

Materialien

15

Thomas Däubler

Nonresponseanalysen der Stichprobe F
des SOEP

Berlin, Mai 2002



DIW Berlin

Deutsches Institut
für Wirtschaftsforschung

Die in diesem Papier vertretenen Auffassungen liegen ausschließlich in der Verantwortung des Verfassers und nicht in der des Instituts.

DIW Berlin
Deutsches Institut
für Wirtschaftsforschung
Königin-Luise-Str. 5
14195 Berlin
Tel. 030 897 89-0
Fax 030 897 89-200
www.diw.de

ISSN 1619-4551

Nonresponseanalysen der Stichprobe F des SOEP

von Thomas Däubler*

1 Einleitung

In den letzten Jahren haben die Äußerungen über sinkende Ausschöpfungsquoten bei Bevölkerungsumfragen ohne Auskunftspflicht in Europa und Nordamerika stark zugenommen. Inzwischen mehren sich auch – trotz den methodischen Schwierigkeiten der Untersuchung des Problems – empirische Belege für diesen internationalen Trend (de Heer 1999, de Leeuw/de Heer 2002). Es kann allerdings kaum von einer einheitlichen Entwicklung gesprochen werden, vielmehr muss nach Ländern, Thematik der Surveys und unterschiedlichen Ausfallgründen differenziert werden.

Obwohl auch in Deutschland seit Jahren steigende Ausfallraten (Anders 1985) beklagt werden, konnte Schnell (1997) keine eindeutigen Hinweise auf einen Anstieg der Nonresponsequote feststellen. Ganz abgesehen davon ist nach wie vor zu betonen, dass eine Ausschöpfungsquote an sich keinesfalls ausreichend ist, um die Qualität einer Stichprobe zu beurteilen, da das Ausmaß einer Verzerrung sich nicht aus der Ausschöpfung allein, sondern nur als Produkt mit der Differenz der Merkmale zwischen Teilnehmern und Nichtteilnehmern ergibt (z.B. Groves 1989: 133). Es muss also versucht werden, diese Differenz zu bestimmen, indem der für den Ausfall verantwortliche Prozess bestmöglich modelliert wird.

Die Innovationsstichprobe des Sozio-oekonomischen Panels (SOEP), die sogenannte „Stichprobe F“, bietet eine sehr gute Gelegenheit, Ausfallprozesse bei einem Survey von Privathaushalten in Deutschland zu untersuchen, da diese Stichprobe mit ungewöhnlich großem Aufwand erhoben und dokumentiert wurde.

* Diese Arbeit entstand während eines Praktikums von Thomas Däubler (Universität Konstanz) am DIW Berlin, Abt. Längsschnittdaten und Mikroanalyse, 14191 Berlin. Email: thomas.daeubler@uni-konstanz.de. Der Autor dankt insbesondere Rainer Pischner (DIW Berlin), Thorsten Schneider (DIW Berlin), Jürgen Schupp (DIW Berlin), Jörg-Peter Schräpler (Ruhr-Universität Bochum und DIW Berlin) sowie Gert G. Wagner (Technische Universität Berlin und DIW Berlin) für hilfreiche Ratschläge und Anregungen. Sein spezieller Dank gilt Don Hedeker (University of Illinois, Chicago) sowie Min Yang (Multilevel Models Project, University of London) für wichtige Hinweise zur Schätzung der verwendeten Mehrebenenmodelle.

Dieses Papier beginnt mit einem kurzen Überblick auf theoretische Erklärungen und empirische Ergebnisse der Nonresponseforschung. Nach einer Beschreibung zum Verlauf der Feldphase der ersten Welle der SOEP-Stichprobe F erfolgen einige deskriptive Auswertungen zur Zusammensetzung von Interviewerstab und ausgefallenen Zielhaushalten. Anschließend wird das Teilnahmeverhalten analysiert, indem die zugrundeliegenden Prozesse (Erreichbarkeit, Befragungsfähigkeit und Kooperationsbereitschaft) auf Basis von Umgebungs- und Interviewermerkmalen modelliert werden. Zu diesem Zweck werden Mehrebenenmodelle verwendet. Außerdem widmet sich ein eigener Abschnitt der Frage, ob es einen Zusammenhang zwischen Erreichbarkeit und inhaltlichen Variablen gibt bzw. wie sich die Zahl der Kontakte auf die Stichprobenstruktur auswirkt.

2 Theorien und Erklärungsfaktoren von Nonresponse

Auf den ersten Blick scheint es eine Vielzahl an verschiedenen Theorien zur Erklärung von Teilnahmeverhalten bei Umfragen zu geben, die das Problem aus je nach Ursprungsdisziplin unterschiedlichem Blickwinkel betrachten. Bei genauerem Hinsehen gilt aber noch heute die Feststellung, dass alle Theorien letzten Endes auf „Kosten-Nutzen-Kalkülen“ rational handelnder Akteure basieren, „wenn dies auch nicht explizit so genannt wird“ (Esser et al. 1989: 120). Beispielsweise stellt auch einer der neuesten Ansätze, die sogenannte „leverage-saliency-theory“ von Groves et al. (2000) nichts anderes als einen Spezialfall einer allgemeinen soziologischen Handlungstheorie dar. Danach nehmen Zielpersonen dann an einem Survey teil, wenn die Teilnahme gegenüber der möglichen Handlungsalternative Nichtteilnahme den zu erwartenden Nutzen vergrößert. Nach Groves et al. (2000) haben bestimmte Merkmale (z.B. Thema, Dauer usw.) eines Surveys unterschiedliche „leverage“ (wörtlich „Hebelkraft“) bei verschiedenen Personen, d.h. sie werden unterschiedlich bewertet. Die Aktivierung der „leverage“ hängt nun wesentlich davon ab, inwieweit das jeweilige Merkmal in der Interaktion zwischen Interviewer und Zielperson durch situationale Faktoren „salient“ gemacht wird, also – wörtlich übersetzt – hervorsteht. Der Kern der Theorie besteht also aus nichts anderem als einem Kosten-Nutzen-Kalkül.

Das Problem liegt nun darin, die Nutzenfunktionen der Zielpersonen zu operationalisieren und zumindest annähernd zu bestimmen. Da diese individuell verschieden sind und gerade bei low-cost-Entscheidungen stark von situationalen Faktoren abhängig sein dürften (Schnell 1997), ist eine Erklärung durch „Globalvariablen“ problematisch. Da aber

zur empirischen Untersuchung des Teilnahmeprozesses in der Regel keine anderen Informationen vorliegen, bleibt nichts anderes übrig, als auf soziodemographische Indikatoren zur impliziten Modellierung der „Nutzenfunktion“ zurückzugreifen. Die Indikatoren halten Groves/Couper (1998: 120) für „nützlich“, solange sie innerhalb der jeweiligen Gruppen auf „shared life experiences“ bezogen sind.

Deshalb sollen zunächst einige Hypothesen zum Teilnahmeverhalten, die in der vorliegenden Studie Verwendung finden, rekapituliert werden. Die Reihenfolge entspricht dabei der mittlerweile standardmäßigen Dreiteilung des Teilnahmeprozesses in Erreichbarkeit, Befragungsfähigkeit und Kooperationsbereitschaft (Brehm 1993: 39f.).

2.1 Erreichbarkeit

Erreichbarkeit eines Zielhaushalts kann als Funktion aus der physischen Zugänglichkeit des Haushaltes (z.B. verschlossene Eingangstore), dem Tagesrhythmus der Haushaltsmitglieder und der Kontaktstrategie des Interviewers aufgefasst werden (Groves/Couper 1998: 81). Daneben können zum vierten Gründe für eine verdeckte Nichterreichbarkeit vorliegen, wenn der Befragte die Tür nicht öffnet, obwohl er anwesend ist, etwa aus Viktimisierungängsten oder um das Interview nicht explizit ablehnen zu wollen. Für sämtliche Faktoren liegen häufig keine Informationen vor, so dass in empirischen Analysen demographische Indikatoren verwendet werden müssen.

Sehr häufig ist ein Zusammenhang zwischen Erreichbarkeit und Stadtentfernung zu beobachten. So zeigt sich, dass Zielpersonen in Großstädten schlechter erreichbar sind.¹ Groves/Couper (1998: 81 f.) geben folgende Erklärungen für das Stadt-Land-Gefälle an: Größere Viktimisierungängste führen zu erschwerter physischer Zugänglichkeit der Häuser und erhöhen die Wahrscheinlichkeit, dass die Tür trotz Anwesenheit nicht geöffnet wird. Eine vermeintlich oder tatsächlich erhöhte Kriminalitätsbelastung kann auch Interviewer davon abhalten, öfter bzw. auch in den Abendstunden diese Gebiete aufzusuchen. Das größere Angebot an Freizeitmöglichkeiten in Städten dürfte ebenso eine Rolle spielen, da die Stadt-Land-Unterschiede bei Groves/Couper (1998: 112) nicht allein auf die größere Zahl an Einpersonenhaushalten zurückzuführen waren. Ähnliche Gründe dürften für die geringere Anzahl an erreichten Haushalten in Häusern mit mehreren Wohnungen verantwortlich sein.

¹ Goyder 1987: 84, Koch 1997: 115, Stoop/Louwen 1999: 9, Schräpler 2000: 131, für einen CATI-Survey Japac/Lundqvist 1999: 3.

Dieses Ergebnis findet sich beispielsweise bei Stoop/Louwen (1999: 9) und Schröppler (2000: 131).

Eine weitere Gruppe, die als schwerer erreichbar gilt, sind Personen mit höherem sozio-ökonomischem Status. Goyder (1987: 84) führt ältere Studien mit diesem Ergebnis an und bestätigt dies auch in eigenen Berechnungen (1987: 108). Hierbei muss aber beachtet werden, dass das Konzept des sozio-ökonomischen Status mehrdimensional ist und sehr unterschiedlich operationalisiert werden kann (vgl. Goyder et al. 1999: 7). Relativ eindeutige empirische Ergebnisse liegen zu Alterseffekten vor. So fanden Koch (1997: 117), Stoop/Louwen (1999: 9) und Schröppler (2000: 131), dass jüngere Personen leichter erreichbar sind als ältere, während Blohm/Diehl (2001: 237) für die deutsche Subpopulation des ALLBUS 96 keine signifikanten Effekte nachweisen konnten. Nicht-Deutsche waren in der Startwelle des SOEP im Jahre 1984 trotz höherer Gesamtausschöpfung bei bivariatem Vergleich schwerer erreichbar als Deutsche (Pannenberget al. 2000: 130). Dieses Ergebnis bestätigen Blohm/Diehl (2001: 231) für den ALLBUS 96 mit ebenfalls bivariaten Berechnungen. Die Resultate könnten allerdings außer auf längere Arbeitszeiten und Heimaturlaube auch auf Alters- und Stadt-Land-Effekte zurückzuführen sein (Blohm/Diehl 2001: 231), was die von Koch (1997: 117) berichtete Instabilität in multivariaten Modellen erklären würde.

Eine entscheidende Rolle bei der Frage, ob eine Zieladresse erreicht werden kann, spielt das Kontaktverhalten des Interviewers, der entscheidet wie oft, zu welchen Zeitpunkten und in welchen Zeitintervallen er Kontaktversuche unternimmt (Purdon et al. 1999: 200). Diese Entscheidungen werden von mehreren Faktoren beeinflusst. Groves/Couper (1998: 274) erwähnen, dass die Arbeitsbelastung des Interviewers den Aufwand für den einzelnen Haushalt reduzieren kann. Lievesley (1986, zitiert nach Morton-Williams 1993: 56) berichtet, dass Interviewer, die neben dem Interviewerjob eine Vollzeittätigkeit ausübten, besonders hohe Kontaktraten vorweisen konnten. Vermutlich handelt es sich um ein Resultat des besseren „timings“ auf Grund der ähnlich strukturierten Tagesabläufe von Interviewer und Zielpersonen. Außerdem gibt es Hinweise, dass Interviewer mit der Zeit lernen, ihr Kontaktverhalten zu optimieren, so dass erfahrenere Interviewer höhere Kontaktquoten erreichen (Morton-Williams 1993: 27). Groves/Couper (1998: 95) konnten dies in ihren „Census-Match“-Studien, in denen zur Analyse des Nonresponse bei US-amerikanischen Haushaltssurveys auf Registerdaten zurückgegriffen wurde, allerdings nicht bestätigen.

2.1.1 Zusammenhang zwischen Erreichbarkeit und inhaltlichen Variablen

Ein anderer Gegenstand von Interesse im Zusammenhang mit der Erreichbarkeit von Zieladressen ist die Frage, ob sich schwer Erreichbare sowohl in den oben erwähnten rein soziodemographischen Variablen als auch in den jeweiligen Variablen von „inhaltlichem“ Interesse, die gewöhnlich als abhängige Variablen analysiert werden, unterscheiden.

Es gibt zahlreiche Studien, die darauf aufbauen, die Fälle einer in einem Survey erhaltenen Stichprobe in Gruppen nach der Erreichbarkeit oder der „Schwierigkeit“ der Realisierung aufzuteilen (einen Überblick bietet Kennickell 2000: 2-4). Auch wenn sämtliche Versuche, daraus Korrekturverfahren für Nichterreichbarkeit bzw. Nonresponse zu entwickeln, auf Grund mehrerer Probleme als gescheitert angesehen werden können (Schnell 1997: 146-152), bleibt dies ein untersuchenswertes Phänomen. Unter anderem kann abgeschätzt werden, ob der Kontaktierungsaufwand ausreichend war.

Kennickell (2000) analysiert Unterschiede zwischen Früh- und Spätantwortern im Survey of Consumer Finances 1998 und stellt leichte Differenzen für einige ökonomische Variablen (meist Einkommenswerte) fest. Ebenfalls leichte Unterschiede zwischen Früh- und Späterreichten konnten Curtin et al. (2000) für einen Telefonsurvey nachweisen. In ihrer Studie zeigten sich die Früh-Erreichten als pessimistischer hinsichtlich der persönlichen und allgemeinen wirtschaftlichen Entwicklung im „Index of Consumer Sentiment“, was Curtin et al. (2000: 420) wesentlich darauf zurückführen, dass die Späterreichten jünger und wohlhabender waren. Keeter et al. (2000) vergleichen zwei Telefonsurveys, die mit unterschiedlichem Aufwand (u.a. hinsichtlich der Kontaktzahl) erhoben wurden. Neben einigen nach den obengenannten Hypothesen zu erwartenden soziodemographischen Differenzen konnten sie nur wenige, eher schwache Abweichungen auf „weichen“ Variablen entdecken. Die Generalisierbarkeit der Ergebnisse ist allerdings beschränkt, da es sich um Fallstudien handelt.

Zusammenfassend lässt sich sagen, dass Unterschiede zwischen Früh- und Späterreichten auf demographischen Variablen wahrscheinlich und bei Einstellungs- und Verhaltensfragen möglich sind, die Differenzen sowie mögliche Verzerrungen der Befunde sich aber in engen Grenzen halten dürften.

2.2 Befragungsfähigkeit

Wird ein Befragungshaushalt erreicht, so muss als nächster Schritt die Befragungsfähigkeit der Zielpersonen gewährleistet sein. Diese hängt vor allem von der sprachlichen Verständigung und der Gesundheit ab.

Um zu vermeiden, dass ausschließlich assimilierte Ausländer an der Befragung teilnehmen, kommen im SOEP neben dem Standardfragebogen auch einige fremdsprachige Versionen bei eventuellen Verständnisproblemen während des Interviews zum Einsatz. Fehlende Sprachkenntnisse spielen dadurch für die Befragungsfähigkeit fast keine Rolle. Vielmehr sind gesundheitliche Probleme dafür verantwortlich, dass Personen nicht zu einem Interview in der Lage sind.² Gesundheitliche Probleme nehmen mit steigendem Alter der Zielpersonen in der Regel zu. Brehm (1993: 49) modelliert den Prozess der Befragungsfähigkeit allein mit Altersvariablen und bestätigt den erwarteten Effekt. Auf dasselbe Ergebnis kommt auch Koch (1997: 117).

2.3 Kooperationsbereitschaft

Die letzte und besonders hohe Hürde auf dem Weg zur Realisierung eines Haushalts bildet das Problem der Kooperationsbereitschaft.

Empirisch gut belegt sind auch hier Unterschiede zwischen Stadt und Land. So berichten unter anderem Koch (1997: 118) und Schräpler (2000: 133), dass Zielpersonen in der Großstadt eher verweigerten. Ihre gleichlautenden Ergebnisse begründen Groves/Couper (1998: 176f.) damit, dass neben Viktimisierungsängsten die große Bevölkerungsdichte („crowding“) dazu führt, den Kontakt mit Fremden zu meiden und nicht auch noch die Privatsphäre in der eigenen Wohnung aufzugeben.

Bei Schräpler (2000: 133) liest man auch über geringere Kooperationsbereitschaft in Gebäuden mit mehreren Wohnungen. Auch diesen Effekt bestätigen Groves/Couper (1998: 139f.). Letztere interpretieren dies dahingehend, dass solche Haustypen als Indikatoren für Wohnumgebungen mit geringerer sozialer Integration anzusehen sind, welche sich wiederum negativ auf die Teilnahmebereitschaft auswirkt. Groves/Couper weisen allerdings nach, dass

² Vgl. in diesem Zusammenhang auch die Panelmortalitätsanalysen zum SOEP von Heller/Schnell (2000).

die Zusammenhänge zwischen Wohnumfeld und Kooperationsbereitschaft bei Kontrolle von soziodemographischen Variablen schwächer werden bzw. verschwinden.

Eine der gängigsten Hypothesen der Nonresponseforschung stellt wohl die Überrepräsentierung mittlerer sozialer Schichten („Mittelstandsbias“) in Bevölkerungsumfragen dar. Das tatsächliche empirische Fundament für diese Behauptung sowohl älterer als auch neuerer Studien ist aber ausgesprochen schwach und inkonsistent (Groves/Couper 1998: 126f.). Schnell (1997: 202f.) führt einen potentiellen Mittelstandsbias allein auf Bildungseffekte zurück, da Personen mit höherem Bildungsabschluss Befragungsthemen für bedeutsam halten und die Konsequenzlosigkeit der Teilnahme besser abschätzen könnten. Hinweise, dass dies zutrifft, finden sich bei Koch (1998: 75-77). Goyder et al. (1999) weisen darauf hin, dass zur Messung des sozio-ökonomischen Status multiple Indikatoren nötig sind und auch Drittvariablen – wie etwa Sponsorship-Effekte – den Zusammenhang zwischen sozialer Schicht und Teilnahmebereitschaft beeinflussen können.

Als essentiell wichtigem Faktor für die Teilnahmeentscheidung wurde in den letzten Jahren der Interaktion zwischen Interviewer und Befragtem („doorstep interaction“) in der Literatur große Beachtung geschenkt. Dabei lassen sich zwei zentrale Konzepte unterscheiden: „tailoring“ und „maintainig interaction“ (Morton-Williams 1993, Snijkers et al. 1999, Hox/de Leeuw 2002). Unter „tailoring“ versteht man den individuellen Zuschnitt der Einleitungssätze auf den jeweiligen Haushalt, wobei der Interviewer versucht, auf Basis aller ihm zur Verfügung stehenden Informationen (dies werden meist nur Eindrücke von der Umgebung und dem Wohnquartier sein)³ die erfolgversprechendste Äußerung seines Anliegens zu wählen. „Maintainig interaction“ bedeutet nichts anderes, als dass der Interviewer den Kontakt zur Zielperson nicht abbrechen lässt und von ihr geäußerte Bedenken ausräumen kann.

Um diese Konzepte erfolgreich anwenden zu können, benötigen die Interviewer fachliche und lokale Kenntnisse sowie Hintergrundwissen über den Lebensstil der Zielpopulation (Groves/Couper 1998: 201).

³ Bei Registerstichproben sind vorab Geschlecht und Alter der Zielperson bekannt.

Vor diesem Hintergrund wird auch plausibel, wieso Interviewermerkmale – wie die Literatur zeigt⁴ – nur bedingt Einfluss auf die Kooperationsbereitschaft haben. Nur wenn diese Variablen nötige Kenntnisse oder in sonst einer Form verbessertes tailoring indizieren, ist ein Zusammenhang mit höheren Kooperationsquoten zu erwarten. Sofern die notwendigen Kenntnisse „on the job“ erworben werden können, sollten erfahrene Interviewer mehr Kontaktierte für eine Teilnahme gewinnen können (Morton-Williams 1993: 27). Allerdings ist kein monotoner Zusammenhang zu erwarten, vielmehr müsste der Kenntniszuwachs mit der Zeit abnehmen. Außerdem können durch lange zurückliegende Schulungen und Motivationsverluste eventuell gegenläufige Effekte auftreten (Groves/Couper 1998: 203f.).

Berichte, dass nebenerwerbstätige Interviewer mehr Zielpersonen für eine Teilnahme gewinnen (Schräpler 2000: 134), können dahin interpretiert werden, dass sie in ihrem Berufsleben soziale Kompetenzen erworben haben. Andererseits treten möglicherweise durch die Mehrfachbelastung, der nebenerwerbstätige Interviewer ausgesetzt sind, negative Effekte auf, wie Koch (1991) belegt.

Weniger im Zusammenhang mit der Theorie der „doorstep interaction“ stehen die Hypothesen zum Einfluss des Geschlechts und Alters der Interviewer. So wird angenommen, dass Frauen und ältere Interviewer mehr Haushalte zu einer Teilnahme überreden können, da sie auf Zielpersonen weniger bedrohlich wirken. Morton-Williams (1993: 104-109) erörtert in diesem Zusammenhang auch, dass Verständnis für die Situation des Interviewers und der Wunsch zu helfen in die Kosten-Nutzen-Erwägungen einfließen. Empirische Belege für höhere Ausschöpfungsquoten (ohne Unterscheidung nach Ausfallgründen) von Frauen finden sich bei Koch (1991). Schräpler (2000: 134) konnte hingegen im SOEP keine Haupteffekte des Geschlechts auf die Kooperationsbereitschaft feststellen, der Einfluss des Alters war nur sehr schwach.

Da Hypothesen über den Einfluss von soziodemographischen Merkmalen der Zielpersonen für das SOEP nur äußerst rudimentär geprüft werden können, wird auf diese nicht näher eingegangen. Für näher Interessierte sei auf die entsprechenden Erörterungen bei Schnell (1997) und Groves/Couper (1998) zu Haushaltsgröße, Geschlecht und Alter sowie bei Blohm/Diehl (2001) zu Nationalität verwiesen.

⁴ vgl. Schnell (1997: 195) und Couper/Groves (1998: 191)

3 Die Erhebung der Stichprobe F des SOEP

3.1 Feldverlauf

Im Jahre 2000 wurde mit der Teilstichprobe F des SOEP angestrebt, 6000 neue Haushalte in das Panel aufzunehmen. Als Grundgesamtheit wurde die Menge aller zum Zeitpunkt der Befragung in Privat-Haushalten lebenden Personen ab 16 Jahren in der BRD definiert. Die Ziehung der Haushalte erfolgte über eine mehrfach geschichtete, mehrstufige Zufallsstichprobe auf Grundlage des ADM-Stichprobensystems. Als Schichtungsmerkmale wurden Bundesländer, Regierungsbezirke, Kreise/kreisfreie Städte und BIK-Gemeindetypen verwendet (Rosenblatt 2001a). Auf der letzten Stufe wurden die Haushalte durch das Random-Route-Verfahren mit Adressenvorlauf, auch als „address-random“ bezeichnet (Kirschner 1984:132, zitiert nach Schnell et al. 1999: 267), ausgewählt.

Im Gegensatz zum weitverbreiteten „standard-random“-Verfahren, bei dem Adressermittlung und Interview simultan stattfinden, werden die zwei Schritte hier unabhängig voneinander ausgeführt. Die Adressen werden nach ihrer Ermittlung dem Befragungsinstitut übergeben, der Interviewer erhält seine Zieladressen dann vom Institut selbst. Dieses Vorgehensweise ist weitaus sicherer vor durch Interviewermanipulationen verursachten Verzerrungen als das „standard-random“-Verfahren.

Tabelle 1 gibt einen Überblick über das Endergebnis der Feldarbeit. Unter den insgesamt 12504 Ausgangsadressen waren 625 als qualitätsneutral anzusehende Ausfälle. Meist existierte kein (Privat-)Haushalt an der zu bearbeitenden Adresse. Bezogen auf das sog. bereinigte Brutto konnten schließlich 51 % der Haushalte für ein Interview gewonnen werden, was im Vergleich mit anderen deutschen Surveys⁵ keineswegs mehr als besonders geringe Ausschöpfung angesehen werden kann. Da das SOEP sowohl Haushalts- als auch Personenmerkmale erhebt, wird jede Person innerhalb eines Haushaltes interviewt. Sind nicht alle Personen eines Haushaltes bereit teilzunehmen, wurde aber mindestens der

⁵ Vgl. Wasmer et al. (1996), Koch (1997), Koch et al. (1998a), Koch (1998b), Koch et al. (2001). Details der Feldarbeit werden in Deutschland kaum öffentlich zugänglich gemacht. So können Ausmaße unterschiedlicher Ausfallgründe nur mit den ALLBUS-Studien verglichen werden. Diese Vergleiche sind jedoch nicht unproblematisch, da Unterschiede im Erhebungsdesign (z. B. Verwendung von Einwohnermeldeamt-Stichproben beim ALLBUS 94, ALLBUS 96, ALLBUS 2000 bzw. einer ADM-Stichprobe mit standard-random-Verfahren beim ALLBUS 98) für die interessierenden Prozesse eine große Rolle spielen.

Haushaltsfragebogen beantwortet, gilt der Haushalt als teilrealisiert. Der Anteil der Teilrealisierungen an allen realisierten Haushalten beträgt 9 %.

Von besonderer Bedeutung ist die Betrachtung der systematischen Ausfälle. Die kleinste Gruppe (3 %) unter den systematischen Ausfällen konstituieren die Befragungsunfähigen.

Tabelle 1: Ergebnis der Feldarbeit

	Anzahl			
Brutto	12504 ¹		12504	
- qualitätsneutraler Ausfall	625		824	

Bereinigtes Brutto	11879	100 %	11663	100 %
<i>abzüglich</i>				
Kein Feldrücklauf	501	4 %	285	
Nicht erreicht	910	8 %		
Nicht befragungsfähig	351	3 %		
Verweigerungen	4057	34 %		

- Summe systemat. Ausfall	5819	49 %		

Realisierte Haushalte	6060	51 %		53 %
In Welle 2 als Fälschungen	8	0,07 %		x
identifizierte Haushalte ²				
Endgültig realisierte Haushalte	6052	51 %		53 %

¹ einschließlich 423 Ausländer-Adressen aus dem verlängertem random walk, der notwendig war, um die Auswahlwahrscheinlichkeit von Personen mit nicht-deutscher Staatsangehörigkeit zu erhöhen.

² Vgl. von Rosenblatt 2001b, S. 9.

Quelle: SOEP (F) 2000, eigene Berechnungen.

Unter der Bezeichnung „kein Feldrücklauf“ sind alle Haushalte zu verstehen, über deren Feldbearbeitung dem durchführenden Institut⁶ keinerlei Informationen vorliegen. Das Ergebnis resultiert für 216 (43 %) der betroffenen 501 Haushalte aus einem Totalausfall des jeweiligen sampling points. Die restlichen 285 Fälle sind einzelne „schwarze Löcher“ in bearbeiteten sampling points. Eine detailliertere Betrachtung dieses Problems erfolgt weiter unten. Da man im ersten Fall nicht sicher und im zweiten Fall sogar ziemlich sicher nicht von der Zufälligkeit des Ausfalls ausgehen kann, werden alle als „kein Feldrücklauf“ klassifizierten Haushalte zu den systematischen Ausfällen gerechnet. Rechnet man hingegen die totalen Sample-Point-Ausfälle zu den qualitätsneutralen Ausfällen, steigt die Gesamtteilnahmequote auf 53 %.

⁶ Die Feldarbeit für das SOEP wird von Infratest Burke Sozialforschung (München) durchgeführt. Das Institut ist auch an der Konzeption des Surveys beteiligt.

Im Vergleich zu 3 % in der Startwelle des SOEP im Jahre 1984 (Schräpler 2000: 122) und zu den Werten in den oben bereits zitierten neueren deutschen Studien ist die Quote der Nichterreichten mit 8 % relativ hoch. Bedenkt man, dass für das SOEP große Anstrengungen in der Feldarbeit unternommen werden und die Interviewer angehalten sind, bis zu 10 Kontaktversuche zu unternehmen, ist diese Zahl bemerkenswert und eine nähere Untersuchung geboten. Der Anteil der eigentlichen Verweigerungen hingegen scheint vergleichsmäßig im Rahmen zu liegen.

3.2 Interviewerstab

Für die Erhebung der Stichprobe F des SOEP wurden zuletzt⁷ 618 Interviewer eingesetzt, wobei die 356 männlichen Personen mit 58 % leicht in der Überzahl waren.

Tabelle 2: Zusammensetzung des Interviewerstabes

<i>Variable</i>	<i>Kategorien</i>	<i>Anteil</i>
Geschlecht	Männlich	58 %
	Weiblich	42 %
Alter	Unter 30	2 %
	30-39	10 %
	40-49	20 %
	50-59	32 %
	60-69	28 %
	70 und älter	8 %
Familienstand	Ledig	11 %
	Verheiratet	73 %
	Geschieden/getrennt/ verwitwet	15 %
Ausbildung	Volksschule	22 %
	Realschule o.ä.	43 %
	Abitur	11 %
	Uni ohne Abschluss	4 %
	Uni mit Abschluss	20 %
Berufstätigkeit: Arbeitszeit	Nicht berufstätig	54 %
	Halbtags berufstätig	15 %
	Voll berufstätig	31 %
Berufstätigkeit	Hausfrau	17 %
	Rentner	28 %
	Student/Azubi/ arbeitslos	9 %
	Halbtags	15 %
	Schichtarbeit	6 %
	Berufstätig bis 16 Uhr	10 %
	Berufstätig bis 17 Uhr	4 %
	Berufstätig bis nach 17 Uhr	1 %
	Gleitzzeit	9 %
Beschäftigungsdauer im Institut	Neu	6 %
	1-3 Jahre	25 %

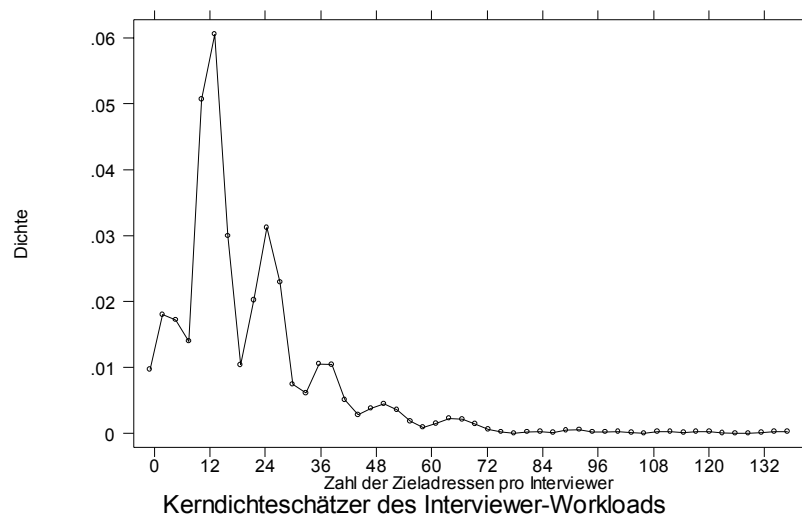
⁷ Alle Analysen, die Interviewereigenschaften berücksichtigen, verwenden Merkmale des Interviewers, der den Haushalt zuletzt bearbeitet hat. Der Datensatz enthält auch noch die Nummer des random-walk-Interviewers sowie des zuerst eingesetzten Interviewers. Zwar ist in 28 % der Fälle der Erstinterviewer nicht mit dem Letztinterviewer identisch, doch kann dessen Einfluss auf das Endergebnis auf Grund der Datenstruktur bislang leider nicht analysiert werden.

	4-7 Jahre	24 %
	8-14 Jahre	25 %
	15 Jahre und länger	20 %
Mitarbeit am SOEP	Neu	61 %
	1-3 Jahre	10 %
	4-6 Jahre	10 %
	7-16 Jahre	20 %
Workload: Zahl der Zieladressen	Bis 12	43 %
	13-24	28 %
	25-36	15 %
	37-48	6 %
	49-66	5 %
	67-135	2 %

Quelle: SOEP (F) 2000, eigene Berechnungen.

Eine Zusammenfassung der soziodemographischen Merkmale des Interviewerstabes bietet Tabelle 2. Die Altersspanne reicht von 18 bis 84 Jahren⁸, das Durchschnittsalter ist mit 54 Jahren verhältnismäßig hoch, da das Umfrageinstitut versucht, möglichst erfahrene Interviewer einzusetzen.

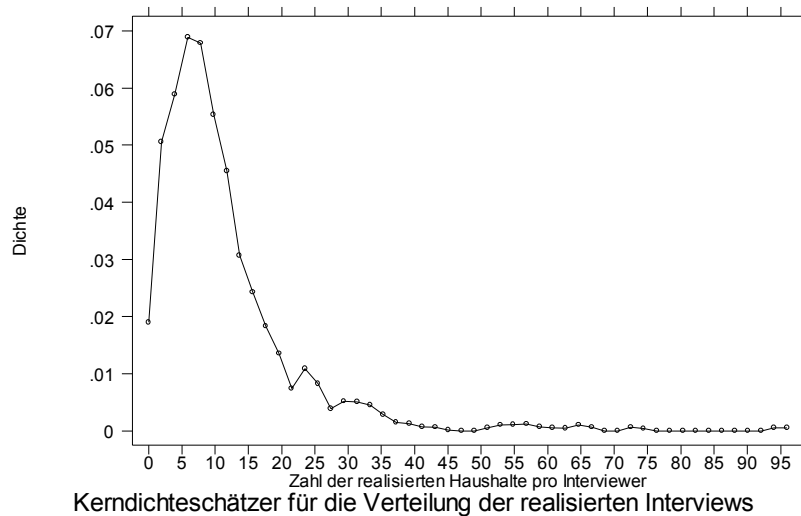
Abbildung 1: Kerndichteschätzer des Interviewer-Workloads



Fast drei Viertel der Interviewer sind verheiratet, 15 % geschieden, getrennt oder verwitwet, nur etwas mehr als ein Zehntel ist ledig. Über einen Volksschulabschluss verfügen 22 %, über einen Realschulabschluss 43 % und über Abitur oder einen höheren Bildungsabschluss 35 % der Interviewer. Die knappe Mehrheit übt keine weitere Erwerbstätigkeit aus, über 30 % sind jedoch neben der Arbeit als Interviewer voll berufstätig. Für 15 % stellt die Beschäftigung einen Nebenjob dar. Eine detailliertere Darstellung der Berufstätigkeit ist ebenfalls Tabelle 2 zu entnehmen. Erwähnenswert erscheint der hohe Anteil an Rentnern im Interviewerstab.

⁸ Der 84-jährige Interviewer ist noch voll einsatzfähig und verfügt über gute Kontakt- und Kooperationsquoten.

Abbildung 2: Kerndichteschätzer der realisierten Interviews pro Interviewer



Nur relativ wenige Personen verfügen über keinelei Arbeitserfahrung im betreffenden Institut, der Median liegt bei sieben Jahren. In Hinsicht auf bereits erfolgte Mitarbeit in vorangegangenen Wellen des SOEP zeigt sich folgendes Bild: Deutlich über die Hälfte der Interviewer ist zum ersten Mal innerhalb dieser Studie beschäftigt. Auch das liegt nahe, da die Aufnahme neuer Befragungspersonen, die noch keine Beziehung zu einem bestimmten Interviewer aufgebaut haben, sehr gut mit der Aufnahme neuer Interviewer in das SOEP verbunden werden kann.

Was den „Workload“ der einzelnen Interviewer angeht, gibt es große Unterschiede: Die Spannweite reicht von einer bis zu 135 zu bearbeitenden Adressen, der Median liegt bei 14 Adressen, unweit der Zahl von 12 pro sampling point gezogenen Adressen. Der durchschnittliche Interviewer erreicht 92 %⁹ seiner Zielhaushalte, wobei er 57 %¹⁰ der erfolgreich kontaktierten Haushalte für eine Teilnahme gewinnen kann. Der Median der realisierten Zielhaushalte pro Interviewer¹¹ beträgt neun, das Maximum ist bei 95 Interviews zu finden.

4 Interne Validierung

Nach der Beschreibung des Feldverlaufs bilden Vergleiche von Gruppen nach Befragungsstatus hinsichtlich soziodemographischer Variablen den nächsten Schritt der

⁹ Bereinigt um qualitätsneutrale Ausfälle und Haushalte mit Endergebnis „kein Feldrücklauf“

¹⁰ Bereinigt um nichtbefragungsfähige Haushalte

¹¹ Bezogen auf die 541 Interviewer, die mindestens einen Haushalt befragen konnten

Untersuchung. Die erste Tabelle stellt die Eigenschaften von Teilnehmern und Ausfällen der Stichprobe F den Charakteristika von denjenigen der Stichproben A und B des SOEP aus dem Jahr 1984 gegenüber. Die zweite Aufstellung gibt einen Überblick allein für Stichprobe F, detailliert aufgeschlüsselt nach Nicht-Erreichten, Nicht-Befragungsfähigen, Verweigerern und Teilnehmern.

Die Informationen über die Vergleichsmerkmale stammen aus verschiedenen Quellen: Stadtentfernung und soziale Schicht sind für alle Gruppen einem speziellen Adressprotokoll entnommen, d.h. sie beruhen auf Schätzungen des Interviewers, der die Adressen während des random walks aufgelistet hat („Quartiersbeschreibung“). Die anderen Variablen beruhen für Teilnehmer auf den Angaben im Personenfragebogen. Für Nichtteilnehmer waren die Interviewer angehalten, im Rahmen einer kleinen Nonresponse-Studie zumindest einige wenige Variablen bei den Zielpersonen selbst oder z.B. durch Proxy-Befragungen der Nachbarn zu ermitteln.

Der Anteil der Nonrespondenten, über die sämtlich erbetenen Informationen vorliegen, (jeweils in Klammern angegeben) ist aber teilweise sehr klein, ebenso kann die Validität der ermittelten Angaben z.B. bei Befragung der Nachbarn angezweifelt werden. Bedenkt man zusätzlich, dass selbst eine ähnliche Randverteilung auf solider Grundlage noch keinerlei Aussage zur Repräsentativität von anderen Variablen (sofern die anderen Variablen nicht hinsichtlich der angegebenen Variablen homogen sind, vgl. Schnell 1993) zulässt, sollten diese deskriptiven Darstellungen mit allergrößter Vorsicht betrachtet werden.

Was Personen mit deutscher Staatsangehörigkeit anbelangt, sind die Unterschiede zwischen Teilnehmern und Ausfällen hinsichtlich der Haushaltsgröße stärker, hinsichtlich Geschlecht und Alter im Vergleich zu der Startwelle 1984 geringer geworden. Der Vorzeichenwechsel hinsichtlich des Arbeiteranteils unter Teilnehmern und Ausfällen scheint bemerkenswert, könnte aber nur darauf zurückzuführen sein, dass die Datenlagen für die Stichproben A und F unterschiedlich sind: Für F liegen bislang¹² nur Berufsinformationen über Erwerbstätige vor.

¹² Die Berufsinformationen von Nicht-Erwerbstätigen werden im Biographiefragebogen erhoben, mit dem Befragte erst im zweiten Teilnahmejahr konfrontiert werden.

Tabelle 3: Merkmale der Teilnehmer/Ausfälle in F im Vergleich mit den Stichproben A (1984 Westdeutsche) und B (1984 Ausländer in der BRD)

	<i>Stichprobe A (1984)</i>		<i>Stichprobe F (2000) Deutsche</i>		<i>Stichprobe B (1984)</i>		<i>Stichprobe F (2000) Ausländer</i>	
	Realisiert	Ausfälle	Realisiert	Ausfälle	Realisiert	Ausfälle	Realisiert	Ausfälle
Fallzahl ges.	4554 (61%)	2906 (39%)	5708 ¹	3440 ¹	1415 (68%)	652 (32 %)	348 ¹	167 ¹
Haushaltsgröße								
Basis (Fallzahl)	4554	1728 (59 %)	5708	2193 (64 %)	1415	261 (40 %)	348	97 (58 %)
1 Person	26 %	30 %	28 %	37 %	15 %	15 %	18 %	27 %
2 Personen	32 %	36 %	36 %	42 %	17 %	18 %	22 %	33 %
3 Personen	19 %	16 %	16 %	11 %	20 %	26 %	19 %	13 %
4 Pers u. mehr	24 %	18 %	20 %	10 %	48 %	42 %	41 %	27 %
Geschlecht des HV ³								
Basis (Fallzahl)	4554	1799 (62 %)	5708	2813 (82 %)	1415	269 (41%)	348	140 (84 %)
Männlich	76 %	67 %	66 %	67 %	92 %	91 %	77 %	84 %
Weiblich	24 %	33 %	34 %	33 %	8 %	9 %	23 %	16 %
Alter des HV ³								
Basis (Fallzahl)	4554	1523 (52 %)	5706	2566 (74 %)	1415	213 (33 %)	348	126 (75 %)
Unter 29	14 %	9 %	8 %	7 %	12 %	9 %	20 %	13 %
30-49	40 %	30 %	42 %	34 %	63 %	62 %	49 %	55 %
50-69	30 %	34 %	35 %	34 %	25 %	29 %	28 %	30 %
70 u. mehr	16 %	27 %	16 %	25 %	0 %	-	3 %	2 %
Stellung des HV ³ im Beruf ²								
Basis (Fallzahl)	4494 (99 %)	894 (31 %)	3323 (58 %) ²	1045 (30 %)	.	.	221 (64 %) ²	67 (40 %)
Arbeiter	39 %	33 %	30 %	35 %	.	.	59 %	75 %
Angestellte	36 %	38 %	46 %	44 %	.	.	26 %	15 %
Beamte	11 %	10 %	9 %	4 %	.	.	1 %	-
Selbständige	11 %	13 %	13 %	16 %	.	.	13 %	9 %
Azubis/Praktikanten	1 %	1 %	1 %	2 %	.	.	1 %	1 %
Sonstige	4 %	6 %

¹ Hier können keine Ausschöpfungen angegeben werden, da die Nationalitätsangaben nur von einem Teil der Ausfälle bekannt sind und somit keine Brutto-Fallzahlen vorliegen.

² Berufsinformationen liegen für die Stichprobe F bei Teilnehmern nur für Erwerbstätige vor. Für die Stichproben A und B sind für Nicht-Erwerbstätige Informationen über den zuletzt ausgeübten Beruf verfügbar und berücksichtigt.

³ Die personenspezifischen Merkmale beziehen sich auf den Haushaltsvorstand. Bei einigen teilrealisierten Haushalten aus der Teilnehmergruppe liegen diese Daten für den Haushaltsvorstand nicht vor, da dieser kein Personeninterview gegeben hat. In diesen Fällen (je 63 für Nationalität, Geschlecht und Alter sowie 33 für Beruf) beziehen sich die Daten auf die Auskunftsperson.

Quelle: SOEP 2000, eigene Berechnungen.

Die Tendenzen für Nicht-Deutsche erscheinen nur teilweise analog zu den Differenzen für Deutsche zu verlaufen. Zwar gibt es Anzeichen, dass die Probleme, Einpersonenhaushalte für die Stichprobe zu gewinnen, ebenfalls im Zeitverlauf zugenommen zu haben. Dasselbe gilt im Gegensatz zur deutschen Population aber auch für jüngere Befragte. Tabelle 3 unterscheidet

allerdings nicht nach verschiedenen Ausfallgründen. Diese Dimension berücksichtigt die folgende Zusammenstellung.

Tabelle 4: Haushalts- und personenspezifische¹ Merkmale in der Bruttostichprobe nach Bearbeitungsergebnis

Merkmal	Kein Rücklauf	Nichterreichte	Nicht interviewbar	Verweigerer	Teilnehmer
	N = 501	N = 910	N = 351	N = 4057	N = 6060
Stadtentfernung					
Datenbasis: Adressprotokoll	100 %	100 %	100 %	100 %	100 %
Stadtmitte Großstadt	20 %	26 %	19 %	16 %	14 %
Stadttrand Großstadt	23 %	36 %	25 %	23 %	23 %
Unter 25 km	22 %	16 %	21 %	23 %	23 %
25 – 50 km	17 %	13 %	22 %	20 %	20 %
50 km u mehr	18 %	9 %	14 %	18 %	21 %
Soziale Schicht					
Datenbasis: Adressprotokoll	100 %	100 %	100 %	100 %	100 %
Unterschicht	2 %	5 %	7 %	6 %	4 %
Untere Mittelschicht	34 %	36 %	33 %	32 %	30 %
Mittlere Mittelschicht	55 %	47 %	52 %	53 %	56 %
Obere Mittelschicht	8 %	10 %	6 %	9 %	8 %
Oberschicht	0 %	1 %	1 %	1 %	1 %
Nationalität des HV¹					
Datenbasis	.	32 % (289)	87 % (304)	74 % (3012)	100 % (6056)
Deutsch	.	94 %	99 %	95 %	94 %
Nicht deutsch	.	6 %	1 %	5 %	6 %
Geschlecht des HV¹					
Datenbasis	.	20 % (185)	86 % (301)	52 % (3034)	100 % (6056)
Männlich	.	70 %	37 %	68 %	67 %
Weiblich	.	30 %	63 %	32 %	33 %
Haushaltsgröße					
Datenbasis	.	17 % (153)	79 % (278)	40 % (2347)	100 % (6060)
1 Person	.	58 %	71 %	37 %	28 %
2 Personen	.	31 %	26 %	41 %	35 %
3 Personen	.	8 %	2 %	11 %	16 %
4 und mehr Personen	.	3 %	1 %	10 %	22 %
Alter des HV¹					
Datenbasis	.	15 % (133)	83 % (291)	47 % (2739)	100 % (6056)
Unter 29	.	15 %	1 %	7 %	8 %
30-49	.	32 %	3 %	35 %	42 %
50-69	.	35 %	12 %	34 %	35 %
70 u. mehr	.	18 %	84 %	24 %	15 %
Stellung des HV¹ im Beruf²					
Datenbasis	.	7 % (66)	11 % (38)	19 % (1126)	58 % ² (3544)
Arbeiter	.	8 %	47 %	37 %	31 %
Angestellte	.	32 %	8 %	42 %	45 %
Beamte	.	5 %	37 %	3 %	9 %
Selbständige	.	50 %	8 %	15 %	13 %
Azubis/Praktikanten	.	6 %	-	2 %	1 %

¹ Die personenspezifischen Merkmale beziehen sich auf den Haushaltsvorstand. Bei einigen teilrealisierten Haushalten aus der Teilnehmergruppe liegen diese Daten für den Haushaltsvorstand nicht vor, da dieser kein Personeninterview gegeben hat. In diesen Fällen (je 63 für Nationalität, Geschlecht und Alter sowie 33 für Beruf) beziehen sich die Daten auf die Auskunftsperson.

² Berufsinformationen liegen für die Stichprobe F bei Teilnehmern nur für Erwerbstätige vor. (vgl. Fußnote 14)

Quelle: SOEP (F) 2000, eigene Berechnungen.

Mit der nötigen Vorsicht lässt sich folgendes zusammenfassen:

Haushalte ohne Feldrücklauf liegen eher in der Großstadt. Nicht befragungsfähig sind vor allem ältere Menschen. Einpersonenhaushalte sind sowohl schwerer erreichbar als auch schwerer für eine Teilnahme zu gewinnen. Während es in der Erreichbarkeit auch die erwarteten Unterschiede zwischen Stadt und Land gibt, gilt dies für die Kooperationsbereitschaft nicht.

Nur unter den Befragungsfähigen gibt es Abweichungen in den Nationalitätsanteilen, und zwar in der eher unerwarteten Richtung. Dass Deutsche weniger befragungsfähig sind, ist ein wichtiger Hinweis darauf, dass Sprachprobleme für die Erhebung des SOEP kein schwerwiegendes Problem darstellen. Der Einsatz verschiedensprachiger Fragebögen stellt also ein sehr effektives Mittel dar, um Ausfälle auf Grund mangelnder Deutschkenntnisse zu verringern.

Die Geschlechtsverteilung ist für Teilnehmer und Verweigerer nicht sehr verschieden, wohingegen die Alters- und Sozialstruktur Differenzen zeigen. Der Anteil der über 70-Jährigen ist unter den Verweigerern deutlich höher als unter Teilnehmern. Arbeiter scheinen eher zu verweigern, während Beamte überdurchschnittlich befragungswillig sind.

5 Detaillierte Analysen des Teilnahmeprozesses

5.1 Feldrücklauf

Im nächsten Schritt soll versucht werden, etwas Licht ins Dunkel des Ergebnisses „kein Feldrücklauf“ zu bringen. Wie oben bereits erwähnt, kann man diesbezüglich Ausfälle ganzer sampling points sowie Ausfälle einzelner Haushalte unterscheiden. Unter die erste Kategorie fallen 216 Zieladressen, also 18 sampling points. Über die betroffenen Zieladressen liegen per definitionem nur die Informationen aus dem Adressprotokoll sowie Angaben über die (nicht)bearbeitenden Interviewer vor.

Von den 18 sampling points liegen sechs in der Stadtmitte einer Großstadt, vier am Stadtrand, zwei im Umkreis bis zu 25 km, drei im Umkreis zwischen 25 und 50 km und drei mehr als 50 km von einer Großstadt entfernt. Bedenkt man, dass nur 16 % aller Zielhaushalte sich in der Stadtmitte einer Großstadt befinden, könnte dies als Hinweis auf einen

systematischen Prozess, wie etwa dass einige wenige Interviewer die Bearbeitung von Großstadtzentren „verweigert“ haben, gewertet werden. Da nicht jeder Interviewer in genau einem sampling point eingesetzt wurde, waren insgesamt 17 Interviewer für die Totalausfälle verantwortlich, d.h. ein Interviewer hat zwei Points nicht bearbeitet. Auffällig ist dabei der hohe Männeranteil (12 von 17 bzw. 71 % gegenüber 59 % unter allen Interviewern) sowie die deutlich geringere mittlere Beschäftigungsdauer bei Infratest (5 Jahre gegenüber 9 Jahren in Hinblick auf alle Interviewer). Außerdem hatten es 12 der 17 Interviewer zum ersten Mal mit dem SOEP zu tun. Es gibt aber durchaus auch Fälle mit überdurchschnittlich langer Beschäftigungsdauer im Institut. Neben gesundheitlichen Problemen der Interviewer sind vielfältige Ursachen (wie zeitliche Beanspruchung für andere Studien) für die sampling-point-Ausfälle denkbar, letzten Endes kann nur spekuliert werden.

Kaum mehr lässt sich über die in ansonsten bearbeiteten sampling points liegenden 285 Haushalte sagen, die als „kein Feldrücklauf“ klassifiziert sind. Nach einer detaillierten Betrachtung lässt sich diese Kategorie nochmals in zwei bzw. drei Subgruppen aufteilen. 45 der 285 betroffenen Zieladressen sind auf Totalausfälle von 14 Interviewern zurückzuführen. Diese Haushalte liegen allerdings in sampling points, in denen mehrere Interviewer gearbeitet haben, so dass nur ein kompletter Ausfall des Interviewers, nicht aber ein Totalausfall des sampling points resultiert. Vier der 14 (28 %) ausgefallenen Klumpen liegen in der Stadtmitte einer Großstadt, 10 der 14 Interviewer sind Männer (71 %). Ist hierin eine gewisse Analogie zu den Ergebnissen bei den Ausfällen der kompletten sampling points zu sehen, gilt dies für die Erfahrungswerte der Interviewer nicht. Sowohl die mittlere Beschäftigungsdauer im Institut als auch die Beteiligung im SOEP liegt über dem Durchschnitt aller Interviewer. Schlussfolgerungen sind hier also sehr unsicher, gerade auch in Anbetracht der geringen Fallzahlen.

Die restlichen 240 Fälle, deren Bearbeitungsstatus unbekannt ist, können alle Interviewern zugeordnet werden, die mindestens eine Zieladresse bearbeitet haben (und das Ergebnis bekannt ist). Anhand des Kriteriums, ob der jeweilige Interviewer zumindest eine Zieladresse *realisieren* konnte, ergeben sich nochmals zwei Untergruppen: Zum einen vier Interviewer ohne jegliches realisiertes Interview mit 14 Fällen ohne Feldrücklauf, zum anderen die übrigen 226 Zielhaushalte als „schwarze Löcher“ von ansonsten „unauffälligen“ 73 Interviewern. Was die erste Gruppe angeht ist bemerkenswert, dass alle vier Interviewer erst kurzzeitig für das Institut tätig sind (dreimal je ein Jahr, einmal zwei Jahre).

Nicht uninteressant ist aber die Frage, was in der zweiten Gruppe zu den Ausfällen führt. Gegenüber der Verteilung unter allen Zieladressen sind Haushalte im Großstadtzentrum in der „Problemgruppe“ unterrepräsentiert (9 % gegenüber 16 %), Haushalte in bis zu 25 km Entfernung vom Zentrum einer Großstadt überrepräsentiert (31 % gegenüber 22 %), während Zieleinheiten mit über 50 km Entfernung zur Großstadt wiederum einen kleineren Anteil (14 % gegenüber 19 %) als in der gesamten Stichprobe aufweisen. Im Vergleich zu den anderen Haustypen ist der Bearbeitungsstatus von Adressen in reinen Wohngebieten seltener ungeklärt (52 % gegenüber 59 %). Weitere nennenswerte Unterschiede in den Anteilen gibt es noch in Hinblick auf Mehrfamilienhäuser (24 % gegenüber 19 %), höhere Gebäude (30 % gegenüber 22 %) und obere Mittelschicht/Oberschicht (15 % gegenüber 9 %). Über die betreffenden 73 Interviewer, die ungeklärte Fälle aufweisen, lassen sich folgende Angaben machen: Die Zahl der Haushalte ohne Feldrücklauf pro Interviewer liegt im Mittel bei drei. Bei 30 Interviewern handelt es sich nur um ein ungeklärtes Interview, ein Interviewer bringt es aber auf 20 nicht bearbeitete Haushalte. Neben dem erhöhten Anteil an halbtags berufstätigen Interviewern (21 % gegenüber 15 %) erweist sich die Gruppe wider Erwarten als überdurchschnittlich erfahren. Die Interviewer sind zwei (bei Betrachtung des arithmetischen Mittels) bzw. ein Jahr (bei Betrachtung des Medians) länger beim Umfrageinstitut beschäftigt als der Durchschnittsinterviewer. Auch ihre Fähigkeit, Haushalte für eine Teilnahme zu gewinnen, gemessen am Anteil der kooperationsbereiten Haushalte unter den erreichten Haushalten pro Interviewer, ist überdurchschnittlich (65,5 % gegenüber 56,8 %). Diese Ergebnisse überraschen durchaus, doch liegt des Rätsels Lösung wohl in einer anderen Variablen verborgen. Der Workload, also die Zahl der zur Bearbeitung übergebenen Zieladressen der 73 Interviewer übertrifft das allgemeine Arbeitsquantum sowohl am arithmetischen Mittel (30 gegenüber 20) als auch am Median gemessen (24 gegenüber 14) deutlich um zehn Adressen. Eine wahrscheinliche Begründung wäre, dass mit zunehmendem Arbeitsaufwand die Wahrscheinlichkeit von Unaufmerksamkeiten organisatorischer Art (z.B. eine Adresse versehentlich nicht zu bearbeiten) zunimmt.

5.2 Modellierung von Erreichbarkeit, Befragungsfähigkeit und Kooperationsbereitschaft

Bei face-to-face-Surveys erfolgt die Zuweisung der Interviewer zu den Befragten nicht zufällig. Da ein Interviewer außerdem in der Regel mehrere Zieladressen zur Bearbeitung erhält, ist die Verwendung von herkömmlichen Regressionsmodellen problematisch, da die

Untersuchungseinheiten nicht unabhängig sind (Schnell 1997: 281-284, Schröpfer 1999: 125-129). Neben der Verletzung der Annahme, dass die Residuen unkorreliert sind, werden außerdem die Standardfehler unterschätzt und damit Koeffizienten fälschlicherweise als signifikant ausgewiesen. Als Ausweg bietet sich der Gebrauch von Mehrebenenmodellen (Goldstein 1995, Engel 1998, Kreft/de Leeuw 1998) an. Diese unterscheiden sich von herkömmlichen Regressionsmodellen durch einen komplexeren Fehlerterm. Im Gegensatz zu Mehrebenenmodellen für lineare Regressionen bringen Schätzverfahren für binäre abhängige Variablen einige Probleme mit sich (Rodríguez/Goldman 2001). Für die vorliegende Studie wurde das Programm MIXOR bzw. MIXREG (Hedeker/Gibbons 1996) in Zusammenhang mit Probit-Varianz-Komponenten-Modellen verwendet.¹³

Als unabhängige Variablen kommen nur die Angaben aus dem Adressprotokoll, die während des random walks aufgelistet wurden und Interviewermerkmale, die für den allergrößten Teil des Stabes¹⁴ vorliegen, zum Einsatz. Sie gehen sämtlich als Dummyvariablen in das Modell ein, da kein linearer Zusammenhang zwischen den unabhängigen und abhängigen Variablen vorliegt, was auch nach den aufgestellten Hypothesen nicht zu erwarten ist. Auf die teilweise vorliegenden Informationen über die Nonrespondenten muss leider verzichtet werden, auch wenn die soziodemographischen Variablen gerade zur Kontrolle der Umgebungs- und Interviewermerkmale wichtig wären.

¹³ Ursprünglich wurden alle Modelle mit den Programmen MIXOR bzw. MIXREG (Hedeker/Gibbons 1996) und MLwiN (Goldstein et al. 1998) berechnet. Während die Ergebnisse für lineare Modelle (Kontaktzahl) für beide Programme fast identisch waren, traten bei den Logit-Modellen (Erreichbarkeit, Kooperationsbereitschaft) häufig auffällige Unterschiede auf. MIXOR verwendet einen Maximum Marginal Likelihood (MML)-Ansatz, während MLwiN unter anderem Quasilikelihood (MQL und PQL)-Verfahren benutzt. Die großen Unterschiede waren wohl darauf zurückzuführen, dass bei Quasilikelihood-Verfahren die Linearisierungsfunktion bei Verwendung des Logit-Links nicht exakt bestimmt werden kann. Dies kann gerade bei kleinen Clustergrößen und extremen (Anteile nahe null oder eins, im vorliegenden Fall vor allem bei der Erreichbarkeitsquote mit .92) abhängigen Variablen (es gibt dann viele Cluster, in denen die abhängige Variable nicht variiert) problematisch sein. Nach Wahl der Probit-Spezifikation, welche exakte Linearisierung erlaubt, wurden die Ergebnisse zwischen den Programmen einheitlicher. Die Erhöhung der Zahl der Quadraturpunkte in MIXOR sorgte ebenfalls für größere Stabilität. Dennoch war die Quasilikelihood-Schätzung der Koeffizienten kleiner als die MML-Schätzung, ein häufig berichtetes Ergebnis (Rodríguez/Goldman 2001). Deshalb fiel die Entscheidung schließlich zu Gunsten von MIXOR, auch wenn über die Validität von MML-Schätzern bei kleinen Clustern noch wenig bekannt ist. Alle aufgeführten binären Modelle sind Berechnungen auf Basis einer Probit-Spezifikation bei 25 Quadraturpunkten und einem Konvergenzkriterium von .001.

An dieser Stelle möchte ich Don Hedeker (University of Illinois, Chicago) sowie Min Yang (Multilevel Models Project, University of London) für ihre hilfreichen Ratschläge danken.

¹⁴ 17 Interviewer haben fehlende Angaben für Geschlecht, Alter, Familienstand, Schulbildung und Beschäftigungsdauer im Institut. Für 24 Interviewer fehlt die Information über die Arbeitszeit; die Dauer der Mitarbeit am SOEP ist in 35 Fällen unbekannt.

Neben der zweifelhaften Generalisierbarkeit der Informationen sind die hohen Item-Nonresponse-Anteile ein Problem.

5.2.1 Erreichbarkeit

Zur Analyse der Erreichbarkeit von Haushalten werden drei Modelle berechnet, wobei die abhängige Variable den Wert eins annimmt, wenn der Haushalt erreicht wurde und den Wert null annimmt, wenn er nicht erreicht werden konnte.

Tabelle 5: Multilevel-Probit-Regression der Erreichbarkeit mit Adressprotokollinformationen und Interviewereigenschaften als erklärenden Variablen

Variablen	Modell 1		Modell 2		Modell 3	
	Koeffizient	Signifikanz	Koeffizient	Signifikanz	Koeffizient	Signifikanz
<i>Fixed effects</i>						
Stadtentfernung (Referenzkategorie: Stadtmitte Großstadt)						
Stadttrand	.207	***	.203	***	.199	***
Bis zu 25 km von Stadtmitte	.396	***	.434	***	.438	***
25 – 50 km von Stadtmitte	.369	***	.356	***	.362	***
Über 50 km von Stadtmitte	.519	***	.476	***	.478	***
Gebäudetyp (Referenzkategorie: Höhere Gebäude)						
Ein-/Zweifamilienhäuser	.276	***	.351	***	.348	***
Mehrfamilienhäuser	.099		.150	**	.150	**
Reihenhäuser	.147	***	.204	***	.199	***
Wohnumfeld (Referenzkategorie: Reines Wohngebiet)						
Hauptsächlich Wohngebiet	-.174	***	-.188	***	-.186	***
Hauptsächlich Gewerbegebiet	-.440	***	-.431	***	-.431	***
Ländliche Bebauung	-.057		-.105		-.102	
Soziale Schicht (Referenzkategorie: Unterschicht)						
Untere Mittelschicht	-.091		-.094		-.104	
Mittlere Mittelschicht	-.074		-.110		-.121	
Obere Mittelschicht/Oberschicht	-.204		-.264	*	-.271	*
Interviewervariablen						
Berufstätigkeit (Referenzkategorie: nicht berufstätig)						
Halbtags berufstätig			-.264	**	-.067	
Ganztags berufstätig			.008		.157	

Workload						
(Referenzkategorie: 1-12 Adressen)						
13-24 Adressen			-173		-200	*
25-36 Adressen			-185		-287	
37-48 Adressen			-376	**	-310	*
49-66 Adressen			-693	***	-688	**
über 66 Adressen			-266		-322	
Beschäftigungsdauer im durchführenden Institut						
(Referenzkategorie: Eintritt im Erhebungsjahr)						
1 – 3 Jahre			.564	***	.449	**
4 – 6 Jahre			.478	**	.354	*
7 – 14 Jahre			.300		.147	
15 Jahre und mehr			.243		.087	
Dauer der Mitarbeit am SOEP						
(Referenzkategorie: bisher keine)						
1 – 3 Jahre			.247		.230	
4 – 6 Jahre			.467	**	.455	***
7 – 16 Jahre			.201		.186	
Geschlecht weiblich						-055
Alter						
(Referenzkategorie: 30-39 Jahre)						
Bis 29 Jahre					-100	
40 – 49 Jahre					-239	
50 – 59 Jahre					-014	
60 – 69 Jahre					.290	
70 und älter					.334	
Konstante	1.544	***	1.261	***	1.299	***
<i>Random effects</i>						
Intercept	.744	***	.733	***	.713	***
Intraklassenkorrelation		.356		.349		.337
McFadden's PseudoR ²		.031		.088		.090
Zahl der Cluster		582		543		543
Zahl der Fälle		11250		10738		10738

* signifikant auf 10 % - Niveau, ** signifikant auf 5 % - Niveau, *** signifikant auf 1 % - Niveau
Quelle: SOEP (F) 2000, eigene Berechnungen.

Modell 1 in Tabelle 5 verwendet lediglich die für alle Haushalte vorliegenden Variablen aus dem Adressprotokoll, nämlich die Entfernung zur nächsten Großstadt, den Gebäudetyp, das Wohnumfeld und die soziale Schicht auf Basis der Schätzung des Adressenauflisters. Ein deutliches Stadt-/Landgefälle ist signifikant. Mit zunehmender Entfernung von der Stadtmitte einer Großstadt steigt die Wahrscheinlichkeit der erfolgreichen Kontaktierung eines Haushalts merklich an. Abgesehen von dem etwas kleineren Koeffizienten der Kategorie von Haushalten, die 25-50 km von der nächsten Großstadt entfernt liegen, ist der Zusammenhang monoton. Ebenfalls besser erreichbar sind Haushalte in Einfamilienhäusern und Reihenhäusern. Vermutlich liegt dies daran, dass dort eher größere Haushalte bzw. Familien mit Kindern wohnen, was für eine häufigere Anwesenheit in der Wohnung spricht.

Signifikant schlechter sind Adressen in Mischgebieten zu kontaktieren. Die prädiktive Kraft des Modells ist gering, die Entropiereduktion des Modells beträgt .03. Bemerkenswert hoch ist der Varianzanteil auf der Interviewerebene – quantifiziert durch die Intraklassenkorrelation –, woraus man schließen kann, dass sich das Kontaktverhalten der Interviewer stark unterscheidet.¹⁵

In Modell 2 werden zusätzlich die Arbeitszeit, die Erfahrungswerte und die Arbeitsbelastung des Interviewers mit aufgenommen. Der Einfluss der Stadtentfernung bleibt höchst signifikant, die Koeffizienten werden allerdings teilweise etwas kleiner. Haushalte mit der größten Distanz zur Stadt sind nach wie vor am besten zu kontaktieren. Die Koeffizienten der Haustyp-Variablen vergrößern sich, zusätzlich wird auch der Dummy für Mehrfamilienhäuser signifikant, wobei die Größenordnung der Koeffizienten den Erwartungen entspricht. Keine allzu großen Änderungen betreffen die Wohnumfeldvariablen. Nur auf 10 % - Niveau signifikant ist der negative Koeffizient der Haushalte aus der oberen Mittelschicht/Oberschicht. Interviewer, die nebenbei halbtags tätig sind, sind schlechtere „Kontaktierer“. Dies mag zum einen an eingeschränkteren Kontaktzeitpunkten, zum anderen daran liegen, dass dies eher Personen sind, welche die Arbeit aus finanziellen Gründen betreiben und daher sehr belastet und weniger motiviert sind. Während ein Workload von 37-48 sowie von 49-66 Zieladressen deutlich negative Auswirkungen hat, trifft dies für die höchste Kategorie nicht zu. Diese Interviewer stellen eine besonders motivierte Subgruppe dar. Wie erwartet zeigt sich der Zusammenhang zwischen Institutserfahrung und Erfolg. Am meisten dazu gelernt wird am Anfang, in der zweiten Kategorie lässt die Wirkung schon nach, ehe die Signifikanz ganz verschwindet. Für die SOEP-Erfahrung erhöht jedoch nur die Zugehörigkeit zur mittleren Kategorie die Kontaktquote. Der Modellfit steigt durch die Hinzunahme der Interviewereigenschaften deutlich auf ca. .09 an; die eingeführten Interviewereigenschaften verringern allerdings die Varianz auf der zweiten Ebene kaum. Folglich sind sie keine besonders guten Indikatoren der Kontaktstrategien.

Modell 3 greift zusätzlich auf die für den Befragten sichtbaren Merkmale der Interviewer zurück, um zu überprüfen, ob eventuell z.B. auf Grund von Viktimisierungsängsten eine Türöffnung unterbleibt. Hierfür lassen sich keine Hinweise

¹⁵ Auf der Ebene einzelner sampling points sind Interviewer- und sampling-point-Effekte im vorliegenden Fall nicht trennbar, da keine „interpenetrierte“ Stichprobe vorliegt. Die wichtigsten Spezifika der sampling-points werden allerdings – über verschiedene sampling points hinweg – durch die unabhängigen Variablen erfasst.

finden. Weder die Geschlechtsvariable noch die Altersdummies sind signifikant von null verschieden. Jedoch verschwindet die Signifikanz der Variable für Teilzeitbeschäftigung. Dies dürfte auf einen Zusammenhang der Berufstätigkeit mit den hinzugenommenen Variablen hindeuten und zeugt von der Instabilität des Effekts.

5.2.1.1 Alternative Modellierung der Erreichbarkeit

Eine zweite Möglichkeit, Erreichbarkeitsmodelle zu schätzen, ist die Operationalisierung der Erreichbarkeit als Zahl der Kontaktversuche, die bis zum ersten erfolgreichen Kontakt nötig sind (Schnell 1997: 219, Groves/Couper 1998: 81). Dies eröffnet außerdem – sofern lediglich die schließlichen Teilnehmer betrachtet werden – den Zugriff auf eine größere Anzahl unabhängiger Variablen.

Der vorliegende Datensatz zur Stichprobe F enthält allerdings nicht die Zahl der Kontakte bis zum erfolgreichen Erstkontakt, sondern nur die Gesamtzahl der Interviewerkontakte. Die Approximierung der Erreichbarkeit über diese Variable könnte dann verzerrt werden, wenn die Zahl jener Kontakte, die *nach erfolgreichem Erstkontakt* bis zum Abschluss der Feldarbeit – gleich mit welchem Ergebnis – am einzelnen Haushalt nötig waren, nicht über alle Haushalte gleich verteilt sind. Dieser Indikator vermischt also potentiell die Konzepte der Erreichbarkeit und Kooperationsbereitschaft. Nach dem deutschen Datenschutzgesetz darf ein Haushalt nach einer explizit erfolgten Verweigerung allerdings nicht nochmals belangt werden. Eine zentrale Nachbearbeitung wie in der ersten Welle des SOEP 1984 fand nicht statt. Problematisch für die Analysen wären aber auch Kontakte, in denen durch tatsächliche oder vorgegebene Zeitprobleme („Geht im Moment nicht, ein andermal“) der Zeitpunkt, zu dem das endgültige Bearbeitungsergebnis feststeht, hinausgezögert wird. Außerdem könnte vermutet werden, dass eine mögliche Verzerrung dadurch entsteht, dass Mehrpersonenhaushalte nach erfolgreichem Erstkontakt öfter besucht werden müssen, bis für alle Personen der Bearbeitungsstatus feststeht bzw. ein Interview vorliegt. Andererseits liegt aber auch nahe, dass Interviewer im eigenen Interesse nach dem Erstkontakt weitere Kontakte zunächst telefonisch verabreden, um die Kontaktzahl sowie Wegekosten möglichst klein halten.

Empirisch lässt sich hier kaum Licht ins Dunkel bringen. Eine lineare Regression der Haushaltsgröße auf die Kontaktzahl zeigt keinen Zusammenhang (F-Test nicht signifikant, R^2 praktisch 0). Dies ist allerdings kein Argument gegen die These, dass Mehrpersonenhaushalte nach der erfolgreichen Kontaktierung mehr Kontakte benötigen. Es könnte vielmehr so

gedeutet werden, dass dies erst dazu beiträgt, den Zusammenhang, der durch die erschwerte Erreichbarkeit von Einpersonenhaushalten entsteht, zu verwischen.¹⁶

5.2.1.1.1 Modell für alle Haushalte

Basis der Schätzungen der Kontaktzahl sind alle Haushalte aus der Bruttostichprobe, sofern sie auf den unabhängigen Variablen keine fehlenden Werte aufweisen. Deshalb sind die folgenden Modelle für alle Haushalte weniger als Schätzung der Erreichbarkeit, sondern explizit der Kontaktzahl anzusehen, da sie auch die nicht erreichbaren Haushalte beinhalten. Eine Beschränkung auf erreichte Haushalte findet sich dann innerhalb der Beschränkung auf realisierte Zieladressen im nächsten Abschnitt wieder. Die beiden folgenden Modelle sind lineare Regressionen auf die Variable Kontaktzahl, die zur Normalisierung der Verteilung box-cox-transformiert wurde.

Tabelle 6: Lineare Multilevel-Regression auf die Zahl der Kontakte für die Bruttostichprobe mit Adressprotokollinformationen und Interviewereigenschaften als erklärenden Variablen

Variablen	Modell 1		Modell 2	
	Koeffizient	Signifikanz	Koeffizient	Signifikanz
<i>Fixed effects</i>				
Stadtentfernung (Referenzkategorie: Stadtmitte Großstadt)				
Stadttrand	.003		.000	
Bis zu 25 km von Stadtmitte	-.035		-.045	*
25 – 50 km von Stadtmitte	-.015		-.015	
Über 50 km von Stadtmitte	-.046		-.061	*
Gebäudetyp (Referenzkategorie: Höhere Gebäude)				
Ein- /Zweifamilienhäuser	-.012		-.009	
Mehrfamilienhäuser	.023		.030	
Reihenhäuser	.014		.009	

¹⁶ In einem anderen Versuch wurde die Kontaktzahl folgendermaßen transformiert: Für alle Haushalte, in denen mehr als 2 Personen leben und die drei und mehr Kontakte aufweisen, wurden zwei Kontakte abgezogen. Rechnet man nun die obengenannte Regression der Haushaltsgröße auf die Kontaktzahl, ergibt sich ein R^2 von 6 %, der t-Wert für die Haushaltsgröße ist mit dem Betrag von –20 besonders groß. Inwieweit diese Transformation aber sachlich zutreffend ist, ist fraglich. Im folgenden wird deshalb mangels sinnvoller Alternativen die Variable in ihrer ursprünglichen Form aus dem Datensatz verwendet.

Wohnumfeld			
(Referenzkategorie: Reines Wohngebiet)			
Hauptsächlich Wohngebiet	.012		.009
Hauptsächlich Gewerbegebiet	.083 *		.084 *
Ländliche Bebauung	-.021		.003
Soziale Schicht			
(Referenzkategorie: Unterschicht)			
Untere Mittelschicht	.029		.008
Mittlere Mittelschicht	.080 **		.055
Obere Mittelschicht/ Oberschicht	.125 ***		.091 **
Interviewervariablen			
Berufstätigkeit			
(Referenzkategorie: nicht berufstätig)			
Halbtags berufstätig			-.054
Ganztags berufstätig			-.040
Workload			
(Referenzkategorie: 1-12 Adressen)			
13-24 Adressen			.078 **
25-36 Adressen			.148 ***
37-48 Adressen			.134 **
49-66 Adressen			.091
über 66 Adressen			.304 ***
Dauer der Mitarbeit am SOEP			
(Referenzkategorie: bisher keine)			
1 – 3 Jahre			-.095 *
4 – 6 Jahre			.057
7 – 16 Jahre			.037
Beschäftigungsdauer im durchführenden Institut			
(Referenzkategorie: Eintritt im Erhebungsjahr)			
1 – 3 Jahre			-.044
4 – 6 Jahre			.074
7 – 14 Jahre			-.019
15 Jahre und mehr			.044
Konstante	.774 ***		.729 ***
<i>Random effects</i>			
Intercept	.108 ***		.101 ***
Intraklassenkorrelation		.270	.256
McFadden's PseudoR ²		.012	.060
Zahl der Cluster		606	567
Zahl der Fälle		11745	11203

* signifikant auf 10 % - Niveau, ** signifikant auf 5 % - Niveau, *** signifikant auf 1 % - Niveau

Quelle: SOEP (F) 2000, eigene Berechnungen.

Modell 1 für alle Haushalte beruht nur auf den Adressprotokollinformationen und ist von äußerst geringer prädiktiver Power (Pseudo R² ca. .01). Ein 5 %- bzw. 10 % - Signifikanz-

niveau erreichen nur die positiven Koeffizienten für mittlere Mittelschicht und obere Mittelschicht/Oberschicht.

Die Vergrößerung des Modells analog zu Modell 2 in der logistischen Regression erreicht immerhin ein Pseudo- R^2 von .06. Die Koeffizienten der sozialen Schicht werden deutlich reduziert, die Kategorie „Mittelschicht“ weist nicht mehr signifikant mehr Kontakte auf als die Referenzkategorie „Unterschicht“. Abgesehen von dem auf 10 %-Niveau signifikanten negativen Koeffizienten der Kategorie „1-3 Jahre“ bezüglich der SOEP-Erfahrung beeinflusst nur die Zahl der Zieladressen die Kontaktzahl auf Interviewerebene. Eher kontraintuitiv ist, dass die Interviewer mit dem geringsten Workload die wenigsten Kontakte versuchen, während die am stärksten Belasteten auch am häufigsten kontaktieren. Dies lässt auf Motivationsgründe schließen. In irgendeiner Form berufstätige Interviewer unterscheiden sich in der Intensität der Kontakte nicht von ihren hauptberuflichen Kollegen.

Vergleicht man die Schätzungen auf die Kontaktzahl mit dem binären Ansatz des Erreichbarkeitskonzepts, so offenbaren sich doch relativ große Unterschiede. In den Umgebungsvariablen erscheinen bei den linearen Ansätzen weit weniger Effekte als signifikant, obwohl der Anteil der Varianz auf Haushaltsebene im Vergleich zu den Probitmodellen zugenommen hat, da der Interviewervarianzanteil von ca. 35 % auf 26 % abgenommen hat. Auf Interviewerebene spielt nur der Workload eine Rolle, nicht jedoch Erfahrung und Berufstätigkeit.

Diese Differenzen legen zwei Schlussfolgerungen nahe. Zum einen ist, wie oben ausgeführt, die Zahl der Kontakte nur eine Komponente der Kontaktstrategie, aber nicht die letztlich entscheidende. So kontaktieren Interviewer mit 37-48 Adressen häufiger, erzielen aber trotzdem schlechtere Kontaktquoten. Die im Vergleich zu den Probit-Modellen kleinere Intraklassenkorrelation deutet darauf hin, dass sich die Interviewer noch auf einer anderen Dimension als der Quantität der Kontakte unterscheiden. Zum zweiten könnten die abweichenden Ergebnisse auch durch das oben angesprochene Problem, dass die abhängige Variable nicht die Zahl der Kontaktversuche bis zum ersten Kontakt misst, verursacht sein.

5.2.1.1.2 Modell nur für teilnehmende Haushalte

Die folgenden Berechnungen für die Zahl der Kontakte beschränken sich auf die teilnehmenden Haushalte. Wie oben bereits erwähnt, erlaubt die Beschränkung nur auf

teilnehmende Haushalte eine Ausweitung der Informationsgrundlage. Will man diese Ergebnisse verallgemeinern, muss man annehmen, dass sich die teilnehmenden Haushalte in Bezug auf die Erreichbarkeit nicht von den nichtteilnehmenden Haushalten unterscheiden. Die Auswahl der unabhängigen Variablen für Modell 1 ist bei Schnell (1997: 221) entlehnt, zusätzlich werden noch drei Dummy-Variablen zum Gesundheitszustand aufgenommen. Alle Personeneigenschaften beziehen sich auf den Haushaltsvorstand.

Zum einen zeigen sich deutliche Alterseffekte: Die Zahl der Kontaktversuche nimmt mit dem Alter des Haushaltsvorstands ab, die Koeffizienten nehmen monoton zu. Hat der Haushaltsvorstand einen Fachschul-, Fachhochschul- oder Universitätsabschluss vorzuweisen, sind gegenüber der Referenzkategorie Lehre ebenfalls mehr Kontakte zu beobachten. Außerdem erhöhen Berufstätigkeit und Wohnen in größeren Gebäuden die Kontaktzahl. Nicht-deutsche Staatsangehörigkeit hat ebenso wie das Vorhandensein von Kindern im Haushalt negativen Einfluss. Außerdem weisen Geschiedene eine höhere Kontaktzahl auf. Signifikanzniveau erreichen weder die Gesundheitsvariablen noch der bei Schnell (1997: 220) betonte Telefonbesitz.

Tabelle 7: Lineare Multilevel-Regression auf die Zahl der Kontakte für die Nettostichprobe mit Befragteneigenschaften sowie Interviewermerkmalen als erklärenden Variablen

Variablen	Modell 1		Modell 2	
	Koeffizient	Signifi- kanz	Koeffizient	Signifi- kanz
<i>Fixed effects</i>				
Altersgruppe (Referenzkategorie: Unter 29)				
30 – 39	-.058		-.054	
40 – 49	-.077	*	-.079	*
50 – 59	-.154	***	-.168	***
60 – 69	-.192	***	-.201	***
70 - 79	-.239	***	-.250	***
80 und älter	-.261	***	-.274	***
Familienstand (Referenzkategorie: Verheiratet und zusammenlebend)				
Verheiratet, getrennt	.019		.020	
Ledig	.010		.004	
Geschieden	.060	**	.069	**
Verwitwet	.038		.038	
Ausbildung (Referenzkategorie: Lehre)				
Fachschule	.053	**	.060	***
Beamte	.074		.079	*
FH	.076	**	.074	**
Uni	.069	**	.062	
Sonstige	-.022		-.011	

Gesundheitszustand (Referenzkategorie: schlecht)		
Sehr gut	-.045	-.066
Gut	-.002	-.021
Zufriedenstellend	.040	.010
Weniger gut	.005	-.015
Einpersonenhaushalt	-.072 **	-.078 **
Größeres Gebäude (mehr als fünf Wohnungen)	.057 **	.060 ***
Großstadt	-.004	.007
Telefonbesitz	.029	.046
Kinder im Haushalt	-.075 ***	-.083 ***
Berufstätigkeit	.131 ***	.121 ***
Nicht-deutsche Staatsangehörigkeit	-.170 ***	-.189 ***
Interviewervariablen		
Berufstätigkeit (Referenzkategorie: nicht berufstätig)		
Halbtags berufstätig		.006
Ganztags berufstätig		-.034
Workload (Referenzkategorie: 1-12 Adressen)		
13-24 Adressen		.031
25-36 Adressen		.074
37-48 Adressen		.053
49-66 Adressen		.012
über 66 Adressen		.192 *
Dauer der Mitarbeit am SOEP (Referenzkategorie: bisher keine)		
1 – 3 Jahre		-.094
4 – 6 Jahre		.258 ***
7 – 16 Jahre		.048
Beschäftigungsdauer im durchführenden Institut (Referenzkategorie: Eintritt im Erhebungsjahr)		
1 – 3 Jahre		-.009
4 – 6 Jahre		.046
7 – 14 Jahre		-.124
15 Jahre und mehr		.019
Konstante	1.110 ***	1.083 ***
<i>Random effects</i>		
Intercept	.123 ***	.115 ***
Intraklassenkorrelation	.286	.274
McFadden's PseudoR ²	.223	.273
Zahl der Cluster	532	501
Zahl der Fälle	4834	4552

Der überraschenderweise leicht negative Koeffizient für Einpersonenhaushalte könnte mit der Kontrolle des Alters in Verbindung stehen, vielleicht aber auch daraufhin deuten, dass

die Variable – wie oben erörtert – nicht die Erreichbarkeit misst. Die Modellgüte erreicht in diesem Zusammenhang mit .22 einen äußerst zufriedenstellenden Wert – gerade auch im Vergleich zu den Berechnungen für mehrere Surveys von Schnell (1997: 220). Weitere als unabhängige Variablen eingeführten Merkmale (Länge der berichteten Tagesarbeitszeit, des Haushaltsvorstands, Arbeit in den Abendstunden, großer Zeitaufwand für Hobbys) waren nicht signifikant und verbesserten entsprechend das Modell nicht merklich.

Modell 2 vergrößert wiederum das Set der unabhängigen Variablen um dieselben Interviewermerkmale wie oben. Für die Eigenschaften des Haushaltsvorstands ergeben sich keine substantiellen Änderungen, außer dass der Dummy für einen akademischen Bildungsabschluss nicht mehr signifikant von null verschieden ist. Auf Interviewerebene werden kaum Effekte sichtbar. Nur Interviewer, die schon vier bis sechs Jahre am SOEP mitarbeiten, unternehmen mehr Kontaktversuche als ihre Kollegen. Die Entropiereduktion steigt auf .27 an, der Varianzanteil auf Interviewerebene entspricht dem selben Wert.

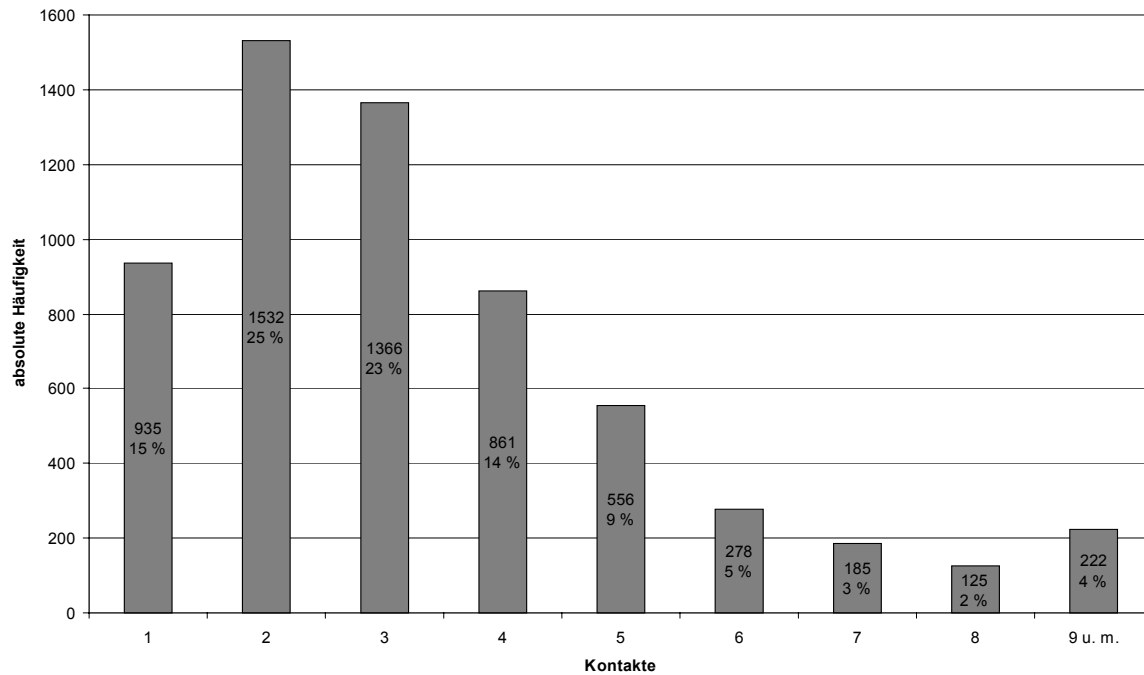
Vergleicht man die vorliegenden Modelle mit dem binären Ansatz (vgl. Abschnitt 5.2.1 oben), so kann man Gemeinsamkeiten und Unterschiede feststellen. Haushalte in höheren Gebäuden weisen mehr Kontakte auf und sind auch schwerer erreichbar. Interviewer, die vier bis sechs Jahre am SOEP mitarbeiten, kontaktieren häufiger und erzielen auch höhere Quoten. Leben in der Großstadt beeinflusst die Kontaktzahl nicht, hat aber auf die Erreichbarkeit großen Einfluss. Die Ergebnisse hinsichtlich der Arbeitsbelastung des Interviewers sind uneinheitlich.

Die festgestellten Unterschiede zwischen dichotomen und linearen Ansätzen können darauf zurückzuführen sein, dass die Kontrolle der soziodemographischen Variablen des Haushaltsvorstands zu Veränderungen in den Schätzungen der Umgebungs- und Interviewervariablen führt.

5.2.2 Der Einfluss der Kontaktzahl auf die Stichprobenstruktur

Wie oben bereits erwähnt werden im SOEP, um eine möglichst hohe Qualität der Stichprobe zu erhalten, weit mehr Kontaktversuche unternommen als es gemeinhin – insbesondere bei kommerziellen Surveys – üblich ist. Einen Überblick über die Verteilung der Kontaktzahl für realisierte Haushalte verschafft die untenstehende Grafik. Man kann sehen, dass ca. ein Viertel der Haushalte mehr als vier Kontakte aufweist.

Zahl der Kontakte
(bezogen auf mind. teilrealisierte Haushalte)



Ob der hohe Aufwand gerechtfertigt ist bzw. ob die Kontaktzahl mit inhaltlichen Variablen zusammenhängt, lässt sich relativ einfach und anschaulich demonstrieren, indem man – für Gruppen getrennt nach der Kontaktzahl – Lageparameter von soziodemographischen und subjektiven Variablen berechnet (vgl. Schnell 1998: 156 f.).

Tabelle 8: Zusammenhang zwischen Kontaktzahl und ausgewählten demographischen und subjektiven Variablen: Spearmans Rho zwischen Kontaktzahl und Anteils- bzw. Mittelwerten (getrennt nach Haushaltsgrößen)

Variable	Haushaltsgrößen					
	Eine Person	Zwei Personen	Drei Personen	Vier Personen	Fünf u. mehr Personen	Über alle Haushaltsgrößen hinweg
Alter	-.95	-.88	-.38	-.72	-.07	-.95
Anteil Arbeiter unter den Berufstätigen	-.62	-.58	-.53	-.97	.07	-.78
Anteil von Hochschulreife unter Schulabschlüssen	.88	.85	.43	.58	-.10	.77
Anteil an Personen mit nichtdeutscher Nationalität	.65	.03	-.43	-.28	-.79	-.42
Anteil an PC-Nutzern (privat)	.77	.95	.78	.73	.40	.98
Anteil an Internetnutzern (privat)	.85	.95	.77	.73	.23	.98
Starkes oder sehr starkes Interesse an Politik	-.17	.10	.73	-.55	-.03	-.03

Haushaltseinkommen	.83	.75	.53	.77	.67	.72
Sorgenangaben Große Sorgen in Hinblick auf...						
Allg. Wirtschaftliche Entwicklung	-.07	-.33	-.77	.32	.58	-.32
Kriminalitäts- entwicklung	-.70	-.72	-.42	.67	-.41	-.60
Ausländer- feindlichkeit	.42	.27	.57	.23	-.15	.35
Zuwanderung	-.53	-.83	-.08	-.08	.00	-.67
Einführung des Euro	-.68	-.73	.15	-.02	.47	-.57
Sicherheit des eigenen Arbeitsplatzes	-.10	-.30	-.30	-.37	-.60	-.78
Zufriedenheits- angaben Skala von 0 (ganz und gar unzufrieden) bis 10 (ganz und gar zufrieden) mit...						
Umweltzustand in der Region	-.23	.03	-.52	-.77	-.38	-.60
Haushalts- einkommen	.47	.30	-.22	-.48	.35	-.03
„alles in allem, mit ihrem Leben“	.02	.07	-.02	-.53	.13	-.27
Zahl der Personen ¹	1668	3957	2073	2148	1055	10901
Anteil in %	15.3	36.3	19.0	19.7	9.7	100

¹ Theoretische Fallzahlen unter der Annahme, dass keine fehlenden Werte vorhanden sind.

Quelle: SOEP (F) 2000, ungewichtet, eigene Berechnungen.

Tabelle 8 gibt entsprechende Ergebnisse wieder. Aufgeführt ist der Rangkorrelationskoeffizient (Spearman's Rho) zwischen der Kontaktzahl und den Anteils- bzw. Mittelwerten innerhalb einer Gruppe mit gleich vielen Kontakten. Sämtlichen Berechnungen in diesem Abschnitt liegt nicht die Haushalts-, sondern die Personenstichprobe mit allen Befragungspersonen eines Haushalts zugrunde, da der Zusammenhang zwischen Kontaktzahl und Struktur der Gesamtstichprobe von Interesse ist. Um eines der obengenannten Probleme mit der Kontaktzahl zu umgehen, wurden die Berechnungen für jede Haushaltsgröße separat durchgeführt.

Wie schon in den Regressionen auf die Kontaktzahl deutlich wurde, sind später Kontaktierte wesentlich jünger. Der Anteil der Arbeiter nimmt mit der Kontaktzahl ab, im Gegenzug nimmt der Anteil von Personen mit Hochschulreife zu. Was den Ausländeranteil betrifft, ist die Richtung des Zusammenhangs über die Haushaltsgrößen hinweg nicht einheitlich. Für Einpersonenhaushalte ist mit zunehmender Zahl der Kontakte erwartungsgemäß ein Anstieg zu verzeichnen. Das Haushaltseinkommen nimmt mit der Kontaktzahl deutlich zu. Die stärkste Korrelation – über alle Haushaltsgrößen betrachtet –

findet sich zwischen Kontaktzahl und sowohl PC- als auch Internetnutzern. Für politisches Interesse hingegen ist das Bild uneinheitlich, während auf Haushaltsebene das Einkommen mit der Kontaktzahl deutlich zunimmt.

Unter den subjektiven Variablen wurden diverse Sorgenangaben (Anteil der Personen mit großen Sorgen gegenüber Personen mit einigen und keinen Sorgen) und Zufriedenheitsskalen ausgewählt. Man kann folgende Feststellungen treffen: Unter Spätkontaktierten machen sich weniger Leute große Sorgen um die Kriminalitätsentwicklung, um die Zuwanderung nach Deutschland, um die Einführung des Euro und um ihren Arbeitsplatz. Der Anteil der Leute mit großen Sorgen um die Ausländerfeindlichkeit in Deutschland nimmt mit der Zahl der Kontakte zu. Unzufriedener sind Leute mit hoher Kontaktzahl was den Umweltzustand ihrer Region anbelangt. Für die Zufriedenheit hinsichtlich des Haushaltseinkommens und der allgemeinen Lebenszufriedenheit ist kein einheitlicher Trend angebar. Alles deutet daraufhin, dass Gebildete und „Kritische“ mehr Kontakte benötigen.

Teilweise unterscheiden sich Vorzeichen und Stärke des Zusammenhangs zwischen Kontaktzahl und den interessierenden Variablen innerhalb der Kategorien kleinerer Haushalte von denen innerhalb der Kategorien größerer Haushalte. Es sei nochmals bemerkt, dass dies an den Problemen mit der Kontaktzahlvariablen liegen kann. Es sollte aber auch bedacht werden, dass Fallzahlen für die größeren Haushalte kleiner sind und damit Ausreißer wahrscheinlicher werden.

Man könnte fragen, ob die Unterschiede in den subjektiven Variablen nicht zum Großteil auf die Unterschiede in den demographischen Variablen zurückzuführen sind. Um dies zu überprüfen, wurden einige einfache multiple lineare bzw. logistische Regressionsmodelle mit soziodemographischen unabhängigen Variablen geschätzt. Zur Kontrolle der Haushaltsgröße werden entsprechende Dummyvariablen eingeführt. Daneben umfasst Set 1 nur das Alter, Set 2 Alter, Schulbildung, Nationalität und Haushaltseinkommen (logarithmiert) als erklärende Variablen. Ziel des Unterfangens ist in diesem Rahmen nicht, inhaltlich vollständige Modelle aufzustellen, sondern zu untersuchen, ob die Kontaktzahl über die demographischen Variablen hinaus einen Einfluss hat. Dies erfolgt über die Ausweitung des Sets der unabhängigen Variablen um die Kontaktzahl und einen anschließenden Vergleich der beiden Modelltypen.

Tabelle 9: Multivariate Modelle zum Zusammenhang zwischen Kontaktzahl und ausgewählten soziodemographischen und subjektiven Variablen

	<i>R² bzw. Pseudo-R² im Grundmodell (in Klammern: Verwendetes Set¹ an unabhängigen Variablen)</i>	<i>Signifikanz der Kontaktzahlvariablen bei Hinzufügung zum Grundmodell (in Klammern: Vorzeichen des Koeffizienten)</i>	<i>Differenz in R² bzw. Pseudo- R² zwischen Grundmodell und Grundmodell mit Kontaktzahl</i>
Anteil Arbeiter unter den Berufstätigen	.005 (Set 1) ¹	*** (-)	.002
Anteil von Hochschulreife unter Schulabschlüssen	.033 (Set 1)	*** (+)	.002
Anteil an PC- Nutzern (privat)	.214 (Set 2) ¹	*** (+)	.002
Anteil an Internetnutzern (privat)	.190 (Set 2)	*** (+)	.002
Starkes oder sehr starkes Interesse an Politik	.093 (Set 2)	n.s.	.000
Haushaltseinkommen	.437 (Set 2)	*** (+)	.004
Sorgenangaben Große Sorgen in Hinblick auf...			
Allg. Wirtschaftliche Entwicklung	.018 (Set 2)	n.s.	.000
Kriminalitätsentwicklung	.028 (Set 2)	n.s.	.000
Ausländerfeindlichkeit	(.014)	n.s.	.000
Zuwanderung	.040 (Set 2)	** (-)	.001
Einführung des Euro	.033 (Set 2)	n.s.	.000
Sicherheit des eigenen Arbeitsplatzes	.057 (Set 2)	n.s.	.000
Zufriedenheitsangaben Skala von 0 (ganz und gar unzufrieden) bis 10 (ganz und gar zufrieden) mit...			
Umweltzustand in der Region	.011 (Set 2)	*** (-)	.002
Haushaltseinkommen	.208 (Set 2)	*** (-)	.003
„alles in allem, mit ihrem Leben“	.057 (Set 2)	*** (-)	.002

¹ Set 1: Alter, Haushaltsgröße, Set 2: Alter, Schulbildung, Nationalität, Haushaltseinkommen (logarithmiert), Haushaltsgröße.
n.s. nicht signifikant, * signifikant auf 10 % - Niveau, ** signifikant auf 5 % - Niveau, *** signifikant auf 1 % - Niveau

Quelle: SOEP (F) 2000, ungewichtet, eigene Berechnungen.

Mit zunehmender Kontaktzahl sinkt der Arbeiteranteil unter den Berufstätigen signifikant, der Hochschulreifeanteil unter den Bildungsabschlüssen nimmt bei steigender Kontaktzahl signifikant zu. Einen positiven Zusammenhang zeigt die Kontaktzahl über die demographischen Variablen hinaus auch mit dem Anteil der PC-/Internetnutzer und dem Haushaltseinkommen. Nicht signifikant wird die Variable Kontaktzahl im Modell für das

Interesse an Politik. Was die Sorgenangaben angeht, erniedrigt die Kontaktzahl nur den Anteil der Leute mit großen Sorgen um die Zuwanderung überzufällig. Somit haben auch einige vermeintlich deutlichere Zusammenhänge diesen Test nicht überstanden. In den Modellen zu den ausgewählten Zufriedenheitsangaben hingegen ist in allen drei Fällen der Einfluss signifikant von null verschieden. Geht man allerdings vom Kriterium der Signifikanztests zur Betrachtung der Differenzen im Modellfit über, um den Effekt der Kontaktzahl abzuschätzen, zeigt sich ein anderes Bild. Die der Kontaktvariablen zuzuschreibende Erklärungskraft ist faktisch gleich null.

Eine andere Möglichkeit, den Einfluss der Kontaktzahl zu prüfen, besteht darin, nicht nach einzelnen durch die Kontaktzahl festgelegten Subgruppen zu trennen, sondern vielmehr hypothetische Stichproben zu bilden, die sich ergeben hätten, wenn nach einer bestimmten Zahl keine weiteren Kontakte mehr unternommen worden wären. Die Schnittstelle wurde hier bei sechs Kontakten angesetzt, die Generierung der Stichprobe beruht auf dem von Curtin et al. (2000: 423f.) vorgeschlagenen Verfahren: Es werden nicht einfach Berechnungen mit den Fällen, die bis zu sechs Kontakte aufweisen, anderen Berechnungen mit allen Fällen gegenübergestellt. Dies hätte nämlich zur Folge, dass sich die Ergebnisse schon allein dadurch ähnlich sind, dass die „kleine“ Stichprobe eine Untermenge der realen Stichprobe ist. Stattdessen wird die reale Stichprobe nach Haushalten zufällig in zwei Unterstichproben aufgeteilt, wobei in diesem Fall eine Unterstichprobe 52.2 %, die andere 47.8 % der Haushalte umfasst. Diese Zahlen kommen zustande, weil die Zahl der Haushalte, die mehr als sechs Kontakte benötigt haben, 8.8 % beträgt. Teilt man nämlich jetzt die reale Stichprobe durch ein Zufallsverfahren in zwei ungefähr gleich große Teile, so umfassen beide Teile ungefähr 4.4 % der Haushalte mit mehr als sechs Kontakten. Von der Stichprobe mit 52.2 % aller Fälle werden nun diese ungefähr 4.4 % der Fälle entfernt. Es resultieren nun zwei unabhängige Stichproben von ungefähr gleicher Größe (ca. 47.8 % aller Fälle). Die „kleine“ Stichprobe besteht aus 2879 Haushalte mit 5210 Personen, die mit bis zu sechs Kontakten realisiert wurden.

Tabelle 10: Vergleich der zwei hypothetischen Stichproben hinsichtlich ausgewählten soziodemographischen und subjektiven Variablen

Variable	Referenzkategorie	% in Referenzkategorie bzw. Mittelwert		Chi2-Pr. Alle Kategorien (df) bzw. t-Test-Pr / zweiseitiger Test (df)
		Unter- stichprobe bis zu 6 Kontakte	Unter- stichprobe Alle Kontakte	
Alter		47.64	47.22	.22 (10409)
Berufsstellung unter den Berufstätigen	Arbeiter	30.2	29.5	.80 (4)
Schulbildung	Hochschulreife	20.1	20.8	.54 (4)
Nationalität	Nicht deutsch	6.7	6.0	.14 (1)
Haushaltseinkommen		4011	4028	.80 (5193)
PC- Nutzung (privat)	Ja	33.8	34.6	.39 (1)
Internetnutzung (privat)	Ja	16.3	16.4	.97 (1)
Interesse an Politik	Überhaupt nicht	15.6	12.8	.00 (1) ***
Sorgenangaben Sorgen in Hinblick auf...				
Allg. Wirtschaftliche Entwicklung	Große Sorgen	22.9	21.9	.45 (2)
Kriminalitätsentwicklung	Große Sorgen	56.1	55.5	.79 (2)
Ausländerfeindlichkeit	Große Sorgen	36.6	36.5	.94 (2)
Zuwanderung	Große Sorgen	29.4	28.7	.66 (2)
Einführung des Euro	Große Sorgen	31.3	31.0	.78 (2)
Sicherheit des eigenen Arbeitsplatzes	Große Sorgen	16.2	15.4	.15 (2)
Zufriedenheits- angaben Skala von 0 (ganz und gar unzufrieden) bis 10 (ganz und gar zufrieden) mit...				
Umweltzustand in der Region		6.75	6.77	.66 (10367)
Haushalts- einkommen		6.52	6.56	.48 (10190)
„alles in allem, mit ihrem Leben“		7.28	7.32	.30 (10384)

Quelle: SOEP (F) 2000, eigene Berechnungen.

Die „große“ Stichprobe, welche auch Fälle mit mehr als sechs Kontakten beinhaltet, umfasst 2910 Haushalte mit 5202 Personen.

Tabelle 10 informiert über die Ergebnisse des Vergleichs. Die Differenzen der Anteilswerte (die entsprechende Referenzkategorie ist jeweils angegeben) sind minimal und liegen meist unter einem Prozent. Ausnahmen sind „überhaupt kein politisches Interesse“ sowie „große Sorgen um die allgemeine wirtschaftliche Entwicklung“. Berechnet man für die Verteilung der nominal- bzw. ordinalskalierten Merkmale in den Stichproben Chi²-Tests, so ergibt sich lediglich ein signifikanter Wert, und zwar für das Interesse an Politik. Dies ist bemerkenswert, da diese Variable in den anderen Berechnungen nicht auffiel. Dies liegt aber daran, dass sich kaum Unterschiede in der dort verwendeten Referenzkategorie „starkes oder sehr starkes Interesse“, wohl aber in den anderen Kategorien finden. In der „großen“

Stichprobe (auch mit Fällen mit mehr als sechs Kontakten) interessieren sich mehr Leute nicht so stark für Politik und weniger Leute überhaupt nicht für Politik. Allerdings muss auch die doch recht große Fallzahl beachtet werden. So erreicht das normierte Cramers V nur einen Wert von .0419, der Zusammenhang ist also nicht besonders groß. Vergleicht man nicht alle Personen, sondern nur Haushaltsvorstände, so gibt es keinen großen Unterschied mehr zwischen den Stichproben, der p-Wert liegt bei .83.¹⁷

Die verglichenen Mittelwerte sind sich sehr ähnlich. Keiner der t-Tests erreicht konventionelles Signifikanzniveau. Außerdem ist zu bedenken, dass die angeführten p-Werte auf unter der Annahme einer einfachen Zufallsstichprobe berechneten t-Werten beruhen. Die Standardfehler werden in komplexen Klumpenstichproben wie dem SOEP unterschätzt, sämtliche angegebenen p-Werte sind also zu gering.

5.2.3 Befragungsfähigkeit

Erwartungsgemäß bildet eine Probit-Regression der Adressprotokollvariablen und Interviewereigenschaften auf die Befragungsfähigkeit kein geeignetes Modell des zugrundeliegenden Prozesses, da gesundheitliche Probleme – im vorliegenden Survey, wie oben schon ausgeführt, die Hauptursache für Befragungsunfähigkeit – zufällig über die unterschiedlichen Quartiersmerkmale und Interviewer verteilt sind. Ein Rückgriff auf die Informationen, die nicht für alle Adressen vorliegen, ist angebracht, da sie für die Gruppe der Nicht-Befragungsfähigen zu einem recht großen Teil (um die 80 %, siehe oben Tabelle 4) vorliegen und der Anteil der Befragungsfähigen in der hier relevanten Stichprobe der mindestens erreichten Haushalte und der Anteil der Befragungsfähigen in der Schätzstichprobe sich erst ab der vierten Dezimalstelle unterscheiden. Auf ein Mehrebenenmodell wird hier verzichtet, da in den Ergebnissen keine großen Unterschiede zu einem Standardmodell zu erwarten sind.¹⁸

¹⁷ Bei diesen Berechnungen war kein einziger Unterschied bei 10 % Irrtumswahrscheinlichkeit signifikant.

¹⁸ Es werden sinnvollerweise keine Interviewereigenschaften mit einbezogen. Was die Klumpung der Stichprobe angeht, dürften die Auswirkungen für den vorliegenden Sachverhalt vernachlässigbar sein.

Tabelle 11: Probit-Regression der Befragungsfähigkeit

<i>Variable</i>	<i>Koeffizient</i>	<i>Signifikanz</i>
Alter (Referenzkategorie.: 70 und älter)		
Unter 29	1.604	***
30-49	1.746	***
50-69	1.256	***
Konstante	1.009	***
Mc Fadden´s PseudoR ²		.252
Fallzahl		8597

* signifikant auf 10 % - Niveau, ** signifikant auf 5 % - Niveau, *** signifikant auf 1 % - Niveau

Quelle: SOEP (F) 2000, eigene Berechnungen.

Wie nicht weiter verwunderlich, kann durch bloßen Rückgriff auf Altersgruppendummies der Prozess sehr gut modelliert werden, da diese unabhängigen Variablen geeignete Prädiktoren für den Gesundheitszustand sind.

Die Altersdummies weisen gegenüber der Referenzkategorie „70 Jahre und älter“ enorm hohe Koeffizienten auf, da fast alle Befragungsunfähigen aus der Referenzkategorie stammen. Man mag einwenden, dass Alter keine Haushaltseigenschaft ist und eine Vermischung von Haushalts- und Personenebene vorliegt. Dies trifft zu, aber zu nicht befragungsfähigen Haushalten liegen leider keine anderen Informationen vor. Theoretisch ist ein Haushalt dann nicht befragungsfähig, wenn alle Haushaltsmitglieder nicht befragungsfähig sind. Deshalb kann man vermuten, dass hauptsächlich Einpersonenhaushalte zu dieser Kategorie gehören. Betrachtet man die Verteilung der Haushaltsgrößen für diese Haushalte (die Information ist mit 278 Haushalten für 79 % der 351 Fälle verfügbar), so sind 71 % Einpersonen- und 25 % Zweipersonenhaushalte. Der gute Fit des Modells erklärt sich womöglich daraus, dass das Alter des Haushaltsvorstands hoch mit dem Alter des Partners korreliert und somit auch die Zweipersonenhaushalte von dem Modell erfasst werden.¹⁹

5.2.4 Kooperationsbereitschaft

Analog zum Vorgehen für die Erreichbarkeit als abhängige Variable erfolgt auch die Modellierung der Kooperationsbereitschaft durch das Schätzen dreier Modelle, die sich der Adressprotokollvariablen und der Interviewereigenschaften bedienen. Die abhängige Variable

¹⁹ Die restlichen neun Fälle, in denen drei oder vier Personen einen befragungsunfähigen Haushalt bilden, stellen Kuriosa dar, vielleicht auch Vercodungsfehler oder Fälschungen.

ist mit eins codiert, wenn ein Haushalt – gegeben positiver Feldrücklauf, Erreichbarkeit und Befragungsfähigkeit – am SOEP teilnimmt. Die abhängige Variable ist null, wenn eine Teilnahme – ebenfalls gegeben positiver Feldrücklauf, Erreichbarkeit und Befragungsfähigkeit – verweigert wird.

In Modell 1 wird erneut ein Zusammenhang mit Entfernung zur nächsten Großstadt sichtbar. Am Stadtrand und in ländlichen Gebieten sind Haushalte kooperationsbereiter, wobei in mehr als 50 km Entfernung zur Großstadt der Koeffizient am größten ist. Ebenfalls kooperationsbereiter sind Haushalte in Ein-/Zweifamilienhäusern. Sozial höher eingestufte Haushalte sind eher für eine Teilnahme zu gewinnen. Der Koeffizient der Dummyvariable für obere Mittelschicht/Oberschicht ist minimal größer als derjenige für die Mittelschicht. Die Ergebnisse sind aber in Relation mit dem Modellfit zu sehen. Die Entropiereduktion ist äußerst klein, die Konstante ist nicht einmal auf 10 %-Niveau von null verschieden. Die Umgebungsvariablen beeinflussen folglich das Entscheidungsverhalten der Zielpersonen kaum. Die Intraklassenkorrelation zwischen den Interviewern ist mit .29 etwas kleiner als beim entsprechenden Modell zur Erreichbarkeit (.36). Das unbeobachtete Interviewerverhalten spielt also bei der Erreichbarkeit eine größere Rolle als bei der Kooperationsbereitschaft.

Modell 2 überprüft zusätzlich, ob das äußere Erscheinen des Interviewers im Zusammenhang mit der Teilnahmeentscheidung steht. Für die Variable Geschlecht trifft dies nicht zu. Frauen erzielen weder bessere noch schlechtere Kooperationsquoten.²⁰ Bei älteren Interviewern – von der Kategorie „40-49 Jahre“ an aufwärts – neigen die Zielhaushalte eher zur Teilnahme. Der Zusammenhang ist allerdings nicht monoton. Über 70-jährige Interviewer weisen einen geringeren Koeffizienten auf als ihre Kollegen im Alter zwischen 50 und 59 sowie zwischen 60 und 69. Die Zusammenhänge hinsichtlich der Umgebungsvariablen zeigen das gleiche Muster wie in Modell 1. Die erklärenden Variablen verbessern das Modell kaum. Weder Pseudo-R² noch der Anteil der Varianz auf Interviewerebene unterliegen substantiellen Veränderungen.

²⁰ Auch Interaktionsterme zwischen Geschlecht und Alter waren nicht signifikant. Für „pretty young female“- oder „little old lady“-Effekte gab es keine Hinweise. Allerdings konnten im vorliegenden Fall keine Interaktionen zwischen Geschlecht der Zielperson und des Interviewers berücksichtigt werden.

Tabelle 12: Multilevel-Probit-Regression der Kooperationsbereitschaft mit Adressprotokollinformationen und Interviewereigenschaften als erklärenden Variablen

Variablen	Modell 1		Modell 2		Modell 3	
	Koeffizient	Signifi- kanz	Koeffizient	Signifi- kanz	Koeffizient	Signifi- kanz
<i>Fixed effects</i>						
Stadtentfernung (Referenzkategorie: Stadtmitte Großstadt)						
Stadttrand	.143	***	.150	***	.162	***
Bis zu 25 km von Stadtmitte	.029		.048		.054	
25 – 50 km von Stadtmitte	.127	**	.122	**	.151	***
Über 50 km von Stadtmitte	.267	***	.264	***	.280	***
Gebäudetyp (Referenzkategorie: Höhere Gebäude)						
Ein- /Zweifamilienhäuser	.087	**	.084	*	.078	*
Mehrfamilienhäuser	.066		.069		.079	*
Reihenhäuser	.076		.075		.058	
Wohnumfeld (Referenzkategorie: Reines Wohngebiet)						
Hauptsächlich Wohngebiet	-.005		-.004		-.016	
Hauptsächlich Gewerbegebiet	.014		.002		-.010	
Ländliche Bebauung	.026		.037		.056	
Soziale Schicht (Referenzkategorie: Unterschicht)						
Untere Mittelschicht	.115	*	.128	**	.108	
Mittlere Mittelschicht	.164	***	.182	***	.151	**
Obere Mittelschicht/ Oberschicht	.174	**	.194	***	.161	**
Interviewervariablen						
Geschlecht weiblich			-.001		-.025	
Alter (Referenzkategorie: 30-39 Jahre)						
Bis 29 Jahre			-.204		-.314	
40 – 49 Jahre			.201	*	.151	
50 – 59 Jahre			.415	***	.407	***
60 – 69 Jahre			.432	***	.446	***
70 und älter			.373	***	.403	***
Dauer der Mitarbeit am SOEP (Referenzkategorie: bisher keine)						
1 – 3 Jahre					-.188	*
4 – 6 Jahre					-.054	
7 – 16 Jahre					-.053	
Beschäftigungsdauer im durchführenden Institut (Referenzkategorie: Eintritt im Erhebungsjahr)						
1 – 3 Jahre					-.032	
4 – 6 Jahre					-.053	
7 – 14 Jahre					-.143	
15 Jahre und mehr					-.104	

Workload				
(Referenzkategorie: 1-12 Adressen)				
13-24 Adressen				-046
25-36 Adressen				.141
37-48 Adressen				.118
49-66 Adressen				.143
über 66 Adressen				.483 **
Konstante	-.108 -		-.443 ***	-.351 **
<i>Random effects</i>				
Intercept	.634 ***		.608 ***	.593 ***
Intraklassenkorrelation	.287		.270	.260
McFadden's PseudoR ²	.013		.018	.046
Zahl der Cluster	577		564	546
Zahl der Fälle	10000		9956	9659

* signifikant auf 10 % - Niveau, ** signifikant auf 5 % - Niveau, *** signifikant auf 1 % - Niveau

Quelle: SOEP (F) 2000, eigene Berechnungen.

Die Erfahrungswerte sowie der Workload der Interviewer komplettieren den Satz der unabhängigen Variablen zu Modell 3. Während die Signifikanzniveaus der Stadtentfernungsvariablen stabil bleiben und die Koeffizienten größer werden, trifft dies für die Haustyp-Dummies nicht zu. Die Signifikanz des Dummies der unteren Mittelschicht verschwindet. Die Alterseffekte bezüglich der Interviewer (über 49-Jährige erzielen höhere Kooperationsbereitschaft) bleiben hochsignifikant. Keinen bzw. nur minimalen Einfluss haben die Erfahrungswerte. Nur Interviewer, die zwischen einem und drei Jahren am SOEP mitarbeiten, unterscheiden sich von den SOEP-„Neulingen“. Das Vorzeichen des Koeffizienten ist kontraintuitiv negativ, allerdings nur auf 10 %-Signifikanzniveau. Erwartungsgemäß spielt der Workload für die Befragungswilligkeit eine geringere Rolle als bei der Erreichbarkeit. Das Feldpersonal mit der größten Arbeitsbelastung ist auch am erfolgreichsten, wobei der Koeffizient der Kategorie „über 66 Zieladressen“ innerhalb des Modells betragsmäßig am größten ist. Modell 3 erreicht nur eine Entropiereduktion von ca. .05. Auch hier verkleinern die expliziten Interviewermerkmale den Anteil der Varianz auf der Interviewerebene kaum.²¹ Individuelles, unbeobachtbares Interviewerverhalten spielt eine größere Rolle als äußere Merkmale des Interviewers.

²¹ In einer hier nicht aufgeführten Schätzung wurden noch die Berufstätigkeit und die Schulbildung des Interviewers zusätzlich miteinbezogen. Es gab allerdings keinerlei Anzeichen, dass sich die Eigenschaften auf die Leistungen der Interviewer auswirken.

6 Zusammenfassung und Schlussfolgerungen

Um mögliche Verzerrungen zu erkennen, wurde versucht, den Teilnahmeprozess für die neue Stichprobe F des SOEP bestmöglich zu modellieren. Zu diesem Zweck wurden Mehrebenenmodelle für die Erreichbarkeit, Kontaktzahl, Befragungsfähigkeit, und Kooperationsbereitschaft geschätzt. Die wesentlichen Ergebnisse werden im folgenden kurz zusammengefasst.

Haushalte am Stadtrand und auf dem Land waren besser erreichbar als solche in der Stadtmitte einer Großstadt. Dasselbe gilt für Zieladressen in kleineren Gebäuden und reinen Wohngebieten. Diese Resultate dürften unter anderem auf unterschiedliche Haushaltsstrukturen (mehr Einpersonenhaushalte in der Großstadt und in größeren Wohneinheiten) und verschiedene Tagesrhythmen der Bewohner zurückzuführen sein. Keinen Einfluss hatte der sozio-ökonomische Status des Zielhaushalts auf die Erreichbarkeit. Darauf, dass auf Grund von Viktimisierungängsten die Tür nicht geöffnet wurde, gab es keine Hinweise, da Alter und Geschlecht des Interviewers nicht in Zusammenhang mit der Erreichbarkeit standen. Stattdessen spielte auf Interviewerebene die Erfahrung eine Rolle. Interviewer, die bereits zwischen einem und drei Jahren für das betreffende Institut gearbeitet hatten, erzielten höhere Kontaktraten als ihre frisch eingestellten Kollegen. Dies bestätigt die Erwartungen, dass die Interviewer ihre Kontaktstrategie bereits zu Beginn ihrer Tätigkeit optimieren. Außerdem erhöhte die Tatsache, bereits zwischen vier und sechs Jahren am SOEP mitzuarbeiten, die Erreichbarkeitsquoten des Feldpersonals. Die Arbeitsbelastung der Interviewer, operationalisiert über die Zahl der zugewiesenen Zieladressen, führte nicht monoton zu schlechteren Kontaktierungsleistungen. Nur Interviewer mit 49-66 Adressen erreichten weniger Haushalte. Vermutlich wirkt neben Selektionseffekten (als erfolgreich bekannten Interviewern werden mehr Adressen zugewiesen) auch die Motivation hier kompensierend. Zwar liegt der Großteil der Varianz auf der Befragtenebene, doch der Anteil auf der Interviewerebene ist mit ungefähr einem Drittel deutlich erkennbar. Das Modell bildet den zugrundeliegenden Prozess nur in mäßiger Qualität ab, die Entropiereduktion ist gering. Demnach können die verwendeten Variablen die relevanten Konzepte, nämlich den Rhythmus, nach dem Haushaltsmitglieder in der Wohnung aufzufinden sind, und die Kontaktstrategie der Interviewer nicht zufriedenstellend abbilden. Dies bedeutet allerdings auch, dass die Verzerrung der Stichprobe auf Grund systematischer Unterschiede der Erreichbarkeit von Haushalten gering ist.

Modelliert man die Kontaktzahl eines Haushalts auf Basis der Nettostichprobe, werden die aus der Literatur bekannten Hypothesen weitestgehend bestätigt. Adressen mit älteren oder nichtdeutschen Haushaltsvorständen und Kindern benötigen weniger Kontakte. Eine höhere Kontaktzahl ist bei berufstätigen und geschiedenen Haushaltsvorständen nachzuweisen, ebenso gibt es Bildungseffekte. Einpersonenhaushalte haben – auf den ersten Blick überraschend – weniger Kontakte, was auf Grund der Kontrolle des Alterseffekts aber durchaus plausibel sein kann. Die Interviewer unterscheiden sich hinsichtlich der Zahl der Kontakte weniger stark als hinsichtlich der Erreichbarkeit der Haushalte allgemein. Dies deutet darauf hin, dass sie sich gerade auch hinsichtlich wichtiger Faktoren wie etwa den Zeitpunkten und den Zeitintervallen der Kontakte unterscheiden.

Analysen, ob sich die Spätkontaktierten für ausgewählte Variablen von den Frühkontaktierten unterscheiden, ergaben zwar deutliche univariate Unterschiede in Bezug auf einige soziodemographische Variablen, Sorgen- und Zufriedenheitsangaben, wenn zwischen nach der Kontaktzahl getrennten Subgruppen der Stichprobe unterschieden wurde. Nach Kontrolle der soziodemographischen Variablen in multivariaten Modellen war der Einfluss der Kontaktzahl allerdings nur noch in einigen Fällen signifikant. Die Ergebnisse aus zwei hypothetischen Stichproben, von denen eine nur Fälle mit bis zu sechs Kontakten und die andere darüber hinaus auch Fälle mit mehr als sechs Kontakten umfasste, unterschieden sich bezüglich ausgewählten soziodemographischen Variablen, Sorgen- und Zufriedenheitsangaben kaum. Mit dem politischen Interesse war nur ein signifikanter Unterschied zwischen den Stichproben festzustellen. Allerdings ist bei der Interpretation des signifikanten Effekts auch die große Fallzahl zu bedenken, bei der schon kleinste Unterschiede signifikant werden. Außerdem war bei Beschränkung auf die Haushaltsvorstände der Unterschied nicht vorhanden. Hier zeigt sich, wie wichtig es ist, dass das SOEP ein Survey ist, der mit einem hohen Aufwand, manifestiert z.B. im adress-random-Verfahren, einer hohen Kontaktzahl und einer langen Feldzeit, erhoben wird. Solche Designspezifika sind für die Qualität einer Stichprobe enorm wichtig. Somit sollte nicht mehr gesagt werden, als dass *im vorliegenden Fall* keine großen Zusammenhänge zwischen Erreichbarkeit und inhaltlichen Variablen nachgewiesen werden konnten.

Die Befragungsfähigkeit kann am besten beschrieben werden. Da nur gesundheitliche Gründe eine Rolle spielen, bietet eine Approximation über Altersvariablen ein aussagekräftiges Modell.

Hinsichtlich der Kooperationsbereitschaft zeigt sich folgendes: Haushalte am Stadtrand und auf dem Land (über 25 km zur nächsten Großstadt) nehmen eher an der Befragung teil. Auf Basis der Variablen aus Interviewerschätzungen gibt es Hinweise auf einen Mittelstandsbias. Sowohl die mittlere Mittelschicht als auch die obere Mittelschicht/Oberschicht nimmt eher am SOEP teil. Ob dies auf einen reinen Bildungseffekt oder sonstige Gründe zurückzuführen ist, kann nicht näher untersucht werden. Auf Seite der Interviewer spielt im Gegensatz zum Geschlecht das Alter eine Rolle. Ältere Interviewer rufen größere Kooperationsbereitschaft hervor. Vermutlich droht bei ihnen weniger die Gefahr, während der „doorstep interaction“ ein falsches Skript auszulösen, oder sie bewegen die Zielpersonen eher dazu, ihnen „einen Gefallen zu tun“. Bemerkenswerterweise waren keine Erfahrungseffekte feststellbar. Entweder unterliegt das Konzept des „tailoring“ bei der Kontaktaufnahme für ein Interview keinem Lernprozess oder er wird durch gegenläufige Effekten wie etwa lange zurückliegende Schulungsmaßnahmen, Motivationsprobleme der Interviewer o.ä. aufgehoben. Die Arbeitsbelastung hängt nur in einer Subgruppe des Feldpersonals mit den Kooperationsquoten zusammen. Die Interviewer mit dem stärksten Workload übertreffen ihre Kollegen deutlich; dies ist als Selektionseffekt anzusehen. Es ist bekannt, dass das Befragungsinstitut den als erfolgreich bekannten Interviewern am meisten Adressen übergibt. Im Vergleich mit der Erreichbarkeit haben die Interviewer zwar etwas weniger Einfluss auf die Kooperationsbereitschaft, der Anteil der Varianz auf Interviewerebene ist mit etwas weniger als 30 % immer noch sehr hoch.

Gerade die Modellierung der Kooperationsbereitschaft kann als gescheitert angesehen werden. Die zur Verfügung stehenden Variablen können die Teilnahmeentscheidung nicht angemessen erklären. Wie in der Einleitung bereits erläutert, war dies auch nicht anders zu erwarten, da die vorhandenen Informationen nur bedingt mit den Kosten-Nutzen-Abwägungen der Zielperson in Verbindung stehen. Der schlechte Fit der Modelle ist freilich in dem Sinne erfreulich, dass er keine Hinweise auf schwerwiegende systematische Ausfälle bei der Stichprobe F des SOEP gibt. Vermutlich sind größtenteils situationale Effekte für die Ausfälle verantwortlich (vgl. Schnell 1997: 213-216).

Literatur

- Anders, M. 1985: Sinkende Ausschöpfungsraten – und was man dagegen tun kann.
In: Kaase, M., Kuchler, M. (Hrsg.): Herausforderungen der Empirischen Sozialforschung.
Mannheim: ZUMA: S. 75-80.
- Blohm, M., Diehl, C. 2001: Wenn Migranten Migranten befragen. Zum Teilnahmeverhalten von Einwanderern bei Bevölkerungsumfragen.
Zeitschrift für Soziologie 30: 223-242.
- Brehm, J. 1993: The phantom respondents. Opinion surveys and political representation.
Ann Arbor: The University of Michigan Press.
- Curtin, R., Presser, S., Singer, E. 2000: The effects of response rate changes on the Index of Consumer Sentiment.
Public Opinion Quarterly 64: 413-428.
- De Heer, W. 1999: International response trends: Results of an international survey.
Journal of Official Statistics 15(2): 129-142.
- De Leeuw, E., de Heer, W. 2002: Trends in household survey nonresponse: A longitudinal and international comparison.
In: Groves, R.M., Dillman, D.A., Eltinge, J.L., Little, R.J.A. (Eds.):
Survey Nonresponse. New York: Wiley. S. 41-54.
- Engel, U. 1998: Einführung in die Mehrebenenanalyse.
Opladen/Wiesbaden: Westdeutscher Verlag.
- Esser, H., Grohmann, H., Müller, W. und Schäffer, K.A. 1989: Mikrozensus im Wandel.
Stuttgart: Metzler-Poeschel.
- Goldstein, H. 1995: Multilevel statistical models. Second edition.
London: Arnold.

Goldstein, H., Rasbash, J., Plewis, J., Draper, D., Browne, W., Yang, M., Woodhouse, G., Healy, M. 1998:

A user's guide to MLwiN.

London: Institute of Education.

Goyder, J. 1987: The silent minority. Nonrespondents on sample surveys.

Cambridge: Polity Press.

Goyder, J., Warriner, K., Coutts, S., Basic, M. 1999: Socio-economic bias in surveys: Some attempts at interpretation.

Paper presented at the International Conference on Survey Nonresponse, October 28-31, 1999, Portland (Oregon), USA.

Groves, R. M. 1989: Survey errors and survey costs.

New York: Wiley.

Groves, R. M., Couper, M. 1998: Nonresponse in household interview surveys.

New York: Wiley.

Groves, R. M., Singer, E., Corning, A. 2000: Leverage-saliency theory of survey participation. Description and illustration.

Public Opinion Quarterly 64: 299-308.

Hedeker, D., Gibbons, R. D. 1996: MIXOR: a computer program for mixed-effects ordinal regression analysis.

Computer Methods and Programs in Biomedicine 49: 157-176.

Heller, G., Schnell, R. 2000: The choir invisible. Zur Analyse der gesundheitsbezogenen Panelmortalität im Sozio-Ökonomischen Panel (SOEP).

in: Helmert, U., Bammann, K., Voges, W., Müller, R. (Hrsg.): Müssen Arme früher sterben? Soziale Ungleichheit und Gesundheit in Deutschland. Weinheim: Juventa. S. 115-134.

Hox, J., de Leeuw, E. 2002: The influence of interviewers' attitude and behavior on household survey nonresponse: An international comparison.

In: Groves, R.M., Dillman, D.A., Eltinge, J.L., Little, R.J.A. (Eds.):

Survey Nonresponse. New York: Wiley. S.103-120.

Japac, L., Lundqvist, P. 1999: Interviewer strategies and attitudes.

Paper presented at the International Conference on Survey Nonresponse, October 28-31, 1999, Portland (Oregon), USA.

Kirschner, H.P. 1984: ALLBUS 1980: Stichprobenplan und Gewichtung.

In: Mayer, U./Schmidt, P. (Hrsg.):

Allgemeine Bevölkerungsumfrage der Sozialwissenschaften. Frankfurt. S. 114- 182.

Keeter, S., Miller, C., Kohut, A., Groves, R. M., Presser, S. 2000: Consequences of reducing nonresponse in a national telephone survey.

Public Opinion Quarterly 64: 125-148.

Kennickell, A. B. 2000: What do the „late“ cases tell us ? Evidence from the 1998 Survey of Consumer Finances.

Paper presented at the International Conference on Survey Nonresponse, October 28-31, 1999, Portland (Oregon), USA. Revised version February 2000.

Kreft, I., de Leeuw, J. 1998: Introducing multilevel modeling.

London: Sage.

Koch, A. 1991: Zum Zusammenhang von Interviewermerkmalen und Ausschöpfungsquoten.

ZUMA-Nachrichten 28: 41-53.

Koch, A. 1997: Teilnahmeverhalten beim ALLBUS 1994.

Kölner Zeitschrift für Soziologie und Sozialpsychologie 49: 98-122.

Koch, A., Kurz, K., Mahr-George, H., Wasmer, M. 1998a: Konzeption und Durchführung der „Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften“ (ALLBUS) 1998.

ZUMA-Arbeitsbericht 99/02. Mannheim: ZUMA.

Koch, A. 1998b: Wenn „mehr“ nicht gleichbedeutend mit „besser“ ist: Ausschöpfungsquoten und Stichprobenverzerrungen in allgemeinen Bevölkerungsumfragen.

ZUMA-Nachrichten 42: 66-90.

Koch A., Wasmer, M., Harkness, J., Scholz, E. 2001: Konzeption und Durchführung der „Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften“ (ALLBUS) 2000.

ZUMA-Arbeitsbericht 2001/05. Mannheim: ZUMA.

Lievesley, D. 1986: Unit nonresponse in interview surveys.

Joint centre for Survey Methods unpublished paper.

London: SCPR.

Morton-Williams, J. 1993: Interviewer Approaches.

Aldershot: Dartmouth.

Pannenberg, M., Pischner, R., Rendtel, U., Spiess, M., Wagner, G. G. 2000: Sampling and weighting.

In: Haisken-DeNew, J. P., Frick, J. R. 2000: Desktop Companion to the German Socio-Economic-Panel Study. Version 4.0. Berlin: DIW, 125-155.

Purdon, S., Campanelli, P., Sturgis, P. 1999: Interviewers' calling strategies on face-to-face interview surveys.

Journal of Official Statistics 15(2): 199-216.

Rodríguez, G., Goldman, N. 2001: Improved estimation for multilevel models with binary response: a case-study.

Journal of the Royal Statistical Society A 164: 339-355.

Rosenblatt, B. v. 2001a: Erprobung innovativer Erhebungskonzepte für Haushalts-Panel-Stichproben. Erstbefragung 2000 der SOEP-Stichprobe F. Methodenbericht.

München: Infratest Sozialforschung.

Rosenblatt, B. v. 2001b: Erprobung innovativer Erhebungskonzepte für Haushalts-Panel-Stichproben. Zweitbefragung 2001 der SOEP-Stichprobe F. Methodenbericht.

München: Infratest Sozialforschung.

Schnell, R. 1993: Homogenität sozialer Kategorien als Voraussetzungen für „Repräsentativität“ und Gewichtungsverfahren.

Zeitschrift für Soziologie 22: 16-32.

Schnell, R. 1997: Nonresponse in Bevölkerungsumfragen.

Opladen: Leske und Budrich.

Schnell, R. 1998: Besuchs- und Berichtsverhalten der Interviewer.

In: Statistisches Bundesamt (Hrsg.): Interviewereinsatz und –qualifikation.

Stuttgart: Metzler und Poeschel.

Schnell, R., Hill P.B., Esser, E. 1999: Methoden der empirischen Sozialforschung.

München: Oldenbourg.

Schräpler, J.P. 1999: Das Befragtenverhalten im Sozio-oekonomischen Panel. Analysen an ausgewählten Beispielen.

Dissertation. Ruhr-Universität Bochum.

Schräpler, J.P. 2000: Was kann man am Beispiel des SOEP bezüglich Nonresponse lernen ?

ZUMA-Nachrichten 46: 117-149.

Snijkers, G., Hoop, J., de Leeuw, E. 1999: Interviewers' tactics for fighting survey nonresponse.

Journal of Official Statistics 15(2): 185-198.

Stoop, I. A. L., Louwen, F. J. 1999: From a silent majority to an articulate minority.

Paper presented at the International Conference on Survey Nonresponse, October 28-31, 1999, Portland (Oregon), USA.

Wasmer, A., Koch, A., Harkness, J., Gabler, S. 1996: Konzeption und Durchführung der „Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften“ (ALLBUS) 1996.

ZUMA-Arbeitsbericht 96/08. Mannheim: ZUMA.