mischungen multivariater Normalverteilungen auf konditionale Mischungen kann die traditionelle Analyse finiter Mischungen multivariater Normalverteilungen so erweitert werden, dass der Erwartungswert  $\mu_k$  und die Kovarianzmatrix  $\Sigma_k$  innerhalb jeder Komponente der Mischverteilung unter der Verwendung spezieller Mittelwert- und Kovarianzstrukturmodelle parametrisiert werden. Beispiele hierfür sind Mischungen von konfirmatorischen faktorenanalytischen Modellen (Yung 1997) und Mischungen von LISREL-Modellen, wie sie von Jedidi et al. (1997) und Stein (1997) behandelt werden. Für die beobachteten Variablen  $\boldsymbol{x}$  und  $\boldsymbol{y}$  wird die gemeinsame multivariate Normalverteilung in jeder Komponente der Mischverteilung vorausgesetzt.

<sup>2</sup> Spezialfälle dieses Modells lassen sich bereits in der Ökonometrie finden. DeSarbo/Cron (1988) analysieren konditionale Mischungen univariater Normalverteilungen. Eine Verallgemeinerung dieses univariaten Regressionsmodells auf multivariate Regressionsmodelle für jede Komponente der Mischverteilung behandeln Jones/McLachlan (1992). Beispiele für die Anwendung der bedingten multivariaten Normalverteilung im Marketing findet man in DeSarbo et al. (1992) und Ramaswamy et al. (1993). Eine Erweiterung dieses multivariaten Regressionsmodells auf simultane Gleichungssysteme für jede Komponente der Mischverteilung behandeln Jedidi et al. (1996).

**2.4 Konditionale Mischungen multivariater Normalverteilungen mit parametrischer Struktur**

Das Modell von Stein (1997) und Arminger et al. (1999) fügt alle Modelle der Abschnitte 2.1 bis 2.3 in Form eines allgemeinen parametrischen Modell für Mischungen von konditional normalverteilten Zufallsvariablen zusammen. Formal ist das Modell definiert als

$$f(y_i | \boldsymbol{x}_i) = \sum_{k=1}^K \pi_k \phi(y_i; \mu_{ik}(\vartheta), \Sigma_k(\vartheta)). \tag{3}$$

In diesem Modell ist  $y_i$  ein Vektor von quantitativen abhängigen Zufallsvariablen und  $\boldsymbol{x}_i$  ein Vektor von Regressoren, der sowohl quantitative als auch qualitative Variablen in Form von Dummyvariablen enthält,  $f(y_i | \boldsymbol{x}_i)$  ist die bedingte Dichtefunktion von  $y_i$  gegeben  $\boldsymbol{x}_i$ . Die mischenden Wahrscheinlichkeiten  $\pi_k$  geben den Anteil der heterogenen Subpopulationen an der Gesamtpopulation an. Es wird angenommen, daß die konditionale Dichte jeder Komponenten der Mischverteilung einer Normalverteilung folgt mit einem Erwartungswert  $\mu_{ik}$  und einer Kovarianzmatrix  $\Sigma_k$ , die spezifisch für jede Komponente der Mischung sind. Der Erwartungswert  $\mu_{ik}$  ist als lineare Regressionsfunktion der erklärenden Variablen in der Form  $\mu_{ik} = \gamma_k + \mathbf{II}_k \boldsymbol{x}_i$  spezifiziert, wobei  $\gamma_k$  den  $p \times 1$  Vektor der Regressionskonstanten und  $\mathbf{II}_k$  eine  $p \times q$  Matrix der Regressionskoeffizienten, die für jede Komponente  $k$  spezifiziert werden, bezeichnet.  $\Sigma_k$  ist die Kovarianzmatrix der Fehler der multivariaten Regression.

Der bedingte Erwartungswert und die bedingte Kovarianzmatrix innerhalb jeder Komponente der Mischverteilung können in Form beliebig spezifizierter konditionaler Mittelwert- und Kovarianzstrukturmodelle parametrisiert werden. Formal ausgedrückt bedeutet das, dass die Regressionsgewichte  $\gamma_k$ ,  $\mathbf{II}_k$  und die Kovarianzmatrix  $\Sigma_k$  durch einen Parametervektor  $\vartheta$  parametrisiert werden, der eine spezifische Struktur für  $\gamma_k$ ,  $\mathbf{II}_k$  und  $\Sigma_k$  angibt. Daher werden diese Modelle als konditionale Mittelwert- und Kovarianzstrukturmodelle für jede Komponente der Mischung bezeichnet. Sie umfassen und verallgemeinern die oben genannten Modelle.

Als ein Beispiel für dieses allgemeine Modell entwickelt Stein (1997) ein konditionales LISREL-Modell, das aus einem komponentenspezifischen Messmodell für die beobachtbaren Variablen  $y_i$

$$y_i^{(k)} = v_k + \Delta_k \eta_i^{(k)} + \varepsilon_i^{(k)} \tag{4}$$

und einem komponentenspezifischen Strukturgleichungsmodellmodell für den latenten Variablenvektor  $\eta_i$

$$\eta_i^{(k)} = B_k \eta_i^{(k)} + \Gamma_k \boldsymbol{x}_i + \zeta_i^{(k)} \tag{5}$$

besteht, wobei  $\eta_i^{(k)} = \eta_i$  gegeben  $\boldsymbol{x}_i$  und  $k$  ist,  $\Delta_k$  ist die Ladungsmatrix der latenten Variablen in der  $k$ -ten Komponente,  $\varepsilon_i^{(k)}$  ist der Vektor der Messfehler in der Komponente  $k$ ,  $v_k$  ist der Vektor der Konstanten des Messmodells in Komponente  $k$ ,  $B_k$  ist die Matrix der Regressionskoeffizienten für die latenten Variablen und  $\Gamma_k$  die Matrix der Einflussgrößen der exogenen beobachtbaren Variablen  $\boldsymbol{x}$  auf die latenten Variablen,  $\zeta_i^{(k)}$  ist der Vektor der Fehler im Strukturgleichungsmodell. Es werden die in der herkömmlichen Mittelwert- und Kovarianzstrukturanalyse herkömmlichen Annahmen für das Messmodell  $E(\varepsilon_i^{(k)}) = 0$ ,  $E(\mu_{ik}^{(k)} \varepsilon_i^{(k)}) = 0$  und  $V(\varepsilon_i^{(k)}) = \Theta_k$  sowie für das Strukturgleichungsmodell  $E(\zeta_i^{(k)}) = 0$ ,  $E(\boldsymbol{x}_i \zeta_i^{(k)}) = 0$  und  $V(\zeta_i^{(k)}) = \Psi_k$  getroffen. Unter den obigen Spezifikationen eines Messmodells und eines Strukturgleichungsmodells lässt sich dann der bedingte Erwartungswert  $\mu_{ik} = \gamma_k + \mathbf{II}_k$  und die konditionale Kovarianzmatrix  $\Sigma_k$  der einzelnen Komponenten in der folgenden Weise parametrisieren:

$$\begin{aligned} \gamma_k &= v_k \\ \mathbf{II}_k &= \Lambda(I - B_k)^{-1} \Gamma_k \boldsymbol{x}_i \\ \Sigma_k &= \Lambda(I - B_k)^{-1} \Psi_k (I - B_k)^{-1} \Lambda' + \Theta_k \end{aligned} \tag{6}$$

**2.5 Schätzung**

In dem unter 2.4 vorgestellten Modell finiter Mischungen konditionaler Mittelwert- und Kovarianzstrukturmodelle sind die mischenden Wahrscheinlichkeiten  $\pi_k$ ,  $k = 1, \dots, K-1$ , die Parameter des multivariaten Regressionsmodells für den bedingten Erwartungswert jeder Komponente der Mischverteilung  $\gamma_k$  und  $\mathbf{II}_k$  sowie die Kovarianzmatrix  $\Sigma_k$  und der Parametervektor  $\vartheta$ , der die Parameter der reduzierten Form  $\gamma_k$ ,  $\mathbf{II}_k$  und  $\Sigma_k$  jeder Komponentenverteilung strukturiert, zu schätzen.

Zur Schätzung der Parameter schlagen Arminger et al. (1999) drei alternative Schätzstrategien vor. Es handelt sich dabei um ein zweistufiges Schätzverfahren (eine Kombination aus E(xpectation)-M(aximization)-Algorithmus und Minimum-Distanz-Schätzung) und zwei einstufige Schätzverfahren (einen direkten EM-Algorithmus und einen Gradienten EM-Algorithmus). Die Schätzalgorithmen sind in dem Computerprogramm MECOSA 3 (Ariminger/Wittenberg 1998) implementiert. Technische Einzelheiten zu den Schätzverfahren sowie

Ergebnisse von Simulationsstudien, die die Anwendbarkeit der drei Algorithmen demonstrieren, lassen sich in Arminger et al. (1999) finden.

In der ersten Stufe der ersten Schätzstrategie werden die mischenden Wahrscheinlichkeiten  $\pi_k$  sowie die Anzahl der Elemente in den einzelnen Komponenten  $n_k$ , die konditionalen Erwartungswerte  $\mu_{ik}$  als lineare Regressionsfunktionen der erklärenden Variablen  $\mathbf{x}_i$  also  $\gamma_k$ ,  $\mathbf{II}_k$  und die Kovarianzmatrizen  $\Sigma_k$  ohne Berücksichtigung der durch  $\vartheta$  aufgeprägten parametrischen Struktur (z. B. in Form eines LISREL-Modells) geschätzt. Dies erfolgt unter Verwendung eines EM-Algorithmus für Regressionsmodelle, der von Jones und McLachlan (1992) entwickelt wurde. Nach der Schätzung der Parameter wird im zweiten Schritt die asymptotische Kovarianzmatrix  $\Omega$  dieser unrestringierten Parameterschätzer geschätzt. Diese asymptotische Kovarianzmatrix ist zum einen zur Berechnung der Teststatistiken und Konfidenzintervalle erforderlich, zum anderen zur Schätzung der Strukturparameter  $\vartheta$  in der zweiten Stufe der Schätzung mit Hilfe der Minimum-Distanz-Schätzung.

Auf den ersten Blick erscheint es nahe liegend, nach der ersten Stufe zur Schätzung der Parameter  $\vartheta$  der strukturellen Form die Multiple-Gruppen-Option, die in Programmen zur Analyse von Mittelwert- und Kovarianzstrukturmodellen implementiert ist, zu verwenden. Dieses Verfahren weist jedoch eine Reihe von Nachteilen auf, so dass die Prozedur zur Schätzung von  $\vartheta$  und zum Testen von Restriktionen von  $\vartheta$  nur bedingt geeignet ist.

Der erste Nachteil ist die Annahme der un konditionalen gemeinsamen Normalverteilung aller Variablen, d. h. auch der Regressoren, sofern sie vorhanden sind. Wie Stein (1997) und Arminger et al. (1999) zeigen, führt die Annahme der un konditionalen gemeinsamen Normalverteilung der unabhängigen und abhängigen Variablen zu inkonsistenten Schätzern der Parameter der mischenden Verteilung  $\pi_k$  und der Parameter der einzelnen Komponenten der Mischverteilung  $\gamma_k$ ,  $\mathbf{II}_k$ ,  $\Sigma_k$  und damit zu inkonsistenten Schätzungen der in  $\vartheta$  enthaltenen Parameter der strukturellen Form, wenn diese Annahme verletzt ist. Dies ist z. B. bei der Verwendung von Dummy-Variablen als Regressoren gegeben. Desweiteren, wie Arminger et al. (1999) zeigen, sind die Schätzungen der Parameter  $\gamma_k$ ,  $\mathbf{II}_k$ ,  $\Sigma_k$  nicht voneinander statistisch unabhängig. Damit ist eine zentrale Annahme der Multiplen-Gruppen-Option verletzt, die besagt, dass die Gruppen voneinander statistisch unabhängig sind. Daher liefert die Multiple-Gruppen-Option von Programmen für Mit-

telwert- und Kovarianzstrukturen selbst bei korrekter Annahme der un konditionalen Normalverteilung aller Variablen inkonsistente Schätzer der Kovarianzmatrix der Parameter und damit falsche Standardfehler, wenn die Voraussetzung der Unabhängigkeit der Stichprobenelemente und damit der Mittelwertvektoren und Kovarianzmatrizen nicht erfüllt ist.

Die zweite Schätzstrategie ist die Verallgemeinerung des von Jedidi et al. (1997) für un konditionale Mischungen von LISREL-Modellen entwickelten direkten EM-Algorithmus, mit dem sowohl die Parameter des Mischverteilungsmodells als auch die Parameter des Mittelwert- und Kovarianzstrukturmodells in einer Stufe geschätzt werden. Dies geschieht in gleicher Weise wie in der ersten Stufe des zweistufigen Schätzverfahrens mit dem Unterschied, dass die Maximierung in dem M-Schritt des EM-Algorithmus unter Berücksichtigung der in  $\vartheta$  enthaltenen Strukturparameter und nicht der Parameter der ersten Stufe  $\gamma_k$ ,  $\mathbf{II}_k$  und  $\Sigma_k$  gebildet wird.

Als dritte Schätzstrategie ist der Gradienten EM-Algorithmus von Becker et al. (1997) zu nennen. Dieser Algorithmus wird in der gleichen Weise wie der direkte EM-Algorithmus durchgeführt mit dem Unterschied, dass die Iteration innerhalb des M-Schrittes des EM-Algorithmus nicht bis zur Konvergenz erfolgt, sondern bereits nach der ersten Iteration zum E-Schritt zurückkehrt. Der Vorteil des Gradienten EM-Algorithmus gegenüber dem direkten EM-Algorithmus ist die schnellere Durchführbarkeit.

Die Entscheidung über die Bevorzugung der ein- oder zweistufigen Schätzstrategie ist von der jeweiligen Problemstellung abhängig. In der Regel ist die Spezifikation von Mittelwert- und Kovarianzstrukturmodelle von unbekanntem Subpopulationen komplizierter als bei Subpopulationen, die von vornherein bekannt sind. In diesen Fällen eignet sich das zweistufige Verfahren, da zunächst der Forscher die unrestringierten Parameterschätzer der ersten Stufe  $\gamma_k$ ,  $\mathbf{II}_k$  und  $\Sigma_k$  des Schätzverfahrens inspizieren kann, um einen Einblick in die Struktur der bis dahin noch unbekanntem Komponenten zu erhalten. Dies ermöglicht ihm, Hypothesen über Beziehungen der Variablen in den einzelnen Komponenten zu formulieren. Ein weiterer Vorteil, der für die Verwendung des zweistufigen Verfahrens spricht, ist die schnellere Berechnung der Parameterschätzer, da im M-Schritt des EM-Algorithmus nur die gewichtete Regressionsanalyse und keine iterative Prozedur eingeht. Darüber hinaus ist als zusätzlicher Vorteil die einfache Implementation eines Testverfahrens zu nennen, mit dem die Anzahl der Komponenten getestet werden kann.

Ein Vorteil der einstufigen Schätzverfahren liegt darin, dass die Anpassung des Strukturmodells durch Zuweisung von Objekten zu anderen Gruppen verbessert werden kann. Dies ist bei der 2-Stufen-Prozedur nicht möglich, da die Gruppierung festgelegt ist, nachdem die Klassifikation von Objekten in der ersten Stufe unter Benutzung des unstrukturierten Modells vorgenommen wurde. Ein weiterer Vorteil des einstufigen Schätzverfahrens ist die direkte Schätzung der in  $\mathcal{D}$  enthaltenen Strukturparameter ohne Berechnung der asymptotischen Kovarianzmatrix der Parameter der ersten Stufe des zweistufigen Verfahrens. Aus diesen Gründen eignet sich dieses Verfahren immer dann, wenn eine Klassifikation der Subpopulationen aufgrund vorgegebener Mittelwert- und Kovarianzstrukturmodelle vorgenommen werden soll. Die Modelle können jedoch nur formuliert werden, wenn die Anzahl der Komponenten bekannt ist.

Als eine Konsequenz der Abwägung von Vor- und Nachteilen der Schätzverfahren ist es empfehlenswert, zunächst die erste Stufe des zweistufigen Schätzverfahrens zu verwenden, um die Anzahl der Komponenten zu testen und die geschätzten Parameter jeder Komponente ( $\gamma_k$ ,  $\Pi_k$  und  $\Sigma_k$ ) auf eine mögliche Parametrisierung zu inspizieren, anschließend zur Spezifikation und Modifikation der Mittelwert- und Kovarianzmodelle die zweite Stufe des zweistufigen Schätzverfahrens verwenden und das endgültige Modell mit Hilfe des Gradienten EM-Algorithmus zu schätzen und zu testen.

## 2.6 Test auf die Anzahl der Komponenten und Konstruktion von Anpassungsindizes

Im Regelfall ist die Anzahl der Komponenten einer Mischverteilung nicht a priori durch theoretische Annahmen festgelegt, sondern wird in der Mischverteilungsanalyse gemeinsam mit den mischenden Wahrscheinlichkeiten geschätzt. Daher ist die Anzahl der Komponenten mit einem statistischen Testverfahren zu überprüfen. Als Test auf die Anzahl  $K = 1, 2, \dots$  der Komponenten in einer finiten Mischung multivariater Normalverteilungen wird üblicherweise der Likelihood-Ratio-Test verwendet. Wie McLachlan/Basford (1988) darstellen, bringt die Anwendung der Likelihood-Ratio-Test-Statistik in Mischverteilungen erhebliche Probleme mit sich, weil die Likelihood-Ratio-Test-Statistik unter der Nullhypothese nicht einer zentralen  $\chi^2$ -Verteilung mit der Anzahl von Freiheitsgraden, die sich aus der Differenz der Anzahl der Parametern in den beiden Hypothesen ergibt, folgt.

Zur Lösung des Problems stehen zwei Möglichkeiten zur Verfügung. Als erste ist ein Verfahren von Wolfe (1971) zu nennen, der eine Modifikation der Zahl der Freiheitsgrade des Likelihood-Ratio-Tests vorschlägt, die geeignet erscheint, die Anzahl der Komponenten in einer Mischung zu testen. Da das Ergebnis von Wolfe (1971) nur auf Simulationsstudien basiert, ist der von ihm vorgeschlagene Test nur als ein Ad-hoc-Test brauchbar. Als Alternative zu der von Wolfe diskutierten Modifikation der Zahl der Freiheitsgrade des Likelihood-Ratio-Tests schlagen Aitkin et al. (1981) die Verwendung des Bootstrap-Ansatzes vor, mit dessen Hilfe die Verteilung der Likelihood-Ratio-Test-Statistik geschätzt werden kann. Arminger et al. (1999) sowie Stein (1997) verwenden zur Überprüfung der Anzahl der Komponenten in einem Mischverteilungsmodell den parametrischen Bootstrap, mit dem sie eine Schätzung der Verteilung der Likelihood-Ratio-Test-Statistik unter der jeweiligen Nullhypothese entwickeln.<sup>3</sup>

Genau wie in der herkömmlichen Mittelwert- und Kovarianzstrukturanalyse wird auch im Fall der Mischverteilungen als Prüfgröße für die Übereinstimmung zwischen Daten und Modell das minimierte Distanzmaß verwendet, das den unter dem Modell berechneten Mittelwertvektor bzw. die Kovarianzmatrix mit den aus den Daten berechneten empirischen Mittelwertvektor bzw. der Kovarianzmatrix vergleicht. Dieses Distanzmaß entspricht der im ADF-Verfahren der herkömmlichen Kovarianzstrukturanalyse berechneten  $\chi^2$ -Statistik und ist unter der Nullhypothese der Übereinstimmung der geschätzten Strukturparametern des Mittelwert- und Kovarianzstrukturmodells und den Parametern des konditionalen Mischverteilungsmodells ohne Restriktionen wiederum zentral  $\chi^2$  verteilt. Der Unterschied besteht darin, daß im Fall der Mischverteilungen sowohl die konditionale Spezifikation als auch die Komponenten der Mischverteilung zu berücksichtigen sind. Zusätzlich werden zur Beurteilung der Anpassungsgüte die von Bentler/Bonett (1980) entwickelten inkrementellen Anpassungsindizes NFI (normed fit index), NNFI (nonnormed fit index) und der von Bentler (1988) entwickelte CFI (comparative fit index) sowie der von Steiger/Lind (1980) entwickelte Lack-of-Fit-Index RMSEA (root mean square error of approximation) adaptiert und weiter entwickelt.<sup>4</sup>

<sup>3</sup> Zu den technischen Einzelheiten siehe Arminger et al. (1999) sowie Stein (1997).

<sup>4</sup> Zu den technischen Einzelheiten siehe Stein (1997).

### 3. Anwendung des Mischverteilungsansatzes zur Aufdeckung unbeobachteter Heterogenität in der Analyse von Lebenszufriedenheit

Zur Illustration des Modells finiter Mischungen konditionaler Mittelwert- und Kovarianzstrukturmodelle und der Schätzprozeduren wird ein Modell zur Analyse der Lebenszufriedenheit aus dem Datensatz des Wohlfahrtssurveys 1993 vorgestellt.<sup>5</sup> In Anlehnung an bisherige Forschungen zur Lebensqualität, in der allgemein davon ausgegangen wird, dass die Zufriedenheit mit dem Leben „alles in allem“ eine zusammenfassende Bewertung der Situation in den einzelnen Lebensbereichen ist (vgl. Glatzer/Zapf 1984; Veenhoven 1993; Berger-Schmitt 1994; Spellerberg 1996) wird im Folgenden die Hypothese geprüft, dass die Zufriedenheit mit dem Leben im Allgemeinen das Ergebnis einer kognitiven Gesamtbilanz der Zufriedenheiten in einzelnen Lebensbereichen ist. Entsprechend wird der Einfluss der Zufriedenheit in ausgewählten Lebensbereichen auf die allgemeine Lebenszufriedenheit untersucht.

Derartige Analysen unter Verwendung von LISREL-Modellen für die Gesamtpopulation lassen sich z. B. bei Hampel (1985) finden. Im Unterschied dazu wird in dieser Untersuchung vermutet, dass die Bevölkerung in heterogene Subpopulationen zerfällt, die unterschiedliche Zusammenhänge in den Zufriedenheiten mit einzelnen Lebensbereichen und der Zufriedenheit mit dem Leben im Allgemeinen aufweisen können.

Zur Erklärung der Zufriedenheit in den einzelnen Lebensbereichen und der allgemeinen Lebenszufriedenheit werden in der Lebensqualitätsforschung verschiedene Determinanten herangezogen. Dazu zählen Statusvariablen wie die berufliche Position und Bildung, demographische Merkmale wie Alter, Geschlecht, Familienstand und das Vorhandensein von Kindern sowie regionale Merkmale. Weitere Einflussgrößen sind die objektiven Lebensbedin-

gungen sowie soziale Vergleichsprozesse, Erwartungshaltungen und Einstellungen, die die Wahrnehmung und Beurteilung der Lebensverhältnisse beeinflussen (vgl. Glatzer 1984; Zapf/Diewald 1984; Spellerberg 1996). Darüber hinaus lässt sich Zufriedenheit auch durch persönliche Merkmale wie Gesundheit oder psychische Stabilität, Lebensereignisse sowie gesellschaftliche Intergration und starke persönlichen Beziehungen erklären (vgl. Veenhoven 1994).

Ein gängiges Verfahren zur Analyse des Einflusses der Determinanten auf einzelne Indikatoren zur Messung der Zufriedenheit in den Lebensbereichen ist die multiple Klassifikationsanalyse (MCA). Anwendungsbeispiele hierfür lassen sich in Diewald/Zapf (1984), Glatzer (1984), Krebs/Schmidt (1996) finden. Für die Analyse der Lebenszufriedenheit lässt sich dieses Verfahren nur bedingt einsetzen. Kritikpunkte sind: Erstens ist nur die Verwendung einer einzelnen abhängigen Variable möglich. Wird der Einfluss auf mehrere abhängige Variablen unter Verwendung der MCA analysiert, müssen diese zuvor zusammengefasst werden. Ein Beispiel hierfür lässt sich in Spellerberg (1996) finden, die den Einfluss sozialer Merkmale und Lebensstile auf die Zufriedenheit mit verschiedenen Lebensbereichen analysiert, wobei die Lebensbereiche zuvor durch Indizes aus einzelnen Items gebildet wurden. Zweitens können nur nominalskalierte Variablen als erklärende Variablen aufgenommen werden. Drittens können keine Messmodelle in Verbindung mit einem Kausalmodell formuliert werden. Es muss daher mit Faktorwerten oder Indizes gearbeitet werden.

Im Unterschied dazu können im vorgeschlagenen Mischverteilungsansatz mit LISREL-Modellen sowohl metrische als auch nichtmetrische erklärende Variablen aufgenommen werden. Darüber hinaus kann die gesamte Komplexität der Möglichkeiten der Mittelwert- und Kovarianzstrukturmodelle ausgeschöpft werden, wie z. B. die Berücksichtigung mehrerer abhängiger Variablen, eine Beziehungsstruktur dieser abhängigen Variablen untereinander und die Verwendung von Messmodellen.

Ohne Berücksichtigung des Mischverteilungsansatzes werden LISREL-Modelle zur gleichzeitigen Berücksichtigung von Messmodellen für die Lebenszufriedenheit, Zusammenhängen zwischen Zufriedenheiten in Lebensbereichen und der Analyse von Effekten von erklärenden Variablen auf diese Zufriedenheiten von Saris/Scherpenzeel (1995) verwendet. Allerdings behandeln Saris/Scherpenzeel (1995) in dieser Arbeit vor allem methodische Fra-

<sup>5</sup> Der Wohlfahrtssurvey ist eine repräsentative Befragung, die darauf ausgerichtet ist, Veränderungen der Lebensqualität in Deutschland zu beobachten. Die Datenerhebung wurde von Infratest, München, im Zeitraum von Februar bis April 1993 durchgeführt. Die Daten wurden durch eine mündliche Befragung mit standardisiertem Fragebogen an 3.062 auf der Basis des ADM-Mastersamples repräsentativ ausgewählten Personen erhoben. 2.046 der Befragten leben in West-Deutschland und 1.016 in Ost-Deutschland. Die der Stichprobe zugrunde gelegte Grundgesamtheit ist definiert als in Privathaushalten lebende deutsche Personen, die zum Zeitpunkt der Erhebung das achtzehnte Lebensjahr vollendet haben.



gen der komparativen internationalen Forschung zur Lebensqualität bezüglich der Vergleichbarkeit und der Reliabilität von Skalen.

Das Problem der Heterogenität, also der Zusammensetzung einer Population aus unterschiedlichen Subpopulationen wird in der Forschung zur Lebensqualität in zwei Arten aufgegriffen. Spellerberg et al. (1992) untersuchen z.B. die nach einem bekannten Kriterium getrennten Gruppen der Ost- und Westdeutschen in Bezug auf unterschiedliche Zusammenhänge zwischen einzelnen Bereichszufriedenheiten und der allgemeinen Lebenszufriedenheit unter Verwendung der multiplen Regressionsanalyse für beide Gruppen. Hingegen führt Spellerberg (1996) in ihrer Analyse des Zusammenhangs von objektiver Lebenslage und subjektivem Wohlbefinden den Lebensstil als intervenierende Variable ein, indem der Einfluss der Zufriedenheit in verschiedenen Lebensbereichen auf die allgemeine Lebenszufriedenheit nach Lebensstilgruppen differenziert wird. Diese Lebensstilgruppen entsprechen unbekanntem Subpopulationen und bewirken damit eine unbeobachtete Heterogenität der analysierten Stichprobe. Zur Aufdeckung der Lebensstilgruppen und Analyse der Zusammenhänge innerhalb der Lebensstilgruppen verwendet Spellerberg (1996) eine Folge statistischer Verfahren. Im ersten Schritt werden Lebensstilgruppen mit Hilfe der Clusteranalyse extrahiert. Im zweiten Schritt werden die Zusammenhänge von Lebensstilen und sozialer Lage mit bivariaten Zusammenhangsmaße dargestellt. Im dritten Schritt wird mit Hilfe multipler linearer Regressionsanalysen überprüft, inwieweit, nach Lebensstilen differenziert, die Zufriedenheit in verschiedenen Lebensbereichen das allgemeine Wohlbefinden erklären kann. Diese Vorgehensweise enthält die in Abschnitt 1 bezüglich der Verwendung der Clusteranalyse und der Multiplen-Gruppen-Option dargestellten Probleme. Sie kann daher durch die Anwendung des Mischverteilungsansatzes mit LISREL-Modellen unmittelbar verbessert werden. Insbesondere können auch latente Variablen direkt im Modell eingesetzt werden.

### 3.1 Auswahl und Beschreibung der Modellvariablen

Zur Identifikation homogener Subpopulationen und Spezifikation von Mittelwert- und Kovarianzstrukturmodellen für jede Komponente werden aus dem Wohlfahrtssurvey Indikatoren zur Messung der Zufriedenheit in verschiedenen Lebensberei-

chen und der Zufriedenheit mit dem Leben im Allgemeinen unter Verwendung einer explorativen Faktorenanalyse mit orthogonaler Rotation ausgewählt. Im Einzelnen handelt es sich bei den Items zur Messung der Zufriedenheiten mit verschiedenen Lebensbereichen um die Fragen: *Wie zufrieden sind Sie mit ... Ihrer Wohnung? (WOHN) ... der Gegend, in der Sie wohnen? (WOHNG) ... der öffentlichen Sicherheit und Bekämpfung der Kriminalität? (SICHER) ... mit dem, was man das „Netz der sozialen Sicherung“ nennt? (SOZNET) ... dem Einkommen Ihres Haushaltes? (HHEIN)... Ihrem Lebensstandard? (STAND)... den Möglichkeiten, sich politisch zu betätigen? (POLIT) ... den demokratischen Einrichtungen in unserem Land? (DEMO)*. Die Indikatoren wurden mit Hilfe einer elfstufigen Skala von 0 (ganz und gar unzufrieden) bis 10 (ganz und gar zufrieden) erhoben.

Als Indikatoren zur Messung der allgemeinen Lebenszufriedenheit werden aus dem Datensatz zwei Fragen ausgewählt. Dies sind zum einen die Frage nach der Zufriedenheit mit den gegenwärtigen Leben (*Was meinen Sie, wie zufrieden sind Sie gegenwärtig – alles in allem – heute mit Ihrem Leben? (LEBEN)*). Der Indikator wurde ebenfalls mit Hilfe der elfstufigen Skala gemessen. Der zweite Indikator ist die Beurteilung der gegenwärtigen Lebensbedingungen zum Zeitpunkt der Befragung (*Wo auf der ersten Leiter würden Sie Ihre gegenwärtigen Lebensbedingungen – alles in allem – einordnen? (BEDING)*), die ebenfalls mit Hilfe einer elfstufigen Skala von 0 (schlechtesten vorstellbaren Lebensbedingungen) bis 10 (besten vorstellbaren Lebensbedingungen) gemessen wurde.

Die Indikatoren zur Messung der Zufriedenheit in den einzelnen Lebensbereichen sowie die Indikatoren zur Messung der allgemeinen Lebenszufriedenheit fungieren als abhängige Variablen in dem Mischverteilungsmodell. Ausgehend von der Überlegung, dass sozialstrukturelle Merkmale möglicherweise die Zufriedenheit in verschiedenen Lebensbereichen und die allgemeine Lebenszufriedenheit beeinflussen, werden aus dem Datensatz vier sozialstrukturelle Variablen ausgewählt und in dem Modell als erklärende Variablen berücksichtigt. Es handelt sich um die dichotome Variable REGION mit den Ausprägungen 0 für *Alte Bundesländer* und 1 für *Neue Bundesländer*, die metrische Variable ALTER IN JAHREN, die dichotome Variable GESCHLECHT mit den Ausprägungen 0 für *weiblich* und 1 für *männlich* und die dichotomisierte Variable SCHULBILDUNG mit den Ausprägungen 0 für *weniger als Fachhochschulreife* und 1 für *mindestens Fachhochschulreife*.

### 3.2 Mischverteilungsanalyse zur Extraktion homogener Typen

Im Folgenden wird der Frage nachgegangen, ob die Stichprobe bezüglich der Struktur der Beziehungen zwischen den Variablen untereinander aus einer homogenen Population stammt oder ob eine Mischung von mindestens zwei unterschiedlichen Subpopulationen vorliegt. Zur Beantwortung dieser Frage müssen folgende statistischen Aufgaben gelöst werden: (a) die Extraktion der verschiedenen Populationen, also der Komponenten der Mischverteilung, (b) die Schätzung der Anzahl der Komponenten, (c) die Schätzung des Anteils der einzelnen Komponenten an der gesamten Verteilung und damit die Stichprobengröße innerhalb der Komponenten, (d) die Schätzung der Regressionskonstanten, Regressionskoeffizienten und Kovarianzmatrix der Residuen sowie die un konditionalen Mittelwerte zur Charakterisierung der Komponenten und (e) die Schätzung der Parameter der konditionalen Mittelwert- und Kovarianzstrukturmodelle.

Die entsprechenden Berechnungen erfolgen mit dem Programm MECOSA 3 (Armingier/Wittenberg 1998). Als Schätzstrategie wird zunächst das zweistufige Schätzverfahren gewählt. In der ersten Stufe werden unter Verwendung des EM-Algorithmus die mischenden Wahrscheinlichkeiten, die Anzahl der Elemente in den einzelnen Komponenten sowie die Regressionskonstanten, Regressionskoeffizienten und Kovarianzmatrix der Residuen für jede Komponente der Mischung geschätzt. Im Anschluss daran wird die asymptotischen Kovarianzmatrix dieser Parameter geschätzt. Durch Implementation eines Ad-hoc Tests wird getestet, aus wie vielen Subpopulationen sich die Grundgesamtheit zusammensetzt.

Nachdem die Ergebnisse vorliegen, wird durch Inspektion der geschätzten Regressionskonstanten, Regressionskoeffizienten und residueller Kovarianzmatrizen ein Einblick in die Struktur der extrahierten Komponenten verschafft. Dies ermöglicht die Formulierung von Hypothesen über den Einfluss der sozialstrukturellen Merkmale auf die Zufriedenheit in den einzelnen Lebensbereichen, die Be-

ziehung der Zufriedenheit in den einzelnen Lebensbereichen untereinander und den Einfluss der Zufriedenheiten in den Lebensbereichen auf die allgemeine Lebenszufriedenheit in den einzelnen Komponenten. Die Hypothesen werden in Form von konditionalen Mittelwert- und Kovarianzstrukturmodellen vom Typ LISREL spezifiziert und die Parameter in der zweiten Stufe des zweistufigen Schätzverfahrens mit Hilfe der Minimum-Distanz-Methode unter Verwendung der asymptotischen Kovarianzmatrix als Gewichtsmatrix aus den unstringierten Parameterschätzern der ersten Stufe geschätzt. Zur Beurteilung der Modellanpassung werden die auf der Minimum-Distanz-Schätzung basierende  $\chi^2$ -Test-Statistik sowie die daraus berechneten Anpassungsindizes NFI, NNFI und CFI und der Lack-of-Fit-Index RMSEA verwendet. Die Parameter des endgültigen Modells werden mit Hilfe des einstufigen Schätzverfahrens unter Verwendung des Gradienten EM-Algorithmus geschätzt.

### 3.3 Anzahl und Größe der Komponenten

Tabelle 1 zeigt die Loglikelihoodfunktion, die Likelihood Ratio Test Statistik und implizites Signifikanzniveau  $\alpha$  der Ad-hoc-Tests für  $H_0 : K = 1$  gegen  $H_1 : K = 2$ ,  $H_0 : K = 2$  gegen  $H_1 : K = 3$  und  $H_0 : K = 3$  gegen  $H_1 : K = 4$ .

Wie aus Tabelle 1 ersichtlich wird, lässt sich in der vorliegenden Stichprobe die Hypothese der Homogenität der Grundgesamtheit nicht halten. Da das implizite Testniveau  $\alpha = 0.999$  des Tests  $H_0 : K = 3$  gegen  $H_1 : K = 4$  größer ist als der kritische Wert 0.005, wird die Nullhypothese nicht abgelehnt. Es wird daher im Folgenden von drei Komponenten der Grundgesamtheit ausgegangen.

Die Anteile (mischenden Wahrscheinlichkeiten) der Komponenten werden als  $\hat{\pi}_1 = 0.2999$ ,  $\hat{\pi}_2 = 0.4158$  und  $\hat{\pi}_3 = 0.2844$  geschätzt. Damit umfasst die erste Komponente 564 Angehörige, die zweite Komponente 782 und die dritte Komponente 536 Angehörige der Stichprobe.

Anzahl $K$ der Komponenten	Loglikelihood-Funktion für $K$	Likelihood-Ratio-Test-Statistik	df	$\alpha$
1	-37942.846			
2	-36976.031	1933.629	210	0.000
3	-33321.743	7308.575	315	0.000
4	-33157.736	328.014	420	0.999

**Tabelle 1** Loglikelihoodfunktionen und Likelihood-Ratio-Statistiken

**Tabelle 2** Mittelwerte aller Indikatoren in den drei Komponenten

Komp	WOHN	WOHNG	SICHER	SOZNET	HHEIN	STAND	POLIT	DEMO	LEBEN	BEDING
1	7.622	7.540	4.575	6.396	7.310	7.524	5.456	5.489	7.856	7.479
2	8.580	8.580	5.069	6.409	7.047	7.373	5.269	5.576	7.835	7.221
3	6.373	6.523	3.742	5.012	4.962	5.879	4.286	4.304	6.765	5.910

### 3.4 Inhaltliche Charakterisierung der Lebenszufriedenheitstypen

Im Folgenden werden die mit Hilfe des Mischverteilungsverfahrens extrahierten Komponenten durch Beschreibung der Unterschiede und Ähnlichkeiten in den Indikatoren zur Messung der Zufriedenheiten mit einzelnen Lebensbereichen und dem Leben im Allgemeinen inhaltlich charakterisiert und als Typen mit unterschiedlichem Grad an Lebenszufriedenheit identifiziert. Zu diesem Zweck werden die un konditionalen Mittelwerte der Indikatoren verwendet. Ihre Aufstellung ergibt das in Tabelle 2 dargestellte Profil für die einzelnen Komponenten.<sup>6</sup>

Bei einem Vergleich der Komponenten fällt zunächst die dritte Komponente auf, die in allen Indikatoren deutlich niedrigere Mittelwerte aufweist. Dies bedeutet, dass in der dritten Komponente die Zufriedenheit in ausgewählten Lebensbereichen und dem Leben im Allgemeinen im Durchschnitt niedriger als in den beiden anderen Komponenten ist. Die ersten beiden Komponenten sind sich in den Mittelwerten sehr ähnlich. Eine Ausnahme sind die Indikatoren „Zufriedenheit mit der eigenen Wohnung“ (WOHN) und „Zufriedenheit mit der Wohngegend“ (WOHNG), in denen die zweite Komponente deutlich höhere Werte aufweist. Die Mittelwerte sind nicht nur höher als in den anderen Komponenten, sondern darüber hinaus extrem hoch. Die erste Komponente weist im Vergleich zu den anderen Komponenten die höchsten Mittelwerte in den Indikatoren zur Messung der Zufriedenheit mit der materiellen Situation „Zufriedenheit mit dem Haushaltseinkommen“ (HHEIN) und „Zufriedenheit mit dem Lebensstandard“ (STAND) auf. Aufgrund des Mittelwertvergleichs können die drei Komponenten inhaltlich charakterisiert werden als: Typ 1: Hohe Zufriedenheit mit der ma-

teriellen Situation, Typ 2: Hohe Zufriedenheit mit den Wohnbedingungen und Typ 3: Geringste Zufriedenheit in allen Bereichen des Lebens und dem Leben im Allgemeinen.

In einem weiteren Schritt werden die drei Zufriedenheitstypen in ihren Unterschieden bezüglich sozialstruktureller Merkmale anhand der Mittelwerte der als unabhängigen Variablen in die Mischverteilungsanalyse eingegangenen Variablen beschrieben. In der Variablen REGION sind große Differenzen zwischen den drei Typen festzustellen. Der Anteil der Personen, die in Ostdeutschland leben, ist in dem dritten Typus mit 43,1 % wesentlich höher als in den beiden anderen Typen. Daraus folgt, dass die Gruppe der Personen, die mit dem Leben unzufriedener sind, sehr stark durch Bewohner der neuen Bundesländer geprägt ist. Am geringsten ist der Anteil der Ostdeutschen in dem zweiten Typus mit 23,66 %, woraus zu schließen ist, dass der Typus des mit den Wohnbedingungen hoch zufriedenen Menschen sehr stark durch Westdeutsche gebildet wird. Der erste Typus nimmt mit einem Anteil von 36 % Ostdeutschen eine mittlere Stellung ein. In den soziodemographischen Merkmalen GESCHLECHT, SCHULBILDUNG und ALTER IN JAHREN unterscheiden sich die drei Typen nur geringfügig. So ist der Anteil der Männer mit 50,89 % in dem durch die „hohe Zufriedenheit mit der materiellen Situation“ charakterisierten Typus im Vergleich zu den anderen etwas stärker vertreten und im Typus mit der „geringsten Zufriedenheit“ mit 45,71 % am geringsten. Personen mit höherer Schulbildung sind im ersten Typus im Durchschnitt überrepräsentiert: Der Anteil von Personen, die mindestens Fachhochschulreife aufweisen, beträgt 26,42 %. Am geringsten ist der Anteil der Personen mit höherer Bildung im zweiten Typus mit 22,38 %. Das Durchschnittsalter ist im zweiten Typus mit ungefähr 42 Jahren am höchsten und im dritten Typus mit 39 Jahren am niedrigsten. Aufgrund der Ergebnisse ist zu schließen: Die Verteilung der sozialstrukturellen Merkmale auf die einzelnen Subpopulationen ergibt eine proportionale Ungleichverteilung der Variablen REGION in die Richtung, dass sich unter den „weniger Zufriedenen“ überproportional viele Ostdeutsche und un-

<sup>6</sup> Die Charakterisierung der einzelnen Subpopulationen durch die Mittelwerte der abhängigen und erklärenden Variablen ist nicht die einzige Möglichkeit, die Subpopulationen zu charakterisieren. Genauso wie in der Clusteranalyse besteht die Möglichkeit, einzelne Personen den Komponenten zuzuordnen und dann zu untersuchen, ob sich die Komponenten bezüglich weitere, bisher nicht betrachteter Merkmale, unterscheiden.

terproportional wenige unter den „mit den Wohnbedingungen hoch Zufriedenen“ befinden.<sup>7</sup>

In einem weiteren Schritt werden die drei Zufriedenheitstypen in ihrer unterschiedlichen Struktur der Beziehungen zwischen den Zufriedenheiten in den ausgewählten Lebensbereichen und der allgemeinen Lebenszufriedenheit sowie des Einflusses der sozialstrukturellen Merkmale auf die Zufriedenheit in einzelne Lebensbereiche und auf die allgemeine Lebenszufriedenheit charakterisiert. Zur Darstellung dienen die Parameterschätzungen der konditionalen LISREL-Modelle.

Als Ergebnis kann für jede Komponente ein Modell mit einer eigenen speziellen Struktur sowohl im Messmodell als auch im Strukturmodell entwickelt werden, das als Gesamtmodell eine hohe Anpassung an die Parameter der ersten Stufe zeigt. Die Minimum-Distanz-Schätzung ergibt eine  $\chi^2$ -Statistik von 348.09 mit 151 Freiheitsgraden.<sup>8</sup> Der Lack-of-Fit-Index RMSEA liegt mit 0.046 unter dem kritischen Wert von 0.05, so dass das Modell gemäß diesem Kriterium eine sehr gute Anpassung („close fit“) im Sinne von Browne/Cudeck (1993) aufweist. Die inkrementellen Anpassungsindizes NFI=0.98, NNFI=0.98 und CFI=0.96 überschreiten den kritischen Wert von 0.9 in erheblichen Maße, so dass insgesamt das Modell auch im Sinne von Bentler (1992) als sehr gut angepasst gelten kann.

In den Abbildungen 1 bis 3 werden die Einflüsse der sozialstrukturellen Merkmale auf die Zufriedenheitskonstrukte sowie die Beziehungen zwi-

<sup>7</sup> Zu dem Ergebnis von Zufriedenheitsdifferenzen zwischen Ost- und Westdeutschen gelangten bereits Landua et al. (1991) bei einer Analyse des Wohlfahrtsurveys 1990. Auch Habich et al. (1999) kommen bei einer Analyse des Wohlfahrtsurveys 1998 zu diesem Ergebnis, obwohl insgesamt eine Annäherung stattgefunden hat. Der Unterschied zu diesen Ost-West-Vergleichen ist darin zu sehen, dass in diesem Fall die latenten Gruppen bezüglich der Verteilung der Variablen REGION beschrieben werden und damit eine andere Betrachtungsweise gewählt wird. Das Ergebnis, dass sich unter den „mit den Wohnbedingungen hoch Zufriedenen“ auch eine beträchtliche Anzahl Ostdeutscher befindet, stimmt prinzipiell mit den Ergebnissen von Müller (1994) überein, der aufgrund einer eigenen Erhebung von 1991 einen hohen Grad der Zufriedenheit mit der Wohnung mit der Wohnung und einen noch höheren Grad der Zufriedenheit mit der Wohnungsgegend der ostdeutschen Bevölkerung feststellt.

<sup>8</sup> Daher wird die Nullhypothese, dass die Parameter, die in der ersten Stufe durch das gewählte Modell generiert wurden, aufgrund der Daten dieser Stichprobe auf dem Testniveau von  $\alpha = 0.05$  zurückgewiesen. Bei der Interpretation der  $\chi^2$ -Statistik ist jedoch zu beachten, dass sie in hohen Maße von der Größe der Stichprobe abhängt.

schen den Konstrukten untereinander (Strukturmodell) und die Wirkung der latenten Konstrukte auf die manifesten Variablen (Messmodell) für die einzelnen Komponenten aufgezeigt. Die latenten Variablen werden durch Ellipsen und die beobachtbaren Variablen durch Rechtecke dargestellt. Dabei werden nur die signifikanten (auf dem Testniveau von  $\alpha = 0.05$  von Null verschiedenen) Koeffizienten dargestellt. Auf die Aufnahme der Residualkomponenten wird aus Gründen der Übersichtlichkeit verzichtet. Die geschätzten Parameterwerte des Modells werden in den Tabellen 3 bis 8 aufgelistet.<sup>9</sup>

### Ergebnisse für die erste Komponente

Abbildung 1 zeigt das Modell für die erste Komponente, die inhaltlich als der Typus mit hoher Zufriedenheit mit der materiellen Situation charakterisiert werden konnte.

Das Modell besteht aus den Messmodellen für *erstens* das Konstrukt „Zufriedenheit mit der materiellen Situation“ (MATERIELL), das durch die Indikatoren „Zufriedenheit mit dem Haushaltseinkommen“ (HHEIN) und „Zufriedenheit mit dem

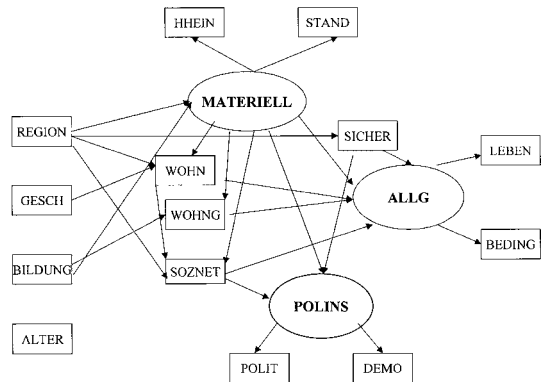


Abb. 1 Signifikante Wirkungszusammenhänge in der ersten Komponente

<sup>9</sup> Es ist zu beachten, dass die Koeffizienten nicht standardisiert sind, da erstens als erklärende Variablen auch Dummyvariablen verwendet werden, die üblicherweise als Verschiebung der Regressionskonstante interpretiert werden, und zweitens ungeklärt ist, auf welche Population sich die Standardisierungen beziehen sollen. Standardisiert man auf der Basis der Subpopulationen, sind die Teilpopulationen nicht mehr vergleichbar, da die Standardabweichungen und damit die verwendeten Skalen populationspezifisch sind. Standardisiert man hingegen auf der Basis der Gesamtpopulation, sind die Koeffizienten nicht mehr normiert.

**Tabelle 3** Messmodell für die erste Komponente

Indikatoren	NU	LAMBDA			THETA				
		WOHN	WOHNG	SICHER	SOZNET	POLINS	MATERIELL	ALLG	
WOHN	7.772 (46.670)	1.000 (-)	-	-	-	-	-	-	-
WOHNG	7.651 (39.851)	-	1.000 (-)	-	-	-	-	-	-
SICHER	4.865 (33.673)	-	-	1.000 (-)	-	-	-	-	-
SOZNET	6.677 (52.787)	-	-	-	1.000 (-)	-	-	-	-
HHEIN	7.656 (70.424)	-	-	-	-	-	1.000 (-)	-	0.521 (8.313)
STAND	7.918 (68.961)	-	-	-	-	-	1.107 (23.057)	-	0.060 (1.043)
POLIT	5.800 (35.779)	-	-	-	-	1.000 (-)	-	-	-
DEMO	5.724 (46.701)	-	-	-	-	0.577 (18.789)	-	-	2.264 (13.655)
LEBEN	8.468 (57.524)	-	-	-	-	-	-	1.000 (-)	0.371 (6.778)
BEDING	8.071 (52.857)	-	-	-	-	-	-	1.016 (13.629)	0.738 (10.263)

Anmerkung: Die in Klammern dargestellten Werte sind die z-Werte für jeden Koeffizienten, die mit Hilfe des MECOSA 3-Programms berechnet wurden.

Lebensstandard“ (STAND) gemessen wird; *zweitens* die „Zufriedenheit mit den politischen Institutionen“ (POLINS), operationalisiert durch die Items „Zufriedenheit mit den Möglichkeiten, sich politisch zu betätigen“ (POLIT) und „Zufriedenheit mit den demokratischen Einrichtungen in unserem Land“ (DEMO) und *drittens* die „Zufriedenheit mit dem Leben im Allgemeinen“ (ALLG) mit den Indikatoren „Zufriedenheit mit dem gegenwärtigen Leben“ (LEBEN) und „Beurteilung der gegenwärtigen Lebensbedingungen“ (BEDING).

Die Items „Zufriedenheit mit dem Netz der sozialen Sicherung“ (SOZNET) und „Zufriedenheit mit der öffentlichen Sicherheit“ (SICHER) konnten nicht einem gemeinsamen Faktor zugeordnet werden, sondern sind als spezifische Faktoren anzusehen. Da diese Faktoren mit den sie messenden Indikatoren identisch sind, werden die Faktorladungen im Messmodell auf den Wert 1 fixiert und bleiben in der graphischen Darstellung unberücksichtigt. Das gleiche gilt für die Indikatoren „Zufriedenheit mit der Wohnung“ (WOHN) und „Zufriedenheit mit der Wohngegend“ (WOHNG).

Tabelle 3 zeigt die geschätzten Regressionskonstanten<sup>10</sup>, die Faktorladungen und die Varianzen der Messfehler<sup>11</sup> im Messmodell. Wie bereits oben erwähnt, sind die ersten vier Faktoren (WOHN, WOHNNG, SICHER und SOZNET) mit den gleichnamigen manifesten Variablen identisch, sodass keine weitere Interpretation erforderlich ist. Der In-

<sup>10</sup> Die Regressionskonstanten der Indikatoren im Messmodell dienen nur zur Anpassung an die Regressionskonstanten der Parameter, die in der ersten Stufe der zweistufigen Schätzprozedur erzielt wurden, da für finite Mischungen generell konditionale Mittelwert- und Kovarianzstrukturmodelle spezifiziert werden müssen. Andernfalls, d.h. ohne Kenntnisse der Mittelwerte in den einzelnen Komponenten lassen sich die Kovarianzmatrizen der einzelnen Komponenten nicht berechnen. Sie sind daher nur der Vollständigkeit wegen aufgeführt und nicht weiter zu interpretieren.

<sup>11</sup> Die Varianzen der Messfehler können dazu verwendet werden, um den durch die latenten Variablen erklärten Anteil der Varianzen der manifesten Variablen, also die Kommunalität, zu schätzen. Auch sie sind nur der Vollständigkeit wegen aufgeführt und nicht weiter zu interpretieren.

Indikator „Zufriedenheit mit seinem Haushaltseinkommen“ (HHEIN) ist die Referenzvariable für den Faktor „Zufriedenheit mit seiner materiellen Situation“ (MATERIELL). Daher ist die entsprechende Faktorladung auf 1 gesetzt. Die Faktorladung des Indikators „Zufriedenheit mit seinem Lebensstandard“ (STAND) auf den Faktor MATERIELL ( $\lambda = 1.107$ ) weist darauf hin, dass in dieser Komponente die Zufriedenheit mit den materiellen Bedingungen geringfügig stärker als mit der Zufriedenheit mit dem Haushaltseinkommen assoziiert wird. Die „Zufriedenheit mit den politischen Institutionen“ (POLINS) ist in der ersten Komponente stärker mit der „Zufriedenheit mit den Möglichkeiten, sich politisch zu betätigen“ (POLIT) (Referenzvariable für POLINS und daher auf 1 fixiert) als mit der „Zufriedenheit mit den demokratischen Einrichtungen“ (DEMO) assoziiert ( $\lambda = 0.577$ ). Der Indikator „Zufriedenheit mit seinem gegenwärtigen Leben“ (LEBEN) ist wie in allen anderen Komponenten die Referenzvariable für den Faktor „Zufriedenheit mit seinem Leben im Allgemeinen“ (ALLG). Damit wird sichergestellt, dass die Bedeutung der allgemeinen Lebenszufriedenheit in allen Komponenten am gleichen Item festgemacht wird. Der Indikator „Beurteilung der gegenwärtigen Lebensbedingungen“ (BEDING) lädt fast gleich stark auf den Faktor „Zufriedenheit mit dem Leben im Allgemeinen“ (ALLG) ( $\lambda = 1.016$ ), was ein Hinweis darauf ist, dass in dieser Komponente die allge-

meine Lebenszufriedenheit etwa gleich stark durch die Zufriedenheit mit dem Leben und mit der Beurteilung der Lebensbedingungen assoziiert ist.

Die geschätzten Parameter des Strukturgleichungsmodells für die erste Komponente sind in den Tabellen 4a und 4b dargestellt. Innerhalb der endogenen Variablen stellt die Zufriedenheit mit der materiellen Situation (MATERIELL) eine dominante Variable dar, die einen positiven Effekt auf fast alle anderen abhängigen Variablen ausübt. Sie beeinflusst sehr stark die Zufriedenheit mit der eigenen Wohnung (WOHN) ( $\beta = 0.466$ ). Je höher die Zufriedenheit mit der materiellen Situation in diesem Typus ausgeprägt ist, desto zufriedener ist er mit seiner eigenen Wohnung. Deutlich schwächer ist dagegen der Einfluss der Zufriedenheit mit der eigenen materiellen Situation auf die Zufriedenheit mit der Wohngegend (WOHNG) in diesem Typus. Zwar steigt auch hier die Zufriedenheit mit der Wohngegend bei wachsender Zufriedenheit mit der materiellen Situation an ( $\beta = 0.193$ ), doch ist dieser Einfluss wesentlich geringer als der vorhergehende. Ferner ist der Einfluss der Zufriedenheit mit der materiellen Situation auf die Zufriedenheit mit dem Netz der sozialen Sicherung (SOZNET) ( $\beta = 0.392$ ) und den politischen Institutionen (POLINS) ( $\beta = 0.412$ ) in diesem Typus erwähnenswert. Je höher die Zufriedenheit mit der eigenen materiellen Situation ist, desto stärker sind die Angehörigen

**Tabelle 4a** Strukturgleichungsmodell für die erste Komponente: Einfluss der Faktoren untereinander

Faktoren	BETA							PSI
	WOHN	WOHNG	SICHER	SOZNET	POLINS	MATERIELL	ALLG	
WOHN	-	-	-	-	-	0.466 (6.135)	-	2.73 (13.51)
WOHNG	-	-	-	-	-	0.193 (2.392)	-	4.02 (13.44)
SICHER	-	-	-	-	-	-	-	4.51 (11.04)
SOZNET	0.221 (4.591)	-	-	-	-	0.392 (4.832)	-	2.84 (13.54)
POLINS	-	-	0.129 (2.633)	0.135 (1.944)	-	0.412 (2.329)	-	4.36 (13.92)
MATERIELL	-	-	-	-	-	-	-	1.38 (9.54)
ALLG	0.075 (3.426)	0.051 (2.864)	0.058 (3.422)	0.086 (4.046)	-	0.462 (12.437)	-	0.20 (4.74)

Anmerkung: Die in Klammern dargestellten Werte sind die z-Werte für jeden Koeffizienten, die mit Hilfe des MECOSA 3-Programms berechnet wurden.

dieser Gruppe auch mit dem Netz der sozialen Sicherheit und den politischen Institutionen in diesem Land zufrieden.

Die allgemeine Lebenszufriedenheit (ALLG) wird sehr stark durch die Zufriedenheit mit der materiellen Situation (MATERIELL) beeinflusst ( $\beta = 0.462$ ). Deutlich geringer hingegen sind die Einflüsse der Zufriedenheit mit der eigenen Wohnung (WOHN), der Wohngegend (WOHNG), der öffentlichen Sicherheit (SICHER), dem sozialen Netz (SOZNET) und den demokratischen Einrichtungen (DEMO) auf die allgemeine Lebenszufriedenheit (ALLG). Die Koeffizienten sind zwar signifikant, aber nicht relevant von Null verschieden.

In Tabelle 4b sind die Einflüsse der sozialstrukturellen

**Tabelle 4b** Strukturgleichungsmodell für die erste Komponente: Einfluss der soziodemographischen Variablen auf die Faktoren

Faktoren	GAMMA			
	REGION	GESCH	BILDUNG	ALTER
WOHN	-0.656 (-3.904)	0.371 (2.388)	-	-
WOHNG	-	-	-0.596 (-2.829)	-
SICHER	-0.997 (-4.257)	-	-	-
SOZNET	-0.342 (-1.765)	-	-	-
DEMO	-	-	-	-
MATERIELL	-0.963 (7.569)	-	0.439 3.183	-
ALLG	-	-	-	-

*Anmerkung:* Die in Klammern dargestellten Werte sind die mit Hilfe des ME-COSA 3-Programms berechneten z-Werte für jeden Koeffizienten.

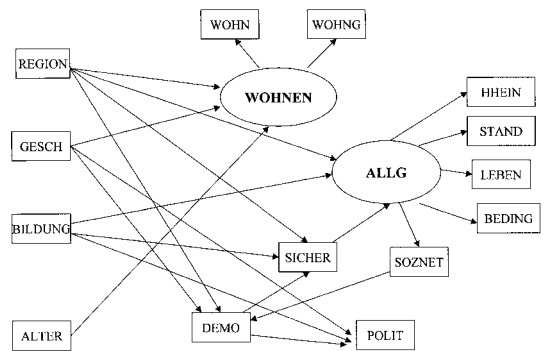
len Variablen auf die Faktoren dargestellt. Betrachtet man die Höhe der Koeffizienten, ist deutlich zu erkennen, dass ein starker negativer Effekt von der Variablen REGION ausgeht. Der negative Koeffizient von REGION auf MATERIELL ( $\beta = -0.963$ ) weist darauf hin, dass in diesem Typus Ostdeutsche mit ihrer materiellen Situation weniger zufrieden sind als Westdeutsche. Der ebenfalls sehr hohe Koeffizient von REGION auf SICHER ( $\beta = -0.997$ ) zeigt an, dass die Ostdeutschen in dieser Gruppe mit der öffentlichen Sicherheit in diesem Land unzufriedener sind als die Westdeutschen. Ebenfalls weniger zufrieden sind sie mit ihrer eigenen Wohnung (WOHN) ( $\beta = -0.656$ ) sowie mit dem Netz der so-

zialen Sicherung (SOZNET) ( $\beta = -0.343$ ). Der positive Koeffizient von GESCH auf WOHN ( $\beta = 0.371$ ) weist auf eine höhere Zufriedenheit der Männer mit ihrer Wohnungen im Vergleich zu den Frauen hin. Der ebenfalls positive Koeffizient von BILDUNG auf MATERIELL ( $\beta = 0.439$ ) zeigt an, dass Personen mit höherer Schulbildung auch mit ihrer materiellen Situation zufriedener sind als Personen mit niedrigerer Schulbildung in dieser Gruppe. Dahingegen weist der negative Koeffizient von BILDUNG auf WOHN ( $\beta = 0.596$ ) auf eine Abnahme der Zufriedenheit mit der Wohngegend bei Personen mit höherer Schulbildung hin.

**Ergebnisse für die zweite Komponente**

Das in Abbildung 2 dargestellte Modell für die zweite Komponente unterscheidet sich im erheblichen Maße von dem der ersten Komponente.

Während in der ersten Komponente die Variablen



**Abb. 2** Signifikante Wirkungszusammenhänge in der zweiten Komponente

„Zufriedenheit mit der Wohnung“ (WOHN) und „Zufriedenheit mit der Wohngegend“ (WOHNG) nicht einem gemeinsamen Faktor zugeordnet werden konnten, laden in der zweiten Komponente die beiden Items auf einen gemeinsamen Faktor, der inhaltlich als „Zufriedenheit mit den Wohnbedingungen“ (WOHNEN) bezeichnet werden kann.

Die Parameterschätzungen für das Messmodell sind in Tabelle 5 aufgeführt. Die Faktorladungen der Indikatoren zur Messung der Wohnbedingungen WOHN und WOHNG sind auf den Wert 1 fixiert. Das ist darauf zurückzuführen, dass die Korrelation zwischen den beiden Indikatoren annähernd perfekt ist ( $r = 0.989$ ) und damit aus statistischer Sicht die Indikatoren identisch sind. Dieser Typus, der bereits aufgrund des Mittelwertvergleichs inhalt-

**Tabelle 5** Messmodell für die zweite Komponente

Indikatoren	NU	LAMBDA					THETA	
		WOHNEN	SICHER	SOZNET	POLIT	DEMO	ALLG	
WOHN	7.930 (24.367)	1.000 (-)	-	-	-	-	-	-
WOHNG	7.997 (20.253)	1.000 (-)	-	-	-	-	-	-
SICHER	5.232 (42.604)	-	1.000 (-)	-	-	-	-	-
SOZNET	6.394 (57.075)	-	-	1.000 (-)	-	-	-	-
HHEIN	6.938 (55.546)	-	-	-	-	-	1.387 (16.976)	1.870 (14.192)
STAND	7.241 (62.666)	-	-	-	-	-	1.376 (20.037)	1.079 (11.420)
POLIT	4.960 (36.932)	-	-	-	1.000 (-)	-	-	-
DEMO	5.443 (42.085)	-	-	-	-	1.000 (-)	-	-
LEBEN	7.714 (85.019)	-	-	-	-	-	1.000 (-)	1.101 (19.571)
BEDING	7.136 (69.930)	-	-	-	-	-	1.154 (17.665)	1.405 (16.125)

Anmerkung: Die in Klammern dargestellten Werte sind die z-Werte für jeden Koeffizienten, die mit Hilfe des MECOSA 3-Programms berechnet wurden.

lich als der Typ mit „hoher Zufriedenheit mit den Wohnbedingungen“ charakterisiert wurde, ist dadurch gekennzeichnet, dass in ihm nur Personen enthalten sind, die in beiden Indikatoren WOHN und WOHNNG (a) die gleichen und (b) extrem hohe Zustimmungswerte aufweisen.

Ferner enthält die zweite Komponente ein Messmodell für die latente Variable „Zufriedenheit mit den allgemeinen Lebensbedingungen“ (ALLG). Die Struktur der Ladungen zeigt, dass die Angehörigen des zweiten Typus unter der Zufriedenheit mit dem Leben im Allgemeinen sowohl die Zufriedenheit mit dem gegenwärtigen Leben (LEBEN) und die Beurteilung der Lebensbedingungen (BEDING) als auch die Zufriedenheit mit dem Haushaltseinkommen (HHEIN) und dem Lebensstandard (STAND) assoziieren. Damit unterscheiden sie sich deutlich von den Angehörigen des ersten Typus, bei denen die Zufriedenheit mit der materiellen Situation zwar einen starken Effekt auf die allgemeine Lebenszufriedenheit ausübt, jedoch ganz klar einen eigenen Faktor bildet.

Als weitere Differenz zur ersten Komponente konnten die Indikatoren DEMO und POLIT nicht auf einen gemeinsamen Faktor reduziert werden, was ein Hinweis darauf ist, dass die Angehörigen dieses Typus die Zufriedenheit mit den demokratischen Einrichtungen deutlich von der Zufriedenheit mit den Möglichkeiten, sich politisch zu betätigen, unterscheiden. Beide Indikatoren werden als separate Variablen betrachtet.

Die geschätzten Parameter des Strukturgleichungsmodells für die zweite Komponente werden in den Tabellen 6a und b aufgelistet. Bedeutend ist der in Tabelle 6a aufgezeigte hohe positive Regressionskoeffizient von der allgemeinen Lebenszufriedenheit (ALLG) auf die Zufriedenheit mit dem sozialen Netz (SOZNET) ( $\beta = 0.857$ ). Dies ist ein Hinweis darauf, dass in diesem Typus die Zufriedenheit mit dem sozialen Netz im erheblichen Maße von der allgemeinen Lebenszufriedenheit beeinflusst wird. Der Regressionskoeffizient von DEMO auf POLIT ( $\beta = 0.678$ ) ist eher als ein Hinweis für eine starke Korrelation zwischen diesen Variablen als eine einseitige Beeinflussung interpretiert werden, da eine



**Tabelle 6a** Strukturgleichungsmodell für die zweite Komponente: Einfluss der Faktoren untereinander

Faktoren	BETA						PSI
	WOHNEN	SICHER	SOZNET	POLIT	DEMO	ALLG	
WOHN	–	–	–	–	–	–	2.44 (16.25)
SICHER	–	–	–	–	0.260 (6.066)	–	5.59 (16.20)
SOZNET	–	–	–	–	–	0.857 (9.541)	3.66 (18.72)
POLIT	–	–	–	–	0.676 (6.357)	–	3.79 (20.26)
DEMO	–	–	0.263 (6.271)	–	–	–	4.13 (5.80)
ALLG	–	0.049 (3.146)	–	–	–	–	0.79 (11.07)

Anmerkung: Die in Klammern dargestellten Werte sind die z-Werte für jeden Koeffizienten, die mit Hilfe des MECOSA 3-Programms berechnet wurden.

**Tabelle 6b** Strukturgleichungsmodell für die zweite Komponente: Einfluß der soziodemographischen Variablen auf die Faktoren

Faktoren	GAMMA			
	REGION	GESCH	BILDUNG	ALTER
WOHNEN	–1.055 (–7.809)	–0.253 (–2.086)	–	0.206 (–3.758)
SICHER	–1.251 (–5.965)	–	0.668 (2.787)	–
SOZNET	–	–	–	–
POLIT	–	0.319 (1.927)	0.339 (1.903)	–
DEMO	–0.926 (–4.247)	0.789 (4.007)	–	–
ALLG	–0.734 (–7.896)	–	0.329 (3.275)	–

Anmerkung: Die in Klammern dargestellten Werte sind die mit Hilfe des ME-COSA 3-Programms berechneten z-Werte für jeden Koeffizienten.

Richtungsänderung die Anpassung des Gesamtmodells nicht verändert. Wenn man die beiden Variablen jedoch als Indikatoren für einen gemeinsamen Faktor betrachtet, wird die Modellanpassung schlechter.

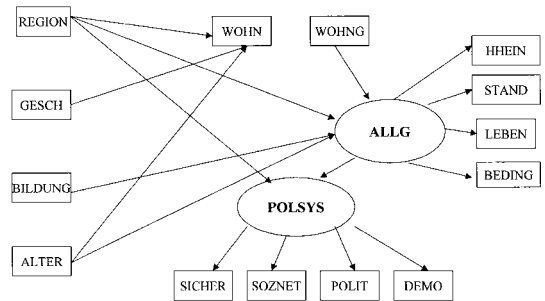
Die Tabelle 6b zeigt einen extrem hohen negativen Koeffizienten von REGION auf WOHNEN ( $\gamma = -1.055$ ). Dies bedeutet, dass in dieser Komponente die Ostdeutschen im Durchschnitt weniger zufriede-

den mit ihren Wohnbedingungen sind als die Westdeutschen. Der ebenfalls extrem hohe negative Effekt von REGION auf SICHER ( $\gamma = -1.251$ ) zeigt an, dass sie auch unzufriedener mit der öffentlichen Sicherheit sind. Dieser Effekt ist deutlich stärker als in der ersten Komponente. Im Unterschied zur ersten Komponente ist in der zweiten Komponente auch ein hoher negativer Koeffizient von REGION auf DEMO ( $\gamma = -0.926$ ) zu verzeichnen, der den Hinweis darauf gibt, dass in dieser Komponente die Ostdeutschen auch mit den demokratischen Einrichtungen in diesem Land unzufriedener als die Westdeutschen sind. Ein direkter Effekt lässt sich in der zweiten Komponente von REGION auf ALLG verzeichnen ( $\gamma = -0.734$ ), der angibt, dass die Ostdeutschen mit dem Leben im allgemeinen, operationalisiert durch die vier oben aufgeführten Indikatoren, unzufriedener als die Westdeutschen sind.

Ein weiteres interessantes Ergebnis ist der Einfluss des Geschlechts (GESCH) auf die Zufriedenheit mit den demokratischen Einrichtungen (DEMO) und den Möglichkeiten, sich in diesem Land politisch zu betätigen (POLIT). Der hohe positive Regressionskoeffizient von GESCH auf DEMO ( $\gamma = 0.789$ ) zeigt an, dass in dieser Komponente Männer mit den demokratischen Einrichtungen in diesem Land zufriedener sind als Frauen. Ferner weist der positive Regressionskoeffizient von GESCH auf POLIT ( $\gamma = 0.319$ ) darauf hin, dass Männer auch mit den Möglichkeiten der politischen Betätigung zufriedener sind. Mit diesen Effekten unterscheidet sich die

zweite Komponente sehr stark von der ersten Komponente, in der keine Unterschiede zwischen Männern und Frauen in Bezug auf diese Variablen festgestellt werden konnten.

Ferner ist der Einfluss der Bildung auf die Zufriedenheit mit der öffentlichen Sicherheit bemerkenswert. Der positive Effekt von BILDUNG auf SICHER ( $\gamma = 0.668$ ) zeigt an, dass in dieser Komponente Personen mit höherer Bildung zufriedener mit der öffentlichen Sicherheit sind als Personen mit niedrigerer Bildung.



**Abb. 3** Signifikante Wirkungszusammenhänge in der dritten Komponente

**Ergebnisse für die dritte Komponente**

In Abbildung 3 sind die signifikanten Wirkungszusammenhänge für die dritte Komponente dargestellt.

Die zugehörigen Parameterschätzer sind in den Tabellen 7 und 8a und 8b aufgelistet. Das Konstrukt „Zufriedenheit mit dem Leben im Allgemeinen“ (ALLG) wird ebenso wie in der zweiten Komponente durch die Indikatoren „Zufriedenheit mit dem Haushaltseinkommen“ (HHEIN), „Zufrieden-

heit mit dem Lebensstandard“ (STAND), „Zufriedenheit mit dem gegenwärtigen Leben“ (LEBEN) und „Beurteilung der gegenwärtigen Lebensbedingungen“ (BEDING) gemessen. Die Struktur der in Tabelle 7 aufgeführten Faktorladungen für das Konstrukt ist der zweiten Komponente ebenfalls sehr ähnlich.

**Tabelle 7** Messmodell für die dritte Komponente

Indikatoren	NU	LAMBDA				THETA
		WOHN	WOHNUNG	POLSYS	ALLG	
WOHN	4.980 (10.432)	1.000 (-)	-	-	-	-
WOHNUNG	6.572 (17.969)	-	1.000 (-)	-	-	-
SICHER	4.531 (16.899)	-	-	0.999 (5.486)	-	5.162 (12.112)
SOZNET	6.034 (19.747)	-	-	1.192 (6.751)	-	3.619 (9.638)
HHEIN	5.686 (15.590)	-	-	-	1.330 (10.749)	3.167 (11.703)
STAND	6.712 (17.975)	-	-	-	1.426 (12.973)	1.710 (9.271)
POLIT	4.927 (18.268)	-	-	1.000 (-)	-	3.988 (11.452)
DEMO	5.504 (4.909)	-	-	1.569 (8.019)	-	2.235 (6.048)
LEBEN	7.355 (26.178)	-	-	-	1.000 (-)	2.024 (13.738)
BEDING	6.605 (21.012)	-	-	-	1.142 (11.038)	2.140 (11.334)

Anmerkung: Die in Klammern dargestellten Werte sind die z-Werte für jeden Koeffizienten, die mit Hilfe des MECOSA 3-Programms berechnet wurden.

Das Modell für die dritte Komponente enthält ein Konstrukt, das in den beiden anderen Komponenten nicht vorhanden ist. Es handelt sich dabei um eine Zufriedenheit, die außerhalb des persönlichen Bereich des Individuums steht und im Bereich des politischen Systems anzusiedeln ist. Dieses Konstrukt wird durch die Indikatoren „Zufriedenheit mit der öffentlichen Sicherheit in diesem Land“ (SICHER), „Zufriedenheit mit dem Netz der sozialen Sicherung“ (SOZNET), „Zufriedenheit mit den Möglichkeiten, sich politisch zu betätigen“ (POLIT) und „Zufriedenheit mit den demokratischen Einrichtungen in diesem Land“ (DEMO) gemessen und als „Zufriedenheit mit dem politischen System“ (POLSYS) bezeichnet. Aus Tabelle 7 wird ersichtlich, dass POLSYS am stärksten mit dem Indikator DEMO verknüpft ist ( $\lambda = 1.596$ ), was ein Hinweis darauf ist, dass in dieser Komponente die Zufriedenheit mit dem politischen System am stärksten durch die Zufriedenheit mit den demokratischen Einrichtungen in diesem Land assoziiert wird.

**Tabelle 8a** Strukturgleichungsmodell für die dritte Komponente: Einfluss der Faktoren untereinander

Faktoren	BETA				PSI
	WOHN	WOHNG	POLSYS	ALLG	
WOHN	–	–	–	–	8.332 (19.984)
WOHNG	–	–	–	–	13.731 (18.611)
POLSYS	–	–	–	0.229 (4.012)	1.080 (4.166)
ALLG	0.106 (5.976)	–	–	–	1.306 (6.833)

Anmerkung: Die in Klammern dargestellten Werte sind die mit Hilfe des MECOSA 3-Programms berechneten z-Wertes für jeden Koeffizienten.

Tabelle 8a zeigt, dass in der Gruppe der „Unzufriedeneren“ die allgemeine Lebenszufriedenheit geringfügig durch die Zufriedenheit mit der Wohngegend beeinflusst wird ( $\beta = 0.106$ ) und von der Zufriedenheit mit der Wohnung unbeeinflusst bleibt. Interessant ist in dieser Gruppe die Beziehung zwischen dem Faktor (ALLG) und dem Faktor (POLSYS), die auf einen Zusammenhang zwischen der allgemeine Lebenszufriedenheit und der Zufriedenheit mit dem politischen System hinweist. Bedenkt man, dass die allgemeine Lebenszufriedenheit durch Lebensbereiche gemessen wird, die der privaten Sphäre des Individuums zuzurechnen sind

und die Zufriedenheit mit dem politischen System durch Bereiche, die der öffentlichen Sphäre zuzurechnen sind, wird der Zusammenhang zwischen den Zufriedenheiten in privaten und öffentlichen Bereichen deutlich. Die Richtung der Kausalbeziehung zeigt jedoch an, dass die Zufriedenheit mit dem politischen System keinen Einfluss auf die allgemeine Lebenszufriedenheit ausübt. Damit wird deutlich, dass in den öffentlichen Lebensbereichen kein Leistungspotential für das subjektive Wohlbefinden liegt. Eine hervorragende Anpassung des Gesamtmodells konnte dann durch die Kausalrichtung von ALLG auf POLSYS erzielt werden. Der Effekt von ALLG auf POLSYS ( $\beta = 0.229$ ) weist darauf hin, dass die Zufriedenheit mit dem Leben im Allgemeinen einen positiven Einfluss auf die Zufriedenheit mit dem politischen System hat. Dies könnte eventuell damit zusammenhängen, dass die allgemeine Lebenszufriedenheit sehr stark durch materielle Bereiche wie die Zufriedenheit mit dem Lebensstandard  $\lambda = 1.426$  und der Zufriedenheit mit dem Haushaltseinkommen  $\lambda = 1.330$  assoziiert ist und dass diese Zufriedenheiten den Einfluss auf die Zufriedenheit mit dem politischen System ausüben. Da es aber keine Rechtfertigung für diese Überlegung gibt, sollte der Effekt von ALLG auf POLSYS eher als Korrelation der allgemeinen Lebenszufriedenheit mit der Zufriedenheit mit dem politischen System aufgefasst werden.

**Tabelle 8b** Strukturgleichungsmodell für die dritte Komponente: Einfluss der soziodemographischen Variablen auf die Faktoren

Faktoren	GAMMA			
	REGION	GESCH	BILDUNG	ALTER
WOHN	–0.976 (–4.151)	–0.949 (–4.010)	–	0.594 (6.003)
WOHNG	–	–	–	–
POLSYS	–0.763 (–4.659)	–	–	–
ALLG	–0.764 (–5.510)	–	0.544 (3.246)	–0.164 (–2.739)

Anmerkung: Die in Klammern dargestellten Werte sind die mit Hilfe des ME-COSA 3-Programms berechneten z-Werte für jeden Koeffizienten.

Aus Tabelle 8b wird ersichtlich, dass auch in dieser Komponente, die inhaltlich als der „Typus mit der geringsten Zufriedenheit mit ausgewählten Lebensbereichen und dem Leben im Allgemeinen“ charakterisiert wurde und durch einen großen Anteil Be-

wohner neuer Bundesländer gekennzeichnet ist, ein starker Effekt von der Variablen REGION ausgeht. Der negative Koeffizient von REGION auf WOHN ( $\gamma = -0.976$ ) zeigt, dass die Bewohner der neuen Bundesländer mit ihrer Wohnung unzufriedener sind als die Bewohner der alten Bundesländer; der negative Koeffizient auf POLSYS, dass sie auch mit dem politischen System weniger zufrieden ( $\gamma = -0.763$ ) und der negative Koeffizient auf ALLG, dass sie mit dem Leben im Allgemeinen unzufriedener ( $\gamma = -0.764$ ) sind. Erwähnenswert ist ferner der Koeffizient von GESCH auf WOHN ( $\gamma = 0.949$ ), der anzeigt, dass die Frauen mit ihrer Wohngegend unzufriedener sind.

## Schlussbetrachtung

Als mögliche Verbesserung der Analysetechniken zur Lösung der aufgezeigten methodischen Probleme werden in diesem Artikel das Modell der finiten Mischung von konditional normalverteilten Zufallsvariablen mit parametrischen Mittelwert- und Kovarianzstrukturen und drei Schätzmethoden zur Schätzung der Parameter des Modells vorgestellt. Dieser Ansatz kann zum einen als Alternative zu den herkömmlichen Clustertechniken verwendet werden. Er ist diesen in vielerlei Hinsicht überlegen: Erstens besteht im Unterschied zur Clusteranalyse die Möglichkeit, die finiten Mischungen auf vorgegebene Variablen (z. B. sozialstrukturelle Variablen) zu konditionieren. Damit können bei der Schätzung unbekannter Gruppen erklärende Variablen berücksichtigt werden. Zweitens können zur Beschreibung der Unterschiede zwischen den Komponenten nicht nur die Mittelwerte sondern auch die Kovarianzmatrizen der untersuchten Variablen verwendet werden. Dies ermöglicht auf einer deskriptiven Ebene, die Beziehung zwischen den Variablen darzustellen. Drittens können Hypothesen über komplexe Wechselbeziehungen zwischen einzelnen (latenten und beobachteten) Variablen mit komponentenspezifischen Strukturgleichungsmodellen formuliert und überprüft werden, indem die bedingten Mittelwerte und Kovarianzmatrizen entsprechend parametrisiert werden. Viertens lässt sich ein Signifikanztest für die Anzahl der Komponenten einer Mischverteilung angeben.

Zum anderen ist das Verfahren eine Alternative zu der herkömmlichen Anwendung statistischer Verfahren wie Mittelwert- und Kovarianzstrukturanalyse, weil es ermöglicht, vorhandene Unterschiede in den Strukturen in unbekanntem Subpopulationen, die auf der Ebene der Gesamtdaten verdeckt

werden, zu ermitteln. Dies gilt nicht nur für die Anwendung der Mittelwert- und Kovarianzstrukturanalyse, sondern für einen Großteil statistischer Verfahren wie z. B. die Regressionsanalyse, bei der sich positive und negative Einflüsse in den Subpopulationen gegenseitig aufheben können, wenn die unbeobachtete Heterogenität unberücksichtigt bleibt.

Dass die Existenz unbeobachteter Heterogenität Analyseergebnisse stark verzerren kann, wird anhand des empirischen Beispiels deutlich. Die Anwendung des Mischverteilungsmodells in Form von Mischungen konditionaler LISREL-Modelle zur Analyse des Zusammenhangs zwischen sozialstrukturellen Variablen, Zufriedenheiten in einzelnen Lebensbereichen und der allgemeinen Lebenszufriedenheit ergibt den klaren Befund einer unbeobachteten Heterogenität. Es konnten drei Typen identifiziert werden, die sich sowohl im Niveau an Zufriedenheit in einzelnen Lebensbereichen und der allgemeinen Lebenszufriedenheit unterscheiden als auch in den Beziehungen zwischen einzelnen Aspekten der Zufriedenheit. So konnte für jeden Typus ein eigenes Modell gefunden werden, das sich nicht nur im Strukturgleichungsmodell sondern auch in den Messmodellen für die Konstrukte von den anderen unterscheidet. Beispiele für ersteres sind die enorme Bedeutung der Zufriedenheit mit der materiellen Situation als Einflussgröße auf einzelne Lebensbereiche und der allgemeinen Lebenszufriedenheit in der ersten Komponente, die inhaltlich auch als der Typus mit hoher Zufriedenheit mit seiner materiellen Situation charakterisiert wird sowie die Beziehung zwischen der allgemeinen Lebenszufriedenheit und der Zufriedenheit mit dem politischen System im Typus des „mit allen Bereichen des Lebens Unzufriedeneren“. Als Beispiele für letzteres sind die unterschiedliche Konnotation des Konstruktes „allgemeinen Lebenszufriedenheit“ in den drei Komponenten zu nennen sowie die identische Assoziation der Zufriedenheit mit den Wohnbedingungen sowohl mit der Zufriedenheit mit der eigenen Wohnung als auch mit den Wohnbedingungen in der zweiten Komponente.

## Literatur

- Aitkin, M. / Anderson, D. / Hinde, J., 1981: Statistical modelling of data on teaching styles. *Journal of the Royal Statistical Society A* 144: 419–461.
- Arminger, G. / Stein, P., 1997: Finite mixtures of covariance structure models with regressors. *Sociological Methods and Research* 26: 148–182.
- Arminger, G. / Stein, P. / Wittenberg, J., 1999: Mixtures of conditional mean- and covariance structure models. *Psychometrika* 4, in Erscheinung.

- Arminger, G. / Wittenberg, J., 1998: MECOSA 3: A program for the analysis of general mean- and covariance structures with non-metric variables, User Guide. Frauenfeld, Switzerland: SLI-AG.
- Becker, M. P. / Yang, I. / Lange, K., 1997: EM-algorithms without missing data. *Statistical Methods in Medical Research* 6: 37–53.
- Bentler, P. M. / Bonett, D. G., 1980: Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin* 88: 588–606.
- Bentler, P. M., 1988: Comparative fit indexes in structural models. *Psychological Bulletin* 107: 238–246.
- Bentler, P. M., 1995: *Theory and Implementation of EQS*. Los Angeles: BMDP Statistical Software, Inc.
- Berger-Schmitt, R., 1994: Niveau und Struktur der Zufriedenheit in der Bundesrepublik Deutschland – Stabilität und Wandel 1978–1988. S. 37–58 in: R. Hauser (Hrsg.), *Mikroanalytische Grundlagen der Gesellschaftspolitik: Ergebnisse aus dem gleichnamigen Sonderforschungsbereich an den Universitäten Frankfurt/M. und Mannheim*, Bd. 1. Berlin: Akademie-Verlag.
- Browne, M. / Cudeck, R., 1993: Alternative ways of assessing model fit. S. 136–162 in: K. A. Bollen / J. S. Long (Hrsg.), *Testing structural equation models*. Newbury Park: Sage.
- Browne, M. / Arminger, G., 1995: Mean- and covariance structure models. S. 185–249 in: G. Arminger / C. C. Clogg / M. E. Sobel (Hrsg.), *Handbook of Statistical Modeling for the Social and Behavioral Sciences*. New York: Plenum.
- Clogg, C. C., 1995: Latent class models. S. 311–359 in: G. Arminger / C. C. Clogg / M. E. Sobel (Hrsg.): *Handbook of Statistical Modeling for the Social and Behavioral Sciences*. New York: Plenum.
- DeSarbo, W. S. / Cron, W. L., 1988: A maximum likelihood methodology for clusterwise linear regression. *Journal of Classification* 5: 249–282.
- DeSarbo, W. S. / Wedel, M. / Vriens, M. / Ramaswamy, V., 1992: Latent class metric conjoint analysis. *Marketing Letters* 3: 273–288.
- DeSarbo, W. S. / Ramaswamy, V. / Reibstein, D. J. / Robinson, W. T., 1993: A latent pooling methodology for regression analysis with limited time series of cross sections: a PIMS data application. *Marketing Science* 12: 103–124.
- Diewald, M. / Zapf, W., 1984: Wohnbedingungen und Wohnzufriedenheit. S. 73–96 in: W. Glatzer / W. Zapf (Hrsg.), *Lebensqualität in der Bundesrepublik. Objektive Lebensbedingungen und subjektives Wohlbefinden*. Frankfurt/M.: Campus.
- Do, K. / McLachlan, G. J., 1984: Estimation of mixing proportions: a case study. *Applied Statistics* 33: 134–140.
- Everitt, B. S. / Hand, D. J. 1981: *Finite mixture distributions*. London: Chapman and Hall.
- Faulbaum, F., 1987: Intergroup comparisons of latent means across waves. *Sociological Methods and Research* 15: 317–335.
- Gibson, W. A., 1966: Latent structure analysis and test theory. S. 78–88 in: P. F. Lazarsfeld / N. W. Henry (Hrsg.): *Readings in mathematical social research*. Cambridge Massachusetts London.
- Glatzer, W., 1984: Einkommensverteilung und Einkommenszufriedenheit. S. 45–72 in: W. Glatzer / W. Zapf (Hrsg.), *Lebensqualität in der Bundesrepublik. Objektive Lebensbedingungen und subjektives Wohlbefinden*. Frankfurt/M.: Campus.
- Goodman, L. A., 1974: The analysis of systems of qualitative variables when some of the variables are unobservable. *Journal of Sociology* 79: 1179–1259.
- Habich, R. / Noll, H. H. / Zapf, W. 1999: Subjektives Wohlbefinden in Ostdeutschland nähert sich westdeutschem Niveau. S. 1–6 in: ZUMA (Hrsg.), *ISI 22*.
- Hampel, J., 1985: Lebenszufriedenheit und Bereichszufriedenheiten. Eine Anwendung der LISREL-Methode. Arbeitspapier Nr. 178 des Sonderforschungsbereich 3, Mikroanalytische Grundlagen der Gesellschaftspolitik. J. W. Goethe-Universität Frankfurt und Universität Mannheim.
- Hildebrand, L., 1983: *Konfirmatorische Analysen von Modellen des Konsumentenverhaltens*. Berlin: Duncker & Humblot.
- Jedidi, K. / Jagpal, H. S. / DeSarbo, W. S., 1997: Finite-mixture structural equation models for response-based segmentation and unobserved heterogeneity. *Marketing Science* 16: 39–59.
- Jedidi, K. / Ramaswamy, V. / DeSarbo, W. S. / Wedel, M., 1996: On Estimating Finite Mixtures of Multivariate Regression and Simultaneous Equation Models. *Structural Equation Modelling* 3: 266–289.
- Jagodzinski, W. / Kühnel, S. M. / Schmidt, P., 1987: Is there a “socratic effect” in nonexperimental panel studies? *Sociological Methods and Research* 15: 259–302.
- Jones, P. N. / McLachlan, G. J., 1992: Fitting finite mixture models in a regression context. *Australian Journal of Statistics* 43: 233–440.
- Jöreskog, K. G. / Sörbom, D., 1993: *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Hillsdale, NJ: Erlbaum.
- Landua, D. / Spellerberg, A. / Habich, R., 1991: Der lange Weg zur Einheit. Unterschiedliche Lebensqualität in den „alten“ und „neuen“ Bundesländern. *WZB-Bericht* P 91–101. Berlin.
- Krebs, D. / Schmidt, P., 1996: Satisfaction in Germany. S. 117–132 in: W. E. Saris / R. Veenhoven / A. C. Scherpenzeel / B. Bunting (Hrsg.), *A Comparative Study of Satisfaction with Life in Europe*. Budapest: Eötvös University Press.
- Lazarsfeld, P. F. / Henry, N. W., 1968: *Latent structure analysis*. Boston.
- McLachlan, G. J. / Basford, K. E., 1988: *Mixture Models*. New York: Marcel Dekker.
- Müller, J., 1994: Lebenslagen, Lebenschancen, soziales Wohlbefinden von Individuen, Haushalten und Familien. S. 63–115 in: D. Wittlich (Hrsg.), *Momente des Umbruchs. Sozialstruktur und Lebensqualität in Ostdeutschland*. Budapest: Berlin: edition sigma.
- Spellerberg, A., 1995: Lebensstile und Lebensqualität in West- und Ostdeutschland. *Angewandte Sozialforschung* 19, 1: 93–106.

- Saris, W. E. / Scherpenzeel, A.C., 1996: Methodological procedures of comparative research. S. 49–78 in: W. E. Saris / R. Veenhoven / A. C. Scherpenzeel / B. Bunting (Hrsg.), *A Comparative Study of Satisfaction with Life in Europe*. Budapest: Eötvös University Press.
- Spellerberg, A., 1996: *Soziale Differenzierung durch Lebensstile. Eine empirische Untersuchung zur Lebensqualität in West- und Ostdeutschland*. Berlin: edition sigma.
- Spellerberg, A. / Landua, D. / Habich, R., 1992: Orientierungen und subjektives Wohlbefinden in West- und Ostdeutschland. S. 249–278 in: W. Glatzer / H. H. Nöll (Hrsg.), *Lebensverhältnisse in Deutschland: Ungleichheit und Angleichung. Soziale Indikatoren, Band XI*. Frankfurt/Main: Suhrkamp.
- Steiger, J. H. / Lind, J. C., 1980: Statistically based tests for the number of common factors, paper presented at the annual meeting of the Psychometric Society. Iowa City: Iowa.
- Stein, P., 1997: *Konstruktion und sozialwissenschaftliche Anwendung finiter Mischungen von Kovarianzstrukturmodellen*. Köln: Eul-Verlag.
- Titterton, D. M. / Smith, A. F. M. / Makov, U. E., 1985: *Statistical analysis of finite mixture distributions*. Chichester: Wiley.
- Veenhoven, R., 1993: *Happiness in Nations. Subjective appreciation of life in 56 nations (1946–1992)* RISBO, *Studies in Social and Cultural Transformation 2*. Erasmus University Rotterdam.
- Veenhoven, R., 1994: Is happiness a trait? Tests of the theory that a better society does not make people any happier. *Social Indicators Research*, 32: 101–160.
- Wolfe, J. H., 1971: A Monte Carlo Study of the sampling distribution of the likelihood ratio for mixtures of multinormal distributions. Technical Bulletin STB 72–2. San Diego: U.S. Naval Personnel and Training Research Laboratory.
- Yung, Y. F. 1997. Finite mixtures in confirmatory factor-analytic models. *Psychometrika* 62: 297–330.