

# **Wie homogen sind Einstellungen gegenüber Ausländern? Zur Aufdeckung und Modellierung unbeobachteter Heterogenität in Umfragedaten**

**Frank Faulbaum, Petra Stein**

## **1. Einleitende Bemerkungen**

Die Aufdeckung der unbeobachteten Heterogenität in Umfragedaten stellt trotz eines inzwischen reichen Repertoires an Klassifikationsverfahren (vgl. z.B. Bacher 1996, Gabler 1997, Hand 1996 u. 1997) sowie neuer Möglichkeiten der Visualisierung komplexer Daten (vgl. z.B. Cleveland 1993) ein weiterhin bestehendes Problem für die statistische Datenanalyse dar. Sie setzt der sinnvollen Anwendung hypothesentestender Verfahren und der Interpretierbarkeit von Resultaten der statistischen Datenanalyse Grenzen, die in ihrer Bedeutung und ihrer Wirkung nicht immer abgeschätzt werden können. Insbesondere in der Panelanalyse (vgl. Hsiao 1995) wurden die Konsequenzen der Existenz unbeobachteter Heterogenität für die statistische Modellierung klar herausgearbeitet und präzisiert. Im Fall der explorativen oder auch konfirmatorischen Faktorenanalyse können in unbekanntem Gruppen möglicherweise vorhandene Unterschiede in den Faktorenstrukturen auf der Ebene der Gesamtdaten vollkommen verdeckt werden. Ein weiteres Beispiel stellt auch die Regressionsanalyse dar, wo sich negative und positive Einflüsse in Untergruppen auf der Ebene der Gesamtdaten gegenseitig aufheben können.

Die am häufigsten angewendeten Standardverfahren für die Entdeckung unbekannter Gruppierungen in Umfragedaten stellen zweifellos die clusteranalytischen Verfahren dar. Diese weisen allerdings eine Reihe bemerkenswerter Unzulänglichkeiten auf. Dazu gehören insbesondere die Vernachlässigung statisti-

scher Verteilungsinformationen, die Beschränkung auf univariate Charakterisierungen der Cluster, etwa durch Mittelwerte, sowie das Fehlen statistischer Tests für die Anzahl der Cluster. Die Clusterbildung geschieht auf der Basis von Distanz- oder Ähnlichkeitsmaßen, aber nicht auf der Basis multivariater Verteilungen der Variablen. Gerade letzteres wäre aber notwendig, wenn unbekannte Gruppen gesucht werden sollen, die sich durch komplexe Abhängigkeitsstrukturen der Variablen, d.h. etwa durch unterschiedliche Kausalmodelle mit manifesten und/oder latenten Variablen voneinander unterscheiden können, da Kausalmodelle üblicherweise als Parametrisierungen der Kovarianzmatrizen und Mittelwertsvektoren multivariater Verteilungen charakterisiert sind.

Schon das Auffinden unbekannter Gruppierungen, die sich durch ihre Meßmodelle voneinander unterscheiden, ist nur denkbar, wenn die multivariaten statistischen Verteilungen und ihre Parametrisierungen bei der Gruppenbildung in Betracht gezogen werden.

Im Falle *beobachteter* Heterogenität, also in dem Fall, wo die Gruppen a priori bekannt sind, stehen die Methoden der Mehrstichprobenanalyse (multiple Gruppenvergleiche) zur Verfügung. Mit Hilfe dieser Analyse wird geprüft, ob sich die Gruppen in bestimmten Aspekten ihrer kausalen Struktur voneinander unterscheiden (vgl. Bentler et al. 1987, Jöreskog 1971, Muthèn 1989, Sörbom 1974). Auch bei ALLBUS-Daten wurden diese Verfahren in der Vergangenheit mehrfach angewendet (vgl. z.B. Faulbaum 1987, 1990, Porst et al. 1987) aber auch in diesem Band (Schmidt/Heyder). Im Fall *unbeobachteter* Heterogenität müssen die Gruppen mit ihren unterschiedlichen Kausalstrukturen erst gefunden werden.

In letzter Zeit wurden im Bereich der Methoden zur Extraktion unbekannter Gruppen bedeutende Fortschritte erzielt. Diese Fortschritte basieren auf Weiterentwicklungen von Mischverteilungsansätzen in Richtung auf die Einbeziehung exogener Variablen und der parametrischen Struktur von Kovarianzmatrizen. Hier sind vor allem die Arbeiten von Stein (1997) sowie Armingier, Stein und Wittenberg (1999) zu nennen, die auf der Basis von Vorarbeiten von Yung (1994) sowie DeSarbo und Cron (1988) entsprechende Erweiterungen des Mischverteilungsansatzes vorgenommen haben. Angewendet wurden diese Ansätze bisher auf Lebensstilvariablen (vgl. Stein 1997), auf die ökologische Orientierung und die Einstellung zum Automobil (Armingier/Stein 1997) und auf Einstellungen zur Integration von Ausländern auf der Basis von Items des ALLBUS 1996 (vgl. Faulbaum/Stein 1998).

In diesem Aufsatz sollen Ergebnisse einer Analyse einiger im ALLBUS 1996 erhobenen Variablen vorgestellt werden, die in einem Zusammenhang zur Integration von Ausländern stehen. Die Ergebnisse sind relevant für viele Arbeiten,

die bisher auf der Basis einer Analyse der ALLBUS-Variablen zur Ausländerfeindlichkeit veröffentlicht wurden (vgl. Blank/Schwarzer 1994, Gehring/Böltken 1985, Kühnel 1987, Leenen 1992, Reuband 1989, vgl. ferner die Beiträge in Bohrnstedt et al. 1987). Die theoretisch/empirische Forschungspraxis in den Untersuchungen zur Ausländerfeindlichkeit ist dabei durchaus nicht unumstritten (vgl. Jaeger 1995). Das Problem beginnt bereits bei der Frage, ob es sich bei diesen Variablen überhaupt um Indikatoren für ein Konstrukt handelt, das in allen sozialen Gruppierungen die gleiche Konnotation besitzt oder ob sich die Stichprobe des ALLBUS eventuell aus noch unbekanntem Untergruppen zusammensetzt, in denen diese Variablen als Indikatoren für unterschiedliche Konstrukte fungieren. Letzteres hätte wiederum Auswirkungen auf die Konstruktvalidität in diesen Gruppen bzw. die Rolle, die das Konstrukt in Kausalmodellen spielt. Die vorliegende Analyse geht davon aus, daß es sich bei den Variablen der Ausländerfeindlichkeit im Grunde um Indikatoren für die Einstellung zur Integration von Ausländern handelt. Untersucht werden soll, inwieweit sich hinter den Einstufungen der Befragten unbekanntes Gruppen verbergen, die hinsichtlich ihrer Einstellung gegenüber der Integration von Ausländern in sich homogen sind, sich aber untereinander deutlich unterscheiden. Zur Lösung des Problems wird der von Stein (1997) sowie Arminger, Stein und Wittenberg (1998) entwickelte Mischverteilungsansatz angewendet.

Im folgenden wird zunächst der methodische Ansatz dargestellt. Dabei werden das Modell finiter Mischungen (Komponenten) von Normalverteilungen sowie einzelne Erweiterungen des Modells auf Mischungen *bedingter* Normalverteilungen dargestellt. Anschließend wird gezeigt, wie die parametrische Struktur der Mittelwerte und der Kovarianzmatrizen, d.h. die Parameter von Strukturgleichungsmodellen, in den Prozeß der Modellsuche eingebunden werden kann. Für die Schätzung der Parameter eines Mischverteilungsmodells unter Einbeziehung der Mittelwerts- und Kovarianzstrukturen in den einzelnen Gruppen werden drei Methoden skizziert. Schließlich werden ein Test auf die Anzahl der Komponenten einer Mischverteilung vorgestellt und die Verwendung von Anpassungsindizes zur Beurteilung der Modellanpassung angesprochen. Ausführlichere Darstellungen finden sich in Stein (1997) sowie Arminger, Stein und Wittenberg (1999). Auf die Darstellung des methodischen Ansatzes erfolgt dessen Anwendung auf die ALLBUS-Daten.

## 2. Methodischer Ansatz

### 2.1 Unbekannte Gruppen als Komponenten finiter Mischungen multivariater Verteilungen

Ein möglicher Ansatz zur Behandlung unbekannter Heterogenität stellt das multivariate Mischverteilungsmodell dar. In diesem Ansatz wird die Analyse unbeobachteter Heterogenität bzw. der Entdeckung unbekannter Gruppen mit komplexen Abhängigkeitsstrukturen als Problem der Extraktion der Komponenten finiter Mischungen multivariater Normalverteilungen (vgl. McLachlan/Basford 1988, Titterington et al. 1985).

Zum besseren Verständnis des Mischverteilungsansatzes sei zunächst der einfache Fall einer *univariaten* Mischverteilung mit nur *zwei* Komponenten betrachtet. Ein Beispiel wäre etwa gegeben, wenn die Einkommensverteilung in der Population aus einer Einkommensverteilung für die Komponente (Gruppe) der Männer und einer Einkommensverteilung für die Komponente (Gruppe) der Frauen zusammengesetzt wäre. Mathematisch kann dieser Fall in der folgenden Weise charakterisiert werden:

$$f(y_i) = \pi_1 f_1(y_i) + \pi_2 f_2(y_i)$$

$\pi_1$  und  $\pi_2$  heißen *mischende Wahrscheinlichkeiten*.  $\pi_1$  stellt die Wahrscheinlichkeit dar, daß ein Element der Population weiblich ist, und  $\pi_2$  die Wahrscheinlichkeit, daß ein Element der Population männlich ist.  $y_i$  ist die Variable *Einkommen* und  $f$  deren Wahrscheinlichkeitsdichtefunktion mit dem Erwartungswert  $\mu$  und der Varianz  $\sigma^2$ .  $f_1$  und  $f_2$  sind bedingte Dichtefunktionen mit den Parametern  $\mu_1$  und  $\sigma_1^2$  bzw.  $\mu_2$  und  $\sigma_2^2$ . In unserem einfachen Beispiel setzt sich die totale Population also aus *zwei Komponenten* zusammen, die durch spezifische Dichtefunktionen und spezifische Wahrscheinlichkeiten ihres Auftretens beschrieben werden können. Für die mischenden Wahrscheinlichkeiten gilt die Restriktion  $\pi_1 + \pi_2 = 1$ .

Das Beispiel bezog sich auf eine Mischverteilung mit zwei Komponenten und nur einer Variablen. Es kann ohne weiteres auf mehr als zwei Komponenten und mehr als eine Variable verallgemeinert werden. In diesem Fall sind die totale Dichtefunktion und die bedingten Dichtefunktionen multivariat und beziehen sich auf einen Vektor von Zufallsvariablen. Ein solcher allgemeinerer Fall läge etwa dann vor, wenn sich die totale Dichtefunktion verschiedener Variablen wie z.B. Lebensstilvariablen in unterschiedliche Gruppen zerlegen ließe, deren

jede sich jeweils durch eine spezifische Wahrscheinlichkeitsdichtefunktion dieser Variablen beschreiben ließe. Die Verteilungsparameter sind bekanntermaßen im multivariaten Fall nicht mehr einzelne Erwartungswerte und Varianzen, sondern Mittelwertsvektoren und Kovarianzmatrizen. Dieser allgemeine Fall läßt sich in der folgenden Weise mathematisch darstellen (vgl. Titterington et al. 1985):

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

In dieser Formel bezeichnen  $K$  die Anzahl der Komponenten (Gruppen),  $\pi_k$  die mischenden Wahrscheinlichkeiten,  $\phi(\mathbf{y}_i; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$  die auf die Komponenten  $k$  bedingten multivariaten Normalverteilungen des Vektors  $\mathbf{y}_i$  von Zufallsvariablen mit Mittelwertsvektoren  $\boldsymbol{\mu}_k$  und Kovarianzmatrizen  $\boldsymbol{\Sigma}_k$  als Verteilungsparametern.

Als Verteilung  $f(\mathbf{y}_i)$  wird eine gemeinsame multivariate Normalverteilung aller Zufallsvariablen des Variablenvektors  $\mathbf{y}_i$  angenommen. Dies bedeutet etwa, daß in dem Fall, wo eine Unterteilung der Variablen in exogene und endogene Variablen vorliegt, auch bezüglich der exogenen Variablen eine Normalverteilung unterstellt werden müßte. Dies stellt für viele Arten exogener Variablen, die nicht selten diskret sind, eine unrealistische Annahme dar. Wird für nicht normalverteilte exogene Variablen irrtümlicherweise eine Normalverteilung unterstellt, so ergeben sich inkonsistente Parameterschätzungen (vgl. Stein 1997). Die strenge Annahme einer multivariaten Normalverteilung aller Variablen läßt sich abschwächen, wenn die Verteilung der endogenen Variablen auf die exogenen Regressoren bedingt wird.

Einen entsprechenden Ansatz formuliert unter Verallgemeinerung eines univariaten Mischverteilungsmodells mit Regressoren von DeSarbo und Cron (1988) Stein (1997; vgl. auch Arminger & Stein 1997). In diesem Fall wird statt der obigen Verteilung von  $f(\mathbf{y}_i)$  die bedingte Verteilung von  $\mathbf{y}_i$  gegeben Regressoren  $\mathbf{x}_i$  betrachtet:

$$f(\mathbf{y}_i | \mathbf{x}_i) = \sum_{k=1}^K \pi_k \phi(\mathbf{y}_i; \boldsymbol{\gamma}_k + \boldsymbol{\Pi}_k \mathbf{x}_i, \boldsymbol{\Sigma}_k)$$

In dieser Formel bezeichnen  $\boldsymbol{\gamma}_k$  den Vektor der Regressionskonstanten und  $\boldsymbol{\Pi}_k$  die Matrix der Regressionskoeffizienten, wobei für den Mittelwertsvektor  $\boldsymbol{\mu}_{ik}$  der bedingten Verteilung von  $\mathbf{y}_i$  gegeben Regressoren  $\mathbf{x}_i$  und die  $k$ -te Komponente gilt, daß  $\boldsymbol{\mu}_{ik} = \boldsymbol{\gamma}_k + \boldsymbol{\Pi}_k \mathbf{x}_i$ .  $\boldsymbol{\Sigma}_k$  ist nunmehr die auf die exogenen Variablen  $\mathbf{x}_i$  bedingte Kovarianzmatrix.

Für die vorliegende Aufgabe, die in den Daten vorhandene unbekannte Heterogenität in den auf exogene Variablen  $x_i$  bedingten Variablen  $y_i$  unter Einbeziehung der kausalanalytischen Strukturen bzw. der Abhängigkeitsstrukturen zwischen latenten und/oder beobachteten Variablen zu erfassen, müssen in die Suche nach unbekanntem Gruppen auch mögliche Parametrisierungen von  $\boldsymbol{\mu}_{ik} = \boldsymbol{\gamma}_k + \boldsymbol{\Pi}_k$  und  $\boldsymbol{\Sigma}_k$  einbezogen werden. Yung (1994) hatte für den Fall ohne bedingende exogene Variablen bereits faktorenanalytische Parametrisierungen behandelt. Stein (1997) hat diesen Fall für bedingte Verteilungen der  $y_i$  auf beliebige Modelle mit latenten Variablen verallgemeinert. Ein mögliches Modell wäre etwa ein konditionales LISREL-Modell, das aus dem Meßmodell

$$y_i^{(k)} = \mathbf{v}_k + \mathbf{A}_k \boldsymbol{\eta}_i^{(k)} + \boldsymbol{\varepsilon}_i^{(k)}$$

und dem Strukturmodell

$$\boldsymbol{\eta}_i^{(k)} = \mathbf{B}_k \boldsymbol{\eta}_i^{(k)} + \boldsymbol{\Gamma}_k x_i + \boldsymbol{\zeta}_i^{(k)}$$

besteht, wobei die in diesen Gleichungen auftretenden Symbole die in der Literatur über Strukturgleichungsmodelle (vgl. Bollen 1989, Loehlin 1998) üblichen Bedeutungen besitzen:

$\boldsymbol{\eta}^{(k)}$  ist ein Vektor latenter Variablen gegeben die Regressorvariablen  $x_i$  in jeder Komponente  $k$ ,  $\mathbf{A}_k$  ist die Ladungsmatrix der latenten Variablen in der  $k$ -ten Komponente,  $\boldsymbol{\varepsilon}_i^{(k)}$  ist der Vektor der Meßfehlervariablen in der Komponente  $k$ ,  $\mathbf{v}_k$  ist der Vektor der Konstante des Meßmodells in Komponente  $k$ ,  $\mathbf{B}_k$  ist die Matrix der Einflußgrößen für die Beziehungen zwischen den latenten  $\boldsymbol{\eta}$ -Variablen in der Komponente  $k$ ,  $\boldsymbol{\Gamma}_k$  ist die Matrix der Einflußgrößen für die Beziehungen zwischen exogenen  $x$ -Variablen und den latenten  $\boldsymbol{\eta}$ -Variablen.

Stein (1997) zeigt, daß unter den obigen Spezifikationen eines Meßmodells und eines Strukturmodells der bedingte Mittelwert  $\boldsymbol{\mu}_{ik} = \boldsymbol{\gamma}_k + \boldsymbol{\Pi}_k$  und die bedingte Kovarianzmatrix  $\boldsymbol{\Sigma}_k$  in der folgenden Weise parametrisiert werden:

$$\boldsymbol{\gamma}_k = \mathbf{v}_k,$$

$$\boldsymbol{\Pi}_k = \mathbf{A}_k (\mathbf{I} - \mathbf{B}_k)^{-1} \boldsymbol{\Gamma}_k,$$

$$\boldsymbol{\Sigma}_k = \mathbf{A}_k (\mathbf{I} - \mathbf{B}_k)^{-1} \boldsymbol{\Psi} (\mathbf{I} - \mathbf{B}_k)^{-1} \mathbf{A}_k' + \boldsymbol{\Theta}_k$$

In der letzten Gleichung sind als zusätzliche Parameter die Kovarianzmatrix  $\boldsymbol{\Psi}$  der  $\boldsymbol{\zeta}$ -Variablen und die Kovarianzmatrix  $\boldsymbol{\Theta}$  der  $\boldsymbol{\varepsilon}$ -Variablen enthalten. Sammeln wir alle Parameter der Strukturgleichungsmodelle in einem Vektor  $\boldsymbol{\vartheta}$ , so läßt sich die Suche nach unbekanntem Gruppen mit spezifischen Abhängigkeitsstrukturen als Schätzung aller Parameter des folgenden Mischverteilungsmodells auffassen:

$$f(y_i | x_i) = \pi_1 \phi(y_i; \boldsymbol{\mu}_{i1}(\boldsymbol{\vartheta}), \boldsymbol{\Sigma}_1(\boldsymbol{\vartheta})) + \pi_2 \phi(y_i; \boldsymbol{\mu}_{i2}(\boldsymbol{\vartheta}), \boldsymbol{\Sigma}_2(\boldsymbol{\vartheta})) \\ \dots + \dots \pi_K \phi(y_i; \boldsymbol{\mu}_{iK}(\boldsymbol{\vartheta}), \boldsymbol{\Sigma}_K(\boldsymbol{\vartheta})).$$

## 2.2 Parameterschätzung

Zur Schätzung der Parameter finiter Mischungen von konditionalen Mittelwert- und Kovarianzstrukturmodellen schlagen Stein (1997) sowie Arminger, Stein und Wittenberg (1999) *drei* alternative Schätzstrategien vor. Die Schätzalgorithmen sind in dem Computerprogramm zur Schätzung allgemeiner Mittelwert- und Kovarianzstrukturmodelle mit metrischen und nichtmetrischen abhängigen Variablen und Mischungen von konditional multivariat normalverteilten Variablen MECOSA 3 (Arminge/Wittenberg 1998) implementiert. Technische Einzelheiten zu den Schätzverfahren sowie Ergebnisse von Simulationsstudien, die die Anwendbarkeit und die spezifischen Vorteile der drei Algorithmen demonstrieren, finden sich in Arminger, Stein und Wittenberg (1999).

Als *erste* Schätzstrategie schlagen Stein (1997) sowie Arminger, Stein und Wittenberg (1999) zur Schätzung der mischenden Wahrscheinlichkeiten und der Parameter konditionaler Mittelwert- und Kovarianzstrukturmodelle ein *zweistufiges* Schätzverfahren vor. In der ersten Stufe werden zunächst die mischenden Wahrscheinlichkeiten  $\pi_k$  und die Anzahl der Elemente in den einzelnen Komponenten sowie die Parameter der Regression der exogenen Variablen  $x_i$  auf  $y_i$ , d.h.  $\boldsymbol{\gamma}_k$  (Vektor der Regressionskonstanten),  $\boldsymbol{\Pi}_k$  (Matrix der Regressionskoeffizienten) und  $\boldsymbol{\Sigma}_k$  (residuale bzw. bedingte Kovarianzmatrix der  $y$ -Variablen) geschätzt. Diese Schätzung erfolgt unter Verwendung eines EM-Algorithmus für Regressionsmodelle, der von Jones und McLachlan (1992) entwickelt wurde. Anschließend wird dann die asymptotische Kovarianzmatrix dieser unrestringierten Parameterschätzer geschätzt. Die Schätzfunktion der asymptotischen Kovarianzmatrix wurde von Arminger, Stein und Wittenberg (1999) abgeleitet. In der zweiten Stufe werden die im Vektor  $\boldsymbol{\vartheta}$  gesammelten Parameter der Strukturgleichungsmodelle in den einzelnen Komponenten geschätzt, die die residualen Kovarianzmatrizen  $\boldsymbol{\Sigma}_k$  parametrisieren. Dabei werden mit Hilfe der im Bereich der Analyse von Kovarianzstrukturen üblichen Minimum-Distanz-Schätzung (vgl. Browne 1982) unter Verwendung der geschätzten asymptotischen Kovarianzmatrix als Gewichtsmatrix die Parameter der Strukturgleichungsmodelle so geschätzt, daß die Abweichungen der aufgrund der Parameter

der Strukturgleichungsmodelle vorhergesagten Parameter der ersten Stufe von den tatsächlichen Schätzungen der Parameter der ersten Stufe ein Minimum ergibt.

Als *zweite* Schätzstrategie schlagen Arminger, Stein und Wittenberg (1999) einen von Jedidi, Jagpal und DeSarbo (1997) für unkonditionale Mischungen von LISREL-Modellen entwickelten direkten EM-Algorithmus vor. Mit Hilfe des direkten EM-Algorithmus werden sowohl die Parameter des Mischverteilungsmodells als auch die Parameter des Mittelwert- und Kovarianzstrukturmodells in einer Stufe geschätzt. Dies geschieht in ähnlicher Weise wie in der ersten Stufe des zweistufigen Schätzverfahrens mit dem einzigen Unterschied, daß die Maximierung im M-Schritt des EM-Algorithmus unter Berücksichtigung der in  $\mathcal{D}$  enthaltenen Strukturparameter und nicht der Parameter des konditionalen Mischverteilungsmodells gebildet wird.

Als *dritte* Schätzstrategie schlagen Arminger, Stein und Wittenberg (1999) den gradienten EM-Algorithmus von Becker, Yang und Lange (1997) vor. 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 Iteration zum E-Schritt zurückkehrt. Der Vorteil des gradienten EM-Algorithmus gegenüber dem direkten EM-Algorithmus ist die schnellere Durchführbarkeit.

### 2.3 Test auf Anzahl der Komponenten und Anpassungsindizes

Im Regelfall ist die Anzahl der Komponenten einer Mischverteilung nicht a priori durch theoretische Vorannahmen festgelegt, sondern wird in der Analyse von Mischverteilungen geschätzt. Als Test auf die Anzahl der Komponenten eines Mischverteilungsmodells würde sich eigentlich ein Likelihood-Ratio-Test anbieten, wobei die Likelihood-Ratio-Statistik  $\lambda$  die Nullhypothese  $H_0: K = G$  Gruppen gegen die Alternativhypothese  $H_1: K = G + 1$  Gruppen prüft. Wie mehrere Autoren festgestellt haben (vgl. MacLachlan/Basford 1988, Kap. 1.10), ist die Anwendung der Likelihood-Ratio-Statistik bei Mischverteilungen problematisch. Die Gründe hierfür sind zusammenfassend in Stein (1997) dargestellt. Deshalb ist man auf Ad-hoc-Methoden oder auf das parametrische Bootstrapverfahren angewiesen, um die Anzahl der Komponenten zu testen. Als Beispiel für ein Ad-hoc Verfahren ist das von Wolfe (1971) zu nennen. Wolfe



fand in einer Simulationsstudie mit zwei Komponenten, daß die Verteilung der Likelihood-Ratio-Statistik  $\lambda$  unter der Nullhypothese  $G = 1$  Komponenten durch eine  $\chi^2_d$  - Verteilung mit  $d = 2(d_1 - d_0)$  Freiheitsgraden approximiert werden kann. Die Nullhypothese wird auf dem Signifikanzniveau  $\alpha$  zurückgewiesen, wenn

$$\lambda > \chi^2_{(G+1)(d_1-d_0)}.$$

Statistisch korrekt, aber sehr aufwendig ist die Anwendung eines parametrischen Bootstrap-Verfahrens zur Schätzung der Verteilung von  $\lambda$ . Anwendungen des parametrischen Bootstraps finden sich in Stein (1997) sowie Arminger, Stein und Wittenberg (1999).

Bei der Anwendung von Anpassungsindizes zur Beurteilung der Anpassungsgüte der Kovarianzstrukturmodelle in den Gruppen muß die besondere Form der Minimum-Distanz-Schätzung für das oben beschriebene konditionale Mischverteilungsmodell berücksichtigt werden. Inkrementelle Anpassungsindizes müssen an die konditionale Modellspezifikation angepaßt werden (vgl. Stein 1997). Bei den inkrementellen Anpassungsindizes NFI (Normed Fit Index), NNFI (Nonormed Fit Index) und CFI (Comparative Fit Index), die im Rahmen der traditionellen Analyse von Kovarianzstrukturen verwendet werden, wird die Anpassung des analysierten Modells mit der des Unabhängigkeitsmodell verglichen, bei dem nur die Diagonalelemente der Kovarianzmatrix der beobachteten Variablen als von Null verschieden angenommen werden. Implizit werden dabei die Mittelwerte freigesetzt, da nur die Matrix der zweiten Momente betrachtet wird. Im Fall einer Mischverteilung werden die Regressionskoeffizienten freigesetzt, die bedingten Kovarianzmatrizen  $\Sigma_k$  werden als Diagonalmatrizen angenommen. Schließlich wird für alle Komponenten angenommen, daß die Regressionskoeffizienten gleich sind. Dies entspricht der Hypothese einer Homogenität der Wirkungen der abhängigen auf die unabhängigen Variablen.

### 3. Unbekannte Gruppierungen von Einstellungen zur Integration von Ausländern

#### 3.1 Problemstellung

Die Frage, die wir durch Anwendung der obigen Methoden beantworten wollen, bezieht sich auf die Einstellungen von Deutschen gegenüber der Integration von Ausländern. Anwendungen der Analyse von Strukturgleichungsmodellen, oder spezifischer, der Methoden des Mehrstichprobenvergleichs in Untersuchungen der Einstellungen gegenüber Ausländern haben bereits eine längere Tradition. 1984 wurden etwa die Resultate einer Test-Retest-Studie zum ALLBUS 1984, die sich mit den Einstellungen gegenüber Gastarbeitern beschäftigte, in einem speziellen Heft der Zeitschrift *Sociological Methods & Research* veröffentlicht. Einige der dort veröffentlichten Beiträge bezogen sich auf Mehrstichprobenvergleiche von Daten eines Dreiwellen-Panels (vgl. Faulbaum 1987, Porst et al. 1987). Spätere Analysen der Einstellungen zu Gastarbeitern mit den gleichen Daten bezogen sich auf die Frage, ob True-Score-Modelle oder faktorenanalytische Modelle zur Analyse von Paneldaten geeigneter sind (vgl. Jagodzinski et al. 1988, Jagodzinski et al. 1990). Neuere Arbeiten befassen sich mit der Frage, ob der Begriff *Gastarbeiter* durch den Begriff in *Deutschland lebende Ausländer* in spezifischen Items des ALLBUS ersetzt werden kann, ohne daß hierdurch die Ergebnisse von Kovarianzstrukturanalysen beeinflußt werden (vgl. Blank/Wasmer 1996). Da diese Studien zu dem Schluß kamen, daß eine solche begriffliche Substitution ohne Folgen für die Datenanalyse bleibt, verwendet der ALLBUS 1996 an Stelle des Begriffs *Gastarbeiter* den Begriff in *Deutschland lebende Ausländer*.

Entsprechend der in Abschnitt 1 dargestellten Ausgangsproblematik steht inhaltlich die Frage nach der unbekanntem Heterogenität der Einstellungen zur Integration von Ausländern im Mittelpunkt. Damit ist konkret die Frage angesprochen, ob sich bisher unbekannte Gruppierungen finden lassen, die sich in den Einstellungskonstrukten und ihren Beziehungen zueinander sowie in den Intensitäten der auf sie wirkenden exogenen Variablen wie sozioökonomische Variablen und Kontextvariablen wie Kontaktintensitäten und Ausländeranteil auf Kreisebene voneinander unterscheiden. Oft, so auch im vorliegenden Band, wird allein die Homogenität in Bezug auf das Erhebungsgebiet untersucht, während in diesem Aufsatz die Homogenität bezüglich des Erhebungsgebiets im Zusammenhang mit weiteren relevanten exogenen Variablen zur Diskussion steht, wobei bei der Suche nach unbekannter Heterogenität zunächst offen ist, ob das Er-

hebungsgebiet überhaupt eine Variable ist, die zur Unterscheidung von Gruppen im Zusammenhang mit Einstellungen zur Ausländerintegration geeignet ist. Das Erhebungsgebiet (Alte Bundesländer/Neue Bundesländer) wurde nicht als exogene Variable in das Modell eingeführt. Vielmehr wurde im Anschluß an die Identifikation der Gruppen eine einfache Kreuztabellierung mit den Variablen Erhebungsgebiet und Gruppenzugehörigkeit durchgeführt, um zu sehen, in welchen Gruppen sich die Anteile eines bestimmten Erhebungsgebiets verstärken.

### 3.2 Beschreibung des Datensatzes und Überblick über die Analyseschritte

Der Datensatz des ALLBUS 1996 enthält 3518 Fälle und 436 Variablen. Von den 3518 Fällen wurden nur die 3309 Befragten mit deutscher Staatsbürgerschaft in der Analyse einbezogen.

Als *endogene* Variablen wurden die folgenden Indikatoren für die Einstellung zur Integration von Ausländern ausgewählt:

V50 (LEBEN): Die in Deutschland lebenden Ausländer sollten ihren Lebensstil ein bißchen besser dem der Deutschen anpassen (1: Stimme überhaupt nicht zu; 7: Stimme voll und ganz zu).

V51 (ARBEIT): Wenn Arbeitsplätze knapp werden, sollte man die in Deutschland lebenden Ausländer wieder in die Heimat zurückschicken (1: Stimme überhaupt nicht zu; 7: Stimme voll und ganz zu).

V52 (POLITIK): Man sollte den in Deutschland tätigen Ausländern jede politische Betätigung in Deutschland untersagen (1: Stimme überhaupt nicht zu; 7: Stimme voll und ganz zu).

V53 (HEIRAT): Die in Deutschland lebenden Ausländer sollten sich ihre Ehepartner unter ihren eigenen Landsleuten auswählen (1: Stimme überhaupt nicht zu; 7: Stimme voll und ganz zu).

V97 (DOPPNAT): Die in Deutschland lebenden Ausländer sollten die Möglichkeit haben, die deutsche Staatsbürgerschaft zu erhalten, ohne daß sie ihre bestehende Staatsbürgerschaft aufgeben müssen, d.h. eine doppelte Staatsbürgerschaft sollte möglich sein (1: Stimme überhaupt nicht zu; 7: Stimme voll und ganz zu).

V98 (SOZLEIST): Die in Deutschland lebenden Ausländer sollten den gleichen Anspruch auf Sozialhilfe und andere Sozialleistungen haben wie die Deutschen (1: Stimme überhaupt nicht zu; 7: Stimme voll und ganz zu).

V99 (WAHLRECHT): Die in Deutschland lebenden Ausländer – egal aus welchen Ländern sie kommen – sollten das kommunale Wahlrecht haben, d.h. sie

sollten an der Wahl der Stadt- oder Gemeindeparlamente teilnehmen dürfen (1: Stimme überhaupt nicht zu; 7: Stimme voll und ganz zu).

V96 (UNTERHALT): Wie wichtig sollte es Ihrer Meinung nach bei der Vergabe der deutschen Staatsbürgerschaft sein, ob die Person für ihren Lebensunterhalt selbst aufkommen kann (1: Überhaupt nicht wichtig; 7: Sehr wichtig).

Die *exogenen* Variablen sind:

V37 (ALTER): Alter in Jahren.

V141 (SEX): Geschlecht des(der) Befragten (0: männlich, 1: weiblich).

V142 (BILD): Allgemeiner Schulabschluß (0: weniger als Fachhochschulreife; 1: mindestens Fachhochschulreife).

V114 (WIRTLAG): Wie beurteilen sie ihre eigene wirtschaftliche Lage heute? (1: sehr gut; 2: Gut; 3: Teils gut/Teils schlecht); 4: Schlecht; 5: Sehr schlecht).

V54 – V57 (Kontaktitems): Haben Sie persönlich Kontakt zu in Deutschland lebenden Ausländern, und zwar

V54 (KONFAM): ...in der eigenen Familie oder näheren Verwandtschaft (0: Ja; 1: Nein)?

V55 (KONARB): ...an Ihrem Arbeitsplatz (0: Ja; 1: Nein)?

V56 (KONNACH): ...in Ihrer Nachbarschaft (0: Ja; 1: Nein)?

V57 (KONFREUND): ...in Ihrem sonstigen Freundes- und Bekanntenkreis (0: Ja; 1: Nein)?

V435 (PROZAUSL): Prozentualer Ausländeranteil auf Kreisebene (1: bis unter 2%; 16: 30%-32%; Abstufungen in 2%-Schritten).

In der Analyse wurde als Methode der Parameterschätzung zunächst das zweistufige Verfahren (EM-Algorithmus und anschließende Minimum-Distanz-Methode; vgl. Abschnitt 2.2) angewendet. Als Begründung für die Anwendung des zweistufigen Verfahrens sind folgende Vorteile gegenüber dem einstufigen Verfahren zu nennen: In der Regel ist die Vorstellung über die Spezifikation von Mittelwert- und Kovarianzstrukturmodellen von unbekanntem Gruppen komplizierter als bei Gruppen, die von vornherein bekannt sind. Ein Vorteil des zweistufigen Verfahrens ist insbesondere darin zu sehen, daß der Forscher die unrestringierten Parameterschätzer der ersten Stufe (Regressionskonstante, Regressionskoeffizienten und residuale Kovarianzmatrix) des Schätzverfahrens zunächst inspizieren kann, um einen Einblick in die Struktur der bis dahin noch unbekanntem Gruppen zu erhalten. Dies ermöglicht ihm, Hypothesen über Beziehungen der Variablen in den einzelnen Gruppen zu formulieren. Ein weiterer Vorteil 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 Im-

plementation eines Testverfahrens zu nennen, mit dem die Anzahl der Komponenten getestet werden kann.

Nach Festlegung der Anzahl der Komponenten und Schätzung der Mittelwert- und Kovarianzstrukturmodelle für die extrahierten Gruppen wurde in einem weiteren Schritt das endgültige Modell mit Hilfe des einstufigen Verfahrens unter Verwendung des gradienten EM-Algorithmus identifiziert. Der Vorteil des einstufigen Verfahrens ist darin zu sehen, daß die Modellanpassung 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 unter Verwendung des unrestringierten Modells vorgenommen wurde. Da dieses Verfahren jedoch nur aufgrund vorgegebener Mittelwert- und Kovarianzstrukturmodelle (und damit auch aufgrund vorgegebener Anzahl der Gruppen) klassifizieren kann, ist es in der vorliegenden Problemstellung nur zur Überprüfung des endgültigen Modells eingesetzt worden.

Unter Verwendung des zweistufigen Verfahrens wurden die Anzahl der Komponenten  $K$  und deren mischende Wahrscheinlichkeiten  $\pi_k$  für die Zugehörigkeit eines Befragten zur Gruppe  $k$  bestimmt und für jede Gruppe die Regressionskoeffizienten der Regressionen aller exogenen Variablen (Alter, Geschlecht, Bildung, Beurteilung der wirtschaftlichen Lage, Items zur Messung des Kontaktes zu Ausländern sowie der prozentuale Ausländeranteil) auf die endogenen Variablen (Items zur Messung der Einstellung gegenüber Ausländern) und die Regressionskonstanten sowie die residualen Kovarianzmatrizen geschätzt. Schließlich wurden die Parameter der Kovarianzstrukturmodelle für jede Gruppe durch Anpassung an die Schätzungen der Regression der exogenen auf die endogenen Variablen bestimmt. Da in diesem Aufsatz die Strukturen der unterschiedlichen Gruppen interessieren, also die Ergebnisse der Anpassung der Strukturgleichungsmodelle im Mittelpunkt stehen, sollen im folgenden hinsichtlich der ersten Stufe nur die Ergebnisse des Tests auf die Anzahl der Gruppen und die mischenden Wahrscheinlichkeiten berichtet werden. Die Analysen wurden mit MECOSA 3 (vgl. Arminger et al. 1996) durchgeführt.

### 3.3 Diskussion der Ergebnisse

#### 3.3.1 Anzahl und Größe der Gruppen

In der hier vorgestellten Analyse wurde das Ad-hoc Verfahren von Wolfe (1971) als Test auf die Anzahl der Komponenten angewendet. Tabelle 1 zeigt die Ergebnisse des Ad-hoc Tests. Jede Zeile der Tabelle enthält die Likelihood-Funktion und die Likelihood-Ratio-Teststatistik für ein spezifisches  $K$ . Der Likelihood-Ratio-Test testet die Nullhypothese  $H_0 : K = G$  gegen die Alternativhypothese  $H_1 : K = G + 1$ . Die Tabelle zeigt, daß für alle  $K$  außer  $K = 4$  die Nullhypothese zurückgewiesen werden muß, da alle Überschreitungswahrscheinlichkeiten  $\alpha$  kleiner als das Signifikanzniveau 0.05 sind. Da für  $K = 4$  gilt, daß  $\alpha = 1.000$ , kann die Nullhypothese  $K = 4$  nicht zurückgewiesen werden. Dies bedeutet, daß die Population in bezug auf die betrachteten Variablen heterogen ist und mutmaßlich aus vier unterschiedlichen Gruppen besteht. Die geschätzten mischenden Wahrscheinlichkeiten sind  $\hat{\pi}_1 = 0.231$ ,  $\hat{\pi}_2 = 0.295$ ,  $\hat{\pi}_3 = 0.375$  und  $\hat{\pi}_4 = 0.098$ . Dies bedeutet bei insgesamt in die Analyse einbezogenen  $n=3009$  Befragten, daß die erste Gruppe aus 696 Befragten, die zweite Gruppe aus 888 Befragten, die dritte Gruppe aus 1129 Befragten und die vierte Gruppe aus 296 Befragten besteht.

Tabelle 1: Ergebnisse des Tests auf Anzahl der Komponenten

Anzahl der Komponenten	Likelihood-Funktion für $K$	Likelihood-Ratio-Teststatistik	df	$\alpha$
1	-40923.508	15555.239	232	0.000
2	-40131.075	15484.964	348	0.000
3	-39091.961	2078.230	464	0.000
4	-38875.948	432.025	580	1.000

### 3.3.2 Beschreibung der Komponenten

Einen ersten Eindruck der Komponenten erhält man, wenn man sich die unkonditionalen Mittelwerte der Variablen, also die Mittelwerte der Variablen der unkonditionalen Mischverteilungen näher ansieht. Tabelle 2 gibt für jede Gruppe einen Überblick über die unkonditionalen Mittelwerte aller Variablen.

Tabelle 2: *Unkonditionale Mittelwerte*

<i>Variablen</i>	<i>Gruppe 1</i>	<i>Gruppe 2</i>	<i>Gruppe 3</i>	<i>Gruppe 4</i>
<i><u>Exogene Variablen:</u></i>				
ALTER	39.606	49.692	49.606	46.872
SEX	0.461	0.496	0.547	0.470
BILD	0.401	0.213	0.148	0.186
WIRTLAG	2.493	2.694	2.697	2.612
KONFAM	0.795	0.860	0.896	0.835
KONARB	0.476	0.698	0.723	0.669
KONNACH	0.638	0.750	0.773	0.740
KONFREUND	0.414	0.634	0.714	0.598
PROZAUSL	4.718	3.698	3.901	4.193
<i><u>Endogene Variablen:</u></i>				
LEBEN	3.890	5.002	5.478	5.071
ARBEIT	1.745	4.144	4.403	3.126
POLITIK	2.399	3.988	4.587	3.690
HEIRAT	1.033	3.548	3.553	2.426
DOPPNAT	4.430	3.291	2.581	3.931
SOZLEIST	5.403	4.140	3.155	7.000
WAHLRECHT	4.838	3.509	2.658	4.456
UNTERHALT	4.069	4.913	7.000	7.000

Was die endogenen Variablen betrifft, so ergibt sich, daß zwei der Komponenten gewissermaßen Extreme auf einem Kontinuum der Einstellung gegenüber der Integration von Ausländern darzustellen scheinen: Gruppe 3 scheint

Ausländer eher als Fremde zu sehen, denen man höchstens einen Besucherstatus zugestehen könnte. Diese Gruppe lehnt jede Form realer Integration ab. Im Gegensatz dazu kann Gruppe 1 durch eine sehr positive Einstellung gegenüber der Integration von Ausländern charakterisiert werden. Das Muster der endogenen Mittelwerte für Gruppe 2 scheint im wesentlichen dem für Gruppe 4 zu folgen, mit dem Unterschied, daß Gruppe 4 nicht so extrem in der Zurückweisung der Ausländerintegration erscheint (niedrigere Werte auf den Variablen LEBEN, ARBEIT, POLITIK und HEIRAT).

Das Mittelwertsmuster von Gruppe 4 ist sehr ähnlich dem von Gruppe 1. Der Unterschied besteht darin, daß in Gruppe 4 noch stärker die Wichtigkeit des eigenen Lebensunterhalts betont wird und einer Egalisierung der Sozialleistungen voll zugestimmt wird. Auch die Forderung nach einer Anpassung des Lebensstils ist bei Gruppe 4 stärker ausgeprägt. Interpretiert man die Zustimmung zu der Forderung, bei der Vergabe der Staatsbürgerschaft zu berücksichtigen, ob die Person für den eigenen Lebensunterhalt aufkommen kann, als weiteren Indikator für die sozioökonomische Integration, so scheint Gruppe 4 zwar die sozioökonomische Integration zu akzeptieren, allerdings unter gleichzeitiger Inkaufnahme der gesellschaftlichen Folgen der Arbeitslosigkeit, wie sie in der vollen Zustimmung zur Egalisierung der Sozialleistungen einschließlich etwa der Sozialhilfe, der Arbeitslosenunterstützung, etc. mit denen der Deutschen zum Ausdruck kommt.

Auf der Basis der un konditionalen Mittelwerte der endogenen Variablen unterscheiden sich die Gruppen vor allem durch das Ausmaß, in dem eine Integration von Ausländern akzeptiert wird. Dabei stellt Gruppe 3 eine Gruppe dar, die durch eine starke Gegnerschaft zu einer Integration von Ausländern charakterisiert werden kann, was durch hohe Mittelwerte auf den endogenen Variablen LEBEN, ARBEIT, POLITIK und, verglichen mit den übrigen Gruppen, auch für HEIRAT sowie durch niedrige Mittelwerte auf den Variablen DOPPNAT, SOZLEIST, WAHLRECHT angezeigt wird. Wir bezeichnen diese Gruppe daher als *hoch anti-integrative Gruppe (HAI)*. Entsprechend charakterisieren wir Gruppe 2 als eine *moderat anti-integrative Gruppe (MAI)*, Gruppe 4 als eine *moderat integrative Gruppe (MI)* und Gruppe 1 als eine *hoch integrative Gruppe (HI)*. Letztere zeichnet sich im Vergleich zu den übrigen Gruppen durch niedrige mittlere Einstufungen der Items LEBEN, ARBEIT, POLITIK und HEIRAT sowie mit hohe Mittelwerte der Variablen DOPPNAT, SOZLEIST, WAHLRECHT und UNTERHALT aus.

Ein Blick auf die exogenen Variablen zeigt, daß die hoch integrative Gruppe (Gruppe 1) im Durchschnitt ein niedrigeres Alter und eine höhere Bildung aufweist sowie über mehr Kontakte zu Ausländern innerhalb der Familie, bei der



Arbeit, in der Nachbarschaft und unter Freunden und Bekannten verfügt als die übrigen Gruppen. Hinzu kommt, daß in dieser Gruppe der prozentuale Anteil der Ausländer auf Kreisebene höher ist. Im Gegensatz dazu weist die hoch anti-integrative Gruppe (Gruppe 3) im Vergleich zu den übrigen Gruppen das höchste durchschnittliche Alter, den höchsten Frauenanteil, die geringste Bildung und die geringsten Kontakte mit Ausländern auf. Der Prozentsatz der Ausländer auf Kreisebene ist geringer als der in den integrativen Gruppen.

Die moderaten Gruppen MAI und MI weisen bezüglich Alter, Bildung sowie beim Kontakt zu Ausländern größere Ähnlichkeit mit der HAI-Gruppe (Gruppe 3) als mit der hoch-integrativen Gruppe HI (Gruppe 1) auf. Allerdings verfügt die moderat integrative Gruppe (Gruppe 4) über mehr Kontakte zu Ausländern in der Familie und unter Freunden und ist gekennzeichnet durch einen höheren Prozentsatz von Ausländern auf Kreisebene.

### 3.3.3 Ergebnisse der Modellanpassung

#### 3.3.3.1 Anpassungsbeurteilung

Im folgenden werden die Ergebnisse der Minimum-Distanz-Schätzung dargestellt, bei der die Strukturparameter der Kovarianzstrukturmodelle so geschätzt werden, daß die Distanz zwischen den geschätzten Werten der Parameter der Regression der exogenen Variablen auf die endogenen Variablen und den durch die Strukturparameter vorhergesagten Werten dieser Regression ein Minimum ergeben.

Hinsichtlich aller Modelle für die einzelnen Gruppen wurde angenommen, daß die endogenen Variablen LEBEN, ARBEIT, POLITIK und HEIRAT vier Indikatoren eines Meßmodells darstellen, das die latente Variable *Einstellung zur Integration von Ausländern* mißt. Diese latente Variable hat den symbolischen Namen INTAUSL. In früheren Analysen (vgl. z.B. Faulbaum 1987) hatte sich gezeigt, daß diese genannten vier Variablen in der Tat ein gutes Meßmodell konstituieren. Weiterhin wird angenommen, daß die Variablen DOPPNAT und WAHLRECHT als zwei Indikatoren einer latenten Variablen *Einstellung zur politischen Integration* aufgefaßt werden können. Diese latente Variablen trägt den Namen INTPOL. Im weiteren Verlauf der Analysen stellte sich dann heraus, daß diese latente Variable eigentlich mit DOPPNAT identifiziert werden kann. Dennoch wurde an einem Meßmodell mit zwei Indikatoren festgehalten; d.h. WAHLRECHT wurde nicht als getrennte Variable behandelt. Eine Berücksich-

tigung der zuletzt geschilderten Version soll späteren Analysen vorbehalten bleiben. Die anderen beiden endogenen Variablen SOZLEIST und UNTERHALT wurden als separate beobachtete Variablen behandelt. Eine weitere gemeinsame Eigenschaft der betrachteten Modelle bestand schließlich darin, daß in der Matrix BETA nur der Einfluß von INTAUSL auf INTPOL freigesetzt wurde. Die Hypothese eines Einflusses von INTPOL auf INTAUSL wurde verworfen, da die INTAUSL als die allgemeinere Einstellung betrachtet wurde. Tabelle 3 enthält die Anpassungsindizes des Gesamtmodells.

*Tabelle 3: Anpassungsindizes des Gesamtmodells*

$\chi^2$	NFI	NNFI	CFI	RMSEA
1235.000	0.997	0.996	0.998	0.041
df=241				

In der Tabelle sind neben dem  $\chi^2$ -Wert des betrachteten Modells, der NFI (Normed Fit Index), NNFI (Nonnormed Fit Index), der CFI (Comparative Fit Index) und der RMSEA (Root Mean Square Error of Approximation) zu sehen. Ein Vergleich von  $\chi^2$  mit dem  $1-\alpha$  Quantil einer zentralen  $\chi^2$ -Verteilung mit  $df = 241$  Freiheitsgraden ergibt, eine Überschreitungswahrscheinlichkeit kleiner als 0,5, was zu einer Zurückweisung des Modells führen müßte. Auf der anderen Seite wäre ein anderes Ergebnis aufgrund des großen Stichprobenumfangs auch nicht zu erwarten gewesen. Angesichts der Größe von NFI, NNFI und CFI nahe 1,000 und angesichts des RMSEA, der unter dem kritischen Wert von 0,05 liegt, kann die Modellanpassung dennoch als sehr akzeptabel bezeichnet werden.

### 3.3.3.2 Komponentenspezifische Ergebnisse im Vergleich

Im folgenden sollen die Modelle der einzelnen Komponenten vorgestellt werden. Die signifikanten Effekte sind in den Tabellen durch ein \* gekennzeichnet. Zunächst sollen die Ergebnisse für die erste Komponente, also die hoch integrative Gruppe 1 (HI), diskutiert werden. Tabelle 4 enthält die Ladungen von  $\mathcal{A}_1$  (LAMBDA), die Meßfehler von  $\Theta_1$  (THETA) und die Konstanten von (NU)

des Meßmodells. Da die Kovarianzmatrizen der Meßfehler in allen Gruppen nur mit den Diagonalelementen besetzt sind, muß über die Fehlerkovarianzen nicht berichtet werden.

Tabelle 4: Meßmodell für Gruppe 1 (Hoch Integrativ) (N=696)

Items	NU	LAMBDA				THETA
		INTAUSL	UNTER- HALT	INTPOL	SOZLEIST	
LEBEN	3.522 (9.793)	0.714* (9.246)	—	—	—	2.817 (12.665)
ARBEIT	1.622 (293.454)	0.473* (7.703)	—	—	—	0.661 (9.5119)
POLITIK	2.047 (10.119)	1.000	—	—	—	1.419* (8.302)
HEIRAT	0.8107 (5.128)	—	—	—	—	—
UNTERHALT	4.076 (22.854)	—	1.000	—	—	0.010 (0.000)
DOPPNAT	4.492 (25.409)	—	—	1.000	—	1.692 (1.077)
SOZLEIST	4.949 (32.487)	—	—	—	1.000	0.010 (0.000)
WAHLRECHT	4.799 (29.292)	—	—	0.170 (1.579)	—	1.766 (6.389)

Es ist zu erkennen, daß die höchsten Ladungen des Meßmodells für die latente Variable INTAUSL bei den manifesten Indikatoren POLITIK und LEBEN auftreten. Die Ladung von HEIRAT wurde in dieser Gruppe auf 0 fixiert, da die Variable in dieser Gruppe eine Varianz von Null aufwies. Die Konnotation dieser latenten Variablen könnte gedeutet werden als Intensität der Forderung an in Deutschland lebende Ausländer, die gesellschaftliche und politische Integrität der Deutschen zu respektieren. Wie bereits oben erwähnt, kann die latente Variable INTPOL mit der gemessenen Variablen DOPPNAT identifiziert werden.

Der  $\beta$  - Koeffizient für den Einfluß von INTAUSL auf INTPOL stellte sich als nicht signifikant heraus ( $\beta_{31} = 0.121$ ;  $z = 0.335$ ). Dieses Ergebnis ist insofern nicht überraschend, als in einer Gruppe, die ohnehin durch eine sehr positive Einstellung zur Ausländerintegration gekennzeichnet ist, eine signifikante Kovariation dieser Variablen nicht mehr zu erwarten ist.

Die Tabellen 5 und 6 zeigen die  $\gamma$ -Koeffizienten für die HI-Gruppe. Man sieht, daß alle Kontaktvariablen, die wirtschaftliche Lage und das Geschlecht signifikante Effekte auf INTAUSL haben. Jede Interpretation dieser Effekte muß die obigen Skalendefinitionen berücksichtigen. So bedeutet der negative Effekt von WIRTLAG, daß mit zunehmend positiver Beurteilung der eigenen wirtschaftlichen Lage die Stärke der Ausprägung auf der latenten Variablen INTAUSL steigt, was bedeutet, daß die Forderung nach einer Respektierung der gesellschaftlichen und politischen Integrität der Deutschen verstärkt wird. Bei der Interpretation der Wirkungen der Kontaktvariablen muß bedacht werden, daß Kontakt mit dem Wert 0 vercodet ist. Ähnlich muß bei der Interpretation der Variablen SEX berücksichtigt, daß die Ausprägung "männlich" mit 0 kodiert ist.

Tabelle 5: Effekte der exogenen Kontaktvariablen in Gruppe 1 (Hoch Integrativ) (N=696)

Faktoren	GAMMA			
	KONFAM	KONARB	KONNACH	KONFREUND
INTAUSL	0.173* (4.573)	0.215* (9.544)	0.399* (9.268)	0.321* (7.999)
UNTERHALT	0.005 (0.035)	0.133 (0.963)	0.575* (3.808)	-0.242* (-10.456)
INTPOL	-0.475* (-2.100)	-0.259 (-1.179)	-0.292 (-1.176)	-0.937* (-3.777)
SOZLEIST	-0.236 (-1.624)	-0.100 (-0.615)	-0.247 (-1.451)	-0.225 (-1.643)

Tabelle 6: Effekte der exogenen Variablen (ohne Kontaktvariablen) in Gruppe I (Hoch Integrativ) (N=696)

Faktoren	GAMMA				
	ALTER	SEX	BILD	WIRTLAG	PROZAUSL
INTAUSL	0.024 (0.656)	-0.557* (-16.171)	-0.034 (-1.170)	-0.230* (-260.069)	0.050 (1.653)
UNTERHALT	-0.046 (-0.297)	-0.725 (-1.319)	0.005 (0.687)	0.128 (1.749)	-0.027 (-0.0116)
INTPOL	0.028 (0.720)	0.815* (3.377)	0.285 (1.216)	0.521* (6.024)	-0.666* (-3.530)
SOZLEIST	0.004 (0.027)	0.105* (4.149)	-0.032 (-0.352)	0.151 (1.012)	-0.582 (-1.811)

Infolgedessen bedeuten die positiven Effekte der Kontaktvariablen auf INTAUSL, daß mit abnehmendem Kontakt die Intensität der Forderung nach Erhalt der gesellschaftlichen und politischen Integrität der Deutschen zunimmt. Der negative Einfluß der Variablen SEX zeigt, daß Männer diesen Respekt stärker fordern als Frauen. Eine Inspektion der übrigen  $\gamma$ -Koeffizienten ergibt negative Effekte von KONFAM, KONFREUND und PROZAUSL auf INTPOL, was bedeutet, daß Befragte mit Kontakt zu Ausländern in der Familie und im Freundeskreis eher der Meinung sind, daß man Ausländern die doppelte Staatsbürgerschaft zugestehen sollte, als solche ohne derartige Kontakte, und daß mit geringer werdendem prozentualen Ausländeranteil auf Kreisebene die Bereitschaft, die doppelte Staatsbürgerschaft zuzugestehen, zunimmt. Festzustellen ist auch ein positiver Effekt von WIRTLAG auf INPOL, d.h. mit zunehmend negativer Einschätzung der heutigen eigenen wirtschaftlichen Lage nimmt die Bereitschaft, die doppelte Staatsbürgerschaft zuzugestehen, zu. Auch die Variable Geschlecht hat einen Einfluß in der Richtung, daß Frauen stärker als Männer für eine doppelte Staatsbürgerschaft votieren.

Was die Einflüsse der exogenen Variablen auf die endogenen beobachteten Variablen UNTERHALT und SOZLEIST angeht, so ist hier vor allem der positive Effekt von KONNACH und der negative Effekt von KONFREUND auf

UNTERHALT zu erwähnen. Dies bedeutet, daß Personen *ohne* Kontakt zu Ausländern in der Nachbarschaft und solche *mit* Kontakt zu Ausländern im Freundes- und Bekanntenkreis stärker die Meinung vertreten, daß es bei Vergabe der deutschen Staatsbürgerschaft wichtig ist, ob die Person für ihren Lebensunterhalt selbst aufkommen kann.

Tabelle 7: Meßmodell für Gruppe 3 (Hoch Anti-Integrativ) (N=1129)

Items	LAMBDA					
	NU	INTAUSL	UNTERHALT	INTPOL	SOZLEIST	THETA
LEBEN	7.546 (11.489)	-0.036 (-1.802)	—	—	—	2.844 (7.904)
ARBEIT	3.211 (343.761)	-2.297* (-4.232)	—	—	—	0.000 (0.000)
POLITIK	4.350 (8.310)	1.000	—	—	—	1.348 (3.137)
HEIRAT	2.219 (6.936)	0.681* (2.369)	—	—	—	4.765 (7.111)
UNTERHALT	6.790 (16.444)	—	1.000	—	—	—
DOPPNAT	1.115 (2.751)	—	—	1.000	—	0.000 (0.003)
SOZLEIST	2.319 (8.329)	—	—	—	1.000	0.010 (0.000)
WAHLRECHT	1.322 (3.499)	—	—	0.218 (1.391)	—	0.000 (1.850)

Vergleichen wir zunächst die eben dargestellten Ergebnisse der hoch integrativen Gruppe 1 (HI) mit denen der hoch anti-integrativen Gruppe 3 (HAI). Die Schätzungen der Parameter des Meßmodells für die HAI-Gruppe ist in Tabelle 7 wiedergegeben. Man sieht, daß die Ladungen von POLITIK und HEI-

RAT auf der latenten Variablen INTAUSL positiv sind, während ARBEIT eine hohe negative Ladung aufweist, die negative Ladung von LEBEN ist nicht signifikant. Das Meßmodell für INTAUSL ist also vollkommen verschieden von dem in der hoch integrativen Gruppe, was auf eine unbeobachtete Heterogenität schon auf der Ebene der Meßmodelle hinweist. Dies heißt aber auch, daß die latente Variable INTAUSL in dieser Gruppe eine andere Bedeutung besitzt. In der HAI-Gruppe liegt die Hauptbetonung des Faktors auf der gesellschaftlichen und politischen Isolation. Unter dieser Interpretation erscheint auch die negative Ladung von ARBEIT verständlich, da eine starke Forderung nach Isolierung der Ausländer inkompatibel mit der Zustimmung zu einem Item ist, dessen Inhalt implizit voraussetzt, daß Ausländer in Deutschland arbeiten könnten.

Vergleicht man die Interpretation der Variablen INTAUSL in der HAI-Gruppe mit der in der HI-Gruppe, so scheint also in der HAI-Gruppe eher der Grad der Forderung nach einer gesellschaftlichen und politischen Isolierung der Ausländer gemessen zu werden (positive Ladungen für LEBEN und POLITIK), während in der HI-Gruppe eher die Einstellung zum Erhalt der gesellschaftlichen und politischen Integrität der Deutschen gemessen wird. Selbstverständlich steht diese Interpretation unter dem Vorbehalt der geringen Indikatorenzahl. Im Vergleich zu diesen Unterschieden sind die Ergebnisse bezüglich der latenten Variablen INTPOL ähnlich.

Der Einfluß von  $\beta_{31}$  von INTAUSL auf INTPOL ist, wie auch in der HI-Gruppe, nicht signifikant ( $\beta_{31} = 0.004$ ;  $z = 0.057$ ). Bezüglich der Einflüsse der  $\gamma$ -Variablen (Tabellen 8 und 9) fallen im Vergleich zu der HI-Gruppe die negativen Effekte von KONFAM und KONFREUND auf INTAUSL auf. Dies bedeutet, daß Personen mit Kontakt in der Familie sowie im Freundes- und Bekanntenkreis stärker für eine Isolierung eintreten. Die Effekte der übrigen Kontaktvariablen gehen in die gleiche Richtung. Der Einfluß von SEX auf INTAUSL zeigt, daß Frauen stärker als Männer in dieser Gruppe für eine solche Isolierung eintreten.

Was die auf die Variable INTPOL bezogenen Effekte angeht, so sind ein negativer Effekt von KONNACH und ein positiver Effekt von BILD zu beobachten. Da INTPOL durch den Indikator DOPPNAT dominiert wird, heißt dies, daß Personen mit Kontakt in der Nachbarschaft stärker einer doppelten Staatsbürgerschaft zustimmen. Letzteres gilt auch mit steigender Bildung.

Weitere signifikante Effekte betreffen die positiven Wirkungen von KONNACH und BILD als auch die negativen Effekte von KONARB und KONFREUND auf SOZLEIST. Offensichtlich sind Personen mit Kontakten zu Ausländern bei der Arbeit oder im Freundeskreis stärker bereit, Ausländern die glei-

chen Sozialleistungen wie den Deutschen zuzugestehen, während der Kontakt in der Nachbarschaft den gegenteiligen Effekt hat.

Tabelle 8: Effekte der exogenen Kontaktvariablen in Gruppe 3 (Hoch Anti-Integrativ) (N=1129)

Faktoren	GAMMA			
	KONFAM	KONARB	KONNACH	KONFREUND
INTAUSL	-0.580*	0.232*	0.431*	-0.226*
	(-4.154)	(4.531)	(2.699)	(-4.387)
UNTERHALT	—	—	—	—
INTPOL	-1.454	0.379	-2.178*	-0.300
	(-1.600)	(0.354)	(-2.366)	(-0.258)
SOZLEIST	-0.100	-0.863*	0.158	-0.899*
	(-0.541)	(-3.285)	(0.464)	(-3.233)

Tabelle 9: Effekte der exogenen Variablen (ohne Kontaktvariablen) für Gruppe 3 (Hoch Anti-Integrativ) (N=1129)

Faktoren	GAMMA				
	ALTER	SEX	BILD	WIRTLAG	PROZAUSL
INTAUSL	0.118	0.145*	0.162	0.023*	0.175
	(0.734)	(2.259)	(0.977)	(2.888)	(1.042)
UNTERHALT	—	—	—	—	—
INTPOL	-0.319	0.587	0.996*	-0.135	-0.040*
	(-1.185)	(1.356)	(2.096)	(-0.726)	(-1.966)
SOZLEIST	-0.100	0.001	0.251	-0.127	2.286*
	(-0.664)	(0.0012)	(1.870)	(-0.328)	(2.636)



Die beiden anderen Gruppen sollte man mit ihren extremeren Gegenstücken vergleichen, d.h. die Gruppe 4 (MI) mit der Gruppe 1 (HI) und die Gruppe 2 (MAI) mit der Gruppe 3 (HAI). Wie ein Vergleich von Tabelle 10 (Meßmodell für MI) mit Tabelle 4 (Meßmodell für HI) zeigt, ist das Meßmodell der Variablen INTAUSL für die Gruppe MI durch eine höhere Ladung der Variablen ARBEIT relativ zur auf 1.0 fixierten Ladung von POLITIK charakterisiert. Die Ladung für HEIRAT ist in dieser Gruppe freigesetzt worden. In dieser Gruppe besteht eher als in der anderen Gruppe die Tendenz, Ausländer als Gastarbeiter und Gäste abzulehnen; d.h. die Ausländer werden in dieser Gruppe stärker unter dem Gastgesichtspunkt gesehen: Gäste kann man, wenn Arbeitsplätze knapp werden, auch wieder in die Heimat zurückschicken.

Tabelle 10: Meßmodell für Gruppe 4 (Moderat Integrativ) (N=296)

Items	NU	LAMBDA				THETA
		INTAUSL	UNTER- HALT	INTPOL	SOZLEIST	
LEBEN	3.485 (14.952)	0.714* (7.957)	—	—	—	2.728 (18.578)
ARBEIT	2.746 (784.476)	0.855* (10.092)	—	—	—	2.410 (14.562)
POLITIK	2.714 (18.521)	1.000	—	—	—	3.190 (15.577)
HEIRAT	1.268 (10.243)	0.919* (8.766)	—	—	—	3.550 (19.400)
UNTERHALT	7.000 (55.848)	—	1.000	—	—	—
DOPPNAT	4.033 (34.181)	—	—	1.000	—	0.148 (0.005)
SOZLEIST	5.347 (51.747)	—	—	—	1.000	—
WAHLRECHT	3.797 (29.278)	—	—	0.012* (2.734)	—	1.429 (4.304)

Die Wirkungen der exogenen Variablen in der Gruppe MI sind in den Tabellen 11 und 12 dargestellt. Die leichte Veränderung in der Konnotation von INTAUSL im Vergleich mit der Gruppe HI könnte für den negativen Effekt von KONNACH auf INTAUSL verantwortlich sein; d.h. die Richtung dieses Effekts wird verständlich, wenn INTAUSL primär als Tendenz verstanden wird, in Ausländern als Gäste abzulehnen. Es erscheint durchaus plausibel, daß Personen ohne Nachbarschaftskontakte in weniger stark Ausländer als Gäste ablehnen. Die geringere Varianz der Variablen ALTER in HI mag der Grund dafür sein, daß es im Vergleich zu MI in HI keinen signifikanten Effekt dieser Variablen gibt. Es gibt im Vergleich zur Gruppe HI ferner keinen Effekt der Kontaktvariablen auf INTPOL in der Gruppe MI.

Es gibt in der Gruppe MI im Vergleich zur Gruppe HI aber signifikante negative Wirkungen von WIRTLAG auf INTAUSL und von ALTER auf INTPOL. Der zuletzt genannte Effekt bedeutet, daß ältere Personen weniger bereit sind, Ausländern eine doppelte Staatsbürgerschaft zuzugestehen. Der fehlende Effekt dieser Variablen in HI könnte durch das erwähnte Varianzargument erklärt werden. Der negative Effekt der wirtschaftlichen Lage auf INTAUSL kann auf dem Hintergrund der obigen Interpretation des Konstrukts INTAUSL so interpretiert werden, daß mit zunehmend negativer Einschätzung der eigenen wirtschaftlichen Lage heute, die Tendenz abnimmt, Ausländer als Gäste abzulehnen.

*Tabelle 11: Effekte der exogenen Kontaktvariablen für Gruppe 4 (Moderat Integrativ) (N=296)*

<i>Faktoren</i>	<i>GAMMA</i>			
	<i>KONFAM</i>	<i>KONARB</i>	<i>KONNACH</i>	<i>KONFREUND</i>
INTAUSL	0.041* (1.967)	0.235* (13.040)	-0.055 (-8.506)	0.656* (10.109)
UNTERHALT	—	—	—	—
INTPOL	0.151 (0.835)	-0.166 (-0.954)	0.169 (1.043)	0.085 (0.383)
SOZLEIST	—	—	—	—

Tabelle 12: Effekte der exogenen Variablen (ohne Kontaktvariablen) für Gruppe 4 (Moderat Integrativ) (N=296)

Faktoren	GAMMA				
	ALTER	SEX	BILD	WIRTLAG	PROZAUSL
INTAUSL	0.062*	-0.455*	0.019	-0.150*	0.353*
	(2.411)	(-8.688)	(0.255)	(-48.831)	(4.362)
UNTERHALT	—	—	—	—	—
INTPOL	-0.776*	0.329*	-0.389	0.406*	-0.098*
	(-4.829)	(3.676)	(-1.602)	(12.712)	(-2.093)
SOZLEIST	—	—	—	—	—

Der Vergleich der Gruppe 2 (MAI) mit der Gruppe 3 (HAI) zeigt, daß die Meßmodelle vollständig verschieden sind. In der Gruppe MAI wird die Konnotation von INTAUSL klar dominiert durch die Variable POLITIK, gefolgt von LEBEN (vgl. Tabelle 13), d.h. die latente Variable hat die Konnotation des "In-Ruhe-Gelassen-Werdens". Höhere Positionierungen auf der latenten Variablen stehen für eine stärkere Forderung, daß Ausländer Deutsche im gesellschaftlichen und politischen Bereich nicht stören sollten, weder durch ihren Lebensstil, noch durch irgendeine politische Betätigung.

Eine bemerkenswerte Konsequenz aus der spezifischen Bedeutung des Konstrukts INTAUSL ergibt sich in der Gruppe MAI für den Einfluß von  $\beta_{31}$  von INTAUSL auf INPOL. Dieser Einfluß ist negativ und signifikant ( $\beta_{31} = -0.666$ ;  $z = -4.593$ ). Nach diesem Befund würde mit einer stärkeren Forderung nach Ungestörtheit die Zustimmung zur doppelten Staatsbürgerschaft abnehmen, ein Effekt, der plausibel erscheint.

Was die Einflüsse der exogenen Variablen anbetrifft (vgl. Tabellen 14 und 15), so sind die Effekte der Kontaktvariablen auf INTAUSL zu denen in der Gruppe HAI entgegengesetzt. Es gibt signifikante positive Effekte von KONFAM und KONFREUND auf INTAUSL, die besagen, daß Personen ohne Kontakt zu Ausländern in der Familie oder im Freundes- und Bekanntenkreis weniger stark die Forderung nach Ungestörtheit vertreten. Auf der anderen Seite haben die Variablen KONARB und KONNACH einen negativen Effekt; d.h. der

Kontakt am Arbeitsplatz und in der Nachbarschaft geht mit einer verstärkten Forderung nach gesellschaftlicher Ungestörtheit einher. Weitere erwähnenswerte Unterschiede zwischen den Gruppen 2 und 3 betreffen die negativen Effekte von KONARB und PROZAUSL auf INTPOL in der Gruppe MAI. Spezifische Effekte in MAI, die in HAI nicht beobachtbar sind, sind die positiven Effekte von KONFREUND und BILD auf SOZLEIST als auch die positiven Effekte von KONARB und KONNACH auf INTPOL.

Tabelle 13: Meßmodell für Gruppe 2 (Moderat Anti-Integrativ) (N=888)

Items	NU	LAMBDA				THETA
		INTAUSL	UNTERHALT	INTPOL	SOZLEIST	
LEBEN	1.469 (9.793)	0.725* (9.670)	—	—	—	2.537 (16.208)
ARBEIT	3.404 (293.454)	0.535* (8.838)	—	—	—	3.013 (15.455)
POLITIK	2.662 (10.119)	1.000	—	—	—	2.296 (10.057)
HEIRAT	2.262 (5.128)	0.695* (10.022)	—	—	—	3.637 (17.336)
UNTERHALT	3.948 (24.426)	—	1.000	—	—	0.010 (0.000)
DOPPNAT	4.467 (28.637)	—	—	1.000	—	0.148 (0.072)
SOZLEIST	4.564 (36.919)	—	—	—	1.000	1.632 (7.068)
WAHLRECHT	4.444 (28.218)	—	—	0.109 (1.331)	—	1.369 (5.157)

Tabelle 14: Effekte der exogenen Kontaktvariablen in Gruppe 2 (Moderat Anti-Integrativ) (N=888)

Faktoren	GAMMA			
	KONFAM	KONARB	KONNACH	KONFREUND
INTAUSL	0.672* (8.318)	-0.088* (-3.830)	-0.352* (-9.539)	0.557* (8.607)
UNTERHALT	0.032 (0.267)	0.078 (0.763)	0.009 (0.069)	0.319* (15.117)
INTPOL	0.149 (0.730)	-0.797* (-4.298)	-0.368* (-2.255)	-0.186 (-0.863)
SOZLEIST	-0.255 (-1.694)	-0.100 (-0.614)	0.061 (0.400)	-0.188 (-1.452)

Tabelle 15: Effekte der exogenen Variablen (ohne Kontaktvariablen) in Gruppe 2 (Moderat Anti-Integrativ) (N=888)

Faktoren	GAMMA				
	ALTER	SEX	BILD	WIRTLAG	PROZAUSL
INTAUSL	0.062 (1.884)	-1.106* (-9.992)	0.039 (0.406)	0.015* (4.579)	-0.057 (-0.584)
UNTERHALT	0.073 (0.564)	-0.248 (-0.514)	0.044* (9.316)	0.030 (0.488)	0.456* (2.054)
INTPOL	-0.078 (-0.7619)	0.034 (0.169)	-0.012 (-0.055)	0.408* (12.500)	-0.171* (-2.612)
SOZLEIST	-0.033 (-0.171)	0.316* (12.494)	0.034 (0.419)	-0.361* (-2.059)	0.236 (0.836)

### 3.3.3.3 Der Einfluß des Erhebungsgebiets

Um einen Einblick über mögliche Zusammenhänge zwischen dem Erhebungsgebiet (Alte Bundesländer/Neue Bundesländer) und der Zugehörigkeit zu den Gruppen 1-4 zu gewinnen, wurde eine Kreuztabellierung vorgenommen. Tabelle 16 zeigt die Resultate. Neben den absoluten Häufigkeiten, sind die Spaltenprozentage in der Tabelle zu sehen.

Tabelle 16: Kreuztabellierung zwischen Erhebungsgebiet und Gruppenzugehörigkeit

Erhebungsgebiet	Gruppenzugehörigkeit			
	Gruppe 1 (Hoch Integrativ)	Gruppe 2 (Moderat Anti-Integrativ)	Gruppe 3 (Hoch Anti-Integrativ)	Gruppe 4 (Moderat Integrativ)
<i>Alte Bundesländer</i>	503 (72.27%)	537 (60.54%)	740(65.54%)	203 (68.58%)
<i>Neue Bundesländer</i>	193 (27.73%)	351 (39.53%)	389 (34.46%)	93 (31.42%)
<i>Summe</i>	696	888	1129	296

Wie zu erkennen ist, gehören Befragte der alten Bundesländer im Vergleich aller Gruppen eher der Gruppe der Hoch Integrativen (Gruppe 1) an. Die neuen Bundesländer dagegen sind in dieser Gruppe am geringsten vertreten. In Bezug auf die übrigen Gruppen sind dagegen keine größeren Unterschiede zu erkennen.

## 4. Zusammenfassung

Die hier vorgestellte Analyse bekannter ALLBUS-Variablen der politischen und gesellschaftlichen Integration von in Deutschland lebenden Ausländern ergab den klaren Befund einer unbeobachteten Heterogenität. Es konnten vier Gruppen gefunden werden, die sich, zumindest tendenziell, durch eine unterschiedlich starke Gegnerschaft gegen eine Integration von Ausländern charakterisieren lie-

Ben. Unterschieden werden konnte eine einer Integration von Ausländern sehr positiv gegenüberstehende Gruppe, die als hoch integrativ (HI) bezeichnet wurde, eine moderat integrative Gruppe (MI), eine moderat anti-integrative Gruppe (MAI) sowie eine hoch anti-integrative Gruppe (HAI).

Allerdings zeigen die Ergebnisse, daß eine erste Charakterisierung der Gruppen durch unconditionale Mittelwerte in der Tat nur unvollständig und eventuell sogar irreführend sein kann, da die auf der Basis dieser Mittelwerte gefundenen Ähnlichkeiten noch nichts über die Ähnlichkeiten der Modelleigenschaften aussagen, insbesondere nicht über die Meßmodelle und die Konnotation der Konstruktvariablen. Die Ergebnisse haben gezeigt, daß es im Sinne der Ausgangsfragestellung (vgl. Abschnitt 1) klare Hinweise für Gruppenunterschiede in den Meßmodellen gibt, die durch die vier Indikatoren LEBEN, ARBEIT, POLITIK und HEIRAT gekennzeichnet sind. Das ursprüngliche Konstrukt der Einstellung gegenüber der Integration von Ausländern ergab in den Gruppen unterschiedliche Konnotationen; d.h. es wurden in den vier Gruppen unterschiedliche Aspekte der Integration gemessen. Wir haben es also schon auf der Ebene der Meßmodelle mit einer unbeobachteten Heterogenität zu tun. Die unterschiedlichen Konstrukte haben wir als Forderung nach Respektierung der gesellschaftlichen und politischen Integrität (HI), Forderung nach gesellschaftlicher und politischer Ungestörtheit (MAI), Forderung nach gesellschaftlicher und politischer Isolation von Ausländern (HAI) und die Tendenz, Ausländer als Gäste zu abzulehnen (MI). Damit ist die Generalisierbarkeit der in der Literatur häufig angenommenen Gültigkeit eines durch die oben genannten vier Indikatoren gekennzeichneten Meßmodells der Einstellung gegenüber Ausländern oder Gastarbeitern (attitude towards guest workers; vgl. die Beiträge in Bohrnstedt et al. 1987) oder der Ausländerfeindlichkeit in Zweifel zu ziehen.

Selbstverständlich sind diese Interpretationen aufgrund der grundsätzlichen Offenheit theoretischer Begriffe als vorläufig anzusehen. Erst die Einbeziehung zusätzlicher Indikatoren in weiteren Untersuchungen könnte diese Interpretationen weiter absichern.

Die Identifikation der Konnotationen der latenten Variablen ist für die Interpretation der Effekte in den Modellen zentral. Letztere muß stets auf die Interpretationen der Konstrukte in den Meßmodellen abgestimmt werden; d.h. es ist eine stete Abstimmung der Kohärenz der Interpretationen erforderlich, eine Aufgabe, die auf den Prozeß einer Konstruktvalidierung in den Gruppen hinausläuft.

Die Interpretationen der Effekte der exogenen Variablen waren im allgemein durchaus nicht unplausibel, obwohl es für eine plausible Interpretation in der Regel auch eine plausible entgegengesetzte Interpretation geben kann. So mag

etwa eine höhere Kontaktintensität ein bereits existierendes Vorurteil oder eine Antipathie gegenüber Ausländern verstärken oder vermindern. In diesem Zusammenhang ist es etwa interessant zu beobachten, daß in der hoch integrativen Gruppe der Sachverhalt keines Kontakts unabhängig von der Art des Kontakts die Forderung nach einer Respektierung der politischen und gesellschaftlichen Integrität der Deutschen verstärkt, also eine Förderung des Kontakts diese Forderung also vermindern könnte, während in der hoch anti-integrativen Gruppe dieser Effekt nur beim Kontakt bei der Arbeit und in der Nachbarschaft zu beobachten ist, und zwar nunmehr bezogen auf die latente Variable der Forderung nach gesellschaftlicher und politischer Isolierung. Im Fall des familiären Kontakts und des Kontakts im Freundes- und Bekanntenkreis gab es in der Gruppe HAI den umgekehrten Effekt.

Bei den gefundenen Konnotationsunterschieden handelt es sich um unterschiedliche semantische Nuancierungen, deren Auswirkungen auf die Gesamtmodelle weiter studiert werden müßten. Auch die weiteren Befunde, über die oben berichtet wurde, müssen in umfassenderen Analysen weiter überprüft werden. Nur sehr wenige der zahlreichen Variablen des ALLBUS 1996 wurden in die Analyse einbezogen. In jedem Fall hat das Verfahren aber zu Ergebnissen geführt, die auf eine klare unbeobachtete Heterogenität der ausländerbezogenen Variablen hinweisen, so daß man nicht davon ausgehen kann, daß das bekannte Meßmodell mit den Variablen LEBEN, ARBEIT, POLITIK und HEIRAT als gegenüber allen Gruppierungen invariant gültig ist.

Ein Vergleich der Gruppen hinsichtlich des Erhebungsgebiets scheint für eine größere Präsenz der alten Bundesländer in der Gruppe der Hoch Integrativen (Gruppe 1) zu sprechen.



## Literatur

- Arminger, G./P. Stein und J. Wittenberg, 1999: Mixtures of conditional mean- and covariance structure models. Eingereicht und zur Veröffentlichung akzeptiert bei Psychometrika.
- Arminger, G./J. Wittenberg und A. Schepers, 1996: MECOSA 3: A program for the analysis of general mean- and covariance structures with non-metric variables, User Guide. Frauenfeld, Switzerland: SLI-AG.
- Bacher, J., 1996: Clusteranalyse. 2. Auflage. München: Oldenbourg.
- Becker, M.P./I. Yang, und K. Lange, 1997: EM-algorithms without missing data. *Statistical Methods in Medical Research* 6: 37-53.
- Bentler, P. M./S.Y. Lee und L.J. Weng, 1987: Multiple population covariance structure analysis under arbitrary distribution theory. *Communications in Statistics-Theory and Methods* 16: 1951-1964.
- Blank, T./S. Schwarzer, 1994: Ist die Gastarbeiterskala noch zeitgemäß? Die Reformulierung der ALLBUS-Skala. *ZUMA-Nachrichten* 34: 97-115.
- Blank T./M. Wasmer, 1996: Gastarbeiter oder Ausländer? Ergebnisse des Splits mit den reformulierten Gastarbeiterfragen im ALLBUS 1994. *ZUMA-Nachrichten* 38: 45-69.
- Bollen, K. A., 1989: *Structural equations with latent variables*. New York: Wiley.
- Bohrnstedt, G.W./ P.Ph. Mohler und W. Müller, 1987: An empirical study of the reliability and stability of survey research items. *Sociological Methods & Research* 15, 3.
- Browne, M.W., 1982: Covariance Structures. In: Hawkins, D.M. (Hrsg.), *Topics in Applied Multivariate Analysis*. London: Cambridge University Press: 72-141.
- DeSarbo, W. S./W.L. Cron, 1988: A maximum likelihood methodology for clusterwise linear regression. *Journal of Classification* 5: 249-282.
- Faulbaum, F., 1987: Intergroup comparisons of latent means across waves. *Sociological Methods and Research* 15: 317-335.
- Faulbaum, F., 1990: Der internationale Vergleich von Meßmodellen unter verallgemeinerten Verteilungsbedingungen. *ZUMA-Nachrichten* 26, 56-71.
- Faulbaum, F./P. Stein, 1998: Conditional mean-and covariance structures in finite mixtures: Theory and example. In: Hox, J./E. de Leeuw (Hrsg.), *Assumptions, Robustness, and Estimation Methods in Multivariate Modeling*. Amsterdam: 173-199.
- Gehring, A./F. Boeltken, 1985: Einstellungen zu Gastarbeitern 1980 und 1984: Ein Vergleich. *ZA-Information* 17: 23-33.
- Hand, D.J., 1996: Classification and computers – shifting the focus. *COMPSTAT 96 – Proceedings in Computational Statistics*. Physika-Verlag.
- Hand, D. J., 1997: *Construction and Assessment of Classification Rules*. London: Wiley.
- Hsiao, C., 1995: Panel analysis for metric data. In: Arminger, G./C.C. Clogg und M.E. Sobel (Hrsg.), *Handbook of Statistical Modeling for the Social and Behavioral Sciences*. New York: Plenum: 77-183.
- Jäger, C., 1995: Theorie und Messung von Ausländerfeindlichkeit. Eine sozialwissenschaftliche Kritik der Forschungspraxis. (Marburger Beiträge zur Sozialwissenschaftlichen Forschung, Nr. 6).
- Jagodzinski, W./S.M. Kühnel und P. Schmidt, 1988: Is the true-score model or the factor model more appropriate? *Sociological Methods & Research*, 17: 158-164.

- Jagodzinski, W./S.M. Kühnel und P. Schmidt, 1990: Searching for parsimony: Are true-score models or factor models more appropriate? *Quality and Quantity*, 24: 447-470.
- Jedidi, K./H.S. Jagpal und W.S. DeSarbo, 1997: Finite-mixture structural equation models for response-based segmentation and unobserved heterogeneity. *Marketing Science* 16: 39-59.
- Jöreskog, K. G., 1971: Simultaneous factor analysis in several populations. *Psychometrika* 57: 409-426.
- Jones, P.N./G.J. McLachlan, 1992: Fitting finite mixture models in a regression context. *Australian Journal of Statistics* 43: 233-440.
- Kühnel, S., 1987: Ein LISREL-Test von multivariaten Mittelwertsdifferenzen bei inhomogenen Varianzen und Kovarianzen. Das Beispiel der Gastarbeiter-Items aus dem ALLBUS 1080 und 1984. *ZA-Information* 21: 52-58.
- Leenen, W.R., 1992: Ausländerfeindlichkeit in Deutschland. Politischer Rechtsruck oder Politikversagen. *Deutschlandarchiv* 10: 1039-1054.
- Loehlin, J.C., 1998: *Latent Variable Models: An Introduction to Factor, Path, and Structural Analysis*. Mahwah: Lawrence Erlbaum Associates.
- McLachlan, G. J./K.E. Basford, 1988: *Mixture Models*. New York: Marcel Dekker.
- Muthèn, B.O., 1989: Latent variable modeling in heterogeneous populations. *Psychometrika* 54, 557-585.
- Porst, R./P. Schmidt und K. Zeitfang, 1987: Comparisons of subgroups by models with multiple indicators. *Sociological Methods & Research* 15: 303-315.
- Reuband, K.-H., 1989: Wächst in der Bundesrepublik die Ausländerfeindlichkeit? *Neue Praxis* 19: 270-274.
- Sörbom, D., 1974: A general method for studying differences in factor means and factor structures between groups. *British Journal of Mathematical and Statistical Psychology* 27: 229-239.
- Stein, P., 1997: *Konstruktion und sozialwissenschaftliche Anwendung finiter Mischungen von Kovarianzstrukturmodellen*. Köln: Eul-Verlag.
- Titterton, D. M./A.F.M. Smith und U.E. Makov, 1985: *Statistical Analysis of Finite Mixture Distributions*. Chichester: Wiley.
- 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., 1994: *Finite mixtures in confirmatory factor-analytic models*. Dissertation, University of California at Los Angeles, Department of Psychology.