Herausgeber:
Zentralarchiv für Empirische Sozialforschung
Universität zu Köln

Das Zentralarchiv ist Mitglied der GESIS

Direktor: Prof. Dr. W. Jagodzinski
Geschäftsführer: E. Mochmann

Postanschrift:
Postfach 410 960
50869 Köln

Hausanschrift:
Bachemer Straße 40
50931 Köln

Telefon:

Zentrale 0221 / 4 76 94	- 0
Telefax	- 44
Redaktion	- 50

Redaktion:
Franz Bauske

E-mail: *name*@za.uni-koeln.de
Internet: http://www.social-science-gesis.de
http://www.za.uni-koeln.de

ISSN: 0723-5607

© Zentralarchiv

Die ZA-INFORMATION erscheint jeweils im Mai und November eines Jahres.

Sie wird kostenlos an Interessenten und Benutzer des Zentralarchivs abgegeben.

Inhaltsverzeichnis

Forschungsnotizen

Die Wirkung materieller Incentives auf den Rücklauf einer schriftlichen Panelbefragung
von ***Kai Arzheimer*** und ***Markus Klein*** 6

Telefonische Befragungen bei alten Menschen: Ausschöpfung, Kontakthäufigkeit und Verweigerungen
von ***Olaf von dem Knesebeck*** und ***Günther Lüschen*** 32

Quoten- und Randomstichproben in der Praxis der Sozialforschung. Gemeinsamkeiten und Unterschiede in der sozialen Zusammensetzung und den Antwortmustern der Befragten.
von ***Karl-Heinz Reuband*** 48

Zum Problem der Abbildung eindimensional konzipierter Konstrukte bei entgegengesetzter Itempolung
von ***Herbert Matschinger*** und ***Dagmar Krebs*** 81

Statische und dynamische Evaluation von Prognosen
von ***René Weber*** 111

Wahlarithmetische Kabinettstückchen:
Mandate im Überfluß oder wie Überhangmandate die Wahl entscheiden
von ***Christian Rindsfüßer*** und ***Susanne Schäfer-Walkmann*** 124

Ethnocentrism in Present Germany:
Some Correlations with Social Integration and Subjective Assessments
by ***Michael Terwey*** 135

The Northern Ireland Agreement: An Explication of a Political Milestone
by ***Bernadette C. Hayes*** and ***Ian McAllister*** 167

Berichte aus dem Archiv

Announcements from the ZA-EUROLAB
by ***Ingvill C. Oedegaard*** 181

'Environment' - 'Medio Ambiente' A special CD-ROM edition of the 1993 ISSP
by ***Rolf Uher*** 185

Empirical Investigation of Social Space
by ***Jörg Blasius*** 189

International conference on Large Scale Data Analysis
May 25-28, 1999 at the Zentralarchiv: Second Call for Papers 191

Analysis of Time Dependent Data: 28th Spring Seminar at the Zentralarchiv
February 22 - March 12, 1999 193

Berichte, Ankündigungen und Mitteilungen

Weiterbildung in statistischer Datenanalyse beim ICPSR Summer Program 1999..... 197

Linux, ein Unix in neuem Gewand
von ***Bruno Hopp*** 198

Buchhinweise

Günter Wiswede: Soziologie. Grundlagen und Perspektiven für den wirtschafts- und sozialwissenschaftlichen Bereich, 3. neubearbeitete Auflage,
Ein Buchhinweis von ***Günter Endruweit*** 201

Wilhelm Bürklin, Markus Klein: Wahlen und Wählerverhalten. Eine Einführung
Ein Buchhinweis von ***Dieter Ohr*** 203

Ekkehard Mochmann, Ingvill C. Oedegaard, Reiner Mauer:
Inventory of National Election Studies in Europe 1945-1995 205

Jürgen Bortz, Gustav Adolf Lienert:
Kurzgefaßte Statistik für die klinische Forschung – Ein praktischer Leitfaden für die Analyse kleiner Stichproben
Ein Buchhinweis von ***Erwin Lautsch*** 207

Bei Beiträgen, die nicht von Mitarbeitern des Zentralarchivs verfaßt wurden, ist die Anschrift der Autoren beim jeweiligen Artikel angegeben. Die Inhalte der Beiträge entsprechen der Meinung der Autoren und geben nicht unbedingt die Ansicht der Redaktion wieder. Alle inhaltlichen Beiträge sind Gegenstand einer Beurteilung durch externe Gutachter.

Mitteilungen der Redaktion

Ausschöpfungsquoten sind ein aktuelles Thema der Sozialforschung. Sinkende Raten werden allenthalben beklagt. ***Kai Arzheimer*** und ***Markus Klein*** haben ein Methodenexperiment gemacht, das die Wirkung von materiellen Anreizen auf die Ausschöpfungsquote bei postalischen Befragungen zeigt. Auch ***Olaf von dem Knesebeck*** und ***Günther Lüschen*** sind auf der Suche nach besseren Ausschöpfungsquoten. Sie prüfen Maßnahmen zur Ausschöpfungsoptimierung bei der telefonischen Befragung älterer Menschen.

Unterscheiden sich Quoten- und Randomstichproben bezüglich der Sozialstruktur der Befragten und der Antwortmuster? Dieser Frage geht ***Karl-Heinz Reuband*** nach und beantwortet u.a. auch, ob in Quotenstichproben die leicht erreichbaren überrepräsentiert sind.

Im Jahr der Bundestagswahl wollen wir uns auch das Thema Wahl vornehmen. Diesmal jedoch nicht mit der Präsentation einer Umfrage. ***Christian Rindsfüßer*** und ***Susanne Schäfer-Walkmann*** beschäftigen sich mit dem Einfluß der Überhangmandate auf die Sitzverteilung im Bundestag. Sie stellen fest, daß z.B. Wahlbeteiligung, Abwanderungstendenzen, Ausländeranteil, im Wahlkreis lebende Kinder und Jugendliche Einfluß auf die Zahl der entstehenden Überhangmandate haben.

Herbert Matschinger und ***Dagmar Krebs*** nehmen Bezug auf die Veröffentlichung von ***Parviz Rostampour*** in der vorigen Ausgabe der ZA-Information. Sie vergleichen die Ergebnisse einer Hauptkomponentenanalyse mit einer Unfoldinganalyse, die nach Meinung der Autoren eine adäquatere Vorgehensweise zur Beurteilung eindimensionaler Konstrukte bei entgegengesetzter Itempolung darstellt.

Bernadette C. Hayes war mehrere Monate Gast im Zentralarchiv im Rahmen des durch die EU geförderten TMR/LSF-Programms. Ihren Vortrag, den sie zum Nordirlandabkommen hielt, hat sie mit ihrem Kollegen ***Ian McAllister*** für uns zu Papier gebracht.

René Weber hat sich auf die Suche nach Möglichkeiten zur Quantifizierung der Güte von Prognosen gemacht. Er empfiehlt einen Mix von Gütemaßen, zu dessen Berechnung er auch ein Programm entwickelt hat.

Die Debatte um die Akzeptanz von Ausländern in Deutschland hat immer noch einen hohen Stellenwert. ***Michael Terwey*** präsentiert hierzu eine Analyse mit ALLBUS-Daten.

Franz Bauske

Die Wirkung materieller Incentives auf den Rücklauf einer schriftlichen Panelbefragung[1]

von Kai Arzheimer[2] und Markus Klein

Zusammenfassung:

Im vorliegenden Aufsatz untersuchen wir den Einfluß materieller Incentives auf die Ausschöpfungs- und die Haltequote einer zweiwelligen schriftlichen Panelbefragung. Dabei zeigt sich, daß Incentives die Ausschöpfungsquote signifikant erhöhen, wobei dieser Effekt in der ersten Panelwelle stärker ausfällt als in der zweiten. Darüber hinaus können wir zeigen, daß kein Hysteresis-Effekt existiert und die beiden Incentives unabhängig voneinander wirken. Effekte der Incentives auf die Stichprobenzusammensetzung lassen sich nicht nachweisen.

Abstract:

In the present paper we investigate the influence that material incentives have on the response rate and the panel mortality of a two-wave mail survey panel. We can show that incentives increase the response rate significantly, this effect being stronger in the first panel wave than in the second. Furthermore, we can demonstrate that there is no hysteresis effect and that the effect of one incentive is independent of the other one. There is no evidence of incentive effects on the sample composition.

1 Einleitung und Problemstellung

Daß materielle Incentives, also kleine Geschenke an die zu interviewenden Personen, die Ausschöpfungsquote schriftlicher Befragungen erhöhen, hat sich in empirischen Untersuchungen immer wieder eindrucksvoll bestätigt (vgl. zuletzt ***Hare et al.*** 1998). Bekannt sind darüber hinaus auch diejenigen Eigenschaften von Incentives, die einem hohen Rücklauf

1 Für wertvolle Hinweise und Anregungen danken wir ***Manu Pötschke***, ***Friederika Priemer*** und ***Annette Schmitt***.

2 ***Kai Arzheimer*** ist wissenschaftlicher Mitarbeiter am Institut für Politikwissenschaft der Johannes Gutenberg-Universität Mainz, 55099 Mainz.

besonders förderlich sind: Wenn das Incentive vor der eigentlichen Befragung gewährt wird, es sich um einen Geldbetrag handelt und sein Wert hoch ist, dann erhöht sich die Rücklaufquote in der Regel besonders stark (***Church*** 1993; ***Fox et al***. 1988; ***Heberlein/Baumgartner*** 1978; ***James/Bolstein*** 1990, 1992; ***Kankuk***/***Berelson*** 1975; ***Linsky*** 1975; ***Nederhof*** 1983; ***Warriner et al.*** 1996; ***Yu***/***Cooper*** 1983). Es gibt darüber hinaus auch empirische Evidenz dafür, daß Incentives ihre rücklaufsteigernde Wirkung nicht nur bei schriftlichen, sondern auch bei mündlichen und telefonischen Interviews entfalten (***Singer*** 1998, ***Willimack et al.*** 1995). Vor dem Hintergrund der in den letzten Jahren zu beobachtenden steigenden Zahl von Interviewverweigerungen bei repräsentativen Bevölkerungsumfragen besitzen diese Befunde eine enorme praktische Bedeutung für die akademische und kommerzielle Umfrageforschung.

Trotz der vergleichsweise umfassenden Forschungsanstrengungen existiert allerdings keine empirische Untersuchung über die Wirkung materieller Incentives im Rahmen von Panelstudien. Diese Lücke versucht der vorliegende Beitrag zu schließen. Am Beispiel einer zweiwelligen schriftlichen Panelbefragung wird im Rahmen eines experimentellen Designs untersucht, wie sich kleine Geschenke an die Befragungspersonen (in unserem Fall eine Telefonkarte im Wert von 6 DM) auf die Rücklaufquote und die Panelmortalität auswirken. Da schriftliche Panels in der Praxis der empirischen Sozialforschung eher eine Ausnahme darstellen, wird mit diesem Aufsatz gleichzeitig auch der Frage nachgegangen, inwieweit schriftliche Panelbefragungen eine Alternative zu den bislang dominierenden Face-to-face-Panels darstellen können (vgl. hierzu auch ***Bauer*** 1988, ***Reuband*** 1998).

2 Theoretischer Rahmen: Austauschtheorie und Reziprozitätsnorm

Daß Incentives ein geeignetes Instrument zur Erhöhung der Ausschöpfungsquote von Umfragen darstellen, scheint bereits aus einem Alltagsverständnis heraus plausibel. Im Rahmen der von ***Don A. Dillman*** (1978) entwickelten *Total Design Method* (TDM) wird der Einsatz von Incentives aus einer umfassenden theoretischen Perspektive heraus empfohlen. Bei der TDM handelt es sich um ein mittlerweile auch in Deutschland hervorragend bewährtes Verfahren zur Optimierung des Rücklaufs und der Datenqualität schriftlich-postalischer Erhebungen (vgl. ***Hippler***/***Seidel*** 1985; ***Niedermayer*** 1989; ***Blasius***/***Reuband*** 1996). Das Interview wird hierbei im Sinne der „Social Exchange Theory“ (***Blau*** 1964; ***Homans*** 1961; ***Thibaut/Kelley*** 1959) als sozialer Austauschprozeß verstanden, der nur dann zustande kommt, wenn für die Befragungsperson der Nutzen der Interviewteilnahme deren Kosten übersteigt. Der Forscher muß folglich bemüht sein, die *Kosten* der Interviewteilnahme für den Befragten möglichst gering zu halten, während diesem der *Nutzen*, den er aus der Interviewteilnahme ziehen kann, möglichst groß erscheinen soll.

Die *Kosten* der Befragungsperson lassen sich durch eine ganze Reihe von Maßnahmen minimieren: Ein klarer und übersichtlicher Fragebogen senkt den Aufwand, der für seine Beantwortung nötig ist, die Verwendung von freigestempelten Rückumschlägen vermeidet Portokosten auf seiten der Befragten und eine überzeugende Versicherung, daß die erhobenen Daten vertraulich behandelt werden, wirkt der Furcht der Befragten vor zukünftigen Nachteilen aus der Interviewteilnahme entgegen (vgl. ***Dillman*** 1978: 14-16).

Die Möglichkeiten auf seiten des Forschers, den *Nutzen* der Interviewteilnahme zu erhöhen, sind allerdings eher beschränkt. Der Nutzen für den Befragten liegt im wesentlichen nur in der Befriedigung, in einer wichtigen Sache „konsultiert" worden zu sein. Bei geschickter Aufmachung des Anschreibens scheint dies für viele Befragte jedoch bereits ein hinreichendes Motiv zur Teilnahme zu sein: „Being consulted has been pointed to by both ***Blau*** and ***Homans*** as a means of providing a reward to people while getting needed information" (***Dillman*** 1978: 13).

Ein etwaiges *Incentive* hingegen kann nicht als Nutzen aus der Interviewteilnahme verstanden werden, da die Befragungsperson dieses auch dann behalten kann, wenn sie die Interviewteilnahme verweigert. Das Incentive eröffnet vielmehr eine soziale Austauschbeziehung mit dem Respondenten: Zunächst ohne Gegenleistung überreicht, fungiert es als ein „symbol of trust" (***Dillman*** 1978: 16) und generiert eine soziale Verpflichtung des Befragten, dem Wunsch nach einem Interview zu entsprechen. Die Interviewteilnahme stiftet so durch die Erfüllung der sog. *Reziprozitätsnorm* dem Befragten einen zusätzlichen Nutzen. Die Reziprozitätsnorm besagt, daß man bemüht sein sollte, freiwillig gewährte Vergünstigungen in irgendeiner Form „zurückzuzahlen" (***Gouldner*** 1960). Der Wert des Incentives sollte vor diesem Hintergrund deshalb auch gegenüber den Befragten nicht als „Vergütung" für die Interviewteilnahme dargestellt werden. „The closer the monetary incentive comes to the value of the service performed, the more the transaction tends to move into the realm of economic exchange and the easier it becomes for many people to refuse it" (***Dillmann*** 1978: 16).

3 Die zugrundeliegende Studie

Die Datenerhebung, über die wir im folgenden berichten werden, wurde im Rahmen des Forschungsprojektes "Die Conjoint Analyse als Instrument der empirischen Wahlforschung. Eine exemplarische Analyse der Hamburger Bürgerschaftswahl 1997 mittels einer schriftlichen Befragung gemäß der ***Dillman***schen Total-Design-Method"[3] durchgeführt.

3 Dieses Forschungsprojekt ist ein Kooperationsprojekt zwischen dem Institut für Politikwissenschaft der Johannes Gutenberg-Universität Mainz und dem Zentralarchiv für Empirische Sozialforschung der Universität zu Köln. Es wird von der Deutschen Forschungsgemeinschaft finanziell gefördert (Geschäftszeichen FA 113/11-1) und steht unter der Leitung von Prof. Dr. ***Jürgen W. Falter*** (Mainz).

Die Datenerhebung startete unmittelbar nach der Bürgerschaftswahl vom 21. September 1997. Um die Kosten der Erhebung niedrig zu halten und den Prozeß der Datenerhebung kontrollieren zu können, haben wir uns für eine schriftliche Befragung nach TDM entschieden, die wir selbst administrierten. Die Notwendigkeit einer Panelstudie ergab sich aus der Tatsache, daß das für unser substantielles Forschungsvorhaben wichtige Konzept der Parteiidentifikation nur dann valide gemessen werden kann, wenn mindestens zwei Meßzeitpunkte vorliegen. Vor diesem Hintergrund haben wir uns für die Durchführung eines zweiwelligen Kurzfristpanels entschieden.

Ein generelles Problem schriftlicher Befragungen besteht darin, daß meist nur Adreß-, aber keine Personenstichproben zur Verfügung stehen. Dies ist aus zwei Gründen problematisch: Zum einen ist mit systematischen Verzerrungen zu rechnen, wenn die Auswahl der Zielperson im Haushalt den Befragten selbst überlassen wird, zum anderen widerspricht die Adressierung des Anschreibens an einen Haushalt den von ***Dillman*** entwickelten Prinzipien der Personalisierung. In unserem Fall waren jedoch die Voraussetzungen für die Ziehung einer Personenstichprobe günstig, da es sich beim Land Hamburg um einen Stadtstaat handelt, für den mit dem Hamburger Amt für zentrale Meldeangelegenheiten eine einzige Meldebehörde zuständig ist. Diese erwies sich als äußerst kooperationsbereit: Sechs Wochen vor dem Wahltermin wurde für uns eine Zufallsstichprobe vom Umfang N=4000 aus dem Melderegister gezogen. Die Adreßinformationen waren somit vollständig und entsprachen dem aktuellen amtlichen Stand. Grund- und Auswahlgesamtheit (die Menge aller Personen, die bei der Bürgerschaftswahl stimmberechtigt waren) stimmten perfekt überein. Zusätzlich zur Adresse wurde uns außerdem das Geburtsjahr, das Geschlecht und die Ortsteilzugehörigkeit[4] der Zielpersonen mitgeteilt. Das Geschlecht wurde benötigt, um die Anschreiben korrekt adressieren zu können. In Verbindung mit dem Geburtsjahr eröffnete es außerdem eine Möglichkeit zu kontrollieren, ob der Fragebogen tatsächlich von der Zielperson ausgefüllt wurde. Die Angabe zur Ortsteilzugehörigkeit schließlich gestattet es, dem Datensatz für spätere Mehrebenenanalysen eine Vielzahl von Kontextinformationen zuzuspielen.

Der Versand erfolgte exakt nach den Vorgaben ***Dillmans*** (vgl. zusammenfassend Tab. 1). In der Woche nach der Wahl erhielten alle Befragten einen zwölfseitigen Fragebogen im Format DIN-A5, der außer den relativ umfangreichen Fragen zur Wahrnehmung und Bewertung politischer Issues, die für die Hauptuntersuchung benötigt wurden, noch einige Standard-Items der politischen Soziologie und die gängigen sozio-demographischen Fragen umfaßte. Dem Fragebogen lag ein aus 10 Karten im Format 5,5 x 9cm bestehendes Kartenspiel bei, auf dem hypothetische Wahlplattformen abgedruckt waren. Die Befragten

4 Hamburg gliedert sich in 180 Ortsteile, für die ***Podszuweit/Schütte*** (1997) eine Vielzahl von Sozialstrukturdaten bereitstellen.

wurden gebeten, diese Karten nach Maßgabe ihrer Wünschbarkeit in eine Rangordnung zu bringen.[5] Da sich diese Aufgabe in den Pretests als für die Befragten sehr schwierig erwies, und wir vor diesem Hintergrund mit einer hohen Rate von Interviewverweigerungen rechneten, lag einem Teil der Fragebögen noch ein kleines Geschenk an die Zielperson bei, mit dem wir die Ausschöpfungsquote zu erhöhen hofften (vgl. Kap. 4).

Das dem Fragebogen beiliegende personalisierte Anschreiben war, ebenso wie alle folgenden, erkennbar von Hand unterzeichnet. Als ein "symbol of trust" wurden die Rückumschläge von Hand vorfrankiert, statt sie mit dem Aufdruck "Gebühr zahlt Empfänger" zu versehen. Um eine Rücklaufkontrolle durchführen zu können, waren alle Rückumschläge durch eine handschriftlich aufgetragene Nummer kenntlich gemacht worden. Die Glaubwürdigkeit der Anonymitätszusicherung haben wir zu erhöhen versucht, indem wir im Anschreiben besonders auf die Tatsache hinwiesen, daß die Datenerfassung und -verarbeitung nicht in Hamburg, sondern in Mainz erfolgt. Dies begründeten wir gegenüber den Befragten mit den Erfordernissen eines effektiven Datenschutzes. Um den Bezug zu Hamburg herzustellen, war auf dem Briefkopf des Anschreibens neben dem Wappen der Hamburger Universität der Hamburger Politikprofessor ***Michael Th. Greven*** als vermeintlicher Projektleiter aufgeführt.[6]

Dem Schema ***Dillmans*** folgend wurde eine Woche später an alle Zielpersonen ein Erinnerungsschreiben verschickt. An jene Personen, die bis zum jeweiligen Zeitpunkt nicht geantwortet hatten, wurden außerdem nach drei und nach sieben Wochen weitere Erinnerungsschreiben versandt, denen ein Ersatzfragebogen sowie ein weiterer frankierter Rückumschlag beilag. Das letzte Erinnerungsschreiben im Rahmen der ersten Panelwelle wurde sieben Wochen nach der Wahl, d.h. Mitte November 1997 verschickt. Rund einen Monat später war der Rücklauf vollständig zum Erliegen gekommen, und die Feldarbeit für die erste Befragungswelle konnte als abgeschlossen angesehen werden.

Um die Befragten nicht zu überlasten, wurde der Beginn der zweiten Feldphase für Mitte Februar angesetzt. Der Fragebogen der zweiten Panelwelle enthielt außer dem für uns besonders wichtigen Parteiidentifikations-Item noch diverse Fragen zur Zufriedenheit mit dem Wahlausgang, zur präferierten Koalition und zur Bewertung des neuen Senats. Insgesamt umfaßte der Fragebogen im Format DIN-A5 lediglich vier Seiten und wurde an alle Personen verschickt, die in der ersten Panelwelle geantwortet hatten. Mit dem Fragebogen wurde außerdem eine mehrseitige, graphisch ansprechend gestaltete Zusammenfassung der Ergebnisse aus der ersten Befragungswelle verschickt.

---

5 Zur Fragestellung und zum Design der Untersuchung vgl. ***Arzheimer/Klein*** 1998.

6 Wir bedanken uns an dieser Stelle herzlich bei ***Michael Th. Greven***, der sich als vermeintlicher Projektleiter zur Verfügung gestellt hat.

Auch in der zweiten Panelwelle wurden die Vorgaben ***Dillmans*** exakt umgesetzt. Die einzige Abweichung bestand darin, daß wir auf das letzte Erinnerungsschreiben (sieben Wochen nach Versand des Fragebogens) verzichteten. Dieses dritte Schreiben hatte in der ersten Welle den Rücklauf nur noch unwesentlich erhöht, aber erhebliche Zusatzkosten verursacht.

Tab. 1: Der Ablauf der Datenerhebung im Rahmen des Projektes "Die Conjoint Analyse als Instrument der empirischen Wahlforschung"

Termin der Hamburger Bürgerschaftswahl	21.09.1997
Ablauf der 1. Welle	
Versand des Anschreibens und des Fragebogens (sowie des Incentives)	25.09.1997
1. Erinnerungsschreiben (ohne Fragebogen)	01.10.1997
2. Erinnerungsschreiben (mit Fragebogen)	17.10.1997
3. Erinnerungsschreiben (mit Fragebogen)	14.11.1997
Ablauf der 2. Welle	
Versand des Anschreibens und des Fragebogens (sowie des Incentives)	26.02.1998
1. Erinnerungsschreiben (ohne Fragebogen)	05.03.1998
2. Erinnerungsschreiben (mit Fragebogen)	19.03.1998

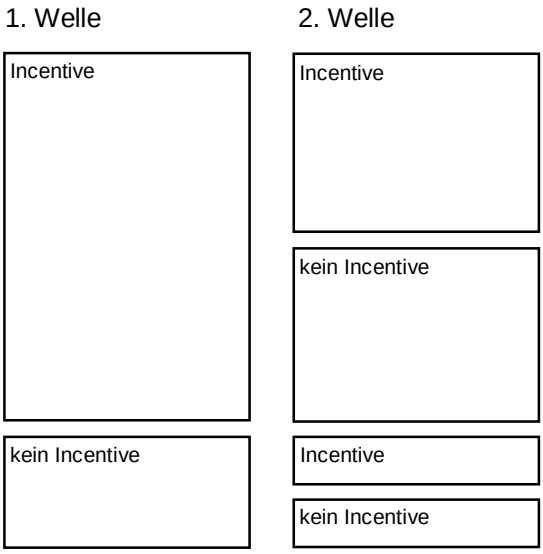
4 Ein experimentelles Design zur Untersuchung des Effekts materieller Incentives auf die Ausschöpfungs- und die Haltequote

Als materielles Incentive haben wir in unserem Forschungsprojekt eine Telefonkarte im Wert von 6 DM eingesetzt, die in einem Geschenkfolder aus Karton mit dem Aufdruck "Ein kleines Dankeschön für Ihre Mithilfe..." verschickt wurde. Das Incentive wurde in jeder Welle jeweils mit dem *ersten* Fragebogen versandt, den Mahnschreiben lag keines mehr bei. Telefonkarten sind aufgrund ihres Gewichtes und Formates hervorragend zum Versand geeignet. Sie haben im Gegensatz zu Kugelschreibern, Feuerzeugen oder ähnlichen "Geschenken" einen für den Befragten klar erkennbaren materiellen Wert und wirken dabei weniger prosaisch als ein beigefügter Geldschein.

Um kontrollieren zu können, ob sich der Einsatz des Incentives lohnt, wurde in die Befragung ein zweistufiges Methodenexperiment integriert: In der ersten Panelwelle erhielten 75 Prozent der Zielpersonen ein Incentive und 25 Prozent nicht. In der zweiten Panelwelle erhielten in jeder der beiden Versuchsgruppen der ersten Panelwelle jeweils 50 Prozent der Zielpersonen ein Incentive und 50 Prozent nicht. Im Ergebnis erhält man so vier Gruppen von Zielpersonen: Eine Gruppe, die in beiden Panelwellen ein Incentive erhielt, zwei wei-

tere Gruppen, die entweder in der ersten oder in der zweiten Welle ein Incentive erhielten, sowie schließlich die Kontrollgruppe, die zu keinem Zeitpunkt ein Incentive erhielt. Das Design des Methodenexperiments ist in Abb. 1 visualisiert.

Abb. 1: Das Design des Methodenexperiments zur Analyse der Wirkung materieller Incentives auf die Ausschöpfungsquote und die Panelmortalität



Mit diesem experimentellen Design lassen sich die folgenden Fragen beantworten:

Wie wirkt sich die Gewährung eines materiellen Incentives in den beiden Panelwellen auf die Geschwindigkeit und die Höhe des Rücklauf aus? (vgl. Kap. 5.1)

Hat ein in der ersten Panelwelle gewährtes Incentive einen Effekt auf die Ausschöpfungsquote der zweiten Panelwelle, d.h. gibt es einen Hysteresis-Effekt? Von einem Hysteresis-Effekt spricht man allgemein immer dann, wenn eine Wirkung anhält, obgleich ihre Ursache verschwunden ist. Bekanntestes Beispiel aus der Physik ist der Magnetismus. Ein Stück Metall behält einen Restmagnetismus, auch wenn das magnetische Feld, welches das Magnetisieren des Metalls bewirkt hat, nicht mehr einwirkt. Übertragen auf unser Beispiel würde das bedeuten, daß bei den Personen, die in der ersten Panelwelle ein Incentive erhielten, die Ausschöpfungsquote der zweiten Panelwelle ebenfalls höher liegt, auch wenn sie in der zweiten Panelwelle kein Incentive mehr erhalten. Die durch ein Incentive begründete Austauschbeziehung wäre in diesem Fall also von dauerhafter Natur (vgl. Kap. 5.2).

Tritt ein Interaktionseffekt zwischen den beiden Incentives auf, d.h., wird die Wirkung des zweiten Incentives verstärkt oder abgeschwächt, wenn die Zielperson schon in der ersten Panelwelle ein Incentive erhalten hat? (vgl. Kap. 5.3)

Verändert der Einsatz von Incentives die Zusammensetzung der letztlich realisierten Stichprobe, d.h. gibt es Bevölkerungsgruppen, in denen Incentives eine stärkere Wirkung besitzen als in anderen Bevölkerungsgruppen, oder wirkt das Incentive in allen Bevölkerungsgruppen in gleicher Weise? (vgl. Kap. 5.4)

5 Der empirische Befund

Tab. 2 zeigt zunächst den Verlauf der Befragung im Überblick. Von den 4000 gelieferten Adressen waren 3874 (96,9 Prozent) verwertbar, 126 Personen waren unbekannt verzogen. Von diesen 3874 Personen beteiligten sich insgesamt 58,5 Prozent an der ersten Panelwelle, 1606 verweigerten die Teilnahme entweder explizit oder reagierten auf keines unserer Anschreiben. Von den 2268 Teilnehmern der ersten Panelwelle standen rund 95 Prozent auch für die zweite Panelwelle zur Verfügung. 57 Personen waren zwischenzeitlich unbekannt verzogen bzw. nicht mehr in Hamburg wohnhaft, 51 Personen hatten die Identifikationsnummer auf dem Rückumschlag der ersten Welle unkenntlich gemacht, so daß für uns nicht zu erkennen war, wer uns geantwortet hatte und in der zweiten Welle wieder angeschrieben werden sollte. Von den verbleibenden 2160 Befragten haben uns 87,1 Prozent ein zweites Mal geantwortet.

Tab. 2: Die Ausschöpfung der Stichprobe im DFG-Projekt “Die Conjoint Analyse als Instrument der empirischen Wahlforschung“

1. Panelwelle		
Bruttostichprobe	4.000	
./. stichprobenneutrale Ausfälle	126	
Bereinigter Stichprobenansatz	3.874	100,0
./. systematische Ausfälle	1.606	
Auswertbare Interviews	2.268	58,5
In der 2. Welle nicht angeschrieben	51	
Bruttostichprobe 2. Welle	2.217	
2. Panelwelle		
Bruttostichprobe	2.217	
./. stichprobenneutrale Ausfälle	57	
Bereinigter Stichprobenansatz	2.160	100,0
./. systematische Ausfälle	278	
Auswertbare Interviews	1.882	87,1
Gesamtauschöpfung über beide Wellen		51,0

5.1 Wie wirken sich materielle Incentives auf die Rücklaufgeschwindigkeit und die Ausschöpfungsquote der ersten und zweiten Panelwelle aus?

Wie variiert nun der Rücklaufprozeß in Abhängigkeit von der Gewährung eines Incentives? In den beiden Abbildungen 2 und 3 ist zunächst die Entwicklung der Rücklaufquoten der beiden Panelwellen in den ersten neunzig Tagen nach Beginn der Feldphase abgetragen, jeweils differenziert nach der Versuchs- und der Kontrollgruppe. In beiden Abbildungen ist zu erkennen, daß der Rücklauf in der Gruppe derjenigen Personen, die ein Incentive erhielten, schneller erfolgt und in eine höhere Ausschöpfungsquote mündet als in der Kontrollgruppe ohne Incentive. Es zeigt sich außerdem, daß die Wirkung des Incentives in der ersten Panelwelle stärker ist als in der zweiten. Letzteres kann sicher damit erklärt werden, daß die Bruttostichprobe der zweiten Panelwelle positiv vorselektiert ist, da sie aus Personen besteht, die bereits in der ersten Panelwelle Antwortbereitschaft zeigten.

Aus der bloßen visuellen Inspektion der beiden Abbildungen 2 und 3 lassen sich aber noch keine stringenten empirischen Aussagen herleiten. Deshalb haben wir im nächsten Schritt die beschriebenen Effekte mit inferenzstatistischen Verfahren geprüft. Zunächst haben wir die Anzahl der Kalendertage zwischen dem ersten Versand der Fragebögen und dem Datum des Poststempels auf dem Umschlag des zurückgesandten Fragebogens ermittelt und verwenden diese als abhängige Variable. In der ersten Panelwelle antworteten die Befragten mit Incentive nach durchschnittlich 15,7 Tagen, die Befragten ohne Incentive nach 18,6 Tagen. Diese Differenz von 2,9 Tagen ist statistisch signifikant. Ein ähnliches Muster zeigt sich in der zweiten Panelwelle. Dort beträgt der entsprechende Zeitraum in der Gruppe mit Incentive 9,0 Tage und in der Gruppe ohne Incentive 10,7 Tage. Auch diese Differenz von 1,7 Tagen ist statistisch signifikant (vgl. Tab. 3.). Der Einsatz eines Incentives erhöht also in beiden Panelwellen die *Geschwindigkeit* des Rücklaufs. Ein schneller Rücklauf ist bei der Durchführung einer Untersuchung nach TDM vor allem deshalb interessant, weil er die Zahl derjenigen Personen, die ein zweites oder drittes Mal angeschrieben werden müssen, reduziert. Damit sinken zugleich die Kosten und der Arbeitsaufwand (vgl. ***Warriner et al.*** 1996: 555f).

Abb. 2: Die Entwicklung der Ausschöpfungsquote der ersten Panelwelle in den ersten 90 Tagen nach Beginn der Feldphase in Abhängigkeit von der Gewährung eines Incentives

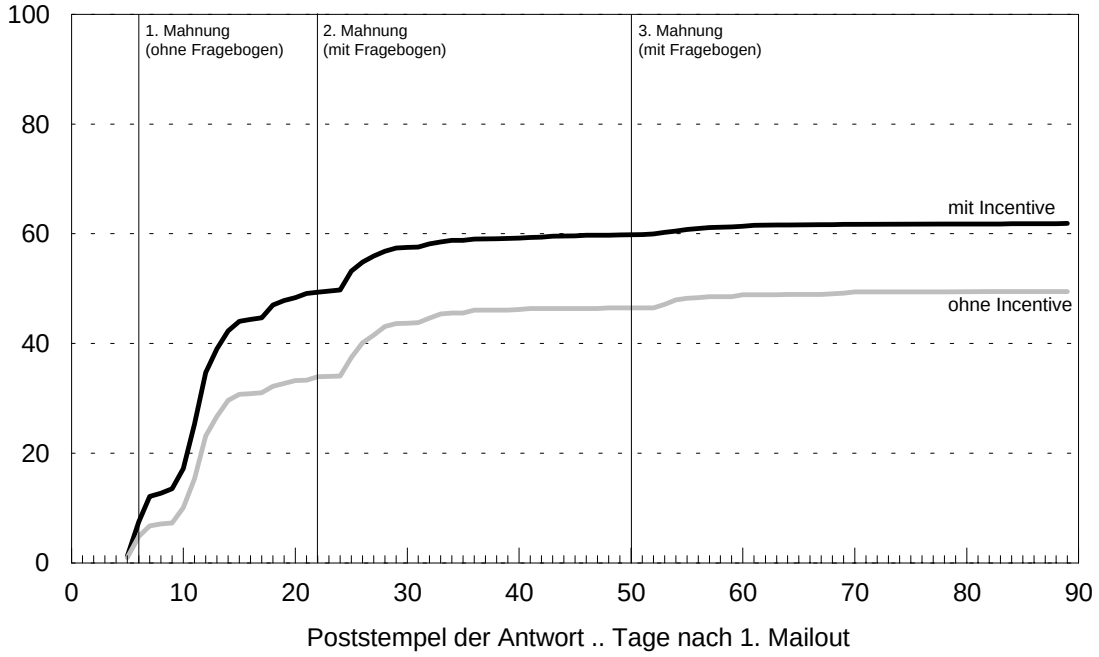
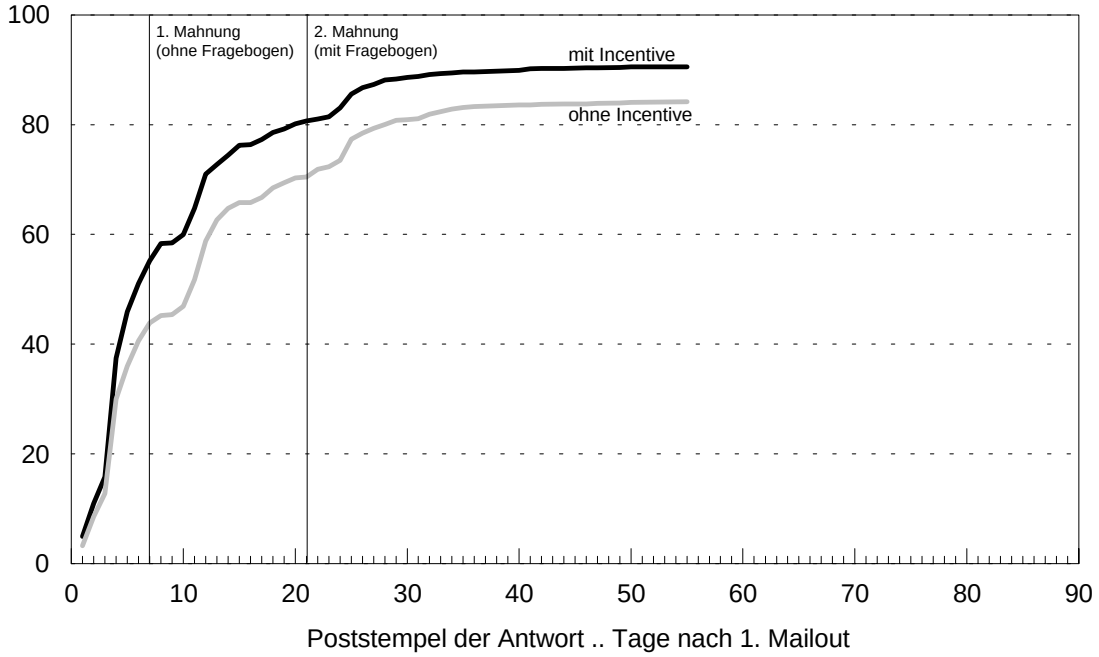


Abb. 3: Die Entwicklung der Ausschöpfungsquote der zweiten Panelwelle in den ersten 90 Tagen nach Beginn der Feldphase in Abhängigkeit von der Gewährung eines Incentives



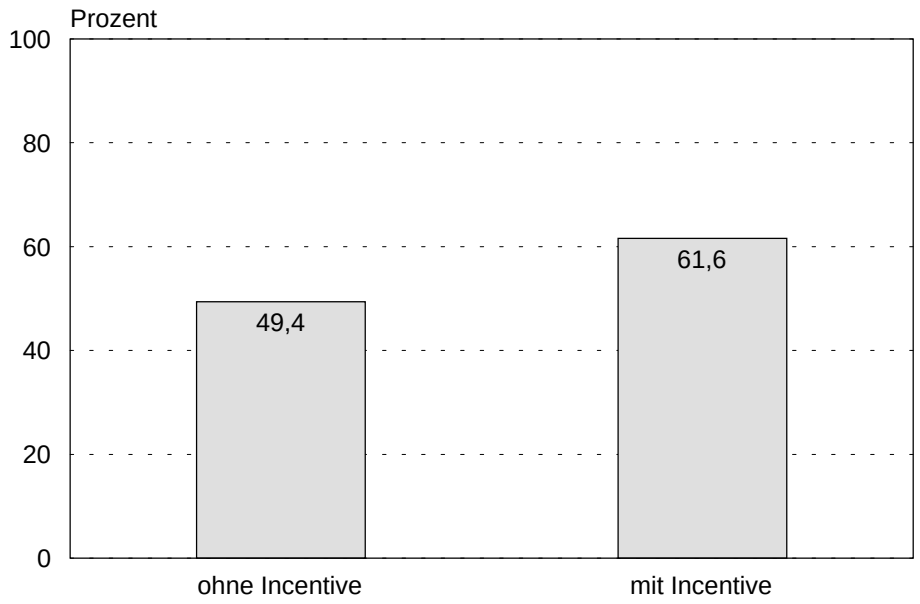
Tab. 3: Der Effekt der Incentives auf die Geschwindigkeit des Rücklaufs in der ersten und zweite Panelwelle - T-Tests für unabhängige Stichproben -

	Mean	SD	df*	T	p
Antwortzeit 1. PW	16,4	11,9	671	4,34	0,000
mit Incentive	15,7	11,3			
ohne Incentive	18,6	13,4			
Antwortzeit 2. PW	9,8	8,6	1812	4,16	0,000
mit Incentive	10,7	8,1			
ohne Incentive	9,0	9,1			

*Freiheitsgrade sind für die unterschiedliche Streuung in den Subgruppen korrigiert

Wichtiger als die Geschwindigkeit des Rücklaufs ist aber dessen *Höhe*. Auch hier lassen sich signifikante Effekte nachweisen: In der ersten Panelwelle betrug die Rücklaufquote in der Versuchsgruppe 61,7 Prozent und in der Kontrollgruppe 49,4 Prozent (vgl. Abb. 4, die genauen Fallzahlen, die diesen Quoten zugrunde liegen sind in Tabelle A1 im Anhang dokumentiert). Die Wahrscheinlichkeit, daß sich eine Zielperson aus der Experimentalgruppe an der Studie beteiligt, lag also um fast ein Viertel höher als die Wahrscheinlichkeit für die Kontrollgruppe. Diese Differenz von 12,3 Prozentpunkten in den Ausschöpfungsquoten ist inhaltlich bedeutsam und statistisch hoch signifikant (Cramers V=0,11, p=0,000).

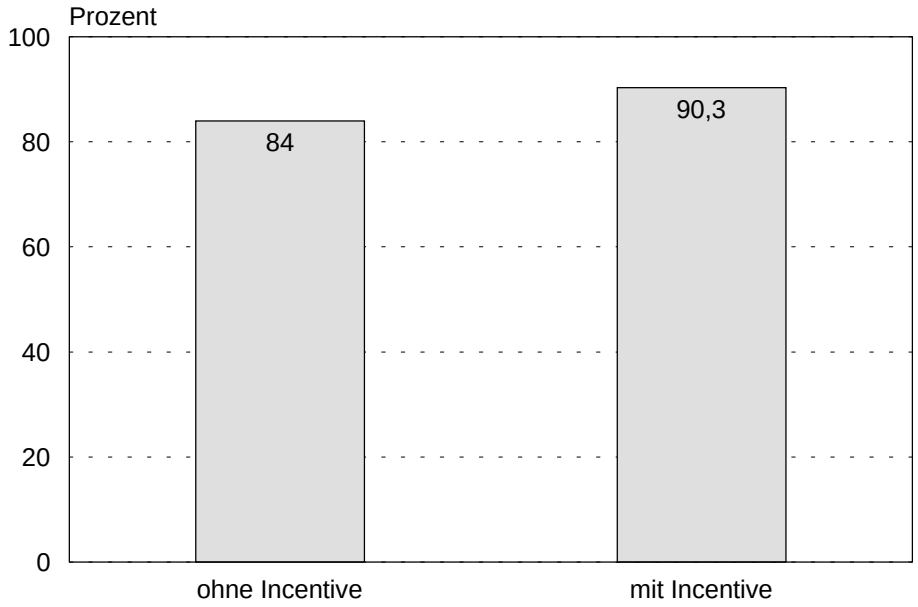
Abb. 4: Der Rücklauf in der ersten Panelwelle in Abhängigkeit von der Gewährung eines Incentives



Auch in der zweiten Panelwelle zeigte sich ein Effekt des Incentives. Hier lagen die entsprechenden Ausschöpfungsquoten bei 84,0 bzw. 90,3 Prozent (vgl. Abb. 5). Das in der

zweiten Panelwelle gewährte Incentive erhöhte die Ausschöpfung also nur noch in geringerem Maße als in der ersten Panelwelle, die Differenz ist gleichwohl statistisch signifikant (Cramers V=0,09, p=0,000).

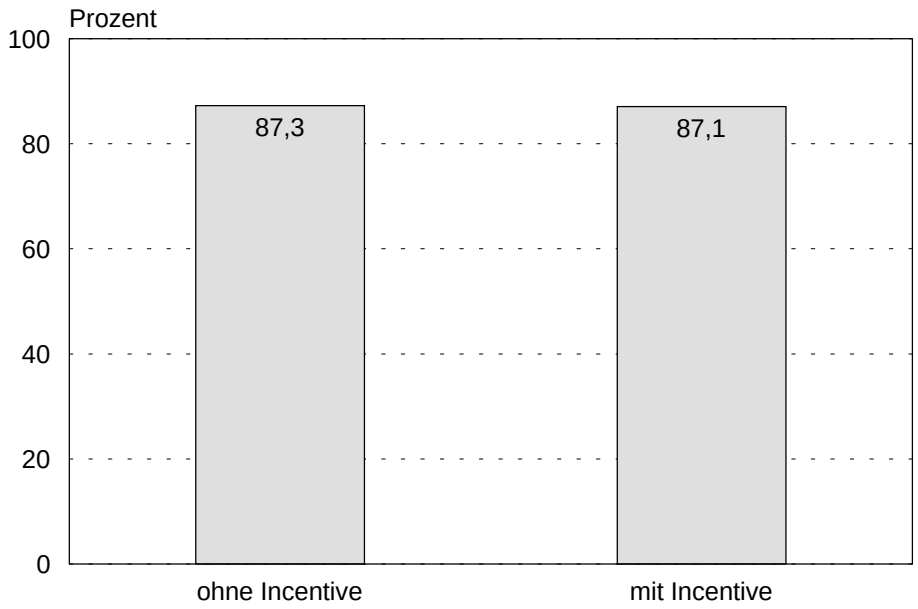
Abb. 5: Die Haltequote in der zweiten Panelwelle in Abhängigkeit von der Gewährung eines Incentives in der zweiten Welle



5.2 Existiert ein Hysteresis-Effekt?

Um die Frage nach der Existenz eines Hysteresis-Effekts beantworten zu können, haben wir in Abb. 6 die Ausschöpfungsquote der zweiten Panelwelle (=Haltequote) in Abhängigkeit von der Gewährung eines Incentives in der ersten Panelwelle abgetragen. Dabei zeigt sich, daß die Haltequote augenscheinlich nicht von der Gewährung eines Incentives in der ersten Panelwelle beeinflußt wird: Bei den Personen, die in der ersten Panelwelle ein Incentive erhielten, beträgt die Haltequote 87,1 Prozent, bei denjenigen Zielpersonen, die kein Incentive erhielten, 87,3 Prozent. Diese geringe Differenz in den Haltequoten ist statistisch nicht signifikant, liegt also im Bereich zufälliger Schwankungen.

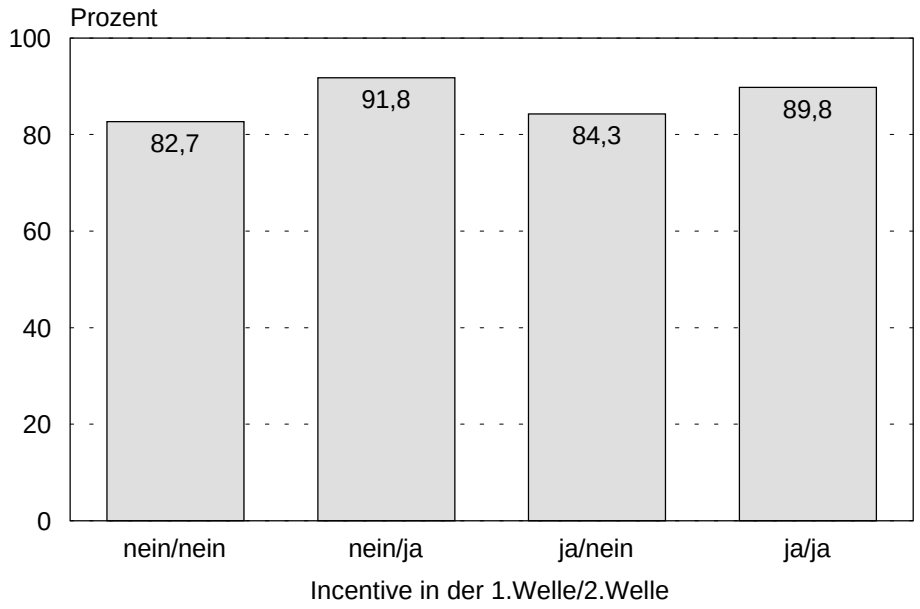
Abb. 6: Die Haltequote in der zweiten Panelwelle in Abhängigkeit von der Gewährung eines Incentives in der ersten Welle



5.3 Existiert ein Interaktionseffekt?

Ein etwaiger Interaktionseffekt zwischen den beiden Incentives müßte in Abbildung 7 zu erkennen sein. Dort haben wir die Haltequote in Abhängigkeit von der Gewährung eines Incentives in der ersten und der zweiten Panelwelle abgetragen. Läge ein wie auch immer gearteter Interaktionseffekt vor, so müßte sich das Verhältnis zwischen den Haltequoten in den ersten beiden Gruppen deutlich von der Relation der Haltequoten zwischen den beiden anderen Gruppen unterscheiden, d.h. der rücklaufsteigernde Effekt des zweiten Incentives müßte durch die Gewährung des ersten Incentives erkennbar modifziert werden. Tatsächlich scheint es so, als habe das zweite Incentive bei den Befragten, die in der ersten Panelwelle keine Telefonkarte erhielten, einen etwas stärkeren Effekt. In diesem Fall erhöht sich die Haltequote durch den Einsatz eines Incentives in der zweiten Welle um 9,1 Prozentpunkte, während diese Erhöhung nur 5,5 Prozentpunkte beträgt, wenn bereits in der ersten Panelwelle ein Incentive gewährt wurde. Diese Unterschiede sind aber so gering, daß an ihrer Signifikanz gezweifelt werden muß.

Abb. 7: Die Haltequote in der zweiten Panelwelle in Abhängigkeit von der Gewährung eines Incentives in der ersten und zweiten Panelwelle



Um die Frage nach der Existenz eines Interaktionseffekts abschließend beantworten zu können, haben wir deshalb für die 2160 Personen, welche die Ausgangsstichprobe der zweiten Panelwelle bildeten, ein log-lineares Modell[7] berechnet. Die abhängige Variable stellt die Teilnahme an der zweiten Panelwelle dar, als erklärende Variablen verwenden wir die beiden Incentives. Die Frage war hierbei, ob es gelingt, die empirischen Daten mit einem Modell hinreichend zu beschreiben, das nur die Haupteffekte der beiden Incentives enthält, nicht aber einen möglichen Interaktionseffekt zwischen ihnen. Wie Tab. 4 zeigt, ist dies der Fall, ein Interaktionseffekt zwischen den beiden Incentives existiert also nicht.[8] Die Interpretation, die sich aus der visuellen Inspektion der Verteilungen ergibt, läßt sich also auch statistisch absichern: Beim Versand zweier Incentives in einem zweiwelligen Kurzfrist-Panel tritt weder ein Hysteresis- noch ein Interaktionseffekt auf. Beide Incentives wirken vollständig unabhängig voneinander.

7 Bei allen von uns berechneten Analysen handelt es sich um Logit-Modelle, d.h. wir unterscheiden analog zum Vorgehen bei der Regressionsanalyse zwischen einer abhängigen Variablen (der Ausschöpfungs- bzw. Haltequote) einerseits und mehreren unabhängigen Variablen andererseits.

8 Anhand dieses Modells läßt sich außerdem auch nochmals bestätigen, daß kein Hysteresis-Effekt existiert: Der Effekt des ersten Incentives ist in dem in Tab. 4 dokumentierten Modell nicht signifikant von null verschieden.

Tab. 4: Der Effekt der beiden Incentives auf die Haltequote
- Ergebnisse eines log-linearen Modells -

	β	p	LR-χ2	Df
Modellgüte		0,279	1,17	1
Konstante	1,947			
Incentive PW 1				
ja	-,-			
nein	-,-			
Incentive PW 2				
ja	0,283			
nein	-0,283			

-,-: Koeffizient ist auf dem .05 Niveau nicht signifikant

5.4 Führen Incentives zu einer veränderten Zusammensetzung der Stichprobe?

Abschließend bleibt noch die Frage zu klären, ob der Einsatz von Incentives zu einer veränderten Zusammensetzung der Stichprobe führt. Ein solcher Einfluß wäre nicht in jedem Fall negativ zu bewerten: Wenn ursprünglich bestehende Stichprobenverzerrungen durch den Einsatz von Incentives ausgeglichen werden könnten, wäre ein solcher Effekt sogar von Vorteil. Würden Incentives aber zusätzliche Stichprobenverzerrungen induzieren, so würde man sich ihren rücklaufsteigernden Effekt mit qualitativ mangelhaften Stichproben erkaufen.

Um diese Fragen beantworten zu können, wollen wir im folgenden untersuchen, ob sich Incentives in allen Bevölkerungsgruppen gleich auswirken, oder ob sie in bestimmten Teilgruppen einen stärkeren bzw. schwächeren Effekt besitzen. Dies können wir anhand der vom Hamburger Amt für zentrale Meldeangelegenheiten zusammen mit den Adressen gelieferten Angaben über die Zielpersonen unserer Umfrage überprüfen. Bei diesen Angaben handelt es sich um das Geschlecht, das Alter und die Ortsteilzugehörigkeit. Die Ortsteilzugehörigkeit ist dabei nicht als Merkmal an sich interessant, sondern weil sie es ermöglicht, Kontextdaten in die Analysen miteinzubeziehen. Als solches Kontextmerkmal werden wir im folgenden das Durchschnittseinkommen[9] des Ortsteils verwenden, in dem die Zielperson lebt. Unsere Erwartung wäre hier, daß das Incentive möglicherweise in ärmeren Ortsteilen eine stärkere Wirkung hat als in reicheren. Da die Incentives in den bei-

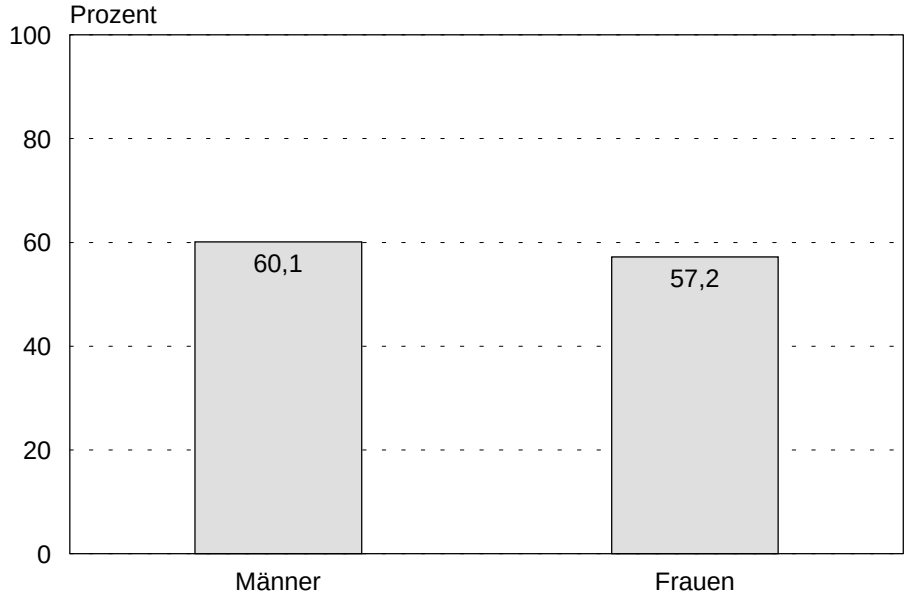
9 Beim Durchschnittseinkommen verwenden wir einen Index, der das Durchschnittseinkommen im Ortstteil zum Durchschnittseinkommen der Gesamtstadt in Relation setzt (Stand 1992). Die Angaben zur Quote der Sozialhilfeempfänger beziehen sich auf das Jahr 1996. Alle Daten sind aus ***Podszuweit/Schütte*** 1997 entnommen.

den Panelwellen unabhängig voneinander wirken und ihr Effekt in der ersten Panelwelle am stärksten ist, wollen wir diese Analysen über den Einfluß von Incentives auf die Stichprobenzusammensetzung auf die erste Welle beschränken.[10]

Betrachten wir zunächst den Rücklauf der ersten Panelwelle in Abhängigkeit von den drei genannten Merkmalen der Zielpersonen. Dadurch gewinnen wir einen Eindruck davon, welche Verzerrungen in der Stichprobe bereits ohne Einsatz von Incentives auftreten. Im folgenden werden wir dann untersuchen, ob Incentives diese Verzerrungen weiter verstärken, abbauen oder aber unverändert lassen. Abb. 8 zeigt den Rücklauf der ersten Welle in Abhängigkeit vom Geschlecht der Zielperson. Dabei zeigt sich, daß Männer (60,1%) mit einer etwas höheren Wahrscheinlichkeit antworten als Frauen (57,2%). Dieser geringe Unterschied von 2,9 Prozentpunkten ist statistisch aber nicht signifikant.

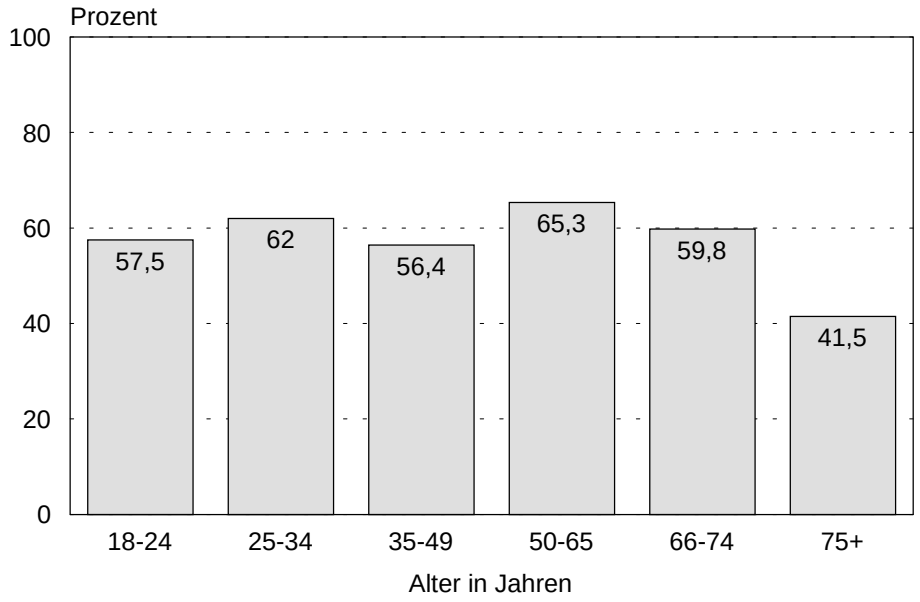
Beim Alter hingegen lassen sich signifikante Unterschiede zwischen den Altersgruppen beobachten (vgl. Abb. 9): Auffällig ist insbesondere der starke Rückgang der Rücklaufquote in der Gruppe der über 75jährigen. Dies ist intuitiv plausibel und entspricht der Erwartung, daß ältere Menschen vor allem aus gesundheitlichen Gründen seltener in der Lage sind, sich an einer Befragung zu beteiligen (vgl. analog dazu die Ergebnisse von ***Gehring/Wagner*** 1998 zur nachlassenden politischen Partizipation im hohen Alter).

Abb. 8: Der Rücklauf in der ersten Panelwelle nach dem Geschlecht



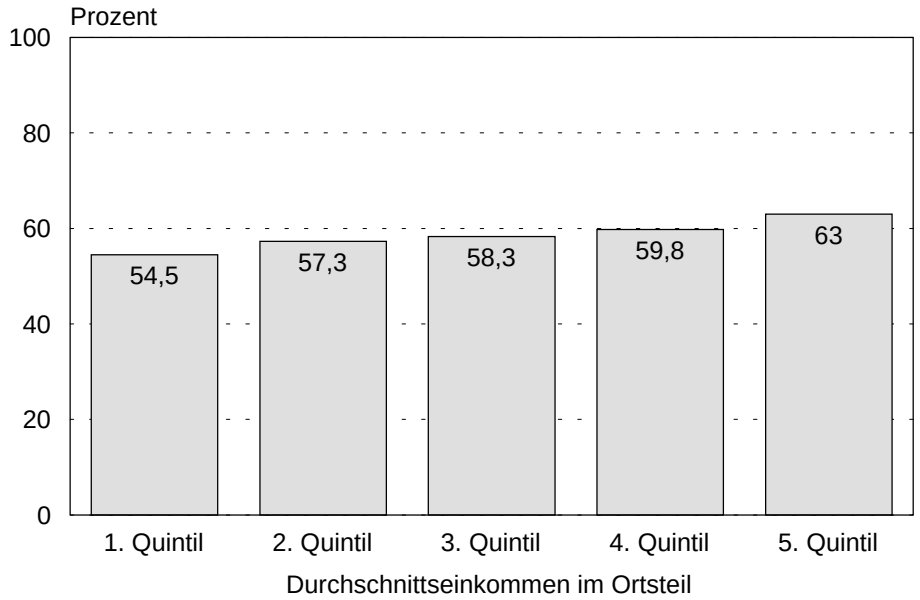
10 Wir haben diese Analysen zu Kontrollzwecken auch für die zweite Panelwelle durchgeführt und kamen dort zu weitgehend identischen Befunden.

Abb. 9: Der Rücklauf in der ersten Panelwelle nach dem Alter



Das Durchschnittseinkommen des Ortsteils, in dem die Zielperson wohnt, ist zunächst einmal natürlich ein Indikator für deren sozialen Kontext. In Ermangelung von Daten über die individuelle ökonomische Situation der Zielperson haben wir das Durchschnittseinkommen des Ortsteils aber als Proxy-Variable für die soziale Lage des Individuums benutzt. In Abb. 10 ist zunächst die Rücklaufquote der ersten Panelwelle in Abhängigkeit vom Durchnittseinkommen des Ortsteils dargestellt. Der besseren Übersichtlichkeit halber haben wir die 180 Hamburger Ortsteile hierfür in Einkommensquintile eingeteilt. Es zeigt sich ein zwar eher schwacher, aber nahezu linearer Zusammenhang, der zudem auf dem 5%-Niveau statistisch signifikant ist: Je höher das Durchschnittseinkommen des Ortsteils, in dem die Zielperson lebt, desto größer ist die Wahrscheinlichkeit, daß sich diese Person an der Umfrage beteiligt.

Abb. 10: Der Rücklauf in der ersten Panelwelle nach dem Durchschnittseinkommen des Ortsteils, in dem die Zielperson lebt



Um nun überprüfen zu können, ob sich Incentives in den verschiedenen Bevölkerungsgruppen unterschiedlich auswirken, haben wir wiederum ein log-lineares Modell berechnet (vgl. Tabelle 4). In dieses gingen die folgenden Variablen ein: Die Teilnahme an der ersten Panelwelle als abhängige Variable sowie das Alter, das Durchschnittseinkommen des Ortsteils und die Gewährung eines Incentives als unabhängige Variablen.[11] Die Frage war hierbei, ob es gelingt, die empirischen Daten über ein Modell zu beschreiben, das nur die Haupteffekte dieser Faktoren beinhaltet und insbesondere ohne Interaktionseffekte zwischen dem Incentive und den anderen unabhängigen Variablen auskommt. Dabei zeigt sich, daß bereits ein Modell ohne Interaktionseffekte nicht mehr signifikant von den empirischen Daten abweicht, was dahingehend interpretiert werden kann, daß sich Incentives in allen untersuchten Gruppen gleich auswirken.

11 Ursprünglich enthielt unser Modell noch das Geschlecht der Zielperson als erklärende Variable. Dadurch ergab sich das Problem, daß die zu analysierende Kontingenztabelle 2 x 2 x 2 x 6 x 5 = 240 Zellen umfaßte, von denen naturgemäß viele gar nicht oder sehr schwach besetzt waren. Unter diesen Bedingungen ist der $\chi 2$-Anpassungstest keine geeignete Grundlage, um die Güte der Modellschätzung zu beurteilen (***Langeheine et al.*** 1996). Da sich zwischen dem Geschlecht und der Teilnahmebereitschaft weder bi- noch multivariat signifikante Zusammenhänge zeigten, haben wir die Variable wieder aus dem Modell entfernt, damit die Zahl der Zellen halbiert wird und so die geschilderten Probleme vermieden werden.

Tab. 5: Der Effekt des ersten Incentives auf den Rücklauf unter Berücksichtigung des Geschlechts, des Alters und des Durchschnittseinkommens des Ortsteils
- Ergebnisse eines log-linearen Modells -

	β	p	LR-χ^2	df
Modellgüte		0,088	122,03	108
Konstante	0,177			
Incentive PW 2				
Ja	0,245			
Nein	-0,245			
Alter				
18-24	-,-			
25-34	0,219			
35-49	-,-			
50-65	0,315			
66-74	-,-			
75++	-0,646			
Durchschnittseinkommen des Ortsteil				
1. Quintil	-0,182			
2. Quintil	-,-			
3. Quintil	-,-			
4. Quintil	-,-			
5. Quintil	0,172			

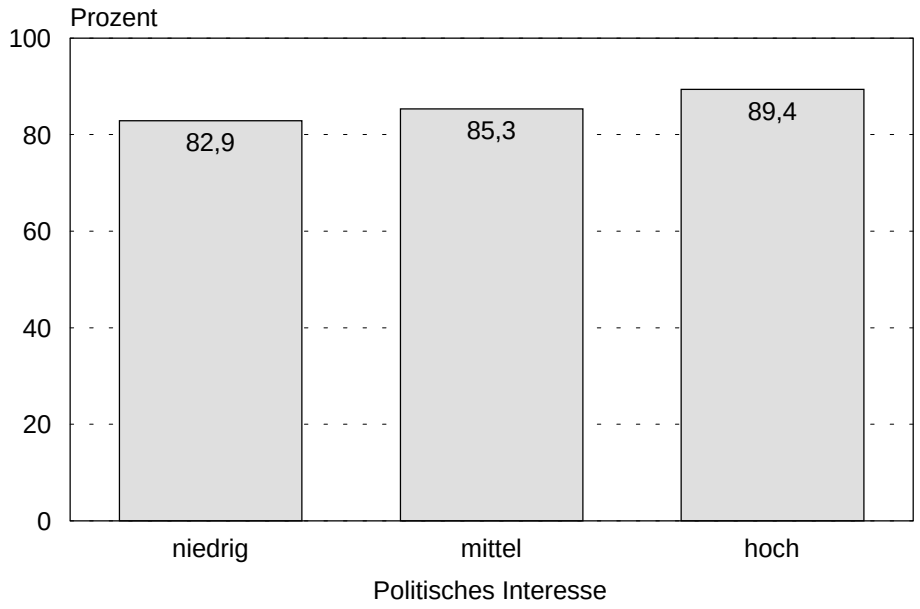
-,-: Koeffizient ist auf dem .05 Niveau nicht signifikant

Die Verwendung von Incentives induziert also keine zusätzlichen Verzerrungen der Stichprobe. Incentives sind allerdings auch nicht in der Lage, die aus der bivariaten Betrachtung bekannten Effekte des Alters und des Durchschnittseinkommens zu kompensieren. So hatten wir erwartet, daß die Wirkung des Incentives durch die Eigenschaft, in einem besser- oder schlechtergestellten Ortsteil zu wohnen, modifiziert wird. Dieser Effekt ist nicht eingetreten. Dies steht in Einklang mit der Untersuchung von ***Warriner et al.***, die keine Wechselwirkungen zwischen dem *individuellen* sozioökonomischen Status der Befragten und dem Effekt materieller Incentives feststellen konnten (1996: 557f). Wir werten dies als ein Indiz dafür, daß die von uns eingangs angesprochene Reziprozitätsnorm tatsächlich eine weitverbreitete Gültigkeit besitzt, unabhängig vom monetären Nutzen, den das Incentive für einen Befragten hat. Für diese Interpretation spricht auch die durch zahlreiche Untersuchungen belegte Tatsache, daß bereits kleine Geldbeträge (im amerikanischen Kontext beispielsweise die schon klassische Dollarnote) den Rücklauf deutlich heben, die Wirkung des Incentives mit steigendem Wert des Geschenkes jedoch nur unterproportional zunimmt (***Fox/Crask/Kim*** 1988; ***James/Bolstein*** 1992; ***Warriner et al.*** 1996).

Die bislang berichteten Analysen bezogen sich ausschließlich auf soziodemographische Variablen. Aus anderen Forschungsarbeiten ist aber bekannt, daß eine der erklärungskräf-

tigsten Determinanten der Interviewteilnahme das Interesse am Thema der Befragung ist. Wie wirkt sich also in unserem Fall das politische Interesse auf die Teilnahmebereitschaft aus? Gelingt es eventuell über die Gewährung von Incentives, auch politisch Nicht-Interessierte zur Interviewteilnahme zu bewegen? Diese Fragen können wir für die erste Panelwelle nicht untersuchen, da das politische Interesse der Zielpersonen naturgemäß zunächst nicht bekannt ist. Da in unserem ersten Fragebogen aber nach dem Interesse an der Politik[12] gefragt wurde, können wir den oben genannten Fragen mit den Daten der zweiten Panelwelle nachgehen. Variiert nun also die Teilnahmebereitschaft an der Wiederholungsbefragung in Abhängigkeit vom politischen Interesse der Zielperson? Zunächst betrachten wir das unabhängig von Incentives. Wie Abb. 11 zeigt, besteht ein schwacher (γ=0,13) positiver Zusammenhang zwischen dem politischen Interesse und der Teilnahme an der Wiederholungsbefragung: Je höher das politische Interesse, desto größer die Wahrscheinlichkeit, daß eine Zielperson sich an der Wiederholungsbefragung beteiligt. Dieser Zusammenhang ist statistisch signifikant (p=0,03).

Abb. 11: Der Rücklauf in der zweiten Panelwelle nach dem politischen Interesse der Zielpersonen

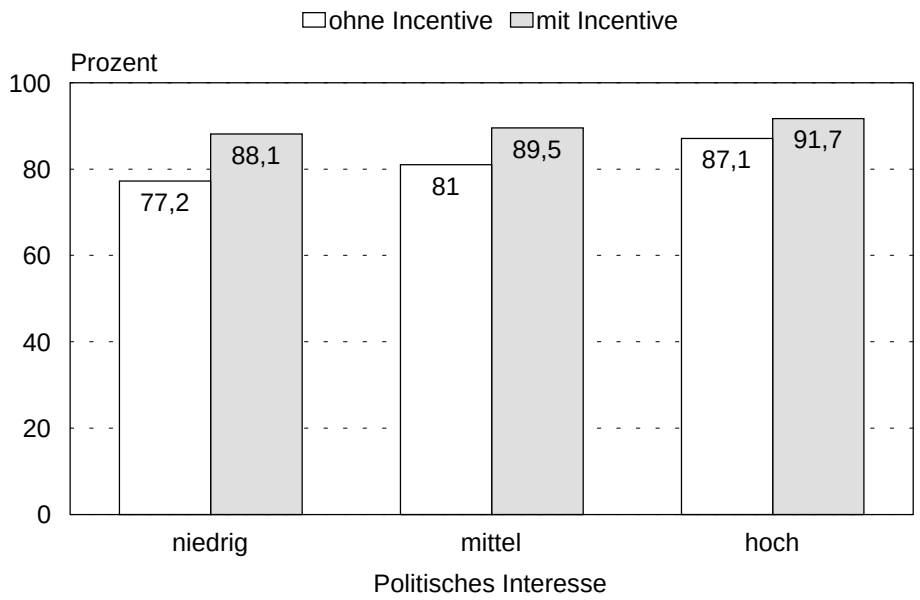


Besteht darüber hinaus auch ein Zusammenhang zwischen der Gewährung eines Incentives und dem Effekt des politischen Interesses auf die Teilnahme an der Wiederholungsbefragung? Dieser Frage gehen wir in Abb. 12 nach. Dort ist der Einfluß des politischen Interes-

12 Das politische Interesse wurde in der Umfrage über eine siebenstufige Ratingskala mit den beiden Polen „Gar nicht“ bis „Sehr stark“ gemessen. Für die hier berichteten Analysen wurde diese Skala wie folgt recodiert: 1,2,3=„niedrig“; 4 =„mittel“; 5,6,7=„hoch“. Die relativen Häufigkeiten dieser drei recodierten Kategorien betragen 18,7%, 23,7% und 57,6%.

ses in Abhängigkeit von der Gewährung eines Incentives in der zweiten Panelwelle dargestellt. Es zeigt sich, daß der Einfluß des politischen Interesses in der Gruppe der Zielpersonen, die kein Incentive erhielten, stärker ausfällt als in der Gruppe der Personen, die ein solches Incentive erhielten. Anders ausgedrückt: Incentives wirken bei politisch Uninteressierten scheinbar stärker als bei politisch Interessierten. In der Gruppe der Personen mit niedrigem politischem Interesse steigt der Rücklauf durch ein Incentive um 10,9 Prozentpunkte, während diese Steigerung in der Gruppe der Personen mit hohem politischem Interesse nur 4,6 Prozentpunkte beträgt. Es scheint also ein Interaktionseffekt zwischen den beiden Größen vorzuliegen. Inwieweit dieser inferenzstatistisch abgesichert werden kann, wollen wir im folgenden im Rahmen eines weiteren log-linearen Modells untersuchen.

Abb. 12: Der Rücklauf in der zweiten Panelwelle nach dem politischen Interesse der Zielpersonen und der Gewährung eines Incentives



In diese Analyse gehen die Teilnahme an der Wiederholungsbefragung als abhängige und die Gewährung eines Incentives sowie das im Rahmen der ersten Befragungswelle gemessene politische Interesse der Zielperson als unabhängige Variablen ein. Auch hier sind wir wieder so vorgegangen, daß wir zunächst versucht haben, die empirischen Daten über ein Modell ohne Interaktionsterme zu beschreiben. Bereits mit diesem Modell erhält man eine sehr gute Modellanpassung (vgl. Tab. 5). Dies bedeutet, daß der in Abb. 12 zu beobachtende Interaktionseffekt noch im Rahmen zufälliger Schwankungen liegt und deshalb statistisch nicht signifikant ist. Es ist aber denkbar, daß dieser Effekt stärker ausfallen und somit die Schwelle der Signifikanz überschreiten würde, wenn die Stichprobe nicht ohnehin positiv vorselektiert wäre – schließlich handelt es sich um Zielpersonen, die durch ihre Antwort in der ersten Welle bereits ein Mindestmaß an Interesse an der Umfrage gezeigt

haben. Es ist also nicht auszuschließen, daß ein ähnliches Design, wenn es für die erste Panelwelle möglich gewesen wäre, dort signifikante Wechselwirkungen zwischen dem politischen Interesse und dem Effekt des Incentives aufgedeckt hätte.

Tab. 6: Der Effekt des zweiten Incentives auf die Haltequote unter Berücksichtigung des politischen Interesses der Zielperson - Ergebnisse eines log-linearen Modells -

	β	p	LR-χ^2	Df
Modellgüte		0,695	0,73	2
Konstante	1,854			
Incentive PW 2				
Ja	0,305			
Nein	-0,305			
Politisches Interesse				
Niedrig	-0,243			
Mittel	-,-			
Hoch	0,314			

-,-: Koeffizient ist auf dem .05 Niveau nicht signifikant

5.5 In welcher Panelwelle sollte das Incentive eingesetzt werden?

Betrachten wir nun abschließend die Gesamtausschöpfung der zweiten Panelwelle, also die Ausschöpfungsquote bezogen auf die der ersten Panelwelle zugrundeliegende Stichprobe (vgl. Abb. 8), die wir durch einfache Multiplikation der jeweiligen Ausschöpfungsquoten der beiden Panelwellen ermittelt haben. Der höchste Gesamtrücklauf ergibt sich mit 55,3 Prozent erwartungsgemäß in der Gruppe von Personen, die in beiden Wellen ein Incentive erhielt. Am niedrigsten ist der Rücklauf hingegen, wenn weder in der ersten noch in der zweiten Panelwelle ein Incentive eingesetzt wurde. Er beträgt dann lediglich 40,9 Prozent. Wird nur in einer der beiden Wellen ein Incentive gewährt, so liegen die Ausschöpfungsquoten zwischen diesen beiden Extremwerten. Obwohl keine Hysteresis-Effekte existieren, erweist es sich dabei als effektiver, das Incentive bereits in der ersten Welle einzusetzen, wenn man sich aus Kostengründen auf *ein* Incentive beschränken muß. Denn der Gesamtrücklauf liegt dann mit 51,9 Prozent deutlich höher, als wenn das Incentive erst in der zweiten Welle gewährt wird. Hier beträgt der Rücklauf nur 45,3 Prozent.

Abb. 13: Die Gesamtausschöpfung im Panel in Abhängigkeit vom ersten und zweiten Incentive

6 Zusammenfassung und Schlußfolgerungen

Unsere Ergebnisse zeigen zunächst, daß auch bei der Durchführung von Panelstudien die schriftliche Befragung nach der Total-Design-Method eine echte Alternative zu den momentan dominierenden mündlichen bzw. telefonischen Befragungsformen darstellt. Selbst in der Gruppe, die keinerlei Incentive erhalten hat, erzielen wir in der ersten Welle eine Ausschöpfungsquote von fast 50 Prozent und eine Haltequote von mehr als 80 Prozent bei der Wiederholungsbefragung. Berücksichtigt man, daß diese Studie in einem großstädtischen Milieu durchgeführt wurde, so ist dieses Ergebnis aus unserer Sicht außerordentlich zufriedenstellend.

Darüber hinaus konnten wir durch unser Methodenexperiment zeigen, daß ein geeignetes materielles Incentive die *Ausschöpfungsquote* schriftlicher Befragung steigert, wobei der rücklaufsteigernde Effekt in der ersten Welle unserer Panelbefragung deutlich größer ausfällt als in der zweiten. Dieser Effekt stellt sich unabhängig vom Alter und vom Geschlecht der Zielperson ein und scheint auch vom sozio-ökonomischen Kontext nicht beeinflußt zu werden.

Wir konnten darüber hinaus zeigen, daß die *Haltequote* in der zweiten Panelwelle von einem in der ersten Panelwelle gewährten Incentive nicht beeinflußt wird, d.h. es tritt kein Hysteresis-Effekt auf. Ebensowenig kommt es zu einer Wechselwirkung zwischen dem ersten und dem zweiten Incentive, d.h. beide Incentives wirken völlig unabhängig voneinander.

Aufgrund der finanziellen Restriktionen wird man sich zumeist auf den Einsatz eines einzigen Incentives beschränken. Unsere Befunde sprechen dafür, daß es effizienter ist, das Incentive bereits in der ersten Panelwelle einzusetzen. Obwohl zwischen beiden Incentives keine Wechselwirkungen auftreten, läßt sich so eine höhere Gesamtausschöpfung erzielen. Zudem reduziert sich die Zahl der Personen, die mehrfach angeschrieben werden müssen, bereits in der ersten Welle, was wegen des dort noch größeren Stichprobenumfangs besonders wünschenswert ist.

Ob sich unsere Befunde auch auf mehrwellige und längerfristig angelegte Panels übertragen lassen, muß aber vorerst eine offene Frage bleiben. Auch ist darauf hinzuweisen, daß Incentives möglicherweise unerwünschte Konsequenzen nach sich ziehen können: Wie unlängst gezeigt wurde, kann der Einsatz von Incentives auf seiten der Befragten zu der Erwartung führen, daß für die Teilnahme an der Befragung generell eine materielle Vergünstigung zu gewähren sei (***Singer***/***Hoewyk***/***Maher*** 1998). Analog zu den von ***Frey*** (1997) beschriebenen Prozessen können offenbar auch im Bereich der empirischen Sozialforschung monetäre Anreize langfristig die intrinsische Motivation der Respondenten zerstören. Wenn aufgrund dieser Effekte langfristig die Bereitschaft, sich unentgeltlich an einem Interview zu beteiligen, weiter sinken würde, würde sich die ohnehin schwierige Situation der Umfrageforschung weiter verschlechtern.

Anhang

Tab. A1: Die Ausschöpfung der Stichprobe in Abhängigkeit von der Gewährung eines materiellen Incentives

1. Panelwelle				
	Ohne Incentive		Mit Incentive	
Bruttostichprobe	1.000		3.000	
./. stichprobenneutrale Ausfälle	28		98	
Bereinigter Stichprobenansatz	972	100,0	2.902	100,0
./. systematische Ausfälle	492		1.114	
Auswertbare Interviews	480	49,4	1.788	61,6
in der 2. Welle nicht angeschrieben	14		37	
Bruttostichprobe 2. Welle	466		1751	

2. Panelwelle								
	Ohne Incentive		Mit Incentive		Ohne Incentive		Mit Incentive	
Bruttostichprobe	228		238		877		874	
./. stichprobenneutrale Ausfälle	3		7		29		18	
Bereinigter Stichprobenansatz	225	100,0	231	100,0	848	100,0	856	100,0
./. systematische Ausfälle	39		19		133		87	
Auswertbare Interviews	186	82,7	212	91,8	715	84,3	769	89,8
Gesamtauschöpfung über beide Wellen		*40,9*		*45,3*		*51,9*		*55,3*

Literatur:

Arzheimer, ***Kai*** / ***Klein***, ***Markus*** 1998: Die Conjoint-Analyse als Instrument der empirischen Wahlforschung. Papier für die Tagung „Wahlen und Wahlforschung: Kontinuität und Wandel“ des Arbeitskreises „Wahlen und politische Einstellungen“ der DVPW vom 14. bis 15. Mai in Mainz. [www.uni-mainz.de/~arzheim/conjoint/Conjoint.html]

Bauer, ***Karde*** 1988: Schriftliche Befragungen im Panel. In: Planung und Analyse, 5, S. 206-209.

Blasius, ***Jörg*** / ***Reuband***, ***Karl-Heinz*** 1996: Postalische Befragungen in der empirischen Sozialforschung. In: Planung und Analyse, 1, S. 35-41.

Blau, ***Peter M.*** 1964: Exchange and Power in Social Life. New York: Wiley.

Brehm, ***John*** 1994: Stubbing our Toes for a Foot in the Door? Prior Contact, Incentives and Survey Response. In: International Journal of Public Opinion Research, 6, S. 45-63.

Bruvold, ***N. T. / Corner, J. M***. 1988: A model for estimating the response rates to a mailed survey. In: Journal of Business Research, 16, S. 101-116.

Christoffersen, ***M. N.*** 1987: The educational bias of mail questionnaires. In: Journal of Official Statistics, 3, S. 459-464.

Church, ***Allan H***. 1993: Estimating The Effect of Incentives on Mail Survey Response Rates: A Meta-Analysis. In: Public Opinion Quarterly, 57, S. 62-79.

Clark, ***G. L. / Kaminski, P. F.*** 1988: How to get more for your money in mail surveys. In: Journal of Business and Industrial Marketing, 3, S. 17-23.

Dalecki, ***M. G.*** / ***Ilvento***, ***T.W.*** / ***Moore***, ***D. E.*** 1988: The effects of multi-wave mailings on the external validity of mail surveys. In: Journal of the Community Development Society, 19, S. 51-70.

De Jonge, ***L.*** / ***Van Veen***, ***W. M. Oppendijk*** / ***Pooters C.*** 1977: Accounting for the speed of response in mail-panel surveys. In: European Research, 5, S. 172-180.

Dillman, ***Don A***. 1978: Mail and telephone surveys: The Total Design Method. New York.

Finn, ***David W.*** / ***Wang***, ***Chih-Kang*** / ***Lamb***, ***Charles W***. 1983: An examination of the effects of sample composition bias in a mail survey. In: Journal of the Marketing Research Society, 25, S. 331-338.

Fox, ***Richard J.*** / ***Crask***, ***Melvin R.*** / ***Kim***, ***Jonghoon*** 1988: Mail survey response rate. A meta-analysis of selected techniques for including response. In: Public Opinion Quarterly, 52, S. 467-491.

Frey, ***Bruno S***. 1997: Markt und Motivation. Wie ökonomische Anreize die (Arbeits-)Moral verdrängen. München.

Gehring, ***Uwe W.*** / ***Wagner***, ***Michael*** 1998: Wahlbeteiligung im hohen Alter. Ergebnisse der Berliner Altersstudie. Manuskript.

Gouldner, ***Alvin*** 1960: The Norm of Reciprocity: A Preliminary Statement. In: American Sociological Review, 25, S. 161-178.

Groves, ***Robert M***. 1992: Understanding the Decision to participate in a survey. In: Public Opinion Quarterly, 56, S. 475-495.

Hare, ***Sheri*** / ***Price***, ***James H.*** / ***Flynn***, ***Michael G.*** / ***King***, ***Keith A***. 1998: Increasing Return Rates of a Mail Survey to Exercise Professionals Using a Modest Monetary Incentive. In: Perceptual and Motor Skills, 86, S. 217-218.

Heberlein, ***Thomas A.*** / ***Baumgartner***, ***Robert*** 1978: Factors Affecting Response Rates To Mailed Questionnaires: A Quantitative Analysis of the Published Literature. In: American Sociological Review, 43, S. 447-462.

Hippler, ***Hans-Jürgen*** 1988: Methodische Aspekte schriftlicher Befragungen: Probleme und Forschungsperspektiven. In: Planung und Analyse, 6, S. 244-248.

Hippler, ***Hans-Jürgen*** / ***Seidel***, ***Kristiane*** 1985: Schriftliche Befragungen bei allgemeinen Bevölkerungs-stichproben - Untersuchungen zur Dillmanschen "Total Design Method".
In: ZUMA-Mitteilungen 16, S. 39-56.

Homans, ***G***. ***C***. 1961: Social Behavior: Ist Elementary Forms. New York: Harcourt, Brace & World.

Huxley, ***Stephen J***. 1980: Predicting response speed in mail surveys. In: Journal of Marketing Research, 17, S. 63-68.

James, ***Jeannine M***. / ***Bolstein***, ***Richard*** 1990: The Effect of Monetary Incentives and Follow-Up Mailings on the Response Rate and Response Quality in Mail Surveys. In: Public Opinion Quarterly, 54, S. 346-361.

James, ***Jeannine M***. / ***Bolstein***, ***Richard*** 1992: Large Monetary Incentives and their Effects on Mail Survey Response Rates. In: Public Opinion Quaterly, 56, S. 442-453.

Kanuk, ***Leslie*** / ***Berenson***, ***Conrad*** 1975: Mail Surveys and Response Rates: A Literature Review.
In: Journal of Marketing Research, 12, S. 440-453.

Langeheine, ***Rolf*** / ***Pannekoek***, ***Jeroen*** / ***van de Pol***, ***Frank*** 1996: Bootstrapping Goodness-of-Fit-Measures in Categorical Data Analysis. In: Sociological Methods & Research, 24, S. 492-516.

Linsky, ***Arnold S***. 1975: Stimulating Responses to Mailed Questionnaires: A Review. In: Public Opinion Quarterly, 39, S. 82-101.

Martinez-Ebers, ***Valerie*** 1997: Using Monetary Incentives with Hard-To-Reach Populations in Panel Surveys. In: International Journal of Public Opinion Research, 9, S. 77-86.

Mowen, ***John C.*** / ***Cialdini***, ***Robert B***. 1980: On implementing the Door-in-the-Face compliance technique in a business context. In: Journal of Marketing Research, 17, S. 253-258.

Nederhof, ***Anton J***. 1983: The Effects of Material Incentives in Mail Surveys: Two Studies.
In: Public Opinion Quarterly, 47, S. 103-111.

Niedermayer, ***Oskar*** 1989: Zur Übertragung der "Total Design Method" auf schriftliche Befragungen in der Bundesrepublik. In: ***Jürgen W. Falter*** / ***Hans Rattinger*** / ***Klaus G***. ***Troitzsch*** [Hrsg.]: Wahlen und politische Einstellungen in der Bundesrepublik Deutschland. Frankfurt a.M. / Bern / New York / Paris, S. 332-348.

Podszuweit, ***Ulrich*** / ***Schütte***, ***Wolfgang*** 1997: Sozialatlas Hamburg 1997. FH Hamburg,
Fachbereich Sozialpädagogik.

Reuband, ***Karl-Heinz*** 1998: Panelmortalität in postalischen Erhebungen und soziale Zusammensetzung der Befragten. Ergebnisse einer allgemeinen Bevölkerungsumfrage. In: Planung und Analyse, 3, S. 16-21.

Singer, ***Eleanor*** 1998: Incentives for Survey Participation: Research on Intended and Unintended Consequences. In: ZUMA-Nachrichten, 42, S. 7-29.

Singer, ***Eleanor*** / ***Hoewyk***, ***Johnvan*** / ***Maher***, ***Mary P***. 1998: Does the Payment of Incentives create Expectation Effects? In: Public Opinion Quarterly, 62, S. 152-164.

Thibaut, ***J.W.*** / ***Kelley***, ***H.H***. 1959: The Social Psychology of Groups. New York: Wiley.

Warriner, ***Keith*** / ***Goyder***, ***John*** / ***Gjertsen***, ***Heidi*** / ***Hohner***, ***Paula*** / ***Mc Spurren***, ***Kathleen*** 1996: Charities, No; Lotteries, No; Cash, Yes. Main Effects and Interactions in a Canadian Incentive Experiment. In: Public Opinion Quarterly, 60, S. 542-562.

Willimack, ***Diane K.*** / ***Schuman***, ***Howard*** / ***Pennell***, ***Beth-Ellen*** / ***Lepkowski***, ***James M.*** 1995***:*** Effects of Prepaid Nonmonetary Incentive on Response Rates and Response Quality in a Face-to-Face Survey.
In: Public Opinion Quarterly, 59, S. 78-92.

Yu, ***Julie*** / ***Cooper***, ***Harris*** 1983: A Quantitative Review of Research Design Effects on Response Rates to Questionnaires. In: Journal of Marketing Research, 20, S. 36-44.

Telefonische Befragungen bei alten Menschen: Ausschöpfung, Kontakthäufigkeit und Verweigerungen

von Olaf von dem Knesebeck[1] und Günther Lüschen[2]

Zusammenfassung

Obgleich telefonische Befragungen in der empirischen Sozialforschung in den letzten Jahren an Bedeutung gewonnen haben, gibt es in Deutschland kaum Erfahrungen mit diesem Instrument im Rahmen von sozialwissenschaftlich-gerontologischen Untersuchungen. Im vorliegenden Beitrag werden Maßnahmen zur Ausschöpfungsoptimierung (Mehrfachkontaktierung von Schwer-Erreichbaren und Konvertierung von Verweigerern) bei einer telefonischen Befragung alter Menschen (60 Jahre und älter) thematisiert. Dabei wird insbesondere der Frage nachgegangen, ob sich leicht erreichbare von schwer erreichbaren Haushalten bzw. Personen und temporäre Verweigerer von kooperationsbereiten Befragungsteilnehmern hinsichtlich soziodemographischer und sozioökonomischer Merkmale unterscheiden. Die Ergebnisse lassen darauf schließen, daß dem Problem der Verweigerung bei der Planung und Durchführung telefonischer Befragungen alter Menschen größere Aufmerksamkeit geschenkt werden sollte als dem Problem der Erreichbarkeit.

Abstract

Although telephone surveys have become more important for empirical research in recent years, there is not much experience in using telephone interviews in socio-gerontological research. In this article measures to optimize response rate (repeated telephone calls and converting temporary refusers) in a telephone survey of older people (≥60 years) are discussed. We focus on the following two questions: first, whether persons who can be reached easily differ from those who are difficult to contact. Secondly we ask whether temporary refusers differ from cooperative respondents as far as their sociodemographic and socioeconomic characteristics are concerned. Results indicate that the problem of refusal

1 Dr. ***Olaf v.d. Knesebeck***, Heinrich-Heine-Universität Düsseldorf, Institut für Medizinische Soziologie, Postfach 10 10 07, 40001 Düsseldorf (Korrespondenzanschrift)

2 Prof. Dr. ***Günther Lüschen***, University of Alabama at Birmingham, Department of Sociology

has higher relevance for the planning and conducting of telephone interviews with older people, compared to the problem of reachability.

1 Einleitung

Im Zuge des demographischen Wandels, der durch eine nachhaltige Zunahme des Anteils alter Menschen gekennzeichnet ist, haben sich in den letzten Jahren verschiedene wissenschaftliche Disziplinen mit den Vorgängen des Alterns und der Lebenssituation alter Menschen auseinandergesetzt (vgl. ***Baltes*** und ***Mittelstraß*** 1992). Gerontologische Fragestellungen gelten als „Zukunftsthema" (***Prahl*** und ***Schroeter*** 1996, S. 9). Für die empirisch-sozialwissenschaftliche Forschung wirft die Beschäftigung mit alten Menschen erhebungsmethodische Fragen auf, welche z.B. die Anwendung der verschiedenen Befragungsarten (persönlich, postalisch, telefonisch) zur Erhebung gerontologischer Daten betreffen. Die wenigen methodischen Untersuchungen, die sich mit Befragungen von in Privathaushalten lebenden alten Menschen beschäftigen, zeigen, daß alle drei Befragungsarten grundsätzlich anwendbar sind (vgl. ***Herzog*** und ***Rogers*** 1988a; ***Brune, Werle*** und ***Hippler*** 1991; ***Tennstedt, Dettling*** und ***McKinlay*** 1992). Bei einem Vergleich zwischen Face-to-Face-, telefonischer und postalischer Befragung in Deutschland ergaben sich in bezug auf das Alter keine statistisch signifikanten Unterschiede zwischen den Erhebungsverfahren (vgl. ***Reuband*** und ***Blasius*** 1996, S. 308). Generell ist bei älteren Befragten allerdings mit höheren Ausfällen zu rechnen (vgl. ***Herzog*** und ***Rogers*** 1988b; ***Tennstedt, Dettling*** und ***McKinlay*** 1992, S. 313; ***Blasius*** und ***Reuband*** 1995; ***Schnell*** 1997a, S. 178). Im allgemeinen wird davon ausgegangen, daß diese Ausfälle vor allem auf Verweigerungen zurückzuführen sind (vgl. ***Bungard*** 1979; ***Nuthmann*** und ***Wahl*** 1996, S. 68f.). ***Herzog*** und ***Roders*** (1988b) haben in diesem Zusammenhang am Beispiel verschiedener Befragungen in Amerika gezeigt, daß der Anteil von Verweigerungen an den Ausfällen zwar in den mittleren und höheren Altersgruppen relativ hoch ist, dann aber bei den Ältesten (75 Jahre und älter) wieder abnimmt. Ähnlich konstatiert ***Schnell***, daß es bei sorgfältig durchgeführten Surveys keine eindeutigen Belege für eine Abnahme der Teilnahmewilligkeit bei älteren Personen gibt (vgl. ***Schnell*** 1997a, S. 181). Da Ausfälle zum Teil auch auf körperlich bedingte Teilnahmeschwierigkeiten (z.B. Hör- und Sehprobleme) bzw. Nichtbefragbarkeit zurückzuführen sind, schlagen einige Autoren vor, eine stellvertretende, mit der Situation des alten Menschen vertraute Person zu befragen (Proxy-Interview, vgl. ***McKinlay*** et al. 1990). Abgesehen von solchen Empfehlungen ist bislang wenig darüber bekannt, wie Ausfällen bei Befragungen alter Menschen begegnet werden kann bzw. welche Bedeutung den verschiedenen Maßnahmen zur Ausschöpfungsoptimierung zukommt. Die folgenden Analysen zur Ausschöpfung, Kontakthäufigkeit und Verweigerung bei telefonischen Befragungen älterer Menschen sollen hierzu einen Beitrag leisten. Der Schwerpunkt liegt dabei auf dem Vergleich zwischen leicht und schwer erreichbaren Haushalten bzw.

Personen sowie auf Unterschieden zwischen temporären Verweigerern und bereitwilligen Befragungsteilnehmern.

2 Telefonische Befragungen: Ausschöpfung, Kontakthäufigkeit und Verweigerungen

Telefonische Befragungen haben in Deutschland in den letzten Jahren an Bedeutung gewonnen. Mit dem Erreichen einer annähernden Vollversorgung der Haushalte mit Telefonanschlüssen in Westdeutschland in den 80er Jahren war die Voraussetzung für einen Einsatz dieser Befragungsart in der empirischen Sozialforschung geschaffen (vgl. ***Frey, Kunz*** und ***Lüschen*** 1990, S. 15; ***Blasius*** und ***Reuband*** 1995, S. 65). Seitdem nimmt die Zahl der Untersuchungen zu, die sich der telefonischen Befragung als Erhebungsmethode bedienen, wobei vermehrt computergestützt vorgegangen wird (CATI). Ausschlaggebend für diesen Bedeutungsgewinn sind vor allem der mit dem Telefoninterview verbundene Zeit- und Kostenvorteil sowie die Möglichkeit einer gegenüber anderen Befragungsformen kontrollierteren Durchführung (vgl. ***Frey, Kunz*** und ***Lüschen*** 1990, S. 199). Demgegenüber läßt die mangelnde Telefondichte, die 1995 noch 66.8% betrug (vgl. ***Schulte*** 1997, S. 165ff.), den Einsatz telefonischer Befragungen in Ostdeutschland nur mit Einschränkungen zu.

Bei telefonischen Befragungen werden häufig Telefonbücher als Grundlage für die Auswahl von Haushalten herangezogen. Bedingung für ein solches Vorgehen ist die Vollständigkeit der Telefonbücher, d.h. es muß sichergestellt sein, daß Stichproben aus Telefonbüchern ein Abbild aller Anschlüsse bzw. Haushalte ermöglichen. Diese Bedingung war bis vor einigen Jahren erfüllt, da der Anteil nicht gelisteter Anschlüsse mit etwa einem Prozent äußerst gering war (vgl. ***Zeh*** 1987, S. 339). Da es seit 1992 den Anschlußinhabern freigestellt wird, ob sie sich in das Telefonbuch eintragen lassen oder nicht, hat sich der Anteil nicht gelisteter Anschlüsse erhöht. Im Jahr 1994 betrug dieser Anteil zwar lediglich 3.7 Prozent (vgl. ***Häder*** 1996, S. 47), jedoch finden sich 1997 schon deutlich höhere Angaben, die für *Westdeutschland* zwischen 4 und 8 Prozent variieren und für *Ostdeutschland* bei 20 Prozent liegen (vgl. ***von der Heyde*** 1997, S. 200; ***Marhenke*** 1997, S. 210). ***Gabler*** und ***Häder*** (1997, S. 7) gehen sogar davon aus, daß circa 20 Prozent aller Telefonnummern in Deutschland nicht in den jeweils aktuellen Telefonbüchern verzeichnet sind. Vor diesem Hintergrund werden vermehrt alternative Stichprobendesigns vorgeschlagen, bei denen „Add-A-Digit“-, „Randomize-Last-Digit“- oder „Random-Digit-Dialing“-Verfahren zur Anwendung kommen (vgl. z.B. ***Schnell*** 1997b, S. 50; ***Gabler*** und ***Häder*** 1997).

In erhebungsmethodischen Untersuchungen werden Ausschöpfungsquoten häufig als ein wichtiges Kriterium für die Beurteilung der wissenschaftlich-methodischen Qualität von Befragungen herangezogen. Die Erfahrungen zeigen, daß die Ausschöpfungsquoten - definiert als das Verhältnis der Zahl der auswertbaren Interviews zur Größe der bereinigten Bruttostichprobe (Antwortrate) - bei telefonischen Befragungen in der Regel zwischen 40

und 60 Prozent liegen (vgl. ***Porst*** 1993, S. 27; ***Blasius*** und ***Reuband*** 1995, S. 67). So ergab sich beispielsweise im Rahmen des von der Forschungsgruppe Wahlen per Telefoninterview durchgeführten Politbarometers zwischen 1989 und 1995 ein Ausschöpfungsmedian von 53% (vgl. ***Schulte*** 1997, S. 158). In einigen Untersuchungen fallen die Ausschöpfungsquoten auch deutlich geringer aus (vgl. ***Porst*** 1991). Auf der anderen Seite haben ***Blasius*** und ***Reuband*** (1995) auf der Basis einer Personenstichprobe aus dem Einwohnermelderegister gezeigt, daß bei sehr hohem Kontaktierungsaufwand Auschöpfungsquoten von 90% bei telefonischen Befragungen erreichbar sind.

Neben expliziten Verweigerungen gehört das Problem der Erreichbarkeit zu den wesentlichen Ausfallgründen bei telefonischen Befragungen. Aufgrund der mit dieser Erhebungsmethode verbundenen direkten, schnellen und preisgünstigen Art der Kontaktierung werden im allgemeinen mehrere Kontaktversuche möglichst zu verschiedenen Zeiten unternommen, um die Ausschöpfung der Stichprobe zu erhöhen (vgl. z.B. ***Wiseman*** und ***McDonald*** 1979). In diesem Zusammenhang sind mehrere Untersuchungen der Frage nachgegangen, wieviele Kontaktversuche zur Ausschöpfungsoptimierung sinnvoll und notwendig sind (vgl. ***Brückner, Hormuth*** und ***Sagawe*** 1982; ***Hormuth*** und ***Brückner*** 1985; ***Jung*** 1990). Dabei ergab sich, daß nach etwa 6 Kontaktversuchen die meisten Nummern kontaktiert werden konnten, „(...) d.h. es kam entweder zum Interview, zur Ablehnung, oder es konnte geklärt werden, daß es sich nicht um einen Zielhaushalt handelte.“ (***Hormuth*** und ***Brückner*** 1985, S. 534) Eine Erhöhung der Kontaktversuche über diese Anzahl würde somit nicht mehr zu einer nennenswerten Verbesserung der Ausschöpfungsquote und besonders für kommerzielle Unternehmen zu kaum vertretbaren Kosten führen. Andere Untersuchungen zu Kontakthäufigkeiten beschäftigen sich mit der Frage, ob sich zwischen leicht und schwer erreichbaren Personen bzw. Haushalten systematische Unterschiede hinsichtlich soziodemographischer Merkmale oder der Antwortmuster zeigen. Auf der Basis solcher Untersuchungen lassen sich Aussagen über die Brauchbarkeit von telefonischen Umfragen mit hoher Nicht-Kontaktrate treffen, bei denen die Anzahl der Kontaktversuche z.B. auf einen begrenzt wird. ***Blasius*** und ***Reuband*** (1995) konnten zeigen, daß eine Begrenzung auf einen Kontaktversuch zu größeren Verzerrungen in der Zusammensetzung der Stichprobe führt, während sich Personen, die zwischen zwei- und fünfmal kontaktiert wurden, kaum unterscheiden. Sie schließen aus ihren Ergebnissen, daß die optimale Anzahl der Kontaktversuche zwischen zwei und fünf liegt.

Neben der Durchführung mehrerer Kontaktversuche besteht ein weiteres Mittel zur Ausschöpfungssteigerung in der nochmaligen Kontaktierung von Verweigerern, mit dem Ziel, diese doch noch zu einer Teilnahme an der Befragung zu bewegen (vgl. ***Reuband*** und ***Blasius*** 1995, S. 7f.). Analysen, die sich mit den Unterschieden zwischen temporären Verweigerern und bereitwilligen Befragungsteilnehmern auseinandersetzen, geben Aufschluß über die Stichprobenqualität von Untersuchungen mit hohen Verweigerungsraten, in denen auf

Bemühungen zur Konvertierung temporärer Verweigerer verzichtet wurde. In einigen Arbeiten werden temporäre Verweigerer auch zur Beschreibung bzw. Vorhersage der Merkmale von endgültigen Verweigerern herangezogen (vgl. ***Smith*** 1983; ***Schnell*** 1997a, S. 152f.).

Im folgenden Beitrag sollen Möglichkeiten und Probleme telefonischer Befragungen bei in Privathaushalten lebenden alten Menschen in Westdeutschland am Beispiel einer Untersuchung in den 10 kreisfreien Städten des Regierungsbezirks Düsseldorf aufgezeigt und diskutiert werden.[3] Dabei werden insbesondere die Ausschöpfung, der Vergleich zwischen leicht und schwer erreichbaren Haushalten bzw. Personen sowie Unterschiede zwischen temporären Verweigerern und bereitwilligen Befragungsteilnehmern thematisiert.

3 Methode

Grundlage der Analysen ist eine im Frühjahr 1994 durchgeführte telefonische Befragung über die Gesundheit und den sozialen Status alter Menschen (GESTALT) (vgl. ***Knesebeck*** 1998).[4] Grundgesamtheit dieser Untersuchung waren alle Personen, die zum Zeitpunkt der Befragung 60 Jahre und älter waren und in einem Privathaushalt in einer der 10 kreisfreien Städte des Regierungsbezirks Düsseldorf lebten. Nach dem Modell einer mehrstufigen reinen Wahrscheinlichkeitsauswahl wurde aus den sechs Telefonbüchern dieser Region eine Stichprobe von insgesamt 3500 Telefonnummern gezogen.[5] Der Auswahlplan zur Gewinnung der Stichprobe wurde durch Generierung von Zufallszahlen mittels eines Statistikprogrammes realisiert.

Bei der telefonischen Kontaktierung wurden bis zu 10 Versuche zu möglichst unterschiedlichen Zeiten unternommen, wobei zunächst durch ein Screening festgestellt wurde, ob und wieviele Personen im betreffenden Haushalt leben, die 60 Jahre oder älter sind. Lebte keine solche Person in diesem Haushalt, so wurde dieser Haushalt aus der Untersuchung ausgeschlossen. Lebte dort mehr als eine solche Person, so wurde auch auf der Haushalts-

---

3 Die sogenannte „Anstaltsbevölkerung“ (***Schnell*** 1991, S.114) ist nicht in die Untersuchung eingeschlossen, d.h. die im folgenden gemachten Aussagen beziehen sich nicht auf Menschen, die in Krankenhäusern, Alten- oder Pflegeheimen leben. Dies ist insofern inhaltlich relevant, als sich dieser Teil der Altenbevölkerung im Hinblick auf soziodemographische Merkmale und den Gesundheitszustand von den in Privathaushalten lebenden alten Menschen unterscheidet, was auch Auswirkungen auf das Teilnahmeverhalten bei Befragungen hat (vgl. ***Salaske*** 1997). Der Anteil der 65-jährigen und Älteren, die in Alten- und Pflegeheimen leben, wird für das Land Nordrhein-Westfalen auf etwa 5 bis 5.5% geschätzt (Ministerium für Arbeit, Gesundheit und Soziales des Landes Nordrhein-Westfalen 1995). Es sei an dieser Stelle darauf hingewiesen, daß sich das Problem des Ausfalls der Anstaltsbevölkerung auch bei den meisten Erhebungen der empirischen Sozialforschung stellt, die Repräsentativität für die ‘allgemeine Bevölkerung’ beanspruchen (vgl. ***Schnell*** 1991).

4 Das Projekt wurde am Sozialwissenschaftlichen Institut, Lehrstuhl für Soziologie III der Heinrich-Heine-Universität Düsseldorf durchgeführt und von der Hans-Böckler-Stiftung gefördert.

5 Zum Zeitpunkt der Untersuchung war der Anteil nicht gelisteter Anschlüsse noch vergleichsweise gering, so daß Telefonbücher als Auswahlgrundlage für Stichproben noch geeignet waren (vgl. Ausführungen unter 2.).

ebene mittels der Geburtstagsmethode (***Salmon*** und ***Nichols*** 1983) ein Wahrscheinlichkeitsauswahlverfahren angewandt, indem nach der Person im Haushalt gefragt wurde, die als letztes Geburtstag hatte.

In der Konzeption der Untersuchung wurde davon ausgegangen, daß nicht mehr alle ausgewählten alten Personen aufgrund von Erkrankungen, Hör- oder Verständnisproblemen zu einem telefonischen Interview in der Lage sind oder sich zu diesem nicht mehr in der Lage fühlen. Um jedoch auch über diese Personen einige Informationen zu bekommen, wurde neben dem „normalen“ Fragebogen, dessen Beantwortung im Durchschnitt 51 Minuten in Anspruch nahm, eine Kurzfassung entwickelt, die von einer, mit der Situation des alten Menschen vertrauten Person, stellvertretend beantwortet wurde.

4 Ergebnisse

4.1 Ausschöpfung

Tabelle 1 zeigt die Stichproben- und Ausfallbeschreibung der telefonischen Befragung, unterteilt nach stichprobenneutralen Ausfällen, Ausfällen vor der Personenauswahl, durchgeführten Interviews und Ausfällen nach der Personenauswahl. Der hohe Anteil stichprobenneutraler Ausfälle (69.1%) ist vor allem durch die Haushalte bedingt, in denen keine Person lebt, die 60 Jahre oder älter ist (54.7%). Inwieweit diese Ausfälle tatsächlich alle stichprobenneutral sind, ist nicht eindeutig zu beurteilen, da die im Rahmen des Screenings gegebene Auskunft, daß im Haushalt keine Person lebt, die 60 Jahre oder älter ist, auch eine ‘versteckte Verweigerung’ sein kann. Der Tabelle ist weiterhin zu entnehmen, daß insgesamt 533 Interviews realisiert werden konnten, von denen 97 anhand des Kurzfragebogens mit einer Referenzperson geführt wurden.

Die Antwortrate beträgt 55.1%, was für telefonische Befragungen in Relation zu den üblicherweise erzielten Werten und angesichts einer Population, die in der Literatur als schwer befragbar gilt, als ein akzeptables Ergebnis erscheint. Eine endgültige Beurteilung dieser Antwortrate ist nur eingeschränkt möglich, da zum einen kaum Erfahrungen mit telefonischen Befragungen dieser Population vorliegen und zum anderen eine Vergleichbarkeit von Ausschöpfungsquoten nur selten gegeben ist, weil über den Berechnungsmodus Uneinigkeit besteht, dieser aber in den seltensten Fällen transparent gemacht wird. Auch die ermittelte Verweigerungsrate von 25.3% (ohne Einbeziehung der Abbrüche 22.8%) liegt im Rahmen der in der Literatur berichteten Werte für Telefonumfragen von 20 bis 30% (vgl. ***Frey, Kunz*** und ***Lüschen*** 1990, S. 40ff.; ***Schulte*** 1997, S. 187; ***Schnell*** 1997a, S. 121), wobei hier ebenfalls das Problem der eingeschränkten Vergleichbarkeit gegeben ist. Die in Tabelle 1 dokumentierte Verweigerungsrate ergab sich unter dem

Tabelle 1: Stichproben- und Ausfallbeschreibung

	Häufigkeit	%
Stichprobenneutrale Ausfälle		
Niemand im Haushalt ist 60 Jahre oder älter	1915	54.7
Kein Haushalt: Betrieb, Behörde usw.	302	8.6
Nummer funktioniert nicht	96	2.7
Anrufbeantworter: Kein HH	39	1.1
Sonstiges: z.B. Fax, Rufnr. hat sich geändert	67	1.9
	2419	69.1
Ausfälle vor der Auswahl		
Kein Kontakt nach 10 Versuchen	165	4.7
Ablehnung: Keine Information über Haushalt	43	1.2
Anrufbeantworter	30	0.9
	238	6.8
Abgeschlossene Interviews		
Interview: Langfassung	436	12.5
Interview: Kurzfassung	97	2.8
Ablehnungen und Abbrüche		
Ablehnung: Haushalt mit Untersuchungseinheit	202	5.8
Abbruch des Interviews	24	0.7
Sonstige Ausfälle nach der Auswahl		
Kein Kontakt nach der Auswahl	9	0.3
Nicht erreichbar: Lange Abwesenheit	4	0.1
Sprachprobleme	14	0.4
Sonstiges: z.B. Krankheit, Verständnisprobleme ohne Referenzperson	57	1.6
	843	24.1
	3500	100.0

Antwortrate: 533/967* = 55.1%

Verweigerungsrate: 245**/967* = 25.3%

* Die Berechnung der bereinigten Bruttostichprobe stellt bei der Berücksichtigung der Ausfälle vor der Auswahl den hohen Anteil der stichprobenneutralen Ausfälle (69.1%) in Rechnung: 843+238-114=967.

** Diese Zahl ergibt sich aus der Summe der Ablehnungen und Abbrüche nach der Auswahl sowie den Ablehnungen vor der Auswahl, bereinigt durch den Anteil der Haushalte ohne Personen im Alter von 60 Jahren und mehr (54.7%): 202+24+43-24= 245.

Einfluß einer postalischen Nachfaßaktion, bei der alle anfänglichen Verweigerer und Abbrecher in einem Brief, der auch ein Präsent (Kugelschreiber) enthielt, nochmals um ihre Teilnahme gebeten wurden. Im Zuge des darauf folgenden, erneuten telefonischen Kontaktes wurde die Verweigerungsrate von über 40% auf die angeführten 25.3% reduziert.

Ein Vergleich der realisierten Stichprobe mit dem Mikrozensus des Jahres 1994 (Quelle: Landesamt für Datenverarbeitung und Statistik Nordrhein-Westfalen) hinsichtlich der soziodemographischen Merkmale Alter, Geschlecht und Familienstand ergibt eine Unter-

repräsentierung der 60- bis 69-jährigen und eine Überrepräsentierung der 70-jährigen und Älteren. Während sich darüber hinaus kaum Unterschiede in bezug auf die Geschlechterverteilung feststellen lassen, sind in der Stichprobe zu wenig Verheiratete und zu viele Verwitwete vertreten. Bei einem Vergleich der realisierten Langinterviews mit dem Mikrozensus zeigt sich, daß die festgestellten Verteilungsdifferenzen in der Altersstruktur auf die Kurzinterviews mit Referenzpersonen zurückzuführen sind: Legt man allein die Langinterviews zugrunde, sind die 80-jährigen und Älteren sogar leicht unterrepräsentiert, d.h. ohne besondere Behandlung der Nicht-Befragbaren ergibt sich hinsichtlich der Altersverteilung ein Ergebnis, das aus anderen Bevölkerungsumfragen bekannt ist (vgl. auch ***Hartmann*** 1990; ***Frey, Kunz*** und ***Lüschen*** 1990, S. 92; ***Schulte*** 1997, S. 160f.; ***Knesebeck*** 1998; zu den mit einem Vergleich von Aggregatstatistik und Stichprobenergebnissen verbundenen Problemen vgl. ***Schnell*** 1997a, S. 135ff.).

4.2 Anzahl der Kontaktversuche[6]

Um die Nicht-Kontaktrate zu reduzieren, wurden bis zu 10 Kontaktversuche an den Wochentagen zwischen 17.00 und 21.00 Uhr sowie am Samstag und Sonntag etwa zwischen 10.00 und 18.00 Uhr unternommen, wobei darauf geachtet wurde, daß die einzelnen Versuche zu möglichst unterschiedlichen Zeiten stattfinden. Tabelle 2 zeigt, daß bei den ersten drei Kontaktversuchen jeweils etwa 20-30% der Langinterviews durchgeführt wurden und nach drei Versuchen fast 80% bzw. nach fünf Versuchen fast 90% der Interviews realisiert werden konnten. Darüber hinaus weisen die Ergebnisse darauf hin, daß die Zahl realisierter Interviews von Versuch zu Versuch immer geringer wird.

Tabelle 2: Verteilung der Kontaktanzahl bei den Langinterviews

Anzahl der Kontakte	N	%	Kum. %
1	140	32.1	32.1
2	108	24.8	56.9
3	89	20.4	77.3
4 bis 5	54	12.4	89.7
6 bis 10	45	10.3	100.0
	436	100.0	

Für die Planung und Durchführung von telefonischen Befragungen ist die Analyse der Unterschiede zwischen leicht und schwer erreichbaren Personen insofern von Bedeutung, als durch diese geklärt werden kann, wieviele Kontaktversuche notwendig sind, um die durch nicht erreichte Personen bzw. Haushalte bedingten Verzerrungen in der Zusammen-

6 In den folgenden Analysen zur Anzahl der Kontaktversuche und zu temporären Verweigerern werden die Langinterviews zugrunde gelegt, da in diese Analysen Variablen eingehen, die im Rahmen der Kurzinterviews mit Referenzpersonen nicht erhoben wurden. Dies hat zur Folge, daß sich die Aussagen auf die alten Menschen beziehen, die zu einem Telefoninterview gesundheitlich in der Lage waren.

setzung der Stichprobe zu reduzieren. Im folgenden wird der Frage nachgegangen, ob sich in der hier untersuchten Altenpopulation ein Zusammenhang zwischen der Anzahl der für ein Interview benötigten Kontaktversuche und ausgewählten soziodemographischen und sozioökonomischen Merkmalen zeigen läßt.

Aus Befragungen der Erwachsenenbevölkerung ist bekannt, daß vor allem solche Personen schwer zu erreichen sind, die allein im Haushalt leben, eine Berufstätigkeit ausüben, ledig und unter 30 Jahre alt sind, wobei sich abzeichnet, daß Personen mit zunehmendem Alter leichter zu erreichen sind (vgl. ***Blasius*** und ***Reuband*** 1995). Ob es auch unter alten Menschen Schwer-Erreichbare gibt und, wenn ja, welche Merkmale diese aufweisen, ist bislang noch nicht untersucht worden. Es ist jedoch zu vermuten, daß die noch Berufstätigen und Jüngeren unter ihnen sich häufiger außerhalb des Haushaltes aufhalten und somit schwerer zu erreichen sind. Darüber hinaus ist davon auszugehen, daß die Tatsache, allein zu leben, im Alter nicht mehr in dem Maße mit herabgesetzter Erreichbarkeit einhergeht, wie dies für die junge Erwachsenenbevölkerung der Fall ist, da sich die Alleinlebenden in Altenpopulationen vornehmlich aus verwitweten hochaltrigen Frauen zusammensetzen.

Tabelle 3: Anzahl der Kontaktversuche und soziodemographische Merkmale (Angaben in Prozent)

	1 Kontakt (n=140)	2 Kontakte (n=108)	3 Kontakte (n=89)	4 - 5 Kontakte (n=54)	6+ Kontakte (n=45)	Gesamt (n=436)
Alter						
60 bis 65	29.5	25.9	15.7	27.8	28.9	25.5
66 bis 70	25.2	25.0	31.5	18.5	24.4	25.5
71 bis 75	18.7	23.1	25.8	29.6	15.6	22.3
76 und älter	26.6	25.9	27.0	24.1	31.1	26.7
Geschlecht						
Weiblich	63.6	67.6	60.7	53.7	73.3	63.8
Männlich	36.4	32.4	39.3	46.3	26.7	36.2
Familienstand						
Verheiratet	47.9	46.3	51.7	40.7	46.7	47.2
Verwitwet	38.6	41.7	38.2	42.6	42.2	40.1
Geschieden/ledig	13.6	12.0	10.1	16.7	11.1	12.6
Anzahl Personen im Haushalt						
1	43.9	43.5	39.8	40.7	42.2	42.4
2	48.9	43.5	50.0	50.0	46.7	47.7
3 und mehr	7.2	13.0	10.2	9.3	11.1	9.9
Berufstätigkeit						
Voll-/teilzeit	5.0	7.4	9.0	5.6	11.1	7.1
Rente/Pension	76.4	78.7	80.9	79.6	68.9	77.5
Hausfrau	18.6	13.9	10.1	14.8	20.0	15.4

Bei keiner Tabelle sind die statistischen Kennwerte (Chi^2 und Cramer's V) auf dem 5%-Niveau signifikant.

Die in Tabelle 3 wiedergegebenen Zusammenhänge sind insgesamt eher schwach; in keinem Fall wird ein Signifikanzniveau von 5% erreicht. Zwar sind offenbar 60- bis 65-jährige etwas leichter und 71- bis 75-jährige etwas schwerer zu erreichen; ein systematischer Zusammenhang zwischen Kontaktanzahl und Alter läßt sich jedoch nicht erkennen. Auch ist trotz starker Schwankungen in den Zellenbesetzungen nicht festzustellen, daß die Erreichbarkeit mit dem Geschlecht kovariiert. Ebenso weist der Familienstand und die Haushaltsgröße keine systematische Beziehung zur Anzahl der Kontaktversuche auf, während Voll- bzw. Teilzeitbeschäftigte tendenziell etwas schwerer zu erreichen sind.

Wie Tabelle 4 zeigt, sind auch die Zusammenhänge zwischen Erreichbarkeit und sozioökonomischen Merkmalen insgesamt eher schwach; allerdings ergeben sich hier einige Tendenzen: Bezogen auf „subjektive“ Indikatoren sozialer Ungleichheit (subjektiv eingeschätzter Rang in der Gesellschaft, Einschätzung des Einflusses in der Familie und subjektiver Grad der Informiertheit) sind untere Statusgruppen tendenziell leichter zu erreichen als obere Statusgruppen, was insbesondere für den ersten Kontaktversuch gilt. Im Hinblick auf die „objektiven“ sozioökonomischen Merkmale *Haushaltseinkommen* und *Vermögen im Haushalt* zeigt sich demgegenüber, daß die mittleren Statusgruppen überdurchschnittlich häufig beim 1. Kontaktversuch erreicht werden, während insbesondere beim *Vermögen* die untere Statusgruppe beim 1. Kontaktversuch, verglichen mit dem Anteil in der Gesamtstichprobe, unterrepräsentiert ist. In bezug auf die Bildung (Schul- und Hochschulzeit in Jahren) und die Einschätzung des Einflusses bei Freunden ergibt sich eine weitgehende Übereinstimmung zwischen der Anteilsverteilung nach dem ersten Kontaktversuch und der Anteilsverteilung in der Gesamtstichprobe. Zusammenfassend läßt sich sagen, daß eine lineare Zu- oder Abnahme der Anteile bei bestimmten Statusgruppen oder Schichten mit zunehmender Kontaktanzahl bei keinem der berücksichtigten sozioökonomischen Merkmale festzustellen ist. Allerdings variiert die Erreichbarkeit der unterschiedlichen Statusgruppen tendenziell bei den ersten 2 bis 3 Kontaktversuchen.

4.3 Temporäre Verweigerer

Zur Verringerung der Verweigerungsrate wurden alle Personen, die ein Interview verweigert oder abgebrochen hatten und von denen die Adresse aufgrund der Angaben im Telefonbuch bekannt war, in einem Brief nochmals um die Teilnahme an der Befragung gebeten. Als zusätzlicher Anreiz war diesem Schreiben ein Kugelschreiber beigelegt. Die angeschriebenen Personen wurden nach diesem Anschreiben erneut telefonisch kontaktiert und um eine Teilnahme gebeten. Auf diese Weise konnten 105 Verweigerer bzw. Abbrecher für die Befragung gewonnen werden. In den folgenden Analysen wird die Frage verfolgt, ob sich diese 105 temporären Verweigerer bzw. Abbrecher von den übrigen „bereitwilligen“ Befragungsteilnehmern im Hinblick auf soziodemographische und sozioökonomische Merkmale unterscheiden.

Tabelle 4: Anzahl der Kontaktversuche und sozioökonomische Merkmale (Angaben in Prozent)

	1 Kontakt (n=140)	2 Kontakte (n=108)	3 Kontakte (n=89)	4 - 5 Kontakte (n=54)	6+ Kontakte (n=45)	Gesamt (n=436)
Schul-/ Hochschulzeit						
bis 8 Jahre	63.0	58.3	64.0	74.0	58.5	62.6
9 bis 10 Jahre	17.0	25.0	11.2	12.0	14.6	17.0
11 bis 13 Jahre	12.6	12.0	11.2	4.0	17.1	11.6
über 13 Jahre	7.4	4.6	13.5	10.0	9.8	8.5
Haushaltsein-kommen						
weniger als 30.000	29.2	32.6	33.8	21.6	31.6	30.2
30.000 bis 35.000	36.7	31.6	31.3	43.1	23.7	33.9
über 35.000	34.2	35.8	35.0	35.3	44.7	35.9
Vermögen im Haushalt						
unter 100.000	40.0	52.8	43.6	41.7	46.2	44.7
100.000 bis unter 500.000	45.2	38.2	35.9	45.8	38.5	40.9
500.000 u. mehr	14.8	9.0	20.5	12.5	15.4	14.4
subj. Rang in der Gesellschaft						
eher oben	18.3	22.3	27.7	30.6	25.0	23.4
in der Mitte	55.7	58.3	47.0	53.1	40.0	52.7
eher unten	26.0	19.4	25.3	16.3	35.0	23.9
Einfluß in der Familie						
wenig/keinen	39.7	34.3	35.6	35.3	30.2	36.0
ziemlich	29.4	27.5	25.3	33.3	41.9	29.8
(sehr) viel	30.9	38.2	39.1	31.4	27.9	34.1
Einfluß bei Freunden						
wenig/keinen	46.6	40.2	46.3	55.1	57.1	47.0
ziemlich	28.2	28.4	29.3	28.6	31.0	28.8
(sehr) viel	25.2	31.4	24.4	16.3	11.9	24.1
Belastung durch Arbeit						
gar nicht	41.2	44.3	41.6	35.2	32.6	40.4
wenig/etwas	25.7	23.6	38.2	35.2	34.9	29.9
ziemlich/sehr	33.1	32.1	20.2	29.6	32.6	29.7
subj. Informa-tionsstatus						
eher schlecht	28.7	25.0	15.9	31.5	23.8	25.0
mittel	22.8	26.9	25.0	22.2	26.2	24.5
eher gut	48.5	48.1	59.1	46.3	50.0	50.5

Bei keiner Tabelle sind die statistischen Kennwerte (Chi^2 und Cramer's V) auf dem 5%-Niveau signifikant.

Sollten sich hier Unterschiede zeigen, so kann daraus abgeleitet werden, daß der Konvertierung von Verweigerern in telefonischen Befragungen alter Menschen zur Reduzierung von Verzerrungen eine besondere Aufmerksamkeit geschenkt werden muß.

Tabelle 5: Temporäre Verweigerer/Abbrecher und soziodemographische Merkmale (Angaben in Prozent)

	nicht verweigert/ abgebrochen (n=331)	verweigert/ abgebrochen (n=105)	Gesamt (n=436)
Alter			
60 bis 65	27.9	18.1	25.5
66 bis 70	26.1	23.8	25.5
71 bis 75	20.3	28.6	22.3
76 und älter	25.8	29.5	26.7
Geschlecht			
Weiblich	65.3	59.0	63.8
Männlich	34.7	41.0	36.2
Familienstand			
Verheiratet	44.4	56.2	47.2
Verwitwet	41.7	35.2	40.1
Geschieden/ledig	13.9	8.6	12.6
Anzahl Personen im Haushalt			
1	44.7	35.2	42.4
2	44.7	57.1	47.7
3 und mehr	10.6	7.6	9.9
Berufstätigkeit			
voll-/teilzeit	7.3	6.7	7.1
Rente/Pension	78.9	73.3	77.5
Hausfrau	13.9	20.0	15.4

Bei keiner Tabelle sind die statistischen Kennwerte (Chi^2 und Cramer's V) auf dem 5%-Niveau signifikant.

Wie Tabelle 5 zu entnehmen ist, sind die ermittelten Zusammenhänge mit soziodemographischen Merkmalen nicht signifikant; trotzdem zeigen sich z.T. beträchtliche Prozentwertunterschiede. So sind bei den konvertierten Verweigerern bzw. Abbrechern über 70-jährige und Männer überrepräsentiert. Zudem verweigern Verheiratete und dementsprechend auch Personen aus 2-Personen-Haushalten vergleichsweise häufig. Hinsichtlich der Berufstätigkeit zeigt sich, daß Rentner und Pensionäre seltener und Hausfrauen häufiger zu den Verweigerern zählen. Betrachtet man die Unterschiede zwischen Kooperationsbereiten und konvertierten Verweigerern in bezug auf das Alter, den Familienstand und die Haushaltsgröße, so scheinen erneute Kontaktaufnahmen zur Konvertierung insgesamt empfehlenswert.

Bei der Analyse des Zusammenhanges zwischen temporärer Verweigerung bzw. Abbruch und sozioökonomischen Merkmalen zeigt sich folgendes Bild: Legt man die hier berücksichtigten „objektiven" sozioökonomischen Merkmale *Bildung*, *Einkommen* und *Vermögen* zugrunde, so verweigern die unteren und oberen Statusgruppen häufiger als die mittleren (vgl. Tabelle 6).

Tabelle 6: Temporäre Verweigerer/Abbrecher und sozioökonomische Merkmale (Angaben in Prozent)

	nicht verweigert/ abgebrochen (n=331)	verweigert/ abgebrochen (n=105)	Gesamt (n=436)
Schul-/Hochschulzeit			
bis 8 Jahre	61.7	66.7	62.6
9 bis 10 Jahre	18.1	13.7	17.0
11 bis 13 Jahre	12.5	8.8	11.6
über 13 Jahre	7.8	10.8	8.5
Haushaltseinkommen			
weniger als 30.000	29.2	33.3	30.2
30.000 bis 35.000	35.8	28.1	33.9
über 35.000	35.1	38.5	35.9
Vermögen im Haushalt			
unter 100.000	42.8	50.5	44.7
100.000 bis unter 500.000	43.8	32.3	40.9
500.000 u. mehr	13.4	17.2	14.4
subj. Rang in der Gesellschaft			
eher oben	23.5	23.2	23.4
in der Mitte	52.4	53.5	52.7
eher unten	24.1	23.2	23.9
Einfluß in der Familie			
wenig/keinen	36.2	35.6	36.0
ziemlich	31.1	25.7	29.8
(sehr) viel	32.7	38.6	34.1
Einfluß bei Freunden			
wenig/keinen	44.9	53.5	47.0
ziemlich	31.1	21.8	28.8
(sehr) viel	23.9	24.8	24.1
Belastung durch Arbeit			
gar nicht	42.0	35.6	40.4
wenig/etwas	28.1	35.6	29.9
ziemlich/sehr	29.9	28.8	29.7
Subjektiver Informationsstatus*			
eher schlecht	25.8	22.3	25.0
mittel	27.1	16.5	24.5
eher gut	47.1	61.2	50.5

* $Chi^2 = 7.03$; df=2; $p<.05$; Cramer's V=.13

Bei keiner der übrigen Tabellen sind die statistischen Kennwerte (Chi^2 und Cramer's V) auf dem 5%-Niveau signifikant.

Hätte man demnach in der Untersuchung darauf verzichtet, Verweigerer durch erneute Kontaktierung zu einer Teilnahme an der Befragung zu bewegen, wären weniger Angehörige unterer und oberer Statusgruppen in der Gesamtstichprobe vertreten gewesen. In bezug auf die „subjektiven" Merkmale lassen die Ergebnisse keine solchen einheitlichen Tendenzen erkennen. So ergibt sich kein Zusammenhang zwischen subjektiv eingeschätztem Rang in der Gesellschaft und temporärer Verweigerung bzw. Abbruch, während z.T. deutliche Unterschiede in den Zellenbesetzungen bei der Einschätzung des Einflusses und der Arbeitsbelastung erkennbar sind. Schließlich ist Tabelle 6 zu entnehmen, daß ein

schwach signifikanter Zusammenhang zwischen dem subjektiven Informationsstatus und temporärer Verweigerung besteht, demzufolge Personen, die sich selbst als eher gut informiert über nationale und internationale Geschehnisse einschätzen, überdurchschnittlich häufig verweigern. Offenbar sind gut informierte alte Menschen eher geneigt, auf der Basis kalter telefonischer Kontaktierung durch Studenten zunächst zu verweigern, sich bei offizieller schriftlicher Information über die Untersuchung (mit dem Briefkopf der Universität Düsseldorf und der Unterschrift des Projektleiters und Lehrstuhlinhabers) dann aber umstimmen zu lassen.

5 Schlußbemerkungen

Im vorliegenden Beitrag wurden auf der Grundlage einer telefonischen Befragung von in Privathaushalten lebenden alten Menschen mögliche Maßnahmen zur Ausschöpfungssteigerung (Mehrfachkontaktierung von schwerer erreichbaren Personen bzw. Haushalten und Konvertierung von Verweigerern) und die Konsequenzen solcher Maßnahmen für die Zusammensetzung der Stichprobe dargestellt.

Die erzielten Ergebnisse lassen darauf schließen, daß bei einem Verzicht auf Bemühungen zur Konvertierung von temporären Verweigerern und bei einer Begrenzung der Feldphase auf einen oder zwei Kontaktversuche mit größeren Verzerrungen in der Zusammensetzung der Stichprobe gerechnet werden muß. Allerdings ist auf der Basis der festgestellten Unterschiede davon auszugehen, daß es nicht notwendig ist, die Feldphase auf bis zu 10 Kontaktversuche auszudehnen. Ähnlich wie bei Befragungen der Allgemeinbevölkerung ergibt sich auch bei Telefoninterviews mit alten Menschen eine relativ günstige Kombination aus Aufwand während der Feldphase, Ausschöpfung und Zusammensetzung der Stichprobe (hinsichtlich der hier untersuchten Merkmale) bei zwischen drei und fünf Kontaktversuchen (vgl. hierzu auch ***Blasius*** und ***Reuband*** 1995, S. 84f.).

Die deutlicheren Unterschiede zwischen temporären Verweigerern und Kooperationsbereiten deuten zudem darauf hin, daß bei der Planung und Durchführung von telefonischen Befragungen alter Menschen dem Problem der Verweigerung größere Aufmerksamkeit zukommen sollte als dem Problem der Erreichbarkeit. Offenbar ist es wichtiger, sich bei telefonischen Befragungen alter Menschen um eine Konvertierung von Verweigerern zu bemühen, als die Nicht-Kontaktrate durch Ausdehnung der Feldphase auf zahlreiche Kontaktversuche zu minimieren. Dies ist um so mehr angezeigt, da der in der Literatur zur Reduzierung von Verweigerungsraten häufig empfohlene briefliche Vorkontakt (vgl. ***Porst*** 1991; ***Schnell*** 1997a, S. 259) bei der Befragung von Alten-Populationen unter Verwendung des Telefonbuchs als Auswahlgrundlage nur mit erheblichem Aufwand zu realisieren ist, weil weit über 50% der Haushalte keine Zielperson enthält. Eine briefliche Nachfaßaktion vor einer erneuten telefonischen Kontaktierung eröffnet die Möglichkeit, die Relevanz

und Nützlichkeit der Untersuchung nochmals kurz zu beschreiben, auf die Wichtigkeit jedes einzelnen Befragten für den Erfolg der Untersuchung hinzuweisen, die untersuchende Institution offiziell kenntlich zu machen und die Verweigerer möglicherweise durch materielle Anreize zu einer Teilnahme zusätzlich zu motivieren.

Literatur:

Baltes, P.B. und ***Mittelstraß, J.***(Hrsg.) (1992): Zukunft des Alterns und gesellschaftliche Entwicklung. Berlin: De Gruyter.

Blasius, J. und ***Reuband, K.-H.*** (1995): Telefoninterviews in der empirischen Sozialforschung: Ausschöpfungsquoten und Antwortqualität. ZA-Information 37, S. 64-87.

Brückner, E.; Hormuth, S. und ***Sagawe, H.*** (1982): Telefoninterviews - ein alternatives Erhebungsverfahren? Ergebnisse einer Pilotstudie. ZUMA-Nachrichten 11, S. 9-36.

Brune, M.; Werle, M. und ***Hippler, H.-J.*** (1991): Probleme bei der Befragung älterer Menschen. Methodische Erfahrungen aus einer schriftlichen Befragung zu Tätigkeitsformen im Ruhestand. ZUMA-Nachrichten 28, S. 73-91.

Bungard, W. (1979): Methodische Probleme bei der Befragung älterer Menschen. Zeitschrift für experimentelle und angewandte Psychologie 26, 2, S. 211-237.

Frey, J.H.; Kunz, G. und ***Lüschen, G.*** (1990): Telefonumfragen in der Sozialforschung. Opladen: Westdeutscher Verlag.

Gabler, S. und ***Häder, S.*** (1997): Überlegungen zu einem Stichprobendesign für Telefonumfragen in Deutschland. ZUMA-Nachrichten 41, S. 7-18.

Häder, S. (1996): Wer sind die „Nonpubs“? Zum Problem anonymer Anschlüsse bei Telefonumfragen. ZUMA-Nachrichten 39, S. 45-68.

Hartmann, P.H. (1990): Wie repräsentativ sind Bevölkerungsumfragen? Ein Vergleich des ALLBUS und des Mikrozensus. ZUMA-Nachrichten 26, S. 7-30.

Herzog, A.R. und ***Rogers, W.L.*** (1988a): Interviewing Older Adults. Public Opinion Quarterly 52, S. 84-99.

Herzog, A.R. und ***Rogers, W.L.*** (1988b): Age and Response Rates to Interview Sample Surveys. Journal of Gerontology: Social Sciences 43, S. 200-205.

Heyde, C. von der (1997): Random-Route und Telefon. Struktur von Telefonhaushalten. In: ***Gabler, S.*** und ***Hoffmeyer-Zlotnik, J.*** (Hrsg.): Stichproben in der Umfragepraxis. Opladen: Westdeutscher Verlag, S. 196-206.

Hormuth, S.E. und ***Brückner, E.*** (1985): Telefoninterviews in Sozialforschung und Sozialpsychologie. Kölner Zeitschrift für Soziologie und Sozialpsychologie 37, S. 526-545.

Jung, M. (1990): Ausschöpfungsprobleme bei repräsentativen Bevölkerungsumfragen. In: ***Lange, U.*** (Hrsg.): Internationaler Vergleich, Sprache und Telefon, Telefonseelsorge und Beratungsdienste, Telefoninterviews. (Telefon und Gesellschaft. Band 2), Berlin: Spiess, S. 386-400.

Knesebeck, O.v.d. (1998): Subjektive Gesundheit im Alter: Soziale, psychische und somatische Einflüsse. Münster: Lit.

Marhenke, W. (1997): Telefonanschlußdaten als Auswahlgrundlage. In: ***Gabler, S.*** und ***Hoffmeyer-Zlotnik, J.*** (Hrsg.): Stichproben in der Umfragepraxis. Opladen: Westdeutscher Verlag, S. 207-220.

McKinlay et al. (1990): Adapting Research Methods for Sociomedical Studies of Older People. Advances in Medical Sociology 1, S. 299-323.

Ministerium für Arbeit, Gesundheit und Soziales des Landes Nordrhein-Westfalen (Hrsg.) (1995): Bedarfsplanung in der kommunalen Altenpolitik und Altenarbeit in Nordrhein-Westfalen.
Band I: Bericht der Forschungsgesellschaft für Gerontologie e.V.

Nuthmann, R. und ***Wahl, H.-W.*** (1996): Methodische Aspekte der Erhebungen der Berliner Altersstudie. In: ***Mayer, K.U***. und ***Baltes, P.B.*** (Hrsg.): Die Berliner Altersstudie. Berlin: Akademie.

Porst, R. (1991): Ausfälle und Verweigerungen bei einer telefonischen Befragung. ZUMA-Nachrichten 29, S. 57-69.

Porst, R. (1993): Ausschöpfungen bei sozialwissenschaftlichen Umfragen. Annäherungen aus der ZUMA-Perspektive. Mannheim: ZUMA-Arbeitsbericht 93/12.

Prahl, H.-W. und ***Schroeter, K.R.*** (1996): Soziologie des Alterns. Paderborn: Schöningh.

Reuband, K.-H. und ***Blasius, J.*** (1995): Verweigerung im Interview als stabiles oder instabiles Handlungsmuster? Ein Vergleich von face-to-face und telefonischen Befragungen. Unveröffentlichtes Manuskript, Dresden und Köln.

Reuband, K.-H. und ***Blasius, J.*** (1996): Face-to-Face-, telefonische und postalische Befragungen. Kölner Zeitschrift für Soziologie und Sozialpsychologie 48, S. 296-318.

Salaske, I. (1997): Die Befragbarkeit von Bewohnern stationärer Alteneinrichtungen unter besonderer Berücksichtigung des Verweigerungsverhaltens. Kölner Zeitschrift für Soziologie und Sozialpsychologie 49, S. 291-305.

Salmon, C.T. und ***Nichols, J.S.*** (1983): The Next-Birthday Method of Respondent Selection. Public Opinion Quarterly 47, S. 270-276.

Schnell, R. (1991): Wer ist das Volk? Zur faktischen Grundgesamtheit bei „allgemeinen Bevölkerungsumfragen“: Undercoverage, Schwererreichbare und Nichtbefragbare. Kölner Zeitschrift für Soziologie und Sozialpsychologie 43, S. 106-137.

Schnell, R. (1997a): Nonresponse in Bevölkerungsumfragen. Ausmaß, Entwicklung und Ursachen. Opladen: Leske+Budrich.

Schnell, R. (1997b): Praktische Ziehung von Zufallsstichproben für Telefon-Surveys. ZA-Information 40, S. 45-59.

Schulte, W. (1997): Telefon- und Face-to-Face-Umfragen und ihre Stichproben. Allgemeine Bevölkerungsumfragen in Deutschland. In: ***Gabler, S.*** und ***Hoffmeyer-Zlotnik, J.*** (Hrsg.): Stichproben in der Umfragepraxis. Opladen: Westdeutscher Verlag, S. 148-195.

Smith, T.W. (1983): The Hidden 25 Percent: An Analysis of Nonresponse on the 1980 General Social Survey. Public Opinion Quarterly 47, S. 386-404.

Tennstedt, S.L.; Dettling, U. und ***McKinlay, J.B.*** (1992): Refusal Rates in a Longitudinal Study of Older People: Implications for Field Methods. Journal of Gerontology: Social Sciences 47, S. 313-318.

Wiseman, F. und ***McDonald, P.*** (1979): Noncontact and Refusal Rates in Consumer Telephone Surveys. Journal of Marketing Research 16, S. 478-484.

Zeh, J. (1987): Stichprobenbildung bei Telefonumfragen. Angewandte Sozialforschung 14, 4, S. 337-347.

Quoten- und Randomstichproben in der Praxis der Sozialforschung. Gemeinsamkeiten und Unterschiede in der sozialen Zusammensetzung und den Antwortmustern der Befragten.

von Karl-Heinz Reuband [1]

Zusammenfassung

Untersucht wird auf der Basis einer postalischen bzw. schriftlichen Umfrage in der wahlberechtigten Bevölkerung der Stadt Dresden, wie sich Random- und Quotenstichproben auf der Ebene der sozialen Zusammensetzung und der Antworten unterscheiden und welche Auswahlstrategien von den Interviewern gewählt werden. Als Interviewer in der Quotenstichprobe wurden Studenten eingesetzt. In den sozialen Merkmalen sowie den Einstellungen und Verhaltensweisen ergeben sich in der Regel große Ähnlichkeiten zwischen den Stichprobenverfahren. Am ehesten treten Unterschiede im Grad sozialer Aktivität auf: Befragte in Quotenstichproben sind sozial aktiver. Nicht die zu Hause leicht Erreichbaren, sondern die sozial Aktiven - die aufgrund ihrer Aktivität in sozialen Netzwerken überproportional sichtbar und zugänglich sind - werden im Vergleich zu Randomstichproben häufiger in Quotenstichproben erfaßt.

Abstract

On the basis of a mail respectively questionnaire survey among the adult population in Dresden a comparison is done between random- and quota samples and a study of the selection processes. Students are used as interviewers in the quota sample. With regard to the social characteristics, attitudes and behavior great similarities emerge. Where differences exist they affect above all the degree of social activity: respondents in quota samples are socially more active. Not the easy to reach, who stay at home, have a greater chance of being recruited into quota samples but the socially active who are visible and accessible in their social networks.

---

1 Dr. ***Karl-Heinz Reuband*** ist Professor am Sozialwissenschaftlichen Institut der Heinrich-Heine-Universität, 40225 Düsseldorf.

1 Quotenstichproben in der Praxis der Sozialforschung

Quotenstichproben stellen in der Bundesrepublik Deutschland in der Markt- und Meinungsforschung ein häufig eingesetztes Auswahlverfahren auch bei der Durchführung groß angelegter Bevölkerungsumfragen dar. Einige der größten und renommiertesten Institute, wie das Institut für Demoskopie oder die GFK, verwenden sie bis heute mehrheitlich oder nahezu ausschließlich. Andere Institute setzten sie in der Vergangenheit lange Zeit ein. Inzwischen ist die Häufigkeit der Nutzung zwar seltener geworden, aber immer noch nennenswert. So wurde in einer Befragung deutscher Markt- und Meinungsforschungsinstitute Ende der 80er Jahre von 92% der Institute angegeben, sie würden - neben anderen Verfahren - auch Quotenstichproben verwenden. Kein anderes Auswahlverfahren wurde derart häufig genannt (***Bausch*** 1990, S. 71).

Die Situation in Deutschland ist in dieser Hinsicht keineswegs einzigartig. Auch in anderen europäischen Ländern - wie Großbritannien und Frankreich - werden Quotenstichproben vor allem in der Markt- und Meinungsforschung häufig verwandt (vgl. ***Taylor*** 1995). In Großbritannien hatten sie gar lange Zeit den Ruf, bei Wahlprognosen den Randomsamples überlegen zu sein. Und in den USA waren Quotenstichproben bis Anfang der 70er Jahre selbst im akademisch hochgeschätzten General Social Survey üblich, und zwar in Form eines "Block-Quota-Samples“, bei dem der Wohnbezirk per Randomstichprobe vorgegeben wird und innerhalb des Bezirks nach Quoten ausgewählt wird (vgl. ***Sudman*** 1966, ***Crespi*** 1988, S. 27).

Auch in der Bundesrepublik wurden Quotenstichproben jüngst noch eingesetzt in Studien, die von akademischen Institutionen oder statistischen Ämtern (bzw. in deren Auftrag) durchgeführt wurden und die einem besonders hohen wissenschaftlichen Standard genügen sollen. Dazu zählen Erhebungen unter alten Menschen (***Bertram*** 1995, S. 431) ebenso wie Jugendstudien (z.B. ***Schmidtchen*** 1997) oder Zeit-Budget-Untersuchungen unter den Bundesbürgern (***Blanke et al***. 1996). Schließlich stützen sich auch grundlegende, periodisch wiederkehrende Studien der statistischen Ämter - wie die Einkommens- und Verbraucherstichprobe - auf Quotensamples (vgl. ***Pöschl*** 1993).

Geschätzt werden Quotenstichproben von Praktikern vor allem aufgrund ihrer im Vergleich zu Randomsamples geringeren Kosten sowie kürzerer Feldlaufzeiten (vgl. ***Moser*** und ***Kalton*** 1971, S. 134, ***Jowell et al***. 1993, S. 240). Dies ist jedoch keine zwingende Begründung. Feldlaufzeiten lassen sich heute mit Telefonbefragungen erheblich reduzieren[2].

2 Die Befragungen werden per Random Last Digit Dialing häufig nach 1-3 Tagen, nicht selten sogar nach einem Tag abgeschlossen. Bei face-to-face Befragungen andererseits wird allein durch den Versand der Befragungsunterlagen bereits viel Zeit verloren. In der Frage der Vergleichbarkeit der Daten dürften telefonische und face-to-face Befragungen sich kaum unterscheiden, wohl aber telefonische von postalischen bzw.

Und die Kosten für Quotenstichproben sind - zumal in der Bundesrepublik - nicht selten denen für Randomsamples gleichgestellt, wenn sie diese nicht gar überschreiten (wie ein Vergleich der Preise, etwa der GfK mit anderen Instituten belegt, vgl. Planung und Analyse Nr. 5, 1998, S. 68). Viele Autoren meinen, Quotenstichproben seien nur als Notlösung einsetzbar, wenn die Kosten für Randomverfahren besonders hoch seien (***Bortz*** 1984, S. 362) - z.B. wenn die relevanten Zielpersonen nur über ein aufwendiges und kostenintensives "screening" im Rahmen von face-to-face Kontakten lokalisiert werden können, oder, wenn ein hohes Maß an Motivation für die Teilnahme erforderlich sei, z.B. bei Konsumentenpanels oder bei der Führung von Haushaltsbüchern (vgl. ***Pöschl*** 1993). Für andere Autoren, wie ***Elisabeth Noelle-Neumann***, sind Quotenverfahren jedoch kein Notbehelf, der über günstigere Feldzeiten oder Preise zu rechtfertigen ist. Quotenverfahren sind vielmehr den Randomverfahren ebenbürtig (wenn sie diesen nicht sogar überlegen sind, vgl. ***Noelle-Neumann*** und ***Petersen*** 1996).

Wie häufig Quotenstichproben in die Forschung Eingang finden, die im universitären Kontext erfolgt, ist unbekannt. Vor allem in Forschungspraktika im Rahmen der Lehre dürften sie relativ häufig sein: der in der Regel große Aufwand einer systematischen Zufallsstichprobe und Ziehung über das Einwohnermelderegister oder per Random-Route-Verfahren entfällt, die Studenten können gleichwohl Erfahrungen in der Sozialforschung sammeln. In den meisten Fällen dürften diese Projekte - abgesehen von universitätsinternen Berichten - freilich zu keinen Veröffentlichungen führen und daher relativ „unsichtbar“ bleiben. Neben der Tatsache, daß es sich um Lehrforschungsprojekte handelt, in denen häufig die erhobenen Daten nur partiell und unvollständig ausgewertet werden, dürfte dafür verantwortlich sein, daß das Quotenverfahren in der akademischen Profession als minderwertig, wenn nicht gar als unwissenschaftlich gilt (vgl. u.a. ***Diekmann*** 1995, S. 338 ff., ***Sahner*** 1995, S. 301, ***Schnell*** 1997). Damit ist eine Situation gegeben, wo ein Stichprobenverfahren in der Praxis einen durchaus bedeutenden Stellenwert hat, es an der akademischen Dignität jedoch fehlt und systematische Bemühungen einer Abschätzung von Vor- und Nachteilen gar nicht erst unternommen werden.

2 Methodische Probleme und offene Fragen

Die Argumente, die üblicherweise gegen den Einsatz von Quotenstichproben vorgebracht werden, sind zahlreich: Informationen über die Grundgesamtheit, die erforderlich sind, liegen häufig nicht vor oder sind nicht aktuell. Nur leicht erkennbare demographische Merkmale können bei den Quotenvorgaben verwendet werden. Und da es sich nicht um eine Zufallsstichprobe handelt, können Stichprobenfehler nicht berechnet werden, die

schriftlichen Befragungen (vgl. ***Reuband*** und ***Blasius*** 1996, ***Reuband***, im Druck, ***Noelle-Neumann*** und ***Petersen***, im Druck).

Ausfallquote bleibt unbekannt.[3] Noch gewichtiger ist, daß mit der Kontrolle einiger ausgewählter Quotenmerkmale noch nichts über die Repräsentanz der anderen sozialen Merkmale sowie der Einstellungen und Verhaltensweisen ausgesagt wird. Einstellungen und Verhaltensweisen werden durch sozialstrukturelle Merkmale nur bedingt prognostiziert (***Schnell*** 1993). Und der Einfluß der sozialstrukturellen Merkmale auf Einstellungen und Verhalten nimmt zudem möglicherweise längerfristig ab (***Schnell*** und ***Kohler*** 1995, vgl. jedoch ***Müller*** 1998). Sollten Quotenstichproben jemals eine gewisse Berechtigung gehabt haben, dann - so scheint es - in der Zukunft immer weniger.

Hinzu kommt, daß es weitgehend dem Interviewer überlassen wird, die Zielpersonen auszusuchen. Sie dürften, so die Vermutung der Kritiker, den Weg des geringsten Aufwandes wählen - so etwa nur in den unteren Stockwerken (z.B. ***Bortz*** 1984, S. 362) - oder bevorzugt Personen in ihrem eigenen Verwandten und Bekanntenkreis befragen. Womöglich gar erstellen sie sich für zukünftige Befragungen eine Kartei von Personen, die sie schon mal befragt haben, über deren soziale Merkmale sie genauer informiert sind und die daher für sie die Basis für zukünftige Auswahlen darstellen kann. Die Kommunikation im Interview wird dadurch eine andere als sie es von der Logik des Interviews her sein sollte. Danach ist es von zentraler Bedeutung für die Qualität der Angaben, daß die Angaben im Interview für den Befragten keine Konsequenzen nach sich ziehen. Und dies ist nur dann gewährleistet, wenn sich Befragte und Interviewer nicht kennen und die Beziehung über die Situation des Interviews hinweg nicht fortgesetzt wird (***Scheuch*** 1973).

Das Argument, die Interviewer würden den Weg des geringsten Aufwandes wählen und - sofern sie von Haus zu Haus gehen - in den unteren Stockwerken bevorzugt befragen, hat sich zumindest in Umfragen des Instituts für Demoskopie als fraglich erwiesen (vgl. ***Noelle-Neumann*** und ***Petersen*** 1996, S. 278). Was das Argument eines Befragten-Stamms angeht, so muß man hier nach Instituten differenzieren. Die Beschränkung auf einen Befragten-Stamm wird nicht von allen Instituten akzeptiert. In den Interviewanweisungen des Instituts für Demoskopie wird geraten, "möglichst keine Ihnen persönlich näher bekannte Person" zu befragen, niemals zwei Personen aus der gleichen Familie und niemanden öfter als einmal im halben Jahr zu befragen. Dies schließt gleichwohl die Empfehlung nicht aus, sich einen Vorrat an "befragungswilligen" Personen zuzulegen, um bei Eilaufträgen gerüstet zu sein. Als befragungswillige Personen gelten jene, bei denen der Termin nach den

3 Das Institut für Demoskopie geht von Stichprobenfehlern wie bei Randomstichproben aus (vgl. ***Noelle-Neumann*** und ***Köcher*** 1993 S. E 32). ***Norval Glenn*** (1977 S. 141) schlägt vor, in Quotenstichproben annähernd die doppelte Fehlerquote zur Ausgangsbasis zu erheben. Was die Ausfälle bei Quotenstichproben angeht, so existieren sie, nur werden sie nicht dokumentiert. Die Aussage, daß es bei Quotenverfahren keine Ausfälle gebe (so ***Pöschl*** 1993 S. 385), ist selbstverständlich falsch. Auch wenn der Interviewer von vornherein Personen auswählt, von denen er meint, sie würden teilnehmen (z.B. im Freundes- und Bekanntenkreis), gibt es Ausfälle durch Verweigerung und Nichterreichen der Zielpersonen. Und selbst wenn es sie nicht geben würde: es bleibt die Vorselektion durch den Interviewer, die notwendigerweise einen breiten Kreis potentieller Interviewpartner ausschließt.

ersten Kontakten nicht zustande kam oder die von Befragten als mögliche Kontaktpersonen genannt wurden (Institut für Demoskopie o.J.).

In ähnlicher Weise schreibt Infratest (welches das Quotenverfahren in rund 25% aller Interviews einsetzt, vgl. ***Bliesch*** 1998, S. 69) in seinen Interviewanweisungen, daß es "empfehlenswert" sei, sich "ohne viel Aufwand auf Quotenuntersuchungen vorzubereiten. Sie können im Kreis Ihrer Verwandten, Bekannten, am Arbeitsplatz oder sonstwo in zwanglosen Gesprächen feststellen, wer vielleicht für spätere Quotenuntersuchungen vorgemerkt werden kann. Das gleiche gilt für fast jedes Interview, das Sie durchführen". Ähnlich wie beim Institut für Demoskopie Allensbach, gibt es eine Beschränkung in der Zahl der hintereinander durchgeführten Interviews. Es dürfen keine Personen innerhalb eines Vierteljahres mehr als einmal interviewt werden (Infratest Burke o.J.).

Andererseits gibt es für andere Institute und Befragungen durchaus Hinweise dafür, daß die Interviewer in erheblichem Maße auf einen eigenen Befragten-Stamm zurückgreifen. So hat ***Wendy Hoag*** am Beispiel einer Shell Jugendstudie dokumentiert, die auf einer Quotenstichprobe basiert und bei der jugendliche Interviewer eingesetzt wurden, daß die meisten der Befragten den Interviewer vorher kannten. Durch diesen Rückgriff auf den Freundes- und Bekanntenkreis stellte sich ein höherer Grad politischer Homogenität ein als er unter den Jüngeren realiter vorhanden sein kann (***Hoag*** 1986). Und in einer GfK Studie auf der Basis einer Quotenstichprobe bekundete eine Mehrheit von über drei Viertel der Befragten, sie seien früher schon mal befragt worden (GfK 1991, S. 15). Gemessen am Anteil in Randomstichproben (Context 1991, S. 14, Forsa 1996), ist dieser Wert zweifelsohne zu hoch und kann nur über die Tatsache erklärt werden, daß die Interviewer auf ihren Befragten-Stamm zurückgriffen.

Doch nicht nur im Hinblick auf die Art der bereits bestehenden Beziehung zur Zielperson können spezifische Effekte aus der Quotenstichprobe erwachsen. Es genügt, daß Personen mit ähnlichen sozialen Merkmalen, wie sie der Interviewer aufweist, kontaktiert werden. Diese Gefahr ist immer dann besonders groß, wenn sich Interviewer ihre Zielpersonen aus dem Kreis derer aussuchen, mit denen sie häufiger Kontakt haben. Der Grund dafür liegt in der Ausdifferenzierung der sozialen Verkehrskreise gemäß sozialen Lagen (vgl. ***Schneider*** 1969, ***Reuband*** 1974). Ist der Interviewerstab - wie allgemein üblich (***Reuband*** 1984) - durch eine Überrepräsentation besser Gebildeter gekennzeichnet und die Bildung nicht als Quotenmerkmal eingeführt, könnte es entsprechend dazu kommen, daß besser Gebildete überproportional als Befragte ausgewählt werden. Die im Vergleich zur Randomstichprobe starke Repräsentanz besser Gebildeter in Quotenstichproben der Eurobarometer-Serie könnte ein Beleg dafür sein (vgl. ***Böltken*** und ***Gehring*** 1984, S. 46).

Anderseits gilt, daß Quotenstichprobe nicht immer gleich Quotenstichprobe ist. Es gibt, wie vor allem Elisabeth ***Noelle-Neumann*** betont hat (1995), eine ganze Reihe unter-

schiedlicher Möglichkeiten, eine Quotenstichprobe zu konstruieren. Je engmaschiger die Quotenvorgaben sind - je größer die Zahl der Quotenmerkmale -, desto geringer sei der Ermessensspielraum der Interviewer und desto größer die Wahrscheinlichkeit, daß sich dieses Verfahren einem Zufallsverfahren annähert. Wenn eine gewisse Fluktuation im Interviewerstab hinzukommt - was normalerweise von den Instituten nicht erstrebt wird, aber die Regel ist -, dürfte auch die Wahrscheinlichkeit reduziert werden, daß bei den Befragungen auf den gleichen, festen Stamm von Befragten zurückgegriffen wird. Umfragen des Instituts für Demoskopie (***Noelle-Neumann*** und ***Köcher*** 1993, S. 1112) erbringen so denn auch eine niedrigere Quote an Interviewerfahrungen als die zuvor zitierte GfK-Studie.

In der Praxis, so behauptet ***Noelle-Neumann***, dürfte es keinen grundlegenden Unterschied für den Prozeß der Auswahl machen, ob ein differenziertes Quotenverfahren oder ein Randomverfahren eingesetzt wird (***Noelle-Neumann*** 1995, S. 279, ***Noelle-Neumann*** und ***Petersen*** 1996 S. 208). Durch die spezifische Vorgabe sozialer Merkmale würden die Interviewer gezwungen, aus dem Kreis der Bekannten und Verwandten herauszutreten. Aus dieser Sicht hätten die jeweils vorgegebenen sozialen Merkmale die primäre Funktion, die Selektivität der Auswahl nach Gutdünken seitens der Interviewer zu reduzieren. Die Frage der Korrelation zwischen sozialen Merkmalen und unterschiedlichem Verhalten bzw. Einstellungen wird von ihr demgegenüber als eher peripher behandelt. Damit wird eine etwas andere Position vertreten als bei den Autoren, die meinen, ein Quotensample würde nur so gut funktionieren wie zwischen den jeweils abhängigen Variablen und den vorgegebenen Quotenmerkmalen enge Korrelationen bestehen (***Jowell et al***. 1993, S. 252, ***Schnell*** 1993, ***Kromrey*** 1998, S. 270).[4]

Gemessen an theoretischen und statistischen Maßstäben stellt die quotengesteuerte Erhebung zweifellos das schlechtere Verfahren dar. Aber ob es auch in der Praxis schlechter ist, bleibt eine andere Frage und kann nur empirisch geklärt werden. Die Zahl der empirischen Arbeiten, die einen Vergleich zwischen beiden Verfahren erlauben, ist für die Bundesrepublik Deutschland ähnlich wie in anderen Ländern (vgl. ***Lynn*** und ***Jowell*** 1996, S. 26) trotz der weiten Verbreitung von Quotenstichproben bemerkenswert gering. Es gibt nur wenige, in der Literatur kaum beachtete, Vergleiche auf der Basis von Primärdaten (***Schmidtchen*** 1962, Institut für Demoskopie 1975, ***Radtke*** und ***Zeh*** 1974, Institut für Demoskopie 1979), publizierter Daten (***Koolwijk*** 1974) sowie Sekundäranalysen (***Birke*** 1998). Daneben finden sich vereinzelt Arbeiten, in denen Untersuchungsbefunde auf der Basis unterschiedlicher Stichprobenverfahren - in der Regel im Kontext eines Vergleichs einzelner thematisch

4 Daß letztere Position nicht ganz unproblematisch ist, läßt sich daran zeigen, daß je nach Einstellungs- und Verhaltensbereich die soziodemographischen Methoden in ihrer Erklärungskraft variieren. Zu Ende gedacht, müßte die Konzeption von engen Zusammenhängen als Voraussetzung der Brauchbarkeit bedeuten, je nach Thema unterschiedliche Quotenkonstellationen vorzugeben (was angesichts der thematischen Heterogenität vieler Studien wenig Sinn macht).

verwandter Fragen - sekundäranalytisch miteinander in Beziehung gesetzt werden (vgl. ***Böltken*** und ***Gehring*** 1984, ***Reuband*** 1992a, 1995). Wenn in diesen Studien Randomstichproben mit Quotenstichproben verglichen werden, handelt es sich in der Regel um Random-Route-Verfahren. Diese weichen in der Reproduktion der sozialen Zusammensetzung der Befragten von Randomstichproben auf der Basis von Einwohnermeldeamtsverzeichnissen ab (***Alt et al***. 1991).

Als wichtigstes Ergebnis der Vergleiche zeigt sich: es finden sich in der Regel keine bedeutsamen Unterschiede auf der Ebene der Randverteilungen zwischen beiden Verfahren (inwieweit dies auch auf der Ebene von Subgruppen gilt, ist weniger untersucht). Die Unterschiede in den soziodemographischen Merkmalen zwischen Random- und Quotenstichproben sind in der Regel gering. Differenziert man stärker nach den sozialen Kategorien als von der Quotierung her vorgegeben, so scheint nach einigen Studien die Randomstichprobe eher der amtlichen Statistik angenähert zu sein als die Quotenstichprobe (***Radtke*** und ***Zeh*** 1974, S. 129, genauere empirische Befunde allerdings werden nicht mitgeteilt). Bei anderen Studien sind die Ergebnisse gemischt. Manche Autoren meinen so denn mit einer gewissen Verwunderung, nach empirischer Prüfung der beiden Stichproben, die Quotenverfahren seien besser und die Random-Route-Verfahren schlechter als vielfach angenommen (so z.B. ***Birke*** 1998, S. 91).

Trotz der überwiegenden Gemeinsamkeiten zwischen den Stichprobenarten sind Unterschiede in der Zusammensetzung der Befragten nicht zu übersehen. Sie betreffen vor allem die sozialen Aktivitätsmuster: Personen mit vielen Freunden und sozialen Aktivitäten sind in Quotenstichproben stärker vertreten (***Schmidtchen*** 1962, ***Koolwijk*** 1974). Ob dies bedeutet, daß Personen mit vielen Kontakten für Befragungen eher aufgeschlossen sind und daher in Quotenstichproben überrepräsentiert sind, oder ob es bedeutet, daß umgekehrt Personen mit vielen Kontakten schwerer erreichbar sind und in Randomstichproben unterrepräsentiert sind, muß als Frage offen bleiben.

Jedenfalls wäre es verfehlt anzunehmen, daß sich die Befragten in Quotenstichproben überproportional aus Personen zusammensetzen, die allein deswegen leicht erreichbar sind, weil sie oft zu Hause sind. Untersuchungen, in denen Quotenstichproben simuliert werden, indem die leicht Erreichbaren in Randomstichproben zum Maßstab genommen werden (z.B. ***Jowell et al.*** 1993, ***Lynn*** 1994, ***Schnell*** 1998), gehen aus diesem Grunde fehl. Die leichte Erreichbarkeit der Person in Quotenstichproben ergibt sich eben nicht aus ihrer Häuslichkeit, sondern aus ihrer *Sichtbarkeit* in sozialen Netzwerken. Je mehr Freunde sie haben, je mehr sie sich außer Haus aufhalten, desto größer die Chance, von einem Interviewer kontaktiert zu werden. Aus dieser Perspektive dürften die Befragten in Quotensamples eher den mobilen Nicht-Erreichbaren ähneln als den Verweigerern, die sich normalerweise gegenüber den Befragten durch weniger Freunde und Bekannte auszeichnen

(vgl. ***Esser*** 1979). Untersuchungen, in denen die Ausfälle in Randomsamples durch Quotensamples kompensiert werden (vgl. ***Scheuch*** 1972, S. 81ff), sollten sich daher bei der Kompensation primär auf die Kategorie der Nicht-Erreichten - und nicht der Verweigerer - beziehen. Letztere stellen eine andere Kategorie von Personen dar.

Nur gelegentlich ergeben sich beim Vergleich von Random- und Quotenstichproben Unterschiede auch über das soziale Aktionsmuster hinaus. Dies ist offenbar dann der Fall, wenn zwischen den interessierenden Variablen und dem sozialen Aktionsmuster ein enger Zusammenhang besteht. So finden sich in einigen Studien höhere Prävalenzwerte für Drogengebrauch in Quotenstichproben als in Randomstichproben. Nähere Analysen belegen, daß die Unterschiede überall dort gehäuft auftreten, wo die Rekrutierung der Befragten primär oder ausschließlich im öffentlichen Raum ansetzt, etwa auf der Straße (gleichgültig, ob auf der Straße dann auch das Interview geführt wird oder die Zielperson in ein Lokal oder Café gebeten wird). Bei einem solchen Vorgehen werden sozial aktive Personen und Personen, die häufig außer Hause sind, bevorzugt in die Stichprobe einbezogen. Drogenkonsumenten weisen diese Eigenschaften überproportional auf. Untersuchungen auf der Basis von Quotenstichproben, die nicht am öffentlichen Raum ansetzen, unterscheiden sich nicht nennenswert von Untersuchungen auf der Basis von Randomstichproben in der Drogenprävalenz (***Reuband*** 1992, 1995, S. 29).

Ein eindeutiges empirisches Urteil über die Qualität der Stichprobenverfahren zu fällen, fällt schwer, weil es externe Kriterien - mit Ausnahme der sozialen Merkmale, die mit dem Mikrozensus oder der Volkszählung verglichen werden können - nicht gibt. Man hat in der Vergangenheit bevorzugt Wahlprognosen als Maßstab genommen und deren Güte am tatsächlichen Wahlausgang bewertet. Wahlprognosen aber sind ein schlechter Maßstab, weil je nach Institut unterschiedliche Gewichtungs- und Schätzverfahren gewählt werden und "last minute swings" das Ergebnis maßgeblich mitbestimmen können. Die große Popularität von Wahlprognosen als Kriterien für die Bewertung von Stichproben ist letztlich wohl nur zu verstehen, weil sie - wie keine anderen Einstellungs- und Verhaltensfragen - an Außenkriterien validiert werden können und weil es in der Vergangenheit jeweils aufgrund fehlgeschlagener Wahlprognosen zu einem Wechsel der Stichprobenverfahren bei den Markt- und Meinungsforschungsinstituten kam. Wahlprognosen sind zu einem zentralen Symbol für die Güte von Stichprobenverfahren und Institutsqualität avanciert (vgl. ***Noelle-Neumann*** 1998, S. 11).

Gemessen an der Genauigkeit der Wahlprognosen sind die Studien, die sich in Deutschland auf Quotenstichproben stützen, nicht schlechter als Studien auf der Basis von Randomstichproben (***Noelle-Neumann*** und ***Petersen*** 1996). Und im Rahmen der Umfragen zur Bundestagswahl 1998 war es das Institut für Demoskopie, das auf der Basis von Quotenstichproben unter verschiedenen Instituten sogar die beste Prognose abgab. Der Befund,

daß die Stichprobenart für die Güte der Wahlprognose nicht zählt, deckt sich mit den Befunden amerikanischer Studien (***Crespi*** 1988, S. 29), was um so bedeutsamer ist, als dort die Abkehr von der Quotenstichprobe stattfand, nachdem die Wahlprognosen in der Wahl von 1948 versagt hatten. In dieser Zeit waren Quotenstichproben verbreitet und ihr "Versagen" wurde auf Eigenheiten des Quotenverfahrens zurückgeführt (***Crespi*** 1988, S. 26).

In den 90er Jahren wiederholte sich die Diskussion um die Stichprobenauswahl bei Wahlprognosen in Großbritannien. Sie setzten ein, als der Eindruck entstand, als hätte vor allem der Rückgriff auf Quotenstichproben zum Versagen der Umfrageinstitute bei der Wahlprognose im Jahr 1992 geführt. Doch es gibt alternative Erklärungen. Diese reichen von unzureichender Quotierung über zu kurzfristige Veränderungen in der Parteipräferenz vor der Wahl bis hin zu Auswirkungen der Schweigespirale (vgl. u.a. ***Jowell et al.*** 1993, ***Moon*** 1994). Sicher ist: Wahlprognosen auf der Basis von Quotenstichproben - vor und nach der Fehlprognose von 1992 - haben sich auch in Großbritannien im allgemeinen bewährt. Von einem generellen Versagen der Quotenstichprobe bei Wahlprognosen kann nicht die Rede sein.

3 Zielsetzung und methodisches Vorgehen

Im folgenden soll im Rahmen einer eigenen Untersuchung geprüft werden, wie sehr sich Random- und Quotenstichproben in der sozialen Zusammensetzung und den Antwortmustern unterscheiden. Darüber hinaus soll geklärt werden, in welcher Weise die Interviewer in Quotenstichproben gemäß den Vorgaben ihre Auswahl treffen und welche Folgen daraus für die Zusammensetzung der Stichprobe erwachsen.

Unsere Untersuchung soll weniger zur Klärung der Frage dienen, ob die Quotenstichprobe in zukünftigen Erhebungen eine legitime Option darstellt, als zur Bestimmung ihrer prinzipiellen Problematik. Je geringer die Unterschiede zwischen Random- und Quotenstichproben sind, desto weniger ist es sinnvoll, Quotenverfahren von vornherein als unbrauchbar einzustufen. Dann erscheint es durchaus als legitim, Umfragen auf der Basis von Quotenstichproben - wie sie etwa in Umfragearchiven in größerem Umfang verfügbar sind - einer Sekundäranalyse zu unterziehen und z.B. zu Zwecken des Zeitvergleiches mit Umfragen auf der Basis von Randomstichproben in Beziehung zu setzen. Auch wenn es an sich sinnvoll sein mag, sich bei Vergleichen auf Studien mit gleichartiger Stichprobenziehung zu stützen (***Reuband*** 1980, S. 146), kommt man bei der Beschreibung und Analyse des Langzeitwandels mitunter nicht umhin, Studien auf der Basis unterschiedlicher Stichprobenverfahren miteinander in Beziehung zu setzen (vgl. z.B. ***Reuband*** 1992b, ***Meulemann*** 1995).

3.1 Methodisches Vorgehen

Die folgende Untersuchung basiert auf einer schriftlichen Befragung der deutschen Wohnbevölkerung der Stadt Dresden ab 18 Jahren. In der Wahl einer schriftlichen statt face-to-face Befragung unterscheidet sie sich von Untersuchungen, in denen normalerweise Quotenstichproben zur Anwendung kommen. Für die Klärung der Effekte unterschiedlicher Auswahlverfahren ist dies jedoch ohne Belang. Solange die Erhebungsverfahren konstant gehalten werden, können die Auswahleffekte bestimmt werden.

Die Randomstichprobe basiert - anders als die sonst üblichen Varianten auf der Grundlage von Random-Route-Verfahren - auf einer Zufallsstichprobe von 1500 wahlberechtigten Personen aus dem Einwohnermelderegister der Stadt. Die Zielpersonen wurden in Anlehnung an die Total Design Method von ***Don Dillman*** (1978) im Rahmen einer postalischen Erhebung angeschrieben und um die Beantwortung des beigelegten 11seitigen Fragebogens, dessen Beantwortungsdauer bei ca. 30-40 Minuten lag, gebeten. Nach bis zu drei Mahnungen wurde - bezogen auf die bereinigte Bruttostichprobe - eine Ausschöpfungsquote von 60% erreicht (vgl. ***Reuband*** 1998b). Dies entspricht für Großstädte einer überproportional guten Quote, höher als sie etwa in den neuen ALLBUS-Erhebungen in Großstädten vergleichbarer Größenordnung zu finden ist (vgl. ***Koch*** 1997). Die Zahl der Befragten liegt bei 887 Personen.

Die Quotenstichprobe, in welcher der gleiche Fragebogen verwandt wurde und die sich auf die gleiche Grundpopulation bezog, wurde unter Mithilfe studentischer Interviewer realisiert. Diese hatten nach vorgegebenen Quotenmerkmalen die Zielpersonen zu suchen und zur Befragung zu motivieren. Wenn immer dies möglich war, sollte der Fragebogen vom Befragten schriftlich beantwortet werden. Nur dort, wo dies nicht praktikabel war - z.B. bei Personen, die auf der Straße befragt wurden -, konnten die Fragen des Fragebogens, dann bei Vorgabe eines Minimalsatzes relevanter Fragen, auch in gekürzter Fassung mündlich vorgelesen werden. Dies kam in einigen wenigen Fällen vor.

Wie häufig die Interviewer beim schriftlichen Ausfüllen des Fragebogens selbst zugegen waren, wissen wir nicht. In dem Aufgabenblatt wurde nur global vermerkt, daß der Fragebogen möglichst schriftlich eingesetzt werden sollte: „Eine Möglichkeit ist hier, Sie geben ihn ab und holen ihn zu einem späteren Zeitpunkt wieder ab.“[5] Die Feldberichte deuten darauf hin, daß die empfohlene Variante bei einem Großteil der Interviews vorkam. Unter

5 Die Praxis, dem Befragten den Fragebogen über einen Interviewer zu überbringen und dann wieder abholen zu lassen, ist eine, die in westlichen Industriegesellschaften in der Vergangenheit unüblich war. Dort ist eine Zustellung per Post die Praxis. In den Ostblockländern hingegen war und ist dieses andere Vorgehen noch verbreitet. In der Bundesrepublik gibt es partielle Analogien im "Drop-Off" Verfahren (wie es im ISSP eingesetzt wird): hier hinterläßt der Interviewer den Fragebogen zur schriftlichen Beantwortung nach Durchführung eines mündlichen Interviews. Er holt ihn dann wieder ab oder der Befragte sendet ihn an das durchführende Institut.

diesen Bedingungen wäre die Situation der Beantwortung mit der in der postalischen Erhebung - Beantwortung ohne Anwesenheit eines Interviewers - vergleichbar. Befragt wurden insgesamt 592 Personen.[6]

Die Feldzeit war weitgehend mit der der Randomstichprobe identisch: die Haupterhebungsphase erstreckt sich auf die Zeit von Februar bis Anfang April 1995. Bei den "Interviewern", die im Rahmen der Quotenstichprobe eingesetzt wurden, handelt es sich um Studenten der Sozialwissenschaften (Soziologie, Politologie, Kommunikationswissenschaften und z.T. auch Erziehungswissenschaften) im ersten und zweiten Semester (davon waren 58% weiblich) im Rahmen einer Pflichtveranstaltung zu den Methoden der empirischen Sozialforschung.[7] Jeder der Teilnehmer hatte fünf Interviews nach individuellen Quotenvorgaben zu erfüllen. Dabei wurde der bei Quotenstichproben üblichen (wenn auch nicht erwünschten) Praxis gemäß den Interviewern die Möglichkeit eingeräumt, nicht nur Fremde, sondern auch Personen im Verwandten- und Bekanntenkreis zu befragen. Im Falle einer Gruppenarbeit wurde erlaubt, die Quotenmerkmale, die man selbst schlecht, aber andere Personen in der Gruppe besser einlösen konnten, zu "tauschen". Die Befragten sollten nicht gleichzeitig Bestandteil der zeitgleich durchgeführten postalischen Erhebung sein.

Die Quoten wurden auf der Basis von veröffentlichten Daten der amtlichen Statistik für die Stadt Dresden errechnet. Alter, Geschlecht und Bildung wurden als Merkmale vorgegeben. Sie wurden als getrennte und nicht als kombinierte Quoten aufgeführt. So wurde angegeben, wieviele der zu Befragenden männlich oder weiblich und wieviele in welcher Alterskategorie sein sollten. Auf eine Kombination dieser Quoten wurde verzichtet, um den Schwierigkeitsgrad nicht allzu groß werden zu lassen (was die Motivation der Teilnehmer reduziert und Fälschungstendenzen erhöht hätte). Es wurde ferner darauf verzichtet, weil geklärt werden sollte, wie sehr sich mit derartigen einfachen Quotenvorgaben - die in vie-

---

6 Einige wenige Interviewer führten, weil sie ihren Quotenvorgaben nicht entsprachen, zusätzliche Interviews durch. Hinzu kommt, daß einige der Interviews nicht in Dresden erhoben wurden. Sie werden hier aus der Analyse ausgelassen. Daher weicht die Zahl der Interviews von dem Erwartungswert (durch 5 teilbar, da jeweils 5 Fragebögen vorgegeben wurden) ab. Die Quotenmerkmale wurden auf der Basis der amtlichen Daten für jeweils 100 Zielpersonen errechnet (das entspricht der Prozentuierungsbasis, auf der Grundlage gerundeter Zahlen). Die so bestimmten individuellen Quotenmerkmale wurden dann in einem weiteren Schritt auf 20 Aufgabenblätter - eines für jeden Interviewer - aufgeteilt. Die Aufgabenblätter waren sukzessiv hintereinander auf einen Stapel gelegt, so daß nach 20 Aufgabenblättern (mit zusammen 100 Zielpersonen) einmal die Quotenstichprobe "abgearbeitet" war. Da die Zahl der Interviewer nicht einem Vielfachen der 100 Zielpersonen entsprach, muß notwendigerweise in der Gesamtstichprobe eine leichte Verzerrung auftreten - bedingt durch die (fiktiven) Zielpersonen, die Bestandteil der Quotenstichproben sind, aber nicht mehr in die Verteilung kamen. Die Aufgabenblätter waren in Stapeln zusammen mit den Fragebögen im Anschluß an eine Vorlesung ausgelegt, die Übernahme der Aufgabenblätter wurde überwacht, so daß eine selektive Zusammenstellung dem Interviewer "genehmer" Quoten entfällt.

7 Die Durchführung der Untersuchung war bei ihnen Bestandteil einer Hausarbeit, die neben einer Klausur in die Benotung des Pflichtscheins in den "Methoden der empirischen Sozialforschung I" einging. In der Hausarbeit sollten die methodischen Erfahrungen bei der Auswahl der Befragten, der Durchführung der Interviews und die Probleme des Fragebogens kurz dargestellt und diskutiert werden.

len Umfragen üblich sind (vgl. ***Noelle-Neumann*** 1995) - Ähnlichkeiten oder Unterschiede zu Randomstichproben auch auf der Ebene nicht-quoterierter Merkmale ergeben. Gerade bei Kritikern des Quotensamples wird dieses potentielle Problem immer wieder thematisiert.

Die Altersquoten waren relativ global gehalten, gemäß den veröffentlichten Statistiken des Statistischen Landesamtes für Dresden: in der Kategorisierung 18-24, 25-44, 45-59, 60 Jahre und älter. Weil der Anteil der Personen mit Hochschulreife nach Daten des Statistischen Amtes in Dresden bei rund einem Fünftel lag, sollte maximal einer der fünf Befragten über die Hochschulreife verfügen. Die entsprechende Vorgabe war im Begleitschreiben als Empfehlung formuliert, nicht als zwingende Vorgabe.

3.2 Verstöße gegen die Quotenvorgaben

Natürlich stellt sich bei einer Untersuchung, die im Rahmen studentischer Ausbildung stattfindet und die in die Benotung eines Pflichtscheins mit eingeht, zwangsläufig die Frage nach der Fälschung von Interviews. Man kann nicht notwendigerweise davon ausgehen, daß alle Studenten - selbst wenn sie sozialwissenschaftliche Fächer studieren und motiviert sein müßten - immer auch an entsprechenden Erfahrungen in der Durchführung empirischer Untersuchungen interessiert sind. Um das Ausmaß der Fälschung und deren Implikation abzuschätzen, haben wir die Teilnehmer der Veranstaltung zu Beginn des Folgesemesters, nach Abgabe der Hausarbeiten, unter anonymen Bedingungen über die von ihnen durchgeführten Interviews befragt. Derartige anonyme Dunkelfelduntersuchungen haben sich in der Vergangenheit bei der Analyse abweichenden, kriminellen Verhaltens als brauchbar erwiesen - mit durchweg hohem Validitätscharakter (vgl. ***Reuband*** 1994, S. 52ff., ***Kreuzer*** 1995). Zwei Drittel der Interviewer waren bei dieser Dunkelfeldbefragung zugegen, von systematischen Verzerrungen in der Zusammensetzung kann nicht ausgegangen werden.[8]

Als wichtigstes Ergebnis der Dunkelfelduntersuchung zeigt sich: Die Datenqualität der Befragung ist im allgemeinen akzeptabel, die Zahl der Totalfälschungen dürfte mit einem Anteil von 3% nicht höher liegen als in qualitativ hochstehenden Umfragen auf Randombasis, wie dem ALLBUS (vgl. ***Koch*** 1995). Häufiger als Totalfälschungen kommen Teilfälschungen des Interviews vor - doch auch sie stellen nicht die Regel, sondern die Ausnahme dar. Rund 17% der durchgeführten Interviews sind davon betroffen. Wenn es zu Teilfälschungen kommt, dann partiell auf der Ebene einzelner Interviews und Fragen und

8 Der Kreis der Teilnehmer in der Folgeveranstaltung ist nicht vollständig mit den der ersten Methodenveranstaltung identisch, weil einige Studenten von der empfohlenen Reihenfolge abwichen und zunächst Methoden II und dann erst Methoden I besuchten. Daher konnten nicht alle Teilnehmer in Methoden I zu ihrer Hausarbeit befragt werden.

nicht notwendigerweise von vornherein mit dem Vorsatz der Täuschung. In vielen Fällen wurde der Befragte aufgrund des geschätzten Alters und der Bildung ausgewählt, und es stellte sich erst am Schluß des Interviews heraus, daß diese Schätzung falsch war und eine Auswahl nicht hätte erfolgen dürfen. In einem solchen Fall steht der Interviewer vor der Wahl, das Interview zu wiederholen oder es in den jeweils vorgegebenen Quotenmerkmalen zu "schönen" (sofern ihm Abweichungen nicht von vornherein explizit eingeräumt werden).

In unserer Dunkelfelduntersuchung wurde die Zahl der gefälschten Interviews und getrennt davon die Art der Fälschungen erfaßt. Weil die Art der Fälschung getrennt erfaßt wurde, kann nicht festgestellt werden, wie viele Interviews nur im Hinblick auf spezifische Merkmale - ob alle oder nur ein einziges - gefälscht wurden. Geschätzt werden kann nur die *maximale* Fehlerquote. Sie beinhaltet die Zahl, die möglich wäre, wenn im Hinblick auf die aufgeführten sozialen Merkmale genauso häufig gefälscht wurde wie gefälschte Interviews insgesamt existieren. In maximal 6% der Fälle wurden Personen außerhalb Dresdens befragt, ohne daß dies im Fragebogen oder der Hausarbeit vermerkt wurde. Geht man davon aus, daß es sich um die unmittelbare Umgebung Dresden handelt (aus der diese Studenten stammen), dürfte diese Verzerrung nicht von allzu großem Nachteil sein, die sozialstrukturellen Merkmalen und Einstellungen der Befragten werden sich nicht allzu sehr von den Befragten in Dresden unterscheiden.

Wenn Altersvorgaben nicht befolgt wurden, handelt es sich in der Regel um triviale Verstöße. In maximal 5% aller Interviews weicht das reale Alter von dem im Interview angegebenen Alter nicht mehr als drei Jahre ab. Größere Abweichungen von den Vorgaben sind mit einem Anteil von 2% selten. Inwiefern dadurch systematische Verzerrungen bedingt wurden, ist ungewiß. Theoretisch könnte die Abweichung mal in die eine und mal in die andere Richtung gehen und sich im Gesamtergebnis ausgleichen. Es könnte aber auch sein, daß in diesen Fällen überproportional häufig jüngere Befragte, die für studentische Interviewer besonders leicht zu erreichen sind, ausgewählt wurden.

Vom Bildungsniveau der Interviewer und ihren Eltern her besteht bei Studenten - besonders in Ostdeutschland, einschließlich Dresden - eine überproportional große Gelegenheitsstruktur für Kontakte mit besser gebildeten Personen (***Wagner*** 1996, S. 19). Die Gefahr, den Weg des geringsten Aufwands zu gehen und diese als Interviewpartner zu bevorzugen, ist entsprechend groß. Doch wenn man den Angaben in der Dunkelfelduntersuchung folgt, wurde auch die Empfehlung bezüglich der Bildung der Befragten weitgehend eingehalten. Bei maximal 11% der Befragten dürften die Angaben zur Bildung "geschönt" worden sein: Der Befragte ist realiter höher gebildet (mit Hochschulreife) als es angegeben wurde. Der tatsächliche Wert auf der Ebene der einzelnen Befragten dürfte, wie spätere Untersuchungen mit detaillierter Erfassung belegen, niedriger liegen (***Reuband*** 1998c). Andere Verfäl-

schungen auf der Ebene der vorgegebenen Quotenmerkmale (z.B. Geschlecht) oder sonstige Angaben kommen nach den Angaben der Interviewer praktisch kaum vor (2%).

Alles in allem ist der Anteil total oder teilweise gefälschter Interviews auf der Ebene der einzelnen Quotenmerkmale nicht groß genug, um von einer bedeutsamen Einschränkung der Datenqualität auszugehen. Die Beeinträchtigungen dürften insgesamt eher gering sein.[9] Davon abgesehen haben wir aufgrund der Dunkelfelduntersuchung die Möglichkeit, die Verfälschungen in ihren Folgen abzuschätzen.

3.3 Auswahlstrategien

Untersuchungen über die Auswahlstrategien bei Quotenstichproben fehlen bislang, obwohl sie seit langer Zeit als wichtig beurteilt werden, schon früh ein Forschungsbedarf konstatiert (vgl. ***Schmidtchen*** 1962) und auch später gelegentlich die Notwendigkeit einer Analyse der "Auswahlprozesse" bei Quoten-Interviews betont wurde. Wir haben in unserer Untersuchung im Anschluß an das Interview über ein Kontaktprotokoll die Art der Rekrutierung und den Ort des Interviews erfaßt. Dabei zeigt sich: unter den Gegebenheiten der Wahlmöglichkeiten entscheiden sich die meisten Interviewer dafür, ihre Befragten aus dem Kreis der ihnen bekannten Personen zu gewinnen. 14% der Befragten stammen aus dem engeren Familienkreis, 6% sind sonstige Verwandte, 23% Freunde und Bekannte und 38% Nachbarn.

Wie sich im Vergleich dazu die bei Umfrageinstituten tätigen Interviewer verhalten, ist ungewiß. Wo immer die Interviewer am eigenen Heimatort befragen dürfen und nicht gezwungen sind, diese in einer anderen Gemeinde zu rekrutieren (wie z.B. beim Institut für Demoskopie und anderen Instituten), dürfte vermutlich eine je nach Institut mal stärkere und mal schwächere analoge Tendenz bestehen.

Der Kreis der Nachbarn umfaßt im Fall der studentischen Interviewer sowohl Nachbarn im Elternhaus als auch Nachbarn am eigenen Wohnsitz. Letzteres kann im Fall des Wohnens im Studentenheim gelegentlich auch den Hausmeister oder anderes Personal mit einschließen. Personen, die dem Interviewer unbekannt sind, werden selten gewählt, und dann in der Regel über andere Personen vermittelt. Dadurch bedingt liegen bereits gewisse Vorinformationen über den sozialen Hintergrund der Zielperson vor.

---

9 Bezogen auf die Interviewer liegt der Anteil der Fälschungen höher: fast die Hälfte hat in der einen oder anderen Form gegen vorgegebene Regeln verstoßen. Dabei beruht der Regelverstoß nicht auf einer prinzipiellen Ablehnung der Aufgabe als vielmehr auf gewissen Schwierigkeiten, die damit verbundenen Auflagen umzusetzen. Es sind primär wohl situationsspezifische Zwänge, die mehrheitlich für die Verstöße verantwortlich sind. In der Mehrzahl der Fälle handelt es sich um ein oder zwei Fragebögen und jeweils dort sind auch nur einzelne Fragen betroffen.

13% der Befragten sind Bekannte von Verwandten oder Bekannten. Nur 5% sind Fremde, die in der Regel auf der Straße oder in Cafés kontaktiert wurden. Der unerwartet niedrige Anteil von Fremden, die auf der Straße oder in Cafés befragt wurden, mag damit zusammenhängen, daß es in der vorliegenden Befragung um die schriftliche Beantwortung eines längeren Fragebogens geht. Kontakte auf der Straße, die einem Zeitdruck unterliegen, eignen sich nicht dafür.

Tabelle 1: Rekrutierungsmuster in der Quotenstichprobe nach Alter der Befragten (in %)

	Alter				
	18-24	25-44	45-59	60 +	Insgesamt
Engerer Familienkreis	-	9	22	18	14
Sonstige Verwandte	3	6	9	6	6
Freunde, gute Bekannte	65	36	9	1	23
Nachbar	24	29	42	53	38
Kontakt über Bekannte	4	12	14	18	13
Fremde	4	8	4	4	5
	100	100	100	100	100
(N=)	(69)	(194)	(153)	(143)	(559)

Welche Quotenvorgaben der Interviewer erhält, hat Konsequenzen für die Rekrutierungsstrategie. Im Fall der Vorgabe für das Geschlecht sind sie am geringsten. Allenfalls eine leichte Tendenz ist erkennbar, nach der weibliche Zielpersonen überproportional aus dem engeren Familienkreis gewählt werden. Bei den anderen Rekrutierungsmustern ergeben sich keine größeren Unterschiede. Grundlegend anders die Situation bei den Merkmalen Alter und Bildung: Wie man Tabelle 1 entnehmen kann, werden die 18-24jährigen, die im Alter dem Interviewer am nächsten stehen, bevorzugt aus dem Freundes- und Bekanntenkreis rekrutiert. Mit zunehmendem Alter der Zielpersonen nimmt diese Rekrutierungsart ab und dafür gewinnt dann die Nachbarschaft an Bedeutung, ebenso - wenn auch auf schwächerem Niveau - die Kontaktaufnahme, vermittelt über andere Personen. Weiterhin läßt sich zeigen, daß die Befragten mit Hauptschulabschluß überproportional aus der Nachbarschaft stammen, Befragte mit Abitur überproportional aus dem Kreis der Freunde und Bekannten (Tabelle 2).

Dies bedeutet: Personen mit ähnlichen sozialen Merkmalen wie der Interviewer werden von ihm bevorzugt aus dem Freundeskreis gewählt, Personen mit sozial heterogenen Merkmalen aus der Nachbarschaft. Die Nachbarschaft eignet sich offenbar am besten als Rekrutierungsfeld für sozial heterogene Kontakte. Und die Ausgangsbedingungen für eine Befragung sind in solchen Beziehungen günstig. Denn Nachbarn kennt man und hat in gewissem Umfang Vorabinformationen zu deren sozialer Lage. Die soziale Beziehung ist seit längerem hergestellt und die Wahrscheinlichkeit einer Ablehnung des Interviews - im Vergleich zu einem Kontakt mit Fremden - gering.

Tabelle 2: Rekrutierungsmuster in der Quotenstichprobe nach Bildung der Befragten (in %)

	Bildung				Insgesamt
	Hauptschule	Realschule	Fachhoch-schulreife	Abitur	
Engerer Familienkreis	16	13	19	9	14
Sonstige Verwandte	5	8	6	4	6
Freunde, gute Bekannte	5	25	18	42	23
Nachbarn	55	34	30	31	38
Kontakt über Bekannte	16	14	18	10	14
Fremde	3	7	10	4	5
	100	100	100	100	100
(N=)	(134)	(194)	(63)	(144)	(535)

Anmerkungen zur Bildungsvariablen: Hauptschule einschließlich POS 8. Klasse, Realschule einschließlich POS 10. Klasse, Fachhochschulreife einschließlich Abschluß Fachoberschule, Abitur einschließlich EOS.

Angenommen werden kann, daß sozial heterogene Nachbarschaften in Ostdeutschland durchweg häufiger sind als in Westdeutschland: Wohnungen konnten schließlich vor der Wende nicht über den freien Markt erworben werden, sie wurden vielmehr vom Wohnungsamt zugewiesen, und eine Tendenz besser Verdienender mit einer Wohnung in "guter Lage" zu versehen, gab es nicht. Nach der Wende vollzog sich mit der Privatisierung des Wohnungsmarktes und der Herausdifferenzierung unterschiedlich guter sozialer Lagen erst allmählich eine soziale Entmischung. So stark wie in Westdeutschland ist sie noch nicht (vgl. ***Harth et al.*** 1998).

Inwieweit die je nach Befragtenmerkmalen unterschiedlichen Rekrutierungsstrategien Konsequenzen für die Qualität der Antworten haben, bedarf der Prüfung im Rahmen zukünftiger Forschung. Daß bestehende enge soziale Beziehungen nicht notwendigerweise zu einer Reduktion der Datenqualität führen müssen - wie oftmals vermutet -, zeigt sich daran, daß die Nennung fiktiver Politiker in solchen Konstellationen seltener vorkommt als in sozial distanzierten Beziehungen (***Reuband*** 1998c). In sozial heterogen zusammengesetzten Interviewerstäben - wie sie bei den kommerziellen Umfrageinstituten üblich sind (***Reuband*** 1984) - dürften sich die differentiellen Rekrutierungsmuster überdies ausgleichen. So wie die jüngeren Interviewer jüngere Befragte aus dem Freundeskreis befragen, so werden die älteren die gleiche Strategie bei ihren Altersgenossen wählen. Denn Freundeskreise sind sozial homogen zusammengesetzt (vg. u.a. ***Schneider*** 1969, ***Reuband*** 1974), was heißt: die Zugangschancen zu Personen mit bestimmten sozialen Merkmalen sind maßgeblich auch eine Funktion der Interviewerstäbe und deren sozialer Zusammensetzung.

4 Ähnlichkeiten und Unterschiede zwischen den Stichproben

4.1 Sozialstrukturelle Merkmale

Wie unterscheiden sich nun die Stichproben in ihrer Zusammensetzung? In einem ersten Schritt soll zunächst untersucht werden, ob sich die Random- und Quotenstichprobe - jeweils bezogen auf die Dresdner Population - in den grundlegenden sozialstrukturellen Merkmalen ähneln und wie ihre Qualität, gemessen an dem Mikrozensus des Statistischen Amtes, einzuschätzen ist. In einem weiteren Schritt sollen dann die Einstellungen und Verhaltensweisen sowie die psychische Befindlichkeit der Befragten in den Vergleich einbezogen werden.

Gemessen an der Altersverteilung und an der Verteilung des Merkmals Geschlecht sind die beiden Stichproben weitgehend deckungsgleich und stimmen mit den amtlichen Unterlagen nahezu überein (Tabelle 3). Dies ist zunächst ein durchaus bedeutsames Ergebnis - immerhin hätte man aufgrund früherer Erfahrungen im Fall postalischer Befragungen selektive Ausfälle in bestimmten Altersgruppen erwarten können. Und im Fall der Quotenstichprobe hätte man aus praktischen Gründen in unserer Studie ebenfalls nicht notwendigerweise eine völlige Übereinstimmung mit der amtlichen Statistik vermuten müssen.[10]

Kombiniert man die Merkmale Alter und Geschlecht, so werden - obwohl die Quotenmerkmale einzeln und nicht in Kombination miteinander vorgegeben wurden - die amtlichen Zahlen in der Quotenstichprobe sehr gut reproduziert (Tabelle 4). In der Altersgruppe der 25-44jährigen und der 45-59jährigen weichen die Werte der Quotenstichprobe - anders als im Fall der Randomstichprobe - von den wahren Werten praktisch nicht ab. In der Altersgruppe der 18-24jährigen sind die Werte der Quotenstichprobe etwas stärker den Werten der amtlichen Statistik angenähert als die Werte der Randomstichprobe. Und in der Altersgruppe der über 60jährigen sind die Abweichungen zwischen den Verfahren gleich groß.

Diese Ähnlichkeiten in der Kombination der Merkmale zwischen Quotenverfahren und amtlicher Statistik sind um so bemerkenswerter, als in der Literatur oftmals die These vertreten wird, daß überall dort, wo keine Quotierung vorgegeben wird, größere Abweichungen aufträten. Wenn man nur einfache Quoten und keine kombinierte Quote vorgäbe - also z.B. nicht angibt, wieviele der Zielpersonen ein bestimmtes Geschlecht *und* Alter haben

10 Schließlich wurden bei der Ausgabe der Aufgabenblätter an die "Interviewer" im Anschluß an die Seminarveranstaltung mehr Begleitschreiben mit Quotenvorgaben ausgelegt als tatsächlich genommen wurden. Die Zahl der potentiellen Interviewer war zu hoch kalkuliert. Daher hätten die Quotenvorgaben, die in das Feld gingen, unter Umständen nur bedingt die ursprüngliche Quotierung reproduziert. Es hätte zudem auch dazu kommen können, weil einige Interviewer nicht alle Interviews in Dresden durchführten (letztere sind hier aus der Betrachtung ausgeklammert). Nach dem Prinzip des Zufalls dürfte dieses zwar keine Konsequenzen für die Dresdner Befragung haben, aber im Fall einer systematischen Auswahl - etwa im Hinblick auf den Ort der Befragung - hätte es Implikationen haben können.

Tabelle 3: Geschlecht, Alter und Familienstand nach Art der Stichprobe und amtlichen Daten (in %)

	Random	Quota	Amtliche Daten
Geschlecht			
Mann	45	47	46
Frau	55	53	54
Alter			
18 - 24	8	12	10
25 - 44	38	35	37
45 - 59	28	28	26
60 +	26	25	26
Familienstand			
Verheiratet	63	52	59
Ledig	22	30	23
Geschieden	8	8	8
Verwitwet	7	9	10

Quelle Amtliche Daten: Landeshauptstadt Dresden 1995, eigene Berechnungen.
Anmerkung: Geschlecht und Alter (in etwas anderer Kategorisierung) waren als Quotenmerkmal vorgegeben, der Familienstand nicht.

Tabelle 4: Männeranteil nach Alter und Stichprobenart (in %)

	Alter				
	18-24	25-44	45-59	60 +	Insgesamt
Random-(Netto)-Stichprobe	42	43	55	38	45
Quotenstichprobe	59	51	48	33	46
Amtliche Daten	52	51	49	35	47
Random-Brutto-Stichprobe	50	48	55	38	48

Quelle Amtliche Daten: Landeshauptstadt Dresden 1995, eigene Berechnungen.

sollen - dann sei die Gefahr fehlerhafter Kombinationen auf Subgruppenebene groß (z.B. ***Bortz*** 1984, S. 342). Dies ist aber nach unseren Daten für die hier verglichenen Merkmale nicht der Fall.

Warum aber, so stellt sich im Anschluß an unsere Befunde die Frage, weichen dann die Ergebnisse der Randomstichprobe von der amtlichen Statistik so sehr ab? Denkbar wäre, daß sie z.T. von der amtlichen Statistik abweichen, weil bereits die Bruttostichprobe aus dem Einwohnerverzeichnis verzerrt ist. Zwar sollte eine derartige Verzerrung bei der Ziehung der Bruttostichproben nicht vorkommen (auf eine entsprechende Prüfung wird in empirischen Untersuchungen demgemäß in der Regel auch verzichtet). Doch ist ein solcher

Fall theoretisch durchaus denkbar. Daher haben wir die entsprechenden Angaben aus der Bruttostichprobe zusätzlich in die Tabelle mit aufgenommen. Man kann ihr entnehmen, daß die Bruttostichprobe in der Tat in manchen Altersgruppen von den Daten des Mikrozensus abweicht und einige der nennenswerten Differenzen in der Nettostichprobe - z.B. bei den 45-59jährigen - darüber erklärt werden könnten. Die Verzerrungen in der Random-Netto-Stichprobe - der Stichprobe der real Befragten - sind also teilweise Folge der Stichprobenziehung. Die Quotenstichprobe unterliegt in unserer Untersuchung keinem vergleichbaren "Bias". Sie ist nahezu identisch mit den Daten der amtlichen Statistik bezüglich des Merkmals Geschlecht.

So sehr auch die Quotenstichprobe in den Merkmalen Alter und Geschlecht der amtlichen Statistik angenähert und in dieser Hinsicht der Randomstichprobe überlegen ist, so wenig ist andererseits garantiert, daß dies auch für andere zentrale Merkmale zutrifft. So ist in unserem Fall die Quotenstichprobe in der Reproduktion der familialen Verhältnisse deutlich der Randomstichprobe unterlegen, Nicht-Verheiratete sind überrepräsentiert. Was bedeutet: Entweder gibt es in der Quotenstichprobe auf seiten der Interviewer eine Tendenz, bevorzugt Personen zu befragen, die nicht verheiratet sind. Oder es gibt Verzerrungen in der Zusammensetzung, die in unserem Globalvergleich auf der Ebene der hier gewählten Altersgruppen zunächst nicht sichtbar werden, aber für die Verteilung des Familienstandes gewichtige Konsequenzen haben.

So könnte man z.B. die Hypothese aufstellen, daß Interviewer dazu neigen, bevorzugt Personen aus dem Freundes- und Bekanntenkreis zu befragen. Unter diesen Umständen hätten altersgleiche Personen, mit denen am häufigsten Kontakte bestehen, die größte Chance, ausgewählt zu werden. Ist die Altersvorgabe relativ global - wie in diesem Fall für die 25-44 jährigen - wäre vorstellbar, daß in dieser Alterskategorie bevorzugt jene Personen aufgenommen werden, die altersmäßig dem Interviewer am nächsten stehen, also besonders jung sind. Der Interviewer würde in diesem Fall der Quotenvorgabe genügen und hätte doch zugleich den Vorteil, auf sein Potential an Kontaktpersonen zurückgreifen zu können.

Eine differenzierte Altersaufgliederung dokumentiert, daß Tendenzen dieser Art tatsächlich vorkommen. Der Anteil der Befragten in der Alterskategorie 25-44, die knapp über 25 Jahre sind, ist in der Quotenstichprobe größer als in der Randomstichprobe. Der Anteil der Verheirateten wird dadurch bedingt reduziert. Doch auch bei den älteren Altersgruppen, bei denen eine analoge Selektion gemäß des Alters nicht erwartet werden kann, sind die Nicht-Verheirateten überrepräsentiert. Entweder ist diese Gruppe von Personen bevorzugtes Ziel der Interviewer - man unterstellt ihnen überproportional viel Zeit für das Interview aufgrund fehlender familialer Verpflichtungen - oder sie erweisen sich mangels anderer Kontakte für Befragungen als besonders aufgeschlossen.

Im Begleitschreiben zur Durchführung der Quotenstichprobe hatten wir den Interviewern empfohlen, es solle unter den befragten Personen allenfalls eine von fünf über einen Fachochschulabschluß oder Abitur verfügen. Diese Vorgabe wurde - wie unsere Daten belegen - nicht immer eingelöst. Sie wurde oftmals nicht eingelöst, weil sich erst nach der Durchführung des Interviews herausstellte, daß man sich in der Bildung des Befragten verschätzt hatte. Man handelte in diesen Fällen auf der Basis unzureichender Vorinformation oder Fehleinschätzung des Bildungshintergrundes in der Situation der Kontaktaufnahme. In der Quotenstichprobe geben 37% der Befragten an, über eine Fachhochschulreife oder das Abitur zu verfügen, gemäß der amtlichen Statistik dürfte der Anteil jedoch nur 18% betragen. Bemerkenswerterweise weicht der Anteil der besser Gebildeten in der Randomstichprobe in dieser Hinsicht vom Anteil in der Quotenstichprobe nicht nennenswert ab. Auch in der Randomstichprobe sind die besser Gebildeten stärker vertreten als im Mikrozensus - ein Befund, der generell die Umfrageforschung im Rahmen von face-to-face-Interviews auszeichnet. Personen mit niedrigerer Bildung stehen der Umfrageforschung distanzierter gegenüber und nehmen seltener an Umfragen teil (***Scheuch*** 1974, ***Pappi*** 1979, ***Hartmann*** und ***Schimpl-Neimanns*** 1992).

Die Überrepräsentation der besser Gebildeten scheint auf den ersten Blick sowohl die Random- als auch die Quotenstichprobe gleichermaßen stark zu betreffen. Tatsächlich jedoch dürfte sie in der Quotenstichprobe noch etwas höher liegen als es unsere Zahlen ausweisen. So wissen wir aus der Dunkelfelduntersuchung, daß einige Interviewer die Bildung nicht immer korrekt vermerkt und die hohe Bildungsqualifikation des Befragten (Abitur, Fachhochschule) verschwiegen haben. Bei *maximal* 11% der Fragebögen ist eine derartige Verfälschung wahrscheinlich (der reale Wert dürfte darunter liegen). Darüber hinaus gibt es auch umfrageinterne Hinweise für eine Verfälschungstendenz. Wenn man davon ausgeht, daß Paarbeziehungen durch eine Tendenz zur Bildungshomogamie gekennzeichnet sind, müßten die Werte für den Befragten und seinen Partner angenähert sein. Dies trifft für die Randomstichprobe auch zu, nicht aber für die Quotenstichprobe. Der Partner weist dort etwas häufiger eine höhere Bildung als der Befragte auf. Weitere Analysen belegen, daß dies für alle Bildungsgruppen - mit Ausnahme derer mit Abitur - gilt. Gleichwohl: Alles in allem sind die Verzerrungen nicht derart groß, daß man von einer bedeutsamen Verfälschung der Daten sprechen könnte.

Aber gilt dies auch, wenn die Daten nach Altersgruppen untergliedert werden? Eine naheliegende Strategie für studentische Interviewer wäre, bei Quotensamples den Weg des geringsten Aufwands zu gehen und die Befragten mit Hochschulreife bevorzugt aus ihrem eigenen studentischen Milieu zu rekrutieren. Der Aufwand ist gering, der Kreis potentieller Befragten groß. Wenn dies der Fall wäre, würden die Befragten mit Fachhochschulreife oder Abitur auf die jüngste Befragtengruppe konzentriert sein. Angesichts der Tatsache, daß Studenten überproportional aus besser gebildeten Kreisen stammen - in Ostdeutsch-

land mehr noch als in Westdeutschland - wäre zudem denkbar, daß die Eltern der Interviewer, und damit Personen im mittleren Alter, eine besonders gute Chance aufweisen, als Befragte in die Erhebung einbezogen zu werden. Träfe dies zu, müßten die besser Gebildeten mit Fachhochschulabschluß oder Abitur im Quotensample stärker als im Randomsample vor allem in den jüngeren und mittleren Altersgruppen zu finden sein. In den höheren Altersgruppen dagegen müßte angesichts der weitgehenden Gleichheit in der Zusammensetzung der Gesamtstichprobe die Quotenstichprobe einen etwas geringeren Anteil aufweisen als in der Randomstichprobe. Dies ist in der Tat der Fall. Die Abweichungen treten vor allem in den jüngeren Altersgruppen konzentriert auf, die mittleren Altersgruppen sind weniger davon betroffen (Tabelle 5). Gleichwohl: Der Effekt ist insgesamt geringer, als man es hätte vermuten können.

Tabelle 5: Bildung nach Alter und Stichprobenart (in %)

	Alter									
	18-24		25-44		45-59		60 +		Insgesamt	
	Random	Quota	Random	Quota	Random	Quota	Random	Quota	Random	Quota
Schüler	2	4	*	*	-	-	-	-	*	*
Ohne Abschluß	-	1	-	1	1	1	1	3	1	1
Hauptschule	3	1	6	7	33	25	55	55	26	23
Realschule	49	40	52	48	28	32	13	16	35	35
Fachhochschulreife	9	4	12	10	17	15	13	12	14	11
Abitur	35	47	27	31	20	29	11	11	21	26
Sonstiges	3	1	3	3	1	5	7	3	3	3
	100	100	100	100	100	100	100	100	100	100
(N=)	(66)	(79)	(307)	(201)	(232)	(157)	(207)	(145)	(812)	(573)

* <0,5 %

Obwohl Unterschiede in der Zusammensetzung nach Schulbildung zwischen den Stichproben existieren, sind sie doch nicht stark genug, um größere Auswertungen auf der Ebene der Erwerbs- und Berufsstruktur zu bedingen. Sowohl bei der Frage zur gegenwärtigen Erwerbstätigkeit und dem erfahrenen Wandel seit der Wende (Tabelle 6) als auch bei der Frage zum zuletzt ausgeübten Beruf (Tabelle 7) lassen sich keine größeren Differenzen nachweisen - auch dort nicht, wo es um Kategorien von Personen geht, zu denen der Interviewer im Alltag einen besonders engen Kontakt hat und man bei einem ausschließlichen Einsatz studentischer Interviewer am häufigsten Verzerrungen in Quotenstichproben erwarten könnte: der Anteil der Schüler und Studenten ist in der Quotenstichprobe nur wenig größer - um die sechs Prozentpunkte - als in der Randomumfrage.

Tabelle 6: Berufstätigkeit im Vergleich zur Zeit vor der Wende nach Stichprobenart (in %)

	Random	Quota	Insgesamt
Noch am selben Arbeitsplatz	19	20	19
Anderer Arbeitsplatz, aber gleicher Betrieb	7	4	6
In anderem Betrieb tätig	26	27	27
Umschulung	2	2	2
ABM	1	1	1
Arbeitslos	7	7	7
Frühruhestand	5	4	5
Rentner	18	15	17
Sonstiges *	15	20	17
	100	100	100

* einschließlich Student, Schüler, vorher nicht berufstätig.

Tabelle 7: Gegenwärtig bzw. zuletzt ausgeübter Beruf nach Stichprobenart (in %)

	Random	Quota	Insgesamt
Ungelernter Arbeiter	6	5	6
Facharbeiter	27	27	27
Höherer Angestellter / Beamter	11	9	10
Unterer, mittlerer Angestellter / Beamter	38	38	38
Akademische, freie Berufe	1	1	1
Selbständig	5	5	5
Schüler / Student	6	12	8
Hausfrau	2	1	3
Sonstiges	3	3	3
	100	100	100

Daß Verfälschungen in nennenswertem Umfang bei den Erwerbs- und Berufsangaben vorgekommen sind, ist nicht anzunehmen. In der Dunkelfelduntersuchung gibt es keine Hinweise dafür, daß Fragen zur Erwerbs- und Berufstätigkeit verändert, womöglich gar der gefälschten Schulbildung "angepaßt" wurden. Bemerkenswert ist schließlich auch, wie sehr sich die Zahlen in der Random- und Quotenstichprobe für jene Gruppen gleichen, die normalerweise nur schwer für Umfragen erreichbar sind und besonders in Quotenstichproben als unterrepräsentiert gelten (vgl. ***Allerbeck*** und ***Hoag*** 1983, S. 25): Arbeitslose oder andere Personengruppen in nicht-stabilen Erwerbsverhältnissen (wie ABM). Würde man den Anteil der Arbeitslosen auf den Anteil derer beziehen, die prinzipiell für eine Erwerbstätigkeit zur Verfügung stehen, käme man in unserer Erhebung auf die gleichen Werte wie in der offiziellen Arbeitslosenstatistik. So sind unter den Personen, die zu den Erwerbspersonen zählen, in der Random- und der Quotenstichprobe jeweils rund 11% arbeitslos, die

Arbeitslosenquote in Dresden beläuft sich im Jahresdurchschnitt 1995 nach den Daten des Landesarbeitsamtes ebenfalls auf 11% (vgl. Statistisches Landesamt des Freistaates Sachsen 1996, S. 209).

Angesichts der weitgehenden Ähnlichkeit in den Merkmalen Schulbildung und Berufstätigkeit - und der Tatsache, daß das Einkommen mit diesen Merkmalen korreliert - erstaunt um so mehr, daß die Angaben zum Netto-Haushalts-Einkommen die beiden Stichproben voneinander divergieren. In der Quotenstichprobe wird durchschnittlich ein geringerer Wert angegeben als in der Randomstichprobe. So geben in der Randomstichprobe 46 % an, unter 3.000,-- DM zu verdienen. In der Quotenstichprobe sind es 58 % (Cramer's V=.18). Ob es sich bei diesem Befund um eine selektive Repräsentation bestimmter Bevölkerungsgruppen handelt oder ein selektives Antwortverhalten, kann hier nicht geklärt werden.

4.2 Einstellungen, Verhalten und psychische Befindlichkeiten

In der bisherigen Analyse haben wir ausgewählte sozialstrukturelle Merkmale der Befragten in die Analyse einbezogen und sowohl Unterschiede als auch Gemeinsamkeiten zwischen den beiden Stichprobenverfahren festgestellt. Ein gewichtiges Argument gegen den Einsatz von Quotenstichproben ist nun aber, daß die sozialstrukturellen Merkmale, die gewöhnlich in der Literatur zur Beurteilung der Güte von Quotenstichproben herangezogen werden, nur begrenzt aussagefähig sind. Die Befragten mögen zwar in den sozialstrukturellen Merkmalen, die in die Quotenvorgaben eingehen, mit der amtlichen Statistik übereinstimmen. Doch über die Übereinstimmung in den anderen Variablen ist damit nichts ausgesagt.

Der vorliegende Fragebogen bietet hierzu ein breites thematisches Spektrum, um diese Frage zu klären - mit Fragen zu den Themen Familie, Jugend, Politik, Kriminalität, wirtschaftliche Lage, soziale und kulturelle Aktivitäten u.a. Für die Analyse haben wir mit nahezu allen Variablen (95 Variablen) eine Kreuztabelle mit der Art des Stichprobenverfahrens durchgeführt und ausgewählte Kennziffern zur Bestimmung der Stichprobeneffekte berechnet: die Prozentpunktdifferenz (PD) zwischen Random- und Quotensample auf der Ebene der Randverteilung sowie Cramer's V. Je größer der Wert ist, desto stärker der Effekt. Werte für Cramer's V unter .10 können als relativ unbedeutend und vernachlässigenswert angesehen werden. Werte über .10 bedürfen der stärkeren Beachtung.

Wie man Tabelle 8 entnehmen kann, sind die Stichprobeneffekte in der Regel relativ gering. In 61% der betrachteten Differenzen liegen sie bei 5% und weniger, in 23% bei 6-9% und in 16% bei 10 und mehr Prozentpunkten. In 67% der Variablen liegt der Wert für Cramer`s V unter .10, in 21% zwischen .10 und .14, in 8% zwischen .15 und .19 und in lediglich 3% der Fälle bei .20 und mehr. Differenziert nach Thematik zeigt sich: in den Einstellungen ebenso wie in der Einstufung psychischer Befindlichkeiten existieren fast

keine Unterschiede. Die Teilnehmer in den beiden Stichproben differieren z.B. nicht in ihrer Lebenszufriedenheit oder dem Vorkommen von Gefühlen des Unglücklichseins. Wenn Unterschiede bestehen, sind es die Teilnehmer an der Quotenbefragung, die sich über die verschiedenen Einstellungsbereiche hinweg eher postmaterialistisch geben: sie bejahen eher das Erziehungsziel "Selbständigkeit und freier Wille" (statt "Gehorsam und Unterordnung"), betonen eher "Selbstverwirklichung", "Lebensgenuß" als Wert und seltener "Sicherheit/Geborgenheit" und sind überdies häufiger Gegner der Todesstrafe. Hinweise für die Annahme, Befragte in Quotenstichproben würden sich eher vor Kriminalität fürchten als Befragte in Randomstichproben, finden sich nicht. Weder bei der Sorge,

Opfer zu werden, noch bei dem üblicherweise zur Messung der Kriminalitätsfurcht eingesetzten sog. „Standardindikator“ ist dies der Fall ("Wie sicher fühlen Sie sich in Ihrer Wohngegend, wenn Sie abends bei Dunkelheit allein auf die Straße gehen?"). Im Gegenteil: die Mitglieder der Quotenstichprobe bekunden tendenziell sogar weniger Furcht (48% vs. 55% "sehr unsicher/etwas unsicher").

Dies ist darum ein wichtiger Befund, weil in der Bundesrepublik bei manchen Vergleichen über die Zeit zum Thema Kriminalitätsfurcht Quoten- und Randomsamples gleichermaßen verwendet werden müssen, um überhaupt Aussagen über Langzeitwandel treffen zu können. Unter diesen Umständen kann es für die Beurteilung von Wandel zentral sein, welcher Wechsel im Stichprobenverfahren mit einer Veränderung in der Furcht einher geht. Wenn die Kriminalitätsfurcht bei Befragten in Quotenstichproben qua Methode niedriger liegt als in Randomsamples, und der Wechsel von einer Quoten- zur Randomstichprobe zeitgleich mit einem Rückgang in der Kriminalitätsfurcht einher geht (vgl. z.B. ***Reuband*** 1992b), dann wird der Wandel in der Furcht einen realen Rückgang repräsentieren. Im umgekehrten Fall, beim Wechsel von einer Random- zur Quotenstichprobe, wäre dagegen bei einer analogen Entwicklung der Kriminalitätsfurcht die Gefahr eines methodischen Artefakts nicht auszuschließen. Denn ein Rückgang wäre in diesem Fall auch aufgrund des Wechsels im Stichprobenverfahren zu erwarten.

Die Variable mit der größten Wirkung in unserer Studie - mit einem Cramer's V- Wert von .23 - betrifft die Beurteilung des Fragebogens selbst, die am Schluß des Interviews erhoben wurde: "Wie oft ist es im vorherigen Fragebogen vorgekommen, daß eine Frage unklar formuliert war und Sie nicht wußten, was gemeint war? Sehr oft, oft, selten, nie." Hier sind es die Teilnehmer an der Quotenbefragung, die sich dezidiert kritischer äußern. Dieser Befund dürfte Eigenheiten der Befragung widerspiegeln. Wie ***Elisabeth Noelle-Neumann*** dargelegt hat (1995), kann es bei der Auswahl der Befragten in Quotenstichproben thematische Selektionseffekte geben: der Interviewer wählt, wenn es sich um eine thematisch homogene Studie handelt, bevorzugt jene Personen, die einen besonders engen Zusammenhang zum Thema aufweisen.

Tabelle 8: Einstellungen/Befindlichkeiten nach Stichprobenart (in %)

	Random	Quota	PD	Cramer's V
Wie gern in Dresden? "Sehr gern"	65	53	- 12	.12
Wie gern in Gegend? "Sehr gern"	40	31	- 9	.10
Besuch mehrmals im Jahr:				
Museen	42	42	0	.06
Theater	39	53	+ 14	.18
Kino	37	50	+ 13	.18
Schaufensterbummel	84	82	- 2	.11
Freunde	79	82	+ 3	.16
Lokale	61	69	+ 7	.18
Liebste Stunden: Arbeit und Freizeit gleich	77	71	- 6	.10
Freizeit, 3 und mehr Stunden	52	47	- 5	.13
Häufigkeit TV 1 x Woche und mehr:				
Unterhaltungssendungen	51	48	- 3	.10
Nachrichtensendungen	94	90	- 4	.14
Politische Sendungen	59	53	- 6	.14
Kriminalfilme	41	41	0	.05
Magazine über Kriminalfälle	32	24	- 8	.13
Vertrauen in Menschen: Ja	41	45	+ 4	.06
Verhältnis zu Nachbarn: helfen uns/befreundet	60	57	- 3	.12
Zahl der Freunde: 3 und mehr	59	69	+ 10	.22
Engster Freund in Dresden	65	66	+ 1	.08
Engster Freund Telefonbesitz	77	74	- 3	.03
Engster Freund Treffen einmal die Woche	32	39	+ 7	.15
Engster Freund Telefonkontakt einmal die Woche	32	41	+ 9	.14
Sorgen "sehr stark /stark":				
Umwelt	24	23	- 1	.08
Kriminalität	19	14	- 5	.07
Arbeitslosigkeit	34	27	- 7	.07
Wohnung	18	18	0	.09
Miete	34	33	- 1	.07
Betrug	17	14	- 3	.05
Familienmitglied arbeitslos	41	40	- 1	.08
Asylanten	33	19	- 14	.22
Bedrohung auf der Straße	47	30	- 17	.19

Fortsetzung Tabelle 8	Random	Quota	PD	Cramer's V
Wichtigkeit von Lebensbereichen "sehr wichtig":				
Eheglück	78	75	- 3	.05
Kinder	61	60	- 1	.02
Freunde	25	36	+ 11	.13
Selbstverwirklichung	29	35	+ 6	.07
Sicherer Arbeitsplatz	62	54	- 8	.09
Gefallen an Arbeit	56	54	- 2	.03
Kollegen	39	33	- 6	.06
Geborgenheit	60	50	- 10	.13
Freiheit	34	37	+3	.05
Sparen	24	17	- 7	.19
Phantasie	19	23	+ 4	.06
Lebensgenuß	21	32	+ 11	.13
Für andere dasein	33	35	+ 2	.06
Attraktives Äußeres	13	12	- 1	.03
Hohes Gehalt	14	15	+ 1	.04
Opfer von Kriminalität innerhalb der letzten 12 Monate	27	24	- 3	.04
Persönlich bedroht innerhalb der letzten 12 Monate	15	17	+ 2	.06
Viktimisierungschancen "sehr wahrscheinlich / wahrscheinlich" in nächsten 12 Monaten:				
Anpöbeln	60	62	+ 2	.07
Diebstahl	58	48	- 10	.03
Raub	28	19	- 9	.10
Körperverletzung	31	22	- 9	.10
Einbruch	28	19	- 9	.09
Vergewaltigung	8	8	0	.04
Wehren gegen Angriff "sehr wahrscheinlich"	47	50	+ 3	.06
Verteidigungschancen "sehr hoch / hoch"	24	24	0	.04
Wohngegend "sehr sicher / ziemlich sicher"	45	52	+ 7	.09
Politisches Interesse "sehr stark / stark"	44	41	- 3	.07
Wirtschaftliche Lage "sehr gut / gut"	51	49	- 2	.05
Kindererziehung zur Selbständigkeit	46	59	+ 13	.09
DDR-System überzeugt / positiv	37	44	+ 7	.09
Wiedervereinigung vollzogen wie geschehen	39	36	- 3	.07
Eigene Lage ohne Wiedervereinigung besser / gleich gut	41	39	-2	.03

Fortsetzung Tabelle 8	Random	Quota	PD	Cramer's V
Einstellungen: Zustimmungen "sehr stark / stark"				
Leben genießen	54	59	+ 5	.08
Mann Geld verdienen	14	12	- 2	.06
Frau: Kind statt Berufstätigkeit	17	15	- 2	.05
Beruf für Frau normal	95	94	0	.02
Eltern immer lieben	73	63	- 10	.11
Eltern für Kinder einstehen	43	43	0	.03
Keine Kinder	42	39	- 3	.05
Wiedervereinigung Gutes für Osten	69	68	- 1	.01
Geduld mit Osten	53	46	- 7	.07
Mehr Opfer für Osten	52	57	+ 5	.06
Vereinigung Nutzen für Westen	65	57	- 8	.08
Sozialismus gute Idee	71	72	+1	.06
Westdeutsche arbeiten im Osten zu ihrem Vorteil	72	65	- 7	.09
Fremd fühlen	35	35	0	.04
Politik zu kompliziert	31	31	0	.03
Stasi Angst immer /gelegentlich	23	21	- 2	.08
Stasi in Nachbarschaft /Arbeit (Kenntnis)	15	17	+ 2	.06
Lebensqualität sehr zufrieden / zufrieden	53	55	+2	.07
Einsam	17	13	- 4	.09
Wohl fühlen	52	52	0	.03
Nicht zufrieden mit Leben	21	16	- 5	.09
Gehetzt fühlen	57	55	- 2	.04
Ungleichheit	17	17	- 5	.11
Tatendrang	70	66	- 4	.07
Zukunftsangst	53	47	- 7	.11
Für Todesstrafe	41	28	- 13	.14
Gesundheit sehr gut / gut	52	56	+ 4	.07
Arbeitsplatz völlig sicher / eher sicher	53	55	+ 2	.04
Parteipräferenz CDU	38	30	- 8	.12
Unklare Fragen nie	42	24	- 18	.23
Befragung gefallen sehr gut / gut	55	41	- 14	.16
Zu erneutem Interview bereit	76	75	- 1	.05

Aufgeführt in Kurzbezeichnung sind die für die Analyse herangezogenen Fragen.

Dies geschieht vermutlich aufgrund der Annahme, daß diese besonders stark für eine Teilnahme zu motivieren sind, nur daß sich in unserem Fall der Focus nicht aus dem Thema, sondern der Aufgabe des Interviews ergibt: der studentische Interviewer, bemüht Informa-

tionen für seine Hausarbeit über methodische Probleme des Fragebogens zu sammeln, ermuntert den Befragten zu entsprechenden Äußerungen und schärft so von vornherein dessen Blick für die gestellte Aufgabe. In einigen Fällen mag er auch mit dem Appell um Hilfe bei der Beurteilung des Fragebogens an die Zielperson herangetreten sein und sie darüber zur Mitarbeit motiviert haben. Wäre die Aufgabe - wie sonst in Quotenumfragen - eine Befragung ungeachtet der methodischen Präferenzen auf seiten des Interviewers gewesen, wäre dieser Effekt vermutlich nicht aufgetreten.

Die Variable, die als nächste in der Effektstärke folgt und mit einem Wert von .22 die Stichproben voneinander unterscheidet, betrifft die Zahl der Freunde. ("Denken Sie nun einmal an enge Freunde/Freundinnen, die Ihnen sehr nahe stehen, nicht an Ihren Ehepartner oder Partner und auch nicht Familienangehörige. Können Sie sagen, wieviel enge Freunde/Freundinnen Sie haben?") In der Quotenstichprobe sind die Personen stärker vertreten, die viele Freunde haben. Daß Teilnehmer an Quotenumfragen generell sozial aktiver sind, deuten weitere Befunde in unserer Untersuchung an, nach denen sie häufiger ins Theater gehen (Cramers V= .18), in das Kino (.18), in ein Lokal/ Restaurant (.19) oder Freunde besuchen (.16). Und auf einer Liste über die Wichtigkeit verschiedener Lebensbereiche wird von ihnen die Aussage "viel mit Freunden zusammen sein" häufiger auch als "sehr wichtig" eingestuft (.13). In dieser Hinsicht reproduzieren unsere Befunde die hier bereits früher referierten Ergebnisse von ***Schmidtchen*** (1962) und ***van Koolwijk*** (1974), denen zufolge die soziale Aktivität der Befragten ein zentrales Unterscheidungsmerkmal für die Zusammensetzung von Random- und Quotenstichproben darstellt.

Unerwartet ist, daß sich in unserer Untersuchung mit einem Wert von .22 die Sorge vor einem Asylantenzuzug "Ich mache mir Sorgen, ... daß immer mehr Asylanten nach Dresden kommen") zur Unterscheidung der beiden Stichproben als ebenso bedeutsam erweist wie die Zahl der Freunde. Warum die Befragten der Quotenstichprobe seltener die Angst vor Asylanten äußern als die Befragten der Randomstichprobe, ist hierbei ungewiß. Eine Möglichkeit liegt darin, daß sich hierin z.T. die bereits zuvor erwähnten postmaterialistischen Einstellungen ausdrücken. Postmaterialisten sind, wie ***Ronald Inglehart*** belegt, insgesamt weniger nationalistisch eingestellt (vgl. ***Inglehart*** 1990, S. 409 ff.). Die Einstellung zu Asylanten könnte - zumal in Ostdeutschland, wo die Ressentiments gegenüber Ausländern weiter verbreitet sind als in Westdeutschland und u.a. im Zusammenhang mit dem Auftreten rechtsradikaler Parteien auch öffentlich thematisiert werden - zu einem zentralen Symbol für unterschiedliche Einstellungskonfigurationen avanciert sein und die beiden materialistischen und postmaterialistischen Lager deutlich separieren.[11]

11 In unserer Untersuchung wurde der Postmaterialismus-Index nicht erhoben, wir können die Hypothese also nicht prüfen. Wir verfügen nur über Aussagen, die als Bestandteil eines postmaterialistischen Wertesyndroms gedeutet werden können.

5 Schlußbemerkungen

Wir haben in unserer Untersuchung, die als explorative Studie zu werten ist, im Rahmen einer schriftlichen Befragung zu prüfen versucht, inwieweit Quoten- und Randomstichproben in ihren Resultaten voneinander abweichen. Dabei wurde festgestellt, daß auf der Aggregatebene der Unterschied zwischen Random- und Quotenstichproben geringer ist als es gewöhnlich in der wissenschaftlichen Literatur vermutet wird. Die Umfragen gleichen einander in den sozialstrukturellen Merkmalen auch dann, wenn die Quoten - wie in unserer Untersuchung - einfach und nicht in kombinierter Weise vorgegeben wurden. Wenn Unterschiede in den sozialstrukturellen Merkmalen zwischen den Stichprobenverfahren auftreten und an der amtlichen Statistik validiert werden können, sind die Folgerungen widersprüchlich: mal erweist sich die Quotenauswahl als brauchbarer, mal die Randomauswahl. Würde man in der Quotenstichprobe jeweils die sozialen Merkmale als Vorgaben mit einbeziehen, die sich in unserer Untersuchung als verzerrt erwiesen haben (wie der Familienstand), würde ihre Qualität zweifellos erhöht. Aus dieser Sicht sind die beobachteten Verzerrungen für Quotenstichproben nicht zwangsläufig. Sie sind auch eine Funktion der Vorgaben. Die Möglichkeit der Korrektur findet freilich ihre Grenzen in der Belastbarkeit des Interviewers. Man kann nicht immer feinere und zusätzliche Quotenvorgaben einführen, weil sonst seine Arbeit so sehr erschwert wird, daß Fälschungen geradezu nahegelegt werden.[12]

Wo zwischen den Stichprobenverfahren im Verhalten größere Unterschiede beobachtet werden können, bestehen sie vor allem in einer stärkeren Selektion sozial aktiver Befragter in Quotenstichproben. Personen mit vielen Kontakten sind offenbar zugänglicher, weil sie in ihrem sozialen Verkehrskreis sichtbarer sind als Personen mit weniger Kontakten. Quotenstichproben scheinen aus dieser Sicht besonders geeignet, den Kreis sozial aktiver Personen einzubeziehen, die in Randomstichproben aufgrund von Nichterreichbarkeit herausfallen. Bei den Verweigerern, die eher durch wenig Kontakte gekennzeichnet sind (vgl. u.a. ***Esser*** 1973), ist die Situation eine andere. Weiterhin deutet sich in unserer Studie an, daß Befragte mit postmaterialistischen Wertorientierungen tendenziell stärker in Quotenstichproben vertreten sind. Inwiefern es sich dabei um eine Überrepräsentation dieser Personengruppe in Quotenstichproben handelt oder eine Unterrepräsentation in Randomstichproben, kann mangels externer Daten über die "wahren" Verhältnisse hier nicht geklärt werden. Ebenso muß offen bleiben, inwiefern sich darin auch die spezifische Zusammensetzung unseres Interviewerstabes - ausschließlich Studenten - widerspiegelt.

---

12 Zu den Quotenplänen, wie sie vom Institut für Demoskopie gewählt werden, siehe ***Noelle-Neumann*** (1995 S. 278). Neben der Gemeindegröße sind dort Alter und Geschlecht (in Kombination) sowie Berufskategorie (in Kombination mit Geschlecht) vorgegeben. In Großbritannien ist die übliche Quotierung eine, die sich auf die Merkmale Geschlecht, Alter und sozialen Status stützt (***Jowell et al.*** 1993 S. 252).

Unsere Untersuchung ist als ein erster Schritt zu einer vertiefenden Analyse des Stellenwertes von Quotenstichproben anzusehen. Im Rahmen weiterer Studien wird zu prüfen sein, welche Auswahlstrategien bei Quotenstichproben von den Interviewern gewählt werden und welche Probleme - etwa in bezug auf die Selektion bestimmter Personengruppen - daraus erwachsen. Zu klären wäre, wie sich Studien mit unterschiedlicher Quotierung und wie sich Random- und Quotenstichproben auf der Ebene von Subgruppen voneinander unterscheiden.[13] Manche Autoren (z.B. ***Scheuch*** 1974) sehen auf der Ebene der Subgruppen die eigentlichen Probleme der Quotenauswahl. Untersuchungen, in denen dieser Frage nachgegangen wird, fehlen indes. Die für die Bundesrepublik verfügbaren Untersuchungen beschränken sich nahezu alle auf einen bloßen Vergleich der Gesamtstichproben.

Schließlich bleibt zu fragen, welchen Einfluß die Zusammensetzung des Interviewerstabes ausübt. Vermutlich werden die Effekte selektiver Rekrutierung, wie wir sie beschrieben haben, reduziert, wenn man einen Interviewerstab wählt, der in den sozialen Merkmalen einem Querschnitt der Bevölkerung ähnelt. Es muß nicht notwendigerweise in allen Merkmalen identisch sein, es genügt - wie in der Praxis der deutschen Sozialforschung und bei Instituten mit Bevorzugung der Quotenauswahl üblich (***Reuband*** 1984) - eine Annäherung. Zu klären wären schließlich auch die Vorzüge und Nachteile der Stichproben und die daraus erwachsenden Möglichkeiten der *komplementären* Nutzung in Abhängigkeit von Problemstellung und Population sowohl im Rahmen der Primär- als auch Sekundäranalyse.

Die Tatsache, daß Quotenstichproben in der deutschen und der ausländischen Sozialforschung weiterhin verbreitet sind und oftmals die einzige Basis für Langzeitvergleiche eröffnen, sollten Grund genug sein, die Eigenheiten von Quotenstichproben genauer zu eruieren statt sie wie bisher zu ignorieren. Quotenstichproben, in der akademischen Sozialforschung allzu oft als unbrauchbar verworfen, bedürfen einer stärkeren systematischen Beachtung, um ihre Möglichkeiten und Probleme besser abschätzen zu können.

Literatur:

Allerbeck, ***K.*** und ***Hoag, W.*** 1983: Zur Auswahlproblematik bei Jugendstudien: ein Plädoyer für die Zufallsauswahl, in: ZA-Information, Nr. 13, S. 24-29.

Alt, C., Bien, W. und ***Krebs, D.*** 1991: Wie zuverlässig ist die Veranschaulichung von Stichprobenverfahren? Random route versus Einwohnermeldeamtstichproben, in: ZUMA Nachrichten, Nr. 28, S. 65 ff.

Bausch, T. 1990: Stichprobenverfahren in der Marktforschung, München.

Bertram, H. (Hrsg.) 1995: Das Individuum und seine Familie Lebensform, Familienbeziehungen und Lebensereignisse im Erwachsenenalter, Opladen: Leske und Budrich, S. 429-455 (Methodischer Anhang).

13 Im vorliegenden Fall verzichten wir auf eine entsprechende Analyse, weil - wie zuvor beschrieben - bei einigen Quotenmerkmalen Verzerrungen auftreten. Subgruppenanalysen stoßen unter diesen Umständen auf Grenzen.

Birke, T. 1998: Quoten-, Random-, Routen- und Einwohnermeldeamt-Stichprobe. Eine vergleichende Analyse. Magisterarbeit im Fach Soziologie, Köln.

Blanke, K., Ehling, M. und ***Schwarz, N.*** 1996: Zeit im Blickfeld. Ergebnisse einer repräsentativen Zeitbudgeterhebung. Schriftenreihe des Bundesministeriums für Familie, Senioren, Frauen und Jugend, Bd. 121, Stuttgart: Kohlhammer.

Bliesch, U. 1998: Interviewerschulung und -kontrolle, in: Statistisches Bundesamt (Hrsg.), Interviewereinsatz und -qualifikation, Schriftenreihe Spektrum Bundesstatistik, Band 11, S. 66-100.

Bortz, J. (unter Mitarbeit von ***Bongers, D.***) 1984: Lehrbuch der empirischen Forschung für Sozialwissenschaftler, Berlin: Springer Verlag

Böltken, F. und ***Gehring, A.*** 1984: Zur Empirie des Postmaterialismus. Quota und Random, Äpfel und Birnen, Kraut und Rüben, in: ZA-Information, Nr. 15, S.38-52.

Context 1991, Folge 25, S. 13-15.

Crespi, L. 1988: Pre-election polling. Sources of accuracay and error. New York: Russel Sage Foundation.

Diekmann, A. 1995: Empirische Sozialforschung. Grundlagen, Methoden, Anwendungen. Reinbek bei Hamburg: Rowohlt.

Dillman, D. 1978: Mail and telefone surveys. New York: Wiley.

Esser, H. 1973: Kooperation und Verweigerung beim Interview, in: ***Erbslöh, E., Esser, H., Reschka, W.*** und ***Schöne, D.***: Studien zum Interview. Meisenheim: Verlag A. Hain, S. 71-141.

Focus 1995, Nr. 40/1995.

Forsa 1996: Umfragen über Umfragen: Ein Jahrzehnt im Überblick, in: Context, Folge 14, S. I - XVI.

GfK 1991: Berufsbild Marktforscher. Das Image der Markt- und Meinungsforschung in Deutschland. Nürnberg: GfK.

Glenn, N. 1977: Cohort analysis. Beverly Hills und London: Sage.

Harth, A., Herlyn, U. und ***Scheller, G.*** 1998: Ausgewählte Segregationsprozesse in ostdeutschen Großstädten, dargestellt am Beispiel der Landeshauptstadt Magdeburg, in: Stadtforschung und Statistik, Heft 2, S. 20-29.

Hartmann, P. und ***Schimpl-Neimanns, B.*** 1992: Sind Sozialstrukturanalysen mit Umfragedaten möglich? Analysen zur Repräsentativität einer Sozialforschungsumfrage, in: Kölner Zeitschrift für Soziologie und Sozialpsychologie, 44, S. 315-340.

Hoag, W. 1986: Der Bekanntenkreis als Universum: Das Quotenverfahren der Shell Studie, in: Kölner Zeitschrift für Soziologie und Sozialpsychologie, 38, S. 123 ff.

Infratest Burke, o.J., Interviewerführung, Basisunterlagen: Auswahlverfahren. München: Infratest Burke.

Inglehart, R. 1990: Culture shift in advanced industrial society. Princeton, N.Y.: Princeton University Press.

Institut für Demoskopie Allensbach, 1975: Allensbacher Werbeträgeranalyse. Experimente. Allensbach: Institut für Demoskopie.

Institut für Demoskopie, 1981: Eine Generation später: Bundesrepublik Deutschland 1953-1979 Allensbach: Institut für Demoskopie.

Institut für Demoskopie, o.J.: ABC der Interviewer. Technische Anweisung zur Durchführung von Interviews. Allensbach: Institut für Demoskopie.

Jowell, R., Hedges, B., Lynn, P., Ferrant, G., Heath, A. 1993: The Polls - A Review. The 1992 British Election: The Failure of the Polls, in: Public Opinion Quarterly, 57:238-263.

Koch, A. 1995: Gefälschte Interviews: Ergebnisse der Interviewerkontrolle beim ALLBUS 1994, in: ZUMA-Nachrichten 36, S. 89 - 105.

Koch, A. 1997: Teilnahmeverhalten beim ALLBUS 1994. Soziodemographische Determinanten von Erreichbarkeit, Befragungsfähigkeit und Kooperationsbereitschaft, in: Kölner Zeitschrift für Soziologie und Sozialpsychologie, 49, S. 98-122.

Koolwijk, J. v. 1974: Das Quotenverfahren, in: ***Koolwijk, J.v.*** und ***Wieken-Mayser, M.*** (Hrsg.): Techniken der empirischen Sozialforschung, Bd. 6, München: Oldenbourg Verlag, S. 81-100.

Kreuzer, A. 1995: Theoretische Überlegungen und Befunde aus Gießener Delinquenzbefragungen, in: ***Kaiser, G.*** und ***Jehle, J.-M.*** (Hrsg.) 1995: Kriminologische Opferforschung. Neue Perspektiven und Erkenntnisse, Teilband II, Verbrechensfurcht und Opferwerdung - Individualopfer und Verarbeitung von Opfererfahrungen, Heidelberg: Kriminalistik Verlag, S. 55-73.

Kromrey, H. 1998: Empirische Sozialforschung, 8. Aufl., Opladen: Leske und Budrich.

Landeshauptstadt Dresden 1993: Bevölkerung am Ort der Hauptwohnung nach Stadtteilen, Geschlecht und Altersgruppen. Dresden: Amt für Statistik und Wahlen.

Landeshauptstadt Dresden 1995: Dresden in Zahlen 1994. Dresden: Amt für Statistik und Wahlen.

Lynn, P. 1994: Evidence of the inaccuracy of quota samples, in: Survey Methods Centre Newsletter, 15 No. 1, S. 20-23.

Lynn, P. und ***Jowell, R.*** 1996: How might opinion polls be improved?: the case for probability sampling, in: Journal of the Royal Statistical Society, A, 159, Part 1, S. 21-28.

Meulemann, H. 1995: Werte und Wertewandel. Zur Identität einer geteilten und wiedervereinten Nation. Weinheim und München.

Moon, N. 1994: The case for quota sampling, in: Survey Methods Centre Newsletter, 15, No. 1, S.9-12.

Moser, C.A. und ***Kalton, C.*** 1971: Survey methods in social investigation. 2nd ed. London.

Müller, W. 1998: Klassenstruktur und Parteiensystem. Zum Wandel der Klassenspaltung im Wahlverhalten, in: Kölner Zeitschrift für Soziologie und Sozialpsychologie, 50, S. 3-46.

Noelle-Neumann, E. und ***Köcher, R.*** 1993: Allensbacher Jahrbuch der Demoskopie 1984-1992, Bd. 9, München: Saur Verlag.

Noelle-Neumann, E. 1994: Methoden der Publizistik und Kommunikationswissenschaft, in: ***Noelle-Neumann, E., Schulz, W. und Wilke, J.*** (Hrsg.): Publizistik Massenkommunikation, Frankfurt a.M.: Fischer, S. 267-307.

Noelle-Neumann, E. und ***Petersen, Th.*** 1996: Alle, nicht jeder. Einführung in die Methoden der Demoskopie. München: dtv.

Noelle-Neumann, E. 1998: Einsatz und Qualifikation der Interviewer als Qualitätsmerkmal der Umfrageforschung., in: Statistisches Bundesamt (Hrsg.), Interviewereinsatz und -qualifikation, Schriftenreihe Spektrum Bundesstatistik, Band 11, S. 10-50.

Pappi, F.U. 1979: Der Beitrag der Umfrageforschung zur Sozialstrukturanalyse, in: ***Pappi, F.U.*** (Hrsg.): Sozialstrukturanalysen mit Umfragedaten, Königstein: A. Hain.

Planung und Analyse 1998, Heft 5, S.68.

Pöschl, H. 1993: Werbung und Beteiligung der Haushalte an der Einkommens- und Verbraucherstichprobe, in: Wirtschaft und Statistik, 6, S. 385-390.

Radtke, G.A. und ***Zeh, C.*** 1974: Random oder Quota? Auswahlverfahren in der Meinungsforschung, in: Politische Meinung. Sonderheft, S. 124-152.

Reuband, K.-H. 1974: Differenzielle Assoziation und soziale Schichtung. Dissertation, Universität Hamburg.

Reuband, K.-H. 1980: Life Histories. Problems and Prospects of Longitudinal Designs. in: ***Clubb, J.M.*** und ***Scheuch, E.K.*** (Hrsg.): Historical Social Research. The Use of Historical and Process Produced Data. Stuttgart: Klett Cotta, S. 135-163.

Reuband, K.-H. 1984: Zur Rekrutierung und sozialen Zusammensetzung von Interviewerstäben. in: ***Meulemann, H.*** und ***Reuband, K.-H.*** (Hrsg.): Soziale Realität im Interview. Empirische Analyse methodischer Probleme. Frankfurt / New York: Campus, S.61-80.

Reuband, K. -H. 1989: On the Use of Self Reports in Measuring Crime among Adults. Methodological Problems and Prospects, in: ***Klein, M.*** (Hrsg.): Cross-National Research in Self Reported Crime and Delinquency. Dordrecht/Boston/London: Kluwer Academic Publishers, S. 89-106.

Reuband, K.-H. 1992a: Drogenkonsum und Drogenpolitik. Deutschland und die Niederlande im Vergleich. Opladen: Leske und Budrich.

Reuband, K.-H. 1992b: Objektive und subjektive Bedrohung durch Kriminalität. Ein Vergleich der Kriminalitätsfurcht in der Bundesrepublik Deutschland und den USA 1965-1990, in: Kölner Zeitschrift für Soziologie und Sozialpsychologie, 44, p. 341-353.

Reuband K.-H. 1994: Soziale Determinanten des Drogengebrauchs. Eine sozialwissenschaftliche Analyse des Gebrauchs weicher Drogen, Opladen: Westdeutscher Verlag.

Reuband, K.-H. 1995: Drogenkonsum und Drogenpolitik in Westeuropa. Epidemiologische Befunde im Vergleich, in: Beilage zur Wochenzeitung Das Parlament, B9/95, S.22 - 31.

Reuband, K.-H. und ***Blasius, J.*** 1996: Face-to-Face, telefonische und postalische Befragungen. Ausschöpfungsquote und Antwortmuster in einer Großstadtstudie, in: Kölner Zeitschrift für Soziologie und Sozialpsychologie, 48, S. 296-318.

Reuband, K.-H. 1998a: Umfragen in Ostdeutschland - Verbreitung, Erscheinungsformen und Bewertung durch die Befragten, in: Planung und Analyse, S. 58-63.

Reuband, K.-H. 1998b: Postalische Befragungen in den neuen Bundesländern. Strategien zur Erhöhung der Rücklaufquote und Auswirkungen auf die Zusammensetzung der Befragten. Unveröffentlichtes Manuskript. Düsseldorf.

Reuband, K.-H. 1998c: "Pseudo-Opinions" in Bevölkerungsumfragen. Wie die Bürger fiktive Politiker beurteilen. Unveröffentlichtes Manuskript. Düsseldorf.

Sahner, H. 1995: Gestern Legitimationsinstrument, heute Politikersatz? Empirische Sozialforschung und politisches System, in: ***Reuband, K.-H., Pappi, F.U.*** und ***Best, H.*** (Hrsg.): Die deutsche Gesellschaft in vergleichender Perspektive. Opladen: Westdeutscher Verlag, S. 287-304.

Scheuch, E.K. 1972: The time-budget-interview, in: ***Szalai, A. et al.*** (Hrsg.): The use of time. Daily activities of urban and suburban populations in twelve countries. Paris: Mouton, S. 67-87.

Scheuch, E.K. 1973: Das Interview in der Sozialforschung, in: ***König, R.*** (Hrsg.): Handbuch der Empirischen Sozialforschung, Bd. II, 3. umgearb. und erw. Auflage, Stuttgart: Enke, S. 66 - 190.

Scheuch, E.K. 1974: Auswahlverfahren in der Sozialforschung, in: ***König, R.*** (Hrsg.), Handbuch der empirischen Sozialforschung, Bd. 3a, 3. Auflage, Stuttgart: Enke.

Schmidtchen, G. 1962: Die repräsentative Quotenauswahl. Bericht über ein Quota-Random-Experiment des Instituts für Demoskopie Allensbach.

Schmidtchen, G. 1997: Wie weit ist der Weg nach Deutschland? Sozialpsychologie der Jugend in der postsozialistischen Welt. Opladen: Leske und Budrich.

Schneider, A. 1969: Expressive Verkehrskreise. Eine empirische Untersuchung zu freundschaftlichen und verwandtschaftlichen Beziehungen. Dissertation Köln.

Schnell, R. 1993: Die Homogenität sozialer Kategorien als Voraussetzung für "Repräsentativität" und Gewichtungsverfahren, in: Zeitschrift für Soziologie, 22, S. 16-32.

Schnell, R. und ***Kohler, U.*** 1995: Empirische Untersuchung einer Individualisierungsthese am Beispiel der Parteipräferenz von 1953-1992, in: Kölner Zeitschrift für Soziologie und Sozialpsychologie, 47, S. 634 ff.

Schnell, R. 1997: Nonresponse in Bevölkerungsumfragen. Ausmaß, Entwicklung, Ursachen. Opladen: Leske und Budrich.

Schnell, R. 1998: Besuchs- und Berichtsverhalten der Interviewer., in: Statistisches Bundesamt (Hrsg.), Interviewereinsatz und -qualifikation, Schriftenreihe Spektrum Bundesstatistik, Band 11, S. 156-170.

Statistisches Landesamt des Freistaates Sachsen, 1996: Statistisches Jahrbuch Sachsen 1996. Kamenz.

Sudman, S. 1966: Probability sampling with quotas, in: Journal of the American Statistical Association, 61, 749-771.

Taylor, H. 1995: Horses for courses: how different countries measure public opinion in very different ways. in: The Public Perspective, 6, No. 2, S. 3-5.

Wagner, B. 1996: Studieren in Ostsachsen. Ergebnisse der 14. Sozialerhebung des Deutschen Studentenwerkes für die Hochschulstandorte Dresden, Zittau und Görlitz. Dresden: Studentenwerk.

Zum Problem der Abbildung eindimensional konzipierter Konstrukte bei entgegengesetzter Itempolung

von Herbert Matschinger [1] und Dagmar Krebs [2]

Zusammenfassung

Eindimensional konzipierte Konstrukte werden häufig durch Variablen mit entgegengesetzter Polung operationalisiert. Die durch Hauptkomponentenanalyse festzustellende Struktur besitzt fast immer eine 2. Achse, die durch vorzeichenidentische Ladungen aller Items gekennzeichnet ist, was zu verschiedenartigsten Interpretation Anlaß gibt (z.B. ***Inkonsistenzachse****). Es wird gezeigt, daß diese Struktur das Resultat eines ganz bestimmten datenerzeugenden Prozesses (ideal point process) ist, der unter Umständen zu eingipfeligen Itemcharakteristiken (ICC) führt. Dies widerspricht der Voraussetzung von wenigstens monotonen ICC's für die Hauptkomponentenanalyse. Anhand künstlicher Daten und eines empirischen Beispiels („Ausländerfeindlichkeit") werden die Ergebnisse der Hauptkomponentenanalysen mit denen einer - eingipfelige ICC's annehmenden - Unfoldinganalyse verglichen. Eine mögliche Interpretation der artifiziellen 2. Achse (quadratische Form der 1. Achse) als* ***Intensitätsachse*** *wird diskutiert.*

Abstract

Unidimensional constructs are frequently measured by oppositely worded items. Principal component analysis (PCA) yields mostly an artificial 2- dimensional structure which leads to different interpretations (e.g. inconsistency axis), since the second axis exhibits equal signs of the loadings. It can be shown that the very structure results from a particular datagenerating mechanism (ideal point process) which produces single- peaked item characteristic curves (ICC), violating the assumption of at least monotonic characteristics for principal component analysis. By means of artificial data and an empirical example („attitudes towards guest-workers") we compare the results between PCA and an Unfoldinganalysis. A possible interpretation of the second axis (quadratic form of the first axis) as an indicator for „intensity" is discussed.

1 Dr. ***Herbert Matschinger*** ist Wissenschaftlicher Mitarbeiter an der Klinik und Poliklinik für Psychiatrie der Universität Leipzig, Johannisallee 20, 04317 Leipzig /e-mail: math@server3.medizin.uni-leipzig.de

2 Dr. ***Dagmar Krebs*** ist Professorin am Institut für Soziologie der Universität Gießen, Karl-Glöckner-Straße 21E, 35394 Gießen /e-mail: dagmar.krebs@sowi.uni-giessen.de

1 Einleitung und Problemstellung

Mit der vorgelegten Arbeit soll versucht werden, ein Dilemma bei der Interpretation der Dimensionsstruktur eines eindimensional konzipierten Konstruktes näher zu beleuchten und Lösungen dafür anzubieten. Das Dilemma tritt nahezu immer dann auf, wenn zur methodischen Behandlung von Akquieszenz Variablen mit entgegengesetzter Polung (im Hinblick auf das abzubildende Konstrukt) Verwendung finden. Aus der Literatur ist bekannt, daß in solchen Fällen die sparsame Abbildung durch Hauptkomponentenanalyse 2 Dimensionen liefert, die - nach meist orthogonaler Rotation - nur die Polung der Variablen repräsentiert (***Sjöberg*** and ***Perrson*** 1979). Verblüffend für den Forscher ist in einem solchen Fall das Resultat, weil damit ”behauptet” wird, daß die durch diese Variablengruppen abgebildeten Teilkonstrukte voneinander unabhängig (unkorrreliert) seien. Dies ist um so unverständlicher, als diese Variablengruppen konzeptionell in den allermeisten Fällen nur gegensätzliche Enden **eines und nur eines** theoretisch begründeten Konstruktes bezeichnen sollen. Die empirisch vorgefundene 2-Dimensionalität wird stets als lästiges Artefakt empfunden und oft einfach ignoriert.

Häufig wird vorgeschlagen, keine Rotation durchzuführen, da die unrotierte Lösung anschaulich interpretierbar sei. Die unrotierte Lösung zeichnet sich fast immer dadurch aus, daß die 2. Achse invariant gleiche Vorzeichen der Ladungen aller Items (unabhängig von der Polung) aufweist. Diese Tatsache führt zu der Interpretation, daß die 2. Achse die **Inkonsistenz** (vergl. ***Rostampour*** 1998) des Antwortverhaltens abbildet, weil vorzeichenidentische Ladungen bei gegensätzlichen Polungen einer anderen inhaltlichen Interpretation nicht unmittelbar zugänglich sind. (Mitunter ist auch die 1. Achse durch gleiche Vorzeichen der Ladungen ausgezeichnet und die 2. Achse bildet das Konstrukt in der gewünschten Weise ab.) Der Verdacht, es hier mit einem Artefakt zu tun zu haben, resultiert aus dem Hinweis, daß bipolar operationalisierte Konstrukte nicht durch Hauptkomponentenanalyse auf ihre sparsame Struktur hin untersucht werden sollten, da es sich dabei um ein inadäquates Modell handelt (***Coombs*** and ***Smith*** 1973; ***van Schuur*** and ***Kiers*** 1994; ***van Schuur*** and ***Kruijtbosch*** 1995; ***Roberts*** and ***Laughlin*** 1996). Schon ***Coombs*** and ***Kao*** (1960) weisen darauf hin, daß bei der Analyse von Ähnlichkeitsdaten immer ein zusätzlicher Faktor gefunden wird, der als Artefakt des Projektionsverfahrens gewertet werden muß. Gegensätzliche Itempolung aber erzeugt genau diese (in vielen Fällen unerwünschte) Bipolarität, auch wenn sich diese durch das Konzept selbst nicht begründen läßt; Lösungsmöglichkeiten und Beispiele finden sich bei ***van Schuur*** (1993, 1997), ***Rost*** (1996). Die Wahl des Abbildungsmodells sollte sich daher am Datentyp orientieren, der durch die spezifische Operationalisierung eines Konstrukts zustandekommt.

1.1 Das Konstrukt ”Ausländerfeindlichkeit”

Um den Gang der Argumentation zu verdeutlichen, werden 11 Items zur Messung von Ausländerfeindlichkeit verwendet werden. Gemessen wurde die Zustimmung zu den Items mittels einer Antwortskala von 1=”stimme überhaupt nicht zu” bis 5=”stimme voll und ganz zu”. Der Wortlaut der Items befindet sich im Anhang. Ausländerfeindlichkeit ist als eindimensionales theoretisches Konstrukt konzipiert, das mittels positiv und negativ formulierter Indikatoren operationalisiert ist, wobei die negativ formulierten Items überwiegen. Die Antwortvorgaben wurden in gleicher Weise, unabhängig von der Polung, eingesetzt. Die Daten stammen aus der Studie ”Gießener Meinungen 1996”, die als telefonische Befragung bei einer Zufallsstichprobe von n=602 Bewohnern der Stadt Gießen durchgeführt wurde.

1.2 Abbildung durch Hauptkomponentenanalyse

Das oben beschriebene Dilemma der Zweidimensionalität als Ergebnis der Hauptkomponentenanalyse (PCA) kann an dem in Tabelle 1 dargestellten Datenbeispiel beobachtet werden. Wie eingangs beschrieben, bilden die zwei Dimensionen nach der Rotation die Polung der Items ab. Die unrotierte Lösung der PCA zeichnet sich bei entgegengesetzt gepolten Items dadurch aus, daß der 1. Faktor das Konstrukt mit den der Itempolung entsprechenden unterschiedlichen Vorzeichen der Ladungen in der gewünschten Weise abbildet, während auf der 2. Achse die Ladungen aller Items (unabhängig von der Polung) gleiche Vorzeichen haben. Wir beobachten damit auch hier das in der Einleitung beschriebene Phänomen.

Tabelle 1: Hauptkomponentenanalyse der 11 Ausländerfeindlichkeitsitems (siehe Anhang)

	Factor 1	Factor 2	Varimax-rotiert	
Eigenwert	4,8	1,1		
V10A	,67544	,17558	-,363	,596
V10B	,72468	,24538	-,350	,680
V10C	,59407	,43541	-,124	,726
V10D	,50620	,27168	-,175	,547
V10E	,56714	,37857	-,145	,666
V10F	,70557	,26299	-,324	-,191
V32B	-,76713	,32750	,779	-,298
V32C	-,68684	,27458	,685	-,280
V32D	-,75379	,24655	,713	-,347
V32E	-,57357	,37411	,672	-,130

1.3 Problemstellungen

Im folgenden sollen 3 Hauptpunkte behandelt werden:

1. Identifikation des Datentyps bei Variablen mit entgegengesetzter Polung.
2. Identifikation des adäquaten Abbildungsverfahrens.
3. Interpretation der durch verallgemeinerte Hauptkomponentenanalyse (siehe Abschnitt 2.2.1) erzeugten 2-dimensionalen Struktur durch den Vergleich von zwei verschiedenen Abbildungsmodellen.

2 Typen von Daten

Es wird immer wieder darauf hingewiesen, daß im Falle der Annahme von ”Effektindikatoren” (***MacCallum*** and ***Browne*** 1993; ***Blalock*** 1971; ***Bollen*** 1984; ***Hotelling*** 1933; ***Bollen*** and ***Lennox*** 1991), also der Abhängigkeit der manifesten Variablen von der nicht direkt beobacht- und meßbaren latenten Variable, zwei Typen von Daten vorliegen können: ”**Dominanzdaten**” oder ”**Ähnlichkeitsdaten**” (***Coombs*** 1964). In Tabelle 2 sind zwei ”perfekte” Datensätze abgebildet, die dem einen oder anderen Modell entsprechen.

Tabelle 2: Beispiele für:

	Dominanzdaten						**Ähnlichkeitsdaten**					
	V1	**V2**	**V3**	**V4**	**V5**	**V6**	**V1**	**V2**	**V3**	**V4**	**V5**	**V6**
P1	**2**	1	1	1	1	1	**2**	**2**	1	1	1	1
P2	**2**	**2**	1	1	1	1	1	**2**	**2**	1	1	1
P3	**2**	**2**	**2**	1	1	1	1	1	**2**	**2**	1	1
P4	**2**	**2**	**2**	**2**	1	1	1	1	1	**2**	**2**	1
P5	**2**	**2**	**2**	**2**	**2**	1	1	1	1	1	**2**	**2**

Jede Zeile enthält eine Beobachtung (Person); der Datensatz ist in der üblichen ”Rechteckform” dargestellt. Kategorie 2 bedeutet für beide Datensätze die ”Zustimmung”, Kategorie 1 das Gegenteil. Der in der linken Seite der Tabelle dargestellte Datensatz zeigt ein perfektes ”Dreiecksmuster”. Die Summe der Zustimmungen ist eine erschöpfende Statistik für die Lage der Respondenten auf **einer** latenten Dimension. Der rechte Teil der Tabelle enthält eine perfekte Parallelogrammstruktur (***Coombs*** 1964: Kap. 4). Diese Art von Daten entsteht normalerweise (aber eben nicht nur) bei sogenannten Auswahldaten, wobei aus einer Menge vorgegebener Stimuli eine - kleinere - Menge ausgewählt (”präferiert”) werden soll (”wähle 2 aus 6”). Hier ist die Summe der ”Lösungen” (Kategorie 2) **keine** erschöpfende Statistik für die Lage des Respondenten auf der latenten Dimension. Selbstverständlich kommen sowohl die Beobachtungen wie auch die Stimuli in einem natürlichen Datensatz nicht in der hier dargestellten idealtypischen Reihenfolge vor und es wird unter anderem zu zeigen sein, wie die ”ideale” Reihenfolge in einem realen Datensatz durch Permutation von Zeilen und Spalten hergestellt werden kann.

2.1 Typen von Verknüpfungen

Bei Dominanzdaten sind die manifesten Variablen mit der latenten Größe wenigstens monoton verknüpft und in den meisten Fällen wird diese Beziehung durch einen einzigen Parameter beschreibbar. Bei Ähnlichkeitsdaten wird dagegen die Lage von Respondenten und Variablen auf einer latenten Präferenzordnung abgebildet, so daß die Antwortwahrscheinlichkeit durch die Ähnlichkeit zwischen den Positionen der Items und der Person auf der latenten Präferenzordnung bestimmt ist. In diesem Fall sind die ICC's (Item-characteristic-curves = Wahrscheinlichkeitsbeziehung zwischen Antwortverhalten und Ausprägung der latenten Variable) eingipfelig. Schon sehr früh ist auf diese unterschiedlichen Antwortprozesse hingewiesen worden, wobei vor allem die Messung von Einstellungen im Vordergrund stand (***Thurstone*** 1928, 1931). Es wird häufig angenommen, daß Einstellungmessung zu Vergleichsprozessen und damit zu Ähnlichkeitsdaten führt. Formal läßt sich der Unterschied zwischen den beiden Modellen darstellen als (***van Schuur*** and ***Kiers*** 1994, S. 98).

Faktoranalytisches Modell: **Unfoldingmodell:**

$$z_{is} = \sum_{j=1}^{r} b_{ij}\theta_{sj} + \zeta_{is} \qquad \text{versus} \qquad z_{is} = \sum_{j=1}^{r} \left|\theta_{sj} - \delta_{ij}\right|^2 + \zeta_{is}$$

wobei:

r	=	Anzahl der Dimensionen
s	=	Beobachtung
i	=	Variable
j	=	Dimension
z_{is}	=	Wert einer Beobachtung s auf einem Indikator i
θ_{sj}	=	Wert einer Beobachtung s auf einer latenten Dimension j
b_{ij}	=	Ladung eines Indikators i auf der latenten Dimension j
δ_{ij}	=	Wert eines Indikators i auf der latenten Dimension j
ζ_{is}	=	Wert einer itemspezifischen Reaktion einer Person s auf einen spezifischen Indikator i (ohne Berücksichtigung der latenten Dimension j)

Man sieht, daß im faktorenanalytischen Modell die Lage der Indikatoren (δ_{ij}) auf den latenten Continua **nicht** expliziter Bestandteil des Modells ist. Dies ist auch nicht notwendig, weil die Linearkombination $\sum b_{ij}\theta_{sj}$ den modellhaft vorausgesetzten, monotonen Zusammenhang zwischen den Variablen und den latenten Dimensionen beschreibt. Bei Ähnlichkeitsdaten gilt jedoch, daß ein Item immer nur im Vergleich zur Position anderer Items und zur Position der (urteilenden) Person auf der latenten Präferenzordnung positio-

niert werden kann. Deshalb sind im Unfolding Modell die Positionen der Indikatoren auf der latenten Dimension δ_{ij} expliziter Bestandteil des Modells, so daß der (das Antwortverhalten leitende) Abstand zwischen Person und Item ermittelt werden kann. Insofern hat auch die Größe z_{is} im Hinblick auf den Vergleich von Beobachtungen auf der latenten Dimension in den beiden Modellen eine unterschiedliche Bedeutung.

Ein weiterer wichtiger Unterschied zwischen dem faktorenanalytischen Modell und dem Unfolding Modell besteht darin, daß die beiden Modelle unterschiedliche Annahmen über Größe und Vorzeichen der Itemkorrelationen machen. Im faktorenanalytischen Modell (Dominanzdaten) sollen die Korrelationen hoch und (bei identischer Polung) positiv sein. Im Unfolding-Modell (Ähnlichkeitsdaten) sind nur jene Stimuli hoch korreliert, die an ähnlichen Stellen der latenten Dimension liegen. Je weiter die Stimuli voneinander entfernt sind, desto kleiner werden die Korrelationen, bis sie schließlich mit umgekehrten Vorzeichen wieder absolut größer werden. Diese Tatsache ist im Zusammenhang mit den Diagnosekriterien für das Vorliegen einer Unfolding-Skala von Bedeutung.

Ein Verfahren zur dimensionalen Abbildung von Ähnlichkeitsdaten ist das Unfolding Modell, welches immer dann eingesetzt werden sollte, wenn die Daten aus einem ”ideal point process” (***Roberts*** and ***Laughlin*** 1996) stammen. Die Rekonstruktion von Präferenzordnungen aus dem beobachteten Material erfolgt immer aus der Relation von Variablen und Befragten auf einer (**gemeinsam gültigen**) Dimension. Sind die Positionen von Befragten und Items bekannt, so ergibt sich die für einen Befragten empirisch feststellbare Rangordnung durch Faltung der Dimension an eben diesem Punkt (siehe z.B. das Titelbild von ***Post*** 1992). Ziel der Analyse ist es, durch ”Entfaltung” (**Unfolding**) der empirischen Ordnungen die gemeinsam gültige Anordnung (”joint- scale”) zu rekonstruieren.

2.2 Methoden der sparsamen Abbildung (Projektion)

Im Zentrum des Interesses steht der Vergleich von Ergebnissen der sparsamen Abbildung unter der einen oder anderen Verknüpfungsannahme. Im folgenden sollen 2 Zugänge kurz beschrieben werden, wobei der Einfachheit halber immer von dichotomen (oder dichotomisierten) Daten ausgegangen wird. Dies ermöglicht eine übersichtlichere Darstellung der Ergebnisse in Scatterplots, da nur relativ wenige Datenpunkte betrachtet werden müssen.

2.2.1 Homogenitätsanalyse (Multiple Korrespondenzanalyse)

Die Problematik der niedrigdimensionalen Projektion soll zunächst an künstlichen Datensätzen und später am Konstrukt ”Ausländerfeindlichkeit” erläutert werden. Als Projektionsverfahren wird - unter der (möglicherweise inadäquaten) Annahme monotoner ICC’s - die sogenannte Homogenitätsanalyse (***Heiser*** and ***Meulmann*** 1994; ***Greenacre*** 1984) ver-

wendet werden (siehe auch ***Nishisato*** 1980, 1994, 1996; ***Gifi*** 1990; ***van de Geer*** 1993a, 1993b). Das Verfahren ist weit häufiger unter der Bezeichnung ”multiple Korrespondenzanalyse” bekannt. ***Cliff, Collins, Zatkin, Gallipeau,*** and ***McCormick*** (1988) stellten ein Ordnungsverfahren dar, mit dem die ”ideale” Ordnung der Zeilen gefunden werden kann. Der dort (***Cliff et. al.*** :87) dargestellte Algorithmus zeigt sich als Variation der ”method of reciprocal averages” (***Hill*** 1973; ***Horst*** 1935; ***Richardson*** and ***Kuder*** 1933), wobei beide Datentypen ”geordnet” werden können. Das Verfahren von ***Cliff*** unterscheidet sich in einigen Punkten von dem im folgenden verwendeten und soll hier nicht näher behandelt werden. Die Homogenitätsanalyse ist als eine spezifische Form der Hauptkomponentenanalyse zu betrachten, wobei Personen und Items (hier Kategorien der Items) im gleichen, möglichst sparsamen Raum abgebildet werden. Das Kriterium für die Bedeutsamkeit der Achsen dieses Raumes ist auch hier der entsprechende Eigenwert, nur wird nicht ein Eigenwert von 1, sondern ein Eigenwert von 1/k (k = Zahl der Items) als untere Grenze für die Bedeutsamkeit angenommen. Das genannte Verfahren liefert sogenannte optimale Quantifizierungen für die Kategorien und optimale Scores für die Respondenten (Zeilen der Datenmatrix). Dabei sind die Kategorienquantifikationen optimal im Sinne der Homogenität: Kategorien, die Ähnliches bezeichnen, stehen an ähnlichen Stellen des Raumes. Die Scores der Respondenten sind optimal im Sinne der Diskriminierungsfähigkeit der Achse zwischen den Beobachtungen (vgl. ***Gifi*** 1990).

Wie bei der ”metrischen” Hauptkomponentenanalyse wird eine lineare Verlustfunktion minimiert. Für die eindimensionale Abbildung gilt:

$$L(\mathbf{x};\varphi) = m^{-1} \sum_j (\mathbf{x} - \varphi_j(\mathbf{h}_j))'(\mathbf{x} - \varphi_j(\mathbf{h}_j)$$

Dabei ist **x** die zunächst unbekannte Hauptkomponente und $\varphi(\mathbf{h})$ eine zu wählende Transformation der Datenmatrix **h**. **m** ist die Zahl der Variablen (also der unabhängigen Spalten) der Indikatormatrix.

Es sei weiterhin :

$$\varphi(\mathbf{h}) = \mathbf{G}_j\mathbf{y}_j$$

wobei $\mathbf{G}_j$ die Indikatormatrix für eine Variable j ist. Dies enthält so viele Spalten wie die Variable Kategorien und in jeder Zeile eine 1 genau dann, wenn die entsprechende Kategorie von einem Respondenten genannt wurde. Alle anderen Spalten sind in dieser Zeile gleich 0. Die horizontal aneinandergereihten $\mathbf{G}_j$ ergeben die Superindikatormatrix **G**, deren Zeilensummen **immer** gleich k sind (k= Zahl der Variablen). Die Spaltensummen entsprechen den Häufigkeiten der Variablenkategorien in der beobachteten Stichprobe. Diese Art der Codierung wird auch ”complete Coding” genannt (***Gifi*** 1990, S. 140ff), weil nicht das

Auftreten, sondern auch das Nichtauftreten einer Kategorie kodiert wird. Dies hat Folgen für die sparsame Abbildung, von denen in Abschnitt 2.3.2 und 3.3 konkret die Rede sein wird. Es sei weiterhin $\mathbf{y}_j$ eine Gewichtungsmatrix, welche die zunächst unbekannten und optimalen Quantifizierungen der arbiträr codierten Kategorien enthält. Die Homogenitätsanalyse minimiert also die folgende Verlustfunktion:

$$L(\mathbf{x};\mathbf{y}) = m^{-1} \sum_j (\mathbf{x} - \mathbf{G}_j\mathbf{y}_j)'(\mathbf{x} - \mathbf{G}_j\mathbf{y}_j)$$

Es sei weiter $\mathbf{D} = \mathrm{diag}(\mathbf{G}_j'\mathbf{G}_j)$ die diagonale Supermatrix der univariaten Randsummen und $\mathbf{C} = \mathbf{G}_j'\mathbf{G}_j$ die Matrix aller bivariaten Assoziationen. Für den normalerweise betrachteten mehrdimensionalen Fall geschieht die Berechnung von **X** und **Y** durch *singular value decomposition* (***Eckart*** and ***Young*** 1936; vergl. auch ***Bekker*** and ***De Leeuw*** 1988) oder durch ”Alternating least sqare”.

Die Zeilen (**X**)- und Spaltenscores (**Y**) stehen dabei in einer eindeutigen Beziehung:

1. Die Quantifikation einer Variablenkategorie ist der Mittelwert aller Personenscores jener Respondenten, die diese Kategorie gewählt haben.
2. Die Quantifikation einer Zeile der Datenmatrix (also eines Befragten), multipliziert mit dem Eigenwert der Achse, ist der Durchschnitt der Quantifikationen jener Kategorien, die von diesem Befragten genannt wurden.

Das Problem der artifiziellen Mehrdimensionalität stellt sich hier genauso wie bei der metrischen PCA. Die folgenden empirischen Analysen werden sich vor allem auf die graphischen Darstellungen sowohl der optimalen Kategorienquantifikationen (**Y**) wie auch der Beobachtungsscores (**X**) - auch ”objectscores” genannt - stützen.

2.2.2 Multiples unidimensionales Unfolding

Zur Analyse von Unfoldingskalen existieren parametrische (siehe z.B. ***Hoijtink*** 1991; ***Hoijtink*** and ***Molenaar*** 1994) wie auch nichtparametrische Ansätze. Zur Bestimmung einer joint-scale bedienen wir uns des multiple unidimensionalen Unfoldingmodells (***van Schuur*** 1993). Die Prozedur gründet auf der Tatsache, daß - wenn die Menge der Items tatsächlich dem Unfoldingmodell genügt - für jedes auf der Dimension angeordnete Triple von Items ein bestimmtes Antwortmuster nicht existieren darf. Ist eine Anordnung der Items A B und C in eben dieser Reihenfolge gegeben, so wird ein Antwortmuster der Form *”A(Zustimmung) B(Ablehnung) C(Zustimmung)”* als **Fehler** deklariert. Natürlich ist die Interpretation als Fehler nur möglich, wenn eine für alle Respondenten gemeinsame Präferenzordnung existiert und daher auch ordinale Urteile der Letzteren anzunehmen sind (***van Schuur*** and ***Kiers*** 1994, S. 104) . Im Gegensatz zu anderen Unfoldingmodellen wird hier

nicht vorausgesetzt, daß die Befragten bezüglich des genauen Ortes der Items auf dem Präferenzkontinuum übereinstimmen; gefordert ist lediglich Übereinstimmung bezüglich der Anordnung. Das Verfahren mit dem diese J-Skala gefunden wird, heißt ”Seriation” oder ”Ordination” (***van Schuur*** and ***Post*** 1990, S. 4).

Für eine Menge von p Items, die einer J-Skala genügt, wird für jedes der p(p-1)(p-2)/6 geordneten Triple die Häufigkeit der Fehler O bestimmt und mit den unter der Bedingung der statistischen Unabhängigkeit zu erwartenden Häufigkeit E verglichen. ***Loevinger's*** H als Maß der Homogenität H=(1-O)/E dient zur Beurteilung des Ergebnisses. Sind die Items vollständig unabhängig, ist H gleich 0. Hat man es mit einer perfekten Unfoldingskala zu tun, so ist der Wert gleich 1. H sollte weder für die Gesamtskala noch für ein Einzelitem den Wert von 0.30 unterschreiten. Wenigstens 3 weitere Matrizen sind zur Beurteilung der Gültigkeit wie auch der Güte der Skala von Bedeutung:

1. **Dominanzmatrix:** Diese quadratische, aber nicht symmetrische Matrix enthält die Anteile jener Befragten , die den Stimulus i, aber **nicht** den Stimulus j gewählt haben. Die Diagonale ist mit 0 besetzt, weil die Stimuli (Kategorien) ”Zustimmung” und ”Ablehnung” bei einem Item nicht gleichzeitig gewählt werden können. In der Dominanzmatrix sind die Items nach ihrer Position auf der joint-scale von Personen und Items geordnet. Von der Diagonalen aus nimmt die Ähnlichkeit in beiden Richtungen ab, d.h. daß der Anteil der Antworten, die Item i zustimmen und gleichzeitig Item j ablehnen, größer werden muß weil die Unähnlichkeit der Items mit zunehmender Entfernung von der Diagonalen zunimmt. Nimmt die Unähnlichkeit zu, sinkt die Wahrscheinlichkeit einer gleichzeitigen zustimmenden Antwort auf Item i und Item j. Gilt die Unfoldingskala, so müssen die Prozentwerte in jeder Zeile von der Diagonale aus nach links und rechts monoton größer werden.
2. **Adjacencymatrix:** Diese untere Dreiecksmatrix enthält den Anteil der Befragten, die **sowohl** den Stimulus i als auch den Stimulus j gewählt haben. Die Items sind nach ihrer Ähnlichkeit geordnet. Da in der Diagonalen die größte Ähnlichkeit besteht, soll - bei Geltung des Modells - der Anteil der gleichzeitigen Zustimmung zu Item i und Item j in Richtung zunehmender Nähe (von links auf die Diagonale hin) größer werden. Gleichzeitig soll natürlich der Anteil der gleichzeitigen Zustimmung zu Item i und Item j in Richtung abnehmender Nähe (vom Item i spaltenweise nach unten=zunehmender Abstand) geringer werden.
3. **Bedingte Adjacencymatrix:** In jeder Zelle dieser Diagnosematrix steht für jeden gewählten Spaltenstimulus j der Anteil der Befragten, die den Zeilenstimulus i gewählt haben. Die Zeilenanordnung repräsentiert grob die ”trace-line” der ICC des entsprechenden Items und trägt daher zur Entscheidung zwischen kumulativen und Ähnlichkeitsdaten bei (***van Schuur*** and ***Post***. 1990 : Kap. 8). Diese Matrix ist das wichtigste Diagnosekriterium für das Vorliegen einer Unfoldingskala.

2.3 Darstellung des Problems an künstlichen Daten

Im ersten Schritt soll der Vergleich der oben eingeführten Abbildungsverfahren mit künstlichen Daten durchgeführt werden. Dabei werden stets dichotome Daten (Zustimmung/ Ablehnung) verwendet. Wenn die Homogenitätsanalyse imstande ist, die implizite Struktur aufzudecken, so sollte nach Sortierung der Zeilen (oder Spalten) nach den Scores für die 1. Achse die Ähnlichkeitsstruktur richtig geordnet hervortreten. Alle Antwortmuster sind dann - mit Rücksicht auf ihre Ähnlichkeit - auf einer Dimension angeordnet, wobei hier der Begriff der Ähnlichkeit als Begriff der Entfernung auf **einer** Präferenzordnung verstanden werden kann; ob diese Ordnung auch als **Rangordnung** verstanden werden kann, muß im konkreten Fall entschieden werden. Die zur ”Entfaltung” des 2. künstlichen Datenbeispiels und auch des empirischen Beispiels ”Ausländerfeindlichkeit” verwendete Methode wurde schon in Punkt 2.2.2 näher erläutert. Die Analysen wurden mit dem Programm **MUDFOLD 2.0** (***van Schuur*** and ***Post*** 1990) bzw. mit dem Modul ”**Categories**” (Prozedur **HOMALS**) von SPSS (6.1.2) durchgeführt. Die Abbildungen wurden mit STATA 5.0 und S-PLUS 3.3 erzeugt.

2.3.1 Problem ”wähle 5 aus 10”

In diesem Beispiel soll schon berücksichtigt werden, daß die spezifische Form der Daten durch die gegensätzliche Polung der Variablen entsteht. Von 10 Items seien fünf in die eine und fünf in die andere Richtung gepolt (In Tab. 3 ”positiv” bzw. ”negativ” genannt). Die Daten formen eine eindeutige Parallelogrammstruktur, was aus dem ersten Teil der Tabelle 3 nicht sofort ersichtlich ist, weil sich die Antwortmuster nicht in der ”richtigen” Reihenfolge befinden. Hier soll noch einmal darauf hingewiesen werden, daß die Summe der Präferenzen **keine** Statistik für die Lage der Respondenten darstellt: diese ist für jede Person genau 5, diskriminiert also nicht zwischen den Beobachtungen. Beispielhaft sei angenommen, daß sich die Variablen schon in der ”richtigen”, d.h. der J-Skala genügenden, Ordnung befinden, was in einem empirischen Datensatz (wie weiter unten dargestellt) selbstverständlich nicht der Fall ist. Bis auf 2 Respondenten zeigen alle Zeilen Zustimmungen sowohl zu positiv wie auch zu negativ gepolten Stimuli und man kann sich die Frage stellen, ob es sich dabei um Inkonsistenz des Antwortverhaltens oder um eine größere ”Toleranzbreite” bei der Beurteilung der Stimuli handelt. Dies muß im empirischen Beispiel weiter unten Gegenstand ausführlicher Überlegungen sein.

Der Plot in 2 Dimensionen (Abb. 1) zeigt einen ”horse-shoe” (***Heiser*** 1985; ***van Rijckevorsel*** 1987; ***Hill*** 1974). Daraus läßt sich ableiten, daß die Variablen eine und nur eine Dimension indizieren: Die 2. Achse ist die quadratische Form, die hier nicht dargestellte 3. Achse die kubische Form der 1. Achse, usw. Die ”Hufeisenform” stellt die Kombination einer

linearen und einer quadratischen Transformation der gleichen Zahlenreihe dar; also der convexen (oder concaven) Funktion:

$$f(y) = (Ay - B)^2$$

Tabelle 3: Projektion des Problems "wähle 5 aus 10"

Originaldatensatz (Code 2 = Zustimmung)

positiv					negativ					Score1	Score2	ID
A	B	C	D	E	F	G	H	I	J			
1	1	1	2	2	2	2	2	1	1	,293	-1,073	P1
1	1	1	1	2	2	2	2	2	1	,878	-,274	P2
1	1	1	1	1	2	2	2	2	2	1,464	1,341	P3
2	2	2	2	2	1	1	1	1	1	-1,464	1,331	P4
1	2	2	2	2	2	1	1	1	1	-,878	-,260	P5
1	1	2	2	2	2	2	1	1	1	-,293	-1,065	P6

Durch Homals sortierter Datensatz (identische Codierung)

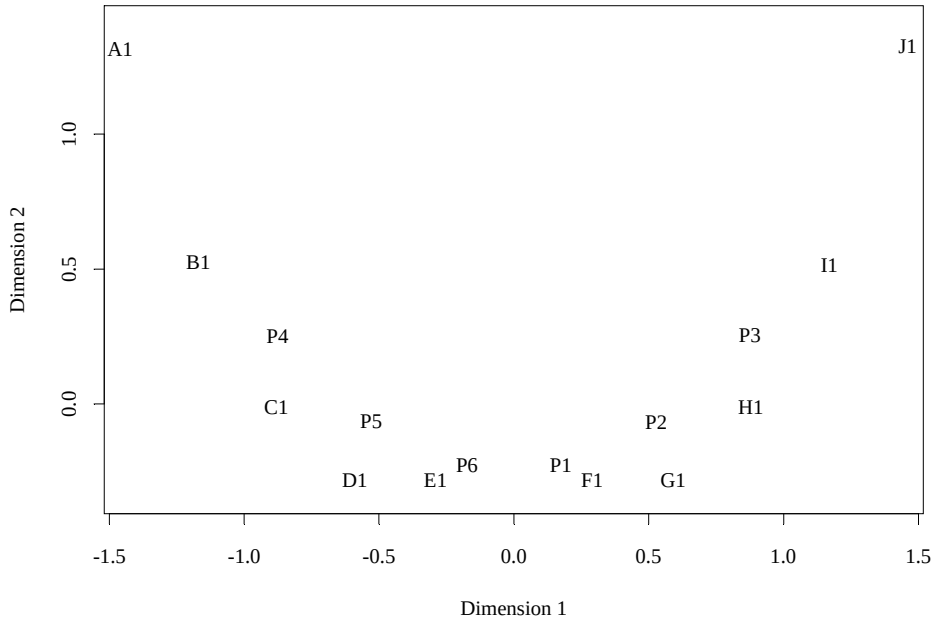
positiv					negativ					Score1	Score2	ID
A	B	C	D	E	F	G	H	I	J			
2	2	2	2	2	1	1	1	1	1	-1,464	1,331	P4
1	2	2	2	2	2	1	1	1	1	-,878	-,260	P5
1	1	2	2	2	2	2	1	1	1	-,293	-1,065	P6
1	1	1	2	2	2	2	2	1	1	,293	-1,073	P1
1	1	1	1	2	2	2	2	2	1	,878	-,274	P2
1	1	1	1	1	2	2	2	2	2	1,464	1,341	P3

"Fehlerfrei" kann die Ordnung der Personen (P4, P5, P6, P1, P2, P3) an der Parabel im 2-dimensionalen Raum abgelesen werden (Abb. 1). Die Präferenzordnung der Items für jede Beobachtung kann durch Projektion der Itemkategorie 1 (gewählt) auf den Beobachtungsvektor gefunden werden (***Davison*** 1977). Die "Personen" (Zeilen der Datenmatrix) kommen immer zwischen den Kategorien 1 (Stimuluspräferenz) der gewählten Items zu liegen. Genau das ist aus Abb. 1 anschaulich zu entnehmen. Aus Platzgründen kann hier das Problem der Rekonstruktion von Ordnungsabhängigkeiten nicht näher erläutert werden. Hierzu wird auf ***Schriever*** (1983, 1985) verwiesen.

Heiser (1981) und ***Cliff et al.*** (1988) zeigte, daß durch Homogenitätsanalyse die oben dargestellte Parallelogrammstruktur abgebildet werden kann (vgl. auch ***van Rijckevorsel*** 1987, S. 72f). Eine einleuchtende Interpretation liefert die Beziehung zwischen dem Vektormodell von ***Tucker*** (***Tucker*** and ***Messick*** 1963) und dem Unfoldingmodell (***Coombs*** 1975). ***Van Schuur*** (1994, S. 100f) faßt zusammen:

> "An efficient set of items can be represented along an arc that is interpretable as an unidimensional subspace in the two-dimensional representation. The point, at which the preference vector intersects the arc can be interpreted as the persons ideal point on the arc, and his or her preference can be specified by unfolding the arc at the ideal point. The unidimensional unfolding representation of a set of items is therefor equivalent to the two-dimensional vector representation of that (efficient) set of items."

Abbildung 1: Gemeinsame Abbildung von Personen und Variablen aus Tabelle 3



Von Bedeutung für das hier behandelte Problem ist die Tatsache, daß die Projektion von Zeilen und Spalten der Datenmatrix von den entsprechenden Randsummen abhängt und damit auch von der Zahl der Zustimmungen eines Befragten. Bis auf Spezialfälle wie in Tabelle 3 (wobei die Zahl k beliebig sein kann, aber für alle Befragten in gleicher Weise gültig zu sein hat) bedingen sich diese Häufigkeiten in gewisser Weise: Befragte mit relativ vielen Zustimmungen ”überlappen” sich auf der Dimension und erzeugen durch ihr Antwortverhalten größere Zustimmungssummen für die entsprechenden Items. Es wir weiter unten gezeigt werden, daß genau die Häufigkeit auf einer der artifiziellen Achsen ablesbar ist.

2.3.2 Problem ”wähle k aus 10”

In den allermeisten Fällen wird aber das Konstrukt ohne die obengenannte strenge Auswahlvorschrift operationalisiert. Die Zahl der Zustimmungen ist dem Respondenten ins Belieben gestellt. Dadurch erhalten wir allerdings unterschiedliche Zustimmungshäufigkeiten für die Befragten, womit die Frage nach der Bedeutung dieser Größe relevant wird. Wie oben dargestellt ist diese Summe (auch wenn alle Items gleichsinnig gepolt werden) nicht notwendig eine Statistik für den Ort der Befragten auf der Dimension. Wir wollen zeigen, wie dichotome Daten dieses Typs sowohl durch Homogenitätsanalyse wie auch durch Unfoldinganalyse auf ihre Struktur hin untersuchen werden können. Tabelle 4 zeigt einen weiteren künstlichen Datensatz, in dem sich die ”Befragten” durch unterschiedlich viele Zustimmungen auszeichnen.

Tabelle 4: Projektion des Problems “wähle k aus 10”

	positiv					negativ					Score1	Score2	Summe der Zustimm.
	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10			
P1	2	1	1	1	1	1	1	1	1	1	,835	-,001	1
P2	2	2	1	1	1	1	1	1	1	1	1,099	,621	2
P3	1	2	2	2	1	1	1	1	1	1	,483	1,942	3
P4	1	1	2	2	2	2	1	1	1	1	-,780	1,506	4
P5	1	1	1	2	2	2	2	2	1	1	-1,618	,019	5
P6	1	1	1	1	2	2	2	2	1	1	-1,374	-,620	4
P7	1	1	1	1	1	2	2	2	1	1	-,894	-,787	3
P8	1	1	1	1	1	1	1	2	2	1	,357	-1,069	2
P9	1	1	1	1	1	1	1	1	2	2	1,066	-1,064	2
P10	1	1	1	1	1	1	1	1	1	2	,828	-,546	1

Gleichzeitig wird angenommen, daß die Respondenten an den extremen Enden einer Präferenzordnung ”präziser” antworten, d.h. ihre Lage auf der für Items und Personen gemeinsamen Skala durch weniger Zustimmungen determiniert ist für Personen, die sich in der Mitte dieser Dimension befinden und daher eine eher neutrale Haltung einnehmen. Hier wird die Häufigkeit der Zustimmungen (unabhängig von der Polung) wieder zu einem substantiell bedeutsamen Indikator, von dem weiter unten ausführlich die Rede sein wird, weil sich in diesem Zusammenhang nicht die Interpretation als ”Intensität”, sondern auch die einer ”Inkonsistenzachse” (***Rostampour*** 1998) anbietet.

2.3.2.1 Homogenitätsanalyse des Problems ”wähle k aus 10”

Die Reduktion auf 2 Dimensionen zeigt ein zunächst verwirrendes Ergebnis (Abb.2). Die Projektion auf die Achsen entspricht nicht der erwarteten Reihenfolge sowohl der Items wie auch der ”Befragten”. Betrachtet man allerdings die Anordnung im Hinblick auf beide Achsen, so ist die bekannte Ordnung unschwer zu erkennen: sowohl die Itemkategorien wie auch die ”Befragten” sind auf einem Kreis angeordnet. Man benötigt offenbar 2 Koordinaten zur Repräsentation der (hier schon bekannten) Ordnung. Betrachtet man die Projektion auf die 2. Achse im rechten Teil von Abb. 3, so sind - bis auf 1,2 und 10 - alle Beobachtungen in der erwarteten Reihenfolge angeordnet. Gerade die ”extremen” Beobachtungen werden eher im Zentrum abgebildet. Dieses Problem wird daher zu einem der wesentlichen Punkte für den Vergleich der beiden Abbildungsverfahren. Abgesehen von den Extrempunkten ist die zur Debatte stehende Dimension auf der 2. Achse abgebildet und es stellt sich die Frage nach der Bedeutung der 1. Achse. Vergleicht man die ”Eigenschaften” sowohl der Zeilen wie auch der Spalten (”Variablen”), die an ähnlichen Stellen der Projektion auf die 1.Achse liegen, so findet man die Häufigkeit der Zustimmungen (Kode 2) als hier einziges Kriterium der Ähnlichkeit.

Abbildung 2: Quantifikation von Zeilen und Spalten der Daten aus Tabelle 4[3]

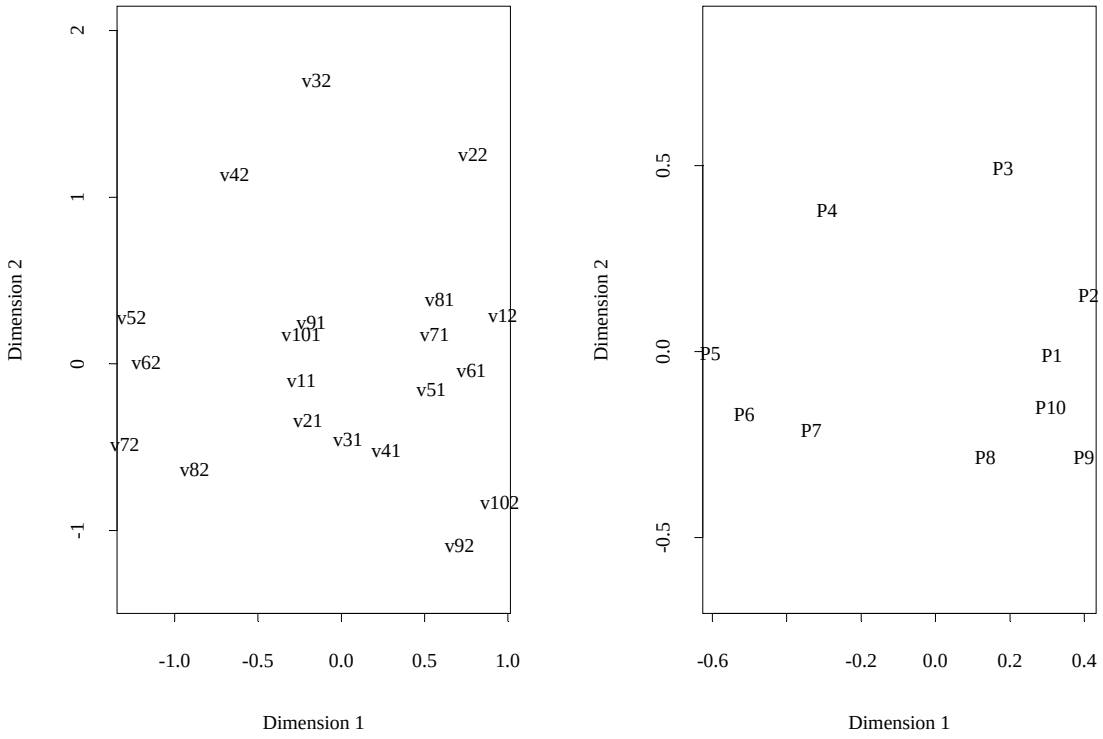


Abb. 3 zeigt diesen (geradezu linearen!) Zusammenhang ganz deutlich. Der Zusammenhang zwischen der 1. Achse und der Summe an Zustimmungen (unabhängig von der Itempolung) ist systematisch zu beobachten.

2.3.2.2 Unfoldinganalyse des Problems ”wähle k aus 10”

Die Ergebnisse des multiplen unidimensionalen Unfolding der Daten aus Tab. 4 zeigen, daß diese als perfekte Unfoldingskala anzusehen sind. Die Diagnosekriterien in Tab. 5 bestätigen diesen Sachverhalt unabhängig vom perfekten Gesamt-H (1.0) (Zur Interpretation der Diagnosematrizen siehe 2.2.2). Sowohl Beobachtungen, wie auch die 10 ”Variablen” sind in der durch die Konstruktion des Datensatzes vorbestimmten Reihenfolge geordnet und es interessiert der Zusammenhang zwischen den Achsen der Homogenitätsanalyse und der Ordnung der Beobachtungen auf der Unfoldingskala. Dabei steht der schon festgestellte Zusammenhang zwischen einer der Achsen (hier der 1.) und der Zahl der Zustimmungen im Vordergrund. Es ist zu berücksichtigen, daß diese Summe keine Statistik für

3 Die Variablen sind mit 1= Ablehnung 2=Zustimmung kodiert

die Präferenz und damit die Position auf einer J-Skala ist. Gerade im vorliegenden Beispiel sind die Extrempositionen durch die gleiche Zahl an Zustimmungen (nämlich eine, bzw. zwei) ausgezeichnet.

Abbildung 3: 1. Achse (Score 1) und Zahl der Zustimmungen aus Tabelle 4[4]

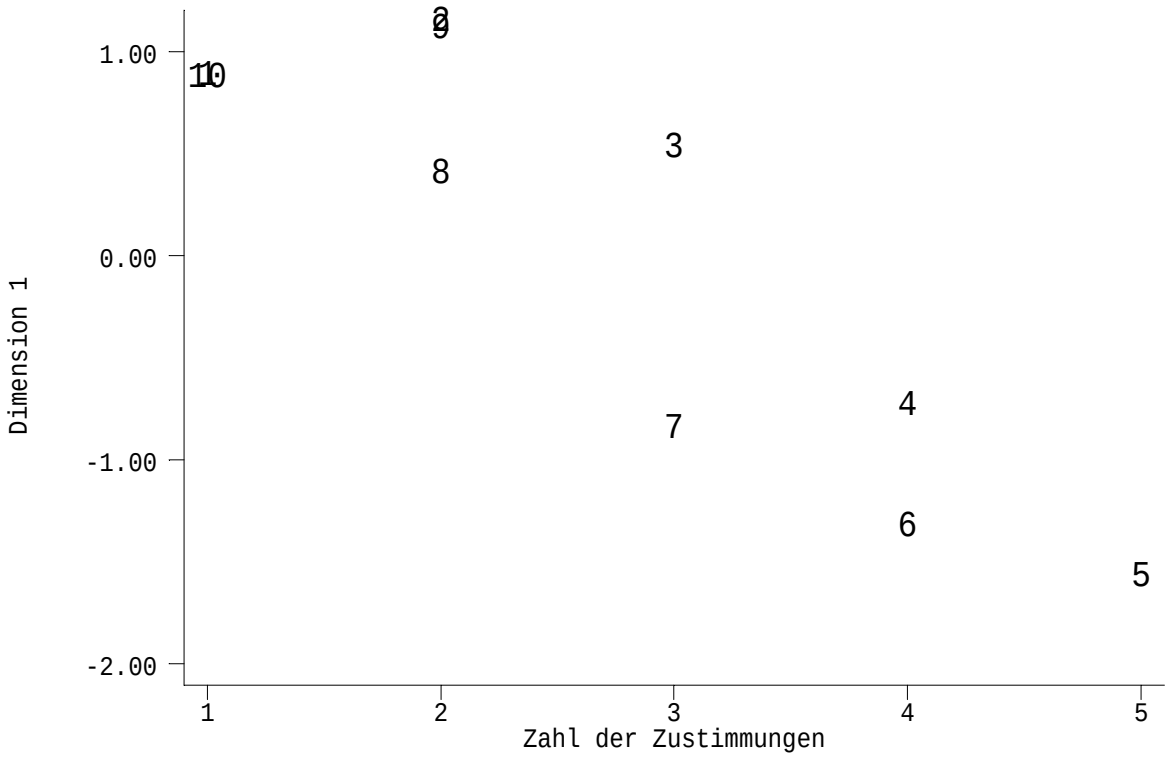


Tabelle 5: Unfolding des Problems “wähle k aus 10“

D O M I N A N C E M A T R I X 1

	A	B	C	D	E	F	G	H	I	J	
A) VAR1	-	10	20	20	20	20	20	20	20	20	**A**
B) VAR2	10	-	10	10	20	20	20	20	20	20	**B**
C) VAR3	20	10	-	0	10	10	20	20	20	20	**C**
D) VAR4	30	20	10	-	10	10	20	20	30	30	**D**
E) VAR5	30	30	20	10	-	0	10	10	30	30	**E**
F) VAR6	40	40	30	20	10	-	10	10	40	40	**F**
G) VAR7	30	30	30	20	10	0	-	0	30	30	**G**
H) VAR8	40	40	40	30	20	10	10	-	30	40	**H**
I) VAR9	20	20	20	20	20	20	20	10	-	10	**I**
J) VAR10	20	20	20	20	20	20	20	20	10	-	**J**
	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	

4 Identifikationszeichen = Zeilennummer

A D J A C E N C Y M A T R I X 1

	A	B	C	D	E	F	G	H	I	J	
A) VAR1	-										A
B) VAR2	10	-									B
C) VAR3	0	10	-								C
D) VAR4	0	10	20	-							D
E) VAR5	0	0	10	20	-						E
F) VAR6	0	0	10	20	30	-					F
G) VAR7	0	0	0	10	20	30	-				G
H) VAR8	0	0	0	10	20	30	30	-			H
I) VAR9	0	0	0	0	0	0	0	10	-		I
J) VAR10	0	0	0	0	0	0	0	0	10	-	J
	A	B	C	D	E	F	G	H	I	J	

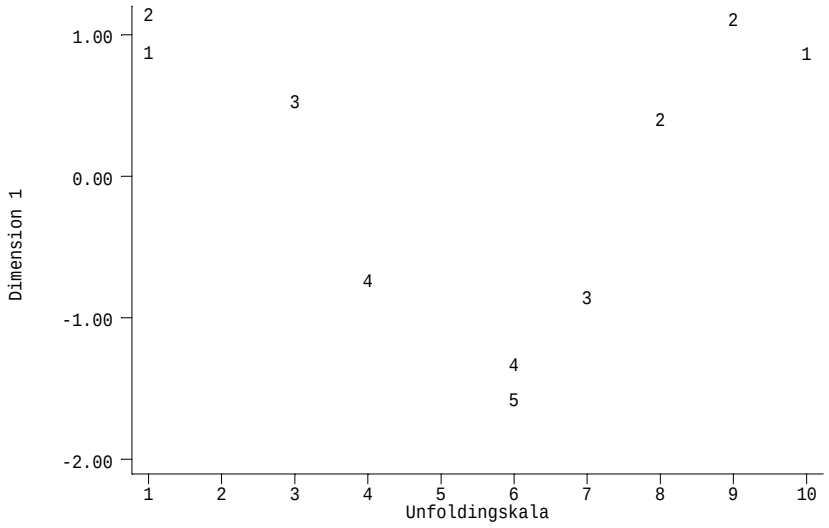
CONDITIONAL ADJACENCY MATRIX 1

	A	B	C	D	E	F	G	H	I	J	
A) VAR1	-	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	A
B) VAR2	0.50	-	0.50	0.33	0.00	0.00	0.00	0.00	0.00	0.00	B
C) VAR3	0.00	0.50	-	0.67	0.33	0.25	0.00	0.00	0.00	0.00	C
D) VAR4	0.00	0.50	1.00	-	0.67	0.50	0.33	0.25	0.00	0.00	D
E) VAR5	0.00	0.00	0.50	0.67	-	0.75	0.67	0.50	0.00	0.00	E
F) VAR6	0.00	0.00	0.50	0.67	1.00	-	1.00	0.75	0.00	0.00	F
G) VAR7	0.00	0.00	0.00	0.33	0.67	0.75	-	0.75	0.00	0.00	G
H) VAR8	0.00	0.00	0.00	0.33	0.67	0.75	1.00	-	0.50	0.00	H
I) VAR9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.25	-	0.50	I
J) VAR10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50	-	J
	A	B	C	D	E	F	G	H	I	J	

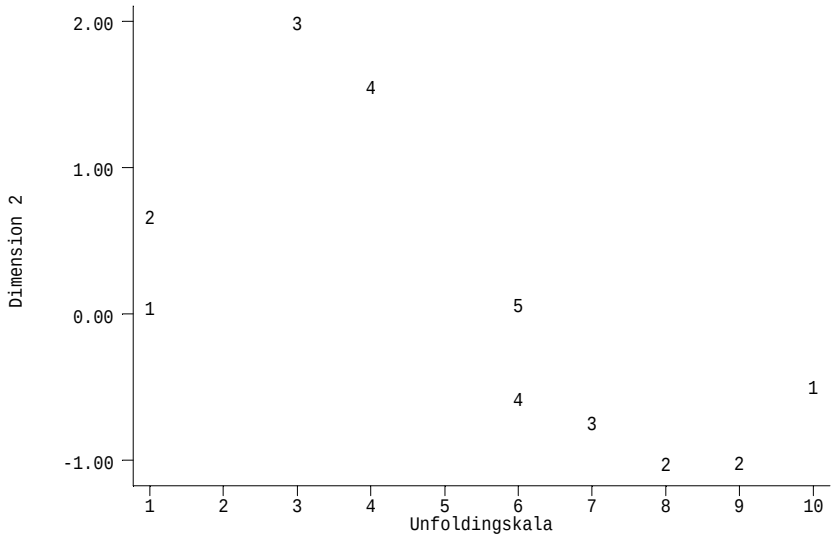
Die Abb. 4 zeigt, daß nur zwischen der 2. Achse und der Unfoldingskala ein nennenswerter (negativer) Zusammenhang besteht. Die monotone Beziehung gilt allerdings nicht für die Extrempositionen der Präferenzordnung, die an ähnlichen Stellen - und zwar eher im Neutralpunkt - der 2. Achse zu liegen kommen. Dies gibt Anlaß zur Vermutung, daß gerade die Extrempositionen der Präferenzordnung durch das Modell für Dominanzdaten inadäquat abgebildet werden. Dieses Phänomen wird uns bei der Analyse des Konstrukts ”Ausländerfeindlichkeit” wiederbegegnen. Der U-förmige Zusammenhang der Unfoldingskala mit der 1. Achse ist nur durch den Zusammenhang mit der Zahl der Zustimmungen bestimmt: Respondenten in der Mitte der Präferenzordnung sind durch mehr Zustimmungen ausgezeichnet als Respondenten mit extremerer Position.

Abbildung 4: Achsen der Homogenitätsanalyse und Unfolding der Daten aus Tabelle 4[5]

a)



b)



3 Strukturanalyse des Konstruktes ”Ausländerfeindlichkeit”

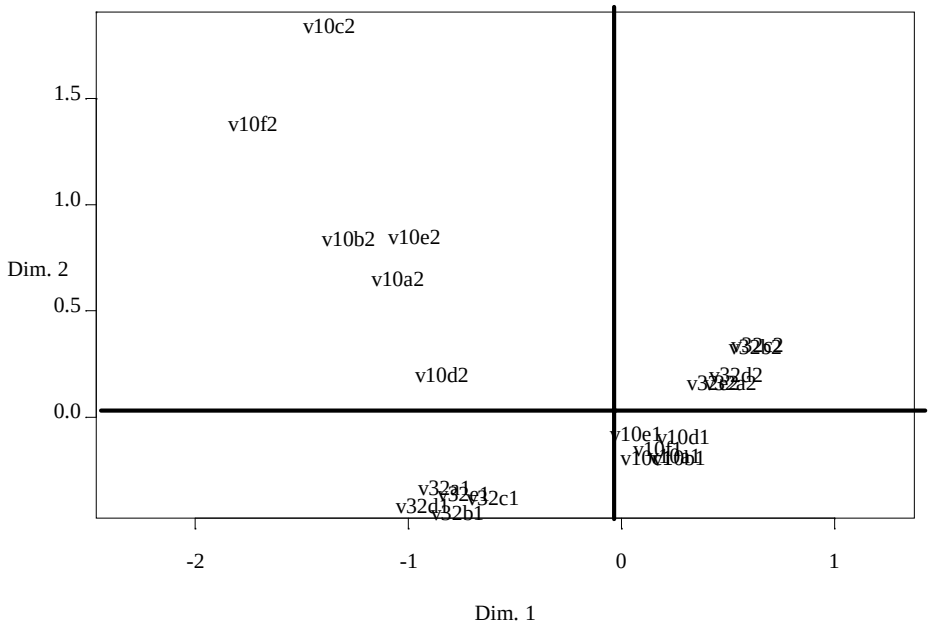
Mit den folgenden Analysen soll das Problem der Dimensionalität am eingangs dargestellten Beispiel der Messung von ”Ausländerfeindlichkeit” weiter beleuchtet werden. Von zentralem Interesse ist dabei die Frage, welches Abbildungsverfahren das eindimensional konzipierte und bipolar operationalisierte Konstrukt adäquat abbildet. Dazu werden die Daten dichotomisiert in ”Zustimmung“ (4 und 5=1) und ”Nicht-Zustimmung” (1,2 und 3= 0). Damit ist gewährleistet, daß für beide Zugänge dieselben Daten verwendet werden können.

5 Identifikationszeichen = Summe der Zustimmungen

3.1 Homogenitätsanalyse der ”Ausländerfeindlichkeit”

Die Projektion der 11 dichotomisierten Items in zwei Dimensionen (Abb. 5) zeigt das schon vermutete Ergebnis: in der oberen Hälfte befinden sich die Zustimmungen (2) zu allen Items, wobei die 2. Achse offensichtlich von den Zustimmungen zu den positiv gepolten Items dominiert wird. In der unteren Hälfte sind alle Kategorien 1 (Ablehnung) angesiedelt. Man sieht auch, daß durch nur geringe Rotation des Achsenkreuzes nach links 2 orthogonale Achsen gebildet werden können, die ausschließlich durch die Polung der Items

Abbildung 5: Homogenitätsanalyse des Konstruktes ”Ausländerfeindlichkeit”



erklärt werden können. Da hier die Struktur der Lösung im Zentrum des Interesses steht, wird auf die tabellarische Darstellung verzichtet (Die Tabellen können bei den Autoren angefordert werden). Die Eigenwerte der Lösung sind für die ersten beiden Dimensionen 0.317 und 0.120; die Annahme einer 2-dimensionalen Struktur also gerechtfertigt.

3.2 Unfoldinganalyse der ”Ausländerfeindlichkeit”

Es ist zu berücksichtigen, daß die Variablen V10c und V10e (vgl. Anhang) nicht in der Analyse aufscheinen, da sie für die Bildung einer J-Skala nicht brauchbar sind. Die Analyse liefert eine plausibel interpretierbare Präferenzordnung der Items von V10f = ”Die Ausländer nehmen den Deutschen die Jobs weg” bis zu V32c = ”Die Ausländer tragen zum Wohlstand unseres Landes bei” (siehe Tab. 6). Extreme Enden der J-Skala sind also Stimuli, die sich auf die ökonomische Situation beziehen. Die Mitte der Skala wird durch eher

unscharf formulierte Items bezeichnet, wie V10d = ”Ausländer sollten nicht die gleichen politischen Rechte haben wie die Deutschen”, bzw. V32e = ” Im allgemeinen kann man Ausländern trauen”. Im wesentlichen ist die Skala durch den ökonomischen Aspekt determiniert und es ist daher auch nicht verwunderlich, daß gerade
negativ formulierte Items ohne direkten Bezug dazu, wie V10c und V10e, nicht auf der J-Skala abgebildet werden können. Während Item V10e vermutlich wegen der negativen Formulierung problematisch ist, bringt Item V10c die Ablehnung einer multikulturellen Gesellschaft und damit eine diffuse Überfremdungsfurcht zum Ausdruck. Dieses Item paßt nicht zu einer inhaltlichen Dimension, die vorrangig den ökonomischen Aspekt der Anwesenheit von Ausländern in der Gesellschaft beinhaltet (vgl. auch die großen Ladungen dieser beiden Items auf dem 2. Faktor der Hauptkomponentenanalyse in Tab. 1). Anhand der Itemordnung ist auch die Ordnung der Respondenten auf der Skala festgelegt: **”Je größer der Score, desto ausländerfreundlicher”.**

Tabelle 6: Unfoldingskala der 9 verbleibenden Items des Konstrukts “Ausländerfeindlichkeit”

	P(I)	H(I)	OBS.	AND EXP. ERRORS	SD(H)	T(H)
F) V10F	0.09	0.60	199	501.0	0.0467	12.9139
B) V10B	0.17	0.53	387	828.4	0.0336	15.8395
A) V10A	0.20	0.52	427	888.0	0.0322	16.1463
D) V10D	0.26	0.48	484	939.4	0.0319	15.1828
K) V32E	0.63	0.52	503	1042.1	0.0277	18.6507
G) V32A	0.62	0.53	493	1042.0	0.0280	18.8306
J) V32D	0.64	0.54	483	1039.1	0.0291	18.4158
H) V32B	0.56	0.55	454	1019.0	0.0319	17.3831
I) V32C	0.49	0.54	431	938.8	0.0358	15.1103

NIND= 585
H= 0.53
T(H)= 25.95

Tabelle 7 enthält die Diagnosekriterien, die deutlich machen, daß die wesentlichsten Bedingungen für eine Unfoldingskala erfüllt sind.

Tabelle 8 enthält eins der wichtigsten Kriterien zur Beurteilung der Frage, ob wir es mit einer und nur einer Dimension im Sinne des unidimensionalen Unfolding, oder doch mit 2 horizontal aneinandergehängten kumulativen Skalen zu tun haben. Die bedingte Adjacencymatrix kann als grobe Schätzung der ”trace-line” (der charakteristischen Kurve) des Zeilenstimulus interpretiert werden und dient auch zu Entscheidung zwischen kumulativen und Ähnlichkeitsdaten (***van Schuur*** and ***Post***. 1990 : Kap. 8). Man sieht aus dieser Tabelle, daß die Annahme unimodaler tracelines eigentlich nicht gerechtfertigt ist (vgl. hier auch das 2. Beispiel mit künstlichen Daten).

Tabelle 7: Diagnosekriterien (Dominanzmatrix und Adjacencymatrix)

D O M I N A N C E M A T R I X 1

	F	B	A	D	K	G	J	H	I	
F) V10F	-	4	4	4	7	**6**	**6**	7	7	**F**
B) V10B	12	-	9	8	10	11	11	12	13	**B**
A) V10A	15	11	-	9	11	12	12	13	14	**A**
D) V10D	21	17	16	-	13	15	16	17	18	**D**
K) V32E	61	56	54	50	-	16	15	19	25	**K**
G) V32A	60	56	55	51	15	-	13	17	23	**G**
J) V32D	61	58	56	54	15	15	-	16	21	**J**
H) V32B	54	51	49	46	11	10	8	-	17	**H**
I) V32C	47	45	43	41	11	10	6	**10**	-	**I**
	F	**B**	**A**	**D**	**K**	**G**	**J**	**H**	**I**	

A D J A C E N C Y M A T R I X 1

	F	B	A	D	K	G	J	H	I	
F) V10F	-									**F**
B) V10B	4	-								**B**
A) V10A	4	8	-							**A**
D) V10D	5	8	10	-						**D**
K) V32E	2	7	9	13	-					**K**
G) V32A	2	5	7	11	47	-				**G**
J) V32D	2	5	7	10	48	48	-			**J**
H) V32B	1	4	6	9	44	45	47	-		**H**
I) V32C	1	4	5	8	38	38	42	**38**	-	**I**
	F	**B**	**A**	**D**	**K**	**G**	**J**	**H**	**I**	

Tabelle 8: Diagnosekriterien (Bedingte Adjacencymatrix)

CONDITIONAL ADJACENCY MATRIX 1

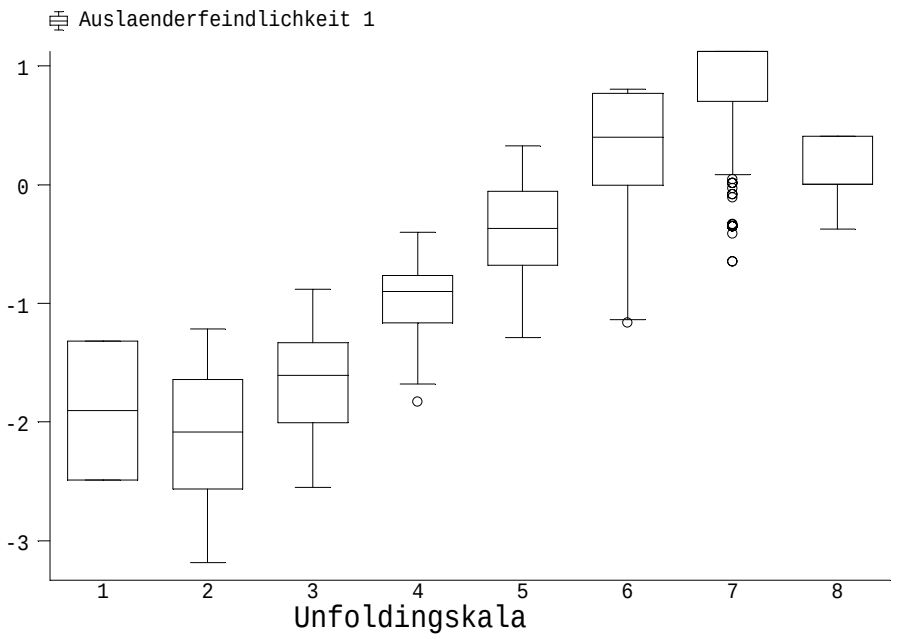
	F	B	A	D	K	G	J	H	I	
F) V10F	-	0.29	0.23	0.20	0.03	0.04	0.04	0.03	0.03	**F**
B) V10B	0.56	-	0.43	0.34	0.11	0.10	0.09	0.09	0.09	**B**
A) V10A	0.50	0.49	-	0.40	0.14	0.12	0.12	0.12	0.12	**A**
D) V10D	0.58	0.51	0.53	-	0.21	0.18	0.16	0.17	0.17	**D**
K) V32E	0.23	0.42	0.46	0.51	-	0.77	0.76	0.80	0.78	**K**
G) V32A	0.27	0.35	0.37	0.44	0.75	-	0.77	0.82	0.79	**G**
J) V32D	0.27	0.35	0.37	0.39	0.76	0.79	-	0.85	0.87	**J**
H) V32B	0.21	0.28	0.33	0.36	0.70	0.73	0.74	-	0.79	**H**
I) V32C	0.19	0.25	0.30	0.32	0.60	0.62	0.67	0.70	-	**I**
	F	**B**	**A**	**D**	**K**	**G**	**J**	**H**	**I**	

Bis auf kleine Unregelmäßigkeiten sind die relativen Häufigkeiten immer monoton verteilt, wobei sie für die negativ formulierten Items von links nach rechts monoton abnehmen und für die positiven Items von rechts nach links monoton ansteigen.

3.3 Vergleich zwischen Unfolding und Homogenitätsanalyse

Da die Zahl der Zustimmungen unabhängig von der Polung offenbar eine zentrale Rolle für die Interpretation der Skalen bildet, sollen hier Hauptkomponentenanalyse, Unfoldingskala und diese Größe simultan betrachtet werden. Zu diesem Zweck wird zunächst die 1. Achse der PCA und die Unfoldingskala gegeneinander geplottet (Abb. 6), wobei zur besseren Darstellung des Zusammenhanges Boxplots für jede der 8 Klassen der Unfoldingskala dienen.

Abbildung 6: Dimension 1 der Homogenitätsanalyse gegen die Unfoldingskala



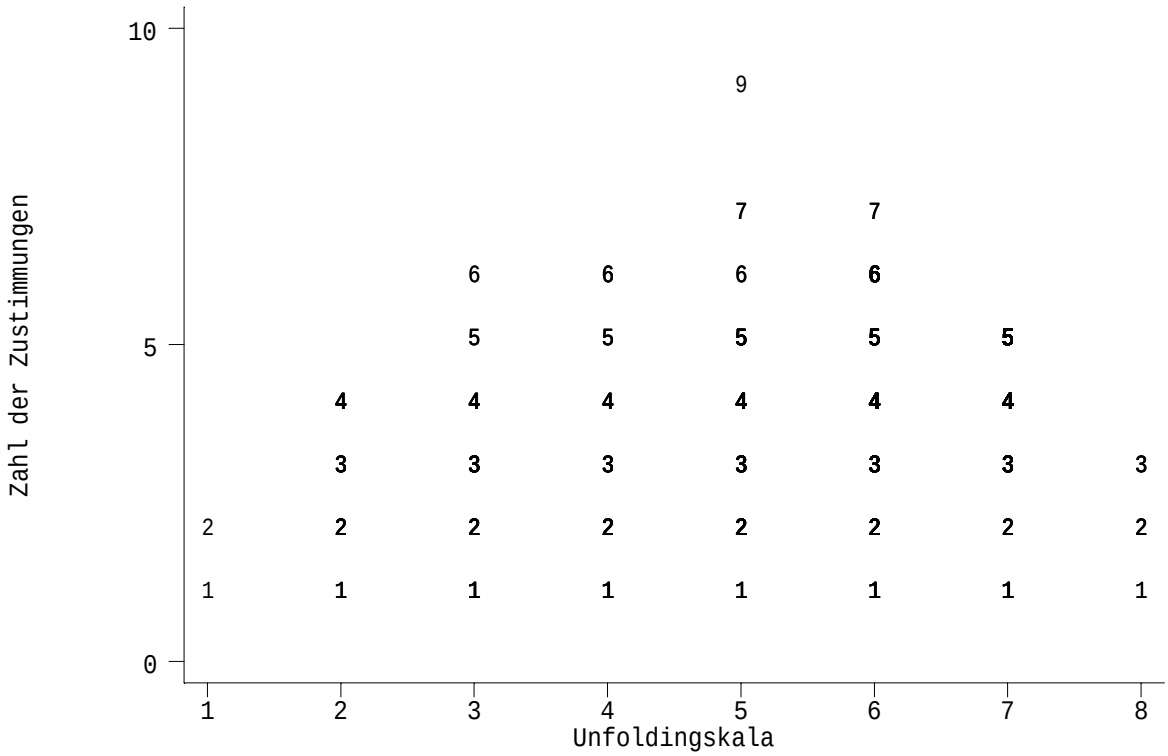
Korreliert man die 1. Achse der Ausländerfeindlichkeit mit der Unfoldingskala, so zeigt sich ein Wert von 0.84; die Graphik zeigt aber einen Zusammenhang in Form einer logistischen Funktion. Dies ist schon für den künstlichen Datensatz zu beobachten gewesen (vgl. Abb. 4a). Man sieht, daß Respondenten mit extremer Lage auf der Unfoldingskala an ähnlicher Stelle der 1. Achse der Hauptkomponentenanalyse zu liegen kommen. Die Mittelwerte der Unfoldingklassen 2-6 liegen aber nahezu auf einer Geraden, so daß im Mittelbereich von einer linearen Beziehung gesprochen werden kann. Die Korrelation der 2. Achse mit der Rangordnung der J-Skala ist gering (**-0.2103**) und es liegt die Vermutung nahe, daß hier wieder eine eindimensionale Ordnung nur in 2 Dimensionen repräsentiert werden kann, wie es schon an den künstlichen Daten gezeigt werden konnte, womit die Annahme, daß die Daten aus einem ”ideal point process” stammen, unterstützt wird. Die Verteilung

selbst liefert aber nicht genügend Hinweise für die Interpretation der 2. Achse. Dazu muß auch die Verteilung der Zahl der Zustimmungen im gleichen Raum betrachtet werden. Dies soll anschaulich durch Scatterplots geschehen, wobei die Zahl der Zustimmungen - nach der Häufigkeit gewichtet - als Zeichen im Plot verwendet wird. Die Gewichtung wird durch die Größe des Zeichens im Plot ausgedrückt (siehe ***Schnell***. 1994 : Kap. 5).

3.4 Unfoldingskala, 2. Achse der Faktoranalyse und Zahl der Zustimmungen

Es existiert ein schwacher nicht-linearer Zusammenhang zwischen der Zahl an Zustimmungen (unabhängig von der Polung) und der Unfoldingskala (Abb. 7). Es ist einleuchtend, das sich an den extremen Enden der Skala ”präzisere” Antwortmuster finden, wobei die Präzision durch die Zahl der Zustimmungen abgebildet wird. In der Mitte der Skala, d.h. im Neutralpunkt der Präferenzordnung, finden sich die unterschiedlichsten Muster mit nahezu jeder möglichen Zahl an Zustimmungen. Im Sinne einer Präferenzordnung sind hier vorwiegend sehr unpräzise antwortende Befragte angesiedelt.

Abbildung 7: Zahl der Zustimmungen in den Klassen der Unfoldingskala



Erst der Zusammenhang aller 3 Kenngrößen liefert letztlich die Kriterien für die Vermutung, daß die 2. Achse nicht als Inkonsistenzachse zu interpretieren ist. Wie schon ausgeführt sind an den ”Rändern” der Unfoldingskala die präziser Antwortenden (d.h. weniger Items zustimmenden) Respondenten zu finden. Die Mitte (3 und 4) wird durch Respon-

denten gekennzeichnet, deren Präferenzbreite wesentlich größer ist (Abb. 8). Man findet hier aber auch Befragte, die mit großer Präzision auf der Unfoldingskala zu verorten sind (eine oder nur 2 Zustimmungen), wobei wieder der Gradient über die 2. Achse der PCA sichtbar wird: Je mehr Zustimmungen, desto größer wird der entsprechende Objektscore (Abb. 9). Die 2. Achse bildet also so etwas wie die Präzision eines Respondenten bei der Einordnung in eine J-Skala ab, nicht aber eine Zustimmungs- oder Ablehnungstendenz (vgl. Abb. 8). Dieses Ergebnis verweist wiederum auf die zu Anfang dieser Arbeit erwähnte Interpretation dieser Achse als Intensitätsachse (***Foa*** 1950; ***Guttman*** 1947, 1954a, 1954b; ***Guttman*** and ***Suchman*** 1947) Diese Interpretation ist sinnvoll, wenn die Zahl der Zustimmungen nicht im kumulativen Sinne, sondern im Sinne der Abbildungs-(Antwort)präzision verstanden wird. Gerade der Begriff der Abbildungspräzision verweist aber auf ein Problem, das im nächsten Abschnitt behandelt werden muß.

Abbildung 8: Dimension 2 der Homogenitätsanalyse gegen die Unfoldingskala[6]

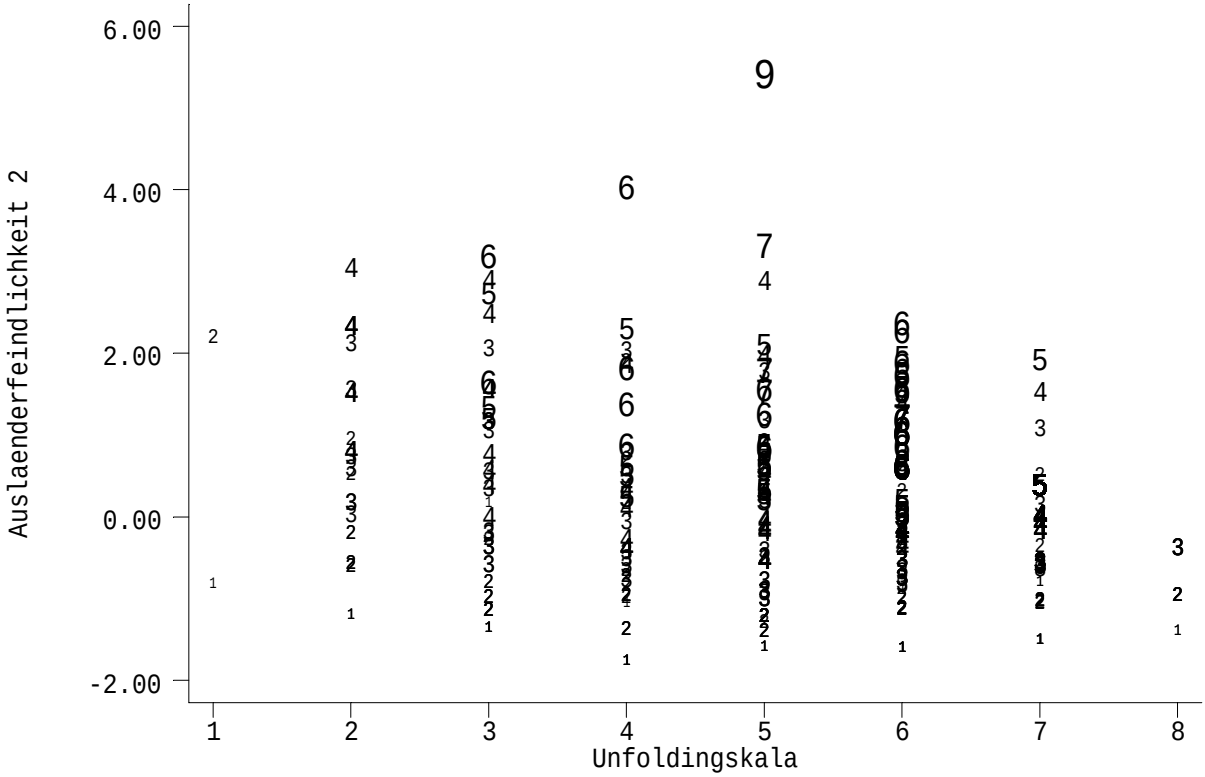
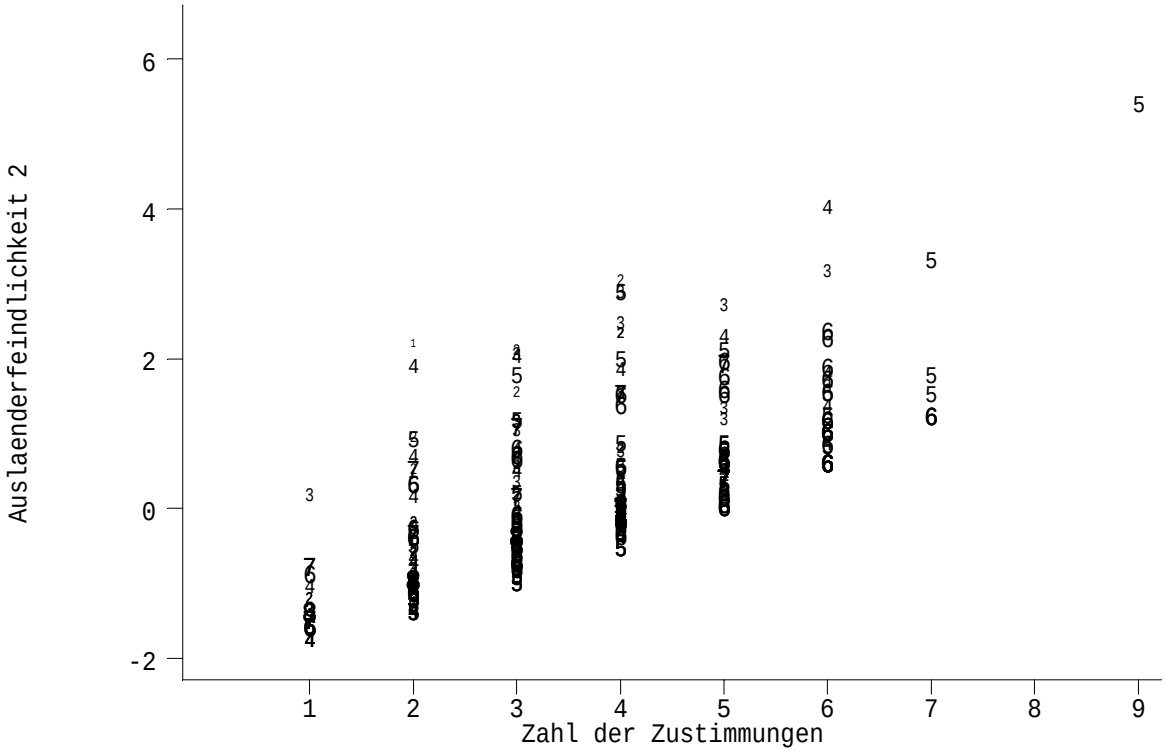


Abb. 9 zeigt den nahezu linearen Zusammenhang zwischen 2. Achse und der Zahl der Zustimmungen. An dieser Stelle sei noch einmal auf die künstlichen Beispiele von Tabelle 2 bzw. 3 zurückverwiesen, weil dort für die Ähnlichkeitsdaten entweder 2 oder 5 Zustimmungen für alle Beobachtungen invariant vorgegeben waren.

6 Identifikationszeichen = Zahl der Zustimmungen; Größe des Zeichens ist proportional zur Häufigkeit.

Abbildung 9: Dimension 2 der Homogenitätsanalyse gegen die Zahl der Zustimmungen[7]



3.5 Beurteilung der Präzision durch Inspektion der Antwortmuster

Die Interpretation der J-Skala kann durch die Inspektion der Antwortmuster bedeutend erleichtert werden. Die inhaltliche Beurteilung liefert dabei vor allem Hinweise, inwieweit die Abbildung des Konstruktes als Präferenzskala gerechtfertigt ist. Daher werden im folgenden für jede Stufe der Unfoldingskala ausgewählte Antwortmuster dargestellt (Tab. 9).

Zunächst erscheint es durchaus einsichtig, daß Respondenten mit extremer Position auf der Unfoldingskala auch extrem positionierten Items zustimmen. Die an diesen Stellen beobachtbaren Muster sind auch ähnlich, einfach weil die hier lokalisierten Befragten nur wenigen Items überhaupt zustimmen. In der Mitte der Skala sind naturgemäß Befragte mit unterschiedlicher Zahl an Zustimmungen zu verzeichnen. Beispielhaft finden wir einen Respondenten auf der 4. Stufe, der ausschließlich dem Item V10d zustimmt. Dieser liegt an der gleichen Stelle wie eine Person, die 6 Items zustimmt. Im Sinne des ”ideal point process” ist diese Anordnung notwendig, weil im ersten Falle ein äußerst ”präzises” Antwortverhalten vorliegt (die Person präferiert ein einziges Item!), wodurch sie an derselben Stelle wie der Stimulus verortet wird. Es lassen sich noch weitere (hier aus Platzgründen nicht aufgeführte) Muster finden, die zwar an gleichen Stellen der Unfoldingskala liegen,

7 Identifikationszeichen = Lage auf der Unfoldingskala; Größe des Zeichens ist proportional zur Häufigkeit.

aber nicht unbedingt zu einer ähnlichen inhaltlichen Interpretation veranlassen. Man kann sich daher berechtigt die Frage stellen, ob die gewonnenen Daten tatsächlich aus einem ”ideal point process” stammen und damit die Abbildung auf eine Unfoldingskala adäquat ist, oder ob die negativ und positiv gepolten Variablen im Sinne der Hauptkomponentenanalyse 2 (möglicherweise korrelierte) Dimensionen aufmachen.

Tabelle 9: Ausgewählte Antwortmuster auf der Präferenzskala (j-scale)

	V10f	**V10b**	**V10a**	**V10d**	**V32e**	**V32a**	**V32d**	**V32b**	**V32c**
1	**ja**	nein	nein	nein	nein	nein	nein	nein	nein
	ja	**ja**	nein	nein	nein	nein	nein	nein	nein
2	**ja**	**ja**	**ja**	**ja**	nein	nein	nein	nein	nein
	ja	nein	**ja**	nein	nein	nein	nein	nein	nein
3	nein	**ja**	nein	**ja**	**ja**	nein	nein	nein	nein
	nein	nein	**ja**	nein	nein	nein	nein	nein	nein
	ja	**ja**	**ja**	**ja**	nein	ja	**ja**	nein	nein
4	**ja**	**ja**	nein	nein	**ja**	**ja**	**ja**	**ja**	nein
	nein	nein	nein	**ja**	nein	nein	nein	nein	nein
5	nein	**ja**	**ja**	nein	**ja**	**ja**	**ja**	**ja**	nein
	nein	nein	nein	nein	**ja**	**ja**	nein	nein	nein
6	nein	nein	**ja**	nein	**ja**	**ja**	**ja**	**ja**	**ja**
	nein	nein	nein	nein	**ja**	**ja**	**ja**	**ja**	nein
7	nein	nein	nein	nein	**ja**	**ja**	**ja**	**ja**	**ja**
	nein	nein	nein	nein	nein	**ja**	**ja**	nein	**ja**
8	nein	nein	nein	nein	nein	nein	nein	**ja**	nein
	nein	nein	nein	nein	nein	nein	**ja**	**ja**	**ja**

4 Diskussion

In den vorangegangenen Abschnitten sind für entgegengesetzt gepolte Items die Ergebnisse von Projektionsverfahren unter der Annahme monotoner sowie eingipfeliger Item-Charakteristiken ausführlich dargestellt und verglichen worden. Im Vordergrund stand das viel diskutierte Phänomen der „überzähligen“ Dimension. „Überzählig“ heißt dabei, daß das üblicherweise verwendete Projektionsverfahren (Hauptkomponentenanalyse) nicht die gewünschte eindimensionale Struktur liefert, sondern wenigstens eine (statistisch relevante) Achse mehr. Diese Achse verleitet zu verschiedenartigsten Interpretationen, weil die Ladungen auf der „überzähligen“ Achse das Vorzeichen der entgegengesetzt gepolten Items nicht berücksichtigen. Wird eine Rotation (meist Varimax) vorgenommen, so zeigt sich fast immer, daß die beiden - entgegengesetzt gepolten - Itemgruppen auf unabhängigen Dimensionen abgebildet werden. Dieses Phänomen widerspricht vollständig den theoretischen Annahmen über die Eindimensionalität des bipolaren Konstrukts und wird daher als höchst unerwünschtes Resultat betrachtet.

Es konnte gezeigt werden, daß die Mehrdimensionalität der Abbildung das Resultat eines ganz spezifischen Datengenerierungsprozesses (Präferenz von Items, die der eigenen Position ähnlich sind) ist, der sowohl bei Auswahldaten als auch bei entgegengesetzt gepolten Items zum Tragen kommt und zu Ähnlichkeitsdaten im Gegensatz zu Dominanzdaten führt. Bei Ähnlichkeitsdaten ist der Einsatz von Unfolding-Modellen im Gegensatz zu Hauptkomponenten-Modellen wünschenswert. An einem Beispiel mit künstlichen Daten konnte gezeigt werden, daß die durch entgegengesetzt gepolte Items entstehende Parallelogramm-Struktur der Daten durchaus aus zwei aneinander gehängten, umgekehrten „Dreiecken“ bestehen kann, d.h. daß die jeweils gleichsinnig gepolten Variablen in sich als Dominanzdatum angesehen werden können. Wie die diagnostischen Kriterien zur Beurteilung der Faltungsskala zeigten, mußte diese Annahme auch im empirischen Beispiel beibehalten werden. Die charakteristischen Kurven aller Items sind monoton, wenngleich in Abhängigkeit von der Itempolung entgegengesetzt gerichtet. Dieser Befund widerspricht jedoch der Annahme, daß die Daten aus einem „ideal point process“ stammen und der Einsatz eines „Unfoldingmodells“ zur eindimensionalen Abbildung gerechtfertigt ist. Mit der Akzeptanz eines derartigen Modells steht oder fällt aber auch die Interpretationsmöglichkeit für die beiden Achsen einer Hauptkomponentenanalyse.

Durch den Vergleich zwischen Unfoldinganalyse und Hauptkomponentenanalyse ließ sich zeigen, daß die Präzision des Antwortverhaltens (die Zahl der Zustimmungen unabhängig von der Polung der Items) auf eine ganz bestimmte Weise durch die Hauptkomponentenanalyse repräsentiert ist. Die zweite Achse der Hauptkomponentenanalyse ist eindeutig mit der Zahl der Zustimmungen verknüpft. Es ist einleuchtend, daß dieser Zusammenhang von der Variation der Zahl der Zustimmungen abhängt. Daher war dieser Zusammenhang für das künstliche Beispiel in Tabelle 3 (pick 5/10) auch nicht beobachtbar, weil die Zahl der Zustimmungen in diesem Beispiel durch die Anweisung festgelegt ist. Des weiteren ist die Repräsentation der Beobachtungen auf der ersten Achse der Hauptkomponentenanalyse und auf der Präferenzordnungsskala identisch, so daß sich unmittelbar die Frage nach dem tatsächlichen Unterschied der beiden Abbildungsverfahren stellen muß. Dieser Unterschied wird am zweiten künstlichen Datenbeispiel deutlich, in dem der Zusammenhang zwischen der Zahl der Zustimmungen und (in diesem Fall der ersten Achse) der Hauptkomponentenanalyse dargestellt wird. An diesem Beispiel kann die grundsätzlich inadäquate Abbildung der auf einer Unfolding-Skala an Extrempositionen befindlichen Antwortmuster durch die „überzählige“ Achse der Hauptkomponentenanalyse aufgezeigt werden.

Zusammenfassend können wir feststellen: Wenn das Unfoldingmodell gültig ist und man gleichzeitig den Zusammenhang zwischen der Zahl der Zustimmungen (unabhängig von der Polung) und einer der beiden Achsen (zumeist der zweiten Achse) der Hauptkomponentenanalyse beobachtet, dann ergibt sich, daß die „überzählige“ Achse keinesfalls die Inkonsistenz des Antwortverhaltens abbildet. Sie ist vielmehr (durchaus im Sinne ***Guttman's***)

als Intensitäts-, beziehungsweise Präzisionsachse zu bezeichnen. Es besteht ein linearer Zusammenhang zwischen der Präzisionsachse und der Zahl der Zustimmungen (in dem Sinne, daß mit zunehmender Zahl an Zustimmungen die Präzision geringer wird) (Abb. 9). Dagegen besteht ein nichtlinearer Zusammenhang zwischen der Faltungsskala und der Präzisionsachse, wobei aus Abb. 8 ersichtlich ist, daß die Zahl der Zustimmungen im mittleren Bereich der Faltungsskala am größten ist. Dieser nichtlineare Zusammenhang ist deshalb zu erwarten, weil Respondenten an extremen Enden der J-Skala normalerweise eine „intensivere“ Einstellung zum Gegenstand (niedriger Wert auf der Präzisionsachse=hohe Präzision) aufweisen und aus diesem Grunde einer geringeren Anzahl an Stimuli ihre Zustimmung geben, als Respondenten, die sich in der Mitte, also am „Neutralpunkt“ der Unfolding-Skala befinden. Damit ist impliziert, daß die Einstellungsdimension nicht für alle Respondenten die gleiche Bedeutung hat. Daß Respondenten im Neutralpunkt der J-Skala mit einer großen Zahl an Zustimmungen antworten, ist u.E. nicht als Inkonsistenz, sondern als Indifferenz anzusehen. Die vorgegebenen Stimuli sind für Respondenten, die im Mittelbereich der Skala mit vielen Zustimmungen antworten, „zu schwach“, um eine explizite Position im Sinne präziser Antworten zu beziehen. Im Sinne eines „ideal point“ Prozesses, der beim Datentypus der Präferenzdaten angenommen wird, weisen diese „neutralen“ Respondenten eine größere „Toleranzbreite“ gegenüber den vorgegebenen Stimuli (Items) auf als Respondenten, die an den Extrempunkten der Skala lokalisiert sind und nur ganz wenigen Items zustimmen können. Zur Verdeutlichung dieser Interpretation ist hier nochmals Abb. 8 heranzuziehen: die Größe des Codes spiegelt die Häufigkeit, mit der eine bestimmte Anzahl an Zustimmungen auftritt. Die „9“ in Abb. 8 steht für häufiges Auftreten von 9 zustimmenden Antworten und hat gleichzeitig einen hohen Wert auf der „überzähligen“ Achse der Hauptkomponentenanalyse (hoher Wert=geringe Intensität bzw. geringe Präzision). Damit handelt es sich bei Respondenten mit vielen Zustimmungen nicht um die neutralen Respondenten, da diese im mittleren Bereich der Unfoldingskala und gleichzeitig im unteren Bereich der „überzähligen“ Achse angesiedelt sind, was an der geringen Häufigkeit einer kleinen Zahl an Zustimmungen deutlich wird. Dagegen wird bei Respondenten, die eine große Zahl an Zustimmungen und gleichzeitig eine geringe Präzision (2. Achse) haben, im Sinne einer Einstellung nichts gemessen, was anschaulich aus Abb. 8 hervorgeht. **Wo jedoch nichts gemessen wird, kann auch keine Inkonsistenz festgestellt werden.** Angemessen ist hier lediglich die Feststellung, daß bei der Personengruppe mit großer Zahl an Zustimmungen und geringer Präzision das Antwortartefakt „Zustimmung“ gemessen wird. Dies aufzuzeigen war unser zentrales Anliegen.

Welche Schlußfolgerungen lassen sich nun für das vorgestellte empirische Beispiel ziehen? Die S-förmige Relation zwischen der ersten (inhaltlich relevanten) Achse der Hauptkomponentenanalyse und der Faltungsskala bleibt Gegenstand weiterführender Überlegungen, weil Respondenten mit präzisem Antwortverhalten am unteren sowie oberen Ende der Faltungsskala durch die erste Hauptkomponente an relativ ähnlichen Stellen abgebildet

werden (Abb. 7). Die Inspektion ausgewählter Antwortmuster an verschiedenen Stellen der Faltungsskala zeigt (Tab. 9), daß extrem positionierte Personen auch immer präzise antworten, d.h. durch wenige (mitunter eine einzige) Zustimmungen gekennzeichnet sind. Das bedeutet aber gleichzeitig, daß die Befragten durch ihre Zahl an Ablehnungen sehr ähnlich sind (ein Phänomen, das aus der für die Homogenitätsanalyse spezifischen Codierung (vgl. in Abschnitt 2.2.1 die Ausführungen über „complete coding“) resultiert). Diese Art der Ähnlichkeit ist jedoch für die Anordnung auf einer Präferenzskala nicht relevant.

Mit der vorliegenden Arbeit konnten wir zeigen, daß im Falle entgegengesetzt gepolter Items immer der Datenerzeugungsprozeß näher untersucht werden sollte, weil man sonst Gefahr läuft, die Item-Gruppen als unterschiedliche Dimensionen eines eindimensional konzipierten Konstrukts anzusehen, oder die zweite Achse einer unrotierten Hauptkomponentenanalyse inadäquat zu interpretieren. Diese Gefahr besteht gerade dann, wenn man sich den Datensatz als zwei horizontal aneinander gehängte kumulative Skalen (Dominanzdaten) vorzustellen hat. In diesem Falle hat man es zwar mit monotonen Item-Charakteristiken zu tun, die für jede Itemgruppe (positiv bzw. negativ formulierte Items) in eine andere Richtung weisen, doch scheint hier die Methode der Entfaltung und die Abbildung von Respondenten und Variablen auf einer gemeinsamen Skala die latente Dimension adäquat abzubilden. Weitere Überlegungen und Untersuchungen, vor allem formaler Natur, zu diesem Problem sind allerdings von Nöten.

Anhang

Ausländerfeindlichkeit			
	negativ formulierte Items		positiv formulierte Items
10 A	Ausländer kommen nur deshalb hierher, um von unserem System sozialer Sicherheit zu profitieren.	32 A	Wir sollten die Ausländer, die hierher kommen, um hier zu leben, freundlich willkommen heißen.
10 B	Wenn die Arbeitslosigkeit steigt, sollten die Ausländer wieder in ihre Heimat zurückgeschickt werden.	32 B	Auch wenn die Arbeitslosigkeit steigt, sollten die Ausländer hierbleiben können.
10 C	Die Ausländer sind eine Bedrohung für unsere Kultur und unsere Bräuche.	32 C	Die Ausländer tragen zum Wohlstand unseres Landes bei.
10 D	Die Ausländer sollten nicht dieselben politischen Rechte haben wie die Deutschen.	32 D	Die Anwesenheit von Ausländern bereichert unsere Gesellschaft.
10 E	Im allgemeinen kann man Ausländern nicht trauen.	32 E	Im allgemeinen kann man Ausländern trauen.
10 F	Die Ausländer nehmen den Deutschen die Jobs weg.		

Literaturliste

Bekker, P. and ***De Leeuw, J.*** (1988). Relations Between Variants of Non-Linear Pricipal Component Analysis. In ***J.L.A. Van Rijckevorsel*** and ***J. De Leeuw*** (Eds.), *Component and Correspondence Analysis.* (pp. 1-31). Chichester, New York Brisbane: Wiley.

Blalock, H.M. (1971). *Causal Models in the Social Sciences.* Chicago, IL: Aldine Press.

Bollen, K.A. (1984). Multiple Indicators: Internal Consistency or no Necessary Relationship ? *Quality and Quantity, 18:*377-385.

Bollen, K.A. and ***Lennox, R.*** (1991). Conventional Wisdom on Measurement: A Structural Equation Perspective. *Psychological Bulletin, 110:* (2). 305-314.

Cliff, N., Collins, L.M., Zatkin, J., Gallipeau, D. and ***McCormick, D.J.*** (1988). An Ordinal Scaling Method for Questionnaire and Other Ordinal Data. *Applied Psychological Measurement, 12:* (1). 83-97.

Coombs, C.H. (1964). *A Theory of Data.* New York, London, Sidney: Wiley.

Coombs, C.H. (1975). A Note on the Relation Between the Vector Model and the Unfolding Model for Preferences. *Psychometrika, 40:* (1). 115-116.

Coombs, C.H. and ***Kao, R.C.*** (1960). On a Connection Between Factor Analysis and Multidimensional Unfolding. *Psychometrika, 25:* (3). 219-231.

Coombs, C.H. and ***Smith, J.E.K.*** (1973). On the Detection of Structure in Attitudes and Developmental Processes. *Psychological Review, 80:* (5). 337-351.

Davison, M.L. (1977). On a Metric, Unidimensional Unfolding Model for Attitudinal and Developmental Data. *Psychometrika, 42:* (4). 523-548.

Eckart, C. and ***Young, G.*** (1936). The Approximation of One Matrix by Another of Lower Rank. *Psychometrika, 1:* (3). 211-218.

Foa, U.G. (1950). Scale and Intensity Analysis in Opinion Research. *International Journal of Opinion and Attitude Research, 4:*192-208.

Gifi, A. (1990). *Nonlinear Multivariate Analysis.* Chichester, New York, Brisbane: Wiley.

Greenacre, M.J. (1984). *Theory and Applications of Correspondence Analysis.* London, Orlando, SanDiego, New York u.a.O.: Academic Press.

Guttman, L. (1947). The Cornell Technique for Scale and Intensity Analysis. *Educational and Psychological Measurement, 7:*247-279.

Guttman, L. (1954a). A New Approach to Factor Analysis: The Radex. In ***P.F. Lazarsfeld*** (Ed.), *Mathematical Thinking in the Social Sciences.* (pp. 258-348). Glencoe, Illinois: The Free Press.

Guttman, L. (1954b). The Principal Components of Scalable Attitudes. In ***P.F. Lazarsfeld*** (Ed.), *Mathematical Thinking in the Social Sciences.* (pp. 216-257). Glencoe, Illinois: The Free Press.

Guttman, L. and ***Suchman, E.A.*** (1947). Intensity and a Zero Point for Attitude Analysis. *American Sociological Review, XII:*57-67.

Heiser, W.J. (, 1981). Unfolding Model of Proximity Data. University of Leiden;

Heiser, W.J. (1985). *Undesired Nonlinearities in Nonlinear Multivariate Analysis.* Leiden: University of Leiden; Departement of Datatheory.

Heiser, W.J. and ***Meulmann, J.J.*** (1994). Homogeneity Analysis: Exploring the Distribution of Variables and Their Nonlinear Relationships. In ***M. Greenacre*** and ***J. Blasius*** (Eds.), *Correspondence Analysis in the Social Sciences: Recent Developments and Applications.* (pp. 179-209). London San Diego u.a.O.: Academic Press.

Hill, M.O. (1973). Reciprocal Averaging: An Eigenvector Method of Ordination. *Journal of Ecology, 61:*237-251.

Hill, M.O. (1974). Correspondence Analysis: A Neglected Multivariate Method. *Applied Statistics, 23:* (3). 340-354.

Hoijtink, H. (1991). The Measurement of Latent Traits by Proximity Items. *Applied Psychological Measurement, 15:* (2). 153-169.

Hoijtink, H. and ***Molenaar, I.*** (1994). An Item Response Model With Single Peaked Item Characteristic Curves: The PARELLA Model. *Quality and Quantity, 28:* 99-116.

Horst, P. (1935). Measuring Complex Attitudes. *Journal of Social Psychology, 6:*369-374.

Hotelling, H. (1933). Analysis of a Complex of Statistical Variables into Principal Components. *Journal of Educational Psychology, 14:*417-441.

Kendall, D.G. (1971). Seriation from Abundance Matrices. In ***F.R. Hodson*** (Ed.), *Mathematics of the Archeological and Historical Sciences.* Edinburgh: Edinburgh University Press.

MacCallum, R.C. and ***Browne, M.W.*** (1993). The Use of Causal Indicators in Covariance Structure Models: Some Practical Issues. *Psychological Bulletin, 114:* (3). 533-541.

Nishisato, S. (1980). *Analysis of Categorical Data: Dual Scaling and its Applications.* Toronto, Canada: University of Toronto Press.

Nishisato, S. (1994). *Elements of Dual Scaling: An Introduction to Practical Data Analysis.* New Jersey: Lawrence Erlbaum Associates,Inc.

Nishisato, S. (1996). Gleaning in the Field of Dual Scaling. *Psychometrika, 61:* (4). 559-599.

Post, W.J. (1992). *Nonparametric Unfolding Models.* Leiden: DSWO Press.

Richardson, M.W. and ***Kuder, G.P.*** (1933). Making a Rating Scala that Measures. *Personnel Journal, 12:*36-40.

Roberts, J.S. and ***Laughlin, J.E.*** (1996). A Unidimensional Item Response Model for Unfolding Responses From a Graded Disagree-Agree Response Scale. *Applied Psychological Measurement, 20:* (3). 231-255.

Rost, J. (1996). *Testtheorie Testkonstruktion.* Bern: Hans Huber.

Rostampour P. (1998). Die Zustimmungs- und Ablehnungstendenzen im Zusammenhang mit der Itemrichtung: positiv vs. negativ formulierte Items. *ZA- Information, 42*: 148-169.

Schnell, R. (1994). *Graphisch gestützte Datenanalyse.* München Wien: Oldenbourg.

Schriever, B.F. (1983). Scaling of Order Dependent Categorical Variables with Correspondence Analysis. *International Statistical Review, 51:*225-238.

Schriever, B.F. (1985). Order Dependence. Free University of Amsterdam;

Sjöberg, L. and ***Perrson, E.*** (1979). The Measurement of Mood. *Scandinavian Journal of Psychology, 20:*1-18.

Thurstone, L.L. (1928). Attitudes Can Be Measured. *American Journal of Sociology, 33:*529-554.

Thurstone, L.L. (1931). The Measurement of Social Attitudes. Journal of Abnormal and Social Psychology, 26:249-269.

Tucker, L.R. and ***Messick, S.J.*** (1963). Individual Difference Model for Multidimensional Scaling. *Psychometrika, 28:*333-367.

van de Geer, J.P. (1993a). *Multivariate Analysis of Categorical Data: Applications.* Newbury Park London New Delhi: Sage Publications.

van de Geer, J.P. (1993b). *Multivariate Analysis of Categorical Data: Theory.* Newbury Park London New Delhi: Sage Publications.

van Rijckevorsel, J.L.A. (1987). *The Application of Fuzzy Coding and Horseshoes in Multiple Correspondence Analysis.* Leiden: DSWO Press.

van Schuur, W.H. (1993). Nonparametrical Unidimensional Unfolding for Multicategory Data. *Political Analysis, 4:*41-74.

van Schuur, W.H. (1997). Intrinsic and Extrinsic Work Values as a Single Unfolding Scale. In ***J. Rost*** and ***R. Langeheine*** (Eds.), *Applications of Latent Trait and Latent Class Models in the Social Sciences.* (pp. 163-171). Münster New York München Berlin: Waxmann.

van Schuur, W.H. and ***Kiers, A.L.H.*** (1994). Why Factor Analysis is Often the Incorrect Model for Analyzing Bipolar Concepts, and What Model to use Instead. *Applied Psychological Measurement, 18:* (2). 97-110.

van Schuur, W.H. and ***Kruijtbosch, M.*** (1995). Measuring Subjective Well- Beeing: Unfolding the Bradburn Affect Balance Scale. *Social Indicators Research, 36:*49-74.

van Schuur, W.H. and ***Post, W.J.*** (1990). *MUDFOLD User's Manual Version 2.0.* Groningen: iec ProGAMMA.

Statische und dynamische Evaluation von Prognosen

von René Weber [1]

Zusammenfassung

Es werden die grundlegenden Möglichkeiten zur Quantifizierung der Prognosegüte dargestellt. Für die Evaluation von horizontalen Prognoseverteilungen wird ein „Gütemaße-Mix“ empfohlen, der unterschiedliche Qualitäten von Prognosen abbilden kann und sich dadurch für differenzierte Prognosevergleiche in der Praxis eignet. Darüber hinaus können Prognosevergleiche informativer gestaltet werden, wenn Prognosen dynamisch evaluiert werden. Eine dynamische Evaluation erhält man, indem man den Prognosehorizont sukzessiv um eine Periode ausdehnt und nach jeder neu hinzukommenden Periode die Prognosegüte neu berechnet. Dadurch ist man in der Lage zu beurteilen, in welchen Zeiträumen ein Prognoseverfahren bessere Prognoseergebnisse erzielt als ein anderes. Anhand von Beispielen wird gezeigt, daß sich die Einschätzung der Güte einer Prognose auf Basis einer statischen und dynamischen Betrachtung unterscheiden kann. Für die statische und dynamische Berechnung der Gütemaße wird ein SAS/IML-Programm mitgeliefert.

Abstract

The fundamental methods to quantify the accuracy of forecasts are presented. A "mix of measures" for the evaluation of horizontal distributions of forecasts is recommended, which can indicate different characteristics of forecasts, and is therefore suitable for detailed comparisons of forecasts in practical applications. Evaluating forecasts dynamically makes comparisons more informative. This can be done by gradually expanding the forecast horizon by one period at a time, which yields a dynamic measure of forecast accuracy. It is thereby possible to assess in what periods one method of forecasting performs better than another. Examples are given to demonstrate that static and dynamic evaluations of forecast accuracy do differ. A SAS/IML-program is supplied that calculates the recommended "mix of measures" statically and dynamically.

1 ***René Weber*** ist Wissenschaftlicher Mitarbeiter am Institut für Psychologie (Lehrstuhl Professor ***Jürgen Bortz***). Fachgebiet Methodenlehre und Statistik. TU Berlin. Franklinstrasse 28/29 - Postfach FR3-8, 10587 Berlin. E-mail: R.Weber@TU-Berlin.De

Quantifizierung der Prognosegüte

Es gibt zwei Möglichkeiten, die Prognosegüte einer Methode zu quantifizieren - entweder auf Basis der vertikalen oder auf Basis der horizontalen Verteilung der Prognosewerte.

Bei der Betrachtung der vertikalen Prognoseverteilung leitet man in analytischer Form Prognoseintervalle (Konfidenzintervalle) her und schätzt diese im allgemeinen anhand der Kovarianzmatrix der Modellresiduen. Je kleiner das Prognoseintervall ist, desto sicherer ist die Prognose bzw. desto besser ist die Prognosegüte. Ist die analytische Herleitung komplex, können Prognoseintervalle mit Hilfe von „Resampling-Methoden“ approximiert werden. Die Güte von Ex-ante-Prognosen kann nur durch die Betrachtung der vertikalen Prognoseverteilung quantifiziert werden.

Die zweite Möglichkeit, die Prognosegüte einer Methode zu quantifizieren, ergibt sich aus der Berechnung eines Kennwertes der horizontalen Prognoseverteilung. Hierbei werden alle resultierenden Prognosewerte mit den tatsächlich beobachteten Werten (die nicht zur Modellidentifikation verwendet wurden!) verglichen und die Unterschiedlichkeit durch eine Maßzahl ausgedrückt. Da hierfür die tatsächlich beobachteten Werte bekannt sein müssen, eignet sich diese Art der Evaluation ausschließlich für eine Ex-post-Prognose. Wegen der einfacheren Berechnung und Interpretation im praktischen Einsatz konzentrieren sich die folgenden Erläuterungen auf die horizontale Prognoseverteilung, d.h. auf die Berechnung von Gütemaßen.

Faktoren der Prognosegüte

Betrachtet man die horizontale Prognoseverteilung, dann beeinflussen hauptsächlich die Faktoren Prognosehorizont, Beschaffenheit der Daten und das gewählte Prognosegütemaß die angegebene Genauigkeit von Prognosen.

Langfristige Prognosen sind in der Regel ungenauer als kurzfristige Prognosen, da die Zeitinvarianzannahme[2] bei einer langfristigen Prognose weniger wahrscheinlich ist als bei einer kurzfristigen Prognose. Unterscheiden sich die zu prognostizierenden Daten z. B. hinsichtlich der Varianz, der Saisonalität oder der Komplexität der Einflußfaktoren, so ist es plausibel, daß die Genauigkeit einer Prognose davon abhängt, wie gut ein Prognosemodell dazu in der Lage ist, die spezifischen Eigenschaften einer Datenbasis zu erfassen.

Ein Irrtum wäre es zu glauben, daß unterschiedliche Prognosegütemaße zu einer übereinstimmenden Beurteilung der verschiedenen Prognosetechniken führen. Fehlprognosen

2 Definition der Zeitinvarianzannahme: Annahme, daß sich die Bedingungen, unter denen ein Prognosemodell identifiziert wurde, im Prognosezeitraum nicht verändern.

können sich durch ein bestimmtes Ausmaß an systematischen und unsystematischen Abweichungs-, Niveau- und Wendepunktfehlern unterscheiden. Die meisten gebräuchlichen und einfach interpretierbaren Gütemaße bewerten jedoch in erster Linie Abweichungsfehler und nur indirekt Niveau- und Wendepunktfehler (diese induzieren Abweichungsfehler). Eine Betrachtung mehrerer Gütemaße, die möglichst unterschiedliche Merkmale aufweisen, um die verschiedenen Arten von Fehlprognosen sensitiv erfassen zu können, würde hier eine differenziertere Betrachtung erlauben. Werden vielfältige Gütemaße verwendet, hat der Anwender die Möglichkeit, durch eine subjektive Gewichtung der Gütemaße diejenige Prognose als die beste zu identifizieren, die seinen eigenen Präferenzen am besten entspricht.

Gütemaße

In der Literatur gibt es eine nahezu unüberschaubare Menge an Gütemaßen [vgl. z. B. ***Khosrawi-Rad, M.D***. 1991]. Um so überraschender ist es, daß bis heute nur wenige Studien existieren, die intensiver die Merkmale und Charakteristiken von Gütemaßen analysieren [wie z.B. ***Cicarelli, J.*** 1982; ***Wright, D.L.*** 1986 und besonders ***Dammers, E.*** 1993].

Eine optimale Auswahl an Gütemaßen sollte sich in erster Linie an den Eigenschaften Sensitivität, Art der Verlustfunktion, Fehlergewichtung, Homogenität, Monotonie, Interpretierbarkeit und Bekanntheit der Maße orientieren. Im einzelnen bedeutet das: Es sollten Maße enthalten sein, die jeweils speziell auf Abweichungs-, Niveau- und Wendepunktfehler sensitiv reagieren. Es sollten symmetrische (Unter- und Überschätzungen werden gleich bewertet) und asymmetrische Verlustfunktionen berücksichtigt werden. Fehlprognosen sollten linear, quadratisch (größere Abweichungen werden stärker gewichtet) und abhängig vom Niveau der beobachteten Datenwerte in die Berechnung unterschiedlicher Gütemaße eingehen. Es sollten homogene und monotone Gütemaße [vgl. ***Dammers, E.*** 1993] vertreten sein, damit Prognosen mit unterschiedlichen Einheiten und auf einer Kardinalskala miteinander verglichen werden können. Die einfache Interpretierbarkeit erleichtert die Anwendung. Bekannte Gütemaße ermöglichen darüber hinaus Prognosevergleiche über verschiedene Studien hinweg.

Sämtliche Eigenschaften können nur in einem „Gütemaße-Mix“ berücksichtigt werden. In Tabelle 1 werden einige wenige Gütemaße wiedergegeben, die für einen Prognosevergleich zu empfehlen sind (P_t: Prognosen zum Zeitpunkt t; O_t: Beobachtete Werte zum Zeitpunkt t; s: Streuung; M: Prognosehorizont; A: Häufigkeit „Wendepunkt (WP) prognostiziert und vorhanden; B: Häufigkeit „WP prognostiziert und nicht vorhanden“; C: Häufigkeit „WP nicht vorhanden und prognostiziert“; D: Häufigkeit „WP nicht vorhanden und nicht prognostiziert“).

Der mittlere Fehler (ME): Der ME zeichnet sich in der Auswahl durch Monotonie aus. Damit können Prognosevergleiche auf einer Kardinalskala erfolgen. Halbiert sich beispielsweise bei zwei zu vergleichenden Prognosen der ME, so kann gesagt werden, daß eine der Prognosen doppelt so gut ist wie die andere. Der ME bringt zum Ausdruck, in welchem Ausmaß die vorhergesagten Werte von den beobachteten Werten abweichen, wobei sich positive und negative Abweichungen gegenseitig ausgleichen können (asymmetrische Verlustfunktion).

Tabelle 1: Gütemaße des „Gütemaße-Mix".

Gütemaß	Formel
Mittlerer Fehler	$ME = \frac{1}{M} \sum (P_t - O_t)$
Mittlerer absoluter prozentualer Fehler	$MAPE = \frac{1}{M} \sum \left\lvert \frac{(P_t - O_t)}{O_t} \right\rvert \cdot 100\%$
Wurzel des mittleren quadr. Fehlers	$RMSE = \sqrt{\frac{1}{M} \sum (P_t - O_t)^2} = \sqrt{MSE}$
Bias-Anteil des MSE (MSE=RMSE^2)	$BIAS = \frac{(\overline{O}_t - \overline{P}_t)^2}{MSE}$; für $MSE \neq 0$
Varianz-Anteil des MSE	$VAR = \frac{(s_O - s_P)^2}{MSE}$; für $MSE \neq 0$
Kovarianz-Anteil des MSE	$KOV = \frac{2 \cdot (1 - R) \cdot s_O \cdot s_P}{MSE}$; für $MSE \neq 0$
Lineare Korrelation	$R = \frac{1}{M} \frac{\sum (P_t - \overline{P}) \cdot (O_t - \overline{O})}{s_P \cdot s_O}$
Wurzel des mittleren proz. Fehlers	$RMSP = \sqrt{\frac{1}{M} \sum \frac{(P_t - O_t)^2}{O_t^2}} \cdot 100\%$
Ballsches Q^2	$Q^2 = 1 - \frac{\sum (O_t - P_t)^2}{\sum (O_t - \overline{O}_{(A)})^2}$
Überzufällige Trefferquote	$ÜTQ = \frac{1}{M} \cdot \left((A + D) - \frac{(A+B)^2 + (C+D)^2}{M} \right)$

Der mittlere absolute prozentuale Fehler (MAPE): Der MAPE wird in erster Linie wegen seiner Fehlergewichtung und seiner anschaulichen Interpretation ausgewählt. Prognosefehler bewertet der MAPE als einziges Maß in der Auswahl linear und zugleich abhängig

vom Niveau der beobachteten Daten. Der MAPE ermittelt, um wieviel Prozent die Prognosen im Mittel von den beobachteten Werten abweichen. Wie alle folgenden Maße besitzt der MAPE eine symmetrische Verlustfunktion.

Die Wurzel aus dem mittleren quadratischen Fehler (RMSE): Der RMSE wird im wesentlichen wegen seiner quadratischen Fehlergewichtung, seiner einfachen Interpretation und seinem hohen Bekanntheitsgrad ausgewählt. Der RMSE liefert die Information, inwieweit die vorhergesagten Werte im Durchschnitt den beobachteten Werten entsprechen, ohne daß sich dabei positive und negative Abweichungen gegenseitig ausgleichen können. Dabei werden größere Abweichungen stärker gewichtet als kleinere Abweichungen.

Bias-, Varianz- und Kovarianz-Anteil (BIAS, VAR, KOV): Diese drei Gütemaße zerlegen den RMSE (bzw. MSE, siehe Tabelle 1) in systematische und unsystematische Fehleranteile. Darüber hinaus besitzt der Bias-Anteil als einziges Maß der Auswahl eine hohe und spezielle Sensitivität gegenüber Niveau-Fehlprognosen. Weicht der Bias-Anteil von Null ab, so ist dies als eine systematische Unter- bzw. Überschätzung der tatsächlich beobachteten Werte durch die Prognosen zu interpretieren (Niveau-Fehlprognose). Weicht der Varianz-Anteil von Null ab, so können die Schwankungen der Beobachtungswerte um ihren mittleren Wert zwar durch die Prognosen erfaßt werden, das Ausmaß dieser Erfassung ist jedoch zu stark bzw. zu schwach (systematische Abweichungs-Fehlprognosen). Prognosen, die weder systematische Niveaufehler noch Abweichungsfehler aufweisen, können sich von den beobachteten Werten nur noch unsystematisch unterscheiden. Diese unsystematischen Abweichungsfehler erfaßt der Kovarianz-Anteil. Bias-, Varianz- und Kovarianz-Anteil sind beschränkt (normiert) zwischen den Werten Null und Eins. Je kleiner der Bias- sowie Varianz-Anteil und je näher der Kovarianz-Anteil bei Eins liegt, desto besser ist die Prognose.

Die Korrelation (R): Die exklusive und hohe Sensitivität der Korrelation bezüglich quantitativer Wendepunkt-Fehlprognosen ist der Hauptgrund für die Auswahl. Von Vorteil bei der Interpretation ist die Beschränktheit und der hohe Bekanntheitsgrad. Je größer R ist ($-1 \leq R \leq 1$), desto mehr Wendepunkte werden in korrekter Richtung und mit korrektem Betrag prognostiziert.

Die Wurzel des mittleren quadr. prozentualen Fehlers (RMSP): Der RMSP wird ausgewählt, da er Prognosefehler sowohl quadratisch als auch in Abhängigkeit vom Ausgangsniveau der Daten gewichtet. Die inhaltliche Interpretation ist deckungsgleich mit der des MAPE. Es wird ermittelt, um wieviel Prozent die Prognosen im Mittel von den beobachteten Werten abweichen.

Das Ballsche (Q^2): Das Ballsche Q^2 liefert als einziges quantitatives Gütemaß der Auswahl einen direkten Vergleich mit einer Trivialprognose (arithmetischer Mittelwert der

beobachteten Werte). Je näher Q^2 bei einem Wert von Eins liegt, desto besser ist die Prognose im Vergleich zur gewählten Trivialprognose. Negative Q^2-Werte indizieren eine überlegene Trivialprognose.

Die überzufällige Trefferquote (ÜTQ): ÜTQ erfaßt exklusiv qualitative Wendepunkt-Fehlprognosen bzw. ausschließlich die Richtung (das Vorzeichen) der Fehlprognosen. Dabei wird ÜTQ um zufällig richtige Prognosen bereinigt. Je größer ÜTQ, desto besser ist die Prognose. Bei gleichwahrscheinlichen Kategorien kann ÜTQ nicht größer als 0.5 bzw. 50% werden. Eine negative überzufällige Trefferquote würde bedeuten, daß eine Zufallsprognose (Trivialprognose) gemäß den Kategorien-Randwahrscheinlichkeiten bessere Ergebnisse liefert als eine zu evaluierende Prognosemethode.

Dynamische Evaluation von Prognosen

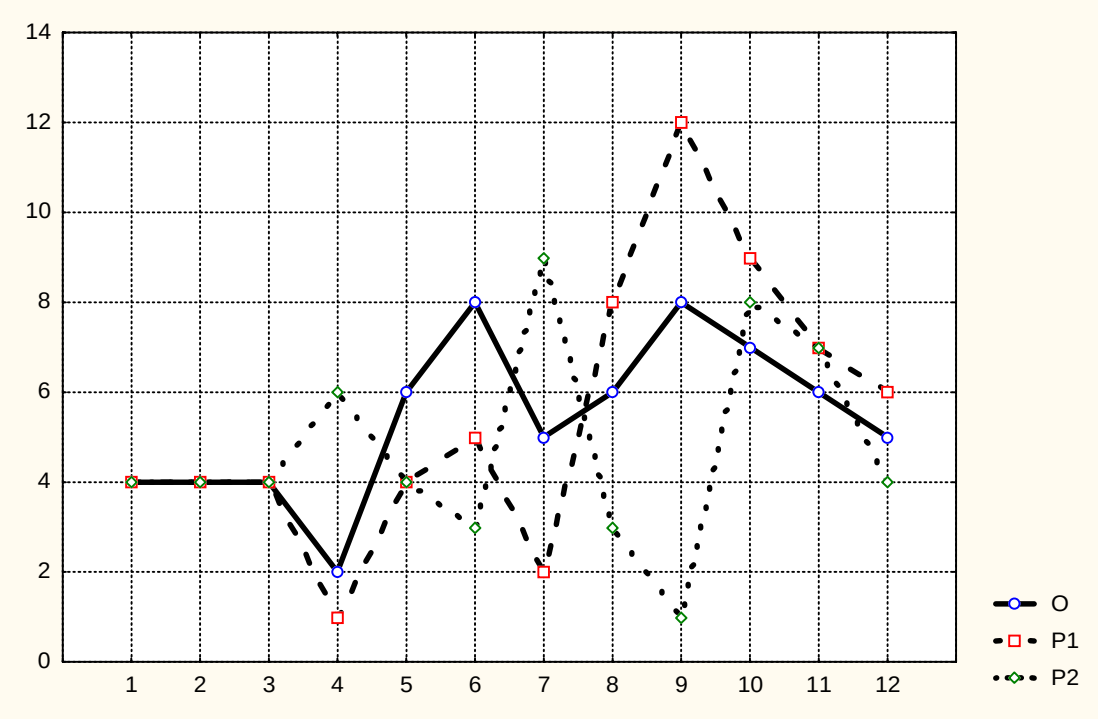
Anhand der empfohlenen Gütemaße ist es möglich, die horizontale Prognoseverteilung in statischer Form zu evaluieren, d.h. sämtliche Prognosewerte werden in einem Schritt bei der Berechnung der Gütemaße einbezogen. Prognosevergleiche in der Praxis werden üblicherweise auf diese Weise vorgenommen. Informativere Vergleiche sind jedoch möglich, wenn man die Prognosegüte in dynamischer Form darstellt. Diese erhält man, wenn man den Prognosehorizont sukzessiv um eine Periode ausdehnt und nach jedem neu hinzukommenden Datum die Prognosegüte neu berechnet. Eine dynamische Prognosegüte kann demnach nur berechnet werden, wenn den zu prognostizierenden Daten eine ordnende Struktur (z. B. die Zeit) zugrundegelegt werden kann. Bei einer größeren Fallzahl kann eine Neuberechnung auch nach einer jeweils festzulegenden Anzahl an Perioden erfolgen, so daß die wesentliche Information einer dynamischen Evaluation komprimierter dargestellt werden kann.

Durch die Dynamisierung kann die Prognosegüte in Abhängigkeit vom Prognosezeitraum betrachtet werden, d.h. man ist in der Lage zu beurteilen, in welchen Zeiträumen ein Prognoseverfahren bessere Prognoseergebnisse liefert als ein anderes. Ein derartiger Vergleich ist insbesondere für die Formulierung eines optimalen Prognosehorizontes einer Methode bei gegebenen Daten geeignet. Zur weiteren Analyse der dynamischen Prognosegüte kann man erneut statistische Verfahren (z. B. zeitreihenanalytische Methoden) heranziehen. Werden beispielsweise Gütemaße, die Abweichungs-Fehlprognosen sensitiv erfassen, dynamisch berechnet, kann darüber hinaus eruiert werden, ob eine schlechte Prognosegüte nur durch singuläre Ausreißer (Strukturbrüche in der dynamischen Darstellung), durch andauernde Fehlprognosen (Trends in der dynamischen Darstellung) oder durch Fehlprognosen nur zu bestimmten Zeitpunkten des Prognosezeitraumes (Trends und Plateaus in der dynamischen Darstellung) bedingt sind. Ebenfalls ist denkbar, daß Prognosen, die statisch evaluiert werden, als gleich gut eingestuft werden, während ein dynamischer

Prognosevergleich deutliche Unterschiede offenbart. Anhand eines Daten-Beispiels und der dynamischen Berechnung des ME, RMSE sowie des BIAS-, VAR- und KOV-Anteils (siehe Tabelle 1) soll dies verdeutlicht werden.

Die Prognosen „P1“ und „P2“ sind bezüglich der beobachteten Werte „O“ miteinander zu vergleichen (siehe Abbildung 1). Aus Gründen der Übersichtlichkeit werden lediglich 12 Beobachtungswerte verwendet. Bei dieser geringen Menge an Beobachtungen könnten grundlegende Prognosequalitäten auch per Augenschein erkannt werden. Bei Prognosen mit einer größeren Menge an Daten ist dies i.d.R. nicht möglich.

Abbildung 1: Prognosebeispiele.



Berechnet man die Prognosegüte statisch, dann erhält man zur Evaluation das Ergebnis in Tabelle 2. Man würde interpretieren, daß P2 sich durch stärkere Abweichungs-Fehlprognosen auszeichnet. Hinsichtlich einer Niveau-Fehlprognose unterscheiden sich beide Prognosen nicht. Insgesamt sind systematische Fehler stärker bei P1 als bei P2 ausgeprägt (siehe Abschnitt „Gütemaße“).

Berechnet man die Prognosegüte dynamisch (siehe Abbildung 2 bis Abbildung 6), dann erhält man zur Evaluation ein differenzierteres Bild der Prognosegüte. Der wesentlich größere mittlere Fehler von Prognose P2 entsteht offenbar erst ab Periode 9 (bedingt durch eine extreme Fehlprognose). Anhand des dynamischen RMSE erkennt man, daß der Unterschied der zwei Prognosen hinsichtlich der Intensität von Abweichungs-Fehlprognosen im

gesamten Prognosezeitraum ungefähr gleich bleibt (Abbildung 3). P2 leistet sich größere Abweichungsfehler als P1.

Tabelle 2: Gütemaße des „Gütemaße-Mix“.

Gütemaß	Prognose P1	Prognose P2
ME	-0.08	0.67
RMSE	2.02	3.19
BIAS	0.00	0.04
VAR	0.39	0.02
KOV	0.61	0.94

Der deutlichste Unterschied zwischen den zwei Prognosen, der anhand der statischen Prognosegüte nicht erkennbar war, besteht allerdings hinsichtlich einer systematischen Niveau-Fehlprognose. Bis zur Periode 8 zeichnet P1 einen erheblich höheren BIAS-Anteil aus, der sich erst ab Periode 9 dem von P2 gleicht (Abbildung 4). Des weiteren ist P1, im Unterschied zu P2, insbesondere zu Beginn und am Ende des Prognosezeitraums durch systematische Abweichungsfehlprognosen gekennzeichnet (Abbildung 5 und Abbildung 6).

Anwendungsbeispiele

Bei jeder empirischen Studie, die mit Prognosen operiert, stellt sich die Frage nach der Güte der Ergebnisse. Dabei sind die Ansprüche an den Informationsgehalt einer durchzuführenden Evaluation unterschiedlich. So mag für einen Wirtschaftswissenschaftler, der den Aktienmarkt bezüglich kurzfristiger Kauf- und Verkaufsignale analysiert, entscheidend sein, ob Prognosen in der Lage sind, Wendepunkte korrekt vorherzusagen. Ob sich Prognosen auch durch einen geringen Abweichungs- und/oder Niveaufehler auszeichnen, ist von untergeordneter Bedeutung. Durch eine dynamische Betrachtung der Prognosegüte ist es darüber hinaus möglich, optimale Prognosezeiträume, z. B. die optimale Vorlaufzeit eines kurzfristigen Handels zu ermitteln.

Abbildung 2: Dynamische Prognosegüte des ME.

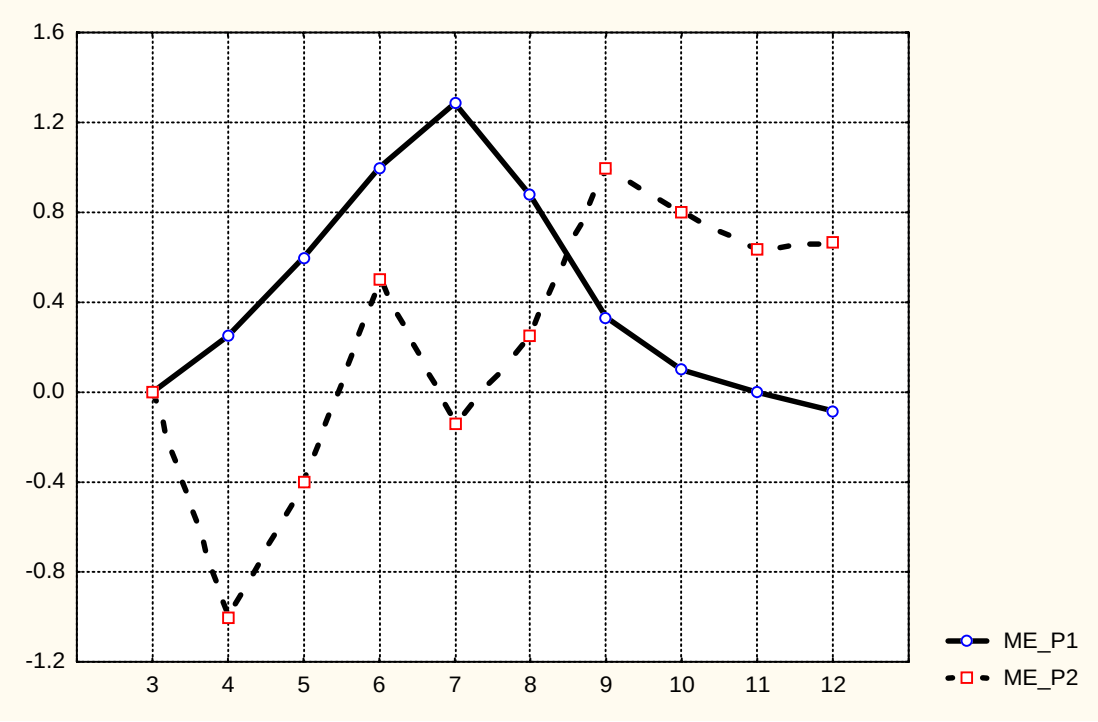


Abbildung 3: Dynamische Prognosegüte des RMSE.

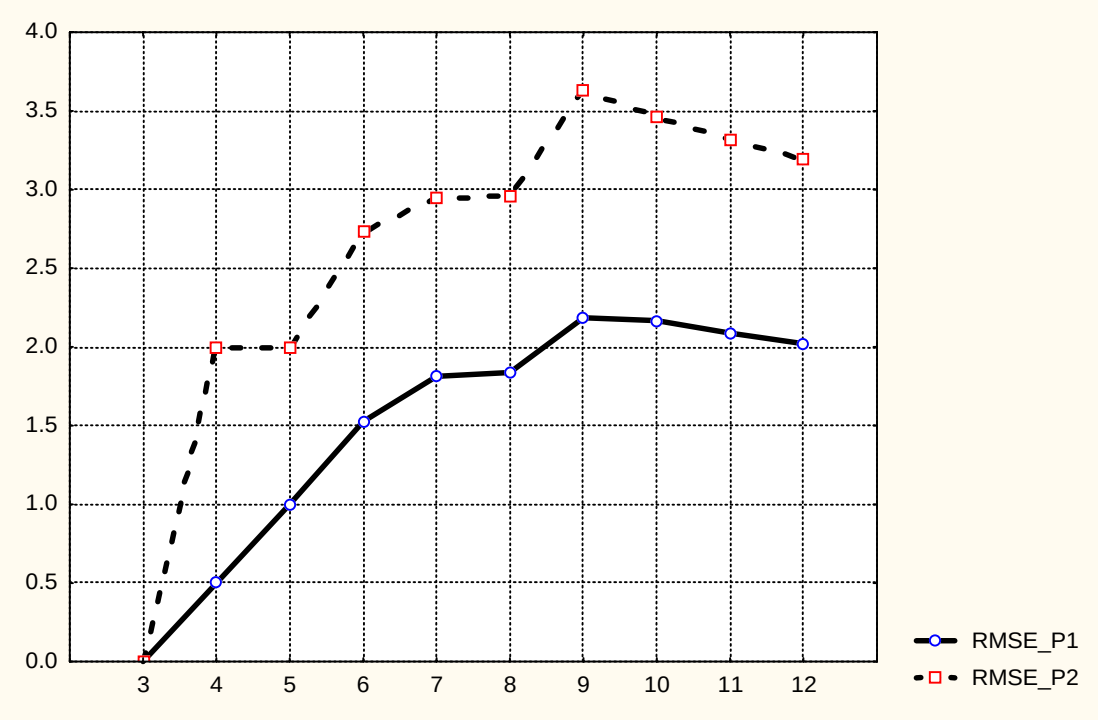


Abbildung 4: Dynamische Prognosegüte des BIAS-Anteils.

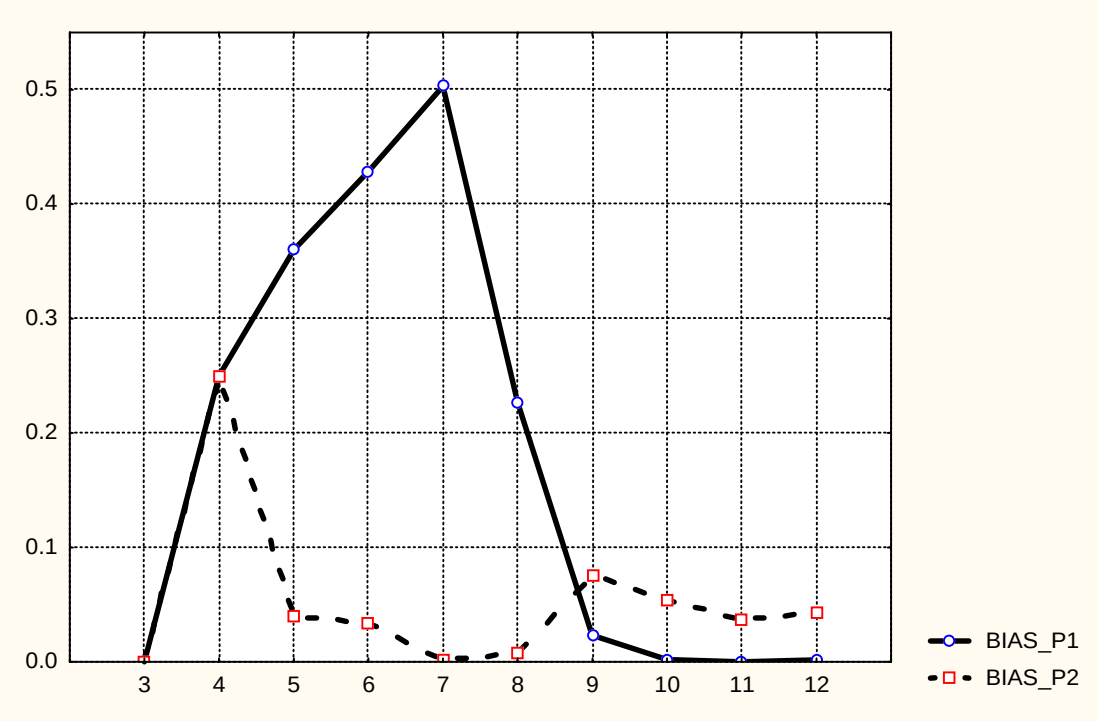


Abbildung 5: Dynamische Prognosegüte des VAR-Anteils.

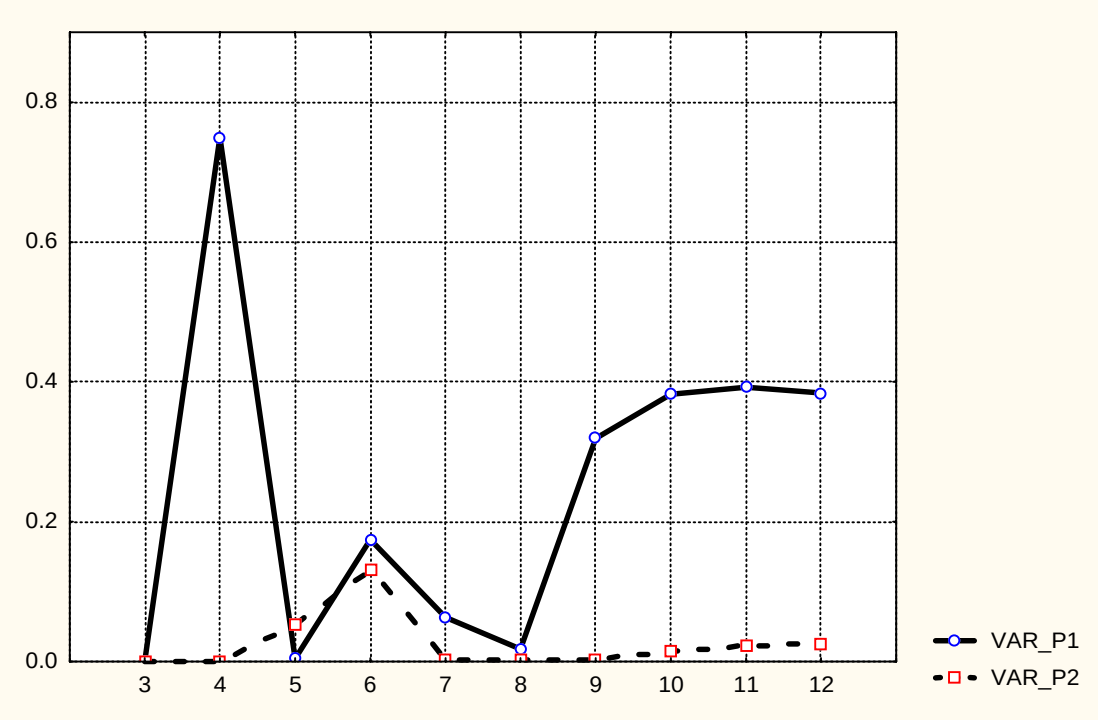
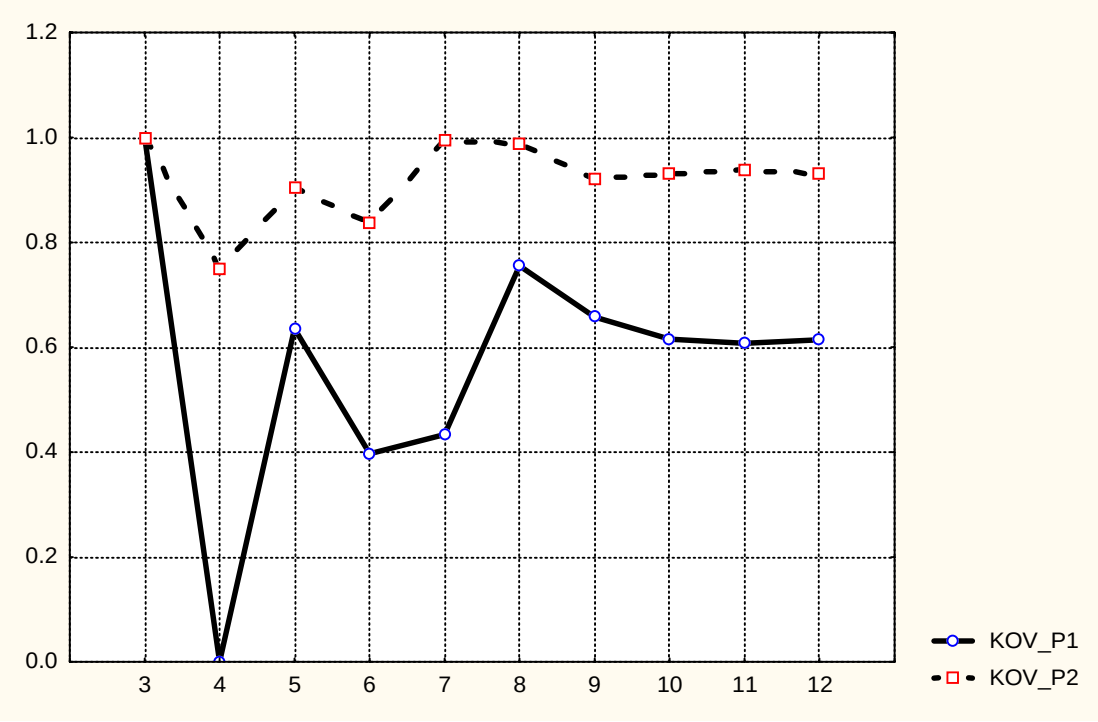


Abbildung 6: Dynamische Prognosegüte des KOV-Anteils.



Auf der anderen Seite könnte für einen empirischen Forscher (z. B. Soziologen, Psychologen, Kommunikationswissenschaftler), der den zukünftigen Verlauf der Fernsehnutzung analysieren und prognostizieren möchte, ein möglichst geringer Abweichungsfehler (eine möglichst gute Modellanpassung im Prognosezeitraum) im Fokus seiner Betrachtungen stehen. Zusätzlich wird hier eine Rolle spielen, ob die Prognose eher zu Unter- oder Überschätzungen der tatsächlichen Werte tendiert und eine Evaluation in der Lage ist, beide Arten der Fehlprognosen zu unterscheiden. Ein Forscher der Erziehungswissenschaften beispielsweise, der den Standpunkt vertritt, daß weniger fernsehen positiv zu beurteilen ist, wird vermutlich Überschätzungen weniger problematisch als Unterschätzungen beurteilen. Ein Forscher, der im Dienste der Programmanbieter arbeitet, und auf dessen Prognosen Preise für Werbezeiten berechnet werden, wird geneigt sein, seine Modelle so zu spezifizieren, daß unvermeidliche Fehlprognosen eher zu Unterschätzungen führen. Darüber hinaus liefert hier eine dynamische Betrachtung der Prognosegüte z. B. Anhaltspunkte dafür, für welchen Zeithorizont Preiszusagen vereinbart werden können bzw. ab welchem Zeitpunkt die Prognosegüte eines spezifizierten Modells Neuverhandlungen nahelegt.

Nicht zuletzt ist eine differenziertere statische und dynamische Evaluation von Prognosen für die Methodiker unter den empirischen Sozialforschern von Bedeutung. Ein Beispiel möge dies verdeutlichen: In den letzten Jahren ist zu beobachten, daß „moderne“ Modellierungs- und Prognosemethoden, die sich verschiedener Verfahren der künstlichen Intelligenz bedienen (z. B. neuronale Netze), zunehmend auch bei Analysen in der empirischen

Sozialforschung eingesetzt werden. Teilweise wird hierbei undifferenziert eine Überlegenheit bezüglich klassischer statistischer Verfahren behauptet, die heute, generell betrachtet, als Mythos einzuschätzen ist. In zahlreichen Analysen zeigt sich, daß der Einsatz derartiger Methoden nur bei bestimmten Datencharakteristiken zu empfehlen ist. Des weiteren offenbart eine differenzierte statische und dynamische Evaluation der Prognoseergebnisse (bzw. der Modellanpassung) in vielen Fällen, daß eine vorhandene Überlegenheit hinsichtlich einer Prognosequalität (siehe oben) durch eine Unterlegenheit in einer anderen Prognosequalität „erkauft“ wird. Eine dynamische Evaluation auf Basis eines „Gütemaße-Mix“ liefert hier dem Forscher ein Instrument, mit dessen Hilfe er entscheiden kann, ob sich für ihn die i.d.R. aufwendigere Anwendung neuer Modellierungs- und Prognosemethoden lohnt.

SAS - Programm

Tabelle 3 (siehe Anhang) liefert die dynamische Berechnung aller in Tabelle 1 angegebenen Gütemaße für die Daten-Vektoren „o_t“ (observed) und „p_t“ (predicted). Umfaßt die dynamische Prognosegüte den gesamten Prognosehorizont, so liefert das Programm zugleich die Prognosegüte in statischer Form (sämtliche Prognosewerte werden dann einbezogen). Die Berechnung der Gütemaße in Matrix-Schreibweise ermöglicht deren einfache Berechnung auch in einer anderen Software. Für andere Daten-Vektoren ist der Programmteil „Daten einlesen“ entsprechend zu modifizieren oder das Programm in Form eines SAS-Makros umzusetzen. Falls gewünscht, kann das Programm per e-mail (R.Weber@TU-Berlin.De) zugesandt werden.

Literatur

Khosrawi-Rad, M.D.: Probleme und Möglichkeiten bei der Definition, Klassifikation, Interpretation und Operationalisierung von Prognose, Prognosemodell, Prognosefehler und Prognosefehlermaße. Kovac-Verlag. Hamburg 1991.

Cicarelli, J.: New method of evaluating the accuracy of economic forecasts. In: Journal of Macroeconomics 4/1982. Seite 469-475.

Wright, D.J.: Evaluation of forecasting methods for decision support. In: International journal of forecasting Vol. 2/1986. Seite 139-152.

Dammers, E.: Measurement in the ex post evaluation of forecasts. In: Quality and Quantity 27/1993. Seite 31-45.

Anhang

Tabelle 3: SAS/IML-Programm zur dynamischen Berechnung der Gütemaße.

```
PROC IML;

* Daten einlesen;
* --------------;
ot = {4,4,4,2,6,8,5,6,8,7,6,5};  * Beobachtete Werte (0);
pt = {4,4,4,1,4,5,2,8,12,9,7,6}; * Prognostizierte Werte (P1);

* Dynamische Berechnung der Masse;
* -------------------------------;
Guete = J(nrow(ot)-2,10+1,0);

DO i = 3 to nrow(ot);

o = ot[1:i]; p = pt[1:i]; m = nrow(o);
meano  = o[:]; meanp  = p[:];
so1    = (o[##] - m*meano**2)/m;
sp1    = (p[##] - m*meanp**2)/m;

IF (so1 < 0)  THEN so1 = 0;
IF (sp1 < 0)  THEN sp1 = 0;

so     = sqrt(so1);
sp     = sqrt(sp1);

IF (so = 0)  THEN so = 0.0001;
IF (sp = 0)  THEN sp = 0.0001;

res     = o-p;
eins    = J(m,1,1);
deltao  = (o[2:m] - o[1:m-1])>=0;
deltap  = (p[2:m] - p[1:m-1])>=0;
n       = nrow(deltao);

ME    = (eins`*res)/m;
MAPE  = (eins`*abs(res/o))/m;
MSE   = (res`*res)/m;
RMSE  = sqrt(MSE);
R     = (((p-meanp)`*(o-meano)/m)/(so*sp));

IF (MSE ^= 0)  THEN BIAS = ((meano - meanp)**2) / MSE;
               ELSE BIAS = 0;
IF (MSE ^= 0)  THEN VAR =  ((so - sp)**2) / MSE;
               ELSE VAR =  0;
IF (MSE ^= 0)  THEN KOV =  (2 * (1 - R) * so * sp) / MSE;
               ELSE KOV =  1;

RMSP  = sqrt((res/o)`*(res/o)/m);
Q2    = 1 - (res`*res) / (m * so**2);
UETQ  = ((deltao=deltap)[+] - (deltao[+]**2 + (n-deltao[+])**2)/n) / n;

Guete[i-2,] = i||ME||MAPE||RMSE||BIAS||VAR||KOV||R||RMSP||Q2||UETQ;

END;

* Output produzieren;
* ------------------;
Namguete = {'Horizont' 'ME' 'MAPE' 'RMSE' 'BIAS' 'VAR' 'KOV' 'R' 'RMSP' 'Q2'
'UETQ'};
PRINT Guete[colname=namguete format=8.3];
CREATE GMASSE FROM guete [COLNAME=Namguete];
APPEND FROM guete;
QUIT;
RUN;
```

Wahlarithmetische Kabinettstückchen: Mandate im Überfluß oder wie Überhangmandate die Wahl entscheiden

von Christian Rindsfüßer und Susanne Schäfer-Walkmann[1]

Zusammenfassung

Überhangmandate nahmen in der Vergangenheit eine Ausnahmestellung im Deutschen Bundestag ein und hatten politisch so gut wie keine Bedeutung. Erst mit der deutschen Vereinigung, der damit verbundenen Neugliederung des Wahlgebietes und weitreichenden Veränderungen in der politischen Landschaft der Bundesrepublik Deutschland bekamen diese überzähligen Mandate eine wichtige Funktion. Sowohl nach der Bundestagswahl 1994 als auch 1998 stabilisierten die Überhangmandate äußerst knappe Mehrheiten. Obgleich das Bundesverfassungsgericht zwischenzeitlich die Verfassungsmäßigkeit der Überhangmandate bejaht hat, bleibt deren stimmgewichtsverzerrende Wirkung bestehen. Der Beitrag simuliert auf der Basis der Wahlergebnisse der Bundestagswahlen von 1994 und 1998 den Einfluß unterschiedlicher institutionell-rechtlicher bzw. wahlarithmetischer Faktoren, die das Zustandekommen und die Zahl der Überhangmandate beeinflussen: die Verteilung der Wahlkreise auf die Bundesländer, eine regional unterschiedliche Wahlbeteiligung und die Sperrklausel.

Abstract

In the past, the so-called „Ueberhangmandate“ (surplus seats) were a rare and low impact phenomenon in the German „Bundestag“. Yet as a result of German reunification, which brought about newly arranged electoral districts and far-reaching changes in the political landscape of the Federal Republic, surplus mandates gained in importance. Both in 1994 and 1998, they increased and stabilized an otherwise bare governing majority. Although

1 ***Christian Rindsfüßer*** (Dipl.-Stat.) ist Wissenschaftlicher Mitarbeiter an der Universität Augsburg, Institut für Volkswirtschaftslehre, Universitätsstraße 16, 86159 Augsburg.
Susanne Schäfer-Walkmann (Dipl.-Pol.) ist Wissenschaftliche Mitarbeiterin an der Universität Augsburg, Institut für Volkswirtschaftslehre, Universitätsstraße 16, 86159 Augsburg und am Internationalen Institut für Empirische Sozialökonomie (INIFES), Haldenweg 23, 86391 Stadtbergen.

the Bundesverfassungsgericht has confirmed the constitutionality of surplus seats, their distorting effects on the voters' decision is undeniable. Based on the results of the last two elections to the Bundestag, the contribution demonstrates the influence of different institutional-legal and arithmetic factors on the occurrence and number of surplus seats: the distribution of electoral districts among the „Laender", regionally differentiated participation rates, and the electoral threshold.

1 Einleitung

„Wahlarithmetische Kabinettstückchen" mit den sogenannten Überhangmandaten sind nichts Neues: Waren 1994 die 16 Überhangmandate (zwölf für die CDU, vier für die SPD) zum ersten Mal ausschlaggebend für die Wahl des Kanzlers und die Regierungsbildung,[2] stellen die diesmal entstandenen 13 Überhangmandate (alle für die SPD) ein beruhigendes Polster für eine Rot-Grüne Regierungsbildung dar.[3] Schließlich hatte der designierte Kanzler ***Gerhard Schröder*** im Vorfeld der Wahl mehrfach die Bedeutung einer „ausreichend stabilen" Mehrheit für die künftige Regierungsarbeit betont. Die knappe Mehrheit von acht Grundmandaten hätte wohl nicht als solide Basis für Rot-Grün bezeichnet werden können.

Diese 13 Überhangmandate machen immerhin 1,94% der Gesamtheit der 669 Sitze im Parlament aus und schlagen sich u.a. auf den „Erfolgswert" der errungenen Stimmen nieder. Beispielsweise benötigte die CDU mit 70.732 Stimmen im Vergleich zur SPD über 3.000 Stimmen mehr, um einen Sitz im Parlament zu erlangen. Noch krasser wirken sich die Verzerrungen hinsichtlich des Erfolgswertes innerhalb ein und derselben Partei aus, sofern diese Überhangmandate erringt: So mußten die Kandidaten der SPD 1998 in den Bundesländern, in denen **keine** Überhangmandate anfielen, durchschnittlich 70.607 (Zweit-)Stimmen erreichen, um ein (Listen-)Mandat zu erringen; dem stehen rechnerisch lediglich 53.365 Zweitstimmen je Mandat in solchen Bundesländern gegenüber, in denen es zu Überhangmandaten gekommen ist. Damit erhält die verfassungsrechtliche wie auch wahlgesetzliche Diskussion über die Überhangmandate sowohl in Fachkreisen als auch einer politisch interessierten Öffentlichkeit neuen Zündstoff.[4]

---

2 Im November 1994 wurde ***Helmut Kohl*** mit einer absoluten Mehrheit von nur einer Abgeordnetenstimme gewählt, obwohl die Koalitionsfraktionen im Bundestag über eine absolute Mehrheit von fünf Abgeordnetenmandaten verfügten; folglich hatten drei Abgeordnete den Kanzler nicht unterstützt. Ohne die zwölf Überhangmandate der CDU wäre eine Wiederwahl Kohls im ersten Wahlgang nicht möglich gewesen (vgl. u.a. ***Gaspers*** 1994 und ***Nicolaus*** 1996).

3 Die Aussagen/Berechnungen für die Bundestagswahl vom 27.09.1998 beruhen auf dem vorläufigen amtlichen Endergebnis, wie es vom Bundeswahlleiter im Internet (www.statistik-bund.de) hinterlegt wurde.

4 Der knappe Wahlausgang zog vor vier Jahren eine ganze Reihe von Einsprüchen beim Wahlprüfungsausschuß des Bundestages gegen die Bundestagswahl vom 16. Oktober 1994 nach sich und mündete – nach einer aufgrund einer Fristversäumnis gescheiterten Organklage der Partei Bündnis 90/Die Grünen – in ein Normenkontrollverfahren des Bundeslandes Niedersachsen zur verfassungsrechtlichen Überprüfung des § 7

2 Überhangmandate: eine Finesse des bundesdeutschen Wahlsystems

Das Zustandekommen von Überhangmandaten resultiert zunächst aus der für das bundesrepublikanische Wahlsystem typischen Verbindung der beiden Prinzipien relative Mehrheitswahl (bei den Erststimmen in den Wahlkreisen) und reine Verhältniswahl (bei den Zweitstimmen). Und obwohl je die Hälfte der 656 Sitze im Parlament über eines der beiden Verfahren bestimmt werden, ist für die Mandatsverteilung im Bundestag (und damit für die Mehrheitsverhältnisse) das Zweitstimmenergebnis prinzipiell entscheidend.[5] Übersteigt die Zahl der Direktmandate die der Listenmandate in einem Bundesland, bleibt die Differenz den Parteien als „Überhang" erhalten. Mit anderen Worten: Überhangmandate entstehen **in den Bundesländern**, ihre verzerrende Wirkung schlägt sich aber **auf Bundesebene** (bis hin zur Koalitionsbildung) nieder.

Neben diesen wahlsystematischen gibt es noch weitere grundlegende Voraussetzungen: Zu Überhangmandaten kommt es nur dann, wenn in einem Bundesland eine Partei eindeutig dominiert, d.h. die Kandidaten dieser Partei den „Löwenanteil" der Wahlkreise (bzw. fast alle Direktmandate) für sich gewinnen. Bei der Bundestagswahl 1994 profitierte davon die CDU in den Ländern Baden-Württemberg (+2 Überhangmandate), Mecklenburg-Vorpommern (+2), Sachsen (+3), Sachsen-Anhalt (+2) und Thüringen (+3), die SPD in den Ländern Brandenburg (+3) und Bremen (+1). Bei der Bundestagswahl 1998 entstanden 12 von 13 Überhangmandaten in den neuen Bundesländern. Neben einem Überhangmandat in Hamburg konnte die SPD in Brandenburg (+3), Mecklenburg-Vorpommern (+2), Sachsen-Anhalt (+4) und Thüringen (+3) zusätzliche Mandate erzielen.[6]

In die gleiche Richtung wirkt das sogenannte Stimmensplitting, denn die Wahrscheinlichkeit, daß Überhangmandate in dem beobachteten Ausmaß entstehen, ist wesentlich geringer, wenn nahezu alle Wähler ihre Erst- und Zweitstimme der gleichen Partei geben oder das Stimmensplitting zwischen den politischen Lagern ausgeglichen wäre.[7] 1994 zog die

---

Abs. 3 Satz 2 Bundeswahlgesetz (im folgenden: BWahlG). In seinem Urteil vom 10. April 1997 hat das Bundesverfassungsgericht (im folgenden: BVerfG) die Verfassungsmäßigkeit von Überhangmandaten positiv beurteilt. Im Vorfeld der Bundestagswahl 1998 wurde ebenfalls ein nur geringer Vorsprung für eine Partei bzw. Koalition prognostiziert, weshalb die Bedeutung möglicher (wahlentscheidender?) Überhangmandate erneut von der Presse aufgegriffen wurde (z.B. in FOCUS 32/1998, S. 30f. oder DER SPIEGEL 34/1994, S. 28).

5 Nach den §§ 6 und 7 BWahlG kommt es dann zu Überhangmandaten, wenn eine Partei in einem Bundesland mehr Direktmandate gewinnt, als ihr nach ihrem Zweitstimmenanteil zustehen (vgl. dazu: ***Nohlen*** 1990, S. 191ff.).

6 Diese Voraussetzung der homogenen Stimmabgabe für eine dominante Partei läßt sich besonders gut am Beispiel Sachsens verdeutlichen, wo im Gegensatz zu vor vier Jahren 1998 **keine** Überhangmandate entstanden: Rein rechnerisch liegt hier die Begründung darin, daß die SPD-Kandidaten diesmal acht der 21 Wahlkreise für sich entscheiden konnten und damit die Homogenität der Stimmenabgabe zugunsten der CDU aufgehoben war.

7 Bei einem Stimmensplitting mit ausgeprägten Parteipräferenzen entsprechen unter Umständen die Zweitstimmen einer Partei nicht mehr den von ihr gewonnenen Wahlkreismandaten, was das Entstehen von Überhangmandaten potentiell verstärkt – den in den Wahlkreisen erfolgreichen „großen" Parteien „fehlen"

CDU aus einer entsprechend getrennten Vergabe der Erst- und Zweitstimme einen größeren Vorteil als die SPD; bis dato war erfahrungsgemäß die Wählerschaft des rechten Lagers eher bereit, der „großen“ Partei die Erststimme zu geben. Bei der Bundestagswahl 1998 weist die Tatsache, daß die SPD bei den Erststimmen einen Zuwachs von 5,5% verzeichnen konnte - gegenüber 4,5% bei den Zweitstimmen - auf tendenzielle Veränderungen im Wählerverhalten hin: Zumindest bei der grünen Wählerschaft scheint die Bereitschaft zugenommen zu haben, in umkämpften Wahlkreisen dem SPD-Kandidaten die Stimme zu geben.

3 Überhangmandate – früher Ausnahmefall, seit drei Wahlperioden die Regel

Ausgenommen in der 4. Wahlperiode (1961; +5, alle CDU) gibt es Überhangmandate gehäuft erst wieder seit der 12. Wahlperiode (1990; +6, alle CDU). Anhand der statistischen Kennziffern läßt sich simulieren, daß im wesentlichen drei Faktoren institutionell-rechtlicher bzw. wahlarithmetischer Art das Zustandekommen von Überhangmandaten in größerer Zahl befördern, zumal diese Faktoren in der Tendenz additiv wirken. Im folgenden werden diese Faktoren sowie deren Wirkung auf das Wahlergebnis anhand fiktiver Annahmen - basierend auf den Datengrundlagen der Bundestagswahlen von 1994 und 1998 - vorgestellt.

3.1 Die Einteilung der Wahlkreise

Die Einteilung der Wahlkreise bei Bundestagswahlen erfolgt grundsätzlich so, daß jeder Wahlkreis ein zusammenhängendes Gebiet zu bilden hat, Ländergrenzen einzuhalten sind und Gemeinde- und Kreisgrenzen soweit als möglich nicht übertreten werden. Abgesehen davon, daß bereits die bislang geltenden, gesetzlich vorgeschriebenen Toleranzmargen von 25% bzw. 33,3% Abweichung[8] zu teilweise erheblichen Verzerrungen führen, ist das Problem noch ein ganz anderes: Bei der Wahlkreiseinteilung wird statt der Anzahl der **Wahlberechtigten** in einem Wahlkreis die Zahl der dort lebenden **Bevölkerung** zugrunde gelegt. Das wiederum bedeutet, daß im jeweiligen Wahlkreis lebende Kinder, Jugendliche und Ausländer in die Berechnungsformel eingehen, obwohl sich diese Bevölkerungsgruppen nicht an der Wahl beteiligen können.[9]

Simuliert man nun eine „gerechtere“ Wahlkreiseinteilung gemäß einer angenommenen weitgehenden „Gleichverteilung“ der Zahl der Wahlberechtigten in den einzelnen Wahlkreisen, so wird deutlich, daß insbesondere die Bundesländer, in denen es 1994 bzw. 1998 zu Überhangmandaten gekommen ist (Baden-Württemberg ausgenommen), tatsächlich zu viele Wahlkreise besitzen (vgl. Tabellen 1a und 1b).

---

gerade im Falle regionaler Konzentration der Wähler auf eine kleinere Partei die Zweitstimmen in den entsprechenden Ländern (vgl. dazu u.a.: ***Gaspers/Herbertz u.a.*** 1998, S. 637).

8 Vgl. § 3 BWahlG.

9 Vgl. dazu ausführlicher ***Gaspers/Herbertz u.a.*** 1998, S. 637.

Tabelle 1a: Überhangmandate bei fiktiver Wahlkreiseinteilung bei Annahme einer identischen Zahl von Wahlberechtigten in allen Wahlkreisen 1994 (Bundesdurchschnitt = 184.305 Wahlberechtigte/Wahlkreis[10])

Bundesland	Wahlkreise BT-Wahl 1994	Abweichung vom Durchschnitt der Wahlberechtigten	"Soll"-Wahlkreise bei "gerechter" Einteilung	Abweichung von der "Zahl der Wahlkreise"	Überhangmandate 1994	Überhangmandate fiktiv 1994
Baden-Württemberg	37	+5,7%	39,1	-2,1	2	4
Brandenburg	*12*	*-12,5%*	*10,5*	*+1,5*	*3*	*1-2*
Bremen	*3*	*-7,8%*	*2,8*	*+0,2*	*1*	*1*
Mecklenburg-Vorpom.	9	-16,9%	7,5	+1,5	2	0-1
Sachsen	21	-7,2%	19,5	+1,5	3	1-2
Sachsen-Anhalt	13	-10,0%	11,7	+1,3	2	1
Thüringen	12	-11,6%	10,6	+1,4	3	2
Summe	**107**	**-**	**101,7**	**5,3**	**16**	**10-13**

Quelle: Ergebnisse der Bundestagswahl 1994; eigene Berechnungen und Darstellung. Normalschrift: Überhangmandate für die CDU, *Kursivsetzung*: Überhangmandate für die SPD.

Tabelle 1b: Überhangmandate bei fiktiver Wahlkreiseinteilung bei Annahme einer identischen Zahl von Wahlberechtigten in allen Wahlkreisen 1998 (Bundesdurchschnitt = 185.092 Wahlberechtigte/Wahlkreis)

Bundesland	Wahlkreise BT-Wahl 1998	Abweichung vom Durchschnitt der Wahlberechtigten	"Soll"-Wahlkreise bei "gerechter" Einteilung	Abweichung von der "Zahl der Wahlkreise"	Überhangmandate 1998	Überhangmandate fiktiv 1998
Brandenburg	12	-9,0%	10,9	+1,1	3	2
Hamburg	7	-6,3%	6,6	+0,4	1	1
Mecklenburg-Vorpom.	9	-16,5%	7,5	+1,5	2	0-1
Sachsen-Anhalt	13	-10,8%	11,6	+1,4	4	3
Thüringen	12	-11,4%	10,6	+1,4	3	2
Summe	**53**	**-**	**47,2**	**5,8**	**13**	**8-9**

Quelle: Ergebnisse der Bundestagswahl 1998; eigene Berechnungen und Darstellung. Alle Überhangmandate für die SPD.

10 Für diese Simulation wurde - entgegen den Bestimmungen des Bundeswahlgesetzes - nicht die durchschnittliche Bevölkerungszahl in einem Wahlkreis zugrunde gelegt, sondern ein fiktiver Durchschnitt der Wahlberechtigten für jeden einzelnen Wahlkreis angenommen. Dazu wurde die Gesamtzahl der Wahlberechtigten (= 60,45 Mio. Wahlberechtigte 1994 bzw. 60,71 Mio. Wahlberechtigte 1998) durch die 328 Wahlkreise geteilt.

Viele der alten Bundesländer weisen im Falle einer derartigen, die Anzahl der Wahlberechtigten pro Wahlkreis berücksichtigenden Simulation eine zu geringe Zahl an Wahlkreisen auf. Baden-Württemberg hätte sowohl 1994 als auch 1998 rein rechnerisch zwei Wahlkreise mehr erhalten müssen, nämlich 39 – und somit die CDU 1994 bis zu vier! Überhangmandate (d.h. zwei Direktmandate zusätzlich zu den bereits erreichten). In Bremen und Hamburg machen sich die geringen Abweichungen zwischen „Ist"-Anzahl und „Soll"-Anzahl an Wahlkreisen nicht entscheidend bemerkbar. Brandenburg, Sachsen, Mecklenburg-Vorpommern, Sachsen-Anhalt und Thüringen verfügten nach diesem Modell über zu viele Wahlkreise - und damit über zu viele Direktmandate. Hier zeigt sich die Wirkung dessen, daß bei der Wahlkreiseinteilung die erhebliche Abwanderung aus den neuen Bundesländern (seit der Wahl 1990) bislang keine Berücksichtigung gefunden hat.[11]

Bei einer Einteilung der Wahlkreise, orientiert an der bundesdurchschnittlichen Zahl der Wahlberechtigten in allen Wahlkreisen, hätten - je nach Rundungsverfahren und Wahljahr - alle neuen Bundesländer je ein bis zwei Überhangmandate zuviel erhalten. Alles in allem wären bei einer „gerechteren" Wahlkreiseinteilung in den neuen Bundesländern 1994 nur ca. 7 statt 13 Überhangmandate entstanden und 1998 ca. 8 statt 12, insgesamt hätte sich die Zahl 1994 und 1998 um wenigstens 4 auf höchstens 13 (1994) bzw. 9 (1998) reduziert.

3.2 Der Faktor Wahlbeteiligung

Als ein zweiter Faktor befördert die Wahlbeteiligung in den einzelnen Bundesländern das Zustandekommen von Überhangmandaten. Genauer gesagt entstehen besonders viele Überhangmandate vorzugsweise in Bundesländern, die in puncto Wahlbeteiligung vom Bundesdurchschnitt in nicht geringem Umfang nach unten abweichen (vgl. Tabellen 2a und 2b). Interessant ist zunächst, daß sich, eine bundesweit gleichverteilte Wahlbeteiligung[12] angenommen, im Modell die Gesamtzahl der Überhangmandate 1994 von 16 auf elf reduziert hätte, 1998 wäre ein Rückgang von 13 auf zehn Überhangmandate zu verzeichnen gewesen.

---

11 Dazu muß bemerkt werden, daß für die Zeit ab der 14. Wahlperiode des Deutschen Bundestages die Regelung des Bundeswahlgesetzes modifiziert wurde. Künftig ist Vorschrift, „daß die Zahl der Wahlkreise in den einzelnen Ländern deren Bevölkerungsanteil soweit wie möglich entsprechen muß und die Bevölkerungszahl eines Wahlkreises von der durchschnittlichen Bevölkerungszahl nicht um mehr als 15 vom Hundert nach oben oder unten abweichen soll; beträgt die Abweichung mehr als 25 vom Hundert, ist eine Neuabgrenzung der betroffenen Wahlkreise vorzunehmen" (***Gaspers***/***Herbertz u.a***. 1998, S. 623). Zwar reduzieren sich die Verzerrungen aufgrund der Toleranzmargen, allerdings wurde erneut die Bevölkerungszahl (und nicht die Zahl der Wahlberechtigten) zugrunde gelegt. Darüber hinaus gibt es keine zwingenden Vorschriften für einen eventuell notwendigen Ausgleich zwischen den Bundesländern.

12 In diesem Modell wurde die Wahlbeteiligung in den einzelnen Bundesländern egalisiert, d.h. die absolute Stimmenzahl der Parteien auf die bundesdurchschnittliche Wahlbeteiligung (gültige Stimmen) von 77,9% (1994) bzw. 81,2% (1998) umgerechnet.

Tabelle 2a: Tatsächliche und fiktive Wahlbeteiligung im Vergleich, BTW 1994 (Annahme: identische Wahlbeteiligung in allen Bundesländern)

Bundesland	Wahlbeteiligung 1994 (gültige Stimmen)	Abweichung vom Durchschnitt (= 77,9%)	Überhangmandate 1994	Überhangmandate fiktiv
Baden-Württemberg	78,7%	0,8%	2	2
Brandenburg	70,8%	-7,1%	*3*	*2*
Bremen	77,3%	-0,6%	*1*	*0*
Mecklenburg-Vorpommern	71,2%	-6,7%	2	1
Sachsen	71,3%	-6,6%	3	2
Sachsen-Anhalt	69,6%	-8,3%	2	1
Thüringen	73,1%	-4,8%	3	3
Summe	**77,9%**	**-**	**16**	**11**

Quelle: Ergebnisse der Bundestagswahl 1994; eigene Berechnungen und Darstellung. Normalschrift: Überhangmandate für die CDU, *Kursivsetzung:* Überhangmandate für die SPD.

Tabelle 2b: Tatsächliche und fiktive Wahlbeteiligung im Vergleich, BTW 1998 (Annahme: identische Wahlbeteiligung in allen Bundesländern)

Bundesland	Wahlbeteiligung 1998 (gültige Stimmen)	Abweichung vom Durchschnitt (= 81,2%)	Überhangmandate 1998	Überhangmandate fiktiv
Brandenburg	75,9%	-5,3%	3	2
Hamburg	80,1%	-1,1%	1	1
Mecklenburg-Vorpom.	78,2%	-3,0%	2	1
Sachsen-Anhalt	76,0%	-5,2%	4	3
Thüringen	81,0%	-0,2%	3	3
Summe	**81,2%**	**-**	**13**	**10**

Quelle: Ergebnisse der Bundestagswahl 1998; eigene Berechnungen und Darstellung. Alle Überhangmandate für die SPD.

Als ein weiterer Aspekt fällt bei dieser Gegenüberstellung die bei beiden Bundestagswahlen deutlich geringere Wahlbeteiligung in den neuen Bundesländern auf - mit den entsprechenden Auswirkungen auf die Mandatsverteilung. Die zahlenmäßig schwankende, jedoch relativ große Gruppe der Nichtwähler (1994: 21,0%; 1998: 17,7%) bereitet Wahlforschern und Parteistrategen in mancherlei Hinsicht Kopfzerbrechen, denn anders als in der Vergangenheit ist das Wählerverhalten in den Demokratien nicht mehr so eindeutig strukturiert und vorhersagbar.[13] Eine Wählerentscheidung bildet sich im Zusammenwirken einer Vielfalt aufeinander bezogener struktureller wie situativer Faktoren, die unterschiedlich ins Gewicht fallen. Über derartige Faktoren kann letztlich nur spekuliert werden: Vielleicht hat

13 Vgl. ***Schultze*** 1992, S. 494ff.

sich ein nicht geringer Teil der unentschlossenen Wähler dafür entschieden, am 27. September nicht zur Wahl zu gehen, vielleicht hat sich die geringere Parteiidentifikation mit den großen Volksparteien (vor allem in den neuen Bundesländern) in der unterdurchschnittlichen Wahlbeteiligung niedergeschlagen bzw. sich darin die Enttäuschung über die Politik bzw. das mangelnde Gespür der politischen Repräsentanten für die soziale Lage und die Probleme in den neuen Bundesländern geäußert oder gar die Personalisierung des Parteienwettbewerbs und der eindeutige Zuschnitt auf die Spitzenkandidaten (bzw. deren mangelnde Verankerung in den neuen Bundesländern) ausgedrückt – die Liste ließe sich sicherlich fortsetzen. Wie auch immer, bezogen auf die hier behandelte Thematik hätte sich die Zahl der Überhangmandate eindeutig reduziert, wenn die Wahlbeteiligung in einem Bundesland nur geringfügig vom Bundesdurchschnitt abweicht; eine regional überdurchschnittliche Wahlbeteiligung hätte schließlich noch stärker dem Entstehen von Überhangmandaten entgegenwirkt.

3.3 Die Wirkung der Sperrklausel

Sperrklauseln wirken generell verzerrend, da ein bestimmter Prozentsatz der abgegebenen Stimmen bei der Berechnung der Mandatsverteilung keine Berücksichtigung findet und sozusagen „unter den Tisch“ fällt.[14] Zusammen mit der sogenannten Grundmandatsklausel (d.h. dem Gewinn der Direktmandate in mindestens drei Wahlkreisen) zeigt die in Deutschland geltende Fünf-Prozent-Sperrklausel ebenfalls Wirkung hinsichtlich des Zustandekommens von Überhangmandaten. In dem folgenden Modell (vgl. Tabellen 3a und 3b) wurde angenommen, daß alle Stimmen, die 1994 bzw. 1998 auf kleinere Parteien entfallen sind, von einer fiktiven „Sonstige“-Partei (1994: 3,6%; 1998: 5,9%) erzielt worden wären. Unter Annahme des Wegfalls der Sperrklausel hätte diese „Sonstige“-Partei ebenfalls an der Mandatsverteilung teilgenommen. In der Folge wären von den 656 Grundmandaten 1994 23 Mandate auf die „Sonstige“-Partei entfallen, 1998 sogar 39 Mandate. Alle anderen regulär im Bundestag vertretenen Parteien hätten entsprechend weniger Mandate erzielt.

Tabelle 3a und 3b verdeutlichen die Wirkung einer solchen Annahme auf die Überhangmandate in den einzelnen Bundesländern: Ohne Sperrklausel 1994 bzw. 1998 hätte sich bei beiden Wahlen die Zahl der Überhangmandate von 16 auf 20 (1994) bzw. 13 auf 16 (1998) noch vergrößert.

14 Die Sperrklausel im Bundeswahlgesetz beispielsweise schließt solche Parteien oder Wählervereinigungen von der Vergabe der Mandate aus, die nicht mindestens fünf Prozent der im gesamten Wahlgebiet abgegebenen gültigen Zweitstimmen errungen haben.

Tabelle 3a: Die Auswirkung der Sperrklausel auf die Zahl der Überhangmandate 1994 (Annahme: Keine Sperrklausel)

Bundesland	Wahlkreise BT-Wahl 1994	Mandatezu-teilung im Modell	Gewonnene Direktman-date 1994	Überhang-mandate 1994	Überhang-mandate fiktiv
Baden-Württemberg	37	34	37	2	3
Brandenburg	*12*	*9*	*12*	*3*	*3*
Bremen	*3*	*2*	*3*	*1*	*1*
Hamburg	*7*	*5*	*6*	*0*	*1*
Mecklenburg-Vorpommern	9	5	7	2	2
Sachsen	21	17	21	3	4
Sachsen-Anhalt	13	8	10	2	2
Thüringen	12	8	12	3	4
Summe	**114**	**88**	**108**	**16**	**20**

Quelle: Ergebnisse der Bundestagswahl 1994; eigene Berechnungen und Darstellung. Normalschrift: Überhangmandate für die CDU, *Kursivsetzung*: Überhangmandate für die SPD.

Tabelle 3b: Die Auswirkung der Sperrklausel auf die Zahl der Überhangmandate 1998 (Annahme: Keine Sperrklausel)

Bundesland	Wahlkreise BT-Wahl 1998	Mandatezu-teilung im Modell	Gewonnene Direktman-date 1998	Überhang-mandate 1998	Überhang-mandate fiktiv
Brandenburg	12	9	12	3	3
Hamburg	7	6	7	1	1
Mecklenburg-Vorpom.	9	5	7	2	2
Sachsen-Anhalt	13	8	13	4	5
Schleswig-Holstein	11	10	11	0	1
Thüringen	12	7	11	3	4
Summe	**64**	**45**	**61**	**13**	**16**

Quelle: Ergebnisse der Bundestagswahl 1998; eigene Berechnungen und Darstellung. Alle Überhangmandate für die SPD.

Rein rechnerisch läßt sich somit ein Zusammenhang zwischen der Sperrklausel und der Gesamtzahl der Überhangmandate aufzeigen: Scheitern nämlich viele Parteien relativ knapp an der Sperrklausel und werden entsprechend bei der Verteilung der Mandate nicht berücksichtigt, reduziert sich die Zahl der Überhangmandate um ein nicht geringes Maß. Umgekehrt bedeutet das: „Wenn mehr Parteien nach Überspringen der Sperrklausel an der Sitzverteilung nach Zweitstimmen teilnehmen, diese Parteien aber keine Wahlkreise erringen, steigt die Wahrscheinlichkeit für das Anfallen von Überhangmandaten ...“.[15]

15 ***Gaspers/Herbertz u.a.*** 1998, S. 637. In einem solchen Fall erzielen die kleineren Parteien (die in der Regel das Direktmandat nicht erzielen können) die Zweitstimmen, die dann den „großen“ Parteien bei der Berechnung der Sitzverteilung fehlen.

Betrachtet man die Faktoren „Einteilung der Wahlkreise“, „Wahlbeteiligung“ und die „Wirkung der Sperrklausel“ zusammen, so wird die Entstehung der hohen Zahl von Überhangmandaten nachvollziehbar. Allein durch eine höhere Wahlbeteiligung in den neuen Bundesländern, verbunden mit einer gerechteren Wahlkreiseinteilung hätte sich die Zahl der Überhangmandate um mehr als die Hälfte reduzieren lassen.

4 Mandate im Überfluß? [16]

Die Fülle der Beschwerden nach der Bundestagswahl 1994 führte hinsichtlich der Überhangmandate für 1998 zu keinen neuen Regelungen – obwohl feststand, daß die Beibehaltung des bisherigen Wahlverfahrens das Entstehen von Überhangmandaten auch bei der Bundestagswahl 1998 begünstigen würde.[17] Je höher deren Zahl, desto mehr wird das an sich vorrangige Prinzip des Verhältniswahlrechts unterlaufen und desto größer sind die Verzerrungen hinsichtlich der Sitzverteilung. Insbesondere kleine Parteien, die nur wenig bzw. keine Aussicht auf den Gewinn von Direktmandaten haben, sind davon betroffen. Angesichts knapper werdender Mehrheiten, nachlassender Wählerbindungen an eine Partei bzw. Unentschlossenheit bis zum Wahltag können Überhangmandate - besonders, wenn sie ähnlich zahlreich sind wie 1994 und 1998 - auch zukünftige Bundestagswahlen entscheiden.

In der Diskussion muß unterschieden werden zwischen wahlarithmetischen Faktoren, die nur wenig beeinflußbar sind und solchen, die entsprechenden Verzerrungen entgegenwirken. Das Stimmensplitting etwa ist die explizite Möglichkeit im bundesdeutschen Wahlrecht, sich mit der Erststimme ausdrücklich für einen bestimmten Kandidaten zu entscheiden, mit der Zweitstimme jedoch einer anderen Partei den Vorzug zu geben. Dem Grunde nach könnte jede der großen Parteien davon profitieren. Dieser Faktor liegt aber ausschließlich in der Hand des Wählers, ebenso wie die grundsätzliche Wahlbereitschaft oder die Verweigerung der Stimmabgabe.

Dennoch hätte man im Vorfeld der Bundestagswahl einige der institutionell-rechtlichen Faktoren ändern können, um derartigen wahlsystematischen Verzerrungen durch Überhangmandate vorzubeugen. Angeboten hätte sich in erster Linie eine Neuordnung der Wahlkreise, jedoch verdeutlichen die „wahlarithmetischen Kabinettstückchen“, daß die bisherige Koalition optimistisch von einer (wenn auch knappen) Wiederwahl ausging und deshalb wahlarithmetisch alles „auf eine Karte“ setzte. Für die SPD rechnete sich 1998 die

16 ***Naundorf*** (1996) bezeichnete die Überhangmandate gar als „überflüssigen Überhang“.

17 So kritisierte beispielsweise der Augsburger Politikwissenschaftler Prof. ***Rainer-Olaf Schultze*** das Urteil des Bundesverfassungsgerichtes als „völlig unsystematisch“, insbesondere hinsichtlich der Aussage, „daß Überhangmandate ab einem Anteil von fünf Prozent an der Gesamtzahl problematisch sind. Fünf Prozent entsprechen 31 oder 32 Überhangmandaten, ein nicht unrealistisches Szenario für eine Bundestagswahl. Dann befänden wir uns in einer schwierigen verfassungsrechtlichen Situation“ (FOCUS 32/1998, S. 31).

verstärkte Erststimmenkampagne gerade in den neuen Bundesländern, die den Bekanntheitsgrad der ostdeutschen Wahlkreiskandidaten erhöhte. Erst die erdrutschartigen Verluste der CDU in den neuen Bundesländern ermöglichten es der SPD, dort die politischen Machtverhältnisse umzudrehen.

Noch ein Wort zur Wahlbeteiligung: Es ist bekannt, daß „Nichtwahl“ eine Möglichkeit ist, seinen Unmut mit der Politik und vor allem auch mit den Parteien auszudrücken. Unter anderem führte die geringe Wahlbeteiligung insbesondere in den neuen Ländern bei der Wahl 1994 und - etwas abgeschwächt - bei der Wahl 1998 jedoch dazu, daß gerade die großen Parteien ihre Position durch Überhangmandate ausbauen konnten. Hier müssen sich die Wähler fragen, ob sie diesen Effekt mit ihrer Stimmenthaltung bewirken wollen. Eine wie auch immer motivierte Wahlenthaltung in den neuen Bundesländern schwächt deren Gewicht im Bundestag insgesamt und führt – unabhängig von der Diskussion um die Überhangmandate – zu einem Abfluß von erzielbaren Mandaten von den neuen in die alten Bundesländer.

Die Überhangmandate – lange Zeit unbeachtet - haben neues Gewicht bekommen, gerade weil auf wahlarithmetische Mittel, wie etwa Ausgleichsmandate, verzichtet wurde. Nach diesem knappen Wahlausgang wird das leidige Thema Überhangmandate auch in Zukunft noch nicht vom Tisch sein. Die anstehende Neueinteilung der Wahlkreise für die nächste Bundestagswahl läßt jedoch hoffen, daß zumindest der Faktor Wahlkreisgröße neutralisiert werden kann.

Literatur

Gaspers, K. 1994:
Grundlagen und Daten der Wahl zum 13. Deutschen Bundestag am 16. Oktober 1994. In: Wirtschaft und Statistik o. Jg., H. 8, S. 622-634.

Gaspers, K. / Herbertz, H.-C. und ***Mitarbeiter*** 1998:
Grundlagen und Daten der Wahl zum 14. Deutschen Bundestag am 27. September 1998. In: Wirtschaft und Statistik, o. Jg., H. 8, S. 623-639.

Naundorf, C. 1996: Der überflüssige Überhang: Reformvorschläge. In: Zeitschrift für Parlamentsfragen 27, H. 3, S. 393-397.

Nicolaus, H. 1996:
Die unzulängliche Rechtfertigung der Überhangmandate: Aufklärungsversuche. In: Zeitschrift für Parlamentsfragen 27, H. 3, S. 383-393.

Nohlen, D. 2.Aufl. 1990:
Wahlrecht und Parteiensystem. Über die politischen Auswirkungen von Wahlsystemen. Opladen.

o.V. 1998:
Überhangmandate. Der Osten entscheidet. In: FOCUS 32/1998 (03.08.1998), S. 30-31.

o.V. 1998:
Überraschender Vorsprung, in: DER SPIEGEL 34/1998 (17.08.1998), S. 28.

Schultze, R.-O. 1992:
Wählerverhalten. In: ***Schmidt***, ***Manfred G.*** (Hrsg.): Die westlichen Länder (Lexikon der Politik, Bd. 3), München, S. 494-510.

Ethnocentrism in Present Germany: Some Correlations with Social Integration and Subjective Assessments

by Michael Terwey

Zusammenfassung:

*Arbeiten von **Durkheim** über Religion und soziale Integration werden als ein Ansatz für die Analyse von Ethnozentrismus in Deutschland verwendet. Dementsprechend ist eine relativ weite, funktionale Definition des Religiösen einzusetzen. Jede Religion enthält Elemente, welche den engeren Kreis offensichtlich religiöser Institutionen überschreiten. Aktive Integration in soziale Gruppen und Identifikation mit ihren Idealen beeinflussen stark die Identitätsfindungen und Einstellungen von Individuen. Die nachfolgenden empirischen Untersuchungen versuchen, Antworten auf verschiedene Fragen zu geben. Sind die unterschiedlichen Stärken von Ethnozentrismus, die sich zwischen Ost- und Westdeutschland feststellen lassen, in Verbindung zu sehen mit den stark differierenden Graden kirchlicher Bindung? Inwieweit stehen politische Linksorientierungen in Verbindung mit der Akzeptanz von Minoritäten? Sind es subjektive ökonomische Unzufriedenheiten, welche das gegenwärtige Ausmaß der ethnischen Intoleranz allgemein bestimmen? Zusätzliche Aufmerksamkeit gilt den "Postmaterialisten", deren Wunschhaltungen auch eine Prioritätssetzung auf die "multikulturelle Gesellschaft" beinhalten mag. Einige Untersuchungsergebnisse deuten auf eine besondere Bedeutung vager Befürchtungen und sozialer Pessimismen für die unterschiedliche Ausbreitung von Ethnozentrismus hin.*

Abstract:

*The work of **Durkheim** about social integration and elementary forms of religious life is taken as a major starting point for the analysis of ethnocentrism in Germany. His functional conception of the religious is not narrowly oriented as in some theories of contemporary sociologists. Every religion contains parts by which it overlaps the circle of obviously religious institutions. Active integration into social groups and identification with their special ideals shape identities and attitudes of individual men. Our empirical investigations deal with several questions. Are the differences in ethnocentrism which we find for*

East- and West Germany explained by differing attachments to Christian churches? To what extent do left-wing ideologies foster tolerance in the New Federal States? Do subjective economic wants account for contemporary ethnocentrism in post-unification Germany? Additional attention should be paid to the 'postmaterialists', whose aspirations may include a priority on the establishment of a 'multicultural society'. Some empirical results suggest the importance of vague anxieties and social pessimism for differing strength of ethnocentric attitudes.

Introduction

For more than a thousand years political affairs were closely connected with Christian ideology and organization. According to ***Durkheim*** (1960), 'God' with somewhat changing traits has served as society hypostasized. A religious importance of such kind may also be attached to the integrating concept of a Nation or 'Volk' (people), but there is no essential sociological difference assumed between an assembly of Christians celebrating biblical events or Jews remembering their Exodus from Egypt and a reunion of citizens commemorating a cherished goodness in their national past (***Durkheim*** 1960). In spite of all 'secularisation', there seems to be a lasting need of a positive ideological 'intégrisme' (cf. amongst others ***Martin*** 1978) combined with some kind of 'transcendences' in the differentiated sense of ***Luckmann*** (1995).

Social relationships are to a large extent constituent for our perception or even construction of reality (***Berger*** and ***Luckmann*** 1974). Many Christians, for instance, imagine or conceive the existence of a God, who is really communicating with men and the existence of a positive community in parish or church. Joining in such a concordant community and attaching respective meanings to the ceremonies in church already exalts beyond the ordinary life. The real event of a somewhat transpersonal communication in practiced church seems to prove to a certain extent that there is true meaning in Christian belief and fellowship. Hence, experiences of solitude, interindividual alienation, hostility and meaninglessness may be expected to be less constituent for the 'Weltanschauung' of Christians in communities than for unchurched people who have not established a functional equivalent. Some kind of 'fait social' or 'Society' in the almost sacral sense of ***Durkheim*** (1960) is proven (cf. also ***Firsching*** 1995). Communal religious practices evolving around more or less portentous 'transcendences' (***Luckmann*** 1995) are suitable for achieving an in-group identification, a positive social identity and an individual self-conception.

A sociological conceptualization of religion, social integration and supplementing worldviews related to economy or moral principles is further exemplified by the following quotation: ‘Man's characteristic privilege is that the bond he accepts is not physical but moral; that is, social. ... the source of the crisis is an abrupt growth of power and wealth. Then,

truly as the conditions of life are changed, the standard according to which needs were regulated can no longer remain the same. ... The scale is upset; but a new scale cannot be immediately improvised. ... Since the relations between various parts of society are necessarily modified, the ideas expressing these relations must change. ... Wealth, on the other hand, by the power it bestows, deceives us into believing that we depend on ourselves only. Reducing the resistance we encounter from objects, it suggests the possibility of unlimited success against them ... Not without reason, therefore, have so many religions dwelt on the advantages and moral value of poverty. It is actually the best school for teaching self-restraint. ... From top to bottom of the ladder, greed is aroused without knowing where to find ultimate foothold. ... A thirst arises for novelties, unfamiliar pleasures, nameless sensations, all of which loose their savor once known. Henceforth one has no strength to endure the least reverse.' (***Durkheim*** 1951: 252 - 256)

Nevertheless, a combination and maximizing of various subjective, personal demands in the sense of *'Wertesynthese'* (synthesis of values; cf. ***Gille*** 1994) may of course endorse many positive ideals. One of the challenging ideals could be a positively integrated, multicultural society, where social security, tolerance, hospitality and the protection of various human rights are guaranteed, while the costs ought to be candidly covered by transfers from security spending, higher taxes for the upper milieus or some other 'skilled' politics.

Ethnocentrism: An intergroup differentiation shaping identification

The basic importance of Christian ideology has been clearly shown by European and American history. Accordingly, it has also been demonstrated, if not to say even overrated, by important sociologists as ***Durkheim*** (1951), ***Weber*** (1993) and ***Tönnies*** (1979; cf. also ***Greeley*** 1995a). As already introduced above, there are some momentous worldviews which are not explicitly attached to a formal church. Union movement, Marxist parties and National Socialism are important examples for organizations proclaiming worldviews which are in some respects similar to traditional religions. An extreme, but very simple example is shown by the quotation of ***Adolf Hitler*** in ***Glock*** and ***Stark*** (1966: XV): 'I believe that I am acting in accordance with the will of the Almighty Creator: by defending myself against the Jew, I am fighting for the work of the Lord.'

In a 'post-modern' period of assumed individualization non-material or material preferences could be dissociated from larger obvious ideologies, but to a certain extent socializing may not succeed without a larger communal reinforcement. In general, we have still to consider individual traits and social group-attachments as constituents of personal identity - and moreover, we have to muse on the intergroup dimension as a medium of externally rooted 'identification'. A basic understanding of social reality is often established by differentiating between an in-group and various out-groups (cf. ***Tajfel*** 1982; ***Wasmer*** et al.

1996). A simple interpretation of the social identity theory of ***Tajfel*** and others could be attached to a nation, a special people (unser Volk) or a church which would be contrasted to various out-groups. The social identity theory proposes that social categorization is employed by individuals not only to systematize and make sense of a complex outer world but also to make sense of oneself. Accordingly, self-concepts are built on both personal identities, who we believe to be as embodied individuals and on an assumed social identity, the groups with which we align or disconnect ourselves. Group (non-) allegiances assist in defining, to ourselves and to others, who we are in terms of our social connections. There are some similarities between the concept of in-group in the sense of social identity theory and the concept of (religious) community proposed by ***Durkheim***, but the latter is more often used with positive connotations as 'society' is assumed to be basically sound (cf. the concept of 'anomie' introduced below).

The most extensive in-group of Germany is originally constituted by the German nationality of the parents or at least one of the parents (ius sanguinis), not by the place of birth as in some other countries. Note, that present out-groups do not necessarily have to be foreigners and that many people who today are testified Germans have formerly immigrated into the German area. Social problems as rising costs for social services may be attributed to some criticized parts inside the nation (e.g. Germans who shun legal work) or to out-grouped ethnics and migrants, as it is still done today. More or less negatively privileged groups whose members often have a German citizenship could be Jews, Roma, Sinti, 'aussiedler'[1] or members of certain religious sects and political groups. The definition of an in-group may not be considered as a vague expression of pride and self-esteem, but also as a means of restriction intended to cope with *moral or economic problems*. A very important example with religious, ethnic and economical connotations is the historical role of Jews, who were held responsible for numerous social and even physical problems.

Today, statements about foreigners are increasingly connected with economic assertions about asylum, social welfare, unemployment, national labor market competition, globalization and criminal violence (***Velling*** 1994; ***Mochmann*** and ***Oedegaard*** 1995; ***Albrecht*** 1997; ***Heitmeyer*** 1997).[2] Diverse impressions about immigrants may be improved or even deteriorated by direct contact. With some groups among the foreigners such contacts can be easier established than with others (e.g. contacts at work cannot be made with ethnics that have yet not entered the labor force), but media reports and political declarations could

1 'Aussiedler' are ethnic Germans being descendants of German farmers and craftspeople who had moved to Eastern Europe in former centuries and came back to Germany after the Second World War.

2 In May 1998 14% of respondents (Politbarometer) named foreigners and asylum as one of the two most important political topics in Germany. The immigration issue has moved into second place, surpassed only by the topic of unemployment (85% of respondents thus designate unemployment as such an especially important issue; Politbarometer 05/98).

counteract personal positive experiences with guest workers, foreign students and other agreeable 'auslander'. It may be brought to notice again that in recent years the nationalities and the motives of immigrants have significantly changed; accordingly the former designation of 'guest workers' (Gastarbeiter) has become more and more untoward for immigrants in general (***Velling*** 1994; ***Rudolph*** 1996; ***Blank*** and ***Schwarzer*** 1994; ***Blank*** and ***Wasmer*** 1996).

Labor market competition and special identification with an ethnic majority group are probably connected. Therefore, negative economic assessments shall be analysed as subjective background for *competition-based ethnocentrism* (cf. ***McCutcheon*** 1998). According to these competition-based theories the highest levels of hostility are supposed to occur among majority group-members in the lower strata of social inequality. Following ***Kornhauser*** (1959) and ***McCutcheon*** (1998), *mass society-based theories of ethnocentrism* could yield alternative explanations for deviating results. In a low-tension mass society membership in smaller communal groups can protect individuals from taking over extreme ideologies. This may be particularly true, if these communities are characterized by tolerant norms of social conduct. In such liberal or compassionate communities we could assume a positive effect of group-integration, with slight reference to Christian tenets or humanist ethics. On the other hand, many communal groups may foster a more restrictive and less tolerant conception of various outsiders (cf. the discussion of social identity theory above and the idea of ‘Gemeinschaft’ as used by ***Tönnies***).

Some survey items about the acceptance of foreigners in Germany

Originally, the term 'guest worker' was used to refer to foreign laborers who had since 1955 been recruited from certain countries in Southern Europe and North Africa (cf. ***Thränhardt*** 1988; ***Velling*** 1994; ***Kühnel*** and ***Terwey*** 1994). For some time, guest workers almost determined the public image of 'foreigners' in West Germany. In contrast to smaller numbers of other immigrants from the Northwest or Central Europe they were often perceived as apparently foreign, and they made up the vast majority of all immigrants till the mid 80s.

The guest workers of the 60s and early 70s were practically invited to immigrate, because there was a shortage of labor supply. They had politically unlimited admission to enter the German labor force and hence the guest workers were having an extremely high employment rate. Later, images about 'foreigners' in Germany were modified due to new subgroups of migrants. Many of the recent foreign migrants are not 'invited guests' who assist in overcoming a shortage of labor supply, but they are unemployed and somehow unwanted refugees dependent on social welfare assistance. The migrants from Eastern Europe as well as the asylum seekers from Asia and Africa are primarily to be mentioned in this

respect. Even though views about 'guest workers' from the time of recruitment after 1955 remained important factors in featuring the public opinion about immigrants, these new streams of immigration nurtured new identifications for out-groups from abroad. Moreover, the immigrants in the early 90s contrast with previous ones by their higher heterogeneity (***Velling*** 1994). East Germans and 'aussiedler', East European labor migrants, asylum seekers and civil war refugees from various places compete with the major former migrant groups. The new quality of immigrants from far abroad also influences their position in the present labor market which is not in a very favorable condition anyway. Whereas according to ius sanguinis 'aussiedler' as ethnic Germans have not been subject to any legal labor market restrictions at all, other migrant groups have been confronted with severe obstacles to find a job.

There is a list of so-called 'guest worker'-items in ALLBUS. These questions are among those ALLBUS-items which have been used for research most frequently (e.g. ***Krauth*** and ***Porst*** 1984; ***Gehring*** and ***Böltken*** 1985; ***Blank*** and ***Schwarzer*** 1994; ***Blank*** and ***Wasmer*** 1996; ***Kühnel*** and ***Terwey*** 1990, 1994; ***Terwey*** 1989). In the following analyses, primarily four of these guest worker items are used as indicators of ethnocentrism (cf. ZA and ZUMA 1996, 1997):

(1) Guest workers (after 1994 Split 2: foreigners living in Germany) should adapt their life style a little better to that of Germans (ADAPTION)

(2) Guest workers (after 1994 Split 2: foreigners living in Germany) should be sent back home when there is a shortage of jobs in Germany. (REMIGRATION)

(3) Guest workers (after 1994 Split 2: foreigners living in Germany) should be forbidden to participate in any political activities. (NO POLITICS)

(4) Guest workers (after 1994 Split 2: foreigners living in Germany) should choose their marriage partners among their countrymen. (ENDOGAMY)[3]

Answers were classified on a seven-point scale: 1. Don't agree at all...7. Agree completely. Moreover, there are four other items dealing with different types of personal contacts with guest workers or foreigners living in Germany. These contact items will later be used in analysis.

---

3 (1) Gastarbeiter (ab 1994 Split 2: die in Deutschland lebenden Ausländer) sollten ihren Lebensstil ein bißchen besser an den der Deutschen anpassen.

(2) Wenn Arbeitsplätze knapp werden, sollte man die Gastarbeiter (ab 1994 Split 2: die in Deutschland lebenden Ausländer) wieder in ihre Heimat zurückschicken.

(3) Man sollte Gastarbeitern (ab 1994 Split 2: den in Deutschland lebenden Ausländern) jede politische Betätigung in Deutschland untersagen.

(4) Gastarbeiter (ab 1994 Split 2: Die in Deutschland lebenden Ausländer) sollten sich ihre Ehepartner unter ihren eigenen Landsleuten auswählen. (cf. ZA and ZUMA 1997)

Growing ethnocentrism in West Germany?

Many political discussions and media-reports have pointed out a growing salience of the immigration issues and a restrengthening of ethnocentrism in present Germany. In the aftermath of economic recession, globalization, worldwide migration and abdication of the socialist regime, several ethnic and even religious conflicts may possibly become more portentous. Muslims appear to be important candidates for out-groupings in a world viewed as sliding into a 'clash of civilizations', where religious and ethnic differences do visibly coincide. In spite of all, a lot of German media coverage, declarations by church officials and many political statements are still aimed at promoting further integration.

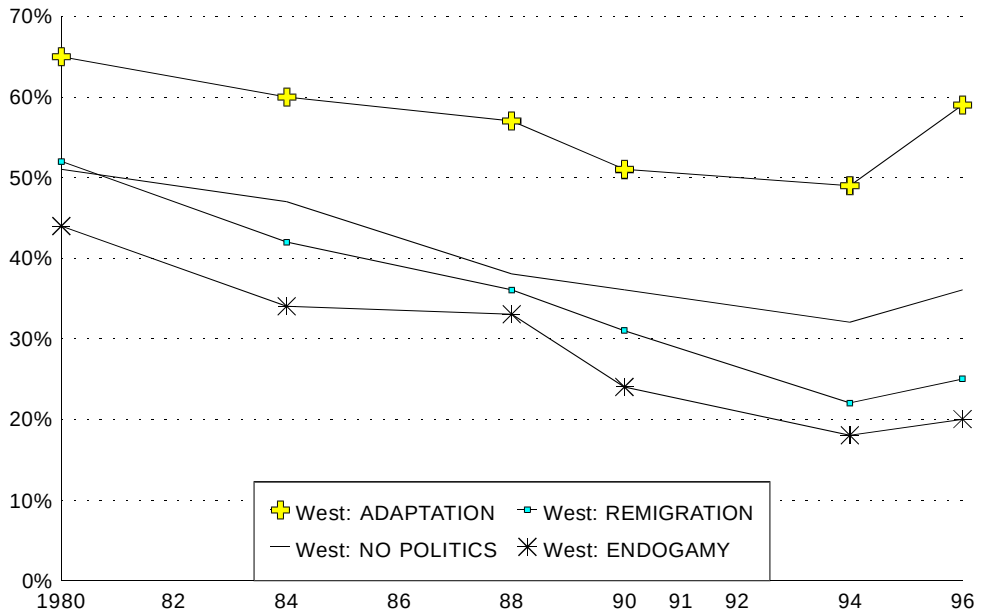
The increasing share of foreign residents in Germany since the 'economic miracle' (*Wirtschaftswunder*) in the 60s has to be recognized as an important demographic change (cf. ***Rudolph*** 1996; ***Fassmann*** and ***Münz*** 1996), which has influenced certain areas more than others. Some selected data from the Federal Bureau of Statistics (Statistisches Bundesamt 1997) may illustrate this:

- In West Germany the proportion of foreigners has risen considerably from 1.2% in 1961 to 10.2% in 1995 (1989: 7.7%)
- In post-socialist East Germany the shares of foreigners remained at notably lower levels of 1.2% in 1989 and about 1.7% in 1995
- In urban areas, the proportion of foreigners is much higher (e.g., Frankfurt (1994): 29.2%, Mannheim 21.1%, Cologne: 19.4%, Hamburg: 15.4%, Berlin: 12.6%)
- All in all, compared to other West European countries Germany has an especially high percentage of foreign residents (about 40% of all foreigners in 10 West European countries live in Germany followed by France with about 21%; ***Fassmann*** and ***Münz*** 1996: 39). This high share of foreign residents may be in part attributed to legal obstacles of civil integration (ius sanguinis).

According to the Statistisches Bundesamt (1997: 40) most foreigners came to the former Bundesrepublik (BRD) in search of employment. This seemed to be appropriate in a prospering post-war economy. In the 'Wirtschaftswunder' some connections between the ethnic, nationalist or religious out-groupings and economic issues seemed to lose priority. The guest workers were often considered to do hard jobs thought to be undesirable for Germans.

As far as our four items for ethnocentrism are concerned, the relationships between the four series remain similar since 1980 (cf. Chart 1). All four indicators show decreases of ethnocentric statements in West Germany until 1994. Considering media coverage and political discussions on that topic, this result seems somewhat surprising. A probably dwindling proportion of the German population may have caused a strong impression of xenophobia by dramatic articulations of hostility(cf. ***Leenen*** 1995).

CHART 1: Support for ADAPTATION, REMIGRATION, NO POLITICS, and ENDOGAMY with regard to foreigners or guest workers



West German respondents (5 - 7 in a 7-point-scale (ALLBUS 1980 - 1996))

However, from 1994 till 1996 the ethnocentric intolerance amongst ALLBUS respondents becomes more frequent again.[4] The 'softer' the indicator, the more it seems to yield a rising tendency of criticism towards immigrants:

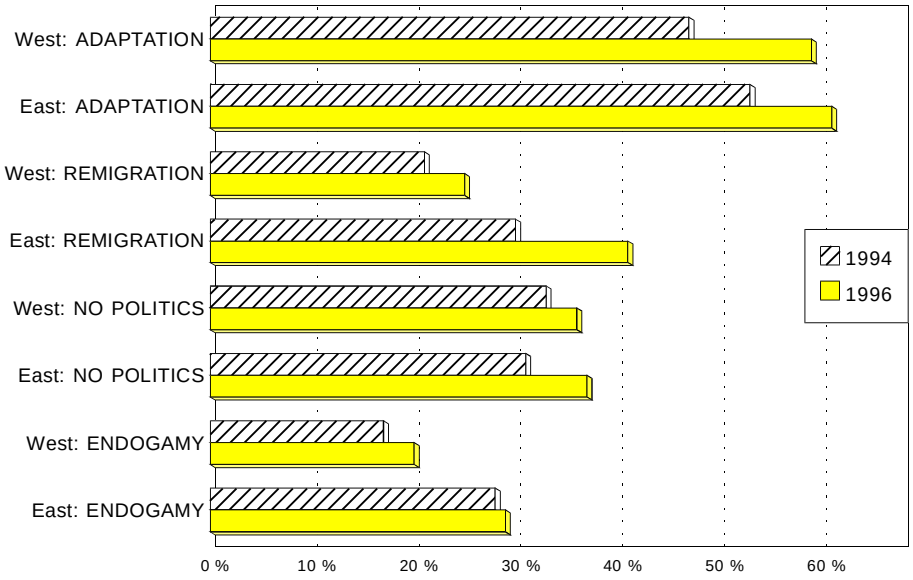
- ENDOGAMY has been supported the least and thus stands for a fairly strict kind of exclusion. The rise from 1994 to 1996 is only two percentage points (18% vs. 20%).
- Since 1984 the demand for REMIGRATION in case of scarce jobs with regard to frequency follows right after endogamy in second place (1994: 22%, 1996: 25%).
- Prohibiting political activities of foreigners has been supported with similar frequencies till 1990. After that NO POLITICS seems to take a slightly different trend than ENDOGAMY and REMIGRATION. More demand for control of political activities may have been caused by some striking actions of small radical groups in Germany (1994: 32%, 1996: 36%).
- Most frequent, however, is the demand for ADAPTATION of life-style as can be seen from Chart 1. The recent increase from 49% to 59% in this somehow soft indicator of ethnocentrism is particularly obvious, but wishing such adaptation may also express some kind of readiness to integrate foreigners under certain conditions.

4 A slightly increasing intolerance of foreigners can also be seen in a time series of another German survey, the 'Politbarometer' (ZA-Nr. 2391).

A basic comparison of ethnocentrism in East and West Germany

Compared to the Old Federal States, Eastern Germany shows at least a similar increase in ethnocentrism from 1994 to 1996.[5] However, the overall level of ethnocentric inclinations is even higher in East Germany though fewer foreigners do really dwell there (cf. Chart 2). Ethnocentrism does not necessarily have to correlate with the number of foreigners living in a country. A small out-group which is viewed as unwelcome, or even a reference group that is hardly present at all, can attract a lot of aversion or serve as an aim for negativity caused by social problems in a transforming society. Particularly, the demand for remigration, when the labor market is tight, has increased strongly from 30% to 41% in the New Federal States. We may presume that in an economic situation often perceived as falling below a proper standard the criticism of outlandish minorities is an outlet of negative feelings. This finding may be seen as some kind of support for competition-based theories of ethnocentrism (cf. additional results in: ***McCutcheon*** 1998). But on the other hand, a constructive ideology providing positive values for integration in the sense of ***Durkheim*** may also be lacking. We shall see later that in ALLBUS 1996 some direct evaluations of individual deprivation do not perform as well in detailed analyses of ethnocentrism as it could have been expected ex ante.

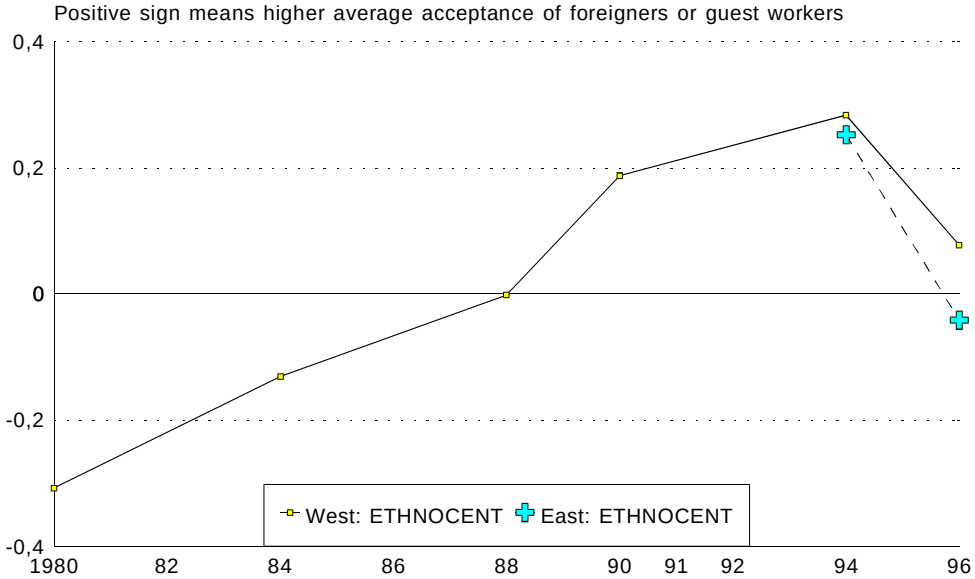
CHART 2: Support of ADAPTATION, REMIGRATION, NO POLITICS, and ENDOGAMY with regard to foreigners in Germany



West and East German Respondents (5 - 7 in a 7-point-scale (ALLBUS 1994, 1996))

5 For 1994 only the survey split that used the wording 'foreigners' has been included in the East-West comparison shown in Chart 2.

CHART 3: Ethnocentrism (ETHNOCENT) in Germany since 1980 (average PCA-Scores for ADAPTATION, REMIGRATION, NO POLITICS, ENDOGAMY combined)



East and West German respondents (ALLBUS 1980 - 1996)

Notwithstanding some singularities of meaning and measurement, all four indicators can be reduced into *one single principal component* or factor by exploratory analysis. This is a pretty stable result for each year of measurement (cf. also a comparison of confirmatory factor analysis and latent class analysis in: ***Kühnel*** and ***Terwey*** 1990). The average scores of a principal component analysis (PCA) across all ALLBUS-data available are shown in Chart 3. Again, the resulting latent variable discloses an increasing acceptance of foreigners up to 1994 and a decline after that. However, the average in 1996 indicates still less ethnocentrism than in 1980, when our measurements were started.

After introducing our basic indicators of ethnocentrism and the latent common component extracted to reduce some of the complexities, we shall proceed with a brief outline of structural differentiations which we expect to be related with ethnocentrism.

Secularization, social disintegration and enlightenment: The case of Germany

Youth in search of positive and negative self-identification is often in the vanguard of social change. Younger people lean towards convictions which they regard as progressive. Since the late 60s, especially the better educated parts of the younger generation have stressed human rights and a continuing enlightenment which ought to free us more and more from uncomprehended traditional forces, which have been considered to be repressive.

Today, traditional churches in Germany are still confronted with a severe loss of members. This withdrawal from traditional organizations proclaiming a Christian worldview is especially significant among young and better-educated people. A lowering percentage of churched people in the general population may be seen as an important indicator of a trend towards increasing secularization in Western Europe. The question is open, whether secularized Europeans replace Christian attitudes by morals which are even closer attached to those human rights protecting outsider minorities or ethnic groups.

Christian adherence was frequently correlated with the acceptance of traditional values in society (***Durkheim*** 1951; ***Inglehart*** 1998; ***Pappi*** and ***Terwey*** 1982). Yet today, churches seem to display a somewhat ambivalent attitude towards the acceptance of liberal human rights. First, there is a large part of conservative church members, who may be more alienated towards immigrants than the more secularized parts of the population. They could even claim some of the ethnocentric parts in the Old Testament as a foundation. Secondly, some Christian representatives play an important role in the public declaring their care for disadvantaged 'neighbors' in this world (cf. the concept of a tolerant community introduced above). The latter position may presently be seen as more corresponding to the commandments of Jesus Christ and his apostles than crusading attitudes and parts of the Old Testament. At first sight, the prevalent general teaching of church in Germany has changed from more rigid and intimidating aspects towards expressions of hope and charity (***Ebertz*** 1993; ***Greeley*** 1995; ***Terwey*** 1993). Such differing attitudes within churches may be correlating with simple demographic disparities. Actually, the older Christians are often expected to be more conservative and ethnically oriented than the younger ones, and we formulate an expectation for members in the bigger German churches (Volkskirchen): *Controlling for age, Christians organized in the predominant churches should be more tolerant of ethnic minorities than non-Christians.*[6]

Anyway, another determinant of ethnic tolerance has to be considered in this context. The effects of secularization as decreasing influence of mainstream-churches would be leading to less liberal ethnic orientations, if they are not counteracted by the main-force of enlightenment, i.e. rational *education*. If maximizing of subjective demands is based on sound material security, it may be expected to have a special stronghold among people with higher credentials and in the superior white collar milieu, which is on average materially rather secure and well-equipped (***Pappi*** and ***Terwey*** 1982). Educational attainment or certified *human capital* is somehow a meritocratic backbone in the white-collar hierarchy and

6 A striking feature is the extraordinarily high degree of unchurched people in Eastern Germany. This result is unparalleled in many other ex-socialist countries for which we have survey data (cf. ***Greeley*** 1995; ***Terwey*** 1998). If our optimistic expectations about reduced ethnocentrism amongst church members are appropriate, the radical secularisation in the former GDR could explain much about the important role that ethnocentrism plays in the New Federal States.

for various reasons also in the milieu of self-employed (cf. ***Terwey*** 1987). There is an important interplay of social class and education in several respects, which is expected to go beyond the sphere of obvious pecuniary inequality. On the other hand, formal education conveys also individual culture, erudition or *'Bildung'* in the sense of German 'Neuhumanismus' and rational *enlightenment*. Both components of education may result in rationalistic substitutes for traditional Christian persuasions (cf. ***Meulemann*** 1985, 1996). Modern education is overtly not turning people to some kind of heathenish transcendentalism and the component of Christian teaching in school is apparently not dominant. We can clearly begin to see the need for a multivariate model as a check for some crossing effects: *People with higher education prove true to the ethics of liberal enlightenment and shun ethnocentric prejudice.*

Table 1 provides us with introductory information about relations between church attachment and ethnocentrism while controlling for some relevant socio-demographic differences. An initial impression of aggregate differences between *East and West Germany* has already been given by Chart 2 and 3, but now we can also see how much individual data variation is explainable at the individual level. The multiple classification analysis (MCA) shows a moderate bivariate effect of eta=.11 which is slightly reduced by the control of other variables also listed in Table 1 (beta=.09). On average the New Federal States deviate .22 or .18 respectively from the Grand Mean (.00) towards higher ethnocentrism.[7]

Corresponding to our expectations, we recognise that *age of the respondents* (eta=.36, beta=.28) is in fact a very important correlate of differential ethnocentrism. Older people are considerably more reluctant about foreigners in Germany. Our assumptions are also definitely confirmed as far as the role of higher *education* is concerned. The impact of education (eta=.41, beta=.24) is similarly strong as the estimated age-effect. Besides, people living in places of larger *community size* (i.e. administrative community size) tend to be moderately less ethnocentric (eta=.16, beta=.09).

Differential objective welfare position is indicated by *social milieus* in Table 1 which can be constituted by direct or indirect relation to the occupational system (cf. ***Geiger*** 1967; ***Terwey*** 1987; 1990).[8] Expectations are confirmed that people in the rather secure and materially well provided milieu of the upper white collars are much less ethnocentric than the farmers and the lower blue collars. A rather high bivariate effect of eta=.29 is substantially reduced by the statistical controls (beta=.12) as levels of education and community size are

7 A gender variable is omitted in Table 1. There is no relevant gender difference as far as this kind of ethnocentrism is concerned.

8 Incomes of the respondents and of their households were also tested as measurements of material welfare. Both types of income showed significant bivariate effects. However, the multivariate impacts proved to be comparatively small.

Table 1: Multiple Classification Analysis for demographic predictors of ethnocentrism factor (ALLBUS 1996, results weighted to adjust the Eastern oversample; positive signs mean higher ethnocentrism)

	n	Unadjusted deviation	eta	Adjusted for independents deviation	beta
Part of Germany					
1 Old Federal States	2454	-,06		-,05	
2 New Federal States	640	,22		,18	
			,11		,09
Age					
1 18-29 years of age	548	-,44		-,35	
2 30-44 years	949	-,27		-,20	
3 45-59 years	845	,12		,10	
4 60 years or more	751	,53		,39	
			,36		,28
Education					
1 Primary education	287	,72		,43	
2 Primary & vocational	1203	,28		,14	
3 Medium Education	901	-,13		-,05	
4 Higher Education	339	-,53		-,26	
5 University, college	364	-,66		-,44	
			,41		,24
Community Size					
1 Less 2000 inhabitants	239	,32		,19	
2 2000-4999	325	,18		,05	
3 5000-19999	739	,10		,07	
4 20000-49999	525	-,02		-,02	
5 50000-99999	256	-,05		-,08	
6 100000-499999	618	-,12		-,04	
7 Over 499999	392	-,28		-,14	
			,16		,09
Social Milieu					
1 Farmers	62	,58		,14	
2 Small self-employed	115	-,28		-,12	
3 Employers	140	-,19		-,06	
4 Upper white collar	672	-,42		-,19	
5 Lower white collar	1014	-,01		,01	
6 Upper blue collar	155	,12		-,02	
7 Lower blue collar	937	,33		,14	
			,29		,12
Denomination					
1 Protestants, EKD	1216	,02		-,01	
2 Other Protestants	36	,37		,03	
3 Catholics	1016	,02		,04	
4 Other religions	45	-,23		-,08	
5 No church at all	781	-,07		-,03	
			,06		,03
Church Attendance					
1 Several times a month	496	,13		-,02	
2 Several times a year	583	,00		-,01	
3 Less often	1039	-,02		,01	
4 Never	975	-,04		,00	
			,06		,01

Multiple R = ,505

to a certain degree accentuated in these milieus. Nevertheless, the impacts of more or less well provided milieus support competition-based explanations of ethnocentrism.

Subsequently, it can be realized from Table 1 that the *denominational groups* inside of Christian churches are a little more ethnocentric than the formally unchurched people. Less tolerance is at first sight especially shown by members of 'evangelische Freikirchen' ('Other Protestants') outside of the Lutheran mainstream (Evangelische Kirche Deutsch-

lands (EKD)). This group of conservative 'Freikirchen' is, however, only pretty small. At least the bivariate description of *church attendance* in unified Germany seems to reveal a systematic pattern, but this globally observed pattern is in contrast to the expectations of enhanced tolerance in the active church-communities. Moreover, the effects of Christian denomination and church attendance tend to vanish in the multivariate part of the analysis in Table 1. These findings will be tested again by several multiple regression in order to see whether the minor effects now shown may disappear completely or may become stronger again. First, we shall turn to another introductory overview of secular and political assessments.

Perceptions of economic situations, anomia and diffuse political support

The subjective evaluation of economic conditions is expected to be one of the most important reasons behind many other conceptions, attitudes and preferences. Note, that according to the rather conservative view of ***Durkheim*** (1951) economic dissatisfaction or greed may result in an anomic worldview, which is partly not rooted in material poverty, but in the stimulation of wishes that can never be completely satisfied (cf. the quotation at the beginning of our text). The aspirations in economically caused anomia exceed legitimate boundaries and impede the livability of individuals or human society.

It has already been indicated that economic dissatisfaction, envies and fears are spreading again in unified Germany up to 1996. In particular, the evaluation of Germany's overall economic situation deteriorated dramatically after a very positive state around 1990. At first sight, Eastern and Western Germany are even very similar in this respect – a phenomenon that would deserve more attention in future analysis. A first step in this direction is a comparison between the assessments of the German economy and of the respondents' own economic situation. These do frequently diverge.

(a) The evaluation of the general situation is a rather volatile opinion on social policy in general. It expresses diffuse political unease, which must not correspond to a real economic scarcity.

(b) The evaluation of the respondent’s own economical situation is generally more stable and frequently connected to the objectively measurable welfare of the respondents (cf. ***Terwey*** 1990, 1997).

Either because of diffuse economical fears or because of subjective material deprivation the acceptance of ethnic strangers may be low. The most important example in German history may be found in the period of the worldwide economic crisis at the end of the 20s, which facilitated the expansion of National Socialist ideology and an anti-Semitic conspiracy theory (cf. ***Popper*** 1974: 123). All in all, our expectations are: *Negative assessments of the general economic situation in Germany are positively related with ethnocentrism. For different reasons, similar expectation may also be appropriate for the effects of the assessments of personal economic situations.*

It has already been shown by empirical analyses that several attitudinal trends in Germany have changed since 1990 (e.g. ***Terwey*** 1995, 1996, 1997). In the course of far reaching societal transformations, previously valid social ideals and habits (*nomos*) loose influence on the individuals (***Durkheim*** 1951, 1960), and a state occurs which lacks effective and positively accepted rules for living. Primarily with regard to society, these insecurities about social norms have been referred to as *'anomie'*. This concept is used with a conservative meaning. Usually, a positive preeminence of social tradition and integration is assumed, which is often beyond the scope of individual intellects and wants. Those individuals who positively accept their given society tend towards '*eunomia*' (***Srole*** 1956). It may be easily understood that a fundamental characteristic of such eunomia is a general social optimism and social attachment. Other individuals who feel that important cultural goals are devalued in society or that there are no more legal means to reach these goals probably come to a pessimistic interpretation of their contemporary society. These people tend towards individual '*anomia*' (***Srole*** 1956; ***Merton*** 1970; ***Friedrichs*** 1997).

If there is some kind of lasting dissatisfaction about society which cannot be traced to easy causes and the solution of the problems appears to be difficult, we all may tend to look for some scapegoat to put the blame on. Racist organizations could provide people in loss of appropriate community with a pseudo-communal surrogate. Therefore, we may expect a positive link between anomia and opposition against some social minorities (cf. ***Srole*** 1956; ***Heitmeyer*** 1997).

We examine anomia as some kind of social pessimism about future and interpersonal commitment using four items from ALLBUS (ZA and ZUMA 1997):

1. No matter what some people say, the situation of common people is not getting better, but worse (WORSE SITUATION).
2. It is hardly fair to bring children into the world with way things look for the future (NO CHILDREN).
3. Most politicians are not at all interested in the problems of common people (CARELESS POLITICIANS)
4. Most people do not really care, what happens to their fellow men (CARELESS MAJORITY).[9]

9 (1) Egal, was manche Leute sagen: Die Situation der einfachen Leute wird nicht besser, sondern schlechter.
(2) So wie die Zukunft aussieht, kann man es kaum noch verantworten, Kinder auf die Welt zu bringen.
(3) Die meisten Politiker interessieren sich in Wirklichkeit gar nicht für die Probleme der einfachen Leute.
(4) Die meisten Leute kümmern sich in Wirklichkeit gar nicht darum, was mit ihren Mitmenschen geschieht.

CHART 4: 'Birth of children is hardly fair with this future' and negative assessments of the German economy in general

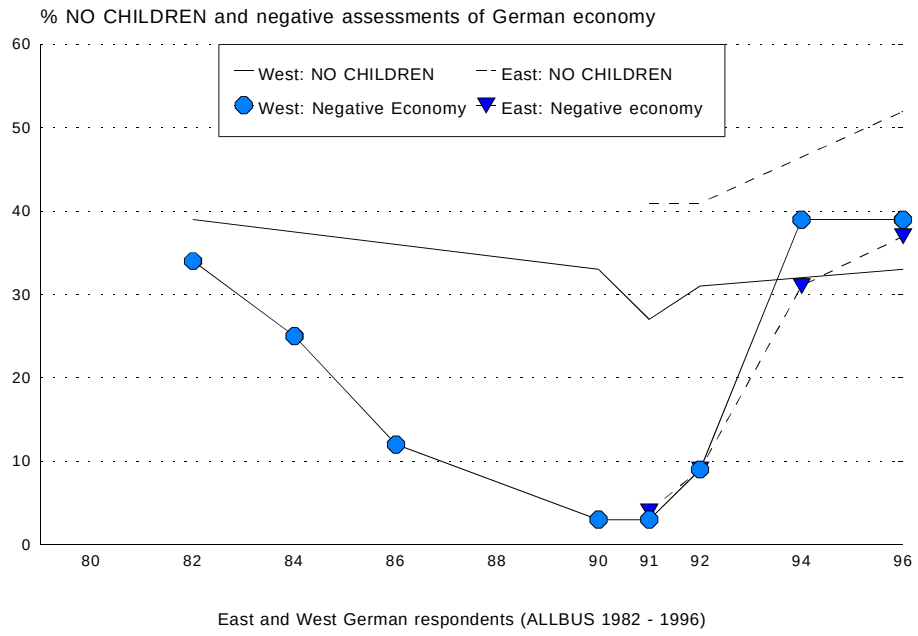


Chart 4 shows recently growing percentages of people who agree with the indicator NO CHILDREN as an example of recent changes in anomia. We may add that NO CHILDREN is the statement which attracts least consent among the four ALLBUS-indicators and reveals, hence, a deeper kind of discouragement. Comparing different years in Chart 4, this profound pessimism seems to be a little bit less frequent in 1991, but spreads afterwards especially in East Germany. On the other hand, the percentages of persons who describe the general economic situation in Germany as bad have grown even more since 1991. Both outlooks seem to covariate only moderately with each other, as far as it can be derived from the aggregated percentages in Chart 4. The other three indicators of anomia, not included in Chart 4, show steeper rising tendencies during the 90s after a slight betterment had eventuated before in Western Germany (cf. ***Terwey*** 1997). In general, anomia and intolerance or out-grouping may be expected to correlate. Our expectation reformulated is: *Anomia is positively correlated with ethnocentrism.*

Table 2: Multiple Classification Analysis for economy assessment, anomia and diffuse political support as predictors of ethnocentrism factor (ALLBUS 1996, results weighted to adjust the Eastern oversample)

	n	Unadjusted deviation	eta	Adjusted for independents deviation	beta
Part of Germany					
1 Old Federal States	1984	-,06		-,04	
2 New Federal States	512	,25		,14	
			,13		,07
WORSE SITUATION					
0 Have other opinion	362	-,15		,05	
1 Have same opinion	2134	,03		-,01	
			,06		,02
NO CHILDREN					
0 Have other opinion	1423	-,21		-,15	
1 Have same opinion	1073	,28		,19	
			,24		,17
CARELESS POLITICIANS					
0 Have other opinion	428	-,38		-,23	
1 Have same opinion	2068	,08		,05	
			,18		,11
CARELESS MAJORITY					
0 Have other opinion	536	-,14		,01	
1 Have same opinion	1960	,04		,00	
			,07		,01
Satisfaction with German democracy					
1 It functions well	236	-,10		,04	
2 Only little change	1660	-,09		-,05	
3 Don't know	77	,44		,32	
4 Much to change	463	,24		,08	
5 Completely to change	61	,50		,13	
			,17		,08
German economic situation					
1 Very good	23	-,39		-,20	
2 Good	305	-,14		-,03	
3 So-so	1158	-,09		-,05	
4 Bad	833	,10		,04	
5 Very bad	177	,41		,21	
			,15		,07
Personal economic situation					
1 Very good	56	-,62		-,44	
2 Good	1221	-,10		-,03	
3 So-so	965	,09		,02	
4 Bad	190	,24		,10	
5 Very bad	64	,49		,24	
			,17		,08

Multiple R = ,322

The MCA-model in Table 2 shows the bivariate correlations and multivariate adjusted effects while the basic East-West-split is included again. Each of the four anomia items discloses higher average ethnocentrism in the group of assenters. A systematic difference is especially found for political alienated ones (CARELESS POLITICIANS; eta=.18, beta=.11) and existential pessimists (NO CHILDREN; eta=.24, beta=.17). The notions of a CARELESS MAJORITY and of a WORSE SITUATION for common people appear to be more unsystematic in this respect. However, our expectations about anomia as a background for ethnocentric out-grouping receive some support.

Satisfaction with the democratic system is also closely related to the expansion of anomia. If we do a factor analysis with the four anomia items in ALLBUS, one single common factor results and the indicator CARELESS POLITICIANS turns out with the strongest loading. Still we should introduce another indicator which has less specific connotations about an assumed misconduct of politicians. Judging the overall *performance of a democracy* is often considered to be an expression of deeply rooted, *diffuse political support* (***Easton*** 1979; ***Terwey*** 1996). A lack of diffuse political support could also be considered to be a supplemental indicator for social disintegration. In fact, those respondents who ask for a change of our democratic system are rather often more ethnocentric than the positively supporting citizens (eta=.17, beta=.08).

Following ***Durkheim*** (1951), subjective economic dissatisfaction and greed are basically linked with social anomie or subjective anomia and it has been pointed out that ethnocentrism and racism have some kind of individual economic background (cf. amongst others ***Krauth*** and ***Porst*** 1984; ***Küchler*** 1994; ***Steußloff*** 1997). To some surprise, results in Table 2 show that subjective assessments of economy have noticeable, but not very strong impacts. Positive perceptions of the general and personal situation are both at least moderately associated with more tolerance (eta=.15 or .17). In our first MCA, the measurement of social inequality by structural milieus was closer related to the varying acceptance of foreigners (cf. eta=.29 in Table 1).

'Postmaterialism': Shifting political preferences

The work of diverse authors on the shift from 'materialist' to 'postmaterialist' preferences has attracted a worldwide attention. A common context of anomia and such value change was proposed by ***Friedrichs*** (1997). Postmaterialist aspirations, which seemed to be safe and sound, but presently fail to meet reality, may be one of the causes for dwindling self-security and growing anomia.

Even though the thesis of flourishing postmaterialism has stimulated the analysis of contemporary societies, it has often been criticized. The author of this article claims almost no

personal interest in the battered issue whether the measurements of postmaterialism are restricted to mere attitudes or reach deeper to the 'real personal values'. The simple index introduced by ***Inglehart*** (1971) can provide – at the very least – an interesting onset for describing political priorities which are in a broader sense to be understood as strains between 'postmodern' worldviews, even though many of ***Maslow***'s ideas remain distant to much of the research since ***Inglehart*** (1971).[10]

In the simple index of ***Inglehart***, postmaterialists are those for whom 'influence on governmental decisions' and 'free expression of opinions' have priority over 'law and order' and 'fighting rising prices'. ALLBUS-data confirm assumptions that West German postmaterialism still increased after a striking period of prosperity (Wirtschaftswunder). The economic recessions which occurred at the end of the sixties and the mid-seventies were quickly overcome to a large extent. Even the pessimistic expectations of many people at the beginning of the eighties were reduced after the new coalition of CDU/CSU and F.D.P. took charge of politics in the economic sphere. However, the social expansion of postmaterialism did stop after the end of the 80s. As the economic burdens increased after German unification, a noticeable re-decline was observed (***Terwey*** 1995). Besides, the New Federal States, which are still in a relatively disadvantageous economic situation, have shown considerably lower levels of postmaterialism than the West.

As postmaterialists according to the simple Inglehart-Index have a strong interest in democratic participation, humanitarian considerations and personal freedom, they may also be supposed to have a positive interest in an universal application of human rights. Therefore, we expect: *People who are described as postmaterialists have less tendency toward ethnic segregation.* Table 3 investigates the relation between ethnocentrism and postmaterialism while some other political attitudes are additionally controlled.

10 According to ***Inglehart***'s interpretation of ***Maslow*** (1954) about a hierarchy of human goals, cohorts who had experienced wars and economic scarcities in the phases preceding the 'economic miracles' (Wirtschaftswunder) would ceteris paribus put relatively high priority on economic security and safety needs. Postmaterialism – flourishing in cohorts who experienced high material affluence – aims at intellectual needs, aesthetic pursuits, humanitarian considerations, and good opportunities for personal self-expression. Amongst others, it is expected to replace most traditional religious orientations in a modern affluent society. Only a part of these assumptions have been examined by empirical analyses.

Table 3: Multiple Classification Analysis: Political preferences predicting ethnocentrism factor (ALLBUS 1996, results weighted to adjust the Eastern oversample)

	n	Unadjusted deviation	Unadjusted eta	Adjusted for independents deviation	Adjusted for independents beta
Part of Germany					
1 Old Federal States	2213	-,06		-,05	
2 New Federal States	591	,22		,18	
			,11		,09
Inglehart-Index					
1 Postmaterialists	648	-,61		-,36	
2 PM-Mixed Type	815	-,07		-,06	
3 Material.-Mixed Type	954	,27		,17	
4 Materialists	387	,50		,32	
			,38		,23
Left-right-selfplacement					
1 Left	57	-,32		-,21	
2	114	-,47		-,23	
3	316	-,41		-,23	
4	351	-,31		-,18	
5	774	-,03		-,05	
6	609	,22		,13	
7	303	,17		,08	
8	171	,47		,37	
9	50	,53		,41	
10 Right	59	,78		,44	
			,30		,18
Government spending for defense					
1 Spend much more	44	1,01		,60	
2 Spend a little more	142	,53		,27	
3 Keep it as it is	680	,20		,10	
4 Spend less	1072	-,11		-,06	
5 Spend much less	866	-,16		-,07	
			,23		,12
Government spending for police					
1 Spend much more	596	,38		,15	
2 Spend a little more	1110	,05		,02	
3 Keep it as it is	951	-,23		-,09	
4 Spend less	125	-,44		-,19	
5 Spend much less	23	-,65		-,31	
			,25		,10
Government spending for pensions					
1 Spend much more	371	,41		,25	
2 Spend a little more	924	,09		,05	
3 Keep it as it is	1379	-,15		-,09	
4 Spend less	120	-,27		-,17	
5 Spend much less	11	-,09		,09	
			,20		,12
Government spending for environment					
1 Spend much more	603	-,31		-,20	
2 Spend a little more	1005	-,03		,01	
3 Keep it as it is	1030	,12		,06	
4 Spend less	153	,58		,34	
5 Spend much less	13	,59		,10	
			,22		,13

cont'd

Table 3 *(continued)***:**

	n	Unadjusted deviation	Unadjusted eta	Adjusted for independents deviation	Adjusted for independents beta
Government spending for culture and art					
1 Spend much more	79	-,09		-,15	
2 Spend a little more	368	-,18		-,14	
3 Keep it as it is	1296	-,05		-,03	
4 Spend less	787	,06		,05	
5 Spend much less	275	,36		,25	
			,14		,10

Multiple R = ,521

Indeed, postmaterialism reveals a remarkable effect (eta=.38, beta=.23), which is even comparable to the effects of education and age, the strongest predictors introduced yet (cf. Table 1). Another strong impact is shown for the left-right-scale, which serves as a general medium for the expression of many political issue and conflict positions. Hence, we try to isolate a left-right component which is probably also included in the Inglehart-Index, and look at the multivariate results for both predictors: Ethnocentrism is also strongly related to left-right-placement with eta=.30 and beta=.18. People who perceive themselves as belonging to the political right wing still tend to be more ethnocentric.

Our conclusion is in this respect that a general left self-orientation provides some people with an important belief-system which fosters pro-immigrant sentiments. Additional separate regressions for East and West Germany show that tolerance is significantly influenced by this ideology in both areas.

Nevertheless, differences between postmaterialists and materialists are at least of comparable relevance as left-right-ideology. It has to be admitted, of course, that the analysis of postmaterialism will profit from the inclusion of further qualifying indicators. The Inglehart-Index contrasts two indicators of preference for democratic influence and liberty with one material issue (inflation) and one security issue (law and order). Priorities for ecology or general culture are at least explicitly neglected in the simple Inglehart-Index. Table 3 shows in addition five indicators of further political preferences. People were asked, whether they would like to increase the amounts of money spent by the government for certain areas. These spending options can help to go a little beyond the somewhat incomplete index-measurement of far-reaching ideals presented to us under the headline 'postmaterialism'. Security issues are often connected with traditional political orientations which are attributed to 'materialism'. According to some results in Table 3, the safety oriented preferences to spend more money for defense and police are significantly connected with higher ethnocentrism. The same tendency is disclosed for the people asking more money to be spent on old-age pensions. On the other hand, postmaterialist options for more investment in conservation of environment and culture or art are corresponding with higher

acceptance of foreigners. Thus the findings for the simple Inglehart-Index in Table 3 are supported and extended in further details. Ethnic tolerance has some significant correlations with so-called postmaterialistic or non-materialistic priorities.

Reported contacts with foreigners and religious tolerance

Besides the four 'guest worker'-items introduced already for the measurement of ethnocentrism, there are four additional items in ALLBUS about *personal contacts* with foreigners living in Germany (ALLBUS 1994/split 1 and 1996) or up to ALLBUS 1994 with guest workers respectively. These questions deal with four different social spheres: FAMILY, WORKING PLACE, NEIGHBORHOOD and FRIENDS (including acquaintances). We shall now investigate these contacts as potential sources for the acceptance of auslander. It is often expected by open-minded scientists and cosmopolitans that the more inter-ethnic contacts are reported, the less out-classing is involved. Mutual knowledge may diminish negative prejudice (***Rippl*** 1995), or people who are in favor of foreigners may be readier to acknowledge even transient forms of company as personal contacts. It is to some extent plausible that contacts often foster better understanding and acceptance: Familiarity breeds consent. But putting this assumption to a test would not be as unimportant as it may appear at first sight. Note the assumptions about tolerance increased by everyday contacts could fail in practice. Such a reverse influence of probably involuntary contacts may be found in many social settings.[11] A finding that ‘familiarity breeds contempt’ would indicate some pretty serious problem for future integration of minorities.

Overall, MCA-results in Table 4 show that people who reported personal contacts with foreigners living in Germany tend to be less ethnocentric. Only a smaller part of the respondents in ALLBUS may have drawn ethnocentric conclusions from these social experiences or have remained intolerant. The strongest predictor among the four contact items is the experience amongst personal friends or acquaintances (eta=.35, beta=.26). The multivariate effect of the East-West-split almost disappears. East Germans report much less personal communications with foreigners, and a larger part of their relative distance towards auslander is statistically correlated with this.

11 ".. experimental studies have .. concluded that only when intergroup contact occurs under highly auspicious circumstances should favorable attitude change be anticipated. ... In nonexperimental context, a major nationwide survey ... found that the racial views of whites whose circle of friends included a black person closely paralleled the views of whites with no black friends ..." (***Sigelman*** and ***Welch*** 1993: 782).

Table 4: Multiple Classification Analysis for contact items and religious teaching preferences as predictors of ethnocentrism factor (ALLBUS 1996, results weighted to adjust the Eastern oversample)

	n	Unadjusted deviation	eta	Adjusted for independents deviation	beta
Part of Germany					
1 Old Federal States	2507	-,06		,01	
2 New Federal States	654	,21		-,05	
			,11		,02
Contacts with foreigners: Family area					
0 No contacts	2644	,06		,02	
1 Yes, have contacts	517	-,33		-,10	
			,15		,04
Contacts with foreigners: Working area					
0 No contacts	1933	,21		,12	
1 Yes, have contacts	1228	-,33		-,19	
			,26		,15
Contacts with foreigners: Neighborhood					
0 No contacts	2185	,11		,02	
1 Yes, have contacts	976	-,25		-,04	
			,16		,03
Contacts with foreigners: Friendship					
0 No contacts	1791	,31		,23	
1 Yes, have contacts	1370	-,40		-,30	
			,35		,26
Preference for religious teachings					
1 Islam lessons too	1042	-,32		-,24	
2 Only Christian lessons	1015	,37		,29	
3 No religious lessons	1104	-,03		-,04	
			,28		,22

Multiple R = ,447

It has to be regrettably acknowledged that we could not yet include more indications of substantive religious beliefs or attitudes in our investigation. But ALLBUS 1996 has at least one additional indicator for the personal opinion about *religious pluralism and tolerance*. The respondents were asked whether they wish that Islam should be taught in German schools as an alternative to Christian teachings. Two further options given in this item are: no religious teaching at all or only Christian lessons.[12] Noticeably, this predictor performs rather well (eta=.28, beta=.22). An exclusive preference for Christian teaching is correlated with latent ethnocentrism and thus this conventional option reveals some intol-

12 The demand for a general restriction of religious teaching is not necessarily an indicator for secularism in general – tolerant secularists would neither mind the opportunity of Christian teachings, nor of Muslim teaching in school. One could personally reject theistic belief, but accept the religious ethics as a valuable choice.

erance. Those secularization supporters who want no religious teaching at all deviate only marginally from the Grand Mean towards slightly reduced ethnocentrism.

An expanded explanation of ethnocentrism

We have introduced some basic socio-demographic differentiations, social group affiliations and a sequence of subjective articulations about the state our society is in or should be in. The results from the separate MCA-tables shall now be tested in comprehensive competition by regression analysis. In addition, extensions of this multi-faceted, but global model will be briefly discussed afterwards.

Amongst socio-demographic predictors, *age* (beta=.16) and *education* (beta=-.13), remain most important in Table 5. The difference between New and Old Federal States is now overtaken completely by the explanatory power of other indicators in this regression. Furthermore, the different levels of ethnocentrism found in various *social milieus* (cf. Table 1) are reduced to a moderate effect for upper white collars (r=-.23, beta=-.03, p=.0779). The size of the place in which respondents live is another structural distinction with significance. Inhabitants of larger towns tend to be more tolerant (r=-.17, beta=-.06, p=.0003). Again the thesis that familiarity breeds contempt is not supported. We have already noted above that many large German cities have particularly high shares of foreigners.

All in all, a statement that socio-demographic differentiations have only moderate effects on ethnocentrism (***Küchler*** 1996: 260) is not confirmed for Germany by our analysis of ALLBUS 1996. Further research will be needed to find out whether this divergence is caused by using different indicators of ethnocentrism or by differently refined indicators of socio-demographic characteristics in international data. We still have to admit that church membership or church attendance seem to be no correlates of ethnocentrism as far as our own analysis is concerned. If anything, right-winged Catholics are inclined towards ethnocentrism, but this effect can be omitted in a parsimonious model without losing much information. Nevertheless, additional investigations will show an interesting variant of this estimation in some respect (cf. below).

The general *left-right-ideology* is still indicatively attached with varying degrees of tolerance (r=.29, beta=.15). Even though its bivariate correlation is even stronger, *postmaterialism* (r=.37, beta=.10) seems to be surpassed at least a little by left-right in multivariate model after controlling for education, age and spending options. These *preferences for governmental spending* (defense, police, environment, old age pensions, culture and art) remain in the extended multivariate model with the defense question being in slight lead (r=-.21, beta=-.08, p=.0000). Overall the signs of the effects are in accordance with assumptions about the positive correlation between postmaterialistic vs. materialistic preferences and level of ethnocentrism.

Table 5: Multiple Regression Analysis with ethnocentrism factor as dependent variable; (positive sign means estimation of higher ethnocentrism; ALLBUS 1996, results weighted to adjust the Eastern oversample)

	r	b	beta	t	p
Age	,38	,010	,16	9,109	,0000
Education	-,41	-,110	-,13	-6,584	,0000
Social milieu: Upper white collar	-,23	-,072	-,03	-1,764	,0779
Size of town	-,17	-,030	-,06	-3,629	,0003
Left-Right-Selfplacement	,29	,085	,15	9,689	,0000
Postmaterialism-Materialism	,37	,102	,10	6,088	,0000
Spending preference: Less defense	-,21	-,082	-,08	-4,910	,0000
Spending preference: Less police	-,24	-,042	-,04	-2,278	,0228
Spending preference: Less old age pensions	-,20	-,060	-,05	-2,991	,0028
Spending preference: Less environmental protection	,21	,063	,06	3,507	,0005
Spending preference: Less culture and art	,13	,062	,06	3,580	,0003
Contacts: Friendship or acquaintance	-,36	-,227	-,11	-6,472	,0000
Contacts: Working area	-,26	-,070	-,03	-2,069	,0386
Contacts: Family area	-,16	-,098	-,04	-2,375	,0176
Contacts: Neighborhood area	-,17	-,052	-,02	-1,506	,1321
Religious lessons: Islam should be taught	-,23	-,159	-,08	-4,344	,0000
Religious lessons: Only Christian lessons	,25	,146	,07	3,885	,0001
Assessment of personal economic situation	,13	,052	,04	2,464	,0138
Assessment of German economic situation	,14	,035	,03	1,780	,0752
Dissatisfaction with democracy	,16	,071	,07	4,261	,0000
Anomia: CARELESS POLITICIANS	,16	,123	,05	2,960	,0031
Anomia: NO CHILDREN	,23	,159	,08	4,810	,0000
Constant		-,906		-5,434	,0000

R = .63 R^2 = .40

Another strong reason for an improved acceptance of foreigners is perceived personal *friendship and acquaintance* (r=-.36, beta=-.11). Contacts at work and in the family prove to be a little less important, while contacts in the neighborhood could even be discarded from this model (r=-.17, beta=-.02, p=.1321). If we eliminated the contact items from the regression in Table 5, R would only drop from .63 to .62. Therefore, we may conclude that these indicators are not indispensable to reach a high multiple correlation.

Still significant is religious tolerance proposed for lessons in schools. Judged by strictly numerical criteria, approved *teaching of Islam* would be the predictor of fifth statistical

importance in building a complex model via stepwise regression, surpassed only by education, age, left-right and personal friendship. In the complex regression (Table 5) it still performs pretty well (r=-.23, beta=-.08, p=.0000). A contrasting supplement is the restrictive demand for exclusive Christian lessons which is still in positive correlation with higher ethnocentrism (r=.25, beta=.07, p=.0001).

Finally, we turn to the remaining indicators of economic dissatisfaction, political support and anomia. To some surprise the *subjective assessments of economy* still do not perform very well – with the opinions of the personal situation performing perhaps slightly better (r=.13, beta=.04, p=.0138) than those of German economy in general (r=.14, beta=.03, p=.0752). These economic interpretations are correlating at least moderately with discontents expressed about the democratic system (r=.24 and r=.23). *Political dissatisfaction* as a measurement of lacking diffuse political support remains a relevant correlate of ethnocentrism in Table 5 (r=.16, beta=.07, p=.0000). Out of the four anomia items two have been retained in extended multivariate regression: CARELESS POLITICIANS and NO CHILDREN. The general anxiety about the future of our children is still more important (r=.23, beta=.08, p=.0000) than the criticism of politicians (r=.16, beta=.05, p=.0000). A connection between ethnocentrism and NO CHILDREN could even remind us of the lamentable catchphrase 'Das Boot ist voll!' (the boat is crowded) demanding more restriction on further immigration and on increase in population.

We shall finally turn to other economic assessments which we have not included in Table 5, either because of their irrelevance for our investigation or because of their close substantial relation with the dependent variable. Expectations about the future of the national and personal economy were also tested (cf. V115, V116 in ALLBUS 1996), but these proved to be almost irrelevant in our context. This finding is corroborated by the moderate influence which the economic expectation of a WORSE SITUATION for common people has shown above. What counts straightaway are *economic threats or costs directly ascribed to auslander* or *social envy* (cf. also ***Durkheim*** 1951; ***Küchler*** 1994, 1996; ***Steußloff*** 1997). Assumptions about foreigners taking away jobs, having bad influence on the housing market or encumbering our social security system are pretty relevant (ALLBUS 1996: V73, V75, V77). These somehow ethnocentric ascriptions raise the variation explained from 40% (cf. Table 5) to 52%. In a period marked by diffuse economic envies and by the weakness of a positive vision or encouraging metaphysics, foreigners are likely to be a target for more or less irrational blames growing out of real competition, vague fears and personal disappointments.

Discussion and extensions

In recent years, the climate of public opinion in Germany has changed notably. Some of the modifications concern political support, the relative importance of so-called materialistic values, economic fears and subjective social integration. Very often personal competence of politicians and other actors in our democracy are put into doubt, even though recent stages in this development should not be overemphasized (cf. ***Parry*** 1976; ***Döring*** 1990, 1992; ***Terwey*** and ***Pollack*** 1997). Nevertheless, it is likely that during phases of subjective dissatisfaction and disintegration certain social minorities (once again) will become a target of negative attributions. Actually, German survey data show a new increase of ethnic alienation. This tendency is yet not very strong, but it deserves close attention. Moreover, intercultural religious tolerance may also be considered to be deficient. While in the Old Federal States at least 40% of all respondents are in favor of teaching Islam in schools, there are only 12% to be found in Eastern Germany.

Objective income differences, social milieus and subjective assessments of economical situations may be considered as important sources of ethnocentrism, but the support we found for these assumptions was only moderate – at least after extended multivariate controls. However, if we put structural milieus and our two basic economic assessments into a simpler regression equation, they explain 10% of the variation in ethnocentrism. An additional larger part of explanation is mediated in this context by educational attainment. Education by itself explains just 5%, but adding education to the simpler regression including social milieus and subjective economic assessments raises the explanatory power from 10% to 23%.

The effect of *education* on ethnocentrism is pretty salient, but it cannot be easily interpreted: Educated people could be less ethnocentric because they have better opportunities to acquire an ethically positive view of our world (effects of *'Bildung', culture or erudition*), but one should also keep in mind that higher education increases the possibility of expressing socially accepted opinions (social desirability; cf. ***Krebs*** and ***Schuessler*** 1987). Anyhow, survey-responses according to social desirability would not completely devaluate our findings about the strong impacts of education, as long as respondents do not consciously deceive.

Another explanation of the strong education-effect is that *inter-ethnic competition in the labor market* is probably less relevant for Germans with higher credentials. Additionally, the special contacts, which better-educated Germans use to have with foreigners, could often happen in advantaged *social settings* shared by similarly educated foreigners. This could lead to more favorable perceptions. We may conclude that at least some important social circumstances of interaction have been controlled in our models (social milieu, community size, age, factual contacts reported in different contexts). A large part of social

class segmentation is partialled out before estimating the multivariate education effect in Table 5. Hence, we feel confident to some extent that a potency of higher education to reduce ethnocentrism is confirmed. Enlightenment via education with its correlates of adopting human rights is ascertained in this investigation.

Neither the ethnocentric traits of the Old Testament, nor the more benevolent compassion of the New Testament have prevailed very much amongst the majority of church members in unified Germany. Only minor differences appear between more or less conservative Christians. This general finding is quite contrary to official declarations about the Christians' stewardship for the weak and the poor. It diverges from some public actions as the 'Kirchenasyl' for particular foreign families whose remigration is proposed by political officials. In general, many contemporary mainstream Christians seem to be at least a little half-hearted about people from abroad. Those dogmatics who have a preference for exclusive Christian lessons in school show a clear tendency towards higher ethnocentrism.

A remarkable correlation between ethnocentrism and anomia was found for political alienation or lack of efficacy, as measured by CARELESS POLITICIANS, and for an existential pessimism which is expressed in the assumption that it is better to have NO CHILDREN any more. If we run *separate regressions for East and West Germany* the effects of anomia prove to be even more important in the New Federal States. The breakdown of socialist communities and the far-reaching transformation after socialism has enhanced social disintegration. Noticing that NO CHILDREN is a rather significant attitudinal correlate of ethnocentrism in East Germany, we may also recall that the basic pessimism stated through this anomia-item has particularly increased in the New Federal States (cf. Chart 4).

Another important information about church is also revealed by using such separate regressions for the Old and New Federal States. Church attendants in the East are slightly less ethnocentric than their more secularized counterparts. Furthermore, the very strong effect of age on ethnocentrism is differentiated in separate regressions. (East: r=.27; West: r=.40). Nevertheless, age is still one of the strongest predictors of ethnocentrism in the New Federal States (same variable used as listed in Table 5).

The details and determinants of these particular differences between both parts of Germany remain to be described in future research. As ***McCutcheon*** (1998) has come to some similar conclusions, using different models and indicators of ethnocentrism, the dissimilarities between the Eastern and Western Christians cannot be considered as simply spurious. They will deserve closer inspection. At least amongst the strongly diminished community of Eastern Christians tolerance seems to be fostered. Anyhow, ethnic adversities in larger parts of the West German population could still turn out worse (e.g. amongst older Germans), if the traditional churches would completely disappear from the social scene.

The notion of a basic worldview or 'Weltanschauung' designates a personal set of elementary, global interpretations concerning sense, potentials and the social sphere a person is part of. In general, 'Weltanschauung' constitutes a somewhat metaphysical reduction of a complex social reality, but it also adds some complexity, because it is always a changing selection and *supplement* of the assumed 'objective truth' (cf. the concept of 'supplément' in French sociology). Subjective views about an appropriate level of material subsistence influence priorities and identifications for some persons, while others would rather free themselves from further striving for such external objects. It may be accepted as a standard that some aspects in reality or very special identifications are preferred to establish a 'Weltanschauung', while others are neglected. As far as ethnocentric tendencies of viewing our world or dealing with it are firmly anchored in certain social groups, or in important individual characteristics, they have to be considered as consequential and integral parts of our social reality.

References:

Albrecht, Günter 1997: Anomie oder Hysterie - oder beides? Die bundesrepublikanische Gesellschaft und ihre Kriminalitätsentwicklung, in: ***Wilhelm Heitmeyer*** (ed.), Was treibt die Gesellschaft auseinander? Bundesrepublik Deutschland: Auf dem Weg von der Konsensus- zur Konfliktgesellschaft. Bd. I, Frankfurt a.M.: Suhrkamp: 506 - 556.

Berger, Peter L. and ***Thomas Luckmann*** 1974: Die gesellschaftliche Konstruktion der Wirklichkeit. Eine Theorie der Wissenssoziologie, Frankfurt: S. Fischer.

Blank, Thomas and ***Stefan Schwarzer*** 1994: Ist die Gastarbeiterskala noch zeitgemäß? Die Reformulierung einer ALLBUS-Skala, in: ZUMA-Nachrichten 34: 97 - 115.

Blank, Thomas and ***Martina Wasmer*** 1996: Gastarbeiter oder Ausländer? Ergebnisse des Splits mit den reformulierten Gastarbeiterfragen im ALLBUS 1994, in: ZUMA-Nachrichten 38: 45 - 69.

Döring, Herbert 1990: Aspekte des Vertrauens in Institutionen: Westeuropa im Querschnitt der internationalen Wertestudie 1981, in: Zeitschrift für Soziologie 19: 73 - 89.

Döring, Herbert 1992: Higher Education and Confidence in Institutions: An Analysis, in: West European Politics 15: 126 - 146.

Durkheim, Emile 1951 (1893): Suicide. A Study in Sociology, Glencoe (Illinois): The Free Press.

Durkheim, Emile 1960[4] (1912): Les formes élémentaires de la vie religieuse, Paris: Presses Universitaires de France.

Easton, David 1979 (1965): A System Analysis of Political Life, Chicago and London: The University of Chicago Press.

Ebertz, Michael N. 1993: Die Zivilisierung Gottes und die Deinstitutionalisierung der 'Gnadenanstalt'. Befunde einer Analyse von eschatologischen Predigten, in: ***Jörg Bergmann***, ***Alois Hahn*** and ***Thomas Luckmann*** (eds.), Religion und Kultur. Sonderheft 33 der Kölner Zeitschrift für Soziologie und Sozialpsychologie, Opladen: Westdeutscher Verlag: 92 - 125.

Fassmann, Heinz and ***Rainer Münz*** 1996: Europäische Migration - Ein Überblick, in: ***Heinz Fassmann*** and ***Rainer Münz*** (eds.), Migration in Europa. Historische Entwicklung, aktuelle Trends und politische Reaktionen, Frankfurt a.M. and New York: Campus: 13 - 52.

Firsching, Horst 1995: Die Sakralisierung der Gesellschaft. Émile Durkheims Soziologie der >Moral< und der >Religion< in der ideenpolitischen Auseinandersetzung der Dritten Republik, in: ***Volkhard Krech*** and ***Hartmann Tyrell*** (eds.), Religionssoziologie um 1900, Würzburg: Ergon: 159 - 193.

Friedrichs, Jürgen 1997: Normenpluralität und abweichendes Verhalten. Eine theoretische und empirische Analyse, in: ***Wilhelm Heitmeyer*** (ed.), Was treibt die Gesellschaft auseinander? Bundesrepublik Deutschland: Auf dem Weg von der Konsensus- zur Konfliktgesellschaft. Bd. I, Frankfurt a.M.: Suhrkamp: 473 - 505.

Gehring, Annekatrin and ***Ferdinand Böltken*** 1985: Einstellungen zu Gastarbeitern 1980 und 1984: Ein Vergleich, in: ZA-Information 17: 23 - 33.

Geiger, Theodor 1967 (1932): Die soziale Schichtung des Deutschen Volkes, Stuttgart: Ferdinand Enke.

Gille, Martina 1994: Wertorientierungen und Wertewandel bei Jugendlichen und jungen Erwachsenen, in: Politische Studien 45/336: 44 - 67.

Glock, Charles Y. and ***Rodney Stark*** 1966: Christian Beliefs and Anti-Semitism, New York and London: Harper & Row.

Greeley, Andrew M. 1995: Religion as Poetry: An Empirical Model, New Brunswick, NJ.

Greeley, Andrew M. (ed.) 1995a: Sociology and Religion: A Collection of Readings, New York: Harper Collins.

Heitmeyer, Wilhelm 1997: Gesellschaftliche Integration, Anomie und ethnisch-interkulturelle Konflikte, in: ***Wilhelm Heitmeyer*** (ed.), Was treibt die Gesellschaft auseinander? Bundesrepublik Deutschland: Auf dem Weg von der Konsensus- zur Konfliktgesellschaft. Bd. I, Frankfurt a.M.: Suhrkamp: 629 - 653.

Inglehart, Ronald 1971: The Silent Revolution in Post-Industrial Societies, in: American Political Science Review 65: 991 - 1017.

Inglehart, Ronald 1998: Modernisierung und Postmodernisierung. Kultureller, wirtschaftlicher und politischer Wandel in 43 Gesellschaften, Frankfurt a.M. and New York: Campus.

Kornhauser, William 1959: The Politics of Mass Society, New York: The Free Press.

Krauth, Cornelia and ***Rolf Porst*** 1984: Sozioökonomische Determinanten von Einstellungen zu Gastarbeitern, in: ***Karl Ulrich Mayer*** and ***Peter Schmidt*** (eds.), Allgemeine Bevölkerungsumfrage der Sozialwissenschaften, Frankfurt a.M. and New York: Campus: 233 - 266.

Krebs, Dagmar and ***Karl F. Schuessler*** 1987: Soziale Empfindungen. Ein interkultureller Skalenvergleich bei Deutschen und Amerikanern, Frankfurt and New York: Campus.

Küchler, Manfred 1994: Germans and 'Others': Racism, Xenophobia, or 'Legitimate Conservativism'?, in: German Politics 3: 47 - 74.

Küchler, Manfred 1996: Xenophobie im internationalen Vergleich, in: ***Jürgen W. Falter***, ***Hans-Gerd Jaschke*** and ***Jürgen R. Winkler***, Rechtsextremismus. Ergebnisse und Perspektiven der Forschung, Opladen: Westdeutscher Verlag: 249 - 262.

Kühnel, Steffen M. and ***Michael Terwey*** 1990: Exploring Attitudes towards Foreign "Guestworkers" in Western Germany, Papier präsentiert auf dem XII. Weltkongreß für Soziologie am 09.7. - 13.7.90 in Madrid.

Kühnel, Steffen M. and ***Michael Terwey*** 1994: Gestörtes Verhältnis? Die Einstellung der Deutschen zu Ausländern in der Bundesrepublik, in: ***Michael Braun*** and ***Peter Ph. Mohler*** (eds.), Blickpunkt Gesellschaft 3. Einstellungen und Verhalten der Bundesbürger, Opladen: Westdeutscher Verlag: 71 - 105.

Leenen, Wolf Rainer 1995: Ausländerfeindlichkeit, fremdenfeindliche Gewalt und politische Öffentlichkeit, in: Deutschland Archiv 28/6: 603 - 624.

Luckmann, Thomas 1995: The New and the Old in Religion, in: ***Andrew M. Greeley*** (ed.), Sociology and Religion: A Collection of Readings, New York: HarperCollins: 232 - 245.

Maslow, Abraham H. 1954: Motivation and Personality, New York: Harper.

Martin, David 1978: The Religious Condition of Europe, in: ***Salvador Giner*** and ***Margaret Scotford Archer*** (eds.), Contemporary Europe. Social Structures and Cultural Patterns, London, Henley and Boston: Routledge & Kegan Paul: 228 - 287.

McCutcheon, Allan L. 1998: Religion and Tolerance of Foreigners: A Comparative Trend Analysis of Anti-Foreigner Sentiment in Post- Unification Germany, Präsentation auf der Tagung: "Religiöser und kirchlicher Wandel in Ostdeutschland", Frankfurt an der Oder, Europa-Universität Viadrina 16.- 18.10.98.

Meulemann, Heiner 1985: Säkularisierung und Politik. Wertwandel und Wertstruktur in der Bundesrepublik Deutschland, in: Politische Vierteljahresschrift 26: 29 - 51.

Meulemann, Heiner 1996: Werte und Wertewandel. Zur Identität einer geteilten und wieder vereinten Nation, Weinheim and München: Juventa.

Merton, Robert K. 1970: Anomie, Anomia und soziale Interaktion - Beziehungen zu abweichenden Verhalten, in: ***Arthur Fischer*** (ed.), Die Entfremdung des Menschen in einer heilen Gesellschaft, München: Juventa: 156 - 161.

Mochmann, Ekkehard and ***Ingvill C. Oedegaard*** 1995: Zwischen Wahrnehmung und Realität: Zur Diskussion der Gewaltentwicklung in der Literatur und in den Medien, in: ***Ekkehard Mochmann*** and ***Uta Gerhardt*** (eds.), Gewalt in Deutschland. Soziale Befunde und Deutungslinien, München: Oldenbourg: 125 - 146.

Pappi, Franz Urban and ***Michael Terwey*** 1982: The German Electorate: Old Cleavages and New Political Conflicts, in: ***Herbert Döring*** and ***Gordon Smith*** (eds.), Party Government and Political Culture in Western Germany, London: Macmillan: 174 - 196.

Parry, Geraint 1976: Trust, Distrust, and Consensus, in: British Journal of Political Science 6: 129 - 142.

Popper, Karl R. 1974: Conjectures and Refutations. The Growth of Scientific Knowledge, London: Routledge and Kegan Paul.

Rippl, Susanne 1995: Vorurteile und persönliche Beziehungen zwischen Ost- und Westdeutschen, in: Zeitschrift für Soziologie 24: 273 - 283.

Rudolph, Hedwig 1996: Die Dynamik der Einwanderung im Nichteinwanderungsland Deutschland, in: ***Heinz Fassmann*** and ***Rainer Münz*** (eds.), Migration in Europa. Historische Entwicklung, aktuelle Trends und politische Reaktionen, Frankfurt a.M. and New York: Campus: 161 - 182.

Sigelman, Lee and ***Susan Welch*** 1993: The Contact Hypothesis Revisited: Black-White Interaction and Positive Racial Attitudes, in: Social Forces 71: 781 - 795.

Srole, Leo 1956: Social Integration and Certain Corollaries - An Exploratory Study, in: American Sociological Review 21: 709 - 716.

Statistisches Bundesamt (ed.) 1997: Datenreport 1997. Zahlen und Fakten über die Bundesrepublik Deutschland, Bonn: Bundeszentrale für politische Bildung.

Steußloff, Hans 1997: Einstellungen der Bevölkerung der Bundesrepublik Deutschland zu Ausländern, in: Berliner Dialog-Hefte 8/4: 6 - 17.

Tajfel, Henri 1982: Gruppenkonflikt und Vorurteil. Entstehung und Funktion sozialer Stereotypen, Bern, Stuttgart and Wien: Huber.

Terwey, Michael 1987: Class Position and Income Inequality: Comparing Results for the Federal Republic with Current U.S. Research, in: ***Wolfgang Teckenberg*** (ed.), Comparative Studies of Social Structure. Recent Research on France, the United States, and the Federal Republic of Germany, Armonk, New York and London: Sharpe: 119 - 171.

Terwey, Michael 1989: Der ALLBUS 1988: Die neue "Allgemeine Bevölkerungsumfrage der Sozialwissenschaften", vorgestellt mit einem Analysebeispiel zum Einstellungswandel gegenüber Gastarbeitern, in: ZA-Information 24: 37 - 49.

Terwey, Michael 1990: Zur Wahrnehmung von wirtschaftlichen Lagen in der Bundesrepublik, in: ***Walter Müller***, ***Peter Ph. Mohler***, ***Barbara Erbslöh*** and ***Martina Wasmer*** (eds.), Blickpunkt Gesellschaft. Einstellungen und Verhalten der Bundesbürger, Opladen: Westdeutscher Verlag: 144 - 171.

Terwey, Michael 1993: Sind Kirche und Religion auf der Verliererstraße? Vergleichende Analysen mit ALLBUS- und ISSP-Daten, in: ZA- Information 32: 95 - 112.

Terwey, Michael 1994: Pluralismus des Glaubens in der Diskussion, in: ZA-Information 35: 110 - 134.

Terwey, Michael 1995: Weltanschauliche Selbstbestimmung und Einstellung zu sozialer Ungleichheit: Unterschiede im Deutschen Post- Sozialismus?, in: ***Heinz Sahner*** and ***Stefan Schwendtner*** (eds.), 27. Kongreß der Deutschen Gesellschaft für Soziologie. Gesellschaften im Umbruch: Sektionen und Arbeitsgruppen, Opladen: Westdeutscher Verlag: 674 - 679.

Terwey, Michael 1996: Demokratiezufriedenheit und Vertrauen: Politische Unterstützung in Westeuropa und im vereinten Deutschland, ZA-Information 39: 94 - 129.

Terwey, Michael 1997: Soziales Desinteresse von Politikern und Mitmenschen: Schwächen im Gemeinsinn Anderer?, in: ZA-Information 41: 120 - 144.

Terwey, Michael 1998: Glaube an ein Leben nach dem Tod, in: ***Detlef Pollack***, ***Irena Borowik*** and ***Wolfgang Jagodzinski*** (eds.), Religiöser Wandel in den Ländern Ost- und Mitteleuropas, Würzburg: ERGON: 179 - 203.

Terwey, Michael and ***Detlef Pollack*** 1997: Institutionenvertrauen in Deutschland - Differenz und Integration, in: ***Karl-Siegbert Rehberg*** (ed.), 28. Kongreß der Deutschen Gesellschaft für Soziologie. Differenz und Integration: Sektionen und Arbeitsgruppen, Opladen: Westdeutscher Verlag: 346 - 351.

Thränhardt, Dietrich 1988: Die Bundesrepublik Deutschland - ein unerklärtes Einwanderungsland, in; Aus Politik und Zeitgeschichte B24/88: 3 - 13.

Tönnies, Ferdinand 1979 (1887): Gemeinschaft und Gesellschaft. Grundbegriffe der reinen Soziologie, Darmstadt: Wissenschaftliche Buchgesellschaft.

Velling, Johannes 1994: Zuwanderer auf dem Arbeitsmarkt: Sind die neuen Migranten die "Gastarbeiter" der neunziger Jahre?, in: ZEW Wirtschaftsanalysen 2 - Quartalshefte des Zentrum für Europäische Wirtschaftsforschung, Mannheim: 261 - 296.

Wasmer, Martina, Achim Koch, Janet Harkness and ***Siegrfried Gabler*** 1996: Konzeption und Durchführung der "Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften" (ALLBUS) 1996, Mannheim: ZUMA-Arbeitsbericht Nr. 96/08.

Weber, Max 1993 (1904/05): Die protestantische Ethik und der "Geist" des Kapitalismus, Bodenheim: Neue Wissenschaftliche Bibliothek.

ZA (Zentralarchiv für Empirische Sozialforschung) and ZUMA (Zentrum für Umfragen, Methoden und Analysen) 1996: ALLBUS 1996 (Studie-Nr. 2800). Codebuch, Köln: Zentralarchiv.

ZA (Zentralarchiv für Empirische Sozialforschung) and ZUMA (Zentrum für Umfragen, Methoden und Analysen) 1997: Kumulierter ALLBUS 1980-96 (Studie-Nr. 1795). Codebuch, Köln: Zentralarchiv.

The Northern Ireland Agreement: An Explication of a Political Milestone [1]

by Bernadette C. Hayes [2] and Ian McAllister [3]

Zusammenfassung

Am Karfreitag, den 10. April 1998 wurde das Nordirlandabkommen von allen politischen Parteien Nordirlands unterzeichnet. Sowohl von den politischen Kommentatoren und politischen Beobachtern wurde das Hauptziel des Abkommens, das Unvereinbare zu vereinbaren, als eine bedeutende Errungenschaft angesehen, nämlich die beiden religiösen Gemeinschaften in Nordirland, die Protestanten und Katholiken, die Unionisten und Nationalisten, die Loyalisten und die Republikaner in einem gerechten und dauerhaften Abkommen, das die unterschiedlichen Traditionen respektiert, zusammenzubringen. In welchem Ausmaß wurde dieses Ziel erreicht? Um eine Antwort auf diese Frage zu finden und ein Verständnis davon zu vermitteln, wie das Abkommen versucht, einen akzeptablen politischen Kompromiß für beide religiösen Gruppen zu finden, richtet sich dieser Artikel auf die folgenden drei Streitfragen: Zunächst werden Hintergrundinformationen im Sinne von aktuellen sozialen und politischen Beziehungen zwischen den beiden religiösen Traditionen präsentiert. Es folgt ein kurzer historischer Überblick über die Intensität und das Ausmaß des Konfliktes, wie auch sein Einfluß auf die laufenden demokratischen Entscheidungen in dieser Provinz. Schließlich werden die eigentlichen Bedingungen des Abkommens, seine öffentliche Akzeptierbarkeit unter den beiden religiösen Gruppen ebenso wie die Chancen für einen friedlichen und andauernden Ausgleich in Nordirland geprüft.

---

1 This article is based on a revised version of a paper presented by ***Bernadette C. Hayes*** at a seminar of the Central Archive for Empirical Social Research, University of Cologne, where she was a visiting researcher from March-May 1998. The research fellowship was funded by the Training and Mobility of Researchers (TMR) programme - Large Scale Facility Activity (LSF) of the European Union.

2 ***Bernadette C. Hayes***, Department of Sociology, Queen's University of Belfast, Belfast BT7 1NN, Northern Ireland.

3 ***Ian McAllister***, Research School of the Social Sciences, Australian National University, Canberra, ACT 0200, Australia.

Abstract

On Good Friday, the 10th April 1998, the Northern Ireland Agreement was formally endorsed by all the major political parties in Northern Ireland. Considered a landmark achievement by both political analysts and commentators alike, the primary purpose of the agreement was to reconcile the irreconcilable; to bring together the two religious communities of Northern Ireland, Protestant and Catholic, Unionist and Nationalist, Loyalist and Republican, in an equitable and lasting accommodation that would respect their differing traditions and heritage. To what extent was this objective met? To answer this question and to provide an understanding of how the agreement sought to find a political compromise acceptable to both religious communities, this article addresses the agreement in terms of the following three issues. First, some background information in terms of the current social and political relations between the two religious traditions in Northern Ireland is presented. This is followed by a brief historical overview of the scale and intensity of the conflict as well as its impact for current democratic governance within the province. Finally, the actual terms of the agreement, its public acceptability among the two religious communities, as well as its chances of achieving a peaceful and lasting settlement within Northern Ireland are examined.

Introduction

After much political drama and extensive media hype, on Good Friday, the 10th April 1998, the Northern Ireland Agreement was formally endorsed by all the major political parties in Northern Ireland. For the first time in Northern Ireland`s long and troubled history, political representatives from both communities - Protestant and Catholic, Unionist and Nationalist, Loyalist and Republican - came together to achieve what must be considered, by any standards, the completion of one of the most difficult but eventually successful negotiations undertaken in this society. United in a common purpose to end the violence and to find a political compromise acceptable to both religious communities, the primary aim of the participants to the agreement was to reconcile the irreconcilable. Above all else, the agreement was a finely-tuned endeavour aimed at satisfying both the differing traditions and aspirations of the two religious communities within Northern Ireland and in doing so, hopefully, to pave the way for democratic governance.

The achievement of this attempted balance was not without controversy, however. No sooner had the agreement been endorsed but an avalanche of criticism was hurled at the participants from all sides. While some political commentators from both sides of the two religious traditions castigated it as a negation of their national traditions and heritage, other, more muted, criticisms stressed the one-sided nature of the agreement in terms of

either a unionist or nationalist advantage. Still other, more vocal commentators, lambasted the agreement as a death knell for democracy and characterised it as nothing more than a sell-out to terrorists, more notably the paramilitary organisations, such as the Irish Republican Army (IRA), on the Catholic or nationalist side. To what extent are these differing interpretations correct? For example, was the Northern Ireland Agreement, in fact, a sell-out to republican terrorists as some Unionist commentators, most notably ***Ian Paisley***, have suggested? Alternatively, to what extent do claims concerning the one-sided nature of the agreement, either in terms of its perceived greater accommodation to either unionist or nationalist aspirations, hold up?

To answer these questions and to provide an understanding of how this agreement sought to achieve a balanced compromise in terms of the differing traditions and aspirations of the two religious communities within Northern Ireland, this article will address the agreement in terms of the following three issues. First, background information in terms of the current social and political relations between the two religious traditions in Northern Ireland is presented. Second, this will be followed by a brief overview of the scale and intensity of the conflict as well as its impact on democratic governance within the province. Finally, the actual terms of the agreement, both in terms of its particular details, its perceived advantages and disadvantages for each of the two religious communities, as well as its chances of achieving a peaceful and lasting political settlement within Northern Ireland is addressed.

Religion, Community Relations and Political Identity

The Northern Irish conflict has traditionally been characterised as a sectarian conflict between two monolithic religious communities, Protestant and Catholic. Although this view has been disputed over the past quarter of century, mainly by Marxists, few social scientists question the basic proposition that the major divide, or the principal source of identity in Northern Ireland, is religious affiliation (see ***Whyte***, 1990, for a review of this literature). Empirical research confirms this proposition (see Table 1). In every census conducted up to 1981, more than nine out of ten residents of Northern Ireland identified either with the Catholic Church or one of the Protestant denominations. In the 1951 census alone, 99.4 per cent of the population fell into these two categories. Although this figure had fallen by 1991, or the most recent census, approximately 85 per cent of the total population still explicitly claim either a Catholic or Protestant religious affiliation. Furthermore, the proportion of individuals who claim a Catholic religious affiliation has grown significantly over the last two decades. Currently, 38.4 per cent of the Northern Ireland population identify themselves as Roman Catholic as compared to 45.6 per cent for their Protestant counterparts.

Table 1: Religious Identification in Northern Ireland, 1961-1991

	Year of Census			
Religious affiliation	1961	1971	1981	1991
Protestant	60.4	55.6	48.4	45.6
Catholic	34.9	31.4	28.0	38.4
Other	2.7	3.6	5.1	5.0
Not Stated	2.0	9.4	18.5	7.3
No religion	---	---	---	3.7
Total	100.0	100.0	100.0	100.0
(N)	(1,425,042)	(1,519,640)	(1,481,959)	(1,577,836)

Note: In the 1991 census, a separate category was introduced which allowed individuals to explicitly identify themselves as having no religious affiliation, or the people whom sociologists of religion have traditionally referred to as "Nones".
Source: Northern Ireland Census, 1961, 1971, 1981 and 1991.

More importantly, however, these religious divisions are not merely nominal but extend to all areas of political and social life within the province. Northern Ireland is an extremely religiously segregated society (see ***Curtice*** and ***Gallagher***, 1990). The two communities are educated separately, are segregated both residentially and in their places of work, and display very low rates of intermarriage. For example, recent estimates from the Social Identity in Northern Ireland Survey undertaken by ***Karen Trew*** at Queen's University of Belfast suggests that only about one in six of the adult population has attended a school where there are pupils of the other religion, a proportion that has remained constant over the past thirty years or so, and that the vast majority claim that all of most of their relatives and neighbours, 85 and 66 per cent respectively, are of the same religion as themselves. Added to this separation are inequalities in life chances. Protestants are, on average, better educated and more affluent than Catholics and are substantially less likely to be unemployed, despite the passage of a series of anti-discrimination laws since the late 1970s (***Smith*** and ***Chambers***, 1991).

However, more so than any other factor, it is disagreement about political identity and the preferred constitutional position of Northern Ireland that separates the two communities. Here preferences are polarised around two competing ideologies: Unionism - support for the retention of the union between Northern Ireland and Great Britain - and, Nationalism - support for a united Ireland. It is these deeply felt and mutually exclusive ideologies that have given Northern Irish politics its distinctive and violent nature. The conflict is further heightened in that this political cleavage coincides to a considerable extent with a religious division (see Table 2). In other words, not only do the vast majority of Protestants (75 per cent) perceive themselves as Unionists, but they are also overwhelming in their support for the maintenance of the Union with Britain (88 per cent).

Table 2: Religious Differences in Political Identity and Constitutional Preferences, 1995

	(Percentages)	
	Protestant	Catholic
Political Identity:		
Unionist	75	0
Neither	25	49
Nationalist	0	51
(N)	(730)	(564)
Constitutional Preferences:		
Remain part of UK	88	34
Reunify Ireland	6	56
Other Option	2	3
Don't know	4	6
Not answered	0	1
(N)	(354)	(290)

Source: Northern Ireland Social Attitudes Survey, 1995.

This uniformity in political identities and constitutional preferences is not as stark within the Catholic community. Nevertheless, Table 2 shows that just over half of all Catholics (51 per cent) see themselves as Nationalists, and a majority (56 per cent) favours a united Ireland. More importantly, however, whereas no Catholic sees him or herself in terms of a Unionist identity, the reverse is true in terms of a Nationalist identity for Protestants. Furthermore, a review of past survey evidence suggests that in contrast to the stability in Protestant views which have never deviated more than three percentage points since the late 1980s, the amount of Catholics who wish to remain within the United Kingdom has fluctuated considerably in the last few years decreasing from 36 per cent in 1993 to just 24 per cent in 1994, only to rise again to 34 per cent in 1995 (***Breen***, 1996).

This sectarian division is further reflected in the political party structure (see Table 3). For example, not only do the overwhelmingly majority of Protestants support the two main Unionist Parties (the Democratic Unionist Party [DUP] led by ***Ian Paisley*** and the Ulster Unionist Party [UUP] led by ***David Trimble***) but Catholics almost exclusively divide their support between the two main nationalist parties (the Social Democratic Labour Party [SDLP] led by ***John Hume*** and Sinn Fein [SF] led by ***Gerry Adams***). Furthermore, whereas no Catholic supports either of the two main unionist parties (DUP and UUP) only 1 per cent of Protestants support Nationalist parties, in this case the more moderate SDLP. Also, although both Protestants and Catholics support the one avowedly non-sectarian party, the Alliance Party of Northern Ireland [APNI] led by ***John Alderdice***, in both cases the level of support it attracts is extremely modest, and it has constantly failed to capture enough votes to win a parliamentary seat at Westminster.

Table 3: Religious Differences in Party Preferences, 1995

	(Percentages)	
	Protestant	Catholic
Democratic Unionist Party (DUP)	19	0
Ulster Unionist Party (UUP)	53	0
Alliance Party Northern Ireland (APNI)	11	9
Social Democratic Labour Party (SDLP)	1	60
Sinn Fein (SF)	0	7
No Party	16	24
(N)	(658)	(507)

Note: The parties are arranged in descending order from the most extreme Unionist parties (DUP) to the most extreme Nationalist parties (SF).
Source: Northern Ireland Social Attitudes Survey, 1995.

In summary, then, Northern Ireland is a deeply divided society. Not only are the two religious communities segregated both geographically and socially, but for the overwhelming majority of the population their political identities and their religious affiliation are almost coterminous. The society is also divided on a range of social factors, all of which reinforce the sectarian divide. As we noted above, there is little or no integration between the communities in education, with the exception of university education; intermarriage is rare and when it occurs, the couple often choose to live outside Northern Ireland; and there is a strong pattern of residential segregation, particularly in the working class areas of Belfast and the other major towns. The net effect is a religion division which is reinforced by divisions within the society as a whole, which are recreated from generation to generation.

Exposure to Political Violence

Accompanying these high levels of religious segregation and political division, Northern Ireland has also experienced a long and sustained period of political violence. By any standards, the exposure of Northern Ireland society to political violence has been massive. For example, since the start of the current Troubles in 1967, there has been: 3,172 deaths (70 per cent of which have been civilian); 24,024 civilian injuries; 35,059 shooting incidents; 14,884 bomb explosions; and 16,365 individuals charged with terrorist offences (see Table 4).

In a population of just over 1.5 million people, the scale and the intensity of the violence put it on the level of a war. If we were to extrapolate the scale of the violence to Britain, with a population of over 57 million, we would find somewhere in the region of 119,000 deaths, one million civilian injuries and 1,315,000 shooting instances. In the United States, the same level of violence would result in just over half a million deaths and close to four

Table 4: The Scale of the Political Violence in Northern Ireland, 1967-1995

	Northern Ireland	Great Britain	United States
Deaths	3,172	119,000	515,000
Civilian injuries	24,024	1,000,000	3,904,000
Shooting incidents	35,059	1,315,000	5,697,000
Bomb explosions	14,884	558,000	2,419,000
Persons charged with terrorist offences	16,365	614,000	2,659,000

Source: ***Bairner***, 1996, pp. 168-171.

million injuries. By any standards, whether in terms of the general loss of life or the extent of civilian casualties, both the level and intensity of the communal violence in Northern Ireland must be considered more like a war than a minor conflict.

It is hardly surprising, then, that substantial numbers of people in Northern Ireland have had significant exposure to various aspects of the violence (see Table 5). The most recently available survey evidence shows that whereas 16 per cent of Protestants and 26 per cent of Catholics claim that they have been caught up in a riot, almost double these numbers - 28 per cent of Protestants and 48 per cent of Catholics - have had some direct exposure to violence. When indirect exposure to the violence is considered, the numbers climb even further, with about two-thirds of individuals within both communities been affected in terms of claiming knowledge of the death or injury of someone they know.

Table 5: Exposure to Political Violence in Northern Ireland, 1995

	(Percentages)	
	Protestant	Catholic
Caught up in Riot	16	26
Direct Exposure		
Intimidated	13	29
House searched	3	24
Family member killed or injured	17	22
Victim, uninjured	5	6
Victim, injured	4	6
[At least one]	[28]	[48]
Indirect Exposure		
Know someone killed	60	63
Know someone injured	56	57
[At least one]	[68]	[69]
(N)	(576)	(366)

Source: Social Identity in Northern Ireland Survey, 1995

There are also some notable differences between the two religious communities in their exposure to different types and intensities of political violence. For example, not only do a significantly higher proportion of Catholics than Protestants claim to have experienced the violence first-hand in terms of their exposure to rioting behaviour, but an even more pronounced differential emerges when a range of other direct violent measures such as self-reported intimidation and house searches are considered. This is not the case when either injury or the loss of life is considered; here, both communities have borne an equal share of the ultimate human cost in terms of Northern Ireland's long and violent history.

The Absence of Democratic Governance

Northern Ireland has never been a competitive democracy in the conventional sense. From the partition of Ireland in 1920 until the suspension of the Stormont parliament in 1972, formal democratic institutions were maintained in terms of a devolved parliament. But the province was not a competitive democracy in the conventional sense, since there was no possibility that the Protestant and unionist majority would ever lose an election, nor that the Catholic and nationalist opposition could ever win. It was a hegemonic state in which one community (Protestant) enjoyed permanent government, while the other community (Catholic) was consigned to permanent political opposition (***O'Leary*** and ***Arthur***, 1990).

Since the introduction of direct rule from Britain in 1972, even limited local democracy has been withdrawn, with the partial (and extremely short-lived) exceptions of the 1973 Assembly and the 1982 Assembly, both of which failed in their goal of re-establishing devolution because of Protestant opposition to power-sharing with their Catholic counterparts. This has had the effect of creating what has been called a democratic deficit, since Northern Ireland affairs are rarely debated at the Westminster parliament in London, and since local government (following its reform in 1973) has few substantive powers. The consequence has been to stifle political debate in Northern Ireland, and to ensure that talented potential political leaders must go elsewhere to follow their ambition.

The legacy of political violence and non-democratic governance is evident in terms of the lack of public support for democratic values in Northern Ireland. The vast majority of people in Northern Ireland have little faith in the democratic process. Recent survey evidence demonstrates that popular support for political efficacy and trust, two of the central concepts in contemporary theories of political participation and democratic politics, is extremely low within this society (***Hayes*** and ***McAllister***, 1996). Irrespective of whether Protestants or Catholics are considered, not only do the vast majority of Northern Ireland citizens express an extremely negative opinion in terms of the responsiveness of their political representatives, but a widespread feeling of alienation and a sense of powerlessness as to how the political system works is by far the most dominant view.

Similar results emerge when political trust, or attitudes concerning the trustworthiness of elected politicians, civil servants and the police force, are considered, although, in this case, significant differences also emerge between Protestants and Catholics in relation to the issue. As a group, Protestants are notably more trusting of these political institutions than Catholics; the average level of trust was 37 per cent among Protestants, but only 19 per cent among Catholics. While an explanation for this religious difference can be found in terms of the consistently more negative views expressed by Catholics concerning the trustworthiness of these political institutions, more so that any other factor, however, it is differences in relation to their attitudes towards the police that set Protestants and Catholics apart on this issue. In contrast to nearly two-thirds of Protestants who believe that the police would not bend the rules to secure a conviction, only just over one quarter of Catholics share this view.

Contrary to other western established democracies, Northern Ireland has never been a truly democratic society in the conventional meaning of the word. Decades of political instability and intransigence, heightened by a seemingly endless stream of violence and death, has done little to negate this political vacuum. Popular democratic values, the underpinnings of any effective set of democratic institutions, are weak and under-developed. It was in the context of these historically-derived circumstances of a bitterly divided and politically disillusioned people that the Good Friday agreement was formally endorsed on 10 April 1998.

The Northern Ireland Agreement

The current peace process has its roots in the 1985 Anglo-Irish Agreement. The Anglo-Irish Agreement established the principle that the government of the Irish Republic had a legitimate role to play in Northern Ireland affairs (see ***O'Leary*** and ***McGarry***, 1993). It was agreed to by the British conservative prime minister ***Margaret Thatcher*** who had reached the conclusion that only extensive collaboration with the government of the Irish Republic would improve the security situation in Northern Ireland. Her assessment of the situation followed the hunger strikes staged by IRA prisoners in 1981-82, during which 10 died, and the bombing of the Grand Hotel in Brighton in 1984, which narrowly missed killing her and other cabinet members. The Anglo-Irish Agreement was opposed by both the Ulster unionists and Irish republicans and it had no direct consequences for establishing democratic institutions in Northern Ireland.

During the 1980s the IRA had also begun to reassess its position (see ***McIntyre***, 1995). Irish republicans had always maintained that the key to a resolution of the Northern Ireland problem was Britain’s continuing presence in Ireland; if Britain indicated its intention to leave Northern Ireland, republicans argued that the Protestant majority would seek common cause with their Catholic neighbours and realise their true Irish heritage and identity. By the late 1980s, two factors had made this position untenable. First, despite discussion of

military and political withdrawal in the mid-1970s, the British government had taken the decision that withdrawal would create more instability and violence than continuing their presence in the province. Britain would therefore remain and try to maintain ‘an acceptable level of violence’. Second, the military efforts of the IRA had failed to secure a British withdrawal, and were unlikely to do so in the future. A new republican strategy to achieve Irish unity was therefore imperative.

This new IRA strategy emerged as ‘the ballot box in one hand and the armalite in the other’, the latter a reference to a rifle favoured by the IRA. Following the IRA hunger strikes and consequent Sinn Fein electoral successes of the early 1980s, the possibility that the IRA/Sinn Fein might replace the SDLP as the electoral representative of the Catholic community appeared feasible for the first time. Such a strategy had the potential to achieve Irish unity by constitutional methods, rather than by extra-constitutional physical force. Several new factors fostered this strategy. First, there had been a generational turnover within the IRA/Sinn Fein leadership; younger, more flexible leaders such as ***Gerry Adams*** and ***Martin McGuinness*** replaced the republican activists of the 1960s and 1970s who saw political change exclusively in terms of physical force. Second, ***John Hume***, the leader of the SDLP, had entered into a dialogue with ***Gerry Adams*** and kept the British government informed of their progress. While ***Hume*** remained a trenchant critic of violence and encountered considerable opposition within his own party for talking to ***Adams***, he remained in regular contact with him throughout the late 1980s and early 1990s.

All of these events culminated in the Downing Street Declaration of 15 December 1993, when the British and Irish governments agreed on a set of principles which would guide their future approach to the Ulster problem. The key part of the declaration was the recognition that Britain had no vested interest in Northern Ireland, and that it was for the people of Ireland themselves to exercise their right of self-determination. But Ulster unionist and British government demands for decommissioning of arms as a prior step before the start of formal talks initially stalled the peace process. The talks eventually commenced in January 1996, under auspices of a former member of the US Senate, ***George Mitchell***, but it was not until the election of the **Blair** Labour government in May 1997, which deliberately blurred the decommissioning issue, that progress began to be made. The talks finally concluded with an agreement on 10 April 1998.

The April 1998 Peace Agreement is based on the principle of parity of esteem: that is, the recognition of the political rights of both Protestants and Catholics in Northern Ireland, and the right to express those rights in political institutions. This principle formed much of the basis of agreement in the Hume-Adams talks. The Agreement has three strands. Strand one involves the creation of an elected assembly in Northern Ireland with a powersharing executive; decisions of the assembly require a majority of both unionists and nationalist before they can be passed. The second strand creates a North-South ministerial council to

bring together ministers from the Irish Republic and Northern Ireland. Strand three creates a British-Irish Council bringing together representatives from all parts of the United Kingdom, including Scotland, Wales, and England, as well as both parts of Ireland. Other parts of the agreement stipulate total decommissioning of weapons within two years of the referendums, the release of paramilitary prisoners, and an inquiry into policing in Northern Ireland.

A crucial part of the Agreement was ratification by the citizens of both parts of Ireland. The two referendums to seek public endorsement for the Peace Agreement were held on 22 May 1998 (Table 6). There was little doubt that the Agreement would win overwhelming support in the Irish Republic and among Ulster Catholics, but the reaction of Ulster Protestants was much more uncertain. The largest party, the Ulster Unionist Party, was split on the issue, with its new leader, ***David Trimble***, supporting the Agreement but encountering trenchant opposition from many of his senior colleagues. The second largest Protestant party, the Democratic Unionists led by ***Ian Paisley***, opposed the Agreement on the grounds that it would undermine the Protestant community's position within the United Kingdom. Just one week prior to the referendum in Northern Ireland, opinion polls were showing that the Protestant community was deeply split on the issue, with a large number of undecided voters.

Table 6: The Result of the Referendums, 22 May 1998

	Northern Ireland		Irish Republic	
	N	Percent	N	Percent
Yes	676,966	71.1	1,442,583	94.4
No	274,879	28.9	85,748	5.6
Total valid votes	951,845	100.0	1,528,331	100.0
Spoiled votes	1,738		17,064	
(Turnout)		(81.1)		(56.3)

Note: In Northern Ireland, the question was: 'Do you support the agreement reached in the multi-party talks on Northern Ireland and set out in Command Paper 3883?; in the Irish Republic: 'Do you approve of the proposal to amend the Constitution contained in the undermentioned Bill, the Nineteenth Amendment of the Constitution Bill, 1998?' Source: http://www.irish-times.com/irish-times/special/peace

In the event, as Table 6 shows, around seven out of every 10 voters supported the Agreement, compared to 94 percent of voters in the Irish Republic. The turnout in Northern Ireland was a very substantial 81 percent, compared to 68 percent in the 1997 Westminster election and 65 percent in the 1996 Forum elections. However, turnout in the Irish Republic was just 56 percent, low by the standards of other referendums and partly a consequence of the referendum being held alongside a referendum on the European Union's Amsterdam agreement, about which the major parties disagreed. The lack of agreement between the parties on the Amsterdam agreement clearly depressed the turnout. Spoiled votes - a traditional mechanism for demonstrating protest - were insignificant in Northern Ireland, but amounted to 1.1 percent of all votes cast in the Irish Republic, and reflect a

republican protest against the south's abandonment of its historical territorial claim over Northern Ireland, one of the key principles of the Agreement.

The split within the Protestant community on the Agreement is clearly demonstrated by the results of an exit poll conducted on polling day in Northern Ireland (**Table 7**). While almost all of the Catholic respondents voted 'yes' to the Agreement, Protestants voting in favour exceeded those voting 'no' by just 2 percent. In terms of voting intention, three quarters of Ulster Unionists voted in favour of the Agreement, compared to just 11 percent of Democratic Unionists. The narrow Protestant majority in favour of the Agreement and the size of the group opposed to it is a major concern. This fundamental division within the Protestant community and the size of the anti-Agreement group does not bode well for the future of the Agreement.

Table 7: Referendum Vote, Voting Intention and Religion

	Yes	No	(N)
All	73	27	(1,677)
Voting intention			
Ulster Unionist Party	75	25	(311)
Democratic Unionist Party	11	89	(207)
Alliance Party	96	4	(137)
SDLP	99	1	(447)
Sinn Fein	98	2	(105)
Religion			
Protestant	51	49	(798)
Catholic	99	1	(715)

Source: Irish Times/Radio Televis Eireann Exit Poll, Friday 22nd May 1998.

With the Agreement ratified by the two electorates, the next stage was to elect an assembly, from which an executive would be drawn representing both communities. These elections were held on 25 June 1998 using the single transferable vote method of proportional representation, based on the 18 parliamentary constituencies used for elections to the Westminster parliament. The party with the largest number of first preference votes was the SDLP, but after transfers, the Ulster Unionist Party emerged with the largest number of seats, 28 of the 108 seats, compared to 24 for the SDLP (see **Table 8**). Sinn Fein attracted 17.7 percent of the vote and won 18 seats, their largest vote to date. In practice, the SDLP now attracts about 55 percent of the Catholic nationalist vote, Sinn Fein 45 percent. The Assembly elections confirm that Northern Ireland is now a four party system, with two Protestant and two Catholic parties attracting eight out of every 10 votes, and occupying 90 of the 108 seats in the Assembly. In each community, the larger of the two parties is the more moderate.

Table 8: The Result of the Assembly Elections, 25 June 1998

	Votes		
	N	Percent	Seats
SDLP	177,963	22.0	24
Ulster Unionist Party	172,225	21.3	28
Democratic Unionist Party	145,917	18.0	20
Sinn Fein	142,858	17.7	18
Alliance	52,636	6.5	6
UK Unionist Party	36,541	4.5	5
Progressive Unionist Party	20,634	2.5	2
Women's Coalition	13,019	1.6	2
Ulster Democratic Party	8,651	1.1	0
Labour	2,729	0.3	0
Others	36,090	4.5	3
Total valid votes	809,263	100.0	108
		(Turnout 70.5)	

Note: Votes are first preference votes. The allocation of seats is based on transfers between the parties using the single transferable vote method of proportional representation, based on 18 constituencies.
Source: http://explorers.whyte.com

Conclusion

Few societies have such deep and abiding divisions as Northern Ireland. What makes Northern Ireland unique is the range of social divisions, and the way in which they reinforce one another and shape political identity. The educational system, patterns of residence, social networks, marriage and much of the labour force remain segregated by religion. Major social groups, which in any society provide the substance of civil society, are also largely segregated on religious grounds. It is hardly surprising that efforts to resolve the Northern Ireland conflict have fallen on the rock of these reinforcing, fundamental societal divisions. To be sure, other societies possess a range of social divisions which reinforce political identity. But few, if any, have the range of reinforcing social divisions that Northern Ireland possesses, and most have a series of competing groups, none of which constitutes a majority. The majoritarian position of the Protestant community has been a fundamental block to political compromise with the Catholic minority.

The 1998 Northern Ireland Agreement, the Referendums and Assembly elections, may collectively represent an advance for Northern Ireland. On the positive side, the paramilitary organisations have been involved in the Agreement for the first time, a fundamental deficiency of previous attempts to resolve the conflict. Moreover, there is substantial evidence that the IRA and their Protestant counterparts have reassessed their military situation and may be prepared to use political methods to try and attain their goals. The Agreement

also represents the positive efforts of many other bodies to try and resolve the conflict, most notably the British and Irish governments, and the ***Clinton*** administration in the United States, all of whom have put pressure on the various groups they are close to in Northern Ireland.

The major problems with the Agreement are twofold. First, it is necessarily an elite accommodation, a consociational agreement common in many European countries. So far as elites retain the confidence of their supporters, elite agreements remain an effective means of resolving political conflict. However, the history of the Protestant community is that political elites who are prepared to compromise on fundamental constitutional issues loose the consensual support of their community. The deep division among Protestants over the referendum is evidence of the potential that exists for moderate Protestant leaders to loose their support. The second problem is that the Agreement does not deal with the basic problems that feed the conflict: the social divisions that reinforce sectarianism and the ambiguity towards the use of physical force for political ends. The new political arrangements will only be able to create genuine cross-community institutions if they exist long enough to establish the confidence of both communities. Many hope that it will achieve that goal, but the weight of history points in the other direction.

References

Bairner, Alan 1996: 'Paramilitarism.' In: ***Arthur Aughey*** and ***Duncan Morrow*** (eds.): Northern Ireland Politics. London: Longman, pp. 159-172.

Breen, Richard 1996: 'Who Wants a United Ireland? Constitutional Preferences among Catholics and Protestants.' In: ***Richard Breen***, ***Paula Devine*** and ***Lizanne Dowds*** (eds.): Social Attitudes in Northern Ireland: The Fifth Report. Belfast: Appletree Press, pp. 33-48.

Curtice, John and ***Tony Gallagher*** 1990: 'The Northern Ireland Dimension.' In: ***Roger Jowell***, ***Sharon Witherspoon*** , ***Lindsay Brook*** and ***Bridget Taylor*** (eds.): British Social Attitudes: The 7th Report. Aldershot: Gower, pp. 183-216.

Hayes, Bernadette C. and ***Ian McAllister*** 1996: 'Public Support for Democratic Values in Northern Ireland.' In: ***Richard Breen***, ***Paula Devine*** and ***Lizanne Dowds*** (eds.): Social Attitudes in Northern Ireland: The Fifth Report. Belfast: Appletree Press, pp. 13-32.

McIntyre, Anthony 1995: 'Modern Irish Republicanism: The Product of British State Strategies.' Irish Political Studies. 10:97-121.

O'Leary, Brendan and ***Paul Arthur*** 1990: 'Northern Ireland as a Site of State and National-Building Failures.' In: ***John McGarry*** and ***Brendan O'Leary*** (eds.): The Future of Northern Ireland. Oxford: Clarendon Press.

O'Leary, Brendan and ***John McGarry*** 1993: The Politics of Antagonism: Understanding Northern Ireland. London: Athlone Press.

Smith, David J. and ***Gerald Chambers*** 1991: Inequality in Northern Ireland. Oxford: Clarendon Press.

Whyte, John 1990. Interpreting Northern Ireland. Oxford: Clarendon Press

Announcements from the ZA-EUROLAB

European Funds Available to Participate in the Training and Mobility of Researchers (TMR) Programme - Large Scale Facilities (LSF) Activity

by Ingvill C. Oedegaard

TMR-LSF funds of the European Union are available for supporting access to Large Scale Facilities in the Social Sciences at the University of Cologne, University of Essex, University of Bergen, CEPS/INSTEAD in Luxembourg, British Library of Political and Economic Science (BLPES) in London and the Library of the European University Institute (EUI) in Florence. The Large Scale Facility located at the Central Archive for Empirical Social Research (ZA) at the University of Cologne, Germany, is a resource and competence centre in comparative social research. It invites researchers to work with comparative data sets, participate in training seminars in social science methods and being trained in data management and archiving in the research environment of the ZA.

Data Available at the Zentralarchiv

The Zentralarchiv provides access to major comparative collections such as the Eurobarometer surveys, the Central and Eastern Eurobarometer, the International Social Survey Programme (ISSP) and the ICORE (International Committee for Research into Elections and Representative Democracy) election studies to national parliaments in Europe. Furthermore, the collections include the Civic Culture Study, the Political Action Surveys, the USIA (United States Information Agency) Studies and the European- and World Values Survey. In addition, the Zentralarchiv holds the political manifestos provided by the "Comparative Manifestos Project" presenting the programmatic profiles of political parties in 20 countries from 1945-1988. Altogether, there are some 4000 data sets and collections available for secondary analysis.

What is Provided by the Zentralarchiv

Access to the LSF will be provided free of charge and will include access to training seminars as well as infrastructural, logistical and scientific support. Within the TMR programme financial support is given to cover international travel and subsistence expenses in accordance with travel expenditure regulations for researchers participating in LSF-activities. Researchers from the European Union or associated states may apply for financial support for a period ideally between one and six months. Researchers from Germany can only be supported to a limited extent.

Who can Participate in the TMR-LSF Activity

All researchers from the EU member states and the associated states Iceland, Norway, Liechtenstein and Israel may apply for participation. They must be entitled to publish the results of their work at the LSF in the open literature. An international board will select applications on the basis of scientific merit through an independent peer review procedure.

How to Apply

If you want to apply please contact:
Ekkehard Mochmann,
Zentralarchiv für Empirische Sozialforschung an der Universität zu Köln,
Bachemer Straße 40, D - 50931 Köln, Germany
e-mail: mochmann@za.uni-koeln.de
Fax: +49-221-47694-77
For further information: http://www.za.uni-koeln.de/eurolab

Bursaries Available for the Spring Seminar in Empirical Social Survey Methods 1999

Within the TMR-LSF Programme of the European Union a limited number of bursaries are available for participation in the spring seminar taking place from 22 February - 12 March 1999. The bursaries cover the participation fee as well as travel costs (first class train or Apex/economic air fare), hotel (max. 120,- DM per night) and subsistence (approx. 39,- DM) per day.

Participants from the EU member states and the associated states Iceland, Norway, Liechtenstein and Israel may apply for support. German citizens can only be supported to the extent they are working at institutions outside Germany, but within the European Union or the associated states.

Participants may apply for one, two or all three weeks. It is also possible to work in the ZA-EUROLAB before or after the spring seminar in order to do analysis. The bursaries are competitive and participants will be selected by an independent peer review panel. Deadline for application is December 31 1998. An application form can be taken from the internet http://www.za.uni-koeln.de/eurolab/. If you have any questions please contact: oedegaard@za.uni-koeln.de.

For further information on the spring seminar 1999 see the announcement in this volume.

Summary Activity Report from the ZA-EUROLAB

During the summer ***Jocelyn Evans***, European University Institute, Italy, continued his stay at the ZA-EUROLAB working on the "Comparative modelling of left-wing vote defection to extreme right parties in Europe" and ***Andries van der Ark***, University of Utrecht, the Netherlands came to work with comparative data sets in latent budget models. Both users gave presentations of their research work to which other social scientists at the University of Cologne were invited. End October a user group with over 25 experts in data management participated in the TMR-LSF programme under the leadership of ***Björn Henrichsen***, Norwegian Social Science Data Services, Norway. The group was part of a European expert meeting on the topic "Survey data in the internet". The participants came from nine different countries. Since the beginning of December a Norwegian/Swedish user group under the leadership of ***Sven Berglund***, University of Örebro, Sweden, is working in the ZA-EUROLAB. The topic of their research work is: "The changing Cleavage Structures in Eastern Europe and the Conditions of European Democracy."

The annual user meeting 1998

The second Annual User Meeting for the participants of the TMR-LSF programme at the ZA-EUROLAB took place on Friday October 2nd 1998.

In the first part of the meeting ***Ekkehard Mochmann*** informed about the "Development and Perspectives of the ZA-EUROLAB and the Infrastructure Access as Objective in the 5th EU Framework Programme. Many of the users were interested in more information on how to apply for EU funds in the social sciences. Thereafter ***Ingvill C. Oedegaard*** presented an overview of access provided in the ZA-EUROLAB during the last year. About 600 access days have been provided so far and another 200 access days have been granted for the next couple of months. Users participating in the programme come from 6 different European countries, another five countries are represented in the allocated 200 access days.

In the second part of the meeting the users introduced their main research interests and research done in the ZA-EUROLAB and answered questions from the audience. The

presentation started with ***Andries van der Ark***, University of Utrecht, the Netherlands (Models on Compositional Data on Comparative Data Sets), followed by ***Judith Bara*** and ***Eric Tanenbaum***, University of Essex, United Kingdom (Monitoring Democratic Five Year Plans: British Manifestos and its Party Platforms), ***Marcel Coenders***, University of Nijmegen, the Netherlands (National Identification and Exclusionism of Ethnic Out-groups), and ***Jocelyn Evans***, European University Institute, Italy (Comparative Modelling of Left-wing Vote Defection to Extreme Right Parties in Europe). Before lunch, **Tor Korneliussen**, University of Bodö, Norway, told how useful the participation in the 1998 ZA-spring seminar had been for consulting other researchers in empirical methods at his home university.

Sylvia Korupp, University of Utrecht, the Netherlands started the second session of the users presentations (Mother's Status and her Influence at her Child's Entry into the Labor Market in the Netherlands). Afterwards, ***Wijbrandt H. van Schuur*** presented his research work (Nonmetric Unidimensional Survey Measurement Using Large scale data sets) followed by ***Bernadette C. Hayes***, Queen's University of Belfast, Northern Ireland (Religious Independents versus Religious Affiliates within established Western Democracies), ***Ursula Tentschert***, University of Vienna, Austria (Mobility in Austria) and ***Matthias Till***, University of Vienna, Austria (The Socio-Economic Situation of Students). Finally, ***Pieter van Wijnen*** presented the progress of his PhD on the Issues and Voting Behaviour in the Netherlands 1971-94. Six Users were not able to participate in the 1998 annual user meeting due to other obligations. The meeting closed with a general discussion of the research presented and the TMR-LSF programme. The evening was spent at a Cologne brewery.

The users emphasised how much they profited from their stay at the ZA-EUROLAB and how interesting and useful it was to get in touch with other scientists at the annual user meetings from which a couple of joint papers and presentations have emerged.

For 1999 there still are vacancies, but it is recommended to apply as soon as possible. For further information concerning the TMR-LSF programme at the Zentralarchiv contact oedegaard@za.uni-koeln.de.

‘Environment’ - ‘Medio Ambiente’

A special CD-ROM edition of the 1993 ISSP

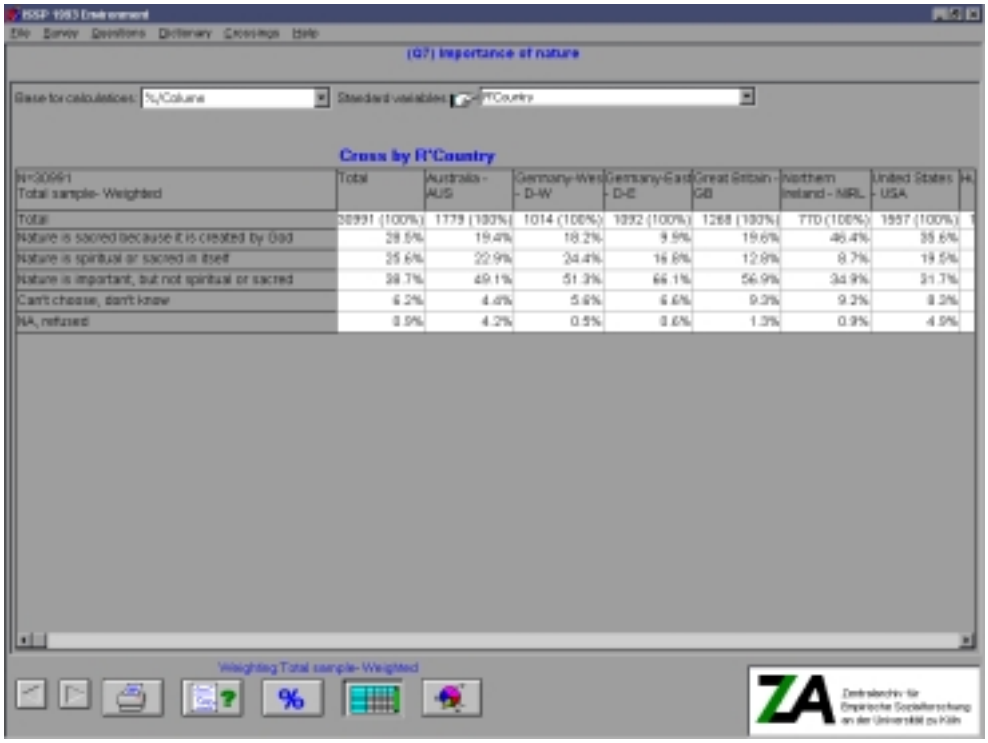
The ‘International Social Survey Programme’ (ISSP), a continuing annual programme of cross-national collaboration between up to 30 participating (in 1998) social science institutions, conducted in 1993 a survey on attitudes toward the ‘Environment’. Data and documentation of this study have been integrated by the Zentralarchiv.

Since 1997 the ZA co-operates with the two Spanish ISSP partners: ASEP (Análisis Sociológicos Económicos y Politícos) and CIS (Centro de Investigaciones Sociológicas). The first CD-ROM produced in co-operation with ASEP/JDS in Madrid is now distributed by ASEP, CIS and the ZA.

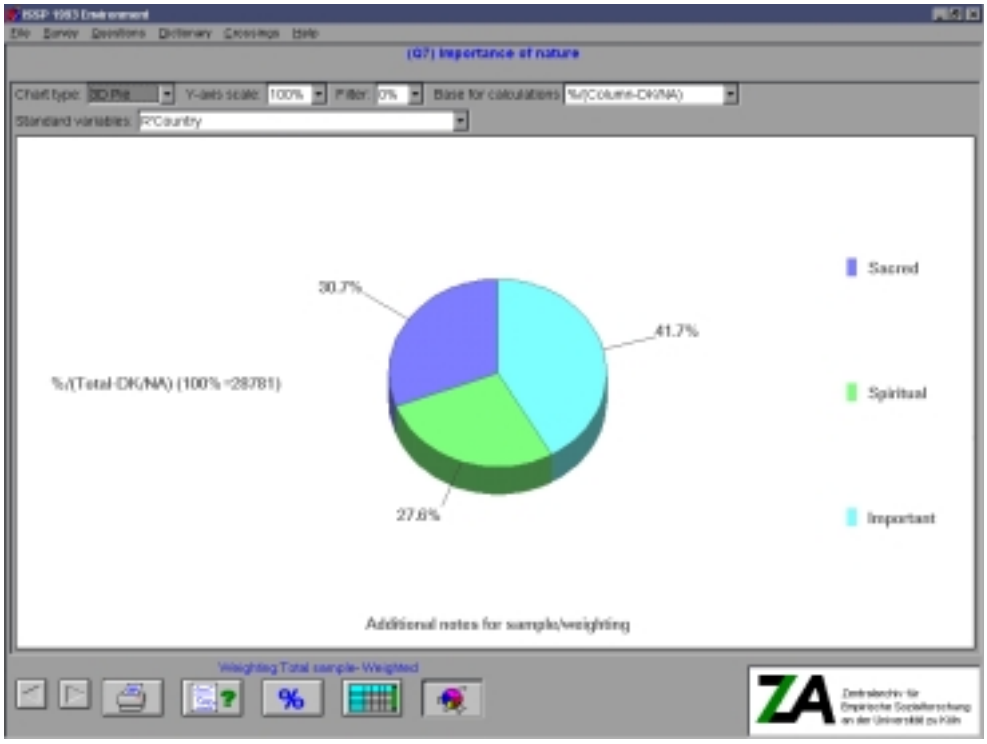
Picture 1

This CD-ROM (1st picture) is designed to provide the community of social science researchers with a source of readily understandable and accessible data deriving from the ISSP 1993 module about environment, with twenty participating countries. It also includes a special environment module featured in five EC countries. The JD Survey Data Explorer is installed as the main data exploration utility, which allows the user to browse on-line the study documentation and the questionnaire index, to print tables, run cross-tabulations and produce graphics just with a mouse-click.

The 2nd picture shows one of the great variety of options which can be used: a standard cross-tabulation (Variable V23 - Q.7 - ‘Importance of Nature’ by V3 ‘countries’), the 3rd picture shows the same variable as a pie-chart (not broken down by countries).



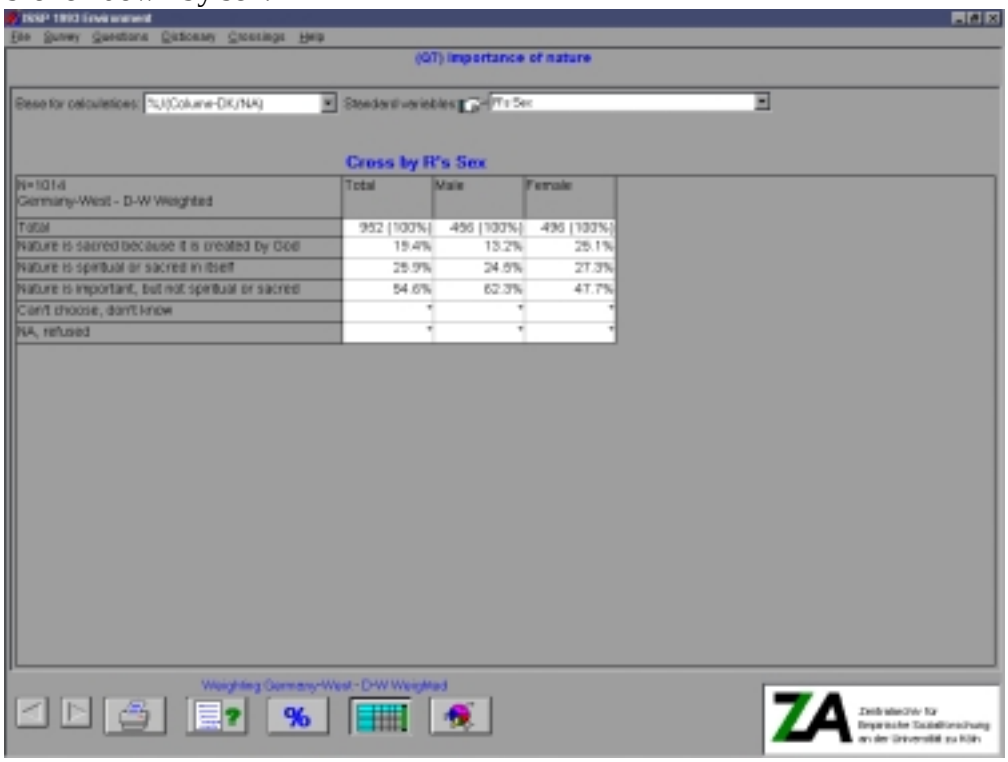
Picture 2



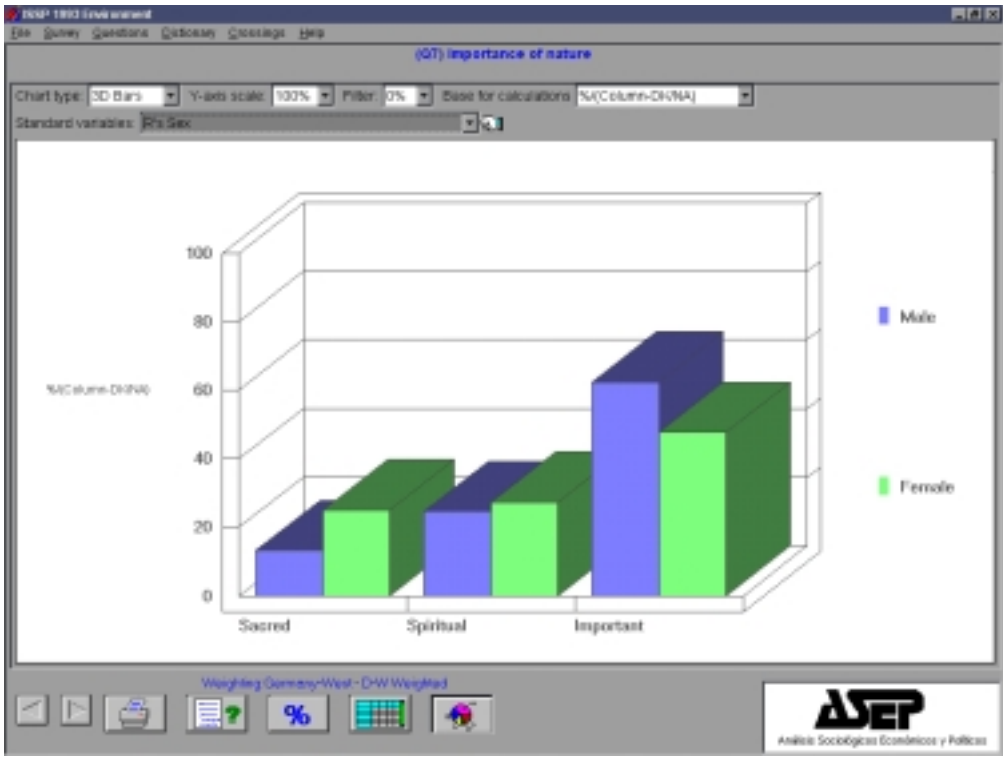
Picture 3

It is also possible to produce results only for special sub-groups of respondents (e.g. by country, sex, age, employment status etc.). The example on pictures 4 and 5 show the same

variable as above (V23 ‘Importance of Nature’) only for respondents from West-Germany broken down by sex.

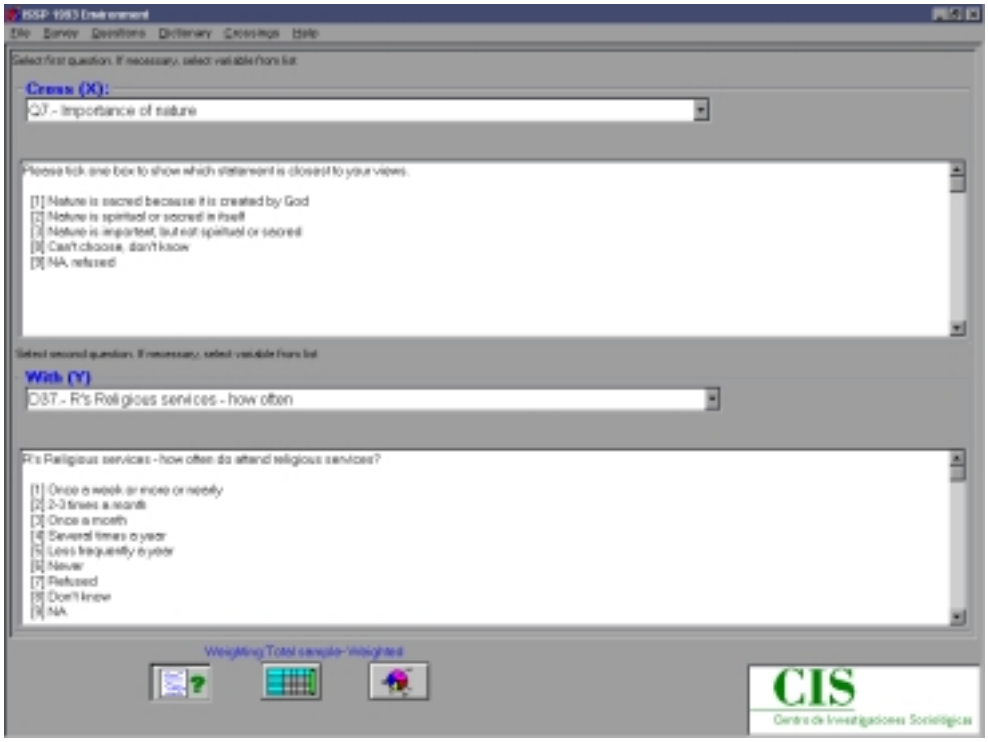


Picture 4

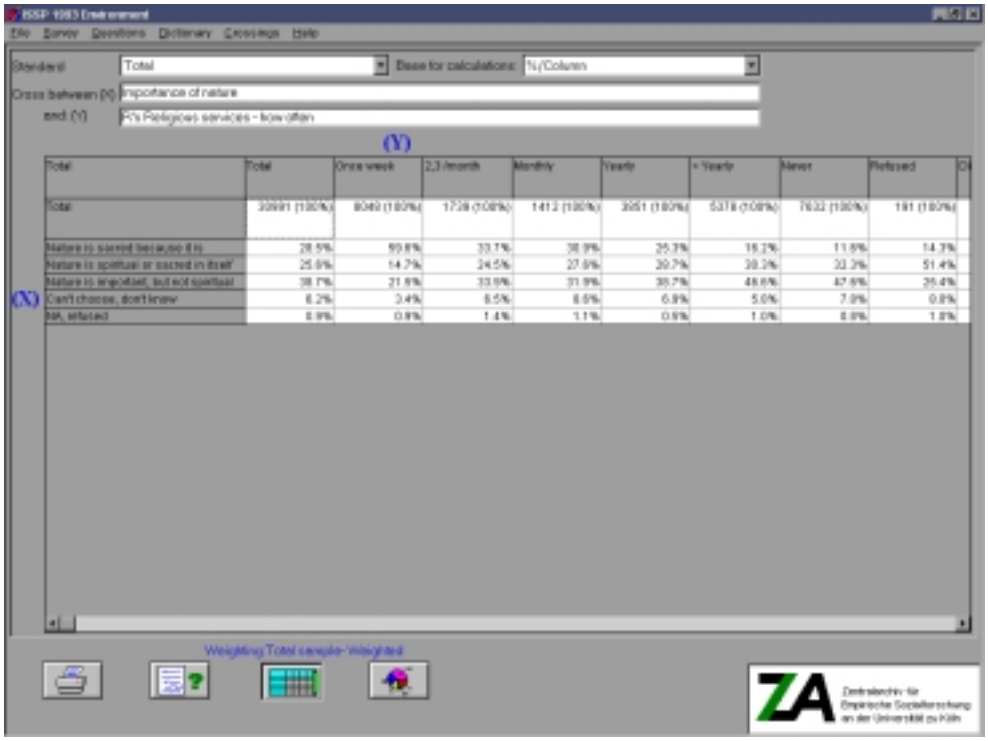


Picture 5

Besides the options shown above it is possible to define and calculate your own cross-tabulations as shown in pictures 6 and 7 (V23 ‘Importance of Nature’ by V277 ‘religious service - how often’)



Picture 6



Picture 7

In addition the CD-ROM includes the SPSS portable file, the machine-readable codebook in ASCII format and the field-questionnaires of all participating countries in PDF format. It is a comprehensive product which can be used on it's own, because of the data-browsing facility, and it can be used for further advanced analysis using SPSS software. The CD-ROM is available at the ZA for the price of US$ 100,--.

Empirical Investigation of Social Space

by Jörg Blasius

Pierre Bourdieu at the Zentralarchiv

From October 7 to 9, 1998, the international conference on "Empirical Investigation of Social Space" was held at the Zentralarchiv für Empirische Sozialforschung. The conference was organised in cooperation with ***Henry Rouanet***, University René Descartes (Paris). The focus of the symposion was the empirical realisation of the theory of the French sociologist ***Pierre Bourdieu***. ***Bourdieu*** divides society into three classes – the upper, middle and lower classes –, which again he subdivides into class fractions. He describes these fractions by means of three types of capital: the economic capital to which belong all sorts of income, the cultural capital which comprises formal education all and other types of "knowledge", e.g. "knowledge of art", and the social capital, i.e. the kind and size of social networks. If the types of capital are taken as latent variables, different groups of society can be described by the composition of their capital. In order to put this theory into empirical practice, ***Bourdieu*** uses correspondence analysis. Both, ***Bourdieu's*** theory as well as correspondence analysis – which too was developed in France – are closely linked to each other, this link being the starting-point of the meeting.

It was the aim of the conference to bring together leading statisticians working in the field of correspondence analysis and well-known social scientists working in the field of ***Bourdieu's*** theory, in order to discuss his ideas. In the three-day conference 70 scientists from 15 countries took part. One participant among the social scientists was ***Pierre Bourdieu*** (Collège de France, Paris), whose books have been translated into almost all important languages and who is the most cited social scientist of the world. In France, ***Bourdieu*** is

also known for his political activities – he was, for instance, personal counsellor of Mitterrand – and his support for striking workers; in Germany he is known for his criticism of the EURO, etc. (cf. respective articles in the newspaper "DIE ZEIT"). The conference program ended with ***Bourdieu's*** talk, which gave him the possibility to refer to the other speakers. His paper linked the applications of his theory and correspondence analysis together. The following open discussion between empirically working social scientists, statisticians and ***Pierre Bourdieu*** came to an end after midnight – in an Italian restaurant.

The following papers were given: The statistical analysis of structured data (***Brigitte Le Roux*** and ***Henry Rouanet***, Paris), Textual data analysis (***Ludovic Lebart***, Paris), Geometric data analysis of sets of matrices (***Michael Greenacre***, Barcelona), Dependency tables: A tool for the visualisation of dependencies in contingency tables and nested line diagrams (***Karl Erich Wolff***, Darmstadt), Formal structures for qualitative theory building (***Rudolf Wille***, Darmstadt and ***Uta Wille***, Zürich), Geometric data analysis of a political space in France today (***Jean Chiche*** and ***Brigitte Le Roux***, Paris), The "Weltanschauungs Space". An attempt to investigate conflicts in mental attitudes by the analysis of popular referendums in Switzerland (***Michael Hermann***, Zürich), Social and attitudinal structure of the visitors of the "Wehrmacht" exhibition (***Peter Schmidt***, Mannheim; ***Norbert Spangenberg***, Frankfurt; ***Karsten Stephan***, Siegen and ***Karl Erich Wolff***, Darmstadt), Exploring the social space of Albania: A stranger's first steps towards familiarisation (***Ivailo Partchev***, Sofia), Ways of improving cluster analyses (***Mart-Jan de Jong*** and ***Jeroen R. M. van de Kamp***, Rotterdam), Skills as cultural capital (***Jürgen Friedrichs*** and ***Jörg Blasius***, Köln), Geometrical and formal approach in order to characterise the quality of employment in the Chilean survey Caracterización Socioeconómica Nacional 1994: A multiple correspondence analysis of structured data (M[a.]***Mercedes Jeria Cáceres***, Santiago de Chile), The field of modern family strategies (***Wolfgang Walter*** and ***Kerstin Hartmann***, Bamberg), Youth values and lifestyles in a post-modern society: An application of HOMALS (***Jacques F. A. Braster***, Rotterdam), The trace of social space (***Martin Munk***, Frederiksberg), The geometric analysis of questionnaires: The lesson of "La Distinction" (***Henry Rouanet***, ***Werner Ackermann*** and ***Brigitte Le Roux***, Paris), Lifestyles of poor people (***Jörg Blasius*** and ***Jürgen Friedrichs***, Köln), Investigating the social space in Luxembourg (***Fernand Fehlen***, Luxemburg), Social structures and their transformations (***Lennart Rosenlund***, Stavanger), Social space and field of higher education. The case of Sweden (***Donald Broady, Mikael Börjesson,*** Uppsala and ***Mikael Palme***, Stockholm), The structure of the French literary field under German Occupation (1940-1944) through correspondence analysis (***Gisèle Sapiro***, Paris), The field of French economists (***Frédéric Lebaron***, Paris), Construction of social space and correspondence analysis (***Pierre Bourdieu***, Paris)), Round Table Discussion: The concept of social space (Moderation: ***Fernand Fehlen***).

International conference on Large Scale Data Analysis

May 25-28, 1999 at the Zentralarchiv: Second Call for Papers

In May 1999, the Zentralarchiv für Empirische Sozialforschung, in collaboration with ***Michael Greenacre*** of the Pompeu Fabra University in Barcelona, is organizing an international conference on the analysis of large data sets. The objective of the conference is to bring together social scientists and statisticians who will have worked on the same sets of empirical data using different methodologies and different philosophies of data analysis.

The methodological topics included in the conference will be: Data quality, Measurement, Sampling and weighting, Equivalence of indicators, Treatment of nonresponses and missing data, Visualisation of categorical data, Data mining and knowledge discovery, Models for categorical data, Structural equation modelling, Multi-level analysis, Scaling techniques

Participants who will present papers at the conference include: ***Erling Andersen*** (Copenhagen), ***Simona Balbi*** (Naples), ***Jaak Billiet*** (Leuven), ***Harold Clarke*** (Texas), ***Tony Coxon*** (Essex), ***Carles Cuadras*** (Barcelona), ***Paul de Boeck*** (Leuven), ***Antoine de Falguerolles*** (Toulouse), ***Edith de Leeuw*** (Amsterdam), ***Rob Eisinga*** (Nijmegen), ***Uwe Engel*** (Potsdam), ***Josep Fortiana*** (Barcelona), ***Michael Friendly*** (Toronto), ***Wolfgang Gaul*** (Karlsruhe), ***John Gower*** (Open University), ***Andrew M. Greeley*** (Chicago), ***Patrick Groenen*** (Leiden***), Jacques Hagenaars*** (Tilburg), ***Joop Hox*** (Utrecht), ***Wolfgang Jagodzinski*** (Köln), ***Ludovic Lebart*** (Paris), ***Allan McCutcheon*** (Nebraska), ***Jacqueline Meulman*** (Leiden), ***Heiner Meulemann*** (Köln), ***Jost Reincke*** (Münster), ***Götz Rohwer*** (Bochum), ***Tamas Rudas*** (Budapest), ***Albert Satorra*** (Barcelona), ***Peter Schmidt*** (Mannheim), ***Victor Thiessen*** (Halifax), ***Karl Erich Wolff*** (Darmstadt), ***Andries van der Ark*** (Tilburg), ***Peter van der Heijden*** (Utrecht), ***Jeroen Vermunt*** (Tilburg) and ***Wijbrandt van Schuur*** (Groningen).

The database on which the conference is centred is from the International Social Survey Programme (ISSP), a voluntary grouping of study teams in over 20 countries. Each country undertakes to include common questions as a supplement to their regular national surveys. The topics change from year to year by agreement, and are replicated every five years or so. Some examples of topics are: the role of government, social inequality, family and changing sex roles, religion, and environment.

The ISSP data to be considered in the conference are available in the form of a compact disk (CD-ROM), which contains data sets for each year from 1985 to 1995, for several countries at a time. Data files are supplied both in ASCII and SPSS format. Codebooks, variable lists, as well as the PDF files of scanned pages of all the original questionnaires are also distributed on the CD-ROM. This database gives a unique opportunity for making comparative social research between different countries on the same topic, as well as investigating trends over time when the study is replicated.

Papers presented at the conference based on analysis and substantive interpretation should satisfy the following two minimum prerequisites: (1) the analysis should have a clearly stated social science or methodological motivation, including a justification for the particular selection of variables; (2) some cross-national comparison must be included in the analysis and the results. We particularly encourage methodologies which permit the interpretation and understanding of large amounts of data.

Prospective participants should contact the Zentralarchiv to request a CD-ROM with the complete database available at present. The cost of this CD-ROM is 50 DM, but this cost will be deducted from the registration fee of those attending the conference. There is a Web page giving details of the ISSP, including codebooks of the surveys, variable lists and crosstabulations of each variable across countries, at the following address:
http://www.za.uni-koeln.de/issp
Some notes on the contents of the CD-ROM are contained in a Word document or PDF file available from the Web page:
http://www.econ.upf.es/databases

Abstracts should be submitted to the organizers by **31 January 1999**, and should be approximately two pages in length, containing the following information: (1) title, (2) author(s), (3) affiliation(s), (4) keywords and (5) a summary, including the aim of study, the surveys/countries/variables used and the statistical methodology employed.

The conference fee will be 150 DM (of which 50 DM will be refunded to those having purchased the CD-ROM). Deadline for registration for the conference (without presenting a paper) is **31 March 1999**.

For further information and requests for data, please contact:

Jörg Blasius
Zentralarchiv für Empirische Sozialforschung, Bachemer Str. 40, D-50931 Cologne
Fax: +49 221 4769444
E-mail: blasius@za.uni-koeln.de
or ***Michael Greenacre*** at e-mail: michael@upf.es

Analysis of Time Dependent Data

28th Spring Seminar at the Zentralarchiv

February 22 - March 12, 1999

The spring seminar is a training course for social scientists interested in advanced techniques of data analysis and in the application of these techniques to data. Participants must have a sound basic knowledge of statistics as well as experience in the handling of PCs and of working with SPSS. The spring seminar comprises lectures, exercises and practical work using personal computers. While in the lectures the logic of models and the corresponding analysis strategies will be explained, during the exercises and in the practical work the participants are given the opportunity to apply these methods to data. In the course of the spring seminar there is also the possibility to get to know the functions and services of the Zentralarchiv which is the German data archive for survey data. As in the past spring seminars, the focus will be on teaching multivariate analysis techniques, this time it will be the analysis of time dependent data. In the first two weeks, for the practical work we will use the data from the German Socio-Economic Panel (GSOEP), in the third week we will use different data sets.

For those who want to keep the data after the seminar, the GSOEP is available directly from the DIW at 50.- DM, please email "soepmail@diw-berlin.de". In the case of not buying the data in advance, one has to sign a document stating that one will use the data for the spring seminar only. If somebody wishes to purchase the data, the price of DM 50.- is a special price for spring seminar participants only. Interested people not registered for the spring seminar have to pay the usual price of 130,- US dollars.

The seminar covers three modules of one week each, to some extent based upon one another, which can be booked either separately or as a block: Introduction to the GSOEP, Panel Analysis, Event History Analysis. In the first week there will be a detailed introduction to the German Socio-Economic Panel. The GSOEP is a panel study which has run every year since 1984, it includes about 7 000 households with about 15 000 persons to be asked. Topics: Longitudinal survey design: aim and main forces of the study, longitudinal and cross-sectional research questions; GSOEP-methodology: subsamples and sample size, the follow-up technology, the concept of new households and new persons, survey instruments and interview methods; developments of sample size and its determinants: attrition and analysis of drop-outs, weighting - cross-sectional and longitudinal, panel conditioning

- a real problem?; spell-data: data files with spell information on income, employment, marital status; the GSOEP in international panel comparability studies: general problems of comparability projects with micro data, panels of other countries such as the panel study of income dynamics (PSID). In the afternoons we will work with the GSOEP data.

The second week deals with the analysis of panel data. Whereas in the first week the GSOEP will be introduced, in the second week we will discuss more advanced problems when analyzing panel data. Topics: panel analysis in longitudinal methodology; analyzing change: a survey of major models; models for univariate and multivariate change; categorical panel data analysis; latent class analysis; structural equation modelling of change; random effects growth models; individuals, households, and regions: multilevel analysis of panel data; analyzing incomplete panel data.

The third week deals with event history data analysis. While panel analysis is applied to data where one has a limited number of time points in which the variables have been developed, event history analysis requires a continuous measuring over time. But there are many variables in the social sciences where one knows the exact point in time when the event happened, for example: birth of a child, marriage and divorce, changing the job situation, and finishing school or university. Topics: time, events, collectives and situations; temporal data structures; distribution functions, densities and rates; non-parametric estimation of distribution-functions and rates; regression models for retrospective data; regression with interval-valued dependent variables; models for the occurrence of events; dynamic regression models; models for several events. As program we will use TDA which is written by the lecturers and available free of charge (www.stat.ruhr-uni-bochum.de).

The lectures will be given in English by:
Dr. ***Joachim Frick*** and Dr. ***Jürgen Schupp***
German Institute for Economic Research (DIW), Berlin
Introduction to the GSOEP
February 22 - 26, 1999

Prof. Dr. ***Uwe Engel***
University of Potsdam
Panel Analysis
March 1 - 5, 1999

Prof. Dr. ***Götz Rohwer*** and Dr. ***Ulrich Pötter***
University of Bochum
Event History Analysis
March 8 - 12, 1999

In addition to the lectures, information on the services of the GESIS and the Zentralarchiv will be given. Together with the lecturers, the Zentralarchiv members ***Jörg Blasius***, ***Markus Klein***, ***Yasemin Niephaus*** and ***Maria Rohlinger*** are going to be the instructors of the afternoon working groups in which the seminar participants will put into practice what they learned in the morning lectures.

The participation fee for the spring seminar is 100,-- DM per week. Students and social scientists out of work pay only half, on presenting an attestation of their status. Participants from Germany might be interested to learn that the spring seminar is recognized as "Bildungsurlaub" and application forms to the employer can be requested from the Zentralarchiv. Travelling and accommodation costs have to be paid by the participants.

A limited number of bursaries is available under the TMR-LSF programme to cover travel and subsistence costs for participants from EU countries (except Germans located in Germany). Please check our Web pages (http://www.za.uni-koeln.de/eurolab) for the conditions of taking part in this programme or contact ***Ingvill Oedegaard*** (oedegaard@za.uni-koeln.de) for further information.

Registration deadline is 29 January 1999. Please use the registration form which you will find at the end of this newsletter and send it to the Zentralarchiv; registration by e-mail is also possible. As the number of participants is limited to 50 persons per week we recommend to register as soon as possible.

Zentralarchiv für Empirische Sozialforschung, Universität zu Köln, Bachemer Str. 40, D-50931 Köln. Tel. ++49-221-4769433 or ++49-221-4703155, Facsimile: ++49-221-4769444, e-mail: blasius@za.uni-koeln.de.

The Zentralarchiv WWW Server provides you with up-to-date information on the spring seminar (http://www.za.uni-koeln.de/events/).

Registration form

Zentralarchiv für Empirische Sozialforschung
University of Cologne, Postfach 41 09 60, D-50869 Köln

Registration for Spring Seminar 1999 (Please use capitals)

Name .....

Address .....

Phone/Facsimile .....

E-mail .....

Studied at .....

Examination.....

Present occupation.....

Institution.....

I want to participate in

- ❑ the whole spring seminar, 22 February - 12 March, 1999
- ❑ parts of the spring seminar

(Please tick off appropriate box)

❑ Dr. ***Joachim Frick*** and ***Dr. Jürgen Schupp***
German Institute for Economic Research (DIW), Berlin
Introduction to the GSOEP
February 22 - 26, 1999

❑ Prof. Dr. ***Uwe Engel***
University of Potsdam
Panel Analysis
March 1 - 5, 1999

❑ Prof. Dr. ***Götz Rohwer*** and Dr. ***Ulrich Pötter***
University of Bochum
Event History Analysis
March 8 - 12, 1999

Date: ..... Signature: .....

Registration deadline: 29 January 1999

Weiterbildung in statistischer Datenanalyse beim ICPSR Summer Program 1999

Im Rahmen der deutschen Mitgliedschaft im Inter-University-Consortium for Political and Social Research (ICPSR) koordiniert das Zentralarchiv die Anmeldung für das jährlich im Juli/August stattfindende 'Summer Program' in Ann Arbor, Michigan, USA. Das achtwöchige Ausbildungsprogramm (kann u.U. auch in zwei vierwöchige terms aufgeteilt werden) besteht aus einer Reihe von Kursen, die nach den Vorkenntnissen in Mathematik, Statistik und Methoden der empirischen Sozialforschung in unterschiedliche Schwierigkeitsstufen ('tracks') eingeteilt ist.
Das Lehrprogramm sieht Vorlesungen und 'Workshops' vor. In diesen Workshops wird der theoretisch/statistische Lehrstoff direkt - im Sinne des learning by doing - mit der Datenauswertung verbunden. Übungsdaten, die gängigen Datenanalysepakete (SPSS, SAS etc.) sowie spezialisierte Einzelprogramme stehen zur Verfügung. Folgende Workshops, darunter auch sehr spezielle Themen, werden angeboten (Auswahl):

First Term
Mathematical Models: Game Theory; Quantitative Historical Analysis; Introduction to Statistics and Data Analysis I; Introduction to Regression Analysis; Regression Analysis; Multivariate Statistical Methods; Scaling and Dimensional Analysis; Maximum Likelihood: Estimation for Generalized Linear Models; Spatial Analysis / Geographic Information Systems
Second Term
Time Series Analysis; Mathematical Models: Rational Choice; Quantitative Analysis of Crime and Criminal Justice; Introduction to Statistics and Data Analysis II; Categorical Data Analysis; "LISREL" Models: General Structural Equations; Advanced Analysis of Variance: Experimental Design; Simultaneous Equation Models; Regression Analysis; Mathematical Statistics

Das Programm ist umfangreicher als hier dargestellt werden kann. Das ICPSR hat das gesamte Ausbildungsangebot in einer Broschüre zusammengestellt. Für Interessenten ist diese beim Zentralarchiv erhältlich. Wenngleich sich die Darlegungen in diesem Heft auf das diesjährige 'Summer Program' beziehen, kann im allgemeinen davon ausgegangen werden, daß sich an der Grundstruktur des Angebotes im kommenden Jahr kaum etwas ändert. Die Programmübersicht für das Jahr 1999 wird in der Regel erst Anfang April verfügbar gemacht. Die Anmeldung erfolgt bis Ende April. Über Möglichkeiten der Finanzierung eines Aufenthaltes in Ann Arbor sollten sich Interessenten bereits zum Ende dieses Jahres orientieren. Ansprechpartner im Zentralarchiv ist, auch für Hinweise auf Förderhilfen:
Maria Rohlinger Telefon: 0221 / 4 76 94 - 45, Sekr. -33
e-mail: rohlinger@za.uni-koeln.de
'Summer Program' im World Wide Web: http://www.icpsr.umich.edu

Linux, ein Unix in neuem Gewand

von Bruno Hopp

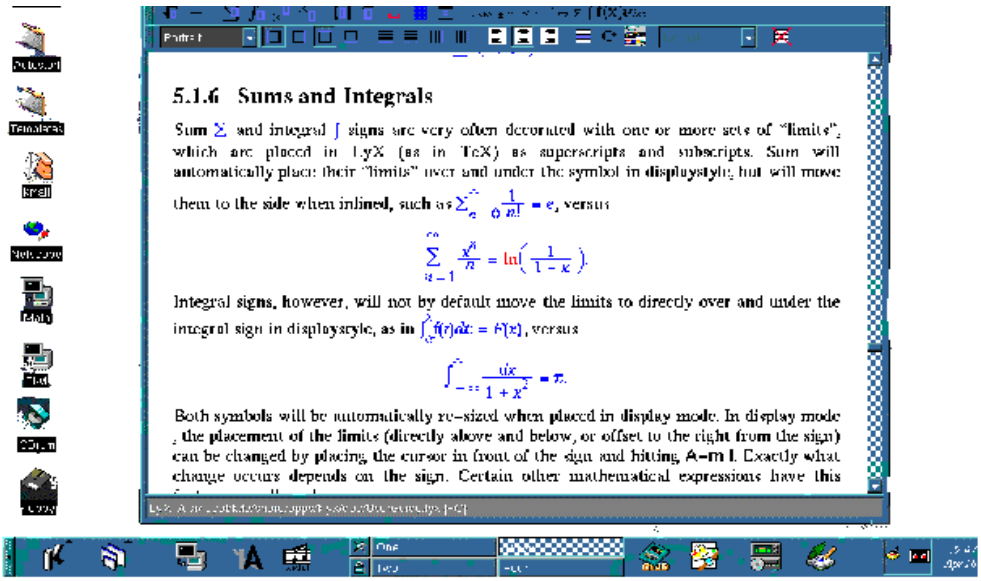
In Client-/Server basierten IP-Netzwerken wird zunehmend als Serverplattform zu Windows NT gegriffen. Langjährige Erfahrungen mit Betriebssystemen des gleichen Herstellers sind ein Kapital, das niemand gern über Bord wirft. Warum ein Blick über den Tellerrand lohnend sein kann, möchte der folgende Artikel aufzeigen.

Linux entstand aus einer Skizze des finnischen Studenten ***Linus Torvalds*** und basierte auf MINIX, einem Unix-Derivat zu Lehrzwecken. Schnell wuchs eine noch kleine engagierte Gemeinde heran, die in unzähligen Stunden Freizeit mit viel Enthusiasmus ein lauffähiges System programmierte. Als der Systemkern einigermaßen stabil wurde und die freie (=unentgeltliche Software), die im Internet schlummerte, nutzbar wurde, wuchs die Zahl der Linux-Anhänger weiter. Zahlreiche Programme unterliegen heute der „GPL“ (GNU General Protection Licence). Sie bedeutet u.a., daß Software kostenfrei entwickelt und nur inkl. des Quellcodes und der Dokumentation weitergegeben werden darf, so daß die Urheberrechte geschützt werden. Hier liegt ein Schlüssel zur ständigen weiteren Verbesserung existierender Software. Ohne Internet resp. Usenet wäre die rasante Entwicklung kaum denkbar gewesen. Das Grundkonzept von Linux: ein an UNIX orientiertes multi-user- und multi-tasking-fähiges Betriebssystem (POSIX-Konformität), ist Ziel und Verpflichtung zugleich. Linux ist lauffähig ab einem i386DX mit nur 4 MB Speicher und ca. 50 MB Plattenplatz (Mindestanforderung), wobei vorhandene Hardware sehr effizient ausgenutzt wird (empfohlene Minimalwerte sind dazu da, daß man sie mindestens verdreifacht, um Freude bei der Arbeit zu haben...).

Neben der freien Verfügbarkeit von Linux fällt auf, daß es zahlreiche „Distributionen“ gibt: sie kosten zwar (vgl. Tabelle am Ende des Beitrages), jedoch gilt der Obulus nicht dem Programmcode oder einem kommerziellen Softwarehersteller, sondern dem Aufwand, Mengen an (freier!) Software aus dem Internet zu einem lauffähigen System zu kombinieren und dies allgemein verständlich zu dokumentieren. Allein im deutschsprachigen Raum gibt es mehr als drei verschiedene Firmen, die dies leisten. Wer bei frei verfügbarem C-/C++-, Fortran-, Pascal-, Ada- und Java-Compiler sowie modernen Scriptsprachen (Perl, Tcl/Tk) Gelegenheit hat, Publikationen mit dem Satzsystem LaTeX zu setzen, emails zu schreiben, Grafiken zu bearbeiten, Datenbanken zu erstellen..., hat ein System, das professionellen Ansprüchen an Produktivität jederzeit gewachsen ist – bei höchster Stabilität. Allein das Grafikprogramm „Gimp“ wird in der Leistung häufig mit dem bekannten „Photoshop“

verglichen. Literaturlisten im BibTeX-Stil, wie sie in (wissenschaftlich geprägten) newsgroups gängig sind, werden direkt editiert. Programme für Texte verfassen, Tabellen berechnen etc. gibt es zuhauf, komplette Officesuiten gleich von mehreren Herstellern – auch wenn mancher Nutzer nach gewisser Gewöhnung die Vorzüge z.B. des emacs neu entdecken wird. Ergänzt wird die Desktop-Funktionalität durch Netzwerkeinbindung: NFS und NIS, SMB-Protokoll für Windows-Netze, IPX für Novellnetze, POP3, Faxe senden und empfangen usw. sind integriert. Wer mag, kann auch exotische Hardware mit seinem Rechner betreiben – das frühere „geht nicht!“ weicht einem „wie kann ich das erreichen?“. Festplattenlose Arbeitsstationen bzw. „X terminals“ holen sich beim Booten alle Information bei einem Server. Was bleibt? Wer früher beim Gedanken an Unix eine Gänsehaut aufgrund unverständlicher Befehlszeilen bekam, sei beruhigt – grafische Oberflächen à la Windows95 und NT (=Fvwm95) sind zahlreich vertreten. Der neue Windowmanager KDE (http://www.kde.org) bietet Entwicklungen, die mancher Marktführer sicher erst in einigen Jahren erfinden wird... Ein weltweit verstreutes Programmiererteam hat nach weniger als drei Jahren den KDE in der Version 1.0 vorgestellt, eine integrierte grafische Umgebung mit Editor, email- und newsclient, Dateimanager, Verwaltungswerkzeugen für Netzwerk etc. Serverfunktionalität ist kein Handicap für eine ansprechende, grafische Arbeitsumgebung mit ausgeklügelter Benutzerführung:

Abbildung: Bildschirm des KDE Windowmanagers mit Klyx



Wer DOS-Applikationen weiter benutzen möchte oder muß, kann dies via „dosemu“, für 16-bittige Windowsanwendungen gibt es ein kostenpflichtiges Modul „Wabi“ (Anbieter: Caldera), welches sogar ältere Word-, Excel- und Access-Versionen zum Zuge kommen läßt. Sicherheit wird unter Linux groß geschrieben: jeder Benutzer bekommt zugewiesen, mit welchen Dateien er bzw. sie normalerweise arbeiten darf, so daß Computerviren wie bei DOS-Systemen (wo ein Benutzer jederzeit alles darf) unbekannt sind. Etwas, das der

Marktführer durch Einführen von sog. „Profilen“ seinen Systemen wieder mühsam beibringen muß.

Eine Quintessenz für sozialwissenschaftliche Anwender? Linux hat sich in wenigen Jahren vom belächelten Hackerspielzeug zu einem ernstzunehmenden Betriebssystem mit beachtlichen Netzwerkfeatures entwickelt. Die wachsende Verfügbarkeit kommerzieller und wissenschaftlicher Anwendungen unter Linux: Datenbanken Yard, Postgres95 oder Adabas D, Bürolösungen Applix und StarOffice, Mathematikpakete Maple V und Mathematica erschließen weitere (nicht nur akademische) Einsatzgebiete. Firmen wie Corel/WordPerfect und Oracle haben angekündigt, ihre wichtigsten Produkte bis Jahresende 1998 für Linux zu portieren. Mit der Verfügbarkeit grafischer Oberflächen (GUIs) hat Linux eine Chance auf breite Akzeptanz auch im PC-Bereich. Organisationen wie die NASA (Einsatz der Debian-Distribution bei Shuttle-Expeditionen) sind durch ihre Vorbildfunktion Wegbereiter für eine ‚Softwarekultur‘ auf neuen Wegen: noch bessere Netzwerkanbindung, stabile Systeme. Erkauft wird soviel Flexibilität mit fehlender professioneller Unterstützung - jeder ist zuerst sein eigener Support, weil hinter Linux selbst keine Firma, kein Konzern steht (auch wenn manches moderne ‚Systemhaus‘ schon Linux-Unterstützung anbietet). Das eigene Zeitbudget zum (Neu-) Lernen ist einziges Hemmnis. Eine Warnung: Linux bietet *volle* IP-Funktionalität wie alle Unix-Varianten - und weil im Internet (fast) alles über IP funktioniert, können verantwortungsvolle Administratoren den Linux-Einsatz in Instituten und Universitäten nur nach konkreten Absprachen genehmigen. Linux ist stark in das Blickfeld professioneller Anwender gerückt, die Wert auf Sicherheit der Daten und Stabilität für die Benutzer legen. Knappe Budgets begünstigen die Verbreitung von preiswerten, nicht kommerziellen Systemen, die gut dokumentiert und leicht zu pflegen sind.

Literatur zum Thema

Diedrich, Oliver: Sixpack. Sechs aktuelle Linux Distributionen im Vergleich. In: c’t, Heft 19/1998, S. 176-190.

Hetze, S./D. Hohndel/M. Müller/O. Kirch: Anwender Handbuch Linux und Leitfaden für die Systemverwaltung. Berlin 1995 (5.Aufl. u.ö.).

Kanera, Timo: Pinguin statt eNTe. In: Linux-Magazin 8/1998, S. 40-47.

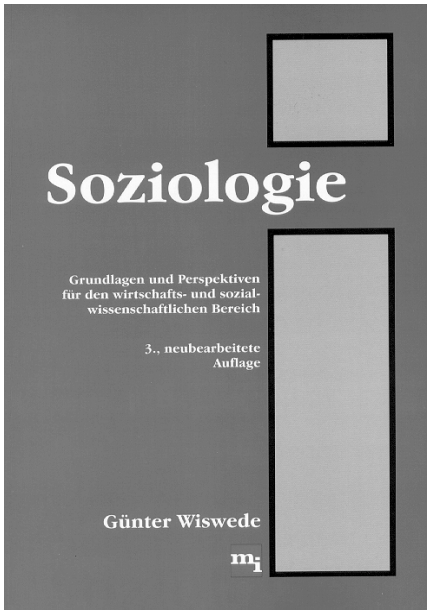
Köhntopp, Kristian: Spinne im Netz. In: c’t, Heft 20/1998, Seite 204-212.

Kofler, Michael: Linux Installation, Konfiguration, Anwendung. Bonn 1998 (3.Aufl.).

Röhrig, Bernhard: Datenbanken mit Linux. Kassel 1998.

Schulze, Martin: Das leisten Linux-Pakete. In: CHIP, Nr. 9/1998 (September), S.126-133.

Buchhinweise



Günter Wiswede: Soziologie. Grundlagen und Perspektiven für den wirtschafts- und sozialwissenschaftlichen Bereich, 3. neubearbeitete Auflage, Landsberg am Lech: Verlag moderne industrie 1998, 452 Seiten, DM 68,--, ISBN 3-478-39203-9.

Ein Buchhinweis von *Günter Endruweit*[1]

Von dem, was die USA im Überfluß haben, haben wir manchmal zu wenig, so auch Lehrbücher zur Einführung in die Soziologie. Das neueste und sicherlich eines der besten von ihnen ist, wiederum überarbeitet, gerade in dritter Auflage erschienen. Es eignet sich ganz hervorragend als Begleitlektüre zu einem Einführungskurs für Studienanfänger.

Das erste Kapitel bringt auf etwa 100 Seiten eine Einführung in allgemeine Aspekte des Faches. Gegenstand, historische Entwicklung und Typen von Grundbegriffen werden dargestellt. Danach werden Wissenschaftstheorie (mit mehreren Richtungen) und Forschungstechniken skizziert, und zwar vor allem in Hinblick darauf, daß der Leser erkennt, von welcher Qualität soziologische Erkenntnisse sein können. Den Abschluß bildet ein sehr knapper (aber i.d.R. folgen zu diesem Thema ja weitere Lehrveranstaltungen) Überblick über das, was manche viel zu verspielt Paradigmen zu nennen pflegen, obwohl es sich um Theorien handeln sollte, die mehr Mühe verlangen.

Die Hauptthemen der Soziologie werden im zweiten Kapitel im wesentlichen an Grundbegriffen expliziert, und zwar unterteilt in mikrosoziologische (Lernen, Interaktion, Rollen und Abweichung auf rund 80 Seiten) und makrosoziologische (Kultur, Ordnung, System, Konflikt, Ungleichheit und Wandel mit zusammen 120 Seiten).

---

1 ***Günter Endruweit*** ist Professor am Institut für Soziologie der Christian-Albrechts-Universität zu Kiel, Olshausenstr. 40, 24118 Kiel.

Diese zur Zeit modische, aber nur bei besonderer Sorgfalt auch erkenntnisfördernde Unterscheidung von terminologischem Groß- und Einzelhandel wird im dritten Kapitel (etwa 100 Seiten) in den einzelnen Unterabschnitten fortgesetzt. Diese behandeln zuerst die Wirtschaft als Teilbereich der Gesellschaft und dann soziale Aspekte einzelner Wirtschaftsfunktionen (Produktion, Koordination und Konsumtion).

Wiswede zielt mit seinem Werk vor allem auf die Wirtschaftswissenschaftler. Das ist sehr berechtigt; denn hier könnte es, wenn eine solche Soziologie für alle zum Pflichtfach gehörte, so manche Koofmich-Type mit beschränktem theoretischem Horizont vor dem Abgleiten in die Asozialität bewahren, wie es jetzt gerade mit manchem Globalisierungsstreben droht. Andererseits ist die Einschränkung des Verfassers zu bescheiden; denn manchem Schöngeist unter den Politikern, Lehrern und Journalisten aus unseren Philosophischen Fakultäten könnte dieses Buch die Augen für die äußerst wichtige gesamtgesellschaftliche Funktion der Wirtschaft öffnen.

Das Buch vermeidet Nachteile der amerikanischen Vorbilder (überflüssige Bebilderung, zu große Vereinfachung von Theorien, ethnozentrische Problemdarstellung durch Alltagsbeispiele, weil oft für college kids gedacht) und mancher deutscher Konkurrenzprodukte (Darstellung letztlich doch nur des eigenen Denkens, während hier die Pluralität deutlich gemacht wird, obwohl auf S. 17 ehrlich die eigene - übrigens mir äußerst vernünftig erscheinende - Position klargemacht wird). Vorteile sind u.a. eine klare, eingehende Gliederung, Stichworte am Rande der Seiten, strukturierender Fettdruck im Text, Literaturhinweise und Kontrollfragen am Ende eines jeden Abschnitts, Ablaufschemata und Begriffshierarchien sowie andere graphische Darstellungen als Gedächtnisstützen. Gegenüber der zweiten Auflage wurde manches gekürzt, einiges davon (so die für unsere oft ahistorisch denkenden Studierenden sicherlich aufschlußreiche Abbildung auf S. 352) bedauerlicherweise; aber die Neuerungen brauchten Platz und sind durchweg nützlich (so die erweiterte Berücksichtigung der Rational-choice-Theorie). Insbesondere aber wird die Soziologie hier nicht als ein Wörterlager für abwechslungsreiche Beschreibungen vorgestellt, sondern als eine empirische Wissenschaft, deren Arbeit dem besseren Verstehen der Gesellschaft dient und damit möglicherweise gar dem sozialen Fortschritt im Sinne einer gezielten Verbesserung ihrer Existenz. Nur damit wird die Soziologie ihren ungeschmälerten Fortbestand erreichen können. Denn in Zeiten von Sparzwang wird ein vernünftiger Souverän zuerst die Gaukler entlassen und zuletzt die Köche.

Wilhelm Bürklin und *Markus Klein*: Wahlen und Wählerverhalten. Eine Einführung, Leske+Budrich, Opladen 1998, 2. Auflage, DM 32,70, ISBN: 3-8100-1978-X.

Ein Buchhinweis von Dieter Ohr[1]

Bei diesem Buch handelt es sich um die vollständig überarbeitete und aktualisierte Neuauflage des 1988 erschienen Buches von ***Wilhelm Bürklin***, das den Titel "Wählerverhalten und Wertewandel" trug. Die Autoren verfolgen den Anspruch, eine Orientierung stiftende Einführung in die Wahlforschung zu geben und gleichzeitig eine Bestandsaufnahme des mittlerweile sehr differenzierten Wissens über den Wähler vorzulegen. Eine solche Bestandsaufnahme ist in mehrfacher Hinsicht wünschenswert. Zum einen sind in den letzten Jahren gravierende Veränderungen im Wahlverhalten selbst zu konstatieren - man denke etwa an die vieldiskutierte Individualisierung oder die gemeinhin behauptete Zunahme an Politikverdrossenheit. Zum anderen gab es eine rasante Entwicklung auch bei den Methoden und Analyseinstrumenten der empirischen Wahlforschung.

Um es vorwegzunehmen: Den beiden Autoren ist es in eindrucksvoller Weise gelungen, ihren Anspruch einzulösen. Mit "Wahlen und Wählerverhalten" liegt nun eine lesenswerte und fundierte Einführung sowohl in die Grundlagen wie auch in aktuelle Forschungsfelder der empirischen Wahlforschung vor, die im deutschen Sprachraum praktisch ohne Konkurrenz ist.

Im Aufbau des Buches lassen sich zwei Teile unterscheiden. Im ersten Teil, der die Kapitel 1 bis 4 umfaßt, werden grundlegende Begriffe, Methoden und Theorien der Wahlforschung vorgestellt und erläutert. Ein Überblick über die wichtigsten Theorien und Konzepte findet sich in Kapitel 2. Formen und Methoden der Wahlforschung werden im 3. Kapitel diskutiert, unter anderem die Aggregatdatenanalyse, die Individualdatenanalyse und Umfrageforschung sowie Verfahren zur Gewinnung von Wahlprognosen und Hochrechnungen. Den für die deutsche Wahlforschung so überaus einflußreichen frühen amerikanischen Wahlstudien, darunter "The People's Choice" von ***Lazarsfeld u.a.*** und "The American Voter" von ***Campbell u.a.*** widmet sich das vierte Kapitel, in dem besonders das sozialpsychologische Grundmodell in seinen Grundzügen und in verschiedenen Spielarten vorgestellt wird. Der zweite Teil des Buches (Kapitel 5 bis 8) stellt im wesentlichen eine syste-

1 Dr. ***Dieter Ohr*** ist wissenschaftlicher Assistent am Institut für Angewandte Sozialforschung der Universität zu Köln, Greinstraße 2, 50939 Köln.

matische Darstellung des Forschungsstands der Wahlforschung dar. Kapitel 5 zeigt auf, in welcher Weise das - dem amerikanischen Kontext entstammende - sozialpsychologische Erklärungsmodell der Wahlforschung auf die Bundesrepublik übertragen wurde. Dabei werden unter anderem thematisiert: die Operationalisierung der Parteiidentifikation, ein zentrales Konzept dieses Ansatzes, die sog. Normalwahlanalyse sowie die in der Disziplin kontrovers diskutierten Thesen einer Individualisierung und eines Dealignment des Wahlverhaltens. Rational-Choice-Theorien des Wählerverhaltens werden im sechsten Kapitel behandelt. Neben den traditionellen Themen des Rational Choice-Ansatzes wie den räumlichen Modellen der Parteienkonkurrenz oder dem sog. Wahlparadoxon wird mit der Theorie des expressiven Wählens auch eine noch relativ neue theoretische Variante vorgestellt, die verspricht, einige zentrale Schwächen des Rational Choice-Ansatzes zu überwinden.

Vor dem Hintergrund eines umfassenden sozialen Wandels und einer damit einhergehenden geringer gewordenen Prägekraft von Sozialstruktur geraten Wertorientierungen als langfristig-stabile Bestimmungsgrößen des Wählerverhaltens stärker in das Blickfeld des wissenschaftlichen Interesses. So ist es nur folgerichtig, daß auch und gerade in einem Lehrbuch zur Wahlforschung die Theorien des gesellschaftlichen Wertewandels eine ausführliche Behandlung erfahren (7. Kapitel). Das achte und letzte Kapitel schließlich präsentiert eine Auswahl wichtiger Diskussionsfelder der gegenwärtigen Wahlforschung, das Problem der sinkenden Wahlbeteiligung, das Wahlverhalten in den neuen Bundesländern, die Wahl rechtsextremer Parteien, der Einfluß der Entwicklung moderner Massenmedien auf Politik und Wahlverhalten sowie die vielfach behauptete Personalisierung von Politik.

Besonders positiv ist an dem Buch hervorzuheben, daß es nicht allein die wichtigsten Theorieansätze und empirischen Befunde auf eine interessante und systematische Weise darstellt. Es wird darüber hinaus auch immer anschaulich demonstriert, welche Entwicklung die in der Wahlforschung verwendeten Meßinstrumente durchlaufen haben und in welcher Weise dies Rückwirkungen auf die theoretische Diskussion und die empirischen Ergebnisse hatte. Sicher gibt es auch an diesem Buch Ansatzpunkte für Kritik. So geben etwa die Passagen zur Aggregatdatenanalyse nur in Teilen das in diesem Gebiet vorherrschende Erkenntnisinteresse und besonders die neueren Entwicklungen wieder. Jedoch ist diese Kritik nur von marginalem Gewicht.

"Wahlen und Wählerverhalten" ist eine vorzügliche Einführung in die Wahlforschung, die nicht nur für die universitäre Lehre und für Praktiker von großem Nutzen sein wird, sondern auch einem breiteren Leserkreis interessante Einsichten in ein für moderne Demokratien wichtiges Forschungsfeld bietet.

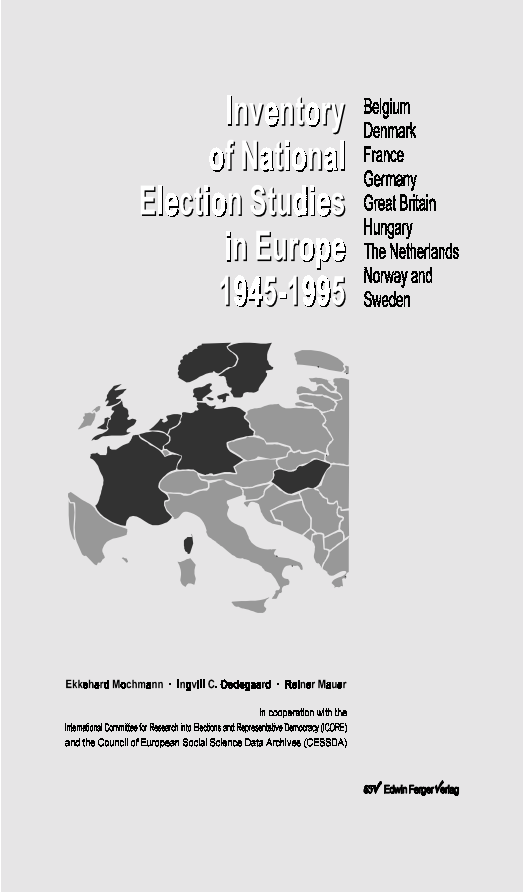
Inventory of National Election Studies in Europe 1945-1995:

Belgium, Denmark, France, Germany, Great Britain, Hungary, The Netherlands, Norway and Sweden.

by ***Ekkehard Mochmann***
Ingvill C. Oedegaard
Reiner Mauer

in co-operation with the
International Committee for Research into Elections and Representative Democracy (ICORE) and the Council of European Social Science Data Archives (CESSDA)

1998 / 422 pages / hardcover / 29 x 21 cm / DM 128.00
ISBN 3-931219-04-6 / Bergisch Gladbach: Edwin Ferger Verlag



This Inventory of National Election Studies in Europe documents election dates and election studies conducted in European countries with well established election research programmes, that co-operate in the International Committee for Research into Elections and Representative Democracy (ICORE). The time period covered extends from the Second World War till 1995.

The purpose of this inventory is to inform about electoral data suitable for longitudinal and/or cross-national election research in Europe. Thus for each country a list of election dates, an overview table containing detailed information about the election study, study descriptions and a list of literature is provided.

Mail to the publisher:
Edwin Ferger Verlag
Eichenhainallee 18 D - 51427 Bergisch Gladbach Germany
Fax: +49-2204-22203
Email: 0220465019@t-online.de

Mail-order form

From **:** _____

_____

_____

_____

Edwin Ferger Verlag
Eichenhainallee 18
51427 Bergisch Gladbach
Germany

Send by mail
or
Fax +49 2204 - 22203

Please send me/us:

No. of copies:

_____ **Inventory of National Election Studies in Europe 1945-1995:**
Belgium, Denmark, France, Germany, Great Britain, Hungary,
The Netherlands, Norway and Sweden.
by ***Ekkehard Mochmann, Ingvill C. Oedegaard, Reiner Mauer***

1998 / 422 pages / hardcover / 29 x 21 cm / DM 128.00
ISBN 3-931219-04-6 / Bergisch Gladbach: Edwin Ferger Verlag

Please add DM 10.00 (int'l postage) for all orders from outside Germany.

Payment options:

- ☐ I have transferred DM_____ to your Giro Account No. 529750005 at Koelner Bank von 1867, Koeln, Germany, (BLZ 37160087)
- ☐ I enclose a cheque for DM _____ made payable to Edwin Ferger Verlag. Eurocheques in DM accepted.
- ☐ Please charge to my EUROCARD / MasterCard: Card No. _____

Card expiry date: _____

Card holder's name: _____

Signature: _____

Jürgen Bortz und *Gustav Adolf Lienert*: Kurzgefaßte Statistik für die klinische Forschung – Ein praktischer Leitfaden für die Analyse kleiner Stichproben. Berlin: Springer, 1998, ISBN 3–540–63738–9, 11 Abb., 90 Tab., 405 Seiten, DM 68,00

Ein Buchhinweis von *Erwin Lautsch*[1]

Obwohl der Titel zunächst für Gesellschaftswissenschaftler fremdartig erscheint, sollte sich der methodisch interessierte Leser der angenehmen Mühe unterziehen, den Text ausführlich zu studieren. Hat man sich einmal mit der Gedankenwelt des „Klinikers“ vertraut gemacht, wird man zu der Feststellung kommen, daß in den Gesellschaftswissenschaften ähnliche Untersuchungsdesigns vorstellbar wären. Insbesondere sollten sich diesbezüglich außer Soziologen und Politikwissenschaftler u.a. auch Psychologen, Erziehungswissenschaftler, Sportwissenschaftler, Marktforscher und in der Evalutionsforschung Tätige angesprochen fühlen. Die vielen praktischen Anwendungen provozieren den Leser über die Möglichkeit nachzudenken, bei eigenen Forschungsansätzen auch ungewohnte methodische Wege zumindest in Erwägung zu ziehen. So könnte man sich vorstellen, daß die eine oder andere Problemstellung, die sonst immer im Verbund einer umfangreichen repräsentativen Bevölkerungsumfrage eingebettet ist, ausgelagert und mittels einem hier vergleichbaren Forschungsdesign empirisch angegangen wird. Als Beispiel könnte hierfür folgende Problemstellung stehen: Mehrere Experten sollen anhand mehrerer vorgegebener Kriterien beurteilen, ob Tageszeitungen durch ihre Kriminalitätsberichterstattung zu einem erhöhten Beunruhigungsgrad in der Bevölkerung beitragen. Auf den Seiten 270–276 findet der Leser ein adäquates Lehrstück aus der Medizin. Oder sollen zwei vermeintlich leistungssteigernde Trainingsmethoden hinsichtlich ihrer Wirksamkeit bei Sportlern verglichen werden, so wäre der auf den Seiten 136–138 beschriebene Untersuchungsplan angezeigt.

Der Leser, der gewohnt ist, mit mittleren und großen Stichprobenumfängen arbeiten zu wollen bzw. zu müssen, findet zahlreiche originelle Anregungen zur Durchführung von Datenanalysen, die ihn wegführen von sogenannten (nicht selten als bedenklich einzuschätzenden) Standardauswertungsroutinen allseits beliebter Statistiksoftware–Pakete. Die Stärke dieses Buches liegt aber gerade darin, daß auch für extrem kleine Stichproben Auswertungsmöglichkeiten angeboten werden, die mehr oder weniger selten Gegenstand der gängigen Fachliteratur sind und schon gar nicht als Software zur Verfügung stehen.

Unbedingt hervorzuheben sind die vielen Möglichkeiten der Datenauswertung, die auf Kontingenztafeln fußen. Diese können entweder das unmittelbare Produkt der Datenerhebung oder mittelbar durch testorientierte Datenaufbereitung entstanden sein.

Mit der Herausgabe der „Kurzgefaßten Statistik“ entsprachen die Autoren einem Wunsch vieler Anwender nonparametrischer Methoden, ein überschaubares und handliches

---

1 Dr. Dr. ***Erwin Lautsch*** ist Privatdozent am Fachbereich 5: Gesellschaftswissenschaften an der Universität Gesamthochschule Kassel, Nora-Platiel-Straße 1, 34127 Kassel.

Arbeitsinstrument Forschenden, Lehrenden und Studierenden zur Verfügung zu stellen, da die vorausgegangene sehr umfangreiche dreibändige Monographie, ***Lienert, G.A.***: Verteilungsfreie Methoden in der Biostatistik, Band 1 (1962[1], 1973[2], 1986[3]), Tafelband (1975) und Band 2 (1978), Meisenheim am Glan: Hain, und das einbändige Lehrbuch, ***Bortz, J.***; ***Lienert, G.A.*** und ***Boehnke, K.***(1990): Verteilungsfreie Methoden in der Biostatistik, Heidelberg: Springer, nicht den erhofften breiten Zuspruch gefunden haben. Das beabsichtigte Anliegen – einen griffigen Leitfaden für Anwender vorzulegen – kann als geglückt bezeichnet werden. Methodisch anspruchsvolle Leser jedoch, die über die Leistungsfähigkeit einzelner Methoden und über den mathematisch–statistischen Hintergrund umfassender informiert werden wollen bzw. nach weiteren Alternativen suchen, sollten auf die von Spezialisten hochgeschätzte Monographie und mindestens auf das Lehrbuch zurückgreifen.

Die kurzgefaßte verteilungsfreie Statistik wächst über den Rahmen eines reinen Statistikbuches dadurch hinaus, da die aufeinander abgestimmte Abhandlung der internen Kapitelgliederungspunkte [1. Zielsetzung (der Methode), 2. Durchführung, 3. Beispiel (–Demonstration bei strenger Einhaltung eines vorgegebenen Schrittfolgemodus) und 4. Anmerkungen, Hinweise etc.] für Anwender vertraute typische Forschungssituationen lebendig werden läßt. Auffallend gut gelungen ist die beabsichtigte Methodenplazierung insbesondere auch dadurch, daß in Anlehnung an häufig anzutreffende Forschungssituationen, „Ketten“ von Problemstellungen/Hypothesen abarbeiten zu müssen, die Wahl der Methoden(n) nicht streng nach formal–statistischen Gliederungskriterien vorgenommen wurde. Entweder geschah dies *innerhalb eines Kapitels* durch die hypothesengeleitete Reihung Globaltest und sogenannter Nachfolgetests (z.B. Allgemeine Kontingenztafelanalyse mit anschließender Trendauswertung und nachgeschobenen Teiltafeltests [S. 88–98]) oder *über mehrere Kapitel*, wenn den entsprechenden Methoden eine relative Eigenständigkeit aus dem Blickfeld des Gesamtkonzeptes des Buches zugebilligt worden ist (z.B. Kap. 2.4.3 und 2.4.4).

Obwohl keine besonderen mathematischen Kenntnisse vorausgesetzt werden, ist der Leser gut beraten, wenn er mit der mathematisch–statistischen Denkweise parametrischer und nonparametrischer Tests bereits grundsätzlich vertraut ist. Die Kenntnisse, die ein statistischer Grundkurs vermittelt, sind dafür eine gute Basis. Ohne Rückfragen bei Statistikern mit einschlägigen Erfahrungen wird es jedoch manchmal nicht weiter gehen.

Für weniger geübte Nutzer wird der Einstieg in die Lektüre durch ein Einführungskapitel (1) erleichtert. In den folgenden drei Kapiteln (2, 3 und 4) werden die Methoden in Abhängigkeit vom Meßtyp behandelt. Methoden in den Kapiteln 5 und 6 sind den Analysezielen: Zusammenhangsanalyse (verteilungsfreie Korrelationstests und –Maße) und Analyse von Urteilsübereinstimmungen gewidmet. Kapitel 7 bleibt Methoden vorbehalten, die dem Konzept der sequentiellen Tests folgen. In Kapitel 8 werden Abfolge– und Zeitreihentests angesprochen, die besondere Aufmerksamkeit verdienen. Mit Kapitel 9 wird ein Ausblick vermittelt. Eine im Anhang befindliche Übersichtstabelle erleichtert eine zielgerichtete Methodenauswahl. Hat man sich einmal an die Art und Weise der Betonung von Textblökken und –abschnitten in Form von Einrahmungen und schattiert eingefärbten Kästchen etc. gewöhnt, empfindet man dies als angenehm bei einem schnellen Orientierungsbedürfnis auf Beispiele, Zielsetzungen und/oder auf markante Anwendungshinweise.