



Diskussionspapiere Discussion Papers

Diskussionspapier Nr. 27

**Kennziffern zur Charakterisierung der Stabilität
von Panel-Erhebungen - Das Beispiel des SOEP**

von Gert Wagner

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

Opinions expressed in this paper are those of the author and do not necessarily reflect views of the Institute.

Diskussionspapier Nr. 27

**Kennziffern zur Charakterisierung der Stabilität
von Panel-Erhebungen - Das Beispiel des SOEP**

von Gert Wagner

Berlin, Juni 1991

Kennziffern zur Charakterisierung der Stabilität von Panel-Erhebungen - Das Beispiel des SOEP

von Gert Wagner

0 Problemaufriß

Panel-Erhebungen messen nicht nur dynamische Entwicklungen, sondern die Stichproben derartiger Wiederholungsbefragungen sind selbst erheblichen Dynamiken unterworfen. Im allgemeinen wird es Panel-Mortalität geben, d.h. daß ein Teil der Befragten des Vorjahres nicht mehr an der Folgerhebung teilnimmt. Über mehrere Wellen hinweg kann der kumulierte "Ausfall" ein erhebliches Ausmaß annehmen. Letztendlich verbirgt sich hinter dieser Entwicklung die Frage, ob bzw. inwieweit die Panel-Stichprobe (noch) für die jeweilige Grundgesamtheit repräsentativ ist.

Diese Frage der Repräsentativität kann genau genommen nur durch möglichst umfangreiche Vergleiche der Panel-Stichprobe mit der Grundgesamtheit bzw. anderen Stichproben aus dieser Grundgesamtheit überprüft werden. Derartige Vergleiche sind mehr oder minder aufwendig, insbesondere dann, wenn sie mit Hilfe von echten Pocket Samples durchgeführt werden sollen, d.h. Stichproben, die zur selben Zeit wie die Panel-Stichprobe gezogen wurden, aber zwischenzeitlich nicht bearbeitet (befragt) wurden¹.

Hilfsweise wird zur Kennzeichnung der Vertrauenswürdigkeit eines Panels die Panel-Stabilität als Anteil der verbliebenen Panelpopulation an der ursprünglichen Stichprobenpopulation gekennzeichnet. Diese "Teilnahme-" bzw. "Stabilitätsquote"

1 Da für eine spezifische Längsschnitts-Erhebung in der Regel keine vergleichbare (amtliche) Datenbasis vorliegt, anhand derer die Repräsentativität aller erhobener Längsschnittsvariablen geprüft werden könnte, sind zur Prüfung von Längsschnitten eigene Kontrollstichproben notwendig, die im Idealfall zur selben Zeit gezogen wurden, wie die 1. Welle des Panels. Derartige "Pocket Sampler" erlauben es z.B. zu testen, ob die Befragten eines Haushaltspanels aufgrund der wiederholten Befragung Verhaltensänderungen aufweisen ("Panel"-Effekt), obwohl sie - gemessen an den üblichen sozialstrukturellen Merkmalen - noch für die Grundgesamtheit repräsentativ sind.

ist auf den ersten Blick ein klares Konzept, daß sich jedoch bei näherem Hinschauen im Detail auch als diskussionswürdig erweist.

Im folgenden wird herausgearbeitet, daß die einfache Berechnung "verbliebene Stichprobengröße im Verhältnis zur Ursprungsstichprobe" die Panel-Stabilität im Hinblick auf analytische Zwecke **unterschätzt**. Am Beispiel des Sozio-oekonomischen Panels (SOEP) (vgl. dazu Projektgruppe Panel 1990) wird gezeigt, daß nur der erhebungsbedingte Ausfall als Panel-Mortalität bei Analysen Schwierigkeiten bereitet.

Bei (internationalen) Vergleichen von Haushaltspanels² sollten nur "demographisch bereinigte" Stabilitätskennziffern benutzt werden.

1 Die 1. Welle der SOEP-Stichprobe

Im Jahre 1984 wurden mit der ersten Welle des SOEP 5921 Haushalte erfolgreich befragt³, in denen 16205 Personen lebten. Von diesen wurden 12245 16jährige und ältere interviewt (45 Personen im Befragungsalter konnten in der ersten Welle nicht befragt werden)⁴.

Üblicherweise wird von den befragten Personen der ersten Welle ausgegangen, um die Panel-Stabilität zu messen (vgl. Rendtel 1990, S. 284 sowie auch Esser et al. 1989, S. 103ff.). In Bezug auf die Repräsentativität der Stichprobe ist allerdings auch eine andere Rechnung möglich:

Ausgangspunkt können die Brutto-Adressen sein, die mit der ersten Welle befragt werden sollten. Da es - wie seit einigen Jahren international üblich - bereits bei der ersten Welle massive Teilnahmeverweigerungen gab, stellen die Interviews der ersten Welle des SOEP bereits nur 56 % der eingesetzten

2 "Haushaltspanels" repräsentieren die Grundgesamtheit der privaten Haushalte eines Gebietes. In neueren Panelerhebungen dieser Art werden in den Haushalten alle erwachsenen Personen befragt, zusätzlich werden "Proxy-Informationen" für die Kinder durch eine gesonderte Befragung des "Haushaltsvorstandes" erhoben.

3 Diese Zahl ist um 48 Haushalte bereinigt, für die sich nach der 2. Welle herausstellte, daß die Angaben der 1. Welle von den Interviewern gefälscht wurden.

4 Auf die im Jahre 1990 gezogene Ost-Stichprobe des SOEP (Basisbefragung '90 in der DDR) wird im folgenden nicht eingegangen, da die zweite Welle noch nicht abgeschlossen ist (vgl. Projektgruppe Panel 1991).

Adressen dar. Schon nach der zweiten Welle waren weniger als 50 % der ursprünglich ausgewählten Adressen in der Stichprobe.

Diese Rechnung berücksichtigt allerdings nicht, daß es sich bei diesen Ausfällen z.T. um "qualitätsneutrale Ausfälle" handelt (z.B. Adressen von Geschäftshaushalten, die nicht zur Grundgesamtheit zählen). Von darüber hinausgehenden Antwortverweigerungen sind weltweit alle Erhebungen betroffen, die auf freiwilliger Grundlage arbeiten (also ohne Teilnahmepflicht) (vgl. auch Esser et al. 1989, S. 104). Für die erste Welle des SOEP konnte anhand von Vergleichen mit bekannten Randverteilungen sowie anhand von Kurzbefragungen der Haushalte, die die Teilnahme am Panel selbst verweigert haben, gezeigt werden, daß die Repräsentativität der Ausgangsstichprobe hinreichend ist. Vorhandene Abweichungen konnten weitgehend anhand einer gebundenen Hochrechnung beseitigt werden (vgl. Rendtel 1991).

Auf jeden Fall sind die Erhebungsprobleme einer ersten Welle nicht panel-spezifisch. Eine Stabilitäts-Kennziffer, die von den gezogenen Brutto-Adressen ausgeht, ist nicht aussagekräftig.

Werden 50 % der ursprünglichen Stichprobengröße bei einem Panel unterschritten, hat dies gleichwohl eine "magische" Signalwirkung, obwohl keinesfalls gesagt werden kann, daß durch die Unterschreitung der 50 %-Marke die Repräsentativität des Samples automatisch gefährdet ist. Dies gilt überhaupt nicht, wenn es sich bei den Ausfällen um reine Zufalls-Ausfälle handelt, die die Strukturen in der Stichprobe nicht verzerren (vgl. Abschnitt 3 unten).

2 Die Stabilität der SOEP-Stichprobe

Ebenso wie bei der Ursprungsstichprobe sind auch Teilnahmeverweigerungen und die Ausfälle während der Laufzeit eines Panels sicherlich nicht rein zufällig; sie sind aber auch nicht streng systematisch.

Akzeptiert man den Grad der Repräsentativität der SOEP-Ausgangsstichprobe (der in einem mindestens ebensolchen Ausmaße

gegeben war, wie dies für übliche Querschnittbefragungen gilt und hingenommen wird), so ergibt sich folgendes Bild der Panel-Stabilität. Dabei wird gemessen, wieviel Prozent der ursprünglich Befragten nach jeder Welle noch im Sample enthalten sind (vgl. Abb. 1). Nach sieben Wellen sind nur noch 61 % der ursprünglich Befragten in der Stichprobe.

Diese bislang übliche Rechnung erkennt allerdings zwei Tatsachen:

- Nicht alle Ausfälle bedeuten ein Analyseproblem; einige Ausfallarten sind im Gegenteil selbst analysierbar und stellen keinen Informationsverlust bezüglich der verbleibenden Population dar.
- Zum einen sind nicht nur über die Befragten der ersten Welle Informationen vorhanden, sondern auch über die damals 15jährigen und jüngeren Haushaltsmitglieder. Für diese gibt es beim SOEP Informationen aus dem Haushaltsfragebogen und dem sogenannten Adressenprotokoll ("Proxy-Interviews");

Führt man dieselbe Rechnung wie in Abb. 1 für alle Haushaltsmitglieder der ersten Welle durch ("Stammpersonen"), so ist der Verlust etwa einen Prozentpunkt geringer und die Gesamtstabilität beträgt bis zur 7. Welle 62 %. Hier wird berücksichtigt, daß ja auch für ursprünglich nicht selbst befragte Kinder auswertbare Längsschnittsinformationen vorliegen. Die relevante Basis zur Berechnung der Stichproben-Stabilität sollten also die "Stammpersonen" der 1. Welle sein.

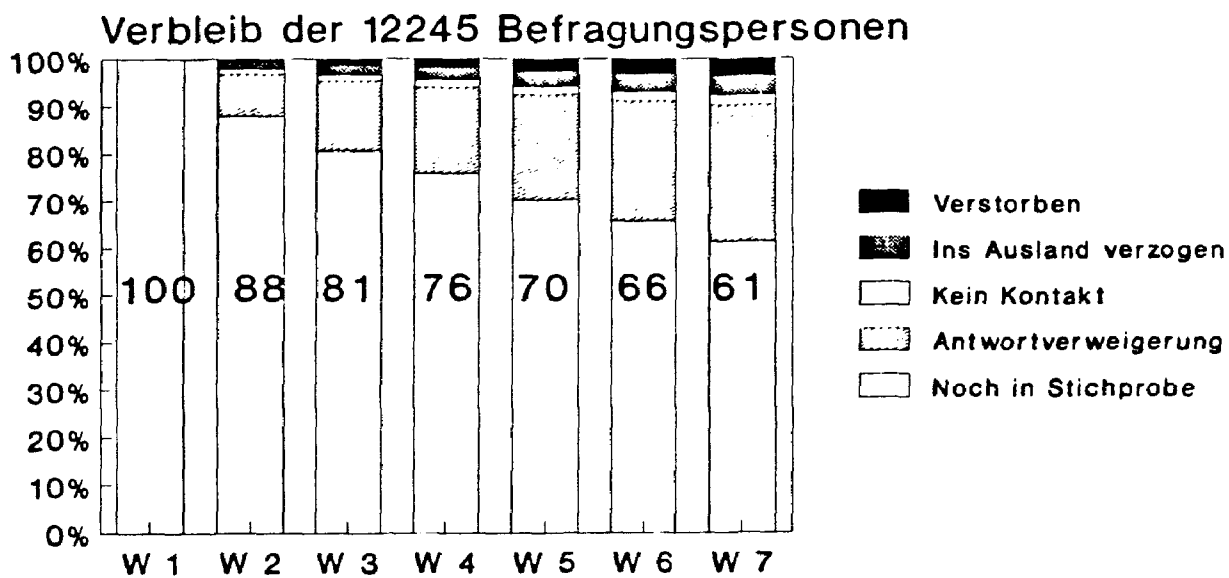
Betrachtet man nun die Ausfallarten, so ergeben sich folgende Unterscheidungen:

- Ausfall wegen Tod (natürliche Mortalität, die die Veränderung in der Grundgesamtheit nachbildet),

Abbildung 1

Entwicklung des Stichprobenbestands

Basis : Alle Befragungspersonen der 1. Welle des SOEP



Datenbasis : Das Sozio-ökonomische Panel
Welle 1 bis Welle 7

- Ausfall wegen Umzug ins Ausland (derartige Personen und Haushalte werden nicht "weiterverfolgt", da die Grundgesamtheit regional eingegrenzt ist auf das Gebiet der Bundesrepublik),
- Ausfälle wegen Nicht-Erreichbarkeit,
- Ausfälle wegen Teilnahmeverweigerungen und
- temporäre Ausfälle aufgrund vorübergehender Nicht-Erreichbarkeit bzw. Interviewbereitschaft.

Ausfälle aufgrund des Todes von Befragten sind kein Informationsverlust, sondern sie sind sogar notwendig, um die Querschnitts-Repräsentativität des Samples zu gewährleisten (wenn - wie dies beim SOEP der Fall ist - auf der anderen Seite die in Starthaushalte ziehende Personen ebenfalls in die Untersuchung einbezogen werden und splitt-offs von Starthaushalten weiterverfolgt werden).

Esser et al. (1989, S. 141) sprechen deswegen von der "Zielstichprobe"⁵, die von einem Panel erreicht werden soll. Wenn diese nur die bundesdeutsche Wohnbevölkerung darstellt, so stellen Umzüge ins Ausland ebenfalls keinen Informationsverlust dar. Der Umzugsprozeß selbst ist mit Hilfe der zuvor erhobenen Daten analysier- und erklärbar (vgl. z.B. Brecht und Michels 1991).

Im Hinblick auf die Querschnittsrepräsentativität eines Panels macht es keinen Sinn, die demographisch bedingten "natürlichen" Ausfälle als Panel-Mortalität zu zählen. Ebenfalls nicht als Panelmortalität können die temporären Ausfälle gezählt werden, da diese nicht nur in späteren Wellen weiter befragt werden, sondern beim SOEP auch für bestimmte zentrale Informationen retrospektive Daten nachbefragt werden.

Wirkliche Informationen gehen nur durch Nicht-Erreichbarkeit und Teilnahmeverweigerungen verloren. Dieser Anteil mißt auch

5 Auch für die seit langem laufende US-amerikanische PSID wird mit der um Todesfälle bereinigten Stabilität argumentiert, vgl. z.B. den PSID-Antrag an die NSF für die Wellen 25 bis 29 (S. 87) Abwanderungen dürften in den USA eine nur geringe Rolle spielen.

den Erfolg der Feldarbeit des Umfrageinstituts.

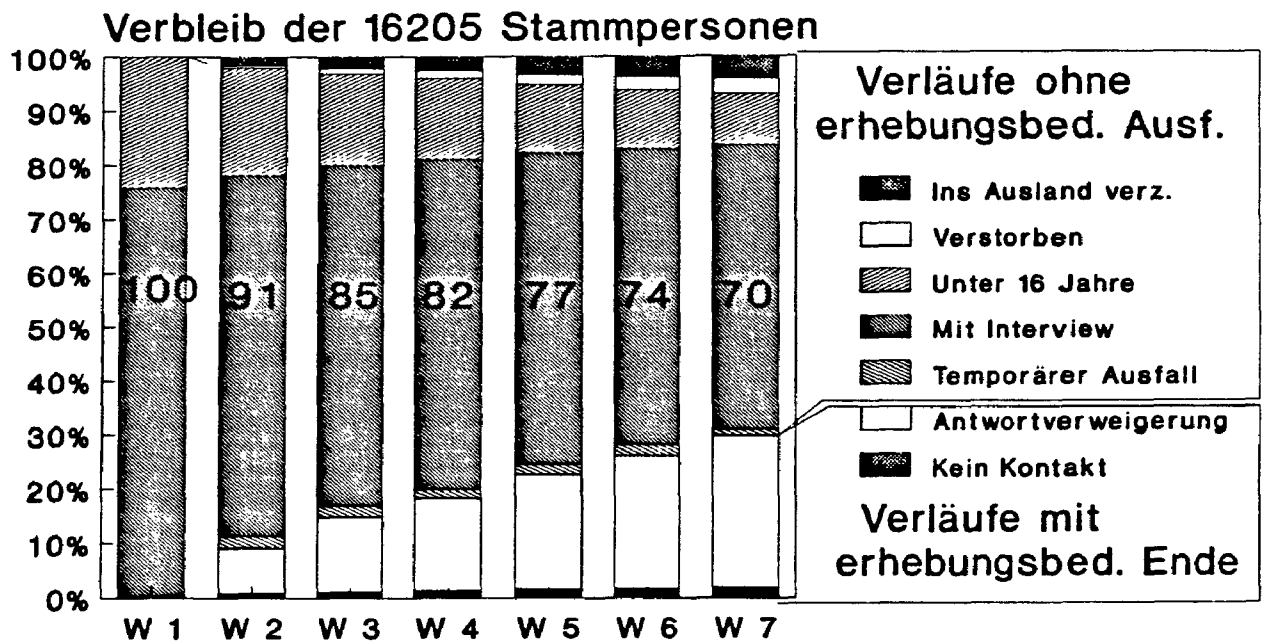
Stellt man in Abbildung 2 die Stabilität der Stichprobe nur anhand der echten Ausfälle (Nicht-Erreichbarkeit und Teilnahmeverweigerung) dar, so ergibt sich mit 70 % ein weit höheres Bild der Stabilität als in Abbildung 1 mit 61 %. Der Unterschied ist deswegen so groß, da im SOEP Haushalte mit ausländischem Haushaltsvorstand überrepräsentiert sind, die italienischer, jugoslawischer, griechischer, spanischer oder türkischer Nationalität sind. Es handelt sich um typische Arbeitsmigranten, deren Rückwanderungsraten relativ hoch sind (vgl. Schulz 1991). Die Abbildungen 3.1 und 3.2 machen die Unterschiede zwischen der Stichprobe A (Deutsche Haushaltsvorstände) und Stichprobe B (Arbeitsmigranten-Haushalte) besonders deutlich.

Eine andere Frage ist die nach der möglichen Beeinträchtigung von Längsschnittanalysen durch Panel-Mortalität. Auch hier bereitet "natürliche" Mortalität durch Tod und Umzug ins Ausland keine Analyseprobleme (im Sinne von Verzerrungen), sofern durch natürliche Mortalität und Umzüge ins Ausland nicht die Fallzahl der Stichprobe derart gemindert wird, daß aus statistischen Gründen (mangelnde Signifikanz der Ergebnisse) keine Analysen mehr sinnvoll sind.

Bei der Beurteilung der Analysemöglichkeiten eines Panels muß darüber hinaus beachtet werden, daß es keineswegs notwendig ist, nur Längsschnittanalysen durchzuführen, in die lediglich Fälle eingehen, die über den gesamten Erhebungszeitraum valide Antworten gegeben haben (complete case analysis). Es gibt verschiedene (multivariate) Methoden, die es erlauben, auch "unvollständige" Beobachtungen von Zuständen und Ereignissen zu analysieren (complete information analysis). Es handelt sich zum einen um die Methode der Ereignisanalyse (vgl. dazu z.B. Brecht und Michels 1991), zum anderen um konventionelle Regressionsmodelle, bei denen durch das "Poolen" von Beobachtungen und eine geeignete Spezifikation der Störterme auch nicht-rechteckige Datenmengen von Befragungseinheiten (Personen, Haushalte) verarbeitet werden können (vgl. zum "Varianzkomponentenmodell" z.B. Löwenbein und Rendtel 1991).

Abbildung 2

Entwicklung Stichprobenbestands Basis : Alle Stammpersonen der 1. Welle



Datenbasis : Das Sozio-ökonomische Panel
Welle 1 bis Welle 7

Abbildung 3.1

Entwicklg. des Bestands von Stichprobe A Basis : Alle Stammpersonen der 1. Welle

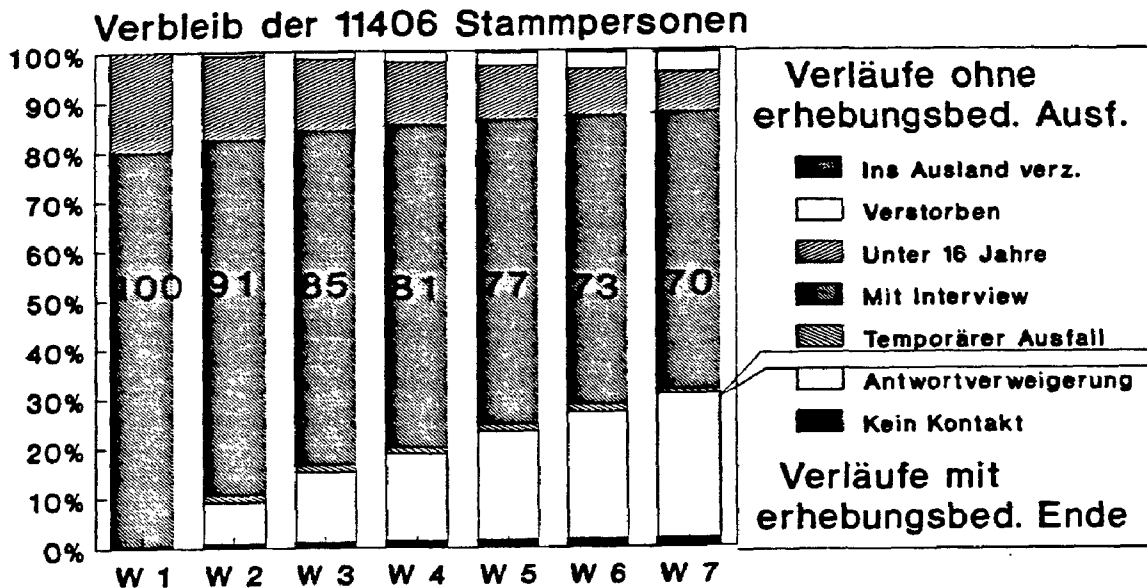


Abbildung 3.2

Entwicklg. des Bestands von Stichprobe B Basis : Alle Stammpersonen der 1. Welle

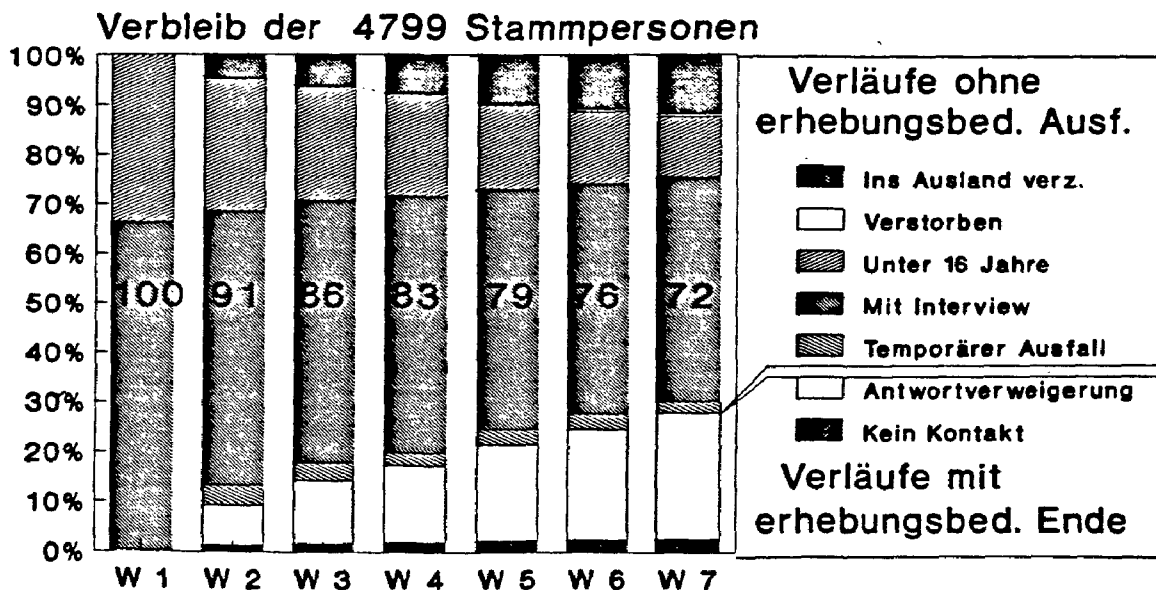


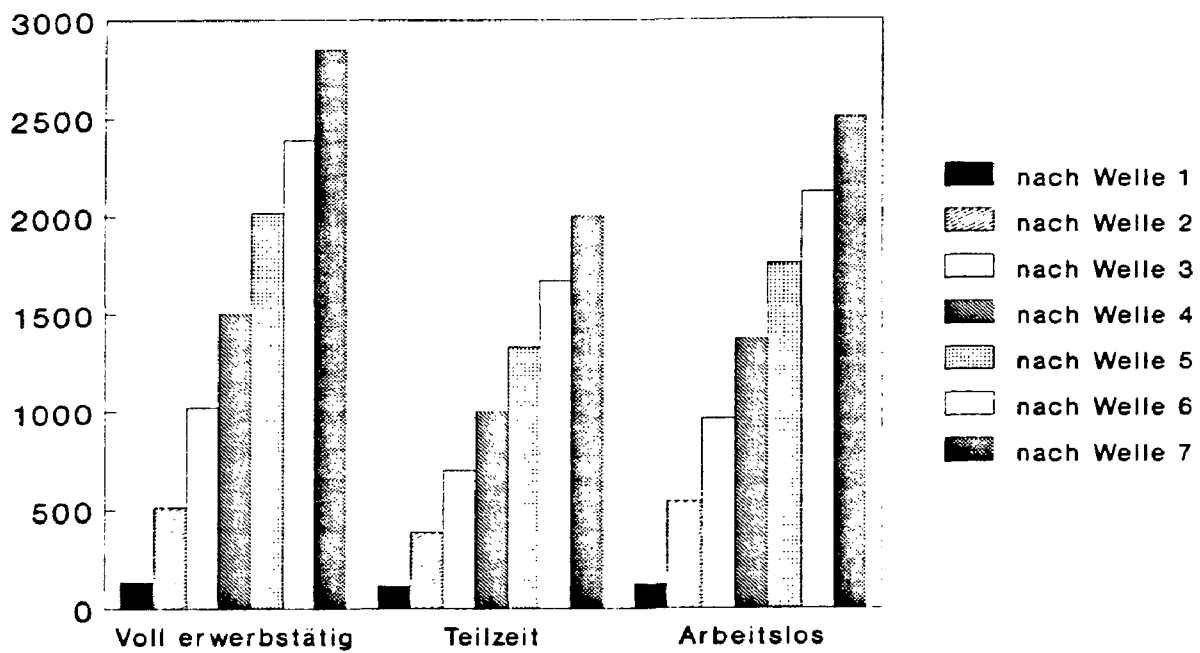
Abbildung 4 zeigt, daß die Zahl von analysierbaren Ereignissen bzw. Zuständen (unzensurierte Spells, bei denen Beginn und Ende eines Zustandes beobachtet wurde) im SOEP von Welle zu Welle steigt (obwohl die Zahl der noch in der Stichprobe befindlichen Stammpersonen und -haushalten sinkt). Selbst für eine so kleine Stichproben-Population wie sozialhilfeabhängige Haushalte (ca. 100 Fälle pro Querschnitt) lassen sich nach fünf Wellen signifikante Ergebnisse für die Determinanten der Sozialhilfebedürftigkeit erzielen (vgl. Voges und Rohwer 1991).

Durch das Poolen von Beobachtungen steigt natürlich mit jeder Welle die Zahl der analysierbaren Einheiten, auch wenn im Querschnitt die Zahl der Beobachtungen jeweils kleiner wird. Sind z.B. in der 1. Welle des SOEP 2 313 deutsche Männer mit Arbeitseinkommen in der Stichprobe enthalten, sind nach fünf Wellen 9 436 Werte analysierbar, obwohl noch 1 061 Arbeitnehmer mit vollständigen Einkommensverläufen (complete cases) im Querschnitt der 5. Welle enthalten sind. Abbildung 5 verdeutlicht diese Zusammenhänge wie folgt.

Während der ersten fünf Wellen gibt es 2 821 Personen, für die mindestens in einer Welle ein Arbeitseinkommen vorliegt (diese Fälle werden gleich 100% gesetzt). Davon hatten bereits in der ersten Welle 2 313 ein Arbeitseinkommen; 508 Fälle waren "systematische Ausfälle", weil sie noch in Ausbildung, arbeitslos oder aus anderen Gründen nicht erwerbstätig waren. Von einem kleinen Teil der Arbeitnehmer wurde keine Einkommensangabe gemacht (item nonresponse). Für die in der schraffierten Fläche (valider Wert) der 1. Säule in Abbildung enthaltenen Arbeitnehmer liegt in der 1. Welle ein Einkommen vor, das aber nicht in allen anderen Wellen beobachtet wird, da diese Personen dann - aus verschiedenen Gründen - ausfallen. Es bleiben schließlich nur 1 601 Fälle, die über fünf Wellen 5 305 Werte ergeben, die als "complete cases" analysierbar sind. Hingegen sind 9 436 Werte von Arbeitseinkommen als "complete information" analysierbar, da auch die "validen Werte" von Personen miteinbezogen werden, die nicht während aller Panel-Wellen Arbeitnehmer sind. Für

Abbildung 4

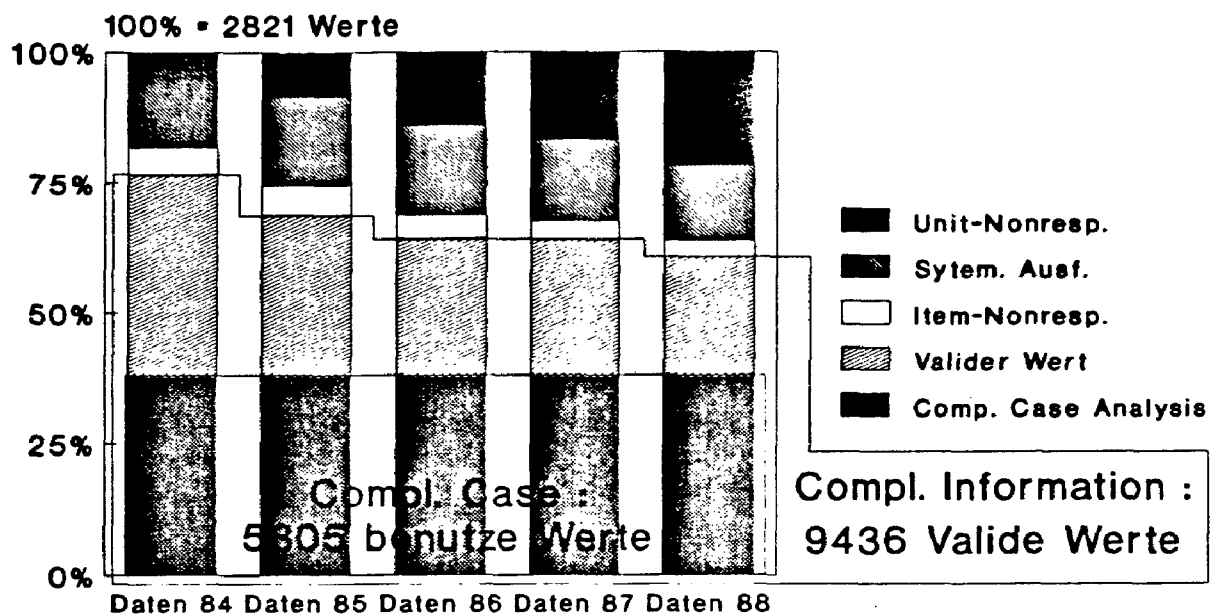
Anzahl unzensurierter Spells im Panel Entwicklung bis zur 7. Welle



Quelle: Das Sozio-ökonomische Panel
Welle 1 bis Welle 7

Abbildung 5: Vergleich von Fallzahlen für die Analyse von
Arbeitseinkommen deutscher Männer im SOEP

Vergleich: Complete Information Analysis und Complete Case Analysis Variable: Bruttoeinkommen



Quelle: Das Sozio-ökonomische Panel
Welle 1 bis Welle 5

entnommen: Löwenbein und Rendtel (1991, S. 160)

die "complete information analysis" kommt es im Laufe des Panels zu einem Zuwachs an Informationen, da ein Teil der "systematischen Ausfälle" der 1. Welle später zu Arbeitnehmern werden. Deswegen gibt es z.B. in der 5. Welle "valide Werte", die nicht in die "complete cases" eingehen, da Einkommensangaben in den Wellen 1 bis 4 mindestens einmal systematisch ausgefallen waren.

Für reine Querschnittsanalysen sieht das Bild des SOEP ebenfalls günstig aus. Durch neu hinzukommende Haushalte (splitt offs von Stammhaushalten) sinkt die Zahl der jeweils in der Stichprobe befindlichen Haushalte und Personen im Querschnitt langsamer als in der Längsschnittsbetrachtung (vgl. Abb. 6.1 und 6.2). Fallzahlprobleme, die gravierender sind als die Fallzahlrestriktionen, wie sie bei den üblichen Umfrage-Stichproben mit ca. 2000 Personen bzw. Haushalten auftreten, wird es beim SOEP noch lange nicht geben. Die Zahl von 2000 Haushalten wird beim inzwischen erreichten Teilnahmeverhalten und der Weiterverfolgung aller Stichprobenhaushalte erwartungsgemäß erst bei Welle 50 unterschritten (Annahme: 2 % Nettoverlust pro Welle)⁶. Bei einem Netto-Verlust von nur 1 % der Haushalte, der angesichts der Entwicklung zwischen Welle 6 und 7 nicht unrealistisch ist, würde erst nach weit über 100 Wellen die Zahl von 2000 Haushalten unterschritten (zusammen mit der neuen Ost-Stichprobe des SOEPs noch weit später).

3 Ausblick

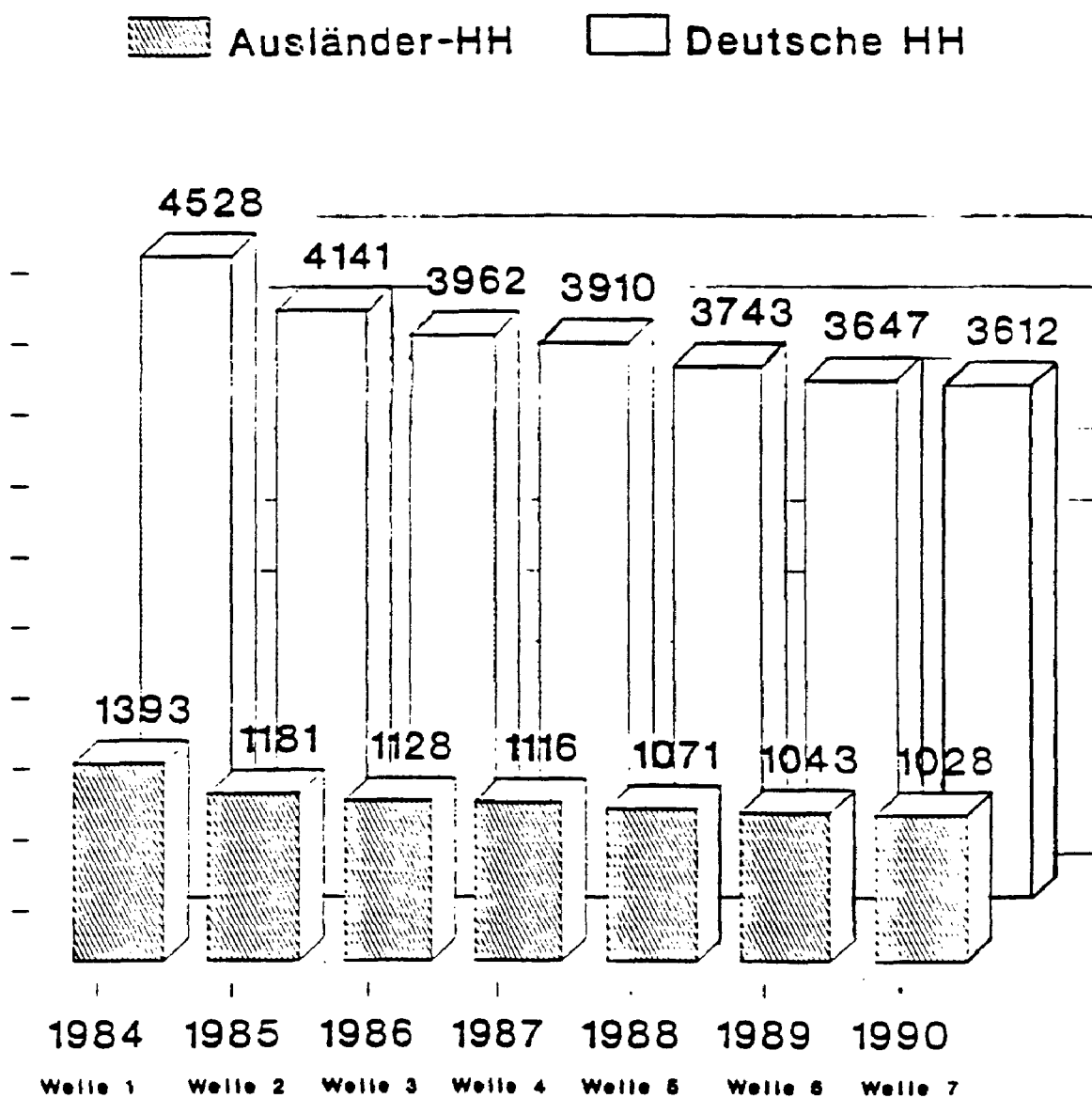
Auf Basis der Abbildung 2 zugrundeliegenden Darstellung wird die "magische Grenze" von weniger als 50 % der Ausgangspopulation vom SOEP deutlich später erreicht werden als in der bislang üblichen Rechnung. In der konventionellen Berechnungsweise wäre dieses Unterschreiten in Welle 9 erwartbar, während es in der angemessenen Rechnung erst in Welle 15 zu erwarten ist. Ob nach Unterschreiten dieser Grenze die Repräsentativität des SOEP gefährdet ist, ist eine freilich viel

6 Da im SOEP die Personen eines Haushaltes nicht statistisch unabhängig sind, sollte auch zur Charakterisierung der kritischen Fallzahl die Zahl der Haushalte herangezogen werden.

Abbildung 6.1

Das Sozio-ökonomische Panel 1984 - 1990

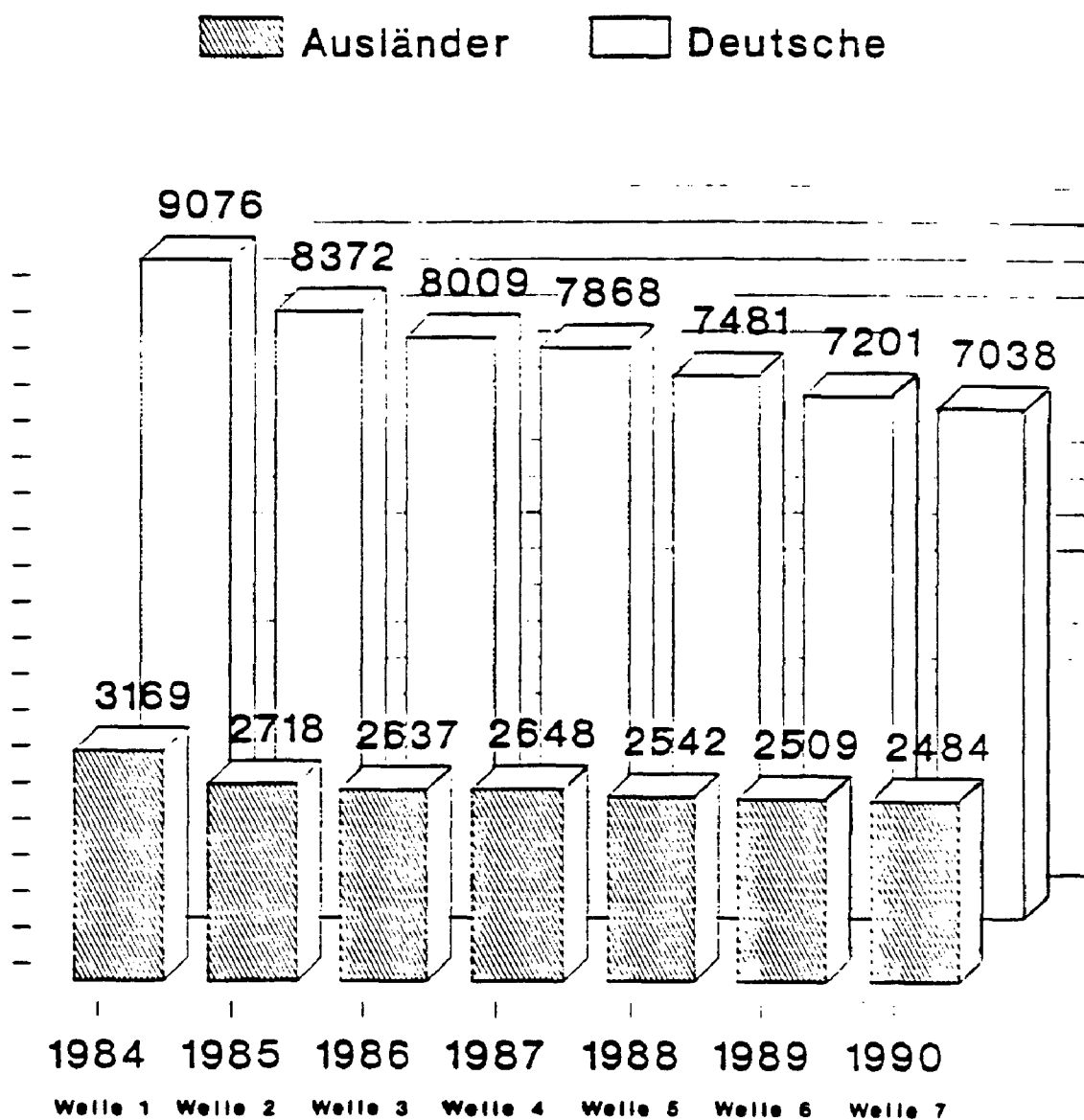
Teilnehmerzahl Haushalte



Fallzahl							
gesamt	5921	5222	5090	5026	4814	4690	4640
Netto-							
verlust		-699	-132	-64	-212	-124	-50

Abbildung 6.2

Das Sozio-ökonomische Panel 1984 - 1990 Teilnehmerzahl Personen



Fallzahl
gesamt

12245 11090 10646 10516 10023 9710 9522

Netto-
verlust

-1155 -444 -130 -493 -313 -188

weitergehende Frage. Sie kann nur durch den Vergleich der dann verbliebenen Stichprobe mit Daten der Grundgesamtheit bzw. anderer Stichproben geprüft werden. Die Ergebnisse der US-amerikanischen Vorbildstudie PSID lassen bezüglich der Repräsentativität langlaufender Panel-Erhebung Optimismus angezeigt sein (vgl. Beckett et al 1988). Darauf weisen auch die Analysen von Rendtel (1990) hin, der zeigen kann, daß im SOEP seit der 3. Welle keine sozialstrukturellen Ausfallmuster mehr beobachtbar sind⁷.

Schließlich sei angemerkt, daß - im Gegensatz zur PSID - die SOEP-Daten standardmäßig so weitergegeben werden, daß jeder Nutzer eigenständige Ausfallanalysen durchführen kann, bzw. bei der Schätzung von Modell-Parametern den Ausfallprozeß modellieren kann. Es sind somit keine spekulativen Betrachtungen über die sozial-strukturelle Repräsentativität der Panel-Stichproben notwendig, sondern diese Repräsentativität des Querschnitts wie des Längsschnitts ist jederzeit von jedem Nutzer nachprüfbar und Ausfallprozesse sind modellierbar⁸.

- 7 Dieser Befund ist auch aus den von Esser et al. (1989, s. 142) benutzten Graphiken ablesbar, die auf den ersten Blick permanente soziale Unterschiede im Ausfallprozeß signalisieren. Ab Welle 3 verlaufen die Ausfallkurven weitgehend parallel, d.h., es gibt keine sozialdifferenziellen Unterschiede im Ausfallverhalten mehr. Hinzu kommt, daß zwischen Welle 1 und 3 einige Personen bereits ihre sozialen Merkmale geändert haben und dadurch die tatsächlichen Unterschiede in der Graphik überzeichnet werden.
- 8 Außerdem ist das Umfrageinstitut "Infratest Sozialforschung" (München) vertraglich verpflichtet, die Durchführung jeder Erhebungswelle mit einem ausführlichen Methodenbericht zu dokumentieren. Diese Berichte können jedem Datennutzer von der Panel-Gruppe im DIW auf Wunsch zur Verfügung gestellt werden.

Literatur

- Beckett, Sean et al. 1988: The Panel Study of Income Dynamics after 14 Years - an evaluation, in: The Journal of Labor Economics, 6 (4), S. 472-492
- Brecht, Beatrix und Paul Michels 1991: Anwendung nicht-parametrischer Schätzverfahren für die Hazardfunktion bei zensierten Daten auf die Aufenthaltsdauer von Gastarbeitern in der Bundesrepublik, Diskussionspapier Nr. 137, Konstanz
- Esser, Harmut et al. 1989: Mikrozensus im Wandel, Stuttgart
- Kasprzyk, Daniel et al. (Hrsg.) 1989: Panel Surveys, New York u.a.
- Lazarsfeld, Paul F. et al. 1957: Die Panel-Befragung, in: R. König (Hg.), Das Interview, Köln, S. 253-268.
- Löwenbein, Oded und Ulrich Rendtel 1991: Selektivität und Panelanalyse, in: U. Rendtel und G. Wagner (Hg.), Lebenslagen im Wandel - Zur Einkommensdynamik in Deutschland seit 1984, Frankfurt, New York, S. 156-187
- Projektgruppe Panel 1990: Das Sozio-ökonomische Panel für die Bundesrepublik Deutschland nach fünf Wellen, in: Vierteljahrshefte zur Wirtschaftsforschung, (2-3), S. 141-151
- Projektgruppe Panel (Hg.) 1991: Lebenslagen im Wandel - Basisdaten und -analysen zur Entwicklung in den Neuen Bundesländern, Frankfurt, New York
- Rendtel, Ulrich 1990: Teilnahmebereitschaft in Panel-Studien - zwischen Beeinflussung, Vertrauen und sozialer Selektion - über die Entwicklung der Antwortbereitschaft im Sozio-ökonomischen Panel, in: Kölner Zeitschrift für Soziologie und Sozialpsychologie, 42 (2), S. 280-299
- Rendtel, Ulrich 1991: Hochrechnung und Stichprobenfehler von Panel-Erhebungen in: Allgemeines Statistisches Archiv, 75(2)
- Schulz, Erika 1991: Die Wanderungen ins Bundesgebiet seit 1984, DIW-Diskussionspapier Nr. 28, Berlin

- Sharma, Sushil K. et al. 1989: Attrition in the Baltimore Longitudinal Study of Aging During the First Twenty Years, in: M.P. Lawton und A.P. Herzog (Hg.), Special Research Methods for Gerontology, Amityville, New York, S. 233-247
- Voges, Wolfgang und Götz Rohwer 1991: Zur Dynamik des Sozialhilfebezugs, in: U. Rendtel und G. Wagner (Hg.), Lebenslagen im Wandel - Zur Einkommensdynamik in Deutschland seit 1984, S. 510-531

Anhang

Entwicklung der SOEP-Stichprobe

Basis: Befragungspersonen der 1. Welle (N=12 245)

	1	2	3	4	5	6	7
Verstorben	0.0	0.6	1.3	1.9	2.6	3.2	3.7
Ins Ausland verzogen	0.0	1.6	2.1	2.6	3.3	3.7	4.0
Kein Kontakt	0.0	1.0	1.3	1.6	1.8	1.9	2.1
Antwortverweigerung	0.0	8.8	14.7	18.2	22.2	25.5	28.9
Noch in Stichprobe	100.0	88.1	80.6	75.8	70.2	65.7	61.2

Basis: Stammpersonen der 1. Welle (N=16 205)

	1	2	3	4	5	6	7
Ins Ausland verzogen	0.0	1.6	2.1	2.5	3.2	3.6	4.0
Verstorben	0.0	0.4	1.0	1.4	2.0	2.5	2.9
Unter 16 Jahre	24.1	19.7	16.9	14.9	12.5	10.8	9.4
Mit Interview	75.5	66.7	63.1	61.3	57.7	54.9	52.8
Temporärer Ausfall	0.3	2.2	2.1	1.5	1.9	2.0	1.2
Antwortverweigerung	0.0	8.2	13.6	16.9	20.9	24.3	27.8
Kein Kontakt	0.0	0.9	1.2	1.5	1.7	1.8	2.0

Basis: Stammpersonen der 1. Welle, Stichprobe A (N=11 406)

	1	2	3	4	5	6	7
Ins Ausland verzogen	0.0	0.3	0.3	0.4	0.5	0.6	0.7
Verstorben	0.0	0.6	1.3	1.9	2.7	3.3	3.9
Unter 16 Jahre	20.0	16.8	14.3	12.6	10.7	9.2	8.1
Mit Interview	79.6	71.8	67.5	65.3	61.6	58.5	56.0
Temporärer Ausfall	0.4	1.5	1.5	1.1	1.5	1.5	0.8
Antwortverweigerung	0.0	8.3	14.0	17.4	21.6	25.3	28.8
Kein Kontakt	0.0	0.8	1.0	1.3	1.5	1.6	1.7

Basis: Stammpersonen der 1. Welle, Stichprobe B (N=4 799)

	1	2	3	4	5	6	7
Ins Ausland verzogen	0.0	4.7	6.2	7.6	9.6	10.8	11.7
Verstorben	0.0	0.1	0.2	0.3	0.4	0.5	0.6
Unter 16 Jahre	33.7	26.6	22.9	20.4	16.9	14.5	12.3
Mit Interview	65.9	55.2	52.6	51.7	48.5	46.4	45.1
Temporärer Ausfall	0.3	4.0	3.6	2.5	3.0	3.2	2.3
Antwortverweigerung	0.0	8.0	12.7	15.5	19.3	22.1	25.3
Kein Kontakt	0.0	1.3	1.6	1.9	2.2	2.4	2.6