

**ZA –Archiv Nummer 2792**

**Wohlfahrtssurvey 1993**

**1993**

## Wohlfahrtssurvey 1993

- Methodenbericht -

Vorgelegt von Infratest Sozialforschung GmbH, München

Martin Kellner-Raum  
Helmut Kuwan  
Bernhard von Rosenblatt

Juni 1993  
925315

Inhalt	Seite
1. Vorbemerkung	3
2. Anlage der Stichprobe	4
3. Durchführung und Ergebnisse der Feldarbeit	9
3.1 Ablauf der Feldarbeit	9
3.2 Ausschöpfung der Stichprobe	9
3.3 Ergebnisse der Interviewerkontrolle	12
3.4 Der Wohlfahrtssurvey aus der Sicht der Interviewer	14
4. Gewichtung und Validierung der Stichprobe	16
4.1 Das Gewichtungsmodell	16
4.2 Der Validierungsansatz	21
4.3 Gewichtungseffekte und Validierungsergebnisse	23
<b>Anhang:</b> Anschreiben Fragebogen	

## 1. Vorbemerkung

Infratest Sozialforschung führt den Wohlfahrtssurvey 1993 im Auftrag des Wissenschaftszentrums Berlin erstmals als gesamtdeutsche Erhebung durch. Das Fragenprogramm ist in den alten und neuen Bundesländern identisch.

Der vorliegende Methodenbericht dokumentiert den Ablauf und die Ergebnisse der Feldarbeit. Er enthält ferner eine ausführliche Beschreibung der Stichprobenanlage und Stichprobenrealisierung.

Die methodischen Verfahren der Feldsteuerung und Stichprobenziehung sind weitgehend identisch mit denen vom Wohlfahrtssurvey 1988.

In Anlehnung an den ALLBUS-Methodenbericht wird die Darstellung von Gewichtungseffekten und Validierungsergebnissen in die Berichterstattung mit aufgenommen. Der Datenprüfung ist ein eigener Bericht gewidmet[1].

Der Methodenbericht wurde in Zusammenarbeit mit der Infratest Statistikabteilung (Christian von der Heyde, Sabine Foeller) erstellt.

---

[1] vgl. Infratest Sozialforschung, Martin Kellner-Raum, Wohlfahrtssurvey 1993 - Dokumentation der Datenprüfung, München, Mai 1993

## 2. Anlage der Stichprobe

### Grundgesamtheit und Auswahlgrundlagen

Die Grundgesamtheit der Untersuchung besteht aus allen zum Zeitpunkt der Befragung in Privathaushalten lebenden deutschen Personen ab 18 Jahren in der Bundesrepublik Deutschland. Die bundesweite Untersuchung wurde als Random-Route-Stichprobe von insgesamt 3.000 Befragten angelegt. Es gab dabei zwei Vorgehensweisen, nach denen die zu befragenden Personen at-random ausgewählt wurden:

- die ADM-Samples mit anschließendem Random-Route und
- das Infratest Mastersample Ost mit anschließendem Random-Route.

Beide Vorgehensweisen führen zu repräsentativen Stichproben der abzubildenden Grundgesamtheit und sind einander weitgehend ähnlich, unterscheiden sich aber im Detail voneinander.

### Erste Auswahlstufe

Die Auswahlseinheiten der ersten Stufe sind sog. Sample-Points. Das sind - im theoretischen Stichprobenansatz - die Adressenklumpen, die von einem Interviewer im Verlauf der Feldarbeit bearbeitet werden sollen.

Vor der Auswahl des Sample-Points werden zunächst alle Gemeinden der **Bundesrepublik West**[2] (ca. 8.500) nach regionalen Kriterien und nach ihrer Größe geschichtet.

---

[2] Im folgenden werden mit "BRD West" die alten Bundesländer und Berlin (West) bezeichnet, mit "BRD Ost" die neuen Länder und Berlin (Ost).

Als regionale Kriterien werden Bundesland, Regierungsbezirk und Kreis verwendet. Als Kriterium der Größe wird nicht die politische Gemeindegröße verwendet, sondern die Differenzierung der Gemeinden nach BIK-Typen, die zusätzlich zur Größe der Gemeinden deren Zentralitätsgrad einschließt.[3]

Auf diese Weise lassen sich für die **BRD-West** rund 1.050 nichtleere Schichten bilden. Das Schichtgewicht besteht aus der Summe der Privathaushalte in den zugehörigen Gemeinden. Für die West-Stichprobe wurden insgesamt 370 Sample-Points proportional zur Verteilung der Haushalte auf Schichten verteilt.

Die Verteilung der Sample-Points auf die Schichten erfolgt mittels einer Allokationsrechnung: Proportional zur Verteilung der Privathaushalte werden 370 Sample-Points auf das Schicht-Tableau verteilt. Anschließend werden die gebrochenen Sample-Point-Zahlen auf ganze Zahlen gerundet, mit Rundungswahrscheinlichkeiten proportional zur Größe der Nachkommastellen (die Rundung ist also nicht kaufmännisch-deterministisch). Die Rundung erfolgt mehrstufig von "außen" nach "innen", d.h. von den höheren Aggregationsebenen Land bzw. BIK-Typ über die Bezirke bzw. BIK-Typen je Land, Bezirk auf die Schichten Kreis x BIK. Auswahl-gesamtheit für die reale Auswahl der Sample Points sind dann aber nicht mehr alle

- 
- [3] Die Abgrenzung der Agglomerationsgebiete erfolgte bisher nach einem Modell, das von dem Statistiker Boustedt in den 50er Jahren entwickelt wurde. Dieses Modell wurde letztmals auf der Basis der Volkszählung 1970 revidiert. Eine Neufassung war also dringend nötig. Diese Revision hat das Beratungsinstitut BIK, Hamburg, übernommen. Seit 1.1.93 sind die Infratest ADM-Netze und das Infratest Mastersample Ost an BIK adaptiert.

Es werden nach BIK folgende Gemeindetypen unterschieden:

500.000 und mehr	(Kernbereich)
500.000 und mehr	(Randzone)
100.000 bis unter 500.000	(Kernbereich)
100.000 bis unter 500.000	(Randzone)
50.000 bis unter 100.000	(Kernbereich)
50.000 bis unter 100.000	(Randzone)
20.000 bis unter 50.000	
5.000 bis unter 20.000	
2.000 bis unter 5.000	
unter 2.000	

Die Größenklassen beziehen sich auf die Einwohnerzahl der jeweiligen Stadtregion. Bei Gemeinden, die keiner Stadtregion angehören, entspricht die Größenklasse der Einwohnerzahl in den politischen Grenzen. Unter Randzone wird die Zusammenfassung des Verdichtungsbereichs und des Übergangsbereichs verstanden.

Gemeinden einer Schicht, sondern die uns zur Verfügung stehenden sog. ADM-Netze.[4] Aus den darin für unser Institut insgesamt verfügbaren rund 3.500 Stimmbezirken wird in systematischer Zufallsauswahl die je Schicht notwendige Zahl von Sample-Points ausgewählt.

Der Übergang von der Gemeinde auf den Sample-Point wird bereits bei der Erstellung der ADM-Stichproben vollzogen, die ebenfalls in der oben geschilderten Weise gebildet werden. Auswahlseinheiten dieser Stichprobenrealisierung sind die (ggf. synthetisierten) Stimmbezirke, d.h. eine Gemeinde fällt dadurch in die Stichprobe, daß einer der in ihr enthaltenen Stimmbezirke ausgewählt wird.

Für die **BRD-Ost** wurde die Stichprobe aus dem Anfang 1990 von Infratest aufgebauten Mastersample Ost gezogen. Der Aufbau des Mastersamples folgte den inzwischen standardisierten Prinzipien der Bildung von ADM-Netzen: Auf die nach Kreisen und den damals verfügbaren 7 Gemeindegrößenklassen[5] geschichtete Grundgesamtheit wurden proportional zur Zahl der Privathaushalte (Schichtgewicht geschätzt aus der Haushaltsverteilung 1981 und der Bevölkerungsverteilung 1989) insgesamt 1.000 Sample-Points verteilt. Eine Gemeinde qualifizierte sich als Mastersample-Gemeinde, wenn mindestens ein Sample-Point auf sie entfiel.

In diesen Gemeinden wurden die Startadressen für den Random-Route-Begehungsweg gezogen.

- 
- [4] Die ADM-Netze sind die den einzelnen Instituten zur Verfügung stehenden Teile der sog. ADM-Stichprobe. Dies ist eine haushaltsproportionale Stichprobe von Stimmbezirken der jeweils letzten Bundestagswahl (in Berlin der Wahl zum Abgeordnetenhaus).

Infratest stehen rund 3.500 Stimmbezirke aus dieser Stichprobe zur Verfügung. Sie dienen als Basis für die Ziehung repräsentativer Haushaltsstichproben. Allerdings werden bei aktuellen Stichproben nicht vollständige Netze des ADM-Samples übernommen (die aus 210 Stimmbezirken bestehen), sondern es wird eine jeweils neu geschichtete Zufallsauswahl aus allen verfügbaren Stimmbezirken gebildet.

- [5] Es werden folgende Gemeindegrößenklassen unterschieden:

500.000	und mehr	
100.000	bis unter	500.000
50.000	bis unter	100.000
20.000	bis unter	50.000
5.000	bis unter	20.000
2.000	bis unter	5.000
	unter	2.000

Die aktuelle Stichprobenbildung im Osten erfolgte analog der Stichprobenbildung West:

Die ca. 7.500 Gemeinden der BRD-Ost wurden nach Kreisen und BIK-Typen geschichtet. Aus den so entstandenen rund 700 nichtleeren Schichten wurden insgesamt 186 Sample-Points proportional zur Haushaltsverteilung bestimmt und entsprechend in systematischer Zufallsauswahl die Startadressen aus dem Mastersample Ost gezogen.

### **Zweite Auswahlstufe**

Die Auswahlseinheiten der zweiten Stufe sind Privathaushalte.

Von der Startadresse (die selbst nicht befragt werden darf) ausgehend wird ein Weg - formal - beschrieben, der eine vom Interviewer unbeeinflusste Adressenauswahl gewährleistet. Die wichtigsten dieser Vorgaben lauten (jeweils ausgehend von der vorgegebenen Startadresse):

- In absteigender Hausnummernfolge auf der gleichen Straßenseite.
- In Mehrfamilienhäusern von oben nach unten, von vorne nach hinten.
- Im gleichen Stockwerk von rechts nach links.

Der Interviewer listet auf diesem Weg jeden dritten Haushalt in seiner Adressenliste auf, bis die vorgegebene Zahl von 9 Adressen erreicht ist. Mit diesen Haushalten hat der Interviewer Kontakt aufzunehmen, um in Erfahrung zu bringen, ob eine (oder mehrere) Zielperson(en) im Haushalt leben. Falls ja, gilt dieser Haushalt als Befragungshaushalt (Brutto).

### **Dritte Auswahlstufe**

Auswahlseinheiten der dritten Stufe sind die Zielpersonen. Die Auswahl erfolgt mittels eines sog. Schwedenschlüssels, der jeder Person im Haushalt, die zur Zielgruppe gehört, die gleiche Auswahlchance zuordnet und dazu führt, daß subjektive Einflüsse auf die Auswahl ausgeschlossen sind.



Zu diesem Zweck ist bei jedem Adreßfeld des Kontaktprotokolls der Auswahl Schlüssel - in ständiger systematischer Rotation - vorgedruckt. Der Interviewer listet in jedem angelaufenen Haushalt die Zielpersonen dem Alter nach geordnet (die älteste Person zuerst) auf. Der Auswahl Schlüssel bestimmt dann, die wievielte Person dieser Reihe interviewt werden muß. Damit ist sichergestellt, daß jede Person im Haushalt die gleiche Chance hat, als Zielperson ausgewählt zu werden.

### **3. Durchführung und Ergebnisse der Feldarbeit**

#### **3.1 Ablauf der Feldarbeit**

Am 18.2.1993 wurde in den alten und neuen Bundesländern mit der Feldarbeit begonnen. Die Feldphase einschließlich Nachbearbeitungszeit endete am 14.4.93. Die gesamte Feldzeit erstreckte sich somit über acht Wochen.

In den neuen Bundesländern und in Ost-Berlin war Infratest Burke Berlin mit der Durchführung der Feldarbeiten betraut.

Die Nachbearbeitung orientierte sich an der Soll-Ist-Verteilung der bis dahin vorliegenden Interviews nach Bundesländern, Regierungsbezirken und Gemeindegrößenklassen und diente damit der Verbesserung der regionalen Stichprobenstruktur.

Zusätzlich wurden Points mit "Totalausfällen" [6] an andere Interviewer vergeben, soweit sie rechtzeitig vor Feldende gemeldet wurden.

#### **3.2 Ausschöpfung der Stichprobe**

Übersicht 1 zeigt die erreichte Ausschöpfung der Stichprobe. Nach Abschluß der Feldarbeit lagen insgesamt 3.062 Interviews vor (Netto-Stichprobe). Der angestrebte Sollwert von 3.000 wurde damit leicht überschritten. Die Ausschöpfung der Stichprobe liegt bei 63,4%. In den alten Bundesländern und West-Berlin wurden 62,9% Ausschöpfung realisiert, in den neuen Bundesländern und Ost-Berlin liegt der Wert mit 64,4% etwas höher.

Die in den alten Bundesländern realisierte Stichprobenausschöpfung liegt damit auf dem gleichen Niveau wie beim Wohlfahrtssurvey 1988.

---

[6] Totalausfälle sind bedingt durch Ausscheiden, Krankheit oder andere Hinderungsgründe des Interviewers. Soweit eine rechtzeitige Mitteilung erfolgt, wird der Point nach Möglichkeit einem anderen verfügbaren und geeigneten Interviewer zugewiesen. Bei Totalausfällen ist die Nachbearbeitung praktisch eine Erstbearbeitung. Ausgefallene Points werden an andere Interviewer übergeben. Die Nachbearbeitung beinhaltet also weder eine Neuziehung von Sample Points noch eine Auflistung zusätzlicher Adressen auf dem Begehungsweg.

## Übersicht 1: Ausschöpfung der Stichprobe

	West		Ost		Gesamt	
	abs.	%	abs.	%	abs.	%
Samplepoints	370		186		556	
Adressen pro Point	9		9		9	
<b>Bruttoansatz</b>	3.330	100	1.674	100	5004	100
Totalausfälle von Samplepoints	54	1.6	90	5.4	144	2.9
Keine Person in der Zielgruppe	17	0.5	1	0.1	18	0.4
Wohnung nicht bewohnt	4	0.1	5	0.3	9	0.2
<b>Verbleibende Adressen</b>	3.255	100.0	1.578	100.0	4.833	100.0
Auskunft über Zielperson verweigert	117	3.6	25	1.6	142	2.9
Im Haushalt niemand angetroffen	251	7.7	110	7.0	361	7.5
Zielperson nicht angetroffen	39	1.2	21	1.3	60	1.2
Zielperson krank oder geistig nicht in der Lage	69	2.1	43	2.7	112	2.3
Zielperson verreist, im Urlaub	85	2.6	43	2.7	128	2.6
Zielperson verweigert aus Zeitgründen	186	5.7	65	4.1	251	5.2
Zielperson verweigert aus anderen Gründen	452	13.9	239	15.1	691	14.3
Sonstige Ausfallgründe	<u>10</u>	<u>0.3</u>	<u>16</u>	<u>1.0</u>	<u>26</u>	<u>0.5</u>
Ausfälle gesamt	1.209	37.1	562	35.6	1.771	36.6
<b>Ausgewertete Interviews</b>	2.046	62.9	1.016	64.4	3.062	63.4

Wie in Kapitel 2 dargestellt, umfaßt der Stichprobenplan 370 Points West und 186 Points Ost. In Übersicht 1 sind als Brutto-Ansatz 3.330 Adressen (West) bzw. 1.674 Adressen (Ost) ausgewiesen. In der BRD West sind 6 Points durch Totalausfälle ausgefallen, in der BRD Ost waren es 10 Points.[7]

### Einsatz der Interviewer

Beim Wohlfahrtssurvey 1993 wurden insgesamt 499 Interviewer eingesetzt, davon 322 im Westen und 177 im Osten.

Übersicht 2 zeigt die Zahl der Interviews pro Interviewer.

**Übersicht 2:** Anzahl der Interviews pro Interviewer

Anzahl der Interviews pro Interviewer	Anzahl der Interviewer (absolut)
1 bis 3	110
4 bis 6	202
7 bis 9	163
10 bis 18	17
19 bis 27	4
28 bis 37	3

95% der Interviewer haben zwischen einem und neun Interviews durchgeführt. 5% haben 10 und mehr Interviews (max. 37) durchgeführt, waren also für mehr als einen Point eingesetzt.

Der Durchschnittswert lag bei 6 realisierten Interviews pro Interviewer. Übersicht 2 zeigt die Streuung.

- 
- (7) Die gegenüber dem Westen höhere Ausfallquote auf Pointebene deckt sich mit unseren Erfahrungen aus anderen Studien: Die Gründe liegen in der geringeren telefonischen Erreichbarkeit und Mobilität der Interviewer und damit einer erschwerten Feldsteuerung. Bei der Untersuchungsanlage wurde dies berücksichtigt.

### 3.3 Ergebnisse der Interviewerkontrolle

#### Interviewerkontrolle West

Von den 2.046 in den alten Bundesländern durchgeführten Interviews wurden 531 mit Kontrollkarten schriftlich nachgeprüft. Weitere 32 Interviews wurden telefonisch kontrolliert. Insgesamt sind also 563 Interviews kontrolliert worden, das entspricht einer Kontrollrate von 28%. In die systematische Kontrolle wurden 92 Interviewer einbezogen. Der Rücklauf (einschließlich telefonischer Kontrolle) entspricht mit 72% dem Durchschnitt aller kontrollierten Studien (West).

In keinem Fall wurde die Durchführung eines Interviews in Abrede gestellt. In vier Fällen stellte es sich heraus, daß die falsche Zielperson im Haushalt befragt wurde. Die betreffenden Interviewer wurden durch ihre Chefinterviewer nachgeschult. Die betreffenden Interviews sind laut Qualitätskontrolle vollständig und verblieben im Datensatz. In sechs Fällen ergaben sich aus den Angaben der befragten Zielpersonen hinsichtlich der Interviewdauer Verdachtsmomente auf Kurzinterviews, die jedoch telefonisch bzw. durch nochmaliges schriftliches Nachfragen bei den Zielpersonen zu Gunsten der Interviewer geklärt werden konnten.

Auf die Frage, wie gerne das Interview gegeben wurde, antworteten die Zielpersonen:

	1993	1988
Gerne	64 %	58 %
Zögernd	29 %	34 %
Ungern	4 %	8 %
Keine Angabe	3 %	-

Dieses Ergebnis ist ein deutlicher Hinweis darauf, daß der Fragebogen von 1993 bei den Zielpersonen eine höhere Zustimmung gefunden hat als der Fragebogen von 1988.

### Interviewerkontrolle Ost

In den neuen Bundesländern wurden 1.016 Interviews durchgeführt. Die Interviewerkontrolle erfolgte schriftlich anhand von Kontrollkarten. In die Kontrolle wurden 42 Interviewer mit 248 durchgeführten Interviews einbezogen. Dies entspricht einer Kontrollrate von 24 % der Interviews.

Der Rücklauf der Kontrollkarten blieb mit 133 Karten (54 %) unter dem Institutsdurchschnitt (65 %). 129 von 133 kontrollierten Interviews wurden bestätigt. In vier Fällen wurde die Kontrollkarte mit dem Vermerk "unzustellbar" zurückgesandt. Die betreffenden Interviewer (3) wurden mit der Bitte um Klärung diese Unstimmigkeiten angeschrieben. Nach einer Adressrecherche konnten die Unstimmigkeiten zugunsten der Interviewer bereinigt werden.

Auf die Frage, wie gerne das Interview gegeben wurde, ergab sich in den neuen Bundesländern folgendes Bild:

	1993
Gerne	62 %
Zögernd	34 %
Ungern	3 %
Keine Angabe	1 %

Das Ergebnis zeigt keinen wesentlichen Unterschied zum Westen.

### 3.4 Der Wohlfahrtssurvey aus der Sicht der Interviewer

Die Infratest Interviewerführung erhebt routinemäßig von den Interviewern anhand sog. "Projektbewertungsskalen" (s. Anhang) Informationen über durchgeführte Projekte.

Mit diesem Instrument können Schwierigkeiten in der Durchführung von Projekten im Feld studienbezogen erhoben und mit dem Institutsdurchschnitt über alle in einem Jahr durchgeführten Projekte in Relation gesetzt werden.

Bei der Projektbewertung werden von den eingesetzten Interviewern fünf Aspekte nach dem jeweiligen Schwierigkeitsgrad beurteilt:[8]

- Thema des Interviews
- Technik des Fragebogens
- Dauer des Interviews
- Fragenverständnis
- Bereitschaft zum Interview.

Übersicht 3 zeigt die Durchschnittswerte der Schwierigkeit nach Themenfeldern für die alten Bundesländer. Es wird deutlich, daß die Bewertung des Wohlfahrtssurvey 1993 für die alten Bundesländer nach allen Kriterien etwas negativer ist als der Durchschnitt der vergleichbaren Infratest Studien der letzten Jahre. Die im Vergleich etwas schlechtere Bewertung ist allerdings für eine Studie mit einer Befragungsdauer von fast 70 Minuten nicht ungewöhnlich. Die größten Abweichungen zeigen sich dann auch bei der Interviewdauer, der Interviewbereitschaft und beim Thema des Interviews.

Leider liegen bisher für die neuen Bundesländer keine Vergleichswerte aus durchgeführten Studien vor. Die Aufnahme und Auswertung der Projektbewertungen befinden sich im Aufbau.

---

[8] Die Interviewer beurteilen die Schwierigkeiten bei der Durchführung dieses Projekts mittels einer Skala mit fünf möglichen Vorgaben:  
 0 = Keine Schwierigkeit; 1 = Geringe Schwierigkeit;  
 2 = Mittlere Schwierigkeit; 3 = Große Schwierigkeit;  
 4 = Sehr große Schwierigkeit

**Übersicht 3:** Durchschnittswerte der Schwierigkeit nach Themenfeldern\*)  
Ergebnisse der Projektbewertung für die alten Bundesländer

Themenfeld	Indexwert Wohlfahrtssurvey alte Bundesländer	durchschnittl. Indexwert
Thema des Interviews	0.8	0.5
Technik des Fragebogens	0.5	0.4
Dauer des Interviews	1.5	1.2
Fragenverständnis	0.5	0.5
Bereitschaft zum Interview	1.9	1.4
Gesamtschwierig- keitsgrad	1.0	0.8

\*) Bezogen auf alle über die Projektbewertungsskala ausgewerteten Infratest-Studien der letzten zwei Jahre in den alten Bundesländern.

Auf einer Skala von

- 0 = Keine Schwierigkeit
- 1 = Geringe Schwierigkeit
- 2 = Mittlere Schwierigkeit
- 3 = Große Schwierigkeit
- 4 = Sehr große Schwierigkeit



#### 4. Gewichtung und Validierung der Stichprobe

Das von Infratest verwendete mehrstufige Gewichtungsverfahren wird im folgenden in der Reihenfolge der Gewichtungsschritte dargestellt. [9]

##### 4.1 Das Gewichtungsmodell

Die durchgeführten Interviews verteilen sich in der Regel nicht proportional zur jeweiligen Grundgesamtheit. Das ist sowohl im Studiendesign begründet, als auch im Verhalten der Befragten und in ihrer Bereitschaft, ein Interview zu geben (vgl. Ausfallgründe lt. Rücklaufmeldung).

Es ist Aufgabe der Gewichtung, designbedingte und ausfallbedingte Disproportionalitäten auszugleichen.

##### Regionalverteilung

In dieser ersten Gewichtungsstufe wird die regionale Verteilung der Interviews proportionalisiert. Basis der Proportionalisierung ist die über die Bevölkerungsfortschreibung geschätzte Zahl der Privathaushalte nach BIK-Typen bzw. Gemeindegrößenklassen, kombiniert mit der Verteilung nach Ländern. Ergebnis dieses Gewichtungsschritts ist eine faktorielle Angleichung der regionalen Struktur der Nettostichprobe an die Schichtungstableaus des ursprünglichen Stichprobenplans.

---

[9] Das Gewichtungsverfahren entspricht dem im ALLBUS 91 und im ALLBUS 92 verwendeten Verfahren. Kapitel 4.1 repliziert daher im wesentlichen den ALLBUS Methodenbericht. Der mehrstufige Gewichtungsvorgang erfolgt kumulativ, das heißt, der nach einem Gewichtungsschritt errechnete Faktor geht als Eingangsfaktor in den jeweils nachfolgenden Gewichtungsschritt ein. (Infratest Sozialforschung, M. Kellner-Raum, H. Kuwan, B.v.Rosenblatt: Methodenbericht zum ALLBUS/ISSP 1992, München 1992)

$$RF_i = \frac{N_{ik} * \sum n_{ik}}{\sum N_{ik} * n_{ik}} \quad \text{wobei}$$

$RF_i$  = Regionalfaktor des Interviews  $i$

$N_{ik}$  = über Fortschreibung geschätzte Zahl der Privathaushalte im Bundesland  $i$  und Boustedtyp (bzw. Gemeindegrößenklasse)  $k$ .

$n_{ik}$  = Zahl der Interviews im Land  $i$  in Gemeinden des Boustedtyps (bzw. der Gemeindegrößenklasse)  $k$ .

### Designgewichtung (Umwandlung)

Das beschriebene Stichprobendesign führt zu einer repräsentativen Haushaltsstichprobe. Die Auswahlwahrscheinlichkeiten der einzelnen Personen sind abhängig von der Haushaltsgröße, d.h. der Zahl der potentiellen Zielpersonen im Haushalt. Je größer diese Zahl ist, desto geringer ist die Auswahlwahrscheinlichkeit für das einzelne Haushaltsmitglied. Die Stichprobe ist demnach - wenn sie als repräsentative Stichprobe von Personen dienen soll - disproportional bezüglich der Haushaltsgröße.

Um diese Disproportionalität auszugleichen, wird die Stichprobe im nachhinein einer sog. Designgewichtung unterzogen. Dabei werden die Auswahlchancen dadurch angepaßt, daß jedes Interview mit der Haushaltsgröße (d.h. der Zahl der Zielpersonen im Haushalt) multipliziert wird. Zusätzlich werden alle Interviews mit einem einheitlichen Faktor multipliziert, der die designgewichtete Interviewzahl auf die Zahl der tatsächlich durchgeführten Interviews reduziert.

Durch dieses Transformation wird die repräsentative Haushaltsstichprobe in eine repräsentative Personenstichprobe umgewandelt.

$$DF_i = ZP_i * RF_i * C_H \quad \text{wobei}$$

$DF_i$  = Designfaktor des Interviews  $i$

$ZP_i$  = Zahl der Zielpersonen im Haushalt  $i$  der Haushaltsstichprobe

$n$  = Fallzahl der Stichprobe

$$C_H = \frac{n}{\sum (ZP_i * RF_i)}$$

### Antreffbarkeit

Personen mit unterschiedlichem Mobilitätsgrad, d.h. aber auch unterschiedlicher Chance, vom Interviewer angetroffen zu werden, verteilen sich nicht gleichmäßig über die Grundgesamtheit. Sie sind wegen dieser Chancenungleichheit in der Stichprobe unterschiedlich repräsentiert. Diese Ungleichgewichtigkeit wird durch faktorielle Anpassung (reziproker Wert der Antreffbarkeit) in der zweiten Gewichtungstufe ausgeglichen.

Das verwendete Modell der Antreffbarkeitsgewichtung wurde von Politz/Simmons entwickelt. Es benutzt die im Interview erfragte Antreffbarkeitschance der befragten Person als Hilfsindikator zur Abschätzung der Befragungsausfälle, die durch Nichtantreffbarkeit bedingt sind. Es handelt sich also um eine (auf den Antreffbarkeitsaspekt begrenzte) Modellierung von Faktoren, die die Wahrscheinlichkeit der Realisierung eines Interviews bestimmen.

$$AF_i = \frac{4}{T_i + 1} * DF_i * C_A \quad \text{wobei}$$

$AF_i$  = Antreffbarkeitsfaktor des Interviews  $i$

$T_i$  = Zahl der Tage, an denen die Zielperson antreffbar war, mit  $T_i = 3$  für 3 oder mehr Tage antreffbar.

$$C_A = \frac{n}{\sum \left( \frac{4}{T_i + 1} * DF_i \right)}$$

### Demographie

In dieser vierten Gewichtungstufe wird die Verteilung der Interviews nach Altersgruppen und Geschlecht proportionalisiert. Basis der Gewichtung sind die entsprechenden Daten der Bevölkerungsfortschreibung.

Hier wird also die realisierte Nettostichprobe an bekannte Verteilungen einer Referenzstatistik angepaßt.

$$EF_{i(jag)} = \frac{N_{jag} * \sum n_{jag}}{\sum N_{jag} * n_{jag}} \quad \text{wobei}$$

$EF_i$  = Endfaktor des Interviews  $i$

$N_{jag}$  = Fortgeschriebene Wohnbevölkerung im Land  $j$  in der Altersgruppe  $a$  mit Geschlecht  $g$

$n_{jag}$  = Gewichtete Fallzahl im Bundesland  $j$  in der Altersklasse  $a$  mit dem Geschlecht  $g$  =  $\sum_i AF_{ijag}$

Ergebnis dieses Gewichtungsschrittes ist eine repräsentative Personenstichprobe, die in der demographischen Struktur nach Altersgruppen und Geschlecht mit der amtlichen Bevölkerungsstatistik übereinstimmt. Der entsprechende Gewichtungsfaktor wird im Datensatz jedes Befragten als "Personenfaktor" ausgewiesen.

### Haushaltsgewichtung (Rückumwandlung)

Für haushaltsbezogene Merkmale (z.B. Haushaltsgröße und -zusammensetzung) ist es oft notwendig oder wünschenswert, nicht eine Personenstichprobe, sondern eine Haushaltsstichprobe zugrunde zu legen. Auch die verfügbaren Referenzstatistiken beziehen sich hier in der Regel auf Haushalte und nicht auf Personen.

Wie oben erläutert, führt das Stichprobendesign des Random-Route-Verfahrens zunächst zu einer Haushaltsstichprobe. Die ungewichteten Zahlen könnten also als Verteilungen einer Haushaltsstichprobe interpretiert werden. Allerdings sind hier die Anpassungen der Stichprobe an die Bevölkerungsstruktur laut amtlicher Statistik noch nicht vorgenommen. Will man konsistente Verteilungen auf der Personen- und Haushaltsebene produzieren, muß man die Stichprobe nach den vorgenommenen Gewichtungsschritten in eine Haushaltsstichprobe re-transformieren.

Diese Rückumwandlung erfolgt auf dieselbe Weise - nur in umgekehrter Richtung - wie vorher die Umwandlung (Designgewichtung), also anhand der Zahl der Zielpersonen im Haushalt.

Ergebnis dieses Gewichtungsschrittes ist eine repräsentative Haushaltsstichprobe. Der entsprechende Gewichtungsfaktor wird im Datensatz jedes Befragten als

"Haushaltsfaktor" ausgewiesen. Zu beachten ist, daß eine Verwendung dieses Faktors nur für haushaltsbezogene Variablen Sinn macht.

### Ost-/West-Proportionalisierung

Bis zu diesem Punkt wurden die WEST- und OST-Stichprobe getrennt gewichtet, so daß sie auch getrennt voneinander ausgewertet werden können.

Um eine gesamtdeutsche Auswertung durchführen zu können, müssen die beiden disproportional angelegten Teilstichproben in einem letzten Gewichtungsschritt auf die tatsächlichen Zahlenverhältnisse proportionalisiert werden:

$$GF_i = EF_i * \frac{N_w}{n_w} * \frac{n_w + n_o}{N_w + N_o} \quad \text{für die Interviews der West-Stichprobe}$$

$$GF_i = EF_i * \frac{N_o}{n_o} * \frac{n_w + n_o}{N_w + N_o} \quad \text{für die Interviews der Ost-Stichprobe,}$$

wobei

$GF_i$  = gemeinsamer Faktor

$N_w$  = fortgeschr. Wohnbevölkerung ab 18 in den alten Bundesländern

$N_o$  = fortgeschr. Wohnbevölkerung ab 18 in den neuen Bundesländern

$n_w$  = Anzahl der Interviews in den alten Bundesländern

$n_o$  = Anzahl der Interviews in den neuen Bundesländern

Als Ergebnis der insgesamt vorgenommenen Gewichtungsschritte liegt ein Datensatz vor, der

- für beide Teilstichproben und
- für die Gesamtstichprobe

in den Merkmalen Bundesland, BIK-Typ (bzw. Gemeindegrößenklasse), Altersgruppe und Geschlecht proportionale Verteilungen zu den entsprechenden Strukturen der Grundgesamtheit, wie sie in der amtlichen Statistik nachgewiesen werden, aufweist.

## 4.2 Der Validierungsansatz

Mit dem folgenden Kapitel wird der Versuch unternommen, einen Beitrag zur Beurteilung der Stichprobenqualität zu leisten. Die Repräsentativität der Stichprobe wird im Vergleich der Stichprobenstruktur mit anderen Datenquellen (Referenzstatistiken) bewertet.[10]

Ausgewählte sozio-demographische und regionale Merkmale der beiden Teilstichproben West und Ost werden im folgenden mit der vorhandenen amtlichen Statistik verglichen.

Dieser Vergleich setzt voraus, daß zum einen in beiden Datenquellen gleiche Merkmale vorhanden sind, zum anderen Effekte unterschiedlicher Frage- und Meßtechniken ausgeschlossen werden können. Diese Bedingungen sind nur für wenige Sozialstrukturmerkmale erfüllt.

Die Validierung erstreckt sich auf folgende Merkmalsverteilungen in der Stichprobe:

- (a) Merkmale, die in die Gewichtung einbezogen wurden:  
Geschlecht, Altersgruppe, Bundesland und Gemeindetyp
- (b) Merkmale, die nicht in die Gewichtung einbezogen wurden:

Erwerbstätige, Familienstand, Stellung im Beruf und Haushaltsgröße.

Nach der Gewichtung dürfen bei den in die Gewichtung einbezogenen Merkmalen keine erheblichen Abweichungen zwischen Stichprobe und Referenzstatistik mehr bestehen. Ziel der vorgenommenen Validierung ist es, zu prüfen, welche Abweichungen zwischen Stichprobe und Referenzstatistik vor der Gewichtung bestehen und zu belegen, daß die Stichprobe nach der Gewichtung tatsächlich der Struktur der Referenzstatistik angeglichen ist. Ferner wird verdeutlicht, welchen Einfluß die Gewichtung auf wesentliche Strukturmerkmale hat - und zwar auch auf solche, die in die Gewichtung nicht einbezogen wurden.

---

[10] Ein alternatives bzw. ergänzendes Verfahren wäre die Validierung über eine Non-Response-Befragung wie beim ALLBUS 1986 (Erbslöh, B./Koch, A., 1988: Die Non-Response-Studie zum ALLBUS 1986, Problemstellung, Design, erste Ergebnisse, in: ZUMA-Nachrichten 22, S. 29-44)

Für Überlegungen zum Zusammenhang von Stichprobenausschöpfung und Stichprobenqualität ("Repräsentativität") siehe Infratest Sozialforschung, M. Kellner-Raum, H. Kuwan, B. v. Rosenblatt: Methodenbericht zum ALLBUS/ISSP 1992, München 1992, S. 19.

Bei den unter (b) aufgezählten Merkmalen ist zu berücksichtigen, daß hier teilweise bereits Effekte unterschiedlicher Definitionen und Meßverfahren auftreten können.

Vorab geben wir eine Übersicht über die Referenzstatistiken, die hier für den Vergleich mit den Stichprobenstrukturen herangezogen werden.

### **Referenzstatistiken**

#### **Teilstichprobe WEST:**

Für die Merkmale Geschlecht, Altersgruppen, und Bundesland dient die Bevölkerungsfortschreibung zum Stand 31.12.1990 als Referenzstatistik. Für die Merkmale Erwerbsstatus, Familienstand, Stellung im Beruf und Haushaltsgröße werden die Vergleichszahlen dem Mikrozensus 1991 entnommen.

Für das Merkmal Gemeindegröße nach BIK wurden die Wahlberechtigten nach dem ADM-Sample als Referenzwert herangezogen.

#### **Teilstichprobe OST:**

Für die Merkmale Geschlecht, Altersgruppe und Bundesland bildet ebenfalls die Bevölkerungsfortschreibung zum 31.12.1990 die Referenzstatistik. Bei den Merkmalen Familienstand und Haushaltsgröße basiert sie auf dem Mikrozensus von 1991.

Als Referenzstatistik für die Merkmale Erwerbsstatus und Stellung im Beruf dient der Arbeitsmarkt Monitor vom November 1992, den Infratest Sozialforschung für die Bundesanstalt für Arbeit bei 10.000 Personen in den neuen Bundesländern durchführt.[11]

Für das Merkmal Schulabschluß liegt keine bevölkerungsbezogene Referenzzahl vor.

---

[11] Da der Arbeitsmarkt Monitor die Grundgesamtheit des Wohlfahrtssurvey (deutsche Bevölkerung ab 18 Jahren) nicht direkt abbildet, sondern die Bevölkerung von 17 bis 65 Jahren repräsentiert, wurde für die Referenzstatistik aus den Daten des Arbeitsmarkt-Monitors eine Hilfsgröße berechnet:  
Referenz = Erwerbstätige 18 bis 65 : Bevölkerung ab 18.  
Ergebnisse des Arbeitsmarkt Monitors für November 1992 werden in den BeitrAB, Bd. 148.6 veröffentlicht. Herausgeber ist die Bundesanstalt für Arbeit.

### 4.3 Gewichtungseffekte und Validierungsergebnisse

In diesem Kapitel werden die Strukturmerkmale der Stichprobe, die im Gewichtungsverfahren an die Referenzstatistik angepaßt wurden, kommentiert. In einem zweiten Abschnitt ist die Validierung der übrigen, nicht an externe Referenzwerte angepaßten Merkmale dargestellt. Am Ende dieses Kapitels sind die Validierungstabellen wiedergegeben, in denen die Verteilungen der in die Validierung einbezogenen Merkmale der jeweiligen Referenzstatistik gegenübergestellt sind.

Die Tabellen folgen den in Kapitel 4.1 beschriebenen Gewichtungsschritten: Zuerst wird für jedes einbezogene Merkmal die Stichprobenstruktur (in Prozent) in jeweils fünf bzw. sechs Tabellenspalten dargestellt. Pro Merkmal sind folgende Verteilungen dokumentiert:

- (1) ungewichtet ("ungew.")
- (2) nach Regionalgewichtung ("regio.")
- (3) nach Designgewichtung bzw. Umwandlung der Haushalts- in eine Personenstichprobe ("umwandl")
- (4) nach Antreffbarkeitsgewichtung ("antreff")
- (5) nach Demographiegewichtung ("demogr.")

Und bei haushaltsbezogenen Merkmalen zusätzlich

- (6) nach Rückumwandlung in eine Haushaltsstichprobe ("HH-gew").

#### Geschlecht

Es ist ein verbreitetes, aber falsches Verfahren, die ungewichteten Stichprobenverteilungen direkt mit der Referenzstatistik zu vergleichen. Dabei wird übersehen, daß die ungewichtete Stichprobe eine Haushalts- und keine Personenstichprobe darstellt, während die Referenzstatistik in der Regel personenbezogen ist. Der korrekte Vergleichswert ist in diesen Fällen erst nach Umwandlung in eine Personenstichprobe gegeben (Spalte "umwandl."). Je mehr ein Merkmal mit der Haushaltsgröße korreliert, umso gravierender wird der Unterschied zwischen ungewichteter und umgewandelter Verteilung.

In der Teilstichprobe Ost weicht der Frauen- bzw. Männeranteil der von der Haushalts- auf die Personenebene umgewandelten Stichprobe von den Referenzwerten ab. Männer sind hier überrepräsentiert. Im Westen stimmt die Stichprobe dagegen sehr gut mit der Referenzstatistik überein. Durch Gewichtung auf die Zahlen der amtlichen Statistik wurden auch im Osten die Strukturen der Referenzstatistik angepaßt.



### Altersgruppen

In den umgewandelten Teilstichproben West und Ost sind die Altersgruppen bis unter 30 Jahre zunächst deutlich unterrepräsentiert - im Osten deutlicher als im Westen (12,7% gegenüber einem Soll-Wert von 22,7% im Osten; 18,2% gegenüber 23,2% im Westen).

Nach dem Ausgleich der schlechteren Antreffbarkeit jüngerer Altersgruppen steigt deren Anteilswert in den Stichproben. Die abschließende Demographiegewichtung gleicht die Struktur der amtlichen Referenzstatistik an.

Für Personen ab 70 Jahren, die in Umfragen normalerweise unterdurchschnittlich erreicht werden, ergibt sich folgendes Bild: Ihr Anteilswert in der ungewichteten (West-)Stichprobe liegt mit 12,4% sehr nahe bei der amtlichen Referenzstatistik (13,5%). Der Anteil der über 70jährigen in der Ost-Stichprobe kommt mit 10,7% der amtlichen Referenz (11,9%) ebenfalls sehr nahe. Dies rührt jedoch zu einem erheblichen Teil von der höheren Auswahlwahrscheinlichkeit von Personen in Ein-Personen-Haushalten her (designbedingte überproportionale Auswahlchance). Dementsprechend sinkt der Anteil der über 70jährigen nach Korrektur dieses Effekts durch die Umwandlung auf 10,2% West bzw. 8,0% Ost. Die Antreffbarkeitsgewichtung "verschlechtert" die Anteile dieser Altersgruppe noch weiter auf West 8,1% und Ost 6,0%. Dies geschieht durch Korrektur der erhöhten Interviewchance der älteren aufgrund ihrer leichteren Antreffbarkeit.

Erst die abschließende Demographiegewichtung bringt den Anteil der älteren Menschen in den beiden Teilstichproben auf den Soll-Wert laut amtlicher Statistik.

### Regionale Verteilung

Die Verteilungen der umgewandelten Stichproben nach Bundesländern weichen nur geringfügig von der Referenzstatistik ab (höchste Abweichung 0,7 Prozentpunkte in Niedersachsen bzw. 0,9 Prozentpunkte in Mecklenburg-Vorpommern). Durch die Gewichtung werden beide Teilstichproben mit der amtlichen Statistik in Übereinstimmung gebracht.

Auch die Verteilung der umgewandelten Stichprobe nach Gemeindetypen weicht kaum von den amtlichen Referenzwerten ab.

Im folgenden werden Merkmale betrachtet, die im Gewichtungsmodell selbst nicht an eine externe Referenzverteilung angepaßt werden. Die Referenzstatistik dient bei diesen Merkmalen lediglich dem nachträglichen Vergleich.

### **Erwerbsstatus**

Für das Merkmal Erwerbstätigkeit wird die Stichprobe nach der Umwandlung in eine Personenstichprobe mit den jeweiligen Referenzwerten verglichen. Der Anteil der Erwerbstätigen liegt sowohl im Westen als auch im Osten unter den Referenzwerten.

Nach Anpassung der geringeren Antreffbarkeit der Erwerbstätigen gegenüber den Nichterwerbstätigen in dem entsprechenden Gewichtungsschritt wird dieses Bild korrigiert. In der Ost-Stichprobe liegt der Anteil der Erwerbstätigen nach der Antreffbarkeitsgewichtung deutlich über dem Referenzwert (58,2% vs. 53,1% Ost, 56,2% vs. 55,5% West). Dies ist auf die leicht verschobene Altersstruktur (die mittleren Altersgruppen mit ihrer hohen Erwerbsquote sind überrepräsentiert) in der Stichprobe zurückzuführen. Die der Antreffbarkeitsgewichtung nachgeschaltete Demographiegewichtung korrigiert diesen Effekt teilweise wieder. In der gewichteten Stichprobenstruktur liegt der Erwerbstätigenanteil im Westen leicht unter, im Osten leicht über dem Wert der Referenzstatistik.

### **Familienstand**

In der Verteilung der verheirateten, verwitweten, geschiedenen und ledigen Befragten in der West- und Ost-Stichprobe spiegelt sich die Verschiebung in der Altersverteilung der Stichprobe. Die mittleren Altersgruppen sind überrepräsentiert, folglich auch der Anteil der Verheirateten in der Stichprobe.

Die Demographiegewichtung bringt die Stichprobenstruktur der Ost-Stichprobe in gute Übereinstimmung mit der Referenzstatistik. In der West-Stichprobe bleibt auch nach Demographiegewichtung der Anteil der Ledigen etwas niedriger und der Anteil der Verheirateten höher als in der Referenzstatistik.

### Haushaltsgröße

Als Referenzwerte werden für Westdeutschland im Mikrozensus 1991[12] 35,1% Einpersonenhaushalte ausgewiesen. Für Ostdeutschland sind es im Mikrozensus 1991 27,6% Einpersonenhaushalte.

Zu beachten ist, daß diese Anteile sich auf die Haushaltsverteilung, nicht auf die Personenverteilung beziehen. Auch die Vergleichszahlen der Stichprobe müssen also auf der Haushaltsebene betrachtet werden.

In der ungewichteten Stichprobe sind die Einpersonenhaushalte zunächst deutlich unterrepräsentiert. Die Gewichtungsschritte auf der Personenebene, insbesondere die Demographiegewichtung, bewirken auf der Haushaltsebene jedoch einen erheblichen Anstieg des Anteils der Einpersonenhaushalte. Obwohl also die Haushaltsgröße durch die Gewichtung selbst nicht an externe Referenzwerte angepaßt wird, liegt der Anteil der Einpersonenhaushalte nach Gewichtung nur noch geringfügig unter den Referenzwerten aus dem Mikrozensus: nämlich bei 32,0% im Westen und 24,9% im Osten.

Ebenfalls zu beachten ist, daß der personenbezogene Anteil der Alleinlebenden sehr viel niedriger ist. Er beträgt in der Stichprobe (nach Gewichtung) 17,0% im Westen bzw. 13,1% im Osten.

---

[12] Der Mikrozensus bezieht sich auf die "wohnberechtigte" Bevölkerung. D.h., daß z.B. Wochenendpendler an ihrem Erst- und ihrem Zweitwohnsitz erfaßt werden. Der Anteil der ausgewiesenen Einpersonenhaushalte dürfte dadurch gegenüber dem üblichen Begriff der Wohnbevölkerung, der auf den Erstwohnsitz abstellt, etwas überhöht sein.

### Stellung im Beruf

Bezogen auf Erwerbstätige ergibt sich in der West-Stichprobe folgendes Bild:

**Übersicht 4** Stellung im Beruf (nach Demographiegewichtung)

WEST	Wohlfahrtssurvey 1993	Mikrozensus 1991[13]
Selbständige	6,9	9,5
Beamte	9,5	9,1
Angestellte	49,4	45,9
Arbeiter	27,3	33,6
Mithelfende	2,6	1,9
in Ausbildung	4,3	--
Erwerbstätige	100,0	100,0

Gemessen am Mikrozensus 1991 sind die Beamten in der West-Stichprobe des Wohlfahrtssurvey gut abgebildet, während Angestellte überrepräsentiert, Arbeiter und Selbständige unterrepräsentiert sind.

Auch für Ostdeutschland liegen Mikrozensusergebnisse für 1991 vor. Wir haben jedoch den Vergleich der Verteilung mit dem Arbeitsmarktmonitor Welle 6 vom November 1992 vorgenommen, da diese Zahlen aktueller sind:

[13] Auszubildende in Angestelltenberufen werden im Mikrozensus als Angestellte ausgewiesen, Auszubildende in Arbeiterberufen als Arbeiter. Im Wohlfahrtssurvey werden Befragte, die in Ausbildung sind, über die Statusabfrage "Stellung im Beruf" erfaßt, so daß eine nachträgliche Zuordnung zur Angestellten- bzw. Arbeiterkategorie nicht möglich ist.

### Übersicht 5 Stellung im Beruf (nach Demographiegewichtung)

OST	Wohlfahrtssurvey 1993	AMM Nov. 1992
Selbständige	5,5	6,9
Beamte	1,0	1,5
Angestellte	52,2	49,9
Arbeiter	35,9	36,8
Mithelfende in Ausbildung	1,7 3,6	0,5 4,5
Erwerbstätige	99,9	100,0

Beide Stichproben stimmen in der Merkmalsverteilung "Stellung im Beruf" gut überein.

### Zusammenfassende Bewertung

Die Stichprobenstruktur des Wohlfahrtssurvey stimmt sowohl bei den Merkmalen, die in die Gewichtung einbezogen wurden, als auch bei nicht in die Gewichtung einbezogenen Merkmalen mit der Referenzstatistik zufriedenstellend überein.

Abweichungen in der Altersstruktur resultieren aus Problemen der Erreichbarkeit bestimmter Bevölkerungsgruppen und den veränderten Haushaltsstrukturen. Durch die vorgenommene Gewichtung konnten diese Abweichungen teilweise oder ganz ausgeglichen werden.