

66

Data Documentation

Documentation of Sample Sizes and Panel Attrition in the German Socio Economic Panel (SOEP) (1984 until 2011)

IMPRESSUM

© DIW Berlin, 2012

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 66

Martin Kroh*

Documentation of Sample Sizes and Panel Attrition in the German Socio Economic Panel (SOEP) (1984 until 2011)

Berlin, September 2012

* DIW Berlin, Socio-Economic Panel Study. mkroh@diw.de
I would like to thank Peter Eckert 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	26
3.1 The Frequency of Successful Follow-Ups	26
3.2 Predicting the Probability of Successful vs. Unsuccessful Follow-Ups in the Year 2011	28
4 Panel Attrition Due to Refusals.....	32
4.1 The Frequency of Participation.....	32
4.2 Predicting the Probability of Re-Interviewing versus Refusal in the Year 2011	34
5 Summary Statistics of the Derived Longitudinal and Cross-Sectional Weights	44
6 References.....	48

Tables

Table 1: The Frequency of Households to be Re-Contacted and the Relative Proportion of Successful Follow-Ups by Subsample and Year.	27
Table 2: Definition of the Regressors of the Logit Model of Unsuccessful Follow-Ups.	28
Table 3a: Estimates of Logit Models of the Probability of Re-Contacting a Household (Relative to Unsuccessful Follow-Up) in 2011.	30
Table 3b: Estimates of Logit Models of the Probability of Re-Contacting a Household (Relative to Unsuccessful Follow-Up) in 2011.	31
Table 4: The Frequency of Re-Contacted Households and the Relative Proportion of Participation by Subsample and Year.	33
Table 5: Definition of the Regressors of the Logit Model of Refusal.	35
Table 6a: Estimates of Logit Models for the Probability of Re-Interviewing a Household (Relative to Refusal) in 2011.	39
Table 6b: Estimates of Logit Models for the Probability of Re-Interviewing a Household (Relative to Refusal) in 2011.	40
Table 6c: Estimates of Logit Models for the Probability of Re-Interviewing a Household (Relative to Refusal) in 2011.	41
Table 6d: Estimates of Logit Models for the Probability of Re-Interviewing a Household (Relative to Refusal) in 2011.	42
Table 7b: Summary Statistics of the Derived Longitudinal Weights at the Household Level for Subsamples E through I (Percentiles of \$HBLEIB up to Wave 28).	46

Figures

Figure 1: The Number of Successful Interviews with Persons by Subsamples A through I, Waves 1 to 28.	4
Figure 2: Comparison of Successful Interviews with Persons and Households (Subsamples A and B), Waves 1 to 28.	5
Figure 3: Comparison of Successful Interviews with Persons and Households (Subsample C), Waves 1 to 22.	6
Figure 4: Comparison of Successful Interviews with Individuals and Households (Subsample D), Waves 1 to 17.	7
Figure 5: Comparison of successful interviews with individuals and households (Subsample E), Waves 1 to 14.	8
Figure 6: Comparison of Successful Interviews with Individuals and Households (Subsample F), Waves 1 to 12.	9
Figure 7: Comparison of Successful Interviews with Individuals and Households (Subsample G), Waves 1 to 10.	10
Figure 8: Comparison of Successful Interviews with Individuals and Households (Subsample H), Waves 1 to 6.	10
Figure 9: All First-Wave Persons (Gross Subsample A). Development up to Wave 28.	11
Figure 10: All First Wave Persons (Gross Subsample B). Development up to Wave 28.	12
Figure 11: All First Wave Persons (Gross Subsample C). Development up to Wave 22.	12
Figure 12: All First Wave Persons (Gross Subsample D). Development up to Wave 17.	13
Figure 13: All First Wave Persons (Gross Subsample E). Development up to Wave 14.	13
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 10.	14
Figure 16: All First Wave Persons (Gross Subsample H). Development up to Wave 6.	15
Figure 17: Entrants and their Participation Behavior (Subsample A).	17
Figure 18: Entrants and their Participation Behavior (Subsample B).	17
Figure 19: Entrants and their Participation Behavior (Subsample C).	18
Figure 20: Entrants and their Participation Behavior (Subsample D).	18
Figure 21: Entrants and their Participation Behavior (Subsample E).	19
Figure 22: Entrants and their Participation Behavior (Subsample F).	19
Figure 23: Entrants and their Participation Behavior (Subsample G).	20
Figure 24: Entrants and their Participation Behavior (Subsample H).	20

Figure 25: **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 26: **Successful Re-Interviewing of First-Wave Respondents by Subsamples D, E, F. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad.....** 22

Figure 27: **Successful Re-Interviewing of First-Wave Respondents by Subsamples G, H, I. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad.....** 23

Figure 28: **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 29: **Successful Re-Interviewing of All First-Wave Respondents by Occupation. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad.** 24

Figure 30: **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 31: **Successful Re-Interviewing of All First-Wave Respondents by Education. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad.** 25

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 28 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 J individually.

The SOEP study 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) and Schonlau et al. (2011). 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 2011 based on the characteristics of households measured in 2010. 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 2011: BBHBLEIB and BBPBLEIB. Based on the inverse of the probability of observing households and persons in 2010, the staying probability in 2011, and additional post-stratification to meet benchmarks of known marginals of the underlying population in 2011, we derive the cross-sectional weights BBHHRF and BBPHRF. 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

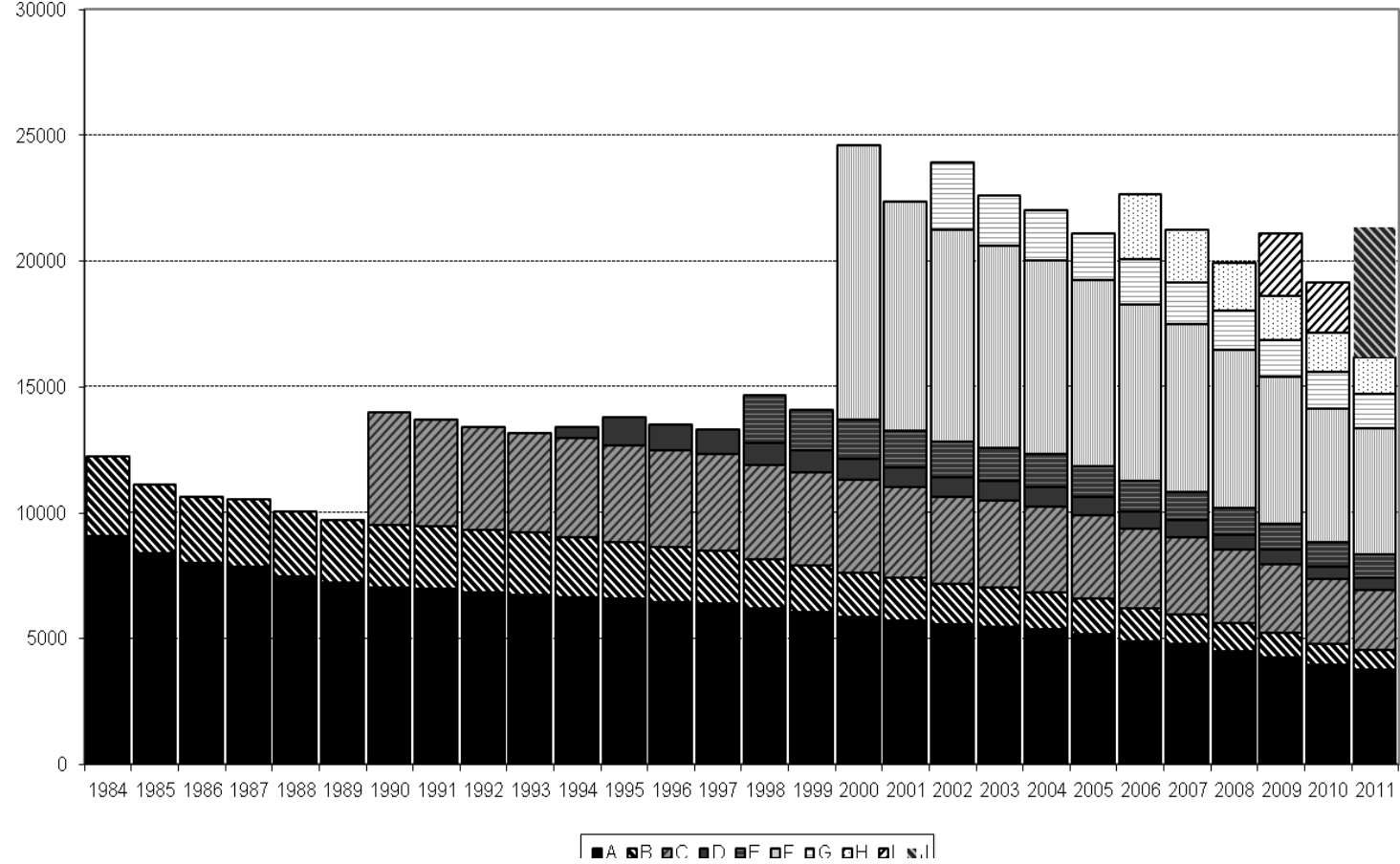
The following figures display the number of successful interviews considering different aspects:

- Figure 1** The Number of Successful Interviews with Persons by Subsamples A through I, Waves 1 to 28 (1984-2011)
- Figure 2** Comparison for Individuals and Households in Subsamples A and B, Waves 1 to 28 (1984 – 2011).
- Figure 3** Comparison for Individuals and Households in Subsample C, Waves 1 to 22, (1990–2011).
- Figure 4** Comparison for Individuals and Households in Subsample D, Waves 1 to 17, (1995–2011).
- Figure 5** Comparison for Individuals and Households in Subsample E, Waves 1 to 14, (1998–2011).
- Figure 6** Comparison for Individuals and Households in Subsample F, Waves 1 to 12, (2000–2011).
- Figure 7** Comparison for Individuals and Households in Subsample G, Waves 1 to 10, (2002-2011).
- Figure 8** Comparison for Individuals and Households in Subsample H, Waves 1 to 6, (2006-2011).

Data Documentation 66

2 Developments in Sample Size

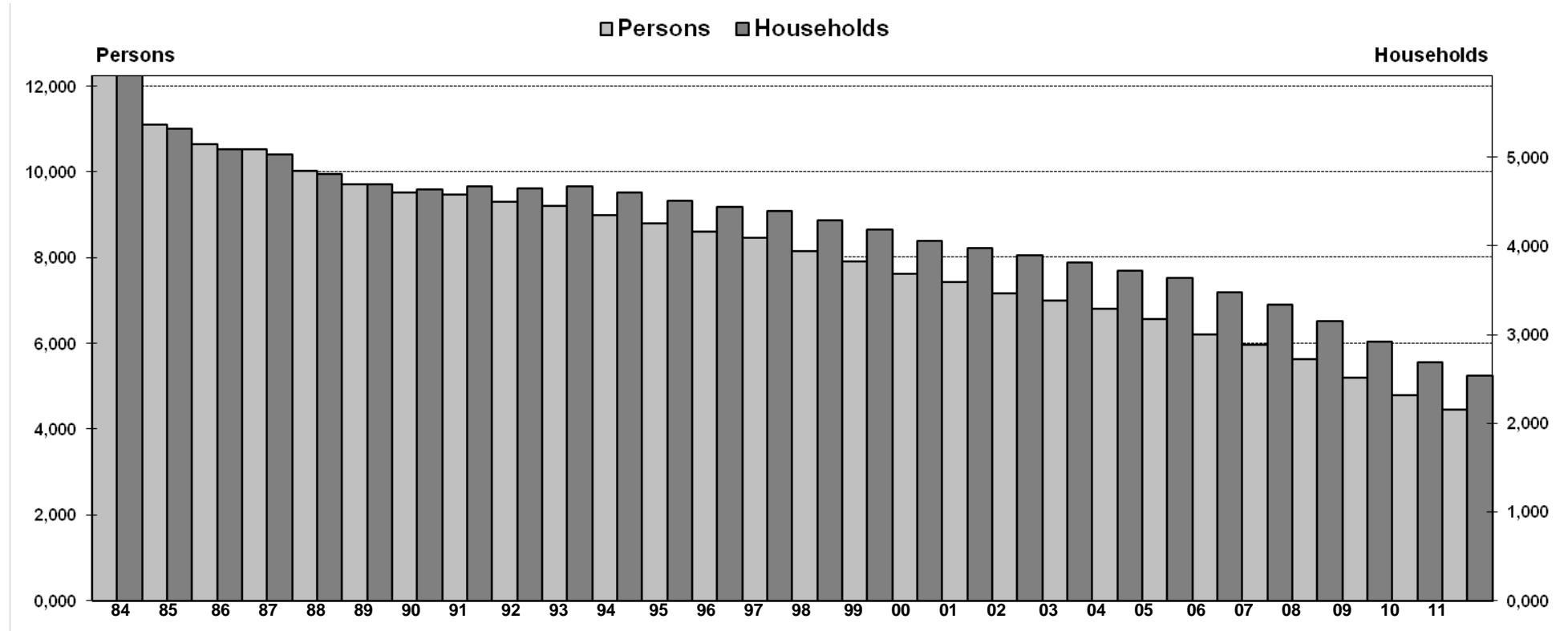
Figure 1: The Number of Successful Interviews with Persons by Subsamples A through J, Waves 1 to 28.



Data Documentation 66

2 Developments in Sample Size

Figure 2: Comparison of Successful Interviews with Persons and Households (Subsamples A and B), Waves 1 to 28

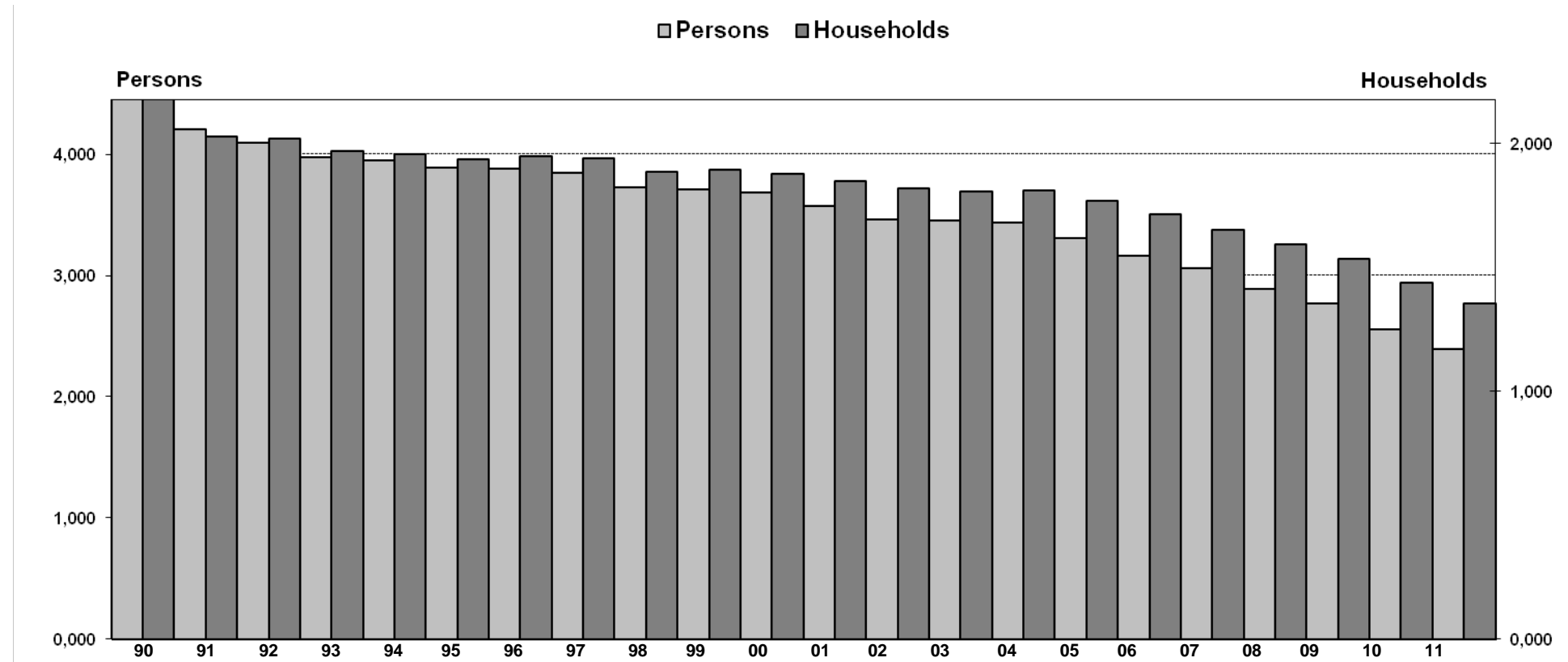


Year	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Persons	12,245	11,090	10,646	10,516	10,023	9,710	9,519	9,467	9,305	9,206	9,001	8,798	8,606	8,467	8,145	7,909	7,623	7,424	7,175	6,999	6,809	6,572	6,198	5,957	5,619	5,197	4,739	4,451
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	2,686	2,539

Data Documentation 66

2 Developments in Sample Size

Figure 3: Comparison of Successful Interviews with Persons and Households (Subsample C), Waves 1 to 22.

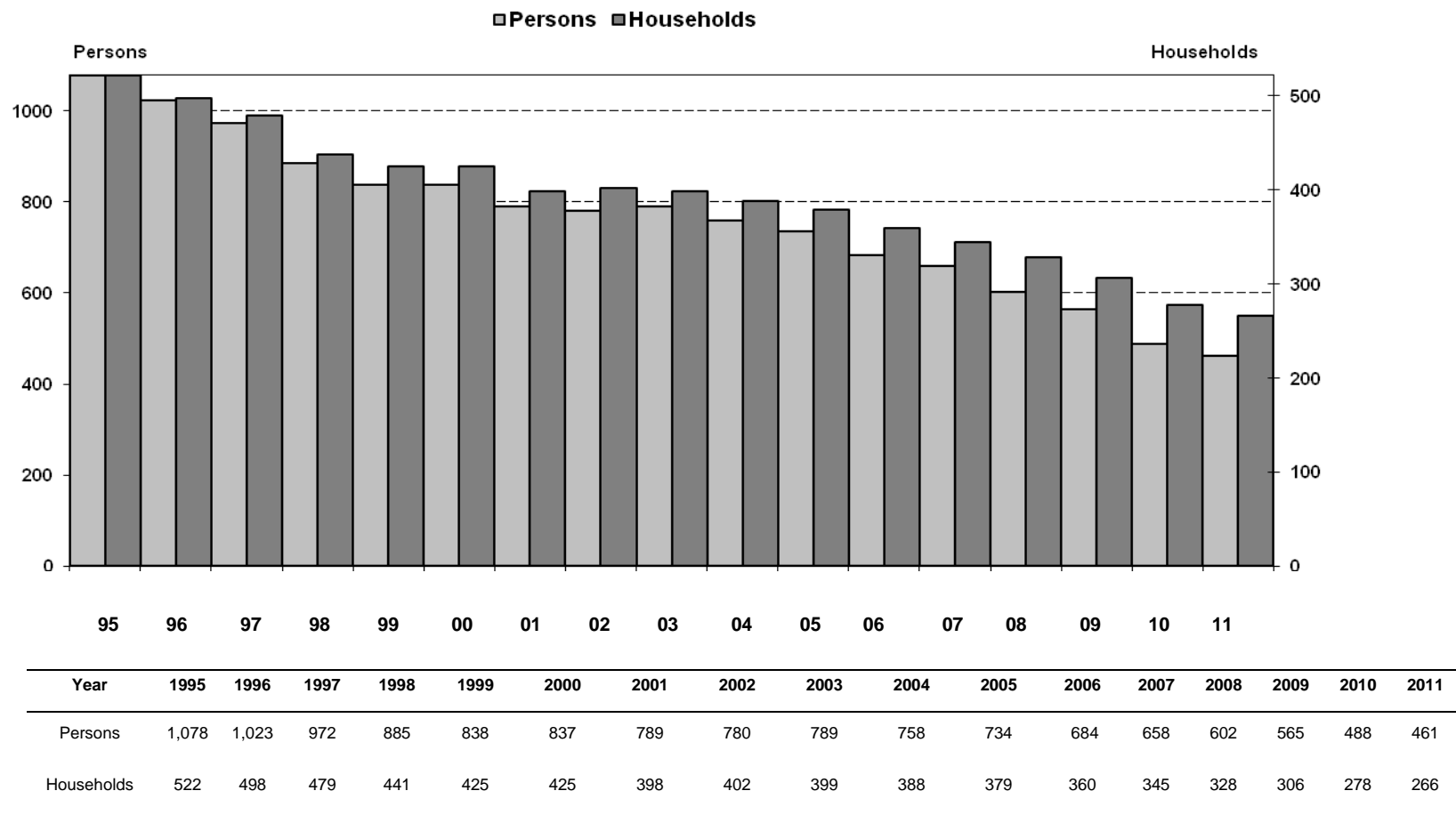


Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Persons	4,453	4,202	4,092	3,973	3,945	3,892	3,882	3,844	3,730	3,709	3,687	3,576	3,466	3,453	3,435	3,304	3,159	3,063	2,889	2,769	2,559	2,392
Households	2,179	2,030	2,020	1,970	1,959	1,938	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	1,437	1,355

Data Documentation 66

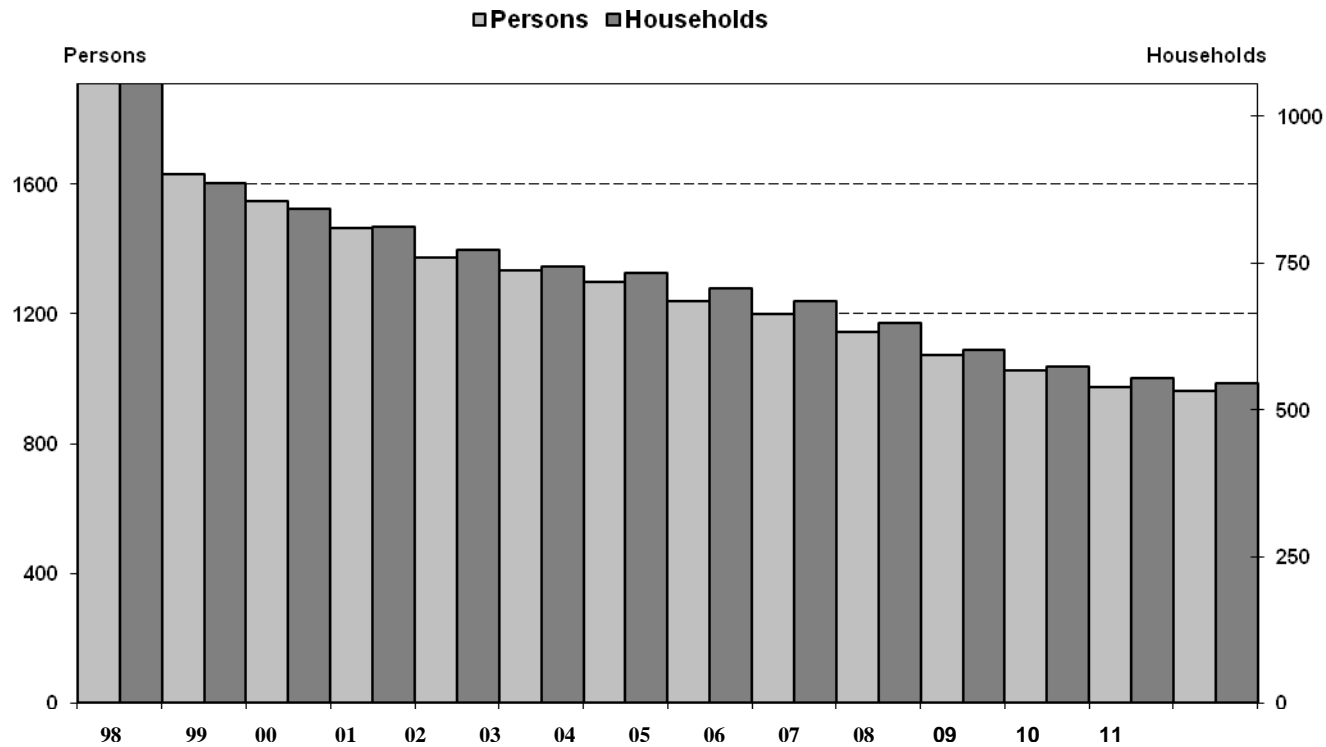
2 Developments in Sample Size

Figure 4: Comparison of Successful Interviews with Individuals and Households (Subsample D), Waves 1 to 17.



Data Documentation 66
2 Developments in Sample Size

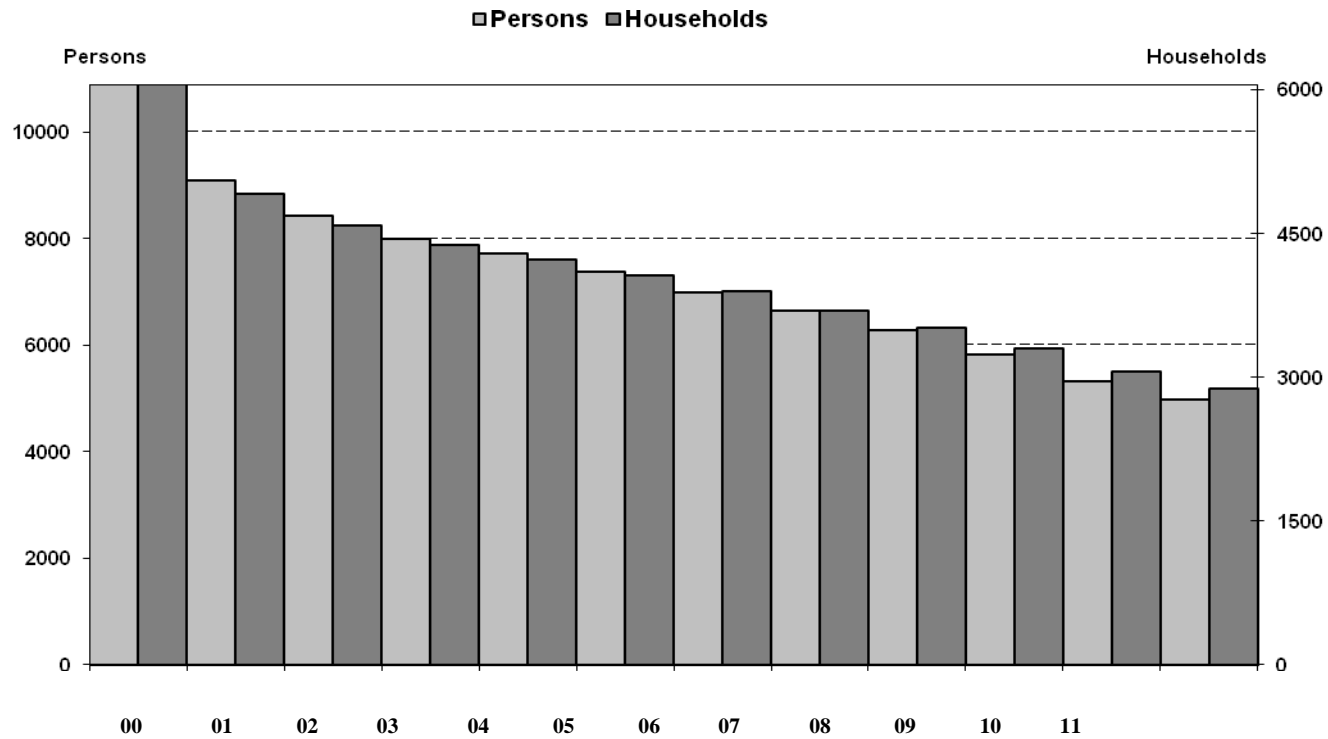
Figure 5: **Comparison of successful interviews with individuals and households (Subsample E), Waves 1 to 14.**



Year	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
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	975	961
Households	1,056	886	842	811	773	744	732	706	686	647	602	574	553	545

Data Documentation 66
2 Developments in Sample Size

Figure 6: Comparison of Successful Interviews with Individuals and Households (Subsample F), Waves 1 to 12.



Year	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Persons	1,0890	9,098	8,427	8,006	7,724	7,371	6,986	6,640	6,274	5,824	5,316	4,984
Households	6,052	4,911	4,586	4,386	4,234	4,070	3,895	3,694	3,513	3,303	3,055	2,885

Data Documentation 66

2 Developments in Sample Size

Figure 7: Comparison of Successful Interviews with Individuals and Households (Subsample G), Waves 1 to 10.

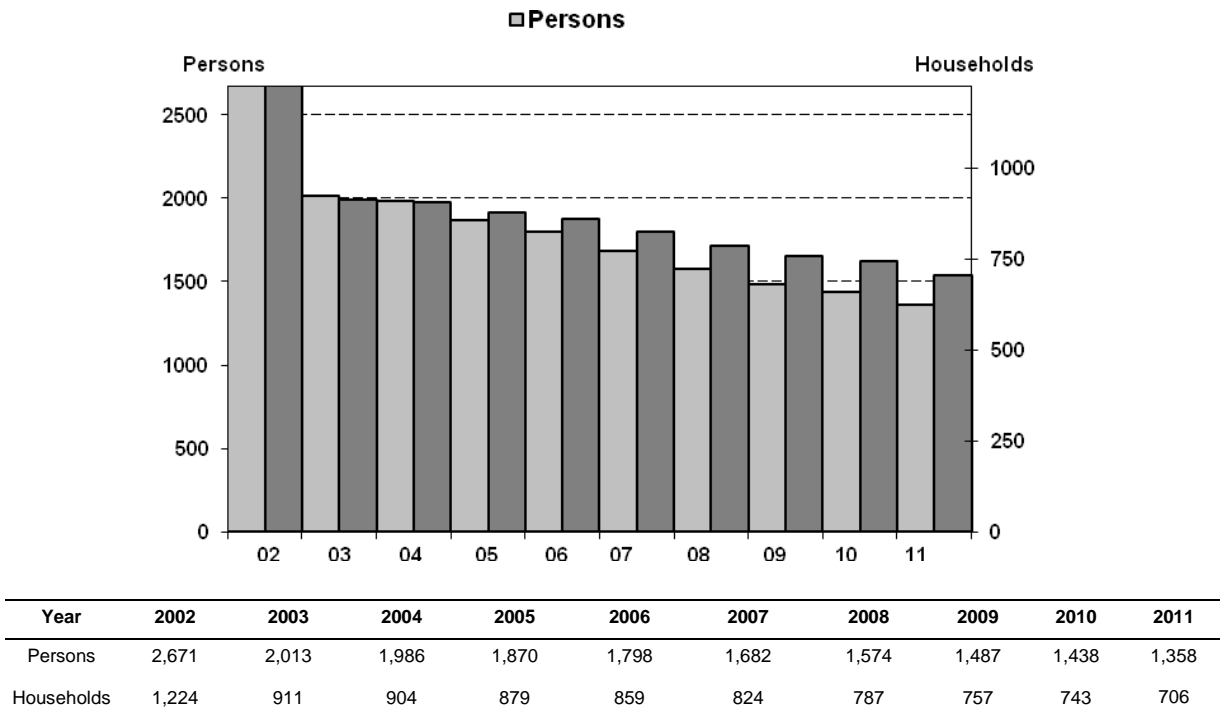
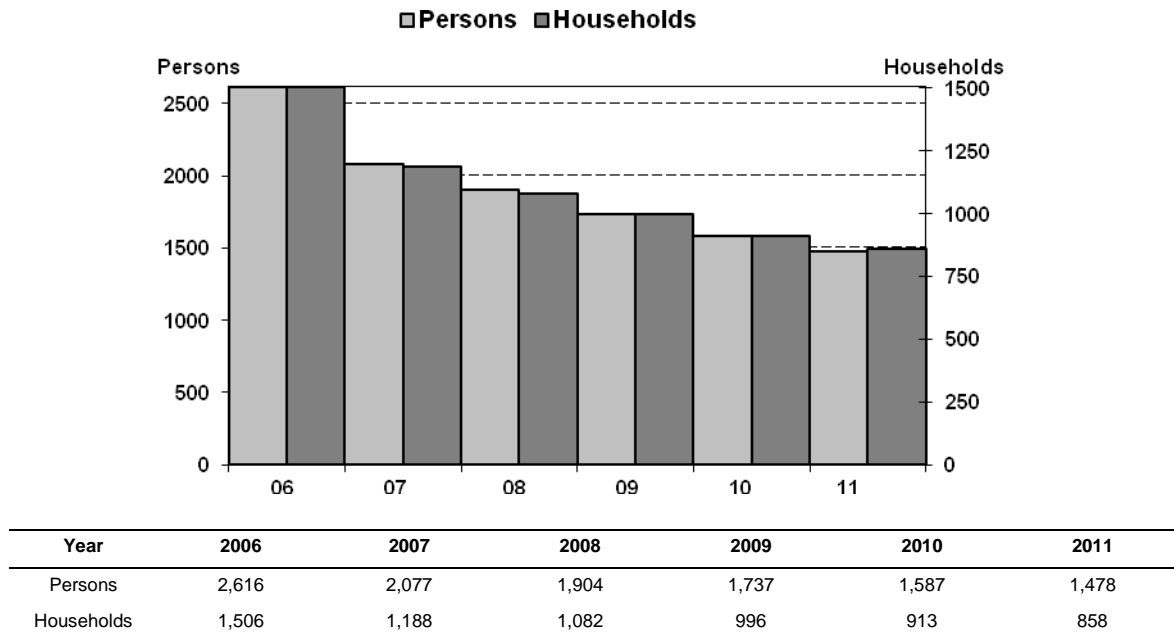


Figure 8: Comparison of Successful Interviews with Individuals and Households (Subsample H), Waves 1 to 6.



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 28.
- Figure 10:** All First Wave Persons in Subsample B. Whereabouts up to Wave 28.
- Figure 11:** All First Wave Persons in Subsample C. Whereabouts up to Wave 22.
- Figure 12:** All First Wave Persons in Subsample D. Whereabouts up to Wave 16.
- Figure 13:** All First Wave Persons in Subsample E. Whereabouts up to Wave 14.
- Figure 14:** All First Wave Persons in Subsample F. Whereabouts up to Wave 12.
- Figure 15:** All First Wave Persons in Subsample G. Whereabouts up to Wave 10.
- Figure 16:** All First Wave Persons in Subsample H. Whereabouts up to Wave 6.

Figure 9: All First-Wave Persons (Gross Subsample A). Development up to Wave 28.

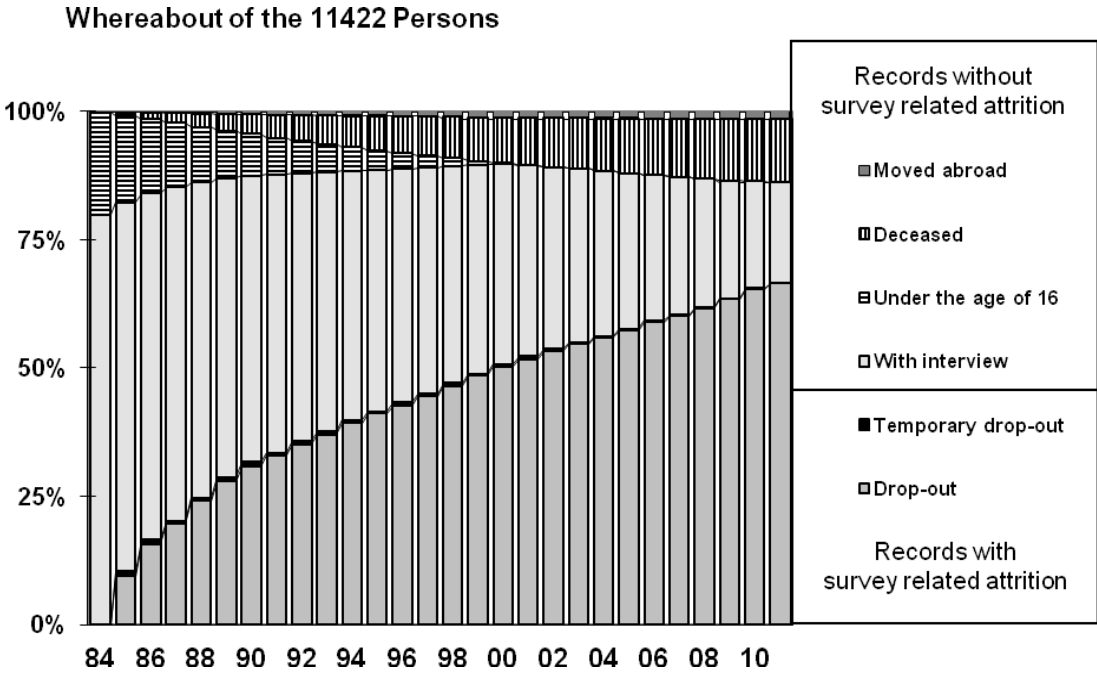


Figure 10: All First Wave Persons (Gross Subsample B). Development up to Wave 28.

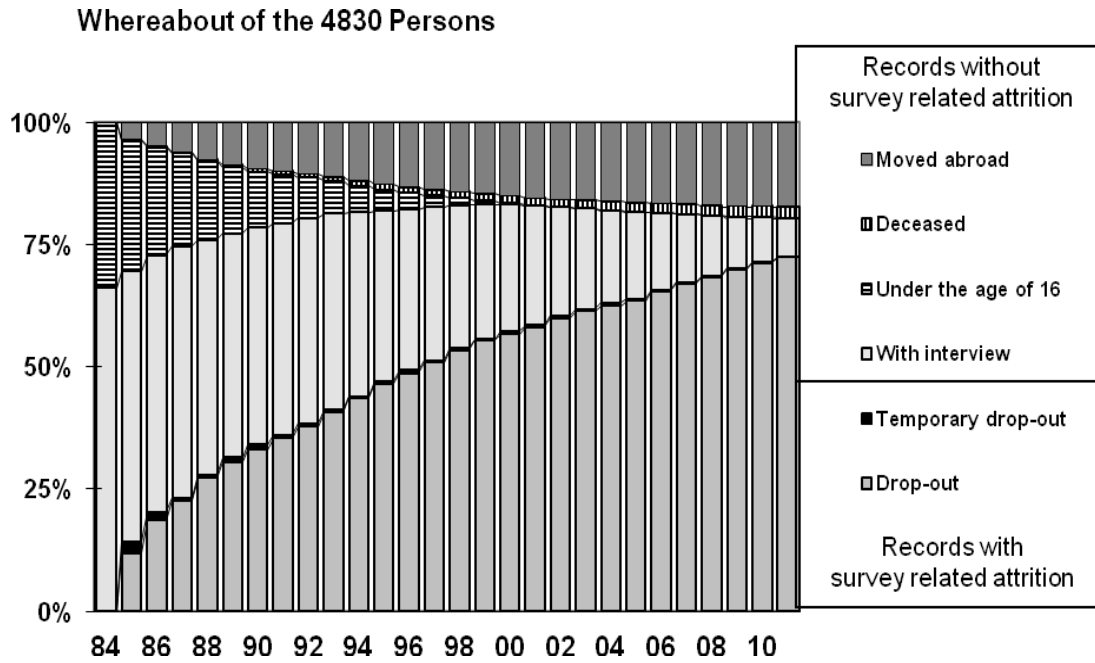


Figure 11: All First Wave Persons (Gross Subsample C). Development up to Wave 22.

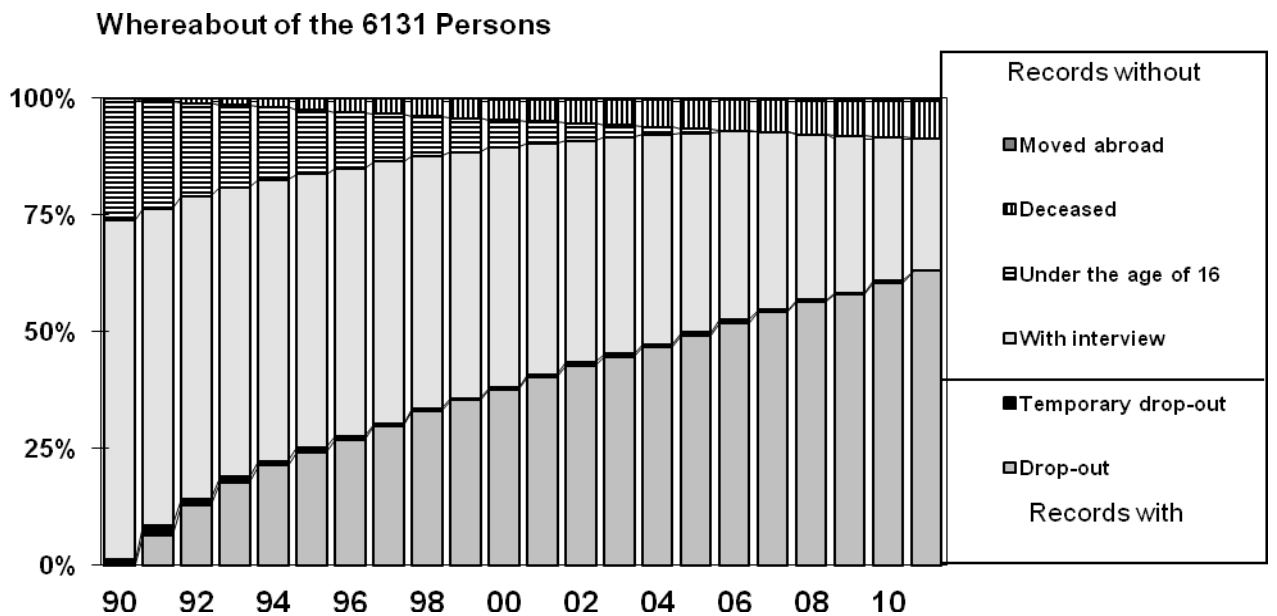


Figure 12: All First Wave Persons (Gross Subsample D). Development up to Wave 17.

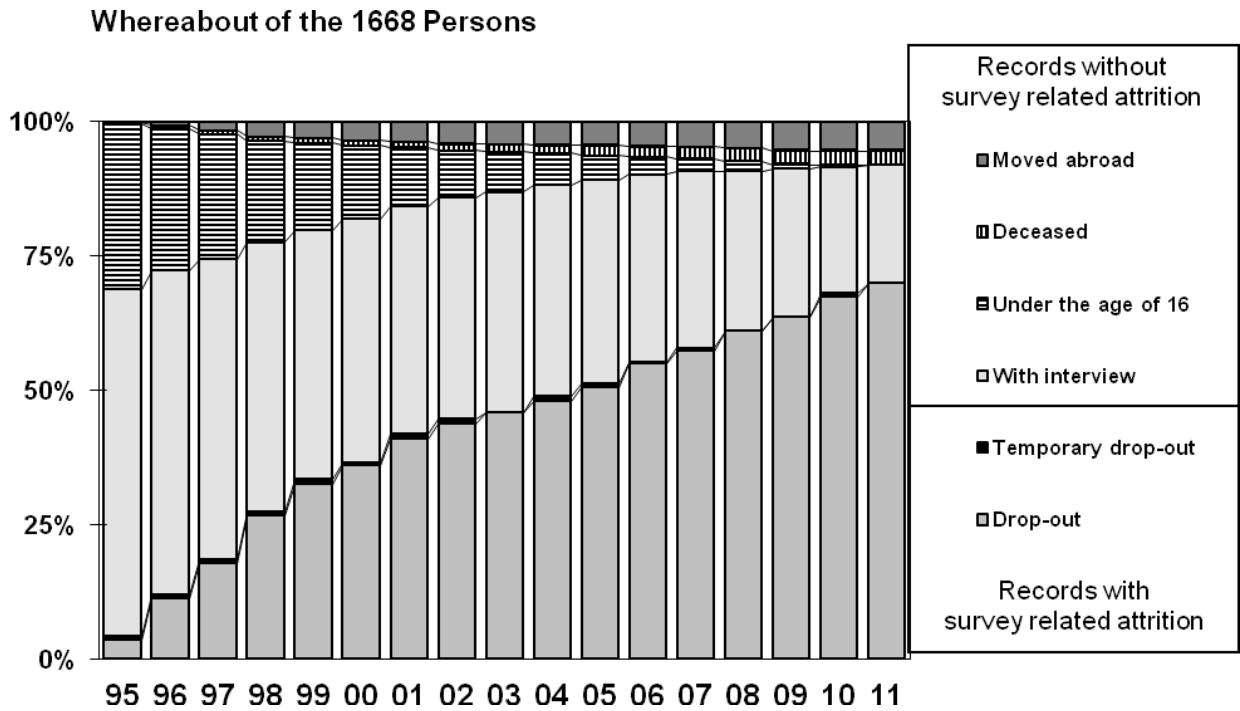


Figure 13: All First Wave Persons (Gross Subsample E). Development up to Wave 14.

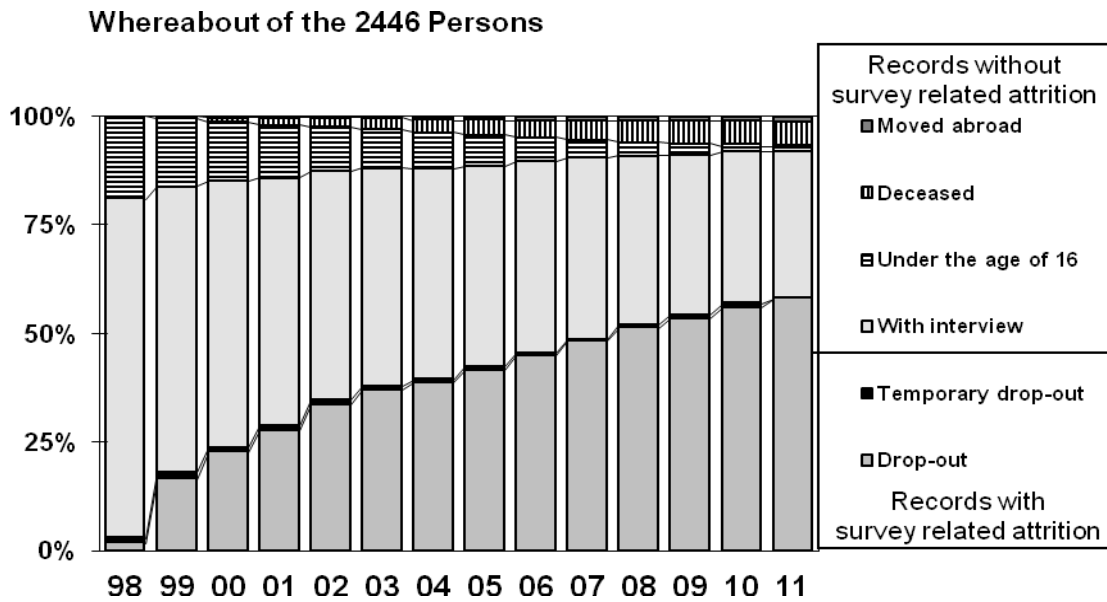


Figure 14: All First Wave Persons (Gross Subsample F). Development up to Wave 12.

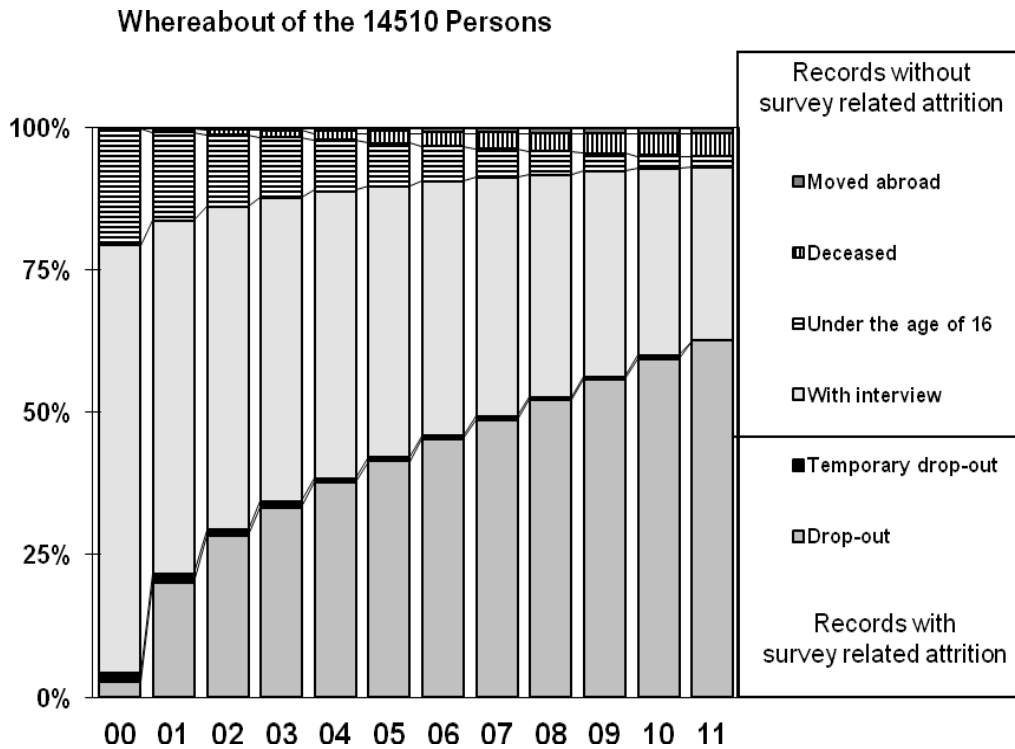


Figure 15: All First Wave Persons (Gross Subsample G). Development up to Wave 10.

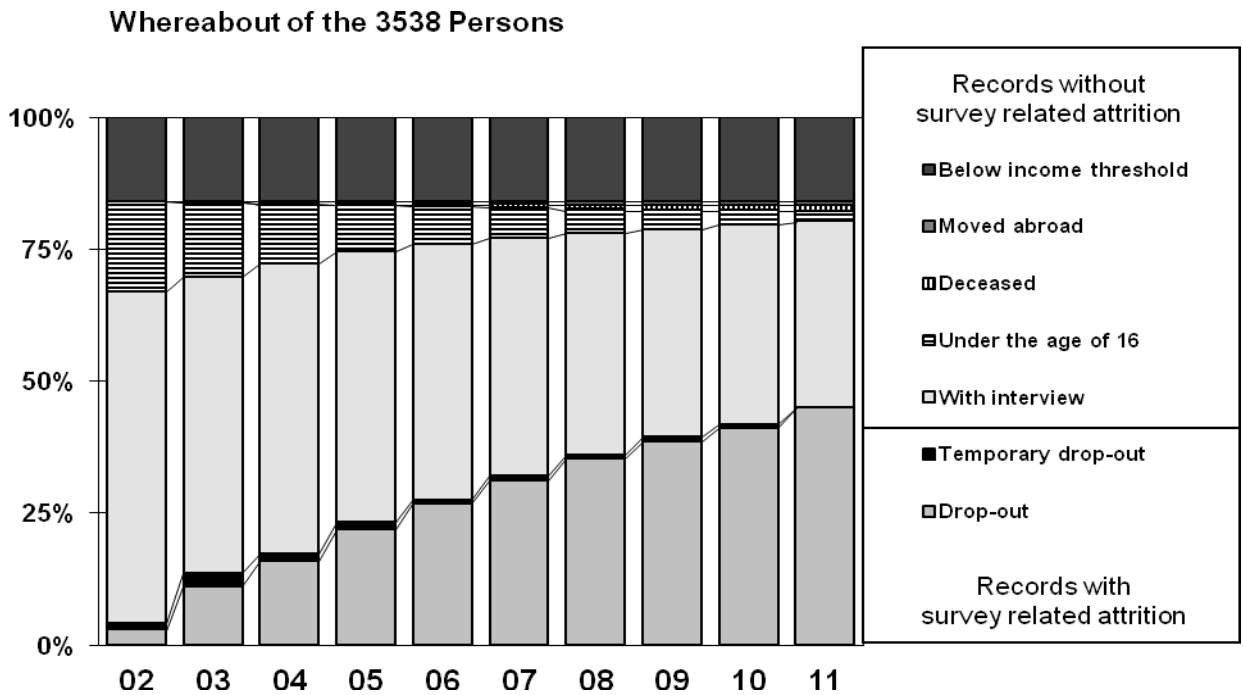
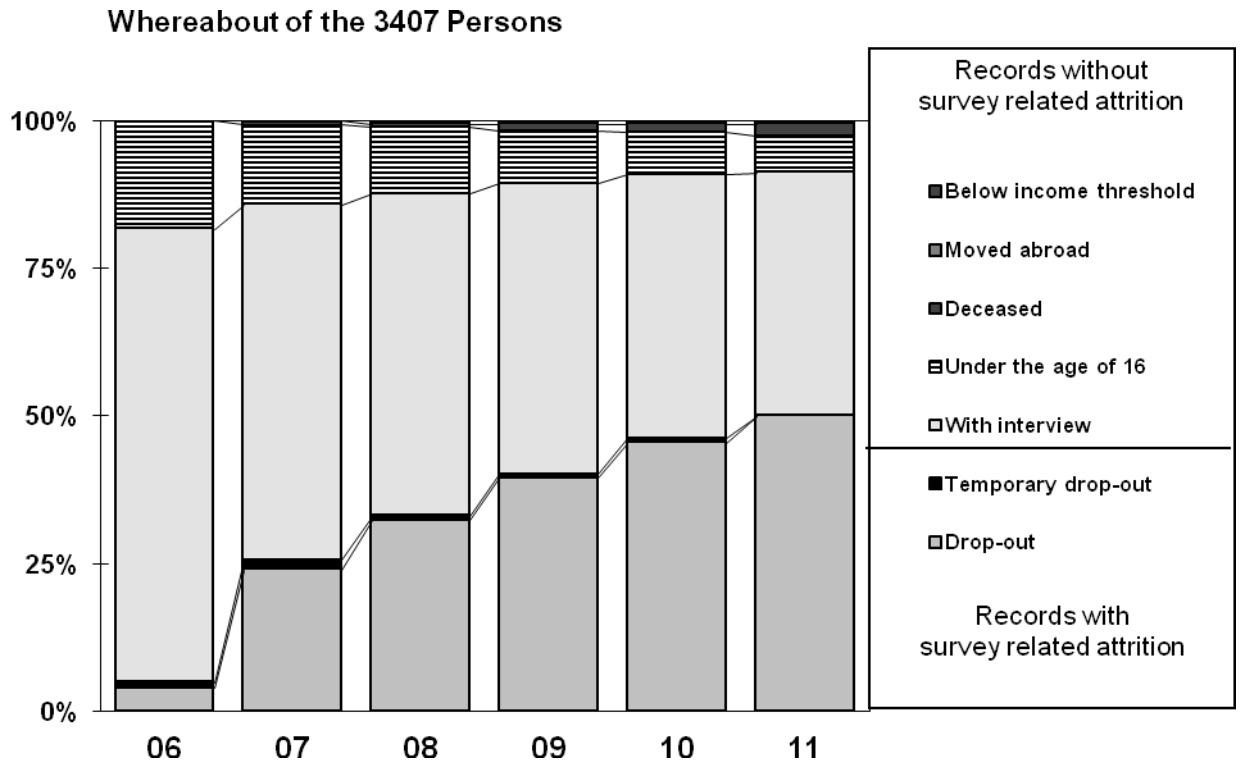


Figure 16: All First Wave Persons (Gross Subsample H). Development up to Wave 6.



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 Subsample A

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

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

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

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

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

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

Figure 24: Entrants Who Were Born or Moved into SOEP Households and Their Participation Behavior in Subsample H

Figure 17: Entrants and their Participation Behavior (Subsample A).

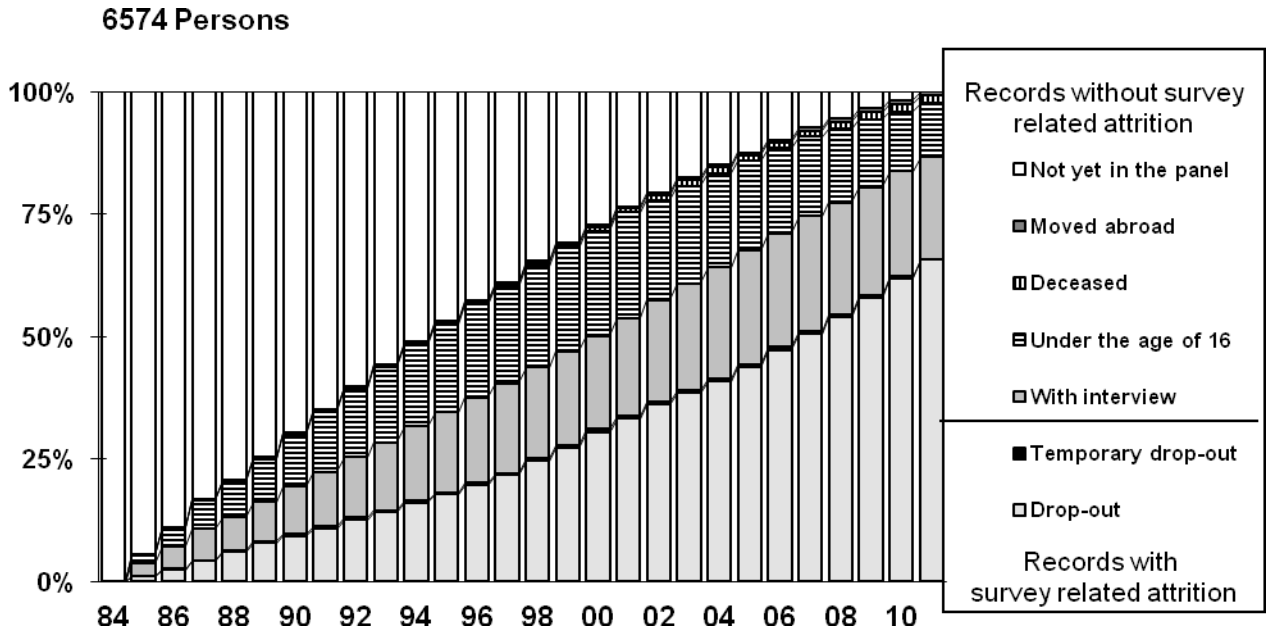


Figure 18: Entrants and their Participation Behavior (Subsample B).

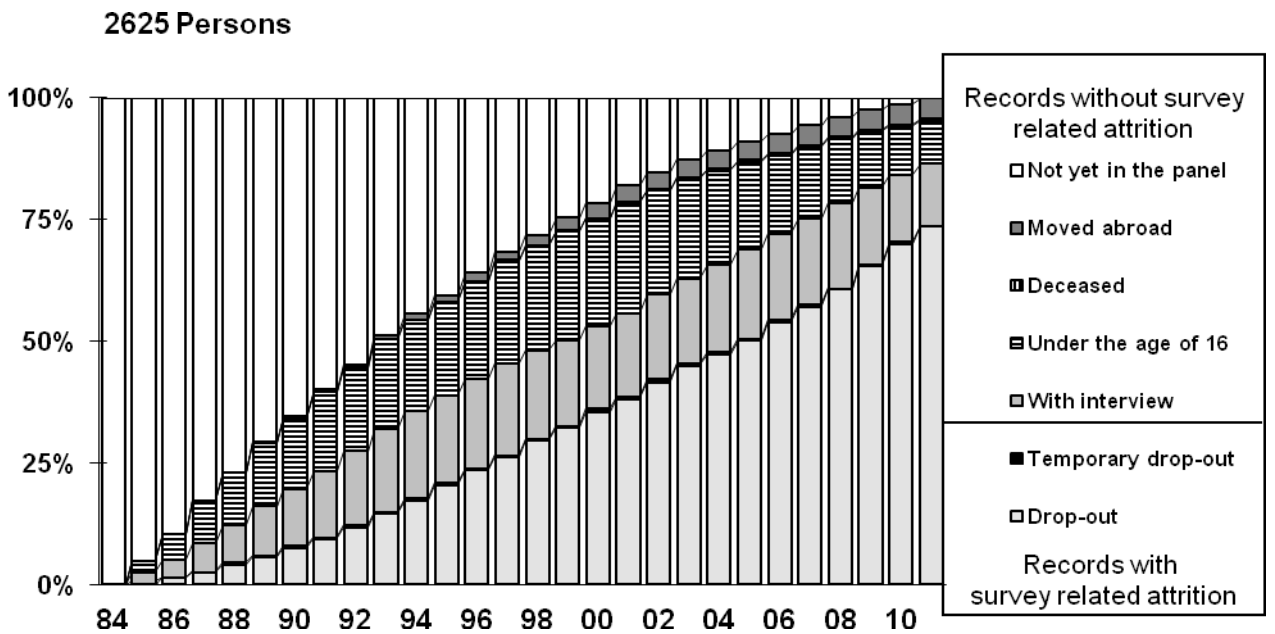


Figure 19: Entrants and their Participation Behavior (Subsample C).

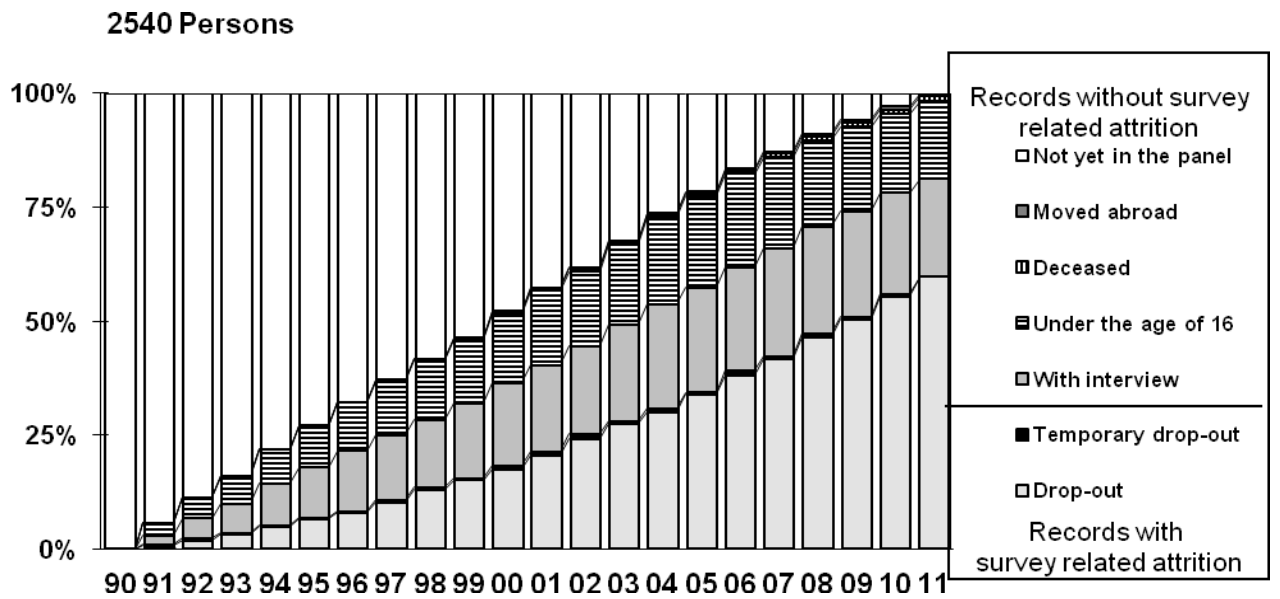


Figure 20: Entrants and their Participation Behavior (Subsample D).

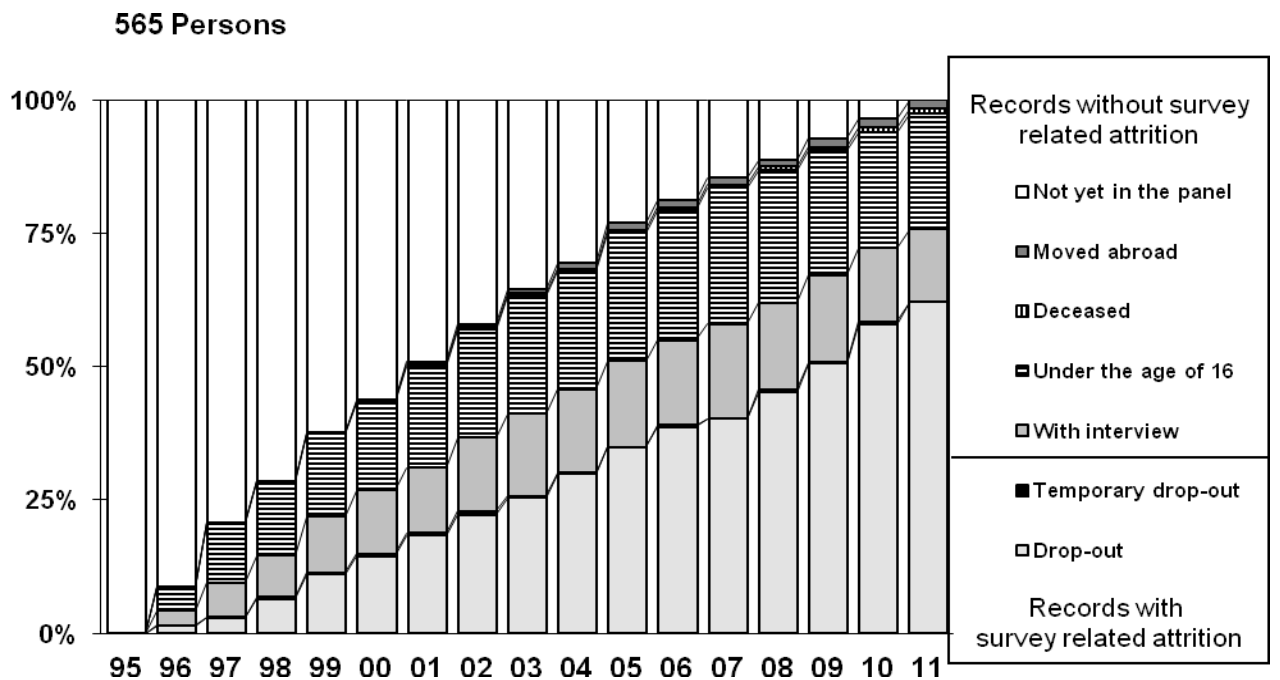


Figure 21: Entrants and their Participation Behavior (Subsample E).

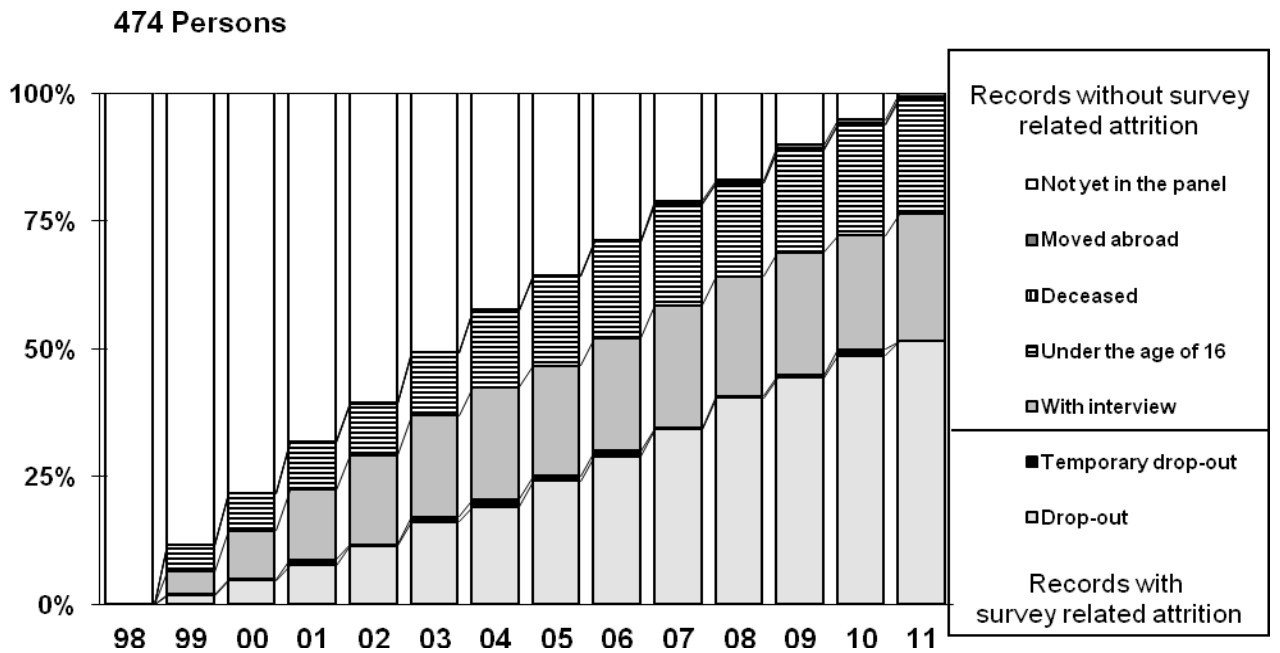


Figure 22: Entrants and their Participation Behavior (Subsample F).

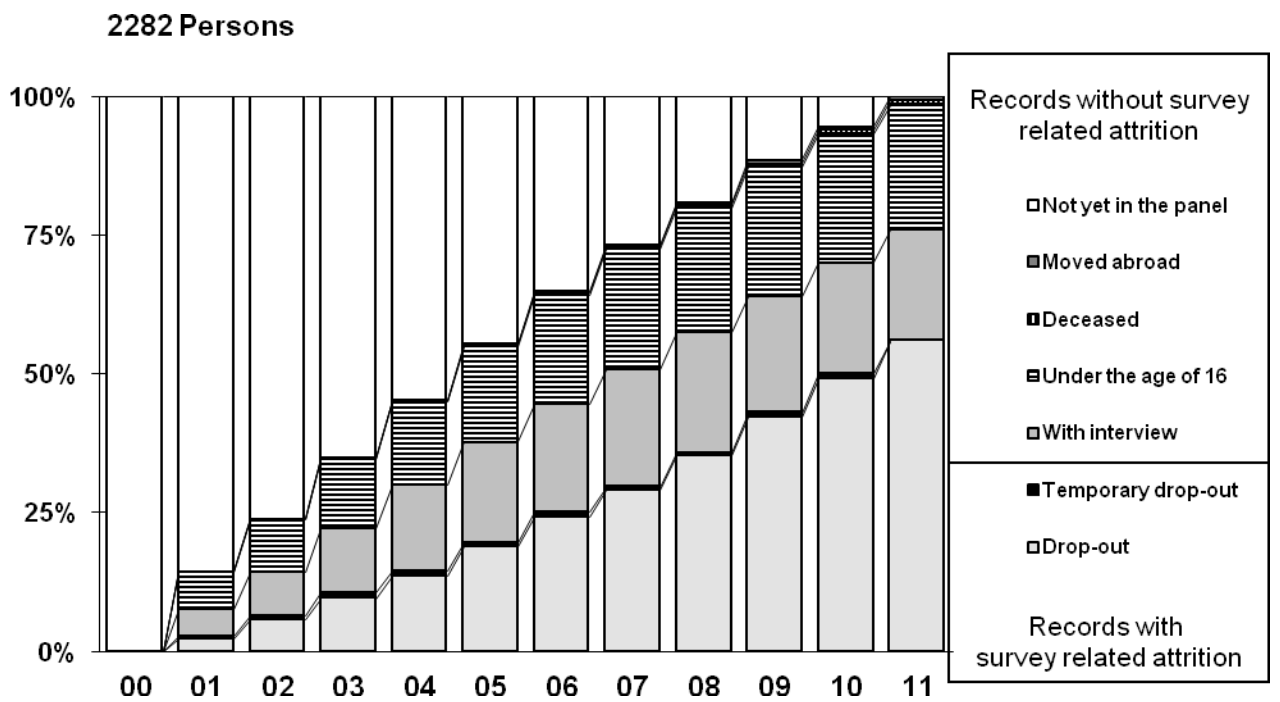


Figure 23: Entrants and their Participation Behavior (Subsample G).

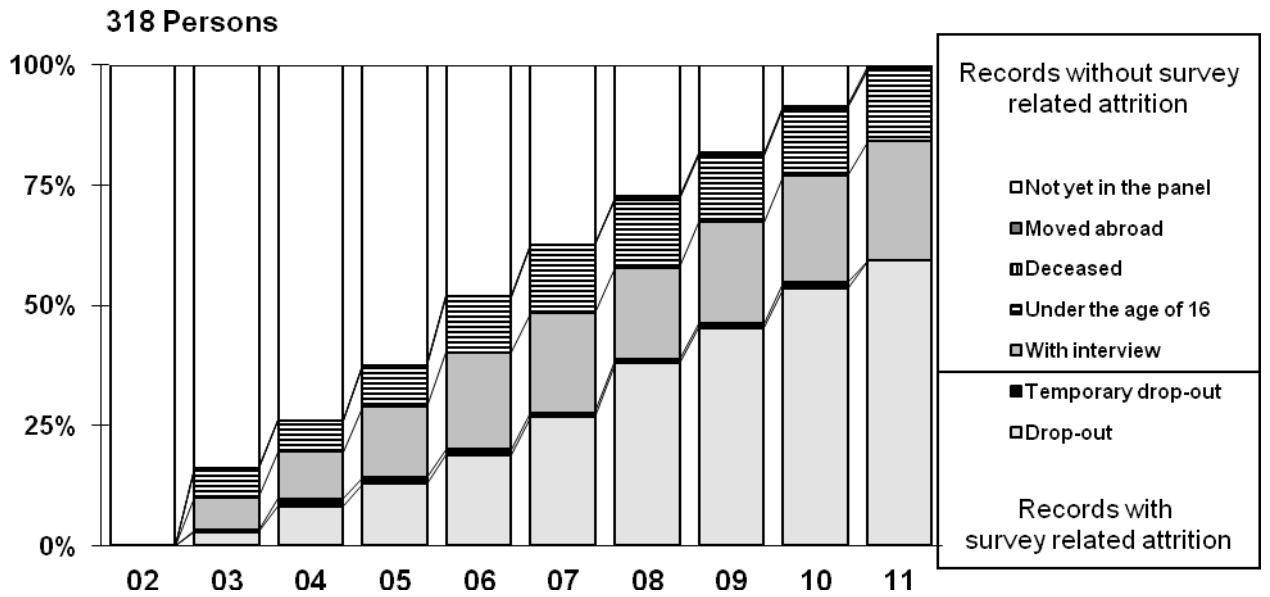
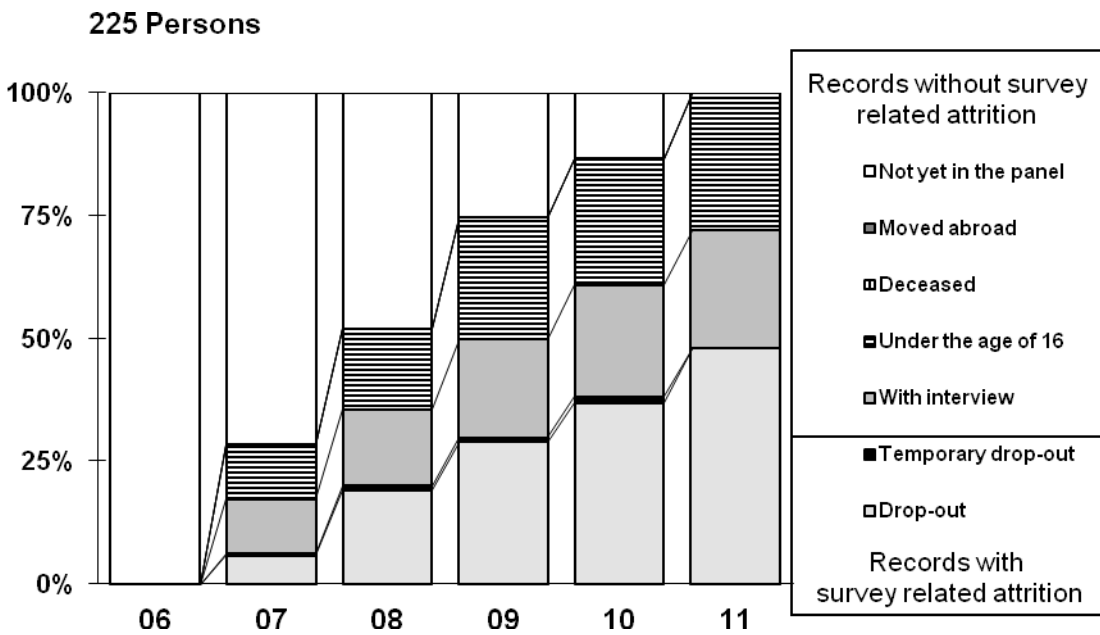


Figure 24: Entrants and their Participation Behavior (Subsample H).



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 25, 26, and 27) and some basic socio-demographic characteristics measured in the year of sampling, such as age, occupation, income, and education (Figures 28 through 31). 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 25), 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 (Figures 26 and 27), 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 25: Successful Re-Interviewing of First-Wave Respondents by Subsamples A, B, C.

Figure 26: Successful Re-Interviewing of First-Wave Respondents by Subsamples D, E, F.

Figure 27: Successful Re-Interviewing of First-Wave Respondents by Subsamples G, H.

Figure 28: Successful Re-Interviewing of All First-Wave Respondents by Age Categories.

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

Figure 30: Successful Re-Interviewing of All First-Wave Respondents by Income Quintiles.

Figure 31: Successful Re-Interviewing of All First-Wave Respondents by Education.

Figure 25: Successful Re-Interviewing of First-Wave Respondents by Subsamples A, B, C. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad.

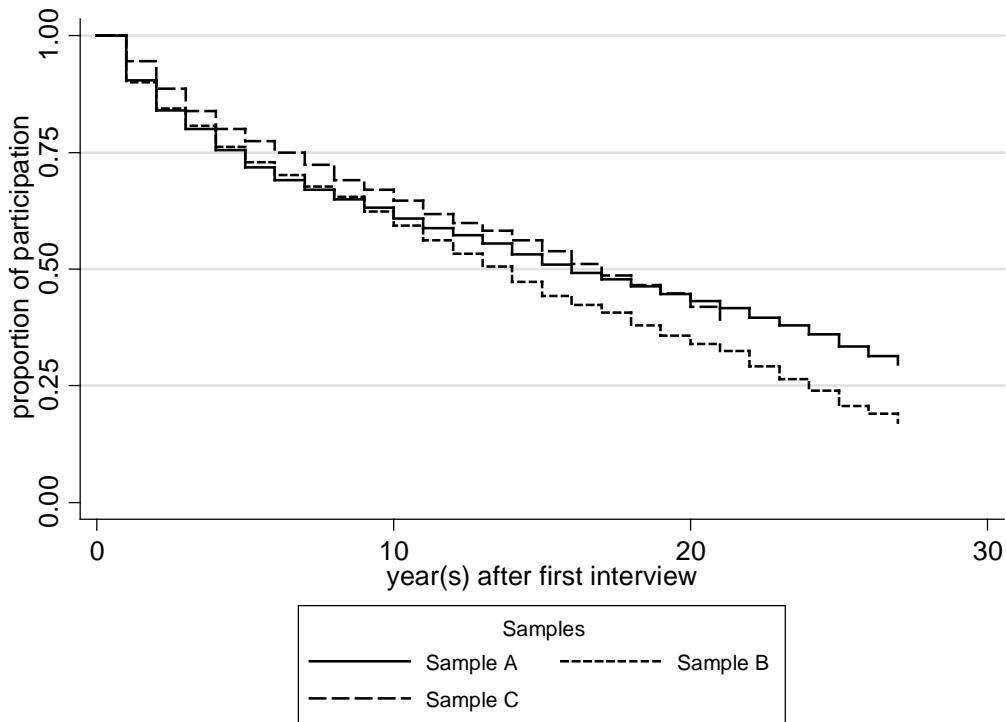


Figure 26: Successful Re-Interviewing of First-Wave Respondents by Subsamples D, E, F. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad.

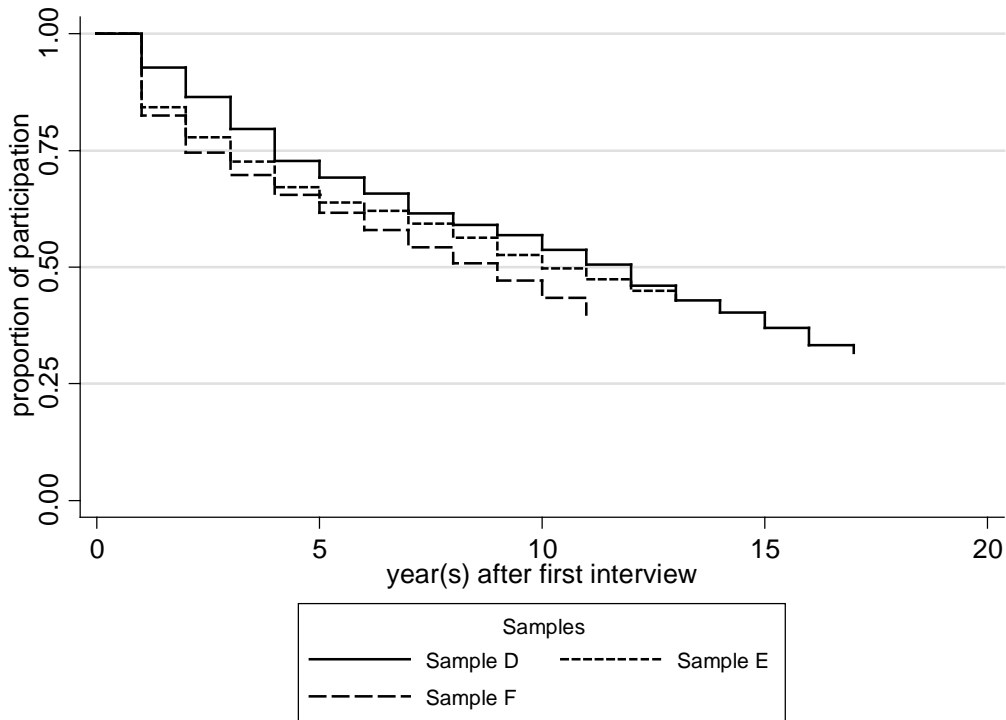


Figure 27: Successful Re-Interviewing of First-Wave Respondents by Subsamples G, H. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad.

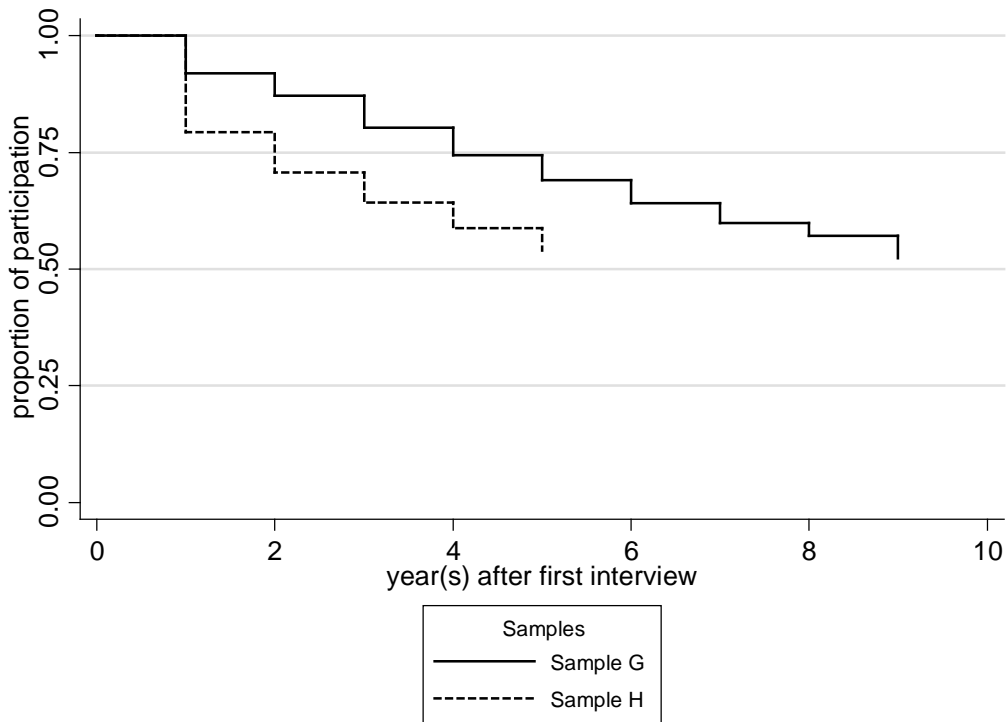


Figure 28: Successful Re-Interviewing of All First-Wave Respondents by Age Categories. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad.

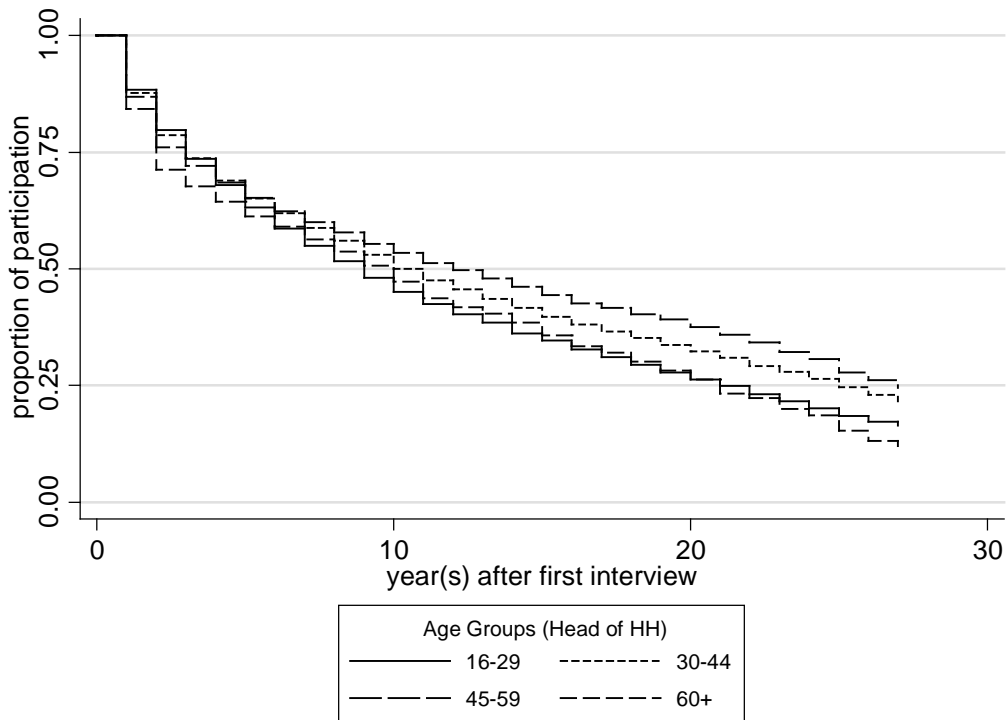


Figure 29: **Successful Re-Interviewing of All First-Wave Respondents by Occupation.**
Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad.

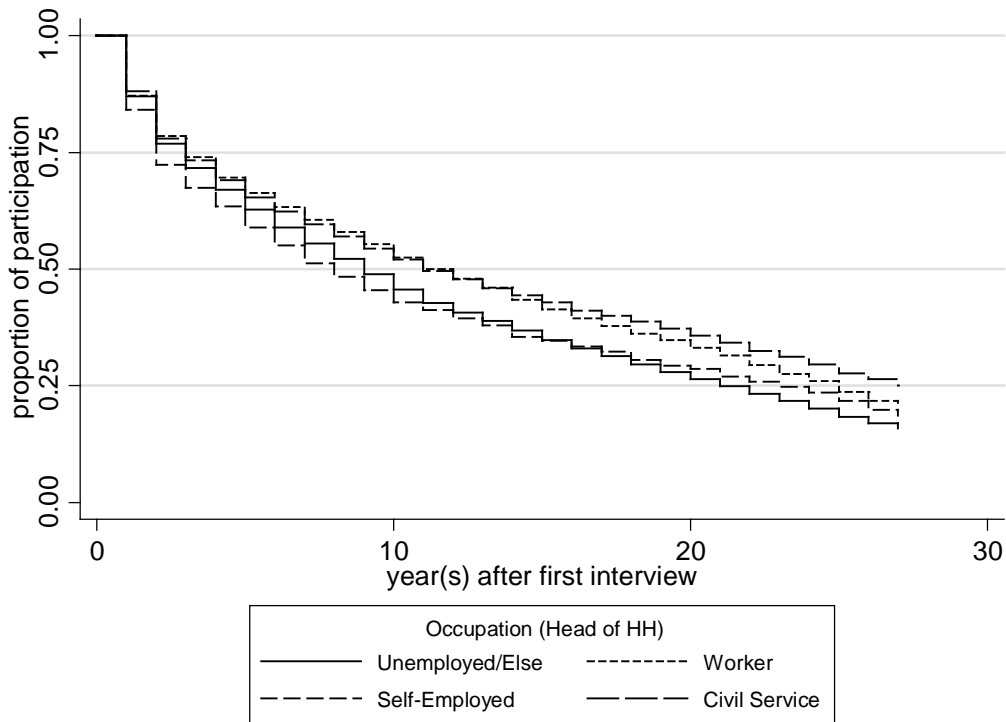


Figure 30: **Successful Re-Interviewing of All First-Wave Respondents by Income Quintiles.**
Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad.

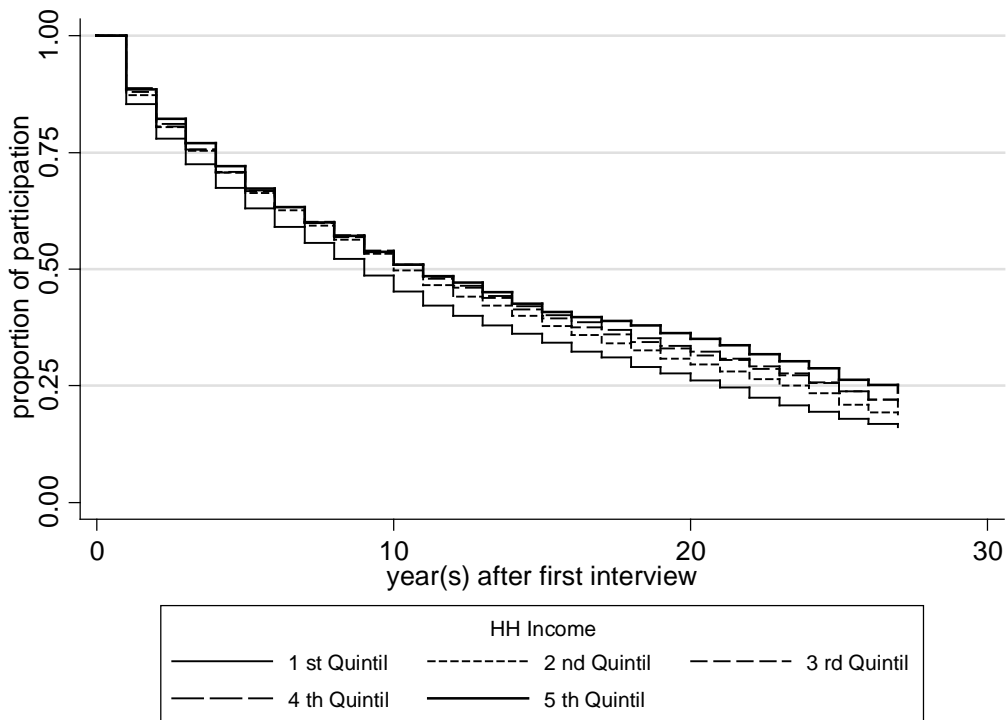
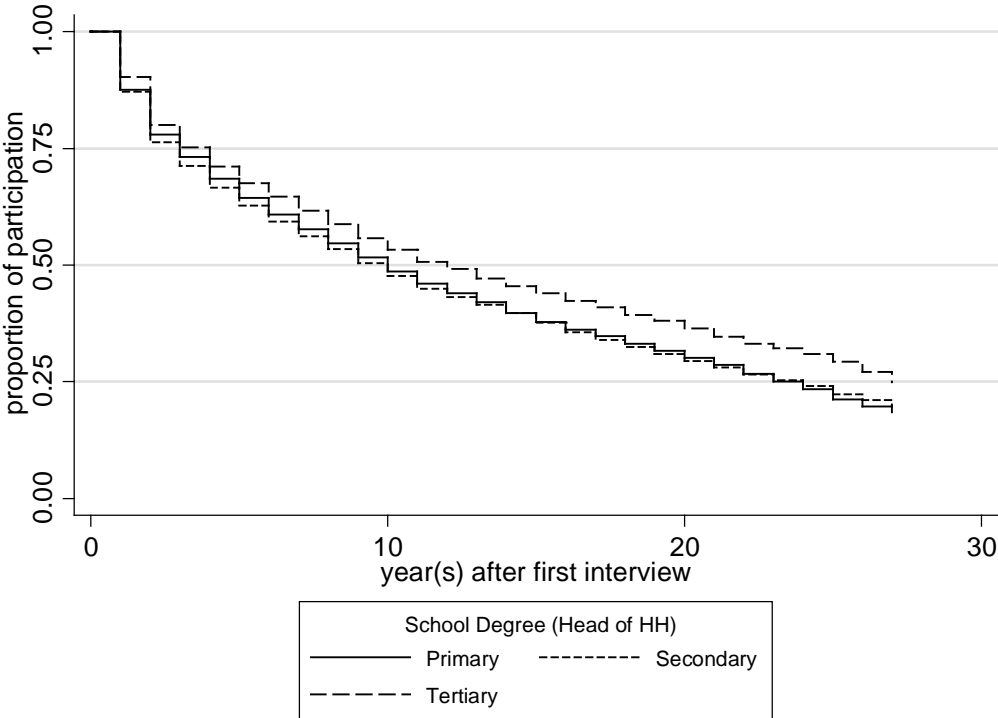


Figure 31: **Successful Re-Interviewing of All First-Wave Respondents by Education. Kaplan-Meier Estimates of Survey-Related Attrition Ignoring Deaths and Moves Abroad.**



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 I and waves 1985 through 2011. 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 66

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		I	
	n	%	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	99.1	1767	99.5	372	99.4	680	99.7	3962	99.4	904	99.6	1326	99.6		
2009	2730	99.7	620	99.2	1695	99.8	351	99.7	636	100	3760	99.6	870	99.5	1145	99.7		
2010	2570	99.8	548	99.3	1627	100	334	99.6	605	99.8	3538	99.5	826	99.9	1059	99.5	1737	98.3
2011	2421	99.8	495	99.0	1541	99.8	302	99.3	589	100	3318	99.6	797	99.6	992	99.5		

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 2011

Based on the household and interview characteristics measured in 2010, we aim at predicting the probability of re-contacting a household relative to unsuccessful follow-up in 2011. 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 2010 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
Interview Characteristics		
Change in Interviewer	Change in Interviewer between Last Waves	0/1
Temp. HH	HH Temporary Drop-Out	0/1
Refusal Related HH	Related HH Drop-Out (Split-Off-HH)	0/1
Move HH	HH Move	0/1
New HH	HH New	0/1
Interviewer Related HH	Same Interviewer in Related HH	0/1
Demographics and Health		
Age below 25	Head of HH Younger than 25 Years	0/1
Age 35-44	Head of HH between 35 and 44 Years	0/1
Single HH	One Person Living in HH	0/1
Couple previous Year	Couple in HH in Previous Year	0/1

Data Documentation 66

3 Panel Attrition Due to Unsuccessful Follow-Ups

Living Apart Together	Partner Living Apart HH	0/1
Non-German HH	One or More Non-Germans Living in HH	0/1
Moving In previous Year	Moving In in the Previous Year	0/1
Underweight	Person Having a BMI Score Less than 18.5	0/1

Financial Situation, Real Estate and Insurance

Unemployed	Head Of HH Status: Unemployed	0/1
Irregular Work	Head of HH Status: Irregular Working	0/1
Benefits	Unemployment Benefit or Social Benefit	0/1
Changing Job Location	Different Job Locations	0/1
Unemployment HH	One or More Persons Unemployed in HH	0/1

Personality Traits and Well-Being

Low Satisfaction	Satisfaction in general low	0/1
------------------	-----------------------------	-----

Data Documentation 66

3 Panel Attrition Due to Unsuccessful Follow-Ups

Table 3a: Estimates of Logit Models of the Probability of Re-Contacting a Household (Relative to Unsuccessful Follow-Up) in 2011.

	Sample A	Sample B	Sample C	Sample D	Sample E	Sample F	Sample G	Sample H
Intercept	2.57	1.57	2.24			3.70	2.18	2.08
<u>Interview Characteristics</u>								
Change in Interv.						-0.65		
Temp. Dropout HH			-0.89					
Refusal Related HH							-0.91	
Move HH						-1.00		
New HH							-0.91	
Interv. Related HH						-0.77		
<u>Demographics and Health</u>								
Age below 25	-0.75					-1.00		
Age 35-44							-0.93	
Single HH						-0.89		
Couple prev. Year			-0.79					
Livi. Apart Togeth.								-0.77
Non-German HH	-0.77							
Moving prev. Year						-1.34		
Underweight			-0.83					
<u>Financial Situation, Real Estate and Insurance</u>								
Unemployed								-0.75
Irregular Work						-0.87		
Benefits						0.43		

Data Documentation 66

3 Panel Attrition Due to Unsuccessful Follow-Ups

Table 3b: **Estimates of Logit Models of the Probability of Re-Contacting a Household (Relative to Unsuccessful Follow-Up) in 2011.**

	Sample A	Sample B	Sample C	Sample D	Sample E	Sample F	Sample G	Sample H
Change Job Locat.	-0.90							
Unemployment HH						-0.91		
<i>Personality, Traits and Well-Being</i>								
Low Satisfaction						-0.88		
<i>Likel. Ratio (Pr > Chisq)</i>	0.45	***	***	***		***	***	***

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

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 66

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		I	
	n	%	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	89.25	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		
2010	2565	87.5	545	80.9	1627	88.3	333	83.5	604	91.6	3523	86.7	825	90.1	1054	86.6	1708	71.3
2011	2417	88.9	491	79.6	1538	88.1	300	88.7	589	92.5	3307	87.2	794	88.9	988	86.8		

n = Number of re-contacted households

% = Percentage of households that participated

4.2 Predicting the Probability of Re-Interviewing versus Refusal in the Year 2011

Based on the household and interview characteristics measured in 2010, and some regional information measured in 2011, we aim at predicting the probability of agreement vs. refusal to participate in the survey by the households that were re-contacted in 2011. 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). In many other cases, personal information are aggregated within the households, for instance, rare events, such as acute medical conditions.

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 2010 are not reported in the present data documentation due to space restrictions, but can be obtained from previous attrition reports.

Data Documentation 66**4 Panel Attrition Due to Refusals****Table 5: Definition of the Regressors of the Logit Model of Refusal.**

Variable	Label	Value
Interview Characteristics		
Female Head of HH	Head of HH Male/Female	0/1
Freshmen	Head of HH's Number of Interviews below Three	0/1
Original Sample Member	Head of HH is Original Sample Member	0/1
New HH	New Split-Off HH with New Address	0/1
Partial Unit Nonresponse	Person(s) in HH did not Participate	0/1
Temporary Drop-Out	Temporary Drop-Out of HH in Previous Year	0/1
Email Disclosed	Email Address Known	0/1
Phone Disclosed	Telephone Number Known	0/1
Change in Interviewer	Change in Interviewer Between Last Waves	0/1
CAPI	Computer Assisted Personal Interview	0/1
SAQ	Self-Administered Questionnaire	0/1
Change in Interview Mode	Change in Interview Mode Between Last Waves	0/1
Temp. Related HH	Related HH Temporary Drop-Out	0/1
Refusal Related HH	Related HH Refusal	0/1
Interviewer Related HH	Same Interviewer in Related HH	0/1
Yugoslavian Head of HH	Head of HH is from a former Yugoslavian country	0/1
Incentives	Monetary Incentives	0/1
Add. Questionnaire Bio	Additional Questionnaire Answered	0/1
Long Interview	Interview Duration Long	0/1
Demographics and Health		
Female Head of HH	Head of HH is Female	0/1
Age 25-34	Head of HH between 25 and 34 Years	0/1
Age 35-44	Head of HH between 35 and 44 Years	0/1
Age 55-64	Head of HH between 55 and 64 Years	0/1
Age 65-74	Head of HH between 65 and 74 Years	0/1
Age 75+	Head of HH Age 75 and Above	0/1

Data Documentation 66

4 Panel Attrition Due to Refusals

Hospital Stay Prev. Year	Head of HH staying in hospital in previous year	0/1
Birth in HH	Baby was Born In HH	0/1
Death in HH	Someone Deceased In HH	0/1
Single HH	One Person Living in HH	0/1
Family HH	Family Living in HH	0/1
Divorced	Head of HH Divorced	0/1
Widowed	Head of HH Widowed	0/1
Moving In	Current Moving In HH	0/1
Couple	Unmarried Couple in HH	0/1
Separation	Current Separation in HH	0/1
Marriage	Marriage in HH	0/1
Previous Couple	Unmarried Couple Previous Year in HH	0/1
Partner Apart Together	Partner Living Apart Together	0/1
Divorce in HH	Someone Divorced in HH	0/1
Health Situation		
Poor Health	General Health Situation Rated as Poorly	0/1
Hospital Stay	At least one HH-Member stayed in Hospital	0/1
Chain Smoker	Chain Smoking	0/1
Care Insurance Benefits	Benefits from Care Insurance in HH	0/1
Overweight	Person Having a BMI Score Higher than 30	0/1
Financial Situation, Real Estate and Insurance		
Subtenant	Subtenant of Dwelling	0/1
Income 25	HH in First Quartile of Income Distribution	0/1
Income 100	HH in Fourth Quartile of Income Distribution	0/1
No Insurances	No Insurances	0/1
Insurances	More than 4 Insurances	0/1
Work and Education		
Commuting	At Least one Member of HH is Commuting	0/1

Data Documentation 66

4 Panel Attrition Due to Refusals

Unemployed	Head of HH is Unemployed	0/1
Irregular Employment	Head of HH Irregularly Employed	0/1
White-Collar Worker	Head of HH is White-Collar Worker	0/1
Civil Servant	Head of HH is a Civil Servant	0/1
Other Employee	Head of HH is Other Employee	0/1
New Job in HH	At least one Person in HH got Employed	0/1
Unemployed in HH	At least one Person in HH is Unemployed	0/1
Not Employed in HH	At least one Person in HH is Not Employed	0/1
Unemployment HH	One or More Persons Unemployed in HH	0/1
All Unemployed in HH	Whole HH is Unemployed	0/1
Low Education	Head of HH has a Low Educational Level	0/1
Personality Traits and Well-Being		
Low Happiness	Head of HH has Low Level of Happiness	0/1
High Happiness	Head of HH has High Level of Happiness	0/1
High Anger	Head of HH has High Level of Anger	0/1
Strong Political Interest	Head of HH has very High or No Political Interest	0/1
Strong Reciprocity	Head of HH has High Level of Reciprocity	0/1
Leisure Time	Head of HH has at Least 3 Hours of Leisure Time per Workday	0/1
Low Risk Aversion	Head of HH Willing to Take Risks	0/1
Building, Area, and Region		
Internet Area	HH Located in Area of High Internet Affinity	0/1
Prosperous Area	HH Located in Area of High Prosperity	0/1
Traditional Area	HH Located in Traditional Area	0/1
Family Area	HH Located in Area with Mostly Families	0/1
Single Area	HH Located in Area with Mostly Singles	0/1
Single Family House Area	HH Located in Area with Mostly Single Family Houses	0/1
Islamic Area	HH Located in Area with High Rate of Islamic Residents	0/1
High Rise Block	HH Located in Area with High Rate of High Rise Buildings	0/1
Low Academics	HH Located in Area of Low Academics Rate	0/1

Data Documentation 66

4 Panel Attrition Due to Refusals

High Turnout	HH Located in Area of High Voter Participation	0/1
Rural Area	HH Located in Mostly Rural Area	0/1
Big Town	HH in Town with More Than 100.000 Inhabitants	0/1
High SPD Share of Vote	HH Located in Area of High Share of Voting "SPD"	0/1
High Grüne Share of Vote	HH Located in Area of High Share of Voting "Grüne"	0/1
High Linke Share of Vote	HH Located in Area of High Share of Voting "Linke"	0/1
Low Purchasing Power	HH in Area with Low of Purchasing-Power (Microm)	0/1
High Purchasing Power	HH in Area with High of Purchasing-Power (Microm)	0/1
North Rhine-Westphalia	HH Located in North Rhine Westphalia	0/1
Meckl. West-Pomerania	HH Located in Mecklenburg West-Pomerania	0/1
Rhineland-P./Saarland	HH Located in Rhineland-Pfalz/Saarland	0/1
Low Saxony/Bremen	HH Located in Low Saxony/Bremen	0/1
Hessen	HH Located in Hessen	0/1
Few Abitur-Graduates	HH Located in Area with Few Abitur-Graduates	0/1
Young Area	HH Located in Area with Low Average Age	0/1

Data Documentation 66

4 Panel Attrition Due to Refusals

Table 6a: Estimates of Logit Models for the Probability of Re-Interviewing a Household (Relative to Refusal) in 2011.

	Sample A	Sample B	Sample C	Sample D	Sample E	Sample F	Sample G	Sample H
Intercept	0.23	0.53	1.60	0.42	2.48	0.46	1.23	-0.80
<i>Interview Characteristics</i>								
Freshmen		-1.74				-0.53		
Original Sample Member	0.23	0.50					0.62	0.86
New HH	-1.00	-0.93		-1.59	-1.76			-0.83
Temporary Drop-Out		-1.75						
Email Disclosed			0.25			0.27		
Phone Disclosed	0.60	0.45		1.38		0.64		0.92
Change in Interviewer	-1.30	-1.48	-1.29	-1.39	-1.93	-1.32	-0.97	-1.33
CAPI						0.15		
SAQ			-0.41			-0.24		
Change Interview Mode						-0.24		
Interviewer Related HH	0.28			0.68			0.52	
Yugoslavian Head of HH		-0.42						
Add. Questionnaire Bio		-0.50					0.97	
Incentives					-0.68			

Data Documentation 66

4 Panel Attrition Due to Refusals

Table 6b: **Estimates of Logit Models for the Probability of Re-Interviewing a Household (Relative to Refusal) in 2011.**

	Sample A	Sample B	Sample C	Sample D	Sample E	Sample F	Sample G	Sample H
<i>Demographic and Relational Characteristics</i>								
Female Head of HH				0.59				
Age 25-34			-0.24				-0.43	
Age 35-44	-0.23						-1.07	
Age 55-64						0.17	0.45	
Age 65-74							0.38	0.30
Age 75+						-0.25		
Birth in HH		0.67						
Death in HH		-1.21						
Single HH	-0.50		-0.65		-0.89	-0.45	-1.34	
Family HH						-0.26		
Divorced	0.27							
Widowed						0.21		
Moving In			-1.02	-1.92		-0.58	-1.33	
Couple	-0.41			0.90				
Separation			-1.18					1.57
Marriage					-1.12			
Previous Couple			-0.46	-1.45	-0.81	-0.32	-1.52	
Divorce in HH	-0.39							
<i>Health Situation</i>								
Poor Health		-0.44						
Hospital Stay	-0.23							
Chain Smoker	0.33							
Care Ins. Benefits	-0.41							
Overweight	0.21							

Data Documentation 66

4 Panel Attrition Due to Refusals

Table 6c: **Estimates of Logit Models for the Probability of Re-Interviewing a Household (Relative to Refusal) in 2011.**

	Sample A	Sample B	Sample C	Sample D	Sample E	Sample F	Sample G	Sample H
<u>Fin. Situation, Real Est., Insurance</u>								
Subtenant		-0.91		-1.02				
Income 25	0.22							
Income 100								0.38
No Insurance			-0.23					
Insurances	0.19							
<u>Work and Education</u>								
Commuting							0.63	
Irregular Employ.							-0.39	
White-Collar Worker	0.45							
Civil Servant	0.24							
Other Employee						0.21		
New Job in HH				-0.63			0.73	
Not Employed in HH	0.22					0.21		
Unemployed HH		0.74						
Low Education				-0.50		0.21		
<u>Personality Traits and Well-Being</u>								
Low Happiness					-0.63			
High Happiness	0.25					0.18		
High Anger	0.26							
Strong Pol. Interest	0.19							
Strong Reciprocity								0.33
Leisure Time			0.31					
Low Risk Aversion							0.28	

Data Documentation 66

4 Panel Attrition Due to Refusals

Table 6d: **Estimates of Logit Models for the Probability of Re-Interviewing a Household (Relative to Refusal) in 2011.**

	Sample A	Sample B	Sample C	Sample D	Sample E	Sample F	Sample G	Sample H
<i>Building, Area and Region</i>								
Internet Area							0.43	
Prosperous Area								-0.31
Traditional Area	0.13				-0.56			
Single Area					0.57			
Single-Fam. House Area			0.21					
Islamic Area					-0.56			
High Rise Block				0.86			-0.56	
Low Academics								
High Turnout								-0.30
Rural Area	-0.16							
Big Town						0.31	0.38	

Data Documentation 66

4 Panel Attrition Due to Refusals

Table 6e: Estimates of Logit Models for the Probability of Re-Interviewing a Household (Relative to Refusal) in 2011.

	Sample A	Sample B	Sample C	Sample D	Sample E	Sample F	Sample G	Sample H
High SPD Share					0.54		0.66	
High Grüne Share						-0.32		
High Linke Share	1.30							
High CDU Share					0.91			
Low Purch. Power		-0.51	0.21					0.41
High Purch. Power					1.33			
N. Rhine- Westphalia		0.49				-0.17		
Meckl. West-Pome.						0.55		
Rhineland-P./Saarl.						0.27		
L. Saxony/Bremen				-0.70				
Hesse		0.96						
Few Abitur-Grad.						0.18		
Young Area								0.44

Likel. Ratio (Pr > Chisq)

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 2011 based on characteristics measured in 2010, BBHBLEIB, lends itself as a longitudinal weighting variable correcting for selective attrition between waves 2010 and 2011. Table 7 reports some sub-sample specific descriptive statistics of the longitudinal weights in each wave.

The product of the cross-sectional weight in 2010, BAHHRF, and the longitudinal weight in 2011, BBHBLEIB, provide the raw data for the cross-sectional weight in 2011. 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 2011. These are at the household level states (Bundesländer), size of the community, household size, and house ownership. At the person level, SOEP weights are also adjusted to the marginal distributions of age, gender, and nationality (Non/German). Table 8 reports sub-sample-specific descriptive statistics of the derived cross-sectional weighting variable BBHHRF and in comparison all previous cross-sectional weights AHHRF through BAHHRF.

Data Documentation 66

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 28).

	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	bahbleib	bbhbleib
<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	1.01	1.00
p50	1.10	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	1.06	1.04
p90	1.22	1.26	1.13	1.20	1.16	1.11	1.09	1.11	1.16	1.15	1.16	1.12	1.13	1.14	1.20	1.15	1.18	1.21	1.14	1.12	1.16	1.22	1.14	1.13	1.25	1.38	1.27
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	2245	2147
<u>sample B</u>																											
p10	1.09	1.10	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	1.01	1.01
p50	1.10	1.10	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	1.10	1.07
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	1.60	1.47	1.55
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	441	391
<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	1.00	1.00	1.01	1.00	1.01	1.00	1.01	1.01
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	1.04	1.05
p90							1.18	1.22	1.17	1.12	1.11	1.15	1.12	1.20	1.10	1.13	1.16	1.21	1.14	1.12	1.15	1.24	1.16	1.18	1.21	1.32	1.24
N							2030	2020	1970	1959	1938	1951	1942	1886	1894	1879	1850	1818	1807	1813	1771	1717	1654	1592	1535	1437	1355
<u>sample D</u>																											
p10												1.00	1.05	1.08	1.05	1.02	1.03	1.00	1.01	1.00	1.00	1.03	1.01	1.02	1.00	1.00	1.01
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	1.01	1.02
p90												1.16	1.09	1.35	1.27	1.10	1.18	1.21	1.09	1.25	1.34	1.44	1.12	1.22	1.16	1.43	1.28
N												396	340	308	300	302	286	289	290	277	273	261	248	231	220	278	266

Data Documentation 66

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 I (Percentiles of \$HBLEIB up to Wave 28).

	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	bahbleib	bbhbleib	
<u>sample E</u>																												
p10															1.00	1.03	1.01	1.01	1.04	1.00	1.01	1.00	1.01	1.00	1.00	1.01	1.00	
p50															1.23	1.07	1.05	1.02	1.04	1.01	1.03	1.03	1.01	1.01	1.04	1.04	1.00	
p90															1.47	1.21	1.25	1.20	1.15	1.08	1.18	1.21	1.16	1.19	1.17	1.25	1.17	
N															886	838	811	773	744	732	706	686	647	602	574	553	546	
<u>sample F</u>																												
p10																	1.08	1.03	1.02	1.02	1.01	1.01	1.02	1.01	1.02	1.01	1.01	
p50																	1.14	1.05	1.04	1.03	1.03	1.03	1.03	1.03	1.03	1.04	1.05	1.05
p90																	1.59	1.46	1.24	1.19	1.17	1.29	1.15	1.14	1.34	1.40	1.31	
N																	4911	4586	4386	4235	4070	3895	3694	3513	3303	3055	2886	
<u>sample G</u>																												
p10																			1.06	1.02	1.03	1.00	1.02	1.01	1.02	1.00	1.00	
p50																			1.10	1.03	1.06	1.04	1.05	1.03	1.04	1.01	1.03	
p90																			1.17	1.25	1.25	1.31	1.17	1.18	1.36	1.24	1.33	
N																			911	904	879	859	824	787	757	743	706	
<u>sample H</u>																												
p10																							1.04	1.01	1.01	1.01	1.00	
p50																							1.16	1.03	1.03	1.04	1.05	
p90																							1.46	1.18	1.22	1.37	1.31	
N																							1188	1082	996	913	857	

Data Documentation 66

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 28).

	ahhrf	bhhrf	chhrf	dhhrf	ehhrf	fhhrf	ghhrf	hhhrf	ihhrf	jhhrf	khhrf	lhhrf	mhhrf	nhhrf
p5	434	496	542	548	534	552	698	679	668	685	706	713	741	753
p10	601	682	758	796	805	824	1072	1050	1038	1057	1102	1134	1182	1216
p25	3777	3887	3602	3536	3550	3607	2217	2342	2332	2404	2405	2318	2333	2342
p50	4721	5082	5302	5406	5637	5857	4604	4698	4648	4679	4674	4376	4364	4331
p75	5661	6413	6826	7036	7545	7876	7048	7168	7154	7248	7282	7004	7043	7068
p90	7123	8458	9243	9592	10369	10855	9954	10268	10510	10760	11211	11082	11360	11873
p95	8308	12397	11125	11433	12537	13278	12364	12983	13647	13989	14604	14797	15283	15892
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	bahhrf	bbhhrf
p5	1067	1047	833	797	519	522	506	509	473	461	465	470	497	450
p10	1403	1390	1132	1084	692	704	693	699	671	672	669	674	702	669
p25	2350	2329	1763	1761	1275	1283	1252	1273	1285	1270	1294	1327	1399	1402
p50	3985	4015	2529	2743	2563	2570	2535	2541	2408	2472	2541	2607	2779	2560
p75	6237	6514	3562	4123	4160	4293	4390	4495	4111	4432	4714	4991	5438	4264
p90	9843	10761	5129	6051	6388	6703	7127	7445	6923	7590	8422	9319	10264	7145
p95	13051	14094	6431	7832	8229	9013	9874	10758	9711	10675	11593	12461	13502	9577
N	7492	7220	13082	11796	12320	11909	11644	11294	12361	11552	10921	10270	9551	12146

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.
- 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.
- Schonlau, Matthias, Watson, Nicole, and Martin Kroh. 2011. Household Survey Panels: How Much do Following Rules Affect Sample Size? *Survey Research Methods* 5: 53-61.
- Spiess, Martin and Ulrich Rendtel. 2000. Combining an Ongoing Panel with a New Cross-Sectional Sample. *DIW Discussion Paper* 198.
- 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.