

Data Documentation

50

Martin Kroh

**Documentation of Sample Sizes and Panel
Attrition in the German Socio Economic
Panel (SOEP) (1984 until 2009)**

Berlin, August 2010

IMPRESSUM

© DIW Berlin, 2010

DIW Berlin
Deutsches Institut für Wirtschaftsforschung
Mohrenstr. 58
10117 Berlin
Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
www.diw.de

ISSN 1861-1532

All rights reserved.
Reproduction and distribution
in any form, also in parts,
requires the express written
permission of DIW Berlin.



Data Documentation 50

Martin Kroh*

Documentation of Sample Sizes and Panel Attrition in the German Socio Economic Panel (SOEP) (1984 until 2009)

Berlin, August 2010

* DIW Berlin, Socio-Economic Panel Study. mkroh@diw.de

I would like to thank Verena Pflieger and Hannes Neiss for excellent research assistance.

Contents

1 Introduction	1
2 Developments in Sample Size.....	2
2.1 Development of the Number of Successful Interviews by Cross-Section	3
2.2 Continuance and Exit: The First Wave Gross Samples and their Participatory Behavior	11
2.3 New Entrants through Birth or Move into SOEP Households and Their Participation Behavior.....	16
2.4 The Risk of Survey-Related Panel Attrition.....	21
3 Panel Attrition Due to Unsuccessful Follow-Ups	25
3.1 The Frequency of Successful Follow-Ups	25
3.2 Predicting the Probability of Successful vs. Unsuccessful Follow-Ups in the Year 2009	27
4 Panel Attrition Due to Refusals.....	29
4.1 The Frequency of Participation.....	29
4.2 Predicting the Probability of Re-Interviewing vs. Refusal in the Year 2009	31
5 Summary Statistics of the Derived Longitudinal and Cross-Sectional Weights	37
6 References	41

Tables

Table 1: The Frequency of Households to be Re-Contacted and the Relative Proportion of Successful Follow-Ups by Subsample and Year.	26
Table 2: Definition of the Regressors of the Logit Model of Unsuccessful Follow-Ups.	27
Table 3: Estimates of Logit Models of the Probability of Re-Contacting a Household (Relative to Unsuccessful Follow-Up) in 2009.	28
Table 4: The Frequency of Re-Contacted Households and the Relative Proportion of Participation by Subsample and Year.	30
Table 5: Definition of the Regressors of the Logit Model of Refusal.	32
Table 6a: Estimates of Logit Models for the Probability of Re-Interviewing a Household (Relative to Refusal) in 2009.	34
Table 6b: Estimates of Logit Models for the Probability of Re-Interviewing a Household (Relative to Refusal) in 2009.	35
Table 6c: Estimates of Logit Models for the Probability of Re-Interviewing a Household (Relative to Refusal) in 2009.	36
Table 7a: Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples A through D (Percentiles of \$HBLEIB up to Wave 26).	38
Table 7b: Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples E through H (Percentiles of \$HBLEIB up to Wave 26).	39
Table 8: Summary Statistics of the Derived Cross-Sectional Weights at the Household Level (Percentiles of \$HHRF up to Wave 26).	40

Figures

Figure 1: The Number of Successful Interviews with Persons by Subsamples A through I, Waves 1 to 26.	4
Figure 2: Comparison of Successful Interviews with Persons and Households (Subsamples A and B), Waves 1 to 26.	5
Figure 3: Comparison of Successful Interviews with Persons and Households (Subsample C), Waves 1 to 20.	6
Figure 4: Comparison of Successful Interviews with Individuals and Households (Subsample D), Waves 1 to 15.	7
Figure 5: Comparison of successful interviews with individuals and households (Subsample E), Waves 1 to 12.	8
Figure 6: Comparison of Successful Interviews with Individuals and Households (Subsample F), Waves 1 to 10.	9
Figure 7: Comparison of Successful Interviews with Individuals and Households (Subsample G), Waves 1 to 8.	10
Figure 8: Comparison of Successful Interviews with Individuals and Households (Subsample H), Waves 1 to 4.	10
Figure 9: All First-Wave Persons (Gross Subsample A). Development up to Wave 26.	12
Figure 10: All First Wave Persons (Gross Subsample B). Development up to Wave 26.	12
Figure 11: All First Wave Persons (Gross Subsample C). Development up to Wave 20.	13
Figure 12: All First Wave Persons (Gross Subsample D). Development up to Wave 15.	13
Figure 13: All First Wave Persons (Gross Subsample E). Development up to Wave 12.	14
Figure 14: All First Wave Persons (Gross Subsample F). Development up to Wave 10.	14
Figure 15: All First Wave Persons (Gross Subsample G). Development up to Wave 8.	15
Figure 16: All First Wave Persons (Gross Subsample H). Development up to Wave 4.	15
Figure 17: Entrants and their Participation Behavior (Subsamples A, B).	17
Figure 18: Entrants and their Participation Behavior (Subsample C).	17
Figure 19: Entrants and their Participation Behavior (Subsample D).	18
Figure 20: Entrants and their Participation Behavior (Subsample E).	18

Figure 21: Entrants and their Participation Behavior (Subsample F).....	19
Figure 22: Entrants and their Participation Behavior (Subsample G).....	19
Figure 23: Entrants and their Participation Behavior (Subsample H).....	20
Figure 24: Successful Re-Interviewing of First-Wave Respondents by Subsamples A, B, C. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad.....	22
Figure 25: Successful Re-Interviewing of First-Wave Respondents by Subsamples D, E, F, G, H. Kaplan-Meier Estimates of Survey- Related Attrition Ignoring Deaths and Moves Abroad.....	22
Figure 26: Successful Re-Interviewing of All First-Wave Respondents by Age Categories. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad.....	23
Figure 27: Successful Re-Interviewing of All First-Wave Respondents by Occupation. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad.	23
Figure 28: Successful Re-Interviewing of All First-Wave Respondents by Income Quintiles. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad.....	24
Figure 29: Successful Re-Interviewing of All First-Wave Respondents by Education. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad.	24

1 Introduction

This data documentation is meant to provide SOEP users with a general overview of the longitudinal development of the survey over the past 26 years and the derivation of weights that compensate for selective panel attrition. In the first section, we report the number of household and personal interviews by cross-section. We do so for the entire SOEP sample as a whole, as well as for sub-samples A through H individually.

In 2009, a new sample I was drawn and interviewed for the first time ($n=1531$ households from the general population). The data dissemination of 2010 contains all household and personal response information on samples A through H as well as this latest sample. Also, the data file DESIGN provides the design weights of sample I (clustering and stratification). Sample I is, however, not yet incorporated in the integrated cross-sectional weighting frame of 2009. This is because sample I in contrast to the existing samples A through H experimentally tests the effectiveness of monetary incentives on participation rates. These experiments thus by definition affect the response probabilities and weights in sample I and require first in depth analysis of their effects before these cases can reliably be integrated in the weighting frame of SOEP. The data distribution 2011 will provide an integration of samples A through I.

The SOEP surveys not only the original sample from the first wave, but also households and persons that entered the survey at later points in time. They enter, for example, when SOEP households split (i.e., individuals move out and form their own households), when people move into SOEP households, and when an original sample member gives birth to a “new sample member”. For a detailed review of the SOEP inclusion rules for new sample units and their treatment within the weighting framework see Spiess et al. (2008). The second section of the present paper on the longitudinal development of the SOEP reports descriptive figures of the participatory behavior of the original sample members and the entrance patterns of new sample members.

Households may leave the survey for several reasons. SOEP’s weighting strategy distinguishes between survey-related reasons and reasons unrelated to the survey (for a detailed description of the SOEP weighting strategy, see Rendtel 1995 and for

a general overview, Haisken-DeNew & Frick 2001). We ignore panel attrition of the latter form due to respondents moving abroad or dying, since these cases technically represent an exit from the underlying population. The second section of this paper provides initial evidence on the risk of survey-related panel attrition in different groups of the original sample units (e.g., in different sub-samples, age, educational, and income groups).

The third section reports in more detail on the occurrence of unsuccessful follow-ups to household addresses by cross-section and sub-sample, and sub-sample-specific regression models of the probability of unsuccessful follow-ups in 2009 based on the characteristics of households measured in 2008. The fourth section does the same for the second form of survey-related attrition: refusals.

Based on the regression models of unsuccessful-follow ups and refusals, we derive predicted observation probabilities. The inverse of the product of these predicted probabilities gives the longitudinal weighting variables for the year 2009: ZHBLEIB and ZPBLEIB. Based on the inverse of the probability of observing households and persons in 2008, the staying probability in 2009, and additional post-stratification to meet benchmarks of known marginals of the underlying population in 2009, we derive the cross-sectional weights ZHHRF and ZPHRF. The final section of this paper documents some summary statistics of the development of the longitudinal and the cross-sectional weights by sub-sample and wave.

2 Developments in Sample Size

With respect to developments in sample size, the following figures focus on (2.1) comparing the number of successful interviews by cross-section, (2.2) providing a longitudinal study of panel attrition in original sample members, (2.3) showing entrance of new sample members by birth / moving into SOEP households and their participation behavior, and (2.4) assessing the risk of survey-related attrition of original sample respondents by social characteristics.

Note that the sample sizes of the English public-use version of SOEP and the German DIW version differ by approximately 5 percent. Five percent of the original

SOEP data was excluded in compliance with German data protection laws, which was accomplished technically by randomly selecting 5 percent of the original wave 1 households and dropping these and the persons living in them from the English public-use version. Hence the difference in sample sizes is not always exactly 5 percent. The sample sizes documented below refer to the original DIW database.

2.1 Development of the Number of Successful Interviews by Cross-Section

The following figures display the number of successful interviews considering different aspects:

- Figure 1** The Number of Successful Interviews with Persons by Subsamples A through I, Waves 1 to 26 (1984-2009)
- Figure 2** Comparison for Individuals and Households in Subsamples A and B, Waves 1 to 26 (1984 – 2009).
- Figure 3** Comparison for Individuals and Households in Subsample C, Waves 1 to 20, (1990–2009).
- Figure 4** Comparison for Individuals and Households in Subsample D, Waves 1 to 15, (1995–2009).
- Figure 5** Comparison for Individuals and Households in Subsample E, Waves 1 to 12, (1998–2009).
- Figure 6** Comparison for Individuals and Households in Subsample F, Waves 1 to 10, (2000–2009).
- Figure 7** Comparison for Individuals and Households in Subsample G, Waves 1 to 8, (2002-2009).
- Figure 8** Comparison for Individuals and Households in Subsample H, Waves 1 to 4, (2006-2009).

Figure 1: The Number of Successful Interviews with Persons by Subsamples A through I, Waves 1 to 26.

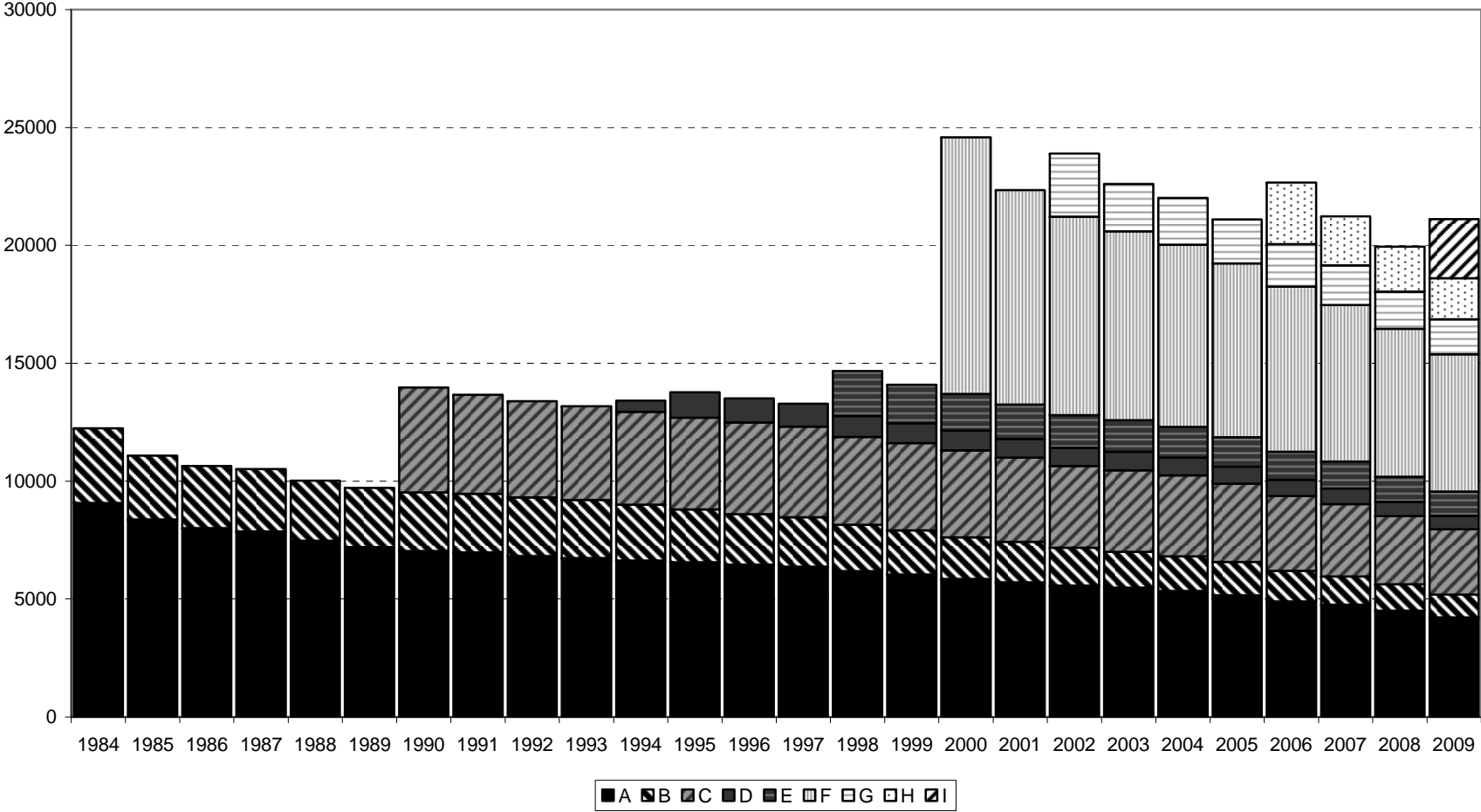
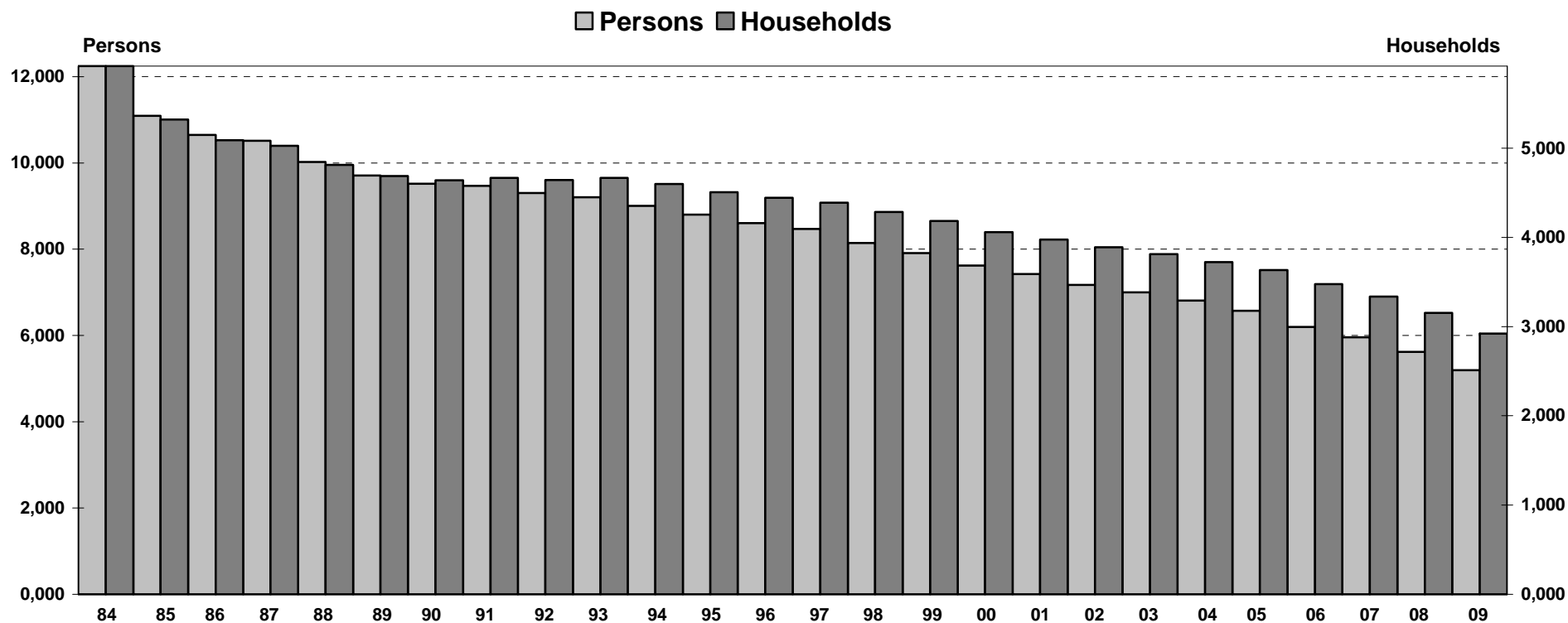
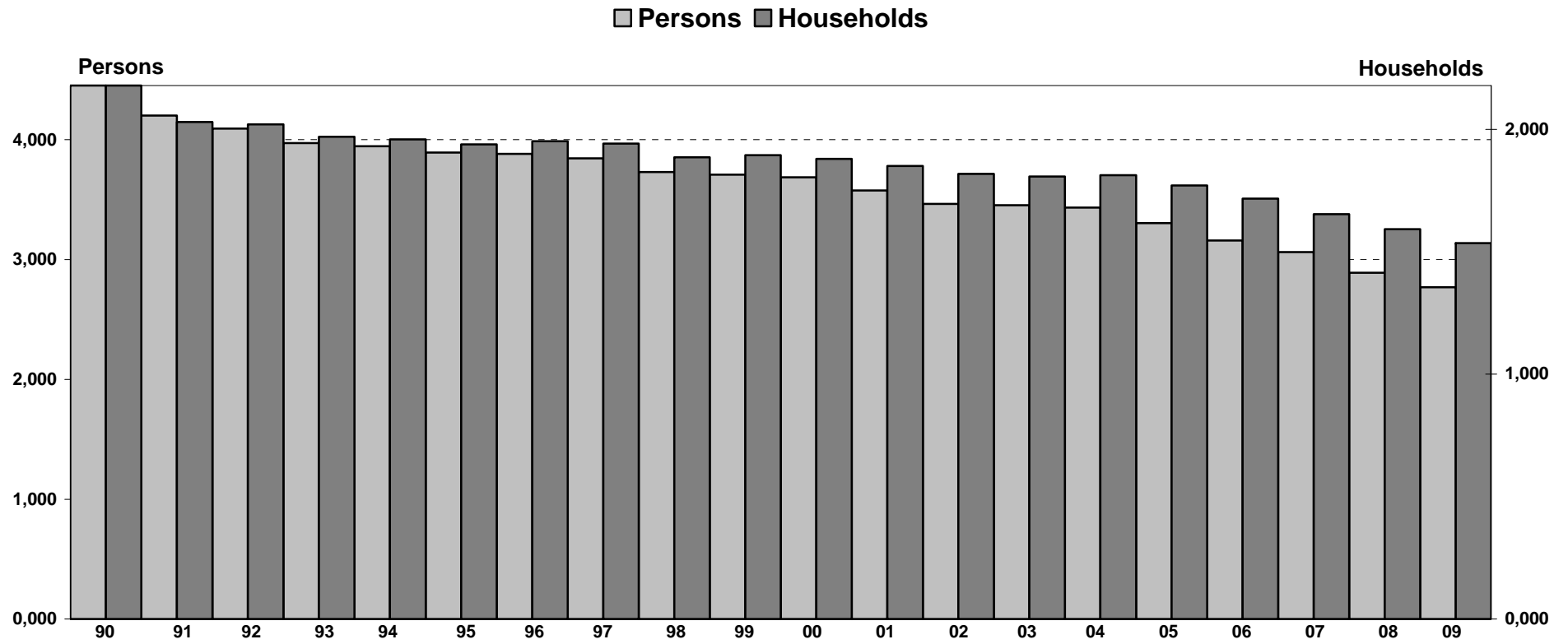


Figure 2: Comparison of Successful Interviews with Persons and Households (Subsamples A and B), Waves 1 to 26



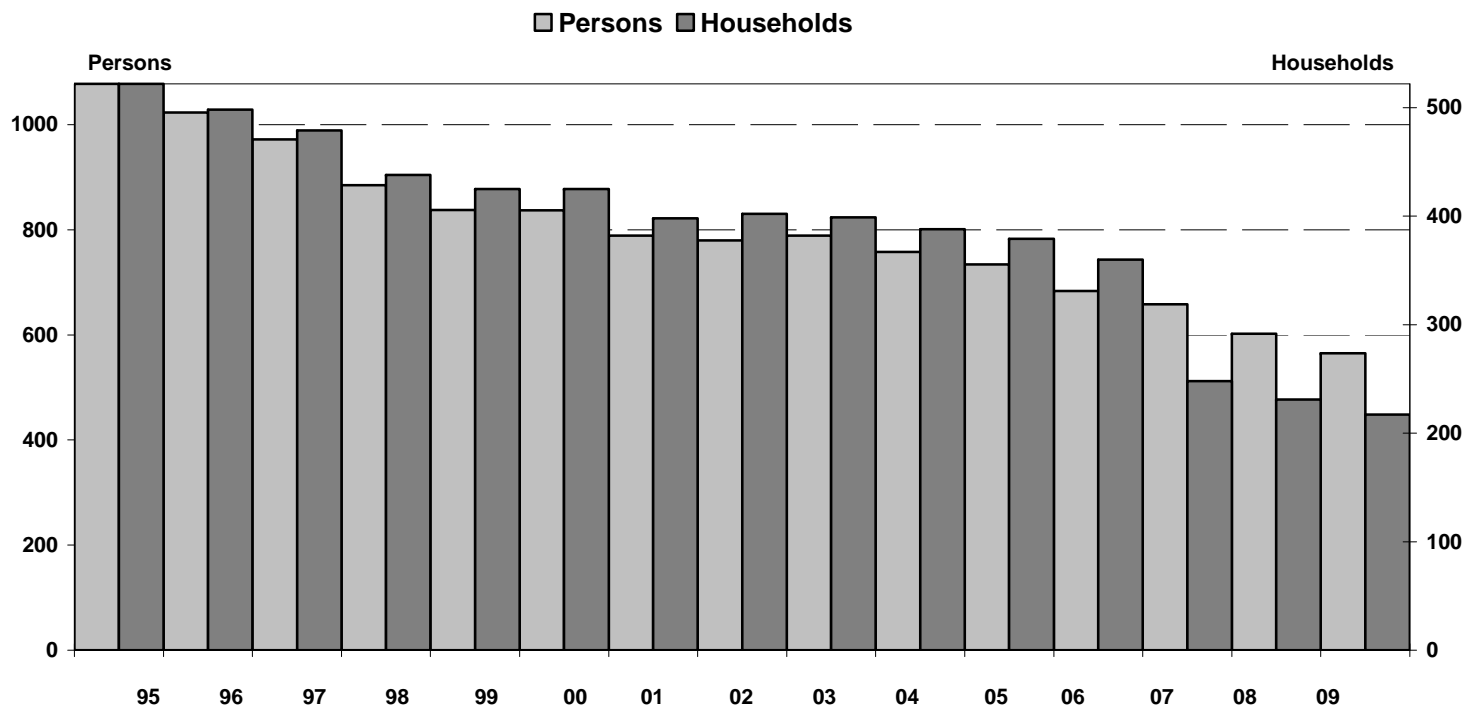
Year	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Persons	12,245	11,090	10,646	10,516	10,023	9,710	9,519	9,467	9,305	9,206	9,001	8,798	8,606	8,467	8,145	7,909	7,623	7,424	7,175	6,999	6,809	6,572	6,198	5,957	5,619	5,197
Households	5,921	5,322	5,090	5,026	4,814	4,690	4,640	4,669	4,645	4,667	4,600	4,508	4,445	4,389	4,285	4,183	4,060	3,977	3,889	3,814	3,724	3,635	3,476	3,337	3,154	2,923

Figure 3: Comparison of Successful Interviews with Persons and Households (Subsample C), Waves 1 to 20.



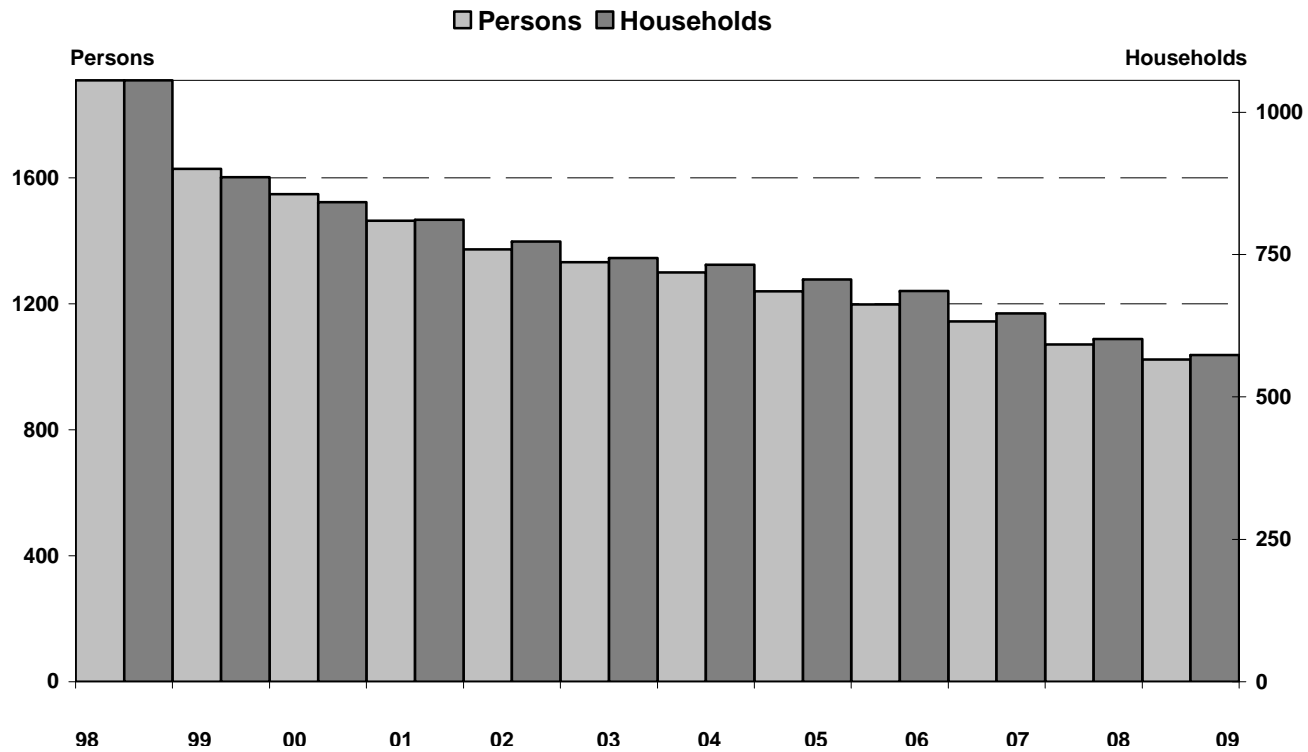
Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Persons	4,453	4,202	4,092	3,973	3,945	3,892	3,882	3,844	3,730	3,709	3,687	3,576	3,466	3,453	3,435	3,304	3,159	3,063	2,889	2,769
Households	2,179	2,030	2,020	1,970	1,959	1,938	1,951	1,942	1,886	1,894	1,879	1,850	1,818	1,807	1,813	1,771	1,717	1,654	1,592	1,535

Figure 4: Comparison of Successful Interviews with Individuals and Households (Subsample D), Waves 1 to 15.



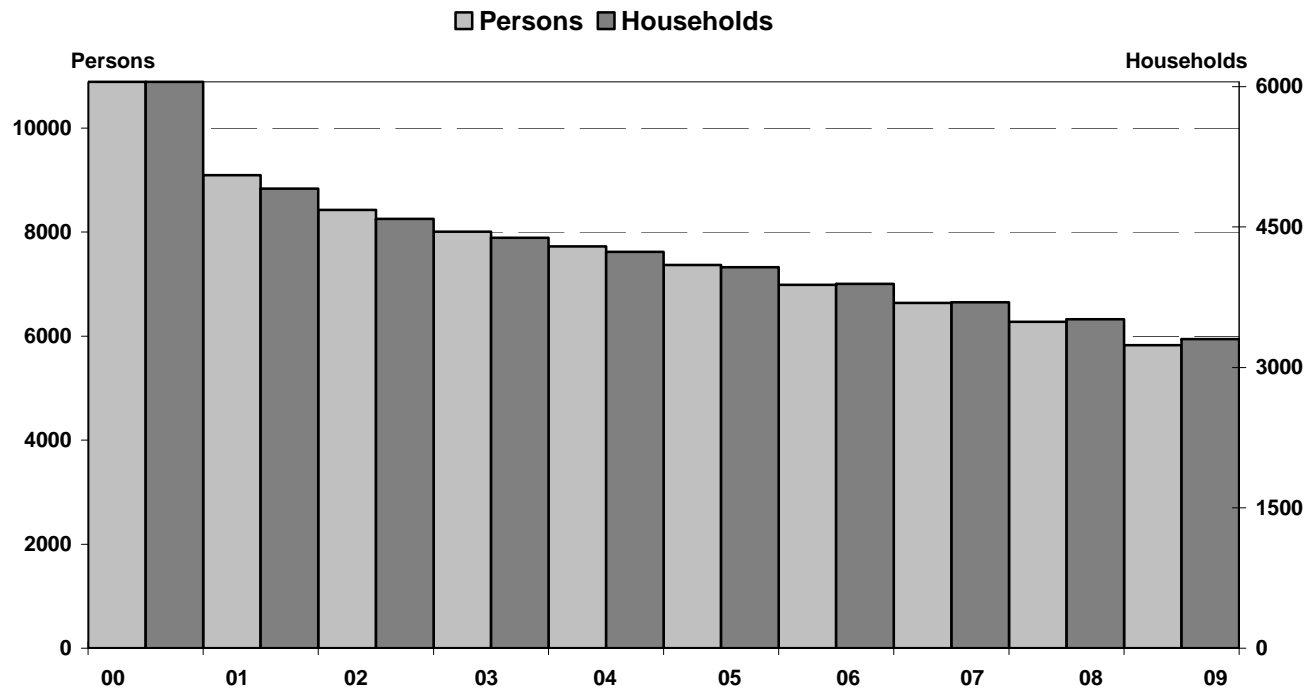
Year	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Persons	1,078	1,023	972	885	838	837	789	780	789	758	734	684	658	602	565
Households	522	498	479	441	425	425	398	402	399	388	379	360	248	231	217

Figure 5: Comparison of successful interviews with individuals and households (Subsample E), Waves 1 to 12.



Year	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Persons	1,910	1,629	1,549	1,464	1,373	1,332	1,300	1,240	1,198	1,144	1,071	1,024
Households	1056	886	842	811	773	744	732	706	686	647	602	574

Figure 6: Comparison of Successful Interviews with Individuals and Households (Subsample F), Waves 1 to 10.

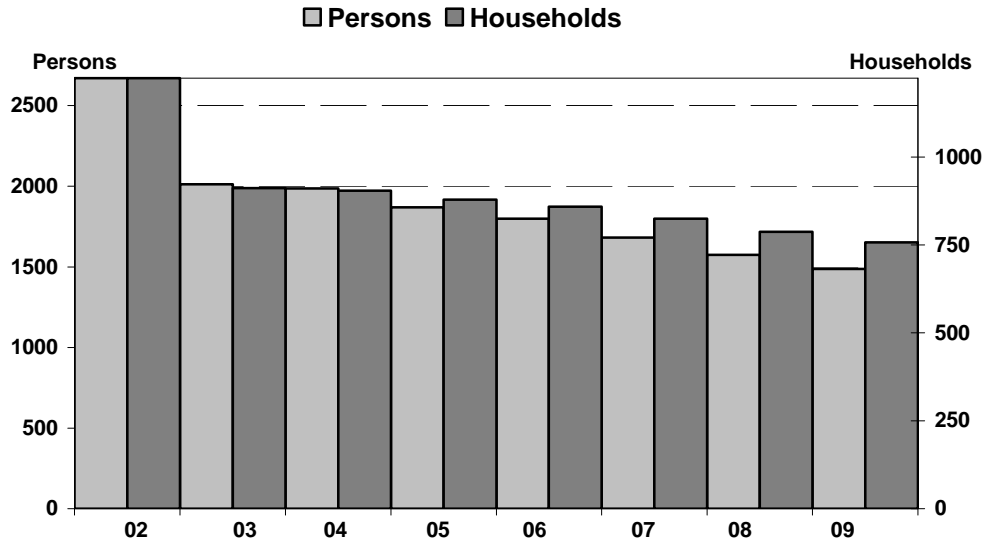


Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Persons	1,0890	9,098	8,427	8,006	7,724	7,371	6,986	6640	6,274	5,824
Households	6,052	4,911	4,586	4,386	4,234	4,070	3,895	3,694	3,513	3,303

Data Documentation 50

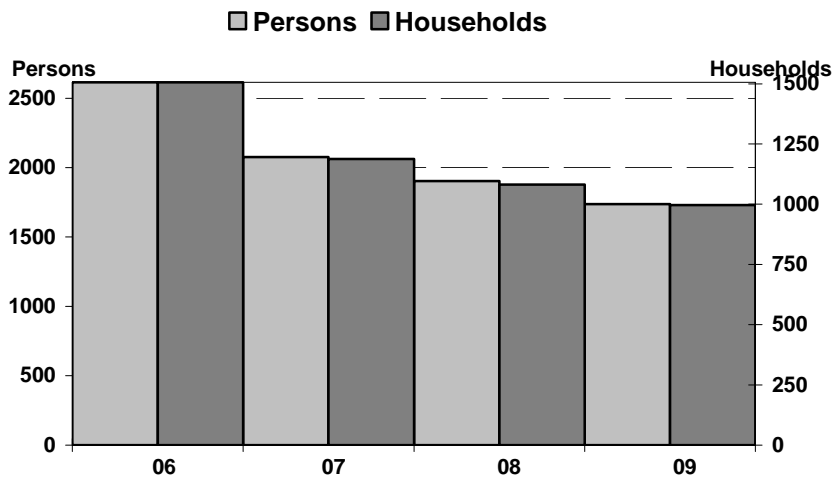
2 Developments in Sample Size

Figure 7: Comparison of Successful Interviews with Individuals and Households (Subsample G), Waves 1 to 8.



Year	2002	2003	2004	2005	2006	2007	2008	2009
Persons	2,671	2,013	1,986	1,870	1,798	1,682	1,574	1,487
Households	1,224	911	904	879	859	824	787	757

Figure 8: Comparison of Successful Interviews with Individuals and Households (Subsample H), Waves 1 to 4.



Year	2006	2007	2008	2009
Persons	2,616	2,077	1,904	1,737
Households	1,506	1,188	1,082	996

2.2 Continuance and Exit: The First Wave Gross Samples and their Participatory Behavior

The following figures display the participation behavior of the first-wave respondents in the subsequent years distinguishing between continued participation, exits due to survey-unrelated attrition, and exits due to survey-related attrition.

Figure 9: All First Wave Persons in Subsample A. Whereabouts up to Wave 26.

Figure 10: All First Wave Persons in Subsample B. Whereabouts up to Wave 26.

Figure 11: All First Wave Persons in Subsample C. Whereabouts up to Wave 20.

Figure 12: All First Wave Persons in Subsample D. Whereabouts up to Wave 15.

Figure 13: All First Wave Persons in Subsample E. Whereabouts up to Wave 12.

Figure 14: All First Wave Persons in Subsample F. Whereabouts up to Wave 10.

Figure 15: All First Wave Persons in Subsample G. Whereabouts up to Wave 8.

Figure 16: All First Wave Persons in Subsample H. Whereabouts up to Wave 4.

Figure 9: All First-Wave Persons (Gross Subsample A). Development up to Wave 26.

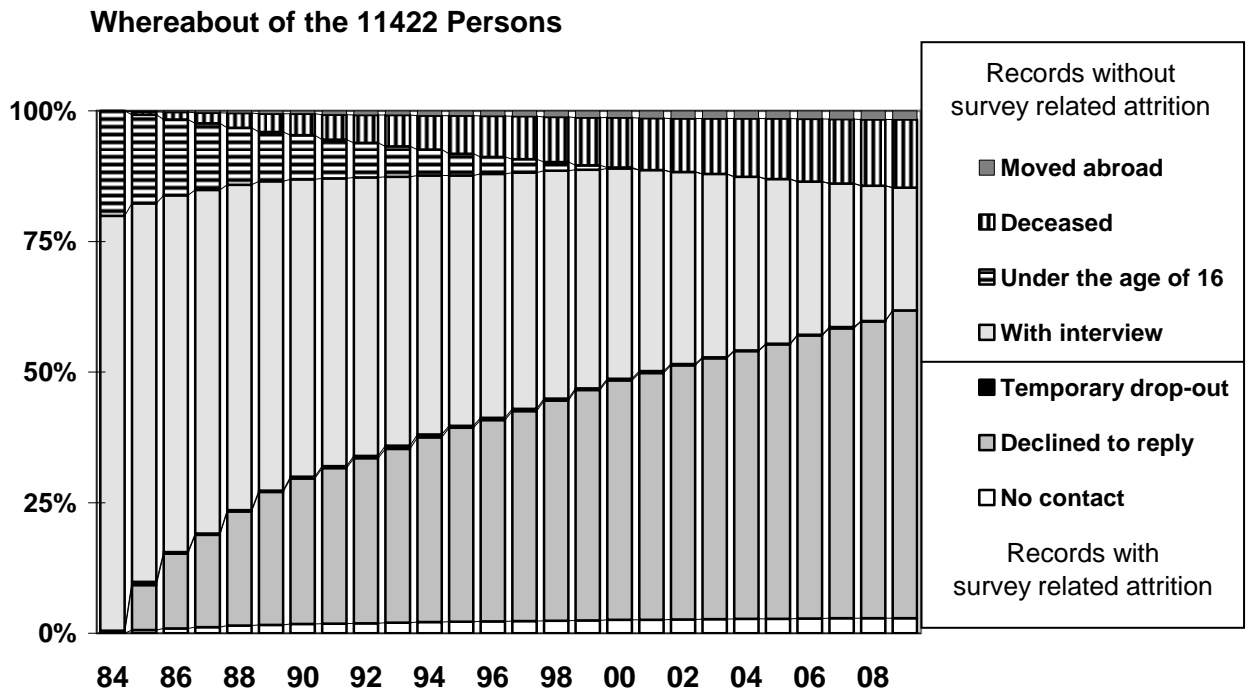


Figure 10: All First Wave Persons (Gross Subsample B). Development up to Wave 26.

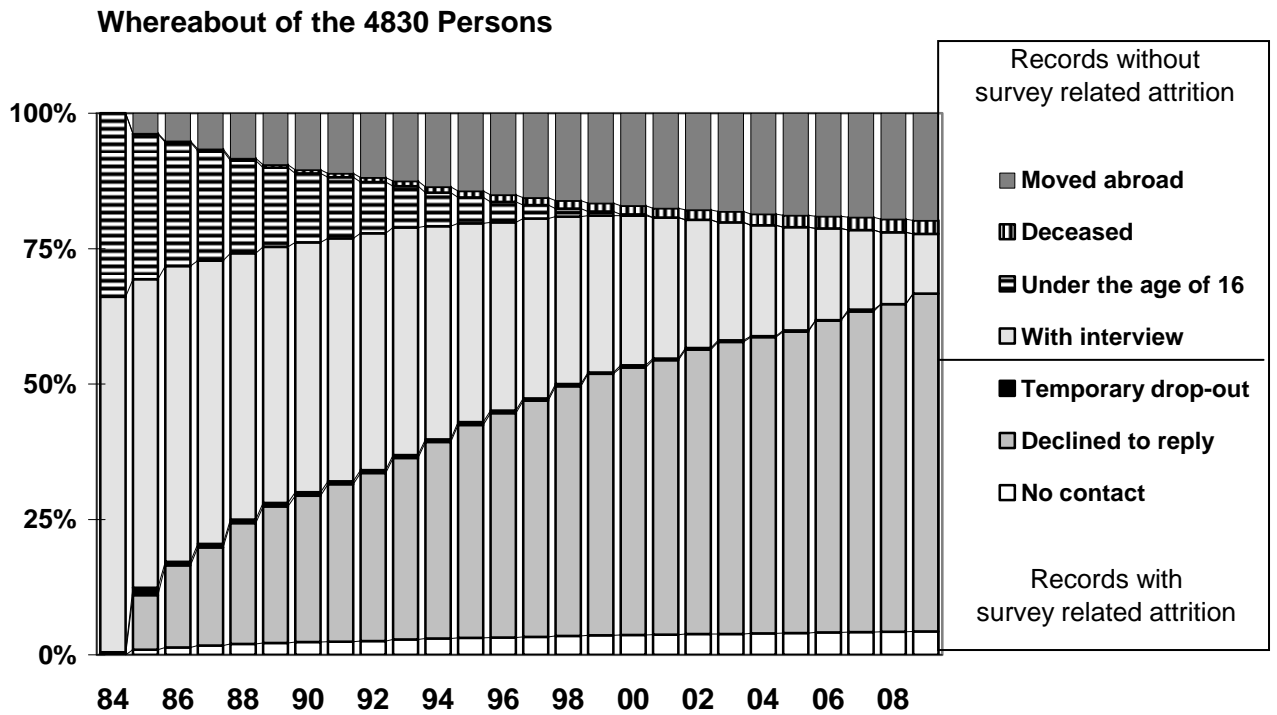


Figure 11: All First Wave Persons (Gross Subsample C). Development up to Wave 20.

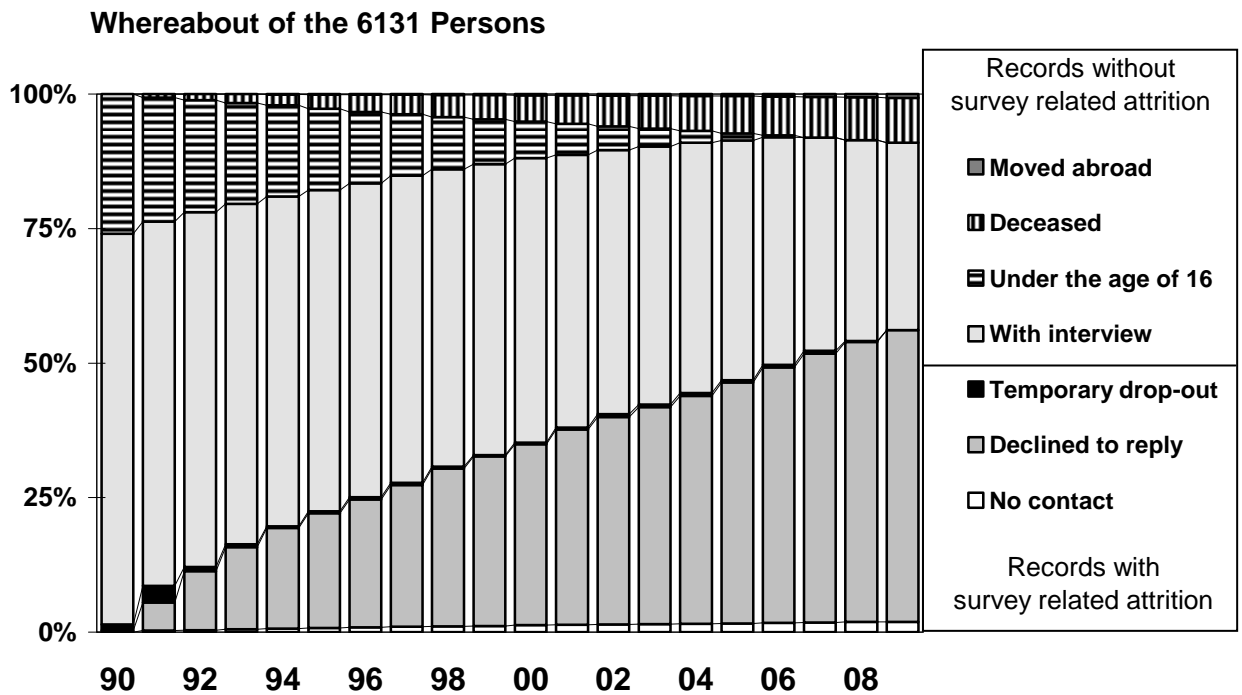


Figure 12: All First Wave Persons (Gross Subsample D). Development up to Wave 15.

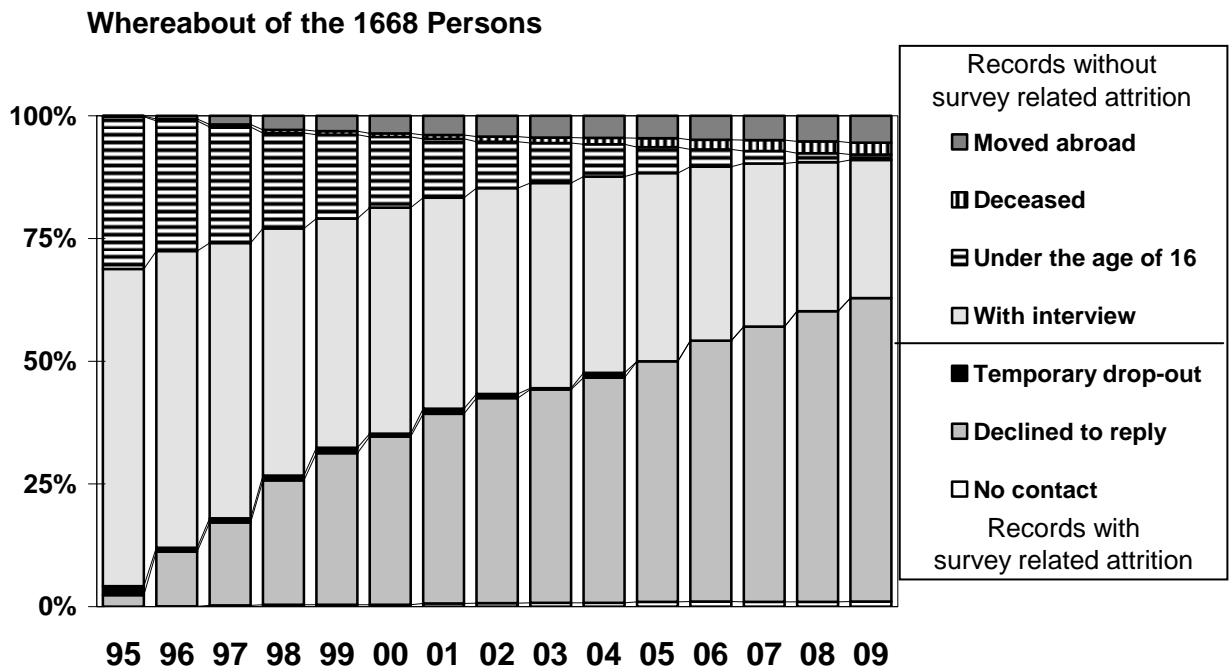


Figure 13: All First Wave Persons (Gross Subsample E). Development up to Wave 12.

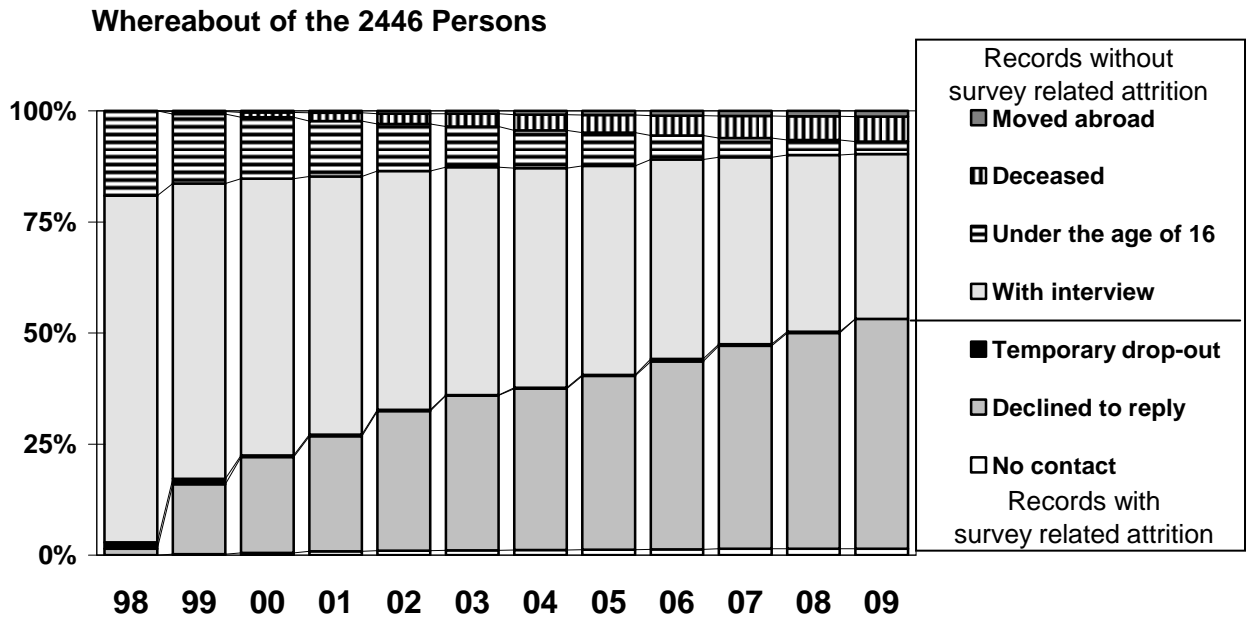


Figure 14: All First Wave Persons (Gross Subsample F). Development up to Wave 10.

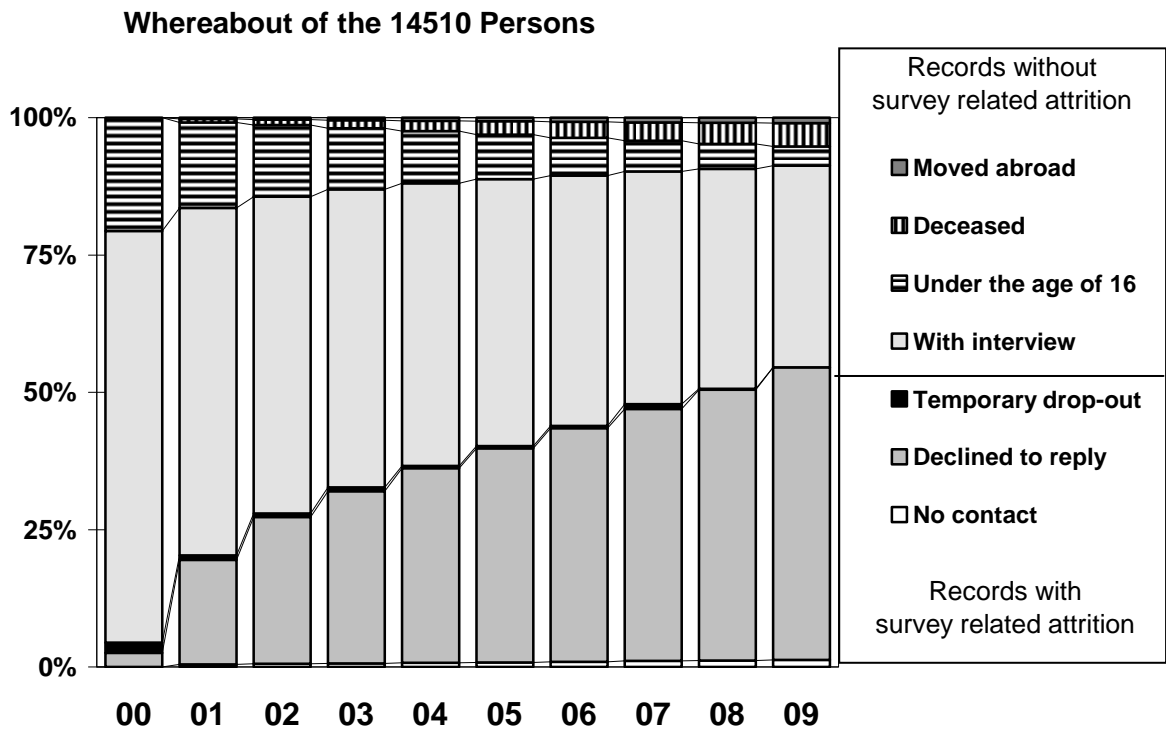


Figure 15: All First Wave Persons (Gross Subsample G). Development up to Wave 8.

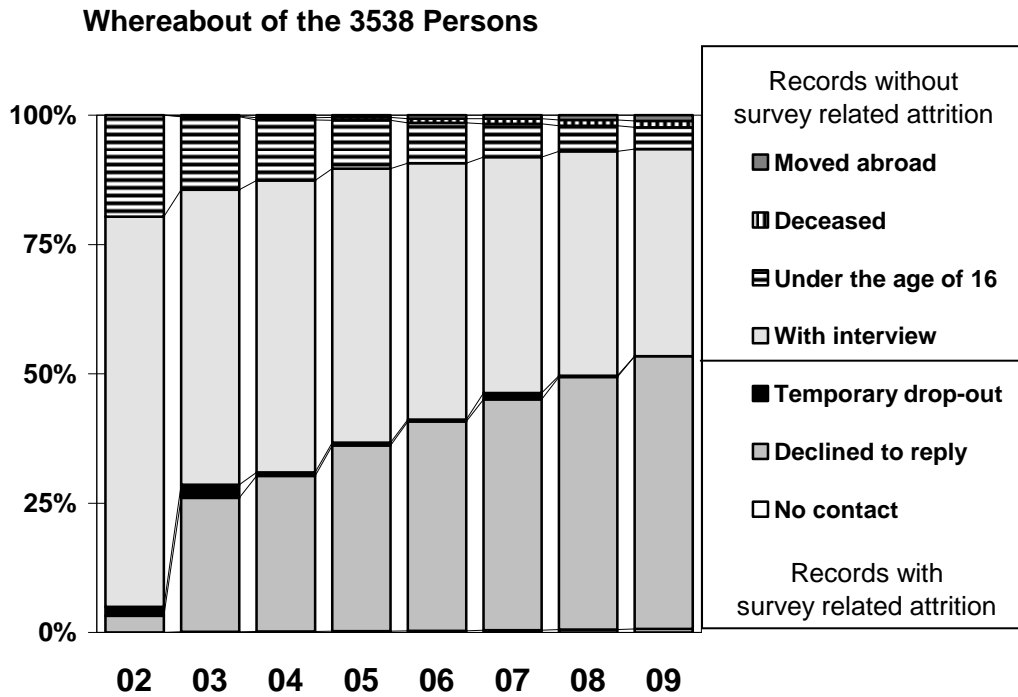
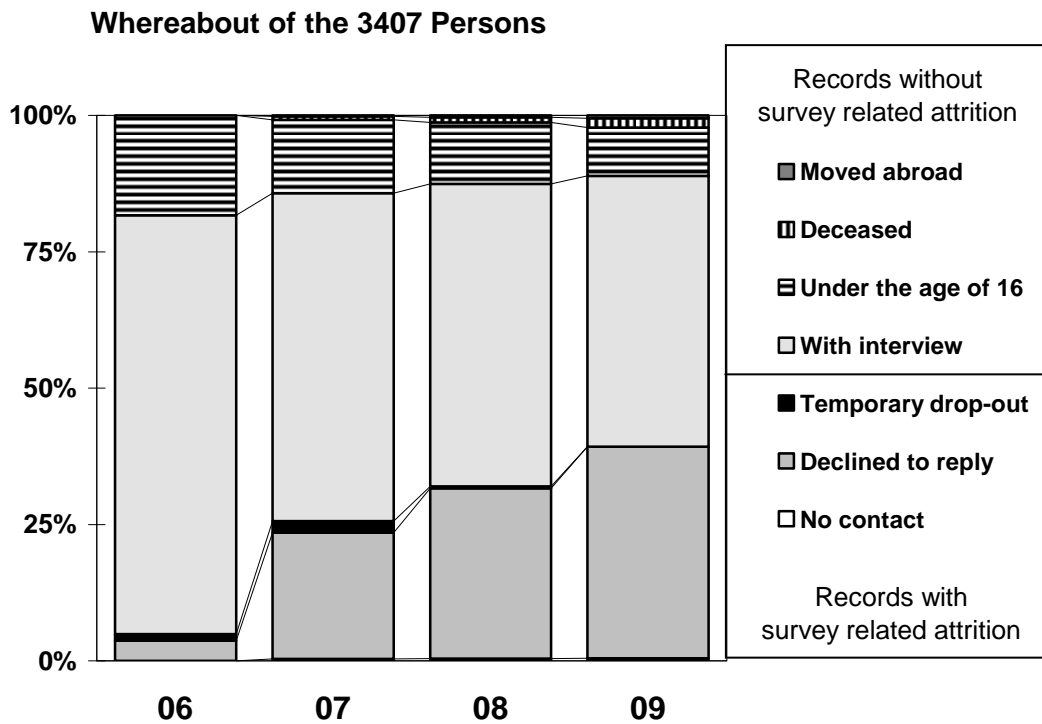


Figure 16: All First Wave Persons (Gross Subsample H). Development up to Wave 4.



2.3 New Entrants through Birth or Move into SOEP Households and Their Participation Behavior

The following figures display the participation behavior of the non-original sample members and their entrance to the ongoing survey, distinguishing between continuation of participation, exits due to survey unrelated attrition, and exits due to survey-related attrition.

Figure 17: Entrants Who Were Born or Moved into SOEP Households and Their Participation Behavior in Subsamples A and B

Figure 18: Entrants Who Were Born or Moved into SOEP Households and Their Participation Behavior in Subsample C

Figure 19: Entrants Who Were Born or Moved into SOEP Households and Their Participation Behavior in Subsample D

Figure 20: Entrants Who Were Born or Moved into SOEP Households and Their Participation Behavior in Subsample E

Figure 21: Entrants Who Were Born or Moved into SOEP Households and Their Participation Behavior in Subsample F

Figure 22: Entrants Who Were Born or Moved into SOEP Households and Their Participation Behavior in Subsample G

Figure 23: Entrants Who Were Born or Moved into SOEP Households and Their Participation Behavior in Subsample H

Figure 17: Entrants and their Participation Behavior (Subsamples A, B).

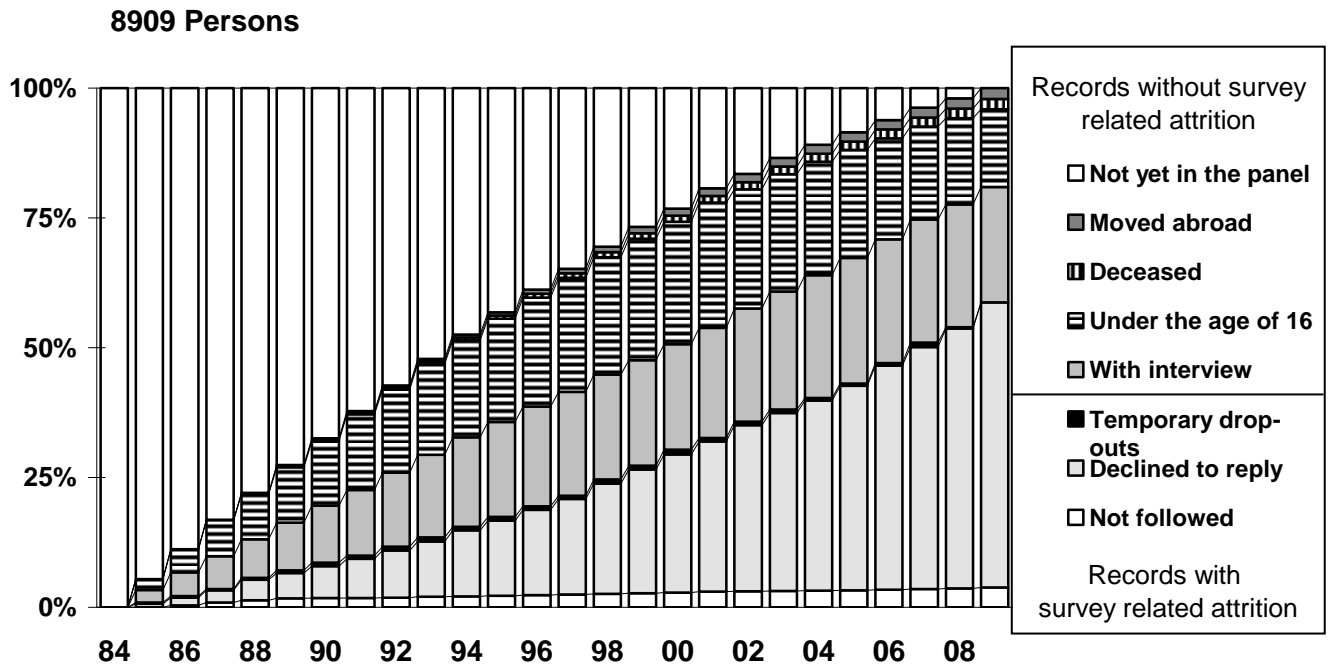


Figure 18: Entrants and their Participation Behavior (Subsample C).

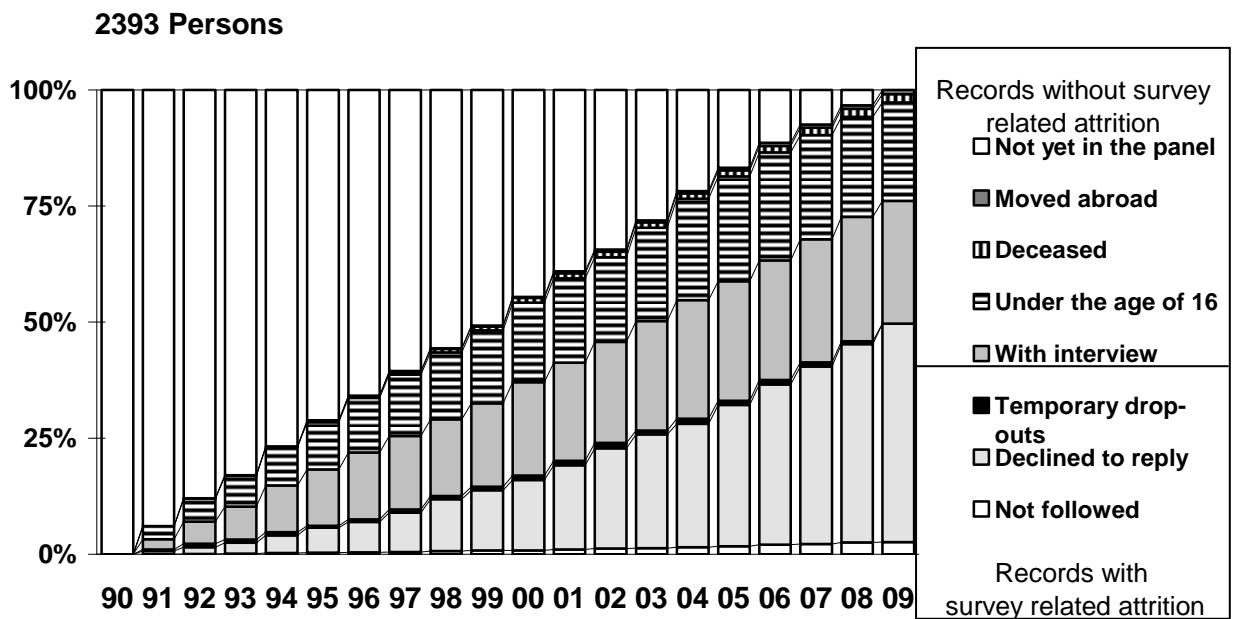


Figure 19: Entrants and their Participation Behavior (Subsample D).

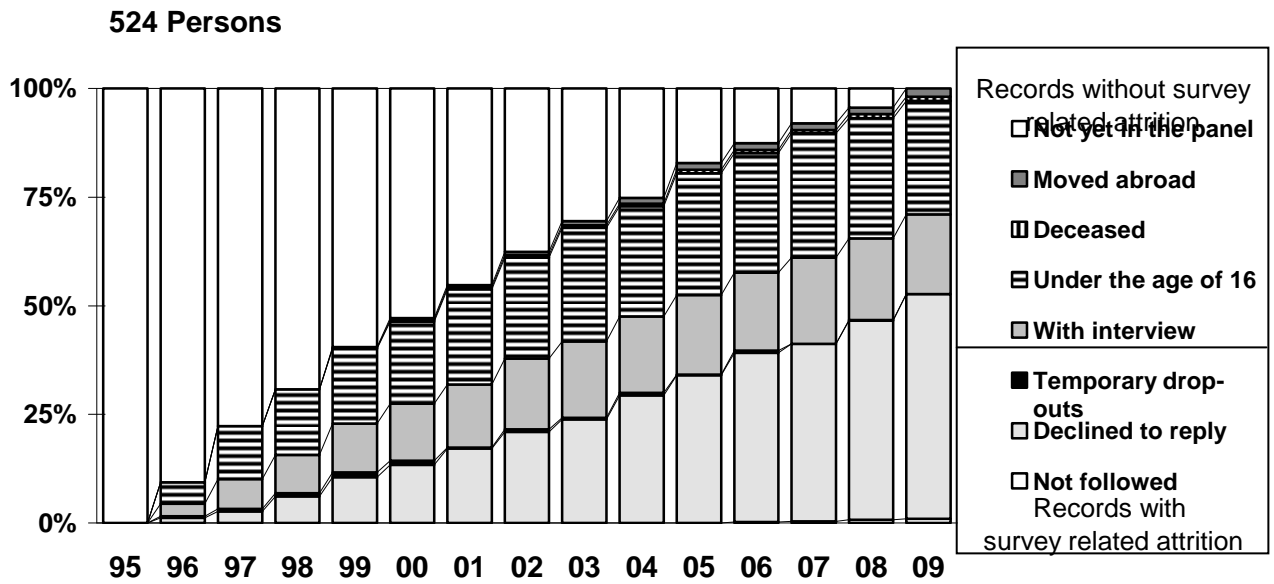


Figure 20: Entrants and their Participation Behavior (Subsample E).

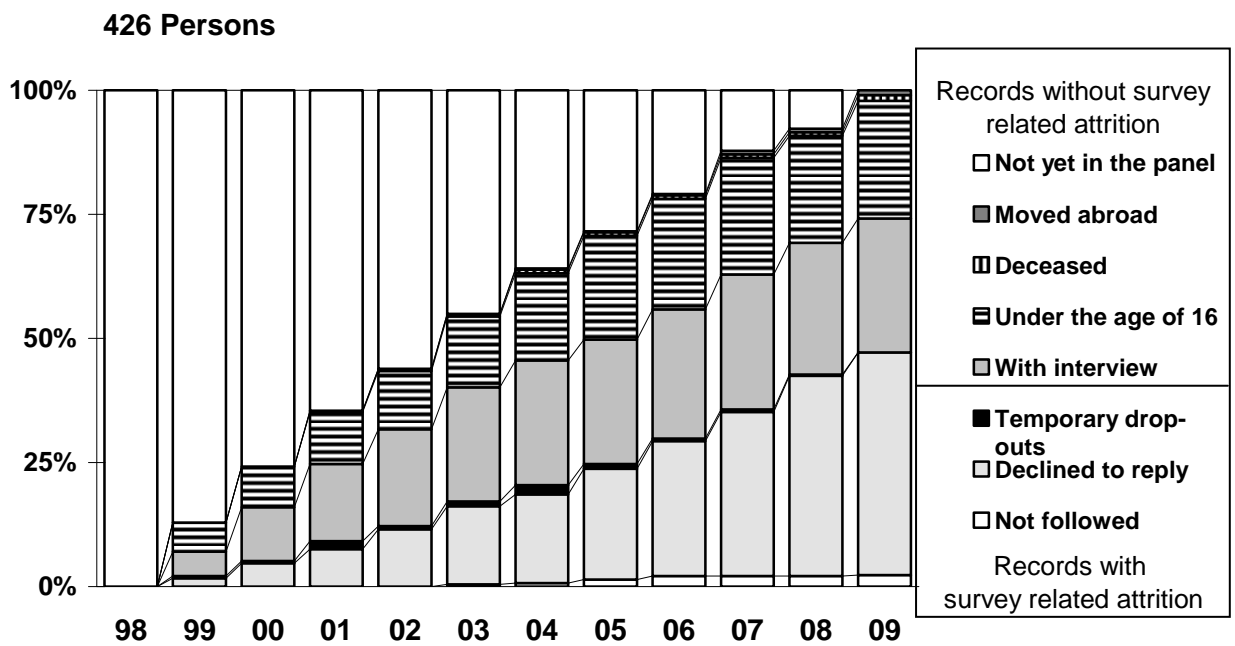


Figure 21: Entrants and their Participation Behavior (Subsample F).

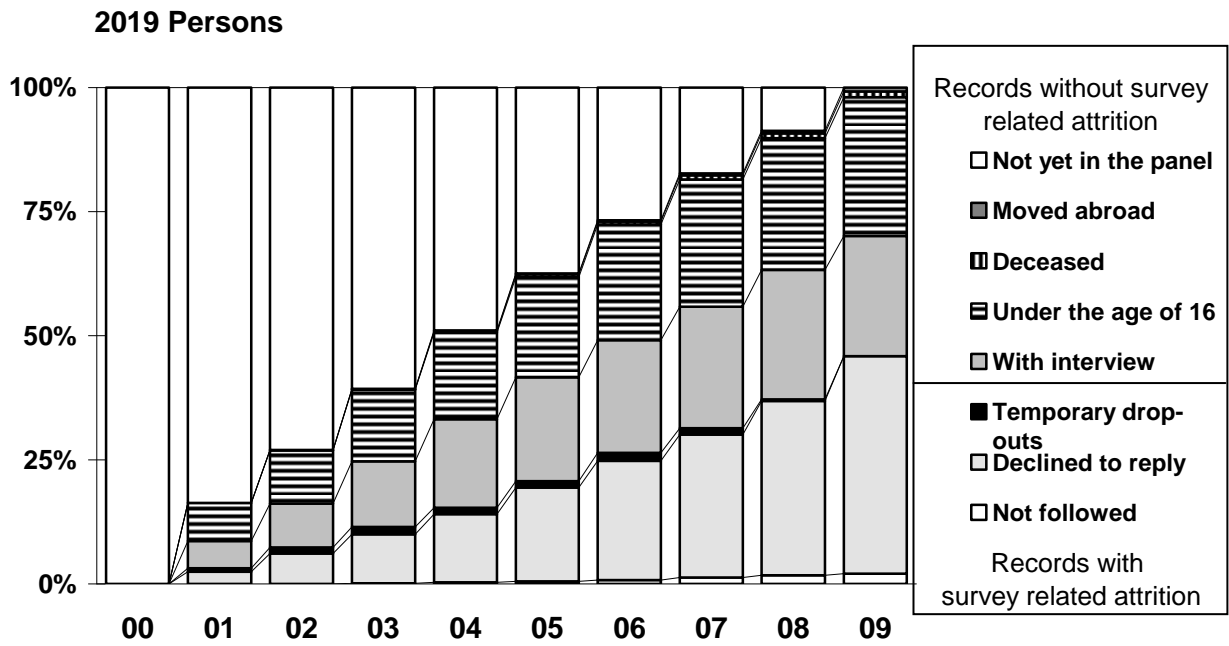


Figure 22: Entrants and their Participation Behavior (Subsample G).

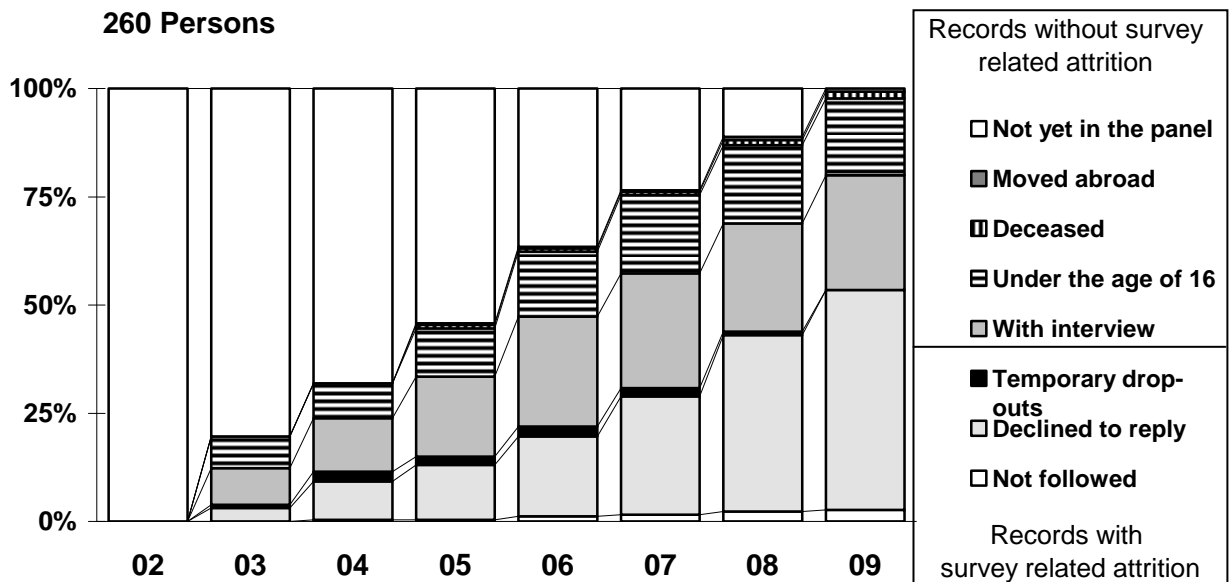
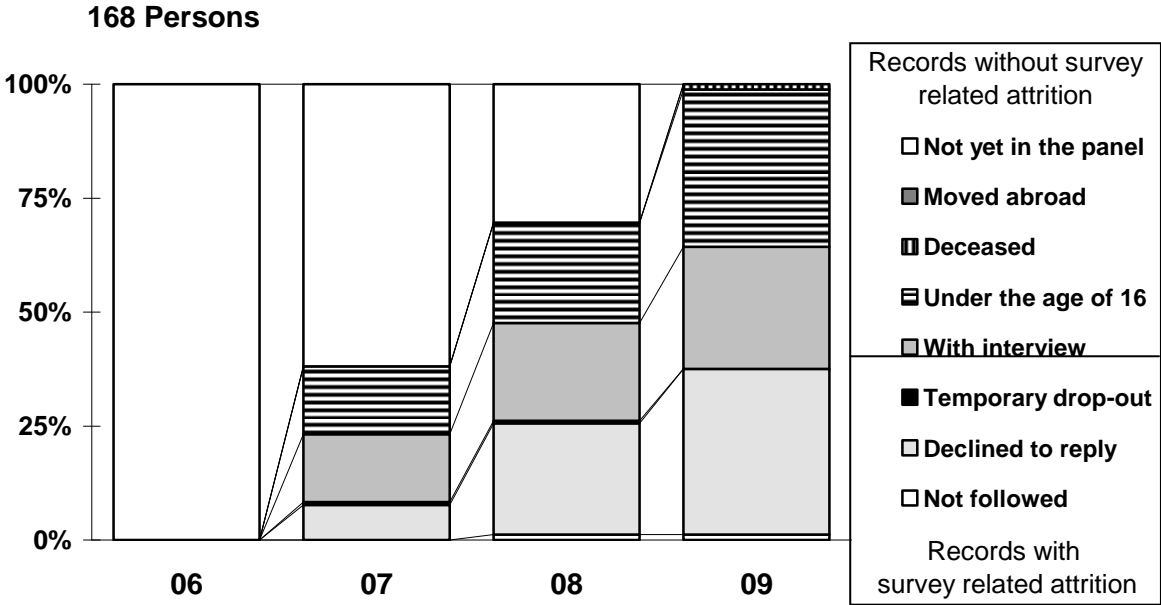


Figure 23: Entrants and their Participation Behavior (Subsample H).

Data Documentation 50

2 Developments in Sample Size



2.4 The Risk of Survey-Related Panel Attrition

The following figures display Kaplan-Meier estimates of the risk of survey related attrition (unsuccessful follow-up and refusal) of the net sample of first-wave respondents thereby ignoring survey unrelated exits (moves abroad and deaths). These figures stratify the drop-out risk in different groups of the sample defined by respondents' sample membership (Figures 24 and 25) and some basic socio-demographic characteristics measured in the year of sampling, such as age, occupation, income, and education (Figures 26 through 29). These unweighted figures show in general only moderate differences in the risk of survey related attrition between groups of the sample. Among the older samples A through C (Figure 24), for instance, first-wave respondents from sample B have a somewhat lower probability of remaining in the survey than respondents from sample A and C. In the more recent samples D through H (Figure 25), first-wave respondents from sample H have a somewhat lower probability of remaining in the survey than respondents from sample F. The latter in turn, have a lower probability of remaining in the survey than respondents from sample G.

Figure 24: Successful Re-Interviewing of First-Wave Respondents by Subsamples A, B, C.

Figure 25: Successful Re-Interviewing of First-Wave Respondents by Subsamples D, E, F, G, H.

Figure 26: Successful Re-Interviewing of All First-Wave Respondents by Age Categories.

Figure 27: Successful Re-Interviewing of All First-Wave Respondents by Occupation.

Figure 28: Successful Re-Interviewing of All First-Wave Respondents by Income Quintiles.

Figure 29: Successful Re-Interviewing of All First-Wave Respondents by Education.

Figure 24: Successful Re-Interviewing of First-Wave Respondents by Subsamples A, B, C. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad.

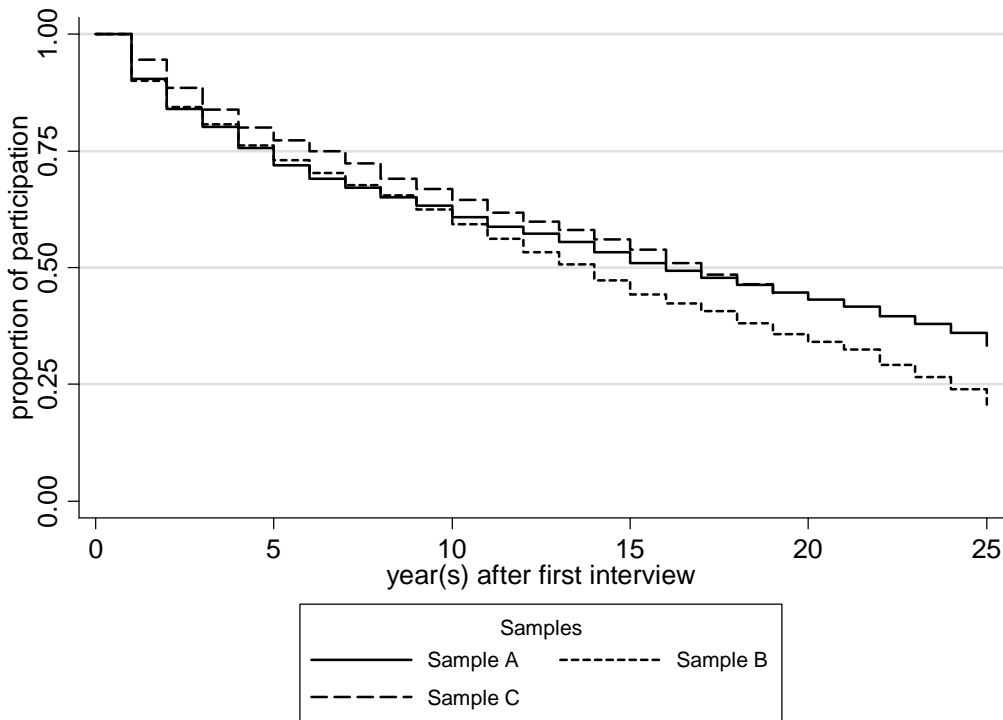


Figure 25: Successful Re-Interviewing of First-Wave Respondents by Subsamples D, E, F, G, H. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad.

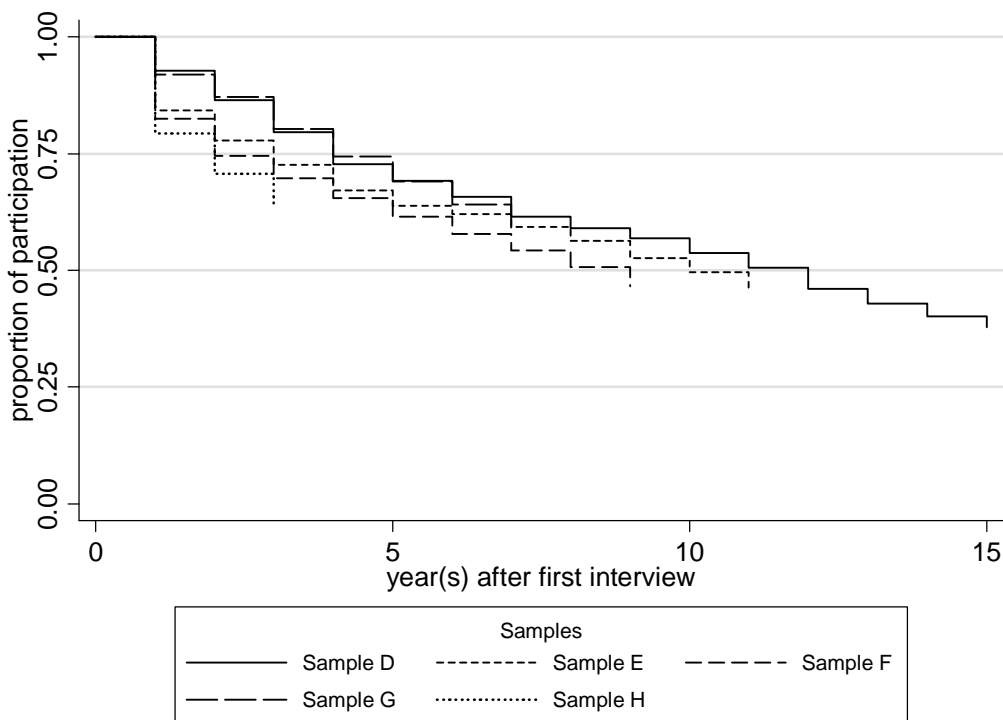


Figure 26: **Successful Re-Interviewing of All First-Wave Respondents by Age Categories. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad.**

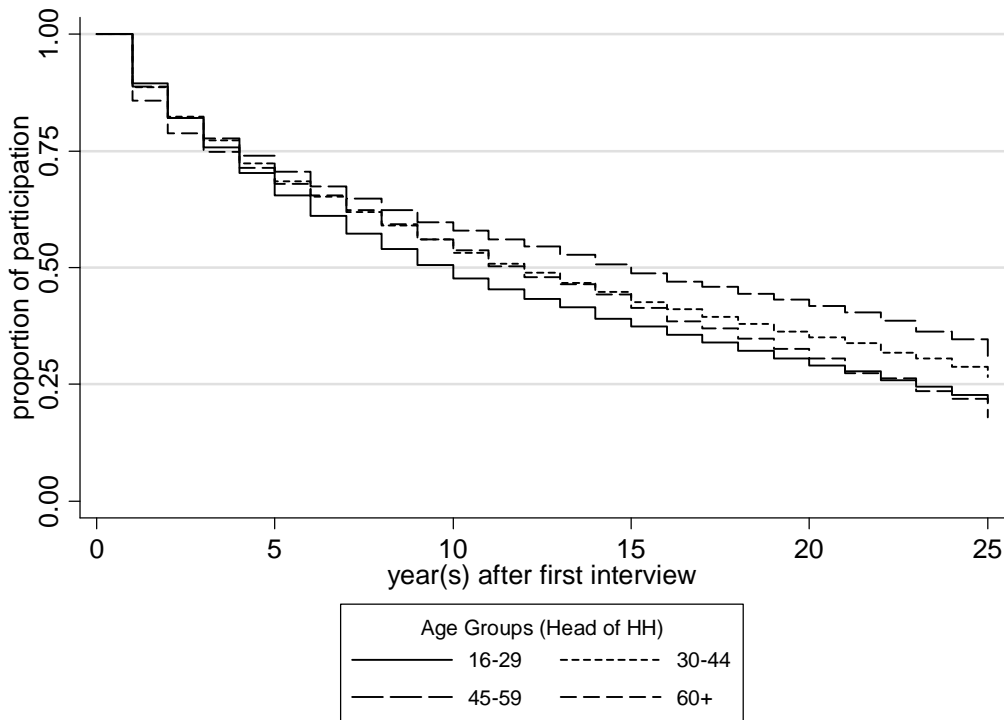


Figure 27: **Successful Re-Interviewing of All First-Wave Respondents by Occupation. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad.**

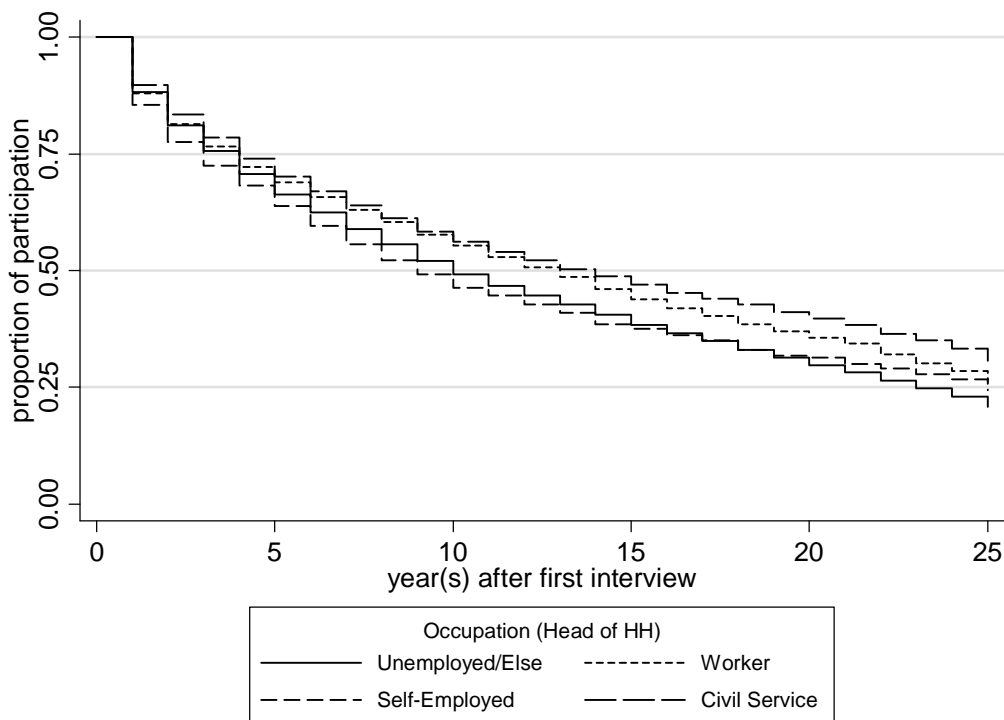


Figure 28: **Successful Re-Interviewing of All First-Wave Respondents by Income Quintiles. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad.**

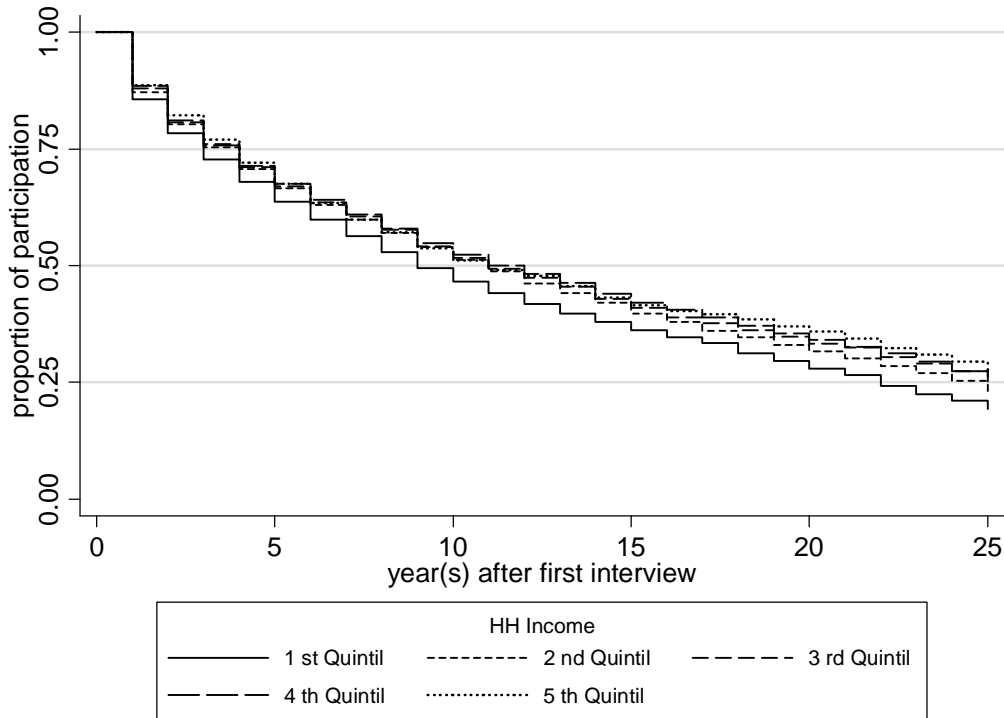
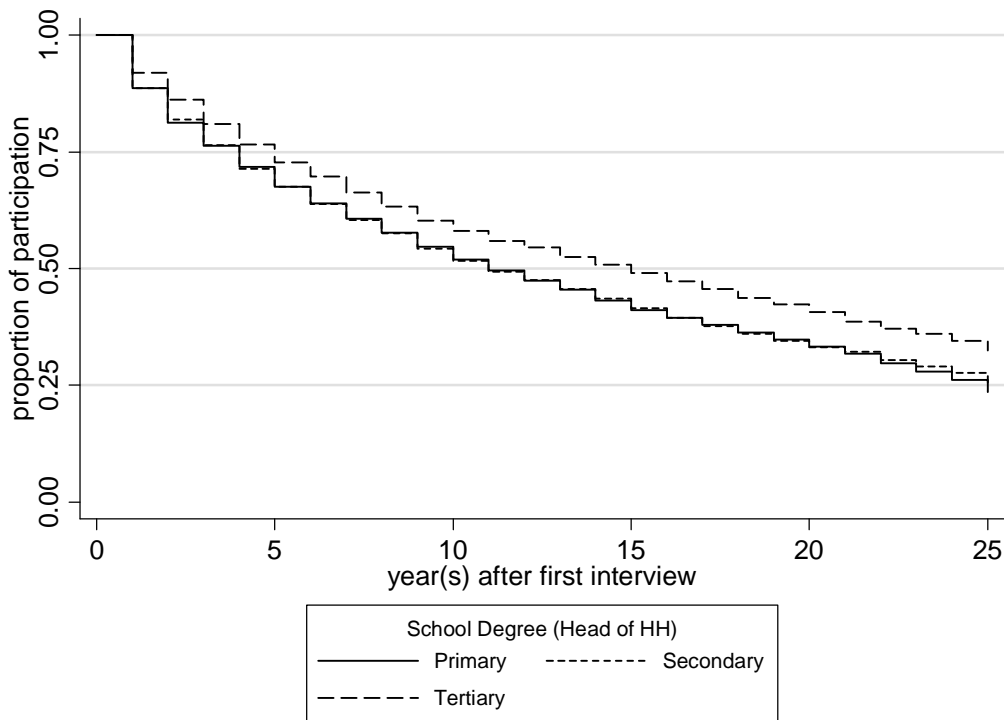


Figure 29: **Successful Re-Interviewing of All First-Wave Respondents by Education. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad.**



3 Panel Attrition Due to Unsuccessful Follow-Ups

In each panel wave, the first step in successful re-interviewing is the relocation of the households of the preceding wave. The fieldwork organization of the SOEP, TNS Infratest Sozialforschung, identifies whether (a) a household still lives at the old address, (b) an entire household has moved or all household members have died, (c) all household members have left the sampling area, and (d) all household members have returned to an existing panel household.

3.1 The Frequency of Successful Follow-Ups

Table 1 displays the number of households of the previous waves that need to be re-contacted and the relative frequency of successful follow-ups in subsamples A through H and waves 1985 through 2009. The re-contact rates refer to all households of the previous wave that still exist in the sampling area plus split-off households. A contact is regarded as successful if the interviewer documented a completed interview or refusal in the address protocol. Moreover, if former household members returned to an existing panel household, this is classified as a successful follow-up.

Data Documentation 50

3 Panel Attrition Due to Unsuccessful Follow-Ups

Table 1: The Frequency of Households to be Re-Contacted and the Relative Proportion of Successful Follow-Ups by Subsample and Year.

Year	A		B		C		D		E		F		G		H	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%
1985	4681	98.5	1370	96.9												
1986	4486	99.0	1325	97.4												
1987	4232	99.1	1220	98.7												
1988	4140	99.2	1191	99.1												
1989	3984	99.1	1157	99.1												
1990	3902	99.2	1124	98.9												
1991	3860	99.5	1151	99.3	2246	98.5										
1992	3845	99.7	1153	99.2	2304	99.5										
1993	3867	99.3	1172	98.7	2227	99.1										
1994	3849	99.3	1150	99.1	2136	99.4										
1995	3784	99.5	1108	99.0	2113	99.6										
1996	3747	99.7	1069	99.3	2104	99.5	544	99.6								
1997	3688	99.6	1038	99.1	2091	99.5	542	99.3								
1998	3667	99.4	1019	99.4	2081	99.4	498	99.4								
1999	3631	99.6	975	99.4	2041	99.7	529	99.1	1100	99.5						
2000	3549	99.6	934	99.5	2028	99.6	467	99.8	968	99.2						
2001	3463	99.6	904	99.5	2036	99.7	454	99.1	922	99.1	6172	99.0				
2002	3406	99.7	877	99.1	2010	99.5	450	99.8	875	99.4	5451	99.5				
2003	3330	99.6	840	99.6	1982	99.6	434	99.5	834	99.3	4965	99.7	1056	99.1		
2004	3260	99.8	803	99.6	1962	99.6	436	99.8	797	99.7	4736	99.6	1010	99.7		
2005	3220	99.8	779	99.4	1959	99.7	429	99.3	783	99.1	4577	99.7	1001	99.7		
2006	3138	99.7	770	99.6	1941	99.4	425	98.8	775	99.1	4401	99.3	995	99.5		
2007	3000	99.7	725	99.5	1834	99.9	387	99.5	727	99.7	4157	99.5	933	99.3	1530	99.5
2008	2856	99.7	676	99.2	1767	99.5	372	99.4	680	99.7	3962	99.4	904	99.6	1326	99.6
2009	2730	99.7	620	99.3	1695	99.8	351	99.7	636	100	3760	99.6	870	99.5	1145	99.7

n = Number of households to be recontacted

% = Percentage of households with successful recontact

3.2 Predicting the Probability of Successful vs. Unsuccessful Follow-Ups in the Year 2009

Based on the household and interview characteristics measured in 2008, we aim at predicting the probability of re-contacting a household relative to unsuccessful follow-up in 2009. Among a very large number of regressors that we tested in preliminary analyses, we identified a smaller number of variables that exert a robust effect on the probability of successful follow-ups ($p < 0.05$). Table 2 describes the regressors and Table 3 reports the subsample-specific estimates of logit models of the probability of re-contacting a household relative to unsuccessful follow-up.

Note that the estimates of regression models of the previous waves 1985 through 2008 are due to space restrictions not reported in the present data documentation, but can be obtained from previous attrition documentations.

Table 2: **Definition of the Regressors of the Logit Model of Unsuccessful Follow-Ups.**

Variable	Label	Value
New HH	New Split-off Household (HH) with New Address	0/1
Moved HH	Change in Address of an Existing HH	0/1
Temporary Drop-Out	Temporary Drop-Out of HH in Previous Year	0/1
East Germany	Sampling Region (0=West, 1=East)	
High Mobility Area	Neighborhood with High Levels of Mobility	0/1

Data Documentation 50

3 Panel Attrition Due to Unsuccessful Follow-Ups

Table 3: Estimates of Logit Models of the Probability of Re-Contacting a Household (Relative to Unsuccessful Follow-Up) in 2009.

	Sample A	Sample B	Sample C	Sample D	Sample E	Sample F	Sample G	Sample H
Intercept	-3.24 (0.62) ***	-2.27 (0.67) ***	-5.23 (0.72) ***	-5.50 (1.00) ***		-4.88 (0.28) ***	-4.68 (0.58) ***	-5.34 (0.62) ***
New HH	-2.29 (0.60) ***						-2.04 (0.58) ***	
Moved HH	-1.62 (0.59)***	-1.59 (0.62) **				-1.68 (0.28) ***		-1.63 (0.62) ***
Temporary Drop-Out		-2.48 (0.75) ***	-2.19 (0.72) ***					
East Germany	-1.08 (0.47) **							
High Mobility Area	-1.15 (0.57) **							
<i>Likelihood Ratio (Pr > 0.45)</i>		****	****	****		****	****	****

Note. *** p < 0.01; ** p < 0.05; * p < 0.10; standard errors in parentheses. **** The specified and the saturated models are the same.

4 Panel Attrition Due to Refusals

In each panel wave, the second step in successful re-interviewing after relocating households from the preceding wave is to obtain each household's confirmation of willingness to participate in the survey. We define successful re-interviewing relative only to survey-related panel attrition, such as refusals, and ignore survey-unrelated attrition, such as deaths and moves abroad, to generate the longitudinal weights.

4.1 The Frequency of Participation

Table 4 displays the participation rates due to refusal by sub-sample and wave. In reverse one can derive the corresponding drop-out rates. Note that we did not distinguish between various types of refusals such as unconditional refusals, refusals due to lack of time or health problems, etc.

Data Documentation 50

4 Panel Attrition Due to Refusals

Table 4: The Frequency of Re-Contacted Households and the Relative Proportion of Participation by Subsample and Year.

Year	A		B		C		D		E		F		G		H	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%	n	%
1985	4611	89.8	1326	89.1												
1986	4442	89.2	1290	87.4												
1987	4194	93.2	1204	92.7												
1988	4105	91.1	1180	90.8												
1989	3949	92.4	1146	91.0												
1990	3871	93.3	1111	92.5												
1991	3842	94.0	1143	92.4	2213	91.7										
1992	3833	93.5	1144	92.7	2290	88.2										
1993	3838	93.9	1156	92.0	2208	89.2										
1994	3821	93.6	1139	89.8	2122	92.3										
1995	3766	93.6	1097	89.5	2101	92.2	634	82.3								
1996	3734	93.3	1061	90.5	2092	93.3	542	91.9								
1997	3674	94.1	1029	90.5	2076	93.6	537	89.2								
1998	3645	92.9	1013	88.6	2066	91.3	523	84.3								
1999	3616	92.0	969	88.5	2030	93.3	495	85.9	1084	81.7						
2000	3535	91.7	929	88.3	2018	93.1	466	91.2	959	87.8						
2001	3448	91.9	899	90.0	2028	91.2	450	88.4	913	88.8	6109	80.4				
2002	3396	92.0	869	88.1	1996	91.1	449	89.5	868	89.1	5420	84.6				
2003	3318	92.6	837	88.6	1974	91.5	432	92.4	828	89.9	4951	88.6	1047	87.0		
2004	3253	92.5	800	89.3	1955	92.7	435	89.2	795	92.1	4719	89.7	1007	89.8		
2005	3214	91.4	774	90.2	1954	90.6	426	89.0	782	90.3	4564	89.2	998	88.1		
2006	3130	90.1	767	85.4	1930	89.0	420	85.7	768	89.3	4370	89.1	990	86.8		
2007	2992	91.0	721	85.2	1832	90.3	385	89.6	725	89.2	4138	89.3	926	89.0	1523	78.0
2008	2850	90.7	671	84.9	1759	90.5	370	88.6	678	88.8	3939	89.2	901	87.3	1321	81.9
2009	2723	89.0	616	81.2	1693	90.7	350	87.4	636	90.3	3746	88.2	866	87.4	1142	87.2

n = Number of re-contacted households

% = Percentage of households that participated

4.2 Predicting the Probability of Re-Interviewing versus Refusal in the Year 2009

Based on the household and interview characteristics measured in 2008, we aim at predicting the probability of agreement vs. refusal to participate in the survey by the households that were re-contacted in 2009. The individual attributes refer in many cases to the head of the household in the previous wave, but for split-off households the attributes refer to the person who moved out of the panel household (in the case of several persons, the first person mentioned in the address protocol).

As in the case of predicting successful follow-ups, we use only model specifications where all included regressors are significantly different from zero. The definition of the regressors is given in Table 5. Table 6 reports the subsample-specific estimates of logit models of the probability of participating relative to refusal. Note that the estimates of regression models of the previous waves 1985 through 2008 are not reported in the present data documentation due to space restrictions, but can be obtained from previous attrition reports.

Data Documentation 50

4 Panel Attrition Due to Refusals

Table 5: **Definition of the Regressors of the Logit Model of Refusal.**

Variable	Label	Value
First-Wave-HH	Household (HH) of the First Wave Sampling	0/1
New-HH	New Split-off HH with New Address	0/1
HH-Move	Existing HH with New Address	0/1
Additional Questioning 2	HH Has Participated in Two or Three Additional Questionings	0/1
Biography	Biography-Questionnaire Completed	0/1
Change in Interviewer	Change in Interviewer Between Last Waves	0/1
Non-Regular Interview	No Regular Personal Interview (e.g. interrupted)	0/1
Short Interview	Interview Duration 1- 15 Minutes (Short = 1)	0/1
Email Disclosed	Email Address Known	0/1
Phone Disclosed	Telephone Number Known	0/1
Partial Unit Nonresponse	One Person in HH did not Participate	0/1
Single HH	One Person Living in HH	0/1
4+ Person HH	Four or More Persons Living in HH	0/1
Gender	Gender of Head of HH (Male = 1)	0/1
Age 25 - 34	Head of HH Age 25- 34	0/1
Age 45 - 54	Head of HH Age 45- 54	0/1
(Age 25-34) * (Old HH)	Interaction Term between Respective Variables	0/1
(Age 45- 54) * (Old HH)	Interaction Term between Respective Variables	0/1
Unmarried	Head of HH Unmarried	0/1
Married	Head of HH Married	0/1
Separation of Couple	Separation of Unmarried Couple	0/1
Separation of Married C.	Separation of Married Couple	0/1
Good Health	Head of HH has Good Health-Status	0/1
Gripstrength-Test	Participation in Gripstrength Test	0/1
Hospital Stay	Head of HH Staying in Hospital in Previous Year	0/1
Expected Life-Satisfaction	Head of HH's Expected Lifesatisfaction in Next Year (High = 0, Low = 1)	0/1

Data Documentation 50

4 Panel Attrition Due to Refusals

Income Not Specified	No Information on HH-Income Available	0/1
Purchasing Power 25	HH in First Quartile of Purchasing-Power Distribution (Microm)	0/1
Purchasing Power 100	HH in Fourth Quartile of Purchasing-Power Distribution (Microm)	0/1
Blue-Collar Worker	Job-Status of Head of HH is Blue-Collar Worker	0/1
Employee	Job-Status of Head of HH is Employee	0/1
Unemployed	Job-Status of Head of HH is Unemployed	0/1
Secondary Education	Highest School Degree of Head of HH is Abitur	0/1
Voc. Edu. not Specified	No Information on Vocational Training of Head of HH	0/1
Low Political Interest	Head of HH Has Very Low or No Political Interest	0/1
East Germany	Sampling Region (West = 0, East = 1)	0/1
Urban Area	City has More than 100 000 Inhabitans	0/1
Poor Area	HH Located in Low Status Area	0/1
Affluent Area (Microm)	HH Located in Area with High Socio-Economic Status	0/1
Middle-Class Area	HH Located in Middle-Class Area	0/1
Hedonistic/Material. Area	HH Located in Hedonistic-Materialistic Area	0/1
Italian Migrant	HH from Italian Migrant Sub-Sample	0/1
Migrant Oversampling	HH from Migrant Sub-Sample	0/1

Data Documentation 50

4 Panel Attrition Due to Refusals

Table 6a: Estimates of Logit Models for the Probability of Re-Interviewing a Household (Relative to Refusal) in 2009.

	Sample A	Sample B	Sample C	Sample D	Sample E	Sample F	Sample G	Sample H
Intercept	0.84 (0.13) ***	1.65 (0.3) ***	1.10 (0.15) ***	0.69 (0.49) ns	0.25 (0.30) ns	1.38 (0.08) ***	0.03 (0.31)ns	1.06 (0.23) ***
First Wave HH						-0.19 (0.08) **		
New HH	-0.50 (0.17) ***	-1.10 (0.34) **	-2.26 (0.64) ***			-0.47 (0.15) ***		-1.02 (0.35) ***
HH Move			-2.08 (0.59) ***			-0.21 (0.10) **		
Additional Questioning 2								0.80 (0.40) **
Biography							-0.86 (0.24) ***	
Change in Interviewer	-0.67 (0.12) ***	-1.2 (0.21) ***			-0.60 (0.28) **	-0.48 (0.10) ***	-0.43 (0.18) **	
Non-Regular Interview	-0.50 (0.06) ***	-1.31 (0.18) ***	-0.62 (0.08) ***		-0.75 (0.17) ***	-1.05 (0.07) ***	-0.90 (0.13) ***	-1.73 (0.15) ***
Short Interview		-0.33 (0.15) **						-0.31 (0.11) ***
Email Disclosed	0.28 (0.07) ***		0.23 (0.10) **					
Phone Disclosed	0.26 (0.11) **	0.45 (0.18) **		1.53 (0.43) ***	1.16 (0.28) ***		0.52 (0.25) **	0.53 (0.22) **
Partial Unit Nonresponse							0.85 (0.30) ***	-0.67 (0.23) ***
Single HH				1.28 (0.52) **				
4+ Person HH	-0.18 (0.08) **							-0.34 (0.13) ***
Gender		-0.29 (0.13)**						
Age 25-34			1.54 (0.61) **			-0.26 (0.08) ***		
Age 45- 54			1.26 (0.61) **					
(Age 25-34)*(Old-HH)			-1.84 (0.62) ***					
(Age45- 54)*(Old-HH)			-1.33 (0.62) **					

Note. *** p < 0.01; ** p < 0.05; * p < 0.10; standard errors in parentheses.

Data Documentation 50

4 Panel Attrition Due to Refusals

Table 6b: Estimates of Logit Models for the Probability of Re-Interviewing a Household (Relative to Refusal) in 2009.

	Sample A	Sample B	Sample C	Sample D	Sample E	Sample F	Sample G	Sample H
Unmarried	-0.18 (0.07) **							
Married								-0.77 (0.28) ***
Separation of Couple			-0.77 (0.30) ***					
Separation of Married C.						-0.38 (0.14) ***		
Good Health	0.18 (0.07) **						0.36 (0.15) **	
Hospital Stay								-0.35 (0.14) **
Gripstrength-Test					-0.32 (0.15) **			
Expected Life-Satisfaction	-0.25 (0.1) *		-0.27 (0.11) **		-0.61 (0.23) ***			
Income Not Specified	-0.28 (0.12) **							
Purchasing Power 25			-0.30 (0.09) ***					
Purchasing Power 100								-0.27 (0.12) **
Blue-Collar Worker					0.53 (0.27) **			
Employee						0.12 (0.06) **		
Unemployed					0.73 (0.36) **			
Secondary Education				-0.64 (0.29) **			0.27 (0.12) **	0.30 (0.12) **
Voc. Ed. not Specified				-0.65 (0.29) **				
Low Political Interest				-0.76 (0.36) **				

Note. *** p < 0.01; ** p < 0.05; * p < 0.10; standard errors in parentheses.

Data Documentation 50

4 Panel Attrition Due to Refusals

Table 6c: Estimates of Logit Models for the Probability of Re-Interviewing a Household (Relative to Refusal) in 2009.

	Sample A	Sample B	Sample C	Sample D	Sample E	Sample F	Sample G	Sample H
East Germany			0.44 (0.11) ***					
Rural Area				-0.77 (0.29) **				
Urban Area		-0.30 (0.13) **						
Poor Area						0.14 (0.06) ***		
Affluent Area	-0.14 (0.06) **							
Middle-Class Area			0.18 (0.09) **				0.26 (0.11) **	
Hedonistic/Material. Area		-0.31 (0.13) **						
Italian Migrant		0.41 (0.17) **						
Migrant Oversampling						-0.41 (0.14) ***		
<i>Likelihood Ratio (Pr > Chisq)</i>	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001	< 0.0001

Note. *** p < 0.01; ** p < 0.05; * p < 0.10; standard errors in parentheses.

5 Summary Statistics of the Derived Longitudinal and Cross-Sectional Weights

Based on the regression models of successful vs. unsuccessful recontacts and agreements vs. refusals to participate, we derive two sets of predicted probabilities, the product of which is the household's "staying probability". The inverse of this probability of staying in the SOEP in 2009 based on characteristics measured in 2008, ZHBLEIB, lends itself as a longitudinal weighting variable correcting for selective attrition between waves 2008 and 2009. Table 7 reports some sub-sample specific descriptive statistics of the longitudinal weights in each wave.

The product of the cross-sectional weight in 2008, YHHRF, and the longitudinal weight in 2009, ZHBLEIB, provide the raw data for the cross-sectional weight in 2009. In a final step, the post-stratification of the cross-sectional weights corrects them to meet benchmarks of known marginals of the underlying population in 2009. This is at the level of households the region of residence (Bundesland), urbanization, household size, and home ownership and on the level of persons age, gender, nationality, and household structure. Table 8 reports sub-sample-specific descriptive statistics of the derived cross-sectional weighting variable ZHHRF and in comparison all previous cross-sectional weights AHHRF through YHHRF.

Data Documentation 50

5 Summary Statistics of the Derived Longitudinal and Cross-Sectional Weights

Table 7a: Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples A through D (Percentiles of \$HBLEIB up to Wave 26).

	bhbleib	chbleib	dhbleib	ehbleib	fhbleib	ghbleib	hhbleib	ihbleib	jhbleib	khbleib	lhbleib	mhbleib	nhbleib	ohbleib	phbleib	qhbleib	rhbleib	shbleib	thbleib	uhbleib	vhbleib	whbleib	xhbleib	yhbleib	zhbleib
<u>sample A</u>																									
p10	1.06	1.04	1.03	1.02	1.03	1.02	1.02	1.01	1.01	1.02	1.01	1.01	1.01	1.02	1.02	1.02	1.02	1.01	1.01	1.01	1.02	1.01	1.01	1.02	1.02
p50	1.1	1.07	1.03	1.04	1.04	1.02	1.02	1.02	1.01	1.02	1.01	1.03	1.02	1.03	1.02	1.02	1.02	1.02	1.03	1.01	1.02	1.04	1.03	1.05	1.05
p90	1.22	1.26	1.13	1.2	1.16	1.11	1.09	1.11	1.16	1.15	1.16	1.12	1.13	1.14	1.2	1.15	1.18	1.21	1.14	1.12	1.16	1.22	1.14	1.13	1.25
N	4141	3962	3910	3743	3647	3612	3613	3585	3603	3577	3526	3485	3458	3387	3325	3240	3168	3123	3072	3010	2937	2821	2723	2584	2423
<u>sample B</u>																									
p10	1.09	1.1	1.03	1.03	1.03	1.04	1.03	1.01	1.02	1.03	1.02	1.04	1.02	1.04	1.04	1.03	1.02	1.04	1.01	1.04	1.05	1.01	1.03	1.01	1.01
p50	1.1	1.1	1.03	1.04	1.04	1.04	1.03	1.03	1.03	1.05	1.05	1.04	1.04	1.07	1.04	1.03	1.02	1.04	1.03	1.04	1.05	1.05	1.07	1.07	1.05
p90	1.26	1.29	1.14	1.22	1.14	1.12	1.16	1.16	1.22	1.22	1.29	1.21	1.29	1.23	1.22	1.18	1.23	1.37	1.31	1.13	1.17	1.33	1.24	1.25	1.6
N	1181	1128	1116	1071	1043	1028	1056	1060	1064	1023	982	960	931	898	858	820	809	766	742	714	698	655	614	570	500
<u>sample C</u>																									
p10							1.03	1.06	1.03	1.02	1.03	1.01	1.02	1.02	1.01	1.01	1.02	1.01	1.01	1	1	1.01	1	1.01	1
p50							1.06	1.06	1.04	1.04	1.03	1.02	1.04	1.02	1.03	1.03	1.02	1.02	1.03	1.01	1.02	1.04	1.03	1.03	1.03
p90							1.18	1.22	1.17	1.12	1.11	1.15	1.12	1.2	1.1	1.13	1.16	1.21	1.14	1.12	1.15	1.24	1.16	1.18	1.21
N							2030	2020	1970	1959	1938	1951	1942	1886	1894	1879	1850	1818	1807	1813	1771	1717	1654	1592	1535
<u>sample D</u>																									
p10												1	1.05	1.08	1.05	1.02	1.03	1	1.01	1	1	1.03	1.01	1.02	1
p50												1.08	1.09	1.08	1.05	1.02	1.03	1.02	1.01	1.01	1.02	1.04	1.04	1.07	1.02
p90												1.16	1.09	1.35	1.27	1.1	1.18	1.21	1.09	1.25	1.34	1.44	1.12	1.22	1.16
N												396	340	308	300	302	286	289	290	277	273	261	248	231	220

Data Documentation 50

5 Summary Statistics of the Derived Longitudinal and Cross-Sectional Weights

Table 7b: Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples E through H (Percentiles of \$HBLEIB up to Wave 26).

	bhbleib	chbleib	dhbleib	ehbleib	fhbleib	ghbleib	hhbleib	ihbleib	jhbleib	khbleib	lhbleib	mhbleib	nhbleib	ohbleib	phbleib	qhbleib	rhbleib	shbleib	thbleib	uhbleib	vhbleib	whbleib	xhbleib	yhbleib	zhbleib	
<u>sample E</u>																										
p10																1	1.03	1.01	1.01	1.04	1	1.01	1	1.01	1	1
p50																1.23	1.07	1.05	1.02	1.04	1.01	1.03	1.03	1.01	1.01	1.04
p90																1.47	1.21	1.25	1.2	1.15	1.08	1.18	1.21	1.16	1.19	1.17
N																886	838	811	773	744	732	706	686	647	602	574
<u>sample F</u>																										
p10																		1.08	1.03	1.02	1.02	1.01	1.01	1.02	1.01	1.02
p50																		1.14	1.05	1.04	1.03	1.03	1.03	1.03	1.03	1.04
p90																		1.59	1.46	1.24	1.19	1.17	1.29	1.15	1.14	1.34
N																		4911	4586	4386	4235	4070	3895	3694	3513	3303
<u>sample G</u>																										
p10																			1.06	1.02	1.03	1	1.02	1.01	1.02	
p50																			1.1	1.03	1.06	1.04	1.05	1.03	1.04	
p90																			1.17	1.25	1.25	1.31	1.17	1.18	1.36	
N																			911	904	879	859	824	787	757	
<u>sample H</u>																										
p10																							1.04	1.01	1.01	
p50																							1.16	1.03	1.03	
p90																							1.46	1.18	1.22	
N																							1188	1082	996	

Data Documentation 50

5 Summary Statistics of the Derived Longitudinal and Cross-Sectional Weights

Table 8: Summary Statistics of the Derived Cross-Sectional Weights at the Household Level (Percentiles of \$HHRF up to Wave 26).

	ahhrf	bhhrf	chhrf	dhhrf	ehhrf	fhhrf	ghhrf	hhhrf	ihhrf	jhhrf	khhrf	lhhrf	mhhrrf	nhhrf
p5	433.41	495.46	541.88	548.69	535.27	551.46	699.77	680.29	669.97	684.16	712.63	702.54	740.17	747.64
p10	600.48	681.65	757.63	795.13	804.65	822.26	1071.36	1048.48	1037.92	1054.62	1101.42	1112.20	1167.67	1204.37
p25	3775.97	3886.69	3615.83	3559.54	3574.10	3643.43	2216.64	2339.59	2331.49	2405.78	2410.23	2413.26	2408.34	2408.43
p50	4719.90	5081.43	5300.95	5405.57	5634.98	5855.85	4605.43	4697.75	4658.16	4690.17	4677.48	4376.16	4356.34	4331.58
p75	5657.32	6412.55	6825.70	7033.14	7544.82	7867.09	7042.49	7167.50	7141.61	7244.68	7279.06	6977.59	7004.35	7020.83
p90	7129.89	8460.74	9243.14	9583.02	10362.78	10851.10	9962.01	10265.52	10499.27	10755.37	11216.65	11069.13	11328.47	11886.85
p95	8305.07	10036.44	11123.83	11430.38	12537.17	13277.36	12363.62	12984.13	13642.81	14003.65	14604.14	14794.33	15257.40	15939.06
N	5921	5322	5090	5026	4814	4690	6819	6699	6665	6637	6559	6768	6699	6621

	ohhrf	phhrf	qhhrf	rhhrf	shhrf	thhrf	uhhrf	vhhrf	whhrf	xhhrf	yhhrf	zhhrf
p5	1057.20	1025.94	839.60	800.91	519.17	522.56	505.68	512.91	471.48	464.13	459.99	463.88
p10	1410.47	1387.50	1123.51	1083.51	696.87	706.24	698.61	705.26	667.67	668.38	667.28	664.18
p25	2364.65	2342.26	1757.51	1752.13	1275.65	1285.39	1260.76	1275.52	1281.83	1269.57	1286.98	1316.40
p50	3954.85	3986.41	2534.35	2742.57	2558.11	2557.68	2523.88	2536.77	2403.33	2458.21	2525.97	2595.46
p75	6235.95	6529.49	3572.20	4139.08	4169.73	4300.34	4390.05	4500.64	4136.26	4451.72	4731.54	4999.79
p90	9884.17	10807.73	5126.70	6039.75	6386.98	6753.57	7171.05	7480.33	6858.64	7591.85	8349.94	9233.39
p95	13169.40	14342.68	6448.46	7832.78	8181.11	9040.54	9880.52	10691.89	9680.28	10574.96	11557.25	12426.70
N	7492	7220	13091	11796	12320	11909	11644	11294	12361	11552	10921	10270

6 References

- Arbeitsgemeinschaft ADM-Stichproben und Bureau Wendt. 1994. Das ADM-Stichproben - System (Stand 1993). In: S. Gabler, S., J. Hoffmeyer-Zlotnik, J. and D. Krebs, D. (eds): *Gewichtung in der Umfragepraxis*. Westdeutscher Verlag, Opladen: 188-203.
- SOEP Group. 2001. The Socio-Economic Panel (SOEP) after more than 15 Years. *Vierteljahrshefte zur Wirtschaftsforschung* 1.
- Haisken-DeNew, John and Joachim Frick (eds.) 2001: *Desktop Companion to the Socio Economic Panel (SOEP)*. Berlin, DIW.
- Pischner, Rainer. 2007. Die Querschnittsgewichtung und die Hochrechnungsfaktoren des Sozio-oekonomischen Panels (SOEP) ab Release 2007 (Welle W): Modifikationen und Aktualisierungen. *DIW Data Documentation* 22.
- Rendtel, Ulrich. 1991. Die Schätzung von Populationswerten in Panelerhebungen. *Allgemeines Statistisches Archiv* 75: 225-244.
- Rendtel, Ulrich. 1995. *Panelausfälle und Panelrepräsentativität*. Frankfurt, Campus Verlag.
- Rendtel, Ulrich, Markus Pannenberg and Stefan Daschke. 1997. Die Gewichtung der Zuwanderer-Stichprobe des Sozio-ökonomischen Panels (SOEP). *Vierteljahrshefte zur Wirtschaftsforschung* 2: 271-286.
- Schupp, Jürgen and Gert G. Wagner. 1995. Die Zuwandererstichprobe des Sozio-oekonomischen Panels (SOEP). *Vierteljahrshefte zur Wirtschaftsforschung* 1.
- Spiess, Martin and Ulrich Rendtel. 2000. Combining an Ongoing Panel with a New Cross-Sectional Sample. *DIW Discussion Paper* 198.
- Spiess, Martin. 2001. Derivation of Design Weights: The Case of the Socio-Economic Panel (SOEP). *DIW Research Notes* 5.
- Spiess, Martin, Kroh, Martin, Pischner, Rainer, and Gert G. Wagner. 2008. On the Treatment of Non-Original Sample Members in the German Household Panel Study (SOEP) - Tracing, Weighting, and Frequencies. *SOEP Paper* 98.
- von Rosenblatt, B., SOEP-Gruppe, TNS Infratest Sozialforschung. 2004. *SOEP 2004 Methodenbericht zum Befragungsjahr 2004 (Welle 21) des Sozio-oekonomischen Panels*.